Chaos, Solitons and Fractals: the interdisciplinary journal of Nonlinear Science, and Nonequilibrium and Complex Phenomena

A Singular Value Decomposition Based Approach to Handle Ill-Conditioning in Optimization Problems with Applications to Portfolio Theory

--Manuscript Draft--

Manuscript Number:	
Article Type:	Full length article
Section/Category:	Interdisciplinary applications of Mathematics
Keywords:	Numerical Stability; Quadratic Programming; Portfolio Optimization
Corresponding Author:	Maria Laura Torrente ITALY
First Author:	Claudia Fassino
Order of Authors:	Claudia Fassino
	Maria Laura Torrente
	Pierpaolo Uberti
Abstract:	We identify a source of numerical instability of quadratic programming problems that is hidden in its linear equality constraints. We propose a new theoretical approach to rewrite the original optimization problem in an equivalent reformulation using the singular value decomposition and substituting the ill-conditioned original matrix of the restrictions with a suitable optimal conditioned one. The proposed novel approach is showed, both empirically and theoretically, to solve ill-conditioning related numerical issues, not only when they depend on bad scaling and are relative easy to handle, but also when they result from almost collinearity and numerically rank-deficient matrices are involved. Furthermore, our strategy looks very promising even when additional inequality constraints are considered in the optimization problem, as it occurs in several practical applications. In this framework, even if no closed form solution is available, we show, through empirical evidence, how the equivalent reformulation of the original problem greatly improves the performances of MatLab's quadratic programming solver and Gurobi. The experimental validation is provided through numerical examples performed on real financial data in the portfolio optimization context.
Suggested Reviewers:	Francesco Cesarone francesco.cesarone@uniroma3.it
	Roy Cerqueti roy.cerqueti@uniroma1.it
Opposed Reviewers:	

Dear Editors,

please find enclosed our manuscript:

Title:
A Singular Value Decomposition Based Approach to Handle Ill-Conditioning in Optimization Problems with Applications to Portfolio Theory

Authors:
Claudia Fassino, Maria-Laura Torrente, Pierpaolo Uberti

which we submit for publication to Chaos, Solitons and Fractals.

Please address all correspondence to:
Maria-Laura Torrente
Dipartimento di Economia, Università di Genova
16126 Genova, Italy
email: marialaura.torrente@economia.unige.it

Yours sincerely,

Maria-Laura Torrente

Title:
A Singular Value Decomposition Based Approach to Handle Ill-Conditioning in Optimization Problems with Applications to Portfolio Theory

Authors:
Claudia Fassino, Maria-Laura Torrente, Pierpaolo Uberti

Highlights:

- Numerical instability of quadratic programming problems
- Ill-conditioned matrix of linear constraints
- Problem's equivalent reformulations using singular value decomposition
- Application to Markowitz portfolio optimization problem

A Singular Value Decomposition Based Approach to Handle Ill-Conditioning in Optimization Problems with Applications to Portfolio Theory

Claudia Fassino[1], Maria-Laura Torrente[2,3], Pierpaolo Uberti[2]

Abstract

We identify a source of numerical instability of quadratic programming problems that is hidden in its linear equality constraints. We propose a new theoretical approach to rewrite the original optimization problem in an equivalent reformulation using the singular value decomposition and substituting the ill-conditioned original matrix of the restrictions with a suitable optimal conditioned one. The proposed novel approach is showed, both empirically and theoretically, to solve ill-conditioning related numerical issues, not only when they depend on bad scaling and are relative easy to handle, but also when they result from almost collinearity and numerically rank-deficient matrices are involved. Furthermore, our strategy looks very promising even when additional inequality constraints are considered in the optimization problem, as it occurs in several practical applications. In this framework, even if no closed form solution is available, we show, through empirical evidence, how the equivalent reformulation of the original problem greatly improves the performances of MatLab®'s quadratic programming solver and Gurobi®. The experimental validation is provided through numerical examples performed on real financial data in the portfolio

Email addresses: `fassino@dima.unige.it` (Claudia Fassino), `marialaura.torrente@economia.unige.it` (Maria-Laura Torrente), `pierpaolo.uberti@unige.it` (Pierpaolo Uberti)

[1]Dipartimento di Matematica, Università degli Studi di Genova, Via Dodecaneso 35, 16146 Genova, Italy

[2]Dipartimento di Economia, Università degli Studi di Genova, Via Vivaldi 5, 16126 Genova, Italy

[3]Corresponding author

Preprint submitted to Chaos, Solitons & Fractals *April 14, 2022*

optimization context.

Keywords: Numerical Stability; Quadratic Programming; Portfolio Optimization

2008 MSC: 15A12, 65F35, 90-08, 91G10

1. Introduction

Quadratic programming problem (QP) minimizes a quadratic objective function under linear equality and/or inequality constraints. These problems are popular in many branches of applied mathematics because of their simple formalization and their great capacity to fit in a stylized way complex real-life problems. Despite the apparent simplicity and tractability, solving a QP becomes hard when the matrix describing the quadratic objective function and/or the matrix describing the linear constraints are ill-conditioned.

In the literature, depending on the various fields of application, a QP is generally analyzed from different points of view; while in the context of the decision sciences a QP is addressed focusing on its objective function, the algebraic approach first puts the emphasis on the role of the constraints. In particular, for a given QP with linear equality constraints, the first approach stresses the role of the objective function, the goal of the decision process, focusing on the conditioning of the associated matrix and giving secondary importance to the restrictions; on the opposite, the second approach pays a special attention to the equality constraints, as the QP is solved by finding the minimum norm solution of the undetermined linear system of the restrictions, see among the others [9, 12, 24]

In this paper, taking advantage of the above mentioned different points of view, we address the problem of the potential numerical instability of a QP elicited by the ill-conditioning of its matrix of restrictions, and propose a method to reduce such phenomenon, by also handling the usually difficult case of almost collinearity in the restrictions. Our choice to concentrate on one of the two possible sources of instability is due to the fact that, in relatively recent years, the conditioning of the quadratic function has been deeply investigated in the literature, especially in the context of the decision sciences, whereas, to the best of our knowledge, no contributions appear in connection to the bad conditioning of the equality constraints.

Our interest for this kind of numerical issues came from the study of the mean-variance portfolio model introduced by [25] and formulated as a QP. Since the publication of such landmark paper, the literature registered an intensive activity on optimal portfolios connected with QP problems (we refer to [13, 20, 27, 31] as non exhaustive list of relevant references). Though Markowitz model is universally known as the starting point of modern portfolio theory, it is also famous for the difficulty of its practical implementation, see [2, 15], and its poor out-of-sample performances, see [8]. The literature identifies in the numerical instability of the model the principal reason of the above mentioned drawbacks. The model's parameters are unknown and, consequently, need to be estimated, leading to computational inaccuracy and uncertainty. In other words, the in-sample mean-variance frontier is a biased estimator of the real efficient frontier, see [17]. Many authors, see for instance [1] and [18], consider the estimation uncertainty the cause of instability in the model. This issue gets worse as the size of the portfolio increases, considering that the number of the parameters to be estimated grows quadratically in the number portfolio's assets, see [28]. Moreover, the solution of the QP depends on the computation of the inverse of the covariance matrix, that is potentially ill-conditioned and close to a singular matrix when the returns of the assets are almost collinear. To improve the numerical stability, a bunch of alternative proposals have been discussed in the literature: among the others we recall the Bayesian approach, see [11], the shrinkage approach, see [22], robust optimization techniques, see [19] and [30], and Lasso techniques, see [4]. As it is clear from the previous enumeration, the literature focuses on the role of the covariance matrix and, consequently, on the objective function of the optimization problem. Unfortunately, even when the approaches mentioned above are able to control the instability depending on the objective function, a potential source of instability remains hidden in the restrictions of the problem, as we highlight in this paper.

For this reason, starting from a QP defined as the minimization of a quadratic objective function under linear equality constraints, we focus on the potential numerical instability elicited by the ill-conditioning of the restrictions matrix, exploiting the closed form solution of the problem, generalizing and improving the approach proposed in [10]. We note that the numerical issues are related to the presence of badly scaled and/or almost collinear restrictions, but while bad scaling is relatively simple to manage, almost collinearity usually constitutes a fatal issue. The approach we propose is showed to be effective also when ill-conditioning depends on almost collinearity, solving apparent unfeasible numerical problems. We propose to substitute the original problem with an equivalent optimal conditioned reformulation, obtained through a singular value decomposition (SVD). Moreover, we show, through a new theoretical result (Theorem 3.1) and two toy examples (Examples 3.1 and 3.2), that the proposed reformulation of the original problem is able to improve the numerical stability. Then we apply our approach to the special case of Markowitz portfolio optimization model, a QP with two linear constraints. It is relevant to notice that the two linear constraints in Markowitz model, the budget constraint and the restriction on portfolio expected return, structurally suffer from almost collinearity and bad scaling. This makes the portfolio optimization problem arising in the financial context so interesting for our analysis that we decided to entirely devote the empirical examples to this case.

Very often in empirical applications additional restrictions to the original problem are needed; for example, it is common in practice to require portfolios to be long only, restricting the solution to a non-negative vector. Also in this case, the proposed equivalent optimal conditioned reformulation of the problem looks very promising. We note that the solution of such QP needs to be computed through a numerical procedure, as no closed form formula is available for it. We compare the performances of MatLab®'s built-in function `quadprog` using Gurobi® as a benchmark representing one of the state-of-the-art solvers. The most relevant evidence of our empirical applications on real financial data is that the proposed equivalent reformulation of the original problem significantly impacts both softwares' performances with respect to the original formulation: the algorithms improve in terms of convergence rate to the optimal solution, number of iterations and computational time.

The paper is organized as follows. In Section 2 the paper's notation and some useful technical results are presented; Section 3 contains the main theoretical results of the paper on the reformulation of the proposed QP and its conditioning; in Section 4 the addressed problem is restricted to the case of two linear equality constraints directly referring to the Markowitz model; Section 5, entirely devoted to the portfolio optimization framework, adapts our proposed reformulation to the QP with additional inequality constraints and offers several examples on real financial data to empirically support its efficiency; finally, Section 6 concludes the paper.

2. Notation and Technical Results

In this section, we first recall some basic notions of linear algebra and the notation useful throughout the paper, see textbooks [3] and [14]. Then, we prove technical results that will be used in Section 3 to reformulate the QP and evaluate the improvements in terms of conditioning.

Let m, n be positive integers and $\mathrm{Mat}_{m\times n}(\mathbb{R})$ be the set of $m \times n$ real matrices. Let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. We denote by A^t the *transpose* of A and by $A^\dagger$ the *Moore-Penrose pseudoinverse* of A. We recall that an *orthogonal* matrix is any square matrix whose columns and rows are orthogonal unit vectors, that is orthonormal vectors. A rectangular matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, with $m > n$, is called *orthonormal* if and only if $A^tA = I_n$. We refer to the *Singular Value Decomposition* (SVD) of A as the unique factorization of the form $A = U\Sigma V^t$, where $U \in \mathrm{Mat}_{m\times m}(R)$, $V \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ are orthogonal matrices and $\Sigma \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ is a rectangular diagonal matrix with diagonal elements, listed in non-increasing order, $\Sigma_{ii} = \sigma_i(A) > 0$, $i = 1, \ldots, r$, and $\Sigma_{ii} = 0$, $i = r+1, \ldots, m$, where $r = \mathrm{rank}(A)$ is the rank of A. The elements

$\sigma_1(A) \geq \ldots \geq \sigma_r(A) > 0$ are the *singular values* of A; in the special case in which A is a $n \times n$ *symmetric* and *positive definite* matrix, the singular values $\sigma_i(A)$, $i = 1, \ldots, n$, are the eigenvalues of A.

Throughout the paper we use the sign function defined over the reals by $\text{sign}(x) = \frac{|x|}{x}$ for each $x \neq 0$ and $\text{sign}(0) = 0$. Further, we use the standard scalar product $\langle \cdot, \cdot \rangle$ and the Euclidean norm (2-norm) $\| \cdot \|$ of $\mathbb{R}^n$ while $1_n = (1, \ldots, 1) \in \text{Mat}_{1\times n}(\mathbb{R})$ denotes the vector of ones.

We recall the definition of condition number of a matrix, see [3, Section 2.1]. Let $A \in \text{Mat}_{m\times n}(\mathbb{R})$ be a full-rank matrix; the (2-norm) *condition number* of A, denoted by $K_2(A)$, is given by:

$$K_2(A) := \|A\| \, \|A^\dagger\| \quad \text{or equivalently} \quad K_2(A) := \frac{\sigma_1(A)}{\sigma_r(A)},$$

with $r = \min\{m, n\} = \text{rank}(A)$. Note that, in the special case $m = n$, the previous relations simply become

$$K_2(A) := \|A\| \, \|A^{-1}\| \quad \text{or equivalently} \quad K_2(A) := \frac{\sigma_1(A)}{\sigma_n(A)}.$$

The condition number is classically used to provide a measure of linear system sensitivity, see [14]. A matrix with a large condition number[4] is said to be *ill-conditioned* and characterized by a warning about the potential numerical instability of the associated linear systems' solution. In this paper the condition number is used as the principal instrument to detect and quantify numerical issues.

In the following we prove some basic results of numerical linear algebra useful for a complete understanding of the paper.

Proposition 2.1. *Let $A = (a_{ij}) \in Mat_{n\times n}(\mathbb{R})$ be a symmetric and positive definite matrix.*

(i) *Let $B \in Mat_{m\times m}(\mathbb{R})$, with $m < n$, be a principal submatrix of A; then*

$$K_2(B) \leq K_2(A).$$

[4] Note that this property depends on the chosen norm and on the definition of "large", for a further study we refer to [14].

(ii) *If* $n = 2$ *then*

$$K_2(A) \geq \frac{\max\{a_{11}, a_{22}\}}{\min\{a_{11}, a_{22}\}}.$$

Proof.

(i) The principal submatrix B is symmetric and positive definite since A is symmetric and positive definite. Let $\lambda_1(A) \geq \ldots \geq \lambda_n(A) > 0$ and $\lambda_1(B) \geq \cdots \geq \lambda_m(B) > 0$ be the eigenvalues of A and B respectively, which coincide with their singular values. Due to a consequence of the *Eigenvalue Interlacing Theorem* [14, Theorem 8.1.7], the following inequalities hold:

$$\lambda_n(A) \leq \lambda_m(B) \qquad \text{and} \qquad \lambda_1(B) \leq \lambda_1(A).$$

Therefore, recalling the definition of condition number, it follows that

$$K_2(B) = \frac{\lambda_1(B)}{\lambda_m(B)} \leq \frac{\lambda_1(A)}{\lambda_n(A)} = K_2(A).$$

(ii) Since A is symmetric and positive definite, then $a_{12} = a_{21}$, $a_{ii} \geq 0$, $i = 1, 2$. Let $\lambda_1(A) \geq \lambda_2(A) > 0$ be the eigenvalues of A which, by definition, satisfy the equation $\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}^2) = 0$. Therefore

$$\begin{aligned} K_2(A) &= \frac{\left(a_{11} + a_{22} + \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}\right)^2}{4(a_{11}a_{22} - a_{12}^2)} \\ &\geq \frac{(a_{11} + a_{22} + \mid a_{11} - a_{22} \mid)^2}{4a_{11}a_{22}}. \end{aligned}$$

Considering the two cases $a_{11} \geq a_{22}$ or $a_{11} < a_{22}$, the right-hand-side of the above inequality becomes $\frac{a_{11}}{a_{22}}$ or $\frac{a_{22}}{a_{11}}$, which proves the proposition.

□

Theorem 2.2. *Let* $A = (a_{ij}) \in Mat_{2\times n}(\mathbb{R})$ *be a full-rank matrix,* $A_i = (a_{ij})_{j=1}^n \in Mat_{1\times n}(\mathbb{R})$ *be the ith row of* A*,* $i = 1, 2$*, and* $s = sign(\langle A_1, A_2\rangle)$*. Assume that* $\|A_1\|^2 = \|A_2\|^2 = \alpha$*. Then the singular values of* A *are* $\sigma_1 = (\alpha + \mid \langle A_1, A_2\rangle \mid)^{\frac{1}{2}}$ *and* $\sigma_2 = (\alpha - \mid \langle A_1, A_2\rangle \mid)^{\frac{1}{2}}$*. Further,* $A = U\Sigma_2 W_2^t$*, where*

$$\Sigma_2 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix},$$

$U \in Mat_{2\times 2}(\mathbb{R})$ is an orthogonal matrix and $W_2 \in Mat_{n\times 2}(\mathbb{R})$ is an orthonormal matrix, defined according to the following cases: if $\langle A_1, A_2\rangle \neq 0$ then

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} s & -s \\ 1 & 1 \end{pmatrix} \quad \textit{and} \quad W_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} \sigma_1^{-1}(sA_1^t + A_2^t) & \sigma_2^{-1}(-sA_1^t + A_2^t) \end{pmatrix}; \quad (1)$$

otherwise, if $\langle A_1, A_2\rangle = 0$ then

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad W_2 = \sigma_1^{-1}\begin{pmatrix} A_1^t & A_2^t \end{pmatrix}. \quad (2)$$

Proof. The singular values of A can be computed as the square roots of the eigenvalues of AA^t; therefore an easy computation yields that σ_1 and σ_2 are the singular values of A. To conclude the proof, it is enough to observe that the matrices U and W_2 as defined in (1) and in (2) are orthogonal and orthonormal matrices respectively, and that, in both cases, $A = U\Sigma_2 W_2^t$. □

3. Optimization problem

We consider the special case of QP: the problem of minimizing a quadratic objective function of several variables subject to linear equality constraints, see [26]. As discussed in the introduction of the paper, this class of optimization problems has been widely studied and proposed in different branches of applied mathematics. In the empirical section we will focus on an application in the financial context for portfolio allocation problems.

Let $Q \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be a symmetric and positive definite matrix, let $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, with $m < n$, be a full-rank matrix and $b \in \mathrm{Mat}_{m\times 1}(\mathbb{R})$ be a column vector. We consider the following problem.

Problem 3.1:

$$\begin{aligned} &\text{Minimize} && x^t Q x \\ &\text{subject to} && Ax = b. \end{aligned}$$

It is immediate to verify that the unique solution of Problem 3.1 is:

$$x = Q^{-1}A^t B^{-1} b \quad (3)$$

with

$$B = AQ^{-1}A^t. \tag{4}$$

For a general approach to the problem's solution see [24]; for the solution in the special case of the financial context see [7]. Though solution (3) is explicitly given, it is straightforward to notice that, in practice, computing it can lead to a high rate of numerical instability, since it depends on the inversion of matrices Q and B. If Q is ill-conditioned and/or almost singular, the definition of matrix B in equation (4) could be meaningless. We enumerated in the Introduction many papers, in the financial context, proposing alternative estimations of the matrix Q to mitigate potential instability and sensitivity of the solution arising from the misspecification of the objective function. For this reasons, in the rest of the paper, we assume that Q is well-conditioned not causing any numerical issue to the solution in equation (3). Such assumption permits to focus on the independent potential source of instability that may naturally arise from the linear constraints $Ax = b$. In order to mitigate this phenomenon, we introduce an alternative equivalent formulation of Problem 3.1.

We consider the singular value decomposition (SVD) of the matrix A:

$$A = U\Sigma W^t, \tag{5}$$

where $U \in \text{Mat}_{m\times m}(R)$, $W \in \text{Mat}_{n\times n}(\mathbb{R})$ are orthogonal matrices and $\Sigma \in \text{Mat}_{m\times n}(\mathbb{R})$ is a rectangular diagonal matrix with $\Sigma_{ii} = \sigma_i(A) > 0$, $i = 1, \dots, m$. Let $\Sigma_m \in \text{Mat}_{m\times m}(\mathbb{R})$ and $W_m \in \text{Mat}_{n\times m}(\mathbb{R})$ be the matrices made up of the first m columns of Σ and W respectively; in particular, note that Σ_m is diagonal and W_m is orthogonal. Then $A = U\Sigma_m W_m^t$ and the equality constraints of Problem 3.1 may also be equivalently written as $W_m^t x = \Sigma_m^{-1}U^t b$. This approach allows us to formulate Problem 3.2 as an alternative equivalent version of Problem 3.1.

Problem 3.2:

$$\begin{array}{ll} \text{Minimize} & x^tQx \\ \text{subject to} & W_m^t x = \Sigma_m^{-1}U^t b. \end{array}$$

Note that Problems 3.1 and 3.2 share the same quadratic objective function x^tQx, whereas the equations of their equality constraints, which define the same geometrical object, are provided through the two different linear systems $Ax = b$ and $W_m^t x = \Sigma_m^{-1}U^t b$. It is immediate to verify that the unique solution of Problem 3.2 is:

$$x = Q^{-1}W_m C^{-1} v$$

with $v = \Sigma_m^{-1}U^t b$ and

$$C = W_m^t Q^{-1} W_m. \tag{6}$$

It is worth noting that the advantage of dealing with the alternative formulation Problem 3.2 stands in the substitution of the linear equality constraints $Ax = b$, which could happen to be ill-conditioned, with $W_m^t x = v$, where W_m is orthonormal. The proposed reformulation of the restrictions is computationally efficient and, in general, not worse in terms of conditioning, substituting to the original matrix A the orthonormal matrix W_m that has, by construction, the minimum condition number. In the following result, in order to support the intuitive idea that such reformulation could increase the numerical stability of the problem solution, we compare the equivalent Problems 3.1 and 3.2. Our analysis is based on the notion of condition number, whose definition is recalled in Section 2.

Theorem 3.1. *In the setting of Problems 3.1 and 3.2 the following inequalities hold:*

$$K_2(C) \leq K_2(Q) \qquad \textit{and} \qquad K_2(B) \leq K_2^2(A)K_2(Q).$$

Further, denoting $C = (c_{ij})_{i,j=1}^m$, it holds:

$$\begin{cases} K_2(B) \geq K_2^2(A) & \textit{if } \frac{c_{11}}{c_{mm}} \geq 1 \\ \\ K_2(B) > \frac{K_2^2(A)}{K_2(Q)} & \textit{if } \frac{1}{K_2^2(A)} \leq \frac{c_{11}}{c_{mm}} < 1. \end{cases}$$

Proof. By formula (6) the matrix C is the principal submatrix of order m of $W^tQ^{-1}W$, where the orthogonal matrix W is defined in (5). Since the matrix Q

is symmetric and positive definite, so it is $W^tQ^{-1}W$ and its condition number is

$$K_2(W^tQ^{-1}W) = K_2(Q^{-1}) = K_2(Q).$$

Using the last equality and Proposition 2.1-(i) applied to $W^tQ^{-1}W$ and C, it follows that $K_2(C) \leq K_2(Q)$.

From (4), using (6) and the equality $A = U\Sigma_m W_m^t$, we get

$$B = U\Sigma_m W_m^t Q^{-1} W_m \Sigma_m U^t = U\Sigma_m C \Sigma_m U^t.$$

Let $G = \Sigma_m C \Sigma_m$; since U is orthogonal, then

$$K_2(B) = K_2(G). \tag{7}$$

Further, since $K_2(\Sigma_m) = K_2(A)$ and $K_2(C) \leq K_2(Q)$, the following inequalities hold:

$$K_2(G) \leq K_2(\Sigma_m)^2 K_2(C) = K_2(A)^2 K_2(C) \leq K_2(A)^2 K_2(Q). \tag{8}$$

From (7) and (8), we get $K_2(B) \leq K_2(A)^2 K_2(Q)$.

For the last part of the proof, we recall that the diagonal elements $\sigma_1 \geq \ldots \geq \sigma_m > 0$ of Σ_m are the singular values of A and that $K_2(A) = \frac{\sigma_1}{\sigma_m}$. Further, we denote the elements of G by g_{ij}, $i, j = 1, \ldots, m$, so that, by definition of G, it holds $g_{ij} = \sigma_i \sigma_j c_{ij}$, for each $i, j = 1, \ldots, m$. We let $I = \{1, m\}$ and consider the two principal submatrices $C_I = (c_{ij})_{i,j \in I}$ and $G_I = (g_{ij})_{i,j \in I}$ of C and G:

$$C_I = \begin{pmatrix} c_{11} & c_{1m} \\ c_{1m} & c_{mm} \end{pmatrix} \qquad G_I = \begin{pmatrix} \sigma_1^2 c_{11} & \sigma_1 \sigma_m c_{1m} \\ \sigma_1 \sigma_m c_{1m} & \sigma_m^2 c_{mm} \end{pmatrix}.$$

From (7) and using Proposition 2.1-(i) applied to G and G_I we get:

$$K_2(B) = K_2(G) \geq K_2(G_I). \tag{9}$$

Further, the hypothesis $\frac{c_{11}}{c_{mm}} \geq \frac{1}{K_2^2(A)} = \frac{\sigma_m^2}{\sigma_1^2}$ implies $\sigma_1^2 c_{11} - \sigma_m^2 c_{mm} \geq 0$, thus applying Proposition 2.1-(ii) to G_I we get

$$K_2(G_I) \geq \frac{\sigma_1^2 c_{11}}{\sigma_m^2 c_{mm}} = \frac{c_{11}}{c_{mm}} K_2^2(A). \tag{10}$$

Relations (9) and (10) yield:

$$K_2(B) \geq \frac{c_{11}}{c_{mm}} K_2^2(A). \tag{11}$$

If $\frac{c_{11}}{c_{mm}} \geq 1$, relation (11) becomes $K_2(B) \geq K_2^2(A)$. Otherwise, if $\frac{1}{K_2^2(A)} \leq \frac{c_{11}}{c_{mm}} < 1$, from Proposition 2.1-(i) applied to C and C_I, and Proposition 2.1-(ii) applied to C_I we get $K_2(C) \geq K_2(C_I) > \frac{c_{mm}}{c_{11}}$, so that

$$\frac{c_{11}}{c_{mm}} > \frac{1}{K_2(C)}. \tag{12}$$

Therefore, using relation (11), (12) and the inequality $K_2(C) \leq K_2(Q)$ it follows that $K_2(B) > \frac{K_2^2(A)}{K_2(Q)}$, which concludes the proof. □

We list in the following remarks some important observations.

Remark 3.1: Under the standing assumption that Q is well-conditioned, we observe that the alternative equivalent formulation 3.2 of Problem 3.1 is always preferable. In fact, from the inequality $K_2(C) \leq K_2(Q)$ of Theorem 3.1, it follows that the matrix C is well-conditioned, and this is completely independent from the conditioning of matrix A. On the other hand, if A is ill-conditioned, so it is the matrix B, as its condition number approximately behaves as the quantity $K_2^2(A)$, a factor which is present in both the provided lower and upper bounds of $K_2(B)$.

Remark 3.2: The ratio $\frac{c_{11}}{c_{mm}} \in (0, +\infty)$ while the union of the alternative conditions given in Theorem 3.1 restricts $\frac{c_{11}}{c_{mm}} \in \left[\frac{1}{K_2^2(A)}, +\infty\right)$. Nevertheless, in the case of interest for our approach, the matrix A of the restrictions is ill-conditioned, its condition number is high and the two intervals become very similar.

Remark 3.3: We note that the proposed equivalent formulation 3.2 of Problem 3.1 could show poor improvements if the matrix A is nearly rank-deficient. In this case, exploiting the notion of *numerical rank* (see [14, 16]), we slightly modify our approach as follows. If $r < m$ is the numerical rank of A, then the equation of the equality constraints in Problem 3.2 becomes $W_r^t x = \Sigma_r^{-1} U_r^t b$, where $U_r \in \mathrm{Mat}_{m\times r}(\mathbb{R})$ and $W_r \in \mathrm{Mat}_{n\times r}(\mathbb{R})$ are the matrices containing the first r columns of U and W respectively, and $\Sigma_r \in \mathrm{Mat}_{r\times r}(\mathbb{R})$ is the leading principal submatrix of order r of Σ. An illustrative instance is provided in Example 3.2.

We end the section with two illustrative toy examples to show, in some simple cases, the effect of the proposed reformulation of the original problem to the numerical stability of the closed form solution.

Example 3.1 is designed on purpose to illustrate the effectiveness of our proposal. Despite its artificial aspect, it results both qualitatively and quantitatively very similar to real financial data scenarios occurring in portfolio optimization applications, as it will be more clear from the examples in Section 5. In Example 3.1 we note that the norms of the two rows of A are almost equal (they differ by less than δ). As a consequence, the problem's numerical issues exclusively depend on the matrix's near collinearity.

Example 3.1: We consider the setting of Problem 3.1 with $n = 3$, $m = 2$ and Q equal to the identity matrix I_3. In this case $K_2(I_3) = 1$, so that the potential numerical instability only depends on the problem's restrictions $Ax = b$. Let $\delta = 10^{-8}$ and

$$A = \begin{pmatrix} 1 & 1 & 1+\delta \\ 1 & 1 & 1 \end{pmatrix} \qquad \text{and} \qquad b = \begin{pmatrix} 1.1 \\ 1 \end{pmatrix}.$$

We note that the singular values of A are 2.4495 and $5.7735 \cdot 10^{-9}$, so that A has full numerical rank; its condition number is $K_2(A) = 4.2426 \cdot 10^{8}$.

Using the provided explicit formula (3) and through symbolic computations, the solution of Problem 3.1 is

$$x^* = \begin{pmatrix} -0.05\delta^{-1} + 0.5 \\ -0.05\delta^{-1} + 0.5 \\ 0.1\delta^{-1} \end{pmatrix} = 10^7 \cdot \begin{pmatrix} -0.5(1 - 10^{-7}) \\ -0.5(1 - 10^{-7}) \\ 1 \end{pmatrix}.$$

Let x_1 and x_2 be the numerical solutions of Problem 3.1 and its equivalent reformulation, Problem 4.1, respectively, computed by MatLab® using formulas

(3) and (19):

$$x_1 = 10^7 \cdot \begin{pmatrix} 0.4290613400000006 \\ 0.4290613400000006 \\ -0.858122660000001 \end{pmatrix} \qquad x_2 = 10^7 \cdot \begin{pmatrix} -0.499999949077899 \\ -0.499999968776829 \\ 1.000000017854728 \end{pmatrix}$$

We note that the condition numbers of the matrices B and C (see (4) and (6)) involved in the closed form solutions are approximately $K_2(B) = 1.8 \cdot 10^{16}$ and $K_2(C) = 1.0$.

By a comparison with x^*, it is evident that x_2 is close to the solution of the proposed minimization problem. Further, as shown below, x_1 does not satisfy the linear constraints $Ax = b$ at all, whereas x_2 exactly satisfies the second linear equality provided by $Ax = b$ and shows a rounding error of about 10^{-9} in the first linear constraint:

$$\mid Ax_1 - b \mid = \begin{pmatrix} 0.9858 \\ 0.8000 \end{pmatrix} \qquad \mid Ax_2 - b \mid = 10^{-8} \cdot \begin{pmatrix} 0.1490 \\ 0 \end{pmatrix}.$$

Note that this example highlights the effectiveness and the improvements in terms of numerical stability related to our proposed reformulation, even when the solution is computed using numerical solvers instead of the available explicit formula. To this aim we consider the well-known mathematical optimization solver functions included in Matlab® and Gurobi®. Starting from Problem 3.1, because of numerical issues, the MatLab®'s function `quadprog` returns a message asserting that no feasible solution is found, while Gurobi® computes the suboptimal solution $(0.3667, 0.3667, 0.3667)^t$. On the contrary, starting from the equivalent reformulation Problem 4.1, both the solvers converge to a solution (approximately) equal to x_2.

Finally, Example 3.2 considers the case of a QP problem whose coefficient matrix of the equality constraints is numerically rank deficient (the numerical rank is computed with respect to the machine precision 10^{-16}).

Example 3.2: We consider the setting of Problem 3.1 with $n = 8$, $m = 6$ and Q equal to the identity matrix I_8. Let $\delta = 10^{-8}$; let $A = (a_{ij}) \in \mathrm{Mat}_{6\times 8}(\mathbb{R})$ be such that all $a_{ij} = 1$ except for the following values: $a_{18} = 1 + \delta$, $a_{38} = 1 - \delta$, $a_{45} = 1 + 3\delta$, $a_{58} = 1 + 4\delta$, $a_{61} = 1 - \delta$. Let $b \in \mathrm{Mat}_{6\times 1}(\mathbb{R})$ be the vector of the arithmetic means of the rows of A. We list the singular values of A:

$$\begin{array}{lll} \sigma_1 = 6.9282 & \sigma_2 = 3.8518 \cdot 10^{-8} & \sigma_3 = 2.3299 \cdot 10^{-8} \\ \sigma_4 = 8.0737 \cdot 10^{-9} & \sigma_5 = 1.2628 \cdot 10^{-24} & \sigma_6 = 4.8463 \cdot 10^{-25}, \end{array}$$

and note that the numerical rank of A is $r = 4$ (such computation has been performed with respect to the machine precision 10^{-16}, see [16]). Let x_1 and x_2 be the solutions of Problem 3.1 and 4.1 (where Remark 3.3 has been employed) computed by MatLab® using formulas (3) and (19):

$$\begin{aligned} x_1 &= (-0.1648, -0.1648, -0.1648, -0.1648, 0.0905, -0.1648, -0.1648, 0.2563)^t \\ x_2 &= (0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250, 0.1250)^t, \end{aligned}$$

from which it is evident that x_2 is the correct solution of the proposed minimization problem.

4. The special case of two linear constraints

In this section we consider the optimization Problem 3.1 in the special case of $m = 2$ linear equality constraints and formulate an alternative version of it.

In Section 3, in the general case, we introduced a constructive approach, essentially relying upon the SVD computation of the coefficient matrix of the problem's linear constraints, to define the equivalent Problem 3.2. In the following, by exploiting Proposition 2.2, which provides an explicit formula for the SVD of any $2 \times n$ full-rank matrix, we introduce an equivalent version of Problem 3.1, holding in the special case $m = 2$.

We consider Problem 3.1 with $m = 2$ and linear equality constraints $DAx = Db$, where

$$D = \begin{pmatrix} \frac{1}{\|A_1\|} & 0 \\ 0 & \frac{1}{\|A_2\|} \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R}) \ , \tag{13}$$

and A_i denotes the ith row of A, $i = 1, 2$. Applying Theorem 2.2 to DA, we get that σ_1 and σ_2, where

$$\sigma_1 = \left(1 + \frac{|\langle A_1, A_2\rangle|}{\|A_1\|\|A_2\|}\right)^{\frac{1}{2}} \qquad \text{and} \qquad \sigma_2 = \left(1 - \frac{|\langle A_1, A_2\rangle|}{\|A_1\|\|A_2\|}\right)^{\frac{1}{2}}, \tag{14}$$

are the singular values of DA and that $DA = U\Sigma_2 W_2^t$, where

$$\Sigma_2 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R}), \tag{15}$$

$U \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ and $W_2 \in \mathrm{Mat}_{n\times 2}(\mathbb{R})$ are orthogonal and orthonormal matrices defined according to the following cases: if $\langle A_1, A_2\rangle \neq 0$ then

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} s & -s \\ 1 & 1 \end{pmatrix} \tag{16}$$

and

$$W_2 = \frac{1}{\sqrt{2}} \left(\sigma_1^{-1} \left(s\tfrac{A_1^t}{\|A_1\|} + \tfrac{A_2^t}{\|A_2\|} \right) \quad \sigma_2^{-1} \left(-s\tfrac{A_1^t}{\|A_1\|} + \tfrac{A_2^t}{\|A_2\|} \right) \right) \tag{17}$$

where $s = \mathrm{sign}(\langle A_1, A_2\rangle)$; otherwise, if $\langle A_1, A_2\rangle = 0$ then

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad W_2 = \left(\tfrac{A_1^t}{\|A_1\|} \quad \tfrac{A_2^t}{\|A_2\|} \right). \tag{18}$$

According to the previous discussion, we introduce an alternative equivalent version of Problem 3.1.

Problem 4.1:

$$\begin{array}{ll} \text{Minimize} & x^t Q x \\ \text{subject to} & W_2^t x = \Sigma_2^{-1} U^t D b. \end{array}$$

where D, Σ_2, U, W_2 are defined in (13), (15), (16), (17) and (18).

The unique solution of Problem 4.1 is:

$$x = Q^{-1} W_2 C^{-1} v \tag{19}$$

with $v = \Sigma_2^{-1} U^t D b$ and $C = W_2^t Q^{-1} W_2$.

4.1. The portfolio selection problem

In this section we recall the mean-variance portfolio model developed by [25] and, based on the results of Section 4, we formulate an equivalent version of the classical portfolio selection problem.

In the setting of portfolio optimization, each entry x_i of the vector x, with $i = 1 \dots n$, denotes the share of the investor's wealth allocated to the ith asset in the portfolio, so that x is the vector of the weights and n is the number of risky investment opportunities. For each $i = 1 \dots n$, the expected return of the ith asset is denoted by μ_i, so that $\mu = (\mu_1, \dots, \mu_n) \in \mathrm{Mat}_{1\times n}(\mathbb{R})$ contains the expected returns. The scalar products $\langle x, 1_n \rangle$ and $\langle x, \mu \rangle$ are respectively equal to the budget constraint, the restriction assuming that all the available wealth is allocated in the portfolio, and the expected return of the portfolio, μ_p in the following. We assume that not all the elements of μ are equal. The covariance matrix of the returns is denoted by $V \in \mathrm{Mat}_{n\times n}(\mathbb{R})$; by construction, V is symmetric and positive definite. Finally, the variance of portfolio's returns is $\sigma_p^2 = x^t V x$.

In this framework, the portfolio selection problem is stated as Problem 4.2 and, equivalently, as Problem 4.3.

Problem 4.2:

$$\begin{aligned} &\text{Minimize} && \sigma_p^2 = x^t V x \\ &\text{subject to} && \begin{pmatrix} \mu \\ 1_n \end{pmatrix} x = \begin{pmatrix} \mu_p \\ 1 \end{pmatrix}. \end{aligned}$$

Problem 4.3:

$$\begin{aligned} &\text{Minimize} && \sigma_p^2 = x^t V x \\ &\text{subject to} && W_2^t x = \Sigma_2^{-1} U^t D \begin{pmatrix} \mu_p \\ 1 \end{pmatrix}. \end{aligned}$$

where $D = \begin{pmatrix} \frac{1}{\|\mu\|} & 0 \\ 0 & \frac{1}{\sqrt{n}} \end{pmatrix}$, $\Sigma_2 = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$, with

$$\sigma_1 = \left(1 + \frac{1}{\sqrt{n}\|\mu\|} \mid \sum_{i=1}^{n} \mu_i \mid \right)^{\frac{1}{2}} \qquad \text{and} \qquad \sigma_2 = \left(1 - \frac{1}{\sqrt{n}\|\mu\|} \mid \sum_{i=1}^{n} \mu_i \mid \right)^{\frac{1}{2}},$$

and $U \in \text{Mat}_{2\times 2}(\mathbb{R})$, $W_2 \in \text{Mat}_{n\times 2}(\mathbb{R})$ are orthogonal and orthonormal matrices defined according to the following cases: if $\sum_{i=1}^n \mu_i \neq 0$ then

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} s & -s \\ 1 & 1 \end{pmatrix} \qquad \text{and} \qquad W_2 = \frac{1}{\sqrt{2}} \left(\sigma_1^{-1} \left(s \frac{\mu^t}{\|\mu\|} + \frac{1_n^t}{\sqrt{n}} \right) \quad \sigma_2^{-1} \left(-s \frac{\mu^t}{\|\mu\|} + \frac{1_n^t}{\sqrt{n}} \right) \right),$$

where $s = \text{sign}\left(\sum_{i=1}^n \mu_i\right)$; otherwise, if $\sum_{i=1}^n \mu_i = 0$ then

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad W_2 = \left(\frac{\mu^t}{\|\mu\|} \quad \frac{1_n^t}{\sqrt{n}} \right) .$$

5. Long-only portfolio optimization and computational results

In this section we focus on Problems 4.2 and 4.3, restricting the application to the portfolio optimization framework, both for its relevance in financial literature and its suitability to highlight the effectiveness of our proposal. To this aim we introduce Problem 5.1, a variation of Problem 4.2 obtained adding the further restriction that all the components x_i are non-negative, i.e. no short positions are allowed in the optimal portfolio. The long-only constraint is so common in practice that it is often not perceived as a constraint and its impact on the performance with respect to long-short strategies has been widely studied, see [6]. Exploiting the results of Section 4.1, the equivalence of Problems 5.1 and 5.2 directly follows from the equivalence of Problems 4.2 and 4.3.

Problem 5.1:

$$\begin{array}{lccll} \text{Minimize} & x^t V x & & & \\ \text{subject to} & \begin{pmatrix} \mu \\ 1_n \end{pmatrix} x & = & \begin{pmatrix} \mu_p \\ 1 \end{pmatrix} & \\ & x_i & \geq & 0 & i = 1, \ldots, n. \end{array}$$

Problem 5.2:

$$\begin{array}{lccll} \text{Minimize} & x^t V x & & & \\ \text{subject to} & W_2^t x & = & \Sigma_2^{-1} U^t D \begin{pmatrix} \mu_p \\ 1 \end{pmatrix} & \\ & x_i & \geq & 0 & i = 1, \ldots, n. \end{array}$$

where the matrices D, Σ_2, U, W_2 are defined in Problem 4.3.

Though Problem 5.1, including also inequality constraints among its restrictions, differs from the optimization problems treated so far, the equivalent reformulation provided by Problem 5.2 could show strong improvements in the solution's numerical behaviour. This is a consequence of the nature of the involved restrictions: if the equality constraints are described by an ill-conditioned system, then small perturbations may cause large variations in the problem's feasible set, with a subsequent significant change in the solution of the optimization problem. This idea is intuitively described by means of the following illustrative example, in which the worst case of almost collinearity between the vectors μ and 1_n is considered.

Example 5.1: We consider the set of constraints of Problem 5.1 in which $\mu = (\mu_i) \in \mathrm{Mat}_{1\times n}(\mathbb{R})$, with $\mu_i = 1$, for each $i = 1, \dots, n-1$, $\mu_n = 1 + \delta$, $\delta \in \mathbb{R}$, and $\mu_p = 1 + \alpha$, $\alpha \in \mathbb{R}$. The set of points verifying the constraints is

$$\mathcal{D} = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \;\mid\; \sum_{i=1}^{n-1} x_i = 1 - \frac{\alpha}{\delta}, x_n = \frac{\alpha}{\delta}, x_i > 0, i = 1, \dots, n \right\}.$$

If $0 < \frac{\alpha}{\delta} < 1$ then $\mathcal{D}$ is not empty and geometrically corresponds to the contraction, with a factor of $1 - \frac{\alpha}{\delta}$, of the $(n-2)$-dimensional standard simplex of $\mathbb{R}^{n-1}$. Especially in the case in which $\frac{\alpha}{\delta}$ is close to 1, a small variation of δ may lead to an extreme change of the set $\mathcal{D}$ (with a relative error on the points coordinates which is proportional to $\frac{\alpha}{\delta - \alpha}$) which eventually becomes the empty set.

A substantial improvement in the numerical stability of the solution of Problem 5.2 is also expected as a consequence of the application of numerical techniques. Indeed, since Problems 5.1 and 5.2 admit no closed form solution, the computation of the optimal portfolio necessarily requires the use of appropriate iterative methods, whose convergence speed and solution accuracy strictly depend on the condition number of the involved matrices (see [14, 23, 29]). In these terms, our approach shares the classical goal of "preconditioning", namely to decrease the condition number so to make the computations faster and more reliable.

In the following examples, the numerical solution is obtained applying the MatLab®'s function `quadprog`, a numerical solver to optimize quadratic objective functions with linear constraints, and the mathematical optimization solver Gurobi®. We use these softwares in their default settings, apart from the maximum number of algorithm's iterations allowed in `quadprog`, which for some computations is set to be equal to 10^6, the maximum number of iterations managed by MatLab®. All computations have been performed on a Quad-Core Intel Core i5 processor (at 2.3 GHz) running macOS.

The following examples are constructed on a dataset containing the historical daily returns from January 3rd, 2000 to September, 17th 2020 of the 10 sectors of the S&P index. The analysis is performed adopting a "rolling-sample" approach, in which the covariance matrix and the vector of expected returns are iteratively computed on the data contained in a sliding window of length $w_e = 240$ days. Specifically, at the ith iteration, we address Problems 5.1 and 5.2 where V and μ are computed using the data collected from ith to $(w_e + i - 1)$th day and the expected return μ_p is randomly chosen within the interval detected by the minimum and the maximum of the elements of μ. We underline that the economic comments on the asset allocation is beyond the scope of the current paper, while we want to stress the fact that numerical issues can be found on real data applications and not only in ad hoc theoretical examples. Moreover, numerical issues appear with relatively small optimization problems; our examples are built on 10-asset portfolios, confirming that portfolio's size is not necessary to generate computational difficulties.

Example 5.2: At each instance of the rolling window, Problems 5.1 and 5.2 are addressed by applying `quadprog`, both in its default setting or extending its maximum number of iterations to 10^6 (we will refer to such setting as `quadprog (MaxIt)`), and using software Gurobi®. The relevant variables of the 5165 computations are gathered in Tables 1 and 2. We briefly summarize the represented fields: *Conv.rate* contains the algorithm's percentage of convergence, while the fields *Time* and *Iterations*, contain the subfields *min*, *max*, *mean*, *std*, *total* which report the minimum value, the maximum value, the arithmetic mean, the standard deviation and the sum of the computational time across all 5165 computations and number of iterations respectively. Note that the computational time refers to the entire time to execute the corresponding problem, including the initial SVD computation.

Table 1: Performances of the solvers `quadprog`, `quadprog(MaxIt)` and Gurobi in case of Example 5.2: results for Problem 5.1.

		`quadprog`	`quadprog`(**MaxIt**)	**Gurobi**
Conv. rate		99.71%	99.98%	100%
Time	max	0.0200 s	0.0300 s	0.1000 s
	mean	0.0053 s	0.0064 s	0.0070 s
	std	0.0059 s	0.0064 s	0.0085 s
	total	27.5500 s	32.8600 s	36.2100 s
Iterations	min	4	4	6
	max	201	769	13
	mean	6.3748	6.4848	9.5332
	std	3.1175	10.7237	0.9368

Tables 1 and 2 contain the performances of the three considered solvers, `quadprog`, `quadprog(MaxIt)` and Gurobi, applied to Problems 5.1 and 5.2 respectively: for all of them the improvements are remarkable. The values of the convergence rate show that, while `Gurobi` always converges to the optimal solution, `quadprog` fails to solve the problem in the original formulation in a significant number of cases. In particular, the frequency of convergence depends on the maximum number of iterations allowed: if we raise it (see the entry corresponding to `quadprog (MaxIt)`), the convergence rate increases, nonetheless there exist still cases in which it fails to converge. On the other hand, in the case of the problem's equivalent reformulation, `quadprog (MaxIt)` shows a convergence rate of 100% making it competitive with Gurobi. Regarding the

Table 2: Performances of the solvers `quadprog`, `quadprog(MaxIt)` and Gurobi in case of Example 5.2: results for Problem 5.2.

		`quadprog`	`quadprog`(**MaxIt**)	**Gurobi**
Conv. rate		99.98%	100%	100%
Time	max	0.0300 s	0.2300 s	0.0700 s
	mean	0.0023 s	0.0018 s	0.0048 s
	std	0.0043 s	0.0039 s	0.0069 s
	total	11.7900 s	9.5100 s	24.7200 s
Iterations	min	4	4	7
	max	201	769	14
	mean	6.3164	6.4263	9.7164
	std	2.9785	10.6847	0.9152

computational time we notice for all the solvers a significant reduction passing from Table 1 to Table 2. In particular, `quadprog` reduces the execution time up to approximately 70%, showing a value which is competitive with that of `Gurobi` which registers a reduction of approximately 30%. The number of iterations seems to be invariant under the problem's reformulation; finally, a direct comparison of its values is inappropriate among different solvers as different algorithms are employed.

In order to stress the role of the restrictions as a source of numerical instability, we modify the allocation problem of the previous example, see Problem 5.1, by substituting the covariance matrix V with the identity I_n. In this case, $K_2(I_n) = 1$ and consequently the eventual numerical instability of the model necessarily depends on the restrictions. This formulation is different from Markowitz portfolio optimization model. Nevertheless, it is common in financial applications to define new portfolio optimization models simply modifying the objective function in Problem 5.1 and maintaining the restrictions. In particular, a bunch of financial literature looks for the portfolio that maximize a given diversification measure, see for example [5]. When we substitute V with I_n in the objective function, we obtain the Herfindahl concentration Index, see among the others [21], that is a very popular measure of concentration.[5] We further note that, independently from the economic interpretation, $x^t I_n x = \|x\|^2$.

Example 5.3: Using the same approach outlined in Example 5.2, Problems 5.1 and 5.2, with $V = I_{10}$, are addressed applying the three solvers `quadprog`, `quadprog (MaxIt)` and the software Gurobi. Tables 3 and 4 summarize the behaviour of the main variables (we refer to Example 5.2 for an explanation of the represented fields).

Table 3: Performances of the solvers `quadprog`, `quadprog(MaxIt)` and Gurobi in case of Example 5.3: results for Problem 5.1.

		`quadprog`	`quadprog`**(MaxIt)**	**Gurobi**
Conv. rate		95.35%	97.31%	100%
Time	max	0.0700 s	7.7600 s	0.0500 s
	mean	0.0045 s	0.0081 s	0.0057 s
	std	0.0057 s	0.1082 s	0.0065 s
	total	23.4800 s	41.8500	29.2000 s
Iterations	min	4	4	7
	max	201	10^6	13
	mean	13.5109	241.4736	8.9400
	std	31.5483	$1.3974 \cdot 10^4$	0.8388

Comparing Tables 3 and 4, we observe that Example 5.3 is qualitatively very similar to Example 5.2, so that the comments and observations expressed in Example 5.2 are still valid. Additionally, it is interesting to notice how the numerical issues are more noticeable in this case, as well as the performance improvements when passing from Problem 5.1 to Problem 5.2. In Table 3 `quadprog`

[5] Deepening the details on the relation between concentration and diversification is beyond the scope of this paper. It is intuitive that minimizing concentration is somehow equivalent to maximize diversification.

Table 4: Performances of the solvers `quadprog`, `quadprog(MaxIt)` and Gurobi in case of Example 5.3: results for Problem 5.2.

		`quadprog`	`quadprog`(MaxIt)	**Gurobi**
Conv. rate		100%	100%	100%
Time	max	0.0200 s	0.0200 s	0.0400 s
	mean	0.0023 s	0.0019 s	0.0039 s
	std	0.0043 s	0.0040 s	0.0057 s
	total	12.0300 s	9.9700 s	20.1000 s
Iterations	min	4	4	7
	max	9	9	13
	mean	5.6705	5.6705	8.9400
	std	0.7947	0.7947	0.8388

shows a convergence rate of about 95.35% which increases to the total success of 100% in Table 4. The performance improvements for all the solvers are also evident when considering the computational time and the number of iterations, whose values conspicuously decrease in the case of the problem's reformulation (in particular, note that the huge values of `quadprog(MaxIt)` entries in Table 3 refer to the occurrence of several non-convergence cases). These results, obtained when the objective function depends on the identity matrix, empirically support our idea of treating the equality restrictions to possibly improve the problem's numerical stability.

6. Conclusions

In this paper we analyze the numerical stability of QP problems focusing our investigation on the restrictions of the problem as a prominent potential source of ill-conditioning. We propose an equivalent version of the original problem that is showed, both theoretically and empirically, to overcome numerical issues related to ill-conditioning. Our approach has the appealing feature to efficiently handle almost collinearity besides the most manageable bad-scaling issue. Strong empirical evidence on real financial data in portfolio optimization context supports the idea that the proposed approach significantly impacts the performances of numerical solvers. We think that the cost of rewriting the model in the proposed equivalent version is negligible if compared to the improvements in terms of numerical stability.

Declarations

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contribution

In contrast with the widespread unpleasant habit to distinguish the individual contribution to the parts of the paper, the authors want to highlight that this paper is the result of a strict collaboration, where each individual expertize has been strictly necessary to the result. The authors are conscious that none of them was able to conceptualize the paper without the contribution of the others.

Data availablility

Data will be made available on reasonable request.

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, *The Review of Financial Studies*, Vol. 4, No 2, 315 – 342 (1991)

[2] Best, M.J., Hlouskova, J., Quadratic programming with transaction costs, *Computers & Operations Research*, Volume 35, Issue 1, 18–33 (2008)

[3] Bini, D., Pan, V.Y., *Polynomial and Matrix Computations - Fundamental Algorithms*, Progress in Theoretical Computer Science, Birkhauser (1994)

[4] Brodie, J., Daubechies, I., De Mol, C., Giannone D., Loris, I., Sparse and stable Markowitz portfolios, *PNAS*, Vol. 106, No. 30, 12267–12272 (2009)

[5] Choueifaty, Y., Coignard, Y., Toward maximum diversification, *Journal of Portfolio Management* **35**, 40 – 51 (2008)

[6] Clarke, R., de Silva, H., Thorley, S., Portfolio Constraints and the Fundamental Law of Active Management, *Financial Analysts Journal*, 58:5, 48 – 66 (2002)

[7] Constantinides, G. M., Malliaris, A.G., *Portfolio Theory*, R. Jarrow et al., Eds., Handbooks in OR & MS, Vol. 9 (1995)

[8] DeMiguel, V., Garlappi, L., Uppal, R., Optimal versus Naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*, 1915 – 1953 (2009)

[9] Donoho, D.L., For most large underdetermined systems of linear equations the minimal $\ell_1$norm solution is also the sparsest solution. *Comm. Pure Appl. Math.*, 59, 797-829 (2006)

[10] Fassino, C., Torrente, M., Uberti, P., Numerical Stability of Optimal Mean Variance Portfolios. In: Corazza M., Gilli M., Perna C., Pizzi C., Sibillo M. (eds) Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer, Cham (2021)

[11] Frost, P.A., Savarino, J.E., An Empirical Bayes Approach to Efficient Portfolio Selection, *The Journal of Financial and Quantitative Analysis*, Vol. 21, No. 3, 293–305 (1986)

[12] Fung, G.M., Mangasarian, O.L., Equivalence of Minimal ℓ_0- and ℓ_p-Norm Solutions of Linear Equalities, Inequalities and Linear Programs for Sufficiently Small p, *J Optim Theory Appl*, **151**, 1–10 (2011)

[13] Gilli, M., Maringer, D., Schumann, E., *Numerical methods and optimization in finance*, Academic Press, (2019)

[14] Golub, G.H., Van Loan, C.F., *Matrix Computations*, Second Edition, The Johns Hopkins University Press, Baltimore-London (1989)

[15] Hirschberger, M., Qi, Y., Steuer, R.E., Large-scale MV efficient frontier computation via a procedure of parametric quadratic programming, *European Journal of Operational Research*, Vol. 204, Issue 3, 581 – 588 (2010)

[16] Hong, Y.P., Pan, C.T., Rank-revealing QR factorizations and the singular value decomposition, *Math. Comp.*, 58 (197), 213 – 232 (1992)

[17] Kan, R., Smith, D.R., The Distribution of the Sample Minimum-Variance Frontier, *Management Science*, Vol. 54, No. 7, 1364–1380 (2008)

[18] Kan, R., Zhou, G., Optimal Portfolio Choice with Parameter Uncertainty, *The Journal of Financial and Quantitative Analysis*, Vol. 42, No. 3, 621–656 (2007)

[19] Kim, J.H., Kim, W. C., Fabozzi, F. J., Recent Developments in Robust Portfolios with a Worst-Case Approach, *J. Optimiz. Theory App.*, Vol. 161, No. 1, 103–121 (2014)

[20] Kolm, P. N., Ttnc, R., Fabozzi, F. J., 60 Years of portfolio optimization: Practical challenges and current trends, *European Journal of Operational Research*, 234(2), 356 – 371 (2014)

[21] Kwoka, J.E., Large firms dominance and price cost margins in manufacturing industries. *Southern Economic Journal* 44, 183 – 189 (1977)

[22] Ledoit, O., Wolf, M., Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *J. Empir. Finance*, Vol. 10, No. 5, 603–621 (2003)

[23] Liesen, J., Tichy, P., Convergence analysis of Krylov subspace methods, *GAMM-Mitteilungen*, Vol. 27, No. 2, 153–173 (2005)

[24] Madych, W.R., Solutions of Undetermined Systems of Linear Equations, *Lectures Notes-Monograph Series*, Vol 20 Spatial Statistics and Imaging, 227 – 238 (1991)

[25] Markowitz, H., Portfolio Selection, *The Journal of Finance*, Vol. 7, No. 1, 77–91 (1952)

[26] Nocedal, J., Wright, S., *Numerical Optimization*, Springer (2006)

[27] Perold, A. F., Large-scale portfolio optimization, *Management science*, 30(10), 1143 – 1160 (1984)

[28] Pflug, G.Ch., Pichler, A., Wozabal, D., The 1/N investment strategy is optimal under high model ambiguity, *Journal of Banking & Finance*, Vol. 36, 410–417 (2012)

[29] Pyzara, A., Bylina B., Bylina, J., The influence of a matrix condition number on iterative methods' convergence, Federated Conference on Computer Science and Information Systems (FedCSIS), 459–464 (2011)

[30] Rosadi, D., Setiawan, E.P., Templ, M., Filzmoser, P., Robust covariance estimators for mean-variance portfolio optimization with transaction lots, *Operations Research Perspectives*, Volume 7 (2020)

[31] Zhang, Y., Li, X., Guo, S., Portfolio selection problems with Markowitz?s mean-variance framework: a review of literature, *Fuzzy Optimization and Decision Making*, 17(2), 125 – 158 (2018)

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: