Fundamental Convex Euclidean Geometry and Semidefinite Programming

Jon Dattorro
Stephen P. Boyd

2002

1 Introduction

Rockafellar's watershed.
Boyd's multitude of applications expressible as convex problem.
Brief history in [Urruty] identifies Hermann Minkowski (1864-1909), Werner Fenchel (1905-1988), Jean-Jacques Moreau (1923-), R. Tyrrell Rockafellar (1935-).

2 Convexity theorems and definitions

There is relatively less published pertaining to *matrix*-valued convex sets and functions. We present results that we may later need. The reader is referred to [1] [2] [3] for a comprehensive treatment of convexity.

2.1 Sets

Definition. *Convex set.* A set $\mathcal{C}$ is convex iff for all $Y, Z \in \mathcal{C}$ and $0 \leq \mu \leq 1$,

$$\mu Y + (1-\mu)Z \in \mathcal{C} \tag{1}$$

Apparent from the definition, a convex set is a connected set.

Theorem. *Intersection.* The intersection of an arbitrary collection of convex sets is convex. [2, §2][1, §2]

Theorem. *Image/inverse image.* Let f be a mapping from $\mathbb{R}^{p\times k}$ to $\mathbb{R}^{m\times n}$. [2, §3]

- The image of a convex set $\mathcal{C}$ under any affine function

$$f(\mathcal{C}) = \{f(X) \mid X \in \mathcal{C}\} \tag{2}$$

  is convex.

- The inverse image of a convex set $\mathcal{F}$,

$$f^{-1}(\mathcal{F}) = \{X \mid f(X) \in \mathcal{F}\} \tag{3}$$

  a single or many-valued mapping, under any affine function f is convex.

Corollary. *Projection.* The orthogonal projection of a convex set on a *subspace*[1] is another convex set. [2, §3]

[1] A nonempty subset of a vector space is called a subspace if every vector of the form $\alpha x + \beta y$ is in the subset whenever x and y are both in the subset, and $\alpha, \beta \in \mathbb{R}$. [4, §2.3] A subspace contains the origin by definition.

Corollary. *Matrix-valued set convexity.* Set $\mathcal{C} \subseteq \mathbb{R}^{m\times n}$ is convex if and only if the set $v^T\mathcal{C}w = <vw^T, \mathcal{C}> \subseteq \mathbb{R}$ is convex for each and every vector $v\in\mathbb{R}^m$ and $w\in\mathbb{R}^n$; more generally, if and only if the set of inner products $<E, \mathcal{C}> \subseteq \mathbb{R}$ is convex for each and every matrix $E\in\mathbb{R}^{m\times n}$.

Proof. This corollary is an application of the image theorem under linear transformation $f(\mathcal{C}) = v^T\mathcal{C}w$. The requirement for *every* v and w precludes trivial transformations. Consider any particular v and w and take any two elements $\mathcal{C}_1$ and $\mathcal{C}_2$ from $\mathcal{C}$. Then

$$v^T\mathcal{C}_1 w \;=\; \operatorname{tr}(wv^T\mathcal{C}_1) \;=\; \mathbf{1}^T\big((vw^T)\circ\mathcal{C}_1\big)\,\mathbf{1} \;=<vw^T, \mathcal{C}_1> \tag{4}$$

where $\circ$ denotes the Hadamard product[2] of matrices [5] [6, §1.1.4], and

$$<A, B> \overset{\Delta}{=} \operatorname{tr}(A^TB) = \operatorname{tr}(BA^T) = \operatorname{tr}(B^TA) = \mathbf{1}^T(A\circ B)\mathbf{1} \tag{5}$$

denotes the inner product of matrices of like size. [1, §2]

The inner products form a convex set which is easy to see by forming convex combinations of them; *viz.*,

$$\begin{aligned}\mu <vw^T, \mathcal{C}_1> + (1-\mu)<vw^T, \mathcal{C}_2> \;&=\; <vw^T, \mu\,\mathcal{C}_1> + <vw^T, (1-\mu)\mathcal{C}_2> \\ &=\; <vw^T, \mu\,\mathcal{C}_1 + (1-\mu)\mathcal{C}_2>\end{aligned} \tag{6}$$

an inner product of vw^T with an element from $\mathcal{C}$ for $0\leq\mu\leq 1$ when $\mathcal{C}$ is convex.

Going the other way, we assume that the inner products form a convex set generated from some set $\mathcal{C}$. Take a convex combination of any two elements from $\mathcal{C}$, then form the inner product,

$$<vw^T, \mu\,\mathcal{C}_1 + (1-\mu)\mathcal{C}_2> \;=\; \mu <vw^T, \mathcal{C}_1> + (1-\mu)<vw^T, \mathcal{C}_2> \tag{7}$$

Because the right-hand side of (7) is a member of a convex set for $0\leq\mu\leq 1$ by assumption, then $\mu\,\mathcal{C}_1 + (1-\mu)\mathcal{C}_2$ must belong to $\mathcal{C}$. Since all such convex combinations belong to $\mathcal{C}$ for any $\mathcal{C}_1$ and $\mathcal{C}_2$ from $\mathcal{C}$, then it must be convex.

The generalization from vw^T to matrix E is trivial. ♦

[2] The Hadamard product is a simple entry-wise product of corresponding entries from two matrices of like size; *id est*, not necessarily square.

Definition. *Symmetric matrix subspace.* Define a subspace of $\mathbb{R}^{N\times N}$: the set of all symmetric $N\times N$ matrices;

$$\mathbb{S}^N \triangleq \left\{A = A^T \in \mathbb{R}^{N\times N}\right\} \tag{8}$$

Definition. *Hollow matrix subspace.* [7] Similarly, define a subspace of $\mathbb{S}^N$: the set of all symmetric $N\times N$ matrices having zero main diagonal;

$$\mathbb{S}^N_\delta \triangleq \left\{A \in \mathbb{S}^N \mid \delta(A) = 0\right\} \tag{9}$$

Definition. *Positive semidefinite cone.* The set of all symmetric positive semidefinite matrices of particular dimension is called the positive semidefinite cone: (*confer* §8.1)

$$\mathbb{S}^N_+ \triangleq \left\{A\in \mathbb{S}^N \mid A \succeq 0\right\} \tag{10}$$

The convex set $\mathbb{S}^N_+$ forms a *proper* cone [1, §2][3] in the vector space $\mathbb{R}^{N(N+1)/2}$ whose dimension is the number of free variables in a symmetric $N\times N$ matrix.

[3]A proper cone is convex, closed, has nonempty interior, [sic] and pointed (contains no line). A proper cone induces vector or matrix inequality (a partial ordering on set $\mathbb{R}^N$ or $\mathbb{S}^N$) by making comparable points and a minimum element well defined.

2.2 Functions

The vector-valued function $f(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{R}^M$ is convex in X if and only if $\text{dom}\, f$ is a convex set and for all $Y, Z \in \text{dom}\, f$ and $0 \leq \mu \leq 1$,

$$f(\mu Y + (1-\mu)Z) \;\preceq\; \mu f(Y) + (1-\mu)f(Z) \tag{11}$$

Since comparison of vectors here is with respect to the nonnegative orthant $\mathbb{R}^M_+$, the same test can be accomplished by separately comparing each element of the vector function. The vector-valued function case is therefore a straightforward generalization of conventional convexity theory for a real function. (See [1, §3] for more details.)

Definition. *Convex function.*
1) *Epigraph.* We define $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ to be a convex function of X iff its epigraph

$$\text{epi}\, g \;\stackrel{\Delta}{=}\; \{(X, t) \mid X \in \text{dom}\, g,\; g(X) \preceq tI\} \;\in\; \mathbb{R}^{p\times k} \times \mathbb{R} \tag{12}$$

forms a convex set.
2) *Inequality form.* A function $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ is convex in X iff $\text{dom}\, g$ is a convex set and for all $Y, Z \in \text{dom}\, g$ and $0 \leq \mu \leq 1$,

$$g(\mu Y + (1-\mu)Z) \;\preceq\; \mu\, g(Y) + (1-\mu)g(Z) \tag{13}$$

We require $g(X) \in \mathbb{S}^M$ because we compare matrix-valued functions with respect to the positive semidefinite cone in the subspace of symmetric matrices (§8.1). The epigraph of a real function is treated in [1] [2] [4]. A convex function is continuous on the relative interior of its domain. [2, §10]

Theorem. *Positive semidefinite matrix.* [1, §2] (*confer*, §5.2) Matrix $A \in \mathbb{S}^M$ is positive semidefinite ($A \succeq 0$) if and only if for each and every $E \succeq 0$,

$$<E,\, A> \;=\; \text{tr}(E^T A) \;\geq\; 0 \tag{14}$$

The product AB of positive (semi)definite matrices A and B is positive (semi)definite when $AB = BA$. [8, §5.2] AB is commutative when AB, A, and

B are symmetric. This product rule applies to longer products. If A is positive (semi)definite, then for all matrices X, X^TAX is positive (semi)definite.

Scalar-Definition. *Matrix-valued function convexity.* $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ is convex in X iff $w^T g(X) w \in \mathbb{R}$ is convex in X for each and every $w \in \mathbb{R}^M$; [1, §3] more generally, iff the real function $<E,\, g(X)>$ is convex in X for each and every $E \succeq 0$.

This general scalar-definition follows directly from (13) and the *positive semidefinite matrix theorem*, with no presumption of differentiability.

Line Theorem. [1, §3] $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ is convex in X if and only if it remains convex on the intersection of any line with its domain.

Definition. *Differentiable convex function.* $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ is convex in X iff $\operatorname{dom} g$ is an open convex set, and its second derivative along every line $X + t\,Y$ that intersects $\operatorname{dom} g$, $g''(X + t\,Y) : \mathbb{R} \rightarrow \mathbb{S}^M$, is positive semidefinite on each point of intersection; *id est*, iff for each and every $X, Y \in \mathbb{R}^{p\times k}$ such that $X + t\,Y \in \operatorname{dom} g$ over some open interval of $t \in \mathbb{R}$,

$$\frac{d^2}{dt^2}\, g(X + t\,Y) \succeq 0 \tag{15}$$

Example. *Matrix inverse.* $g(X) = X^{-1}$ on $\{X \in \mathbb{S}^M \mid X \succ 0\}$. For all $Y \in \mathbb{S}^M$,

$$\frac{d^2}{dt^2}\, g(X + t\,Y) \;=\; 2(X + t\,Y)^{-1} Y (X + t\,Y)^{-1} Y (X + t\,Y)^{-1} \;\succeq\; 0 \tag{16}$$

on some open interval of $t \in \mathbb{R}$ such that $X + t\,Y \succ 0$. Hence, $g(X)$ is convex in X. This result is extensible;[4] $\operatorname{tr} X^{-1}$ is convex on $X \succ 0$. [5, §7.6, Prob.2]

Example. *Matrix exponential.* $g(X) = e^X$ on $\{X \in \mathbb{S}^M \mid X \succ 0\}$. ...

There are more methods to determine function convexity, [1, §3] [2] each of them efficient when appropriate.

---

[4] $d/dt \operatorname{tr} g(X + t\,Y) = \operatorname{tr} d/dt\, g(X + t\,Y)$. (App. H)

2.2.1 Quasiconvex functions

Quasiconvex functions are useful in practical problem solving because they are *unimodal*; a global minimum is guaranteed to exist over the function domain or over any convex set in the function domain. In terms of *sublevel set*, their definition is elegant and parallels the epigraph-form definition for convex functions:

Definition. *Quasiconvex function.*
1) *Sublevel set.* We define $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ to be a quasiconvex function of matrix X iff $\operatorname{dom} g$ is a convex set and for each and every $\nu \in \mathbb{R}$ the corresponding sublevel set

$$\mathcal{L}_\nu \triangleq \left\{X \in \mathbb{R}^{p\times k} \mid g(X) \preceq \nu I\right\} \tag{17}$$

is convex.
2) *Inequality form.* $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ is a quasiconvex function of matrix X iff $\operatorname{dom} g$ is a convex set and for all $Y, Z \in \operatorname{dom} g$ and $0 \leq \mu \leq 1$,

$$g(\mu Y + (1-\mu)Z) \preceq \max\{g(Y),\, g(Z)\} \tag{18}$$

A quasiconvex function is not necessarily continuous.

Definition. *Differentiable quasiconvex function.* Declare $t \in \mathbb{R}$ variable and $\operatorname{dom} g$ an open and necessarily convex set. Then $g(X) : \mathbb{R}^{p\times k} \rightarrow \mathbb{S}^M$ is quasiconvex in X if wherever in the domain the *directional derivative*[5] [9, §A.5] [10] [11] [12, §2.3.4] becomes zero, the second directional derivative is positive definite there [1, §3] in the same direction Y; *id est*, if

$$\left.\frac{d}{dt}\right|_{t=0} g(X + t\,Y) = 0 \quad \Rightarrow \quad \left.\frac{d^2}{dt^2}\right|_{t=0} g(X + t\,Y) \succ 0 \tag{19}$$

for each and every point $X \in \operatorname{dom} g$ and all nonzero $Y \in \mathbb{R}^{p\times k}$.

Conversely, if $g(X)$ is quasiconvex then for each and every $X \in \operatorname{dom} g$ and all $Y \in \mathbb{R}^{p\times k}$,

$$\left.\frac{d}{dt}\right|_{t=0} g(X + t\,Y) = 0 \quad \Rightarrow \quad \left.\frac{d^2}{dt^2}\right|_{t=0} g(X + t\,Y) \succeq 0 \tag{20}$$

[5] By using a generalization of the Taylor series in Appendix A, we extend the traditional definition of directional derivative so that *direction* may be indicated by a vector or matrix.

2.2.2 Salient properties of convex and quasiconvex functions

1. g convex $\Leftrightarrow$ $-g$ concave.
   g quasiconvex $\Leftrightarrow$ $-g$ quasiconcave.

2. Convexity $\Rightarrow$ quasiconvexity.
   Concavity $\Rightarrow$ quasiconcavity.

3. The definition of quasiconvexity is exactly like that of convexity for the scalar-definition and for the line theorem in §2.2. [1, §3]

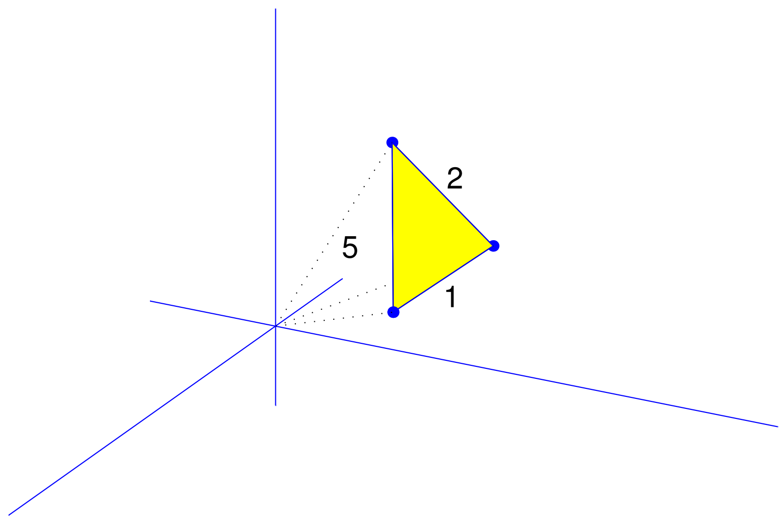

Figure 1: Convex hull of three points ($N=3$) is shaded in $\mathbb{R}^3$ ($n=3$). Dotted lines are imagined vectors to points.

3 Euclidean Distance Matrix

Euclidean space is a finite-dimensional vector space having both a metric and inner-product defined on it. A Euclidean distance matrix, an EDM $\in \mathbb{R}_+^{N\times N}$, is an exhaustive table of distance-squared between points taken by pair from a list of N points in Euclidean space $\mathbb{R}^n$. Each point is labelled ordinally, hence the row or column index of an EDM, i or $j \in 1\ldots N$, individually addresses all the points in the list.

Consider the following example of an EDM for the case $N = 3$:

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} 0 & d_{12} & d_{13} \\ d_{12} & 0 & d_{23} \\ d_{13} & d_{23} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 5 \\ 1 & 0 & 4 \\ 5 & 4 & 0 \end{bmatrix} \qquad (21)$$

Observe that D has N^2 entries but only $N(N-1)/2$ pieces of information. In Figure 1 we show three points in $\mathbb{R}^3$ that can be arranged in a list to correspond to D in (21). Such a list is not unique because any rotation, reflection, or offset of the points in Figure 1 would produce the same D.

3.1 Metric space requirements

For $i,j \in 1 \dots N$, the Euclidean distance between points x_i and x_j must satisfy the axiomatic requirements imposed by any metric space: [13]

1. $\sqrt{d_{ij}} \geq 0,\ i \neq j$ — nonnegativity
2. $\sqrt{d_{ij}} = 0,\ i = j$ — self-distance
3. $\sqrt{d_{ij}} = \sqrt{d_{ji}}$ — symmetry
4. $\sqrt{d_{ij}} \leq \sqrt{d_{ik}} + \sqrt{d_{kj}}\ ,\ i \neq j \neq k$ — triangle inequality

where $\sqrt{d_{ij}}$ is the Euclidean metric in $\mathbb{R}^n$ (§3.3). Suppose we let d_{ij} denote distance-squared[6] and the i,j^{th} entry of an EDM D. Then all entries of an EDM must be in concord with the axioms: specifically, each entry must be nonnegative, the main diagonal must be zero,[7] and an EDM must be symmetric. The fourth axiom provides upper and lower bounds for each entry (§9); loose bounds when $N > 3$. Axiom 4 is more generally true when there are no restrictions on indices i, j, k but furnishes no new information.

[6] Despite the terminology "distance-squared", the definition adopted throughout is

$$\sqrt{d} \triangleq |d|^{1/2}\, e^{\imath\, \mathrm{Arg}(d)/2}$$

where $\imath = \sqrt{-1}$. Thus we consistently make the presumption: $\sqrt{d_{kj}} \geq 0 \Leftrightarrow d_{kj} \geq 0$.

[7] What we call zero self-distance, Marsden calls *nondegeneracy*. [14, §1.6]

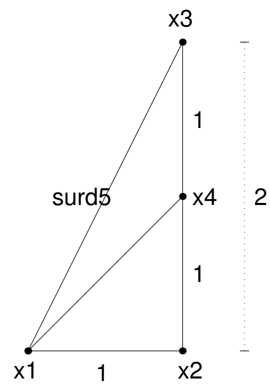

Figure 2: Four axioms of Euclidean space are not a recipe for reconstruction of this polyhedron.

3.2 $\exists$ *fifth Euclidean requirement*

The four axioms of Euclidean space provide insufficient information to reconstruct a *convex polyhedron* more complex than a triangle, from incomplete distance information. Yet any list of points or the vertices of any polyhedron must conform to the axioms.

Example. *Triangle.* Consider the EDM in (21), but missing one of its entries:

$$D = \begin{bmatrix} 0 & 1 & d_{13} \\ 1 & 0 & 4 \\ d_{31} & 4 & 0 \end{bmatrix} \tag{22}$$

Can we determine the unknown entries of D by applying the axioms? (§9) Axiom 1 demands $\sqrt{d_{13}}, \sqrt{d_{31}} \geq 0$, axiom 2 requires the main diagonal be zero, while axiom 3 makes $\sqrt{d_{31}} = \sqrt{d_{13}}$. The fourth axiom tells us

$$1 \leq \sqrt{d_{13}} \leq 3 \tag{23}$$

Indeed, described over that closed interval $[1, 3]$ is a family of triangular polyhedra whose angle at vertex x_2 varies from 0 to π radians.

Example. *Small completion problem.* Now consider the polyhedron in Figure 2 formed from an unknown list of four points $\{x_1, x_2, x_3, x_4\}$. The corresponding EDM less one critical piece of information, d_{14}, is given by

$$D = \begin{bmatrix} 0 & 1 & 5 & d_{14} \\ 1 & 0 & 4 & 1 \\ 5 & 4 & 0 & 1 \\ d_{14} & 1 & 1 & 0 \end{bmatrix} \tag{24}$$

From axiom 4 we may write a few inequalities for the two triangles common to d_{14}; we find

$$\sqrt{5}-1 \;\leq\; \sqrt{d_{14}} \;\leq\; 2 \tag{25}$$

We cannot further narrow those loose bounds on $\sqrt{d_{14}}$ using only the four axioms. (App. C.1.1) Yet there is only one possible choice for $\sqrt{d_{14}}$ because points x_2, x_3, x_4 must be collinear. All other values of $\sqrt{d_{14}}$ in the interval $[\sqrt{5}-1,\, 2]$ specify impossible distances in any dimension; *id est*, in this example the triangle inequality axiom does *not* yield an interval for $\sqrt{d_{14}}$ over which a family of convex polyhedra can be reconstructed.

We will return to this simple example to illustrate more elegant methods of solution in §5.4 and §6.4.1.

3.2.1 Lookahead

There must exist at least one requirement more than the four axioms of Euclidean space that makes them altogether necessary and sufficient to reconstruct convex polyhedra. We will return to that question in §6. One of our early objectives is to determine the matrix criteria which subsume all the Euclidean axioms and requirements. Once found (§5.3), we will see there is a bridge from convex polyhedra to EDMs.[8]

We digress to review some invaluable concepts and to link the axioms to matrix criteria.

---

[8]From an EDM, a generating list (§4.5, §4.6) can be found (§5.3.4) correct to within an offset, rotation, and reflection (§3.4).

3.3 EDM definition

Ascribe points in a list $\{x_\ell \in \mathbb{R}^n,\ \ell = 1 \ldots N\}$ to the columns of a matrix X;

$$X = [x_1 \ \cdots \ x_N] \in \mathbb{R}^{n\times N} \tag{26}$$

When D is an EDM, its entries d_{ij} must be related to those points constituting the list by the Euclidean metric-squared:

$$\begin{aligned} d_{ij} &= \|x_i - x_j\|^2 = (x_i - x_j)^T(x_i - x_j) = \|x_i\|^2 + \|x_j\|^2 - 2x_i^T x_j \\ &= \begin{bmatrix} x_i^T & x_j^T \end{bmatrix} \begin{bmatrix} I & -I \\ -I & I \end{bmatrix} \begin{bmatrix} x_i \\ x_j \end{bmatrix} \end{aligned} \tag{27}$$

Thus each entry d_{ij} is a convex (§2.2) quadratic function of $\begin{bmatrix} x_i \\ x_j \end{bmatrix} \in \mathbb{R}^{2n}$. [1, §3] [2, §6] In terms of matrices, the collection of all Euclidean distance matrices $\mathbf{EDM}^N$ is a convex subset of $\mathbb{R}_+^{N\times N}$, hence not a subspace;

$$0 \in \mathbf{EDM}^N \subseteq \left(\mathbb{S}_\delta^N \cap \mathbb{R}_+^{N\times N}\right) \subset \mathbb{S}^N \tag{28}$$

$D \in \mathbf{EDM}^N$ must be expressible as a function of some X; *id est*, it must have the form

$$\mathcal{D}(X) = \delta(X^T X)\mathbf{1}^T + \mathbf{1}\delta^T(X^T X) - 2X^T X \ \in \mathbf{EDM}^N \tag{29}$$

where the main diagonal of $A \in \mathbb{R}^{N\times N}$ is denoted[9]

$$\delta(A) \in \mathbb{R}^N \tag{30}$$

and

$$X^T X = \begin{bmatrix} \|x_1\|^2 & x_1^T x_2 & x_1^T x_3 & \cdots & x_1^T x_N \\ x_2^T x_1 & \|x_2\|^2 & x_2^T x_3 & \cdots & x_2^T x_N \\ x_3^T x_1 & x_3^T x_2 & \|x_3\|^2 & \ddots & x_3^T x_N \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ x_N^T x_1 & x_N^T x_2 & x_N^T x_3 & \cdots & \|x_N\|^2 \end{bmatrix} \in \mathbb{R}^{N\times N} \tag{31}$$

[9] When linear function $\delta()$ operates on a square matrix, it returns a vector composed of all the entries from the main diagonal. On a diagonal matrix $\Lambda \in \mathbb{R}^{N\times N}$,

$$\delta(\delta(\Lambda)) \stackrel{\Delta}{=} \Lambda \in \mathbb{R}^{N\times N}$$

Operating on a vector, $\delta()$ naturally returns a diagonal matrix.

Conversely, $\mathcal{D}(X)$ in (29) will make an EDM for any $X \in \mathbb{R}^{n\times N}$, but $\mathcal{D}(X)$ is not a convex function of X (§3.5). $\mathcal{D}(X)$ is the matrix definition of EDM and so conforms to the Euclidean axioms:

Nonnegativity of EDM entries (axiom 1, §3.1) is obvious from the distance-square definition (27), and so assumed to hold for any D expressible in the form $\mathcal{D}(X)$ in (29).

When we say D is an EDM, reading from (29), it implicitly means the main diagonal must be zero (axiom 2, self-distance) and D must be symmetric (axiom 3); $\delta(D) = 0$ and $D^T = D$ are necessary matrix criteria.

3.3.1 Inner-product form

Equivalent to (27) is [15, §1-7][8, §3.2]

$$\begin{aligned} d_{ij} &= d_{ik} + d_{kj} - 2\sqrt{d_{ik}d_{kj}}\cos\theta_{ikj} \\ &= \begin{bmatrix} \sqrt{d_{ik}} & \sqrt{d_{kj}} \end{bmatrix} \begin{bmatrix} 1 & -e^{\imath\theta_{ikj}} \\ -e^{-\imath\theta_{ikj}} & 1 \end{bmatrix} \begin{bmatrix} \sqrt{d_{ik}} \\ \sqrt{d_{kj}} \end{bmatrix} \end{aligned} \tag{32}$$

called the *law of cosines*, where $\imath = \sqrt{-1}$, i, k, j are positive integers, and θ_{ikj} is the angle at vertex x_k formed by vectors $x_i - x_k$ and $x_j - x_k$;

$$\cos\theta_{ikj} = \frac{\frac{1}{2}(d_{ik} + d_{kj} - d_{ij})}{\sqrt{d_{ik}d_{kj}}} = \frac{(x_i - x_k)^T(x_j - x_k)}{\|x_i - x_k\|\,\|x_j - x_k\|} \tag{33}$$

$d_{ij}\left(\begin{bmatrix} \sqrt{d_{ik}} \\ \sqrt{d_{kj}} \end{bmatrix}\right)$ is a convex (§2.2) quadratic function on $\mathbb{R}^2_+$. $d_{ij}(\theta_{ikj})$ is a quasiconvex function (§2.2.1) [1, §3] minimized over domain $-\pi \leq \theta_{ikj} \leq \pi$ when $\theta_{ikj} = 0$, we have the Pythagorean theorem when $\theta_{ikj} = \pm\pi/2$, and $d_{ij}(\theta_{ikj})$ is maximized when $\theta_{ikj} = \pm\pi$;

$$\begin{array}{ll} d_{ij} = \left(\sqrt{d_{ik}} + \sqrt{d_{kj}}\right)^2, & \theta_{ikj} = \pm\pi \\ d_{ij} = d_{ik} + d_{kj}\,, & \theta_{ikj} = \pm\frac{\pi}{2} \\ d_{ij} = \left(\sqrt{d_{ik}} - \sqrt{d_{kj}}\right)^2, & \theta_{ikj} = \;\;0 \end{array} \tag{34}$$

so

$$|\sqrt{d_{ik}} - \sqrt{d_{kj}}| \;\leq\; \sqrt{d_{ij}} \;\leq\; \sqrt{d_{ik}} + \sqrt{d_{kj}} \tag{35}$$

Hence the triangle inequality, axiom 4 (§3.1, *confer* (101a)) of Euclidean space, holds for any EDM D.

We may construct the inner-product form of the EDM definition for matrices by evaluating (32) for $k=1$:

$$\mathcal{D}(\Theta) \triangleq \begin{bmatrix} 0 \\ \delta(\Theta^T\Theta) \end{bmatrix} \mathbf{1}^T + \mathbf{1}\left[0 \quad \delta^T(\Theta^T\Theta)\right] - 2\begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & \Theta^T\Theta \end{bmatrix} \in \mathbf{EDM}^N \tag{36}$$

for which all Euclidean axioms hold, and where

$$\Theta^T\Theta =$$

$$\begin{bmatrix} d_{12} & \sqrt{d_{12}d_{13}}\cos\theta_{213} & \sqrt{d_{12}d_{14}}\cos\theta_{214} & \cdots & \sqrt{d_{12}d_{1N}}\cos\theta_{21N} \\ \sqrt{d_{12}d_{13}}\cos\theta_{213} & d_{13} & \sqrt{d_{13}d_{14}}\cos\theta_{314} & \cdots & \sqrt{d_{13}d_{1N}}\cos\theta_{31N} \\ \sqrt{d_{12}d_{14}}\cos\theta_{214} & \sqrt{d_{13}d_{14}}\cos\theta_{314} & d_{14} & \ddots & \sqrt{d_{14}d_{1N}}\cos\theta_{41N} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \sqrt{d_{12}d_{1N}}\cos\theta_{21N} & \sqrt{d_{13}d_{1N}}\cos\theta_{31N} & \sqrt{d_{14}d_{1N}}\cos\theta_{41N} & \cdots & d_{1N} \end{bmatrix}$$

$$\in \mathbb{R}^{N-1\times N-1} \tag{37}$$

The entries of $\Theta^T\Theta$ result from inner products as in (33); *id est*,

$$\Theta = [x_2 - x_1 \quad x_3 - x_1 \quad \cdots \quad x_N - x_1] \in \mathbb{R}^{n\times N-1} \tag{38}$$

Like $\mathcal{D}(X)$ (29), $\mathcal{D}(\Theta)$ will make an EDM for any $\Theta\in\mathbb{R}^{n\times N-1}$ and is neither a convex function of Θ (§3.5.1). Scrutinizing $\Theta^T\Theta$ we find that because of the choice $k=1$, distances therein are all with respect to point x_1. Similarly, angles in $\Theta^T\Theta$ are between all vector pairs having vertex x_1. Yet picking arbitrary θ_{i1j} and d_{ij} to fill $\Theta^T\Theta$ will not necessarily make an EDM. We deduce that knowledge of inter-point distance is equivalent to knowledge of distance and angle from the perspective of one point, x_1 in our case. The total amount of information in $\Theta^T\Theta$, $N(N-1)/2$, is unchanged[10] with respect to EDM D.

[10] The reason for the amount $O(N^2)$ information is because of the *relative* measurements. The use of a fixed reference in the measurement of angles and distances would reduce the required information but is antithetical. In the particular case $n=2$, for example, ordering all points x_ℓ in a length-N list by increasing angle of vector $x_\ell - x_1$ with respect to $x_2 - x_1$, θ_{i1j} becomes equivalent to $\sum_{k=i}^{j-1}\theta_{k,1,k+1} \leq 2\pi$ and the amount of information is reduced to $2N-3$; rather, $O(N)$.

3.4 Rotation, reflection, offset invariance

When D is an EDM, there exist an infinite number of corresponding N-point lists in Euclidean space. All those lists are related by rotation, reflection, and offset (*translation* or shift).

3.4.1 Offset

If there were a common offset among all the x_ℓ, it would be cancelled in the formation of each d_{ij}. Proof follows directly from (27). Knowing that offset α in advance, we may remove it from the list in X by subtracting $\alpha\mathbf{1}^T$. Then it stands to reason by definition (29) of an EDM, for any offset $\alpha \in \mathbb{R}^n$,

$$\mathcal{D}(X - \alpha\mathbf{1}^T) = \mathcal{D}(X) \tag{39}$$

In words, inter-point distances are unaffected by translation; EDM D is *offset invariant*. When $\alpha = x_1$, in particular,

$$\mathcal{D}(X - x_1\mathbf{1}^T) \;=\; \mathcal{D}(X - Xe_1\mathbf{1}^T) \;=\; \mathcal{D}\big(X\,[\,\mathbf{0} \quad \sqrt{2}V_{\mathcal{N}}\,]\big) \;=\; \mathcal{D}(X) \tag{40}$$

where we introduce the full-rank *skinny* matrix[11]

$$V_{\mathcal{N}} \;\triangleq\; \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & -1 & \cdots & -1 \\ 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix} \;=\; \frac{1}{\sqrt{2}}\begin{bmatrix} -\mathbf{1}^T \\ I \end{bmatrix} \in \mathbb{R}^{N\times N-1} \tag{41}$$

having range

$$\mathcal{R}(V_{\mathcal{N}}) = \mathcal{N}(\mathbf{1}^T), \quad \mathcal{N}(\mathbf{1}^T) \perp \mathcal{R}(\mathbf{1}) \tag{42}$$

(and nullspace $\mathcal{N}(V_{\mathcal{N}}) = 0$), and where

$$e_1 \triangleq \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{43}$$

For the inner-product form EDM definition, $\mathcal{D}(\Theta)$ is not offset invariant in the following sense: For $\alpha \in \mathbb{R}^n$ it is generally true that

$$\mathcal{D}(\Theta - \alpha\mathbf{1}^T) \neq \mathcal{D}(\Theta) \tag{44}$$

[11] "Skinny" meaning more rows than columns. In §D.2, properties of $V_{\mathcal{N}}$ can be found.

3.4.2 Rotation/Reflection

Rotation of the list in X about some arbitrary point, or reflection through some hyperplane is accomplished via $QX - \alpha\mathbf{1}^T$, where Q is an *orthogonal matrix*. [8] We rightfully expect

$$\mathcal{D}(QX - \alpha\mathbf{1}^T) = \mathcal{D}\big(Q(X - \alpha\mathbf{1}^T)\big) = \mathcal{D}(QX) = \mathcal{D}(X) \tag{45}$$

Because $\mathcal{D}(X)$ is offset invariant, we may safely ignore offset and consider only the impact of matrices that pre-multiply X. Inter-point distances are unaffected by rotation or reflection; we say, EDM D is *rotation/reflection invariant*. Proof follows from the fact,[12] $Q^T = Q^{-1} \Rightarrow X^TQ^TQX = X^TX$. So (45) follows directly from (29).

The class of pre-multiplying matrices for which inter-point distances are unaffected is a little more broad than orthogonal matrices. Looking at EDM definition (29), it appears that any matrix Q_o such that

$$X^T Q_o^T Q_o X = X^T X \tag{46}$$

will have the property

$$\mathcal{D}(Q_o X) = \mathcal{D}(X) \tag{47}$$

An example is $Q_o \in \mathbb{R}^{m\times n}$ $(m > n)$ having orthonormal columns.

Likewise, $\mathcal{D}(\Theta)$ (36) is rotation/reflection invariant;

$$\mathcal{D}(Q\Theta) = \mathcal{D}(\Theta) \tag{48}$$

so (46) and (47) would similarly apply.

3.4.3 Invariance conclusion

In the construction of an EDM, absolute rotation, reflection, or offset information is lost. Reconstruction of point position, the list in X, can be guaranteed correct only in the affine dimension r; *id est*, in relative position.

[12] The characteristic $Q^{-1} = Q^T$ defines an orthogonal matrix Q. Curiously, Q^T is itself an orthogonal matrix.

3.5 $-V_{\mathcal{N}}^T\mathcal{D}(X)V_{\mathcal{N}}$ convexity

In §3.3 we saw that the EDM entries $d_{ij}\left(\begin{bmatrix} x_i \\ x_j \end{bmatrix}\right)$ are convex quadratic functions. Yet $-\mathcal{D}(X)$ (29) is not a quasiconvex function of matrix X because the second directional derivative (§2.2.1)

$$-\frac{d^2}{dt^2}\bigg|_{t=0}\mathcal{D}(X+t\,Y) = 2\left(-\delta(Y^TY)\mathbf{1}^T - \mathbf{1}\delta^T(Y^TY) + 2Y^TY\right) \tag{49}$$

is *indefinite* for any $Y\in\mathbb{R}^{n\times N}$ because the main diagonal is zero. [6, §4.2.8] [5, §7.1, prob.2] Hence $-\mathcal{D}(X)$ can neither be convex in X.

The outcome is different when instead we consider

$$-V_{\mathcal{N}}^T\mathcal{D}(X)V_{\mathcal{N}} = 2V_{\mathcal{N}}^TX^TXV_{\mathcal{N}} \tag{50}$$

where $V_{\mathcal{N}}$ is defined in (41). (50) meets the criterion for convexity in §2.2 over its domain which is all of $\mathbb{R}^{n\times N}$; *viz.*, for any Y,

$$-\frac{d^2}{dt^2}V_{\mathcal{N}}^T\mathcal{D}(X+t\,Y)V_{\mathcal{N}} \;=\; 4V_{\mathcal{N}}^TY^TYV_{\mathcal{N}} \succeq 0 \tag{51}$$

$-V_{\mathcal{N}}^T\mathcal{D}(X)V_{\mathcal{N}}$ is therefore a convex quadratic function of X that achieves its minimum, with respect to the positive semidefinite cone (§2.1), at $X=0$.

3.5.1 Inner-product form convexity

In §3.3.1 we saw that d_{ij} is a convex quadratic function of $\begin{bmatrix} \sqrt{d_{ik}} \\ \sqrt{d_{kj}} \end{bmatrix}$ and a quasiconvex function of θ_{ikj}. Here the situation for the inner-product form $\mathcal{D}(\Theta)$ (36) of the EDM definition is identical to that in §3.5: $-\mathcal{D}(\Theta)$ is not a quasiconvex function of Θ by the same reasoning, and

$$-V_{\mathcal{N}}^T\mathcal{D}(\Theta)V_{\mathcal{N}} = \Theta^T\Theta \tag{52}$$

is a convex quadratic function of Θ on the domain $\mathbb{R}^{n\times N-1}$ achieving its minimum at $\Theta=0$.

3.6 Injectivity of $\mathcal{D}$

The EDM definitions $\mathcal{D}(X)$ (29) and $\mathcal{D}(\Theta)$ (36) are many-to-one maps (§3.4) to the same range, a convex subset of subspace $\mathbb{S}^N_\delta$ called the *EDM cone* (§8.3);

$$\begin{aligned}\mathbf{EDM}^N &= \left\{\mathcal{D}(X) : \mathbb{R}^{n\times N} \rightarrow \mathbb{S}^N_\delta \mid X \in \mathbb{R}^{n\times N},\quad n\in 1,2,\ldots\right\} \\ &= \left\{\mathcal{D}(\Theta) : \mathbb{R}^{n\times N-1} \rightarrow \mathbb{S}^N_\delta \mid \Theta \in \mathbb{R}^{n\times N-1},\ n\in 1,2,\ldots\right\}\end{aligned} \tag{53}$$

Substituting (52) back into the inner-product form EDM definition (36), EDM $D \triangleq \mathcal{D}(\Theta)$ may be decomposed:

$$\mathcal{D}(D) \triangleq D = \begin{bmatrix} 0 \\ \delta(-V_{\mathcal{N}}^T D V_{\mathcal{N}}) \end{bmatrix} \mathbf{1}^T + \mathbf{1}\begin{bmatrix}0 & \delta^T(-V_{\mathcal{N}}^T D V_{\mathcal{N}})\end{bmatrix} - 2\begin{bmatrix} 0 & \mathbf{0}^T \\ \mathbf{0} & -V_{\mathcal{N}}^T D V_{\mathcal{N}} \end{bmatrix} \tag{54}$$

Defining $\mathcal{D}(D)$ on the left-hand side, we invented a new function. Unlike $\mathcal{D}(X)$ or $\mathcal{D}(\Theta)$, $\mathcal{D}(D)$ is not a definition of EDM; rather, it is an identity predicated upon D being EDM. $\mathcal{D}(D)$ is an injective (one-to-one) map of the EDM cone onto itself. Yet when the domain is instead $\mathbb{S}^N_\delta$, $\mathcal{D}(D)$ becomes an injective map onto that same space $\mathbb{S}^N_\delta$. Proof follows directly from the fact that linear function $\mathcal{D}(D)$ on $\mathbb{S}^N_\delta$ has no nullspace in its domain. [16, §A.1]

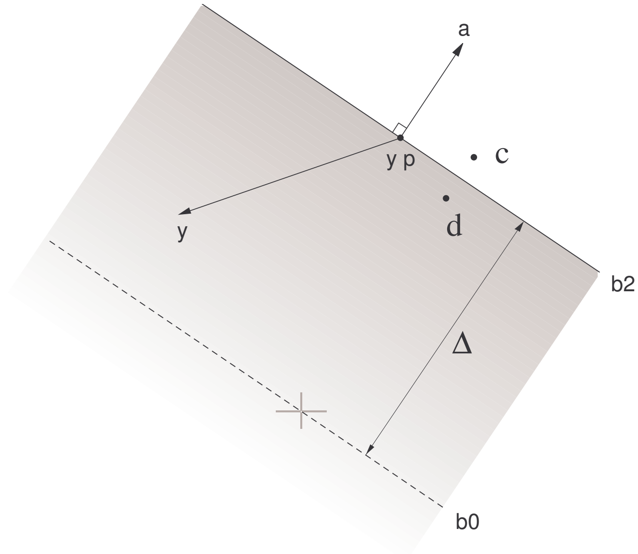

Figure 3: Hyperplane $a^T(y-y_o)=0$ delimiting halfspaces in $\mathbb{R}^2$.

4 Basic convex geometry

4.1 Halfspace, Hyperplane

A two-dimensional affine set in $\mathbb{R}^n$ is called a *plane*. An $n-1$-dimensional affine set in $\mathbb{R}^n$ is called a *hyperplane*. [2] [3] A hyperplane can be a subspace, but not necessarily. Any hyperplane h in $\mathbb{R}^n$ can be described as the solution set to $a^Ty = b$,

$$h(a,b) \;=\; \{y \mid a^Ty = b\} \;=\; \{y \mid a^T(y-y_o) = 0\} \;=\; \{y_o + Z\xi \mid \xi \in \mathbb{R}^{n-1}\} \;\subset\; \mathbb{R}^n \tag{55}$$

assuming the hyperplane has normal $a \in \mathbb{R}^n$ not equal to 0 as in Figure 3.

All solutions y constituting the hyperplane are offset from the nullspace of a^T by the same constant vector y_o which is any particular solution to $a^Ty = b$; *id est*, $y = y_o + Z\xi$ where the columns of $Z \in \mathbb{R}^{n \times n-1}$ hold a basis for the nullspace $\mathcal{N}(a^T)$.

$\mathbb{R}^n$ is divided into two *halfspaces* by any hyperplane. The resulting closed halfspaces may be described

$$\mathcal{H}(a,b)_- \;=\; \{y \mid a^Ty \le b\} \;=\; \{y \mid a^T(y-y_o) \le 0\} \tag{56}$$
$$\mathcal{H}(a,b)_+ \;=\; \{y \mid a^Ty \ge b\} \;=\; \{y \mid a^T(y-y_o) \ge 0\} \tag{57}$$

Visualization is easier if we say $b = a^T y_o$. Then for any vector $y - y_o$ that makes an obtuse angle with normal a, y will lie in the halfspace $\mathcal{H}_-$ on one side of the hyperplane (shaded in Figure 3), while acute angles denote the other side $\mathcal{H}_+$.

Another useful halfspace description comes about when we consider all the points in $\mathbb{R}^n$ closer to c than to d or equidistant, in the Euclidean sense;

$$\{y \mid \|y - c\| \leq \|y - d\|\} \tag{58}$$

This description, in terms of distance, is resolved with the hyperplane description (56) by squaring both sides of the inequality;

$$\left\{y \mid (d-c)^T y \leq \frac{\|d\|^2 - \|c\|^2}{2}\right\} = \left\{y \mid (d-c)^T \left(y - \frac{d+c}{2}\right) \leq 0\right\} \tag{59}$$

Definition. *Supporting hyperplane.* The boundary of a closed halfspace that contains set $\mathcal{C}$ is called a supporting hyperplane $\hbar$ to $\mathcal{C}$ when it contains at least one point of $\mathcal{C}$. [2, §11] For example, given point y_o on the boundary of $\mathcal{C}$ and $a \neq 0$, if

$$\{y \in \mathcal{C} \mid a^T(y - y_o) \leq 0\} = \mathcal{C} \tag{60}$$

then

$$\hbar(a, a^T y_o) = \{y \mid a^T(y - y_o) = 0\} \tag{61}$$

describes a supporting hyperplane to $\mathcal{C}$ at y_o. When the supporting hyperplane contains only one point of $\mathcal{C}$, the hyperplane is *strictly* supporting and termed *tangent* to $\mathcal{C}$. [17, §25/6]

Theorem. *Halfspaces.* [1, §2] [3, §A.4.2(b)] A closed convex set is equivalent to the intersection of all halfspaces that contain it.

4.2 Extreme direction, extreme point, vertex

Need a figure here to show extreme and exposed points...
An extreme point x_ε of a convex set $\mathcal{C}$ is a point belonging to $\mathcal{C}$ that is not expressible as a *convex combination* of points in $\mathcal{C}$ distinct from x_ε; *id est*, for all $x_1, x_2 \in \mathcal{C}$,

$$x_\varepsilon \neq \mu x_1 + (1-\mu) x_2, \ \mu \in [0,1], \text{ for any } x_1, x_2 \neq x_\varepsilon \tag{62}$$

The point of tangency in $\mathcal{C}$ with a strictly supporting hyperplane identifies an extreme point, but not *vice versa.*[13]

Definition. *Ray.* The set

$$\{\zeta a + b \mid \zeta \geq 0,\ a \neq 0\} \tag{63}$$

defines a *halfline* called a ray having base $b \in \mathbb{R}^M$ and direction $a \in \mathbb{R}^M$.

The concept *extreme direction* arises in connection with the geometric object called the *convex cone* (§4.4, §8). Informally, an extreme direction corresponds to an *edge* of a convex cone. An extreme direction is formally represented by a ray emanating from the origin. Rockafellar suggests the mnemonic, "extreme point at infinity". [2, §18]

Theorem. *Extremes.* [2, §18] Any closed convex set containing no lines can be expressed as the convex hull of all its extreme points and directions.

Definition. *Face, exposed face, facet, edge, vertex.* A *face* of a convex set $\mathcal{C}$ is a convex subset $\mathcal{F} \subseteq \mathcal{C}$ such that every closed line segment in $\mathcal{C}$ having relative interior point in $\mathcal{F}$, has both endpoints in $\mathcal{F}$. [2, §18] The zero-dimensional faces of $\mathcal{C}$ are the extreme points. The empty set and $\mathcal{C}$ itself are faces of $\mathcal{C}$.

$\mathcal{F}$ is an *exposed face* of convex set $\mathcal{C}$ iff there is a supporting hyperplane $\hbar$ to $\mathcal{C}$ such that [3, §A.2.4]

$$\mathcal{F} = \mathcal{C} \cap \hbar \tag{64}$$

An *exposed point* is a zero-dimensional exposed face, and the definition of *vertex*; it is an extreme point but not *vice versa.* An edge is a one-dimensional face of a convex set; [3, §A.2.3] it is not necessarily exposed.

A *facet* is an $(n-1)$-dimensional exposed face of convex set $\mathcal{C}$ of affine dimension n. [18] [3, §A.2.3, §A.2.4]

Definition. *Boundary.*

[13] The point of tangency with a strictly supporting hyperplane is the same as an exposed point, which is an extreme point. But an extreme point is not necessarily exposed. [3, §A.2.4]

4.3 Affine dimension, affine hull

The lower bound on Euclidean dimension consistent with an EDM D is called the *embedding* [7] or affine dimension. The affine dimension r is the dimension of the smallest *affine set* in $\mathbb{R}^n$ (empty set, point, line, plane, hyperplane, subspace, $\mathbb{R}^n$) that contains the list (26) in $X \in \mathbb{R}^{n\times N}$; r is the same as the dimension of the subspace parallel to that affine set. [2, §1] That affine set in which the points are embedded is unique and is called the *affine hull* [1, §2];

$$\mathcal{A} \triangleq \operatorname{aff}\{x_\ell\,,\ \ell = 1\ldots N\} \;=\; \operatorname{aff} X \;=\; \{Xa \mid a^T\mathbf{1} = 1\} \;\subseteq\; \mathbb{R}^n \tag{65}$$

4.4 Convex cone, conic hull

A set $\mathcal{K}$ is called a *cone* if ζA belongs to $\mathcal{K}$ whenever $A \in \mathcal{K}$ and $\zeta \geq 0$. We call the set $\mathcal{C}$ a *convex* cone iff

$$A, B \in \mathcal{C} \;\Rightarrow\; \zeta A + \xi B \in \mathcal{C} \;\text{ for all } \zeta, \xi \geq 0 \tag{66}$$

If A lies on the relative boundary of $\mathcal{C}$, then the ray $\{\zeta A \mid \zeta \geq 0\}$ also belongs to the boundary and is called an *extreme direction* [2, §18] of $\mathcal{C}$. A convex cone is a narrower but more familiar class of cone equivalently described as the non-empty intersection of a number of hyperplanes (through the origin) and halfspaces each of whose delimiting hyperplane (Figure 3) passes through the origin.[14] Esoteric examples of convex cones include the point at the origin, any line through the origin, any ray having the origin as base, any subspace, any halfspace delimited by a hyperplane through the origin, and $\mathbb{R}^n$. When a convex cone has a vertex, it resides at the origin.

In terms of a point list contained in the columns of X (26), the *conic hull* is defined:

$$\mathcal{C} \triangleq \textbf{Cone}\{x_\ell\,,\ \ell = 1\ldots N\} \;=\; \{Xa \mid a \succeq 0\} \;\subseteq\; \mathbb{R}^n \tag{67}$$

The conic hull of a set forms a convex cone; the smallest convex cone that contains the set.

---

[14] The number of hyperplanes and halfspaces constituting a convex cone is possibly but not necessarily infinite. When the number is finite, the convex cone is termed polyhedral. The convex cone is a closed set by the definition of halfspace in (56) and (57).

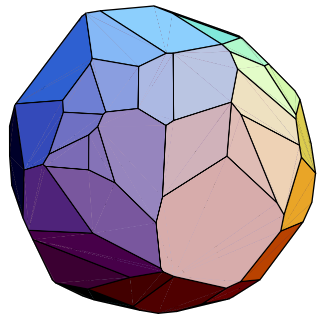

Figure 4: Convex hull of a random list of points in $\mathbb{R}^3$. Some points from the generating list reside in the relative interior of this convex polyhedron. [18, *Convex Polyhedron,Avis-Fukuda*]

4.5 Convex hull

The *convex hull* [1, §2][2] of any finite-length list (or set) of points in Euclidean space $\mathbb{R}^n$ forms a unique *relatively closed* [4][15][14] convex polyhedron whose vertices constitute some subset of that list; *e.g.*, Figure 4,

$$\mathcal{P} \stackrel{\Delta}{=} \text{conv}\{x_\ell\,,\ \ell = 1 \ldots N\} \;=\; \text{conv}\ X \;=\; \{Xa \mid a^T\mathbf{1} = 1,\ a \succeq 0\} \;\subseteq\; \mathbb{R}^n \tag{68}$$

Observe the notation $a \succeq 0$.[16] The *relative boundary* and *relative interior* [1, §2][17][2] of polyhedron $\mathcal{P}$ constitute the convex hull $\mathcal{P}$; the smallest *convex set* (§2.1) that contains the list in X. Given $\mathcal{P}$, the *generating list* $\{x_\ell\}$ is not unique.

[15] When a set $\mathcal{C}$ is relatively closed, it means closed relative to the affine hull of $\mathcal{C}$.

[16] For symmetric matrices, the notation $A \succeq B$ denotes comparison on the positive semidefinite cone (§2.1). For vectors, $a \succeq b$ denotes comparison on the nonnegative orthant while $\geq$ is reserved for scalar comparison as in $a^Ty \geq b$.

[17] The relative interior of a set $\mathcal{C}$ is the interior relative to the affine hull of $\mathcal{C}$. (Likewise for the relative boundary.)

4.6 Convex polyhedron

We define convex polyhedra to be of any dimension and to comprise all affine sets, *polyhedral cones*,[18] line segments, rays, and halfspaces. Not all convex polyhedra are bounded, hence neither can they all be described by a convex hull as in (68). Boyd and Vandenberghe [1, §2][19] propose a universal description of polyhedra in terms of that same generating list (26) in $X \in \mathbb{R}^{n\times N}$, which encompasses all three classes of convex object:

$$\{Xa \mid a_{1:m}^T \mathbf{1} = 1,\ a_{1:m} \succeq 0,\ 1 \leq m \leq N\} \tag{69}$$

$a_{1:m}$ denoting the truncated a-vector,

$$a_{1:m} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} \tag{70}$$

From (65), (67), and (68), we deduce the conditions that may be applied to the coefficients left unspecified in (69);

$$\left.\begin{array}{lcl} \text{affine sets} & \longrightarrow & a^T\mathbf{1} = 1 \\ \text{convex cones} & \longrightarrow & a \succeq 0 \end{array}\right\} \longleftarrow \text{bounded polyhedra} \tag{71}$$

There is an equivalent halfspace/hyperplane description of convex polyhedra: A polyhedron is the intersection of a *finite* number of halfspaces and hyperplanes; [3]

$$\{y \mid Ay \preceq b,\ Cy = d\} \subseteq \mathbb{R}^n \tag{72}$$

where A and C generally denote matrices. Each row of A is a vector normal to the hyperplane delimiting one halfspace, while each row of C is a vector normal to a hyperplane. A polyhedron thus described must be a convex set.[20](§2.1) When b and d in (72) are zero, the result is a polyhedral cone. Conversion between the halfspace/hyperplane (72) and convex hull descriptions (68)(69) is nontrivial, in general.[21] [19]

[18] In Rockafellar's terminology, *finitely-generated convex cones*; [2, §19] *id est*, all sets that can be described, using matrix coefficients, $\{y \mid Ay \preceq 0,\ Cy = 0\}$. (*confer* (72))

[19] Their definition is designed to enhance Rockafellar's [2, §19] by facilitating intersection with affine sets.

[20] We consider only convex polyhedra, but acknowledge the existence of concave polyhedra. [18, *Kepler-Poinsot Solid*]

[21] The conversion is easy for simplices (§5.3.2). [1, §2]

A vertex of a convex polyhedron is the same as a zero-dimensional exposed face.[22] (§4.2) [3, §A.2.4] Hence a vertex always resides on the relative boundary of a convex polyhedron. [17, §25/6] In Figure 4 the vertices are located at the intersection of three or more facets. Not all members of a generating list become vertices of the corresponding polyhedron; certainly true for (68) and (69), some list members reside in the polyhedron's relative interior. Conversely, when (68) applies, the convex hull of the vertices is a polyhedron identical to the convex hull of the generating list.

4.7 Embedding in the affine hull

The affine hull (65) of the x_ℓ in X is identical to the affine hull of the polyhedron $\mathcal{P}$ formed from all convex combinations of the x_ℓ as in (68); [1, §2][2, §17]

$$\mathcal{A} = \operatorname{aff} X = \operatorname{aff} \mathcal{P} \tag{73}$$

Comparing definitions (65) and (68), it becomes obvious that the x_ℓ and their convex hull $\mathcal{P}$ are embedded in their unique affine hull $\mathcal{A}$;

$$\mathcal{A} \supseteq \mathcal{P} \supseteq \{x_\ell\} \tag{74}$$

We define the dimension r of the affine hull $\mathcal{A}$ to be the same as the dimension of the convex hull $\mathcal{P}$ [2, §2], but r is not necessarily equal to the rank of X.[23] For the particular example illustrated in Figure 1, $\mathcal{P}$ is the triangle plus its relative interior while its three vertices constitute the entire list in X. The affine hull $\mathcal{A}$ is the unique plane that contains the triangle, so $r=2$ in that example while the rank of X is 3. Were there only two points in Figure 1, then the affine hull would instead be the unique line passing through them; r would become 1 while the rank would be 2.

[22]For the polyhedron in $\mathbb{R}^3$ from Figure 4, the edges are one-dimensional exposed faces while the two-dimensional exposed faces are the facets.

[23] $\operatorname{rank} X \leq \min\{n, N\}$, $r \leq \min\{n, N-1\}$ (84).

4.8 Determining affine dimension

Affine dimension r is important because we lose any absolute offset component common to all the generating x_ℓ in $\mathbb{R}^n$ when reconstructing convex polyhedra given only distance information. (§3.4) To calculate r, we first eliminate any offset that serves to increase dimensionality of the subspace required to contain $\mathcal{P}$; subtracting $\alpha \in \mathcal{A}$ from every list member will work,

$$X - \alpha\mathbf{1}^T \tag{75}$$

translating $\mathcal{A}$ to the origin: [24]

$$\mathcal{A} - \alpha \;=\; \operatorname{aff}(X - \alpha\mathbf{1}^T) \;=\; \operatorname{aff}(X) - \alpha \tag{76}$$

$$\mathcal{P} - \alpha \;=\; \operatorname{conv}(X - \alpha\mathbf{1}^T) \;=\; \operatorname{conv}(X) - \alpha \tag{77}$$

which follow from their definitions. Because (73) and (74) translate,

$$\mathbb{R}^n \;\supseteq\; \mathcal{A} - \alpha \;=\; \operatorname{aff}(X - \alpha\mathbf{1}^T) \;=\; \operatorname{aff}(\mathcal{P} - \alpha) \;\supseteq\; \mathcal{P} - \alpha \;\supseteq\; \{x_\ell - \alpha\} \tag{78}$$

where from the previous relations it is easily shown

$$\operatorname{aff}(\mathcal{P} - \alpha) = \operatorname{aff}(\mathcal{P}) - \alpha \tag{79}$$

Translating $\mathcal{A}$ neither changes its dimension nor the dimension of the embedded polyhedron $\mathcal{P}$;

$$r \;\triangleq\; \dim\mathcal{A} \;=\; \dim(\mathcal{A} - \alpha) \;\triangleq\; \dim(\mathcal{P} - \alpha) \;=\; \dim\mathcal{P} \tag{80}$$

For any $\alpha \in \mathbb{R}^n$, (76)-(80) remain true. [2, pg.4, pg.12] Yet when $\alpha \in \mathcal{A}$, the affine set $\mathcal{A} - \alpha$ becomes a unique subspace of $\mathbb{R}^n$ in which the $\{x_\ell - \alpha\}$ and their convex hull $\mathcal{P} - \alpha$ are embedded (78), and whose dimension is more easily calculated.

---

[24] We might choose the *geometric center* of the x_ℓ; [7][20]

$$\alpha = \alpha_g = X b_g = \frac{1}{N} X\mathbf{1} \in \mathcal{P} \subseteq \mathcal{A}$$

If we were to associate a point-mass m_ℓ with each of the points x_ℓ in a list, then their *center of mass* (or *gravity*) would be $(\sum x_\ell\, m_\ell)/\sum m_\ell$. The geometric center is the same as the center of mass under the assumption of uniform mass density across points. [10] The geometric center always lies in the convex hull; (68) *id est*, $\alpha_g \in \mathcal{P}$ because $b_g^T\mathbf{1} = 1$ and $b_g \succeq 0$.

Example. *Translating first list-member to origin.* Subtracting $\alpha = x_1$ ($x_1 \in \mathcal{P} \subseteq \mathcal{A}$) from every list member will translate $\mathcal{A}$ and $\mathcal{P}$ and, in particular, x_1 to the origin in $\mathbb{R}^n$; *viz.*,

$$X - x_1\mathbf{1}^T \;=\; X - Xe_1\mathbf{1}^T \;=\; X(I - e_1\mathbf{1}^T) \;=\; X\left[\begin{array}{cc}\mathbf{0} & \sqrt{2}V_{\mathcal{N}}\end{array}\right] \tag{81}$$

where $V_{\mathcal{N}}$ is defined in (41), and e_1 in (43). Applying (78) to (81),

$$\mathbb{R}^n \supseteq \mathcal{R}(XV_{\mathcal{N}}) = \mathcal{A} - x_1 = \operatorname{aff}(X - x_1\mathbf{1}^T) = \operatorname{aff}(\mathcal{P} - x_1) \supseteq \mathcal{P} - x_1 \ni \mathbf{0} \tag{82}$$

where $XV_{\mathcal{N}} \in \mathbb{R}^{n\times N-1}$. Hence

$$r = \dim \mathcal{R}(XV_{\mathcal{N}}) \tag{83}$$

The general fact

$$r \le \min\{n, N-1\} \tag{84}$$

is evident from (81) but can be visualized in the example illustrated in Figure 1. There we imagine a vector from the origin to each point in the list. Those three vectors are linearly independent in $\mathbb{R}^3$, but the affine dimension r is 2 because the three points lie in a plane. When that plane is translated to the origin, it becomes the only subspace of dimension $r = 2$ which can contain the translated triangular polyhedron.

4.9 Affine dimension *vs.* rank

Now, suppose D is an EDM as in (29) and we pre-multiply by $-V_{\mathcal{N}}^T$ and post-multiply by $V_{\mathcal{N}}$. Then because $V_{\mathcal{N}}^T\mathbf{1} = 0$ (42),

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} \;=\; 2V_{\mathcal{N}}^T X^T X V_{\mathcal{N}} \in \mathbb{R}^{N-1\times N-1} \tag{85}$$

Consequent to (39) for any α,

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} \;=\; 2V_{\mathcal{N}}^T (X - \alpha\mathbf{1}^T)^T (X - \alpha\mathbf{1}^T) V_{\mathcal{N}} \tag{86}$$

Similarly multiplying the inner-product form EDM definition (36),

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} \;=\; \Theta^T\Theta \in \mathbb{R}^{N-1\times N-1} \tag{87}$$

Theorem. *Rank, nullspace, and number of* 0 *eigenvalues.* For any $A\in\mathbb{R}^{m\times p}$,

$$\operatorname{rank}(A) + \dim\mathcal{N}(A) = p \tag{88}$$

by *conservation of dimension.* [5, §0.4] For any square matrix ($p = m$), the number of 0 eigenvalues is at least equal to $\dim\mathcal{N}(A)$. For any diagonalizable matrix ($p = m$), the number of 0 eigenvalues is exactly equal to $\dim\mathcal{N}(A)$.

For any matrix A, $\operatorname{rank} A^TA = \operatorname{rank} A$. [5, §0.4][25] Hence

$$\operatorname{rank} V_{\mathcal{N}}^T D V_{\mathcal{N}} \;=\; \operatorname{rank} X V_{\mathcal{N}} \;=\; \operatorname{rank}\Theta \;=\; r \tag{89}$$

By conservation of dimension,

$$r + \dim\mathcal{N}(V_{\mathcal{N}}^T D V_{\mathcal{N}}) \;=\; r + \dim\mathcal{N}(\Theta) \;=\; N-1 \tag{90}$$

We summarize the affine dimension:

$$\begin{array}{rcl} r & \stackrel{\Delta}{=} & \dim(\mathcal{P}-\alpha) \;=\; \dim\mathcal{P} \;=\; \dim\operatorname{conv} X \\ & = & \dim(\mathcal{A}-\alpha) \;=\; \dim\mathcal{A} \;=\; \dim\operatorname{aff} X \\ & = & \operatorname{rank}(X - x_1\mathbf{1}^T) \\ & = & \operatorname{rank} X V_{\mathcal{N}} \\ & = & \left.\begin{array}{l} \operatorname{rank} V_{\mathcal{N}}^T D V_{\mathcal{N}} \\ \operatorname{rank}\Theta \\ \operatorname{rank}\Lambda\,,\ (125) \end{array}\right\} D \in \mathbf{EDM}^N \\ & \leq & \min\{n, N-1\} \end{array} \tag{91}$$

[25] For $A\in\mathbb{R}^{m\times p}$, $\mathcal{N}(A^TA) = \mathcal{N}(A)$. [8, §3.3]

5 Metric *vs.* matrix criteria

5.1 Nonnegativity axiom 1

When D is an EDM (29), then it is apparent from (85) that

$$2V_{\mathcal{N}}^T X^T X V_{\mathcal{N}} \;=\; -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \tag{92}$$

because for any A, $A^T A \succeq 0$.[26] Kreyszig notes that axioms 2 through 4 (§3.1) together imply nonnegativity of the d_{ij}. [13, §1.1, prob.15] We claim that nonnegativity is enforced primarily by the matrix inequality (92); *id est*,

$$\left.\begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \\ \delta(D) = 0 \\ D^T = D \end{array}\right\} \Rightarrow\; d_{ij} \geq 0,\; i \neq j \tag{93}$$

(The matrix criterion to enforce strict positivity differs by a stroke of the pen. (96))

We now support our claim: If any matrix $A \in \mathbb{R}^{m\times m}$ is positive semidefinite, then its main diagonal $\delta(A) \in \mathbb{R}^m$ must have all nonnegative entries. [6, §4.2] Given $\delta(D) = 0$ and $D^T = D$,

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} =$$

$$\begin{bmatrix} d_{12} & \frac{1}{2}(d_{12}+d_{13}-d_{23}) & \frac{1}{2}(d_{1,i+1}+d_{1,j+1}-d_{i+1,j+1}) & \cdots & \frac{1}{2}(d_{12}+d_{1N}-d_{2N}) \\ \frac{1}{2}(d_{12}+d_{13}-d_{23}) & d_{13} & \frac{1}{2}(d_{1,i+1}+d_{1,j+1}-d_{i+1,j+1}) & \cdots & \frac{1}{2}(d_{13}+d_{1N}-d_{3N}) \\ \frac{1}{2}(d_{1,j+1}+d_{1,i+1}-d_{j+1,i+1}) & \frac{1}{2}(d_{1,j+1}+d_{1,i+1}-d_{j+1,i+1}) & d_{1,i+1} & \ddots & \frac{1}{2}(d_{14}+d_{1N}-d_{4N}) \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{1}{2}(d_{12}+d_{1N}-d_{2N}) & \frac{1}{2}(d_{13}+d_{1N}-d_{3N}) & \frac{1}{2}(d_{14}+d_{1N}-d_{4N}) & \cdots & d_{1N} \end{bmatrix}$$

$$\in \mathbb{R}^{N-1\times N-1} \tag{94}$$

where row,column indices $i,j \in 1 \ldots N-1$. [21] It follows that

$$\left.\begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \\ \delta(D) = 0 \\ D^T = D \end{array}\right\} \Rightarrow\; \delta(-V_{\mathcal{N}}^T D V_{\mathcal{N}}) = \begin{bmatrix} d_{12} \\ d_{13} \\ \vdots \\ d_{1N} \end{bmatrix} \succeq 0 \tag{95}$$

[26] For $A \in \mathbb{R}^{m\times n}$, $A^T A \succeq 0 \;\Leftrightarrow\; y^T A^T A y \;=\; \|Ay\|^2 \;\geq\; 0$ for all $y \in \mathbb{R}^n$. When A is full-rank skinny or square, $A^T A \succ 0$.

Multiplication of $V_{\mathcal{N}}$ by any permutation matrix Ξ has null effect on its range and nullspace. In other words, any permutation of the rows or columns of $V_{\mathcal{N}}$ produces a basis for $\mathcal{N}(\mathbf{1}^T)$; *id est*, $\mathcal{R}(\Xi_r V_{\mathcal{N}}) = \mathcal{R}(V_{\mathcal{N}}\Xi_c) = \mathcal{R}(V_{\mathcal{N}}) = \mathcal{N}(\mathbf{1}^T)$. Hence, $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \Leftrightarrow -V_{\mathcal{N}}^T \Xi_r^T D \Xi_r V_{\mathcal{N}} \succeq 0$ $(\Leftrightarrow -\Xi_c^T V_{\mathcal{N}}^T D V_{\mathcal{N}} \Xi_c \succeq 0)$. Various permutation matrices[27] will sift the remaining d_{ij} similarly to (95) thereby proving their nonnegativity. Hence $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$ is a sufficient test of the first axiom (§3.1) of Euclidean space, nonnegativity. ♦

5.1.1 Strict positivity

Should we require the points in $\mathbb{R}^n$ to be distinct, then entries of D off the main diagonal must be *strictly* positive $\{d_{ij} > 0,\ i \neq j\}$, and only those entries along the main diagonal of D are zero. By similar argument, the strict matrix inequality is a sufficient test of strict positivity of the Euclidean metric-squared;

$$\left.\begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succ 0 \\ \delta(D) = 0 \\ D^T = D \end{array}\right\} \Rightarrow\ d_{ij} > 0,\ i \neq j \tag{96}$$

[27]The rule of thumb is: If $\Xi_r(i,1) = 1$, then $\delta(-V_{\mathcal{N}}^T \Xi_r^T D \Xi_r V_{\mathcal{N}}) \in \mathbb{R}^{N-1}$ is some permutation of the i^{th} row or column of D excepting the 0 entry from the main diagonal.

5.2 Triangle inequality axiom 4

In light of Kreyszig's observation [13, §1.1, prob.15] that metric space axioms 2 through 4 (§3.1) imply axiom 1, the nonnegativity criterion (93) suggests that the matrix inequality $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$ might somehow take on the role of the triangle inequality; *id est*,

$$\left.\begin{array}{r} \delta(D) = 0 \\ D^T = D \\ -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \end{array}\right\} \Rightarrow \sqrt{d_{ij}} \;\leq\; \sqrt{d_{ik}} + \sqrt{d_{kj}}\;,\;\; i \neq j \neq k \tag{97}$$

We now show that is indeed the case: Let T be the *leading principal submatrix* in $\mathbb{R}^{2\times 2}$ of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ (upper left 2×2 submatrix from (94));

$$T \triangleq \left[\begin{array}{cc} d_{12} & \frac{1}{2}(d_{12}+d_{13}-d_{23}) \\ \frac{1}{2}(d_{12}+d_{13}-d_{23}) & d_{13} \end{array}\right] \tag{98}$$

A bit obscure from standard texts [8, §6.3] [5] [6] but true that T must be positive (semi)definite whenever $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ is. More generally,

Theorem. *Principal submatrix.* (App. F, *confer* (14))

- $A \in \mathbb{S}^M$ is positive definite if and only if all the principal submatrices of dimension less than M are positive definite and $\det A$ is positive.
- $A \in \mathbb{S}^M$ is positive semidefinite if and only if all principal submatrices of dimension less than M are positive semidefinite and $\det A$ is nonnegative.

Now we have,

$$\begin{array}{c} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \;\Rightarrow\; T \succeq 0 \;\Leftrightarrow\; \sigma_1 \geq \sigma_2 \geq 0 \\ -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succ 0 \;\Rightarrow\; T \succ 0 \;\Leftrightarrow\; \sigma_1 > \sigma_2 > 0 \end{array} \tag{99}$$

where σ_1 and σ_2 are the eigenvalues of T, real due only to symmetry of T:

$$\begin{array}{l} \sigma_1 = \frac{1}{2}\left(d_{12} + d_{13} + \sqrt{d_{23}^2 - 2(d_{12}+d_{13})d_{23} + 2(d_{12}^2 + d_{13}^2)}\right) \in \mathbb{R} \\ \sigma_2 = \frac{1}{2}\left(d_{12} + d_{13} - \sqrt{d_{23}^2 - 2(d_{12}+d_{13})d_{23} + 2(d_{12}^2 + d_{13}^2)}\right) \in \mathbb{R} \end{array} \tag{100}$$

Nonnegativity of eigenvalue σ_1 is guaranteed by only the nonnegativity of the d_{ij} which in turn is guaranteed by the matrix inequality (93). The inequality between the eigenvalues in (99) follows from only the realness of the d_{ij}. Since σ_1 always exceeds or equals σ_2, conditions for the positive (semi)definiteness of submatrix T can be completely determined by examining σ_2, the smaller of its two eigenvalues. A triangle inequality is made apparent when we express T eigenvalue nonnegativity in terms of D matrix entries; *viz.*,

$$\begin{gathered} T \succeq 0 \quad\Leftrightarrow\quad \det T = \sigma_1\sigma_2 \geq 0\,, \quad d_{12}\,, d_{13} \geq 0 \\ \Leftrightarrow \\ d_{ij} \in \mathbb{R}, \quad \sigma_2 \geq 0 \\ \Leftrightarrow \\ d_{12}\,, d_{13}\,, d_{23} \geq 0 \\ \text{and} \\ |\sqrt{d_{12}} - \sqrt{d_{23}}\,| \;\leq\; \sqrt{d_{13}} \;\leq\; \sqrt{d_{12}} + \sqrt{d_{23}} \qquad \text{(a)} \end{gathered} \tag{101}$$

Triangle inequality (101a) (*confer* (105)), in terms of three entries from D, is equivalent to §3.1 axiom 4

$$\begin{aligned} \sqrt{d_{13}} &\leq \sqrt{d_{12}} + \sqrt{d_{23}} \\ \sqrt{d_{23}} &\leq \sqrt{d_{12}} + \sqrt{d_{13}} \\ \sqrt{d_{12}} &\leq \sqrt{d_{13}} + \sqrt{d_{23}} \end{aligned} \tag{102}$$

for the corresponding points x_1, x_2, x_3 from some length-N list.[28]

[28] Accounting for symmetry axiom 3 (§3.1), axiom 4 demands three inequalities be satisfied per one of type (101a). The first of those inequalities in (102) is self evident from (101a), while the two remaining follow from the left-hand side of (101a) and the fact for scalars, $|a| \leq b \Leftrightarrow a \leq b$ and $-a \leq b$.

5.2.1 Comment

Given D whose dimension N is greater than or equal to 3, in total there are $N!/(3!(N-3)!)$ distinct triangle inequalities like (35) that must be satisfied, of which each d_{ij} is involved in $N-2$, and each point x_i is in $(N-1)(N-2)/2$. We have so far revealed only one of those triangle inequalities, (101a) which came from T (98). Yet we claim if $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$ then all triangle inequalities will be satisfied simultaneously;

$$|\sqrt{d_{ik}} - \sqrt{d_{kj}}\,| \;\leq\; \sqrt{d_{ij}} \;\leq\; \sqrt{d_{ik}} + \sqrt{d_{kj}}\,, \quad i<k<j \tag{103}$$

(There are no more.) To verify our claim, we must show that the matrix inequality $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$ is a sufficient test of all the triangle inequalities; more efficient, we add, for larger N.

5.2.2 Shore

The columns of $\Xi_r V_{\mathcal{N}} \Xi_c$ hold a basis for $\mathcal{N}(\mathbf{1}^T)$ when Ξ_r and Ξ_c are permutation matrices. In other words, any permutation of the rows or columns of $V_{\mathcal{N}}$ leaves its range and nullspace unchanged; *id est*, $\mathcal{R}(\Xi_r V_{\mathcal{N}} \Xi_c) = \mathcal{R}(V_{\mathcal{N}}) = \mathcal{N}(\mathbf{1}^T)$ (42). Hence, two distinct matrix inequalities can be equivalent tests of the positive semidefiniteness of D on $\mathcal{R}(V_{\mathcal{N}})$; *id est*, $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \Leftrightarrow -(\Xi_r V_{\mathcal{N}} \Xi_c)^T D (\Xi_r V_{\mathcal{N}} \Xi_c) \succeq 0$. By properly choosing the permutation matrices,[29] the leading principal submatrix $T_\Xi \in \mathbb{R}^{2\times 2}$ of $-(\Xi_r V_{\mathcal{N}} \Xi_c)^T D (\Xi_r V_{\mathcal{N}} \Xi_c)$ may be loaded with the entries of D needed to test any particular triangle inequality (similarly to (94)-(101)). Because all the triangle inequalities can be individually tested using a test equivalent to the lone matrix inequality $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$, it logically follows that the lone matrix inequality tests all those triangle inequalities simultaneously. We conclude that $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$ is a sufficient test of the fourth axiom (§3.1) of Euclidean space, triangle inequality. ♦

[29] To individually test triangle inequality $|\sqrt{d_{ik}} - \sqrt{d_{kj}}\,| \leq \sqrt{d_{ij}} \leq \sqrt{d_{ik}} + \sqrt{d_{kj}}$ for particular i,k,j, set $\Xi_r(i,1) = \Xi_r(k,2) = \Xi_r(j,3) = 1$, and $\Xi_c = I$.

5.2.3 Strict triangle inequality

Without exception, all the inequalities in (101) and (102) can be made strict while their equivalences remain true. The then strict inequality (101a) or (102) may be interpreted as a *strict triangle inequality* under which collinear arrangement of points is not allowed. [17, §24/6] Hence by similar reasoning, $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succ 0$ is a sufficient test of all the strict triangle inequalities; *id est*,

$$\left.\begin{array}{r}\delta(D)=0\\ D^T=D\\ -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succ 0\end{array}\right\} \Rightarrow \sqrt{d_{ij}} \;<\; \sqrt{d_{ik}}+\sqrt{d_{kj}}\,,\;\; i\neq j\neq k \tag{104}$$

5.2.4 Affine dimension reduction in two dimensions

The leading principal 2×2 submatrix T of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ has largest eigenvalue σ_1 (100) which is a convex function of D.[30] σ_1 can never be zero unless $d_{12}=d_{13}=d_{23}=0$. σ_1 can never be negative so long as d_{ij} are nonnegative. The remaining eigenvalue σ_2 is a concave function of D that becomes zero only at the upper and lower bounds of inequality (101a) and its equivalent forms: (*confer* (103))

$$\begin{array}{rcccl}
|\sqrt{d_{12}}-\sqrt{d_{23}}\,| & \leq & \sqrt{d_{13}} & \leq & \sqrt{d_{12}}+\sqrt{d_{23}} \qquad \text{(a)}\\
 & & \Leftrightarrow & & \\
|\sqrt{d_{12}}-\sqrt{d_{13}}\,| & \leq & \sqrt{d_{23}} & \leq & \sqrt{d_{12}}+\sqrt{d_{13}} \qquad \text{(b)}\\
 & & \Leftrightarrow & & \\
|\sqrt{d_{13}}-\sqrt{d_{23}}\,| & \leq & \sqrt{d_{12}} & \leq & \sqrt{d_{13}}+\sqrt{d_{23}} \qquad \text{(c)}
\end{array} \tag{105}$$

In between those bounds, σ_2 is strictly positive; otherwise, it would be negative but prevented by the condition $T \succeq 0$.

When σ_2 becomes zero, it means that triangle $\triangle_{123}$ has collapsed to a line segment; a potential reduction in affine dimension r. The same logic is valid for any principal 2×2 submatrix of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$, hence applicable to other triangles.

[30] The maximum eigenvalue of any symmetric matrix is always a convex function of its entries, while the minimum eigenvalue is always concave. [1, §3] In our particular case, say $\underline{d} \triangleq \begin{bmatrix} d_{12}\\ d_{13}\\ d_{23}\end{bmatrix} \in \mathbb{R}^3$. Then the Hessian $\nabla^2\sigma_1(\underline{d}) \succeq 0$ certifies convexity whereas $\nabla^2\sigma_2(\underline{d}) \preceq 0$ certifies concavity. Each Hessian has rank equal to 1. The respective gradients $\nabla\sigma_1(\underline{d})$ and $\nabla\sigma_2(\underline{d})$ are nowhere zero.

5.3 Bridge: Convex polyhedra to EDMs

The criteria for the existence of an EDM include, by definition (29) (36), the axioms imposed upon its entries d_{ij} by a Euclidean space. From §5.1 and §5.2, we know there is a relationship of matrix criteria to those axioms. Here is a snapshot of what we are sure thus far: for $i,j \in 1 \ldots N$, (*confer* §3.1)

$$\begin{array}{l} \sqrt{d_{ij}} \geq 0,\ i \neq j \\ \sqrt{d_{ij}} = 0,\ i = j \\ \sqrt{d_{ij}} = \sqrt{d_{ji}} \\ \sqrt{d_{ij}}\ \leq\ \sqrt{d_{ik}} + \sqrt{d_{kj}}\,,\ i \neq j \neq k \end{array} \quad \Leftarrow \quad \begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \\ \delta(D) = 0 \\ D^T = D \end{array} \quad \Leftarrow \quad D \in \mathbf{EDM}^N \tag{106}$$

At this moment, we have no converse for (106). As of concern in §3.2, we have yet to establish the metric requirements beyond the four Euclidean axioms which allow D to be identified as an EDM, or which facilitate polyhedron or list reconstruction from an incomplete EDM. Our present goal is to establish the necessary and sufficient matrix criteria that will subsume all the Euclidean axioms and any further requirements[31] for all $N > 1$; *id est*,

$$\begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \\ D \in \mathbb{S}_\delta^N \end{array} \quad \Leftrightarrow \quad D \in \mathbf{EDM}^N \tag{107}$$

From (42) $\mathcal{R}(V_{\mathcal{N}}) = \mathcal{N}(\mathbf{1}^T)$, so (107) is the same as

$$\begin{array}{r} -z^T D z \geq 0 \\ \mathbf{1}^T z = 0 \\ D \in \mathbb{S}_\delta^N \end{array} \quad \Leftrightarrow \quad D \in \mathbf{EDM}^N \tag{108}$$

[31]In 1935, Schoenberg [21] first extolled expansion (94) showing that nonnegativity of $-y^T V_{\mathcal{N}}^T D V_{\mathcal{N}}\, y \geq 0$, for all $y \in \mathbb{R}^{N-1}$, is necessary and sufficient for D to be an EDM. (Symmetry and zero self-distance predicate (94).)

5.3.1 Geometric criterion

We derive matrix criteria for D to be an EDM, validating (107) using simple geometry: distance to the polyhedron formed by the convex hull of a list of points in Euclidean space $\mathbb{R}^n$.

EDM assertion. D is a Euclidean distance matrix if and only if $D\in\mathbb{S}_\delta^N$ and distances-squared from the origin

$$\{\|p(y)\|^2 = -y^T V_{\mathcal{N}}^T D V_{\mathcal{N}}\, y \mid y \in \mathcal{S} - \beta\} \tag{109}$$

correspond to points p in some relatively closed convex polyhedron

$$\mathcal{P} - \alpha \;=\; \{p(y) \mid y \in \mathcal{S} - \beta\} \tag{110}$$

having N or fewer vertices embedded in an r-dimensional subspace $\mathcal{A} - \alpha$ of $\mathbb{R}^n$, where $\alpha \in \mathcal{A} = \operatorname{aff}\mathcal{P}$, and where the domain of linear function $p(y)$ is the *unit simplex* $\mathcal{S}$ shifted such that its vertex at the origin is translated to $-\beta$ in $\mathbb{R}^{N-1}$. When $\beta = 0$, $\alpha = x_1$.

5.3.2 Unit simplex

The unit simplex, a peculiar convex subset of the *nonnegative orthant*,

$$\mathcal{S} \triangleq \{s \mid s \succeq 0,\ \mathbf{1}^T s \leq 1\} \subset \mathbb{R}_+^{N-1} \tag{111}$$

is itself a regular convex polyhedron having nonempty interior, N vertices, and dimension $N-1$. [1, §2]

$$\dim \mathcal{S} \;=\; N-1 \tag{112}$$

The origin supplies one vertex, while heads of the *standard basis* [5] [8] $\{e_i\,,\ i=1\ldots N-1\}$ in $\mathbb{R}^{N-1}$ constitute those remaining;[32]

$$\mathcal{S} \;=\; \operatorname{conv}\{0,\ \{e_i\,,\ i=1\ldots N-1\}\} \tag{113}$$

In terms of $V_\mathcal{N}$, the unit simplex can be represented equivalently;

$$\mathcal{S} \;=\; \{s \in \mathbb{R}^{N-1} \mid \sqrt{2}V_\mathcal{N}\, s \succeq -e_1\} \tag{114}$$

where e_1 is as in (43).

Incidental to the EDM assertion, shifting the simplex domain in $\mathbb{R}^{N-1}$ translates the polyhedron $\mathcal{P}$ in $\mathbb{R}^n$. Indeed, there is a one-to-one correspondence between vertices of the unit simplex and members of the list generating $\mathcal{P}$;

$$\begin{array}{ccccc} p & : & \mathbb{R}^{N-1} & \rightarrow & \mathbb{R}^n \\ & & & & \\ p\left(\left\{\begin{array}{c} -\beta \\ e_1-\beta \\ e_2-\beta \\ \vdots \\ e_{N-1}-\beta \end{array}\right\}\right) & & & = & \left\{\begin{array}{c} x_1-\alpha \\ x_2-\alpha \\ x_3-\alpha \\ \vdots \\ x_N-\alpha \end{array}\right\} \end{array} \tag{115}$$

[32] In $\mathbb{R}^0$ the unit simplex is the point at the origin, in $\mathbb{R}$ the unit simplex is the line segment $[0,1]$, in $\mathbb{R}^2$ it is a triangle and its relative interior, in $\mathbb{R}^3$ it is the convex hull of a tetrahedron [22], in $\mathbb{R}^4$ it is the convex hull of a pentatope. [18] The *unit* simplex is a special case of the general class of polyhedra called *simplex*: [2]

$$\left\{\operatorname{conv}\{v_i\} \mid \dim \operatorname{aff}\{v_i\,,\ i=0\ldots k\} = k,\ v_i \in \mathbb{R}^{N-1},\ N-1 \geq k\right\}$$

5.3.3 Proof of EDM assertion

$\Longrightarrow$ To validate the EDM assertion (§5.3.1) in the forward direction, we must demonstrate that if D is an EDM then each distance-squared $\|p(y)\|^2$ described by (109) corresponds to a point p in some embedded polyhedron $\mathcal{P}-\alpha$. Assume that D is indeed an EDM; *id est*, D comes from a list of N unknown points in Euclidean space $\mathbb{R}^n$; $D=\mathcal{D}(X)$ for $X\in\mathbb{R}^{n\times N}$ as in (29). Since D is offset invariant (§3.4), we may shift the affine hull $\mathcal{A}$ of those unknown points to the origin as in (75). Then take any point p in their convex hull; (68)

$$\mathcal{P}-\alpha \;=\; \{p=(X-Xb\,\mathbf{1}^T)a \mid a^T\mathbf{1}=1,\; a\succeq 0\} \tag{116}$$

where $\alpha=Xb\in\mathcal{A} \;\Leftrightarrow\; b^T\mathbf{1}=1$. Solutions to $a^T\mathbf{1}=1$ are:[33]

$$a=e_1+\sqrt{2}V_{\mathcal{N}}\,s \tag{117}$$

where $s\in\mathbb{R}^{N-1}$ and e_1 is as in (43). Similarly, $b=e_1+\sqrt{2}V_{\mathcal{N}}\beta$.

$$\begin{aligned}\mathcal{P}-\alpha &= \{p=X(I-(e_1+\sqrt{2}V_{\mathcal{N}}\beta)\mathbf{1}^T)(e_1+\sqrt{2}V_{\mathcal{N}}\,s)\mid \sqrt{2}V_{\mathcal{N}}\,s\succeq -e_1\}\\ &= \{p=X\sqrt{2}V_{\mathcal{N}}(s-\beta)\mid \sqrt{2}V_{\mathcal{N}}\,s\succeq -e_1\}\end{aligned} \tag{118}$$

which describes the domain of $p(s)$ as the unit simplex $\mathcal{S}$; [22][1, §2]

$$\mathcal{S}=\{s\mid \sqrt{2}V_{\mathcal{N}}\,s\succeq -e_1\}\subset\mathbb{R}_+^{N-1} \tag{119}$$

Making the substitution $y\leftarrow s-\beta$,

$$\mathcal{P}-\alpha \;=\; \{p=\sqrt{2}XV_{\mathcal{N}}\,y \mid y\in\mathcal{S}-\beta\} \tag{120}$$

Point p belongs to a convex polyhedron $\mathcal{P}-\alpha$ embedded in an r-dimensional subspace of $\mathbb{R}^n$ because the convex hull of any list forms a polyhedron, [1, §2][2] and because the translated affine hull $\mathcal{A}-\alpha$ contains the translated polyhedron $\mathcal{P}-\alpha$ (78) and the origin (when $\alpha\in\mathcal{A}$), and because $\mathcal{A}$ has dimension r by definition (80). Now, any distance-squared from the origin to the polyhedron $\mathcal{P}-\alpha$ can be formulated

$$\{p^Tp \;=\; \|p\|^2 \;=\; 2y^TV_{\mathcal{N}}^TX^TXV_{\mathcal{N}}\,y \mid y\in\mathcal{S}-\beta\} \tag{121}$$

Applying (85) to (121) we get (109).

[33] The solutions a constitute a hyperplane orthogonal to the vector $\mathbf{1}$, and offset from the origin in $\mathbb{R}^N$ by any particular solution; in this case, $a=e_1$. Since $\mathcal{R}(V_{\mathcal{N}})=\mathcal{N}(\mathbf{1}^T)$ and $\mathcal{N}(\mathbf{1}^T)\perp\mathcal{R}(\mathbf{1})$, $V_{\mathcal{N}}\,s$ is a hyperplane orthogonal to $\mathbf{1}$ for $s\in\mathbb{R}^{N-1}$.

$\Longleftarrow$ To validate the assertion in the reverse direction, we must show that if each distance-squared $\|p(y)\|^2$ on the shifted unit simplex $\mathcal{S}-\beta$ corresponds to a point $p(y)$ in some embedded polyhedron $\mathcal{P}-\alpha$, then D is an EDM. The r-dimensional subspace $\mathcal{A}-\alpha \subseteq \mathbb{R}^n$ is spanned by

$$p(\mathcal{S}-\beta) = \mathcal{P}-\alpha \tag{122}$$

because $\mathcal{A}-\alpha = \operatorname{aff}(\mathcal{P}-\alpha) \supseteq \mathcal{P}-\alpha$ (78). So, outside the domain $\mathcal{S}-\beta$ of linear function $p(y)$, the simplex complement $\overline{\mathcal{S}-\beta} \subset \mathbb{R}^{N-1}$ must contain the domain of the distance-squared $\|p(y)\|^2 = p^T(y)p(y)$ to the remaining points in the subspace $\mathcal{A}-\alpha$; *id est*, to the polyhedron's relative exterior $\overline{\mathcal{P}-\alpha}$. For $\|p(y)\|^2$ to be nonnegative on the entire subspace $\mathcal{A}-\alpha$, $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ must be positive semidefinite and is assumed symmetric;[34]

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} \triangleq \Phi^T\Phi \tag{123}$$

where $\Phi \in \mathbb{R}^{m\times N-1}$ for some $m \geq r$. Because $p(\mathcal{S}-\beta)$ is a convex polyhedron, it is necessarily a set of linear combinations of points from some length-N list because every convex polyhedron having N or fewer vertices can be generated that way (§4.6). [2, §19] Equivalent to (109) are

$$\{p^Tp \mid p \in \mathcal{P}-\alpha\} \;=\; \{p^Tp = y^T\Phi^T\Phi\, y \mid y \in \mathcal{S}-\beta\} \tag{124}$$

Because $p \in \mathcal{P}-\alpha$ may be found by factoring (124), the list Φ is found by factoring (123). A unique EDM can be constructed from that list using the inner-product form definition $\mathcal{D}(\Theta)|_{\Theta=\Phi}$ (36). That EDM will be identical to D if $\delta(D)=0$ by injectivity of $\mathcal{D}(D)$ (54). ♦

[34]The antisymmetric part $\left(-V_{\mathcal{N}}^T D V_{\mathcal{N}} - (-V_{\mathcal{N}}^T D V_{\mathcal{N}})^T\right)/2$ is annihilated by $\|p(y)\|^2$. $A^T = A \succeq 0 \;\Leftrightarrow\; A = R^TR$ for some real matrix R. [8, §6.3]

5.3.4 List reconstruction

At the stage of reconstruction, $D\in\mathbf{EDM}^N$ and we wish to find a generating list (§4.5, §4.6) for $\mathcal{P}-\alpha$ by factoring positive semidefinite $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ (123) as suggested in §5.3.3. One way to factor (123) is via *diagonalization* of symmetric matrices; [8][5]

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} \triangleq Q\Lambda Q^T \tag{125}$$

$$Q\Lambda Q^T \succeq 0 \;\Leftrightarrow\; \Lambda \succeq 0 \tag{126}$$

where $Q\in\mathbb{R}^{N-1\times N-1}$ is an orthogonal matrix containing eigenvectors while $\Lambda\in\mathbb{R}^{N-1\times N-1}$ is a diagonal matrix containing corresponding nonnegative eigenvalues. From the diagonalization, identify the list using (85);

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} \;=\; 2V_{\mathcal{N}}^T X^T X V_{\mathcal{N}} \;\triangleq\; Q\sqrt{\Lambda}\,Q_o^T Q_o\sqrt{\Lambda}Q^T \tag{127}$$

where $\sqrt{\Lambda}\,Q_o^T Q_o\sqrt{\Lambda} \triangleq \Lambda = \sqrt{\Lambda}\sqrt{\Lambda}$, and where $Q_o\in\mathbb{R}^{n\times N-1}$ is unknown as is its dimension n. Rotation/reflection is accounted for by Q_o yet only its first r columns are necessarily orthonormal.[35] Assuming $y\in\mathcal{S}$ then $p=\sqrt{2}XV_{\mathcal{N}}\,y$ $=Q_o\sqrt{\Lambda}\,Q^T y$ in $\mathbb{R}^n$ belongs to the polyhedron $\mathcal{P}-x_1$ whose generating point list constitutes (81) the columns of

$$\begin{bmatrix}\mathbf{0} & \sqrt{2}XV_{\mathcal{N}}\end{bmatrix} \;=\; \begin{bmatrix}\mathbf{0} & Q_o\sqrt{\Lambda}\,Q^T\end{bmatrix}\in\mathbb{R}^{n\times N} \tag{128}$$

If we like, we may choose n to be

$$\operatorname{rank} V_{\mathcal{N}}^T D V_{\mathcal{N}} \;=\; \operatorname{rank} Q_o\sqrt{\Lambda}\,Q^T \;=\; \operatorname{rank}\Lambda \;=\; r \tag{129}$$

which is the smallest n possible[36] because $XV_{\mathcal{N}}$ has rank $r\leq n$. (89) The simplest choice for Q_o is $[I \;\; \mathbf{0}]\in\mathbb{R}^{r\times N-1}$ where $r\leq N-1$. (84)

[35] Q_o is not necessarily an orthogonal matrix. Q_o is constrained such that only its first r columns are necessarily orthonormal because there are only r nonzero eigenvalues in Λ when $V_{\mathcal{N}}^T D V_{\mathcal{N}}$ has rank r (§4.9). The remaining columns of Q_o are arbitrary.

[36] If we write $Q^T=\begin{bmatrix} q_1^T \\ \vdots \\ q_{N-1}^T\end{bmatrix}$ as row-wise eigenvectors, $\Lambda=\begin{bmatrix}\lambda_1 & & & & 0\\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & 0\end{bmatrix}$ in terms of eigenvalues, and $Q_o=\begin{bmatrix}q_{o_1} & \cdots & q_{o_{N-1}}\end{bmatrix}$ as column vectors, then $Q_o\sqrt{\Lambda}\,Q^T=\sum\limits_{i=1}^{r}\sqrt{\lambda_i}\,q_{o_i}\,q_i^T$ is a sum of r linearly independent rank-one matrices. Hence the summation has rank r.

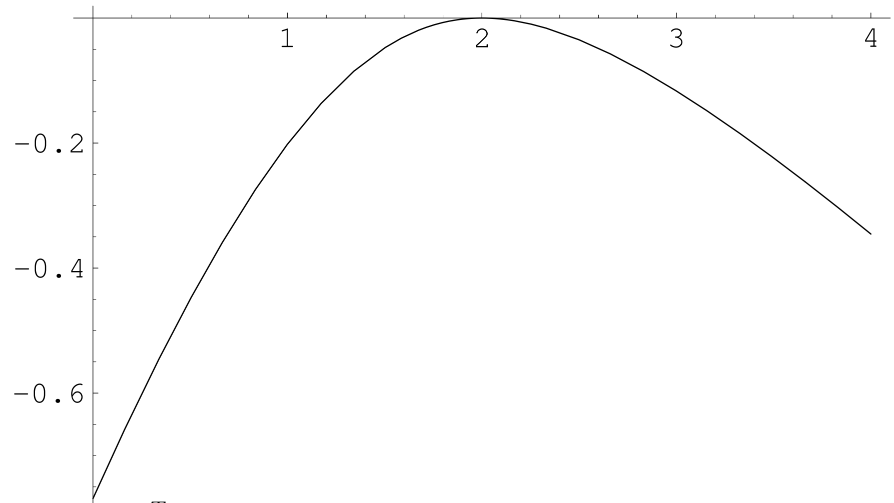

Figure 5: $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ is positive semidefinite for only one value of d_{14}; 2.

Given the list (128) found from the diagonalization of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$, we might wish to verify it. Because of offset and rotation/reflection invariance (§3.4), EDM D can be uniquely constructed from that list by calculating: (29)

$$\mathcal{D}(X) = \mathcal{D}(X[\mathbf{0}\ \ \sqrt{2}V_{\mathcal{N}}]) = \mathcal{D}(Q_o[\mathbf{0}\ \ \sqrt{\Lambda}\,Q^T]) = \mathcal{D}([\mathbf{0}\ \ \sqrt{\Lambda}\,Q^T]) \qquad (130)$$

5.4 Example revisited

We now apply the necessary and sufficient criteria (107), for an EDM, to an earlier problem.

Example. *Small completion problem.* Continuing the example in §3.2 which pertains to Figure 2 where $N=4$, d_{14} is ascertainable from the matrix inequality $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$. Because all distances in (24) are known except $\sqrt{d_{14}}$, we may simply calculate the minimum eigenvalue of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ over a range of d_{14} as in Figure 5. We observe a unique value of d_{14} satisfying (107) where the abscissa is tangent to the minimum eigenvalue. Since the minimum eigenvalue of a symmetric matrix is known to be a concave function (§5.2.4), there are no other satisfying values of d_{14}.

5.5 EDM indefiniteness

Any symmetric positive semidefinite matrix having a zero entry on its main diagonal must be zero along the entire row and column to which that zero belongs. [6, §4.2.8] [5, §7.1, prob.2] In other words, when $D \in \mathbf{EDM}^N$, there can be no positive nor negative semidefinite EDM except the zero matrix because $\mathbf{EDM}^N \subseteq \mathbb{S}_\delta^N$ (28) and

$$\mathbb{S}_\delta^N \cap \mathbb{S}_+^N = \mathbf{0} \tag{131}$$

the zero matrix. So, there can be no factorization $D = A^T A$ nor $-D = A^T A$. [8, §6.3] Hence the eigenvalues of an EDM are neither all nonnegative nor all nonpositive; an EDM is indefinite.

5.5.1 EDM eigenvalues

For any symmetric $-D$, we can find its rank and characterize its eigenvalues by *congruence transformation*: [8, §6.3]

$$-C^T D C \;\triangleq\; -\begin{bmatrix} V_{\mathcal{N}}^T \\ \frac{1}{\sqrt{2}}\mathbf{1}^T \end{bmatrix} D \begin{bmatrix} V_{\mathcal{N}} & \dfrac{1}{\sqrt{2}}\mathbf{1} \end{bmatrix} \;=\; -\begin{bmatrix} V_{\mathcal{N}}^T D V_{\mathcal{N}} & \frac{1}{\sqrt{2}} V_{\mathcal{N}}^T D \mathbf{1} \\ \frac{1}{\sqrt{2}}\mathbf{1}^T D V_{\mathcal{N}} & \frac{1}{2}\mathbf{1}^T D \mathbf{1} \end{bmatrix} \tag{132}$$

Because $C \triangleq \begin{bmatrix} V_{\mathcal{N}} & \frac{1}{\sqrt{2}}\mathbf{1} \end{bmatrix} \in \mathbb{R}^{N \times N}$ has full rank,

$$\operatorname{rank} D \;=\; \operatorname{rank} C^T D C \tag{133}$$

The congruence $-C^T D C$ has the same number of positive, zero, and negative eigenvalues as $-D$. Further, if we denote the eigenvalues of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ by σ_i, $i \in 1 \ldots N-1$, the eigenvalues of $-C^T D C$ by ζ_i, $i \in 1 \ldots N$, and we arrange each respective set of eigenvalues in decreasing order, then by theory of *interlacing eigenvalues for bordered symmetric matrices*, [5, §4.3]

$$\zeta_N \leq \sigma_{N-1} \leq \zeta_{N-1} \leq \sigma_{N-2} \leq \cdots \leq \sigma_2 \leq \zeta_2 \leq \sigma_1 \leq \zeta_1 \tag{134}$$

When $D \in \mathbf{EDM}^N$, $\sigma_i \geq 0$ for all i [5, §7.2] because $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$, as we now know. That means the congruence must have $N-1$ nonnegative eigenvalues; $\zeta_i \geq 0$, $i \in 1 \ldots N-1$. The remaining eigenvalue ζ_N cannot be nonnegative because then $-D$ would be positive semidefinite, an impossibility; so $\zeta_N < 0$. By congruence, $-D$ must therefore have one and only one negative eigenvalue:

$$\begin{aligned} \partial_i &\geq 0\,, \;\; i \in 1 \ldots N-1 \\ \partial_N &= -\sum_{i=1}^{N-1} \partial_i \;<\; 0 \end{aligned} \tag{135}$$

where $\partial_i\,,\ i \in 1 \ldots N$, are the eigenvalues of $-D$; their sum must be zero only because $\operatorname{tr} D = 0$. [8, §5.1]

6 Fifth Euclidean requirement

We continue now with the question raised in §3.2 regarding the necessity for at least one requirement more than the four Euclidean axioms (§3.1) to reconstruct a convex polyhedron or its generating list from incomplete distance information. There we saw that the Euclidean axioms become insufficient when the number of points N exceeds three, regardless of affine dimension.

6.1 Path not followed

6.1.1 $N = 4$

An intuitively appealing way to augment the Euclidean axioms is to recognize that the three-dimensional analog to triangle & distance is tetrahedron & facet-area. Each of the four facets of a tetrahedron is a triangle and its relative interior. Suppose we identify each facet of a non-regular tetrahedron by its area-squared: A_1, A_2, A_3, A_4. Then analogous to axiom 4, we may write an area inequality for the facets

$$\sqrt{A_i} \;\leq\; \sqrt{A_j} + \sqrt{A_k} + \sqrt{A_l}\,, \quad i \neq j \neq k \neq l \in 1,2,3,4 \tag{136}$$

which is a generalized triangle inequality [13, §1.1] that follows from [23] [18, *Law of Cosines*]

$$\sqrt{A_i} = \sqrt{A_j}\cos\phi_{ij} + \sqrt{A_k}\cos\phi_{ik} + \sqrt{A_l}\cos\phi_{il} \tag{137}$$

where ϕ_{ij} is the *dihedral* angle at the common edge between triangular facets i and j. Conveniently, the area of the i^{th} triangle has a formula in terms of $D_i \in \mathbf{EDM}^{N-1}$, the EDM corresponding to that particular triangle; *id est*, from the *Cayley-Menger determinant* [18],

$$A_i = \frac{(-1)^N}{2^{N-1}(N-1)!^2}\det\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & D_i \end{bmatrix} \tag{138}$$

where D_i is a principal submatrix of $D \in \mathbf{EDM}^N$, the EDM of the whole tetrahedron. The number of principal 3×3 submatrices of D is, of course, equal to the number of triangles in the tetrahedron; $N!/(3!(N-3)!)$.

6.1.2 $N=5$

Moving to the next level, we may encounter an object called the *polychoron*, a polyhedron in four dimensions.[37] The analog to triangle & distance is now polychoron & facet-volume. The polychoron has five facets, $N!/(4!(N-4)!)$, each of them a non-regular tetrahedron whose *volume*-squared is calculated using the same formula; (138) where D is now the EDM of the polychoron and D_i is the EDM of the i^{th} tetrahedron. We could then write another inequality in terms of facet volume, like (136) instead having four terms on the right-hand side, and so on.

6.2 Recapitulate

From (92) we learned that the matrix inequality $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0$ is a necessary test for D to be an EDM. In §5.3.1 the connection between convex polyhedra and EDMs was pronounced by the EDM assertion. The matrix inequality together with $D \in \mathbb{S}_\delta^N$ became a sufficient test in §5.3.3 when we demanded that every convex polyhedron have a corresponding EDM. The matrix criteria in (107) for the existence of an EDM are therefore necessary and sufficient for all $N>1$ (App. B), and subsume all the Euclidean requirements.

In the particular case $N=3$, $-V_{\mathcal{N}}^T D V_{\mathcal{N}} = T$ defined in (98). So $T \succeq 0$ and $D \in \mathbb{S}_\delta^{\mathbf{3}}$ are the necessary and sufficient conditions for D to be an EDM. From (101), the triangle inequality is then the only Euclidean constraint on the bounds of the necessarily nonnegative d_{ij}; and those bounds are tight. That means the four axioms of Euclidean space (§3.1) are necessary and sufficient requirements for D to be an EDM in the case $N=3$; *id est*, the fifth Euclidean requirement (§3.2) becomes coincident with axiom 4: for $i,j \in 1,2,3$

$$\begin{array}{l} \sqrt{d_{ij}} \geq 0,\ i \neq j \\ \sqrt{d_{ij}} = 0,\ i = j \\ \sqrt{d_{ij}} = \sqrt{d_{ji}} \\ \sqrt{d_{ij}} \leq \sqrt{d_{ik}} + \sqrt{d_{kj}}\,,\ i \neq j \neq k \end{array} \quad\Leftrightarrow\quad \begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0 \\ D \in \mathbb{S}_\delta^{\mathbf{3}} \end{array} \quad\Leftrightarrow\quad D \in \mathbf{EDM}^{\mathbf{3}} \tag{139}$$

Yet the four axioms become insufficient when $N>3$.

[37] The simplest regular polychoron is called a pentatope. A pentahedron is a three-dimensional object having five vertices.

6.3 Derivation

The correspondence between the triangle inequality and the EDM was developed in §5.2 where a triangle inequality (101a) was revealed within the leading principal 2×2 submatrix of positive semidefinite $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$. Our choice of the *leading* principal submatrix was arbitrary; actually, a unique triangle inequality like (35) corresponds to any one of the $(N-1)!/(2!(N-1-2)!)$ principal 2×2 submatrices.[38] Because all principal submatrices are positive semidefinite iff $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ is (App. F), then assuming $-V_{\mathcal{N}}^T D V_{\mathcal{N}} \in \mathbb{S}^{\mathbf{3}}$ and $D \in \mathbb{S}_\delta^{\mathbf{4}}$ it is sufficient to show that all d_{ij} are nonnegative (1×1 principal submatrices), all triangle inequalities are satisfied (2×2), and $\det(-V_{\mathcal{N}}^T D V_{\mathcal{N}})$ is nonnegative (3×3). When $N=4$, in other words, the nonnegative determinant is the fifth and last requirement for $D\in\mathbf{EDM}^N$.

6.3.1 Nonnegative determinant

When $D\in\mathbb{S}_\delta^{\mathbf{4}}$, from (94)

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}} = \begin{bmatrix} d_{12} & \frac{1}{2}(d_{12}+d_{13}-d_{23}) & \frac{1}{2}(d_{12}+d_{14}-d_{24}) \\ \frac{1}{2}(d_{12}+d_{13}-d_{23}) & d_{13} & \frac{1}{2}(d_{13}+d_{14}-d_{34}) \\ \frac{1}{2}(d_{12}+d_{14}-d_{24}) & \frac{1}{2}(d_{13}+d_{14}-d_{34}) & d_{14} \end{bmatrix} \tag{140}$$

By (87) when $D\in\mathbf{EDM}^{\mathbf{4}}$, $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ is equal to the inner-product (37),

$$\Theta^T\Theta = \begin{bmatrix} d_{12} & \sqrt{d_{12}d_{13}}\cos\theta_{213} & \sqrt{d_{12}d_{14}}\cos\theta_{214} \\ \sqrt{d_{12}d_{13}}\cos\theta_{213} & d_{13} & \sqrt{d_{13}d_{14}}\cos\theta_{314} \\ \sqrt{d_{12}d_{14}}\cos\theta_{214} & \sqrt{d_{13}d_{14}}\cos\theta_{314} & d_{14} \end{bmatrix} \tag{141}$$

Because Euclidean space has both a metric and an inner-product defined on it, the more concise inner-product form of the determinant is admitted;

$$\det(\Theta^T\Theta) = -d_{12}d_{13}d_{14}\big(\cos^2\theta_{213}+\cos^2\theta_{214}+\cos^2\theta_{314} - 2\cos\theta_{213}\cos\theta_{214}\cos\theta_{314} - 1\big) \tag{142}$$

[38] There are fewer principal 2×2 submatrices than triangles when $N>3$ because there are $N!/(3!(N-3)!)$ triangles made by point triples.

The determinant is nonnegative when

$$\begin{array}{rcccl}\cos\theta_{214}\cos\theta_{314}-\sqrt{\sin^2\theta_{214}\sin^2\theta_{314}} & \leq & \cos\theta_{213} & \leq & \cos\theta_{214}\cos\theta_{314}+\sqrt{\sin^2\theta_{214}\sin^2\theta_{314}}\\ \cos\theta_{213}\cos\theta_{314}-\sqrt{\sin^2\theta_{213}\sin^2\theta_{314}} & \leq & \cos\theta_{214} & \leq & \cos\theta_{213}\cos\theta_{314}+\sqrt{\sin^2\theta_{213}\sin^2\theta_{314}}\\ \cos\theta_{213}\cos\theta_{214}-\sqrt{\sin^2\theta_{213}\sin^2\theta_{214}} & \leq & \cos\theta_{314} & \leq & \cos\theta_{213}\cos\theta_{214}+\sqrt{\sin^2\theta_{213}\sin^2\theta_{214}}\end{array} \tag{143}$$

which simplifies, for $0\leq\theta_{ikj}\leq\pi$, to

$$\begin{array}{rcccl}\cos(\theta_{214}+\theta_{314}) & \leq & \cos\theta_{213} & \leq & \cos(\theta_{214}-\theta_{314})\\ \cos(\theta_{213}+\theta_{314}) & \leq & \cos\theta_{214} & \leq & \cos(\theta_{213}-\theta_{314})\\ \cos(\theta_{213}+\theta_{214}) & \leq & \cos\theta_{314} & \leq & \cos(\theta_{213}-\theta_{214})\end{array} \tag{144}$$

Analogously to the triangle inequality (105), the determinant is zero upon any equality in (144). Because point labelling is arbitrary, the fifth requirement must apply to each of the $N=4$ points as though each were in turn labelled x_1. Hence there exists a more general form of the fifth requirement:

Axiom. Fifth Euclidean requirement. *Angle inequality.* Augmenting the axioms of the metric in Euclidean space $\mathbb{R}^n$, the inequality for all $i\neq j\neq k\neq l\in 1\ldots N$

$$\begin{array}{c}\cos(\theta_{ikl}+\theta_{lkj}) \;\leq\; \cos\theta_{ikj} \;\leq\; \cos(\theta_{ikl}-\theta_{lkj}) \qquad (a)\\ 0\leq\theta_{ikl}\,,\theta_{lkj}\,,\theta_{ikj}\leq\pi\end{array} \tag{145}$$

where $\theta_{ikj}=\theta_{jki}$ is an angle between vectors at vertex x_k as defined in (33), must be satisfied at each point x_k. Inequality (145) can be equivalently written linearly as a triangle inequality, but between angles; *viz.*,

$$\begin{array}{cl}|\theta_{ikl}-\theta_{lkj}| \;\leq\; \theta_{ikj} \;\leq\; \theta_{ikl}+\theta_{lkj} & (a)\\ \theta_{ikl}+\theta_{lkj}+\theta_{ikj}\;\leq\;2\pi & (b)\\ 0\leq\theta_{ikl}\,,\theta_{lkj}\,,\theta_{ikj}\leq\pi & \end{array} \tag{146}$$

▲

Just as the triangle inequality is the ultimate test for reconstruction of only three points, the fifth Euclidean requirement is the ultimate test for only four. When the number of points N equals 4, the triangle inequality remains a necessary although penultimate test;

$$\begin{array}{l}\text{four Euclidean axioms (§3.1).}\\ \text{(145) or (146)}\end{array} \;\Leftrightarrow\; \begin{array}{r}-V_{\mathcal{N}}^T D V_{\mathcal{N}} \succeq 0\\ D \in \mathbb{S}_\delta^{\mathbf{4}}\end{array} \;\Leftrightarrow\; D \in \mathbf{EDM^4} \tag{147}$$

6.3.2 Beyond the fifth requirement

When the number of points N exceeds 4, the four Euclidean axioms and the angle inequality together become insufficient conditions for reconstruction. In other words, the angle inequality remains a necessary Euclidean requirement at each x_k, and becomes a sufficient test of the positive semidefiniteness of all the principal 3×3 submatrices in $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$. The angle inequality can be considered a test of integrity for every purported tetrahedron.

When $N = 5$ in particular, the angle inequality becomes the penultimate Euclidean requirement while nonnegativity of $\det(\Theta^T\Theta)$ corresponds, by induction, to the sixth and last Euclidean requirement (but its expression becomes unwieldy), and so on.

6.4 Affine dimension reduction in three dimensions

The determinant of any $M\times M$ matrix is equal to the product of its M eigenvalues. [8, §5.1] When $N = 4$ and $\det(\Theta^T\Theta)$ is zero, that means one or more eigenvalues of $\Theta^T\Theta \in \mathbb{R}^{3\times3}$ are zero. The determinant will go to zero whenever equality is attained on either side of (145a), (146a), or (146b), meaning that a tetrahedron has collapsed to a lower affine dimension; *id est*, $r = \operatorname{rank}\Theta^T\Theta = \operatorname{rank}\Theta$ is reduced below $N-1$ exactly by the number of zero eigenvalues (§4.9).

Therefore, in solving completion problems of any size N where one or more entries of an EDM are unknown, the dimension r of the affine hull required to contain the unknown points is potentially reduced by selecting distances to attain equality in (145a) or (146a) or (146b).

6.4.1 Example again revisited

We now apply the fifth Euclidean requirement to an earlier problem.

Example. *Small completion problem.* Returning again to the example in §3.2 which pertains to Figure 2 where $N=4$ (*confer* §5.4), d_{14} is ascertainable from the fifth Euclidean requirement. Because all distances in (24) are known except $\sqrt{d_{14}}$, $\cos\theta_{123}=0$ and $\theta_{324}=0$ are results from identity (33). Applying (145),

$$\begin{array}{rcccl} \cos(\theta_{123}+\theta_{324}) & \leq & \cos\theta_{124} & \leq & \cos(\theta_{123}-\theta_{324}) \\ 0 & \leq & \cos\theta_{124} & \leq & 0 \end{array} \tag{148}$$

It follows from (33) that d_{14} can only be 2. Because equality is attained in (148), the affine dimension r cannot exceed $N-2$, as explained.

7 Closest EDM

Can be solved analytically.
Do example 2.
Mention that this was thought to be a hard problem in 80's, mention Gower.

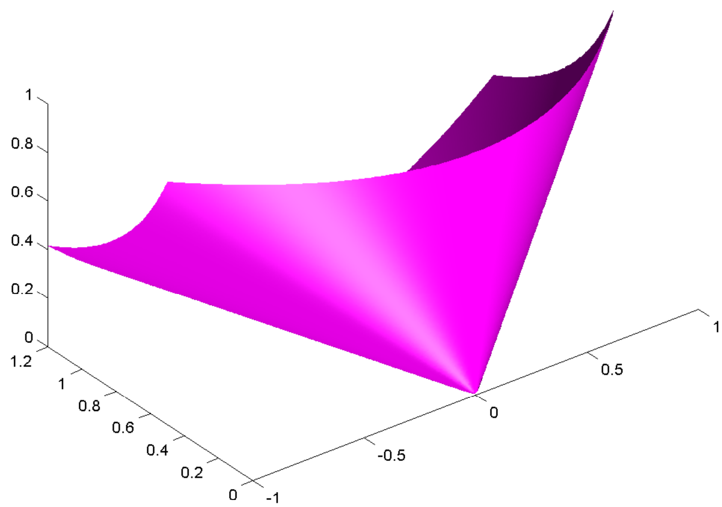

Figure 6: Boundary of positive semidefinite cone of $\mathbb{S}^2$ matrices plotted in $\mathbb{R}^3$. Courtesy, Alexandre W. d'Aspremont.

8 PSD and EDM cones

The geometric object called the convex cone was introduced in §4.4. The set of all EDMs forms a convex cone in $\mathbb{S}^N$ because they satisfy the defining equation (66). That will become more apparent after we examine the convex cone of positive semidefinite symmetric matrices, also in $\mathbb{S}^N$. From §5.5, we know that the EDM cone does not intersect the PSD cone except at the origin, their only vertex. Even so, we will still be able to relate the two cones.

8.1 Positive semidefinite (PSD) cone

From the definition of positive semidefiniteness, $A \succeq 0 \Leftrightarrow y^T A y \geq 0$ for all $y \in \mathbb{R}^M$. To see where the infinite number of halfspaces comes from, imagine for each particular y the product $y^T A y$ is a linear function of the matrix entries... show for M=2

The PSD cone (10) is a geometric object unique to each particular value of M in $\mathbb{R}^{M(M+1)/2}$. For $M=2$ the PSD cone is a three-dimensional object, semi-infinite in expanse, whose boundary is illustrated, but shown truncated, in Figure 6. For $M=3$ the PSD cone is six-dimensional.

8.1.1 Rank along boundary of PSD cone

All the positive semidefinite matrices having at least one 0 eigenvalue reside on the boundary of the PSD cone in any dimension. The boundary is delimited by all the supporting hyperplanes defined by the scalar inequality.

Prove that in the case of the PSD cone, if A belongs to the boundary of $\mathcal{C}$, then the ray generated by A must travel along the boundary of the convex cone. Use the 0 eigenvalue theorem, the fact that ζA must belong to the cone for all $\zeta \geq 0$ by definition, and the fact that scaling A simply scales its eigenvalues. Therefore there exist rays base 0 through the boundary of the PSD cone in any dimension.

In Figure 6, the only matrix having two 0 eigenvalues lies at the origin.

$$\operatorname{rank} A + \operatorname{rank} B \;\geq\; \operatorname{rank}(A+B) \;\geq\; \max\{\operatorname{rank} A,\, \operatorname{rank} B\} \tag{149}$$

Left-hand side from [5, §0.4] looks like some sort of triangle inequality, and holds for all A, B. rank is a quasiconcave function (§2.2.1) on $\mathbb{S}_+^N$ because the right-hand side has the form of (18), sort of.

But let us instead explain this inequality more intuitively: *Nullspace.*

$$z^T(A+B)z = z^TAz + z^TBz \tag{150}$$

Matrices symmetric. Number of zero eigenvalues equals dimension of nullspace by theorem... Nullspace of right-hand side is intersection of two nullspaces which is smaller than or same as the nullspace of $A+B$. When nullspace diminishes, rank increases.

Range. Same explanation, alternate point of view. Imagine the polyhedral cone generated by the extreme directions connoted by A and B;

$$\mathcal{C} = \{\zeta A + \xi B \;\text{ for all } \zeta, \xi \geq 0\} \tag{151}$$

Now suppose A and B each lie on the boundary of the PSD cone but are not collinear with the origin. Then the polyhedral cone $\mathcal{C}$ is a slice of the PSD cone containing A, B, and the origin. An example of such a slice is shown in Figure 7. When both ζ and ξ are nonzero, $\zeta A + \xi B$ resides in the relative interior of the polyhedral cone, but interior to the PSD cone. Hence, the sum must correspond to a matrix having higher rank. If one or the other coefficient is 0, then...

In higher dimensions, the statements above remain true...

8.1.2 Polyhedral cones within

Definition. *Set slice.* We define a *slice* of a set $\mathcal{C}$ as the intersection of $\mathcal{C}$ with a (two-dimensional) plane. When $\mathcal{C}$ is convex, the slice remains convex by the Intersection theorem (§2.1).

In the particular case of positive semidefinite matrices, (66) is certainly true; [5, §7.1] that suggests:

Theorem. *Cone intersection.* [2, §2] The intersection of an arbitrary collection of convex cones is a convex cone.

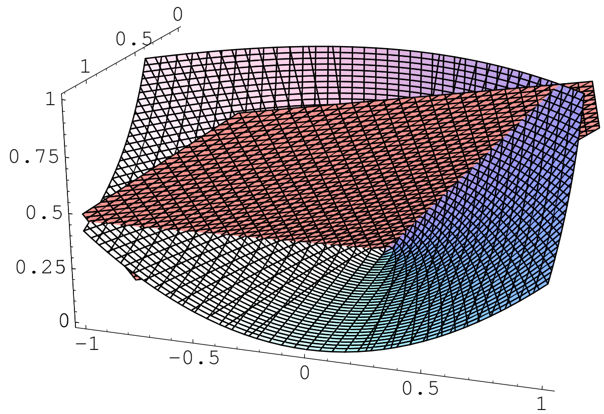

Figure 7:

Slice decomposition. The PSD cone of any dimension comprises an infinite number of two-dimensional polyhedral cones revealed by slicing it with any plane through the origin. The extreme directions (§4.4) of the polyhedral cones are extreme directions of the PSD cone.

Indeed, any slice through the origin of the PSD cone is a polyhedral cone whose extreme directions are sketched in Figure 8 for a number of arbitrarily chosen slices. The rays shown emanating from the origin all lie along the boundary of the PSD cone and along the relative boundary of the corresponding polyhedral cone. Hence the extreme directions of the polyhedral cones are also extreme directions of the PSD cone. In Figure 8, any two rays are extreme directions of some slice through the origin.

By (66) it follows that there exist rays emanating from the origin which travel along the boundary of the PSD cone of any dimension, because each of those rays is an extreme direction of some polyhedral cone made by slicing the PSD cone using a plane through the origin.

Alternate construction of PSD cone is convex hull of all the extreme directions and the origin, by the extremes theorem in §4.2.

8.2 Hyperplanes hinged on PSD cone

(66) is the inspiration for this section. From Sweep.nb,
Show that all rank 1 3x3 and 2x2 matrices, and all rank 2 3x3 matrices can

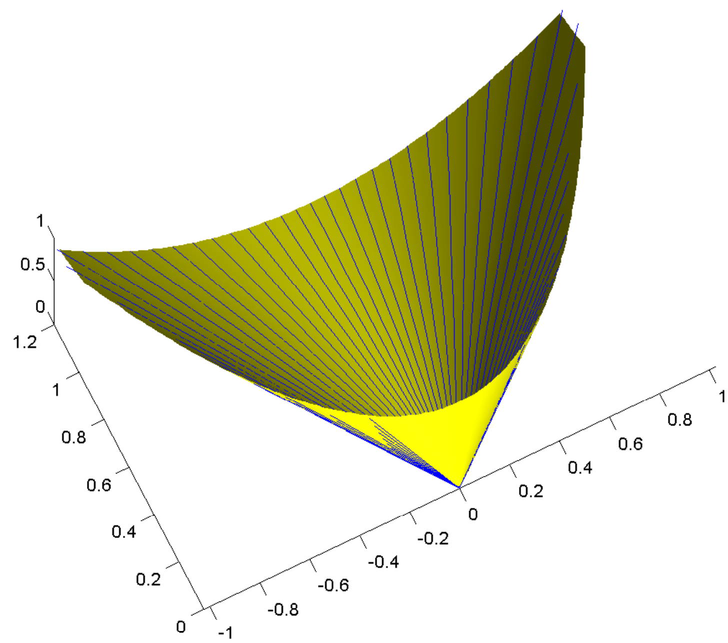

Figure 8: Showing extreme directions.

be described by matrix on pg.141 of nascence notebook. Plot a slice of the 3-dimensional surface of the 6-dimensional PSD cone for particular s1 and s3; it is described completely by the nullspace of the hinge equations pg.140 ibid.

8.3 EDM cone

We may rewrite (107) slightly to emphasize that D belongs to the EDM cone when $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ belongs to the positive semidefinite cone.

Definition. *Cone of Euclidean distance matrices*; $\mathbf{EDM}^N$. The set of all Euclidean distance matrices forms a geometric object unique to each particular value of N in $\mathbb{R}^{N(N-1)/2}$, called the EDM cone: (*confer* (53))

$$\begin{array}{r} -V_{\mathcal{N}}^T D V_{\mathcal{N}} \in \mathbb{S}_+^N \\ D \in \mathbb{S}_\delta^N \end{array} \;\Leftrightarrow\; D \in \mathbf{EDM}^N \tag{152}$$

As shown in §5.5, there can be no positive nor negative semidefinite EDM, so the EDM cone and the positive semidefinite cone of symmetric matrices cannot intersect except at the origin;

$$\mathbf{EDM}^N \cap \mathbb{S}_+^N = 0 \tag{153}$$

Assuming D_1 and D_2 belong to $\mathbb{S}_\delta^N$, then $D_2 - D_1$ belongs to the EDM cone iff $-V_{\mathcal{N}}^T(D_2 - D_1)V_{\mathcal{N}} \succeq 0$ by (107).

$$D_1 \underset{\mathbf{EDM}^N}{\preceq} D_2 \quad\Leftrightarrow\quad \left\{ \begin{array}{c} -V_{\mathcal{N}}^T D_1 V_{\mathcal{N}} \preceq -V_{\mathcal{N}}^T D_2 V_{\mathcal{N}} \\ D_1\,, D_2 \in \mathbb{S}_\delta^N \end{array} \right. \tag{154}$$

From (108), any matrix V in place of $V_{\mathcal{N}}$ such that $\mathcal{R}(V) = \mathcal{N}(\mathbf{1}^T)$ will satisfy (154).

The *dual positive semidefinite cone* is defined

$$\mathbb{S}_+^{M*} = \{Y \in \mathbb{S}^M \mid \operatorname{tr} X^T Y \geq 0 \text{ for all } X \in \mathbb{S}_+^M\} = \mathbb{S}_+^M \tag{155}$$

The positive semidefinite cone is self-dual.
Find the dual EDM cone.

9 EDM completion

Intriguing is the question of whether the list in X may be reconstructed given an incomplete EDM. We have already examined this problem for a small example given in §3.2 and then revisited in §5.4 and §6.4.1. When the number of points N exceeds 4, it is no longer convenient to use the fifth Euclidean requirement as we did in §6.4.1; we need a more general method.

Other researchers [7][24] have formulated this completion problem in a non-convex way. We will utilize the convexity of $-V_{\mathcal{N}}^T\mathcal{D}(X)V_{\mathcal{N}}$ (§3.5) to reconstruct the list.

10 Least squares problem solving via EDM

11 Map of the USA

A fundamental application of EDMs is to reconstruct relative point position given only the EDM. To draw the map of the USA will be a good illustration of our findings thus far, although this presentation is certainly not made for the first time.(citation)

some rows of X found by way of (ref) can (always) be truncated.

12 Optical character recognition

13 Spectral analysis

13.1 Fourier series

The set of all symmetric matrices $\mathbb{S}^M$ forms a subspace in $\mathbb{R}^{M\times M}$, so for it there exists a standard orthonormal basis;

$$E_{ij} = \begin{cases} e_i e_i^T\,, & i=j=1\ldots M \\ \frac{1}{\sqrt{2}}\left(e_i e_j^T + e_j e_i^T\right), & 1\leq i<j\leq M \end{cases} \tag{156}$$

where there are $M(M+1)/2$ standard basis matrices $E_{ij}\in\mathbb{S}^M$ formed from the standard basis vectors $e_i\in\mathbb{R}^M$. The inner product $<E_{ij}\,,\,\mathcal{C}_1>$ becomes a coefficient of orthogonal projection, and any element of $\mathcal{C}$ can be written as a Fourier series [25, §2]

$$\mathcal{C}_1 = \sum_{\substack{i,j\,=\,1 \\ j\,\geq\,i}}^{M} <E_{ij}\,,\,\mathcal{C}_1>\; E_{ij} \tag{157}$$

which, in this case, is a projection on the standard basis matrices. From the set-convexity corollary in §2.1, it follows that $<E_{ij}\,,\,\mathcal{C}>$ is a convex set when $\mathcal{C}$ is.

Because any symmetric matrix can be diagonalized [8, §5.6], $\mathcal{C}_1$ has a decomposition in terms of its *eigenmatrices* $q_i q_i^T$ and eigenvalues λ_i,

$$\mathcal{C}_1 \;=\; Q\Lambda Q^T = \sum_{i=1}^{M}\lambda_i\, q_i q_i^T \tag{158}$$

where $\Lambda\in\mathbb{S}^M$ is a diagonal matrix having $\delta(\Lambda)_i=\lambda_i$ and $Q=[q_1\cdots q_M]\in\mathbb{S}^M$ is an orthogonal matrix containing eigenvectors. If we rotate the standard basis matrices into alignment with the eigenmatrices of $\mathcal{C}_1$ by applying a traditional *similarity transformation*, [8, §5.6]

$$QE_{ij}Q^T = \begin{cases} q_i q_i^T\,, & i=j=1\ldots M \\ \frac{1}{\sqrt{2}}\left(q_i q_j^T + q_j q_i^T\right), & 1\leq i<j\leq M \end{cases} \tag{159}$$

a remarkable thing happens to the Fourier series:

$$\begin{aligned}
\mathcal{C}_1 &= \sum_{\substack{i,j=1 \\ j \geq i}}^{M} < QE_{ij}Q^T,\, \mathcal{C}_1 > \; QE_{ij}Q^T \\
&= \sum_{i=1}^{M} < q_i q_i^T,\, \mathcal{C}_1 > \; q_i q_i^T \\
&= \sum_{i=1}^{M} \lambda_i \, q_i q_i^T
\end{aligned} \tag{160}$$

The eigenvalues are clearly the coefficients of orthogonal projection of any symmetric matrix on its eigenmatrices. The remaining $M(M-1)/2$ Fourier series coefficients for $i \neq j$ are zeroed by the projection. Each element of $\mathcal{C}$ generally brings a different eigenmatrix, so unfortunately the set-convexity corollary does not apply to eigenvalues corresponding to the subspace of all symmetric matrices. The set of all symmetric rank-one positive semidefinite matrices does not form a subspace, hence there can be no standard basis for it.[39] The set of all *circulant* matrices (§13.3) forms a subspace whose members all have the same eigenvectors; the orthogonal basis from the DFT.

$$\max_{UU^T = I_k} \operatorname{tr} UU^T \mathcal{C}_1 = \sum_{i=1}^{k} \lambda_i \tag{161}$$

$$\operatorname{tr} QQ^T \mathcal{C}_1 = < QQ^T,\, \mathcal{C}_1 > = \sum_{i=1}^{M} \lambda_i \tag{162}$$

$\mathcal{C}_1$ is projected onto the range of Q.

[39] The sum of any two nonzero matrices of the form ww^T has rank one or two; indeed, for A and $B \in \mathbb{S}_+^M$ [5, §0.4] [1, §3],

$$\operatorname{rank} A + \operatorname{rank} B \;\geq\; \operatorname{rank}(A+B) \;\geq\; \max\{\operatorname{rank} A,\, \operatorname{rank} B\}$$

13.2 DFT

The discrete Fourier transform (DFT) is a staple of the digital signal processing community. [26] In essence, the DFT is a correlation of a windowed *sequence* (or *discrete signal*) with exponentials whose frequencies are equally spaced on the unit circle.[40] The DFT of the sequence $\{f(i) \in \mathbb{R},\ i=0 \ldots n-1\}$ is, in traditional form,[41]

$$F(k) = \sum_{i=0}^{n-1} f(i)\, e^{-j\, i 2\pi k/n} \tag{163}$$

for $k = 0 \ldots n-1$ and $j = \sqrt{-1}$. The implicit window on $f(i)$ in (163) is rectangular. The values $\{F(k) \in \mathbb{C},\ k=0 \ldots n-1\}$ are considered a spectral analysis of the sequence $f(i)$; *id est*, the $F(k)$ are amplitudes of exponentials which when combined, give back the original sequence,

$$f(i) = \frac{1}{n} \sum_{k=0}^{n-1} F(k)\, e^{j\, i 2\pi k/n} \tag{164}$$

The argument of F, the index k, corresponds to the discrete frequencies $2\pi k/n$ of the exponentials $e^{j\, i 2\pi k/n}$ in the synthesis equation (164).

The matrix form of the DFT is written

$$F = Wf \tag{165}$$

where $F = [F(k),\ k=0 \ldots n-1]$, $f = [f(i),\ i=0 \ldots n-1]$, and the *DFT matrix* is [27]

$$W = W^T = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & e^{-j2\pi k/n} & e^{-j4\pi k/n} & \cdots & e^{-j(n-1)2\pi k/n} \\ 1 & e^{-j4\pi k/n} & e^{-j8\pi k/n} & \cdots & e^{-j(n-1)4\pi k/n} \\ 1 & e^{-j6\pi k/n} & e^{-j12\pi k/n} & \cdots & e^{-j(n-1)6\pi k/n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & e^{-j(n-1)2\pi k/n} & e^{-j(n-1)4\pi k/n} & \cdots & e^{-j(n-1)^2 2\pi k/n} \end{bmatrix} \in \mathbb{R}^{n \times n} \tag{166}$$

[40] That is the unit circle in the z plane; $z = e^{sT}$ where $s = \sigma + j\omega$ is the traditional Laplace frequency, ω is the Fourier frequency in radians $2\pi f$, while T is the sample period.

[41] The convention is lowercase for the sequence and uppercase for its transform.

Direct implementation of (163) would require on the order of n^2 operations for large n. Similarly, the IDFT is

$$f = \frac{1}{n} W^H F \tag{167}$$

where the superscript H denotes the conjugate transpose of the DFT matrix. The solution to the computational problem of evaluating the DFT for large n culminated in the development of the fast Fourier transform (FFT) algorithm whose intensity is proportional to $n \log(n)$. [26]

13.3 Circulant matrices

Circulant matrices. Eigenvectors always same. Means set convexity corollary may be applied IF circulant makes a subspace. Idea is: Interpolating between any two circulant matrices would interpolate the known eigenvalues. Cite Gray.

For C circulant, where the first row is some time sequence $c_0, c_1, c_2, \ldots,$ let $X = C^T$. Then

$$\mathcal{D}(C^T) \;=\; k_i \mathbf{1}\mathbf{1}^T - \,2CC^T \tag{168}$$

where $k_i = 2\delta(CC^T)_i$ is any one of the diagonal entries which are all identical for circulant matrices. This is classical relationship between autocorrelation and similarity function where CC^T takes on the role of autocorrelation.

13.4 DFT via EDM

The DFT (163) is separable in the real and the imaginary part; meaning, the analysis exhibits no dependency between the two parts when the sequence is real; *viz.*,

$$F(k) = \sum_{i=0}^{n-1} f(i)\,\cos(i2\pi k/n) - j\sum_{i=0}^{n-1} f(i)\,\sin(i2\pi k/n) \tag{169}$$

It follows then, to relate the DFT to our work with EDMs, we should separately consider the Euclidean distance-squared between the sequence and each part of the complex exponentials. Augmenting the real list $\{x_\ell \in \mathbb{R}^n,\ \ell = 1 \ldots N\}$ will be the new imaginary list $\{y_\ell \in \mathbb{R}^n,\ \ell = 1 \ldots N\}$, where

$$\begin{aligned} x_1 \;=\; y_1 &\;\overset{\Delta}{=}\; [f(i),\ i=0\ldots n-1] \\ x_\ell &\overset{\Delta}{=} [\ \ \cos(i2\pi(\ell-2)/n),\ i=0\ldots n-1], \quad \ell = 2\ldots N \\ y_\ell &\overset{\Delta}{=} [-\sin(i2\pi(\ell-2)/n),\ i=0\ldots n-1], \quad \ell = 2\ldots N \end{aligned} \tag{170}$$

where $N = n+1$, and where the [ ] bracket notation means a vector made from a sequence. The row-1 entries (columns $\ell = 2\ldots N$) of EDM D^x are

$$\begin{aligned} d^x_{1\ell} &= \|x_\ell - x_1\|^2 \\ &= \sum_{i=0}^{n-1} (\cos(i2\pi(\ell-2)/n) \;-\; f(i))^2 \\ &= \sum_{i=0}^{n-1} \cos^2(i2\pi(\ell-2)/n) \;+\; f^2(i) \;-\; 2f(i)\cos(i2\pi(\ell-2)/n) \\ &= \tfrac{1}{4}\Big(2n+1+\tfrac{\sin(2\pi(\ell(2n-1)+2)/n)}{\sin(2\pi(\ell-2)/n)}\Big) \;+\; \tfrac{1}{n}\sum_{k=0}^{n-1} |F(k)|^2 \;-\; 2\,\Re F(\ell-2) \end{aligned} \tag{171}$$

where $\Re$ takes the real part of its argument, and where the Fourier summation is from the Parseval relation [25][26][27][28] for the DFT.[42] For the imaginary list we have a separate EDM D^y whose row-1 entries (columns

[42] The Fourier summation $\sum |F(k)|^2/n$ replaces $\sum f^2(i)$; we arbitrarily chose not to mix domains. Some physical systems, such as Magnetic Resonance Imaging devices, naturally produce signals originating in the Fourier domain. [29]

$\ell = 2 \dots N$) are

$$
\begin{aligned}
d^y_{1\ell} &= \|y_\ell - y_1\|^2 \\
&= \sum_{i=0}^{n-1} (\sin(i2\pi(\ell-2)/n) \; + \; f(i))^2 \\
&= \sum_{i=0}^{n-1} \sin^2(i2\pi(\ell-2)/n) \; + \; f^2(i) \; + \; 2f(i)\sin(i2\pi(\ell-2)/n) \\
&= \tfrac{1}{4}(2n - 1 - \tfrac{\sin(2\pi(\ell(2n-1)+2)/n)}{\sin(2\pi(\ell-2)/n)}) \; + \; \tfrac{1}{n}\sum_{k=0}^{n-1} |F(k)|^2 \; - \; 2\,\Im F(\ell-2)
\end{aligned}
\tag{172}
$$

where $\Im$ takes the imaginary part of its argument. In the remaining rows ($m = 2 \dots N,\ m < \ell$) of these two EDMs, D^x and D^y, we have[43]

$$
\begin{aligned}
d^x_{m\ell} &= \|x_\ell - x_m\|^2 \\
&= \sum_{i=0}^{n-1} (\cos(i2\pi(\ell-2)/n) \; - \; \cos(i2\pi(m-2)/n))^2 \\
&= \tfrac{1}{4}(4n + 2 + \tfrac{\sin(2\pi(\ell(2n-1)+2)/n)}{\sin(2\pi(\ell-2)/n)} + \tfrac{\sin(2\pi(m(2n-1)+2)/n)}{\sin(2\pi(m-2)/n)}) \\
d^y_{m\ell} &= \|y_\ell - y_m\|^2 \\
&= \sum_{i=0}^{n-1} (\sin(i2\pi(\ell-2)/n) \; - \; \sin(i2\pi(m-2)/n))^2 \\
&= \tfrac{1}{4}(4n - 2 - \tfrac{\sin(2\pi(\ell(2n-1)+2)/n)}{\sin(2\pi(\ell-2)/n)} - \tfrac{\sin(2\pi(m(2n-1)+2)/n)}{\sin(2\pi(m-2)/n)})
\end{aligned}
\tag{173}
$$

We observe from these distance-squared equations that only the first row and column of each EDM depends upon the sequence $f(i)$ itself. The remaining entries depend only upon the sequence length n.

To relate the EDMs D^x and D^y to the DFT in a useful way, consider the possibility of finding the inverse DFT (IDFT) via either EDM. For reasonable values of N, the number of EDM entries N^2 can become prohibitively large. Yet the DFT is subject to the same kind of computational intensity. It is neither our purpose nor goal to invent a fast algorithm for doing this, we simply present an example of finding an IDFT by way of the EDM. The technique we use was developed in §5.3.1:

[43] $\lim_{i\to 2} \sin(2\pi(i(2n-1) + 2)/n) \,/\, \sin(2\pi(i-2)/n) \; = \; 2n - 1$

14 Self similarity

Given the length-M real sequence $f(i)$, $i=0\ldots M-1$, we define the *self-similarity function*;[44]

$$A(\ell) \triangleq \frac{1}{2}\sum_{i=0}^{n-1}(f(i)-f(i-\ell))^2 \tag{174}$$

where $n\leq M$ is the window length.

$$x_{i+1} \triangleq f(i) \in \mathbb{R} \tag{175}$$

$$d_{ij} = \begin{cases} (x_i - x_j)^2, & 1 \leq i,j \leq M \\ 0 & \text{otherwise} \end{cases} \tag{176}$$

$$A(\ell) = \frac{1}{2}\sum_{i=1}^{n} d_{i,i-\ell} \tag{177}$$

which is a sum of some diagonal of the EDM D. To select the ij^{th} element of $\sqrt{D}$,

$$\sqrt{d_{ij}} \triangleq e_i^T\sqrt{D}e_j \;=\; e_j^T\sqrt{D}e_i \;=\; \operatorname{tr}(e_j e_i^T\sqrt{D}) = <e_i e_j^T,\, \sqrt{D}> \tag{178}$$

where here,

$$D \triangleq \sqrt{D}\circ\sqrt{D} \tag{179}$$

and where $\circ$ denotes the Hadamard product (§2.1). For $1\leq i<j\leq M$

$$\sqrt{2d_{ij}} = \sqrt{2}\,\operatorname{tr}(e_j e_i^T\sqrt{D}) = \frac{1}{\sqrt{2}}\operatorname{tr}\left((e_i e_j^T + e_j e_i^T)\sqrt{D}\right) \tag{180}$$

is a coefficient of orthogonal projection of $\sqrt{D}$ on a member of the orthonormal basis for the vector space $\mathbb{S}^M$.

$$A(\ell) = \sum_{i=1}^{n}\operatorname{tr}(e_i e_{i-\ell}^T D) \;=\; \operatorname{tr}\sum_{i=1}^{n} e_i e_{i-\ell}^T\, D \tag{181}$$

is the Parseval [25][26][27][28] relation giving the total energy of the projection.

[44] which is the progenitor of autocorrelation. When $n\to\infty$, $A(\ell)=K-2R(\ell)$ where, for some constant K, $R(\ell)$ is the autocorrelation function.

A Directional derivatives, gradients, Hessian, linear matrix inequality, matrix-valued Taylor series

A.1 First directional derivative

Assume that a differentiable function $g(X) : \mathbb{R}^{K\times L} \rightarrow \mathbb{R}^{M\times N}$ has continuous first and second-order gradients ∇g and $\nabla^2 g$ over $\operatorname{dom} g$ which is an open set. We seek simple expressions for the first and second directional derivatives, both in $\mathbb{R}^{M\times N}$, dubbed $\overset{\rightarrow Y}{dg}$ and $\overset{\rightarrow Y}{dg^2}$ respectively.

Assuming that the limit exists, we may state the partial derivative of the mn^{th} entry of g with respect to the kl^{th} entry of X;

$$\frac{\partial g_{mn}(X)}{\partial X_{kl}} = \lim_{\Delta t\rightarrow 0} \frac{g_{mn}(X + \Delta t\, e_k e_l^T) - g_{mn}(X)}{\Delta t} \in \mathbb{R} \tag{182}$$

where e_k is the k^{th} standard basis vector in $\mathbb{R}^K$ and e_l is the l^{th} standard basis vector in $\mathbb{R}^L$. The total number of partial derivatives equals $KLMN$ while the gradient is defined in their terms; the mn^{th} entry of the gradient is

$$\nabla g_{mn}(X) \triangleq \begin{bmatrix} \frac{\partial g_{mn}(X)}{\partial X_{11}} & \frac{\partial g_{mn}(X)}{\partial X_{12}} & \cdots & \frac{\partial g_{mn}(X)}{\partial X_{1L}} \\ \frac{\partial g_{mn}(X)}{\partial X_{21}} & \frac{\partial g_{mn}(X)}{\partial X_{22}} & \cdots & \frac{\partial g_{mn}(X)}{\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_{mn}(X)}{\partial X_{K1}} & \frac{\partial g_{mn}(X)}{\partial X_{K2}} & \cdots & \frac{\partial g_{mn}(X)}{\partial X_{KL}} \end{bmatrix} \in \mathbb{R}^{K\times L} \tag{183}$$

while the gradient is defined

$$\begin{aligned} \nabla g(X) &\triangleq \begin{bmatrix} \nabla g_{11}(X) & \nabla g_{12}(X) & \cdots & \nabla g_{1N}(X) \\ \nabla g_{21}(X) & \nabla g_{22}(X) & \cdots & \nabla g_{2N}(X) \\ \vdots & \vdots & & \vdots \\ \nabla g_{M1}(X) & \nabla g_{M2}(X) & \cdots & \nabla g_{MN}(X) \end{bmatrix} \in \mathbb{R}^{M\times N\times K\times L} \\ &= \begin{bmatrix} \frac{\partial g(X)}{\partial X_{11}} & \frac{\partial g(X)}{\partial X_{12}} & \cdots & \frac{\partial g(X)}{\partial X_{1L}} \\ \frac{\partial g(X)}{\partial X_{21}} & \frac{\partial g(X)}{\partial X_{22}} & \cdots & \frac{\partial g(X)}{\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g(X)}{\partial X_{K1}} & \frac{\partial g(X)}{\partial X_{K2}} & \cdots & \frac{\partial g(X)}{\partial X_{KL}} \end{bmatrix} \in \mathbb{R}^{K\times L\times M\times N} \end{aligned} \tag{184}$$

By simply rotating our perspective of the four-dimensional representation of the gradient matrix, we found a second equivalent expression in (184). If $M, N>1$ and $L=1$, then the entries of ∇g could be written into the cells of an MNK-size cube like Rubik's. [18] If $L>1$, then the entries of ∇g are matrices requiring a fourth dimension for their representation.

When the limit for $\Delta t \in \mathbb{R}$ exists, it is easy to show by substitution of variables in (182)

$$\frac{\partial g_{mn}(X)}{\partial X_{kl}} Y_{kl} = \lim_{\Delta t \to 0} \frac{g_{mn}(X + \Delta t\, Y_{kl}\, e_k e_l^T) - g_{mn}(X)}{\Delta t} \in \mathbb{R} \tag{185}$$

which may be interpreted as the change in g_{mn} at X when the change in X_{kl} is equal to Y_{kl}, the kl^{th} entry of $Y \in \mathbb{R}^{K\times L}$. Because the total change in $g_{mn}(X)$ due to Y is the sum of change with respect to each and every X_{kl}, the mn^{th} entry of the directional derivative is the corresponding total differential [10, §15.8]

$$\begin{aligned}
dg_{mn}(X)|_{dX\to Y} &= \sum_{k,l} \frac{\partial g_{mn}(X)}{\partial X_{kl}} Y_{kl} = \operatorname{tr}\left(\nabla g_{mn}(X)^T Y\right) &(186)\\
&= \sum_{k,l} \lim_{\Delta t \to 0} \frac{g_{mn}(X + \Delta t\, Y_{kl}\, e_k e_l^T) - g_{mn}(X)}{\Delta t} &(187)\\
&= \lim_{\Delta t \to 0} \frac{g_{mn}(X + \Delta t\, Y) - g_{mn}(X)}{\Delta t} &(188)\\
&= \left.\frac{d}{dt}\right|_{t=0} g_{mn}(X + t\, Y) &(189)
\end{aligned}$$

for $t \in \mathbb{R}$. Equation (188) is called the *Gateaux differential* [9, §A.5] [3, §D.2.1] that may be understood as the change in g_{mn} at X when the change in X is equal in *magnitude* and *direction* to Y which is assumed finite.[45] Hence the directional derivative,

[45] Although Y is a matrix, we may regard it as a vector in $\mathbb{R}^{KL}$.

$$
\overset{\rightarrow Y}{dg}(X) \triangleq \left.\begin{bmatrix} dg_{11}(X) & dg_{12}(X) & \cdots & dg_{1N}(X) \\ dg_{21}(X) & dg_{22}(X) & \cdots & dg_{2N}(X) \\ \vdots & \vdots & & \vdots \\ dg_{M1}(X) & dg_{M2}(X) & \cdots & dg_{MN}(X) \end{bmatrix}\right|_{dX\rightarrow Y} \in \mathbb{R}^{M\times N}
$$

$$
= \begin{bmatrix} \operatorname{tr}\left(\nabla g_{11}(X)^T Y\right) & \operatorname{tr}\left(\nabla g_{12}(X)^T Y\right) & \cdots & \operatorname{tr}\left(\nabla g_{1N}(X)^T Y\right) \\ \operatorname{tr}\left(\nabla g_{21}(X)^T Y\right) & \operatorname{tr}\left(\nabla g_{22}(X)^T Y\right) & \cdots & \operatorname{tr}\left(\nabla g_{2N}(X)^T Y\right) \\ \vdots & \vdots & & \vdots \\ \operatorname{tr}\left(\nabla g_{M1}(X)^T Y\right) & \operatorname{tr}\left(\nabla g_{M2}(X)^T Y\right) & \cdots & \operatorname{tr}\left(\nabla g_{MN}(X)^T Y\right) \end{bmatrix}
$$

$$
= \begin{bmatrix} \sum\limits_{k,l} \frac{\partial g_{11}(X)}{\partial X_{kl}} Y_{kl} & \sum\limits_{k,l} \frac{\partial g_{12}(X)}{\partial X_{kl}} Y_{kl} & \cdots & \sum\limits_{k,l} \frac{\partial g_{1N}(X)}{\partial X_{kl}} Y_{kl} \\ \sum\limits_{k,l} \frac{\partial g_{21}(X)}{\partial X_{kl}} Y_{kl} & \sum\limits_{k,l} \frac{\partial g_{22}(X)}{\partial X_{kl}} Y_{kl} & \cdots & \sum\limits_{k,l} \frac{\partial g_{2N}(X)}{\partial X_{kl}} Y_{kl} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{k,l} \frac{\partial g_{M1}(X)}{\partial X_{kl}} Y_{kl} & \sum\limits_{k,l} \frac{\partial g_{M2}(X)}{\partial X_{kl}} Y_{kl} & \cdots & \sum\limits_{k,l} \frac{\partial g_{MN}(X)}{\partial X_{kl}} Y_{kl} \end{bmatrix} \tag{190}
$$

from which it follows

$$
\overset{\rightarrow Y}{dg}(X) = \sum_{k,l} \frac{\partial g(X)}{\partial X_{kl}} Y_{kl} \tag{191}
$$

Yet for all $X \in \operatorname{dom} g$, any $Y \in \mathbb{R}^{K\times L}$, and some open interval of $t \in \mathbb{R}$,

$$
g(X + t\,Y) = g(X) \;+\; t\,\overset{\rightarrow Y}{dg}(X) \;+\; o(t^2) \tag{192}
$$

which is the first-order Taylor expansion about X. [12, §2.3.4] [9, §A.5] Differentiation with respect to t and subsequent t-zeroing isolates the second term of the expansion. Thus differentiating and zeroing $g(X + t\,Y)$ in t is an equivalent operation to individually differentiating and zeroing every entry $g_{mn}(X + t\,Y)$ as in (189). So the directional derivative becomes

$$
\overset{\rightarrow Y}{dg}(X) = \left.\frac{d}{dt}\right|_{t=0} g(X + t\,Y) \tag{193}
$$

[30, §2.1, §5.4.5] which is simplest.

A.2 Linear Matrix Inequality

When $\partial g(X)/\partial X_{kl}$ are instead fixed constant matrices in $\mathbb{S}^M$ and when $Y\in\mathbb{R}^{K\times L}$ is variable, (191) has the form from a class of matrix called *linear matrix inequality*. More generally,

$$G(y) \triangleq G_0 + \sum_{i=1}^{m} G_i\, y_i \succeq 0 \tag{194}$$

where $y\in\mathbb{R}^m$ is variable and the m fixed matrices $G_i\in\mathbb{S}^M$ are given, is a linear matrix inequality (LMI). [31] Hence the directional derivative is related to that matrix class.

A.3 Second directional derivative

By similar argument, it turns out that the second directional derivative is equally simple. Given $g(X):\mathbb{R}^{K\times L}\rightarrow\mathbb{R}^{M\times N}$,

$$\nabla\frac{\partial g_{mn}(X)}{\partial X_{kl}} = \frac{\partial \nabla g_{mn}(X)}{\partial X_{kl}} \triangleq \begin{bmatrix} \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{11}} & \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{12}} & \cdots & \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{1L}} \\ \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{21}} & \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{22}} & \cdots & \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{K1}} & \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{K2}} & \cdots & \frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\partial X_{KL}} \end{bmatrix} \in \mathbb{R}^{K\times L} \tag{195}$$

$$\begin{aligned} \nabla^2 g_{mn}(X) &\triangleq \begin{bmatrix} \nabla\frac{\partial g_{mn}(X)}{\partial X_{11}} & \nabla\frac{\partial g_{mn}(X)}{\partial X_{12}} & \cdots & \nabla\frac{\partial g_{mn}(X)}{\partial X_{1L}} \\ \nabla\frac{\partial g_{mn}(X)}{\partial X_{21}} & \nabla\frac{\partial g_{mn}(X)}{\partial X_{22}} & \cdots & \nabla\frac{\partial g_{mn}(X)}{\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \nabla\frac{\partial g_{mn}(X)}{\partial X_{K1}} & \nabla\frac{\partial g_{mn}(X)}{\partial X_{K2}} & \cdots & \nabla\frac{\partial g_{mn}(X)}{\partial X_{KL}} \end{bmatrix} \in \mathbb{R}^{K\times L\times K\times L} \\ &= \begin{bmatrix} \frac{\partial\nabla g_{mn}(X)}{\partial X_{11}} & \frac{\partial\nabla g_{mn}(X)}{\partial X_{12}} & \cdots & \frac{\partial\nabla g_{mn}(X)}{\partial X_{1L}} \\ \frac{\partial\nabla g_{mn}(X)}{\partial X_{21}} & \frac{\partial\nabla g_{mn}(X)}{\partial X_{22}} & \cdots & \frac{\partial\nabla g_{mn}(X)}{\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial\nabla g_{mn}(X)}{\partial X_{K1}} & \frac{\partial\nabla g_{mn}(X)}{\partial X_{K2}} & \cdots & \frac{\partial\nabla g_{mn}(X)}{\partial X_{KL}} \end{bmatrix} \end{aligned} \tag{196}$$

By rotating our perspective, we get several views of the second-order gradient.[46]

$$\nabla^2 g(X) = \begin{bmatrix} \nabla^2 g_{11}(X) & \nabla^2 g_{12}(X) & \cdots & \nabla^2 g_{1N}(X) \\ \nabla^2 g_{21}(X) & \nabla^2 g_{22}(X) & \cdots & \nabla^2 g_{2N}(X) \\ \vdots & \vdots & & \vdots \\ \nabla^2 g_{M1}(X) & \nabla^2 g_{M2}(X) & \cdots & \nabla^2 g_{MN}(X) \end{bmatrix} \in \mathbb{R}^{M\times N\times K\times L\times K\times L}$$

$$= \begin{bmatrix} \nabla\frac{\partial g(X)}{\partial X_{11}} & \nabla\frac{\partial g(X)}{\partial X_{12}} & \cdots & \nabla\frac{\partial g(X)}{\partial X_{1L}} \\ \nabla\frac{\partial g(X)}{\partial X_{21}} & \nabla\frac{\partial g(X)}{\partial X_{22}} & \cdots & \nabla\frac{\partial g(X)}{\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \nabla\frac{\partial g(X)}{\partial X_{K1}} & \nabla\frac{\partial g(X)}{\partial X_{K2}} & \cdots & \nabla\frac{\partial g(X)}{\partial X_{KL}} \end{bmatrix} \in \mathbb{R}^{K\times L\times M\times N\times K\times L} \tag{197}$$

$$= \begin{bmatrix} \frac{\partial\nabla g(X)}{\partial X_{11}} & \frac{\partial\nabla g(X)}{\partial X_{12}} & \cdots & \frac{\partial\nabla g(X)}{\partial X_{1L}} \\ \frac{\partial\nabla g(X)}{\partial X_{21}} & \frac{\partial\nabla g(X)}{\partial X_{22}} & \cdots & \frac{\partial\nabla g(X)}{\partial X_{2L}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial\nabla g(X)}{\partial X_{K1}} & \frac{\partial\nabla g(X)}{\partial X_{K2}} & \cdots & \frac{\partial\nabla g(X)}{\partial X_{KL}} \end{bmatrix} \in \mathbb{R}^{K\times L\times K\times L\times M\times N}$$

Assuming that the limits exist, we may state the partial derivative of the mn^{th} entry of g with respect to the kl^{th} and ij^{th} entries of X;

$$\frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\,\partial X_{ij}} = \tag{198}$$

$$\lim_{\Delta\tau,\Delta t\to 0} \frac{g_{mn}(X+\Delta t\, e_k e_l^T + \Delta\tau\, e_i e_j^T) - g_{mn}(X+\Delta t\, e_k e_l^T) - \big(g_{mn}(X+\Delta\tau\, e_i e_j^T) - g_{mn}(X)\big)}{\Delta\tau\,\Delta t}$$

Differentiating (185) and then scaling by Y_{ij},

$$\frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\,\partial X_{ij}} Y_{kl} Y_{ij} = \lim_{\Delta t\to 0} \frac{\partial g_{mn}(X+\Delta t\, Y_{kl}\, e_k e_l^T) - \partial g_{mn}(X)}{\partial X_{ij}\,\Delta t} Y_{ij} \tag{199}$$

$$= \lim_{\Delta\tau,\Delta t\to 0} \frac{g_{mn}(X+\Delta t\, Y_{kl}\, e_k e_l^T + \Delta\tau\, Y_{ij}\, e_i e_j^T) - g_{mn}(X+\Delta t\, Y_{kl}\, e_k e_l^T) - \big(g_{mn}(X+\Delta\tau\, Y_{ij}\, e_i e_j^T) - g_{mn}(X)\big)}{\Delta\tau\,\Delta t}$$

[46] When $M=N=L=1$, the second-order gradient is traditionally called the Hessian.

which can be proved by substitution of variables in (198). The mn^{th} second-order total differential due to Y is

$$d^2 g_{mn}(X)|_{dX\to Y} = \sum_{i,j}\sum_{k,l}\frac{\partial^2 g_{mn}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} = \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{mn}(X)^T Y\right)^T Y\right) \tag{200}$$

$$= \sum_{i,j}\lim_{\Delta t\to 0}\frac{\partial g_{mn}(X+\Delta t\,Y)-\partial g_{mn}(X)}{\partial X_{ij}\,\Delta t}Y_{ij} \tag{201}$$

$$= \lim_{\Delta t\to 0}\frac{g_{mn}(X+2\Delta t\,Y)-2g_{mn}(X+\Delta t\,Y)+g_{mn}(X)}{\Delta t^2} \tag{202}$$

$$= \left.\frac{d^2}{dt^2}\right|_{t=0} g_{mn}(X+t\,Y) \tag{203}$$

Although equality between (200) and (202) is not proven here, (202) is the second-order Gateaux differential (*confer* (188)); algebraically verifiable.[47] Hence the second directional derivative,

$$\overset{\to Y}{dg^2}(X) \triangleq \left.\begin{bmatrix} d^2g_{11}(X) & d^2g_{12}(X) & \cdots & d^2g_{1N}(X) \\ d^2g_{21}(X) & d^2g_{22}(X) & \cdots & d^2g_{2N}(X) \\ \vdots & \vdots & & \vdots \\ d^2g_{M1}(X) & d^2g_{M2}(X) & \cdots & d^2g_{MN}(X) \end{bmatrix}\right|_{dX\to Y} \in \mathbb{R}^{M\times N}$$

$$= \begin{bmatrix} \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{11}(X)^T Y\right)^T Y\right) & \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{12}(X)^T Y\right)^T Y\right) & \cdots & \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{1N}(X)^T Y\right)^T Y\right) \\ \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{21}(X)^T Y\right)^T Y\right) & \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{22}(X)^T Y\right)^T Y\right) & \cdots & \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{2N}(X)^T Y\right)^T Y\right) \\ \vdots & \vdots & & \vdots \\ \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{M1}(X)^T Y\right)^T Y\right) & \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{M2}(X)^T Y\right)^T Y\right) & \cdots & \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g_{MN}(X)^T Y\right)^T Y\right) \end{bmatrix}$$

$$= \begin{bmatrix} \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{11}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} & \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{12}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} & \cdots & \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{1N}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} \\ \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{21}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} & \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{22}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} & \cdots & \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{2N}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} \\ \vdots & \vdots & & \vdots \\ \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{M1}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} & \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{M2}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} & \cdots & \sum\limits_{i,j}\sum\limits_{k,l}\frac{\partial^2 g_{MN}(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} \end{bmatrix} \tag{204}$$

[47] *Mathematica* is capable of symbolic partial differentiation and limiting operations on a specified g. [32]

from which it follows

$$\overset{\to Y}{dg^2}(X) = \sum_{i,j}\sum_{k,l}\frac{\partial^2 g(X)}{\partial X_{kl}\,\partial X_{ij}}Y_{kl}Y_{ij} = \sum_{i,j}\frac{\partial}{\partial X_{ij}}\,\overset{\to Y}{dg}(X)\,Y_{ij} \tag{205}$$

Yet for all $X \in \operatorname{dom} g$, any $Y \in \mathbb{R}^{K\times L}$, and some open interval of $t \in \mathbb{R}$,

$$g(X + t\,Y) = g(X) \;+\; t\,\overset{\to Y}{dg}(X) \;+\; \frac{1}{2}t^2\,\overset{\to Y}{dg^2}(X) \;+\; o(t^3) \tag{206}$$

which is the second-order Taylor expansion about X. [9, §A.5][12, §2.3.4] Differentiation twice with respect to t and subsequent t-zeroing isolates the third term of the expansion. Thus differentiating and zeroing $g(X + t\,Y)$ in t is an equivalent operation to individually differentiating and zeroing every entry $g_{mn}(X + t\,Y)$ as in (203). So the second directional derivative becomes

$$\overset{\to Y}{dg^2}(X) = \left.\frac{d^2}{dt^2}\right|_{t=0} g(X + t\,Y) \tag{207}$$

[30, §2.1, §5.4.5] which is again simplest.

A.4 Taylor series

Series expansions of the differentiable matrix-valued function $g(X)$ of matrix argument, were given earlier in (192) and (206). The *mean value theorem* from Calculus insures a finite number of terms in the series. [10] Assuming $g(X)$ has continuous first, second, and third order gradients over the open set $\operatorname{dom} g$, then for all $X \in \operatorname{dom} g$, any $Y \in \mathbb{R}^{K\times L}$, and some open interval of $t \in \mathbb{R}$, the complete Taylor series is expressed,

$$g(X+t\,Y) = g(X) \;+\; t\,\overset{\to Y}{dg}(X) \;+\; \frac{1}{2!}t^2\,\overset{\to Y}{dg^2}(X) \;+\; \frac{1}{3!}t^3\,\overset{\to Y}{dg^3}(X) \;+\; o(t^4) \tag{208}$$

which is the third-order expansion about X.

In the case of a real-valued function $g(X) : \mathbb{R}^{K\times L} \to \mathbb{R}$, all the directional derivatives are in $\mathbb{R}$:

$$\overset{\to Y}{dg}(X) = \operatorname{tr}\left(\nabla g(X)^T Y\right) \tag{209}$$

$$\overset{\to Y}{dg^2}(X) = \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g(X)^T Y\right)^T Y\right) = \operatorname{tr}\left(\nabla \overset{\to Y}{dg}(X)^T Y\right) \tag{210}$$

$$\overset{\to Y}{dg^3}(X) = \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla \operatorname{tr}\left(\nabla g(X)^T Y\right)^T Y\right)^T Y\right) = \operatorname{tr}\left(\nabla \overset{\to Y}{dg^2}(X)^T Y\right) \tag{211}$$

In the case $g(X) : \mathbb{R}^K \to \mathbb{R}$ has vector argument, they simplify:

$$\overset{\to Y}{dg}(X) = \nabla g(X)^T Y \tag{212}$$

$$\overset{\to Y}{dg^2}(X) = Y^T \nabla^2 g(X) Y \tag{213}$$

$$\overset{\to Y}{dg^3}(X) = \nabla\left(Y^T \nabla^2 g(X) Y\right)^T Y \tag{214}$$

and so on, where the symmetric second-order gradient matrix $\nabla^2 g(X)$ is called the *Hessian* while its transpose is known as the *Jacobian*.

A.5 Correspondence of gradient to derivative

From the equations for directional derivative, we can derive a relationship between the gradient with respect to matrix X and the derivative with respect to real variable t: Removing from (193) the evaluation at $t=0$ (rigorously justifiable), we find an expression for the directional derivative of $g(X)$ in the direction Y evaluated along a line, parameterized by t, in direction Y;

$$\overset{\rightarrow Y}{dg}(X+t\,Y) = \frac{d}{dt}g(X+t\,Y) \tag{215}$$

In the important case of a real-valued function $g(X) : \mathbb{R}^{K\times L} \rightarrow \mathbb{R}$ and the first directional derivative, from (209) we have, therefore,

$$\frac{d}{dt}g(X+t\,Y) = \operatorname{tr}\left(\nabla_X g(X+t\,Y)^T Y\right) \tag{216}$$

Likewise removing the evaluation at $t=0$ from (207),

$$\overset{\rightarrow Y}{dg^2}(X+t\,Y) = \frac{d^2}{dt^2}g(X+t\,Y) \tag{217}$$

we can find a similar relationship between the second-order gradient and the second derivative: From (213), the case where $g(X) : \mathbb{R}^{K} \rightarrow \mathbb{R}$ has vector argument,

$$\frac{d^2}{dt^2}g(X+t\,Y) = Y^T \nabla_X^2 g(X+t\,Y)Y \tag{218}$$

Example. $f(X) = w^T X^T X w,\ X \in \mathbb{R}^{K\times L},\ w \in \mathbb{R}^L$. Applying (216),

$$\frac{d}{dt}w^T(X+t\,Y)^T(X+t\,Y)w \;=\; w^T(X^TY + Y^TX + 2t\,Y^TY)w \tag{219}$$

which (using the tables in Appendix H) is equivalent to

$$\operatorname{tr}\left(\nabla_X f(X+t\,Y)^T Y\right) \;=\; \operatorname{tr}\left(2ww^T(X^T + t\,Y^T)Y\right) \tag{220}$$

$$\;=\; w^T(2X^TY + 2t\,Y^TY)w \tag{221}$$

because $w^T(X^TY + Y^TX)w = 2w^TX^TYw$.

For this particular example, it would be easy to extract $\nabla f(X)$ from (219) knowing (216); suggesting an alternate method to find the gradient.

B $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ nesting

From (98) observe that $T = -V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow 3}$. In fact, the leading principal submatrices of $-V_{\mathcal{N}}^T D V_{\mathcal{N}}$ form a nested sequence, nested by inclusion, whose members are each positive semidefinite [6][8][5] and have the same form as T; *id est*,

$$\begin{aligned}
\{-V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow 2} &= [d_{12}] \in \mathbb{S}_+\,, && \text{(a)}\\
-V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow 3} &= \begin{bmatrix} d_{12} & \frac{1}{2}(d_{12}+d_{13}-d_{23}) \\ \frac{1}{2}(d_{12}+d_{13}-d_{23}) & d_{13} \end{bmatrix} = T \in \mathbb{S}_+^2\,, && \text{(b)}\\
-V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow 4} &= \begin{bmatrix} d_{12} & \frac{1}{2}(d_{12}+d_{13}-d_{23}) & \frac{1}{2}(d_{12}+d_{14}-d_{24}) \\ \frac{1}{2}(d_{12}+d_{13}-d_{23}) & d_{13} & \frac{1}{2}(d_{13}+d_{14}-d_{34}) \\ \frac{1}{2}(d_{12}+d_{14}-d_{24}) & \frac{1}{2}(d_{13}+d_{14}-d_{34}) & d_{14} \end{bmatrix}, && \text{(c)}\\
&\ \vdots \\
-V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow i} &= \begin{bmatrix} -V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow i-1} & \nu(i) \\ \\ \nu^T(i) & d_{1i} \end{bmatrix} \in \mathbb{S}_+^{i-1}, && \text{(d)}\\
&\ \vdots \\
-V_{\mathcal{N}}^T D V_{\mathcal{N}} &= \begin{bmatrix} -V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow N-1} & \nu(N) \\ \\ \nu^T(N) & d_{1N} \end{bmatrix} \in \mathbb{S}_+^{N-1} && \text{(e)}\\
\}
\end{aligned} \tag{222}$$

where

$$\nu(i) \triangleq \frac{1}{2}\begin{bmatrix} d_{12}+d_{1i}-d_{2i} \\ d_{13}+d_{1i}-d_{3i} \\ \vdots \\ d_{1,i-1}+d_{1i}-d_{i-1,i} \end{bmatrix} \in \mathbb{R}^{i-2}, \quad i > 2 \tag{223}$$

and where $-V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N\leftarrow 1}$ is undefined. Bordered symmetric matrices in the form (222d) are known to have *intertwined* [8, §6.4] (or *interlaced* [5, §4.3]) eigenvalues; (*confer* 5.5.1) that means, for the particular submatrices (222a) and (222b),

$$\sigma_2 \leq d_{12} \leq \sigma_1 \tag{224}$$

where d_{12} is the eigenvalue of the submatrix (222a), and σ_1, σ_2 are the eigenvalues of T (222b)(98). Intertwining in (224) predicts that should d_{12} become zero, then σ_2 must go to zero.[48] The eigenvalues are similarly intertwined for submatrices (222b) and (222c);

$$\gamma_3 \leq \sigma_2 \leq \gamma_2 \leq \sigma_1 \leq \gamma_1 \tag{225}$$

where $\gamma_1, \gamma_2, \gamma_3$ are the eigenvalues of submatrix (222c). Intertwining likewise predicts that should σ_2 become zero (a possibility revealed in Appendix C.1), then γ_3 must go to zero. Combining our results so far for $N = 2, 3, 4$: (224)(225)

$$\gamma_3 \leq \sigma_2 \leq d_{12} \leq \sigma_1 \leq \gamma_1 \tag{226}$$

The preceding logic extends by induction through the remaining members of the sequence (222).

[48]If d_{12} were zero, then eigenvalue σ_2 becomes zero (100) because d_{13} must then be equal to d_{23}; *id est*, $d_{12} = 0 \Leftrightarrow x_1 = x_2$. (26)

C Schur Complement

Consider the *Schur complement*: Given $A^T = A$ and $C^T = C$, then

$$G = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0 \qquad \Big| \qquad G = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succ 0$$

$$\begin{array}{ll|l} \Leftrightarrow\ A \succeq 0, \quad B^T(I - AA^\dagger) = 0, \quad C - B^T A^\dagger B \succeq 0 & & \Leftrightarrow\ A \succ 0, \quad C - B^T A^{-1} B \succ 0 \\ \Leftrightarrow\ C \succeq 0, \quad B(I - CC^\dagger) = 0, \quad A - BC^\dagger B^T \succeq 0 & & \Leftrightarrow\ C \succ 0, \quad A - BC^{-1} B^T \succ 0 \end{array} \tag{227}$$

where $A^\dagger$ denotes the Moore-Penrose (pseudo-)inverse, and where $C - B^T A^{-1} B$ is called the Schur complement of A in G, while $A - BC^{-1}B^T$ is the Schur complement of C in G. [31][33, §4.8]

C.1 Tightening the triangle inequality

For example, from Appendix B we identify

$$G \triangleq -V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N \leftarrow 4} \tag{228}$$

$$A \triangleq T = -V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N \leftarrow 3} \tag{229}$$

both positive semidefinite by assumption, $B = \nu(4)$ defined in (223), and $C = d_{14}$. Using the latter non-strict form of (227), $C \geq 0$ by assumption (§5.1) and $CC^\dagger = I$. So by theory of *positive semidefinite ordering of eigenvalues* [5, §7.7, prob.1]

$$-V_{\mathcal{N}}^T D V_{\mathcal{N}}|_{N \leftarrow 4} \succeq 0 \ \Leftrightarrow\ T \succeq d_{14}^{-1} \nu(4) \nu^T(4) \ \Rightarrow \begin{cases} \sigma_1 \geq d_{14}^{-1} \|\nu(4)\|^2 \\ \sigma_2 \geq 0 \end{cases} \tag{230}$$

where $\{d_{14}^{-1}\|\nu(4)\|^2, 0\}$ are the eigenvalues of $d_{14}^{-1}\nu(4)\nu^T(4)$ and σ_1, σ_2 are the eigenvalues of T.

C.1.1 Example revisitation

Applying the inequality for σ_1 in (230) to the example in §3.2, Figure 2, the lower bound on $\sqrt{d_{14}}$, 1.236 in (25), is tightened to 1.289. The correct value of $\sqrt{d_{14}}$ to three significant figures is 1.414.

D The V matrices

It will become convenient to define a matrix V that arises naturally as a consequence of translating the geometric center to the origin. Instead of $X-\alpha_g\mathbf{1}^T$ we may write XV; *viz.*,

$$X-\alpha_g\mathbf{1}^T \;=\; X-\frac{1}{N}X\mathbf{1}\mathbf{1}^T \;=\; X(I-\frac{1}{N}\mathbf{1}\mathbf{1}^T) \;=\; XV\in\mathbb{R}^{n\times N} \tag{231}$$

Obviously,

$$V=V^T=I-\frac{1}{N}\mathbf{1}\mathbf{1}^T\in\mathbb{R}^{N\times N} \tag{232}$$

which is an elementary matrix. (App. D.1)

$$V_w=\begin{bmatrix} \frac{-1}{\sqrt{N}} & \frac{-1}{\sqrt{N}} & \cdots & \frac{-1}{\sqrt{N}} \\ 1+\frac{-1}{N+\sqrt{N}} & \frac{-1}{N+\sqrt{N}} & \cdots & \frac{-1}{N+\sqrt{N}} \\ \frac{-1}{N+\sqrt{N}} & \ddots & \cdots & \frac{-1}{N+\sqrt{N}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-1}{N+\sqrt{N}} & \frac{-1}{N+\sqrt{N}} & \cdots & 1+\frac{-1}{N+\sqrt{N}} \end{bmatrix}\in\mathbb{R}^{N\times N-1} \tag{233}$$

Thus far, we have three auxiliary matrices V, $V_{\mathcal{N}}$, and V_w that share some common attributes listed in Table D. Full-rank skinny matrix V_w is distinguished in so far as V can be expressed in terms of it; [24]

$$V=V_wV_w^T \tag{234}$$

but $V_w^TV_w=I$.[49]

Table D

	$\dim V_.$	$\operatorname{rank} V_.$	$\mathcal{R}(V_.)$	$\mathcal{N}(V_.^T)$	$V_.^TV_.$	$V_.V_.^T$
V	$N\times N$	$N-1$	$\mathcal{N}(\mathbf{1}^T)$	$\mathcal{R}(\mathbf{1})$	V	V
$V_{\mathcal{N}}$	$N\times(N-1)$	$N-1$	$\mathcal{N}(\mathbf{1}^T)$	$\mathcal{R}(\mathbf{1})$	$\frac{1}{2}(I+\mathbf{1}\mathbf{1}^T)$	$\frac{1}{2}\begin{bmatrix} N-1 & -\mathbf{1}^T \\ -\mathbf{1} & I \end{bmatrix}$
V_w	$N\times(N-1)$	$N-1$	$\mathcal{N}(\mathbf{1}^T)$	$\mathcal{R}(\mathbf{1})$	I	V

[49] The fact that V can be expressed as in (234) shows that V is a projection matrix since all projection matrices P can be expressed in the form $P=QQ^T$ where Q has orthonormal columns.

D.1 Elementary matrix

A matrix of the form
$$E = I - \zeta uv^T \tag{235}$$
where $\zeta \in \mathbb{R}$ and where u and v are vectors of the same dimension, is called an *elementary* matrix, or a *rank-one modification of the identity.* [33] Any elementary matrix in $\mathbb{R}^{N\times N}$ has $N-1$ eigenvalues equal to 1, corresponding to eigenvectors that span $\mathcal{N}(uv^T)$. The remaining eigenvalue is $1-\zeta v^T u$ having corresponding eigenvector u.

$$E^{-1} = I - \beta uv^T \tag{236}$$
where $\beta = \zeta \,/ \bigl(\zeta u^T v - 1\bigr)$.

For the particular elementary matrix V, the remaining eigenvalue equals 0. Because $V = V^T$ is diagonalizable, the number of zero eigenvalues must be equal to $\dim \mathcal{N}(V^T) = 1$, and because $V^T\mathbf{1} = 0$, then $\mathcal{N}(V^T) = \mathcal{R}(\mathbf{1})$. Because $V = V^T$ and $V^2 = V$, the elementary matrix V is a projection matrix on its range $\mathcal{R}(V) = \mathcal{N}(\mathbf{1}^T)$ having dimension $N-1$.

$$\frac{1}{2}\,\mathrm{tr}(-V^T D V) \;=\; \frac{1}{2}\,\mathrm{tr}(-V_{\mathcal{N}}^{\dagger} D V_{\mathcal{N}}) \;=\; \frac{1}{2N}\mathbf{1}^T D\,\mathbf{1} \;=\; \frac{1}{2N}\,\mathrm{tr}(\mathbf{1}\mathbf{1}^T D) \;=\; \frac{1}{N}\sum_{i,j} d_{ij} \tag{237}$$

Any matrix $A \in \mathbb{S}^N$ that can be expressed
$$A = k_1 I + k_2 \mathbf{1}\mathbf{1}^T \tag{238}$$
for any $k_1, k_2 \in \mathbb{R}$, will have $\mathrm{tr}(A^T D)$ proportional to $\sum d_{ij}$. If k_1 is $1-\rho$ while k_2 equals $\rho \in \mathbb{R}$, then as long as $-1/(N\text{–}1) < \rho < 1$, all the eigenvalues of A are guaranteed to be positive and therefore A is guaranteed to be positive definite. [Reznik]

D.2 $V_{\mathcal{N}}$

1. $V_{\mathcal{N}}^T \mathbf{1} = 0$
2. $V_{\mathcal{N}}^{\dagger} = \sqrt{2}\left[-\frac{1}{N}\mathbf{1} \quad I - \frac{1}{N}\mathbf{1}\mathbf{1}^T\right] \in \mathbb{R}^{N-1\times N}$
3. $V_{\mathcal{N}}^{\dagger} V_{\mathcal{N}} = I$

4. $V_{\mathcal{N}} V_{\mathcal{N}}^{\dagger} \triangleq V^T = V$

5. $[V_{\mathcal{N}} \quad \frac{1}{\sqrt{2}}\mathbf{1}]^{-1} = \begin{bmatrix} V_{\mathcal{N}}^{\dagger} \\ \frac{\sqrt{2}}{N}\mathbf{1}^T \end{bmatrix}$

6. $V_{\mathcal{N}}^{\dagger}\mathbf{1} = 0$

7. $(I - e_1\mathbf{1}^T)V_{\mathcal{N}} = \begin{bmatrix} \mathbf{0} & \sqrt{2}V_{\mathcal{N}} \end{bmatrix} V_{\mathcal{N}} = V_{\mathcal{N}}$

E loose ends

Minimization of d from the sequence $\{a, b, c, d\}$ where $a \leq b \leq c \leq d$ is equivalent to $\min \|\{a, b, c, d\}\|_\infty$.

F Principal submatrix

Here we prove the *principal submatrix theorem* in §5.2.

A principal submatrix of a square matrix $A \in \mathbb{R}^{M\times M}$ is formed by first making a rectangular matrix S composed of any subset of columns from the identity matrix. For example, for $M=4$,

$$S \triangleq [e_1\ e_2] \in \mathbb{R}^{M\times 2} \tag{239}$$

where e_i identifies the i^{th} standard basis vector. Then S is applied

$$A(1:2, 1:2) = S^T A S \in \mathbb{R}^{2\times 2} \tag{240}$$

using Golub's notation for extracting part of a matrix [6] (identical to MATLAB), yielding the leading principal 2×2 submatrix in this example.

There are

1. $M!/(1!(M-1)!)$ principal 1×1 submatrices,
2. $M!/(2!(M-2)!)$ principal 2×2 submatrices,
3. $\vdots$

and so on.

Any symmetric matrix A can be diagonalized by its eigenvectors q_i and eigenvalues λ_i. Hence,

$$y^T A y = \sum_i \lambda_i (q_i^T y)^2$$

It is well known that A is positive semidefinite iff all $\lambda_i \geq 0$. All the principal submatrices of A are formed by loading y with various patterns of ones and zeros. [8, §6.3] The eigen-decomposition demands that any principal submatrix selected by y be positive semidefinite whenever A is.

G Volume of a convex polyhedron

Strang [8] pg.212

No method is known for computing the volume of a general convex polyhedron [34, p.173].

Algorithm: many out there on internet.

Volume is a concept in $\mathbb{R}^3$; in higher dimensions called "content". The volume of a tetrahedron is 1/3 the product of the area of the base times the height. [17] Volume of any simplex is proportional to the Cayley-Menger determinant.

How to know that a particular distance refers to an interior point? Find the list. If any list member can be expressed as a convex combination of the others, then it is an interior point. [2]

Semidefinite program to minimize volume as in §6.1.

Volume of a polyhedron is log concave function of halfspace description parameter. [Reader, pg.78] Volume is log of piecewise linear concave function. -Boyd

H Derivatives and gradients

Citations...Graham, Chen, Brooke,

- $a, b, x, y \in \mathbb{R}^k$, $A, B, X, Y \in \mathbb{R}^{p\times k}$, $\mu \in \mathbb{R}$ unless otherwise noted.

- x^μ means $\delta(\delta(x)^\mu)$ for $\mu \in \mathbb{R}$; *id est*, element-wise exponentiation.
  δ is the main diagonal operator (30). $x^0 \triangleq \mathbf{1}$, $X^0 \triangleq I$.

- $\frac{d}{dx} \triangleq \begin{bmatrix} \frac{d}{dx_1} \\ \vdots \\ \frac{d}{dx_k} \end{bmatrix}$, $\log x$, $\sin x$, *etcetera*, are maps $f : \mathbb{R}^k \rightarrow \mathbb{R}^k$ that maintain dimension.

- The symmetric part of square matrix A is $(A + A^T)/2$;
  The antisymmetric part is $(A - A^T)/2$.

- When algebraically proving results for symmetric matrices, it is critical to take the gradient with respect to the *non*symmetric matrix first and then substitute the symmetric elements.

Table 1: **Algebraic**

$\nabla_x \, Ax + b = A^T$	
$\nabla_x \, x^T A + b^T = A$	
$\nabla_x \, x^T A x + 2x^T B y + y^T C y = (A + A^T)x + 2By$	
$\nabla_x \, w^T x^T x w = 2xw^T w$	$\nabla_X \, w^T X^T X w = 2Xww^T$
$\nabla_x \, w^T x x^T w = 2ww^T x$	$\nabla_X \, w^T X X^T w = 2ww^T X$
	$\frac{d}{dt}(X + t\,Y) = Y$
	$\frac{d}{dt} B^T (X + t\,Y)^{-1} A = -B^T (X + t\,Y)^{-1} Y (X + t\,Y)^{-1} A$
	$\frac{d^2}{dt^2} B^T (X + t\,Y)^{-1} A = 2B^T (X + t\,Y)^{-1} Y (X + t\,Y)^{-1} Y (X + t\,Y)^{-1} A$
	$\frac{d}{dt}(X + t\,Y)^T A (X + t\,Y) = Y^T A X + X^T A Y + 2t\,Y^T A Y$

Table 2: **Trace**

$\nabla_x\, x^T y = \nabla_x\, y^T x = y$	$\nabla_X \operatorname{tr}(X^T Y) = \nabla_X \operatorname{tr}(Y^T X) = \nabla_X \operatorname{tr}(X Y^T) = Y$
	$\nabla_X \operatorname{tr}(Y^T X X^T Y) = 2YY^T X$
	$\nabla_X \operatorname{tr}(Y^T X^T X Y) = 2XYY^T$
	$\nabla_X\, a^T X b = \nabla_X\, b^T X^T a = \nabla_X \operatorname{tr}(b^T X^T a) = \nabla_X \operatorname{tr}(X^T a b^T) = a b^T$
$\nabla_x\, \mu x = \mu I$	$\nabla_X \operatorname{tr} \mu X = \nabla_X\, \mu \operatorname{tr} X = \mu I$
$\frac{d}{dx} x^{-1} = -x^{-2}$	$\nabla_X \operatorname{tr} X^{-1} = -X^{-2T}$
$\frac{d}{dx} x^{\mu} = \mu x^{\mu-1}$	$\nabla_X \operatorname{tr} X^{\mu} = \mu X^{(\mu-1)T}$
	$\frac{d}{dt} \operatorname{tr}(X + t\,Y) = \operatorname{tr} Y$
	$\frac{d}{dt} \operatorname{tr} g(X + t\,Y) = \operatorname{tr} \frac{d}{dt} g(X + t\,Y)$

Table 3: **Log determinant**.
Argument of det is positive definite but not necessarily symmetric.

$\frac{d}{dx}\log x = x^{-1}\,, \qquad x \succ 0$	$\nabla_X \log\det X = X^{-T}$
$\frac{d}{dx}\log x^{-1} = -x^{-1}\,, \quad x \succ 0$	$\nabla_X \log\det X^{-1} = -X^{-T}$
$\frac{d}{dx}\log x^{\mu} = \mu x^{-1}\,, \qquad x \succ 0$	$\nabla_X \log\det X^{\mu} = \mu X^{-T}$
	$\frac{d}{dt}\log\det(X+t\,Y) = \operatorname{tr}\left((X+t\,Y)^{-1}Y\right)$

Table 4: **Determinant**

$$\nabla_X \det X = \det(X) X^{-T}$$

$$\nabla_X \det X^{-1} = -\det(X^{-1}) X^{-T}$$

$$\nabla_X \det X^{\mu} = \mu \det(X^{\mu}) X^{-T}$$

$$\nabla_X (\det X)^{\mu \text{ or } 1/n} =$$

$$\tfrac{d}{dt} \det(X + t\,Y) = \det(X + t\,Y) \operatorname{tr}\left((X + t\,Y)^{-1} Y\right)$$

Table 5: **Exponential**

$$\nabla_X \operatorname{tr} e^{YX} = e^{YX} Y, \qquad X, Y \in \mathbb{S}^p$$

$$\nabla_X e^{\operatorname{tr}(Y^T X)} =$$

$$\tfrac{d}{dt} e^{tY} = e^{tY} Y = Y e^{tY}$$

$$\tfrac{d}{dt} e^{X + t\,Y} = e^{X + t\,Y} Y = Y e^{X + t\,Y}, \qquad XY = YX$$

I $-\mathcal{D}(X)$ Quasiconvexity -WRONG-

To establish quasiconvexity of $-\mathcal{D}(X)$ (29)[50] requires convexity of its sublevel set $\mathcal{L}_\nu$ for all $\nu\in\mathbb{R}$ (§2.2.1);

$$\mathcal{L}_\nu = \left\{X\in\mathbb{R}^{n\times N} \mid -\mathcal{D}(X) \preceq \nu I\right\} \tag{241}$$

$\mathcal{L}_\nu$ is empty (yet convex [2, §2]) when $\nu<0$ because $-\mathcal{D}(X)$ is indefinite (§5.5). When $\nu\geq 0$, $\mathcal{L}_\nu \ni 0$.

Definition. *Subspace* $\mathbb{S}^N_{\delta\kappa}$. We introduce the subspace of all symmetric matrices in $\mathbb{R}^{N\times N}$ having constant main diagonal: *confer* (9)

$$\mathbb{S}^N_\delta \subset \mathbb{S}^N_{\delta\kappa} \triangleq \left\{A\in\mathbb{S}^N \mid \delta(A)=\kappa\mathbf{1},\ \kappa\in\mathbb{R}\right\} \subseteq \mathbb{S}^N \tag{242}$$

$\mathcal{D}(X)\in\mathbb{S}^N_\delta$ (28) while $\nu I+\mathcal{D}(X)\in\mathbb{S}^N_{\delta\kappa}$. For $\mathbb{S}^N_{\delta\kappa}$, there exists a nontrivial intersection with the positive semidefinite cone; $\mathbb{S}^N_{\delta\kappa}\cap\mathbb{S}^N_+ \neq \mathbf{0}$ or $\emptyset$.

Make an affine function $f_\nu(Y):\mathbb{S}^N\to\mathbb{S}^N_{\delta\kappa}$ from the EDM definition (29);

$$\begin{aligned} f_\nu(Y) &\triangleq \nu I+\mathcal{D}(X)|_{Y\leftarrow X^TX} \\ &= \nu I+\delta(Y)\mathbf{1}^T+\mathbf{1}\delta^T(Y)-2Y \end{aligned} \tag{243}$$

(Like $\mathcal{D}(D)$ (54), $f_\nu(Y)$ on $\mathbb{S}^N_\delta$ is injective.) In terms of f_ν, the sublevel set for $-\mathcal{D}(X)$ is

$$\mathcal{L}_\nu = \left\{X\in\mathbb{R}^{n\times N} \mid Y=X^TX,\ f_\nu(Y)\succeq 0\right\} \tag{244}$$

Because $X^TX\succeq 0$ for all $X\in\mathbb{R}^{n\times N}$ [8, §6.3], we can define a convex set closely related to $\mathcal{L}_\nu$;

$$\mathcal{Y}_\nu \triangleq \{Y\in\mathbb{S}^N_+ \mid f_\nu(Y)\succeq 0\} \subseteq \mathbb{S}^N_+ \tag{245}$$

$\mathcal{Y}_\nu$ is convex because it is the intersection of the positive semidefinite cone with the inverse image of that same cone $\mathbb{S}^N_+$ under affine f_ν; *id est*, the intersection of two convex sets [2, §2], having properties:

[50] $-\mathcal{D}(X)$ is continuous but not a *convex* function of matrix X because $-\nabla^2_t\mathcal{D}(X_o+tY_o)$ evaluated at all points on any line X_o+tY_o through its domain [1, §3] is indefinite, where $t\in\mathbb{R}$ is variable, and X_o and $Y_o\in\mathbb{R}^{n\times N}$.

- When $\nu \geq 0$, $\mathcal{Y}_\nu \ni 0$.
- $\lim\limits_{\nu\to\infty} \mathcal{Y}_\nu = \mathbb{S}_+^N$
- $\mathcal{Y}_\nu$ is empty when $\nu < 0$ because nonnegativity of the main diagonal is a necessary condition for symmetric positive semidefinite matrices. [6, §4.2.8]

The level set $\mathcal{L}_\nu$ is the inverse image of $\mathcal{Y}_\nu$ under a matrix-valued convex quadratic function h; *confer* (244),

$$\mathcal{L}_\nu \;=\; h^{-1}(\mathcal{Y}_\nu) \mid h(X) = X^T X \tag{246}$$

Because $\mathcal{Y}_\nu$ is a convex set, if $\mathcal{Y}_\nu \times \mathcal{L}_\nu$ were a convex set then $\mathcal{L}_\nu$ would be convex by implication. [1, §2] We wonder if $\mathcal{L}_\nu$ is convex because $\mathcal{Y}_\nu$ is convex and contains the *minimum element* [1, §2] (zero) with respect to the range $\mathbb{S}_+^N$ of quadratic function h.
Trying to show $z^T\mathcal{Y}_\nu z \times (\mathcal{L}_\nu \cap (X_o + t\,Y_o))$ *is convex for any* X_o*,* Y_o*, and* z*.*

To show that $\mathcal{L}_\nu$ is convex, we simultaneously invoke the scalar-definition of convex functions and the line theorem: (§2.2) For any X_o and $Y_o \in \mathbb{R}^{n\times N}$, every $z \in \mathbb{R}^n$, and variable $t \in \mathbb{R}$,

$$z^T h(X_o + t\,Y_o) z \;=\; z^T X_o^T X_o\, z + t\, z^T(X_o^T Y_o + Y_o^T X_o) z + t^2 z^T Y_o^T Y_o\, z \tag{247}$$

is convex and quadratic in t having minimum value $h_t^\star(z)$;

$$0 \;\leq\; h_t^\star(z) \;\triangleq\; z^T X_o^T X_o\, z - \frac{\left(z^T(X_o^T Y_o + Y_o^T X_o) z\right)^2}{4\, z^T Y_o^T Y_o\, z} \;\leq\; z^T h(X_o + t\,Y_o) z \tag{248}$$

$z^T h(X_o + tY_o) z$ is a two-to-one map from $\mathbb{R}$ to $\mathbb{R}_+$;

$$z^T h(X_o + t\,Y_o) z \;:\; \pm\, t - \frac{z^T(X_o^T Y_o + Y_o^T X_o) z}{2\, z^T Y_o^T Y_o\, z} \;\rightarrow\; h_t^\star(z) + t^2 z^T Y_o^T Y_o\, z \tag{249}$$

The lower bound 0 is the minimum element of $z^T\mathcal{Y}_\nu z$, with respect to $\mathbb{R}_+$, and tight; for example when $X_o = 0$, then $h_t^\star(z) = 0$ while

$$z^T h(t\,Y_o) z \;:\; \pm t \;\rightarrow\; t^2 z^T Y_o^T Y_o\, z \tag{250}$$

achieves the lower bound at $t = 0$. If for every z, $z^T\mathcal{Y}_\nu z$ is a convex set containing the zero element, then from (250) the intersection of its inverse image under $z^T h z$ with any line must be connected.

J Acknowledgements

Boyd, Smith, Osgood, WinEdt, MiKTEX, *Mathematica*,

References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* `http://www.stanford.edu/~boyd/cvxbook`

[2] R. Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, 1997. First published in 1970.

[3] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of Convex Analysis.* Springer-Verlag, 2001.

[4] David G. Luenberger. *Optimization by Vector Space Methods.* Wiley, 1969.

[5] Roger A. Horn and Charles R. Johnson. *Matrix Analysis.* Cambridge University Press, 1987.

[6] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* Johns Hopkins, third edition, 1996.

[7] Michael W. Trosset. Distance matrix completion by numerical optimization. *Computational Optimization and Applications*, 17(1):11–22, October 2000.

[8] Gilbert Strang. *Linear Algebra and its Applications.* Harcourt Brace, third edition, 1988.

[9] Dimitri P. Bertsekas. *Nonlinear Programming.* Athena Scientific, second edition, 1999.

[10] George B. Thomas, Jr. *Calculus and Analytic Geometry.* Addison-Wesley, fourth edition, 1972.

[11] Gilbert Strang. *Calculus.* Wellesley-Cambridge Press, 1992.

[12] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical Optimization.* Academic Press, 1999.

[13] Erwin Kreyszig. *Introductory Functional Analysis with Applications.* Wiley, 1989.

[14] Jerrold E. Marsden and Michael J. Hoffman. *Elementary Classical Analysis.* Freeman, second edition, 1995.

[15] William Wooton, Edwin F. Beckenbach, and Frank J. Fleming. *Modern Analytic Geometry*. Houghton Mifflin, 1975.

[16] John B. Conway. *A Course in Functional Analysis*. Springer-Verlag, second edition, 1990.

[17] Paul J. Kelly and Norman E. Ladd. *Geometry*. Scott, Foresman and Company, 1965.

[18] Eric W. Weisstein. Mathworld – *A Wolfram Web Resource*. `http://mathworld.wolfram.com/search`

[19] D. Avis and K. Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Discrete and Computational Geometry*, 8:295–313, 1992.

[20] John Clifford Gower. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra and its Applications*, 67:81–97, 1985. `http://www.convexoptimization.com/TOOLS/Gower1.pdf`

[21] Isaac J. Schoenberg. Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert". *Annals of Mathematics*, 36(3):724–732, July 1935. `http://www.convexoptimization.com/TOOLS/Schoenberg2.pdf`

[22] George B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, 1998.

[23] Jung Rye Lee. The law of cosines in a tetrahedron. *Journal of the Korea Society of Mathematical Education Series B: The Pure and Applied Mathematics*, 4(1):1–6, 1997.

[24] Abdo Y. Alfakih, Amir Khandani, and Henry Wolkowicz. Solving Euclidean distance matrix completion problems via semidefinite programming. *Computational Optimization and Applications*, 12(1):13–30, January 1999. `http://citeseer.ist.psu.edu/alfakih97solving.html`

[25] Martin Vetterli and Jelena Kovačević. *Wavelets and Subband Coding*. Prentice-Hall, 1995.

[26] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, 1989.

[27] Boaz Porat. *A Course in Digital Signal Processing*. Wiley, 1997.

[28] Robert M. Gray and Joseph W. Goodman. *Fourier Transforms, An Introduction for Engineers*. Kluwer Academic Publishers, 1995.

[29] Graham A. Wright. An introduction to magnetic resonance. `www.sunnybrook.utoronto.ca:8080/~gawright/main_mr.html` , 2002.

[30] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.

[31] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[32] Stephen Wolfram. *The Mathematica Book*. Wolfram Media/Cambridge University Press, third edition, 1996.

[33] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Numerical Linear Algebra and Optimization*, volume 1. Addison-Wesley, 1991.

[34] C. S. Ogilvy. *Excursions in Geometry*. Dover, 1990. Citation: *Proceedings of the CUPM Geometry Conference*, Mathematical Association of America, No.16 (1967), p.21.