9

Optimizing allocations

The classical approach to allocation optimization discussed in the second part of the book assumes that the distribution of the market is known. The sample-based allocation, discussed in the previous chapter, is a two-step process: first the market distribution is estimated and then the estimate is inputted in the classical allocation optimization problem. Since this process leverages the estimation error, portfolio managers, traders and professional investors in a broader sense mistrust these two-step "optimal" approaches and prefer to resort to ad-hoc recipes, or trust their prior knowledge/experience.

In this chapter we discuss allocation strategies that account for estimation risk within the allocation decision process. These strategies must be optimal according to the evaluation criteria introduced in the previous chapter: in other words, the overall opportunity cost of these strategies must be as low as possible.

The main reasons why estimation risk plays such an important role in financial applications is the extreme sensitivity of the optimal allocation function to the unknown parameters that determine the distribution of the market. In Section 9.1 we use the Bayesian approach to estimation to limit this sensitivity. We present Bayesian allocations in terms of the predictive distribution of the market, as well as the classical-equivalent Bayesian allocation, which relies on Bayes-Stein shrinkage estimators of the market parameters. The Bayesian approach provides a mechanism that mixes the positive features of the prior allocation and the sample-based allocation: the estimate of the market is shrunk towards the investor's prior in a self-adjusting way and the overall opportunity cost is reduced.

In Section 9.2 we present the Black-Litterman approach to control the extreme sensitivity of the optimal allocation function to the unknown market parameters. Like the Bayesian approach, the Black-Litterman methodology makes use of Bayes' rule. In this case the market is directly shrunk towards the investor's prior views, rather than indirectly through the market parameters. We present the theory in a general context, performing the computations explicitly in the case of normally distributed markets. Then we apply those results to the mean-variance framework. Finally we propose a methodology to assess and tweak the investor's prior views.

In Section 9.3 we present Michaud's resampling technique. The rationale behind this approach consists in limiting the extreme sensitivity of the optimal allocation function to the market parameters by averaging several sample-based allocations in different scenarios. After presenting the resampled allocation in both the mean-variance and in a more general setting, we discuss the advantages and the limitations of this technique.

In Section 9.4 we discuss robust allocation decisions. Rather than trying to limit the sensitivity of the optimal allocation function, the robust approach aims at determining the "best" allocation in the presence of estimation risk, according to the evaluation criteria discussed in Chapter 8. In other words, robust allocations minimize the opportunity cost over a reasonable set of potential markets. The conceptually intuitive robust approach is hard to implement in the general case. Therefore, we resort to the two-step mean-variance framework: under suitable assumptions for the investment constraints the optimal allocations solve a second-order cone programming problem: as a result, the optimal allocations can be efficiently determined numerically.

In Section 9.5 we blend the optimality properties of the robust approach with the smoothness and self-adjusting nature of the Bayesian approach. Indeed, the robust approach presents only two disadvantages: the possible markets considered in the robust optimization are defined quite arbitrarily and the investor's prior views are not taken into account. By means of the Bayesian posterior we can select naturally a notable set of markets and smoothly blend the investor's experience with the information from the market. We present first the robust Bayesian method in a general context, showing how this approach includes the previous allocation strategies as limit cases. Then we apply the general theory to the two-step mean-variance framework, discussing the self-adjusting mechanism of robust Bayesian allocations strategies.

9.1 Bayesian allocation

Consider the optimal allocation function (8.30), which for each value of the market parameters $\boldsymbol{\theta}$ maximizes the investor's satisfaction given his investment constraints:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\theta}}}{\operatorname{argmax}} \left\{ \mathcal{S}_{\boldsymbol{\theta}}(\boldsymbol{\alpha}) \right\}. \tag{9.1}$$

Since the true value $\boldsymbol{\theta}^{\mathrm{t}}$ of the market parameters is not known, the truly optimal allocation cannot be implemented. Furthermore, as discussed in Chapter 8, the allocation function (9.1) is extremely sensitive to the input parameters $\boldsymbol{\theta}$: a slightly wrong input can give rise to a very large opportunity cost.

In this section we use the Bayesian approach to parameter estimation to define allocation decisions whose opportunity cost is not as large.

9.1.1 Utility maximization

Expected utility has been historically the first and most prominent approach to model the investor's preferences. Therefore Bayesian theory was first applied to allocation problems in the context of expected utility maximization, see Zellner and Chetty (1965), and Bawa, Brown, and Klein (1979).

We recall from Section 5.4 that in the expected utility framework the investor's index of satisfaction is modeled by the certainty-equivalent ensuing from an increasing utility function u:

$$\mathcal{S}(\boldsymbol{\alpha}) \equiv u^{-1}\left(\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}\right)\right\}\right). \tag{9.2}$$

In this expression the investor's objective Ψ, namely absolute wealth, relative wealth, net profits, or other specifications, is a linear function of the allocation and the market vector: $\Psi \equiv \boldsymbol{\alpha}'\mathbf{M}$. The market vector $\mathbf{M}$ is a simple affine function of the market prices at the investment horizon: its distribution can be represented in terms of a probability density function $f_{\boldsymbol{\theta}}(\mathbf{m})$ which is fully determined by a set of market parameters $\boldsymbol{\theta}$.

Due to (5.99), in this context the optimal allocation function (9.1) can be expressed equivalently as follows:

$$\begin{aligned}\boldsymbol{\alpha}(\boldsymbol{\theta}) &\equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\theta}}}{\operatorname{argmax}}\left\{\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}^{\boldsymbol{\theta}}\right)\right\}\right\} \\ &= \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\theta}}}{\operatorname{argmax}}\left\{\int u\left(\boldsymbol{\alpha}'\mathbf{m}\right) f_{\boldsymbol{\theta}}(\mathbf{m})\, d\mathbf{m}\right\}.\end{aligned} \tag{9.3}$$

Consider an investor with exponential utility function. His expected utility reads:

$$\begin{aligned}\mathrm{E}\left\{u\left(\Psi_{\boldsymbol{\alpha}}^{\boldsymbol{\theta}}\right)\right\} &= -\,\mathrm{E}\left\{e^{-\frac{1}{\zeta}\boldsymbol{\alpha}'\mathbf{M}}\right\} \\ &= -\int e^{-\frac{1}{\zeta}\boldsymbol{\alpha}'\mathbf{m}} f_{\boldsymbol{\theta}}(\mathbf{m})\, d\mathbf{m} \equiv -\phi_{\boldsymbol{\theta}}\left(\frac{i}{\zeta}\boldsymbol{\alpha}\right),\end{aligned} \tag{9.4}$$

where $\phi_{\boldsymbol{\theta}}$ denotes the characteristic function of the market vector. Assume that the market is normally distributed. From (2.157) the characteristic function reads:

$$\phi_{\boldsymbol{\xi},\boldsymbol{\Phi}}(\mathbf{x}) = e^{i\boldsymbol{\xi}'\mathbf{x}-\frac{1}{2}\mathbf{x}'\boldsymbol{\Phi}\mathbf{x}}. \tag{9.5}$$

Then the allocation optimization (9.3) becomes:

$$\boldsymbol{\alpha}(\boldsymbol{\xi},\boldsymbol{\Phi}) \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\xi},\boldsymbol{\Phi}}}{\operatorname{argmax}}\left\{-e^{-\frac{1}{\zeta}\left(\boldsymbol{\xi}'\boldsymbol{\alpha}-\frac{1}{2\zeta}\boldsymbol{\alpha}'\boldsymbol{\Phi}\boldsymbol{\alpha}\right)}\right\}. \tag{9.6}$$

This problem is clearly equivalent to the maximization of the certainty equivalent (8.4).

The optimal allocation function (9.3) is extremely sensitive to the unknown market parameters $\boldsymbol{\theta}$.

On the other hand, in the Bayesian framework the unknown parameters $\boldsymbol{\theta}$ are a random variable whose possible outcomes are described by the posterior probability density function $f_{\text{po}}(\boldsymbol{\theta})$. Assume that the investment constraints in the allocation function (9.3) do not depend on the unknown parameters $\boldsymbol{\theta}$. In order to smoothen the sensitivity of the allocation function to the parameters it is quite natural to consider the weighted average of the argument of the optimization (9.3) over all the possible outcomes of the market parameters:

$$\overline{\boldsymbol{\alpha}} \equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}}{\operatorname{argmax}} \left\{ \int \mathrm{E}\left\{ u\left(\Psi_{\boldsymbol{\alpha}}^{\boldsymbol{\theta}}\right)\right\} f_{\text{po}}(\boldsymbol{\theta})\, d\boldsymbol{\theta} \right\}. \tag{9.7}$$

The posterior distribution of the parameters depends on both the information on the market i_T and the investor's experience e_C, see (7.15). Consider the *predictive distribution* of the market, which is defined in terms of the posterior distribution of the parameters as follows:

$$f_{\text{prd}}(\mathbf{m}; i_T, e_C) \equiv \int f_{\boldsymbol{\theta}}(\mathbf{m})\, f_{\text{po}}(\boldsymbol{\theta}; i_T, e_C)\, d\boldsymbol{\theta}. \tag{9.8}$$

This expression is indeed a probability density function, i.e. it satisfies (2.5) and (2.6). Like the posterior distribution of the parameters, also the predictive distribution of the market depends on both information and experience: it describes the statistical features of the market vector $\mathbf{M}$, keeping into account that the value of $\boldsymbol{\theta}$ is not known with certainty, i.e., accounting for estimation risk.

Using the definition of the predictive density in the average allocation (9.7) and exchanging the order of integration it is immediate to check that the average allocation can be written as follows:

$$\begin{aligned} \boldsymbol{\alpha}_{\mathrm{B}}[i_T, e_C] &= \underset{\boldsymbol{\alpha} \in \mathcal{C}}{\operatorname{argmax}} \left\{ \int u(\boldsymbol{\alpha}'\mathbf{m})\, f_{\text{prd}}(\mathbf{m}; i_T, e_C)\, d\mathbf{m} \right\} \\ &\equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}}{\operatorname{argmax}} \left\{ \mathrm{E}\left\{ u\left(\Psi_{\boldsymbol{\alpha}}^{i_T, e_C}\right)\right\}\right\}. \end{aligned} \tag{9.9}$$

This is the *Bayesian allocation decision,* which maximizes the expected utility of the investor' objective, where the expectation is computed according to the predictive distribution of the market. In other words, the Bayesian allocation decision is the standard Von Neumann-Morgenstern optimal allocation where instead of the unknown market distribution we use its predictive distribution.

Since the predictive distribution accounts for estimation risk and includes the investor's experience, so does the Bayesian allocation decision.

Assume that in our example (9.5) the covariance $\boldsymbol{\Phi}$ is known, and that the posterior distribution of the expected value is normal:

$$\boldsymbol{\xi} \sim \mathrm{N}\left(\boldsymbol{\xi}_1[i_T, e_C], \frac{\boldsymbol{\Phi}}{T_1}\right). \tag{9.10}$$

When $\mathbf{\Phi}$ is known, this specification is consistent with the posterior (7.32).

We show in Appendix www.9.7 that the predictive distribution of the normal market (9.5) with the normal posterior for the parameters (9.10) is also normal:

$$\phi_{\text{prd}}\left(\mathbf{x}; i_T, e_C\right) = e^{i\mathbf{x}'\boldsymbol{\xi}_1[i_T,e_C]-\frac{1}{2}\mathbf{x}'\frac{1+T_1}{T_1}\mathbf{\Phi}\mathbf{x}}. \tag{9.11}$$

Therefore, from (9.4) the Bayesian allocation decision reads:

$$\boldsymbol{\alpha}_{\text{B}} \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}}{\operatorname{argmax}} \left\{ -e^{-\frac{1}{\zeta}\left(\boldsymbol{\alpha}'\boldsymbol{\xi}_1-\frac{1+T_1}{2\zeta T_1}\boldsymbol{\alpha}'\mathbf{\Phi}\boldsymbol{\alpha}\right)} \right\}. \tag{9.12}$$

Allocation decisions based on the predictive distribution continue to find applications in finance, see for instance Jorion (1986). See also Pastor (2000) and Pastor and Stambaugh (2002) for applications based on explicit factor models.

9.1.2 Classical-equivalent maximization

Consider the more general case where the investment constraints in the optimal allocation function (9.1) depend on the unknown parameters $\boldsymbol{\theta}$, or the investor's satisfaction cannot be modeled by the certainty-equivalent. Then the Bayesian allocation (9.9) is not a viable option.

To generalize the Bayesian approach to this context, instead of averaging the distribution of the market by means of the predictive distribution (9.8) we average the distribution of the market parameters that feed the optimal allocation function. In other words, we replace the true unknown market parameters in (9.1) with a classical-equivalent estimator $\widehat{\boldsymbol{\theta}}_{\text{ce}}$, such as the expected value of the posterior distribution (7.5) or the mode of the posterior distribution (7.6). This way we obtain the *classical-equivalent Bayesian allocation decision*:

$$\begin{aligned} \boldsymbol{\alpha}_{\text{ce}}\left[i_T, e_C\right] &\equiv \boldsymbol{\alpha}\left(\widehat{\boldsymbol{\theta}}_{\text{ce}}\left[i_T, e_C\right]\right) \\ &\equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\widehat{\boldsymbol{\theta}}_{\text{ce}}[i_T,e_C]}}{\operatorname{argmax}} \left\{\mathcal{S}_{\widehat{\boldsymbol{\theta}}_{\text{ce}}[i_T,e_C]}\left(\boldsymbol{\alpha}\right)\right\}. \end{aligned} \tag{9.13}$$

This allocation decision depends through the classical-equivalent estimate on both the market information available i_T and the investor's experience e_C.

Consider the leading example (8.18), where we assumed that the market consists of equity-like securities for which the linear returns are market invariants:

$$\mathbf{L}_t \equiv \operatorname{diag}\left(\mathbf{P}_{t-\tau}\right)^{-1}\mathbf{P}_t - \mathbf{1}. \tag{9.14}$$

We assume as in (8.19) that the linear returns are normally distributed:

$$\mathbf{L}_t|\boldsymbol{\mu},\boldsymbol{\Sigma} \sim \mathrm{N}(\boldsymbol{\mu},\boldsymbol{\Sigma}). \tag{9.15}$$

This is the multivariate normal Bayesian model (7.16).

The available information on the market is represented by the time series of the past linear returns, see (8.41). As in (7.19) this information can be summarized by the sample mean of the observed linear returns (8.79), their sample covariance (8.80) and the length of the time series:

$$i_T \equiv \left\{\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}; T\right\}. \tag{9.16}$$

As in (7.27) the investor's experience is summarized by the following parameters:

$$e_C \equiv \{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0; T_0, \nu_0\}. \tag{9.17}$$

As in (7.20)-(7.21) the investor's experience is modeled as a normal-inverse-Wishart distribution:

$$\boldsymbol{\mu}|\boldsymbol{\Sigma} \sim \mathrm{N}\left(\boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}}{T_0}\right), \quad \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_0, \frac{\boldsymbol{\Sigma}_0^{-1}}{\nu_0}\right). \tag{9.18}$$

The classical-equivalent estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are (7.35) and (7.38), which we report here:

$$\widehat{\boldsymbol{\mu}}_{\mathrm{ce}}(i_T, e_C) = \frac{T_0\boldsymbol{\mu}_0 + T\widehat{\boldsymbol{\mu}}}{T_0 + T}, \tag{9.19}$$

$$\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}(i_T, e_C) = \frac{1}{\nu_0 + T + N + 1}\left[\nu_0\boldsymbol{\Sigma}_0 + T\widehat{\boldsymbol{\Sigma}}\right. \tag{9.20}$$
$$\left. + \frac{(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}})(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}})'}{\frac{1}{T} + \frac{1}{T_0}}\right].$$

In our leading example the optimal allocation function is (8.32). Substituting the classical-equivalent estimators into the functional expression of the optimal allocation function we obtain the classical-equivalent Bayesian allocation:

$$\boldsymbol{\alpha}_{\mathrm{ce}} \equiv [\mathrm{diag}(\mathbf{p}_T)]^{-1}\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}^{-1}\left(\zeta\widehat{\boldsymbol{\mu}}_{\mathrm{ce}} + \frac{w_T - \zeta\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}^{-1}\widehat{\boldsymbol{\mu}}_{\mathrm{ce}}}{\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}^{-1}\mathbf{1}}\mathbf{1}\right). \tag{9.21}$$

9.1.3 Evaluation

To evaluate the classical-equivalent Bayesian allocation we proceed as in Chapter 8, computing the distribution of the opportunity cost as the underlying market parameters $\boldsymbol{\theta}$ vary in a suitable stress test range $\boldsymbol{\Theta}$, which in this case is naturally defined as the domain of the posterior distribution.

Therefore, for each value $\boldsymbol{\theta}$ of the market parameters in the domain $\boldsymbol{\Theta}$ of the posterior distribution we compute the optimal allocation function $\boldsymbol{\alpha}(\boldsymbol{\theta})$ as defined in (9.1). Then we compute as in (8.31) the optimal level of satisfaction if $\boldsymbol{\theta}$ are the underlying market parameters, namely $\overline{\mathcal{S}}(\boldsymbol{\theta})$.

In our leading example the optimal allocation function is (8.32) and the respective optimal level of satisfaction is (8.33).

Next, for each value $\boldsymbol{\theta}$ of the market parameters in the stress test set $\boldsymbol{\Theta}$ we randomize as in (8.48) the information from the market i_T, generating a distribution of information scenarios $I_T^{\boldsymbol{\theta}}$ that depends on the assumption $\boldsymbol{\theta}$ on the market parameters. This way the classical-equivalent estimator becomes a random variable:

$$\widehat{\boldsymbol{\theta}}_{\mathrm{ce}}[i_T, e_C] \mapsto \widehat{\boldsymbol{\theta}}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]. \tag{9.22}$$

We stress that the distribution of this random variable is determined by the underlying assumption $\boldsymbol{\theta}$ on market parameters.

In our example, we replace i_T, i.e. the specific observations of the past linear returns, with a set $I_T^{\boldsymbol{\mu},\boldsymbol{\Sigma}}$ of T independent and identically distributed variables (9.15). This way the sample mean and the sample covariance become random variables distributed according to (8.85) and (8.86) respectively. As a result, the classical-equivalent estimators (9.19) and (9.20) become random variables, whose distribution can be simulated by a large number J of Monte Carlo scenarios as in (8.88):

$$_j\widehat{\boldsymbol{\mu}}_{\mathrm{ce}}^{\boldsymbol{\mu},\boldsymbol{\Sigma}}, \quad _j\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}^{\boldsymbol{\mu},\boldsymbol{\Sigma}}, \qquad j = 1, \ldots, J. \tag{9.23}$$

Notice that this distribution depends on the assumption $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ on the market parameters.

In turn, the classical-equivalent Bayesian allocation decision (9.13) yields a random variable whose distribution depends on the underlying market parameters:

$$\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right] \equiv \boldsymbol{\alpha}\left(\widehat{\boldsymbol{\theta}}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right). \tag{9.24}$$

In our example we substitute (9.23) in (9.21), obtaining J allocations $_j\boldsymbol{\alpha}_{\mathrm{ce}}^{\boldsymbol{\mu},\boldsymbol{\Sigma}}$.

Next we compute as in (8.23) the satisfaction $\mathcal{S}_{\boldsymbol{\theta}}\left(\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right)$ ensuing from each scenario of the classical-equivalent Bayesian allocation decision (9.24) under the assumption $\boldsymbol{\theta}$ for the market parameters, which, we recall, is a random variable. Similarly, from (8.26) and expressions such as (8.35) we compute the cost of the classical-equivalent Bayesian allocation decision violating the constraints $\mathcal{C}_{\boldsymbol{\theta}}^{+}\left(\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right)$ in each scenario, which is also a random variable.

In our example we proceed as in (8.90)-(8.91).

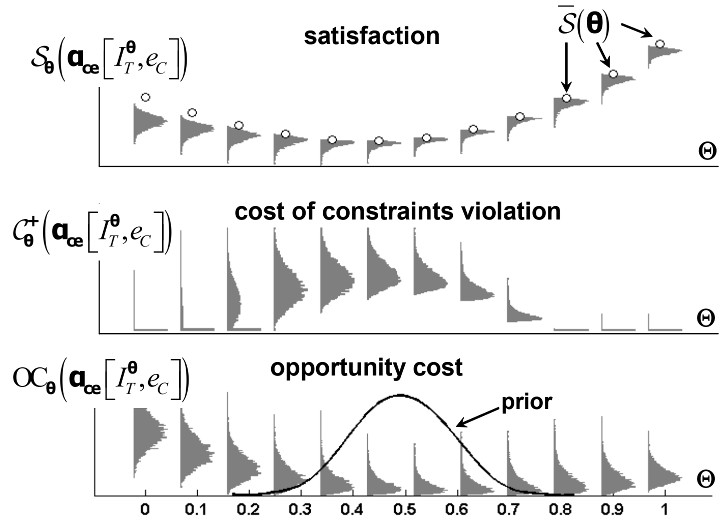

Fig. 9.1. Bayesian classical-equivalent allocation: evaluation

Then we compute the opportunity cost (8.53) of the classical-equivalent Bayesian allocation under the assumption $\boldsymbol{\theta}$ for the market parameters, which is the difference between the satisfaction from the unattainable optimal allocation and the satisfaction from the classical-equivalent Bayesian allocation, plus the cost of the classical-equivalent Bayesian allocation violating the constraints:

$$\begin{aligned} \mathrm{OC}_{\boldsymbol{\theta}}\left(\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right) \equiv \overline{\mathcal{S}}\left(\boldsymbol{\theta}\right) - \mathcal{S}_{\boldsymbol{\theta}}\left(\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right) \\ + \mathcal{C}_{\boldsymbol{\theta}}^{+}\left(\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right). \end{aligned} \tag{9.25}$$

Finally, as in (8.57) we let the market parameters $\boldsymbol{\theta}$ vary in the stress test range $\boldsymbol{\Theta}$, analyzing the opportunity cost of the classical-equivalent Bayesian allocation as a function of the underlying market parameters:

$$\boldsymbol{\theta} \mapsto \mathrm{OC}_{\boldsymbol{\theta}}\left(\boldsymbol{\alpha}_{\mathrm{ce}}\left[I_T^{\boldsymbol{\theta}}, e_C\right]\right). \tag{9.26}$$

If the distribution of the opportunity cost (9.26) is tightly peaked around a positive value very close to zero for all the markets $\boldsymbol{\theta}$ in the stress test range $\boldsymbol{\Theta}$, in particular it is close to zero in all the scenarios in correspondence of the true, yet unknown, value $\boldsymbol{\theta}^{\mathrm{t}}$. In this case the classical-equivalent Bayesian allocation decision is guaranteed to perform well and is close to optimal.

In our example we proceed as in (8.94)-(8.97), see Figure 9.1 and compare with Figure 8.4. Refer to symmys.com for more details on these plots.

9.1.4 Discussion

As discussed in Section 7.1.2, due to (7.4) the classical-equivalent estimator is a shrinkage estimator of the market parameters. Indeed it is a Bayes-Stein shrinkage estimator, where the shrinkage target is represented by the investor's prior experience $\boldsymbol{\theta}_0$. When the information available in the market is much larger than the investor's confidence in his experience, i.e. $T \gg C$, the classical-equivalent estimator converges to the sample estimate $\widehat{\boldsymbol{\theta}}$. On the other hand, when the investor's confidence in his experience is much larger than the information from the market, i.e. $C \gg T$, the classical-equivalent estimator shrinks to the prior $\boldsymbol{\theta}_0$.

Therefore, when $T \gg C$, the classical-equivalent Bayesian allocation (9.13) tends to the sample-based allocation (8.81). On the other hand, when $C \gg T$, the classical-equivalent Bayesian allocation tends to the prior allocation (8.64) which is fully determined by the prior parameters inputted by the investor and completely disregards the information from the market.

In the general case, the classical-equivalent Bayesian allocation is a blend of the sample-based allocation and the allocation determined by the prior. In other words, the classical-equivalent Bayesian allocation strategy can be interpreted as a "shrinkage" of the sample-based allocation towards the investor's prior/experience, where the amount of shrinkage is adjusted naturally by the relation between the amount information T and the confidence level C.

We recall from (8.53) that the opportunity cost of an allocation decision can be interpreted as the loss of an estimator. The same way as shrinkage estimators are a little more biased but less inefficient than sample estimators and thus display a lower error, so classical-equivalent Bayesian allocations generate opportunity costs that are less scattered than in the case of the sample-based strategy, at least for those values of the market parameters close to the prior assumption.

We see this in Figure 9.1, which refers to the classical-equivalent Bayesian allocation (9.21). Compare this figure with the evaluation of the prior allocation in Figure 8.3 and with the evaluation of the sample-based allocation in Figure 8.4.

The market parameters vary as in (8.58)-(8.59), i.e. the market is determined by the overall level of correlation. We plot the distribution of the prior overall correlation as implied by (9.18), which we compute by means of simulations.

Since the Bayesian estimate includes the investor's experience, the classical-equivalent Bayesian allocation automatically yields better results when the stress test (9.26) is run in the neighborhood of the prior assumptions on the market parameters, although coincidentally the cost of constraints violation is larger in the same region.

9.2 Black-Litterman allocation

Consider the optimal allocation function (8.30), which for each value of the market parameters $\boldsymbol{\theta}$ maximizes the investor's satisfaction given his investment constraints:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\theta}}}{\operatorname{argmax}} \left\{ \mathcal{S}_{\boldsymbol{\theta}}(\boldsymbol{\alpha}) \right\}. \tag{9.27}$$

Since the true value $\boldsymbol{\theta}^{\mathrm{t}}$ of the market parameters is not known, the truly optimal allocation cannot be implemented. Furthermore, as discussed in Chapter 8, the allocation function (9.66) is extremely sensitive to the input parameters $\boldsymbol{\theta}$: a slightly wrong input can give rise to a very large opportunity cost.

Like the Bayesian approach, the approach to asset allocation of Black and Litterman (1990) applies Bayes' rule to limit the sensitivity of the optimal allocation function to the input parameters. Nevertheless, the Black-Litterman framework differs from the classical-equivalent approach in that in the classical-equivalent approach the estimates of the market parameters are shrunk toward the investor's prior, whereas in the Black-Litterman approach it is the market distribution that is shrunk toward the investor's prior[1].

We present first the theory for the general case, where the market is described by a generic distribution and the investor can express views on any function of the market. Then we detail the computations that lead to the Black-Litterman allocation decision for the case where the investor expresses views on linear combinations of a normally distributed market.

9.2.1 General definition

Consider a market represented by the multivariate random variable $\mathbf{X}$. This could be the set of market invariants, or directly the set of market prices at the investment horizon, or any other variable that directly or indirectly fully determines the market.

Assume that it is possible to determine the distribution of this random variable, as represented for instance by the probability density function $f_{\mathbf{X}}$, by means of a reliable model/estimation technique. We call this the "official" distribution of the market. For instance, we could estimate this distribution by one of the techniques discussed in Chapter 4, or by means of general equilibrium arguments.

[1] The interpretation in terms of shrinkage of market parameters is also possible, see He and Litterman (2002).

Consider for example the case where the market X is represented by the daily return on the S&P 500 index, and suppose that X is normally distributed:

$$X \sim \mathrm{N}\left(\mu, \sigma^2\right). \tag{9.28}$$

We represent this distribution on the horizontal axis in Figure 9.2.

The distribution $f_{\mathbf{X}}$ is affected by estimation risk. To smoothen the effect of estimation risk, the statistician asks the investor's opinion on the market. The opinion is the investor's view on the outcome of the market $\mathbf{X}$. The investor's opinion is not a one-shot statement: the investor must be an expert, must have built a track-record and will be asked an opinion on a regular basis.

When asked by the statistician, the investor assesses that the outcome of the market is $\mathbf{V}$, a random variable that, possibly depending on the market scenario, is larger or smaller than the value $\mathbf{X}$ predicted by the "official" model. In other words, when the variable $\mathbf{X}$ assumes a specific value $\mathbf{x}$, the investor believes that the real outcome differs from $\mathbf{x}$ by a random amount. Therefore, the view $\mathbf{V}$ is a perturbation of the "official" outcome, and as such it is expressed as a conditional distribution $\mathbf{V}|\mathbf{x}$. The choice of the model for this conditional distribution, as represented for instance by the probability density function $f_{\mathbf{V}|\mathbf{x}}$, reflects the statistician's confidence in the investor.

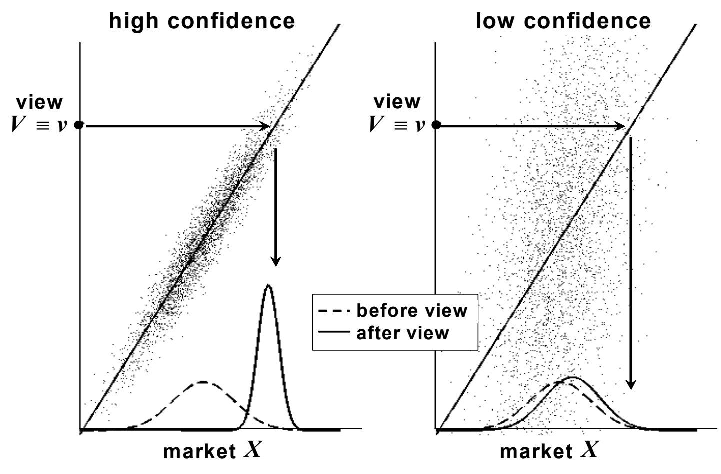

Fig. 9.2. Black-Litterman approach to market estimation

For example, the investor's opinion on the return of the S&P 500 index could be modeled as a normal perturbation to the "official" distribution:

$$V|x \sim \mathrm{N}\left(x, \phi^2\right). \tag{9.29}$$

If the statistician considers the investor unreliable, i.e. if he assumes that the investor's view will significantly depart from the "official" distribution (9.28) on a regular basis, he will choose a large value for the conditional standard deviation ϕ of the view. Viceversa, if the statistician trusts the investor he will model the view with a low value of ϕ.

In Figure 9.2 we see that when the confidence is high, the investor's statement is very close to the "official" distribution (a tight clouds of points). Viceversa, when the confidence is low, the cloud is very scattered.

More in general, the investor's opinion might regard a specific area of expertise of the market. In other words, instead of regarding directly the market $\mathbf{X}$, the view refers to a generic multivariate function $\mathbf{g}(\mathbf{X})$ on the market. Therefore the conditional model for the view becomes of the form $\mathbf{V}|\mathbf{x} \equiv \mathbf{V}|\mathbf{g}(\mathbf{x})$ and is represented for instance by the respective conditional probability density function $f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}$.

Once the model has been set up, the statistician will ask the investor's opinion. The investor will produce a specific number $\mathbf{v}$, namely his prediction on $\mathbf{V}$.

At this point the statistician processes the above inputs and computes the distribution of the market conditioned on the investor's opinion $\mathbf{X}|\mathbf{v}$. The representation of this distribution in terms of its probability density function follows from Bayes' rule (2.43), which in this context reads:

$$f_{\mathbf{X}|\mathbf{v}}(\mathbf{x}|\mathbf{v}) = \frac{f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})}{\int f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}(\mathbf{v}|\mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}}. \tag{9.30}$$

In our example the distribution of the market conditioned on the investor's view is normal:

$$X|v \sim \mathrm{N}\left(\widetilde{\mu}\left(v, \phi^2\right), \widetilde{\sigma}^2\left(\phi^2\right)\right). \tag{9.31}$$

This is a specific instance of the result (9.44), which we discuss below in a more general context. The parameters $(\widetilde{\mu}, \widetilde{\sigma})$ depend on the view v and the confidence in the view ϕ^2.

We see in Figure 9.2 that when the confidence is high the view has a large impact on the new distribution, which shrinks substantially towards the investor's statement. Indeed, when the cloud representing the joint distribution is tight, knowledge of one coordinate (the view) almost completely determines the other (the market). When the confidence is low, the market distribution is almost unaffected by the investor's statement.

To summarize, in order to include the investor's view in the "official" market model, we proceed as follows: we start from the "official" distribution of the market $f_{\mathbf{X}}$; then we determine the investor's area of expertise, i.e. a function $\mathbf{g}$ of the market; then we specify a model $f_{\mathbf{V}|\mathbf{g}(\mathbf{x})}$ for the conditional distribution of the investor's view given the market; then we record the investor's input, i.e. the specific value $\mathbf{v}$ of his view; finally we compute the conditional distribution (9.30) of the market given the investor's view.

At this point we can define the *Black-Litterman allocation decision* as the optimal allocation function (9.27) computed using the market (9.30) determined by the view:

$$\boldsymbol{\alpha}_{\mathrm{BL}}[\mathbf{v}] \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\mathbf{v}}}{\operatorname{argmax}} \{\mathcal{S}_{\mathbf{v}}(\boldsymbol{\alpha})\}. \tag{9.32}$$

Unlike in the other allocation strategies discussed in this chapter, the dependence of the Black-Litterman allocation on the contingent realization of the information i_T is not explicit.

Suppose that the market consists of the S&P500, whose return is X, and a risk-free security with null return. Assume that the investor has a budget w_T. Then an allocation is fully determined by the relative weight $\omega \equiv \alpha/w_T$ of the investment α in the risky security.

Assume that the investor's objective is final wealth, that his index of satisfaction is the expected value, and that he is bound by the no-short sale constraint. Then the Black-Litterman allocation reads:

$$\omega_{\mathrm{BL}}[v] \equiv \underset{0\leq\omega\leq 1}{\operatorname{argmax}} \{\omega\widetilde{\mu}\}, \tag{9.33}$$

where $\widetilde{\mu}$ is the expected value in (9.31).

9.2.2 Practicable definition: linear expertise on normal markets

Black and Litterman (1990) compute and discuss the analytical solution to (9.30) in a specific, yet quite general, case, see also Black and Litterman (1992).

First of all, the "official" model for the N-dimensional market vector $\mathbf{X}$ is assumed normal[1]:

$$\mathbf{X} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{9.34}$$

To illustrate, we consider an institution that adopts the RiskMetrics model to optimize the allocation of an international fund that invests in the following six stock indices: Italy, Spain, Switzerland, Canada, US and Germany. In this case the market are the daily compounded returns:

$$\mathbf{C} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{9.35}$$

Notice that this corresponds to the standard distributional assumption in Black and Scholes (1973).

[1] In the original paper the market is represented by the linear returns on a set of securities and the parameters $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ satisfy a general equilibrium model.

The expected value of the daily returns is assumed zero:

$$\boldsymbol{\mu} \equiv (0,0,0,0,0,0)' . \tag{9.36}$$

The covariance matrix of the daily returns on the above asset classes is estimated by exponential smoothing of the observed daily returns and is made publicly available by RiskMetrics. The matrix in our example was estimated in August 1999. Its decomposition in terms of standard deviations and correlations reads respectively:

$$\sqrt{\operatorname{diag}(\boldsymbol{\Sigma})} \equiv 0.01 \times (1.34, 1.52, 1.53, 1.55, 1.82, 1.97)' \tag{9.37}$$

and (we report only the non-trivial elements)

$$\operatorname{Cor}\{\mathbf{C}\} = \begin{pmatrix} \cdot & 54\% & 62\% & \mathbf{25}\% & 41\% & 59\% \\ \cdot & \cdot & 69\% & \mathbf{29}\% & 36\% & \mathbf{83}\% \\ \cdot & \cdot & \cdot & \mathbf{15}\% & 46\% & 65\% \\ \cdot & \cdot & \cdot & \cdot & \mathbf{47}\% & \mathbf{39}\% \\ \cdot & \cdot & \cdot & \cdot & \cdot & 38\% \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} . \tag{9.38}$$

Second, the investor's area of expertise is a linear function of the market:

$$\mathbf{g}(\mathbf{x}) \equiv \mathbf{P}\mathbf{x}, \tag{9.39}$$

where $\mathbf{P}$ is the "pick" matrix: each of its K rows is an N-dimensional vector that corresponds to one view and selects the linear combination of the market involved in that view.

The specification (9.39) is very flexible, in that the investor does not necessarily need to express views on all the market variables. Furthermore, views do not necessarily need to be expressed in absolute terms for each market variable considered, as any linear combination of the market constitutes a potential view.

A fund manager might assess absolute views on three markets: the Spanish, the Canadian and the German index. Therefore, the "pick" matrix reads:

$$\mathbf{P} \equiv \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} . \tag{9.40}$$

Notice from (9.38) that the Spanish and the German markets are highly correlated (83%) and that the Canadian index is relatively independent of the other markets.

Third, the conditional distribution of the investor's views given the outcome of the market is assumed normal:

$$\mathbf{V}|\mathbf{Px} \sim \mathrm{N}(\mathbf{Px}, \boldsymbol{\Omega}), \tag{9.41}$$

where the symmetric and positive matrix $\boldsymbol{\Omega}$ denotes the statistician's confidence in the investor's opinion.

A particularly convenient choice for the uncertainty matrix is

$$\boldsymbol{\Omega} \equiv \left(\frac{1}{c} - 1\right) \mathbf{P}\boldsymbol{\Sigma}\mathbf{P}', \tag{9.42}$$

where c is a positive scalar. This corresponds to an "empirical Bayesian" approach: the statistician gives relatively speaking more leeway to the investor's assessment on those combinations that are more volatile according to the official market model (9.34). The scalar c tweaks the absolute confidence in the investor's skills, see Figure 9.2. The case $c \to 0$ gives rise to an infinitely disperse distribution of the views: this means that the investor's views have no impact, i.e. the investor is not trusted. The case $c \to 1$ gives rise to an infinitely peaked distribution of the views: this means that the investor is trusted completely over the official market model. The case $c \equiv 1/2$ corresponds to the situation where the investor is trusted as much as the official market model.

In our example we define $\boldsymbol{\Omega}$ as in (9.42), where we set $c \equiv 1/2$.

Fourth, the investor is asked his opinion on his area of expertise. This will turn into a specific value $\mathbf{v}$ of the views $\mathbf{V}$.

The fund manager assesses that the Spanish index will remain unvaried, the Canadian stock index will score a negative return of 2% and the German index will experience a positive change of 2%. Therefore the views read:

$$\mathbf{v} \equiv 0.01 \times (0, -2, 2)'. \tag{9.43}$$

By means of Bayes' rule (9.30) it is possible to compute the distribution of the market conditioned on the investor's views. We show in Appendix www.9.3 that the *Black-Litterman distribution* is normal:

$$\mathbf{X}|\mathbf{v} \sim \mathrm{N}(\boldsymbol{\mu}_{\mathrm{BL}}, \boldsymbol{\Sigma}_{\mathrm{BL}}), \tag{9.44}$$

where the expected values read:

$$\boldsymbol{\mu}_{\mathrm{BL}}(\mathbf{v}, \boldsymbol{\Omega}) \equiv \boldsymbol{\mu} + \boldsymbol{\Sigma}\mathbf{P}' \left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega}\right)^{-1} (\mathbf{v} - \mathbf{P}\boldsymbol{\mu}); \tag{9.45}$$

and the covariance matrix reads:

$$\boldsymbol{\Sigma}_{\mathrm{BL}}(\boldsymbol{\Omega}) \equiv \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{P}' \left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega}\right)^{-1} \mathbf{P}\boldsymbol{\Sigma}. \tag{9.46}$$

Notice that the expression of the covariance is not affected by the value of the views $\mathbf{v}$. This is a peculiarity of the normal setting.

The expression of the Black-Litterman market distribution can be used to determine the optimal asset allocation that includes the investor's views.

In our example we consider an investor who has an initial budget w_T, and who is subject to the full-investment and the no-short-sale constraints:

$$\mathcal{C}: \ \boldsymbol{\alpha}'\mathbf{p}_T = w_T, \quad \boldsymbol{\alpha} \geq \mathbf{0}. \tag{9.47}$$

Furthermore, we assume that the investor's objective is final wealth:

$$\Psi_{\boldsymbol{\alpha}} \equiv \boldsymbol{\alpha}'\mathbf{P}_{T+\tau}. \tag{9.48}$$

In order to determine the optimal allocation we consider the two-step mean-variance framework. First we compute the efficient frontier(6.74), which in this context reads:

$$\begin{aligned} \boldsymbol{\alpha}(v) &\equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmax}}\, \boldsymbol{\alpha}' \operatorname{E}\{\mathbf{P}_{T+\tau}\} \\ &\text{subject to } \begin{cases} \boldsymbol{\alpha}'\mathbf{p}_T = w_T \\ \boldsymbol{\alpha} \geq \mathbf{0} \\ \boldsymbol{\alpha}' \operatorname{Cov}\{\mathbf{P}_{T+\tau}\}\, \boldsymbol{\alpha} = v. \end{cases} \end{aligned} \tag{9.49}$$

To compute the market inputs, namely $\operatorname{E}\{\mathbf{P}_{T+\tau}\}$ and $\operatorname{Cov}\{\mathbf{P}_{T+\tau}\}$, we need the characteristic function (2.157) of the Black-Litterman distribution (9.44) of the compounded returns:

$$\phi_{\mathbf{C}}(\boldsymbol{\omega}) = e^{i\boldsymbol{\mu}'_{\mathrm{BL}}\boldsymbol{\omega} - \frac{1}{2}\boldsymbol{\omega}'\boldsymbol{\Sigma}_{\mathrm{BL}}\boldsymbol{\omega}}. \tag{9.50}$$

Dropping "BL" from the notation, from (3.95) the expected values of the prices read:

$$\begin{aligned} \operatorname{E}\left\{P^{(n)}_{T+\tau}\right\} &= P^{(n)}_T \phi_{\mathbf{C}}\left(-i\boldsymbol{\delta}^{(n)}\right) \\ &= P^{(n)}_T e^{\left(\mu_n + \frac{\Sigma_{nn}}{2}\right)}. \end{aligned} \tag{9.51}$$

Similarly, from (3.96) we obtain the covariance matrix of the market:

$$\begin{aligned} \operatorname{Cov}\left\{P^{(m)}_{T+\tau}, P^{(n)}_{T+\tau}\right\} &= P^{(m)}_T P^{(n)}_T \phi_{\mathbf{C}}\left(-i\boldsymbol{\delta}^{(m)} - i\boldsymbol{\delta}^{(n)}\right) \\ &\quad - \operatorname{E}\left\{P^{(m)}_{T+\tau}\right\} \operatorname{E}\left\{P^{(n)}_{T+\tau}\right\} \\ &= P^{(m)}_T P^{(n)}_T e^{(\mu_m + \mu_n)} e^{\frac{1}{2}(\Sigma_{mm} + \Sigma_{nn})} \left(e^{\Sigma_{mn}} - 1\right). \end{aligned} \tag{9.52}$$

Formulas (9.51) and (9.52) yield the inputs of the mean-variance optimization as functions of the Black-Litterman parameters (9.45) and (9.46). Substituting these expressions in (9.49) we obtain for any level of variance v the respective efficient allocation that includes the investor's views, see Figure 9.4. In a second stage the investor chooses the efficient portfolio that best suits his profile, as in Figure 6.23.

9.2.3 Evaluation

The Black-Litterman approach can, but does not need to, rely on the contingent historical information i_T available when the investment decision is made. Indeed, this approach blends two models for the market, namely the investor's and the official models: these models can be based on historical information, or they can rely on prior information, or other rationales, such as general equilibrium arguments, etc. Therefore we cannot apply the approach discussed in Section 8.1 to the evaluation of the Black-Litterman allocation.

On the other hand, the expected value $\boldsymbol{\mu}_{\mathrm{BL}}$ tilted by the views $\mathbf{v}$ according to the Black-Litterman formula (9.45) might be in strong contrast with the value $\boldsymbol{\mu}$ that appears in the official market model (9.34). In this section we discuss a technique to measure this difference and tweak the most extreme views accordingly, see also Fusai and Meucci (2003). Notice that we only need to consider the tilted expected values, since the explicit value of the views $\mathbf{v}$ does not enter the expression for the covariance matrix (9.46).

First we recall the definition (1.35) of z-score, widely used by practitioners: the distance of a suspicious value x of the random variable X from the accepted expected value μ divided by the standard deviation σ of X. In a multivariate environment the z-score becomes the Mahalanobis distance (2.61).

Under the normal hypothesis (9.34) for the official market model, the square Mahalanobis distance of the market $\mathbf{X}$ from its expected value $\boldsymbol{\mu}$ through the metric induced by its covariance $\boldsymbol{\Sigma}$ is distributed as a chi-square with N degrees of freedom:

$$M^2 \equiv (\mathbf{X}-\boldsymbol{\mu})' \, \boldsymbol{\Sigma}^{-1} (\mathbf{X}-\boldsymbol{\mu}) \sim \chi^2_N, \tag{9.53}$$

see Appendix www.7.1.

In our context the "suspicious" value is the Black-Litterman vector of expected values $\boldsymbol{\mu}_{\mathrm{BL}}$. If we consider $\boldsymbol{\mu}_{\mathrm{BL}}$ as a realization of the random variable $\mathbf{X}$, we can compute the respective realization of the square Mahalanobis distance accordingly:

$$m^2_{\mathbf{v}} \equiv (\boldsymbol{\mu}_{\mathrm{BL}}(\mathbf{v})-\boldsymbol{\mu})' \, \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_{\mathrm{BL}}(\mathbf{v})-\boldsymbol{\mu}) \,. \tag{9.54}$$

Intuitively, if the square distance $m^2_{\mathbf{v}}$ is small, the views are not too far from the market model and the consistence of the Black-Litterman expectations with the market model is high. In turn, the realization $m^2_{\mathbf{v}}$ of the random variable M^2 can be considered small if M^2 is likely to be larger than $m^2_{\mathbf{v}}$.

Therefore, we define the index of consistence $C(\mathbf{v})$ of the Black-Litterman expectations with the market model as the probability that the random variable M^2 is larger than the realization $m^2_{\mathbf{v}}$:

$$C(\mathbf{v}) \equiv \mathbb{P}\left(M^2 \geq m^2_{\mathbf{v}}\right) = 1 - F^{\mathrm{Ga}}_{N,1}\left(m^2_{\mathbf{v}}\right). \tag{9.55}$$

In this expression $F^{\mathrm{Ga}}_{N,1}$ represents the cumulative density function of the chi-square distribution with N degrees of freedom, which is a special case of the gamma cumulative density function (1.111).

In the extreme case where the realization $m^2_{\mathbf{v}}$ is zero, i.e. when $\boldsymbol{\mu}_{\mathrm{BL}}$ coincide with the model value $\boldsymbol{\mu}$, the random variable M^2 is certainly larger than the realized value and thus the consistence of the Black-Litterman expectations with the market model is total, i.e. one. As the realized value $m^2_{\mathbf{v}}$ increases, i.e. as $\boldsymbol{\mu}_{\mathrm{BL}}$ drifts apart from the model value $\boldsymbol{\mu}$, the random variable M^2 becomes less and less likely to be larger than the observed value and the consistence of the Black-Litterman expectations with the market model decreases.

We remark that the consistence C of the Black-Litterman expectations with the market model plays a dual role with the statistician's confidence c in the investor that appears in (9.42). Indeed, when the confidence c in the investor is zero, the views are ignored and the Black-Litterman distribution becomes the market distribution. Therefore the Mahalanobis distance of the Black-Litterman model from the official market model becomes null and the consistence $C(\mathbf{v})$ of the Black-Litterman expectations with the market model is total. As the confidence c in the investor increases, so does the Mahalanobis distance of the Black-Litterman model from the official market model and thus the consistence C of the Black-Litterman expectations with the market decreases.

When the overall consistence (9.55) is below an agreed threshold, often a slight shift in only one of the views suffices to boost the consistence level. Therefore, another natural problem is how to detect the "boldest" views, and how to fix them accordingly. To solve this problem, we compute the sensitivity of the consistence index to the views. From the chain rule of calculus, this sensitivity reads:

$$\begin{aligned} \frac{\partial C(\mathbf{v})}{\partial \mathbf{v}} &= \frac{dC}{dm^2} \frac{\partial m^2}{\partial \boldsymbol{\mu}_{\mathrm{BL}}} \frac{\partial \boldsymbol{\mu}_{\mathrm{BL}}}{\partial \mathbf{v}} \\ &= -2 f^{\mathrm{Ga}}_{N;1}\left(m^2_{\mathbf{v}}\right) \left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega}\right)^{-1} \mathbf{P}\left(\boldsymbol{\mu}_{\mathrm{BL}} - \boldsymbol{\mu}\right). \end{aligned} \tag{9.56}$$

In this expression $f^{\mathrm{Ga}}_{N;1}$ is the probability density function of the chi-square distribution with N degrees of freedom, which is a special case of the gamma probability density function (1.110).

In order to tweak the views, the investor simply needs to compute (9.56) and find the entry with the largest absolute value. If that entry is positive (negative), the respective view must be increased (decreased) slightly.

To illustrate, we apply this recipe to our example. We start with the views (9.43), which we report here:

$$\mathbf{v} \equiv 0.01 \times (0, -2, 2)'. \tag{9.57}$$

The consistence index (9.55) and the consistence sensitivities (9.56) read respectively:

$$C = 93.8\%, \quad \frac{\partial C}{\partial \mathbf{v}} = (8.1, 5.6, -9.0)'. \tag{9.58}$$

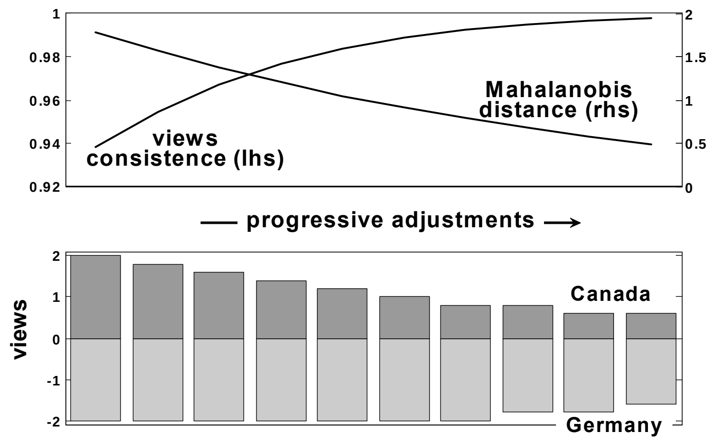

Fig. 9.3. Black-Litterman approach: views assessment

The consistence index is relatively insensitive to the second view on Canada, although it is of the same magnitude as the third view on Germany, namely 200 basis points. On the other hand the first view on Spain, which is apparently innocuous, has a larger effect on the consistence index: this is not unexpected, since the second view refers to a relatively independent market, whereas the first and third views state contrasting opinions on highly correlated markets, see (9.38).

Suppose that a consistence of at least 95% is required. To reach this level one should fine-tune, and actually decrease, the third view on the German index. It turns out that a 20 basis point shift, that changes (9.57) as follows

$$\mathbf{v} = 0.01 \times (0, -2, 1.8)' , \tag{9.59}$$

brings the overall consistence above the desired level:

$$C = 95.4\%. \tag{9.60}$$

In Figure 9.3 we see the effect on the consistence index of progressively reducing the boldness of the views: in the lower plot we display different views on the performance of Canada and Germany starting from the initial views +2% and -2% respectively; in the upper plot of the figure we report the progressively increasing consistence index (9.55) corresponding to less and less extreme views, along with the respective progressively decreasing square Mahalanobis distance (9.54) between the Black-Litterman expectations and the market expectations.

9.2.4 Discussion

The Black-Litterman approach might at first seem a little cumbersome. Why model the views as random variables conditioned on the market, when we could model them as deterministic functions of the market? In other words, instead of (9.41) we could more easily define the views as a function of the market $\mathbf{V} \equiv \mathbf{PX}$, and take the investor's input as a specific value $\mathbf{v}$ on which to condition the distribution of the market. This amounts to computing directly the conditional distribution of the market $\mathbf{X}|\mathbf{PX} \equiv \mathbf{v}$.

As we show in Appendix www.9.4 the conditional distribution of the market is normal:

$$\mathbf{X}|\mathbf{PX} \equiv \mathbf{v} \sim \mathrm{N}\left(\boldsymbol{\mu}_C, \boldsymbol{\Sigma}_C\right), \tag{9.61}$$

where the conditional expected values read:

$$\boldsymbol{\mu}_C \equiv \boldsymbol{\mu} + \boldsymbol{\Sigma}\mathbf{P}' \left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1} \left(\mathbf{v} - \mathbf{P}\boldsymbol{\mu}\right); \tag{9.62}$$

and the conditional covariance matrix reads:

$$\boldsymbol{\Sigma}_C \equiv \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{P}' \left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}'\right)^{-1} \mathbf{P}\boldsymbol{\Sigma}. \tag{9.63}$$

It is immediate to check that, as expected, this distribution is degenerate on the views:

$$\mathbf{PX}|\mathbf{PX} \equiv \mathbf{v} \sim \mathrm{N}\left(\mathbf{v}, \mathbf{0}\right). \tag{9.64}$$

Indeed, by definition of conditional distribution, the views $\mathbf{PX} = \mathbf{v}$ are supposed to take place with certainty. This is the reason why the direct conditional approach to modeling the views is not appropriate: the conditional approach yields a too "spiky" distribution. Therefore, since the allocation optimization process is very sensitive to the input parameters, when the optimal allocations are computed directly according to the conditional model, the resulting portfolios are extremely different from those computed according to the "official" market model and often give rise to corner solutions, see Figure 9.4.

Instead, the Black-Litterman distribution (9.44) blends smoothly the "official" market model (9.34) with the investor's blunt opinion, represented by the conditional distribution (9.61).

Indeed, the conditional distribution represents an extreme case of the Black-Litterman distribution, namely the case where the scatter matrix $\boldsymbol{\Omega}$ is null, i.e. the statistician's confidence in the investor's views is total. On the other hand, the "official" market model represents the opposite extreme case of the the Black-Litterman distribution, namely the case where the scatter matrix $\boldsymbol{\Omega}$ is infinite, i.e. the statistician's confidence in the investor is null:

$$\mathbf{X} \sim \mathrm{N}\left(\boldsymbol{\mu}_{\mathrm{BL}}, \boldsymbol{\Sigma}_{\mathrm{BL}}\right) \begin{array}{l} \nearrow \quad \mathbf{X} \sim \mathrm{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \qquad \left(\boldsymbol{\Omega} \to \infty\right) \\ \\ \searrow \quad \mathbf{X} \sim \mathrm{N}\left(\boldsymbol{\mu}_C, \boldsymbol{\Sigma}_C\right) \qquad \left(\boldsymbol{\Omega} \to \mathbf{0}\right). \end{array} \tag{9.65}$$

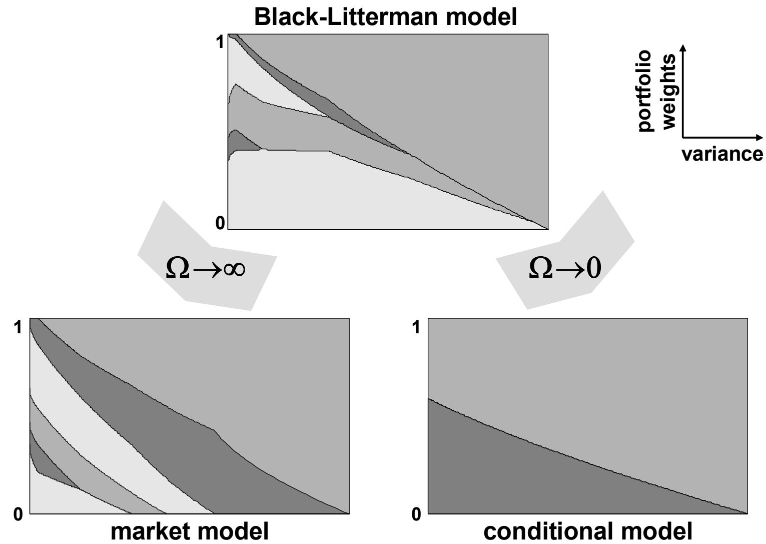

Fig. 9.4. Black-Litterman approach: sensitivity to the input parameters

For the intermediate cases, as the confidence in the investor's views decreases, the Black-Litterman distribution smoothly shifts away from the conditional model towards the "official" market model. This mechanism lessens the effect of the input parameters on the final allocations.

In Figure 9.4 we plot the efficient portfolios in terms of their relative weights computed according to the Black-Litterman distribution as in (9.49). We consider the general Black-Litterman distribution, as well as its limit cases, namely the "official" market model (9.34) and the distribution conditioned on the investor's views (9.61). Notice that the conditional distribution gives rise to corner solutions, i.e. highly concentrated portfolios.

9.3 Resampled allocation

Consider the optimal allocation function (8.30), which for each value of the market parameters $\boldsymbol{\theta}$ maximizes the investor's satisfaction given his investment constraints:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\theta}}}{\operatorname{argmax}} \{\mathcal{S}_{\boldsymbol{\theta}}(\boldsymbol{\alpha})\}. \tag{9.66}$$

Since the true value $\boldsymbol{\theta}^{\mathrm{t}}$ of the market parameters is not known, the truly optimal allocation cannot be implemented. Furthermore, as discussed in Chapter

8, the optimal allocation function is extremely sensitive to the input parameters $\boldsymbol{\theta}$: a slightly wrong input can give rise to a very large opportunity cost.

Unlike the Bayesian and the Black-Litterman approaches, where the above problem is tackled by smoothing the estimate of the input parameters before the optimization in (9.66), the resampling technique averages the outputs of a set of optimizations.

We present first the original resampled frontier of Michaud (1998), U.S. Patent No. 6,003,018, which refers to the mean-variance setting, see also Scherer (2002). Then we discuss its extension to generic markets and preferences.

9.3.1 Practicable definition: the mean-variance setting

We recall that the mean-variance approach is a two-step simplification of an allocation problem: the investor first determines a set of mean-variance efficient allocations and then selects among those allocations the one that better suits him.

The assumptions of the original resampling recipe are the following: first, the investor's objective admits the mean-variance formulations in terms of linear returns and relative weights, see Section 6.3.4; second, the market consists of equity-like securities for which the linear returns are market invariants, see Section 3.1.1; third, the investment horizon and the estimation interval coincide, see Section 6.5.4; fourth, the investment constraints are such that the dual formulation is correct, see Section 6.5.3; fifth, the constraints do not depend on unknown market parameters.

Under the above assumptions the mean-variance problem can be written as in (6.147), which in the dual formulation (6.146) reads:

$$\mathbf{w}^{(i)} = \underset{\substack{\mathbf{w}\in\mathcal{C} \\ \mathbf{w}'\boldsymbol{\mu}\geq e^{(i)}}}{\operatorname{argmin}}\, \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}, \quad i = 1,\ldots,I. \tag{9.67}$$

In this expression $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the expected values and the covariances of the linear returns of the securities relative to the investment horizon; the set $\left\{e^{(1)},\ldots,e^{(I)}\right\}$ is a significative grid of target expected values; and $\mathcal{C}$ is the set of investment constraints.

To determine the efficient portfolio weights (9.67) the resampling recipe follows these steps.

Step 1. Estimate the inputs ${}_0\widehat{\boldsymbol{\mu}}$ and ${}_0\widehat{\boldsymbol{\Sigma}}$ of the mean-variance framework from the analysis of the observed time series i_T of the past linear returns:

$$i_T \equiv \{\mathbf{l}_1,\ldots,\mathbf{l}_T\}\,. \tag{9.68}$$

This can be done for instance, but not necessarily, by means of the sample estimators (8.79) and (8.80).

Step 2a. Consider the time series i_T as the realization of a set of market invariants, i.e. independent and identically distributed returns:

$$I_T \equiv \{\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_T\}. \tag{9.69}$$

Step 2b. Make assumptions on the distribution generating the returns (9.69), for instance assuming normality, and set the estimated parameters as the true parameters that determine the distribution of the returns:

$$\mathbf{L}_t \sim \mathrm{N}\left({}_0\widehat{\boldsymbol{\mu}}, {}_0\widehat{\boldsymbol{\Sigma}}\right). \tag{9.70}$$

Step 2c. Resample a large number Q of Monte Carlo scenarios of realizations of (9.69) from the distribution (9.70):

$${}_q i_T \equiv \{{}_q\mathbf{l}_1, \ldots, {}_q\mathbf{l}_T\}, \quad q = 1, \ldots, Q. \tag{9.71}$$

Step 3. Estimate the inputs ${}_q\widehat{\boldsymbol{\mu}}$ and ${}_q\widehat{\boldsymbol{\Sigma}}$ of the mean-variance framework from the resampled time series (9.71) as in Step 1.

Step 4a. Compute the global minimum-variance portfolio from each of the resampled inputs:

$${}_q\mathbf{w}_{MV} = \underset{\mathbf{w} \in \mathcal{C}}{\operatorname{argmin}}\, \mathbf{w}'{}_q\widehat{\boldsymbol{\Sigma}}\mathbf{w}, \quad q = 1, \ldots, Q. \tag{9.72}$$

Step 4b. Compute the respective estimated expected value in each scenario:

$${}_q\underline{e} \equiv {}_q\mathbf{w}'_{MV}{}_q\widehat{\boldsymbol{\mu}}, \quad q = 1, \ldots, Q. \tag{9.73}$$

Step 4c. Compute the maximum estimated expected value in each scenario:

$${}_q\overline{e} \equiv \max\left\{{}_q\widehat{\boldsymbol{\mu}}'\boldsymbol{\delta}^{(1)}, \ldots, {}_q\widehat{\boldsymbol{\mu}}'\boldsymbol{\delta}^{(N)}\right\}, \quad q = 1, \ldots, Q, \tag{9.74}$$

where $\boldsymbol{\delta}$ is the canonical basis (A.15).

Step 4d. For each scenario q determine a grid $\left\{{}_q e^{(1)}, \ldots, {}_q e^{(I)}\right\}$ of equally-spaced target expected values as follows:

$$\begin{aligned} {}_q e^{(1)} &\equiv {}_q\underline{e} \\ &\vdots \\ {}_q e^{(i)} &\equiv {}_q\underline{e} + \frac{{}_q\overline{e} - {}_q\underline{e}}{I-1}(i-1) \\ &\vdots \\ {}_q e^{(I)} &\equiv {}_q\overline{e}. \end{aligned} \tag{9.75}$$

Step 4e. Solve the mean-variance dual problem (9.67) for all the Monte Carlo scenarios $q = 1, \ldots, Q$ and all the target expected values $i = 1, \ldots, I$:

$${}_q\mathbf{w}^{(i)} = \underset{\substack{\mathbf{w} \in \mathcal{C} \\ \mathbf{w}'{}_q\widehat{\boldsymbol{\mu}} \geq {}_q e^{(i)}}}{\operatorname{argmin}}\, \mathbf{w}'{}_q\widehat{\boldsymbol{\Sigma}}\mathbf{w}. \tag{9.76}$$

Step 5. Define the *resampled efficient frontier* as the average of the above allocations, possibly rejecting some outliers:

$$\mathbf{w}_{\text{rs}}^{(i)} \equiv \frac{1}{Q} \sum_{q=1}^{Q} {}_{q}\mathbf{w}^{(i)}, \quad i = 1, \ldots, I, \tag{9.77}$$

where "rs" stands for "resampled".

Step 6. Compute the efficient allocations from the respective relative weights:

$$\boldsymbol{\alpha}_{\text{rs}}^{(i)} \equiv w_T \operatorname{diag}(\mathbf{p}_T)^{-1} \mathbf{w}_{\text{rs}}^{(i)}, \quad i = 1, \ldots, I, \tag{9.78}$$

where w_T is the initial budget.

Following the steps 1-6 we obtain a set of allocations, namely (9.78), from which the investor can choose according to his preferences.

9.3.2 General definition

It is not difficult to generalize the rationale behind the resampled frontier to a more general setting, which does not necessarily rely on the two-step mean-variance approach. We modify the steps 1-6 that led to the resampled frontier respectively as follows.

Step 0. Instead of the expected values and the covariances of the linear returns, in general the market at the investment horizon is determined by a set of parameters $\boldsymbol{\theta}$, which steer the parametric distribution of the market invariants $\mathbf{X}_t^{\boldsymbol{\theta}}$, see (8.17).

Step 1. Using one of the techniques discussed in Chapter 4, estimate the parameters ${}_0\widehat{\boldsymbol{\theta}} \equiv \widehat{\boldsymbol{\theta}}[i_T]$ from the available time series of the market invariants:

$$i_T \equiv \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}. \tag{9.79}$$

We stress that the market invariants are not necessarily the linear returns: depending on the market, they could be for instance changes in yield to maturity, or other quantities, see Section 3.1.

Step 2. Generate a large number of Monte Carlo realizations of the time series ${}_q i_T$ of the market invariants, assuming that the distribution underlying the market invariants in Step 0 is determined by the estimated values. In other words, generate a large number Q of Monte Carlo realizations:

$${}_q i_T \equiv \{{}_q\mathbf{x}_1, \ldots, {}_q\mathbf{x}_T\}, \quad q = 1, \ldots, Q, \tag{9.80}$$

from the following set of random variables:

$$I_T^{{}_0\widehat{\boldsymbol{\theta}}} \equiv \left\{\mathbf{X}_1^{{}_0\widehat{\boldsymbol{\theta}}}, \ldots, \mathbf{X}_T^{{}_0\widehat{\boldsymbol{\theta}}}\right\}. \tag{9.81}$$

Step 3. In each scenario q estimate as in Step 1 the parameters ${}_q\widehat{\boldsymbol{\theta}} \equiv \widehat{\boldsymbol{\theta}}[{}_q i_T]$ from the resampled time series (9.80).

Step 4. Instead of determining the efficient frontier, in general the investor maximizes his primary index of satisfaction given his constraints, which depend on the market parameters, see (9.66). Therefore replace the optimization (9.76) with the following expression:

$$ {}_q\boldsymbol{\alpha} \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{{}_q\widehat{\boldsymbol{\theta}}}}{\operatorname{argmax}} \left\{ \mathcal{S}_{{}_q\widehat{\boldsymbol{\theta}}} \left(\boldsymbol{\alpha}\right) \right\}, \tag{9.82} $$

for all the Monte Carlo scenarios $q = 1, \ldots, Q$.

Step 5. Determine the resampled allocation by averaging the Monte Carlo optimal allocations:

$$ \boldsymbol{\alpha}_{\mathrm{rs}} \equiv \frac{1}{Q} \sum_{q=1}^{Q} {}_q\boldsymbol{\alpha}. \tag{9.83} $$

We stress that the resampled allocation is a decision $\boldsymbol{\alpha}_{\mathrm{rs}}\left[i_T\right]$, which depends on the available information (9.79) through the following chain, which summarizes the whole resampling technique:

$$ i_T \overset{\text{estimate}}{\longmapsto} {}_0\widehat{\boldsymbol{\theta}} \overset{\text{resample}}{\longmapsto} {}_q i_T \overset{\text{estimate}}{\longmapsto} {}_q\widehat{\boldsymbol{\theta}} \overset{\text{optimize}}{\longmapsto} {}_q\boldsymbol{\alpha} \overset{\text{average}}{\longmapsto} \boldsymbol{\alpha}_{\mathrm{rs}}. \tag{9.84} $$

We can further simplify the generic definition of the resampled allocation by avoiding the above sequential steps 1-5. Indeed the q-th scenario of the resampled allocation (9.82) is the optimal allocation function $\boldsymbol{\alpha}\left(\boldsymbol{\theta}\right)$ defined in (9.66) applied to the estimate from the q-th scenario of the Monte-Carlo-generated time series:

$$ {}_q\boldsymbol{\alpha} = \boldsymbol{\alpha}\left(\widehat{\boldsymbol{\theta}}\left[{}_q i_T\right]\right). \tag{9.85} $$

Furthermore, the q-th scenario of the time series ${}_q i_T$ is a realization of the random variable $I_T^{\widehat{\boldsymbol{\theta}}[i_T]}$, see (9.81). Therefore the average of the Monte Carlo scenarios (9.83) is the expectation of the allocations induced by the random variable $I_T^{\widehat{\boldsymbol{\theta}}[i_T]}$. In other words, the general definition of the *resampled allocation* can be summarized as follows:

$$ \boldsymbol{\alpha}_{\mathrm{rs}}\left[i_T\right] \equiv \mathrm{E}\left\{\boldsymbol{\alpha}\left(\widehat{\boldsymbol{\theta}}\left[I_T^{\widehat{\boldsymbol{\theta}}[i_T]}\right]\right)\right\}, \tag{9.86} $$

where "rs" stands for "resampled". This is indeed an allocation decision, which processes the currently available information, see (8.38).

In all the cases of practical interest, the resampled allocation cannot be computed in analytical closed form from the definition (9.86). Therefore, to implement the resampling technique we need to follow all the steps in (9.84).

Consider a random vector $\mathbf{u}$ distributed as follows:

$$ \mathbf{u} \sim \mathrm{N}\left(\widehat{\boldsymbol{\mu}}\left[i_T\right], \frac{1}{T}\widehat{\boldsymbol{\Sigma}}\left[i_T\right]\right), \tag{9.87} $$

where $\widehat{\boldsymbol{\mu}}\left[i_T\right]$ and $\widehat{\boldsymbol{\Sigma}}\left[i_T\right]$ are the sample mean and covariance of the linear returns (8.79) and (8.80) respectively. Now consider the positive and symmetric random matrix $\mathbf{V}$ distributed as follows:

$$T\mathbf{V}^{-1} \sim \mathrm{W}\left(T-1, \widehat{\boldsymbol{\Sigma}}\left[i_T\right]\right). \tag{9.88}$$

Furthermore, assume that $\mathbf{w}$ and $\mathbf{V}$ are independent. From (8.85) and (8.86) we obtain the distribution of the sample estimators applied to the time series distributed according to the estimated parameters:

$$\widehat{\boldsymbol{\mu}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right] \stackrel{d}{=} \mathbf{u}, \quad \widehat{\boldsymbol{\Sigma}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right] \stackrel{d}{=} \mathbf{V}^{-1}. \tag{9.89}$$

In our leading example the optimal allocation function is (8.32):

$$\boldsymbol{\alpha}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right) = \left[\operatorname{diag}\left(\mathbf{p}_T\right)\right]^{-1} \boldsymbol{\Sigma}^{-1}\left(\zeta\boldsymbol{\mu} + \frac{w_T - \zeta\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}\mathbf{1}\right). \tag{9.90}$$

Therefore the allocations induced by the random variable $I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}$ read:

$$\begin{aligned}\boldsymbol{\alpha}\left(\widehat{\boldsymbol{\mu}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right], \widehat{\boldsymbol{\Sigma}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right]\right) \stackrel{d}{=}\ & \zeta\left[\operatorname{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathbf{V}\mathbf{u} \\ & + \frac{w_T - \zeta\mathbf{1}'\mathbf{V}\mathbf{u}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\left[\operatorname{diag}\left(\mathbf{p}_T\right)\right]^{-1}\mathbf{V}\mathbf{1}.\end{aligned} \tag{9.91}$$

In turn the resampled allocation, which is the expected value of the above allocations, reads:

$$\begin{aligned}\boldsymbol{\alpha}_{\mathrm{rs}}\left[i_T\right] &\equiv \mathrm{E}\left\{\boldsymbol{\alpha}\left(\widehat{\boldsymbol{\mu}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right], \widehat{\boldsymbol{\Sigma}}\left[I_T^{\widehat{\boldsymbol{\mu}}[i_T],\widehat{\boldsymbol{\Sigma}}[i_T]}\right]\right)\right\} \\ &= \left[\operatorname{diag}\left(\mathbf{p}_T\right)\right]^{-1}\left(\zeta\,\mathrm{E}\left\{\mathbf{V}\widehat{\boldsymbol{\mu}} - \frac{\mathbf{1}'\mathbf{V}\widehat{\boldsymbol{\mu}}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\mathbf{V}\mathbf{1}\right\} + w_T\,\mathrm{E}\left\{\frac{\mathbf{V}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}\right\}\right),\end{aligned} \tag{9.92}$$

see Appendix www.9.1.

The expectations in (9.92) are not known in analytical form. Therefore we generate a large number Q of Monte Carlo scenarios from (9.88):

$${}_q\mathbf{V}^{i_T}, \quad q = 1, \ldots, Q, \tag{9.93}$$

where we emphasized that the distribution that generates the Monte Carlo scenarios is determined by the available time series of market invariants i_T. Then we compute the resampled allocation (9.92) as follows:

$$\begin{aligned}\boldsymbol{\alpha}_{\mathrm{rs}}\left[i_T\right] \equiv \left[\operatorname{diag}\left(\mathbf{p}_T\right)\right]^{-1} & \left(\frac{\zeta}{Q}\sum_{q=1}^{Q} {}_q\mathbf{V}^{i_T}\widehat{\boldsymbol{\mu}}\right. \\ & \left. - \frac{\zeta}{Q}\sum_{q=1}^{Q}\frac{\mathbf{1}'{}_q\mathbf{V}^{i_T}\widehat{\boldsymbol{\mu}}}{\mathbf{1}'{}_q\mathbf{V}^{i_T}\mathbf{1}}{}_q\mathbf{V}\mathbf{1} + \frac{w_T}{Q}\sum_{q=1}^{Q}\frac{{}_q\mathbf{V}^{i_T}\mathbf{1}}{\mathbf{1}'{}_q\mathbf{V}^{i_T}\mathbf{1}}\right).\end{aligned} \tag{9.94}$$

Notice that the resampled allocation depends on the available time series of market invariants i_T, because this determines the Monte Carlo simulations (9.93) through (9.87) and (9.88).

In Figure 9.5 we display the resampled allocation $\boldsymbol{\alpha}_{\mathrm{rs}}[i_T]$ along with the sample-based allocation that $\boldsymbol{\alpha}_{\mathrm{s}}[i_T]$ in the plane of the coordinates that determine the investor's satisfaction and constraints, see (8.25) and (8.36):

$$v \equiv \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T)\, \boldsymbol{\Sigma} \operatorname{diag}(\mathbf{p}_T)\, \boldsymbol{\alpha} \tag{9.95}$$

$$e \equiv \boldsymbol{\alpha}' \operatorname{diag}(\mathbf{p}_T)\,(\mathbf{1} + \boldsymbol{\mu})\,. \tag{9.96}$$

In the specific case plotted in the figure the resampling process generates an allocation with less opportunity cost than the sample-based allocation. Furthermore the resampled allocation satisfies the constraints, as opposed to the sample-based allocation, compare with Figure 8.6 and the respective discussion. Nonetheless, we remark that there is no guarantee that this will always be the case, see the discussion below in Section 9.3.4.

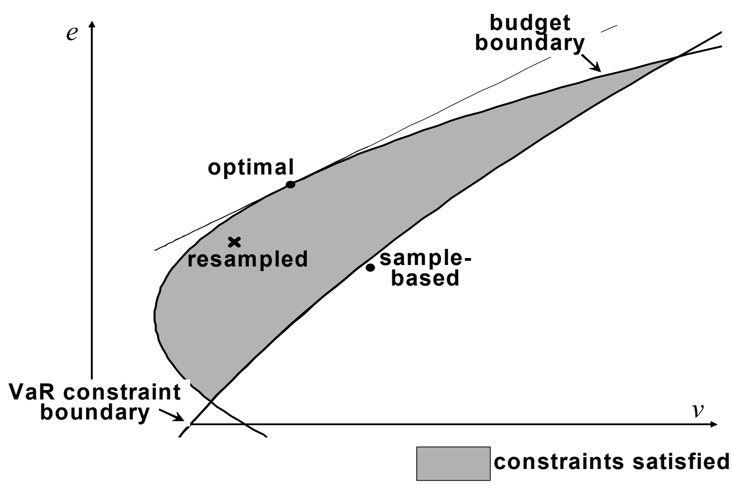

Fig. 9.5. Resampled allocation: comparison with sample-based allocation

9.3.3 Evaluation

To evaluate the sample-based allocation decision we should proceed in principle as in Chapter 8.

First we should consider a set $\boldsymbol{\Theta}$ of market parameters that is broad enough to most likely include the true, unknown value $\boldsymbol{\theta}^{\mathrm{t}}$.

For each value $\boldsymbol{\theta}$ of the market parameters in the stress test set $\boldsymbol{\Theta}$ we should compute the optimal allocation function, see $\boldsymbol{\alpha}(\boldsymbol{\theta})$ (9.66). Then we should compute as in (8.31) the optimal level of satisfaction if $\boldsymbol{\theta}$ are the underlying market parameters, namely $\overline{\mathcal{S}}(\boldsymbol{\theta})$.

Next, we should randomize as in (8.48) the information from the market i_T, generating a distribution of information scenarios that depends on the assumption $\boldsymbol{\theta}$ on the market parameters:

$$I_T^{\boldsymbol{\theta}} \equiv \left\{ \mathbf{X}_1^{\boldsymbol{\theta}}, \ldots, \mathbf{X}_T^{\boldsymbol{\theta}} \right\}. \tag{9.97}$$

Then we should compute the resampled allocation (9.86) from the randomized information, obtaining the random variable $\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right]$.

Next we should compute as in (8.23) the satisfaction $\mathcal{S}_{\boldsymbol{\theta}} \left(\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right] \right)$ ensuing from each scenario of the resampled allocation decision under the assumption $\boldsymbol{\theta}$ for the market parameters, which, we recall, is a random variable. Similarly, from (8.26) and expressions such as (8.35) we should compute the cost of the resampled allocation decision violating the constraints $\mathcal{C}_{\boldsymbol{\theta}}^{+} \left(\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right] \right)$ in each scenario, which is also a random variable.

Then we should compute the opportunity cost (8.53) of the resampled allocation under the assumption $\boldsymbol{\theta}$ for the market parameters, namely the random variable defined as the difference between the optimal unattainable level of satisfaction and the satisfaction from the resampled allocation, plus the cost of the resampled allocation violating the constraints:

$$\mathrm{OC}_{\boldsymbol{\theta}} \left(\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right] \right) \equiv \overline{\mathcal{S}}(\boldsymbol{\theta}) - \mathcal{S}_{\boldsymbol{\theta}} \left(\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right] \right) + \mathcal{C}_{\boldsymbol{\theta}}^{+} \left(\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right] \right). \tag{9.98}$$

Finally, as in (8.57) we should compute the opportunity cost of the resampled allocation as a function of the underlying market parameters:

$$\boldsymbol{\theta} \mapsto \mathrm{OC}_{\boldsymbol{\theta}} \left(\boldsymbol{\alpha}_{\mathrm{rs}} \left[I_T^{\boldsymbol{\theta}} \right] \right), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta}. \tag{9.99}$$

The resampled allocation would be suitable if the opportunity cost turns out tightly distributed around a value close to zero for all the market parameters $\boldsymbol{\theta}$ in the stress test range $\boldsymbol{\Theta}$.

Unfortunately, the above evaluation cannot be done. Indeed, in practice, the randomization (9.97) is performed by generating a large number J of Monte Carlo realizations of the time series of the market invariants:

$$\boldsymbol{\theta} \overset{\text{stress test}}{\mapsto} {}^{j}i_T \equiv \left\{ {}^{j}\mathbf{x}_1, \ldots, {}^{j}\mathbf{x}_T \right\}, \quad j = 1, \ldots, J. \tag{9.100}$$

In turn, in each scenario j the resampled allocation is obtained by implementing a second Monte Carlo simulation as in (9.84). In other words, the distribution of the opportunity cost as a function of the assumptions on the underlying parameters (9.99) is obtained trough the following chain of steps:

$$\begin{aligned} \boldsymbol{\theta} &\overset{\text{stress test}}{\mapsto} {}^{j}i_T \overset{\text{estimate}}{\mapsto} {}^{j}_{0}\widehat{\boldsymbol{\theta}} \overset{\text{resample}}{\mapsto} {}^{j}_{q}i_T \overset{\text{estimate}}{\mapsto} {}^{j}_{q}\widehat{\boldsymbol{\theta}} \\ &\overset{\text{optimize}}{\mapsto} {}^{j}_{q}\boldsymbol{\alpha} \overset{\text{average}}{\mapsto} {}^{j}\boldsymbol{\alpha}_{\mathrm{rs}} \overset{\text{evaluate}}{\mapsto} \mathrm{OC}_{\boldsymbol{\theta}} \left({}^{j}\boldsymbol{\alpha}_{\mathrm{rs}} \right). \end{aligned} \tag{9.101}$$

To implement this chain we need to solve an optimization problem for each Monte Carlo scenario q stemming from another Monte Carlo scenario j: the computational burden of this operation is prohibitive.

9.3.4 Discussion

The resampling technique is very innovative. It displays several advantages but also a few drawbacks, see also Markowitz and Usmen (2003) and Ceria and Stubbs (2004).

In the first place, intuitively the expectation in the definition (9.86) of the resampled allocation decision reduces the sensitivity to the market parameters, and thus it gives rise to a less disperse opportunity cost than the sample-based allocation decision. Nonetheless, the proof of this statement for generic markets and preferences is not obvious.

Furthermore, the expectation in the definition (9.86) of the resampled allocation can give rise to resampled allocations that violate the investment constraints, not only in the case where the constraints depend on the unknown market parameters. For instance, consider the constraint (8.15) of not investing in more than M of the N securities in the market: each allocation ${}_q\boldsymbol{\alpha}$ in the average (9.83) satisfies this constraint, but the ensuing resampled allocation does not.

Finally, it is very hard to stress test the performance of this technique due to the excessive computational burden, see (9.101) and comments thereafter.

9.4 Robust allocation

So far the pursuit of optimal allocation strategies has focused on fixing the excessive sensitivity to the input parameters of the optimal allocation function. The robust approach aims directly at determining the "best" allocation, according to the evaluation criteria discussed in Chapter 8.

First we formalize the intuitive definition of robust allocation decisions for general markets and preferences. Then, in order to compute the solution of a robust allocation problem in practice, we resort to the two-step mean-variance framework.

9.4.1 General definition

Consider the opportunity cost of a generic allocation $\boldsymbol{\alpha}$ that satisfies the investment constraints, which is defined in (8.37) as the difference between the maximum possible satisfaction and the actual satisfaction provided by the given allocation:

$$\mathrm{OC}_{\boldsymbol{\theta}}(\boldsymbol{\alpha}) \equiv \overline{\mathcal{S}}(\boldsymbol{\theta}) - \mathcal{S}_{\boldsymbol{\theta}}(\boldsymbol{\alpha}) . \tag{9.102}$$

According to the discussion in Section 8.1, since the true value of the market parameters $\boldsymbol{\theta}$ is not known, an allocation is optimal if it gives rise to a minimal opportunity cost for all the values of the market parameters in an uncertainty range $\boldsymbol{\Theta}$ that is broad enough to most likely include the true, unknown value $\boldsymbol{\theta}^{\mathrm{t}}$ of the market parameters. This way in particular the opportunity cost is guaranteed to be low in correspondence of the unknown value $\boldsymbol{\theta}^{\mathrm{t}}$.

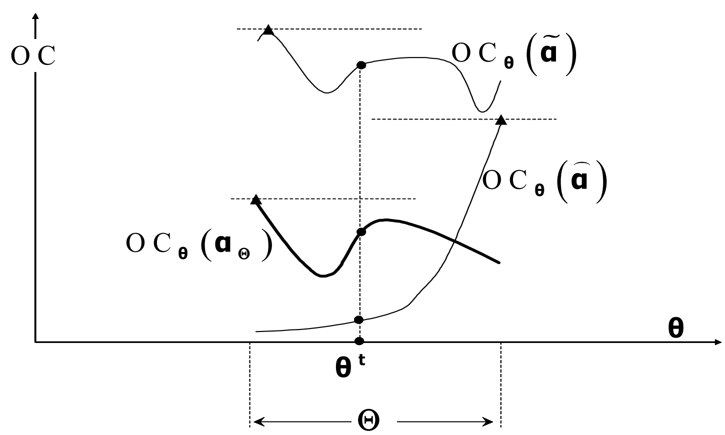

Fig. 9.6. Opportunity cost as function of the market parameters

The robust approach aims precisely at determining an allocation $\boldsymbol{\alpha}$ such that the opportunity cost is uniformly minimal for all the values $\boldsymbol{\theta}$ in the uncertainty range $\boldsymbol{\Theta}$. To make sure that the opportunity cost is uniformly low for all the values $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}$ we take a conservative approach and monitor its maximum over the range $\boldsymbol{\Theta}$, see Figure 9.6. Furthermore, we require that the allocation $\boldsymbol{\alpha}$ satisfies the constraints for all the values $\boldsymbol{\theta}$ in the given range $\boldsymbol{\Theta}$, a condition which we denote as follows:

$$\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\Theta}} \equiv \{\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\theta}} \text{ for all } \boldsymbol{\theta} \in \boldsymbol{\Theta}\}. \tag{9.103}$$

In other words, we consider the allocation such that the maximum opportunity cost (9.102) on the given range is the lowest possible:

$$\boldsymbol{\alpha}_{\boldsymbol{\Theta}} \equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}_{\boldsymbol{\Theta}}}{\operatorname{argmin}} \left\{ \max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \left\{ \overline{\mathcal{S}}(\boldsymbol{\theta}) - \mathcal{S}_{\boldsymbol{\theta}}(\boldsymbol{\alpha}) \right\} \right\}. \tag{9.104}$$

Notice that this allocation in general does not give rise the least possible opportunity cost in correspondence of the true parameters $\boldsymbol{\theta}^{\mathrm{t}}$, although the damage is guaranteed to be contained, see Figure 9.6.

The allocation (9.104) and its quality depend on the choice for the uncertainty range $\boldsymbol{\Theta}$ of the market parameters, see Figure 9.7. The smaller the

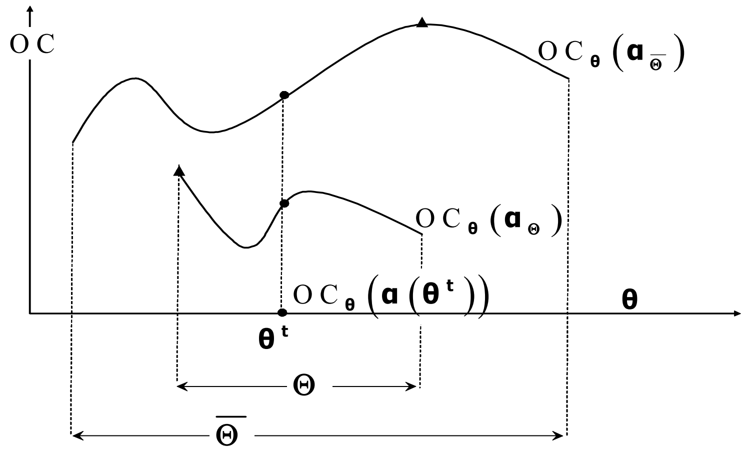

Fig. 9.7. Quality of robust allocation as function of the uncertainty range

range $\boldsymbol{\Theta}$, the lower the maximum value of the opportunity cost generated by $\boldsymbol{\alpha}_{\boldsymbol{\Theta}}$ and thus the higher the quality of $\boldsymbol{\alpha}_{\boldsymbol{\Theta}}$. Indeed, in the limit case where the evaluation set is the single true value $\boldsymbol{\theta}^{\mathrm{t}}$ the ensuing allocation (9.104) becomes the truly optimal solution (8.39), which gives rise to a null opportunity cost. As we expand the evaluation set $\boldsymbol{\Theta}$, the opportunity cost of the best allocation (9.104), although it is uniformly the least among all the possible allocations, increases.

To summarize, we built a recipe to pursue the best allocation by accounting for estimation risk: first, determine an uncertainty range $\boldsymbol{\Theta}$ of market parameters that contains the true parameter $\boldsymbol{\theta}^{\mathrm{t}}$, and yet it is as small as possible; then solve the optimization (9.104).

Consider our leading example where satisfaction is determined by the certainty-equivalent of an exponential utility function and the investor has a full-investment budget constraint and a value at risk constraint. Assume that we determined a suitable range $\boldsymbol{\Theta}$ for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. The allocation recipe (9.104) reads in this context:

$$\boldsymbol{\alpha}_{\boldsymbol{\Theta}} \equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ \max_{\boldsymbol{\mu},\boldsymbol{\Sigma}\in\boldsymbol{\Theta}} \left\{ \overline{\mathrm{CE}}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right) - \mathrm{CE}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(\boldsymbol{\alpha}\right) \right\} \right\} \tag{9.105}$$
$$\text{subject to } \begin{cases} \boldsymbol{\alpha}'\mathbf{p}_T = w_T \\ \mathrm{Var}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(\boldsymbol{\alpha}\right) \leq \gamma w_T, \text{ for all } \boldsymbol{\mu},\boldsymbol{\Sigma} \in \boldsymbol{\Theta}, \end{cases}$$

where the explicit expression of the certainty equivalent and the VaR are provided in (8.25), (8.28) and (8.33).

In order to be confident that the range $\boldsymbol{\Theta}$ contains $\boldsymbol{\theta}^{\mathrm{t}}$ and yet it is as small as possible we need to collect information from the market. Just like a generic estimator (8.78) associates with the available information i_T a *value* $\widehat{\boldsymbol{\theta}}$ that suitably represents a quantity of interest, so we can use the available information to determine a suitable *range* of values, which we call the *uncertainty set*, or the *robustness set*:

$$i_T \mapsto \widehat{\boldsymbol{\Theta}}\left[i_T\right]. \tag{9.106}$$

There exists a variety of methods to perform this operation, which generalize the theory of point estimation discussed in Chapter 4. We discuss in Section 9.5 one of these methods, which relies on the Bayesian approach to parameter estimation.

For instance, consider a market where the linear returns of the N securities are independent and normally distributed:

$$\mathbf{L}_t \sim \mathrm{N}\left(\boldsymbol{\mu}^{\mathrm{t}}, \boldsymbol{\Sigma}^{\mathrm{t}}\right), \tag{9.107}$$

where $\boldsymbol{\mu}^{\mathrm{t}}$ and $\boldsymbol{\Sigma}^{\mathrm{t}}$ are the true expected values and covariance matrix respectively.

Assume that the covariance $\boldsymbol{\Sigma}^{\mathrm{t}}$ is known. We have to determine a suitable uncertainty set $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}$ for $\boldsymbol{\mu}$ such that we can be confident that the true parameter $\boldsymbol{\mu}^{\mathrm{t}}$ lies within its boundaries. Consider the sample estimator $\widehat{\boldsymbol{\mu}}\left[i_T\right]$ defined in (8.79), and define the uncertainty set $\widehat{\boldsymbol{\Theta}}$ as follows:

$$\widehat{\boldsymbol{\Theta}}_{\mu}\left[i_T\right] \equiv \left\{\boldsymbol{\mu} \text{ such that } \operatorname{Ma}^2\left(\boldsymbol{\mu}, \widehat{\boldsymbol{\mu}}\left[i_T\right], \boldsymbol{\Sigma}^{\mathrm{t}}\right) \leq \frac{Q_{\chi^2_N}\left(p\right)}{T}\right\}, \tag{9.108}$$

where Ma is the Mahalanobis distance (2.61) of $\boldsymbol{\mu}$ from $\widehat{\boldsymbol{\mu}}$ induced by the metric $\boldsymbol{\Sigma}^{\mathrm{t}}$; $Q_{\chi^2_N}\left(p\right)$ is the quantile of the chi-square distribution with N degrees of freedom (1.109) for a confidence level $p \in (0,1)$; and T is the number of observations in the time series of the returns that we use to estimate $\widehat{\boldsymbol{\mu}}$.

The set (9.108) is an ellipsoid centered in $\widehat{\boldsymbol{\mu}}$, with shape determined by $\boldsymbol{\Sigma}^{\mathrm{t}}$ and with radius proportional to $1/\sqrt{T}$, see $(A.73)$ and comments thereafter. As we show in Appendix www.9.2 the following result holds for the probability that the range (9.108) captures the true expected values:

$$\mathbb{P}\left\{\boldsymbol{\mu}^{\mathrm{t}} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}\left[i_T\right]\right\} = p. \tag{9.109}$$

In other words, with a confidence p that we can set arbitrarily, the true parameter $\boldsymbol{\mu}^{\mathrm{t}}$ lies within the set (9.108): as we require a higher confidence, the quantile in (9.108) increases, and so does the size of the ellipsoid. As intuition suggests, for a given confidence p, the more information is available, i.e. the larger the number of observations T in the time series of the returns, the smaller the uncertainty ellipsoid.

By letting the evaluation range in the optimization problem (9.104) be determined by currently available information as in (9.106), we obtain the definition of the *robust allocation decision*:

$$\boldsymbol{\alpha}_{\mathrm{r}}\left[i_T\right] \equiv \underset{\boldsymbol{\alpha} \in \mathcal{C}_{\widehat{\boldsymbol{\Theta}}[i_T]}}{\operatorname{argmin}} \left\{ \max_{\boldsymbol{\theta} \in \widehat{\boldsymbol{\Theta}}[i_T]} \left\{ \overline{\mathcal{S}}\left(\boldsymbol{\theta}\right) - \mathcal{S}_{\boldsymbol{\theta}}\left(\boldsymbol{\alpha}\right) \right\} \right\}. \tag{9.110}$$

This is indeed a decision, which processes the currently available information, see (8.38).

The smaller the uncertainty set $\widehat{\boldsymbol{\Theta}}$ in (9.110), the less conservative the investor from the point of view of estimation risk. Indeed, in the limit where the robustness set consists of only one point, namely the point estimate $\widehat{\boldsymbol{\theta}}$, the robust allocation decision becomes the sample-based allocation decision (8.81). Nevertheless, we stress that if the uncertainty set is very likely to include the true unknown parameters, the smaller the uncertainty set, the better the quality of the robust allocation, see Figure 9.7.

In Appendix www.9.5 we show that using the uncertainty set (9.108) in (9.105) the ensuing robust allocation decision solves the following problem:

$$\boldsymbol{\alpha}_{\mathrm{r}} \equiv \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ \max_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_{\mu}} \left\{ \begin{array}{l} \boldsymbol{\mu}'\mathbf{T}\boldsymbol{\mu} + \frac{w_T}{A}\mathbf{1}'\left(\boldsymbol{\Sigma}^{\mathrm{t}}\right)^{-1}\boldsymbol{\mu} \\ -\boldsymbol{\alpha}'\operatorname{diag}\left(\mathbf{p}_T\right)\boldsymbol{\mu} + \frac{1}{2\zeta}\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|^2 \end{array} \right\} \right\}, \tag{9.111}$$

subject to:

$$\begin{cases} \boldsymbol{\alpha}'\mathbf{p}_T = w_T \\ \sqrt{2}\tau\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\| \leq \widehat{\boldsymbol{\mu}}'\operatorname{diag}\left(\mathbf{p}_T\right)\boldsymbol{\alpha} + \gamma w_T \\ \qquad\qquad + \left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|^2 - \frac{\sqrt{Q_N(p)/T}}{\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\boldsymbol{\alpha}\right\|}. \end{cases} \tag{9.112}$$

In this expression

$$\begin{aligned} A &\equiv \mathbf{1}'\left(\boldsymbol{\Sigma}^{\mathrm{t}}\right)^{-1}\mathbf{1} \\ \tau &\equiv \operatorname{erf}^{-1}\left(2c-1\right) \\ \mathbf{T} &\equiv \frac{\zeta}{2}\left(\boldsymbol{\Sigma}^{\mathrm{t}}\right)^{-1}\left(\mathbf{I} - \frac{1}{A}\mathbf{1}\mathbf{1}'\left(\boldsymbol{\Sigma}^{\mathrm{t}}\right)^{-1}\right); \end{aligned} \tag{9.113}$$

and $\boldsymbol{\Lambda}$ and $\mathbf{E}$ are the eigenvalues and the eigenvectors respectively of the following spectral decomposition:

$$\operatorname{diag}\left(\mathbf{p}_T\right)\boldsymbol{\Sigma}^{\mathrm{t}}\operatorname{diag}\left(\mathbf{p}_T\right) \equiv \mathbf{E}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\mathbf{E}'. \tag{9.114}$$

The maximization for a given $\boldsymbol{\alpha}$ in (9.111) is satisfied by the tangency condition of ellipsoidal contours in the variable $\boldsymbol{\mu}$ with a fixed ellipsoid: this problem does not admit analytical solutions, as it is a modification of the spectral equation, see (A.68). Therefore, the second optimization, namely the minimization in (9.111) cannot be performed. Furthermore, the VaR constraint in (9.112) is not a conic constraint. Therefore the solution of the robust allocation is not numerically tractable, see Section 6.2.

9.4.2 Practicable definition: the mean-variance setting

Although the rationale behind the robust allocation decision is conceptually simple, solving the min-max optimization (9.110) is close to impossible even under simple assumptions on preferences, markets and constraints, as we have seen in the example (9.111)-(9.112).

Therefore robust allocation is tackled in practice within the two-step mean-variance framework. This is not surprising, since we resorted to the mean-variance approximation even in the classical setting that disregards estimation risk. When the robust allocation problem is set in the mean-variance framework we can apply recent results on robust optimization, see El Ghaoui and Lebret (1997) and Ben-Tal and Nemirovski (1995), see also Ben-Tal and Nemirovski (2001).

We recall from Section 6.3 that the mean-variance approach is a two-step simplification of a generic allocation problem: the investor first determines a set of mean-variance efficient allocations and then he selects among those allocations the one that better suits him.

We assume that the investment constraints $\mathcal{C}$ do not depend on the unknown market parameters and are such that the inequality version (6.144) of the mean-variance problem applies, see Section 6.5.3. In this setting the mean-variance problem can be written as follows:

$$\boldsymbol{\alpha}^{(i)} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}}\, \boldsymbol{\alpha}'\boldsymbol{\mu} \tag{9.115}$$
$$\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ \boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha} \leq v^{(i)}. \end{cases}$$

In this expression $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the expected value and the covariance matrix respectively of the market vector $\mathbf{M}$:

$$\boldsymbol{\mu} \equiv \mathrm{E}\{\mathbf{M}\}, \quad \boldsymbol{\Sigma} \equiv \operatorname{Cov}\{\mathbf{M}\}; \tag{9.116}$$

the market vector $\mathbf{M}$ in turn is the affine transformation of the prices at the investment horizon $\mathbf{P}_{T+\tau}$ which together with the allocation vector $\boldsymbol{\alpha}$ determines the investor's objective $\Psi \equiv \boldsymbol{\alpha}'\mathbf{M}$, see (5.10); the set $\{v^{(1)}, \ldots, v^{(I)}\}$ is a significative grid of target variances of the investor's objective.

According to (9.110), the robust version of the mean-variance problem (9.115) reads:

$$\boldsymbol{\alpha}_{\mathrm{r}}^{(i)} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{ \min_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}} \left\{ \boldsymbol{\alpha}' \boldsymbol{\mu} \right\} \right\} \tag{9.117}$$

$$\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ \max\limits_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \left\{ \boldsymbol{\alpha}' \boldsymbol{\Sigma} \boldsymbol{\alpha} \right\} \leq v^{(i)}, \end{cases}$$

where $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}$ are uncertainty sets for the market parameters (9.116) that are estimated from the available information i_T. Depending on the specification of these uncertainty sets, the resulting robust problem assumes different forms.

- **Known covariances, elliptical set for expected values**

A possible specification for the uncertainty sets assumes an ellipsoidal shape for the uncertainty on the parameter $\boldsymbol{\mu}$ and no uncertainty for $\boldsymbol{\Sigma}$:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{ \boldsymbol{\mu} \text{ such that } \operatorname{Ma}^2 \left(\boldsymbol{\mu}, \mathbf{m}, \mathbf{T} \right) \leq q^2 \right\} \tag{9.118}$$

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} \equiv \widehat{\boldsymbol{\Sigma}}. \tag{9.119}$$

In this expression $\widehat{\boldsymbol{\Sigma}}$ is a point estimate of $\boldsymbol{\Sigma}$; $\mathbf{m}$ is an N-dimensional vector; $\mathbf{T}$ is an $N \times N$ symmetric and positive matrix; Ma is the Mahalanobis distance (2.61) of $\boldsymbol{\mu}$ from $\mathbf{m}$ induced by the metric $\mathbf{T}$; and

$$q^2 \equiv Q_{\chi^2_N} \left(p \right) \tag{9.120}$$

is the quantile of the chi-square distribution with N degrees of freedom (1.109) for a confidence level $p \in (0, 1)$.

Ceria and Stubbs (2004) consider the following specification in (9.118):

$$\mathbf{m} \equiv \widehat{\boldsymbol{\mu}} \left[i_T \right], \ \mathbf{T} \text{ exogenous,} \tag{9.121}$$

where $\widehat{\boldsymbol{\mu}}$ is a sample-based estimator of the true parameter.

De Santis and Foresi (2002) blend a market model with the investor's views by specifying the parameters in (9.118) in terms of the Black-Litterman posterior distribution (9.44):

$$\mathbf{m} \equiv \boldsymbol{\mu}_{\mathrm{BL}}, \quad \mathbf{T} \equiv \boldsymbol{\Sigma}_{\mathrm{BL}}. \tag{9.122}$$

The uncertainty set (9.118) is an ellipsoid centered in $\mathbf{m}$ whose shape is determined by $\mathbf{T}$, see (A.73) and comments thereafter. The rationale behind the assumption (9.118) is that the uncertainty about $\boldsymbol{\mu}$ is approximately normally distributed:

$$\boldsymbol{\mu} \sim \mathrm{N} \left(\mathbf{m}, \mathbf{T} \right), \tag{9.123}$$

see also (9.108). In Appendix www.7.1 we show that in this case the following result holds for the probability that the range captures the true expected values:

$$\mathbb{P}\left\{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_{\mu}\right\} = p. \tag{9.124}$$

If the investor considers small ellipsoids by setting p close to zero, he is little worried about missing the true expected values in the optimization (9.117). In other words, he is very aggressive as far as estimation risk is concerned. On the other hand, if the the investor sets p close to one, he is very cautious from the point of view of estimation risk.

As we discuss in Section 9.4.3, if the investment constraints $\mathcal{C}$ are sufficiently regular, the optimization (9.117) simplifies to a second-order cone programming problem and thus the robust frontier can be computed numerically.

- **Box set for expected values, elliptical set for covariances**

An alternative specification of the uncertainty sets in the robust optimization (9.117) is adopted by Goldfarb and Iyengar (2003). The uncertainty set for the expected values is of the box-form:

$$\widehat{\boldsymbol{\Theta}}_{\mu} \equiv \left\{\boldsymbol{\mu} \text{ such that } \underline{\boldsymbol{\mu}} \leq \boldsymbol{\mu} \leq \overline{\boldsymbol{\mu}}\right\}. \tag{9.125}$$

The uncertainty set for the covariance matrix follows from a K-factor model such as (3.119), where factors and perturbations are uncorrelated. In other words, the uncertainty set for the covariance matrix is specified as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} \equiv \left\{\mathbf{B}\mathbf{G}\mathbf{B}' + \operatorname{diag}(\mathbf{d})\right\}. \tag{9.126}$$

In this expression $\underline{\mathbf{d}} \leq \mathbf{d} \leq \overline{\mathbf{d}}$; the covariance $\mathbf{G}$ of the factors is assumed known, and each row $\mathbf{b}_{(n)}$ of the $N \times K$ matrix of the factor loadings $\mathbf{B}$ belongs to an ellipsoid such as $(A.73)$:

$$\mathbf{b}_{(n)} \in \mathcal{E}_n, \quad n = 1, \ldots, N. \tag{9.127}$$

As it turns out, when the investment constraints $\mathcal{C}$ are sufficiently regular, this specification also gives rise to a second-order cone programming problem. Therefore the robust frontier can computed numerically, see Section 6.2.

- **Box set for expected values, box set for covariances**

A third possible specification of the uncertainty sets in (9.117) is provided by Halldorsson and Tutuncu (2003), who assume box-sets for all the parameters:

$$\widehat{\boldsymbol{\Theta}}_{\mu} \equiv \left\{\boldsymbol{\mu} \text{ such that } \underline{\boldsymbol{\mu}} \leq \boldsymbol{\mu} \leq \overline{\boldsymbol{\mu}}\right\} \tag{9.128}$$

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} \equiv \left\{\boldsymbol{\Sigma} \succeq \mathbf{0} \text{ such that } \underline{\boldsymbol{\Sigma}} \leq \boldsymbol{\Sigma} \leq \overline{\boldsymbol{\Sigma}}\right\}, \tag{9.129}$$

where the notation $\boldsymbol{\Sigma} \succeq \mathbf{0}$ stands for symmetric and positive matrices. Under further assumptions on the investment constraints $\mathcal{C}$, the ensuing robust mean-variance problem can be cast in the form of a saddle-point search and solved numerically with an interior-point algorithm.

9.4.3 Discussion

As we show in Appendix www.9.6, the robust mean-variance problem (9.117) under the specifications (9.118)-(9.119) for the robustness sets can be written equivalently as follows:

$$\boldsymbol{\alpha}_{\mathrm{r}}^{(i)} = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} \left\{ \boldsymbol{\alpha}'\mathbf{m} - q\sqrt{\boldsymbol{\alpha}'\mathbf{T}\boldsymbol{\alpha}} \right\} \tag{9.130}$$
$$\text{subject to } \begin{cases} \boldsymbol{\alpha} \in \mathcal{C} \\ \boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha} \leq v^{(i)}. \end{cases}$$

If the investment constraints $\mathcal{C}$ are regular enough, this problem can be cast in the form of a second-order cone programming problem (6.55), see Appendix www.9.6. Therefore the robust frontier can computed numerically.

The robust efficient frontier (9.130) represents a two-parameter family of allocations, i.e. a surface, determined by the target variance v, which represents market risk, and the size of the uncertainty ellipsoid, which is directly related to q and represents aversion to estimation risk, see (9.124) and comments thereafter.

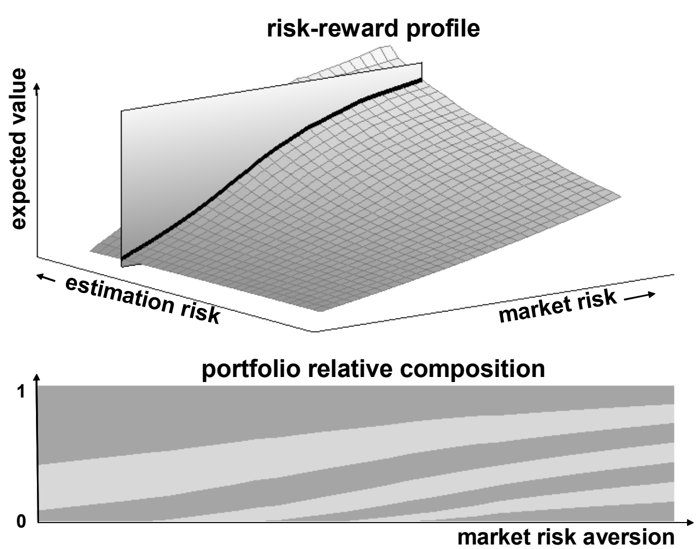

Fig. 9.8. Robust efficient allocations: fixed aversion to estimation risk

We can parameterize the robust surface (9.130) equivalently in terms of a market risk Lagrange multiplier $\gamma_m \geq 0$ and an estimation risk multiplier $\gamma_e \equiv q$ as follows:

$$\boldsymbol{\alpha}_r(\gamma_m, \gamma_e) = \underset{\boldsymbol{\alpha} \in \mathcal{C}}{\operatorname{argmax}} \left\{ \boldsymbol{\alpha}'\mathbf{m} - \gamma_m \sqrt{\boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha}} - \gamma_e \sqrt{\boldsymbol{\alpha}'\mathbf{T}\boldsymbol{\alpha}} \right\}. \quad (9.131)$$

This way we obtain the following interpretation of the two-parameter robust frontier: investors balance the trade off between the expected value of their objective, represented by the term $\boldsymbol{\alpha}'\mathbf{m}$, and risk. Risk appears in two forms: market risk, represented by the market volatility $\sqrt{\boldsymbol{\alpha}'\widehat{\boldsymbol{\Sigma}}\boldsymbol{\alpha}}$, and estimation risk, represented by the estimation uncertainty $\sqrt{\boldsymbol{\alpha}'\mathbf{T}\boldsymbol{\alpha}}$.

Larger values of the multiplier γ_m give rise to allocations that suit investors who are more averse to market risk: therefore we can interpret γ_m as a market risk aversion parameter. Similarly, larger values of the multiplier γ_e give rise to allocations that suit investors who are more averse to estimation risk: therefore we can interpret γ_e as an estimation risk aversion parameter.

In Figure 9.8 we compute the robust efficient frontier for a market of $N \equiv 7$ securities, under the standard constraints of no short-selling, namely $\boldsymbol{\alpha} \geq \mathbf{0}$, and of full investment of the initial budget w_T, namely $\boldsymbol{\alpha}'\mathbf{p}_T = w_T$

The top plot displays the expected value $\mathbf{m}'\boldsymbol{\alpha}_{\mathrm{r}}$ of the robust efficient surface (9.131) as a function of the aversion to market risk γ_m and of the aversion to estimation risk γ_e.

The bottom plot displays the robust allocations $\boldsymbol{\alpha}_{\mathrm{r}}(\gamma_m, \overline{\gamma_e})$ in terms of the relative portfolio weights for a given level of estimation risk, i.e. for a fixed value $\overline{\gamma_e}$: these are the allocations that correspond to the "slice" of the robust surface in the top portion of the figure.

Similarly, in the top plot in Figure 9.9 we display the expected value $\mathbf{m}'\boldsymbol{\alpha}_{\mathrm{r}}$ of the robust efficient surface as a function of the aversion to market risk γ_m and of the aversion estimation risk γ_e.

The bottom plot displays the robust allocations $\boldsymbol{\alpha}_{\mathrm{r}}(\overline{\gamma_m}, \gamma_e)$ in terms of the relative portfolio weights for a given level of market risk, i.e. for a fixed value $\overline{\gamma_m}$: these are the allocations that correspond to the "slice" of the robust surface in the top portion of the figure.

9.5 Robust Bayesian allocation

Robust allocation decision are optimal over a whole range of market parameters, because by construction they minimize the opportunity cost over the given range. Nevertheless, in the classical approach, the choice of the robustness range is quite arbitrary.

Using the Bayesian approach to estimation we can naturally identify a suitable robustness range for the market parameters: robust Bayesian allocation decisions account for estimation risk over a range of market parameters that includes both the available information and the investor's experience according to a self-adjusting mechanism.

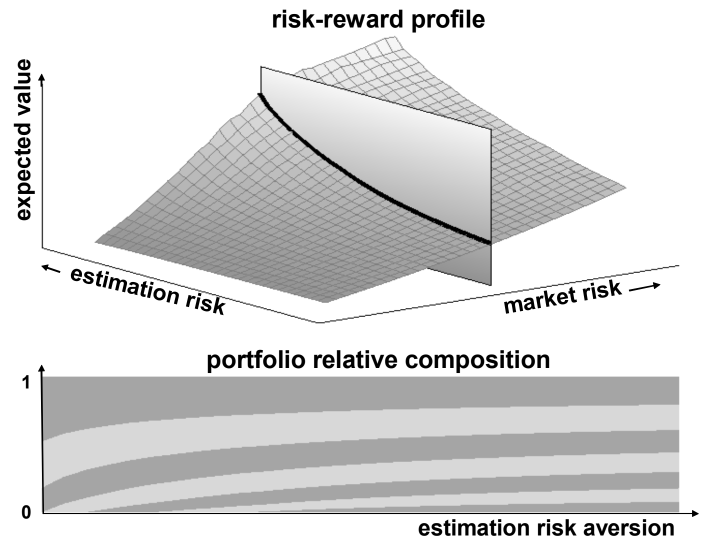

Fig. 9.9. Robust efficient allocations: fixed aversion to market risk

9.5.1 General definition

The robust allocation decision (9.110) minimizes the opportunity cost due to estimation risk uniformly over the uncertainty set $\widehat{\boldsymbol{\Theta}}$ for market parameters. The choice of the uncertainty set is crucial for the success of the respective allocation strategy: on the one hand $\widehat{\boldsymbol{\Theta}}$ should be as small as possible, in order to keep the maximum possible opportunity cost low; on the other hand $\widehat{\boldsymbol{\Theta}}$ should be as large as possible, in order to most likely include the true unknown parameters.

The Bayesian framework defines uncertainty sets in a natural way. Indeed, in the Bayesian framework the unknown market parameters $\boldsymbol{\theta}$ are random variables. The likelihood that the parameters assume given values is described by the posterior probability density function $f_{\text{po}}(\boldsymbol{\theta})$, which is determined by the available information i_T and by the investor's experience e_C, see Figure 7.1. The region where the posterior distribution displays a higher concentration deserves more attention than the tails of the distribution: this region is a natural choice for the uncertainty set $\widehat{\boldsymbol{\Theta}}$.

From the discussion in Section 7.1.2, the region where the posterior distribution displays a higher concentration is represented by the location-dispersion ellipsoid of the market parameters (7.10), see Figure 7.2:

$$\widehat{\boldsymbol{\Theta}}^q [i_T, e_C] \equiv \left\{ \boldsymbol{\theta} : \left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\text{ce}} \right)' \mathbf{S}_{\boldsymbol{\theta}}^{-1} \left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_{\text{ce}} \right) \leq q^2 \right\}. \tag{9.132}$$

In this expression S is the dimension of the vector $\boldsymbol{\theta}$; $\widehat{\boldsymbol{\theta}}_{\text{ce}}$ is a classical-equivalent estimator of the market parameters, such as the expected value (7.5) or the mode (7.6); and $\mathbf{S}_{\boldsymbol{\theta}}$ is a scatter matrix for the market parameters, such as the covariance matrix (7.7) or the modal dispersion (7.8).

Using the Bayesian location-dispersion ellipsoid (9.132) as the uncertainty set for the robust allocation decision (9.110) we obtain the *robust Bayesian allocation decision*:

$$\boldsymbol{\alpha}_{\text{rB}}\left[i_T, e_C\right] \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\widehat{\boldsymbol{\Theta}}^q\left[i_T,e_C\right]}}{\operatorname{argmin}}\left\{\max_{\boldsymbol{\theta}\in\widehat{\boldsymbol{\Theta}}^q\left[i_T,e_C\right]}\left\{\overline{\mathcal{S}}\left(\boldsymbol{\theta}\right)-\mathcal{S}_{\boldsymbol{\theta}}\left(\boldsymbol{\alpha}\right)\right\}\right\}. \tag{9.133}$$

This decision minimizes the maximum possible opportunity cost of an allocation that satisfies the investment constraints for all the markets within the location-dispersion ellipsoid.

The robust Bayesian allocation decision is indeed a decision, as it processes the currently available information i_T as in (8.38) through the ellipsoid (9.132). Furthermore, the robust Bayesian allocation decision also processes the investor's experience e_C within a sound statistical framework. Finally, the robust Bayesian allocation decision also depends on the radius factor q. From (7.11) and (7.12) we can interpret q as the investor's aversion to estimation risk: the smaller q, the smaller the ellipsoid, the higher the chances that the true value of the market parameters are not included within the boundaries of the uncertainty set.

The interplay among the available information i_T, the investor's experience e_C and the investor's aversion to estimation risk q shapes the uncertainty set (9.132) and thus the robust Bayesian allocation decision (9.133) in a self-adjusting way.

Due to (7.4), when the confidence C in the investor's experience e_C is very large compared to the amount of information T from the market, the posterior distribution becomes extremely peaked around the prior $\boldsymbol{\theta}_0$. Therefore, no matter the aversion to estimation risk q, the robustness set (9.132) shrinks to the point $\boldsymbol{\theta}_0$, see the discussion in Section 7.1.2. In other words, the robust Bayesian allocation decision (9.133) becomes:

$$\boldsymbol{\alpha}_{\text{p}} \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\boldsymbol{\theta}_0}}{\operatorname{argmax}}\left\{\mathcal{S}_{\boldsymbol{\theta}_0}\left(\boldsymbol{\alpha}\right)\right\}. \tag{9.134}$$

This is a prior allocation decision, see (8.64).

Similarly, due to (7.4), when the amount T of information on the market i_T is very large compared to the confidence C in the investor's experience e_C, the posterior distribution becomes extremely peaked around its classical-equivalent estimator, which is determined by the sample i_T. Therefore, no matter the aversion to estimation risk q, the robustness set (9.132) shrinks to a point, namely the sample estimate $\widehat{\boldsymbol{\theta}}\left[i_T\right]$, see the discussion in Section 7.1.2. In other words the robust Bayesian allocation decision (9.133) becomes:

$$\boldsymbol{\alpha}_{\mathrm{s}}\left[i_T\right] \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\widehat{\boldsymbol{\theta}}[i_T]}}{\operatorname{argmax}}\left\{\mathcal{S}_{\widehat{\boldsymbol{\theta}}[i_T]}\left(\boldsymbol{\alpha}\right)\right\}. \tag{9.135}$$

This is the sample-based allocation decision, see (8.81).

When the aversion to estimation risk q in the definition of the robustness set (9.132) tends to zero, the radius of the ellipsoid shrinks to zero and thus the ellipsoid degenerates to a point, its center, which is the classical-equivalent estimator $\widehat{\boldsymbol{\theta}}_{\mathrm{ce}}$. Therefore the robust Bayesian allocation decision (9.133) becomes:

$$\boldsymbol{\alpha}_{\mathrm{ce}}\left[i_T, e_C\right] \equiv \underset{\boldsymbol{\alpha}\in\mathcal{C}_{\widehat{\boldsymbol{\theta}}_{\mathrm{ce}}[i_T,e_C]}}{\operatorname{argmax}}\left\{\mathcal{S}_{\widehat{\boldsymbol{\theta}}_{\mathrm{ce}}[i_T,e_C]}\left(\boldsymbol{\alpha}\right)\right\}. \tag{9.136}$$

This is the classical-equivalent Bayesian allocation decision, see (9.13).

For all the intermediate cases, the robust Bayesian allocation decision smoothly blends the information from the market with the investor's experience, at the same time accounting for estimation risk, within a sound, self-adjusting statistical framework.

9.5.2 Practicable definition: the mean-variance setting

The conceptually simple robust Bayesian allocation decision (9.133) cannot be computed in practice even under simple assumptions on preferences, markets and constraints. Therefore, it must be implemented within the two-step mean-variance framework, where the investor first determines a set of efficient allocations and then selects among those allocations the one that best suits him.

We assume that the investment constraints $\mathcal{C}$ do not depend on the unknown market parameters and are such that the inequality version (6.144) of the mean-variance problem applies. Furthermore, it is convenient to set up the mean-variance problem in terms of relative weights and linear returns, see Section 6.3.4.

With these settings the mean-variance problem can be written as follows:

$$\mathbf{w}^{(i)} = \underset{\mathbf{w}}{\operatorname{argmax}}\, \mathbf{w}'\boldsymbol{\mu} \tag{9.137}$$
$$\text{subject to } \begin{cases} \mathbf{w}\in\mathcal{C} \\ \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \leq v^{(i)}, \end{cases}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ represent the expected values and the covariances of the linear returns on the securities relative to the investment horizon:

$$\boldsymbol{\mu} \equiv \mathrm{E}\left\{\mathbf{L}_{T+\tau,\tau}\right\}, \quad \boldsymbol{\Sigma} \equiv \mathrm{Cov}\left\{\mathbf{L}_{T+\tau,\tau}\right\}; \tag{9.138}$$

and the set $\left\{v^{(1)}, \ldots, v^{(I)}\right\}$ is a significative grid of target variances of the return on the portfolio.

According to (9.133), the robust Bayesian version of the mean-variance problem (9.137) reads:

$$\mathbf{w}_{\text{rB}}^{(i)} = \underset{\mathbf{w}}{\operatorname{argmax}} \left\{ \min_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}} \{\mathbf{w}'\boldsymbol{\mu}\} \right\} \tag{9.139}$$

$$\text{subject to } \begin{cases} \mathbf{w} \in \mathcal{C} \\ \max_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\} \leq v^{(i)}, \end{cases}$$

where $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}$ are location-dispersion ellipsoids for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ respectively, defined in terms of the Bayesian posterior distribution of these parameters.

In order to specify the posterior distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ we make a few further assumptions, see also Meucci (2005): first, the market consists of equity-like securities for which the linear returns are market invariants, see Section 3.1.1; second, the investment horizon and the estimation interval coincide, see Section 6.5.4; third, the linear returns are normally distributed:

$$\mathbf{L}_{t,\tau} | \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{9.140}$$

Furthermore, we model the investor's prior experience as a normal-inverse-Wishart distribution:

$$\boldsymbol{\mu} | \boldsymbol{\Sigma} \sim \mathrm{N}\left(\boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}}{T_0}\right), \quad \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_0, \frac{\boldsymbol{\Sigma}_0^{-1}}{\nu_0}\right). \tag{9.141}$$

We recall from Section 7.2 that $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ represents the investor's experience on the parameters. On the other hand, (T_0, ν_0) represents the respective confidence. Therefore the investor's experience is summarized in:

$$e_C \equiv \{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0; T_0, \nu_0\}. \tag{9.142}$$

Under the above hypotheses it is possible to compute the posterior distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ analytically, see Section 7.2. The information from the market is summarized by the sample mean and the sample covariance of the past realizations of the linear returns, namely

$$\widehat{\boldsymbol{\mu}} \equiv \frac{1}{T} \sum_{t=1}^{T} \mathbf{l}_{t,\tau}, \quad \widehat{\boldsymbol{\Sigma}} \equiv \frac{1}{T} \sum_{t=1}^{T} (\mathbf{l}_{t,\tau} - \widehat{\boldsymbol{\mu}})(\mathbf{l}_{t,\tau} - \widehat{\boldsymbol{\mu}}), \tag{9.143}$$

plus the length of the time-series:

$$i_T \equiv \left\{\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}; T\right\}. \tag{9.144}$$

The posterior distribution, like the prior distribution (9.141), is also normal-inverse-Wishart, where the respective parameters read:

$$T_1 [i_T, e_C] \equiv T_0 + T \tag{9.145}$$

$$\boldsymbol{\mu}_1 [i_T, e_C] \equiv \frac{1}{T_1} [T_0 \boldsymbol{\mu}_0 + T \widehat{\boldsymbol{\mu}}] \tag{9.146}$$

$$\nu_1 [i_T, e_C] \equiv \nu_0 + T \tag{9.147}$$

$$\boldsymbol{\Sigma}_1 [i_T, e_C] \equiv \frac{1}{\nu_1} \left[\nu_0 \boldsymbol{\Sigma}_0 + T \widehat{\boldsymbol{\Sigma}} + \frac{(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}})(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}})'}{\frac{1}{T} + \frac{1}{T_0}} \right]. \tag{9.148}$$

The uncertainty set for $\boldsymbol{\mu}$ is the location-dispersion ellipsoid (7.37) of the marginal posterior distribution of $\boldsymbol{\mu}$:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{ \boldsymbol{\mu} : \ (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_{\mathrm{ce}})' \, \mathbf{S}_{\boldsymbol{\mu}}^{-1} \, (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_{\mathrm{ce}}) \leq q_{\boldsymbol{\mu}}^2 \right\}. \tag{9.149}$$

In this expression $q_{\boldsymbol{\mu}}$ is the radius factor that represents aversion to estimation risk for $\boldsymbol{\mu}$; $\widehat{\boldsymbol{\mu}}_{\mathrm{ce}}$ is the classical-equivalent estimator of $\boldsymbol{\mu}$, which from (7.35) reads explicitly:

$$\widehat{\boldsymbol{\mu}}_{\mathrm{ce}} \left[i_T, e_C \right] = \boldsymbol{\mu}_1; \tag{9.150}$$

and $\mathbf{S}_{\boldsymbol{\mu}}$ is the scatter matrix for $\boldsymbol{\mu}$, which from (7.36) reads explicitly:

$$\mathbf{S}_{\boldsymbol{\mu}} \left[i_T, e_C \right] = \frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} \boldsymbol{\Sigma}_1. \tag{9.151}$$

The uncertainty set for $\boldsymbol{\Sigma}$ is the location-dispersion ellipsoid (7.40) of the marginal posterior distribution of $\boldsymbol{\Sigma}$:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}} \equiv \left\{ \boldsymbol{\Sigma} : \ \operatorname{vech} \left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}} \right]' \mathbf{S}_{\boldsymbol{\Sigma}}^{-1} \operatorname{vech} \left[\boldsymbol{\Sigma} - \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}} \right] \leq q_{\boldsymbol{\Sigma}}^2 \right\}. \tag{9.152}$$

In this expression vech is the operator that stacks the columns of a matrix skipping the redundant entries above the diagonal; $q_{\boldsymbol{\Sigma}}$ is the radius factor that represents aversion to estimation risk for $\boldsymbol{\Sigma}$; $\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}$ is the classical-equivalent estimator of $\boldsymbol{\Sigma}$, which from (7.38) reads explicitly:

$$\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}} \left[i_T, e_C \right] = \frac{\nu_1}{\nu_1 + N + 1} \boldsymbol{\Sigma}_1; \tag{9.153}$$

and $\mathbf{S}_{\boldsymbol{\Sigma}}$ is the scatter matrix for vech $[\boldsymbol{\Sigma}]$. From (7.39) the scatter matrix reads explicitly as follows:

$$\mathbf{S}_{\boldsymbol{\Sigma}} \left[i_T, e_C \right] = \frac{2\nu_1^2}{(\nu_1 + N + 1)^3} \left(\mathbf{D}_N' \left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1} \right) \mathbf{D}_N \right)^{-1}, \tag{9.154}$$

where $\mathbf{D}_N$ is the duplication matrix (A.113) and $\otimes$ is the Kronecker product (A.95).

9.5.3 Discussion

In Appendix www.9.8 we show that the robust Bayesian mean-variance problem (9.139) with the robustness uncertainty sets specified as in (9.149) and (9.152) simplifies as follows:

$$\mathbf{w}_{\mathrm{rB}}^{(i)} = \underset{\mathbf{w}}{\operatorname{argmax}} \left\{ \mathbf{w}' \boldsymbol{\mu}_1 - \gamma_\mu \sqrt{\mathbf{w}' \boldsymbol{\Sigma}_1 \mathbf{w}} \right\} \tag{9.155}$$

$$\text{subject to } \begin{cases} \mathbf{w} \in \mathcal{C} \\ \mathbf{w}' \boldsymbol{\Sigma}_1 \mathbf{w} \leq \gamma_{\Sigma}^{(i)}, \end{cases}$$

where:

$$\gamma_{\mu} \equiv \sqrt{\frac{q_{\boldsymbol{\mu}}^{2}}{T_{1}} \frac{\nu_{1}}{\nu_{1}-2}} \tag{9.156}$$

$$\gamma_{\Sigma}^{(i)} \equiv \frac{v^{(i)}}{\frac{\nu_{1}}{\nu_{1}+N+1}+\sqrt{\frac{2 \nu_{1}^{2} q_{\boldsymbol{\Sigma}}^{2}}{\left(\nu_{1}+N+1\right)^{3}}}}. \tag{9.157}$$

This maximization is in the same form as the robust allocation decision (9.130). Like that problem, under regularity assumption for the constraints $\mathcal{C}$ also this maximization can be can be cast in the form of a second-order cone programming problem (6.55). Therefore the robust Bayesian frontier (9.155) can computed numerically.

The original robust Bayesian mean-variance problem (9.139) with the robustness uncertainty sets (9.149) and (9.152) is parametrized by the aversion to estimation risk for the expected values, represented by $q_{\boldsymbol{\mu}}$, the aversion to estimation risk for the covariances, represented by $q_{\boldsymbol{\Sigma}}$, and the exposure to market risk, represented by $v^{(i)}$. Therefore, in principle, the robust Bayesian mean-variance efficient frontier should constitute a three-dimensional surface in the N-dimensional space of the allocations.

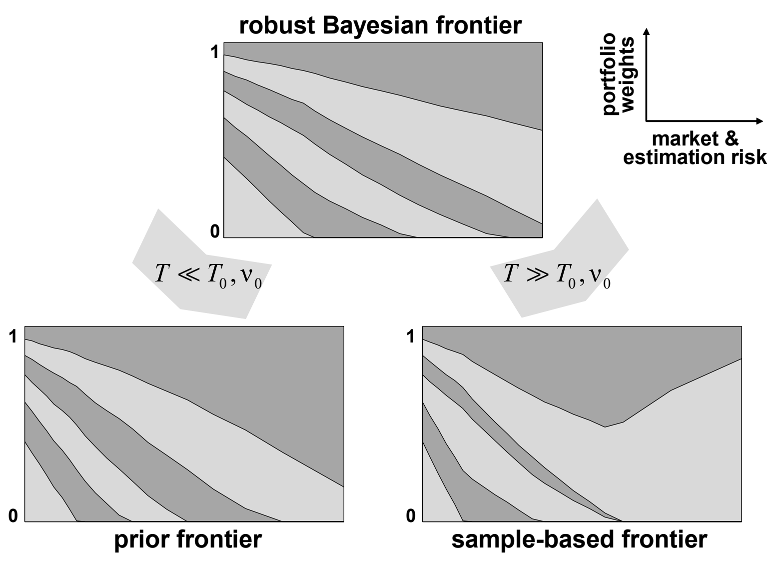

Fig. 9.10. Robust Bayesian mean-variance efficient allocations

Nevertheless, the efficient allocations (9.155) can be parametrized equivalently in terms of one single positive multiplier λ as follows:

$$\mathbf{w}_{\mathrm{rB}}(\lambda) = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmax}} \left\{ \mathbf{w}'\boldsymbol{\mu}_1 - \lambda\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_1\mathbf{w}} \right\}. \tag{9.158}$$

The multiplier λ is determined by the scalars (9.156) and (9.157). It is easy to check that the value of λ is directly related to the aversion to estimation risk $(q_{\boldsymbol{\mu}}, q_{\boldsymbol{\Sigma}})$ and inversely related to the exposure to market risk $v^{(i)}$. Accordingly, the term under the square root in (9.158) represents both estimation and market risk and the coefficient λ represents aversion to both types of risk.

In other words, the a-priori three-dimensional robust Bayesian efficient frontier collapses to a line. Hence the robust Bayesian mean-variance efficient frontier is conceptually similar to, and just as parsimonious as, the classical mean-variance efficient frontier (9.137). Nevertheless, in the classical setting the coefficient of risk aversion only refers to market risk, whereas in the robust Bayesian setting the coefficient of risk aversion blends aversion to both market risk and estimation risk.

From (9.145)-(9.148) the expected values $\boldsymbol{\mu}_1$ and the covariance matrix $\boldsymbol{\Sigma}_1$ in (9.158) are self-adjusting mixtures of the classical estimators $\left(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}\right)$ and of the prior parameters $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. In particular, when the number of observations T is large with respect to the confidence levels T_0 and ν_0 in the investor's prior, the expected values $\boldsymbol{\mu}_1$ tend to the sample mean $\widehat{\boldsymbol{\mu}}$ and the covariance matrix $\boldsymbol{\Sigma}_1$ tends to the sample covariance $\widehat{\boldsymbol{\Sigma}}$. Therefore we obtain a sample-based efficient frontier:

$$\mathbf{w}_{\mathrm{s}}(\lambda) = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmax}} \left\{ \mathbf{w}'\widehat{\boldsymbol{\mu}} - \lambda\sqrt{\mathbf{w}'\widehat{\boldsymbol{\Sigma}}\mathbf{w}} \right\}. \tag{9.159}$$

Similarly, when the confidence levels T_0 and ν_0 in the investor's prior are large with respect to the number of observations T, the expected values $\boldsymbol{\mu}_1$ tend to the prior $\boldsymbol{\mu}_0$ and the covariance matrix $\boldsymbol{\Sigma}_1$ tends to the prior $\boldsymbol{\Sigma}_0$. Therefore we obtain a prior efficient frontier that disregards any information from the market:

$$\mathbf{w}_{\mathrm{p}}(\lambda) = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmax}} \left\{ \mathbf{w}'\boldsymbol{\mu}_0 - \lambda\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_0\mathbf{w}} \right\}. \tag{9.160}$$

Consider a market of $N \equiv 6$ stocks from the utilities sector of the S&P 500. We estimate the sample mean and covariance from a database of weekly returns. We specify the prior with an equilibrium argument, as in (8.58)-(8.59), where we assume a correlation of 0.5.

Suppose that the investor is bound by the standard budget constraint $\mathbf{w}'\mathbf{1} = 1$ and the standard no-short-sale constraint $\mathbf{w} \geq \mathbf{0}$.

In Figure 9.10 we plot the general robust Bayesian efficient frontier (9.158) and the limit cases (9.159) and (9.160), refer to `symmys.com` for more details.