Updated May 24, 2007

Attilio Meucci

Risk and Asset Allocation

Technical Appendices

Contents

Preface

This booklet "Technical Appendices" complements the textbook "Risk and Asset Allocation" - Springer Quantitative Finance, by Attilio Meucci.
This booklet can be downloaded from the book's website:

`www.symmys.com` > Book > Downloads > Technical Appendices.

Each chapter of the booklet "Technical Appendices" refers to the respective chapter in the textbook "Risk and Asset Allocation".

This booklet cross references the formulas in the textbook. In order not to generate confusion when referencing the formulas in this booklet versus the formulas in the textbook, the numbering of the formulas in this booklet are preceded by a "T".

Also notice that in the textbook the notation, say, "Appendix www.2.4" refers to Chapter 2, Section 4 of this booklet (which is located on the web at the above address). On the other hand the notation, say, "Appendix B.3" refers to the mathematical Appendix B, Section 3, at the end of the textbook.

Any feedback on the "Technical Appendices", on "Risk and Asset Allocation", as well as on the materials available at `www.symmys.com` are highly appreciated: please visit this webiste to contact the author.

New York City, May 2007,

Attilio Meucci

1

Technical appendix to Chapter 1

1.1 Distribution of functions of random variables

Consider a random variable Y defined as an invertible function g of a random variable X as follows:

$$X \mapsto Y \equiv g(X). \tag{$T1.1$}$$

We compute the representations of the distribution of Y in terms of the representations of the distributions of X.

- Probability density function.

  By the definition (1.3) of the pdf f_Y we have:

$$\begin{aligned} f_Y(y)\,dy &= \mathbb{P}\{Y \in [y, y+dy]\} = \mathbb{P}\{g(X) \in [y, y+dy]\} \\ &= \mathbb{P}\left\{X \in \left[g^{-1}(y), g^{-1}(y+dy)\right]\right\} \\ &= \int_{g^{-1}(y)}^{g^{-1}(y+dy)} f_X(x)\,dx, \end{aligned} \tag{$T1.2$}$$

where the second to last equality follows from the invertibility of the function g. On the other hand, from a Taylor expansion we obtain:

$$g^{-1}(y+dy) = g^{-1}(y) + \frac{1}{g'(g^{-1}(y))}dy. \tag{$T1.3$}$$

Substituting ($T1.3$) in ($T1.2$) we obtain:

$$\begin{aligned} f_Y(y)\,dy &= \int_{g^{-1}(y)}^{g^{-1}(y)+\frac{1}{g'(g^{-1}(y))}dy} f_X(x)\,dx \\ &= f_X\left(g^{-1}(y)\right)\left|\frac{1}{g'(g^{-1}(y))}\right| dy, \end{aligned} \tag{$T1.4$}$$

which yields the desired result:

$$f_Y(y) = \frac{f_X\left(g^{-1}(y)\right)}{\left|g'\left(g^{-1}(y)\right)\right|}. \tag{T1.5}$$

- Cumulative distribution function.

By the definition (1.7) of the cumulative distribution function F_Y we have:

$$\begin{aligned} F_Y(y) &\equiv \mathbb{P}\{Y \leq y\} = \mathbb{P}\{g(X) \leq y\} \\ &= \mathbb{P}\left\{X \leq g^{-1}(y)\right\} \\ &= F_X\left(g^{-1}(y)\right), \end{aligned} \tag{T1.6}$$

where the second to last equality follows from the invertibility of the function g, under the assumption that g is an increasing function of its argument. In case g is a decreasing function of its argument we obtain similarly:

$$\begin{aligned} F_Y(y) &\equiv \mathbb{P}\{Y \leq y\} = \mathbb{P}\{g(X) \leq y\} \\ &= \mathbb{P}\left\{X \geq g^{-1}(y)\right\} = 1 - \mathbb{P}\left\{X \leq g^{-1}(y)\right\} \\ &= 1 - F_X\left(g^{-1}(y)\right). \end{aligned} \tag{T1.7}$$

- Quantile.

Also the quantile of a transformed variable is particularly simple to compute in terms of the quantile of the original variable. Indeed consider the following series of identities that follow from the definition (1.7) of the cumulative distribution function F_Y:

$$F_Y\left(g\left(Q_X(p)\right)\right) = \mathbb{P}\left\{Y \leq g\left(Q_X(p)\right)\right\} = \mathbb{P}\left\{X \leq Q_X(p)\right\} = p, \tag{T1.8}$$

where the second to last equality follows from the invertibility of the function g, under the assumption that g is an increasing function of its argument. By applying the definition (1.17) of the quantile Q_Y to the leftmost and rightmost terms we obtain:

$$Q_Y(p) = g\left(Q_X(p)\right), \tag{T1.9}$$

In the case where g is a decreasing function of its argument we obtain similarly:

$$F_Y\left(g\left(Q_X(p)\right)\right) = \mathbb{P}\left\{Y \leq g\left(Q_X(p)\right)\right\} = \mathbb{P}\left\{X \geq Q_X(p)\right\} = 1 - p. \tag{T1.10}$$

By applying the definition (1.17) of the quantile Q_Y to the leftmost and rightmost terms we obtain:

$$Q_Y(p) = g\left(Q_X(1-p)\right). \tag{T1.11}$$

1.2 Distribution of positive affine transformations of random variables

Consider as a special case of $(T1.1)$ a generic positive affine transformation:

$$X \mapsto Y \equiv g(X) \equiv m + sX, \tag{T1.12}$$

where $s > 0$.

In this case we have:

$$g^{-1}(y) = \frac{y-m}{s}, \quad g'(x) = s. \tag{T1.13}$$

From $(T1.5)$ the probability density function of the transformed variable reads:

$$f_Y(y) = \frac{1}{s} f_X\left(\frac{y-m}{s}\right). \tag{T1.14}$$

From $(T1.6)$ the cumulative density function of the transformed variable reads:

$$F_Y(y) = F_X\left(\frac{y-m}{s}\right), \tag{T1.15}$$

From $(T1.9)$ the quantile of the transformed variable reads:

$$Q_Y(p) = m + sQ_X(p), \tag{T1.16}$$

In the case of affine transformations we can also compute the characteristic function of the transformed variable in terms of the original characteristic function. Indeed from the definition (1.12) of the characteristic function:

$$\begin{aligned} \phi_Y(\omega) &\equiv \mathrm{E}\left\{e^{i\omega Y}\right\} = \mathrm{E}\left\{e^{i\omega(m+sX)}\right\} \\ &= e^{i\omega m}\,\mathrm{E}\left\{e^{is\omega X}\right\}. \end{aligned} \tag{T1.17}$$

And therefore we obtain:

$$\phi_Y(\omega) = e^{i\omega m}\phi_X(s\omega). \tag{T1.18}$$

1.3 Distribution of the exponential of random variables

Consider as a special case of $(T1.1)$ the exponential transformation

$$X \mapsto Y \equiv g(X) \equiv e^X. \tag{T1.19}$$

In this case we have:

$$g^{-1}(y) = \ln(y), \quad g'(x) = e^x. \tag{T1.20}$$

From $(T1.5)$ the probability density function of the transformed variable reads:

$$f_Y(y) = \frac{1}{y} f_X(\ln(y)). \tag{$T1.21$}$$

From $(T1.6)$ the cumulative density function of the transformed variable reads:

$$F_Y(y) = F_X(\ln(y)). \tag{$T1.22$}$$

From $(T1.9)$ the quantile of the transformed variable reads:

$$Q_Y(p) = e^{Q_X(p)}. \tag{$T1.23$}$$

1.4 Affine equivariance of standard summary statistics

The expected value (1.25) is an affine equivariant parameter of location, i.e. it satisfies (1.22). Indeed

$$\begin{aligned} \mathrm{E}\{m+sX\} &\equiv \int_{\mathbb{R}} (m+sx) f_X(x)\,dx \qquad (T1.24) \\ &= m\int_{\mathbb{R}} f_X(x)\,dx + s\int_{\mathbb{R}} x f_X(x)\,dx \\ &\equiv m + s\,\mathrm{E}\{X\}, \end{aligned}$$

The median (1.26) is an affine equivariant parameter of location, i.e. it satisfies (1.22). Indeed, this is a specific case of $(T1.16)$ which states that the quantile, and thus in particular the median, is invariant with respect to any invertible transformation:

$$\begin{aligned} \mathrm{Med}\{m+sX\} &\equiv Q_{m+sX}\left(\frac{1}{2}\right) = m + sQ_X\left(\frac{1}{2}\right) \qquad (T1.25) \\ &\equiv m + s\,\mathrm{Med}\{X\}. \end{aligned}$$

By the same argument, the range (1.37) is an affine equivariant parameter of dispersion, i.e. it satisfies (1.32). Indeed if s is positive we obtain:

$$\begin{aligned} \mathrm{Ran}\{m+sX\} &\equiv Q_{m+sX}(\overline{p}) - Q_{m+sX}(\underline{p}) \\ &= [m + sQ_X(\overline{p})] - [m + sQ_X(\underline{p})] \qquad (T1.26) \\ &= s\left[Q_X(\overline{p}) - Q_X(\underline{p})\right] \\ &\equiv s\,\mathrm{Ran}\{X\} \end{aligned}$$

The reader can derive from $(T1.11)$ the proof in the case $s < 0$.

The mode (1.30) is an affine equivariant parameter of location, i.e. it satisfies (1.22). From $(T1.14)$ we know the density $f_Y(y)$ and therefore:

$$\begin{aligned}\mathrm{Mod}\{m+sX\} &\equiv \max_{y\in\mathbb{R}}\{f_{m+sX}(y)\} = \max_{y\in\mathbb{R}}\left\{\frac{1}{s}f_X\left(\frac{y-m}{s}\right)\right\} \\ &= \max_{y\in\mathbb{R}}\left\{f_X\left(\frac{y-m}{s}\right)\right\} = m + s\max_{x\in\mathbb{R}}\{f_X(x)\} \\ &\equiv m + s\,\mathrm{Mod}\{X\}.\end{aligned} \qquad (T1.27)$$

1.5 Expected value vs. median of symmetrical distributions

If a probability density function is symmetrical around $\widetilde{x}$, the area underneath the pdf on the half-line $(-\infty, \widetilde{x}]$ is the same as the area underneath the pdf on the half-line $[\widetilde{x}, +\infty)$, and thus they must both equal $1/2$. Thus from the definition of cumulative density function (1.7) we have:

$$F_X(\widetilde{x}) \equiv \int_{-\infty}^{\widetilde{x}} f_X(x)\,dx = \frac{1}{2}. \qquad (T1.28)$$

In turn, from the definition of the median (1.26) this implies that the symmetry point $\widetilde{x}$ is the median:

$$\mathrm{Med}\{X\} \equiv Q_X\left(\frac{1}{2}\right) = \widetilde{x}. \qquad (T1.29)$$

On the other hand for the expected value (1.25) we have:

$$\begin{aligned}\mathrm{E}\{X\} &\equiv \int_{\mathbb{R}} x f_X(x)\,dx = \widetilde{x} + \int_{\mathbb{R}} (x-\widetilde{x}) f_X(x)\,dx \\ &= \widetilde{x} + \int_{\mathbb{R}} u f_X(\widetilde{x}+u)\,du = \widetilde{x}.\end{aligned} \qquad (T1.30)$$

The last inequality follows since due to (1.28) we can write:

$$\int_{-\infty}^{0} u f_X(\widetilde{x}+u)\,du = -\int_{0}^{+\infty} u f_X(\widetilde{x}-u)\,du, \qquad (T1.31)$$

and thus the last integral in $(T1.30)$ is null.

1.6 Relation between characteristic function and moments

Assume that the characteristic function of a random variable X is analytical, i.e. it can be recovered entirely from its Taylor expansion. Then we can consider its expansion around the origin:

$$\phi_X(\omega) = 1 + (i\omega)\,\mathrm{RM}_1^X + \cdots + \frac{(i\omega)^k}{k!}\,\mathrm{RM}_k^X + \cdots . \qquad (T1.32)$$

In this expansion the generic coefficient RM_k^X is defined in terms of the derivatives of the characteristic function as follows:

$$\mathrm{RM}_k^X \equiv i^{-k} \left. \frac{d^k \phi_X(\omega)}{d\omega^k} \right|_{\omega=0} . \qquad (T1.33)$$

By performing the derivatives on the definition (1.12) of the characteristic function we obtain:

$$\frac{d^k}{d\omega^k}\{\phi_X(\omega)\} \equiv \frac{d^k}{d\omega^k}\left\{\int_{\mathbb{R}} e^{i\omega x} f_X(x)\,dx\right\} \qquad (T1.34)$$

$$= i^k \int_{\mathbb{R}} e^{i\omega x} x^k f_X(x)\,dx. \qquad (T1.35)$$

Therefore

$$\left. \frac{d^k \phi_X(\omega)}{d\omega^k} \right|_{\omega=0} = i^k \int_{\mathbb{R}} x^k f_X(x)\,dx = i^k\,\mathrm{E}\left\{X^k\right\}. \qquad (T1.36)$$

Substituting this in ($T1.33$) shows that RM_k^X is the k-th raw moment defined in (1.47):

$$\mathrm{RM}_k^X \equiv \mathrm{E}\left\{X^k\right\}. \qquad (T1.37)$$

Therefore any raw moment, and in particular the expected value of X, can be easily computed by differentiating the characteristic function.

On the other hand, the generic central moment of order k defined in (1.48) is a function of the raw moments of order up to k:

$$\mathrm{CM}_k^X = \sum_{j=0}^{k} \frac{k!\,(-1)^{k-j}}{j!\,(k-j)!}\,\mathrm{RM}_j^X \left(\mathrm{RM}_1^X\right)^{k-j}, \qquad (T1.38)$$

see e.g. Abramowitz and Stegun (1974) and Papoulis (1984). Therefore, after a few algebraic manipulations it is straightforward to derive the expression of any central moment. In particular, for the first moments we have:

$$\begin{aligned} \mathrm{CM}_2^X &= -\left(\mathrm{RM}_1^X\right)^2 + \mathrm{RM}_2^X \qquad (T1.39)\\ \mathrm{CM}_3^X &= 2\left(\mathrm{RM}_1^X\right)^3 - 3\left(\mathrm{RM}_1^X\right)\left(\mathrm{RM}_2^X\right) + \mathrm{RM}_3^X \\ \mathrm{CM}_4^X &= -3\left(\mathrm{RM}_1^X\right)^4 + 6\left(\mathrm{RM}_1^X\right)^2\left(\mathrm{RM}_2^X\right) - 4\,\mathrm{RM}_1^X\,\mathrm{RM}_3^X + \mathrm{RM}_4^X \end{aligned}$$

These expressions in turn allow to easily compute variance, standard deviation, skewness and kurtosis.

1.7 Histogram vs. pdf

Consider a set of T random i.i.d. variables

$$X_t \stackrel{d}{=} X, \quad t = 1, \ldots T. \tag{T1.40}$$

Consider the realizations of the above variables:

$$i_T \equiv \{x_1, \ldots, x_T\}. \tag{T1.41}$$

Consider the empirical distribution $\mathrm{Em}(i_T)$ stemming from the realization i_T, as defined in (1.119).

The Glivenko-Cantelli theorem (4.34) states that, under a few mild conditions, the empirical distribution converges to the true distribution of X as the number of observations T goes to infinity.

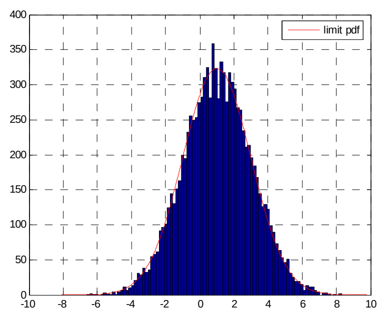

Fig. 1.1. Histogram vs. probability density function

In terms of the pdf, the Glivenko-Cantelli theorem reads:

$$f_{i_T} \equiv \frac{1}{T} \sum_{t=1}^{T} \delta^{(x_t)} \underset{T \to \infty}{\longrightarrow} f_X. \tag{T1.42}$$

Consider the histogram of the empirical pdf, where the width of all the bins is Δ. Denoting $\#_i^{\Delta}$ the number of points included in the generic i-th bin, the following relation holds:

$$\#_i^{\Delta} \equiv T \int_{x_i - \frac{\Delta}{2}}^{x_i + \frac{\Delta}{2}} f_{i_T}(y)\, dy \underset{T \to \infty}{\overset{\Delta \to 0}{\longrightarrow}} f_X(x_i)\, T\Delta. \tag{T1.43}$$

Therefore the histogram represents a regularized version of the true pdf, rescaled by the factor $T\Delta$. In Figure 1.1 we show the case of the normal distribution.

2

Technical appendix to Chapter 2

2.1 Distribution of grades

We recall that the grade of X is defined as follows:

$$U \equiv F_X(X), \tag{T2.1}$$

where F_X is the cumulative distribution function (1.7) of the variable X. To prove that

$$U \sim \mathrm{U}([0,1]), \tag{T2.2}$$

the standard uniform distribution defined in (1.54), we have to show that:

$$\mathbb{P}\{U \leq u\} = \begin{cases} 0 \text{ if } u \leq 0 \\ u \text{ if } u \in [0,1] \\ 1 \text{ if } u \geq 1. \end{cases} \tag{T2.3}$$

We first observe that by the definition of the cumulative function (1.7) the variable Y always lies in the interval $[0,1]$, therefore

$$\begin{aligned} \mathbb{P}\{U \leq u\} &= 0 \text{ if } u \leq 0 \\ \mathbb{P}\{U \leq u\} &= 1 \text{ if } u \geq 1. \end{aligned} \tag{T2.4}$$

As for the remaining cases, from the definition of the quantile function (1.17) we obtain:

$$\begin{aligned} \mathbb{P}\{U \leq u\} &= \mathbb{P}\{F_X(X) \leq u\} = \mathbb{P}\{X \leq Q_X(u)\} \\ &= F_X(Q_X(u)) = u. \end{aligned} \tag{T2.5}$$

This proves (T2.2).

On the other hand, if (T2.2) holds, then for any random variable X we have:

$$Q_X(U) \stackrel{d}{=} X, \tag{T2.6}$$

where $\stackrel{d}{=}$ means "has the same distribution as". Indeed:

$$\begin{aligned}\mathbb{P}\{Q_X(U) \leq x\} &= \mathbb{P}\{U \leq F_X(x)\} = \mathbb{P}\{F_X(X) \leq F_X(x)\} \\ &= \mathbb{P}\{X \leq x\}.\end{aligned} \quad (T2.7)$$

2.2 Distribution of invertible functions of random variables

Define a random variable $\mathbf{Y}$ as an invertible, *increasing* function $\mathbf{g}$ of a random variable $\mathbf{X}$

$$\mathbf{X} \mapsto \mathbf{Y} \equiv \mathbf{g}(\mathbf{X}), \quad (T2.8)$$

meaning that each entry $y_n \equiv g_n(\mathbf{x})$ is a non-decreasing function of any of the arguments $(x_1, \ldots, x_N)$. We compute the representations of the distribution of $\mathbf{Y}$ in terms of the representations of the distributions of $\mathbf{X}$.

- Probability density function

  From the definition of the pdf (2.4) we can write:

$$\begin{aligned}f_{\mathbf{Y}}(\mathbf{y})\, d\mathbf{y} &\equiv \mathbb{P}\{\mathbf{g}(\mathbf{X}) \in [\mathbf{y}, \mathbf{y} + d\mathbf{y}]\} \\ &= \mathbb{P}\left\{\mathbf{X} \in \left[\mathbf{g}^{-1}(\mathbf{y}), \mathbf{g}^{-1}(\mathbf{y} + d\mathbf{y})\right]\right\} \\ &= \int_{[\mathbf{g}^{-1}(\mathbf{y}), \mathbf{g}^{-1}(\mathbf{y}+d\mathbf{y})]} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}.\end{aligned} \quad (T2.9)$$

On the other hand, from a first order Taylor expansion we obtain:

$$\mathbf{g}^{-1}(\mathbf{y} + d\mathbf{y}) \approx \mathbf{g}^{-1}(\mathbf{y}) + \left[\mathbf{J}^{g}\left(\mathbf{g}^{-1}(\mathbf{y})\right)\right]^{-1} d\mathbf{y}, \quad (T2.10)$$

where the *Jacobian* $\mathbf{J}^{g}$ of a function $\mathbf{g}$ is defined as follows:

$$\mathbf{J}^{g}_{mn}(\mathbf{x}) \equiv \frac{\partial g_m(\mathbf{x})}{\partial x_n}. \quad (T2.11)$$

Therefore,

$$\begin{aligned}f_{\mathbf{Y}}(\mathbf{y})\, d\mathbf{y} &= \int_{\left[\mathbf{g}^{-1}(\mathbf{y}), \mathbf{g}^{-1}(\mathbf{y}) + [\mathbf{J}^{g}(\mathbf{g}^{-1}(\mathbf{y}))]^{-1} d\mathbf{y}\right]} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \\ &= f_{\mathbf{X}}\left(\mathbf{g}^{-1}(\mathbf{y})\right) \left|\left[\mathbf{J}^{g}\left(\mathbf{g}^{-1}(\mathbf{y})\right)\right]^{-1}\right| d\mathbf{y},\end{aligned} \quad (T2.12)$$

where the determinant accounts for the difference in volume between the infinitesimal parallelotope with sides $d\mathbf{y}$ and the infinitesimal parallelotope with sides $d\mathbf{x}$, see $(A.34)$. Therefore using $(A.83)$ we obtain the desired result:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}\left(\mathbf{g}^{-1}(\mathbf{y})\right)}{\left|\mathbf{J}^{g}\left(\mathbf{g}^{-1}(\mathbf{y})\right)\right|}. \quad (T2.13)$$

Notice that to compute the probability density function of the variable $\mathbf{Y}$, we do not need to assume that the function $\mathbf{g}$ be increasing. Indeed, as long as $\mathbf{g}$ is invertible, it suffices to replace the absolute value of the determinant in ($T2.12$). Thus in this slightly more general case we obtain:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}\left(\mathbf{g}^{-1}(\mathbf{y})\right)}{\sqrt{\left|\mathbf{J}^{g}\left(\mathbf{g}^{-1}(\mathbf{y})\right)\right|^{2}}}. \qquad (T2.14)$$

- Cumulative distribution function.

From the definition (2.9) of the cumulative distribution function F_Y we have:

$$\begin{aligned} F_{\mathbf{Y}}(\mathbf{y}) &\equiv \mathbb{P}\{\mathbf{Y} \leq \mathbf{y}\} = \mathbb{P}\{\mathbf{g}(\mathbf{X}) \leq \mathbf{y}\} \\ &= \mathbb{P}\left\{\mathbf{X} \leq \mathbf{g}^{-1}(\mathbf{y})\right\} \\ &= F_{\mathbf{X}}\left(\mathbf{g}^{-1}(\mathbf{y})\right). \end{aligned} \qquad (T2.15)$$

2.3 Results on copulas

As an application of ($T2.13$), we consider the random variable $\mathbf{U}$ defined by the following transformation

$$\mathbf{X} \mapsto \mathbf{U} \equiv \mathbf{g}(\mathbf{X}), \qquad (T2.16)$$

where $\mathbf{g}$ is defined component-wise in terms of the cdf F_{X_n} of the the generic n-th component X_n:

$$g_n(x_1, \ldots, x_N) \equiv F_{X_n}(x_n). \qquad (T2.17)$$

This is an invertible increasing transformation. From (1.17) the inverse of this transformation is the component-wise quantile:

$$g_n^{-1}(u_1, \ldots, u_N) \equiv Q_{X_n}(u_n), \qquad (T2.18)$$

By definition, the copula of $\mathbf{X}$ is the distribution of $\mathbf{U}$

Since the probability density function is the derivative of the cumulative distribution function, the Jacobian ($T2.11$) of the transformation reads:

$$\mathbf{J} = \operatorname{diag}(f_{X_1}, \ldots, f_{X_N}), \qquad (T2.19)$$

and thus from ($A.42$) its determinant is

$$|\mathbf{J}| = f_{X_1} \cdots f_{X_N}. \qquad (T2.20)$$

Therefore from ($T2.13$) the probability density function of the copula reads:

$$f_{\mathbf{U}}(\mathbf{u}) = \frac{f_{\mathbf{X}}\left(Q_{X_1}(u_1), \ldots, Q_{X_N}(u_N)\right)}{f_{X_1}\left(Q_{X_1}(u_1)\right) \cdot \cdots \cdot f_{X_N}\left(Q_{X_N}(u_N)\right)}. \qquad (T2.21)$$

As for the cumulative distribution function of the copula, from $(T2.15)$ we obtain:

$$F_{\mathbf{U}}(u_1, \ldots, u_N) = F_{\mathbf{X}}\left(Q_{X_1}(u_1), \ldots, Q_{X_N}(u_N)\right). \qquad (T2.22)$$

We now prove the invariance of the copula under a generic increasing transformation:

$$\mathbf{X} \mapsto \mathbf{Y} \equiv \mathbf{h}(\mathbf{X}). \qquad (T2.23)$$

From $(T2.22)$ the copulas of $\mathbf{X}$ and $\mathbf{Y}$ are the same if and only if the following is true:

$$F_{\mathbf{Y}}\left(Q_{Y_1}(u_1), \ldots, Q_{Y_N}(u_N)\right) = F_{\mathbf{X}}\left(Q_{X_1}(u_1), \ldots, Q_{X_N}(u_N)\right). \qquad (T2.24)$$

Indeed, from $(T2.15)$ we have:

$$F_{\mathbf{Y}}(y_1, \ldots, y_N) = F_{\mathbf{X}}\left(g_1^{-1}(y_1), \ldots, g_N^{-1}(y_N)\right). \qquad (T2.25)$$

On the other hand, the invariance property of the quantile $(T1.9)$, reads in this context as follows:

$$Q_{Yn}(u_n) = g_n\left(Q_{X_n}(u_n)\right). \qquad (T2.26)$$

Substituting $(T2.26)$ in $(T2.25)$ yields $(T2.24)$.

2.4 Distribution of affine transformations of a random variable

Consider a generic random variable $\mathbf{X}$ and the new random variable $\mathbf{Y}$ defined as an *invertible* affine transformation of $\mathbf{X}$

$$\mathbf{X} \mapsto \mathbf{Y} \equiv \mathbf{g}(\mathbf{X}) \equiv \mathbf{m} + \mathbf{B}\mathbf{X}, \qquad (T2.27)$$

where $\mathbf{a}$ is an N-dimensional vector and $\mathbf{B}$ is an invertible matrix.

In this case the Jacobian $(T2.11)$ is

$$\mathbf{J}^{\mathbf{g}} \equiv \mathbf{B}. \qquad (T2.28)$$

From $(A.82)$ and $(A.84)$ we see that

$$\left|\mathbf{J}^{\mathbf{g}}\right|^2 = |\mathbf{B}|^2 = |\mathbf{B}|\,|\mathbf{B}'| = \left|\mathbf{B}\mathbf{B}'\right|. \qquad (T2.29)$$

Therefore according to $(T2.14)$ the probability density function of $\mathbf{Y}$ reads:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}\left(\mathbf{B}^{-1}(\mathbf{y}-\mathbf{m})\right)}{\sqrt{|\mathbf{B}\mathbf{B}'|}}. \tag{T2.30}$$

In the case of affine transformations we can also compute the characteristic function of the transformed variable in terms of the original characteristic function. Indeed from the definition (2.13) of the characteristic function:

$$\begin{aligned}\phi_{\mathbf{Y}}(\boldsymbol{\omega}) &\equiv \mathrm{E}\left\{e^{i\boldsymbol{\omega}'\mathbf{Y}}\right\} = \mathrm{E}\left\{e^{i\boldsymbol{\omega}'(\mathbf{m}+\mathbf{B}\mathbf{X})}\right\} \\ &= e^{i\boldsymbol{\omega}'\mathbf{m}}\,\mathrm{E}\left\{e^{i(\mathbf{B}'\boldsymbol{\omega})'\mathbf{X}}\right\}.\end{aligned} \tag{T2.31}$$

And therefore we obtain:

$$\phi_{\mathbf{Y}}(\boldsymbol{\omega}) = e^{i\boldsymbol{\omega}'\mathbf{m}}\phi_{\mathbf{X}}(\mathbf{B}'\boldsymbol{\omega}). \tag{T2.32}$$

Consider now the case where the affine transformation is not invertible, i.e.

$$\operatorname{rank}(\mathbf{B}) \neq N \equiv \dim(\mathbf{X}). \tag{T2.33}$$

In particular, we consider linear combinations of random variables:

$$\Psi \equiv \mathbf{b}'\mathbf{X}. \tag{T2.34}$$

The distribution of Ψ is the marginal distribution of any invertible affine transformation that extends (T2.34):

$$\mathbf{Y} \equiv \begin{pmatrix} \Psi \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix} \equiv \mathbf{B}\mathbf{X}, \tag{T2.35}$$

For example, we can extend (T2.34) defining $\mathbf{B}$ as follows:

$$\mathbf{B} \equiv \begin{pmatrix} b_1 & (b_2, \ldots, b_N) \\ \mathbf{0}_{N-1} & \mathbf{I}_{N-1} \end{pmatrix}, \tag{T2.36}$$

where $\mathbf{0}_{N-1}$ is an $(N-1)$-dimensional column vector of zeros and $\mathbf{I}_{N-1}$ is the $(N-1)$-dimensional identity matrix.

The probability density function of (T2.34) is obtained by integrating out of (T2.30) the dependence on the ancillary variables (T2.35) as in (2.22):

$$\begin{aligned}f_{\Psi}(\psi) &= \int_{\mathbb{R}^{N-1}} f_{\mathbf{Y}}(\psi, y_2, \ldots, y_N)\, dy_2 \cdots dy_N \\ &= \frac{1}{\sqrt{|\mathbf{B}\mathbf{B}'|}} \int_{\mathbb{R}^{N-1}} f_{\mathbf{X}}\left(\mathbf{B}^{-1}\mathbf{y}\right) dy_2 \cdots dy_N.\end{aligned} \tag{T2.37}$$

Nevertheless, it is in general very difficult to perform this last step, as it involves a multiple integration. For instance, if $b_1 \neq 0$ we can choose the extension $\mathbf{B}$ according to ($T2.36$) we obtain:

$$\mathbf{B}^{-1} = \begin{pmatrix} \frac{1}{b_1} & -\frac{(b_2,\ldots,b_N)}{b_1} \\ \mathbf{0}_{N-1} & \mathbf{I}_{N-1} \end{pmatrix}, \tag{$T2.38$}$$

and thus from ($T2.37$) the probability density function of ($T2.34$) reads:

$$\begin{aligned} f_{\Psi}(\psi) = \frac{1}{\sqrt{b_1^2}} \int_{\mathbb{R}^{N-1}} f_{\mathbf{X}} \bigg(& \frac{\psi}{b_1} - \frac{b_2}{b_1} y_2 \\ & - \cdots - \frac{b_N}{b_1} y_N, y_2, \ldots, y_N \bigg) dy_2 \cdots dy_N . \end{aligned} \tag{$T2.39$}$$

On the other hand, the characteristic function of ($T2.34$) is obtained by setting to zero in ($T2.32$) the dependence on the ancillary variables ($T2.35$) as in (2.24):

$$\begin{aligned} \phi_{\psi}(\omega) &= \phi_{\mathbf{Y}}(\omega, \mathbf{0}_{N-1}) \\ &= \phi_{\mathbf{X}} \left(\mathbf{B}' \begin{pmatrix} \omega \\ \mathbf{0}_{N-1} \end{pmatrix} \right). \end{aligned} \tag{$T2.40$}$$

For instance, if we choose the extension $\mathbf{B}$ according to ($T2.36$) we obtain from ($T2.40$) that the characteristic function of ($T2.34$) reads:

$$\phi_{\psi}(\omega) = \phi_{\mathbf{X}}(\omega \mathbf{b}). \tag{$T2.41$}$$

2.5 Affine equivariance of mode and modal dispersion

Consider a generic invertible affine transformation

$$\mathbf{Y} = \mathbf{a} + \mathbf{B}\mathbf{X} \tag{$T2.42$}$$

of the N-dimensional random variable $\mathbf{X}$ as in ($T2.27$). From ($T2.30$) we derive the vector of the first order derivatives of the pdf of $\mathbf{Y}$ in terms of the pdf of $\mathbf{X}$:

$$\frac{\partial f_{\mathbf{Y}}(\mathbf{y})}{\partial \mathbf{y}} = \frac{(\mathbf{B}')^{-1}}{\sqrt{|\mathbf{B}\mathbf{B}'|}} \left. \frac{\partial f_{\mathbf{X}}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{B}^{-1}(\mathbf{y}-\mathbf{a})}. \tag{$T2.43$}$$

Deriving further, we obtain the matrix of the second order derivatives:

$$\frac{\partial^2 f_{\mathbf{Y}}(\mathbf{y})}{\partial \mathbf{y} \partial \mathbf{y}'} = \frac{(\mathbf{B}')^{-1}}{\sqrt{|\mathbf{B}\mathbf{B}'|}} \left. \frac{\partial^2 f_{\mathbf{X}}}{\partial \mathbf{x} \partial \mathbf{x}'} \right|_{\mathbf{x}=\mathbf{B}^{-1}(\mathbf{y}-\mathbf{a})} \mathbf{B}^{-1}. \tag{$T2.44$}$$

By its definition (2.52), the mode $\text{Mod}\{\mathbf{X}\}$ is the maximum. Thus it is determined by the following first order condition:

$$\left.\frac{\partial f_{\mathbf{X}}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}} = \mathbf{0}. \tag{$T2.45$}$$

First we prove that the mode satisfies is affine equivariant, i.e. it satisfies (2.51), which in this context reads:

$$\text{Mod}\{\mathbf{a}+\mathbf{B}\mathbf{X}\} = \mathbf{a}+\mathbf{B}\,\text{Mod}\{\mathbf{X}\}, \tag{$T2.46$}$$

Indeed, from ($T2.43$) and ($T2.45$) we obtain:

$$\left.\frac{\partial f_{\mathbf{Y}}}{\partial \mathbf{y}}\right|_{\mathbf{y}=\mathbf{a}+\mathbf{B}\,\text{Mod}\{\mathbf{X}\}} = \frac{(\mathbf{B}')^{-1}}{\sqrt{|\mathbf{B}\mathbf{B}'|}} \left.\frac{\partial f_{\mathbf{X}}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}} = \mathbf{0}. \tag{$T2.47$}$$

Since by the definition ($T2.45$) of mode we have

$$\left.\frac{\partial f_{\mathbf{Y}}}{\partial \mathbf{y}}\right|_{\mathbf{y}=\text{Mod}\{\mathbf{Y}\}} = \mathbf{0}, \tag{$T2.48$}$$

The result ($T2.46$) follows.

Now we prove that the modal dispersion (2.65) satisfies is affine equivariant, i.e. it satisfies (2.64) and it is symmetric and positive definite. From its definition, the modal dispersion of a generic random variable $\mathbf{X}$ reads:

$$\begin{aligned}
\text{MDis}\{\mathbf{X}\} &\equiv -\left(\left.\frac{\partial^2 \ln f_{\mathbf{X}}}{\partial \mathbf{x}\partial \mathbf{x}'}\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}}\right)^{-1}. \\
&= -\left(\left.\frac{\partial}{\partial \mathbf{x}}\left[\frac{1}{f_{\mathbf{X}}}\frac{\partial f_{\mathbf{X}}}{\partial \mathbf{x}'}\right]\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}}\right)^{-1} \\
&= -\left(\left.\frac{1}{f_{\mathbf{X}}}\frac{\partial^2 f_{\mathbf{X}}}{\partial \mathbf{x}\partial \mathbf{x}'} - \frac{1}{f_{\mathbf{X}}^2}\frac{\partial f_{\mathbf{X}}}{\partial \mathbf{x}}\frac{\partial f_{\mathbf{X}}}{\partial \mathbf{x}'}\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}}\right)^{-1} \\
&= -f_{\mathbf{X}}(\text{Mod}\{\mathbf{X}\}) \left.\frac{\partial^2 f_{\mathbf{X}}}{\partial \mathbf{x}\partial \mathbf{x}'}\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}}^{-1},
\end{aligned} \tag{$T2.49$}$$

Therefore from ($T2.44$) and ($T2.46$) we obtain:

$$\begin{aligned}
\text{MDis}\{\mathbf{Y}\} &= -f_{\mathbf{Y}}(\text{Mod}\{\mathbf{Y}\}) \left.\frac{\partial^2 f_{\mathbf{Y}}}{\partial \mathbf{y}\partial \mathbf{y}'}\right|_{\mathbf{y}=\text{Mod}\{\mathbf{Y}\}}^{-1} \\
&= -f_{\mathbf{Y}}(\text{Mod}\{\mathbf{Y}\}) \left[\frac{(\mathbf{B}')^{-1}}{\sqrt{|\mathbf{B}\mathbf{B}'|}} \left.\frac{\partial^2 f_{\mathbf{X}}}{\partial \mathbf{x}\partial \mathbf{x}'}\right|_{\mathbf{x}=\mathbf{B}^{-1}(\text{Mod}\{\mathbf{Y}\}-\mathbf{a})} \mathbf{B}^{-1}\right]^{-1} \\
&= -f_{\mathbf{Y}}(\text{Mod}\{\mathbf{Y}\}) \left[\frac{(\mathbf{B}')^{-1}}{\sqrt{|\mathbf{B}\mathbf{B}'|}} \left.\frac{\partial^2 f_{\mathbf{X}}}{\partial \mathbf{x}\partial \mathbf{x}'}\right|_{\mathbf{x}=\text{Mod}\{\mathbf{X}\}} \mathbf{B}^{-1}\right]^{-1}.
\end{aligned} \tag{$T2.50$}$$

Using ($T2.30$) and ($T2.46$) this expression becomes:

$$\begin{aligned}\mathrm{MDis}\{\mathbf{Y}\} &= -\frac{f_{\mathbf{X}}(\mathrm{Mod}\{\mathbf{X}\})}{\sqrt{|\mathbf{BB}'|}}\left[\frac{(\mathbf{B}')^{-1}}{\sqrt{|\mathbf{BB}'|}}\left.\frac{\partial^2 f_{\mathbf{X}}}{\partial\mathbf{x}\partial\mathbf{x}'}\right|_{\mathbf{x}=\mathrm{Mod}\{\mathbf{X}\}}\mathbf{B}^{-1}\right]^{-1} \quad (T2.51)\\ &= -f_{\mathbf{X}}(\mathrm{Mod}\{\mathbf{X}\})\,\mathbf{B}\left[\left.\frac{\partial^2 f_{\mathbf{X}}}{\partial\mathbf{x}\partial\mathbf{x}'}\right|_{\mathbf{x}=\mathrm{Mod}\{\mathbf{X}\}}\right]^{-1}\mathbf{B}'.\end{aligned}$$

Finally, using ($T2.49$) we obtain:

$$\mathrm{MDis}\{\mathbf{Y}\} = \mathbf{B}\,\mathrm{MDis}\{\mathbf{X}\}\,\mathbf{B}'. \quad (T2.52)$$

This proves that the modal dispersion is affine equivariant.

It is immediate to check from the definition (2.65) that the modal dispersion is a symmetric matrix. Furthermore, the mode is a maximum for the log-pdf, and therefore the matrix of the second derivatives of the log-pdf at the mode is negative definite. Therefore, the modal dispersion is positive definite. Affine equivariance, symmetry and positivity make the modal dispersion a scatter matrix.

2.6 Affine equivariance of expected value and covariance

Consider a generic affine transformation

$$\mathbf{X} \mapsto \widetilde{\mathbf{Y}} \equiv \widetilde{\mathbf{a}} + \widetilde{\mathbf{B}}\mathbf{X}, \quad (T2.53)$$

where $\widetilde{\mathbf{a}}$ is a K-dimensional vector and $\widetilde{\mathbf{B}}$ is a non-invertible $K \times N$ matrix.

First, we prove here the affine equivariance (2.51) of the expected value under generic affine transformations, i.e.

$$\mathrm{E}\left\{\widetilde{\mathbf{a}} + \widetilde{\mathbf{B}}\mathbf{X}\right\} = \widetilde{\mathbf{a}} + \widetilde{\mathbf{B}}\,\mathrm{E}\{\mathbf{X}\}. \quad (T2.54)$$

Adding $(N-K)$ non-collinear rows $\overline{\mathbf{B}}$ to $\widetilde{\mathbf{B}}$, $(N-K)$ elements $\overline{\mathbf{a}}$ to $\widetilde{\mathbf{a}}$ and denoting $\overline{\mathbf{Y}}$ a set of $(N-K)$ ancillary random variables as follows

$$\mathbf{Y} \equiv \begin{pmatrix}\widetilde{\mathbf{Y}}\\ \overline{\mathbf{Y}}\end{pmatrix}, \quad \mathbf{a} \equiv \begin{pmatrix}\widetilde{\mathbf{a}}\\ \overline{\mathbf{a}}\end{pmatrix}, \quad \mathbf{B} \equiv \begin{pmatrix}\widetilde{\mathbf{B}}\\ \overline{\mathbf{B}}\end{pmatrix}, \quad (T2.55)$$

we extend the transformation ($T2.53$) to an invertible affine transformation as in ($T2.27$):

$$\mathbf{X} \mapsto \mathbf{Y} \equiv \mathbf{a} + \mathbf{B}\mathbf{X}. \quad (T2.56)$$

From the definition of expected value and using ($T2.30$) we obtain:

$$\begin{aligned}
\mathrm{E}\left\{\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}\right\} &\equiv \int_{\mathbb{R}^K} \widetilde{\mathbf{y}} f_{\widetilde{\mathbf{Y}}}\left(\widetilde{\mathbf{y}}\right) d\widetilde{\mathbf{y}} = \int_{\mathbb{R}^N} \widetilde{\mathbf{y}} f_{\widetilde{\mathbf{Y}},\overline{\mathbf{Y}}}\left(\widetilde{\mathbf{y}},\overline{\mathbf{y}}\right) d\widetilde{\mathbf{y}} d\overline{\mathbf{y}} \\
&= \int_{\mathbb{R}^N} \widetilde{\mathbf{y}} f_{\mathbf{Y}}\left(\mathbf{y}\right) d\mathbf{y} = \int_{\mathbb{R}^N} \widetilde{\mathbf{y}} \frac{f_{\mathbf{X}}\left(\mathbf{B}^{-1}\left(\mathbf{y}-\mathbf{a}\right)\right)}{\sqrt{\left|\mathbf{B}\mathbf{B}'\right|}} d\mathbf{y} \qquad (T2.57) \\
&= \int_{\mathbb{R}^N} \left(\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{x}\right) \frac{f_{\mathbf{X}}\left(\mathbf{x}\right)}{\sqrt{\left|\mathbf{B}\mathbf{B}'\right|}} d\mathbf{y}.
\end{aligned}$$

With the change of variable $\mathbf{y} \equiv \mathbf{a} + \mathbf{B}\mathbf{x}$ we obtain:

$$\begin{aligned}
\mathrm{E}\left\{\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}\right\} &= \int_{\mathbb{R}^N} \left(\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{x}\right) \frac{f_{\mathbf{X}}\left(\mathbf{x}\right)}{\sqrt{\left|\mathbf{B}\mathbf{B}'\right|}} \left|\mathbf{B}\right| d\mathbf{x} \\
&= \widetilde{\mathbf{a}} + \widetilde{\mathbf{B}} \int_{\mathbb{R}^N} \mathbf{x} f_{\mathbf{X}}\left(\mathbf{x}\right) d\mathbf{x} \qquad (T2.58) \\
&= \widetilde{\mathbf{a}} + \widetilde{\mathbf{B}}\, \mathrm{E}\left\{\mathbf{X}\right\},
\end{aligned}$$

which proves ($T2.54$).

Now we prove here the affine equivariance (2.64) of the covariance matrix under generic affine transformations, i.e.

$$\mathrm{Cov}\left\{\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}\right\} = \widetilde{\mathbf{B}}\, \mathrm{Cov}\left\{\mathbf{X}\right\} \widetilde{\mathbf{B}}'. \qquad (T2.59)$$

From the definition of covariance (2.67) and the equivariance of the expected value ($T2.54$) we obtain:

$$\begin{aligned}
\mathrm{Cov}\left\{\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}\right\} &\equiv \mathrm{E}\left\{\left(\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}-\mathrm{E}\left\{\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}\right\}\right)\left(\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}-\mathrm{E}\left\{\widetilde{\mathbf{a}}+\widetilde{\mathbf{B}}\mathbf{X}\right\}\right)'\right\} \\
&= \mathrm{E}\left\{\widetilde{\mathbf{B}}\left(\mathbf{X}-\mathrm{E}\left\{\mathbf{X}\right\}\right)\left(\mathbf{X}-\mathrm{E}\left\{\mathbf{X}\right\}\right)' \widetilde{\mathbf{B}}'\right\} \qquad (T2.60) \\
&= \widetilde{\mathbf{B}}\, \mathrm{Cov}\left\{\mathbf{X}\right\} \widetilde{\mathbf{B}}',
\end{aligned}$$

where the last equality follows from the linearity of the expectation operator (B.56).

This proves that the covariance is an affine equivariant operator. It is immediate to check from the definition (2.68) that the covariance matrix is symmetric. Furthermore we proved in (B.68) that the covariance matrix is positive. Alternatively, from the affine equivariance ($T2.59$) we obtain:

$$\mathbf{a}'\, \mathrm{Cov}\left\{\mathbf{X}\right\} \mathbf{a} = \mathrm{Cov}\left\{\mathbf{a}'\mathbf{X}\right\} = \mathrm{Var}\left\{\mathbf{a}'\mathbf{X}\right\} \geq 0. \qquad (T2.61)$$

Affine equivariance, symmetry and positivity make the covariance a scatter matrix.

2.7 Regularized call and put option payoffs

The reader is advised to quickly review AppendixB.4 in the main text before going through the sequel. From the definition $(B.49)$ of regularization, the regularized profile of the call option is the convolution of the exact profile (2.36) with the approximate Dirac delta $(B.18)$. Therefore, from the definition of convolution $(B.43)$ we obtain

$$\begin{aligned} C_\epsilon(x) &\equiv \left[C * \delta_\epsilon^{(0)}\right](x) \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{+\infty} \max(y-K,0)\, e^{-\frac{1}{2\epsilon^2}(x-y)^2} dy \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{K}^{+\infty} (y-K)\, e^{-\frac{1}{2\epsilon^2}(y-x)^2} dy \qquad (T2.62) \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{K-x}^{+\infty} (u+x-K)\, e^{-\frac{1}{2\epsilon^2}u^2} du \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{K-x}^{+\infty} u e^{-\frac{1}{2\epsilon^2}u^2} du + \frac{1}{\sqrt{2\pi}\epsilon}(x-K) \int_{K-x}^{+\infty} e^{-\frac{1}{2\epsilon^2}u^2} du \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{K-x}^{+\infty} \frac{d}{du}\left[-\epsilon^2 e^{-\frac{1}{2\epsilon^2}u^2}\right] du + \frac{(x-K)}{2}\left[\frac{2}{\sqrt{\pi}} \int_{\frac{K-x}{\sqrt{2\epsilon^2}}}^{+\infty} e^{-z^2} dz\right], \end{aligned}$$

where in the last line we performed the change of variable $u/\sqrt{2\epsilon^2} \equiv z$. Using the relation $(B.78)$ between the complementary error function and the error function, as well as $(B.76)$, i.e. the fact that the error function is odd, we obtain:

$$C_\epsilon(x) = \frac{(x-K)}{2}\left(1 + \operatorname{erf}\left(\frac{x-K}{\sqrt{2\epsilon^2}}\right)\right) + \frac{\epsilon}{\sqrt{2\pi}} e^{-\frac{1}{2\epsilon^2}(x-K)^2}. \qquad (T2.63)$$

Similarly, for the put option (2.113) we obtain:

$$\begin{aligned} P_\epsilon(x) &\equiv \left[P * \delta_\epsilon^{(0)}\right](x) \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{+\infty} -\min(y-K,0)\, e^{-\frac{1}{2\epsilon^2}(x-y)^2} dy \\ &= -\frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{K} (y-K)\, e^{-\frac{1}{2\epsilon^2}(y-x)^2} dy \qquad (T2.64) \\ &= -\frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{K-x} (u+x-K)\, e^{-\frac{1}{2\epsilon^2}u^2} du \\ &= -\frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{K-x} u e^{-\frac{1}{2\epsilon^2}u^2} du - \frac{1}{\sqrt{2\pi}\epsilon}(x-K) \int_{-\infty}^{K-x} e^{-\frac{1}{2\epsilon^2}u^2} du \\ &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{K-x} \frac{d}{du}\left[\epsilon^2 e^{-\frac{1}{2\epsilon^2}u^2}\right] du - \frac{(x-K)}{2}\left[\frac{2}{\sqrt{\pi}} \int_{-\infty}^{\frac{K-x}{\sqrt{2\epsilon^2}}} e^{-z^2} dz\right], \end{aligned}$$

where in the last line we performed the change of variable $u/\sqrt{2\epsilon^2} \equiv z$. Using the definition of the error function (B.75) and the fact that it is an odd function we obtain:

$$P_\epsilon(x) = -\frac{(x-K)}{2}\left(1 - \operatorname{erf}\left(\frac{x-K}{\sqrt{2\epsilon^2}}\right)\right) + \frac{\epsilon}{\sqrt{2\pi}} e^{-\frac{1}{2\epsilon^2}(x-K)^2}. \qquad (T2.65)$$

2.8 The equation of the enshrouding rectangle

We prove that for a generic $n = 1, \ldots, N$ the two hyperplanes described by the following equation

$$x_n = \mathrm{E}\{X_n\} \pm \mathrm{Sd}\{X_n\} \qquad (T2.66)$$

are tangent to the ellipsoid.$\mathcal{E}_{\mathrm{E},\mathrm{Cov}}$ defined in (2.75).

First we consider its implicit representation

$$g(\mathbf{x}) = 0, \qquad (T2.67)$$

where from (2.75) the function g is defined as follows:

$$g(\mathbf{x}) \equiv (\mathbf{x} - \mathrm{E})' \operatorname{Cov}^{-1} (\mathbf{x} - \mathrm{E}) - 1. \qquad (T2.68)$$

To find the tangency condition of the ellipsoid with the rectangle we compute the gradient of the implicit representation of $\mathcal{E}_{\mathrm{E},\mathrm{Cov}}$

$$\frac{\partial g}{\partial \mathbf{x}} = 2 \operatorname{Cov}^{-1} (\mathbf{x} - \mathrm{E}). \qquad (T2.69)$$

Since the generic n-th side of the rectangle is perpendicular to the n-th axis, when the gradient is parallel to the n-th axis the rectangle is tangent to the ellipsoid. Therefore, to find the tangency condition we must impose the following condition:

$$\operatorname{Cov}^{-1} (\mathbf{x} - \mathrm{E}) = \alpha \boldsymbol{\delta}^{(n)}, \qquad (T2.70)$$

where α is some scalar that we have to compute and $\boldsymbol{\delta}^{(n)}$ is the n-th element of the canonical basis of $\mathbb{R}^N$, see (A.15). To compute α we substitute (T2.70) in the implicit equation (T2.67) of the ellipsoid:

$$\begin{aligned} 1 &= (\mathbf{x} - \mathrm{E})' \operatorname{Cov}^{-1} (\mathbf{x} - \mathrm{E}) \\ &= \left(\alpha \operatorname{Cov} \boldsymbol{\delta}^{(n)}\right)' \operatorname{Cov}^{-1} \left(\alpha \operatorname{Cov} \boldsymbol{\delta}^{(n)}\right) \\ &= \alpha^2 \operatorname{Var}\{X_n\}, \end{aligned} \qquad (T2.71)$$

so that

$$\alpha = \pm \frac{1}{\mathrm{Sd}\{X_n\}}. \qquad (T2.72)$$

Substituting ($T2.72$) back in ($T2.70$) and then again in ($T2.67$) yields

$$\begin{aligned} 1 &= (\mathbf{x} - \mathrm{E})' \operatorname{Cov}^{-1} (\mathbf{x} - \mathrm{E}) \\ &= (\mathbf{x} - \mathrm{E})' \left(\pm \frac{1}{\operatorname{Sd}\{X_n\}} \boldsymbol{\delta}^{(n)} \right) = \pm \frac{x_n - \mathrm{E}\{X_n\}}{\operatorname{Sd}\{X_n\}}, \end{aligned} \tag{T2.73}$$

which proves ($T2.66$).

2.9 Chebyshev's inequality

Consider a generic vector $\mathbf{v}$ and a generic symmetric and positive matrix $\mathbf{U}$. These define an ellipsoid $\mathcal{E}^q_{\mathbf{v},\mathbf{U}}$ as in (2.87). Therefore

$$\begin{aligned} q^2 \mathbb{P}\left\{ \mathbf{X} \notin \mathcal{E}^q_{\mathbf{v},\mathbf{U}} \right\} &= \int_{\mathbb{R}^N / \mathcal{E}^q_{\mathbf{v},\mathbf{U}}} q^2 f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \\ &\leq \int_{\mathbb{R}^N / \mathcal{E}^q_{\mathbf{v},\mathbf{U}}} (\mathbf{x} - \mathbf{v})' \mathbf{U}^{-1} (\mathbf{x} - \mathbf{v}) f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \\ &\leq \int_{\mathbb{R}^n} (\mathbf{x} - \mathbf{v})' \mathbf{U}^{-1} (\mathbf{x} - \mathbf{v}) f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \\ &= \mathrm{E}\left\{ (\mathbf{x} - \mathbf{v})' \mathbf{U}^{-1} (\mathbf{x} - \mathbf{v}) \right\} \\ &\equiv a(\mathbf{v}, \mathbf{U}). \end{aligned} \tag{T2.74}$$

Notice that we can re-write $a(\mathbf{v}, \mathbf{U})$ as follows:

$$a(\mathbf{v}, \mathbf{U}) = \operatorname{tr}\left(\mathrm{E}\left\{ (\mathbf{X} - \mathbf{v})(\mathbf{X} - \mathbf{v})' \right\} \mathbf{U}^{-1} \right). \tag{T2.75}$$

From this we obtain:

$$a(\mathrm{E}, \operatorname{Cov}) = \operatorname{tr}\left(\operatorname{Cov} \operatorname{Cov}^{-1} \right) = \operatorname{tr}(\mathbf{I}_N) = N. \tag{T2.76}$$

Now we prove that the minimum of ($T2.75$) is ($T2.76$). In other words, among all possible vectors $\mathbf{v}$ and symmetric, positive matrices $\mathbf{U}$ such that

$$|\mathbf{U}| = |\operatorname{Cov}\{\mathbf{X}\}| \tag{T2.77}$$

the minimum value of ($T2.75$) is achieved by the choice $\mathbf{v} \equiv \mathrm{E}\{\mathbf{X}\}$ and $\mathbf{U} \equiv \operatorname{Cov}\{\mathbf{X}\}$.

Consider an arbitrary vector $\mathbf{u}$ and a perturbation

$$\mathbf{v} \mapsto \mathbf{v} + \eta \mathbf{u} \tag{T2.78}$$

If $\mathbf{v}$ minimizes ($T2.75$), in the limit $\eta \to 0$ we must have:

$$\begin{aligned} 0 &= \mathrm{E}\left\{ (\mathbf{X} - (\mathbf{v} + \eta \mathbf{u}))' \mathbf{U}^{-1} (\mathbf{X} - (\mathbf{v} + \eta \mathbf{u})) \right\} \\ &\quad - \mathrm{E}\left\{ (\mathbf{X} - \mathbf{v})' \mathbf{U}^{-1} (\mathbf{X} - \mathbf{v}) \right\} \\ &\approx -2\eta\, \mathrm{E}\left\{ \mathbf{u}' \mathbf{U}^{-1} (\mathbf{x} - \mathbf{v}) \right\} = -2\eta \mathbf{u}' \mathbf{U}^{-1} (\mathrm{E}\{\mathbf{X}\} - \mathbf{v}), \end{aligned} \tag{T2.79}$$

and therefore we must have

$$\mathbf{v} \equiv \mathrm{E}\{\mathbf{X}\} \tag{$T2.80$}$$

Now consider an arbitrary perturbation

$$\mathbf{U} \mapsto \mathbf{U}(\mathbf{I} + \epsilon\mathbf{B}), \tag{$T2.81$}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{B}$ is a matrix that preserves the volumes. From (A.77) this means:

$$|\mathbf{U}(\mathbf{I} + \epsilon\mathbf{B})| = |\mathbf{U}|. \tag{$T2.82$}$$

In the limit of small perturbations $\epsilon \to 0$, from (A.122) that this condition becomes:

$$\mathrm{tr}(\mathbf{B}) = 0. \tag{$T2.83$}$$

If $(\mathrm{E}\{\mathbf{X}\}, \mathbf{U})$ minimize (T2.75), in the limit $\epsilon \to 0$ we must have

$$\begin{aligned} 0 &= \mathrm{E}\left\{(\mathbf{X} - \mathrm{E}\{\mathbf{X}\})' [\mathbf{U}(\mathbf{I} + \epsilon\mathbf{B})]^{-1} (\mathbf{X} - \mathrm{E}\{\mathbf{X}\})\right\} \\ &\quad - \mathrm{E}\left\{(\mathbf{X} - \mathrm{E}\{\mathbf{X}\})' \mathbf{U}^{-1} (\mathbf{X} - \mathrm{E}\{\mathbf{X}\})\right\} \\ &= \mathrm{tr}\left(\mathrm{Cov}\{\mathbf{X}\} [\mathbf{U}(\mathbf{I} + \epsilon\mathbf{B})]^{-1}\right) - \mathrm{tr}\left(\mathrm{Cov}\{\mathbf{X}\}\mathbf{U}^{-1}\right) \\ &\equiv \mathrm{tr}\left(\mathrm{Cov}\{\mathbf{X}\}(\mathbf{I} - \epsilon\mathbf{B})\mathbf{U}^{-1}\right) - \mathrm{tr}\left(\mathrm{Cov}\{\mathbf{X}\}\mathbf{U}^{-1}\right) \\ &= -\epsilon\,\mathrm{tr}\left(\mathrm{Cov}\{\mathbf{X}\}\mathbf{B}\mathbf{U}^{-1}\right) \\ &= -\epsilon\,\mathrm{tr}\left(\mathbf{B}\mathbf{U}^{-1}\mathrm{Cov}\{\mathbf{X}\}\right) \end{aligned} \tag{$T2.84$}$$

To summarize, from (T2.83) and (T2.84) we must have:

$$\mathrm{tr}(\mathbf{B}) = 0 \Rightarrow \mathrm{tr}\left(\mathbf{B}\mathbf{U}^{-1}\mathrm{Cov}\{\mathbf{X}\}\right) = 0, \tag{$T2.85$}$$

which is only true if $\mathbf{U}$ is proportional to the covariance, i.e.

$$\mathbf{U} = \alpha\,\mathrm{Cov}\{\mathbf{X}\}, \tag{$T2.86$}$$

for some scalar α. Given the normalization (T2.77) we obtain the desired result.

2.10 Relation between characteristic function and moments

Assume that the characteristic function of a random variable $\mathbf{X}$ is analytical, i.e. it can be recovered entirely from its Taylor expansion. Then we can consider its expansion around zero:

$$\phi_{\mathbf{X}}(\boldsymbol{\omega}) = 1 + i \sum_{n=1}^{N} \omega_n \, \mathrm{RM}_n^{\mathbf{X}} + \cdots . \qquad (T2.87)$$
$$+ \frac{i^k}{k!} \sum_{n_1,\ldots,n_k=1}^{N} (\omega_{n_1} \cdots \omega_{n_k}) \, \mathrm{RM}_{n_1 \cdots n_k}^{\mathbf{X}} + \cdots$$

In this expansion the generic coefficient $\mathrm{RM}_{n_1 \cdots n_k}^{\mathbf{X}}$ is defined in terms of the derivatives of the characteristic function as follows:

$$\mathrm{RM}_{n_1 \cdots n_k}^{\mathbf{X}} \equiv i^{-k} \left. \frac{\partial^k \phi_{\mathbf{X}}(\boldsymbol{\omega})}{\partial \omega_{n_1} \cdots \partial \omega_{n_k}} \right|_{\omega=0} . \qquad (T2.88)$$

By performing the derivatives on the definition (2.13) of the characteristic function we obtain:

$$\frac{\partial^k}{\partial \omega_{n_1} \cdots \partial \omega_{n_k}} \{\phi_{\mathbf{X}}(\boldsymbol{\omega})\} \equiv \frac{\partial^k}{\partial \omega_{n_1} \cdots \partial \omega_{n_k}} \left\{ \int_{\mathbb{R}^N} e^{i\boldsymbol{\omega}'\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} \right\} \qquad (T2.89)$$
$$= i^k \int_{\mathbb{R}^N} x_{n_1} \cdots x_{n_k} e^{i\boldsymbol{\omega}'\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x}.$$

Therefore

$$\left. \frac{\partial^k \phi_{\mathbf{X}}(\boldsymbol{\omega})}{\partial \omega_{n_1} \cdots \partial \omega_{n_k}} \right|_{\boldsymbol{\omega}=\mathbf{0}} = i^k \int_{\mathbb{R}^N} x_{n_1} \cdots x_{n_k} f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} \qquad (T2.90)$$
$$= i^k \, \mathrm{E}\{X_{n_1} \cdots X_{n_k}\} .$$

Substituting this in (T2.88) shows that $\mathrm{RM}_{n_1 \cdots n_k}^{\mathbf{X}}$ is the k-th raw moment defined in (2.91):

$$\mathrm{RM}_{n_1 \cdots n_k}^{\mathbf{X}} \equiv \mathrm{E}\{X_{n_1} \cdots X_{n_k}\} . \qquad (T2.91)$$

Therefore any raw moment can be easily computed by differentiating the characteristic function. In particular, from (T2.88) we obtain the expected value, which is the raw moment of order one:

$$\mathrm{E}\{X_n\} = \mathrm{RM}_n^{\mathbf{X}} = \frac{1}{i} \left. \frac{\partial \phi_{\mathbf{X}}(\boldsymbol{\omega})}{\partial \omega_n} \right|_{\omega=0} . \qquad (T2.92)$$

On the other hand, the k-th central moment

$$\mathrm{CM}_{n_1 \cdots n_k}^{\mathbf{X}} \equiv \mathrm{E}\{(X_{n_1} - \mathrm{E}\{X_{n_1}\}) \cdots (X_{n_k} - \mathrm{E}\{X_{n_k}\})\} . \qquad (T2.93)$$

is a function of the raw moments of order up to k, a generalization of (T1.38). Similarly k-th raw moment is a function of the central moments of order up to k. These statements follow by expanding the products in (T2.93) and inverting the ensuing triangular transformation. See also David and Barton (1962) for an interesting approach based on differentiation.

In particular for the covariance matrix, which is the central moment of order two, we obtain:

$$\begin{aligned} \text{Cov}\{X_m, X_n\} &= \text{CM}^{\mathbf{X}}_{mn} \\ &= \text{RM}^{\mathbf{X}}_{mn} - \text{RM}^{\mathbf{X}}_{m} \, \text{RM}^{\mathbf{X}}_{n} \,, \end{aligned} \tag{T2.94}$$

where the second raw moment follows from $(T2.88)$:

$$\text{RM}^{\mathbf{X}}_{mn} = - \left. \frac{\partial^2 \phi_{\mathbf{X}}(\boldsymbol{\omega})}{\partial \omega_m \partial \omega_n} \right|_{\omega=0} . \tag{T2.95}$$

2.11 Results on the uniform distribution

Assume that the random variable $\mathbf{X}$ is uniformly distributed on the unit sphere in $\mathbb{R}^N$:

$$\mathbf{X} \sim \text{U}\left(\mathcal{E}_{\mathbf{0}_N, \mathbf{I}_N}\right) . \tag{T2.96}$$

The probability density function of $\mathbf{X}$ reads:

$$\frac{1}{V_N} \mathbb{I}_{\mathcal{E}_{\mathbf{0},\mathbf{I}}}(\mathbf{x}) \,, \tag{T2.97}$$

where V_N is the volume of the unit sphere in $\mathbb{R}^N$:

$$V_N \equiv \frac{\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2} + 1\right)}, \tag{T2.98}$$

see Fang, Kotz, and Ng (1990), p. 74.

In Fang, Kotz, and Ng (1990), p. 75 we find the expression of the marginal probability density function of the last $(N-K)$ entries of $(T2.96)$ which reads:

$$f(x_{K+1}, \ldots, x_N) = \frac{\Gamma\left(\frac{N+2}{2}\right)}{\Gamma\left(\frac{K+2}{2}\right) \pi^{\frac{N-K}{2}}} \left(1 - \sum_{n=K+1}^{N} x_n^2\right)^{\frac{K}{2}} , \tag{T2.99}$$

where Γ is the gamma function $(B.80)$ and

$$\sum_{n=K+1}^{N} x_n^2 \leq 1. \tag{T2.100}$$

Therefore, the marginal distribution is not uniform. Notice that (2.151), which we computed explicitly, is a special case of $(T2.99)$, as follows immediately from $(B.81)$ and $(B.82)$.

To compute the characteristic function, using the result (2.9) in Fang, Kotz, and Ng (1990) and $(T2.99)$ the characteristic function of $(T2.96)$ reads:

$$\begin{aligned}\phi(\boldsymbol{\omega}) &\equiv \mathrm{E}\left\{e^{i\boldsymbol{\omega}'\mathbf{X}}\right\} = \mathrm{E}\left\{e^{i\sqrt{\boldsymbol{\omega}'\boldsymbol{\omega}}X_N}\right\} \\ &= \int_{-\infty}^{+\infty} e^{i\sqrt{\boldsymbol{\omega}'\boldsymbol{\omega}}x_N} f(x_N)\,dx_N \\ &= \frac{\Gamma\left(\frac{N+2}{2}\right)}{\Gamma\left(\frac{N+1}{2}\right)\pi^{\frac{1}{2}}} \int_{-1}^{+1} \cos\left(\sqrt{\boldsymbol{\omega}'\boldsymbol{\omega}}x\right)\left(1-x^2\right)^{\frac{N-1}{2}} dx \\ &\quad + i\frac{\Gamma\left(\frac{N+2}{2}\right)}{\Gamma\left(\frac{N+1}{2}\right)\pi^{\frac{1}{2}}} \int_{-\infty}^{+\infty} \sin\left(\sqrt{\boldsymbol{\omega}'\boldsymbol{\omega}}x\right)\left(1-x^2\right)^{\frac{N-1}{2}} dx.\end{aligned} \tag{T2.101}$$

The last term vanishes due to the symmetry of $\left(1-x^2\right)$ around the origin. From $(B.89)$ and $(B.82)$ we have:

$$B\left(\frac{1}{2},\frac{N+1}{2}\right) = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)} = \frac{\sqrt{\pi}\,\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)}. \tag{T2.102}$$

Therefore the characteristic function reads:

$$\phi(\boldsymbol{\omega}) = \frac{2}{B\left(\frac{1}{2},\frac{N+1}{2}\right)} \int_0^{+1} \cos\left(\sqrt{\boldsymbol{\omega}'\boldsymbol{\omega}}x\right)\left(1-x^2\right)^{\frac{N-1}{2}} dx. \tag{T2.103}$$

To compute the moments, we represent $\mathbf{X}$ as follows:

$$\mathbf{X} = R\mathbf{U}, \tag{T2.104}$$

where from (2.259) $R \equiv \|\mathbf{X}\|$ and $\mathbf{U} \equiv \mathbf{X}/\|\mathbf{X}\|$ are independent and $\mathbf{U}$ is uniformly distributed on the *surface* of the unit ball $\mathcal{E}_{\mathbf{0}_N,\mathbf{I}_N}$. Therefore from $(T2.207)$ we obtain:

$$\mathrm{E}\{\mathbf{X}\} = \mathrm{E}\{R\}\,\mathrm{E}\{\mathbf{U}\} = \mathbf{0}. \tag{T2.105}$$

Similarly:

$$\mathrm{Cov}\{\mathbf{X}\} = \mathrm{E}\left\{R^2\mathbf{U}\mathbf{U}'\right\} = \mathrm{E}\left\{R^2\right\}\mathrm{Cov}\{\mathbf{U}\}. \tag{T2.106}$$

From Fang, Kotz, and Ng (1990), p. 75, the pdf of $R \equiv \|\mathbf{X}\|$ reads:

$$f_R(r) = N r^{N-1}\mathbb{I}_{[0,1]}(r), \tag{T2.107}$$

where $\mathbb{I}$ is the indicator function $(B.72)$. Therefore:

$$\mathrm{E}\left\{R^k\right\} = \int_0^1 r^k N r^{N-1} dr = N\int_0^1 r^{N+k-1} dr = \frac{N}{N+k}. \tag{T2.108}$$

Therefore, using $(T2.208)$ we obtain:

$$\mathrm{Cov}\{\mathbf{X}\} = \frac{\mathbf{I}_N}{N+2}. \tag{T2.109}$$

More in general, we can obtain any moment by applying $(T2.212)$

$$\begin{aligned}\mathrm{CM}^{\mathbf{X}}_{m_1\cdots m_k} &= \mathrm{CM}^{R\mathbf{U}}_{m_1\cdots m_k} = \mathrm{RM}^{R\mathbf{U}}_{m_1\cdots m_k} \\ &= \mathrm{E}\left\{RU_{n_1}\cdots RU_{n_k}\right\} \\ &= \mathrm{E}\left\{R^k\right\}\mathrm{E}\left\{U_{n_1}\cdots U_{n_k}\right\} \\ &= \frac{N}{N+k}\mathrm{E}\left\{U_{n_1}\cdots U_{n_k}\right\},\end{aligned} \qquad (T2.110)$$

and then using (T2.206).

With the transformation:

$$\mathbf{X} \mapsto \mathbf{Y} \equiv \boldsymbol{\mu} + \mathbf{B}\mathbf{X}, \qquad (T2.111)$$

where $\mathbf{B}\mathbf{B}' \equiv \boldsymbol{\Sigma}$ we obtain a variable $\mathbf{Y}$ that is uniformly distributed on the ellipsoid $\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$:

$$\mathbf{Y} \sim \mathrm{U}\left(\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\right). \qquad (T2.112)$$

Therefore, the probability density function of the is obtained by applying (T2.13) to (T2.97). Similarly, the characteristic function of the uniform distribution on the ellipsoid $\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ is obtained by applying (T2.32) to (T2.103). Notice that there is a typo in Fang, Kotz, and Ng (1990). The expected value of the uniform distribution on the ellipsoid $\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ is obtained by applying (2.56) to (T2.105) and the covariance is obtained by applying (2.71) to (T2.109).

2.12 Results on the normal distribution

Characteristic function of the normal distribution

Consider a univariate standard normal variable:

$$X \sim \mathrm{N}(0,1). \qquad (T2.113)$$

Its characteristic function reads:

$$\begin{aligned}\phi(\omega) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega x}e^{-\frac{x^2}{2}}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left(x^2-2i\omega x\right)}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\left[(x-i\omega)^2+\omega^2\right]}dx \\ &= e^{-\frac{1}{2}\omega^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-i\omega)^2}d(x-i\omega) \\ &= e^{-\frac{1}{2}\omega^2}\end{aligned} \qquad (T2.114)$$

Consider now a set of N independent standard normal variables $\mathbf{X}$. By definition, their juxtaposition is a standard N-dimensional normal random vector:

$$\mathbf{X} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}). \qquad (T2.115)$$

Therefore:

$$\begin{aligned} \phi(\boldsymbol{\omega}) &\equiv \mathrm{E}\left\{e^{i\boldsymbol{\omega}'\mathbf{X}}\right\} = \prod_{n=1}^{N} \mathrm{E}\left\{e^{i\omega_n X_n}\right\} \\ &= \prod_{n=1}^{N} e^{-\frac{1}{2}\omega_n^2} = e^{-\frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\omega}}. \end{aligned} \qquad (T2.116)$$

With the transformation:

$$\mathbf{X} \mapsto \mathbf{Y} \equiv \boldsymbol{\mu} + \mathbf{B}\mathbf{X}, \qquad (T2.117)$$

where $\mathbf{B}\mathbf{B}' \equiv \boldsymbol{\Sigma}$ we obtain a generic multivariate normal random vector:

$$\mathbf{Y} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \qquad (T2.118)$$

The characteristic function of ($T2.118$) is obtained by applying ($T2.32$) to ($T2.116$).

Probability density function of the copula of the bivariate normal distribution

From (2.30), the pdf of the copula reads

$$f^{\mathrm{N}}(u_1, u_2) = \frac{f^{\mathrm{N}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(Q^{\mathrm{N}}_{\mu_1,\sigma_1^2}(u_1), Q^{\mathrm{N}}_{\mu_2,\sigma_2^2}(u_2)\right)}{f^{\mathrm{N}}_{\mu_1,\sigma_1^2}\left(Q^{\mathrm{N}}_{\mu_1,\sigma_1^2}(u_1)\right) f^{\mathrm{N}}_{\mu_2,\sigma_2^2}\left(Q^{\mathrm{N}}_{\mu_2,\sigma_2^2}(u_2)\right)}, \qquad (T2.119)$$

where Q is the quantile (1.70) of the marginal normal one-dimensional distribution:

$$Q^{\mathrm{N}}_{\mu,\sigma^2}(u) = \mu + \sqrt{2\sigma^2}\,\mathrm{erf}^{-1}(2u-1). \qquad (T2.120)$$

From the expression (2.170) of the two dimensional joint normal pdf $f^{\mathrm{N}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ we obtain:

$$f^{\mathrm{N}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(Q^{\mathrm{N}}_{\mu_1,\sigma_1^2}(u_1), Q^{\mathrm{N}}_{\mu_2,\sigma_2^2}(u_2)\right) = \frac{\left(\sigma_1^2\sigma_2^2\left(1-\rho^2\right)\right)^{-\frac{1}{2}}}{2\pi} e^{-\frac{1}{2}\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{(1-\rho^2)}}, \qquad (T2.121)$$

where

$$z_i \equiv \sqrt{2}\,\mathrm{erf}^{-1}(2u_i - 1). \qquad (T2.122)$$

On the other hand, from the expression (1.67) of the marginal pdf we obtain:

$$f^{\mathrm{N}}_{\mu_i,\sigma_i^2}\left(Q^{\mathrm{N}}_{\mu_i,\sigma_i^2}(u_i)\right) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}} e^{-\frac{z_i^2}{2}}. \qquad (T2.123)$$

Therefore

$$f^{\mathrm{N}}(u_1,u_2)=\frac{1}{\sqrt{1-\rho^2}}\exp\left(g_\rho(u_1,u_2)\right) \qquad (T2.124)$$

where

$$g_\rho(u_1,u_2)\equiv-\begin{pmatrix}\operatorname{erf}^{-1}(2u_1-1)\\ \operatorname{erf}^{-1}(2u_2-1)\end{pmatrix}'\left(\begin{pmatrix}1&\rho\\ \rho&1\end{pmatrix}^{-1}-\begin{pmatrix}1&0\\ 0&1\end{pmatrix}\right)\begin{pmatrix}\operatorname{erf}^{-1}(2u_1-1)\\ \operatorname{erf}^{-1}(2u_2-1)\end{pmatrix} \qquad (T2.125)$$

2.13 Results on the matrix-valued normal distribution

Assume:

$$\mathbf{X}\sim\mathrm{N}(\mathbf{M},\boldsymbol{\Sigma},\mathbf{S}). \qquad (T2.126)$$

From the definition (2.180) of this distribution and the definition of the normal pdf (2.156) we have:

$$f(\mathbf{X})\equiv f(\operatorname{vec}(\mathbf{X}))\equiv(2\pi)^{-\frac{NK}{2}}\left|\mathbf{S}_K\otimes\boldsymbol{\Sigma}_N\right|^{-\frac{1}{2}} e^{-\frac{1}{2}(\operatorname{vec}(\mathbf{X})-\operatorname{vec}(\mathbf{M}))'(\mathbf{S}_K\otimes\boldsymbol{\Sigma}_N)^{-1}(\operatorname{vec}(\mathbf{X})-\operatorname{vec}(\mathbf{M}))} \qquad (T2.127)$$

From the property (A.102) of the Kronecker product we can write

$$\left|\mathbf{S}_K\otimes\boldsymbol{\Sigma}_N\right|^{-\frac{1}{2}}=\left|\mathbf{S}_K\right|^{-\frac{N}{2}}\left|\boldsymbol{\Sigma}_N\right|^{-\frac{K}{2}}, \qquad (T2.128)$$

Furthermore, from the property (A.101) of the Kronecker product, we can write

$$(\mathbf{S}_K\otimes\boldsymbol{\Sigma}_N)^{-1}=\mathbf{S}_K^{-1}\otimes\boldsymbol{\Sigma}_N^{-1}.$$

Therefore (T2.127) can be written as follows:

$$f(\mathbf{X})=(2\pi)^{-\frac{NK}{2}}\left|\mathbf{S}_K\right|^{-\frac{N}{2}}\left|\boldsymbol{\Sigma}_N\right|^{-\frac{K}{2}} e^{-\frac{1}{2}\left\{(\operatorname{vec}(\mathbf{X})-\operatorname{vec}(\mathbf{M}))'\left(\mathbf{S}_K^{-1}\otimes\boldsymbol{\Sigma}_N^{-1}\right)(\operatorname{vec}(\mathbf{X})-\operatorname{vec}(\mathbf{M}))\right\}}. \qquad (T2.129)$$

On the other hand, defining

$$\mathbf{Y}\equiv\mathbf{X}-\mathbf{M},\quad \boldsymbol{\Omega}_N\equiv\boldsymbol{\Sigma}_N^{-1},\quad \boldsymbol{\Phi}_K\equiv\mathbf{S}_K^{-1}, \qquad (T2.130)$$

and recalling the definition (A.96) of the Kronecker product, and the definition (A.104) of the "vec" operator, the term in curly brackets in (T2.129) can be written as follows:

$$\begin{aligned}\{\cdots\} &\equiv \operatorname{vec}(\mathbf{Y})' (\mathbf{\Phi}_K \otimes \mathbf{\Omega}_N) \operatorname{vec}(\mathbf{Y}) \\ &\equiv \left(\mathbf{Y}'_{(1)} \cdots \mathbf{Y}'_{(K)}\right) \begin{pmatrix} \Phi_{11}\mathbf{\Omega} & \cdots & \Phi_{1K}\mathbf{\Omega} \\ \vdots & \ddots & \vdots \\ \Phi_{K1}\mathbf{\Omega} & \cdots & \Phi_{KK}\mathbf{\Omega} \end{pmatrix} \begin{pmatrix} \mathbf{Y}_{(1)} \\ \vdots \\ \mathbf{Y}_{(K)} \end{pmatrix} \\ &= \sum_k \mathbf{Y}'_{(k)} (\Phi_{kj}\mathbf{\Omega}) \mathbf{Y}_{(j)} \\ &= \sum_{n.m,k,j} Y_{nk} \Phi_{kj} \Omega_{nm} Y_{mj} \\ &= \sum_{n.m,k,j} \Omega_{mn} Y_{nk} \Phi_{kj} Y_{mj} \\ &= \operatorname{tr}\left\{\mathbf{\Omega}\mathbf{Y}\mathbf{\Phi}\mathbf{Y}'\right\} = \operatorname{tr}\left\{\mathbf{\Phi}\mathbf{Y}'\mathbf{\Omega}\mathbf{Y}\right\}\end{aligned} \quad (T2.131)$$

Therefore ($T2.129$) can be written as follows:

$$f(\mathbf{X}) \equiv (2\pi)^{-\frac{NK}{2}} |\mathbf{\Sigma}_N|^{-\frac{K}{2}} |\mathbf{S}_K|^{-\frac{N}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{\mathbf{S}_K^{-1}(\mathbf{X}-\mathbf{M})'\mathbf{\Sigma}_N^{-1}(\mathbf{X}-\mathbf{M})\right\}}. \quad (T2.132)$$

To prove the role of the matrices $\mathbf{\Sigma}_N$ and $\mathbf{S}_K$, consider two generic N-dimensional columns $\mathbf{X}_{(j)}$ and $\mathbf{X}_{(k)}$ among the K that compose the random matrix $\mathbf{X}$. The (m,n)-entry of the $N \times N$ covariance matrix between the two columns $\mathbf{X}_{(j)}$ and $\mathbf{X}_{(k)}$ can be written as follows:

$$\begin{aligned}\left[\operatorname{Cov}\left\{\mathbf{X}_{(j)}, \mathbf{X}_{(k)}\right\}\right]_{m,n} &= \operatorname{Cov}\left\{X^{(C)}_{(j-1)N+m}, X^{(C)}_{(k-1)N+n}\right\} \\ &= (\mathbf{S}_K \otimes \mathbf{\Sigma}_N)_{(j-1)N+m,(k-1)N+n} \\ &= \begin{pmatrix} S_{11}\mathbf{\Sigma} & \cdots & S_{1K}\mathbf{\Sigma} \\ \vdots & \ddots & \vdots \\ S_{K1}\mathbf{\Sigma} & \cdots & S_{KK}\mathbf{\Sigma} \end{pmatrix}_{(j-1)N+m,(k-1)N+n} \\ &= S_{j,k}\Sigma_{m,n}.\end{aligned} \quad (T2.133)$$

This proves that if

$$\mathbf{X} \sim \mathrm{N}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{S}) \quad (T2.134)$$

then

$$\operatorname{Cov}\left\{\mathbf{X}_{(j)}, \mathbf{X}_{(k)}\right\} = S_{j,k}\mathbf{\Sigma}. \quad (T2.135)$$

On the other hand, from the following identities

$$\begin{aligned}f(\mathbf{X}) &\equiv (2\pi)^{-\frac{NK}{2}} |\mathbf{\Sigma}_N|^{-\frac{K}{2}} |\mathbf{S}_K|^{-\frac{N}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{\mathbf{S}_K^{-1}(\mathbf{X}-\mathbf{M})'\mathbf{\Sigma}_N^{-1}(\mathbf{X}-\mathbf{M})\right\}} \\ &= (2\pi)^{-\frac{NK}{2}} |\mathbf{S}_K|^{-\frac{N}{2}} |\mathbf{\Sigma}_N|^{-\frac{K}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{\mathbf{\Sigma}_N^{-1}(\mathbf{X}-\mathbf{M})\mathbf{S}_K^{-1}(\mathbf{X}-\mathbf{M})'\right\}},\end{aligned} \quad (T2.136)$$

we see that if

$$\mathbf{X} \sim \mathrm{N}(\mathbf{M}, \mathbf{\Sigma}, \mathbf{S}) \quad (T2.137)$$

then

$$\mathbf{X}' \sim \mathrm{N}\left(\mathbf{M}', \mathbf{S}, \mathbf{\Sigma}\right) \tag{T2.138}$$

Using (*T*2.135) and the fact that the columns of $\mathbf{X}'$ are the rows of $\mathbf{X}$ we thus obtain

$$\mathrm{Cov}\left\{\mathbf{X}^{(m)}, \mathbf{X}^{(n)}\right\} = \Sigma_{mn}\mathbf{S}. \tag{T2.139}$$

2.14 Results on the matrix-valued Student t distribution

The generalization of the Student t distribution to matrix-variate random variables was studied by Dickey (1967). Our definition of the probability density function (2.199) corresponds in the notation of Dickey (1967) to the following special case:

$$p \equiv N, \quad q \equiv K, \quad m \equiv \nu + N, \quad \mathbf{Q} \equiv \nu\mathbf{S}, \quad \mathbf{P} \equiv \mathbf{\Sigma}^{-1}. \tag{T2.140}$$

If

$$\mathbf{X} \sim \mathrm{St}\left(\nu, \mathbf{M}, \mathbf{\Sigma}, \mathbf{S}\right), \tag{T2.141}$$

then

$$\mathrm{E}\left\{\mathbf{X}\right\} = \mathbf{M} \tag{T2.142}$$

$$\mathrm{Cov}\left\{\mathbf{X}_{(j)}, \mathbf{X}_{(k)}\right\} = \frac{\nu}{\nu-2} S_{jk}\mathbf{\Sigma} \tag{T2.143}$$

$$\mathrm{Cov}\left\{\mathbf{X}^{(m)}, \mathbf{X}^{(n)}\right\} = \frac{\nu}{\nu-2} \Sigma_{mn}\mathbf{S}. \tag{T2.144}$$

We show here that the Student t distribution yields the normal distribution when the degrees of freedom tend to infinity. In other words, we prove the following result:

$$\mathrm{St}\left(\infty, \mathbf{M}, \mathbf{\Sigma}, \mathbf{S}\right) = \mathrm{N}\left(\mathbf{M}, \mathbf{\Sigma}, \mathbf{S}\right), \tag{T2.145}$$

where the term on the left hand side is the matrix-variate Student t distribution (2.198) and the term on the right hand side is the matrix-variate normal distribution (2.181). The above result immediately proves the specific vector-variate case. Indeed, from (2.183) and (2.201) we obtain:

$$\begin{aligned} \mathrm{St}\left(\infty, \mathbf{m}, \mathbf{\Sigma}\right) &= \mathrm{St}\left(\infty, \mathbf{m}, \mathbf{\Sigma}, 1\right) = \mathrm{N}\left(\mathbf{M}, \mathbf{\Sigma}, 1\right) \\ &= \mathrm{N}\left(\mathbf{m}, \mathbf{\Sigma}\right). \end{aligned} \tag{T2.146}$$

In turn, since the vector-variate pdf (2.188) case generalizes the one-dimensional pdf (1.86) we also obtain

$$\mathrm{St}\left(\infty, m, \sigma^2\right) = \mathrm{N}\left(m, \sigma^2\right). \tag{T2.147}$$

To prove (*T*2.145) we start using (*A*.122) in the definition (2.199) of the pdf of a matrix-valued Student distribution $\mathrm{St}\left(\nu, \mathbf{M}, \mathbf{\Sigma}, \mathbf{S}\right)$. In the limit $\nu \to \infty$ we obtain:

$$
\begin{aligned}
f(\mathbf{X}) &\equiv \gamma\,|\boldsymbol{\Sigma}|^{-\frac{K}{2}}\,|\mathbf{S}|^{-\frac{N}{2}}\left|\mathbf{I}_K+\mathbf{S}^{-1}(\mathbf{X}-\mathbf{M})'\frac{\boldsymbol{\Sigma}^{-1}}{\nu}(\mathbf{X}-\mathbf{M})\right|^{-\frac{\nu}{2}} \qquad (T2.148)\\
&\approx \gamma\,|\boldsymbol{\Sigma}|^{-\frac{K}{2}}\,|\mathbf{S}|^{-\frac{N}{2}}\left(1+\frac{1}{\nu}\operatorname{tr}\left(\mathbf{S}^{-1}(\mathbf{X}-\mathbf{M})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})\right)\right)^{-\frac{\nu}{2}},
\end{aligned}
$$

where γ is normalization constant (2.200), which we report here:

$$
\gamma(\nu)\equiv(\nu\pi)^{-\frac{NK}{2}}\frac{\Gamma\left(\frac{\nu+N}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{\Gamma\left(\frac{\nu-1+N}{2}\right)}{\Gamma\left(\frac{\nu-1}{2}\right)}\cdots\frac{\Gamma\left(\frac{\nu-K+1+N}{2}\right)}{\Gamma\left(\frac{\nu-K+1}{2}\right)}. \qquad (T2.149)
$$

Using the following limit (see e.g. Rudin (1976)):

$$
e^x=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n, \qquad (T2.150)
$$

we can then write

$$
\begin{aligned}
f^{\mathrm{St}}_{\nu\to\infty,\boldsymbol{\mu},\boldsymbol{\Sigma},\mathbf{S}}(\mathbf{X}) &\approx \gamma\,|\boldsymbol{\Sigma}|^{-\frac{K}{2}}\,|\mathbf{S}|^{-\frac{N}{2}} \qquad (T2.151)\\
&\left[\left(1+\frac{1}{\nu}\operatorname{tr}\left(\mathbf{S}^{-1}(\mathbf{X}-\mathbf{M})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})\right)\right)^{\nu}\right]^{-\frac{1}{2}}\\
&\approx \gamma\,|\boldsymbol{\Sigma}|^{-\frac{K}{2}}\,|\mathbf{S}|^{-\frac{N}{2}}\,e^{-\frac{1}{2}\operatorname{tr}\left(\mathbf{S}^{-1}(\mathbf{X}-\mathbf{M})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})\right)}.
\end{aligned}
$$

Turning now to the normalization constant ($T2.149$), the following approximation holds in the limit $n\to\infty$, see e.g. Graham, Knuth, and Patashnik (1994) or `mathworld.com`:

$$
\Gamma\left(n+\frac{1}{2}\right)\approx\sqrt{n}\,\Gamma(n). \qquad (T2.152)
$$

Applying this result recursively we obtain in the limit $n\to\infty$ the following approximation:

$$
\Gamma\left(\frac{n+N}{2}\right)\approx\left(\frac{n}{2}\right)^{\frac{N}{2}}\Gamma\left(\frac{n}{2}\right). \qquad (T2.153)
$$

Applying this to the normalization constant ($T2.149$) we obtain in the limit $\nu\to\infty$ the following approximation:

$$
\begin{aligned}
\gamma(\nu\to\infty) &\approx (\nu\pi)^{-\frac{NK}{2}}\left(\frac{\nu}{2}\right)^{\frac{N}{2}}\cdots\left(\frac{\nu-K+1}{2}\right)^{\frac{N}{2}}\\
&\approx (\nu\pi)^{-\frac{NK}{2}}\left(\frac{\nu}{2}\right)^{\frac{NK}{2}} \qquad (T2.154)\\
&= (2\pi)^{-\frac{NK}{2}}.
\end{aligned}
$$

Thus in the limit $\nu\to\infty$ the pdf of the matrix-variate Student t distribution $\mathrm{St}(\nu,\mathbf{M},\boldsymbol{\Sigma},\mathbf{S})$ reads:

$$f^{\text{St}}_{\nu\to\infty,\boldsymbol{\mu},\boldsymbol{\Sigma},\mathbf{S}}(\mathbf{X}) \to (2\pi)^{-\frac{NK}{2}} |\boldsymbol{\Sigma}|^{-\frac{K}{2}} |\mathbf{S}|^{-\frac{N}{2}} \left[e^{\operatorname{tr}\left(\mathbf{S}^{-1}(\mathbf{X}-\mathbf{M})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\mathbf{M})\right)} \right]^{-\frac{1}{2}}, \tag{T2.155}$$

which is the pdf (2.182) of the matrix-variate normal distribution $\mathrm{N}(\mathbf{M}, \boldsymbol{\Sigma}, \mathbf{S})$.

2.15 Results on the Cauchy distribution

The logarithm of the Cauchy probability density function (2.209) reads:

$$\ln f^{\text{Ca}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x}) = \gamma - \frac{N+1}{2} \ln\left(1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right) \tag{T2.156}$$

The first order derivative of the log-Cauchy probability density function reads:

$$\frac{\partial \ln f^{\text{Ca}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x})}{\partial \mathbf{x}} = -(N+1) \frac{\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}. \tag{T2.157}$$

Setting this expression to zero we obtain the mode:

$$\operatorname{Mod}\{\mathbf{X}\} = \boldsymbol{\mu}. \tag{T2.158}$$

The Hessian of the log-Cauchy probability density function reads:

$$\begin{aligned}
\frac{\partial^2 \ln f^{\text{Ca}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'} &= -(N+1) \frac{\partial}{\partial \mathbf{x}} \frac{(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}}{1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})} \\
&= -(N+1) \frac{1}{1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})} \frac{\partial}{\partial \mathbf{x}} \left[(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}\right] \\
&\quad -(N+1) \left[\frac{\partial}{\partial \mathbf{x}} \frac{1}{1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}\right] (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} \\
&= -(N+1) \frac{\boldsymbol{\Sigma}^{-1}}{1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})} \\
&\quad -(N+1) - \frac{2\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}}{\left(1 + (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right)^2}.
\end{aligned} \tag{T2.159}$$

Evaluating this expression in the mode (T2.158) we obtain:

$$\left.\frac{\partial^2 f^{\text{Ca}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'}\right|_{\mathbf{x}=\operatorname{Mod}\{\mathbf{X}\}} = -(N+1)\boldsymbol{\Sigma}^{-1}. \tag{T2.160}$$

Therefore the modal dispersion (2.65) reads:

$$\operatorname{MDis}\{\mathbf{X}\} \equiv -\left(\left.\frac{\partial^2 f^{\text{Ca}}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'}\right|_{\mathbf{x}=\operatorname{Mod}\{\mathbf{X}\}}\right)^{-1} = \frac{1}{N+1}\boldsymbol{\Sigma} \tag{T2.161}$$

2.16 Results on log-distributions

Assume the distribution of the random variable $\mathbf{X}$ is known, and it is represented by its pdf $f_{\mathbf{X}}$ or its characteristic function $\phi_{\mathbf{X}}$. Consider a new random variable $\mathbf{Y}$ defined as follows:

$$\mathbf{Y} \equiv e^{\mathbf{X}}, \tag{T2.162}$$

where the exponential is defined component-wise. This is a transformation $\mathbf{g}$ of the form (T2.8), which reads component-wise as follows:

$$g_n(x_1, \ldots, x_N) \equiv e^{x_n}. \tag{T2.163}$$

The inverse transformation $\mathbf{g}^{-1}$ reads component-wise:

$$g_n^{-1}(y_1, \ldots, y_N) \equiv \ln(y_n). \tag{T2.164}$$

The Jacobian (T2.11) reads:

$$\mathbf{J}^{\mathbf{g}} = \operatorname{diag}(e^{x_1}, \ldots, e^{x_N}), \tag{T2.165}$$

and thus from (A.42) its determinant reads:

$$|\mathbf{J}^{\mathbf{g}}| = \prod_{n=1}^{N} e^{x_n}. \tag{T2.166}$$

We have to evaluate (T2.166) in $\mathbf{x} = \mathbf{g}^{-1}(\mathbf{y})$. Therefore from ($T$2.164) we obtain

$$\left|\mathbf{J}^{g}\left(\mathbf{g}^{-1}(\mathbf{y})\right)\right| = \prod_{n=1}^{N} y_n \tag{T2.167}$$

Therefore from (T2.13) the pdf of the lognormal distribution reads:

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{f_{\mathbf{X}}(\ln(\mathbf{y}))}{\prod_{n=1}^{N} y_n}. \tag{T2.168}$$

Now we compute the raw moments (T2.91) of the log-variable $\mathbf{Y}$:

$$\begin{aligned} \mathrm{RM}^{\mathbf{Y}}_{n_1 \cdots n_k} &\equiv \mathrm{E}\{Y_{n_1} \cdots Y_{n_k}\} \\ &= \mathrm{E}\left\{e^{X_{n_1}} \cdots e^{X_{n_k}}\right\} = \mathrm{E}\left\{e^{X_{n_1} + \cdots + X_{n_k}}\right\} \\ &= \mathrm{E}\left\{\exp\left(i\boldsymbol{\omega}'_{n_1 \cdots n_k}\mathbf{X}\right)\right\}, \end{aligned} \tag{T2.169}$$

where the vector $\boldsymbol{\omega}$ is defined in terms of the canonical basis (A.15) as follows:

$$\boldsymbol{\omega}_{n_1 \cdots n_k} \equiv \frac{1}{i}\left(\boldsymbol{\delta}^{(n_1)} + \cdots + \boldsymbol{\delta}^{(n_k)}\right). \tag{T2.170}$$

Comparing with (2.13), we realize that the last term in (T2.169) is the characteristic function of $\mathbf{X}$. Therefore we obtain:

$$\mathrm{RM}^{\mathbf{Y}}_{n_1 \cdots n_k} = \phi_{\mathbf{X}} \left(\boldsymbol{\omega}_{n_1 \cdots n_k} \right) .. \quad (T2.171)$$

The central moments can be obtained from the raw moments as discussed in Appendix www.2.10

In particular, if $\mathbf{X}$ is normally distributed:

$$\mathbf{X} \sim \mathrm{N} \left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \right), \quad (T2.172)$$

then by definition $\mathbf{Y}$ is lognormally distributed:

$$\mathbf{Y} \sim \mathrm{LogN} \left(\boldsymbol{\mu}, \boldsymbol{\Sigma} \right). \quad (T2.173)$$

Therefore from (2.156) and (T2.168) we immediately obtain the pdf of the lognormal distribution:

$$f^{\mathrm{LogN}}_{\boldsymbol{\mu}, \Sigma} \left(\mathbf{y} \right) = \frac{(2\pi)^{-\frac{N}{2}} \left| \boldsymbol{\Sigma} \right|^{-\frac{1}{2}}}{\prod_{n=1}^{N} y_n} e^{-\frac{1}{2} (\ln(\mathbf{y}) - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\ln(\mathbf{y}) - \boldsymbol{\mu})}. \quad (T2.174)$$

Furthermore, from (2.157) and (T2.171) we obtain the expression of the raw moments of the lognormal distribution:

$$\begin{aligned} \mathrm{RM}^{\mathbf{Y}}_{n_1 \cdots n_k} = e^{\boldsymbol{\mu}' \left(\boldsymbol{\delta}^{(n_1)} + \cdots + \boldsymbol{\delta}^{(n_k)} \right)} \\ e^{\frac{1}{2} \left(\boldsymbol{\delta}^{(n_1)} + \cdots + \boldsymbol{\delta}^{(n_k)} \right)' \boldsymbol{\Sigma} \left(\boldsymbol{\delta}^{(n_1)} + \cdots + \boldsymbol{\delta}^{(n_k)} \right)}. \end{aligned} \quad (T2.175)$$

In particular the expected value, which is the first raw moment, reads

$$\mathrm{E} \left\{ Y_n \right\} = \mathrm{RM}^{\mathbf{Y}}_{n} = e^{\mu_n + \frac{\Sigma_{nn}}{2}}. \quad (T2.176)$$

The second raw moment reads:

$$\mathrm{E} \left\{ Y_m Y_n \right\} = \mathrm{RM}^{\mathbf{Y}}_{mn} = e^{\mu_m + \mu_n + \frac{\Sigma_{mm}}{2} + \frac{\Sigma_{nn}}{2} + \Sigma_{mn}}. \quad (T2.177)$$

Therefore the covariance matrix reads:

$$\begin{aligned} \mathrm{Cov} \left\{ X_m, X_n \right\} &= \mathrm{E} \left\{ Y_m Y_n \right\} - \mathrm{E} \left\{ Y_m \right\} \mathrm{E} \left\{ Y_n \right\} \\ &= e^{\mu_m + \mu_n + \frac{\Sigma_{mm}}{2} + \frac{\Sigma_{nn}}{2}} \left(e^{\Sigma_{mn}} - 1 \right). \end{aligned} \quad (T2.178)$$

2.17 Results on the Wishart distribution

Relation between Wishart and gamma distribution

If $\mathbf{W}$ is Wishart distributed as in (2.223), then from (2.222) for any conformable matrix $\mathbf{A}$ we have:

$$\begin{aligned}\mathbf{AWA}' &= \mathbf{AX}_1\mathbf{X}_1'\mathbf{A}' + \cdots + \mathbf{AX}_\nu\mathbf{X}_\nu'\mathbf{A}' \\ &= \mathbf{Y}_1'\mathbf{Y}_1' + \cdots + \mathbf{Y}_\nu\mathbf{Y}_\nu' \\ &\sim \mathrm{W}\left(\nu; \mathbf{A\Sigma A}'\right).\end{aligned} \qquad (T2.179)$$

since

$$\mathbf{Y}_t \equiv \mathbf{AX}_t \sim \mathrm{N}\left(\mathbf{0}; \mathbf{A\Sigma A}'\right). \qquad (T2.180)$$

In particular, we can reconcile the multivariate Wishart with the one-dimensional gamma distribution by choosing $\mathbf{A} \equiv \mathbf{a}'$, a row vector. In that case each term in the sum is normally distributed as follows:

$$Y_t \equiv \boldsymbol{\alpha}'\mathbf{X}_t \sim \mathrm{N}\left(0; \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\right). \qquad (T2.181)$$

Therefore from (1.106).

$$\mathbf{a}'\mathbf{Wa} \sim \mathrm{Ga}\left(\nu, \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\right). \qquad (T2.182)$$

The pdf of the inverse-Wishart distribution

Assume as in (2.232) that the matrix $\mathbf{Z}$ has an inverse-Wishart distribution:

$$\mathbf{Z} \sim \mathrm{IW}\left(\nu, \boldsymbol{\Psi}\right). \qquad (T2.183)$$

By definition

$$\mathbf{Z} \equiv \mathbf{g}\left(\mathbf{W}\right) \equiv \mathbf{W}^{-1}, \qquad (T2.184)$$

where

$$\mathbf{W} \sim \mathrm{W}\left(\nu, \boldsymbol{\Psi}^{-1}\right). \qquad (T2.185)$$

Then from ($T2.14$) in Appendix www.2.2:

$$f^{\mathrm{IW}}_{\nu,\boldsymbol{\Psi}}\left(\mathbf{Z}\right) = \frac{f^{\mathrm{W}}_{\nu,\boldsymbol{\Psi}^{-1}}\left(\mathbf{g}^{-1}\left(\mathbf{Z}\right)\right)}{\sqrt{\left|\mathbf{J}^{g}\left(\mathbf{g}^{-1}\left(\mathbf{Z}\right)\right)\right|^2}}. \qquad (T2.186)$$

Using the following result in Magnus and Neudecker (1999) that applies to any invertible $N \times N$ matrix $\mathbf{Q}$:

$$\left|\frac{\partial \mathbf{Q}^{-1}}{\partial \mathbf{Q}}\right| = (-1)^{\frac{N(N+1)}{2}} \left|\mathbf{Q}\right|^{-(N+1)}, \qquad (T2.187)$$

we derive:

$$\begin{aligned} f^{\mathrm{IW}}_{\nu,\boldsymbol{\Psi}}\left(\mathbf{Z}\right) &= \left|\mathbf{Z}\right|^{-(N+1)} f^{\mathrm{W}}_{\nu,\boldsymbol{\Psi}^{-1}}\left(\mathbf{Z}^{-1}\right) \\ &= \left|\mathbf{Z}\right|^{-(N+1)} \frac{1}{\kappa} \left|\boldsymbol{\Psi}^{-1}\right|^{-\frac{\nu}{2}} \left|\mathbf{Z}^{-1}\right|^{\frac{\nu-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Z}^{-1}\right)} \\ &= \frac{1}{\kappa}\left|\boldsymbol{\Psi}\right|^{\frac{\nu}{2}} \left|\mathbf{Z}\right|^{-\frac{\nu+N+1}{2}} e^{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}\mathbf{Z}^{-1}\right)}. \end{aligned} \qquad (T2.188)$$

2.18 Results on elliptical distributions

The family of ellipsoids centered in $\boldsymbol{\mu}$ with shape $\boldsymbol{\Sigma}$ are described by the following implicit equations:

$$\mathrm{Ma}\left(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = u, \tag{$T2.189$}$$

where Ma is the Mahalanobis distance of the point $\mathbf{x}$ from $\boldsymbol{\mu}$ through the metric $\boldsymbol{\Sigma}$, as defined in (2.61).and $u \in (0, \infty)$, see ($A.73$).

If the pdf $f_{\mathbf{X}}$ is constant on those ellipsoids then it must be of the form:

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(\mathbf{x}\right) = h\left[\mathrm{Ma}^2\left(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)\right], \tag{$T2.190$}$$

where h is a positive function, such that the normalization condition (2.6) is satisfied, i.e.

$$\int_{\mathbb{R}^N} h\left[\mathrm{Ma}^2\left(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)\right] d\mathbf{x} = 1. \tag{$T2.191$}$$

Suppose we have determined such a function h. From ($T2.30$), changing $\boldsymbol{\mu}$ into a generic parameter $\widetilde{\boldsymbol{\mu}}$ does not affect the normalization condition, and therefore the ensuing pdf is still the pdf of an elliptical distribution centered in $\widetilde{\boldsymbol{\mu}}$. On the other hand, if we change $\boldsymbol{\Sigma}$ into a generic dispersion parameter $\widetilde{\boldsymbol{\Sigma}}$, in order to preserve the normalization condition we have to rescale ($T2.190$) accordingly:

$$f_{\widetilde{\boldsymbol{\mu}},\widetilde{\boldsymbol{\Sigma}}}\left(\mathbf{x}\right) = \sqrt{\frac{\left|\widetilde{\boldsymbol{\Sigma}}\right|}{\left|\boldsymbol{\Sigma}\right|}} h\left[\mathrm{Ma}^2\left(\mathbf{x}, \widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\Sigma}}\right)\right]. \tag{$T2.192$}$$

Therefore, it is more convenient to replace ($T2.190$) with the following specification:

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(\mathbf{x}\right) = \frac{1}{\sqrt{\left|\boldsymbol{\Sigma}\right|}} g\left[\mathrm{Ma}^2\left(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)\right], \tag{$T2.193$}$$

in such a way that the same functional form g is viable for any location and dispersion parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

To summarize, the pdf of an elliptical distribution is of the form ($T2.193$), where g is any positive function that satisfies:

$$\int_0^\infty y^{\frac{N-2}{2}} g\left(y\right) dy < \infty, \tag{$T2.194$}$$

see Fang, Kotz, and Ng (1990), p. 35.

Moments of elliptical distributions

First we follow Fang, Kotz, and Ng (1990) to compute the moments of a random variable $\mathbf{U}$ uniformly distributed on the surface of the unit ball.

Consider a standard multivariate normal variable

$$\mathbf{X} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_N). \qquad (T2.195)$$

We can write:

$$\mathbf{X} = \|\mathbf{X}\| \mathbf{U}. \qquad (T2.196)$$

From (2.259) we have that $\|\mathbf{X}\|$ and $\mathbf{U} \equiv \mathbf{X}/\|\mathbf{X}\|$ are independent and $\mathbf{U}$ uniformly distributed on the surface of the unit ball. Then:

$$\begin{aligned} \mathrm{E}\left\{\prod_{i=1}^N X_i^{2s_i}\right\} &= \mathrm{E}\left\{\prod_{i=1}^N (\|\mathbf{X}\| U_i)^{2s_i}\right\} \qquad (T2.197) \\ &= \mathrm{E}\left\{\left(\prod_{i=1}^N \|\mathbf{X}\|^{2s_i}\right)\left(\prod_{i=1}^N U_i^{2s_i}\right)\right\} \\ &= \mathrm{E}\left\{\|\mathbf{X}\|^{2s}\right\} \mathrm{E}\left\{\prod_{i=1}^N U_i^{2s_i}\right\}, \end{aligned}$$

where

$$s \equiv \sum_{i=1}^N s_i. \qquad (T2.198)$$

Thus

$$\mathrm{E}\left\{\prod_{i=1}^N U_i^{2s_i}\right\} = \frac{\prod_{i=1}^N \mathrm{E}\left\{X_i^{2s_i}\right\}}{\mathrm{E}\left\{\|\mathbf{X}\|^{2s}\right\}}. \qquad (T2.199)$$

On the other hand, for a standard normal variable X_i we have:

$$\mathrm{E}\left\{X_i^{2s_i}\right\} = \frac{(2s_i)!}{2^{s_i} s_i!}, \qquad (T2.200)$$

see e.g. `mathworld.com` and references therein. For a standard multivariate normal variable $\mathbf{X}$ we have

$$\begin{aligned} \mathrm{E}\left\{\|\mathbf{X}\|^{2s}\right\} &= \mathrm{E}\left\{\left(X_1^2 + \cdots + X_N^2\right)^s\right\} \qquad (T2.201) \\ &= \mathrm{E}\left\{Y^s\right\}. \end{aligned}$$

Therefore from (1.109) we see that ($T2.201$) is the s-th raw moment of a chi-square distribution with N degrees of freedom and thus, see e.g. `mathworld.com` and references therein, we have:

$$\mathrm{E}\left\{\|\mathbf{X}\|^{2s}\right\} = \frac{\Gamma\left(\frac{N}{2} + s\right) 2^s}{\Gamma\left(\frac{N}{2}\right)}. \qquad (T2.202)$$

From ($B.82$) we have:

$$\Gamma\left(\frac{N}{2}+s\right) = \Gamma\left(\frac{N+2s}{2}\right) = \frac{(N+2s-2)\,(N+2\,(s-1))\cdots n_0\sqrt{\pi}}{2^{\frac{N+2s-1}{2}}}. \quad (T2.203)$$

Defining:

$$x^{[s]} \equiv x\,(x+1)\cdots(x+s-1) \quad (T2.204)$$

we can write (T2.202) as follows

$$\begin{aligned} \mathrm{E}\left\{\|\mathbf{X}\|^{2s}\right\} &= (N+2\,(s-1))\cdots(N+2)\,N \\ &= 2^s\left(\frac{N}{2}+(s-1)\right)\cdots\left(\frac{N}{2}+1\right)\frac{N}{2} \\ &= 2^s\left(\frac{N}{2}\right)^{[s]}. \end{aligned} \quad (T2.205)$$

Substituting (T2.200) and (T2.205) in (T2.199) we obtain:

$$\mathrm{E}\left\{\prod_{i=1}^{N} U_i^{2s_i}\right\} = \frac{1}{\left(\frac{N}{2}\right)^{[s]}}\prod_{i=1}^{N}\frac{(2s_i)!}{4^{s_i}s_i!}, \quad (T2.206)$$

which is Formula (3.6) in Fang, Kotz, and Ng (1990).

In particular

$$\mathrm{E}\{\mathbf{U}\} = \mathbf{0} \quad (T2.207)$$

and

$$\mathrm{Cov}\{\mathbf{U}\} = \frac{\mathbf{I}_N}{N}, \quad (T2.208)$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix.

Consider now a generic elliptical random variable $\mathbf{X}$ with location parameter $\boldsymbol{\mu}$ and scatter parameter $\boldsymbol{\Sigma}$. To compute its central moments we write:

$$\mathbf{X} \equiv \boldsymbol{\mu} + R\mathbf{A}\mathbf{U}, \quad (T2.209)$$

where

$$\begin{aligned} \mathbf{A}\mathbf{A}' &\equiv \boldsymbol{\Sigma} \\ \mathbf{U} &\equiv \frac{\mathbf{A}^{-1}(\mathbf{X}-\boldsymbol{\mu})}{\|\mathbf{A}^{-1}(\mathbf{X}-\boldsymbol{\mu})\|} \\ R &\equiv \left\|\mathbf{A}^{-1}(\mathbf{X}-\boldsymbol{\mu})\right\|. \end{aligned} \quad (T2.210)$$

From (2.259), R is independent of $\mathbf{U}$, which is uniformly distributed on the surface of the unit ball. Using (T2.207) we obtain:

$$\begin{aligned} \mathrm{E}\{\mathbf{X}\} &= \mathrm{E}\{\boldsymbol{\mu}+R\mathbf{A}\mathbf{U}\} = \boldsymbol{\mu} + \mathrm{E}\{R\}\,\mathbf{A}\,\mathrm{E}\{\mathbf{U}\} \\ &= \boldsymbol{\mu}. \end{aligned} \quad (T2.211)$$

Using (2.93) we obtain:

$$\begin{aligned} \mathrm{CM}^{\mathbf{X}}_{m_1\cdots m_k} &= \mathrm{CM}^{\boldsymbol{\mu}+R\mathbf{AU}}_{m_1\cdots m_k} = \mathrm{CM}^{\mathbf{A}R\mathbf{U}}_{m_1\cdots m_k} \\ &= \sum_{n_1,\ldots,n_k=1}^{N} A_{m_1n_1}\cdots A_{m_kn_k}\,\mathrm{CM}^{R\mathbf{U}}_{m_1\cdots m_k} \\ &= \sum_{n_1,\ldots,n_k=1}^{N} A_{m_1n_1}\cdots A_{m_kn_k}\,\mathrm{RM}^{R\mathbf{U}}_{m_1\cdots m_k} \\ &= \sum_{n_1,\ldots,n_k=1}^{N} A_{m_1n_1}\cdots A_{m_kn_k}\,\mathrm{E}\left\{RU_{n_1}\cdots RU_{n_k}\right\} \\ &= \mathrm{E}\left\{R^k\right\}\sum_{n_1,\ldots,n_k=1}^{N} A_{m_1n_1}\cdots A_{m_kn_k}\,\mathrm{E}\left\{U_{n_1}\cdots U_{n_k}\right\}. \end{aligned} \quad (T2.212)$$

Substituting (T2.206) in (T2.212) yields the desired result.

2.19 Results on stable distributions

Consider a normally distributed random variable:

$$\mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \quad (T2.213)$$

Consider the spectral decomposition (A.70) of the covariance matrix:

$$\boldsymbol{\Sigma} \equiv \mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{E}', \quad (T2.214)$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix of the eigenvalues of $\mathbf{S}$:

$$\boldsymbol{\Lambda} \equiv \mathrm{diag}(\lambda_1, \ldots, \lambda_N); \quad (T2.215)$$

and $\mathbf{E}$ is the juxtaposition of the eigenvectors of $\mathbf{S}$:

$$\mathbf{E} \equiv \left(\mathbf{e}^{(1)}|\cdots|\mathbf{e}^{(N)}\right), \quad (T2.216)$$

Define the N vectors $\left\{\mathbf{v}^{(1)}, \ldots \mathbf{v}^{(N)}\right\}$ as follows:

$$\left(\mathbf{v}^{(1)}|\ldots|\mathbf{v}^{(N)}\right) \equiv \mathbf{V} \equiv \mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}}. \quad (T2.217)$$

These vectors and their opposite belong to the surface of the ellipsoid $\mathcal{E}_{\mathbf{0},\boldsymbol{\Sigma}}$ with shape parameter $\boldsymbol{\Sigma}$ centered in zero defined in (A.73). Indeed

$$\begin{aligned} \left[\pm\mathbf{v}^{(m)}\right]'\boldsymbol{\Sigma}^{-1}\left[\pm\mathbf{v}^{(n)}\right] &= \left[\mathbf{v}^{(m)}\right]'\boldsymbol{\Sigma}^{-1}\left[\mathbf{v}^{(n)}\right] \\ &= \left[\mathbf{V}'\boldsymbol{\Sigma}^{-1}\mathbf{V}\right]_{mn} \\ &= \left[\left(\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{E}'\right)\mathbf{E}\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'\left(\mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}}\right)\right]_{mn} \\ &= [\mathbf{I}]_{mn}. \end{aligned} \quad (T2.218)$$

Thus, in particular:

$$\left[\pm\mathbf{v}^{(n)}\right]' \boldsymbol{\Sigma}^{-1} \left[\pm\mathbf{v}^{(n)}\right] = 1. \tag{$T2.219$}$$

Consider the following measure:

$$m \equiv \frac{1}{4} \sum_{n=1}^{N} \left(\delta^{(\mathbf{v}_n)} + \delta^{(-\mathbf{v}_n)} \right), \tag{$T2.220$}$$

where $\delta^{(\mathbf{x})}$ is the Dirac delta centered in $\mathbf{x}$ as defined in (B.16). Due to (T2.219) this measure satisfies (2.284) and thus it is defined on the surface of the ellipsoid. Also, it trivially satisfies (2.283) and thus it is symmetrical. Furthermore,

$$\begin{aligned} \int_{\mathbb{R}^N} \mathbf{s}\mathbf{s}' m(\mathbf{s})\, d\mathbf{s} &\equiv \frac{1}{4} \int_{\mathbb{R}^N} \mathbf{s}\mathbf{s}' \sum_{n=1}^{N} \left(\delta^{(\mathbf{v}_n)} + \delta^{(-\mathbf{v}_n)} \right)(\mathbf{s})\, d\mathbf{s} \\ &= \frac{1}{2} \sum_{n=1}^{N} \mathbf{v}_n \mathbf{v}_n' = \frac{1}{2} \mathbf{V}\mathbf{V}' = \frac{1}{2} \mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{\Lambda}^{\frac{1}{2}} \mathbf{E}' \\ &= \frac{1}{2} \boldsymbol{\Sigma} \end{aligned} \tag{$T2.221$}$$

Therefore

$$\begin{aligned} \int_{\mathbb{R}^N} \left|\boldsymbol{\omega}'\mathbf{s}\right|^2 m(\mathbf{s})\, d\mathbf{s} &= \int_{\mathbb{R}^N} (\boldsymbol{\omega}'\mathbf{s})(\mathbf{s}'\boldsymbol{\omega})\, m(\mathbf{s})\, d\mathbf{s} \\ &= \boldsymbol{\omega}' \left(\int_{\mathbb{R}^N} \mathbf{s}\mathbf{s}' m(\mathbf{s})\, d\mathbf{s} \right) \boldsymbol{\omega} \\ &= \frac{1}{2} \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega}. \end{aligned} \tag{$T2.222$}$$

This shows that the characteristic function (2.157) of the normal distribution can be written as

$$\begin{aligned} \phi^{\mathrm{N}}_{\boldsymbol{\mu},\Sigma}(\boldsymbol{\omega}) &= e^{i\boldsymbol{\mu}'\boldsymbol{\omega} - \frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega}} \\ &= e^{i\boldsymbol{\mu}'\boldsymbol{\omega}} \exp\left(- \int_{\mathbb{R}^N} \left|\boldsymbol{\omega}'\mathbf{s}\right|^2 m(\mathbf{s})\, d\mathbf{s} \right) \end{aligned} \tag{$T2.223$}$$

3

Technical appendix to Chapter 3

3.1 Properties of the ATMF implied volatility

Substituting the definition (3.48) of the ATMF strike in the Black-Scholes pricing formula (3.41) we obtain:

$$\begin{aligned} C_t^{(K_t,E)} &= C^{BS}\left(E-t, K_t, U_t, Z_t^{(E)}, \sigma_t^{(K,E)}\right) \\ &= \frac{1}{2}U_t\left(1+\operatorname{erf}\left(\frac{\sqrt{E-t}}{2}\sigma_t^{(K,E)}\right)\right) \\ &\quad -\frac{1}{2}U_t\left(1+\operatorname{erf}\left(-\frac{\sqrt{E-t}}{2}\sigma_t^{(K,E)}\right)\right) \\ &= U_t \operatorname{erf}\left(\frac{\sqrt{E-t}}{\sqrt{8}}\sigma_t^{(K,E)}\right) \end{aligned} \tag{T3.1}$$

Therefore

$$\sigma_t^{(K_t,E)} = \sqrt{\frac{8}{E-t}}\operatorname{erf}^{-1}\left(\frac{C_t^{(K_t,E)}}{U_t}\right) \tag{T3.2}$$

Using the following Maple command:

```
> taylor(RootOf(erf(y)=x,y),x=0, 3);
```

we can perform a third-order Taylor expansion of the inverse error function:

$$\operatorname{erf}^{-1}(x) = \frac{\pi^{1/2}}{2}x + \frac{\pi^{3/2}}{24}x^3 + \cdots \tag{T3.3}$$

Since the term in the argument of the inverse error function in (T3.2) is of the order of a few percentage points, we can stop at the first order, obtaining (3.51):

$$\sigma_t^{(K_t,E)} \approx \sqrt{\frac{2\pi}{E-t}}\frac{C_t^{(K_t,E)}}{U_t}. \tag{T3.4}$$

Nevertheless, we could easily proceed to higher orders.

3.2 Distribution of the sum of independent variables

Consider two variables $\mathbf{X}_A$ and $\mathbf{X}_B$ whose joint probability density function is $f_{\mathbf{X}_A,\mathbf{X}_B}$. Consider now the sum of these variables. The pdf of the sum is

$$f_{\mathbf{X}_A+\mathbf{X}_B}(\mathbf{y}) = \int_{\mathbb{R}} f_{\mathbf{X}_A,\mathbf{X}_B}(\mathbf{y}-\mathbf{x},\mathbf{x})\,d\mathbf{x}. \tag{T3.5}$$

This result follows from:

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y})\,d\mathbf{y} &= \mathbb{P}\{\mathbf{Y}\in[\mathbf{y},\mathbf{y}+d\mathbf{y}]\} = \mathbb{P}\{\mathbf{X}_A+\mathbf{X}_B\in[\mathbf{y},\mathbf{y}+d\mathbf{y}]\} \\ &= \mathbb{P}\{\mathbf{X}_A\in[\mathbf{y}-\mathbf{X}_B,\mathbf{y}-\mathbf{X}_B+d\mathbf{y}]\} \\ &= \left(\int_{\mathbb{R}} f_{\mathbf{X}_A,\mathbf{X}_B}(\mathbf{y}-\mathbf{x}_B,\mathbf{x}_B)\,d\mathbf{x}_B\right)d\mathbf{y}. \end{aligned} \tag{T3.6}$$

If the variables $\mathbf{X}$ and $\mathbf{Y}$ are independent, the joint pdf is the product of the marginal pdf:

$$f_{\mathbf{X}_A,\mathbf{X}_B}(\mathbf{x}_A,\mathbf{x}_B) = f_{\mathbf{X}_A}(\mathbf{x}_A)\,f_{\mathbf{X}_B}(\mathbf{x}_B) \tag{T3.7}$$

Therefore the pdf of the sum ($T3.5$) becomes

$$f_{\mathbf{X}_A+\mathbf{X}_B}(\mathbf{y}) = \int_{\mathbb{R}} f_{\mathbf{X}_A}(\mathbf{y}-\mathbf{x})\,f_{\mathbf{X}_B}(\mathbf{x})\,d\mathbf{x}. \tag{T3.8}$$

We see that this is the convolution ($B.43$) of the marginal pdf:

$$f_{\mathbf{X}_A+\mathbf{X}_B} = f_{\mathbf{X}_A} * f_{\mathbf{X}_B} \tag{T3.9}$$

This is the representation of the distribution of the sum of two independent variables in terms of the probability density function.

Repeating the above argument we obtain that the sum of any number of *independent and identically distributed* random variables reads:

$$f_{\mathbf{X}_1+\cdots+\mathbf{X}_T} = f_{\mathbf{X}} * \cdots * f_{\mathbf{X}}, \tag{T3.10}$$

where $f_{\mathbf{X}}$ is the common pdf of each generic variable $\mathbf{X}_t$.

The representation in terms of the characteristic function is much easier. Indeed using the factorization (2.48) of the characteristic function of independent variables we obtain:

$$\begin{aligned} \phi_{\mathbf{X}_1+\cdots+\mathbf{X}_T}(\boldsymbol{\omega}) &\equiv \mathrm{E}\left\{e^{i\boldsymbol{\omega}'(\mathbf{X}_1+\cdots+\mathbf{X}_T)}\right\} \\ &= \mathrm{E}\left\{e^{i\boldsymbol{\omega}'\mathbf{X}_1}\cdots e^{i\boldsymbol{\omega}'\mathbf{X}_T}\right\} \\ &= \mathrm{E}\left\{e^{i\boldsymbol{\omega}'\mathbf{X}_1}\right\}\cdots\mathrm{E}\left\{e^{i\boldsymbol{\omega}'\mathbf{X}_T}\right\}. \end{aligned} \tag{T3.11}$$

Therefore

$$\phi_{\mathbf{X}_1+\cdots+\mathbf{X}_T}(\boldsymbol{\omega}) = \left(\phi_{\mathbf{X}}(\boldsymbol{\omega})\right)^T, \tag{T3.12}$$

where $\phi_{\mathbf{X}}$ is the common characteristic function of each generic variable $\mathbf{X}_t$.

This result is not surprising. Indeed, we recall that from (2.14) that the characteristic function of a distribution is the Fourier transform see (B.34) of the probability density function of that distribution. Therefore:

$$\phi_{\mathbf{X}_1+\cdots+\mathbf{X}_T} = \mathcal{F}\left[f_{\mathbf{X}_1+\cdots+\mathbf{X}_T}\right]. \tag{$T3.13$}$$

Therefore, using the expression of the pdf of the sum (T3.10) and the relation between convolution and the Fourier transform (B.45) we obtain:

$$\begin{aligned} \phi_{\mathbf{X}_1+\cdots+\mathbf{X}_T} &= \mathcal{F}\left[f_{\mathbf{X}} * \cdots * f_{\mathbf{X}}\right] = \left(\mathcal{F}\left[f_{\mathbf{X}}\right]\right)^T \\ &= \left(\phi_{\mathbf{X}}\left(\boldsymbol{\omega}\right)\right)^T, \end{aligned} \tag{$T3.14$}$$

which is again (T3.12).

Formula (T3.12) also provides a faster way to compute the probability density function. Indeed we only need to apply once the inverse Fourier transform $\mathcal{F}^{-1}$ as defined in (B.40), if the distribution of $\mathbf{X}$ is known through its characteristic function:

$$f_{\mathbf{X}_1+\cdots+\mathbf{X}_T} = \mathcal{F}^{-1}\left[\left(\phi_{\mathbf{X}}\left(\boldsymbol{\omega}\right)\right)^T\right]. \tag{$T3.15$}$$

In case the distribution of $\mathbf{X}$ is known through its pdf we only need to apply once the inverse Fourier transform $\mathcal{F}^{-1}$ and once the Fourier transform $\mathcal{F}$:

$$f_{\mathbf{X}_1+\cdots+\mathbf{X}_T} = \mathcal{F}^{-1}\left[\left(\mathcal{F}\left[f_{\mathbf{X}}\right]\right)^T\right]. \tag{$T3.16$}$$

3.3 The "square-root rule"

We recall from (3.64) the relation between the investment-horizon characteristic function and the estimation interval characteristic function

$$\phi_{\mathbf{X}_{T,\tau}} = \left(\phi_{\mathbf{X}_{T,\widetilde{\tau}}}\right)^{\frac{\tau}{\widetilde{\tau}}}. \tag{$T3.17$}$$

The first derivative of the characteristic function reads:

$$\begin{aligned} \frac{\partial \phi_{\mathbf{X}_{T,\tau}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}} &= \frac{\partial \left(\phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)\right)^{\frac{\tau}{\widetilde{\tau}}}}{\partial \boldsymbol{\omega}} \\ &= \frac{\tau}{\widetilde{\tau}}\left(\phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)\right)^{\frac{\tau}{\widetilde{\tau}}-1} \frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}}. \end{aligned} \tag{$T3.18$}$$

The second derivative of the characteristic function reads:

$$\begin{aligned} \frac{\partial^2 \phi_{\mathbf{X}_{T,\tau}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'} &= \frac{\partial}{\partial \boldsymbol{\omega}}\left[\frac{\tau}{\widetilde{\tau}}\left(\phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)\right)^{\frac{\tau}{\widetilde{\tau}}-1} \frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}'}\right] \\ &= \frac{\tau}{\widetilde{\tau}}\left(\frac{\tau}{\widetilde{\tau}}-1\right)\left(\phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)\right)^{\frac{\tau}{\widetilde{\tau}}-2} \frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}} \frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}'} \\ &\quad + \frac{\tau}{\widetilde{\tau}}\left(\phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)\right)^{\frac{\tau}{\widetilde{\tau}}-1} \frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}\left(\boldsymbol{\omega}\right)}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'}. \end{aligned} \tag{$T3.19$}$$

From ($T2.88$), evaluating these derivatives in the origin and using

$$\phi_{\mathbf{X}}(\mathbf{0}) \equiv \mathrm{E}\left\{e^{i\mathbf{X}'\mathbf{0}}\right\} = 1, \tag{T3.20}$$

we obtain for the first raw moment:

$$\mathrm{E}\{\mathbf{X}_{T,\tau}\} = -i\frac{\partial \phi_{\mathbf{X}_{T,\tau}}(\mathbf{0})}{\partial \boldsymbol{\omega}} = \frac{\tau}{\widetilde{\tau}}\left(-i\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}}\right); \tag{T3.21}$$

and for the second raw moment:

$$\begin{aligned}
\mathrm{E}\left\{\mathbf{X}_{T,\tau}\mathbf{X}'_{T,\tau}\right\} &= -\frac{\partial^2 \phi_{\mathbf{X}_{T,\tau}}(\mathbf{0})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'} \qquad (T3.22)\\
&= -\frac{\tau}{\widetilde{\tau}}\left(\frac{\tau}{\widetilde{\tau}}-1\right)\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}}\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}'} - \frac{\tau}{\widetilde{\tau}}\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'}.
\end{aligned}$$

Therefore for the covariance we obtain:

$$\begin{aligned}
\mathrm{Cov}\{\mathbf{X}_{T,\tau}\} &= \mathrm{E}\left\{\mathbf{X}_{T,\tau}\mathbf{X}'_{T,\tau}\right\} - \mathrm{E}\{\mathbf{X}_{T,\tau}\}\,\mathrm{E}\{\mathbf{X}_{T,\tau}\}'\\
&= -\frac{\tau}{\widetilde{\tau}}\left(\frac{\tau}{\widetilde{\tau}}-1\right)\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}}\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}'} \qquad (T3.23)\\
&\quad -\frac{\tau}{\widetilde{\tau}}\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'} + \left(\frac{\tau}{\widetilde{\tau}}\right)^2\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}}\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}'}\\
&= \frac{\tau}{\widetilde{\tau}}\left(\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}}\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}'} - \frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'}\right)\\
&= \frac{\tau}{\widetilde{\tau}}\left(-\left(-i\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}}\right)\left(-i\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}'}\right) + \left(-\frac{\partial \phi_{\mathbf{X}_{T,\widetilde{\tau}}}(\mathbf{0})}{\partial \boldsymbol{\omega}\partial \boldsymbol{\omega}'}\right)\right)
\end{aligned}$$

Using again ($T2.88$) in ($T3.21$) we obtain:

$$\mathrm{E}\{\mathbf{X}_{T,\tau}\} = \frac{\tau}{\widetilde{\tau}}\,\mathrm{E}\{\mathbf{X}_{T,\widetilde{\tau}}\}; \tag{T3.24}$$

Using again ($T2.88$) in ($T3.23$) we obtain:

$$\begin{aligned}
\mathrm{Cov}\{\mathbf{X}_{T,\tau}\} &= \frac{\tau}{\widetilde{\tau}}\left(-\mathrm{E}\{\mathbf{X}_{T,\widetilde{\tau}}\}\,\mathrm{E}\{\mathbf{X}_{T,\widetilde{\tau}}\}' + \mathrm{E}\left\{\mathbf{X}_{T,\widetilde{\tau}}\mathbf{X}'_{T,\widetilde{\tau}}\right\}\right) \qquad (T3.25)\\
&= \frac{\tau}{\widetilde{\tau}}\left(\mathrm{Cov}\{\mathbf{X}_{T,\widetilde{\tau}}\}\right).
\end{aligned}$$

The statement in the main text "More in general, a multiplicative relation such as ($T3.24$) or ($T3.25$) holds for all the raw moments and all the central moments, when they are defined" is incorrect: it only holds for the expected value and the covariance.

3.4 Results on regression dimension reduction

Regression in the general case

From the definition (3.116) of the generalized r-square and (3.120), the regression factor loadings minimizes the following quantity:

$$\begin{aligned} M &\equiv \mathrm{E}\left\{(\mathbf{X}-\mathbf{BF})'(\mathbf{X}-\mathbf{BF})\right\} \\ &= \sum_{n,k,j} \mathrm{E}\left\{(X_n - B_{nk}F_k)(X_n - B_{nj}F_j)\right\} \\ &= \sum_n \mathrm{E}\left\{X_n^2\right\} - \sum_{n,k} \mathrm{E}\left\{B_{nk}F_k X_n\right\} \\ &\quad - \sum_{n,j} \mathrm{E}\left\{X_n B_{nj} F_j\right\} + \sum_{n,k,j} \mathrm{E}\left\{B_{nk}F_k B_{nj}F_j\right\} \\ &= \sum_n \mathrm{E}\left\{X_n^2\right\} - 2\sum_{n,k} B_{nk}\,\mathrm{E}\left\{X_n F_k\right\} \\ &\quad + \sum_{n,k,j} B_{nj}B_{nk}\,\mathrm{E}\left\{F_k F_j\right\}. \end{aligned} \tag{T3.26}$$

Therefore the first order conditions with respect to B_{sl} read:

$$\begin{aligned} 0_{sl} &= \frac{\partial M}{\partial B_{sl}} \\ &= 2\,\mathrm{tr}\left\{\mathrm{E}\left\{\mathbf{XF}'\right\}\mathbf{B}'\right\} + \mathrm{tr}\left\{\mathbf{B}\,\mathrm{E}\left\{\mathbf{FF}'\right\}\mathbf{B}'\right\} \\ &= -2\,\mathrm{E}\left\{X_s F_l\right\} + \frac{\partial}{\partial B_{sl}}\left(\sum_{n,k,j} B_{nj}B_{nk}\,\mathrm{E}\left\{F_k F_j\right\}\right) \\ &= -2\,\mathrm{E}\left\{X_s F_l\right\} + 2\sum_k B_{sk}\,\mathrm{E}\left\{F_k F_l\right\}. \end{aligned} \tag{T3.27}$$

In matrix notation these equations read:

$$\mathbf{0}_{N\times K} = -2\,\mathrm{E}\left\{\mathbf{XF}'\right\} + 2\mathbf{B}\,\mathrm{E}\left\{\mathbf{FF}'\right\}, \tag{T3.28}$$

whose solution is:

$$\mathbf{B}_r = \mathrm{E}\left\{\mathbf{XF}'\right\}\mathrm{E}\left\{\mathbf{FF}'\right\}^{-1}. \tag{T3.29}$$

The residuals read:

$$\begin{aligned} \mathbf{U} &\equiv \mathbf{X} - \mathbf{B}_r\mathbf{F} \\ &= \mathbf{X} - \mathrm{E}\left\{\mathbf{XF}'\right\}\mathrm{E}\left\{\mathbf{FF}'\right\}^{-1}\mathbf{F} \\ &= \mathbf{X} - \left(\mathrm{Cov}\left\{\mathbf{X},\mathbf{F}\right\} - \mathrm{E}\left\{\mathbf{X}\right\}\mathrm{E}\left\{\mathbf{F}'\right\}\right)\mathrm{E}\left\{\mathbf{FF}'\right\}^{-1}\mathbf{F}. \end{aligned} \tag{T3.30}$$

In general, the residuals do not have zero expected value:

$$\begin{aligned} \mathrm{E}\{\mathbf{U}\} &= \mathrm{E}\left\{\mathbf{X} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathbf{F}\right\} \\ &= \mathrm{E}\{\mathbf{X}\} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathrm{E}\{\mathbf{F}\}\,. \end{aligned} \tag{T3.31}$$

Furthermore, in general the residuals are correlated:

$$\begin{aligned} \mathrm{Cov}\{\mathbf{U},\mathbf{F}\} &= \mathrm{E}\{\mathbf{U}\mathbf{F}'\} - \mathrm{E}\{\mathbf{U}\}\,\mathrm{E}\{\mathbf{F}'\} \\ &= \mathrm{E}\left\{\left(\mathbf{X} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathbf{F}\right)\mathbf{F}'\right\} \\ &\quad - \mathrm{E}\left\{\mathbf{X} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathbf{F}\right\}\mathrm{E}\{\mathbf{F}'\} \\ &= \mathrm{E}\{\mathbf{X}\mathbf{F}'\} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathrm{E}\{\mathbf{F}\mathbf{F}'\} \\ &\quad - \mathrm{E}\{\mathbf{X}\}\,\mathrm{E}\{\mathbf{F}'\} + \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathrm{E}\{\mathbf{F}\}\,\mathrm{E}\{\mathbf{F}'\} \\ &= \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathrm{E}\{\mathbf{F}\}\,\mathrm{E}\{\mathbf{F}'\} - \mathrm{E}\{\mathbf{X}\}\,\mathrm{E}\{\mathbf{F}'\}\,. \end{aligned} \tag{T3.32}$$

This expression is zero if

$$\mathrm{E}\{\mathbf{F}\} = \mathbf{0}. \tag{T3.33}$$

Recalling (T3.29), we can also express the covariance of the residuals with the factor as follows

$$\begin{aligned} \mathrm{Cov}\{\mathbf{U},\mathbf{F}\} &= \mathrm{E}\{\mathbf{U}\mathbf{F}'\} - \mathrm{E}\{\mathbf{U}\}\,\mathrm{E}\{\mathbf{F}'\} \\ &= \mathrm{E}\left\{\left(\mathbf{X} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathbf{F}\right)\mathbf{F}'\right\} \\ &\quad - \mathrm{E}\left\{\mathbf{X} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathbf{F}\right\}\mathrm{E}\{\mathbf{F}'\} \\ &= \mathrm{E}\{\mathbf{X}\mathbf{F}'\} - \mathrm{E}\{\mathbf{X}\}\,\mathrm{E}\{\mathbf{F}'\} \\ &\quad + \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\left(\mathrm{E}\{\mathbf{F}\}\,\mathrm{E}\{\mathbf{F}'\} - \mathrm{E}\{\mathbf{F}\mathbf{F}'\}\right) \\ &= \mathrm{Cov}\{\mathbf{X},\mathbf{F}\} - \mathrm{E}\{\mathbf{X}\mathbf{F}'\}\,\mathrm{E}\{\mathbf{F}\mathbf{F}'\}^{-1}\mathrm{Cov}\{\mathbf{F}\} \\ &= \mathrm{Cov}\{\mathbf{X},\mathbf{F}\} - \mathbf{B}_r\,\mathrm{Cov}\{\mathbf{F}\}\,. \end{aligned} \tag{T3.34}$$

Regression with constant among factors

Assume one of the factors is a constant as in (3.126). Then the linear model (3.119) becomes

$$\mathbf{X} \equiv \mathbf{a} + \mathbf{G}\mathbf{F} + \mathbf{U}. \tag{T3.35}$$

In order to maximize the generalized r-square (3.116) we have to minimize the following expression:

$$\begin{aligned} M &\equiv \mathrm{E}\left\{[\mathbf{X} - (\mathbf{a} + \mathbf{G}\mathbf{F})]'\,[\mathbf{X} - (\mathbf{a} + \mathbf{G}\mathbf{F})]\right\} \\ &= \mathrm{E}\{\mathbf{X}'\mathbf{X}\} + \mathbf{a}'\mathbf{a} + \mathrm{E}\{\mathbf{F}'\mathbf{G}'\mathbf{G}\mathbf{F}\} \\ &\quad + 2\mathbf{a}'\mathbf{G}\,\mathrm{E}\{\mathbf{F}\} - 2\mathbf{a}'\,\mathrm{E}\{\mathbf{X}\} - 2\,\mathrm{E}\{\mathbf{X}'\mathbf{G}\mathbf{F}\}\,. \end{aligned} \tag{T3.36}$$

We re-write (T3.36) emphasizing the terms containing $\mathbf{a}$:

$$M = \cdots + \sum_j (a_j)^2 + 2 \sum_{j,k} a_j G_{jk} \operatorname{E}\{F_k\} - 2 \sum_j a_j \operatorname{E}\{X_j\} \qquad (T3.37)$$

Setting to zero the first order derivative with respect to a_j we obtain

$$a_j = \operatorname{E}\{X_j\} - \sum_k G_{jk} \operatorname{E}\{F_k\} \qquad (T3.38)$$

which in matrix notation yields

$$\mathbf{a}_r = \operatorname{E}\{\mathbf{X}\} - \mathbf{G}_r \operatorname{E}\{\mathbf{F}\} \qquad (T3.39)$$

Now we re-write ($T3.36$) emphasizing the terms containing $\mathbf{G}$:

$$\begin{aligned} M &= \cdots + \sum_{jkl} \operatorname{E}\{F_j G_{kj} G_{kl} F_l\} \qquad (T3.40) \\ &\quad +2 \sum_{jk} a_j G_{jk} \operatorname{E}\{F_k\} - 2 \sum_{jk} \operatorname{E}\{X_j G_{jk} F_k\} \\ &= \sum_{jkl} \operatorname{E}\{F_k G_{jk} G_{jl} F_l\} + 2 \sum_{jk} a_j G_{jk} \operatorname{E}\{F_k\} - 2 \sum_{jk} \operatorname{E}\{X_j G_{jk} F_k\} \end{aligned}$$

Setting to zero the first order derivative with respect to G_{jk} and using ($T3.38$) we obtain:

$$\begin{aligned} 0 &= \sum_l \operatorname{E}\{F_k G_{jl} F_l\} + a_j \operatorname{E}\{F_k\} - \operatorname{E}\{X_j F_k\} \qquad (T3.41) \\ &= \sum_l \operatorname{E}\{F_k G_{jl} F_l\} + \left(\operatorname{E}\{X_j\} - \sum_l G_{jl} \operatorname{E}\{F_l\} \right) \operatorname{E}\{F_k\} - \operatorname{E}\{X_j F_k\} \\ &= \left(\sum_l \operatorname{E}\{F_k G_{jl} F_l\} - \sum_l \operatorname{E}\{G_{jl} F_l\} \operatorname{E}\{F_k\} \right) \\ &\quad - (\operatorname{E}\{X_j F_k\} - \operatorname{E}\{X_j\} \operatorname{E}\{F_k\}) \\ &= \sum_l \operatorname{Cov}\{G_{jl} F_l, F_k\} - \operatorname{Cov}\{X_j, F_k\} \\ &= \operatorname{Cov}\left\{ [\mathbf{GF}]_j , F_k \right\} - \operatorname{Cov}\{X_j, F_k\} . \end{aligned}$$

In matrix notation this expression reads:

$$\mathbf{G} \operatorname{Cov}\{\mathbf{F}\} = \operatorname{Cov}\{\mathbf{X}, \mathbf{F}\} , \qquad (T3.42)$$

which implies

$$\mathbf{G}_r = \operatorname{Cov}\{\mathbf{X}, \mathbf{F}\} \operatorname{Cov}\{\mathbf{F}\}^{-1} . \qquad (T3.43)$$

Substituting ($T3.38$) and ($T3.42$) in ($T3.35$) we find the expression of the recovered invariants:

$$\widetilde{\mathbf{X}}_r \equiv \mathrm{E}\{\mathbf{X}\} + \mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F}\}^{-1}(\mathbf{F} - \mathrm{E}\{\mathbf{F}\}). \tag{T3.44}$$

The residuals read:

$$\mathbf{U}_r \equiv \mathbf{X} - \widetilde{\mathbf{X}}_r = \overline{\mathbf{X}} - \mathbf{G}_r\overline{\mathbf{F}},$$

where

$$\overline{\mathbf{X}} \equiv \mathbf{X} - \mathrm{E}\{\mathbf{X}\}, \quad \overline{\mathbf{F}} \equiv \mathbf{F} - \mathrm{E}\{\mathbf{F}\}. \tag{T3.45}$$

Therefore the residuals have zero expected value.

The covariance of the residuals with the factors reads:

$$\begin{aligned} \mathrm{Cov}\{\mathbf{U}_r,\mathbf{F}\} &= \mathrm{E}\left\{\left[\overline{\mathbf{X}} - \mathbf{G}_r\overline{\mathbf{F}}\right]\overline{\mathbf{F}}'\right\} \\ &= \mathrm{E}\left\{\overline{\mathbf{X}}\overline{\mathbf{F}}'\right\} - \mathbf{G}_r\,\mathrm{E}\left\{\overline{\mathbf{F}}\overline{\mathbf{F}}'\right\} \\ &= \mathrm{Cov}\{\mathbf{X},\mathbf{F}\} - \mathbf{G}_r\,\mathrm{Cov}\{\mathbf{F}\} \end{aligned} \tag{T3.46}$$

Therefore using $(T3.43)$ we obtain:

$$\begin{aligned} \mathrm{Cov}\{\mathbf{U}_r,\mathbf{F}\} &= \mathrm{Cov}\{\mathbf{X},\mathbf{F}\} - \mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F}\}^{-1}\,\mathrm{Cov}\{\mathbf{F}\} \\ &= \mathbf{0} \end{aligned} \tag{T3.47}$$

The covariance of the residual reads:

$$\begin{aligned} \mathrm{Cov}\{\mathbf{U}_r\} &= \mathrm{E}\left\{\left[\overline{\mathbf{X}} - \mathbf{G}_r\overline{\mathbf{F}}\right]\left[\overline{\mathbf{X}} - \mathbf{G}_r\overline{\mathbf{F}}\right]'\right\} \\ &\mathrm{E}\left\{\overline{\mathbf{X}}\overline{\mathbf{X}}'\right\} - 2\,\mathrm{E}\left\{\overline{\mathbf{X}}\overline{\mathbf{F}}'\right\}\mathbf{G}_r' + \mathbf{G}_r\,\mathrm{E}\left\{\overline{\mathbf{F}}\overline{\mathbf{F}}'\right\}\mathbf{G}_r' \\ &= \mathrm{Cov}\{\mathbf{X},\mathbf{X}\} - 2\,\mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathbf{G}_r' + \mathbf{G}_r\,\mathrm{Cov}\{\mathbf{F},\mathbf{F}\}\,\mathbf{G}_r' \end{aligned} \tag{T3.48}$$

Therefore using $(T3.43)$ we obtain:

$$\mathrm{Cov}\{\mathbf{U}_r\} = \mathrm{Cov}\{\mathbf{X},\mathbf{X}\} - \mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F},\mathbf{F}\}^{-1}\,\mathrm{Cov}\{\mathbf{F},\mathbf{X}\} \tag{T3.49}$$

PCA analysis of regression

First of all we consider a scale independent model. We recall from (1.35) that the z-scores of the variables are defined as follows:

$$\mathbf{Z}_X \equiv \mathbf{D}_X^{-1}(\mathbf{X} - \mathrm{E}\{\mathbf{X}\}) \tag{T3.50}$$

where

$$\mathbf{D}_X \equiv \mathrm{diag}(\mathrm{Sd}\{X_1\},\ldots,\mathrm{Sd}\{X_N\}). \tag{T3.51}$$

Left-multiplying $(T3.44)$ by $\mathbf{D}_X^{-1}$ we obtain the recovered z-score of the original variable $\mathbf{X}$:

$$\begin{aligned}\widetilde{\mathbf{Z}}_X &\equiv \mathbf{D}_X^{-1}\left(\widetilde{\mathbf{X}}_r - \mathrm{E}\{\mathbf{X}\}\right) \\ &= \mathbf{D}_X^{-1}\,\mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F}\}^{-1}(\mathbf{F}-\mathrm{E}\{\mathbf{F}\}) \\ &= \mathrm{Cov}\left\{\mathbf{D}_X^{-1}\mathbf{X},\mathbf{F}\right\}\mathrm{Cov}\{\mathbf{F}\}^{-1}(\mathbf{F}-\mathrm{E}\{\mathbf{F}\})\end{aligned} \tag{T3.52}$$

Consider the spectral decomposition (2.76) of the covariance of the factors:

$$\mathrm{Cov}\{\mathbf{F}\} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}', \tag{T3.53}$$

where $\mathbf{E}$ is the juxtaposition of the eigenvectors:

$$\mathbf{E} \equiv \left(\mathbf{e}^{(1)},\ldots,\mathbf{e}^{(K)}\right), \tag{T3.54}$$

and satisfies $\mathbf{E}\mathbf{E}' = \mathbf{I}_K$, the identity matrix; and $\boldsymbol{\Lambda}$ is the diagonal matrix of the eigenvalues sorted in decreasing order:

$$\boldsymbol{\Lambda} \equiv \mathrm{diag}(\lambda_1,\ldots,\lambda_K). \tag{T3.55}$$

With the spectral decomposition we can always rotate the factors in such a way that they are uncorrelated. Indeed the rotated factors $\mathbf{E}'\underline{\mathbf{F}}$ satisfy:

$$\mathrm{Cov}\{\mathbf{E}'\mathbf{F}\} = \mathbf{E}'\,\mathrm{Cov}\{\mathbf{F}\}\,\mathbf{E} = \mathbf{E}'\mathbf{E}\boldsymbol{\Lambda}\mathbf{E}'\mathbf{E} = \boldsymbol{\Lambda}, \tag{T3.56}$$

which is diagonal. Now consider the z-scores of the rotated factors:

$$\mathbf{Z}_F \equiv \boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{F}-\mathrm{E}\{\mathbf{F}\}), \tag{T3.57}$$

which are uncorrelated and have unit standard deviation:

$$\mathrm{Cov}\{\mathbf{Z}_F\} = \boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'\mathbf{E}\boldsymbol{\Lambda}\mathbf{E}'\mathbf{E}\boldsymbol{\Lambda}^{-\frac{1}{2}} = \mathbf{I}_K. \tag{T3.58}$$

From (T3.52) the recovered z-score of the original variable $\mathbf{X}$ reads:

$$\begin{aligned}\widetilde{\mathbf{Z}}_X &= \mathrm{Cov}\left\{\mathbf{D}_X^{-1}\mathbf{X},\mathbf{F}\right\}\mathbf{E}\boldsymbol{\Lambda}^{-1}\mathbf{E}'(\mathbf{F}-\mathrm{E}\{\mathbf{F}\}) \\ &= \mathrm{Cov}\left\{\mathbf{D}_X^{-1}\mathbf{X},\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{F}-\mathrm{E}\{\mathbf{F}\})\right\}\left[\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{F}-\mathrm{E}\{\mathbf{F}\})\right] \\ &= \mathrm{Cor}\{\mathbf{X},\mathbf{E}'\mathbf{F}\}\left[\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{F}-\mathrm{E}\{\mathbf{F}\})\right],\end{aligned} \tag{T3.59}$$

On the other hand, the generalized r-square defined in (3.116) reads in this context:

$$\begin{aligned}R^2\left\{\mathbf{X},\widetilde{\mathbf{X}}_r\right\} &\equiv R^2\left\{\mathbf{Z}_X,\widetilde{\mathbf{Z}}_X\right\} \\ &\equiv 1-\frac{\mathrm{E}\left\{\left(\mathbf{Z}_X-\widetilde{\mathbf{Z}}_X\right)'\left(\mathbf{Z}_X-\widetilde{\mathbf{Z}}_X\right)\right\}}{\mathrm{tr}\{\mathrm{Cov}\{\mathbf{Z}_X\}\}} \\ &= 1-\frac{a}{N}.\end{aligned} \tag{T3.60}$$

The term in the numerator can be written as follows:

$$\begin{aligned} a &\equiv \mathrm{E}\left\{\left(\mathbf{Z}_X-\widetilde{\mathbf{Z}}_X\right)'\left(\mathbf{Z}_X-\widetilde{\mathbf{Z}}_X\right)\right\} \\ &= \mathrm{E}\left\{\left[\mathbf{D}_X^{-1}\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)\right]'\mathbf{D}_X^{-1}\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)\right\} \\ &= \mathrm{E}\left\{\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)'\mathbf{D}_X^{-1}\mathbf{D}_X^{-1}\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)\right\} \\ &= \mathrm{tr}\left(\mathbf{D}_X^{-1}\mathbf{D}_X^{-1}\,\mathrm{E}\left\{\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)\left(\mathbf{X}-\widetilde{\mathbf{X}}\right)'\right\}\right) \end{aligned} \tag{T3.61}$$

using (T3.49) this becomes:

$$\begin{aligned} a &\equiv \mathrm{tr}\left(\mathbf{D}_X^{-1}\mathbf{D}_X^{-1}\left[\mathrm{Cov}\{\mathbf{X},\mathbf{X}\}-\mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F},\mathbf{F}\}^{-1}\,\mathrm{Cov}\{\mathbf{F},\mathbf{X}\}\right]\right) \\ &= \mathrm{tr}\left(\mathbf{D}_X^{-1}\mathbf{D}_X^{-1}\,\mathrm{Cov}\{\mathbf{X},\mathbf{X}\}\right) \\ &\quad -\mathrm{tr}\left(\mathbf{D}_X^{-1}\mathbf{D}_X^{-1}\,\mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F},\mathbf{F}\}^{-1}\,\mathrm{Cov}\{\mathbf{F},\mathbf{X}\}\right) \\ &= \mathrm{tr}\left(\mathrm{Cor}\{\mathbf{X},\mathbf{X}\}\right) \\ &\quad -\mathrm{tr}\left(\mathbf{D}_X^{-1}\mathbf{D}_X^{-1}\,\mathrm{Cov}\{\mathbf{X},\mathbf{F}\}\,\mathrm{Cov}\{\mathbf{F},\mathbf{F}\}^{-1}\,\mathrm{Cov}\{\mathbf{F},\mathbf{X}\}\right) \\ &= N-\mathrm{tr}\left(\mathrm{Cov}\left\{\mathbf{D}_X^{-1}\mathbf{X},\mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{Z}_F\right\}\mathbf{E}\boldsymbol{\Lambda}^{-1}\mathbf{E}'\,\mathrm{Cov}\left\{\mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{Z}_F,\mathbf{D}_X^{-1}\mathbf{X}\right\}\right) \\ .&= N-\mathrm{tr}\left(\mathrm{Cov}\left\{\mathbf{D}_X^{-1}\mathbf{X},\mathbf{Z}_F\right\}\mathrm{Cov}\left\{\mathbf{Z}_F,\mathbf{D}_X^{-1}\mathbf{X}\right\}\right) \\ &\quad N-\mathrm{tr}\left(\mathrm{Cor}\{\mathbf{X},\mathbf{E}'\mathbf{F}\}\,\mathrm{Cor}\{\mathbf{E}'\mathbf{F},\mathbf{X}\}\right) \end{aligned} \tag{T3.62}$$

Therefore the r-square (T3.60) reads:

$$\begin{aligned} R^2\left\{\mathbf{Z}_X,\widetilde{\mathbf{Z}}_X\right\} &= 1-\frac{N-\mathrm{tr}\left(\mathrm{Cor}\{\mathbf{X},\mathbf{E}'\mathbf{F}\}\,\mathrm{Cor}\{\mathbf{E}'\mathbf{F},\mathbf{X}\}\right)}{N} \\ &= \frac{1}{N}\mathrm{tr}\left(\mathrm{Cor}\{\mathbf{X},\mathbf{E}'\mathbf{F}\}\,\mathrm{Cor}\{\mathbf{E}'\mathbf{F},\mathbf{X}\}\right). \end{aligned} \tag{T3.63}$$

3.5 Results on PCA dimension reduction

Recovered invariants as projection

The PCA-recovered invariants (3.160) read:

$$\widetilde{\mathbf{X}}_p \equiv \mathbf{a}+\mathbf{G}\mathbf{X}, \tag{T3.64}$$

where

$$\mathbf{a} \equiv \left(\mathbf{I}_N-\mathbf{E}_K\mathbf{E}_K'\right)\mathrm{E}\{\mathbf{X}\}, \tag{T3.65}$$

and

$$\mathbf{G} \equiv \mathbf{E}_K \mathbf{E}_K'. \tag{T3.66}$$

Consider a generic point $\mathbf{x}$ in $\mathbb{R}^N$. Since the eigenvectors of the covariance matrix are a basis of $\mathbb{R}^N$ we can express $\mathbf{x}$ as follows:

$$\mathbf{x} \equiv \mathrm{E}\{\mathbf{X}\} + \sum_{n=1}^{N} \alpha_n \mathbf{e}^{(n)}, \tag{T3.67}$$

for suitable coefficients $\{\alpha_1, \ldots, \alpha_N\}$, where $\mathbf{e}^{(n)}$ denotes the n-th eigenvector.

To prove that (T3.64) represents the projection on the hyperplane of maximal variation generated by the first K principal axes we need to prove the following relation:

$$\mathbf{a} + \mathbf{G}\mathbf{x} = \mathrm{E}\{\mathbf{X}\} + \sum_{n=1}^{K} \alpha_n \mathbf{e}^{(n)}. \tag{T3.68}$$

By substituting (T3.65), (T3.66) and (T3.67) in the left hand side of the above relation we obtain:

$$\begin{aligned} \mathbf{a} + \mathbf{G}\mathbf{x} &\equiv (\mathbf{I}_N - \mathbf{E}_K \mathbf{E}_K') \, \mathrm{E}\{\mathbf{X}\} \\ &\quad + \mathbf{E}_K \mathbf{E}_K' \left(\mathrm{E}\{\mathbf{X}\} + \sum_{n=1}^{N} \alpha_n \mathbf{e}^{(n)} \right) \\ &= \mathrm{E}\{\mathbf{X}\} + \sum_{n=1}^{N} \alpha_n \mathbf{E}_K \mathbf{E}_K' \mathbf{e}^{(n)}. \end{aligned} \tag{T3.69}$$

Therefore in order prove our statement it suffices to prove that if $n \leq K$ then the following holds:

$$\mathbf{E}_K \mathbf{E}_K' \mathbf{e}^{(n)} = \mathbf{e}^{(n)}, \tag{T3.70}$$

and if $n > K$ then the following holds:

$$\mathbf{E}_K \mathbf{E}_K' \mathbf{e}^{(n)} = \mathbf{0}. \tag{T3.71}$$

Both statements follow from the definition (3.157) of $\mathbf{E}_K$, which implies:

$$\mathbf{E}_K \mathbf{E}_K' \mathbf{e}^{(n)} \equiv \left(\mathbf{e}^{(1)} | \cdots | \mathbf{e}^{(K)} \right) \begin{pmatrix} \left[\mathbf{e}^{(1)}\right]' \mathbf{e}^{(n)} \\ \vdots \\ \left[\mathbf{e}^{(K)}\right]' \mathbf{e}^{(n)} \end{pmatrix}, \tag{T3.72}$$

and the fact that $\mathbf{E}$ is orthonormal:

$$\mathbf{E}\mathbf{E}' = \mathbf{I}_N, \tag{T3.73}$$

where $\mathbf{I}$ is the identity matrix.

Results on the residual

The residual of the PCA dimension reduction reads:

$$\begin{aligned}\mathbf{U}_p &\equiv \mathbf{X} - \widetilde{\mathbf{X}}_p \equiv \left(\mathbf{I}_N - \mathbf{E}_K \mathbf{E}_K'\right)\left(\mathbf{X} - \mathrm{E}\left\{\mathbf{X}\right\}\right) \\ &= \mathbf{R}_K \mathbf{R}_K' \left(\mathbf{X} - \mathrm{E}\left\{\mathbf{X}\right\}\right),\end{aligned} \tag{T3.74}$$

where $\mathbf{R}_K$ is the juxtaposition of the last $(N-K)$ eigenvectors:

$$\mathbf{R}_K \equiv \left(\mathbf{e}^{(K+1)} | \cdots | \mathbf{e}^{(N)}\right). \tag{T3.75}$$

This matrix satisfies

$$\mathbf{R}_K \mathbf{R}_K' \mathbf{R}_K \mathbf{R}_K' = \mathbf{R}_K \mathbf{R}_K'. \tag{T3.76}$$

Indeed, from $(T3.73)$ and the definition of $\mathbf{R}_K$ we obtain:

$$\mathbf{R}_K \mathbf{R}_K' \mathbf{e}^{(n)} \equiv \left(\mathbf{e}^{(K+1)} | \cdots | \mathbf{e}^{(N)}\right) \begin{pmatrix} \left[\mathbf{e}^{(K+1)}\right]' \mathbf{e}^{(n)} \\ \vdots \\ \left[\mathbf{e}^{(N)}\right]' \mathbf{e}^{(n)} \end{pmatrix}. \tag{T3.77}$$

Therefore, if $n > K$ then:

$$\mathbf{R}_K \mathbf{R}_K' \mathbf{e}^{(n)} = \mathbf{e}^{(n)}, \tag{T3.78}$$

and if $n \leq K$ then:

$$\mathbf{R}_K \mathbf{R}_K' \mathbf{e}^{(n)} = \mathbf{0}. \tag{T3.79}$$

Since the set of eigenvectors is a basis in $\mathbb{R}^N$, $(T3.78)$ and $(T3.79)$ prove $(T3.76)$.

Therefore the term in the numerator of the generalized r-square (3.116) of the PCA dimension reduction reads:

$$\begin{aligned}M &\equiv \mathrm{E}\left\{\left(\mathbf{X} - \widetilde{\mathbf{X}}_p\right)' \left(\mathbf{X} - \widetilde{\mathbf{X}}_p\right)\right\} \\ &= \mathrm{E}\left\{\left(\mathbf{X} - \mathrm{E}\left\{\mathbf{X}\right\}\right)' \mathbf{R}_K \mathbf{R}_K' \mathbf{R}_K \mathbf{R}_K' \left(\mathbf{X} - \mathrm{E}\left\{\mathbf{X}\right\}\right)\right\} \\ &= \mathrm{E}\left\{\left(\mathbf{X} - \mathrm{E}\left\{\mathbf{X}\right\}\right)' \mathbf{R}_K \mathbf{R}_K' \left(\mathbf{X} - \mathrm{E}\left\{\mathbf{X}\right\}\right)\right\} \\ &= \mathrm{tr}\left(\mathrm{Cov}\left\{\mathbf{R}_K' \mathbf{X}\right\}\right),\end{aligned} \tag{T3.80}$$

On the other hand from the definition $(T3.75)$ of $\mathbf{R}_K$ we obtain:

$$\mathbf{E}' \mathbf{R}_K = \left(\boldsymbol{\delta}^{(K+1)} | \cdots | \boldsymbol{\delta}^{(N)}\right), \tag{T3.81}$$

where $\boldsymbol{\delta}^{(n)}$ is the n-th element of the canonical basis $(A.15)$. Therefore

$$\begin{aligned}\mathrm{Cov}\left\{\mathbf{R}_K' \mathbf{X}\right\} &= \mathbf{R}_K' \,\mathrm{Cov}\left\{\mathbf{X}\right\} \mathbf{R}_K \\ &= \mathbf{R}_K' \mathbf{E} \boldsymbol{\Lambda} \mathbf{E}' \mathbf{R}_K \\ &= \mathrm{diag}\left(\lambda_{K+1}, \ldots, \lambda_N\right)\end{aligned} \tag{T3.82}$$

Substituting ($T3.82$) in ($T3.80$) we obtain:

$$M = \sum_{n=K+1}^{N} \lambda_n. \qquad (T3.83)$$

The term in the denominator of the generalized r-square (3.116) is the sum of all the eigenvalues. This follows from (3.149) and ($A.67$). Therefore, the generalize r-square reads:

$$R^2 \left\{ \mathbf{X}, \widetilde{\mathbf{X}}_p \right\} = 1 - \frac{\sum_{n=K+1}^{N} \lambda_n}{\sum_{n=1}^{N} \lambda_n.} = \frac{\sum_{n=1}^{K} \lambda_n}{\sum_{n=1}^{N} \lambda_n.}. \qquad (T3.84)$$

The residual ($T3.74$) clearly has zero expected value. Similarly, the factors

$$\mathbf{F}_p \equiv \mathbf{E}_K' (\mathbf{X} - \mathrm{E}\{\mathbf{X}\}) \qquad (T3.85)$$

have zero expected value. From ($T3.78$) and ($T3.79$) we obtain:

$$\mathbf{R}_K \mathbf{R}_K' \mathbf{E} = \left(\mathbf{0} | \cdots | \mathbf{0} | \mathbf{e}^{(K+1)} | \cdots | \mathbf{e}^{(N)} \right). \qquad (T3.86)$$

Similarly:

$$(\mathbf{E}' \mathbf{E}_K) = \left(\mathbf{e}^{(1)} | \cdots | \mathbf{e}^{(K)} \right). \qquad (T3.87)$$

Therefore the covariance of the residuals with the factors reads:

$$\begin{aligned} \mathrm{Cov}\{\mathbf{U}_p, \mathbf{F}_p\} &= \mathrm{E}\left\{ \mathbf{U}_p \mathbf{F}_p' \right\} \qquad (T3.88) \\ &= \mathrm{E}\left\{ \mathbf{R}_K \mathbf{R}_K' (\mathbf{X} - \mathrm{E}\{\mathbf{X}\}) (\mathbf{X} - \mathrm{E}\{\mathbf{X}\})' \mathbf{E}_K \right\} \\ &= \mathbf{R}_K \mathbf{R}_K' \,\mathrm{Cov}\{\mathbf{X}\}\, \mathbf{E}_K \\ &= (\mathbf{R}_K \mathbf{R}_K' \mathbf{E}) \,\mathbf{\Lambda}\, (\mathbf{E}' \mathbf{E}_K) \\ &= \left(\mathbf{0} | \cdots | \mathbf{0} | \mathbf{e}^{(K+1)} | \cdots | \mathbf{e}^{(N)} \right) \left(\lambda_1 \mathbf{e}^{(1)} | \cdots | \lambda_K \mathbf{e}^{(K)} \right) \\ &= \mathbf{0}_{N \times K}, \end{aligned}$$

where the last equality follows from $\mathbf{E}\mathbf{E}' = \mathbf{I}_N$.

3.6 Spectral basis in the continuum

In order to better capture the analogies between the continuum and the discrete case, we advise the reader to refer to Appendix B at the end of the book, and in particular to the (rationale behind) Tables B.4, B.11 and B.20.

First of all, we consider a generic Toeplitz operator S defined on $L_2(\mathbb{R})$, i.e. an operator whose kernel representation $S(x, y)$ vanishes fast enough at infinity and satisfies:

$$S(x + z, y) = S(x, y - z). \qquad (T3.89)$$

Suppose that the operator admits a one-dimensional eigenvalue/eigenfunction pair, i.e. there exist a number λ_ω and a function

$$S\left[e^{(\omega)}\right] = \lambda_\omega e^{(\omega)}, \qquad (T3.90)$$

where the function is unique up to a constant. Using Table B.4, the spectral equation $(T3.90)$ reads explicitly as follows:

$$\int_{\mathbb{R}} S(x,y)\, e^{(\omega)}(y)\, dy = \lambda_\omega e^{(\omega)}(x). \qquad (T3.91)$$

First of all we determine the generic form of such an eigenfunction, if it exists.

Expanding in Taylor series the spectral basis and using the spectral equation we obtain

$$\begin{aligned} \lambda_\omega e^{(\omega)}(x+dx) &= \int_{\mathbb{R}} S(x+dx,y)\, e^{(\omega)}(y)\, dy \\ &= \int_{\mathbb{R}} S(x,y-dx)\, e^{(\omega)}(y)\, dy \qquad (T3.92) \\ &= \int_{\mathbb{R}} \left[S(x,y) - \partial_y S(x,y)\, dx\right] e^{(\omega)}(y)\, dy. \end{aligned}$$

On the other hand from another Taylor expansion we obtain:

$$\lambda_\omega e^{(\omega)}(x+dx) = \lambda_\omega \left[e^{(\omega)}(x) + \frac{de^{(\omega)}}{dx} dx\right]. \qquad (T3.93)$$

Therefore, integrating by parts and using the assumption that the matrix S vanishes at infinity, we obtain the following identity:

$$\begin{aligned} \lambda_\omega \frac{de^{(\omega)}}{dx} &= -\int_{\mathbb{R}} \left(\partial_y S(x,y)\right) e^{(\omega)}(y)\, dy \qquad (T3.94) \\ &= -\int_{\mathbb{R}} \partial_y \left[S(x,y)\, e^{(\omega)}(y)\right] dy + \int_{\mathbb{R}} S(x,y)\, \partial_y e^{(\omega)}(y)\, dy \\ &= \int_{\mathbb{R}} S(x,y)\, \partial_y e^{(\omega)}(y)\, dy, \end{aligned}$$

or, in terms of the (one-dimensional) derivative operator $(B.25)$:

$$S\left[\mathcal{D}e^{(\omega)}\right] = \lambda_\omega \left[\mathcal{D}e^{(\omega)}\right], \qquad (T3.95)$$

Therefore $\mathcal{D}e^{(\omega)}$ is an eigenvector relative to the same eigenvalue as $e^{(\omega)}$.

Now assume that the Toeplitz operator S is symmetric and positive. Similarly to $(A.51)$ the operator is *symmetric* if its kernel is symmetric across the diagonal, i.e.

$$S(x,y) = \overline{S(y,x)}. \qquad (T3.96)$$

Similarly to $(A.52)$ the operator is *positive* if for any function v in its domain the following is true:

$$\langle v, S[v]\rangle \equiv \int_{\mathbb{R}} \overline{v(x)} \int_{\mathbb{R}} S(x,y)\, v(y)\, dy dx \geq 0. \qquad (T3.97)$$

In this case we can restate the spectral theorem in the continuum making use of the formal substitutions in Tables B.4, B.11 and B.20: if the kernel representation S of a linear operator satisfies $(T3.96)$ and $(T3.97)$, then the operator admits an orthogonal basis of *eigenfunctions*.

In other words, then there exists a set of functions $\left\{e^{(\omega)}(\cdot)\right\}_{\omega\in\mathbb{R}}$ and a set of positive values $\{\lambda_\omega\}_{\omega\in\mathbb{R}}$ such that $(T3.90)$ holds, which is the equivalent of $(A.53)$ in the continuous setting of functional analysis.

Furthermore, the set of eigenfunctions satisfies the equivalent of $(A.54)$ and $(A.56)$, i.e.

$$\left\langle e^{(\omega)}, e^{(\psi)}\right\rangle \equiv \int_{\mathbb{R}} e^{(\omega)}(x)\, \overline{e^{(\psi)}(x)} dx = 2\pi\delta^{(\omega)}(\psi), \qquad (T3.98)$$

where we chose a slightly more convenient normalization constant.

Consider the operator E represented by the following kernel:

$$E(y,\omega) \equiv e^{(\omega)}(y). \qquad (T3.99)$$

This is the equivalent of $(A.62)$, i.e. it is a (rescaled) unitary operator, the same way as $(A.62)$ is a rotation. Indeed:

$$\begin{aligned}
\|Eg\|^2 &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{(\omega)}(x)\, g(\omega)\, d\omega\right) \overline{\left(\int_{\mathbb{R}} e^{(\psi)}(x)\, g(\psi)\, d\psi\right)} dx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{(\omega)}(x)\, \overline{e^{(\psi)}(x)} dx\right) g(\omega)\, \overline{g(\psi)} d\psi d\omega \qquad (T3.100) \\
&= 2\pi \int_{\mathbb{R}}\int_{\mathbb{R}} \delta^{(\omega)}(\psi)\, g(\omega)\, \overline{g(\psi)} d\psi d\omega \\
&= 2\pi \int_{\mathbb{R}} g(\omega)\, \overline{g(\omega)} d\omega = 2\pi \|g\|^2
\end{aligned}$$

By means of the spectral theorem we can explicitly compute the eigenfunctions and the eigenvalues of a positive and symmetric Toeplitz operator.

First of all from $(T3.90)$, $(T3.95)$ and the fact that in the spectral theorem to each eigenvalue corresponds only one eigenvector, we obtain that the following relation must hold:

$$\frac{de^{(\omega)}(x)}{dx} = g_\omega e^{(\omega)}(x), \qquad (T3.101)$$

for some constant g_ω that might depend on ω. The general solution to this equation is

$$e^{(\omega)}(x) = A_\omega e^{g_\omega x}. \tag{T3.102}$$

To determine this constant, we compare the normalization condition ($T3.98$) with ($B.41$) obtaining:

$$e^{(\omega)}(x) = e^{i\omega x}. \tag{T3.103}$$

To compute the eigenvalues of S we substitute ($T3.103$) in ($T3.91$) and we re-write the spectral equation:

$$\lambda_\omega e^{i\omega x} = \int_{\mathbb{R}} S(x, x+z) e^{i\omega(x+z)} dz = e^{i\omega x} \int_{\mathbb{R}} S(x, x+z) e^{i\omega z} dz \tag{T3.104}$$

Now recall that S is Toeplitz and thus it is fully determined by its cross-diagonal section:

$$S(x, x+z) = S(0, z) \equiv h(z), \tag{T3.105}$$

where h is symmetric around the origin. Therefore we only need to evaluate ($T3.104$) at $x = 0$, which yields:

$$\lambda_\omega = \int_{\mathbb{R}} h(z) e^{i\omega z} dz \tag{T3.106}$$

In other words, the eigenvalues as a function of the frequency ω are the Fourier transform of the cross-diagonal section of the kernel representation ($T3.105$) of the operator:

$$\lambda_\omega = \mathcal{F}[h](\omega) \tag{T3.107}$$

In particular, if

$$h(z) \equiv \sigma^2 e^{-\gamma|z|} \tag{T3.108}$$

then

$$\begin{aligned}
\lambda_\omega &= \sigma^2 \int_{\mathbb{R}} e^{-\gamma|z|} \cos(\omega z)\, dz + i\sigma^2 \int_{\mathbb{R}} e^{-\gamma|z|} \sin(\omega z)\, dz \\
&= 2\sigma^2 \int_0^{+\infty} e^{-\gamma z} \cos(\omega z)\, dz + 0 \\
&= \frac{2\sigma^2 \gamma}{\gamma^2 + \omega^2}.
\end{aligned} \tag{T3.109}$$

3.7 Numerical Market Projection

Here we show how to perform the operations (3.65) by means of the fast Fourier transform in the standard case where analytical results are not available. The idea draws on Albanese, Jackson, and Wiberg (2003), the proof relies heavily on Xi Chen's contribution.

Approximating the probability density function

Consider a random variable X with pdf f_X. We approximate the pdf with a histogram of N bins:

$$f_X(x) \approx \sum_{n=1}^{N} f_n 1_{\Delta_n}(x), \qquad (T3.110)$$

The bins $\Delta_1, \ldots, \Delta_N$ are defined as follows. First of all, we define the bins' width:

$$h \equiv \frac{2a}{N}, \qquad (T3.111)$$

where a is a large enough real number and N is an even larger integer number. Now, consider a grid of equally spaced points:

$$\begin{aligned} \xi_1 &\equiv -a + h \\ &\vdots \\ \xi_n &\equiv -a + nh \\ &\vdots \\ \xi_{N-1} &\equiv a - h. \end{aligned} \qquad (T3.112)$$

Then for $n = 1, \ldots, N-1$ we define Δ_n as the interval of length h that surrounds symmetrically the point ξ_n:

$$\Delta_n \equiv \left(\xi_n - \frac{h}{2}, \xi_n + \frac{h}{2}\right]. \qquad (T3.113)$$

For $n = N$ we define the interval as follows:

$$\Delta_N \equiv \left(-a, -a + \frac{h}{2}\right] \cup \left(a - \frac{h}{2}, a\right]. \qquad (T3.114)$$

This wraps the real line around a circle where the point $-a$ coincides with the point a.

As far as the coefficients f_n in (T3.110) are concerned, for all $n = 1, \ldots, N$ they are defined as follows:

$$f_n \equiv \frac{1}{h} \int_{\Delta_n} f(x)\, dx. \qquad (T3.115)$$

We collect the discretized pdf values f_n into a vector $\mathbf{f}_X$.

Approximating the characteristic function

We need to compute the characteristic function:

$$\phi_X(\omega) \equiv \int_{\mathbb{R}} e^{i\omega x} f_X(x)\, dx. \qquad (T3.116)$$

Using (T3.110) and

$$\frac{1}{h} \int_{\mathbb{R}} g(x)\, 1_{\Delta_n}(x)\, dx \approx g(-a + nh), \qquad (T3.117)$$

we can approximate the characteristic function as follows:

$$\begin{aligned}\phi_X(\omega) &\approx \sum_{n=1}^{N} f_n \int_{\mathbb{R}} e^{i\omega x} 1_{\Delta_n}(x)\, dx \\ &\approx \sum_{n=1}^{N} f_n h e^{i\omega(-a+nh)} = \sum_{n=1}^{N} f_n h e^{-\frac{2\pi i}{N}\frac{\omega a}{\pi}\left(\frac{N}{2}-n\right)}.\end{aligned} \tag{T3.118}$$

In particular, we can evaluate the approximate characteristic function at the points:

$$\omega_r \equiv -(r-1)\frac{\pi}{a}, \tag{T3.119}$$

obtaining:

$$\begin{aligned}\phi_X(\omega_r) &\approx \sum_{n=1}^{N} f_n h e^{-\frac{2\pi i}{N}(r-1)\left(n-\frac{N}{2}\right)} \\ &= \sum_{n=1}^{N} f_n h e^{-\frac{2\pi i}{N}(r-1)n} e^{\pi i(r-1)} \\ &= e^{\pi i(r-1)} h e^{-\frac{2\pi i}{N}(r-1)} \sum_{n=1}^{N} f_n e^{-\frac{2\pi i}{N}(r-1)n} e^{\frac{2\pi i}{N}(r-1)} \\ &= e^{\pi i(r-1)\left(1-\frac{2}{N}\right)} h \sum_{n=1}^{N} f_n e^{-\frac{2\pi i}{N}(r-1)(n-1)}.\end{aligned} \tag{T3.120}$$

Finally, since N is supposed to be very large we can finally write:

$$\phi_X(\omega_r) \approx e^{\pi i(r-1)} h \sum_{n=1}^{N} f_n e^{-\frac{2\pi i}{N}(r-1)(n-1)}. \tag{T3.121}$$

The discrete Fourier transform

Consider now the discrete Fourier transform (DFT), an invertible matrix operation $\mathbf{f} \mapsto \mathbf{p}$ which is defined component-wise as follows:

$$p_r(\mathbf{f}) \equiv \sum_{n=1}^{N} f_n e^{-\frac{2\pi i}{N}(r-1)(n-1)}. \tag{T3.122}$$

Its inverse, the inverse discrete Fourier transform (IDFT), is the matrix operation $\mathbf{p} \mapsto \mathbf{f}$ which is defined component-wise as follows:

$$f_n(\mathbf{p}) \equiv \frac{1}{N} \sum_{r=1}^{N} p_r e^{\frac{2\pi i}{N}(r-1)(n-1)}. \tag{T3.123}$$

Comparing (T3.121) with (T3.122) we see that the approximate cf is a simple multiplicative function of the DFT of the discretized pdf $\mathbf{f}$.

$$\phi_X(\omega_r) \approx e^{\pi i(r-1)} h p_r(\mathbf{f}_X). \quad (T3.124)$$

Now consider the random variable:

$$Y \equiv X_1 + \cdots + X_T, \quad (T3.125)$$

where $X_1, \ldots, X_T$ are i.i.d. copies of X. The cf of Y satisfies the identity $\phi_Y \equiv \phi_X^T$, see (3.64). Therefore

$$\phi_Y(\omega_r) \approx e^{\pi i(r-1)T} h^T (p_r(\mathbf{f}_X))^T. \quad (T3.126)$$

On the other hand, from (T3.124), the relation between the cf ϕ_Y and the discrete pdf $\mathbf{f}_Y$ is:

$$\phi_Y(\omega_r) \approx e^{\pi i(r-1)} h p_r(\mathbf{f}_Y), \quad (T3.127)$$

Therefore

$$p_r(\mathbf{f}_Y) \approx e^{\pi i(r-1)(T-1)} h^{T-1} (p_r(\mathbf{f}_X))^T. \quad (T3.128)$$

The values $p_r(\mathbf{f}_Y)$ can now be fed into the IDFT (T3.123) to yield the discretized pdf $\mathbf{f}_Y$ of Y as defined in (T3.125).

4

Technical appendix to Chapter 4

4.1 Geometric interpretation of nonparametric estimators

From $(A.77)$ the volume of the ellipsoid $\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ is proportional to $\sqrt{|\boldsymbol{\Sigma}|}$. Therefore, defining $\boldsymbol{\Omega} \equiv \boldsymbol{\Sigma}^{-1}$, the optimization problem (4.48) becomes:

$$\left(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Omega}}\right) \equiv \underset{(\boldsymbol{\mu},\boldsymbol{\Omega})\in\mathcal{C}}{\operatorname{argmin}} |\boldsymbol{\Omega}| , \qquad (T4.1)$$

where the constraints read:

$$\mathcal{C}_1 : \frac{1}{T}\sum_{t=1}^{T} (\mathbf{x}_t - \boldsymbol{\mu})' \boldsymbol{\Omega} (\mathbf{x}_t - \boldsymbol{\mu}) = 1. \qquad (T4.2)$$

$$\mathcal{C}_2 : \boldsymbol{\Omega} \text{ symmetric, positive} \qquad (T4.3)$$

We solve neglecting $\mathcal{C}_2$ and we check later that $\mathcal{C}_2$ is satisfied. The Lagrangian reads:

$$\mathcal{L} \equiv |\boldsymbol{\Omega}| - \lambda \left[\frac{1}{T}\sum_{t=1}^{T} (\mathbf{x}_t - \boldsymbol{\mu})' \boldsymbol{\Omega} (\mathbf{x}_t - \boldsymbol{\mu}) - 1\right] \qquad (T4.4)$$

The first order condition with respect to $\boldsymbol{\mu}$ is

$$\mathbf{0}_{N\times 1} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\mu}} = \frac{2}{T}\sum_{t=1}^{T} \boldsymbol{\Omega} (\mathbf{x}_t - \boldsymbol{\mu}) \qquad (T4.5)$$

From which we see that the optimal $\widehat{\boldsymbol{\mu}}$ is the sample mean:

$$\widehat{\boldsymbol{\mu}} \equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}_t \equiv \widehat{\mathrm{E}}. \qquad (T4.6)$$

As for the first order condition with respect to $\boldsymbol{\Omega}$, we see from $(A.125)$ that if $\mathbf{A}$ is symmetric, then the following identity holds:

$$\frac{\partial \ln |\mathbf{A}|}{\partial \mathbf{A}} = \mathbf{A}^{-1}. \tag{$T4.7$}$$

Therefore

$$\mathbf{0}_{N\times N} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\Omega}} = \boldsymbol{\Omega}^{-1} - \frac{\lambda}{T}\sum_{t=1}^{T} (\mathbf{x}_t - \boldsymbol{\mu})(\mathbf{x}_t - \boldsymbol{\mu})', \tag{$T4.8$}$$

from which we see that the optimal $\widehat{\boldsymbol{\Omega}}$ satisfies

$$\widehat{\boldsymbol{\Omega}} = \left(\frac{\lambda}{T}\sum_{t=1}^{T} (\mathbf{x}_t - \boldsymbol{\mu})(\mathbf{x}_t - \boldsymbol{\mu})'\right)^{-1} \equiv \frac{1}{\lambda}\widehat{\mathrm{Cov}}^{-1} \tag{$T4.9$}$$

To compute the Lagrange multiplier λ we re-write the constraint ($T4.2$) as follows:

$$1 = \mathrm{tr}\left[\widehat{\boldsymbol{\Omega}}\widehat{\mathrm{Cov}}\right] = \mathrm{tr}\left[\frac{1}{\lambda}\mathbf{I}_N\right] = \frac{N}{\lambda}, \tag{$T4.10$}$$

from which $\lambda = N$.

To prove (4.53) we simply write the first order conditions, which read:

$$\mathbf{0}_{N\times K} = \sum_{t=1}^{T} -\left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right)\mathbf{f}_t'. \tag{$T4.11$}$$

The solution to this set of equations are the OLS factor loadings (4.52)

4.2 MLE estimators for elliptic variables

Location and dispersion

First of all we need two general results. Define

$$M_t^2 \equiv (\mathbf{x}_t - \boldsymbol{\mu})' \boldsymbol{\Omega} (\mathbf{x}_t - \boldsymbol{\mu}), \tag{$T4.12$}$$

where $\boldsymbol{\Omega}$ is a positive symmetric matrix and $\boldsymbol{\mu}$ any vector. It is easy to check the following:

$$\frac{\partial M_t^2}{\partial \boldsymbol{\mu}} = -2\boldsymbol{\Omega}(\mathbf{x}_t - \boldsymbol{\mu}) \tag{$T4.13$}$$

$$\frac{\partial M_t^2}{\partial \boldsymbol{\Omega}} = (\mathbf{x}_t - \boldsymbol{\mu})(\mathbf{x}_t - \boldsymbol{\mu})' \tag{$T4.14$}$$

Now assume that the distribution of the invariants is elliptical:

$$\mathbf{X} \sim \mathrm{El}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g), \tag{$T4.15$}$$

To compute the MLE estimators $\widehat{\boldsymbol{\mu}}[i_T]$ and $\widehat{\boldsymbol{\Sigma}}[i_T]$ we have to maximize the likelihood function (4.66) over the following parameter set

$$\mathbf{\Theta} \equiv \mathbb{R}^N \times \{\text{symmetric, positive, } N \times N \text{ matrices}\} . \qquad (T4.16)$$

First of all it is equivalent, though easier, to maximize the logarithm of the likelihood function. Secondly we neglect the constraint that $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ lie in $\mathbf{\Theta}$ and verify ex-post that the unconstrained solution belongs to $\mathbf{\Theta}$. Third, it is easier to compute the ML estimators of $\boldsymbol{\mu}$ and $\mathbf{\Omega} \equiv \mathbf{\Sigma}^{-1}$. The MLE estimator of $\mathbf{\Sigma}$ is simply the inverse of the estimator of $\mathbf{\Omega}$ by the invariance property (4.70) of the ML estimators.

From (4.74) the log-likelihood reads:

$$\begin{aligned} \ln\left(f_{\boldsymbol{\theta}}\left(i_T\right)\right) &= \sum_{t=1}^{T} \ln f_{\boldsymbol{\theta}}\left(\mathbf{x}_t\right) \\ &= \frac{T}{2} \ln\left|\det\left(\mathbf{\Omega}\right)\right| + \sum_{t=1}^{T} \ln\left[g\left(M_t^2\right)\right] . \end{aligned} \qquad (T4.17)$$

The first order conditions with respect to $\boldsymbol{\mu}$ read:

$$\begin{aligned} \mathbf{0}_{N\times 1} &= \frac{\partial}{\partial \boldsymbol{\mu}}\left[\ln\left(f_{\boldsymbol{\theta}}\left(i_T\right)\right)\right] \\ &= \frac{\partial}{\partial \boldsymbol{\mu}}\left[\sum_{t=1}^{T} \ln f_{\boldsymbol{\theta}}\left(\mathbf{x}_t\right)\right] \\ &= \frac{\partial}{\partial \boldsymbol{\mu}}\left[\sum_{t=1}^{T} \ln\left[g\left(M_t^2\right)\right]\right] \\ &= \sum_{t=1}^{T} \frac{g'\left(M_t^2\right)}{g\left(M_t^2\right)} \frac{\partial M_t^2}{\partial \boldsymbol{\mu}} = \sum_{t=1}^{T} w_t \mathbf{\Omega}\left(\mathbf{x}_t - \boldsymbol{\mu}\right) , \end{aligned} \qquad (T4.18)$$

where we used (T4.13) and we defined:

$$w_t \equiv -2\frac{g'\left(M_t^2\right)}{g\left(M_t^2\right)}. \qquad (T4.19)$$

The solution to this equations is

$$\widehat{\boldsymbol{\mu}} = \frac{\sum_{t=1}^{T} w_t \mathbf{x}_t}{\sum_{s=1}^{T} w_s}. \qquad (T4.20)$$

The first order conditions with respect to $\mathbf{\Omega}$ reads

$$\begin{aligned} \mathbf{0}_{N\times N} &= \frac{\partial \ln\left(f_{\boldsymbol{\theta}}\left(i_T\right)\right)}{\partial \mathbf{\Omega}} = \frac{\partial \sum_{t=1}^{T} \ln f_{\boldsymbol{\theta}}\left(\mathbf{x}_t\right)}{\partial \mathbf{\Omega}} \\ &= \frac{T}{2} \frac{\partial \ln\left|\det\left(\mathbf{\Omega}\right)\right|}{\partial \mathbf{\Omega}} + \sum_{t=1}^{T} \frac{g'\left(M_t^2\right)}{g\left(M_t^2\right)} \frac{\partial M_t^2}{\partial \mathbf{\Omega}} \\ &= \frac{T}{2}\mathbf{\Omega}^{-1} - \frac{1}{2}\sum_{t=1}^{T} w_t \left(\mathbf{x}_t - \boldsymbol{\mu}\right)\left(\mathbf{x}_t - \boldsymbol{\mu}\right)' , \end{aligned} \qquad (T4.21)$$

where in the last row we used $(T4.14)$ and the fact that from $(A.125)$ for a symmetric matrix $\boldsymbol{\Omega}$ we have:

$$\frac{\partial \ln|\boldsymbol{\Omega}|}{\partial \boldsymbol{\Omega}} = \boldsymbol{\Omega}^{-1}. \qquad (T4.22)$$

Thus the solution to $(T4.21)$ reads:

$$\widehat{\boldsymbol{\Sigma}} \equiv \widehat{\boldsymbol{\Omega}}^{-1} = \frac{1}{T}\sum_{t=1}^{T} w_t \left(\mathbf{x}_t - \boldsymbol{\mu}\right)\left(\mathbf{x}_t - \boldsymbol{\mu}\right)' \qquad (T4.23)$$

This matrix is symmetric and positive definite, and thus the unconstrained optimization is correct.

Explicit factors

Consider the following linear model with explicit-factors for the invariants:

$$\mathbf{X} \equiv \mathbf{B}\mathbf{f} + \mathbf{U}. \qquad (T4.24)$$

Assume that the conditional distribution of the perturbations is elliptical:

$$\mathbf{U}_t|\mathbf{f}_t \sim \mathrm{El}\left(\mathbf{0}, \boldsymbol{\Sigma}, g\right). \qquad (T4.25)$$

From the property (2.270) of elliptical distribution this implies that the conditional distribution of the invariants is elliptical with the same density generator

$$\mathbf{X}_t|\mathbf{f}_t \sim \mathrm{El}\left(\mathbf{B}\mathbf{f}_t, \boldsymbol{\Sigma}, g\right). \qquad (T4.26)$$

To compute the MLE estimators $\widehat{\mathbf{B}}\left[i_T\right]$ and $\widehat{\boldsymbol{\Sigma}}\left[i_T\right]$, we proceed as for the location-dispersion parameters. We define $\boldsymbol{\Omega} \equiv \boldsymbol{\Sigma}^{-1}$ and we maximize the log-likelihood function:

$$\ln\left(f_{\boldsymbol{\theta}}\left(i_T\right)\right) \equiv \frac{T}{2}\ln\left|\det\left(\boldsymbol{\Omega}\right)\right| + \sum_{t=1}^{T}\ln\left[g\left(M_t^2\right)\right], \qquad (T4.27)$$

where

$$M_t^2 \equiv \left(\mathbf{x}_t - \mathbf{B}\mathbf{f}_t\right)' \boldsymbol{\Omega}\left(\mathbf{x}_t - \mathbf{B}\mathbf{f}_t\right). \qquad (T4.28)$$

The first order conditions with respect to $\mathbf{B}$ read

$$\begin{aligned} \mathbf{0}_{N\times K} &= \frac{\partial}{\partial \mathbf{B}}\left[\ln\left(f_{\boldsymbol{\theta}}\left(i_T\right)\right)\right] \\ &= \sum_{t=1}^{T}\frac{g'\left(M_t^2\right)}{g\left(M_t^2\right)}\frac{\partial M_t^2}{\partial \mathbf{B}} \\ &= \sum_{t=1}^{T} w_t \boldsymbol{\Omega}\left(\mathbf{x}_t - \mathbf{B}\mathbf{f}_t\right)\mathbf{f}_t', \end{aligned} \qquad (T4.29)$$

where

$$w_t \equiv -2\frac{g'\left(M_t^2\right)}{g\left(M_t^2\right)}. \qquad (T4.30)$$

The solution to (T4.29) reads:

$$\mathbf{B} = \left[\sum_{t=1}^{T} w_t \mathbf{x}_t \mathbf{f}_t'\right]\left[\sum_{t=1}^{T} w_t \mathbf{f}_t \mathbf{f}_t'\right]^{-1} \qquad (T4.31)$$

The first order conditions with respect to $\mathbf{\Omega}$ follow like (T4.21) and yield (4.93).

4.3 MLE estimators of location-dispersion (normal case)

Independence of sample mean and sample covariance

Assume:

$$\mathbf{X}_t \sim \mathrm{N}\left(\boldsymbol{\mu}, \mathbf{\Sigma}\right). \qquad (T4.32)$$

Consider the following variables:

$$\begin{aligned} \widehat{\boldsymbol{\mu}} &\equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{X}_t \\ \mathbf{U}_1 &\equiv \mathbf{X}_1 - \widehat{\boldsymbol{\mu}} \\ &\vdots \\ \mathbf{U}_T &\equiv \mathbf{X}_T - \widehat{\boldsymbol{\mu}} \end{aligned} \qquad (T4.33)$$

The joint characteristic function of $\{\mathbf{U}_1, \ldots, \mathbf{U}_T, \widehat{\boldsymbol{\mu}}\}$ reads:

$$\begin{aligned} \phi &\equiv \phi_{\mathbf{U}_1,\ldots,\mathbf{U}_T,\widehat{\boldsymbol{\mu}}}\left(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_T, \boldsymbol{\tau}\right) \\ &= \mathrm{E}\left\{e^{i\left(\sum_{t=1}^{T} \boldsymbol{\omega}_t' \mathbf{U}_t + \boldsymbol{\tau}'\widehat{\boldsymbol{\mu}}\right)}\right\} \\ &= \mathrm{E}\left\{e^{i\left[\sum_{t=1}^{T} \boldsymbol{\omega}_t'\left(\mathbf{X}_t - \frac{1}{T}\sum_{s=1}^{T}\mathbf{X}_s\right) + \boldsymbol{\tau}'\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{X}_t\right)\right]}\right\} \\ &= \mathrm{E}\left\{e^{i\left(\sum_{t=1}^{T}\left(\boldsymbol{\omega}_t + \frac{\boldsymbol{\tau}}{T} - \frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s\right)'\mathbf{X}_t\right)}\right\} \end{aligned} \qquad (T4.34)$$

From the independence of the invariants we can factor the characteristic function as follows:

$$\begin{aligned} \phi &= \prod_{t=1}^{T} \mathrm{E}\left\{e^{i\left(\boldsymbol{\omega}_t + \frac{\boldsymbol{\tau}}{T} - \frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s\right)'\mathbf{X}_t}\right\} \qquad (T4.35) \\ &= \prod_{t=1}^{T} \phi_{\mathbf{X}_t}\left(\boldsymbol{\omega}_t - \frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s + \frac{\boldsymbol{\tau}}{T}\right). \end{aligned}$$

Since $\mathbf{X}_t$ is normal ($T4.32$), from (2.157) we have

$$\phi_{\mathbf{X}_t}(\boldsymbol{\omega}) = e^{i\boldsymbol{\mu}'\boldsymbol{\omega}-\frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega}}. \qquad (T4.36)$$

Therefore:

$$\begin{aligned}\phi &= e^{i\sum_{t=1}^{T}\boldsymbol{\mu}'\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s+\frac{\boldsymbol{\tau}}{T}\right)} \\ &\quad e^{\sum_{t=1}^{T}-\frac{1}{2}\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s+\frac{\boldsymbol{\tau}}{T}\right)'\boldsymbol{\Sigma}\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s+\frac{\boldsymbol{\tau}}{T}\right)}\end{aligned} \qquad (T4.37)$$

In the last expression a few terms simplify:

$$\begin{aligned}\sum_{t=1}^{T}\boldsymbol{\mu}'\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s+\frac{\boldsymbol{\tau}}{T}\right) &= \boldsymbol{\mu}'\boldsymbol{\tau} \\ \sum_{t=1}^{T}\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s\right)'\boldsymbol{\Sigma}\frac{\boldsymbol{\tau}}{T} &= 0.\end{aligned} \qquad (T4.38)$$

Therefore the joint characteristic function factors into the following product:

$$\phi_{\mathbf{U}_1,\ldots,\mathbf{U}_T,\widehat{\boldsymbol{\mu}}}(\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_T,\boldsymbol{\tau}) = \psi(\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_T)\,\chi(\boldsymbol{\tau}), \qquad (T4.39)$$

where

$$\begin{aligned}\psi(\boldsymbol{\omega}_1,\ldots,\boldsymbol{\omega}_T) &\equiv e^{\sum-\frac{1}{2}\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s\right)'\boldsymbol{\Sigma}\left(\boldsymbol{\omega}_t-\frac{1}{T}\sum_{s=1}^{T}\boldsymbol{\omega}_s\right)} \\ \chi(\boldsymbol{\tau}) &\equiv e^{i\boldsymbol{\mu}'\boldsymbol{\tau}-\frac{1}{2T}\boldsymbol{\tau}'\boldsymbol{\Sigma}\boldsymbol{\tau}}.\end{aligned} \qquad (T4.40)$$

This proves the variables $\{\mathbf{U}_1,\ldots,\mathbf{U}_T\}$ and $\widehat{\boldsymbol{\mu}}$ are independent. In particular the sample covariance matrix

$$\widehat{\boldsymbol{\Sigma}} \equiv \frac{1}{T}\sum_{t=1}^{T}\mathbf{U}_t\mathbf{U}_t' \qquad (T4.41)$$

is independent of $\widehat{\boldsymbol{\mu}}$.

Distribution and estimation error of the sample mean

From ($T4.32$) we obtain:

$$\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu} \sim \mathrm{N}\left(\mathbf{0},\frac{\boldsymbol{\Sigma}}{T}\right), \qquad (T4.42)$$

and thus from (2.222) and (2.223):

$$(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})' \sim \mathrm{W}\left(1,\frac{\boldsymbol{\Sigma}}{T}\right). \qquad (T4.43)$$

Therefore from (2.227) the estimation error reads:

$$\begin{aligned}\mathrm{Err}\{\widehat{\boldsymbol{\mu}},\boldsymbol{\mu}\} &\equiv \mathrm{E}\left\{(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})'(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})\right\} \\ &= \mathrm{tr}\left(\mathrm{E}\left\{(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})'\right\}\right) \\ &= \frac{1}{T}\mathrm{tr}(\boldsymbol{\Sigma})\end{aligned} \qquad (T4.44)$$

Distribution and estimation error of the sample covariance

First of all we notice that the estimator of the covariance of $\mathbf{X}$ is the same as the estimator of the covariance of $\mathbf{Y} \equiv \mathbf{X} + \mathbf{b}$ for any $\mathbf{b}$. Indeed defining

$$\widehat{\boldsymbol{\nu}} \equiv \frac{1}{T} \sum_{t=1}^{T} \mathbf{Y}_t \tag{$T4.45$}$$

we easily verify that

$$\widehat{\boldsymbol{\mu}} = \widehat{\boldsymbol{\nu}} + \mathbf{b} \tag{$T4.46$}$$

and thus:

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_{\mathbf{Y}} &\equiv \frac{1}{T} \sum_{t=1}^{T} \left(\mathbf{Y}_t - \widehat{\boldsymbol{\nu}}\right) \left(\mathbf{Y}_t - \widehat{\boldsymbol{\nu}}\right) \\ &= \frac{1}{T} \sum_{t=1}^{T} \left(\mathbf{X}_t + \mathbf{b} - (\widehat{\boldsymbol{\mu}} + \mathbf{b})\right) \left(\mathbf{X}_t + \mathbf{b} - (\widehat{\boldsymbol{\mu}} + \mathbf{b})\right)' \\ &= \frac{1}{T} \sum_{t=1}^{T} \left(\mathbf{X}_t - \widehat{\boldsymbol{\mu}}\right) \left(\mathbf{X}_t - \widehat{\boldsymbol{\mu}}\right)' \\ &\equiv \widehat{\boldsymbol{\Sigma}}_{\mathbf{X}} \end{aligned} \tag{$T4.47$}$$

Therefore we can assume here that $\boldsymbol{\mu} \equiv \mathbf{0}$ in ($T4.32$).

Consider

$$\begin{aligned} \mathbf{W} &\equiv T\widehat{\boldsymbol{\Sigma}} + T\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}' \\ &= \sum_{t=1}^{T} \mathbf{X}_t \mathbf{X}_t', \end{aligned} \tag{$T4.48$}$$

where the last equality follows from substitution of the definitions ($T4.33$) and ($T4.41$) in ($T4.48$). From the above proved independence of $\widehat{\boldsymbol{\Sigma}}$ and $\widehat{\boldsymbol{\mu}}$ the characteristic function of $\mathbf{W}$ must be the product of the characteristic function of $T\widehat{\boldsymbol{\Sigma}}$ and the characteristic function of $T\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'$. Therefore

$$\phi_{T\widehat{\boldsymbol{\Sigma}}}\left(\boldsymbol{\Omega}\right) = \frac{\phi_{\mathbf{W}}\left(\boldsymbol{\Omega}\right)}{\phi_{T\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'}\left(\boldsymbol{\Omega}\right)} \tag{$T4.49$}$$

On the one hand from ($T4.32$) and (2.223) we obtain that $\mathbf{W}$ is Wishart distributed with the following parameters:

$$\mathbf{W} \sim \mathrm{W}\left(T, \boldsymbol{\Sigma}\right), \tag{$T4.50$}$$

and thus from (2.226) its characteristic function reads:

$$\phi_{\mathbf{W}}\left(\boldsymbol{\Omega}\right) = \frac{1}{\left|\mathbf{I} - 2i\boldsymbol{\Sigma}\boldsymbol{\Omega}\right|^{T/2}} \tag{$T4.51$}$$

On the other hand, the characteristic function of $T\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'$ reads:

$$\begin{aligned}\phi_{T\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'}(\boldsymbol{\Omega}) &\equiv \mathrm{E}\left\{e^{i\,\mathrm{tr}\left(\left[T\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'\right]\boldsymbol{\Omega}\right)}\right\} \\ &= \mathrm{E}\left\{e^{i\,\mathrm{tr}\left(\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'\left[T\boldsymbol{\Omega}\right]\right)}\right\} \\ &= \phi_{\widehat{\boldsymbol{\mu}}\widehat{\boldsymbol{\mu}}'}(T\boldsymbol{\Omega}) = \frac{1}{\left|\mathbf{I}-2i\boldsymbol{\Sigma}\boldsymbol{\Omega}\right|^{1/2}}\end{aligned} \qquad (T4.52)$$

Substituting ($T4.51$) and ($T4.52$) in ($T4.49$) we obtain

$$\phi_{T\widehat{\boldsymbol{\Sigma}}}(\boldsymbol{\Omega}) = \frac{1}{\left|\mathbf{I}-2i\boldsymbol{\Sigma}\boldsymbol{\Omega}\right|^{(T-1)/2}} \qquad (T4.53)$$

which from (2.226) shows that

$$T\widehat{\boldsymbol{\Sigma}} \sim \mathrm{W}(T-1,\boldsymbol{\Sigma}). \qquad (T4.54)$$

In particular:

$$\begin{aligned}\mathrm{E}\left\{\widehat{\boldsymbol{\Sigma}}[I_T]\right\}-\boldsymbol{\Sigma} &= \frac{1}{T}\mathrm{E}\left\{T\widehat{\boldsymbol{\Sigma}}[I_T]\right\}-\boldsymbol{\Sigma} \\ &= \frac{T-1}{T}\boldsymbol{\Sigma}-\boldsymbol{\Sigma} \\ &= -\frac{1}{T}\boldsymbol{\Sigma}\end{aligned} \qquad (T4.55)$$

Therefore the bias reads:

$$\begin{aligned}\mathrm{Bias}^2\left(\widehat{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}\right) &\equiv \mathrm{tr}\left\{\left(\mathrm{E}\left\{\widehat{\boldsymbol{\Sigma}}[I_T]\right\}-\boldsymbol{\Sigma}\right)^2\right\} \\ &= \frac{1}{T^2}\mathrm{tr}\left\{\boldsymbol{\Sigma}^2\right\}\end{aligned} \qquad (T4.56)$$

As for the error, from its definition we obtain:

$$\begin{aligned}\mathrm{Err}^2_{\boldsymbol{\mu},\Sigma}\left(\widehat{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}\right) &\equiv \mathrm{E}\left\{\mathrm{tr}\left[\left(\widehat{\boldsymbol{\Sigma}}[I_T]-\boldsymbol{\Sigma}\right)^2\right]\right\} \\ &\mathrm{E}\left\{\sum_{m,n}\left[\boldsymbol{\Sigma}[I_T]-\boldsymbol{\Sigma}\right]_{mn}\left[\widehat{\boldsymbol{\Sigma}}[I_T]-\boldsymbol{\Sigma}\right]_{nm}\right\} \\ &\sum_{m,n}\mathrm{E}\left\{\left[\widehat{\boldsymbol{\Sigma}}[I_T]-\boldsymbol{\Sigma}\right]^2_{mn}\right\} \\ &= \sum_{m,n}\mathrm{E}\left\{\left(\widehat{\Sigma}_{mn}-\boldsymbol{\Sigma}_{mn}\right)^2\right\} = \sum_{m,n}\mathrm{Err}^2_{\boldsymbol{\mu},\Sigma}\left(\widehat{\Sigma}_{mn},\Sigma_{mn}\right) \\ &= \sum_{m,n}\left[\mathrm{Bias}^2_{\boldsymbol{\mu},\Sigma}\left(\widehat{\Sigma}_{mn},\Sigma_{mn}\right)+\mathrm{Inef}^2_{\boldsymbol{\mu},\Sigma}\left(\widehat{\Sigma}_{mn}\right)\right]\end{aligned} \qquad (T4.57)$$

using (4.106) and (4.107) this becomes

$$\begin{aligned}\mathrm{Err}^2_{\boldsymbol{\mu},\Sigma}\left(\widehat{\boldsymbol{\Sigma}},\boldsymbol{\Sigma}\right) &= \sum_{m,n}\left(\frac{1}{T^2}\Sigma^2_{mn}+\frac{T-1}{T^2}\Sigma_{mm}\Sigma_{nn}+\frac{T-1}{T^2}\Sigma^2_{mn}\right)\\ &= \frac{1}{T}\sum_{m,n}\Sigma^2_{mn}+\frac{T-1}{T^2}\sum_{m,n}\Sigma_{mm}\Sigma_{nn} \qquad (T4.58)\\ &= \frac{1}{T}\left(\mathrm{tr}\left(\boldsymbol{\Sigma}^2\right)+\left(1-\frac{1}{T}\right)\left[\mathrm{tr}\left(\boldsymbol{\Sigma}\right)\right]^2\right)\end{aligned}$$

4.4 MLE estimators of factor loadings (normal case)

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

We make the i.i.d. normal hypothesis:

$$\mathbf{X}_t|\mathbf{f}_t \sim \mathrm{N}\left(\mathbf{B}\mathbf{f}_t, \boldsymbol{\Sigma}\right). \qquad (T4.59)$$

Notice that in the spirit of explicit factors models, the dependent variables $\mathbf{X}_t$ are random variables, whereas the factors $\mathbf{f}_t$ are considered observed numbers. In other words, we derive all the distributions *conditioned* on knowledge of the factors.

We derive here the joint distribution of the sample factor loadings

$$\widehat{\mathbf{B}} \equiv \widehat{\boldsymbol{\Sigma}}_{XF}\widehat{\boldsymbol{\Sigma}}_F^{-1}, \qquad (T4.60)$$

where

$$\widehat{\boldsymbol{\Sigma}}_{XF} \equiv \frac{1}{T}\sum_{t=1}^T \mathbf{X}_t\mathbf{f}_t', \quad \widehat{\boldsymbol{\Sigma}}_F \equiv \frac{1}{T}\sum_{t=1}^T \mathbf{f}_t\mathbf{f}_t'; \qquad (T4.61)$$

and the sample covariance

$$\widehat{\boldsymbol{\Sigma}} \equiv \frac{1}{T}\sum_{t=1}^T\left(\mathbf{X}_t-\widehat{\mathbf{B}}\mathbf{f}_t\right)\left(\mathbf{X}_t-\widehat{\mathbf{B}}\mathbf{f}_t\right)'. \qquad (T4.62)$$

Notice that the invariants $\mathbf{X}$ are random variables, whereas the factors $\mathbf{f}$ are not.

From the normal hypothesis ($T4.59$) the joint pdf of the time series

$$I_T \equiv \{\mathbf{X}_1,\ldots,\mathbf{X}_T|\mathbf{f}_1,\ldots,\mathbf{f}_T\} \qquad (T4.63)$$

in terms of the factor loadings $\mathbf{B}$ and the dispersion parameter $\boldsymbol{\Omega} \equiv \boldsymbol{\Sigma}^{-1}$ reads:

$$\begin{aligned}f\left(i_T\right) &= \gamma_1\left|\boldsymbol{\Omega}\right|^{\frac{T}{2}} e^{-\frac{1}{2}\sum_{t=1}^T(\mathbf{x}_t-\mathbf{B}\mathbf{f}_t)'\boldsymbol{\Omega}(\mathbf{x}_t-\mathbf{B}\mathbf{f}_t)} \qquad (T4.64)\\ &= \gamma_1\left|\boldsymbol{\Omega}\right|^{\frac{T}{2}} e^{-\frac{1}{2}\mathrm{tr}\left\{\boldsymbol{\Omega}\sum_{t=1}^T(\mathbf{x}_t-\mathbf{B}\mathbf{f}_t)(\mathbf{x}_t-\mathbf{B}\mathbf{f}_t)'\right\}}.\end{aligned}$$

The term in curly brackets can be written as

$$\{\cdots\} = \mathbf{\Omega A}, \tag{T4.65}$$

where:

$$\begin{aligned}
\mathbf{A} \equiv & \sum_{t=1}^{T} (\mathbf{x}_t - \mathbf{B}\mathbf{f}_t)(\mathbf{x}_t - \mathbf{B}\mathbf{f}_t)' \\
= & \sum_{t=1}^{T} \left[\left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right) + \left(\widehat{\mathbf{B}}\mathbf{f}_t - \mathbf{B}\mathbf{f}_t\right)\right]\left[\left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right) + \left(\widehat{\mathbf{B}}\mathbf{f}_t - \mathbf{B}\mathbf{f}_t\right)\right]' \\
= & \sum_{t=1}^{T} \left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right)\left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right)' + \left(\widehat{\mathbf{B}} - \mathbf{B}\right)\left(\sum_{t=1}^{T} \mathbf{f}_t \mathbf{f}_t\right)\left(\widehat{\mathbf{B}} - \mathbf{B}\right)' \\
& + \sum_{t=1}^{T} \left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right)\left(\widehat{\mathbf{B}}\mathbf{f}_t - \mathbf{B}\mathbf{f}_t\right)' \\
& T\widehat{\mathbf{\Sigma}} + \left(\widehat{\mathbf{B}} - \mathbf{B}\right) T\widehat{\mathbf{\Sigma}}_F \left(\widehat{\mathbf{B}} - \mathbf{B}\right)' + \mathbf{0}.
\end{aligned} \tag{T4.66}$$

In this expression the last term vanishes, since:

$$\begin{aligned}
\sum_{t=1}^{T} \left(\mathbf{x}_t - \widehat{\mathbf{B}}\mathbf{f}_t\right)\left(\widehat{\mathbf{B}}\mathbf{f}_t - \mathbf{B}\mathbf{f}_t\right)' = & \sum_{t=1}^{T} \mathbf{x}_t \mathbf{f}_t' \widehat{\mathbf{B}}' + \sum_{t=1}^{T} \widehat{\mathbf{B}}\mathbf{f}_t \mathbf{f}_t' \mathbf{B}' \\
& - \sum_{t=1}^{T} \mathbf{x}_t \mathbf{f}_t' \mathbf{B}' - \sum_{t=1}^{T} \widehat{\mathbf{B}}\mathbf{f}_t \mathbf{f}_t' \widehat{\mathbf{B}}' \\
= & \, T\widehat{\mathbf{\Sigma}}_{XF}\widehat{\mathbf{B}}' + \widehat{\mathbf{B}} T\widehat{\mathbf{\Sigma}}_F \mathbf{B}' - T\widehat{\mathbf{\Sigma}}_{XF}\mathbf{B}' - T\widehat{\mathbf{B}}\widehat{\mathbf{\Sigma}}_F \widehat{\mathbf{B}}' \\
= & \, T\widehat{\mathbf{\Sigma}}_{XF}\widehat{\mathbf{\Sigma}}_F^{-1}\widehat{\mathbf{\Sigma}}_{XF}' + T\widehat{\mathbf{\Sigma}}_{XF}\mathbf{B}' \\
& - T\widehat{\mathbf{\Sigma}}_{XF}\mathbf{B}' - T\widehat{\mathbf{\Sigma}}_{XF}\widehat{\mathbf{\Sigma}}_F^{-1}\widehat{\mathbf{\Sigma}}_{XF}' \\
= & \, \mathbf{0}
\end{aligned} \tag{T4.67}$$

Substituting ($T4.66$) in the curly brackets ($T4.65$) in ($T4.64$) we obtain

$$f(i_T) = \gamma_1 |\mathbf{\Omega}|^{\frac{T}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{T\mathbf{\Omega}\left[\widehat{\mathbf{\Sigma}} + (\widehat{\mathbf{B}} - \mathbf{B})\widehat{\mathbf{\Sigma}}_F (\widehat{\mathbf{B}} - \mathbf{B})'\right]\right\}} \tag{T4.68}$$

We can factor the above expression as follows:

$$f(i_T) = f\left(i_T | \widehat{\mathbf{B}}, T\widehat{\mathbf{\Sigma}}\right) f\left(\widehat{\mathbf{B}}, T\widehat{\mathbf{\Sigma}}\right) \tag{T4.69}$$

where

$$f\left(i_T | \widehat{\mathbf{B}}, T\widehat{\mathbf{\Sigma}}\right) \equiv \gamma_2 \left|\widehat{\mathbf{\Sigma}}\right|^{-\frac{T-K-N-1}{2}} \tag{T4.70}$$

and $f\left(\widehat{\mathbf{B}}, T\widehat{\mathbf{\Sigma}}\right)$ factors as follows:

$$f\left(\widehat{\mathbf{B}}, T\widehat{\mathbf{\Sigma}}\right) = f\left(\widehat{\mathbf{B}}\right) f\left(T\widehat{\mathbf{\Sigma}}\right), \quad (T4.71)$$

where

$$f\left(\widehat{\mathbf{B}}\right) \equiv \gamma_2 \left|T\mathbf{\Omega}\right|^{\frac{K}{2}} \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{(T\mathbf{\Omega})\left(\widehat{\mathbf{B}}-\mathbf{B}\right)\widehat{\mathbf{\Sigma}}_F\left(\widehat{\mathbf{B}}-\mathbf{B}\right)'\right\}} \quad (T4.72)$$

and

$$f\left(T\widehat{\mathbf{\Sigma}}\right) \equiv \gamma_3 \left|\mathbf{\Omega}\right|^{\frac{T-K}{2}} \left|T\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}T\widehat{\mathbf{\Sigma}}\right)}. \quad (T4.73)$$

Expression ($T4.72$) is of the form (2.182). Therefore the OLS factor loadings are have a matrix-valued normal distribution:

$$\widehat{\mathbf{B}} \sim \mathrm{N}\left(\mathbf{B}, (T\mathbf{\Omega})^{-1}, \widehat{\mathbf{\Sigma}}_F^{-1}\right). \quad (T4.74)$$

Since $\mathbf{\Omega} \equiv \mathbf{\Sigma}^{-1}$ this means:

$$\widehat{\mathbf{B}} \sim \mathrm{N}\left(\mathbf{B}, \frac{\mathbf{\Sigma}}{T}, \widehat{\mathbf{\Sigma}}_F^{-1}\right). \quad (T4.75)$$

Also, expression ($T4.73$) is the pdf (2.224) of a Wishart distribution, and thus:

$$T\widehat{\mathbf{\Sigma}} \sim \mathrm{W}\left(T-K, \mathbf{\Sigma}\right). \quad (T4.76)$$

Finally, from the factorization ($T4.71$) we see that $T\widehat{\mathbf{\Sigma}}$, and thus $\widehat{\mathbf{\Sigma}}$, is independent of $\widehat{\mathbf{B}}$.

4.5 Shrinkage estimator of location

First we prove *Stein's lemma*. Consider an N-dimensional normal variable

$$\mathbf{X} \sim \mathrm{N}\left(\boldsymbol{\mu}, \mathbf{I}_N\right), \quad (T4.77)$$

where $\mathbf{I}_N$ is the N-dimensional identity matrix. Consider a smooth function of N variables g. Then

$$\mathrm{E}\left\{g\left(\mathbf{X}\right)\left(X_n - \mu_n\right)\right\} = \mathrm{E}\left\{\frac{\partial g\left(\mathbf{X}\right)}{\partial x_n}\right\} \quad (T4.78)$$

From the definition of expected value for a normal distribution we have:

$$\begin{aligned}\mathrm{E}\left\{g\left(\mathbf{X}\right)\left(X_n - \mu_n\right)\right\} &= \int_{\mathbb{R}^N} g\left(x_1, \ldots, x_n, \ldots, x_N\right)\left(x_n - \mu_n\right) \\ &\quad (2\pi)^{-\frac{N}{2}} e^{-\frac{1}{2}\sum_k (x_k - \mu_k)^2} d\mathbf{x} \quad (T4.79) \\ &= \int_{-\infty}^{+\infty} \left(x_n - \mu_n\right) G\left(x_n\right) \frac{e^{-\frac{(x_n-\mu_n)^2}{2}}}{\sqrt{2\pi}} dx_n,\end{aligned}$$

where we defined G as follows:

$$G(x) \equiv \int_{\mathbb{R}^{N-1}} g(x_1, \ldots, x, \ldots, x_N) \quad (T4.80)$$
$$(2\pi)^{-\frac{N-1}{2}} e^{-\frac{1}{2}\sum_{k\neq n}(x_k-\mu_k)^2} dx_1 \cdots dx_{n-1} dx_{n+1} \cdots dx_N.$$

Notice that

$$dG(x) \equiv \int_{\mathbb{R}^{N-1}} \frac{\partial g(\mathbf{x})}{\partial x_n} (2\pi)^{-\frac{N-1}{2}} e^{-\frac{1}{2}\sum_{k\neq n}(x_k-\mu_k)^2} d\mathbf{x}. \quad (T4.81)$$

Replacing the variables in the integral ($T4.79$) as follows:

$$u \equiv G(x_n) \quad (T4.82)$$
$$v \equiv -e^{-\frac{(x_n-\mu_n)^2}{2}}, \quad (T4.83)$$

we get

$$\mathrm{E}\{g(\mathbf{X})(X_n - \theta_n)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u dv \quad (T4.84)$$
$$= \frac{1}{\sqrt{2\pi}} \left(uv|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} v du \right)$$

The first term vanishes. Replacing ($T4.82$) and ($T4.83$) in the second term and using ($T4.81$) we obtain:

$$\mathrm{E}\{g(\mathbf{X})(X_n - \theta_n)\} = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} -e^{-\frac{(x_n-\mu_n)^2}{2}} \quad (T4.85)$$
$$\int_{\mathbb{R}^{N-1}} \frac{\partial g(\mathbf{x})}{\partial x_n} (2\pi)^{-\frac{N-1}{2}} e^{-\frac{1}{2}\sum_{k\neq n}(x_k-\mu_k)^2} d\mathbf{x}$$
$$= \int_{\mathbb{R}^{N}} \frac{\partial g(\mathbf{x})}{\partial x_n} (2\pi)^{-\frac{N}{2}} e^{-\frac{1}{2}\sum_{k}(x_k-\mu_k)^2} d\mathbf{x}$$
$$= \mathrm{E}\left\{ \frac{\partial g(\mathbf{X})}{\partial x_n} \right\}$$

and the result follows.

Consider now a set of T i.i.d. multivariate normal random variables

$$\mathbf{X}_t \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (T4.86)$$

where the dimension of each random variable is:

$$N > 2. \quad (T4.87)$$

Consider the following shrinkage estimator

$$\widehat{\boldsymbol{\delta}}_a \equiv \left(1 - \frac{a}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})}\right) \widehat{\boldsymbol{\mu}} + \frac{a}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})} \mathbf{b}, \qquad (T4.88)$$

where $\widehat{\boldsymbol{\mu}}$ is the sample mean

$$\widehat{\boldsymbol{\mu}} \equiv \frac{1}{T} \sum_t \mathbf{X}_t \sim \mathrm{N}\left(\boldsymbol{\mu}, \frac{\boldsymbol{\Sigma}}{T}\right); \qquad (T4.89)$$

where $\mathbf{b}$ is any constant vector; and where a is any scalar such that:

$$0 < a < \frac{2}{T} \left(\mathrm{tr}\,(\boldsymbol{\Sigma}) - 2\lambda_1\right), \qquad (T4.90)$$

where λ_1 is the largest eigenvalue of the matrix $\boldsymbol{\Sigma}$.

From the definition (4.134) of error, we have:

$$\begin{aligned} [\mathrm{Err}\,(\boldsymbol{\delta}, \boldsymbol{\mu})]^2 &= \mathrm{E}\left\{[\boldsymbol{\delta} - \boldsymbol{\mu}]' [\boldsymbol{\delta} - \boldsymbol{\mu}]\right\} \qquad (T4.91) \\ &= \mathrm{E}\left\{\left[\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu} - \frac{a(\widehat{\boldsymbol{\mu}} - \mathbf{b})}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})}\right]' \left[\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu} - \frac{a(\widehat{\boldsymbol{\mu}} - \mathbf{b})}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})}\right]\right\} \\ &= [\mathrm{Err}\,(\widehat{\boldsymbol{\mu}}, \boldsymbol{\mu})]^2 + a^2\, \mathrm{E}\left\{\frac{1}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})}\right\} \\ &\quad -2a\, \mathrm{E}\left\{\frac{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})}\right\} \end{aligned}$$

We proceed now to simplify the expression of the last expectation in ($T4.91$). Consider the principal component decomposition ($A.70$) of the matrix $\boldsymbol{\Sigma}$ in ($T4.89$):

$$\boldsymbol{\Sigma} \equiv \mathbf{E} \boldsymbol{\Lambda} \mathbf{E}' \qquad (T4.92)$$

and define the following vector of independent normal variables:

$$\mathbf{Y} \equiv \sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' \widehat{\boldsymbol{\mu}} \sim \mathrm{N}\,(\boldsymbol{\nu}, \mathbf{I}), \qquad (T4.93)$$

where

$$\boldsymbol{\nu} \equiv \sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' \boldsymbol{\mu}, \quad \mathbf{c} \equiv \sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' \mathbf{b} \qquad (T4.94)$$

Then the term in curly brackets in the last expectation in ($T4.91$) reads:

$$\begin{aligned} \frac{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})} &= \frac{\left[\sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' (\widehat{\boldsymbol{\mu}} - \mathbf{b})\right]' \boldsymbol{\Lambda} \left[\sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})\right]}{\left[\sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' (\widehat{\boldsymbol{\mu}} - \mathbf{b})\right]' \boldsymbol{\Lambda} \left[\sqrt{T} \boldsymbol{\Lambda}^{-\frac{1}{2}} \mathbf{E}' (\widehat{\boldsymbol{\mu}} - \mathbf{b})\right]} \\ &= \frac{(\mathbf{Y} - \mathbf{c})' \boldsymbol{\Lambda} (\mathbf{Y} - \boldsymbol{\nu})}{(\mathbf{Y} - \mathbf{c})' \boldsymbol{\Lambda} (\mathbf{Y} - \mathbf{c})} \qquad (T4.95) \\ &= \sum_{j=1}^{N} g_j(\mathbf{Y}) (Y_j - \nu_j), \end{aligned}$$

where

$$g_j(\mathbf{y}) \equiv \frac{(Y_j - c_j)\lambda_j}{(\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c})}. \qquad (T4.96)$$

Applying the rules of calculus we compute:

$$\begin{aligned}\frac{\partial g_j(\mathbf{y})}{\partial y_j} &= \frac{\lambda_j}{(\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c})} \qquad (T4.97)\\ &\quad + (y_j - c_j)\lambda_j \frac{d}{dy_j}\frac{1}{(\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c})}\\ &= \frac{\lambda_j}{(\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c})} - \frac{2\lambda_j^2 (y-c_j)^2}{\left[(\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c})\right]^2}\\ &= \frac{\lambda_j (\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c}) - 2\lambda_j^2 (y-c_j)^2}{\left[(\mathbf{y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{y}-\mathbf{c})\right]^2},\end{aligned}$$

Therefore, using Stein's lemma ($T4.78$), we obtain for the last expectation in ($T4.91$):

$$\begin{aligned}\mathrm{E}\left\{\frac{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})}{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\mathbf{b})}\right\} &= \sum_{j=1}^{N}\mathrm{E}\left\{g_j(\mathbf{Y})(Y_j-\nu_j)\right\}\\ &= \sum_{j=1}^{N}\mathrm{E}\left\{\frac{\partial g_j(\mathbf{Y})}{\partial y_j}\right\} \qquad (T4.98)\\ &= \sum_{j=1}^{N}\mathrm{E}\left\{\frac{\lambda_j(\mathbf{Y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{Y}-\mathbf{c}) - 2\lambda_j^2(Y_j-c_j)^2}{\left[(\mathbf{Y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{Y}-\mathbf{c})\right]^2}\right\}\\ &= \mathrm{E}\left\{\frac{\mathrm{tr}(\boldsymbol{\Lambda})}{(\mathbf{Y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{Y}-\mathbf{c})} - 2\frac{(\mathbf{Y}-\mathbf{c})'\boldsymbol{\Lambda}\boldsymbol{\Lambda}(\mathbf{Y}-\mathbf{c})}{\left[(\mathbf{Y}-\mathbf{c})'\boldsymbol{\Lambda}(\mathbf{Y}-\mathbf{c})\right]^2}\right\}\end{aligned}$$

Since from the definitions ($T4.93$)-($T4.94$) we obtain

$$\mathbf{Y}-\mathbf{c} = \sqrt{T}\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\widehat{\boldsymbol{\mu}}-\mathbf{b}), \qquad (T4.99)$$

we can simplify ($T4.98$) as follows:

$$\begin{aligned}\mathrm{E}\left\{\frac{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})}{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\mathbf{b})}\right\} &= \mathrm{E}\left\{\frac{\mathrm{tr}(\boldsymbol{\Lambda})}{T}\frac{1}{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\mathbf{b})}\right. \qquad (T4.100)\\ &\quad \left. - 2\frac{(\widehat{\boldsymbol{\mu}}-\mathbf{b})\mathbf{E}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\mathbf{E}'(\widehat{\boldsymbol{\mu}}-\mathbf{b})}{\left[(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\mathbf{b})\right]^2}\right\}\\ &= \mathrm{E}\left\{\frac{1}{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\mathbf{b})}\left(\frac{N\overline{\lambda}}{T} - \frac{2}{T}\frac{(\widehat{\boldsymbol{\mu}}-\mathbf{b})\boldsymbol{\Sigma}(\widehat{\boldsymbol{\mu}}-\mathbf{b})}{(\widehat{\boldsymbol{\mu}}-\mathbf{b})'(\widehat{\boldsymbol{\mu}}-\mathbf{b})}\right)\right\},\end{aligned}$$

where

$$\overline{\lambda} \equiv \frac{\operatorname{tr}(\mathbf{\Lambda})}{N} \tag{T4.101}$$

is the average of the eigenvalues.

From the relation ($A.68$) on the largest eigenvalue λ_1 in ($T4.92$) we obtain:

$$\frac{(\widehat{\boldsymbol{\mu}} - \mathbf{b})\, \mathbf{\Sigma}\, (\widehat{\boldsymbol{\mu}} - \mathbf{b})}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})} \leq \lambda_1. \tag{T4.102}$$

Therefore, substituting ($T4.100$) in ($T4.91$), using ($T4.102$) and recalling ($T4.90$) we obtain the following relation for the error:

$$\begin{aligned}[\operatorname{Err}(\boldsymbol{\delta}, \boldsymbol{\mu})]^2 = {} & [\operatorname{Err}(\widehat{\boldsymbol{\mu}}, \boldsymbol{\mu})]^2 + \operatorname{E}\left\{ \frac{a}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})} \right. \\ & \left. \left(a - 2\frac{N}{T}\overline{\lambda} + \frac{4}{T} \frac{(\widehat{\boldsymbol{\mu}} - \mathbf{b})\, \mathbf{\Sigma}\, (\widehat{\boldsymbol{\mu}} - \mathbf{b})}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})} \right) \right\} \\ \leq {} & [\operatorname{Err}(\widehat{\boldsymbol{\mu}}, \boldsymbol{\mu})]^2 + \operatorname{E}\left\{ \frac{a\left(a - \frac{2}{T}\left(N\overline{\lambda} - 2\lambda_1\right)\right)}{(\widehat{\boldsymbol{\mu}} - \mathbf{b})' (\widehat{\boldsymbol{\mu}} - \mathbf{b})} \right\} \\ \leq {} & [\operatorname{Err}(\widehat{\boldsymbol{\mu}}, \boldsymbol{\mu})]^2 \end{aligned} \tag{T4.103}$$

In particular, the lowest upper bound is reached at

$$a \equiv \frac{1}{T}\left(N\overline{\lambda} - 2\lambda_1\right). \tag{T4.104}$$

4.6 Shrinkage estimator of covariance

The number of non-zero eigenvalues of the sample covariance $\widehat{\mathbf{\Sigma}}$ is its rank. Thus we want to prove:

$$\operatorname{rank}\left(\widehat{\mathbf{\Sigma}}\right) < T \tag{T4.105}$$

we re-write the definition of the sample covariance as follows

$$\widehat{\mathbf{\Sigma}} = \frac{1}{T}\mathbf{X}'_{T\times N}\left(\mathbf{I}_T - \frac{1}{T}\mathbf{1}_T\mathbf{1}'_T\right)\mathbf{X}_{T\times N}, \tag{T4.106}$$

where $\mathbf{X}_{T\times N}$ is the matrix of past observations, $\mathbf{I}$ is the identity matrix and $\mathbf{1}$ is a vector of ones. From this expression and the property ($A.22$) of the rank operator we obtain

$$\operatorname{rank}\left(\widehat{\mathbf{\Sigma}}\right) \leq \operatorname{rank}\left(\mathbf{I}_T - \frac{1}{T}\mathbf{1}_T\mathbf{1}'_T\right) = T - 1. \tag{T4.107}$$

We denote as $\lambda_n(\mathbf{S})$ the n-th eigenvalue of the symmetric and positive matrix $\mathbf{S}$. We want to prove:

$$\frac{\lambda_N\left(\widehat{\mathbf{\Sigma}}^S\right)}{\lambda_1\left(\widehat{\mathbf{\Sigma}}^S\right)} > \frac{\lambda_N\left(\widehat{\mathbf{\Sigma}}\right)}{\lambda_1\left(\widehat{\mathbf{\Sigma}}\right)}. \qquad (T4.108)$$

We notice that the highest eigenvalue of the shrinkage estimator satisfies

$$\lambda_1\left(\widehat{\mathbf{\Sigma}}^S\right) < \lambda_1\left(\widehat{\mathbf{\Sigma}}\right). \qquad (T4.109)$$

To show this, we first prove that for arbitrary positive symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ and positive number α and β we have

$$\lambda_1\left(\alpha\mathbf{A}+\beta\mathbf{B}\right) \leq \alpha\lambda_1\left(\mathbf{A}\right)+\beta\lambda_1\left(\mathbf{B}\right). \qquad (T4.110)$$

This is true because from $(A.68)$ the largest eigenvalue of a matrix $\mathbf{A}$ satisfies:

$$\begin{aligned}\lambda_1\left(\alpha\mathbf{A}+\beta\mathbf{B}\right) &= \max_{\mathbf{v}'\mathbf{v}=1} \mathbf{v}'\left(\alpha\mathbf{A}+\beta\mathbf{B}\right)\mathbf{v} \qquad (T4.111)\\ &\leq \alpha \max_{\mathbf{v}'\mathbf{v}=1} \mathbf{v}'\mathbf{A}\mathbf{v} + \beta \max_{\mathbf{v}'\mathbf{v}=1} \mathbf{v}'\mathbf{B}\mathbf{v}\\ &= \alpha\lambda_1\left(\mathbf{A}\right)+\beta\lambda_1\left(\mathbf{B}\right)\end{aligned}$$

Therefore

$$\begin{aligned}\lambda_1\left(\widehat{\mathbf{\Sigma}}^S\right) &\equiv \lambda_1\left(\left(1-\alpha\right)\widehat{\mathbf{\Sigma}}+\alpha\widehat{\mathbf{C}}\right) \qquad (T4.112)\\ &\leq \left(1-\alpha\right)\lambda_1\left(\widehat{\mathbf{\Sigma}}\right)+\alpha\lambda_1\left(\widehat{\mathbf{C}}\right)\\ &= \lambda_1\left(\widehat{\mathbf{\Sigma}}\right)-\alpha\left[\lambda_1\left(\widehat{\mathbf{\Sigma}}\right)-\lambda_1\left(\widehat{\mathbf{C}}\right)\right] < \lambda_1\left(\widehat{\mathbf{\Sigma}}\right),\end{aligned}$$

where the last inequality follows from:

$$\lambda_1\left(\widehat{\mathbf{\Sigma}}\right) > \lambda_1\left(\widehat{\mathbf{C}}\right) \equiv \frac{1}{N}\sum_{n=1}^{N}\lambda_n\left(\widehat{\mathbf{\Sigma}}\right). \qquad (T4.113)$$

Similarly for the least eigenvalue we have

$$\lambda_N\left(\widehat{\mathbf{\Sigma}}^S\right) > \lambda_N\left(\widehat{\mathbf{\Sigma}}\right), \qquad (T4.114)$$

which follows from the above argument and the reverse identities $(A.69)$ and

$$\lambda_N\left(\widehat{\mathbf{\Sigma}}\right) < \lambda_N\left(\widehat{\mathbf{C}}\right) \equiv \frac{1}{N}\sum_{n=1}^{N}\lambda_n\left(\widehat{\mathbf{\Sigma}}\right). \qquad (T4.115)$$

4.7 Influence functions of common estimators

Estimators as implicit functionals

Consider the estimator $\widetilde{\boldsymbol{\theta}}\left[f_{i_T}\right]$, which is an implicit functional of the empirical pdf f_{i_T} defined by the following equation:

$$\mathbf{0} = \int_{\mathbb{R}^N} \boldsymbol{\psi}\left(\mathbf{x}, \widetilde{\boldsymbol{\theta}}[h]\right) h(\mathbf{x})\, d\mathbf{x}. \qquad (T4.116)$$

where h is a generic function. We consider the function $h_\epsilon \equiv (1-\epsilon) f_{\mathbf{X}} + \epsilon\delta^{(\mathbf{y})}$. Deriving in zero ($T4.116$) with respect to ϵ we obtain

$$\begin{aligned}
\mathbf{0} &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_{\mathbb{R}^N} \boldsymbol{\psi}\left(\mathbf{x}, \widetilde{\boldsymbol{\theta}}[h_\epsilon]\right)\left[(1-\epsilon) f_{\mathbf{X}}(\mathbf{x}) + \epsilon\delta^{(\mathbf{y})}(\mathbf{x})\right] d\mathbf{x} \\
&= \int_{\mathbb{R}^N} \left.\frac{\partial \boldsymbol{\psi}(\mathbf{x},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\widetilde{\boldsymbol{\theta}}[f_{\mathbf{X}}]} \left.\frac{d\widetilde{\boldsymbol{\theta}}[h_\epsilon]}{d\epsilon}\right|_{\epsilon=0} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \qquad (T4.117) \\
&\quad + \int_{\mathbb{R}^N} \boldsymbol{\psi}\left(\mathbf{x}, \widetilde{\boldsymbol{\theta}}[f_{\mathbf{X}}]\right)\left[-f_{\mathbf{X}}(\mathbf{x}) + \delta^{(\mathbf{y})}(\mathbf{x})\right] d\mathbf{x} \\
&= \int_{\mathbb{R}^N} \left.\frac{\partial \boldsymbol{\psi}(\mathbf{x},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\widetilde{\boldsymbol{\theta}}[f_{\mathbf{X}}]} f(\mathbf{x})\, d\mathbf{x} \left.\frac{d\widetilde{\boldsymbol{\theta}}[h_\epsilon]}{d\epsilon}\right|_{\epsilon=0} + \boldsymbol{\psi}\left(\mathbf{y}, \widetilde{\boldsymbol{\theta}}[f_{\mathbf{X}}]\right)
\end{aligned}$$

On the other hand, from the definition (4.185) of influence function we obtain:

$$\mathrm{IF}\left(\mathbf{y}, f_{\mathbf{X}}, \widehat{\boldsymbol{\theta}}\right) \equiv \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(\widetilde{\boldsymbol{\theta}}[h_\epsilon] - \widetilde{\boldsymbol{\theta}}[h_0]\right) = \left.\frac{d\widetilde{\boldsymbol{\theta}}[h_\epsilon]}{d\epsilon}\right|_0. \qquad (T4.118)$$

Therefore

$$\mathrm{IF}\left(\mathbf{y}, f_{\mathbf{X}}, \widehat{\boldsymbol{\theta}}\right) = -\left[\int_{\mathbb{R}^N} \left.\frac{\partial \boldsymbol{\psi}(\mathbf{x},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\widetilde{\boldsymbol{\theta}}[f_{\mathbf{X}}]} f(\mathbf{x})\, d\mathbf{x}\right]^{-1} \boldsymbol{\psi}\left(\mathbf{y}, \widetilde{\boldsymbol{\theta}}[f_{\mathbf{X}}]\right). \qquad (T4.119)$$

Estimators as explicit functionals

Sample estimators of the unknown quantity $G[f_{\mathbf{X}}]$ are by definition explicit functionals of the empirical pdf:

$$\widetilde{G}[f_{i_T}] \equiv G[f_{i_T}]. \qquad (T4.120)$$

Therefore from its defintion (4.185) the influence function reads:

$$\mathrm{IF}\left(\mathbf{y}, f_{\mathbf{X}}, \widehat{G}\right) \equiv \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(G\left[(1-\epsilon) f_{\mathbf{X}} + \epsilon\delta^{(\mathbf{y})}\right] - G[f_{\mathbf{X}}]\right), \qquad (T4.121)$$

where $\mathbf{y}$ is an arbitrary point. Now consider the function:

$$h_\epsilon \equiv (1-\epsilon) f_{\mathbf{X}} + \epsilon\delta^{(\mathbf{y})}. \qquad (T4.122)$$

The influence function can be written:

$$\mathrm{IF}\left(\mathbf{y}, f_{\mathbf{X}}, \widehat{G}\right) \equiv \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(G[h_\epsilon] - G[h_0]\right) = \left.\frac{dG[h_\epsilon]}{d\epsilon}\right|_0. \qquad (T4.123)$$

- Sample mean

Consider the functional associated with the sample mean $\widehat{\boldsymbol{\mu}}$, which reads:

$$\widetilde{\boldsymbol{\mu}}\left[h\right] \equiv \int_{\mathbb{R}^N} \mathbf{x} h\left(\mathbf{x}\right) d\mathbf{x}. \tag{T4.124}$$

From (T4.123) the influence function reads:

$$\operatorname{IF}\left(\mathbf{y}, f, \widehat{\boldsymbol{\mu}}\right) \equiv \left.\frac{d\widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]}{d\epsilon}\right|_0. \tag{T4.125}$$

First we compute:

$$\begin{aligned}
\widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right] &\equiv \int_{\mathbb{R}^N} \mathbf{x} h_\epsilon\left(\mathbf{x}\right) d\mathbf{x} \\
&= \int_{\mathbb{R}^N} \mathbf{x}\left(\left(1-\epsilon\right) f\left(\mathbf{x}\right) + \epsilon\delta^{(\mathbf{y})}\left(\mathbf{x}\right)\right) d\mathbf{x} \\
&= \left(1-\epsilon\right) \int_{\mathbb{R}^N} \mathbf{x} f\left(\mathbf{x}\right) d\mathbf{x} + \epsilon\mathbf{y} \\
&= \operatorname{E}\left\{\mathbf{X}\right\} + \epsilon\left(-\operatorname{E}\left\{\mathbf{X}\right\} + \mathbf{y}\right).
\end{aligned} \tag{T4.126}$$

From this and (T4.125) we derive:

$$\operatorname{IF}\left(\mathbf{y}, f, \widehat{\boldsymbol{\mu}}\right) = -\operatorname{E}\left\{\mathbf{X}\right\} + \mathbf{y}. \tag{T4.127}$$

- Sample covariance

Consider the fucntional associated with the sample covariance $\widehat{\boldsymbol{\Sigma}}$, which reads:

$$\widetilde{\boldsymbol{\Sigma}}\left[h\right] \equiv \int_{\mathbb{R}^N} \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h\right]\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h\right]\right)' h\left(\mathbf{x}\right) d\mathbf{x}. \tag{T4.128}$$

From (T4.123) the influence function reads:

$$\operatorname{IF}\left(\mathbf{y}, f, \widehat{\boldsymbol{\Sigma}}\right) \equiv \lim_{\epsilon\longrightarrow 0} \frac{1}{\epsilon}\left(\widetilde{\boldsymbol{\Sigma}}\left[h_\epsilon\right] - \widetilde{\boldsymbol{\Sigma}}\left[h_0\right]\right) = \left.\frac{d\widetilde{\boldsymbol{\Sigma}}\left[h_\epsilon\right]}{d\epsilon}\right|_0. \tag{T4.129}$$

First we compute:

$$\begin{aligned}
\widetilde{\boldsymbol{\Sigma}}\left[h_\epsilon\right] &\equiv \int_{\mathbb{R}^N} \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)' h_\epsilon\left(\mathbf{x}\right) d\mathbf{x} \\
&= \int_{\mathbb{R}^N} \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)' \left(\left(1-\epsilon\right) f\left(\mathbf{x}\right) + \epsilon\delta^{(\mathbf{y})}\left(\mathbf{x}\right)\right) d\mathbf{x} \\
&= \left(1-\epsilon\right) \int_{\mathbb{R}^N} \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)' f\left(\mathbf{x}\right) d\mathbf{x} \\
&\quad + \epsilon\left(\mathbf{y} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)\left(\mathbf{y} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)'
\end{aligned} \tag{T4.130}$$

Deriving this expression with respect to ϵ we obtain:

$$\begin{aligned}
\operatorname{IF}\left(\mathbf{y}, f, \widehat{\mathbf{\Sigma}}\right) &= \left.\frac{d\widetilde{\mathbf{\Sigma}}\left[h_\epsilon\right]}{d\epsilon}\right|_0 \\
&= -\int_{\mathbb{R}^N} \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_0\right]\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_0\right]\right)' f\left(\mathbf{x}\right) d\mathbf{x} \\
&\quad + (1-0) \int_{\mathbb{R}^N} \left.\frac{d}{d\epsilon}\right|_0 \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)' f\left(\mathbf{x}\right) d\mathbf{x} \\
&\quad + \left(\mathbf{y} - \widetilde{\boldsymbol{\mu}}\left[h_0\right]\right)\left(\mathbf{y} - \widetilde{\boldsymbol{\mu}}\left[h_0\right]\right)' \\
&\quad + 0 \times \left.\frac{d}{d\epsilon}\right|_0 \left(\mathbf{y} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)\left(\mathbf{y} - \widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]\right)'
\end{aligned} \tag{T4.131}$$

Using $\widetilde{\boldsymbol{\mu}}\left[h_0\right] = \operatorname{E}\left\{\mathbf{X}\right\}$ this means:

$$\begin{aligned}
\operatorname{IF}\left(\mathbf{y}, f, \widehat{\mathbf{\Sigma}}\right) &= -\operatorname{Cov}\left\{\mathbf{X}\right\} \\
&\quad - \int_{\mathbb{R}^N} 2 \left.\frac{d\widetilde{\boldsymbol{\mu}}\left[h_\epsilon\right]}{d\epsilon}\right|_0 \left(\mathbf{x} - \operatorname{E}\left\{\mathbf{X}\right\}\right) f\left(\mathbf{x}\right) d\mathbf{x} \\
&\quad + \left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)\left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)'
\end{aligned} \tag{T4.132}$$

Now using ($T4.127$) we obtain:

$$\begin{aligned}
\operatorname{IF}\left(\mathbf{y}, f, \widehat{\mathbf{\Sigma}}\right) &= -\operatorname{Cov}\left\{\mathbf{X}\right\} \\
&\quad -2 \int_{\mathbb{R}^N} \left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)\left(\mathbf{x} - \operatorname{E}\left\{\mathbf{X}\right\}\right)' f\left(\mathbf{x}\right) d\mathbf{x} \\
&\quad + \left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)\left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)'
\end{aligned} \tag{T4.133}$$

The term in the middle is null:

$$\begin{aligned}
\mathbf{Z} &\equiv \int_{\mathbb{R}^N} \left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)\left(\mathbf{x} - \operatorname{E}\left\{\mathbf{X}\right\}\right)' f\left(\mathbf{x}\right) d\mathbf{x} \\
&= \mathbf{y} \int_{\mathbb{R}^N} \left(\mathbf{x} - \operatorname{E}\left\{\mathbf{X}\right\}\right)' f\left(\mathbf{x}\right) d\mathbf{x} - \operatorname{E}\left\{\mathbf{X}\right\} \int_{\mathbb{R}^N} \left(\mathbf{x} - \operatorname{E}\left\{\mathbf{X}\right\}\right)' f\left(\mathbf{x}\right) d\mathbf{x} \\
&= \mathbf{0}
\end{aligned} \tag{T4.134}$$

Therefore:

$$\operatorname{IF}\left(\mathbf{y}, f, \widehat{\mathbf{\Sigma}}\right) = -\operatorname{Cov}\left\{\mathbf{X}\right\} + \left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)\left(\mathbf{y} - \operatorname{E}\left\{\mathbf{X}\right\}\right)' \tag{T4.135}$$

- OLS factor loadings

For the OLS factor loadings $\widehat{\mathbf{B}}$ define $\mathbf{z} \equiv (\mathbf{x}, \mathbf{f})$ the set of invariants and factors. We have

$$\mathbf{G}\left[f_{\mathbf{Z}}\right] = \left[\int \mathbf{x}\mathbf{f}' f_{\mathbf{Z}}\left(\mathbf{z}\right) d\mathbf{z}\right] \left[\int \mathbf{f}\mathbf{f}' f_{\mathbf{Z}}\left(\mathbf{z}\right) d\mathbf{z}\right]^{-1} \tag{T4.136}$$

Consider a point $\mathbf{w} \equiv \left(\widetilde{\mathbf{x}}, \widetilde{\mathbf{f}}\right)$. We have:

$$\mathbf{G}\left[f_{\mathbf{Z}} + \epsilon\left(\delta^{(\mathbf{w})} - f_{\mathbf{Z}}\right)\right] = (\mathbf{A} + \epsilon\mathbf{B})(\mathbf{C} + \epsilon\mathbf{D})^{-1}, \tag{T4.137}$$

where

$$\begin{aligned}
\mathbf{A} &\equiv \int \mathbf{x}\mathbf{f}' f_{\mathbf{Z}} d\mathbf{z} = \mathrm{E}\left\{\mathbf{X}\mathbf{F}'\right\} \\
\mathbf{B} &\equiv \int \mathbf{x}\mathbf{f}'\left(\delta^{(\mathbf{w})} - f_{\mathbf{Z}}\right) d\mathbf{z} = \widetilde{\mathbf{x}}\widetilde{\mathbf{f}}' - \mathrm{E}\left\{\mathbf{X}\mathbf{F}'\right\} \\
\mathbf{C} &\equiv \int \mathbf{f}\mathbf{f}' f_{\mathbf{Z}} d\mathbf{z} = \mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\} \\
\mathbf{D} &\equiv \int \mathbf{f}\mathbf{f}'\left(\delta^{(\mathbf{w})} - f_{\mathbf{Z}}\right) d\mathbf{z} = \widetilde{\mathbf{f}}\widetilde{\mathbf{f}}' - \mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\}
\end{aligned} \tag{T4.138}$$

Since

$$\begin{aligned}
(\mathbf{C} + \epsilon\mathbf{D})^{-1} &= \left(\mathbf{C}\left(\mathbf{I} + \epsilon\mathbf{C}^{-1}\mathbf{D}\right)\right)^{-1} \\
&= \left(\mathbf{I} + \epsilon\mathbf{C}^{-1}\mathbf{D}\right)^{-1}\mathbf{C}^{-1} \\
&\approx \left(\mathbf{I} - \epsilon\mathbf{C}^{-1}\mathbf{D}\right)\mathbf{C}^{-1}
\end{aligned} \tag{T4.139}$$

we have

$$\begin{aligned}
\mathbf{G}\left[f_{\mathbf{Z}} + \epsilon\left(\delta^{(\mathbf{w})} - f_{\mathbf{Z}}\right)\right] &\approx (\mathbf{A} + \epsilon\mathbf{B})\left(\mathbf{C}^{-1} - \epsilon\mathbf{C}^{-1}\mathbf{D}\mathbf{C}^{-1}\right) \\
&\approx \mathbf{A}\mathbf{C}^{-1} + \epsilon\left(\mathbf{B}\mathbf{C}^{-1} - \mathbf{A}\mathbf{C}^{-1}\mathbf{D}\mathbf{C}^{-1}\right).
\end{aligned} \tag{T4.140}$$

Therefore

$$\begin{aligned}
\mathrm{IF}\left(\mathbf{y}, f_{\mathbf{X}}, \widehat{\mathbf{B}}\right) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon}\left(\mathbf{G}\left[f_{\mathbf{Z}} + \epsilon\left(\delta^{(\mathbf{w})} - f_{\mathbf{Z}}\right)\right] - \mathbf{G}\left[f_{\mathbf{Z}}\right]\right) \\
&= \left(\mathbf{B}\mathbf{C}^{-1} - \mathbf{A}\mathbf{C}^{-1}\mathbf{D}\mathbf{C}^{-1}\right) \\
&= \left[\widetilde{\mathbf{x}}\widetilde{\mathbf{f}}' - \mathrm{E}\left\{\mathbf{X}\mathbf{F}'\right\}\right]\mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\}^{-1} \\
&\quad - \mathrm{E}\left\{\mathbf{X}\mathbf{F}'\right\}\mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\}^{-1}\left[\widetilde{\mathbf{f}}\widetilde{\mathbf{f}}' - \mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\}\right]\mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\}^{-1} \\
&= \left(\widetilde{\mathbf{x}}\widetilde{\mathbf{f}}' - \mathbf{B}\widetilde{\mathbf{f}}\widetilde{\mathbf{f}}'\right)\mathrm{E}\left\{\mathbf{F}\mathbf{F}'\right\}^{-1}
\end{aligned} \tag{T4.141}$$

4.8 Missing data: the E-M algorithm

Suppose we are at step u. The invariants are normally distributed with the following parameters:

$$\mathbf{X}_t \sim \mathrm{N}\left(\boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right) \tag{T4.142}$$

Consider the first non-central moment of $\mathbf{X}_t$ conditional on the observations:

$$\mu_{t,n}^{(u)} \equiv \mathrm{E}\left\{X_{t,n}|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\}. \tag{T4.143}$$

From (2.165), for the observed values we have:

$$\boldsymbol{\mu}_{t,\mathrm{obs}(t)}^{(u)} = \mathbf{x}_{t,\mathrm{obs}(t)}; \tag{T4.144}$$

and for the missing values we have:

$$\begin{aligned}\boldsymbol{\mu}_{t,\mathrm{mis}(t)}^{(u)} &= \boldsymbol{\mu}_{\mathrm{mis}(t)}^{(u)} \\ &\quad+\boldsymbol{\Sigma}_{\mathrm{mis}(t),\mathrm{obs}(t)}^{(u)}\left(\boldsymbol{\Sigma}_{\mathrm{obs}(t),\mathrm{obs}(t)}^{(u)}\right)^{-1}\left(\mathbf{x}_{t,\mathrm{obs}(t)} - \boldsymbol{\mu}_{\mathrm{obs}(t)}^{(u)}\right).\end{aligned} \tag{T4.145}$$

Consider now the second non-central conditional moment:

$$\mathbf{S}_t^{(u)} \equiv \mathrm{E}\left\{\mathbf{X}_t\mathbf{X}_t'|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\} \tag{T4.146}$$

This matrix consists of three sub-components:

$$\begin{aligned}\mathbf{S}_{t,\mathrm{obs}(t),\mathrm{obs}(t)}^{(u)} &\equiv \mathrm{E}\left\{\mathbf{X}_{t,\mathrm{obs}(t)}\mathbf{X}_{t,\mathrm{obs}(t)}'|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\} \\ &= \mathbf{x}_{t,\mathrm{obs}(t)}\mathbf{x}_{t,\mathrm{obs}(t)}' = \left[\boldsymbol{\mu}_{t,\mathrm{obs}(t)}^{(u)}\right]\left[\boldsymbol{\mu}_{t,\mathrm{obs}(t)}^{(u)}\right]';\end{aligned} \tag{T4.147}$$

$$\begin{aligned}\mathbf{S}_{t,\mathrm{mis}(t),\mathrm{obs}(t)}^{(u)} &\equiv \mathrm{E}\left\{\mathbf{X}_{t,\mathrm{mis}(t)}\mathbf{X}_{t,\mathrm{obs}(t)}'|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\} \\ &= \mathrm{E}\left\{\mathbf{X}_{t,\mathrm{mis}(t)}|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\}\mathbf{x}_{t,\mathrm{obs}(t)}' \\ &= \boldsymbol{\mu}_{t,\mathrm{mis}(t)}^{(u)}\mathbf{x}_{t,\mathrm{obs}(t)}' = \left[\boldsymbol{\mu}_{t,\mathrm{mis}(t)}^{(u)}\right]\left[\boldsymbol{\mu}_{t,\mathrm{obs}(t)}^{(u)}\right]';\end{aligned} \tag{T4.148}$$

and

$$\begin{aligned}\mathbf{S}_{t,\mathrm{mis}(t),\mathrm{mis}(t)}^{(u)} &\equiv \mathrm{E}\left\{\mathbf{X}_{t,\mathrm{mis}(t)}\mathbf{X}_{t,\mathrm{mis}(t)}'|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\} \\ &= \mathrm{E}\left\{\mathbf{X}_{t,\mathrm{mis}(t)}|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\}\mathrm{E}\left\{\mathbf{X}_{t,\mathrm{mis}(t)}|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\}' \\ &\quad+\mathrm{Cov}\left\{\mathbf{X}_{t,\mathrm{mis}(t)}|\mathbf{x}_{t,\mathrm{obs}(t)}, \boldsymbol{\mu}^{(u)}, \boldsymbol{\Sigma}^{(u)}\right\} \\ &= \left[\boldsymbol{\mu}_{t,\mathrm{mis}(t)}^{(u)}\right]\left[\boldsymbol{\mu}_{t,\mathrm{mis}(t)}^{(u)}\right]' + \boldsymbol{\Sigma}_{\mathrm{mis}(t),\mathrm{mis}(t)}^{(u)} \\ &\quad-\boldsymbol{\Sigma}_{\mathrm{mis}(t),\mathrm{obs}(t)}^{(u)}\left(\boldsymbol{\Sigma}_{\mathrm{obs}(t),\mathrm{obs}(t)}^{(u)}\right)^{-1}\boldsymbol{\Sigma}_{\mathrm{obs}(t),\mathrm{mis}(t)}^{(u)}.\end{aligned} \tag{T4.149}$$

In other words, defining the matrix $\mathbf{C}$ as

$$\mathbf{C}_{t,\mathrm{obs}(t),\mathrm{mis}(t)}^{(u)} = \mathbf{0}, \quad \mathbf{C}_{t,\mathrm{obs}(t),\mathrm{obs}(t)}^{(u)} = \mathbf{0}, \tag{T4.150}$$

and otherwise

$$\mathbf{C}_{t,\text{mis}(t),\text{mis}(t)}^{(u)} = \boldsymbol{\Sigma}_{\text{mis}(t),\text{mis}(t)}^{(u)} - \boldsymbol{\Sigma}_{\text{mis}(t),\text{obs}(t)}^{(u)} \left(\boldsymbol{\Sigma}_{\text{obs}(t),\text{obs}(t)}^{(u)}\right)^{-1} \boldsymbol{\Sigma}_{\text{obs}(t),\text{mis}(t)}^{(u)}, \tag{T4.151}$$

we can write

$$\mathbf{S}_t^{(u)} = \left[\boldsymbol{\mu}_t^{(u)}\right]\left[\boldsymbol{\mu}_t^{(u)}\right]' + \mathbf{C}_t^{(u)} \tag{T4.152}$$

Now we can update the estimate of the unconditional first moment as the sample mean of the conditional first moments:

$$\boldsymbol{\mu}^{(u+1)} \equiv \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\mu}_t^{(u)}. \tag{T4.153}$$

Similarly we can update the estimate of the unconditional second moment as the sample mean of the conditional second moments:

$$\mathbf{S}^{(u+1)} \equiv \frac{1}{T}\sum_{t=1}^{T} \mathbf{S}_t^{(u)}. \tag{T4.154}$$

The estimate of the covariance then follows from (T2.94):

$$\begin{aligned}\boldsymbol{\Sigma}^{(u+1)} &\equiv \mathbf{S}^{(u+1)} - \left[\boldsymbol{\mu}^{(u+1)}\right]\left[\boldsymbol{\mu}^{(u+1)}\right]' \\ &\approx \mathbf{S}^{(u+1)} - \left[\boldsymbol{\mu}^{(u)}\right]\left[\boldsymbol{\mu}^{(u)}\right]'.\end{aligned} \tag{T4.155}$$

From the definition (T4.153) and (T4.154) this is equivalent to

$$\boldsymbol{\Sigma}^{(u+1)} \equiv \frac{1}{T}\sum_{t=1}^{T}\left(\boldsymbol{\mu}_t^{(u)} - \boldsymbol{\mu}^{(u)}\right)\left(\boldsymbol{\mu}_t^{(u)} - \boldsymbol{\mu}^{(u)}\right)' + \mathbf{C}_t^{(u)}. \tag{T4.156}$$

With the above results we obtain the routine described in the text.

5

Technical appendix to Chapter 5

5.1 Gamma approximation of the investor's objective

Consider the generic second-order approximation (3.108) for the N prices of the securities in terms of the underlying K-dimensional market invariants $\mathbf{X}$, which we report here:

$$P_{T+\tau}^{(n)} \approx g^{(n)}(\mathbf{0}) + \mathbf{X}' \, \partial_{\mathbf{x}} g^{(n)} \Big|_{\mathbf{x}=\mathbf{0}} + \frac{1}{2} \mathbf{X}' \, \partial^2_{\mathbf{xx}} g^{(n)} \Big|_{\mathbf{x}=\mathbf{0}} \mathbf{X}, \tag{T5.1}$$

where $n = 1, \ldots, N$. From (5.11) the market is an invertible affine transformation of the prices, i.e.

$$\begin{aligned} M^{(n)} \approx a_n &+ \sum_{m=1}^{N} B_{nm} g^{(m)}(\mathbf{0}) + \sum_{m=1}^{N} B_{nm} \mathbf{X}' \, \partial_{\mathbf{x}} g^{(m)} \Big|_{\mathbf{x}=\mathbf{0}} \\ &+ \frac{1}{2} \sum_{m=1}^{N} B_{nm} \mathbf{X}' \, \partial^2_{\mathbf{xx}} g^{(m)} \Big|_{\mathbf{x}=\mathbf{0}} \mathbf{X}. \end{aligned} \tag{T5.2}$$

In turn, from (5.10) the objective is a linear combination of the market:

$$\begin{aligned} \Psi_{\boldsymbol{\alpha}} &\equiv \sum_{n=1}^{N} \alpha_n M^{(n)} \\ &\approx \sum_{n=1}^{N} \alpha_n a_n + \sum_{n=1}^{N} \alpha_n \sum_{m=1}^{N} B_{nm} g^{(m)}(\mathbf{0}) \\ &\quad + \sum_{n=1}^{N} \alpha_n \sum_{m=1}^{N} B_{nm} \left[\mathbf{X}' \, \partial_{\mathbf{x}} g^{(m)} \Big|_{\mathbf{x}=\mathbf{0}} \right] \\ &\quad + \frac{1}{2} \sum_{n=1}^{N} \alpha_n \sum_{m=1}^{N} B_{nm} \left[\mathbf{X}' \, \partial^2_{\mathbf{xx}} g^{(m)} \Big|_{\mathbf{x}=\mathbf{0}} \mathbf{X} \right].. \end{aligned} \tag{T5.3}$$

In other words,

$$\Psi_{\boldsymbol{\alpha}} \approx \Xi_{\boldsymbol{\alpha}} \equiv \theta_{\boldsymbol{\alpha}} + \boldsymbol{\Delta}_{\boldsymbol{\alpha}}' \mathbf{X} + \frac{1}{2} \mathbf{X}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \mathbf{X}, \tag{T5.4}$$

where

$$\theta_{\boldsymbol{\alpha}} \equiv \sum_{n=1}^{N} \alpha_n a_n + \sum_{n,m=1}^{N} \alpha_n B_{nm} g^{(m)}(\mathbf{0}).$$

$$\boldsymbol{\Delta}_{\boldsymbol{\alpha}} \equiv \sum_{n,m=1}^{N} \alpha_n B_{nm}\, \partial_{\mathbf{x}} g^{(m)} \Big|_{\mathbf{x}=\mathbf{0}}. \tag{T5.5}$$

$$\boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \equiv \sum_{n,m=1}^{N} \alpha_n B_{nm}\, \partial^2_{\mathbf{xx}} g^{(m)} \Big|_{\mathbf{x}=\mathbf{0}}$$

Assume now that the K-dimensional invariants $\mathbf{X}$ be normally distributed as in (5.29). Then we can compute explicitly the characteristic function of the approximate objective ($T5.4$). Defining:

$$\mathbf{Z} \equiv \mathbf{X} - \boldsymbol{\mu} \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}),$$

from ($T5.4$) we have:

$$\begin{aligned} \Xi_{\boldsymbol{\alpha}} &= \theta_{\boldsymbol{\alpha}} + \boldsymbol{\Delta}_{\boldsymbol{\alpha}}' (\boldsymbol{\mu} + \mathbf{Z}) + \frac{1}{2} (\boldsymbol{\mu} + \mathbf{Z})' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} (\boldsymbol{\mu} + \mathbf{Z}) \\ &= \theta_{\boldsymbol{\alpha}} + \boldsymbol{\Delta}_{\boldsymbol{\alpha}}' \boldsymbol{\mu} + \boldsymbol{\Delta}_{\boldsymbol{\alpha}}' \mathbf{Z} + \frac{1}{2} \boldsymbol{\mu}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \boldsymbol{\mu} + \boldsymbol{\mu}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \mathbf{Z} + \frac{1}{2} \mathbf{Z}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \mathbf{Z} \\ &= b_{\boldsymbol{\alpha}} + \mathbf{w}_{\boldsymbol{\alpha}}' \mathbf{Z} + \frac{1}{2} \mathbf{Z}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \mathbf{Z} \end{aligned} \tag{T5.6}$$

where

$$b_{\boldsymbol{\alpha}} \equiv \theta_{\boldsymbol{\alpha}} + \boldsymbol{\Delta}_{\boldsymbol{\alpha}}' \boldsymbol{\mu} + \frac{1}{2} \boldsymbol{\mu}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \boldsymbol{\mu} \tag{T5.7}$$

$$\mathbf{w}_{\boldsymbol{\alpha}} \equiv \boldsymbol{\Delta}_{\boldsymbol{\alpha}} + \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \boldsymbol{\mu} \tag{T5.8}$$

We performing the Cholesky decomposition of the covariance

$$\boldsymbol{\Sigma} \equiv \mathbf{B}\mathbf{B}' \tag{T5.9}$$

and the principal component decomposition of the following symmetric matrix:

$$\mathbf{B}' \boldsymbol{\Gamma}_{\boldsymbol{\alpha}} \mathbf{B} = \mathbf{E} \boldsymbol{\Lambda} \mathbf{E}', \tag{T5.10}$$

where $\mathbf{E}\mathbf{E}' = \mathbf{I}$ and $\boldsymbol{\Lambda}$ is the diagonal matrix of the eigenvalues. We define:

$$\mathbf{C} \equiv \mathbf{B}\mathbf{E} \tag{T5.11}$$

and we introduce the multivariate standard normal variable

$$\mathbf{Y} \equiv \mathbf{C}^{-1}\mathbf{Z} \sim \mathrm{N}\left(\mathbf{0}, \mathbf{E}'\mathbf{B}^{-1}\boldsymbol{\Sigma}\left(\mathbf{B}'\right)^{-1}\mathbf{E}\right) = \mathrm{N}\left(\mathbf{0}, \mathbf{I}_K\right). \tag{T5.12}$$

Finally we define:

$$\boldsymbol{\delta} \equiv \mathbf{C}'\mathbf{w}. \tag{T5.13}$$

In these terms and dropping the dependence on $\boldsymbol{\alpha}$ from the notation, (T5.6) becomes:

$$\begin{aligned} \Xi_{\boldsymbol{\alpha}} &= b + \mathbf{w}'\mathbf{C}\mathbf{C}^{-1}\mathbf{Z} + \frac{1}{2}\mathbf{Z}'\left(\mathbf{C}'\right)^{-1}\mathbf{C}'\boldsymbol{\Gamma}\mathbf{C}\mathbf{C}^{-1}\mathbf{Z} \\ &= b + \boldsymbol{\delta}'\mathbf{Y} + \frac{1}{2}\mathbf{Y}\mathbf{E}'\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\mathbf{E}\mathbf{Y} \\ &= b + \boldsymbol{\delta}'\mathbf{Y} + \frac{1}{2}\mathbf{Y}\mathbf{E}'\mathbf{E}\boldsymbol{\Lambda}\mathbf{E}'\mathbf{E}\mathbf{Y} \\ &= b + \sum_{k=1}^{K}\left(\delta_k Y_k + \frac{1}{2}\lambda_k Y_k^2\right), \end{aligned} \tag{T5.14}$$

As in Feuerverger and Wong (2000), we compute analytically the characteristic function of $\Xi_{\boldsymbol{\alpha}}$:

$$\begin{aligned} \phi_{\Xi}(\omega) &\equiv \mathrm{E}\left\{e^{i\Xi\omega}\right\} \\ &= \int_{\mathbb{R}^N} e^{i\omega\left[b+\sum_{k=1}^{K}\left(\delta_k y_k + \frac{\lambda_k}{2}y_k^2\right)\right]} f(\mathbf{y})\, d\mathbf{y} \end{aligned} \tag{T5.15}$$

where f is the standard normal density (2.156), which factors into the product of the marginal densities:

$$f(\mathbf{y}) = \prod_{k=1}^{K}\sqrt{2\pi}e^{-\frac{1}{2}y_k^2} \tag{T5.16}$$

Therefore the characteristic function (T5.15) becomes

$$\phi_{\Xi}(\omega) = e^{i\omega b}\prod_{k=1}^{K} G\left(\delta_k, \lambda_k\right), \tag{T5.17}$$

where

$$G(\delta, \lambda) \equiv \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega\left[\delta y + \frac{\lambda}{2}y^2\right]} e^{-\frac{y^2}{2}}\, dy \tag{T5.18}$$

Since

$$
\begin{aligned}
i\omega\delta y-\frac{1-i\omega\lambda}{2}y^{2} &= -\frac{1-i\omega\lambda}{2}\left(y^{2}-\frac{i\omega\delta}{\frac{1-i\omega\lambda}{2}}y\right) \\
&= -\frac{1-i\omega\lambda}{2}\left[y-\frac{i\omega\delta}{2\frac{1-i\omega\lambda}{2}}\right]^{2} \\
&\quad +\frac{1-i\omega\lambda}{2}\left(\frac{i\omega\delta}{2\frac{1-i\omega\lambda}{2}}\right)^{2} \\
&= -\frac{1-i\omega\lambda}{2}\left[y-\frac{i\omega\delta}{1-i\omega\lambda}\right]^{2}-\frac{(\omega\delta)^{2}}{2\,(1-i\omega\lambda)},
\end{aligned}
\tag{T5.19}
$$

we obtain:

$$
\begin{aligned}
G(\delta,\lambda) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{1-i\omega\lambda}{2}\left[y-\frac{i\omega\delta}{1-i\omega\lambda}\right]^{2}-\frac{(\omega\delta)^{2}}{2(1-i\omega\lambda)}}dy \\
&= \sqrt{\frac{1}{1-i\omega\lambda}}e^{-\frac{(\omega\delta)^{2}}{2(1-i\omega\lambda)}} \\
&\quad \int_{-\infty}^{+\infty}\frac{1}{\sqrt{\frac{2\pi}{1-i\omega\lambda}}}e^{-\frac{1-i\omega\lambda}{2}\left[y-\frac{i\omega\delta}{1-i\omega\lambda}\right]^{2}}dy \\
&= \frac{1}{\sqrt{1-i\lambda\omega}}e^{-\frac{\delta^{2}\omega^{2}}{2(1-i\lambda\omega)}}.
\end{aligned}
\tag{T5.20}
$$

Substituting this back into (T5.17) we finally obtain the expression of the characteristic function

$$
\begin{aligned}
\phi_{\Xi}(\omega) &= \frac{e^{i\omega b}}{\sqrt{\prod_{k=1}^{K}(1-i\omega\lambda_{k})}}e^{-\frac{1}{2}\sum_{k=1}^{K}\frac{\delta_{k}^{2}\omega^{2}}{(1-i\lambda_{k}\omega)}} \\
&= \left|\mathbf{I}_{K}-i\omega\boldsymbol{\Lambda}\right|^{-\frac{1}{2}}e^{i\omega b}e^{-\frac{1}{2}\boldsymbol{\delta}'(\mathbf{I}-i\omega\boldsymbol{\Lambda})^{-1}\boldsymbol{\delta}}
\end{aligned}
\tag{T5.21}
$$

Notice that substituting (T5.10) and (T5.9) we obtain:

$$
\begin{aligned}
\left|\mathbf{I}_{K}-i\omega\boldsymbol{\Lambda}\right| &= \left|\mathbf{E}\left(\mathbf{I}_{K}-i\omega\boldsymbol{\Lambda}\right)\mathbf{E}'\right| = \left|\mathbf{I}_{K}-i\omega\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\right| \\
&= \left|\left(\mathbf{B}'\right)^{-1}\left(\mathbf{I}_{K}-i\omega\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\right)\mathbf{B}'\right|D \\
&= \left|\mathbf{I}_{K}-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right|.
\end{aligned}
\tag{T5.22}
$$

On the other hand, substituting (T5.13) and (T5.11) we obtain:

$$
\begin{aligned}
\boldsymbol{\delta}'\left(\mathbf{I}-i\omega\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\delta} &= \mathbf{w}'\mathbf{C}\left(\mathbf{I}_{K}-i\omega\boldsymbol{\Lambda}\right)^{-1}\mathbf{C}'\mathbf{w} \\
&= \mathbf{w}'\mathbf{B}\mathbf{E}\left(\mathbf{I}_{K}-i\omega\boldsymbol{\Lambda}\right)^{-1}\mathbf{E}'\mathbf{B}'\mathbf{w} \\
&= \mathbf{w}'\mathbf{B}\left(\mathbf{E}\left(\mathbf{I}_{K}-i\omega\boldsymbol{\Lambda}\right)\mathbf{E}'\right)^{-1}\mathbf{B}'\mathbf{w}
\end{aligned}
\tag{T5.23}
$$

Substituting again (T5.10) and (T5.9) this reads:

$$\begin{aligned}\boldsymbol{\delta}'\left(\mathbf{I}-i\omega\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\delta} &= \mathbf{w}'\mathbf{B}\left(\mathbf{I}_K-i\omega\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{w} \\ &= \mathbf{w}'\mathbf{B}\mathbf{B}'\left(\mathbf{B}'\right)^{-1}\left(\mathbf{I}_K-i\omega\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{w} \\ &= \mathbf{w}'\mathbf{B}\mathbf{B}'\left(\mathbf{B}'\right)^{-1}\left(\mathbf{I}_K-i\omega\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\right)^{-1}\mathbf{B}'\mathbf{w} \\ &= \mathbf{w}'\boldsymbol{\Sigma}\left(\left(\mathbf{I}_K-i\omega\mathbf{B}'\boldsymbol{\Gamma}\mathbf{B}\right)\mathbf{B}'\right)^{-1}\mathbf{B}'\mathbf{w} \\ &= \mathbf{w}'\boldsymbol{\Sigma}\left(\mathbf{B}'-i\omega\mathbf{B}'\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right)^{-1}\mathbf{B}'\mathbf{w} \\ &= \mathbf{w}'\boldsymbol{\Sigma}\left(\mathbf{I}_K-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right)^{-1}\left(\mathbf{B}'\right)^{-1}\mathbf{B}'\mathbf{w} \\ &= \mathbf{w}'\boldsymbol{\Sigma}\left(\mathbf{I}_K-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right)^{-1}\mathbf{w}\end{aligned} \qquad (T5.24)$$

Therefore, substituting ($T5.22$) and ($T5.24$) in ($T5.21$) we obtain the characteristic function of the approximate objective:

$$\phi_{\Xi}\left(\omega\right)=\left|\mathbf{I}_K-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right|^{-\frac{1}{2}}e^{i\omega b}e^{-\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\left(\mathbf{I}_K-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right)^{-1}\mathbf{w}}. \qquad (T5.25)$$

Finally, substituting ($T5.7$) and ($T5.8$) we obtain:

$$\begin{aligned}\phi_{\Xi}\left(\omega\right)=&\left|\mathbf{I}_K-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right|^{-\frac{1}{2}}e^{i\omega\left(\theta+\boldsymbol{\Delta}'\boldsymbol{\mu}+\frac{1}{2}\boldsymbol{\mu}'\boldsymbol{\Gamma}\boldsymbol{\mu}\right)} \\ &e^{-\frac{1}{2}\left(\boldsymbol{\Delta}+\boldsymbol{\Gamma}\boldsymbol{\mu}\right)'\boldsymbol{\Sigma}\left(\mathbf{I}_K-i\omega\boldsymbol{\Gamma}\boldsymbol{\Sigma}\right)^{-1}\left(\boldsymbol{\Delta}+\boldsymbol{\Gamma}\boldsymbol{\mu}\right)}.\end{aligned} \qquad (T5.26)$$

To compute the moments of the approximate objective $\Xi_{\boldsymbol{\alpha}}$ we use ($T1.33$). In order to do this, we write the characteristic function as follows:

$$\phi_{\Xi_{\boldsymbol{\alpha}}}\left(\omega\right)=v^{-\frac{1}{2}}e^{u}, \qquad (T5.27)$$

where

$$u\left(\omega\right)\equiv i\omega b-\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\left(\mathbf{I}-i\omega\mathbf{V}\right)^{-1}\mathbf{w} \qquad (T5.28)$$

$$v\left(\omega\right)\equiv\left|\mathbf{I}-i\omega\mathbf{V}\right|. \qquad (T5.29)$$

and

$$\mathbf{V}\equiv\boldsymbol{\Gamma}\boldsymbol{\Sigma}. \qquad (T5.30)$$

To show how this works we explicitly compute the first three derivatives. It is easy to implement this approach systematically up to any order with by programming a software package such as Mathematica®. The first three derivatives of the characteristic function read:

$$\phi'_{\Xi}(\omega) = -\frac{1}{2}v^{-\frac{3}{2}}v'e^{u} + v^{-\frac{1}{2}}e^{u}u'$$
$$\phi''_{\Xi}(\omega) = \frac{3}{4}v^{-\frac{5}{2}}(v')^{2}e^{u} - \frac{1}{2}v^{-\frac{3}{2}}v''e^{u}$$
$$-v^{-\frac{3}{2}}v'e^{u}u' + v^{-\frac{1}{2}}e^{u}(u')^{2} + v^{-\frac{1}{2}}e^{u}u''$$
$$\phi'''_{\Xi}(\omega) = -\frac{15}{8}v^{-\frac{7}{2}}(v')^{3}e^{u} + \frac{3}{4}v^{-\frac{5}{2}}2v'v''e^{u} + \frac{3}{4}v^{-\frac{5}{2}}(v')^{2}e^{u}u' \tag{T5.31}$$
$$+\frac{3}{4}v^{-\frac{5}{2}}v'v''e^{u} - \frac{1}{2}v^{-\frac{3}{2}}v'''e^{u} - \frac{1}{2}v^{-\frac{3}{2}}v''e^{u}u'$$
$$+\frac{3}{2}v^{-\frac{5}{2}}(v')^{2}e^{u}u' - v^{-\frac{3}{2}}v''e^{u}u' - v^{-\frac{3}{2}}v'e^{u}(u')^{2} - v^{-\frac{3}{2}}v'e^{u}u''$$
$$-\frac{1}{2}v^{-\frac{3}{2}}v'e^{u}(u')^{2} + v^{-\frac{1}{2}}e^{u}(u')^{3} + 2v^{-\frac{1}{2}}e^{u}u'u''$$
$$-\frac{1}{2}v^{-\frac{3}{2}}v'e^{u}u'' + v^{-\frac{1}{2}}e^{u}u'u'' + v^{-\frac{1}{2}}e^{u}u'''$$

These expressions depend on the first three derivatives of u and v, which we obtain by applying the following generic rules that apply for any conformable matrices $\mathbf{M}, \mathbf{A}, \mathbf{B}$:

$$\frac{d\mathbf{M}^{-1}(\omega)}{d\omega} = \mathbf{M}^{-1}\frac{d\mathbf{M}(\omega)}{d\omega}\mathbf{M}^{-1} \tag{T5.32}$$
$$\frac{d(\mathbf{A}\mathbf{M}(\omega)\mathbf{B})}{d\omega} = \mathbf{A}\frac{d(\mathbf{M}(\omega))}{d\omega}\mathbf{B} \tag{T5.33}$$
$$\frac{d|\mathbf{A}+\omega\mathbf{B}|}{d\omega} = |\mathbf{A}+\omega\mathbf{B}|\,\mathrm{tr}\left[(\mathbf{A}+\omega\mathbf{B})^{-1}\mathbf{B}\right], \tag{T5.34}$$

where ($T5.32$) follows from ($A.126$), ($T5.33$) follows from a term by term expression of the product $\mathbf{AMB}$ and ($T5.34$) follows from ($A.124$).

Using these formulas in ($T5.28$) we obtain

$$u(\omega) \equiv i\omega b - \frac{1}{2}\mathbf{w}'\mathbf{\Sigma}(\mathbf{I}-i\omega\mathbf{V})^{-1}\mathbf{w}$$
$$u'(\omega) = ib - \frac{1}{2}\mathbf{w}'\mathbf{\Sigma}(\mathbf{I}-i\omega\mathbf{V})^{-1}[-i\mathbf{V}](\mathbf{I}-i\omega\mathbf{V})^{-1}\mathbf{w} \tag{T5.35}$$
$$u''(\omega) = -\mathbf{w}'\mathbf{\Sigma}(\mathbf{I}-i\omega\mathbf{V})^{-1}[-i\mathbf{V}](\mathbf{I}-i\omega\mathbf{V})^{-1}[-i\mathbf{V}](\mathbf{I}-i\omega\mathbf{V})^{-1}\mathbf{w}$$
$$u'''(\omega) = -3\mathbf{w}'\mathbf{\Sigma}(\mathbf{I}-i\omega\mathbf{V})^{-1}[-i\mathbf{V}](\mathbf{I}-i\omega\mathbf{V})^{-1}$$
$$[-i\mathbf{V}](\mathbf{I}-i\omega\mathbf{V})^{-1}[-i\mathbf{V}](\mathbf{I}-i\omega\mathbf{V})^{-1}\mathbf{w}$$

and

$$
\begin{aligned}
v(\omega) &\equiv |\mathbf{I} - i\omega\boldsymbol{\Sigma}| \\
v(\omega)' &= |\mathbf{I} + i\omega\boldsymbol{\Sigma}| \operatorname{tr}\left[(\mathbf{I} + i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right] \\
v''(\omega) &= |\mathbf{I} + i\omega\boldsymbol{\Sigma}| \operatorname{tr}\left[(\mathbf{I} + i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right] \operatorname{tr}\left[(\mathbf{I} + i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right] \\
&\quad + |\mathbf{I} + i\omega\boldsymbol{\Sigma}| \operatorname{tr}\left[(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(-i\boldsymbol{\Sigma})(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right] \\
v'''(\omega) &= |\mathbf{I} + i\omega\boldsymbol{\Sigma}| \left(\operatorname{tr}\left[(\mathbf{I} + i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right]\right)^3 \qquad (T5.36) \\
&\quad +3\,|\mathbf{I} + i\omega\boldsymbol{\Sigma}| \operatorname{tr}\left[(\mathbf{I} + i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right] \\
&\quad \operatorname{tr}\left[(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(-i\boldsymbol{\Sigma})(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right] \\
&\quad + |\mathbf{I} + i\omega\boldsymbol{\Sigma}| \operatorname{tr}\left[(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}[-i\boldsymbol{\Sigma}](\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(-i\boldsymbol{\Sigma})(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right. \\
&\quad \left. + (\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(-i\boldsymbol{\Sigma})(\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}[-i\boldsymbol{\Sigma}](\mathbf{I} - i\omega\boldsymbol{\Sigma})^{-1}(i\boldsymbol{\Sigma})\right]
\end{aligned}
$$

Evaluating these derivatives in $\omega = 0$ we obtain

$$
\begin{aligned}
u &\equiv -\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \\
u' &= i\left(b + \frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{V}\mathbf{w}\right) \qquad (T5.37) \\
u'' &= \mathbf{w}'\boldsymbol{\Sigma}\mathbf{V}^2\mathbf{w} \\
u''' &= -i3\mathbf{w}'\boldsymbol{\Sigma}\mathbf{V}^3\mathbf{w}
\end{aligned}
$$

and

$$
\begin{aligned}
v &\equiv 1 \\
v' &= i\operatorname{tr}(\mathbf{V}) \qquad (T5.38) \\
v'' &= -\left[\operatorname{tr}(\mathbf{V})\right]^2 + \operatorname{tr}\left(\mathbf{V}^2\right) \\
v''' &= -i\left[\operatorname{tr}(\mathbf{V})\right]^3 + 3i\operatorname{tr}(\mathbf{V})\operatorname{tr}\left(\mathbf{V}^2\right) - 2i\operatorname{tr}\left(\mathbf{V}^3\right)
\end{aligned}
$$

These values must be substituted in ($T5.31$) to yield the expressions of the first three non-central moments as in Appendix www.1.6.

For example, the first moment is

$$
\begin{aligned}
\mathrm{E}\left\{\Xi_{\boldsymbol{\alpha}}\right\} &= i^{-1}\phi'_{\Xi_{\boldsymbol{\alpha}}}(0) = i^{-1}\left(-\frac{1}{2}v^{-\frac{3}{2}}v'e^u + v^{-\frac{1}{2}}e^u u'\right) \qquad (T5.39) \\
&= e^{-\frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}}\left(\left(b + \frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{V}\mathbf{w}\right) - \frac{1}{2}\operatorname{tr}(\mathbf{V})\right).
\end{aligned}
$$

Finally, the explicit dependence on allocation comes from substituting in the final expression ($T5.7$), ($T5.8$) and ($T5.30$).

5.2 Properties of generic indices of satisfaction

Consistence with weak dominance

We follow a personal communication by D. Tasche to see that estimability and sensibility imply consistence with weak dominance.

Assume that $\Psi_{\boldsymbol{\alpha}}$ weakly dominates $\Psi_{\boldsymbol{\beta}}$. From the definition (5.36), this means that for all $u \in (0,1)$ the following inequality holds:

$$Q_{\Psi_{\boldsymbol{\alpha}}}(u) \geq Q_{\Psi_{\boldsymbol{\beta}}}(u) . \qquad (T5.40)$$

We want to prove that if $\mathcal{S}$ is estimable and sensible, then

$$\mathcal{S}(\boldsymbol{\alpha}) \geq \mathcal{S}(\boldsymbol{\beta}) . \qquad (T5.41)$$

From the definition of estimability (5.52) the index must be a function of the distribution of the objective, as represented, say, by the cdf:

$$\mathcal{S}(\boldsymbol{\alpha}) = G\left[F_{\Psi_{\boldsymbol{\alpha}}}\right] . \qquad (T5.42)$$

From (2.27) the random variable X defined below satisfies has the same distribution as the objective:

$$X_{\boldsymbol{\alpha}} \equiv Q_{\Psi_{\boldsymbol{\alpha}}}(U) \stackrel{d}{=} \Psi_{\boldsymbol{\alpha}}, \quad U \sim \mathrm{U}([0,1]) \qquad (T5.43)$$

Therefore both variables share the same cumulative distribution function:

$$F_{\Psi_{\boldsymbol{\alpha}}}(\psi) = F_{X_{\boldsymbol{\alpha}}}(\psi) , \qquad \psi \in \mathbb{R}. \qquad (T5.44)$$

Thus

$$\mathcal{S}(\boldsymbol{\alpha}) = G\left[F_{\Psi_{\boldsymbol{\alpha}}}\right] = G\left[F_{X_{\boldsymbol{\alpha}}}\right] . \qquad (T5.45)$$

A similar result holds for the second allocation:

$$\mathcal{S}(\boldsymbol{\beta}) = G\left[F_{\Psi_{\boldsymbol{\beta}}}\right] = G\left[F_{X_{\boldsymbol{\beta}}}\right] . \qquad (T5.46)$$

On the other hand, from ($T5.40$) in all scenarios $X_{\boldsymbol{\alpha}} \geq X_{\boldsymbol{\beta}}$, i.e. $X_{\boldsymbol{\alpha}}$ strongly dominates $X_{\boldsymbol{\beta}}$. Therefore, from the sensibility of $\mathcal{S}$ we must have

$$G\left[F_{X_{\boldsymbol{\alpha}}}\right] \geq G\left[F_{X_{\boldsymbol{\beta}}}\right] . \qquad (T5.47)$$

This and ($T5.45$)-($T5.46$) in turn imply the desired result:

$$\mathcal{S}(\boldsymbol{\alpha}) = G\left[F_{X_{\boldsymbol{\alpha}}}\right] \geq G\left[F_{X_{\boldsymbol{\beta}}}\right] = \mathcal{S}(\boldsymbol{\beta}) . \qquad (T5.48)$$

Constancy

Assume that an index of satisfaction is translation invariant

$$\Psi_{\mathbf{g}} \equiv 1 \Rightarrow \mathcal{S}\left(\boldsymbol{\alpha} + \lambda \mathbf{g}\right) = \mathcal{S}\left(\boldsymbol{\alpha}\right) + \lambda; \tag{T5.49}$$

and positive homogeneous

$$\mathcal{S}\left(\lambda \boldsymbol{\alpha}\right) = \lambda \mathcal{S}\left(\boldsymbol{\alpha}\right), \quad \text{for all } \lambda \geq 0. \tag{T5.50}$$

Then it displays the constancy feature. Indeed assume that $\mathbf{b}$ is a deterministic allocation $\Psi_{\mathbf{b}} \equiv \psi_{\mathbf{b}}$ then Indeed

$$\begin{aligned} \mathcal{S}\left(\boldsymbol{\alpha} + \lambda \mathbf{b}\right) &= \mathcal{S}\left(\psi_{\mathbf{b}}\left(\frac{\boldsymbol{\alpha}}{\psi_{\mathbf{b}}} + \lambda \frac{\mathbf{b}}{\psi_{\mathbf{b}}}\right)\right) \\ &= \psi_{\mathbf{b}} \mathcal{S}\left(\frac{\boldsymbol{\alpha}}{\psi_{\mathbf{b}}} + \lambda \frac{\mathbf{b}}{\psi_{\mathbf{b}}}\right) \\ &= \psi_{\mathbf{b}}\left[\mathcal{S}\left(\frac{\boldsymbol{\alpha}}{\psi_{\mathbf{b}}}\right) + \lambda\right] \\ &= \mathcal{S}\left(\psi_{\mathbf{b}} \frac{\boldsymbol{\alpha}}{\psi_{\mathbf{b}}}\right) + \lambda \psi_{\mathbf{b}} \\ &= \mathcal{S}\left(\boldsymbol{\alpha}\right) + \lambda \psi_{\mathbf{b}} \end{aligned} \tag{T5.51}$$

In particular

$$\mathcal{S}\left(\lambda \mathbf{b}\right) = \mathcal{S}\left(\mathbf{0} + \lambda \mathbf{b}\right) = \mathcal{S}\left(\mathbf{0}\right) + \lambda \psi_{\mathbf{b}} = \lambda \psi_{\mathbf{b}}. \tag{T5.52}$$

From the homogeneity $\mathcal{S}\left(\mathbf{0}\right) \equiv \mathbf{0}$. Therefore, setting $\lambda \equiv 1$ we obtain

$$\mathcal{S}\left(\mathbf{b}\right) = \psi_{\mathbf{b}}, \tag{T5.53}$$

which is the constancy property (5.62).

5.3 Properties of the certainty-equivalent

Consistence with stochastic dominance

To see that increasing utility implies consistence with weak dominance, with a change of variable we write expected utility as follows:

$$\mathrm{E}\left\{u\left(\Psi\right)\right\} = \int_{-\infty}^{+\infty} u\left(\psi\right) f_{\psi}\left(\psi\right) d\psi = \int_{0}^{1} u\left(Q_{\Psi}\left(s\right)\right) ds. \tag{T5.54}$$

Now, assume that the following inequality holds:

$$Q_{\Psi_{\boldsymbol{\alpha}}}\left(s\right) \geq Q_{\Psi_{\boldsymbol{\beta}}}\left(s\right), \quad \text{for all } s \in \left(0,1\right). \tag{T5.55}$$

Then

$$\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} \geq \mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\beta}}\right)\right\}. \qquad (T5.56)$$

Due to (5.99) this also implies that

$$\mathrm{CE}\left(\boldsymbol{\alpha}\right) \geq \mathrm{CE}\left(\boldsymbol{\beta}\right), \qquad (T5.57)$$

which shows that (5.109) holds true.

Positive homogeneity

From its definition

$$\mathrm{CE}\left(\boldsymbol{\alpha}\right) \equiv u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right), \qquad (T5.58)$$

the certainty-equivalent is positive homogeneous if it satisfies:

$$u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\lambda\boldsymbol{\alpha}}\right)\right\}\right) = \lambda u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right). \qquad (T5.59)$$

Since from (5.16) the objective is positive homogeneous

$$\Psi_{\lambda\boldsymbol{\alpha}} = \lambda\Psi_{\boldsymbol{\alpha}}, \qquad (T5.60)$$

we obtain an equation that does not depend on the allocation:

$$u^{-1}\left(\mathrm{E}\left\{u\left(\lambda\Psi\right)\right\}\right) = \lambda u^{-1}\left(\mathrm{E}\left\{u\left(\Psi\right)\right\}\right). \qquad (T5.61)$$

Assume the utility function is of the power type:

$$u\left(\psi\right) \equiv \psi^{\beta}. \qquad (T5.62)$$

Then, using (T1.14) and the change of variable $y \equiv \lambda\psi$, we obtain:

$$\begin{aligned}\mathrm{E}\left\{u\left(\lambda\Psi\right)\right\} &= \int_{\mathbb{R}} y^{\beta} f_{\lambda\Psi}\left(y\right) dy = \int_{\mathbb{R}} y^{\beta}\frac{1}{\lambda} f_{\Psi}\left(\frac{y}{\lambda}\right) dy \\ &= \lambda^{\beta}\int_{\mathbb{R}} \psi^{\beta} f_{\Psi}\left(\psi\right) d\psi = \lambda^{\beta}\,\mathrm{E}\left\{u\left(\Psi\right)\right\}.\end{aligned} \qquad (T5.63)$$

Since the inverse of the power utility function (T5.62) reads:

$$u^{-1}\left(z\right) = z^{\frac{1}{\beta}}, \qquad (T5.64)$$

from (T5.63) we obtain:

$$\begin{aligned}u^{-1}\left(\mathrm{E}\left\{u\left(\lambda\Psi\right)\right\}\right) &= \left[\mathrm{E}\left\{u\left(\lambda\Psi\right)\right\}\right]^{\frac{1}{\beta}} = \left[\lambda^{\beta}\,\mathrm{E}\left\{u\left(\Psi\right)\right\}\right]^{\frac{1}{\beta}} \\ &= \lambda\left[\mathrm{E}\left\{u\left(\Psi\right)\right\}\right]^{\frac{1}{\beta}} = \lambda u^{-1}\left(\mathrm{E}\left\{u\left(\Psi\right)\right\}\right),\end{aligned} \qquad (T5.65)$$

which is (T5.61) and thus concludes the proof.

Translation invariance

From its definition

$$\mathrm{CE}(\boldsymbol{\alpha}) \equiv u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right) \tag{$T5.66$}$$

the certainty-equivalent is translation invariant if it satisfies:

$$\Psi_{\mathbf{b}} \equiv 1 \Rightarrow u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}+\lambda\mathbf{b}}\right)\right\}\right) = u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right) + \lambda. \tag{$T5.67$}$$

Since from (5.17) the objective is additive

$$\Psi_{\boldsymbol{\alpha}+\lambda\mathbf{b}} = \Psi_{\boldsymbol{\alpha}} + \lambda, \tag{$T5.68$}$$

we obtain an equation that does not depend on the allocation:

$$u^{-1}\left(\mathrm{E}\left\{u\left(\Psi+\lambda\right)\right\}\right) = u^{-1}\left(\mathrm{E}\left\{u\left(\Psi\right)\right\}\right) + \lambda. \tag{$T5.69$}$$

Assume the utility function is exponential:

$$u = -e^{-\beta\psi}. \tag{$T5.70$}$$

Then, using ($T1.14$) and the change of variable $y \equiv \psi + \lambda$ we obtain:

$$\begin{aligned}\mathrm{E}\left\{u\left(\Psi+\lambda\right)\right\} &= \int -e^{-\beta y} f_{\Psi+\lambda}(y)\, dy = \int -e^{-\beta y} f_{\Psi}(y-\lambda)\, dy \\ &= e^{-\beta\lambda} \int -e^{-\beta\psi} f_{\Psi}(\psi)\, d\psi = e^{-\beta\lambda}\, \mathrm{E}\left\{u\left(\Psi\right)\right\},\end{aligned} \tag{$T5.71$}$$

On the other hand the inverse of the exponential utility function reads:

$$u^{-1}(z) = -\frac{\ln(-z)}{\beta}. \tag{$T5.72$}$$

Therefore

$$\begin{aligned}u^{-1}\left(\mathrm{E}\left\{u\left(\Psi+\lambda\right)\right\}\right) &= -\frac{1}{\beta}\ln\left(-\left[\mathrm{E}\left\{u\left(\Psi+\lambda\right)\right\}\right]\right) \\ &= -\frac{1}{\beta}\ln\left(e^{-\beta\lambda}\left[-\mathrm{E}\left\{u\left(\Psi\right)\right\}\right]\right) \\ &= \lambda - \frac{1}{\beta}\left[\ln\left(\left[-\mathrm{E}\left\{u\left(\Psi\right)\right\}\right]\right)\right] \\ &= \lambda + u^{-1}\left(\mathrm{E}\left\{u\left(\Psi\right)\right\}\right),\end{aligned} \tag{$T5.73$}$$

which is ($T5.69$) and thus concludes the proof.

Risk aversion/propensity

Since we only deal with increasing utility functions from (5.99) to prove risk aversion we can prove equivalently the following implication:

$$\Psi_{\mathbf{b}} = \psi_{\mathbf{b}}, \quad \mathrm{E}\{\Psi_{\mathbf{f}}\} = 0 \Rightarrow \mathrm{E}\{u(\Psi_{\mathbf{b}})\} \geq \mathrm{E}\{u(\Psi_{\mathbf{b+f}})\}. \tag{$T5.74$}$$

On the one hand, we have the following chain of identity.

$$\begin{aligned} \mathrm{E}\{u(\Psi_{\mathbf{b}})\} &= u(\psi_{\mathbf{b}}) = u(\mathrm{E}\{\Psi_{\mathbf{b}}\}) \\ &= u(\mathrm{E}\{\Psi_{\mathbf{b}} + \Psi_{\mathbf{f}}\}) = u(\mathrm{E}\{\Psi_{\mathbf{b+f}}\}), \end{aligned} \tag{$T5.75$}$$

On the other hand, Jensen's inequality states that for any random variable Ψ the following is true if and only if u is concave:

$$u(\mathrm{E}\{\Psi\}) \geq \mathrm{E}\{u(\Psi)\}. \tag{$T5.76$}$$

Therefore if and only if u is concave we obtain the following result:

$$\mathrm{E}\{u(\Psi_{\mathbf{b}})\} \geq \mathrm{E}\{u(\Psi_{\mathbf{b+f}})\}, \tag{$T5.77$}$$

which proves ($T5.74$).

A similar proof links convexity of the utility function with risk propensity and linearity of the utility function with risk neutrality.

To compute the risk premium in the case of small bets, from a second order Taylor expansion and using the assumptions

$$\Psi_{\mathbf{b}} = \psi_{\mathbf{b}}, \quad \mathrm{E}\{\Psi_{\mathbf{f}}\} = 0 \tag{$T5.78$}$$

we obtain:

$$\begin{aligned} u(\mathrm{CE}(\mathbf{b}+\mathbf{f})) &\equiv \mathrm{E}\{u(\Psi_{\mathbf{b+f}})\} = \mathrm{E}\{u(\psi_{\mathbf{b}} + \Psi_{\mathbf{f}})\} \\ &\approx u(\psi_{\mathbf{b}}) + \frac{\mathrm{Var}\{\Psi_{\mathbf{f}}\}}{2} u''(\psi_{\mathbf{b}}), \end{aligned} \tag{$T5.79$}$$

On the other hand from the definition of risk premium (5.85), which we report here

$$\mathrm{RP}(\mathbf{b},\mathbf{f}) \equiv \mathrm{CE}(\mathbf{b}) - \mathrm{CE}(\mathbf{b}+\mathbf{f}). \tag{$T5.80$}$$

and the constancy of the certainty-equivalent

$$\mathrm{CE}(\mathbf{b}) = \psi_{\mathbf{b}} \tag{$T5.81$}$$

we obtain

$$\mathrm{CE}(\mathbf{b}+\mathbf{f}) = \psi_{\mathbf{b}} - \mathrm{RP}(\mathbf{b},\mathbf{f}), \tag{$T5.82$}$$

where $\mathrm{RP}(\mathbf{b},\mathbf{f})$ is a small quantity. Therefore

$$\begin{aligned} u(\mathrm{CE}(\mathbf{b}+\mathbf{f})) &= u(\psi_{\mathbf{b}} - \mathrm{RP}(\mathbf{b},\mathbf{f})) \\ &\approx u(\psi_{\mathbf{b}}) - u'(\psi_{\mathbf{b}}) \mathrm{RP}(\mathbf{b},\mathbf{f}) \end{aligned} \tag{$T5.83$}$$

Thus from ($T5.79$) and ($T5.83$) we obtain:

$$\mathrm{RP}(\mathbf{b},\mathbf{f}) \approx -\frac{u''(\psi_{\mathbf{b}})}{u'(\psi_{\mathbf{b}})} \frac{\mathrm{Var}\{\Psi_{\mathbf{f}}\}}{2}. \tag{$T5.84$}$$

Dependence on allocation: approximation in terms of the moments of the objective

To compute the approximate expression of the certainty-equivalent in terms of the moments of the allocation we first expand the utility function around an arbitrary value $\widetilde{\psi}$:

$$\begin{aligned} u(\psi) = {} & u\left(\widetilde{\psi}\right) + u'\left(\widetilde{\psi}\right)\left(\psi - \widetilde{\psi}\right) + \frac{1}{2}u''\left(\widetilde{\psi}\right)\left(\psi - \widetilde{\psi}\right)^2 \\ & + \frac{1}{3!}u'''\left(\widetilde{\psi}\right)\left(\psi - \widetilde{\psi}\right)^3 + \cdots . \end{aligned} \quad (T5.85)$$

Taking expectations and pivoting the expansion around the objective's expected value

$$\widetilde{\psi} \equiv \mathrm{E}\{\Psi\} \quad (T5.86)$$

we obtain:

$$\mathrm{E}\{u(\Psi)\} \approx u(\mathrm{E}\{\Psi\}) + \frac{u''(\mathrm{E}\{\Psi\})}{2}\mathrm{Var}\{\Psi\}, \quad (T5.87)$$

where the term in the first derivative cancels out. On the other hand, another Taylor expansion yields:

$$u^{-1}(z + \epsilon) \approx u^{-1}(z) + \frac{1}{u'\left(u^{-1}(z)\right)}\epsilon. \quad (T5.88)$$

Substituting ($T5.87$) in ($T5.88$) we obtain:

$$\begin{aligned} u^{-1}(\mathrm{E}\{u(\Psi)\}) & \approx u^{-1}(u(\mathrm{E}\{\Psi\})) + \frac{u''(\mathrm{E}\{\Psi\})}{2u'\left(u^{-1}(u(\mathrm{E}\{\Psi\}))\right)}\mathrm{Var}\{\Psi\} \\ & = \mathrm{E}\{\Psi\} + \frac{u''}{2u'}(\mathrm{E}\{\Psi\})\mathrm{Var}\{\Psi\} \end{aligned} \quad (T5.89)$$

so that the first order approximation reads:

$$\mathrm{CE}(\boldsymbol{\alpha}) \approx \mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} - \frac{\mathrm{A}(\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\})}{2}\mathrm{Var}\{\Psi_{\boldsymbol{\alpha}}\}. \quad (T5.90)$$

First order sensitivity analysis

To compute the marginal contribution of the allocation $\boldsymbol{\alpha}$ to the certainty-equivalent we use the chain rule of calculus. First we derive the following result:

$$\begin{aligned} \partial_{\boldsymbol{\alpha}} \mathrm{E}\{u(\Psi_{\boldsymbol{\alpha}})\} & = \partial_{\boldsymbol{\alpha}}\left[\int_{\mathbb{R}^N} u(\psi_{\boldsymbol{\alpha}}) f_{\mathbf{M}}(\mathbf{m})\, d\mathbf{m}\right] \\ & = \int_{\mathbb{R}^N} \partial_{\boldsymbol{\alpha}}(\psi_{\boldsymbol{\alpha}})\, u'(\psi_{\boldsymbol{\alpha}}) f_{\mathbf{M}}(\mathbf{m})\, d\mathbf{m} \\ & = \mathrm{E}\{\partial_{\boldsymbol{\alpha}}(\Psi_{\boldsymbol{\alpha}})\, u'(\Psi_{\boldsymbol{\alpha}})\} \end{aligned} \quad (T5.91)$$

From this and the chain rule we obtain:

$$\begin{aligned}\partial_{\boldsymbol{\alpha}} \mathrm{CE}(\boldsymbol{\alpha}) &= \partial_{\boldsymbol{\alpha}} \left[u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right)\right] \\ &= \frac{du^{-1}}{dz}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right) \partial_{\boldsymbol{\alpha}} \mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} \\ &= \frac{1}{u'\left(u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right)\right)} \partial_{\boldsymbol{\alpha}} \mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} \\ &= \frac{\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right) \partial_{\boldsymbol{\alpha}}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}}{u'\left(\mathrm{CE}(\boldsymbol{\alpha})\right)},\end{aligned} \tag{T5.92}$$

For example consider the case where the objective are the net gains:

$$\Psi_{\boldsymbol{\alpha}} \equiv \boldsymbol{\alpha}'\left(\mathbf{P}_{T+\tau} - \mathbf{P}_T\right). \tag{T5.93}$$

Then the market vector (5.10) reads:

$$\mathbf{M} \equiv \mathbf{P}_{T+\tau} - \mathbf{P}_T, \tag{T5.94}$$

and the partial derivative of the objective (T5.93) reads:

$$\partial_{\boldsymbol{\alpha}}\left(\Psi_{\boldsymbol{\alpha}}\right) = \mathbf{M}. \tag{T5.95}$$

Assume that the utility is the error function:

$$u(\psi) \equiv \operatorname{erf}\left(\frac{\psi}{\sqrt{2\eta}}\right). \tag{T5.96}$$

Then the first derivative reads:

$$u'(\psi) = \sqrt{\frac{2}{\pi\eta}} e^{-\frac{1}{2\eta}\psi^2} \tag{T5.97}$$

Assume that the markets are normally distributed:

$$\mathbf{P}_{T+\tau} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{T5.98}$$

Then the market vector (5.10) is normally distributed:

$$\mathbf{M} \sim \mathrm{N}(\boldsymbol{\nu}, \boldsymbol{\Sigma}), \quad \boldsymbol{\nu} \equiv \boldsymbol{\mu} - \mathbf{P}_T. \tag{T5.99}$$

Thus the numerator in (T5.92) reads:

$$\begin{aligned}\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right) \partial_{\boldsymbol{\alpha}}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} &= \sqrt{\frac{2}{\pi\eta}} \mathrm{E}\left\{e^{\left(-\frac{1}{2\eta}\left(\boldsymbol{\alpha}'\mathbf{M}\right)^2\right)} \mathbf{M}\right\} \\ &= \sqrt{\frac{2}{\pi\eta}} \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \int_{\mathbb{R}^N} \mathbf{m} e^{-\frac{1}{2}\mathbf{m}'\left(\frac{\boldsymbol{\alpha}\boldsymbol{\alpha}'}{\eta}\right)\mathbf{m}} e^{-\frac{1}{2}(\mathbf{m}-\boldsymbol{\nu})'\boldsymbol{\Sigma}^{-1}(\mathbf{m}-\boldsymbol{\nu})} d\mathbf{m} \\ &= \sqrt{\frac{2}{\pi\eta}} \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \int_{\mathbb{R}^N} \mathbf{m} e^{-\frac{1}{2}D(\mathbf{m})} d\mathbf{m},\end{aligned} \tag{T5.100}$$

The term in the last exponential can be written as follows:

$$\begin{aligned} D(\mathbf{m}) &\equiv \mathbf{m}'\left(\frac{\boldsymbol{\alpha}\boldsymbol{\alpha}'}{\eta}\right)\mathbf{m}+(\mathbf{m}-\boldsymbol{\nu})'\,\boldsymbol{\Sigma}^{-1}(\mathbf{m}-\boldsymbol{\nu}) \\ &=(\mathbf{m}-\boldsymbol{\xi})'\,\boldsymbol{\Phi}^{-1}(\mathbf{m}-\boldsymbol{\xi})+\boldsymbol{\nu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu}-\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}, \end{aligned} \tag{T5.101}$$

where

$$\boldsymbol{\xi}\equiv\left[\frac{\boldsymbol{\alpha}\boldsymbol{\alpha}'}{\eta}+\boldsymbol{\Sigma}^{-1}\right]^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu},\quad \boldsymbol{\Phi}\equiv\left[\frac{\boldsymbol{\alpha}\boldsymbol{\alpha}'}{\eta}+\boldsymbol{\Sigma}^{-1}\right]^{-1}. \tag{T5.102}$$

Therefore from (T5.92) we obtain:

$$\begin{aligned} \partial_{\boldsymbol{\alpha}}\,\mathrm{CE}(\boldsymbol{\alpha}) &= \frac{\mathrm{E}\{u'(\Psi_{\boldsymbol{\alpha}})\,\partial_{\boldsymbol{\alpha}}(\Psi_{\boldsymbol{\alpha}})\}}{u'(\mathrm{CE}(\boldsymbol{\alpha}))} \\ &= \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}}e^{\frac{\eta}{2}\,\mathrm{CE}(\boldsymbol{\alpha})^2}\int_{\mathbb{R}^N}\mathbf{m}e^{-\frac{1}{2}D(\mathbf{m})}d\mathbf{m} \\ &= \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}}e^{\frac{\eta}{2}\,\mathrm{CE}(\boldsymbol{\alpha})^2}e^{-\frac{1}{2}\left[\boldsymbol{\nu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu}-\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}\right]}\frac{(2\pi)^{\frac{N}{2}}}{|\boldsymbol{\Phi}|^{-\frac{1}{2}}} \\ &\quad \int_{\mathbb{R}^N}\mathbf{m}\frac{|\boldsymbol{\Phi}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}}e^{-\frac{1}{2}(\mathbf{m}-\boldsymbol{\xi})'\boldsymbol{\Phi}^{-1}(\mathbf{m}-\boldsymbol{\xi})}d\mathbf{m} \\ &= \gamma(\boldsymbol{\alpha})\,\boldsymbol{\xi}, \end{aligned} \tag{T5.103}$$

where from (T5.102) we obtain:

$$\begin{aligned} \gamma(\boldsymbol{\alpha}) &\equiv \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{|\boldsymbol{\Phi}|^{-\frac{1}{2}}}e^{\frac{\eta}{2}\,\mathrm{CE}(\boldsymbol{\alpha})^2}e^{-\frac{1}{2}\left[\boldsymbol{\nu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu}-\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}\right]} \\ &= \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{\left|\frac{1}{\eta}\boldsymbol{\alpha}\boldsymbol{\alpha}'+\boldsymbol{\Sigma}^{-1}\right|^{\frac{1}{2}}}e^{\frac{\eta}{2}\,\mathrm{CE}(\boldsymbol{\alpha})^2}e^{-\frac{1}{2}\left[\boldsymbol{\nu}'\left(\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left[\frac{1}{\eta}\boldsymbol{\alpha}\boldsymbol{\alpha}'+\boldsymbol{\Sigma}^{-1}\right]^{-1}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\nu}\right]} \end{aligned} \tag{T5.104}$$

Using again (T5.102) and then (T5.99) we finally obtain:

$$\partial_{\boldsymbol{\alpha}}\,\mathrm{CE}(\boldsymbol{\alpha})=\gamma(\boldsymbol{\alpha})\left[\frac{1}{\eta}\boldsymbol{\alpha}\boldsymbol{\alpha}'+\boldsymbol{\Sigma}^{-1}\right]^{-1}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}-\mathbf{P}_T), \tag{T5.105}$$

Second-order sensitivity analysis

To compute the second derivative of the certainty-equivalent, first we derive the following result:

$$\begin{aligned}\partial_{\boldsymbol{\alpha}}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} &= \partial_{\boldsymbol{\alpha}}\left[\int_{\mathbb{R}^N} u'\left(\psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\psi_{\boldsymbol{\alpha}}\right) f_{\mathbf{M}}\left(\mathbf{m}\right) d\mathbf{m}\right] \\ &= \int_{\mathbb{R}^N} f_{\mathbf{M}}\left(\mathbf{m}\right)\left[u'\left(\psi_{\boldsymbol{\alpha}}\right)\partial^2_{\boldsymbol{\alpha}\boldsymbol{\alpha}'}\left(\psi_{\boldsymbol{\alpha}}\right)\right. \\ &\quad \left. + u''\left(\psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}}\left(\psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\psi_{\boldsymbol{\alpha}}\right)\right] d\mathbf{m} \\ &= \mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial^2_{\boldsymbol{\alpha}\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right) + u''\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}}\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\end{aligned} \tag{T5.106}$$

From this, ($T5.92$) and the chain rule of calculus we obtain:

$$\begin{aligned}\partial^2_{\boldsymbol{\alpha}\boldsymbol{\alpha}'}\,\mathrm{CE}\left(\boldsymbol{\alpha}\right) &= \partial_{\boldsymbol{\alpha}}\left[\frac{\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}\right] \\ &= \partial_{\boldsymbol{\alpha}}\left[\frac{1}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}\right]\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} \\ &\quad + \frac{1}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}\partial_{\boldsymbol{\alpha}}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} \\ &= \left[-\frac{u''\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}{\left[u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)\right]^2}\frac{\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}\right]\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\} \\ &\quad + \frac{\mathrm{E}\left\{u'\left(\Psi_{\boldsymbol{\alpha}}\right)\partial^2_{\boldsymbol{\alpha}\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right) + u''\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}}\left(\Psi_{\boldsymbol{\alpha}}\right)\partial_{\boldsymbol{\alpha}'}\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}\end{aligned} \tag{T5.107}$$

Since the market is linear in the allocation:

$$\Psi_{\boldsymbol{\alpha}} = \boldsymbol{\alpha}'\mathbf{M}, \tag{T5.108}$$

we obtain:

$$\partial^2_{\boldsymbol{\alpha}\boldsymbol{\alpha}'}\,\mathrm{CE}\left(\boldsymbol{\alpha}\right) = \frac{\mathrm{E}\left\{u''\left(\boldsymbol{\alpha}'\mathbf{M}\right)\mathbf{M}\mathbf{M}'\right\} - u''\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)\mathbf{w}\mathbf{w}'}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)} \tag{T5.109}$$

where

$$\mathbf{w} \equiv \mathrm{E}\left\{\frac{u'\left(\boldsymbol{\alpha}'\mathbf{M}\right)}{u'\left(\mathrm{CE}\left(\boldsymbol{\alpha}\right)\right)}\mathbf{M}\right\} \tag{T5.110}$$

The denominator in ($T5.109$) is always positive. On the other hand, the numerator in ($T5.109$) can take on any sign, depending on the local curvature of the utility function. Therefore the convexity of the certainty-equivalent is not determined.

5.4 Properties of the quantile-based index of satisfaction

Constancy

Assume that an allocation $\mathbf{b}$ gives rise to a deterministic objective $\psi_{\mathbf{b}}$. Then from ($B.22$) the probability density function of the objective is the Dirac delta centered at $\psi_{\mathbf{b}}$ and from $(B.53)$ the cumulative distribution function is the Heaviside function $(B.74)$ with step at $\psi_{\mathbf{b}}$, which is not invertible. Thus the quantile is not defined.

Nonetheless, we can obtain the quantile using the smoothing technique (1.20). Indeed from $(B.18)$ the regularized pdf of the objective reads:

$$f_{\Psi_{\mathbf{b}};\epsilon}(\psi) \equiv \left(\delta^{(\psi_{\mathbf{b}})} * \delta_{\epsilon}^{(0)}\right)(\psi) = \frac{1}{\sqrt{2\pi}\epsilon} e^{-\frac{(\psi-\psi_{\mathbf{b}})^2}{2\epsilon^2}}. \qquad (T5.111)$$

The respective regularized cdf reads:

$$\begin{aligned} F_{\Psi_{\mathbf{b}};\epsilon}(\psi) &= \frac{1}{\sqrt{2\pi}\epsilon} \int_{-\infty}^{\psi} e^{-\frac{(x-\psi_{\mathbf{b}})^2}{2\epsilon^2}} dx \qquad (T5.112) \\ &= \frac{1}{2}\left(\frac{2}{\sqrt{\pi}} \int_{-\infty}^{\frac{\psi-\psi_{\mathbf{b}}}{\sqrt{2}\epsilon}} e^{-y^2} dy\right) \\ &= \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\psi-\psi_{\mathbf{b}}}{\sqrt{2}\epsilon}\right)\right), \end{aligned}$$

where we used the change of variable $y \equiv (x - \psi_{\mathbf{b}})/\sqrt{2}\epsilon$.

The regularized quantile of the objective is the inverse of the regularized cumulative distribution function:

$$Q_{\Psi_{\mathbf{b}};\epsilon}(s) \equiv F_{\Psi_{\mathbf{b}};\epsilon}^{-1}(s) = \psi_{\mathbf{b}} + \sqrt{2}\epsilon \operatorname{erf}^{-1}(2s-1). \qquad (T5.113)$$

For small ϵ the regularized quantile satisfies:

$$Q_{\Psi_{\mathbf{b}};\epsilon}(s) \approx \psi_{\mathbf{b}}, \quad \text{for all } s \in (0,1), \qquad (T5.114)$$

and the approximation becomes exact in the limit $\epsilon \to 0$. Thus the quantile (5.159) satisfies:

$$\Psi_{\mathbf{b}} = \psi_{\mathbf{b}} \Rightarrow \mathrm{Q}_c(\mathbf{b}) \equiv Q_{\Psi_{\mathbf{b}}}(1-c) = \psi_{\mathbf{b}}, \qquad (T5.115)$$

which is the constancy property (5.62) in this context.

Homogeneity, translation invariance, additive co-monotonicity

Consider the s-quantile of the variable X, defined implicitly as in (1.18) by

$$\mathbb{P}\{X \leq Q_X\} = s. \qquad (T5.116)$$

If h is an increasing function then $(T5.116)$ is equivalent to the following identity:

$$\mathbb{P}\{h(X) \leq h(Q_X)\} = s. \qquad (T5.117)$$

On the other hand, the s-quantile $Q_{h(X)}(s)$ of the variable $h(X)$ is defined implicitly by:

$$\mathbb{P}\left\{h(X) \leq Q_{h(X)}\right\} = s \tag{T5.118}$$

Since (*T*5.117) and (*T*5.118) hold for any s we obtain the general result for any increasing function h:

$$Q_{h(X)}(s) = h\left(Q_X(s)\right) \tag{T5.119}$$

As special cases, consider $h(X) \equiv \lambda X$, where $\lambda > 0$. Then (*T*5.119) implies

$$Q_{\lambda X}(s) = \lambda Q_X(s). \tag{T5.120}$$

Now consider $h(X) \equiv X + \lambda$. Then (*T*5.119) implies

$$Q_{X+\lambda}(s) = Q_X(s) + \lambda. \tag{T5.121}$$

Finally consider $h(X) \equiv X + g(X)$, where g is an increasing function Then applying repeatedly (*T*5.119) we obtain:

$$Q_{X+g(X)}(s) = Q_X(s) + g\left(Q_X(s)\right) = Q_X(s) + Q_{g(X)}(s). \tag{T5.122}$$

Expression (*T*5.120) and the positive homogeneity of the objective (5.16) prove the positive homogeneity of the quantile-based index of satisfaction:

$$\begin{aligned} \mathrm{Q}_c(\lambda\boldsymbol{\alpha}) &\equiv Q_{\Psi_{\lambda\boldsymbol{\alpha}}}(1-c) = Q_{\lambda\Psi_{\boldsymbol{\alpha}}}(1-c) \\ &= \lambda Q_{\Psi_{\boldsymbol{\alpha}}}(1-c) \equiv \lambda\,\mathrm{Q}_c(\boldsymbol{\alpha}). \end{aligned} \tag{T5.123}$$

Expression (*T*5.121) and the additivity of the objective (5.17) prove the translation-invariance of the quantile-based index of satisfaction:

$$\begin{aligned} \mathrm{Q}_c(\boldsymbol{\alpha} + \lambda\mathbf{b}) &\equiv Q_{\Psi_{\boldsymbol{\alpha}+\lambda\mathbf{b}}}(1-c) = Q_{\Psi_{\boldsymbol{\alpha}}+\lambda}(1-c) \\ &= Q_{\Psi_{\boldsymbol{\alpha}}}(1-c) + \lambda \equiv \mathrm{Q}_c(\boldsymbol{\alpha}) + \lambda. \end{aligned} \tag{T5.124}$$

Expression (*T*5.122) and the additivity of the objective (5.17) prove the additive co-monotonicity of the quantile-based index of satisfaction:

$$\begin{aligned} \mathrm{Q}_c(\boldsymbol{\alpha} + \boldsymbol{\delta}) &\equiv Q_{\Psi_{\boldsymbol{\alpha}+\boldsymbol{\delta}}}(1-c) = Q_{\Psi_{\boldsymbol{\alpha}}+\Psi_{\boldsymbol{\delta}}}(1-c) \\ &= Q_{\Psi_{\boldsymbol{\alpha}}}(1-c) + Q_{\Psi_{\boldsymbol{\delta}}}(1-c) \\ &\equiv \mathrm{Q}_c(\boldsymbol{\alpha}) + \mathrm{Q}_c(\boldsymbol{\delta}). \end{aligned} \tag{T5.125}$$

Cornish-Fisher expansion

The Cornish-Fisher expansion (5.179) states that the quantile of the objective $\Psi_{\boldsymbol{\alpha}}$ can be approximated in terms of the quantile $z(s)$ of the standard normal distribution and the first three moments as follows:

$$\begin{aligned} Q_{\Psi_{\boldsymbol{\alpha}}}(s) \approx &\left[\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} - \frac{\mathrm{CM}_3\{\Psi_{\boldsymbol{\alpha}}\}}{6\,\mathrm{Var}\{\Psi_{\boldsymbol{\alpha}}\}}\right] \\ &+ \mathrm{Sd}\{\Psi_{\boldsymbol{\alpha}}\}\,z(s) + \frac{\mathrm{CM}_3\{\Psi_{\boldsymbol{\alpha}}\}}{6\,\mathrm{Var}\{\Psi_{\boldsymbol{\alpha}}\}} z^2(s), \end{aligned} \tag{T5.126}$$

where CM_3 is the third central moment. Using ($T1.39$) to express the central moments in terms of the raw moments we obtain the approximate expression of the quantile of the objective:

$$\begin{aligned} \mathrm{Q}_c(\boldsymbol{\alpha}) &\equiv Q_{\Psi_{\boldsymbol{\alpha}}}(1-c) \approx Q_{\Psi_{\boldsymbol{\alpha}}}(1-c) \\ &\approx A(\boldsymbol{\alpha}) + B(\boldsymbol{\alpha})\, z(1-c) + C(\boldsymbol{\alpha})\, z^2(1-c) \end{aligned} \tag{T5.127}$$

where

$$\begin{aligned} A &\equiv \mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} - \frac{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}^3\right\} - 3\,\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}^2\right\}\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} + 2\,\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}^3}{6\left(\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}^2\} - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}^2\right)} \\ B &\equiv \sqrt{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}^2\} - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}^2} \\ C &\equiv \frac{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}^3\right\} - 3\,\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}^2\right\}\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} + 2\,\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}^3}{6\left(\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}^2\} - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}^2\right)} \end{aligned} \tag{T5.128}$$

To obtain the explicit analytical expression of these coefficients as functions of the allocation $\boldsymbol{\alpha}$ we use the derivatives of the characteristic function of the objective as discussed in Appendix www.5.1.

First-order sensitivity analysis

The following proof is adapted from Gourieroux, Laurent, and Scaillet (2000). From the definition of quantile (1.18), the quantile-based index of satisfaction (5.159) is defined implicitly as follows:

$$1 - c = \mathbb{P}\{\Psi_{\boldsymbol{\alpha}} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\} = \mathbb{P}\{\boldsymbol{\alpha}'\mathbf{M} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\} \tag{T5.129}$$

Defining

$$X_n \equiv \sum_{j \neq n} \alpha_j M_j \tag{T5.130}$$

we see that $\mathrm{Q}(\boldsymbol{\alpha})$ is defined implicitly as follows in terms of the joint pdf f of (X_n, M_n):

$$\begin{aligned} 1 - c &= \mathbb{P}\{X_n + a_n M_n \leq \mathrm{Q}\} \\ &= \int \left[\int_{-\infty}^{\mathrm{Q} - \alpha_n m_n} f(x_n, m_n)\, dx_n\right] dm_n. \end{aligned} \tag{T5.131}$$

Since in general

$$\begin{aligned} \frac{\partial}{\partial a}\int_{-\infty}^{g(a)} f(x)\, dx &= \lim_{\delta a \to 0} \frac{1}{\delta a}\int_{g(a)}^{g(a) + \frac{dg}{da}\delta a} f(x)\, dx \\ &= f(g(a))\, \frac{dg(a)}{da}, \end{aligned} \tag{T5.132}$$

differentiating both sides of (T5.131) with respect to α_n we obtain:

$$0=\int f\left(\mathrm{Q}-\alpha_n M_n, m_n\right)\left(\frac{\partial\,\mathrm{Q}}{\partial \alpha_n}-m_n\right) dm_n \tag{T5.133}$$

or

$$\begin{aligned}\frac{\partial\,\mathrm{Q}}{\partial \alpha_n} &= \frac{\int m_n f\left(\mathrm{Q}-\alpha_n M_n, m_n\right) dm_n}{\int f\left(\mathrm{Q}-\alpha_n M_n, m_n\right) dm_n} \\ &= \mathrm{E}\left\{M_n | X_n=\mathrm{Q}(\boldsymbol{\alpha})-\alpha_n M_n\right\}\end{aligned} \tag{T5.134}$$

Therefore, substituting back the definition (T5.130) we obtain

$$\frac{\partial\,\mathrm{Q}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}=\mathrm{E}\left\{\mathbf{M} | \boldsymbol{\alpha}'\mathbf{M}=\mathrm{Q}(\boldsymbol{\alpha})\right\} \tag{T5.135}$$

Second-order sensitivity analysis

Consider now a small perturbation of the allocation $\boldsymbol{\alpha}$ in the direction of the j-th security:

$$\boldsymbol{\beta}=\boldsymbol{\alpha}+\epsilon\boldsymbol{\delta}^{(j)}. \tag{T5.136}$$

We derive the second derivatives from the definition

$$\partial_{ij}^2\,\mathrm{Q}(\boldsymbol{\alpha}) \equiv \lim_{\epsilon \to 0} \frac{1}{\epsilon}\left[\partial_i\,\mathrm{Q}(\boldsymbol{\beta})-\partial_i\,\mathrm{Q}(\boldsymbol{\alpha})\right]. \tag{T5.137}$$

From (T5.135) and (T5.136) we see that

$$\begin{aligned}\partial_i\,\mathrm{Q}(\boldsymbol{\beta}) &= \mathrm{E}\left\{M_i | \boldsymbol{\alpha}'\mathbf{M}+\epsilon M_j=\mathrm{Q}\left(\boldsymbol{\alpha}+\epsilon\boldsymbol{\delta}^{(j)}\right)\right\} \\ &\approx \mathrm{E}\left\{M_i | \boldsymbol{\alpha}'\mathbf{M}+\epsilon M_j=\mathrm{Q}(\boldsymbol{\alpha})+\epsilon\partial_j\,\mathrm{Q}(\boldsymbol{\alpha})\right\} \\ &= \mathrm{E}\left\{M_i | \boldsymbol{\alpha}'\mathbf{M}-\mathrm{Q}(\boldsymbol{\alpha})+\epsilon\left[M_j-\partial_j\,\mathrm{Q}(\boldsymbol{\alpha})\right]=0\right\} \\ &= \mathrm{E}\left\{M_i | \boldsymbol{\alpha}'\mathbf{M}-\mathrm{Q}(\boldsymbol{\alpha})+\epsilon\left[M_j-\mathrm{E}\left\{M_j | \boldsymbol{\alpha}'\mathbf{M}=\mathrm{Q}(\boldsymbol{\alpha})\right\}\right]=0\right\}\end{aligned} \tag{T5.138}$$

In other words, defining the variables

$$Z \equiv \boldsymbol{\alpha}'\mathbf{M}-\mathrm{Q}(\boldsymbol{\alpha}), \qquad Y \equiv M_j-\mathrm{E}\left\{M_j | Z=0\right\}, \qquad X \equiv M_i \tag{T5.139}$$

we can write:

$$\partial_i\,\mathrm{Q}(\boldsymbol{\beta}) \approx \mathrm{E}\left\{X | Z+\epsilon Y=0\right\}. \tag{T5.140}$$

Also notice that in this notation

$$\partial_i\,\mathrm{Q}(\boldsymbol{\alpha})=\mathrm{E}\left\{X | Z=0\right\} \tag{T5.141}$$

Consider the joint pdf f of (X, Y, Z). The conditional expectation in (T5.140) can be computed in terms of the conditional pdf, which reads:

$$f\left(x,y|Z+\epsilon Y=0\right)=f\left(x,y|z=-\epsilon y\right)=\frac{f\left(x,y,-\epsilon y\right)}{\int f\left(x,y,-\epsilon y\right)dxdy}. \qquad (T5.142)$$

Therefore ($T5.140$) becomes:

$$\begin{aligned}
\partial_i \,\mathrm{Q}\left(\boldsymbol{\beta}\right) &\approx \frac{\int\int xf\left(x,y,-\epsilon y\right)dxdy}{\int\int f\left(x,y,-\epsilon y\right)dxdy} \qquad (T5.143)\\
&\approx \frac{\int\int xf\left(x,y,0\right)dxdy-\epsilon\int\int xy\partial_z f\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy-\epsilon\int\int y\partial_z f\left(x,y,0\right)dxdy}\\
&= \left[\int\int xf\left(x,y,0\right)dxdy-\epsilon\int\int xy\partial_z f\left(x,y,0\right)dxdy\right]\\
&\quad \left[\left(\int\int f\left(x,y,0\right)dxdy\right)\left(1-\epsilon\frac{\int\int y\partial_z f\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy}\right)\right]^{-1}\\
&\approx \left[\int\int xf\left(x,y,0\right)dxdy-\epsilon\int\int xy\partial_z\left[\ln f\left(x,y,0\right)\right]f\left(x,y,0\right)dxdy\right]\\
&\quad \left(\int\int f\left(x,y,0\right)dxdy\right)^{-1}\\
&\quad \left[1+\epsilon\frac{\int\int y\partial_z\left[\ln f\left(x,y,0\right)\right]f\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy}\right]\\
&\approx \frac{\int\int xf\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy}-\epsilon\frac{\int\int xy\partial_z\left[\ln f\left(x,y,0\right)\right]f\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy}\\
&\quad +\epsilon\frac{\int\int y\partial_z\left[\ln f\left(x,y,0\right)\right]f\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy}\frac{\int\int xf\left(x,y,0\right)dxdy}{\int\int f\left(x,y,0\right)dxdy}
\end{aligned}$$

Thus

$$\begin{aligned}
\partial_i \,\mathrm{Q}\left(\boldsymbol{\beta}\right) &\approx \mathrm{E}\left\{X|Z=0\right\}-\epsilon\left[\mathrm{E}\left\{XY\partial_z\left[\ln f\left(X,Y,0\right)\right]|Z=0\right\}\right.\\
&\quad \left.-\,\mathrm{E}\left\{X|Z=0\right\}\mathrm{E}\left\{Y\partial_z\left[\ln f\left(X,Y,0\right)\right]|Z=0\right\}\right]\\
&= \mathrm{E}\left\{X|Z=0\right\}-\epsilon\left[\mathrm{Cov}\left\{X,Y\partial_z\left[\ln f\left(X,Y,0\right)\right]|Z=0\right\}\right] \qquad (T5.144)\\
&= \mathrm{E}\left\{X|Z=0\right\}-\epsilon\left[\mathrm{Cov}\left\{X,Y\partial_z\left[\ln f\left(X,Y|0\right)+\ln f_Z\left(0\right)\right]|Z=0\right\}\right]\\
&= \mathrm{E}\left\{X|Z=0\right\}-\epsilon\left[\mathrm{Cov}\left\{X,Y\partial_z\left[\ln f\left(X,Y|0\right)\right]|Z=0\right\}\right.\\
&\quad \left.+\,\partial_z\left[\ln f_Z\left(0\right)\right]\mathrm{Cov}\left\{X,Y|Z=0\right\}\right] \qquad (T5.145)
\end{aligned}$$

On the other hand

$$\begin{aligned}
\partial_z \left[\mathrm{Cov}\left\{ X, Y | Z = z \right\} \right] &= \partial_z \left[\mathrm{E}\left\{ XY | Z = z \right\} - \mathrm{E}\left\{ X | Z = z \right\} \mathrm{E}\left\{ Y | Z = z \right\} \right] \\
&= \partial_z \left[\int \int xyf\left(x, y | z\right) dxdy \right. \qquad (T5.146) \\
&\quad \left. - \int \int xf\left(x, y | z\right) dxdy \int \int yf\left(x, y | z\right) dxdy \right] \\
&= \int \int xy \partial_z f\left(x, y | z\right) dxdy \\
&\quad - \mathrm{E}\left\{ Y | Z = z \right\} \partial_z \left[\mathrm{E}\left\{ X | Z = z \right\} \right] \\
&\quad - \mathrm{E}\left\{ X | Z = z \right\} \int \int y \partial_z f\left(x, y | z\right) dxdy \\
&= \int \int xy \partial_z \left[\ln f\left(x, y | z\right) \right] f\left(x, y | z\right) dxdy \\
&\quad - \mathrm{E}\left\{ Y | Z = z \right\} \partial_z \left[\mathrm{E}\left\{ X | Z = z \right\} \right] \\
&\quad - \mathrm{E}\left\{ X | Z = z \right\} \int \int y \partial_z \left[\ln f\left(x, y | z\right) \right] f\left(x, y | z\right) dxdy \\
&= \mathrm{E}\left\{ XY \partial_z \left[\ln f\left(X, Y | z\right) \right] | Z = z \right\} \\
&\quad - \mathrm{E}\left\{ X | Z = z \right\} \mathrm{E}\left\{ Y \partial_z \left[\ln f\left(X, Y | z\right) \right] | Z = z \right\} \\
&\quad - \mathrm{E}\left\{ Y | Z = z \right\} \partial_z \left[\mathrm{E}\left\{ X | Z = z \right\} \right] \\
&= \mathrm{Cov}\left\{ X, Y \partial_z \left[\ln f\left(X, Y | z\right) \right] | Z = z \right\} \\
&\quad - \mathrm{E}\left\{ Y | Z = z \right\} \partial_z \left[\mathrm{E}\left\{ X | Z = z \right\} \right],
\end{aligned}$$

which shows that

$$\begin{aligned}
\mathrm{Cov}\left\{ X, Y \partial_z \left[\ln f\left(X, Y | z\right) \right] | Z = z \right\} = \partial_z \left[\mathrm{Cov}\left\{ X, Y | Z = z \right\} \right] \qquad (T5.147) \\
+ \mathrm{E}\left\{ Y | Z = z \right\} \partial_z \left[\mathrm{E}\left\{ X | Z = z \right\} \right]
\end{aligned}$$

Therefore (T5.144) becomes

$$\begin{aligned}
\partial_i \mathrm{Q}\left(\boldsymbol{\beta}\right) \approx\ & \mathrm{E}\left\{ X | Z = 0 \right\} \qquad (T5.148) \\
& - \epsilon \left[\partial_z \left[\mathrm{Cov}\left\{ X, Y | Z = 0 \right\} \right] + \mathrm{E}\left\{ Y | Z = 0 \right\} \partial_z \left[\mathrm{E}\left\{ X | Z = 0 \right\} \right] \right. \\
& \left. + \partial_z \left[\ln f_Z\left(0\right) \right] \mathrm{Cov}\left\{ X, Y | Z = 0 \right\} \right]
\end{aligned}$$

In this expression, from (T5.139)

$$\mathrm{E}\left\{ Y | Z = 0 \right\} = \mathrm{E}\left\{ M_j - \mathrm{E}\left\{ M_j | Z = 0 \right\} | Z = 0 \right\} = 0 \qquad (T5.149)$$

Therefore

$$\begin{aligned}
\partial_i \mathrm{Q}\left(\boldsymbol{\beta}\right) \approx \mathrm{E}\left\{ X | Z = 0 \right\} - \epsilon \left[\partial_z \left[\mathrm{Cov}\left\{ X, Y | z = 0 \right\} \right] \right. \qquad (T5.150) \\
\left. + \partial_z \left[\ln f_Z\left(0\right) \right] \mathrm{Cov}\left\{ X, Y | z = 0 \right\} \right]
\end{aligned}$$

and finally, from (T5.141) we obtain

$$\begin{aligned}\partial^2_{ij}\,\mathrm{Q}\left(\boldsymbol{\alpha}\right) \equiv & -\partial_z\left[\mathrm{Cov}\left\{X,Y|z=0\right\}\right] \\ & -\partial_z\left[\ln f_Z\left(0\right)\right]\mathrm{Cov}\left\{X,Y|z=0\right\}\end{aligned} \quad (T5.151)$$

Substituting again the definitions ($T5.139$) in this formula we obtain:

$$\begin{aligned}\frac{\partial^2\,\mathrm{Q}\left(\boldsymbol{\alpha}\right)}{\partial\boldsymbol{\alpha}'\partial\boldsymbol{\alpha}} &= -\partial_z\left[\mathrm{Cov}\left\{\mathbf{M},\mathbf{M}-\mathrm{E}\left\{\mathbf{M}|\boldsymbol{\alpha}'\mathbf{M}=\mathrm{Q}\left(\boldsymbol{\alpha}\right)\right\}|z=0\right\}\right] \\ &\quad -\frac{\partial\ln f_Z\left(0\right)}{\partial z}\mathrm{Cov}\left\{\mathbf{M},\mathbf{M}-\mathrm{E}\left\{\mathbf{M}|\boldsymbol{\alpha}'\mathbf{M}=\mathrm{Q}\left(\boldsymbol{\alpha}\right)\right\}|z=0\right\} \\ &= -\partial_z\left[\mathrm{Cov}\left\{\mathbf{M}|z=0\right\}\right]-\frac{\partial\ln f_Z\left(0\right)}{\partial z}\mathrm{Cov}\left\{\mathbf{M}|z=0\right\} \\ &= -\left.\frac{\partial\,\mathrm{Cov}\left\{\mathbf{M}|\boldsymbol{\alpha}'\mathbf{M}=z\right\}}{\partial z}\right|_{z=\mathrm{Q}(\boldsymbol{\alpha})} \\ &\quad -\frac{\partial\ln f_{\boldsymbol{\alpha}'\mathbf{M}}\left(\mathrm{Q}\left(\boldsymbol{\alpha}\right)\right)}{\partial z}\mathrm{Cov}\left\{\mathbf{M}|\boldsymbol{\alpha}'\mathbf{M}=\mathrm{Q}\left(\boldsymbol{\alpha}\right)\right\}\end{aligned} \quad (T5.152)$$

To discuss the sign of the second derivative in the normal case we need the following result:

$$\begin{aligned}\boldsymbol{\Sigma}\left(\mathbf{I}-\frac{\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Sigma}}{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}}\right)\geq 0 &\Leftrightarrow \boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\beta}\geq\frac{\boldsymbol{\beta}'\boldsymbol{\Sigma}\boldsymbol{\alpha}\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\beta}}{\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}} \\ &\Leftrightarrow \left\langle\boldsymbol{\alpha},\boldsymbol{\alpha}\right\rangle\left\langle\boldsymbol{\beta},\boldsymbol{\beta}\right\rangle\geq\left|\left\langle\boldsymbol{\alpha},\boldsymbol{\beta}\right\rangle\right|^2\end{aligned} \quad (T5.153)$$

where

$$\left\langle\boldsymbol{\alpha},\boldsymbol{\beta}\right\rangle\equiv\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\beta}. \quad (T5.154)$$

The last row in ($T5.153$) is true because of the Cauchy-Schwartz inequality ($A.8$).

5.5 Properties of spectral indices of satisfaction

Spectral representation

We consider weighted averages of the expected shortfall for different confidence levels. From the definition (5.207) of expected shortfall this means

$$\begin{aligned}\mathrm{Spc}\left(\boldsymbol{\alpha}\right) &\equiv \int_0^1 \mathrm{ES}_c\left(\boldsymbol{\alpha}\right)w\left(c\right)dc \\ &= \int_0^1 \frac{1}{1-c}\left[\int_0^{1-c} Q_{\Psi_{\boldsymbol{\alpha}}}\left(s\right)ds\right]w\left(c\right)dc,\end{aligned} \quad (T5.155)$$

where

$$w\left(c\right)\geq 0, \qquad \int_0^1 w\left(c\right)dc = 1. \quad (T5.156)$$

Equivalently:

$$\begin{aligned}\int_0^1 \mathrm{ES}_c(\boldsymbol{\alpha})\, w(c)\, dc &= \int_0^1 \left[\int_0^c Q_{\Psi_{\boldsymbol{\alpha}}}(s)\, ds\right] \frac{w(c)}{c} dc \\ & \int_0^1 \left[\int_0^c Q_{\Psi_{\boldsymbol{\alpha}}}(s)\, \frac{w(c)}{c} ds\right] dc \\ & \int_0^1 Q_{\Psi_{\boldsymbol{\alpha}}}(s) \left[\int_s^1 \frac{w(c)}{c} dc\right] ds \\ & \int_0^1 Q_{\Psi_{\boldsymbol{\alpha}}}(s)\, \phi(s)\, ds,\end{aligned} \qquad (T5.157)$$

where

$$\phi(s) \equiv \int_s^1 \frac{w(x)}{x} dx \qquad (T5.158)$$

On the one hand, from

$$\phi'(s) = -\frac{w(s)}{s} \qquad (T5.159)$$

we obtain:

$$\begin{aligned}\int_0^1 [s\phi(s)]'\, ds &= \int_0^1 \phi(s)\, ds + \int_0^1 s\phi' ds \\ &= \int_0^1 \phi(s)\, ds - \int_0^1 w(s)\, ds.\end{aligned} \qquad (T5.160)$$

On the other hand, from $(T5.158)$ we obtain $\phi(1) = 0$ and thus

$$\int_0^1 [s\phi(s)]'\, ds = \left. s\phi(s)\right|_0^1 = 0. \qquad (T5.161)$$

Therefore

$$\int_0^1 \phi(s)\, ds = 1. \qquad (T5.162)$$

Finally, from $(T5.159)$ we also obtain $\phi' \leq 0$.

Spectral indices of satisfaction and risk aversion

Consider a fair game, i.e. an allocation $\mathbf{f}$ such that

$$\mathrm{E}\{\Psi_{\mathbf{f}}\} = 0 \qquad (T5.163)$$

Then a fortiori, for any $s \in (0,1)$ the following is true:

$$\mathrm{E}\{\Psi_{\mathbf{f}} | \Psi_{\mathbf{f}} \leq Q_{\Psi_{\mathbf{f}}}(s)\} \leq 0 \qquad (T5.164)$$

Thus from the definition of expected shortfall (5.208) for any confidence level $\mathrm{ES}_c(\boldsymbol{\alpha}) \leq 0$. Since the expected shortfall generates the spectral indices of satisfaction, the satisfaction derived from any fair game is negative whenever satisfaction is measured with a spectral index.

Cornish-Fisher expansion

The Cornish-Fisher expansion (5.179) states that the quantile of the approximate objective $\Xi_{\boldsymbol{\alpha}}$ can be approximated in terms of the quantile $z(s)$ of the standard normal distribution and the first three moments as follows:

$$Q_{\Xi_{\boldsymbol{\alpha}}}(s) \approx \left[\mathrm{E}\{\Xi_{\boldsymbol{\alpha}}\} - \frac{\mathrm{CM}_3\{\Xi_{\boldsymbol{\alpha}}\}}{6\,\mathrm{Var}\{\Xi_{\boldsymbol{\alpha}}\}}\right] + \mathrm{Sd}\{\Xi_{\boldsymbol{\alpha}}\}\, z(s) + \frac{\mathrm{CM}_3\{\Xi_{\boldsymbol{\alpha}}\}}{6\,\mathrm{Var}\{\Xi_{\boldsymbol{\alpha}}\}} z^2(s), \tag{T5.165}$$

where CM_3 is the third central moment. Using (T1.39) to express the central moments in terms of the raw moments we obtain the approximate expression of the quantile of the objective:

$$Q_{\Psi_{\boldsymbol{\alpha}}}(s) \approx Q_{\Xi_{\boldsymbol{\alpha}}}(s) \approx A(\boldsymbol{\alpha}) + B(\boldsymbol{\alpha})\, z(s) + C(\boldsymbol{\alpha})\, z^2(s) \tag{T5.166}$$

where (A, B, C) are defined in (5.181). To obtain the explicit analytical expression of these coefficients as functions of the allocation $\boldsymbol{\alpha}$ we use the derivatives of the characteristic function of the objective as discussed in Appendix www.5.1. To obtain the spectral index of satisfaction we apply (T5.166) to its definition (5.223), obtaining:

$$\begin{aligned}\mathrm{Spc}_{\phi}(\boldsymbol{\alpha}) &\equiv \int_0^1 \phi(s)\, Q_{\Psi_{\boldsymbol{\alpha}}}(s)\, ds \\ &\approx A(\boldsymbol{\alpha}) + B(\boldsymbol{\alpha}) \int_0^1 \phi(s)\, z(s)\, ds + C(\boldsymbol{\alpha}) \int_0^1 \phi(s)\, z^2(s)\, ds.\end{aligned} \tag{T5.167}$$

Extreme value theory

Define the variable

$$Z \equiv \mathrm{Q}_c(\boldsymbol{\alpha}) - \Psi_{\boldsymbol{\alpha}}. \tag{T5.168}$$

and (5.182) we obtain

$$\begin{aligned}1 - L_{\mathrm{Q}_c(\boldsymbol{\alpha})}(z) &= 1 - \mathbb{P}\{\Psi_{\boldsymbol{\alpha}} - \mathrm{Q}_c(\boldsymbol{\alpha}) \leq -z | \Psi_{\boldsymbol{\alpha}} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\} \\ &= 1 - \mathbb{P}\{\mathrm{Q}_c(\boldsymbol{\alpha}) - \Psi_{\boldsymbol{\alpha}} \geq z | \Psi_{\boldsymbol{\alpha}} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\} \\ &= \mathbb{P}\{\mathrm{Q}_c(\boldsymbol{\alpha}) - \Psi_{\boldsymbol{\alpha}} \leq z | \Psi_{\boldsymbol{\alpha}} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\} \\ &= \mathbb{P}\{Z \leq z | Z \geq 0\}.\end{aligned} \tag{T5.169}$$

This is the cdf of Z conditioned on $Z \geq 0$. If the confidence level c is high, from (5.184) this cdf is approximated by $G_{\xi,v}$. Thus

$$\mathrm{E}\{Z | Z \geq 0\} \approx \int_0^{\infty} z \frac{dG_{\xi,v}(z)}{dz} dz = \frac{v}{1-\xi}, \tag{T5.170}$$

where the last result can be found in Embrechts, Klueppelberg, and Mikosch (1997). On the other hand, from the definition (5.208) of expected shortfall we derive

$$\begin{aligned}\mathrm{ES}_c(\boldsymbol{\alpha}) &= \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}} | \Psi_{\boldsymbol{\alpha}} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\right\} \\ &= \mathrm{Q}_c(\boldsymbol{\alpha}) + \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}} - \mathrm{Q}_c(\boldsymbol{\alpha}) | \Psi_{\boldsymbol{\alpha}} \leq \mathrm{Q}_c(\boldsymbol{\alpha})\right\} \\ &= \mathrm{Q}_c(\boldsymbol{\alpha}) - \mathrm{E}\left\{Z | Z \geq 0\right\}\end{aligned} \qquad (T5.171)$$

Therefore and the result follows.

First-order sensitivity analysis

We compute the expression of the first derivative of the expected shortfall. The result for a generic spectral measure follows from (T5.155) and the definition (T5.159) of the weights in terms of the spectrum. Here we adapt from Bertsimas, Lauprete, and Samarov (2004). First we define:

$$X_n \equiv \sum_{j \neq n} \alpha_j M_j. \qquad (T5.172)$$

From the definition (5.208) of expected shortfall we obtain

$$\begin{aligned}\frac{\partial\, \mathrm{ES}_c(\boldsymbol{\alpha})}{\partial \alpha_n} &= \frac{\partial}{\partial \alpha_n}\left[\frac{1}{1-c}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x + \alpha_n m)\, \mathbb{I}_{x \leq \mathrm{Q}_c - \alpha_n m}(x, m)\right. \\ &\quad \left. f_{X_n, M_n}(x, m)\, dx dm\right] \qquad (T5.173) \\ &= \frac{1}{1-c}\int_{-\infty}^{+\infty}\frac{\partial}{\partial \alpha_n}\int_{-\infty}^{\mathrm{Q}_c - \alpha_n m}(x + \alpha_n m) f_{X_n, M_n}(x, m)\, dx dm\end{aligned}$$

Using (T5.132) this becomes

$$\begin{aligned}\frac{\partial\, \mathrm{ES}_c(\boldsymbol{\alpha})}{\partial \alpha_n} = \frac{1}{1-c}\int_{-\infty}^{+\infty} &\left(\frac{\partial\, \mathrm{Q}_c(\boldsymbol{\alpha})}{\partial \alpha_n} - m\right) \qquad (T5.174) \\ &\mathrm{Q}_c(\boldsymbol{\alpha}) f_{X_n, M_n}(\mathrm{Q}_c(\boldsymbol{\alpha}) - \alpha_n m, m)\, dm \\ &+ \frac{1}{1-c}\int_{-\infty}^{+\infty}\int_{-\infty}^{\mathrm{Q}_c - \alpha_n m} m f_{X_n, M_n}(x, m)\, dx dm\end{aligned}$$

On the other hand:

$$\begin{aligned}0 = \frac{\partial(1-c)}{\partial \alpha_n} &= \frac{\partial}{\partial \alpha_n}\int_{-\infty}^{+\infty}\int_{-\infty}^{\mathrm{Q}_c - \alpha_n m} f_{X_n, M_n}(x, m)\, dx dm \qquad (T5.175) \\ &\int_{-\infty}^{+\infty}\left(\frac{\partial\, \mathrm{Q}_c(\boldsymbol{\alpha})}{\partial \alpha_n} - m\right) f_{X_n, M_n}(\mathrm{Q}_c(\boldsymbol{\alpha}) - \alpha_n m, m)\, dm.\end{aligned}$$

Therefore

$$\frac{\partial \operatorname{ES}_c(\boldsymbol{\alpha})}{\partial \alpha_n} = \frac{1}{1-c}\int_{-\infty}^{+\infty}\int_{-\infty}^{\mathrm{Q}_c-\alpha_n m} m f_{X_n,M_n}(x,m)\,dxdm \quad (T5.176)$$
$$= \frac{1}{1-c}\int\int_{\Psi_{\boldsymbol{\alpha}}\leq \mathrm{Q}_c} m f_{X_n,M_n}(x,m)\,dxdm$$

which is the desired result.

Second-order sensitivity analysis

Now we compute the expression of the second derivative of the expected shortfall. The result for a generic spectral measure follows from ($T5.155$) and the definition ($T5.159$) of the weights in terms of the spectrum. We adapt the proof from Rau-Bredow (2002). In the notation ($T5.172$) the second derivative is:

$$\frac{\partial^2 \operatorname{ES}_c(\boldsymbol{\alpha})}{\partial \alpha_n \partial \boldsymbol{\alpha}'} = \frac{\partial}{\partial \alpha_n}\operatorname{E}\{\mathbf{M}'|X_n+\alpha_n M_n \leq \mathrm{Q}_c(\boldsymbol{\alpha})\} \quad (T5.177)$$
$$= \frac{\partial}{\partial \alpha_n}\left[\int \mathbf{m}' f_{\mathbf{M}}(\mathbf{m}|X_n+\alpha_n M_n \leq \mathrm{Q}_c(\boldsymbol{\alpha}))\,d\mathbf{m}\right].$$

In this expression the conditional density reads:

$$f_{\mathbf{M}}(\mathbf{m}|X_n+\alpha_n M_n \leq \mathrm{Q}_c) = \frac{\int_{-\infty}^{\mathrm{Q}_c-\alpha_n m_n} f_{X_n,\mathbf{M}}(x,\mathbf{m})\,dx}{\int\int_{-\infty}^{\mathrm{Q}_c-\alpha_n m_n} f_{X_n,\mathbf{M}}(x,\mathbf{m})\,dxd\mathbf{m}} \quad (T5.178)$$
$$= \frac{\int_{-\infty}^{\mathrm{Q}_c-\alpha_n m_n} f_{X_n,\mathbf{M}}(x,\mathbf{m})\,dx}{\int_{x_n+\alpha_n m_n\leq \mathrm{Q}_c} f_{X_n,\mathbf{M}}(x,\mathbf{m})\,dxd\mathbf{m}}$$
$$= \frac{\int_{-\infty}^{\mathrm{Q}_c-\alpha_n m_n} f_{X_n,\mathbf{M}}(x,\mathbf{m})\,dx}{1-c}.$$

Therefore using ($T5.132$) we obtain:

$$\frac{\partial^2 \operatorname{ES}_c(\boldsymbol{\alpha})}{\partial \alpha_n \partial \boldsymbol{\alpha}'} = \frac{1}{1-c}\left[\int \mathbf{m}'\frac{\partial}{\partial \alpha_n}\int_{-\infty}^{\mathrm{Q}_c-\alpha_n m_n} f_{X_n,\mathbf{M}}(x,\mathbf{m})\,dxd\mathbf{m}\right] \quad (T5.179)$$
$$= \frac{1}{1-c}\left[\int \mathbf{m}'\left[\frac{\partial \mathrm{Q}_c(\boldsymbol{\alpha})}{\partial \alpha_n} - m_n\right] f_{X_n,\mathbf{M}}(\mathrm{Q}_c(\boldsymbol{\alpha})-\alpha_n m_n,\mathbf{m})\,d\mathbf{m}\right]$$

Using Bayes' rule

$$f_{X_n,\mathbf{M}}(\mathrm{Q}-\alpha_n m_n,\mathbf{m}) = f_{X_n+\alpha_n m_n,\mathbf{M}}(\mathrm{Q},\mathbf{m}) \quad (T5.180)$$
$$= f_{\mathbf{M}|X_n+\alpha_n m_n}(\mathbf{m}|X_n+\alpha_n m_n=\mathrm{Q})\, f_{X_n+\alpha_n m_n}(\mathrm{Q})$$

and recalling that $X_n+\alpha_n m_n = \Psi_{\boldsymbol{\alpha}}$, the above expression becomes

$$\begin{aligned}\frac{\partial^2 \operatorname{ES}_c(\boldsymbol{\alpha})}{\partial \alpha_n \partial \boldsymbol{\alpha}'} &= \frac{1}{1-c}\left\{\int \mathbf{m}'\left(\frac{\partial \operatorname{Q}_c(\boldsymbol{\alpha})}{\partial \alpha_n} - m_n\right)\right. \\ &\qquad \left. f_{\mathbf{M}|\Psi_{\boldsymbol{\alpha}}}(\mathbf{m}|\Psi_{\boldsymbol{\alpha}} = \operatorname{Q}_c(\boldsymbol{\alpha}))\, f_{\Psi_{\boldsymbol{\alpha}}}(\operatorname{Q}_c(\boldsymbol{\alpha}))\, d\mathbf{m}\right\} \\ &= \frac{f_{\Psi_{\boldsymbol{\alpha}}}(\operatorname{Q}_c(\boldsymbol{\alpha}))}{1-c}\frac{\partial \operatorname{Q}_c(\boldsymbol{\alpha})}{\partial \alpha_n}\operatorname{E}\{\mathbf{M}'|\Psi_{\boldsymbol{\alpha}} = \operatorname{Q}_c(\boldsymbol{\alpha})\} \\ &\quad - \frac{f_{\Psi_{\boldsymbol{\alpha}}}(\operatorname{Q}_c(\boldsymbol{\alpha}))}{1-c}\operatorname{E}\{M_n\mathbf{M}'|\Psi_{\boldsymbol{\alpha}} = \operatorname{Q}_c(\boldsymbol{\alpha})\}\end{aligned} \tag{T5.181}$$

Recalling (5.188) and using vector notation we obtain

$$\begin{aligned}\frac{\partial^2 \operatorname{ES}_c(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'} &= \frac{f_{\Psi_{\boldsymbol{\alpha}}}(\operatorname{Q}_c(\boldsymbol{\alpha}))}{1-c}\operatorname{E}\{\mathbf{M}|\boldsymbol{\alpha}'\mathbf{M} = \operatorname{Q}_c(\boldsymbol{\alpha})\}\operatorname{E}\{\mathbf{M}'|\boldsymbol{\alpha}'\mathbf{M} = \operatorname{Q}_c(\boldsymbol{\alpha})\} \\ &\quad - \frac{f_{\Psi_{\boldsymbol{\alpha}}}(\operatorname{Q}_c(\boldsymbol{\alpha}))}{1-c}\operatorname{E}\left\{\mathbf{M}\mathbf{M}'|\boldsymbol{\alpha}'\mathbf{M} = \operatorname{Q}_c(\boldsymbol{\alpha})\right\} \\ &= -\frac{f_{\Psi_{\boldsymbol{\alpha}}}(\operatorname{Q}_c(\boldsymbol{\alpha}))}{1-c}\operatorname{Cov}\{\mathbf{M}|\boldsymbol{\alpha}'\mathbf{M} = \operatorname{Q}_c(\boldsymbol{\alpha})\}\end{aligned} \tag{T5.182}$$

5.6 A note on extreme value theory (EVT)

To estimate the parameters ξ and v using the MATLAB function `gpfit` proceed as follows. Define the excess as the following random variable

$$Z \equiv \widetilde{\psi} - \Psi_{\boldsymbol{\alpha}}|\Psi_{\boldsymbol{\alpha}} \leq \widetilde{\psi}. \tag{T5.183}$$

Notice that the cdf of Z satisfies

$$\begin{aligned}F_Z(z) &\equiv \mathbb{P}\{Z \leq z\} \\ &= \mathbb{P}\left\{\widetilde{\psi} - \Psi_{\boldsymbol{\alpha}} \leq z|\Psi_{\boldsymbol{\alpha}} \leq \widetilde{\psi}\right\} \\ &= \mathbb{P}\left\{\Psi_{\boldsymbol{\alpha}} \geq \widetilde{\psi} - z|\Psi_{\boldsymbol{\alpha}} \leq \widetilde{\psi}\right\} \\ &= 1 - \mathbb{P}\left\{\Psi_{\boldsymbol{\alpha}} \leq \widetilde{\psi} - z|\Psi_{\boldsymbol{\alpha}} \leq \widetilde{\psi}\right\} \\ &\equiv 1 - L_{\widetilde{\psi}}(z),\end{aligned} \tag{T5.184}$$

where in the last row we used (5.182). From (5.183) we obtain:

$$F_Z(z) \approx G_{\xi,v}(z). \tag{T5.185}$$

The function `xi_v=gpfit(Excess)`, attempts to fit (T5.185), where `Excess` are the realizations of the random variable (T5.183).

6

Technical appendix to Chapter 6

6.1 Maximum achievable certainty-equivalent with exponential utility

From the expression (6.21) of the satisfaction index, which we report here:

$$\mathrm{CE}(\boldsymbol{\alpha}) \equiv \boldsymbol{\xi}'\boldsymbol{\alpha} - \frac{1}{2\zeta}\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}, \qquad (T6.1)$$

and the constraint (6.24) we obtain the Lagrangian:

$$\mathcal{L} \equiv \boldsymbol{\xi}'\boldsymbol{\alpha} - \frac{1}{2\zeta}\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha} - \lambda\left(\boldsymbol{\alpha}'\mathbf{p}_T - w_T\right). \qquad (T6.2)$$

We neglect in the Lagrangian the second constraint (6.26), which from (6.22) and (6.24) reads:

$$\boldsymbol{\xi}'\boldsymbol{\alpha} - \mathrm{erf}^{-1}(c)\sqrt{\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}} \geq (1-\gamma)\, w_T. \qquad (T6.3)$$

We verify ex-post that the constraint is automatically satisfied.

From the first-order conditions on the Lagrangian we obtain:

$$\boldsymbol{\alpha} = \zeta\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} + \gamma\boldsymbol{\Phi}^{-1}\mathbf{p}_T, \qquad (T6.4)$$

where γ is a suitable scalar.

To compute γ we notice that the maximization of ($T6.2$) is the same as (6.70), where the objective is given by $\mathbf{M} \equiv \mathbf{P}_{T+\tau}$ and the constraint is (6.94), with $\mathbf{d} \equiv \mathbf{p}_T$ and $c \equiv w_T$. Thus the solution must be of the form (6.97). Recalling the definitions (6.99) of $\boldsymbol{\alpha}_{MV}$ and (6.100) of $\boldsymbol{\alpha}_{SR}$ respectively, and defining the scalar

$$\theta \equiv \frac{e - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{SR}}\} - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}} \qquad (T6.5)$$

we rewrite (6.97) as follows:

$$\boldsymbol{\alpha}=\theta\frac{w_T\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}+(1-\theta)\,\frac{w_T\boldsymbol{\Phi}^{-1}\mathbf{p}_T}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{p}_T}. \tag{T6.6}$$

By comparing (T6.4) with (T6.6) we obtain:

$$\theta=\frac{\zeta}{w_T}\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} \tag{T6.7}$$

and thus

$$\gamma=(1-\theta)\,\frac{w_T}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}=\frac{w_T-\zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}. \tag{T6.8}$$

Substituting this expression back into (T6.4) we obtain the optimal allocation:

$$\boldsymbol{\alpha}^*=\zeta\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}+\frac{w_T-\zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\boldsymbol{\Phi}^{-1}\mathbf{P}_T. \tag{T6.9}$$

Notice that the optimal allocation (T6.9), lies on the efficient frontier. i.e. on the hyperbola in Figure 6.11, which in our context becomes Figure 6.1.

When the risk propensity ζ is zero we obtain the minimum variance portfolio $\boldsymbol{\alpha}_{MV}$. As the risk propensity ζ tends to infinity, the solution departs from the "belly" of the hyperbola along the upper branch of the hyperbola, passing through the maximum Sharpe ratio portfolio $\boldsymbol{\alpha}_{SR}$.

The VaR constraint (T6.3) is satisfied automatically if two the confidence required c is not too high and the margin γ is not too small. Indeed consider the following equation:

$$\boldsymbol{\xi}'\boldsymbol{\alpha}-\operatorname{erf}^{-1}(c)\sqrt{\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}}=(1-\gamma)\,w_T. \tag{T6.10}$$

This is a straight line through the origin in Figure 6.1. If $\operatorname{erf}^{-1}(c)$ is not larger than the maximum Sharpe ratio, i.e. the slope of the line through the origin and the portfolio $\boldsymbol{\alpha}_{SR}$, and if γ is large enough, then all the portfolios above the straight line on the frontier satisfy the VaR constraint. These portfolios correspond to the choice (6.7) for suitable choices of the extremes.

To compute the maximum achievable index of satisfaction, we replace (T6.9) in (T6.1):

$$\begin{aligned}
\mathrm{CE}(\boldsymbol{\alpha}^{*}) \equiv & \left[\boldsymbol{\xi}'\boldsymbol{\alpha}^{*} - \frac{1}{2\zeta}\boldsymbol{\alpha}^{*\prime}\boldsymbol{\Phi}\boldsymbol{\alpha}^{*}\right] \\
= & \boldsymbol{\xi}'\left(\zeta\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} + \frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\boldsymbol{\Phi}^{-1}\mathbf{P}_T\right) \\
& -\frac{1}{2\zeta}\left(\zeta\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} + \frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\boldsymbol{\Phi}^{-1}\mathbf{P}_T\right)'\boldsymbol{\Phi} \\
& \left(\zeta\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} + \frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\boldsymbol{\Phi}^{-1}\mathbf{P}_T\right) \\
= & \zeta\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} + \left(\frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\right)\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\mathbf{P}_T \qquad (T6.11) \\
& -\frac{\zeta}{2}\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} - \frac{1}{2\zeta}\left(\frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\right)^2\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T \\
& -\frac{1}{2}\left(\frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\right)\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathrm{CE}(\boldsymbol{\alpha}^{*}) = & \frac{\zeta}{2}\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi} + \frac{1}{2}\left(\frac{w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}\right)\boldsymbol{\xi}'\boldsymbol{\Phi}^{-1}\mathbf{P}_T \qquad (T6.12) \\
& -\frac{1}{2\zeta}\frac{\left(w - \zeta\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\boldsymbol{\xi}\right)^2}{\mathbf{P}_T'\boldsymbol{\Phi}^{-1}\mathbf{P}_T}
\end{aligned}$$

6.2 Results on constrained optimization

QCQP as special case of SOCP

From the spectral decomposition, the original quadratic programming problem:

$$\begin{aligned}
\mathbf{z}^{*} \equiv & \underset{\mathbf{z}}{\operatorname{argmin}}\left\{\mathbf{z}'\mathbf{S}_{(0)}\mathbf{z} + 2\mathbf{u}_{(0)}'\mathbf{z} + v_{(0)}\right\}, \qquad (T6.13) \\
& \text{s.t.} \begin{cases} \mathbf{A}\mathbf{z} = \mathbf{a} \\ \mathbf{z}'\mathbf{S}_{(j)}\mathbf{z} + 2\mathbf{u}_{(j)}'\mathbf{z} + v_{(j)} \leq 0, \end{cases}
\end{aligned}$$

for $j = 1, \ldots, J$. can be written equivalently as follows:

$$\begin{aligned}
\mathbf{z}^{*} \equiv & \underset{\mathbf{z}}{\operatorname{argmin}}\left\{\left\|\boldsymbol{\Lambda}_{(0)}^{1/2}\mathbf{E}_{(0)}'\mathbf{z} + \boldsymbol{\Lambda}_{(0)}^{-1/2}\mathbf{E}_{(0)}'\mathbf{u}_{(0)}\right\|^2 + v_{(0)} - \mathbf{u}_{(0)}\mathbf{S}_{(0)}^{-1}\mathbf{u}_{(0)}\right\} \qquad (T6.14) \\
& \text{s.t.} \begin{cases} \mathbf{A}\mathbf{z} = \mathbf{a} \\ \left\|\boldsymbol{\Lambda}_{(j)}^{1/2}\mathbf{E}_{(j)}'\mathbf{z} + \boldsymbol{\Lambda}_{(j)}^{-1/2}\mathbf{E}_{(j)}'\mathbf{u}_{(j)}\right\|^2 \leq \mathbf{u}_{(j)}\mathbf{S}_{(j)}^{-1}\mathbf{u}_{(j)} - v_{(j)}, \end{cases} .
\end{aligned}$$

for $j = 1, \ldots, J$. This problem in turn is equivalent to:

$$\mathbf{z}^{*} \equiv \underset{\mathbf{z}}{\operatorname{argmin}} \left\{ \left\| \boldsymbol{\Lambda}_{(0)}^{1/2} \mathbf{E}_{(0)}^{\prime} \mathbf{z} + \boldsymbol{\Lambda}_{(0)}^{-1/2} \mathbf{E}_{(0)}^{\prime} \mathbf{u}_{(0)} \right\|^{2} \right\} \tag{T6.15}$$

$$\text{s.t.} \begin{cases} \mathbf{A}\mathbf{z} = \mathbf{a} \\ \left\| \boldsymbol{\Lambda}_{(j)}^{1/2} \mathbf{E}_{(j)}^{\prime} \mathbf{z} + \boldsymbol{\Lambda}_{(j)}^{-1/2} \mathbf{E}_{(j)}^{\prime} \mathbf{u}_{(j)} \right\|^{2} \leq \mathbf{u}_{(j)} \mathbf{S}_{(j)}^{-1} \mathbf{u}_{(j)} - v_{(j)}, \end{cases}$$

for $j = 1, \ldots, J$. Introducing a new variable t this problem is equivalent to:

$$(\mathbf{z}^{*}, t^{*}) \equiv \underset{(\mathbf{z},t)}{\operatorname{argmin}} \{t\} \tag{T6.16}$$

$$\text{s.t.} \begin{cases} \mathbf{A}\mathbf{z} = \mathbf{a} \\ \left\| \boldsymbol{\Lambda}_{(0)}^{1/2} \mathbf{E}_{(0)}^{\prime} \mathbf{z} + \boldsymbol{\Lambda}_{(0)}^{-1/2} \mathbf{E}_{(0)}^{\prime} \mathbf{u}_{(0)} \right\| \leq t \\ \left\| \boldsymbol{\Lambda}_{(1)}^{1/2} \mathbf{E}_{(1)}^{\prime} \mathbf{z} + \boldsymbol{\Lambda}_{(1)}^{-1/2} \mathbf{E}_{(1)}^{\prime} \mathbf{u}_{(1)} \right\| \leq \sqrt{\mathbf{u}_{(1)} \mathbf{S}_{(1)}^{-1} \mathbf{u}_{(1)} - v_{(1)}} \\ \vdots \\ \left\| \boldsymbol{\Lambda}_{(J)}^{1/2} \mathbf{E}_{(J)}^{\prime} \mathbf{z} + \boldsymbol{\Lambda}_{(J)}^{-1/2} \mathbf{E}_{(J)}^{\prime} \mathbf{u}_{(J)} \right\| \leq \sqrt{\mathbf{u}_{(J)} \mathbf{S}_{(J)}^{-1} \mathbf{u}_{(J)} - v_{(J)}} \end{cases}$$

6.3 Feasible set and MV efficient frontier

To solve

$$\boldsymbol{\alpha}(v) \equiv \underset{\boldsymbol{\alpha}^{\prime}\mathbf{d}=c, \operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\}=v}{\operatorname{argmax}} \operatorname{E}\{\Psi_{\boldsymbol{\alpha}}\}, \tag{T6.17}$$

we first compute the feasible set in the space of moments of the objective function $(v, e) = (\operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\}, \operatorname{E}\{\Psi_{\boldsymbol{\alpha}}\})$.

We consider the general case where $\operatorname{E}\{\mathbf{M}\}$ and $\mathbf{d}$ are not collinear. First we prove that any level of expected value $e \in \mathbb{R}$ is attainable. This is true if for any value e there exists an $\boldsymbol{\alpha}$ such that:

$$e = \operatorname{E}\{\Psi_{\boldsymbol{\alpha}}\} = \boldsymbol{\alpha}^{\prime} \operatorname{E}\{\mathbf{M}\} \tag{T6.18}$$

$$c = \boldsymbol{\alpha}^{\prime}\mathbf{d}. \tag{T6.19}$$

In turn, this is true if we can solve the following system for an arbitrary value of e:

$$\begin{pmatrix} \operatorname{E}\{M_j\} & \operatorname{E}\{M_k\} \\ b_j & b_k \end{pmatrix} \begin{pmatrix} \alpha_j \\ \alpha_k \end{pmatrix} = \begin{pmatrix} e - \sum_{n \neq j,k} \alpha_n \operatorname{E}\{M_k\} \\ c - \sum_{n \neq j,k} \alpha_n b_n \end{pmatrix}. \tag{T6.20}$$

Since $\operatorname{E}\{\mathbf{M}\}$ and $\mathbf{d}$ are not collinear we can always find two indices (j, k) such that the matrix on the left-hand side of ($T6.20$) is invertible. Therefore, we can fix arbitrarily e and all the entries of $\boldsymbol{\alpha}$ that appear on the right hand side of ($T6.20$) and solve for the remaining two entries on the left-hand side of ($T6.20$).

Now we prove that if a point (v, e) is feasible, so is any point $(v + \gamma, e)$, where γ is any positive number. Indeed, if we make any of the entries on the right hand side of (T6.20) go to infinity and solve for the remaining two entries on the left-hand side of (T6.20) the variance of the ensuing allocations satisfies the constraints and tends to infinity. For continuity, all the points between (v, e) and $(+\infty, e)$ are covered.

Therefore the feasible set can only be bounded on the left of the (v, e) plane. To find out if that boundary exists, we fix a generic expected value e and compute the minimum variance achievable that satisfies the affine constraint. Therefore, we minimize the following unconstrained Lagrangian:

$$\begin{aligned}\mathcal{L}(\boldsymbol{\alpha}, \lambda, \mu) &\equiv \operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\} - \lambda(\boldsymbol{\alpha}'\mathbf{d} - c) - \mu(\operatorname{E}\{\Psi_{\boldsymbol{\alpha}}\} - e). \\ &= \boldsymbol{\alpha}' \operatorname{Cov}\{\mathbf{M}\}\boldsymbol{\alpha} - \lambda(\boldsymbol{\alpha}'\mathbf{d} - c) - \mu(\boldsymbol{\alpha}' \operatorname{E}\{\mathbf{M}\} - e).\end{aligned} \tag{T6.21}$$

The first-order conditions yield:

$$\mathbf{0} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}} = 2\operatorname{Cov}\{\mathbf{M}\}\boldsymbol{\alpha} - \lambda\mathbf{d} - \mu \operatorname{E}\{\mathbf{M}\} \tag{T6.22}$$

in addition to the two constraints

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}}{\partial \lambda} = \boldsymbol{\alpha}'\mathbf{d} - c \\ 0 &= \frac{\partial \mathcal{L}}{\partial \mu} = \boldsymbol{\alpha}' \operatorname{E}\{\mathbf{M}\} - e,\end{aligned} \tag{T6.23}$$

From (T6.22) the solution reads

$$\boldsymbol{\alpha} = \frac{\lambda}{2}\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d} + \frac{\mu}{2}\operatorname{Cov}\{\mathbf{M}\}^{-1}\operatorname{E}\{\mathbf{M}\}. \tag{T6.24}$$

The Lagrange multipliers can be obtained as follows: First, we define four scalar constants:

$$\begin{array}{ll} A \equiv \mathbf{d}' \operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d} & B \equiv \mathbf{d}' \operatorname{Cov}\{\mathbf{M}\}^{-1}\operatorname{E}\{\mathbf{M}\} \\ C \equiv \operatorname{E}\{\mathbf{M}\}' \operatorname{Cov}\{\mathbf{M}\}^{-1}\operatorname{E}\{\mathbf{M}\} & D \equiv AC - B^2 \end{array} \tag{T6.25}$$

Left-multiplying the solution (T6.24) by $\mathbf{d}'$ and using the first constraint in (T6.23) we obtain:

$$\begin{aligned}c = \mathbf{d}'\boldsymbol{\alpha} &= \frac{\lambda}{2}\mathbf{d}' \operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d} \\ &\quad + \frac{\mu}{2}\mathbf{d}' \operatorname{Cov}\{\mathbf{M}\}^{-1}\operatorname{E}\{\mathbf{M}\} \\ &= \frac{\lambda}{2}A + \frac{\mu}{2}B.\end{aligned} \tag{T6.26}$$

Similarly, left-multiplying the solution (T6.24) by $\operatorname{E}\{\mathbf{M}\}'$ and using the second constraint in (T6.23) we obtain:

$$\begin{aligned} e = \mathrm{E}\{\mathbf{M}\}'\boldsymbol{\alpha} &= \frac{\lambda}{2}\mathrm{E}\{\mathbf{M}\}'\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d} \\ &+\frac{\mu}{2}\mathrm{E}\{\mathbf{M}\}'\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathrm{E}\{\mathbf{M}\} \\ &= \frac{\lambda}{2}B+\frac{\mu}{2}C \end{aligned} \quad (T6.27)$$

Now we can invert ($T6.27$) and ($T6.26$) obtaining:

$$\lambda = \frac{2cC-2eB}{D}, \qquad \mu = \frac{2eA-2cB}{D} \quad (T6.28)$$

Finally, left-multiplying ($T6.22$) by $\boldsymbol{\alpha}'$ we obtain:

$$\begin{aligned} 0 &= 2\boldsymbol{\alpha}'\operatorname{Cov}\{\mathbf{M}\}\boldsymbol{\alpha}-\lambda\boldsymbol{\alpha}'\mathbf{d}-\mu\boldsymbol{\alpha}'\mathrm{E}\{\mathbf{M}\} \\ &= 2\operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\}-\lambda c-\mu e \\ &= 2\left(\operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\}-\frac{cC-eB}{D}c-\frac{eA-cB}{D}e\right). \end{aligned} \quad (T6.29)$$

This shows that the boundary $v(e) \equiv \operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\}$ exists. Collecting the terms in e we obtain its equation:

$$v = \frac{A}{D}e^2 - \frac{2cB}{D}e + \frac{c^2C}{D}, \quad (T6.30)$$

which shows that the feasible set is bounded on the left by a parabola. In the space of the coordinates $(d, e) = (\operatorname{Sd}\{\Psi_{\boldsymbol{\alpha}}\}, \mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\})$ the parabola ($T6.30$) becomes a hyperbola:

$$d^2 = \frac{A}{D}e^2 - \frac{2cB}{D}e + \frac{c^2C}{D}, \quad (T6.31)$$

The allocations $\boldsymbol{\alpha}$ that give rise to the boundary parabola ($T6.30$) are obtained from ($T6.24$) by substituting the Lagrange multipliers ($T6.28$):

$$\begin{aligned} \boldsymbol{\alpha} &= \frac{cC-eB}{D}\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d}+\frac{eA-cB}{D}\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathrm{E}\{\mathbf{M}\} \\ &= \frac{(cC-eB)A}{D}\frac{\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d}}{\mathbf{d}'\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathbf{d}} \\ &+\frac{(eA-cB)B}{D}\frac{\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathrm{E}\{\mathbf{M}\}}{\mathbf{d}'\operatorname{Cov}\{\mathbf{M}\}^{-1}\mathrm{E}\{\mathbf{M}\}} \end{aligned} \quad (T6.32)$$

If $c \neq 0$ we can write ($T6.32$) as:

$$\boldsymbol{\alpha} = (1-\gamma(\boldsymbol{\alpha}))\boldsymbol{\alpha}_{MV} + \gamma(\boldsymbol{\alpha})\boldsymbol{\alpha}_{SR}, \quad (T6.33)$$

where the scalar γ is defined as:

$$\gamma(\boldsymbol{\alpha}) \equiv \frac{\left(\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} A - cB\right) B}{cD} \tag{T6.34}$$

and $(\boldsymbol{\alpha}_{MV}, \boldsymbol{\alpha}_{SR})$ are two specific portfolios defined as follows:

$$\boldsymbol{\alpha}_{MV} \equiv \frac{c \operatorname{Cov}\left\{\mathbf{M}\right\}^{-1} \mathbf{d}}{\mathbf{d}' \operatorname{Cov}\left\{\mathbf{M}\right\}^{-1} \mathbf{d}} \tag{T6.35}$$

$$\boldsymbol{\alpha}_{SR} \equiv \frac{c \operatorname{Cov}\left\{\mathbf{M}\right\}^{-1} \mathrm{E}\left\{\mathbf{M}\right\}}{\mathbf{d}' \operatorname{Cov}\left\{\mathbf{M}\right\}^{-1} \mathrm{E}\left\{\mathbf{M}\right\}}. \tag{T6.36}$$

Portfolio (T6.35) corresponds to the case $\gamma = 0$. From the expression for γ in (T6.34) and from the expression for the Lagrange multipliers in (T6.28) we see that $\boldsymbol{\alpha}_{MV}$ is the allocation that corresponds to the case where the Lagrange multiplier μ is zero in (T6.24). From the original Lagrangian (T6.21), if $\mu = 0$ the ensuing allocation is the minimum-variance portfolio. From (T6.30), or by direct computation we derive the coordinates of $\boldsymbol{\alpha}_{MV}$ in the space of moments:

$$v_{MV} \equiv \operatorname{Var}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\} = \frac{c^2}{A}, \quad e_{MV} \equiv \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\} = \frac{cB}{A}. \tag{T6.37}$$

Portfolio (T6.36) corresponds to the case $\gamma = 1$. This is the allocation on the feasible boundary that corresponds to the highest Sharpe ratio. Indeed, by direct computation we derive the coordinates of $\boldsymbol{\alpha}_{SR}$ in the space of moments:

$$v_{SR} \equiv \operatorname{Var}\left\{\Psi_{\boldsymbol{\alpha}_{SR}}\right\} = \frac{c^2 C}{B^2}, \quad e_{SR} \equiv \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{SR}}\right\} = \frac{cC}{B}, \tag{T6.38}$$

On the other hand the highest Sharpe ratio is the steepness of the straight line tangent to the hyperbola (T6.31), which we obtain by maximizing its analytical expression as a function of the expected value:

$$\operatorname{SR}(e) \equiv \frac{e}{d(e)} = \frac{e}{\sqrt{\frac{A}{D}e^2 - \frac{2cB}{D}e + \frac{c^2 C}{D}}}. \tag{T6.39}$$

The first-order conditions with respect to e show that the maximum of the Sharpe ratio is reached at (T6.38).

It is immediate to check that the ratio e/v is the same for both portfolio (T6.37) and portfolio (T6.38), and thus the two allocations lie on the same radius from the origin in the (v, e) plane.

As for the expression of the scalar γ in (T6.34), since

$$\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{SR}}\right\} - \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\} = \frac{cC}{B} - \frac{cB}{A} = \frac{cD}{AB} \tag{T6.40}$$

we can simplify it as follows:

$$\begin{aligned}\gamma &\equiv \frac{\left(\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} A-c B\right) B}{c D}=\frac{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} A B}{c D}-\frac{B^2}{D} \\ &=\frac{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}}{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{SR}}\right\}-\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\}}-\frac{\left(\frac{cB}{A}\right)}{\left(\frac{cD}{AB}\right)} \\ &=\frac{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}-\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\}}{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{SR}}\right\}-\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\}},\end{aligned} \qquad (T6.41)$$

which shows that the upper (lower) branch of the boundary parabola is spanned by the positive (negative) values of γ.

To consider the case $c=0$ we take the limit $c \rightarrow 0$ in the above results. The boundary (T6.30) of the feasible set in the coordinates $(v, e)=\left(\operatorname{Var}\left\{\Psi_{\boldsymbol{\alpha}}\right\}, \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}\right)$ is still a parabola:

$$v=\frac{A}{D} e^2; \qquad (T6.42)$$

whereas in the space of coordinates $(s, e)=\left(\operatorname{Sd}\left\{\Psi_{\boldsymbol{\alpha}}\right\}, \mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}\right)$ the boundary degenerates from the hyperbola (T6.31) into two straight lines:

$$d(e)=\pm \sqrt{\frac{A}{D}} e. \qquad (T6.43)$$

As for the allocations that generate this boundary, taking the limit $c \rightarrow 0$ in (T6.33) and recalling the definitions (T6.34), (T6.35) and (T6.36) we obtain:

$$\begin{aligned}\boldsymbol{\alpha} &=\lim_{c \rightarrow 0}\left[\boldsymbol{\alpha}_{MV}+\gamma(\boldsymbol{\alpha})\left(\boldsymbol{\alpha}_{SR}-\boldsymbol{\alpha}_{MV}\right)\right] \\ &=\lim_{c \rightarrow 0}\left[\gamma(\boldsymbol{\alpha})\left(\boldsymbol{\alpha}_{SR}-\boldsymbol{\alpha}_{MV}\right)\right] \\ &=\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} \frac{\operatorname{Cov}\{\mathbf{M}\}^{-1}}{D}(A \mathrm{E}\{\mathbf{M}\}-B \mathbf{d}) \\ &=\zeta(\boldsymbol{\alpha}) \operatorname{Cov}\{\mathbf{M}\}^{-1}(A \mathrm{E}\{\mathbf{M}\}-B \mathbf{d}),\end{aligned} \qquad (T6.44)$$

where the scalar ζ is defined as follows

$$\zeta(\boldsymbol{\alpha}) \equiv \frac{\mathrm{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}}{D}. \qquad (T6.45)$$

The upper (lower) branch of the boundary parabola is spanned by the positive (negative) values of ζ.

With the geometry of the feasible set at hand, we can move on to compute the mean-variance curve (T6.17): fixing a level of variance v and maximizing the expected value in the feasible set means hitting the upper branch of the parabola (T6.30). Therefore if $c \neq 0$ the mean-variance curve reads:

$$\boldsymbol{\alpha} \equiv(1-\gamma) \boldsymbol{\alpha}_{MV}+\gamma \boldsymbol{\alpha}_{SR}, \qquad \gamma>0. \qquad (T6.46)$$

if $c=0$ the mean-variance curve reads:

$$\boldsymbol{\alpha} \equiv \zeta \frac{\operatorname{Cov}\{\mathbf{M}\}^{-1}(\mathrm{E}\{\mathbf{M}\}-\mathbf{d})}{\mathbf{d}^{\prime} \operatorname{Cov}\{\mathbf{M}\}^{-1} \mathrm{E}\{\mathbf{M}\}}, \qquad \zeta \operatorname{sign}(B)>0. \qquad (T6.47)$$

6.4 The effect on the MV efficient frontier of market correlations

In the case of $N = 2$ assets the $(N-1)$-dimensional affine constraint (6.94) determines a line

$$\alpha_1 = \frac{c}{b_1} - \alpha_2 \frac{b_2}{b_1}, \tag{T6.48}$$

which corresponds to the feasible set. Defining

$$\alpha \equiv \alpha_2, \quad \widetilde{c} \equiv \frac{c}{b_1}, \quad \widetilde{b} \equiv \frac{b_2}{b_1} \tag{T6.49}$$

The investor's objective reads

$$\begin{aligned} \Psi_\alpha &= \alpha_1 M_1 + \alpha_2 M_2 \\ &= \left(\widetilde{c} - \alpha \widetilde{b}\right) M_1 + \alpha M_2 \\ &= \widetilde{c} M_1 + \alpha \left(M_2 - \widetilde{b} M_1\right). \end{aligned} \tag{T6.50}$$

Its expected value reads

$$e \equiv \mathrm{E}\{\Psi_\alpha\} = \widetilde{c}\,\mathrm{E}\{M_1\} + \alpha \left(\mathrm{E}\{M_2\} - \widetilde{b}\,\mathrm{E}\{M_1\}\right). \tag{T6.51}$$

For the standard deviation we have the general expression:

$$\begin{aligned} d^2 &\equiv [\mathrm{Sd}\{\Psi_\alpha\}]^2 = \boldsymbol{\alpha}' \,\mathrm{Cov}\{\mathbf{M}\}\, \boldsymbol{\alpha} \\ &= \left(\widetilde{c} - \alpha \widetilde{b}\right)^2 [\mathrm{Sd}\{M_1\}]^2 + \alpha^2 [\mathrm{Sd}\{M_2\}]^2 \\ &\quad + 2\alpha \left(\widetilde{c} - \alpha \widetilde{b}\right) \rho\, \mathrm{Sd}\{M_1\}\, \mathrm{Sd}\{M_2\}, \end{aligned} \tag{T6.52}$$

where $\rho \equiv \mathrm{Cor}\{M_1, M_2\}$.

From (T6.51) and (T6.52) we derive the coordinates of a full allocation in the first asset, which corresponds to $\alpha = 0$:

$$e^{(1)} = \widetilde{c}\,\mathrm{E}\{M_1\}, \qquad d^{(1)} = \widetilde{c}\,\mathrm{Sd}\{M_1\}; \tag{T6.53}$$

and a full allocation in the second asset, which corresponds to $\alpha = \widetilde{c}/\widetilde{b}$:

$$e^{(2)} = \frac{\widetilde{c}}{\widetilde{b}}\,\mathrm{E}\{M_2\}, \qquad d^{(2)} = \frac{\widetilde{c}}{\widetilde{b}}\,\mathrm{Sd}\{M_2\}. \tag{T6.54}$$

Without loss of generality, we make the assumption:

$$e^{(1)} < e^{(2)}, \qquad d^{(1)} < d^{(2)} \tag{T6.55}$$

In this notation we can more conveniently re-express the expected value (T6.51) of a generic allocation as follows:

$$e = e^{(1)} + \frac{\alpha \widetilde{b}}{\widetilde{c}} \left(e^{(2)} - e^{(1)} \right). \qquad (T6.56)$$

As for the standard deviation ($T6.52$) we obtain:

$$\begin{aligned} d^2 = & \left(1 - \frac{\alpha \widetilde{b}}{\widetilde{c}}\right)^2 \left[d^{(1)}\right]^2 + \left(\frac{\alpha \widetilde{b}}{\widetilde{c}}\right)^2 \left[d^{(2)}\right]^2 \\ & + 2\left(1 - \frac{\alpha \widetilde{b}}{\widetilde{c}}\right) \rho \frac{\alpha \widetilde{b}}{\widetilde{c}} d^{(1)} d^{(2)}, \end{aligned} \qquad (T6.57)$$

If $\rho = 1$, ($T6.57$) simplifies to:

$$d^2 = \left[\left(1 - \frac{\alpha \widetilde{b}}{\widetilde{c}}\right) d^{(1)} + \frac{\alpha \widetilde{b}}{\widetilde{c}} d^{(2)} \right]^2, \qquad (T6.58)$$

which as long as

$$\alpha \geq -\frac{d^{(1)}}{d^{(2)} - d^{(1)}} \frac{\widetilde{c}}{\widetilde{b}} \qquad (T6.59)$$

in turn simplifies to

$$d = \left(1 - \frac{\alpha \widetilde{b}}{\widetilde{c}}\right) d^{(1)} + \frac{\alpha \widetilde{b}}{\widetilde{c}} d^{(2)}. \qquad (T6.60)$$

This expression coupled with ($T6.56$) yield the allocation curve in the case $\rho = 1$:

$$e = e^{(1)} + \left(d - d^{(1)}\right) \frac{e^{(2)} - e^{(1)}}{d^{(2)} - d^{(1)}}, \qquad (T6.61)$$

which is a line through the coordinates of the two securities. When the allocation α is such that ($T6.59$) holds as an equality, we obtain a zero-variance portfolio whose expected value from ($T6.56$) reads:

$$e = e^{(1)} - d^{(1)} \frac{e^{(2)} - e^{(1)}}{d^{(2)} - d^{(1)}} < e^{(1)}. \qquad (T6.62)$$

Notice from ($T6.59$) that this situation corresponds to a negative position in the second asset.

If $\rho = -1$, ($T6.57$) simplifies to

$$d^2 = \left[\left(1 - \frac{\alpha \widetilde{b}}{\widetilde{c}}\right) d^{(1)} - \frac{\alpha \widetilde{b}}{\widetilde{c}} d^{(2)} \right]^2, \qquad (T6.63)$$

which as long as

$$\alpha \geq \frac{d^{(1)}}{d^{(1)} + d^{(2)}} \frac{\widetilde{c}}{\widetilde{b}} \qquad (T6.64)$$

in turn simplifies to

$$d = -d^{(1)} + \frac{\alpha \widetilde{b}}{\widetilde{c}} \left(d^{(1)} + d^{(2)} \right). \tag{T6.65}$$

This expression coupled with (T6.56) yield the allocation curve in the case $\rho = -1$:

$$e = e^{(1)} + \left(d + d^{(1)} \right) \frac{e^{(2)} - e^{(1)}}{d^{(2)} + d^{(1)}}. \tag{T6.66}$$

When the allocation α is such that (T6.64) holds as an equality, we obtain a zero-variance portfolio whose expected value from (T6.56) reads:

$$e = e^{(1)} + d^{(1)} \frac{e^{(2)} - e^{(1)}}{d^{(1)} + d^{(2)}} > e^{(1)}. \tag{T6.67}$$

Notice that in this situation from (T6.64) the allocation in the second asset is positive and from (T6.48) so is the allocation in the first asset:

$$\alpha_1 = \widetilde{c} \left(1 - \frac{d^{(1)}}{d^{(1)} + d^{(2)}} \right). \tag{T6.68}$$

6.5 The geometry of total-return- and benchmark-allocation

Total return efficient allocations in the plane of relative coordinates

Here we show that the efficient frontier is a translation of the relative frontier in the plane of expected variance / expected value of relative returns. From (6.193) the generic portfolio (6.175) on the efficient frontier satisfies:

$$\operatorname{Var}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}\} = \frac{A}{D} \operatorname{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}\}^2 - \frac{2wB}{D} \operatorname{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}\} + \frac{w^2 C}{D} \tag{T6.69}$$

which can be re-written as follows:

$$\begin{aligned} \operatorname{Var}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} - \operatorname{Var}\{\Psi_{\boldsymbol{\beta}}\} + 2\operatorname{Cov}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}, \Psi_{\boldsymbol{\beta}}\} &= \frac{A}{D} \left(\operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} + \operatorname{E}\{\Psi_{\boldsymbol{\beta}}\} \right)^2 \\ &\quad - \frac{2wB}{D} \left(\operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} + \operatorname{E}\{\Psi_{\boldsymbol{\beta}}\} \right) + \frac{w^2 C}{D} \end{aligned} \tag{T6.70}$$

Expanding the products and rearranging, we obtain

$$\begin{aligned} \operatorname{Var}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} &= -2\operatorname{Cov}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}, \Psi_{\boldsymbol{\beta}}\} + \frac{A}{D} \operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\}^2 \\ &\quad + \operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} \left[\frac{2A}{D} \operatorname{E}\{\Psi_{\boldsymbol{\beta}}\} - \frac{2wB}{D} \right] \\ &\quad + \frac{A}{D} \operatorname{E}\{\Psi_{\boldsymbol{\beta}}\}^2 - \frac{2wB}{D} \operatorname{E}\{\Psi_{\boldsymbol{\beta}}\} + \frac{w^2 C}{D} + \operatorname{Var}\{\Psi_{\boldsymbol{\beta}}\} \end{aligned} \tag{T6.71}$$

From (6.194), (6.99) and (6.100) we obtain for a generic allocation $\boldsymbol{\alpha}$:

$$\begin{aligned}\frac{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}-\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{SR}}\}-\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}} &= \frac{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}-\frac{w\,\mathrm{E}\{\mathbf{P}_{T+\tau}\}'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\mathbf{P}_T}{\mathbf{P}_T'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\mathbf{P}_T}}{\frac{w\,\mathrm{E}\{\mathbf{P}_{T+\tau}\}'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\,\mathrm{E}\{\mathbf{P}_{T+\tau}\}}{\mathbf{P}_T'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\,\mathrm{E}\{\mathbf{P}_{T+\tau}\}}-\frac{w\,\mathrm{E}\{\mathbf{P}_{T+\tau}\}'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\mathbf{P}_T}{\mathbf{P}_T'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\mathbf{P}_T}} \\ &= \frac{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}-\frac{wB}{A}}{\frac{wC}{B}-\frac{wB}{A}} = \frac{\left(\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}-\frac{wB}{A}\right)BA}{wD} \\ &= \frac{\left(\mathrm{E}\{\Psi_{\boldsymbol{\alpha}-\boldsymbol{\beta}}\}+\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}\right)BA-wB^2}{wD} \qquad (T6.72)\end{aligned}$$

Using this result, from (6.175) and the budget constraint $\boldsymbol{\beta}'\mathbf{P}_T = w$ the covariance reads:

$$\begin{aligned}\mathrm{Cov}\{\Psi_{\boldsymbol{\beta}},\Psi_{\widetilde{\boldsymbol{\alpha}}}\} &= \boldsymbol{\beta}'\,\mathrm{Cov}\{\mathbf{P}_{T+\tau}\} \qquad (T6.73)\\ &\quad\left(\boldsymbol{\alpha}_{MV}+\frac{\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}\}-\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{SR}}\}-\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}\left(\boldsymbol{\alpha}_{SR}-\boldsymbol{\alpha}_{MV}\right)\right) \\ &= \frac{w\boldsymbol{\beta}'\mathbf{P}_T}{A}+\frac{\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}}\}-\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{SR}}\}-\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}\left(\frac{w\boldsymbol{\beta}'\,\mathrm{E}\{\mathbf{P}_{T+\tau}\}}{B}-\frac{w\boldsymbol{\beta}'\mathbf{P}_T}{A}\right) \\ &= \frac{w^2}{A}+\frac{\left(\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}+\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}\right)BA-wB^2}{D}\left(\frac{\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{B}-\frac{w}{A}\right)\end{aligned}$$

Substituting (T6.72) in (T6.73) we obtain:

$$\begin{aligned}\mathrm{Var}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\} &= -2\left[\frac{w^2}{A}+\frac{\left(\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}+\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}\right)BA-wB^2}{D}\left(\frac{\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{B}-\frac{w}{A}\right)\right] \\ &\quad+\frac{A}{D}\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}^2+\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}\left(\frac{2A}{D}\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}-\frac{2wB}{D}\right) \\ &\quad+\frac{A}{D}\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}^2-\frac{2wB}{D}\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}+\frac{w^2C}{D}+\mathrm{Var}\{\Psi_{\boldsymbol{\beta}}\} \\ &= -\frac{2w^2}{A}-\frac{2\,\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}BA}{D}\left(\frac{\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{B}-\frac{w}{A}\right) \qquad (T6.74)\\ &\quad-\frac{2\,\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}BA}{D}\left(\frac{\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{B}-\frac{w}{A}\right)+\frac{2wB^2}{D}\left(\frac{\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{B}-\frac{w}{A}\right) \\ &\quad+\frac{A}{D}\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}^2+\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}\left(\frac{2A}{D}\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}-\frac{2wB}{D}\right) \\ &\quad+\frac{A}{D}\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}^2-\frac{2wB}{D}\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}+\frac{w^2C}{D}+\mathrm{Var}\{\Psi_{\boldsymbol{\beta}}\}\end{aligned}$$

The first-degree terms in $\mathrm{E}\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\}$ cancel, and other terms simplify to yield the following expression:

$$\operatorname{Var}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} = \frac{A}{D}\operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\}^2 + w^2\left(\frac{C}{D} - \frac{2}{A} - \frac{2B^2}{DA}\right) - \frac{A}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}^2 + \frac{2wB}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\} + \operatorname{Var}\left\{\Psi_{\boldsymbol{\beta}}\right\} \quad (T6.75)$$

From the definition of D in (6.194)this simplifies further into:

$$\operatorname{Var}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} = \frac{A}{D}\operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\}^2 - \frac{w^2C}{D} - \frac{A}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}^2 + \frac{2wB}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\} + \operatorname{Var}\left\{\Psi_{\boldsymbol{\beta}}\right\} \quad (T6.76)$$

or

$$\operatorname{Var}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\} = \frac{A}{D}\operatorname{E}\left\{\Psi_{\widetilde{\boldsymbol{\alpha}}-\boldsymbol{\beta}}\right\}^2 + \delta_{\boldsymbol{\beta}} \quad (T6.77)$$

where

$$\delta_{\boldsymbol{\beta}} \equiv \operatorname{Var}\left\{\Psi_{\boldsymbol{\beta}}\right\} - \frac{A}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}^2 + \frac{2wB}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\} - \frac{w^2C}{D}. \quad (T6.78)$$

Since the benchmark is not necessarily mean-variance efficient from (6.193) we have that $\delta_{\boldsymbol{\beta}} \geq 0$ and the equality holds if and only if the benchmark is mean-variance efficient.

Benchmark-relative efficient allocation in the plane of absolute coordinates

From the equation of the relative frontier (6.199) which we re-write here

$$\operatorname{Var}\left\{\Psi_{\boldsymbol{\alpha}-\boldsymbol{\beta}}\right\} = \frac{A}{D}\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}-\boldsymbol{\beta}}\right\}^2, \quad (T6.79)$$

and the linearity of the objective $\Psi_{\boldsymbol{\alpha}-\boldsymbol{\beta}} = \Psi_{\boldsymbol{\alpha}} - \Psi_{\boldsymbol{\beta}}$ we obtain

$$\operatorname{Var}\left\{\Psi_{\boldsymbol{\alpha}}\right\} = 2\operatorname{Cov}\left\{\Psi_{\boldsymbol{\alpha}}, \Psi_{\boldsymbol{\beta}}\right\} - \operatorname{Var}\left\{\Psi_{\boldsymbol{\beta}}\right\} + \frac{A}{D}\left(\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}^2 + \operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}^2 - 2\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\}\operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}\right), \quad (T6.80)$$

From (6.194), (6.99) and (6.100) we obtain for a generic allocation $\boldsymbol{\alpha}$:

$$\begin{aligned}
\frac{\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} - \operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}}{\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}_{SR}}\right\} - \operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}_{MV}}\right\}} &= \frac{\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} - \operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}}{\frac{w\operatorname{E}\{\mathbf{P}_{T+\tau}\}'\operatorname{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\operatorname{E}\{\mathbf{P}_{T+\tau}\}}{\mathbf{P}_T'\operatorname{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\operatorname{E}\{\mathbf{P}_{T+\tau}\}} - \frac{w\operatorname{E}\{\mathbf{P}_{T+\tau}\}'\operatorname{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\mathbf{P}_T}{\mathbf{P}_T'\operatorname{Cov}\{\mathbf{P}_{T+\tau}\}^{-1}\mathbf{P}_T}} \\
&= \frac{\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} - \operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}}{\frac{wC}{B} - \frac{wB}{A}} \\
&= \frac{\left(\operatorname{E}\left\{\Psi_{\boldsymbol{\alpha}}\right\} - \operatorname{E}\left\{\Psi_{\boldsymbol{\beta}}\right\}\right)BA}{wD} \quad (T6.81)
\end{aligned}$$

Using this result, from (6.190) and the budget constraint $\boldsymbol{\beta}'\mathbf{P}_T = w$ the covariance reads:

$$\begin{aligned}\operatorname{Cov}\{\Psi_{\boldsymbol{\beta}}, \Psi_{\boldsymbol{\alpha}}\} &= \boldsymbol{\beta}' \operatorname{Cov}\{\mathbf{P}_{T+\tau}\}\left(\boldsymbol{\beta} + \frac{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} - \mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{SR}}\} - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}(\boldsymbol{\alpha}_{SR} - \boldsymbol{\alpha}_{MV})\right) \\ &= \operatorname{Var}\{\Psi_{\boldsymbol{\beta}}\} \qquad (T6.82) \\ &\quad + \frac{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} - \mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{\mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{SR}}\} - \mathrm{E}\{\Psi_{\boldsymbol{\alpha}_{MV}}\}}\left(\frac{w\boldsymbol{\beta}' \mathrm{E}\{\mathbf{P}_{T+\tau}\}}{B} - \frac{w\boldsymbol{\beta}'\mathbf{P}_T}{A}\right) \\ &= \operatorname{Var}\{\Psi_{\boldsymbol{\beta}}\} + \frac{(\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} - \mathrm{E}\{\Psi_{\boldsymbol{\beta}}\})BA}{D}\left(\frac{\mathrm{E}\{\Psi_{\boldsymbol{\beta}}\}}{B} - \frac{w}{A}\right)\end{aligned}$$

Substituting this into (T6.80) and simplifying we obtain:

$$\operatorname{Var}\{\Psi_{\boldsymbol{\alpha}}\} = \frac{A}{D}\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\}^2 - \frac{2Bw}{D}\mathrm{E}\{\Psi_{\boldsymbol{\alpha}}\} + \frac{w^2C}{D} + \delta_{\boldsymbol{\beta}}, \qquad (T6.83)$$

where $\delta_{\boldsymbol{\beta}}$ is defined in (T6.78).

6.6 Formulation of MV in terms of returns

If the investor has a positive initial budget $W_T > 0$, then maximizing $\mathrm{E}\{W_{T+\tau}\}$ in the original mean-variance problem (6.68) is equivalent to maximizing $\mathrm{E}\{W_{T+\tau}\}/W_T$. On the other hand, as $\operatorname{Var}\{W_{T+\tau}\}$ spans all the real numbers v, so does $\operatorname{Var}\{W_{T+\tau}\}/W_T^2$. Therefore, given that initial wealth W_T is not a random variable and given the definition of linear return on wealth (6.81), the original mean-variance problem (6.68) is equivalent to the following expression:

$$\boldsymbol{\alpha}(v) \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}, \operatorname{Var}\left\{L^W_{T,\tau}\right\}=v}{\operatorname{argmax}} \mathrm{E}\left\{L^W_{T,\tau}\right\}. \qquad (T6.84)$$

To solve (T6.84) we notice that the linear return on wealth is a function of relative weights and linear returns on securities:

$$\begin{aligned}1 + L^W_{T,\tau} \equiv \frac{W_{T+\tau}}{W_T} &= \frac{\sum_{n=1}^N \alpha_n P^{(n)}_{T+\tau}}{W_T} \qquad (T6.85) \\ &= \sum_{n=1}^N \frac{\alpha_n P^{(n)}_T}{W_T}\frac{P^{(n)}_{T+\tau}}{P^{(n)}_T} = \sum_{n=1}^N w_n\left(1 + L^{(n)}_{T,\tau}\right) \\ &= 1 + \mathbf{w}'\mathbf{L}_{T,\tau},\end{aligned}$$

where we have used the budget constraint $\boldsymbol{\alpha}'\mathbf{P}_T = W_T$ and the following identity:

$$\sum_{n=1}^N w_n = \sum_{n=1}^N \frac{\alpha_n P^{(n)}_T}{\boldsymbol{\alpha}'\mathbf{P}_T} = 1. \qquad (T6.86)$$

Therefore for the expected value we have:

$$\mathrm{E}\left\{L_{T,\tau}^{W}\right\} = \mathrm{E}\left\{\mathbf{w}'\mathbf{L}_{T,\tau}\right\} = \mathbf{w}'\,\mathrm{E}\left\{\mathbf{L}_{T,\tau}\right\} \tag{T6.87}$$

Similarly, for the variance we obtain:

$$\mathrm{Var}\left\{L_{T,\tau}^{W}\right\} = \mathrm{Var}\left\{\mathbf{w}'\mathbf{L}_{T,\tau}\right\} = \mathbf{w}'\,\mathrm{Cov}\left\{\mathbf{L}_{T,\tau}\right\}\mathbf{w} \tag{T6.88}$$

Therefore, due to (T6.87) and (T6.88) problem (T6.84) reads:

$$\boldsymbol{\alpha}(v) \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C},\,\mathbf{w}'(\boldsymbol{\alpha})\,\mathrm{Cov}\{\mathbf{L}_{T,\tau}\}\mathbf{w}(\boldsymbol{\alpha})=v}{\operatorname{argmax}} \mathbf{w}'(\boldsymbol{\alpha})\,\mathrm{E}\left\{\mathbf{L}_{T,\tau}\right\}, \tag{T6.89}$$

where $\mathbf{w}$ as a function of $\boldsymbol{\alpha}$ is obtained by inverting (6.86). On the other hand, it is easier to convert the constraints $\mathcal{C}$ that hold for $\boldsymbol{\alpha}$ into constraints that hold for $\mathbf{w}$ (for ease of exposition we keep denoting them as $\mathcal{C}$) and then maximize (T6.89) with respect to $\mathbf{w}$:

$$\mathbf{w}(v) \equiv \underset{\mathbf{w}\in\mathcal{C},\,\mathbf{w}'\,\mathrm{Cov}\{\mathbf{L}_{T,\tau}\}\mathbf{w}=v}{\operatorname{argmax}} \mathbf{w}'\,\mathrm{E}\left\{\mathbf{L}_{T,\tau}\right\}, \tag{T6.90}$$

The original mean-variance curve is simply $\boldsymbol{\alpha}(v) \equiv \boldsymbol{\alpha}(\mathbf{w}(v))$ obtained from (6.86).

7

Technical appendix to Chapter 7

7.1 Mahalanobis square distance of normal variables

Consider a generic multivariate normal random variable:

$$\mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) . \tag{T7.1}$$

Consider the spectral decomposition (3.149) of the covariance matrix:

$$\boldsymbol{\Sigma} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}', \tag{T7.2}$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix of the respective eigenvalues sorted in decreasing order:

$$\boldsymbol{\Lambda} \equiv \operatorname{diag}(\lambda_1, \ldots, \lambda_N) . \tag{T7.3}$$

and the matrix $\mathbf{E}$ is the juxtaposition of the eigenvectors, which represents a rotation:

$$\mathbf{E} \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(N)}\right) . \tag{T7.4}$$

Now consider the new random variable:

$$\mathbf{Y} \equiv \boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{X} - \boldsymbol{\mu}) . \tag{T7.5}$$

From (2.163) we obtain

$$\mathbf{Y} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}) . \tag{T7.6}$$

Therefore from (1.106) and (1.109) it follows:

$$\sum_{n=1}^{N} Y_n^2 \sim \chi_N^2 . \tag{T7.7}$$

On the other hand

$$\sum_{n=1}^{N} Y_n^2 = \mathbf{Y}'\mathbf{Y} \tag{T7.8}$$
$$= \left[\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{X}-\boldsymbol{\mu})\right]'\left[\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}'(\mathbf{X}-\boldsymbol{\mu})\right]$$
$$= (\mathbf{X}-\boldsymbol{\mu})'\mathbf{E}\boldsymbol{\Lambda}^{-1}\mathbf{E}'(\mathbf{X}-\boldsymbol{\mu})$$
$$= (\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}).$$

Therefore for the Mahalanobis distance (2.61) of the variable $\mathbf{X}$ from the point $\boldsymbol{\mu}$ through the metric $\boldsymbol{\Sigma}$ we obtain:

$$\mathrm{Ma}^2(\mathbf{X},\boldsymbol{\mu},\boldsymbol{\Sigma}) \equiv (\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \sim \chi_N^2. \tag{T7.9}$$

From the definition (1.7) of cumulative distribution function:

$$F_{\chi_N^2}(q) \equiv \mathbb{P}\left\{\mathrm{Ma}^2(\mathbf{X},\boldsymbol{\mu},\boldsymbol{\Sigma}) \leq q\right\} \tag{T7.10}$$
$$= \mathbb{P}\left\{(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \leq q\right\}.$$

By applying the quantile function (1.17) to both sides of the above equality we obtain:

$$p = \mathbb{P}\left\{(\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \leq (q_N^p)^2\right\}, \tag{T7.11}$$

where q_N^p is the square root of the quantile of the chi-square distribution with N degrees of freedom relative to a confidence level p:

$$q_N^p \equiv \sqrt{Q_{\chi_N^2}(p)}. \tag{T7.12}$$

Therefore, from the definition of the ellipsoid:

$$\mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}^{q} \equiv \left\{\mathbf{x} \text{ such that } (\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \leq q^2\right\} \tag{T7.13}$$

we obtain:

$$\mathbb{P}\left\{\mathbf{X} \in \mathcal{E}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}^{q_N^p}\right\} = p. \tag{T7.14}$$

7.2 NIW location-dispersion: posterior distribution

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

By the NIW (normal-inverse-Wishart) assumption (7.20)-(7.21) on the prior

$$\boldsymbol{\Omega} \equiv \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_0, (\nu_0\boldsymbol{\Sigma}_0)^{-1}\right) \tag{T7.15}$$

and

$$\boldsymbol{\mu}|\boldsymbol{\Omega} \sim \mathrm{N}\left(\boldsymbol{\mu}_0, (T_0\boldsymbol{\Omega})^{-1}\right). \tag{T7.16}$$

Thus from (2.156) and (2.224) the joint prior pdf of $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ is

$$\begin{aligned} f_{\text{pr}}(\boldsymbol{\mu},\boldsymbol{\Omega}) &= f_{\text{pr}}(\boldsymbol{\mu}|\boldsymbol{\Omega})\, f_{\text{pr}}(\boldsymbol{\Omega}) \\ &= \gamma_1 |\boldsymbol{\Omega}|^{\frac{1}{2}}\, e^{-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_0)'(T_0\boldsymbol{\Omega})(\boldsymbol{\mu}-\boldsymbol{\mu}_0)} \\ &\quad |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}}\, |\boldsymbol{\Omega}|^{\frac{\nu_0-N-1}{2}}\, e^{-\frac{1}{2}\operatorname{tr}(\nu_0\boldsymbol{\Sigma}_0\boldsymbol{\Omega})}. \end{aligned} \tag{T7.17}$$

As for the pdf of current information (7.13), from (4.102) the sample mean is normally distributed

$$\widehat{\boldsymbol{\mu}} \sim \mathrm{N}\left(\boldsymbol{\mu}, [T\boldsymbol{\Omega}]^{-1}\right) \tag{T7.18}$$

and from (4.103) the distribution of the sample covariance is

$$T\widehat{\boldsymbol{\Sigma}} \sim \mathrm{W}(T-1,\boldsymbol{\Sigma}) \tag{T7.19}$$

and these variables are independent. Therefore from (2.156) and (2.224) the pdf of current information from time series $f(i_T|\boldsymbol{\mu},\boldsymbol{\Omega})$ as summarized by $i_T \equiv \left(\widehat{\boldsymbol{\mu}}, T\widehat{\boldsymbol{\Sigma}}\right)$ conditioned on knowledge of the parameters $(\boldsymbol{\mu},\boldsymbol{\Omega})$ reads:

$$\begin{aligned} f(i_T|\boldsymbol{\mu},\boldsymbol{\Omega}) &= \gamma_2 |\boldsymbol{\Omega}|^{\frac{1}{2}}\, e^{-\frac{1}{2}(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})'(T\boldsymbol{\Omega})(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu})} \\ &\quad |\boldsymbol{\Omega}|^{\frac{T-1}{2}} \left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} e^{-\frac{1}{2}\operatorname{tr}\left(T\boldsymbol{\Omega}\widehat{\boldsymbol{\Sigma}}\right)} \end{aligned} \tag{T7.20}$$

Thus, after trivial regrouping and simplifications, the joint pdf of current information and the parameters reads:

$$\begin{aligned} f(i_T,\boldsymbol{\mu},\boldsymbol{\Omega}) &= f(i_T|\boldsymbol{\mu},\boldsymbol{\Omega})\, f_{\text{pr}}(\boldsymbol{\mu},\boldsymbol{\Omega}) = \\ &\gamma_3 e^{-\frac{1}{2}\left\{(\boldsymbol{\mu}-\boldsymbol{\mu}_0)'(T_0\boldsymbol{\Omega})(\boldsymbol{\mu}-\boldsymbol{\mu}_0)+(\boldsymbol{\mu}-\widehat{\boldsymbol{\mu}})'(T\boldsymbol{\Omega})(\boldsymbol{\mu}-\widehat{\boldsymbol{\mu}})\right\}} \\ &\left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}}\, |\boldsymbol{\Omega}|^{\frac{T+\nu_0-N}{2}} \\ &e^{-\frac{1}{2}\operatorname{tr}\left(T\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Omega}+\nu_0\boldsymbol{\Sigma}_0\boldsymbol{\Omega}\right)} \end{aligned} \tag{T7.21}$$

After expanding and rearranging, the terms in the curly brackets in the second row can be re-written as follows:

$$\{\cdots\} = (\boldsymbol{\mu}-\boldsymbol{\mu}_1)'\, T_1\boldsymbol{\Omega}\,(\boldsymbol{\mu}-\boldsymbol{\mu}_1) + \operatorname{tr}(\boldsymbol{\Phi}\boldsymbol{\Omega}) \tag{T7.22}$$

where

$$\begin{aligned} T_1 &\equiv T_0 + T \\ \boldsymbol{\mu}_1 &\equiv \frac{T_0\boldsymbol{\mu}_0 + T\widehat{\boldsymbol{\mu}}}{T_0+T} \\ \boldsymbol{\Phi} &\equiv \frac{TT_0}{T_0+T}(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)(\widehat{\boldsymbol{\mu}}-\boldsymbol{\mu}_0)' \end{aligned} \tag{T7.23}$$

Therefore, defining

$$\boldsymbol{\Sigma}_1 \equiv \frac{T\widehat{\boldsymbol{\Sigma}} + \nu_0 \boldsymbol{\Sigma}_0 + \boldsymbol{\Phi}}{\nu_1}, \tag{T7.24}$$

where ν_1 is a number yet to be defined, we can re-write the joint pdf (T7.21) as follows:

$$\begin{aligned} f(i_T, \boldsymbol{\mu}, \boldsymbol{\Omega}) = \gamma_3 e^{-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)' T_1 \boldsymbol{\Omega}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)} \\ \left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}} |\boldsymbol{\Omega}|^{\frac{T+\nu_0-N}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1 \boldsymbol{\Sigma}_1 \boldsymbol{\Omega})} \end{aligned} \tag{T7.25}$$

At this point we can perform the integration over $(\boldsymbol{\mu}, \boldsymbol{\Omega})$ to find the marginal pdf $f(i_T)$

$$\begin{aligned} f(i_T) &= \int f(i_T, \boldsymbol{\mu}, \boldsymbol{\Omega})\, d\boldsymbol{\mu} d\boldsymbol{\Omega} \\ &= \gamma_4 \int \left\{ \int \gamma_5 |\boldsymbol{\Omega}|^{\frac{1}{2}} e^{-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)' T_1 \boldsymbol{\Omega}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)} d\boldsymbol{\mu} \right\} \\ &\quad \left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}} |\boldsymbol{\Omega}|^{\frac{T+\nu_0-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1 \boldsymbol{\Sigma}_1 \boldsymbol{\Omega})} d\boldsymbol{\Omega} \\ &= \gamma_4 \int \left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}} |\boldsymbol{\Omega}|^{\frac{T+\nu_0-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1 \boldsymbol{\Sigma}_1 \boldsymbol{\Omega})} d\boldsymbol{\Omega}, \end{aligned} \tag{T7.26}$$

where we have used the fact that the term in curly brackets is the integral of a normal pdf (2.156) over the entire space and thus sum to one. Defining now

$$\nu_1 \equiv T + \nu_0 \tag{T7.27}$$

we write (T7.26) as follows:

$$\begin{aligned} f(i_T) &= \gamma_6 \left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}} |\boldsymbol{\Sigma}_1|^{-\frac{\nu_1}{2}} \\ &\quad \left\{ \int \gamma_7 |\boldsymbol{\Sigma}_1|^{\frac{\nu_1}{2}} |\boldsymbol{\Omega}|^{\frac{\nu_1-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1 \boldsymbol{\Sigma}_1 \boldsymbol{\Omega})} d\boldsymbol{\Omega} \right\} \\ &= \gamma_6 \left|\widehat{\boldsymbol{\Sigma}}\right|^{\frac{T-N-2}{2}} |\boldsymbol{\Sigma}_0|^{\frac{\nu_0}{2}} |\boldsymbol{\Sigma}_1|^{-\frac{\nu_1}{2}}, \end{aligned} \tag{T7.28}$$

where we have used the fact that the term in curly brackets is the integral of a Wishart pdf (2.224) over the entire space and thus sum to one.

Finally, we obtain the posterior pdf (7.15) by dividing the joint pdf (T7.25) by the marginal pdf (T7.28):

$$\begin{aligned} f_{\text{po}}(\boldsymbol{\mu}, \boldsymbol{\Omega}) &\equiv \frac{f(i_T, \boldsymbol{\mu}, \boldsymbol{\Omega})}{f(i_T)} \\ &\quad \gamma_7 |\boldsymbol{\Omega}|^{\frac{1}{2}} e^{-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)' T_1 \boldsymbol{\Omega}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)} \\ &\quad |\boldsymbol{\Sigma}_1|^{\frac{\nu_1}{2}} |\boldsymbol{\Omega}|^{\frac{\nu_1-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1 \boldsymbol{\Sigma}_1 \boldsymbol{\Omega})} \end{aligned} \tag{T7.29}$$

From (2.156) and (2.224) we see that this means:

$$\boldsymbol{\mu}|\boldsymbol{\Omega} \sim \mathrm{N}\left(\boldsymbol{\mu}_1, [T_1\boldsymbol{\Omega}]^{-1}\right) \tag{$T7.30$}$$

and

$$\boldsymbol{\Omega} \sim \mathrm{W}\left(\nu_1, (\nu_1\boldsymbol{\Sigma}_1)^{-1}\right). \tag{$T7.31$}$$

In other words,

$$\boldsymbol{\mu}|\boldsymbol{\Sigma} \sim \mathrm{N}\left(\boldsymbol{\mu}_1, \frac{\boldsymbol{\Sigma}^{-1}}{T_1}\right) \tag{$T7.32$}$$

and

$$\boldsymbol{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_1, \frac{\boldsymbol{\Sigma}_1^{-1}}{\nu_1}\right). \tag{$T7.33$}$$

7.3 NIW location-dispersion: mode and modal dispersion

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

We consider the notation for the NIW (normal-inverse-Wishart) assumptions (7.32)-(7.33) on the posterior, although of course the proof applies verbatim to the prior, or any NIW distribution. Thus assume

$$\boldsymbol{\Sigma}^{-1} \equiv \boldsymbol{\Omega} \sim \mathrm{W}\left(\nu_1, (\nu_1\boldsymbol{\Sigma}_1)^{-1}\right). \tag{$T7.34$}$$

and

$$\boldsymbol{\mu}|\boldsymbol{\Omega} \sim \mathrm{N}\left(\boldsymbol{\mu}_1, [T_1\boldsymbol{\Omega}]^{-1}\right) \tag{$T7.35$}$$

The parameter in this context are

$$\boldsymbol{\theta} \equiv \left(\boldsymbol{\mu}', \mathrm{vech}\left[\boldsymbol{\Omega}\right]'\right)'. \tag{$T7.36$}$$

From (2.156) and (2.224) the joint NIW (normal-inverse-Wishart) probability density function of $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$ reads:

$$\begin{aligned} f(\boldsymbol{\theta}) &= f(\boldsymbol{\mu}|\boldsymbol{\Omega})\, f(\boldsymbol{\Omega}) \\ &= \gamma_1 \left|\boldsymbol{\Omega}\right|^{\frac{\nu_1 - N}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1\boldsymbol{\Sigma}_1\boldsymbol{\Omega})} e^{-\frac{T_1}{2}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)'\boldsymbol{\Omega}(\boldsymbol{\mu}-\boldsymbol{\mu}_1)} \end{aligned} \tag{$T7.37$}$$

To determine the mode of this distribution

$$\widetilde{\boldsymbol{\theta}} \equiv \left(\widetilde{\boldsymbol{\mu}}', \mathrm{vech}\left[\widetilde{\boldsymbol{\Omega}}\right]'\right)' \tag{$T7.38$}$$

we impose the first-order conditions on the logarithm of the joint probability density function ($T7.37$).

$$\ln f \equiv \gamma_2 + \frac{\nu_1 - N}{2} \ln |\boldsymbol{\Omega}| - \frac{1}{2} \operatorname{tr} \left\{ \nu_1 \boldsymbol{\Sigma}_1 \boldsymbol{\Omega} + T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)(\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \boldsymbol{\Omega} \right\}. \tag{T7.39}$$

Computing the first variation and using $(A.124)$ we obtain:

$$\begin{aligned} d \ln f \equiv & -\frac{1}{2} \operatorname{tr} \left\{ \left(-(\nu_1 - N) \boldsymbol{\Omega}^{-1} + \nu_1 \boldsymbol{\Sigma}_1 + T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)(\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \right) d\boldsymbol{\Omega} \right\} \\ & - \operatorname{tr} \left\{ T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \boldsymbol{\Omega} d\boldsymbol{\mu} \right\}. \end{aligned} \tag{T7.40}$$

Therefore

$$d \ln f \equiv \operatorname{tr} \{ \mathbf{G}_{\boldsymbol{\Omega}} d\boldsymbol{\Omega} \} + \operatorname{tr} \{ \mathbf{G}_{\boldsymbol{\mu}} d\boldsymbol{\mu} \}, \tag{T7.41}$$

where

$$\mathbf{G}_{\boldsymbol{\Omega}} \equiv \frac{1}{2} \left[(\nu_1 - N) \boldsymbol{\Omega}^{-1} - \nu_1 \boldsymbol{\Sigma}_1 - T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)(\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \right] \tag{T7.42}$$

$$\mathbf{G}_{\boldsymbol{\mu}} \equiv -T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \boldsymbol{\Omega} \tag{T7.43}$$

Using $(A.120)$ and the duplication matrix $(A.113)$ to get rid of the redundancies of $d\boldsymbol{\Omega}$ in $(T7.41)$ we obtain:

$$d \ln f = \operatorname{vec} [\mathbf{G}'_{\boldsymbol{\Omega}}]' \mathbf{D}_N \operatorname{vech} [d\boldsymbol{\Omega}] + \operatorname{vec} \left[\mathbf{G}'_{\boldsymbol{\mu}} \right]' \operatorname{vec} [d\boldsymbol{\mu}] \tag{T7.44}$$

Therefore from $(A.116)$ and $(A.118)$ we obtain:

$$\frac{\partial \ln f}{\partial \boldsymbol{\mu}} = \operatorname{vec} \left[\mathbf{G}'_{\boldsymbol{\mu}} \right] = -T_1 \boldsymbol{\Omega} (\boldsymbol{\mu} - \boldsymbol{\mu}_1). \tag{T7.45}$$

Similarly, from $(A.116)$ and $(A.118)$ we obtain:

$$\begin{aligned} \frac{\partial \ln f}{\partial \operatorname{vech} [\boldsymbol{\Omega}]} &= \mathbf{D}'_N \operatorname{vec} [\mathbf{G}'_{\boldsymbol{\Omega}}] \\ &= \frac{1}{2} \mathbf{D}'_N \operatorname{vec} \left[(\nu_1 - N) \boldsymbol{\Omega}^{-1} - \nu_1 \boldsymbol{\Sigma}_1 - T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)(\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \right]. \end{aligned} \tag{T7.46}$$

Applying the first-order conditions to $(T7.45)$ and $(T7.46)$ we obtain the mode of the location parameter:

$$\widetilde{\boldsymbol{\mu}} \equiv \boldsymbol{\mu}_1 \tag{T7.47}$$

and the mode of the dispersion parameter:

$$\widetilde{\boldsymbol{\Omega}}^{-1} = \frac{\nu_1}{\nu_1 - N} \boldsymbol{\Sigma}_1. \tag{T7.48}$$

To compute the modal dispersion we differentiate $(T7.40)$. Using $(A.126)$ the second differential reads:

$$\begin{aligned} d(d \ln f) = & -\frac{1}{2} \operatorname{tr} \left\{ \left((\nu_1 - N) \boldsymbol{\Omega}^{-1} (d\boldsymbol{\Omega}) \boldsymbol{\Omega}^{-1} + 2T_1 d\boldsymbol{\mu} (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \right) d\boldsymbol{\Omega} \right\} \\ & - \operatorname{tr} \{ T_1 d\boldsymbol{\mu}' \boldsymbol{\Omega} d\boldsymbol{\mu} \} - \operatorname{tr} \left\{ T_1 (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' d\boldsymbol{\Omega} d\boldsymbol{\mu} \right\} \\ = & -\frac{\nu_1 - N}{2} \operatorname{tr} \left\{ \boldsymbol{\Omega}^{-1} (d\boldsymbol{\Omega}) \boldsymbol{\Omega}^{-1} d\boldsymbol{\Omega} \right\} \\ & - T_1 d\boldsymbol{\mu}' \boldsymbol{\Omega} d\boldsymbol{\mu} - 2T_1 \operatorname{tr} \left\{ (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' d\boldsymbol{\Omega} d\boldsymbol{\mu} \right\}. \end{aligned} \tag{T7.49}$$

The first term can be expressed using (A.107), (A.106) the duplication matrix (A.113) to get rid of the redundancies of $d\mathbf{\Omega}$ as follows:

$$\begin{aligned}\operatorname{tr}\left\{\mathbf{\Omega}^{-1}\left(d\mathbf{\Omega}\right)\mathbf{\Omega}^{-1}d\mathbf{\Omega}\right\} &= \operatorname{vec}\left[d\mathbf{\Omega}\right]'\operatorname{vec}\left[\mathbf{\Omega}^{-1}\left(d\mathbf{\Omega}\right)\mathbf{\Omega}^{-1}\right] \\ &= \operatorname{vec}\left[d\mathbf{\Omega}\right]\left(\mathbf{\Omega}^{-1}\otimes\mathbf{\Omega}^{-1}\right)\operatorname{vec}\left[d\mathbf{\Omega}\right] \qquad (T7.50) \\ &= \operatorname{vec}\left[d\mathbf{\Omega}\right]\left(\mathbf{\Omega}^{-1}\otimes\mathbf{\Omega}^{-1}\right)\operatorname{vec}\left[d\mathbf{\Omega}\right] \\ &= \operatorname{vech}\left[d\mathbf{\Omega}\right]\mathbf{D}_N'\left(\mathbf{\Omega}^{-1}\otimes\mathbf{\Omega}^{-1}\right)\mathbf{D}_N\operatorname{vech}\left[d\mathbf{\Omega}\right].\end{aligned}$$

We are interested in the Hessian evaluated in the mode, where (T7.47) holds and thus the last term in (T7.49) cancels:

$$\begin{aligned}d\left(d\ln f\right)\big|_{\widetilde{\boldsymbol{\mu}},\widetilde{\mathbf{\Omega}}} = &-\frac{(\nu_1 - N)}{2}\operatorname{vech}\left[d\mathbf{\Omega}\right]\mathbf{D}_N'\left(\widetilde{\mathbf{\Omega}}^{-1}\otimes\widetilde{\mathbf{\Omega}}^{-1}\right)\mathbf{D}_N\operatorname{vech}\left[d\mathbf{\Omega}\right] \\ &-T_1 d\boldsymbol{\mu}'\widetilde{\mathbf{\Omega}}d\boldsymbol{\mu}. \qquad (T7.51)\end{aligned}$$

Therefore from (A.117) and (A.121) and substituting back (T7.48) we obtain:

$$\left.\frac{\partial^2 \ln f}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}'}\right|_{\widetilde{\boldsymbol{\mu}},\widetilde{\mathbf{\Omega}}} = -T_1\frac{\nu_1 - N}{\nu_1}\mathbf{\Sigma}_1^{-1} \qquad (T7.52)$$

$$\left.\frac{\partial^2 \ln f}{\partial\operatorname{vech}\left[\mathbf{\Omega}\right]\partial\boldsymbol{\mu}'}\right|_{\widetilde{\boldsymbol{\mu}},\widetilde{\mathbf{\Omega}}} = \mathbf{0}_{(N(N+1)/2)^2\times N^2} \qquad (T7.53)$$

$$\left.\frac{\partial^2 \ln f}{\partial\operatorname{vech}\left(\mathbf{\Omega}\right)\partial\operatorname{vech}\left(\mathbf{\Omega}\right)'}\right|_{\widetilde{\boldsymbol{\mu}},\widetilde{\mathbf{\Omega}}} = -\frac{1}{2}\frac{\nu_1^2}{\nu_1 - N}\mathbf{D}_N'\left(\mathbf{\Sigma}_1\otimes\mathbf{\Sigma}_1\right)\mathbf{D}_N \qquad (T7.54)$$

Finally the modal dispersion reads:

$$\begin{aligned}\operatorname{MDis}\left\{\boldsymbol{\theta}\right\} &\equiv \left(-\left.\frac{\partial^2 \ln f}{\partial\left(\boldsymbol{\mu},\operatorname{vech}\left(\mathbf{\Omega}\right)\right)\partial\left(\boldsymbol{\mu},\operatorname{vech}\left(\mathbf{\Omega}\right)\right)'}\right|_{\widetilde{\boldsymbol{\mu}},\widetilde{\mathbf{\Omega}}}\right)^{-1} \qquad (T7.55) \\ &= \begin{pmatrix}\mathbf{S}_{\boldsymbol{\mu}} & \mathbf{0}_{N^2\times(N(N+1)/2)^2} \\ \mathbf{0}_{(N(N+1)/2)^2\times N^2} & \mathbf{S}_{\mathbf{\Sigma}}\end{pmatrix},\end{aligned}$$

where

$$\mathbf{S}_{\boldsymbol{\mu}} \equiv \frac{1}{T_1}\frac{\nu_1}{\nu_1 - N}\mathbf{\Sigma}_1 \qquad (T7.56)$$

$$\mathbf{S}_{\mathbf{\Sigma}} \equiv \frac{2}{\nu_1}\frac{\nu_1 - N}{\nu_1}\left[\mathbf{D}_N'\left(\mathbf{\Sigma}_1\otimes\mathbf{\Sigma}_1\right)\mathbf{D}_N\right]^{-1}. \qquad (T7.57)$$

7.4 IW dispersion: mode and modal dispersion

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

We consider the notation for the IW (inverse-Wishart) assumptions (7.33) on the posterior, although of course the proof applies verbatim to the prior, or any IW distribution. Thus assume

$$\mathbf{\Sigma} \sim \mathrm{IW}\left(\nu_1, \nu_1 \mathbf{\Sigma}_1\right). \tag{$T7.58$}$$

From (2.233) the probability density function of $\mathbf{\Sigma}$ reads:

$$f\left(\mathbf{\Sigma}\right) = \frac{1}{\kappa} \left|\nu_1 \mathbf{\Sigma}_1\right|^{\frac{\nu_1}{2}} \left|\mathbf{\Sigma}\right|^{-\frac{\nu_1+N+1}{2}} e^{-\frac{1}{2}\operatorname{tr}\left(\nu_1 \mathbf{\Sigma}_1 \mathbf{\Sigma}^{-1}\right)}. \tag{$T7.59$}$$

To determine the mode of this distribution we impose the first-order conditions on the logarithm of the joint probability density function (T7.59).

$$\ln f \equiv \gamma_2 - \frac{\nu_1 + N + 1}{2} \ln\left|\mathbf{\Sigma}\right| - \frac{1}{2}\operatorname{tr}\left\{\nu_1 \mathbf{\Sigma}_1 \mathbf{\Sigma}^{-1}\right\}. \tag{$T7.60$}$$

Computing the first variation and using (A.124) and (A.126):

$$\begin{aligned} d\ln f &= -\frac{\nu_1+N+1}{2}\operatorname{tr}\left(\mathbf{\Sigma}^{-1} d\mathbf{\Sigma}\right) \\ &\quad +\frac{1}{2}\operatorname{tr}\left\{\nu_1 \mathbf{\Sigma}_1 \mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\right\} \\ &= \operatorname{tr}\left(\frac{1}{2}\left(\nu_1 \mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1} - \left(\nu_1+N+1\right)\mathbf{\Sigma}^{-1}\right) d\mathbf{\Sigma}\right) \end{aligned} \tag{$T7.61$}$$

where

$$\mathbf{G} \equiv \frac{1}{2}\left(\nu_1 \mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1} - \left(\nu_1+N+1\right)\mathbf{\Sigma}^{-1}\right). \tag{$T7.62$}$$

Using (A.120) and the duplication matrix (A.113) to get rid of the redundancies of $d\mathbf{\Sigma}$ we obtain:

$$d\ln f = \operatorname{vec}\left[\mathbf{G}'\right]' \mathbf{D}_N \operatorname{vech}\left[d\mathbf{\Sigma}\right] \tag{$T7.63$}$$

Therefore from (A.116) and (A.118) we obtain:

$$\begin{aligned} \frac{\partial \ln f}{\partial \operatorname{vech}\left[\mathbf{\Sigma}\right]} &= \mathbf{D}_N' \operatorname{vec}\left[\mathbf{G}'\right] \\ &= \frac{1}{2}\mathbf{D}_N' \operatorname{vec}\left[\nu_1 \mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1} - \left(\nu_1+N+1\right)\mathbf{\Sigma}^{-1}\right]. \end{aligned} \tag{$T7.64$}$$

Applying the first-order conditions to (T7.64) we obtain the mode:

$$\mathrm{Mod}_{i_T, e_C} = \frac{\nu_1}{\nu_1+N+1}\operatorname{vech}\left[\mathbf{\Sigma}_1\right]. \tag{$T7.65$}$$

To compute the modal dispersion we differentiate (T7.61). Using (A.126) we obtain:

$$\begin{aligned} d\left(d\ln f\right) = & -\frac{1}{2}\operatorname{tr}\left(\nu_1 \mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}d\mathbf{\Sigma}\right) \\ & -\frac{1}{2}\operatorname{tr}\left(\nu_1 \mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}d\mathbf{\Sigma}\right) \\ & +\frac{1}{2}\operatorname{tr}\left(\left(\nu_1+N+1\right)\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}d\mathbf{\Sigma}\right) \\ = & -\nu_1\operatorname{tr}\left(\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\right) \\ & +\frac{\nu_1+N+1}{2}\operatorname{tr}\left(\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\right) \end{aligned} \tag{T7.66}$$

Using $(A.107)$ and $(A.106)$ and the duplication matrix $(A.113)$ to get rid of the redundancies of $d\mathbf{\Sigma}$ we can write:

$$\begin{aligned} a & \equiv \operatorname{tr}\left(\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\right) \\ & = \operatorname{vec}\left[d\mathbf{\Sigma}\right]'\operatorname{vec}\left[\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\right] \\ & = \operatorname{vec}\left[d\mathbf{\Sigma}\right]'\left(\mathbf{\Sigma}^{-1}\otimes\mathbf{\Sigma}^{-1}\right)\operatorname{vec}\left[d\mathbf{\Sigma}\right] \\ & = \operatorname{vech}\left[d\mathbf{\Sigma}\right]'\mathbf{D}_N'\left(\mathbf{\Sigma}^{-1}\otimes\mathbf{\Sigma}^{-1}\right)\mathbf{D}_N\operatorname{vech}\left[d\mathbf{\Sigma}\right]. \end{aligned} \tag{T7.67}$$

Similarly, using $(A.107)$ and $(A.106)$ and the duplication matrix $(A.113)$ to get rid of the redundancies of $d\mathbf{\Sigma}$ we can write:

$$\begin{aligned} b & \equiv \operatorname{tr}\left(\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\right) \\ & = \operatorname{vec}\left[d\mathbf{\Sigma}\right]'\operatorname{vec}\left[\mathbf{\Sigma}^{-1}\left(d\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\right] \\ & = \operatorname{vec}\left[d\mathbf{\Sigma}\right]'\left(\left(\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\right)\otimes\mathbf{\Sigma}^{-1}\right)\operatorname{vec}\left[d\mathbf{\Sigma}\right] \\ & = \operatorname{vech}\left[d\mathbf{\Sigma}\right]'\mathbf{D}_N'\left(\left(\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\right)\otimes\mathbf{\Sigma}^{-1}\right)\mathbf{D}_N\operatorname{vech}\left[d\mathbf{\Sigma}\right]. \end{aligned} \tag{T7.68}$$

Therefore, substituting $(T7.67)$ and $(T7.68)$ in $(T7.66)$ we obtain

$$\begin{aligned} d\left(d\ln f\right) & = -\nu_1 b + \frac{\nu_1+N+1}{2}a \\ & = \operatorname{vech}\left[d\mathbf{\Sigma}\right]'\mathbf{H}\operatorname{vech}\left[d\mathbf{\Sigma}\right], \end{aligned} \tag{T7.69}$$

where

$$\begin{aligned} \mathbf{H} \equiv & -\nu_1\mathbf{D}_N'\left(\left(\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_1\mathbf{\Sigma}^{-1}\right)\otimes\mathbf{\Sigma}^{-1}\right)\mathbf{D}_N \\ & +\frac{\nu_1+N+1}{2}\mathbf{D}_N'\left(\mathbf{\Sigma}^{-1}\otimes\mathbf{\Sigma}^{-1}\right)\mathbf{D}_N \end{aligned} \tag{T7.70}$$

We are interested in the Hessian evaluated in the mode, where $(T7.65)$ holds, i.e. in the point

$$\mathbf{\Sigma} \equiv \frac{\nu_1}{\nu_1+N+1}\mathbf{\Sigma}_1. \tag{T7.71}$$

In this point

$$d\left(d\ln f\right)\big|_{\text{Mod}} = \operatorname{vech}\left[d\mathbf{\Sigma}\right]' \mathbf{H}|_{\text{Mod}} \operatorname{vech}\left[d\mathbf{\Sigma}\right], \qquad (T7.72)$$

where

$$\begin{aligned}\mathbf{H}|_{\text{Mod}} \equiv & -\nu_1 \left(\frac{\nu_1+N+1}{\nu_1}\right)^3 \mathbf{D}_N' \left(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{\Sigma}_1^{-1}\right)\mathbf{D}_N \qquad (T7.73)\\ & + \frac{\nu_1+N+1}{2}\left(\frac{\nu_1+N+1}{\nu_1}\right)^2 \mathbf{D}_N' \left(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{\Sigma}_1^{-1}\right)\mathbf{D}_N\\ & = -\frac{1}{2}\frac{\left(\nu_1+N+1\right)^3}{\nu_1^2}\mathbf{D}_N' \left(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{\Sigma}_1^{-1}\right)\mathbf{D}_N.\end{aligned}$$

Therefore

$$\left.\frac{\partial^2 \ln f\left[\mathbf{\Sigma}\right]}{\partial \operatorname{vech}\left[\mathbf{\Sigma}\right]\partial \operatorname{vech}\left[\mathbf{\Sigma}\right]'}\right|_{\text{Mod}} = -\frac{1}{2}\frac{\left(\nu_1+N+1\right)^3}{\nu_1^2}\mathbf{D}_N' \left(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{\Sigma}_1^{-1}\right)\mathbf{D}_N \qquad (T7.74)$$

From the definition of modal dispersion (2.65)

$$\begin{aligned}\text{MDis} & \equiv -\left(\left.\frac{\partial^2 \ln f\left[\mathbf{\Sigma}\right]}{\partial \operatorname{vech}\left[\mathbf{\Sigma}\right]\partial \operatorname{vech}\left[\mathbf{\Sigma}\right]'}\right|_{\text{Mod}}\right)^{-1} \qquad (T7.75)\\ & = \frac{2\nu_1^2}{\left(\nu_1+N+1\right)^3}\left(\mathbf{D}_N' \left(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{\Sigma}_1^{-1}\right)\mathbf{D}_N\right)^{-1}\end{aligned}$$

7.5 NIW location-dispersion: marginal distribution of location

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

We consider the notation for the NIW (normal-inverse-Wishart) assumptions (7.20)-(7.21) on the prior, although of course the proof applies verbatim to the posterior, or any NIW distribution. Thus we assume:

$$\mathbf{\Omega} \equiv \mathbf{\Sigma}^{-1} \sim \text{W}\left(\nu_0, \left(\nu_0\mathbf{\Sigma}_0\right)^{-1}\right) \qquad (T7.76)$$

and

$$\boldsymbol{\mu}|\mathbf{\Omega} \sim \text{N}\left(\boldsymbol{\mu}_0, \left(T_0\mathbf{\Omega}\right)^{-1}\right), \qquad (T7.77)$$

From (2.156) and (2.224) the joint prior pdf of $\boldsymbol{\mu}$ and $\mathbf{\Omega}$ is

$$\begin{aligned}f\left(\boldsymbol{\mu},\mathbf{\Omega}\right) & = f\left(\boldsymbol{\mu}|\mathbf{\Omega}\right) f\left(\mathbf{\Omega}\right) \qquad (T7.78)\\ & = \gamma_1 \left|\mathbf{\Omega}\right|^{\frac{1}{2}} e^{-\frac{1}{2}\left(\boldsymbol{\mu}-\boldsymbol{\mu}_0\right)'\left(T_0\mathbf{\Omega}\right)\left(\boldsymbol{\mu}-\boldsymbol{\mu}_0\right)}\\ & \quad \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Omega}\right|^{\frac{\nu_0-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}\left(\nu_0\mathbf{\Sigma}_0\mathbf{\Omega}\right)},\end{aligned}$$

To determine the unconditional pdf of $\boldsymbol{\mu}$ we have to compute the marginal in ($T7.78$). Defining

$$\boldsymbol{\Sigma}_2 \equiv \nu_0 \boldsymbol{\Sigma}_0 + T_0 \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right) \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right)' \tag{$T7.79$}$$

we obtain

$$\begin{aligned} f(\boldsymbol{\mu}) &\equiv \int f(\boldsymbol{\mu}, \boldsymbol{\Omega})\, d\boldsymbol{\Omega} \\ &= \int \gamma_1 \left|\boldsymbol{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\boldsymbol{\Omega}\right|^{\frac{\nu_0 - N}{2}} e^{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}_2 \boldsymbol{\Omega})} d\boldsymbol{\Omega} \\ &= \gamma_2 \left|\boldsymbol{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\boldsymbol{\Sigma}_2\right|^{-\frac{\nu_0+1}{2}} \\ &\quad \left\{ \int \gamma_3 \left|\boldsymbol{\Sigma}_2\right|^{\frac{\nu_0+1}{2}} \left|\boldsymbol{\Omega}\right|^{\frac{\nu_0 - N}{2}} e^{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}_2 \boldsymbol{\Omega})} d\boldsymbol{\Omega} \right\} \\ &= \gamma_2 \left|\boldsymbol{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\boldsymbol{\Sigma}_2\right|^{-\frac{\nu_0+1}{2}}, \end{aligned} \tag{$T7.80$}$$

where we have used the fact that the term in curly brackets is the integrals of the Wishart pdf (2.224) over the entire space and thus it sums to one.

Thus substituting again ($T7.79$) we obtain that the marginal pdf ($T7.80$) reads:

$$f(\boldsymbol{\mu}) = \gamma_2 \left|\boldsymbol{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\nu_0 \boldsymbol{\Sigma}_0 + T_0 \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right) \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right)'\right|^{-\frac{\nu_0+1}{2}}, \tag{$T7.81$}$$

From ($A.91$) we obtain the following identity:

$$\begin{aligned} \left|\boldsymbol{\Sigma}\right|^{\frac{\nu_0}{2}} \left|\boldsymbol{\Sigma} + \mathbf{v}\mathbf{v}'\right|^{-\frac{\nu_0+1}{2}} &= \left|\boldsymbol{\Sigma}\right|^{\frac{\nu_0}{2}} \left[\left|\boldsymbol{\Sigma}\right| \left|\mathbf{I} + \boldsymbol{\Sigma}^{-1}\mathbf{v}\mathbf{v}'\right|\right]^{-\frac{\nu_0+1}{2}} \\ &= \left|\boldsymbol{\Sigma}\right|^{\frac{1}{2}} \left|\mathbf{I} + \boldsymbol{\Sigma}^{-1}\mathbf{v}\mathbf{v}'\right|^{-\frac{\nu_0+1}{2}} \\ &= \left|\boldsymbol{\Sigma}\right|^{-\frac{1}{2}} \left(1 + \mathbf{v}'\boldsymbol{\Sigma}^{-1}\mathbf{v}\right)^{-\frac{\nu_0+1}{2}} \end{aligned} \tag{$T7.82$}$$

Applying this result to

$$\mathbf{v} \equiv \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right) \sqrt{T_0} \tag{$T7.83$}$$

$$\boldsymbol{\Sigma} \equiv \nu_0 \boldsymbol{\Sigma}_0, \tag{$T7.84$}$$

we reduce ($T7.81$) to the following expression:

$$f(\boldsymbol{\mu}) = \gamma_4 \left|\frac{\boldsymbol{\Sigma}_0}{T_0}\right|^{-\frac{1}{2}} \left|1 + \frac{1}{\nu_0} \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right)' \left(\frac{\boldsymbol{\Sigma}_0}{T_0}\right)^{-1} \left(\boldsymbol{\mu} - \boldsymbol{\mu}_0\right)\right|^{-\frac{\nu_0+1}{2}}. \tag{$T7.85$}$$

By comparison with (2.188) we see that this is a multivariate Student t distribution with the following parameters:

$$\boldsymbol{\mu} \sim \operatorname{St}\left(\nu_0, \boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}_0}{T_0}\right). \tag{$T7.86$}$$

7.6 NIW factor loadings-dispersion: posterior distribution

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

By the NIW (normal-inverse-Wishart) assumptions (4.129)-(7.59) on the prior we have:

$$\mathbf{\Omega} \equiv \mathbf{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_0, (\nu_0 \mathbf{\Sigma}_0)^{-1}\right) \tag{T7.87}$$

and

$$\mathbf{B}|\mathbf{\Sigma} \sim \mathrm{N}\left(\mathbf{B}_0, \frac{\mathbf{\Sigma}}{T_0}, \mathbf{\Sigma}_{F,0}^{-1}\right), \tag{T7.88}$$

or, in terms of $\mathbf{\Omega} \equiv \mathbf{\Sigma}^{-1}$:

$$\mathbf{B}|\mathbf{\Omega} \sim \mathrm{N}\left(\mathbf{B}_0, (T_0 \mathbf{\Omega})^{-1}, \mathbf{\Sigma}_{F,0}^{-1}\right). \tag{T7.89}$$

Thus from (2.182) and (2.224) the joint prior pdf of $\mathbf{B}$ and $\mathbf{\Omega}$ is:

$$\begin{aligned} f_{\mathrm{pr}}(\mathbf{B}, \mathbf{\Omega}) &= f_{\mathrm{pr}}(\mathbf{B}|\mathbf{\Omega}) f_{\mathrm{pr}}(\mathbf{\Omega}) \\ &= \gamma_1 |\mathbf{\Omega}|^{\frac{K}{2}} |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} e^{-\frac{1}{2}\mathrm{tr}\left\{(T_0\mathbf{\Omega})(\mathbf{B}-\mathbf{B}_0)\mathbf{\Sigma}_{F,0}(\mathbf{B}-\mathbf{B}_0)'\right\}} \\ &\quad |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} |\mathbf{\Omega}|^{\frac{\nu_0-N-1}{2}} e^{-\frac{1}{2}\mathrm{tr}(\nu_0 \mathbf{\Sigma}_0 \mathbf{\Omega})}, \end{aligned} \tag{T7.90}$$

The current information conditioned on the parameters $\mathbf{B}$ and $\mathbf{\Omega}$ is summarized by the OLS factor loadings and sample covariance.

$$i_T \equiv \left\{\widehat{\mathbf{B}}, \widehat{\mathbf{\Sigma}}\right\}. \tag{T7.91}$$

In (4.129) we showed that $\widehat{\mathbf{B}}$ is distributed as follows:

$$\widehat{\mathbf{B}} \sim \mathrm{N}\left(\mathbf{B}, \frac{\mathbf{\Sigma}}{T}, \widehat{\mathbf{\Sigma}}_F^{-1}\right), \tag{T7.92}$$

and in (4.130) we showed that $T\widehat{\mathbf{\Sigma}}$ is distributed as follows:

$$T\widehat{\mathbf{\Sigma}} \sim \mathrm{W}(T-K, \mathbf{\Sigma}). \tag{T7.93}$$

Furthermore, we showed that $\widehat{\mathbf{B}}$ and $T\widehat{\mathbf{\Sigma}}$ are independent. Therefore, from (2.182) and (2.224) we have

$$\begin{aligned} f(i_T|\mathbf{B}, \mathbf{\Omega}) &= f\left(\widehat{\mathbf{B}}|\mathbf{B}, \mathbf{\Omega}\right) f\left(T\widehat{\mathbf{\Sigma}}|\mathbf{B}, \mathbf{\Omega}\right) \\ &= \gamma_2 |T\mathbf{\Omega}|^{\frac{K}{2}} \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} e^{-\frac{1}{2}\mathrm{tr}\left\{(T\mathbf{\Omega})\left(\widehat{\mathbf{B}}-\mathbf{B}\right)\widehat{\mathbf{\Sigma}}_F\left(\widehat{\mathbf{B}}-\mathbf{B}\right)'\right\}} \\ &\quad |\mathbf{\Omega}|^{\frac{T-K}{2}} \left|T\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} e^{-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Omega} T\widehat{\mathbf{\Sigma}}\right)}. \end{aligned} \tag{T7.94}$$

Thus, after trivial regrouping and simplifications, the joint pdf of current information and the parameters reads:

$$\begin{aligned} f\left(i_T,\mathbf{B},\mathbf{\Omega}\right) &= f\left(i_T|\mathbf{B},\mathbf{\Omega}\right) f_{\mathrm{pr}}\left(\mathbf{B},\mathbf{\Omega}\right) \qquad (T7.95)\\ &= \gamma_3 \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} \left|\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} \left|\mathbf{\Sigma}_{F,0}\right|^{\frac{N}{2}} \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Omega}\right|^{\frac{T+K+\nu_0-N-1}{2}} \\ &\quad e^{-\frac{1}{2}\operatorname{tr}\left\{\mathbf{\Omega}(\mathbf{B}-\mathbf{B}_0)T_0\mathbf{\Sigma}_{F,0}(\mathbf{B}-\mathbf{B}_0)'+\mathbf{\Omega}\left(\widehat{\mathbf{B}}-\mathbf{B}\right)T\widehat{\mathbf{\Sigma}}_F\left(\widehat{\mathbf{B}}-\mathbf{B}\right)'\right\}} \\ &\quad e^{-\frac{1}{2}\operatorname{tr}\left(\mathbf{\Omega}\left(T\widehat{\mathbf{\Sigma}}+\nu_0\mathbf{\Sigma}_0\right)\right)} \end{aligned}$$

We show below that the terms in curly brackets in $(T7.95)$ can be re-written as follows:

$$\{\cdots\} = T_1\mathbf{\Omega}\left(\mathbf{B}-\mathbf{B}_1\right)\mathbf{\Sigma}_{F,1}\left(\mathbf{B}-\mathbf{B}_1\right)' + \mathbf{\Omega}\mathbf{\Phi}, \qquad (T7.96)$$

where

$$T_1 \equiv T_0 + T \qquad (T7.97)$$

$$\mathbf{\Sigma}_{F,1} \equiv \frac{T_0\mathbf{\Sigma}_{F,0}+T\widehat{\mathbf{\Sigma}}_F}{T_1} \qquad (T7.98)$$

$$\mathbf{B}_1 \equiv \left(\mathbf{B}_0T_0\mathbf{\Sigma}_{F,0}+\widehat{\mathbf{B}}T\widehat{\mathbf{\Sigma}}_F\right)\left(T_0\mathbf{\Sigma}_{F,0}+T\widehat{\mathbf{\Sigma}}_F\right)^{-1} \qquad (T7.99)$$

$$\begin{aligned} \mathbf{\Phi} &\equiv \mathbf{B}_0T_0\mathbf{\Sigma}_{F,0}\mathbf{B}_0' + \widehat{\mathbf{B}}T\widehat{\mathbf{\Sigma}}_F\widehat{\mathbf{B}}' \qquad (T7.100)\\ &\quad -\left(\mathbf{B}_0T_0\mathbf{\Sigma}_{F,0}+\widehat{\mathbf{B}}T\widehat{\mathbf{\Sigma}}_F\right)\left(T_0\mathbf{\Sigma}_{F,0}+T\widehat{\mathbf{\Sigma}}_F\right)^{-1}\left(T_0\mathbf{\Sigma}_{F,0}\mathbf{B}_0'+T\widehat{\mathbf{\Sigma}}_F\widehat{\mathbf{B}}'\right). \end{aligned}$$

Indeed, defining $\mathbf{D} \equiv T_0\mathbf{\Sigma}_{F,0}$ and $\mathbf{C} \equiv T\widehat{\mathbf{\Sigma}}_F$ we can write the expression in curly brackets $(T7.96)$ as $\mathbf{\Omega}\mathbf{A}$, where

$$\begin{aligned} \mathbf{A} &\equiv \left(\mathbf{B}-\mathbf{B}_0\right)\mathbf{D}\left(\mathbf{B}-\mathbf{B}_0\right)' + \left(\mathbf{B}-\widehat{\mathbf{B}}\right)\mathbf{C}\left(\mathbf{B}-\widehat{\mathbf{B}}\right)' \\ &= \mathbf{BDB}' + \mathbf{B}_0\mathbf{DB}_0' - 2\mathbf{BDB}_0' + \mathbf{BCB}' + \widehat{\mathbf{B}}\mathbf{C}\widehat{\mathbf{B}}' - 2\mathbf{BC}\widehat{\mathbf{B}}' \\ &= \mathbf{B}\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}' + \mathbf{B}_0\mathbf{DB}_0' - 2\mathbf{BDB}_0' + \widehat{\mathbf{B}}\mathbf{C}\widehat{\mathbf{B}}' - 2\mathbf{BC}\widehat{\mathbf{B}}' \qquad (T7.101)\\ &= \mathbf{B}\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}' + \mathbf{B}_1\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}_1' - 2\mathbf{B}\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}_1' \\ &\quad +\mathbf{B}_0\mathbf{DB}_0' - 2\mathbf{BDB}_0' + \widehat{\mathbf{B}}\mathbf{C}\widehat{\mathbf{B}}' - 2\mathbf{BC}\widehat{\mathbf{B}}' \\ &\quad -\mathbf{B}_1\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}_1' + 2\mathbf{B}\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}_1' \\ &= \left(\mathbf{B}-\mathbf{B}_1\right)\left(\mathbf{D}+\mathbf{C}\right)\left(\mathbf{B}-\mathbf{B}_1\right)' \\ &\quad +\mathbf{B}_0\mathbf{DB}_0' + \widehat{\mathbf{B}}\mathbf{C}\widehat{\mathbf{B}}' - 2\mathbf{BDB}_0' - 2\mathbf{BC}\widehat{\mathbf{B}}' \\ &\quad -\mathbf{B}_1\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}_1' + 2\mathbf{B}\left(\mathbf{D}+\mathbf{C}\right)\mathbf{B}_1' \end{aligned}$$

defining

$$\mathbf{B}_1 \equiv \left(\mathbf{B}_0\mathbf{D}+\widehat{\mathbf{B}}\mathbf{C}\right)\left(\mathbf{D}+\mathbf{C}\right)^{-1} \qquad (T7.102)$$

the above simplifies to

$$\begin{aligned}\mathbf{A} = & \left(\mathbf{B}-\mathbf{B}_1\right)\left(\mathbf{D}+\mathbf{C}\right)\left(\mathbf{B}-\mathbf{B}_1\right)' \\ & +\mathbf{B}_0\mathbf{D}\mathbf{B}_0' + \widehat{\mathbf{B}}\mathbf{C}\widehat{\mathbf{B}}' - \left(\mathbf{B}_0\mathbf{D}+\widehat{\mathbf{B}}\mathbf{C}\right)\left(\mathbf{D}+\mathbf{C}\right)^{-1}\left(\mathbf{B}_0\mathbf{D}+\widehat{\mathbf{B}}\mathbf{C}\right)',\end{aligned} \tag{T7.103}$$

which proves the result.

Substituting (T7.96) in (T7.95) and defining

$$\nu_1 \equiv T + \nu_0 \tag{T7.104}$$

$$\mathbf{\Sigma}_1 \equiv \frac{T\widehat{\mathbf{\Sigma}} + \nu_0\mathbf{\Sigma}_0 + \mathbf{\Phi}}{\nu_1} \tag{T7.105}$$

we obtain:

$$\begin{aligned}f\left(i_T,\mathbf{B},\mathbf{\Omega}\right) = & f\left(i_T|\mathbf{B},\mathbf{\Omega}\right) f_{\text{pr}}\left(\mathbf{B},\mathbf{\Omega}\right) \\ = & \gamma_3 \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} \left|\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} \left|\mathbf{\Sigma}_{F,0}\right|^{\frac{N}{2}} \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Omega}\right|^{\frac{T+K+\nu_0-N-1}{2}} \\ & e^{-\frac{1}{2}\operatorname{tr}\left\{T_1\mathbf{\Omega}(\mathbf{B}-\mathbf{B}_1)\mathbf{\Sigma}_{F,1}(\mathbf{B}-\mathbf{B}_1)'\right\}} e^{-\frac{1}{2}\operatorname{tr}(\mathbf{\Omega}\nu_1\mathbf{\Sigma}_1)}.\end{aligned} \tag{T7.106}$$

At this point we can perform the integration over $(\mathbf{B},\mathbf{\Omega})$ to determine the marginal pdf $f\left(i_T\right)$

$$\begin{aligned}f\left(i_T\right) = & \int f\left(i_T,\mathbf{B},\mathbf{\Omega}\right) d\mathbf{B}d\mathbf{\Omega} \\ = & \gamma_4 \int \left\{\int \gamma_5 \left|\mathbf{\Omega}\right|^{\frac{K}{2}} \left|\mathbf{\Sigma}_{F,1}\right|^{\frac{N}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{T_1\mathbf{\Omega}(\mathbf{B}-\mathbf{B}_1)\mathbf{\Sigma}_{F,1}(\mathbf{B}-\mathbf{B}_1)'\right\}} d\mathbf{B}\right\} \\ & \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} \left|\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} \left|\mathbf{\Sigma}_{F,1}\right|^{-\frac{N}{2}} \left|\mathbf{\Sigma}_{F,0}\right|^{\frac{N}{2}} \\ & \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Omega}\right|^{\frac{\nu_1-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\mathbf{\Omega}\boldsymbol{\nu}_1\mathbf{\Sigma}_1)} d\mathbf{\Omega} \\ = & \gamma_4 \int \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} \left|\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} \left|\mathbf{\Sigma}_{F,1}\right|^{-\frac{N}{2}} \left|\mathbf{\Sigma}_{F,0}\right|^{\frac{N}{2}} \\ & \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Omega}\right|^{\frac{\nu_1-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\mathbf{\Omega}\boldsymbol{\nu}_1\mathbf{\Sigma}_1)} d\mathbf{\Omega},\end{aligned} \tag{T7.107}$$

where we used the fact that the expression in curly brackets is the integral of the pdf of a matrix-valued normal distribution (2.182) over the entire space and thus sums to one. Thus we can write (T7.107) as follows:

$$\begin{aligned}f\left(i_T\right) = & \gamma_5 \left\{\int \gamma_6 \left|\mathbf{\Sigma}_1\right|^{\frac{\nu_1}{2}} \left|\mathbf{\Omega}\right|^{\frac{\nu_1-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\nu_1\mathbf{\Sigma}_1\mathbf{\Omega})} d\mathbf{\Omega}\right\} \\ & \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} \left|\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} \left|\mathbf{\Sigma}_{F,1}\right|^{-\frac{N}{2}} \left|\mathbf{\Sigma}_{F,0}\right|^{\frac{N}{2}} \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Sigma}_1\right|^{-\frac{\nu_1}{2}} \\ = & \gamma_5 \left|\widehat{\mathbf{\Sigma}}_F\right|^{\frac{N}{2}} \left|\widehat{\mathbf{\Sigma}}\right|^{\frac{T-K-N-1}{2}} \left|\mathbf{\Sigma}_{F,1}\right|^{-\frac{N}{2}} \left|\mathbf{\Sigma}_{F,0}\right|^{\frac{N}{2}} \left|\mathbf{\Sigma}_0\right|^{\frac{\nu_0}{2}} \left|\mathbf{\Sigma}_1\right|^{-\frac{\nu_1}{2}},\end{aligned} \tag{T7.108}$$

where we used the fact that the term in curly brackets is the integral of the pdf of a Wishart distribution (2.224) over the entire space and thus sums to one.

Finally, we obtain the posterior pdf (7.15) by dividing the joint pdf by the marginal pdf:

$$\begin{aligned} f_{\text{po}}(\mathbf{B},\boldsymbol{\Omega}) &\equiv \frac{f(i_T,\mathbf{B},\boldsymbol{\Omega})}{f(i_T)} \\ &= \gamma_7 \left|\boldsymbol{\Sigma}_{F,1}\right|^{\frac{N}{2}} \left|\boldsymbol{\Sigma}_1\right|^{\frac{\nu_1}{2}} \left|\boldsymbol{\Omega}\right|^{\frac{K+\nu_1-N-1}{2}} \\ &\quad e^{-\frac{1}{2}\operatorname{tr}\left\{T_1\boldsymbol{\Omega}(\mathbf{B}-\mathbf{B}_1)\boldsymbol{\Sigma}_{F,1}(\mathbf{B}-\mathbf{B}_1)'\right\}} e^{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Omega}\nu_1\boldsymbol{\Sigma}_1)} \\ &= \gamma_8 \left|\boldsymbol{\Omega}\right|^{\frac{K}{2}} \left|\boldsymbol{\Sigma}_{F,1}\right|^{\frac{N}{2}} e^{-\frac{1}{2}\operatorname{tr}\left\{T_1\boldsymbol{\Omega}(\mathbf{B}-\mathbf{B}_1)\boldsymbol{\Sigma}_{F,1}(\mathbf{B}-\mathbf{B}_1)'\right\}} \\ &\quad \gamma_9 \left|\boldsymbol{\Sigma}_1\right|^{\frac{\nu_1}{2}} \left|\boldsymbol{\Omega}\right|^{\frac{\nu_1-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Omega}\nu_1\boldsymbol{\Sigma}_1)} \end{aligned} \tag{T7.109}$$

From (2.182) and (2.224) this proves that:

$$\mathbf{B}|\boldsymbol{\Omega} \sim \mathrm{N}\left(\mathbf{B}_1, (\boldsymbol{\Omega}T_1)^{-1}, \boldsymbol{\Sigma}_{F,1}^{-1}\right) \tag{T7.110}$$

and

$$\boldsymbol{\Omega} \sim \mathrm{W}\left(\nu_1, (\nu_1\boldsymbol{\Sigma}_1)^{-1}\right). \tag{T7.111}$$

Recalling that $\boldsymbol{\Sigma}^{-1} \equiv \boldsymbol{\Omega}$ this means:

$$\boldsymbol{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_1, \frac{\boldsymbol{\Sigma}_1^{-1}}{\nu_1}\right) \tag{T7.112}$$

and

$$\mathbf{B}|\boldsymbol{\Sigma} \sim \mathrm{N}\left(\mathbf{B}_1, \frac{\boldsymbol{\Sigma}}{T_1}, \boldsymbol{\Sigma}_{F,1}^{-1}\right). \tag{T7.113}$$

7.7 NIW factor loadings-dispersion: mode and modal dispersion

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants. We consider the notation for the NIW (normal-inverse-Wishart) assumptions (7.71)-(7.72) on the posterior, although of course the proof applies verbatim to the prior, or any NIW distribution. Thus assume

The parameter in this context are

$$\boldsymbol{\theta} \equiv \left(\operatorname{vec}[\mathbf{B}]', \operatorname{vech}[\boldsymbol{\Omega}]'\right)'. \tag{T7.114}$$

Assume that $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are joint NIW (normal-inverse-Wishart) distributed, i.e.:

$$\boldsymbol{\Sigma}^{-1} \equiv \boldsymbol{\Omega} \sim \mathrm{W}\left(\nu_1, (\nu_1\boldsymbol{\Sigma}_1)^{-1}\right). \tag{T7.115}$$

and

$$\mathbf{B}|\boldsymbol{\Omega} \sim \mathrm{N}\left(\mathbf{B}_1, (T_1\boldsymbol{\Omega})^{-1}, \boldsymbol{\Sigma}_{F,1}^{-1}\right) \tag{T7.116}$$

From from (2.182) and (2.224) the joint NIW (normal-inverse-Wishart) probability density function of $\mathbf{B}$ and $\mathbf{\Omega}$ reads:

$$\begin{aligned} f(\mathbf{B}, \mathbf{\Omega}) &= f(\mathbf{B}|\mathbf{\Omega}) f(\mathbf{\Omega}) \\ &= \gamma_1 |\mathbf{\Sigma}_{F,1}|^{\frac{N}{2}} |\mathbf{\Sigma}_1|^{\frac{\nu_1}{2}} |\mathbf{\Omega}|^{\frac{\nu_1+K-N-1}{2}} \\ &\quad e^{-\frac{1}{2}\operatorname{tr}\left\{\mathbf{\Omega}\left[(\mathbf{B}-\mathbf{B}_1)T_1\mathbf{\Sigma}_{F,1}(\mathbf{B}-\mathbf{B}_1)'+\nu_1\mathbf{\Sigma}_1\right]\right\}} \end{aligned} \tag{T7.117}$$

To determine the mode of this distribution

$$\widetilde{\boldsymbol{\theta}} \equiv \left(\operatorname{vec}\left[\widetilde{\mathbf{B}}\right]', \operatorname{vech}\left[\widetilde{\mathbf{\Omega}}\right]'\right)' \tag{T7.118}$$

we impose the first-order condition on the logarithm of the joint probability density function ($T7.117$):

$$\begin{aligned} \ln f \equiv{} & \gamma_2 + \frac{\nu_1 + K - N - 1}{2} \ln |\mathbf{\Omega}| \\ & -\frac{1}{2} \operatorname{tr}\left\{\left[(\mathbf{B} - \mathbf{B}_1) T_1 \mathbf{\Sigma}_{F,1} (\mathbf{B} - \mathbf{B}_1)' + \nu_1 \mathbf{\Sigma}_1\right] \mathbf{\Omega}\right\} \end{aligned} \tag{T7.119}$$

To compute the first variation we use ($A.124$) obtaining:

$$\begin{aligned} d \ln f \equiv{} & \frac{1}{2} \operatorname{tr}\left\{\left[(\nu_1 + K - N - 1) \mathbf{\Omega}^{-1} - \mathbf{A}\right] d\mathbf{\Omega}\right\} \\ & - \operatorname{tr}\left\{T_1 \mathbf{\Sigma}_{F,1} (\mathbf{B} - \mathbf{B}_1)' \mathbf{\Omega} d\mathbf{B}\right\}, \end{aligned} \tag{T7.120}$$

where

$$\mathbf{A} \equiv (\mathbf{B} - \mathbf{B}_1) T_1 \mathbf{\Sigma}_{F,1} (\mathbf{B} - \mathbf{B}_1)' + \nu_1 \mathbf{\Sigma}_1 \tag{T7.121}$$

Therefore

$$d \ln f \equiv \operatorname{tr}\{\mathbf{G}_{\mathbf{\Omega}} d\mathbf{\Omega}\} + \operatorname{tr}\{\mathbf{G}_{\mathbf{B}} d\mathbf{B}\}, \tag{T7.122}$$

where

$$\mathbf{G}_{\mathbf{\Omega}} \equiv \frac{1}{2}\left[(\nu_1 + K - N - 1) \mathbf{\Omega}^{-1} - \mathbf{A}\right] \tag{T7.123}$$

$$\mathbf{G}_{\mathbf{B}} \equiv -T_1 \mathbf{\Sigma}_{F,1} (\mathbf{B} - \mathbf{B}_1)' \mathbf{\Omega} \tag{T7.124}$$

Using ($A.120$) and the duplication matrix ($A.113$) to get rid of the redundancies of $d\mathbf{\Omega}$ in ($T7.122$) we obtain:

$$d \ln f = \operatorname{vec}[\mathbf{G}'_{\mathbf{\Omega}}]' \mathbf{D}_N \operatorname{vech}[d\mathbf{\Omega}] + \operatorname{vec}[\mathbf{G}'_{\mathbf{B}}]' \operatorname{vec}[d\mathbf{B}] \tag{T7.125}$$

Therefore from ($A.116$) and ($A.118$) we obtain:

$$\frac{\partial \ln f}{\partial \operatorname{vec}[\mathbf{B}]} = \operatorname{vec}[\mathbf{G}'_{\mathbf{B}}] = -T_1 \operatorname{vec}[\mathbf{\Omega}(\mathbf{B} - \mathbf{B}_1) \mathbf{\Sigma}_{F,1}]. \tag{T7.126}$$

Similarly, from ($A.116$) and ($A.118$) we obtain:

$$\frac{\partial \ln f}{\partial \operatorname{vech}[\mathbf{\Omega}]} = \mathbf{D}_N' \operatorname{vec}[\mathbf{G}_{\mathbf{\Omega}}'] \qquad (T7.127)$$
$$= \frac{1}{2}\mathbf{D}_N' \operatorname{vec}\left[(\nu_1 + K - N - 1)\,\mathbf{\Omega}^{-1} - \mathbf{A}\right].$$

Applying the first-order conditions to ($T7.126$) and ($T7.127$) and re-substituting ($T7.121$) we obtain the mode of the factor loadings:

$$\widetilde{\mathbf{B}} \equiv \mathbf{B}_1, \qquad (T7.128)$$

and the mode of the dispersion parameter:

$$\widetilde{\mathbf{\Omega}}^{-1} = \frac{\nu_1}{\nu_1 + K - N - 1}\mathbf{\Sigma}_1. \qquad (T7.129)$$

To compute the modal dispersion we differentiate ($T7.120$). Using ($A.126$) the second differential reads:

$$\begin{aligned} d\,(d \ln f) &= -\frac{1}{2}\operatorname{tr}\left\{\left[(\nu_1 + K - N - 1)\,\mathbf{\Omega}^{-1}(d\mathbf{\Omega})\,\mathbf{\Omega}^{-1} + 2T_1 d\mathbf{B}\mathbf{\Sigma}_{F,1}(\mathbf{B} - \mathbf{B}_1)'\right] d\mathbf{\Omega}\right\} \\ &\quad - \operatorname{tr}\{T_1\mathbf{\Sigma}_{F,1} d\mathbf{B}'\mathbf{\Omega} d\mathbf{B}\} - \operatorname{tr}\left\{T_1\mathbf{\Sigma}_{F,1}(\mathbf{B} - \mathbf{B}_1)' d\mathbf{\Omega} d\mathbf{B}\right\} \\ &= -\frac{\nu_1 + K - N - 1}{2}\operatorname{tr}\left\{\mathbf{\Omega}^{-1}(d\mathbf{\Omega})\,\mathbf{\Omega}^{-1} d\mathbf{\Omega}\right\} \qquad (T7.130) \\ &\quad - T_1 \operatorname{tr}\{\mathbf{\Sigma}_{F,1} d\mathbf{B}'\mathbf{\Omega} d\mathbf{B}\} - 2T_1 \operatorname{tr}\left\{\mathbf{\Sigma}_{F,1}(\mathbf{B} - \mathbf{B}_1)' d\mathbf{\Omega} d\mathbf{B}\right\}, \end{aligned}$$

The first term in ($T7.130$) can be expressed using ($A.107$), ($A.106$) the duplication matrix ($A.113$) to get rid of the redundancies of $d\mathbf{\Omega}$ as follows:

$$\begin{aligned} \operatorname{tr}\left\{\mathbf{\Omega}^{-1}(d\mathbf{\Omega})\,\mathbf{\Omega}^{-1} d\mathbf{\Omega}\right\} &= \operatorname{vec}[d\mathbf{\Omega}]' \operatorname{vec}\left[\mathbf{\Omega}^{-1}(d\mathbf{\Omega})\,\mathbf{\Omega}^{-1}\right] \\ &= \operatorname{vec}[d\mathbf{\Omega}]\left(\mathbf{\Omega}^{-1} \otimes \mathbf{\Omega}^{-1}\right)\operatorname{vec}[d\mathbf{\Omega}] \qquad (T7.131) \\ &= \operatorname{vec}[d\mathbf{\Omega}]\left(\mathbf{\Omega}^{-1} \otimes \mathbf{\Omega}^{-1}\right)\operatorname{vec}[d\mathbf{\Omega}] \\ &= \operatorname{vech}[d\mathbf{\Omega}]\,\mathbf{D}_N'\left(\mathbf{\Omega}^{-1} \otimes \mathbf{\Omega}^{-1}\right)\mathbf{D}_N \operatorname{vech}[d\mathbf{\Omega}]. \end{aligned}$$

Similarly, the second term in ($T7.130$) can be expressed using ($A.107$), ($A.106$), ($A.108$) and ($A.109$):

$$\begin{aligned} \operatorname{tr}\{\mathbf{\Sigma}_{F,1} d\mathbf{B}'\mathbf{\Omega} d\mathbf{B}\} &= \operatorname{tr}\{d\mathbf{B}\mathbf{\Sigma}_{F,1} d\mathbf{B}'\mathbf{\Omega}\} \qquad (T7.132) \\ &= \operatorname{vec}[d\mathbf{B}']' \operatorname{vec}[\mathbf{\Sigma}_{F,1} d\mathbf{B}'\mathbf{\Omega}] \\ &= (\mathbf{K}_{KN}\operatorname{vec}[d\mathbf{B}])'(\mathbf{\Omega} \otimes \mathbf{\Sigma}_{F,1})(\mathbf{K}_{KN}\operatorname{vec}[d\mathbf{B}]) \\ &= \operatorname{vec}[d\mathbf{B}]'\,\mathbf{K}_{NK}(\mathbf{\Omega} \otimes \mathbf{\Sigma}_{F,1})\,\mathbf{K}_{KN}\operatorname{vec}[d\mathbf{B}]. \end{aligned}$$

Since we are interested in the Hessian evaluated in the mode, where ($T7.128$) holds. Therefore the last term in ($T7.130$) cancels and we can express the second differential ($T7.130$) as follows:

$$\begin{aligned} d\,(d \ln f)|_{\widetilde{\mathbf{B}},\widetilde{\mathbf{\Omega}}} &= -\frac{\nu_1 + K - N - 1}{2}\operatorname{vech}[d\mathbf{\Omega}]\,\mathbf{D}_N'\left(\widetilde{\mathbf{\Omega}}^{-1} \otimes \widetilde{\mathbf{\Omega}}^{-1}\right)\mathbf{D}_N \operatorname{vech}[d\mathbf{\Omega}] \\ &\quad - T_1 \operatorname{vec}[d\mathbf{B}]'\,\mathbf{K}_{NK}\left(\widetilde{\mathbf{\Omega}} \otimes \mathbf{\Sigma}_{F,1}\right)\mathbf{K}_{KN}\operatorname{vec}[d\mathbf{B}]. \qquad (T7.133) \end{aligned}$$

Therefore from $(A.117)$ and $(A.121)$ and substituting back $(T7.129)$ we obtain:

$$\left.\frac{\partial^2 \ln f}{\partial \operatorname{vec}[\mathbf{B}]\, \partial \operatorname{vec}[\mathbf{B}]'}\right|_{\widetilde{\mathbf{B}},\widetilde{\mathbf{\Omega}}} = -T_1 \frac{\nu_1 + K - N - 1}{\nu_1} \mathbf{K}_{NK} \left(\mathbf{\Sigma}_1^{-1} \otimes \mathbf{\Sigma}_{F,1}\right) \mathbf{K}_{KN}$$

$$\left.\frac{\partial^2 \ln f}{\partial \operatorname{vech}[\mathbf{\Omega}]\, \partial \operatorname{vec}[\mathbf{B}]'}\right|_{\widetilde{\mathbf{B}},\widetilde{\mathbf{\Omega}}} = \mathbf{0}_{(N(N+1)/2)^2 \times (NK)^2} \qquad (T7.134)$$

$$\left.\frac{\partial^2 \ln f}{\partial \operatorname{vech}(\mathbf{\Omega})\, \partial \operatorname{vech}(\mathbf{\Omega})'}\right|_{\widetilde{\mathbf{B}},\widetilde{\mathbf{\Omega}}} = -\frac{1}{2} \frac{\nu_1^2}{\nu_1 + K - N - 1} \mathbf{D}_N' \left(\mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_1\right) \mathbf{D}_N$$

Finally the modal dispersion reads:

$$\begin{aligned} \operatorname{MDis}\{\boldsymbol{\theta}\} &\equiv \left(- \left.\frac{\partial^2 \ln f}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right|_{\widetilde{\boldsymbol{\theta}}} \right)^{-1} \qquad (T7.135) \\ &= \begin{pmatrix} \mathbf{S}_{\mathbf{B}} & \mathbf{0}_{(NK)^2 \times (N(N+1)/2)^2} \\ \mathbf{0}_{(N(N+1)/2)^2 \times (NK)^2} & \mathbf{S}_{\mathbf{\Sigma}} \end{pmatrix}, \end{aligned}$$

where using $(A.109)$ and $(A.101)$ we have:

$$\mathbf{S}_{\mathbf{B}} \equiv \frac{1}{T_1} \frac{\nu_1}{\nu_1 + K - N - 1} \mathbf{K}_{NK} \left(\mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_{F,1}^{-1}\right) \mathbf{K}_{KN} \qquad (T7.136)$$

$$\mathbf{S}_{\mathbf{\Sigma}} \equiv \frac{2}{\nu_1} \frac{\nu_1 + K - N - 1}{\nu_1} \left[\mathbf{D}_N' \left(\mathbf{\Sigma}_1 \otimes \mathbf{\Sigma}_1\right) \mathbf{D}_N\right]^{-1}. \qquad (T7.137)$$

7.8 NIW factor loadings-dispersion: marginal distribution of factor loadings

First of all, a comment on the notation to follow: we will denote here $\gamma_1, \gamma_2, \ldots$ simple normalization constants.

We consider the notation for the NIW (normal-inverse-Wishart) assumptions (4.129)-(7.59) on the prior, although of course the proof applies verbatim to the posterior, or any NIW distribution. Thus we assume:

$$\mathbf{\Omega} \equiv \mathbf{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_0, (\nu_0 \mathbf{\Sigma}_0)^{-1}\right) \qquad (T7.138)$$

and

$$\mathbf{B}|\mathbf{\Omega} \sim \mathrm{N}\left(\mathbf{B}_0, (T_0 \mathbf{\Omega})^{-1}, \mathbf{\Sigma}_{F,0}^{-1}\right). \qquad (T7.139)$$

From (2.182) and (2.224) the joint prior pdf of $\mathbf{B}$ and $\mathbf{\Omega}$ is:

$$\begin{aligned} f(\mathbf{B}, \mathbf{\Omega}) &= f(\mathbf{B}|\mathbf{\Omega})\, f(\mathbf{\Omega}) \\ &= \gamma_1 |\mathbf{\Omega}|^{\frac{K}{2}} |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} e^{-\frac{1}{2} \operatorname{tr}\left\{(T_0 \mathbf{\Omega})(\mathbf{B} - \mathbf{B}_0) \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)'\right\}} \qquad (T7.140) \\ &\quad |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} |\mathbf{\Omega}|^{\frac{\nu_0 - N - 1}{2}} e^{-\frac{1}{2} \operatorname{tr}(\nu_0 \mathbf{\Sigma}_0 \mathbf{\Omega})} \\ &= \gamma_1 |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} |\mathbf{\Omega}|^{\frac{\nu_0 + K - N - 1}{2}} \\ &\quad e^{-\frac{1}{2} \operatorname{tr}\left\{\mathbf{\Omega}\left((\mathbf{B} - \mathbf{B}_0) T_0 \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)' + \nu_0 \mathbf{\Sigma}_0\right)\right\}}. \end{aligned}$$

To determine the unconditional pdf of $\mathbf{B}$ we have to compute the marginal in ($T7.140$). Defining

$$\mathbf{\Sigma}_2 \equiv (\mathbf{B} - \mathbf{B}_0) T_0 \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)' + \nu_0 \mathbf{\Sigma}_0 \tag{$T7.141$}$$

we obtain

$$\begin{aligned} f(\mathbf{B}) &\equiv \int f(\mathbf{B}, \mathbf{\Omega})\, d\mathbf{\Omega} \\ &= \int \gamma_1 |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} |\mathbf{\Omega}|^{\frac{\nu_0+K-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}\{\mathbf{\Omega}\mathbf{\Sigma}_2\}} d\mathbf{\Omega} \\ &= \gamma_2 |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} |\mathbf{\Sigma}_2|^{-\frac{\nu_0+K}{2}} \\ &\quad \left\{ \int \gamma_3 |\mathbf{\Sigma}_2|^{\frac{\nu_0+K}{2}} |\mathbf{\Omega}|^{\frac{\nu_0+K-N-1}{2}} e^{-\frac{1}{2}\operatorname{tr}\{\mathbf{\Omega}\mathbf{\Sigma}_2\}} d\mathbf{\Omega} \right\} \\ &= \gamma_2 |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} |\mathbf{\Sigma}_2|^{-\frac{\nu_0+K}{2}} \end{aligned} \tag{$T7.142$}$$

where we have used the fact that the term in curly brackets is the integrals of the Wishart pdf (2.224) over the entire space and thus it sums to one.

Thus substituting again ($T7.141$) we obtain that the marginal pdf ($T7.142$) reads:

$$\begin{aligned} f(\mathbf{B}) = \gamma_2 |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} |\mathbf{\Sigma}_0|^{\frac{\nu_0}{2}} \\ \left| \nu_0 \mathbf{\Sigma}_0 + (\mathbf{B} - \mathbf{B}_0) T_0 \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)' \right|^{-\frac{\nu_0+K}{2}}, \end{aligned} \tag{$T7.143$}$$

From ($A.91$) we obtain the following general identity:

$$\begin{aligned} |\mathbf{\Sigma}|^{\frac{\nu_0}{2}} \left| \mathbf{\Sigma} + \mathbf{V}\mathbf{V}' \right|^{-\frac{\nu_0+K}{2}} &= |\mathbf{\Sigma}|^{\frac{\nu_0}{2}} \left[|\mathbf{\Sigma}| \left| \mathbf{I}_N + \mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{V}' \right| \right]^{-\frac{\nu_0+K}{2}} \\ &= |\mathbf{\Sigma}|^{-\frac{K}{2}} \left| \mathbf{I}_N + \mathbf{\Sigma}^{-1}\mathbf{V}\mathbf{V}' \right|^{-\frac{\nu_0+K}{2}} \\ &= |\mathbf{\Sigma}|^{-\frac{K}{2}} \left| \mathbf{I}_K + \mathbf{V}'\mathbf{\Sigma}^{-1}\mathbf{V} \right|^{-\frac{\nu_0+K}{2}} \end{aligned} \tag{$T7.144$}$$

Applying this result to

$$\begin{aligned} \mathbf{V} &\equiv (\mathbf{B} - \mathbf{B}_0) \mathbf{P}_{F,0} \\ \mathbf{\Sigma} &\equiv \nu_0 \mathbf{\Sigma}_0, \end{aligned} \tag{$T7.145$}$$

where $\mathbf{P}_{F,0}$ satisfies

$$T_0 \mathbf{\Sigma}_{F,0} \equiv \mathbf{P}_{F,0} \mathbf{P}_{F,0}', \tag{$T7.146$}$$

we reduce ($T7.143$) to the following expression:

$$\begin{aligned} f(\mathbf{B}) = \gamma_4 |\mathbf{\Sigma}_{F,0}|^{\frac{N}{2}} |\nu_0 \mathbf{\Sigma}_0|^{-\frac{K}{2}} \\ \left| \mathbf{I}_N + (\nu_0 \mathbf{\Sigma}_0)^{-1} (\mathbf{B} - \mathbf{B}_0) T_0 \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)' \right|^{-\frac{\nu_0+K}{2}} \end{aligned} \tag{$T7.147$}$$

Applying again $(A.91)$ we obtain:

$$
\begin{aligned}
f(\mathbf{B}) = \gamma_4 &\left|\mathbf{\Sigma}_{F,0}^{-1}\right|^{-\frac{N}{2}} \left|\nu_0 \mathbf{\Sigma}_0\right|^{-\frac{K}{2}} \\
&\left|\mathbf{I}_K + T_0 \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)' (\nu_0 \mathbf{\Sigma}_0)^{-1} (\mathbf{B} - \mathbf{B}_0)\right|^{-\frac{\nu_0 + K}{2}} \\
= \gamma_5 &\left|\mathbf{\Sigma}_{F,0}^{-1}\right|^{-\frac{N}{2}} \left|(\nu_0 + K - N)^{-1} \nu_0 \mathbf{\Sigma}_0\right|^{-\frac{K}{2}} \\
&\left|\mathbf{I}_K + T_0 \mathbf{\Sigma}_{F,0} (\mathbf{B} - \mathbf{B}_0)' \frac{\left[(\nu_0 + K - N)(\nu_0 \mathbf{\Sigma}_0)^{-1}\right]}{(\nu_0 + K - N)} (\mathbf{B} - \mathbf{B}_0)\right|^{-\frac{\nu_0 + K}{2}}
\end{aligned}
\tag{T7.148}
$$

Comparing with (2.199) we see that this is the pdf of a matrix-variate Student t distribution with the following parameters:

$$
\mathbf{B} \sim \mathrm{St}\left(\nu_0 + K - N, \mathbf{B}_0, \frac{\nu_0}{\nu_0 + K - N} \mathbf{\Sigma}_0, (T_0 \mathbf{\Sigma}_{F,0})^{-1}\right). \tag{T7.149}
$$

7.9 Results on the determination of the prior

Allocation-implied parameters

Consider the constraints

$$
\mathcal{C}_1 : \ \boldsymbol{\alpha}' \mathbf{p}_T = w_T \tag{T7.150}
$$

and

$$
\mathcal{C}_2 : \ \underline{\mathbf{b}} \leq \mathbf{B} \boldsymbol{\alpha} \leq \overline{\mathbf{b}}, \tag{T7.151}
$$

where B is a $K \times N$ matrix and $\left(\underline{\mathbf{b}}, \overline{\mathbf{b}}\right)$ are K-dimensional vectors. From (7.91) we obtain the allocation function:

$$
\boldsymbol{\alpha}(\boldsymbol{\mu}, \mathbf{\Sigma}) \equiv \underset{\substack{\boldsymbol{\alpha}' \mathbf{p}_T = w_T \\ \underline{\mathbf{b}} \leq \mathbf{B} \boldsymbol{\alpha} \leq \overline{\mathbf{b}}}}{\operatorname{argmax}} \left\{ \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T)(\mathbf{1} + \boldsymbol{\mu}) - \frac{1}{2\zeta} \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \mathbf{\Sigma} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha} \right\}. \tag{T7.152}
$$

We can solve this problem by means of Lagrange multipliers. We defining the Lagrangian:

$$
\begin{aligned}
\mathcal{L} \equiv\ & \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T)(\mathbf{1} + \boldsymbol{\mu}) - \frac{1}{2\zeta} \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \mathbf{\Sigma} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha} \\
& - \lambda \boldsymbol{\alpha}' \mathbf{p}_T - \left(\overline{\boldsymbol{\lambda}} - \underline{\boldsymbol{\lambda}}\right)' \mathbf{B} \boldsymbol{\alpha},
\end{aligned}
\tag{T7.153}
$$

where λ is the multiplier relative to the equality constraint $\boldsymbol{\alpha}' \mathbf{p}_T = w_T$ and $\left(\overline{\boldsymbol{\lambda}}, \underline{\boldsymbol{\lambda}}\right)$ are the multipliers relative to the additional inequality constraints $(T7.151)$ and satisfy the Kuhn-Tucker conditions:

$$\overline{\boldsymbol{\lambda}}, \underline{\boldsymbol{\lambda}} \geq \mathbf{0} \tag{T7.154}$$

$$\sum_{n=1}^{N} \underline{\lambda}_k B_{kn} \underline{b}_n = \sum_{n=1}^{N} \overline{\lambda}_k B_{kn} \overline{b}_n = 0, \quad k = 1, \ldots, K. \tag{T7.155}$$

Therefore, defining:

$$\widetilde{\boldsymbol{\mu}} \equiv \boldsymbol{\mu} - [\operatorname{diag}(\mathbf{p}_T)]^{-1} \mathbf{B}' \left(\overline{\boldsymbol{\lambda}} - \underline{\boldsymbol{\lambda}}\right), \tag{T7.156}$$

we can write the Lagrangian as follows:

$$\mathcal{L} = \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T)(\mathbf{1} + \widetilde{\boldsymbol{\mu}}) - \lambda \boldsymbol{\alpha}' - \frac{1}{2\zeta} \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\Sigma} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha}. \tag{T7.157}$$

This is the Lagrangian of the optimization ($T7.152$) with the constraints ($T7.150$) but without the constraints ($T7.151$). Its solution is (6.39). After substituting (7.90) in that expression we obtain the respective allocation function:

$$\begin{aligned} (\widetilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}) &\mapsto \boldsymbol{\alpha}(\widetilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}) \\ &\equiv [\operatorname{diag}(\mathbf{p}_T)]^{-1} \boldsymbol{\Sigma}^{-1} \left(\zeta \widetilde{\boldsymbol{\mu}} + \frac{w_T - \zeta \mathbf{1}' \boldsymbol{\Sigma}^{-1} \widetilde{\boldsymbol{\mu}}}{\mathbf{1}' \boldsymbol{\Sigma}^{-1} \mathbf{1}} \mathbf{1} \right). \end{aligned} \tag{T7.158}$$

This can be inverted, by pinning down specific values $\overline{\boldsymbol{\Sigma}}$ for the covariance matrix and solving the ensuing implicit equation:

$$\widetilde{\boldsymbol{\mu}} - \frac{\mathbf{1}' \overline{\boldsymbol{\Sigma}}^{-1} \widetilde{\boldsymbol{\mu}}}{\mathbf{1}' \overline{\boldsymbol{\Sigma}}^{-1} \mathbf{1}} \mathbf{1} = \frac{1}{\zeta} \left(\overline{\boldsymbol{\Sigma}} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha} - \frac{w_T}{\mathbf{1}' \overline{\boldsymbol{\Sigma}}^{-1} \mathbf{1}} \mathbf{1} \right). \tag{T7.159}$$

From the inverse function

$$\boldsymbol{\alpha} \mapsto \widetilde{\boldsymbol{\mu}}(\boldsymbol{\alpha}) \tag{T7.160}$$

and from ($T7.156$) the implied returns that include the constraints ($T7.151$) read:

$$\boldsymbol{\mu}_c = \widetilde{\boldsymbol{\mu}}(\boldsymbol{\alpha}) + [\operatorname{diag}(\mathbf{p}_T)]^{-1} \mathbf{B}' \left(\overline{\boldsymbol{\lambda}} - \underline{\boldsymbol{\lambda}}\right) \tag{T7.161}$$

Likelihood maximization

In our example (7.91), consider an investor who has no risk propensity, i.e. such that $\zeta \to 0$ in his exponential utility function. Then the quadratic term becomes overwhelming in the index of satisfaction, which become independent of the expected returns:

$$\mathrm{CE}_{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}) \approx -\frac{1}{2\zeta} \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\Sigma} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha}. \tag{T7.162}$$

Assume there exists a budget constraint

$$\mathcal{C}_1 : \ \boldsymbol{\alpha}' \mathbf{p}_T = w_T, \tag{T7.163}$$

And consider the no-short-sale constraint:

$$\mathcal{C}_2 : \ \boldsymbol{\alpha} \geq \mathbf{0}. \tag{T7.164}$$

The allocation function $\boldsymbol{\alpha}(\boldsymbol{\Sigma})$ follows in terms of the Lagrangian:

$$\boldsymbol{\alpha}(\boldsymbol{\Sigma}), \lambda^*, \underline{\boldsymbol{\lambda}}^* = \underset{\boldsymbol{\alpha}, \lambda, \underline{\boldsymbol{\lambda}}}{\operatorname{argmin}} \mathcal{L}(\boldsymbol{\alpha}, \lambda, \underline{\boldsymbol{\lambda}}), \tag{T7.165}$$

where

$$\mathcal{L}(\boldsymbol{\alpha}, \lambda, \underline{\boldsymbol{\lambda}}) \equiv \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\Sigma} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha} - \lambda \boldsymbol{\alpha}' \mathbf{p}_T - \boldsymbol{\alpha}' \underline{\boldsymbol{\lambda}}. \tag{T7.166}$$

From the first-order conditions on the Lagrangian we obtain:

$$\boldsymbol{\alpha} = \lambda \operatorname{diag}(\mathbf{p}_T)^{-1} \boldsymbol{\Sigma}^{-1} \mathbf{1} + \operatorname{diag}(\mathbf{p}_T)^{-1} \boldsymbol{\Sigma}^{-1} \operatorname{diag}(\mathbf{p}_T)^{-1} \underline{\boldsymbol{\lambda}}, \tag{T7.167}$$

together with the Kuhn-Tucker conditions

$$\underline{\boldsymbol{\lambda}} \geq \mathbf{0}, \quad \underline{\lambda}_n = 0 \iff \alpha_n > 0. \tag{T7.168}$$

Since prices are positive, the allocation is positive if

$$\boldsymbol{\Sigma}^{-1} \mathbf{1} \geq \mathbf{0}, \tag{T7.169}$$

where the inequality is meant entry by entry.

7.10 Monte Carlo Markov Chain (MCMC) generation of posterior distribution

References for the discussion below can be found e.g. in Chib and Greenberg (1995). We recall that according to Bayesian theory, first we must specify a flexible parametric family for the market, which might include fat tails and skewness, as represented by its pdf:

$$\mathbf{X} | \boldsymbol{\theta} \Leftrightarrow f_{\mathbf{X}}(\mathbf{x}_t | \boldsymbol{\theta}). \tag{T7.170}$$

Then as in (7.13) the likelihood of the data

$$i_T \equiv \{\mathbf{x}_1, \ldots, \mathbf{x}_T\} \tag{T7.171}$$

reads:

$$l(i_T | \boldsymbol{\theta}) \equiv \prod_{t=1}^{T} f_{\mathbf{X}}(\mathbf{x}_t | \boldsymbol{\theta}). \tag{T7.172}$$

Assume a prior for the parameters $f_0(\boldsymbol{\theta})$. Then posterior distribution of the parameters reads:

$$\pi\left(\boldsymbol{\theta}|i_T\right) \equiv \frac{l\left(i_T|\boldsymbol{\theta}\right) f_0\left(\boldsymbol{\theta}\right)}{\int l\left(i_T|\boldsymbol{\theta}\right) f_0\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta}}, \tag{T7.173}$$

see (7.15).

To generate samples from the posterior distribution we use a Metropolis-Hastings Markov Chain Monte Carlo (MCMC) algorithm. First we select a one-parameter family of candidate-generating densities $q\left(\boldsymbol{\xi}, \boldsymbol{\theta}\right)$, which satisfies $\int q\left(\boldsymbol{\xi}, \boldsymbol{\theta}\right) d\boldsymbol{\theta} = 1$. Then we define the function

$$\alpha\left(\boldsymbol{\theta}, \boldsymbol{\xi}\right) \equiv \min\left\{\frac{l\left(i_T|\boldsymbol{\xi}\right) f_0\left(\boldsymbol{\xi}\right) q\left(\boldsymbol{\xi}, \boldsymbol{\theta}\right)}{l\left(i_T|\boldsymbol{\theta}\right) f_0\left(\boldsymbol{\theta}\right) q\left(\boldsymbol{\theta}, \boldsymbol{\xi}\right)}, 1\right\}. \tag{T7.174}$$

In particular, if we choose a symmetric function $q\left(\boldsymbol{\xi}, \boldsymbol{\theta}\right) = q\left(\boldsymbol{\theta}, \boldsymbol{\xi}\right)$, this function simplifies to

$$\alpha\left(\boldsymbol{\theta}, \boldsymbol{\xi}\right) \equiv \min\left\{\frac{l\left(i_T|\boldsymbol{\xi}\right) f_0\left(\boldsymbol{\xi}\right)}{l\left(i_T|\boldsymbol{\theta}\right) f_0\left(\boldsymbol{\theta}\right)}, 1\right\}. \tag{T7.175}$$

A convenient rule of thumb is to pick q to be the multivariate normal density

$$q\left(\mathbf{z}, \mathbf{b}\right) \Leftrightarrow \mathrm{N}\left(\mathbf{b}, \kappa \operatorname{diag}\left(\mathbf{b}\right)\right), \tag{T7.176}$$

where $\kappa \approx 10^{-2}$. Then the algorithm proceeds as follows:

- Step 0. Set $j \equiv 0$ and generate a starting point for the parameters $\boldsymbol{\theta}^{(j)}$.
- Step 1. Generate $\boldsymbol{\xi}$ from $q\left(\boldsymbol{\theta}^{(j)}, \cdot\right)$ and u from $\mathrm{U}\left([0,1]\right)$.
- Step 2. If $u \leq \alpha\left(\boldsymbol{\theta}^{(j)}, \boldsymbol{\xi}\right)$ set $\boldsymbol{\theta}^{(j+1)} \equiv \boldsymbol{\xi}$, else set $\boldsymbol{\theta}^{(j+1)} \equiv \boldsymbol{\theta}^{(j)}$.
- Step 3. If convergence is achieved go to Step 4, else go to Step 1.
- Step 4. Disregard the first samples and return the remaining $\boldsymbol{\theta}^{(j)}$.

8

Technical appendix to Chapter 8

8.1 Optimal allocation as function of invariant parameters

Replacing the market parameters (8.21) in the certainty-equivalent (8.25) we obtain:

$$\mathcal{S} = \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{P}_T)(\mathbf{1}+\boldsymbol{\mu}) - \frac{1}{2\zeta}\boldsymbol{\alpha}' \operatorname{diag}(\mathbf{P}_T)\,\boldsymbol{\Sigma} \operatorname{diag}(\mathbf{P}_T)\,\boldsymbol{\alpha}. \qquad (T8.1)$$

Substituting in this expression the optimal allocation (8.32), which we report here:

$$\boldsymbol{\alpha} = \zeta\,[\operatorname{diag}(\mathbf{P}_T)]^{-1}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\,[\operatorname{diag}(\mathbf{P}_T)]^{-1}\,\boldsymbol{\Sigma}^{-1}\mathbf{1}, \qquad (T8.2)$$

we obtain:

$$\begin{aligned}
\mathcal{S} &= (\mathbf{1}+\boldsymbol{\mu})' \left[\zeta\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\boldsymbol{\Sigma}^{-1}\mathbf{1}\right] \qquad (T8.3)\\
&\quad -\frac{1}{2\zeta}\left[\zeta\boldsymbol{\mu}' + \frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}'\right]\left[\zeta\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\boldsymbol{\Sigma}^{-1}\mathbf{1}\right]\\
&= \zeta\,(\mathbf{1}+\boldsymbol{\mu})'\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\,(\mathbf{1}+\boldsymbol{\mu})'\,\boldsymbol{\Sigma}^{-1}\mathbf{1}\\
&\quad -\frac{1}{2\zeta}\left(\zeta^2\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \zeta\frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right.\\
&\quad \left. +\zeta\frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \left(\frac{w_T - \zeta \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)^2 \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}\right).
\end{aligned}$$

Defining

$$A \equiv \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}, \quad B \equiv \mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}, \quad C \equiv \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \qquad (T8.4)$$

the above expression simplifies as follows:

$$\begin{aligned}\mathcal{S} &= \zeta B + \zeta C + \frac{w_T - \zeta B}{A} A + \frac{w_T - \zeta B}{A} B \\ &\quad - \frac{1}{2\zeta}\left[\zeta^2 C + \zeta \frac{w_T - \zeta B}{A} B + \zeta \frac{w_T - \zeta B}{A} B + \left(\frac{w_T - \zeta B}{A}\right)^2 A\right] \\ &= \zeta B + \zeta C + w_T - \zeta B + \frac{w_T - \zeta B}{A} B \\ &\quad - \frac{1}{2}\left[\zeta C + 2\frac{w_T - \zeta B}{A} B + \frac{-2Bw_T + \zeta B^2}{A}\right] - \frac{1}{2\zeta}\frac{w_T^2}{A} \\ &= \frac{1}{2}\zeta\left(C - \frac{B^2}{A}\right) + w_T\left(1 + \frac{B}{A} - \frac{1}{2\zeta}\frac{w_T}{A}\right).\end{aligned} \tag{T8.5}$$

8.2 Statistical significance of sample allocation

Consider the independent sample estimators (8.85) and (8.86):

$$\widehat{\boldsymbol{\mu}} \sim \mathrm{N}\left(\boldsymbol{\mu}, \frac{1}{T}\boldsymbol{\Sigma}\right), \quad T\widehat{\boldsymbol{\Sigma}} \sim \mathrm{W}\left(T-1, \boldsymbol{\Sigma}\right). \tag{T8.6}$$

where W denotes the Wishart distribution.

Define:

$$\widehat{v} \equiv \boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right) \widehat{\boldsymbol{\Sigma}} \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\alpha}. \tag{T8.7}$$

From (2.230) the distribution of this random variable satisfies

$$T\widehat{v} \sim \mathrm{Ga}\left(T-1, \boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\Sigma} \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\alpha}\right), \tag{T8.8}$$

where Ga denotes the gamma distribution. Thus from (1.113) the expected value of $\widehat{v}$ reads:

$$\mathrm{E}\left\{\widehat{v}\right\} = \frac{T-1}{T}\boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\Sigma} \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\alpha}, \tag{T8.9}$$

and from (1.114) the inefficiency of $\widehat{v}$ reads:

$$\mathrm{Sd}\left\{\widehat{v}\right\} = \sqrt{2\frac{T-1}{T^2}}\boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\Sigma} \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\alpha}. \tag{T8.10}$$

Similarly, define

$$\widehat{e} \equiv \boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right)\left(\mathbf{1} + \widehat{\boldsymbol{\mu}}\right) \tag{T8.11}$$

From (2.163) we obtain:

$$\widehat{e} \sim \mathrm{N}\left(\boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right)\left(\mathbf{1} + \boldsymbol{\mu}\right), \frac{\boldsymbol{\alpha}' \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\Sigma} \operatorname{diag}\left(\mathbf{P}_T\right) \boldsymbol{\alpha}}{T}\right). \tag{T8.12}$$

Therefore, from (2.158) the expected value of $\widehat{e}$ reads:

$$\mathrm{E}\{\widehat{e}\} = \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{P}_T)(\mathbf{1} + \boldsymbol{\mu}), \tag{T8.13}$$

and from (2.159) the inefficiency of $\widehat{e}$ reads:

$$\mathrm{Sd}\{\widehat{e}\} = \sqrt{\frac{\boldsymbol{\alpha}' \operatorname{diag}(\mathbf{P}_T)\, \boldsymbol{\Sigma} \operatorname{diag}(\mathbf{P}_T)\, \boldsymbol{\alpha}}{T}}. \tag{T8.14}$$

Furthermore, since $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ are independent, so are $\widehat{v}$ and $\widehat{e}$.

9

Technical appendix to Chapter 9

9.1 Results on the resampled allocation

$$\begin{aligned}\boldsymbol{\alpha}_{\mathrm{rs}}\left[i_T\right] &\equiv \mathrm{E}\left\{\boldsymbol{\alpha}_{\mathrm{s}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right]\right\} &(T9.1)\\ &= \mathrm{E}\left\{\zeta\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathbf{V}\mathbf{w}\right\} + \mathrm{E}\left\{\frac{w_T}{\mathbf{1}'\mathbf{V}\mathbf{1}}\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathbf{V}\mathbf{1}\right\}\\ &\quad - \mathrm{E}\left\{\frac{\zeta\mathbf{1}'\mathbf{V}\mathbf{w}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathbf{V}\mathbf{1}\right\}\\ &= \zeta\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathrm{E}\left\{\mathbf{V}\right\}\mathrm{E}\left\{\mathbf{w}\right\} + w_T\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathrm{E}\left\{\frac{\mathbf{V}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\right\}\\ &\quad -\zeta\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathrm{E}\left\{\mathbf{w}\right\}'\mathrm{E}\left\{\frac{\mathbf{V}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\mathbf{V}\mathbf{1}\right\}\\ &= \zeta\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathrm{E}\left\{\mathbf{V}\right\}\widehat{\boldsymbol{\mu}} + w_T\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathrm{E}\left\{\frac{\mathbf{V}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\right\}\\ &\quad -\zeta\left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\widehat{\boldsymbol{\mu}}'\,\mathrm{E}\left\{\frac{\mathbf{V}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\mathbf{V}\mathbf{1}\right\}.\end{aligned}$$

Therefore

$$\begin{aligned}\boldsymbol{\alpha}_{\mathrm{rs}}\left[i_T\right] &= \left[\mathrm{diag}\left(\mathbf{p}_T\right)\right]^{-1}\left(\zeta\left(\mathrm{E}\left\{\mathbf{V}\widehat{\boldsymbol{\mu}}\right\} - \mathrm{E}\left\{\frac{\mathbf{1}'\mathbf{V}\widehat{\boldsymbol{\mu}}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\mathbf{V}\mathbf{1}\right\}\right)\right. &(T9.2)\\ &\quad \left. + \, w_T\,\mathrm{E}\left\{\frac{\mathbf{V}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\right\}\right).\end{aligned}$$

9.2 Probability bounds for the sample mean

From (8.85) the sample estimator is distributed as follows:

$$\widehat{\boldsymbol{\mu}} \sim \mathrm{N}\left(\boldsymbol{\mu}^{\mathrm{t}}, \frac{\boldsymbol{\Sigma}^{\mathrm{t}}}{T}\right), \tag{T9.3}$$

where $\boldsymbol{\mu}^{\mathrm{t}}$ and $\boldsymbol{\Sigma}^{\mathrm{t}}$ are the true underlying parameters. Therefore from Appendix www.7.1 we have:

$$\left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}^{\mathrm{t}}\right)' \left(\frac{\boldsymbol{\Sigma}^{\mathrm{t}}}{T}\right)^{-1} \left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}^{\mathrm{t}}\right) \sim \chi^2_N. \tag{T9.4}$$

From the definition (1.7) of cumulative distribution function:

$$\begin{aligned} F_{\chi^2_N}(T\gamma) &\equiv \mathbb{P}\left\{\left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}^{\mathrm{t}}\right)' \left(\frac{\boldsymbol{\Sigma}^{\mathrm{t}}}{T}\right)^{-1} \left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}^{\mathrm{t}}\right) \leq T\gamma\right\} \\ &= \mathbb{P}\left\{\left(\boldsymbol{\mu}^{\mathrm{t}} - \widehat{\boldsymbol{\mu}}\right)' \left(\boldsymbol{\Sigma}^{\mathrm{t}}\right)^{-1} \left(\boldsymbol{\mu}^{\mathrm{t}} - \widehat{\boldsymbol{\mu}}\right) \leq \gamma\right\}. \end{aligned} \tag{T9.5}$$

By applying the quantile function (1.17) to both sides of the above equality we obtain:

$$p = \mathbb{P}\left\{\left(\boldsymbol{\mu}^{\mathrm{t}} - \widehat{\boldsymbol{\mu}}\right)' \left(\boldsymbol{\Sigma}^{\mathrm{t}}\right)^{-1} \left(\boldsymbol{\mu}^{\mathrm{t}} - \widehat{\boldsymbol{\mu}}\right) \leq \frac{Q_{\chi^2_N}(p)}{T}\right\}. \tag{T9.6}$$

Therefore, considering the set

$$\boldsymbol{\Theta}[i_T] \equiv \left\{\boldsymbol{\mu} \in \mathbb{R}^N \text{ such that } \operatorname{Ma}^2\left(\boldsymbol{\mu}, \widehat{\boldsymbol{\mu}}[i_T], \boldsymbol{\Sigma}^{\mathrm{t}}\right) \leq \frac{Q_{\chi^2_N}(p)}{T}\right\} \tag{T9.7}$$

The following result holds:

$$\mathbb{P}\left\{\boldsymbol{\mu}^{\mathrm{t}} \in \boldsymbol{\Theta}[i_T]\right\} = p. \tag{T9.8}$$

9.3 The Black-Litterman approach

First of all we prove the general Bayes' rule (9.30), which we report here:

$$f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}|\mathbf{v}) = \frac{f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})}{\int f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}}. \tag{T9.9}$$

Indeed, by the definition of the conditional density we have

$$f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}|\mathbf{v}) \equiv \frac{f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v})}{f_{\mathbf{V}}(\mathbf{v})}, \tag{T9.10}$$

where $f_{\mathbf{X},\mathbf{V}}$ is the joint distribution of $\mathbf{X}$ and $\mathbf{V}$ and

$$f_{\mathbf{V}}(\mathbf{v}) \equiv \int f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v})\, d\mathbf{x} \tag{T9.11}$$

is the marginal pdf of $\mathbf{V}$. On the other hand, by the definition of the conditional density we also have:

$$f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v}) = f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x}). \tag{T9.12}$$

Thus

$$\begin{aligned} f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}|\mathbf{v}) &\equiv \frac{f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v})}{\int f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v})\, d\mathbf{x}} \\ &= \frac{f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})}{\int f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}}, \end{aligned} \tag{T9.13}$$

as claimed.

In the Black-Litterman setting, the marginal pdf of $\mathbf{X}$ is assumed normal:

$$f_{\mathbf{X}}(\mathbf{x}) \equiv \frac{|\boldsymbol{\Sigma}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{T9.14}$$

and the conditional pdf of $\mathbf{V}$ given $\mathbf{PX} = \mathbf{Px}$ is:

$$f_{\mathbf{V}|\mathbf{Px}}(\mathbf{v}|\mathbf{x}) \equiv \frac{|\boldsymbol{\Omega}|^{-\frac{1}{2}}}{(2\pi)^{\frac{K}{2}}} e^{-\frac{1}{2}(\mathbf{v}-\mathbf{Px})'\boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{Px})}. \tag{T9.15}$$

Thus the joint pdf of $\mathbf{V}$ and $\mathbf{X}$ reads:

$$\begin{aligned} f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v}) &= f_{\mathbf{V}|\mathbf{Px}}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x}) \\ &\propto |\boldsymbol{\Sigma}|^{-\frac{1}{2}} |\boldsymbol{\Omega}|^{-\frac{1}{2}} e^{-\frac{1}{2}\left[(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})+(\mathbf{v}-\mathbf{Px})'\boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{Px})\right]}. \end{aligned} \tag{T9.16}$$

Expanding the expression in square brackets in $(T9.16)$ we obtain:

$$\begin{aligned} [\cdots] &= (\mathbf{x}-\boldsymbol{\mu})'\, \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) + (\mathbf{v}-\mathbf{Px})'\, \boldsymbol{\Omega}^{-1}(\mathbf{v}-\mathbf{Px}) \\ &= \mathbf{x}'\boldsymbol{\Sigma}^{-1}\mathbf{x} - 2\mathbf{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{v} - 2\mathbf{x}'\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v} + \mathbf{x}'\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{Px} \\ &= \mathbf{x}'\left(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\mathbf{x} - 2\mathbf{x}'\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}\right] + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{v} \end{aligned} \tag{T9.17}$$

We define $\widetilde{\boldsymbol{\mu}}$ in such a way that the following holds:

$$\left[\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}\right] \equiv \left(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\widetilde{\boldsymbol{\mu}}. \tag{T9.18}$$

This implies

$$\widetilde{\boldsymbol{\mu}}(\mathbf{v}) \equiv \left(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}\right). \tag{T9.19}$$

Using $(T9.18)$ we easily re-write $(T9.17)$ as follows:

$$\begin{aligned}[\cdots] &= \mathbf{x}'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\mathbf{x}-2\mathbf{x}'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\widetilde{\boldsymbol{\mu}} \\ &\quad +\widetilde{\boldsymbol{\mu}}'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\widetilde{\boldsymbol{\mu}}+\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{v} \\ &\quad -\widetilde{\boldsymbol{\mu}}'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\widetilde{\boldsymbol{\mu}} \\ &= (\mathbf{x}-\widetilde{\boldsymbol{\mu}})'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)(\mathbf{x}-\widetilde{\boldsymbol{\mu}})+\alpha. \end{aligned} \tag{T9.20}$$

where

$$\alpha \equiv \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{v}-\widetilde{\boldsymbol{\mu}}'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\widetilde{\boldsymbol{\mu}} \tag{T9.21}$$

Substituting the definition ($T9.19$) in this expression we obtain:

$$\begin{aligned}\alpha &= \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{v} \\ &\quad -\left(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}+\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}\right) \\ &= \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{v}-\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ &\quad -\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{P}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v} \\ &\quad +2\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v} \\ &= \mathbf{v}'\left\{\boldsymbol{\Omega}^{-1}-\boldsymbol{\Omega}^{-1}\mathbf{P}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}'\boldsymbol{\Omega}^{-1}\right\}\mathbf{v} \\ &\quad +2\mathbf{v}'\boldsymbol{\Omega}^{-1}\mathbf{P}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \\ &\quad +\boldsymbol{\mu}'\left(\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\mu} \end{aligned} \tag{T9.22}$$

Using the identity ($A.90$) we write the expression in curly brackets as follows:

$$\boldsymbol{\Omega}^{-1}-\boldsymbol{\Omega}^{-1}\mathbf{P}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\mathbf{P}'\boldsymbol{\Omega}^{-1}=\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1} \tag{T9.23}$$

Also, we define $\widetilde{\mathbf{v}}$ in such a way that

$$\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\widetilde{\mathbf{v}}=-\boldsymbol{\Omega}^{-1}\mathbf{P}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \tag{T9.24}$$

Therefore

$$\begin{aligned}\alpha &= \mathbf{v}'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\mathbf{v}-2\mathbf{v}'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\widetilde{\mathbf{v}} \qquad (T9.25) \\ &\quad +\widetilde{\mathbf{v}}'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\widetilde{\mathbf{v}}+\boldsymbol{\mu}'\left(\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\mu} \\ &\quad -\widetilde{\mathbf{v}}'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\widetilde{\mathbf{v}} \qquad (T9.26) \\ &= (\mathbf{v}-\widetilde{\mathbf{v}})'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}(\mathbf{v}-\widetilde{\mathbf{v}})+\phi, \end{aligned}$$

where

$$\begin{aligned}\phi &\equiv \boldsymbol{\mu}'\left(\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\mu} \\ &\quad -\widetilde{\mathbf{v}}'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\widetilde{\mathbf{v}} \end{aligned} \tag{T9.27}$$

From $(T9.24)$ we see that

$$\widetilde{\mathbf{v}} = -\left(\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)\boldsymbol{\Omega}^{-1}\mathbf{P}\left(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \qquad (T9.28)$$

does not depend on either $\mathbf{v}$ or $\mathbf{x}$. Therefore neither does ϕ in $(T9.27)$. Substituting $(T9.25)$ back in $(T9.20)$ the expression in square brackets in $(T9.16)$ reads

$$\begin{aligned}[\cdots] &= \left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}(\mathbf{v})\right)'\left(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\left(\mathbf{x} - \widetilde{\boldsymbol{\mu}}(\mathbf{v})\right) \\ &\quad + \left(\mathbf{v} - \widetilde{\mathbf{v}}\right)'\left(\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}\left(\mathbf{v} - \widetilde{\mathbf{v}}\right) + \phi. \end{aligned} \qquad (T9.29)$$

Therefore $(T9.16)$ becomes:

$$\begin{aligned} f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v}) &\propto |\boldsymbol{\Sigma}|^{-\frac{1}{2}}|\boldsymbol{\Omega}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\widetilde{\boldsymbol{\mu}}(\mathbf{v}))'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)(\mathbf{x}-\widetilde{\boldsymbol{\mu}}(\mathbf{v}))} \\ &\qquad e^{-\frac{1}{2}(\mathbf{v}-\widetilde{\mathbf{v}})'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}(\mathbf{v}-\widetilde{\mathbf{v}})} \\ &= \left|\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right|^{\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\widetilde{\boldsymbol{\mu}}(\mathbf{v}))'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)(\mathbf{x}-\widetilde{\boldsymbol{\mu}}(\mathbf{v}))} \\ &\qquad \left|\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{v}-\boldsymbol{\gamma})'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}(\mathbf{v}-\boldsymbol{\gamma})}, \end{aligned} \qquad (T9.30)$$

where the last equality follows from

$$\frac{\left|\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right|}{|\boldsymbol{\Sigma}|\,|\boldsymbol{\Omega}|\,\left|\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right|} = 1, \qquad (T9.31)$$

which in turn follows from an application of $(A.91)$:

$$\begin{aligned} |\boldsymbol{\Sigma}|\,|\boldsymbol{\Omega}|\left|\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right| &= \left|\boldsymbol{\Sigma}\left(\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)\right|\,|\boldsymbol{\Omega}| \\ &= \left|\mathbf{I} + \boldsymbol{\Sigma}\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right|\,|\boldsymbol{\Omega}| \\ &= \left|\mathbf{I} + \boldsymbol{\Omega}^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right|\,|\boldsymbol{\Omega}| = \left|\boldsymbol{\Omega}\left(\mathbf{I} + \boldsymbol{\Omega}^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)\right| \\ &= \left|\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right| \end{aligned} \qquad (T9.32)$$

To summarize, from $(T9.30)$ we see that

$$f_{\mathbf{X},\mathbf{V}}(\mathbf{x},\mathbf{v}) \propto f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}|\mathbf{v})\, f_{\mathbf{V}}(\mathbf{v}), \qquad (T9.33)$$

where

$$\begin{aligned} f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}|\mathbf{v}) &\propto \left|\boldsymbol{\Sigma}^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right|^{\frac{1}{2}} \\ &\qquad e^{-\frac{1}{2}(\mathbf{x}-\widetilde{\boldsymbol{\mu}}(\mathbf{v}))'\left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)(\mathbf{x}-\widetilde{\boldsymbol{\mu}}(\mathbf{v}))} \end{aligned} \qquad (T9.34)$$

and

$$f_{\mathbf{V}}(\mathbf{v}) \propto \left|\boldsymbol{\Omega} + \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{v}-\boldsymbol{\gamma})'\left(\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1}(\mathbf{v}-\boldsymbol{\gamma})}. \qquad (T9.35)$$

Since $(T9.34)$ and $(T9.35)$ are normal probability density functions, it follows that the random variable $\mathbf{X}$ conditioned on $\mathbf{V} = \mathbf{v}$ is normally distributed:

$$\mathbf{X}|\mathbf{V}=\mathbf{v}\sim \mathrm{N}\left(\widetilde{\boldsymbol{\mu}},\widetilde{\boldsymbol{\Sigma}}\right), \tag{T9.36}$$

and so is the marginal distribution of $\mathbf{V}$:

$$\mathbf{V}\sim \mathrm{N}\left(\widetilde{\mathbf{v}},\widetilde{\boldsymbol{\Omega}}\right), \tag{T9.37}$$

The expected value $\widetilde{\boldsymbol{\mu}}(\mathbf{v})$ in ($T9.36$) is defined in ($T9.19$) and, using ($A.90$), it reads:

$$\begin{aligned}\widetilde{\boldsymbol{\mu}}(\mathbf{v}) &\equiv \left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}\right). \\ &= \left(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\mathbf{P}'\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\right)\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{v}\right) \\ &= \boldsymbol{\mu}+\boldsymbol{\Sigma}\mathbf{P}'\left(\boldsymbol{\Omega}^{-1}-\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\boldsymbol{\Omega}^{-1}\right)\mathbf{v} \\ &\quad -\boldsymbol{\Sigma}\mathbf{P}'\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\mu}\end{aligned} \tag{T9.38}$$

Noticing that

$$\boldsymbol{\Omega}^{-1}-\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\boldsymbol{\Omega}^{-1}=\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}, \tag{T9.39}$$

which can be easily checked left-multiplying both sides by $\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)$, the expression for the expected value $\widetilde{\boldsymbol{\mu}}(\mathbf{v})$ in ($T9.36$) can be further simplified as follows:

$$\widetilde{\boldsymbol{\mu}}(\mathbf{v})=\boldsymbol{\mu}+\boldsymbol{\Sigma}\mathbf{P}'\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}(\mathbf{v}-\mathbf{P}\boldsymbol{\mu}) \tag{T9.40}$$

Similarly, from ($T9.34$) and using ($A.90$) the covariance matrix in ($T9.36$) reads:

$$\begin{aligned}\widetilde{\boldsymbol{\Sigma}} &\equiv \left(\boldsymbol{\Sigma}^{-1}+\mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}\right)^{-1} \\ &= \boldsymbol{\Sigma}-\boldsymbol{\Sigma}\mathbf{P}'\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\end{aligned} \tag{T9.41}$$

On the other hand, the expected value $\widetilde{\mathbf{v}}$ in ($T9.37$) is defined in ($T9.28$) and from ($T9.35$) the covariance matrix in ($T9.37$) reads:

$$\widetilde{\boldsymbol{\Omega}}\equiv\boldsymbol{\Omega}+\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' \tag{T9.42}$$

9.4 Investor's certain views in the Black-Litterman approach

First of all we complete the $K\times N$ matrix $\mathbf{P}$ to a non-singular $N\times N$ matrix

$$\mathbf{S}\equiv\begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}, \tag{T9.43}$$

where $\mathbf{Q}$ is an *arbitrary* full-rank $(N-K)\times N$ matrix. It will soon become evident that the choice of $\mathbf{Q}$ is irrelevant.

Then we compute the probability density of the following random variable:

$$\mathbf{Y} \equiv \mathbf{SX} = \begin{pmatrix} \mathbf{QX} \\ \mathbf{PX} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{Y}_A \\ \mathbf{Y}_B \end{pmatrix} \tag{T9.44}$$

It is immediate to check that $\mathbf{Y}$ is normal:

$$\mathbf{Y} \sim \mathrm{N}(\boldsymbol{\nu}, \mathbf{T}) \tag{T9.45}$$

where

$$\boldsymbol{\nu} \equiv \begin{pmatrix} \boldsymbol{\nu}_A \\ \boldsymbol{\nu}_B \end{pmatrix} = \begin{pmatrix} \mathbf{Q}\boldsymbol{\mu} \\ \mathbf{P}\boldsymbol{\mu} \end{pmatrix} \tag{T9.46}$$

and

$$\mathbf{T} \equiv \begin{pmatrix} \mathbf{T}_{AA} & \mathbf{T}_{AB} \\ \mathbf{T}_{BA} & \mathbf{T}_{BB} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{Q\Sigma Q}' & \mathbf{Q\Sigma P}' \\ \mathbf{P\Sigma Q}' & \mathbf{P\Sigma P}' \end{pmatrix}. \tag{T9.47}$$

At this point we can compute the conditional probability density. From (2.164) we obtain

$$\mathbf{Y}_A|\mathbf{y}_B \sim \mathrm{N}(\boldsymbol{\xi}, \mathit{\Phi}) \tag{T9.48}$$

where the expected values and covariance read explicitly

$$\begin{aligned} \boldsymbol{\xi} &\equiv \boldsymbol{\nu}_A + \mathbf{T}_{AB}\mathbf{T}_{BB}^{-1}(\mathbf{y}_B - \boldsymbol{\nu}_B) \\ \boldsymbol{\Phi} &\equiv \mathbf{T}_{AA} - \mathbf{T}_{AB}\mathbf{T}_{BB}^{-1}\mathbf{T}_{BA} \end{aligned} \tag{T9.49}$$

Substituting the investor's opinion $\mathbf{y}_B = \mathbf{Px} = \mathbf{v}$ in (*T*9.49) we obtain therefore that the expected value of the whole vector $\mathbf{y}$ is

$$\begin{aligned} \mathrm{E}\{\mathbf{Y}|\mathbf{Y}_B = \mathbf{v}\} &= \begin{pmatrix} \mathrm{E}\{\mathbf{Y}_A|\mathbf{Y}_B = \mathbf{v}\} \\ \mathbf{v} \end{pmatrix} = \\ &\begin{pmatrix} \mathbf{Q}\boldsymbol{\mu} + \mathbf{Q\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\mu}) \\ \mathbf{v} \end{pmatrix} \end{aligned} \tag{T9.50}$$

Recalling that $\mathbf{Y} = \mathbf{SX}$ and rewriting $\mathbf{q}$ as $\mathbf{P}\boldsymbol{\mu} + \mathbf{P\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}(\mathbf{v} - \mathbf{Pv})$ we can express (*T*9.50) as follows:

$$\mathbf{S}\,\mathrm{E}\{\mathbf{X}|\mathbf{PX} = \mathbf{v}\} = \mathbf{S}\left(\boldsymbol{\mu} + \mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\mu})\right). \tag{T9.51}$$

Since $\mathbf{S}$ is invertible we finally obtain

$$\mathrm{E}\{\mathbf{X}|\mathbf{PX} = \mathbf{v}\} = \boldsymbol{\mu} + \mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}(\mathbf{v} - \mathbf{P}\boldsymbol{\mu}), \tag{T9.52}$$

which is the expression of the conditional expectation that we were looking for.

As for the covariance matrix, substituting the investor's opinion $\mathbf{Y}_B = \mathbf{PX} = \mathbf{v}$ in (*T*9.49) we obtain therefore that the expected value of the whole vector $\mathbf{y}$ is

$$\begin{aligned}\operatorname{Cov}\{\mathbf{Y}|\mathbf{Y}_B=\mathbf{v}\} &= \begin{pmatrix}\operatorname{Cov}\{\mathbf{Y}_A|\mathbf{Y}_B=\mathbf{v}\} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\\ &= \begin{pmatrix}\mathbf{\Phi} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \qquad (T9.53)\\ &= \begin{pmatrix}\mathbf{Q\Sigma Q}'-\mathbf{Q\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma Q}' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}.\end{aligned}$$

Recalling that $\mathbf{Y}=\mathbf{SX}$ we can write

$$a \equiv \mathbf{S}\operatorname{Cov}\{\mathbf{X}|\mathbf{PX}=\mathbf{v}\}\mathbf{S}' \qquad (T9.54)$$

as follows

$$\begin{aligned}a &= \operatorname{Cov}\{\mathbf{SX}|\mathbf{PX}=\mathbf{v}\}\\ &= \begin{pmatrix}\mathbf{Q}\left(\mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}\right)\mathbf{Q}' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \qquad (T9.55)\\ &= \begin{pmatrix}\mathbf{Q}\left(\mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}\right)\mathbf{Q}' & \mathbf{Q}\left(\mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}\right)\mathbf{P}'\\ \mathbf{P}\left(\mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}\right)\mathbf{Q}' & \mathbf{P}\left(\mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}\right)\mathbf{P}'\end{pmatrix}\\ &= \begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}\left(\mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}\right)\begin{pmatrix}\mathbf{Q}\\ \mathbf{P}\end{pmatrix}'\end{aligned}$$

Since $\mathbf{S}$ is invertible we can pre- and post- multiply (T9.54) by $\mathbf{S}^{-1}$ and finally obtain:

$$\operatorname{Cov}\{\mathbf{X}|\mathbf{PX}=\mathbf{v}\} = \mathbf{\Sigma}-\mathbf{\Sigma P}'\left(\mathbf{P\Sigma P}'\right)^{-1}\mathbf{P\Sigma}, \qquad (T9.56)$$

which is the expression of the conditional covariance that we were looking for.

9.5 Computations for the robust version of the leading example

From (8.33), (8.25) and (8.29) we obtain:

$$\boldsymbol{\alpha}^* \equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\left\{\max_{\boldsymbol{\mu}\in\widehat{\boldsymbol{\Theta}}_p}\left\{\begin{array}{c}\frac{\zeta}{2}\left(\boldsymbol{\mu}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}-\frac{1}{A}\left(\mathbf{1}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}\right)^2\right)\\ +w_T\left(1+\frac{1}{A}\mathbf{1}'\mathbf{\Sigma}^{-1}\boldsymbol{\mu}-\frac{w_T}{\zeta}\frac{1}{2A}\right)\\ -\boldsymbol{\alpha}'\operatorname{diag}(\mathbf{p}_T)(\mathbf{1}+\boldsymbol{\mu})+\frac{1}{2\zeta}\boldsymbol{\alpha}'\mathbf{\Phi}\boldsymbol{\alpha}\end{array}\right\}\right\} \qquad (T9.57)$$

$$\text{s.t. } \begin{cases}\boldsymbol{\alpha}'\mathbf{p}_T = w_T\\ (1-\gamma)\,w_T-\boldsymbol{\alpha}'\operatorname{diag}(\mathbf{p}_T)(\mathbf{1}+\boldsymbol{\mu})+\sqrt{2\boldsymbol{\alpha}'\mathbf{\Phi}\boldsymbol{\alpha}}\lambda\le 0, \text{ for all } \boldsymbol{\mu}\in\widehat{\boldsymbol{\Theta}}_p\end{cases}$$

where

$$\begin{aligned}A &\equiv \mathbf{1}'\mathbf{\Sigma}^{-1}\mathbf{1},\ \lambda\equiv\operatorname{erf}^{-1}(2c-1) \qquad (T9.58)\\ \mathbf{\Phi} &\equiv \operatorname{diag}(\mathbf{p}_T)\,\mathbf{\Sigma}\operatorname{diag}(\mathbf{p}_T).\end{aligned}$$

and

$$\widehat{\boldsymbol{\Theta}}_p \equiv \left\{ \boldsymbol{\mu} \text{ such that } (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}) \leq \frac{Q_N(p)}{T} \right\}, \tag{T9.59}$$

Using the budget constraint this becomes:

$$\begin{aligned} \boldsymbol{\alpha}^* \equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} & \left\{ \max_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_p} \left\{ \begin{array}{c} \frac{\zeta}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{\zeta}{2A} \left(\mathbf{1}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right)^2 + \frac{w_T}{A} \left(\mathbf{1}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right) \\ -\boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\mu} + \frac{1}{2\zeta} \boldsymbol{\alpha}' \boldsymbol{\Phi} \boldsymbol{\alpha} \end{array} \right\} \right\} \\ \text{s.t.} & \left\{ \begin{array}{l} \boldsymbol{\alpha}' \mathbf{p}_T = w_T \\ \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\mu} \geq \sqrt{2 \boldsymbol{\alpha}' \boldsymbol{\Phi} \boldsymbol{\alpha}} \lambda - \gamma w, \text{ for all } \boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_p \end{array} \right. \end{aligned} \tag{T9.60}$$

or:

$$\begin{aligned} \boldsymbol{\alpha}^* \equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} & \left\{ \max_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_p} \left\{ \begin{array}{c} \frac{\zeta}{2} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{\zeta}{2A} \left(\mathbf{1}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right)^2 + \frac{w_T}{A} \left(\mathbf{1}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right) \\ -\boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\mu} + \frac{1}{2\zeta} \boldsymbol{\alpha}' \boldsymbol{\Phi} \boldsymbol{\alpha} \end{array} \right\} \right\} \\ \text{s.t.} & \left\{ \begin{array}{l} \boldsymbol{\alpha}' \mathbf{p} = w \\ \max_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_p} \left\{ \sqrt{2 \boldsymbol{\alpha}' \boldsymbol{\Phi} \boldsymbol{\alpha}} \lambda - \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\mu} \right\} \leq \gamma w. \end{array} \right. \end{aligned} \tag{T9.61}$$

The second maximization

$$\max_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_p} \left\{ \sqrt{2 \boldsymbol{\alpha}' \boldsymbol{\Phi} \boldsymbol{\alpha}} \lambda - \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\mu} \right\} \leq \gamma w \tag{T9.62}$$

is maximization constrained on an ellipsoid of contour surfaces that describe parallel hyperplanes. The tangency condition is achieved when the gradients are parallel. For the gradient of the ellipsoid we have

$$\mathbf{g}_{\widehat{\boldsymbol{\Theta}}_p} \propto \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}). \tag{T9.63}$$

For the gradient of the hyperplane we have:

$$\mathbf{g}_H \propto \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha}. \tag{T9.64}$$

Therefore the maximum is achieved when there exists a ρ such that:

$$\boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}) = \rho \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha} \tag{T9.65}$$

Since $\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_p$ we have

$$\begin{aligned} \frac{Q_N(p)}{T} &= (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}) \\ &= (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})' \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}) \\ &= \rho^2 \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\Sigma} \operatorname{diag}(\mathbf{p}_T) \boldsymbol{\alpha} \end{aligned} \tag{T9.66}$$

Therefore

$$\rho = \pm\sqrt{\frac{Q_N(p)}{T}\frac{1}{\boldsymbol{\alpha}'\operatorname{diag}(\mathbf{p}_T)\boldsymbol{\Sigma}\operatorname{diag}(\mathbf{p}_T)\boldsymbol{\alpha}}} \tag{T9.67}$$

Substituting in (T9.65) we obtain

$$\boldsymbol{\mu} = \widehat{\boldsymbol{\mu}} - \sqrt{\frac{Q_N(p)/T}{\boldsymbol{\alpha}'\operatorname{diag}(\mathbf{p}_T)\boldsymbol{\Sigma}\operatorname{diag}(\mathbf{p}_T)\boldsymbol{\alpha}}}\boldsymbol{\Sigma}\operatorname{diag}(\mathbf{p}_T)\boldsymbol{\alpha}, \tag{T9.68}$$

where the choice of the sign follows from the maximization (T9.62).

Therefore the original problem (T9.61) reads:

$$\boldsymbol{\alpha}^* \equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmin}}\left\{\max_{\boldsymbol{\mu}\in\widehat{\boldsymbol{\Theta}}_p}\left\{\boldsymbol{\mu}'\mathbf{T}\boldsymbol{\mu} + \frac{w}{A}\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \boldsymbol{\alpha}'\operatorname{diag}(\mathbf{p})\boldsymbol{\mu} + \frac{1}{2\zeta}\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}\right\}\right\} \tag{T9.69}$$
$$\text{s.t.}\begin{cases}\boldsymbol{\alpha}'\mathbf{p} = w\\ \sqrt{2\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}}\lambda + \sqrt{\frac{Q_N(p)/T}{\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}}} - \boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha} - \boldsymbol{\alpha}'\operatorname{diag}(\mathbf{p}_T)\widehat{\boldsymbol{\mu}} \leq \gamma w.\end{cases},$$

where

$$\mathbf{T} \equiv \frac{\zeta}{2}\boldsymbol{\Sigma}^{-1} - \frac{\zeta}{2A}\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}'\boldsymbol{\Sigma}^{-1}. \tag{T9.70}$$

9.6 Computations for the robust mean-variance problem

Taking into account the elliptical/certain specifications (9.118)-(9.119), the robust mean-variance problem (9.117) can be written as follows:

$$\boldsymbol{\alpha}_{\mathrm{r}}^{(i)} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}}\left\{\min_{\boldsymbol{\mu}\in\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}}\{\boldsymbol{\alpha}'\boldsymbol{\mu}\}\right\} \tag{T9.71}$$
$$\text{subject to }\begin{cases}\boldsymbol{\alpha}\in\mathcal{C}\\ \boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha} \leq v_i,\end{cases}$$

where

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{\boldsymbol{\mu} : (\boldsymbol{\mu}-\mathbf{m})'\mathbf{T}^{-1}(\boldsymbol{\mu}-\mathbf{m}) \leq q^2\right\}. \tag{T9.72}$$

Consider the spectral decomposition (A.70) of the shape parameter:

$$\mathbf{T} \equiv \mathbf{E}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\mathbf{E}'. \tag{T9.73}$$

Then:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{\boldsymbol{\mu} : (\boldsymbol{\mu}-\mathbf{m})'\mathbf{E}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\Lambda}^{-1/2}\mathbf{E}'(\boldsymbol{\mu}-\mathbf{m}) \leq q^2\right\}. \tag{T9.74}$$

Define the variable

$$\mathbf{u} \equiv \frac{1}{q}\boldsymbol{\Lambda}^{-1/2}\mathbf{E}'(\boldsymbol{\mu}-\mathbf{m}), \tag{T9.75}$$

which implies

$$\boldsymbol{\mu} \equiv \mathbf{m} + q\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}. \tag{T9.76}$$

Then

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{\mathbf{m} + q\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u} : \mathbf{u}'\mathbf{u} \leq 1\right\}. \tag{T9.77}$$

We can express the minimization in (T9.71) as follows:

$$\begin{aligned} \min_{\boldsymbol{\mu}\in\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}} \{\boldsymbol{\alpha}'\boldsymbol{\mu}\} &= \min_{\mathbf{u}'\mathbf{u}\leq 1} \left\{\boldsymbol{\alpha}'\left(\mathbf{m} + q\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}\right)\right\} \\ &= \boldsymbol{\alpha}'\mathbf{m} + q \min_{\mathbf{u}'\mathbf{u}\leq 1} \left\{\boldsymbol{\alpha}'\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}\right\} \\ &= \boldsymbol{\alpha}'\mathbf{m} + q \min_{\mathbf{u}'\mathbf{u}\leq 1} \left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}, \mathbf{u}\right\rangle, \end{aligned} \tag{T9.78}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product (A.5). This scalar product reaches a minimum when the vector $\mathbf{u}$ is opposite to the other term in the product:

$$\widetilde{\mathbf{u}} \equiv -\frac{\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}}{\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|}, \tag{T9.79}$$

and the respective minimum reads:

$$\begin{aligned} \min_{\mathbf{u}'\mathbf{u}\leq 1} \left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}, \mathbf{u}\right\rangle &= \left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}, \widetilde{\mathbf{u}}\right\rangle \\ &= \left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}, -\frac{\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}}{\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|}\right\rangle \\ &= -\frac{1}{\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|}\left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}, \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\rangle \\ &= -\frac{1}{\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|}\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|^2 \\ &= -\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|. \end{aligned} \tag{T9.80}$$

Substituting (T9.80) in (T9.78), the original problem (T9.71) reads:

$$\begin{aligned} \boldsymbol{\alpha}_{\mathrm{r}}^{(i)} &= \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{\boldsymbol{\alpha}'\mathbf{m} - q\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|\right\} \\ &\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ \boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha} \leq v_i. \end{cases} \end{aligned} \tag{T9.81}$$

Equivalently, from (T9.73) we can write:

$$\begin{aligned} \boldsymbol{\alpha}_{\mathrm{r}}^{(i)} &= \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{\boldsymbol{\alpha}'\mathbf{m} - q\sqrt{\boldsymbol{\alpha}'\mathbf{T}\boldsymbol{\alpha}}\right\} \\ &\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ \boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha} \leq v_i. \end{cases} \end{aligned} \tag{T9.82}$$

To put the problem ($T9.81$) in the SOCP form (6.55) we introducing an auxiliary variable z:

$$\left(\boldsymbol{\alpha}_{\mathrm{r}}^{(i)}, z_{\mathrm{r}}^{(i)}\right) = \underset{\boldsymbol{\alpha}, z}{\operatorname{argmax}} \left\{\boldsymbol{\alpha}'\mathbf{m} - z\right\} \qquad (T9.83)$$

$$\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ q\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\| \leq z \\ \boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha} \leq v_i. \end{cases}$$

Furthermore, considering the spectral decomposition ($A.70$) of the estimate of the covariance

$$\widehat{\boldsymbol{\Sigma}} \equiv \mathbf{F}\boldsymbol{\Gamma}^{1/2}\boldsymbol{\Gamma}^{1/2}\mathbf{F}', \qquad (T9.84)$$

we can write

$$\boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha} = \left\langle \boldsymbol{\Gamma}^{1/2}\mathbf{F}'\boldsymbol{\alpha}, \boldsymbol{\Gamma}^{1/2}\mathbf{F}'\boldsymbol{\alpha} \right\rangle. \qquad (T9.85)$$

Therefore the mean-variance problem ($T9.83$) can be written as follows:

$$\left(\boldsymbol{\alpha}_{\mathrm{r}}^{(i)}, z_{\mathrm{r}}^{(i)}\right) = \underset{\boldsymbol{\alpha}, z}{\operatorname{argmax}} \left\{\boldsymbol{\alpha}'\mathbf{m} - z\right\} \qquad (T9.86)$$

$$\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ q\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\| \leq z \\ \left\|\boldsymbol{\Gamma}^{1/2}\mathbf{F}'\boldsymbol{\alpha}\right\| \leq \sqrt{v_i}. \end{cases}$$

If the investment constraints $\mathcal{C}$ are regular enough, this problem is in the SOCP form (6.55).

9.7 Restating the robust mean-variance problem in SeDuMi format

Let us define

$$\mathbf{x} \equiv (\boldsymbol{\alpha}', z)' \qquad (T9.87)$$

and let us assume that $\mathcal{C}$ represents the full-budget constraints and the long-only constraints:

$$\sum_{n=1}^{N} x_n = 1 \qquad (T9.88)$$

$$x_n \geq 0, \quad n = 1, \ldots, N \qquad (T9.89)$$

We redefine the following quantities to re-express our problem

- Target

$$\mathbf{b} \equiv (-\mathbf{m}', 1)' \qquad (T9.90)$$

- Long-only and budget constraints:

$$\mathbf{D}_{lo} \equiv [\mathbf{I}_N | \mathbf{0}_N] \quad (T9.91)$$

$$\mathbf{f}_{lo} \equiv \mathbf{0}_N \quad (T9.92)$$

$$\mathbf{D}_{b1} \equiv [\mathbf{1}'_N | 0] \quad (T9.93)$$

$$f_{b1} \equiv -1 \quad (T9.94)$$

$$\mathbf{D}_{b2} \equiv [-\mathbf{1}'_N | 0] \quad (T9.95)$$

$$f_{b2} \equiv 1 \quad (T9.96)$$

$$\mathbf{D} \equiv \begin{pmatrix} \mathbf{D}_{lo} \\ \mathbf{D}_{b1} \\ \mathbf{D}_{b2} \end{pmatrix}' \quad (T9.97)$$

$$\mathbf{f} \equiv \begin{pmatrix} \mathbf{f}_{lo} \\ f_{b1} \\ f_{b2} \end{pmatrix} \quad (T9.98)$$

- Estimation error

$$\mathbf{A}'_1 \equiv \left[q \mathbf{\Lambda}^{1/2} \mathbf{E}' | \mathbf{0}_N \right] \quad (T9.99)$$

$$\mathbf{b}'_1 \equiv [\mathbf{0}'_N | 1] \quad (T9.100)$$

$$d_1 \equiv 0 \quad (T9.101)$$

$$\mathbf{c}_1 \equiv \mathbf{0}_N \quad (T9.102)$$

- Variance

$$\mathbf{A}'_2 \equiv \left[\mathbf{\Gamma}^{1/2} \mathbf{F}' | \mathbf{0}_N \right] \quad (T9.103)$$

$$\mathbf{b}'_2 \equiv \left[\mathbf{0}'_{N+1} \right] \quad (T9.104)$$

$$d_2 \equiv \sqrt{v_i} \quad (T9.105)$$

$$\mathbf{c}_2 \equiv \mathbf{0}_N \quad (T9.106)$$

Then our problem ($T9.86$) reads:

$$\mathbf{x}^* = \underset{\mathbf{y}}{\operatorname{argmin}} \{\mathbf{b}'\mathbf{x}\} \quad (T9.107)$$

subject to

$$\mathbf{D}'\mathbf{x} + \mathbf{f} \geq \mathbf{0} \quad (T9.108)$$

$$\|\mathbf{A}'_1 \mathbf{x} + \mathbf{c}_1\| \leq \mathbf{b}'_1 \mathbf{x} + d_1 \quad (T9.109)$$

$$\|\mathbf{A}'_2 \mathbf{x} + \mathbf{c}_2\| \leq \mathbf{b}'_2 \mathbf{x} + d_2 \quad (T9.110)$$

This problem is in the standard SeDuMi format.

9.8 Normal predictive distribution

More in general, consider

$$\begin{aligned} f_{\mathbf{M}}(\mathbf{m}) &\equiv \int f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\mathbf{m})\, f_{\boldsymbol{\nu},\boldsymbol{\Phi}}(\boldsymbol{\mu})\, d\boldsymbol{\mu} \\ &= \int \frac{e^{-\frac{1}{2}(\mathbf{m}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{m}-\boldsymbol{\mu})}}{(2\pi)^{\frac{N}{2}}\sqrt{|\boldsymbol{\Sigma}|}} \frac{e^{-\frac{1}{2}(\boldsymbol{\mu}-\boldsymbol{\nu})'\boldsymbol{\Phi}^{-1}(\boldsymbol{\mu}-\boldsymbol{\nu})}}{(2\pi)^{\frac{N}{2}}\sqrt{|\boldsymbol{\Phi}|}} d\boldsymbol{\mu} \\ &= \frac{(2\pi)^{-N}}{\sqrt{|\boldsymbol{\Sigma}|}\sqrt{|\boldsymbol{\Phi}|}} \int e^{-\frac{1}{2}a} d\boldsymbol{\mu}, \end{aligned} \tag{T9.111}$$

where

$$\begin{aligned} a &\equiv (\mathbf{m}-\boldsymbol{\mu})'\,\boldsymbol{\Sigma}^{-1}(\mathbf{m}-\boldsymbol{\mu}) + (\boldsymbol{\mu}-\boldsymbol{\nu})'\,\boldsymbol{\Phi}^{-1}(\boldsymbol{\mu}-\boldsymbol{\nu}) \\ &= \mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m} + \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - 2\mathbf{m}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \boldsymbol{\mu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\mu} + \boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} - 2\boldsymbol{\mu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\ &= \boldsymbol{\mu}'\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)\boldsymbol{\mu} - 2\boldsymbol{\mu}'\left(\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right) + \mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m} + \boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \end{aligned} \tag{T9.112}$$

Defining

$$\mathbf{b} \equiv \left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right) \tag{T9.113}$$

we can write

$$\begin{aligned} a = &(\boldsymbol{\mu}-\mathbf{b})'\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)(\boldsymbol{\mu}-\mathbf{b}) \\ &-\mathbf{b}'\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)\mathbf{b} + \mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m} + \boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \end{aligned} \tag{T9.114}$$

Therefore (T9.111) becomes:

$$\begin{aligned} f_{\mathbf{M}}(\mathbf{m}) = &\frac{(2\pi)^{-N}}{\sqrt{|\boldsymbol{\Sigma}|}\sqrt{|\boldsymbol{\Phi}|}} e^{-\frac{1}{2}\left[-\mathbf{b}'\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)\mathbf{b}+\mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right]} \\ &(2\pi)^{\frac{N}{2}}\left|\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right|^{-\frac{1}{2}} \int \frac{e^{-\frac{1}{2}(\boldsymbol{\mu}-\mathbf{b})'\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)(\boldsymbol{\mu}-\mathbf{b})}}{(2\pi)^{\frac{N}{2}}\left|\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right|^{-\frac{1}{2}}} d\boldsymbol{\mu} \\ = &\gamma_2 e^{-\frac{1}{2}c} \end{aligned} \tag{T9.115}$$

where γ_2 is a normalization constant and

$$\begin{aligned}
c &\equiv -\mathbf{b}'\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)\mathbf{b}+\mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\
&= -\left(\mathbf{m}'\boldsymbol{\Sigma}^{-1}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\right)\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right) \\
&\quad +\mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \qquad (T9.116) \\
&= -\left(\mathbf{m}+\boldsymbol{\Sigma}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right)'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\left(\mathbf{m}+\boldsymbol{\Sigma}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right) \\
&\quad +\mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\
&= -\mathbf{m}'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\mathbf{m}-\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\
&\quad -2\mathbf{m}'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}+\mathbf{m}'\boldsymbol{\Sigma}^{-1}\mathbf{m}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\
&= \mathbf{m}'\left[\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right]\mathbf{m} \\
&\quad -2\mathbf{m}'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\
&\quad -\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}+\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}
\end{aligned}$$

Defining

$$\begin{aligned}
\mathbf{T} &\equiv \boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1} \qquad (T9.117) \\
\mathbf{Tg} &\equiv \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\
h &\equiv \boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}-\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}
\end{aligned}$$

we obtain

$$\begin{aligned}
c &= \mathbf{m}'\mathbf{Tm}-2\mathbf{m}'\mathbf{Tg}+h \qquad (T9.118) \\
&= \mathbf{m}'\mathbf{Tm}-2\mathbf{m}'\mathbf{Tg}+\mathbf{g}'\mathbf{Tg}-\mathbf{g}'\mathbf{Tg}+h \\
&= (\mathbf{m}-\mathbf{g})'\,\mathbf{T}\,(\mathbf{m}-\mathbf{g})-\mathbf{g}'\mathbf{Tg}+h
\end{aligned}$$

Since

$$\begin{aligned}
\mathbf{g} = &\left[\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right]^{-1} \qquad (T9.119) \\
&\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}
\end{aligned}$$

the term $h-\mathbf{g}'\mathbf{Tg}$ cancels (???):

$$\begin{aligned} h-\mathbf{g}'\mathbf{T}\mathbf{g} &= \boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}-\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\ &\quad -\left[\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}\right]' \\ &\quad \left[\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right]^{-1} \\ &\quad \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\ &= \boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}-\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu} \\ &\quad -\boldsymbol{\nu}'\boldsymbol{\Phi}^{-1}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1} \\ &\quad \left[\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right]^{-1} \\ &\quad \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Phi}^{-1}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{\nu}. \end{aligned} \tag{T9.120}$$

Therefore

$$f_{\mathbf{M}}(\mathbf{m})=\gamma_2 e^{-\frac{1}{2}(\mathbf{m}-\mathbf{g})'\mathbf{T}(\mathbf{m}-\mathbf{g})}. \tag{T9.121}$$

or in other words

$$\mathbf{M}\sim \mathrm{N}\left(\mathbf{g},\mathbf{T}^{-1}\right). \tag{T9.122}$$

In our example

$$\boldsymbol{\Phi}\equiv\frac{\boldsymbol{\Sigma}}{T}. \tag{T9.123}$$

Therefore

$$\begin{aligned} \mathbf{g} &= \left[\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+T\boldsymbol{\Sigma}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1}\right]^{-1} \\ &\quad \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+T\boldsymbol{\Sigma}^{-1}\right)^{-1}T\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu} \\ &= \left[\boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\frac{1}{1+T}\right]^{-1}\frac{T}{1+T}\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu} \\ &= \frac{1+T}{T}\boldsymbol{\Sigma}\frac{T}{1+T}\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu} \\ &= \boldsymbol{\nu} \end{aligned} \tag{T9.124}$$

and

$$\begin{aligned} \mathbf{T} &= \boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\Sigma}^{-1}+T\boldsymbol{\Sigma}^{-1}\right)^{-1}\boldsymbol{\Sigma}^{-1} \\ &= \boldsymbol{\Sigma}^{-1}-\boldsymbol{\Sigma}^{-1}\frac{1}{1+T} \\ &= \frac{T}{1+T}\boldsymbol{\Sigma}^{-1} \end{aligned} \tag{T9.125}$$

9.9 Computations for the Bayesian robust mean-variance problem

The robustness uncertainty set for $\boldsymbol{\mu}$

Consider the ellipsoid (9.149):

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}\left[i_T, e_C\right] \equiv \left\{\boldsymbol{\mu} : \left(\boldsymbol{\mu} - \boldsymbol{\mu}_1\right)' \boldsymbol{\Sigma}_1^{-1} \left(\boldsymbol{\mu} - \boldsymbol{\mu}_1\right) \leq \frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right\}. \qquad (T9.126)$$

Consider the spectral decomposition of the dispersion parameter:

$$\boldsymbol{\Sigma}_1 \equiv \mathbf{F} \boldsymbol{\Gamma}^{1/2} \boldsymbol{\Gamma}^{1/2} \mathbf{F}', \qquad (T9.127)$$

where $\mathbf{F}$ is the juxtaposition $(A.62)$ of the eigenvectors and $\boldsymbol{\Gamma}$ is the diagonal matrix $(A.65)$ of the eigenvalues.

We can write $(T9.126)$ as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{\boldsymbol{\mu} : \left(\boldsymbol{\mu} - \boldsymbol{\mu}_1\right)' \mathbf{F} \boldsymbol{\Gamma}^{-1/2} \boldsymbol{\Gamma}^{-1/2} \mathbf{F}' \left(\boldsymbol{\mu} - \boldsymbol{\mu}_1\right) \leq \frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right\}. \qquad (T9.128)$$

Define the new variable:

$$\mathbf{u} \equiv \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2\right)^{-1/2} \boldsymbol{\Gamma}^{-1/2} \mathbf{F}' \left(\boldsymbol{\mu} - \boldsymbol{\mu}_1\right), \qquad (T9.129)$$

which implies

$$\boldsymbol{\mu} = \boldsymbol{\mu}_1 + \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2\right)^{1/2} \mathbf{F} \boldsymbol{\Gamma}^{1/2} \mathbf{u}, \qquad (T9.130)$$

we can write $(T9.128)$ as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{\boldsymbol{\mu}_1 + \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2\right)^{1/2} \mathbf{F} \boldsymbol{\Gamma}^{1/2} \mathbf{u}, \quad \mathbf{u}'\mathbf{u} \leq 1 \right\}.$$

Since

$$\boldsymbol{\omega}'\boldsymbol{\mu} = \left\langle \boldsymbol{\omega}, \boldsymbol{\mu}_1 + \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_N^{p_{\boldsymbol{\mu}}}\right)^{1/2} \mathbf{F} \boldsymbol{\Gamma}^{1/2} \mathbf{u} \right\rangle \qquad (T9.131)$$

$$\langle \boldsymbol{\omega}, \boldsymbol{\mu}_1 \rangle + \left\langle \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_N^{p_{\boldsymbol{\mu}}}\right)^{1/2} \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega}, \mathbf{u} \right\rangle$$

we have

$$\min_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}} \left\{\boldsymbol{\omega}'\boldsymbol{\mu}\right\} = \langle \boldsymbol{\omega}, \boldsymbol{\mu}_1 \rangle \qquad (T9.132)$$

$$+ \min_{\mathbf{u}'\mathbf{u} \leq 1} \left\langle \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_N^{p_{\boldsymbol{\mu}}}\right)^{1/2} \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega}, \mathbf{u} \right\rangle$$

$$= \boldsymbol{\omega}'\boldsymbol{\mu}_1 - \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_N^{p_{\boldsymbol{\mu}}}\right)^{1/2} \left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega} \right\|$$

The robustness uncertainty set for $\boldsymbol{\Sigma}$

Consider the ellipsoid (9.152):

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} \equiv \left\{ \boldsymbol{\Sigma} : \ \operatorname{vech}\left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]' \mathbf{S}_{\boldsymbol{\Sigma}}^{-1} \operatorname{vech}\left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right] \leq q_{\boldsymbol{\Sigma}}^2 \right\}, \qquad (T9.133)$$

where

$$\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\left[i_T, e_C\right] = \frac{\nu_1}{\nu_1 + N + 1} \boldsymbol{\Sigma}_1; \qquad (T9.134)$$

and $\mathbf{S}_{\boldsymbol{\Sigma}}$ is the dispersion parameter of $\boldsymbol{\Sigma}$:

$$\mathbf{S}_{\boldsymbol{\Sigma}}\left[i_T, e_C\right] = \frac{2\nu_1^2}{\left(\nu_1 + N + 1\right)^3} \left(\mathbf{D}_N' \left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}\right) \mathbf{D}_N\right)^{-1}. \qquad (T9.135)$$

Consider the spectral decomposition of the rescaled dispersion parameter (T7.75);

$$\left(\mathbf{D}_N' \left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}\right) \mathbf{D}_N\right)^{-1} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}', \qquad (T9.136)$$

where $\mathbf{E}$ is the juxtaposition (A.62) of the eigenvectors:

$$\mathbf{E} \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(N(N+1)/2)}\right); \qquad (T9.137)$$

and $\boldsymbol{\Lambda}$ is the diagonal matrix (A.65) of the eigenvalues:

$$\boldsymbol{\Lambda} \equiv \operatorname{diag}\left(\lambda_1, \ldots, \lambda_{N(N+1)/2}\right). \qquad (T9.138)$$

We can write (T9.133) as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} \equiv \left\{ \operatorname{vech}\left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]' \mathbf{E}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{\Lambda}^{-1/2}\mathbf{E}' \operatorname{vech}\left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right] \leq \frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1 + N + 1\right)^3} \right\}. \qquad (T9.139)$$

Define the new variable:

$$\mathbf{u} \equiv \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1 + N + 1\right)^3} \right)^{-1/2} \boldsymbol{\Lambda}^{-1/2}\mathbf{E}' \operatorname{vech}\left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right], \qquad (T9.140)$$

which implies

$$\operatorname{vech}\left[\boldsymbol{\Sigma}\right] \equiv \operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right] + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1 + N + 1\right)^3} \right)^{1/2} \mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}. \qquad (T9.141)$$

we can write (T9.139) as follows

$$\begin{aligned} \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} &\equiv \left\{ \operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right] + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1 + N + 1\right)^3} \right)^{1/2} \mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}, \quad \mathbf{u}'\mathbf{u} \leq 1 \right\} \qquad (T9.142) \\ &= \left\{ \operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right] + \sum_{s=1}^{N(N+1)/2} \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2 \lambda_s}{\left(\nu_1 + N + 1\right)^3} \right)^{1/2} \mathbf{e}^{(s)} u_s, \quad \mathbf{u}'\mathbf{u} \leq 1 \right\} \end{aligned}$$

Each eigenvector $\mathbf{e}^{(s)}$ represents the non-redundant entries of a matrix. To consider all the elements we simply multiply by the duplication matrix (A.113).

Then from (T9.141) we obtain:

$$\begin{aligned}\boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega} &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}') \operatorname{vec}[\boldsymbol{\Sigma}] \\ &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}') \mathbf{D}_N \operatorname{vech}[\boldsymbol{\Sigma}] \\ &= \left\langle \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')', \operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\text{ce}}\right] + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3}\right)^{1/2} \mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u} \right\rangle \\ &= \left\langle \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')', \operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\text{ce}}\right] \right\rangle \\ &\quad + \left\langle \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')', \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3}\right)^{1/2} \mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u} \right\rangle \\ &= \boldsymbol{\omega}'\widehat{\boldsymbol{\Sigma}}_{\text{ce}}\boldsymbol{\omega} \\ &\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3}\right)^{1/2} \left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')', \mathbf{u} \right\rangle\end{aligned} \tag{T9.143}$$

Therefore

$$\begin{aligned}\max_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \{\boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega}\} &= \boldsymbol{\omega}'\widehat{\boldsymbol{\Sigma}}_{\text{ce}}\boldsymbol{\omega} \\ &\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3}\right)^{1/2} \max_{\mathbf{u}'\mathbf{u} \leq 1} \left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')', \mathbf{u} \right\rangle \\ &= \boldsymbol{\omega}'\widehat{\boldsymbol{\Sigma}}_{\text{ce}}\boldsymbol{\omega} \\ &\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3}\right)^{1/2} \left\| \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \right\|.\end{aligned} \tag{T9.144}$$

Substituting (T9.134) this becomes

$$\begin{aligned}\max_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \boldsymbol{\omega}'\boldsymbol{\Sigma}\boldsymbol{\omega} &= \frac{\nu_1}{\nu_1 + N + 1}\boldsymbol{\omega}'\boldsymbol{\Sigma}_1\boldsymbol{\omega} \\ &\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3}\right)^{1/2} \left\| \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}') \right\|\end{aligned} \tag{T9.145}$$

To simplify this expression, consider the pseudo inverse $\widetilde{\mathbf{D}}$ of the duplication matrix:

$$\widetilde{\mathbf{D}}_N \mathbf{D}_N = \mathbf{I}_{N(N+1)/2}. \tag{T9.146}$$

It is possible to show that

$$\left(\mathbf{D}_N' \left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}\right) \mathbf{D}_N\right)^{-1} = \widetilde{\mathbf{D}}_N \left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}\right)^{-1} \widetilde{\mathbf{D}}_N' \tag{T9.147}$$

and

$$(\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, \mathbf{D}_N \widetilde{\mathbf{D}}_N = (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\,, \tag{T9.148}$$

see Magnus and Neudecker (1999).

Now consider the square of the norm in (T9.145). Using (T9.147) and (T9.148) we obtain:

$$\begin{aligned} a &\equiv \left\| \boldsymbol{\Lambda}^{1/2} \mathbf{E}' \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \right\|^2 \\ &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, \mathbf{D}_N \mathbf{E} \boldsymbol{\Lambda}^{1/2} \boldsymbol{\Lambda}^{1/2} \mathbf{E}' \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \\ &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, \mathbf{D}_N \left(\mathbf{D}_N' \left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1} \right) \mathbf{D}_N \right)^{-1} \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \\ &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, \mathbf{D}_N \widetilde{\mathbf{D}}_N (\boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_1)\, \widetilde{\mathbf{D}}_N' \mathbf{D}_N' (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \\ &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, \mathbf{D}_N \widetilde{\mathbf{D}}_N (\boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_1) \left[(\boldsymbol{\omega}' \otimes \boldsymbol{\omega}') \left(\mathbf{D}_N \widetilde{\mathbf{D}}_N \right) \right]' \\ &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, (\boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_1)\, (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \end{aligned} \tag{T9.149}$$

Using (A.100) this becomes:

$$\begin{aligned} a &= (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')\, (\boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_1)\, (\boldsymbol{\omega}' \otimes \boldsymbol{\omega}')' \\ &= (\boldsymbol{\omega}' \boldsymbol{\Sigma}_1 \boldsymbol{\omega}) \otimes (\boldsymbol{\omega}' \boldsymbol{\Sigma}_1 \boldsymbol{\omega}) \\ &= (\boldsymbol{\omega}' \boldsymbol{\Sigma}_1 \boldsymbol{\omega})^2 \end{aligned} \tag{T9.150}$$

Therefore (T9.145) yields:

$$\begin{aligned} \max_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega} &= \frac{\nu_1}{\nu_1 + N + 1} \boldsymbol{\omega}' \boldsymbol{\Sigma}_1 \boldsymbol{\omega} \\ &\quad + \left(\frac{2 \nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3} \right)^{1/2} (\boldsymbol{\omega}' \boldsymbol{\Sigma}_1 \boldsymbol{\omega}) \\ &= \left[\frac{\nu_1}{\nu_1 + N + 1} + \left(\frac{2 \nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3} \right)^{1/2} \right] (\boldsymbol{\omega}' \boldsymbol{\Sigma}_1 \boldsymbol{\omega})\,. \end{aligned} \tag{T9.151}$$

Equivalently, recalling (T9.127) we can write (T9.145) as follows:

$$\max_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \{ \boldsymbol{\omega}' \boldsymbol{\Sigma} \boldsymbol{\omega} \} = \left[\frac{\nu_1}{\nu_1 + N + 1} + \left(\frac{2 \nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1 + N + 1)^3} \right)^{1/2} \right] \left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega} \right\|^2 \tag{T9.152}$$

The mean-variance problem

Substituting (T9.132) and (T9.152) in (9.139), and defining

$$\gamma_{\mu} \equiv \left(\frac{1}{T_1} \frac{q_{\boldsymbol{\mu}}^2 \nu_1}{\nu_1 - 2} \right)^{1/2} \tag{T9.153}$$

$$\gamma_{\Sigma}^{(i)} \equiv \frac{v^{(i)}}{\frac{\nu_1}{\nu_1+N+1} + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{(\nu_1+N+1)^3} \right)^{1/2}}, \tag{T9.154}$$

the Bayesian robust mean-variance problem reads:

$$\begin{aligned} \boldsymbol{\omega}_{\mathrm{Br}}^{(i)} = \underset{\boldsymbol{\omega}}{\operatorname{argmax}} \left\{ \boldsymbol{\omega}' \boldsymbol{\mu}_1 - \gamma_{\mu} \left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega} \right\| \right\} \\ \text{s.t.} \ \left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega} \right\| \leq \sqrt{\gamma_{\Sigma}^{(i)}}. \end{aligned} \tag{T9.155}$$

This is equivalent to:

$$\left(\boldsymbol{\omega}_{\mathrm{Br}}^{(i)}, z^* \right) = \underset{\boldsymbol{\omega} \in \mathcal{C}, z}{\operatorname{argmax}} \left\{ \boldsymbol{\omega}' \boldsymbol{\mu}_1 - z \right\} \tag{T9.156}$$

subject to

$$\left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega} \right\| \leq z / \gamma_{e\mu} \tag{T9.157}$$

$$\left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \boldsymbol{\omega} \right\| \leq \sqrt{\gamma_{\Sigma}^{(i)}}. \tag{T9.158}$$

This problem is in the SOCP form (6.55).