The cosine measure relative to a subspace

Charles Audet · Warren Hare · Gabriel Jarry-Bolduc

January 17, 2024

Abstract The *cosine measure* was introduced in 2003 to quantify the richness of a finite positive spanning sets of directions in the context of derivative-free directional methods. A positive spanning set is a set of vectors whose nonnegative linear combinations span the whole space. The present work extends the definition of cosine measure. In particular, the paper studies cosine measures relative to a subspace, and proposes a deterministic algorithm to compute it. The paper also studies the situation in which the set of vectors is infinite. The extended definition of the cosine measure might be useful for subspace decomposition methods.

Keywords Positive spanning set · positive basis · cosine measure · gradient approximation · subspace decomposition

1 Introduction

Given a set of vectors D in $\mathbb{R}^n$, the concepts of *spanning*, *linear independence*, and *basis*, are considered foundational to linear algebra. Closely related, but less well studied, are the ideas of *positive spanning*, *positive linear independence*, and *positive basis* [7]. All of these notions can be defined considering a set's properties relative to $\mathbb{R}^n$ or relative to a linear subspace [20].

In addition to the intrinsic mathematical interest in these concepts, positive bases have been shown to be fundamental to the convergence analysis in *direct-search methods* in *derivative-free optimization*. The first occurrence dates from 1996 [18] and more recent presentations are found

Charles Audet
GERAD and Département de mathématiques et génie industriel, Polytechnique Montréal, Montréal, Québec, Canada. Audet's research is partially funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, Discovery Grant #2020-04448. ORCID 0000-0002-3043-5393
E-mail: charles.audet@gerad.ca

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, Discover Grant #2018-03865. ORCID 0000-0002-4240-3903
E-mail: warren.hare@ubc.ca

Gabriel Jarry-Bolduc
Department of Mathematics and Statistics, Saint Francis Xavier University, Antigonish, Nova Scotia, Canada.
ORCID 0000-0002-1827-8508
E-mail: gabjarry@alumni.ubc.ca

in the textbooks [3] and [6]. A key tool in this analysis is the *cosine measure*, introduced in [15]. Algorithms to compute the cosine measure are provided in [10,21]. Results on the maximal value of the cosine measure for positive bases of $\mathbb{R}^n$ are developed in [12,19]. The notion of cosine measure is also briefly investigated for *positive k-spanning set* of $\mathbb{R}^n$ in [11]. Until now, this cosine measure was only deeply studied for positive spanning sets of $\mathbb{R}^n$. In this paper, we explore the *cosine measure relative to a subspace*, which effectively provides a measure of how well a subset of the positive spanning set considered explores that subspace. The results herein will be of high value in understanding the *subspace decomposition methods* which are recently gaining traction in [2,5,9,14,17,26].

The main goals of this paper are to introduce the notion of cosine measure relative to a subspace, to demonstrate its value in understanding positive spanning properties of a set, and to provide a deterministic algorithm to compute it. As a secondary goal, we provide several novel results examining the case where the set of vectors is infinite.

The remainder of this paper is organized as follows. The notation is presented in Section 2 and background results that are necessary to understand this paper are provided. This includes *positive spanning*, *positive linear independence*, *positive basis*, *cosine measure*, *cosine vector set*, and *active set*. These final three definitions are extended to include "relative to a subspace". In Section 3, properties of a positive spanning set of a linear subspace are investigated. It is shown that many known results regarding $\mathbb{R}^n$ are easily adapted to working in a subspace or working with an infinite set. In Section 4, the notion of cosine measure relative to a subspace is explored. Several results are provided that link the positive span, cosine measure, and cosine measure relative to a subspace. The case where the cosine measure relative to a subspace is equal to 0 is examined. The section concludes with new results showing how the cosine measure relative to a subspace can be used to provide a general error bound on the true gradient of a smooth function. In Section 5, a deterministic algorithm to compute the cosine measure relative to the span of the set is provided and proven to return the correct results. In Section 6, the main results of this paper are summarized.

2 Notation and Preliminaries

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 the dimension of the vector is clear, we may omit the subscript. The i^{th} coordinate vector in $\mathbb{R}^n$ is denoted e_i.

We denote by $B_n(x^0;\Delta)$ the *open ball centered about $x^0 \in \mathbb{R}^n$ with radius* $0 < \Delta < \infty$ and by $\overline{B}_n(x^0;\Delta)$ the *closed ball centered about x^0 with radius Δ*. That is

$$B_n(x^0;\Delta) = \left\{x \in \mathbb{R}^n : \|x - x^0\| < \Delta\right\} \quad \text{and} \quad \overline{B}_n(x^0;\Delta) = \left\{x \in \mathbb{R}^n : \|x - x^0\| \leq \Delta\right\}.$$

Given a set of vectors $D \subseteq \mathbb{R}^n$ (possibly infinite), the cardinality of D is denoted by $|D|$. The radius of the smallest ball centered at the origin containing the set D is denoted by Δ_D and given by

$$\Delta_D = \sup_{d \in D} \|d\|. \tag{1}$$

When useful, in this paper, a set of vectors is represented as a matrix where each column represents a vector in the set. That is

$$D = \{d_1, d_2, \ldots, d_m\} \text{ is interchangable with } D = \begin{bmatrix} d_1 \ d_2 \ \ldots \ d_m \end{bmatrix}.$$

The *dimension* of a linear subspace $L \subseteq \mathbb{R}^n$ is denoted by $\dim(L)$. A *trivial* subspace $L \subset \mathbb{R}^n$ is a set of the form $L = \{\mathbf{0}_n\}$ or $L = \emptyset$. Similarly, a set D in $\mathbb{R}^n$ is said to be trivial if $D = \{\mathbf{0}_n\}$ or $D = \emptyset$. In this paper, the linear subspace considered are assumed to be nontrivial.

We now focus on the key definitions studied in this paper. In the remaining of this paper, the word linear may be omitted when discussing the notion of span, subspace and basis. All these notions are defined for the linear case.

Definition 1 (span, positive span) Let $D \subseteq \mathbb{R}^n$.

i. The *span* of D is denoted by $\operatorname{span}(D)$ and defined by

$$\operatorname{span}(D) = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{k} \alpha_i d_i, k \in \mathbb{N}, d_i \in D, \alpha_i \in \mathbb{R} \right\}.$$

ii. The *positive span* of D is denoted by $\operatorname{pspan}(D)$ and defined by

$$\operatorname{pspan}(D) = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^{k} \lambda_i d_i, k \in \mathbb{N}, d_i \in D, \lambda_i \geq 0 \right\}.$$

Note that $\operatorname{span}(D)$ is always a linear subspace of $\mathbb{R}^n$. Moreover, L is a linear subspace of $\mathbb{R}^n$ if and only if $L = \operatorname{span}(L)$.

The *projection* of a vector $v \in \mathbb{R}^n$ onto a linear subspace L will be denoted by $\operatorname{Proj}_L v$. We will be particularly interested in the projection onto $\operatorname{span}(D)$. For ease of writing, we provide this with the special notation $\mathrm{P}_D = \operatorname{Proj}_{\operatorname{span}(D)}$. In Section 5, we shall make use of the well-known formula

$$\mathrm{P}_D\, v = DD^\dagger v,$$

where $D^\dagger$ denotes the *Moore–Penrose pseudoinverse of* D [23, Chapter 7].

Definition 2 (spanning, positive spanning) Let $D \subseteq \mathbb{R}^n$ and L be a subspace of $\mathbb{R}^n$.

i. The set D is said to be a *spanning set of* L, or *spans* L, if and only if $\operatorname{span}(D) = L$.
ii. The set D is said to be a *positive spanning set of* L, or *positively spans* L, if and only if $\operatorname{pspan}(D) = L$.

It is easy to prove that, if D (positively) spans L, then D must be a subset of L. Indeed, one always has

$$D \subseteq \operatorname{pspan}(D) \subseteq \operatorname{span}(D).$$

The term (positively) independent is used to mean that removing any vector from the set changes the (positive) spanning properties of the set.

Definition 3 (independence, positive independence) Let $D \subseteq \mathbb{R}^n$.

i. The set D is said to be *dependent* if and only if there exists a vector $d \in D$ such that $d \in \operatorname{span}(D \setminus \{d\})$. Conversely, D is said to be *independent* if and only if $\operatorname{span}(D \setminus \{d\}) \neq \operatorname{span}(D)$ for any $d \in D$.
ii. The set D is said to be *positively dependent* if and only if there exists a vector $d \in D$ such that $d \in \operatorname{pspan}(D \setminus \{d\})$. Conversely, D is said to be *positively independent* if and only if $\operatorname{pspan}(D \setminus \{d\}) \neq \operatorname{pspan}(D)$ for any $d \in D$.

Finally, we introduce the notion of a (positive) basis of a subspace.

Definition 4 (basis, positive basis) Let $D \subseteq \mathbb{R}^n$ and L be a subspace of $\mathbb{R}^n$.

i. The set D is a *basis of* L if and only if D is independent with $\mathrm{span}(D) = L$.
ii. The set D is a *positive basis of* L if and only if D is positively independent with $\mathrm{pspan}(D) = L$.

It is well known that if D is a basis of L then $|D| = \dim(L)$. It is shown in [20, Cor. 2.4 & Thm 6.6] that the minimal cardinality of a positive basis of a subspace L is $\dim(L) + 1$ and the maximal cardinality is $2\dim(L)$. Positive bases of these sizes are called *minimal positive bases* and *maximal positive bases* (respectively). Positive bases with cardinality strictly between $\dim(L) + 1$ and $2\dim(L)$ are called *intermediate positive bases.*

We end this section by recalling the definition of cosine measure and defining the cosine measure relative to a subspace. The cosine measure is valuable tool to quantify how well a set covers the space $\mathbb{R}^n$. We will see that the cosine measure relative to a subspace extends the value of this tool to work with a subspace of $\mathbb{R}^n$. We begin with the definition of cosine measure (as given in [12,21]) and the corresponding cosine vector set (from [10]).

Definition 5 (cosine measure) Let $D \subseteq \mathbb{R}^n$ be a nonempty finite set of nonzero vectors. The *cosine measure* of D is defined by

$$\mathrm{cm}\,(D) \;=\; \min_{\substack{u \in \mathbb{R}^n \\ \|u\|=1}} \max_{d \in D} \frac{u^\top d}{\|d\|},$$

and the *cosine vector set of* D, denoted by $\mathrm{cV}(D)$, is defined by

$$\mathrm{cV}(D) \;=\; \operatorname*{argmin}_{\substack{u \in \mathbb{R}^n \\ \|u\|=1}} \max_{d \in D} \frac{u^\top d}{\|d\|}.$$

Note that the original definition of the cosine measure requires D to be finite. Another limitation of the cosine measure is that is it focuses on how well the set covers the entire space $\mathbb{R}^n$. As a result, if D is a positive spanning set of a proper subspace, then the cosine measure will always return $\mathrm{cm}\,(D) = 0$ (Corollary 23 herein). To address these limitations, we introduce the *cosine measure relative to a subspace*, which provides information on how well a set covers a linear subspace. Simultaneously, we define a corresponding cosine vector set and allow for both concepts to be well-defined for infinite sets.

However, since an empty set, or the zero vector, provides no coverage of $\mathbb{R}^n$, the cosine measure relative to a subspace is defined for nonempty sets of nonzero vectors. If a set is empty, the cosine measure (relative to a subspace) should be assumed to be undefined. If a set contains the zero vector, then one should remove the vector and work with the remaining set. If the remaining set is empty, then the result would again be undefined.

For the remainder of this paper we shall assume that the set D is a nonempty set of nonzero vectors.

Definition 6 (cosine measure relative to L) Let $L \subseteq \mathbb{R}^n$ be a nontrivial linear subspace and let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors. The *cosine measure of D relative to L* is denoted by $\mathrm{cm}_L(D)$ and defined by

$$\mathrm{cm}_L(D) \;=\; \min_{\substack{u \in L \\ \|u\|=1}} \sup_{d \in D} \frac{u^\top d}{\|d\|}, \tag{2}$$

and the *cosine vector set of D relative to L*, denoted by $\mathrm{cV}_L(D)$, is defined by

$$\mathrm{cV}_L(D) \;=\; \operatorname*{argmin}_{\substack{u \in L \\ \|u\|=1}} \sup_{d \in D} \frac{u^\top d}{\|d\|}.$$

If $L = \mathbb{R}^n$ and $|D|$ is finite, then the cosine measure relative to L returns the classical cosine measure: $\mathrm{cm}_{\mathbb{R}^n}(\cdot) = \mathrm{cm}(\cdot)$. Indeed, if D is finite, then the sup can be replaced by max in (2). If D is a trivial set or if L is a trivial subspace of $\mathbb{R}^n$, then the cosine measure and cosine vector set (relative to L) are undefined. When D is a nonempty set of nonzero vectors and L is a nontrivial subspace, the cosine measure relative to a subspace returns a value in $[-1, 1]$. Besides, the cosine vector set of D relative to L is well-defined and nonempty. Notice that when D is an infinite set, then the cosine vector set may be infinite.

Remark 7 Given a nonempty set of nonzero vectors D and a nontrivial linear subspace L in $\mathbb{R}^n$ such that $D \not\subseteq L$, one might consider projecting each normalized vector $d/\|d\| \in D$ onto L before computing the cosine measure of D relative to L. Denote by $\widetilde{D}$ the set obtained from these projections and by $\widetilde{d}$ a vector in the set $\widetilde{D}$. Note that the set $\widetilde{D}$ might contain the zero vector. If the computations are done using the projected set $\widetilde{D}$, then it is crucial to keep the zero vector in the set $\widetilde{D}$. The cosine measure of D relative to L can be computed as follows whenever we work from the projected set $\widetilde{D}$ rather than using the original set D:

$$\begin{aligned} \mathrm{cm}_L(D) &= \min_{\substack{u \in L \\ \|u\|=1}} \sup_{d \in D} \frac{u^\top d}{\|d\|} \\ &= \min_{\substack{u \in L \\ \|u\|=1}} \sup_{d \in D} u^\top \left(\mathrm{Proj}_L \frac{d}{\|d\|} + \left(\frac{d}{\|d\|} - \mathrm{Proj}_L \frac{d}{\|d\|} \right) \right) \\ &= \min_{\substack{u \in L \\ \|u\|=1}} \sup_{\widetilde{d} \in \widetilde{D}} u^\top \widetilde{d}, \end{aligned} \tag{3}$$

as $u^\top \left(\frac{d}{\|d\|} - \mathrm{Proj}_L \frac{d}{\|d\|} \right) = 0$ for any $u \in L$.

The following example shows the importance of keeping the zero vector in the projected set $\widetilde{D}$ if the computations are done from the projected set $\widetilde{D}$.

Example 8 Let $D = \{e_1, e_2\} \subseteq \mathbb{R}^2$ and $L = \{x \in \mathbb{R}^2 : x_2 = 0\}$. Then

$$\mathrm{cm}_L(D) = 0.$$

The projection of e_1 onto L is equal to e_1 and the projection of e_2 onto L is equal to $\mathbf{0}$. We obtain the projected set $\widetilde{D} = \{e_1, \mathbf{0}\}$. Computing the cosine measure of D relative L as defined in Equation (3), we obtain $\mathrm{cm}_L(D) = 0$. However, if one removes the zero vector from $\widetilde{D}$, the final result would be

$$\min_{\substack{u \in L \\ \|u\|=1}} \sup_{\widetilde{d} \in \widetilde{D} \setminus \{\mathbf{0}\}} u^\top \widetilde{d} = -1 \neq \mathrm{cm}_L(D).$$

We shall often be interested in the special case when $L = \mathrm{span}(D)$. For ease of writing, when the linear subspace considered is the span of a set, we only write the set name and omit the word span in the subscript. More precisely,

$$\mathrm{cm}_D(\cdot) = \mathrm{cm}_{\mathrm{span}(D)}(\cdot) \quad \text{and} \quad \mathrm{cV}_D(\cdot) = \mathrm{cV}_{\mathrm{span}(D)}(\cdot).$$

Finally, some analysis will require examining the active set of the cosine measure relative to L.

Definition 9 (active set) Let $L \subseteq \mathbb{R}^n$ be a nontrivial linear subspace and $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors.

The *active set for D at a cosine vector $u \in \mathrm{cV}_L(D)$ relative to L* is denoted by $\mathcal{A}_L(D,u)$ and defined by

$$\mathcal{A}_L(D,u) \;=\; \left\{ d \in D : \frac{d^\top u}{\|d\|} = \mathrm{cm}_L(D) \right\}.$$

The *active set for D relative to L* is denoted by $\overline{\mathcal{A}}_L(D)$, and defined by

$$\overline{\mathcal{A}}_L(D) \;=\; \bigcup_{u \in \mathrm{cV}_L(D)} \mathcal{A}_L(D,u).$$

As above, we are mostly interested in the special case of $L = \mathrm{span}(D)$. For this case we use the special notation

$$\mathcal{A}_D(\cdot,\cdot) = \mathcal{A}_{\mathrm{span}(D)}(\cdot,\cdot).$$

3 Properties of a positive spanning set of a linear subspace

In this section, properties of positive spanning sets of a linear subspace are investigated. The results are developed to consider both the cases where $|D|$ is finite and $|D|$ is infinite. We begin by showing that a set D contained in a linear subspace L is a positive spanning set of L if and only if $\mathrm{pspan}(D) = \mathrm{span}(D) = L$.

Lemma 10 *Let $L \subseteq \mathbb{R}^n$ be a linear subspace. Then D positively spans L if and only if* $\mathrm{pspan}(D) = \mathrm{span}(D) = L$.

Proof ($\Rightarrow$) Suppose that D positively spans L; i.e., $\mathrm{pspan}(D) = L$. This implies $D \subseteq L$ and therefore $\mathrm{pspan}(D) \subseteq \mathrm{span}(D) \subseteq L = \mathrm{pspan}(D)$.
($\Leftarrow$) Conversely, if $\mathrm{pspan}(D) = \mathrm{span}(D) = L$, then D positively spans L by definition. □

We next show that if $|D|$ is infinity and the positive span is a linear subspace, then the positive span can be created through a finite subset of D. We use the following lemma.

Lemma 11 *[4, Theorem 2.11] Let $D \subseteq \mathbb{R}^n$ be a nonempty (possibly infinite) set of nonzero vectors. Then D contains a basis of* $\mathrm{span}(D)$.

Corollary 12 *Let $D \subseteq \mathbb{R}^n$ be a nonempty (possibly infinite) set of nonzero vectors and L be a linear subspace of $\mathbb{R}^n$. If $\mathrm{pspan}(D) = L$, then there exists a finite subset of vectors $C \subseteq D$ such that $\mathrm{pspan}(C) = L$.*

Proof If D is a finite set, then take $C = D$.

Suppose D is an infinite set and $\dim(L) \geq 1$. By Lemma 10, $\mathrm{pspan}(D) = L$ implies $L = \mathrm{span}(D)$. By Lemma 11, the set D contains a basis of $\mathrm{span}(D)$. Denote this basis by $B = \{b_1, b_2, \ldots, b_m\}$ where $m = \dim(\mathrm{span}(D))$. Define $v = -\sum_{i=1}^m b_i$. Since $\mathrm{pspan}(D) = \mathrm{span}(D)$, the vector v can be written as

$$v = \alpha_1 d_1 + d_2 + \cdots + \alpha_k d_k$$

where $d_i \in D$ and k is an integer greater or equal than 1. Since the $\mathrm{pspan}(B \cup \{v\}) = \mathrm{span}(D)$ [20, Theorem 5.1], we have

$$\mathrm{pspan}(B \cup \{d_1, d_2 \ldots, d_k\}) = \mathrm{pspan}(B \cup \{v\}) = \mathrm{span}(D) = L.$$

The set $B \cup \{d_1, d_2 \ldots, d_k\}$ is finite, so the proof is complete. □

Note that the above result requires D to be a positive spanning set. Indeed, consider the *floating ring set* defined by $D = \{x \in \mathbb{R}^3 : (x_1)^2 + (x_2)^2 = 1, x_3 = 1\}$. It can be shown that the set D is positively independent [13]. As such, for any finite subset $C \subseteq D$, we have $\operatorname{pspan}(C) \neq \operatorname{pspan}(D)$.

This inspires the following result about positive independence for an infinite set of vectors.

Corollary 13 *Let $D \subseteq \mathbb{R}^n$ be an infinite set of vectors. If D is positively independent, then* $\operatorname{pspan}(D) \neq \operatorname{span}(D)$.

Proof Let E be a finite subset of D with $\operatorname{pspan}(E) = \operatorname{span}(D)$. If D is positively independent, then $\operatorname{pspan}(D) = \operatorname{pspan}(E)$ creates a contradiction. □

Notice that Corollary 13 is specific to infinite sets. Indeed, a finite set of positively independent vectors D such that $\operatorname{pspan}(D) = \operatorname{span}(D)$ is a positive basis.

We now extend [20, Theorem 2.5] to the cases of infinite sets.

Theorem 14 *Let $D \subseteq \mathbb{R}^n$ be a nonempty (possibly infinite) set of nonzero vectors. Then the following are equivalent.*

(i) The set D is a positive spanning set of $\operatorname{span}(D)$.
(ii) For any $d \in D$, $-d$ is in $\operatorname{pspan}(D \setminus \{d\})$.

Moreover, these imply that there exists a finite subset $C \subseteq D$ such that $\operatorname{pspan}(C) = \operatorname{pspan}(D)$.

Suppose C is a finite subset of D such that $\operatorname{pspan}(C) = \operatorname{pspan}(D)$ (if D is finite, then $C = D$). Let $s = |C| \geq 1$. Then the following are equivalent.

(iii) The set D is a positive spanning set of $\operatorname{span}(D)$.
(iv) The set C is a positive spanning set of $\operatorname{span}(D)$.
(v) There exists $\alpha \in \mathbb{R}^s$, such that $\alpha > 0$ and $C\alpha = \mathbf{0}_n$.
(vi) There exists $\beta \in \mathbb{R}^s$, such that $\beta \geq 1$ and $C\beta = \mathbf{0}_n$.
(vii) There exists $\gamma \in \mathbb{R}^s$ such that $\gamma \geq 0$ and $C\gamma = -C\mathbf{1}_s$.

where items (v), (vi), and (vii), interpret C as a matrix in $\mathbb{R}^{n\times s}$.

Proof Parts (i) and (ii) are equivalent by [20, Theorem 2.5]. Suppose that D is a positive spanning set of $\operatorname{span}(D)$. By Corollary 12, part (i) implies that there exists a finite subset $C \subseteq D$ such that $\operatorname{pspan}(C) = \operatorname{span}(D)$. The proof of equivalence of parts (iii) to (vii) is identical to the proof for finite sets provided in [20, Theorem 2.5]. □

Our next example shows the importance of $\operatorname{pspan}(C) = \operatorname{pspan}(D)$ in the second half of Theorem 14.

Example 15 Consider $D_\infty = \{d : \|d\| = 1, d_1 \geq 0\} \subseteq \mathbb{R}^2$. The set $C = \{e_2, -e_2\}$ is contained in D_∞, but $\operatorname{pspan}(C) \neq \operatorname{pspan}(D_\infty)$. Notice that C satisfies parts (v), (vi), and (vii) of Theorem 14, but C is not a positive spanning set of $\operatorname{span}(D_\infty)$.

In practice, Theorem 14(vii) is useful to decide if a given set is a positive spanning set. Theorem 14(vii) can also be used to show how to extend a set that is not a positive spanning set into a positive spanning set by adding only one vector.

Proposition 16 *Let $D \subseteq \mathbb{R}^n$ be a nonempty (possibly infinite) set of nonzero vectors. Suppose $\operatorname{pspan}(D) \neq \operatorname{span}(D)$. Let $m = \dim(\operatorname{span}(D))$ and $B = \{b_1, \cdots, b_m\} \subseteq D$ be a basis of $\operatorname{span}(D)$. Define the vector $w = -\sum_{j=1}^m b_j$. Then*

(i) the set $D' = D \cup \{w\}$ is a positive spanning set of $\operatorname{span}(D)$, and

(ii) the set $D'' = D \cup -B$ *is a positive spanning set of* $\operatorname{span}(D)$.

Proof (i) Let $C = B \cup \{w\}$. By [3, Thm 6.4], $\operatorname{pspan}(C) = \operatorname{span}(B) = \operatorname{span}(D)$. We also have that the sum of the vectors of C is the null vector. Therefore, Theorem 14(v) ensures that $\operatorname{pspan}(D') = \operatorname{span}(D)$.
(ii) The result follows similarly using $C = B \cup (-B)$. □

One of the most useful properties of a positive spanning set of $\mathbb{R}^n$ is that for any nonzero vector $v \in \mathbb{R}^n$, there exists a vector d in the positive spanning set for which

$$v^\top d > 0.$$

It follows that given $f \in \mathcal{C}^1$, a positive spanning set of $\mathbb{R}^n$ contains a descent direction of f at any point $x^0 \in \mathbb{R}^n$ where $\nabla f(x^0) \neq \mathbf{0}_n$. This result can be generalized to the linear subspaces and infinite sets. To do this, we apply the following lemma.

Lemma 17 *Let* $D \subseteq \mathbb{R}^n$ *be a nonempty (possibly infinite) set of nonzero vectors. If* $\operatorname{pspan}(D) \neq \operatorname{span}(D)$, *then all the vectors in* D *are contained in a closed half-space of* $\operatorname{span}(D)$.

Proof Follows immediately from [22, Corollary 11.7.3]. □

Proposition 18 *Let* $D \subseteq \mathbb{R}^n$ *be a nonempty (possibly infinite) set of nonzero vectors. Then* $\operatorname{pspan}(D) = \operatorname{span}(D)$ *if and only if for any nonzero vector* $v \in \operatorname{span}(D)$ *there exists a vector* $d \in D$ *such that*

$$v^\top d > 0. \tag{4}$$

Proof ($\Rightarrow$) Suppose $\operatorname{pspan}(D) = \operatorname{span}(D)$. Let v be a nonzero vector in $\operatorname{span}(D)$. As $v \in \operatorname{pspan}(D)$, v can be written as

$$v = \alpha_1 d_1 + \alpha_2 d_2 + \cdots + \alpha_m d_m,$$

where $\alpha_j \geq 0$ for all $j \in \{1, 2, \ldots, m\}$. It follows that

$$0 < v^\top v = \alpha_1 d_1^\top v + \alpha_2 d_2^\top v + \cdots + \alpha_m d_m^\top v.$$

Hence, we must have $d_j^\top v > 0$ for at least one $j \in \{1, 2, \ldots, m\}$.
($\Leftarrow$) Conversely, suppose that given any $v \in \operatorname{span}(D) \setminus \{\mathbf{0}_n\}$ there exists a $d \in D$ such that $v^\top d > 0$. If $\operatorname{pspan}(D) \neq \operatorname{span}(D)$, then by Lemma 17 there exists $h \in \operatorname{span}(D) \setminus \{\mathbf{0}_n\}$ such that all vectors in D are contained in $\{w \in \operatorname{span}(D) : w^\top h \leq 0\}$. This leads to a contradiction by taking $v = h$. □

In Proposition 18, if we let the nonzero vector v be in $\mathbb{R}^n$ rather than $\operatorname{span}(D)$, then the result does not necessarily hold. The forward direction is true if we replace the strict inequality in (4) by an inequality. Proposition 19 provides a proof of this claim.

Proposition 19 *Let* $D \subseteq \mathbb{R}^n$ *be a nonempty (possibly infinite) set of nonzero vectors. If* $\operatorname{pspan}(D) = \operatorname{span}(D)$, *then for any vector* $v \in \mathbb{R}^n$, *there exists a vector* $d \in D$ *such that*

$$v^\top d \geq 0.$$

Proof Let $v \in \mathbb{R}^n$ and note that $v^\top d = (\mathrm{P}_D\, v)^\top d$ for all $d \in D$. If $\mathrm{P}_D\, v = \mathbf{0}_n$, then $v^\top d = 0$ for all $d \in D$. If $\mathrm{P}_D\, v \neq \mathbf{0}_n$, then by Proposition 18 there exists $d \in D$ such that $(\mathrm{P}_D\, v)^\top d > 0$. □

The converse of Proposition 19 is not necessarily true. For example, consider the set $D = \{e_1, -e_1, e_2\} \subseteq \mathbb{R}^2$. For this set, for any vector $v \in \mathbb{R}^2$, there exists $d \in D$ such that $v^\top d \geq 0$, however $\operatorname{pspan}(D) \neq \operatorname{span}(D)$.

In the next section, the notion of cosine measure relative to a linear subspace is investigated.

4 Cosine measure relative to a subspace

Definition 6 extends the definition of the cosine measure to include the idea of a cosine measure relative to a subspace. In this section we explore the basic properties of this new definition. As we are most interested in the cosine measure relative to $\mathrm{span}(D)$, many results focus on $\mathrm{cm}_D(D)$.

Since the definition of the cosine measure assumes D is finite, results involving the cosine measure include the assumptions that D in finite.

We begin with the obvious relationship between the cosine measure and the cosine measure relative to a subspace.

Proposition 20 *Let $D \subseteq \mathbb{R}^n$ be a nonempty finite set of nonzero vectors, L be a nontrivial linear subspace of $\mathbb{R}^n$. Then*

$$\mathrm{cm}(D) \leq \mathrm{cm}_L(D)$$

with equality if and only if $\mathrm{cV}(D) \cap L \neq \emptyset$.

Proof Since L is a nontrivial subspace of $\mathbb{R}^n$, the definitions of the cosine measures ensure that

$$\mathrm{cm}(D) \;=\; \min_{\substack{u \in \mathbb{R}^n \\ \|u\|=1}} \max_{d \in D} \frac{u^\top d}{\|d\|} \;\leq\; \min_{\substack{u \in L \\ \|u\|=1}} \max_{d \in D} \frac{u^\top d}{\|d\|} \;=\; \mathrm{cm}_L(D).$$

Now, suppose $\mathrm{cm}(D) = \mathrm{cm}_L(D)$. Let $u^* \in \mathrm{cV}_L(D)$. Then u^* must be in $\mathrm{cV}(D)$, so $\mathrm{cV} \cap L \neq \emptyset$. Conversely, suppose $\mathrm{cV}(D) \cap L \neq \emptyset$. Let $u^\# \in \mathrm{cV}(D) \cap L$. Then

$$\mathrm{cm}(D) \;=\; \max_{d \in D} \frac{(u^\#)^\top d}{\|d\|} \geq \min_{\substack{u \in L \\ \|u\|=1}} \max_{d \in D} \frac{u^\top d}{\|d\|} \;=\; \mathrm{cm}_L(D).$$

Since $\mathrm{cm}(D) \leq \mathrm{cm}_L(D)$, we must have $\mathrm{cm}(D) = \mathrm{cm}_L(D)$. ◻

4.1 Relating $\mathrm{pspan}(D)$ and $\mathrm{cm}_D(D)$

We now turn our attention to how the cosine measure relative to a subspace provides knowledge about positive spanning properties. We begin with an example computing the cosine measure relative to a subspace. We will return to this example after each result to illustrate what the information the cosine measure relative to a subspace provides. In the following example, it is relatively easy to find the exact value of the cosine measure relative to a subspace since the subspaces considered are either one dimensional or two dimensionals. A deterministic algorithm will be provided in Section 5.

Example 21 Consider the sets of directions $D_1 = \{e_1, -e_1, e_2, -e_2\} \subseteq \mathbb{R}^3$ and $D_2 = \{e_1, -e_1, e_2\} \subseteq \mathbb{R}^3$ and the subspaces $L = \{x \in \mathbb{R}^3 : 4x_1 = 3x_2, x_3 = 0\}$ and $M = \{x \in \mathbb{R}^3 : x_1 = x_3 = 0\}$. First, note that

$$\begin{aligned} &\mathrm{cm}\,(D_1) = 0, \qquad \mathrm{cV}(D_1) = \{e_3, -e_3\} \\ \text{and} \qquad &\mathrm{cm}\,(D_2) = 0, \qquad \mathrm{cV}(D_2) = \{v = (0, v_2, v_3) \,:\, \|v\| = 1,\, v_2 \leq 0\}. \end{aligned}$$

The shaded regions in Figure 1 represent the positive span of D_1 on the left and of D_2 on the right. Both figures also show the subspaces L and M.

There are exactly two unit vectors in L:

$$u^1 = \begin{bmatrix} 0.6 \\ 0.8 \\ 0 \end{bmatrix} \quad \text{and} \quad u^2 = \begin{bmatrix} -0.6 \\ -0.8 \\ 0 \end{bmatrix}.$$

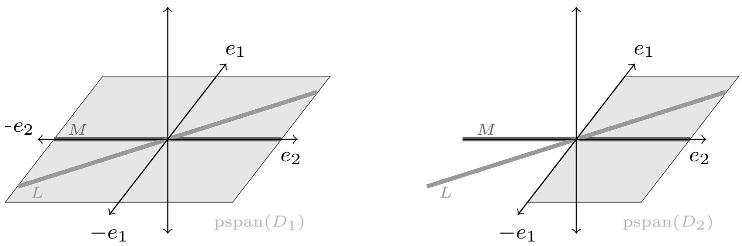

Fig. 1 Illustrations of the positive spans of D_1 and D_2

The cosine measures and cosine vectors relative to the subspace L are

$$\begin{array}{rrcll} & \mathrm{cm}_L(D_1) &=& \min\left\{\max\limits_{d\in D_1} u^\top d \,:\, u\in\{u^1,u^2\}\right\} = 0.8, & \mathrm{cV}_L(D_1) = \{u^1,u^2\} \\ \text{and} & \mathrm{cm}_L(D_2) &=& \min\left\{\max\limits_{d\in D_2} u^\top d \,:\, u\in\{u^1,u^2\}\right\} = 0.6, & \mathrm{cV}_M(D_2) = \{u^2\}. \end{array}$$

In the subspace M, there are exactly two unit vectors: e_2 and $-e_2$. We obtain

$$\begin{array}{rrcll} & \mathrm{cm}_M(D_1) &=& \min\left\{\max\limits_{d\in D_1} u^\top d \,:\, u\in\{e_2,-e_2\}\right\} = 1, & \mathrm{cV}_L(D_1) = \{e_2,-e_2\} \\ \text{and} & \mathrm{cm}_M(D_2) &=& \min\left\{\max\limits_{d\in D_2} u^\top d \,:\, u\in\{e_2,-e_2\}\right\} = 0, & \mathrm{cV}_M(D_2) = \{-e_2\}. \end{array}$$

Considering the subspaces $\mathrm{span}(D_1)$ and $\mathrm{span}(D_2)$, we find

$$\begin{array}{rll} & \mathrm{cm}_{D_1}(D_1) = 1/\sqrt{2}, & \mathrm{cV}_{D_1}(D_1) = \{(v_1,v_2,0) \,:\, |v_1| = |v_2| = 1/\sqrt{2}\} \\ \text{and} & \mathrm{cm}_{D_2}(D_2) = 0, & \mathrm{cV}_{D_2}(D_2) = \{-e_2\}. \end{array}$$

Given a nontrivial finite set $D \subseteq \mathbb{R}^n$, it is known that $\mathrm{cm}(D) > 0$ if and only if D positively spans $\mathbb{R}^n$ [21, Theorem 4.2]. Note that since $\mathrm{pspan}(D) \subseteq \mathrm{span}(D)$, the previous result can be expressed as follows: given a nontrivial finite set $D \subseteq \mathbb{R}^n$, we have $\mathrm{cm}(D) > 0$ if and only if $\mathrm{pspan}(D) = \mathrm{span}(D)$. Thus, neither D_1 nor D_2 in Example 21 is a positive spanning set of $\mathbb{R}^3$. The cosine measure provides no further information.

The next proposition shows the ability of the cosine measure relative to a subspace to detect if the subspace is contained in the positive span of D.

Proposition 22 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors and $L \subseteq \mathrm{span}(D)$ be a nontrivial linear subspace. Then $L \subseteq \mathrm{pspan}(D)$ if and only if $\mathrm{cm}_L(D) > 0$. In particular, $\mathrm{pspan}(D) = \mathrm{span}(D)$ if and only if $\mathrm{cm}_D(D) > 0$.*

Proof ($\Rightarrow$) Suppose $L \subseteq \mathrm{pspan}(D)$. Let $u^* \in \mathrm{cV}_L(D)$. Since $u^* \in L$, we have that u^* can be expressed as $u^* = \alpha_1 d_1 + \alpha_2 d_2 + \cdots + \alpha_m d_m$, where $\alpha_j \geq 0$ for all $j \in \{1, 2, \ldots, m\}$. Similar to Proposition 18, $1 = (u^*)^\top u^* = \alpha_1 d_1^\top u^* + \alpha_2 d_2^\top u^* + \cdots + \alpha_m d_m^\top u^*$ implies $d_j^\top v > 0$ for at least one $j \in \{1, 2, \ldots, m\}$. Thus $\mathrm{cm}_L(D) > 0$.

($\Leftarrow$) Conversely, suppose $\mathrm{cm}_L(D) > 0$. This implies that given any $v \in L$ with $\|v\| = 1$ there exists $d \in D$ such that $v^\top d > 0$. If L is not a subset of $\mathrm{pspan}(D)$, then $\mathrm{pspan}(D) \neq \mathrm{span}(D)$, so applying Lemma 10 in the same manner as Proposition 18 leads to the same contradiction. Thus $L \subseteq \mathrm{pspan}(D)$.

The final statement comes from setting $L = \mathrm{span}(D)$. □

In Example 21, $\mathrm{cm}_L(D_1) > 0, \mathrm{cm}_L(D_2) > 0$, $\mathrm{cm}_M(D_1) > 0$; thus $L \subseteq \mathrm{pspan}(D_1)$, $L \subseteq \mathrm{pspan}(D_2)$, and $M \subseteq \mathrm{pspan}(D_1)$. Also, $\mathrm{cm}_{D_1}(D_1) > 0$, so $\mathrm{pspan}(D_1) = \mathrm{span}(D_1)$. However, $\mathrm{cm}_M(D_2) \leq 0$, so M is not a subspace of $\mathrm{pspan}(D_2)$. Also, $\mathrm{cm}_{D_2}(D_2) = 0$, so $\mathrm{pspan}(D_2) \neq \mathrm{span}(D_2)$.

Next we prove the claim in the introduction: "if D is a positive spanning set of a *proper* subspace, then the cosine measure will always return 0".

Corollary 23 *Let $D \subseteq \mathbb{R}^n$ be a nonempty finite set of nonzero vectors. Suppose* $\mathrm{span}(D) \neq \mathbb{R}^n$. *If* $\mathrm{pspan}(D) = \mathrm{span}(D)$, *then* $\mathrm{cm}(D) = 0$.

Proof Select any $v \in \mathbb{R}^n$ with $\|v\| = 1$. Let $v_D = \mathrm{P}_D\, v$ and $v_{D^\perp} = v - v_D$. Since $\mathrm{pspan}(D) = \mathrm{span}(D)$, Proposition 22 implies $\mathrm{cm}_D(D) > 0$, which further implies

$$\max_{d \in D} \frac{(v_D)^\top d}{\|d\|} \;\geq\; 0,$$

with equality only if $v_D = \mathbf{0}_n$. Since $v_{D^\perp}$ is in the orthogonal subspace to D, $(v_{D^\perp})^\top d = 0$ for all $d \in D$. Thus,

$$\max_{d \in D} \frac{(v)^\top d}{\|d\|} \;=\; \max_{d \in D} \frac{(v_D)^\top d}{\|d\|} + \max_{d \in D} \frac{(v_{D^\perp})^\top d}{\|d\|} \;\geq\; 0,$$

with equality only if $v_D = \mathbf{0}_n$. Selecting v to be in the orthogonal subspace to $\mathrm{span}(D)$ now demonstrates $\mathrm{cm}(D) = 0$. □

In Example 21, $\mathrm{span}(D_1) \neq \mathbb{R}^3$ and $\mathrm{cm}(D_1) = 0$, so Corollary 23 allows for the possibility that $\mathrm{pspan}(D_1) = \mathrm{span}(D_1)$. However, notice that this is not sufficient to ensure that result, as D_2 in Example 21 also has $\mathrm{cm}(D_2) = 0$, but $\mathrm{pspan}(D_2) \neq \mathrm{span}(D_2)$. Hence, Corollary 23 cannot be made into an 'if and only if' statement.

The next theorem further investigates the relation between the value of the cosine measure and the value of the cosine measure relative to the subspace $\mathrm{span}(D)$.

Theorem 24 *Let D be a nonempty finite set of nonzero vectors in $\mathbb{R}^n$. If* $\mathrm{cm}(D) \neq \mathrm{cm}_D(D)$, *then* $\mathrm{pspan}(D) = \mathrm{span}(D) \neq \mathbb{R}^n$.

Proof If $\mathrm{cm}(D) \neq \mathrm{cm}_D(D)$, then Proposition 20 using $L = \mathbb{R}^n$ implies that $\mathrm{span}(D) \neq \mathbb{R}^n$.

For eventual contradiction, suppose $\mathrm{cm}(D) \neq \mathrm{cm}_D(D)$ and $\mathrm{pspan}(D) \neq \mathrm{span}(D)$. From Proposition 20, $\mathrm{cm}(D) \neq \mathrm{cm}_D(D)$ implies that $\mathrm{cm}(D) < \mathrm{cm}_D(D)$. From Proposition 22, $\mathrm{pspan}(D) \neq \mathrm{span}(D)$ implies that $\mathrm{cm}_D(D) \leq 0$. Hence, we have

$$\mathrm{cm}(D) < \mathrm{cm}_D(D) \leq 0. \tag{5}$$

Let $v \in \mathrm{cV}(D)$ and notice that Proposition 20 implies $v \notin \mathrm{span}(D)$. Let $v_D = \mathrm{P}_D\, v$ and notice that $v^\top d = v_D^\top d$ for all $d \in D$. This implies that

$$\mathrm{cm}(D) = \max_{d \in D} v^\top \frac{d}{\|d\|} = \max_{d \in D} v_D^\top \frac{d}{\|d\|},$$

which further implies $v_D \neq \mathbf{0}_n$. Define $\ell = \|v_D\|$. Considering $v_D \in \mathrm{span}(D)$ and $\|\frac{1}{\ell} v_D\| = 1$, we obtain

$$\mathrm{cm}_D(D) \leq \max_{d \in D} \left(\frac{1}{\ell} v_D\right)^\top \frac{d}{\|d\|} = \frac{1}{\ell}\,\mathrm{cm}(D) \;<\; \mathrm{cm}(D),$$

where the last inequality comes from $0 < \ell < 1$ and $\mathrm{cm}(D) < 0$. This contradicts inequality (5), so the initial supposition cannot hold. □

Returning to Example 21, notice that $\mathrm{cm}(D_1) = \mathrm{cm}_{D_1}(D_1)$, and indeed we have $\mathrm{span}(D_1) = \mathrm{pspan}(D_1) \neq \mathbb{R}^3$. In Example 21, we find the $\mathrm{cm}(D_2) = \mathrm{cm}_{D_2}(D_2) = 0$, so Theorem 24 does not apply. In the next section, we explore what $\mathrm{cm}_D(D) = 0$ tells us about D.

4.2 Consequence of $\mathrm{cm}_D(D) = 0$

In this section, we focus exclusively on properties of $\mathrm{cm}_D(D)$. Since the results do not involve $\mathrm{cm}(D)$, we no longer require the assumption that D is a finite set. However, for the sake of simplicity, we continue to assume that D is a nonempty set of nonzero vectors.

To explore the consequences of $\mathrm{cm}_D(D) = 0$ we first begin by examining the other extreme case: $\mathrm{cm}_D(D) = 1$.

Proposition 25 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors. Define $U = \{u \in \mathrm{span}(D) : \|u\| = 1\}$ and $\widehat{D} = \mathrm{cl}\{d/\|d\| : d \in D\}$, where* cl *denotes the closure of the set. Then the following are equivalent.*

(i) $U = \widehat{D}$,
(ii) $\mathrm{cm}_D(D) = 1$,
(iii) $\mathrm{cV}_D(D) = U$,
(iv) $\overline{\mathcal{A}}_D(D) = U$.

Proof Clearly (i) implies (ii), (iii), and (iv). Therefore we only need to show (ii) implies (i). Suppose $\mathrm{cm}_D(D) = 1$. By construction, $\widehat{D} \subseteq U$. Let $u^* \in U$. By definition of the cosine measure relative to span(D),

$$1 = \mathrm{cm}_D(D) = \min_{u \in \mathrm{span}(D) \|u\|=1} \sup_{d \in D} \frac{u^\top d}{\|d\|} = \min_{\substack{u \in \mathrm{span}(D) \\ \|u\|=1}} \sup_{\hat{d} \in \widehat{D}} u^\top \hat{d} \leq \sup_{\hat{d} \in \widehat{D}} (u^*)^\top \hat{d} \leq \|u^*\| = 1,$$

where the last inequality is true by Cauchy–Schwarz. Moreover, the last inequality is an equality if and only if $\hat{d} = u^*$. Since equality holds across the above, this implies $u^* \in \widehat{D}$. Thus $U \subseteq \widehat{D}$, which provides (i) and the proof is complete. □

An important consequence of Proposition 25 is that if $|D| = 2$, then either $\mathrm{cm}_D(D) = 1$ or $\mathrm{cm}_D(D) < 0$.

Corollary 26 *Let $D \subseteq \mathbb{R}^n$ be a set of nonzero vectors with exactly 2 vectors. Then either $\mathrm{cm}_D(D) = 1$ or $\mathrm{cm}_D(D) < 0$.*

Proof Let $D = \{d_1, d_2\}$. Set $u = -\left(\frac{d_1}{\|d_1\|} + \frac{d_2}{\|d_2\|}\right)$. If $u = \mathbf{0}_n$, then Proposition 25 creates $\widehat{D} = \{\frac{d_1}{\|d_1\|}, -\frac{d_1}{\|d_1\|}\} = U$, so $\mathrm{cm}_D(D) = 1$. If $u \neq \mathbf{0}_n$, then set $u^* = u/\|u\|$ and notice

$$\mathrm{cm}_D(D) \leq \max\left\{\frac{(u^*)^\top d_1}{\|d_1\|}, \frac{(u^*)^\top d_2}{\|d_2\|}\right\} = \frac{1}{\|u\|}\left(-1 - \frac{d_1^\top d_2}{\|d_1\|\|d_2\|}\right) < 0.$$

□

We now begin our examination of the consequences of $\mathrm{cm}_D(D) = 0$. Our goal is to show that $\mathrm{cm}_D(D) = 0$ if and only if D contains a finite nontrivial subset that is a positive spanning. The proof is reductionist in nature and uses the following lemma, which shows that if $\mathrm{cm}_D(D) = 0$, then there exists a subset $V \subseteq D$ that is strictly smaller than D and has $\mathrm{cm}_V(V) \geq 0$.

Lemma 27 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors. If $\mathrm{cm}_D(D) = 0$, then there exists a finite subset $V \subseteq D$ such that $1 < |V| < |D|$ and $\mathrm{cm}_V(V) \geq 0$.*

Proof Let $u^* \in \mathrm{cV}_D(D)$. Define $H = \{v \in \mathrm{span}(D) : (u^*)^\top v = 0\}$ and $V = D \cap H$. Since $\mathrm{cm}_D(D) = 0$, by definition there exists at least one vector in V.

If $|D| = |V|$, then $V = D$, which implies $(u^*)^\top d = 0$ for all $d \in D$. This yields $(u^*)^\top v = 0$ for all $v \in \mathrm{span}(D)$, which implies $u^* = \mathbf{0}_n$ contradicting $\|u^*\| = 1$. Thus, $|V| < |D|$.

Finally, for eventual contradiction, suppose $\mathrm{cm}_V(V) < 0$. This implies that there exists $v^* \in \mathrm{span}(V)$ such that $(v^*)^\top d < 0$ for all $d \in V$. For $\epsilon > 0$, define

$$u' = u^* + \epsilon v^* \in \mathrm{span}(D).$$

Given any $d \in V$, we have

$$(u')^\top d = (u^* + \epsilon v^*)^\top d = (u^*)^\top d + (\epsilon v^*)^\top d < 0,$$

as $(u^*)^\top d = 0$ and $(\epsilon v^*)^\top d < 0$. Given $d \in D \setminus V$, we have

$$(u')^\top d = (u^*)^\top d + (\epsilon v^*)^\top d \leq \max_{d \in D \setminus V} \left(\|d\| \frac{(u^*)^\top d}{\|d\|} \right) + (\epsilon v^*)^\top d.$$

Since $u^* \in \mathrm{cV}_D(D)$, we have

$$\max_{d \in D \setminus V} \left(\|d\| \frac{(u^*)^\top d}{\|d\|} \right) < 0.$$

Thus, for ϵ sufficiently small $(u')^\top d < 0$ for all $d \in D$. The vector u' contradicts $\mathrm{cm}_D(D) = 0$, and therefore we must have $\mathrm{cm}_V(V) \geq 0$.

Finally, $|V| > 1$, as if $|V| = 1$, then $\mathrm{cm}_V(V) = -1$. If V is not finite, then apply Theorem 14 to reduce to a finite subset. □

We now present the main result for this subsection.

Theorem 28 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors. Suppose* $\mathrm{pspan}(D) \neq \mathrm{span}(D)$. *Then* $\mathrm{cm}_D(D) = 0$ *if and only if D contains a nonempty finite proper subset V such that* $\mathrm{pspan}(V) = \mathrm{span}(V)$.

Proof First, note that $\mathrm{pspan}(D) \neq \mathrm{span}(D)$ implies $\mathrm{cm}_D(D) \leq 0$.

($\Leftarrow$) Suppose there exists a nonempty proper subset $V \subset D$ such that $\mathrm{pspan}(V) = \mathrm{span}(V)$. Let $u^* \in \mathrm{cV}_D(D)$ and define $u^*_V = P_V u^*$ and $u^*_{V^\perp} = u^* - u^*_V$. Since $u^*_{V^\perp}$ is in the orthogonal subspace to V, we have

$$\sup_{d \in V} \frac{(u^*)^\top v}{\|v\|} = \sup_{d \in V} \frac{(u^*_V)^\top v}{\|v\|} \leq \sup_{d \in D} \frac{(u^*_V)^\top d}{\|d\|} = \sup_{d \in D} \frac{(u^*)^\top d}{\|d\|} - \sup_{d \in D} \frac{(u^*_{V^\perp})^\top d}{\|d\|} \leq \mathrm{cm}_D(D) \leq 0. \quad (6)$$

If $u^*_V \neq \mathbf{0}_n$, then Proposition 18 would imply the existence of $v \in V$ with $(u^*_V)^\top v > 0$, therefore equation (6) implies that $u^*_V = \mathbf{0}_n$. Substituting $u^*_V = \mathbf{0}_n$ in equation (6) shows $\mathrm{cm}_D(D) = 0$.

($\Rightarrow$) Suppose $\mathrm{cm}_D(D) = 0$. Without loss of generality, we assume D is finite. (Indeed, if D is not finite, then apply Theorem 14 to drop to a finite set.) Let $m = |D|$. Note that $m \geq 3$, as $|D| = 1$ implies $\mathrm{cm}_D(D) = -1$ and $|D| = 2$ implies $\mathrm{cm}_D(D) = 1$ or $\mathrm{cm}_D(D) < 0$.

By Lemma 27, there exists a nonempty proper subset $D_1 \subsetneq D$ such that $1 < |D_1| < m$ and $\mathrm{cm}_{D_1}(D_1) \geq 0$. If $\mathrm{cm}_{D_1}(D_1) > 0$, then $V = D_1$ is our desired set. If $\mathrm{cm}_{D_1}(D_1) = 0$, then $|D_1| > 2$ and therefore we can repeat the procedure as necessary to generate a nonempty proper subset $D_k \subset D_{k-1}$ such that $2 \leq |D_k| < |D_{k-1}| \leq m - k$ and $\mathrm{cm}_{D_k}(D_k) \geq 0$. This process must terminate before $k = m - 1$ or a contradiction is created. When the procedure is terminated we have $\mathrm{cm}_{D_k}(D_k) > 0$, so $V = D_k$ is our desired set. □

The previous theorem can be adapted to the cosine measure of D or reformulated to discuss dimensions of subspaces.

Corollary 29 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors.*

(i) Suppose $\operatorname{pspan}(D) \neq \mathbb{R}^n$. Then $\operatorname{cm}(D) = 0$ if and only if D contains a nonempty proper subset V such that $\operatorname{pspan}(V) = \operatorname{span}(V)$.
(ii) Suppose $\operatorname{pspan}(D) \neq \operatorname{span}(D)$. Then $\operatorname{cm}_D(D) = 0$ if and only if D contains a positive basis of a linear subspace $L \subset \operatorname{span}(D)$ with $1 \leq \dim(L) < \dim(\operatorname{span}(D))$.
(iii) Suppose $\operatorname{pspan}(D) \neq \operatorname{span}(D)$. Then $\operatorname{cm}_D(D) = 0$ if and only if D contains a minimal positive basis of a linear subspace $L \subset \operatorname{span}(D)$ with $1 \leq \dim(L) < \dim(\operatorname{span}(D))$.

Proof Item (i) is immediate from Theorem 28. Item (ii) results from rephrasing Theorem 28 in terms of positive bases. Item (iii) follows from [24, Theorem 1], where it is shown that a positive basis of a linear subspace can be partitioned to minimal positive bases [24, Theorem 1]. □

4.3 Bounding the norm of the gradient in directional direct-search methods

Directional methods such as the Generalized Pattern Search (GPS) [25] and the Mesh Adaptive Direct Search (MADS)[1] algorithms are foundational methods in derivative-free optimization [3,6]. Their convergence rely on applying poll steps, evaluating $f(x^k + d)$, $d \in D$ where D is a positive basis for $\mathbb{R}^n$, to seek improvement in the objective function. An important result is that in the case of a failed poll step (i.e., $f(x^k) \leq f(x^k + d)$ for all $d \in D$), the cosine measure of D combined with the radius of D provides an error bound on the gradient of f (assuming ∇f is Lipschitz continuous). We provide a formal statement below from [6], but recommend seeing [8, 16] for alternate presentations.

Given the nature of directional direct-search algorithm, in this subsection we continue to assume that $\mathbf{0}_n \notin D$.

Theorem 30 *[6, Theorem 2.8] Let $D \subseteq \mathbb{R}^n$ be a nonempty finite subset of nonzero vectors with radius Δ_D. Let $f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $x^0 \in \operatorname{dom} f$. Suppose D positively spans $\mathbb{R}^n$ and $f(x^0) \leq f(x^0 + d)$ for all $d \in D$. If ∇f is Lipschitz continuous with constant $L_{\nabla f} \geq 0$ in an open set containing the ball $B_n(x^0; \Delta_D)$, then*

$$\|\nabla f(x^0)\| \leq \frac{1}{2} L_{\nabla f} \operatorname{cm}(D)^{-1} \Delta_D.$$

We next present two extensions of this result. Theorem 31(i) directly extends Theorem 30 to allow for subspaces and infinite sets. (Setting $L = \mathbb{R}^n$ and D finite in Theorem 31(i) reproduces Theorem 30.) Theorem 31(ii) presents a new result, demonstrating a stronger error bound in the situation where the set D contains symmetry.

Theorem 31 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors. Suppose the radius of D is finite ($0 < \Delta_D < \infty$). Let $f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $x^0 \in \operatorname{dom} f$. In addition, suppose $\operatorname{pspan}(D) = \operatorname{span}(D)$ and $f(x^0) \leq f(x^0 + d)$ for all $d \in D$.*

(i) If ∇f is Lipschitz continuous with constant $L_{\nabla f} \geq 0$ in an open set containing the ball $\overline{B}_n(x^0; \Delta_D)$, then

$$\| \operatorname{P}_D \nabla f(x^0)\| \leq \frac{1}{2} L_{\nabla f} \operatorname{cm}_D(D)^{-1} \Delta_D. \tag{7}$$

(ii) Suppose in addition that for each $d \in D$ there exists $\alpha_d > 0$ such that $-\alpha_d d \in D$. If $\nabla^2 f$ is Lipschitz continuous with constant $L_{\nabla^2 f} \geq 0$ in an open set containing the ball $\overline{B}_n(x^0; \Delta_D)$, then

$$\| \mathrm{P}_D \nabla f(x^0)\| \leq \frac{1}{3}\alpha_{\max} L_{\nabla^2 f} \,\mathrm{cm}_D(D)^{-1} \Delta_D^2, \tag{8}$$

where $\alpha_{\max} = \sup_{d \in D}\{\alpha_d\}$.

Proof (i) Let $v = -\mathrm{P}_D \nabla f(x^0)$. If $v \neq \mathbf{0}_n$, then by definition, we have

$$\mathrm{cm}_D(D) \leq \sup_{d \in D} \frac{v^\top d}{\|v\|\|d\|}$$

and therefore there exists a $d \in D$ such that

$$\mathrm{cm}_D(D)\|v\|\|d\| \leq v^\top d.$$

If $v = \mathbf{0}_n$, the above holds trivially. Therefore, there exists a vector $d \in D$ such that

$$\mathrm{cm}_D(D)\| \mathrm{P}_D \nabla f(x^0)\|\|d\| \leq -\left(\mathrm{P}_D \nabla f(x^0)\right)^\top d. \tag{9}$$

Applying Taylor's Theorem and the assumption that $f(x^0) - f(x^0 + d) \leq 0$ for all $d \in D$, we have

$$-\left(\mathrm{P}_D \nabla f(x^0)\right)^\top d = -\nabla f(x^0)^\top d = f(x^0) - f(x^0 + d) + R_1(x^0; d) \leq R_1(x^0; d) \tag{10}$$

where R_1 is the first-order remainder term. Combining equations (9) and (10), then noting that $|R_1(x^0; d)| \leq \frac{1}{2} L_{\nabla f}\|d\|^2$, yields

$$\mathrm{cm}_D(D)\| \mathrm{P}_D \nabla f(x^0)\|\|d\| \leq R_1(x^0; d) \leq \frac{1}{2} L_{\nabla f}\|d\|^2.$$

Since $\mathrm{span}(D) = \mathrm{pspan}(D)$, we know that $\mathrm{cm}_D(D) > 0$. Equation (7) follows using the fact $\|d\| \leq \Delta_D$ for all $d \in D$.

(ii) Suppose that for each $d \in D$ there exists $\alpha_d > 0$ such that $-\alpha_d d \in D$ and let $\alpha_{\max} = \sup_{d \in D}\{\alpha_d\}$. If $\alpha_{\max} = \infty$, then the result holds trivially, so we assume $\alpha_{\max} < \infty$.

Taylor's Theorem and the assumption that $f(x^0) - f(x^0 + d) \leq 0$ for all $d \in D$, yields

$$\begin{aligned} -\nabla f(x^0)^\top d &= f(x^0) - f(x^0 + d) + \frac{1}{2} d^\top \nabla^2 f(x^0) d + R_2(x^0; d), \\ -\nabla f(x^0)^\top d &\leq \frac{1}{2} d^\top \nabla^2 f(x) d + R_2(x^0; d) \end{aligned} \tag{11}$$

and similarly,

$$-\nabla f(x^0)^\top(-\alpha_d d) \leq \frac{\alpha_d^2}{2} d^\top \nabla^2 f(x^0) d + R_2(x^0; -\alpha_d d), \tag{12}$$

where R_2 is the second-order remainder term. Multiplying (11) by α^2 and subtracting (12), we get

$$-(\alpha_d^2 + \alpha_d)\nabla f(x^0)^\top d = -(\alpha_d^2 + \alpha_d)\left(\mathrm{P}_D \nabla f(x^0)\right)^\top d \leq \alpha_d^2 R_2(x^0; d) - R_2(x^0; -\alpha_d d) \tag{13}$$

Therefore, combining equations (9) and (13), then noting that $|R_2(x^0;d)| \leq \frac{1}{6} L_{\nabla^2 f} \|d\|^3$, we find

$$\begin{aligned}(\alpha_d^2 + \alpha_d) \operatorname{cm}_D(D) \| \operatorname{P}_D \nabla f(x^0) \| \|d\| &\leq \alpha_d^2 R_2(x^0;d) - R_2(x^0; -\alpha_d d) \\ &\leq \alpha_d^2 \frac{L_{\nabla^2 f}}{6} \|d\|^3 + \frac{L_{\nabla^2 f}}{6} \| - \alpha_d d\|^3 \\ &\leq (\alpha_d^2 + \alpha_d^3) \frac{L_{\nabla^2 f}}{3} \|d\|^3\end{aligned}$$

Using the bounds $\|d\| \leq \Delta_D$ and $\alpha_d \leq \alpha_{\max}$, we obtain (8). □

Theorem 31 could be used in directional direct-search methods in several different manners. One obvious example would be to use it to generate a stopping condition, particularly in the case where $\operatorname{span}(D) = \operatorname{pspan}(D)$. In the case of reduced subspace methods where $\operatorname{span}(D) \neq \operatorname{pspan}(D)$, Theorem 31 could be used to create flags indicating when it is time switch to a different subspace.

While convergence of the directional direct-search methods requires the use of positive spanning sets, it is easy to conceive of an implementation that does not enforce positive spanning sets at every iteration. The following corollary demonstrates how Theorem 31 might be used to help determine next steps in the case of a failed poll step where D is not a positive spanning set.

Corollary 32 *Let $D \subseteq \mathbb{R}^n$ be a nonempty set of nonzero vectors. Suppose the radius of D is finite ($0 < \Delta_D < \infty$). Let $f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $x^0 \in \operatorname{dom} f$. In addition, suppose $\operatorname{pspan}(D) \neq \operatorname{span}(D)$ and $f(x^0) \leq f(x^0 + d)$ for all $d \in D$.*

Let $B = \{d_1, d_2 \ldots, d_m\} \subseteq D$ be a basis of $\operatorname{span}(D)$. Define

$$w = -\Delta_D \frac{\sum_{j=1}^m d_j}{\| \sum_{j=1}^m d_j \|}, \quad D' = D \cup \{w\}, \quad \textit{and} \quad D'' = D \cup -D.$$

(i) If ∇f is Lipschitz continuous with constant $L_{\nabla f} \geq 0$ in an open set containing the ball $\overline{B}_n(x^0; \Delta_D)$, then at least one of the following holds:

$$f(x^0 + w) < f(x^0) \quad \text{or} \quad \| \operatorname{P}_D \nabla f(x^0) \| \leq \frac{1}{2} L_{\nabla f} \operatorname{cm}_D(D')^{-1} \Delta_D. \tag{14}$$

(ii) If $\nabla^2 f$ is Lipschitz continuous with constant $L_{\nabla^2 f} \geq 0$ in an open set containing the ball $\overline{B}_n(x^0; \Delta_D)$, then then at least one of the following holds

$$f(x^0 - d) < f(x^0) \textit{ for some } d \in D \quad \text{or} \quad \| \operatorname{P}_D \nabla f(x^0) \| \leq \frac{1}{3} L_{\nabla^2 f} \operatorname{cm}_D(D'')^{-1} \Delta_D^2. \tag{15}$$

Proof (i) Suppose $f(x^0) \leq f(x^0 + w)$. By Proposition 16, D' is a positive spanning set. The result follows from Theorem 31(i), noting that the scaling of w makes $\Delta_{D'} = \Delta_D$.
(ii) Suppose $f(x^0) \leq f(x^0 - d)$ for all $d \in D$. By Proposition 16, D'' is a positive spanning set. The result follows from Theorem 31, noting that $\Delta_{D''} = \Delta_D$ and $\alpha_d = 1$ for all $d \in D$. □

5 Computing the cosine measure relative to a subspace

Section 4 established the value of the cosine measure relative to a subspace. In this section, we investigate how to compute $\operatorname{cm}_D(D)$ for a nonempty finite set of nonzero vectors. In Example 21, we were able to compute the cosine measure relative to a subspace since the subspace considered were either 1-dimensional or 2-dimensional. In this section, we provide a general deterministic algorithm to compute the cosine measure relative to a subspace. Note that Algorithm 1 assumes

D is finite and $\mathbf{0}_n \notin D$. Reasons for both are clear. If D is infinite, then we must confront the challenge of how to express the set in a manner suitable for algorithmic use. If we allowed $\mathbf{0}_n \in D$, then the first step of the algorithm would simply become remove $\mathbf{0}_n$ from D.

Algorithm 1 is a modification of the algorithm in [10] to allow for subspaces and for the case where D is not a positive spanning set. To allow for the scenario where D is not a positive spanning set, Algorithm 1 begins by checking if $\operatorname{pspan}(D) = \operatorname{span}(D)$. Recall that this can be done by solving a linear program using Theorem 14(vii). Within the algorithm `ps` is used as a flag to store whether D is a positive spanning set (`ps` = 1) or not (`ps` = -1). This flag is used in line (2.1) to control whether γ_B is positive or negative. The notation $\mathbf{G}(B)$ represents the *Gram matrix* of B. That is $\mathbf{G}(B) = B^\top B$.

Algorithm 1: The cosine measure of a finite set D relative to $\operatorname{span}(D)$

Given a nonempty finite set D of q nonzero vectors in $\mathbb{R}^n$,

0. Normalize: set $D \leftarrow \{d/\|d\| : d \in D\}$.

1. Determine if $\operatorname{pspan}(D) = \operatorname{span}(D)$ **and define**

(1.1) $\quad \mathtt{ps} = \begin{cases} 1, \text{ if } \operatorname{pspan}(D) = \operatorname{span}(D) \\ -1, \text{ if } \operatorname{pspan}(D) \neq \operatorname{span}(D). \end{cases}$

2. Let $m = \dim(\operatorname{span}(D)) \geq 1$. **For all bases** B **of** $\operatorname{span}(D)$ **contained in** D, **compute**

(2.1) $\quad \gamma_B = (\mathtt{ps})\dfrac{1}{\sqrt{\mathbf{1}_m{}^\top \mathbf{G}(B)^{-1}\mathbf{1}_m}}$ (pos. if $\operatorname{pspan}(D) = \operatorname{span}(D)$, neg. otherwise),

(2.2) $\quad u_B = \gamma_B \left(B^\top\right)^\dagger \mathbf{1}_m$ (the unit vector associated to γ_B),

(2.3) $\quad p_B = \left[[p_B]_1 \cdots [p_B]_q\right] = (u_B)^\top D$ (the dot product vector),

(2.4) $\quad \mathring{p}_B = \max\limits_{1\leq j\leq q} [p_B]_j$ (the maximum value in p_B).

3. Return

(3.1) $\quad \operatorname{cm}_D(D) = \begin{cases} \min\limits_{B\subseteq D} \mathring{p}_B, & \text{if } \mathtt{ps} = 1 \quad (\operatorname{pspan}(D) = \operatorname{span}(D)), \\ \min\left\{\min\limits_{B\subseteq D} \mathring{p}_B, 0\right\}, & \text{if } \mathtt{ps} = -1 \quad (\operatorname{pspan}(D) \neq \operatorname{span}(D)). \end{cases}$

(3.2) $\quad \operatorname{cV}_D(D) = \begin{cases} \{u_B : \mathring{p}_B = \operatorname{cm}_D(D)\}, & \text{if } \operatorname{cm}_D(D) \neq 0, \\ \{u \in \operatorname{span}(D) : D^\top u \leq \mathbf{0}_q, \|u\| = 1\}, & \text{if } \operatorname{cm}_D(D) = 0. \end{cases}$

To prove that the algorithm returns the correct cosine measure and the cosine vector set, we begin by introducing several results that can be viewed as a generalization of the results in [10, 19]. The following lemma is an adaptation of Lemma 1 in [19].

Lemma 33 *Let* $B = \left[d_1 \cdots d_m\right]$ *be a basis of* $S_B = \operatorname{span}(B)$ *in* $\mathbb{R}^n$ *written in matrix form where* $1 \leq m \leq n$ *and where each* d_j *is a unit vector. Then there exist a unit vectors* $u_B \in S_B$ *such that*

$$(u_B)^\top d_1 = \cdots = (u_B)^\top d_m = \gamma_B,$$

and

$$((-1)u_B)^\top d_1 = \cdots = ((-1)u_B)^\top d_m = (-1)\gamma_B.$$

where

$$\gamma_B = \frac{1}{\sqrt{\mathbf{1}_m^\top \mathbf{G}(B)^{-1}\mathbf{1}_m}} > 0. \tag{16}$$

Moreover,

$$u_B = \gamma_B (B^\top)^\dagger \mathbf{1}_m. \tag{17}$$

The proof is essentially identical to the proof of [19, Lemma 1] by replacing $\mathbb{R}^n$ with S_B.

Lemma 34 *Let $B = \{d_1, \cdots, d_m\}$ be a basis of $S_B = \operatorname{span}(B)$ in $\mathbb{R}^n$ and where each d_j is a unit vector. Suppose u is a unit vector in S_B such that $u^\top d_1 = \cdots = u^\top d_m = \alpha > 0$. Then $\alpha = \gamma_B$, where γ_B is defined as in (16).*

The previous lemma can be proved using a similar process than the proof of Lemma 13 in [10]. Next, we recall a lemma that will be useful to prove the key theorem of this section.

Lemma 35 *[10, Lemma 16] Let $\epsilon \neq 0$ and let u and v be unit vectors in $\mathbb{R}^n$. Then*

(i) $\|u + \epsilon v\| = 1$ *if and only if* $\epsilon = -2u^\top v$, *and*
(ii) $\|u + \epsilon v\| < 1$ *implies* $\|u - \epsilon v\| > 1$.
(iii) Assume $\|u \pm \epsilon v\| \neq 0$. *Then*

$$\frac{u \pm \epsilon v}{\|u \pm \epsilon v\|} = u \iff v = \pm u.$$

Theorem 36 *Let D be a nonempty set of nonzero vectors in $\mathbb{R}^n$ with $\operatorname{cm}_D(D) \neq 0$. Let $u_B \in \operatorname{cV}_D(D)$. Then*

$$\operatorname{span}(\mathcal{A}_D(D, u_B)) = \operatorname{span}(D).$$

Proof Without loss of generality, assume that all vectors d in D are unit vectors. Suppose that $\operatorname{span}(\mathcal{A}_D(D, u_B)) \neq \operatorname{span}(D)$, i.e., the rank of $\mathcal{A}_D(D, u_B)$ is strictly less than $\dim(\operatorname{span}(D))$. This implies that the kernel of $\mathcal{A}_D(D, u_B)$ is nonempty. Let $v \in \operatorname{span}(D)$ be a unit vector in the kernel of $\mathcal{A}_D(D, u_B)$. This means that $d^\top v = 0$ for all d in $\mathcal{A}_D(D, u_B)$.

Notice that, if $d \in D \setminus \mathcal{A}_D(D, u_B)$, then

$$d^\top u_B < \operatorname{cm}_D(D).$$

Consider the vector $u_B + \epsilon v \in \operatorname{span}(D)$. Since $\operatorname{cm}_D(D) \neq 0$, it follows that $u_B \neq \pm v$ as $u_B^\top d \neq 0$ for all $d \in \mathcal{A}_D(D, u_B)$. Hence, using Lemma 35(iii), for sufficiently small $\epsilon > 0$ and not equal to $|-2u_B^\top v|$, we have

$$\frac{d^\top (u_B \pm \epsilon v)}{\|u_B \pm \epsilon v\|} < \operatorname{cm}_D(D)$$

for all $d \in D \setminus \mathcal{A}_D(D, u_B)$. Moreover, since $d^\top v = 0$, it follows that

$$\frac{d^\top (u_B \pm \epsilon v)}{\|u_B \pm \epsilon v\|} = \frac{d^\top u_B}{\|u_B \pm \epsilon v\|} \pm 0 = \frac{\operatorname{cm}_D(D)}{\|u_B \pm \epsilon v\|}$$

for all $d \in \mathcal{A}_D(D, u_B)$ and where $\operatorname{cm}_D(D) \neq 0$ by assumption. By Lemma 35(i), $\epsilon \neq -2u_B^\top v$ implies that $\|u_B + \epsilon v\| \neq 1$. By Lemma 35(ii), if $\|u_B + \epsilon v\| < 1$, then $\|u_B - \epsilon v\| > 1$. Select w in $\{u_B + \epsilon v, u_B - \epsilon v\}$ such that $\|w\| > 1$. Then

$$\frac{d^\top w}{\|w\|} < \operatorname{cm}_D(D)$$

for all $d \in D$. This contradicts the definition of cosine measure.

Therefore, $\operatorname{span}(D) \subseteq \operatorname{span}(\mathcal{A}_D(D, u_B)) \subseteq \operatorname{span}(D)$, and the result follows. □

Theorem 36 can be viewed as an extension of Proposition 17 in [10] for the following two reasons: the linear subspace $\mathrm{span}(D)$ is considered rather than the whole space $\mathbb{R}^n$; and the assumption that the set D considered is a positive spanning set is deleted and replaced by the more general condition that $\mathrm{cm}_D(D) \neq 0$.

The following corollary follows from Theorem 36 and the fact that a spanning set of a vector space contains a basis of the vector space [4, Theorem 2.11].

Corollary 37 *Let D be a nonempty set of nonzero vectors in $\mathbb{R}^n$ such that $\mathrm{cm}_D(D) \neq 0$. Let $u_B \in \mathrm{cV}(D)$. Then $\mathcal{A}_D(D, u_B)$ contains a basis of* $\mathrm{span}(D)$.

We are now ready to show that Algorithm 1 returns the desired values.

Theorem 38 *Let $D = \{d_1, \dots, d_q\}$ be a set of $q \geq 1$ nonzero vectors in $\mathbb{R}^n$. Then Algorithm 1 returns $\mathrm{cm}_D(D)$ and $\mathrm{cV}_D(D)$.*

Proof Without loss of generality, assume that all vectors d_j are unit vectors. Since $q \geq 1$, $\mathrm{cV}_D(D) \neq \emptyset$. Let $u_B \in \mathrm{cV}_D(D)$. Define

$$\mathtt{ps} = \begin{cases} 1, \text{ if } \mathrm{pspan}(D) = \mathrm{span}(D) \\ -1, \text{ if } \mathrm{pspan}(D) \neq \mathrm{span}(D). \end{cases}$$

Case (i) $\mathtt{ps} = \mathtt{1}$. Suppose $\mathtt{ps} = \mathtt{1}$. By Proposition 22, this implies $\mathrm{cm}_D(D) > 0$. Let $u_B \in \mathrm{cV}_D(D)$. By Corollary 37, $\mathcal{A}_D(D, u_B)$ contains a basis of $\mathrm{span}(D)$. Without loss of generality, let this basis (written in matrix form) be $B_* = \begin{bmatrix} d_1 \cdots d_m \end{bmatrix}$ where $m = \dim(\mathrm{span}(D)) \geq 1$. Hence,

$$\mathrm{cm}_D(D) = d_1^\top u_B = \cdots = d_m^\top u_B > 0.$$

By Lemma 34,

$$\mathrm{cm}_D(D) = \gamma_{B_*} = (1)\frac{1}{\sqrt{\mathbf{1}_m^\top \mathbf{G}(B_*)^{-1}\mathbf{1}_m}}$$

Note that $\mathring{p}_{B_*} = \max_{1\leq j\leq q} d_j^\top u_B = \gamma_{B_*}$ since $\gamma_{B_*} = \mathrm{cm}_D(D)$. Therefore, we have

$$\mathrm{cm}_D(D) = \min_{B\subseteq D} \mathring{p}_B = \mathring{p}_{B_*}.$$

Taking all the vectors u_B associated to $\mathring{p}_B$ such that $\mathrm{cm}_D(D) = \mathring{p}_B$ in Step (3.2) returns the complete set $\mathrm{cV}_D(D)$.

Case (ii) $\mathtt{ps} = -\mathtt{1}$. Suppose $\mathtt{ps} = -\mathtt{1}$. Then either $\mathrm{cm}_D(D) < 0$ or $\mathrm{cm}_D(D) = 0$. When $\mathrm{cm}_D(D) < 0$, using any $u_B \in \mathrm{cV}_D(D)$, Corollary 37 guarantees that the $\mathcal{A}_D(D, u_B)$ contains a basis of $\mathrm{span}(D)$. A similar process than the previous case shows that the cosine measure is equal to

$$\mathrm{cm}_D(D) = \min_{B\subseteq D} \mathring{p}_B.$$

If $\mathrm{cm}_D(D) = 0$, then Corollary 37 does not apply, so $\min_{B\subseteq D} \mathring{p}_B$ can be 0 or strictly positive. (If $\mathrm{cm}_D(D) = 0$, then $\mathrm{cm}_D(D) \leq \min_{B\subseteq D} \mathring{p}_B$ implies that this value cannot be negative.) Therefore,

$$\min\left\{\min_{B\subseteq D} \mathring{p}_B, 0\right\}$$

returns the exact cosine measure when $\mathtt{ps} = -\mathtt{1}$.

When $\mathrm{cm}_D(D) < 0$, then the first branch of Step (3.2) returns the complete vector set $\mathrm{cV}_D(D)$ by Theorem 36. When $\mathrm{cm}_D(D) = 0$, then finding all unit vectors u in $\mathrm{span}(D)$ such that $D^\top u \leq \mathbf{0}_q$ provides the complete vector set, which is the second branch in Step (3.2). □

6 Conclusion

This paper introduces the definitions of the cosine measure relative to a subspace, and the related cosine vector set relative to a subspace. These definitions not only generalize the cosine measure and cosine vector set to allow for working in subspaces, but also generalize these ideas to work for infinite sets.

Novel results demonstrate the value of these new definitions for working with sets that are not positive spanning. Proposition 16 shows that any nonempty set of vectors can be extended to positive spanning set of its span by adding at most one vector to the set. Section 4 provides several properties of the cosine measure relative to a subspace. Theorem 24 shows that if the cosine measure of a set D differs from the cosine measure relative to the span of D, then D is a positive spanning set of its span and $\operatorname{span}(D)$ must be a proper subspace of $\mathbb{R}^n$. Theorem 28 proves that the cosine measure relative to $\operatorname{span}(D)$ is equal to zero if and only if the set D contains a positive spanning set of the span of a proper nonempty subset of D. Theorem 31 uses the notion of cosine measure relative to a subspace to define two error bounds on the projected gradient of a smooth function. In the case where the set is not a positive spanning set of its span, Corollary 32 introduces results that could be valuable when the poll step of a derivative-free algorithm fails. Lastly, a deterministic algorithm is proposed to compute the cosine measure relative to its span. The algorithm is designed to accept a non-positive spanning set as an input. Combined, these results demonstrate that the cosine measure relative to a subspace is a valuable and practical tool to quantify the positive spanning properties of a set relative to a subspace. On a final note, an implementation of Algorithm 1 in MATLAB is available upon request.

References

1. C. Audet and J. Dennis, *Mesh adaptive direct search algorithms for constrained optimization*, SIAM Journal on Optimization, 17 (2006), pp. 188–217.
2. C. Audet, J. Dennis, Jr., and S. Le Digabel, *Parallel space decomposition of the mesh adaptive direct search algorithm*, SIAM Journal on Optimization, 19 (2008), pp. 1150–1170.
3. C. Audet and W. Hare, *Derivative-Free and Blackbox Optimization*, Springer Series in Operations Research and Financial Engineering, Springer, Cham, Switzerland, 2017.
4. W. Brown, *A Second Course in Linear Algebra*, Wiley, 1988.
5. C. Cartis and L. Roberts, *Scalable subspace methods for derivative-free nonlinear least-squares optimization*, Mathematical Programming, 199 (2023), pp. 461–524.
6. A. Conn, K. Scheinberg, and L. Vicente, *Introduction to Derivative-Free Optimization*, MOS-SIAM Series on Optimization, SIAM, Philadelphia, 2009.
7. C. Davis, *Theory of positive linear dependence*, American Journal of Mathematics, 76 (1954), pp. 733–746.
8. E. Dolan, R. Lewis, and V. Torczon, *On the local convergence of pattern search*, SIAM Journal on Optimization, 14 (2003), pp. 567–583.
9. S. Gratton, C. Royer, L. Vicente, and Z. Zhang, *Direct search based on probabilistic feasible descent for bound and linearly constrained problems*, Computational Optimization and Applications, 72 (2019), pp. 525–559.
10. W. Hare and G. Jarry-Bolduc, *A deterministic algorithm to compute the cosine measure of a finite positive spanning set*, Optimization Letters, 14 (2020), pp. 1305–1316.
11. 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.
12. W. Hare, G. Jarry-Bolduc, and C. Planiden, *Nicely structured positive bases with maximal cosine measure*, Optimization Letters, (2023), pp. 1–21.
13. W. Hare and H. Song, *On the cardinality of positively linearly independent sets*, Optimization Letters, 10 (2016), pp. 649–654.
14. M. Kimiaei, A. Neumaier, and P. Faramarzi, *New subspace method for unconstrained derivative-free optimization*, ACM Transactions on Mathematical Software, 49 (2023).
15. 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.

16. 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.
17. D. Kozak, S. Becker, A. Doostan, and L. Tenorio, *A stochastic subspace approach to gradient-free optimization in high dimensions*, Computational Optimization and Applications, 79 (2021), pp. 339–368.
18. R. Lewis and V. Torczon, *Rank ordering and positive bases in pattern search algorithms*, Tech. Rep. TR96-71, ICASE, NASA Langley Research Center, 1996.
19. G. Nævdal, *Positive bases with maximal cosine measure*, Optimization Letters, (2018), pp. 1–8.
20. R. Regis, *On the properties of positive spanning sets and positive bases*, Optimization and Engineering, 17 (2016), pp. 229–262.
21. ———, *On the properties of the cosine measure and the uniform angle subspace*, Computational Optimization and Applications, 78 (2021), pp. 915–952.
22. R. Rockafellar, *Convex analysis*, vol. 28.;28;, Princeton University Press, Princeton, New Jersey, 1970.
23. S. Roman, *Advanced Linear Algebra*, vol. 135, Springer, New York, 3rd ed., 2007.
24. Z. Romanowicz, *Geometric structure of positive bases in linear spaces*, Applicationes Mathematicae, 19 (1987), pp. 557–567.
25. V. Torczon, *On the convergence of pattern search algorithms*, SIAM Journal on Optimization, 7 (1997), pp. 1–25.
26. P. Xie and Y. Yuan, *A new two-dimensional model-based subspace method for large-scale unconstrained derivative-free optimization: 2d-mosub*, arXiv preprint arXiv:2309.14855, (2023).