Robust Bayesian Allocation

Attilio Meucci[1]
attilio_meucci@symmys.com

this version: May 12, 2011
latest revision of article and code at `http://symmys.com/node/102`

Abstract

Using the Bayesian posterior distribution of the market parameters we define self-adjusting uncertainty regions for the robust mean-variance problem. Under a normal-inverse-Wishart conjugate assumption for the market, the ensuing robust Bayesian mean-variance optimal portfolios are shrunk by the aversion to estimation risk toward the global minimum variance portfolio.

After discussing the theory, we test robust Bayesian allocations in a simulation study and in an application to the management of sectors of the S&P 500.

Fully commented code is available at http://symmys.com/node/102

JEL Classification: C1, G11

Keywords: estimation risk, Bayesian estimation, MCMC, robust optimization, location-dispersion ellipsoid, classical equivalent, shrinkage, global minimum variance portfolio, equally-weighted portfolio, quantitative portfolio management, asset allocation

[1] The author is grateful to Ralf Werner and Jordi Serra

1 Introduction

The classical approach to asset allocation is a two-step process: first the market distribution is estimated, then an optimization is performed, as if the estimated distribution were the true market distribution. Since this is not the case, the classical "optimal" allocation is not truly optimal. More importantly, since the optimization process is extremely sensitive to the input parameters, the sub-optimality due to estimation risk can be dramatic, see Jobson and Korkie (1980), Best and Grauer (1991), Chopra and Ziemba (1993).

Bayesian theory provides a way to limit the sensitivity of the final allocation to the input parameters by shrinking the estimate of the market parameters toward the investor's prior, see Bawa, Brown, and Klein (1979), Jorion (1986), Pastor and Stambaugh (2002).

Similarly, the approach of Black and Litterman (1990) uses Bayes' rule to shrink the general equilibrium distribution of the market toward the investor's views.

The theory of robust optimization provides a different approach to dealing with estimation risk: the investor chooses the best allocation in the worst market within a given uncertainty range, see Goldfarb and Iyengar (2003), Halldorsson and Tutuncu (2003), Ceria and Stubbs (2004).

Robust allocations are guaranteed to perform adequately for all the markets within the given uncertainty range. Nevertheless, the choice of this range is quite arbitrary. Furthermore, the investor's prior knowledge, a key ingredient in any allocation decision, is not taken in consideration.

Using the Bayesian approach to estimation we can naturally identify a suitable uncertainty range for the market parameters, namely the location-dispersion ellipsoid of their posterior distribution. Robust Bayesian allocations are the solutions to a robust optimization problem that uses as uncertainty range the Bayesian location-dispersion ellipsoid. Similarly to robust allocations, these allocations account for estimation risk over a whole range of market parameters. Similarly to Bayesian decisions, these allocations include the investor's prior knowledge in the optimization process within a sound and self-adjusting statistical framework.

Robust Bayesian allocations are discussed in De Santis and Foresi (2002), where the Bayesian setting is provided by the Black-Litterman posterior distribution. Here we consider robust Bayesian decisions that also account for the estimation error in the covariances and that, unlike in the Black-Litterman framework, explicitly process the information from the market, namely the observed time series of the past returns. As it turns out, the multi-parameter, non-conically constrained mean-variance optimization simplifies to a parsimonious Bayesian efficient frontier that resembles the classical frontier, except that the classical parameterization in terms of the exposure to market risk becomes in this context a parameterization in terms of the exposure to both market risk and estimation risk.

In Section 2 we introduce the general robust Bayesian mean-variance approach to asset allocation. In Section 3 we compute explicitly the Bayesian location-dispersion ellipsoids for the market parameters under the standard normal-inverse-Wishart conjugate assumption for the market. In Section 4 we solve the ensuing mean-variance problem and we compute the robust Bayesian optimal portfolios. In this section we also perform two empirical tests, a simulation-based test and a back-test on portfolios of sectors of the S&P 500. Section 6 concludes.

An appendix details all the technicalities. Fully commented code is available at http://symmys.com/node/102

2 General robust Bayesian mean-variance framework

The classical mean-variance problem reads:

$$\mathbf{w}^{(i)} = \underset{\mathbf{w}}{\operatorname{argmax}}\, \mathbf{w}'\boldsymbol{\mu} \quad \text{subject to } \begin{cases} \mathbf{w} \in \mathcal{C} \\ \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} \leq v^{(i)}. \end{cases} \tag{1}$$

In this expression $\mathbf{w}$ is the N-dimensional vector of relative portfolio weights; $\mathcal{C}$ is a set of investment constraints; the set $\left\{v^{(1)}, \ldots, v^{(I)}\right\}$ is a significative grid of target variances of the return on the portfolio, where the return at time t for a horizon τ of an asset that at the generic time t trades at the price P_t is defined as $R_{t,\tau} \equiv P_t/P_{t-\tau} - 1$; and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ represent respectively the expected values and the covariances of the returns on the N securities in the market relative to the investment horizon:

$$\boldsymbol{\mu} \equiv \mathrm{E}\left\{\mathbf{R}_{T+\tau,\tau}\right\}, \quad \boldsymbol{\Sigma} \equiv \mathrm{Cov}\left\{\mathbf{R}_{T+\tau,\tau}\right\}. \tag{2}$$

The robust version of the mean-variance problem (1) reads:

$$\mathbf{w}^{(i)} = \underset{\mathbf{w}}{\operatorname{argmax}} \left\{ \min_{\boldsymbol{\mu} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}} \left\{\mathbf{w}'\boldsymbol{\mu}\right\} \right\} \quad \text{subject to } \begin{cases} \mathbf{w} \in \mathcal{C} \\ \max\limits_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}} \left\{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\right\} \leq v^{(i)}, \end{cases} \tag{3}$$

where $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}$ are suitable uncertainty regions for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ respectively.

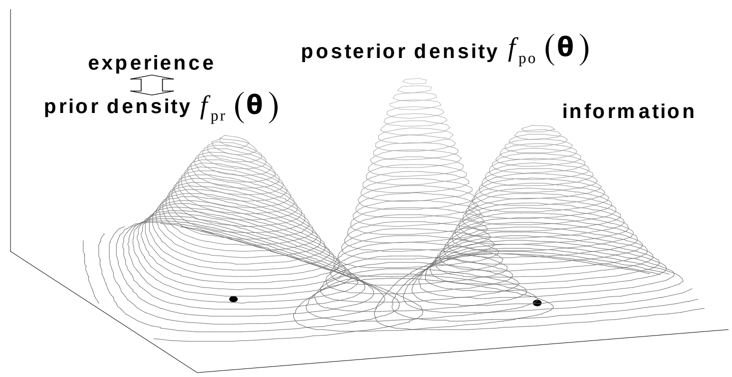

Figure 1: Bayesian approach to parameter estimation

The Bayesian framework defines uncertainty sets in a natural way. Indeed, in the Bayesian framework the unknown generic market parameters $\boldsymbol{\theta}$ (e.g. $\boldsymbol{\mu}$ or $\boldsymbol{\Sigma}$) are random variables. The likelihood that the parameters assume given values is described by the posterior probability density function $f_{\mathrm{po}}(\boldsymbol{\theta})$, which is determined by the information available at the time the allocation takes place and by the investor's experience and respective confidence, modeled in terms of a prior probability density function $f_{\mathrm{pr}}(\boldsymbol{\theta})$, see Figure 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, which in the Bayesian literature is known as credibility set.

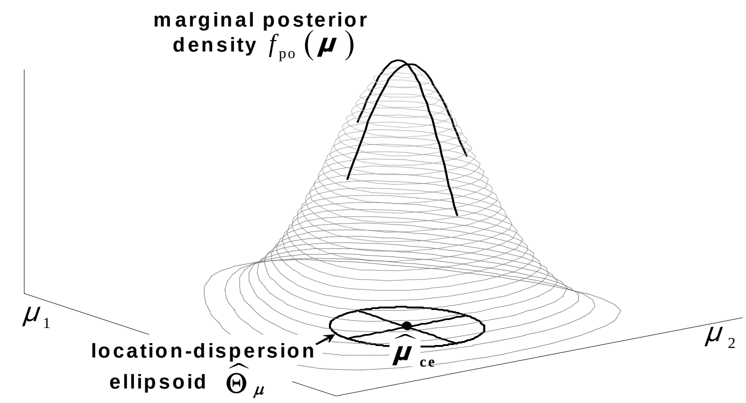

Figure 2: Bayesian posterior distribution and uncertainty set

Consider first $\boldsymbol{\mu}$. A credibility set, i.e. a region where the posterior distribution displays a higher concentration, can be represented by the location-dispersion ellipsoid of the marginal posterior distribution of $\boldsymbol{\mu}$, see Figure 2 for the case $N \equiv 2$:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{ \boldsymbol{\mu} : \ (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_{\text{ce}})' \, \mathbf{S}_{\boldsymbol{\mu}}^{-1} \, (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}_{\text{ce}}) \leq q_{\boldsymbol{\mu}}^2 \right\}. \tag{4}$$

In this expression $q_{\boldsymbol{\mu}}$ is the radius factor for the ellipsoid; $\widehat{\boldsymbol{\mu}}_{\text{ce}}$ is a classical-equivalent estimator such as the expected value or the mode of the marginal posterior distribution of $\boldsymbol{\mu}$; and $\mathbf{S}_{\boldsymbol{\mu}}$ is a scatter matrix such as the covariance matrix or the modal dispersion of the marginal posterior distribution of $\boldsymbol{\mu}$.

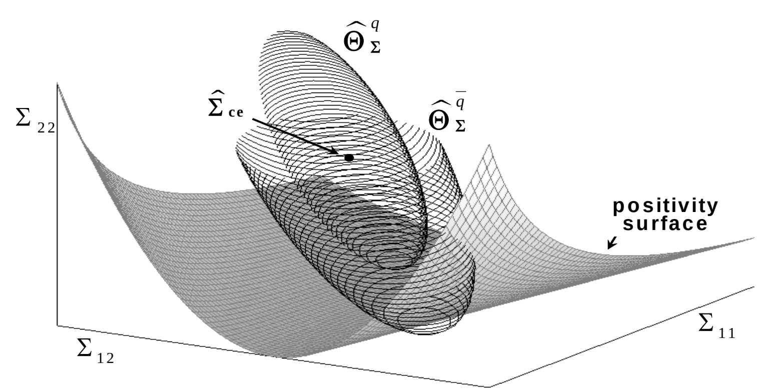

Figure 3: Bayesian location-dispersion ellipsoid for covariance

Similarly, consider $\mathbf{\Sigma}$. A credibility set, i.e. a region where the posterior distribution displays a higher concentration, can be represented by the location-dispersion ellipsoid of the marginal posterior distribution of $\mathbf{\Sigma}$:

$$\widehat{\mathbf{\Theta}}_{\mathbf{\Sigma}} \equiv \left\{ \mathbf{\Sigma} : \ \operatorname{vech}\left[\mathbf{\Sigma} - \widehat{\mathbf{\Sigma}}_{\mathrm{ce}}\right]' \mathbf{S}_{\mathbf{\Sigma}}^{-1} \operatorname{vech}\left[\mathbf{\Sigma} - \widehat{\mathbf{\Sigma}}_{\mathrm{ce}}\right] \leq q_{\mathbf{\Sigma}}^{2} \right\}. \tag{5}$$

In this expression vech is the operator that stacks the columns of a matrix skipping the redundant entries above the diagonal; $q_{\mathbf{\Sigma}}$ is the radius factor for the ellipsoid; $\widehat{\mathbf{\Sigma}}_{\mathrm{ce}}$ is a classical-equivalent estimator such as the expected value or the mode of the marginal posterior distribution of $\mathbf{\Sigma}$; and $\mathbf{S}_{\boldsymbol{\mu}}$ is a scatter matrix such as the covariance matrix or the modal dispersion of the marginal posterior distribution of vech$[\mathbf{\Sigma}]$. The matrices $\mathbf{\Sigma}$ in this ellipsoid are always symmetric, because the vech operator only spans the non-redundant elements of a matrix. When the radius factor $q_{\mathbf{\Sigma}}$ is not too large the matrices $\mathbf{\Sigma}$ in this ellipsoid are also positive definite: indeed, positivity is a continuous property and $\widehat{\mathbf{\Sigma}}_{\mathrm{ce}}$ is positive definite, see Figure 3 for the case $N \equiv 2$, which is completely determined by three entries.

The robust Bayesian mean-variance approach to allocation consists in using the Bayesian elliptical uncertainty sets (4) and (5) in the robust mean-variance allocation problem (3). Notice that this problem is parametrized by the radius factor $q_{\boldsymbol{\mu}}$, the radius factor $q_{\mathbf{\Sigma}}$ and the target variance $v^{(i)}$.

The term $q_{\boldsymbol{\mu}}$ represents aversion to estimation risk for the expected values $\boldsymbol{\mu}$: considering a larger ellipsoid increases the chances of including the true, unknown value of $\boldsymbol{\mu}$ within the ellipsoid. The value of $q_{\boldsymbol{\mu}}$ is typically set according to the quantile of the chi-square distribution with N degrees of freedom:

$$q_{\boldsymbol{\mu}}^{2} \equiv Q_{\chi_{N}^{2}}\left(p_{\boldsymbol{\mu}}\right). \tag{6}$$

Indeed, under the normal hypothesis for the posterior of $\boldsymbol{\mu}$ the square Mahalanobis distance in (4) is chi-square distributed. A confidence level $p_{\boldsymbol{\mu}} \equiv 5\%$ corresponds to an investor who is not too concerned with poorly estimating $\boldsymbol{\mu}$; a confidence level $p_{\boldsymbol{\mu}} \equiv 25\%$ corresponds to a conservative investor who is very concerned with properly estimating $\boldsymbol{\mu}$.

Similarly, the term $q_{\mathbf{\Sigma}}$ represents aversion to estimation risk for the covariances $\mathbf{\Sigma}$: considering a larger ellipsoid increases the chances of including the true, unknown value of $\mathbf{\Sigma}$ within the ellipsoid. The value of $q_{\mathbf{\Sigma}}$ can be set according to the quantile of the chi-square distribution with $N(N+1)/2$ degrees of freedom:

$$q_{\mathbf{\Sigma}}^{2} \equiv Q_{\chi_{N(N+1)/2}^{2}}\left(p_{\mathbf{\Sigma}}\right). \tag{7}$$

This follows from a heuristic argument. Indeed, for a ν-dimensional random variable $\mathbf{Z}$ with fully arbitrary distribution the following identity for the Mahalanobis distance holds

$$M \equiv (\mathbf{Z} - \mathrm{E}\{\mathbf{Z}\})' \operatorname{Cov}\{\mathbf{Z}\}^{-1} (\mathbf{Z} - \mathrm{E}\{\mathbf{Z}\})' \equiv \sum_{i=1}^{\nu} Y_i^2, \tag{8}$$

where $\mathbf{Y} \equiv \mathbf{C}^{-1} (\mathbf{Z} - \mathrm{E}\{\mathbf{Z}\})$ and $\mathbf{C}$ is the Cholesky decomposition of the covariance $\mathrm{Cov}\{\mathbf{Z}\} \equiv \mathbf{C}\mathbf{C}'$. It is easy to check that $\mathrm{E}\{Y_i\} = 0$, $\mathrm{Var}\{Y_i\} = 1$. Furthermore, all the Y_i are uncorrelated, although they are not necessarily independent of each other. Therefore we expect M to behave similarly to a chi-square distribution with ν degrees of freedom, especially when $\nu = N(N+1)/2$ is rather large. A confidence level $p_{\boldsymbol{\Sigma}} \equiv 5\%$ corresponds to an investor who is not too concerned with poorly estimating $\boldsymbol{\Sigma}$; a confidence level $p_{\boldsymbol{\Sigma}} \equiv 25\%$ corresponds to a conservative investor who is very concerned with properly estimating $\boldsymbol{\Sigma}$.

Finally, the term $v^{(i)}$ represents exposure to market risk. As a result, in principle, the robust Bayesian mean-variance efficient frontier should constitute a three-dimensional surface in the N-dimensional space of the allocations, parametrized by $q_{\boldsymbol{\mu}}$, $q_{\boldsymbol{\Sigma}}$ and v.

3 Specification of a market model

In order to determine the Bayesian elliptical uncertainty sets (4) and (5) we need to compute the posterior distributions of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. To do so, we need to make parametric assumptions on the market dynamics and model our prior knowledge of the parameters that steer the market dynamics. In this section we consider "conjugate" assumptions for the dynamics and the prior which give rise to analytical expressions for the posterior distribution of the market.

We make the following assumptions: first, the market consists of equity-like securities for which the returns are independently and identically distributed across time; second, the estimation interval is the same as the investment horizon; third, the returns are normally distributed:

$$\mathbf{R}_{t,\tau} | \boldsymbol{\mu}, \boldsymbol{\Sigma} \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{9}$$

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

$$\boldsymbol{\mu} | \boldsymbol{\Sigma} \sim \mathrm{N}\left(\boldsymbol{\mu}_0, \frac{\boldsymbol{\Sigma}}{T_0}\right), \quad \boldsymbol{\Sigma}^{-1} \sim \mathrm{W}\left(\nu_0, \frac{\boldsymbol{\Sigma}_0^{-1}}{\nu_0}\right). \tag{10}$$

In this expression $(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ represent the investor's experience on the parameters, whereas (T_0, ν_0) represent the respective confidence.

Under the above hypotheses it is possible to compute the posterior distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ analytically, see Aitchison and Dunsmore (1975) or Meucci (2005) for all the computations. First of all, the information from the market is summarized in the sample mean and the sample covariance of the past realizations of the returns:

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

The posterior distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, like the prior distribution (10), is also normal-inverse-Wishart, where the respective parameters read:

$$\begin{aligned} T_1 &\equiv T_0 + T && (11) \\ \boldsymbol{\mu}_1 &\equiv \frac{1}{T_1}\left[T_0\boldsymbol{\mu}_0 + T\widehat{\boldsymbol{\mu}}\right] && (12) \\ \nu_1 &\equiv \nu_0 + T && (13) \\ \boldsymbol{\Sigma}_1 &\equiv \frac{1}{\nu_1}\left[\nu_0\boldsymbol{\Sigma}_0 + T\widehat{\boldsymbol{\Sigma}} + \frac{(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}})(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}})'}{\frac{1}{T} + \frac{1}{T_0}}\right]. && (14) \end{aligned}$$

The marginal posterior distribution of $\boldsymbol{\mu}$ is multivariate Student t. From its expression it is possible to compute explicitly the classical-equivalent estimator and the scatter matrix that appear in the location-dispersion ellipsoid (4), see Meucci (2005):

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_{\mathrm{ce}} &= \boldsymbol{\mu}_1 && (15) \\ \mathbf{S}_{\boldsymbol{\mu}} &= \frac{1}{T_1}\frac{\nu_1}{\nu_1 - 2}\boldsymbol{\Sigma}_1. && (16) \end{aligned}$$

It is also possible to compute explicitly the classical-equivalent estimator and the scatter matrix of the inverse-Wishart marginal posterior distribution of $\boldsymbol{\Sigma}$ that appear in the location-dispersion ellipsoid (5), see Meucci (2005):

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}} &= \frac{\nu_1}{\nu_1 + N + 1}\boldsymbol{\Sigma}_1 && (17) \\ \mathbf{S}_{\boldsymbol{\Sigma}} &= \frac{2\nu_1^2}{(\nu_1 + N + 1)^3}\left(\mathbf{D}_N'\left(\boldsymbol{\Sigma}_1^{-1} \otimes \boldsymbol{\Sigma}_1^{-1}\right)\mathbf{D}_N\right)^{-1}. && (18) \end{aligned}$$

In this expression $\mathbf{D}_N$ is the duplication matrix that reinstates the redundant entries above the diagonal of a symmetric matrix, see Magnus and Neudecker (1999), and $\otimes$ is the Kronecker product.

We stress that the above specifications are flexible enough to describe "simple" priors on "simple" markets. However, as far as the views are concerned, the normal-inverse-Wishart assumption (10) prevents the specification of different confidence levels on correlations and standard deviations. As far as the market is concerned, the conditionally normal i.i.d. assumption (9) is sufficient to model the returns on asset classes such as stock sectors or mutual funds at medium investment horizons. At short horizons, e.g. intra-day, or with different asset classes, e.g. individual stocks or hedge funds, fat tails and market asymmetries play a major role. Furthermore, independence across time is no longer valid, as the effect of volatility clustering becomes important.

In these more complex situations where the conjugate normal-inverse Wishart assumption is not suitable to model the market dynamics and the views, the posterior distribution of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and the location-dispersion ellipsoids must be computed numerically by means of Markov chain Monte Carlo techniques, see e.g. Geweke (2005).

4 Optimal portfolios

In the technical appendix we prove that the robust Bayesian mean-variance problem ensuing from the specifications (9)-(10), i.e. the robust problem (3) where the Bayesian elliptical uncertainty sets (4) and (5) are specified by (15)-(18), simplifies as follows:

$$\mathbf{w}_{\mathrm{rB}}^{(i)} = \underset{\substack{\mathbf{w}\in\mathcal{C} \\ \mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}\leq\gamma_{\Sigma}^{(i)}}}{\operatorname{argmax}} \left\{\mathbf{w}'\boldsymbol{\mu}_1 - \gamma_\mu\sqrt{\mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}}\right\} \tag{19}$$

where:

$$\gamma_\mu \equiv \sqrt{\frac{q_{\boldsymbol{\mu}}^2}{T_1}\frac{\nu_1}{\nu_1-2}} \tag{20}$$

$$\gamma_{\Sigma}^{(i)} \equiv \frac{v^{(i)}}{\frac{\nu_1}{\nu_1+N+1}+\sqrt{\frac{2\nu_1^2 q_{\mathbf{\Sigma}}^2}{(\nu_1+N+1)^3}}}. \tag{21}$$

[2]Under standard regularity assumptions for the investment constraints $\mathcal{C}$ the maximization (19) can be can be cast in the form of a second-order cone programming problem. Therefore the robust Bayesian frontier can be computed numerically, see Ben-Tal and Nemirovski (2001).

Alternatively, the above result shows that the generic three-dimensional robust Bayesian efficient frontier degenerates to a one-parameter family, which can be parametrized by a single positive multiplier λ as follows:

$$\mathbf{w}_{\mathrm{rB}}^{(i)} \in \mathbf{w}(\lambda) = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmax}} \left\{\mathbf{w}'\boldsymbol{\mu}_1 - \lambda\sqrt{\mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}}\right\}. \tag{22}$$

In other words, the a-priori three-dimensional robust Bayesian efficient frontier collapses to a line. We stress that in markets other than (9) and with prior structures more general than (10) the robust Bayesian efficient frontier, which must be computed by means of Markov chain Monte Carlo techniques, remains three-dimensional.

Notice the self-adjusting nature of the Bayesian setting. From (11)-(14) the expected values $\boldsymbol{\mu}_1$ and the covariance matrix $\mathbf{\Sigma}_1$ that determine the efficient allocations (22) are mixtures of the investor's prior knowledge $(\boldsymbol{\mu}_0, \mathbf{\Sigma}_0)$ and of information from the market $\left(\widehat{\boldsymbol{\mu}}, \widehat{\mathbf{\Sigma}}\right)$: the balance between these components is steered by the relative weight of the confidence in the investor's prior parameters, represented respectively by (T_0, ν_0), and the amount of information, represented by number of observations T in the time series.

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 $\boldsymbol{\mu}_1$

[2]We do not need to worry about the positivity condition in Figure 3, because under our assumptions the max of $\mathbf{\Sigma}$ in (3) always lies in the positivity region.

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:

$$\mathbf{w}(\lambda) = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmax}} \left\{ \mathbf{w}'\widehat{\boldsymbol{\mu}} - \lambda\sqrt{\mathbf{w}'\widehat{\boldsymbol{\Sigma}}\mathbf{w}} \right\}. \tag{23}$$

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:

$$\mathbf{w}(\lambda) = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmax}} \left\{ \mathbf{w}'\boldsymbol{\mu}_0 - \lambda\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_0\mathbf{w}} \right\}. \tag{24}$$

The robust Bayesian frontier (22), as well as its limit cases (23) and (24), is apparently purely Bayesian, i.e. non-robust. However, estimation risk enters the picture through the multiplier λ, which is determined by the scalars (20) and (21). It is easy to check that the value of λ is directly related to the aversion to estimation risk for $\boldsymbol{\mu}$, namely $q_{\boldsymbol{\mu}}$, and to the aversion to estimation risk for $\boldsymbol{\Sigma}$, namely $q_{\boldsymbol{\Sigma}}$, and inversely related to the exposure to market risk $v^{(i)}$.

Accordingly, the term under the square root in (22) represents both estimation risk and market risk and the coefficient λ represents aversion to both types of risk. In the classical or purely Bayesian setting "risk" only refers to market risk, whereas in the robust Bayesian setting "risk" blends both market risk and estimation risk. As the aversion to estimation risk or market risk increases, investors ride down the frontier and shrink their allocation towards the global minimum variance portfolio:

$$\mathbf{w}_{MV} = \underset{\mathbf{w}\in\mathcal{C}}{\operatorname{argmin}} \left\{ \mathbf{w}'\boldsymbol{\Sigma}_1\mathbf{w} \right\}. \tag{25}$$

The minimum-variance portfolio plays a special role in the shrinkage literature, see Jagannathan and Ma (2003). Shrinking to the global minimum variance portfolio in order to diminish estimation risk is also explored in Kan and Zhou (2006), where this choice is imposed exogenously.

5 Empirical results

In this section we run two empirical tests. In both cases we assume that the investor is bound by the standard budget constraint $\mathbf{w}'\mathbf{1} = 1$ and no-short-sale constraint $\mathbf{w} \geq \mathbf{0}$. Fully commented code is available at http://symmys.com/node/102

The first test is a simulation study: we consider an artificial normal market and we evaluate the robust Bayesian allocations in the true mean-variance plane of coordinates, i.e. the plane of coordinates that can be computed when the true parameters underlying the market distribution are known. We assume $N \equiv 40$ asset classes and $T \equiv 52$ observations in the time series. We factor the covariance matrix $\mathbf{\Sigma}$ into a homogeneous correlation matrix

$$\mathbf{C} \equiv \begin{pmatrix} 1 & \theta & & \cdots & \theta \\ \theta & \ddots & & & \vdots \\ \vdots & & & \ddots & \theta \\ \theta & \cdots & & \theta & 1 \end{pmatrix},$$

where $\theta \equiv 0.7$, and equally-spaced volatilities:

$$\underline{\sigma} \equiv 0.1, \underline{\sigma} + \Delta, \ldots, \overline{\sigma} - \Delta, \overline{\sigma} \equiv 0.4.$$

We define the expected returns $\boldsymbol{\mu}$ in such a way that a mean-variance optimization would yield an equally-weighed portfolio:

$$\boldsymbol{\mu} \equiv 2.5 \times \mathbf{\Sigma} \frac{\mathbf{1}}{N}.$$

We define the prior as follows: $\mathbf{\Sigma}_0$ is equal to the sample covariance on the principal diagonal and is null otherwise; as for the prior expected returns we set

$$\boldsymbol{\mu}_0 \equiv 0.5 \times \mathbf{\Sigma}_0 \frac{\mathbf{1}}{N};$$

We set the confidence in the prior as twice the confidence in the sample: $T_0 \equiv \nu_0 \equiv 2T$. We set the aversion to estimation risk in (20)-(21) as in (6) and (7), where $p_{\boldsymbol{\mu}} \equiv p_{\mathbf{\Sigma}} \equiv 0.1$. We set the target variance in (21) as the sample variance of an aggressive portfolio whose sample expected return is 4/5 of the maximum sample expected return.

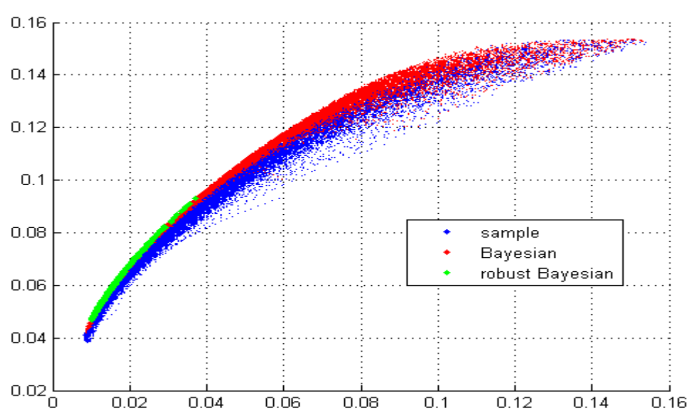

Figure 4: Shrinkage effect of robust Bayesian decisions

We run a Monte Carlo experiment and we report in Figure 4 the results. The sample-based allocation is broadly scattered in inefficient regions. The purely

Bayesian portfolios are shrunk toward the prior: therefore they are less scattered and more efficient, although the prior differs significantly from the true market parameters. The robust-Bayesian allocations is a subset of the Bayesian frontier that is further shrunk toward the global minimum variance portfolio and even more closely tight to the right of the efficient frontier.

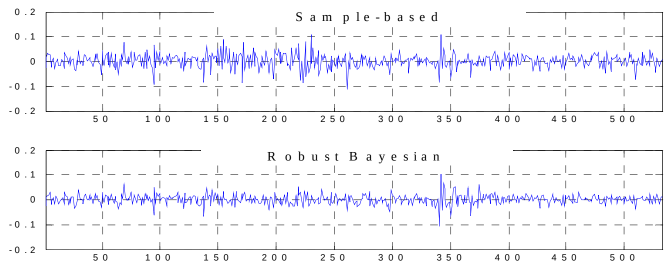

Figure 5: Weekly returns of sample-based and robust Bayesian allocations

In the second experiment we compute optimal robust Bayesian portfolios of sectors of the S&P 500. We consider an investment horizon of one week and thus we base our estimates on weekly returns. All the assumptions are the same as in the simulation tests, except of course that the true market parameters are unknown. Similarly to the extensive analysis of DeMiguel, Garlappi, and Uppal (2005) we consider rolling estimates of the market parameters over a period of one year. Then we compare the ensuing robust Bayesian allocation with the purely sample-based allocation.

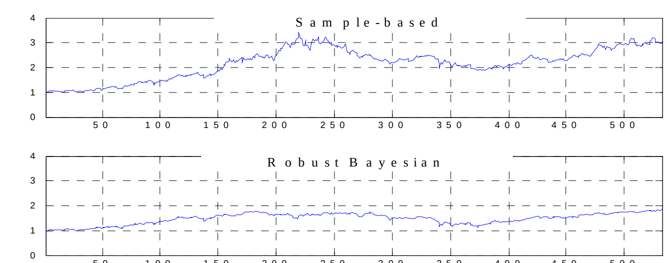

Figure 6: Cumulative P&L of sample-based and robust Bayesian allocations

In Figure 5 we plot the weekly returns of the two strategies and in Figure 6 the respective cumulative P&L. As expected, the robust Bayesian allocation is more conservative. Although the final gain of the sample-based allocation is larger, its maximum drawdown largely exceed that of the robust Bayesian strategy. We stress that this example is one of the least favorable to the robust Bayesian approach. In other cases, such as using more than one year of data in the rolling estimation of the market parameters, the robust Bayesian strategy clearly dominates the sample-based allocation, which incur substantial losses.

6 Conclusions

The robust Bayesian approach to allocation displays the optimality features of robust optimization as well as the self-adjusting Bayesian mechanisms that account for the investor's prior knowledge within a sound statistical framework.

Under fairly standard assumptions for the market, the robust Bayesian mean-variance optimal portfolios reduce to a purely Bayesian optimal portfolio which is shrunk toward the Bayesian minimum variance portfolio by the aversion to estimation risk.

As customary in the Bayesian framework, the above Bayesian portfolios are a mix of purely sample-based allocations, which completely disregard the investor's prior knowledge, and purely prior allocations, which completely disregard the information from the market. The interplay between these two components is steered by the relative weight of the confidence in the prior with respect to the amount of information from the market.

Please refer to the code available at http://symmys.com/node/102 for more details.

References

Aitchison, J., and I. R. Dunsmore, 1975, *Statistical Prediction Analysis* (Cambridge University Press).

Bawa, V. S., S. J. Brown, and R. W. Klein, 1979, *Estimation Risk and Optimal Porfolio Choice* (North Holland).

Ben-Tal, A., and A. Nemirovski, 2001, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications* (Society for Industrial and Applied Mathematics).

Best, M. J., and R. R. Grauer, 1991, On the sensitivity of mean-variance-efficient portfolios to changes in asset means: Some analytical and computational results, *Review of Financial Studies* 4, 315–342.

Black, F., and R. Litterman, 1990, Asset allocation: combining investor views with market equilibrium, *Goldman Sachs Fixed Income Research*.

Ceria, S., and R. A. Stubbs, 2004, Incorporating estimation errors into portfolio selection: Robust efficient frontiers, *Axioma Inc. Technical Report*.

Chopra, V., and W. T. Ziemba, 1993, The effects of errors in means, variances, and covariances on optimal portfolio choice, *Journal of Portfolio Management* pp. 6–11.

De Santis, G., and S. Foresi, 2002, Robust optimization, *Goldman Sachs Technical Report*.

DeMiguel, V., L. Garlappi, and R. Uppal, 2005, How inefficient is the 1/n asset-allocation strategy?, *Working Paper*.

Geweke, J., 2005, *Contemporary Bayesian Econometrics and Statistics* (Wiley).

Goldfarb, D., and G. Iyengar, 2003, Robust portfolio selection problems, *Mathematics of Operations Research* 28, 1–38.

Halldorsson, B. V., and R. H. Tutuncu, 2003, An interior-point method for a class of saddle-point problems, *Journal of Optimization Theory and Applications* 116, 559–590.

Jagannathan, R., and T. Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1683.

Jobson, J. D., and B. Korkie, 1980, Estimation for Markowitz efficient portfolios, *Journal of the American Statistical Association* 75, 544–554.

Jorion, P., 1986, Bayes-Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis* 21, 279–291.

Kan, R., and G. Zhou, 2006, Optimal estimation for economic gains: Portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis*.

Magnus, J. R., and H. Neudecker, 1999, *Matrix Differential Calculus with Applications in Statistics and Econometrics, Revised Edition* (Wiley).

Meucci, A., 2005, *Risk and Asset Allocation* (Springer) Available at http://symmys.com.

Pastor, L., and R. F. Stambaugh, 2002, Investing in equity mutual funds, *Journal of Financial Economics* 63, 351–380.

7 Technical appendix

In this appendix we show how to simplify the max-min expression for the efficient frontier and how to reduce the problem in the SOCP form.

7.1 Worst-case scenario for $\boldsymbol{\mu}$

Consider the ellipsoid (4) under the specifications (15)-(16):

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{ \boldsymbol{\mu} : \ (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \, \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_1) \leq \frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right\}. \tag{26}$$

Consider the spectral decomposition of the dispersion parameter:

$$\boldsymbol{\Sigma}_1 \equiv \mathbf{F} \boldsymbol{\Gamma}^{1/2} \boldsymbol{\Gamma}^{1/2} \mathbf{F}', \tag{27}$$

where $\boldsymbol{\Gamma}$ is the diagonal matrix of the eigenvalues sorted in decreasing order and $\mathbf{F}$ is the juxtaposition of the respective eigenvectors.

We can write (26) as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{ \boldsymbol{\mu} : (\boldsymbol{\mu} - \boldsymbol{\mu}_1)' \, \mathbf{F} \boldsymbol{\Gamma}^{-1/2} \boldsymbol{\Gamma}^{-1/2} \mathbf{F}' (\boldsymbol{\mu} - \boldsymbol{\mu}_1) \leq \frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right\}. \tag{28}$$

Define the new variable:

$$\mathbf{u} \equiv \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{-1/2} \boldsymbol{\Gamma}^{-1/2} \mathbf{F}' (\boldsymbol{\mu} - \boldsymbol{\mu}_1), \tag{29}$$

which implies

$$\boldsymbol{\mu} = \boldsymbol{\mu}_1 + \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{1/2} \mathbf{F} \boldsymbol{\Gamma}^{1/2} \mathbf{u}. \tag{30}$$

We can write (28) as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}} \equiv \left\{ \boldsymbol{\mu}_1 + \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{1/2} \mathbf{F} \boldsymbol{\Gamma}^{1/2} \mathbf{u}, \quad \mathbf{u}'\mathbf{u} \leq 1 \right\}. \tag{31}$$

Since

$$\begin{aligned} \mathbf{w}'\boldsymbol{\mu} &= \left\langle \mathbf{w}, \boldsymbol{\mu}_1 + \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{1/2} \mathbf{F} \boldsymbol{\Gamma}^{1/2} \mathbf{u} \right\rangle \\ &= \langle \mathbf{w}, \boldsymbol{\mu}_1 \rangle + \left\langle \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{1/2} \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \mathbf{w}, \mathbf{u} \right\rangle, \end{aligned} \tag{32}$$

we obtain:

$$\begin{aligned} \min_{\boldsymbol{\Sigma} \in \widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}} \{\mathbf{w}'\boldsymbol{\mu}\} &= \langle \mathbf{w}, \boldsymbol{\mu}_1 \rangle + \min_{\mathbf{u}'\mathbf{u} \leq 1} \left\langle \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{1/2} \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \mathbf{w}, \mathbf{u} \right\rangle \\ &= \mathbf{w}'\boldsymbol{\mu}_1 - \left(\frac{1}{T_1} \frac{\nu_1}{\nu_1 - 2} q_{\boldsymbol{\mu}}^2 \right)^{1/2} \left\| \boldsymbol{\Gamma}^{1/2} \mathbf{F}' \mathbf{w} \right\|. \end{aligned} \tag{33}$$

Recalling (27) this becomes:

$$\min_{\boldsymbol{\Sigma}\in\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\mu}}}\left\{\mathbf{w}'\boldsymbol{\mu}\right\}=\mathbf{w}'\boldsymbol{\mu}_1-\left(\frac{1}{T_1}\frac{\nu_1}{\nu_1-2}q_{\boldsymbol{\mu}}^2\right)^{1/2}\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_1\mathbf{w}}. \tag{34}$$

7.2 Worst-case scenario for $\boldsymbol{\Sigma}$

Consider the ellipsoid (5):

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}\equiv\left\{\boldsymbol{\Sigma}:\ \operatorname{vech}\left[\boldsymbol{\Sigma}-\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]'\mathbf{S}_{\boldsymbol{\Sigma}}^{-1}\operatorname{vech}\left[\boldsymbol{\Sigma}-\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]\leq q_{\boldsymbol{\Sigma}}^2\right\}, \tag{35}$$

under the specifications (17)-(18). Consider the spectral decomposition of the rescaled dispersion parameter (18):

$$\left(\mathbf{D}_N'\left(\boldsymbol{\Sigma}_1^{-1}\otimes\boldsymbol{\Sigma}_1^{-1}\right)\mathbf{D}_N\right)^{-1}\equiv\mathbf{E}\boldsymbol{\Lambda}\mathbf{E}', \tag{36}$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix of the eigenvalues sorted in decreasing order and $\mathbf{E}$ is the juxtaposition of the respective eigenvectors. We can write (35) as follows:

$$\begin{aligned}\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}\ \equiv\ &\left\{\operatorname{vech}\left[\boldsymbol{\Sigma}-\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]'\mathbf{E}\boldsymbol{\Lambda}^{-1/2}\right. \\ &\left.\boldsymbol{\Lambda}^{-1/2}\mathbf{E}'\operatorname{vech}\left[\boldsymbol{\Sigma}-\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]\leq\frac{2\nu_1^2q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right\}.\end{aligned} \tag{37}$$

Define the new variable:

$$\mathbf{u}\equiv\left(\frac{2\nu_1^2q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{-1/2}\boldsymbol{\Lambda}^{-1/2}\mathbf{E}'\operatorname{vech}\left[\boldsymbol{\Sigma}-\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right], \tag{38}$$

which implies:

$$\operatorname{vech}\left[\boldsymbol{\Sigma}\right]\equiv\operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]+\left(\frac{2\nu_1^2q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}. \tag{39}$$

We can write (37) as follows:

$$\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}\equiv\left\{\operatorname{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right]+\left(\frac{2\nu_1^2q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u},\quad\mathbf{u}'\mathbf{u}\leq1\right\}. \tag{40}$$

From (39) we obtain:

$$\begin{aligned}
\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} &= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathrm{vec}\left[\boldsymbol{\Sigma}\right] \\
&= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathbf{D}_N\,\mathrm{vech}\left[\boldsymbol{\Sigma}\right] \\
&= \Big\langle \mathbf{D}_N'\left(\mathbf{w}'\otimes\mathbf{w}'\right)', \mathrm{vech}\left[\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\right] \\
&\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\mathbf{E}\boldsymbol{\Lambda}^{1/2}\mathbf{u}\Big\rangle \\
&= \mathbf{w}'\widehat{\boldsymbol{\Sigma}}_{\mathrm{ce}}\mathbf{w} \\
&\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N'\left(\mathbf{w}'\otimes\mathbf{w}'\right)', \mathbf{u}\right\rangle .
\end{aligned} \tag{41}$$

Substituting (17) in (41) we obtain:

$$\begin{aligned}
\max_{\boldsymbol{\Sigma}\in\widehat{\boldsymbol{\Theta}}_{\boldsymbol{\Sigma}}}\left\{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\right\} &= \frac{\nu_1}{\nu_1+N+1}\mathbf{w}'\boldsymbol{\Sigma}_1\mathbf{w} \\
&\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\max_{\mathbf{u}'\mathbf{u}\leq 1}\left\langle \boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N'\left(\mathbf{w}'\otimes\mathbf{w}'\right)', \mathbf{u}\right\rangle \\
&= \frac{\nu_1}{\nu_1+N+1}\mathbf{w}'\boldsymbol{\Sigma}_1\mathbf{w} \\
&\quad + \left(\frac{2\nu_1^2 q_{\boldsymbol{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\left\|\boldsymbol{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}_N'\left(\mathbf{w}'\otimes\mathbf{w}'\right)'\right\| .
\end{aligned} \tag{42}$$

To simplify this expression, consider the pseudo inverse $\widetilde{\mathbf{D}}$ of the duplication matrix:

$$\widetilde{\mathbf{D}}_N\mathbf{D}_N = \mathbf{I}_{N(N+1)/2}. \tag{43}$$

It is possible to show that:

$$\left(\mathbf{D}_N'\mathbf{A}\mathbf{D}_N\right)^{-1} = \widetilde{\mathbf{D}}_N\mathbf{A}^{-1}\widetilde{\mathbf{D}}_N' \tag{44}$$

and

$$\left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathbf{D}_N\widetilde{\mathbf{D}}_N = \left(\mathbf{w}'\otimes\mathbf{w}'\right), \tag{45}$$

see Magnus and Neudecker (1999).

Now consider the square of the norm in (42). Using (44) and (45) we obtain:

$$
\begin{aligned}
a &\equiv \left\| \mathbf{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}'_N \left(\mathbf{w}'\otimes\mathbf{w}'\right)' \right\|^2 \\
&= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathbf{D}_N\mathbf{E}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{E}'\mathbf{D}'_N\left(\mathbf{w}'\otimes\mathbf{w}'\right)' \\
&= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathbf{D}_N\left(\mathbf{D}'_N\left(\mathbf{\Sigma}_1^{-1}\otimes\mathbf{\Sigma}_1^{-1}\right)\mathbf{D}_N\right)^{-1}\mathbf{D}'_N\left(\mathbf{w}'\otimes\mathbf{w}'\right)' \\
&= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathbf{D}_N\widetilde{\mathbf{D}}_N\left(\mathbf{\Sigma}_1\otimes\mathbf{\Sigma}_1\right)\widetilde{\mathbf{D}}'_N\mathbf{D}'_N\left(\mathbf{w}'\otimes\mathbf{w}'\right)' \\
&= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\mathbf{D}_N\widetilde{\mathbf{D}}_N\left(\mathbf{\Sigma}_1\otimes\mathbf{\Sigma}_1\right)\left[\left(\mathbf{w}'\otimes\mathbf{w}'\right)\left(\mathbf{D}_N\widetilde{\mathbf{D}}_N\right)\right]' \\
&= \left(\mathbf{w}'\otimes\mathbf{w}'\right)\left(\mathbf{\Sigma}_1\otimes\mathbf{\Sigma}_1\right)\left(\mathbf{w}'\otimes\mathbf{w}'\right)' \\
&= \left(\mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}\right)\otimes\left(\mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}\right) = \left(\mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}\right)^2.
\end{aligned} \tag{46}
$$

Therefore (42) yields:

$$
\max_{\mathbf{\Sigma}\in\widehat{\mathbf{\Theta}}_{\mathbf{\Sigma}}} \mathbf{w}'\mathbf{\Sigma}\mathbf{w} = \left[\frac{\nu_1}{\nu_1+N+1} + \left(\frac{2\nu_1^2 q_{\mathbf{\Sigma}}^2}{\left(\nu_1+N+1\right)^3}\right)^{1/2}\right]\left(\mathbf{w}'\mathbf{\Sigma}_1\mathbf{w}\right). \tag{47}
$$

7.3 Robust Bayesian mean-variance problem

Substituting (34) and (47) in (3) and using the definitions (20) and (21) we obtain the robust Bayesian mean-variance problem (19).

Recalling (27), we can write (19) as follows:

$$
\mathbf{w}^{(i)} = \underset{\mathbf{w}}{\operatorname{argmax}}\left\{\mathbf{w}'\boldsymbol{\mu}_1 - \gamma_\mu\left\|\mathbf{\Gamma}^{1/2}\mathbf{F}'\mathbf{w}\right\|\right\} \tag{48}
$$

$$
\text{subject to } \begin{cases} \mathbf{w}\in\mathcal{C} \\ \left\|\mathbf{\Gamma}^{1/2}\mathbf{F}'\mathbf{w}\right\| \leq \sqrt{\gamma_\Sigma^{(i)}}. \end{cases}
$$

This is equivalent to:

$$
\left(\mathbf{w}^{(i)}, z^*\right) = \underset{\mathbf{w}}{\operatorname{argmax}}\left\{\mathbf{w}'\boldsymbol{\mu}_1 - z\right\} \tag{49}
$$

$$
\text{subject to } \begin{cases} \mathbf{w}\in\mathcal{C} \\ \left\|\mathbf{\Gamma}^{1/2}\mathbf{F}'\mathbf{w}\right\| \leq z/\gamma_\mu \\ \left\|\mathbf{\Gamma}^{1/2}\mathbf{F}'\mathbf{w}\right\| \leq \sqrt{\gamma_\Sigma^{(i)}}. \end{cases}
$$

If the investment constraints $\mathcal{C}$ are at most quadratic this is a second order cone programming problem.