

Numerical resolution of cone-constrained eigenvalue problems

A. PINTO DA COSTA¹ and ALBERTO SEEGER²

¹Universidade Técnica de Lisboa, Instituto Superior Técnico
Departamento de Engenharia Civil e Arquitectura and ICIST
Avenida Rovisco Pais, 1049-001 Lisboa, Portugal

²University of Avignon, Department of Mathematics, 33 rue Louis Pasteur
84000 Avignon, France

E-mails: apcosta@civil.ist.utl.pt / alberto.seeger@univ-avignon.fr

Abstract. Given a convex cone K and matrices A and B , one wishes to find a scalar λ and a nonzero vector x satisfying the complementarity system

$$K \ni x \perp (Ax - \lambda Bx) \in K^+.$$

This problem arises in mechanics and in other areas of applied mathematics. Two numerical techniques for solving such kind of cone-constrained eigenvalue problem are discussed, namely, the Power Iteration Method and the Scaling and Projection Algorithm.

Mathematical subject classification: 65F15, 90C33.

Key words: complementarity condition, generalized eigenvalue problem, power iteration method, scaling and projection algorithm.

1 Introduction

The Euclidean space \mathbb{R}^n is equipped with the inner product $\langle y, x \rangle = y^T x$ and the associated norm $\| \cdot \|$. The dimension n is greater than or equal to 2. Orthogonality with respect to $\langle \cdot, \cdot \rangle$ is indicated by means of the symbol \perp . Through-

out this work, one assumes that:

$$\begin{aligned} K & \text{ is a pointed closed convex cone in } \mathbb{R}^n, \\ A \text{ and } B & \text{ are real matrices of size } n \times n, \\ \langle x, Bx \rangle & > 0 \text{ for all } x \in K \setminus \{0\}. \end{aligned} \tag{1}$$

We are interested in solving numerically the following cone-constrained generalized eigenvalue problem. As usual, K^+ indicates the positive dual cone of K .

Problem 1.

$$\begin{aligned} \text{Find } \lambda \in \mathbb{R} \text{ and a nonzero vector } x \in \mathbb{R}^n \text{ such that} \\ K \ni x \perp (Ax - \lambda Bx) \in K^+. \end{aligned}$$

Problem 1 arises in mechanics [9, 10] and in various areas of applied mathematics [4, 5, 6, 12, 13]. The specific meanings of A and B depend on the context. It is important to underline that, in the present work, the matrices A and B are not necessarily symmetric.

An interesting particular case of Problem 1 is the so-called Pareto eigenvalue problem, whose precise formulation reads as follows.

Problem 2.

$$\begin{aligned} \text{Find } \lambda \in \mathbb{R} \text{ and a nonzero vector } x \in \mathbb{R}^n \text{ such that} \\ \mathbb{R}_+^n \ni x \perp (Ax - \lambda x) \in \mathbb{R}_+^n. \end{aligned}$$

For the general theory behind Problems 1 and 2, the reader is conveyed to the references [7, 11, 15, 16]. This note focuses only on numerical issues. To be more precise, we discuss two specific algorithms for solving cone-constrained eigenvalue problems, namely, the Scaling and Projection Algorithm (SPA) and the Power Iteration Method (PIM). Both techniques have some features in common, but they are different in spirit. A third algorithm, based on optimization techniques, is formulated and left open for future examination.

2 Numerical experience with the SPA

All computational tests are carried out for the particular choice $B = I_n$, but the formulation of the SPA is given for a general matrix B as in (1). The SPA is an iterative method designed for solving Problem 1 written in the following equivalent form:

Problem 3.

Find $\lambda \in \mathbb{R}$ and vectors $x, y \in \mathbb{R}^n$ such that

$$\begin{aligned}(A - \lambda B)x &= y \\ K \ni x \perp y &\in K^+ \\ \phi(x) &= 1.\end{aligned}\tag{2}$$

The normalization condition (2) prevents x from being equal to 0. Adding such condition does not change the solution set of the original cone-constrained generalized eigenvalue problem. As example of function ϕ , one may think of

$$\begin{aligned}\phi(x) &= \|x\|, \\ \phi(x) &= \sqrt{\langle x, Bx \rangle}, \\ \phi(x) &= \langle e, x \rangle \quad \text{with } e \in \text{int}(K^+),\end{aligned}\tag{3}$$

but other options are also possible. Our favorite choice is the linear function (3) because in such a case the set

$$K_\phi = \{x \in K : \phi(x) = 1\}$$

is not just compact, but also convex. In the parlance of the theory of convex cones, the set K_ϕ corresponds to a “base” of K . Unless explicitly stated otherwise, we assume that the choice (3) is in force.

The SPA generates a sequence $\{x^t\}_{t \geq 0}$ in K_ϕ , a sequence $\{\lambda_t\}_{t \geq 0}$ of real numbers, and a sequence $\{y^t\}_{t \geq 0}$ of vectors in \mathbb{R}^n . The detailed description of the algorithm is as follows:

- *Initialization:* Fix a positive scaling factor s . Pick $u \in K \setminus \{0\}$ and let $x^0 = u/\phi(u)$.

- *Iteration*: One has a current point x^t in K_ϕ . Compute the Rayleigh-Ritz ratio

$$\lambda_t = \langle x^t, Ax^t \rangle / \langle x^t, Bx^t \rangle,$$

the vector $y^t = Ax^t - \lambda_t Bx^t$, and the projection

$$v^t = \Pi_K[x^t - sy^t].$$

The projection v^t is necessarily in $K \setminus \{0\}$. Proceed now to a ϕ -normalization:

$$x^{t+1} = v^t / \phi(v^t).$$

One could perfectly well consider a scaling factor that is updated at each iteration, but in this section we stick to the case of a fixed scaling factor. The case of a variable scaling factor will be commented in Section 4. The next convergence result is taken from [11].

Theorem 1. *If $\{x^t\}_{t \geq 0}$ converges to a certain vector \bar{x} , then*

(a) $\{\lambda_t\}_{t \geq 0}$ converges to $\bar{\lambda} = \langle \bar{x}, A\bar{x} \rangle / \langle \bar{x}, B\bar{x} \rangle$.

(b) $\{y^t\}_{t \geq 0}$ converges to $\bar{y} = A\bar{x} - \bar{\lambda}B\bar{x}$.

(c) $(\bar{x}, \bar{\lambda}, \bar{y})$ solves Problem 3.

2.1 Applying the SPA to the Pareto eigenvalue problem

The Pareto eigenvalue problem already exhibits many of the mathematical difficulties arising in the context of a general convex cone K . While dealing with the Pareto cone \mathbb{R}_+^n one chooses

$$\phi(x) = x_1 + x_2 + \dots + x_n \tag{4}$$

as normalizing function. This corresponds to the linear choice (3) with $e = (1, 1, \dots, 1)^T$ belonging to the interior of \mathbb{R}_+^n .

A few words on terminology are appropriate at this point in time. One usually refers to

$$\sigma_{\text{pareto}}(A) = \{ \lambda \in \mathbb{R} : (x, \lambda) \text{ solves Problem 2} \}$$

as the Pareto spectrum (or set of Pareto eigenvalues) of A . The map $\sigma_{\text{pareto}}: \mathbb{M}_n \rightarrow 2^{\mathbb{R}}$ has some similarities with the classical spectral map, but the differences are even more numerous. It deserves to be stressed that

$$\pi_n = \max_{A \in \mathbb{M}_n} \text{card} [\sigma_{\text{pareto}}(A)]$$

is finite, but grows exponentially with respect to the dimension n of the underlying Euclidean space. Thus, a matrix of a relatively small size could have a very large number of Pareto eigenvalues. In view of this observation, it is simply not reasonable to try to compute all the elements of $\sigma_{\text{pareto}}(A)$. A more realistic goal is finding just one or a bunch of them.

2.1.1 Testing on a 2×2 matrix with 3 Pareto eigenvalues

We start by considering the 2×2 matrix

$$A = \begin{bmatrix} 8 & -1 \\ 3 & 4 \end{bmatrix} \tag{5}$$

whose Pareto spectrum is displayed in Table 1. For ease of visualization, the Pareto eigenvalues are arranged in increasing order. The corresponding solutions are named starting from S_1 up to S_3 .

Sol. name	λ	x_1	x_2	y_1	y_2
S_1	5	0.25	0.75	0	0
S_2	7	0.5	0.5	0	0
S_3	8	1	0	0	3

Table 1 – Solution set of the Pareto eigenproblem associated with matrix (5).

For this small size matrix, the SPA was able to detect all the solutions. Table 2 shows which solution was detected as function of a scaling parameter s varying in steps of 0.1 units.

A first conclusion that can be drawn from this example is that the detected solution does not need to be in the same face of the Pareto cone as the initial vector u . Which one is the detected solution depends on the initial point u and, perhaps more significantly, on the size of the scaling factor. This and subsequent

s	Sol.	# Its.
0.1	S_1	75
0.2	S_1	35
0.3	S_1	20
0.4	S_1	13

s	Sol.	# Its.
0.5	S_1	6
0.6	S_1	12
0.7	S_1	16
0.8	S_1	37

s	Sol.	# Its.
0.9	S_1	82
1.0	S_2	2
1.1	S_3	5
1.2	S_3	4

s	Sol.	# Its.
1.3	S_3	4
1.4	S_3	4
1.5 to 10	S_3	3

Table 2 – Influence of the scaling factor s on the detected solution and on the number of iterations required by the SPA to achieve convergence. The matrix A under consideration is (5) and the initial point is $u = (0, 1)$.

tests show that no changes in the detected solution and in the number of iterations occur when the parameter s is taken beyond a certain threshold value.

Observe that the solutions S_1 and S_3 were found by several of the values assumed by s , whereas the solution S_2 was detected only with the specific choice $s = 1$. In short, some Pareto eigenvalues are more likely to be found than others.

Remark 1. For a given u , the detected solution and the number of iterations required by the SPA to achieve convergence may be very sensitive with respect to the parameter s . In Table 2 one observes substantial changes while passing from $s = 0.9$ to $s = 1.1$.

2.1.2 Testing on a 3×3 matrix with 9 Pareto eigenvalues

We now test the SPA on the matrix

$$A = \begin{bmatrix} 8 & -1 & 4 \\ 3 & 4 & 1/2 \\ 2 & -1/2 & 6 \end{bmatrix} \tag{6}$$

whose Pareto spectrum is displayed in Table 3. This specific matrix has 9 Pareto eigenvalues and, probably, this is the largest possible cardinality for the Pareto

Sol. name	λ	x_1	x_2	x_3	y_1	y_2	y_3
S_1	4.133975	0	1	0.267949	0.071797	0	0
S_2	4.602084	0.185257	1	0.092628	0	0	0
S_3	5	0.333333	1	0	0	0	0.166667
S_4	5.866025	0	0.267949	1	3.732051	0	0
S_5	6	0	0	1	4	0.5	0
S_6	7	1	1	0	0	0	1.5
S_7	8	1	0	0	0	3	2
S_8	9.397916	1	0.602084	0.5	0	0	0
S_9	10	1	0	0.5	0	3.25	0

Table 3 – Solution set of the Pareto eigenproblem associated with matrix (6).

spectrum of a 3×3 matrix. We know for sure that a 3×3 matrix cannot have 11 or more Pareto eigenvalues. We have not found yet a 3×3 matrix with 10 Pareto eigenvalues and seriously doubt that such a matrix exists.

The challenge that (6) offers to the SPA is not the “large” number of Pareto eigenvalues, but the fact that some of them are hard to detect numerically. In Table 4 and in the sequel, the notation S_j/m indicates that the solution S_j was found within m iterations and the symbol ∞ indicates that convergence did not occur within 2000 iterations. Some of the conclusions that can be drawn from Table 4 are:

- i) Given an initial point u , it is possible to recover several solutions to the Pareto eigenvalue problem by letting the scaling factor s to assume different values. One should not expect however to recover all the solutions (unless of course one changes also the initial point).
- ii) Given an initial point u , there are values of s for which the SPA does not converge. It is very difficult to identify the convergence region

$$S(u) \equiv \text{values of } s \text{ for which SPA converges when initialized at } u$$

for the scaling factor, and to see what kind of structure such a region possesses. Table 4 shows that $S(u)$ is not necessarily an interval.

	Scaling factor s				
u	0.3	0.4	0.5	0.6	0.7
(0, 1, 0)	$S_2/24$	$S_3/186$	$S_1/316$	∞	∞
(0, 1, 1)	$S_1/26$	$S_1/18$	$S_1/12$	$S_1/8$	$S_1/12$
(1, 1, 1)	$S_1/28$	$S_1/19$	$S_1/13$	$S_1/8$	$S_1/13$
u	0.8 \rightarrow 1.6	1.7	1.8	1.9	2 \rightarrow 10
(0, 1, 0)	∞	$S_9/4$	$S_9/4$	$S_9/3$	$S_9/3$
(0, 1, 1)	∞	$S_9/10$	$S_5/14$	$S_5/6$	$S_9/4$
(1, 1, 1)	∞	$S_9/11$	$S_5/15$	$S_5/7$	$S_9/5$

Table 4 – Influence of the initial point u and the scaling factor s on the detected solution and on the number of iterations required by the SPA to achieve convergence. The matrix under consideration is (6).

2.1.3 Testing on random matrices and random initial points

Table 5 has been produced as follows. For each dimension $n \in \{2, 3, \dots, 9\}$, one generates a collection of 1000 stochastically independent random matrices with uniform distribution on the hypercube $[-1, 1]^{n \times n}$. For each one of these 1000 random matrices, one generates a collection of 100 stochastically independent random initial vectors with uniform distribution on the hypercube $[0, 1]^n$. One applies the SPA to each one of the $10^5 = 1000 \times 100$ random pairs (A, u) and counts the number of times in which convergence occurs within 2000 iterations.

As one can see from Table 5, the best performances of the SPA are obtained when the scaling factor s lies in the interval $[0.1, 0.4]$. When s is taken in such range, the percentages of convergence are all above 90 per cent. This good news speaks favorably of the SPA. The same experience as in Table 5 is carried out for $n \in \{10, 20, 30, 100\}$, but with a different set of values for the scaling factor s . Table 6 displays the obtained results. One sees that the overall performance of the SPA diminishes as the dimension n gets larger. For $n = 100$ the best score of convergence is 78%. This is still acceptable.

Table 7 reports on the case $n \in \{200, 300, 400, 500, 1000\}$. In this range of dimensions, the performance of the SPA is a bit less satisfactory. For instance, when $n = 1000$ the best score of convergence is 62%.

n	Scaling factor s							
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
2	100	100	100	100	100	100	100	100
3	99	99	99	99	98	99	99	99
4	98	98	97	98	97	97	95	94
5	97	97	96	96	95	96	93	89
6	94	95	94	94	95	91	89	84
7	94	93	95	94	93	90	85	75
8	94	92	91	93	90	86	79	63
9	93	93	93	91	89	84	68	52

Table 5 – SPA applied to a sample of 10^5 random pairs (A, u) . The figures on the table refer to percentages of convergence. The best performances for each $n \in \{2, 3, \dots, 9\}$ are indicated with a bold font.

n	Scaling factor s							
	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.2
10	33	51	70	85	90	93	91	91
20	44	66	80	86	87	89	88	87
30	53	72	85	83	85	87	86	80
100	75	76	77	76	78	76	78	64

Table 6 – SPA applied to a sample of 10^5 random pairs (A, u) . The figures on the table refer to percentages of convergence. The best performances for each $n \in \{10, 20, 30, 100\}$ are indicated with a bold font.

n	Scaling factor s							
	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
200	66	72	70	71	72	72	69	69
300	65	70	69	67	71	65	63	64
400	66	66	65	67	62	62	60	54
500	64	67	63	63	64	61	57	50
1000	58	62	60	54	45	×	×	×

Table 7 – SPA applied to a sample of 10^5 random pairs (A, u) . The figures on the table refer to percentages of convergence. The best performances for each $n \in \{200, 300, 400, 500, 1000\}$ are indicated with a bold font. The cases indicated with a cross were not tested.

2.2 Applying the SPA to a circular eigenvalue problem

A circular eigenvalue problem is a prototype of eigenvalue problem with constraints described by a nonpolyhedral cone:

Problem 4.

Find $\lambda \in \mathbb{R}$ and a nonzero vector $x \in \mathbb{R}^n$ such that

$$L_\theta \ni x \perp (Ax - \lambda x) \in L_\theta^+. \tag{7}$$

Here, L_θ stands for the n -dimensional circular cone whose half-aperture angle is $\theta \in]0, \pi/2[$, and whose revolution axis is the ray generated by the canonical vector $e_n = (0, \dots, 0, 1)^T$. In short,

$$L_\theta = \{x \in \mathbb{R}^n : (\cos \theta)\|x\| \leq x_n\}. \tag{8}$$

A circular eigenvalue problem differs substantially from a Pareto eigenvalue problem, the main difference being that a circular spectrum

$$\sigma_\theta(A) = \{\lambda \in \mathbb{R} : (x, \lambda) \text{ solves Problem 4}\}$$

does not need to be finite. As shown in [16], the set $\sigma_\theta(A)$ could perfectly well contain one or more intervals of positive length.

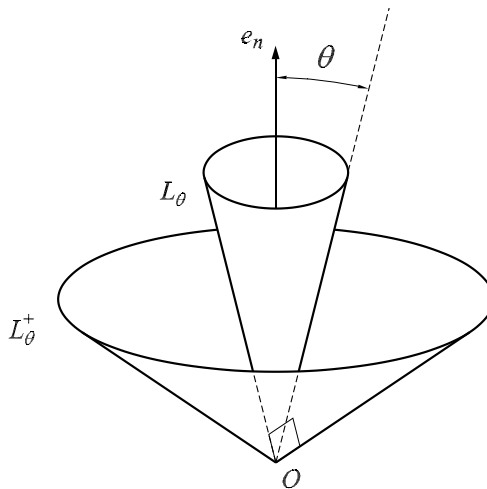


Figure 1 – The circular cone L_θ and its dual L_θ^+ .

As seen in Figure 1, the dual cone of L_θ is again a circular cone. In fact, as reported in [3], one has the general relation $L_\theta^+ = L_{(\pi/2)-\theta}$. The projection of $z \in \mathbb{R}^n$ onto L_θ can be evaluated by using the explicit formula

$$\Pi_{L_\theta}[z] = \begin{cases} z & \text{if } z \in L_\theta \\ 0 & \text{if } z \in -L_\theta^+ \\ v & \text{if } z \in \mathbb{R}^n \setminus (L_\theta \cup -L_\theta^+) \end{cases}$$

with

$$v_i = z_i - \frac{\sqrt{\|z\|^2 - z_n^2} - z_n \tan \theta}{1 + [\tan \theta]^2} \frac{z_i}{\sqrt{\|z\|^2 - z_n^2}} \quad \forall i \in \{1, \dots, n - 1\},$$

$$v_n = z_n + \frac{\sqrt{\|z\|^2 - z_n^2} - z_n \tan \theta}{1 + [\tan \theta]^2} \tan \theta.$$

The above expression of v can be found in [1] or obtained simply by working out the Karush-Kuhn-Tucker optimality conditions for the convex minimization problem

$$\begin{aligned} &\text{minimize} \quad \|x - z\|^2 \\ &x \in \mathbb{R}^n \text{ s.t. } (\cos \theta)\|x\| \leq x_n. \end{aligned}$$

For dealing with the circular eigenvalue problem, one uses

$$\phi(x) = x_n \tag{9}$$

as normalization function. This corresponds to the linear choice (3) with vector $e = e_n$. Note that e_n belongs to the interior of L_θ^+ .

A convex cone whose half-aperture angle is 90 degrees is no longer pointed, so the SPA implemented with the linear normalizing function (9) could run into numerical troubles if θ increases beyond a certain threshold.

We are interested in examining the performance of the SPA as function of the half-aperture angle θ of the circular cone. The numerical experiment reported in Table 8 concerns a sample of 10^5 random pairs (A, u) generated in the same way as in Tables 5, 6, and 7. The SPA is tested with three different values of s and with values of θ incremented in steps of 10 degrees.

Some of the conclusions that can be drawn from Table 8 are:

θ	Scaling factor s		
	0.05	0.1	1
5	100	100	100
15	100	100	100
25	99	99	99
35	94	99	96
45	88	97	93
55	78	93	86
65	69	89	83
75	63	83	78
85	57	80	74

θ	Scaling factor s		
	0.05	0.1	1
5	100	100	100
15	100	100	100
25	100	100	100
35	100	100	63
45	100	100	1
55	98	99	0
65	91	91	0
75	74	73	0
85	40	54	0

θ	Scaling factor s		
	0.05	0.1	1
5	100	100	100
15	100	100	100
25	100	100	100
35	100	100	53
45	100	100	0
55	100	100	0
65	93	95	0
75	72	74	0
85	39	50	0

Table 8 – SPA applied to a sample of 10^5 random pairs (A, u) . One considers the cases $n = 3$ (left), $n = 50$ (middle), and $n = 100$ (right). The figures on the tables refer to percentages of convergence.

- i) For each tested scaling factor s and dimension n , one observes that the rate of success of the SPA decreases as the half-aperture angle increases. The best results are obtained for circular cones with moderate half-aperture angle. As predicted by common sense, dealing with circular cones with half-aperture angle near 90 degrees is problematic.
- ii) As we already saw in previous numerical experiences, a right choice of scaling factor is crucial for obtaining a good performance of the SPA. As a general rule, when dealing with large dimensional problems one should reduce the size of the parameter s . As one can see in the middle table and in the right table, the choice $s = 1$ is appropriate only if the half-aperture angle does not exceed 25 degrees. Beyond this angular threshold, the performance of the SPA deteriorates very quickly. By contrast, if one uses a smaller scaling factor, say $s = 0.1$, then the SPA works perfectly well in a wider range of half-aperture angles.

3 Theory and numerical experience with the PIM

The power iteration method is a well known algorithm for finding the largest eigenvalue in modulus of a given matrix. To fix the ideas, consider the case of

an $n \times n$ real matrix M with real eigenvalues $\lambda_1, \dots, \lambda_n$ such that

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|.$$

The classical power iteration method consists in generating a sequence $\{x^t\}_{t \geq 0}$ by selecting an initial point x^0 and applying the recursive formula

$$x^{t+1} = \frac{Mx^t}{\phi(Mx^t)}, \quad (10)$$

where ϕ is typically the Euclidean norm $\|\cdot\|$ or a certain linear function. One expects the sequence $\{x^t\}_{t \geq 0}$ to converge to an eigenvector u_1 associated to the eigenvalue λ_1 . See, for instance, Section 13.4 in Schatzman's book [14] for a precise formulation of the convergence result. The limit u_1 satisfies necessarily the normalization condition $\phi(u_1) = 1$. The eigenvalue λ_1 is obtained by using either the relation $\lambda_1 = \phi(Mu_1)$ or the Rayleigh-Ritz ratio $\lambda_1 = \langle u_1, Au_1 \rangle / \|u_1\|^2$.

Remark 2. If one wishes to find a different eigenvalue, then one chooses a convenient shifting parameter $\beta \in \mathbb{R}$ and applies the iteration scheme (10) to the shifted matrix $M - \beta I_n$. By the way, the so-called inverse power iteration method consists in applying (10) to the inverse of the matrix $M - \beta I_n$.

It is not difficult to adjust the power iteration method to the case of a generalized eigenvalue problem involving a pair (A, B) of matrices. However, incorporating a cone K into the picture is a matter that requires a more careful thinking. For the sake of clarity in the discussion, we leave aside the matrix B and concentrate on K . The problem at hand reads as follows.

Problem 5.

Find $\lambda \in \mathbb{R}$ and a nonzero vector $x \in \mathbb{R}^n$ such that

$$K \ni x \perp (Ax - \lambda x) \in K^+.$$

In principle, the usual spectrum of A has very little to do with

$$\sigma_K(A) = \{\lambda \in \mathbb{R} : (x, \lambda) \text{ solves Problem 5}\},$$

the latter set being called the K -spectrum (or set of K -eigenvalues) of A . However, regardless of the choice of K , one always has the inclusion

$$\sigma_K(A) \subset [\lambda_{\min}(A_{\text{sym}}), \lambda_{\max}(A_{\text{sym}})], \tag{11}$$

with $\lambda_{\min}(A_{\text{sym}})$ and $\lambda_{\max}(A_{\text{sym}})$ denoting, respectively, the smallest and the largest eigenvalue of the symmetric part

$$A_{\text{sym}} = \frac{A + A^T}{2}$$

of the matrix A . The computation of the extremal eigenvalues of a symmetric matrix presents no difficulty.

Lemma 1. *Consider any $\beta \in \mathbb{R}$ such that*

$$\beta > \lambda_{\max}(A_{\text{sym}}). \tag{12}$$

Then, the following conditions are equivalent:

- (a) x is a K -eigenvector of A .
- (b) x and $\Pi_K[\beta x - Ax]$ are nonzero vectors pointing in the same direction.

Proof. For all $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^n$, one has the general identity

$$\beta x - Ax = (\beta - \lambda)x - (Ax - \lambda x).$$

This relation is at the core of the proof. Let x be a K -eigenvector of A and let λ be the associated K -eigenvalue. In particular, x is a nonzero vector in K . Furthermore, the combination of (11) and (12) shows that $\gamma = \beta - \lambda$ is a positive scalar. By applying Moreau's theorem [8] to the orthogonal decomposition

$$\beta x - Ax = \underbrace{\gamma x}_{\in K} - \underbrace{Ax - \lambda x}_{\in K^+},$$

one obtains

$$\gamma x = \Pi_K[\beta x - Ax]. \tag{13}$$

This confirms that x and $\Pi_K[\beta x - Ax]$ are nonzero vectors pointing in the same direction. Conversely, suppose that $x \neq 0$ and that (13) holds for some positive scalar γ . Then, one can draw the following conclusions. Firstly, $x \in K$ because $\gamma x \in K$ and $\gamma > 0$. Secondly, in view of Moreau's theorem, the equality (13) yields

$$\begin{aligned}\gamma x - (\beta x - Ax) &\in K^+, \\ \langle \gamma x, \gamma x - (\beta x - Ax) \rangle &= 0.\end{aligned}$$

We have proven in this way that x is a nonzero vector satisfying the complementarity system

$$K \ni x \perp (Ax - (\beta - \gamma)x) \in K^+.$$

In other words, x is a K -eigenvector of A and $\lambda = \beta - \gamma$ is the associated K -eigenvalue. \square

Condition (b) in Lemma 1 provides the inspiration for the formulation of the PIM in the context of a cone-constrained eigenvalue problem. The detailed description of this algorithm is as follows:

- *Initialization:* Select a shifting parameter β as in (12). Pick a nonzero vector $u \in K$ and let $x^0 = u/\phi(u)$.
- *Iteration:* One has a current point x^t in K_ϕ . Compute first the projection

$$v^t = \Pi_K[\beta x^t - Ax^t], \quad (14)$$

and then the ϕ -normalized vector

$$x^{t+1} = v^t/\phi(v^t). \quad (15)$$

As was the case with the SPA, a sequence generated by the PIM remains in the compact set K_ϕ . The emphasis of the PIM is the computation of K -eigenvectors. If one wishes, at each iteration one can also evaluate the Rayleigh-Ritz ratio

$$\lambda_t = \langle x^t, Ax^t \rangle / \|x^t\|^2,$$

but this is not really necessary.

Theorem 2. *Let $\{x^t\}_{t \geq 0}$ be generated by the PIM.*

- (a) *The general term x^t is well defined, i.e., the projection v^t is different from 0.*
- (b) *If $\{x^t\}_{t \geq 0}$ converges to a certain \bar{x} , then \bar{x} is a K -eigenvector of A .*
- (c) *The K -eigenvalue of A associated to \bar{x} is given by*

$$\bar{\lambda} = \langle \bar{x}, A\bar{x} \rangle / \|\bar{x}\|^2 = \beta - \phi(\Pi_K[\beta\bar{x} - A\bar{x}]). \tag{16}$$

Proof. Suppose, on the contrary, that $v^t = 0$ at some iteration t . From the usual characterization of a projection, one has

$$\langle \beta x^t - Ax^t, w \rangle \leq 0 \quad \forall w \in K.$$

In particular, $\langle \beta x^t - Ax^t, x^t \rangle \leq 0$. Since $x^t \neq 0$, one obtains

$$\beta \leq \frac{\langle x^t, Ax^t \rangle}{\|x^t\|^2} \leq \lambda_{\max}(A_{\text{sym}}),$$

a contradiction with assumption (12). This takes care of (a). Suppose now that $\{x^t\}_{t \geq 0}$ converges to \bar{x} . By passing to the limit in (14) one sees that $\{v^t\}_{t \geq 0}$ converges to

$$\bar{v} = \Pi_K[\beta\bar{x} - A\bar{x}].$$

As done before with each term v^t , one can also prove that \bar{v} is different from 0. By passing now to the limit in (15), one gets

$$\bar{x} = \bar{v} / \phi(\bar{v}).$$

Since $\bar{\gamma} = \phi(\bar{v})$ is a positive scalar, the nonzero vectors \bar{x} and \bar{v} are pointing in the same direction. In view of Lemma 1, one concludes that \bar{x} is a K -eigenvector of A . The first equality in (16) is obvious, while the second one is implicit in the proof of Lemma 1. □

The inequality (12) has been helpful at several stages. Lemma 1 and Theorem 2 remain true if one works under the weaker assumption

$$\beta > \sup_{\substack{x \in K \\ \|x\|=1}} \langle x, Ax \rangle,$$

but the evaluation of the above supremum could be a bothersome task. From a purely theoretical point of view, one could even work under the assumption

$$\beta > \sup_{\lambda \in \sigma_K(A)} \lambda. \tag{17}$$

In practice, it makes no sense relying on (17) because $\sigma_K(A)$ is unknown a priori.

3.1 Applying the PIM to the Pareto eigenvalue problem

3.1.1 Testing on a small size matrix

We start by testing the PIM on the matrix (6). Recall that (6) is a nonsymmetric 3×3 matrix with 9 Pareto eigenvalues. All of them are positive and the largest one is $\lambda = 10$. A straightforward computation shows that

$$A_{\text{sym}} = \begin{bmatrix} 8 & 1 & 3 \\ 1 & 4 & 0 \\ 3 & 0 & 6 \end{bmatrix}, \quad \lambda_{\max}(A_{\text{sym}}) = 10.2682.$$

The symbol ∞ in Table 9 means, as usual, that convergence does not occur within 2000 iterations. Observe that number of iterations required by the PIM to achieve convergence increases as β goes further and further away from the threshold value (17). So, it is not a good idea to take β excessively large.

u	Shifting parameter β								
	5	6	7	8	9	10	11	20	50
(0, 1, 0)	E	∞	∞	$S_2/22$	$S_2/9$	$S_2/10$	$S_2/14$	$S_2/47$	$S_2/143$
(0, 1, 1)	E	$S_1/8$	$S_1/21$	$S_1/31$	$S_1/41$	$S_1/50$	$S_1/60$	$S_1/140$	$S_1/391$
(1, 1, 1)	E	E	E	$S_1/32$	$S_1/42$	$S_1/51$	$S_1/61$	$S_1/143$	$S_1/397$

Table 9 – Influence of the initial point u and the parameter β on the detected solution and on the number of iterations required by the PIM to achieve convergence. The matrix under consideration is (6).

On the other hand, the symbol E in Table 9 indicates that an error occurs while performing the first iteration of the PIM. To be more precise, the PIM must stop because projecting onto K produces the zero vector. This unfavorable

situation does not contradict the statement (a) of Theorem 2. What happens is simply that the value of β under consideration does not satisfy the assumption (12), not even the weaker condition (17). In connection with this issue, it is worthwhile mentioning that if one uses a parameter β smaller than the right-hand side of (17), then the PIM may still work in the sense that each projection v^t may be different from 0 and $\{x^t\}$ may converge to a K -eigenvector of A . Logically speaking, (17) is a sufficient condition for the PIM to work properly, but such condition is not necessary.

In Table 10 one applies the PIM to the matrix (6), but one considers a sample of 10^8 random initial points. The question that is being addressed is the following one: to which Pareto eigenvalue of (6) is the PIM most likely to converge?

β	Obtained solution										% of success
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	None	
10	48.6	0	11.3	0	26.4	0	13.7	0	0	0	100.0
15	52.4	0	10.6	0	22.6	0	14.4	0	0	0	100.0
20	47.1	0	10.4	0	21.8	0	14.6	0	0	6.1	93.9
25	47.2	0	10.4	0	21.4	0	14.6	0	0	6.5	93.5

Table 10 – PIM applied to the matrix (6). The figures on the table refer to percentages of convergence to each particular solution. A sample of 10^8 random initial points is being considered.

3.1.2 Testing on random matrices and random initial points

The same kind of numerical experiment as in Section 2.1.3 is now carried out with the PIM. To avoid excessive repetitions, we consider only the case of a dimension n ranging over $\{3, 6, 12\}$.

Since one works with a sample of random pairs (A, u) , it is not entirely clear how to choose the parameter β in order to optimize the performance of the PIM. In Table 11 we explore what happens with the choices

$$\beta_k = \lambda_{\max}(A_{\text{sym}}) + (k/2) - 1 \quad \text{with } k = 1, 2, 3, 4. \tag{18}$$

Note that β_1 is smaller than $\lambda_{\max}(A_{\text{sym}})$. The heuristic selection rule (18) may look strange at first sight, but it is based on statistical information concerning

the distribution of the random variable $\lambda_{\max}(A_{\text{sym}})$ when $n \in \{3, 6, 12\}$.

n	Shifting parameter β			
	β_1	β_2	β_3	β_4
3	81	99	98	98
6	92	94	95	95
12	89	91	89	91

Table 11 – PIM applied to a sample of 10^5 random pairs (A, u) . The best performances for each dimension $n \in \{3, 6, 12\}$ are indicated with a bold font.

The figures on Table 11 refer to percentages of success. The word “success” means, of course, that the PIM has worked properly, failure corresponding to a case of lack of convergence or to the obtention of a projection equal to 0. As one can see, the percentages of success of the PIM do not change too much when β ranges over (18).

In Figure 2 we enlarge the field of values taken by the parameter β . The idea is identifying the largest interval over which β leads to a favorable and stable rate of success.

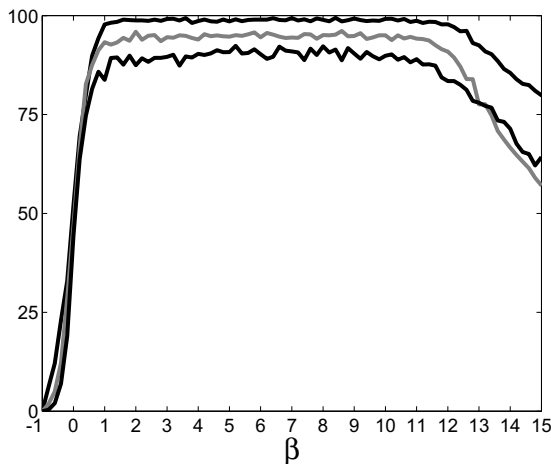


Figure 2 – PIM applied to a sample of 10^5 random pairs (A, u) . The parameter β is increased in steps of 0.2 units from -1 to 15 . The upper, intermediate, and lower curves correspond to the cases $n = 3$, $n = 6$, and $n = 12$, respectively. The vertical axis indicates percentage of success.

One observes a very interesting plateau phenomenon: for each dimension n , there is a corresponding range $[\beta_n^{\text{lower}}, \beta_n^{\text{upper}}]$ of values for β in which the PIM performs at its top level. The height of each plateau depends on the dimension: the height decreases as n increases. Note also that the interval $[\beta_n^{\text{lower}}, \beta_n^{\text{upper}}]$ is almost identical for the three tested values of n .

4 SPA versus PIM

From a conceptual point of view, the SPA and the PIM are different methods. However, they share in common at least two important steps: a projection onto the cone K and a ϕ -normalization procedure.

Are the SPA and the PIM two different algorithms after all? The answer is yes if the scaling factor s and the shifting parameter β are not allowed to change from one iteration to the next. If one is more flexible and considers instead a sequence $\{s_t\}_{t \geq 0}$ of scaling factors and a sequence $\{\beta_t\}_{t \geq 0}$ of shifting parameters, then one can reformulate the SPA as a PIM, and viceversa.

Proposition 1. *Let $B = I_n$. The following statements are true:*

- (a) *Implementing the t -th iteration of the PIM with $\beta_t > \langle x^t, Ax^t \rangle / \|x^t\|^2$ produces the same outcome as implementing the t -th iteration of the SPA with*

$$s_t = \left[\beta_t - \frac{\langle x^t, Ax^t \rangle}{\|x^t\|^2} \right]^{-1}. \tag{19}$$

- (b) *Conversely, implementing the t -th iteration of the SPA with $s_t > 0$ produces the same outcome as implementing the t -th iteration of the PIM with*

$$\beta_t = \frac{1}{s_t} + \frac{\langle x^t, Ax^t \rangle}{\|x^t\|^2}. \tag{20}$$

Proof. The projection map Π_K and the normalizing function ϕ are both positively homogeneous. Hence, ϕ -normalizing the projected vector $\Pi_K[x^t - s_t y^t]$ produces the same result as ϕ -normalizing the projected vector $\Pi_K[(1/s_t)x^t - y^t]$. It suffices then to observe that

$$\begin{aligned} (1/s_t)x^t - y^t = \beta_t x^t - Ax^t &\iff (1/s_t)x^t - (Ax^t - \lambda_t x^t) = \beta_t x^t - Ax^t \\ &\iff (1/s_t) + \lambda_t = \beta_t \end{aligned} \tag{21}$$

with $\lambda_t = \langle x^t, Ax^t \rangle / \|x^t\|^2$. Of course, one rewrites the equality (21) in the form (19) for proving part (a), and in the form (20) for proving part (b). \square

4.1 Numerical experience with a sequence $\{s_t\}_{t \geq 0}$

We have tested again the SPA on random matrices and random initial points, but this time we have used a variable scaling factor s_t . The attention is focused on dimensions not exceeding $n = 30$. Inspired by the relation (19) and the experimental data reported in Figure 2, we choose a scaling factor s_t obeying to the feedback rule

$$s_t = \left[\beta - \frac{\langle x^t, Ax^t \rangle}{\|x^t\|^2} \right]^{-1} \quad \text{with} \quad \beta = 6. \tag{22}$$

The word “feedback” is used here to indicate that s_t is updated by taking in account the current value x^t of the state vector x .

Table 12 reports on the outcome of this experiment. The most important lesson of this experiment is the following one: the SPA implemented with the feedback rule performs equally well as the SPA implemented with the constant scaling factor s that is optimal for the prescribed dimension. For instance, when $n = 9$, the feedback rule leads to a rate of success of 93%, a percentage that matches the rate of success of the SPA implemented with the optimal choice $s = 0.1$ (cf. Table 5).

n	3	6	9	12	15	18	21	24	27	30
%	99	95	93	90	89	89	87	88	85	86

Table 12 – For each $n \in \{3, 6, 9, \dots, 30\}$, the SPA was applied to a sample of 10^5 random pairs (A, u) . The variable scaling factor used was (22). Figures correspond to percentages of convergence.

5 By way of conclusion

Our main comments on the SPA and the PIM are as follows:

1. The SPA generates a bounded sequence $\{(x^t, \lambda_t, y^t)\}_{t \geq 0}$ that depends on a certain initial point u and a certain scaling parameter $s > 0$. If the sequence converges, then its limit $(\bar{x}, \bar{\lambda}, \bar{y})$ is a solution to Problem 3.

2. The SPA does not aim at finding all the solutions to Problem 3. This is not just because the number of solutions could be very large, but also because some solutions could be extremely hard to be detected. It happens in practice that some solutions are hard to be found even if the SPA is initialized at a convenient initial point. The existence of such “rare” solutions is somewhat intrinsic to the nature of the cone-constrained eigenvalue problem. An interesting open question is understanding why some solutions are so hard to be detected and others are not. An appropriate theoretical analysis remains to be done.
3. The SPA was extensively tested in the Paretian case. One of the main conclusions of Section 2.1 is that the scaling factor $s > 0$ must be chosen with care if one wishes the SPA to produce a convergent sequence. Extensive numerical experimentation has been carried out for Pareto eigenvalue problems with random data. Specifically, the entries of the matrix A and the entries of the initial vector u were considered as independent random variables uniformly distributed over the interval $[-1, 1]$. For each dimension n , there is a convenient range of values for the scaling factor s . A good choice of s is one for which the SPA converges in a vast majority of cases, see Tables 5, 6, and 7.
4. The PIM has the advantage of relying on a parameter β that is not too sensitive with respect to dimensionality. Thanks to the plateau phenomenon described in Section 3.1.2, there is no need of putting too much effort in selecting the shifting parameter β . One just needs to make sure that β lies in a certain range that is easily identifiable.

We end this work by mentioning an open alternative. Observe that Problem 3 is about finding a triplet (x, λ, y) that lies in the closed convex cone $Q = K \times \mathbb{R} \times K^+$, and that solves the system of nonlinear equations

$$\begin{aligned} Ax - \lambda Bx - y &= 0, \\ \langle x, y \rangle &= 0, \\ \phi(x) - 1 &= 0. \end{aligned}$$

In principle, one could apply any method for solving systems of nonlinear equations on closed convex cones. Do such methods exist in the literature? Are they efficient? These questions open the way to a broad discussion that deserves a detailed treatment in a separate publication.

A natural idea that comes to mind is finding a global minimum over Q of the residual term

$$f(x, \lambda, y) = \mu_1 \|Ax - \lambda Bx - y\|^2 + \mu_2 \langle x, y \rangle^2 + [\phi(x) - 1]^2,$$

where μ_i are positive scalars interpreted as weight coefficients. For instance, Friedlander et al., [2] use a residual minimization technique for solving a so-called horizontal linear complementarity problem. However, in our particular context, one should not be too optimistic about residual minimization. Indeed, this technique may produce a local minimum that is not a global one.

Last but not the least, our original motivation was solving a very specific cone-constrained eigenvalue problem arising in mechanics (namely, detecting directional instabilities in elastic systems with frictional contact). The problem at hand is highly structured and will be treated in full extent in a forthcoming work. However, we want to mention here some features that could be of interest for non-specialists:

- Firstly, the cone defining the constraints is a Cartesian product $K = K_1 \times \dots \times K_N$ of a large number of closed convex cones living in small dimensional Euclidean spaces. As a consequence of this, the vector $x = (x_{[1]}, \dots, x_{[N]})$ breaks into N blocks and Problem 1 leads to a collection of subproblems (coupled, in general).
- Secondly, some of the blocks $x_{[r]}$ are unconstrained, i.e., $K_r = \mathbb{R}^{n_r}$. This means that K is unpointed, a situation that is preferable to avoid. One way to overcome this difficulty is by applying our algorithms to the constrained subproblems and by treating the unconstrained ones as in classical numerical linear algebra. Of course, all this requires some necessary adjustments.
- Thirdly, the constrained blocks $x_{[r]}$ obey to the unilateral contact law and the Coulomb friction law. Typically, K_r is a closed convex cone dependent

on the coefficient of friction. It is also possible to accommodate the more general case of an anisotropic friction law.

Specially structured matrices are not reflected by the random data used in this paper. Our primary concern has been focussing in problems for which the matrix A has no structure whatsoever.

Remark 3. The results presented in Tables 6 and 7 were produced with the IST Cluster, which is a Cluster 1350 computing system from IBM. It has a total of 70 computing nodes each with 2 POWER dual core processors at 2.3 GHz, with 2GB per core. All the nodes are running the AIX 5.3 operating system. The parallel processes were connected using library MPI (Message Passing Interface). The programs developed during this study were written in FORTRAN.

Acknowledgments. We thank Prof. Moitinho de Almeida (Departamento de Engenharia Civil e Arquitectura, IST, Portugal) for the parallelization of one of the programs developed for this study.

REFERENCES

- [1] H.H. Bauschke and S.G. Kruk, *Reflection-projection method for convex feasibility problems with an obtuse cone*. J. Optim. Theory Appl., **120** (2004), 503–531.
- [2] A. Friedlander, J.M. Martínez and S.A. Santos, *Solution of linear complementarity problems using minimization with simple bounds*. J. Global Optim., **6** (1995), 253–267.
- [3] J.L. Goffin, *The relaxation method for solving systems of linear inequalities*. Math. Oper. Res., **5** (1980), 388–414.
- [4] A. Iusem and A. Seeger, *On pairs of vectors achieving the maximal angle of a convex cone*. Math. Program., **104** (2005), 501–523.
- [5] J.J. Júdice, M. Raydan, S.S. Rosa and S.A. Santos, *On the solution of the symmetric eigenvalue complementarity problem by the spectral projected gradient algorithm*. Numer. Algor., **47** (2008), 391–407.
- [6] J.J. Júdice, I.M. Ribeiro and H.D. Sherali, *The eigenvalue complementarity problem*. Comput. Optim. Appl., **37** (2007), 139–156.
- [7] P. Lavelledieu and A. Seeger, *Existence de valeurs propres pour les systèmes multivoques: résultats anciens et nouveaux*. Annal. Sci. Math. Québec, **25** (2000), 47–70.
- [8] J.J. Moreau, *Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires*. C.R. Acad. Sci. Paris, **255** (1962), 238–240.

- [9] A. Pinto da Costa, I.N. Figueiredo, J.J. Júdice and J.A.C. Martins, *A complementarity eigenproblem in the stability of finite dimensional elastic systems with frictional contact*. Complementarity: Applications, Algorithms and Extensions, 67–83. Applied Optimization Series, vol. 50. Eds. M. Ferris, O. Mangasarian and J.S. Pang, Kluwer Acad. Publ., Dordrecht. (1999).
- [10] A. Pinto da Costa, J.A.C. Martins, I.N. Figueiredo and J.J. Júdice, *The directional instability problem in systems with frictional contacts*. Comput. Methods Appl. Mech. Engrg., **193** (2004), 357–384.
- [11] A. Pinto da Costa and A. Seeger, *Cone-constrained eigenvalue problems: theory and algorithms*. Comput. Optim. Appl., **43/44** (2009), in press.
- [12] M. Queiroz, J.J. Júdice and C. Humes, *The symmetric eigenvalue complementarity problem*. Math. Comp., **73** (2004), 1849–1863.
- [13] P. Quittner, *Spectral analysis of variational inequalities*. Comm. Math. Univ. Carolinae, **27** (1986), 605–629.
- [14] M. Schatzman, *Analyse Numérique*, Dunod, Paris. (2001).
- [15] A. Seeger, *Eigenvalue analysis of equilibrium processes defined by linear complementarity conditions*. Linear Algebra Appl., **292** (1999), 1–14.
- [16] A. Seeger and M. Torki, *On eigenvalues induced by a cone constraint*. Linear Algebra Appl., **372** (2003), 181–206.