Comput Optim Appl
DOI 10.1007/s10589-009-9260-7

Sensitivity analysis and calibration of the covariance matrix for stable portfolio selection

Vincent Guigues

Received: 10 April 2007

Abstract We recommend an implementation of the Markowitz problem to generate stable portfolios with respect to perturbations of the problem parameters. The stability is obtained proposing novel calibrations of the covariance matrix between the returns that can be cast as convex or quasiconvex optimization problems. A statistical study as well as a sensitivity analysis of the Markowitz problem allow us to justify these calibrations. Our approach can be used to do a global and explicit sensitivity analysis of a class of quadratic optimization problems. Numerical simulations finally show the benefits of the proposed calibrations using real data.

Keywords Markowitz model · Sensitivity analysis · Covariance matrix estimation · Quadratic programming · Semidefinite programming

1 Introduction

We are interested in the stability of the portfolio solution of the Markowitz problem [12] and of a generalisation of this problem taking into account the transaction costs [6]. The Markowitz approach today remains both the simplest and the most general portfolio selection model. However, the estimation of the problem parameters, the mean return vector ρ and the covariance matrix Q between the returns over the investment period, is a complicated task. For instance, it is pointed out in [1, 2], that if we use the empirical estimations of the parameters, the portfolio's composition is traditionally very sensitive to changes in the returns. Our approach takes into account the numerical risk that is linked with the first step of estimating the statistical quantities by introducing an intermediate step between this first step of statistical estimation and the second step of selection. This intermediate step can be interpreted

V. Guigues (✉)
Laboratoire de Modélisation et Calcul, IMAG, BP 53, 38041 Grenoble Cedex 9, France
e-mail: vincent.guigues@imag.fr

as a filter or as a numerical regularization of the statistical estimations and results in a new calibration of the covariance matrix. This calibration thus focuses on the defaults of the initial estimation of the covariance matrix. This initial estimation depends on the model for the returns: i.i.d. as in [11] or slowly varying mean and covariance matrix as in [7].

Our paper is organized as follows. The second section of the paper briefly recalls the Markowitz model and the problem of estimating its parameters. It also gives a few properties of the Markowitz model useful for our study. To control portfolio stability, given two portfolios x_1^* and x_2^* obtained for the values (ρ_1, Q_1) and (ρ_2, Q_2) of the parameters, we would like to bound from above $\|x_2^* - x_1^*\|_1$ or $\|x_2^* - x_1^*\|_2$ in terms of $\|Q_2 - Q_1\|$ and $\|\rho_2 - \rho_1\|$. Notice that contrary to $\|x_2^* - x_1^*\|_2$, $\|x_2^* - x_1^*\|_1$ has a physical interpretation; it represents the portfolio composition variation, but the bounds we obtain on $\|x_2^* - x_1^*\|_2$ allow us to justify some existing covariance matrix calibrations such as [13] (which was motivated by numerical observations) and the calibrations we introduce in Sect. 4. The third section is thus devoted to a sensitivity analysis of the Markowitz problems [12] and [6]. Three different versions of the Markowitz model are studied. Since these three models can all be cast as quadratic optimization problems satisfying the Slater assumption, we already know from [5] that the solutions are locally radially Lipschitz, though in [5] the Lipschitz constant is not explicit. On the contrary, our sensitivity analysis aims at finding explicit and global bounds. For the version where the return constraint is aggregated in the objective, we show that the solutions are radially Lipschitz with respect to the parameters. We then study a version of the problem integrating a return constraint without transactions costs as in [12] and with transaction costs as in [6]. Roughly speaking, the sensitivity analysis of all models tends to show that the portfolios generated using the Markowitz model will be stable with respect to small perturbations of the parameters if the lowest eigenvalue of the estimated covariance matrix and at least one mean return are sufficiently large. The sensitivity analysis, through Theorems 3.1 and 3.2, is thus the theoretical support for the stable covariance matrix calibrations we propose in Sect. 4. Numerical simulations in Sect. 5 show that one of the calibrations we propose leads to the most stable portfolios (among a set of competing calibration methods) while providing performing portfolios.

2 Markowitz model, sources of instabilities and statistical framework

2.1 Markowitz mean-variance model

We recall the formulations of [6, 12]. The Markowitz model is a portfolio optimization model corresponding to a single investment over a given investment period of H time steps. Given n risky assets and a risk-free asset, the Markowitz model gives the proportion of the different assets composing the optimal portfolio. The return r_i of each asset i over the investment period is unknown. The standard mean-variance Markowitz model uses the first and second moments of the distribution of the returns. Therefore, the probability distribution of the returns r over the investment period is characterized by a vector of expected returns $\mathbb{E}[r] = \rho$ and a covariance matrix between the returns Q such that $Q = \mathbb{E}[(r - \rho)(r - \rho)^\top]$. A portfolio is then given by a vector $x \in \mathbb{R}^n$ of risky asset weights. The weight of the risk-free asset (whose return is ρ_0) is $x_0 = 1 - x^\top \mathbf{e}$, where in this expression, and in what follows, $\mathbf{e}$ is a vector with all components equal to one. Hence, the expected total return of the portfolio is $\mathbb{E}[x^\top r + x_0\rho_0] = x^\top \rho + x_0\rho_0$ and the risk of the investment is defined by the variance of the total return of the portfolio $\mathbb{E}[(x^\top r - x^\top \rho)^2] = x^\top Qx$.

The optimal portfolio is then a solution of the following problem $P(k, \rho, Q)$ parameterized by k, ρ and Q:

$$P(k, \rho, Q) \quad \begin{cases} \min \frac{1}{2} x^\top Qx - k\, x^\top(\rho - \rho_0\mathbf{e}) \\ x \in \Delta_n, \end{cases}$$

where $k \geq 0$ depends on the investor's risk aversion and $\Delta_n = \{x \in \mathbb{R}^n \mid x^\top \mathbf{e} \leq 1, x \geq 0\}$ denotes the unit simplex. The model simultaneously tries to minimize the variance of the portfolio return and to maximize the expected return of the portfolio over the investment period.

Another approach is based on a target value ℓ for the expected return and yields the following problem $P'(\ell, \rho, Q)$:

$$P'(\ell, \rho, Q) \quad \begin{cases} \min \frac{1}{2} x^\top Qx \\ x^\top(\rho - \rho_0\mathbf{e}) \geq \ell - \rho_0, \ x \in \Delta_n. \end{cases}$$

Finally, it is also possible to take transaction costs into account as in [6]. In [6], the i-th component x_i of a portfolio $x = (x_1, \dots, x_n)$ gives the amount invested in asset i, the amount x_0 being invested in the risk-free asset. We introduce the following notation:

- x_i^-: the initial value of i-th asset before the rebalancing of the portfolio;
- y_i: the amount of risky asset i we sell at the beginning of the period, with the corresponding transaction cost μ_i $(0 < \mu_i < 1)$;
- z_i: the amount of risky asset i we buy at the beginning of the period, with the corresponding transaction cost ν_i $(0 < \nu_i < 1)$.

The set of portfolios is then defined by the following constraints:

$$\begin{cases} x_i = x_i^- - y_i + z_i, \quad i = 1, \dots, n, \\ x_0 = x_0^- + \sum_{i=1}^n (1 - \mu_i)\, y_i - \sum_{i=1}^n (1 + \nu_i)\, z_i, \\ x \geq 0, \quad x_0 \geq 0, \quad y \geq 0, \quad z \geq 0, \end{cases}$$

where $(x^-, x_0^-) \geq 0$ and $(x^-, x_0^-) \neq 0$. Notice that if (x^-, x_0^-) was null, the only admissible portfolio would be $x = 0$. The Markowitz problem taking into account the transaction costs then reads:

$$P''(\ell, \rho, Q) \quad \begin{cases} \min \frac{1}{2} x^\top Qx \\ \rho^\top x + \rho_0 \left(x_0^- + (\mathbf{e} - \mu)^\top y - (\mathbf{e} + \nu)^\top z\right) \geq \ell\, (\mathbf{e}^\top x^- + x_0^-), \\ x + y - z = x^-, \\ (\mathbf{e} + \nu)^\top z - (\mathbf{e} - \mu)^\top y \leq x_0^-, \\ x \geq 0, \quad y \geq 0, \quad z \geq 0. \end{cases} \tag{1}$$

The return constraint in P' (resp. P'') is equivalent to $x^\top \rho + x_0\rho_0 \geq \ell$ (resp. $x^\top \rho + \rho_0 x_0 \geq \ell\ (\mathbf{e}^\top x^- + x_0^-)$); meaning indeed that ℓ is a target mean return. Also if x^* (resp. (x^*, y^*, z^*)) is an optimal solution of problem P' (resp. P'') then the weight (resp. the amount) of the risk-free asset is $x_0^* = 1 - \mathbf{e}^\top x^*$ (resp. $x_0^* = x_0^- + (\mathbf{e} - \mu)^\top y^* - (\mathbf{e} + \nu)^\top z^*$). From now on, we use the following hypotheses:

H1. The covariance matrix Q is positive definite.

H2. For problem P', $0 < \rho_0 < \ell$, and for problem P'', $0 < \rho_0 < \frac{\ell(\mathbf{e}^\top x^- + x_0^-)}{(\mathbf{e}-\mu)^\top x^- + x_0^-}$.

H3. There exists $\kappa > 0$ such that for problem P', for at least one component i, $\rho(i) > \ell + \kappa$, and for problem P'', for at least one component i, we have $\rho(i) > \frac{(1+\nu_i)}{(\mathbf{e}-\mu)^\top x^- + x_0^-}(\ell + \kappa)(\mathbf{e}^\top x^- + x_0^-)$. Also, for P' and P'', vectors ρ and $\mathbf{e}$ are linearly independent.

In what follows, we say that problem $P(k, \rho_1, Q_1)$, $P'(\ell, \rho_1, Q_1)$ or $P''(\ell, \rho_1, Q_1)$ satisfies hypotheses H1, H2 and H3 if the above hypotheses H1, H2 and H3 are satisfied replacing ρ by ρ_1 and Q by Q_1.

A few comments on hypotheses H1, H2 *and* H3 The covariance matrix Q is always positive semidefinite. Hypothesis H1 is needed for the sensitivity analysis but is also consistent with the commonly used assumption of arbitrage free markets. Indeed, if Q had a null eigenvalue with eigenvector v, the portfolio $x = \frac{v}{v^\top \mathbf{e}}$ (if we allow for short sellings) would be risk-free. We would then have the illusion of being able to invest without risk on risky assets.

If hypothesis H2 does not hold for $P'(\ell, \rho, Q)$ or $P''(\ell, \rho, Q)$, then an optimal strategy consists of investing everything in the risk-free asset.

Condition H3 is not too demanding: it requires a mean return $\rho(i)$ to be sufficiently large. For instance, for problem P', it requires a mean return to be strictly greater than the target mean return ℓ; but for problem P' to be feasible, there must be at least one asset i such that $\rho(i) \geq \ell$. For P'', hypothesis H3 implies that at least one asset has mean return strictly greater than ℓ and guarantees that the portfolio obtained investing all the money in asset i satisfies the return constraint i.e., has a mean return greater than $\ell(\mathbf{e}^\top x^- + x_0^-)$. Hypothesis H3 also allows us to show the Slater assumption for P' and P''. Finally, notice that hypotheses H2 and H3 for problem P' can be obtained replacing μ and ν by 0 (there are no transaction costs) in H2 and H3 for P''.

2.2 A few properties of the Markowitz model

We give a few properties of the Markowitz model that will be useful for our sensitivity analysis. Since the objective function of problem $P'(\ell, \rho, Q)$ (resp. $P''(\ell, \rho, Q)$) is defined everywhere, and bounded from below on the polyhedral and nonempty feasible set, both primal problem P' (resp. P'') and its dual are equivalent to each other. We will thus be able to either work on problem P' or P'' directly or on their duals.

Lemma 2.1 *A constraint of a convex problem that is not active at the optimum can be removed without changing the optimal value.*

Proof Let us write the convex problem under the form:

$$\mathcal{P}_1 \quad \begin{cases} \min h(x) \\ g_i(x) \le 0, \quad i \in J. \end{cases}$$

Let us denote by X_1 the feasible set of $\mathcal{P}_1$, x_1 the minimizer of h over X_1 and h_1 the optimal value of $\mathcal{P}_1$. Let us consider a non-active constraint at the optimum with index $i_0 \in J$. We thus have $g_{i_0}(x_1) < 0$. We show that $\mathcal{P}_1$ is equivalent to the problem of minimizing h over the set $X_2 = \{x \in \mathbb{R}^n \mid g_i(x) \le 0, i \in J \backslash i_0\}$.

Since $X_1 \subseteq X_2$, the minimum h_2 of h over X_2 is clearly less than or equal to h_1. We show that in fact, for all $x \in X_2, h(x) \ge h_1$ (which will imply that $h_2 \ge h_1$ and that the two problems have the same optimal values). Let $x \in X_2$. If $g_{i_0}(x) \le 0$ then $x \in X_1$ and $h(x) \ge h_1$ by definition of x_1. Contrarily, if $g_{i_0}(x) > 0$, since $g_{i_0}(x_1) < 0$ and since g_{i_0} is continuous, the intermediate value theorem gives the existence of $t^* \in]0, 1[$ such that $g_{i_0}(t^* x_1 + (1 - t^*)x) = 0$. Besides, from the convexity of the set X_2, it follows that $x_0 = t^* x_1 + (1 - t^*)x \in X_2$ (since x_1 and x are in X_2). This implies $x_0 \in X_1$ and $h(x_0) \ge h_1$. Finally, since h is convex, we obtain $h_1 \le h(x_0) \le t^* h_1 + (1 - t^*)h(x)$. □

Lemma 2.2 *Consider problems $P'(\ell, \rho, Q)$ and $P''(\ell, \rho, Q)$ and suppose that Assumptions* H1, H2 *and* H3 *are satisfied for $P'(\ell, \rho, Q)$ and $P''(\ell, \rho, Q)$. The following holds*:

(i) *The Slater condition of qualification of constraints is satisfied for P' and P''.*
(ii) *The return constraint is active at the optimal solution x^* : $(\rho - \rho_0\, \mathbf{e})^\top x^* = \ell - \rho_0$ for problem P' and $\rho^\top x^* + \rho_0 x_0^* = \ell\, (\mathbf{e}^\top x^- + x_0^-)$ for problem P''.*

Proof Let us show (i) for P'. From H3, we can find an index i such that $\rho(i) > \ell$. Let $\varepsilon > 0$ and let us define the portfolio $x \in \mathbb{R}^n$ by $x_i = 1 - n\varepsilon$ and $x_k = \varepsilon$ for $k \ne i$. We have $x^\top \mathbf{e} < 1$ and if $\varepsilon < \frac{1}{n}$, we also have $x > 0$. Finally, since $x^\top(\rho - \rho_0\, \mathbf{e}) = \rho(i) - \rho_0 + a\varepsilon$, for some $a \in \mathbb{R}$, we can choose ε sufficiently small in such a way that $x^\top(\rho - \rho_0\, \mathbf{e}) > \ell - \rho_0$ and thus that no constraint is active at x. We now show (i) for P''. Let i be such that $\rho(i) > \frac{(1+\nu_i)}{(\mathbf{e}-\mu)^\top x^- + x_0^-}(\ell + \kappa)(\mathbf{e}^\top x^- + x_0^-)$. Let $\varepsilon > 0$ and let $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ be such that $x = x^- - y + z$ and

$$\begin{cases} \text{if } k \ne i \text{ and } x_k^- = 0, & \text{then } y_k = \varepsilon \text{ and } z_k = 2\varepsilon, \\ \text{if } k \ne i \text{ and } x_k^- > 0, & \text{then } y_k = x_k^- \text{ and } z_k = \varepsilon, \\ \text{finally, } y_i = x_i^- + \varepsilon, & \text{and } z_i \text{ is such that } \quad x_0 = \varepsilon. \end{cases} \tag{2}$$

The amount z_i can be expressed as $z_i = \frac{1}{1+\nu_i}(x_0^- + \sum_{j=1}^n (1 - \mu_j)x_j^-) + a\varepsilon$, for some $a \in \mathbb{R}$ and we have $x_i = -\varepsilon + z_i$ and $\rho^\top x + \rho_0 x_0 = \frac{\rho(i)}{1+\nu_i}(x_0^- + \sum_{j=1}^n (1 - \mu_j)x_j^-) + a'\varepsilon$, for some $a' \in \mathbb{R}$. Since $(x^-, x_0^-) \ge 0$, with $(x^-, x_0^-) \ne 0$, and since H3 holds, we can choose ε sufficiently small to have $z_i > 0$, $x_i > 0$ and $\rho^\top x + \rho_0 x_0 > \ell(\mathbf{e}^\top x^- + x_0^-)$. No inequality constraint is thus active at (x, y, z).

Let us now prove (ii). First, from (i), the feasible set of both P' and P'' is not empty (and compact) and both P and P' have optimal solutions that satisfy the return constraint. Now by contradiction, suppose the return constraint is not active at the optimum for P' and P''. Then, since H1 holds, using Lemma 2.1, we could remove this constraint for convex problems P and P' without changing the optimal value and the solution of problem P' would be $x^* = 0$. But $x = 0$ does not satisfy the return constraint since H2 holds so the return constraint is active for P'. For problem P'', $(x^* = 0, x_0^* = x_0^- + \sum_{j=1}^n (1 - \mu_j) x_j^-, y^* = x^-, z^* = 0)$, would be a feasible point and the objective function at this point is 0. We would thus necessarily have $x^* = 0$ for problem P'' and the optimal value of P'' would be 0. However, the return constraint cannot be satisfied with $x = 0$. Indeed, the maximal return that can be obtained with $x = 0$ is the optimal value of the following optimization problem:

$$\begin{cases} \max \rho_0 (x_0^- + (\mathbf{e} - \mu)^\top y - (\mathbf{e} + \nu)^\top z) \\ y - z = x^-, \quad y \geq 0, \quad z \geq 0, \\ (\mathbf{e} + \nu)^\top z - (\mathbf{e} - \mu)^\top y \leq x_0^-. \end{cases} \tag{3}$$

Since the optimal value of the above optimization problem (3) is $\rho_0 (x_0^- + \sum_{j=1}^n (1 - \mu_j) x_j^-)$ (obtained with $y_j = x_j^-, z_j = 0$), and since H2 holds, the return constraint cannot be satisfied for P'' with $x = 0$. Thus the return constraint cannot be removed from P'' neither and it is also active for P''. □

Notice that if the optimal solution x^* of $P'(\ell, \rho, Q)$ satisfies $x_i^* > 0$ for $i = 1, \ldots, n$, then it suffices to apply the KKT Theorem (pp. 305–306 of [9]) to get an explicit expression of x^*. We also have an explicit expression of the solution if short sellings are allowed for $P(k, \rho, Q)$ and $P'(\ell, \rho, Q)$, i.e., if the constraints $(x, x_0) \geq 0$ are removed. Indeed, in this case, problems $P(k, \rho, Q)$ and $P'(\ell, \rho, Q)$ amount to solving problems $\tilde{P}(k, \rho, Q)$ and $\tilde{P}'(\ell, \rho, Q)$ below:

$$\tilde{P}(k, \rho, Q) \quad \begin{cases} \min \frac{1}{2} x^\top Q x - k\, x^\top (\rho - \rho_0 \mathbf{e}) \\ x \in \mathbb{R}^n, \end{cases}$$

$$\tilde{P}'(\ell, \rho, Q) \quad \begin{cases} \min \frac{1}{2} x^\top Q x \\ x^\top (\rho - \rho_0 \mathbf{e}) \geq \ell - \rho_0. \end{cases}$$

Lemma 2.3 *If Q is positive definite, if $\rho_0 < \ell$ and if ρ and $\mathbf{e}$ are linearly independent, then optimal solutions to $\tilde{P}(k, \rho, Q)$ and $\tilde{P}'(\ell, \rho, Q)$ are respectively given by*:

$$x^*(k, \rho, Q) = k Q^{-1} (\rho - \rho_0 \mathbf{e}) \quad \textit{and}$$

$$x^*(\ell, \rho, Q) = \frac{\ell - \rho_0}{(\rho - \rho_0 \mathbf{e})^\top Q^{-1} (\rho - \rho_0 \mathbf{e})} Q^{-1} (\rho - \rho_0 \mathbf{e}).$$

We conclude this section discussing the sources of instability of the composition of the portfolios.

2.3 Sources of instabilities and statistical framework

The sources of instability are the parameters of the model, i.e., the mean return vector ρ and the covariance matrix Q. The stability of the portfolio selection process thus depends on the calibration of ρ and Q. More precisely, the next section will provide a desirable property of the calibrated covariance matrix for stability.

We will thus focus on covariance matrix calibration for portfolio selection and will do this study in two statistical frameworks for the underlying process of returns:

(A) The case of i.i.d. returns.
(B) The case of a weakly stationary process for the returns where the mean ρ and the covariance matrix Q slowly vary in time as in [7] (see details below).

Though many papers study the calibration of the covariance matrix of stock returns assuming i.i.d. returns, this assumption may only be valid on short periods of time. It is thus of interest to consider model (B) above which is more realistic for stock returns on arbitrary time periods.

Let r_t, $t = 1, \ldots, T$, be T observations of the returns, available the day of the investment. When the returns are i.i.d., the traditional estimations of the mean and of the covariance matrix are the empirical mean $\hat{\rho}$ and the empirical covariance matrix $\hat{Q}$ defined by

$$\hat{\rho} = \frac{1}{T}\sum_{t=1}^{T} r_t \quad \text{and} \quad \hat{Q} = \frac{1}{T}\sum_{t=1}^{T} (r_t - \hat{\rho})(r_t - \hat{\rho})^\top.$$

Some criticisms are commonly formulated on this estimation $\hat{Q}$. The rank of the empirical covariance matrix is less than or equal to T so if $n \geq T + 1$, this matrix is not invertible. If the number of assets n is close to the number of available observations per asset T, then the total number of parameters to estimate is close to the total number of observations which is problematic. In practice, we realize that even if the number of observations T per asset is much greater than the number of assets, the estimated covariance matrix is ill-conditioned. Taking for instance the assets of the Dow Jones (from January 1999 to January 2002), we observed that in most cases, using different samples of size $T = 900$, about one half of the eigenvalues of the empirical covariance matrix is nearly 0 and the condition number is around 10^7.

With model (B) above (see [7]), we suppose the returns follow the quite general and distribution-free model

$$r_t = \rho_t + \zeta_t, \quad \text{with } \mathbb{E} r_t = \rho_t, \qquad \mathbb{E}\zeta_t\zeta_t^\top = Q_t \succeq 0,$$

where ζ_t are independent random vectors in $\mathbb{R}^n$ with a mean of zero. We also suppose that for some $\sigma > 0$, $\mathbb{E}\|r_t\|_\infty^4 \leq \sigma^4$. Let τ be the investment date and H be the investment horizon. Using this model for the returns and if there is an interval of local time homogeneity, then a procedure is detailed in [7] to determine adaptive estimations $\hat{\rho}$ and $\hat{Q}$ of the H time steps mean return $\rho = \rho_\tau$ over the investment period and of the covariance matrix $Q = Q_\tau$ between the H time steps returns. An interval of local time homogeneity is an interval where ρ_t and Q_t slowly vary on this interval. A more precise definition of this interval can be found in [7]. The adaptive estimations are the empirical estimations of the mean and of the covariance matrix when using only the data of the interval of homogeneity. The criticisms formulated above for the empirical covariance matrix are thus valid for the adaptive covariance matrix replacing T by the length of this interval.

However, if the empirical or adaptive (depending on the statistical context) estimations have known defaults, they contain information and permit, not only to give bounds on the errors we make using them, but also to give a reasonable estimation of the solution [7]. Moreover, in the case when the returns are i.i.d., the empirical covariance matrix also has nice properties such as being maximum likelihood under normality. By definition, in this framework, it is thus the most likely covariance matrix given the data. We thus propose to take as a starting point of the estimation of the Markowitz model parameters, the empirical or adaptive (depending on the context) estimations. In what follows, these estimations will be denoted by $\hat{\rho}$ and $\hat{Q}$ for respectively the mean and the covariance matrix. We will explain in Sect. 4 how to correct this estimation $\hat{Q}$ of the covariance matrix. To this aim, we start with a sensitivity analysis of the Markowitz problem.

3 Sensitivity analysis of the Markowitz problem

We fix nominal values k (or ℓ) and (ρ_1, Q_1) for the parameters of the Markowitz problem, and consider the corresponding optimization problem as the unperturbed problem. For a given perturbation (ρ_2, Q_2) of parameters (ρ_1, Q_1), we consider the corresponding perturbed Markowitz problem, the parameter k (or ℓ) remaining fixed. The objective function of the unperturbed and perturbed problems will respectively be denoted by f_1 and f_2 (whose expressions may differ, depending on the Markowitz problem studied). We denote the solution of $P(k, \rho_i, Q_i)$ or $P'(\ell, \rho_i, Q_i)$ by x_i^* (it is unique because Q_i is positive definite) and a solution of $P''(\ell, \rho_i, Q_i)$ by (x_i^*, y_i^*, z_i^*). Finally, in what follows, $\mathcal{S}_n(\mathbb{R})$ is the set of real symmetric matrices of size n and for $X \in \mathcal{S}_n(\mathbb{R})$, $X \succeq 0$ (resp. $X \succ 0$) means the real symmetric matrix X is positive semidefinite (resp. positive definite).

In [1, 2], a sensitivity analysis of P is done through a parametric quadratic programming formulation but in a simplified setting: without risk-free asset and considering Q fixed. In [5], Daniel shows that under the Slater Assumption (which holds for problems P' and P'' due to Lemma 2.2), solutions to a general quadratic optimization problem are locally radially Lipschitz, but without providing an explicit Lipschitz constant.

Our contribution is to provide global bounds that are explicit functions of the parameters. The study can be extended to the sensitivity analysis of quadratic optimization problems.

3.1 Sensitivity analysis of problem P

The feasible set of problem P is fixed when ρ and Q vary. Since f_1 satisfies a second order growth condition on Δ_n, we can apply the following proposition to obtain the sensitivity of the solutions.

Proposition 3.1 (Proposition 4.32, p. 287 in [3].) *Let us consider the two optimization problems*

$$\mathcal{P}_1 \quad \begin{cases} \min f_1(x) \\ x \in X \end{cases} \qquad \textit{and} \quad \mathcal{P}_2 \quad \begin{cases} \min f_2(x) \\ x \in X, \end{cases}$$

where $f_1, f_2 : X \to \mathbb{R}$. *Let* S_1 *be the set of solutions of* $\mathcal{P}_1$ *and let* x_2^* *be a solution of problem* $\mathcal{P}_2$. *If* (i) f_1 *satisfies a second-order growth condition on* X ($\exists\, c > 0$ *such that for every* $x \in X$ *and* $x_1^* \in S_1$, $f_1(x) \geq f_1(x_1^*) + c\|x - x_1^*\|^2$) *and* (ii) *the function* $f_2(\cdot) - f_1(\cdot)$ *is Lipschitz continuous with modulus* β *on* X, *then*

$$dist(x_2^*, S_1) \leq \frac{\beta}{c}.$$

Definition 3.1 For any symmetric matrix Q, let $\beta(Q)$ be such that the quadratic function $x^\top Qx$ is $\beta(Q)$-strongly convex with respect to $\| . \|_1$, i.e.,

$$\beta(Q) = \inf_{x \neq 0} \frac{x^\top Qx}{\|x\|_1^2}.$$

We will make use of the following lemma:

Lemma 3.1 *Let* $Q \in \mathcal{S}_n(\mathbb{R})$, *then* $\sup_{x \in \Delta_n} \|Qx\|_2 = \max_i \|C_i(Q)\|_2$, *where* $C_i(Q)$ *is the* i*-th column of* Q.

Proof Let us denote by $\tilde{q}_{ij}$ the elements of the matrix $Q^\top Q$. Then $\tilde{q}_{ii} = \sum_{j=1}^n {q_{ji}}^2 = \|C_i(Q)\|_2^2$. Hence, if e_i, $i = 1, \dots, n$, are the vectors of the canonical basis:

$$\begin{aligned} \sup_{x \in \Delta_n} \|Qx\|_2 &= \sup_{x \in \Delta_n} (x^\top Q^\top Qx)^{\frac{1}{2}} = \max_i (e_i^\top Q^\top Qe_i)^{\frac{1}{2}} \\ &= \max_i (\tilde{q}_{ii})^{\frac{1}{2}} = \max_i \|C_i(Q)\|_2. \end{aligned}$$

The second equality comes from the convexity of the problem: the maximum is attained at an extremal point of the feasible set. □

The following theorem provides a sensitivity analysis of problem P:

Theorem 3.1 *Consider problem* $P(k, \rho_1, Q_1)$ *and its perturbed version* $P(k, \rho_2, Q_2)$. *Let Assumption* H1 *hold for these problems. For* $i = 1, 2$, *if* x_i^* *is the solution of* $P(k, \rho_i, Q_i)$, *then*:

$$|f_2(x_2^*) - f_1(x_1^*)| \leq \frac{1}{2}\|Q_2 - Q_1\|_\infty + k\|\rho_2 - \rho_1\|_\infty, \tag{4}$$

$$\|x_2^* - x_1^*\|_1 \leq \frac{2}{\max(\beta(Q_1), \beta(Q_2))}(\|Q_2 - Q_1\|_\infty + k\|\rho_2 - \rho_1\|_\infty), \tag{5}$$

$$\|x_2^* - x_1^*\|_2 \le \frac{2}{\max(\lambda_{\min}(Q_1), \lambda_{\min}(Q_2))}\Big(\max_i \|C_i(Q_2 - Q_1)\|_2 + k\|\rho_2 - \rho_1\|_2\Big), \tag{6}$$

where $C_i(Q)$ is the i-th column of Q.

Proof Let us show (4). We suppose $f_2(x_2^*) \ge f_1(x_1^*)$ (the other case is symmetric). In this case, $|f_2(x_2^*) - f_1(x_1^*)| = f_2(x_2^*) - f_1(x_1^*) = f_2(x_2^*) - f_2(x_1^*) + f_2(x_1^*) - f_1(x_1^*)$. But since $x_1^* \in \Delta_n$, by definition of x_2^*, $f_2(x_2^*) - f_2(x_1^*) \le 0$. Thus,

$$\begin{aligned} |f_2(x_2^*) - f_1(x_1^*)| &\le \frac{x_1^{*\top}(Q_2 - Q_1)x_1^*}{2} - k(\rho_2 - \rho_1)^\top x_1^* \\ &\le \frac{\|x_1^*\|_1^2 \|Q_2 - Q_1\|_\infty}{2} + k\|\rho_2 - \rho_1\|_\infty \|x_1^*\|_1 \end{aligned}$$

with $\|x_1^*\|_1 \le 1$. Let us now show (5). First note that the objective function f_1 of the Markowitz problem $P(k, \rho_1, Q_1)$ satisfies a second-order growth condition on Δ_n:

$$\exists c > 0,\ \forall x \in \Delta_n \quad f_1(x) \ge f_1(x_1^*) + c\|x - x_1^*\|_1^2.$$

Indeed, a second-order Taylor series expansion of f_1 at x_1^* gives:

$$f_1(x) = f_1(x_1^*) + (x - x_1^*)^\top \nabla f_1(x_1^*) + \frac{1}{2}(x - x_1^*)^\top \nabla^2 f_1(x_1^*)(x - x_1^*),$$

where $\nabla f_1(x_1^*) = Q_1 x_1^* - k(\rho_1 - \rho_0 \mathbf{e})$ and $\nabla^2 f_1(x_1^*) = Q_1$. The first-order optimality conditions give $(x - x_1^*)^\top \nabla f_1(x_1^*) \ge 0$ for all $x \in \Delta_n$. On the other hand:

$$(x - x_1^*)^\top \nabla^2 f_1(x_1^*)(x - x_1^*) \ge \beta(Q_1)\|x - x_1^*\|_1^2.$$

Hence, (3.1) is satisfied with $c = \frac{\beta(Q_1)}{2}$ and $c > 0$ since $Q_1 \succ 0$ (hypothesis H1). It remains to show that the function $h(\cdot) = f_2(\cdot) - f_1(\cdot)$ is Lipschitz continuous on Δ_n which is straightforward. Indeed, since h is continuous and differentiable, we can use the mean value theorem to get:

$$\forall (x, y) \in \Delta_n\ |h(x) - h(y)| \le \sup_{x \in \Delta_n} (\|\nabla h(x)\|_\infty)\|x - y\|_1.$$

Further, for all $x \in \Delta_n$:

$$\|\nabla h(x)\|_\infty = \|(Q_2 - Q_1)x - k(\rho_2 - \rho_1)\|_\infty \le \|Q_2 - Q_1\|_\infty + k\|\rho_2 - \rho_1\|_\infty = \beta.$$

We then apply Proposition 3.1 to obtain $\|x_2^* - x_1^*\|_1 \le \frac{2}{\beta(Q_1)}(\|Q_2 - Q_1\|_\infty + k\|\rho_2 - \rho_1\|_\infty)$. Exchanging the role of x_1, f_1, ρ_1, Q_1, and x_2, f_2, ρ_2, Q_2, we can also show that $\|x_2^* - x_1^*\|_1 \le \frac{2}{\beta(Q_2)}(\|Q_2 - Q_1\|_\infty + k\|\rho_2 - \rho_1\|_\infty)$ and (5) follows. We can then show (6) following the proof of (5) and applying Lemma 3.1. □

Notice that the use of norm $\|.\|_1$ gives a bound with $\beta(Q_1)$ instead of $\lambda_{\min}(Q_1)$, the latter being easily computed.

3.2 Sensitivity analysis of problems P' and P''

The method we use for the sensitivity analysis of problems P' and P'' consists of introducing the dual problem obtained dualizing the return constraint and to work on this dual problem which is equivalent to the primal problem. Thus, the inner minimization problem solved to compute the value of the dual function for fixed λ, has a fixed feasible set. We then write the first order optimality conditions for this problem and bound the Lagrange multipliers. Notice that the Slater assumption for problems P' and P'' (which holds, due to Lemma 2.2) is a necessary and sufficient condition for the set of Lagrange multipliers to be bounded (Theorem 2.3.2, p. 312 of [9]).

Theorem 3.2 *Consider problem* $P'(\ell,\rho_1,Q_1)$ (*resp.* $P''(\ell,\rho_1,Q_1)$) *and its perturbed version* $P'(\ell,\rho_2,Q_2)$ (*resp.* $P''(\ell,\rho_2,Q_2)$). *Let Assumptions* H1, H2 *and* H3 *hold for these problems and let* $\kappa=\min(\kappa_1,\kappa_2)$ *where* κ_i *is a value of* κ *such that* H3 *holds for* $P'(\ell,\rho_i,Q_i)$ (*resp.* $P''(\ell,\rho_i,Q_i)$). *For* $i=1,2$, *if* x_i^* *is the solution of* $P'(\ell,\rho_i,Q_i)$ (*resp. if* (x_i^*,y_i^*,z_i^*) *is a solution of* $P''(\ell,\rho_i,Q_i)$), *then* $\|x_2^*-x_1^*\|_1$ (*resp.* $\|\frac{x_2^*-x_1^*}{\boldsymbol{e}^\top x^-+x_0^-}\|_1$) *is bounded from above by*

$$\frac{\|Q_2-Q_1\|_\infty}{2\,\beta(Q_1)}+\frac{\sqrt{\|Q_2-Q_1\|_\infty^2+\frac{2}{\kappa}(\|Q_1\|_\infty+\|Q_2\|_\infty)\beta(Q_1)\|\rho_2-\rho_1\|_\infty}}{2\,\beta(Q_1)},\tag{7}$$

and $\|x_2^*-x_1^*\|_2$ (*resp.* $\|\frac{x_2^*-x_1^*}{\boldsymbol{e}^\top x^-+x_0^-}\|_2$) *is bounded from above by*

$$\frac{\max_i\|C_i(Q_2-Q_1)\|_2}{2\,\lambda_{\min}(Q_1)}+\frac{\sqrt{\max_i\|C_i(Q_2-Q_1)\|_2^2+\frac{2}{\kappa}(\|Q_1\|_\infty+\|Q_2\|_\infty)\lambda_{\min}(Q_1)\|\rho_2-\rho_1\|_\infty}}{2\,\lambda_{\min}(Q_1)}.\tag{8}$$

Upper bound (7) (*resp.* (8)) *is valid replacing* $\beta(Q_1)$ (*resp.* $\lambda_{\min}(Q_1)$) *by* $\beta(Q_2)$ (*resp.* $\lambda_{\min}(Q_2)$).

Smaller upper bounds, though more involved, are given in the Appendix in the proof of this theorem. The following result is then a corollary of this theorem.

Corollary 3.1 *Consider problem* $P'(\ell,\rho_1,Q_1)$ (*resp.* $P''(\ell,\rho_1,Q_1)$) *and its perturbed version* $P'(\ell,\rho_2,Q_2)$ (*resp.* $P''(\ell,\rho_2,Q_2)$). *Let Assumptions* H1, H2 *and* H3 *hold for these problems. For* $i=1,2$, *if* x_i^* *is the solution of* $P'(\ell,\rho_i,Q_i)$ (*resp. if* (x_i^*,y_i^*,z_i^*) *is a solution of* $P''(\ell,\rho_i,Q_i)$), *then* $\|x_2^*-x_1^*\|_2$ (*resp.* $\|\frac{x_2^*-x_1^*}{\boldsymbol{e}^\top x^-+x_0^-}\|_2$) *is bounded from above by*

$$\frac{\max_i\|C_i(Q_2-Q_1)\|_2}{\max(\lambda_{\min}(Q_1),\lambda_{\min}(Q_2))}+\frac{\sqrt{(\|Q_1\|_\infty+\|Q_2\|_\infty)\|\rho_2-\rho_1\|_\infty}}{\sqrt{2\kappa}\max(\lambda_{\min}(Q_1),\lambda_{\min}(Q_2))}.\tag{9}$$

Proposition 4.37, p. 291 of [3] gives a local sensitivity analysis for a generic optimization problem where both the objective function and the feasible set vary. If $C(\ell,\rho)$ is the feasible set of $P'(\ell,\rho,Q)$ or $P''(\ell,\rho,Q)$, the upper bound provided for $\|x_2^*-x_1^*\|$ by this proposition depends on the Hausdorff distance $\mathrm{Haus}(C(\ell,\rho_1),C(\ell,\rho_1)\cap C(\ell,\rho_2))$. Using Hoffman bound [10] yields an upper bound of the kind $\tau(\rho_1,\rho_2)\|\rho_2-\rho_1\|$ for the Hausdorff distance, but since $\tau(\rho_1,\rho_2)$ is unknown, the bound is still not explicit and local. For problem P', the (strong) Slater assumption implies Robinson's constraint qualification. Proposition 4.41 of [3] can thus be applied to get

$$\exists K>0,\ \text{such that}\quad \mathrm{Haus}(C(\ell,\rho_1),C(\ell,\rho_1)\cap C(\ell,\rho_2))\le K\|\rho_2-\rho_1\|,$$

but here again K is not explicit and the analysis is local.

We can extend the results of this section to study the sensitivity analysis of such quadratic optimization problems:

$$\begin{cases}\min \frac{1}{2}x^\top Qx+c^\top x\\ x^\top f_j=b_j,\quad j=1,\dots,m_1,\\ x\in X,\end{cases}$$

where X is a nonempty closed convex set and the parameters $f_j, j=1,\dots,m_1$, c in $\mathbb{R}^n$, $b\in\mathbb{R}^{m_1}$ and $Q\succ 0$ are parameters of problems from this class. We assume that the set X can be described by a set of inequalities of the kind $h_j(x)\le 0$, $j=1,\dots,m_2$ with given convex differentiable functions h_j. We also suppose that there exists $M>0$ such that for all $x\in X$ and every j, $\|\nabla h_j(x)\|_\infty\le M$. No equality constraints describe the set X and we suppose the Slater assumption holds. In this case, as was done for Theorem 3.2, we can introduce the dual problem obtained by dualizing the constraints $x^\top f_j=b_j$, $j=1,\dots,m_1$, bound from above the optimal Lagrange multipliers and give an explicit and global bound for $\|x_2(Q_2,c_2,f_1^2,\dots,f_{m_1}^2,b_1^2,\dots,b_{m_1}^2)-x_1(Q_1,c_1,f_1^1,\dots,f_{m_1}^1,b_1^1,\dots,b_{m_1}^1)\|_1$.

4 Stable calibration of the covariance matrix

This section focuses on stable calibrations of the covariance matrix of stock returns. We first explain what we mean by stable calibration and justify this objective.

4.1 Motivations

We can view the portfolio selection step as a black box taking as inputs the mean return vector and the covariance matrix, and providing as an output a portfolio. The composition of the portfolio will be stable with respect to the inputs if small perturbations of these inputs produce small changes in the portfolio composition. In particular, small perturbations in the observations of the returns which induce estimations of the mean return and covariance matrix satisfying hypotheses H1, H2 and H3, should result in small perturbations in the selected portfolio. Such a behavior is especially of interest for three basic reasons:

- First, it is interesting per se, as portfolio managers prefer stable portfolios: the portfolios obtained using closed values $(\hat{\rho}_1, \hat{Q}_1)$ and $(\hat{\rho}_2, \hat{Q}_2)$ of the estimated parameters should be close.
- Second, if the inputs we use are close to the true unknown inputs, and if the selection step is stable, the composition of the portfolio it produces should be close to that of the true (unknown) optimal portfolio.
- Finally, when portfolios are rebalanced, the more stable the composition is, the less the transaction costs.

We start with some observations useful for all the stabilization methods we introduce next.

4.2 Preliminary observations

Stability for $\tilde{P}(k,\rho,Q)$ If short sellings are allowed for $P(k,\rho,Q)$, we obtain problem $\tilde{P}(k,\rho,Q)$, and from Lemma 2.3, the optimal solution is $x^*(k,\rho,Q) = kQ^{-1}(\rho-\rho_0\mathbf{e})$ which implies $\|x^*(k,\rho,Q)\|_2 \le \frac{k\|\rho-\rho_0\mathbf{e}\|_2}{\lambda_{\min}(Q)}$. Thus if $\lambda_{\min}(Q) \ge \frac{k\|\rho-\rho_0\mathbf{e}\|_2}{r}$ for some $0<r<1$, then $x^*(k,\rho,Q) \in \mathcal{B}(0,r) = \{x \mid \|x\|_2 \le r\}$. In particular, if $\lambda_{\min}(Q_1) \ge \frac{k\|\rho_1-\rho_0\mathbf{e}\|_2}{r}$ and $\lambda_{\min}(Q_2) \ge \frac{k\|\rho_2-\rho_0\mathbf{e}\|_2}{r}$, then $x_1 \in \mathcal{B}(0,r)$, $x_2 \in \mathcal{B}(0,r)$, and $\|x_2-x_1\|_2 \le 2r$. If ρ is bounded and M is such that $\|\rho-\rho_0\mathbf{e}\|_2 \le M$, then if $\lambda_{\min}(Q_1) \ge \frac{kM}{r}$ and $\lambda_{\min}(Q_2) \ge \frac{kM}{r}$, we have $x_1 \in \mathcal{B}(0,r)$ and $x_2 \in \mathcal{B}(0,r)$. Increasing sufficiently the smallest eigenvalue of the covariance matrix thus appears as a way of stabilizing the selection step for $\tilde{P}(k,\rho,Q)$. More precisely, if this smallest eigenvalue is greater than $\frac{kM}{r}$, for some $0<r<1$, we enforce the solutions to stay in the ball $\mathcal{B}(0,r)$. In particular, this forbids any component of x to be greater than r.

Stability for $\tilde{P}'(\ell,\rho,Q)$ If short sellings are allowed for $P(\ell,\rho,Q)$, we obtain problem $\tilde{P}(\ell,\rho,Q)$ and using Lemma 2.3 we obtain the bound $\|x^*(\ell,\rho,Q)\|_2 \le \frac{\ell-\rho_0}{\|\rho-\rho_0\mathbf{e}\|_2}\frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)}$ for the optimal solution $x^*(\ell,\rho,Q)$. If κ in hypothesis H3 for $P(\ell,\rho,Q)$ is sufficiently large and if the condition number of Q is sufficiently small, more precisely if

$$\kappa \ge \frac{(\ell-\rho_0)(1-r)}{r} > 0 \quad \text{and} \quad \frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} \le \left(\frac{\ell-\rho_0+\kappa}{\ell-\rho_0}\right) r,$$

for some $0<r<1$, then $x^*(\ell,\rho,Q) \in \mathcal{B}(0,r)$. However, since H2 holds, we will never have $x^*(\ell,\rho,Q)=0$.

Stability for $P(k,\rho,Q)$ For $P(k,\rho,Q)$, if the mean return vector is bounded i.e., if $\|\rho_1\|_2 \le M$ and $\|\rho_2\|_2 \le M$, then using (6), if Q is fixed and such that $\lambda_{\min}(Q) \ge \frac{4kM}{r}$, for some $0<r<1$, we have $\|x_2^*-x_1^*\|_2 \le r$ and we guarantee stability. More generally, if I_n is the $n\times n$ identity matrix, we have $\lim_{\lambda\to\infty}\|x(k,\rho,Q+\lambda I_n)\|_2 = 0$. Thus for any $0<r<1$, we can find $\lambda_0(\rho,Q)>0$ such that if $\lambda \ge \lambda_0(\rho,Q)$ then $x(k,\rho,Q+\lambda I_n) \in \mathcal{B}(0,r)$. Since $\lambda_{\min}(Q+\lambda I_n) = \lambda_{\min}(Q)+\lambda$, increasing this way the smallest eigenvalue of Q (replacing Q by $Q+\lambda I_n$, for λ chosen sufficiently large) thus yields stability for P.

Stability for $P'(\ell,\rho,Q)$ and $P''(\ell,\rho,Q)$ For problem P' (resp. P''), we have for $\|x_2^*-x_1^*\|_2$ (resp $\|\frac{x_2^*-x_1^*}{\mathbf{e}^\top x^-+x_0^-}\|_2$), the upper bound (9). The first term in this upper bound (9) can be arbitrarily small for perturbations of the covariance matrix of a given range ($\max_i \|C_i(Q_2-Q_1)\|_2 \le k$ for some fixed $k>0$) and increasing sufficiently the smallest eigenvalue of Q_1 or Q_2 (for instance for diagonal matrices Q_1 and $Q_2 = Q_1+\varepsilon I_n$, with $\lambda_{\min}(Q_1)$ sufficiently large). However, since for any matrix Q, we have $\|Q\|_\infty \ge \frac{\lambda_{\min}(Q)}{n}$, the second term in (9) is bounded from below by $\sqrt{\frac{\|\rho_2-\rho_1\|_\infty}{2\kappa n}}$, which can be large for large perturbations of ρ. A way to allow the second term in (9) to be small is to choose κ large enough and to consider perturbations of the mean return of a given range ($\|\rho_2-\rho_2\|_2 \le k$ for some fixed $k>0$). For the parameter κ to have a significant value, at least one mean return must have a value significantly larger than the target return ℓ, or, equivalently, the target return ℓ must be chosen significantly smaller than at least one mean return (while being larger than ρ_0).

Remark 4.1 The observations above indicate that under hypotheses H1, H2 and H3, to stabilize the selection steps $\tilde{P}(k,\rho,Q)$, $\tilde{P}'(\ell,\rho,Q)$, and $P(k,\rho,Q)$ the smallest eigenvalue of the covariance matrix Q should have a significant value. For models $P'(\ell,\rho,Q)$ and $P''(\ell,\rho,Q)$, to obtain stability, we should choose κ sufficiently large, take a large value for the smallest eigenvalue of the covariance matrix, and consider small perturbations.

In Sect. 2.3, we underlined the degeneracy of the empirical and adaptive estimations of the covariance matrix. In [4], it is also shown that the smallest eigenvalues of the empirical covariance matrix are underestimated. The above Remark 4.1 combined with these observations indicate that the empirical and adaptive estimations should not only be corrected for stability but also to avoid numerical problems and obtain more relevant statistical estimations.

It can be noticed that the recommendations of Remark 4.1 impose for P' and P'' conditions on the mean return vector through hypotheses H2 and H3 (where in particular κ is involved). We now intend to propose ways of exploiting the recommendations made in this remark on the covariance matrix. The general idea is to look for a matrix close to $\hat{Q}$ that enhances the stability properties of the model. A compromise will also have to be found between efficiency and stability.

4.3 Closest covariance matrix to $\hat{Q}$

In [11], they provide a consistent estimation of the parameter α^* such that $\alpha^* F + (1-\alpha^*)\hat{Q}$ (where F is a single-index covariance matrix and $\hat{Q}$ is the empirical covariance matrix) is the closest matrix to the matrix Q. In [8], they compute the nearest correlation matrix to the empirical covariance matrix.

We also propose to look for the closest covariance matrix to the matrix $\hat{Q}$ (the empirical or adaptive) but additionally requiring this matrix to satisfy three constraints ensuring, in particular, that the resulting matrix is positive definite. To introduce these constraints, we need the Frobenius scalar product $\langle .,. \rangle$ defined by

$$\forall X, Y \in \mathcal{S}_n(\mathbb{R}), \quad \langle X, Y\rangle = Tr(XY),$$

where $Tr(X)$ is the trace of the matrix X. The first constraint $X \succeq \alpha I$, with $\alpha > 0$, is equivalent to $\lambda_{\min}(X) \geq \alpha$. The parameter α represents an arbitrary threshold for the smallest eigenvalue of the estimated covariance matrix. This constraint is thus a way of exploiting Remark 4.1. In particular, it guarantees that the smallest eigenvalue of the calibrated covariance matrix is positive as the assumption of arbitrage free markets require. The second constraint $\langle I_n, X\rangle = \langle I_n, \hat{Q}\rangle$, ensures the conservation of the empirical or "adaptive" total risk. Finally, we choose m portfolios $q_i, i = 1, \dots, m$. We can estimate the variance $\hat{\sigma}_i^2$ of the portfolio q_i return and require that $\hat{\sigma}_i^2$ is equal to the estimation $q_i^\top X q_i$ of the variance of the portfolio q_i return, obtained using the covariance matrix X. If we suppose the return process is stationary, all the data will be needed to compute $\hat{\sigma}_i^2$. Under local time homogeneity only the data of the homogeneity interval is used. This yields the following problem:

$$\begin{cases} \min \|X - \hat{Q}\|_F \\ \langle I_n, X\rangle = \langle I_n, \hat{Q}\rangle, & \text{(a)} \\ \langle q_i q_i^\top, X\rangle = \hat{\sigma}_i^2, \quad i = 1, \dots, m, & \text{(b)} \\ X \succeq \alpha I, & \text{(c)} \end{cases} \tag{10}$$

where for $X \in \mathcal{S}_n(\mathbb{R})$, $\|X\|_F$ denotes the Frobenius norm of X, i.e., $\|X\|_F = \sqrt{\langle X, X\rangle} = \sqrt{Tr(X^2)}$. This problem can be expressed as a quadratic-semidefinite program and solved via interior point methods ([14] for instance).

In what follows, this method of correction of the matrix $\hat{Q}$ will be called C_1. We can also consider particular cases of this method. If the constraints (a) and (b) are removed (calibration C_2) and if the spectral decomposition of $\hat{Q}$ is $\hat{Q} = \sum_{i=1}^n \lambda_i(\hat{Q}) v_i v_i^\top$, where v_i is the i-th eigenvector of the matrix $\hat{Q}$ associated to the eigenvalue $\lambda_i(\hat{Q})$, then the solution of problem (10) is $X = \sum_{i=1}^n \max(\lambda_i(\hat{Q}), \alpha) v_i v_i^\top$. Another particular case where we have an explicit solution is the case where (a) is removed, $\alpha = 0$ and the portfolios chosen for the constraints (b) constitute an orthonormal basis of eigenvectors of the matrix $\hat{Q}$ (calibration C_3).

Proposition 4.1 *Consider optimisation problem* (10) *where* (a) *is removed*, $m = n$ *is the dimension of the matrix* $\hat{Q}$, $\alpha = 0$ *and the vectors* q_i *constitute an orthonormal basis of eigenvectors of the matrix* $\hat{Q}$. *Then the solution of* (10) *is given by*: $X^* = \sum_{i=1}^n \hat{\sigma}_i^2 q_i q_i^\top$.

Proof The Slater hypothesis being satisfied, $(X^*, Z^*, (\mu_i^*)_{1\leq i\leq n})$ constitutes a primal-dual solution of problem (10) if and only if:

$$\begin{cases} X^* \succeq 0, \quad Z^* \succeq 0, \quad \langle X^*, Z^*\rangle = 0, & \text{(a}'\text{)} \\ \langle q_i q_i^\top, X^*\rangle = \hat{\sigma}_i^2, & \text{(b}'\text{)} \\ X^* = \hat{Q} + Z^* - \sum_{i=1}^n \mu_i^* q_i q_i^\top. & \text{(c}'\text{)} \end{cases}$$

Conditions (a′) give $X^* Z^* = 0$ and since $X^* \succ 0$, we have $Z^* = 0$. Condition (c′) is thus satisfied with $\mu_i^* = \lambda_i(\hat{Q}) - \hat{\sigma}_i^2$ where $\lambda_i(\hat{Q})$ is the eigenvalue of the matrix $\hat{Q}$

associated to the eigenvector q_i. Finally, (b′) is satisfied:

$$\langle q_i q_i^\top, X^*\rangle = \sum_{j=1}^n \hat{\sigma}_j^2 Tr(q_j q_j^\top q_i q_i^\top) = \hat{\sigma}_i^2 Tr(q_i q_i^\top) = \hat{\sigma}_i^2 \|q_i\|_2^2 = \hat{\sigma}_i^2. \qquad \square$$

Remark 4.2 An interesting feature of the calibration in Proposition 4.1 is that in particular it corrects the estimation of the risk in directions where the risk is not well evaluated with $\hat{Q}$. These directions correspond to the eigenvectors associated to the smallest and highest eigenvalues.

Finally, we could also remove the constraints (b) from (10) (calibration C_4).

4.4 Maximizing the lowest eigenvalue

The calibrations introduced in the previous subsection depend on the choice of the parameter α and on the portfolios q_i. No natural choice seems to prevail for these parameters. In this section, we instead intend to present a systematic calibration of the covariance matrix. This calibration uses additional statistical information and more directly exploits the results of Sect. 3 to allow for stability.

The statistical information (coming from [7]) provides functions $\eta_\rho(\lambda, n, T)$ and $\eta_Q(\lambda, n, T)$ such that the events

$$\|\hat{\rho} - \rho\|_\infty \le \eta_\rho(\lambda, n, T) \quad \text{and} \quad \|\hat{Q} - Q\|_\infty \le \eta_Q(\lambda, n, T) \tag{11}$$

hold with probabilities functions of a positive parameter λ, of the number of risky assets n and of the number of observations T used for estimation. With a slight abuse of notation, in (11) we have used for the estimators of the mean and of the covariance matrix the same notation as the estimations. Parameter λ can be chosen in such a way that the probability that (11) holds is arbitrarily high [7]. Our idea is then to use this information and Remark 4.1 to maximize the lowest eigenvalue of Q using the box constraints on the covariance matrix given in (11). The quantity $\eta_Q(\lambda, n, T)$ is thus chosen in such a way that with a large probability the event $\|\hat{Q} - Q\|_\infty \le \eta_Q(\lambda, n, T)$ holds. This way, the set

$$E = \{Q \mid \|\hat{Q} - Q\|_\infty \le \eta_Q(\lambda, n, T)\}, \tag{12}$$

where $\hat{Q}$ is the empirical (or adaptive) estimation of the covariance matrix, is a confidence area for the covariance matrix Q with a given confidence level. The quantity $\eta_Q(\lambda, n, T)$ can also be seen as a user defined parameter that would control the size of the search zone around $\hat{Q}$.

Since Q is a covariance matrix, we also impose $Q \succeq 0$. Hence we come to the following problem:

$$\begin{cases} \max \lambda_{\min}(Q) \\ \|Q - \hat{Q}\|_\infty \le \eta_Q(\lambda, n, T), \quad Q \succeq 0. \end{cases} \tag{13}$$

This is a nondifferentiable convex optimization problem. We transform it into the SDP program (14) below which can be efficiently solved with interior point methods:

$$\begin{cases} \min \ (-u) \\ V(i,j) + u\delta_{ij} + Y(i,j) = \eta_Q(\lambda,n,T) + \hat{Q}(i,j), \\ W(i,j) - u\delta_{ij} - Y(i,j) = \eta_Q(\lambda,n,T) - \hat{Q}(i,j), \\ V(i,j) \geq 0, \quad W(i,j) \geq 0, \quad Y \succeq 0, \end{cases} \tag{14}$$

where δ_{ij} is the Kronecker symbol. The covariance matrix Q is then given by $Y^* + u^* I$ with Y^* and u^* the optimal values of Y and u in (14). We will denote by C_5 this calibration of the covariance matrix.

4.5 Best condition number

We saw in Sect. 4.2 that for stability in problem $\tilde{P}'(\ell,\rho,Q)$, it is desirable to have a small condition number for the estimated covariance matrix. Moreover, it is noticed in [4] that the largest eigenvalues of the empirical covariance matrix are overestimated and the lowest underestimated (and it is also the case of the adaptive estimation), yielding to a large condition number. We can thus try to find the best condition number for the covariance matrix, while imposing the same box constraints as before on the components of this matrix. The covariance matrix Q thus solves:

$$\begin{cases} \min \frac{\lambda_{\max}(Q)}{\lambda_{\min}(Q)} \\ \|Q - \hat{Q}\|_\infty \leq \eta_Q(\lambda,n,T), \quad Q \succeq 0, \end{cases} \tag{15}$$

where we recall that $\eta_Q(\lambda,n,T)$ is such that E defined in (12) is a confidence area for Q with a given confidence level. The above problem (15) is a quasiconvex problem. It is equivalent to solve:

$$\begin{cases} \min t \\ s \leq \lambda_{\min}(Q), \\ v \geq \lambda_{\max}(Q), \\ v \leq ts, \\ \|Q - \hat{Q}\|_\infty \leq \eta_Q(\lambda,n,T), \quad Q \succeq 0. \end{cases} \tag{16}$$

We can then find a solution of this problem by dichotomy.

5 Numerical results

5.1 Stability tests

The goal of this section is to illustrate, via simulations on real data (the 30 assets of the Dow Jones), the influence of the increase of the smallest eigenvalue of the empirical or adaptive covariance matrix on the sensitivity of the composition of the

Table 1 Condition number of the solution Q^* of problems (14) and (16) for fixed $\hat{Q}$ and different values of η_Q

Method	η_Q^1	η_Q^2	η_Q^3
C_5	80.24	16.25	6.91
Min Cond	70.85	9.06	2.29

portfolios. We also compare the behaviors of the optimal portfolios obtained using the empirical covariance matrix or the adaptive covariance matrix $\hat{Q}$ and their corrections C_2 and C_5. The Markowitz problem (1) was solved using the Mosek optimization library and optimization problem (13) using the SeDuMi library.

5.1.1 Reducing the condition number

We first illustrate the magnitude of the condition number reduction using the calibrations introduced in Sects. 4.4 and 4.5. We choose an empirical covariance matrix $\hat{Q}$ with condition number 1.11×10^6. We then compute the condition number of different matrices Q solutions of (14) (calibration C_5) and (16) (calibration denoted by "Min Cond") for the following values of η_Q: $\eta_Q^1 = 0.01\lambda_{\max}(\hat{Q})$, $\eta_Q^2 = 0.05\lambda_{\max}(\hat{Q})$, and $\eta_Q^3 = 0.1\lambda_{\max}(\hat{Q})$. The results are reported in Table 1.

The condition number thus significantly decreases even if only small variations of the entries of $\hat{Q}$ are allowed. Both calibrations yield close condition numbers in this example.

5.1.2 Evolution of the portfolio composition in time

To observe the influence of the increase of $\lambda_{\min}(\hat{Q})$ on the behavior of the portfolios, we conduct the following experiment: A first investment is done on January 2, 1999 (we denote this date by t_0); the investment horizon is 60 days, the yearly risk-free rate is 5% and the target return for these 60 days is $\ell = 2.5\%$. The portfolio is then regularly rebalanced every 60 days for dates $t_j = t_0 + 60j$, $j = 1, \ldots, 11$. For each investment date t_j, the empirical estimations $\hat{\rho}_j$ and $\hat{Q}_j$ of the mean and of the covariance matrix are computed. We want to analyse the influence of the parameter α of the method C_2 on the stability of the composition of the portfolios. At each date t_j, we compute the correction of the matrix $\hat{Q}_j$ using calibration C_2 and the values $\alpha_j(i)$ of α given by $\alpha_j(i) = 10^{i-7}\lambda_{\max}(\hat{Q}_j)$ for $i = 1, \ldots, 6$. Let $\hat{Q}_j^i$ be the correction of matrix $\hat{Q}_j$ for the value $\alpha_j(i)$ of α. We denote by x_j^i the solution of problem $P'(\ell, \hat{\rho}_j, \hat{Q}_j^i)$. We then compute

$$p(i) = \frac{1}{11}\sum_{j=0}^{10} \|x_{j+1}^i - x_j^i\|_1.$$

The evolution of $p(i)$ with i is shown in Fig. 1 which follows.

Hence, the increasing of $\lambda_{\min}(\hat{Q})$ tends to stabilize the composition of the portfolios in this example. This has in fact been observed using different starting dates t_0, different target returns and different risk-free rates.

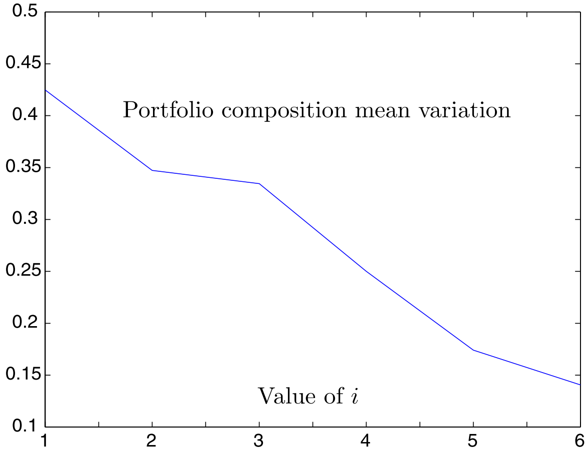

Fig. 1 Sensitivity of the portfolio composition mean variation as a function of α

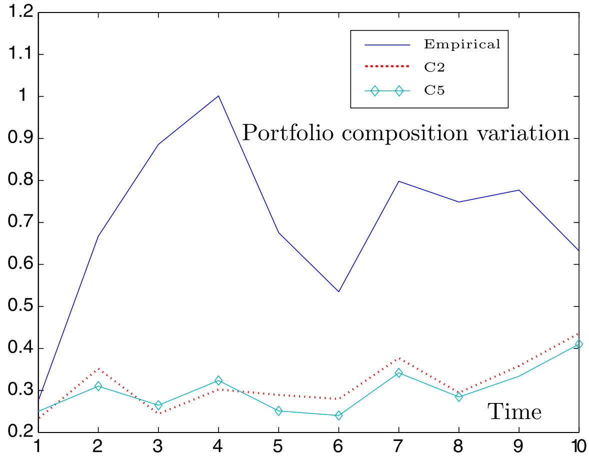

Fig. 2 Variation of the portfolio composition in time for three different calibration methods

We now compare the "Empirical," C_2 and C_5 methods. We call "Empirical," the method using the empirical estimations of the parameters. If $\hat{Q}$ is the empirical covariance matrix, we choose $\alpha = 0.01\lambda_{\max}(\hat{Q})$ for method C_2, and $\eta_Q = \alpha$ for method C_5. The date of the first investment is January 2, 1999 (date denoted by t_0), the investment horizon is still 60 days, the target return is 4%, and the yearly risk-free rate is 5%. The portfolios are regularly rebalanced every 60 days from t_0. For the i-th rebalancing, we determine a portfolio x_M^i for each method M. Figure 2 represents the evolution of $(\|x_M^i - x_M^{i-1}\|_1)_{i\geq 2}$ as a function of i and for each method. This experiment also tends to show that the increase in $\lambda_{\min}(\hat{Q})$ permits the stability of the portfolio composition. The C_2 and C_5 methods seem to be particularly stable in this example. For these methods, the modification of the composition of the optimal portfolio is always less important than the "Empirical" method. The same experiment was conducted using different values for the parameters of the Markowitz model. We used different starting dates t_0, different investment horizons (60 and 40 days) and

Table 2 Portfolio composition mean variation when the mean returns change

"Empirical"	C_2	C_5
0.0119	0.0060	0.0058

different target returns (2, 3 and 4%). In all the simulations, the C_2 and C_5 methods were the most stable, always leading to less important modifications of the portfolio composition than the "Empirical" method.

5.1.3 *Influence of the perturbations of the mean returns on the optimal portfolio composition*

We fix a date t_0 (January 2, 1999) and for each method M ($M =$ "Empirical", C_2, C_5), we estimate (ρ, Q) by $(\hat{\rho}, \hat{Q}_M)$ [$\hat{\rho}$ is the empirical mean of the returns and $\hat{Q}_M$ is the estimation of the covariance matrix using method M]. From these estimations, we can compute the optimal portfolio x_M associated with method M and using model P'. We then make n (i.e. 30) iterations. At iteration i, we envisage four perturbations which consist of replacing $\hat{\rho}(i)$ by $\hat{\rho}(i) \pm 0.05|\hat{\rho}(i)|$, $\hat{\rho}(i) \pm 0.1|\hat{\rho}(i)|$. At iteration i, each perturbation j produces a portfolio x_M^{ij} for method M. A comparison of $\frac{1}{30*4}\sum_{i,j} \|x_M - x_M^{ij}\|_1$ can then be made for all methods M. This experiment was repeated 400 times (using an increasing number of historical data) and gave the average results given in Table 2. We observe that the perturbation of ρ does not change the composition of the portfolio much in these cases. Method C_5 is the most stable with respect to perturbations of the mean return vector in this experiment.

5.2 Diversification of the portfolios

We noticed on various simulations that the use of the corrected covariance matrices tends to diversify the portfolios much more than if the empirical or adaptive covariance matrix was used. To obtain diversified portfolios, portfolio managers traditionally introduce box constraints on the components of the portfolio. It is interesting to notice that corrections C_1 and C_3 seem to provide diversified portfolios without changing the constraints of the problem.

5.3 Comparison of the calibrations of the covariance matrix on real data

We compute the optimal portfolios which would have been obtained by investing in the assets of the Dow Jones from January 2, 1995 to June 30, 2004 and rebalancing the portfolio every H days. The yearly risk-free rate is 1%, the transaction costs are 0.5% and the yearly target return is $\ell = 10\%$. We measure the influence of the corrections of the adaptive covariance matrix (see Sect. 2.3) introduced in Sect. 4. The parameters of the adaptive method are chosen a posteriori (see [7] for further details). The result of these experiments, conducted using different values of H, is given in Table 3. In this table, we call Rdt the return of a method over the investment period. $\bar{R}$ and σ are the empirical mean and standard deviation of the sample of the H day return of the portfolio. We notice that the corrections of the adaptive method tend to provide portfolios whose returns are larger and give standard deviations that are close to each other.

Table 3 Comparison of different calibrations of the covariance matrix using the assets of the Dow Jones (from January 1995 to June 2004), a risk-free asset and the Markowitz model P''

	$H = 15$ days			$H = 30$ days			$H = 60$ days		
Method	Rdt	$\bar{R}$	σ	Rdt	$\bar{R}$	σ	Rdt	$\bar{R}$	σ
Adaptive	2.47	1.0057	0.0184	2.4444	1.0113	0.0253	3.8672	1.0386	0.1082
C_1	2.63	1.0061	0.0210	2.8044	1.0131	0.0304	4.1250	1.0409	0.1138
C_2	2.50	1.0057	0.0184	2.5363	1.0117	0.0257	4.0898	1.0398	0.1045
C_3	2.64	1.0062	0.0257	2.7134	1.0130	0.0387	4.1549	1.0414	0.1152
C_4	2.52	1.0058	0.0183	2.5591	1.0118	0.0257	4.1487	1.0401	0.1044
C_5	2.58	1.0059	0.0185	2.6058	1.0121	0.0262	4.4440	1.0421	0.1075

6 Conclusion

We first introduced a sensitivity analysis for different versions of the Markowitz model. Using the quite general model given in [7] for the returns, we then proposed strategies to compute stable portfolios using the Markowitz model.

One of our calibrations of the covariance matrix (the one proposed in Sect. 4.4) has shown its efficiency numerically speaking, beating all the other methods in most of the stability tests done while providing performing portfolios. This calibration shows the importance of the condition number of the estimated covariance matrix. Indeed, a lowest eigenvalue of the covariance matrix close to 0 (as is the case for the adaptive covariance matrix) is absurd financially speaking, and yields numerical problems to solve the Markowitz problem. On the contrary, our proposed covariance matrices are not ill-conditioned: they are positive definite matrices as the constraints require.

Appendix

In this Appendix, we show Theorem 3.2. To show this theorem, we will make use of the following lemma:

Lemma A.1 *Let $f, g_1, \dots, g_m : \mathbb{R}^n \to \mathbb{R}$ be convex functions, and let X be a convex subset of $\mathbb{R}^n$. Let us consider the convex primal problem $\mathcal{P}$ below*

$$\mathcal{P} \quad \begin{cases} \min f(x) \\ g(x) \equiv (g_1(x), \dots, g_m(x)) \leq 0, \\ x \in X, \end{cases} \qquad \textit{and the dual problem } \mathcal{D} \quad \begin{cases} \max \theta(\lambda) \\ \lambda \geq 0, \end{cases}$$

where

$$\theta(\lambda) = \begin{cases} \min f(x) + \lambda^\top g(x) \\ x \in X. \end{cases} \tag{17}$$

Let the Slater condition hold for $\mathcal{P}$ (there exists $x \in X$ such that $g_j(x) < 0$, $j = 1, \dots, m$) and let us suppose that f is bounded from below on $\{x \mid g(x) \leq 0, x \in X\}$.

Let $S^*_{\mathcal{P}}$ and $S^*_{\mathcal{D}}$ be respectively the set of solutions of $\mathcal{P}$ and $\mathcal{D}$ and for fixed λ, let $S^*(\lambda)$ be the set of solutions of (17). Then for any $\lambda^* \in S^*_{\mathcal{D}}$, we have $S^*_{\mathcal{P}} \subset S^*(\lambda^*)$.

Proof Let us take $\lambda^* \in S^*_{\mathcal{D}}$. The hypotheses of the Convex Duality Theorem apply and for any $x^* \in S^*_{\mathcal{P}}$, the optimal value $f(x^*)$ of primal problem $\mathcal{P}$ and the optimal value $\theta(\lambda^*)$ of dual problem $\mathcal{D}$ coincide. Moreover, by definition of $\theta(\lambda^*)$, since $x^* \in X$, we have $\theta(\lambda^*) \le f(x^*) + g(x^*)^\top \lambda^*$. This gives $f(x^*) \le f(x^*) + g(x^*)^\top \lambda^*$, i.e., $g(x^*)^\top \lambda^* \ge 0$. But since $\lambda^* \ge 0$ and $g(x^*) \le 0$, this implies $g(x^*)^\top \lambda^* = 0$. We thus have, using once again the definition of $\theta(\lambda^*)$:

$$\theta(\lambda^*) = f(x^*) = f(x^*) + g(x^*)^\top \lambda^* \le f(x) + g(x)^\top \lambda^*, \quad \forall x \in X.$$

Since, $x^* \in X$, this shows that x^* is a minimizer of $f(x) + g(x)^\top \lambda^*$ over X, i.e., that $x^* \in S^*(\lambda^*)$. □

Proof of Theorem 3.2 For convenience, we use the notation $\bar{\rho}_1 = \rho_1 - \rho_0 \mathbf{e}$, $\bar{\rho}_2 = \rho_2 - \rho_0 \mathbf{e}$ and $\bar{\ell} = \ell - \rho_0$. For $i = 1, 2$, let x_i^* be the solution of $P'(\ell, \rho_i, Q_i)$. Let us first show that (7) and (8) are upper bounds for respectively $\|x_2^* - x_1^*\|_1$ and $\|x_2^* - x_1^*\|_2$.

Let $\lambda \in \mathbb{R}$, let

$$\theta_i(\lambda) = \begin{cases} \inf \frac{1}{2} x^\top Q_i x + \lambda(\bar{\ell} - x^\top \bar{\rho}_i) \\ x \in \Delta_n, \end{cases} \tag{18}$$

be the dual function of the problem $P'(\ell, \rho_i, Q_i)$ where only the uncertain constraint has been dualized, and let λ_i^* be an optimal solution of the dual problem consisting of solving $\max_{\lambda \in \mathbb{R}_+} \theta_i(\lambda)$. Both primal problem $P'(\ell, \rho_i, Q_i)$ and its dual problem are equivalent to each other and have the same optimal value. The hypotheses of Lemma A.1 hold for primal problem $P'(\ell, \rho_i, Q_i)$ and its dual problem. Since the objective function of $P'(\ell, \rho_i, Q_i)$ is strictly convex, the set of solutions of this problem is reduced to x_i^*. Also, for any fixed λ, since the objective function of problem (18) is strictly convex, the solution to (18) is unique and denoted by $x(\lambda)$. For problem $P'(\ell, \rho_i, Q_i)$, Lemma A.1 thus tells us that $x_i^* = x(\lambda_i^*)$. From the optimality of $x(\lambda_i^*) = x_i^*$, we then have for $i = 1, 2$:

$$\forall x \in \Delta_n, \quad (x - x_i^*)^\top (Q_i x_i^* - \lambda_i^* \bar{\rho}_i) \ge 0.$$

Since x_1^* and x_2^* are in Δ_n we can use the previous inequality for $x = x_2^*, i = 1$ and $x = x_1^*, i = 2$, which gives:

$$\begin{cases} (x_2^* - x_1^*)^\top (Q_1 x_1^* - \lambda_1^* \bar{\rho}_1) \ge 0 \\ (x_1^* - x_2^*)^\top (Q_2 x_2^* - \lambda_2^* \bar{\rho}_2) \ge 0. \end{cases} \tag{19}$$

Adding the inequalities (19) and rearranging the terms we get:

$$(x_2^* - x_1^*)^\top Q_1 (x_2^* - x_1^*) \le (x_2^* - x_1^*)^\top (Q_1 - Q_2) x_2^* + R \tag{20}$$

with $R = (x_2^* - x_1^*)^\top (-\lambda_1^* \bar{\rho}_1 + \lambda_2^* \bar{\rho}_2)$. Since for $i = 1, 2$, $x_i^{*\top} \bar{\rho}_i = \bar{\ell}$, we have $(x_2^* - x_1^*)^\top (-\lambda_1^* \bar{\rho}_1 + \lambda_2^* \bar{\rho}_2) = (\bar{\rho}_2 - \bar{\rho}_1)^\top (-\lambda_2^* x_1^* + \lambda_1^* x_2^*)$. Plugging this result in

(20) and observing that $\|x_1^*\|_1 \le 1$ and $\|x_2^*\|_1 \le 1$, we obtain:

$$\beta(Q_1)\|x_2^* - x_1^*\|_1{}^2 \le \|Q_2 - Q_1\|_\infty \|x_2^* - x_1^*\|_1 + \|\rho_2 - \rho_1\|_\infty(\lambda_1^* + \lambda_2^*). \quad (21)$$

It remains to bound the multipliers λ_i^*. First, we can bound from below the optimal value of $P'(\ell, \rho_i, Q_i)$ by 0, i.e., $\theta_i(\lambda_i^*) \ge 0$. Let e_j, $j = 1, \dots, n$, be the vectors of the canonical basis. From H3, for $i = 1, 2$, there exists $j_i \in 1, \dots, n$, such that $\rho_i(j_i) > \ell + \kappa$, with $\kappa > 0$. Since for $i = 1, 2$ we have $e_{j_i} \in \Delta_n$, by definition of the dual function, for $i = 1, 2$:

$$\forall \lambda \quad \theta_i(\lambda) \le \frac{1}{2} e_{j_i}{}^\top Q_i e_{j_i} + \lambda(\bar{\ell} - \bar{\rho}_i(j_i)). \quad (22)$$

Using (22) for $\lambda = \lambda_i^*$ and since $\theta_i(\lambda_i^*) \ge 0$, we have:

$$\kappa \lambda_i^* \le \lambda_i^*(\rho_i(j_i) - \ell) \le \frac{1}{2} Q_i(j_i, j_i) \le \frac{\|Q_i\|_\infty}{2}. \quad (23)$$

We thus have for λ_i^* the upper bound $\lambda_i^* \le \frac{\|Q_i\|_\infty}{2\kappa}$. If we plug these bounds for λ_1^* and λ_2^* in (21), we see that $P(\|x_2^* - x_1^*\|_1) \le 0$, P being the second-order polynomial defined by $P(x) = \beta(Q_1)x^2 - \|Q_2 - Q_1\|_\infty x - \frac{(\|Q_1\|_\infty + \|Q_2\|_\infty)}{2\kappa}\|\rho_2 - \rho_1\|_\infty$. Thus, $\|x_2^* - x_1^*\|_1$ is lower or equal to the largest root of P, which shows (7).

Exchanging x_1^*, ρ_1, Q_1 and x_2^*, ρ_2, Q_2, we then obtain for $\|x_2^* - x_1^*\|_1$ the upper bound (7) with $\beta(Q_1)$ replaced with $\beta(Q_2)$.

Let us now show that (8) is an upper bound for $\|x_2^* - x_1^*\|_2$. Using (20), the upper bound $\lambda_i^* \le \frac{\|Q_i\|_\infty}{2\kappa}$ for λ_i^*, and since $x_2^* \in \Delta_n$, we obtain:

$$\begin{aligned}&\lambda_{\min}(Q_1)^2 \|x_2^* - x_1^*\|_2^2 \\ &\quad \le \|x_2^* - x_1^*\|_2 \max_{x \in \Delta_n} \|(Q_2 - Q_1)x\|_2 + \frac{(\|Q_1\|_\infty + \|Q_2\|_\infty)}{2\kappa}\|\rho_2 - \rho_1\|_\infty.\end{aligned}$$

Using Lemma 3.1 we then see that $P(\|x_2^* - x_1^*\|_2) \le 0$ where $P(x) = \lambda_{\min}(Q_1)x^2 - \max_i \|C_i(Q_2 - Q_1)\|_2 x - \frac{(\|Q_1\|_\infty + \|Q_2\|_\infty)}{2\kappa}\|\rho_2 - \rho_1\|_\infty$ and we conclude as before.

However, we could have obtained smaller upper bounds, though more involved. These upper bounds could be obtained using the above proofs of (7) and (8) and using a smaller upper bound for λ_i^*. This upper bound for λ_i^* is obtained as follows.

We first improve the lower bound on the optimal value of $P'(\ell, \rho_i, Q_i)$. More precisely, we have for this optimal value, the lower bound $\frac{1}{2} y_i{}^\top Q_i y_i$ where y_i is the solution of the following relaxed problem:

$$\begin{cases} \min \frac{1}{2} y^\top Q_i y \\ \bar{\rho}_i^\top y = \bar{\ell}. \end{cases} \quad (24)$$

Hence we have:

$$\theta_i(\lambda_i^*) \ge \frac{1}{2} y_i^\top Q_i y_i. \quad (25)$$

Further, for $i = \{1, 2\}$, there can be various indexes j_i such that $\bar{\rho}_i(j_i) > \bar{\ell}$. We thus have for $i = \{1, 2\}$ and for every index j such that $\bar{\rho}_i(j) > \bar{\ell}$:

$$\forall \lambda \quad \theta_i(\lambda) \le \frac{1}{2} {e_j}^\top Q_i e_j + \lambda(\bar{\ell} - \bar{\rho}_i(j)). \tag{26}$$

Using (24) and (25) with $\lambda = \lambda_i^*$ one has:

$$\lambda_i^* \le \frac{1}{2} \min_{\rho_i(j) > \ell} \frac{1}{\rho_i(j) - \ell} (Q_i(j, j) - {y_i}^\top Q_i y_i). \tag{27}$$

The solution of (25) is given by $y_i = \frac{\bar{\ell}}{\bar{\rho}_i^\top {Q_i}^{-1} \bar{\rho}_i} {Q_i}^{-1} \bar{\rho}_i$. Finally, plugging this expression of y_i into (27) gives the following improved upper bound for λ_i^*:

$$\lambda_i^* \le \frac{1}{2} \min_{\rho_i(j) > \ell} \frac{1}{\rho_i(j) - \ell} \left(Q_i(j, j) - \frac{\bar{\ell}^2}{\bar{\rho}_i^\top {Q_i}^{-1} \bar{\rho}_i} \right).$$

If (x_i^*, y_i^*, z_i^*) is a solution of $P''(\ell, \rho_i, Q_i)$, we now show that (7) and (8) are upper bounds for respectively $\| \frac{x_2^* - x_1^*}{\mathbf{e}^\top x^- + x_0^-} \|_1$ and $\| \frac{x_2^* - x_1^*}{\mathbf{e}^\top x^- + x_0^-} \|_2$.

The feasible set of P'' is the intersection of the hyperplane defined by the return constraint (this constraint is active, see Lemma 2.2) and a set defined by the remaining constraints that we will denote by $Y(\mu, \nu, x^-)$. Let here $\bar{\ell} = \ell\,(\mathbf{e}^\top x^- + x_0^-) - \rho_0 x_0^-$, let

$$W = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

be the vector of decision variables, let W_i^* be a solution of $P''(\ell, \rho_i, Q_i)$, let $\lambda \in \mathbb{R}$, and let

$$\theta_i(\lambda) = \begin{cases} \inf \frac{1}{2} x^\top Q_i x + \lambda \left(\bar{\ell} - x^\top \rho_i - \rho_0 \,(\mathbf{e} - \mu)^\top y + \rho_0 \,(\mathbf{e} + \nu)^\top z \right) \\ W = (x, y, z)^\top \in Y(\mu, \nu, x^-), \end{cases} \tag{28}$$

be the dual function of problem $P''(\ell, \rho_i, Q_i)$ where only the return constraint has been dualized. Let us also introduce the dual problem $\max_{\lambda \ge 0} \theta_i(\lambda)$. Primal problem $P''(\ell, \rho_i, Q_i)$ and its dual are equivalent to each other and have the same optimal value. Also, using Lemma A.1 (whose hypotheses are satisfied for P''), there is an optimal solution λ_i^* to the dual problem and a solution $W(\lambda_i^*)$ to problem (28) for $\lambda = \lambda_i^*$, such that $W_i^* = W(\lambda_i^*)$. From the optimality of $W(\lambda_i^*)$, we get:

$$\forall W = (x, y, z)^\top \in Y(\mu, \nu, x^-), \quad (W - W_i^*)^\top \begin{pmatrix} Q_i\, x_i^* - \lambda_i^* \rho_i \\ \lambda_i^* \rho_0 (\mu - \mathbf{e}) \\ \lambda_i^* \rho_0 (\nu + \mathbf{e}) \end{pmatrix} \ge 0.$$

Using the previous inequality for $W = W_2^*, i = 1$ and $W = W_1^*, i = 2$, we get:

$$\begin{cases} (x_2^* - x_1^*)^\top (Q_1\, x_1^* - \lambda_1^* \rho_1) + \lambda_1^* \rho_0 \left((y_2^* - y_1^*)^\top (\mu - \mathbf{e}) + (\nu + \mathbf{e})^\top (z_2^* - z_1^*) \right) \ge 0 \\ (x_1^* - x_2^*)^\top (Q_2\, x_2^* - \lambda_2^* \rho_2) + \lambda_2^* \rho_0 \left((y_1^* - y_2^*)^\top (\mu - \mathbf{e}) + (\nu + \mathbf{e})^\top (z_1^* - z_2^*) \right) \ge 0. \end{cases}$$

Adding the two previous inequalities and rearranging the terms we get:

$$(x_1^* - x_2^*)^\top Q_1 (x_1^* - x_2^*) \\ \leq (x_2^* - x_1^*)^\top (Q_1 - Q_2) x_2^* + (x_2^* - x_1^*)^\top (\lambda_2^* \rho_2 - \lambda_1^* \rho_1) + M, \qquad (29)$$

with

$$M = \rho_0 (\lambda_1^* - \lambda_2^*)((y_2^* - y_1^*)^\top (\mu - \mathbf{e}) + (z_2^* - z_1^*)^\top (\nu + \mathbf{e})).$$

Since the return constraint is active, we have, for $i = 1, 2$,

$$x_i^{*\top} \rho_i + \rho_0 (x_0^- + (\mathbf{e} - \mu)^\top y_i^* - (\nu + \mathbf{e})^\top z_i^*) = \ell (\mathbf{e}^\top x^- + x_0^-).$$

Thus, $M = (\lambda_1^* - \lambda_2^*)(x_2^{*\top} \rho_2 - x_1^{*\top} \rho_1)$. Plugging this result in (29) and observing that for any $W = (x, y, z)^\top \in Y(\mu, \nu, x^-)$ we have $\|x\|_1 \leq \mathbf{e}^\top x^- + x_0^-$, (which implies $\|x_i^*\|_1 \leq \mathbf{e}^\top x^- + x_0^-$ for $i = 1, 2$), we then have:

$$\beta(Q_1) \|x_2^* - x_1^*\|_1^2 \\ \leq (\|x_2^* - x_1^*\|_1 \|Q_2 - Q_1\|_\infty + (\lambda_1^* + \lambda_2^*) \|\rho_2 - \rho_1\|_\infty)(\mathbf{e}^\top x^- + x_0^-). \quad (30)$$

It remains to bound from above the Lagrange multipliers λ_i^*. We can bound from below the optimal value of $P''(\ell, \rho_i, Q_i)$ by 0. Thus, we have $\theta_i(\lambda_i^*) \geq 0$. From hypothesis H3, for $i = 1, 2$ there exists j_i such that $\rho_i(j_i) > \frac{(1+\nu_{j_i})}{(\mathbf{e}-\mu)^\top x^- + x_0^-}(\ell + \kappa)(\mathbf{e}^\top x^- + x_0^-)$. Let $\varepsilon > 0$ and let us then introduce for $i = 1, 2$, the point $W_i = (x_i, y_i, z_i)^\top \in Y(\mu, \nu, x^-)$ defined replacing i by j_i in (2). We thus have, $x_i = x^- - y_i + z_i$ and

$$\begin{cases} \text{if } k \neq j_i \text{ and } x_k^- = 0, & y_i(k) = \varepsilon, z_i(k) = 2\varepsilon, \\ \text{if } k \neq j_i \text{ and } x_k^- > 0, & y_i(k) = x_k^-, z_i(k) = \varepsilon, \\ \text{finally } y_i(j_i) = x_{j_i}^- + \varepsilon \text{ and } z_i(j_i) \text{ is such that} & x_i(0) = \varepsilon. \end{cases}$$

By definition of the dual function, we then have

$$\forall \lambda, \quad \theta_i(\lambda) \leq \frac{1}{2} x_i^\top Q_i x_i + \lambda(\ell (\mathbf{e}^\top x^- + x_0^-) - \rho_i^\top x_i - \rho_0 x_i(0)). \qquad (31)$$

We have $\rho_i^\top x_i + \rho_0 x_i(0) = \frac{\rho_i(j_i)}{1+\nu_{j_i}}(x_0^- + (\mathbf{e} - \mu)^\top x^-) + a_i' \varepsilon$, for some $a_i' \in \mathbb{R}$. As was done in the proof of Lemma 2.2, since H3 holds, we can then choose ε sufficiently small to have

$$\rho_i^\top x_i + \rho_0 x_i(0) > (\ell + \kappa)(\mathbf{e}^\top x^- + x_0^-). \qquad (32)$$

Using (31) with $\lambda = \lambda_i^*$, (32), and since $\theta_i(\lambda_i^*) \geq 0$ we then get:

$$\lambda_i^* \kappa (\mathbf{e}^\top x^- + x_0^-) \leq \frac{1}{2} \|Q_i\|_\infty \|x_i\|_1^2 \leq \frac{1}{2} \|Q_i\|_\infty (\mathbf{e}^\top x^- + x_0^-)^2.$$

This gives for λ_i^* the upper bound $\lambda_i^* \leq \frac{\|Q_i\|_\infty}{2\kappa}(\mathbf{e}^\top x^- + x_0^-)$. Plugging this bound in (30), we see that $P(\|\frac{x_2^*-x_1^*}{\mathbf{e}^\top x^- + x_0^-}\|_1) \leq 0$, where

$$P(x) = \beta(Q_1)x^2 - \|Q_2 - Q_1\|_\infty x - \frac{(\|Q_1\|_\infty + \|Q_2\|_\infty)}{2\kappa}\|\rho_2 - \rho_1\|_\infty.$$

Consequently, $\|\frac{x_2^*-x_1^*}{\mathbf{e}^\top x^- + x_0^-}\|_1$ is lower than or equal to the largest root of P which is given by (7).

We finally show that for problem P'', $\|\frac{x_2^*-x_1^*}{\mathbf{e}^\top x^- + x_0^-}\|_2$ is bounded from above by (8). We first have

$$\begin{aligned}(x_2^* - x_1^*)^\top (Q_1 - Q_2)\, x_2^* &\leq (\mathbf{e}^\top x^- + x_0^-)\|x_2^* - x_1^*\|_2 \max_{x \in \Delta_n} \|(Q_2 - Q_1)x\|_2, \\ &\leq (\mathbf{e}^\top x^- + x_0^-)\|x_2^* - x_1^*\|_2 \max_i \|C_i(Q_2 - Q_1)\|_2, \qquad (33)\end{aligned}$$

using Lemma 3.1. Using (29) and (33) we then obtain $P(\|\frac{x_2^*-x_1^*}{\mathbf{e}^\top x^- + x_0^-}\|_2) \leq 0$, now with

$$P(x) = \lambda_{\min}(Q_1)x^2 - \max_i \|C_i(Q_2 - Q_1)\|_\infty x - \frac{(\|Q_1\|_\infty + \|Q_2\|_\infty)}{2\kappa}\|\rho_2 - \rho_1\|_\infty$$

and we can conclude as before. □

Acknowledgements The author is grateful to Anatoli Juditski of the "Laboratoire de Modélisation et Calcul" of University Joseph Fourier for helpful discussions and to François Oustry who suggested calibration (10).

References

1. Best, M.J., Grauer, R.R.: On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. Rev. Financ. Stud. **4**(2), 315–342 (1991)
2. Best, M.J., Grauer, R.R.: Sensitivity analysis for mean-variance portfolio problems. Manag. Sci. **37**(8), 980–989 (1991)
3. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer Series in Operations Research and Financial Engineering. Springer, New York (2000)
4. Bouchaud, J.-P., Cizeau, P., Laloux, L., Potters, M.: Noise dressing of financial correlation matrices. Phys. Rev. Lett. **83**(7), 1467–1470 (1999)
5. Daniel, J.W.: Stability of the solution of definite quadratic programs. Math. Program. **5**(1), 41–53 (1973)
6. Dantzig, G.B., Infanger, G.: Multi-stage stochastic linear programs for portfolio optimization. Ann. Oper. Res. **45**(1), 59–76 (1993)
7. Guigues, V.: Mean and covariance matrix adaptive estimation for a weakly stationary process. Application in stochastic optimization. Stat. Decis. **26**, 109–143 (2008)
8. Higham, N.: Computing the nearest symmetric correlation matrix-a problem from finance. IMA J. Numer. Anal. **22**(3), 329–343 (2002)
9. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I and II. Springer, Berlin (1993)
10. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. J. Res. Natl. Bureau Stand. **49**(4), 263–265 (1952)

« COAP 10589 layout: Small Extended v.1.2 reference style: mathphys file: coap9260.tex (Loreta) aid: 9260 doctopic: OriginalPaper class: spr-small-v1.1 v.2009/05/07 Prn:2009/06/05; 13:57 p. 27/27»

11. Ledoit, O.: Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. J. Empir. Finance **10**, 603–621 (2003)
12. Markowitz, H.: Portfolio selection. J. Finance **7**(1), 77–91 (1952)
13. NDiaye, P., Oustry, F.: Kato covariance matrix. Raise partner internal report (2002)
14. Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11–12**, 625–653 (1999)