
The cosine measure of a function at a point

Warren Hare · Gabriel Jarry-Bolduc · Chayne Planiden

August 7, 2026

Abstract The cosine measure of a set of vectors in $\mathbb{R}^n$ measures how well the set covers all directions in $\mathbb{R}^n$. It identifies the direction furthest, in angle, from the set. It is used in the convergence theory of various optimization algorithms, but also highlights interesting geometric properties of sets. For example, the cosine measure of a set S is greater than zero if and only if given any $\mathcal{C}^1$ function f at a point $\mathbf{x}$ where the gradient is nonzero, S must contain a descent direction of f at $\mathbf{x}$. In this paper, we examine the question of what can be said when the function f is non-differentiable or if it has a gradient equal to the zero vector. To examine these cases, we introduce the novel concept of the *cosine measure of a function* at a point. This value provides an infimum on the value of the cosine measure that a set of vectors requires to guarantee it contains a descent direction of the function at the point of interest. We present mathematical theory around this concept, including examples showing that the cosine measure of a smooth function can have any value in $[-1, 1]$. We further present algorithms to compute the cosine measure of a function, and examples demonstrating the algorithm on smooth and nonsmooth functions. These results also shed light on the cosine measure of infinite sets and nonconvex cones.

Keywords Cosine measure · Descent directions · Infinite sets · Finite-max function · ℓ_1 norm

1 Introduction

In *Derivative-free optimization* (DFO), *direct search methods* usually rely on *positive spanning sets* to explore the space around an incumbent solution. A notion closely related to positive

Warren Hare

Department of Mathematics, University of British Columbia, Kelowna, British Columbia, Canada.

Hare's research is partially funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada,

Discovery Grant #2023-03555. ORCID 0000-0002-4240-3903

E-mail: warren.hare@ubc.ca

Gabriel Jarry-Bolduc

Department of Mathematics and Statistics, American University of Sharjah, Sharjah, United Arab Emirates.

ORCID 0000-0002-1827-8508

E-mail: gabjarry@alumni.ubc.ca

Chayne Planiden

School of Mathematics and Physics, University of Wollongong, Wollongong, NSW, 2500, Australia.

Research supported by University of Wollongong. ORCID 0000-0002-0412-8445

E-mail: chayne@uow.edu.au

spanning sets is the concept of *positive bases*. A positive basis can be thought of as a minimal positive spanning set. In other words, if any one vector in a positive basis is removed, then it is no longer a positive spanning set. The properties of positive spanning sets and positive bases have been studied in [1, 6, 8, 9, 12, 18, 23, 28, 32, 33, 36, 37]. In [15, 29, 30], the notion of *positive k -spanning set* is studied and its value in DFO is briefly discussed.

One property of positive spanning sets (that make them useful in DFO methods) is that if the gradient of a function at a point is well-defined and not equal to the zero vector, any positive spanning set contains at least one descent direction (and at least one ascent direction) for the function at the point. To decide whether a given finite set of vectors is a positive spanning set of $\mathbb{R}^n$, the *cosine measure* can be computed [26]. It is known that a set of vectors is a positive spanning set of $\mathbb{R}^n$ if and only if the cosine measure is greater than zero [2]. The cosine measure is a real number between -1 and 1 inclusively. Its value quantifies the largest angle that can be created between the set and an arbitrary vector in $\mathbb{R}^n$; with values near 1 representing that this largest angle is close to 0 (i.e., $\cos(\theta) \approx 1$). The structure of positive bases with maximal cosine measure is investigated in [16, 24, 31]. The convergence theory supporting a direct search method usually assumes that the cosine measure values of the positive spanning sets utilized are bounded away from zero [7]. The cosine measure is also used to define the notion of *complexity measure* that arises in complexity bounds for direct search methods [5, 10].

The main goal of this paper is to investigate the value of the cosine measure necessary to guarantee that a set of vectors contains a descent direction of a real-valued function f . When the gradient of f at a point is well-defined and not equal to the zero vector, then a set of vectors with cosine measure value greater than zero guarantees that the set contains a descent direction of f at the point. This is not necessarily true if the gradient is equal to the zero vector or if the gradient does not exist. The notion of *cosine measure of a function* emerges from the investigation of the two latter cases. We present theoretical results and examples exploring this notion.

A deterministic algorithm for computing the cosine measure of a finite positive spanning set was proposed in [14]. A few months later, an alternative approach to compute the cosine measure of a finite set of vectors was introduced in [34]. Further research has revealed that computing the cosine measure exactly is an NP-hard problem [25], and proposed heuristics to compute the value quickly in high dimension [19]. Recently, the notion of cosine measure has been extended to consider *cosine measure relative to a subspace* and an algorithm for its computation is presented in [2]. All of the aforementioned algorithms begin with the assumption that the set of vectors is finite.

A secondary result of this current work is the development of algorithms to determine cosine measures where the set contains infinitely many vectors. This leads to algorithms for computing the cosine measure of a classical type of nonsmooth function: the finite-max function. In this case, the Clarke subdifferential at the point of interest is the convex hull of the active gradients, in line with the general gradient-limit representation of Clarke subdifferentials for stratifiable functions established in [11]. The set of non-descent directions for this type of function at the point of interest is the polar of the cone generated by the active gradients. The cosine measure of the function at a certain point quantifies the angular size of this cone, measuring how difficult it is for a finite set of directions to contain a descent direction. Whereas gradient-sampling methods [4] address this difficulty by approximating the Clarke subdifferential through randomly sampled gradients, the approach in this paper is deterministic.

The remainder of this paper is organized as follows. In Section 2, the notation is clarified, and background results and necessary definitions to understand the remainder of this paper are provided. In Section 3, the concept of cosine measure of a function is defined and properties of this novel definition are introduced. Examples are provided showing the cosine measure of a function can take any value between -1 and 1 inclusively when the gradient is equal to the zero

vector. In Section 4, pseudo-codes to compute the cosine measure of infinite sets are introduced. In Section 5, we show how to find the value of the cosine measure of a finite-max function. In particular, a closed-form solution is provided for the ℓ_1 -norm function. Lastly, the main results of this paper are summarized, and future research directions are proposed in Section 6.

2 Preliminaries

In this paper, we follow the standard notation found in [35]. A vector is written as a column vector and denoted with a bold-face lower-case letter. A matrix is denoted with a bold-face capital letter. The transpose of matrix $\mathbf{A}$ is denoted by $\mathbf{A}^\top$. We work in finite-dimensional space $\mathbb{R}^n$ with inner product $\mathbf{x}^\top \mathbf{y} = \sum_{i=1}^n x_i y_i$ and induced norm $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}}$. The identity matrix in $\mathbb{R}^{n \times n}$ is denoted by $\mathbf{I}_n$. We use $\mathbf{e}_i \in \mathbb{R}^n$ where $i \in \{1, 2, \dots, n\}$, to denote the standard unit basis vectors in $\mathbb{R}^n$, i.e. the i^{th} column of $\mathbf{I}_n$. The zero vector in $\mathbb{R}^n$ is denoted by $\mathbf{0}_n$ and the vector of all ones in $\mathbb{R}^n$ is denoted by $\mathbf{1}_n$. When there is no ambiguity about the dimensions of the matrices or vectors considered, the subscript n referring to the dimension may be dropped.

The notation $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is used to refer to a real-valued function f where the codomain is $\mathbb{R}$ and $\text{dom } f$ is assumed to be a non-empty subset of $\mathbb{R}^n$ where f is well-defined. The interior of a set $\mathcal{S}$ is denoted by $\text{int } \mathcal{S}$ and the closure of the set $\mathcal{S}$ is denoted by $\text{cl } \mathcal{S}$. The column space (span) of a set $\mathcal{S}$ is denoted by $\text{span}(\mathcal{S})$. the set of all unit vectors in $\mathbb{R}^n$ is denoted by $\mathbb{S}^n$. The definition of positive span follows.

Definition 1 (Positive span) Let $\mathcal{S}$ be a non-empty set in $\mathbb{R}^n$. The *positive span* of $\mathcal{S}$ is denoted by $\text{pspan}(\mathcal{S})$ and defined by

$$\text{pspan}(\mathcal{S}) = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{d}_i, k \in \mathbb{N}, \mathbf{d}_i \in \mathcal{S}, \lambda_i \geq 0 \right\}.$$

The *normalized positive span* of $\mathcal{S}$ is denoted by $\overline{\text{pspan}}(\mathcal{S})$ and defined by $\overline{\text{pspan}}(\mathcal{S}) = \text{pspan}(\mathcal{S}) \cap \mathbb{S}^n$.

Note that the positive span of $\mathcal{S}$ is equivalent to the *conical hull* of $\mathcal{S}$. When $\mathcal{S}$ is finite, $\text{pspan}(\mathcal{S})$ is a *polyhedral cone*. The set $\mathcal{S}$ is said to be a *positive spanning set* of $\mathbb{R}^n$, or to *positively span* $\mathbb{R}^n$, if and only if $\text{pspan}(\mathcal{S}) = \mathbb{R}^n$. It is known that the cardinality of a positive spanning set of $\mathbb{R}^n$ is equal to or greater than $n + 1$. A finite positive spanning set of $\mathbb{R}^n$ with cardinality s will be denoted by $\mathbb{P}_{n,s}$.

The following definition of cosine measure makes it possible to consider infinite sets of vectors. The max operator in the original definition (see [26]) is replaced by a supremum operator.

Definition 2 (Cosine measure) [2] Let $\mathcal{S} \subseteq \mathbb{R}^n$ be a non-empty (possibly infinite) set of unit vectors. The *cosine measure* of $\mathcal{S}$ is defined by

$$\text{cm}(\mathcal{S}) = \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \sup_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d},$$

and the *cosine vector set* of $\mathcal{S}$, denoted by $\text{cV}(\mathcal{S})$, is defined by

$$\text{cV}(\mathcal{S}) = \arg \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \sup_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d}.$$

In the case where $\mathcal{S}$ is empty, we define the cosine measure of $\mathcal{S}$ to be equal to -1 and the cosine vector set of $\mathcal{S}$ to be the empty set.

Definition 3 (active set) Let $\mathcal{S} \subseteq \mathbb{R}^n$ be a non-empty set of unit vectors.

The *active set* for $\mathcal{S}$ at a cosine vector $\mathbf{u} \in \text{cV}(\mathcal{S})$ is denoted by $\mathcal{A}(\mathcal{S}, \mathbf{u})$ and defined by

$$\mathcal{A}(\mathcal{S}, \mathbf{u}) = \{\mathbf{d} \in \mathcal{S} : \mathbf{d}^\top \mathbf{u} = \text{cm}(\mathcal{S})\}.$$

When $\mathcal{S}$ is finite and $\text{pspan}(\mathcal{S}) = \mathbb{R}^n$, it is known that $\mathcal{A}(\mathcal{S}, \mathbf{u})$ contains a basis of $\mathbb{R}^n$ for any $\mathbf{u} \in \text{cV}(\mathcal{S})$ [14, Cor.18]. The following result clarifies the relation between cosine measures and positive spanning sets of $\mathbb{R}^n$.

Proposition 4 Let $\mathcal{S}$ be a non-empty (possibly infinite) set of unit vectors. Then $\text{cm}(\mathcal{S}) > 0$ if and only if $\mathcal{S}$ is a positive spanning set of $\mathbb{R}^n$.

Proof The proof follows immediately from [2, Proposition 18]. $\square$

Next, we recall the definition of descent direction.

Definition 5 (Descent direction) Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathbf{x} \in \text{dom } f$. A unitary vector $\mathbf{d} \in \mathbb{R}^n$ is said to be a *descent direction* of f at $\mathbf{x}$ if and only if there exists a scalar $\bar{t} > 0$ such that $f(\mathbf{x} + t\mathbf{d}) < f(\mathbf{x})$ for all $t \in (0, \bar{t}]$.

As mentioned in Section 1, an important property of positive spanning sets is that if ∇f is well-defined at $\mathbf{x} \in \text{dom } f$ and $\nabla f(\mathbf{x}) \neq \mathbf{0}_n$, we have $-\nabla f(\mathbf{x})^\top \mathbf{d} > 0$ for at least one vector $\mathbf{d} \in \mathbb{P}_{n,s}$. In other words, there exists at least one descent direction of f at $\mathbf{x}$ in a positive spanning set of $\mathbb{R}^n$ [7]. The statement does not extend to non-differentiable functions or points where the gradient is equal to the zero vector, as the next two examples illustrate.

Example 6 Consider the function $f(\mathbf{x}) = \max\{|x_1|, |x_2|\}$ at the point $\mathbf{x}_0 = [1 \ 1]^\top$ and the set $\mathbb{P}_{2,4} = \{\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_1, -\mathbf{e}_2\}$. Then $\mathbb{P}_{2,4}$ is a positive spanning set of $\mathbb{R}^n$ with $\text{cm}(\mathbb{P}_{2,4}) = 1/\sqrt{2}$. However, $\mathbb{P}_{2,4}$ does not contain a descent direction of f at $\mathbf{x}_0$.

Example 7 Consider the function $f(\mathbf{x}) = (x_1 - x_2)^2 - (x_1 + x_2)^2$ at the point $\mathbf{x}_0 = \mathbf{0}$ and the set $\mathbb{P}_{2,4} = \{\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_1, -\mathbf{e}_2\}$. Then, $\mathbb{P}_{2,4}$ is a positive spanning set of $\mathbb{R}^n$ with $\text{cm}(\mathbb{P}_{2,4}) = 1/\sqrt{2}$. However, $\mathbb{P}_{2,4}$ does not contain a descent direction of f at $\mathbf{x}_0$.

Notice that in Examples 6 and 7, if a set of vectors with cosine measure strictly greater than $1/\sqrt{2}$ is utilized, then the set will contain a descent direction of f at the given respective point. We conclude that a “sufficiently good” set of vectors in terms of cosine measure may be able to find a descent direction in the non-differentiable case or when the gradient is equal to the zero vector. The question of how to quantify “sufficiently good” leads to the definition of *cosine measure of a function*.

3 Cosine measure of a function

In this section, we introduce the definition of cosine measure of a function. We will see that it provides information on the quality that a set of vectors must have to ensure a descent direction of f is found at $\mathbf{x}$. We investigate the properties of the cosine measure of a function, and provide examples in $\mathbb{R}^2$ that show the cosine measure of a function can take any value between -1 and 1 (inclusive) when the gradient is equal to the zero vector or does not exist.

Henceforth, the set of all unit vectors in $\mathbb{R}^n$ is denoted by $\mathbb{S}^n$. The set of all unit descent directions of f at $\mathbf{x}$ is denoted by $D(f; \mathbf{x})$. The complement of $D(f; \mathbf{x})$ in $\mathbb{S}^n$ is denoted by $D^c(f; \mathbf{x})$, i.e.,

$$\begin{aligned} D(f; \mathbf{x}) &= \{\mathbf{d} \in \mathbb{S}^n : \mathbf{d} \text{ is a descent direction of } f \text{ at } \mathbf{x}\}, \\ D^c(f; \mathbf{x}) &= \mathbb{S}^n \setminus D(f; \mathbf{x}). \end{aligned}$$

We simply write D and D^c when it is clear from the context what the function f and the point of interest $\mathbf{x}$ are.

Definition 8 (Cosine measure and cosine vector set of f at $\mathbf{x}$) Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathbf{x} \in \text{dom } f$. The cosine measure of f at $\mathbf{x}$ is denoted by $\text{cm}(f; \mathbf{x})$ and defined by

$$\text{cm}(f; \mathbf{x}) = \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d}$$

provided $D^c \neq \emptyset$. If $D^c = \emptyset$, then $\text{cm}(f; \mathbf{x})$ is defined to be -1 . The cosine vector set of f at $\mathbf{x}$ is denoted by $\text{cV}(f; \mathbf{x})$ and defined by

$$\text{cV}(f; \mathbf{x}) = \arg \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d}$$

provided $D^c \neq \emptyset$. If $D^c = \emptyset$, then $\text{cV}(f; \mathbf{x})$ is defined to be $\emptyset$.

In the previous definition, note that the supremum operator may be replaced by a max operator by considering the set $\text{cl } D^c$ rather than D^c . It follows that when $D^c \neq \emptyset$,

$$\text{cm}(f; \mathbf{x}) = \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \max_{\mathbf{d} \in \text{cl } D^c} \mathbf{u}^\top \mathbf{d} \quad \text{and} \quad \text{cV}(f; \mathbf{x}) = \arg \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \max_{\mathbf{d} \in \text{cl } D^c} \mathbf{u}^\top \mathbf{d}.$$

Note that $\text{cV}(f; \mathbf{x})$ is non-empty if and only if $D^c(f; \mathbf{x})$ is non-empty. Observe that, when $D^c \neq \emptyset$, the definition of the cosine measure of f at $\mathbf{x}$ can be reformulated in terms of the definition of cosine measure of a set:

$$\text{cm}(f; \mathbf{x}) = \text{cm}(D^c) = \text{cm}(\text{cl } D^c).$$

In words, the cosine measure of f at $\mathbf{x}$ is equal to the cosine measure of the set of all unit non-descent directions of f at $\mathbf{x}$.

Theorem 10 clarifies the relation between the cosine measure of a function at a point and the cosine measure of a set of vectors. First, we recall a result about the effect of an orthogonal matrix on a set of unit vectors.

Lemma 9 Let $\mathcal{S} \subseteq \mathbb{R}^n$ be a closed set of unit vectors and $R \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Define the closed set $\mathcal{S}_R = \{R\mathbf{d} : \mathbf{d} \in \mathcal{S}\}$. Then $\text{cm}(\mathcal{S}) = \text{cm}(\mathcal{S}_R)$.

Proof This is an immediate consequence of [34, Thm 3.4]. □

Theorem 10 Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathbf{x} \in \text{dom } f$ and let $\mathcal{S} \subseteq \mathbb{R}^n$ be a closed set of unit vectors (possibly infinite). If $\text{cm}(\mathcal{S}) > \text{cm}(f; \mathbf{x})$, then $\mathcal{S}$ contains a descent direction of f at $\mathbf{x}$. Moreover, $\mathcal{S}_R = \{R\mathbf{d} : \mathbf{d} \in \mathcal{S}\}$ contains a descent direction of f at $\mathbf{x}$ for any orthogonal matrix $R \in \mathbb{R}^{n \times n}$.

Proof Suppose $\text{cm}(\mathcal{S}) > \text{cm}(f; \mathbf{x}) \geq -1$. Since $\text{cm}(\mathcal{S}) > -1$, the set $\mathcal{S}$ contains at least two vectors (so $\mathcal{S} \neq \emptyset$).

If D^c is empty, then $D = \mathbb{S}^n$ and it follows that all of the vectors in $\mathcal{S}$ are descent directions of f at $\mathbf{x}$.

Suppose D^c is non-empty. Let $\mathbf{u} \in \text{cV}(f; \mathbf{x})$. We have

$$\sup_{\mathbf{v} \in \mathcal{S}} \mathbf{u}^\top \mathbf{v} \geq \text{cm}(\mathcal{S}) > \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} = \text{cm}(f; \mathbf{x}).$$

Let $\mathbf{v}^* \in \mathcal{S}$ be a vector attaining the supremum $\sup_{\mathbf{v} \in \mathcal{S}} \mathbf{u}^\top \mathbf{v}$. Note that such a vector exists in $\mathcal{S}$, since $\mathcal{S}$ is a closed set. If $\mathbf{v}^* \in D^c$, then

$$\text{cm}(f; \mathbf{x}) = \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} \geq \mathbf{u}^\top \mathbf{v}^* \geq \text{cm}(\mathcal{S}),$$

which is a contradiction to the assumption $\text{cm}(\mathcal{S}) > \text{cm}(f; \mathbf{x})$. Therefore, $\mathbf{v}^* \in D$, so is a descent direction of f at $\mathbf{x}$.

The second results follows from Lemma 9. $\square$

It is tempting to believe that given a set $\mathcal{S}$, function f , and point $\mathbf{x}$; if $\mathcal{S}_R$ contains a descent direction for any orthogonal matrix R , then $\text{cm}(\mathcal{S}) \geq \text{cm}(f; \mathbf{x})$. But this is not necessarily true, as the following example illustrates.

Example 11 Consider the function $f(\mathbf{x}) = x_1^2 - x_2^2$ at $\mathbf{x}_0 = \mathbf{0}$, and the minimal (optimal) positive basis $\mathcal{S} = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ -\frac{1}{2} \end{bmatrix} \right\}$. Note that $\mathcal{S}_R$ contains a descent direction of f at $\mathbf{x}_0$ for any orthogonal matrix $R \in \mathbb{R}^{2 \times 2}$. (This is visualized in Figure 1.) The cosine measure of $\mathcal{S}$ is $\text{cm}(\mathcal{S}) = 1/2$, but $\text{cm}(f; \mathbf{0}) = 1/\sqrt{2}$.

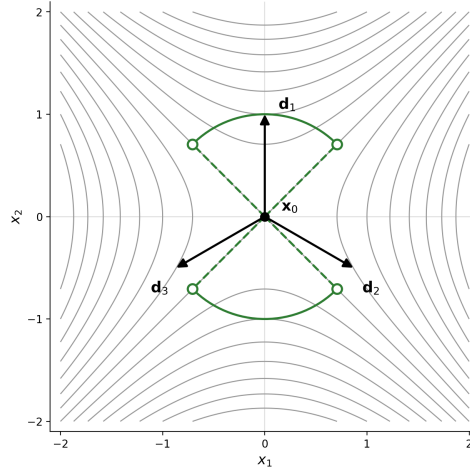


Fig. 1: Any rotation or reflection of the set $\mathcal{S}$ contains a descent direction of f at $\mathbf{x}_0 = \mathbf{0}$, but $\text{cm}(\mathcal{S}) = 1/2 < 1/\sqrt{2} = \text{cm}(f; \mathbf{0})$. The two green spherical caps represent the descent directions of f at $\mathbf{x}_0$.

From the previous example, we conclude that it is possible to create a set $\mathcal{S}$ with lower cosine measure than $\text{cm}(f; \mathbf{x})$ and identify a descent direction even when the set is rotated. However, it is also possible to create a set $\mathcal{S}$ with cosine measure equal to $\text{cm}(f; \mathbf{x})$ such that $\mathcal{S}$ does not contain a descent direction of f at $\mathbf{x}$, provided $\text{cm}(f; \mathbf{x}) > -1$.

Proposition 12 *Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathbf{x} \in \text{dom } f$. Let $\text{cm}(f; \mathbf{x}) > -1$. Then there exists a non-empty set of vectors $\mathcal{S}$ such that $\text{cm}(\mathcal{S}) = \text{cm}(f; \mathbf{x})$ and $\mathcal{S}$ does not contain a descent direction of f at $\mathbf{x}$.*

Proof Take $\mathcal{S} = D^c$. $\square$

We now provide several results about the value of the cosine measure of f at $\mathbf{x}$. In examining Theorem 13(iii), recall that, assuming $\text{pspan}(\text{cl } \mathcal{S}) \neq \mathbb{R}^n$, we have $\text{cm}(\text{cl } \mathcal{S}) = 0$ if and only if $\text{cl } \mathcal{S}$ contains a non-empty proper subset V such that $\text{pspan}(V) = \text{span}(V)$ [2, Corollary 29]. This implies that $\text{cl } \mathcal{S}$ contains a positive spanning set of a linear subspace $L \subset \mathbb{R}^n$ with $1 \leq \dim(L) < n$.

Theorem 13 *Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathbf{x} \in \text{dom } f$. Then the following hold:*

- (i) $\text{cm}(f; \mathbf{x}) = -1$ if and only if $|D^c| \leq 1$,
- (ii) $\text{cm}(f; \mathbf{x}) \leq 0$ if and only if D^c is contained in a closed half-space of $\mathbb{R}^n$ (whose boundary contains the origin),
- (iii) if $\text{pspan}(\text{cl } D^c) \neq \mathbb{R}^n$, then $\text{cm}(f; \mathbf{x}) = 0$ if and only if $\text{cl } D^c$ contains a positive spanning set of a linear subspace $L \subset \mathbb{R}^n$ where $1 \leq \dim(L) < n$,
- (iv) $\text{cm}(f; \mathbf{x}) = 1$ if and only if $\text{cl } D^c = \mathbb{S}^n$.

Proof (i) Suppose $\text{cm}(f; \mathbf{x}) = -1$ and $|D^c| > 1$. Let $\mathbf{d}_1$ and $\mathbf{d}_2$ be two distinct unit vectors in D^c . Then there exists a unit vector $\mathbf{u} \in \mathbb{R}^n$ such that $\sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} = -1$. This means that $\mathbf{u}^\top \mathbf{d}_1 = \mathbf{u}^\top \mathbf{d}_2 = -1$. This implies $\mathbf{d}_1 = -\mathbf{u}$ and $\mathbf{d}_2 = -\mathbf{u}$, a contradiction.

Conversely, suppose $|D^c| \leq 1$. If $|D^c| = 0$, then the result holds by definition. If $|D^c| = 1$, then take $\mathbf{u} = -\mathbf{d}$ where $\mathbf{d}$ is the vector in D^c .

(ii) Suppose $\text{cm}(f; \mathbf{x}) \leq 0$. If $\text{cV}(f; \mathbf{x})$ is empty, then this means $D^c = \emptyset$ and the result holds. Suppose $\mathbf{u} \in \text{cV}(f; \mathbf{x})$. Then

$$\text{cm}(f; \mathbf{x}) = \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} \leq 0.$$

Let H be the hyperplane in $\mathbb{R}^n$ with normal vector $\mathbf{u}$ passing through the origin. Then D^c is contained in the closed half space $H^- = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{u}^\top \mathbf{v} \leq 0\}$.

Conversely, suppose D^c is contained in a closed half-space defined by a hyperplane H . Let $\mathbf{u}$ be a unit normal vector to the hyperplane H such that

$$\sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} \leq 0.$$

Therefore, we must have $\text{cm}(f; \mathbf{x}) \leq \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} \leq 0$.

(iii) The result follows immediately from [2, Corollary 29].

(iv) Suppose $\text{cm}(f; \mathbf{x}) = 1$ and $\text{cl } D^c \neq \mathbb{S}^n$. This means there exists a unit vector $\mathbf{v}$ such that $\sup_{\mathbf{d} \in D^c} \mathbf{v}^\top \mathbf{d} = \max_{\mathbf{d} \in \text{cl } D^c} \mathbf{v}^\top \mathbf{d} < 1$, a contradiction.

Conversely, suppose $\text{cl } D^c = \mathbb{S}^n$. Let $\mathbf{u} \in \text{cV}(f; \mathbf{x})$. Then $\mathbf{u} \in \text{cl } D^c$. Hence,

$$\text{cm}(f; \mathbf{x}) = \sup_{\mathbf{d} \in D^c} \mathbf{u}^\top \mathbf{d} = \max_{\mathbf{d} \in \text{cl } D^c} \mathbf{u}^\top \mathbf{d} = \mathbf{u}^\top \mathbf{u} = 1.$$

□

When f is a single-variable function there are only three possibilities.

- (i) The set D^c is empty. In this case $\text{cm}(f; \mathbf{x}) = -1$.
- (ii) The set D^c contains one direction ($D^c = \{1\}$ or $D^c = \{-1\}$). In this case $\text{cm}(f; \mathbf{x}) = -1$.
- (iii) The set D^c contains two directions ($D^c = \{-1, 1\}$). In this case $\text{cm}(f; \mathbf{x}) = 1$.

Theorem 13 can be reformulated as follows for a single-variable function:

- (i) $\text{cm}(f; \mathbf{x}) = -1$ if and only if $\text{cl } D^c = D^c \neq \mathbb{S}^1$,
- (ii) $\text{cm}(f; \mathbf{x}) = 1$ if and only if $\text{cl } D^c = D^c = \mathbb{S}^1$.

The next theorem provides a result about the value of the cosine measure of a function at a differentiable point where $\nabla f(\mathbf{x}) \neq \mathbf{0}_n, n \geq 2$.

Theorem 14 (Cosine measure at a differentiable point where $\nabla f(\mathbf{x}) \neq \mathbf{0}$) *Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ with $n \geq 2$. Suppose f is differentiable at $\mathbf{x} \in \text{int dom } f$. If $\nabla f(\mathbf{x}) \neq \mathbf{0}$, then $\text{cm}(f; \mathbf{x}) = 0$ and $\text{cV}(f; \mathbf{x}) = \left\{ -\frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|} \right\}$.*

Proof Suppose $\nabla f(\mathbf{x}) \neq \mathbf{0}$. Consider the hyperplane passing through the origin and defined by $H = \{\mathbf{v} \in \mathbb{R}^n : -\nabla f(\mathbf{x})^\top \mathbf{v} = 0\}$. Then the unit vectors in H with initial point at the origin are in $\text{cl } D^c$. This means that $\text{cl } D^c$ contains a positive spanning set of a linear subspace with dimension $n - 1$. Since $\text{pspan}(\text{cl } D^c)$ is not equal to $\mathbb{R}^n$, $\text{cm}(f; \mathbf{x}) = 0$ by Theorem 13. Since the unit vector $\mathbf{u} \in \mathbb{R}^n$ such that $\mathbf{u}^\top \mathbf{d} \leq 0$ for all $\mathbf{d} \in \text{cl } D^c$ is unique, we get $\text{cV}(f; \mathbf{x}) = \left\{ -\frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|} \right\}$. $\square$

Note that the previous theorem assumes $n \geq 2$. For a single-variable function, $\text{cm}(f; \mathbf{x}) = -1$ whenever $f'(x) \neq 0$.

The converse of Theorem 14 is not necessarily true. Examples where $\text{cm}(f; \mathbf{x}) = 0$ and $\nabla f(\mathbf{x}) = \mathbf{0}$ can be created. The following section provides examples for the case where the gradient is equal to the zero vector. We will see that anything can happen when $\nabla f(\mathbf{x}) = \mathbf{0}$ (or when f is not differentiable at $\mathbf{x}$).

3.1 Cosine measure of a function when the gradient is equal to the zero vector

In this section, we provide examples in $\mathbb{R}^2$ showing that the cosine measure of a function may take any value between -1 and 1 inclusively. The value of the cosine measure of a function is calculated in each example. Note that the set of non-descent directions of a function at a point $\mathbf{x}$, $D^c(f; \mathbf{x})$, may not be a finite set of directions. In [2], a deterministic algorithm is provided to compute the cosine measure (relative to a subspace) of any non-empty finite set of nonzero vectors. Hence, Algorithm 1 from [2] can be applied directly to compute the cosine measure of a function when D^c is finite. In Section 4, we show how to compute the cosine measure of a real-valued function in $\mathbb{R}^n$. Alternately, by switching to polar coordinates, computing the cosine measure in $\mathbb{R}^2$ can be reduced to single-variable optimization. In all of the following examples this computation is easy, so omitted.

We begin with an example where $\nabla f(\mathbf{x}_0) = \mathbf{0}$ and $\text{cm}(f; \mathbf{x}_0) = -1$.

Example 15 ($\text{cm}(f; \mathbf{x}_0) = -1$) Consider the function $f : \text{dom } f = \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} = [x_1 \ x_2]^\top \mapsto -x_1^2 - x_2^2$ at the point $\mathbf{x}_0 = \mathbf{0}$. Note that D^c is empty. Therefore, $\text{cm}(f; \mathbf{x}_0) = -1$.

The next example illustrates $\nabla f(\mathbf{x}_0) = \mathbf{0}$ with $\text{cm}(f; \mathbf{x}_0) \in (-1, 0)$.

Example 16 ($\text{cm}(f; \mathbf{x}_0) \in (-1, 0)$) Consider the class of functions

$$f_\alpha : \text{dom } f = \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} \mapsto \begin{cases} -\alpha x_1^2 - x_2^2, & x_2 \leq 0 \\ -\alpha x_1^2 + x_2^2, & \text{otherwise,} \end{cases}$$

where $\alpha \in (0, \infty)$, at the point $\mathbf{x}_0 = \mathbf{0}$.

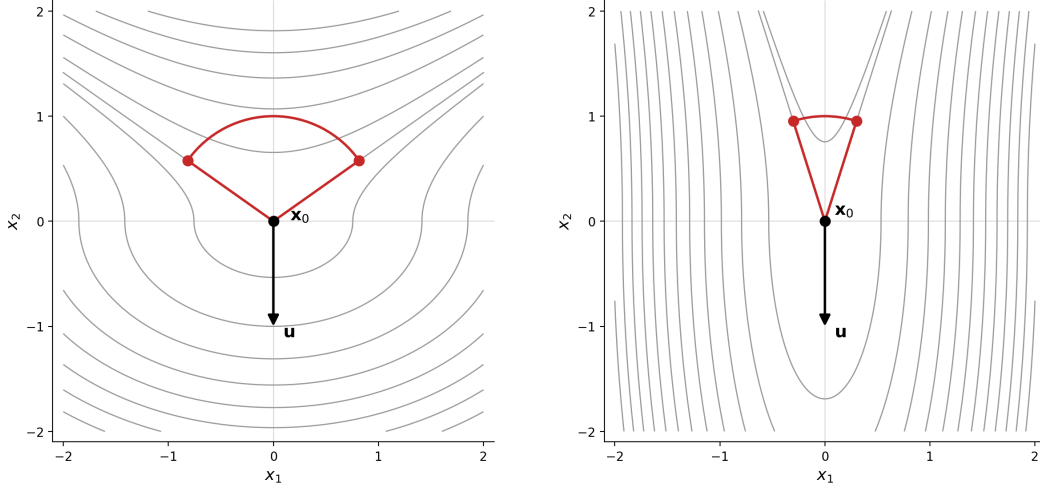


Fig. 2: The set D^c for f_α at $\mathbf{x}_0 = \mathbf{0}$ is represented by the red spherical cap. On the left, $\alpha = 0.5$, and on the right $\alpha = 10$.

Note that f_α is differentiable at $\mathbf{x}_0 = \mathbf{0}$. The contour lines at $f(\mathbf{x}) = 0$ are given by the equations $x_2 = \sqrt{\alpha}|x_1|$. As a vector-valued function, the contour lines at $f(\mathbf{x}) = 0$ are $L_1(t) = \left[\frac{1}{\sqrt{1+\alpha}} \sqrt{\frac{\alpha}{1+\alpha}} \right]^\top t$ and $L_2(t) = \left[\frac{-1}{\sqrt{1+\alpha}} \sqrt{\frac{\alpha}{1+\alpha}} \right]^\top t$, where $t \geq 0$.

We find that the set D^c is $\overline{\text{pspan}} \left\{ \left[\frac{1}{\sqrt{1+\alpha}} \sqrt{\frac{\alpha}{1+\alpha}} \right]^\top, \left[\frac{-1}{\sqrt{1+\alpha}} \sqrt{\frac{\alpha}{1+\alpha}} \right]^\top \right\}$. We obtain

$$\text{cm}(f_\alpha; \mathbf{0}) = -\sqrt{\frac{\alpha}{1+\alpha}}, \quad \text{and the cosine vector is } \mathbf{u} = [0 \ -1]^\top.$$

As a function of α , $\text{cm}(f_\alpha; \mathbf{0})$ is a continuous function on $(0, \infty)$ with

$$\lim_{\alpha \rightarrow 0^+} -\sqrt{\frac{\alpha}{1+\alpha}} = 0 \quad \text{and} \quad \lim_{\alpha \rightarrow \infty} -\sqrt{\frac{\alpha}{1+\alpha}} = -1.$$

Hence, $\text{cm}(f_\alpha; \mathbf{0})$ can take any value in the interval $(-1, 0)$.

We next illustrate $\nabla f(\mathbf{x}_0) = \mathbf{0}$ with $\text{cm}(f; \mathbf{x}_0) = 0$.

Example 17 ($\text{cm}(f; \mathbf{x}_0) = 0$) Consider the function $f : \text{dom } f = \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} = [x_1 \ x_2]^\top \mapsto -x_2^2$ at the point $\mathbf{x}_0 = \mathbf{0}$. As a vector-valued function, the contour line at $f(\mathbf{x}) = 0$ is given by $L(t) = [1 \ 0]^\top t$. We find that the set D^c is equal to $D^c = \{[1 \ 0]^\top, [-1 \ 0]^\top\}$ and $\text{cm}(f; \mathbf{0}) = 0$.

The following example provides a class of functions where $\nabla f(\mathbf{x}_0) = \mathbf{0}$ with $\text{cm}(f; \mathbf{x}_0) \in (0, 1)$.

Example 18 ($\text{cm}(f; \mathbf{x}_0) \in (0, 1)$) Consider the class of functions

$$f_\beta : \text{dom } f = \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} \mapsto -\beta x_1^2 + x_2^2,$$

where $\beta > 0$, at the point $\mathbf{x}_0 = \mathbf{0}$. For any values of $\beta > 0$, we have $\nabla f_\beta(\mathbf{0}) = \mathbf{0}$. As a vector-valued function, the contour lines at $f(\mathbf{x}) = 0$ have equation $L_1(t) = \left[\frac{1}{\sqrt{1+\beta}} \sqrt{\frac{\beta}{1+\beta}} \right]^\top t$, and

$L_2(t) = \left[\frac{1}{\sqrt{1+\beta}} - \sqrt{\frac{\beta}{1+\beta}} \right] t$. Let $\mathbf{d}_1 = \left[\frac{1}{\sqrt{1+\beta}} \sqrt{\frac{\beta}{1+\beta}} \right]^\top$ and $\mathbf{d}_2 = \left[\frac{1}{\sqrt{1+\beta}} - \sqrt{\frac{\beta}{1+\beta}} \right]^\top$. The set D^c can be written as $D^c = \overline{\text{pspan}}\{\mathbf{d}_1, -\mathbf{d}_2\} \cup \overline{\text{pspan}}\{-\mathbf{d}_1, \mathbf{d}_2\}$. We obtain

$$\text{cm}(f; \mathbf{0}) = \frac{1}{\sqrt{1+\beta}}$$

and the cosine vectors are $\mathbf{u}_1 = \mathbf{e}_1$ and $\mathbf{u}_2 = -\mathbf{e}_1$. As a function of β , $\text{cm}(f; \mathbf{0})$ is continuous on $(0, \infty)$. Additionally,

$$\lim_{\beta \rightarrow 0} \frac{1}{\sqrt{1+\beta}} = 1 \quad \text{and} \quad \lim_{\beta \rightarrow \infty} \frac{1}{\sqrt{1+\beta}} = 0.$$

This shows we can pick $\beta > 0$ to create a function where the cosine measure of f at $\mathbf{x}_0 = \mathbf{0}$ can take any value in $(0, 1)$.

Finally, we provide an example where $\nabla f(\mathbf{x}_0) = \mathbf{0}$ and $\text{cm}(f; \mathbf{x}_0) = 1$.

Example 19 ($\text{cm}(f; \mathbf{x}_0) = 1$) Consider the function $f : \text{dom } f = \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} \mapsto -x_1^3 + x_2^2$ at the point $\mathbf{x}_0 = \mathbf{0}$. The contour lines at $f(\mathbf{x}) = 0$ are given by the equations $x_2 = \pm x_1^{1.5}$. The only descent direction at $\mathbf{x}_0 = \mathbf{0}$ is $\mathbf{d} = \mathbf{e}_1$. Thus $D^c = \mathbb{S}^2 \setminus \{\mathbf{e}_1\}$ and $\text{cm}(f; \mathbf{0}) = 1$.

In the next section, pseudo-codes are provided showing how the cosine measure of a function may be computed for functions. The examples above are all sufficiently simple that these codes can be applied (by hand) to confirm the cosine measures.

4 Computing the cosine measure of a function

In this section, we show how to compute the cosine measure of a function at a point. The pseudo-code is divided in two cases: $\text{cl } D^c(f; \mathbf{x})$ is contained in a closed half-space of $\mathbb{R}^n$ and $\text{cl } D^c(f; \mathbf{x})$ is not contained in a closed half-space of $\mathbb{R}^n$. To make notation tighter, we write D^c and D to represent $D^c(f; \mathbf{x})$ and $D(f; \mathbf{x})$, respectively.

4.1 The case where $\text{cl } D^c$ is contained in a closed half-space of $\mathbb{R}^n$

We begin by considering the case where $\text{cl } D^c$ is contained in a closed half-space of $\mathbb{R}^n$. In this case, the cosine measure value is non-positive. We assume that the closure of the set of non-descent directions can be written in the form

$$\text{cl } D^c = \bigcup_{i=1}^k \overline{\text{pspan}}(\mathcal{S}_i) \tag{1}$$

where each $\mathcal{S}_i$ is a finite set of nonzero vectors. This structure allows $\text{cl } D^c$ to contain infinite vectors and does not force the cone $\{\mathbf{v} : \mathbf{v} = \lambda \mathbf{d}, \mathbf{d} \in \text{cl } D^c, \lambda \geq 0\}$ to be convex. Recall the definition of the polar cone of a set $\mathcal{S}$.

Definition 20 Let $\mathcal{S} \subseteq \mathbb{R}^n$ be non-empty. The *polar cone* of $\mathcal{S}$ is denoted by $\text{polar}(\mathcal{S})$ and defined by

$$\text{polar}(\mathcal{S}) = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v}^\top \mathbf{d} \leq 0 \text{ for all } \mathbf{d} \in \mathcal{S}\}.$$

Next, we present the pseudo-code for this case, then we show that it returns the exact cosine measure and cosine vector set.

Algorithm 1: Computing $\text{cm}(D^c)$ when $\text{cl } D^c$ is contained in a closed half-space of $\mathbb{R}^n$

Given a set of non-descent directions contained in a closed half-space of $\mathbb{R}^n$ with its closure written in the form

$$\text{cl } D^c = \bigcup_{i=1}^k \overline{\text{pspan}}(\mathcal{S}_i),$$

where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$.

0. Special cases and initialization:

Define the normalized set

$$\mathcal{S} = \bigcup_{i=1}^k \{\mathbf{d}/\|\mathbf{d}\| : \mathbf{d} \in \mathcal{S}_i, i = 1, 2, \dots, k\}.$$

Define $q = |\mathcal{S}|$.

If $q = 0$, then return $\text{cm}(D^c) = -1$ and $\text{cV}(D^c) = \emptyset$.

1. Optimize:

Solve the second-order conic program (SOCP)

$$(1.1) \quad (t^*, \mathbf{u}^*) \in \arg \min_{t, \mathbf{u}} \{t : \|\mathbf{u}\| \leq 1, \mathbf{u}^\top \mathbf{d} \leq t \quad \forall \mathbf{d} \in \mathcal{S}\}$$

2. Compute cosine measure and the cosine vector set: return

$$(2.1) \quad \text{cm}(D^c) = t^*.$$

If $\text{cm}(D^c) < 0$, then return

$$(2.2.1) \quad \text{cV}(D^c) = \{\mathbf{u}^*\}.$$

Else ($\text{cm}(D^c) = 0$),

$$(2.2.2) \quad \text{cV}(D^c) = \text{polar}(\mathcal{S}) \cap \mathbb{S}^n.$$

Before proving that Algorithm 1 returns the correct output, we make a few remarks. Note that Algorithm 1 requires solving a SOCP. In MATLAB, for example, this can be done with the function `coneprog`. Finally, we note that, by [33, Theorem 3.2], each $\mathcal{S}_i$ can be reduced to a positive linearly independent subset of $\mathcal{S}_i$ with maximal cardinality. In [33, Section 4], Regis refers to this as a *frame* of $\text{pspan}(\mathcal{S})$. The vectors of a frame are sometimes called the *generators*. If all generators of a frame for $\text{pspan}(\mathcal{S})$ lie in an open half-space, then $\text{pspan}(\mathcal{S})$ is said to be *pointed*. Note that the frame of a pointed polyhedral cone is unique (up to positive scaling). Reducing each $\mathcal{S}_i$ to a frame could reduce the size of the second-order conic program in Step 1 and the (possible) cost of finding the polar cone of $\mathcal{S}$, resulting in faster completion of the algorithm.

The following theorem provides the main result to prove Algorithm 1 returns the correct results.

Theorem 21 *Let $\mathcal{T} = \bigcup_{i=1}^k \overline{\text{pspan}}(\mathcal{S}_i)$, where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$, and let $\mathcal{S} = \bigcup_{i=1}^k \mathcal{S}_i$. If $\mathcal{T}$ is contained in a closed half-space of $\mathbb{R}^n$, then*

$$\text{cm}(\mathcal{T}) = \text{cm}(\mathcal{S}) \quad \text{and} \quad \text{cV}(\mathcal{T}) = \text{cV}(\mathcal{S}).$$

Proof Since $\mathcal{T}$ is contained in a closed half-space, $\mathcal{S}$ is contained in a closed half-space. This means that $\text{cm}(\mathcal{T}) \leq 0$ and $\text{cm}(\mathcal{S}) \leq 0$. Without loss of generality, assume all vectors in $\mathcal{S}$ are unit vectors. Since $\mathcal{S} \subseteq \mathcal{T}$, we have $\text{cm}(\mathcal{S}) \leq \text{cm}(\mathcal{T})$.

For a unit vector $\mathbf{u} \in \mathbb{R}^n$, write $f_{\mathcal{S}}(\mathbf{u}) = \max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d}$ and $f_{\mathcal{T}}(\mathbf{u}) = \max_{\mathbf{d} \in \mathcal{T}} \mathbf{u}^\top \mathbf{d}$. We claim that $f_{\mathcal{T}}(\mathbf{u}) = f_{\mathcal{S}}(\mathbf{u})$ whenever $f_{\mathcal{S}}(\mathbf{u}) \leq 0$. Indeed, let $\mathbf{d} \in \mathcal{T}$. Say $\mathbf{d} \in \overline{\text{pspan}}(\mathcal{S}_i)$ with

$\mathcal{S}_i = \{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_q\}$. We may write $\mathbf{d} = \sum_{j=1}^q \alpha_j \mathbf{d}_j$ with $\alpha_j \geq 0$. By the triangle inequality, we know $\sum_{j=1}^q \alpha_j \geq \|\mathbf{d}\| = 1$. We obtain

$$\mathbf{u}^\top \mathbf{d} = \sum_{j=1}^q \alpha_j \mathbf{u}^\top \mathbf{d}_j \leq \left(\sum_{j=1}^q \alpha_j \right) f_{\mathcal{S}}(\mathbf{u}) \leq f_{\mathcal{S}}(\mathbf{u}),$$

where the last inequality uses $f_{\mathcal{S}}(\mathbf{u}) \leq 0$. Since $\mathbf{d} \in \mathcal{T}$ was arbitrary, $f_{\mathcal{S}}(\mathbf{u})$ is an upper bound for $\mathbf{u}^\top \mathbf{d}$ over $\mathcal{T}$. Hence, $f_{\mathcal{T}}(\mathbf{u}) \leq f_{\mathcal{S}}(\mathbf{u})$. Clearly, $f_{\mathcal{S}}(\mathbf{u}) \leq f_{\mathcal{T}}(\mathbf{u})$ since $\mathcal{S} \subseteq \mathcal{T}$. This proves the claim.

For any $\mathbf{u}_{\mathcal{S}} \in \text{cV}(\mathcal{S})$, the previous result gives $\text{cm}(\mathcal{T}) \leq f_{\mathcal{T}}(\mathbf{u}_{\mathcal{S}}) = f_{\mathcal{S}}(\mathbf{u}_{\mathcal{S}}) = \text{cm}(\mathcal{S}) \leq 0$. Therefore $\text{cm}(\mathcal{T}) = \text{cm}(\mathcal{S})$.

Finally, if $\mathbf{u} \in \text{cV}(\mathcal{T})$, then $\text{cm}(\mathcal{S}) \leq f_{\mathcal{S}}(\mathbf{u}) \leq f_{\mathcal{T}}(\mathbf{u}) = \text{cm}(\mathcal{T}) = \text{cm}(\mathcal{S})$, so $\mathbf{u} \in \text{cV}(\mathcal{S})$. Conversely, if $\mathbf{u} \in \text{cV}(\mathcal{S})$, then $f_{\mathcal{S}}(\mathbf{u}) = \text{cm}(\mathcal{S}) \leq 0$, and we get $f_{\mathcal{T}}(\mathbf{u}) = \text{cm}(\mathcal{S}) = \text{cm}(\mathcal{T})$. That is, $\mathbf{u} \in \text{cV}(\mathcal{T})$. $\square$

Theorem 21 shows that the cosine measure of the (possibly) infinite set $\bigcup_{i=1}^k \overline{\text{pspan}}(\mathcal{S}_i)$ can be found by computing the cosine measure of the finite set $\mathcal{S} = \bigcup_{i=1}^k \mathcal{S}_i$. This allows the application of the techniques from [2]. We now provide three results on the cosine vector set.

Lemma 22 *Let $\mathcal{S}$ be a non-empty set of unit vectors in $\mathbb{R}^n$. If $\text{cm}(\mathcal{S}) < 0$, then $\text{cV}(\mathcal{S})$ contains exactly one vector.*

Proof Since $\mathcal{S}$ is a non-empty set, $\text{cV}(\mathcal{S})$ contains at least one vector. We show that it contains exactly one by contradiction. Suppose $\text{cm}(\mathcal{S}) = \alpha < 0$ and $\mathbf{u}, \mathbf{v}$ are two distinct unitary vectors contained in $\text{cV}(\mathcal{S})$. This means $\mathbf{u}^\top \mathbf{d} \leq \alpha$ and $\mathbf{v}^\top \mathbf{d} \leq \alpha$ for all $\mathbf{d} \in \mathcal{S}$. Consider the unit vector $\frac{\mathbf{u}+\mathbf{v}}{\|\mathbf{u}+\mathbf{v}\|}$. Note that $\|\mathbf{u} + \mathbf{v}\| \neq 0$ since $\|\mathbf{u} + \mathbf{v}\| = 0$ implies $\mathbf{v} = -\mathbf{u}$ and this would imply $\mathbf{v}^\top \mathbf{d} = -\mathbf{u}^\top \mathbf{d} \geq -\alpha > 0$. Then for any $\mathbf{d} \in \mathcal{S}$, we obtain

$$\left(\frac{\mathbf{u} + \mathbf{v}}{\|\mathbf{u} + \mathbf{v}\|} \right)^\top \mathbf{d} \leq \frac{2\alpha}{\|\mathbf{u} + \mathbf{v}\|} < \alpha,$$

as $\|\mathbf{u} + \mathbf{v}\| < \|\mathbf{u}\| + \|\mathbf{v}\| = 2$ and $\alpha < 0$. This is a contradiction to the assumption that $\text{cm}(\mathcal{S}) = \alpha$. $\square$

Next, we prove that the vector $\mathbf{u}^*$ found while solving the SOCP in Step (1.1) is a unitary vector whenever $\text{cm}(D^c) < 0$.

Lemma 23 *Let $\mathcal{S}$ be a finite set of unit vectors. Suppose that the convex relaxation of the cosine measure problem returns a negative value, i.e.,*

$$\min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\| \leq 1}} \max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d} < 0. \quad (2)$$

If $\mathbf{u}$ is a vector solving (2), then $\|\mathbf{u}\| = 1$.

Proof By way of contradiction, suppose $\mathbf{u}$ solves (2). If $\|\mathbf{u}\| = 0$, then $\mathbf{u}^\top \mathbf{d} = 0$ for all $\mathbf{d} \in \mathcal{S}$, a contradiction.

Suppose $0 < \|\mathbf{u}\| < 1$. Then

$$\left(\frac{\mathbf{u}}{\|\mathbf{u}\|} \right)^\top \mathbf{d} < \mathbf{u}^\top \mathbf{d} < 0$$

for all $\mathbf{d} \in \mathcal{S}$. It follows that

$$\max_{\mathbf{d} \in \mathcal{S}} \left(\frac{\mathbf{u}}{\|\mathbf{u}\|} \right)^\top \mathbf{d} < \max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d},$$

a contradiction to the assumption that $\mathbf{u}$ solves (2). $\square$

Next, we show that the cosine vector set of D^c is equal to $\text{polar}(D^c) \cap \mathbb{S}^n$ whenever $\text{cm}(D^c) = 0$.

Proposition 24 *Let $\mathcal{S}$ be a finite set of unit vectors in $\mathbb{R}^n$. If $\text{cm}(\mathcal{S}) = 0$, then $\text{cV}(\mathcal{S}) = \text{polar}(\mathcal{S}) \cap \mathbb{S}^n$.*

Proof Let $\mathbf{u} \in \text{cV}(\mathcal{S})$. Then $\max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d} = 0$ and so $\mathbf{u}^\top \mathbf{d} \leq 0$ for all $\mathbf{d} \in \mathcal{S}$. Hence, $\mathbf{u} \in \text{polar}(\mathcal{S}) \cap \mathbb{S}^n$. Now, suppose $\mathbf{u} \in \text{polar}(\mathcal{S}) \cap \mathbb{S}^n$. This implies $\max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d} \leq 0$. Since $\text{cm}(\mathcal{S}) = 0$, we must have $\max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d} = 0$ (otherwise, it contradicts the definition of cosine measure). It follows that $\mathbf{u} \in \text{cV}(\mathcal{S})$. $\square$

Note that

$$\text{polar}\left(\bigcup_{i=1}^k \text{pspan}(\mathcal{S}_i)\right) = \bigcap_{i=1}^k \text{polar}(\text{pspan}(\mathcal{S}_i)) = \bigcap_{i=1}^k \text{polar}(\mathcal{S}_i) = \text{polar}(\mathcal{S}) \quad (3)$$

where $\mathcal{S} = \bigcup_{i=1}^k \mathcal{S}_i$. The polar cone is finitely generated (since $\text{cl}(D^c)$ is assumed to have the form described in (1)). A frame for the polar cone can be found using the MATLAB toolbox MPT3 [22]. We conclude this section by showing that Algorithm 1 returns the correct value of cosine measure and the complete cosine vector set.

Theorem 25 *Given a set of non-descent directions contained in a closed half-space of $\mathbb{R}^n$ written in the form $\text{cl } D^c = \bigcup_{i=1}^k \overline{\text{pspan}}(\mathcal{S}_i)$ where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$, Algorithm 1 returns the correct cosine measure value and the complete cosine vector set.*

Proof Lemma 33 of [2] notes that if the cosine measure is nonpositive, then the convex relaxation of the cosine measure problem is exact, i.e., if $\text{cm}(\mathcal{S}) \leq 0$, then

$$\text{cm}(\mathcal{S}) = \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\| \leq 1}} \max_{\mathbf{d} \in \mathcal{S}} \mathbf{u}^\top \mathbf{d}. \quad (4)$$

Step 1 is a standard reformulation of Problem 4. Step (2.1) extracts $\text{cm}(D^c)$, given by t^* .

If $\text{cm}(D^c) = 0$, Step (2.2.2) returns the full cosine vector set of D^c by Proposition 24 and (3). If $\text{cm}(D^c) < 0$, Step (2.2.1) returns the unique cosine vector of D^c by Lemmas 22 and 23. $\square$

4.2 The case where $\text{cl } D^c$ is not contained in a closed half-space of $\mathbb{R}^n$

Next, we consider the case where $\text{cl } D^c$ is not contained in a closed half-space of $\mathbb{R}^n$. In this case, the cosine measure is strictly positive. To find the value of cosine measure and the cosine vector set, we will work from the *spherical complement* of $\text{cl } D^c$, $D^\circ = \{\mathbf{x} \in \mathbb{S}^n : \mathbf{x} \notin \text{cl } D^c\}$. In this case, we assume that the spherical complement of the closure of the non-descent directions can be written in the form

$$D^\circ = \mathbb{S}^n \cap \bigcup_{i=1}^k \text{int } \text{pspan}(\mathcal{S}_i), \quad (5)$$

where $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) = \{\mathbf{0}\}$ for all $i \neq j$. Intuitively, D° is providing a proxy for the interior of D relative to the unit sphere (it is not the interior, nor is it the relative interior, as both of those are empty). Notice that $\text{cl } D^c$ not contained in a closed half-space does not imply that D° is contained in a closed half-space. Indeed, the function and point given in Example 11

has $\text{cl } D^c = \{\mathbf{x} \in \mathbb{R}^2 : |x_1| \geq |x_2|, x_1^2 + x_2^2 = 1\}$ and $D^o = \{\mathbf{x} \in \mathbb{R}^2 : |x_1| < |x_2|, x_1^2 + x_2^2 = 1\}$, neither of which are contained in a closed half-space.

If $\text{cl } D^c$ is the entire unit sphere, then $\text{cm}(D^c) = 1$ by Theorem 13. This creates a special case in Algorithm 2 that we elaborate in the next lemma.

Lemma 26 *Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathbf{x} \in \text{int dom } f$. Then the following are equivalent:*

- (i) $\text{cl } D^c = \mathbb{S}^n$
- (ii) $D^o = \emptyset$,
- (iii) $\text{cm}(D^c) = 1$,
- (iv) $\text{cV}(D^c) = \mathbb{S}^n$.

Proof The result follows immediately from [2, Prop. 2]. $\square$

Algorithm 2 proposes a pseudo-code to compute the cosine measure and the cosine vector set for a function when $\text{cl } D^c$ is not contained in a closed half-space of $\mathbb{R}^n$.

Algorithm 2: Computing $\text{cm}(D^c)$ when D^c is not contained in a closed half-space

Given a set of non-descent directions not contained in a closed half-space of $\mathbb{R}^n$ with its spherical complement written in the form

$$D^o = \mathbb{S}^n \cap \bigcup_{i=1}^k \text{int pspan}(\mathcal{S}_i),$$

where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$ and $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) = \{\mathbf{0}\}$ for all $i \neq j$.

0. Special cases and initialization:

If $D^o = \emptyset$, return $\text{cm}(D^c) = 1$, $\text{cV}(D^c) = \mathbb{S}^n$.

Else define the index set of nontrivial $\mathcal{S}_i$, $J = \{i : \text{int pspan}(\mathcal{S}_i) \neq \emptyset\} \subseteq \{1, 2, \dots, k\}$.

1. Outer loop: For each $i \in J$ do the following.

Compute the set of positive unit normals to the $(n-1)$ -dim. faces of $\text{pspan}(\mathcal{S}_i)$

$$(1.1) \quad \mathcal{P}_i = \left\{ \mathbf{p} : \begin{array}{l} \|\mathbf{p}\| = 1, \mathbf{p}^\top \mathbf{d} \geq 0 \text{ for all } \mathbf{d} \in \mathcal{S}_i, \\ \dim(\text{span}(\mathbf{d} : \mathbf{p}^\top \mathbf{d} = 0, \mathbf{d} \in \mathcal{S}_i)) = n - 1 \end{array} \right\}.$$

Solve the second-order conic program (SOCP)

$$(1.2) \quad (t_i^*, \mathbf{u}_i^*) \in \arg \max_{t, \mathbf{u}} \{t : \|\mathbf{u}\| \leq 1, \mathbf{p}^\top \mathbf{u} \geq t \quad \forall \mathbf{p} \in \mathcal{P}_i\}$$

2. Compute the cosine measure and cosine vector set: return

$$(2.1) \quad \text{cm}(D^c) = \min_{i \in J} \sqrt{1 - (t_i^*)^2},$$

$$(2.2) \quad \text{cV}(D^c) = \left\{ \mathbf{u}_i^* : t_i^* = \max_{j \in J} t_j^*, \quad i \in J \right\}.$$

In words, Algorithm 2 considers the pointed polyhedral cones $C_i = \text{pspan}(\mathcal{S}_i)$. For each C_i , it finds the unit vector $\mathbf{u}_i^*$ which is as far as possible from the $(n-1)$ dimensional faces of C_i . To do that, the algorithm finds the positive normal vectors for all faces of C_i . Working with the set of positive normals $\mathcal{P}_i$ associated to C_i , the problem requires to find the unit vector $\mathbf{u}_i^*$ which is as close as possible (in angle) to the farthest positive normal vectors in $\mathcal{P}_i$. Once it is done for all pointed polyhedral cone C_i , the cosine measure is obtained by taking the minimum value of $\sin(\varphi_i)$ where $0 < \varphi_i < \pi/2$ is the angle between $\mathbf{u}_i^*$ and an active vector $\mathbf{p}_i \in \mathcal{P}_i$.

Before proving that Algorithm 2 returns the correct outputs, we make a few more remarks.

First, the assumption that $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) = \{\mathbf{0}\}$ is necessary to obtain the exact value of cosine measure. For example, suppose

$$D^o = \{(x_1, x_2) : 0 < x_1, -x_1 < x_2, \|\mathbf{x}\| = 1\}.$$

The desired description would use $\{(\mathbf{e}_1 - \mathbf{e}_2)/\sqrt{2}, \mathbf{e}_2\}$, while an undesirable description would use $\{(\mathbf{e}_1 - \mathbf{e}_2)/\sqrt{2}, (\mathbf{e}_1 + \mathbf{e}_2)/\sqrt{2}\}$ and $\{\mathbf{e}_1, \mathbf{e}_2\}$ (see Figure 3). The correct representation gives

$$\text{cm}(D^c) = \cos(3\pi/8) \approx 0.3827.$$

However, the “overlapping” description gives $\text{cm}(D^c) = \cos(\pi/4) \approx 0.7071$. If $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) \neq \{\mathbf{0}\}$ for some $i \neq j$, then the value of cosine measure may be overestimated.

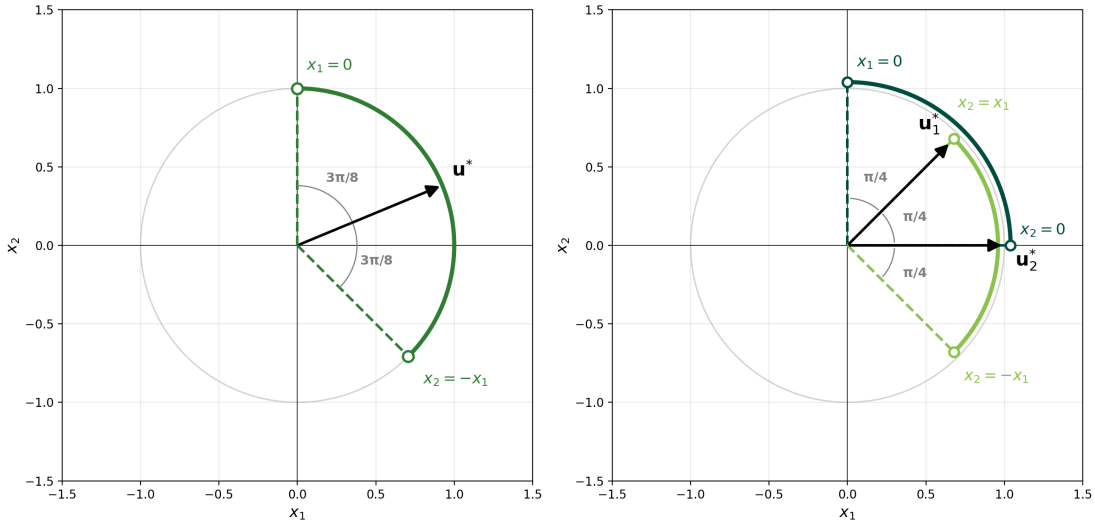


Fig. 3: On the left, the set $\{(x_1, x_2) : 0 < x_1, -x_1 < x_2, \|\mathbf{x}\| = 1\}$ is represented by one spherical cap. The cosine vector $\mathbf{u}^*$ is found using Algorithm 2. On the right, the set is decomposed into two overlapping spherical caps. A vector $\mathbf{u}^*$ is computed for each spherical cap. This yields an incorrect (overestimated) value of the cosine measure.

Examining the special cases (Step 0), we notice that if Algorithm 2 is applied to a single variable function (i.e., $n = 1$), then “ $\text{cl } D^c$ is not in a closed half-space” implies that $\text{cl } D^c = \mathbb{S}^1$, so the special case $D^o = \emptyset$ is applied. Also note, in Step 0, that if $J = \emptyset$, then the special case $D^o = \emptyset$ would have been triggered. Therefore, if the algorithm continues to Step 1, then we must have $n \geq 2$ and $J \neq \emptyset$.

Like Algorithm 1, if desired, each $\mathcal{S}_i$ could be reduced to its unique frame at the beginning of Step 1. This can simplify the process to identify all boundary faces of a pointed polyhedral cone.

In Step 1, the algorithm defines $\mathcal{P}_i$ to be the set of positive unit normal vectors to the $n - 1$ dimensional faces of $C_i = \text{pspan}(\mathcal{S}_i)$. To justify this, first notice that $\dim \text{pspan}(\mathcal{S}_i) = n$, as otherwise $\text{int } \text{pspan}(\mathcal{S}_i) = \emptyset$ and the set would have been ignored in Step 1. As $\dim \text{pspan}(\mathcal{S}_i) = n$, the set $\mathcal{P}_i$ can be created by taking subsets of $\mathcal{S}_i$ of size $n - 1$ using the (generalized) cross product to create $\tilde{\mathbf{p}}$ then checking if either $\mathbf{p} = \tilde{\mathbf{p}}/\|\tilde{\mathbf{p}}\|$ or $\mathbf{p} = -\tilde{\mathbf{p}}/\|\tilde{\mathbf{p}}\|$ satisfies $\mathbf{p}^\top \mathbf{d} \geq 0$. It is possible

to identify the same face of $\text{pspan}(\mathcal{S}_i)$ more than once by considering sets of $n - 1$ linearly independent vectors contained in $\mathcal{S}_i$. For example, consider the set $\mathcal{S}_i = \{\pm \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \subset \mathbb{R}^3$. The linearly independent sets $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $\{-\mathbf{e}_1, \mathbf{e}_2\}$ both find the unit normal vector $\mathbf{p} = \mathbf{e}_3$. Similarly, the linearly independent sets $\{\mathbf{e}_1, \mathbf{e}_3\}$ and $\{-\mathbf{e}_1, \mathbf{e}_3\}$ both find the unit normal vector $\mathbf{p} = \mathbf{e}_2$. As such, when constructing $\mathcal{P}_i$ it would be wise to implement a strategy to detect and remove repeated vectors. Notice that this process is identifying the facets of a polyhedral cone, which can grow exponential in dimension.

Step 1 of Algorithm 2 then requires solving a second-order conic program for each C_i . Note that the constraint $\|\mathbf{u}\| \leq 1$ in Step (1.2) relaxes the sphere constraint $\|\mathbf{u}\| = 1$. Since $t_i^* > 0$ for every pointed full-dimensional cone, scaling $\mathbf{u}$ toward the unit sphere strictly improves the objective function t . Hence, every optimal solution satisfies $\|\mathbf{u}_i^*\| = 1$. Furthermore, note that $\mathbf{u}_i^*$ must be in $\text{int pspan}(\mathcal{S}_i)$ since $\mathbf{p}^\top \mathbf{u}_i^* > 0$ for all $\mathbf{p} \in \mathcal{P}_i$ (this is used in Theorem 29).

We now focus on proving that Algorithm 2 returns the correct output.

Lemma 27 *Suppose D^c is a set of non-descent directions not contained in a closed half-space of $\mathbb{R}^n$ with its spherical complement written in the form*

$$D^\circ = \mathbb{S}^n \cap \bigcup_{i=1}^k \text{int pspan}(\mathcal{S}_i),$$

where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$ and $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) = \{\mathbf{0}\}$ for $i \neq j$. For each $i \in \{1, 2, \dots, k\}$ define $\mathcal{S}_i^\circ = \mathbb{S}^n \cap \text{int pspan}(\mathcal{S}_i)$. Then, for all i ,

$$0 < \text{cm}((\mathcal{S}_i^\circ)^c) \leq 1,$$

with $\text{cm}((\mathcal{S}_i^\circ)^c) = 1$ if and only if $\mathcal{S}_i^\circ = \emptyset$.

Proof Since D^c is not contained in a closed half-space, it follows from Theorem 13 that $0 < \text{cm}(D^c) \leq \text{cm}((\mathcal{S}_i^\circ)^c)$ for all i . The second statement follows from Lemma 26. $\square$

Next, we show that the cosine measure of D^c is obtained by considering the appropriate single set $\mathcal{S}_i$. This greatly simplifies the remaining proofs.

Lemma 28 *Suppose D^c is a set of non-descent directions not contained in a closed half-space of $\mathbb{R}^n$ such that $\text{cl } D^c \neq \mathbb{S}^n$ and with its spherical complement written in the form*

$$D^\circ = \mathbb{S}^n \cap \bigcup_{i=1}^k \text{int pspan}(\mathcal{S}_i),$$

where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$ and $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) = \{\mathbf{0}\}$ for $i \neq j$. For each $i \in \{1, 2, \dots, k\}$ define $\mathcal{S}_i^\circ = \mathbb{S}^n \cap \text{int pspan}(\mathcal{S}_i)$. Then there exists $i \in \{1, 2, \dots, k\}$ such that

$$\text{cm}(D^c) = \text{cm}((\mathcal{S}_i^\circ)^c).$$

Proof Note that $\mathcal{S}_i^\circ$ is non-empty for at least one i since $\text{cl}(D^c) \neq \mathbb{S}^n$. Without loss of generality, assume that $\mathcal{S}_i^\circ \neq \emptyset$ for all $i \in \{1, 2, \dots, k\}$. Since $D^c \subseteq (\mathcal{S}_i^\circ)^c$ for every i , we get $\text{cm}(D^c) \leq \text{cm}((\mathcal{S}_i^\circ)^c)$ for all i .

For the reverse inequality, let $\mathbf{u}^* \in \text{cV}(D^c)$ and define $\gamma = \text{cm}(D^c) < 1$. Consider the open spherical cap

$$\mathcal{C} = \{\mathbf{d} \in \mathbb{S}^n : \mathbf{d}^\top \mathbf{u}^* > \gamma\}.$$

Since $\gamma < 1$, we must have $\mathbf{u}^* \in \mathcal{C}$, so $\mathcal{C} \neq \emptyset$. By definition of $\text{cm}(D^c)$, every $\mathbf{d} \in D^c$ satisfies $\mathbf{d}^\top \mathbf{u}^* \leq \gamma$. Hence, $\mathcal{C} \cap D^c = \emptyset$ and it follows $\mathcal{C} \subseteq \bigcup_{i=1}^k \mathcal{S}_i^\circ$. Since $\text{pspan}(\mathcal{S}_i) \cap \text{pspan}(\mathcal{S}_j) = \{\mathbf{0}\}$ for $i \neq j$, and $\mathcal{C}$ is connected, there exists a unique index i^* such that $\mathcal{C} \subseteq \mathcal{S}_{i^*}^\circ$. Therefore, every $\mathbf{d} \in (\mathcal{S}_{i^*}^\circ)^c$ satisfies $\mathbf{d}^\top \mathbf{u}^* \leq \gamma$. This implies

$$\text{cm}((\mathcal{S}_{i^*}^\circ)^c) \leq \max_{\mathbf{d} \in (\mathcal{S}_{i^*}^\circ)^c} \mathbf{d}^\top \mathbf{u}^* \leq \gamma = \text{cm}(D^c).$$

□

Lemma 28 allows us to simplify the remaining proofs, as it allows us to consider each $\text{pspan}(\mathcal{S}_i)$ one at a time. That is, it suffices to show that $(t_i^*, \mathbf{u}_i^*)$, from Step 1 of Algorithm 2, provides the cosine measure and a cosine vector for $(\mathcal{S}_i^\circ)^c$. In Step 2, we simply take the minimum of $\sqrt{1 - (t_i^*)^2}$ over i to obtain the cosine measure of D^c .

Theorem 29 Suppose D^c is a set of non-descent directions not contained in a closed half-space of $\mathbb{R}^n$ with its spherical complement written in the form

$$D^c = \mathbb{S}^n \cap \bigcup_{i=1}^k \text{int pspace}(\mathcal{S}_i),$$

where each $\mathcal{S}_i$ is a finite set of nonzero vectors in $\mathbb{R}^n$ and $\text{pspace}(\mathcal{S}_i) \cap \text{pspace}(\mathcal{S}_j) = \{\mathbf{0}\}$ for $i \neq j$.

For each $i \in \{1, 2, \dots, k\}$ define $\mathcal{S}_i^\circ = \mathbb{S}^n \cap \text{int pspace}(\mathcal{S}_i)$. Let $\mathcal{P}_i \subset \mathbb{R}^n$ be the finite set containing the unit normal vector $\mathbf{p}_{j_i}$ for each boundary hyperplane H_{j_i} of $\text{pspace}(\mathcal{S}_i)$ such that $\mathbf{p}_{j_i}^\top \mathbf{x} \geq 0$ for all $\mathbf{x} \in \text{pspace}(\mathcal{S}_i)$ and for all $j_i \in \{1, 2, \dots, m_i\}, i \in \{1, 2, \dots, k\}$. If $\mathcal{S}_i^\circ \neq \emptyset$, then

$$\text{cm}((\mathcal{S}_i^\circ)^c) = \max_{\mathbf{p} \in \mathcal{P}_i} \sqrt{1 - (\mathbf{p}^\top \mathbf{u}_i^*)^2},$$

where $\mathbf{u}_i^*$ is the unique unit vector solving the SOCP in Step (1.2) of Algorithm 2. Consequently, Algorithm 2 returns the cosine measure and the cosine vector set for D^c .

Proof Consider i such that $\mathcal{S}_i^\circ \neq \emptyset$. Note that this implies $\text{cm}((\mathcal{S}_i^\circ)^c) < 1$. Define the normalized frontier of $\text{pspace}(\mathcal{S}_i)$ via

$$\delta \text{pspace}(\mathcal{S}_i) = (\text{cl pspace}(\mathcal{S}_i) \setminus \text{int pspace}(\mathcal{S}_i)) \cap \mathbb{S}^n.$$

Note that

$$\begin{aligned} \text{cm}((\mathcal{S}_i^\circ)^c) &= \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \sup_{\mathbf{d} \in (\mathcal{S}_i^\circ)^c} \mathbf{d}^\top \mathbf{u} \\ &= \min_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \max_{\mathbf{d} \in \text{cl}((\mathcal{S}_i^\circ)^c)} \mathbf{d}^\top \mathbf{u} \\ &= \min_{\mathbf{u} \in \mathcal{S}_i^\circ} \max_{\mathbf{d} \in \text{cl}((\mathcal{S}_i^\circ)^c)} \mathbf{d}^\top \mathbf{u} && (\text{since } \text{cm}(D^c) < 1, \text{ so } \mathbf{u} \text{ cannot be in } (\mathcal{S}_i^\circ)^c) \\ &= \min_{\mathbf{u} \in \mathcal{S}_i^\circ} \max_{\substack{\mathbf{d} \in \text{cl pspace}(\mathcal{S}_i) \setminus \text{int pspace}(\mathcal{S}_i) \\ \|\mathbf{d}\|=1}} \mathbf{d}^\top \mathbf{u} \\ &= \min_{\mathbf{u} \in \mathcal{S}_i^\circ} \max_{\mathbf{d} \in \delta \text{pspace}(\mathcal{S}_i)} \mathbf{d}^\top \mathbf{u}, \end{aligned} \tag{6}$$

since the maximum is attained by a unit vector on the boundary of $\text{pspace}(\mathcal{S}_i)$.

Let $\mathbf{u}_i^* \in \mathcal{S}_i^\circ$ be a solution to (6) and let $H_{1_i}, \dots, H_{m_i}$ be the boundary hyperplanes of $\text{pspan}(\mathcal{S}_i)$. Denote the projection of the vector $\mathbf{u}_i^*$ onto the hyperplane H_{j_i} by $\text{Proj}_{H_{j_i}} \mathbf{u}_i^*$. Also, denote a boundary hyperplane containing a unit vector $\mathbf{d} \in \mathbb{R}^n$ by $H_{\mathbf{d}}$. We have

$$\begin{aligned} 0 &< \max_{\mathbf{d} \in \delta \text{pspan}(\mathcal{S}_i)} \mathbf{d}^\top \mathbf{u}_i^* = \max_{\mathbf{d} \in \delta \text{pspan}(\mathcal{S}_i)} \mathbf{d}^\top (\text{Proj}_{H_{\mathbf{d}}} \mathbf{u}_i^*) \\ &= \max_{j_i \in \{1, 2, \dots, m_i\}} \left(\frac{\text{Proj}_{H_{j_i}} \mathbf{u}_i^*}{\|\text{Proj}_{H_{j_i}} \mathbf{u}_i^*\|} \right)^\top \text{Proj}_{H_{j_i}} \mathbf{u}_i^* \\ &= \max_{j_i \in \{1, 2, \dots, m_i\}} \|\text{Proj}_{H_{j_i}} \mathbf{u}_i^*\|. \end{aligned}$$

Note that $\text{Proj}_{H_{j_i}} \mathbf{u}_i^*$ need not lie in $\text{pspan}(\mathcal{S}_i)$. However, the maximum of $\|\text{Proj}_{H_{j_i}} \mathbf{u}_i^*\|$ over j_i is attained at a hyperplane realizing the minimal distance from $\mathbf{u}_i^*$, for which $\text{Proj}_{H_{j_i}} \mathbf{u}_i^* \in \text{pspan}(\mathcal{S}_i)$. Next, note that

$$\max_{j_i \in \{1, 2, \dots, m_i\}} \|\text{Proj}_{H_{j_i}} \mathbf{u}_i^*\| = \max_{j_i \in \{1, 2, \dots, m_i\}} \sin(\varphi_{\mathbf{p}_{j_i}})$$

where $0 < \varphi_{\mathbf{p}_{j_i}} < \pi/2$ is the angle between the positive normal vector $\mathbf{p}_{j_i}$ associated to the hyperplane H_{j_i} and the vector $\mathbf{u}_i^*$. Since $(\mathbf{u}_i^*)^\top \mathbf{p}_{j_i} = \cos(\varphi_{\mathbf{p}_{j_i}})$ and $\sin(\varphi_{\mathbf{p}_{j_i}}) = \sqrt{1 - \cos^2(\varphi_{\mathbf{p}_{j_i}})}$, we obtain

$$\text{cm}((\mathcal{S}_i^\circ)^c) = \max_{\mathbf{p} \in \mathcal{P}_i} \sqrt{1 - (\mathbf{p}^\top \mathbf{u}_i^*)^2}.$$

Now, since $\mathcal{P}_i$ is a non-empty compact set, and $\phi(t) = \sqrt{1 - t^2}$ is a strictly decreasing continuous function on $0 < t < 1$, we have

$$\min_{\mathbf{u} \in \mathcal{S}_i^\circ} \max_{\mathbf{p} \in \mathcal{P}_i} \sqrt{1 - (\mathbf{p}^\top \mathbf{u})^2} = \sqrt{1 - \left(\max_{\mathbf{u} \in \mathcal{S}_i^\circ} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u} \right)^2}.$$

Moreover, the two problems have identical solution sets

$$\arg \min_{\mathbf{u} \in \mathcal{S}_i^\circ} \max_{\mathbf{p} \in \mathcal{P}_i} \sqrt{1 - (\mathbf{p}^\top \mathbf{u})^2} = \arg \max_{\mathbf{u} \in \mathcal{S}_i^\circ} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u}.$$

Next, we prove that

$$\max_{\mathbf{u} \in \mathcal{S}_i^\circ} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u} = \max_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u}. \quad (7)$$

Indeed, since $\mathcal{S}_i^\circ$ is a subset of $\mathbb{S}^n$, we get

$$\max_{\mathbf{u} \in \mathcal{S}_i^\circ} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u} \leq \max_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u}.$$

Let $f(\mathbf{u}) = \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u}$ and $\mathbf{u}^*$ be a unit vector achieving the maximum of the right hand side of (7). Then for any $\tilde{\mathbf{u}} \in \mathcal{S}_i^\circ$, we have $f(\mathbf{u}^*) \geq f(\tilde{\mathbf{u}}) > 0$. Note that $f(\mathbf{u}^*) > 0$ implies $\mathbf{u}^*$ is in $\mathcal{S}_i^\circ$. We get

$$\max_{\tilde{\mathbf{u}} \in \mathcal{S}_i^\circ} f(\tilde{\mathbf{u}}) \geq f(\mathbf{u}^*) = \max_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} f(\mathbf{u}).$$

Besides, since the optimal value of $f(\mathbf{u})$ is positive, every maximizer over $\mathbb{S}^n$ satisfies $\mathbf{p}^\top \mathbf{u}^* > 0$. Hence, both sides of (7) have the same solution set.

Finally, using similar arguments as in Lemmas 22, 23 and Lemma 33 in [2], we obtain that the convex relaxation of the problem

$$\max_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\|=1}} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u} \quad (8)$$

may be considered:

$$\max_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\| \leq 1}} \min_{\mathbf{p} \in \mathcal{P}_i} \mathbf{p}^\top \mathbf{u}. \quad (9)$$

The convex relaxation returns the unique unit vector solving the nonconvex program (8). Step (1.2) of Algorithm 2 is a reformulation of the convex program (9). In the case where D° consists of several polyhedral cones $\text{pspan}(\mathcal{S}_i)$, the correctness of Step 2 follows from Lemma 28. $\square$

Remark 30 Given a full-dimensional pointed polyhedral cone $C = \text{pspan}(\mathcal{S}) \subsetneq \mathbb{R}^n$, we could define the *sine measure of C* (denoted by $\text{sm}(C)$) by

$$\text{sm}(C) = \min_{\substack{\mathbf{u} \in C \\ \|\mathbf{u}\|=1}} \max_{\mathbf{p} \in \mathcal{P}} \sqrt{1 - (\mathbf{p}^\top \mathbf{u})^2} = \min_{\substack{\mathbf{u} \in C \\ \|\mathbf{u}\|=1}} \max_{\mathbf{p} \in \mathcal{P}} \sin(\varphi) = \sqrt{1 - \left(\max_{\substack{\mathbf{u} \in \mathbb{R}^n \\ \|\mathbf{u}\| \leq 1}} \min_{\mathbf{p} \in \mathcal{P}} \mathbf{p}^\top \mathbf{u} \right)^2} \quad (10)$$

where $0 < \varphi < \pi/2$ is the angle between $\mathbf{u}$ and $\mathbf{p}$, and where $\mathcal{P}$ is the finite set of positive unit normal vectors for the boundary hyperplanes of C . That is, find the unit vector in C that is as far as possible from its nearest boundary hyperplane. Equivalently, in terms of $\mathcal{P}$, find the unit vector which is as close as possible (in angle) to the farthest normal in $\mathcal{P}$. A unit vector $\mathbf{u}$ solving the sine measure problem is called the *sine vector of C*. It is worth mentioning that the sine vector of C is equivalent to the *incenter* of a *proper cone* K denoted by $\pi_{\text{inc}(K)}$ and introduced in [20, 21].

Under this terminology, Step (1.2) is the second-order conic program to solve the sine measure problem for each $\text{pspan}(\mathcal{S}_i)$. From Theorem 29, each $\mathbf{u}_i^*$ found while solving the SOCP is a unit vector in $\text{int psan}(\mathcal{S}_i)$. However, it should be noted that even though the problem is an SOCP, the number of constraints for this problem grows with the number of boundary hyperplanes to C . Indeed, determining the constraints for the SOCP is equivalent to identifying the faces of a pointed polyhedral cone from its extreme rays (which can grow exponentially).

5 Computing the cosine measure of the finite-max function

In this section, we consider a nonsmooth function of the form

$$F = \max\{f_1, f_2, \dots, f_\ell\}, \quad \text{where } f_i \in \mathcal{C}^1 \quad \text{for all } i \in \{1, 2, \dots, \ell\}. \quad (11)$$

We will see that computing the cosine measure of F at $\mathbf{x}$ relies on the active gradients. Indeed, the set $\mathcal{P}$ defined in Step (1.1) of Algorithm 2 is exactly the negative of the set of normalized active gradients at the point $\mathbf{x}$. A particular type of finite-max function, the ℓ_1 norm is considered to conclude this section. We begin by recalling the active index set at $\mathbf{x}$.

Definition 31 (Active set of indices) Let $F : \mathbb{R}^n \rightarrow \mathbb{R}$ be defined by

$$F(\mathbf{x}) = \max_{i \in \{1, 2, \dots, \ell\}} f_i(\mathbf{x}),$$

where $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is in $\mathcal{C}^1$. The *active index set* of F at $\mathbf{x}$ is denoted by $I(\mathbf{x})$ and defined by

$$I(\mathbf{x}) = \{i \in \{1, 2, \dots, \ell\} : f_i(\mathbf{x}) = F(\mathbf{x})\}.$$

Note that $I(\mathbf{x})$ is non-empty for any $\mathbf{x}$.

The set of non-descent directions of F at $\mathbf{x}$ is given by

$$D^c = \bigcup_{i \in I(\mathbf{x})} \{\mathbf{d} : \nabla f_i(\mathbf{x})^\top \mathbf{d} \geq 0, \|\mathbf{d}\| = 1\}.$$

The spherical complement of $\text{cl } D^c$ in $\mathbb{S}^n$ is equal to

$$D^o = \left\{ \bigcap_{i \in I(\mathbf{x})} \{\mathbf{d} : \nabla f_i(\mathbf{x})^\top \mathbf{d} < 0, \|\mathbf{d}\| = 1\} \right\} \cap \mathbb{S}^n. \quad (12)$$

When $|I(\mathbf{x})| = 1$ and the gradient of the active function is not the zero vector, by Theorem 14, we obtain that $\text{cm}(D^c) = 0$ and $\text{cV}(D^c) = -\nabla f_{i^*}(\mathbf{x}) / \|\nabla f_{i^*}(\mathbf{x})\|$ where $I(\mathbf{x}) = \{i^*\}$.

If one of the active gradient vectors is equal to the zero vector, then the set of non-descent directions of F at $\mathbf{x}$ can take any form. In particular, examples may be created (as we did in Section 3.1) showing that in this case the cosine measure of F at $\mathbf{x}$ may take any value in $[-1, 1]$. For this reason, we assume that the set of active gradient vectors does not contain the zero vector in the following algorithm.

Algorithm 3: Computing $\text{cm}(D^c)$ of a max function

Given a function of the form $F = \max\{f_1, f_2, \dots, f_\ell\}$, where $f_i \in \mathcal{C}^1$ for all $i \in \{1, 2, \dots, \ell\}$ and a point $\mathbf{x} \in \mathbb{R}^n$.

0. Special cases and initialization:

Define the set of active indices $I(\mathbf{x}) = \{i \in \{1, 2, \dots, \ell\} : f_i(\mathbf{x}) = F(\mathbf{x})\}$.

If $\nabla f_i(\mathbf{x}) = \mathbf{0}$ for any $i \in I(\mathbf{x})$, stop (cosine measure cannot be computed).

If $|I(\mathbf{x})| = 1$, return $\text{cm}(D^c) = 0$, $\text{cV}(D^c) = -\nabla f_{i^*}(\mathbf{x}) / \|\nabla f_{i^*}(\mathbf{x})\|$, where $\{i^*\} = I(\mathbf{x})$.

1. SOCP set up and solution:

Define the set of active negative gradients of F at $\mathbf{x}$:

$$(1.1) \quad \mathcal{P} = \left\{ -\frac{\nabla f_i(\mathbf{x})}{\|\nabla f_i(\mathbf{x})\|} : i \in I(\mathbf{x}) \right\}$$

Solve the second-order conic program (SOCP)

$$(1.2) \quad (t^*, \mathbf{u}^*) \in \arg \max_{t, \mathbf{u}} \{t : \|\mathbf{u}\| \leq 1, \mathbf{p}^\top \mathbf{u} \geq t \quad \forall \mathbf{p} \in \mathcal{P}\}$$

2. Compute the cosine measure and cosine vector set: return

$$(2.1) \quad \text{cm}(D^c) = \sqrt{1 - (t^*)^2},$$

$$(2.2) \quad \text{cV}(D^c) = \begin{cases} \{\mathbf{u}^*\} & \text{if } t^* > 0, \\ \mathbb{S}^n & \text{if } t^* = 0. \end{cases}$$

Notice that Algorithm 3 is Algorithm 2 with a few simplifications. The set of positive normals is now given by the set of normalized active negative gradients at $\mathbf{x}$. The outer loop is deleted since the set of descent directions of F at $\mathbf{x}$ can be represented by a single polyhedral cone C (intersecting with $\mathbb{S}^n$) as described in (12). Note that it is possible for C to be empty. For example, consider $f_1 : \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} \mapsto x_1, f_2 : \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} \mapsto -x_1$ at the point $\mathbf{x} = \mathbf{0}$. Then $F = |x_1|$ and we have $\nabla f_1(\mathbf{0}) = \mathbf{e}_1$ and $\nabla f_2(\mathbf{0}) = -\mathbf{e}_1$. In this case, the SOCP returns $t = 0$ and we get $\text{cm}(D^c) = 1, \text{cV}(D^c) = \mathbb{S}^2$. Thus we have the following corollary.

Corollary 32 *Given a function of the form $F = \max\{f_1, f_2, \dots, f_\ell\}$, where $f_i \in \mathcal{C}^1$ for all $i \in \{1, 2, \dots, \ell\}$ and a point $\mathbf{x} \in \mathbb{R}^n$, Algorithm 3 returns the correct cosine measure and cosine vector set.*

We end with two examples demonstrating Algorithm 3.

Example 33 Consider two functions with domain $\mathbb{R}^2$ defined by

$$f_1(\mathbf{x}) = x_1 + x_2, \quad f_2(\mathbf{x}) = -x_1 + x_2.$$

Let $F(\mathbf{x}) = \max\{x_1 + x_2, -x_1 + x_2\} = |x_1| + x_2$, and consider the point $\mathbf{x}_0 = \mathbf{0}$. From the contour plot (see Figure 4), it is easy to confirm that $\text{cm}(F; \mathbf{0}) = 1/\sqrt{2}$ and $\text{cV}(F; \mathbf{0}) = \{[0 \ -1]^\top\}$.

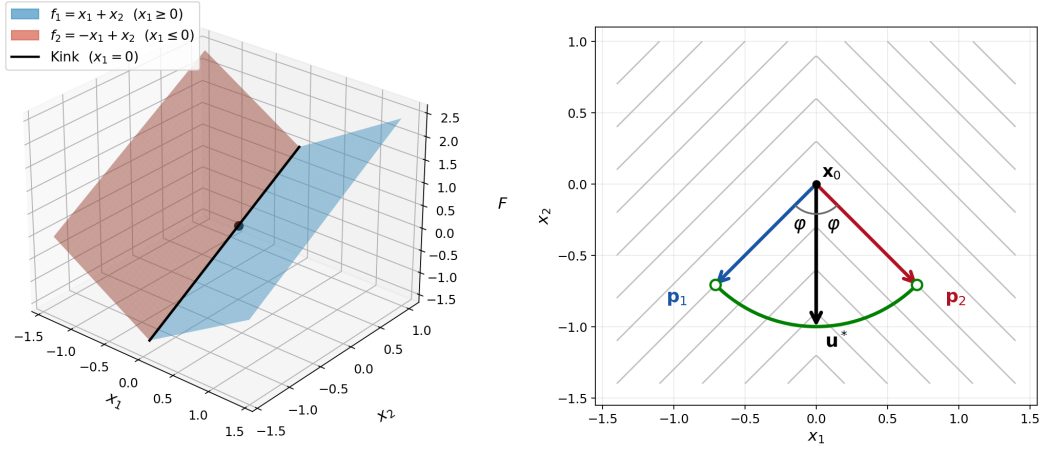


Fig. 4: Graph and the descent directions of $F = \max\{x_1 + x_2, -x_1 + x_2\}$ at $\mathbf{x}_0 = \mathbf{0}$ (represented by the green spherical cap).

The active set of indices at $\mathbf{x}_0 = \mathbf{0}$ is $I(\mathbf{x}_0) = \{1, 2\}$, so the normalized active negative gradient set is

$$\mathcal{P} = \left\{ \mathbf{p}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \mathbf{p}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

Solving the SOCP,

$$\begin{aligned} \max_{t, \mathbf{u}} \left\{ t : \begin{array}{l} \|\mathbf{u}\| \leq 1 \\ \mathbf{p}^\top \mathbf{u} \geq t \quad \forall \mathbf{p} \in \mathcal{P} \end{array} \right\} &= \max_{t, \mathbf{u}} \left\{ t : \begin{array}{l} \|\mathbf{u}\| \leq 1 \\ -u_1/\sqrt{2} - u_2/\sqrt{2} \geq t \\ u_1/\sqrt{2} - u_2/\sqrt{2} \geq t \end{array} \right\} \\ &= \max_{t, \mathbf{u}} \left\{ t : \begin{array}{l} \|\mathbf{u}\| \leq 1 \\ -u_2 \geq \sqrt{2}t \end{array} \right\}, \end{aligned}$$

we obtain

$$\text{cm}(F; \mathbf{0}) = \sqrt{1 - (1/\sqrt{2})^2} = 1/\sqrt{2} \quad \text{and} \quad \mathbf{u}^* = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \text{cV}(F; \mathbf{0}).$$

In Example 33, the cosine vector $\mathbf{u}^*$ is also the direction of steepest descent. In general, this is not necessarily true. For example, considering $F(\mathbf{x}) = \max\{x_1, 2x_2\}$ at the point $\mathbf{x}_0 = \mathbf{0}$, we find $\mathbf{u}^* = -\frac{1}{\sqrt{2}} [1 \ 1]^\top$ (and $\text{cm}(D^c) = 1/\sqrt{2}$) but the direction of steepest descent, denoted by $\mathbf{u}_{sd}$, is $-\frac{1}{\sqrt{5}} [2 \ 1]^\top$. Figure 5 illustrates both vectors on the contour plot of F .

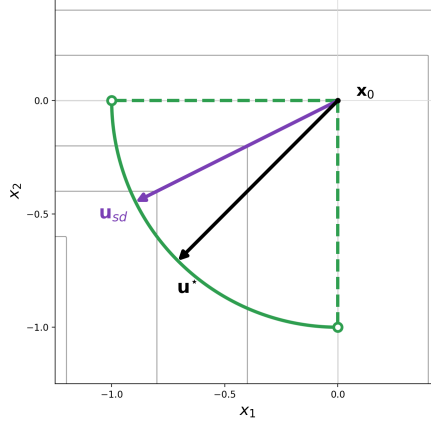


Fig. 5: The cosine vector $\mathbf{u}^*$ of F at $\mathbf{x}_0 = \mathbf{0}$ is not equal to the steepest descent direction $\mathbf{u}_{sd}$

Our second example considers the common regularization function ℓ_1 . Recall, in $\mathbb{R}^n$,

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| = \max_{\boldsymbol{\sigma} \in \{-1,1\}^n} \boldsymbol{\sigma}^\top \mathbf{x},$$

where the notation $\{-1,1\}^n$ means that each entry of the n -dimensional vector $\boldsymbol{\sigma}$ is either -1 or 1 . Therefore, $\|\mathbf{x}\|_1$ is the maximum of 2^n linear functions. Each linear function has the form $f_{\boldsymbol{\sigma}}(\mathbf{x}) = \boldsymbol{\sigma}^\top \mathbf{x} \in \mathcal{C}^1$ and gradient equal to $\nabla f_{\boldsymbol{\sigma}} = \boldsymbol{\sigma} \neq \mathbf{0}_n$. Algorithm 3 provides us the tools to compute the cosine measure of this function as a simple formula of the number of zeros in the point $\mathbf{x}_0$.

Theorem 34 (Cosine Measure of the ℓ_1 norm at $\mathbf{x}_0$) Let $F(\mathbf{x}) = \|\mathbf{x}\|_1$ and let $\mathbf{x}_0 \in \mathbb{R}^n, n \geq 2$. Let m be the number of 0s in $\mathbf{x}_0$. Then

$$\text{cm}(F; \mathbf{x}_0) = \sqrt{\frac{m}{n}} \quad \text{and} \quad \text{cV}(F; \mathbf{x}_0) = \begin{cases} \mathbb{S}^n & \text{if } m = n \\ \frac{-\mathbf{s}}{\sqrt{n-m}} & \text{if } m < n, \end{cases}$$

where $\mathbf{s} \in \{-1,0,1\}^n$ is defined by

$$s_j = \text{sgn}((\mathbf{x}_0)_j) = \begin{cases} +1 & (\mathbf{x}_0)_j > 0, \\ 0 & (\mathbf{x}_0)_j = 0, \\ -1 & (\mathbf{x}_0)_j < 0, \end{cases} \quad j \in \{1, 2, \dots, n\}. \quad (13)$$

Proof If $m = n$, then $\mathbf{x}_0 = \mathbf{0}$ and the result follows from Theorem 13.

Suppose $m < n$ and define the nonzero and zero index sets respectively by

$$N = \{j : (\mathbf{x}_0)_j \neq 0\}, \quad Z = \{j : (\mathbf{x}_0)_j = 0\}, \quad |N| = n - m, \quad |Z| = m.$$

The active negative gradient set at $\mathbf{x}_0$ can be written as

$$\tilde{\mathcal{P}} = \left\{ \mathbf{s} : \begin{cases} \sigma_i = -s_i & i \in N, \\ \sigma_i = \pm 1 & i \in Z \end{cases} \right\}.$$

Since $\|\mathbf{s}\| = \sqrt{n}$, the normalized active negative gradient set at $\mathbf{x}_0$ can be written as

$$\mathcal{P} = \left\{ \mathbf{p} : \begin{array}{ll} p_i = -s_i/\sqrt{n} & i \in N, \\ p_i = \pm 1/\sqrt{n} & i \in Z \end{array} \right\}.$$

Thus, the SOCP in Algorithm 3 becomes

$$\begin{aligned} & \max_{t, \mathbf{u}} \{ t : \|\mathbf{u}\| \leq 1, \mathbf{p}^\top \mathbf{u} \geq t \quad \forall \mathbf{p} \in \mathcal{P} \} \\ &= \max_{t, \mathbf{u}} \left\{ t : \|\mathbf{u}\| \leq 1, -\sum_{i \in N} s_i u_i - \sum_{i \in Z} |u_i| \geq \sqrt{n}t \right\} \\ &= \max_{t, \mathbf{u}} \left\{ t : \|\mathbf{u}\| \leq 1, -\sum_{i \in N} s_i u_i \geq \sqrt{n}t, \quad u_i = 0 \quad \forall i \in Z \right\}. \end{aligned}$$

This is solved when

$$u_i^* = \begin{cases} \frac{-s_i}{\sqrt{n-m}} & \text{if } i \in N \\ 0 & \text{if } i \in Z \end{cases} \quad \text{and} \quad t^* = \frac{\sqrt{n-m}}{\sqrt{n}}.$$

Applying Step 2 of Algorithm 3, we find

$$\text{cm}(F; \mathbf{x}_0) = \sqrt{1 - (t^*)^2} = \sqrt{\frac{m}{n}} \quad \text{and} \quad \text{cV}(F; \mathbf{x}_0) = \begin{cases} \mathbb{S}^n & \text{if } m = n \\ \frac{-\mathbf{s}}{\sqrt{n-m}} & \text{if } m < n, \end{cases}.$$

□

Combining Theorem 34 with Theorem 10, we gain insight regarding the cosine measure required to ensure a set contains a descent direction of the ℓ_1 norm at a given point.

Corollary 35 *Let $F : \mathbb{R}^n \rightarrow \mathbb{R} : \mathbf{x} \mapsto \|\mathbf{x}\|_1, n \geq 2$. Suppose $\mathbf{x}_0$ has m zero entries. If E is a set of vectors in $\mathbb{R}^n$ with cosine measure $\text{cm}(E) > \frac{\sqrt{m}}{\sqrt{n}}$, then E contains a descent direction for F at $\mathbf{x}_0$. (If $m = n$, this is impossible.)*

6 Conclusion

Several future research directions have emerged from this work. An analysis of the finite-max function using approximate gradients to find the cosine measure value could be explored. In practice, the exact gradients are usually unknown and using approximation of the gradients might be necessary. Gradients can be approximated using *generalized simplex gradients* [13, 17] and bounds between the exact cosine measure and the approximate cosine measure could be developed. Another obvious research directions is to investigate the cosine measure of other classical nonsmooth functions. For instance, the ℓ_1 norm function of $A\mathbf{x} - \mathbf{b}$, the maximum eigenvalue function, the nuclear norm function and the Hinge loss are reasonable choices.

The cardinality of active set of vectors for a sine vector (or a cosine vector) is still unclear. It is known that the active set may not contain a basis of $\mathbb{R}^n$ in general. Depending on the properties of the set D^c , it might be possible to develop results on the cardinality of the active set.

Another important topic to investigate is the cosine measure of a set $\text{pspan}(\mathcal{S})$ such that $\text{pspan}(\mathcal{S})$ has an infinite frame. In this paper, we have assumed that all frames are finite. However,

this is not always the case. One of the most famous examples of an infinite frame is the *Lorentz cone*: in $\mathbb{R}^3$, $\mathcal{L} = \text{pspan}(\mathcal{S})$ where $\mathcal{S} = \{\mathbf{v} \in \mathbb{R}^3 : \mathbf{v} = [\cos(\theta) \sin(\theta) 1]^\top, \theta \in [0, 2\pi)\}$. In some cases, it might be possible to develop results related to cosine measure. Beyond such convex cones, nonconvex semialgebraic sets with explicitly characterized geometric structure provide another promising source of infinite directional families. In particular, bilinear constraint sets, including hyperbolas, possess continuously varying tangent and normal directions that could naturally be studied from the perspective of cosine measures [3, 27]. These examples may help clarify when cosine measures associated with infinite families are computable and how the geometry of the underlying set influences their values and active directions.

Acknowledgments The authors would like to thank Manish Krishan Lal for his valuable comments and helpful suggestions. The authors would also like to thank Dana Abukhalaf for her assistance in proving Proposition 24.

References

1. C. AUDET, *A short proof on the cardinality of maximal positive bases*, Optimization Letters, 5 (2011), pp. 191–194.
2. C. AUDET, W. HARE, AND G. JARRY-BOLDUC, *The cosine measure relative to a subspace*, 2024.
3. H. BAUSCHKE, M. KRISHAN LAL, AND X. WANG, *Projections onto hyperbolas or bilinear constraint sets in hilbert spaces*, Journal of Global Optimization, 86 (2023), pp. 25–36.
4. J. BURKE, A. LEWIS, AND M. OVERTON, *Approximating subdifferentials by random sampling of gradients*, Mathematics of Operations Research, 27 (2002), pp. 567–584.
5. B. CAVARRETTA, F. GOYENS, C. ROYER, AND F. YGER, *Complexity guarantees and polling strategies for riemannian direct-search methods*, 2025.
6. A. CONN, K. SCHEINBERG, AND L. VICENTE, *Introduction to derivative-free optimization*, vol. 8, Siam, 2009.
7. ———, *Introduction to Derivative-Free Optimization*, SIAM, 2009.
8. D. CORNAZ, S. KERLEAU, AND C. ROYER, *A characterization of positive spanning sets with ties to strongly connected digraphs*, 2024.
9. C. DAVIS, *Theory of positive linear dependence*, American Journal of Mathematics, 76 (1954), pp. 733–746.
10. M. DODANGEH, L. VICENTE, AND Z. ZHANG, *On the optimal order of worst case complexity of direct search*, Optimization Letters, 10 (2016), pp. 699–708.
11. D. DRUSVYATSKIY, A. IOFFE, AND A. LEWIS, *Clarke subgradients for directionally Lipschitzian stratifiable functions*, Mathematics of Operations Research, 40 (2015), pp. 328–349.
12. W. HARE, *Compositions of convex functions and fully linear models*, Optimization Letters, 11 (2017), pp. 1217–1227.
13. W. HARE AND G. JARRY-BOLDUC, *Calculus identities for generalized simplex gradients: Rules and applications*, SIAM Journal on Optimization, 30 (2020), pp. 853–884.
14. ———, *A deterministic algorithm to compute the cosine measure of a finite positive spanning set*, Optimization Letters, 14 (2020), pp. 1305–1316.
15. W. HARE, G. JARRY-BOLDUC, S. KERLEAU, AND C. ROYER, *Using orthogonally structured positive bases for constructing positive k -spanning sets with cosine measure guarantees*, Linear Algebra and its Applications, 680 (2024), pp. 183–207.
16. W. HARE, G. JARRY-BOLDUC, AND C. PLANIDEN, *Nicely structured positive bases with maximal cosine measure*, Optimization Letters, (2023), pp. 1–21.
17. W. HARE, G. JARRY-BOLDUC, AND C. PLANIDEN, *Error bounds for overdetermined and underdetermined generalized centred simplex gradients*, IMA Journal of Numerical Analysis, 42 (2020), pp. 744–770.
18. W. HARE AND H. SONG, *On the cardinality of positively linearly independent sets*, Optimization Letters, 10 (2016), pp. 649–654.
19. W. HARE AND S. SUN, *On the computation of the cosine measure in high dimensions*, 2025.
20. R. HENRION AND A. SEEGER, *Inradius and circumradius of various convex cones arising in applications*, Set-Valued and Variational Analysis, 18 (2010), pp. 317–339.
21. ———, *On properties of different notions of centers for convex cones*, Set-Valued and Variational Analysis, 18 (2010), pp. 205–231.
22. M. HERCEG, M. KVASNICA, C. JONES, AND M. MORARI, *Multi-parametric toolbox 3.0*, in 2013 European Control Conference (ECC), 2013, pp. 502–510.
23. G. JARRY-BOLDUC, *Numerical Analysis for Derivative-Free Optimization*, PhD thesis, University of British Columbia, 2023.

24. G. JARRY-BOLDUC, P. NADEAU, AND S. SINGH, *Uniform simplex of an arbitrary orientation*, Optimization Letters, (2019), pp. 1–11.
25. C. JONES AND M. MCPARTLON, *Spherical discrepancy minimization and algorithmic lower bounds for covering the sphere*, 2020, pp. 874–891.
26. T. KOLDA, R. LEWIS, AND V. TORCZON, *Optimization by direct search: New perspectives on some classical and modern methods*, SIAM Review, 45 (2003), pp. 385–482.
27. M. KRISHAN LAL, *Nonconvex Projections Arising in Bilinear Mathematical Models*, PhD thesis, University of British Columbia, 2023.
28. R. LEWIS AND V. TORCZON, *Rank ordering and positive bases in pattern search algorithms*, Tech. Rep. TR96-71, Institute for Computer Applications in Science and Engineering, 1996.
29. D. MARCUS, *Minimal positive 2-spanning sets of vectors*, Proceedings of the American Mathematical Society, 82 (1981), pp. 165–172.
30. ———, *Gale diagrams of convex polytopes and positive spanning sets of vectors*, Discrete applied mathematics, 9 (1984), pp. 47–67.
31. G. NÆVDAL, *Positive bases with maximal cosine measure*, Optimization Letters, (2018), pp. 1–8.
32. J. REAY, *Unique minimal representations with positive bases*, The American Mathematical Monthly, 73 (1966), pp. 253–261.
33. R. REGIS, *On the properties of positive spanning sets and positive bases*, Optimization and Engineering, 17 (2016), pp. 229–262.
34. ———, *On the properties of the cosine measure and the uniform angle subspace*, Computational Optimization and Applications, 78 (2021), pp. 915–952.
35. R. ROCKAFELLAR AND R. WETS, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1998.
36. Z. ROMANOWICZ, *Geometric structure of positive bases in linear spaces*, Applicationes Mathematicae, 19 (1987), pp. 557–567.
37. G. SHEPHARD, *Diagrams for positive bases*, Journal of the London Mathematical Society, 2 (1971), pp. 165–175.