



**HAL**  
open science

# A dual approach to semidefinite least-squares problems

Jérôme Malick

► **To cite this version:**

Jérôme Malick. A dual approach to semidefinite least-squares problems. [Research Report] RR-4212, INRIA. 2001. inria-00072409v1

**HAL Id: inria-00072409**

**<https://inria.hal.science/inria-00072409v1>**

Submitted on 24 May 2006 (v1), last revised 25 Mar 2013 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*An efficient dual algorithm to solve  
conic least-square problems*

Jérôme MALICK

**N° 4212**

Juin 2001

THÈME 4



*Rapport  
de recherche*



# An efficient dual algorithm to solve conic least-square problems

Jérôme MALICK

Thème 4 — Simulation et optimisation  
de systèmes complexes  
Projet Numopt

Rapport de recherche n° 4212 — Juin 2001 — 20 pages

**Abstract:** We propose a general method to solve the problem which we call “conic least-square”: to project, in a Euclidean space, onto the intersection of a closed convex cone and of affine hyperplanes. We are particularly interested in the case of symmetric positive semidefinite matrices, which has many applications in finance.

First, we study in a novative way the differentiability properties of the distance to a closed convex cone. Then we propose a partial Lagrangian dualization of the conic least-square problem, which leads us to a convex differentiable dual problem. We assess this approach with some numerical experiments.

**Key-words:** Convex optimization, eigenvalue optimization, semidefinite optimization, Moreau orthogonal decomposition, Moreau-Yosida regularization, Lagrangian duality.

# Un algorithme dual efficace pour résoudre les problèmes de moindres carrés coniques

**Résumé :** Nous proposons une méthode générale pour résoudre le problème que nous appelons ‘de moindres carrés conique’ : calculer la projection, dans un espace euclidien, sur l’intersection d’un cône convexe fermé et d’hyperplans affines. Nous nous intéressons de manière toute particulière au cas du cône des matrices symétriques semi-définies positives, dont les applications sont nombreuses en finance.

Tout d’abord, nous étudions de manière originale les propriétés de différentiabilité de la fonction distance à un cône convexe fermé. Nous proposons ensuite une dualisation lagrangienne partielle du problème des moindres carrés conique, qui aboutit à un problème dual convexe et différentiable. Nous illustrons la méthode par quelques résultats numériques.

**Mots-clés :** Optimisation convexe, optimisation des valeurs propres, régularisation de Moreau-Yosida, projection sur un convexe fermé, décomposition de Moreau, dualité lagrangienne, optimisation semi-définie positive.

# 1 Introduction, motivation

## 1.1 To find the best conic approximation

In this paper we study the projection onto convex subsets of a cone in a Euclidean space in general, and provide a particular treatment for the cone of positive semidefinite matrices. We call conic least-squares the class of problems consisting in computing the projection onto the convex set defined by the intersection of the considered cone and affine hyperplanes. When considering the particular semidefinite case, we speak of semidefinite least-square.

Problems of these classes, notably semidefinite least-squares, arise in many different fields of numerical and applied mathematics. For instance, a “good” approximation covariance matrix between  $n$  assets, which plays a key role in portfolio risk analysis, could be obtained from a first estimate by solving a semidefinite least-square (this is developed in Subsection 6.4).

The aim of this paper is twofold:

1. to analyse, in an original way, properties of distance functions to a closed convex cone, and notably to give explicit formulae for the semidefinite case;
2. to propose an efficient dual algorithm to solve conic least-square problems based on a partial Lagrangian dualization, using these properties.

The paper is organized as follows. After presenting notation and illustrations of the general conic least-square problem, we focus, in Section 2, on the case where there are no affine constraints. We recover known properties of distance functions to a cone with convex analysis tools. Section 3 is devoted to a similar study of the semidefinite least-square problem, using results on spectral functions. In Section 4, we show that a partial Lagrangian dualization of constraints yields a dual problem which is convex and differentiable. An efficient algorithm (which we call the conic dual algorithm) is proposed in Section 5 to tackle the general conic least-square problem. Computational results, comparison with existing techniques and applications of the semidefinite version of this algorithm are presented in Section 6.

## 1.2 Basic notation and terminology

Our notation is classical. It closely follows that of [LO99] and [HUL93]. For more clarity, let us recall some of them which we will constantly use.

The general framework of this paper is a *Euclidean space*, say  $\mathbb{R}^p$ , equipped with a scalar product  $\langle \cdot, \cdot \rangle_p$ . We will denote by  $\|\cdot\|_p$  the associated norm. When no confusion is possible, we often drop the dependence on the dimension  $p$ . Let us recall some well-known results (see among others [HUL93, Chap. III]). The projection onto a closed convex set  $C$  is well-defined: the projection of a point  $c \in \mathbb{R}^p$  is the unique  $x_c \in C$  which satisfies the following variational inequality (sometimes called the angle-characterization of the projection)

$$\langle c - x_c, x - x_c \rangle \leq 0 \quad \text{for all } x \in C. \quad (1)$$

We remind also that the projection is a 1-Lipschitzian function. Besides a subset  $\mathcal{K}$  of  $\mathbb{R}^p$  is said to be a cone if the half line  $\{\alpha x : \alpha \geq 0\}$  is entirely contained in  $\mathcal{K}$  whenever  $x \in \mathcal{K}$ . In

this paper,  $\mathcal{K}$  will stand for a *closed convex cone* of  $\mathbb{R}^p$ . The projection onto  $\mathcal{K}$ , denoted by  $p_{\mathcal{K}} : \mathbb{R}^p \longrightarrow \mathcal{K}$ , is then well-defined and 1-Lipschitzian.

We consider, in particular, the Euclidian space of  $n \times n$  symmetric matrices  $\mathcal{S}_n$  (here  $p = \frac{n(n+1)}{2}$ ), equipped with the restriction of the ordinary dot-product in  $\mathbb{R}^{n^2}$ , also called Frobenius scalar product

$$\forall A, B \in \mathbb{R}^{n^2} \quad \langle A, B \rangle_{n^2} := \text{tr} A^T B = \sum_{i,j=1}^n A_{ij} B_{ij},$$

where  $\text{tr}(X)$  stands for the trace of the matrix  $X$ . Recall also that, for  $C \in \mathcal{S}_n$ , we have the two following basic properties. First, all eigenvalues of  $C$  are real; we will denote them by  $\lambda_i(C)$ , and rank them in nonincreasing order

$$\lambda_1(C) \geq \lambda_2(C) \geq \dots \geq \lambda_n(C).$$

In the remainder of this paper,  $\lambda(C)$  will stand for  $[\lambda_1(C), \dots, \lambda_n(C)]^T$ . Secondly,  $C$  is diagonalisable in an orthonormal basis  $\mathcal{E}_C = \{e_i(C)\}_{i=1,\dots,n}$  of  $\mathbb{R}^n$  formed by eigenvectors of  $C$ . Here the orthogonal matrix  $W_C$  stands for the transform matrix from the canonical basis to  $\mathcal{E}_C$ ,

$$C = W_C (\text{Diag} \lambda(C)) W_C^T = \sum_{i=1}^n \lambda_i(C) e_i(C) e_i(C)^T.$$

This last expression will be called the *spectral decomposition* of  $C$ . We will often drop the dependence on  $C$  from our notation.

In  $\mathcal{S}_n$  we are interested in the cone  $\mathcal{K} = \mathcal{S}_n^+$  (resp.  $\mathcal{S}_n^-$ ) of positive semidefinite (resp. negative semidefinite) matrices. Observe that  $\mathcal{S}_n^+ = \{X \in \mathcal{S}_n : \lambda_n(X) \geq 0\}$ , where the smallest eigenvalue  $\lambda_n(X) = \min_{\|u\|=1} u^T X u$  appears as a concave continuous function of  $X$ . This shows that the cone  $\mathcal{S}_n^+$  (and also  $\mathcal{S}_n^- = -\mathcal{S}_n^+$ ) is convex and closed. We will often use the notation  $X \succeq 0$  to express that a (symmetric) matrix  $X$  lies in  $\mathcal{S}_n^+$ .

To end up this subsection, we give below a very short glossary of symbols and terms which may not be universal:

- the adjoint of a linear mapping  $A$  which will be denoted by  $A^*$ ;
- for any vector  $x$  in  $\mathbb{R}^n$ ,  $\text{Diag} x$  will denote the diagonal matrix with the vector  $x$  on the main diagonal. Let also its adjoint operator  $\text{diag} : \mathcal{S}_n \rightarrow \mathbb{R}^n$  be defined by  $\text{diag}(A) = [a_{11}, \dots, a_{nn}]^T$ ;
- $I_p$  denotes as usual the identity matrix in  $\mathbb{R}^p$ ;
- $a^+ := \max\{0, a\}$  stands for the nonnegative part of  $a \in \mathbb{R}$  and  $a^- := \min\{0, a\}$  the nonpositive; we thus have  $a = a^+ + a^-$  and  $|a| = a^+ - a^-$ .

### 1.3 Formulation

The problems we will focus on in this paper can be expressed as follows. Let  $\mathcal{K}$  be a closed convex cone of  $\mathbb{R}^p$ . Let a vector  $c \in \mathbb{R}^p$ , a vector  $b \in \mathbb{R}^m$  and a linear operator  $\mathcal{A} : \mathbb{R}^p \longrightarrow \mathbb{R}^m$  be given. We want to compute the projection of the vector  $c$  onto the closed convex subset of  $\mathbb{R}^p$  formed by the intersection of  $\mathcal{K}$  and affine hyperplanes defined by  $\mathcal{A}$ . In others words, we

want to get the best conic approximation of  $c$  which satisfies affine constraint defined by  $\mathcal{A}$ . Our goal is thus to find an “efficient” algorithm to solve optimization programs of the type:

$$\begin{cases} \inf & \frac{1}{2} \|x - c\|^2 \\ & \mathcal{A}x = b \\ & x \in \mathcal{K}. \end{cases}$$

The very first remark is that, if the feasible domain is nonempty, there exists a unique  $x^*$  which achieves the above infimum. In what follows, we therefore use the notation  $\min$  rather than  $\inf$  for the conic least-square problem. As we are in a Euclidean space, each component function of  $\mathcal{A}$  can also be expressed as a scalar product. This means there exist  $m$  elements  $a_i \in \mathbb{R}^p$  such that, for any  $x \in \mathbb{R}^p$  there holds

$$\mathcal{A}(x) = \begin{bmatrix} \langle a_1, x \rangle \\ \vdots \\ \langle a_m, x \rangle \end{bmatrix}. \quad (2)$$

Therefore an equivalent formulation is, for  $b$ ,  $a_i$  and  $c$  given:

$$\begin{cases} \min & \frac{1}{2} \|x - c\|^2 \\ & \langle a_i, x \rangle = b_i, \quad i = 1, \dots, m \\ & x \in \mathcal{K}. \end{cases}$$

Here we give 2 particular instances of conic least-square problems.

**Example (Semidefinite least-square).** This is a very important subclass of our problem, that we will develop precisely in Section 6. We will notably detail an algorithm to solve problems of this class. Besides the computation of a “good” approximation of the covariance matrix of  $n$  assets can be expressed as a semidefinite least-square problem (see Subsection 6.4).

Here the cone  $\mathcal{K}$  is  $\mathcal{S}_n^+$ . A optimization program is written under a semidefinite least-square form if there are a matrix  $C \in \mathcal{S}_n$ , a vector  $b \in \mathbb{R}^m$  and a linear operator  $\mathcal{A} : \mathcal{S}_n \rightarrow \mathbb{R}^m$  such that:

$$\begin{cases} \min & \frac{1}{2} \|X - C\|^2 \\ & \mathcal{A}X = b \\ & X \succeq 0. \end{cases}$$

If  $\mathcal{A}$  is expressed via its  $m$  component functions, consider  $m$  symmetric matrices which transform the program into:

$$\begin{cases} \min & \frac{1}{2} \|X - C\|^2 \\ & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ & X \succeq 0. \end{cases}$$

In many instances, each  $A_i$  is often a  $\{0, 1\}$ -matrix with only 2 nonzero entries. The corresponding entries of  $X$  are then fixed: a structure is imposed to the matrix.  $\square$

**Example (Preconditioning problem).** In another application,  $\mathcal{K}$  is  $\mathcal{S}_n^+ \times \mathcal{S}_n$ .



A difficult problem in linear algebra consists in looking for a preconditioner  $P$  approximating the inverse of a symmetric matrix  $A$  and with the same inertia as  $A$ . This problem amounts to solve following optimization problem:

$$\begin{cases} \min & \frac{1}{2} \|(2I_n - Q) - X\|^2 + \alpha \|P\|^2 \\ & \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m \\ & X = PA + AP - Q, P \in \mathcal{S}_n \\ & X \succeq 0, \end{cases}$$

for a given  $\alpha$  (see [FO01]). A direct way to handle this problem is then to consider the cone  $\mathcal{K} = \mathcal{S}_n^+ \times \mathcal{S}_n$  in  $\mathcal{S}_n \times \mathcal{S}_n$ , equipped with the Euclidean product-norm  $(\|\cdot\|^2 + \|\cdot\|^2)^{\frac{1}{2}}$ , and to transform the problem into a  $\mathcal{S}_n^+ \times \mathcal{S}_n$  least-square one. We denote by  $A \oplus B$  the couple  $(A, B) \in \mathcal{S}_n \times \mathcal{S}_n$ , and similarly  $\mathcal{A}_1 \oplus \mathcal{A}_2$  a linear operator on  $\mathcal{S}_n \times \mathcal{S}_n$ . For instance the objective function is then written as  $\frac{1}{2} \|(2I_n - Q) \oplus 0 - X \oplus \sqrt{\alpha}P\|^2$ . Scaling  $P$  to eliminate  $\sqrt{\alpha}$ , we obtain an  $\mathcal{S}_n^+ \times \mathcal{S}_n$ -least-square problem:

$$\begin{cases} \min & \frac{1}{2} \|C \oplus 0 - X \oplus P\|^2 \\ & \langle A_i \oplus 0, X \oplus P \rangle = b_i, \quad i = 1, \dots, m \\ & \mathcal{B}(X \oplus P) = Q \\ & X \oplus P \in \mathcal{S}_n^+ \times \mathcal{S}_n, \end{cases}$$

where  $C := (2I_n - Q)$  and  $\mathcal{B}$  has been introduced to express simply affine constraints. Denoting by  $\mathcal{L}_A : P \mapsto \frac{1}{\sqrt{\alpha}}(PA + AP)$  the Lyapunov operator, we introduce the operator  $\mathcal{B}$  on  $\mathcal{S}_n \times \mathcal{S}_n$  defined by  $\mathcal{B} := I_{\mathcal{S}_n} \oplus 0 - 0 \oplus \mathcal{L}_A$ . The constraint  $X - \frac{1}{\sqrt{\alpha}}(PA + AP) + Q = 0$  is then expressed as  $\mathcal{B}(X \oplus P) = Q$ .  $\square$

## 2 Projection onto a closed convex cone

To begin with, we isolate the problem of computing the projection  $p_{\mathcal{K}}(c)$  of a fixed  $c \in \mathbb{R}^p$  onto a closed convex cone  $\mathcal{K}$ . The aim of this section is to cast a fresh glance on this known problem.

### 2.1 Moreau's theorem

The projection onto  $\mathcal{K}$  enjoys properties which come close to those of the projection onto a subspace. The attached subset, which plays the role of the orthogonal subspace, is the *polar cone*  $\mathcal{K}^\circ$  of  $\mathcal{K}$ , defined ([HUL93, Chap. III]) by:

$$\mathcal{K}^\circ := \{s \in \mathbb{R}^p : \langle s, x \rangle \leq 0 \quad \forall x \in \mathcal{K}\}.$$

A very first observation is that  $\mathcal{K}^\circ$  is also closed, convex. An important property is that  $(\mathcal{K}^\circ)^\circ = \mathcal{K}$ . It is clear that the polar cone of a linear subspace is its orthogonal subspace. There is then a decomposition result which somehow generalizes the decomposition of a vector space into the direct sum of a (finite-dimensional) subspace and its orthogonal.

**Theorem 2.1 (Moreau)** *Let  $\mathcal{K}$  be a closed convex cone. For the three elements  $x$ ,  $x_1$  and  $x_2$  in  $\mathbb{R}^p$ , the two properties below are equivalent:*

- (i)  $x = x_1 + x_2$  with  $x_1 \in \mathcal{K}$ ,  $x_2 \in \mathcal{K}^\circ$  and  $\langle x_1, x_2 \rangle = 0$ ,  
(ii)  $x_1 = p_{\mathcal{K}}(x)$  and  $x_2 = p_{\mathcal{K}^\circ}(x)$ .

A proof of this theorem is partly given in ([HUL93, Chap. III]). For simplicity and completeness, we give here a simple proof just based on the angle-characterization (1) of the projection onto a closed convex set.

**Proof.** (i)  $\Rightarrow$  (ii). We will prove that  $x_1$  is the projection of  $x$  onto  $\mathcal{K}$ . By changing the roles of  $x_1$  and  $x_2$  in this proof,  $x_2 = p_{\mathcal{K}^\circ}(x)$  is straightforward. Since  $x_2 = x - x_1$  lies in  $\mathcal{K}^\circ$ , we have

$$\forall y \in \mathcal{K}, \quad \langle x - x_1, y \rangle \leq 0. \quad (3)$$

Besides we have by assumption  $\langle x - x_1, x_1 \rangle = 0$ , so that

$$\forall y \in \mathcal{K}, \quad \langle x - x_1, y - x_1 \rangle \leq 0.$$

Since  $x_1$  lies in  $\mathcal{K}$ , this exactly says from (1) that  $x_1 = p_{\mathcal{K}}(x)$ .

(ii)  $\Rightarrow$  (i). We prove first that  $\langle x - x_1, x_1 \rangle = 0$  and next that  $x_2 = x - x_1$ . Since  $\mathcal{K}$  is a cone, take (1) for  $y = 0$  and  $y = 2x_1$  to get two inequalities which imply

$$\langle x - x_1, x_1 \rangle = 0. \quad (4)$$

Then from (1)

$$0 \geq \langle x - x_1, y - x_1 \rangle = \langle x - x_1, y \rangle \quad \text{for all } y \in \mathcal{K}^\circ.$$

Hence  $x - x_1 \in \mathcal{K}^\circ$ . Besides, since  $x_1 \in \mathcal{K}$  we have for  $y \in \mathcal{K}^\circ$

$$\langle x - (x - x_1), y - (x - x_1) \rangle = \langle x_1, y \rangle - \langle x_1, x - x_1 \rangle = \langle x_1, y \rangle \leq 0.$$

Invoke the uniqueness of the projection onto  $\mathcal{K}^\circ$  to get  $x = x_1 + x_2$ . This ends the proof.  $\square$

The decomposition  $x = p_{\mathcal{K}}(x) + p_{\mathcal{K}^\circ}(x)$  will be called the *polar decomposition* of  $x$ . Notice that we then have:

$$\|x\|^2 = \|p_{\mathcal{K}}(x) + p_{\mathcal{K}^\circ}(x)\|^2 = \|p_{\mathcal{K}}(x)\|^2 + \|p_{\mathcal{K}^\circ}(x)\|^2. \quad (5)$$

## 2.2 Moreau-Yosida regularization

In this subsection, we study variational properties of the half squared distance to  $\mathcal{K}$ , which will appear naturally in Section 4:

$$\begin{aligned} d_{\mathcal{K}} : \mathbb{R}^p &\longrightarrow \mathbb{R} \\ x &\longmapsto \min_{y \in \mathcal{K}} \frac{1}{2} \|x - y\|^2. \end{aligned} \quad (6)$$

Its properties are stated in Theorems 2.2 and 2.3. They are not new, and can be proved with basic tools. Here we show that they are close to convex analysis by the Moreau-Yosida regularization.

We start by observing that there is another useful expression of  $d_{\mathcal{K}}$ . By definition the above minimum is reached at the unique point  $p_{\mathcal{K}}(x)$ . Then Theorem 2.1 yields

$$d_{\mathcal{K}}(x) = \frac{1}{2} \|x - p_{\mathcal{K}}(x)\|^2 = \frac{1}{2} \|p_{\mathcal{K}^\circ}(x)\|^2. \quad (7)$$

To guide intuition, consider an easy case: when  $\mathcal{K} = E$  is a linear subspace of  $\mathbb{R}^p$ . The orthogonal subspace  $E^\perp$  is the polar cone and the projection is then a linear mapping. Therefore  $d_E$  is differentiable (actually  $C^\infty$ ) on  $\mathbb{R}^p$  and  $\nabla d_E(x) = p_{E^\perp}(x)$  for  $x \in \mathbb{R}^p$ . In a general framework, this last fact persists; we prove it with the help of the *Moreau-Yosida regularization*, described in [HUL93, Chap. XV and Example 3.4.4 of Chap. XI]. For a convex function  $f$  on  $\mathbb{R}^p$ , we define the Moreau-Yosida regularization of  $f$  to be the function

$$x \longmapsto \min_{y \in \mathbb{R}^p} \left\{ f(y) + \frac{1}{2} \|x - y\|^2 \right\}.$$

**Theorem 2.2** *Let  $\mathcal{K}$  be a closed convex cone in  $\mathbb{R}^p$ . Then the function defined by (7) is a convex differentiable function mapping  $\mathbb{R}^p$  to  $\mathbb{R}$ , whose gradient is for  $x \in \mathbb{R}^p$*

$$\nabla d_{\mathcal{K}}(x) = p_{\mathcal{K}^\circ}(x). \quad (8)$$

*Furthermore the gradient function is Lipschitz continuous with a Lipschitz constant equal to 1.*

**Proof.** Let  $I_{\mathcal{K}}$  be the indicator function of  $\mathcal{K}$  (whose values are 0 on  $\mathcal{K}$  and  $+\infty$  elsewhere). The theorem is just Theorem 4.1.4 of [HUL93, Chap. XV] written in our case, since  $d_{\mathcal{K}}$  can be interpreted as the Moreau-Yosida regularization of  $I_{\mathcal{K}}$

$$d_{\mathcal{K}}(x) = \min_{y \in \mathbb{R}^p} \left\{ \frac{1}{2} \|x - y\|^2 + I_{\mathcal{K}}(y) \right\}.$$

We get in particular  $\nabla d_{\mathcal{K}}(x) = x - p_{\mathcal{K}}(x) = p_{\mathcal{K}^\circ}(x)$ . The Lipschitz property of the gradient mapping is given in general by the quoted theorem. However this is here a triviality since the gradient is a projection.  $\square$

When  $\mathcal{K} = E$ , the gradient  $\nabla d_E(x) = p_{E^\perp}(x)$  is a linear mapping, hence clearly differentiable over  $\mathbb{R}^p$ . This property is “almost” kept by passing to the general case.

**Theorem 2.3** *Let  $\mathcal{K}$  be a closed convex cone in  $\mathbb{R}^p$ . Then the function defined by (7) is twice differentiable almost everywhere on  $\mathbb{R}^p$ .*

Because  $\nabla d_{\mathcal{K}} = p_{\mathcal{K}^\circ}$  is (globally) Lipschitzian on  $\mathbb{R}^p$ , this statement is direct consequence of the following general result.

**Theorem 2.4 (Rademacher’s Theorem, 9.60 in [RW98])** *Let  $F$  be a function locally Lipschitzian on  $\mathbb{R}^p$ . Let  $\mathcal{D}$  be the subset of  $\mathbb{R}^p$  consisting of the points where  $F$  is differentiable. Then  $\mathbb{R}^p \setminus \mathcal{D}$  is a negligible set. In particular  $\mathcal{D}$  is dense into  $\mathbb{R}^p$ .*

We cannot hope to have more than the almost everywhere differentiability in the general framework. A counterexample is given in Theorem 3.8 where  $d_{\mathcal{S}_n^+}$  is shown *not* to be twice differentiable everywhere on  $\mathcal{S}_n$ .

### 3 Projection onto $\mathcal{S}_n^+$

In this section, we consider the semidefinite least-square problem without any affine constraint. In Subsection 3.1, we recall (and prove in an original way) our major theorem: an explicit formula for the projection onto  $\mathcal{S}_n^+$ . In Subsection 3.2, we get (and go further) variational results on  $d_{\mathcal{S}_n^+}$  and  $d_{\mathcal{S}_n^-}$  by applying the corresponding general properties of Subsection 2.2, and also by considering these mappings from another point of view.

#### 3.1 The projections

We give in this subsection the theorem which expresses the projection of  $C$  onto  $\mathcal{S}_n^+$ . This theorem is actually not new: it is a well-known result of linear algebra established by Eckart, Young and Higham (see [HS95] for discussions on this topic). The proof they gave is based on linear algebra results. Here we will see that this theorem is a straightforward consequence of the convex analysis-based Theorem 2.1.

To a symmetric matrix  $C$ , associate  $C_+$  the “positive semidefinite part” of  $C$  (negative eigenvalues are set to zero)

$$C_+ := \sum_{i=1}^n \lambda_i^+ e_i e_i^T = W \begin{bmatrix} \max\{\lambda_1, 0\} & & \\ & \ddots & \\ & & \max\{\lambda_n, 0\} \end{bmatrix} W^T; \quad (9)$$

likewise the “negative semidefinite part” is

$$C_- := \sum_{i=1}^n \lambda_i^- e_i e_i^T = W \begin{bmatrix} \min\{\lambda_1, 0\} & & \\ & \ddots & \\ & & \min\{\lambda_n, 0\} \end{bmatrix} W^T. \quad (10)$$

We now give two results of linear algebra in  $\mathcal{S}_n$ . The first lemma is well known; it says that the polar cone of the cone of positive semidefinite matrices is the cone of negative semidefinite matrices. The second lemma is a simple (but crucial) remark to apply Moreau’s theorem.

**Lemma 3.1 (Lemma 3.1 in [LO99])** *The cone  $\mathcal{S}_n^+$  is antipolar, i.e.  $(\mathcal{S}_n^+)^o = -\mathcal{S}_n^+ = \mathcal{S}_n^-$ . We also have  $(\mathcal{S}_n^-)^o = \mathcal{S}_n^+$ . Thereby there notably holds, for any  $A \in \mathcal{S}_n^+$  and for any  $B \in \mathcal{S}_n^-$ ,  $\langle A, B \rangle \leq 0$ .*

**Lemma 3.2** *Let  $C \in \mathcal{S}_n$ . Consider  $C_+ \in \mathcal{S}_n^+$  and  $C_- \in \mathcal{S}_n^-$  defined by (9) and (10) respectively. Then there holds*

$$\langle C_+, C_- \rangle = 0.$$

**Proof.** This just results from the orthonormality of the family  $\{e_i\}_{i=1,\dots,n}$  of eigenvectors of  $C$ . In fact

$$C_+ C_- = \left( \sum_{\lambda_i < 0} \lambda_i e_i e_i^T \right) \left( \sum_{\lambda_j > 0} \lambda_j e_j e_j^T \right) = 0.$$

then the trace is zero too, i.e.  $\text{tr}(C_+ C_-) = \langle C_+, C_- \rangle = 0$ .  $\square$

We are now in position to state and prove the main result of the subsection.

**Theorem 3.3 (Eckart-Young-Higham)** *Let  $C \in \mathcal{S}_n$ . Then the projection  $p_{\mathcal{S}_n^+}(C)$  of  $C$  onto  $\mathcal{S}_n^+$  is the matrix  $C_+$ , defined by (9). Likewise the projection  $p_{\mathcal{S}_n^-}(C)$  of  $C$  onto the polar cone  $(\mathcal{S}_n^+)^o = \mathcal{S}_n^-$  is  $C_-$  defined by (10).*

**Proof.** It can be seen as a nice, direct application of Moreau's theorem (Theorem 2.1). It indeed suffices to notice that  $C = C_+ + C_-$ , that  $C_+ \in \mathcal{S}_n^+$  and  $C_- \in \mathcal{S}_n^- = (\mathcal{S}_n^+)^o$  by Lemma 3.1, and then Theorem 2.1 and Lemma 3.2 yields the results.  $\square$

This theorem justifies the names “positive[negative] semidefinite part” of a symmetric matrix. What made things simpler was to consider the projection onto  $\mathcal{S}_n^+$  and onto its polar cone simultaneously. We will continue to use this trick.

## 3.2 On distance functions

Similarly to (6), we will denote by  $d_+$  [resp.  $d_-$ ] the half squared distance to  $\mathcal{S}_n^+$  [resp.  $\mathcal{S}_n^-$ ]. The work done in Subsection 2.2 enables us to get easily:

**Theorem 3.4** *The functions  $d_+$  and  $d_-$  are convex, differentiable on  $\mathcal{S}_n$  and their gradient functions are Lipschitzian. For  $C \in \mathcal{S}_n$  we have*

$$\nabla d_+(C) = C_- \quad \nabla d_-(C) = C_+.$$

**Proof.** Apply Theorem 2.2 with  $\mathcal{K} = \mathcal{S}_n^+$  and  $\mathcal{K} = \mathcal{S}_n^-$ . Use Theorem 3.3 to express the gradient.  $\square$

The aim of this subsection is to recover this theorem with some matrix analysis results. After that, we will show that this viewpoint enables to precise the localisation of the twice differentiability.

From (7),  $d_+$  and  $d_-$  have another interesting expressions

$$\begin{aligned} d_+(C) &= \frac{1}{2} \|C_-\|^2 = \frac{1}{2} \sum_{i=1}^n (\lambda_i(C)^-)^2, \\ d_-(C) &= \frac{1}{2} \|C_+\|^2 = \frac{1}{2} \sum_{i=1}^n (\lambda_i(C)^+)^2. \end{aligned} \quad (11)$$

The functions  $d_+$  and  $d_-$  depend only on the eigenvalues of  $C$ . Functions of this kind are called *spectral functions*. A spectral function  $f$  can be written as

$$f(C) = F(\lambda(C)),$$

where  $F$  is a symmetric function (i.e. invariant under coordinate permutations) mapping  $\mathbb{R}^n$  to  $\mathbb{R}$ . For our two spectral functions, the symmetric functions in question are

$$D_+(x_1, \dots, x_n) = \frac{1}{2} \sum_{i=1}^n \min\{0, x_i\}^2 \quad D_-(x_1, \dots, x_n) = \frac{1}{2} \sum_{i=1}^n \max\{0, x_i\}^2, \quad (12)$$

since we have  $d_+(C) = D_+(\lambda(C))$  and  $d_-(C) = D_-(\lambda(C))$ . Let us recall theorems about spectral functions we need.

**Theorem 3.5 (Corollary 2.7 of [Lew96a])** *If the function  $F : \mathbb{R}^m \rightarrow \mathbb{R}$  is symmetric convex then the matrix function  $F \circ \lambda : \mathcal{S}_n \rightarrow \mathbb{R}$  is also convex.*

**Theorem 3.6 (Theorem 1.1 of [Lew96b])** *Let  $A \in \mathcal{S}_n$  and suppose  $\lambda(A)$  belongs to the domain of the symmetric function  $F$ . Then  $F$  is differentiable at the point  $\lambda(A)$  if and only if  $F \circ \lambda$  is differentiable at the point  $A$ . In that case we have the formula*

$$\nabla(f \circ \lambda)(A) = U(\text{Diag} \nabla F(\lambda(A)))U^T,$$

for any orthogonal matrix  $U$  satisfying  $A = U(\text{Diag}(\lambda(A)))U^T$ .

**Theorem 3.7 (Theorem 3.3 of [LS])** *Let  $A \in \mathcal{S}_n$ . The symmetric function  $F$  is twice differentiable at the point  $\lambda(A)$  if and only if the spectral function  $F \circ \lambda$  is twice differentiable at the matrix  $A$ .*

These three major theorems yield Theorem 3.4, and a little more.

**Proof (Another proof of Theorem 3.4).** Let  $C \in \mathcal{S}_n$  with spectral decomposition  $C = \sum_{i=1}^n \lambda_i e_i e_i^T$ . As sums of convex functions,  $D_+$  and  $D_-$  defined at (12) are convex on  $\mathbb{R}^m$ . Then Theorem 3.5 brings the convexity of  $d_+$  and  $d_-$  on  $\mathcal{S}_n$ . Clearly  $D_+$  and  $D_-$  are also differentiable and

$$\nabla D_+ = [\min(0, x_1), \dots, \min(0, x_n)]^T \quad \nabla D_- = [\max(0, x_1), \dots, \max(0, x_n)]^T. \quad (13)$$

Theorem 3.6 yields that  $d_+$  and  $d_-$  are differentiable on  $\mathcal{S}_n$  and that a formulation of the gradients at  $C$  is

$$\nabla d_+(C) = \sum_{i=1}^n \lambda_i^- e_i e_i^T = C_- \quad \nabla d_-(C) = \sum_{i=1}^n \lambda_i^+ e_i e_i^T = C_+.$$

Finally the Lipschitz continuity of the gradients is a consequence of Theorem 3.3 which says that they are projections, thus Lipschitz continuous.  $\square$

There is actually *no* chance that  $\nabla d_+$  be also differentiable everywhere: just take the counterexample in one dimension  $d_1 : x \mapsto \frac{1}{2} \min(0, x)^2$ . Its derivative function is  $d_1'(x) = \min(0, x)$  which is clearly not differentiable at 0. We actually have a more precise result which somehow generalizes the previous example.

**Theorem 3.8** *The functions  $d_+$  and  $d_-$  are twice differentiable at  $A \in \mathcal{S}_n$  if and only if  $A$  is invertible.*

**Proof.** We write the proof for  $d_+$ , it is exactly the same for  $d_-$ . From (13), we see that the gradient of  $D_+$  is differentiable at  $x = (x_i)_i$  if and only if each  $x_i$  is nonzero. Theorem 3.7 enables us to say that  $d_+$  is twice differentiable if and only if 0 is not an eigenvalue of  $A$ , i.e. if and only if  $A$  is invertible.  $\square$

## 4 Lagrangian duality

We propose in this section a partial Lagrangian dualization of the general conic least-square problem. Our basic idea is to treat in two different ways the two different kinds of constraints:

on one hand affine constraints in  $\mathbb{R}^p$  and on the other hand semidefinite constraints. The technique is to dualize only affine constraints, forming a *partial Lagrangian*.

The adjoint operator of  $\mathcal{A}$ , denoted as usual by  $\mathcal{A}^*$ , will come into play. When  $\mathcal{A}$  has the form (2), it is easy to see that:

$$\begin{aligned} \mathcal{A}^* : \mathbb{R}^m &\longrightarrow \mathbb{R}^p \\ y &\longmapsto \sum_{i=1}^m y_i a_i \end{aligned}$$

All the present paper relies upon the following statement It motivates the developments of previous sections and will give birth to our computational methods.

**Theorem 4.1** *Consider the generic conic least-square problem on  $\mathbb{R}^p$*

$$(primal) \begin{cases} \min & \frac{1}{2} \|x - c\|_p^2 \\ x \in \mathcal{K}, & \mathcal{A}x = b, \end{cases}$$

which is our primal problem. Form the partial Lagrangian depending on two variables (the primal variable  $x$  which lies in  $\mathcal{K}$  and the dual variable  $y$  which lies in the constraint space  $\mathbb{R}^m$ )

$$L(x; y) := \frac{1}{2} \|c - x\|_p^2 - \langle y, \mathcal{A}x - b \rangle_m. \quad (14)$$

Define the corresponding dual function

$$\theta(y) := \min_{x \in \mathcal{K}} L(x; y), \quad (15)$$

which results in the dual problem on the constraint space  $\mathbb{R}^m$

$$(dual) \begin{cases} \sup & \theta(y) \\ y \in \mathbb{R}^m. \end{cases}$$

The dual function has the following expressions

$$\begin{aligned} \theta(y) &= -\frac{1}{2} \|p_{\mathcal{K}}(c + \mathcal{A}^*y)\|_p^2 + \frac{1}{2} \|c\|_p^2 + \langle y, b \rangle_m \\ &= -d_{\mathcal{K}^\circ}(c + \mathcal{A}^*y) + \frac{1}{2} \|c\|_p^2 + \langle y, b \rangle_m. \end{aligned} \quad (16)$$

**Proof.** Let us transform the partial Lagrangian to isolate the function  $d_{\mathcal{K}}$  of (6):

$$\begin{aligned} L(x; y) &= \frac{1}{2} \|c - x\|_p^2 - \langle \mathcal{A}^*y, x \rangle_p + \langle y, b \rangle_m \\ &= \frac{1}{2} \|(c + \mathcal{A}^*y) - x\|_p^2 - \left(\frac{1}{2} \|\mathcal{A}^*y\|_p^2 + \langle C, \mathcal{A}^*y \rangle_p\right) + \langle b, y \rangle_m \\ &= \frac{1}{2} \|(c + \mathcal{A}^*y) - x\|_p^2 - \left(\frac{1}{2} \|\mathcal{A}^*y + c\|_p^2 - \frac{1}{2} \|c\|_p^2\right) + \langle b, y \rangle_m. \end{aligned}$$

Now get an expression of  $\theta$ . For any fixed  $y \in \mathbb{R}^m$ :

$$\begin{aligned} \theta(y) &:= \min_{x \in \mathcal{K}} L(x; y) \\ &= d_{\mathcal{K}}(c + \mathcal{A}^*y) - \frac{1}{2} \|\mathcal{A}^*y + c\|_p^2 + \frac{1}{2} \|c\|_p^2 + \langle y, b \rangle_m. \end{aligned}$$

Simplify with (7) and (5):

$$\begin{aligned}\theta(y) &= \frac{1}{2} \|p_{\mathcal{K}^\circ}(c + \mathcal{A}^*y)\|_p^2 - \frac{1}{2} \|\mathcal{A}^*y + c\|_p^2 + \frac{1}{2} \|c\|_p^2 + \langle y, b \rangle_m \\ &= -\frac{1}{2} \|p_{\mathcal{K}}(c + \mathcal{A}^*y)\|_p^2 + \frac{1}{2} \|c\|_p^2 + \langle y, b \rangle_m \\ &= -d_{\mathcal{K}^\circ}(c + \mathcal{A}^*y) + \frac{1}{2} \|c\|_p^2 + \langle y, b \rangle_m.\end{aligned}$$

We therefore obtain the expected formulations of  $\theta$ .  $\square$

Notice that we know the unique point in  $\mathcal{K}$  which achieves the minimum in (15) for  $y \in \mathbb{R}^p$ . In the remainder of the paper, we denote it by  $x(y)$ :

$$x(y) := \operatorname{argmin}_{x \in \mathcal{K}} L(x; y) = p_{\mathcal{K}}(c + \mathcal{A}^*y). \quad (17)$$

In others words there holds

$$\theta(y) = L(x(y); y). \quad (18)$$

Yet  $\theta$  has a strong structure which could be exploited for algorithmic perspectives. In fact, the dual function inherits the properties of  $d_{\mathcal{K}^\circ}$  studied in Section 2. Let us sum them up.

**Theorem 4.2** *The function  $\theta$  of (16) satisfies the properties below:*

- (i)  $\theta$  is concave,
- (ii) and differentiable,
- (iii)  $\nabla\theta$  is Lipschitz continuous.

Besides its gradient  $\nabla\theta(y) = -\mathcal{A}\{p_{\mathcal{K}}(c + \mathcal{A}^*y)\} + b$  is differentiable almost everywhere.

**Proof.** The dual function, as a minimum of concave functions of  $y$ , is concave by construction. Besides, with equation (16), according to results on  $d_{\mathcal{K}^\circ}$  (Theorem 2.2 and Theorem 2.3 for  $\mathcal{K}^\circ$ ),  $\theta$  is differentiable, its gradient is

$$\begin{aligned}\nabla\theta(y) &= -\mathcal{A}\{\nabla d_{\mathcal{K}^\circ}(c + \mathcal{A}^*y)\} + b \\ &= -\mathcal{A}\{p_{\mathcal{K}}(c + \mathcal{A}^*y)\} + b,\end{aligned} \quad (19)$$

and is twice differentiable almost everywhere on  $\mathbb{R}^m$ .  $\square$

The dual problem is thus a *convex-differentiable* optimization problem, which can be simply formulated as

$$\begin{cases} \inf \frac{1}{2} \|p_{\mathcal{K}}(c + \mathcal{A}^*y)\|_p^2 - \langle b, y \rangle_m \\ y \in \mathbb{R}^m. \end{cases} \quad (20)$$

We therefore know some techniques to tackle it in practice (see next section). To end up this section, we illustrate the last two results with the semidefinite case.

**Example (Semidefinite least-square).** The key point for this class is that we have an easy-to-compute formulation of the projection onto  $\mathcal{S}_n^+$  (Theorem 3.3). From (20), solving the corresponding dual problem turns out to solving the following convex-differentiable program

$$\begin{cases} \min \|(C + \mathcal{A}^*y)_+\|^2 - b^T y \\ y \in \mathbb{R}^m, \end{cases}$$

where  $\mathcal{A}^*(y) = \sum_{i=1}^n y_i A_i$ . The dual objective function is concave and differentiable, its gradient  $-\mathcal{A}\{(C + \mathcal{A}^*y)_+\} + b$  is Lipschitz continuous.  $\square$



## 5 Conic Dual Algorithm

We want to solve the primal problem, i.e. to find  $x^*$  which achieves the minimum of the distance from  $c \in \mathbb{R}^m$  and satisfies constraints. The structure of the primal is not easy to exploit directly. On the other hand, its dual problem is more strongly structured (Theorem 4.2). We show that one way to avoid difficulties is to exploit the dual problem.

We assume in this section that there is a solution  $y^*$  of the dual problem: the dual function is bounded from above and its supremum is actually a maximum, achieved at  $y^*$ . Actually this assumption is not really a restriction since in practice it is always satisfied. We are therefore in the following primal-dual situation

$$(primal) \left\{ \begin{array}{l} \min \\ x \in \mathcal{K}, \mathcal{A}x = b \end{array} \right. \frac{1}{2} \|x - c\|^2 \quad (dual) \left\{ \begin{array}{l} \max \\ y \in \mathbb{R}^m. \end{array} \right. \theta(y) \quad (21)$$

with  $\theta$  expressed by (16). The only difference from the expression of Theorem 4.1 is the maximum instead of the supremum in the dual problem.

### 5.1 From dual to primal solution

In this subsection we suppose that we are able to get efficiently  $y^*$ . We show that in this case the primal problem is solved indeed. We start by mentioning that each value of the dual function gives a lower bound on the primal objective function.

**Lemma 5.1 (weak duality)** *For all dual-feasible points (i.e.  $y \in \mathbb{R}^m$ ) and for all primal-feasible points (i.e.  $x \in \mathcal{K}$  such that  $\mathcal{A}x = b$ ), there holds*

$$\theta(y) \leq \frac{1}{2} \|c - x\|^2.$$

**Proof.** By definition (15), we have  $\theta(y) \leq L(x; y)$ . Since  $x$  is primal-feasible, we have from (14) the desired inequality. This is actually a general result, called *the weak duality theorem* (see [HUL93, Chap. XII]).  $\square$

**Theorem 5.2** *Assume the existence of a dual solution  $y^*$ . Then the solution  $x^*$  of the primal problem is given by*

$$x^* = p_{\mathcal{K}}(c + \mathcal{A}^*y^*). \quad (22)$$

**Proof.** From Theorem 4.2,  $\theta$  is concave and differentiable, then at  $y^*$  which achieves its maximum, its gradient is zero. By equations (17) and (19), this results in  $\mathcal{A}x(y^*) = b$ , i.e.  $x(y^*)$  is primal-feasible. Then we have by (18):

$$\theta(y^*) = L(x(y^*); y^*) = \frac{1}{2} \|c - x(y^*)\|^2 \quad (23)$$

By Lemma 5.1,  $\theta(y)$  is a lower bound of the objective function of the primal. Equation (23) means that this lower bound is reached at the primal-feasible  $x(y^*)$ . Thus that point is the minimum and

$$x^* = x(y^*) = p_{\mathcal{K}}(c + \mathcal{A}^*y^*).$$

This thus ends the proof.  $\square$

This theorem says that there is *no duality gap* between the primal and the dual. This is expressed by equation (23): the value of the primal function at its minimum and the one of the dual at its maximum are the same.

It is worth mentioning that there is a particular case yielding both existence of  $y^*$  and absence of a duality gap. This is when the primal *Slater condition* is satisfied (see [HUL93, chap XII]). This condition is that feasibility of the primal constraints is preserved despite perturbations of  $b$ . It corresponds to the existence of a point *strictly feasible* of the primal: there exists  $x$  satisfying  $\mathcal{A}x = b$  and lying in the interior of  $\mathcal{K}$ , assumed nonempty.

## 5.2 How to get a dual solution

Since  $\theta$  has some nice properties, a quasi-Newton algorithm provides us with an algorithm to solve the dual, i.e. to compute an approximation  $\tilde{y}$  of  $y^*$ . A classical efficient quasi-Newton algorithm is the one called BFGS with Wolfe line-search (see [BGLS97]). In the rest of the paper, when talking about a quasi-Newton algorithm we refer to BFGS.

**Algorithm 5.3 (Conic Dual Algorithm)** *Consider the pair of primal-dual problems of (21). Let a black-box perform the following task:*

- (i) compute  $\mathcal{A}^*y$  for given  $y \in \mathbb{R}^m$ ;
- (ii) compute  $\mathcal{A}x$  for given  $x \in \mathbb{R}^p$ ;
- (iii) compute  $p_{\mathcal{K}}(z)$  for given  $z \in \mathbb{R}^p$ .

*Use a quasi-Newton optimization code to minimize  $\theta$  on  $\mathbb{R}^m$ . With the help of the above black-box, this code generates a maximizing sequence  $(y_k)_k$  together with the corresponding:*

- (i)  $x_k = p_{\mathcal{K}}(c + \mathcal{A}^*y_k)$ ;
- (ii)  $\nabla\theta(y_k) = -\mathcal{A}x_k + b$ ;
- (iii)  $\theta(y_k) = -\frac{1}{2}\|x_k\|_p^2 + \langle y_k, b \rangle_m$ .

To implement the above algorithm the only thing we basically need is to compute  $p_{\mathcal{K}}$ . In other words, the key point to solve the general conic least-square problem (i.e. to compute the projection onto the intersection of  $\mathcal{K}$  and affine hyperplans) is to know how to solve the problem without affine constraints (i.e. to compute the projection onto  $\mathcal{K}$ ). This means that this algorithm is efficient if the difficulty is brought by the addition of affine constraints. For instance, this is the case for the two examples of Subsection 1.3. We have for these examples an easy-to-compute expression of the projections: the projection of  $C$  onto  $\mathcal{S}_n^+$  is  $C_+$  and the projection of  $C \oplus D$  onto  $\mathcal{S}_n^+ \times \mathcal{S}_n$  is  $C_+ \oplus D$ . It is then possible to run the algorithm on them.

Besides, the BFGS algorithm is known to be convergent when the objective function is convex and has a Lipschitz gradient (Theorem 3.9 of [BGLS97]); this is the case here from Theorem 4.2. Then the conic dual algorithm converges also. Here is a convenient convergence result:

**Theorem 5.4** *Let the Slater assumption hold. Then Algorithm 5.3 provides an approximation of  $x^*$ : for any  $\epsilon > 0$ , there is  $k$  such that  $\|x_k - x^*\| \leq \epsilon$ . Besides the computation cost is given by*

$$(\text{cost of } p_{\mathcal{K}} + \text{cost of } p_{\mathcal{A}^*} + \text{cost of } p_{\mathcal{A}} + m^2) \times k.$$

**Proof.** From the Slater assumption, the dual optimal set is bounded ([Roc70]). Then  $(y_k)$  is bounded. For some dual solution  $y^*$  and  $k$  large enough,  $\|y_k - y^*\| \leq \epsilon/\|\mathcal{A}^*\|$  (see [Pow75]). Now from Lipschitzian property of the projection we can write

$$\|x^* - x_k\| = \|p_{\mathcal{K}}(c + \mathcal{A}^*y^*) - p_{\mathcal{K}}(c + \mathcal{A}^*y_k)\| \leq \|\mathcal{A}^*\| \|y^* - y_k\| \leq \epsilon.$$

As for the computation effort, observe that quasi-Newton requires for each iteration  $m^2$  cost to compute  $\nabla\theta$  and  $\theta$ . We obtain the expected result.  $\square$

The superlinear (or even linear) convergence is *not* guaranteed because  $\theta$  is *not* twice differentiable on  $\mathbb{R}^m$ . According to the state of the art, the superlinear convergence is actually unlikely. It will thus be fruitful to design a superlinearly convergent version of this algorithm to proceed as less iterations as possible. To get deeper results it is necessary to refine the first order information and a way to do this is to use some local second order information. This is the subject of current studies, based on Theorem 3.8.

## 6 Illustration: semidefinite least-square

In this section, we will focus on the case where  $\mathcal{K}$  is the cone of positive semidefinite  $n \times n$  symmetric matrices. In what follows, this class of semidefinite least-squares problem will be denoted by *(sdl)*. We present, in Subsection 6.1, the natural first idea to solve *(sdl)*. We then compare it to the semidefinite dual algorithm. In Subsection 6.3, we enlarge the field of use of the dual algorithm. We end by mentioning one of the numerous perspectives of applications of this algorithm in finance.

### 6.1 A direct approach

A natural idea is to use known techniques to tackle *(sdl)* directly. The problem is indeed a quadratic-semidefinite program ([BTN98]) and then an  $\epsilon$ -solution can be found in polynomial-time using SDP optimization tools (primal-dual interior-points method, see [SVW00]). Let us explain how this can be done and then why this is not a good idea, for large problems.

The (nonlinear) objective function can actually be pushed into constraints with the classical trick consisting in introducing  $t \in \mathbb{R}$  subject to  $\|X - C\|^2 \leq t^2$ . A *(sdl)*-equivalent optimization program is

$$\begin{cases} \min & t \\ & \|X - C\| \leq t \\ & \mathcal{A}X = b \\ & X \succeq 0. \end{cases}$$

The problem can thus be seen as a quadratic-semidefinite program with a Lorentz cone (or quadratic cone) in  $\mathbb{R}^{\frac{n(n+1)}{2}+1}$ . The problem can then be solved with modern interior-points methods solvers, SDPPACK or SEDUMI for instance ([stu99]).

In interior points method, a Newton step per iteration must be computed in the space  $\mathcal{S}_n$ ; this costs  $(\frac{n(n+1)}{2})^3 \approx n^6$  (for more details about computation costs for these methods see [Nay00]).

It is generally thought that problems with matrices of order  $n \geq 100$  cannot be solved using interior point approach. Moreover, in practice, to use such softwares imposes to fit in a datastructure which makes the number of variables get bigger: the practical cost is still more expensive. For large matrices ( $n \geq 100$ ) not even one iteration can be achieved.

In our study, we solved instances of (*sdl*s) using SEDUMI for a problem with  $C \in \mathcal{S}_{60}$  and  $b \in \mathbb{R}^{10}$ ; see Subsection 6.2 for more experiments. SEDUMI is now over three years old and so more efficient interior-point codes exist. However, the gain in efficiency in using news codes is likely not to be significant enough for our purposes.

## 6.2 Numerical results

Consider now the semidefinite version of the conic dual algorithm (Algorithm 5.3). By Theorem 3.3, the computation of the projection onto  $\mathcal{S}_n^+$  is nothing more than the computation of the eigenvalues of the considered matrix. This could be done using symmetric QR factorization ([GvL83]) whose computation cost is of order  $n^3$ . From Theorem 5.4, we then know that the cost of one iteration is now  $n^3 + mn^2 + m^2$ . In the worst case  $m = n(n+1)/2$  (every entry of  $X$  is fixed), the cost is in  $n^4$ . Actually the number of affine constraints is rarely high in applications (typically  $m \leq 10$ ). We can consider that the cost is  $n^3$  per iteration. Our algorithm is then better in theory than the quadratic semidefinite algorithm proposed in the last subsection (for precised computation costs, see [Nay00]). This is also checked in practice (see Figure 1).

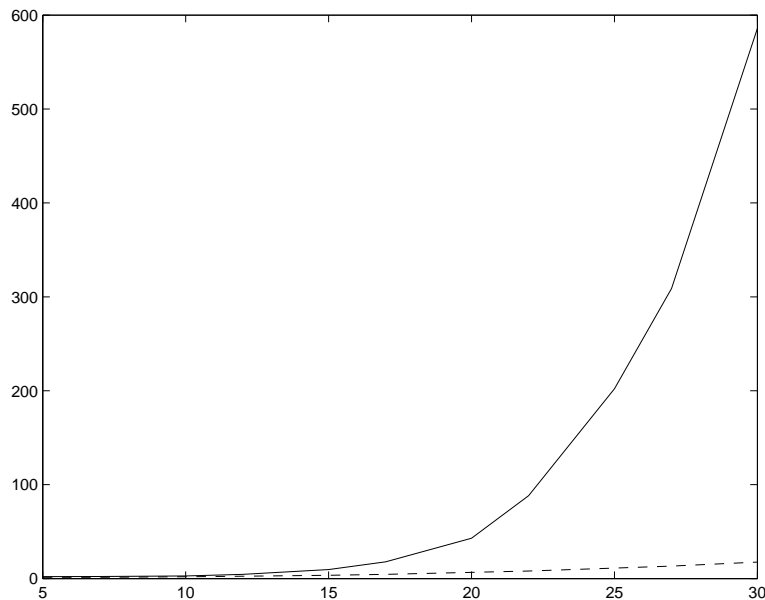


Figure 1: Comparison of computation effort between conic dual algorithm and SEDUMI. The full line stands for results of SEDUMI, the dashed line for the semidefinite dual algorithm.

We compute a prototype of semidefinite conic dual algorithm with the MATLAB-like numerical mathematics software SCILAB ([sci]). Figure 1 compares both approaches on (*sdl*s) with full matrices of size from 5 to 30 with 5 affines constraints. The full line represents results of SEDUMI, the dashed line, those of semidefinite dual algorithm. The vertical unit is one second of

cputime on a Sparc Ultra 1.

Furthermore, even though *linear convergence* is not guaranteed, it is observed in practice: the property

$$\frac{\theta(y_{k+1}) - \theta(y^*)}{\theta(y_k) - \theta(y^*)} < \text{cst} < 1,$$

occurred in all of our experiments, using different kinds of matrices (we changed size, sparsity pattern, eigenvalue distribution). The constant depends of course on the matrix in question.

We solved instances of *(sdl)* using the prototype coded in SCILAB with *sizes up to  $n=500$  and  $m=20$* . Beyond these values, troubles come from the projection on  $\mathcal{S}_n^+$ ; a suitable algorithm may be needed. It is thought that it is possible to reach higher size if an efficient eigenvalues decomposition is used.

### 6.3 Possible extensions

We mention that the general pattern can be enlarged. First, we show it is possible to treat affine constraints with nonsymmetric  $A_i$ . Second, we see that constraints of the kind  $X \succeq \alpha I_n$  (i.e.  $X$  is in an *affine cone*) are also accepted.

**Nonsymmetric  $A_i$ .** The symmetry assumption of  $A_i$  in *(sdl)* is actually useless: each  $A_i \in \mathbb{R}^{n^2}$  can be replaced with  $\tilde{A}_i \in \mathcal{S}_n$ . Let  $A \in \mathbb{R}^{n^2}$ ,  $X \in \mathcal{S}_n^+$  and observe that  $\langle A, X \rangle = \langle A^T, X \rangle$ . This is quite easy to check:

$$\langle A, X \rangle = \text{tr}(A^T X) = \text{tr}(X^T A) = \text{tr}(X A) = \text{tr}(A X) = \langle A^T, X \rangle.$$

So we can replace conditions  $\langle A_i, X \rangle = b_i$  with  $A_i \in \mathbb{R}^{n^2}$  by the equivalent constraint  $\langle \tilde{A}_i, X \rangle = b_i$  where  $\tilde{A}_i = \frac{A_i + A_i^T}{2}$ . Accordingly we will restrict ourselves in theory to symmetric  $A_i$ 's.

**Affine cone.** Needless to say, the following variante of *(sdl)*

$$\begin{cases} \min & \frac{1}{2} \|X - C\|^2 \\ & X \succeq M, \quad \mathcal{A}X = b, \end{cases}$$

enters the same framework via the change of variable  $Z = X - M$ . Note that this problem without affine constraint has  $(C - M)_+ + M$  for explicit solution. The case  $M = \alpha I_n$  deserves an additional comment: having the spectral decomposition  $C = W(\text{Diag}\lambda(C))W^T$ , let  $k$  be an integer satisfying  $\lambda_k(C) \geq \alpha \geq \lambda_{k+1}(C)$ , then the projection onto the affine cone  $\{X \in \mathcal{S}_n : X \succeq \alpha I_n\}$  is

$$W \begin{bmatrix} \text{Diag}[\lambda_1(C), \dots, \lambda_k(C)]^T & \\ & \alpha I_{n-k} \end{bmatrix} W^T.$$

This result is very natural. The projection onto  $\mathcal{S}_n^+$  consists in truncating negative eigenvalues to zero. To project onto  $\{X \in \mathcal{S}_n : X \succeq \alpha I_n\}$ , we truncate eigenvalues below  $\alpha$ .

### 6.4 An application: calibration of volatilities

In this subsection we outline the problem of finding a “good” approximation  $\bar{Q}$  of the theoretical covariance matrix  $Q$  between  $n$  assets; it turns out to be a semidefinite least-square problem.

This matrix  $\bar{Q}$  is used to have a robust estimation of the “ex-ante” risk of any possible portfolio among these  $n$  assets.

Let  $\tilde{Q}$  be a first estimate of the true covariance matrix  $Q$ . For instance  $\tilde{Q}$  can be the empirical estimate after  $k$  days. The point is that  $\tilde{Q}$  cannot be successfully used in optimization programs because it is often ill-conditioned (see [GPB99] and [LCPB00]). The reason is that the large [resp. small] eigenvalues of  $Q$  are over[resp. under]estimated (one underestimates the risk associated to some investments and one overestimates the risk of others). An extreme case is when the number of observations  $k$  is too small,  $\tilde{Q}$  is rank-deficient (some investments are considered with no risk) and its conditioning is  $+\infty$ . We may then impose  $X \succeq \alpha I_n$  for some selected  $\alpha > 0$ . Eventually we are lead to solve the so called volatility calibration problem, which is a shifted (*sdl*s):

$$\begin{cases} \min & \frac{1}{2} \|X - \tilde{Q}\|^2 \\ & X \succeq \alpha I_n \\ & \langle I_n, X \rangle = \text{tr}(\tilde{Q}) \\ & \langle A_i, X \rangle = \sigma_i^2, \end{cases}$$

where  $\sigma_i^2$  represent “ex-post” volatilities of well-chosen portfolios (see [A. 00] for further information about calibration).

## Acknowledgement

My first grateful thanks go to my supervisor Claude Lemaréchal for his advice and his everyday help. I am also really indebted to Francois Oustry for the fruitful discussions we had together.

## References

- [A. 00] A. d’Aspremont. A full-rank market correlation calibration procedure for the market model of interest rates. BNP-Paribas, 2000.
- [BGLS97] J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, and C. Sagastizábal. *Optimisation Numérique: aspects théoriques et pratiques*. Springer Verlag, Paris, 1997.
- [BTN98] A. Ben-Tal and A. Nemirovski. *Convex optimization in engineering*, 1998.
- [FO01] S. Filipowski and F. Oustry. Sdp applied to preconditionning of symmetric matrices. Technical report, Boeing, 2001.
- [GPB99] S. Galluccio, M. Potters, and J.-P. Bouchaux. Rational decisions, random matrices and spin glasses. *Physical Review Letters*, 83(7), 1999.
- [GvL83] G.H. Golub and C.F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, 1983.
- [HS95] U. Helmke and M. Shayman. Critical points of matrix least squares distance functions. *Linear Algebra and its Applications*, 215:1–19, 1995.
- [HUL93] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer Verlag, Heidelberg, 1993.

- [LCPB00] L. Laloux, P. Cizeau, M. Potters, and J.-P. Bouchaux. Random matrix theory and financial correlation. 2000. To appear in the proceedings of the 1999 Dublin conference published by World Scientific I.J.T.A.F.
- [Lew96a] A.S. Lewis. Convex analysis on the hermitian matrices. *SIAM Journal on Optimization*, 6(1):164–177, 1996.
- [Lew96b] A.S. Lewis. Derivatives of spectral functions. *Mathematics of Operations Research*, 21:576–588, 1996.
- [LO99] C. Lemaréchal and F. Oustry. Semi-definite relaxations and lagrangian duality with application to combinatorial optimization. Rapport de Recherche 3710, INRIA, 1999.
- [LS] A. Lewis and H. Sendov. Twice differentiable spectral functions. to appear.
- [Nay00] M. Nayakkankuppam. Optimization over symmetric cones. 2000. Thesis.
- [Pow75] M.J.D. Powell. Convergence properties of a class of minimization algorithms. In O.L. Mangasarian, R.R. Meyer, and S.M. Robinson, editors, *Nonlinear Programming 2*, pages 1–27. Academic Press, New York, 1975.
- [Roc70] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [RW98] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis*. Springer Verlag, Heidelberg, 1998.
- [sci] Scilab’s home page. <http://www-rocq.inria.fr/scilab>.
- [stu99] Using sedum 1.0x, a matlab toolbox for optimization over symmetric cones, 1999. <http://www.unimaas.nl/sturm>.
- [SVW00] R. Saigal, L. Vandenberghe, and H. Wolkowicz. *Handbook of Semidefinite-Programming*. Kluwer, 2000.



---

Unité de recherche INRIA Rhône-Alpes

655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique

615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)

Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

---

Éditeur

INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)

<http://www.inria.fr>

ISSN 0249-6399