Positive bases with maximal cosine measure

Geir Nævdal¹

Received: 2 February 2018 / Accepted: 28 September 2018 / Published online: 5 October 2018
© The Author(s) 2018

Abstract

Positive spanning sets and positive bases are important in the construction of derivative-free optimization algorithms. The convergence properties of the algorithms might be tied to the cosine measure of the positive basis that is used, and having higher cosine measure might in general be preferable. In this paper, the upper bound of the cosine measure for certain positive bases in $\mathbb{R}^n$ are found. In particular if the size of the positive basis is $n + 1$ (the minimal positive bases) the maximal value of the cosine measure is $1/n$. A straightforward corollary is that the maximal cosine measure for any positive spanning set of size $n + 1$ is $1/n$. If the size of a positive basis is $2n$ (the maximal positive bases) the maximal cosine measure is $1/\sqrt{n}$. In all the cases described, the positive bases achieving these upper bounds are characterized.

Keywords Positive basis · Positive spanning set · Cosine measure

1 Introduction

The seminal work on positive bases by Davis [6] established many of their properties. The last few decades positive bases have received renewed attention due to their use in derivative-free optimization. For example, Conn et al. [5] which gives an introduction to derivative-free optimization introduces positive spanning sets and positive bases and the cosine measure (if necessary, see definitions below) as part of the background before starting on developing any optimization routines. The survey paper [10] also introduces these topics early on. Positive bases are also introduced as a background before constructing a couple of direct search methods in the recent book [3]. The usefulness of positive bases within derivative-free optimization was first pointed out in [11].

The convergence properties of several derivative free algorithms depends on the cosine measure of a positive spanning set, and it is preferable to design the positive

✉ Geir Nævdal
gena@norce-research.no

¹ NORCE Norwegian Research Centre, Postboks 22, 5838 Bergen, Norway

spanning set having its cosine measure as large as possible. In a recent paper, the problem of determining positive spanning sets with maximal cosine measure is described as “widely open” [7]. The present paper shows that the maximal cosine measure for a positive spanning set for $\mathbb{R}^n$ of size $n + 1$ is $1/n$. The minimal size of a positive spanning set for $\mathbb{R}^n$ is $n + 1$, which means that such a positive spanning set is necessarily a positive basis, and it is called a *minimal positive basis*. (It is trivial to show that the minimal size of a positive basis of $\mathbb{R}^n$ is $n + 1$.) Already C. Davis [6] established that the maximal size of a positive basis for $\mathbb{R}^n$ is $2n$ and such a positive basis is called a *maximal positive basis*. To show that the maximal size is $2n$ is not trivial, but a short proof using ideas from optimization theory was given by Audet [2]. In [7], the cosine measure of a positive spanning set for $\mathbb{R}^n$ of size $2n$ is conjectured to always be less than or equal to $1/\sqrt{n}$. The present paper shows that this is indeed the case for the maximal positive bases (which is a subset of the positive spanning sets of size $2n$).

Both the minimal and maximal positive bases with maximal cosine measure turn out to be given as examples of positive bases in [5]. The minimal bases of unit vectors with cosine measure $1/n$ are the positive bases $\mathbf{v}_1, \dots, \mathbf{v}_{n+1} \in \mathbb{R}^n$ where $\mathbf{v}_i^T \mathbf{v}_j = -1/n$ if $i \neq j$. The maximal positive bases for $\mathbb{R}^n$ consisting of the standard basis vectors and their negatives have cosine measure $1/\sqrt{n}$.

In [1] the authors show the existence of a positive basis $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ with $\mathbf{v}_i^T \mathbf{v}_j = -1/n$ if $i \neq j$ and show how to construct such a positive basis motivated by an optimization problem occurring in molecular geometry. An application using positive bases and the cosine measure in improving the performance of a direct search algorithm is shown in [4].

To be more formal, let us provide the definitions of positive spanning sets, positive bases and their cosine measure. A *positive spanning set* is a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r \in \mathbb{R}^n$ with the property that any $\mathbf{v} \in \mathbb{R}^n$ is in its positive span, that is $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_r \mathbf{v}_r$ for some $\alpha_1, \dots, \alpha_r \geq 0$. The set is a positive basis if none of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ can be written as a positive sum of the remaining vectors (i.e. the vectors are positively independent). Basic properties of positive spanning sets and positive bases can be found in e.g., [5,6,12]. The size of the spanning set is equal to the number of vectors it contains, which is r in the definition given above. For $\mathbb{R}^n$, $n \geq 3$ it has been shown that there exists positively independent set of vectors of an arbitrary number of elements (shown independently in [8,12]).

For convenience, only positive spanning sets without the zero vector is considered. Then the cosine measure of a positive spanning set or of a positive basis D is defined by

$$\text{cm}(D) = \min_{0 \neq \mathbf{u} \in \mathbb{R}^n} \max_{\mathbf{d} \in D} \frac{\mathbf{u}^T \mathbf{d}}{\|\mathbf{u}\| \|\mathbf{d}\|},$$

see, e.g., [5].

The paper is organized as follows. Fundamental material on positive bases and positive semidefinite matrices are presented in Sect. 2. Section 3 gives the main results of the paper. A short outlook is provided in Sect. 4.

2 Some background about positive semidefinite matrices

The ideas used to provide the proofs of my results are essentially from two sources. Obviously, the properties of positive spanning sets and positive bases would be exploited. The other source of auxiliary results are from the theory of positive semidefinite matrices and is covered by [9, Chapter 7]. The results presented in this section should be well-known in their respective domains, but are included here for easy reference.

The results concerning positive spanning sets and positive bases are introduced as needed. As mentioned in the introduction, references for this topic are e.g., [5,6,12]. Here we will only point to one property, that is given in all the three references above, that will be used in the proof of Theorem 1. This result is one of the characterizations of positive spanning sets and states that for any positive spanning set $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^n$ it holds that for any $\mathbf{w} \in \mathbb{R}^n$ there exists an index i ($1 \leq i \leq r$) such that $\mathbf{w}^T \mathbf{v}_i > 0$. Obviously, the inequality can also be reversed since the result also must hold for $-\mathbf{w}$.

Recall that an $n \times n$ matrix $\mathbf{G}$ is *positive semidefinite* if it holds that $\mathbf{x}^T \mathbf{G} \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ ($\mathbf{x} \neq \mathbf{0}$). If the inequality is strict, the matrix is *positive definite*. In the proof of Theorem 1 we will utilize the fact that any principal submatrix of a positive semidefinite matrix is again positive semidefinite. In particular, this means that if all the diagonal entries of a positive semidefinite matrix are one, then the off-diagonal entries will be in the interval $[-1, 1]$. (If an index set $1 \leq i_1 < i_2 < \dots < i_k \leq n$, $\mathbf{B}$ is a principal submatrix of $\mathbf{A} \in \mathbb{R}^{n \times n}$ if $\mathbf{B} \in \mathbb{R}^{k \times k}$ is formed by selecting the columns and rows $i_1 < i_2 < \dots < i_k$ of $\mathbf{A}$. The inheritance property follows easily by restricting the vector $\mathbf{x}$ in the definition above to the same index set.)

Concerning positive semidefinite matrices the properties of the Gram matrix of a set of vectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ will be useful. The Gram matrix is defined as

$$\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_r) = \begin{pmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \dots & \mathbf{v}_1^T \mathbf{v}_r \\ \vdots & & \vdots \\ \mathbf{v}_r^T \mathbf{v}_1 & \dots & \mathbf{v}_r^T \mathbf{v}_r \end{pmatrix},$$

Later on, the form $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_r)$ is used to denote such a Gram matrix. The Gram matrices are necessarily positive semidefinite (they can be factorized using the vectors that define them). If $\mathbf{v}_1, \dots, \mathbf{v}_r$ are linearly independent, the matrix $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_r)$ is non-singular (i.e. positive definite). If the vectors are linearly dependent, the rank of $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_r)$ is equal to the dimension of the (sub)space the vectors span.

The proof of Theorem 1 utilizes the fact that if the $(n+1) \times (n+1)$ matrix, $\mathbf{G}$, written as a 2×2 block matrix,

$$\mathbf{G} = \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 1 \end{pmatrix}$$

where $\mathbf{A}$ is a singular $n \times n$ matrix and $\mathbf{b} \in \mathbb{R}^n$, is a positive semidefinite matrix, then $\mathbf{b}$ is orthogonal to the kernel of $\mathbf{A}$. This can be proven by noting that if $\mathbf{b}$ is not orthogonal to the kernel of $\mathbf{A}$, then we can select $\mathbf{c} \in \mathbb{R}^n$ from the kernel of $\mathbf{A}$ (i.e.,

$\mathbf{A}\mathbf{c} = \mathbf{0}$), such that $\mathbf{c}^T \mathbf{b} = 1$. Then it follows that

$$(\mathbf{c}^T \quad -1) \begin{pmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{b}^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ -1 \end{pmatrix} = -1,$$

contradicting the assumption that $\mathbf{G}$ is a positive semidefinite matrix.

The following technical lemma will be used in the proof of Theorem 2. (Note that the bound for α that can be obtained directly from $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)$ would be useful.)

Lemma 1 *Let $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ be linearly independent unit vectors. Let $\mathbf{e} \in \mathbb{R}^n$ be the vector having all its entries equal to one. Then there exists a unit vector $\mathbf{u} \in \mathbb{R}^n$ such that $\mathbf{u}^T \mathbf{v}_i = \alpha$ for $i = 1, \dots, n$, where $0 < \alpha = 1/\sqrt{\mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e}}$. Moreover, $\alpha \leq \sqrt{\mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) \mathbf{e}}/n$.*

Proof Utilizing the linear independence of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ there exists a vector $\tilde{\mathbf{u}}$ such that $\tilde{\mathbf{u}}^T \mathbf{v}_i = 1$ for $i = 1, \dots, n$. The vector $\tilde{\mathbf{u}}$ can be scaled to a unit vector $\mathbf{u}$ such that $\mathbf{u}^T \mathbf{v}_i = \alpha > 0$ for $i = 1, \dots, n$. To calculate α consider the extension of the Gram matrix $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)$ to $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{u})$ which must be positive semidefinite matrix of rank n (i.e. it is singular) having the form

$$\begin{pmatrix} \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) & \alpha \mathbf{e} \\ \alpha \mathbf{e}^T & 1 \end{pmatrix}. \quad (1)$$

By multiplying the matrix above with

$$\begin{pmatrix} \mathbf{I} & -\alpha \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}$$

from the right and its transpose from the left (here $\mathbf{I}$ and $\mathbf{0}$ denote the identity and zero matrices with appropriate dimension) it follows that the matrix given in (1) has the same rank as

$$\begin{pmatrix} \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) & \mathbf{0} \\ \mathbf{0} & 1 - \alpha^2 \mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e} \end{pmatrix}.$$

The matrix above is positive semidefinite with rank n if

$$1 - \alpha^2 \mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e} = 0,$$

which has the positive solution

$$\alpha = 1/\sqrt{\mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e}}.$$

The inequality $(\mathbf{e}^T \mathbf{e})^2 \leq (\mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e})(\mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) \mathbf{e})$ follows from the Cauchy–Schwarz inequality (see e.g., [9, Theorem 5.1.4]) using the vectors $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{1/2} \mathbf{e}$ and $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1/2} \mathbf{e}$. This provides the bound $\alpha \leq \sqrt{\mathbf{e}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) \mathbf{e}}/n$. $\square$

3 Results

Now it is time to present the main results of the paper.

Theorem 1 *Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n+1}$ be a (minimal) positive basis for $\mathbb{R}^n$. Then the following holds:*

1. *Its cosine measure is bounded above by $1/n$.*
2. *The positive bases of unit vectors $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ with the property that*

$$\mathbf{v}_i^T \mathbf{v}_j = -1/n$$

for $i \neq j$ are the only positive bases of unit vectors that has cosine measure $1/n$.

Proof We first prove (1). Without loss of generality we assume that the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ are unit vectors. Form the Gram matrix

$$\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_{n+1}) = \begin{pmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \dots & \mathbf{v}_1^T \mathbf{v}_{n+1} \\ \vdots & & \vdots \\ \mathbf{v}_{n+1}^T \mathbf{v}_1 & \dots & \mathbf{v}_{n+1}^T \mathbf{v}_{n+1} \end{pmatrix}.$$

The Gram matrix $\mathbf{G}$ is a positive semidefinite matrix with rank n with ones on the diagonal. Since $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ is a positive spanning set there exists $\alpha_1, \dots, \alpha_{n+1} \geq 1$ such that $\sum_{i=1}^{n+1} \alpha_i \mathbf{v}_i = 0$. This implies that the vector $(\alpha_1, \dots, \alpha_{n+1})^T$ is an eigenvector of $\mathbf{G}$ with eigenvalue 0. Without loss of generality we can, if necessary, reorder the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ and scale the α 's such that $1 = \alpha_1 \leq \alpha_i, i > 1$.

Let $\mathbf{u} \in \mathbb{R}^n$ be a unit vector and extend the Gram matrix $\mathbf{G} \in \mathbb{R}^{(n+1) \times (n+1)}$ to a Gram matrix

$$\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_{n+1}, \mathbf{u}) = \begin{pmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \dots & \mathbf{v}_1^T \mathbf{v}_{n+1} & \mathbf{v}_1^T \mathbf{u} \\ \mathbf{v}_2^T \mathbf{v}_1 & \dots & \mathbf{v}_2^T \mathbf{v}_{n+1} & \mathbf{v}_2^T \mathbf{u} \\ \vdots & & \vdots & \vdots \\ \mathbf{v}_{n+1}^T \mathbf{v}_1 & \dots & \mathbf{v}_{n+1}^T \mathbf{v}_{n+1} & \mathbf{v}_{n+1}^T \mathbf{u} \\ \mathbf{u}^T \mathbf{v}_1 & \dots & \mathbf{u}^T \mathbf{v}_{n+1} & \mathbf{u}^T \mathbf{u} \end{pmatrix}. \quad (2)$$

The matrix $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_{n+1}, \mathbf{u})$ is a positive semidefinite matrix, again with rank n . Since $\mathbf{v}_2, \dots, \mathbf{v}_{n+1}$ is a basis, there exists a vector $\mathbf{u}$ as described above where

$$\mathbf{v}_2^T \mathbf{u} = \mathbf{v}_3^T \mathbf{u} = \dots = \mathbf{v}_{n+1}^T \mathbf{u} = \gamma > 0.$$

Since $\gamma > 0$, it follows that $\beta = \mathbf{v}_1^T \mathbf{u} < 0$ since $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ is a positive basis. Since $\mathbf{v}_2, \dots, \mathbf{v}_{n+1}$ form a basis for $\mathbb{R}^n$ the latter statement is true for any γ , but $\mathbf{u}$ will be a unit vector only for one positive value of γ . We can obtain a bound for γ by noting that the matrix

$$\begin{pmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \mathbf{v}_1^T \mathbf{v}_2 & \cdots & \mathbf{v}_1^T \mathbf{v}_{n+1} & \beta \\ \mathbf{v}_2^T \mathbf{v}_1 & \mathbf{v}_2^T \mathbf{v}_2 & \cdots & \mathbf{v}_2^T \mathbf{v}_{n+1} & \gamma \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{v}_{n+1}^T \mathbf{v}_1 & \mathbf{v}_{n+1}^T \mathbf{v}_2 & \cdots & \mathbf{v}_{n+1}^T \mathbf{v}_{n+1} & \gamma \\ \beta & \gamma & \cdots & \gamma & 1 \end{pmatrix}$$

is a positive semidefinite matrix only if the vector $(\beta \ \gamma \ \dots \ \gamma)^T$ is orthogonal to the kernel of $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_{n+1})$. This means that $\beta\alpha_1 + \sum_{j=2}^{n+1} \gamma\alpha_j = 0$. Since $\alpha_1 = 1$, solving for γ gives $\gamma = \frac{-\beta}{\sum_{j=2}^{n+1} \alpha_j}$. Since $0 > \beta \geq -1$ and $\alpha_j \geq 1$ we obtain the bound $\gamma \leq \frac{1}{n}$ which shows that the cosine measure of $\mathbf{v}_1, \dots, \mathbf{v}_{n+1}$ is bounded above by $1/n$. The fact that $\beta \geq -1$ follows for instance by applying the Cauchy–Schwarz inequality to $\mathbf{v}_1^T \mathbf{u}$.

Proof of (2): $\gamma = 1/n$ only if $\alpha_2 = \dots = \alpha_n = 1$ and $\beta = -1$. In the first part of the proof, the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n+1}$ was, if necessary, reordered such that $\mathbf{v}_1$ was chosen such that $\alpha_1 = 1$. It is clear that in this case any of the vectors $\mathbf{v}_i$ can be interchanged with $\mathbf{v}_1$. Moreover, since all principal submatrices of a positive semidefinite matrix are again positive semidefinite this can only hold if $\mathbf{v}_i^T \mathbf{v}_j = -1/n$ for all $i \neq j$. This follows by considering the principal submatrix

$$\mathbf{G}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}) = \begin{pmatrix} 1 & \mathbf{v}_i^T \mathbf{v}_j & -1 \\ \mathbf{v}_j^T \mathbf{v}_i & 1 & 1/n \\ -1 & 1/n & 1 \end{pmatrix}$$

of the matrix given in (2). Since we have, $\mathbf{u}^T \mathbf{v}_i = -1$ it follows that $\mathbf{v}_i = -\mathbf{u}$. But then $\mathbf{u}^T \mathbf{v}_j = 1/n$ implies that $\mathbf{v}_j^T \mathbf{v}_i^T = -1/n$. (The fact that $\det \mathbf{G}(\mathbf{v}_i, \mathbf{v}_j, \mathbf{u}) = -(\mathbf{v}_i^T \mathbf{v}_j)^2 - (2/n)(\mathbf{v}_i^T \mathbf{v}_j) - (1/n^2) = -(\mathbf{v}_i^T \mathbf{v}_j + (1/n))^2$ must be nonnegative, since the determinant of a positive semidefinite matrix is always nonnegative, also provides the result.)

That there indeed exists a positive basis of unit vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n+1} \in \mathbb{R}^n$ such that $\mathbf{v}_i^T \mathbf{v}_j = -1/n$ for all i, j is proven in [1,5]. $\square$

Since every positive spanning set of $\mathbb{R}^n$ of size $n + 1$ must be a positive basis, we also can state that:

Corollary 1 *The maximal cosine measure for any positive spanning set of size $n + 1$ in $\mathbb{R}^n$ is $1/n$.*

Now we proceed to show the result for maximal positive bases. The maximal positive bases in $\mathbb{R}^n$ has size $2n$ and it consists of a set of n linearly independent vectors and their negatives (which can be multiplied by an arbitrary positive constant). The structure for maximal positive bases is conveniently stated in [12, Theorem 6.3]. They can be listed as

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, -\delta_1 \mathbf{v}_1, -\delta_2 \mathbf{v}_2, \dots, -\delta_n \mathbf{v}_n$$

where $\delta_i > 0$ for $1 \leq i \leq n$.

Theorem 2 *Let*

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, -\delta_1 \mathbf{v}_1, -\delta_2 \mathbf{v}_2, \dots, -\delta_n \mathbf{v}_n$$

where $\delta_i > 0$ for $1 \leq i \leq n$ be a (maximal) positive basis for $\mathbb{R}^n$. Then the following holds:

1. The cosine measure is bounded above by $1/\sqrt{n}$.
2. The bound is attained if and only if $\mathbf{v}_i^T \mathbf{v}_j = 0$ for $i \neq j$ and $1 \leq i, j \leq n$.

Proof By the definition of the cosine measure we will need the scaled positive basis vector $\frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$. Therefore we can without loss of generality we assume that $\mathbf{v}_1, \dots, \mathbf{v}_n$ are unit vectors and $\delta_i = 1$ for $1 \leq i \leq n$. The proof proceeds by showing that by making a proper selection of $\mathbf{v}_i$ or $-\mathbf{v}_i$ for $1 \leq i \leq n$ there exist a unit vector $\mathbf{u}$ with the property that $\mathbf{u}^T \mathbf{v}_1 = \mathbf{u}^T \mathbf{v}_2 = \dots = \mathbf{u}^T \mathbf{v}_n \leq 1/\sqrt{n}$.

That there exists a vector $\mathbf{u}$ such that $\mathbf{u}^T \mathbf{v}_1 = \mathbf{u}^T \mathbf{v}_2 = \dots = \mathbf{u}^T \mathbf{v}_n = \beta$ for any β follows immediately since the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent. If $\beta = 0$, the vector $\mathbf{u} = \mathbf{0}$. Otherwise, $\mathbf{u} \neq \mathbf{0}$, and scaling $\mathbf{u}$ to be a unit vector of course also scales β . The choice of β that provides a solution $\mathbf{u}$ that is a unit vector can be found by considering the Gram matrix $\mathbf{G}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n, \mathbf{u})$. This matrix has the required properties if $\mathbf{u}^T \mathbf{u} = 1$ and $\mathbf{u}^T \mathbf{v}_1 = \mathbf{u}^T \mathbf{v}_2 = \dots = \mathbf{u}^T \mathbf{v}_n = \beta$ if $\beta = 1/\sqrt{\mathbf{e}_n^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n)^{-1} \mathbf{e}_n}$ where $\mathbf{e}_n$ is the vector of length n with all its entries equal to one. The required upper bound will follow from Lemma 1 if we can select the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that $\mathbf{e}_n^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) \mathbf{e}_n \leq n$ by replacing some $\mathbf{v}_i$'s by $-\mathbf{v}_i$ if necessary.

We will prove by induction that this is possible for all $k \leq n$ to construct $\mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_k)$ such that $\mathbf{e}_k^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_k) \mathbf{e}_k \leq k$. For $k = 1$ there is nothing to prove, since $\mathbf{G}(\mathbf{v}_1) = 1$. For $k = 2$ the result follows by replacing $\mathbf{v}_2$ with $-\mathbf{v}_2$ if $\mathbf{v}_1^T \mathbf{v}_2 > 0$. Let us assume that $\mathbf{e}_{k-1}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1}) \mathbf{e}_{k-1} \leq k - 1$. Since

$$\mathbf{e}_k^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_k) \mathbf{e}_k = \mathbf{e}_{k-1}^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_{k-1}) \mathbf{e}_{k-1} + 2 \left(\mathbf{v}_k^T \mathbf{v}_1 + \dots + \mathbf{v}_k^T \mathbf{v}_{k-1} \right) + 1,$$

we see that the induction step is completed by replacing $\mathbf{v}_k$ with $-\mathbf{v}_k$ if $(\mathbf{v}_k^T \mathbf{v}_1 + \dots + \mathbf{v}_k^T \mathbf{v}_{k-1}) = \mathbf{v}_k^T (\mathbf{v}_1 + \dots + \mathbf{v}_{k-1}) > 0$, but keeping $\mathbf{v}_k$ otherwise.

To prove the second part of the theorem, one needs to do the minor change in the provided construction that one numbers the vectors such that $\mathbf{v}_1^T \mathbf{v}_2 \neq 0$, if no such pair of vectors exists, the proof is done. Otherwise, select $\mathbf{v}_1$ and $\mathbf{v}_2$ such that $\mathbf{v}_1^T \mathbf{v}_2 < 0$. This leads to $\mathbf{e}_2^T \mathbf{G}(\mathbf{v}_1, \mathbf{v}_2) \mathbf{e}_2 < 2$. Now we can repeat the induction step shown above, but we will have $\mathbf{e}_k^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_k) \mathbf{e}_k < k$, which in the end leads to $\mathbf{e}_n^T \mathbf{G}(\mathbf{v}_1, \dots, \mathbf{v}_n) \mathbf{e}_n < n$, and therefore we get $\beta < 1/\sqrt{n}$ by Lemma 1. $\square$

4 Discussion

This paper have presented the upper bounds for the cosine measures for the minimal and maximal positive bases. Obviously, the results provided here should be

complemented. One direction would be to find the maximal cosine measure for positive spanning sets consisting of $2n$ vectors in $\mathbb{R}^n$ as discussed in [7]. One can also look for extension of the results provided here to positive bases for $\mathbb{R}^n$ of size r for $n + 1 < r < 2n$.

That the results here cover the maximal and minimal positive bases is in good alignment with the fact that the structure in these two cases are better described than positive bases of intermediate size. The structures for the two cases discussed here are for instance provided in [12, Section 6].

Concerning the minimal cosine measures, both [5, 10] provides examples where the cosine measure becomes arbitrarily close to zero.

Acknowledgements The author is grateful for the efforts made by two anonymous reviewers in carefully reading the paper and providing suggestions that improved the final version. The author acknowledges the Research Council of Norway and the industry partners, ConocoPhillips Skandinavia AS, Aker BP ASA, Eni Norge AS, Total E&P Norge AS, Equinor ASA, Neptune Energy Norge AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, Wintershall Norge AS, and DEA Norge AS, of The National IOR Centre of Norway for support.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Alberto, P., Nogueira, F., Rocha, H., Vicente, L.N.: Pattern search methods for user-provided points: application to molecular geometry problems. *SIAM J. Optim.* **14**(4), 1216–1236 (2004)
2. Audet, C.: A short proof on the cardinality of maximal positive bases. *Optim. Lett.* **5**, 191–194 (2011)
3. Audet, C., Hare, W.: *Derivative-Free and Blackbox Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, Berlin (2017)
4. Audet, C., Ianni, A., Digabel, S.L., Tribes, C.: Reducing the number of function evaluations in mesh adaptive direct search algorithms. *SIAM J. Optim.* **24**(2), 621–642 (2014)
5. Conn, A.R., Scheinberg, K., Vicente, L.N.: *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on optimization, vol. 8. Society of Industrial and Applied Mathematics and the Mathematical Programming Society, Philadelphia (2009)
6. Davis, C.: Theory of positive linear dependence. *Am. Math. J.* **76**(4), 733–746 (1954)
7. Dodangeh, M., Vicente, L.N., Zhang, Z.: On the optimal order of worst case complexity of direct search. *Optim. Lett.* **10**, 699–708 (2016)
8. Hare, W., Song, H.: On the cardinality of positively linearly independent sets. *Optim. Lett.* **10**(4), 649–654 (2016)
9. Horn, R.A., Johnson, C.R.: *Matrix Analysis*, 2nd edn. Cambridge University Press, Cambridge (2013)
10. Kolda, T.G., Lewis, R.M., Torczon, V.: Optimization by direct search: new perspectives on some classical and modern methods. *SIAM Rev.* **45**(3), 385–482 (2003)
11. Lewis, R.M., Torczon, V.: Rank ordering and positive bases in pattern search algorithms. NASA Contractor Report 201628, ICASE Report No. 96-71, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center (1999)
12. Regis, R.G.: On the properties of positive spanning sets and positive bases. *Optim. Eng.* **17**, 229–262 (2016)