TOP (2016) 24:66–87
DOI 10.1007/s11750-015-0382-z

ORIGINAL PAPER

Critical angles between two convex cones II. Special cases

Alberto Seeger[1] · **David Sossa**[2]

Received: 27 March 2014 / Accepted: 23 March 2015 / Published online: 3 April 2015

Abstract The concept of critical angle between two linear subspaces has applications in statistics, numerical linear algebra and other areas. Such concept has been abundantly studied in the literature. Part I of this work is an attempt to build up a theory of critical angles for a pair of closed convex cones. The need of such theory is motivated, among other reasons, by some specific problems arising in regression analysis of cone-constrained data, see Tenenhaus in (Psychometrika 53:503–524, 1988). Angle maximization and/or angle minimization problems involving a pair of convex cones are at the core of our discussion. Such optimization problems are nonconvex in general and their numerical resolution offers a number of challenges. Part II of this work focusses on the practical computation of the maximal angle between specially structured cones.

Keywords Nonconvex optimization · Maximal angle · Critical angle · Convex cones · Topheavy cones · Ellipsoidal cones · Cones of matrices

Mathematics Subject Classification 15A18 · 15A48 · 52A40 · 90C26 · 90C33

D. Sossa is supported by CONICYT, Chile.

✉ Alberto Seeger
alberto.seeger@univ-avignon.fr

David Sossa
dsossa@dim.uchile.cl

1 Département de Mathématiques, Université d'Avignon, 33 rue Louis Pasteur, 84000 Avignon, France

2 Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile

1 Introduction

The concept of critical angle (or principal angle) between two linear subspaces has a long history going back to a late 19-th century work by the French mathematician Camille Jordan (cf. Jordan 1875). Hotelling (1936) used the concept of critical angle between linear subspaces for analyzing the canonical correlation between two sets of vector-valued variables. More recent applications of critical angles between linear subspaces include robust optimization (cf. Mohammadi 2014), linear stochastic models and ARMA models (cf. De Cock and De Moor 2000, 2002), pattern recognition and machine learning (cf. Shashua and Wolf 2003). Part I of this work discusses the concept of critical angle for a pair of closed convex cones. The present paper focusses the attention on the practical computation of the maximal angle between specially structured cones.

Let $(\mathbb{X}, \langle\cdot,\cdot\rangle)$ be a Euclidian space of dimension at least two and let $\mathcal{C}(\mathbb{X})$ be the set of nontrivial closed convex cones in $\mathbb{X}$. That a closed convex cone is nontrivial means that it is different from the zero cone and different from the whole space. The maximal angle of a pair (P, Q) of nontrivial closed convex cones in $\mathbb{X}$ is defined by

$$\Theta(P,Q) := \max_{u\in P\cap S_{\mathbb{X}},\, v\in Q\cap S_{\mathbb{X}}} \arccos\langle u, v\rangle, \tag{1}$$

where $S_{\mathbb{X}}$ stands for the unit sphere of $\mathbb{X}$. A pair $(u, v) \in \mathbb{X}^2$ solving the angle maximization problem (1) is called an *antipodal pair* of (P, Q). Antipodal pairs always exist, but they are not unique in general. A necessary condition for (u, v) to be an antipodal pair of (P, Q) is that

$$\begin{cases} u \in P \cap S_{\mathbb{X}}, \\ v \in Q \cap S_{\mathbb{X}}, \\ v - \langle u, v\rangle u \in P^*, \\ u - \langle u, v\rangle v \in Q^*, \end{cases} \tag{2}$$

where P^* and Q^* are the positive dual cones of P and Q, respectively. If the system (2) holds, then (u, v) is called a *critical pair* of (P, Q) and $\arccos\langle u, v\rangle$ is called a *critical angle* of (P, Q). The adjective *proper* is added to a critical pair (u, v) and the corresponding critical angle if u and v are not collinear. One refers to the set

$$\Gamma(P,Q) := \{\arccos\langle u, v\rangle : (u, v) \text{ satisfies } (2)\}$$

as the *angular spectrum* of (P, Q). Angular spectra have usually a finite cardinality, but not always.

In a similar way, one can treat the angle minimization problem

$$\Psi(P,Q) := \min_{u\in P\cap S_{\mathbb{X}},\, v\in Q\cap S_{\mathbb{X}}} \arccos\langle u, v\rangle. \tag{3}$$

Angle minimization problems like (3) arise in a number of applications, for instance in the theory of exponential dichotomies for linear ODEs (cf. Obert 1991) and in regression analysis of ordinal data (cf. Tenenhaus 1988).

Example 1.1 Let $\mathbb{S}_n$ denote the unit sphere of $\mathbb{R}^n$. As in Tenenhaus (1988, Section 2), consider two convex polyhedral cones P and Q, generated, respectively, by the sets of vectors $\{g_1, \ldots, g_p\} \subseteq \mathbb{R}^n$ and $\{h_1, \ldots, h_q\} \subseteq \mathbb{R}^n$. Canonical analysis of the two cones P and Q is the search for two vectors $u_0 \in P \cap \mathbb{S}_n$ and $v_0 \in Q \cap \mathbb{S}_n$ maximizing the square of $\langle u, v\rangle$ over $u \in P \cap \mathbb{S}_n$ and $v \in Q \cap \mathbb{S}_n$. According to the relative position of the cones P and Q, the problem is to find the smallest or the largest angle between the two cones.

Example 1.2 A frequent problem arising in the analysis of cone-constrained linear systems is to check if the kernel of a matrix A intersects nontrivially a closed convex cone Q, i.e., if the system

$$Ax = 0, \quad x \in Q \tag{4}$$

admits a nonzero solution $x \in \mathbb{X}$. By homogeneity, one may suppose that the solution x is sought on the unit sphere $S_{\mathbb{X}}$. Clearly, the problem at hand is equivalent to check if the minimal angle $\vartheta_0 := \Psi(\ker A, Q)$ is equal to zero. If ϑ_0 is positive but nearly zero, then it is possible to find a unit vector x that solves (4) within a certain tolerance level, i.e., after changing Q by a slightly larger cone.

There is no loss of generality in focussing the attention just on angle maximization. Indeed, one readily sees that

$$\Psi(P, Q) = \pi - \Theta(P, -Q). \tag{5}$$

Part I of this work (cf. Seeger and Sossa 2015) establishes various geometric and analytic results concerning antipodality, criticality and angular spectra. The present paper focusses on the computation of the maximal angle between specially structured cones. It is worthwhile stressing that (1) and (3) are nonconvex optimization problems. The organization of the paper is as follows. Section 2 discusses the case in which P and Q are revolution cones. One gives explicit formulas for computing all the critical angles. Section 3 discusses the case in which P and Q are topheavy cones. The class of topheavy cones is quite large and includes in particular the ℓ^p-cones and the ellipsoidal cones. Section 4 concerns the computation of the maximal angle between two cones of matrices. A large portion of this section is devoted to a difficult question arising in numerical linear algebra: how large can be the angle between a positive semidefinite symmetric matrix and a symmetric matrix that is nonnegative entrywise?

1.1 Preliminary material

A critical pair of (P, Q) may not solve the angle maximization problem (1). However, each component of a critical pair is a solution to a certain optimization problem. The details are explained below.

Proposition 1.3 *Let $P, Q \in \mathcal{C}(\mathbb{X})$. Then, (u, v) is a critical pair of (P, Q) if and only if*

$$\begin{cases} u \ \textit{minimizes} \ \langle \cdot, v - \langle u, v\rangle u\rangle \ \textit{on} \ P \cap S_{\mathbb{X}}, \\ v \ \textit{minimizes} \ \langle u - \langle u, v\rangle v, \cdot\rangle \ \textit{on} \ Q \cap S_{\mathbb{X}}. \end{cases} \tag{6}$$

Proof The proof is immediate. The key observation is that u is orthogonal to $v-\langle u, v\rangle u$ and that v is orthogonal to $u - \langle u, v\rangle v$. □

There are many alternative characterizations of criticality. The characterization (6) will be used later in a number of occasions. Beware that (6) is a weaker than

$$\begin{cases} u \text{ minimizes } \langle \cdot, v\rangle \text{ on } P \cap S_{\mathbb{X}}, \\ v \text{ minimizes } \langle u, \cdot\rangle \text{ on } Q \cap S_{\mathbb{X}}. \end{cases} \tag{7}$$

A pair (u, v) as in (7) is said to be a *Nash antipodal pair* of (P, Q). Nash antipodality is a property that lies between criticality and antipodality. The following easy result is recorded just for convenience. It concerns the maximal angle between a convex cone and its dual.

Proposition 1.4 *Let* $K \in \mathcal{C}(\mathbb{X})$. *Then,* $\Theta(K, K^*) = \pi/2$.

Proof Since $\langle u, v\rangle \geq 0$ for all $u \in K$ and $v \in K^*$, it is clear that

$$\Theta\left(K, K^*\right) \leq \pi/2. \tag{8}$$

Since $K \cup -K^*$ is not the whole space $\mathbb{X}$, there exists a nonzero vector $z \notin K \cup -K^*$. Let $\Pi_K(z)$ denote the projection of z onto K. Moreau's orthogonal decomposition theorem implies that $\Pi_K(z)$ and $\Pi_{K^*}(-z)$ are nonzero orthogonal vectors. Hence,

$$u := \frac{\Pi_K(z)}{\|\Pi_K(z)\|} \in K, \quad v := \frac{\Pi_K^*(-z)}{\|\Pi_K^*(-z)\|} \in K^*$$

are orthogonal unit vectors. This proves that (8) is in fact an equality. □

2 Critical angles in a pair of revolution cones

Revolutions cones, also called circular cones, are amongst the simplest and most common non-polyhedral convex cones used in mathematics. By definition, a revolution cone in $\mathbb{X}$ is a closed convex cone of the form

$$\mathrm{Rev}(\phi, b) := \{x \in \mathbb{X} : \langle b, x\rangle \geq \|x\| \cos\phi\},$$

where $b \in S_{\mathbb{X}}$ defines the revolution axis and $\phi \in [0, \pi/2]$ corresponds to the half-aperture angle. The next theorem shows that a pair of revolution cones has at most three critical angles. It provides also explicit formulas for computing each critical angle.

Theorem 2.1 *Let* $P = \mathrm{Rev}(\phi_1, b_1)$ *and* $Q = \mathrm{Rev}(\phi_2, b_2)$, *with* $b_1, b_2 \in S_{\mathbb{X}}$ *and* $\phi_1, \phi_2 \in [0, \pi/2]$. *Then*

$$\Gamma(P, Q) = \begin{cases} \{0, \alpha_1, \pi\} & \text{if } \alpha_1 \geq 0,\ \alpha_2 \geq \pi, \\ \{0, \alpha_1, \alpha_2\} & \text{if } \alpha_1 \geq 0,\ \alpha_2 < \pi, \\ \{\pi\} & \text{if } \alpha_1 < 0,\ \alpha_2 \geq \pi, \\ \{\alpha_2\} & \text{if } \alpha_1 < 0,\ \alpha_2 < \pi, \end{cases} \tag{9}$$

where

$$\alpha_1 := \phi_1 + \phi_2 - \arccos\langle b_1, b_2\rangle,$$
$$\alpha_2 := \phi_1 + \phi_2 + \arccos\langle b_1, b_2\rangle.$$

Proof If $\mathbb{X}$ is a two-dimensional space, then (9) is obtained by arguments of planar geometry. Suppose that $\mathbb{X}$ is of dimension at least three. The improper critical angles of (P, Q) are easy to identify. Indeed,

$$\begin{aligned} 0 \in \Gamma(P, Q) &\Leftrightarrow P \cap Q \neq \{0\} \quad \Leftrightarrow \quad \alpha_1 \geq 0, \\ \pi \in \Gamma(P, Q) &\Leftrightarrow P \cap -Q \neq \{0\} \quad \Leftrightarrow \quad \alpha_2 \geq \pi. \end{aligned}$$

So, one just needs to detect the proper critical angles of (P, Q). We claim that

$$\begin{cases} \text{if } (u, v) \text{ is a proper critical pair of } (P, Q), \\ \text{then } L := \operatorname{span}\{u, v\} \text{ contains } b_1 \text{ and } b_2. \end{cases} \tag{10}$$

This claim will be shown in a moment. As a consequence of (10), one gets the following planar reduction principle: (u, v) is a proper critical pair of (P, Q) if and only if (u, v) is a proper critical pair of $(P \cap L, Q \cap L)$, where

$$P \cap L := \{x \in L : \langle b_1, x\rangle \geq \|x\| \cos \phi_1\},$$
$$Q \cap L := \{x \in L : \langle b_2, x\rangle \geq \|x\| \cos \phi_2\},$$

are viewed as revolution cones in the two-dimensional space L. In other words, one is back to a planar setting. We now prove (10). Consider a proper critical pair (u, v) of (P, Q) and set $\lambda := \langle u, v\rangle$. We distinguish between two cases:
Case 1: $\phi_1, \phi_2 > 0$. This case is the most interesting one. By combining Proposition 1.3 and the boundary principle for proper critical pairs (cf. Seeger and Sossa 2015, Theorem 2.3), one knows that u and v solve

$$\begin{cases} \text{minimize } \langle x, v - \lambda u\rangle \\ \langle b_1, x\rangle = \cos \phi_1 \\ \|x\|^2 = 1 \end{cases} \tag{11}$$

and

$$\begin{cases} \text{minimize } \langle u - \lambda v, y\rangle \\ \langle b_2, y\rangle = \cos \phi_2 \\ \|y\|^2 = 1, \end{cases} \tag{12}$$

respectively. Hence, there exist Lagrange multipliers μ_1, γ_1 for the problem (11) and Lagrange multipliers μ_2, γ_2 for the problem (12), such that

$$v - \lambda u = \gamma_1 b_1 + \mu_1 u, \tag{13}$$
$$u - \lambda v = \gamma_2 b_2 + \mu_2 v. \tag{14}$$

The multiplication of (13) by u and (14) by v yields

$$\mu_1 + \gamma_1 \cos\phi_1 = 0, \quad \mu_2 + \gamma_2 \cos\phi_2 = 0,$$

respectively. Hence,

$$\begin{aligned}(\gamma_1 \cos\phi_1 - \lambda)\, u + v &= \gamma_1 b_1,\\ u + (\gamma_2 \cos\phi_2 - \lambda)\, v &= \gamma_2 b_2.\end{aligned} \tag{15}$$

Observe that $\gamma_1 \neq 0$ and $\gamma_2 \neq 0$, because (u, v) is proper. This proves that $b_1, b_2 \in L$. *Case 2*: $\phi_1 = 0, \phi_2 > 0$. In this case, $P = \vec{b}_1 := \{tb_1 : t \geq 0\}$. Hence, $u = b_1$ and $b_1 \in L$. That $b_2 \in L$ follows from (15) and the fact that $\gamma_2 \neq 0$. □

Corollary 2.2 *Let P, Q be two revolution cones as in Theorem 2.1. Then*

$$\Theta(P, Q) := \min\{\pi, \arccos\langle b_1, b_2\rangle + \phi_1 + \phi_2\}, \tag{16}$$

$$\Psi(P, Q) := \max\{0, \arccos\langle b_1, b_2\rangle - \phi_1 - \phi_2\}. \tag{17}$$

Proof Formula (16) is a direct consequence of Theorem 2.1. Formula (17) is obtained by combining (5) and (16). □

3 Maximal angle between two topheavy cones

A topheavy cone in $\mathbb{R}^{n+1}$ is a closed convex cone of the form

$$\operatorname{epi} f := \left\{(\xi, t) \in \mathbb{R}^{n+1} : f(\xi) \leq t\right\},$$

where f is a norm on $\mathbb{R}^n$. Topheavy cones are pointed, have nonempty interior, and enjoy a number of useful properties. By way of example, one can mention the duality formula $(\operatorname{epi} f)^* = \operatorname{epi}(f_*)$, where f_* denotes the dual norm of f. Topheavy cones have been studied under various points of view in the literature, see for instance Fiedler and Haynsworth (1973), Lyubich (1995) and Seeger (2011). The next proposition explains how to compute the maximal angle between two topheavy cones.

Proposition 3.1 *Let f and g be norms on $\mathbb{R}^n$. Then,* $\cos\left[\Theta(\operatorname{epi} f, \operatorname{epi} g)\right]$ *is equal to the optimal value of the minimization problem*

$$\begin{cases} \text{minimize } \xi \odot \eta \\ \|\xi\|^2 + [f(\xi)]^2 = 1 \\ \|\eta\|^2 + [g(\eta)]^2 = 1, \end{cases} \tag{18}$$

where $\odot$ stands for the "product" operation given by

$$\xi \odot \eta := \langle \xi, \eta\rangle + \left[1 - \|\xi\|^2\right]^{1/2} \left[1 - \|\eta\|^2\right]^{1/2}.$$

Proof The cones epi f and epi g have nonempty interior. The first and second components of an antipodal pair must be sought on the boundary of epi f and on the boundary of epi g, respectively. Hence, $\cos[\Theta(\mathrm{epi}\, f, \mathrm{epi}\, g)]$ is equal to the optimal value of the minimization problem

$$\begin{cases} \text{minimize } \langle \xi, \eta \rangle + ts \\ f(\xi) = t, \\ g(\eta) = s, \\ \|\xi\|^2 + t^2 = 1, \\ \|\eta\|^2 + s^2 = 1. \end{cases}$$

It suffices now to get rid of the variables t and s. □

To derive an explicit solution to the problem (18), one needs of course a bit more information on the norms f and g. The following definition proves to be useful.

Definition 3.2 Two norms f, g on $\mathbb{R}^n$ are lower correlated if the minimization problems

$$\alpha_f := \min\{f(x) : \|x\| = 1\}, \tag{19}$$
$$\alpha_g := \min\{g(x) : \|x\| = 1\}, \tag{20}$$

have a solution in common.

For instance, a positive multiple of $\|\cdot\|$ is lower correlated to any norm on $\mathbb{R}^n$. Without further ado, we state:

Theorem 3.3 *Let f, g be lower correlated norms on $\mathbb{R}^n$. Then*

$$\Theta(\mathrm{epi}\, f, \mathrm{epi}\, g) = \arccos\left(\frac{\alpha_f}{\left[1+\alpha_f^2\right]^{1/2}}\right) + \arccos\left(\frac{\alpha_g}{\left[1+\alpha_g^2\right]^{1/2}}\right). \tag{21}$$

The above maximal angle is attained for instance with the unit vectors

$$(\xi_0, t_0) := \left(\frac{w}{\left[1+\alpha_f^2\right]^{1/2}}, \frac{\alpha_f}{\left[1+\alpha_f^2\right]^{1/2}}\right) \in \mathrm{epi}\, f,$$
$$(\eta_0, s_0) := \left(\frac{-w}{\left[1+\alpha_g^2\right]^{1/2}}, \frac{\alpha_g}{\left[1+\alpha_g^2\right]^{1/2}}\right) \in \mathrm{epi}\, g,$$

where w is any vector taken from the set

$$S(f, g) := \left[\mathrm{argmin}_{\|x\|=1} f(x)\right] \bigcap \left[\mathrm{argmin}_{\|x\|=1} g(x)\right].$$

Proof Let $w \in S(f, g)$. Since w is a solution to (19) and $-w$ is a solution to (20), the vectors

$$\xi_0 := \frac{w}{\left[1+\alpha_f^2\right]^{1/2}}, \quad \eta_0 := \frac{-w}{\left[1+\alpha_g^2\right]^{1/2}}$$

solve, respectively,

$$\gamma_f := \max\left\{\|\xi\| : \|\xi\|^2 + [f(\xi)]^2 = 1\right\},$$
$$\gamma_g := \max\left\{\|\eta\| : \|\eta\|^2 + [g(\eta)]^2 = 1\right\}.$$

Let (ξ, η) be any pair satisfying the equality constraints in (18). Then

$$\|\xi\| \le \|\xi_0\| = \gamma_f = \left(1+\alpha_f^2\right)^{-1/2},$$
$$\|\eta\| \le \|\eta_0\| = \gamma_g = \left(1+\alpha_g^2\right)^{-1/2},$$

and

$$\left\langle \|\xi\|^{-1}\xi, \|\eta\|^{-1}\eta \right\rangle \ge \left\langle \|\xi_0\|^{-1}\xi_0, \|\eta_0\|^{-1}\eta_0 \right\rangle = -1.$$

Hence,

$$\xi \odot \eta \;\ge\; \xi_0 \odot \eta_0 \;=\; -\left[1+\alpha_f^2\right]^{-1/2}\left[1+\alpha_g^2\right]^{-1/2} + \left[1-\gamma_f^2\right]^{1/2}\left[1-\gamma_g^2\right]^{1/2}.$$

In other words, (ξ_0, η_0) solves (18) and

$$\cos\left[\Theta(\operatorname{epi} f, \operatorname{epi} g)\right] \;=\; \frac{\alpha_f\,\alpha_g - 1}{\left[1+\alpha_f^2\right]^{1/2}\left[1+\alpha_g^2\right]^{1/2}}.$$

The last equality is an equivalent way of writing (21). □

Remark 3.4 In view of Seeger (2011, Theorem 5.2), formula (21) can be rewritten as:

$$\Theta(\operatorname{epi} f, \operatorname{epi} g) = \frac{\theta_{\max}(\operatorname{epi} f) + \theta_{\max}(\operatorname{epi} g)}{2}, \tag{22}$$

where $\theta_{\max}(K)$ denotes the maximal angle of $K \in \mathcal{C}(\mathbb{X})$. Formula (22) is consistent with geometric intuition, but one must remember that f and g are assumed to be lower correlated. Indeed, formula (22) may fail if one drops the lower correlation assumption.

As particular example of topheavy cone, one may consider the ℓ^p-cone

$$L_p := \left\{(\xi, t) \in \mathbb{R}^{n+1} : \|\xi\|_p \leq t\right\}.$$

Here, $p \in [1, \infty]$ and $\|\cdot\|_p$ stands for the ℓ^p-norm in $\mathbb{R}^n$. Of special interest are the cases $p = 1$, $p = 2$, and $p = \infty$. One gets

$$\begin{aligned}
&\Theta(L_1, L_1) = \pi/2, \quad \Theta(L_2, L_2) = \pi/2,\\
&\Theta(L_1, L_2) = \pi/2, \quad \Theta(L_1, L_\infty) = \pi/2,\\
&\Theta(L_\infty, L_\infty) = 2 \arccos [1+n]^{-1/2},\\
&\Theta(L_2, L_\infty) = \arccos [1+n]^{-1/2} + \pi/4.
\end{aligned}$$

These formulas are obtained using the proposition stated below.

Proposition 3.5 *The following statements hold:*

(a) *Let* $p, q \in [2, \infty]$. *Then*

$$\Theta(L_p, L_q) = \arccos \left[1 + n^{(p-2)/p}\right]^{-1/2} + \arccos \left[1 + n^{(q-2)/q}\right]^{-1/2}.$$

(b) *Let* $p \in [1, \infty]$ *and* $q \in [1, p_*]$, *where* p_* *is given by* $p^{-1} + p_*^{-1} = 1$. *Then* $\Theta(L_p, L_q) = \pi/2$.

Proof Part (a). If p and q are both in $[2, \infty]$, then the norms $f(x) := \|x\|_p$ and $g(x) := \|x\|_q$ are lower correlated. Indeed,

$$\widehat{\mathbf{1}}_n := \frac{1}{\sqrt{n}} (1, \ldots, 1)^T \in \left[\operatorname{argmin}_{\|x\|=1} \|x\|_p\right] \bigcap \left[\operatorname{argmin}_{\|x\|=1} \|x\|_q\right].$$

So, it suffices to substitute $\alpha_f = \|\widehat{\mathbf{1}}_n\|_p$ and $\alpha_g = \|\widehat{\mathbf{1}}_n\|_q$ into (21).
Part (b). For all $p \in [1, \infty]$, one has the duality formula $L_p^* = L_{p_*}$ (cf. Lyubich 1995, Proposition 3.1). Hence, Proposition 1.4 yields $\Theta(L_p, L_{p_*}) = \pi/2$. On the other hand, by applying Theorem 3.3 one gets $\Theta(L_1, L_1) = \pi/2$. Hence, for $q \in [1, p_*]$, one obtains

$$\pi/2 = \Theta(L_1, L_1) \leq \Theta(L_p, L_q) \leq \Theta(L_p, L_{p_*}) = \pi/2. \tag{23}$$

Of course, in (23) one uses the fact that the family $\{L_p\}_{p \geq 1}$ is nondecreasing with respect to set inclusion. $\square$

3.1 Maximal angle between two ellipsoidal cones

An ellipsoidal cone in $\mathbb{R}^{n+1}$ is a closed convex cone of the form

$$E_A := \left\{(\xi, t) \in \mathbb{R}^{n+1} : \sqrt{\langle \xi, A\xi \rangle} \leq t\right\},$$

where A is a positive definite symmetric matrix of order n. Hence, an ellipsoidal cone is a particular instance of a topheavy cone. It is easy to see that the norms

$$f(x) = \sqrt{\langle x, Ax\rangle}, \quad g(x) = \sqrt{\langle x, Bx\rangle} \tag{24}$$

are lower correlated if the eigenspaces

$$E_{\min}(A) := \{x \in \mathbb{R}^n : Ax = \lambda_1(A)x\}$$
$$E_{\min}(B) := \{x \in \mathbb{R}^n : Bx = \lambda_1(B)x\}$$

have a nonzero vector in common. Here, $\lambda_1(A)$ stands for the smallest eigenvalue of A.

Proposition 3.6 *Let A, B be positive definite symmetric matrices of order n. Then,* $\cos[\Theta(E_A, E_B)]$ *is equal to the optimal value of the minimization problem*

$$\begin{cases} \text{minimize } \xi \odot \eta \\ \|\xi\|^2 + \langle \xi, A\xi\rangle = 1 \\ \|\eta\|^2 + \langle \eta, B\eta\rangle = 1. \end{cases}$$

Furthermore, if the eigenspaces $E_{\min}(A)$ and $E_{\min}(B)$ have nonzero vector in common, then

$$\Theta(E_A, E_B) = \arccos\left[\frac{\lambda_1(A)}{1+\lambda_1(A)}\right]^{1/2} + \arccos\left[\frac{\lambda_1(B)}{1+\lambda_1(B)}\right]^{1/2}.$$

Proof It suffices to apply Proposition 3.1 and Theorem 3.3 to the norms mentioned in (24). Note that

$$\alpha_f^2 = \min_{\|x\|=1} \langle x, Ax\rangle = \lambda_1(A).$$

Similarly, $\alpha_g^2 = \lambda_1(B)$. □

3.2 An ellipsoidal cone versus a nonnegative orthant

The next proposition gives a formula for computing the maximal angle between an ellipsoidal cone and a nonnegative orthant. One uses the notation $\mu_{\min}(C)$ to indicate the smallest Pareto eigenvalue of a square matrix C (cf. Seeger 1999). From the general theory of Pareto eigenvalues, one knows that

$$\mu_{\min}(C) = \min_{\substack{\|\eta\|=1 \\ \eta \geq 0}} \langle \eta, C\eta\rangle \tag{25}$$

whenever the matrix C is symmetric. The notation $\eta \geq 0$ indicates that each component of $\eta \in \mathbb{R}^n$ is nonnegative.

Proposition 3.7 *Let A be a positive definite symmetric matrix of order n. Let $C := -(I_n + A)^{-1}$ with I_n denoting the identity matrix of order n. Then*

$$\Theta(E_A, \mathbb{R}^{n+1}_+) = \arccos\left(-\sqrt{-\mu_{\min}(C)}\right).$$

Furthermore, $((\xi, t), (\eta, s))$ is an antipodal pair of $(E_A, \mathbb{R}^{n+1}_+)$ if and only if

$$\begin{cases} \eta \text{ is a solution to (25)}, \\ s = 0, \\ \xi = [-\langle \eta, C\eta\rangle]^{-1/2}\, C\eta, \\ t = 1 + [\mu_{\min}(C)]^{-1}\|C\eta\|^2. \end{cases}$$

Proof The term $c := \cos[\Theta(E_A, \mathbb{R}^{n+1}_+)]$ corresponds to the optimal value of the minimization problem

$$\begin{cases} \text{minimize } \ \langle \xi, \eta\rangle + ts \\ \langle \xi, A\xi\rangle^{1/2} = t, \\ \|\xi\|^2 + t^2 = 1, \\ \|\eta\|^2 + s^2 = 1, \\ \eta \geq 0,\, s \geq 0. \end{cases} \tag{26}$$

Clearly, $s = 0$ at the minimum. By getting rid of the variable t, the problem (26) is converted into

$$\begin{cases} \text{minimize } \ \langle \xi, \eta\rangle \\ \|\xi\|^2 + \langle \xi, A\xi\rangle = 1, \\ \|\eta\|^2 = 1, \ \eta \geq 0 \end{cases}$$

The change of variables $\gamma = (I_n + A)^{1/2}\xi$ leads to

$$c = \min_{\substack{\|\eta\|=1 \\ \eta\geq 0}} \min_{\|\gamma\|=1} \langle (I_n + A)^{-1/2}\eta, \gamma\rangle.$$

Since the inner minimization problem is solved by

$$\gamma = -\|(I_n + A)^{-1/2}\eta\|^{-1}\,(I_n + A)^{-1/2}\eta,$$

one gets

$$\begin{aligned} -c &= \max_{\substack{\|\eta\|=1 \\ \eta\geq 0}} \|(I_n + A)^{-1/2}\eta\| = \left[\max_{\substack{\|\eta\|=1 \\ \eta\geq 0}} \langle \eta, (I_n + A)^{-1}\eta\rangle\right]^{1/2} \\ &= \left[-\min_{\substack{\|\eta\|=1 \\ \eta\geq 0}} \langle \eta, C\eta\rangle\right]^{1/2} = [-\mu_{\min}(C)]^{1/2}. \end{aligned}$$

This completes the proof of the proposition. □

3.3 An ellipsoidal cone versus a ray

The next proposition explains how to compute the maximal angle between an ellipsoidal cone E_A and a ray $\vec{v} := \{tv : t \geq 0\}$.

Proposition 3.8 *Let A be a positive definite symmetric matrix of order n and $v := (\eta, s)$ be a unit vector in $\mathbb{R}^{n+1}$. Then*

(a) $\cos[\Theta(E_A, \vec{v}\,)]$ *is equal to the optimal value of the nonconvex minimization problem*
$$\begin{cases} \text{minimize } \langle \xi, \eta\rangle + s[1 - \|\xi\|^2]^{1/2} \\ \|\xi\|^2 + \langle \xi, A\xi\rangle = 1. \end{cases}$$

(b) *Under the additional assumption $\langle \eta, A^{-1}\eta\rangle^{1/2} > s > 0$, one can write*
$$\cos\left[\Theta(E_A, \vec{v}\,)\right] = -s \min\left\{\|x - b\| : \langle x, Mx\rangle \leq 1\right\}, \tag{27}$$
where $b := (1/s)(I_n + A)^{-1/2}\eta$ and $M := (I_n + A)^{1/2}A^{-1}(I_n + A)^{1/2}$.

Proof The proof of (a) is as in Proposition 3.1, so one concentrates on (b). For notational convenience, one writes
$$\begin{aligned} f(\xi) &:= \langle \xi, A\xi\rangle^{1/2} = \|A^{1/2}\xi\|, \\ F(\xi) &:= \langle \xi, (I_n + A)\xi\rangle^{1/2} = \|(I_n + A)^{1/2}\xi\|. \end{aligned}$$

Note that f and F are norms on $\mathbb{R}^n$. Let $\gamma := \cos[\Theta(E_A, \vec{v}\,)]$. One has
$$\begin{aligned} \gamma &= \min_{\substack{(\xi,t)\in E_A \\ \|\xi\|^2+t^2=1}} \{\langle \xi, \eta\rangle + ts\} \\ &= \min_{\substack{f(\xi)=t \\ \|\xi\|^2+t^2=1}} \{\langle \xi, \eta\rangle + ts\} \qquad (28) \\ &= \min_{F(\xi)=1} \{\langle \xi, \eta\rangle + sf(\xi)\}, \qquad (29) \end{aligned}$$

where (28) is a consequence of Seeger and Sossa (2015, Theorem 2.3). Since $E_A^* = E_{A^{-1}}$, the condition $\langle \eta, A^{-1}\eta\rangle^{1/2} > s$ amounts to saying that (η, s) does not belong to dual cone of E_A, i.e., there exists a unit vector $(\tilde{\xi}, \tilde{t}) \in E_A$ such that $\langle \tilde{\xi}, \eta\rangle + \tilde{t}s < 0$. Hence, $\gamma < 0$ and, by a positive homogeneity argument, the equality constraint in (29) can be written as an inequality constraint. In other words, one has
$$\gamma = \min_{F(\xi)\leq 1} \{\langle \xi, \eta\rangle + sf(\xi)\}. \tag{30}$$

Observe that (30) is a convex minimization problem. Using standard rules of convex analysis, one can show that
$$\gamma = -s \min_{f^\circ(z)\leq 1} F^\circ\left(z - s^{-1}\eta\right), \tag{31}$$

where

$$f^\circ(\mu) = \langle \mu, A^{-1}\mu\rangle^{1/2} \;=\; \|A^{-1/2}\mu\|,$$
$$F^\circ(\mu) = \langle \mu, (I_n + A)^{-1}\mu\rangle^{1/2} \;=\; \|(I_n + A)^{-1/2}\mu\|$$

are the polar norms of f and F, respectively. One can view the minimization problem in (31) as a dual version of (30). To complete the proof of (b), it remains to introduce in (31) the change of variables $x = (I_n + A)^{-1/2}z$. □

The minimization problem on the right-hand side of (27) is about finding the minimal distance from a point to an ellipsoid. The numerical resolution of such a projection problem offers no difficulty.

4 Critical angles between two cones of matrices

Let the space $\mathrm{Sym}(n)$ of symmetric matrices of order n be equipped with the trace inner product $\langle A, B\rangle = \mathrm{tr}(AB)$. This section concerns the analysis of critical angles in a pair of convex cones in $\mathrm{Sym}(n)$. For notational convenience, one writes $\mathcal{CS}(n) := \mathcal{C}(\mathrm{Sym}(n))$ and introduces the symbol $\mathcal{O}_n$ to indicate the set of orthogonal matrices of order n. A nonempty set $\mathcal{P}$ in the space $\mathrm{Sym}(n)$ is said to be *orthogonally invariant* if

$$A \in \mathcal{P} \;\;\Rightarrow\;\; U^T AU \in \mathcal{P} \;\;\text{for all}\;\; U \in \mathcal{O}_n.$$

For instance, the SDP cone

$$\mathcal{P}_n := \{A \in \mathrm{Sym}(n) : A \text{ is positive semidefinite}\}$$

is orthogonally invariant.

Proposition 4.1 *Suppose that at least one of the cones $\mathcal{P}, \mathcal{Q} \in \mathcal{CS}(n)$ is orthogonally invariant. Let (A, B) be a critical pair of $(\mathcal{P}, \mathcal{Q})$. Then A and B commute, i.e., $AB = BA$.*

Proof Suppose, for instance, that $\mathcal{P}$ is orthogonally invariant. Write $\lambda := \langle A, B\rangle$. By Proposition 1.3, one knows that A minimizes the linear form $\langle B - \lambda A, \cdot\rangle$ on

$$\mathcal{P}^\diamond := \{X \in \mathcal{P} : \|X\| = 1\}.$$

Since $\mathcal{P}^\diamond$ is an orthogonally invariant set, the commutation principle stated in Iusem and Seeger (2007, Lemma 4) implies that $A(B - \lambda A) = (B - \lambda A)A$. This leads to $AB = BA$. □

There is a rich literature devoted to the analysis of orthogonally invariant sets. One knows, for instance, that $\mathcal{P} \in \mathcal{CS}(n)$ is orthogonally invariant if and only if there exists a permutation invariant cone $P \in \mathcal{C}(\mathbb{R}^n)$ such that

$$\mathcal{P} = \lambda^{-1}(P) := \{A \in \mathrm{Sym}(n) : \lambda(A) \in P\}. \tag{32}$$

Here and in the sequel, the notation $\lambda(A)$ stands for the vector of eigenvalues of A arranged in nondecreasing order, i.e.,

$$\lambda_1(A) \leq \cdots \leq \lambda_n(A).$$

The cone P in the representation formula (32) is unique and given by

$$P = \{x \in \mathbb{R}^n : \mathrm{Diag}(x) \in \mathcal{P}\},$$

where $\mathrm{Diag}(x)$ is the diagonal matrix whose entries on the diagonal are the components of x. One refers to P as the permutation invariant cone associated with $\mathcal{P}$.

Theorem 4.2 *Let $\mathcal{P}, \mathcal{Q} \in \mathcal{CS}(n)$ be orthogonally invariant and $P, Q \in \mathcal{C}(\mathbb{R}^n)$ be the associated permutation invariant cones. Then, $\Gamma(\mathcal{P}, \mathcal{Q}) = \Gamma(P, Q)$ and $\Theta(\mathcal{P}, \mathcal{Q}) = \Theta(P, Q)$. Furthermore, the following statements are equivalent:*

(a) *(A, B) is a critical (respectively, antipodal) pair of $(\mathcal{P}, \mathcal{Q})$,*
(b) *There exist a critical (respectively, antipodal) pair (u, v) of (P, Q) and a matrix $U \in \mathcal{O}_n$ such that $A = U\mathrm{Diag}(u)U^T$ and $B = U\mathrm{Diag}(v)U^T$.*

Proof The proof follows similar steps as in Iusem and Seeger (2007, Theorem 5). □

Example 4.3 By way of example, consider the problem of finding the critical angles between SDP cone $\mathcal{P}_n$ and the cone

$$\mathcal{D}_n := \{A \in \mathrm{Sym}(n) : \lambda_n(A) \leq \mathrm{tr}(A)\}.$$

Both cones are orthogonally invariant. The associated permutation invariant cones are $\mathbb{R}^n_+$ and

$$D_n := \{x \in \mathbb{R}^n : \max\{x_1, \ldots, x_n\} \leq x_1 + \cdots + x_n\},$$

respectively. Beware that $\mathcal{P}_n$ and $\mathcal{D}_n$ are non-polyhedral cones in $\mathrm{Sym}(n)$, so a direct computation of $\Gamma(\mathcal{P}_n, \mathcal{D}_n)$ could be difficult. Computing $\Gamma(\mathbb{R}^n_+, D_n)$ is much easier, because $\mathbb{R}^n_+$ and D_n are simplicial cones in $\mathbb{R}^n$. Table 1 is filled by using Theorem 7.3 in Seeger and Sossa (2015).

Remark 4.4 It is possible to derive an explicit formula for the maximal angle between $\mathcal{P}_n$ and $\mathcal{D}_n$. One gets

$$\Theta(\mathcal{P}_n, \mathcal{D}_n) = \arccos\left(\frac{2-n}{\sqrt{n^2 - 3n + 3}}\right) \tag{33}$$

for all $n \geq 2$. For obtaining (33), we exploit the fact that D_n is a polyhedral cone generated by n linearly independent unit vectors, namely, the permutations of the vector

$$w = \frac{1}{\sqrt{n^2 - 3n + 3}}\,(2 - n, 1, \ldots, 1)^T.$$

Table 1 Critical angles between $\mathcal{P}_n$ and $\mathcal{D}_n$

n	$\cos\theta$	θ
3	1	0
	$-1/\sqrt{3}$	0.6959π
4	1	0
	$-1/\sqrt{5}$	0.6476π
	$-2/\sqrt{7}$	0.7728π
5	1	0
	$-1/\sqrt{7}$	0.6234π
	$-\sqrt{2}/\sqrt{5}$	0.7180π
	$-3/\sqrt{13}$	0.8128π

4.1 The SDP cone versus the cone of nonnegative matrices

We now address the difficult problem of estimating the maximal angle between the SDP cone $\mathcal{P}_n$ and the cone

$$\mathcal{N}_n := \{B \in \mathrm{Sym}(n) : B \text{ is nonnegative entrywise}\}.$$

Such problem was raised in a recent paper by Goldberg and Shaked-Monderer (2014). The following facts are known, see Goldberg and Shaked-Monderer (2014) for the parts (a) and (d).

Proposition 4.5 *One has:*

(a) $\Theta(\mathcal{P}_2, \mathcal{N}_2) = \Theta(\mathcal{P}_3, \mathcal{N}_3) = \Theta(\mathcal{P}_4, \mathcal{N}_4) = (3/4)\pi$.
(b) *The pair of matrices achieving the maximal angle* $\Theta(\mathcal{P}_2, \mathcal{N}_2)$ *is unique and given by*

$$(A, B) = \left(\begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}, \begin{bmatrix} 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix} \right).$$

(c) $\{\Theta(\mathcal{P}_n, \mathcal{N}_n)\}_{n\geq 2}$ *is nondecreasing. More generally,* $\Gamma(\mathcal{P}_n, \mathcal{N}_n) \subseteq \Gamma(\mathcal{P}_{n+1}, \mathcal{N}_{n+1})$.
(d) $\lim_{n\to\infty} \Theta(\mathcal{P}_n, \mathcal{N}_n) = \pi$.

The next theorem lists various conditions that are necessary for antipodality in $(\mathcal{P}_n, \mathcal{N}_n)$. We start by writing a linear algebra result concerning the smallest eigenvalue of a nonnegative symmetric matrix.

Lemma 4.6 *Let* $B \in \mathcal{N}_n$. *Then,* $\sqrt{2}\,\lambda_1(B) + \|B\| \geq 0$, *with equality if and only if*

$$\begin{cases} \lambda_1(B) + \lambda_n(B) = 0, \\ \lambda_2(B) = 0, \ldots, \lambda_{n-1}(B) = 0. \end{cases} \tag{34}$$

Proof To alleviate the notation, we write $\lambda_i := \lambda_i(B)$ for all $i \in \{1, \ldots, n\}$. Since B is nonnegative entrywise, the spectral radius

$$\rho(B) := \max_{1 \le i \le n} |\lambda_i|$$

of B is an eigenvalue of B. Hence, $\rho(B) = \lambda_n \ge -\lambda_1$. On the other hand,

$$[\lambda_1^2 + \cdots + \lambda_n^2]^{1/2} \ge [\lambda_1^2 + \lambda_n^2]^{1/2} \ge \frac{|\lambda_1| + |\lambda_n|}{\sqrt{2}} \ge \frac{\lambda_n - \lambda_1}{\sqrt{2}}.$$

It follows that

$$\|B\| + \sqrt{2}\lambda_1 \ge \left(\frac{\lambda_n + \lambda_1}{\sqrt{2}}\right) \ge 0.$$

This completes the proof of the lemma. □

Theorem 4.7 *Let $n \ge 3$. The following conditions are necessary for (A, B) to be an antipodal pair of $(\mathcal{P}_n, \mathcal{N}_n)$:*

(a) *A is not in $\mathcal{N}_n$ and B is not in $\mathcal{P}_n$.*
(b) $B = A^{\mathcal{N}_n} := \|\Pi_{\mathcal{N}_n}(-A)\|^{-1}\Pi_{\mathcal{N}_n}(-A)$.
(c) $A = B^{\mathcal{P}_n} := \|\Pi_{\mathcal{P}_n}(-B)\|^{-1}\Pi_{\mathcal{P}_n}(-B)$.
(d) $B_{i,i} = 0$ *for all* $i \in \{1, \ldots, n\}$.
(e) $AB = BA$.
(f) $\operatorname{rank}(B) \ge 2$.
(g) $\operatorname{rank}(A) = \operatorname{card}\{i : \lambda_i(B) < 0\} \le n - 2$.

Proof Part (a). This is because the angle between A and B is at least $(3/4)\pi$.
Part (b). Note that B solves the minimization problem

$$f(A) := \min\{\langle A, Y\rangle : Y \in \mathcal{N}_n, \|Y\| = 1\}.$$

This problem has clearly a unique solution, namely, the matrix $Y = A^{\mathcal{N}_n}$ whose entries are

$$Y_{i,j} = -(1/c)\min\{0, A_{i,j}\}$$

with

$$c := \|\Pi_{\mathcal{N}_n}(-A)\| = \left[\sum_{i,j=1}^{n} \left(\min\{0, A_{i,j}\}\right)^2\right]^{1/2}.$$

Part (c). Similarly, A solves the minimization problem

$$g(B) := \min\{\langle X, B\rangle : X \in \mathcal{P}_n, \|X\| = 1\}. \tag{35}$$

We claim that $B^{\mathcal{P}_n}$ is the unique solution to (35). Since A and B commute, there exist an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathbb{R}^n$ and a unit vector $\gamma \in \mathbb{R}^n_+$ such that

$$A = \sum_{i=1}^{n} \gamma_i x_i x_i^T, \quad B = \sum_{i=1}^{n} \lambda_i(B) x_i x_i^T. \tag{36}$$

One has

$$\begin{aligned} \langle \gamma, \lambda(B) \rangle = \langle A, B \rangle = g(B) &= \min_{\substack{\|\xi\|=1 \\ \xi \geq 0}} \sum_{i=1}^{n} \left\langle \xi_i x_i x_i^T, B \right\rangle \\ &= \min_{\substack{\|\xi\|=1 \\ \xi \geq 0}} \langle \xi, \lambda(B) \rangle. \end{aligned} \tag{37}$$

Hence, γ solves the minimization problem (37). But such problem admits a unique solution, which can be computed explicitly in terms of the $\lambda_i(B)$'s. One gets

$$\gamma_i = -(1/d) \min\{0, \lambda_i(B)\}, \tag{38}$$

with

$$d := \left[\sum_{i=1}^{n} (\min\{0, \lambda_i(B)\})^2 \right]^{1/2}.$$

By combining (36) and (38) one sees that, up to normalization, A is the projection of $-B$ onto $\mathcal{P}_n$.
Part (d). This is a consequence of (b).
Part (e). This is a consequence of Proposition 4.1.
Part (f). As a consequence of (d), at least two eigenvalues of B are different from 0.
Part (g). Let r be the number of negative eigenvalues of B. Then, thanks to (36) and (38), one has

$$A = -(1/d) \sum_{i=1}^{r} \lambda_i(B) x_i x_i^T$$

with $d = \left[\sum_{k=1}^{r} \lambda_k^2(B)\right]^{1/2}$. In particular, the rank$(A) = r$. In the remaining part of the proof, we use the notation $\lambda_i := \lambda_i(B)$. One has rank$(A) \leq n - 1$, because A must be on the boundary of $\mathcal{P}_n$. Suppose that rank$(A) = n - 1$. One must arrive to a contradiction. From the proof of (c), one sees that $\lambda_i < 0$ for all $i \in \{1, \ldots, n-1\}$ and

$$\langle A, B \rangle = -\left(\lambda_1^2 + \cdots + \lambda_{n-1}^2\right)^{1/2}.$$

On the other hand, one has

$$\lambda_1 + \cdots + \lambda_n = 0, \quad \lambda_1^2 + \cdots + \lambda_n^2 = 1, \quad \lambda_n > 0.$$

One gets in this way

$$\lambda_n = \frac{1}{\sqrt{2}} \left[1 + 2 \sum_{1 \le i < j \le n-1} \lambda_i \lambda_j \right]^{1/2} > \frac{1}{\sqrt{2}} \tag{39}$$

and $\langle A, B \rangle = -[1 - \lambda_n^2]^{1/2} > -1/\sqrt{2}$, contradicting the inequality $\Theta(\mathcal{P}_n, \mathcal{N}_n) \ge (3/4)\pi$. □

The next corollary fully settles the case $n = 3$.

Corollary 4.8 (A, B) *is an antipodal pair of* $(\mathcal{P}_3, \mathcal{N}_3)$ *if and only if*

$$A = xx^T, \quad B = \frac{1}{\sqrt{2}} \left(yy^T - xx^T \right)$$

with $x, y \in \mathbb{R}^3$ *such that*

$$\begin{cases} \|x\| = 1, \ \|y\| = 1, \ \langle x, y \rangle = 0, \\ y_i y_j \ge x_i x_j \ \textit{for}\ 1 \le i \le j \le 3. \end{cases} \tag{40}$$

Proof Let (A, B) be an antipodal pair of $(\mathcal{P}_3, \mathcal{N}_3)$. Theorem 4.7(g) implies that $\operatorname{rank}(A) = 1$. Hence, $A = xx^T$ with $\|x\| = 1$. Using Lemma 4.6, one gets

$$\begin{aligned} \cos[\Theta(\mathcal{P}_3, \mathcal{N}_3)] &= \min_{\|u\|=1} \min_{\substack{B \in \mathcal{N}_3 \\ \|B\|=1}} \langle uu^T, B \rangle \\ &= \min_{\substack{B \in \mathcal{N}_3 \\ \|B\|=1}} \lambda_1(B) = -1/\sqrt{2}. \end{aligned}$$

The second part of the corollary is obtained using (34). □

Remark 4.9 If t, s are nonnegative reals such that $t^2 + s^2 = 1$, then

$$x = \left(1/\sqrt{2}\right) (t, s, -1)^T, \quad y = \left(1/\sqrt{2}\right) (t, s, 1)^T$$

satisfy (40) and lead to the antipodal pair

$$(A, B) = \left(\frac{1}{2} \begin{bmatrix} t^2 & ts & -t \\ ts & s^2 & -s \\ -t & -s & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & t \\ 0 & 0 & s \\ t & s & 0 \end{bmatrix} \right).$$

Hence, the number of antipodal pairs of $(\mathcal{P}_3, \mathcal{N}_3)$ is not finite, not even countable.

From the proof of Theorem 4.7, one sees that

$$\cos\left[\Theta(\mathcal{P}_n,\mathcal{N}_n)\right] = \min\{f(A) : A \in \mathcal{P}_n,\ \|A\| = 1\} \tag{41}$$

$$= \min\{g(B) : B \in \mathcal{N}_n,\ \|B\| = 1\} \tag{42}$$

with

$$f(A) = -\left[\sum_{i,j=1}^{n} \left(\min\{0, A_{i,j}\}\right)^2\right]^{1/2},$$

$$g(B) = -\left[\sum_{i=1}^{n} (\min\{0, \lambda_i(B)\})^2\right]^{1/2}.$$

The minimization problems (41) and (42) are hard to solve in practice, because they are nonconvex and nonsmooth. However, the variational formulas (41) and (42) are useful to obtain lower bounds for $\Theta(\mathcal{P}_n,\mathcal{N}_n)$.

Example 4.10 Consider for instance the case $n = 5$. The nonnegative matrix B

$$B = \frac{1}{\sqrt{10}}\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

has unit norm and its eigenvalues are

$$\lambda_1(B) = \lambda_2(B) = \frac{-1-\sqrt{5}}{2\sqrt{10}},\quad \lambda_3(B) = \lambda_4(B) = \frac{-1+\sqrt{5}}{2\sqrt{10}},\quad \lambda_5(B) = \frac{2}{\sqrt{10}}.$$

Hence

$$g(B) = -\left[\left(\frac{-1-\sqrt{5}}{2\sqrt{10}}\right)^2 + \left(\frac{-1-\sqrt{5}}{2\sqrt{10}}\right)^2\right]^{1/2} = -\frac{1+\sqrt{5}}{2\sqrt{5}}$$

and

$$\Theta(\mathcal{P}_5,\mathcal{N}_5) \geq \arccos\left(-\frac{1+\sqrt{5}}{2\sqrt{5}}\right) \approx 0.7575\pi.$$

Intensive numerical testing suggests that the above inequality is an equality, but we do not have a formal proof of this fact. The strict inequality $\Theta(\mathcal{P}_5,\mathcal{N}_5) > (3/4)\pi$ was already observed in Goldberg and Shaked-Monderer (2014).

The next proposition is a complement to Theorem 4.7. It applies to the case $n \geq 5$ only.

Proposition 4.11 *Suppose that $n \geq 5$. Let (A, B) be an antipodal pair of $(\mathcal{P}_n, \mathcal{N}_n)$. Then,*

$$\lambda_2(B) < 0 < \lambda_{n-1}(B).$$

In particular, $\operatorname{rank}(B) \geq 4$ *and* $\operatorname{rank}(A) \geq 2$.

Proof Let $\lambda_i := \lambda_i(B)$ for all $i \in \{1, \ldots, n\}$. Suppose that $\lambda_{n-1} \leq 0$. One must arrive to a contradiction. One has $\lambda_i \leq 0$ for all $i \in \{1, \ldots, n-1\}$. The inequality in (39) is not strict, but holds in the form "$\geq$". One gets in such a case

$$\langle A, B \rangle = -\sqrt{1 - \lambda_n^2} \geq -1/\sqrt{2},$$

which contradicts the inequality $\Theta(\mathcal{P}_n, \mathcal{N}_n) > (3/4)\pi$, cf. Example 4.10. Hence, $\lambda_{n-1} > 0$. If $\lambda_2 \geq 0$, then

$$-\lambda_1 = \lambda_2 + \cdots + \lambda_n \geq \lambda_{n-1} + \lambda_n > \lambda_n,$$

contradicting the fact that $\lambda_n = \rho(B)$. $\square$

It is quite difficult to obtain an explicit formula for $\Theta(\mathcal{P}_n, \mathcal{N}_n)$ when $n \geq 5$. Some words on the numerical estimation of $\Theta(\mathcal{P}_n, \mathcal{N}_n)$ are in order. As a consequence of Seeger and Sossa (2015, Theorem 3.2), one can write

$$\Theta(\mathcal{P}_n, \mathcal{N}_n) = 2 \arccos \sqrt{2 t_n},$$

where t_n denotes the optimal value of the nonlinear program

$$\begin{cases} \text{minimize } f_0(Z, t) := t, \\ (Z, t) \in \operatorname{Sym}(n) \times \mathbb{R}, \\ f_1(Z, t) := \frac{1}{2} [\operatorname{dist}(Z, \mathcal{P}_n)]^2 - t \leq 0, \\ f_2(Z, t) := \frac{1}{2} [\operatorname{dist}(Z, -\mathcal{N}_n)]^2 - t \leq 0, \\ f_3(Z, t) := \|Z\|^2 - 1 = 0. \end{cases} \tag{43}$$

The gradients of f_0, f_1, f_2, and f_3 are all easily computable. For instance, the partial gradients of f_1 and f_2 with respect to Z are given by

$$\langle \nabla_Z f_1(Z, t), D \rangle = D - \Pi_{\mathcal{P}_n}(D),$$
$$\langle \nabla_Z f_2(Z, t), D \rangle = D + \Pi_{\mathcal{N}_n}(-D).$$

Projecting onto $\mathcal{P}_n$ and $\mathcal{N}_n$ offers no difficulty. Table 2 has been filled by solving (43) with the help of the package "fmincon" of MATLAB. This is done for each

Table 2 Lower bound for $\Theta(\mathcal{P}_n, \mathcal{N}_n)$

n	$\Theta(\mathcal{P}_n, \mathcal{N}_n)$	n	$\Theta(\mathcal{P}_n, \mathcal{N}_n)$	n	$\Theta(\mathcal{P}_n, \mathcal{N}_n)$	n	$\Theta(\mathcal{P}_n, \mathcal{N}_n)$	n	$\Theta(\mathcal{P}_n, \mathcal{N}_n)$
5	$0.7575\,\pi$	6	$0.7575\,\pi$	7	$0.7575\,\pi$	8	$0.7607\,\pi$	9	$0.7607\,\pi$
10	$0.7609\,\pi$	11	$0.7626\,\pi$	12	$0.7649\,\pi$	13	$0.7649\,\pi$	14	$0.7658\,\pi$
15	$0.7677\,\pi$	16	$0.7699\,\pi$	17	$0.7699\,\pi$	18	$0.7699\,\pi$	19	$0.7703\,\pi$
20	$0.7719\,\pi$	21	$0.7719\,\pi$	22	$0.7719\,\pi$	23	$0.7722\,\pi$	24	$0.7735\,\pi$
25	$0.7735\,\pi$	26	$0.7735\,\pi$	27	$0.7739\,\pi$	28	$0.7750\,\pi$	29	$0.7750\,\pi$

$n \in \{5, \ldots, 29\}$. Since (43) is a nonconvex optimization problem, we are not sure if "fmincon" is yielding a global solution or just a local solution. For this reason, we are rather conservative and consider the figures in Table 2 only as lower bounds for $\Theta(\mathcal{P}_n, \mathcal{N}_n)$. These figures have been rounded down to four decimal places.

Remark 4.12 Consider a dimension n of the form $n = (q + 1)(q^3 + 1)$, with q being a prime power. It has been shown by Goldberg and Shaked-Monderer (2014) that

$$\Theta(\mathcal{P}_n, \mathcal{N}_n) \geq \arccos\left(-\frac{\sqrt{q^2+1}}{q+1}\right). \tag{44}$$

The lower bound (44) has the merit of being easily computable, but it applies only to special choices of n and it is not optimal in general. Consider for instance the choice $q = 2$, which corresponds to the first prime power. The inequality (44) becomes

$$\Theta(\mathcal{P}_{27}, \mathcal{N}_{27}) \geq \arccos\left(-\sqrt{5}/3\right) \approx 0.7677\,\pi,$$

but Table 2 yields the better lower bound $\Theta(\mathcal{P}_{27}, \mathcal{N}_{27}) \geq 0.7739\,\pi$.

References

De Cock K, De Moor B (2002) Subspace angles between ARMA models. Syst Control Lett 46:265–270

De Cock K, De Moor B (2000) Subspace angles between linear stochastic models. In: Proceedings 39th IEEE conference on decision and control, Sydney, Australia

Fiedler M, Haynsworth E (1973) Cones which are topheavy with respect to a norm. Linear Multilinear Algebra 1:203–211

Goldberg F, Shaked-Monderer N (2014) On the maximal angle between copositive matrices. Elec J Linear Algebra 27:837–850

Hotelling H (1936) Relations between two sets of variates. Biometrika 28:321–377

Iusem A, Seeger A (2007) Angular analysis of two classes of non-polyhedral convex cones: the point of view of optimization theory. Comput Applied Math 26:191–214

Jordan C (1875) Essai sur la géométrie à n dimensions. Bull Soc Math France 3:103–174

Lyubich Y (1995) Perron-Frobenius theory for finite-dimensional spaces with a hyperbolic cone. Linear Algebra Appl 220:283–309

Mohammadi B (2014) Principal angles between subspaces and reduced order modelling accuracy in optimization. Struct Multidiscip Optim 50:237–252

Obert DG (1991) The angle between two cones. Linear Algebra Appl 144:63–70

Seeger A (1999) Eigenvalue analysis of equilibrium processes defined by linear complementarity conditions. Linear Algebra Appl 292:1–14

Seeger A (2011) Epigraphical cones I. J Convex Anal 18:1171–1196

Seeger A, Sossa D (2015) Critical angles bewteen two convex cones. Part I: general theory. TOP, online since. doi:10.1007/s11750-015-0375-y

Shashua A, Wolf L (2003) Learning over sets using kernel principal angles. J Mach Learn Res 4:913–931

Tenenhaus M (1988) Canonical analysis of two convex polyhedral cones and applications. Psychometrika 53:503–524