Chapter 2

Convex geometry

Convexity has an immensely rich structure and numerous applications. On the other hand, almost every "convex" idea can be explained by a two-dimensional picture.

−Alexander Barvinok [27, p.vii]

We study convex geometry because it is the easiest of geometries. For that reason, much of a practitioner's energy is expended seeking invertible transformation of problematic sets to convex ones.

As convex geometry and linear algebra are inextricably bonded by linear inequality (*asymmetry*), we provide much background material on linear algebra (especially in the appendices) although a reader is assumed comfortable with [348] [350] [218] or any other intermediate-level text. The essential references to convex analysis are [215] [325]. The reader is referred to [347] [27] [410] [43] [63] [322] [377] for a comprehensive treatment of convexity. There is relatively less published pertaining to convex matrix-valued functions. [231] [219, §6.6] [312]

2.1 Convex set

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

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

Under that defining condition on μ, the linear sum in (1) is called a *convex combination* of Y and Z. If Y and Z are points in real finite-dimensional Euclidean *vector space* [243] [419] $\mathbb{R}^n$ or $\mathbb{R}^{m \times n}$ (matrices), then (1) represents the closed line segment joining them. Line segments are thereby convex sets; $\mathcal{C}$ is convex iff the line segment connecting any two points in $\mathcal{C}$ is itself in $\mathcal{C}$. Apparent from this definition: a convex set is a connected set. [274, §3.4, §3.5] [43, p.2] A convex set can, but does not necessarily, contain the *origin* $\mathbf{0}$.

An *ellipsoid* centered at $x = a$ (Figure **15** p.39), given matrix $C \in \mathbb{R}^{m\times n}$

$$\{x\in\mathbb{R}^n \mid \|C(x-a)\|^2 = (x-a)^{\mathrm{T}} C^{\mathrm{T}} C(x-a) \leq 1\} \tag{2}$$

is a good icon for a convex set.[2.1]

2.1.1 subspace

A nonempty subset $\mathcal{R}$ of real Euclidean vector space $\mathbb{R}^n$ is called a *subspace* (§2.5) if every vector[2.2] of the form $\alpha x + \beta y$, for $\alpha , \beta \in \mathbb{R}$, is in $\mathcal{R}$ whenever vectors x and y are. [266, §2.3] A subspace is a convex set containing the origin, by definition. [325, p.4] Any subspace is therefore open in the sense that it contains no boundary, but closed in the sense [274, §2]

$$\mathcal{R} + \mathcal{R} = \mathcal{R} \tag{3}$$

It is not difficult to show

$$\mathcal{R} = -\mathcal{R} \tag{4}$$

as is true for any subspace $\mathcal{R}$, because $x\in\mathcal{R} \Leftrightarrow -x\in\mathcal{R}$. Given any $x\in\mathcal{R}$

$$\mathcal{R} = x + \mathcal{R} \tag{5}$$

Intersection of an arbitrary collection of subspaces remains a subspace. Any subspace not constituting the entire *ambient vector space* $\mathbb{R}^n$ is a *proper subspace*; *e.g*,[2.3] any line (of infinite extent) through the origin in two-dimensional Euclidean space $\mathbb{R}^2$. The vector space $\mathbb{R}^n$ is itself a conventional subspace, inclusively, [243, §2.1] although not proper.

2.1.2 linear independence

Arbitrary given vectors in Euclidean space $\{\Gamma_i \in \mathbb{R}^n ,\ i = 1 \ldots N\}$ are *linearly independent* (l.i.) if and only if, for all $\zeta \in \mathbb{R}^N$ $(\zeta_i \in \mathbb{R})$

$$\Gamma_1 \zeta_1 + \cdots + \Gamma_{N-1} \zeta_{N-1} - \Gamma_N \zeta_N = \mathbf{0} \tag{6}$$

has only the *trivial solution* $\zeta = \mathbf{0}$; in other words, iff no vector from the given set can be expressed as a linear combination of those remaining.

Geometrically, two nontrivial vector subspaces are linearly independent iff they intersect only at the origin.

2.1.2.1 preservation of linear independence

(*confer* §2.4.2.4, §2.10.1) Linear transformation preserves linear dependence. [243, p.86] Conversely, linear independence can be preserved under linear transformation. Given $Y = [\, y_1 \;\; y_2 \; \cdots \; y_N \,] \in \mathbb{R}^{N\times N}$, consider the mapping

$$T(\Gamma) : \mathbb{R}^{n\times N} \to \mathbb{R}^{n\times N} \;\triangleq\; \Gamma Y \tag{7}$$

[2.1] Ellipsoid semiaxes are eigenvectors of $C^{\mathrm{T}}C$ whose lengths are inverse square root eigenvalues. This particular definition is slablike (Figure **13**) in $\mathbb{R}^n$ when C has nontrivial nullspace.

[2.2] A *vector* is assumed, throughout, to be a column vector.

[2.3] We substitute abbreviation *e.g* in place of the Latin *exempli gratia*.

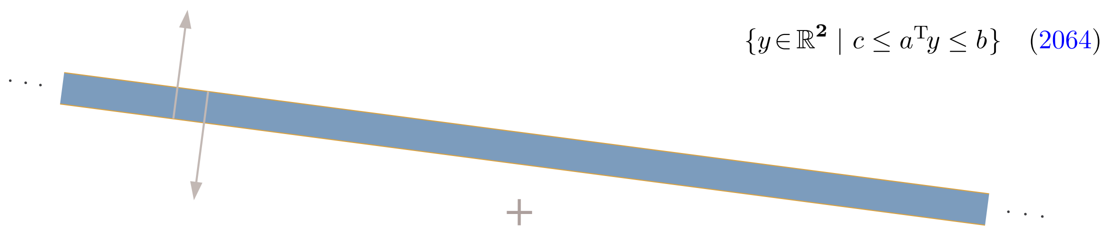


Figure 13: A *slab* is a convex Euclidean body infinite in extent but not affine. Illustrated in $\mathbb{R}^2$, it may be constructed by intersecting two opposing halfspaces whose bounding hyperplanes are parallel but not coincident. Because number of halfspaces used in its construction is finite, slab is a *polyhedron* (§2.12). (Cartesian axes + and vector inward-normal, to each halfspace-boundary, are drawn for reference.)

whose domain is the set of all matrices $\Gamma\in\mathbb{R}^{n\times N}$ holding a linearly independent set columnar. Linear independence of $\{\Gamma y_i\in\mathbb{R}^n,\ i=1\ldots N\}$ demands, by definition, there exist no nontrivial solution $\zeta\in\mathbb{R}^N$ to

$$\Gamma y_1\,\zeta_i + \cdots + \Gamma y_{N-1}\,\zeta_{N-1} - \Gamma y_N\,\zeta_N = \mathbf{0} \tag{8}$$

By factoring out Γ, we see that triviality is ensured by linear independence of $\{y_i\in\mathbb{R}^N\}$.

2.1.3 Orthant:

name given to a closed convex set that is the higher-dimensional generalization of *quadrant* from the classical Cartesian partition of $\mathbb{R}^2$; a *Cartesian cone*. The most common is the nonnegative orthant $\mathbb{R}^n_+$ or $\mathbb{R}^{n\times n}_+$ (analogue to quadrant I) to which membership denotes nonnegative vector- or matrix-entries respectively; *e.g*,

$$\mathbb{R}^n_+ \triangleq \{x\in\mathbb{R}^n \mid x_i\geq 0\ \ \forall\, i\} \tag{9}$$

The nonpositive orthant $\mathbb{R}^n_-$ or $\mathbb{R}^{n\times n}_-$ (analogue to quadrant III) denotes negative and 0 entries. Orthant convexity[2.4] is easily verified by definition (1).

2.1.4 affine set

A nonempty *affine set* (from the word *affinity*) is any subset of $\mathbb{R}^n$ that is a translation of some subspace. Any affine set is convex and open so contains no boundary: *e.g*, empty set $\emptyset$, point, line, plane, *hyperplane* (§2.4.2), subspace, *etcetera*. The intersection of an arbitrary collection of affine sets remains affine.

[2.4] All orthants are selfdual simplicial cones. (§2.13.5.1, §2.12.3.1.1)

2.1.4.0.1 Definition. *Affine subset.*
We analogize *affine subset* to subspace,[2.5] defining it to be any nonempty affine set of vectors; an affine subset of $\mathbb{R}^n$. △

For some *parallel*[2.6] subspace $\mathcal{R}$ and any point $x \in \mathcal{A}$

$$\begin{aligned} \mathcal{A} \text{ is affine } \Leftrightarrow \quad \mathcal{A} &= x + \mathcal{R} \\ &= \{y \mid y - x \in \mathcal{R}\} \end{aligned} \tag{10}$$

Affine hull of a set $\mathcal{C} \subseteq \mathbb{R}^n$ (§2.3.1) is the smallest affine set containing it.

2.1.5 dimension

Dimension of an arbitrary set $\mathcal{S}$ is Euclidean dimension of its affine hull; [410, p.14]

$$\dim \mathcal{S} \triangleq \dim \operatorname{aff} \mathcal{S} = \dim \operatorname{aff}(\mathcal{S} - s)\ , \quad s \in \mathcal{S} \tag{11}$$

the same as dimension of the subspace parallel to that affine set $\operatorname{aff} \mathcal{S}$ when nonempty. Hence dimension (of a set) is synonymous with *affine dimension.* [215, A.2.1]

2.1.6 empty set *versus* empty interior

Emptiness $\emptyset$ of a set is handled differently than *interior* in the classical literature. It is common for a nonempty convex set to have empty interior; *e.g*, paper in the real world:

- An ordinary flat sheet of paper is a nonempty convex set having empty interior in $\mathbb{R}^3$ but nonempty interior relative to its affine hull.

2.1.6.1 relative interior

Although it is always possible to pass to a smaller ambient Euclidean space where a nonempty set acquires an interior [27, §II.2.3], we prefer the qualifier *relative* which is the conventional fix to this ambiguous terminology.[2.7] So we distinguish *interior* from *relative interior* throughout: [347] [410] [377]

- Classical interior $\operatorname{int} \mathcal{C}$ is defined as a union of points: x is an interior point of $\mathcal{C} \subseteq \mathbb{R}^n$ if there exists an open ball of dimension n and nonzero radius centered at x that is contained in $\mathcal{C}$.

- Relative interior $\operatorname{rel\,int} \mathcal{C}$ of a convex set $\mathcal{C} \subseteq \mathbb{R}^n$ is interior relative to its affine hull.[2.8]

[2.5] The popular term *affine subspace* is an oxymoron.
[2.6] Two affine sets are *parallel* when one is a translation of the other. [325, p.4]
[2.7] Superfluous mingling of terms as in *relatively nonempty set* would be an unfortunate consequence. From the opposite perspective, some authors use the term *full* or *full-dimensional* to describe a set having nonempty interior.
[2.8] Likewise for *relative boundary* (§2.1.7.2), although *relative closure* is superfluous. [215, §A.2.1]

(a)

(b) $\mathbb{R}^2$

(c)

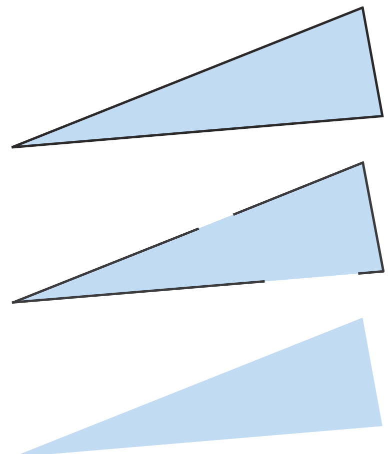


Figure 14: **(a)** Closed convex set. **(b)** Neither open, closed, or convex. Yet PSD cone can remain convex in absence of certain boundary components (§2.9.2.9.3). Nonnegative orthant with origin excluded (§2.6) and positive orthant with origin adjoined [325, p.49] are convex. **(c)** Open convex set.

Thus defined, it is common (though confusing) for $\operatorname{int}\mathcal{C}$ the interior of $\mathcal{C}$ to be empty while its relative interior is not: this happens whenever dimension of its affine hull is less than dimension of the ambient space $(\dim \operatorname{aff}\mathcal{C} < n\,;$ *e.g*, were $\mathcal{C}$ paper$)$ or in the exception when $\mathcal{C}$ is a single point; [274, §2.2.1]

$$\operatorname{rel\,int}\{x\} \triangleq \operatorname{aff}\{x\} = \{x\}\ , \qquad \operatorname{int}\{x\} = \emptyset\ , \qquad x \in \mathbb{R}^n \tag{12}$$

In any case, *closure* of the relative interior of a convex set $\mathcal{C}$ always yields closure of the set itself;

$$\overline{\operatorname{rel\,int}\mathcal{C}} = \overline{\mathcal{C}} \tag{13}$$

Closure is invariant to translation. If $\mathcal{C}$ is convex then $\operatorname{rel\,int}\mathcal{C}$ and $\overline{\mathcal{C}}$ are convex. [215, p.24] If $\mathcal{C}$ has nonempty interior, then

$$\operatorname{rel\,int}\mathcal{C} = \operatorname{int}\mathcal{C} \tag{14}$$

Given the intersection of convex set $\mathcal{C}$ with affine set $\mathcal{A}$

$$\operatorname{rel\,int}(\mathcal{C}\cap\mathcal{A}) = \operatorname{rel\,int}(\mathcal{C})\cap\mathcal{A} \quad \Leftarrow \quad \operatorname{rel\,int}(\mathcal{C})\cap\mathcal{A} \neq \emptyset \tag{15}$$

Because an affine set $\mathcal{A}$ is open

$$\operatorname{rel\,int}\mathcal{A} = \mathcal{A} \tag{16}$$

2.1.7 classical boundary

(*confer* §2.1.7.2) *Boundary* of a set $\mathcal{C}$ is the closure of $\mathcal{C}$ less its interior;

$$\partial\mathcal{C} = \overline{\mathcal{C}} \setminus \operatorname{int}\mathcal{C} \tag{17}$$

[56, §1.1] which follows from the fact

$$\overline{\operatorname{int}\mathcal{C}} = \overline{\mathcal{C}} \quad\Leftrightarrow\quad \partial\operatorname{int}\mathcal{C} = \partial\mathcal{C} \tag{18}$$

and presumption of nonempty interior.[2.9] Implications are:

- $\operatorname{int}\mathcal{C} = \overline{\mathcal{C}} \setminus \partial\mathcal{C}$
- a bounded open set has *boundary* defined but not contained in the set
- interior of an open set is equivalent to the set itself;

from which an open set is defined: [274, p.109]

$$\mathcal{C} \text{ is } open \;\Leftrightarrow\; \operatorname{int}\mathcal{C} = \mathcal{C} \tag{19}$$

$$\mathcal{C} \text{ is } closed \;\Leftrightarrow\; \overline{\operatorname{int}\mathcal{C}} = \mathcal{C} \tag{20}$$

The set illustrated in Figure **14**b is not open because it is not equivalent to its interior, for example, it is not closed because it does not contain its boundary, and it is not convex because it does not contain all convex combinations of its boundary points.

2.1.7.1 Line intersection with boundary

A line can intersect the boundary of a convex set in any dimension at a point demarcating the line's entry to the set interior. On one side of that entry-point along the line is the exterior of the set, on the other side is the set interior. In other words,

- starting from any point of a convex set, a move toward the interior is an immediate entry into the interior. [27, §II.2]

When a line intersects the interior of a convex body in any dimension, the boundary appears to the line to be as thin as a point. This is intuitively plausible because, for example, a line intersects the boundary of the ellipsoids in Figure **15** at a point in $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$. Such thinness is a remarkable fact when pondering visualization of convex *polyhedra* (§2.12, §5.14.3) in four Euclidean dimensions, for example, having boundaries constructed from other three-dimensional convex polyhedra called *faces.*

We formally define *face* in (§2.6). For now, we observe the boundary of a convex body to be entirely constituted by all its faces of dimension lower than the body itself. Any face of a convex set is convex. For example: The ellipsoids in Figure **15** have boundaries composed only of zero-dimensional faces. The two-dimensional slab in Figure **13** is an unbounded *polyhedron* having one-dimensional faces making its boundary. The three-dimensional bounded polyhedron in Figure **22** has zero-, one-, and two-dimensional polygonal faces constituting its boundary.

[2.9] Otherwise, for $x \in \mathbb{R}^n$ as in (12), [274, §2.1-§2.3]

$$\overline{\operatorname{int}\{x\}} = \overline{\emptyset} = \emptyset$$

the empty set is both open and closed.

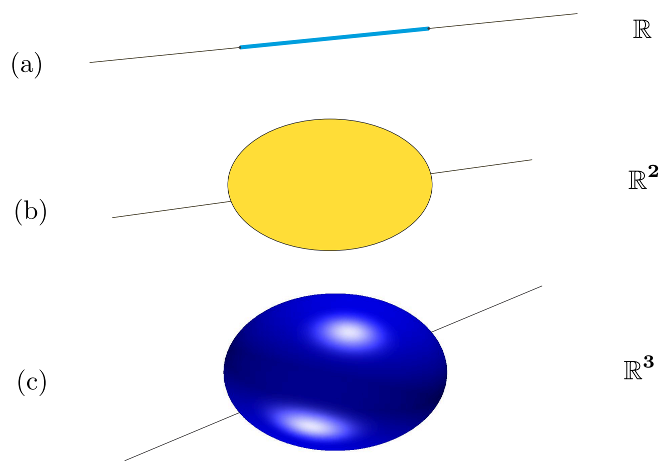


Figure 15: **(a)** Ellipsoid in $\mathbb{R}$ is a line segment whose boundary comprises two points. Intersection of line with ellipsoid in $\mathbb{R}$, **(b)** in $\mathbb{R}^2$, **(c)** in $\mathbb{R}^3$. Each ellipsoid illustrated has entire boundary constituted by zero-dimensional faces; in fact, by *vertices* (§2.6.1.0.1). Intersection of line with boundary is a point at entry to interior. These same facts hold in higher dimension.

2.1.7.1.1 Example. *Intersection of line with boundary in* $\mathbb{R}^{\mathbf{6}}$.
The convex cone of positive semidefinite matrices $\mathbb{S}^{\mathbf{3}}_{+}$ (§2.9), in the ambient subspace of symmetric matrices $\mathbb{S}^{\mathbf{3}}$ (§2.2.2.0.1), is a six-dimensional Euclidean body in *isometrically isomorphic* $\mathbb{R}^{\mathbf{6}}$ (§2.2.1). Boundary of the positive semidefinite cone, in this dimension, comprises faces having only the dimensions 0, 1, and 3; *id est*, $\{\rho(\rho+1)/2,\ \rho=0,1,2\}$.

Unique minimum-distance projection PX (§E.9) of any point $X\in\mathbb{S}^{\mathbf{3}}$ on that cone $\mathbb{S}^{\mathbf{3}}_{+}$ is known in closed form (§7.1.2). Given, for example, $\lambda\in\operatorname{int}\mathbb{R}^{\mathbf{3}}_{+}$ and *diagonalization* (§A.5.1) of exterior point

$$X = Q\Lambda Q^{\mathrm{T}} \in \mathbb{S}^{\mathbf{3}}, \quad \Lambda \triangleq \begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \lambda_2 & \\ \mathbf{0} & & -\lambda_3 \end{bmatrix} \tag{21}$$

where $Q\in\mathbb{R}^{\mathbf{3}\times\mathbf{3}}$ is an *orthogonal matrix,* then the projection on $\mathbb{S}^{\mathbf{3}}_{+}$ in $\mathbb{R}^{\mathbf{6}}$ is

$$PX = Q\begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \lambda_2 & \\ \mathbf{0} & & 0 \end{bmatrix}Q^{\mathrm{T}} \in \mathbb{S}^{\mathbf{3}}_{+} \tag{22}$$

This positive semidefinite matrix PX nearest X thus has rank 2, found by discarding all negative eigenvalues in Λ. The line connecting these two points is $\{X + (PX-X)t \mid t\in\mathbb{R}\}$ where $t=0 \Leftrightarrow X$ and $t=1 \Leftrightarrow PX$. Because this line intersects the boundary of the *positive semidefinite cone* $\mathbb{S}^{\mathbf{3}}_{+}$ at point PX and passes through its interior (by assumption), then the matrix corresponding to an infinitesimally positive perturbation of t there should reside interior to the cone (rank 3). Indeed, for ε an arbitrarily small positive constant,

$$X + (PX-X)t|_{t=1+\varepsilon} = Q(\Lambda+(P\Lambda-\Lambda)(1+\varepsilon))Q^{\mathrm{T}} = Q\begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \lambda_2 & \\ \mathbf{0} & & \varepsilon\lambda_3 \end{bmatrix}Q^{\mathrm{T}} \in \operatorname{int}\mathbb{S}^{\mathbf{3}}_{+} \tag{23}$$

$\square$

2.1.7.1.2 Example. *Tangential line intersection with boundary.*
A higher-dimensional boundary $\partial\mathcal{C}$ of a convex Euclidean body $\mathcal{C}$ is simply a dimensionally larger set through which a line can pass when it does not intersect the body's interior. Still, for example, a line existing in five or more dimensions may pass *tangentially* (intersecting no point interior to $\mathcal{C}$ [235, §15.3]) through a single point relatively interior to a three-dimensional face on $\partial\mathcal{C}$. Let's understand why by inductive reasoning.

Figure **16**a shows a vertical line-segment whose boundary comprises its two endpoints. For a line to pass through the boundary tangentially (intersecting no point relatively interior to the line-segment), it must exist in an ambient space of at least two dimensions. Otherwise, the line is confined to the same one-dimensional space as the line-segment and must pass along the segment to reach the end points.

Figure **16**b illustrates a two-dimensional ellipsoid whose boundary is constituted entirely by zero-dimensional faces. Again, a line must exist in at least two dimensions to tangentially pass through any single arbitrarily chosen point on the boundary (without intersecting the ellipsoid interior).

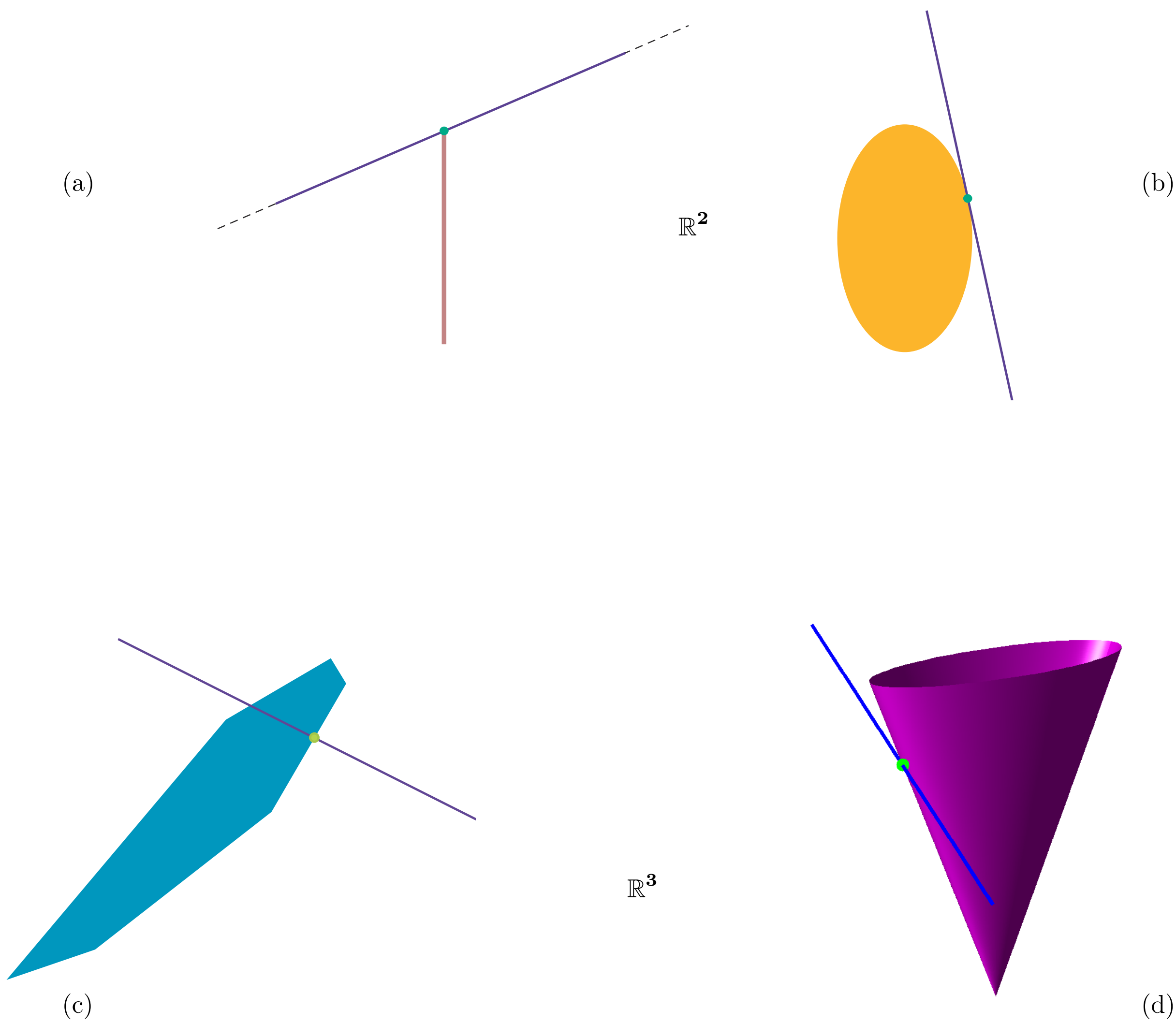


Figure 16: Line tangential: **(a) (b)** to relative interior of a zero-dimensional face in $\mathbb{R}^2$, **(c) (d)** to relative interior of a one-dimensional face in $\mathbb{R}^3$.

Now let's move to an ambient space of three dimensions. Figure **16**c shows a polygon rotated into three dimensions. For a line to pass through its zero-dimensional boundary (one of its *vertices*) tangentially, it must exist in at least the two dimensions of the polygon. But for a line to pass tangentially through a single arbitrarily chosen point in the relative interior of a one-dimensional face on the boundary as illustrated, it must exist in at least three dimensions.

Figure **16**d illustrates a solid circular cone (drawn truncated) whose one-dimensional faces are halflines emanating from its pointed end (*vertex*). This cone's boundary is constituted solely by those one-dimensional halflines. A line may pass through the boundary tangentially, striking only one arbitrarily chosen point relatively interior to a one-dimensional face, if it exists in at least the three-dimensional ambient space of the cone.

From these few examples, way deduce a general rule (without proof):

- A line may pass tangentially through a single arbitrarily chosen point relatively interior to a k-dimensional face on the boundary of a convex Euclidean body if the line exists in dimension at least equal to $k+2$.

Now the interesting part, with regard to Figure **22** showing a bounded polyhedron in $\mathbb{R}^3$; call it $\mathcal{P}$: A line existing in at least four dimensions is required in order to pass tangentially (without hitting $\operatorname{int}\mathcal{P}$) through a single arbitrary point in the relative interior of any two-dimensional polygonal face on the boundary of polyhedron $\mathcal{P}$. Now imagine that polyhedron $\mathcal{P}$ is itself a three-dimensional face of some other polyhedron in $\mathbb{R}^4$. To pass a line tangentially through polyhedron $\mathcal{P}$ itself, striking only one point from its relative interior $\operatorname{rel\,int}\mathcal{P}$ as claimed, requires a line existing in at least five dimensions.[2.10]

It is not too difficult to deduce:

- A line may pass through a single arbitrarily chosen point interior to a k-dimensional convex Euclidean body (hitting no other interior point) if that line exists in dimension at least equal to $k+1$.

In layman's terms, this means: a being capable of navigating four spatial dimensions (one Euclidean dimension beyond our physical reality) could see inside three-dimensional objects. □

2.1.7.2 Relative boundary

The classical definition of *boundary* of a set $\mathcal{C}$ presumes nonempty interior:

$$\partial\mathcal{C} = \overline{\mathcal{C}} \setminus \operatorname{int}\mathcal{C} \qquad (17)$$

More suitable to study of convex sets is the *relative boundary*; defined [215, §A.2.1.2]

$$\operatorname{rel}\partial\mathcal{C} \triangleq \overline{\mathcal{C}} \setminus \operatorname{rel\,int}\mathcal{C} \qquad (24)$$

boundary relative to affine hull of $\mathcal{C}$.

[2.10] This rule can help determine whether there exists unique solution to a convex optimization problem whose *feasible set* is an intersecting line; *e.g*, the *trilateration* problem (§5.4.2.2.8).

In the exception when $\mathcal{C}$ is a single point $\{x\}$, (12)

$$\operatorname{rel}\partial\{x\} = \overline{\{x\}}\backslash\{x\} = \emptyset\,, \qquad x\in\mathbb{R}^n \tag{25}$$

A bounded convex polyhedron (§2.3.2, §2.12.0.0.1) in subspace $\mathbb{R}$, for example, has boundary constructed from two points, in $\mathbb{R}^2$ from at least three line segments, in $\mathbb{R}^3$ from convex polygons, while a convex *polychoron* (a bounded polyhedron in $\mathbb{R}^4$ [412]) has boundary constructed from three-dimensional convex polyhedra. A halfspace is partially bounded by a hyperplane; its interior therefore excludes that hyperplane. An affine set has no relative boundary.

2.1.8 intersection, sum, difference, product

2.1.8.0.1 Theorem. *Intersection.* [325, §2, thm.6.5]
Intersection of an arbitrary collection of convex sets $\{\mathcal{C}_i\}$ is convex. For a finite collection of N sets, a necessarily nonempty intersection of relative interior $\bigcap_{i=1}^N \operatorname{rel}\operatorname{int}\mathcal{C}_i = \operatorname{rel}\operatorname{int}\bigcap_{i=1}^N \mathcal{C}_i$ equals relative interior of intersection. And for a possibly infinite collection, $\bigcap\overline{\mathcal{C}_i} = \overline{\bigcap\mathcal{C}_i}$. ⋄

In converse this theorem is implicitly false insofar as a convex set can be formed by the intersection of sets that are not. Unions of convex sets are generally not convex. [215, p.22]

Vector sum of two convex sets $\mathcal{C}_1$ and $\mathcal{C}_2$ is convex [215, p.24]

$$\mathcal{C}_1 + \mathcal{C}_2 = \{x+y \mid x\in\mathcal{C}_1\,,\; y\in\mathcal{C}_2\} \tag{26}$$

but not necessarily closed unless at least one set is closed and bounded.

By additive inverse, we can similarly define *vector difference* of two convex sets

$$\mathcal{C}_1 - \mathcal{C}_2 = \{x-y \mid x\in\mathcal{C}_1\,,\; y\in\mathcal{C}_2\} \tag{27}$$

which is convex. Applying this definition to nonempty convex set $\mathcal{C}_1$, its selfdifference $\mathcal{C}_1-\mathcal{C}_1$ is generally nonempty, nontrivial, and convex; *e.g*, for any *convex cone* $\mathcal{K}$, (§2.7.2) the set $\mathcal{K}-\mathcal{K}$ constitutes its affine hull. [325, p.15]

Cartesian product of convex sets

$$\mathcal{C}_1\times\mathcal{C}_2 = \left\{\begin{bmatrix} x\\ y\end{bmatrix} \mid x\in\mathcal{C}_1\,,\; y\in\mathcal{C}_2\right\} = \begin{bmatrix}\mathcal{C}_1\\ \mathcal{C}_2\end{bmatrix} \tag{28}$$

remains convex. The converse also holds; *id est*, a Cartesian product is convex iff each set is. [215, p.23]

Convex results are also obtained for scaling $\kappa\,\mathcal{C}$ of a convex set $\mathcal{C}$, rotation/reflection $Q\,\mathcal{C}$, or translation $\mathcal{C}+\alpha$; each similarly defined.

Given any operator T and convex set $\mathcal{C}$, we are prone to write $T(\mathcal{C})$ meaning

$$T(\mathcal{C}) \triangleq \{T(x) \mid x\in\mathcal{C}\} \tag{29}$$

Given linear operator T, it therefore follows from (26),

$$\begin{aligned} T(\mathcal{C}_1+\mathcal{C}_2) &= \{T(x+y) \mid x\in\mathcal{C}_1\,,\; y\in\mathcal{C}_2\}\\ &= \{T(x)+T(y) \mid x\in\mathcal{C}_1\,,\; y\in\mathcal{C}_2\}\\ &= T(\mathcal{C}_1)+T(\mathcal{C}_2)\end{aligned} \tag{30}$$

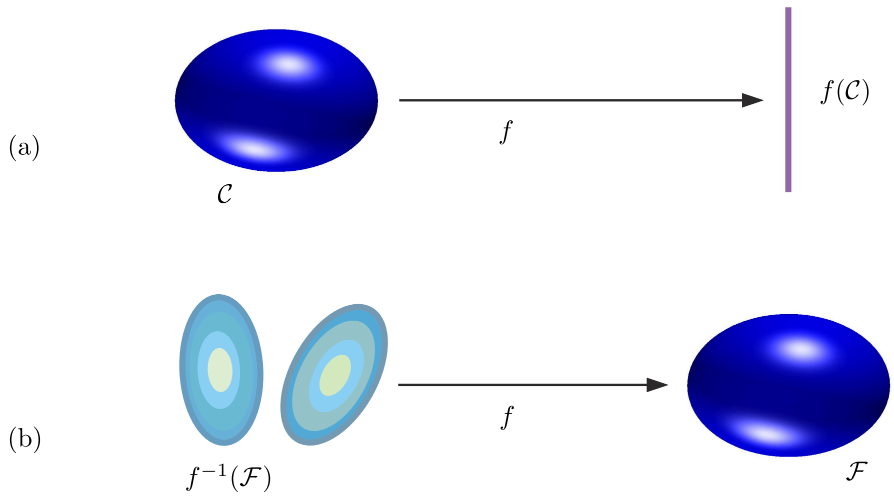


Figure 17: **(a)** Image of convex set in domain of any convex function f is convex, but there is no converse. **(b)** Inverse image under convex function f.

2.1.9 inverse image

While *epigraph* (§3.5) of a convex function must be convex, it generally holds that inverse image (Figure **17**) of a convex function is not. The most prominent examples to the contrary are affine functions (§3.4):

2.1.9.0.1 Theorem. *Inverse image.* [325, §3]
Let f be a mapping from $\mathbb{R}^{p\times k}$ to $\mathbb{R}^{m\times n}$.

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

$$f(\mathcal{C}) = \{f(X) \mid X \in \mathcal{C}\} \subseteq \mathbb{R}^{m\times n} \tag{31}$$

  is convex.

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

$$f^{-1}(\mathcal{F}) = \{X \mid f(X) \in \mathcal{F}\} \subseteq \mathbb{R}^{p\times k} \tag{32}$$

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

In particular, any affine transformation of an affine set remains affine. [325, p.8] Ellipsoids are invariant to any [*sic*] affine transformation.

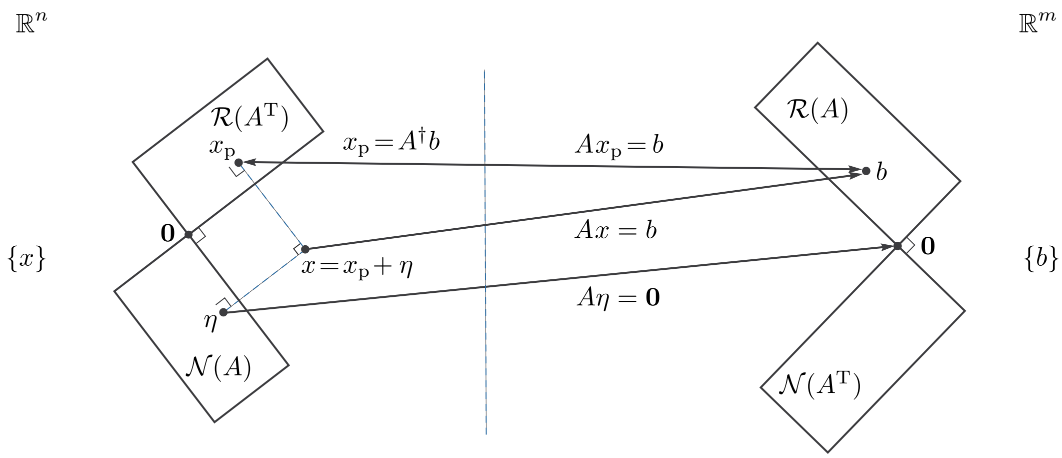


Figure 18: (*confer* Figure **175**) Action of linear map represented by $A \in \mathbb{R}^{m\times n}$: Component of vector x in nullspace $\mathcal{N}(A)$ maps to origin while component in rowspace $\mathcal{R}(A^{\mathrm{T}})$ maps to range $\mathcal{R}(A)$. For any $A \in \mathbb{R}^{m\times n}$, $A^{\dagger}Ax = x_{\mathrm{p}}$ and $AA^{\dagger}Ax = b$ (§E) and inverse image of $b \in \mathcal{R}(A)$ is a nonempty affine set: $x_{\mathrm{p}} + \mathcal{N}(A)$.

Although not precluded, this *inverse image theorem* does not require a uniquely invertible mapping f . Figure **18**, for example, mechanizes inverse image under a general linear map. Example 2.9.1.0.2 and Example 3.5.0.0.2 offer further applications.

Each converse of this two-part theorem is generally false; *id est*, given f affine, a convex image $f(\mathcal{C})$ does not imply that set $\mathcal{C}$ is convex, and neither does a convex inverse image $f^{-1}(\mathcal{F})$ imply set $\mathcal{F}$ is convex. A counterexample, invalidating a converse, is easy to visualize when the affine function is an orthogonal *projector* [348] [266]:

2.1.9.0.2 Corollary. *Projection on subspace.*[2.11] (2037) [325, §3]
Orthogonal projection of a convex set on a subspace or nonempty affine set is another convex set. ⋄

Again, the converse is false. Shadows, for example, are umbral projections that can be convex when the body providing the shade is not.

2.2 Vectorized-matrix inner product

Euclidean space $\mathbb{R}^n$ comes equipped with a vector inner-product

$$\langle y\,, z\rangle \triangleq y^{\mathrm{T}}z = \|y\|\|z\|\cos\psi \tag{33}$$

[2.11] For hyperplane representations see §2.4.2. For projection of convex sets on hyperplanes see [410, §6.6]. A nonempty affine set is called an *affine subset* (§2.1.4.0.1). Orthogonal projection of points on affine subsets is reviewed in §E.4.

where ψ (1004) represents angle (in radians) between vectors y and z. We prefer those angle brackets to connote a geometric rather than algebraic perspective; *e.g*, vector y might represent a hyperplane normal (§2.4.2). Two vectors are *orthogonal* (*perpendicular*) to one another if and only if their inner product vanishes (iff ψ is an odd multiple of $\frac{\pi}{2}$);

$$y \perp z \;\Leftrightarrow\; \langle y\,,z\rangle = 0 \tag{34}$$

When orthogonal vectors each have unit *norm,* then they are *orthonormal.* A vector inner-product defines Euclidean norm (vector 2-norm, §A.7.1)

$$\|y\|_2 = \|y\| \triangleq \sqrt{y^{\mathrm T} y}\;, \qquad \|y\| = 0 \;\Leftrightarrow\; y = \mathbf{0} \tag{35}$$

For linear operator A, its *adjoint* $A^{\mathrm T}$ is a linear operator defined by [243, §3.10]

$$\langle y\,, A^{\mathrm T} z\rangle \triangleq \langle Ay\,, z\rangle \tag{36}$$

For linear operation on a vector, represented by real matrix A, the adjoint operator $A^{\mathrm T}$ is its *transposition.* This operator is *selfadjoint* when $A = A^{\mathrm T}$.

Vector inner-product for matrices is calculated just as it is for vectors; by first transforming a matrix in $\mathbb{R}^{p\times k}$ to a vector in $\mathbb{R}^{pk}$ by concatenating its columns in the *natural order.* For lack of a better term, we shall call that linear *bijective* (one-to-one and onto [243, App.A1.2]) transformation *vectorization.* For example, the vectorization of $Y = [\,y_1\; y_2 \cdots\, y_k\,] \in \mathbb{R}^{p\times k}$ [182] [344] is

$$\operatorname{vec} Y \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} \in \mathbb{R}^{pk} \tag{37}$$

Then the vectorized-matrix inner-product is trace of matrix inner-product; for $Z \in \mathbb{R}^{p\times k}$, [63, §2.6.1] [215, §0.3.1] [421, §8] [384, §2.2]

$$\langle Y\,,\, Z\rangle \;\triangleq\; \operatorname{tr}(Y^{\mathrm T} Z) \;=\; \operatorname{vec}(Y)^{\mathrm T} \operatorname{vec} Z \tag{38}$$

where (§A.1.1)

$$\operatorname{tr}(Y^{\mathrm T} Z) = \operatorname{tr}(Z Y^{\mathrm T}) = \operatorname{tr}(Y Z^{\mathrm T}) = \operatorname{tr}(Z^{\mathrm T} Y) = \mathbf{1}^{\mathrm T}(Y \circ Z)\mathbf{1} \tag{39}$$

and where $\circ$ denotes the *Hadamard product* [2.12] of matrices [174, §1.1.4]. The adjoint $A^{\mathrm T}$ operation on a matrix can therefore be defined in like manner:

$$\langle Y\,,\, A^{\mathrm T} Z\rangle \triangleq \langle AY\,,\, Z\rangle \tag{40}$$

Take any element $\mathcal{C}_1$ from a matrix-valued set in $\mathbb{R}^{p\times k}$, for example, and consider any particular dimensionally compatible real vectors v and w. Then vector inner-product of $\mathcal{C}_1$ with $vw^{\mathrm T}$ is

$$\langle vw^{\mathrm T}, \mathcal{C}_1\rangle \;=\; \langle v\,, \mathcal{C}_1 w\rangle \;=\; v^{\mathrm T}\mathcal{C}_1 w \;=\; \operatorname{tr}(wv^{\mathrm T}\mathcal{C}_1) \;=\; \mathbf{1}^{\mathrm T}\big((vw^{\mathrm T})\circ \mathcal{C}_1\big)\,\mathbf{1} \tag{41}$$

[2.12] Hadamard product is a simple entrywise product of corresponding entries from two matrices of like size; *id est*, not necessarily square. A commutative operation, the Hadamard product can be extracted from within a Kronecker product. [218, p.475]

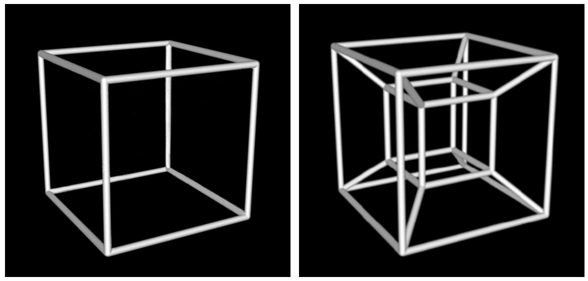


Figure 19: **(a)** Cube in $\mathbb{R}^3$ projected on paper-plane $\mathbb{R}^2$. Subspace projection operator is not an isomorphism because new adjacencies are introduced. **(b)** *Tesseract* is a projection of hypercube in $\mathbb{R}^4$ on $\mathbb{R}^3$.

Further, linear bijective vectorization is *distributive* with respect to Hadamard product; *id est*,

$$\operatorname{vec}(Y \circ Z) = \operatorname{vec}(Y) \circ \operatorname{vec}(Z) \tag{42}$$

2.2.0.0.1 Example. *Application of inverse image theorem.*
Suppose set $\mathcal{C} \subseteq \mathbb{R}^{p\times k}$ were convex. Then for any particular vectors $v \in \mathbb{R}^p$ and $w \in \mathbb{R}^k$, the set of vector inner-products

$$\mathcal{Y} \triangleq v^{\mathrm{T}}\mathcal{C}w \;=\; \langle vw^{\mathrm{T}}, \mathcal{C}\rangle \subseteq \mathbb{R} \tag{43}$$

is convex. It is easy to show directly that convex combination of elements from $\mathcal{Y}$ remains an element of $\mathcal{Y}$.[2.13] Instead given convex set $\mathcal{Y}$, $\mathcal{C}$ must be convex consequent to *inverse image theorem* 2.1.9.0.1.

More generally, vw^{T} in (43) may be replaced with any particular matrix $Z \in \mathbb{R}^{p\times k}$ while convexity of set $\langle Z, \mathcal{C}\rangle \subseteq \mathbb{R}$ persists. Further, by replacing v and w with any particular respective matrices U and W of dimension compatible with all elements of convex set $\mathcal{C}$, then set $U^{\mathrm{T}}\mathcal{C}W$ is convex by the *inverse image theorem* because it is a linear mapping of $\mathcal{C}$. □

2.2.1 Frobenius'

2.2.1.0.1 Definition. *Isomorphic.*
An *isomorphism* of a vector space is a transformation equivalent to a linear bijective mapping. Image and inverse image under the transformation operator are then called *isomorphic vector spaces.* △

[2.13] To verify that, take any two elements $\mathcal{C}_1$ and $\mathcal{C}_2$ from the convex matrix-valued set $\mathcal{C}$, and then form the vector inner-products (43) that are two elements of $\mathcal{Y}$ by definition. Now make a convex combination of those inner products; *videlicet*, for $0 \leq \mu \leq 1$

$$\mu\,\langle vw^{\mathrm{T}}, \mathcal{C}_1\rangle \;+\; (1-\mu)\,\langle vw^{\mathrm{T}}, \mathcal{C}_2\rangle \;=\; \langle vw^{\mathrm{T}},\; \mu\,\mathcal{C}_1 + (1-\mu)\mathcal{C}_2\rangle$$

The two sides are equivalent by linearity of inner product. The right-hand side remains a vector inner-product of vw^{T} with an element $\mu\mathcal{C}_1 + (1-\mu)\mathcal{C}_2$ from the convex set $\mathcal{C}$; hence, it belongs to $\mathcal{Y}$. Since that holds true for any two elements from $\mathcal{Y}$, then it must be a convex set. ♦

Isomorphic vector spaces are characterized by preservation of *adjacency*; *id est*, if v and w are points connected by a line segment in one vector space, then their images will be connected by a line segment in the other. Two Euclidean bodies may be considered isomorphic if there exists an isomorphism, of their vector spaces, under which the bodies correspond. [386, §I.1] Projection (§E) is not an isomorphism, Figure **19** for example; hence, perfect reconstruction (inverse projection) is generally impossible without additional information.

When $Z = Y \in \mathbb{R}^{p\times k}$ in (38), *Frobenius' norm* is resultant from vector inner-product; (*confer* (1781))

$$\begin{aligned}\|Y\|_{\mathrm F}^2 \;&=\; \|\operatorname{vec} Y\|_2^2 \;=\; \langle Y\,,\,Y\rangle \;=\; \operatorname{tr}(Y^{\mathrm T} Y)\\ &=\; \sum_{i,j} Y_{ij}^2 \;=\; \sum_i \lambda(Y^{\mathrm T} Y)_i \;=\; \sum_i \sigma(Y)_i^2\end{aligned} \tag{44}$$

where $\lambda(Y^{\mathrm T} Y)_i$ is the i^{th} eigenvalue of $Y^{\mathrm T} Y$, and $\sigma(Y)_i$ the i^{th} singular value of Y. Were Y a *normal matrix* (§A.5.1), then $\sigma(Y) = |\lambda(Y)|$ [432, §8.1] thus

$$\|Y\|_{\mathrm F}^2 \;=\; \sum_i \lambda(Y)_i^2 \;=\; \|\lambda(Y)\|_2^2 \;=\; \langle \lambda(Y)\,,\,\lambda(Y)\rangle = \langle Y\,,\,Y\rangle \tag{45}$$

The converse (45) $\Rightarrow$ normal matrix Y also holds. [218, §2.5.4]

Frobenius' norm is the Euclidean norm of vectorized matrices. Because the metrics are equivalent, for $X \in \mathbb{R}^{p\times k}$

$$\|\operatorname{vec} X - \operatorname{vec} Y\|_2 = \|X - Y\|_{\mathrm F} \tag{46}$$

and because vectorization (37) is a linear bijective map, then vector space $\mathbb{R}^{p\times k}$ is isometrically isomorphic with vector space $\mathbb{R}^{pk}$ in the Euclidean sense and vec is an *isometric isomorphism* of $\mathbb{R}^{p\times k}$. Because of this Euclidean structure, all known results from convex analysis in Euclidean space $\mathbb{R}^n$ carry over directly to the space of real matrices $\mathbb{R}^{p\times k}$; *e.g*, norm function convexity (§3.2).

2.2.1.1 Injective linear operators

Injective mapping (transformation) means one-to-one mapping; synonymous with *uniquely invertible* linear mapping on Euclidean space.

- Linear injective mappings are fully characterized by lack of nontrivial nullspace.

2.2.1.1.1 Definition. *Isometric isomorphism.*
An isometric isomorphism of a vector space having a *metric* defined on it is a linear bijective mapping T that preserves distance; *id est*, for all $x, y \in \operatorname{dom} T$

$$\|Tx - Ty\| = \|x - y\| \tag{47}$$

Then isometric isomorphism T is called a *bijective isometry.* $\triangle$

Unitary linear operator $Q : \mathbb{R}^k \to \mathbb{R}^k$, represented by orthogonal matrix $Q \in \mathbb{R}^{k\times k}$ (§B.5.2), is an isometric isomorphism; *e.g*, discrete Fourier transform via (889). Suppose $T(X) = UXQ$ is a bijective isometry where U is a dimensionally compatible *orthonormal*

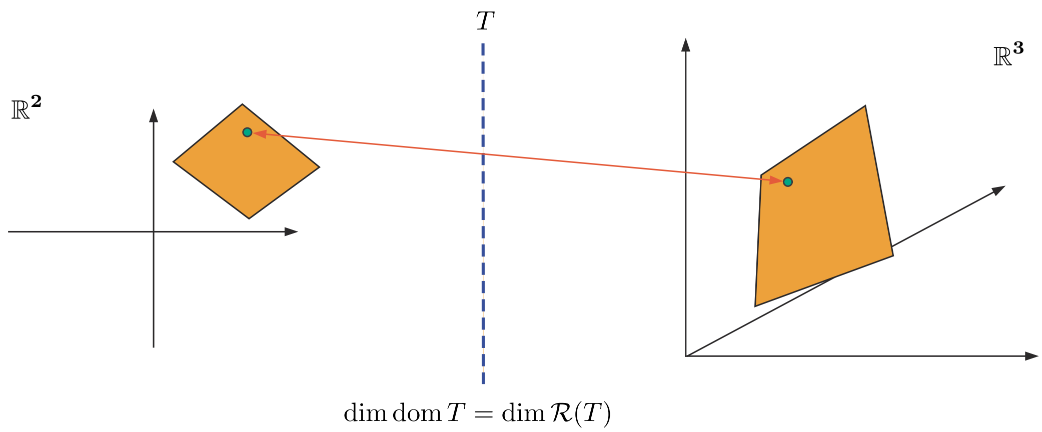


Figure 20: Linear injective mapping $Tx = Ax : \mathbb{R}^2 \to \mathbb{R}^3$ of Euclidean body remains two-dimensional under mapping represented by skinny full-rank matrix $A \in \mathbb{R}^{3\times 2}$; two bodies are isomorphic by Definition 2.2.1.0.1.

matrix.[2.14] Then we also say Frobenius' norm is *orthogonally invariant*; meaning, for $X, Y \in \mathbb{R}^{p\times k}$

$$\|U(X-Y)Q\|_{\mathrm{F}} = \|X-Y\|_{\mathrm{F}} \tag{48}$$

Yet isometric operator $T : \mathbb{R}^2 \to \mathbb{R}^3$, represented by $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ on $\mathbb{R}^2$, is injective but not a *surjective* map to $\mathbb{R}^3$. [243, §1.6, §2.6] This operator T can therefore be a bijective isometry only with respect to its range.

Any linear injective transformation on Euclidean space is uniquely invertible on its range. In fact, any linear injective transformation has a range whose dimension equals that of its domain. In other words, for any invertible linear transformation T [*ibidem*]

$$\dim \operatorname{dom}(T) = \dim \mathcal{R}(T) \tag{49}$$

e.g, T represented by skinny-or-square full-rank matrices. (Figure **20**) An important consequence of this fact is:

- Affine dimension, of any n-dimensional Euclidean body in domain of operator T, is invariant to linear injective transformation.

2.2.1.1.2 Example. *Noninjective linear operators.*
Mappings in Euclidean space created by noninjective linear operators can be characterized in terms of an orthogonal projector (§E). Consider noninjective linear operator

[2.14] any matrix U whose columns are orthonormal with respect to each other ($U^{\mathrm{T}}U = I$); these include the orthogonal matrices.

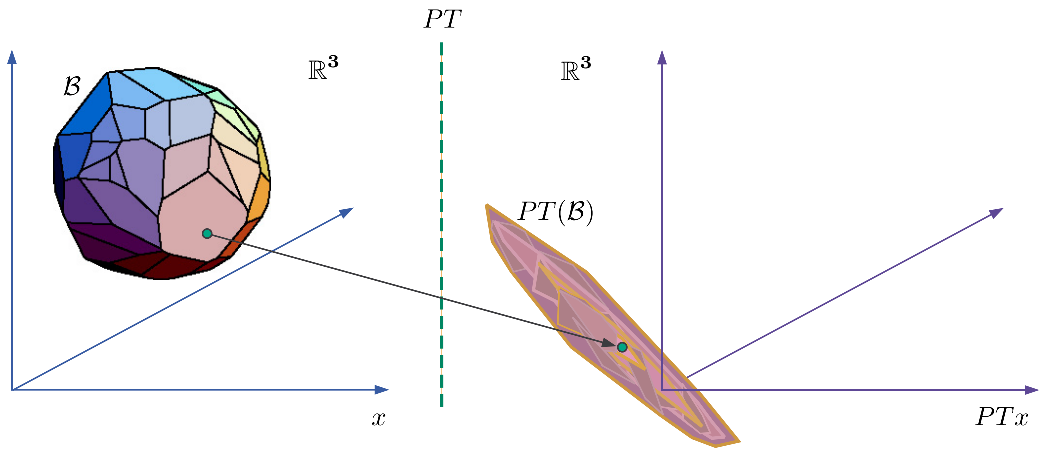


Figure 21: Linear noninjective mapping $PTx = A^\dagger A x : \mathbb{R}^3 \to \mathbb{R}^3$ of three-dimensional Euclidean body $\mathcal{B}$ has affine dimension 2 under projection on rowspace of fat full-rank matrix $A \in \mathbb{R}^{2\times 3}$. Set of coefficients of orthogonal projection $T\mathcal{B} = \{Ax \mid x \in \mathcal{B}\}$ is isomorphic with projection $P(T\mathcal{B})$ [*sic*].

$Tx = Ax : \mathbb{R}^n \to \mathbb{R}^m$ represented by fat matrix $A \in \mathbb{R}^{m\times n}$ $(m < n)$. What can be said about the nature of this m-dimensional mapping?

Concurrently, consider injective linear operator $Py = A^\dagger y : \mathbb{R}^m \to \mathbb{R}^n$ where $\mathcal{R}(A^\dagger) = \mathcal{R}(A^{\mathrm{T}})$. $P(Ax) = PTx$ achieves projection of vector x on the row space $\mathcal{R}(A^{\mathrm{T}})$. (§E.3.1) This means vector Ax can be succinctly interpreted as coefficients of orthogonal projection.

Pseudoinverse matrix $A^\dagger$ is skinny and full-rank, so operator Py is a linear *bijection* with respect to its range $\mathcal{R}(A^\dagger)$. By Definition 2.2.1.0.1, image $P(T\mathcal{B})$ of projection $PT(\mathcal{B})$ on $\mathcal{R}(A^{\mathrm{T}})$ in $\mathbb{R}^n$ must therefore be isomorphic with the set of projection coefficients $T\mathcal{B} = \{Ax \mid x \in \mathcal{B}\}$ in $\mathbb{R}^m$ and have the same affine dimension by (49). To illustrate, we present a three-dimensional Euclidean body $\mathcal{B}$ in Figure **21** where any point x in the nullspace $\mathcal{N}(A)$ maps to the origin. □

2.2.2 Symmetric matrices

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

$$\mathbb{S}^M \triangleq \left\{A \in \mathbb{R}^{M\times M} \mid A = A^{\mathrm{T}}\right\} \subseteq \mathbb{R}^{M\times M} \tag{50}$$

This subspace comprising symmetric matrices $\mathbb{S}^M$ is isomorphic with the vector space $\mathbb{R}^{M(M+1)/2}$ whose dimension is the number of free variables in a symmetric $M\times M$ matrix. The *orthogonal complement* [348] [266] of $\mathbb{S}^M$ is

$$\mathbb{S}^{M\perp} \triangleq \left\{A \in \mathbb{R}^{M\times M} \mid A = -A^{\mathrm{T}}\right\} \subset \mathbb{R}^{M\times M} \tag{51}$$

the subspace of *antisymmetric* matrices in $\mathbb{R}^{M\times M}$; *id est*,

$$\mathbb{S}^M \oplus \mathbb{S}^{M\perp} = \mathbb{R}^{M\times M} \tag{52}$$

where unique vector sum $\oplus$ is defined on page 666. △

All antisymmetric matrices are *hollow* by definition (have $\mathbf{0}$ main diagonal). Any square matrix $A\in\mathbb{R}^{M\times M}$ can be written as a sum of its symmetric and antisymmetric parts: respectively,

$$A = \frac{1}{2}(A + A^{\mathrm{T}}) + \frac{1}{2}(A - A^{\mathrm{T}}) \tag{53}$$

The symmetric part is orthogonal in $\mathbb{R}^{M^2}$ to the antisymmetric part; *videlicet*,

$$\operatorname{tr}\big((A + A^{\mathrm{T}})(A - A^{\mathrm{T}})\big) = 0 \tag{54}$$

In the ambient space of real matrices, the antisymmetric matrix subspace can be described

$$\mathbb{S}^{M\perp} = \left\{\frac{1}{2}(A - A^{\mathrm{T}}) \mid A\in\mathbb{R}^{M\times M}\right\} \subset \mathbb{R}^{M\times M} \tag{55}$$

because any matrix in $\mathbb{S}^M$ is orthogonal to any matrix in $\mathbb{S}^{M\perp}$. Further confined to the ambient subspace of symmetric matrices, $\mathbb{S}^{M\perp}$ would become trivial.

2.2.2.1 Isomorphism of symmetric matrix subspace

When a matrix is symmetric in $\mathbb{S}^M$, we may still employ the vectorization transformation (37) to $\mathbb{R}^{M^2}$; vec , an isometric isomorphism. We might instead choose to realize in the lower-dimensional subspace $\mathbb{R}^{M(M+1)/2}$ by ignoring redundant entries (below the main diagonal) during transformation. Such a realization would remain isomorphic but not isometric. Lack of isometry is a spatial distortion due now to disparity in metric between $\mathbb{R}^{M^2}$ and $\mathbb{R}^{M(M+1)/2}$. To realize isometrically in $\mathbb{R}^{M(M+1)/2}$, we must make a correction: For $Y = [Y_{ij}]\in\mathbb{S}^M$ we take symmetric vectorization [231, §2.2.1]

$$\operatorname{svec} Y \triangleq \begin{bmatrix} Y_{11} \\ \sqrt{2}Y_{12} \\ Y_{22} \\ \sqrt{2}Y_{13} \\ \sqrt{2}Y_{23} \\ Y_{33} \\ \vdots \\ Y_{MM} \end{bmatrix} \in \mathbb{R}^{M(M+1)/2} \tag{56}$$

where all entries off the main diagonal have been scaled. Now for $Z\in\mathbb{S}^M$

$$\langle Y\,,\,Z\rangle \triangleq \operatorname{tr}(Y^{\mathrm{T}}Z) = \operatorname{vec}(Y)^{\mathrm{T}}\operatorname{vec} Z = \mathbf{1}^{\mathrm{T}}(Y\circ Z)\mathbf{1} = \operatorname{svec}(Y)^{\mathrm{T}}\operatorname{svec} Z \tag{57}$$

Then because the metrics become equivalent, for $X\in\mathbb{S}^M$

$$\|\operatorname{svec} X - \operatorname{svec} Y\|_2 = \|X - Y\|_{\mathrm{F}} \tag{58}$$

and because symmetric vectorization (56) is a linear bijective mapping, then svec is an isometric isomorphism of the symmetric matrix subspace. In other words, $\mathbb{S}^M$ is isometrically isomorphic with $\mathbb{R}^{M(M+1)/2}$ in the Euclidean sense under transformation svec .

The set of all symmetric matrices $\mathbb{S}^M$ forms a proper subspace in $\mathbb{R}^{M\times M}$, so for it there exists a standard orthonormal *basis* in isometrically isomorphic $\mathbb{R}^{M(M+1)/2}$

$$\{E_{ij}\in\mathbb{S}^M\} = \left\{\begin{array}{ll} e_i e_i^{\mathrm T}\,, & i=j=1\ldots M \\ \frac{1}{\sqrt{2}}\Big(e_i e_j^{\mathrm T} + e_j e_i^{\mathrm T}\Big)\,, & 1\leq i<j\leq M \end{array}\right\} \tag{59}$$

where $M(M+1)/2$ *standard basis matrices* E_{ij} are formed from the *standard basis vectors*

$$e_i = \left[\left\{\begin{array}{ll} 1\,, & i=j \\ 0\,, & i\neq j \end{array}\right.\,,\quad j=1\ldots M\right]\in\mathbb{R}^M \tag{60}$$

Thus we have a basic *orthogonal expansion* for $Y\in\mathbb{S}^M$

$$Y = \sum_{j=1}^{M}\sum_{i=1}^{j}\langle E_{ij}\,,Y\rangle\, E_{ij} \tag{61}$$

whose unique coefficients

$$\langle E_{ij}\,,Y\rangle \;=\; \left\{\begin{array}{ll} Y_{ii}\,, & i=1\ldots M \\ Y_{ij}\sqrt{2}\,, & 1\leq i<j\leq M \end{array}\right. \tag{62}$$

correspond to entries of the symmetric vectorization (56).

2.2.3 Symmetric hollow subspace

2.2.3.0.1 Definition. *Hollow subspaces.* [371]
Define a subspace of $\mathbb{R}^{M\times M}$: the convex set of all (real) symmetric $M\times M$ matrices having $\mathbf{0}$ main diagonal;

$$\mathbb{R}_h^{M\times M} \triangleq \left\{A\in\mathbb{R}^{M\times M} \mid A=A^{\mathrm T},\;\; \delta(A)=\mathbf{0}\right\}\subset\mathbb{R}^{M\times M} \tag{63}$$

where the main diagonal of $A\in\mathbb{R}^{M\times M}$ is denoted (§A.1)

$$\delta(A)\in\mathbb{R}^M \tag{1504}$$

Operating on a vector, linear operator δ naturally returns a diagonal matrix; $\delta(\delta(A))$ is a diagonal matrix. Operating recursively on a vector $\Lambda\in\mathbb{R}^N$ or diagonal matrix $\Lambda\in\mathbb{S}^N$, operator $\delta(\delta(\Lambda))$ returns Λ itself;

$$\delta^2(\Lambda) \;\equiv\; \delta(\delta(\Lambda)) \;=\; \Lambda \tag{1506}$$

The subspace $\mathbb{R}_h^{M\times M}$ (63) comprising (real) symmetric hollow matrices is isomorphic with subspace $\mathbb{R}^{M(M-1)/2}$; its orthogonal complement is

$$\mathbb{R}_h^{M\times M\perp} \triangleq \left\{A\in\mathbb{R}^{M\times M} \mid A=-A^{\mathrm T}+2\delta^2(A)\right\}\subseteq\mathbb{R}^{M\times M} \tag{64}$$

the subspace of *antisymmetric antihollow* matrices in $\mathbb{R}^{M\times M}$; *id est*,

$$\mathbb{R}_h^{M\times M} \oplus \mathbb{R}_h^{M\times M\perp} = \mathbb{R}^{M\times M} \tag{65}$$

Yet defined instead as a proper subspace of ambient $\mathbb{S}^M$

$$\mathbb{S}_h^M \triangleq \left\{A\in\mathbb{S}^M \mid \delta(A) = \mathbf{0}\right\} \equiv \mathbb{R}_h^{M\times M} \subset \mathbb{S}^M \tag{66}$$

the orthogonal complement $\mathbb{S}_h^{M\perp}$ of *symmetric hollow subspace* $\mathbb{S}_h^M$,

$$\mathbb{S}_h^{M\perp} \triangleq \left\{A\in\mathbb{S}^M \mid A=\delta^2(A)\right\} \subseteq \mathbb{S}^M \tag{67}$$

called *symmetric antihollow subspace,* is simply the subspace of diagonal matrices; *id est*,

$$\mathbb{S}_h^M \oplus \mathbb{S}_h^{M\perp} = \mathbb{S}^M \tag{68}$$

$\triangle$

Any matrix $A\in\mathbb{R}^{M\times M}$ can be written as a sum of its symmetric hollow and antisymmetric antihollow parts: respectively,

$$A = \left(\frac{1}{2}(A+A^{\mathrm T}) - \delta^2(A)\right) + \left(\frac{1}{2}(A-A^{\mathrm T}) + \delta^2(A)\right) \tag{69}$$

The symmetric hollow part is orthogonal to the antisymmetric antihollow part in $\mathbb{R}^{M^2}$; *videlicet*,

$$\operatorname{tr}\left(\left(\frac{1}{2}(A+A^{\mathrm T}) - \delta^2(A)\right)\left(\frac{1}{2}(A-A^{\mathrm T}) + \delta^2(A)\right)\right) = 0 \tag{70}$$

because any matrix in subspace $\mathbb{R}_h^{M\times M}$ is orthogonal to any matrix in the antisymmetric antihollow subspace

$$\mathbb{R}_h^{M\times M\perp} = \left\{\frac{1}{2}(A-A^{\mathrm T}) + \delta^2(A) \mid A\in\mathbb{R}^{M\times M}\right\} \subseteq \mathbb{R}^{M\times M} \tag{71}$$

of the ambient space of real matrices; which reduces to the diagonal matrices in the ambient space of symmetric matrices

$$\mathbb{S}_h^{M\perp} = \left\{\delta^2(A) \mid A\in\mathbb{S}^M\right\} = \left\{\delta(u) \mid u\in\mathbb{R}^M\right\} \subseteq \mathbb{S}^M \tag{72}$$

In anticipation of their utility with Euclidean distance matrices (EDMs) in §5, for symmetric hollow matrices we introduce the linear bijective vectorization dvec that is the natural analogue to symmetric matrix vectorization svec (56): for $Y = [Y_{ij}]\in\mathbb{S}_h^M$

$$\operatorname{dvec} Y \triangleq \sqrt{2}\begin{bmatrix} Y_{12} \\ Y_{13} \\ Y_{23} \\ Y_{14} \\ Y_{24} \\ Y_{34} \\ \vdots \\ Y_{M-1,M} \end{bmatrix} \in \mathbb{R}^{M(M-1)/2} \tag{73}$$

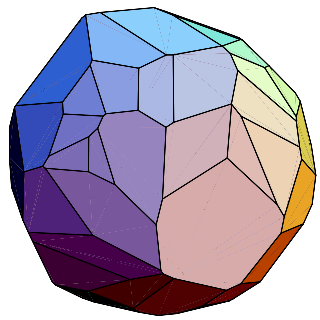


Figure 22: Convex hull of a random list of points in $\mathbb{R}^3$. Some points from that generating list reside interior to this *convex polyhedron* (§2.12). [412, *Convex Polyhedron*] (Avis-Fukuda-Mizukoshi)

Like svec, dvec is an isometric isomorphism on the symmetric hollow subspace. For $X \in \mathbb{S}_h^M$

$$\|\operatorname{dvec} X - \operatorname{dvec} Y\|_2 = \|X - Y\|_F \tag{74}$$

The set of all symmetric hollow matrices $\mathbb{S}_h^M$ forms a proper subspace in $\mathbb{R}^{M\times M}$, so for it there must be a standard orthonormal basis in isometrically isomorphic $\mathbb{R}^{M(M-1)/2}$

$$\{E_{ij} \in \mathbb{S}_h^M\} = \left\{ \frac{1}{\sqrt{2}}\left(e_i e_j^{\mathrm{T}} + e_j e_i^{\mathrm{T}}\right), \quad 1 \le i < j \le M \right\} \tag{75}$$

where $M(M-1)/2$ standard basis matrices E_{ij} are formed from the standard basis vectors $e_i \in \mathbb{R}^M$.

The *symmetric hollow majorization corollary* A.1.2.0.2 characterizes eigenvalues of symmetric hollow matrices.

2.3 Hulls

We focus on the affine, convex, and conic hulls: convex sets that may be regarded as kinds of Euclidean container or vessel united with its interior.

2.3.1 Affine hull, affine dimension

Affine dimension of any set in $\mathbb{R}^n$ is the dimension of the smallest affine set (empty set, point, line, plane, hyperplane (§2.4.2), translated subspace, $\mathbb{R}^n$) that contains it. For nonempty sets, affine dimension is the same as dimension of the subspace parallel to that affine set. [325, §1] [215, §A.2.1]

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

$$X = [\, x_1 \, \cdots \, x_N \,] \in \mathbb{R}^{n\times N} \tag{76}$$

In particular, we define *affine dimension* r of the N-point list X as dimension of the smallest affine set in Euclidean space $\mathbb{R}^n$ that contains X ;

$$r \triangleq \dim \operatorname{aff} X \tag{77}$$

Affine dimension r is a lower bound sometimes called *embedding dimension.* [371] [201] That affine set $\mathcal{A}$ in which those points are embedded is unique and called the *affine hull* [347, §2.1];

$$\begin{aligned} \mathcal{A} \;\triangleq\; & \operatorname{aff}\{x_\ell \in \mathbb{R}^n ,\;\ell = 1 \dots N\} && = \;\operatorname{aff} X \\ =\;& x_1 + \mathcal{R}\{x_\ell - x_1\,,\;\ell = 2 \dots N\} && = \;\{Xa \mid a^{\mathrm{T}}\mathbf{1} = 1\} \;\subseteq\; \mathbb{R}^n \end{aligned} \tag{78}$$

for which we call list X a set of *generators.* Hull $\mathcal{A}$ is parallel to subspace

$$\mathcal{R}\{x_\ell - x_1\,,\;\ell = 2 \dots N\} \;=\; \mathcal{R}(X - x_1\mathbf{1}^{\mathrm{T}}) \;\subseteq\; \mathbb{R}^n \tag{79}$$

where

$$\mathcal{R}(A) \;=\; \{Ax \mid \forall\, x\} \qquad (142)$$

Given some arbitrary set $\mathcal{C}$ and any $x \in \mathcal{C}$

$$\operatorname{aff}\mathcal{C} = x + \operatorname{aff}(\mathcal{C} - x) \tag{80}$$

where $\operatorname{aff}(\mathcal{C} - x)$ is a subspace.

$$\operatorname{aff}\emptyset \triangleq \emptyset \tag{81}$$

The affine hull of a point x is that point itself;

$$\operatorname{aff}\{x\} = \{x\} \tag{82}$$

Affine hull of two distinct points is the unique line through them. (Figure **23**) The affine hull of three noncollinear points in any dimension is that unique plane containing the points, and so on. The subspace of symmetric matrices $\mathbb{S}^m$ is the affine hull of the cone of positive semidefinite matrices; (§2.9)

$$\operatorname{aff}\mathbb{S}^m_+ = \mathbb{S}^m \tag{83}$$

2.3.1.0.1 Example. *Affine hull of rank-*1 *correlation matrices.* [234]
The set of all $m \times m$ *rank-*1 *correlation matrices* is defined by all the binary vectors y in $\mathbb{R}^m$ (*confer* §5.9.1.0.1)

$$\{yy^{\mathrm{T}} \in \mathbb{S}^m_+ \mid \delta(yy^{\mathrm{T}}) = \mathbf{1}\} \tag{84}$$

Affine hull of the rank-1 correlation matrices is equal to the set of normalized symmetric matrices; *id est,*

$$\operatorname{aff}\{yy^{\mathrm{T}} \in \mathbb{S}^m_+ \mid \delta(yy^{\mathrm{T}}) = \mathbf{1}\} \;=\; \{A \in \mathbb{S}^m \mid \delta(A) = \mathbf{1}\} \tag{85}$$

□

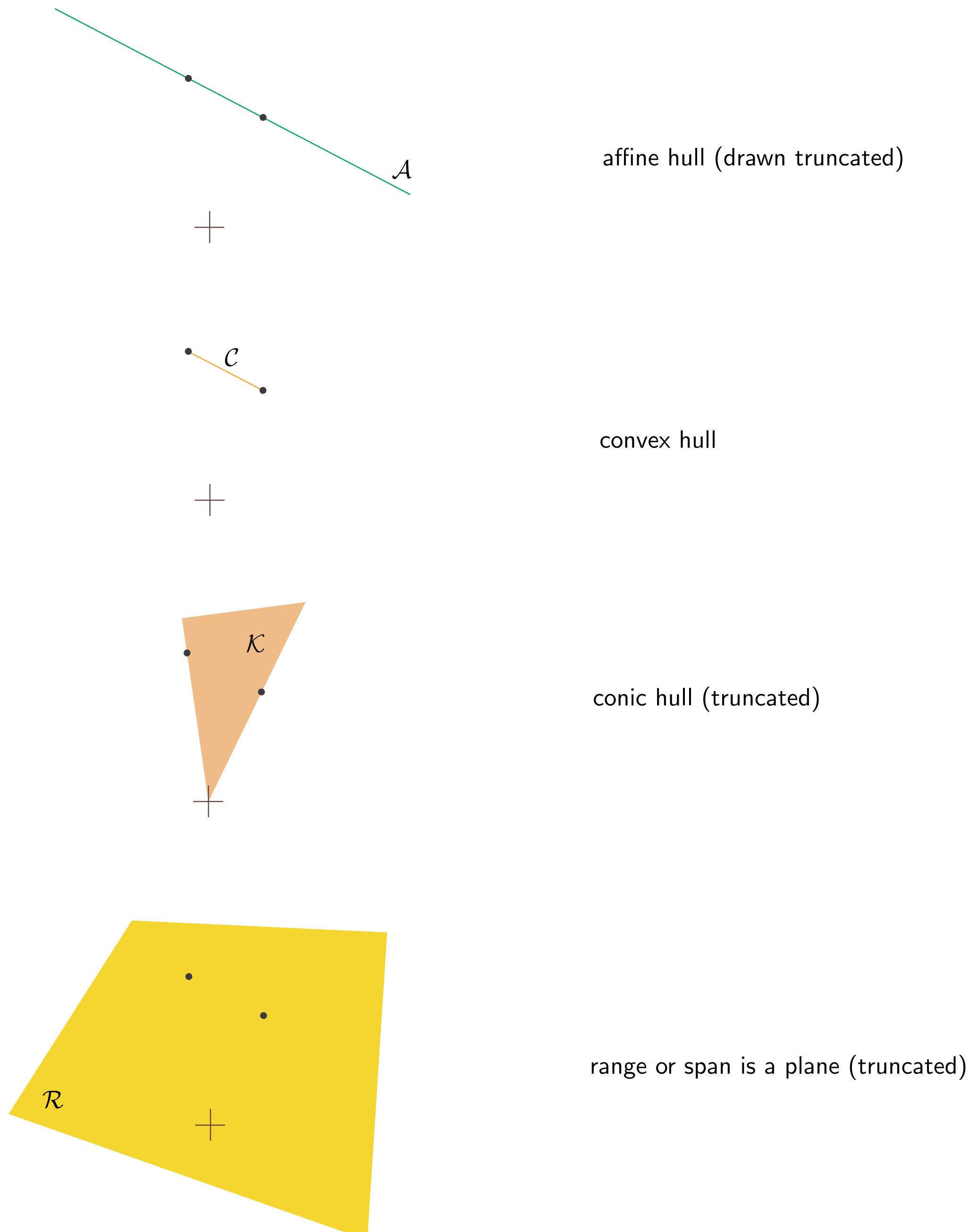


Figure 23: Given two points in Euclidean vector space of any dimension, their various hulls are illustrated. Each hull is a subset of range; generally, $\mathcal{A}, \mathcal{C}, \mathcal{K} \subseteq \mathcal{R} \ni \mathbf{0}$. (Cartesian axes drawn for reference.)

2.3.1.0.2 Exercise. *Affine hull of correlation matrices.*
Prove (85) via definition of affine hull. Find the convex hull instead. ▼

2.3.1.1 Partial order induced by $\mathbb{R}_+^N$ and $\mathbb{S}_+^M$

Notation $a \succeq 0$ means vector a belongs to nonnegative orthant $\mathbb{R}_+^N$ while $a \succ 0$ means vector a belongs to the nonnegative orthant's interior $\operatorname{int}\mathbb{R}_+^N$. $a \succeq b$ denotes comparison of vector a to vector b on $\mathbb{R}^N$ with respect to the nonnegative orthant; *id est*, $a \succeq b$ means $a-b$ belongs to the nonnegative orthant but neither a or b is necessarily nonnegative. With particular respect to the nonnegative orthant, $a \succeq b \;\Leftrightarrow\; a_i \geq b_i \;\; \forall\, i$ (369).

More generally, $a \succeq_{\mathcal{K}} b$ denotes comparison with respect to pointed closed convex cone $\mathcal{K}$, whereas comparison with respect to the cone's interior is denoted $a \succ_{\mathcal{K}} b$. But equivalence with entrywise comparison does not generally hold, and neither a or b necessarily belongs to $\mathcal{K}$. (§2.7.2.2)

The symbol $\geq$ is reserved for scalar comparison on the real line $\mathbb{R}$ with respect to the nonnegative real line $\mathbb{R}_+$ as in $a^{\mathrm{T}}y \geq b$. Comparison of matrices with respect to the positive semidefinite cone $\mathbb{S}_+^M$, like $I \succeq A \succeq 0$ in Example 2.3.2.0.1, is explained in §2.9.0.1.

2.3.2 Convex hull

The *convex hull* [215, §A.1.4] [325] of any bounded[2.15] list or set of N points $X \in \mathbb{R}^{n\times N}$ forms a unique bounded convex *polyhedron* (*confer* §2.12.0.0.1) whose vertices constitute some subset of that list;

$$\mathcal{P} \;\triangleq\; \operatorname{conv}\{x_\ell \;,\; \ell=1\ldots N\} \;=\; \operatorname{conv} X \;=\; \{Xa \mid a^{\mathrm{T}}\mathbf{1} = 1,\; a \succeq 0\} \subseteq \mathbb{R}^n \tag{86}$$

Union of relative interior and relative boundary (§2.1.7.2) of the polyhedron comprise its convex hull $\mathcal{P}$, the smallest closed convex set that contains the list X ; *e.g*, Figure **22**. Given $\mathcal{P}$, the *generating list* $\{x_\ell\}$ is not unique. But because every bounded polyhedron is the convex hull of its vertices, [347, §2.12.2] the vertices of $\mathcal{P}$ comprise a *minimal set* of generators.

Given some arbitrary set $\mathcal{C} \subseteq \mathbb{R}^n$, its convex hull $\operatorname{conv}\mathcal{C}$ is equivalent to the smallest convex set containing it. (*confer* §2.4.1.1.1) The convex hull is a subset of the affine hull;

$$\operatorname{conv}\mathcal{C} \;\subseteq\; \operatorname{aff}\mathcal{C} \;=\; \operatorname{aff}\overline{\mathcal{C}} \;=\; \overline{\operatorname{aff}\mathcal{C}} \;=\; \operatorname{aff}\operatorname{conv}\mathcal{C} \tag{87}$$

Any closed bounded convex set $\mathcal{C}$ is equal to the convex hull of its boundary;

$$\mathcal{C} = \operatorname{conv}\partial\mathcal{C} \tag{88}$$

$$\operatorname{conv}\emptyset \triangleq \emptyset \tag{89}$$

[2.15] An arbitrary set $\mathcal{C}$ in $\mathbb{R}^n$ is *bounded* iff it can be contained in a Euclidean ball having finite radius. [120, §2.2] (*confer* §5.7.3.0.1) The smallest ball containing $\mathcal{C}$ has radius $\inf\limits_x \sup\limits_{y\in\mathcal{C}} \|x-y\|$ and center $x^\star$ whose determination is a convex problem because $\sup\limits_{y\in\mathcal{C}} \|x-y\|$ is a convex function of x ; but the supremum may be difficult to ascertain.

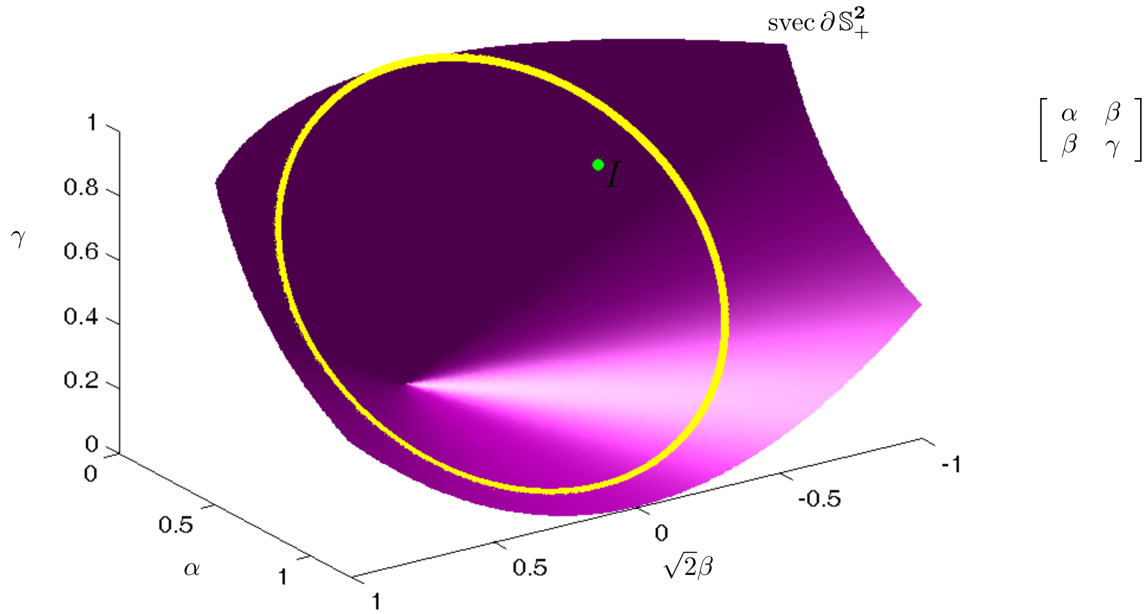


Figure 24: Two Fantopes. Circle (radius $1/\sqrt{2}$), shown here on boundary of positive semidefinite cone $\mathbb{S}_+^2$ in isometrically isomorphic $\mathbb{R}^3$ from Figure **46**, comprises boundary of a Fantope (90) in this dimension ($k=1,\ N=2$). Lone point illustrated is Identity matrix I, interior to PSD cone, and is that Fantope corresponding to $k=2,\ N=2$. (View is from inside PSD cone looking toward origin.)

2.3.2.0.1 Example. *Hull of* rank-k *projection matrices.* [155] [304] [12, §4.1] [310, §3] [255, §2.4] [256] Convex hull of the set comprising outer product of orthonormal matrices has equivalent expression: for $1 \leq k \leq N$ (§2.9.0.1)

$$\operatorname{conv}\left\{UU^{\mathrm T} \mid U \in \mathbb{R}^{N\times k},\ U^{\mathrm T}U = I\right\} = \left\{A \in \mathbb{S}^N \mid I \succeq A \succeq 0\,,\ \langle I\,,\ A\rangle = k\right\} \subset \mathbb{S}^N_+ \tag{90}$$

This important convex body we call *Fantope* (after mathematician Ky Fan). In case $k=1$, there is slight simplification: ((1710), Example 2.9.2.7.1)

$$\operatorname{conv}\left\{UU^{\mathrm T} \mid U \in \mathbb{R}^{N},\ U^{\mathrm T}U = 1\right\} = \left\{A \in \mathbb{S}^N \mid A \succeq 0\,,\ \langle I\,,\ A\rangle = 1\right\} \tag{91}$$

This particular Fantope is called *spectahedron.* [*sic*] [163, §5.1] In case $k=N$, the Fantope is Identity matrix I. More generally, the set

$$\left\{UU^{\mathrm T} \mid U \in \mathbb{R}^{N\times k},\ U^{\mathrm T}U = I\right\} \tag{92}$$

comprises the extreme points (§2.6.0.0.1) of its convex hull. By (1552), each and every extreme point $UU^{\mathrm T}$ has only k nonzero eigenvalues λ and they all equal 1; *id est*, $\lambda(UU^{\mathrm T})_{1:k} = \lambda(U^{\mathrm T}U) = \mathbf{1}$. So Frobenius' norm of each and every extreme point equals the same constant

$$\|UU^{\mathrm T}\|_{\mathrm F}^2 = k \tag{93}$$

Each extreme point simultaneously lies on the boundary of the positive semidefinite cone (when $k<N$, §2.9) and on the boundary of a *hypersphere* of dimension $k(N-\frac{k}{2}+\frac{1}{2})$ and radius $\sqrt{k(1-\frac{k}{N})}$ centered at $\frac{k}{N}I$ along the ray (base $\mathbf{0}$) through the Identity matrix I in isomorphic vector space $\mathbb{R}^{N(N+1)/2}$ (§2.2.2.1).

Figure **24** illustrates extreme points (92) comprising the boundary of a Fantope, the boundary of a *disc* corresponding to $k=1$, $N=2$; but that circumscription does not hold in higher dimension. (§2.9.2.8) □

2.3.2.0.2 Example. *Nuclear norm ball*: *convex hull of rank-1 matrices.*
From (91), in Example 2.3.2.0.1, we learn that the convex hull of normalized symmetric rank-1 matrices is a slice of the positive semidefinite cone. In §2.9.2.7 we find the convex hull of all symmetric rank-1 matrices to be the entire positive semidefinite cone.

In the present example we abandon symmetry; instead posing, what is the convex hull of bounded nonsymmetric rank-1 matrices:

$$\operatorname{conv}\{uv^{\mathrm T} \mid \|uv^{\mathrm T}\| \leq 1\,,\ u \in \mathbb{R}^m,\ v \in \mathbb{R}^n\} = \{X \in \mathbb{R}^{m\times n} \mid \sum_i \sigma(X)_i \leq 1\} \tag{94}$$

where $\sigma(X)$ is a vector of singular values. (Since $\|uv^{\mathrm T}\| = \|u\|\|v\|$ (1701), norm of each vector constituting a *dyad* $uv^{\mathrm T}$ (§B.1) in the hull is effectively bounded above by 1.)

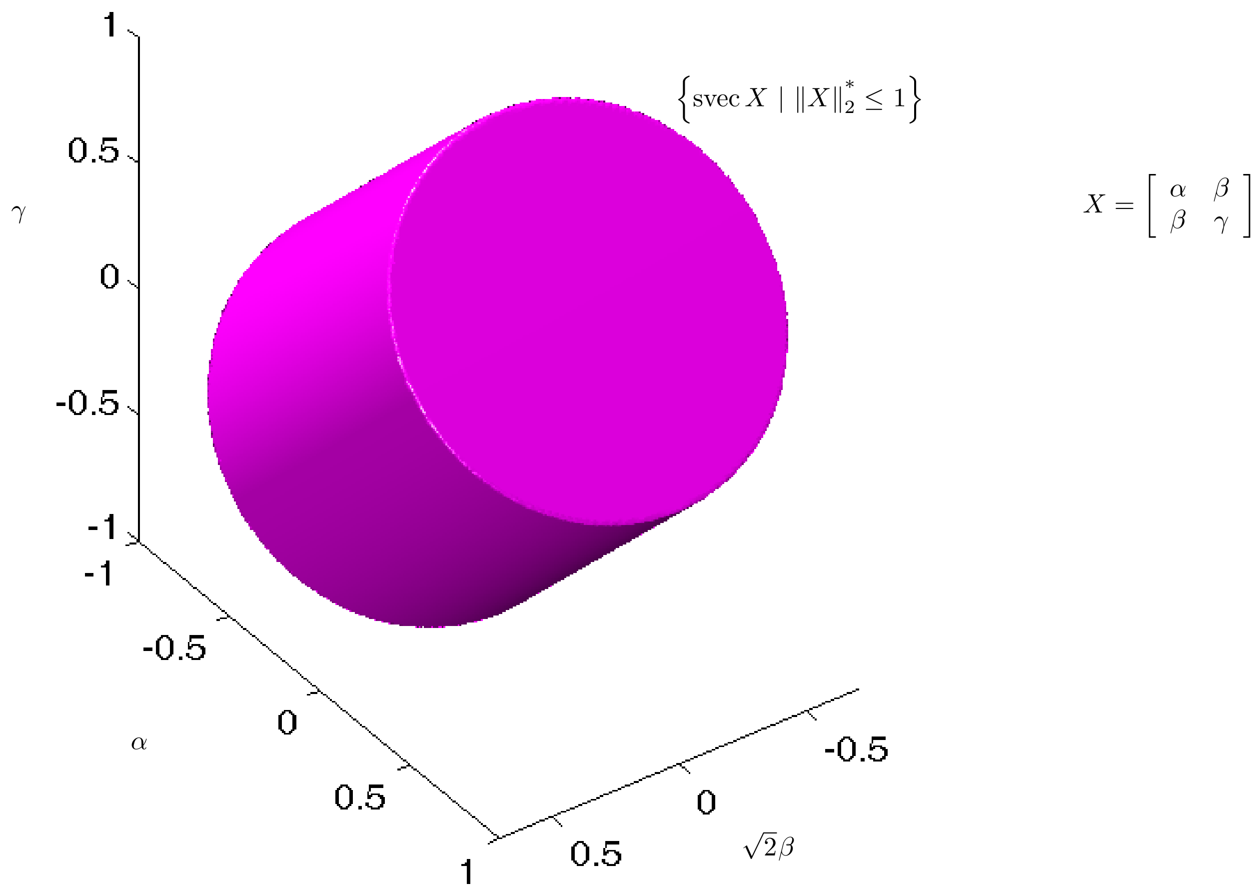


Figure 25: *Nuclear norm* is a sum of singular values; $\|X\|_2^* \triangleq \sum_i \sigma(X)_i$. Nuclear norm ball, in the subspace of 2×2 symmetric matrices, is a truncated cylinder in isometrically isomorphic $\mathbb{R}^3$.

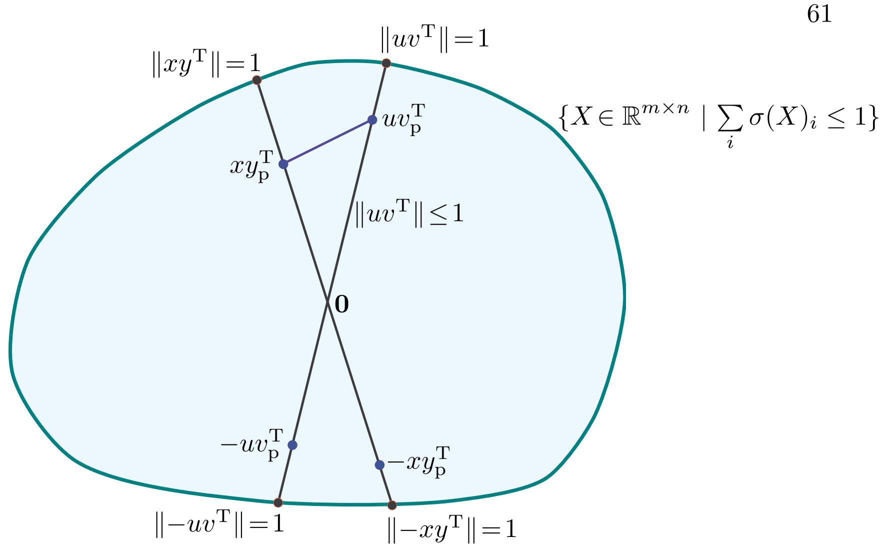


Figure 26: $uv_{\mathrm{p}}^{\mathrm{T}}$ is a convex combination of normalized dyads $\|\pm uv^{\mathrm{T}}\|=1$; similarly for $xy_{\mathrm{p}}^{\mathrm{T}}$. Any point in line segment joining $xy_{\mathrm{p}}^{\mathrm{T}}$ to $uv_{\mathrm{p}}^{\mathrm{T}}$ is expressible as a convex combination of two to four points indicated on boundary.

Proof. ($\Leftarrow$) Suppose $\sum\sigma(X)_i \leq 1$. Decompose $X=U\Sigma V^{\mathrm{T}}$ by SVD (§A.6) where $U=[u_1 \ldots u_{\min\{m,n\}}]\in\mathbb{R}^{m\times\min\{m,n\}}$, $V=[v_1 \ldots v_{\min\{m,n\}}]\in\mathbb{R}^{n\times\min\{m,n\}}$, and whose sum of singular values is $\sum\sigma(X)_i=\operatorname{tr}\Sigma=\kappa\leq 1$. Then we may write $X=\sum\frac{\sigma_i}{\kappa}\sqrt{\kappa}u_i\sqrt{\kappa}v_i^{\mathrm{T}}$ which is a convex combination of dyads each of whose norm does not exceed 1. (Srebro)

($\Rightarrow$) Now suppose we are given a convex combination of dyads $X=\sum\alpha_i\,u_i v_i^{\mathrm{T}}$ such that $\sum\alpha_i=1$, $\alpha_i\geq 0\ \forall i$, and $\|u_i v_i^{\mathrm{T}}\|\leq 1\ \forall i$. Then by triangle inequality for singular values [219, cor.3.4.3] $\sum\sigma(X)_i\leq\sum\sigma(\alpha_i\,u_i v_i^{\mathrm{T}})=\sum\alpha_i\|u_i v_i^{\mathrm{T}}\|\leq\sum\alpha_i$. ♦

Given any particular dyad $uv_{\mathrm{p}}^{\mathrm{T}}$ in the convex hull, because its polar $-uv_{\mathrm{p}}^{\mathrm{T}}$ and every convex combination of the two belong to that hull, then the unique line containing those two points $\pm uv_{\mathrm{p}}^{\mathrm{T}}$ (their affine combination (78)) must intersect the hull's boundary at the normalized dyads $\{\pm uv^{\mathrm{T}} \mid \|uv^{\mathrm{T}}\|=1\}$. Any point formed by convex combination of dyads in the hull must therefore be expressible as a convex combination of dyads on the boundary: Figure **26**,

$$\operatorname{conv}\{uv^{\mathrm{T}} \mid \|uv^{\mathrm{T}}\|\leq 1\,,\ u\in\mathbb{R}^m,\ v\in\mathbb{R}^n\}\equiv\operatorname{conv}\{uv^{\mathrm{T}} \mid \|uv^{\mathrm{T}}\|=1\,,\ u\in\mathbb{R}^m,\ v\in\mathbb{R}^n\} \tag{95}$$

id est, dyads may be normalized and the hull's boundary contains them;

$$\begin{aligned}\partial\{X\in\mathbb{R}^{m\times n} \mid \sum_i\sigma(X)_i\leq 1\} &= \{X\in\mathbb{R}^{m\times n} \mid \sum_i\sigma(X)_i=1\}\\ &\supseteq \{uv^{\mathrm{T}} \mid \|uv^{\mathrm{T}}\|=1\,,\ u\in\mathbb{R}^m,\ v\in\mathbb{R}^n\}\end{aligned} \tag{96}$$

Normalized dyads constitute the set of extreme points (§2.6.0.0.1) of this nuclear norm ball (*confer* Figure **25**) which is, therefore, their convex hull. □

2.3.2.0.3 Exercise. *Convex hull of outer product.*
Describe the interior of a Fantope.
Find the convex hull of nonorthogonal projection matrices (§E.1.1):

$$\{UV^{\mathrm T} \mid U\in\mathbb{R}^{N\times k},\ V\in\mathbb{R}^{N\times k},\ V^{\mathrm T}U=I\} \tag{97}$$

Find the convex hull of nonsymmetric matrices bounded under some norm:

$$\{UV^{\mathrm T} \mid U\in\mathbb{R}^{m\times k},\ V\in\mathbb{R}^{n\times k},\ \|UV^{\mathrm T}\|\leq 1\} \tag{98}$$

▼

2.3.2.0.4 Example. *Permutation polyhedron.* [217] [335] [277]
A *permutation matrix* Ξ is formed by interchanging rows and columns of Identity matrix I. Since Ξ is square and $\Xi^{\mathrm T}\Xi=I$, the set of all permutation matrices is a *proper subset* of the nonconvex *manifold* of orthogonal matrices (§B.5). In fact, the only orthogonal matrices having all nonnegative entries are permutations of the Identity:

$$\Xi^{-1}=\Xi^{\mathrm T}, \qquad \Xi\geq\mathbf{0} \tag{99}$$

And the only positive semidefinite permutation matrix is the Identity. [350, §6.5 prob.20]

Regarding the permutation matrices as a set of points in Euclidean space, its convex hull is a bounded polyhedron (§2.12) described (Birkhoff, 1946)

$$\operatorname{conv}\{\Xi=\Pi_i(I\in\mathbb{S}^n)\in\mathbb{R}^{n\times n},\ i=1\ldots n!\}=\{X\in\mathbb{R}^{n\times n} \mid X^{\mathrm T}\mathbf{1}=\mathbf{1},\ X\mathbf{1}=\mathbf{1},\ X\geq\mathbf{0}\} \tag{100}$$

where Π_i is a linear operator here representing the i^{th} permutation. This polyhedral hull, whose $n!$ vertices are the permutation matrices, is also known as the set of *doubly stochastic matrices.* The permutation matrices are the minimal cardinality (fewest nonzero entries) doubly stochastic matrices. The only orthogonal matrices belonging to this polyhedron are the permutation matrices.

It is remarkable that $n!$ permutation matrices can be described as the extreme points (§2.6.0.0.1) of a bounded polyhedron, of affine dimension $(n-1)^2$, that is itself described by $2n$ equalities ($2n-1$ linearly independent equality constraints in n^2 nonnegative variables). By *Carathéodory's theorem,* conversely, any doubly stochastic matrix can be described as a convex combination of at most $(n-1)^2+1$ permutation matrices. [218, §8.7] [56, thm.1.2.5] This polyhedron, then, can be a device for *relaxing* an integer, combinatorial, or Boolean optimization problem.[2.16] [69] [300, §3.1] □

2.3.2.0.5 Example. *Convex hull of orthonormal matrices.* [28, §1.2]
Consider rank-k matrices $U\in\mathbb{R}^{n\times k}$ such that $U^{\mathrm T}U=I$. These are the orthonormal matrices; a closed bounded submanifold, of all orthogonal matrices, having dimension $nk-\frac{1}{2}k(k+1)$ [53]. Their convex hull is expressed, for $1\leq k\leq n$

$$\begin{aligned}\operatorname{conv}\{U\in\mathbb{R}^{n\times k} \mid U^{\mathrm T}U=I\} &= \{X\in\mathbb{R}^{n\times k} \mid \|X\|_2\leq 1\}\\ &= \{X\in\mathbb{R}^{n\times k} \mid \|X^{\mathrm T}a\|\leq\|a\|\ \ \forall\, a\in\mathbb{R}^n\}\end{aligned} \tag{101}$$

[2.16] *Relaxation* replaces an *objective function* with its *convex envelope* or expands a feasible set to one that is convex. Dantzig first showed in 1951 that, by this device, the so-called *assignment problem* can be formulated as a linear program. [334] [27, §II.5]

By Schur complement (§A.4), the *spectral norm* $\|X\|_2$ constraining largest singular value $\sigma(X)_1$ can be expressed as a semidefinite constraint

$$\|X\|_2 \leq 1 \;\Leftrightarrow\; \begin{bmatrix} I & X \\ X^{\mathrm{T}} & I \end{bmatrix} \succeq 0 \tag{102}$$

because of equivalence $X^{\mathrm{T}}X \preceq I \Leftrightarrow \sigma(X) \preceq \mathbf{1}$ with singular values. (1655) (1539) (1540)

When $k = n$, matrices U are orthogonal and their convex hull is called the *spectral norm ball* which is the set of all *contractions*. [219, p.158] [346, p.313] The orthogonal matrices then constitute the extreme points (§2.6.0.0.1) of this hull. Hull intersection with the nonnegative orthant $\mathbb{R}_+^{n\times n}$ contains the permutation polyhedron (100). □

2.3.3 Conic hull

In terms of a finite-length point list (or set) arranged columnar in $X \in \mathbb{R}^{n\times N}$ (76), its conic hull is expressed

$$\mathcal{K} \;\triangleq\; \operatorname{cone}\{x_\ell\,,\ \ell = 1 \dots N\} \;=\; \operatorname{cone} X \;=\; \{Xa \mid a \succeq 0\} \;\subseteq\; \mathbb{R}^n \tag{103}$$

id est, every nonnegative combination of points from the list. Conic hull of any finite-length list forms a *polyhedral cone* [215, §A.4.3] **(**§2.12.1.0.1; *e.g*, Figure **53**a**)**; the smallest closed convex cone (§2.7.2) that contains the list.

By convention, the aberration [347, §2.1]

$$\operatorname{cone}\emptyset \triangleq \{\mathbf{0}\} \tag{104}$$

Given some arbitrary set $\mathcal{C}$, it is apparent

$$\operatorname{conv}\mathcal{C} \subseteq \operatorname{cone}\mathcal{C} \tag{105}$$

2.3.4 Vertex-description

The conditions in (78), (86), and (103) respectively define an *affine combination, convex combination,* and *conic combination* of elements from the set or list. Whenever a Euclidean body can be described as some hull or span of a set of points, then that representation is loosely called a *vertex-description* and those points are called *generators.*

2.4 Halfspace, Hyperplane

A two-dimensional affine subset is called a *plane.* An $(n-1)$-dimensional affine subset of $\mathbb{R}^n$ is called a *hyperplane.* [325] [215] Every hyperplane partially bounds a *halfspace* (which is convex, but not affine, and the only nonempty convex set in $\mathbb{R}^n$ whose complement is convex and nonempty).

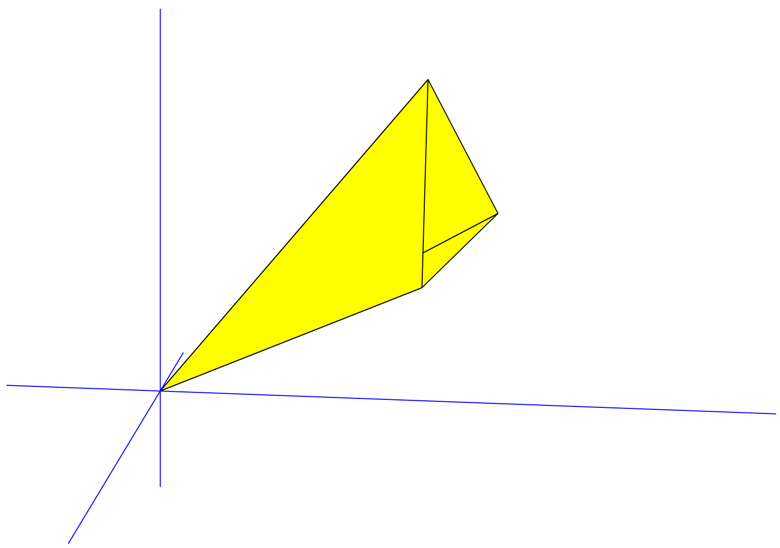


Figure 27: A simplicial cone (§2.12.3.1.1) in $\mathbb{R}^{\mathbf{3}}$ whose boundary is drawn truncated; constructed using $A\in\mathbb{R}^{\mathbf{3}\times\mathbf{3}}$ and $C=\mathbf{0}$ in (286). By the most fundamental definition of a cone (§2.7.1), entire boundary can be constructed from an aggregate of rays emanating exclusively from the origin. Each of three extreme directions corresponds to an edge (§2.6.0.0.3); they are conically, affinely, and linearly independent for this cone. Because this set is polyhedral, exposed directions are in one-to-one correspondence with extreme directions; there are only three. Its extreme directions give rise to what is called a *vertex-description* of this polyhedral cone; simply, the conic hull of extreme directions. Obviously this cone can also be constructed by intersection of three halfspaces; hence the equivalent *halfspace-description.*

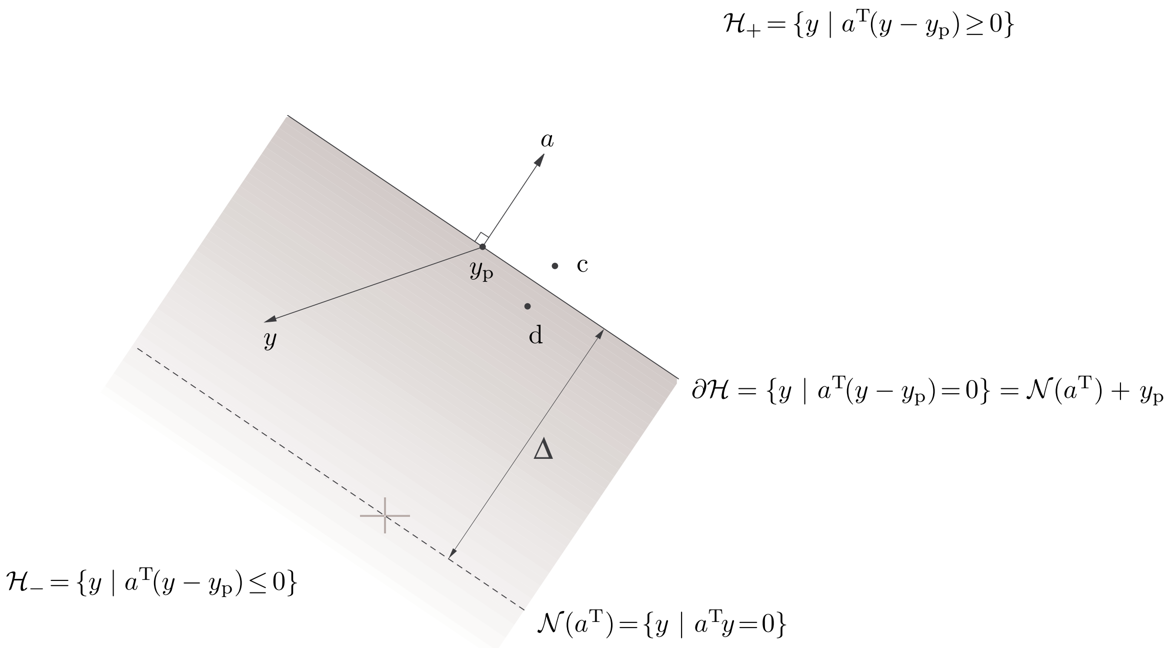


Figure 28: Hyperplane illustrated $\partial\mathcal{H}$ is a line partially bounding halfspaces $\mathcal{H}_-$ and $\mathcal{H}_+$ in $\mathbb{R}^{\mathbf{2}}$. Shaded is a rectangular piece of semiinfinite $\mathcal{H}_-$ with respect to which vector a is outward-normal to bounding hyperplane; vector a is inward-normal with respect to $\mathcal{H}_+$. Halfspace $\mathcal{H}_-$ contains nullspace $\mathcal{N}(a^{\mathrm{T}})$ (dashed line through origin) because $a^{\mathrm{T}}y_{\mathrm{p}} > 0$. Hyperplane, halfspace, and nullspace are each drawn truncated. Points c and d are equidistant from hyperplane, and vector $\mathrm{c} - \mathrm{d}$ is normal to it. Δ is distance from origin to hyperplane.

2.4.1 Halfspaces $\mathcal{H}_+$ and $\mathcal{H}_-$

Euclidean space $\mathbb{R}^n$ is partitioned in two by any hyperplane $\partial\mathcal{H}$; *id est*, $\mathcal{H}_- + \mathcal{H}_+ = \mathbb{R}^n$. The resulting (closed convex) halfspaces, both partially bounded by $\partial\mathcal{H}$, may be described

$$\mathcal{H}_- = \{y \mid a^{\mathrm T}y \leq b\} = \{y \mid a^{\mathrm T}(y-y_{\mathrm p}) \leq 0\} \subset \mathbb{R}^n \tag{106}$$

$$\mathcal{H}_+ = \{y \mid a^{\mathrm T}y \geq b\} = \{y \mid a^{\mathrm T}(y-y_{\mathrm p}) \geq 0\} \subset \mathbb{R}^n \tag{107}$$

where nonzero vector $a\in\mathbb{R}^n$ is an *outward-normal* to the hyperplane partially bounding $\mathcal{H}_-$ while an *inward-normal* with respect to $\mathcal{H}_+$. For any vector $y-y_{\mathrm p}$ that makes an obtuse angle with normal a , vector y will lie in the halfspace $\mathcal{H}_-$ on one side (shaded in Figure **28**) of the hyperplane while acute angles denote y in $\mathcal{H}_+$ on the other side.

An equivalent more intuitive representation of a halfspace comes about when we consider all the points in $\mathbb{R}^n$ closer to point $\mathrm d$ than to point $\mathrm c$ or equidistant, in the Euclidean sense; from Figure **28**,

$$\mathcal{H}_- = \{y \mid \|y-\mathrm d\| \leq \|y-\mathrm c\|\} \tag{108}$$

This representation, in terms of proximity, is resolved with the more conventional representation of a halfspace (106) by squaring both sides of the inequality in (108);

$$\mathcal{H}_- = \left\{y \mid (\mathrm c-\mathrm d)^{\mathrm T}y \leq \frac{\|\mathrm c\|^2-\|\mathrm d\|^2}{2}\right\} = \left\{y \mid (\mathrm c-\mathrm d)^{\mathrm T}\left(y-\frac{\mathrm c+\mathrm d}{2}\right) \leq 0\right\} \tag{109}$$

2.4.1.1 PRINCIPLE 1: Halfspace-description of convex sets

The most fundamental principle in convex geometry follows from the *geometric Hahn-Banach theorem* [266, §5.12] [19, §1] [143, §I.1.2] which guarantees any closed convex set to be an intersection of halfspaces.

2.4.1.1.1 Theorem. *Halfspaces.* [215, §A.4.2b] [43, §2.4]
A closed convex set in $\mathbb{R}^n$ is equivalent to the intersection of all halfspaces that contain it. ⋄

Intersection of multiple halfspaces in $\mathbb{R}^n$ may be represented using a matrix constant A

$$\bigcap_i \mathcal{H}_{i-} = \{y \mid A^{\mathrm T}y \preceq b\} = \{y \mid A^{\mathrm T}(y-y_{\mathrm p}) \preceq 0\} \tag{110}$$

$$\bigcap_i \mathcal{H}_{i+} = \{y \mid A^{\mathrm T}y \succeq b\} = \{y \mid A^{\mathrm T}(y-y_{\mathrm p}) \succeq 0\} \tag{111}$$

where b is now a vector, and the i^{th} column of A is normal to a hyperplane $\partial\mathcal{H}_i$ partially bounding $\mathcal{H}_i$. By the *halfspaces theorem,* intersections like this can describe interesting convex Euclidean bodies such as polyhedra and cones, giving rise to the term *halfspace-description.*

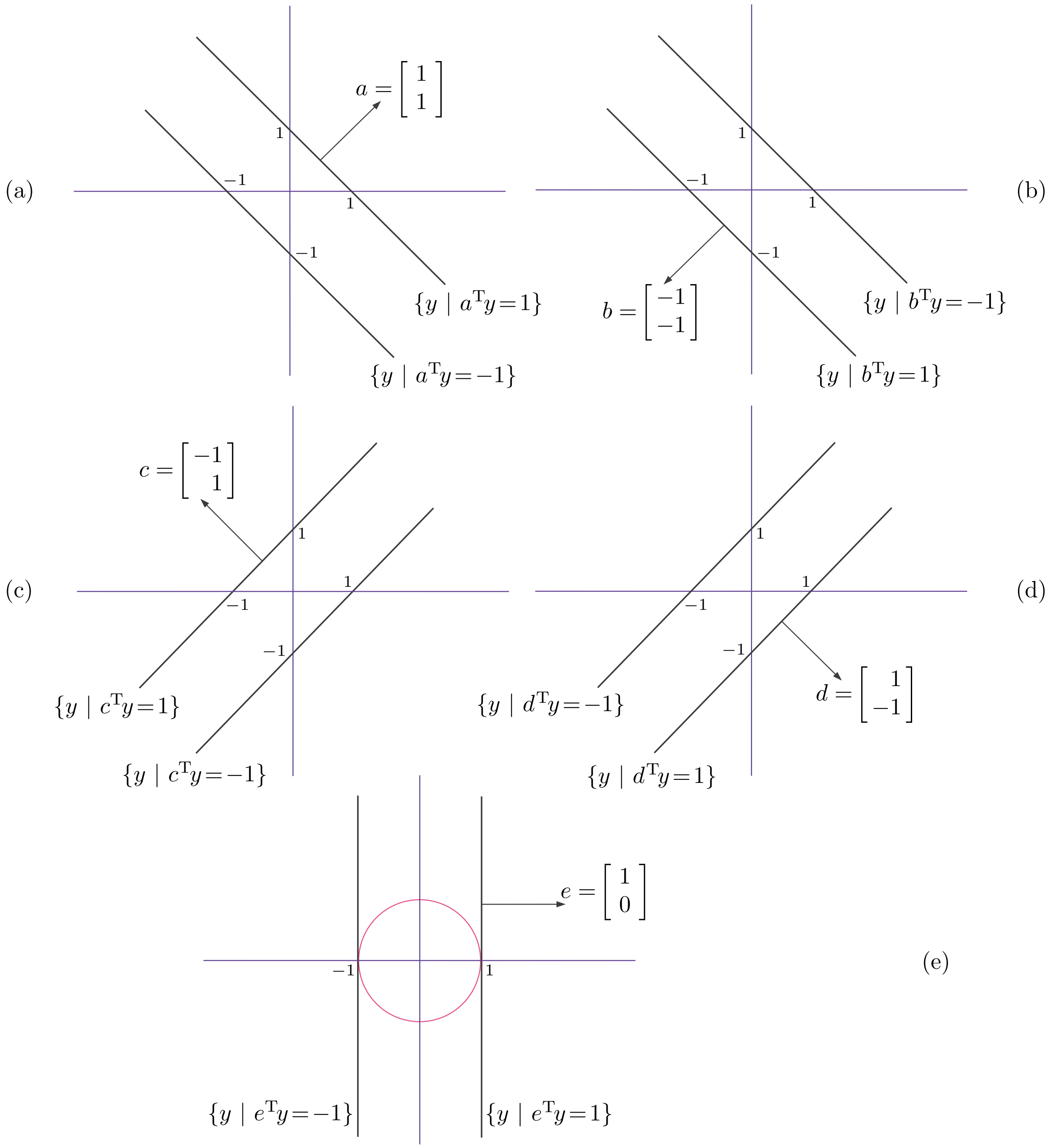


Figure 29: **(a)-(d)** Hyperplanes in $\mathbb{R}^2$ (truncated) redundantly emphasize: hyperplane movement opposite to its normal direction minimizes vector inner-product. This concept is exploited to attain analytical solution of linear programs by visual inspection; *e.g*, §2.4.2.6.2, §2.5.1.2.2, §3.4.0.0.2, [63, exer.4.8-exer.4.20]. Each graph is also interpretable as contour plot of a real affine function of two variables as in Figure **77**. **(e)** $|\beta|/\|\alpha\|$ from $\partial\mathcal{H} = \{x \mid \alpha^{\mathrm{T}}x = \beta\}$ represents radius of hypersphere about $\mathbf{0}$ supported by any hyperplane with same ratio |inner product|/norm.

2.4.2 Hyperplane $\partial\mathcal{H}$ representations

Every hyperplane $\partial\mathcal{H}$ is an affine set parallel to an $(n-1)$-dimensional subspace of $\mathbb{R}^n$; it is itself a subspace if and only if it contains the origin.

$$\dim \partial\mathcal{H} = n-1 \tag{112}$$

so a hyperplane is a point in $\mathbb{R}$, a line in $\mathbb{R}^{\mathbf{2}}$, a plane in $\mathbb{R}^{\mathbf{3}}$, and so on. Every hyperplane can be described as the intersection of complementary halfspaces; [325, §19]

$$\partial\mathcal{H} \;=\; \mathcal{H}_- \cap \mathcal{H}_+ \;=\; \{y \mid a^{\mathrm{T}}y \leq b\;,\; a^{\mathrm{T}}y \geq b\} \;=\; \{y \mid a^{\mathrm{T}}y = b\} \tag{113}$$

a halfspace-description. Assuming *normal* $a\in\mathbb{R}^n$ to be nonzero, then any hyperplane in $\mathbb{R}^n$ can be described as the solution set to vector equation $a^{\mathrm{T}}y=b$ (illustrated in Figure **28** and Figure **29** for $\mathbb{R}^{\mathbf{2}}$);

$$\partial\mathcal{H} \triangleq \{y \mid a^{\mathrm{T}}y = b\} = \{y \mid a^{\mathrm{T}}(y-y_{\mathrm{p}}) = 0\} = \{Z\xi + y_{\mathrm{p}} \mid \xi\in\mathbb{R}^{n-1}\} \subset \mathbb{R}^n \tag{114}$$

All solutions y constituting the hyperplane are *offset* from the nullspace of a^{T} by the same constant vector $y_{\mathrm{p}}\in\mathbb{R}^n$ that is any particular solution to $a^{\mathrm{T}}y=b$; *id est*,

$$y = Z\xi + y_{\mathrm{p}} \tag{115}$$

where the columns of $Z\in\mathbb{R}^{n\times n-1}$ constitute a basis for $\mathcal{N}(a^{\mathrm{T}})=\{x\in\mathbb{R}^n \mid a^{\mathrm{T}}x=0\}$ the nullspace.[2.17]

Conversely, given any point y_{p} in $\mathbb{R}^n$, the unique hyperplane containing it having normal a is the affine set $\partial\mathcal{H}$ (114) where b equals $a^{\mathrm{T}}y_{\mathrm{p}}$ and where a basis for $\mathcal{N}(a^{\mathrm{T}})$ is arranged in Z columnar. Hyperplane dimension is apparent from dimension of Z; that hyperplane is parallel to the span of its columns.

2.4.2.0.1 Exercise. *Hyperplane scaling.*
Given normal y, draw a hyperplane $\{x\in\mathbb{R}^{\mathbf{2}} \mid x^{\mathrm{T}}y=1\}$. Suppose $z=\frac{1}{2}y$. On the same plot, draw the hyperplane $\{x\in\mathbb{R}^{\mathbf{2}} \mid x^{\mathrm{T}}z=1\}$. Now suppose $z=2y$, then draw the last hyperplane again with this new z. What is the apparent effect of scaling normal y?

▼

2.4.2.0.2 Example. *Distance from origin to hyperplane.*
Given the (shortest) distance $\Delta\in\mathbb{R}_+$ from the origin to a hyperplane having normal vector a, we can find its representation $\partial\mathcal{H}$ by dropping a perpendicular. The point thus found is the orthogonal projection of the origin on $\partial\mathcal{H}$ (§E.5.0.0.5), equal to $a\Delta/\|a\|$ if the origin is known *a priori* to belong to halfspace $\mathcal{H}_-$ (Figure **28**), or equal to $-a\Delta/\|a\|$ if the origin belongs to halfspace $\mathcal{H}_+$; *id est*, when $\mathcal{H}_- \ni \mathbf{0}$

$$\partial\mathcal{H} \;=\; \left\{y \mid a^{\mathrm{T}}(y - a\Delta/\|a\|) = 0\right\} \;=\; \left\{y \mid a^{\mathrm{T}}y = \|a\|\Delta\right\} \tag{116}$$

[2.17] We will find this expression for y in terms of nullspace of a^{T} (more generally, of matrix A (143)) to be a useful trick (a practical device) for eliminating affine equality constraints, much as we did here.

or when $\mathcal{H}_+ \ni \mathbf{0}$

$$\partial\mathcal{H} \;=\; \left\{y \mid a^{\mathrm T}(y + a\Delta/\|a\|) = 0\right\} \;=\; \left\{y \mid a^{\mathrm T}y = -\|a\|\Delta\right\} \tag{117}$$

Knowledge of only distance Δ and normal a thus introduces ambiguity into the hyperplane representation. □

2.4.2.1 Matrix variable

Any halfspace in $\mathbb{R}^{mn}$ may be represented using a matrix variable. For variable $Y \in \mathbb{R}^{m\times n}$, given constants $A \in \mathbb{R}^{m\times n}$ and $b = \langle A\,,Y_{\mathrm p}\rangle \in \mathbb{R}$

$$\mathcal{H}_- \;=\; \{Y \in \mathbb{R}^{mn} \mid \langle A\,,Y\rangle \le b\} \;=\; \{Y \in \mathbb{R}^{mn} \mid \langle A\,,Y - Y_{\mathrm p}\rangle \le 0\} \tag{118}$$

$$\mathcal{H}_+ \;=\; \{Y \in \mathbb{R}^{mn} \mid \langle A\,,Y\rangle \ge b\} \;=\; \{Y \in \mathbb{R}^{mn} \mid \langle A\,,Y - Y_{\mathrm p}\rangle \ge 0\} \tag{119}$$

Recall vector inner-product from §2.2: $\langle A\,,Y\rangle = \operatorname{tr}(A^{\mathrm T}Y) = \operatorname{vec}(A)^{\mathrm T}\operatorname{vec}(Y)$.

Hyperplanes in $\mathbb{R}^{mn}$ may, of course, also be represented using matrix variables.

$$\partial\mathcal{H} \;=\; \{Y \mid \langle A\,,Y\rangle = b\} \;=\; \{Y \mid \langle A\,,Y - Y_{\mathrm p}\rangle = 0\} \subset \mathbb{R}^{mn} \tag{120}$$

Vector a from Figure **28** is normal to the hyperplane illustrated. Likewise, nonzero vectorized matrix A is normal to hyperplane $\partial\mathcal{H}$;

$$A \perp \partial\mathcal{H} \text{ in } \mathbb{R}^{mn} \tag{121}$$

2.4.2.2 Vertex-description of hyperplane

Any hyperplane in $\mathbb{R}^n$ may be described as affine hull of a minimal set of points $\{x_\ell \in \mathbb{R}^n,\ \ell = 1 \ldots n\}$ arranged columnar in a matrix $X \in \mathbb{R}^{n\times n}$: (78)

$$\begin{aligned}
\partial\mathcal{H} &= \operatorname{aff}\{x_\ell \in \mathbb{R}^n,\ \ell = 1\ldots n\}\ , && \dim\operatorname{aff}\{x_\ell\ \forall\,\ell\} = n-1\\
&= \operatorname{aff} X\ , && \dim\operatorname{aff} X = n-1\\
&= x_1 + \mathcal{R}\{x_\ell - x_1\ ,\ \ell = 2\ldots n\}\ , && \dim\mathcal{R}\{x_\ell - x_1\ ,\ \ell = 2\ldots n\} = n-1\\
&= x_1 + \mathcal{R}(X - x_1\mathbf{1}^{\mathrm T})\ , && \dim\mathcal{R}(X - x_1\mathbf{1}^{\mathrm T}) = n-1
\end{aligned} \tag{122}$$

where

$$\mathcal{R}(A) \;=\; \{Ax \mid \forall\,x\} \qquad (142)$$

2.4.2.3 Affine independence, minimal set

For any particular affine set, a *minimal set* of points constituting its vertex-description is an affinely independent generating set and *vice versa.*

Arbitrary given points $\{x_i \in \mathbb{R}^n,\ i = 1\ldots N\}$ are *affinely independent* (a.i.) if and only if, over all $\zeta \in \mathbb{R}^N \ni \zeta^{\mathrm T}\mathbf{1} = 1,\ \zeta_k = 0 \in \mathbb{R}$ (*confer* §2.1.2)

$$x_i\,\zeta_i + \cdots + x_j\,\zeta_j - x_k = \mathbf{0}\,, \qquad i \ne \cdots \ne j \ne k = 1\ldots N \tag{123}$$

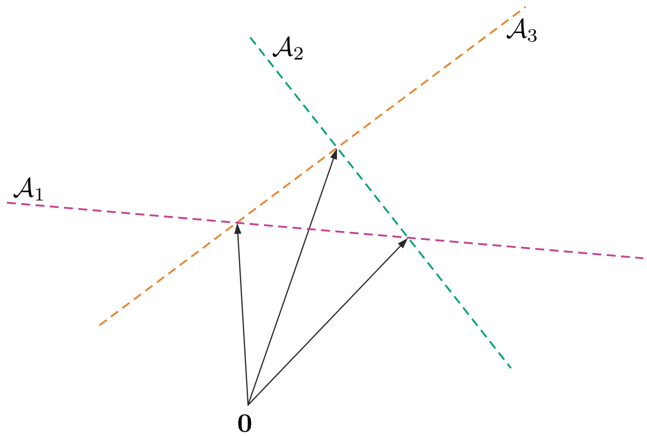


Figure 30: Of three points illustrated, any one particular point does not belong to affine hull $\mathcal{A}_i$ ($i \in 1,2,3,$ each drawn truncated) of points remaining. Three corresponding vectors in $\mathbb{R}^{\mathbf{2}}$ are, therefore, affinely independent (but neither linearly or conically independent).

has no solution ζ; in words, iff no point from the given set can be expressed as an affine combination of those remaining. We deduce

$$\text{l.i.} \;\Rightarrow\; \text{a.i.} \tag{124}$$

Consequently, $\{x_i\,,\; i=1\ldots N\}$ is an affinely independent set if and only if $\{x_i - x_1\,,\; i=2\ldots N\}$ is a linearly independent (l.i.) set. [221, §3] (Figure **30**) This is equivalent to the property that the columns of $\begin{bmatrix} X \\ \mathbf{1}^{\mathrm{T}} \end{bmatrix}$ (for $X \in \mathbb{R}^{n\times N}$ as in (76)) form a linearly independent set. [215, §A.1.3]

Two nontrivial affine subsets are affinely independent iff their intersection is empty $\{\emptyset\}$ or, analogously to subspaces, they intersect only at a point.

2.4.2.4 Preservation of affine independence

Independence in the linear (§2.1.2.1), affine, and conic (§2.10.1) senses can be preserved under linear transformation. Suppose a matrix $X \in \mathbb{R}^{n\times N}$ (76) holds an affinely independent set in its columns. Consider a transformation on the domain of such matrices

$$T(X) : \mathbb{R}^{n\times N} \to \mathbb{R}^{n\times N} \;\triangleq\; XY \tag{125}$$

where fixed matrix $Y \triangleq [\,y_1 \;\; y_2 \;\cdots\; y_N\,] \in \mathbb{R}^{N\times N}$ represents linear operator T. Affine independence of $\{Xy_i \in \mathbb{R}^n,\; i=1\ldots N\}$ demands (by definition (123)) there exist no solution $\zeta \in \mathbb{R}^N \ni \zeta^{\mathrm{T}}\mathbf{1} = 1,\; \zeta_k = 0,$ to

$$Xy_i\,\zeta_i + \cdots + Xy_j\,\zeta_j - Xy_k = \mathbf{0}\,, \qquad i \neq \cdots \neq j \neq k = 1\ldots N \tag{126}$$

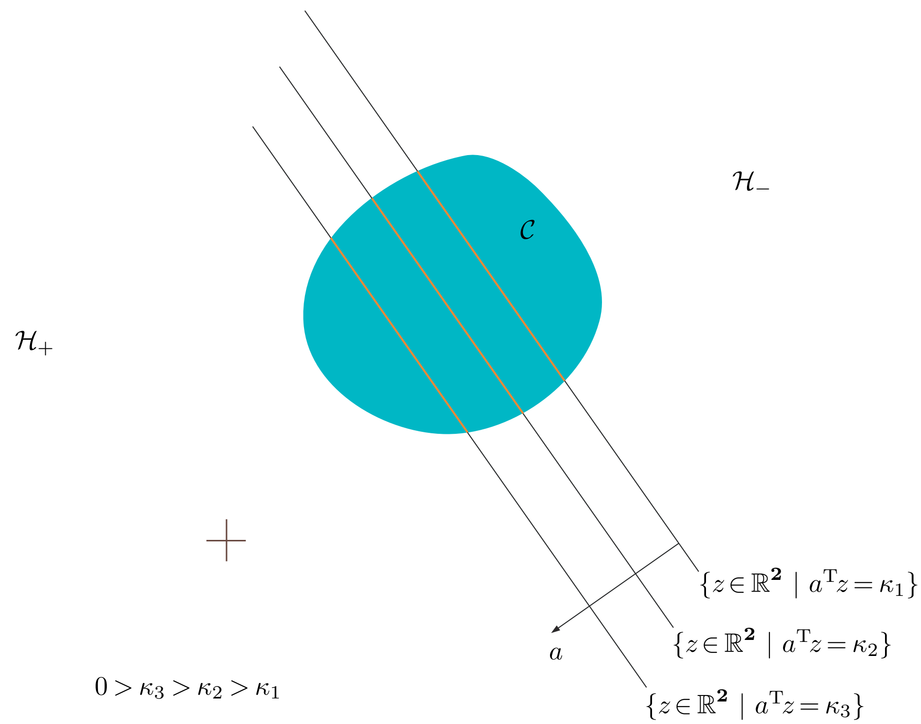


Figure 31: (*confer* Figure **77**) Each linear contour, of equal inner product in vector z with normal a, represents i^{th} hyperplane in $\mathbb{R}^2$ parametrized by scalar κ_i. Inner product κ_i increases in direction of normal a. In convex set $\mathcal{C} \subset \mathbb{R}^2$, i^{th} line segment $\{z \in \mathcal{C} \mid a^{\mathrm{T}}z = \kappa_i\}$ represents intersection with hyperplane. (Cartesian axes for reference.)

By factoring out X, we see that is ensured by affine independence of $\{y_i \in \mathbb{R}^N\}$ and by $\mathcal{R}(Y) \cap \mathcal{N}(X) = \mathbf{0}$ where

$$\mathcal{N}(A) = \{x \mid Ax = \mathbf{0}\} \qquad (143)$$

2.4.2.5 Affine maps

Affine transformations preserve affine hulls. Given any affine mapping T of vector spaces and some arbitrary set $\mathcal{C}$ [325, p.8]

$$\operatorname{aff}(T\mathcal{C}) = T(\operatorname{aff}\mathcal{C}) \qquad (127)$$

2.4.2.6 PRINCIPLE 2: Supporting hyperplane

The second most fundamental principle of convex geometry also follows from the *geometric Hahn-Banach theorem* [266, §5.12] [19, §1] that guarantees existence of at least one

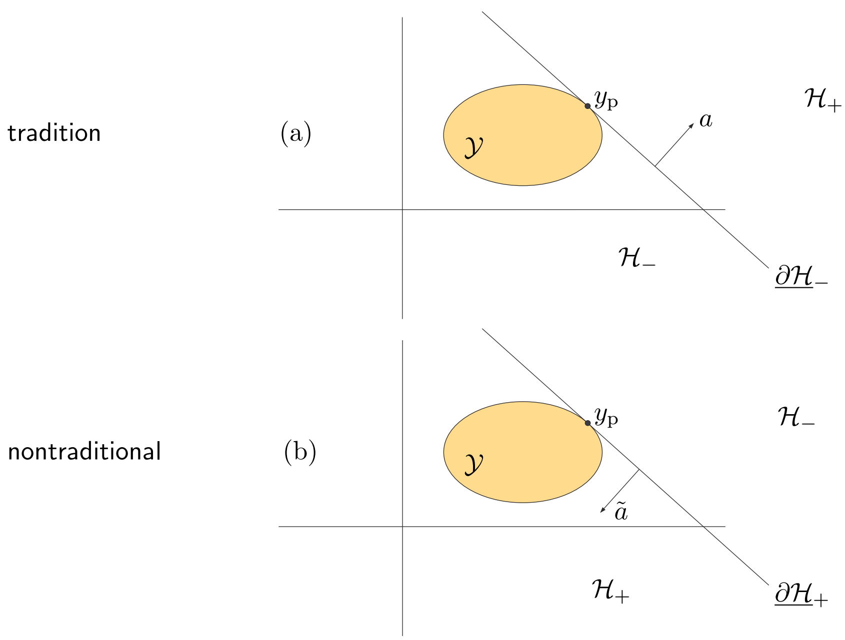


Figure 32: **(a)** Hyperplane $\underline{\partial\mathcal{H}}_-$ (128) supporting closed set $\mathcal{Y} \subset \mathbb{R}^{\mathbf{2}}$. Vector a is inward-normal to hyperplane with respect to halfspace $\mathcal{H}_+$, but outward-normal with respect to set $\mathcal{Y}$. A supporting hyperplane can be considered the limit of an increasing sequence in the normal-direction like that in Figure **31**. **(b)** Hyperplane $\underline{\partial\mathcal{H}}_+$ nontraditionally supporting $\mathcal{Y}$. Vector $\tilde{a}$ is inward-normal to hyperplane now with respect to both halfspace $\mathcal{H}_+$ and set $\mathcal{Y}$. Tradition [215] [325] recognizes only positive normal polarity in support function $\sigma_{\mathcal{Y}}$ as in (129); *id est*, normal a , figure (a). But both interpretations of supporting hyperplane are useful.

hyperplane in $\mathbb{R}^n$ supporting a full-dimensional convex set[2.18] at each point on its boundary.

The partial boundary $\partial\mathcal{H}$ of a halfspace that contains arbitrary set $\mathcal{Y}$ is called a *supporting hyperplane* $\underline{\partial\mathcal{H}}$ to $\mathcal{Y}$ when the hyperplane contains at least one point of $\overline{\mathcal{Y}}$. [325, §11]

2.4.2.6.1 Definition. *Supporting hyperplane* $\underline{\partial\mathcal{H}}$.
Assuming set $\mathcal{Y}$ and some normal $a\neq\mathbf{0}$ reside in opposite halfspaces[2.19] (Figure **32**a), then a hyperplane supporting $\mathcal{Y}$ at point $y_{\rm p}\in\partial\mathcal{Y}$ is described

$$\underline{\partial\mathcal{H}}_- = \left\{y \mid a^{\rm T}(y-y_{\rm p})=0\,,\quad y_{\rm p}\in\overline{\mathcal{Y}}\,,\quad a^{\rm T}(z-y_{\rm p})\leq 0\quad\forall\, z\in\overline{\mathcal{Y}}\right\} \tag{128}$$

Given only normal a, the hyperplane supporting $\mathcal{Y}$ is equivalently described

$$\underline{\partial\mathcal{H}}_- = \left\{y \mid a^{\rm T}y=\sup\{a^{\rm T}z\mid z\in\mathcal{Y}\}\right\} \tag{129}$$

where real function

$$\sigma_{\mathcal{Y}}(a) = \sup\{a^{\rm T}z\mid z\in\mathcal{Y}\} \qquad (554)$$

is called the *support function* for $\mathcal{Y}$.

Another equivalent but nontraditional representation[2.20] for a supporting hyperplane is obtained by reversing polarity of normal a; (1772)

$$\begin{aligned}\underline{\partial\mathcal{H}}_+ &= \left\{y \mid \tilde{a}^{\rm T}(y-y_{\rm p})=0\,,\quad y_{\rm p}\in\overline{\mathcal{Y}}\,,\quad \tilde{a}^{\rm T}(z-y_{\rm p})\geq 0\quad\forall\, z\in\overline{\mathcal{Y}}\right\}\\ &= \left\{y \mid \tilde{a}^{\rm T}y=-\inf\{\tilde{a}^{\rm T}z\mid z\in\mathcal{Y}\} = \sup\{-\tilde{a}^{\rm T}z\mid z\in\mathcal{Y}\}\right\}\end{aligned} \tag{130}$$

where normal $\tilde{a}$ and set $\mathcal{Y}$ both now reside in $\mathcal{H}_+$ (Figure **32**b).

When a supporting hyperplane contains only a single point of $\overline{\mathcal{Y}}$, that hyperplane is termed *strictly supporting*.[2.21] △

A full-dimensional set that has a supporting hyperplane at every point on its boundary, conversely, is convex. A convex set $\mathcal{C}\subset\mathbb{R}^n$, for example, can be expressed as the intersection of all halfspaces partially bounded by hyperplanes supporting it; *videlicet*, [266, p.135]

$$\overline{\mathcal{C}} = \bigcap_{a\in\mathbb{R}^n}\left\{y \mid a^{\rm T}y\leq\sigma_{\mathcal{C}}(a)\right\} \tag{131}$$

by the *halfspaces theorem* (§2.4.1.1.1).

There is no geometric difference between supporting hyperplane $\underline{\partial\mathcal{H}}_+$ or $\underline{\partial\mathcal{H}}_-$ or $\underline{\partial\mathcal{H}}$ and[2.22] an ordinary hyperplane $\partial\mathcal{H}$ coincident with them.

[2.18] It is customary to speak of a hyperplane supporting set $\mathcal{C}$ but not containing $\mathcal{C}$; called *nontrivial support*. [325, p.100] Hyperplanes in support of lower-dimensional bodies are admitted.

[2.19] Normal a belongs to $\mathcal{H}_+$ by definition.

[2.20] useful for constructing the *dual cone*; *e.g*, Figure **59**b. Tradition would instead have us construct the *polar cone*; which is, the negative dual cone.

[2.21] Rockafellar terms a strictly supporting hyperplane *tangent* to $\mathcal{Y}$ if it is unique there; [325, §18, p.169] a definition we do not adopt because our only criterion for tangency is intersection exclusively with a relative boundary. Hiriart-Urruty & Lemaréchal [215, p.44] (*confer* [325, p.100]) do not demand any tangency of a supporting hyperplane.

[2.22] If vector-normal polarity is unimportant, we may instead signify a supporting hyperplane by $\underline{\partial\mathcal{H}}$.

2.4.2.6.2 Example. *Minimization over hypercube.*
Consider minimization of a linear function over a *hypercube,* given vector c

$$\begin{array}{ll}\underset{x}{\text{minimize}} & c^{\mathrm{T}}x\\ \text{subject to} & -\mathbf{1}\preceq x\preceq\mathbf{1}\end{array}\tag{132}$$

This convex optimization problem is called a *linear program*[2.23] because the *objective*[2.24] of minimization $c^{\mathrm{T}}x$ is a linear function of variable x and the constraints describe a *polyhedron* (intersection of a finite number of halfspaces and hyperplanes).

Any vector x satisfying the constraints is called a *feasible solution.* Applying graphical concepts from Figure **29**, Figure **31**, and Figure **32**, $x^\star = -\operatorname{sgn}(c)$ is an *optimal solution* to this minimization problem but is not necessarily unique. It generally holds for optimization problem solutions:

$$\text{optimal} \;\Rightarrow\; \text{feasible}\tag{133}$$

Because an optimal solution always exists at a hypercube vertex (§2.6.1.0.1) regardless of value of nonzero vector c in (132) [98, p.158] [16, p.2], mathematicians see this geometry as a means to relax a discrete problem (whose desired solution is integer or combinatorial, *confer* Example 4.2.3.1.1). [255, §3.1] [256] □

2.4.2.6.3 Exercise. *Unbounded below.*
Suppose instead we minimize over the unit hypersphere in Example 2.4.2.6.2; $\|x\|\leq 1$. What is an expression for optimal solution now? Is that program still linear?

Now suppose minimization of absolute value in (132). Are the following programs equivalent for some arbitrary real convex set $\mathcal{C}$? (*confer* (516))

$$\begin{array}{ll}\underset{x\in\mathbb{R}}{\text{minimize}} & |x|\\ \text{subject to} & -1\leq x\leq 1\\ & x\in\mathcal{C}\end{array} \quad\equiv\quad \begin{array}{ll}\underset{\alpha\;,\;\beta}{\text{minimize}} & \alpha+\beta\\ \text{subject to} & 1\geq\beta\geq 0\\ & 1\geq\alpha\geq 0\\ & \alpha-\beta\in\mathcal{C}\end{array}\tag{134}$$

Many optimization problems of interest and some methods of solution require nonnegative variables. The method illustrated below splits a variable into parts; $x=\alpha-\beta$ (extensible to vectors). Under what conditions on vector a and scalar b is an optimal solution $x^\star$ negative infinity?

$$\begin{array}{ll}\underset{\alpha\in\mathbb{R}\;,\;\beta\in\mathbb{R}}{\text{minimize}} & \alpha-\beta\\ \text{subject to} & \beta\geq 0\\ & \alpha\geq 0\\ & a^{\mathrm{T}}\begin{bmatrix}\alpha\\ \beta\end{bmatrix}=b\end{array}\tag{135}$$

Minimization of the *objective function* entails maximization of β. ▼

[2.23] The term *program* has its roots in economics. It was originally meant with regard to a plan or to efficient organization or systematization of some industrial process. [98, §2]

[2.24] The *objective* is the function that is argument to minimization or maximization.

2.4.2.7 PRINCIPLE 3: Separating hyperplane

The third most fundamental principle of convex geometry again follows from the *geometric Hahn-Banach theorem* [266, §5.12] [19, §1] [143, §I.1.2] that guarantees existence of a hyperplane separating two nonempty convex sets in $\mathbb{R}^n$ whose relative interiors are nonintersecting. *Separation* intuitively means each set belongs to a halfspace on an opposing side of the hyperplane. There are two cases of interest:

1) If the two sets intersect only at their relative boundaries (§2.1.7.2), then there exists a separating hyperplane $\underline{\partial\mathcal{H}}$ containing the intersection but containing no points relatively interior to either set. If at least one of the two sets is open, conversely, then the existence of a separating hyperplane implies the two sets are nonintersecting. [63, §2.5.1]

2) A *strictly separating hyperplane* $\partial\mathcal{H}$ intersects the closure of neither set; its existence is guaranteed when intersection of the closures is empty and at least one set is bounded. [215, §A.4.1]

2.4.3 Angle between hyperspaces

Given halfspace-descriptions, *dihedral* angle between hyperplanes or halfspaces is defined as the angle between their defining normals. Given normals a and b respectively describing $\partial\mathcal{H}_a$ and $\partial\mathcal{H}_b$, for example

$$\measuredangle(\partial\mathcal{H}_a\,,\partial\mathcal{H}_b) \triangleq \arccos\left(\frac{\langle a\,,\,b\rangle}{\|a\|\,\|b\|}\right) \text{ radians} \tag{136}$$

2.5 Subspace representations

There are two common forms of expression for Euclidean subspaces, both coming from elementary linear algebra: *range form* $\mathcal{R}$ and *nullspace form* $\mathcal{N}$; `a.k.a,` vertex-description and halfspace-description respectively.

The fundamental vector subspaces associated with a matrix $A\in\mathbb{R}^{m\times n}$ [348, §3.1] are ordinarily related by orthogonal complement

$$\mathcal{R}(A^{\mathrm{T}}) \perp \mathcal{N}(A)\,, \qquad \mathcal{N}(A^{\mathrm{T}}) \perp \mathcal{R}(A) \tag{137}$$

$$\mathcal{R}(A^{\mathrm{T}}) \oplus \mathcal{N}(A) = \mathbb{R}^n\,, \qquad \mathcal{N}(A^{\mathrm{T}}) \oplus \mathcal{R}(A) = \mathbb{R}^m \tag{138}$$

and of dimension:

$$\dim\mathcal{R}(A^{\mathrm{T}}) \;=\; \dim\mathcal{R}(A) \;=\; \operatorname{rank} A \;\leq\; \min\{m\,,n\} \tag{139}$$

with complementarity (`a.k.a` *conservation of dimension*)

$$\dim\mathcal{N}(A) = n - \operatorname{rank} A\;, \qquad \dim\mathcal{N}(A^{\mathrm{T}}) = m - \operatorname{rank} A \tag{140}$$

These equations (137)-(140) comprise the *fundamental theorem of linear algebra.* [348, p.95, p.138]

From these four fundamental subspaces, the rowspace and range identify one form of subspace description (*vertex-description* (§2.3.4) or *range form*)

$$\mathcal{R}(A^{\mathrm{T}}) \triangleq \operatorname{span} A^{\mathrm{T}} = \{A^{\mathrm{T}}y \mid y\in\mathbb{R}^m\} = \{x\in\mathbb{R}^n \mid A^{\mathrm{T}}y=x\,,\;\; y\in\mathcal{R}(A)\} \tag{141}$$

$$\mathcal{R}(A) \triangleq \operatorname{span} A = \{Ax \mid x\in\mathbb{R}^n\} = \{y\in\mathbb{R}^m \mid Ax=y\,,\;\; x\in\mathcal{R}(A^{\mathrm{T}})\} \tag{142}$$

while the nullspaces identify the second common form (*halfspace-description* (113) or *nullspace form*)

$$\mathcal{N}(A) \triangleq \{x\in\mathbb{R}^n \mid Ax=\mathbf{0}\} = \{x\in\mathbb{R}^n \mid x\perp\mathcal{R}(A^{\mathrm{T}})\} \tag{143}$$

$$\mathcal{N}(A^{\mathrm{T}}) \triangleq \{y\in\mathbb{R}^m \mid A^{\mathrm{T}}y=\mathbf{0}\} = \{y\in\mathbb{R}^m \mid y\perp\mathcal{R}(A)\} \tag{144}$$

Range forms (141) (142) are realized as the respective span of the column vectors in matrices A^{T} and A, whereas nullspace form (143) or (144) is the solution set to a linear equation similar to hyperplane definition (114). Yet because matrix A generally has multiple rows, halfspace-description $\mathcal{N}(A)$ is actually the intersection of as many hyperplanes through the origin; for (143), each row of A is normal to a hyperplane while each row of A^{T} is a normal for (144).

2.5.0.0.1 Exercise. *Subspace algebra.*
Given

$$\mathcal{R}(A)+\mathcal{N}(A^{\mathrm{T}}) = \mathcal{R}(B)+\mathcal{N}(B^{\mathrm{T}}) = \mathbb{R}^m \tag{145}$$

prove

$$\mathcal{R}(A)\supseteq\mathcal{N}(B^{\mathrm{T}}) \;\Leftrightarrow\; \mathcal{N}(A^{\mathrm{T}})\subseteq\mathcal{R}(B) \tag{146}$$

$$\mathcal{R}(A)\supseteq\mathcal{R}(B) \;\Leftrightarrow\; \mathcal{N}(A^{\mathrm{T}})\subseteq\mathcal{N}(B^{\mathrm{T}}) \tag{147}$$

e.g, Theorem A.3.1.0.6. ▼

2.5.1 Subspace or affine subset...

Any particular vector subspace $\mathcal{R}_{\mathrm{p}}$ can be described as nullspace $\mathcal{N}(A)$ of some matrix A or as range $\mathcal{R}(B)$ of some matrix B.

More generally, we have the choice of expressing an $n-m$-dimensional affine subset of $\mathbb{R}^n$ as the intersection of m hyperplanes, or as the offset span of $n-m$ vectors:

2.5.1.1 ...as hyperplane intersection

Any affine subset $\mathcal{A}$ of dimension $n-m$ can be described as an intersection of m hyperplanes in $\mathbb{R}^n$; given *fat* $(m\leq n)$ *full-rank* (rank $=\min\{m\,,n\}$) matrix

$$A \triangleq \begin{bmatrix} a_1^{\mathrm{T}} \\ \vdots \\ a_m^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{m\times n} \tag{148}$$

and vector $b\in\mathbb{R}^m$,

$$\mathcal{A} \triangleq \{x\in\mathbb{R}^n \mid Ax=b\} = \bigcap_{i=1}^{m}\{x \mid a_i^{\mathrm{T}}x=b_i\} \tag{149}$$

a halfspace-description. (113)

For example: The intersection of any two independent[2.25] hyperplanes in $\mathbb{R}^3$ is a line, whereas three independent hyperplanes intersect at a point. In $\mathbb{R}^4$, the intersection of two independent hyperplanes is a plane (Example 2.5.1.2.1), whereas three hyperplanes intersect at a line, four at a point, and so on. $\mathcal{A}$ describes a subspace whenever $b=\mathbf{0}$ in (149).

For $n>k$

$$\mathcal{A}\cap\mathbb{R}^k = \{x\in\mathbb{R}^n \mid Ax=b\}\cap\mathbb{R}^k = \bigcap_{i=1}^{m}\left\{x\in\mathbb{R}^k \mid a_i(1\!:\!k)^{\mathrm{T}}x=b_i\right\} \tag{150}$$

The result in §2.4.2.2 is extensible; *id est*, any affine subset $\mathcal{A}$ also has a vertex-description:

2.5.1.2 ... as span of nullspace basis

Alternatively, we may compute a basis for nullspace of matrix A (§E.3.1) and then equivalently express affine subset $\mathcal{A}$ as its span plus an offset: Define

$$Z \triangleq \operatorname{basis}\mathcal{N}(A)\in\mathbb{R}^{n\times n-\operatorname{rank}A} \tag{151}$$

so $AZ=\mathbf{0}$. Then we have a vertex-description in Z,

$$\mathcal{A} = \{x\in\mathbb{R}^n \mid Ax=b\} = \left\{Z\xi+x_{\mathrm{p}} \mid \xi\in\mathbb{R}^{n-\operatorname{rank}A}\right\}\subseteq\mathbb{R}^n \tag{152}$$

the offset span of $n-\operatorname{rank}A$ column vectors, where x_{p} is any particular solution to $Ax=b$; *e.g*, $\mathcal{A}$ describes a subspace whenever $x_{\mathrm{p}}=\mathbf{0}$.

2.5.1.2.1 Example. *Intersecting planes in 4-space.*
Two planes can intersect at a point in four-dimensional Euclidean vector space. It is easy to visualize intersection of two planes in three dimensions; a line can be formed. In four dimensions it is harder to visualize. So let's resort to the tools acquired.

Suppose an intersection of two hyperplanes in four dimensions is specified by a fat full-rank matrix $A_1\in\mathbb{R}^{2\times 4}$ ($m=2$, $n=4$) as in (149):

$$\mathcal{A}_1 \triangleq \left\{x\in\mathbb{R}^4 \;\middle|\; \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\end{bmatrix}x=b_1\right\} \tag{153}$$

The nullspace of A_1 is two dimensional (from Z in (152)), so $\mathcal{A}_1$ represents a plane in four dimensions. Similarly define a second plane in terms of $A_2\in\mathbb{R}^{2\times 4}$:

$$\mathcal{A}_2 \triangleq \left\{x\in\mathbb{R}^4 \;\middle|\; \begin{bmatrix} a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{bmatrix}x=b_2\right\} \tag{154}$$

[2.25] Any number of hyperplanes are called *independent* when defining normals are linearly independent. This misuse departs from independence of two affine subsets that demands intersection only at a point or not at all. (§2.1.4.0.1)

If the two planes are affinely independent and intersect, they intersect at a point because $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ is invertible;

$$\mathcal{A}_1 \cap \mathcal{A}_2 = \left\{ x \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\} \tag{155}$$

□

2.5.1.2.2 Exercise. *Linear program.*
Minimize a hyperplane over affine set $\mathcal{A}$ in the nonnegative orthant

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^{\mathrm{T}} x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \tag{156}$$

where $\mathcal{A} = \{x \mid Ax = b\}$. Two cases of interest are drawn in Figure **33**. Graphically illustrate and explain optimal solutions indicated in the caption. Why is $\alpha^\star$ negative in both cases? Is there solution on the vertical axis? What causes objective unboundedness in the latter case (b)? Describe all vectors c that would yield finite optimal objective in (b).

Graphical solution to linear program

$$\begin{array}{ll} \underset{x}{\text{maximize}} & c^{\mathrm{T}} x \\ \text{subject to} & x \in \mathcal{P} \end{array} \tag{157}$$

is illustrated in Figure **34**. Bounded set $\mathcal{P}$ is an intersection of many halfspaces. Why is optimal solution $x^\star$ not aligned with vector c as in Cauchy-Schwarz inequality (2172)?

▼

2.5.2 Intersection of subspaces

The intersection of nullspaces associated with two matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{k \times n}$ can be expressed most simply as

$$\mathcal{N}(A) \cap \mathcal{N}(B) \;=\; \mathcal{N}\!\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) \;\triangleq\; \{x \in \mathbb{R}^n \mid \begin{bmatrix} A \\ B \end{bmatrix} x = \mathbf{0}\} \tag{158}$$

nullspace of their rowwise concatenation.

Suppose the columns of a matrix Z constitute a basis for $\mathcal{N}(A)$ while the columns of a matrix W constitute a basis for $\mathcal{N}(BZ)$. Then [174, §12.4.2]

$$\mathcal{N}(A) \cap \mathcal{N}(B) = \mathcal{R}(ZW) \tag{159}$$

If each basis is orthonormal, then the columns of ZW constitute an orthonormal basis for the intersection.

In the particular circumstance A and B are each positive semidefinite [22, §6], or in the circumstance A and B are two linearly independent dyads (§B.1.1), then

$$\mathcal{N}(A) \cap \mathcal{N}(B) = \mathcal{N}(A + B)\,, \quad \left\{ \begin{array}{l} A\,, B \in \mathbb{S}_+^M \\ \quad \textbf{or} \\ A + B = u_1 v_1^{\mathrm{T}} + u_2 v_2^{\mathrm{T}} \quad \text{(l.i.)} \end{array} \right. \tag{160}$$

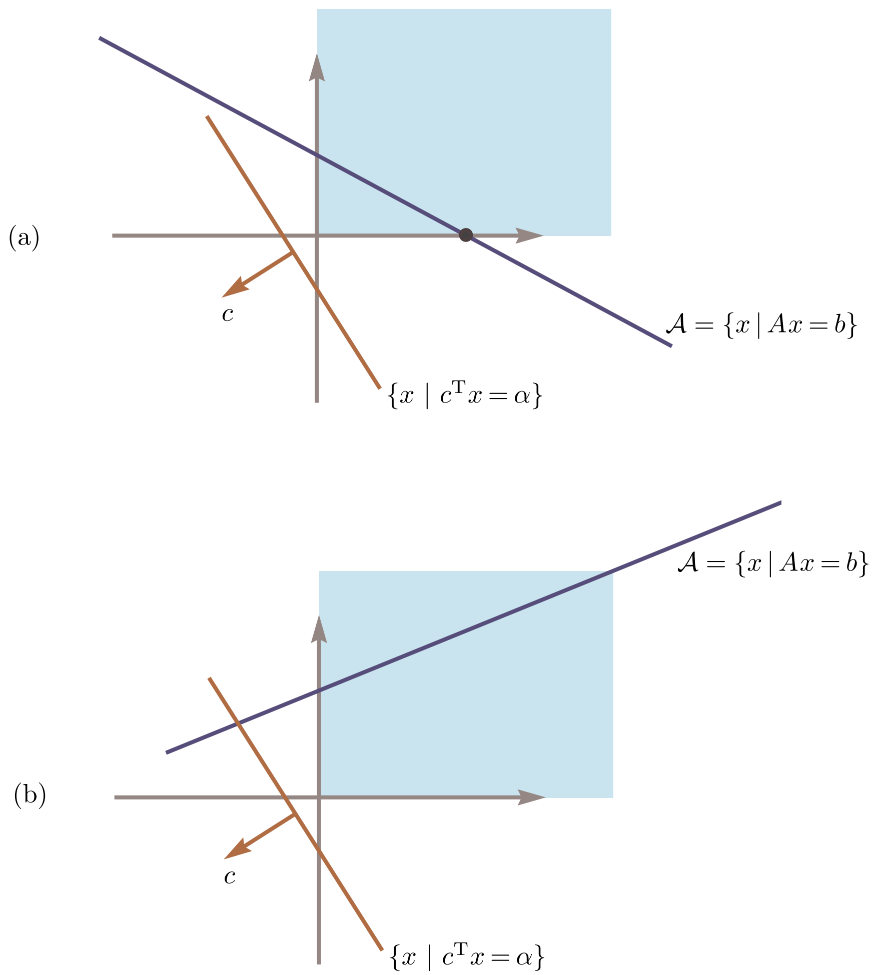


Figure 33: Minimizing hyperplane over affine set $\mathcal{A}$ in nonnegative orthant $\mathbb{R}^{\mathbf{2}}_+$ whose extreme directions (§2.8.1) are the nonnegative Cartesian axes. Solutions are visually ascertainable: **(a)** Optimal solution is $\bullet$. **(b)** Optimal objective $\alpha^\star = -\infty$.

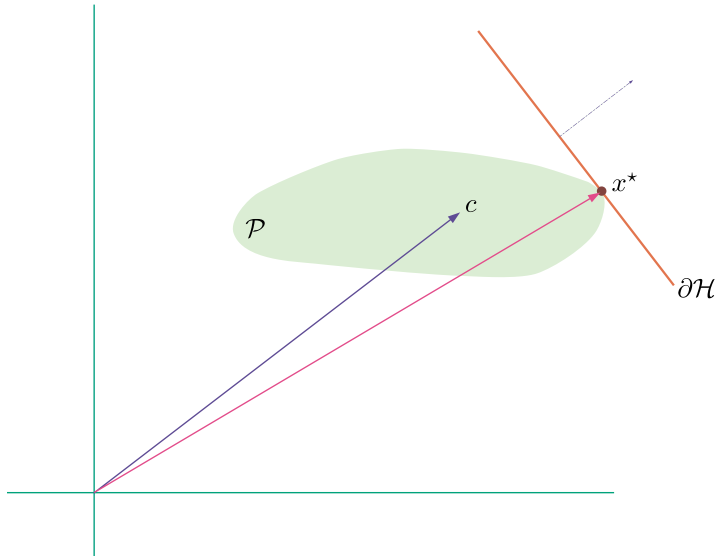


Figure 34: Maximizing hyperplane $\partial\mathcal{H}$, whose normal is vector $c\in\mathcal{P}$, over polyhedral set $\mathcal{P}$ in $\mathbb{R}^2$ is a linear program (157). Optimal solution $x^\star$ at $\bullet$.

2.5.3 Visualization of matrix subspaces

Fundamental subspace relations, such as

$$\mathcal{R}(A^{\mathrm{T}}) \perp \mathcal{N}(A)\,, \qquad \mathcal{N}(A^{\mathrm{T}}) \perp \mathcal{R}(A) \tag{137}$$

are partially defining. But to aid visualization of involved geometry, it sometimes helps to vectorize matrices. For any square matrix A, $s\in\mathcal{N}(A)$, and $w\in\mathcal{N}(A^{\mathrm{T}})$

$$\langle A\,,\, ss^{\mathrm{T}}\rangle = 0\,, \qquad \langle A\,,\, ww^{\mathrm{T}}\rangle = 0 \tag{161}$$

because $s^{\mathrm{T}}A\,s = w^{\mathrm{T}}A\,w = 0$. This innocuous observation becomes a sharp instrument for visualization of diagonalizable matrices (§A.5.1): for rank-ρ matrix $A\in\mathbb{R}^{M\times M}$

$$A \;=\; S\Lambda S^{-1} = [\,s_1 \cdots s_M\,]\;\Lambda \begin{bmatrix} w_1^{\mathrm{T}} \\ \vdots \\ w_M^{\mathrm{T}} \end{bmatrix} = \sum_{i=1}^{M} \lambda_i\, s_i w_i^{\mathrm{T}} \tag{1636}$$

where nullspace eigenvectors are real by Theorem A.5.0.0.1 and where (§B.1.1)

$$\begin{aligned} \mathcal{R}\{s_i\in\mathbb{R}^M \mid \lambda_i\!=\!0\} \;&=\; \mathcal{R}\left(\sum_{i=\rho+1}^{M} s_i s_i^{\mathrm{T}}\right) \;=\; \mathcal{N}(A) \\ \mathcal{R}\{w_i\in\mathbb{R}^M \mid \lambda_i\!=\!0\} \;&=\; \mathcal{R}\left(\sum_{i=\rho+1}^{M} w_i w_i^{\mathrm{T}}\right) \;=\; \mathcal{N}(A^{\mathrm{T}}) \end{aligned} \tag{162}$$

Define an unconventional basis among column vectors of each summation:

$$\begin{aligned}\operatorname{basis}\mathcal{N}(A) &\subseteq \sum_{i=\rho+1}^{M} s_i s_i^{\mathrm T} \subseteq \mathcal{N}(A)\\ \operatorname{basis}\mathcal{N}(A^{\mathrm T}) &\subseteq \sum_{i=\rho+1}^{M} w_i w_i^{\mathrm T} \subseteq \mathcal{N}(A^{\mathrm T})\end{aligned} \tag{163}$$

We shall regard a *vectorized subspace* as vectorization of any $M\times M$ matrix whose columns comprise an *overcomplete* basis for that subspace; *e.g*, §E.3.1

$$\begin{aligned}\operatorname{vec}\operatorname{basis}\mathcal{N}(A) &= \operatorname{vec}\sum_{i=\rho+1}^{M} s_i s_i^{\mathrm T}\\ \operatorname{vec}\operatorname{basis}\mathcal{N}(A^{\mathrm T}) &= \operatorname{vec}\sum_{i=\rho+1}^{M} w_i w_i^{\mathrm T}\end{aligned} \tag{164}$$

By this reckoning, $\operatorname{vec}\operatorname{basis}\mathcal{R}(A) = \operatorname{vec} A$ but is not unique. Now, because

$$\left\langle A\,,\ \sum_{i=\rho+1}^{M} s_i s_i^{\mathrm T}\right\rangle = 0\,, \qquad \left\langle A\,,\ \sum_{i=\rho+1}^{M} w_i w_i^{\mathrm T}\right\rangle = 0 \tag{165}$$

then vectorized matrix A is normal to a hyperplane (of dimension M^2-1) that contains both vectorized nullspaces (each of whose dimension is $M-\rho$);

$$\operatorname{vec} A \perp \operatorname{vec}\operatorname{basis}\mathcal{N}(A)\,, \qquad \operatorname{vec}\operatorname{basis}\mathcal{N}(A^{\mathrm T}) \perp \operatorname{vec} A \tag{166}$$

These vectorized subspace orthogonality relations represent a departure (absent $^{\mathrm T}$) from fundamental subspace relations (137) stated at the outset.

2.6 Extreme, Exposed

2.6.0.0.1 Definition. *Extreme point.*
An extreme point x_ε of a convex set $\mathcal{C}$ is a point, belonging to its closure $\overline{\mathcal{C}}$ [43, §3.3], that is not expressible as a convex combination of points in $\overline{\mathcal{C}}$ distinct from x_ε ; *id est*, for $x_\varepsilon \in \overline{\mathcal{C}}$ and all $x_1, x_2 \in \overline{\mathcal{C}} \setminus x_\varepsilon$

$$\mu\, x_1 + (1-\mu)x_2 \neq x_\varepsilon\,, \quad \mu\in[0,1] \tag{167}$$

△

In other words, x_ε is an extreme point of $\mathcal{C}$ if and only if x_ε is not a point relatively interior to any line segment in $\overline{\mathcal{C}}$. [377, §2.10]

Borwein & Lewis offer: [56, §4.1.6] An extreme point of a convex set $\mathcal{C}$ is a point x_ε in $\overline{\mathcal{C}}$ whose *relative complement* $\overline{\mathcal{C}}\setminus x_\varepsilon$ is convex.

The set consisting of a single point $\mathcal{C}=\{x_\varepsilon\}$ is itself an extreme point.

2.6.0.0.2 Theorem. *Extreme existence.* [325, §18.5.3] [27, §II.3.5]
A nonempty closed convex set containing no lines has at least one extreme point. ⋄

2.6.0.0.3 Definition. *Face, edge.* [215, §A.2.3]

- A *face* $\mathcal{F}$ of convex set $\mathcal{C}$ is a convex subset $\mathcal{F}\subseteq\overline{\mathcal{C}}$ such that every closed line segment $\overline{x_1 x_2}$ in $\overline{\mathcal{C}}$, having a relatively interior point $(x\in\operatorname{rel\,int}\overline{x_1 x_2})$ in $\mathcal{F}$, has both endpoints in $\mathcal{F}$. The zero-dimensional faces of $\mathcal{C}$ constitute its extreme points. The empty set $\emptyset$ and $\overline{\mathcal{C}}$ itself are conventional faces of $\mathcal{C}$. [325, §18]

- All faces $\mathcal{F}$ are extreme sets by definition; *id est*, for $\mathcal{F}\subseteq\overline{\mathcal{C}}$ and all $x_1, x_2\in\overline{\mathcal{C}}\backslash\mathcal{F}$

$$\mu\, x_1 + (1-\mu)x_2 \notin \mathcal{F}\;, \quad \mu\in[0,1] \tag{168}$$

- A one-dimensional face of a convex set is called an *edge*. △

Dimension of a face is the penultimate number of affinely independent points (§2.4.2.3) belonging to it;

$$\dim\mathcal{F} = \sup_\rho \dim\{x_2 - x_1\;,\; x_3 - x_1\;,\;\ldots\;,\; x_\rho - x_1 \mid x_i\in\mathcal{F}\,,\; i=1\ldots\rho\} \tag{169}$$

The point of intersection in $\overline{\mathcal{C}}$ with a strictly supporting hyperplane identifies an extreme point, but not *vice versa.* The nonempty intersection of any supporting hyperplane with $\overline{\mathcal{C}}$ identifies a face, in general, but not *vice versa.* To acquire a converse, the concept *exposed face* requires introduction:

2.6.1 Exposure

2.6.1.0.1 Definition. *Exposed face, exposed point, vertex, facet.* [215, §A.2.3, §A.2.4]

- $\mathcal{F}$ is an *exposed face* of an n-dimensional convex set $\mathcal{C}$ iff there is a supporting hyperplane $\underline{\partial\mathcal{H}}$ to $\overline{\mathcal{C}}$ such that

$$\mathcal{F} = \overline{\mathcal{C}} \cap \underline{\partial\mathcal{H}} \tag{170}$$

  Only faces of dimension -1 through $n-1$ can be exposed by a hyperplane.

- An *exposed point,* the definition of *vertex,* is equivalent to a zero-dimensional exposed face; the point of intersection with a strictly supporting hyperplane.

- A *facet* is an $(n-1)$-dimensional exposed face of an n-dimensional convex set $\mathcal{C}$; facets exist in one-to-one correspondence with the $(n-1)$-dimensional faces.[2.26]

- $\overline{\{\text{exposed points}\}} \;=\; \{\text{extreme points}\}$
  $\{\text{exposed faces}\} \;\subseteq\; \{\text{faces}\}$ △

[2.26] This coincidence occurs simply because the facet's dimension is the same as the dimension of the supporting hyperplane exposing it.

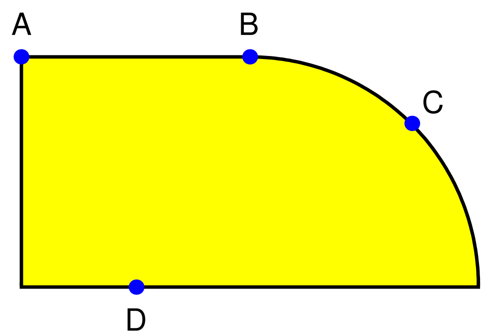


Figure 35: Closed convex set in $\mathbb{R}^{\mathbf{2}}$. Point A is exposed hence extreme; a classical vertex. Point B is extreme but not an exposed point. Point C is exposed and extreme; zero-dimensional exposure makes it a vertex. Point D is neither an exposed or extreme point although it belongs to a one-dimensional exposed face. [215, §A.2.4] [347, §3.6] Closed face $\overline{\mathrm{AB}}$ is exposed; a facet. The arc is not a conventional face, yet it is composed entirely of extreme points. Union of all rotations of this entire set about its vertical edge produces another convex set in three dimensions having no edges; but that convex set produced by rotation about horizontal edge containing D has edges.

2.6.1.1 Density of exposed points

For any closed convex set $\mathcal{C}$, its exposed points constitute a *dense* subset of its extreme points; [325, §18] [352] [347, §3.6, p.115] dense in the sense [412] that closure of that subset yields the set of extreme points.

For the convex set illustrated in Figure **35**, point B cannot be exposed because it relatively bounds both the facet $\overline{\mathrm{AB}}$ and the closed quarter circle, each bounding the set. Since B is not relatively interior to any line segment in the set, then B is an extreme point by definition. Point B may be regarded as the limit of some sequence of exposed points beginning at vertex C.

2.6.1.2 Face transitivity and algebra

Faces of a convex set enjoy transitive relation. If $\mathcal{F}_1$ is a face (an extreme set) of $\mathcal{F}_2$ which in turn is a face of $\mathcal{F}_3$, then it is always true that $\mathcal{F}_1$ is a face of $\mathcal{F}_3$. (The parallel statement for exposed faces is false. [325, §18]) For example, any extreme point of $\mathcal{F}_2$ is an extreme point of $\mathcal{F}_3$; in this example, $\mathcal{F}_2$ could be a face exposed by a hyperplane supporting polyhedron $\mathcal{F}_3$. [239, def.115/6 p.358] Yet it is erroneous to presume that a face, of dimension 1 or more, consists entirely of extreme points. Nor is a face of dimension 2 or more entirely composed of edges, and so on.

For the polyhedron in $\mathbb{R}^3$ from Figure **22**, for example, the nonempty faces exposed by a hyperplane are the vertices, edges, and facets; there are no more. The zero-, one-, and two-dimensional faces are in one-to-one correspondence with the exposed faces in that example.

2.6.1.3 Smallest face

Define the smallest face $\mathcal{F}$, that contains some element G, of a convex set $\mathcal{C}$:

$$\mathcal{F}(\mathcal{C}\ni G) \tag{171}$$

videlicet, $\overline{\mathcal{C}} \supset \operatorname{rel\,int}\mathcal{F}(\mathcal{C}\ni G)\ni G$. An affine set has no faces except itself and the empty set. The smallest face, that contains G, of intersection of convex set $\mathcal{C}$ with an affine set $\mathcal{A}$ [255, §2.4] [256]

$$\mathcal{F}\big((\mathcal{C}\cap\mathcal{A})\ni G\big) = \mathcal{F}(\mathcal{C}\ni G)\cap\mathcal{A} \tag{172}$$

equals intersection of $\mathcal{A}$ with the smallest face, that contains G, of set $\mathcal{C}$.

2.6.1.4 Conventional boundary

(*confer* §2.1.7.2) Relative boundary

$$\operatorname{rel}\partial\mathcal{C} = \overline{\mathcal{C}}\setminus\operatorname{rel\,int}\mathcal{C} \tag{24}$$

is equivalent to:

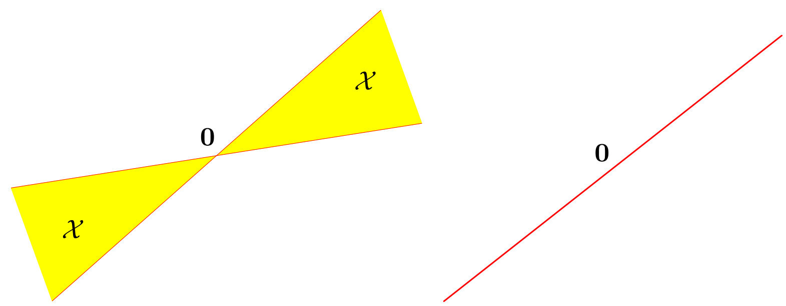
(a)  (b)

Figure 36: **(a)** Two-dimensional nonconvex cone drawn truncated. Boundary of this cone is itself a cone. Each half is itself a convex cone. **(b)** This convex cone (drawn truncated) is a line through the origin in any dimension. It has no relative boundary, while its relative interior comprises entire line.

2.6.1.4.1 Definition. *Conventional boundary of convex set.* [215, §C.3.1]
The relative boundary $\partial\mathcal{C}$ of a nonempty convex set $\mathcal{C}$ is the union of all exposed faces of $\overline{\mathcal{C}}$. $\triangle$

Equivalence to (24) comes about because it is conventionally presumed that any supporting hyperplane, central to the definition of exposure, does not contain $\mathcal{C}$. [325, p.100] Any face $\mathcal{F}$ of convex set $\mathcal{C}$ (that is not $\mathcal{C}$ itself) belongs to $\operatorname{rel}\partial\mathcal{C}$. (§2.8.2.1)

2.7 Cones

In optimization, convex cones achieve prominence because they generalize subspaces. Most compelling is the projection analogy: Projection on a subspace can be ascertained from projection on its orthogonal complement (Figure **180**), whereas projection on a closed convex cone can be determined from projection instead on its *algebraic complement* **(**§2.13, Figure **181**, §E.9.2**)**; called the *polar cone.*

2.7.0.0.1 Definition. *Ray.*
The one-dimensional set

$$\{\zeta\Gamma + B \mid \zeta \geq 0\,,\ \Gamma \neq \mathbf{0}\} \subset \mathbb{R}^n \tag{173}$$

defines a *halfline* called a *ray* in nonzero *direction* $\Gamma\in\mathbb{R}^n$ having *base* $B\in\mathbb{R}^n$. When $B=\mathbf{0}$, a ray is the conic hull of direction Γ; hence a convex cone. $\triangle$

Relative boundary of a single ray, base $\mathbf{0}$ in any dimension, is the origin because that is the union of all exposed faces not containing the entire set. Its relative interior is the ray itself excluding the origin.

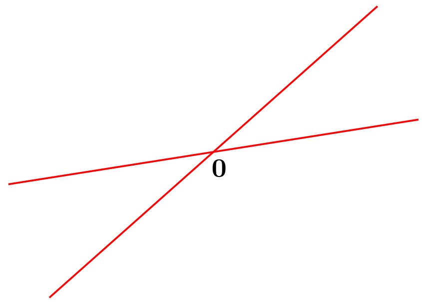


Figure 37: This nonconvex cone in $\mathbb{R}^{\mathbf{2}}$ is a pair of lines through the origin. [266, §2.4] Because the lines are linearly independent, they are algebraic complements whose vector sum is $\mathbb{R}^{\mathbf{2}}$ a convex cone.

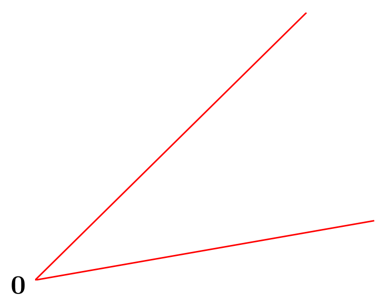


Figure 38: Boundary of a convex cone in $\mathbb{R}^{\mathbf{2}}$ is a nonconvex cone; a pair of rays emanating from the origin.

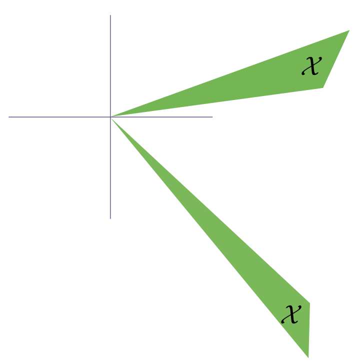


Figure 39: Union of two pointed closed convex cones in $\mathbb{R}^{\mathbf{2}}$ is nonconvex cone $\mathcal{X}$.

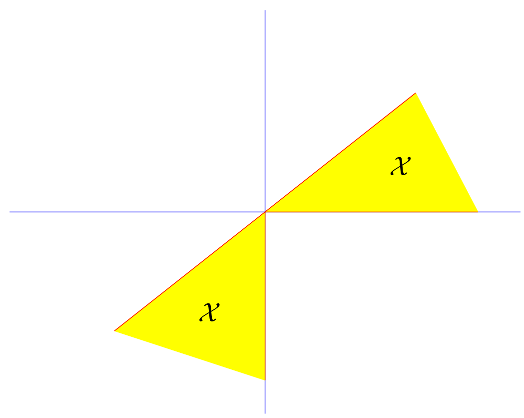


Figure 40: Truncated nonconvex cone $\mathcal{X}=\{x\in\mathbb{R}^{\mathbf{2}} \mid x_1\geq x_2\ ,\ x_1 x_2\geq 0\}$. Boundary is also a cone. [266, §2.4] (Cartesian axes drawn for reference.) Each half (about the origin) is itself a convex cone.

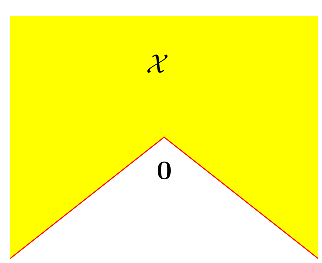


Figure 41: Nonconvex cone $\mathcal{X}$ drawn truncated in $\mathbb{R}^{\mathbf{2}}$. Boundary is also a cone. [266, §2.4] Cone exterior is convex cone.

2.7.1 Cone defined

A set $\mathcal{X}$ is called, simply, *cone* if and only if

$$\Gamma \in \mathcal{X} \;\Rightarrow\; \zeta\Gamma \in \overline{\mathcal{X}} \;\text{ for all }\; \zeta \geq 0 \tag{174}$$

where $\overline{\mathcal{X}}$ denotes closure of cone $\mathcal{X}$; *e.g*, Figure **38**, Figure **39**. An example of nonconvex cone is the union of two opposing quadrants: $\mathcal{X} = \{x \in \mathbb{R}^{\mathbf{2}} \mid x_1 x_2 \geq 0\}$. [410, §2.5] Similar examples are Figure **36** and Figure **40**.

All cones obey (174) and can be defined by an aggregate of rays emanating exclusively from the origin. Hence all closed cones contain the origin $\mathbf{0}$ and are unbounded, excepting the simplest cone $\{\mathbf{0}\}$. The empty set $\emptyset$ is not a cone, but its conic hull is;

$$\operatorname{cone}\emptyset \;=\; \{\mathbf{0}\} \qquad\qquad (104)$$

2.7.2 Convex cone

We call set $\mathcal{K}$ a *convex cone* iff

$$\Gamma_1\,, \Gamma_2 \in \mathcal{K} \;\Rightarrow\; \zeta\Gamma_1 + \xi\Gamma_2 \in \overline{\mathcal{K}} \;\text{ for all }\; \zeta\,, \xi \geq 0 \tag{175}$$

id est, if and only if any conic combination of elements from $\mathcal{K}$ belongs to its closure. Apparent from this definition, $\zeta\Gamma_1 \in \overline{\mathcal{K}}$ and $\xi\Gamma_2 \in \overline{\mathcal{K}}$ $\forall\, \zeta\,, \xi \geq 0$; meaning, $\mathcal{K}$ is a cone. Set $\mathcal{K}$ is convex since, for any particular $\zeta\,, \xi \geq 0$

$$\mu\,\zeta\Gamma_1 \,+\, (1-\mu)\,\xi\Gamma_2 \in \overline{\mathcal{K}} \quad \forall\, \mu \in [0,1] \tag{176}$$

because $\mu\,\zeta\,, (1-\mu)\,\xi \geq 0$. Obviously,

$$\{\mathcal{X}\} \supset \{\mathcal{K}\} \tag{177}$$

the set of all convex cones is a proper subset of all cones. The set of convex cones is a narrower but more familiar class of cone, any member of which can be equivalently described as the intersection of a possibly (but not necessarily) infinite number of hyperplanes (through the origin) and halfspaces whose bounding hyperplanes pass through the origin; a halfspace-description (§2.4). Convex cones need not be full-dimensional.

More familiar convex cones are *Lorentz cone* (*confer* Figure **49**)[2.27]

$$\mathcal{K}_\ell = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R} \mid \|x\|_\ell \leq t \right\} \;, \qquad \ell = 2 \tag{178}$$

and polyhedral cone (§2.12.1.0.1); *e.g*, any orthant generated by Cartesian half-axes (§2.1.3). Esoteric examples of convex cones include the point at the origin, any line through the origin, any ray having the origin as base such as the nonnegative real line $\mathbb{R}_+$ in subspace $\mathbb{R}$, any halfspace partially bounded by a hyperplane through the origin, the positive semidefinite cone $\mathbb{S}^M_+$ (191), the cone of Euclidean distance matrices $\mathbb{EDM}^N$ (978) (§6), *completely positive* semidefinite matrices $\{CC^{\mathrm{T}} \mid C \geq \mathbf{0}\}$ [41, p.71], any subspace, and Euclidean vector space $\mathbb{R}^n$.

[2.27] `a.k.a:` *second-order cone, quadratic cone, circular cone* (§2.9.2.8.1), unbounded *ice-cream cone* united with its interior.

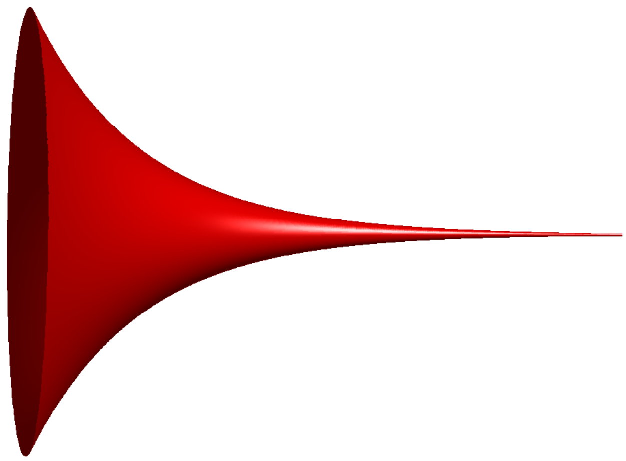


Figure 42: Not a cone; ironically, the three-dimensional *flared horn* (with or without its interior) resembling mathematical symbol $\succ$ denoting strict cone membership and partial order.

2.7.2.1 cone invariance

More Euclidean bodies are cones, it seems, than are not.[2.28] The convex cone class of Euclidean body is invariant to scaling, linear and single- or many-valued inverse linear transformation, vector summation, and Cartesian product, but is not invariant to translation. [325, p.22]

2.7.2.1.1 Theorem. *Cone intersection (nonempty).*

- Intersection of an arbitrary collection of convex cones is a convex cone. [325, §2, §19]
- Intersection of an arbitrary collection of closed convex cones is a closed convex cone. [274, §2.3]
- Intersection of a finite number of polyhedral cones (Figure **53** p.124, §2.12.1.0.1) remains a polyhedral cone. ⋄

The property *pointedness* is ordinarily associated with a convex cone but, strictly speaking,

- pointed cone $\not\Leftrightarrow$ convex cone (Figure **38**, Figure **39**)

[2.28] *confer* Figures: **27 36 37 38 39 40 41 42 44 46 53 58 61 63 64 66 67 68 69 70 152 165 191**

2.7.2.1.2 Definition. *Pointed convex cone.* (*confer* §2.12.2.2)
A convex cone $\mathcal{K}$ is *pointed* iff it contains no line. Equivalently, $\mathcal{K}$ is not pointed iff there exists any nonzero direction $\Gamma \in \overline{\mathcal{K}}$ such that $-\Gamma \in \overline{\mathcal{K}}$. If the origin is an extreme point of $\overline{\mathcal{K}}$ or, equivalently, if

$$\overline{\mathcal{K}} \cap -\overline{\mathcal{K}} = \{\mathbf{0}\} \tag{179}$$

then $\mathcal{K}$ is pointed, and *vice versa.* [347, §2.10] A convex cone is pointed iff the origin is the smallest nonempty face of its closure. △

Then a pointed closed convex cone, by principle of separating hyperplane (§2.4.2.7), has a strictly supporting hyperplane at the origin. The simplest and only bounded [410, p.75] convex cone $\mathcal{K} = \{\mathbf{0}\} \subseteq \mathbb{R}^n$ is pointed, by convention, but not full-dimensional. Its relative boundary is the empty set $\emptyset$ (25) while its relative interior is the point $\mathbf{0}$ itself (12). The pointed convex cone that is a halfline, emanating from the origin in $\mathbb{R}^n$, has relative boundary $\mathbf{0}$ while its relative interior is the halfline itself excluding $\mathbf{0}$. Pointed are any Lorentz cone, cone of Euclidean distance matrices $\mathbb{EDM}^N$ in symmetric hollow subspace $\mathbb{S}^N_h$, and positive semidefinite cone $\mathbb{S}^M_+$ in ambient $\mathbb{S}^M$.

2.7.2.1.3 Theorem. *Pointed cones.* [56, §3.3.15, exer.20]
A closed convex cone $\mathcal{K} \subset \mathbb{R}^n$ is pointed if and only if there exists a normal α such that the set

$$\mathcal{C} \triangleq \{x \in \mathcal{K} \mid \langle x\,, \alpha\rangle = 1\} \tag{180}$$

is closed, bounded, and $\mathcal{K} = \operatorname{cone} \mathcal{C}$. Equivalently, $\mathcal{K}$ is pointed if and only if there exists a vector β normal to a hyperplane strictly supporting $\mathcal{K}$; *id est*, for some positive scalar ϵ

$$\langle x\,, \beta\rangle \geq \epsilon\|x\| \quad \forall\, x \in \mathcal{K} \tag{181}$$

⋄

If closed convex cone $\mathcal{K}$ is not pointed, then it has no extreme point.[2.29] Yet a pointed closed convex cone has only one extreme point [43, §3.3]: the exposed point residing at the origin; its vertex. Pointedness is invariant to Cartesian product by (179). And from the *cone intersection theorem* it follows that an intersection of convex cones is pointed if at least one of the cones is; implying, each and every nonempty exposed face of a pointed closed convex cone is a pointed closed convex cone.

2.7.2.2 Pointed closed convex cone induces partial order

Relation $\preceq$ represents *partial order* on some set if that relation possesses[2.30]

reflexivity $(x \preceq x)$

antisymmetry $(x \preceq z\,,\ z \preceq x \Rightarrow x = z)$

transitivity $(x \preceq y\,,\ y \preceq z \Rightarrow x \preceq z)$, $\qquad (x \preceq y\,,\ y \prec z \Rightarrow x \prec z)$

[2.29] nor does it have extreme directions (§2.8.1).

[2.30] A set is *totally ordered* if it further obeys a comparability property of the relation: for each and every x and y from the set, $x \preceq y$ or $y \preceq x$; *e.g*, one-dimensional real vector space $\mathbb{R}$ is the smallest unbounded totally ordered and connected set.

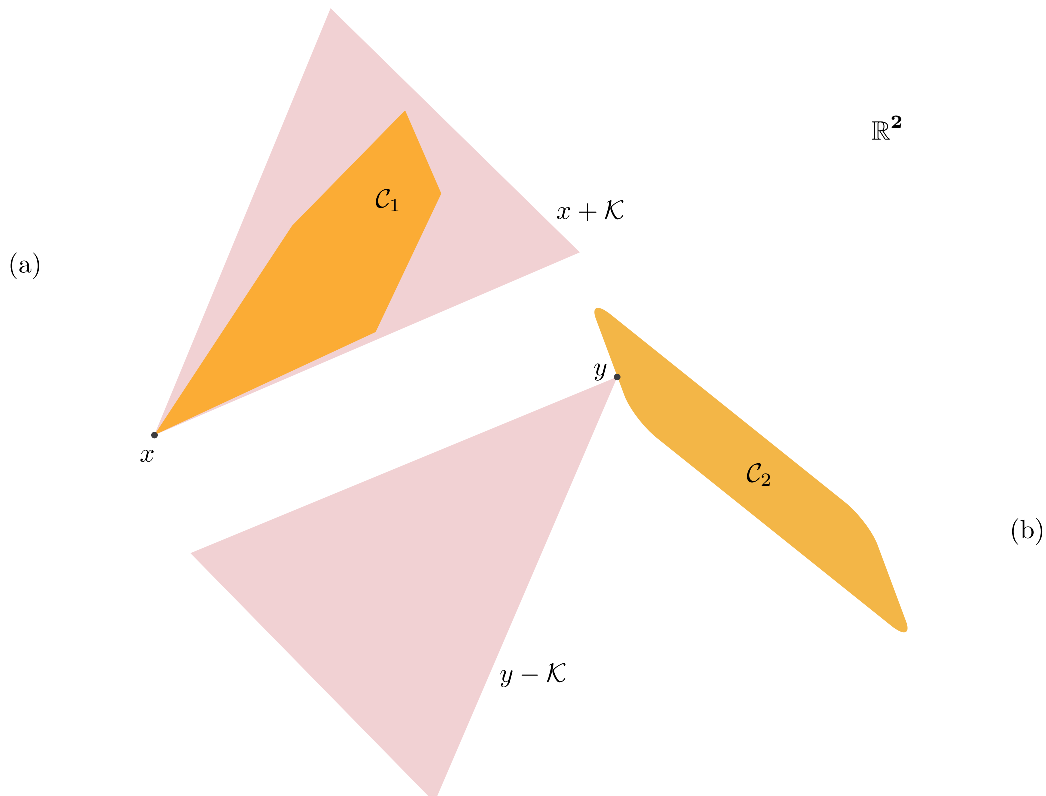


Figure 43: (*confer* Figure **73**) **(a)** Point x is the minimum element of set $\mathcal{C}_1$ with respect to cone $\mathcal{K}$ because cone translated to $x \in \mathcal{C}_1$ contains entire set. (Cones drawn truncated.) **(b)** Point y is a minimal element of set $\mathcal{C}_2$ with respect to cone $\mathcal{K}$ because negative cone translated to $y \in \mathcal{C}_2$ contains only y. These concepts, minimum/minimal, become equivalent under a *total order*.

A pointed closed convex cone $\mathcal{K}$ induces partial order on $\mathbb{R}^n$ or $\mathbb{R}^{m\times n}$, [22, §1] [341, p.7] essentially defined by vector or matrix inequality;

$$x \underset{\mathcal{K}}{\preceq} z \quad\Leftrightarrow\quad z-x\in\mathcal{K} \tag{182}$$

$$x \underset{\mathcal{K}}{\prec} z \quad\Leftrightarrow\quad z-x\in\operatorname{rel\,int}\mathcal{K} \tag{183}$$

Neither x or z is necessarily a member of $\mathcal{K}$ for these relations to hold. Only when $\mathcal{K}$ is a nonnegative orthant $\mathbb{R}^n_+$ do these inequalities reduce to ordinary entrywise comparison (§2.13.4.2.3) while partial order lingers. Inclusive of that special case, we ascribe nomenclature *generalized inequality* to comparison with respect to a pointed closed convex cone.

We say two points x and y are *comparable* when $x\preceq y$ or $y\preceq x$ with respect to pointed closed convex cone $\mathcal{K}$. Visceral mechanics of actually comparing points, when cone $\mathcal{K}$ is not an orthant, are well illustrated in the example of Figure **67** which relies on the equivalent membership-interpretation in definition (182) or (183).

Comparable points and the *minimum element* of some vector- or matrix-valued partially ordered set are thus well defined, so nonincreasing sequences with respect to cone $\mathcal{K}$ can therefore converge in this sense: Point $x\in\mathcal{C}$ is the (unique) minimum element of set $\mathcal{C}$ with respect to cone $\mathcal{K}$ iff for each and every $z\in\mathcal{C}$ we have $x\preceq z$; equivalently, iff $\mathcal{C}\subseteq x+\mathcal{K}$.[2.31]

A closely related concept, *minimal element,* is useful for partially ordered sets having no minimum element: Point $x\in\mathcal{C}$ is a minimal element of set $\mathcal{C}$ with respect to pointed closed convex cone $\mathcal{K}$ if and only if $(x-\mathcal{K})\cap\mathcal{C}=x$. (Figure **43**) No uniqueness is implied here, although implicit is the assumption: $\dim\mathcal{K}\geq\dim\operatorname{aff}\mathcal{C}$. In words, a point that is a minimal element is smaller (with respect to $\mathcal{K}$) than any other point in the set to which it is comparable.

Further properties of partial order with respect to pointed closed convex cone $\mathcal{K}$ are not defining:

$$\text{homogeneity}\quad (x\preceq y\,,\;\lambda\geq 0\Rightarrow\lambda x\preceq\lambda z),\qquad (x\prec y\,,\;\lambda>0\Rightarrow\lambda x\prec\lambda z)$$

$$\text{additivity}\quad (x\preceq z\,,\;u\preceq v\Rightarrow x+u\preceq z+v),\qquad (x\prec z\,,\;u\preceq v\Rightarrow x+u\prec z+v)$$

2.7.2.2.1 Definition. *Proper cone*: a cone that is

- pointed
- closed
- convex
- full-dimensional. △

[2.31] Borwein & Lewis [56, §3.3 exer.21] ignore possibility of equality to $x+\mathcal{K}$ in this condition, and require a second condition: …and $\mathcal{C}\subset y+\mathcal{K}$ *for some* y *in* $\mathbb{R}^n$ *implies* $x\in y+\mathcal{K}$.

A proper cone remains proper under injective linear transformation. [94, §5.1] Examples of proper cones are the positive semidefinite cone $\mathbb{S}^M_+$ in the ambient space of symmetric matrices (§2.9), the nonnegative real line $\mathbb{R}_+$ in vector space $\mathbb{R}$, or any orthant in $\mathbb{R}^n$, and the set of all coefficients of univariate degree-n polynomials nonnegative on interval $[0,1]$ [63, exmp.2.16] or univariate degree-$2n$ polynomials nonnegative over $\mathbb{R}$ [63, exer.2.37].

2.8 Cone boundary

Every hyperplane supporting a convex cone contains the origin. [215, §A.4.2] Because any supporting hyperplane to a convex cone must therefore itself be a cone, then from the *cone intersection theorem* (§2.7.2.1.1) it follows:

2.8.0.0.1 Lemma. *Cone faces.* [27, §II.8]
Each nonempty exposed face of a convex cone is a convex cone. ⋄

2.8.0.0.2 Theorem. *Proper-cone boundary.*
Suppose a nonzero point Γ lies on the boundary $\partial\mathcal{K}$ of proper cone $\mathcal{K}$ in $\mathbb{R}^n$. Then it follows that the ray $\{\zeta\Gamma \mid \zeta \geq 0\}$ also belongs to $\partial\mathcal{K}$. ⋄

Proof. By virtue of its propriety, a proper cone guarantees existence of a strictly supporting hyperplane at the origin. [325, cor.11.7.3][2.32] Hence the origin belongs to the boundary of $\mathcal{K}$ because it is the zero-dimensional exposed face. The origin belongs to the ray through Γ, and the ray belongs to $\mathcal{K}$ by definition (174). By the *cone faces lemma,* each and every nonempty exposed face must include the origin. Hence the closed line segment $\overline{\mathbf{0}\Gamma}$ must lie in an exposed face of $\mathcal{K}$ because both endpoints do by Definition 2.6.1.4.1. That means there exists a supporting hyperplane $\partial\underline{\mathcal{H}}$ to $\mathcal{K}$ containing $\overline{\mathbf{0}\Gamma}$. So the ray through Γ belongs both to $\mathcal{K}$ and to $\partial\underline{\mathcal{H}}$. $\partial\underline{\mathcal{H}}$ must therefore expose a face of $\mathcal{K}$ that contains the ray; *id est,*

$$\{\zeta\Gamma \mid \zeta \geq 0\} \subseteq \mathcal{K} \cap \partial\underline{\mathcal{H}} \subset \partial\mathcal{K} \tag{184}$$

♦

Proper cone $\{\mathbf{0}\}$ in $\mathbb{R}^0$ has no boundary (24) because (12)

$$\operatorname{rel\,int}\{\mathbf{0}\} = \{\mathbf{0}\} \tag{185}$$

The boundary of any proper cone in $\mathbb{R}$ is the origin.

The boundary of any convex cone whose dimension exceeds 1 can be constructed entirely from an aggregate of rays emanating exclusively from the origin.

[2.32] Rockafellar's corollary yields a supporting hyperplane at the origin to any convex cone in $\mathbb{R}^n$ not equal to $\mathbb{R}^n$.

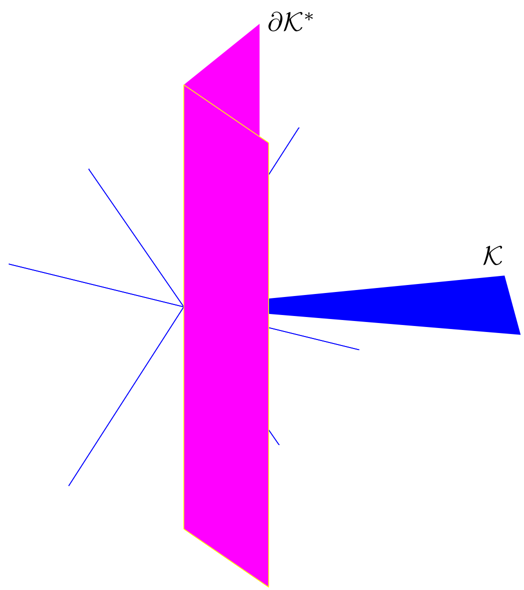


Figure 44: $\mathcal{K}$ is a pointed polyhedral cone not full-dimensional in $\mathbb{R}^3$ (drawn truncated in a plane parallel to the floor upon which you stand). Dual cone $\mathcal{K}^*$ is a *wedge* whose truncated boundary is illustrated (drawn perpendicular to the floor). In this particular instance, $\mathcal{K} \subset \operatorname{int} \mathcal{K}^*$ (excepting the origin). (Cartesian coordinate axes drawn for reference.)

2.8.1 Extreme direction

The property *extreme direction* arises naturally in connection with the pointed closed convex cone $\mathcal{K}\subset\mathbb{R}^n$, being analogous to extreme point. [325, §18, p.162][2.33] An extreme direction Γ_ε of pointed $\mathcal{K}$ is a vector corresponding to an edge that is a ray $\{\zeta\,\Gamma_\varepsilon\in\mathcal{K} \mid \zeta\geq 0\}$ emanating from the origin.[2.34] Nonzero direction Γ_ε in pointed $\mathcal{K}$ is extreme if and only if

$$\zeta_1\Gamma_1+\zeta_2\Gamma_2 \neq \Gamma_\varepsilon \quad \forall\,\zeta_1\,,\zeta_2\geq 0\,,\quad \forall\,\Gamma_1\,,\Gamma_2\in\mathcal{K}\backslash\{\zeta\,\Gamma_\varepsilon\in\mathcal{K} \mid \zeta\geq 0\} \tag{186}$$

In words, an extreme direction in a pointed closed convex cone is the direction of a ray, called an *extreme ray,* that cannot be expressed as a conic combination of directions of any rays in the cone distinct from it.

An extreme ray is a one-dimensional face of $\mathcal{K}$. By (105), extreme direction Γ_ε is not a point relatively interior to any line segment in $\mathcal{K}\backslash\{\zeta\,\Gamma_\varepsilon\in\mathcal{K} \mid \zeta\geq 0\}$. Thus, by analogy, the corresponding extreme ray $\{\zeta\,\Gamma_\varepsilon\in\mathcal{K} \mid \zeta\geq 0\}$ is not a ray relatively interior to any *plane segment*[2.35] in $\mathcal{K}$.

2.8.1.1 extreme distinction, uniqueness

An extreme *direction* is unique, but its vector representation Γ_ε is not because any positive scaling of it produces another vector in the same (extreme) direction. Hence an extreme direction is *unique* to within a positive scaling. When we say extreme directions are *distinct,* we are referring to distinctness of rays containing them. Nonzero vectors of various length in the same extreme direction are therefore interpreted to be identical extreme directions.[2.36]

The extreme directions of the polyhedral cone in Figure **27** (p.64), for example, correspond to its three edges. For any pointed polyhedral cone, there is a one-to-one correspondence of one-dimensional faces with extreme directions.

The extreme directions of the positive semidefinite cone (§2.9) comprise the infinite set of all symmetric rank-one matrices. [22, §6] [211, §III] It is sometimes prudent to instead consider the less infinite but complete normalized set, for $M>0$ (*confer* (235))

$$\{zz^{\mathrm{T}}\in\mathbb{S}^M \mid \|z\|=1\} \tag{187}$$

The positive semidefinite cone in one dimension $M=1$, $\mathbb{S}_+$ the nonnegative real line, has one extreme direction belonging to its relative interior; an idiosyncrasy of dimension 1.

Pointed closed convex cone $\mathcal{K}=\{\mathbf{0}\}$ has no extreme direction because extreme directions are nonzero by definition.

- If closed convex cone $\mathcal{K}$ is not pointed, then it has no extreme directions and no vertex. [22, §1]

Conversely, pointed closed convex cone $\mathcal{K}$ is equivalent to the convex hull of its vertex and all its extreme directions. [325, §18, p.167] That is the practical utility of extreme direction; to facilitate construction of polyhedral sets, apparent from the *extremes theorem*:

[2.33] We diverge from Rockafellar's extreme direction: "extreme point at infinity".

[2.34] An edge (§2.6.0.0.3) of a convex cone is not necessarily a ray. A convex cone may contain an edge that is a line; *e.g*, a wedge-shaped polyhedral cone ($\mathcal{K}^*$ in Figure **44**).

[2.35] A planar fragment; in this context, a planar cone.

[2.36] Like vectors, an extreme direction can be identified with the Cartesian point at the vector's head with respect to the origin.

2.8.1.1.1 Theorem. (Klee) *Extremes.* [347, §3.6] [325, §18, p.166] (*confer* §2.3.2, §2.12.2.0.1) Any closed convex set containing no lines can be expressed as the convex hull of its extreme points and extreme rays. ⋄

It follows that any element of a convex set containing no lines may be expressed as a linear combination of its extreme elements; *e.g*, §2.9.2.7.1.

2.8.1.2 generators

In the narrowest sense, *generators* for a convex set comprise any collection of points and directions whose convex hull constructs the set.

When the *extremes theorem* applies, the extreme points and directions are called generators of a convex set. An arbitrary collection of generators for a convex set includes its extreme elements as a subset; the set of extreme elements of a convex set is a minimal set of generators for that convex set. Any polyhedral set has a minimal set of generators whose *cardinality* is finite.

When the convex set under scrutiny is a closed convex cone, conic combination of generators during construction is implicit as shown in Example 2.8.1.2.1 and Example 2.10.2.0.1. So, a vertex at the origin (if it exists) becomes benign.

We can, of course, generate affine sets by taking the affine hull of any collection of points and directions. We broaden, thereby, the meaning of generator to be inclusive of all kinds of hulls.

Any hull of generators is loosely called a *vertex-description.* (§2.3.4) Hulls encompass subspaces, so any basis constitutes generators for a vertex-description; span basis $\mathcal{R}(A)$.

2.8.1.2.1 Example. *Application of extremes theorem.*
Given an extreme point at the origin and N extreme rays $\{\zeta\Gamma_i\,,\ i=1\ldots N \mid \zeta\geq 0\}$ (§2.7.0.0.1), denoting the i^{th} extreme direction by $\Gamma_i\in\mathbb{R}^n$, then their convex hull (86) is

$$\begin{aligned}\mathcal{P} &= \left\{[\mathbf{0}\ \ \Gamma_1\ \ \Gamma_2\cdots\Gamma_N]\,a\,\zeta \mid a^{\mathrm{T}}\mathbf{1}=1,\ a\succeq 0,\ \zeta\geq 0\right\}\\ &= \left\{[\Gamma_1\ \ \Gamma_2\cdots\Gamma_N]\,a\,\zeta \mid a^{\mathrm{T}}\mathbf{1}\leq 1,\ a\succeq 0,\ \zeta\geq 0\right\}\\ &= \left\{[\Gamma_1\ \ \Gamma_2\cdots\Gamma_N]\,b \mid b\succeq 0\right\}\subset\mathbb{R}^n\end{aligned}\tag{188}$$

a closed convex set that is simply a conic hull like (103). □

2.8.2 Exposed direction

2.8.2.0.1 Definition. *Exposed point & direction of pointed convex cone.* [325, §18] (*confer* §2.6.1.0.1)

- When a convex cone has a vertex, an exposed point, it resides at the origin; there can be only one.
- In the closure of a pointed convex cone, an *exposed direction* is the direction of a one-dimensional exposed face that is a ray emanating from the origin.
- {exposed directions} $\subseteq$ {extreme directions} △

For a proper cone in vector space $\mathbb{R}^n$ with $n \geq 2$, we can say more:

$$\overline{\{\text{exposed directions}\}} = \{\text{extreme directions}\} \tag{189}$$

It follows from Lemma 2.8.0.0.1 for any pointed closed convex cone, there is one-to-one correspondence of one-dimensional exposed faces with exposed directions; *id est*, there is no one-dimensional exposed face that is not a ray base $\mathbf{0}$.

The pointed closed convex cone $\mathbb{EDM}^{\mathbf{2}}$, for example, is a ray in isomorphic subspace $\mathbb{R}$ whose relative boundary (§2.6.1.4.1) is the origin. The conventionally exposed directions of $\mathbb{EDM}^{\mathbf{2}}$ constitute the empty set $\emptyset \subset \{\text{extreme direction}\}$. This cone has one extreme direction belonging to its relative interior; an idiosyncrasy of dimension 1.

2.8.2.1 Connection between boundary and extremes

2.8.2.1.1 Theorem. *Exposed.* [325, §18.7] (*confer* §2.8.1.1.1)
Any closed convex set $\mathcal{C}$ containing no lines (and whose dimension is at least 2) can be expressed as closure of the convex hull of its exposed points and exposed rays. ⋄

From Theorem 2.8.1.1.1,

$$\left.\begin{aligned} \operatorname{rel}\partial\mathcal{C} &= \overline{\mathcal{C}} \setminus \operatorname{rel}\operatorname{int}\mathcal{C} && (24)\\ &= \overline{\operatorname{conv}\{\text{exposed points and exposed rays}\}} \setminus \operatorname{rel}\operatorname{int}\mathcal{C} \\ &= \operatorname{conv}\{\text{extreme points and extreme rays}\} \setminus \operatorname{rel}\operatorname{int}\mathcal{C} \end{aligned}\right\} \tag{190}$$

Thus each and every extreme point of a convex set (that is not a point) resides on its relative boundary, while each and every extreme direction of a convex set (that is not a halfline and contains no line) resides on its relative boundary because extreme points and directions of such respective sets do not belong to relative interior by definition.

The relationship between extreme sets and the relative boundary actually goes deeper: Any face $\mathcal{F}$ of convex set $\mathcal{C}$ (that is not $\mathcal{C}$ itself) belongs to $\operatorname{rel}\partial\mathcal{C}$, so $\dim\mathcal{F} < \dim\mathcal{C}$. [325, §18.1.3]

2.8.2.2 Converse *caveat*

It is inconsequent to presume that each and every extreme point and direction is necessarily exposed, as might be erroneously inferred from the *conventional boundary definition* (§2.6.1.4.1); although it can correctly be inferred: each and every extreme point and direction belongs to some exposed face.

Arbitrary points residing on the relative boundary of a convex set are not necessarily exposed or extreme points. Similarly, the direction of an arbitrary ray, base $\mathbf{0}$, on the boundary of a convex cone is not necessarily an exposed or extreme direction. For the polyhedral cone illustrated in Figure **27**, for example, there are three two-dimensional exposed faces constituting the entire boundary, each composed of an infinity of rays. Yet there are only three exposed directions.

Neither is an extreme direction on the boundary of a pointed convex cone necessarily an exposed direction. Lift the two-dimensional set in Figure **35**, for example, into three dimensions such that no two points in the set are collinear with the origin. Then its conic hull can have an extreme direction B on the boundary that is not an exposed direction, illustrated in Figure **45**.

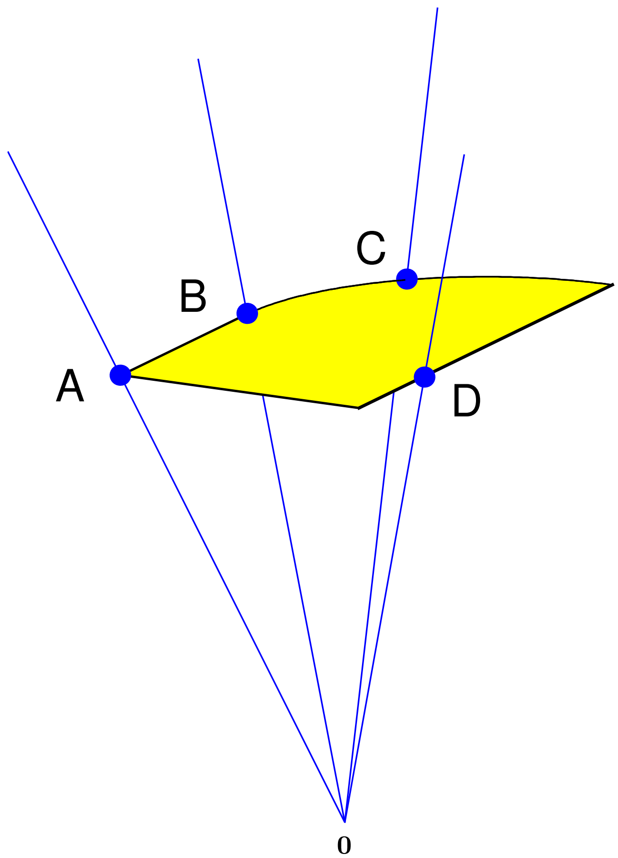


Figure 45: Properties of extreme points carry over to extreme directions. [325, §18] Four rays (drawn truncated) on boundary of conic hull of two-dimensional closed convex set from Figure 35 lifted to $\mathbb{R}^3$. Ray through point A is exposed hence extreme. Extreme direction B on cone boundary is not an exposed direction, although it belongs to the exposed face cone$\{A, B\}$. Extreme ray through C is exposed. Point D is neither an exposed or extreme direction although it belongs to a two-dimensional exposed face of the conic hull.

2.9 Positive semidefinite (PSD) cone

The cone of positive semidefinite matrices studied in this section is arguably the most important of all non-polyhedral cones whose facial structure we completely understand.

−Alexander Barvinok [27, p.78]

2.9.0.0.1 Definition. *Positive semidefinite cone.*
The set of all symmetric positive semidefinite matrices of particular dimension M is called the *positive semidefinite cone*:

$$\begin{aligned} \mathbb{S}_+^M &\triangleq \left\{A \in \mathbb{S}^M \mid A \succeq 0\right\} \\ &= \left\{A \in \mathbb{S}^M \mid y^{\mathrm{T}}A\,y \geq 0 \;\; \forall \|y\| = 1\right\} \\ &= \bigcap_{\|y\|=1} \left\{A \in \mathbb{S}^M \mid \langle yy^{\mathrm{T}} ,\, A\rangle \geq 0\right\} \\ &\equiv \{A \in \mathbb{S}_+^M \mid \operatorname{rank} A \leq M\} \end{aligned} \tag{191}$$

formed by the intersection of an infinite number of halfspaces (§2.4.1.1) in vectorized variable[2.37] A, each halfspace having partial boundary containing the origin in isomorphic $\mathbb{R}^{M(M+1)/2}$. It is a unique immutable proper cone in the ambient space of symmetric matrices $\mathbb{S}^M$.

The positive definite (full-rank) matrices comprise the cone interior

$$\begin{aligned} \operatorname{int} \mathbb{S}_+^M &= \left\{A \in \mathbb{S}^M \mid A \succ 0\right\} \\ &= \left\{A \in \mathbb{S}^M \mid y^{\mathrm{T}}A\,y > 0 \;\; \forall \|y\| = 1\right\} \\ &= \bigcap_{\|y\|=1} \left\{A \in \mathbb{S}^M \mid \langle yy^{\mathrm{T}} ,\, A\rangle > 0\right\} \\ &= \{A \in \mathbb{S}_+^M \mid \operatorname{rank} A = M\} \end{aligned} \tag{192}$$

while all singular positive semidefinite matrices (having at least one 0 eigenvalue) reside on the cone boundary (Figure **46**); (§A.7.5)

$$\begin{aligned} \partial \mathbb{S}_+^M &= \left\{A \in \mathbb{S}^M \mid A \succeq 0\,,\; A \nsucc 0\right\} \\ &= \left\{A \in \mathbb{S}^M \mid \min\{\lambda(A)_i \,,\; i=1 \ldots M\} = 0\right\} \\ &= \left\{A \in \mathbb{S}_+^M \mid \langle yy^{\mathrm{T}} ,\, A\rangle = 0 \;\; \text{for some } \|y\| = 1\right\} \\ &= \{A \in \mathbb{S}_+^M \mid \operatorname{rank} A < M\} \end{aligned} \tag{193}$$

where $\lambda(A) \in \mathbb{R}^M$ holds the eigenvalues of A. △

The only symmetric positive semidefinite matrix in $\mathbb{S}_+^M$ having M 0-eigenvalues resides at the origin. (§A.7.3.0.1)

[2.37] infinite in number when $M > 1$. Because $y^{\mathrm{T}}A\,y = y^{\mathrm{T}}A^{\mathrm{T}}y$, matrix A is almost always assumed symmetric. (§A.2.1)

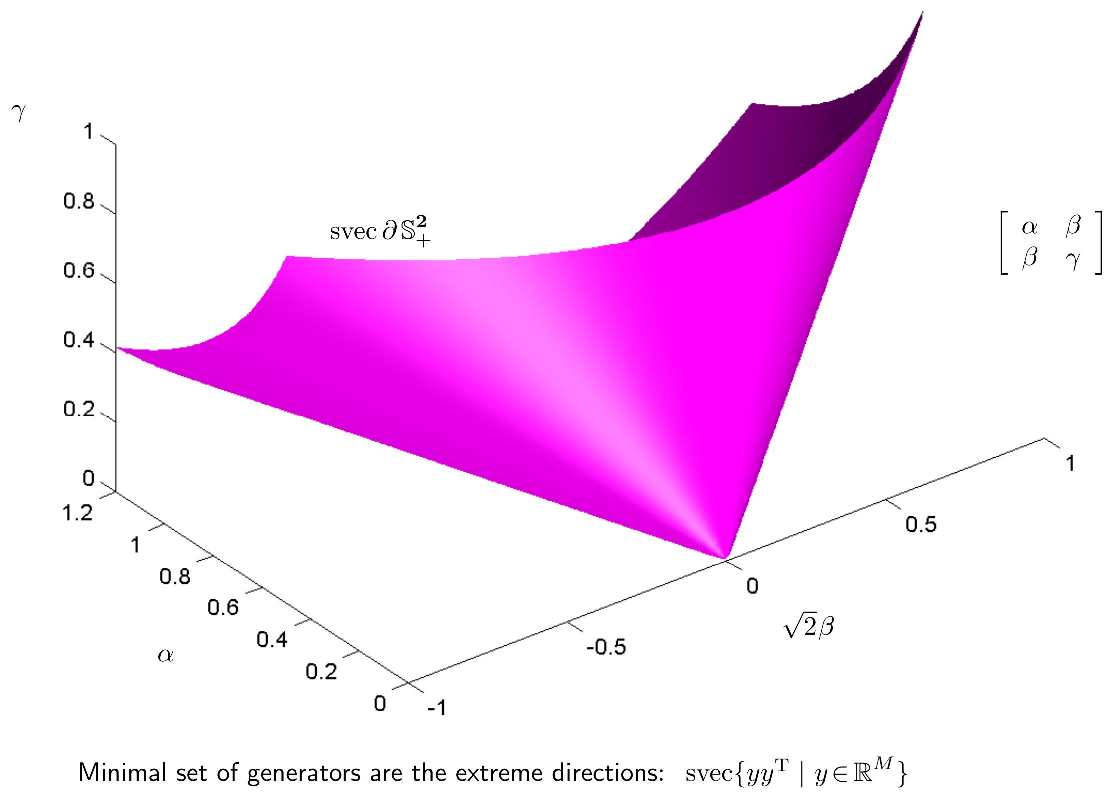


Figure 46: (d'Aspremont) Truncated boundary of PSD cone in $\mathbb{S}^2$ plotted in isometrically isomorphic $\mathbb{R}^3$ via svec (56); 0-contour of smallest eigenvalue (193). Lightest shading is closest, darkest shading is farthest and inside shell. Entire boundary can be constructed from an aggregate of rays (§2.7.0.0.1) emanating exclusively from origin: $\left\{\kappa^2[\,z_1^2 \;\; \sqrt{2}z_1z_2 \;\; z_2^2\,]^{\mathrm{T}} \mid \kappa\in\mathbb{R}\,,\; z\in\mathbb{R}^2\right\}$. A circular cone in this dimension (§2.9.2.8), each and every ray on boundary corresponds to an extreme direction but such is not the case in any higher dimension (*confer* Figure **27**). PSD cone geometry is not as simple in higher dimensions [27, §II.12] although PSD cone is selfdual (376) in ambient real space of symmetric matrices. [211, §II] PSD cone has no two-dimensional face in any dimension, its only extreme point residing at $\mathbf{0}$.

2.9.0.1 Membership

Observe notation $A \succeq 0$ denoting a positive semidefinite matrix;[2.38] meaning (*confer* §2.3.1.1), matrix A belongs to the positive semidefinite cone in the subspace of symmetric matrices whereas $A \succ 0$ denotes membership to that cone's interior. (§2.13.2) Notation $A \succ 0$, denoting a positive definite matrix, can be read: *symmetric matrix A exceeds the origin with respect to the positive semidefinite cone interior.* These notations further imply that coordinates [*sic*] for orthogonal expansion of a positive (semi)definite matrix must be its (nonnegative) positive eigenvalues (§2.13.7.1.1, §E.6.4.1.1) when expanded in its *eigenmatrices* (§A.5.0.3); *id est*, eigenvalues must be (nonnegative) positive.

Generalizing comparison on the real line, the notation $A \succeq B$ denotes comparison with respect to the positive semidefinite cone; (§A.3.1) *id est*, $A \succeq B \Leftrightarrow A - B \in \mathbb{S}_+^M$ but neither matrix A or B necessarily belongs to the positive semidefinite cone. Yet, (1575) $A \succeq B\,,\;\; B \succeq 0 \;\Rightarrow\; A \succeq 0\,;$ *id est*, $A \in \mathbb{S}_+^M$. (*confer* Figure **67**)

2.9.0.1.1 Example. *Equality constraints in semidefinite program* (687).
Employing properties of partial order (§2.7.2.2) for the pointed closed convex positive semidefinite cone, it is easy to show, given $A + S = C$

$$\begin{array}{l} S \succeq 0 \;\Leftrightarrow\; A \preceq C \\ S \succ 0 \;\Leftrightarrow\; A \prec C \end{array} \tag{194}$$

□

2.9.1 Positive semidefinite cone is convex

The set of all positive semidefinite matrices forms a convex cone in the ambient space of symmetric matrices because any pair satisfies definition (175); [218, §7.1] *videlicet*, for all $\zeta_1\,, \zeta_2 \geq 0$ and each and every $A_1\,, A_2 \in \mathbb{S}^M$

$$\zeta_1 A_1 + \zeta_2 A_2 \succeq 0 \;\Leftarrow\; A_1 \succeq 0\,,\; A_2 \succeq 0 \tag{195}$$

a fact easily verified by the definitive test for positive semidefiniteness of a symmetric matrix (§A):

$$A \succeq 0 \Leftrightarrow x^{\mathrm{T}} A\, x \geq 0 \;\text{ for each and every } \|x\| = 1 \tag{196}$$

id est, for $A_1\,, A_2 \succeq 0$ and each and every $\zeta_1\,, \zeta_2 \geq 0$

$$\zeta_1\, x^{\mathrm{T}} A_1\, x + \zeta_2\, x^{\mathrm{T}} A_2\, x \geq 0 \quad \text{for each and every normalized } x \in \mathbb{R}^M \tag{197}$$

The convex cone $\mathbb{S}_+^M$ is more easily visualized in the isomorphic vector space $\mathbb{R}^{M(M+1)/2}$ whose dimension is the number of free variables in a symmetric $M \times M$ matrix. When $M = 2$ the PSD cone is semiinfinite in expanse in $\mathbb{R}^3$, having boundary illustrated in Figure **46**. When $M = 3$ the PSD cone is six-dimensional, and so on.

[2.38] the same as *nonnegative definite matrix*.

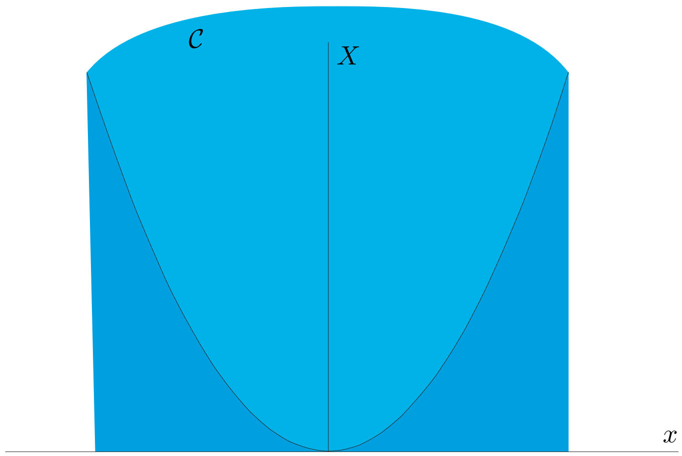


Figure 47: Convex set $\mathcal{C} = \{X \in \mathbb{S} \times x \in \mathbb{R} \mid X \succeq xx^{\mathrm{T}}\}$ drawn truncated.

2.9.1.0.1 Example. *Sets from maps of positive semidefinite cone.*
The set

$$\mathcal{C} = \{X \in \mathbb{S}^n \times x \in \mathbb{R}^n \mid X \succeq xx^{\mathrm{T}}\} \tag{198}$$

is convex because it has *Schur-form*; (§A.4)

$$X - xx^{\mathrm{T}} \succeq 0 \quad \Leftrightarrow \quad f(X\,,\,x) \triangleq \begin{bmatrix} X & x \\ x^{\mathrm{T}} & 1 \end{bmatrix} \succeq 0 \tag{199}$$

e.g, Figure **47**. Set $\mathcal{C}$ is the inverse image (§2.1.9.0.1) of $\mathbb{S}_+^{n+1}$ under affine mapping f. The set $\{X \in \mathbb{S}^n \times x \in \mathbb{R}^n \mid X \preceq xx^{\mathrm{T}}\}$ is not convex, in contrast, having no Schur-form. Yet for fixed $x = x_{\mathrm{p}}$, the set

$$\{X \in \mathbb{S}^n \mid X \preceq x_{\mathrm{p}} x_{\mathrm{p}}^{\mathrm{T}}\} \tag{200}$$

is simply the negative semidefinite cone shifted to $x_{\mathrm{p}} x_{\mathrm{p}}^{\mathrm{T}}$. $\square$

2.9.1.0.2 Example. *Inverse image of positive semidefinite cone.*
Now consider finding the set of all matrices $X \in \mathbb{S}^N$ satisfying

$$AX + B \succeq 0 \tag{201}$$

given $A\,, B \in \mathbb{S}^N$. Define the set

$$\mathcal{X} \triangleq \{X \mid AX + B \succeq 0\} \subseteq \mathbb{S}^N \tag{202}$$

which is the inverse image of the positive semidefinite cone under affine transformation $g(X) \triangleq AX + B$. Set $\mathcal{X}$ must therefore be convex by Theorem 2.1.9.0.1.

Yet we would like a less amorphous characterization of this set, so instead we consider its vectorization (37) which is easier to visualize:

$$\operatorname{vec} g(X) \;=\; \operatorname{vec}(AX) + \operatorname{vec} B \;=\; (I \otimes A)\operatorname{vec} X + \operatorname{vec} B \tag{203}$$

where

$$I \otimes A \triangleq Q\Lambda Q^{\mathrm{T}} \in \mathbb{S}^{N^2} \tag{204}$$

is block-diagonal formed by *Kronecker product* (§A.1.1 *no.*31, §D.1.2.1). Assign

$$\begin{aligned} x &\triangleq \operatorname{vec} X \in \mathbb{R}^{N^2} \\ b &\triangleq \operatorname{vec} B \in \mathbb{R}^{N^2} \end{aligned} \tag{205}$$

then make the equivalent problem: Find

$$\operatorname{vec} \mathcal{X} = \{x \in \mathbb{R}^{N^2} \mid (I \otimes A)x + b \in \mathcal{K}\} \tag{206}$$

where

$$\mathcal{K} \triangleq \operatorname{vec} \mathbb{S}_+^N \tag{207}$$

is a proper cone isometrically isomorphic with the positive semidefinite cone in the subspace of symmetric matrices; the vectorization of every element of $\mathbb{S}_+^N$. Utilizing the diagonalization (204),

$$\begin{aligned} \operatorname{vec} \mathcal{X} &= \{x \mid \Lambda Q^{\mathrm{T}} x \in Q^{\mathrm{T}}(\mathcal{K} - b)\} \\ &= \{x \mid \Phi Q^{\mathrm{T}} x \in \Lambda^{\dagger} Q^{\mathrm{T}}(\mathcal{K} - b)\} \subseteq \mathbb{R}^{N^2} \end{aligned} \tag{208}$$

where † denotes matrix *pseudoinverse* (§E) and

$$\Phi \triangleq \Lambda^{\dagger}\Lambda \tag{209}$$

is a diagonal projection matrix whose entries are either 1 or 0 (§E.3). We have the complementary sum

$$\Phi Q^{\mathrm{T}} x + (I - \Phi)Q^{\mathrm{T}} x = Q^{\mathrm{T}} x \tag{210}$$

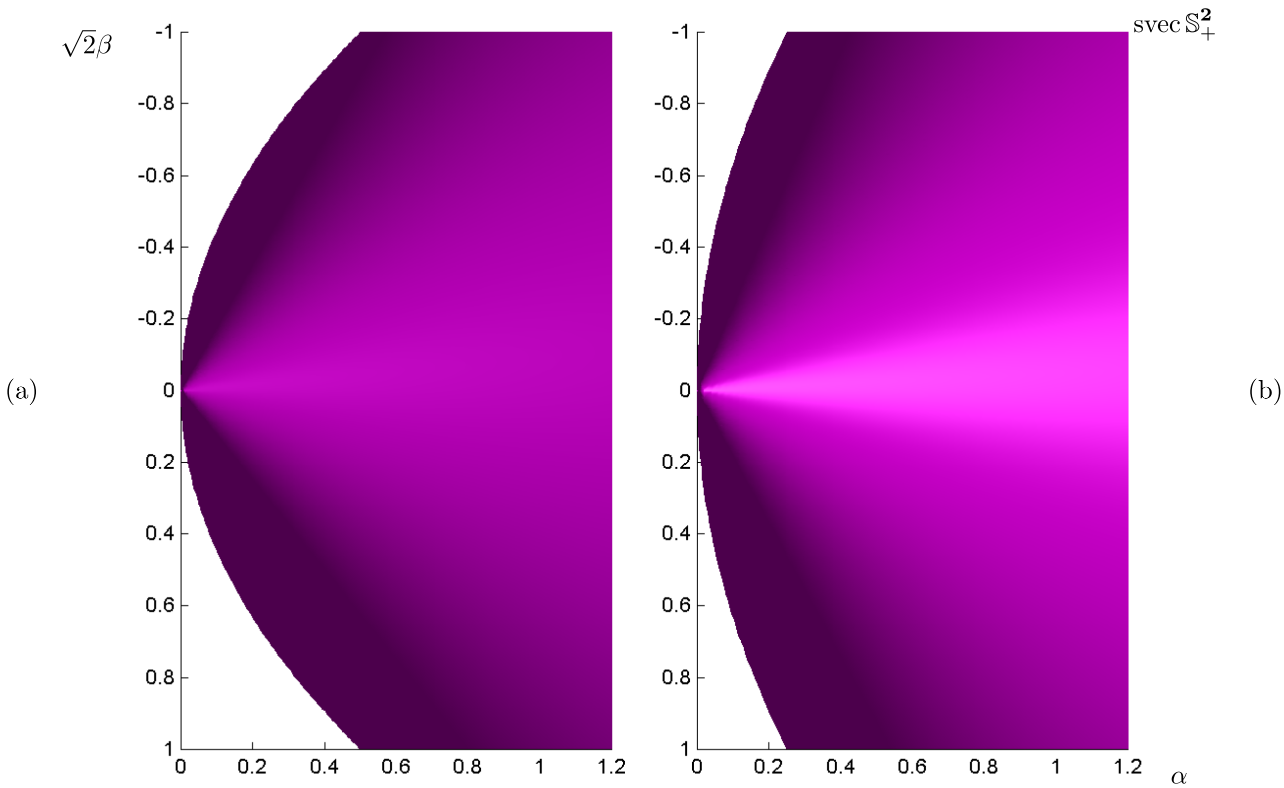
So, adding $(I - \Phi)Q^{\mathrm{T}} x$ to both sides of the membership within (208) admits

$$\begin{aligned} \operatorname{vec} \mathcal{X} &= \{x \in \mathbb{R}^{N^2} \mid Q^{\mathrm{T}} x \in \Lambda^{\dagger} Q^{\mathrm{T}}(\mathcal{K} - b) + (I - \Phi)Q^{\mathrm{T}} x\} \\ &= \{x \mid Q^{\mathrm{T}} x \in \Phi\big(\Lambda^{\dagger} Q^{\mathrm{T}}(\mathcal{K} - b)\big) \oplus (I - \Phi)\mathbb{R}^{N^2}\} \\ &= \{x \in Q\Lambda^{\dagger} Q^{\mathrm{T}}(\mathcal{K} - b) \oplus Q(I - \Phi)\mathbb{R}^{N^2}\} \\ &= (I \otimes A)^{\dagger}(\mathcal{K} - b) \,\oplus\, \mathcal{N}(I \otimes A) \end{aligned} \tag{211}$$

where we used the facts: linear function $Q^{\mathrm{T}} x$ in x on $\mathbb{R}^{N^2}$ is a bijection, and $\Phi\Lambda^{\dagger} = \Lambda^{\dagger}$.

$$\operatorname{vec} \mathcal{X} = (I \otimes A)^{\dagger} \operatorname{vec}(\mathbb{S}_+^N - B) \,\oplus\, \mathcal{N}(I \otimes A) \tag{212}$$

In words, set $\operatorname{vec} \mathcal{X}$ is the vector sum of the translated PSD cone (linearly mapped onto the rowspace of $I \otimes A$ (§E)) and the nullspace of $I \otimes A$ (synthesis of fact from §A.6.3 and §A.7.3.0.1). Should $I \otimes A$ have no nullspace, then $\operatorname{vec} \mathcal{X} = (I \otimes A)^{-1} \operatorname{vec}(\mathbb{S}_+^N - B)$ which is the expected result. □



$$\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$$

Figure 48: **(a)** Projection of truncated PSD cone $\mathbb{S}^{\mathbf{2}}_{+}$, truncated above $\gamma = 1$, on $\alpha\beta$-plane in isometrically isomorphic $\mathbb{R}^{\mathbf{3}}$. View is from above with respect to Figure 46. **(b)** Truncated above $\gamma = 2$. From these plots we might infer, for example, line $\left\{[0 \;\; 1/\sqrt{2} \;\; \gamma\,]^{\mathrm{T}} \mid \gamma \in \mathbb{R}\right\}$ intercepts PSD cone at some large value of γ; in fact, $\gamma = \infty$.

2.9.2 Positive semidefinite cone boundary

For any symmetric positive semidefinite matrix A of rank ρ, there must exist a rank ρ matrix Y such that A be expressible as an outer product in Y; [348, §6.3]

$$A = YY^{\mathrm{T}} \in \mathbb{S}_+^M\ , \quad \operatorname{rank} A = \operatorname{rank} Y = \rho\ , \quad Y \in \mathbb{R}^{M\times\rho} \tag{213}$$

Then the boundary of the positive semidefinite cone may be expressed

$$\partial\mathbb{S}_+^M = \left\{A \in \mathbb{S}_+^M \mid \operatorname{rank} A < M\right\} = \left\{YY^{\mathrm{T}} \mid Y \in \mathbb{R}^{M\times M-1}\right\} \tag{214}$$

Because the boundary of any convex body is obtained with closure of its relative interior (§2.1.7, §2.1.7.2), from (192) we must also have

$$\begin{aligned}\mathbb{S}_+^M = \overline{\left\{A \in \mathbb{S}_+^M \mid \operatorname{rank} A = M\right\}} &= \overline{\left\{YY^{\mathrm{T}} \mid Y \in \mathbb{R}^{M\times M},\ \operatorname{rank} Y = M\right\}} \\ &= \left\{YY^{\mathrm{T}} \mid Y \in \mathbb{R}^{M\times M}\right\}\end{aligned} \tag{215}$$

2.9.2.1 rank ρ subset of the positive semidefinite cone

For the same reason (closure), this applies more generally; for $0 \le \rho \le M$

$$\overline{\left\{A \in \mathbb{S}_+^M \mid \operatorname{rank} A = \rho\right\}} = \left\{A \in \mathbb{S}_+^M \mid \operatorname{rank} A \le \rho\right\} \tag{216}$$

For easy reference, we give such generally nonconvex sets a name: *rank ρ subset of a positive semidefinite cone*. For $\rho < M$ this subset, nonconvex for $M > 1$, resides on the positive semidefinite cone boundary.

2.9.2.1.1 Exercise. *Closure and rank ρ subset.*
Prove equality in (216). ▼

For example,

$$\partial\mathbb{S}_+^M = \overline{\left\{A \in \mathbb{S}_+^M \mid \operatorname{rank} A = M-1\right\}} = \left\{A \in \mathbb{S}_+^M \mid \operatorname{rank} A \le M-1\right\} \tag{217}$$

In $\mathbb{S}^{\mathbf{2}}$, each and every ray on the boundary of the positive semidefinite cone in isomorphic $\mathbb{R}^{\mathbf{3}}$ corresponds to a symmetric rank-1 matrix (Figure **46**), but that does not hold in any higher dimension.

2.9.2.2 Subspace tangent to open rank ρ subset

When the positive semidefinite cone subset in (216) is left unclosed as in

$$\mathcal{M}(\rho) \triangleq \left\{A \in \mathbb{S}_+^N \mid \operatorname{rank} A = \rho\right\} \tag{218}$$

then we can specify a subspace tangent to the positive semidefinite cone at a particular member of manifold $\mathcal{M}(\rho)$. Specifically, the subspace $\mathcal{R}_{\mathcal{M}}$ tangent to manifold $\mathcal{M}(\rho)$ at $B\in\mathcal{M}(\rho)$ [202, §5, prop.1.1]

$$\mathcal{R}_{\mathcal{M}}(B) \triangleq \{XB + BX^{\mathrm{T}} \mid X\in\mathbb{R}^{N\times N}\} \subseteq \mathbb{S}^N \tag{219}$$

has dimension

$$\dim \operatorname{svec} \mathcal{R}_{\mathcal{M}}(B) \;=\; \rho\left(N - \frac{\rho-1}{2}\right) \;=\; \rho(N-\rho) + \frac{\rho(\rho+1)}{2} \tag{220}$$

Tangent subspace $\mathcal{R}_{\mathcal{M}}$ contains no member of the positive semidefinite cone $\mathbb{S}_+^N$ whose rank exceeds ρ.

Subspace $\mathcal{R}_{\mathcal{M}}(B)$ is a hyperplane supporting $\mathbb{S}_+^N$ when $B\in\mathcal{M}(N-1)$. Another good example of tangent subspace is given in §E.7.2.0.2 by (2115); $\mathcal{R}_{\mathcal{M}}(\mathbf{1}\mathbf{1}^{\mathrm{T}}) = \mathbb{S}_c^{N\perp}$, orthogonal complement to the *geometric center subspace*. (Figure **162** p.470)

2.9.2.3 Faces of PSD cone, their dimension *versus* rank

Each and every face of the positive semidefinite cone, having dimension less than that of the cone, is exposed. [262, §6] [231, §2.3.4] Because each and every face of the positive semidefinite cone contains the origin (§2.8.0.0.1), each face belongs to a subspace of dimension the same as the face.

Define $\mathcal{F}(\mathbb{S}_+^M \ni A)$ (171) as the smallest face, that contains a given positive semidefinite matrix A, of positive semidefinite cone $\mathbb{S}_+^M$. Then matrix A, having ordered diagonalization $A = Q\Lambda Q^{\mathrm{T}}\in\mathbb{S}_+^M$ (§A.5.1), is relatively interior to[2.39] [27, §II.12] [120, §31.5.3] [255, §2.4] [256]

$$\begin{aligned}
\mathcal{F}\Big(\mathbb{S}_+^M \ni A\Big) \;&=\; \{X\in\mathbb{S}_+^M \mid \mathcal{N}(X) \supseteq \mathcal{N}(A)\} \\
&=\; \{X\in\mathbb{S}_+^M \mid \langle Q(I-\Lambda\Lambda^\dagger)Q^{\mathrm{T}},\, X\rangle = 0\} \\
&=\; \{Q\Lambda\Lambda^\dagger\Psi\Lambda\Lambda^\dagger Q^{\mathrm{T}} \mid \Psi\in\mathbb{S}_+^M\} \\
&=\; Q\Lambda\Lambda^\dagger\,\mathbb{S}_+^M\Lambda\Lambda^\dagger Q^{\mathrm{T}} \\
&\simeq\; \mathbb{S}_+^{\operatorname{rank} A}
\end{aligned} \tag{221}$$

which is isomorphic with convex cone $\mathbb{S}_+^{\operatorname{rank} A}$; *e.g*, $Q\,\mathbb{S}_+^M Q^{\mathrm{T}} = \mathbb{S}_+^M$. The larger the nullspace of A, the smaller the face. (140) Thus dimension of the smallest face that contains given matrix A is

$$\dim \mathcal{F}\Big(\mathbb{S}_+^M \ni A\Big) \;=\; \operatorname{rank}(A)\big(\operatorname{rank}(A)+1\big)/2 \tag{222}$$

in isomorphic $\mathbb{R}^{M(M+1)/2}$, and each and every face of $\mathbb{S}_+^M$ is isomorphic with a positive semidefinite cone having dimension the same as the face. Observe: not all dimensions are represented, and the only zero-dimensional face is the origin. The positive semidefinite cone has no facets, for example.

[2.39] For $X\in\mathbb{S}_+^M$, $A = Q\Lambda Q^{\mathrm{T}}\in\mathbb{S}_+^M$, show $\mathcal{N}(X)\supseteq\mathcal{N}(A) \;\Leftrightarrow\; \langle Q(I-\Lambda\Lambda^\dagger)Q^{\mathrm{T}},\, X\rangle = 0$.
Given $\langle Q(I-\Lambda\Lambda^\dagger)Q^{\mathrm{T}},\, X\rangle = 0 \;\Leftrightarrow\; \mathcal{R}(X)\perp\mathcal{N}(A)$. (§A.7.4)
(⇒) Assume $\mathcal{N}(X)\supseteq\mathcal{N}(A)$, then $\mathcal{R}(X)\perp\mathcal{N}(X)\supseteq\mathcal{N}(A)$.
(⇐) Assume $\mathcal{R}(X)\perp\mathcal{N}(A)$, then $X\,Q(I-\Lambda\Lambda^\dagger)Q^{\mathrm{T}} = \mathbf{0} \;\Rightarrow\; \mathcal{N}(X)\supseteq\mathcal{N}(A)$. ♦

2.9.2.3.1 Table: Rank k *versus* dimension of $\mathbb{S}^3_+$ faces

	k	$\dim \mathcal{F}(\mathbb{S}^3_+ \ni \text{rank-}k \text{ matrix})$
	0	0
boundary	≤ 1	1
	≤ 2	3
interior	≤ 3	6

For positive semidefinite cone $\mathbb{S}^2_+$ in isometrically isomorphic $\mathbb{R}^3$ depicted in Figure 46, rank-2 matrices belong to the interior of that face having dimension 3 (the entire closed cone), rank-1 matrices belong to relative interior of a face having dimension[2.40] 1, and the only rank-0 matrix is the point at the origin (the zero-dimensional face).

2.9.2.3.2 Exercise. *Bijective isometry.*

Prove that the smallest face of positive semidefinite cone $\mathbb{S}^M_+$, containing a particular full-rank matrix A having ordered diagonalization $Q\Lambda Q^{\mathrm{T}}$, is the entire cone: *id est*, prove $Q\,\mathbb{S}^M_+ Q^{\mathrm{T}} = \mathbb{S}^M_+$ from (221). ▼

2.9.2.4 rank-k face of PSD cone

Any rank-$k < M$ positive semidefinite matrix A belongs to a face, of positive semidefinite cone $\mathbb{S}^M_+$, described by intersection with a hyperplane: for ordered diagonalization of $A = Q\Lambda Q^{\mathrm{T}} \in \mathbb{S}^M_+ \ni \operatorname{rank}(A) = k < M$

$$\begin{aligned}
\mathcal{F}\left(\mathbb{S}^M_+ \ni A\right) &= \{X \in \mathbb{S}^M_+ \mid \langle Q(I - \Lambda\Lambda^\dagger)Q^{\mathrm{T}},\, X\rangle = 0\} \\
&= \left\{X \in \mathbb{S}^M_+ \;\middle|\; \left\langle Q\left(I - \begin{bmatrix} I \in \mathbb{S}^k & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}\right)Q^{\mathrm{T}},\, X\right\rangle = 0\right\} \\
&= \mathbb{S}^M_+ \cap \underline{\partial\mathcal{H}}_+ \\
&\simeq \mathbb{S}^k_+
\end{aligned} \tag{223}$$

Faces are doubly indexed: continuously indexed by orthogonal matrix Q, and discretely indexed by rank k. Each and every orthogonal matrix Q makes projectors $Q(:\,,k+1\!:\!M)Q(:\,,k+1\!:\!M)^{\mathrm{T}}$ indexed by k, in other words, each projector describing a normal[2.41] $\operatorname{svec}\big(Q(:\,,k+1\!:\!M)Q(:\,,k+1\!:\!M)^{\mathrm{T}}\big)$ to a supporting hyperplane $\underline{\partial\mathcal{H}}_+$ (containing the origin) exposing a face (§2.11) of the positive semidefinite cone containing rank-k (and less) matrices.

2.9.2.4.1 Exercise. *Simultaneously diagonalizable means commutative.*

Given diagonalization of rank-$k \leq M$ positive semidefinite matrix $A = Q\Lambda Q^{\mathrm{T}}$ and any particular $\Psi \succeq 0$, both in $\mathbb{S}^M$ from (221), show how $I - \Lambda\Lambda^\dagger$ and $\Lambda\Lambda^\dagger\Psi\Lambda\Lambda^\dagger$ share a complete set of eigenvectors. ▼

[2.40] The boundary constitutes all the one-dimensional faces, in $\mathbb{R}^3$, which are rays emanating from the origin.

[2.41] Any vectorized nonzero matrix $\in \mathbb{S}^M_+$ is normal to a hyperplane supporting $\mathbb{S}^M_+$ (§2.13.1) because PSD cone is selfdual. Normal on boundary exposes nonzero face by (329) (330).

2.9.2.5 PSD cone face containing principal submatrix

A *principal submatrix* of a matrix $A\in\mathbb{R}^{M\times M}$ is formed by discarding any particular subset of its rows and columns having the same indices. There are $M!/(1!(M-1)!)$ principal 1×1 submatrices, $M!/(2!(M-2)!)$ principal 2×2 submatrices, and so on, totaling 2^M-1 principal submatrices including A itself. Principal submatrices of a symmetric matrix are symmetric. A given symmetric matrix has rank ρ iff it has a nonsingular principal $\rho\times\rho$ submatrix but none larger. [315, §5-10] By loading vector y in test $y^{\mathrm T}Ay$ (§A.2) with various binary patterns, it follows that any principal submatrix must be positive (semi)definite whenever A is (Theorem A.3.1.0.4). If positive semidefinite matrix $A\in\mathbb{S}_+^M$ has principal submatrix of dimension ρ with rank r, then $\operatorname{rank}A\leq M-\rho+r$ by (1625).

Because each and every principal submatrix of a positive semidefinite matrix in $\mathbb{S}^M$ is positive semidefinite, then each principal submatrix belongs to a certain face of positive semidefinite cone $\mathbb{S}_+^M$ by (222). Of special interest are full-rank positive semidefinite principal submatrices, for then description of smallest face becomes simpler. We can find the smallest face, that contains a particular complete full-rank principal submatrix of A, by embedding that submatrix in a $\mathbf{0}$ matrix of the same dimension as A: Were Φ a binary diagonal matrix

$$\Phi=\delta^2(\Phi)\in\mathbb{S}^M\,,\qquad \Phi_{ii}\in\{0\,,1\} \tag{224}$$

having diagonal entry 0 corresponding to a discarded row and column from $A\in\mathbb{S}_+^M$, then any principal submatrix[2.42] so embedded can be expressed $\Phi A\Phi$; *id est*, for an embedded principal submatrix $\Phi A\Phi\in\mathbb{S}_+^M \ni \operatorname{rank}\Phi A\Phi=\operatorname{rank}\Phi\leq\operatorname{rank}A$

$$\begin{aligned}\mathcal{F}\Big(\mathbb{S}_+^M\ni\Phi A\Phi\Big) &= \{X\in\mathbb{S}_+^M \mid \mathcal{N}(X)\supseteq\mathcal{N}(\Phi A\Phi)\}\\ &= \{X\in\mathbb{S}_+^M \mid \langle I-\Phi\,,\,X\rangle=0\}\\ &= \{\Phi\Psi\Phi \mid \Psi\in\mathbb{S}_+^M\}\\ &\simeq \mathbb{S}_+^{\operatorname{rank}\Phi}\end{aligned} \tag{225}$$

Smallest face that contains an embedded principal submatrix, whose rank is not necessarily full, may be expressed like (221): For embedded principal submatrix $\Phi A\Phi\in\mathbb{S}_+^M \ni \operatorname{rank}\Phi A\Phi\leq\operatorname{rank}\Phi$, apply ordered diagonalization instead to

$$\hat\Phi^{\mathrm T}A\,\hat\Phi=U\Upsilon U^{\mathrm T}\in\mathbb{S}_+^{\operatorname{rank}\Phi} \tag{226}$$

where $U^{-1}=U^{\mathrm T}$ is an orthogonal matrix and $\Upsilon=\delta^2(\Upsilon)$ is diagonal. Then

$$\begin{aligned}\mathcal{F}\Big(\mathbb{S}_+^M\ni\Phi A\Phi\Big) &= \{X\in\mathbb{S}_+^M \mid \mathcal{N}(X)\supseteq\mathcal{N}(\Phi A\Phi)\}\\ &= \{X\in\mathbb{S}_+^M \mid \langle\hat\Phi U(I-\Upsilon\Upsilon^\dagger)U^{\mathrm T}\hat\Phi^{\mathrm T}+I-\Phi\,,\,X\rangle=0\}\\ &= \{\hat\Phi U\Upsilon\Upsilon^\dagger\Psi\Upsilon\Upsilon^\dagger U^{\mathrm T}\hat\Phi^{\mathrm T} \mid \Psi\in\mathbb{S}_+^{\operatorname{rank}\Phi}\}\\ &\simeq \mathbb{S}_+^{\operatorname{rank}\Phi A\Phi}\end{aligned} \tag{227}$$

[2.42] To express a leading principal submatrix, for example, $\Phi=\begin{bmatrix} I & \mathbf{0}\\ \mathbf{0}^{\mathrm T} & \mathbf{0}\end{bmatrix}$.

where binary diagonal matrix Φ is partitioned into nonzero and zero columns by permutation $\Xi \in \mathbb{R}^{M\times M}$;

$$\Phi\Xi^{\mathrm T} \triangleq [\,\hat\Phi \;\; \mathbf{0}\,] \in \mathbb{R}^{M\times M}, \quad \operatorname{rank}\hat\Phi = \operatorname{rank}\Phi\,, \quad \Phi = \hat\Phi\hat\Phi^{\mathrm T} \in \mathbb{S}^M, \quad \hat\Phi^{\mathrm T}\hat\Phi = I \tag{228}$$

Any embedded principal submatrix may be expressed

$$\Phi A \Phi = \hat\Phi\hat\Phi^{\mathrm T} A\,\hat\Phi\hat\Phi^{\mathrm T} \in \mathbb{S}^M_+ \tag{229}$$

where $\hat\Phi^{\mathrm T} A\,\hat\Phi \in \mathbb{S}_+^{\operatorname{rank}\Phi}$ extracts the principal submatrix whereas $\hat\Phi\hat\Phi^{\mathrm T} A\,\hat\Phi\hat\Phi^{\mathrm T}$ embeds it.

2.9.2.5.1 Example. *Smallest face containing disparate elements.*
Smallest face formula (221) can be altered to accommodate a union of points $\{A_i \in \mathbb{S}^M_+\}$:

$$\mathcal{F}\left(\mathbb{S}^M_+ \supset \bigcup_i A_i\right) = \left\{X \in \mathbb{S}^M_+ \;\middle|\; \mathcal{N}(X) \supseteq \bigcap_i \mathcal{N}(A_i)\right\} \tag{230}$$

To see that, imagine two vectorized matrices A_1 and A_2 on diametrically opposed sides of the positive semidefinite cone $\mathbb{S}^{\mathbf 2}_+$ boundary pictured in Figure **46**. Regard $\operatorname{svec} A_1$ as normal to a hyperplane in $\mathbb{R}^{\mathbf 3}$ containing a vectorized basis for its nullspace: $\operatorname{svec}\operatorname{basis}\mathcal{N}(A_1)$ (§2.5.3). Similarly, there is a second hyperplane containing $\operatorname{svec}\operatorname{basis}\mathcal{N}(A_2)$ having normal $\operatorname{svec} A_2$. While each hyperplane is two-dimensional, each nullspace has only one affine dimension because A_1 and A_2 are rank-1. Because our interest is only that part of the nullspace in the positive semidefinite cone, then by

$$\langle X\,,\,A_i\rangle = 0 \;\Leftrightarrow\; XA_i = A_iX = \mathbf{0}\,, \qquad X, A_i \in \mathbb{S}^M_+ \tag{1681}$$

we may ignore the fact that vectorized nullspace $\operatorname{svec}\operatorname{basis}\mathcal{N}(A_i)$ is a proper subspace of the hyperplane. We may think instead in terms of whole hyperplanes because equivalence (1681) says that the positive semidefinite cone effectively filters that subset of the hyperplane, whose normal is A_i , constituting $\mathcal{N}(A_i)$.

And so hyperplane intersection makes a line intersecting the positive semidefinite cone $\mathbb{S}^{\mathbf 2}_+$ but only at the origin. In this hypothetical example, smallest face containing those two matrices therefore comprises the entire cone because every positive semidefinite matrix has nullspace containing $\mathbf{0}$. The smaller the intersection, the larger the smallest face. □

2.9.2.5.2 Exercise. *Disparate elements.*
Prove that (230) holds for an arbitrary set $\{A_i \in \mathbb{S}^M_+ \;\forall\, i \in \mathcal{I}\}$. One way is by showing $\bigcap \mathcal{N}(A_i) \cap \mathbb{S}^M_+ = \operatorname{conv}(\{A_i\})^\perp \cap \mathbb{S}^M_+$; with perpendicularity $^\perp$ as in (372).[2.43] ▼

2.9.2.6 face of all PSD matrices having same principal submatrix

Now we ask what is the smallest face of the positive semidefinite cone containing all matrices having a complete principal submatrix in common; in other words, that face containing all PSD matrices (of any rank) with particular entries fixed − the smallest

[2.43] Hint: (1681) (2014).

face containing all PSD matrices whose fixed entries correspond to some given embedded principal submatrix $\Phi A\Phi$. To maintain generality,[2.44] we move an extracted principal submatrix $\hat{\Phi}^{\mathrm{T}}A\,\hat{\Phi}\in\mathbb{S}_+^{\operatorname{rank}\Phi}$ into leading position via permutation Ξ from (228): for $A\in\mathbb{S}_+^M$

$$\Xi A\Xi^{\mathrm{T}} \triangleq \begin{bmatrix} \hat{\Phi}^{\mathrm{T}}A\,\hat{\Phi} & B \\ B^{\mathrm{T}} & C \end{bmatrix} \in \mathbb{S}_+^M \tag{231}$$

By properties of partitioned PSD matrices in §A.4.0.1,

$$\operatorname{basis}\mathcal{N}\!\left(\begin{bmatrix} \hat{\Phi}^{\mathrm{T}}A\,\hat{\Phi} & B \\ B^{\mathrm{T}} & C \end{bmatrix}\right) \supseteq \begin{bmatrix} \mathbf{0} \\ I - CC^{\dagger} \end{bmatrix} \tag{232}$$

Hence $\mathcal{N}(\Xi X\Xi^{\mathrm{T}}) \supseteq \mathcal{N}(\Xi A\Xi^{\mathrm{T}}) \not\supseteq \operatorname{span}\begin{bmatrix}\mathbf{0}\\ I\end{bmatrix}$ in a smallest face $\mathcal{F}$ formula[2.45] because all PSD matrices, given fixed principal submatrix, are admitted: Define a set of all PSD matrices

$$\mathcal{S} \triangleq \left\{ A = \Xi^{\mathrm{T}}\begin{bmatrix} \hat{\Phi}^{\mathrm{T}}A\,\hat{\Phi} & B \\ B^{\mathrm{T}} & C \end{bmatrix}\Xi \succeq 0 \;\middle|\; B\in\mathbb{R}^{\operatorname{rank}\Phi\times M-\operatorname{rank}\Phi},\; C\in\mathbb{S}_+^{M-\operatorname{rank}\Phi} \right\} \tag{233}$$

having fixed embedded principal submatrix $\Phi A\Phi = \Xi^{\mathrm{T}}\begin{bmatrix} \hat{\Phi}^{\mathrm{T}}A\,\hat{\Phi} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}\Xi$. So

$$\begin{aligned}
\mathcal{F}\Big(\mathbb{S}_+^M \supseteq \mathcal{S}\Big) &= \Big\{X\in\mathbb{S}_+^M \mid \mathcal{N}(X)\supseteq\mathcal{N}(\mathcal{S})\Big\} \\
&= \{X\in\mathbb{S}_+^M \mid \langle \hat{\Phi}U(I-\Upsilon\Upsilon^{\dagger})U^{\mathrm{T}}\hat{\Phi}^{\mathrm{T}},\, X\rangle = 0\} \\
&= \left\{\Xi^{\mathrm{T}}\begin{bmatrix} U\Upsilon\Upsilon^{\dagger} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & I \end{bmatrix}\Psi\begin{bmatrix} \Upsilon\Upsilon^{\dagger}U^{\mathrm{T}} & \mathbf{0} \\ \mathbf{0}^{\mathrm{T}} & I \end{bmatrix}\Xi \;\middle|\; \Psi\in\mathbb{S}_+^M\right\} \\
&\simeq \mathbb{S}_+^{M-\operatorname{rank}\Phi+\operatorname{rank}\Phi A\Phi}
\end{aligned} \tag{234}$$

$\Xi = I$ whenever $\Phi A\Phi$ denotes a leading principal submatrix. Smallest face of the positive semidefinite cone, containing all matrices having the same full-rank principal submatrix ($\Upsilon\Upsilon^{\dagger} = I$, $\Upsilon\succeq 0$), is the entire cone (Exercise 2.9.2.3.2).

2.9.2.7 Extreme directions of positive semidefinite cone

Because the positive semidefinite cone is pointed (§2.7.2.1.2), there is a one-to-one correspondence of one-dimensional faces with extreme directions in any dimension M; *id est*, because of the *cone faces lemma* (§2.8.0.0.1) and direct correspondence of exposed faces to faces of $\mathbb{S}_+^M$, it follows: there is no one-dimensional face of the positive semidefinite cone that is not a ray emanating from the origin.

Symmetric dyads constitute the set of all extreme directions: For $M>1$

$$\{yy^{\mathrm{T}}\in\mathbb{S}^M \mid y\in\mathbb{R}^M\} \subset \partial\mathbb{S}_+^M \tag{235}$$

[2.44] to fix any principal submatrix; not only leading principal submatrices.
[2.45] meaning, more pertinently, $I-\Phi$ is dropped from (227).

this superset of extreme directions (infinite in number, *confer* (187)) for the positive semidefinite cone is, generally, a subset of the boundary. By *extremes theorem* 2.8.1.1.1, the convex hull of extreme rays and origin is the positive semidefinite cone: (§2.8.1.2.1)

$$\text{conv}\{yy^{\mathrm T}\in\mathbb{S}^M \mid y\in\mathbb{R}^M\} = \left\{\sum_{i=1}^{\infty} b_i\, z_i z_i^{\mathrm T} \mid z_i\in\mathbb{R}^M,\ b\succeq 0\right\} = \mathbb{S}_+^M \tag{236}$$

For two-dimensional matrices ($M=2$, Figure **46**)

$$\{yy^{\mathrm T}\in\mathbb{S}^{\mathbf 2} \mid y\in\mathbb{R}^{\mathbf 2}\} = \partial\mathbb{S}_+^{\mathbf 2} \tag{237}$$

while for one-dimensional matrices, in exception, ($M=1$, §2.7)

$$\{yy\in\mathbb{S} \mid y\neq\mathbf{0}\} = \text{int}\,\mathbb{S}_+ \tag{238}$$

Each and every extreme direction $yy^{\mathrm T}$ makes the same angle with the Identity matrix in isomorphic $\mathbb{R}^{M(M+1)/2}$, dependent only on dimension; *videlicet*,[2.46]

$$\measuredangle(yy^{\mathrm T},\, I) = \arccos\frac{\langle yy^{\mathrm T},\, I\rangle}{\|yy^{\mathrm T}\|_{\mathrm F}\ \|I\|_{\mathrm F}} = \arccos\left(\frac{1}{\sqrt{M}}\right) \quad \forall\, y\in\mathbb{R}^M \tag{239}$$

This means the positive semidefinite cone broadens in higher dimension.

2.9.2.7.1 Example. *Positive semidefinite matrix from extreme directions.*
Diagonalizability (§A.5) of symmetric matrices yields the following results:

Any positive semidefinite matrix (1539) in $\mathbb{S}^M$ can be written in the form

$$A \;=\; \sum_{i=1}^{M}\lambda_i\, z_i z_i^{\mathrm T} \;=\; \hat A\hat A^{\mathrm T} \;=\; \sum_i \hat a_i\hat a_i^{\mathrm T} \succeq\; 0\,, \qquad \lambda\succeq 0 \tag{240}$$

a conic combination of linearly independent extreme directions ($\hat a_i\hat a_i^{\mathrm T}$ or $z_i z_i^{\mathrm T}$ where $\|z_i\|=1$), where λ is a vector of eigenvalues.

If we limit consideration to all symmetric positive semidefinite matrices bounded via unity trace

$$\mathcal{C} \triangleq \{A\succeq 0 \mid \text{tr}\,A=1\} \qquad (91)$$

then any matrix A from that set may be expressed as a convex combination of linearly independent extreme directions;

$$A = \sum_{i=1}^{M}\lambda_i\, z_i z_i^{\mathrm T} \in \mathcal{C}\;, \qquad \mathbf{1}^{\mathrm T}\lambda = 1\,, \quad \lambda\succeq 0 \tag{241}$$

Implications are:

1. set $\mathcal{C}$ is convex (an intersection of PSD cone with hyperplane),
2. because the set of eigenvalues corresponding to a given square matrix A is unique (§A.5.0.1), no single eigenvalue can exceed 1; *id est*, $I\succeq A$
3. and the converse holds: set $\mathcal{C}$ is an instance of Fantope (91). □

2.46 Analogy with respect to the *EDM cone* is considered in [201, p.162] where it is found: angle is not constant. Extreme directions of the EDM cone can be found in §6.4.3.2. The cone's axis is $-E=\mathbf{1}\mathbf{1}^{\mathrm T}-I$ (1183).

2.9.2.7.2 Exercise. *Extreme directions of positive semidefinite cone.*
Prove, directly from definition (186), that symmetric dyads (235) constitute the set of all extreme directions of the positive semidefinite cone. ▼

2.9.2.8 Positive semidefinite cone is generally not circular

Extreme angle equation (239) suggests that the positive semidefinite cone might be invariant to rotation about its axis of revolution; *id est*, a circular cone. We investigate this now:

2.9.2.8.1 Definition. *Circular cone*:[2.47]
a pointed closed convex cone having hyperspherical sections orthogonal to its *axis of revolution* about which the cone is invariant to rotation. △

A *conic section* is the intersection of a cone with any hyperplane. In three dimensions, an intersecting plane perpendicular to a circular cone's axis of revolution produces a section bounded by a circle. (Figure **49**) A prominent example of a circular cone in convex analysis is Lorentz cone (178). We also find that the positive semidefinite cone and cone of Euclidean distance matrices are circular cones, but only in low dimension.

The positive semidefinite cone has axis of revolution that is the ray (base $\mathbf{0}$) through the Identity matrix I. Consider a set of normalized extreme directions of the positive semidefinite cone: for some arbitrary positive constant $a\in\mathbb{R}_+$

$$\{yy^{\mathrm T}\in\mathbb{S}^M \mid \|y\| = \sqrt{a}\} \subset \partial\mathbb{S}_+^M \tag{242}$$

The distance from each extreme direction to the axis of revolution is radius

$$\mathrm{R} \triangleq \inf_c \|yy^{\mathrm T} - cI\|_{\mathrm F} = a\sqrt{1-\frac{1}{M}} \tag{243}$$

which is the distance from $yy^{\mathrm T}$ to $\frac{a}{M}I$; the length of vector $yy^{\mathrm T}-\frac{a}{M}I$.

Because distance R (in a particular dimension) from the axis of revolution to each and every normalized extreme direction is identical, the extreme directions lie on the boundary of a hypersphere in isometrically isomorphic $\mathbb{R}^{M(M+1)/2}$. From Example 2.9.2.7.1, the convex hull (excluding vertex at the origin) of the normalized extreme directions is a conic section

$$\mathcal{C} \triangleq \operatorname{conv}\{yy^{\mathrm T} \mid y\in\mathbb{R}^M,\ y^{\mathrm T}y=a\} = \mathbb{S}_+^M \cap \{A\in\mathbb{S}^M \mid \langle I\,,\,A\rangle = a\} \tag{244}$$

orthogonal to Identity matrix I;

$$\left\langle \mathcal{C}-\frac{a}{M}I\ ,\ I\right\rangle = \operatorname{tr}(\mathcal{C}-\frac{a}{M}I) = 0 \tag{245}$$

[2.47] A circular cone is assumed convex throughout, although not so by other authors. We also assume a *right* circular cone.

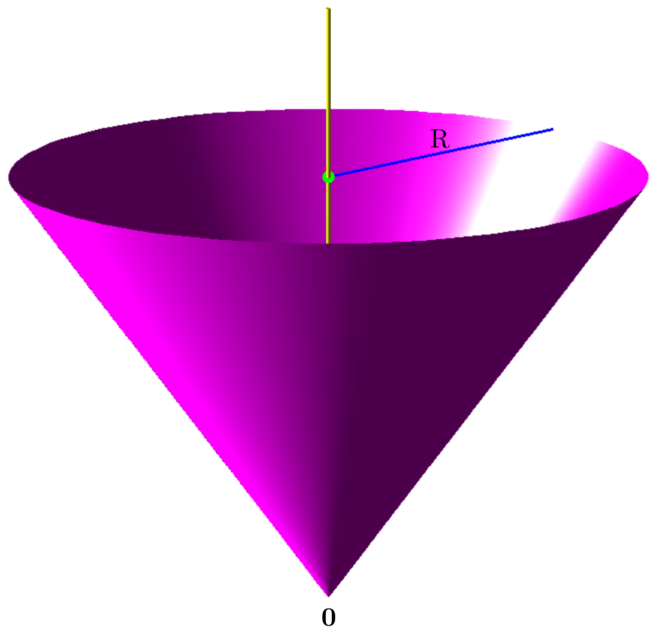


Figure 49: This solid circular cone in $\mathbb{R}^3$ continues upward infinitely. Axis of revolution is illustrated as vertical line through origin. R represents radius: distance measured from an extreme direction to axis of revolution. Were this a Lorentz cone, any plane slice containing axis of revolution would make a right angle.

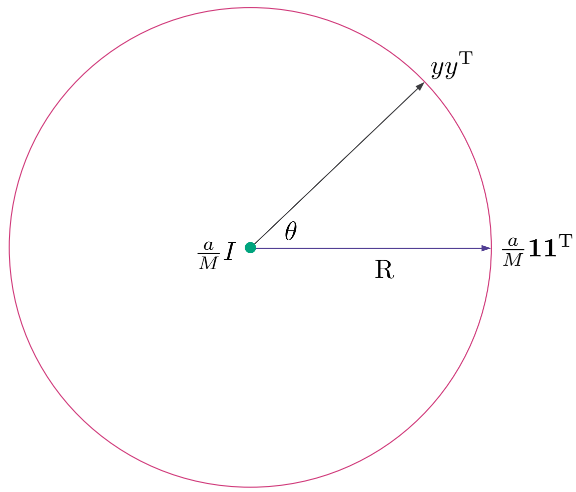


Figure 50: Illustrated is a section, perpendicular to axis of revolution, of circular cone from Figure **49**. Radius R is distance from any extreme direction to axis at $\frac{a}{M}I$. Vector $\frac{a}{M}\mathbf{1}\mathbf{1}^{\mathrm{T}}$ is an arbitrary reference by which to measure angle θ.

Proof. Although the positive semidefinite cone possesses some characteristics of a circular cone, we can show it is not by demonstrating shortage of extreme directions; *id est*, some extreme directions corresponding to each and every angle of rotation about the axis of revolution are nonexistent: Referring to Figure **50**, [419, §1-7]

$$\cos\theta = \frac{\left\langle \frac{a}{M}\mathbf{1}\mathbf{1}^{\mathrm{T}} - \frac{a}{M}I \;,\; yy^{\mathrm{T}} - \frac{a}{M}I \right\rangle}{a^2(1-\frac{1}{M})} \tag{246}$$

Solving for vector y we get

$$a\big(1 + (M-1)\cos\theta\big) = (\mathbf{1}^{\mathrm{T}}y)^2 \tag{247}$$

which does not have real solution $\forall\, 0 \leq \theta \leq 2\pi$ in every matrix dimension M. ⧫

From the foregoing proof we can conclude that the positive semidefinite cone might be circular but only in matrix dimensions 1 and 2. Because of a shortage of extreme directions, conic section (244) cannot be hyperspherical by the *extremes theorem* (§2.8.1.1.1, Figure **45**).

2.9.2.8.2 Exercise. *Circular semidefinite cone.*

Prove the positive semidefinite cone to be circular in matrix dimensions 1 and 2 while it is a rotation of Lorentz cone (178) in matrix dimension 2 .[2.48] ▼

[2.48] Hint: Given cone $\left\{\begin{bmatrix} \alpha & \beta/\sqrt{2} \\ \beta/\sqrt{2} & \gamma \end{bmatrix} \mid \sqrt{\alpha^2+\beta^2} \leq \gamma \right\}$, show $\frac{1}{\sqrt{2}}\begin{bmatrix} \gamma+\alpha & \beta \\ \beta & \gamma-\alpha \end{bmatrix}$ is a vector rotation that is positive semidefinite under the same inequality.

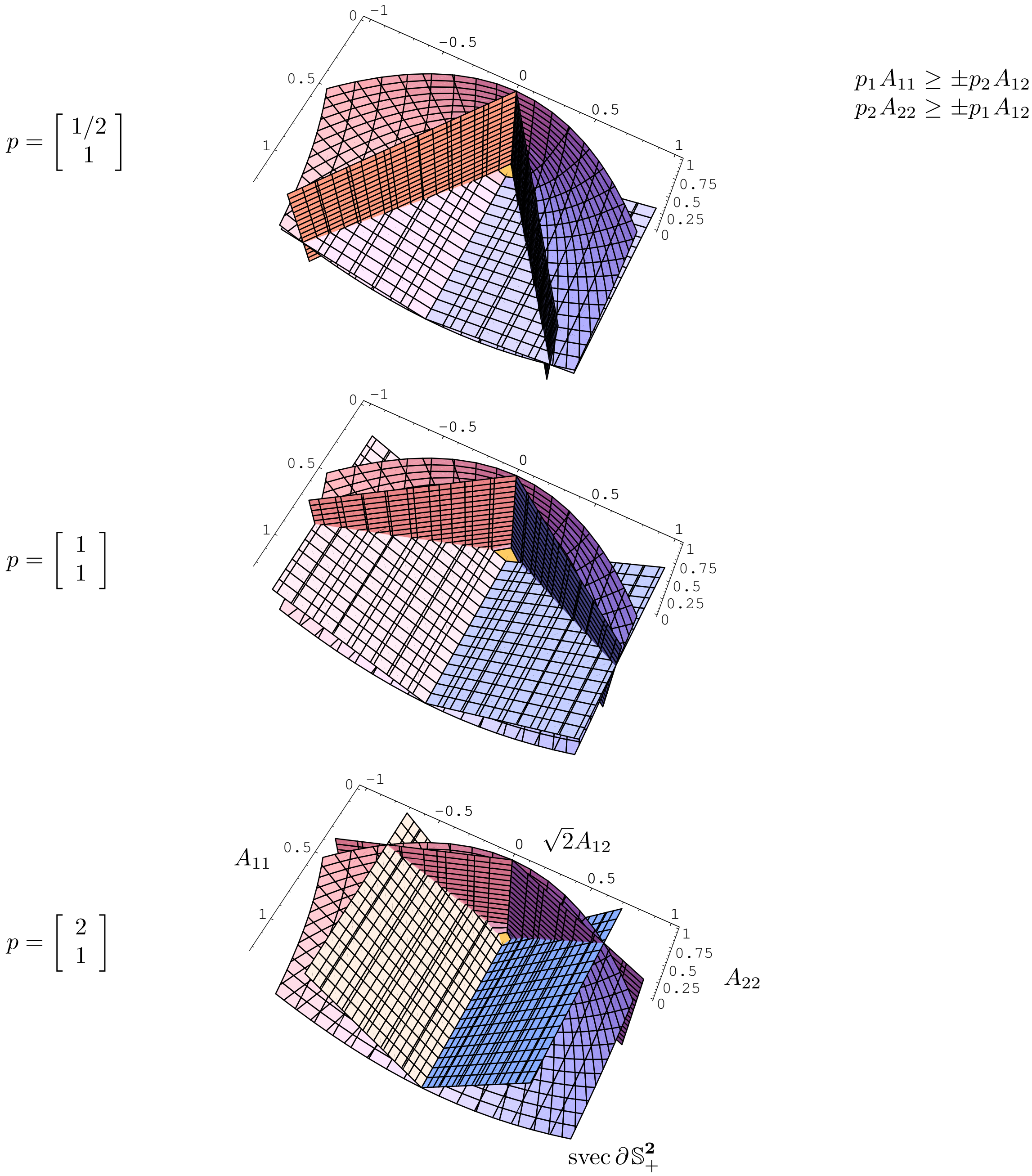


Figure 51: Proper polyhedral cone $\mathcal{K}$, created by intersection of halfspaces, inscribes PSD cone in isometrically isomorphic $\mathbb{R}^3$ as predicted by *Geršgorin discs theorem* for $A = [A_{ij}] \in \mathbb{S}^2$. Hyperplanes supporting $\mathcal{K}$ intersect along boundary of PSD cone. Four extreme directions of $\mathcal{K}$ coincide with extreme directions of PSD cone.

2.9.2.8.3 Example. *Positive semidefinite cone inscription in three dimensions.*

Theorem. *Geršgorin discs.* [218, §6.1] [383] [271, p.140]
For $p\in\mathbb{R}^m_+$ given $A=[A_{ij}]\in\mathbb{S}^m$, then all eigenvalues of A belong to the union of m closed intervals on the real line;

$$\lambda(A) \in \bigcup_{i=1}^m \left\{ \xi\in\mathbb{R} \;\middle|\; |\xi - A_{ii}| \leq \varrho_i \triangleq \frac{1}{p_i}\sum_{\substack{j=1\\ j\neq i}}^m p_j|A_{ij}| \right\} = \bigcup_{i=1}^m [A_{ii}-\varrho_i\;,\;A_{ii}+\varrho_i] \tag{248}$$

Furthermore, if a union of k of these m [intervals] forms a connected region that is disjoint from all the remaining $n-k$ [intervals], then there are precisely k eigenvalues of A in this region. ⋄

To apply the theorem to determine positive semidefiniteness of symmetric matrix A, we observe that for each i we must have

$$A_{ii} \geq \varrho_i \tag{249}$$

Suppose

$$m = 2 \tag{250}$$

so $A\in\mathbb{S}^{\mathbf{2}}$. Vectorizing A as in (56), $\operatorname{svec} A$ belongs to isometrically isomorphic $\mathbb{R}^{\mathbf{3}}$. Then we have $m2^{m-1}=4$ inequalities, in the matrix entries A_{ij} with Geršgorin parameters $p=[p_i]\in\mathbb{R}^{\mathbf{2}}_+$,

$$\begin{aligned} p_1 A_{11} &\geq \pm p_2 A_{12} \\ p_2 A_{22} &\geq \pm p_1 A_{12} \end{aligned} \tag{251}$$

which describe an intersection of four halfspaces in $\mathbb{R}^{m(m+1)/2}$. That intersection creates the proper polyhedral cone $\mathcal{K}$ (§2.12.1) whose construction is illustrated in Figure **51**. Drawn truncated is the boundary of the positive semidefinite cone $\operatorname{svec}\mathbb{S}^{\mathbf{2}}_+$ and the bounding hyperplanes supporting $\mathcal{K}$.

Created by means of Geršgorin discs, $\mathcal{K}$ always belongs to the positive semidefinite cone for any nonnegative value of $p\in\mathbb{R}^m_+$. Hence any point in $\mathcal{K}$ corresponds to some positive semidefinite matrix A. Only the extreme directions of $\mathcal{K}$ intersect the positive semidefinite cone boundary in this dimension; the four extreme directions of $\mathcal{K}$ are extreme directions of the positive semidefinite cone. As p_1/p_2 increases in value from 0, two extreme directions of $\mathcal{K}$ sweep the entire boundary of this positive semidefinite cone. Because the entire positive semidefinite cone can be swept by $\mathcal{K}$, the system of linear inequalities

$$Y^{\mathrm{T}}\operatorname{svec} A \triangleq \begin{bmatrix} p_1 & \pm p_2/\sqrt{2} & 0 \\ 0 & \pm p_1/\sqrt{2} & p_2 \end{bmatrix} \operatorname{svec} A \succeq 0 \tag{252}$$

(when made dynamic) can replace a semidefinite constraint $A\succeq 0$; *id est*, for

$$\mathcal{K} = \{z \mid Y^{\mathrm{T}}z \succeq 0\} \subset \operatorname{svec}\mathbb{S}^m_+ \tag{253}$$

given p where $Y\in\mathbb{R}^{m(m+1)/2\times m2^{m-1}}$

$$\operatorname{svec} A \in \mathcal{K} \;\Rightarrow\; A\in\mathbb{S}^m_+ \tag{254}$$

but

$$\exists\, p \ni Y^{\mathrm{T}} \operatorname{svec} A \succeq 0 \;\Leftrightarrow\; A \succeq 0 \tag{255}$$

In other words, *diagonal dominance* [218, p.349, §7.2.3]

$$A_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^{m} |A_{ij}| \;, \quad \forall\, i = 1 \ldots m \tag{256}$$

is generally only a sufficient condition for membership to the PSD cone. But by dynamic weighting p in this dimension, diagonal dominance was made necessary and sufficient.

□

In higher dimension ($m > 2$), boundary of the positive semidefinite cone is no longer constituted completely by its extreme directions (symmetric rank-one matrices); its geometry becomes intricate. How all the extreme directions can be swept by an inscribed polyhedral cone,[2.49] similarly to the foregoing example, remains an open question.

2.9.2.8.4 Exercise. *Dual inscription.*
Find dual proper polyhedral cone $\mathcal{K}^*$ from Figure **51**. ▼

2.9.2.9 Boundary constituents of the positive semidefinite cone

2.9.2.9.1 Lemma. *Sum of positive semidefinite matrices.* (*confer* (1555))
For $A, B \in \mathbb{S}_+^M$

$$\operatorname{rank}(A + B) = \operatorname{rank}(\mu A + (1 - \mu) B) \tag{257}$$

over open interval $(0, 1)$ of μ. ◇

Proof. Any positive semidefinite matrix belonging to the PSD cone has an eigenvalue decomposition that is a positively scaled sum of linearly independent symmetric dyads. By the *linearly independent dyads definition* in §B.1.1.0.1, rank of the sum $A + B$ is equivalent to the number of linearly independent dyads constituting it. Linear independence is insensitive to further positive scaling by μ. The assumption of positive semidefiniteness prevents annihilation of any dyad from the sum $A + B$. ♦

2.9.2.9.2 Example. *Rank function quasiconcavity.* (*confer* §3.8)
For $A, B \in \mathbb{R}^{m \times n}$ [218, §0.4]

$$\operatorname{rank} A + \operatorname{rank} B \;\geq\; \operatorname{rank}(A + B) \tag{258}$$

that follows from the fact [348, §3.6]

$$\dim \mathcal{R}(A) + \dim \mathcal{R}(B) \;=\; \dim \mathcal{R}(A + B) + \dim\big(\mathcal{R}(A) \cap \mathcal{R}(B)\big) \tag{259}$$

For $A, B \in \mathbb{S}_+^M$

$$\operatorname{rank} A + \operatorname{rank} B \;\geq\; \operatorname{rank}(A + B) \;\geq\; \min\{\operatorname{rank} A \,,\, \operatorname{rank} B\} \tag{1555}$$

[2.49] It is not necessary to sweep the entire boundary in higher dimension.

that follows from the fact

$$\mathcal{N}(A+B) = \mathcal{N}(A) \cap \mathcal{N}(B)\ , \qquad A, B \in \mathbb{S}_+^M \tag{160}$$

Rank is a *quasiconcave* function on $\mathbb{S}_+^M$ because the right-hand inequality in (1555) has the concave form (645); *videlicet*, Lemma 2.9.2.9.1. □

From this example we see, unlike convex functions, *quasiconvex* functions are not necessarily continuous. (§3.8) We also glean:

2.9.2.9.3 Theorem. *Convex subsets of positive semidefinite cone.*
Subsets of the positive semidefinite cone $\mathbb{S}_+^M$, for $0 \le \rho \le M$

$$\mathbb{S}_+^M(\rho) \triangleq \{X \in \mathbb{S}_+^M \mid \operatorname{rank} X \ge \rho\} \tag{260}$$

are pointed convex cones, but not closed unless $\rho = 0$; *id est*, $\mathbb{S}_+^M(0) = \mathbb{S}_+^M$. ◇

Proof. Given ρ, a subset $\mathbb{S}_+^M(\rho)$ is convex if and only if convex combination of any two members has rank at least ρ. That is confirmed by applying identity (257) from Lemma 2.9.2.9.1 to (1555); *id est*, for $A, B \in \mathbb{S}_+^M(\rho)$ on closed interval $[0,1]$ of μ

$$\operatorname{rank}(\mu A + (1-\mu)B) \ \ge\ \min\{\operatorname{rank} A\,,\ \operatorname{rank} B\} \tag{261}$$

It can similarly be shown, almost identically to proof of the lemma, any conic combination of A, B in subset $\mathbb{S}_+^M(\rho)$ remains a member; *id est*, $\forall\, \zeta, \xi \ge 0$

$$\operatorname{rank}(\zeta A + \xi B) \ \ge\ \min\{\operatorname{rank}(\zeta A)\,,\ \operatorname{rank}(\xi B)\} \tag{262}$$

Therefore, $\mathbb{S}_+^M(\rho)$ is a convex cone. ♦

Another proof of convexity can be made by projection arguments:

2.9.2.10 Projection on $\mathbb{S}_+^M(\rho)$

Because these cones $\mathbb{S}_+^M(\rho)$ indexed by ρ (260) are convex, projection on them is straightforward. Given a symmetric matrix H having diagonalization $H \triangleq Q\Lambda Q^{\mathrm{T}} \in \mathbb{S}^M$ (§A.5.1) with eigenvalues Λ arranged in nonincreasing order, then its *Euclidean projection* (minimum-distance projection) on $\mathbb{S}_+^M(\rho)$

$$P_{\mathbb{S}_+^M(\rho)} H = Q\Upsilon^\star Q^{\mathrm{T}} \tag{263}$$

corresponds to a map of its eigenvalues:

$$\Upsilon^\star_{ii} = \begin{cases} \max\{\epsilon\,,\ \Lambda_{ii}\}\ , & i=1\ldots\rho \\ \max\{0\,,\ \Lambda_{ii}\}\ , & i=\rho+1\ldots M \end{cases} \tag{264}$$

where ϵ is positive but arbitrarily close to 0.

2.9.2.10.1 Exercise. *Projection on open convex cones.*
Prove (264) using Theorem E.9.2.0.1. ▼

Because each $H \in \mathbb{S}^M$ has unique projection on $\mathbb{S}^M_+(\rho)$ (despite possibility of repeated eigenvalues in Λ), we may conclude it is a convex set by the *Bunt-Motzkin theorem* (§E.9.0.0.1).

Compare (264) to the well-known result regarding Euclidean projection on a rank ρ subset of the positive semidefinite cone (§2.9.2.1)

$$\mathbb{S}^M_+ \setminus \mathbb{S}^M_+(\rho+1) \;=\; \{X \in \mathbb{S}^M_+ \mid \operatorname{rank} X \leq \rho\} \qquad (216)$$

$$P_{\mathbb{S}^M_+ \setminus \mathbb{S}^M_+(\rho+1)} H = Q\,\Upsilon^\star Q^{\mathrm{T}} \tag{265}$$

As proved in §7.1.4, this projection of H corresponds to the eigenvalue map

$$\Upsilon^\star_{ii} = \begin{cases} \max\{0\,,\,\Lambda_{ii}\}\,, & i=1\ldots\rho \\ 0\,, & i=\rho+1\ldots M \end{cases} \qquad (1430)$$

Together these two results (264) and (1430) mean: A higher-rank solution to projection on the positive semidefinite cone lies arbitrarily close to any given lower-rank projection, but not *vice versa*. Were the number of nonnegative eigenvalues in Λ known *a priori* not to exceed ρ, then these two different projections would produce identical results in the limit $\epsilon \to 0$.

2.9.2.11 Uniting constituents

Interior of the PSD cone $\operatorname{int} \mathbb{S}^M_+$ is convex by Theorem 2.9.2.9.3, for example, because all positive semidefinite matrices having rank M constitute the cone interior.

All positive semidefinite matrices of rank less than M constitute the cone boundary; an amalgam of positive semidefinite matrices of different rank. Thus each nonconvex subset of positive semidefinite matrices, for $0 < \rho < M$

$$\{Y \in \mathbb{S}^M_+ \mid \operatorname{rank} Y = \rho\} \tag{266}$$

having rank ρ successively 1 lower than M, appends a nonconvex constituent to the cone boundary; but only in their union is the boundary complete: (*confer* §2.9.2)

$$\partial \mathbb{S}^M_+ = \bigcup_{\rho=0}^{M-1} \{Y \in \mathbb{S}^M_+ \mid \operatorname{rank} Y = \rho\} \tag{267}$$

The composite sequence, the cone interior in union with each successive constituent, remains convex at each step; *id est*, for $0 \leq k \leq M$

$$\bigcup_{\rho=k}^{M} \{Y \in \mathbb{S}^M_+ \mid \operatorname{rank} Y = \rho\} \tag{268}$$

is convex for each k by Theorem 2.9.2.9.3.

2.9.2.12 Peeling constituents

Proceeding the other way: To peel constituents off the complete positive semidefinite cone boundary, one starts by removing the origin; the only rank-0 positive semidefinite matrix. What remains is convex. Next, the extreme directions are removed because they constitute all the rank-1 positive semidefinite matrices. What remains is again convex, and so on. Proceeding in this manner eventually removes the entire boundary leaving, at last, the convex interior of the PSD cone; all the positive definite matrices.

2.9.2.12.1 Exercise. *Difference* $A-B$.
What about a difference of matrices A,B belonging to the PSD cone? Show:

- Difference of any two points on the boundary belongs to the boundary or exterior.
- Difference $A-B$, where A belongs to the boundary while B is interior, belongs to the exterior. ▼

2.9.3 Barvinok's proposition

Barvinok posits existence and quantifies an upper bound on rank of a positive semidefinite matrix belonging to the intersection of the PSD cone with an affine subset:

2.9.3.0.1 Proposition. *Affine intersection with PSD cone.* [27, §II.13] [25, §2.2]
Consider finding a matrix $X\in\mathbb{S}^N$ satisfying

$$X\succeq 0\,,\qquad \langle A_j\,,\,X\rangle=b_j\,,\quad j=1\ldots m \tag{269}$$

given nonzero linearly independent (vectorized) $A_j\in\mathbb{S}^N$ and real b_j. Define the affine subset

$$\mathcal{A}\triangleq\{X\mid\langle A_j\,,\,X\rangle=b_j\,,\;j=1\ldots m\}\subseteq\mathbb{S}^N \tag{270}$$

If the intersection $\mathcal{A}\cap\mathbb{S}_+^N$ is nonempty given a number m of equalities, then there exists a matrix $X\in\mathcal{A}\cap\mathbb{S}_+^N$ such that

$$\operatorname{rank}X\,(\operatorname{rank}X+1)/2\le m \tag{271}$$

whence the upper bound[2.50]

$$\operatorname{rank}X\le\left\lfloor\frac{\sqrt{8m+1}-1}{2}\right\rfloor \tag{272}$$

Given desired rank instead, equivalently,

$$m<(\operatorname{rank}X+1)(\operatorname{rank}X+2)/2 \tag{273}$$

[2.50] §4.1.2.2 contains an intuitive explanation. This bound is itself limited above, of course, by N; a *tight* limit corresponding to an interior point of $\mathbb{S}_+^N$.

An extreme point of $\mathcal{A} \cap \mathbb{S}_+^N$ satisfies (272) and (273). (*confer* §4.1.2.2) A matrix $X \triangleq R^{\mathrm{T}}R$ is an extreme point if and only if the smallest face, that contains X, of $\mathcal{A} \cap \mathbb{S}_+^N$ has dimension 0; [255, §2.4] [256] *id est*, iff

$$\begin{aligned}&\dim \mathcal{F}\Big((\mathcal{A} \cap \mathbb{S}_+^N) \ni X\Big)\\ &= \operatorname{rank}(X)\big(\operatorname{rank}(X)+1\big)/2 - \operatorname{rank}\big[\operatorname{svec} RA_1R^{\mathrm{T}} \;\; \operatorname{svec} RA_2R^{\mathrm{T}} \cdots \operatorname{svec} RA_mR^{\mathrm{T}}\,\big]\end{aligned} \tag{274}$$

(171) equals 0 in isomorphic $\mathbb{R}^{N(N+1)/2}$.

Now the intersection $\mathcal{A} \cap \mathbb{S}_+^N$ is assumed bounded: Assume a given nonzero upper bound ρ on rank, a number of equalities

$$m = (\rho+1)(\rho+2)/2 \tag{275}$$

and matrix dimension $N \geq \rho+2 \geq 3$. If the intersection is nonempty and bounded, then there exists a matrix $X \in \mathcal{A} \cap \mathbb{S}_+^N$ such that

$$\operatorname{rank} X \leq \rho \tag{276}$$

This represents a tightening of the upper bound; a reduction by exactly 1 of the bound provided by (272) given the same specified number m (275) of equalities; *id est*,

$$\operatorname{rank} X \leq \frac{\sqrt{8m+1}-1}{2} - 1 \tag{277}$$

◇

2.10 Conic independence (c.i.)

In contrast to extreme direction, the property *conically independent direction* is more generally applicable; inclusive of all closed convex cones (not only pointed closed convex cones). Arbitrary given directions $\{\Gamma_i \in \mathbb{R}^n,\ i=1\ldots N\}$ comprise a *conically independent set* if and only if (*confer* §2.1.2, §2.4.2.3)

$$\Gamma_i\,\zeta_i + \cdots + \Gamma_j\,\zeta_j - \Gamma_\ell = \mathbf{0}\,, \qquad i \neq \cdots \neq j \neq \ell = 1\ldots N \tag{278}$$

has no solution $\zeta \in \mathbb{R}_+^N$ ($\zeta_i \in \mathbb{R}_+$); in words, iff no direction from the given set can be expressed as a conic combination of those remaining; *e.g*, Figure **52** [391, *conic independence* test (278) MATLAB]. Arranging any set of generators for a particular closed convex cone in a matrix columnar,

$$X \triangleq [\,\Gamma_1 \;\; \Gamma_2 \;\; \cdots \;\; \Gamma_N\,] \in \mathbb{R}^{n\times N} \tag{279}$$

then this test of conic independence (278) may be expressed as a set of linear *feasibility problems*: for $\ell = 1\ldots N$

$$\begin{array}{rl}\text{find} & \zeta \in \mathbb{R}^N\\ \text{subject to} & X\zeta = \Gamma_\ell\\ & \zeta \succeq 0\\ & \zeta_\ell = 0\end{array} \tag{280}$$

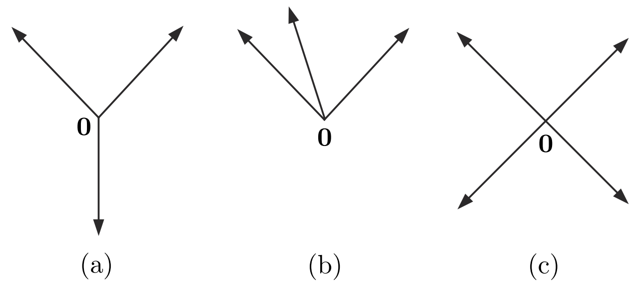


Figure 52: Vectors in $\mathbb{R}^2$: **(a)** affinely and conically independent, **(b)** affinely independent but not conically independent, **(c)** conically independent but not affinely independent. None of the examples exhibits linear independence. (In general, a.i. $\not\Leftrightarrow$ c.i.)

If feasible for any particular ℓ , then the set is not conically independent.

To find all conically independent directions from a set via (280), generator Γ_ℓ must be removed from the set once it is found (feasible) conically dependent on remaining generators in X . So, to continue testing remaining generators when Γ_ℓ is found to be dependent, Γ_ℓ must be discarded from matrix X before proceeding. A generator Γ_ℓ that is instead found conically independent of remaining generators in X , on the other hand, is conically independent of any subset of remaining generators. A c.i. set thus found is not necessarily unique.

It is evident that linear independence (l.i.) of N directions implies their conic independence;

- l.i. $\Rightarrow$ c.i.

which suggests, number of l.i. generators in the columns of X cannot exceed number of c.i. generators. Denoting by $\mathbf{k}$ the number of conically independent generators contained in X , we have the most fundamental rank inequality for convex cones

$$\dim \operatorname{aff} \mathcal{K} = \dim \operatorname{aff}[\,\mathbf{0}\ \ X\,] = \operatorname{rank} X \leq \mathbf{k} \leq N \tag{281}$$

Whereas N directions in n dimensions can no longer be linearly independent once N exceeds n , conic independence remains possible:

2.10.0.0.1 Table: Maximum number of c.i. directions

dimension n	$\sup \mathbf{k}$ (pointed)	$\sup \mathbf{k}$ (not pointed)
0	0	0
1	1	2
2	2	4
3	∞	∞
$\vdots$	$\vdots$	$\vdots$

Assuming veracity of this table, there is an apparent vastness between two and three dimensions. The finite numbers of conically independent directions indicate:

- Convex cones in dimensions $0, 1$, and 2 must be polyhedral. (§2.12.1)

Conic independence is certainly one convex idea that cannot be completely explained by a two-dimensional picture as Barvinok suggests [27, p.vii].

From this table it is also evident that dimension of Euclidean space cannot exceed the number of conically independent directions possible;

- $n \le \sup \mathbf{k}$

2.10.1 Preservation of conic independence

Independence in the linear (§2.1.2.1), affine (§2.4.2.4), and conic senses can be preserved under linear transformation. Suppose a matrix $X \in \mathbb{R}^{n\times N}$ (279) holds a conically independent set columnar. Consider a transformation on the domain of such matrices

$$T(X) : \mathbb{R}^{n\times N} \to \mathbb{R}^{n\times N} \triangleq XY \tag{282}$$

where fixed matrix $Y \triangleq [\, y_1 \;\; y_2 \; \cdots \; y_N \,] \in \mathbb{R}^{N\times N}$ represents linear operator T. Conic independence of $\{Xy_i \in \mathbb{R}^n,\ i=1\ldots N\}$ demands, by definition (278),

$$Xy_i\,\zeta_i + \cdots + Xy_j\,\zeta_j - Xy_\ell = \mathbf{0}\,, \qquad i \neq \cdots \neq j \neq \ell = 1 \ldots N \tag{283}$$

have no solution $\zeta \in \mathbb{R}^N_+$. That is ensured by conic independence of $\{y_i \in \mathbb{R}^N\}$ and by $\mathcal{R}(Y) \cap \mathcal{N}(X) = \mathbf{0}$; seen by factoring out X.

2.10.1.1 linear maps of cones

[22, §7] If $\mathcal{K}$ is a convex cone in Euclidean space $\mathcal{R}$ and T is any linear mapping from $\mathcal{R}$ to Euclidean space $\mathcal{M}$, then $T(\mathcal{K})$ is a convex cone in $\mathcal{M}$ and $x \preceq y$ with respect to $\mathcal{K}$ implies $T(x) \preceq T(y)$ with respect to $T(\mathcal{K})$. If $\mathcal{K}$ is full-dimensional in $\mathcal{R}$, then so is $T(\mathcal{K})$ in $\mathcal{M}$.

If T is a linear bijection, then $x \preceq y \;\Leftrightarrow\; T(x) \preceq T(y)$. If $\mathcal{K}$ is pointed, then so is $T(\mathcal{K})$. And if $\mathcal{K}$ is closed, so is $T(\mathcal{K})$. If $\mathcal{F}$ is a face of $\mathcal{K}$, then $T(\mathcal{F})$ is a face of $T(\mathcal{K})$.

Linear bijection is only a sufficient condition for pointedness and closedness; convex polyhedra (§2.12) are invariant to any linear or inverse linear transformation [27, §I.9] [325, p.44, thm.19.3].

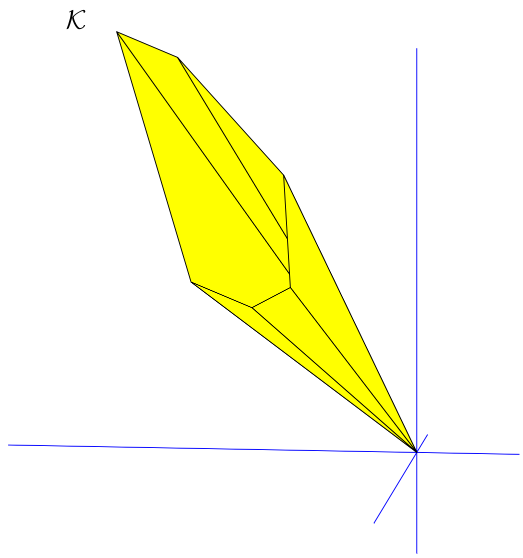
2.10.2 Pointed closed convex $\mathcal{K}$ & conic independence

The following bullets can be derived from definitions (186) and (278) in conjunction with the *extremes theorem* (§2.8.1.1.1):

The set of all extreme directions from a pointed closed convex cone $\mathcal{K} \subset \mathbb{R}^n$ is not necessarily a linearly independent set, yet it must be a conically independent set; (compare Figure **27** on page 64 with Figure **53**a)

- {extreme directions} $\Rightarrow$ {c.i.}

(a)



(b)

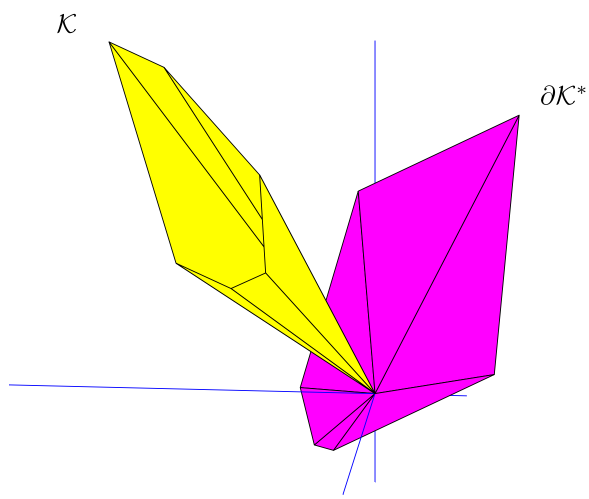


Figure 53: **(a)** A pointed polyhedral cone (drawn truncated) in $\mathbb{R}^3$ having six facets. The extreme directions, corresponding to six edges emanating from the origin, are generators for this cone; not linearly independent but they must be conically independent. **(b)** The boundary of dual cone $\mathcal{K}^*$ (drawn truncated) is now added to the drawing of same $\mathcal{K}$. $\mathcal{K}^*$ is polyhedral, proper, and has the same number of extreme directions as $\mathcal{K}$ has facets.

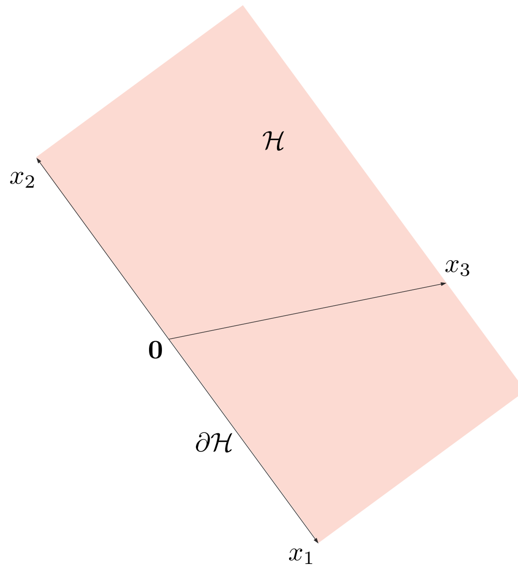


Figure 54: Minimal set of generators $X = [\, x_1 \;\, x_2 \;\, x_3 \,] \in \mathbb{R}^{\mathbf{2\times 3}}$ (not extreme directions) for halfspace about origin; affinely and conically independent.

When a conically independent set of directions from pointed closed convex cone $\mathcal{K}$ is known to comprise generators, conversely, then all directions from that set must be extreme directions of the cone;

- $\{$extreme directions$\} \Leftrightarrow \{$c.i. generators of pointed closed convex $\mathcal{K}\}$

Barker & Carlson [22, §1] call the extreme directions a *minimal generating set* for a pointed closed convex cone. A minimal set of generators is therefore a conically independent set of generators, and *vice versa*,[2.51] for a pointed closed convex cone.

An arbitrary collection of n or fewer distinct extreme directions from pointed closed convex cone $\mathcal{K} \subset \mathbb{R}^n$ is not necessarily a linearly independent set; *e.g*, dual extreme directions (482) from Example 2.13.11.0.3.

- $\{\leq n$ extreme directions in $\mathbb{R}^n\} \nRightarrow \{$l.i.$\}$

Linear dependence of few extreme directions is another convex idea that cannot be explained by a two-dimensional picture as Barvinok suggests [27, p.vii]; indeed, it only first comes to light in four dimensions! But there is a converse: [347, §2.10.9]

- $\{$extreme directions$\} \Leftarrow \{$l.i. generators of closed convex $\mathcal{K}\}$

[2.51] This converse does not hold for nonpointed closed convex cones as Table **2.10.0.0.1** implies; *e.g*, ponder four conically independent generators for a plane ($n=2$, Figure **52**).

2.10.2.0.1 Example. *Vertex-description of halfspace $\mathcal{H}$ about origin.*
From $n+1$ points in $\mathbb{R}^n$ we can make a vertex-description of a convex cone that is a halfspace $\mathcal{H}$, where $\{x_\ell \in \mathbb{R}^n,\ \ell=1\ldots n\}$ constitutes a minimal set of generators for a hyperplane $\partial\mathcal{H}$ through the origin. An example is illustrated in Figure **54**. By demanding the augmented set $\{x_\ell \in \mathbb{R}^n,\ \ell=1\ldots n+1\}$ be affinely independent (we want vector x_{n+1} not parallel to $\partial\mathcal{H}$), then

$$\begin{aligned}\mathcal{H} &= \bigcup_{\zeta\geq 0}(\zeta\, x_{n+1} + \partial\mathcal{H})\\ &= \{\zeta\, x_{n+1} + \operatorname{cone}\{x_\ell \in \mathbb{R}^n,\ \ell=1\ldots n\} \mid \zeta\geq 0\}\\ &= \operatorname{cone}\{x_\ell \in \mathbb{R}^n,\ \ell=1\ldots n+1\}\end{aligned} \tag{284}$$

a union of parallel hyperplanes. Cardinality is one step beyond dimension of the ambient space, but $\{x_\ell\ \forall\,\ell\}$ is a minimal set of generators for this convex cone $\mathcal{H}$ which has no extreme elements. □

2.10.2.0.2 Exercise. *Enumerating conically independent directions.*
Do Example 2.10.2.0.1 in $\mathbb{R}$ and $\mathbb{R}^{\mathbf{3}}$ by drawing two figures corresponding to Figure **54** and enumerating $n+1$ conically independent generators for each. Describe a nonpointed polyhedral cone in three dimensions having more than 8 conically independent generators. (*confer* Table **2.10.0.0.1**) ▼

2.10.3 Utility of conic independence

Perhaps the most useful application of conic independence is determination of the intersection of closed convex cones from their halfspace-descriptions, or representation of the sum of closed convex cones from their vertex-descriptions.

$\bigcap \mathcal{K}_i$ A halfspace-description for the intersection of any number of closed convex cones $\mathcal{K}_i$ can be acquired by pruning normals; specifically, only the conically independent normals from the aggregate of all the halfspace-descriptions need be retained.

$\sum \mathcal{K}_i$ Generators for the sum of any number of closed convex cones $\mathcal{K}_i$ can be determined by retaining only the conically independent generators from the aggregate of all the vertex-descriptions.

Such conically independent sets are not necessarily unique or minimal.

2.11 When extreme means exposed

For any convex full-dimensional polyhedral set in $\mathbb{R}^n$, distinction between the terms *extreme* and *exposed* vanishes [347, §2.4] [120, §2.2] for faces of all dimensions except n; their meanings become equivalent as we saw in Figure **22** (discussed in §2.6.1.2). In other words, each and every face of any polyhedral set (except the set itself) can be exposed by a hyperplane, and *vice versa*; *e.g*, Figure **27**.

Lewis [262, §6] [231, §2.3.4] claims nonempty extreme proper subsets and the exposed subsets coincide for $\mathbb{S}^n_+$; *id est*, each and every face of the positive semidefinite cone, whose dimension is less than dimension of the cone, is exposed. A more general discussion of cones having this property can be found in [358]; *e.g*, Lorentz cone (178) [21, §II.A].

2.12 Convex polyhedra

Every polyhedron, such as the convex hull (86) of a bounded list X, can be expressed as the solution set of a finite system of linear equalities and inequalities, and *vice versa.* [120, §2.2]

2.12.0.0.1 Definition. *Convex polyhedra, halfspace-description.*
A convex polyhedron is the intersection of a finite number of halfspaces and hyperplanes;

$$\mathcal{P} = \{y \mid Ay \succeq b\,,\; Cy = d\} \subseteq \mathbb{R}^n \tag{285}$$

where coefficients A and C generally denote matrices. Each row of C is a vector normal to a hyperplane, while each row of A is a vector inward-normal to a hyperplane partially bounding a halfspace. △

By the *halfspaces theorem* in §2.4.1.1.1, a polyhedron thus described is a closed convex set possibly not full-dimensional; *e.g*, Figure **22**. Convex polyhedra[2.52] are finite-dimensional comprising all affine sets (§2.3.1, §2.1.4), polyhedral cones, line segments, rays, halfspaces, convex polygons, *solids* [239, def.104/6 p.343], polychora, *polytopes,*[2.53] *etcetera.*

It follows from definition (285) by exposure that each face of a convex polyhedron is a convex polyhedron.

Projection of any polyhedron on a subspace remains a polyhedron. More generally, image and inverse image of a convex polyhedron under any linear transformation remains a convex polyhedron; [27, §I.9] [325, thm.19.3] the foremost consequence being, invariance of polyhedral set closedness.

When b and d in (285) are $\mathbf{0}$, the resultant is a polyhedral cone. The set of all polyhedral cones is a subset of convex cones:

2.12.1 Polyhedral cone

From our study of cones, we see: the number of intersecting hyperplanes and halfspaces constituting a convex cone is possibly but not necessarily infinite. When the number is finite, the convex cone is termed *polyhedral.* That is the primary distinguishing feature between the set of all convex cones and polyhedra; all polyhedra, including polyhedral cones, are *finitely generated* [325, §19]. (Figure **55**) We distinguish polyhedral cones in the set of all convex cones for this reason, although all convex cones of dimension 2 or less are polyhedral.

[2.52] We consider only convex polyhedra throughout, but acknowledge the existence of concave polyhedra. [412, *Kepler-Poinsot Solid*]

[2.53] Some authors distinguish bounded polyhedra via the designation *polytope.* [120, §2.2]

2.12.1.0.1 Definition. *Polyhedral cone, halfspace-description.*[2.54] (*confer* (103))
A polyhedral cone is the intersection of a finite number of halfspaces and hyperplanes about the origin;

$$\begin{aligned}
\mathcal{K} &= \{y \mid Ay \succeq 0\,,\; Cy = \mathbf{0}\} \;\subseteq\; \mathbb{R}^n && \text{(a)}\\
&= \{y \mid Ay \succeq 0\,,\; Cy \succeq 0\,,\; Cy \preceq 0\} && \text{(b)}\\
&= \left\{ y \mid \begin{bmatrix} A \\ C \\ -C \end{bmatrix} y \succeq 0 \right\} && \text{(c)}
\end{aligned} \tag{286}$$

where coefficients A and C generally denote matrices of finite dimension. Each row of C is a vector normal to a hyperplane containing the origin, while each row of A is a vector inward-normal to a hyperplane containing the origin and partially bounding a halfspace. △

A polyhedral cone thus defined is closed, convex (§2.4.1.1), has only a finite number of generators (§2.8.1.2), and can be not full-dimensional. (Minkowski) Conversely, any finitely generated convex cone is polyhedral. (Weyl) [347, §2.8] [325, thm.19.1]

2.12.1.0.2 Exercise. *Unbounded convex polyhedra.*
Illustrate an unbounded polyhedron that is not a cone or its translation. ▼

From the definition it follows that any single hyperplane through the origin, or any halfspace partially bounded by a hyperplane through the origin is a polyhedral cone. The most familiar example of polyhedral cone is any quadrant (or orthant, §2.1.3) generated by Cartesian half-axes. Esoteric examples of polyhedral cone include the point at the origin, any line through the origin, any ray having the origin as base such as the nonnegative real line $\mathbb{R}_+$ in subspace $\mathbb{R}$, polyhedral flavor (proper) Lorentz cone (316), any subspace, and $\mathbb{R}^n$. More polyhedral cones are illustrated in Figure **53** and Figure **27**.

2.12.2 Vertices of convex polyhedra

By definition, a vertex (§2.6.1.0.1) always lies on the relative boundary of a convex polyhedron. [239, def.115/6 p.358] In Figure **22**, each vertex of the polyhedron is located at an intersection of three or more facets, and every edge belongs to precisely two facets [27, §VI.1 p.252]. In Figure **27**, the only vertex of that polyhedral cone lies at the origin.

The set of all polyhedral cones is clearly a subset of convex polyhedra and a subset of convex cones (Figure **55**). Not all convex polyhedra are bounded; evidently, neither can they all be described by the convex hull of a bounded set of points as defined in (86). Hence a universal vertex-description of polyhedra in terms of that same finite-length list X (76):

[2.54] Rockafellar [325, §19] proposes affine sets be handled via complementary pairs of affine inequalities; *e.g*, antisymmetry $Cy \succeq d$ and $Cy \preceq d$ which can present severe difficulty to some *interior-point methods* of numerical solution.

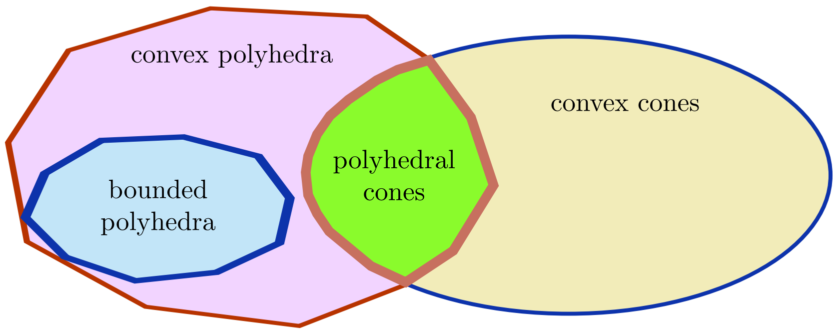


Figure 55: Polyhedral cones are finitely generated, unbounded, and convex.

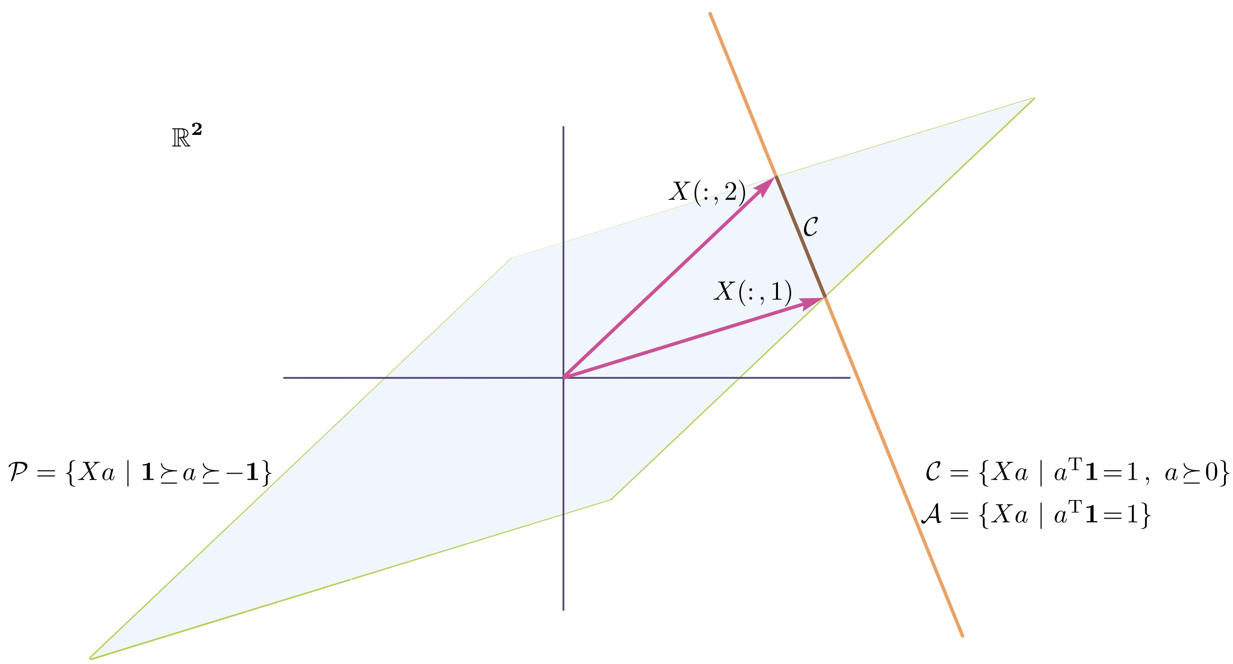


Figure 56: A polyhedron's generating list X does not necessarily constitute its vertices. Polyhedron $\mathcal{P}$ is a parallelogram, polyhedron $\mathcal{A}$ is a line, while polyhedron $\mathcal{C}$ is a line segment in $\mathcal{P}$. Were vector a instead unbounded, $\mathcal{P}$ would become the subspace $\mathcal{R}(X)$.

2.12.2.0.1 Definition. *Convex polyhedra, vertex-description.*
Denote upper u and lower ℓ real vector bounds and a truncated N-dimensional a-vector

$$a_{i:j} = \begin{bmatrix} a_i \\ \vdots \\ a_j \end{bmatrix} \tag{287}$$

By discriminating a suitable finite-length *generating list* (or set) arranged columnar in $X\in\mathbb{R}^{n\times N}$, then any particular polyhedron may be described

$$\mathcal{P} = \left\{Xa \mid a_{1:k}^{\mathrm T}\mathbf{1} = 1\,,\; u\succeq a_{m:N}\succeq \ell\,,\; \{1\ldots k\}\cup\{m\ldots N\} = \{1\ldots N\}\right\} \tag{288}$$

where $0\leq k\leq N$ and $1\leq m\leq N+1$. Setting $k=0$ removes the affine equality condition. Setting $m=N+1$ removes the inequality. △

Coefficient indices in (288) may or may not be overlapping. From (78), (86), (103), and (142), we summarize how the coefficient conditions may be applied;

$$\begin{array}{lcl} \text{affine set} & \longrightarrow & \left.\begin{array}{l} a_{1:k}^{\mathrm T}\mathbf{1} = 1 \\ a_{m:N}\succeq 0 \end{array}\right\} \longleftarrow \quad \text{convex hull } (m\leq k) \\ \text{polyhedral cone} & \longrightarrow & \\ \text{subspace} & \longrightarrow & \infty\succeq a\succeq -\infty \end{array} \tag{289}$$

It is always possible to describe a convex hull in a region of overlapping indices because, for $1\leq m\leq k\leq N$

$$\{a_{m:k} \mid a_{m:k}^{\mathrm T}\mathbf{1} = 1,\; a_{m:k}\succeq 0\} \subseteq \{a_{m:k} \mid a_{1:k}^{\mathrm T}\mathbf{1} = 1,\; a_{m:N}\succeq 0\} \tag{290}$$

Generating list members are not unique nor necessarily vertices of the corresponding polyhedron; *e.g*, Figure **56**. Indeed, for convex hull (86) (a special case of (288)), some subset of list members may reside in the polyhedron's relative interior. Conversely, convex hull of the vertices and extreme rays of a polyhedron is identical to the convex hull of any list generating that polyhedron; that is, *extremes theorem* 2.8.1.1.1.

2.12.2.1 Vertex-description of polyhedral cone

Given closed convex cone $\mathcal{K}$ in a subspace of $\mathbb{R}^n$ having any set of generators for it arranged in a matrix $X\in\mathbb{R}^{n\times N}$ as in (279), then that cone is described setting $m=1$ and $k=0$ in vertex-description (288):

$$\mathcal{K} \;=\; \operatorname{cone} X \;=\; \{Xa \mid a\succeq 0\} \;\subseteq\; \mathbb{R}^n \tag{103}$$

a conic hull of N generators.

2.12.2.2 Pointedness

(§2.7.2.1.2) [347, §2.10] Assuming all generators constituting the columns of $X\in\mathbb{R}^{n\times N}$ are nonzero, polyhedral cone $\mathcal{K}$ is pointed if and only if there is no nonzero $a\succeq 0$ that solves $Xa=\mathbf{0}$; *id est*, iff

$$\begin{array}{rl} \text{find} & a \\ \text{subject to} & Xa = \mathbf{0} \\ & \mathbf{1}^{\mathrm T}a = 1 \\ & a\succeq 0 \end{array} \tag{291}$$

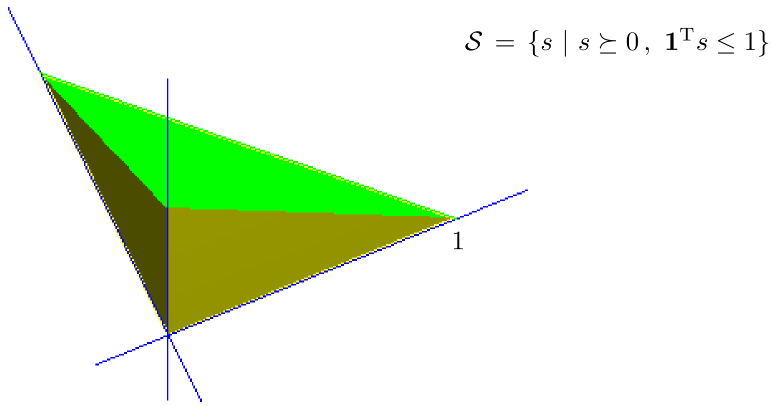


Figure 57: Unit simplex $\mathcal{S}$ in $\mathbb{R}^{\mathbf{3}}$ is a unique solid tetrahedron but not *regular*.

is infeasible or iff $\mathcal{N}(X) \cap \mathbb{R}^N_+ = \mathbf{0}$.[2.55] Otherwise, the cone will contain at least one line and there can be no vertex; *id est*, the cone cannot otherwise be pointed. Any subspace, Euclidean vector space $\mathbb{R}^n$, or any halfspace are examples of nonpointed polyhedral cone; hence, no vertex. This null-pointedness criterion $Xa = \mathbf{0}$ means that a pointed polyhedral cone is invariant to linear injective transformation.

Examples of pointed polyhedral cone $\mathcal{K}$ include: the origin, any $\mathbf{0}$-based ray in a subspace, any two-dimensional V-shaped cone in a subspace, any orthant in $\mathbb{R}^n$ or $\mathbb{R}^{m\times n}$; *e.g*, nonnegative real line $\mathbb{R}_+$ in vector space $\mathbb{R}$.

2.12.3 Unit simplex

A peculiar subset of the nonnegative orthant with halfspace-description

$$\mathcal{S} \triangleq \{s \mid s \succeq 0\,,\ \mathbf{1}^{\mathrm T} s \le 1\} \subseteq \mathbb{R}^n_+ \tag{292}$$

is a unique bounded convex full-dimensional polyhedron called *unit simplex* (Figure **57**) having $n+1$ facets, $n+1$ vertices, and dimension

$$\dim \mathcal{S} = n \tag{293}$$

The origin supplies one vertex while heads of the *standard basis* [218] [348] $\{e_i\,,\ i=1\ldots n\}$ in $\mathbb{R}^n$ constitute those remaining;[2.56] thus its vertex-description:

$$\begin{aligned} \mathcal{S} &= \operatorname{conv}\{\mathbf{0},\ \{e_i\,,\ i=1\ldots n\}\} \\ &= \left\{[\,\mathbf{0}\ \ e_1\ \ e_2\ \cdots\ e_n\,]\,a \mid a^{\mathrm T}\mathbf{1} = 1\,,\ a \succeq 0\right\} \end{aligned} \tag{294}$$

[2.55] If $\operatorname{rank} X = n$, then the dual cone $\mathcal{K}^*$ (§2.13.1) is pointed. (309)

[2.56] 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 $\mathbb{R}^{\mathbf{2}}$ it is a triangle and its relative interior, in $\mathbb{R}^{\mathbf{3}}$ it is the convex hull of a tetrahedron (Figure **57**), in $\mathbb{R}^{\mathbf{4}}$ it is the convex hull of a *pentatope* [412], and so on.

2.12.3.1 Simplex

The unit simplex comes from a class of general polyhedra called *simplex,* having vertex-description: [98] [325] [410] [120] given $n \geq k$

$$\operatorname{conv}\{x_\ell \in \mathbb{R}^n \mid \ell = 1 \ldots k+1,\ \dim \operatorname{aff}\{x_\ell\} = k\} \tag{295}$$

So defined, a simplex is a closed bounded convex set possibly not full-dimensional. Examples of simplices, by increasing affine dimension, are: a point, any line segment, any triangle and its relative interior, a general tetrahedron, any five-vertex polychoron, and so on.

2.12.3.1.1 Definition. *Simplicial cone.*
A proper polyhedral (§2.7.2.2.1) cone $\mathcal{K}$ in $\mathbb{R}^n$ is called *simplicial* iff $\mathcal{K}$ has exactly n extreme directions; [21, §II.A] equivalently, iff proper $\mathcal{K}$ has exactly n linearly independent generators contained in any given set of generators. △

- simplicial cone ⇒ proper polyhedral cone

There are an infinite variety of simplicial cones in $\mathbb{R}^n$; *e.g*, Figure **27**, Figure **58**, Figure **68**. Any orthant is simplicial, as is any rotation thereof.

2.12.4 Converting between descriptions

Conversion between halfspace- (285) (286) and vertex-description (86) (288) is nontrivial, in general, [16] [120, §2.2] but the conversion is easy for simplices. [63, §2.2.4] Nonetheless, we tacitly assume the two descriptions of polyhedra to be equivalent. [325, §19 thm.19.1] We explore conversions in §2.13.4, §2.13.9, and §2.13.11:

2.13 Dual cone & generalized inequality & biorthogonal expansion

These three concepts, dual cone, generalized inequality, and biorthogonal expansion, are inextricably melded; meaning, it is difficult to completely discuss one without mentioning the others. The dual cone is critical in tests for convergence by contemporary primal/dual methods for numerical solution of conic problems. [430] [294, §4.5] For unique minimum-distance projection on a closed convex cone $\mathcal{K}$, the negative dual cone $-\mathcal{K}^*$ plays the role that orthogonal complement plays for subspace projection.[2.57] (§E.9.2, Figure **181**) Indeed, $-\mathcal{K}^*$ is the algebraic complement in $\mathbb{R}^n$;

$$\mathcal{K} \boxplus -\mathcal{K}^* = \mathbb{R}^n \tag{2140}$$

where $\boxplus$ denotes unique orthogonal vector sum.

[2.57] Namely, projection on a subspace is ascertainable from projection on its orthogonal complement (Figure **180**).

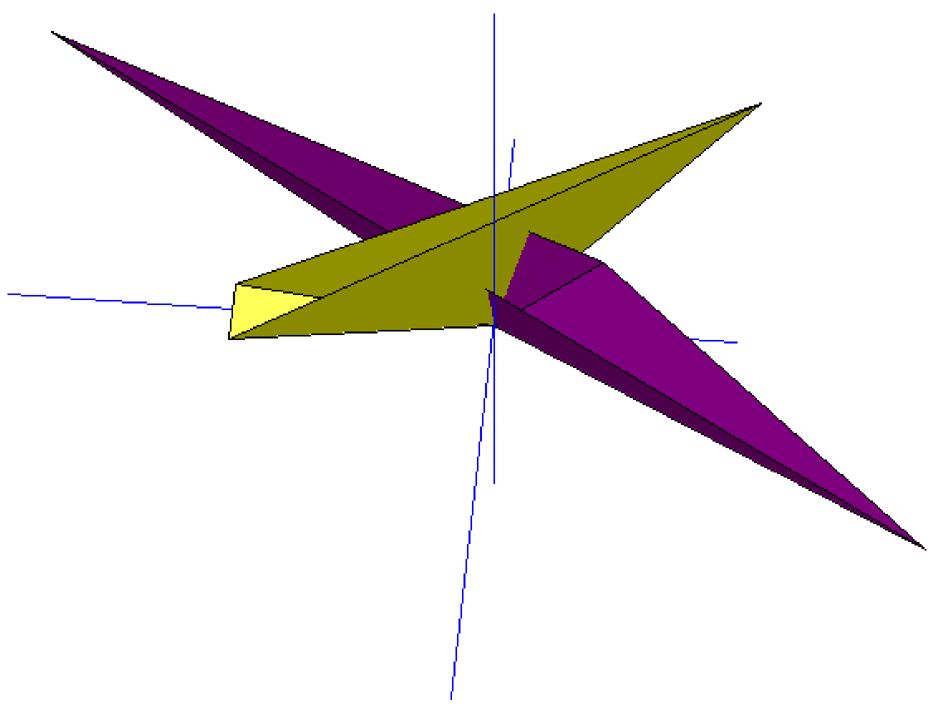
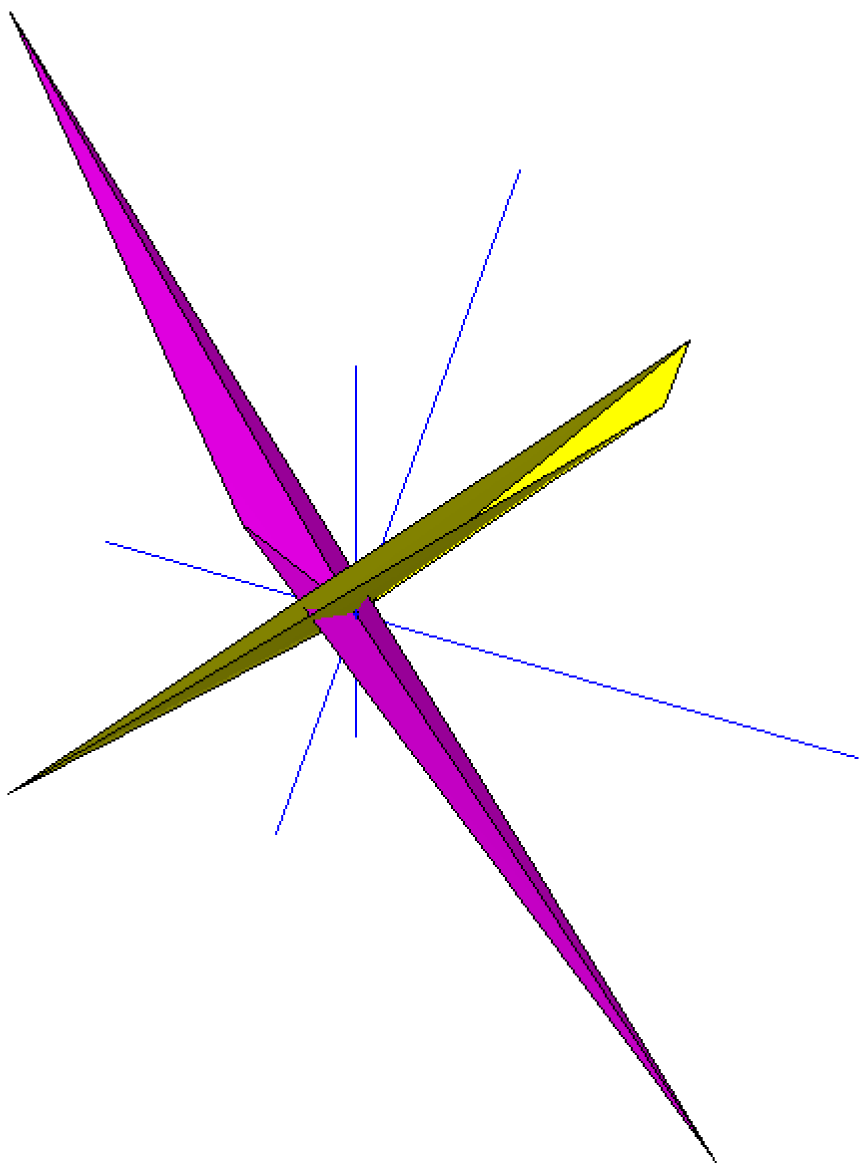


Figure 58: Two views of a simplicial cone and its dual in $\mathbb{R}^3$ (second view on next page). Semiinfinite boundary of each cone is truncated for illustration. Each cone has three facets (*confer* §2.13.11.0.3). (Cartesian axes are drawn for reference.)



One way to think of a pointed closed convex cone is as a new kind of coordinate system whose basis is generally nonorthogonal; a conic system, very much like the familiar Cartesian system whose analogous cone is the first quadrant (the nonnegative orthant). Generalized inequality $\succeq_{\mathcal{K}}$ is a formalized means to determine membership to any pointed closed convex cone $\mathcal{K}$ (§2.7.2.2) whereas *biorthogonal expansion* is, fundamentally, an expression of coordinates in a pointed conic system whose axes are linearly independent but not necessarily orthogonal. When cone $\mathcal{K}$ is the nonnegative orthant, then these three concepts come into alignment with the Cartesian prototype: biorthogonal expansion becomes orthogonal expansion, the dual cone becomes identical to the orthant, and generalized inequality obeys a total order entrywise.

2.13.1 Dual cone

For any set $\mathcal{K}$ (convex or not), the *dual cone* [116, §4.2]

$$\mathcal{K}^* \triangleq \{y \in \mathbb{R}^n \mid \langle y\,,\, x\rangle \geq 0 \text{ for all } x \in \mathcal{K}\} \tag{296}$$

is a unique cone[2.58] that is always closed and convex because it is an intersection of halfspaces (§2.4.1.1.1). Each halfspace has inward-normal x, belonging to $\mathcal{K}$, and boundary containing the origin; *e.g*, Figure **59**a.

When cone $\mathcal{K}$ is convex, there is a second and equivalent construction: Dual cone $\mathcal{K}^*$ is the union of each and every vector y inward-normal to a hyperplane supporting $\mathcal{K}$ (§2.4.2.6.1); *e.g*, Figure **59**b. When $\mathcal{K}$ is represented by a halfspace-description such as (286), for example, where

$$A \triangleq \begin{bmatrix} a_1^{\mathrm{T}} \\ \vdots \\ a_m^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{m\times n}\,, \qquad C \triangleq \begin{bmatrix} c_1^{\mathrm{T}} \\ \vdots \\ c_p^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{p\times n} \tag{297}$$

then the dual cone can be represented as the conic hull

$$\mathcal{K}^* = \operatorname{cone}\{a_1\,, \ldots,\, a_m\,,\, \pm c_1\,, \ldots,\, \pm c_p\} \tag{298}$$

a vertex-description, because each and every conic combination of normals from the halfspace-description of $\mathcal{K}$ yields another inward-normal to a hyperplane supporting $\mathcal{K}$.

$\mathcal{K}^*$ can also be constructed pointwise using projection theory from §E.9.2: for $P_{\mathcal{K}}x$ the Euclidean projection of point x on closed convex cone $\mathcal{K}$

$$-\mathcal{K}^* = \{x - P_{\mathcal{K}}x \mid x \in \mathbb{R}^n\} = \{x \in \mathbb{R}^n \mid P_{\mathcal{K}}x = \mathbf{0}\} \tag{2141}$$

2.13.1.0.1 Exercise. *Manual dual cone construction.*
Perhaps the most instructive graphical method of dual cone construction is cut-and-try. Find the dual of each polyhedral cone from Figure **60** by using dual cone equation (296).
▼

[2.58] The dual cone is the negative polar cone defined by many authors; $\mathcal{K}^* = -\mathcal{K}^\circ$. [215, §A.3.2] [325, §14] [42] [27] [347, §2.7]

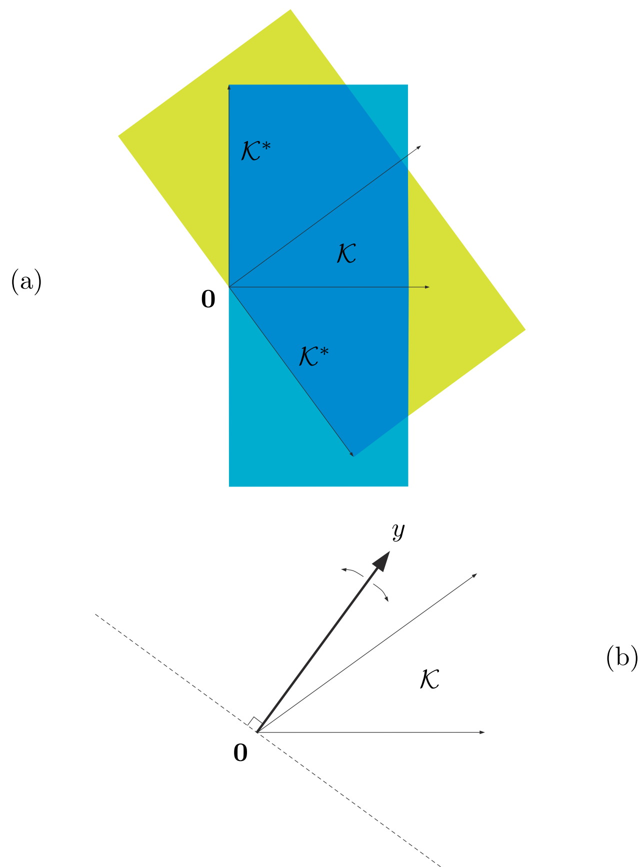
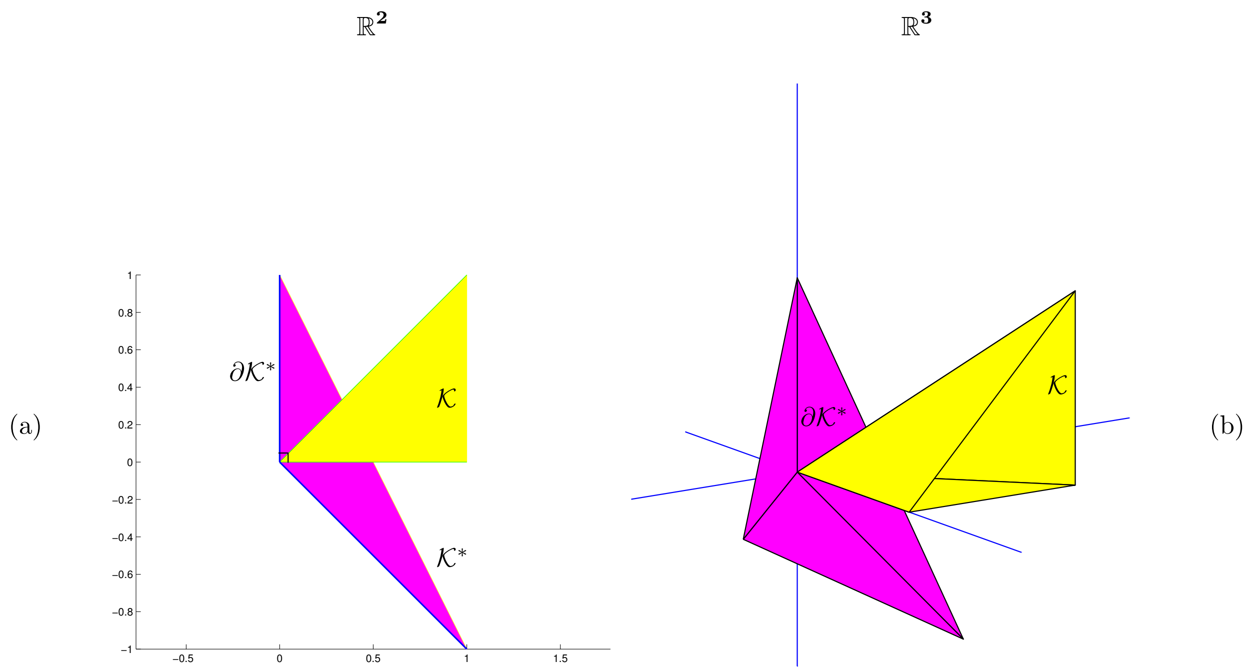


Figure 59: Two equivalent constructions of dual cone $\mathcal{K}^*$ in $\mathbb{R}^2$: **(a)** Showing construction by intersection of halfspaces about $\mathbf{0}$ (drawn truncated). Only those two halfspaces whose bounding hyperplanes have inward-normal corresponding to an extreme direction of this pointed closed convex cone $\mathcal{K} \subset \mathbb{R}^2$ need be drawn; by (368). **(b)** Suggesting construction by union of inward-normals y to each and every hyperplane $\underline{\partial\mathcal{H}_+}$ supporting $\mathcal{K}$. This interpretation is valid when $\mathcal{K}$ is convex because existence of a supporting hyperplane is then guaranteed (§2.4.2.6).



$$x \in \mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; y \in \mathcal{G}(\mathcal{K}^*) \tag{365}$$

Figure 60: Dual cone construction by right angle. Each extreme direction of a proper polyhedral cone is orthogonal to a facet of its dual cone, and *vice versa,* in any dimension. (§2.13.6.1) **(a)** This characteristic guides graphical construction of dual cone in two dimensions: It suggests finding dual-cone boundary ∂ by making right angles with extreme directions of polyhedral cone. The construction is then pruned so that each dual boundary vector does not exceed $\pi/2$ radians in angle with each and every vector from polyhedral cone. Were dual cone in $\mathbb{R}^2$ to narrow, Figure **61** would be reached in limit. **(b)** Same polyhedral cone and its dual continued into three dimensions. (*confer* Figure **68**)

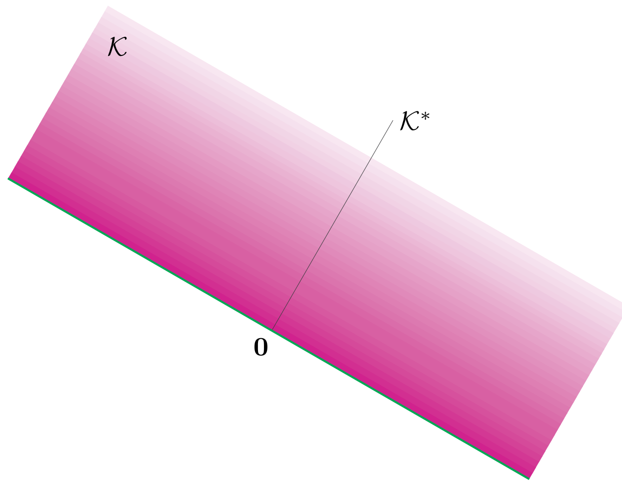


Figure 61: Polyhedral cone $\mathcal{K}$ is a halfspace about origin in $\mathbb{R}^{\mathbf{2}}$. Dual cone $\mathcal{K}^*$ is a ray base $\mathbf{0}$, hence not full-dimensional in $\mathbb{R}^{\mathbf{2}}$; so $\mathcal{K}$ cannot be pointed, hence has no extreme directions. (Both convex cones appear truncated.)

2.13.1.0.2 Exercise. *Dual cone definitions.*
What is $\{x \in \mathbb{R}^n \mid x^{\mathrm{T}} z \geq 0 \;\; \forall\, z \in \mathbb{R}^n\}$?
What is $\{x \in \mathbb{R}^n \mid x^{\mathrm{T}} z \geq 1 \;\; \forall\, z \in \mathbb{R}^n\}$?
What is $\{x \in \mathbb{R}^n \mid x^{\mathrm{T}} z \geq 1 \;\; \forall\, z \in \mathbb{R}^n_+\}$? ▼

As defined, dual cone $\mathcal{K}^*$ exists even when the affine hull of the original cone is a proper subspace; *id est*, even when the original cone is not full-dimensional.[2.59]

To further motivate our understanding of the dual cone, consider the ease with which convergence can be ascertained in the following optimization problem (301p):

2.13.1.0.3 Example. *Dual problem.* (*confer* §4.1)
Duality is a powerful and widely employed tool in applied mathematics for a number of reasons. First, the dual program is always convex even if the primal is not. Second, the number of variables in the dual is equal to the number of constraints in the primal which is often less than the number of variables in the primal program. Third, the maximum value achieved by the dual problem is often equal to the minimum of the primal. [317, §2.1.3] When not equal, the dual always provides a bound on the primal optimal objective. For convex problems, the dual variables provide necessary and sufficient optimality conditions:

[2.59] Rockafellar formulates dimension of $\mathcal{K}$ and $\mathcal{K}^*$. [325, §14.6.1] His monumental work *Convex Analysis* has not one figure or illustration. See [27, §II.16] for illustration of Rockafellar's *recession cone* [43].

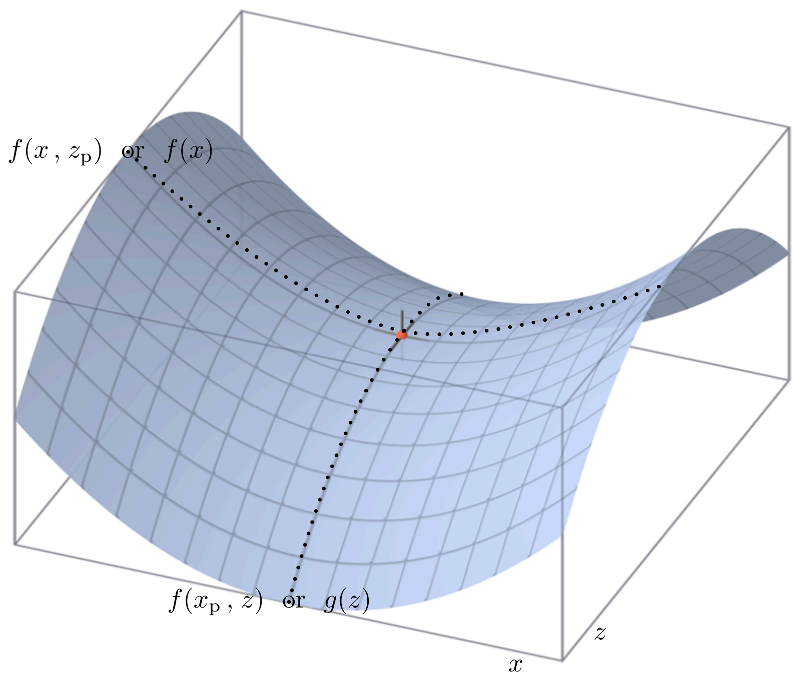


Figure 62: (Drawing by Lucas V. Barbosa.) This serves as mnemonic icon for primal and dual problems, although objective functions from conic problems (301p) (301d) are linear. When problems are *strong duals,* duality gap is 0 ; meaning, functions $f(x)$, $g(z)$ (dotted) *kiss* at saddle value as depicted at center. Otherwise, dual functions never meet $\big(f(x) > g(z)\big)$ by (299).

Essentially, *Lagrange duality theory* concerns representation of a given optimization problem as half of a *minimax problem*. [325, §36] [63, §5.4] Given any real function $f(x\,,z)$

$$\underset{x}{\text{minimize}}\,\underset{z}{\text{maximize}}\, f(x\,,z) \geq \underset{z}{\text{maximize}}\,\underset{x}{\text{minimize}}\, f(x\,,z) \tag{299}$$

always holds. When

$$\underset{x}{\text{minimize}}\,\underset{z}{\text{maximize}}\, f(x\,,z) = \underset{z}{\text{maximize}}\,\underset{x}{\text{minimize}}\, f(x\,,z) \tag{300}$$

we have *strong duality* and then a *saddle value* [166] exists. (Figure **62**) [322, p.3] Consider primal conic problem (p) (over cone $\mathcal{K}$) and its corresponding *dual problem* (d): [311, §3.3.1] [255, §2.1] [256] given vectors $\alpha\,,\beta$ and matrix constant C

$$\text{(p)}\quad \begin{array}{ll} \underset{x}{\text{minimize}} & \alpha^{\mathrm{T}}x \\ \text{subject to} & x \in \mathcal{K} \\ & Cx = \beta \end{array} \qquad \begin{array}{ll} \underset{y\,,\,z}{\text{maximize}} & \beta^{\mathrm{T}}z \\ \text{subject to} & y \in \mathcal{K}^* \\ & C^{\mathrm{T}}z + y = \alpha \end{array} \quad \text{(d)} \tag{301}$$

Observe: the dual problem is also conic, and its objective function value never exceeds that of the primal;

$$\begin{aligned} \alpha^{\mathrm{T}}x &\geq \beta^{\mathrm{T}}z \\ x^{\mathrm{T}}(C^{\mathrm{T}}z + y) &\geq (Cx)^{\mathrm{T}}z \\ x^{\mathrm{T}}y &\geq 0 \end{aligned} \tag{302}$$

which holds by definition (296). Under the sufficient condition that (301p) is a *convex problem*[2.60] satisfying *Slater's condition* (p.249), then equality

$$x^{\star\mathrm{T}}y^{\star} = 0 \tag{303}$$

is achieved; which is necessary and sufficient for optimality (§2.13.10.1.5); each problem (p) and (d) attains the same optimal value of its objective and each problem is called a *strong dual* to the other because the *duality gap* (optimal primal−dual objective difference) becomes 0. Then (p) and (d) are together equivalent to the minimax problem

$$\begin{array}{lll} \underset{x,y,z}{\text{minimize}} & \alpha^{\mathrm{T}}x - \beta^{\mathrm{T}}z & \\ \text{subject to} & x \in \mathcal{K}\,, & y \in \mathcal{K}^* \\ & Cx = \beta\,, & C^{\mathrm{T}}z + y = \alpha \end{array} \qquad \text{(p)}-\text{(d)} \tag{304}$$

whose optimal objective always has the saddle value 0 (regardless of the particular convex cone $\mathcal{K}$ and other problem parameters). [378, §3.2] Thus determination of convergence for either primal or dual problem is facilitated.

Were convex cone $\mathcal{K}$ polyhedral (§2.12.1), then problems (p) and (d) would be linear programs. Selfdual nonnegative orthant $\mathcal{K}$ yields the primal prototypical linear program and its dual. Were $\mathcal{K}$ a positive semidefinite cone, then problem (p) has the form of prototypical *semidefinite program* (687) with (d) its dual. It is sometimes possible to solve a primal problem by way of its dual; advantageous *when the dual problem is easier to solve than the primal problem, for example, because it can be solved analytically, or has some special structure that can be exploited.* [63, §5.5.5] (§4.2.3.1) □

[2.60] In this context, problems (p) and (d) are convex if $\mathcal{K}$ is a convex cone.

2.13.1.1 Key properties of dual cone

- For any cone, $(-\mathcal{K})^* = -\mathcal{K}^*$
- For any cones $\mathcal{K}_1$ and $\mathcal{K}_2$, $\quad \mathcal{K}_1 \subseteq \mathcal{K}_2 \;\Rightarrow\; \mathcal{K}_1^* \supseteq \mathcal{K}_2^*$ [347, §2.7]
- (Cartesian product) For closed convex cones $\mathcal{K}_1$ and $\mathcal{K}_2$, their Cartesian product $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2$ is a closed convex cone, and
$$\mathcal{K}^* = \mathcal{K}_1^* \times \mathcal{K}_2^* \tag{305}$$
where each dual is determined with respect to a cone's ambient space.
- (conjugation) [325, §14] [116, §4.5] [347, p.52] When $\mathcal{K}$ is any convex cone, dual of the dual cone equals closure of the original cone;
$$\mathcal{K}^{**} = \overline{\mathcal{K}} \tag{306}$$
is the intersection of all halfspaces about the origin that contain $\mathcal{K}$. Because $\mathcal{K}^{***} = \mathcal{K}^*$ always holds,
$$\mathcal{K}^* = (\overline{\mathcal{K}})^* \tag{307}$$
When convex cone $\mathcal{K}$ is closed, then dual of the dual cone is the original cone; $\mathcal{K}^{**} = \mathcal{K} \;\Leftrightarrow\; \mathcal{K}$ is a closed convex cone: [347, p.53, p.95]
$$\mathcal{K} = \{x \in \mathbb{R}^n \mid \langle y\,,\, x\rangle \geq 0 \;\; \forall\, y \in \mathcal{K}^*\} \tag{308}$$
- If any cone $\mathcal{K}$ is full-dimensional, then $\mathcal{K}^*$ is pointed;
$$\mathcal{K} \text{ full-dimensional } \Rightarrow \mathcal{K}^* \text{ pointed} \tag{309}$$
If the closure of any convex cone $\mathcal{K}$ is pointed, conversely, then $\mathcal{K}^*$ is full-dimensional;
$$\overline{\mathcal{K}} \text{ pointed } \Rightarrow \mathcal{K}^* \text{ full-dimensional} \tag{310}$$
Given that a cone $\mathcal{K} \subset \mathbb{R}^n$ is closed and convex, $\mathcal{K}$ is pointed if and only if $\mathcal{K}^* - \mathcal{K}^* = \mathbb{R}^n$; *id est*, iff $\mathcal{K}^*$ is full-dimensional. [56, §3.3 exer.20]
- (vector sum) [325, thm.3.8] For convex cones $\mathcal{K}_1$ and $\mathcal{K}_2$
$$\mathcal{K}_1 + \mathcal{K}_2 = \operatorname{conv}(\mathcal{K}_1 \cup \mathcal{K}_2) \tag{311}$$
is a convex cone.
- (dual vector-sum) [325, §16.4.2] [116, §4.6] For convex cones $\mathcal{K}_1$ and $\mathcal{K}_2$
$$\mathcal{K}_1^* \cap \mathcal{K}_2^* \;=\; (\mathcal{K}_1 + \mathcal{K}_2)^* \;=\; (\mathcal{K}_1 \cup \mathcal{K}_2)^* \tag{312}$$

- (closure of vector sum of duals)[2.61] For closed convex cones $\mathcal{K}_1$ and $\mathcal{K}_2$

$$(\mathcal{K}_1 \cap \mathcal{K}_2)^* = \overline{\mathcal{K}_1^* + \mathcal{K}_2^*} = \overline{\operatorname{conv}(\mathcal{K}_1^* \cup \mathcal{K}_2^*)} \tag{313}$$

[347, p.96] where operation closure becomes superfluous under the sufficient condition $\mathcal{K}_1 \cap \operatorname{int}\mathcal{K}_2 \neq \emptyset$ [56, §3.3 exer.16, §4.1 exer.7].

- (Krein-Rutman) Given closed convex cones $\mathcal{K}_1 \subseteq \mathbb{R}^m$ and $\mathcal{K}_2 \subseteq \mathbb{R}^n$ and any linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, then provided $\operatorname{int}\mathcal{K}_1 \cap A\mathcal{K}_2 \neq \emptyset$ [56, §3.3.13, *confer* §4.1 exer.9]

$$(A^{-1}\mathcal{K}_1 \cap \mathcal{K}_2)^* = A^{\mathrm{T}}\mathcal{K}_1^* + \mathcal{K}_2^* \tag{314}$$

where dual of cone $\mathcal{K}_1$ is with respect to its ambient space $\mathbb{R}^m$ and dual of cone $\mathcal{K}_2$ is with respect to $\mathbb{R}^n$, where $A^{-1}\mathcal{K}_1$ denotes inverse image (§2.1.9.0.1) of $\mathcal{K}_1$ under mapping A, and where A^{T} denotes adjoint operator. The particularly important case $\mathcal{K}_2 = \mathbb{R}^n$ is easy to show: for $A^{\mathrm{T}}A = I$

$$\begin{aligned}(A^{\mathrm{T}}\mathcal{K}_1)^* &\triangleq \{y \in \mathbb{R}^n \mid x^{\mathrm{T}}y \geq 0 \;\; \forall\, x \in A^{\mathrm{T}}\mathcal{K}_1\}\\ &= \{y \in \mathbb{R}^n \mid (A^{\mathrm{T}}z)^{\mathrm{T}}y \geq 0 \;\; \forall\, z \in \mathcal{K}_1\}\\ &= \{A^{\mathrm{T}}w \mid z^{\mathrm{T}}w \geq 0 \;\; \forall\, z \in \mathcal{K}_1\}\\ &= A^{\mathrm{T}}\mathcal{K}_1^*\end{aligned} \tag{315}$$

- $\mathcal{K}$ is proper if and only if $\mathcal{K}^*$ is proper.

- $\mathcal{K}$ is polyhedral if and only if $\mathcal{K}^*$ is polyhedral. [347, §2.8]

- $\mathcal{K}$ is simplicial if and only if $\mathcal{K}^*$ is simplicial. (§2.13.9.2) A simplicial cone and its dual are proper polyhedral cones (Figure **68**, Figure **58**), but not the converse.

- $\mathcal{K} \boxplus -\mathcal{K}^* = \mathbb{R}^n \Leftrightarrow \mathcal{K}$ is closed and convex. (2140)

- Any direction in a proper cone $\mathcal{K}$ is normal to a hyperplane separating $\mathcal{K}$ from $-\mathcal{K}^*$.

2.13.1.2 Examples of dual cone

When $\mathcal{K}$ is $\mathbb{R}^n$, $\mathcal{K}^*$ is the point at the origin, and *vice versa.*

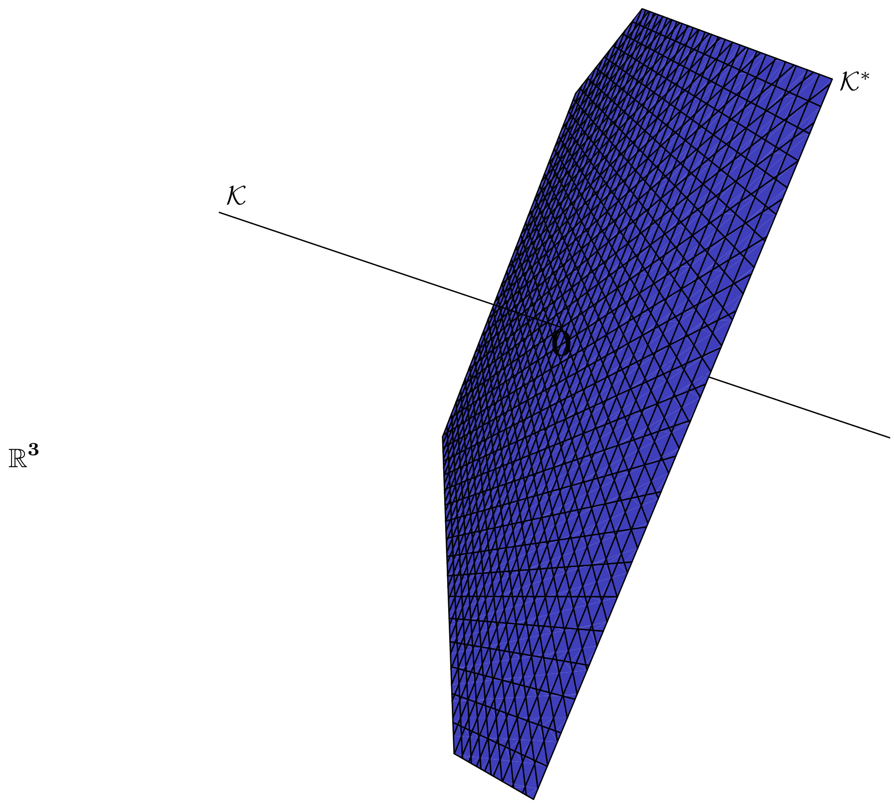
When $\mathcal{K}$ is a subspace, $\mathcal{K}^*$ is its orthogonal complement, and *vice versa.* (§E.9.2.1, Figure **63**)

When cone $\mathcal{K}$ is a halfspace in $\mathbb{R}^n$ with $n > 0$ (Figure **61** for example), the dual cone $\mathcal{K}^*$ is a ray (base $\mathbf{0}$) belonging to that halfspace but orthogonal to its bounding hyperplane (that contains the origin), and *vice versa.*

[2.61] These parallel analogous results for subspaces $\mathcal{R}_1, \mathcal{R}_2 \subseteq \mathbb{R}^n$; [116, §4.6]

$$\begin{aligned}(\mathcal{R}_1 + \mathcal{R}_2)^\perp &= \mathcal{R}_1^\perp \cap \mathcal{R}_2^\perp\\ (\mathcal{R}_1 \cap \mathcal{R}_2)^\perp &= \overline{\mathcal{R}_1^\perp + \mathcal{R}_2^\perp}\end{aligned}$$

$\mathcal{R}^{\perp\perp} = \mathcal{R}$ for any subspace $\mathcal{R}$.



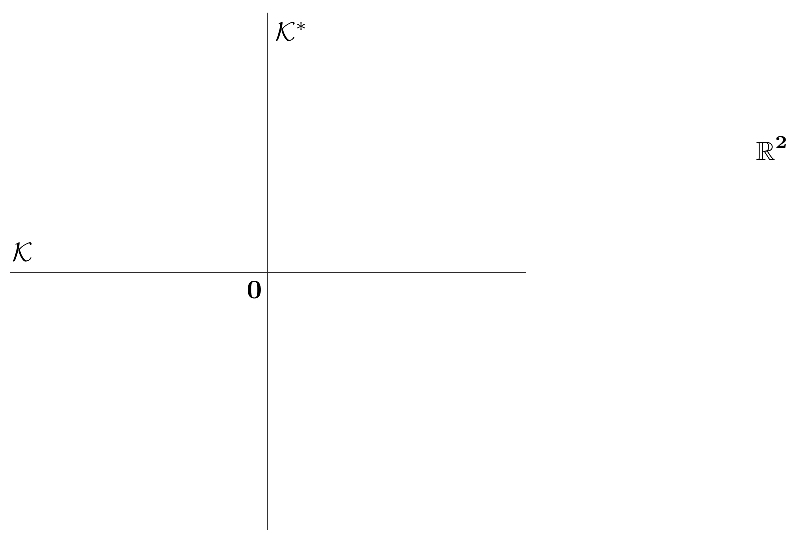


Figure 63: When convex cone $\mathcal{K}$ is any one Cartesian axis, its dual cone is the convex hull of all axes remaining; its orthogonal complement. In $\mathbb{R}^3$, dual cone $\mathcal{K}^*$ (drawn tiled and truncated) is a hyperplane through origin; its normal belongs to line $\mathcal{K}$. In $\mathbb{R}^2$, dual cone $\mathcal{K}^*$ is a line through origin while convex cone $\mathcal{K}$ is that line orthogonal.

When convex cone $\mathcal{K}$ is a closed halfplane in $\mathbb{R}^3$ (Figure **64**), it is neither pointed or full-dimensional; hence, the dual cone $\mathcal{K}^*$ can be neither full-dimensional or pointed.

When $\mathcal{K}$ is any particular orthant in $\mathbb{R}^n$, the dual cone is identical; *id est*, $\mathcal{K}=\mathcal{K}^*$.

When $\mathcal{K}$ is any quadrant in subspace $\mathbb{R}^2$, $\mathcal{K}^*$ is a *wedge*-shaped polyhedral cone in $\mathbb{R}^3$; *e.g*, for $\mathcal{K}$ equal to quadrant I, $\mathcal{K}^*=\begin{bmatrix}\mathbb{R}^2_+\\ \mathbb{R}\end{bmatrix}$.

When $\mathcal{K}$ is a polyhedral flavor Lorentz cone (*confer* (178))

$$\mathcal{K}_\ell = \left\{\begin{bmatrix} x\\ t\end{bmatrix} \in \mathbb{R}^n\times\mathbb{R} \mid \|x\|_\ell \leq t\right\}, \qquad \ell\in\{1,\,\infty\} \tag{316}$$

its dual is the proper cone [63, exmp.2.25]

$$\mathcal{K}_q \;=\; \mathcal{K}_\ell^* \;=\; \left\{\begin{bmatrix} x\\ t\end{bmatrix} \in \mathbb{R}^n\times\mathbb{R} \mid \|x\|_q \leq t\right\}, \qquad \ell\in\{1,\,2,\,\infty\} \tag{317}$$

where $\|x\|_\ell^* = \|x\|_q$ is that norm dual to $\|x\|_\ell$ determined via solution to $1/\ell + 1/q = 1$.[2.62] Figure **66** illustrates $\mathcal{K}=\mathcal{K}_1$ and $\mathcal{K}^*=\mathcal{K}_1^*=\mathcal{K}_\infty$ in $\mathbb{R}^2\times\mathbb{R}$.

2.13.2 Abstractions of *Farkas' lemma*

2.13.2.0.1 Corollary. *Generalized inequality and membership relation.* [215, §A.4.2] Let $\mathcal{K}$ be any closed convex cone and $\mathcal{K}^*$ its dual, and let x and y belong to a vector space $\mathbb{R}^n$. Then

$$y\in\mathcal{K}^* \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; x\in\mathcal{K} \tag{318}$$

which is, merely, a statement of fact by definition of dual cone (296). By closure we have conjugation: [325, thm.14.1]

$$x\in\mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; y\in\mathcal{K}^* \tag{319}$$

which may be regarded as a simple translation of *Farkas' lemma* [144] as in [325, §22] to the language of convex cones, and a generalization of the well-known Cartesian cone fact

$$x\succeq 0 \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; y\succeq 0 \tag{320}$$

for which implicitly $\mathcal{K}=\mathcal{K}^*=\mathbb{R}^n_+$ the nonnegative orthant.

Membership relation (319) is often written instead as *dual generalized inequalities,* when $\mathcal{K}$ and $\mathcal{K}^*$ are pointed closed convex cones,

$$x \underset{\mathcal{K}}{\succeq} 0 \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; y \underset{\mathcal{K}^*}{\succeq} 0 \tag{321}$$

meaning, coordinates for biorthogonal expansion of x (§2.13.7.1.2, §2.13.8) [384] must be nonnegative when x belongs to $\mathcal{K}$. Conjugating,

$$y \underset{\mathcal{K}^*}{\succeq} 0 \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; x \underset{\mathcal{K}}{\succeq} 0 \tag{322}$$

⋄

[2.62] Dual norm is not a conjugate or dual function.

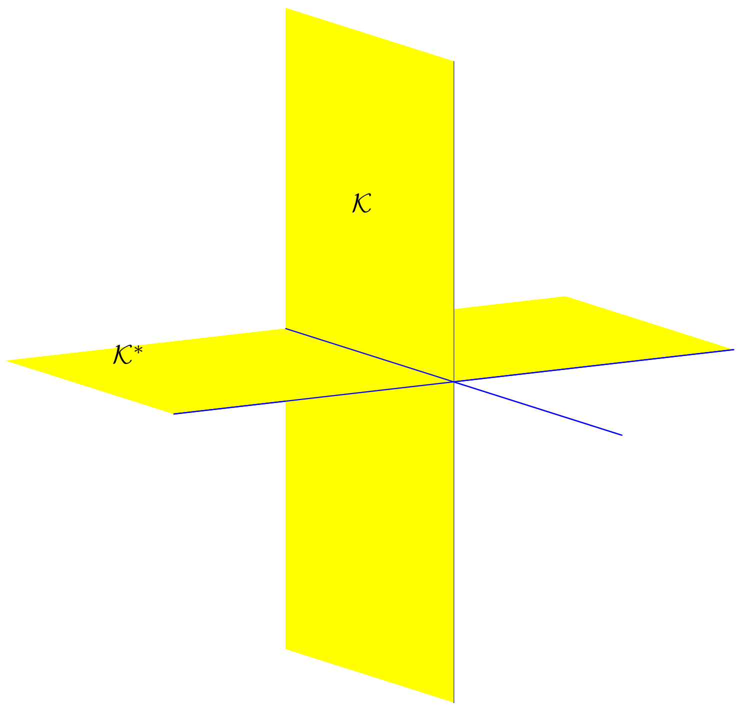


Figure 64: $\mathcal{K}$ and $\mathcal{K}^*$ are halfplanes in $\mathbb{R}^3$; *blades.* Both semiinfinite convex cones appear truncated. Each cone is like $\mathcal{K}$ from Figure **61** but embedded in a two-dimensional subspace of $\mathbb{R}^3$. (Cartesian coordinate axes drawn for reference.)

When pointed closed convex cone $\mathcal{K}$ is not polyhedral, coordinate axes for biorthogonal expansion asserted by the corollary are taken from extreme directions of $\mathcal{K}$; expansion is assured by *Carathéodory's theorem* (§E.6.4.1.1).

We presume, throughout, the obvious:

$$\begin{array}{c} x \in \mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; y \in \mathcal{K}^* \qquad (319) \\ \Leftrightarrow \\ x \in \mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \;\text{ for all }\; y \in \mathcal{K}^*,\; \|y\| = 1 \end{array} \tag{323}$$

2.13.2.0.2 Exercise. *Dual generalized inequalities.*
Test Corollary 2.13.2.0.1 and (323) graphically on the two-dimensional polyhedral cone and its dual in Figure **60**. ▼

(*confer* §2.7.2.2) When pointed closed convex cone $\mathcal{K}$ is implicit from context:

$$\begin{array}{l} x \succeq 0 \;\Leftrightarrow\; x \in \mathcal{K} \\ x \succ 0 \;\Leftrightarrow\; x \in \operatorname{rel\,int} \mathcal{K} \end{array} \tag{324}$$

Strict inequality $x \succ 0$ means coordinates for biorthogonal expansion of x must be positive when x belongs to $\operatorname{rel\,int} \mathcal{K}$. Strict membership relations are useful; *e.g*, for any proper cone[2.63] $\mathcal{K}$ and its dual $\mathcal{K}^*$

$$x \in \operatorname{int} \mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \;\text{ for all }\; y \in \mathcal{K}^*,\; y \neq \mathbf{0} \tag{325}$$

$$x \in \mathcal{K}\,,\; x \neq \mathbf{0} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \;\text{ for all }\; y \in \operatorname{int} \mathcal{K}^* \tag{326}$$

Conjugating, we get the dual relations:

$$y \in \operatorname{int} \mathcal{K}^* \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \;\text{ for all }\; x \in \mathcal{K}\,,\; x \neq \mathbf{0} \tag{327}$$

$$y \in \mathcal{K}^*,\; y \neq \mathbf{0} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \;\text{ for all }\; x \in \operatorname{int} \mathcal{K} \tag{328}$$

Boundary-membership relations for proper cones are also useful:

$$x \in \partial\mathcal{K} \;\Leftrightarrow\; \exists\; y \neq \mathbf{0} \;\ni\; \langle y\,,\,x\rangle = 0\,,\; y \in \mathcal{K}^*,\; x \in \mathcal{K} \tag{329}$$

$$y \in \partial\mathcal{K}^* \;\Leftrightarrow\; \exists\; x \neq \mathbf{0} \;\ni\; \langle y\,,\,x\rangle = 0\,,\; x \in \mathcal{K}\,,\; y \in \mathcal{K}^* \tag{330}$$

which are consistent; *e.g*, $x \in \partial\mathcal{K} \;\Leftrightarrow\; x \notin \operatorname{int}\mathcal{K}$ **and** $x \in \mathcal{K}$.

2.13.2.0.3 Example. *Linear inequality.* [354, §4] (*confer* §2.13.5.1.1)
Consider a given matrix A and closed convex cone $\mathcal{K}$. By membership relation we have

$$\begin{array}{rl} Ay \in \mathcal{K}^* & \Leftrightarrow\; x^{\mathrm{T}}A\,y \geq 0 \;\; \forall\, x \in \mathcal{K} \\ & \Leftrightarrow\; y^{\mathrm{T}}z \geq 0 \;\; \forall\, z \in \{A^{\mathrm{T}}x \mid x \in \mathcal{K}\} \\ & \Leftrightarrow\; y \in \{A^{\mathrm{T}}x \mid x \in \mathcal{K}\}^* \end{array} \tag{331}$$

This implies

$$\{y \mid Ay \in \mathcal{K}^*\} = \{A^{\mathrm{T}}x \mid x \in \mathcal{K}\}^* \tag{332}$$

[2.63] An open cone $\mathcal{K}$ is admitted to (325) and (328) by (19).

When $\mathcal{K}$ is the *selfdual* nonnegative orthant (§2.13.5.1), for example, then

$$\{y \mid Ay \succeq 0\} = \{A^{\mathrm{T}}x \mid x \succeq 0\}^* \tag{333}$$

and the dual relation

$$\{y \mid Ay \succeq 0\}^* = \{A^{\mathrm{T}}x \mid x \succeq 0\} \tag{334}$$

comes by a theorem of Weyl (p.128) that yields closedness for any vertex-description of a polyhedral cone. $\square$

2.13.2.1 Null certificate, Theorem of the alternative

If in particular $x_{\mathrm{p}} \notin \mathcal{K}$ a closed convex cone, then construction in Figure **59**b suggests there exists a supporting hyperplane (having inward-normal belonging to dual cone $\mathcal{K}^*$) separating x_{p} from $\mathcal{K}$; indeed, (319)

$$x_{\mathrm{p}} \notin \mathcal{K} \;\Leftrightarrow\; \exists\, y \in \mathcal{K}^* \;\ni\; \langle y\,,\, x_{\mathrm{p}}\rangle < 0 \tag{335}$$

Existence of any one such y is a certificate of null membership. From a different perspective,

$$\begin{gathered} x_{\mathrm{p}} \in \mathcal{K} \\ \text{or in the alternative} \\ \exists\, y \in \mathcal{K}^* \;\ni\; \langle y\,,\, x_{\mathrm{p}}\rangle < 0 \end{gathered} \tag{336}$$

By *alternative* is meant: these two systems are incompatible; one system is feasible while the other is not.

2.13.2.1.1 Example. *Theorem of the alternative for linear inequality.*
Myriad alternative systems of linear inequality can be explained in terms of pointed closed convex cones and their duals.

Beginning from the simplest Cartesian dual generalized inequalities (320) (with respect to nonnegative orthant $\mathbb{R}^m_+$),

$$y \succeq 0 \;\Leftrightarrow\; x^{\mathrm{T}}y \geq 0 \;\text{ for all }\; x \succeq 0 \tag{337}$$

Given $A \in \mathbb{R}^{n\times m}$, we make vector substitution $y \leftarrow A^{\mathrm{T}}y$

$$A^{\mathrm{T}}y \succeq 0 \;\Leftrightarrow\; x^{\mathrm{T}}A^{\mathrm{T}}y \geq 0 \;\text{ for all }\; x \succeq 0 \tag{338}$$

Introducing a new vector by calculating $b \triangleq Ax$ we get

$$A^{\mathrm{T}}y \succeq 0 \;\;\Leftrightarrow\;\; b^{\mathrm{T}}y \geq 0\,,\;\; b = Ax \;\text{ for all }\; x \succeq 0 \tag{339}$$

By complementing sense of the scalar inequality:

$$\begin{gathered} A^{\mathrm{T}}y \succeq 0 \\ \text{or in the alternative} \\ b^{\mathrm{T}}y < 0\,,\;\; \exists\, b = Ax\,,\;\; x \succeq 0 \end{gathered} \tag{340}$$

If one system has a solution, then the other does not; define a convex cone $\mathcal{K} = \{y \mid A^{\mathrm{T}} y \succeq 0\}$, then $y \in \mathcal{K}$ or in the alternative $y \notin \mathcal{K}$.

Scalar inequality $b^{\mathrm{T}} y < 0$ is movable to the other side of alternative (340), but that requires some explanation: From results in Example 2.13.2.0.3, the dual cone is $\mathcal{K}^* = \{Ax \mid x \succeq 0\}$. From (319) we have

$$\begin{array}{rcl} y \in \mathcal{K} & \Leftrightarrow & b^{\mathrm{T}} y \geq 0 \;\text{ for all }\; b \in \mathcal{K}^* \\ A^{\mathrm{T}} y \succeq 0 & \Leftrightarrow & b^{\mathrm{T}} y \geq 0 \;\text{ for all }\; b \in \{Ax \mid x \succeq 0\} \end{array} \tag{341}$$

Given some b vector and $y \in \mathcal{K}$, then $b^{\mathrm{T}} y < 0$ can only mean $b \notin \mathcal{K}^*$. An alternative system is therefore simply $b \in \mathcal{K}^*$: [215, p.59] (Farkas/Tucker)

$$\begin{array}{c} A^{\mathrm{T}} y \succeq 0\,, \quad b^{\mathrm{T}} y < 0 \\ \text{or in the alternative} \\ b = Ax\,, \quad x \succeq 0 \end{array} \tag{342}$$

Geometrically this means: either vector b belongs to convex cone $\mathcal{K}^*$ or it does not. When $b \notin \mathcal{K}^*$, then there is a vector $y \in \mathcal{K}$ normal to a hyperplane separating point b from cone $\mathcal{K}^*$.

For another example, from membership relation (318) with affine transformation of dual variable we may write: Given $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$

$$b - Ay \in \mathcal{K}^* \quad \Leftrightarrow \quad x^{\mathrm{T}}(b - Ay) \geq 0 \qquad \forall\, x \in \mathcal{K} \tag{343}$$

$$A^{\mathrm{T}} x = \mathbf{0}\,, \quad b - Ay \in \mathcal{K}^* \quad \Rightarrow \quad x^{\mathrm{T}} b \geq 0 \qquad \forall\, x \in \mathcal{K} \tag{344}$$

From membership relation (343), conversely, suppose we allow any $y \in \mathbb{R}^m$. Then because $-x^{\mathrm{T}} A\, y$ is unbounded below, $x^{\mathrm{T}}(b - Ay) \geq 0$ implies $A^{\mathrm{T}} x = \mathbf{0}$: for $y \in \mathbb{R}^m$

$$A^{\mathrm{T}} x = \mathbf{0}\,, \quad b - Ay \in \mathcal{K}^* \;\Leftarrow\; x^{\mathrm{T}}(b - Ay) \geq 0 \qquad \forall\, x \in \mathcal{K} \tag{345}$$

In toto,

$$b - Ay \in \mathcal{K}^* \quad \Leftrightarrow \quad x^{\mathrm{T}} b \geq 0\,, \quad A^{\mathrm{T}} x = \mathbf{0} \quad \forall\, x \in \mathcal{K} \tag{346}$$

Vector x belongs to cone $\mathcal{K}$ but is also constrained to lie in a subspace of $\mathbb{R}^n$ specified by an intersection of hyperplanes through the origin $\{x \in \mathbb{R}^n \mid A^{\mathrm{T}} x = \mathbf{0}\}$. From this, alternative systems of generalized inequality with respect to pointed closed convex cones $\mathcal{K}$ and $\mathcal{K}^*$

$$\begin{array}{c} Ay \underset{\mathcal{K}^*}{\preceq} b \\ \text{or in the alternative} \\ x^{\mathrm{T}} b < 0\,, \quad A^{\mathrm{T}} x = \mathbf{0}\,, \quad x \underset{\mathcal{K}}{\succeq} 0 \end{array} \tag{347}$$

derived from (346) simply by taking the complementary sense of the inequality in $x^{\mathrm{T}} b$. These two systems are alternatives; if one system has a solution, then the other does not.[2.64] [325, p.201]

By invoking a strict membership relation between proper cones (325), we can construct a more exotic alternative strengthened by demand for an interior point;

$$b - Ay \underset{\mathcal{K}^*}{\succ} 0 \quad\Leftrightarrow\quad x^{\mathrm{T}}b > 0\,,\quad A^{\mathrm{T}}x = \mathbf{0} \quad \forall\, x \underset{\mathcal{K}}{\succeq} 0\,,\quad x \neq \mathbf{0} \tag{348}$$

From this, alternative systems of generalized inequality [63, pp.50,54,262]

$$\begin{gathered} Ay \underset{\mathcal{K}^*}{\prec} b \\ \text{or in the alternative} \\ x^{\mathrm{T}}b \leq 0\,,\quad A^{\mathrm{T}}x = \mathbf{0}\,,\quad x \underset{\mathcal{K}}{\succeq} 0\,,\quad x \neq \mathbf{0} \end{gathered} \tag{349}$$

derived from (348) by taking complementary sense of the inequality in $x^{\mathrm{T}}b$.

And from this, alternative systems with respect to the nonnegative orthant attributed to Gordan in 1873: [175] [56, §2.2] substituting $A \leftarrow A^{\mathrm{T}}$ and setting $b = \mathbf{0}$

$$\begin{gathered} A^{\mathrm{T}}y \prec 0 \\ \text{or in the alternative} \\ Ax = \mathbf{0}\,,\quad x \succeq 0\,,\quad \|x\|_1 = 1 \end{gathered} \tag{350}$$

Ben-Israel collects related results from Farkas, Motzkin, Gordan, and Stiemke in *Motzkin transposition theorem*. [34] □

2.13.3 Optimality condition

(*confer* §2.13.10.1) The general first-order necessary and sufficient condition for optimality of solution $x^\star$ to a minimization problem ((301p) for example) with real differentiable convex objective function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is [324, §3]

$$\nabla f(x^\star)^{\mathrm{T}}(x - x^\star) \geq 0 \quad \forall\, x \in \mathcal{C}\,,\ \ x^\star \in \mathcal{C} \tag{351}$$

where $\mathcal{C}$ is a convex *feasible set*,[2.65] and where $\nabla f(x^\star)$ is the *gradient* (§3.6) of f with respect to x evaluated at $x^\star$. In words, negative gradient is normal to a hyperplane supporting the feasible set at a point of optimality. (Figure **71**)

[2.64] If solutions at $\pm\infty$ are disallowed, then the alternative systems become instead *mutually exclusive* with respect to nonpolyhedral cones. Simultaneous infeasibility of the two systems is not precluded by mutual exclusivity; called a *weak alternative*. Ye provides an example illustrating simultaneous infeasibility with respect to the positive semidefinite cone: $x \in \mathbb{S}^2$, $y \in \mathbb{R}$, $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and $b = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ where $x^{\mathrm{T}}b$ means $\langle x\,,\, b\rangle$. A better strategy than disallowing solutions at $\pm\infty$ is to demand an interior point as in (349) or Lemma 4.2.1.1.2. Then question of simultaneous infeasibility is moot.

[2.65] presumably nonempty set of all variable values satisfying all given problem constraints; the set of *feasible solutions.*

Direct solution to *variation inequality* (351), instead of the corresponding minimization, spawned from *calculus of variations.* [266, p.178] [143, p.37] One solution method solves an equivalent fixed point-of-projection problem

$$x = P_{\mathcal{C}}\big(x - \nabla f(x)\big) \tag{352}$$

that follows from a necessary and sufficient condition for projection on convex set $\mathcal{C}$ (Theorem E.9.1.0.2)

$$P\big(x^\star - \nabla f(x^\star)\big) \in \mathcal{C}\ , \qquad \langle x^\star - \nabla f(x^\star) - x^\star,\ x - x^\star\rangle \leq 0 \quad \forall\, x \in \mathcal{C} \tag{2125}$$

Proof of equivalence [388, *Complementarity problem*] is provided by Németh. Given minimum-distance projection problem

$$\begin{array}{cl} \underset{x}{\text{minimize}} & \frac{1}{2}\|x - y\|^2 \\ \text{subject to} & x \in \mathcal{C} \end{array} \tag{353}$$

on convex feasible set $\mathcal{C}$ for example, the equivalent fixed point problem

$$x \;=\; P_{\mathcal{C}}\big(x - \nabla f(x)\big) \;=\; P_{\mathcal{C}}(y) \tag{354}$$

is solved in one step.

In the unconstrained case ($\mathcal{C} = \mathbb{R}^n$), optimality condition (351) reduces to the classical rule (p.212)

$$\nabla f(x^\star) = \mathbf{0}\,, \quad x^\star \in \operatorname{dom} f \tag{355}$$

which can be inferred from the following application:

2.13.3.0.1 Example. *Optimality for equality constrained problem.*
Given a real differentiable convex function $f(x) : \mathbb{R}^n \to \mathbb{R}$ defined on domain $\mathbb{R}^n$, a fat full-rank matrix $C \in \mathbb{R}^{p\times n}$, and vector $d \in \mathbb{R}^p$, the convex optimization problem

$$\begin{array}{cl} \underset{x}{\text{minimize}} & f(x) \\ \text{subject to} & Cx = d \end{array} \tag{356}$$

is characterized by the well-known necessary and sufficient optimality condition [63, §4.2.3]

$$\nabla f(x^\star) + C^{\mathrm{T}}\nu = \mathbf{0} \tag{357}$$

where $\nu \in \mathbb{R}^p$ is the eminent *Lagrange multiplier.* [323] [266, p.188] [245] In other words, solution $x^\star$ is optimal if and only if $\nabla f(x^\star)$ belongs to $\mathcal{R}(C^{\mathrm{T}})$.

Via membership relation, we now derive condition (357) from the general first-order condition for optimality (351): For problem (356)

$$\mathcal{C} \triangleq \{x \in \mathbb{R}^n \mid Cx = d\} \;=\; \{Z\xi + x_{\mathrm{p}} \mid \xi \in \mathbb{R}^{n - \operatorname{rank} C}\} \tag{358}$$

is the feasible set where $Z \in \mathbb{R}^{n \times n - \operatorname{rank} C}$ holds basis $\mathcal{N}(C)$ columnar, and x_{p} is any particular solution to $Cx = d$. Since $x^\star \in \mathcal{C}$, we arbitrarily choose $x_{\mathrm{p}} = x^\star$ which yields an equivalent optimality condition

$$\nabla f(x^\star)^{\mathrm{T}} Z\xi \geq 0 \quad \forall\, \xi \in \mathbb{R}^{n - \operatorname{rank} C} \tag{359}$$

when substituted into (351). But this is simply half of a membership relation where the cone dual to $\mathbb{R}^{n-\operatorname{rank} C}$ is the origin in $\mathbb{R}^{n-\operatorname{rank} C}$. We must therefore have

$$Z^{\mathrm{T}}\nabla f(x^\star) = \mathbf{0} \;\Leftrightarrow\; \nabla f(x^\star)^{\mathrm{T}} Z\xi \geq 0 \quad \forall\, \xi \in \mathbb{R}^{n-\operatorname{rank} C} \tag{360}$$

meaning, $\nabla f(x^\star)$ must be orthogonal to $\mathcal{N}(C)$. These conditions

$$Z^{\mathrm{T}}\nabla f(x^\star) = \mathbf{0}\,, \qquad Cx^\star = d \tag{361}$$

are necessary and sufficient for optimality. □

2.13.4 Discretization of membership relation

2.13.4.1 Dual halfspace-description

Halfspace-description of dual cone $\mathcal{K}^*$ is equally simple as vertex-description

$$\mathcal{K} \;=\; \operatorname{cone}(X) \;=\; \{Xa \mid a \succeq 0\} \;\subseteq\; \mathbb{R}^n \qquad (103)$$

for corresponding closed convex cone $\mathcal{K}$: By definition (296), for $X \in \mathbb{R}^{n\times N}$ as in (279), (*confer* (286))

$$\begin{aligned}
\mathcal{K}^* \;&=\; \left\{y \in \mathbb{R}^n \mid z^{\mathrm{T}} y \geq 0 \text{ for all } z \in \mathcal{K}\right\}\\
&=\; \left\{y \in \mathbb{R}^n \mid z^{\mathrm{T}} y \geq 0 \text{ for all } z = Xa\,,\ a \succeq 0\right\}\\
&=\; \left\{y \in \mathbb{R}^n \mid a^{\mathrm{T}} X^{\mathrm{T}} y \geq 0\,,\ a \succeq 0\right\}\\
&=\; \left\{y \in \mathbb{R}^n \mid X^{\mathrm{T}} y \succeq 0\right\}
\end{aligned} \tag{362}$$

that follows from the *generalized inequality and membership corollary* (320). The semi-infinity of tests specified by all $z \in \mathcal{K}$ has been reduced to a set of generators for $\mathcal{K}$ constituting the columns of X; *id est*, the test has been discretized.

Whenever cone $\mathcal{K}$ is known to be closed and convex, the conjugate statement must also hold; *id est*, given any set of generators for dual cone $\mathcal{K}^*$ arranged columnar in Y, then the consequent vertex-description of dual cone connotes a halfspace-description for $\mathcal{K}$: [347, §2.8]

$$\mathcal{K}^* = \{Ya \mid a \succeq 0\} \;\Leftrightarrow\; \mathcal{K}^{**} = \mathcal{K} = \left\{z \mid Y^{\mathrm{T}} z \succeq 0\right\} \tag{363}$$

2.13.4.2 First dual-cone formula

From these two results (362) and (363) we deduce a general principle:

- From any [*sic*] given vertex-description (103) of closed convex cone $\mathcal{K}$, a halfspace-description of the dual cone $\mathcal{K}^*$ is immediate by matrix transposition (362); conversely, from any given halfspace-description (286) of $\mathcal{K}$, a dual vertex-description is immediate (363). [325, p.122]

Various other converses are just a little trickier. (§2.13.9, §2.13.11)

We deduce further: For any polyhedral cone $\mathcal{K}$, the dual cone $\mathcal{K}^*$ is also polyhedral and $\mathcal{K}^{**} = \mathcal{K}$. [347, p.56]

The *generalized inequality and membership corollary* is discretized in the following theorem inspired by (362) and (363):

2.13.4.2.1 Theorem. *Discretized membership.* (*confer* §2.13.2.0.1)[2.66]
Given any set of generators, (§2.8.1.2) denoted by $\mathcal{G}(\mathcal{K})$ for closed convex cone $\mathcal{K} \subseteq \mathbb{R}^n$, and any set of generators denoted $\mathcal{G}(\mathcal{K}^*)$ for its dual such that

$$\mathcal{K} = \operatorname{cone} \mathcal{G}(\mathcal{K}) \;, \qquad \mathcal{K}^* = \operatorname{cone} \mathcal{G}(\mathcal{K}^*) \tag{364}$$

then discretization of the *generalized inequality and membership corollary* (p.144) is necessary and sufficient for certifying cone membership: for x and y in vector space $\mathbb{R}^n$

$$x \in \mathcal{K} \;\Leftrightarrow\; \langle \gamma^*, x\rangle \geq 0 \;\text{ for all }\; \gamma^* \in \mathcal{G}(\mathcal{K}^*) \tag{365}$$

$$y \in \mathcal{K}^* \;\Leftrightarrow\; \langle \gamma \,, y\rangle \geq 0 \;\text{ for all }\; \gamma \in \mathcal{G}(\mathcal{K}) \tag{366}$$

⋄

Proof. $\mathcal{K}^* = \{\mathcal{G}(\mathcal{K}^*)a \mid a \succeq 0\}$. $y \in \mathcal{K}^* \Leftrightarrow y = \mathcal{G}(\mathcal{K}^*)a$ for some $a \succeq 0$. $x \in \mathcal{K} \Leftrightarrow \langle y \,, x\rangle \geq 0 \;\forall\, y \in \mathcal{K}^* \Leftrightarrow \langle \mathcal{G}(\mathcal{K}^*)a \,, x\rangle \geq 0 \;\forall\, a \succeq 0$ (319). $a \triangleq \sum_i \alpha_i e_i$ where e_i is the i^{th} member of a standard basis of possibly infinite cardinality. $\langle \mathcal{G}(\mathcal{K}^*)a \,, x\rangle \geq 0 \;\forall\, a \succeq 0 \Leftrightarrow \sum_i \alpha_i \langle \mathcal{G}(\mathcal{K}^*)e_i \,, x\rangle \geq 0 \;\forall\, \alpha_i \geq 0 \Leftrightarrow \langle \mathcal{G}(\mathcal{K}^*)e_i \,, x\rangle \geq 0 \;\forall\, i$. Conjugate relation (366) is similarly derived. ◆

2.13.4.2.2 Exercise. *Discretized dual generalized inequalities.*
Test Theorem 2.13.4.2.1 on Figure **60**a using extreme directions there as generators.

▼

From the *discretized membership theorem* we may further deduce a more surgical description of closed convex cone that prescribes only a finite number of halfspaces for its construction when polyhedral: (Figure **59**a)

$$\mathcal{K} = \{x \in \mathbb{R}^n \mid \langle \gamma^*, x\rangle \geq 0 \;\text{ for all }\; \gamma^* \in \mathcal{G}(\mathcal{K}^*)\} \tag{367}$$

$$\mathcal{K}^* = \{y \in \mathbb{R}^n \mid \langle \gamma \,, y\rangle \geq 0 \;\text{ for all }\; \gamma \in \mathcal{G}(\mathcal{K})\} \tag{368}$$

2.13.4.2.3 Exercise. *Partial order induced by orthant.*
When comparison is with respect to the nonnegative orthant $\mathcal{K} = \mathbb{R}^n_+$, then from the *discretized membership theorem* it directly follows:

$$x \preceq z \;\Leftrightarrow\; x_i \leq z_i \;\;\forall\, i \tag{369}$$

Generate simple counterexamples demonstrating that this equivalence with entrywise inequality holds only when the underlying cone inducing partial order is the nonnegative orthant; *e.g*, explain Figure **65**. ▼

[2.66] Stated in [22, §1] without proof for pointed closed convex case.

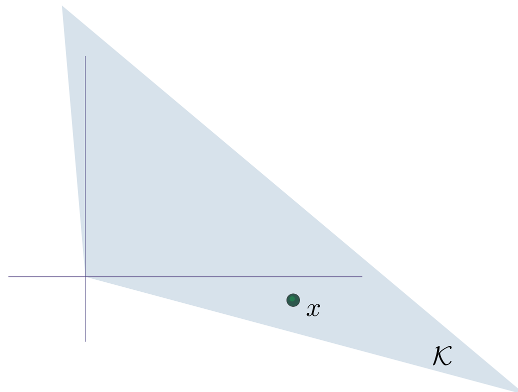


Figure 65: $x \succeq 0$ with respect to $\mathcal{K}$ but not with respect to nonnegative orthant $\mathbb{R}^{\mathbf{2}}_+$ (pointed convex cone $\mathcal{K}$ drawn truncated).

2.13.4.2.4 Example. *Boundary membership to proper polyhedral cone.*
For a polyhedral cone, test (329) of boundary membership can be formulated as a linear program. Say proper polyhedral cone $\mathcal{K}$ is specified completely by generators that are arranged columnar in

$$X = [\,\Gamma_1 \ \cdots \ \Gamma_N\,] \in \mathbb{R}^{n\times N} \tag{279}$$

id est, $\mathcal{K} = \{X a \mid a \succeq 0\}$. Then membership relation

$$c \in \partial\mathcal{K} \;\Leftrightarrow\; \exists\, y \neq \mathbf{0} \ni \langle y\,,c\rangle = 0\,,\ y \in \mathcal{K}^*,\ c \in \mathcal{K} \tag{329}$$

may be expressed[2.67]

$$\begin{array}{cl} \underset{a\,,\,y}{\text{find}} & y \neq \mathbf{0} \\ \text{subject to} & c^{\mathrm{T}} y = 0 \\ & X^{\mathrm{T}} y \succeq 0 \\ & X a = c \\ & a \succeq 0 \end{array} \tag{370}$$

This linear feasibility problem has a solution iff $c \in \partial\mathcal{K}$. $\square$

2.13.4.3 smallest face of pointed closed convex cone

Given nonempty convex subset $\mathcal{C}$ of a convex set $\mathcal{K}$, the *smallest face* of $\mathcal{K}$ containing $\mathcal{C}$ is equivalent to intersection of all faces of $\mathcal{K}$ that contain $\mathcal{C}$. [325, p.164] By (308), membership relation (329) means that each and every point on boundary $\partial\mathcal{K}$ of proper cone $\mathcal{K}$ belongs to a hyperplane supporting $\mathcal{K}$ whose normal y belongs to dual cone $\mathcal{K}^*$. It follows that the smallest face $\mathcal{F}$, containing $\mathcal{C} \subset \partial\mathcal{K} \subset \mathbb{R}^n$ on boundary of proper cone $\mathcal{K}$, is the intersection of all hyperplanes containing $\mathcal{C}$ whose normals are in $\mathcal{K}^*$;

$$\mathcal{F}(\mathcal{K} \supset \mathcal{C}) = \{x \in \mathcal{K} \mid x \perp \mathcal{K}^* \cap \mathcal{C}^\perp\} \tag{371}$$

[2.67] Nonzero $y \in \mathbb{R}^n$ may be realized in many ways; *e.g*, via constraint $\mathbf{1}^{\mathrm{T}} y = 1$ if $c \neq \mathbf{1}$.

where

$$\mathcal{C}^{\perp} \triangleq \{y\in\mathbb{R}^n \mid \langle z\,,\,y\rangle=0 \;\; \forall\, z\in\mathcal{C}\} \tag{372}$$

When $\mathcal{C}\cap\operatorname{int}\mathcal{K}\neq\emptyset$ then $\mathcal{F}(\mathcal{K}\supset\mathcal{C})=\mathcal{K}$.

2.13.4.3.1 Example. *Finding smallest face of cone.*
Suppose polyhedral cone $\mathcal{K}$ is completely specified by generators arranged columnar in

$$X = [\,\Gamma_1 \;\cdots\; \Gamma_N\,] \in \mathbb{R}^{n\times N} \tag{279}$$

To find its smallest face $\mathcal{F}(\mathcal{K}\ni c)$ containing a given point $c\in\mathcal{K}$, by the *discretized membership theorem* 2.13.4.2.1, it is necessary and sufficient to find generators for the smallest face. We may do so one generator at a time:[2.68] Consider generator Γ_i. If there exists a vector $z\in\mathcal{K}^*$ orthogonal to c but not to Γ_i, then Γ_i cannot belong to the smallest face of $\mathcal{K}$ containing c. Such a vector z can be realized by a linear feasibility problem:

$$\begin{array}{rl} \text{find} & z\in\mathbb{R}^n \\ \text{subject to} & c^{\mathrm{T}}z=0 \\ & X^{\mathrm{T}}z\succeq 0 \\ & \Gamma_i^{\mathrm{T}}z=1 \end{array} \tag{373}$$

If there exists a solution z for which $\Gamma_i^{\mathrm{T}}z=1$, then

$$\Gamma_i \;\not\perp\; \mathcal{K}^*\cap c^{\perp} = \{z\in\mathbb{R}^n \mid X^{\mathrm{T}}z\succeq 0\,,\; c^{\mathrm{T}}z=0\} \tag{374}$$

so $\Gamma_i\notin\mathcal{F}(\mathcal{K}\ni c)$; solution z is a certificate of null membership. If this problem is infeasible for generator $\Gamma_i\in\mathcal{K}$, conversely, then $\Gamma_i\in\mathcal{F}(\mathcal{K}\ni c)$ by (371) and (362) because $\Gamma_i\perp\mathcal{K}^*\cap c^{\perp}$; in that case, Γ_i is a generator of $\mathcal{F}(\mathcal{K}\ni c)$.

Since the constant in constraint $\Gamma_i^{\mathrm{T}}z=1$ is arbitrary positively, then by *theorem of the alternative* there is correspondence between (373) and (347) admitting the alternative linear problem: for a given point $c\in\mathcal{K}$

$$\begin{array}{rl} \underset{a\in\mathbb{R}^N,\; \mu\in\mathbb{R}}{\text{find}} & a\,,\,\mu \\ \text{subject to} & \mu c-\Gamma_i = X a \\ & a\succeq 0 \end{array} \tag{375}$$

Now if this problem is feasible (bounded) for generator $\Gamma_i\in\mathcal{K}$, then (373) is infeasible and $\Gamma_i\in\mathcal{F}(\mathcal{K}\ni c)$ is a generator of the smallest face that contains c. □

2.13.4.3.2 Exercise. *Finding smallest face of pointed closed convex cone.*
Show that formula (371) and algorithms (373) and (375) apply more broadly; *id est*, a full-dimensional cone $\mathcal{K}$ is an unnecessary condition.[2.69] ▼

2.13.4.3.3 Exercise. *Smallest face of positive semidefinite cone.*
Derive (221) from (371). ▼

[2.68] When finding a smallest face, generators of $\mathcal{K}$ in matrix X may not be diminished in number (by discarding columns) until all generators of the smallest face have been found.

[2.69] Hint: A hyperplane, with normal in $\mathcal{K}^*$, containing cone $\mathcal{K}$ is admissible.

2.13.5 Dual PSD cone and generalized inequality

The *dual positive semidefinite cone* $\mathcal{K}^*$ is confined to $\mathbb{S}^M$ by convention;

$$\mathbb{S}_+^{M^*} \triangleq \{Y \in \mathbb{S}^M \mid \langle Y , X\rangle \geq 0 \text{ for all } X \in \mathbb{S}_+^M\} = \mathbb{S}_+^M \tag{376}$$

The positive semidefinite cone is selfdual in the ambient space of symmetric matrices [63, exmp.2.24] [40] [211, §II]; $\mathcal{K} = \mathcal{K}^*$.

Dual generalized inequalities with respect to the positive semidefinite cone in the ambient space of symmetric matrices can therefore be simply stated: (Fejér)

$$X \succeq 0 \;\Leftrightarrow\; \operatorname{tr}(Y^{\mathrm{T}}X) \geq 0 \text{ for all } Y \succeq 0 \tag{377}$$

Membership to this cone can be determined in the isometrically isomorphic Euclidean space $\mathbb{R}^{M^2}$ via (38). (§2.2.1) By the two interpretations in §2.13.1, positive semidefinite matrix Y can be interpreted as inward-normal to a hyperplane supporting the positive semidefinite cone.

The fundamental statement of positive semidefiniteness, $y^{\mathrm{T}}Xy \geq 0 \;\forall y$ (§A.3.0.0.1), evokes a particular instance of these dual generalized inequalities (377):

$$X \succeq 0 \;\Leftrightarrow\; \langle yy^{\mathrm{T}} , X\rangle \geq 0 \;\; \forall\, yy^{\mathrm{T}}(\succeq 0) \tag{1531}$$

Discretization (§2.13.4.2.1) allows replacement of positive semidefinite matrices Y with this minimal set of generators comprising the extreme directions of the positive semidefinite cone (§2.9.2.7).

2.13.5.1 selfdual cones

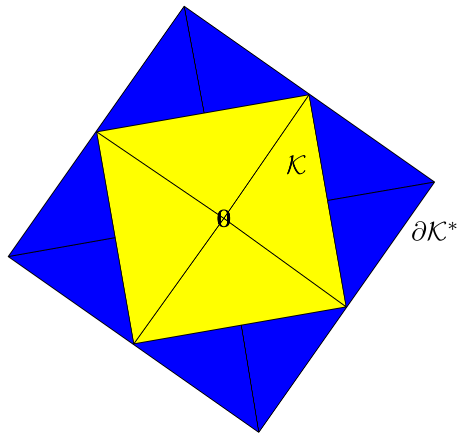
From (131) (a consequence of the *halfspaces theorem,* §2.4.1.1.1), where the only finite value of the support function for a convex cone is 0 [215, §C.2.3.1], or from discretized definition (368) of the dual cone we get a rather self evident characterization of selfdual cones:

$$\mathcal{K} = \mathcal{K}^* \quad\Leftrightarrow\quad \mathcal{K} = \bigcap_{\gamma \in \mathcal{G}(\mathcal{K})} \{y \mid \gamma^{\mathrm{T}} y \geq 0\} \tag{378}$$

In words: Cone $\mathcal{K}$ is *selfdual* iff its own extreme directions are inward-normals to a (minimal) set of hyperplanes bounding halfspaces whose intersection constructs it. This means each extreme direction of $\mathcal{K}$ is normal to a hyperplane exposing one of its own faces; a necessary but insufficient condition for selfdualness (Figure **66**, for example).

Selfdual cones are necessarily full-dimensional. [31, §I] Their most prominent representatives are the orthants (Cartesian cones), the positive semidefinite cone $\mathbb{S}_+^M$ in the ambient space of symmetric matrices (376), and Lorentz cone (178) [21, §II.A] [63, exmp.2.25]. In three dimensions, a plane containing the axis of revolution of a selfdual cone (and the origin) will produce a *slice* whose boundary makes a right angle.

(a)



(b)

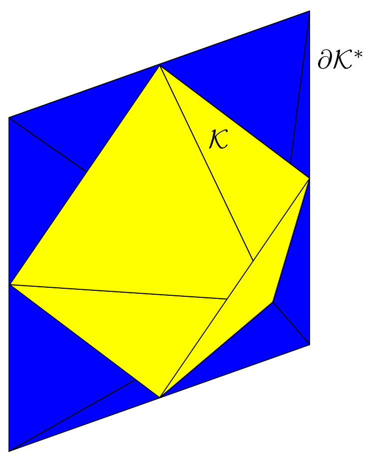


Figure 66: Two (truncated) views of a polyhedral cone $\mathcal{K}$ and its dual in $\mathbb{R}^3$. Each of four extreme directions from $\mathcal{K}$ belongs to a face of dual cone $\mathcal{K}^*$. Cone $\mathcal{K}$, shrouded by its dual, is symmetrical about its axis of revolution. Each pair of diametrically opposed extreme directions from $\mathcal{K}$ makes a right angle. An orthant (or any rotation thereof; a simplicial cone) is not the only selfdual polyhedral cone in three or more dimensions; [21, §2.A.21] *e.g*, consider an *equilateral* having five extreme directions. In fact, every selfdual polyhedral cone in $\mathbb{R}^3$ has an odd number of extreme directions. [23, thm.3]

2.13.5.1.1 Example. *Linear matrix inequality.* (*confer* §2.13.2.0.3)
Consider a peculiar vertex-description for a convex cone $\mathcal{K}$ defined over a positive semidefinite cone (instead of a nonnegative orthant as in definition (103)): for $X \in \mathbb{S}^n_+$ given $A_j \in \mathbb{S}^n$, $j=1\ldots m$

$$\begin{aligned}\mathcal{K} &= \left\{ \begin{bmatrix} \langle A_1\,,\,X\rangle \\ \vdots \\ \langle A_m\,,\,X\rangle \end{bmatrix} \mid X\succeq 0 \right\} \subseteq \mathbb{R}^m \\ &= \left\{ \begin{bmatrix} \operatorname{svec}(A_1)^{\mathrm T} \\ \vdots \\ \operatorname{svec}(A_m)^{\mathrm T} \end{bmatrix} \operatorname{svec} X \mid X\succeq 0 \right\} \\ &\triangleq \{A \operatorname{svec} X \mid X\succeq 0\}\end{aligned} \tag{379}$$

where $A\in\mathbb{R}^{m\times n(n+1)/2}$, and where symmetric vectorization svec is defined in (56). Cone $\mathcal{K}$ is indeed convex because, by (175)

$$A \operatorname{svec} X_{\mathrm{p}_1}\,,\, A \operatorname{svec} X_{\mathrm{p}_2} \in \mathcal{K} \;\Rightarrow\; A(\zeta \operatorname{svec} X_{\mathrm{p}_1} + \xi \operatorname{svec} X_{\mathrm{p}_2}) \in \mathcal{K} \;\text{ for all } \zeta\,,\xi \geq 0 \tag{380}$$

since a nonnegatively weighted sum of positive semidefinite matrices must be positive semidefinite. (§A.3.1.0.2) Although matrix A is finite-dimensional, $\mathcal{K}$ is generally not a polyhedral cone (unless $m=1$ or 2) simply because $X\in\mathbb{S}^n_+$.

> **Theorem.** *Inverse image closedness.* [215, prop.A.2.1.12] [325, thm.6.7]
> Given affine operator $g:\mathbb{R}^m\to\mathbb{R}^p$, convex set $\mathcal{D}\subseteq\mathbb{R}^m$, and convex set $\mathcal{C}\subseteq\mathbb{R}^p \ni g^{-1}(\operatorname{rel\,int}\mathcal{C})\neq\emptyset$, then
>
> $$\operatorname{rel\,int} g(\mathcal{D}) = g(\operatorname{rel\,int}\mathcal{D})\,,\quad \operatorname{rel\,int} g^{-1}\mathcal{C} = g^{-1}(\operatorname{rel\,int}\mathcal{C})\,,\quad \overline{g^{-1}\mathcal{C}} = g^{-1}\overline{\mathcal{C}} \tag{381}$$
> ⋄

By this theorem, relative interior of $\mathcal{K}$ may always be expressed

$$\operatorname{rel\,int}\mathcal{K} = \{A \operatorname{svec} X \mid X\succ 0\} \tag{382}$$

Because $\dim(\operatorname{aff}\mathcal{K}) = \dim(A \operatorname{svec}\mathbb{S}^n)$ (127) then, provided the vectorized A_j matrices are linearly independent,

$$\operatorname{rel\,int}\mathcal{K} = \operatorname{int}\mathcal{K} \tag{14}$$

meaning, cone $\mathcal{K}$ is full-dimensional $\Rightarrow$ dual cone $\mathcal{K}^*$ is pointed by (309). Convex cone $\mathcal{K}$ can be closed, by this corollary:

> **Corollary.** *Cone closedness invariance.* [57, §3] [58, §3]
> Given linear operator $A:\mathbb{R}^p\to\mathbb{R}^m$ and closed convex cone $\mathcal{X}\subseteq\mathbb{R}^p$, convex cone $\mathcal{K}=A(\mathcal{X})$ is closed $\left(\overline{A(\mathcal{X})} = A(\mathcal{X})\right)$ if and only if
>
> $$\mathcal{N}(A)\cap\mathcal{X} = \{\mathbf{0}\} \qquad \textbf{or} \qquad \mathcal{N}(A)\cap\mathcal{X} \nsubseteq \operatorname{rel}\partial\mathcal{X} \tag{383}$$
>
> Otherwise, $\overline{\mathcal{K}} = \overline{A(\mathcal{X})} \supseteq A(\overline{\mathcal{X}}) \supseteq A(\mathcal{X})$. [325, thm.6.6] ⋄

If matrix A has no nontrivial nullspace, then $A\,\text{svec}\,X$ is an isomorphism in X between cone $\mathbb{S}^n_+$ and range $\mathcal{R}(A)$ of matrix A; (§2.2.1.0.1, §2.10.1.1) sufficient for convex cone $\mathcal{K}$ to be closed and have relative boundary

$$\text{rel}\,\partial\mathcal{K} = \{A\,\text{svec}\,X \mid X \succeq 0\,,\;\; X \nsucc 0\} \tag{384}$$

Now consider the (closed convex) dual cone:

$$\begin{aligned}\mathcal{K}^* &= \{y \mid \langle z\,,\,y\rangle \geq 0 \;\text{ for all }\; z\in\mathcal{K}\} \subseteq \mathbb{R}^m\\ &= \{y \mid \langle z\,,\,y\rangle \geq 0 \;\text{ for all }\; z = A\,\text{svec}\,X\,,\;\; X\succeq 0\}\\ &= \{y \mid \langle A\,\text{svec}\,X\,,\;y\rangle \geq 0 \;\text{ for all }\; X\succeq 0\}\\ &= \left\{y \mid \langle \text{svec}\,X\,,\;A^{\mathrm{T}}y\rangle \geq 0 \;\text{ for all }\; X\succeq 0\right\}\\ &= \left\{y \mid \;\text{svec}^{-1}(A^{\mathrm{T}}y) \succeq 0\right\}\end{aligned} \tag{385}$$

that follows from (377) and leads to an equally peculiar halfspace-description

$$\mathcal{K}^* = \{y\in\mathbb{R}^m \mid \sum_{j=1}^m y_j A_j \succeq 0\} \tag{386}$$

The summation inequality with respect to positive semidefinite cone $\mathbb{S}^n_+$ is known as *linear matrix inequality.* [61] [165] [278] [381] Although we already know that the dual cone is convex (§2.13.1), *inverse image theorem* 2.1.9.0.1 certifies convexity of $\mathcal{K}^*$ which is the inverse image of positive semidefinite cone $\mathbb{S}^n_+$ under linear transformation $g(y) \triangleq \sum y_j A_j$. And although we already know that the dual cone is closed, it is certified by (381). By the *inverse image closedness theorem,* dual cone relative interior may always be expressed

$$\text{rel int}\,\mathcal{K}^* = \{y\in\mathbb{R}^m \mid \sum_{j=1}^m y_j A_j \succ 0\} \tag{387}$$

Function $g(y)$ on $\mathbb{R}^m$ is an isomorphism when the vectorized A_j matrices are linearly independent; hence, uniquely invertible. Inverse image $\mathcal{K}^*$ must therefore have dimension equal to $\dim\!\big(\mathcal{R}(A^{\mathrm{T}}) \cap \text{svec}\,\mathbb{S}^n_+\big)$ (49) and relative boundary

$$\text{rel}\,\partial\mathcal{K}^* = \{y\in\mathbb{R}^m \mid \sum_{j=1}^m y_j A_j \succeq 0\,,\;\; \sum_{j=1}^m y_j A_j \nsucc 0\} \tag{388}$$

When this dimension equals m, then dual cone $\mathcal{K}^*$ is full-dimensional

$$\text{rel int}\,\mathcal{K}^* = \text{int}\,\mathcal{K}^* \qquad (14)$$

which implies: closure of convex cone $\mathcal{K}$ is pointed (309). □

2.13.6 Dual of pointed polyhedral cone

In a subspace of $\mathbb{R}^n$, now we consider a pointed polyhedral cone $\mathcal{K}$ given in terms of its extreme directions Γ_i arranged columnar in

$$X = [\,\Gamma_1 \;\; \Gamma_2 \;\cdots\; \Gamma_N\,] \in \mathbb{R}^{n\times N} \qquad (279)$$

The *extremes theorem* (§2.8.1.1.1) provides the vertex-description of a pointed polyhedral cone in terms of its finite number of extreme directions and its lone vertex at the origin:

2.13.6.0.1 Definition. *Pointed polyhedral cone, vertex-description.*
Given pointed polyhedral cone $\mathcal{K}$ in a subspace of $\mathbb{R}^n$, denoting its i^{th} extreme direction by $\Gamma_i \in \mathbb{R}^n$ arranged in a matrix X as in (279), then that cone may be described: (86) (*confer* (188) (292))

$$\begin{aligned} \mathcal{K} &= \left\{ [\,\mathbf{0} \;\; X\,]\, a\,\zeta \mid a^{\mathrm{T}}\mathbf{1} = 1,\; a \succeq 0,\; \zeta \geq 0 \right\} \\ &= \left\{ X a\,\zeta \mid a^{\mathrm{T}}\mathbf{1} \leq 1,\; a \succeq 0,\; \zeta \geq 0 \right\} \\ &= \left\{ X b \mid b \succeq 0 \right\} \subseteq \mathbb{R}^n \end{aligned} \tag{389}$$

that is simply a conic hull (like (103)) of a finite number N of directions. Relative interior may always be expressed

$$\operatorname{rel\,int} \mathcal{K} = \{ X b \mid b \succ 0 \} \subset \mathbb{R}^n \tag{390}$$

but identifying the cone's relative boundary in this manner

$$\operatorname{rel} \partial \mathcal{K} = \{ X b \mid b \succeq 0\,,\; b \nsucc 0 \} \tag{391}$$

holds only when matrix X represents a bijection onto its range; in other words, some coefficients meeting lower bound zero ($b \in \partial \mathbb{R}^N_+$) do not necessarily provide membership to relative boundary of cone $\mathcal{K}$. $\triangle$

Whenever cone $\mathcal{K}$ is pointed, closed, and convex (not only polyhedral), then dual cone $\mathcal{K}^*$ has a halfspace-description in terms of the extreme directions Γ_i of $\mathcal{K}$:

$$\mathcal{K}^* = \left\{ y \mid \gamma^{\mathrm{T}} y \geq 0 \;\text{ for all }\; \gamma \in \{\Gamma_i\,,\; i = 1 \ldots N\} \subseteq \operatorname{rel} \partial \mathcal{K} \right\} \tag{392}$$

because when $\{\Gamma_i\}$ constitutes any set of generators for $\mathcal{K}$, the discretization result in §2.13.4.1 allows relaxation of the requirement $\forall\, x \in \mathcal{K}$ in (296) to $\forall\, \gamma \in \{\Gamma_i\}$ directly.[2.70] That dual cone so defined is unique, identical to (296), polyhedral whenever the number of generators N is finite

$$\mathcal{K}^* = \left\{ y \mid X^{\mathrm{T}} y \succeq 0 \right\} \subseteq \mathbb{R}^n \qquad (362)$$

and is full-dimensional because $\mathcal{K}$ is assumed pointed. But $\mathcal{K}^*$ is not necessarily pointed unless $\mathcal{K}$ is full-dimensional (§2.13.1.1).

2.13.6.1 Facet normal & extreme direction

We see from (362) that the conically independent generators of cone $\mathcal{K}$ (namely, the extreme directions of pointed closed convex cone $\mathcal{K}$ constituting the N columns of X) each define an inward-normal to a hyperplane supporting dual cone $\mathcal{K}^*$ (§2.4.2.6.1) and exposing a dual facet when N is finite. Were $\mathcal{K}^*$ pointed and finitely generated, then by closure the conjugate statement would also hold; *id est*, the extreme directions of pointed $\mathcal{K}^*$ each define an inward-normal to a hyperplane supporting $\mathcal{K}$ and exposing a facet when N is finite. Examine Figure **60** or Figure **66**, for example.

We may conclude, the extreme directions of proper polyhedral $\mathcal{K}$ are respectively orthogonal to the facets of $\mathcal{K}^*$; likewise, the extreme directions of proper polyhedral $\mathcal{K}^*$ are respectively orthogonal to the facets of $\mathcal{K}$.

[2.70] The extreme directions of $\mathcal{K}$ constitute a minimal set of generators. Formulae and conversions to vertex-description of the dual cone are in §2.13.9 and §2.13.11.

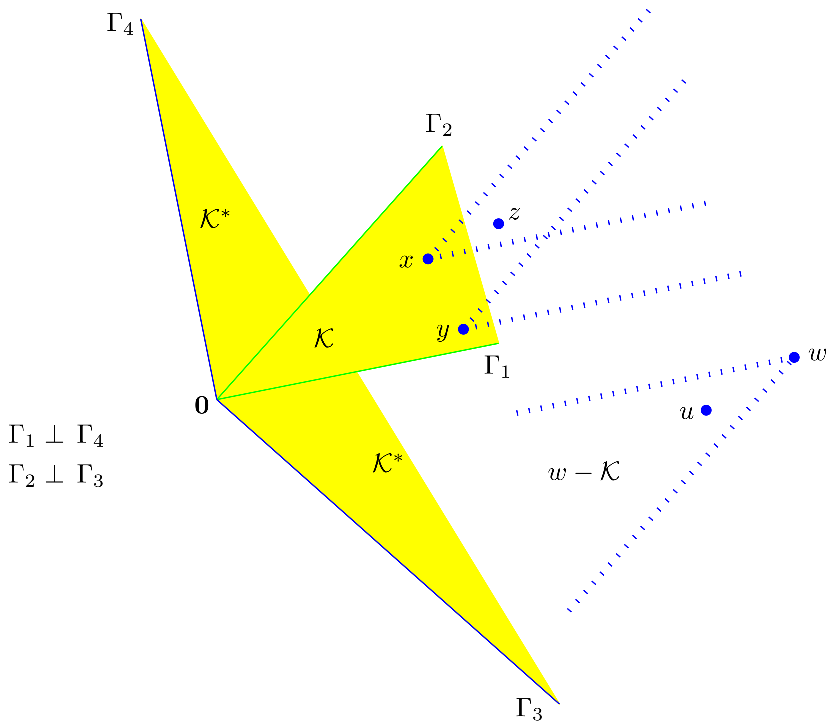


Figure 67: (*confer* Figure **177**) Simplicial cone $\mathcal{K} \in \mathbb{R}^{\mathbf{2}}$ and its dual $\mathcal{K}^*$ drawn truncated. Conically independent generators Γ_1 and Γ_2 constitute extreme directions of $\mathcal{K}$ while Γ_3 and Γ_4 constitute extreme directions of $\mathcal{K}^*$. Dotted ray-pairs bound translated cones $\mathcal{K}$. Point x is comparable to point z (and *vice versa*) but not to y; $z \succeq_{\mathcal{K}} x \Leftrightarrow z - x \in \mathcal{K} \Leftrightarrow z - x \succeq_{\mathcal{K}} 0$ iff $\exists$ nonnegative coordinates for biorthogonal expansion of $z - x$. Point y is not comparable to z because z does not belong to $y \pm \mathcal{K}$. Translating a negated cone is quite helpful for visualization: $u \preceq_{\mathcal{K}} w \Leftrightarrow u \in w - \mathcal{K} \Leftrightarrow u - w \preceq_{\mathcal{K}} 0$. Points need not belong to $\mathcal{K}$ to be comparable; *e.g*, all points less than w (w.r.t $\mathcal{K}$) belong to $w - \mathcal{K}$.

2.13.7 Biorthogonal expansion by example

2.13.7.0.1 Example. *Relationship to dual polyhedral cone.*
Simplicial cone $\mathcal{K}$ illustrated in Figure **67** induces a partial order on $\mathbb{R}^2$. All points greater than x with respect to $\mathcal{K}$, for example, are contained in the translated cone $x+\mathcal{K}$. The extreme directions Γ_1 and Γ_2 of $\mathcal{K}$ do not make an orthogonal set; neither do extreme directions Γ_3 and Γ_4 of dual cone $\mathcal{K}^*$; rather, we have the *biorthogonality condition* [384]

$$\begin{array}{c}\Gamma_4^{\mathrm{T}}\Gamma_1 = \Gamma_3^{\mathrm{T}}\Gamma_2 = 0\\ \Gamma_3^{\mathrm{T}}\Gamma_1 \neq 0\,, \quad \Gamma_4^{\mathrm{T}}\Gamma_2 \neq 0\end{array} \tag{393}$$

Biorthogonal expansion of $x\in\mathcal{K}$ is then

$$x \;=\; \Gamma_1\,\frac{\Gamma_3^{\mathrm{T}}x}{\Gamma_3^{\mathrm{T}}\Gamma_1} + \Gamma_2\,\frac{\Gamma_4^{\mathrm{T}}x}{\Gamma_4^{\mathrm{T}}\Gamma_2} \tag{394}$$

where $\Gamma_3^{\mathrm{T}}x/(\Gamma_3^{\mathrm{T}}\Gamma_1)$ is the nonnegative coefficient of nonorthogonal projection (§E.6.1) of x on Γ_1 in the direction orthogonal to Γ_3 (y in Figure **177** p.619), and where $\Gamma_4^{\mathrm{T}}x/(\Gamma_4^{\mathrm{T}}\Gamma_2)$ is the nonnegative coefficient of nonorthogonal projection of x on Γ_2 in the direction orthogonal to Γ_4 (z in Figure **177**); they are coordinates in this nonorthogonal system. Those coefficients must be nonnegative $x\succeq_{\mathcal{K}} 0$ because $x\in\mathcal{K}$ (324) and $\mathcal{K}$ is simplicial.

If we ascribe the extreme directions of $\mathcal{K}$ to the columns of a matrix

$$X \triangleq [\,\Gamma_1 \;\; \Gamma_2\,] \tag{395}$$

then we find that the pseudoinverse transpose matrix

$$X^{\dagger\mathrm{T}} = \left[\,\Gamma_3\frac{1}{\Gamma_3^{\mathrm{T}}\Gamma_1} \quad \Gamma_4\frac{1}{\Gamma_4^{\mathrm{T}}\Gamma_2}\,\right] \tag{396}$$

holds the extreme directions of the dual cone. Therefore

$$x = XX^{\dagger}x \qquad (402)$$

is biorthogonal expansion (394) (§E.0.1), and biorthogonality condition (393) can be expressed succinctly (§E.1.1)[2.71]

$$X^{\dagger}X = I \qquad (403)$$

Expansion $w=XX^{\dagger}w$, for any particular $w\in\mathbb{R}^n$ more generally, is unique w.r.t X if and only if the extreme directions of $\mathcal{K}$ populating the columns of $X\in\mathbb{R}^{n\times N}$ are linearly independent; *id est*, iff X has no nullspace. □

2.13.7.0.2 Exercise. *Visual comparison of real sums.*
Given $y\preceq x$ with respect to the nonnegative orthant, draw a figure showing a negated shifted orthant (like the cone in Figure **67**) illustrating why $\mathbf{1}^{\mathrm{T}}y\leq\mathbf{1}^{\mathrm{T}}x$ for y and x anywhere in $\mathbb{R}^2$. Incorporate two hyperplanes into your drawing, one containing y and another containing x with reference to Figure **29**. Does this result hold in higher dimension? ▼

[2.71] Possibly confusing is the fact that formula $XX^{\dagger}x$ is simultaneously: the orthogonal projection of x on $\mathcal{R}(X)$ (2013), and a sum of nonorthogonal projections of $x\in\mathcal{R}(X)$ on the range of each and every column of full-rank X skinny-or-square (§E.5.0.0.2).

2.13.7.1 Pointed cones and biorthogonality

Biorthogonality condition $X^\dagger X = I$ from Example 2.13.7.0.1 means Γ_1 and Γ_2 are linearly independent generators of $\mathcal{K}$ (§B.1.1.1); generators because every $x \in \mathcal{K}$ is their conic combination. From §2.10.2 we know that means Γ_1 and Γ_2 must be extreme directions of $\mathcal{K}$.

A biorthogonal expansion is necessarily associated with a pointed closed convex cone; pointed, otherwise there can be no extreme directions (§2.8.1). We will address biorthogonal expansion with respect to a pointed polyhedral cone not full-dimensional in §2.13.8.

2.13.7.1.1 Example. *Expansions implied by diagonalization.* (*confer* §6.4.3.2.1)
When matrix $X \in \mathbb{R}^{M\times M}$ is *diagonalizable* (§A.5),

$$X \;=\; S\Lambda S^{-1} \;=\; [\, s_1 \cdots s_M \,]\; \Lambda \begin{bmatrix} w_1^{\mathrm{T}} \\ \vdots \\ w_M^{\mathrm{T}} \end{bmatrix} \;=\; \sum_{i=1}^{M} \lambda_i \, s_i w_i^{\mathrm{T}} \tag{1636}$$

coordinates for biorthogonal expansion are its eigenvalues λ_i (contained in diagonal matrix Λ) when expanded in S;

$$X = SS^{-1}X \;=\; [\, s_1 \cdots s_M \,] \begin{bmatrix} w_1^{\mathrm{T}} X \\ \vdots \\ w_M^{\mathrm{T}} X \end{bmatrix} = \sum_{i=1}^{M} \lambda_i \, s_i w_i^{\mathrm{T}} \tag{397}$$

Coordinate values depend upon geometric relationship of X to its linearly independent eigenmatrices $s_i w_i^{\mathrm{T}}$. (§A.5.0.3, §B.1.1)

- Eigenmatrices $s_i w_i^{\mathrm{T}}$ are linearly independent dyads constituted by right and left eigenvectors of diagonalizable X and are generators of some pointed polyhedral cone $\mathcal{K}$ in a subspace of $\mathbb{R}^{M\times M}$.

When S is real and X belongs to that polyhedral cone $\mathcal{K}$, for example, then coordinates of expansion (the eigenvalues λ_i) must be nonnegative.

When $X = Q\Lambda Q^{\mathrm{T}}$ is symmetric, coordinates for biorthogonal expansion are its eigenvalues when expanded in Q; *id est*, for $X \in \mathbb{S}^M$

$$X \;=\; QQ^{\mathrm{T}}X \;=\; \sum_{i=1}^{M} q_i \, q_i^{\mathrm{T}} X \;=\; \sum_{i=1}^{M} \lambda_i \, q_i q_i^{\mathrm{T}} \;\in\; \mathbb{S}^M \tag{398}$$

becomes an orthogonal expansion with *orthonormality condition* $Q^{\mathrm{T}}Q = I$ where λ_i is the i^{th} eigenvalue of X, q_i is the corresponding i^{th} eigenvector arranged columnar in orthogonal matrix

$$Q = [\, q_1 \;\; q_2 \;\; \cdots \;\; q_M \,] \in \mathbb{R}^{M\times M} \tag{399}$$

and where eigenmatrix $q_i q_i^{\mathrm{T}}$ is an extreme direction of some pointed polyhedral cone $\mathcal{K} \subset \mathbb{S}^M$ and an extreme direction of the positive semidefinite cone $\mathbb{S}_+^M$.

- Orthogonal expansion is a special case of biorthogonal expansion of $X \in \operatorname{aff} \mathcal{K}$ occurring when polyhedral cone $\mathcal{K}$ is any rotation about the origin of an orthant belonging to a subspace.

Similarly, when $X = Q\Lambda Q^{\mathrm{T}}$ belongs to the positive semidefinite cone in the subspace of symmetric matrices, coordinates for orthogonal expansion must be its nonnegative eigenvalues (1539) when expanded in Q; *id est*, for $X \in \mathbb{S}_+^M$

$$X = QQ^{\mathrm{T}}X = \sum_{i=1}^{M} q_i\, q_i^{\mathrm{T}} X = \sum_{i=1}^{M} \lambda_i\, q_i q_i^{\mathrm{T}} \in \mathbb{S}_+^M \tag{400}$$

where $\lambda_i \geq 0$ is the i^{th} eigenvalue of X. This means matrix X simultaneously belongs to the positive semidefinite cone and to the pointed polyhedral cone $\mathcal{K}$ formed by the conic hull of its eigenmatrices. □

2.13.7.1.2 Example. *Expansion respecting nonpositive orthant.*
Suppose $x \in \mathcal{K}$ any orthant in $\mathbb{R}^n$.[2.72] Then coordinates for biorthogonal expansion of x must be nonnegative; in fact, absolute value of the Cartesian coordinates.

Suppose, in particular, x belongs to the nonpositive orthant $\mathcal{K} = \mathbb{R}_-^n$. Then biorthogonal expansion becomes orthogonal expansion

$$x = XX^{\mathrm{T}}x = \sum_{i=1}^{n} -e_i(-e_i^{\mathrm{T}}x) = \sum_{i=1}^{n} -e_i|e_i^{\mathrm{T}}x| \in \mathbb{R}_-^n \tag{401}$$

and the coordinates of expansion are nonnegative. For this orthant $\mathcal{K}$ we have orthonormality condition $X^{\mathrm{T}}X = I$ where $X = -I$, $e_i \in \mathbb{R}^n$ is a standard basis vector, and $-e_i$ is an extreme direction (§2.8.1) of $\mathcal{K}$.

Of course, this expansion $x = XX^{\mathrm{T}}x$ applies more broadly to domain $\mathbb{R}^n$, but then the coordinates each belong to all of $\mathbb{R}$. □

2.13.8 Biorthogonal expansion, derivation

Biorthogonal expansion is a means for determining coordinates in a pointed conic coordinate system characterized by a nonorthogonal basis. Study of nonorthogonal bases invokes pointed polyhedral cones and their duals; extreme directions of a cone $\mathcal{K}$ are assumed to constitute the *basis* while those of the dual cone $\mathcal{K}^*$ determine coordinates.

Unique biorthogonal expansion with respect to $\mathcal{K}$ relies upon existence of its linearly independent extreme directions: Polyhedral cone $\mathcal{K}$ must be pointed; then it possesses extreme directions. Those extreme directions must be linearly independent to uniquely represent any point in their span.

We consider nonempty pointed polyhedral cone $\mathcal{K}$ possibly not full-dimensional; *id est*, we consider a basis spanning a subspace. Then we need only observe that section of dual cone $\mathcal{K}^*$ in the affine hull of $\mathcal{K}$ because, by *expansion* of x, membership $x \in \operatorname{aff} \mathcal{K}$ is

[2.72] An orthant is simplicial and selfdual.

implicit and because any breach of the ordinary dual cone into ambient space becomes irrelevant (§2.13.9.3). *Biorthogonal expansion*

$$x = XX^{\dagger}x \in \operatorname{aff}\mathcal{K} = \operatorname{aff}\operatorname{cone}(X) \tag{402}$$

is expressed in the extreme directions $\{\Gamma_i\}$ of $\mathcal{K}$ arranged columnar in

$$X = [\,\Gamma_1 \;\; \Gamma_2 \;\; \cdots \;\; \Gamma_N\,] \in \mathbb{R}^{n\times N} \qquad (279)$$

under assumption of *biorthogonality*

$$X^{\dagger}X = I \tag{403}$$

where † denotes matrix pseudoinverse (§E). We therefore seek, in this section, a vertex-description for $\mathcal{K}^*\cap\operatorname{aff}\mathcal{K}$ in terms of linearly independent dual generators $\{\Gamma_i^*\}\subset\operatorname{aff}\mathcal{K}$ in the same finite quantity[2.73] as the extreme directions $\{\Gamma_i\}$ of

$$\mathcal{K} = \operatorname{cone}(X) = \{Xa \mid a\succeq 0\} \subseteq \mathbb{R}^n \qquad (103)$$

We assume the quantity of extreme directions N does not exceed the dimension n of ambient vector space because, otherwise, expansion w.r.t $\mathcal{K}$ could not be unique; *id est*, assume N linearly independent extreme directions hence $N\leq n$ (X *skinny*[2.74]-or-square full-rank). In other words, fat full-rank matrix X is prohibited by uniqueness because of existence of an infinity of right inverses;

- polyhedral cones whose extreme directions number in excess of the ambient space dimension are precluded in biorthogonal expansion.

2.13.8.1 $x\in\mathcal{K}$

Suppose x belongs to $\mathcal{K}\subseteq\mathbb{R}^n$. Then $x=Xa$ for some $a\succeq 0$. Coordinate vector a is unique only when $\{\Gamma_i\}$ is a linearly independent set.[2.75] Vector $a\in\mathbb{R}^N$ can take the form $a=Bx$ if $\mathcal{R}(B)=\mathbb{R}^N$. Then we require $Xa=XBx=x$ and $Bx=BXa=a$. The pseudoinverse $B=X^{\dagger}\in\mathbb{R}^{N\times n}$ (§E) is suitable when X is skinny-or-square and full-rank. In that case $\operatorname{rank}X=N$, and for all $c\succeq 0$ and $i=1\ldots N$

$$a\succeq 0 \;\Leftrightarrow\; X^{\dagger}Xa\succeq 0 \;\Leftrightarrow\; a^{\mathrm T}X^{\mathrm T}X^{\dagger\mathrm T}c\geq 0 \;\Leftrightarrow\; \Gamma_i^{\mathrm T}X^{\dagger\mathrm T}c\geq 0 \tag{404}$$

The penultimate inequality follows from the *generalized inequality and membership corollary,* while the last inequality is a consequence of that corollary's discretization (§2.13.4.2.1).[2.76] From (404) and (392) we deduce

$$\mathcal{K}^*\cap\operatorname{aff}\mathcal{K} = \operatorname{cone}(X^{\dagger\mathrm T}) = \{X^{\dagger\mathrm T}c \mid c\succeq 0\}\subseteq\mathbb{R}^n \tag{405}$$

[2.73] When $\mathcal{K}$ is contained in a proper subspace of $\mathbb{R}^n$, the ordinary dual cone $\mathcal{K}^*$ will have more generators in any minimal set than $\mathcal{K}$ has extreme directions.

[2.74] "Skinny" meaning thin; more rows than columns.

[2.75] Conic independence alone (§2.10) is insufficient to guarantee uniqueness.

[2.76]

$$\begin{aligned} a\succeq 0 \;\Leftrightarrow\; a^{\mathrm T}X^{\mathrm T}X^{\dagger\mathrm T}c\geq 0 \quad &\forall\,(c\succeq 0 \;\Leftrightarrow\; a^{\mathrm T}X^{\mathrm T}X^{\dagger\mathrm T}c\geq 0 \quad \forall\,a\succeq 0)\\ &\forall\,(c\succeq 0 \;\Leftrightarrow\; \Gamma_i^{\mathrm T}X^{\dagger\mathrm T}c\geq 0 \quad \forall\,i) \end{aligned} \qquad \blacklozenge$$

is the vertex-description for that section of $\mathcal{K}^*$ in the affine hull of $\mathcal{K}$ because $\mathcal{R}(X^{\dagger\mathrm{T}}) = \mathcal{R}(X)$ by definition of the pseudoinverse. From (309), we know $\mathcal{K}^* \cap \operatorname{aff}\mathcal{K}$ must be pointed if $\operatorname{rel\,int}\mathcal{K}$ is logically assumed nonempty with respect to $\operatorname{aff}\mathcal{K}$.

Conversely, suppose full-rank skinny-or-square matrix $(N \le n)$

$$X^{\dagger\mathrm{T}} \triangleq [\,\Gamma_1^* \;\; \Gamma_2^* \, \cdots \, \Gamma_N^*\,] \in \mathbb{R}^{n\times N} \tag{406}$$

comprises the extreme directions $\{\Gamma_i^*\} \subset \operatorname{aff}\mathcal{K}$ of the dual-cone intersection with the affine hull of $\mathcal{K}$.[2.77] From the *discretized membership theorem* and (313) we get a partial dual to (392); *id est*, assuming $x \in \operatorname{aff\,cone} X$

$$x \in \mathcal{K} \quad\Leftrightarrow\quad \gamma^{*\mathrm{T}} x \ge 0 \;\text{ for all }\; \gamma^* \in \{\Gamma_i^*,\, i=1\ldots N\} \subset \partial\mathcal{K}^* \cap \operatorname{aff}\mathcal{K} \tag{407}$$

$$\Leftrightarrow\quad X^\dagger x \succeq 0 \tag{408}$$

that leads to a partial halfspace-description,

$$\mathcal{K} \;=\; \left\{x \in \operatorname{aff\,cone} X \mid X^\dagger x \succeq 0\right\} \tag{409}$$

For $\gamma^* = X^{\dagger\mathrm{T}} e_i$, any $x = Xa$, and for all i we have $e_i^{\mathrm{T}} X^\dagger X a = e_i^{\mathrm{T}} a \ge 0$ only when $a \succeq 0$. Hence $x \in \mathcal{K}$.

When X is full-rank, then unique biorthogonal expansion of $x \in \mathcal{K}$ becomes (402)

$$x \;=\; XX^\dagger x \;=\; \sum_{i=1}^{N} \Gamma_i\, \Gamma_i^{*\mathrm{T}} x \tag{410}$$

whose *coordinates* $a = \Gamma_i^{*\mathrm{T}} x$ must be nonnegative because $\mathcal{K}$ is assumed pointed, closed, and convex. Whenever X is full-rank, so is its pseudoinverse $X^\dagger$. (§E) In the present case, the columns of $X^{\dagger\mathrm{T}}$ are linearly independent and generators of the dual cone $\mathcal{K}^* \cap \operatorname{aff}\mathcal{K}$; hence, the columns constitute its extreme directions. (§2.10.2) That section of the dual cone is itself a polyhedral cone (by (286) or the *cone intersection theorem,* §2.7.2.1.1) having the same number of extreme directions as $\mathcal{K}$.

2.13.8.2 $x \in \operatorname{aff}\mathcal{K}$

The extreme directions of $\mathcal{K}$ and $\mathcal{K}^* \cap \operatorname{aff}\mathcal{K}$ have a distinct relationship; because $X^\dagger X = I$, then for $i,j = 1\ldots N$, $\Gamma_i^{\mathrm{T}} \Gamma_i^* = 1$, while for $i \ne j$, $\Gamma_i^{\mathrm{T}} \Gamma_j^* = 0$. Yet neither set of extreme directions, $\{\Gamma_i\}$ nor $\{\Gamma_i^*\}$, is necessarily orthogonal. This is a biorthogonality condition, precisely, [384, §2.2.4] [218] implying each set of extreme directions is linearly independent. (§B.1.1.1)

Intuitively, any nonnegative vector a is a conic combination of the standard basis $\{e_i \in \mathbb{R}^N\}$; $a \succeq 0 \Leftrightarrow a_i\, e_i \succeq 0$ for all i. The last inequality in (404) is a consequence of the fact that $x = Xa$ may be any extreme direction of $\mathcal{K}$, in which case a is a standard basis vector; $a = e_i \succeq 0$. Theoretically, because $c \succeq 0$ defines a pointed polyhedral cone (in fact, the nonnegative orthant in $\mathbb{R}^N$), we can take (404) one step further by discretizing c:

$$a \succeq 0 \;\Leftrightarrow\; \Gamma_i^{\mathrm{T}} \Gamma_j^* \ge 0 \;\text{ for } i,j = 1\ldots N \;\Leftrightarrow\; X^\dagger X \ge \mathbf{0}$$

In words, $X^\dagger X$ must be a matrix whose entries are each nonnegative.

[2.77] When closed convex cone $\mathcal{K}$ is not full-dimensional, $\mathcal{K}^*$ has no extreme directions.

Biorthogonal expansion therefore applies more broadly; meaning, for any $x \in \operatorname{aff} \mathcal{K}$, vector x can be uniquely expressed $x = Xb$ where $b \in \mathbb{R}^N$ because $\operatorname{aff} \mathcal{K}$ contains the origin. Thus, for any such $x \in \mathcal{R}(X)$ (*confer* §E.1.1), biorthogonal expansion (410) becomes $x = XX^\dagger Xb = Xb$.

2.13.9 Formulae finding dual cone

2.13.9.1 Pointed $\mathcal{K}$, dual, X skinny-or-square full-rank

We wish to derive expressions for a convex cone and its ordinary dual under the general assumptions: pointed polyhedral $\mathcal{K}$ denoted by its linearly independent extreme directions arranged columnar in matrix X such that

$$\operatorname{rank}(X \in \mathbb{R}^{n\times N}) \;=\; N \;\triangleq\; \dim \operatorname{aff} \mathcal{K} \;\leq\; n \tag{411}$$

The vertex-description is given:

$$\mathcal{K} = \{Xa \mid a \succeq 0\} \subseteq \mathbb{R}^n \tag{412}$$

from which a halfspace-description for the dual cone follows directly:

$$\mathcal{K}^* = \{y \in \mathbb{R}^n \mid X^{\mathrm{T}} y \succeq 0\} \tag{413}$$

By defining a matrix

$$X^\perp \triangleq \operatorname{basis} \mathcal{N}(X^{\mathrm{T}}) \tag{414}$$

(a columnar basis for the orthogonal complement of $\mathcal{R}(X)$), we can say

$$\operatorname{aff} \operatorname{cone} X \;=\; \operatorname{aff} \mathcal{K} \;=\; \{x \mid X^{\perp\mathrm{T}} x = \mathbf{0}\} \tag{415}$$

meaning $\mathcal{K}$ lies in a subspace, perhaps $\mathbb{R}^n$. Thus a halfspace-description

$$\mathcal{K} = \{x \in \mathbb{R}^n \mid X^\dagger x \succeq 0\,,\; X^{\perp\mathrm{T}} x = \mathbf{0}\} \tag{416}$$

and a vertex-description[2.78] from (313)

$$\mathcal{K}^* = \{\, [\, X^{\dagger\mathrm{T}} \;\; X^\perp \;\; -X^\perp\,] b \mid b \succeq 0\,\} \subseteq \mathbb{R}^n \tag{417}$$

These results are summarized for a pointed polyhedral cone, having linearly independent generators, and its ordinary dual:

Cone Table 1	$\mathcal{K}$	$\mathcal{K}^*$
vertex-description	X	$X^{\dagger\mathrm{T}}$, $\pm X^\perp$
halfspace-description	$X^\dagger$, $X^{\perp\mathrm{T}}$	X^{T}

[2.78] These descriptions are not unique. A vertex-description of the dual cone, for example, might use four conically independent generators for a plane (§2.10.0.0.1, Figure 52) when only three would suffice.

2.13.9.2 Simplicial case

When a convex cone is simplicial (§2.12.3), Cone Table **1** simplifies because then $\operatorname{aff}\operatorname{cone} X = \mathbb{R}^n$: For square X and assuming simplicial $\mathcal{K}$ such that

$$\operatorname{rank}(X \in \mathbb{R}^{n\times N}) \;=\; N \;\triangleq\; \dim\operatorname{aff}\mathcal{K} \;=\; n \tag{418}$$

we have

Cone Table S	$\mathcal{K}$	$\mathcal{K}^*$
vertex-description	X	$X^{\dagger\mathrm{T}}$
halfspace-description	$X^{\dagger}$	X^{T}

For example, vertex-description (417) simplifies to

$$\mathcal{K}^* \;=\; \{X^{\dagger\mathrm{T}} b \mid b \succeq 0\} \subset \mathbb{R}^n \tag{419}$$

Now, because $\dim\mathcal{R}(X) = \dim\mathcal{R}(X^{\dagger\mathrm{T}})$, (§E) dual cone $\mathcal{K}^*$ is simplicial whenever $\mathcal{K}$ is.

2.13.9.3 Cone membership relations in a subspace

It is obvious by definition (296) of ordinary dual cone $\mathcal{K}^*$, in ambient vector space $\mathcal{R}$, that its determination instead in subspace $\mathcal{S} \subseteq \mathcal{R}$ is identical to its intersection with $\mathcal{S}$; *id est*, assuming closed convex cone $\mathcal{K} \subseteq \mathcal{S}$ and $\mathcal{K}^* \subseteq \mathcal{R}$

$$(\mathcal{K}^* \text{ were ambient } \mathcal{S}) \;\equiv\; (\mathcal{K}^* \text{ in ambient } \mathcal{R}) \cap \mathcal{S} \tag{420}$$

because

$$\{y \in \mathcal{S} \mid \langle y\,,\,x\rangle \geq 0 \text{ for all } x \in \mathcal{K}\} \;=\; \{y \in \mathcal{R} \mid \langle y\,,\,x\rangle \geq 0 \text{ for all } x \in \mathcal{K}\} \;\cap\; \mathcal{S} \tag{421}$$

From this, a constrained membership relation for the ordinary dual cone $\mathcal{K}^* \subseteq \mathcal{R}$, assuming $x, y \in \mathcal{S}$ and closed convex cone $\mathcal{K} \subseteq \mathcal{S}$

$$y \in \mathcal{K}^* \cap \mathcal{S} \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \text{ for all } x \in \mathcal{K} \tag{422}$$

By closure in subspace $\mathcal{S}$ we have conjugation (§2.13.1.1):

$$x \in \mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle \geq 0 \text{ for all } y \in \mathcal{K}^* \cap \mathcal{S} \tag{423}$$

This means membership determination in subspace $\mathcal{S}$ requires knowledge of dual cone only in $\mathcal{S}$. For sake of completeness, for proper cone $\mathcal{K}$ with respect to subspace $\mathcal{S}$ (*confer* (325))

$$x \in \operatorname{int}\mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \text{ for all } y \in \mathcal{K}^* \cap \mathcal{S}\,,\;\; y \neq \mathbf{0} \tag{424}$$

$$x \in \mathcal{K}\,,\;\; x \neq \mathbf{0} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \text{ for all } y \in \operatorname{int}\mathcal{K}^* \cap \mathcal{S} \tag{425}$$

(By closure, we also have the conjugate relations.) Yet when $\mathcal{S}$ equals $\operatorname{aff}\mathcal{K}$ for $\mathcal{K}$ a closed convex cone

$$x \in \operatorname{rel\,int}\mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \text{ for all } y \in \mathcal{K}^* \cap \operatorname{aff}\mathcal{K}\,,\;\; y \neq \mathbf{0} \tag{426}$$

$$x \in \mathcal{K}\,,\;\; x \neq \mathbf{0} \;\Leftrightarrow\; \langle y\,,\,x\rangle > 0 \text{ for all } y \in \operatorname{rel\,int}(\mathcal{K}^* \cap \operatorname{aff}\mathcal{K}) \tag{427}$$

2.13.9.4 Subspace $\mathcal{S} = \operatorname{aff}\mathcal{K}$

Assume now a subspace $\mathcal{S}$ that is the affine hull of cone $\mathcal{K}$: Consider again a pointed polyhedral cone $\mathcal{K}$ denoted by its extreme directions arranged columnar in matrix X such that

$$\operatorname{rank}(X\in\mathbb{R}^{n\times N}) \;=\; N \;\triangleq\; \dim\operatorname{aff}\mathcal{K} \;\leq\; n \tag{411}$$

We want expressions for the convex cone and its dual in subspace $\mathcal{S}=\operatorname{aff}\mathcal{K}$:

Cone Table A	$\mathcal{K}$	$\mathcal{K}^*\cap\operatorname{aff}\mathcal{K}$
vertex-description	X	$X^{\dagger\mathrm{T}}$
halfspace-description	$X^{\dagger}\,,\ X^{\perp\mathrm{T}}$	$X^{\mathrm{T}},\ X^{\perp\mathrm{T}}$

When $\dim\operatorname{aff}\mathcal{K}=n$, this table reduces to Cone Table **S**. These descriptions facilitate work in a proper subspace. The subspace of symmetric matrices $\mathbb{S}^N$, for example, often serves as ambient space.[2.79]

2.13.9.4.1 Exercise. *Conically independent columns and rows.*
We suspect the number of conically independent columns (rows) of X to be the same for $X^{\dagger\mathrm{T}}$, where † denotes matrix pseudoinverse (§E). Prove whether it holds that the columns (rows) of X are c.i. $\Leftrightarrow$ the columns (rows) of $X^{\dagger\mathrm{T}}$ are c.i. ▼

2.13.9.4.2 Example. *Monotone nonnegative cone.* [63, exer.2.33] [372, §2]
Simplicial cone (§2.12.3.1.1) $\mathcal{K}_{\mathcal{M}+}$ is the cone of all nonnegative vectors having their entries sorted in nonincreasing order:

$$\begin{aligned}\mathcal{K}_{\mathcal{M}+} &\triangleq \{x \mid x_1 \geq x_2 \geq \cdots \geq x_n \geq 0\} \subseteq \mathbb{R}^n_+\\ &= \{x \mid (e_i - e_{i+1})^{\mathrm{T}}x \geq 0,\ i=1\ldots n{-}1,\ e_n^{\mathrm{T}}x \geq 0\}\\ &= \{x \mid X^{\dagger}x \succeq 0\}\end{aligned} \tag{428}$$

a halfspace-description where e_i is the i^{th} standard basis vector, and where[2.80]

$$X^{\dagger\mathrm{T}} \triangleq [\, e_1{-}e_2 \quad e_2{-}e_3 \quad \cdots \quad e_n \,] \in \mathbb{R}^{n\times n} \tag{429}$$

For any vectors x and y, simple algebra demands

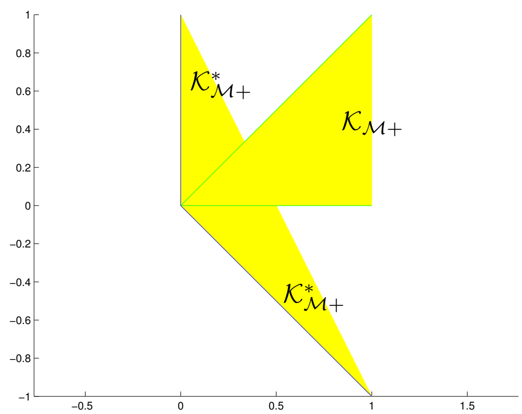
$$\begin{aligned}x^{\mathrm{T}}y \;=\; \sum_{i=1}^n x_i\, y_i \;=\; &(x_1-x_2)y_1 + (x_2-x_3)(y_1+y_2) + (x_3-x_4)(y_1+y_2+y_3) + \cdots\\ &+ (x_{n-1}-x_n)(y_1+\cdots+y_{n-1}) + x_n(y_1+\cdots+y_n)\end{aligned} \tag{430}$$

Because $x_i - x_{i+1} \geq 0\ \forall\, i$ by assumption whenever $x\in\mathcal{K}_{\mathcal{M}+}$, we can employ dual generalized inequalities (322) with respect to the selfdual nonnegative orthant $\mathbb{R}^n_+$ to find

[2.79] The dual cone of positive semidefinite matrices $\mathbb{S}^{N*}_+ = \mathbb{S}^N_+$ remains in $\mathbb{S}^N$ by convention, whereas the ordinary dual cone would venture into $\mathbb{R}^{N\times N}$.

[2.80] With $X^\dagger$ in hand, we might concisely scribe the remaining vertex- and halfspace-descriptions from the tables for $\mathcal{K}_{\mathcal{M}+}$ and its dual. Instead we use dual generalized inequalities in their derivation.

(a)



(b)

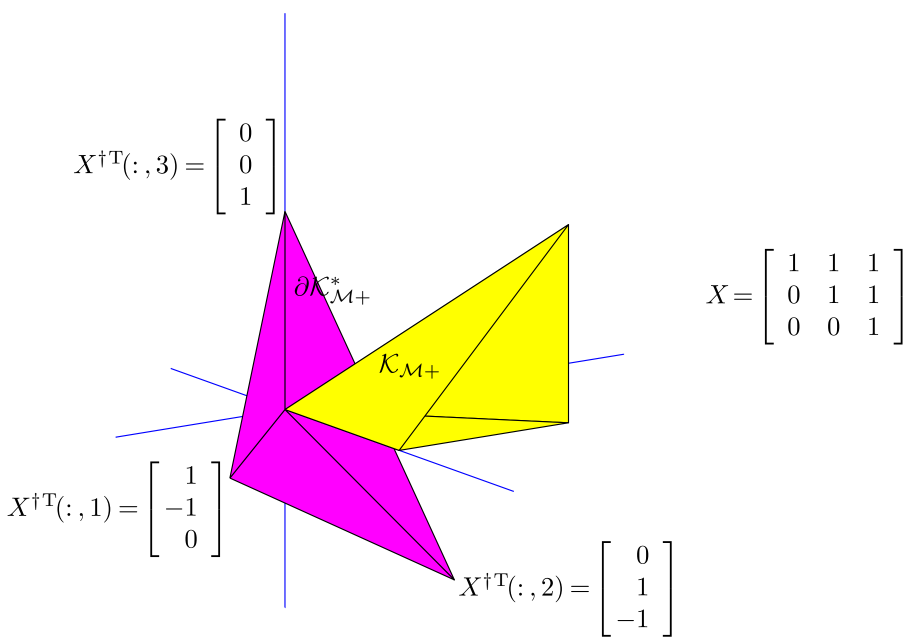


Figure 68: Simplicial cones. **(a)** Monotone nonnegative cone $\mathcal{K}_{\mathcal{M}+}$ and its dual $\mathcal{K}^*_{\mathcal{M}+}$ (drawn truncated) in $\mathbb{R}^2$. **(b)** Monotone nonnegative cone and boundary of its dual (both drawn truncated) in $\mathbb{R}^3$. Extreme directions of $\mathcal{K}^*_{\mathcal{M}+}$ are indicated.

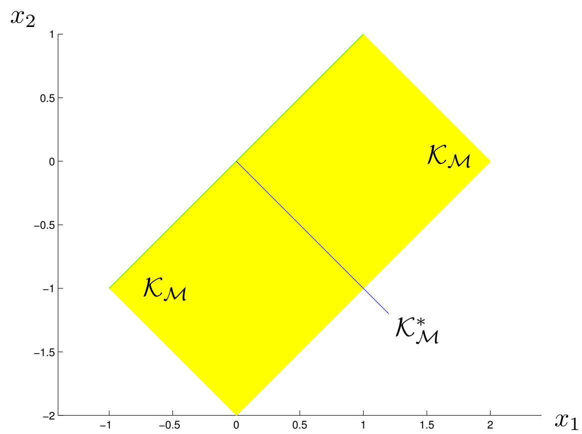


Figure 69: Monotone cone $\mathcal{K}_\mathcal{M}$ and its dual $\mathcal{K}^*_\mathcal{M}$ (drawn truncated) in $\mathbb{R}^2$.

the halfspace-description of dual monotone nonnegative cone $\mathcal{K}^*_{\mathcal{M}+}$. We can say $x^{\mathrm{T}}y \geq 0$ for all $X^\dagger x \succeq 0$ [*sic*] if and only if

$$y_1 \geq 0\,, \quad y_1 + y_2 \geq 0\,, \quad \ldots\,, \quad y_1 + y_2 + \cdots + y_n \geq 0 \tag{431}$$

id est,

$$x^{\mathrm{T}}y \geq 0 \;\; \forall\, X^\dagger x \succeq 0 \quad \Leftrightarrow \quad X^{\mathrm{T}}y \succeq 0 \tag{432}$$

where

$$X = [\, e_1 \quad e_1 + e_2 \quad e_1 + e_2 + e_3 \quad \cdots \quad \mathbf{1}\,] \in \mathbb{R}^{n\times n} \tag{433}$$

Because $X^\dagger x \succeq 0$ connotes membership of x to pointed $\mathcal{K}_{\mathcal{M}+}$, then by (296) the dual cone we seek comprises all y for which (432) holds; thus its halfspace-description

$$\mathcal{K}^*_{\mathcal{M}+} = \{y \underset{\mathcal{K}^*_{\mathcal{M}+}}{\succeq} 0\} = \{y \mid \textstyle\sum_{i=1}^k y_i \geq 0\,,\; k = 1 \ldots n\} = \{y \mid X^{\mathrm{T}}y \succeq 0\} \subset \mathbb{R}^n \tag{434}$$

The monotone nonnegative cone and its dual are simplicial, illustrated for two Euclidean spaces in Figure **68**.

From §2.13.6.1, the extreme directions of proper $\mathcal{K}_{\mathcal{M}+}$ are respectively orthogonal to the facets of $\mathcal{K}^*_{\mathcal{M}+}$. Because $\mathcal{K}^*_{\mathcal{M}+}$ is simplicial, the inward-normals to its facets constitute the linearly independent rows of X^{T} by (434). Hence the vertex-description for $\mathcal{K}_{\mathcal{M}+}$ employs the columns of X in agreement with Cone Table **S** because $X^\dagger = X^{-1}$. Likewise, the extreme directions of proper $\mathcal{K}^*_{\mathcal{M}+}$ are respectively orthogonal to the facets of $\mathcal{K}_{\mathcal{M}+}$ whose inward-normals are contained in the rows of $X^\dagger$ by (428). So the vertex-description for $\mathcal{K}^*_{\mathcal{M}+}$ employs the columns of $X^{\dagger\mathrm{T}}$. □

2.13.9.4.3 Example. *Monotone cone.* (Figure **69**, Figure **70**)
Full-dimensional but not pointed, the monotone cone is polyhedral and defined by the halfspace-description

$$\mathcal{K}_\mathcal{M} \triangleq \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \cdots \geq x_n\} = \{x \in \mathbb{R}^n \mid X^{*\mathrm{T}}x \succeq 0\} \tag{435}$$

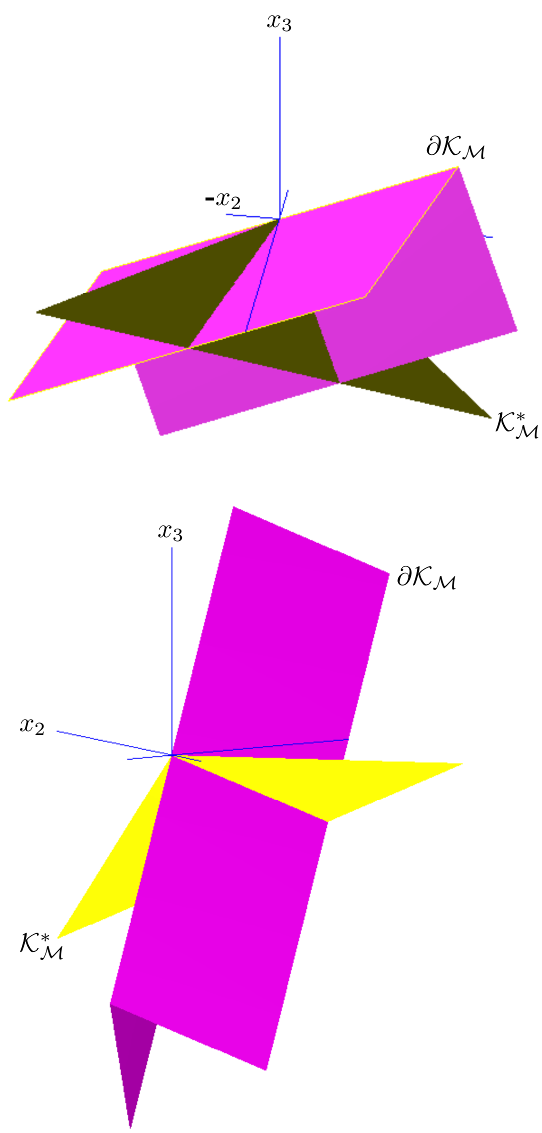


Figure 70: Two views of monotone cone $\mathcal{K}_\mathcal{M}$ and its dual $\mathcal{K}^*_\mathcal{M}$ (drawn truncated) in $\mathbb{R}^3$. Monotone cone is not pointed. Dual monotone cone is not full-dimensional. (Cartesian coordinate axes are drawn for reference.)

Its dual is therefore pointed but not full-dimensional;

$$\mathcal{K}_{\mathcal{M}}^* = \{\, X^* b \triangleq [\, e_1 - e_2 \;\; e_2 - e_3 \;\cdots\; e_{n-1} - e_n \,]\, b \mid b \succeq 0 \,\} \subset \mathbb{R}^n \tag{436}$$

the dual cone vertex-description where the columns of X^* comprise its extreme directions. Because dual monotone cone $\mathcal{K}_{\mathcal{M}}^*$ is pointed and satisfies

$$\operatorname{rank}(X^* \in \mathbb{R}^{n \times N}) \;=\; N \;\triangleq\; \dim \operatorname{aff} \mathcal{K}^* \;\leq\; n \tag{437}$$

where $N = n-1$, and because $\mathcal{K}_{\mathcal{M}}$ is closed and convex, we may adapt Cone Table **1** (p.166) as follows:

Cone Table 1*	$\mathcal{K}^*$	$\mathcal{K}^{**} = \mathcal{K}$
vertex-description	X^*	$X^{*\dagger\mathrm{T}}$, $\pm X^{*\perp}$
halfspace-description	$X^{*\dagger}$, $X^{*\perp\mathrm{T}}$	$X^{*\mathrm{T}}$

The vertex-description for $\mathcal{K}_{\mathcal{M}}$ is therefore

$$\mathcal{K}_{\mathcal{M}} = \{[\, X^{*\dagger\mathrm{T}} \quad X^{*\perp} \quad -X^{*\perp} \,]a \mid a \succeq 0\} \subset \mathbb{R}^n \tag{438}$$

where $X^{*\perp} = \mathbf{1}$ and

$$X^{*\dagger} = \frac{1}{n}\begin{bmatrix} n-1 & -1 & -1 & \cdots & -1 & -1 & -1 \\ n-2 & n-2 & -2 & \ddots & \cdots & -2 & -2 \\ \vdots & n-3 & n-3 & \ddots & -(n-4) & \vdots & -3 \\ 3 & \vdots & n-4 & \ddots & -(n-3) & -(n-3) & \vdots \\ 2 & 2 & \cdots & \ddots & 2 & -(n-2) & -(n-2) \\ 1 & 1 & 1 & \cdots & 1 & 1 & -(n-1) \end{bmatrix} \in \mathbb{R}^{n-1 \times n} \tag{439}$$

while

$$\mathcal{K}_{\mathcal{M}}^* = \{y \in \mathbb{R}^n \mid X^{*\dagger} y \succeq 0\,,\; X^{*\perp\mathrm{T}} y = \mathbf{0}\} \tag{440}$$

is the dual monotone cone halfspace-description. □

2.13.9.4.4 Exercise. *Inside the monotone cones.*
Mathematically describe the respective interior of the monotone nonnegative cone and monotone cone. In three dimensions, also describe the relative interior of each face. ▼

2.13.9.5 More pointed cone descriptions with equality condition

Consider pointed polyhedral cone $\mathcal{K}$ having a linearly independent set of generators and whose subspace membership is explicit; *id est,* we are given the ordinary halfspace-description

$$\mathcal{K} = \{x \mid Ax \succeq 0\,,\; Cx = \mathbf{0}\} \subseteq \mathbb{R}^n \tag{286a}$$

where $A\in\mathbb{R}^{m\times n}$ and $C\in\mathbb{R}^{p\times n}$. This can be equivalently written in terms of nullspace of C and vector ξ :

$$\mathcal{K} = \{Z\xi \in \mathbb{R}^n \mid AZ\xi \succeq 0\} \tag{441}$$

where $\mathcal{R}(Z\in\mathbb{R}^{n\times n-\operatorname{rank}C}) \triangleq \mathcal{N}(C)$. Assuming (411) is satisfied

$$\operatorname{rank} X \;\triangleq\; \operatorname{rank}\Big((AZ)^\dagger \in \mathbb{R}^{n-\operatorname{rank}C\times m}\Big) \;=\; m-\ell \;=\; \dim\operatorname{aff}\mathcal{K} \;\leq\; n-\operatorname{rank}C \tag{442}$$

where ℓ is the number of conically dependent rows in AZ which must be removed to make $\hat{A}Z$ before the Cone Tables become applicable.[2.81] Then results collected there admit assignment $\hat{X}\triangleq(\hat{A}Z)^\dagger\in\mathbb{R}^{n-\operatorname{rank}C\times m-\ell}$, where $\hat{A}\in\mathbb{R}^{m-\ell\times n}$, followed with linear transformation by Z. So we get the vertex-description, for full-rank $(\hat{A}Z)^\dagger$ skinny-or-square,

$$\mathcal{K} = \{Z(\hat{A}Z)^\dagger b \mid b \succeq 0\} \tag{443}$$

From this and (362) we get a halfspace-description of the dual cone

$$\mathcal{K}^* = \{y\in\mathbb{R}^n \mid (Z^{\mathrm{T}}\hat{A}^{\mathrm{T}})^\dagger Z^{\mathrm{T}} y \succeq 0\} \tag{444}$$

From this and Cone Table **1** (p.166) we get a vertex-description, (1974)

$$\mathcal{K}^* = \{[\,Z^{\dagger\mathrm{T}}(\hat{A}Z)^{\mathrm{T}} \quad C^{\mathrm{T}} \quad -C^{\mathrm{T}}\,]c \mid c \succeq 0\} \tag{445}$$

Yet because

$$\mathcal{K} = \{x \mid Ax \succeq 0\} \cap \{x \mid Cx = \mathbf{0}\} \tag{446}$$

then, by (313), we get an equivalent vertex-description for the dual cone

$$\begin{aligned}\mathcal{K}^* &= \overline{\{x \mid Ax \succeq 0\}^* + \{x \mid Cx = \mathbf{0}\}^*}\\ &= \{[\,A^{\mathrm{T}} \quad C^{\mathrm{T}} \quad -C^{\mathrm{T}}\,]b \mid b \succeq 0\}\end{aligned} \tag{447}$$

from which the conically dependent columns may, of course, be removed.

2.13.10 Dual cone-translate

(§E.10.3.2.1) First-order optimality condition (351) inspires a dual-cone variant: For any set $\mathcal{K}$, the negative dual of its translation by any $a\in\mathbb{R}^n$ is

$$\begin{aligned}-(\mathcal{K}-a)^* \;&=\; \{y\in\mathbb{R}^n \mid \langle y\,,\,x-a\rangle \leq 0 \text{ for all } x\in\mathcal{K}\} \;\triangleq\; \mathcal{K}^\perp(a)\\ &=\; \{y\in\mathbb{R}^n \mid \langle y\,,\,x\rangle \leq 0 \text{ for all } x\in\mathcal{K}-a\}\end{aligned} \tag{448}$$

a closed convex cone called *normal cone* to $\mathcal{K}$ at point a. From this, a new membership relation like (319):

$$y \in -(\mathcal{K}-a)^* \;\Leftrightarrow\; \langle y\,,\,x-a\rangle \leq 0 \text{ for all } x \in \mathcal{K} \tag{449}$$

and by closure the conjugate, for closed convex cone $\mathcal{K}$

$$x \in \mathcal{K} \;\Leftrightarrow\; \langle y\,,\,x-a\rangle \leq 0 \text{ for all } y \in -(\mathcal{K}-a)^* \tag{450}$$

[2.81] When the conically dependent rows (§2.10) are removed, the rows remaining must be linearly independent for the Cone Tables (p.19) to apply.

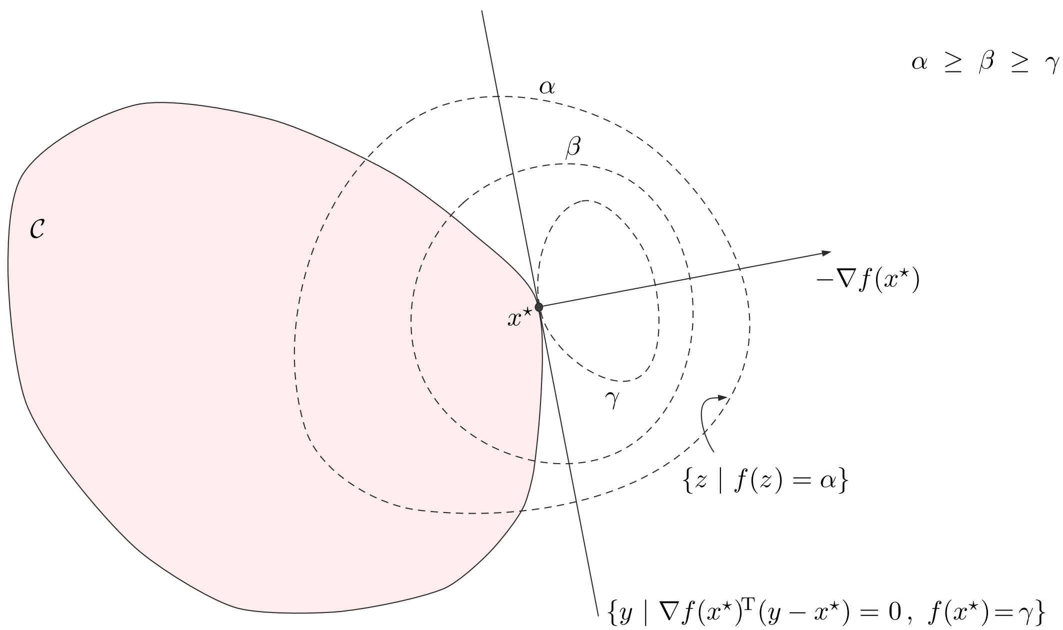


Figure 71: (*confer* Figure **82**) Shown is a plausible contour plot in $\mathbb{R}^2$ of some arbitrary differentiable convex real function $f(x)$ at selected levels α, β, and γ; *id est*, contours of equal level f (*level sets*) drawn dashed in function's domain. From results in §3.6.2 (p.218), gradient $\nabla f(x^\star)$ is normal to γ-*sublevel set* $\mathcal{L}_\gamma f$ (560) by Definition E.9.1.0.1. From §2.13.10.1, function is minimized over convex set $\mathcal{C}$ at point $x^\star$ iff negative gradient $-\nabla f(x^\star)$ belongs to normal cone to $\mathcal{C}$ there. In circumstance depicted, normal cone is a ray whose direction is coincident with negative gradient. So, gradient is normal to a hyperplane supporting both $\mathcal{C}$ and the γ-sublevel set.

2.13.10.1 first-order optimality condition - restatement

(*confer* §2.13.3) The general first-order necessary and sufficient condition for optimality of solution $x^\star$ to a minimization problem with real differentiable convex objective function $f(x):\mathbb{R}^n\to\mathbb{R}$ over convex feasible set $\mathcal{C}$ is [324, §3]

$$-\nabla f(x^\star)\in-(\mathcal{C}-x^\star)^*\,,\qquad x^\star\in\mathcal{C}\tag{451}$$

id est, the negative gradient (§3.6) belongs to the normal cone to $\mathcal{C}$ at $x^\star$ as in Figure **71**.

2.13.10.1.1 Example. *Normal cone to orthant.*
Consider proper cone $\mathcal{K}=\mathbb{R}^n_+$, the selfdual nonnegative orthant in $\mathbb{R}^n$. The normal cone to $\mathbb{R}^n_+$ at $a\in\mathcal{K}$ is (2208)

$$\mathcal{K}^\perp_{\mathbb{R}^n_+}(a\in\mathbb{R}^n_+)=-(\mathbb{R}^n_+-a)^*=-\mathbb{R}^n_+\cap a^\perp\;,\quad a\in\mathbb{R}^n_+\tag{452}$$

where $-\mathbb{R}^n_+=-\mathcal{K}^*$ is the algebraic complement of $\mathbb{R}^n_+$, and $a^\perp$ is the orthogonal complement to range of vector a. This means: When point a is interior to $\mathbb{R}^n_+$, the normal cone is the origin. If $n_{\rm p}$ represents number of nonzero entries in vector $a\in\partial\mathbb{R}^n_+$, then $\dim(-\mathbb{R}^n_+\cap a^\perp)=n-n_{\rm p}$ and there is a complementary relationship between the nonzero entries in vector a and the nonzero entries in any vector $x\in-\mathbb{R}^n_+\cap a^\perp$. □

2.13.10.1.2 Example. *Optimality conditions for conic problem.*
Consider a convex optimization problem having real differentiable convex objective function $f(x):\mathbb{R}^n\to\mathbb{R}$ defined on domain $\mathbb{R}^n$

$$\begin{array}{cl}\underset{x}{\text{minimize}} & f(x)\\ \text{subject to} & x\in\mathcal{K}\end{array}\tag{453}$$

Let's first suppose that the feasible set is a pointed polyhedral cone $\mathcal{K}$ possessing a linearly independent set of generators and whose subspace membership is made explicit by fat full-rank matrix $C\in\mathbb{R}^{p\times n}$; *id est*, we are given the halfspace-description, for $A\in\mathbb{R}^{m\times n}$

$$\mathcal{K}\;=\;\{x\mid Ax\succeq0\,,\;Cx=\mathbf{0}\}\subseteq\mathbb{R}^n\tag{286a}$$

(We'll generalize to any convex cone $\mathcal{K}$ shortly.) Vertex-description of this cone, assuming $(\hat{A}Z)^\dagger$ skinny-or-square full-rank, is

$$\mathcal{K}=\{Z(\hat{A}Z)^\dagger b\mid b\succeq0\}\tag{443}$$

where $\hat{A}\in\mathbb{R}^{m-\ell\times n}$, ℓ is the number of conically dependent rows in AZ (§2.10) which must be removed, and $Z\in\mathbb{R}^{n\times n-\operatorname{rank}C}$ holds basis $\mathcal{N}(C)$ columnar.

From optimality condition (351),

$$\nabla f(x^\star)^{\rm T}\big(Z(\hat{A}Z)^\dagger b-x^\star\big)\geq0\quad\forall\,b\succeq0\tag{454}$$

$$-\nabla f(x^\star)^{\rm T}Z(\hat{A}Z)^\dagger(b-b^\star)\leq0\quad\forall\,b\succeq0\tag{455}$$

because

$$x^\star \triangleq Z(\hat{A}Z)^\dagger b^\star \in \mathcal{K} \tag{456}$$

From membership relation (449) and Example 2.13.10.1.1

$$\begin{array}{c}\langle -(Z^{\mathrm T}\hat{A}^{\mathrm T})^\dagger Z^{\mathrm T}\nabla f(x^\star)\,,\, b-b^\star\rangle \leq 0 \;\text{ for all }\; b\in\mathbb{R}_+^{m-\ell}\\ \Leftrightarrow\\ -(Z^{\mathrm T}\hat{A}^{\mathrm T})^\dagger Z^{\mathrm T}\nabla f(x^\star)\in -\mathbb{R}_+^{m-\ell}\cap b^{\star\perp}\end{array} \tag{457}$$

Then equivalent necessary and sufficient conditions for optimality of conic problem (453) with feasible set $\mathcal{K}$ are: (*confer* (361))

$$(Z^{\mathrm T}\hat{A}^{\mathrm T})^\dagger Z^{\mathrm T}\nabla f(x^\star) \underset{\mathbb{R}_+^{m-\ell}}{\succeq} 0\,,\qquad b^\star \underset{\mathbb{R}_+^{m-\ell}}{\succeq} 0\,,\qquad \nabla f(x^\star)^{\mathrm T} Z(\hat{A}Z)^\dagger b^\star = 0 \tag{458}$$

expressible, by (444),

$$\nabla f(x^\star)\in\mathcal{K}^*,\qquad x^\star\in\mathcal{K}\,,\qquad \nabla f(x^\star)^{\mathrm T}x^\star = 0 \tag{459}$$

This result (459) actually applies more generally to any convex cone $\mathcal{K}$ comprising the feasible set: Necessary and sufficient optimality conditions are in terms of objective gradient

$$-\nabla f(x^\star)\in -(\mathcal{K}-x^\star)^*\,,\qquad x^\star\in\mathcal{K} \qquad (451)$$

whose membership to normal cone, assuming only cone $\mathcal{K}$ convexity,

$$-(\mathcal{K}-x^\star)^* \;=\; \mathcal{K}^\perp_{\mathcal{K}}(x^\star\in\mathcal{K}) \;=\; -\mathcal{K}^*\cap x^{\star\perp} \qquad (2208)$$

equivalently expresses conditions (459).

When $\mathcal{K}=\mathbb{R}_+^n$, in particular, then $C=\mathbf{0}$, $A=Z=I\in\mathbb{S}^n$; *id est*,

$$\begin{array}{cl}\underset{x}{\text{minimize}} & f(x)\\ \text{subject to} & x \underset{\mathbb{R}_+^n}{\succeq} 0\end{array} \tag{460}$$

Necessary and sufficient optimality conditions become (*confer* [63, §4.2.3])

$$\nabla f(x^\star) \underset{\mathbb{R}_+^n}{\succeq} 0\,,\qquad x^\star \underset{\mathbb{R}_+^n}{\succeq} 0\,,\qquad \nabla f(x^\star)^{\mathrm T}x^\star = 0 \tag{461}$$

equivalent to condition (329)[2.82] (under nonzero gradient) for membership to the nonnegative orthant boundary $\partial\mathbb{R}_+^n$. □

2.13.10.1.3 Example. *Complementarity problem.* [226]

A complementarity problem in nonlinear function f is nonconvex

$$\begin{array}{rl}\text{find} & z\in\mathcal{K}\\ \text{subject to} & f(z)\in\mathcal{K}^*\\ & \langle z\,,\, f(z)\rangle = 0\end{array} \tag{462}$$

[2.82] and equivalent to well-known Karush-Kuhn-Tucker (KKT) optimality conditions [63, §5.5.3] because the dual variable becomes gradient $\nabla f(x)$.

yet bears strong resemblance to (459) and to Moreau's decomposition (2144) on page 639 for projection P on mutually polar cones $\mathcal{K}$ and $-\mathcal{K}^*$. Identify a sum of mutually orthogonal projections $x \triangleq z - f(z)$; in Moreau's terms, $z = P_{\mathcal{K}}x$ and $-f(z) = P_{-\mathcal{K}^*}x$. Then $f(z) \in \mathcal{K}^*$ (§E.9.2.2 *no.*4) and z is a solution to the complementarity problem iff it is a fixed point of

$$z = P_{\mathcal{K}}x = P_{\mathcal{K}}\big(z - f(z)\big) \tag{463}$$

Given that a solution exists, existence of a fixed point would be guaranteed by theory of *contraction.* [243, p.300] But because only *nonexpansivity* (Theorem E.9.3.0.1) is achievable by a projector, uniqueness cannot be assured. [219, p.155] Elegant proofs of equivalence between complementarity problem (462) and fixed point problem (463) are provided by Németh [398, *Fixed point problems*]. □

2.13.10.1.4 Example. *Linear complementarity problem.* [91] [288] [329]
Given matrix $B \in \mathbb{R}^{n\times n}$ and vector $q \in \mathbb{R}^n$, a prototypical complementarity problem on the nonnegative orthant $\mathcal{K} = \mathbb{R}^n_+$ is linear in $w = f(z)$:

$$\begin{array}{rl} \text{find} & z \succeq 0 \\ \text{subject to} & w \succeq 0 \\ & w^{\mathrm T} z = 0 \\ & w = q + Bz \end{array} \tag{464}$$

This problem is not convex when both vectors w and z are variable.[2.83] Notwithstanding, this linear complementarity problem can be solved by identifying $w \leftarrow \nabla f(z) = q + Bz$ then substituting that gradient into (462)

$$\begin{array}{rl} \text{find} & z \in \mathcal{K} \\ \text{subject to} & \nabla f(z) \in \mathcal{K}^* \\ & \langle z\,,\, \nabla f(z)\rangle = 0 \end{array} \tag{465}$$

which is simply a restatement of optimality conditions (459) for conic problem (453). Suitable $f(z)$ is the quadratic objective from convex problem

$$\begin{array}{rl} \underset{z}{\text{minimize}} & \frac{1}{2} z^{\mathrm T} B z + q^{\mathrm T} z \\ \text{subject to} & z \succeq 0 \end{array} \tag{466}$$

which means $B \in \mathbb{S}^n_+$ should be (symmetric) positive semidefinite for solution of (464) by this method. Then (464) has solution iff (466) does. □

[2.83] But if one of them is fixed, then the problem becomes convex with a very simple geometric interpretation: Define the affine subset

$$\mathcal{A} \triangleq \{y \in \mathbb{R}^n \mid By = w - q\}$$

For $w^{\mathrm T} z$ to vanish, there must be a complementary relationship between the nonzero entries of vectors w and z; *id est*, $w_i z_i = 0\ \forall i$. Given $w \succeq 0$, then z belongs to the convex set of solutions:

$$z \in -\mathcal{K}^{\perp}_{\mathbb{R}^n_+}(w \in \mathbb{R}^n_+) \cap \mathcal{A} = \mathbb{R}^n_+ \cap w^{\perp} \cap \mathcal{A}$$

where $\mathcal{K}^{\perp}_{\mathbb{R}^n_+}(w)$ is the normal cone to $\mathbb{R}^n_+$ at w (452). If this intersection is nonempty, then the problem is solvable.

2.13.10.1.5 Exercise. *Optimality for equality constrained conic problem.*
Consider a conic optimization problem like (453) having real differentiable convex objective function $f(x):\mathbb{R}^n\to\mathbb{R}$

$$\begin{array}{cl}\underset{x}{\text{minimize}} & f(x)\\ \text{subject to} & Cx=d\\ & x\in\mathcal{K}\end{array}\tag{467}$$

minimized over convex cone $\mathcal{K}$ but, this time, constrained to affine set $\mathcal{A}=\{x\mid Cx=d\}$. Show, by means of first-order optimality condition (351) or (451), that necessary and sufficient optimality conditions are: (*confer* (459))

$$\begin{array}{r}x^\star\in\mathcal{K}\\ Cx^\star=d\\ \nabla f(x^\star)+C^{\mathrm{T}}\nu^\star\in\mathcal{K}^*\\ \langle\nabla f(x^\star)+C^{\mathrm{T}}\nu^\star,\,x^\star\rangle=0\end{array}\tag{468}$$

where $\nu^\star$ is any vector[2.84] satisfying these conditions. ▼

2.13.11 Proper nonsimplicial $\mathcal{K}$, dual, X fat full-rank

Since conically dependent columns can always be removed from X to construct $\mathcal{K}$ or to determine $\mathcal{K}^*$ [391], then assume we are given a set of N conically independent generators (§2.10) of an arbitrary proper polyhedral cone $\mathcal{K}$ in $\mathbb{R}^n$ arranged columnar in $X\in\mathbb{R}^{n\times N}$ such that $N>n$ (fat) and $\operatorname{rank}X=n$. Having found formula (419) to determine the dual of a simplicial cone, the easiest way to find a vertex-description of proper dual cone $\mathcal{K}^*$ is to first decompose $\mathcal{K}$ into simplicial parts $\mathcal{K}_i$ so that $\mathcal{K}=\bigcup\mathcal{K}_i$.[2.85] Each component simplicial cone in $\mathcal{K}$ corresponds to some subset of n linearly independent columns from X. The key idea, here, is how the extreme directions of the simplicial parts must remain extreme directions of $\mathcal{K}$. Finding the dual of $\mathcal{K}$ amounts to finding the dual of each simplicial part:

2.13.11.0.1 Theorem. *Dual cone intersection.* [347, §2.7]
Suppose proper cone $\mathcal{K}\subset\mathbb{R}^n$ equals the union of M simplicial cones $\mathcal{K}_i$ whose extreme directions all coincide with those of $\mathcal{K}$. Then proper dual cone $\mathcal{K}^*$ is the intersection of M dual simplicial cones $\mathcal{K}_i^*$; *id est*,

$$\mathcal{K}=\bigcup_{i=1}^{M}\mathcal{K}_i\quad\Rightarrow\quad\mathcal{K}^*=\bigcap_{i=1}^{M}\mathcal{K}_i^*\tag{469}$$

⋄

[2.84] an optimal dual variable, these optimality conditions are equivalent to the KKT conditions [63, §5.5.3].

[2.85] That proposition presupposes, of course, that we know how to perform simplicial decomposition efficiently; also called "triangulation". [321] [189, §3.1] [190, §3.1] Existence of multiple simplicial parts means expansion of $x\in\mathcal{K}$, like (410), can no longer be unique because number N of extreme directions in $\mathcal{K}$ exceeds dimension n of the space.

Proof. For $X_i \in \mathbb{R}^{n\times n}$, a complete matrix of linearly independent extreme directions (p.125) arranged columnar, corresponding simplicial $\mathcal{K}_i$ (§2.12.3.1.1) has vertex-description

$$\mathcal{K}_i = \{X_i c \mid c \succeq 0\} \tag{470}$$

Now suppose,

$$\mathcal{K} \;=\; \bigcup_{i=1}^{M} \mathcal{K}_i \;=\; \bigcup_{i=1}^{M} \{X_i c \mid c \succeq 0\} \tag{471}$$

The union of all $\mathcal{K}_i$ can be equivalently expressed

$$\mathcal{K} \;=\; \left\{ [\,X_1\; X_2 \cdots X_M\,] \begin{bmatrix} a \\ b \\ \vdots \\ c \end{bmatrix} \;\mid\; a\,, b \ldots c \succeq 0 \right\} \tag{472}$$

Because extreme directions of the simplices $\mathcal{K}_i$ are extreme directions of $\mathcal{K}$ by assumption, then

$$\mathcal{K} \;=\; \{\,[\,X_1\; X_2 \cdots X_M\,]\,d \mid d \succeq 0\,\} \tag{473}$$

by the *extremes theorem* (§2.8.1.1.1). Defining $X \triangleq [\,X_1\; X_2 \cdots X_M\,]$ (with any redundant [*sic*] columns optionally removed from X), then $\mathcal{K}^*$ can be expressed ((362), Cone Table **S** p.167)

$$\mathcal{K}^* \;=\; \{y \mid X^{\mathrm{T}} y \succeq 0\} \;=\; \bigcap_{i=1}^{M} \{y \mid X_i^{\mathrm{T}} y \succeq 0\} \;=\; \bigcap_{i=1}^{M} \mathcal{K}_i^* \tag{474}$$

♦

To find the extreme directions of the dual cone, first we observe that some facets of each simplicial part $\mathcal{K}_i$ are common to facets of $\mathcal{K}$ by assumption, and the union of all those common facets comprises the set of all facets of $\mathcal{K}$ by design. For any particular proper polyhedral cone $\mathcal{K}$, the extreme directions of dual cone $\mathcal{K}^*$ are respectively orthogonal to the facets of $\mathcal{K}$. (§2.13.6.1) Then the extreme directions of the dual cone can be found among inward-normals to facets of the component simplicial cones $\mathcal{K}_i$; those normals are extreme directions of the dual simplicial cones $\mathcal{K}_i^*$. From the theorem and Cone Table **S** (p.167),

$$\mathcal{K}^* \;=\; \bigcap_{i=1}^{M} \mathcal{K}_i^* \;=\; \bigcap_{i=1}^{M} \{X_i^{\dagger\mathrm{T}} c \mid c \succeq 0\} \tag{475}$$

The set of extreme directions $\{\Gamma_i^*\}$ for proper dual cone $\mathcal{K}^*$ is therefore constituted by those conically independent generators, from the columns of all the dual simplicial matrices $\{X_i^{\dagger\mathrm{T}}\}$, that do not violate discrete definition (362) of $\mathcal{K}^*$;

$$\{\Gamma_1^*\,, \Gamma_2^* \ldots \Gamma_N^*\} \;=\; \text{c.i.}\Big\{X_i^{\dagger\mathrm{T}}(:,j)\,,\; i=1\ldots M\,,\; j=1\ldots n \mid X_i^{\dagger}(j,:)\Gamma_\ell \geq 0,\; \ell=1\ldots N\Big\} \tag{476}$$

where c.i. denotes selection of only the conically independent vectors from the argument set, argument $(:,j)$ denotes the j^{th} column while $(j,:)$ denotes the j^{th} row, and $\{\Gamma_\ell\}$ constitutes the extreme directions of $\mathcal{K}$. Figure **53**b (p.124) shows a cone and its dual found via this algorithm.

2.13.11.0.2 Example. *Dual of $\mathcal{K}$ nonsimplicial in subspace* $\operatorname{aff}\mathcal{K}$ *.*
Given conically independent generators for pointed closed convex cone $\mathcal{K}$ in $\mathbb{R}^4$ arranged columnar in

$$X = [\,\Gamma_1\ \ \Gamma_2\ \ \Gamma_3\ \ \Gamma_4\,] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{bmatrix} \tag{477}$$

having $\dim\operatorname{aff}\mathcal{K} = \operatorname{rank}X = 3$, (281) then performing the most inefficient simplicial decomposition in $\operatorname{aff}\mathcal{K}$ we find

$$\begin{array}{ll} X_1 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, & X_2 = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \\ X_3 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}, & X_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix} \end{array} \tag{478}$$

The corresponding dual simplicial cones in $\operatorname{aff}\mathcal{K}$ have generators respectively columnar in

$$\begin{array}{ll} 4X_1^{\dagger\mathrm{T}} = \begin{bmatrix} 2 & 1 & 1 \\ -2 & 1 & 1 \\ 2 & -3 & 1 \\ -2 & 1 & -3 \end{bmatrix}, & 4X_2^{\dagger\mathrm{T}} = \begin{bmatrix} 1 & 2 & 1 \\ -3 & 2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & -3 \end{bmatrix} \\ 4X_3^{\dagger\mathrm{T}} = \begin{bmatrix} 3 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & -2 & 3 \\ -1 & -2 & -1 \end{bmatrix}, & 4X_4^{\dagger\mathrm{T}} = \begin{bmatrix} 3 & -1 & 2 \\ -1 & 3 & -2 \\ -1 & -1 & 2 \\ -1 & -1 & -2 \end{bmatrix} \end{array} \tag{479}$$

Applying algorithm (476) we get

$$[\,\Gamma_1^*\ \ \Gamma_2^*\ \ \Gamma_3^*\ \ \Gamma_4^*\,] = \frac{1}{4}\begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 2 & -1 & -2 \\ 1 & -2 & -1 & 2 \\ -3 & -2 & -1 & -2 \end{bmatrix} \tag{480}$$

whose rank is 3, and is the known result;[2.86] a conically independent set of generators for that pointed section of the dual cone $\mathcal{K}^*$ in $\operatorname{aff}\mathcal{K}$; *id est*, $\mathcal{K}^* \cap \operatorname{aff}\mathcal{K}$. □

[2.86]These calculations proceed so as to be consistent with [122, §6]; as if the ambient vector space were proper subspace $\operatorname{aff}\mathcal{K}$ whose dimension is 3. In that ambient space, $\mathcal{K}$ may be regarded as a proper cone. Yet that author (from the citation) erroneously states dimension of the ordinary dual cone to be 3 ; it is, in fact, 4.

2.13.11.0.3 Example. *Dual of proper polyhedral $\mathcal{K}$ in $\mathbb{R}^4$.*
Given conically independent generators for a full-dimensional pointed closed convex cone $\mathcal{K}$

$$X = [\,\Gamma_1 \;\; \Gamma_2 \;\; \Gamma_3 \;\; \Gamma_4 \;\; \Gamma_5\,] = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & -1 & 0 \end{bmatrix} \tag{481}$$

we count $5!/\big((5-4)!\,4!\big)=5$ component simplices.[2.87] Applying algorithm (476), we find the six extreme directions of dual cone $\mathcal{K}^*$ (with $\Gamma_2 = \Gamma_5^*$)

$$X^* = [\,\Gamma_1^* \;\; \Gamma_2^* \;\; \Gamma_3^* \;\; \Gamma_4^* \;\; \Gamma_5^* \;\; \Gamma_6^*\,] = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 1 & 0 & 0 \end{bmatrix} \tag{482}$$

which means, (§2.13.6.1) this proper polyhedral $\mathcal{K} = \operatorname{cone}(X)$ has six (three-dimensional) facets generated $\mathcal{G}$ by its {extreme directions}:

$$\mathcal{G}\left\{\begin{matrix} \mathcal{F}_1 \\ \mathcal{F}_2 \\ \mathcal{F}_3 \\ \mathcal{F}_4 \\ \mathcal{F}_5 \\ \mathcal{F}_6 \end{matrix}\right\} = \left\{\begin{matrix} \Gamma_1 & \Gamma_2 & \Gamma_3 & & \\ \Gamma_1 & \Gamma_2 & & & \Gamma_5 \\ \Gamma_1 & & & \Gamma_4 & \Gamma_5 \\ \Gamma_1 & & \Gamma_3 & \Gamma_4 & \\ & & \Gamma_3 & \Gamma_4 & \Gamma_5 \\ & \Gamma_2 & \Gamma_3 & & \Gamma_5 \end{matrix}\right\} \tag{483}$$

whereas dual proper polyhedral cone $\mathcal{K}^*$ has only five:

$$\mathcal{G}\left\{\begin{matrix} \mathcal{F}_1^* \\ \mathcal{F}_2^* \\ \mathcal{F}_3^* \\ \mathcal{F}_4^* \\ \mathcal{F}_5^* \end{matrix}\right\} = \left\{\begin{matrix} \Gamma_1^* & \Gamma_2^* & \Gamma_3^* & \Gamma_4^* & & \\ \Gamma_1^* & \Gamma_2^* & & & & \Gamma_6^* \\ \Gamma_1^* & & & \Gamma_4^* & \Gamma_5^* & \Gamma_6^* \\ & & \Gamma_3^* & \Gamma_4^* & \Gamma_5^* & \\ & \Gamma_2^* & \Gamma_3^* & & \Gamma_5^* & \Gamma_6^* \end{matrix}\right\} \tag{484}$$

Six two-dimensional cones, having generators respectively $\{\Gamma_1^* \;\; \Gamma_3^*\}$ $\{\Gamma_2^* \;\; \Gamma_4^*\}$ $\{\Gamma_1^* \;\; \Gamma_5^*\}$ $\{\Gamma_4^* \;\; \Gamma_6^*\}$ $\{\Gamma_2^* \;\; \Gamma_5^*\}$ $\{\Gamma_3^* \;\; \Gamma_6^*\}$, are relatively interior to dual facets; so cannot be two-dimensional faces of $\mathcal{K}^*$ (by Definition 2.6.0.0.3).

We can check this result (482) by reversing the process; we find $6!/\big((6-4)!\,4!\big)-3=12$ component simplices in the dual cone.[2.88] Applying algorithm (476) to those simplices returns a conically independent set of generators for $\mathcal{K}$ equivalent to (481). □

2.13.11.0.4 Exercise. *Reaching proper polyhedral cone interior.*
Name two extreme directions Γ_i of cone $\mathcal{K}$ from Example 2.13.11.0.3 whose convex hull passes through that cone's interior. Explain why. Are there two such extreme directions of dual cone $\mathcal{K}^*$? ▼

[2.87] There are no linearly dependent combinations of three or four extreme directions in the primal cone.
[2.88] Three combinations of four dual extreme directions are linearly dependent; they belong to the dual facets. But there are no linearly dependent combinations of three dual extreme directions.

2.13.12 coordinates in proper nonsimplicial system

A natural question pertains to whether a theory of unique coordinates, like biorthogonal expansion w.r.t pointed closed convex $\mathcal{K}$, is extensible to proper cones whose extreme directions number in excess of ambient spatial dimensionality.

2.13.12.0.1 Theorem. *Conic coordinates.*
With respect to vector v in some finite-dimensional Euclidean space $\mathbb{R}^n$, define a coordinate $t_v^\star$ of point x in full-dimensional pointed closed convex cone $\mathcal{K}$

$$t_v^\star(x) \triangleq \sup\{t\in\mathbb{R} \mid x - t\,v\in\mathcal{K}\} \tag{485}$$

Given points x and y in cone $\mathcal{K}$, if $t_v^\star(x)=t_v^\star(y)$ for each and every extreme direction v of $\mathcal{K}$ then $x=y$. ⋄

Conic coordinate definition (485) acquires its heritage from conditions (375) for generator membership to a smallest face. Coordinate $t_v^\star(c)=0$, for example, corresponds to unbounded μ in (375); indicating, extreme direction v cannot belong to the smallest face of cone $\mathcal{K}$ that contains c.

2.13.12.0.2 Proof. Vector $x-t^\star v$ must belong to the cone boundary $\partial\mathcal{K}$ by definition (485). So there must exist a nonzero vector λ that is inward-normal to a hyperplane supporting cone $\mathcal{K}$ and containing $x-t^\star v$; *id est*, by boundary membership relation for full-dimensional pointed closed convex cones (§2.13.2)

$$x-t^\star v \in \partial\mathcal{K} \;\Leftrightarrow\; \exists\,\lambda\neq\mathbf{0} \ni \langle\lambda\,,\,x-t^\star v\rangle = 0\,,\quad \lambda\in\mathcal{K}^*,\quad x-t^\star v\in\mathcal{K} \tag{329}$$

where

$$\mathcal{K}^* = \{w\in\mathbb{R}^n \mid \langle v\,,\,w\rangle \geq 0 \text{ for all } v\in\mathcal{G}(\mathcal{K})\} \tag{368}$$

is the full-dimensional pointed closed convex dual cone. The set $\mathcal{G}(\mathcal{K})$, of possibly infinite cardinality N, comprises generators for cone $\mathcal{K}$; *e.g*, its extreme directions which constitute a minimal generating set. If $x-t^\star v$ is nonzero, any such vector λ must belong to the dual cone boundary by conjugate boundary membership relation

$$\lambda\in\partial\mathcal{K}^* \;\Leftrightarrow\; \exists\,x-t^\star v\neq\mathbf{0} \ni \langle\lambda\,,\,x-t^\star v\rangle = 0\,,\quad x-t^\star v\in\mathcal{K}\,,\quad \lambda\in\mathcal{K}^* \tag{330}$$

where

$$\mathcal{K} = \{z\in\mathbb{R}^n \mid \langle\lambda\,,\,z\rangle \geq 0 \text{ for all } \lambda\in\mathcal{G}(\mathcal{K}^*)\} \tag{367}$$

This description of $\mathcal{K}$ means: cone $\mathcal{K}$ is an intersection of halfspaces whose inward-normals are generators of the dual cone. Each and every face of cone $\mathcal{K}$ (except the cone itself) belongs to a hyperplane supporting $\mathcal{K}$. Each and every vector $x-t^\star v$ on the cone boundary must therefore be orthogonal to an extreme direction constituting generators $\mathcal{G}(\mathcal{K}^*)$ of the dual cone.

To the i^{th} extreme direction $v=\Gamma_i\in\mathbb{R}^n$ of cone $\mathcal{K}$, ascribe a coordinate $t_i^\star(x)\in\mathbb{R}$ of x from definition (485). On domain $\mathcal{K}$, the mapping

$$t^\star(x) = \begin{bmatrix} t_1^\star(x) \\ \vdots \\ t_N^\star(x) \end{bmatrix} : \mathbb{R}^n\to\mathbb{R}^N \tag{486}$$

has no nontrivial nullspace. Because $x - t^\star v$ must belong to $\partial\mathcal{K}$ by definition, the mapping $t^\star(x)$ is equivalent to a convex problem (separable in index i) whose objective (by (329)) is tightly bounded below by 0 :

$$\begin{array}{rll} t^\star(x) \;\equiv\; \arg\underset{t\in\mathbb{R}^N}{\text{minimize}} & \sum\limits_{i=1}^{N}\Gamma_{j(i)}^{*\mathrm{T}}(x - t_i\Gamma_i) & \\ \text{subject to} & x - t_i\Gamma_i \in \mathcal{K}\;, & i=1\ldots N \end{array} \tag{487}$$

where index $j\in\mathcal{I}$ is dependent on i and where (by (367)) $\lambda = \Gamma_j^* \in \mathbb{R}^n$ is an extreme direction of dual cone $\mathcal{K}^*$ that is normal to a hyperplane supporting $\mathcal{K}$ and containing $x - t_i^\star\Gamma_i$. Because extreme-direction cardinality N for cone $\mathcal{K}$ is not necessarily the same as for dual cone $\mathcal{K}^*$, index j must be judiciously selected from a set $\mathcal{I}$.

To prove injectivity when extreme-direction cardinality $N > n$ exceeds spatial dimension, we need only show mapping $t^\star(x)$ to be invertible; [139, thm.9.2.3] *id est*, x is recoverable given $t^\star(x)$:

$$\begin{array}{rll} x \;=\; \arg\underset{\tilde{x}\in\mathbb{R}^n}{\text{minimize}} & \sum\limits_{i=1}^{N}\Gamma_{j(i)}^{*\mathrm{T}}(\tilde{x} - t_i^\star\Gamma_i) & \\ \text{subject to} & \tilde{x} - t_i^\star\Gamma_i \in \mathcal{K}\;, & i=1\ldots N \end{array} \tag{488}$$

The feasible set of this nonseparable convex problem is an intersection of translated full-dimensional pointed closed convex cones $\bigcap_i \mathcal{K} + t_i^\star\Gamma_i$. The objective function's linear part describes movement in normal-direction $-\Gamma_j^*$ for each of N hyperplanes. The optimal point of hyperplane intersection is the unique solution x when $\{\Gamma_j^*\}$ comprises n linearly independent normals that come from the dual cone and make the objective vanish. Because the dual cone $\mathcal{K}^*$ is full-dimensional, pointed, closed, and convex by assumption, there exist N extreme directions $\{\Gamma_j^*\}$ from $\mathcal{K}^* \subset \mathbb{R}^n$ that span $\mathbb{R}^n$. So we need simply choose N spanning dual extreme directions that make the optimal objective vanish. Because such dual extreme directions preexist by (329), $t^\star(x)$ is invertible.

Otherwise, in the case $N \leq n$, $t^\star(x)$ holds coordinates for biorthogonal expansion. Reconstruction of x is therefore unique. ♦

2.13.12.1 reconstruction from conic coordinates

The foregoing proof of the *conic coordinates theorem* is not constructive; it establishes existence of dual extreme directions $\{\Gamma_j^*\}$ that will reconstruct a point x from its coordinates $t^\star(x)$ via (488), but does not prescribe the index set $\mathcal{I}$. There are at least two computational methods for specifying $\{\Gamma_{j(i)}^*\}$: one is combinatorial but sure to succeed, the other is a geometric method that searches for a minimum of a nonconvex function. We describe the latter:

Convex problem (P)

$$\text{(P)}\quad \begin{array}{ll} \underset{t\in\mathbb{R}}{\text{maximize}} & t \\ \text{subject to} & x - t\,v \in \mathcal{K} \end{array} \qquad\qquad \begin{array}{ll} \underset{\lambda\in\mathbb{R}^n}{\text{minimize}} & \lambda^{\mathrm{T}}x \\ \text{subject to} & \lambda^{\mathrm{T}}v = 1 \\ & \lambda \in \mathcal{K}^* \end{array} \quad\text{(D)} \tag{489}$$

is equivalent to definition (485) whereas convex problem (D) is its dual;[2.89] meaning, primal and dual optimal objectives are equal $t^\star = \lambda^{\star\mathrm{T}} x$ assuming Slater's condition (p.249) is satisfied. Under this assumption of strong duality, $\lambda^{\star\mathrm{T}}(x - t^\star v) = t^\star(1 - \lambda^{\star\mathrm{T}} v) = 0$; which implies, the primal problem is equivalent to

$$\begin{array}{cl} \underset{t\in\mathbb{R}}{\text{minimize}} & \lambda^{\star\mathrm{T}}(x - t\,v) \\ \text{subject to} & x - t\,v \in \mathcal{K} \end{array} \qquad (\mathrm{p}) \tag{490}$$

while the dual problem is equivalent to

$$\begin{array}{cl} \underset{\lambda\in\mathbb{R}^n}{\text{minimize}} & \lambda^{\mathrm{T}}(x - t^\star v) \\ \text{subject to} & \lambda^{\mathrm{T}} v = 1 \\ & \lambda \in \mathcal{K}^* \end{array} \qquad (\mathrm{d}) \tag{491}$$

Instead given coordinates $t^\star(x)$ and a description of cone $\mathcal{K}$, we propose inversion by alternating solution of respective primal and dual problems

$$\begin{array}{cll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & \sum\limits_{i=1}^{N} \Gamma_i^{*\mathrm{T}}(x - t_i^\star \Gamma_i) & \\ \text{subject to} & x - t_i^\star \Gamma_i \in \mathcal{K}\ , & i = 1 \ldots N \end{array} \tag{492}$$

$$\begin{array}{cll} \underset{\Gamma_i^*\in\mathbb{R}^n}{\text{minimize}} & \sum\limits_{i=1}^{N} \Gamma_i^{*\mathrm{T}}(x^\star - t_i^\star \Gamma_i) & \\ \text{subject to} & \Gamma_i^{*\mathrm{T}} \Gamma_i = 1\ , & i = 1 \ldots N \\ & \Gamma_i^* \in \mathcal{K}^*\ , & i = 1 \ldots N \end{array} \tag{493}$$

where dual extreme directions Γ_i^* are initialized arbitrarily and ultimately ascertained by the alternation. Convex problems (492) and (493) are iterated until convergence which is guaranteed by virtue of a monotonically nonincreasing real sequence of objective values. Convergence can be fast. The mapping $t^\star(x)$ is uniquely inverted when the necessarily nonnegative objective vanishes; *id est*, when $\Gamma_i^{*\mathrm{T}}(x^\star - t_i^\star \Gamma_i) = 0 \;\; \forall\, i$. Here, a zero objective can occur only at the true solution x. But this global optimality condition cannot be guaranteed by the alternation because the common objective function, when regarded in both primal x and dual Γ_i^* variables simultaneously, is generally neither quasiconvex or monotonic. (§3.8.0.0.3)

Conversely, a nonzero objective at convergence is a certificate that inversion was not performed properly. A nonzero objective indicates that a global minimum of a multimodal objective function could not be found by this alternation. That is a flaw in this particular iterative algorithm for inversion; not in theory.[2.90] A numerical remedy is to reinitialize the Γ_i^* to different values.

[2.89] Form a Lagrangian associated with primal problem (P):

$$\begin{array}{ll} \mathfrak{L}(t\,,\lambda) = t + \lambda^{\mathrm{T}}(x - t\,v) = \lambda^{\mathrm{T}} x + t(1 - \lambda^{\mathrm{T}} v)\ , & \lambda \underset{\mathcal{K}^*}{\succeq} 0 \\ \sup\limits_t \mathfrak{L}(t\,,\lambda) = \lambda^{\mathrm{T}} x\ , & 1 - \lambda^{\mathrm{T}} v = 0 \end{array}$$

Dual variable (Lagrange multiplier [266, p.216]) λ generally has a nonnegative sense $\succeq$ for primal maximization with any cone membership constraint, whereas λ would have a nonpositive sense $\preceq$ were the primal instead a minimization problem with a cone membership constraint.

[2.90] The Proof 2.13.12.0.2, that suitable dual extreme directions $\{\Gamma_j^*\}$ always exist, means that a global optimization algorithm would always find the zero objective of alternation (492) (493); hence, the unique inversion x. But such an algorithm can be combinatorial.