MPE — Volume 2, pp. 499-522
Reprints available directly from the publisher
Photocopying permitted by license only

LINEAR PROGRAMMING WITH POSITIVE SEMI-DEFINITE MATRICES*

J. B. LASSERRE†

LAAS-CNRS, 7 Avenue du Colonel Roche, 31 077 Toulouse Cédex, France.

(Received 31 July 1995)

We consider the general linear programming problem over the cone of positive semi-definite matrices. We first provide a simple sufficient condition for existence of optimal solutions and absence of a duality gap without requiring existence of a strictly feasible solution. We then simply characterize the analogues of the standard concepts of linear programming, i.e., extreme points, basis, reduced cost, degeneracy, pivoting step as well as a Simplex-like algorithm.

KEYWORDS: Positive semi-definite matrices, linear programming, simplex algorithm

1. INTRODUCTION

Optimization with positive semi-definite matrices has been investigated by several researchers in the last twenty years (e.g. [7]) with many applications, particularly in Control Theory (see [2]). It seems generally admitted that the *interior points* methods originated by Nesterov and Nemirovskii (see [12]) are particularly efficient compared to *gradient*-type or *cutting-planes*-type methods like [8]. However, surprisingly enough, we are not aware of any Simplex-like algorithm, and to our knowledge, the analogues of the standard notions of *basic* solution, *basis*, *reduced-cost*, and *degeneracy* in standard Linear Programming (LP) have not been investigated.

Our contribution in this paper is to provide:

- a simple sufficient condition for absence of a duality gap between the primal and dual problems. It is not of the *Slater*-type for it does not require knowledge of a strictly feasible point as in [2], [12], which can be difficult in many problems.
- As in standard LP, a simple characterization of the *extreme points* (or *basic* solutions), as well as the analogues of *reduced cost*, *basis* of a basic solution, and *degenerate* basic solutions. Those concepts bring some insight into the structure of the extreme points. In particular, the number of zero-eigenvalues of a feasible solution is crucial in characterizing a basic solution as well as degeneracy, pretty much like the number of zeros in a feasible solution in standard LP. The reduced cost also provides a means to check optimality as in standard LP.

*A shortened incomplete version has been presented at the 34th IEEE CDC conference in New Orleans.

†Email: lasserre@laas.fr

- a *sufficient* condition as well as a *necessary and sufficient* condition of optimality of a primal feasible solution. In many cases, those conditions are also easy to check via a simple traditional linear programming problem.
- a Simplex-like algorithm

so that all the basic concepts in standard LP have their analogues in positive semi-definite programming.

2. NOTATION AND PRELIMINARIES

We introduce the following notation:

- X: the vector space ($\equiv R^{N(N+1)/2}$) of (N, N) symmetric real-matrices.
- Y: ($\equiv X$) its dual.
- $\langle .,. \rangle$: the usual duality bracket

$$\langle x, y \rangle = \mathrm{trace}(x \cdot y)\ \forall (x, y) \in X \times Y,$$

where $x \cdot y$ stands for the usual matrix multiplication (x and y being now considered as matrices with N^2 elements).

- Z: the vector space ($\equiv R^{M(M+1)/2}$) of (M, M) symmetric real-matrices.
- I_X, I_Z: the identity matrices in X and Z respectively.
- P: the (closed) positive cone in X of semi-definite positive matrices.
- "$\succcurlyeq$": the partial ordering in X (same notation in Z) where $x \succcurlyeq y \Leftrightarrow x - y \in P$.
- $\{p_k(x)\}_1^r$: the r normalized zero-eigenvectors of $x \in X$.
- $P(x)$: the (N, r)-matrix whose columns are the r vectors $\{p_k(x)\}$. $P(x) \equiv 0$ if $x \in X$ is not singular.
- $R(x)$: the eigenprojection matrix associated to the zero-eigenvalue of $x \in X$. $R(x) := \sum_{i=1}^{r} p_k(x) p_k^T(x)$. $R(x) \equiv 0$ if $x \in X$ is not singular.
- $Q(x)$: the matrix whose columns are the eigenvectors associated with the positive eigenvalues of $x \in X$.
- $\|x\|^2$: the norm of $x \in X$, derived from the scalar product $\langle x, x \rangle$. Equipped with this norm, $(X, \|.\|)$ is a Hilbert space.
- n, m: $n := N(N+1)/2$; $m := M(M+1)/2$.

Thus, in the sequel, when $x \in X$ appears with a dot "." in a multiplication it should be understood as an (N, N) matrix and not a $N(N+1)/2$ vector (the same applies for $v \in Z$).

For some given $c \in X$ and $b \in Z$, consider the following *linear programming* problem

$$IP \begin{cases} \sup \langle c, x \rangle \\ Bx \preccurlyeq b \\ x \succcurlyeq 0 \end{cases} \tag{1}$$

and its dual

$$IP^* \begin{cases} \inf\langle b, u\rangle \\ B^*u \quad \succcurlyeq c \\ u \quad \succcurlyeq 0 \end{cases} \tag{2}$$

where B^*: $Z \rightarrow X$ is the adjoint linear mapping defined as

$$\langle B^*y, x\rangle : = \langle y, Bx\rangle \; x \in X, y \in Z.$$

A Lyapunov inequality $A^T{\cdot}x + x{\cdot}A \preccurlyeq 0$, or a Riccati-type matrix inequality $A^T.x + x.A + x.CR^{-1}\ C^T.x + Q \preccurlyeq 0$ that can be represented as

$$\begin{bmatrix} A^T.x + x.A & | & x.C \\ \text{---} & & \text{---} \\ C^T.x & | & 0 \end{bmatrix} \preccurlyeq \begin{bmatrix} -Q & | & 0 \\ \text{---} & | & \text{---} \\ 0 & | & R \end{bmatrix}$$

are examples of such operators B. For instance, if Bx stands for $A^T.x + x.A$ then $B^*\ y$ stands for $A.y + y.A^T$. In our context, since the spaces X, Z are R^n and R^m we consider the usual topology.

IP is said to be *consistent* if it has a feasible solution. It is said to have *finite value* if its optimal value is finite.

In case where b is diagonal and Bx is diagonal when x is constrained to be diagonal, i.e., the constraint $x \succcurlyeq 0$ becomes $x \geq 0$, then *IP* (and therefore IP^*) reduces to a standard LP problem. We will then show that the notions derived in the sequel reduce to the usual ones in standard LP.

3. LINEAR PROGRAMMING

In contrast to standard LP, several issues such as solvability, absence of duality gap and strong duality are not always trivial and must be investigated. Indeed, the cone of positive semi-definite matrices is not *polyhedral*, and thus the celebrated Farkas Lemma is true only under some closedness assumption (automatically satisfied when the cone is polyhedral) [4], [3], [6] (see [11] for a new Farkas Lemma without this closedness assumption). In this section, we provide a simple sufficient condition for solvability of both *IP* and IP^* as well as absence of a duality gap.

We also provide a simple characterization of basic solutions and optimal solutions. Moreover, we simply derive the analogues of the *reduced-cost* as well as the *basis* of an extreme point (or basic solution) so that all the basic concepts of standard LP have their analogues in positive semi-definite programming.

3.1. Existence of Optimal Solutions

Conditions for existence of optimal solutions as well as absence of a duality gap between *IP* and *IP** have been given in [2], [12]. They are of the *Slater*-type, i.e., they require knowledge of a strictly feasible point x in *IP* or y in *IP**. However, note that the interior-point conditions in [2] are a special case of Theorem 3.13 in [1]. In some applications, checking this strictly feasibility condition may be as difficult as solving the original problem.

We exhibit a simple sufficient condition, namely $c \succ 0$ *and IP consistent with finite value* (not necessarily *strictly* feasible) that guarantees solvability of *IP* as well as absence of a duality gap. This condition is easier to check in practice, and arises in many examples from Control Theory.

THEOREM 3.1. *Assume that* $c \succ 0$ *and IP is consistent with finite value. Then, IP is solvable and there is no duality gap.*

Proof. By using notation from [1] we define the set

$$D: = \{(Bx + z, \langle c, x\rangle)|x, z \succcurlyeq 0\}$$

Assume that for a sequence $\{x_n, z_n\}_n$, the sequence $\{(Bx_n + z_n, \langle c, x_n\rangle)\}$ converges to some (d, a), i.e.,

$$Bx_n + z_n \to d, \text{ and } \langle c, x_n\rangle \to a$$

Since $c \succ 0$, $c \succcurlyeq \alpha I_X$ for some positive scalar α so that for $x \succcurlyeq 0$,

$$\langle c, x\rangle \geq \alpha\langle I_X, x\rangle \geq \alpha\|x\|$$

This because $\langle I_X, x\rangle = \text{trace}(x)$ and

$$\sum_{i=1}^{n} \lambda_i \geq \sqrt{\sum_{i=1}^{n} \lambda_i^2}.$$

Therefore,

$$\langle c, x_n\rangle \to a \Rightarrow \alpha\|x_n\| \leq 2a$$

for n large enough. Thus, for n large enough the sequence $\{x_n\}$ is in a compact set. Therefore $Bx_n \to Bx$ for some subsequence and thus, $z_n \to z: = b - Bx$ for that same subsequence. Since the positive cones in X and Z are closed, $x, z \succcurlyeq 0$. Moreover, $\langle c, x_n\rangle \to \langle c, x\rangle = a$ for the same subsequence. Hence, D is closed. Now, since *IP* is consistent with finite value, from Theorem 3.22 in [1], *IP* is solvable and there is no duality gap. □

Remark 3.1: A sufficient condition for *IP* to have a finite value is e.g. to assume that *IP** is consistent, i.e. has a (not necessarily strict) feasible solution. One may note that "$c > 0$ *and IP* consistent*" is also a sufficient condition for absence of a duality gap since it implies that *IP** is strictly feasible (by a careful use of two theorems of the alternative). However, the optimal value may be $-\infty$ if *IP* is not feasible. Thus, $c > 0$ and feasibility (not necessarily strict) of *both* programs ensures absence of a duality gap and solvability of *IP*.

3.2. Basic Solutions

What is a basic feasible solution in *IP*? Let us transform the inequality program *IP* into an equality program *EP* by adding a slack matrix variable $z \in Z$, $z \geqslant 0$ so that $Bx \geqslant b \Leftrightarrow Bx + z = b \Leftrightarrow T(x, z) = b$. The criterion becomes $\langle c, x\rangle + \langle 0, z\rangle$ and *IP* and *IP** are now the linear programs

$$EP \begin{cases} \sup\langle c, x\rangle \\ Bx + z = b \\ x, z \geqslant 0 \end{cases} \quad \text{and } EP^* \begin{cases} \inf\langle b, u\rangle \\ B^*u - w = c \\ u, w \geqslant 0 \end{cases}$$

Remark 3.2: Remember that B is an operator, not necessarily in matrix notation. For instance if Bx stands for $A^T.x + x.A$ (where x is understood now as a square symmetric (N, N) matrix), then in matrix-vector notation, $Bx + z: = Sx + z$ where S is some matrix and x and z are vectors in X and Z. Similarly, given some matrix $R(x)$, the constraint $u.R(x) = 0$ can be written $R'(x)u$ in matrix-vector notation, for some appropriate matrix $R'(x)$.

For any feasible solution (x, z) to *EP* let

$$D(x, z): = \{(u, v) \in X \times Z | \exists\lambda > 0, (x, z) \pm \lambda(u, v) \geqslant 0,\} \tag{3}$$

and let $N(T)$ be the null space of T, i.e.,

$$N(T): = \{(u, v) \in X \times Z | Bu + v = 0\}. \tag{4}$$

Then, by definition, (x, z) is a basic feasible solution if and only if

$$D(x, z) \cap N(T) = \{0\}. \tag{5}$$

PROPOSITION 3.2. *If EP is solvable then there exists an optimal basic solution.*
This is because $\forall(x, z) \in X \times Z$, $D(x, z) \cap N(T)$ is always of finite dimension (see e.g. [1]).

3.3. Characterization of a Basic Solution

It is clear that if x and z have no zero eigenvalue then (x, z) is not an extremal point since it is always possible to find a vector (u, v) such that

$$Bu + v = 0; (x, z) \pm \lambda(u, v) \succcurlyeq 0$$

for λ sufficiently small. Therefore, necessarily, if (x, z) is a basic solution, then x and/or z have at least one zero eigenvalue.

THEOREM 3.3. *x is a basic solution if and only if*

$$Bu + v = 0, u.R(x) = 0, v.R(z) = 0 \Rightarrow (u, v) = (0, 0). \tag{6}$$

Proof. From (5), it suffices to prove that

$$D(x, z) := \{(u, v) \in X \times Z \mid u.R(x) = 0, v.R(z) = 0.\}$$

Equation to be placed

Using the Jordan decomposition,

$$x = [Q(x)|P(x)] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} [Q(x)|P(x)]^T$$

and u is partitioned accordingly, i.e.,

$$u = [Q(x)|P(x)] \begin{bmatrix} u_{11} & u_{12}^T \\ u_{12} & u_{22} \end{bmatrix} [Q(x)|P(x)]^T.$$

Then, $x \pm \lambda u \succcurlyeq 0$ for some $\lambda > 0$ if and only if $u_{12} = 0$ and $u_{22} = 0$. But then, $u.P(x) = 0$, i.e. all the r zero-eigenvectors of x are also zero-eigenvectors of u which in turn implies $u.R(x) = 0$ (and similarly $v.R(z) = 0$).

Conversely, assume that $u.R(x) = 0$ and $v.R(z) = 0$. As $R(x) = \sum_{i=1}^{r} p_i(x)p_i(x)^T$, then

$$u.R(x) = 0 \Rightarrow u.R(x).p_i(x) = 0, i = 1, \cdots r \Rightarrow [\sum_{j=1}^{r} (u.p_j(x))p_j^T(x)]p_i(x) = u.p_i(x)$$

$$= 0 \quad i = 1, \cdots r$$

Thus, the $p_i(x)$ are still zero-eigenvectors of $x + \rho u$ for any scalar ρ and, by continuity, the positive eigenvalues of $x + \rho u$ are still positive provided ρ is small enough, which implies $x \pm \rho u \succcurlyeq 0$ for $\rho > 0$ and small enough. The proof is similar for $z + \rho v$.□

Note that this condition is easily verified via checking a rank condition.

Analogy with standard LP. To check the meaning of the characterization (6) in standard LP, assume that in *EP*, Bx and b are diagonal matrices when x and z are constrained to be diagonal matrices so that the constraints $x, z \succcurlyeq 0$ reduce to $x, z \geq 0$. Let x', z' and b' be now vectors in R^N, R^M and R^M that identify the diagonal elements of x, z and b. The constraint $Bx + z = b$ reduce to $Hx' + z' = b'$ for some $(M, M + N)$-matrix H (see Remark 3.2). Also, the scalar product $\langle c, x\rangle$ can be written $c'^T x'$ for some suitable vector c'. Therefore, the linear program $\{\sup\langle c, x\rangle | \, Bx + z = b | \, x, z \geqslant 0\}$ reads now

$$\{\max c'^T x' | \, Hx' + z' = b'; x', z' \geq 0\}. \tag{7}$$

x being diagonal, its zero-eigenvalues correspond to its zero diagonal elements. Therefore, its eigenprojection matrix $R(x)$ is diagonal. All the entries are zero except the diagonal elements (equal to 1) corresponding to the zero diagonal elements of x. The same is true for z. For instance,

$$x = \begin{bmatrix} x_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_4 \end{bmatrix}, \text{ with } x_1 x_4 > 0 \Rightarrow R(x) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence, the constraint $u.R(x) = 0$ amounts to set some columns of u at zero, and similarly for $v.R(z)$. In the above example the second and third columns of u must be zero. In addition, to keep the constraint $x, z \geq 0$ always valid, the feasible directions (u, v) must also be diagonal matrices, so that finally, $u.R(x) = 0$ and $v.R(z) = 0$ reduce to $x_i > 0 \Rightarrow u_i = 0$ and $z_i > 0 \Rightarrow v_i = 0$. Therefore, the characterization of a basic solution (x, z) in (6) reduces to

$$Hu' + v' = 0; \; u_i^{'} = 0 \textit{ if } x_i^{'} = 0, \; v_i^{'} = 0 \textit{ if } z_i^{'} = 0 \Rightarrow (u', v') = (0, 0) \tag{8}$$

(where u' and v' are vectors of same dimension as x' and z').

Since $u_i^{'} = 0$ if $x_i^{'} = 0$ and $v_i^{'} = 0$ if $z_i^{'} = 0$ describe $D(x', z')$, i.e., the set of directions (u', v') that preserve nonnegativity of $(x', z') \pm \lambda(u', v')$ the above relation simply means that (x', z') is an extreme point.

3.4. Reduced-Cost and Basis of a Basic Solution

Let (x, z) be a basic solution of *EP* and let $S(x, z) := D(x, z) + N(T)$. In matrix-vector notation, (6) is equivalent to

$$M(x, z) \begin{bmatrix} u \\ v \end{bmatrix} = 0$$

with the matrix $M(x, z)$ being the matrix of the system

$$Su + v = 0: R'(x)u = 0; R'(z)v = 0.$$

(see Remark 3.2).

PROPOSITION 3.4. *At a basic solution* (x, z), *for any* $(y, w) \in S(x, z)$, *the system*

$$\begin{bmatrix} Bu + v & = By + w \\ u.R(x) & = 0 \\ v.R(z) & = 0 \end{bmatrix}$$

or equivalently, in matrix notation,

$$M(x, z)\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} T(y, w) \\ 0 \\ 0 \end{bmatrix}$$

has a unique solution (u, v). $M(x, z)$ *is called the matrix of the basis at* (x, z).

Let T_{xz} *be the linear operator on* $S(x, z)$ *that maps* (y, w) *to the unique solution* (u, v), *i.e.* $(u, v) = T_{xz}(y, w)$. *The vector* $\pi \in X \times Z$, $\pi := (I - T_{xz})^* (c, 0)$ *is the reduced cost at* (x, z), *i.e.,* (x, z) *is optimal if and only if* $\langle \pi, (y, w)\rangle \leq 0$ *for all* $(y, w) \geqslant 0$ *in* $S(x, z)$, *i.e.,* $-\pi$ *is a nonnegative linear functional on* $S(x, z)$.

Proof. We proceed as in [1]. Let $(y, w) \in S(x, z)$, i.e., $(y, w) := (y_1, w_1) + (y_2, w_2)$, $(y_1, w_1) \in D(x, z)$ and $(y_2, w_2) \in N(T)$. Then, since (x, z) is basic, from (2), we conclude that (y_1, w_1) is the unique solution in $D(x, z)$ to $Bu + v = By + w$. Thus, (x, z) is the unique solution of $M(x, z)(u, v) = [b, 0, 0]^T$ which singles out $M(x, z)$ as the matrix of the *basis* at (x, z). Let us see how the change in cost can be expressed when moving to another solution.

Let $D'(x, z)$ be any direct complement of $D(x, z)$ in $S(x, z)$ and consider any other feasible solution $(x, z) + (x', z')$.

$(x', z') \in N(T)$, so that $(x', z') \in S(x, z)$ and can be written $(x_1, z_1) + (x_2, z_2)$ with $(x_1, z_1) \in D(x, z)$ and $(x_2, y_2) \in D'(x, z)$. Moreover, $Bx_1 + z_1 = -Bx_2 - z_2$. Thus, by definition of $D(x, z)$, (x_1, z_1) is the unique solution in $D(x, z)$ to

$$\begin{bmatrix} Bu + v & = -Bx_2 - z_2 \\ u.R(x) & = 0 \\ v.R(z) & = 0 \end{bmatrix}$$

Let T_{xz}: $S(x, z) \rightarrow D(x, z)$, with $T_{xz}(y, w) = (u, v)$ where $M(x, z)(u, v) = [T(y, w), 0, 0]$. The change in cost at this other arbitrary solution $(x, z) + (x', z')$ is

$$\begin{aligned} \langle (c, 0), (x', z')\rangle &= \langle (c, 0), (x_1, z_1) + (x_2, z_2)\rangle \\ &= \langle (c, 0), (I - T_{xz})(x_2, z_2)\rangle \\ &= \langle (I - T_{xz})^* (c, 0), (x_2, z_2)\rangle \\ &= \langle \pi, (x_2, z_2)\rangle \end{aligned}$$

Note that π is defined *only on* $S(x, z)$. Let us check that π has all the properties of a *reduced cost*.

Since $N(T)$ is also a direct complement of $D(x, z)$ in $S(x, z)$ if we note P_D the projection on $D(x, z)$ and P_N the corresponding projection on $N(T)$ we also have

$$(I - T_{xz})(y, w) = 0, (y, w) \in D(x, z) \Rightarrow \langle \pi, (y, w) \rangle = 0.$$

and thus $\langle \pi, P_D(y, w) \rangle = 0$, $(y, w) \in S(x, z)$. Moreover (just make $(x_1, z_1) = (0, 0)$),

$$\langle (c, 0), (y, w) \rangle = \langle \pi, (y, w) \rangle, (y, w) \in N(T) \Rightarrow \langle \pi, (y, w) \rangle = \langle \pi, P_N(y, w) \rangle, (y, w) \in S(x, z).$$

so that for all $(y, w) \in S(x, z)$,

$$\langle (c, 0), P_N(y, w) \rangle = \langle P_N^*(c, 0), (y, w) \rangle = \langle \pi, P_D(y, w) + P_N(y, w) \rangle = \langle \pi, (y, w) \rangle \Leftrightarrow \pi = P_N^*(c, 0)$$

For ease of notation, when $(x, z) \in X \times Z$, then $(x, z) \geqslant 0$ means $x \geqslant 0$, $z \geqslant 0$.

i) Now, suppose that there exists $(x', z') \in S(x, z)$, $(x', z') \geqslant 0$ such that $\langle \pi, (x', z') \rangle > 0$. By definition of $D(x, z)$ there is some $\lambda > 0$ such that $(x, z) - \lambda(x_1, z_1) \geqslant 0$, $(x_1, z_1) \in D(x, z)$ and hence,

$$(x, z) - \lambda(x_1, z_1) + \lambda(x', z') \geqslant 0$$

$$T((x, z) - \lambda(x_1, z_1) + \lambda(x', z')) = b$$

Note that $(x_1, z_1) \neq (x', z')$ since $\langle \pi, (x_1, z_1) \rangle = 0$. In addition

$$\begin{aligned} \langle (c, 0), (x, z) - \lambda(x_1, z_1) + \lambda(x', z') \rangle &= \langle (c, 0), (x, z) \rangle + \lambda \langle \pi, (x', z') - (x_1, z_1) \rangle \\ &= \langle (c, 0), (x, z) \rangle + \lambda \langle \pi, (x', z') \rangle \\ &> \langle (c, 0), (x, z) \rangle \end{aligned}$$

and (x, z) is not optimal.

ii) On the other hand, suppose that $\langle \pi, (y, w) \rangle \leq 0$ for all $(y, w) \in S(x, z)$. If (x', z') is some other feasible solution, then $((x, z) - (x', z')) = a \in N(T)$ so that $(x, z) - a = (x', z') \in S(x, z)$. Therefore,

$$\begin{aligned} \langle (c, 0), (x, z) \rangle - \langle (c, 0), (x', z') \rangle &= \langle (c, 0), (x, z) - (x', z') \rangle \\ &= \langle (c, 0), P_N((x, z) - (x', z')) \rangle \\ &= \langle P_N^*(c, 0), (x, z) - (x', z') \rangle \end{aligned}$$

$$= \langle \pi, (x, z) - (x', z') \rangle$$

$$\geq 0$$

where the last inequality follows from $\langle \pi, (x, z) \rangle = 0$, and $\langle \pi, (x', z') \rangle \leq 0$.□

Degeneracy. Degeneracy at a *basic* solution (x, z) occurs when $S(x, z)$ is only a subspace of $X \times Z$. This situation is the analogue of degeneracy in standard LP where at least one variable in the basis is equal to zero. This happens when the number of zero-eigenvalues of x and z is large enough. In this case, $u.R(x) = 0$ and $v.R(z) = 0$ impose more conditions than needed to ensure $N(T) \cap D(x, z) = \{0\}$, i.e., the matrix $M(x, z)$ is not a square invertible matrix but a rectangular matrix of full rank. In the Appendix, two examples of degenerate basic solutions are presented.

On the other hand, to be a basic solution, a minimum number of conditions $u.R(x) = 0$ and/or $v.R(z) = 0$ are needed to ensure uniqueness of the solution to $Bu + v = a$. Therefore, a basic solution necessarily has a certain number of zero-eigenvalues. Still in the example in the Appendix, $Bu + v = 0$ is a (10, 16) system with rank 10. u and v are (3.3) and (4.4)-matrices respectively. At least 6 additional constraints on u and v are needed for (x, z) to be basic. At a basic solution with $x > 0$, if z has a simple zero-eigenvalue, $v.R(z) = 0$ imposes only 4 constraints so that necessarily z has at least a double zero-eigenvalue for which $v.R(z) = 0$ provides more than 6 constraints which in turn implies that this basic solution must be degenerate.

Thus, Theorem 3.3 provides some insight into the structure of a basic solution. The number of zero-eigenvalues of x and z is crucial in the characterization of basic solution as well as degenerate solutions, pretty much like the number of zeros in a solution in standard LP.

Analogy with standard LP. Again, consider the case where EP reduces to the standard linear program (7) At a basic solution (x, z) corresponds an extreme point (x', z') and

$$Bu + v = 0;\ u.R(x) = 0;\ v.R(z) = 0 \Rightarrow (u, v) = 0.\ \textit{i.e.},\ M(x, z)\begin{bmatrix} u \\ v \end{bmatrix} = 0 \Rightarrow (u, v) = 0.$$

To the matrix $M(x, z)$ corresponds a (M, p)-matrix M', obtained from $[H|I]$ be selecting the columns corresponding to positive x_i' and z_i'. M' has full rank and is invertible when square, i.e., M' is the usual basis in standard LP. The case where M' is a rectangular (M, p)-matrix with $(p < M)$ is precisely the case of a degenerate vertex (x', z').

At a vertex (x', z'), $D(x', z')$ is the set of (u', v') such that

$$x_i' = 0 \Rightarrow u_i' = 0;\ z_i' = 0 \Rightarrow v_i' = 0$$

(see (8)). Thus, $S(x', z') \neq R^{M+N}$ if M' is rectangular. For example, concatenate M' with a $(M, M - p)$-matrix Q by adding columns of zero x_i's and z_i's so that $B' := [M'|Q]$ is a square (M, M)-invertible matrix and

$$[H|I] = [M'|Q|N'] = [B'|N']$$

B' is then a basis of the degenerate vertex (x', z'). A straightforward calculation yields

$$S(x', z') := \{(u', v') \in R^{M+N} \mid [u' + B'^{-1} N' v']_k = 0, k = p + 1, \cdots M\} \neq R^{M+N}. \tag{9}$$

Computing π. To compute π, just note that

$$\langle (c, 0), T_{xz}(y, w)\rangle = \left\langle \begin{bmatrix} \xi \\ h \\ w \end{bmatrix}, \begin{bmatrix} T(y, w) \\ 0 \\ 0 \end{bmatrix} \right\rangle$$

where (ξ, h, ψ) solves

$$\begin{aligned} B^* \xi + \mathcal{R}(x)h &= c \\ \xi + \mathcal{R}(z)\psi &= 0 \end{aligned}$$

with $\mathcal{R}(x)$ (resp. $\mathcal{R}(z)$) being the adjoint of the operator $u \rightarrow u.R(x)$ (resp. $v \rightarrow v.R(z)$), so that π is just $(c - B^* \xi, -\xi)$. Non-uniqueness of (ξ, h, ψ) is irrelevant since $\langle(\xi, h, \psi), (T(y, w), 0, 0)\rangle$ is constant over the solutions to the above system.

Analogy with standard LP. Again, consider the case where *EP* reduces to the standard LP (7). The reduced cost π is the vector $(c - B^* \xi, -\xi)$. At a vertex (x', z'), the above computation of $\xi \in R^M$ yields

$$\xi^T H^i = c_{x_i'} \textit{ if } x_i' > 0;\ \xi_k = 0 \textit{ if } z_k' > 0;\ \xi^T H^i + \nu_i = c_{x_i'} \textit{ if } x_i' = 0;\ \xi_i + \mu_i = 0 \textit{ if } z_i' = 0$$

where H^i denotes the column in H corresponding to the variable x_i'. As before, partition the matrix $[H|I]$ into $[M'|N']$ and denote $c'_{M'}$, the vector of components of c' for which $x_i' > 0$ and 0 if $z_i' > 0$. Let $c'_{N'}$, be the vector of the remaining components. Observe that

$$\xi^T M' = c'_{M'},$$

or equivalently $\xi = c'^T_{M'} M'^{-1}$ in the non-degenerate case. Partition also π into $[\pi_{M'}|\pi_{N'}]$ so that

$$\pi_{M'} = 0;\ \pi_{N'} = c'^T_{N'} - c'^T_{M'} M'^{-1} N' \tag{10}$$

which corresponds to the definition of the reduced cost in standard LP. In this case, $S(x, z)$ is just R^{M+N} so that if $-\pi$ is a nonnegative linear functional on $S(x, z)$, we retrieve the test of optimality of a basic solution in standard LP, i.e., $c'^T_{N'} - c'^T_{M'} M'^{-1} N' \leq 0$.

In the degenerate case in standard LP, M' is rectangular but can be complemented with columns of zero x_i's and z_i's until a square invertible matrix is obtained. The reduced cost π can still be computed in the same manner. But an anti-cycling rule is then necessary for the cost might remain constant at the next iteration of the Simplex algorithm (see next section).

3.5. Finding a Feasible Direction of Improvement

Assume that we are at some basic feasible solution (x, z). From what we have just seen, it suffices to find $(x', z') \in S(x, z)$, $(x', z') \succcurlyeq 0$ such that $\langle \pi, (x', z') \rangle > 0$. This is the analogue of the Simplex criterion to select a nonbasic variable that will enter the basis. Then, there will be some $\lambda > 0$ such that

$$(y, w) := (x, z) - \lambda(x_1, z_1) + \lambda(x', z') \succcurlyeq 0$$

$$T(y, w) := T((x, z) - \lambda(x_1, z_1) + \lambda(x', z')) = b$$

Equation to be placed

with $(x_1, z_1) = P_D(x', z')$, so that (y, w) is a new feasible solution with strictly better value.

1. At a non-degenerate vertex

In this case $S(x, z) = X \times Z$. Let $\pi \in X \times Z$, the reduced cost at (x, z), be written (π_1, π_2) with $\pi_1 \in X$, $\pi_2 \in Z$. Then (x, z) non-optimal implies $\pi_1 \not\preccurlyeq 0$ or $\pi_2 \not\preccurlyeq 0$. Thus, say π_1, has at least one strictly positive eigenvalue and its Jordan decomposition yields

$$\pi_1 = Q \begin{bmatrix} \Lambda & | & 0 \\ -- & & -- \\ 0 & | & S \end{bmatrix} Q^T$$

for some nonsingular matrix Q and where $\Lambda \neq 0$ is a diagonal matrix with only nonnegative entries and S is a diagonal matrix with only negative entries. If we select x' as

$$x' = Q \begin{bmatrix} \Theta & | & 0 \\ -- & & -- \\ 0 & | & 0 \end{bmatrix} Q^T$$

with Θ any diagonal matrix with positive entries, then $\langle \pi, (x', 0) \rangle = \langle \pi_1, x' \rangle > 0$ and thus, in the direction $(u, v) := (x', 0) - P_D(x', 0)$ one may strictly improve the objective.

2. At a degenerate vertex

In this case, $S(x, z) \neq X \times Z$. A simple characterization of $S(x, z)$ is as follows: $S(x, z) := D(x, z) + N(T)$ so that

$$(y, w) \in S(x, z) \Leftrightarrow \exists(u_1, v_1), (u_2, v_2) \ s.t. \left[\begin{array}{lcl} y & = & u_1 + u_2 \\ w & = & v_1 + v_2 \\ Bu_2 + v_2 & = & 0 \\ u_1.R(x) & = & 0 \\ v_1.R(z) & = & 0 \end{array}\right.$$

or in other words

$$S(x, z) := \{(y, w) \in N \times Z | \langle \xi_i, By + w \rangle = 0, i = 1, \cdots s\}.$$

where the $\{(\xi_i, h_i, \psi_i)\}$, $i = 1, \ldots s$, form a basis of the subspace

$$\begin{aligned} B^* \xi + \mathcal{R}(x)h &= 0 \\ \xi + \mathcal{R}(z)\psi &= 0 \end{aligned}$$

Hence, finding $(x', z') \in S(x, z)$ such that $\langle \pi, (x', z') \rangle \geq 0$ reduces to solve

$$P \left[\begin{array}{ll} \max\langle c - B^* \xi, x' \rangle - \langle \xi, z' \rangle & \\ \langle \xi_i, Bx' + z' \rangle & = 0, i = 1, \cdots s \\ x', z' & \geqslant 0 \end{array}\right]$$

a small dimension problem in general. Again, the direction $(u, v) = (x', z') - P_D(x', z')$ is a feasible direction of improvement. In the program P above, one sees why the reduced-cost vector $\pi = (c - B^* \xi, -\xi)$ need not be unique and can be any solution to

$$B^*\xi + \mathcal{R}(x)h = c; \ \xi + \mathcal{R}(z)\psi = 0$$

since $\langle \xi_i, Bx' + z' \rangle = 0$, $i = 1, \ldots s$.

Pivoting procedure

Once (x', z') has been calculated, one moves to $(x_1, z_1) := (x + \lambda^* u, z + \lambda^* v)$ until some eigenvalue of x_1 or z_1 becomes strictly negative. If this does not happen then EP is unbounded.

If (x_1, z_1) is a basic solution, then the pivoting procedure is terminated. If not, then by Theorem 3.3, there exists (u, v) such that

$$\begin{aligned} Bu + v &= 0 \\ u.R(x_1) &= 0 \\ v.R(z_1) &= 0 \\ \langle c, u \rangle &\geq 0 \end{aligned}$$

and we can move in the direction (u, v) to another solution without decreasing the objective and until some positive eingenvalue becomes zero. We end up with a basic solution in a finite number of this step since the number of zero-eigenvalues strictly increases by at least one at each step. In the simplex-like algorithm described later, we will use another feasible direction of improvement.

Analogy with standard LP. Again consider (7). At a non-degenerate vertex (x', z'), we have seen that the reduced cost π is $[0, c'^T_{N'} - c'^T_{M'}, M'^{-1} N']$ (see (10)). Thus, selecting $(u', v') \geqslant 0$ in $S(x, z)$ such that $\langle \pi, (u', v') \rangle > 0$ reduces to find a strictly positive component of $c'^T_N - c'^T_{M'}, M'^{-1} N'$, i.e., to select a non-basic variable with positive reduced-cost. The direction $(u, v){:} = (u', v') - P_D(u', v')$ is given by $(-M'^{-1} Nv', v')$ and is a direction of strict improvement.

At a degenerate vertex in standard LP, one avoids considering $S(x, z)$ and looks for a nonnegative (u', v') in R^{M+N} such that $\langle \pi, (u', v') \rangle > 0$, i.e., exactly as in the non-degenerate case. However, a strict improvement is not guaranteed and an anti-cycling rule (e.g., the lexicographic rule) is necessary. The direction of improvement calculated in the degenerate case and translated in standard LP would yield the following:

At the degenerate vertex (x', z') the matrix $[H|I]$ is partitioned into $[B'|N']{:} = [M'|Q|N']$ with $B' = [M'|Q]$ as basis (see previous section). Then, one would compute (u', v') in solving

$$\max\{(c'^T_{N'} - c'^T_{B'} B'^{-1} N')v' \mid [u' + B'^{-1} N'v']_k = 0 \; k = p + 1, \cdots M; \quad u', v' \geq 0\} \tag{11}$$

(see (9)). The direction $(u, v){:} = (-B'^{-1} Nv', v')$ is a direction of strict improvement since the zero components of (x', z') would increase. In the non-degenerate case, in the direction of improvement calculated, one would end up at another basic solution since only one non-basic variable may become positive and at least one basic variable will become null. In the degenerate case, if q basic variables are null then (u', v') solution of (11) has at most $q + 1$ positive components (to the q constraints in (11) we must add a normalizing constraint) so that in the direction $(-B'^{-1} Nv', v')$ we end up at a solution with at most M positive variables, i.e., still a basic solution. This is a difference with semi-definite programming, where in the direction of improvement we may end up at a non-basic solution, in which case an extra step is required. Note that in standard LP, this procedure would avoid the anti-cycling rule needed in the Simplex algorithm and would guarantee a strict increase in the objective.

3.6. Characterization of Optimal Solutions

We have seen one necessary and sufficient condition for optimality of a basic solution (x, z) through its reduced cost π. We also have two other direct characterizations of an optimal solution without using π.

THEOREM 3.5. *At a feasible solution (x, z), consider the following linear program $P_1(x, z)$*

$$P_1\,(x, z) \begin{cases} max\langle c, u\rangle \\ Bu + v \quad = 0 \\ \langle R(x), u\rangle \geq 0 \\ \langle R(z), v\rangle \geq 0 \end{cases} \tag{12}$$

If the optimal value of P_1 is 0, then (x, z) is an optimal solution.

Proof. If the optimal value of P_1 (x, z) is zero, then by a standard Farkas Lemma, the system

$$\begin{aligned} B^* y + \lambda R(x) &= -c \\ y + \nu R(z) &= 0 \\ \lambda, \nu &\geq 0 \end{aligned}$$

has a feasible solution (y, λ, ν). Indeed, Farkas Lemma is valid here for in P_1 (x, z) we have an homogeneous system of equalities and inequalities which describe a convex *polyhedral* cone. Multiplying both sides by (x, z) yields

$$\langle x, B^* y + c\rangle + \lambda\langle x, R(x)\rangle = 0$$

$$\langle z, y\rangle + \nu\langle z, R(z)\rangle = 0$$

but we know that $x.R(x) = 0$ and $z.R(z) = 0$ so that we finally have

$$\langle x, B^* y + c\rangle = 0, \langle z, y\rangle = 0.$$

We also have $y \leqslant 0$ since $R(z) \geqslant 0$ and $\nu \geq 0$. Also, $B^* y + c \leqslant 0$ since $R(x) \geqslant 0$ and $\lambda \geq 0$. Therefore, (x, z) and $-y$ are two feasible solutions in EP and EP^* respectively, and satisfy

$$Bx \leqslant b;\ - B^* y \geqslant c;\ x, - y \geqslant 0 \text{ and } \langle x, B^* y + c\rangle = 0, \tag{13}$$

i.e., *complementarity slackness* holds, which in turn proves that x and $-y$ are optimal solutions (see [1]).□

Remark 3.3: $P_1(x, z)$ reduces to a standard LP problem by replacing a constraint $\langle a, x\rangle$ with the usual scalar product $\bar{a}^T x$ for an appropriate $\bar{a}$.

Note that we only have a sufficient condition for x to be optimal. The next characterization of optimality is a necessary *and* sufficient condition. We first make the following assumption

ASSUMPTION A: *There exists $x_0 \succ 0$ such that $Bx_0 \prec b$.*

THEOREM 3.6. *Assume that A holds. At a feasible solution (x, z), consider the following linear program $P_2(x, z)$:*

$$P_2(x, z) \begin{cases} max\langle c, u\rangle & \\ Bu + v & = 0 \\ P^T(x).u.P(x) & \succcurlyeq 0 \\ P^T(z).v.P(z) & \succcurlyeq 0 \end{cases} \tag{14}$$

(x, z) is an optimal solution if and only if the optimal value of $P_2(x, z)$ is zero.

Proof. i) if. Assume the optimal value of $P_2(x, z)$ is zero and (x, z) is not optimal. Then consider any other feasible solution (x', z') with strictly better value.

Obviously, the direction $(u, v){:} = (x' - x, z' - z)$ satisfies $Bu + v = 0$ and $\langle c, u\rangle > 0$. In addition, as the set Ω of feasible solutions in EP is a convex set, $(x + \rho u) \succcurlyeq 0$ and $(z + \rho v) \succcurlyeq 0$ for all $\rho \leq 1$. Now, if $P^T(x).u.P(x) \not\succcurlyeq 0$, it means that one zero-eigenvalue of x has a negative derivative in the direction u. Therefore, for $\rho > 0$ small enough, $(x + \rho u) \not\succcurlyeq 0$, a contradiction. The same argument holds true for $P^T(z).v.P(z)$. Thus, (u, v) is feasible in $P_2(x, z)$ and $\langle c, u\rangle > 0$, a contradiction. Therefore, (x, z) must be optimal.

ii) only if. Assume that (x, z) is optimal and the optimal value of $P_2(x, z)$ is not zero (hence $+\infty$). Any matrix u can be written $u^+ - u^-$ with $u^+ \succcurlyeq 0$ and $u^- \succcurlyeq 0$ and we note $|u|{:} = u^+ + u^-$. Consider the following perturbation $P_2(x, z, \epsilon)$ of $P_2(x, z)$ ($\equiv P_2(x, z, 0)$).

$$P_2(x, z, \epsilon) \begin{cases} \max\langle c, u\rangle & \\ Bu + v & = 0 \\ P^T(x).u.P(x) & \succcurlyeq \epsilon I_1 \\ P^T(z).v.P(z) & \succcurlyeq \epsilon I_2 \\ \langle I_X, |u|\rangle + \langle I_Z, |v|\rangle & \leq 1 \end{cases} \tag{15}$$

where I_1 (resp. I_2) is the identity matrix of same dimension as $P^T(x).u.P(x)$ (resp. $P^T(z).v.P(z)$) and where the last constraint is a normalizing constraint to avoid an infinite value. $P_2(x, z, \epsilon)$ is consistent (i.e., has a feasible solution) for ϵ small enough. Indeed, under Assumption A, the direction $(u_0, v_0){:} = (x_0 - x, z_0 - z)$ satisfies $Bu_0 + v_0 = 0$. Moreover, since x_0 has strictly positive eigenvalues, $u_0.R(x) \neq 0$ and $P^T(x).u_0.P(x) \succ 0$. Otherwise, if some eigenvalue of $P^T(x).u_0.P(x)$ is nonpositive, then in the direction u_0, the smallest eigenvalue will become negative (as a concave function with nonpositive derivative and because $u_0.R(x) \neq 0$). Similarly, $P^T(z).v_0.P(z) \succ 0$. Thus, (u_0, v_0) is feasible for $P_2(x, z, \epsilon)$ when ϵ is sufficiently small since then, $P^T(x).u_0.P(x) \succcurlyeq \epsilon I_1$ and $P^T(z).v_0.P(z) \succcurlyeq \epsilon I_2$.

$P_2(x, z, 0)$ (with a normalizing constraint) has a finite positive optimal value δ. In addition, the feasible set being compact, $P_2(x, z, 0)$ has no duality gap so that for a small perturbation of the right-hand side, its optimal value (well-defined since $P_2(x, z, \epsilon)$ is consistent) is still positive, i.e., $P_2(x, z, \epsilon) > \delta/2$ for ϵ small enough. Pick (u, v) in $P_2(x, z, \epsilon)$ with $\langle c, u\rangle > \delta/2$. In the direction (u, v), $(x + \rho u, z + \rho v)$ is still feasible for ρ small enough since $P^T(x).u.P(x) \succ 0$ and $P^T(z).v.P(z) \succ 0$. In addition, one may strictly improve the criterion, a contradiction with (x, z) optimal.□

Remark 3.4: In $P_2(x, z)$, the constraint $P^T(x).u.P(x) \succcurlyeq 0$ states that the derivatives (in the direction u) of the zero-eigenvalues of x are nonnegative (see [13]). Thus, one may replace $P^T(x).u.P(x)$ by $R(x).u.R(x)$ since these derivatives are also the eigenvalues of $R(x).u.R(x)$ in the subspace $R(x)X$ (see [10], [13]). This fact will be used later.

4. A SIMPLEX-LIKE ALGORITHM

In this section we first describe a conceptual Simplex-like algorithm for positive semi-definite programming. By Simplex-like we mean that with this algorithm one moves from one vertex to another one like in the standard Simplex algorithm.

Assume that we are at some current basic solution (x, z).

Step 0. Compute the reduced cost $\pi = (c - B^* \xi - \xi)$ where (ξ, h, ψ) solves:

$$B^*\xi + \mathcal{R}(x)h = c: \xi + \mathcal{R}(z)\psi = 0.$$

If $\langle \pi, (x', z')\rangle \leq 0$ for all $(x', z') \succcurlyeq 0$ in $S(x, z)$ then STOP. (x, z) is optimal.

Step 1. compute $(x', z') \succcurlyeq 0$ in $S(x, z)$ such that $\langle \pi, (x', z')\rangle > 0$ and a feasible direction of descent (u, v): $= (x', z') - P_D(x', z')$, i.e., solve

$$\begin{bmatrix} Bu + v = 0 \\ u.R(x) = x'.R(x) \\ v.R(z) = v'.R(z) \end{bmatrix} \Leftrightarrow \begin{bmatrix} Bu' + v' = Bx' + z' \\ u'.R(x) = 0 \\ v'.R(z) = 0 \end{bmatrix}$$

which has a unique solution.

Step 2. Move in the direction (u, v) until some eigenvalue of (x_1, z_1): $= (x + \lambda^* u, z + \lambda^* v)$ becomes strictly negative. Check if (x_1, z_1) is basic. If basic then set (x, z): $= (x_1, z_1)$ and go to Step 0. Otherwise go to Step 3.

Step 3. Compute (u, v) such that

$$\begin{aligned} Bu + v &= 0 \\ u.R(x_1) &= 0 \\ v.R(z_1) &= 0 \\ \langle c, u\rangle &\geq 0 \end{aligned}$$

If $(u, v) = (0, 0)$ then (x_1, z_1) is basic and go to Step 0. Otherwise, move in the direction (u, v) until some positive eigenvalue of $(x_1 + \rho u, z_1 + \rho v)$ becomes zero (the zero eigenvalues remain zero). Set $(x_1, z_1) := (x_1 + \rho u, z_1 + \rho v)$ and go to Step 3.

Note that the inner loop in Step 3 stops after a finite number of iterations since the number of zero-eigenvalues strictly increases by at least one at each iteration.

One may also simply replace Step 0 and Step 1 by the single step:

Step 0–1. Compute a direction (u, v) that solves

$$P_2(x, z) \begin{cases} \max\langle c, u\rangle & \\ Bu + z & = 0 \\ P^T(x).u.P(x) & \succcurlyeq 0 \\ P^T(z).v.P(z) & \succcurlyeq 0 \\ \langle I_X, |u|\rangle + \langle I_Z, |v|\rangle & \leq 1 \end{cases}$$

If the optimal value is zero STOP, (x, z) is optimal. If not then go to Step 2.

(In fact compute (u, v) with the perturbed version $P_2(x, z, \epsilon)$ as in (15) to ensure a strict increase of the criterion). Note that the feasible direction problem $P_2(x, z, \epsilon)$ is of the same type as the original problem except the matrices $P^T(x).u.P(x)$ and $P^T(z).v.P^T(z)$ are of much smaller size, i.e., r and r', the respective numbers of zero-eigenvalues of x and z. Moreover, optimality in $P_2(x, z)$ is not required since a feasible direction (u, v) with $\langle c, u\rangle > 0$ is sufficient.

Convergence of the Simplex algorithm. In contrast to standard LP, only *asymptotic* convergence might be expected since the number of extreme points is not finite in general. In addition, the above algorithm requires some care for at a point (x, z) where x (and/or z) has r (> 1) zero-eigenvalues, the matrix $P(x)$ (as well as $P(z)$) is not continuous in general. On the other hand, the *total projection* for the 0-*group* $\sum_{k=1}^{r'} R_k(x')$ is continuous in a neighborhood of x (see [10]) (when the r zero-eigenvalues of x split into r' (eventually multiple) eigenvalues). In the direction u, the derivatives of the r repeated zero-eigenvalues of x are the repeated eigenvalues of $P^T(x).u.P(x)$, but also the repeated eigenvalues of $R(x).u.R(x)$ in the subspace $R(x)X$. (see [10]).

The above algorithm may not converge due to the lack of continuity of $P(x)$. One might have a sequence of iterates (x_k, z_k) converging to some (x, z) but (x, z) not being optimal. For instance, the number of zero-eigenvalues of x_k and z_k might be p and $p + 1$ for (x, z). The corresponding positive $(p + 1)th$ eigenvalue of x_k or z_k decreases to zero, but the direction of descent at (x_k, z_k), computed in $P_2(x_k, z_k)$ is such that a zero eigenvalue of x_k or z_k increases a little and becomes again zero, whereas the $(p + 1)th$ decreases slowly. The stepsize is smaller and smaller and one remains stuck to a nonoptimal solution (x, z).

To avoid a *jamming* phenomenon, one should replace the above step 0–1 by the new steps:

Step 0: set $\epsilon = \epsilon_0 > 0$.

Step 1: Let $\lambda_k(x)$, $k = 1,..N$ (resp. $\lambda_k(z)$) denote the eigenvalues of x and define the sets:

$$\lambda_\epsilon(x) := \{\lambda_k(x) | \lambda_k(x) \leq \epsilon\}. \ \lambda_\epsilon(z) := \{\lambda_k(z) | \lambda_k(z) \leq \epsilon\}.$$

Let $R^j_\epsilon(x)$ be $\sum_1^j R_k(x)$ where the $\{R_k(x)\}$ are the eigenprojections of the (eventually multiple) eigenvalues $\lambda_k(x) \in \lambda_\epsilon(x)$ (with same definition for $R^j_\epsilon(z)$).

$$P'_2(x, z, \epsilon) \begin{cases} \max\langle c, u\rangle & \\ Bu + z & = -\epsilon(\alpha B I_X + \beta I_Z) \\ R^j_\epsilon(x).u.R^j_\epsilon(x) & \succcurlyeq 0 \\ R^j_\epsilon(z).v.R^j_\epsilon(z) & \succcurlyeq 0 \\ \langle I_X, |u|\rangle + \langle I_Z, |v|\rangle & \leq 1 \end{cases}$$

If $\langle c, u\rangle > \epsilon$ go to step 2 otherwise set $\epsilon := \epsilon/2$ and go to step 1.

As in Theorem 3.6, one may show that if (x, z) is not optimal, then $P'_2(x, z, 0)$ as well as $P'_2(x, z, \epsilon)$ have positive value for ϵ small enough. The α and β parameters are chosen according to [7], e.g., $\alpha \geq \|\Lambda(x)^{-1}\|.\|u\|^2$ where $\Lambda(x)$ is the diagonal matrix of the positive eigenvalues of x. Since $\|u\|$ and $\|v\|$ are bounded in $P'_2(x, z, \epsilon)$ it suffices to take α and β large enough. Then, from e.g. [7], $(x + \gamma u + \gamma^2\alpha I_X, z + \gamma v + \gamma^2 \beta I_Z) \succcurlyeq 0$ along this trajectory for γ small enough. Thus, from a solution (u, v) in $P'_2(x, z, \epsilon)$, $B(u/2^k + (\epsilon/2^k)\alpha I_X) + (v/2^k + (\epsilon/2^k)\beta I_Z) = 0$ so that when 2^k is large enough, $x' := x + (\epsilon/2^k)u + (\epsilon^2/2^k)\alpha I_X \succ x + (\epsilon/2^k)u + (\epsilon/2^k)^2\alpha I_X \succcurlyeq 0$ and $\langle c, x'\rangle > \langle c, x\rangle$ (since $\langle c, u\rangle > 0$). Thus, $P'_2(x, z, \epsilon)$ provides us with a feasible trajectory along which one strictly improves the criterion.

THEOREM 4.1. *Assume that EP has finite value and there exists $x_0 \succ 0$, such that $Bx_0 \prec b$. The above Simplex-like algorithm converges asymptotically to an optimal solution.*

Proof. Consider a sequence (x^k, z^k) generated by the above algorithm and let ϵ_k be the corresponding value of ϵ at the end of step 1. The sequence of associated costs increases monotonically and is bounded from above so that it converges to some value g^*. Moreover, as $c \succ 0$, $c \succcurlyeq \gamma I_X$ for some scalar γ so that x^k is bounded. Consequently, z^k is also bounded. In addition, the optimal solution (u_k, v_k) in $P'_2(x^k, z^k, \epsilon_k)$ is also bounded. Thus, let (x^*, z^*), (u^*, v^*), η be any accumulation point of the sequence $\{(x^k, z^k), (u_k, v_k), \epsilon_k\}$, (r, r') being the number of zero-eigenvalues of (x^*, z^*).

We first prove that $\eta = 0$. Indeed, assume not, i.e., there is some subsequence still denoted $\{(x^k, z^k)\}$ such that $\epsilon_k \to \eta > 0$. We may assume that η is still small enough so that $R^j_\eta(x^*)$ is the eigenprojection corresponding to the r zero-eigenvalues of x^* (and similarly for z^*), i.e., $\lambda_\eta(x^*)$ contains only the r zero-eigenvalues of x^*. For k large enough, the eigenvalues in $\lambda_{\epsilon_k}(x^k)$ correspond to the r zero-eigenvalues of x^* (and similarly for z^*). Thus, by a continuity argument (the total eigenprojection of the 0-group is continuous), u^* is feasible in $P'_2(x^*, z^*, \eta)$. Then, for some integer p, $x^* + (\eta/2^p)u^* + (\eta^2/2^p)\alpha I_X \succ x^* + (\eta/2^p)u^* + (\eta/2^p)^2\alpha I_X \succcurlyeq 0$, and thus, again by a continuity argument, $x^k + (\epsilon_k/2^p)u_k + (\epsilon_k^2/2^p)\alpha I_X \succcurlyeq 0$ (and similarly for z^k) for k large enough in the subsequence, so that the step-size is bounded from below in the subsequence considered, a contradiction with the convergence of the sequence.

Hence, $\epsilon_k \to 0$ for a subsequence still denoted $\{\epsilon_k\}$. Assume that the optimal value of $P'_2(x^*, z^*, 0)$ is positive, say δ. With similar arguments as in Theorem 3.6, $P'_2(x^*, z^*, 0)$ has no duality gap so that the optimal value δ_ϵ of $P_2'(x^*, z^*, \epsilon)$ satisfies $\delta_\epsilon \geq \delta/2$ for ϵ small enough. As soon as $\epsilon \leq \bar{\epsilon}$ (for some $\bar{\epsilon} > 0$) and x^k is close enough to x^*, the eigenvalues in $\lambda_\epsilon(x^k)$ correspond to the r zero-eigenvalues of x^* (and similarly for z^*). Thus, by a continuity argument (remember that the total eigenprojection of the 0-group is continuous), the optimal value of $P_2'(x^k, z^k, \epsilon) \geq \delta/4$ for all $\epsilon \leq \bar{\epsilon}$ and k large enough which implies that ϵ_k is bounded from below, a contradiction.

Thus, the optimal value of $P_2'(x^*, z^*, 0)$ is zero, which in turn implies that (x^*, z^*) is optimal (see Remark 3.4).□

5. CONCLUSION

We have presented analogues for positive semi-definite programming of the basic concepts in standard LP. For practical purposes and in comparison with other existing methods like cutting planes or interior point methods as in e.g., [8], [12], the efficiency of the Simplex-like algorithm remains to be proved, and is a topic of further research beyond the scope of this paper.

References

1. E. J. Anderson and P. Nash, *Linear programming in infinite dimensional spaces*, John Wiley & Sons, 1987.
2. S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, vol 15, 1994.
3. A. Ben Israel, *Linear Equations and Inequalities on Finite Dimensional, Real or Complex, Vector Spaces: A Unified Theory*, J. Math. Anal. Appl., 27 (1969), pp. 367–389.
4. A. Berman and A. Ben-Israel, *More on Linear Inequalities with applications to Matrix Theory*, J. Math. Anal. Appl., 33 (1971), pp. 482–496.
5. B. D. Craven and J. J. Koliha, *Generalizations of Farkas' Theorem*, SIAM J. Math. Anal., 8 (1977), pp. 983–997.
6. R. Fletcher, *Semi-definite matrix constraints in optimization*, SIAM J. Contr. Optim., 23 (1985), pp. 493–513.
7. J. C. Geromel, P. L. D. Peres and J. Bernussou, *On a convex parameter space method for linera control design of uncertain systems*, SIAM J. Contr. Optim., 29 (1991), pp. 381–402.
8. J. B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer-Verlag, New York, 1993.
9. T. Kato, *Perturbation Theory for Linear Operators, 2nd* ed., Springer-Verlag, 1984.
10. J. B. Lasserre, *A new Farkas Lemma for positive semi-definite matrices*, IEEE Trans. Aut. Contr., 40 (1995), pp. 1131–1133.
11. Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, vol 13, 1994.
12. M. L. Overton and R. S. Womersley, *On minimizing the spectral radius of a nonsymmetric matrix function: Optimality conditions and duality theory*. SIAM J. Matr. Anal. Appl., 9 (1988), pp. 473–498.

APPENDIX

Below is an example in Control Theory, to illustrate various notions explored in the paper. This example is taken from a Tutorial workshop *Convex Optimization Techniques in Robust Control*, held at the IEEE-CDC conference, San Antonio, 1993.

Consider the problem

$$\text{minimize } trace(x)$$

$$A^T x + xA + xBB^T x + Q \preccurlyeq 0$$

with

$$A = \begin{bmatrix} -1 & -2 & 1 \\ 3 & 2 & 1 \\ 1 & -2 & -1 \end{bmatrix} B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} Q = \begin{bmatrix} 1 & -1 & 0 \\ -1 & -3 & -12 \\ 0 & -12 & -36 \end{bmatrix}$$

or equivalently

$$\text{minimize } \langle I, x \rangle$$

$$\begin{pmatrix} A^T x + xA & xB \\ B^T x & 0 \end{pmatrix} \preccurlyeq \begin{pmatrix} -Q & 0 \\ 0 & I \end{pmatrix}$$

where x is a (3, 3) symmetric real-valued matrix.

Provided that (A, B) is stabilizable, the solution x^* is the stabilizing solution of the algebraic Riccati equation

$$A^T x + xA + xBB^T x + Q = 0.$$

(see [2]). As the optimal solution x^* is negative definite, we may also solve equivalently

$$\text{maximize } \langle I, x \rangle$$

$$\begin{pmatrix} A^T x + xA & xB \\ B^T x & 0 \end{pmatrix} - z = \begin{pmatrix} Q & 0 \\ 0 & -I \end{pmatrix}$$

$$x, z \succcurlyeq 0$$

If we now write the above program in matrix notation with x a vector in R^6 and z a vector in R^{10} we have to solve

$$\begin{aligned} &\max \langle I, x \rangle \\ &Hx - z = b \\ &x, z \succcurlyeq 0 \end{aligned}$$

with now

$$H = \begin{bmatrix} -2 & 6 & 2 & 0 & 0 & 0 \\ -2 & 1 & -2 & 3 & 1 & 0 \\ 1 & 1 & -2 & 0 & 3 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 4 & -4 & 0 \\ 0 & 1 & -2 & 1 & 1 & -2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 2 & -2 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ -3 \\ -12 \\ 0 \\ -36 \\ 0 \\ -1 \end{bmatrix}$$

1- Basic solutions, basis and reduced-cost

Consider for example the feasible solution

$$x = \begin{bmatrix} 5.3714 & 4.5281 & -2.7642 \\ 4.5281 & 4.2145 & -3.1901 \\ -2.7642 & -3.1901 & 3.8641 \end{bmatrix}$$

$$z = \begin{bmatrix} 9.8976 & 9.7671 & 9.7217 & 2.6072 \\ 9.7671 & 14.5059 & 15.3528 & 1.3380 \\ 9.7217 & 15.3528 & 16.3631 & 1.0998 \\ 2.6072 & 1.3380 & 1.0998 & 1.0000 \end{bmatrix}$$

where trace$(x) = 13.4500$. z has one double zero-eigenvalue with corresponding (rank 2) eigenprojection matrix

$$R(z) = \begin{bmatrix} 0.1992 & -0.1136 & 0.0140 & -0.3826 \\ -0.1136 & 0.5781 & -0.4781 & 0.0487 \\ 0.0140 & -0.4781 & 0.4317 & 0.1285 \\ -0.3826 & 0.0487 & 0.1285 & 0.7911 \end{bmatrix}$$

i.e., $\mathrm{R}(z) = \mathrm{P}_1(z)\,\mathrm{P}_1^{\mathrm{T}}(z) + \mathrm{P}_2(z)\,\mathrm{P}_2^{\mathrm{T}}(z)$ where $P_1(z)$ and $P_2(z)$ are the two zero-eigenvectors of z.

$$[P_1(z), P_2(z)] = \begin{bmatrix} -0.0231 & 0.4457 \\ 0.7287 & -0.2171 \\ -0.6570 & -0.0027 \\ -0.1920 & -0.8685 \end{bmatrix}$$

The condition $v.R(z) = 0$ reduces to $v.P_1(z) = 0$ and $v.P_2(z) = 0$, or in matrix notation $Sv = 0$ with

$$S = \begin{bmatrix} -0.023 & 0.728 & -0.657 & -0.192 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.023 & 0 & 0 & 0.729 & -0.657 & -0.192 & 0 & 0 & 0 \\ 0 & 0 & -0.023 & 0 & 0 & 0.729 & 0 & -0.657 & -0.192 & 0 \\ 0 & 0 & 0 & -0.023 & 0 & 0 & 0.729 & 0 & -0.657 & -0.192 \\ 0.446 & -0.217 & -0.003 & -0.868 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.446 & 0 & 0 & -0.217 & -0.003 & -0.868 & 0 & 0 & 0 \\ 0 & 0 & 0.4460 & 0 & -0.217 & 0 & -0.003 & -0.868 & 0 & \\ 0 & 0 & 0 & 0.446 & 0 & 0 & -0.217 & 0 & -0.003 & -0.868 \end{bmatrix}$$

Hence, the basis at the solution (x, z) is the (rank 16) (18, 16)-matrix

$$M(x, z) = \begin{bmatrix} H & | & -I \\ - & - & - \\ 0 & | & S \end{bmatrix}$$

The subspace $S(x, z)$ is easy to characterize. It suffices to look at the conditions on (y, w) to ensure that the system

$$y = x_1 + x_2;\ u = z_1 + z_2;\ Hx_2 - z_2 = 0;\ Sz_1 = 0$$

has at least a solution (x_1, z_1), (x_2, z_2). The conditions are simply

$$h \in R^8,\ H^T S^T h = 0 \Rightarrow h^T Sw = 0$$

or equivalently $WSw = 0$ where the rows of W span the subspace $\{h \in R^8 \mid H^T S^T h = 0\}$.

To compute the reduced cost (see Section 3.4), find any solution to the system

$$\begin{aligned} H^T \xi &= c \\ -\xi + S^T h &= 0 \end{aligned}$$

and the reduced-cost π is just the vector $(c - H^T \xi, \xi) = (0, \xi)$. The equivalent of the simplex criterion reduces to find (y, w) such that

$$\langle (0, \xi), (y, w) \rangle > 0;\ WSw = 0;\ y, w \geqslant 0$$

and $(u, v) = (y, w) - P_D(y, w)$ is a direction of strict ascent. $P_D(y, w)$ is the projection on $D(x, z)$ of (y, w), i.e., the above (x_1, z_1) in the decomposition of (y, w) (guaranteed to exist).

A solution (x, z) with $x \succ 0$ and where z has a single zero-eigenvalue is not a basic solution since the corresponding matrix $M(x, z)$ will be (14, 16) and will have rank less than 16. This is the case for example at the solution

$$x = \begin{bmatrix} 5.1935 & 4.4068 & -2.3644 \\ 4.4068 & 4.2061 & -2.6182 \\ -2.3644 & -2.6182 & 3.8641 \end{bmatrix}$$

$$z = \begin{bmatrix} 10.3247 & 9.7486 & 10.3384 & 2.8291 \\ 9.7486 & 12.6703 & 14.9953 & 1.7885 \\ 10.3384 & 14.9953 & 18.3065 & 1.4996 \\ 2.8291 & 1.7885 & 1.4996 & 1.0000 \end{bmatrix}$$

where $\text{trace}(x) = 13.2636$. Indeed, one can find a direction (u, v) such that $M(x, z)(u, v) = 0$ and $\text{trace}(u) > 0$ which yields to the previous basic solution.

A nondegenerate basic solution will have a square invertible matrix $M(x, z)$ which could be the case if for example, the conditions $u.R(x) = 0$ and $v.R(z) = 0$ yield 6 constraints on the variables u and v and the matrix $M(x, z)$ is invertible.

One may also check that at the optimal solution

$$x^* = \begin{bmatrix} 6.3542 & 5.8895 & -2.2046 \\ 5.8895 & 6.2855 & -2.2201 \\ -2.2046 & -2.2201 & 6.0770 \end{bmatrix}$$

$$z^* = \begin{bmatrix} 17.2194 & 15.2267 & 16.0696 & 4.1496 \\ 15.2267 & 13.4644 & 14.2101 & 3.6694 \\ 16.0696 & 14.2101 & 14.9966 & 3.8724 \\ 4.1496 & 3.6694 & 3.8724 & 1.0000 \end{bmatrix}$$

with $\text{trace}(x^*) = 18.7168$, the matrix z^* has a triple zero-eigenvalue so that the corresponding matrix $M(x^*, z^*)$ has rank 16 which proves that (x^*, z^*) is also a basic optimal solution.