Chapter 7

Proximity problems

In the "extremely large-scale case" (N *of order of tens and hundreds of thousands*), [iteration cost $O(N^3)$] *rules out all advanced convex optimization techniques, including all known polynomial time algorithms.*

−Arkadi Nemirovski, 2004 ⇊

A problem common to various sciences is to find the Euclidean distance matrix (EDM) $D \in \mathbb{EDM}^N$ closest in some sense to a given complete matrix of measurements H under a constraint on affine dimension $0 \leq r \leq N-1$ (§2.3.1, §5.7.1.1); rather, r is bounded above by desired affine dimension ρ.

7.0.1 Measurement matrix H

Ideally, we want a given matrix of measurements $H \in \mathbb{R}^{N \times N}$ to conform with the first three Euclidean metric properties (§5.2); to belong to the intersection of the orthant of nonnegative matrices $\mathbb{R}_+^{N \times N}$ with the symmetric hollow subspace $\mathbb{S}_h^N$ (§2.2.3.0.1). Geometrically, we want H to belong to the polyhedral cone (§2.12.1.0.1)

$$\mathcal{K} \triangleq \mathbb{S}_h^N \cap \mathbb{R}_+^{N \times N} \tag{1463}$$

Yet in practice, H can possess significant measurement uncertainty (noise).

Sometimes realization of an optimization problem demands that its input, the given matrix H, possess some particular characteristics; perhaps symmetry and hollowness or nonnegativity. When that H given does not have the desired properties, then we must impose them upon H prior to optimization:

- When *measurement matrix* H is neither symmetric or hollow, taking its symmetric hollow part is equivalent to orthogonal projection on the symmetric hollow subspace $\mathbb{S}_h^N$.

- When measurements of distance in H are negative, zeroing negative entries effects unique minimum-distance projection on the orthant of nonnegative matrices $\mathbb{R}_+^{N \times N}$ in isomorphic $\mathbb{R}^{N^2}$ (§E.9.2.2.3).

7.0.1.1 Order of imposition

Since convex cone $\mathcal{K}$ (1463) is the intersection of an orthant with a subspace, we want to project on that subset of the orthant belonging to the subspace; on the nonnegative orthant in the symmetric hollow subspace that is, in fact, the intersection. For that reason alone, unique minimum-distance projection of H on $\mathcal{K}$ (that member of $\mathcal{K}$ closest to H in isomorphic $\mathbb{R}^{N^2}$ in the Euclidean sense) can be attained by first taking its symmetric hollow part, and only then clipping negative entries of the result to 0; *id est*, there is only one correct *order of projection,* in general, on an orthant intersecting a subspace:

- project on the subspace, then project the result on the orthant in that subspace. (*confer* §E.9.5)

In contrast, order of projection on an intersection of subspaces is arbitrary.

That order of projection rule applies more generally, of course, to intersection of any convex set $\mathcal{C}$ with any subspace. Consider the *proximity problem*[7.1] over convex feasible set $\mathbb{S}_h^N \cap \mathcal{C}$ given nonsymmetric nonhollow $H \in \mathbb{R}^{N\times N}$:

$$\begin{array}{cl}\underset{B\in\mathbb{S}_h^N}{\text{minimize}} & \|B-H\|_{\mathrm{F}}^2 \\ \text{subject to} & B\in\mathcal{C}\end{array} \tag{1464}$$

a convex optimization problem. Because the symmetric hollow subspace $\mathbb{S}_h^N$ is orthogonal to the antisymmetric antihollow subspace $\mathbb{R}_h^{N\times N\perp}$ (§2.2.3), then for $B\in\mathbb{S}_h^N$

$$\operatorname{tr}\left(B^{\mathrm{T}}\left(\frac{1}{2}(H-H^{\mathrm{T}})+\delta^2(H)\right)\right)=0 \tag{1465}$$

so the objective function is equivalent to

$$\|B-H\|_{\mathrm{F}}^2 \equiv \left\|B-\left(\frac{1}{2}(H+H^{\mathrm{T}})-\delta^2(H)\right)\right\|_{\mathrm{F}}^2 + \left\|\frac{1}{2}(H-H^{\mathrm{T}})+\delta^2(H)\right\|_{\mathrm{F}}^2 \tag{1466}$$

This means the antisymmetric antihollow part of given matrix H would be ignored by minimization with respect to symmetric hollow variable B under Frobenius' norm; *id est*, minimization proceeds as though given the symmetric hollow part of H.

This action of Frobenius' norm (1466) is effectively a Euclidean projection (minimum-distance projection) of H on the symmetric hollow subspace $\mathbb{S}_h^N$ prior to minimization. Thus minimization proceeds inherently following the correct order for projection on $\mathbb{S}_h^N\cap\mathcal{C}$. Therefore we may either assume $H\in\mathbb{S}_h^N$, or take its symmetric hollow part prior to optimization.

7.0.1.2 Flagrant input error under nonnegativity demand

More pertinent to the optimization problems presented herein where

$$\mathcal{C}\triangleq\mathbb{EDM}^N\subseteq\mathcal{K} = \mathbb{S}_h^N\cap\mathbb{R}_+^{N\times N} \tag{1467}$$

then should some particular realization of a proximity problem demand input H be nonnegative, and were we only to zero negative entries of a nonsymmetric nonhollow input H prior to optimization, then the ensuing projection on $\mathbb{EDM}^N$ would be guaranteed incorrect (out of order).

[7.1] There are two equivalent interpretations of projection (§E.9): one finds a set normal, the other, minimum distance between a point and a set. Here we realize the latter view.

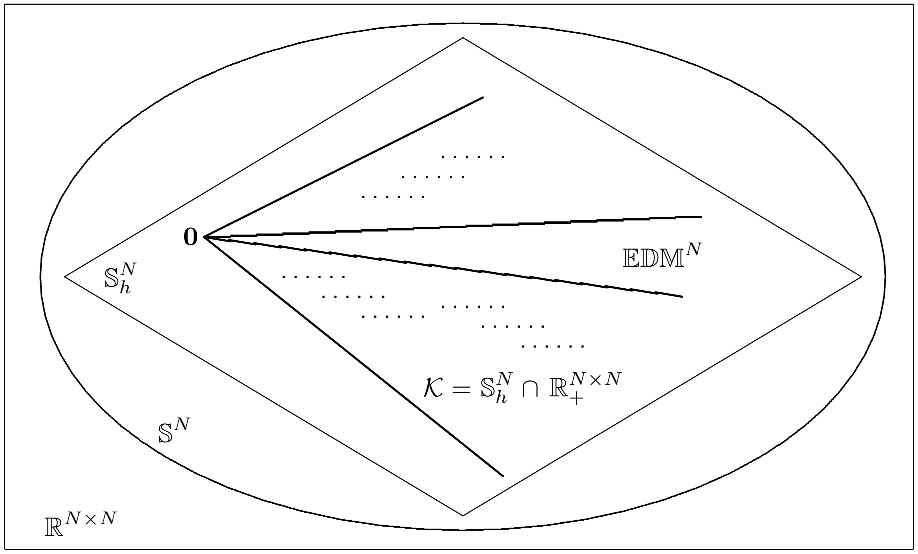

Figure 179: Pseudo-Venn diagram: EDM cone $\mathbb{EDM}^N$ belongs to intersection of symmetric hollow subspace with nonnegative orthant; $\mathbb{EDM}^N \subseteq \mathcal{K}$ (1048). $\mathbb{EDM}^N$ cannot exist outside of $\mathbb{S}_h^N$, but $\mathbb{R}_+^{N\times N}$ does.

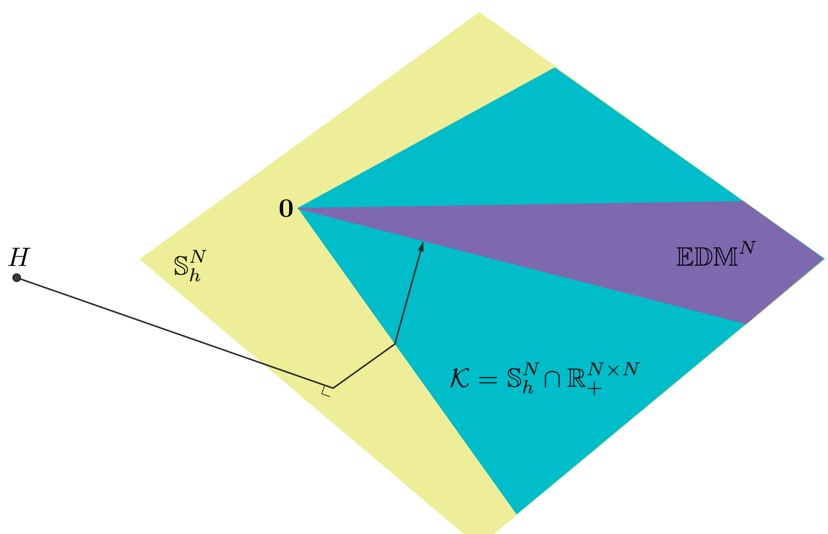

Figure 180: Pseudo-Venn diagram from Figure **179** showing elbow placed in path of projection of H on $\mathbb{EDM}^N \subset \mathbb{S}_h^N$ by an optimization problem demanding nonnegative input matrix H. The first two line segments, leading away from H, result from correct order of projection required to provide nonnegative H prior to optimization. Were H nonnegative, its projection on $\mathbb{S}_h^N$ would instead belong to $\mathcal{K}$; making the elbow disappear. (*confer* Figure **197**)

Now comes a surprising fact: Even were we to correctly follow the order of projection rule so as to provide $H \in \mathcal{K}$ prior to optimization, then the ensuing projection on $\mathbb{EDM}^N$ will be incorrect whenever input H has negative entries and some proximity problem demands nonnegative input H.

This is best understood referring to Figure **179**: Suppose nonnegative input H is demanded, and then the problem realization correctly projects its input first on $\mathbb{S}_h^N$ and then directly on $\mathcal{C} = \mathbb{EDM}^N$. That demand for nonnegativity effectively requires imposition of $\mathcal{K}$ on input H prior to optimization so as to obtain correct order of projection (on $\mathbb{S}_h^N$ first). Yet such an imposition prior to projection on $\mathbb{EDM}^N$ generally introduces an *elbow* into the path of projection (illustrated in Figure **180**) caused by the technique itself; that being, a particular proximity problem realization requiring nonnegative input.

Any procedure, for imposition of nonnegativity on input H, can only be incorrect in this circumstance. There is no resolution unless input H is guaranteed nonnegative with no tinkering. Otherwise, we have no choice but to employ a different problem realization; one not demanding nonnegative input.

7.0.2 Least lower bound

Most of the problems we encounter in this chapter have the general form:

$$\begin{array}{cl} \underset{B}{\text{minimize}} & \|B - A\|_{\mathrm{F}} \\ \text{subject to} & B \in \mathcal{C} \end{array} \tag{1468}$$

where $A \in \mathbb{R}^{m \times n}$ is given data. This particular objective denotes Euclidean projection (§E) of vectorized matrix A on the set $\mathcal{C}$ which may or may not be convex. When $\mathcal{C}$ is convex, then projection is unique minimum-distance because Frobenius' norm square is a strictly convex function of variable B and because the optimal solution is the same regardless of the square (524). When $\mathcal{C}$ is a subspace, then the direction of projection is orthogonal to $\mathcal{C}$.

Denoting by $A \triangleq U_A \Sigma_A Q_A^{\mathrm{T}}$ and $B \triangleq U_B \Sigma_B Q_B^{\mathrm{T}}$ their full singular value decompositions (whose singular values are always nonincreasingly ordered (§A.6)), there exists a tight lower bound on the objective over the manifold of orthogonal matrices;

$$\|\Sigma_B - \Sigma_A\|_{\mathrm{F}} \leq \inf_{U_A, U_B, Q_A, Q_B} \|B - A\|_{\mathrm{F}} \tag{1469}$$

This least lower bound holds more generally for any orthogonally invariant norm on $\mathbb{R}^{m \times n}$ (§2.2.1) including the Frobenius and spectral norm [370, §II.3]. [233, §7.4.51]

7.0.3 Problem approach.

stress/sstress problems traditionally posed in terms of point position $\{x_i \in \mathbb{R}^n,\ i = 1 \ldots N\}$

$$\underset{\{x_i\}}{\text{minimize}} \sum_{i,\, j \,\in\, \mathcal{I}} (\|x_i - x_j\| - h_{ij})^2 \tag{1470}$$

$$\underset{\{x_i\}}{\text{minimize}} \sum_{i,\, j \,\in\, \mathcal{I}} (\|x_i - x_j\|^2 - h_{ij})^2 \tag{1471}$$

(where $\mathcal{I}$ is an abstract set of indices and h_{ij} is given data) are everywhere converted herein to the distance-square variable D or to Gram matrix G; the Gram matrix acting as bridge between position and distance. (That conversion is performed regardless of whether known data is complete.) Then the techniques of chapter 5 or chapter 6 are applied to find relative or absolute position. This approach is taken because we prefer introduction of rank constraints into convex problems rather than searching a googol of local minima in nonconvex problems like (1471) or (1470) [118] (§3.9.0.0.3, §7.2.2.7.1).

7.0.4 Three prevalent proximity problems

There are three statements of the closest-EDM problem prevalent in the literature, the multiplicity due primarily to choice of projection on the EDM *versus* positive semidefinite (PSD) cone and vacillation between the distance-square variable d_{ij} *versus* absolute distance $\sqrt{d_{ij}}$. In their most fundamental form, the three prevalent proximity problems are (1472.1), (1472.2), and (1472.3): [386] for $D \triangleq [d_{ij}]$ and $\sqrt[\circ]{D} \triangleq [\sqrt{d_{ij}}\,]$

$$
\begin{array}{cll}
(1) & \begin{array}{cl}\underset{D}{\text{minimize}} & \|-V(D-H)V\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \leq \rho \\ & D \in \mathbb{EDM}^N\end{array} & \begin{array}{cl}\underset{\sqrt[\circ]{D}}{\text{minimize}} & \|\sqrt[\circ]{D}-H\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \leq \rho \\ & \sqrt[\circ]{D} \in \sqrt{\mathbb{EDM}^N}\end{array} \quad (2) \\
(3) & \begin{array}{cl}\underset{D}{\text{minimize}} & \|D-H\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \leq \rho \\ & D \in \mathbb{EDM}^N\end{array} & \begin{array}{cl}\underset{\sqrt[\circ]{D}}{\text{minimize}} & \|-V(\sqrt[\circ]{D}-H)V\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \leq \rho \\ & \sqrt[\circ]{D} \in \sqrt{\mathbb{EDM}^N}\end{array} \quad (4)
\end{array}
\tag{1472}
$$

where we have made explicit an imposed upper bound ρ on affine dimension

$$r = \operatorname{rank} V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} = \operatorname{rank} VDV \qquad (1203)$$

that is benign when $\rho = N-1$ or H were realizable with $r \leq \rho$. Problems (1472.2) and (1472.3) are Euclidean projections of vectorized matrix H on an EDM cone, whereas problems (1472.1) and (1472.4) are Euclidean projections of vectorized matrix $-VHV$ on a PSD cone.[7.2] (§6.3) Problem (1472.4) is not posed in the literature because it has limited theoretical foundation.[7.3]

Analytical solution to (1472.1) is known in closed form for any bound ρ and any auxiliary matrix V although, as the problem is stated, it is a convex optimization only in the case $\rho = N-1$. We show, in §7.1.4, how (1472.1) becomes a convex optimization problem for any ρ when transformed to the spectral domain. When expressed as a function of point list in a matrix X as in (1470), problem (1472.2) becomes a variant of what is known in statistics literature as a *stress problem.* [57, p.34] [116] [401] Problem (1472.3) is a rank-constrained *sstress problem,* whereas (1472.1) is equivalent to a rank-constrained *strain problem.* [117, §5][7.4] Problems (1472.2) and (1472.3) are convex optimization problems in D for the case $\rho = N-1$ wherein (1472.3) becomes equivalent to (1471). Even with the rank constraint removed from (1472.2), we will see that the convex problem remaining inherently minimizes affine dimension.

Generally speaking, each problem in (1472) produces a different result because there is no isometry relating them. Of the various auxiliary V-matrices (§B.4), the geometric centering matrix V (1071) appears in the literature most often although $V_{\mathcal{N}}$ (1055) is the auxiliary matrix naturally consequent to Schoenberg's seminal exposition [355]. Substitution of any auxiliary matrix or its pseudoinverse into these problems produces another valid problem.

Substitution of $V_{\mathcal{N}}$ for V in (1472.1), in particular, produces a different result because

$$
\begin{array}{cl}\underset{D}{\text{minimize}} & \|-V_{\mathcal{N}}^{\mathrm{T}}(D-H)V_{\mathcal{N}}\|_{\mathrm{F}}^2 \\ \text{subject to} & D \in \mathbb{EDM}^N\end{array}
\tag{1473}
$$

[7.2] Because $-VHV$ is orthogonal projection of $-H$ on the geometric center subspace $\mathbb{S}_c^N$ (§E.7.2.0.2), problems (1472.1) and (1472.4) may be interpreted as oblique (nonminimum distance) projections of $-H$ on a positive semidefinite cone.

[7.3] $D \in \mathbb{EDM}^N \Rightarrow \sqrt[\circ]{D} \in \mathbb{EDM}^N,\ -V\sqrt[\circ]{D}V \in \mathbb{S}_+^N$ (§5.10)

[7.4] Equivalence to de Leeuw's strain problem statement is established for $\rho = N-1$ via (1825) (41) (46).

finds D to attain Euclidean distance of vectorized $-V_{\mathcal{N}}^{\mathrm{T}}HV_{\mathcal{N}}$ to the positive semidefinite cone in isometrically isomorphic subspace $\mathbb{R}^{N(N-1)/2}$, whereas

$$\begin{array}{cl}\underset{D}{\text{minimize}} & \|-V(D-H)V\|_{\mathrm{F}}^2\\ \text{subject to} & D\in\mathbb{EDM}^N\end{array}\tag{1474}$$

attains Euclidean distance of vectorized $-VHV$ to the positive semidefinite cone in ambient isometrically isomorphic $\mathbb{R}^{N(N+1)/2}$; quite different projections[7.5] regardless of whether affine dimension is constrained. But substitution of auxiliary matrix $V_{\mathcal{W}}^{\mathrm{T}}$ (§B.4.3) or $V_{\mathcal{N}}^{\dagger}$ yields the same result as (1472.1) because $V=V_{\mathcal{W}}V_{\mathcal{W}}^{\mathrm{T}}=V_{\mathcal{N}}V_{\mathcal{N}}^{\dagger}$; *id est*,

$$\begin{aligned}\|-V(D-H)V\|_{\mathrm{F}}^2 &= \|-V_{\mathcal{W}}V_{\mathcal{W}}^{\mathrm{T}}(D-H)V_{\mathcal{W}}V_{\mathcal{W}}^{\mathrm{T}}\|_{\mathrm{F}}^2 = \|-V_{\mathcal{W}}^{\mathrm{T}}(D-H)V_{\mathcal{W}}\|_{\mathrm{F}}^2\\ &= \|-V_{\mathcal{N}}V_{\mathcal{N}}^{\dagger}(D-H)V_{\mathcal{N}}V_{\mathcal{N}}^{\dagger}\|_{\mathrm{F}}^2 = \|-V_{\mathcal{N}}^{\dagger}(D-H)V_{\mathcal{N}}\|_{\mathrm{F}}^2\end{aligned}\tag{1475}$$

We see no compelling reason to prefer one particular auxiliary V-matrix over another. Each has its own coherent interpretations; *e.g*, §5.4.2, §6.6, §B.4.5. Neither can we say that any particular problem formulation produces generally better results than another.[7.6]

7.1 First prevalent problem: Projection on PSD cone

This first problem

$$\left.\begin{array}{cl}\underset{D}{\text{minimize}} & \|-V_{\mathcal{N}}^{\mathrm{T}}(D-H)V_{\mathcal{N}}\|_{\mathrm{F}}^2\\ \text{subject to} & \operatorname{rank}V_{\mathcal{N}}^{\mathrm{T}}DV_{\mathcal{N}}\leq\rho\\ & D\in\mathbb{EDM}^N\end{array}\right\}\quad\text{Problem 1}\tag{1476}$$

poses Euclidean projection of $-V_{\mathcal{N}}^{\mathrm{T}}HV_{\mathcal{N}}$ (in subspace $\mathbb{S}^{N-1}$) on a generally nonconvex subset (when $\rho<N-1$) of a positive semidefinite cone boundary $\partial\mathbb{S}_+^{N-1}$ whose elemental matrices have rank no greater than desired affine dimension ρ (§5.7.1.1). Problem 1 finds the closest EDM D in the sense of Schoenberg. (1068) [355] As it is stated, this optimization problem is convex only when desired affine dimension is largest $\rho=N-1$ although its analytical solution is known [292, thm.14.4.2] for all nonnegative $\rho\leq N-1$.[7.7]

We assume only that the given measurement matrix H is symmetric;[7.8]

$$H\in\mathbb{S}^N\tag{1477}$$

Arranging the eigenvalues λ_i of $-V_{\mathcal{N}}^{\mathrm{T}}HV_{\mathcal{N}}$ in nonincreasing order for all i $\big(\lambda_i\geq\lambda_{i+1}$ with corresponding i^{th} eigenvector $v_i\big)$, then an optimal solution to Problem 1 is [400, §2]

$$-V_{\mathcal{N}}^{\mathrm{T}}D^\star V_{\mathcal{N}} = \sum_{i=1}^{\rho}\max\{0\,,\,\lambda_i\}\,v_iv_i^{\mathrm{T}}\tag{1478}$$

[7.5] Isomorphism $T(Y)=V_{\mathcal{N}}^{\dagger\mathrm{T}}YV_{\mathcal{N}}^{\dagger}$ onto $\mathbb{S}_c^N=\{VXV\mid X\in\mathbb{S}^N\}$ relates the map in (1473) to that in (1474), but is not an isometry. This behavior may be observed via Matlab program `isedm()` (provided on *Wıκımization* [439]) that solves (1472.1) for any desired upper bound on affine dimension ρ and allows selection of auxiliary matrix V or $V_{\mathcal{N}}$.

[7.6] All four problem formulations (1472) produce identical results when affine dimension r, implicit to a realizable measurement matrix H, does not exceed desired affine dimension ρ; because, the optimal objective value will vanish ($\|\star\|=0$).

[7.7] being first pronounced in the context of multidimensional scaling by Mardia [291] in 1978 who attributes the generic result (§7.1.2) to Eckart & Young, 1936 [149].

[7.8] Projection, in Problem 1, is on a rank ρ subset of positive semidefinite cone $\mathbb{S}_+^{N-1}$ (§2.9.2.1) in the subspace of symmetric matrices $\mathbb{S}^{N-1}$. It is wrong here to zero the main diagonal of given H because first projecting H, on the symmetric hollow subspace, places an elbow in the path of projection in Problem 1. (Figure 180) where

$$-V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}} \triangleq \sum_{i=1}^{N-1} \lambda_i v_i v_i^{\mathrm{T}} \in \mathbb{S}^{N-1} \tag{1479}$$

is an eigenvalue decomposition and

$$D^\star \in \mathbb{EDM}^N \tag{1480}$$

is an optimal Euclidean distance matrix.

In §7.1.4 we show how to transform Problem 1 to a convex optimization problem for any ρ.

7.1.1 Closest-EDM Problem 1, convex case

7.1.1.0.1 Proof. *Solution* (1478), *convex case.*
When desired affine dimension is unconstrained, $\rho = N-1$, the rank function disappears from (1476) leaving a convex optimization problem; a simple unique minimum-distance projection on positive semidefinite cone $\mathbb{S}_+^{N-1}$: *videlicet*

$$\begin{array}{cl}\underset{D\in\mathbb{S}_h^N}{\text{minimize}} & \|-V_{\mathcal{N}}^{\mathrm{T}}(D-H)V_{\mathcal{N}}\|_{\mathrm{F}}^2 \\ \text{subject to} & -V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \succeq 0\end{array} \tag{1481}$$

by (1068). Because

$$\mathbb{S}^{N-1} = -V_{\mathcal{N}}^{\mathrm{T}} \mathbb{S}_h^N V_{\mathcal{N}} \qquad (1172)$$

then the necessary and sufficient conditions for projection in isometrically isomorphic $\mathbb{R}^{N(N-1)/2}$ on selfdual (387) positive semidefinite cone $\mathbb{S}_+^{N-1}$ are:[7.9] (§E.9.2.0.1) (1767) (*confer* (2252))

$$\begin{array}{l} -V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} \succeq 0 \\ -V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} \left(-V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} + V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}}\right) = \mathbf{0} \\ -V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} + V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}} \succeq 0 \end{array} \tag{1482}$$

Symmetric $-V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}}$ is diagonalizable hence decomposable in terms of its eigenvectors v and eigenvalues λ as in (1479). Therefore (*confer* (1478))

$$-V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} = \sum_{i=1}^{N-1} \max\{0\,,\ \lambda_i\} v_i v_i^{\mathrm{T}} \tag{1483}$$

satisfies (1482), optimally solving (1481). To see that, recall: these eigenvectors constitute an orthogonal set and

$$-V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} + V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}} = -\sum_{i=1}^{N-1} \min\{0\,,\ \lambda_i\} v_i v_i^{\mathrm{T}} \tag{1484}$$

♦

[7.9] The Karush-Kuhn-Tucker (KKT) optimality conditions, [318, p.328] [66, §5.5.3] for problem (1481), are identical to these conditions for projection on a convex cone.

7.1.2 generic problem, projection on PSD cone

Prior to determination of $D^\star$, analytical solution (1478) to Problem 1 is equivalent to solution of a generic rank-constrained projection problem: Given desired affine dimension ρ and

$$A \triangleq -V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}} = \sum_{i=1}^{N-1} \lambda_i \, v_i v_i^{\mathrm{T}} \in \mathbb{S}^{N-1} \qquad (1479)$$

Euclidean projection on a rank ρ subset of a positive semidefinite cone (on a generally nonconvex subset of the PSD cone boundary $\partial\mathbb{S}_+^{N-1}$ when $\rho < N-1$)

$$\left.\begin{array}{cl}\underset{B\in\mathbb{S}^{N-1}}{\text{minimize}} & \|B-A\|_{\mathrm{F}}^2\\ \text{subject to} & \operatorname{rank} B \leq \rho\\ & B \succeq 0\end{array}\right\} \text{ Generic 1} \qquad (1485)$$

has well known optimal solution (Eckart & Young) [149]

$$B^\star \triangleq -V_{\mathcal{N}}^{\mathrm{T}} D^\star V_{\mathcal{N}} = \sum_{i=1}^{\rho} \max\{0\,,\,\lambda_i\}\, v_i v_i^{\mathrm{T}} \in \mathbb{S}^{N-1} \qquad (1478)$$

Once optimal $B^\star$ is found, the technique of §5.12 can be used to determine a uniquely corresponding optimal Euclidean distance matrix $D^\star$; a unique correspondence by injectivity arguments in §5.6.2.

7.1.2.1 Projection on rank ρ subset of PSD cone

Because (1172) provides invertible mapping to the generic problem, then Problem 1

$$\begin{array}{cl}\underset{D\in\mathbb{S}_h^N}{\text{minimize}} & \|-V_{\mathcal{N}}^{\mathrm{T}}(D-H)V_{\mathcal{N}}\|_{\mathrm{F}}^2\\ \text{subject to} & \operatorname{rank} V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \leq \rho\\ & -V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \succeq 0\end{array} \qquad (1486)$$

is truly a Euclidean projection of vectorized $-V_{\mathcal{N}}^{\mathrm{T}} H V_{\mathcal{N}}$ on that generally nonconvex subset of symmetric matrices (belonging to positive semidefinite cone $\mathbb{S}_+^{N-1}$) having rank no greater than desired affine dimension ρ;[7.10] called *rank ρ subset*: (269)

$$\mathbb{S}_+^{N-1} \setminus \mathbb{S}_+^{N-1}(\rho+1) = \{X \in \mathbb{S}_+^{N-1} \mid \operatorname{rank} X \leq \rho\} \qquad (224)$$

7.1.3 Choice of spectral cone

Spectral projection substitutes projection on a polyhedral cone, containing a complete set of eigenspectra (§5.11.1.0.3), in place of projection on a convex set of diagonalizable matrices; *e.g*, (1499). In this section we develop a method of spectral projection for constraining rank of positive semidefinite matrices in a proximity problem like (1485). We will see why an orthant turns out to be the best choice of spectral cone, and why presorting is critical.

Define a nonlinear permutation-operator

$$\pi(x) : \mathbb{R}^n \to \mathbb{R}^n \qquad (1487)$$

that sorts its vector argument x into nonincreasing order; `a.k.a,` *presorting function.*

[7.10] Recall: affine dimension is a lower bound on embedding (§2.3.1), equal to dimension of the smallest affine set in which points from a list X corresponding to an EDM D can be embedded.

7.1.3.0.1 Definition. *Spectral projection.*
Let R be an orthogonal matrix and Λ a nonincreasingly ordered diagonal matrix of eigenvalues. *Spectral projection* means unique minimum-distance projection of a rotated (R, §B.5.5) nonincreasingly ordered (π) vector (δ) of eigenvalues

$$\pi\big(\delta(R^{\mathrm{T}}\Lambda R)\big) \tag{1488}$$

on a polyhedral cone containing all eigenspectra corresponding to a rank ρ subset of a PSD cone (§2.9.2.1) or the EDM cone (in Cayley-Menger form, §5.11.2.3). △

In the simplest and most common case, projection on a positive semidefinite cone, orthogonal matrix R equals I (§7.1.4.0.1) and diagonal matrix Λ is ordered during diagonalization (§A.5.1). Then spectral projection simply means projection of $\delta(\Lambda)$ on a subset of the nonnegative orthant, as we shall now ascertain:

It is curious how nonconvex Problem 1 has such a simple analytical solution (1478). Although solution to generic problem (1485) is well known since 1936 [149], its equivalence was observed in 1997 [400, §2] to projection of an ordered vector of eigenvalues (in diagonal matrix Λ) on a subset of the monotone nonnegative cone (§2.13.10.4.2)

$$\mathcal{K}_{\mathcal{M}+} = \{\upsilon \mid \upsilon_1 \geq \upsilon_2 \geq \cdots \geq \upsilon_{N-1} \geq 0\} \subseteq \mathbb{R}_+^{N-1} \tag{438}$$

Of interest, momentarily, is only the smallest convex subset of the monotone nonnegative cone $\mathcal{K}_{\mathcal{M}+}$ containing every nonincreasingly ordered eigenspectrum corresponding to a rank ρ subset of positive semidefinite cone $\mathbb{S}_+^{N-1}$; *id est*,

$$\mathcal{K}_{\mathcal{M}+}^{\rho} \triangleq \{\upsilon\in\mathbb{R}^{\rho} \mid \upsilon_1 \geq \upsilon_2 \geq \cdots \geq \upsilon_{\rho} \geq 0\} \subseteq \mathbb{R}_+^{\rho} \tag{1489}$$

a pointed polyhedral cone, a ρ-dimensional convex subset of the monotone nonnegative cone $\mathcal{K}_{\mathcal{M}+}\subseteq\mathbb{R}_+^{N-1}$ having property, for λ denoting eigenspectra,

$$\begin{bmatrix} \mathcal{K}_{\mathcal{M}+}^{\rho} \\ \mathbf{0} \end{bmatrix} = \pi(\lambda(\text{rank } \rho \text{ subset})) \subseteq \mathcal{K}_{\mathcal{M}+}^{N-1} \triangleq \mathcal{K}_{\mathcal{M}+} \tag{1490}$$

For each and every elemental eigenspectrum

$$\gamma \in \lambda(\text{rank } \rho \text{ subset}) \subseteq \mathbb{R}_+^{N-1} \tag{1491}$$

of the rank ρ subset (ordered or unordered in λ), there is a nonlinear surjection $\pi(\gamma)$ onto $\mathcal{K}_{\mathcal{M}+}^{\rho}$.

7.1.3.0.2 Exercise. *Smallest spectral cone.*
Prove that there is no convex subset of $\mathcal{K}_{\mathcal{M}+}$ smaller than $\mathcal{K}_{\mathcal{M}+}^{\rho}$ containing every ordered eigenspectrum corresponding to the rank ρ subset of a positive semidefinite cone (§2.9.2.1). ▼

7.1.3.0.3 Proposition. (Hardy-Littlewood-Pólya) *Inequalities.* [208, §X] [59, §1.2]
Any vectors σ and γ in $\mathbb{R}^{N-1}$ satisfy a tight inequality

$$\pi(\sigma)^{\mathrm{T}}\pi(\gamma) \;\geq\; \sigma^{\mathrm{T}}\gamma \;\geq\; \pi(\sigma)^{\mathrm{T}}\Xi\,\pi(\gamma) \tag{1492}$$

where Ξ is the order-reversing permutation matrix defined in (1920), and permutator $\pi(\gamma)$ is a nonlinear function that sorts vector γ into nonincreasing order thereby providing the greatest upper bound and least lower bound with respect to every possible sorting. ⋄

7.1.3.0.4 Corollary. *Monotone nonnegative sort.*
Any given vectors $\sigma\,,\gamma\in\mathbb{R}^{N-1}$ satisfy a tight Euclidean distance inequality

$$\|\pi(\sigma)-\pi(\gamma)\| \;\leq\; \|\sigma-\gamma\| \tag{1493}$$

where nonlinear function $\pi(\gamma)$ sorts vector γ into nonincreasing order thereby providing the least lower bound with respect to every possible sorting. ⋄

Given $\gamma\in\mathbb{R}^{N-1}$

$$\inf_{\sigma\in\mathbb{R}_+^{N-1}}\|\sigma-\gamma\| \;=\inf_{\sigma\in\mathbb{R}_+^{N-1}}\|\pi(\sigma)-\pi(\gamma)\| \;=\inf_{\sigma\in\mathbb{R}_+^{N-1}}\|\sigma-\pi(\gamma)\| \;=\inf_{\sigma\in\mathcal{K}_{\mathcal{M}+}}\|\sigma-\pi(\gamma)\| \tag{1494}$$

Yet for γ representing an arbitrary vector of eigenvalues, because

$$\inf_{\sigma\in\left[\begin{smallmatrix}\mathbb{R}_+^\rho\\ \mathbf{0}\end{smallmatrix}\right]}\|\sigma-\gamma\|^2 \;\geq \inf_{\sigma\in\left[\begin{smallmatrix}\mathbb{R}_+^\rho\\ \mathbf{0}\end{smallmatrix}\right]}\|\sigma-\pi(\gamma)\|^2 \;=\inf_{\sigma\in\left[\begin{smallmatrix}\mathcal{K}_{\mathcal{M}+}^\rho\\ \mathbf{0}\end{smallmatrix}\right]}\|\sigma-\pi(\gamma)\|^2 \tag{1495}$$

then projection of γ on the eigenspectra of a rank ρ subset can be tightened simply by presorting γ into nonincreasing order.

Proof. Simply because $\pi(\gamma)_{1:\rho}\succeq\pi(\gamma_{1:\rho})$

$$\begin{aligned}
\inf_{\sigma\in\left[\begin{smallmatrix}\mathbb{R}_+^\rho\\ \mathbf{0}\end{smallmatrix}\right]}\|\sigma-\gamma\|^2 &= \gamma_{\rho+1:N-1}^{\mathrm{T}}\gamma_{\rho+1:N-1} + \inf_{\sigma\in\mathbb{R}_+^{N-1}}\|\sigma_{1:\rho}-\gamma_{1:\rho}\|^2\\
&= \gamma^{\mathrm{T}}\gamma + \inf_{\sigma\in\mathbb{R}_+^{N-1}}\sigma_{1:\rho}^{\mathrm{T}}\sigma_{1:\rho} - 2\sigma_{1:\rho}^{\mathrm{T}}\gamma_{1:\rho}\\
&\geq \gamma^{\mathrm{T}}\gamma + \inf_{\sigma\in\mathbb{R}_+^{N-1}}\sigma_{1:\rho}^{\mathrm{T}}\sigma_{1:\rho} - 2\sigma_{1:\rho}^{\mathrm{T}}\pi(\gamma)_{1:\rho}\\
\inf_{\sigma\in\left[\begin{smallmatrix}\mathbb{R}_+^\rho\\ \mathbf{0}\end{smallmatrix}\right]}\|\sigma-\gamma\|^2 &\geq \inf_{\sigma\in\left[\begin{smallmatrix}\mathbb{R}_+^\rho\\ \mathbf{0}\end{smallmatrix}\right]}\|\sigma-\pi(\gamma)\|^2
\end{aligned} \tag{1496}$$

◆

7.1.3.1 Orthant is best spectral cone for Problem 1

This means unique minimum-distance projection of γ on the nearest spectral member of the rank ρ subset is tantamount to presorting γ into nonincreasing order. Only then does unique spectral projection on a subset $\mathcal{K}_{\mathcal{M}+}^\rho$ of the monotone nonnegative cone become equivalent to unique spectral projection on a subset $\mathbb{R}_+^\rho$ of the nonnegative orthant (which is simpler); in other words, unique minimum-distance projection of sorted γ on the nonnegative orthant in a ρ-dimensional subspace of $\mathbb{R}^N$ is indistinguishable from its projection on the subset $\mathcal{K}_{\mathcal{M}+}^\rho$ of the monotone nonnegative cone in that same subspace.

7.1.4 Closest-EDM Problem 1, "nonconvex" case

Proof of solution (1478), for projection on a rank ρ subset of positive semidefinite cone $\mathbb{S}_+^{N-1}$, can be algebraic in nature. [400, §2] Here we derive that known result but instead using a more geometric argument via spectral projection on a polyhedral cone (subsuming the proof in §7.1.1). In so doing, we demonstrate how nonconvex Problem 1 is transformed to a convex optimization:

7.1.4.0.1 Proof. *Solution* (1478), *nonconvex case.*
As explained in §7.1.2, we may instead work with the more facile generic problem (1485). With diagonalization of unknown

$$B \triangleq U\Upsilon U^{\mathrm{T}} \in \mathbb{S}^{N-1} \tag{1497}$$

given desired affine dimension $0 \leq \rho \leq N-1$ and diagonalizable

$$A \triangleq Q\Lambda Q^{\mathrm{T}} \in \mathbb{S}^{N-1} \tag{1498}$$

having eigenvalues in Λ arranged in nonincreasing order, by (51) the generic problem is equivalent to

$$\begin{array}{cl} \underset{B\in\mathbb{S}^{N-1}}{\text{minimize}} & \|B-A\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} B \leq \rho \\ & B \succeq 0 \end{array} \quad\equiv\quad \begin{array}{cl} \underset{R\,,\,\Upsilon}{\text{minimize}} & \|\Upsilon - R^{\mathrm{T}}\Lambda R\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank}\Upsilon \leq \rho \\ & \Upsilon \succeq 0 \\ & R^{-1} = R^{\mathrm{T}} \end{array} \tag{1499}$$

where

$$R \triangleq Q^{\mathrm{T}} U \in \mathbb{R}^{N-1\times N-1} \tag{1500}$$

is a bijection in U on the set of orthogonal matrices. We propose solving (1499) by instead solving the problem sequence:

$$\begin{array}{ll} \begin{array}{cl} \underset{\Upsilon}{\text{minimize}} & \|\Upsilon - R^{\mathrm{T}}\Lambda R\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank}\Upsilon \leq \rho \\ & \Upsilon \succeq 0 \end{array} & \text{(a)} \\ \begin{array}{cl} \underset{R}{\text{minimize}} & \|\Upsilon^\star - R^{\mathrm{T}}\Lambda R\|_{\mathrm{F}}^2 \\ \text{subject to} & R^{-1} = R^{\mathrm{T}} \end{array} & \text{(b)} \end{array} \tag{1501}$$

Problem (1501a) is equivalent to:

(1) orthogonal projection of $R^{\mathrm{T}}\Lambda R$ on an $N-1$-dimensional subspace of isometrically isomorphic $\mathbb{R}^{N(N-1)/2}$ containing $\delta(\Upsilon) \in \mathbb{R}_+^{N-1}$

(2) nonincreasingly ordering the result,

(3) unique minimum-distance projection of the ordered result on $\begin{bmatrix} \mathbb{R}_+^\rho \\ \mathbf{0} \end{bmatrix}$ (§E.9.5).

Projection on that $N-1$-dimensional subspace amounts to zeroing $R^{\mathrm{T}}\Lambda R$ at all entries off the main diagonal; thus, the equivalent sequence leading with a spectral projection:

$$\begin{array}{ll} \begin{array}{cl} \underset{\Upsilon}{\text{minimize}} & \|\,\delta(\Upsilon) - \pi\big(\delta(R^{\mathrm{T}}\Lambda R)\big)\,\|^2 \\ \text{subject to} & \delta(\Upsilon) \in \begin{bmatrix} \mathbb{R}_+^\rho \\ \mathbf{0} \end{bmatrix} \end{array} & \text{(a)} \\ \begin{array}{cl} \underset{R}{\text{minimize}} & \|\Upsilon^\star - R^{\mathrm{T}}\Lambda R\|_{\mathrm{F}}^2 \\ \text{subject to} & R^{-1} = R^{\mathrm{T}} \end{array} & \text{(b)} \end{array} \tag{1502}$$

Because any permutation matrix is an orthogonal matrix, $\delta(R^{\mathrm{T}}\Lambda R) \in \mathbb{R}^{N-1}$ can always be arranged in nonincreasing order without loss of generality; hence, permutation operator π .

Unique minimum-distance projection of vector $\pi\big(\delta(R^{\mathrm T}\Lambda R)\big)$ on the ρ-dimensional subset $\left[\begin{array}{c}\mathbb{R}^\rho_+\\ \mathbf{0}\end{array}\right]$ of nonnegative orthant $\mathbb{R}^{N-1}_+$ requires: (§E.9.2.0.1)

$$\begin{aligned}\delta(\Upsilon^\star)_{\rho+1:N-1} &= \mathbf{0}\\ \delta(\Upsilon^\star) &\succeq 0\\ \delta(\Upsilon^\star)^{\mathrm T}\big(\delta(\Upsilon^\star)-\pi(\delta(R^{\mathrm T}\Lambda R))\big) &= 0\\ \delta(\Upsilon^\star)-\pi(\delta(R^{\mathrm T}\Lambda R)) &\succeq 0\end{aligned}\tag{1503}$$

which are necessary and sufficient conditions. Any value $\Upsilon^\star$ satisfying conditions (1503) is optimal for (1502a). So

$$\delta(\Upsilon^\star)_i = \left\{\begin{array}{ll}\max\left\{0\,,\ \pi\big(\delta(R^{\mathrm T}\Lambda R)\big)_i\right\}, & i=1\ldots\rho\\ 0\,, & i=\rho+1\ldots N-1\end{array}\right.\tag{1504}$$

specifies an optimal solution. The lower bound on the objective with respect to R in (1502b) is tight: by (1469)

$$\|\,|\Upsilon^\star|-|\Lambda|\,\|_{\mathrm F}\ \leq\ \|\Upsilon^\star-R^{\mathrm T}\Lambda R\|_{\mathrm F}\tag{1505}$$

where $|\ |$ denotes absolute entry-value. For selection of $\Upsilon^\star$ as in (1504), this lower bound is attained when (*confer* §C.4.2.2)

$$R^\star = I\tag{1506}$$

which is the known solution. ♦

7.1.4.1 significance

Importance of this well-known [149] optimal solution (1478) for projection on a rank ρ subset of a positive semidefinite cone should not be dismissed:

- Problem 1 (1476) and its generic form (1485), as stated, are generally nonconvex. Their known analytical solution encompasses projection on a rank ρ subset (224) of a positive semidefinite cone (generally, a nonconvex subset of its boundary) from either the exterior or interior of that cone.[7.11] By problem transformation to the spectral domain, projection on a rank ρ subset becomes a convex optimization problem.

- This solution is closed form.

- This solution is equivalent to projection on a polyhedral cone in the spectral domain (spectral projection §7.1.3.0.1, projection on a spectral cone §5.11.1.0.2); a necessary and sufficient condition (§A.3.1) for membership of a symmetric matrix to a rank ρ subset of a positive semidefinite cone (§2.9.2.1).

- A minimum-distance projection, on a rank ρ subset of a positive semidefinite cone, is a positive semidefinite matrix orthogonal (in the Euclidean sense) to direction of projection[7.12] because $U^\star = Q$ in (1500).

- For the convex case problem (1481), this solution is always unique. Otherwise, distinct eigenvalues (multiplicity 1) in Λ guarantee uniqueness of this solution by the reasoning in §A.5.0.1.[7.13]

[7.11] Projection on the boundary from the interior, of a convex Euclidean body, is generally a nonconvex problem. (§E.9.1.1.2)

[7.12] But Theorem E.9.2.0.1, for unique projection on a closed convex cone, does not apply here because direction of projection is not necessarily a member of the dual PSD cone. This occurs, for example, whenever positive eigenvalues are truncated.

[7.13] Uncertainty of uniqueness prevents the erroneous conclusion that a rank ρ subset (224) were a convex body by the *Bunt-Motzkin theorem* (§E.9.0.0.1).

7.1.4.2 list projection interpretation

Because $-VDV\frac{1}{2}=X^{\mathrm T}X$ when point list X is geometrically centered, $X\mathbf{1}=\mathbf{0}$, Problem 1 can be equivalently restated: by (1070)

$$(1)\quad \begin{array}{ll}\underset{D}{\text{minimize}} & \|-V(D-H)V\|_{\mathrm F}^2\\ \text{subject to} & \operatorname{rank} VDV\leq\rho\\ & D\in\mathbb{EDM}^N\end{array} \quad\equiv\quad \underset{X\in\mathbb{R}^{\rho\times N}}{\text{minimize}}\ \|X^{\mathrm T}X-Y^{\mathrm T}Y\|_{\mathrm F}^2 \quad (\mathrm G) \tag{1507}$$

where $Y\in\mathbb{R}^{n\times N}$ comprises geometrically centered point list estimates ($Y=YV$) whose dimensionality is to be reduced (by best fit) to

$$\rho\leq\eta\triangleq\min\{n\,,\,N\} \tag{1508}$$

We call (1507.G) the *Gram-form* Problem 1; it may be interpreted as minimum-distance projection of $Y^{\mathrm T}Y\in\mathbb{S}^N$ on a rank ρ subset (§2.9.2.1) of the PSD cone in isometrically isomorphic $\mathbb{R}^{N(N+1)/2}$. Geometrically centered Y remains centered, postprojection, because the subspace $\mathbb{S}_c^N$ of symmetric geometrically centered matrices $VY^{\mathrm T}YV$ (1151) is invariant to projection on a positive semidefinite cone by Lemma 6.8.1.1.1.

Orthogonal projection of estimates Y , on span of ρ principal eigenvectors of $YY^{\mathrm T}\in\mathbb{S}^n$, provides unique (not rotation invariant) optimal $X^\star$ in the sense

$$\begin{array}{ll}\underset{X\in\mathbb{R}^{n\times N}}{\text{minimize}} & \|X^{\mathrm T}X-Y^{\mathrm T}Y\|_{\mathrm F}^2\\ \text{subject to} & \operatorname{rank}(X^{\mathrm T}X)\leq\rho\end{array} = \begin{array}{ll}\underset{X\in\mathbb{R}^{n\times N}}{\text{minimize}} & \|XX^{\mathrm T}-YY^{\mathrm T}\|_{\mathrm F}^2\\ \text{subject to} & \operatorname{rank}(XX^{\mathrm T})\leq\rho\end{array} = \sum_{i=\rho+1}^{\eta}\lambda(Y^{\mathrm T}Y)_i^2 \tag{1509}$$

where $\operatorname{rank}(X^{\mathrm T}X)=\operatorname{rank}(XX^{\mathrm T})$ (1634). Defining nonincreasingly ordered diagonalization $YY^{\mathrm T}\triangleq Q_n\Lambda_nQ_n^{\mathrm T}\in\mathbb{S}^n$, then orthogonal projection of Y is (§E.3.2)[7.14]

$$X^\star = Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}Y\in\mathbb{R}^{n\times N} \tag{1510}$$

So

$$\begin{aligned}\|X^\star X^{\star\mathrm T}-YY^{\mathrm T}\|_{\mathrm F}^2 &= \|Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}YY^{\mathrm T}-YY^{\mathrm T}\|_{\mathrm F}^2\\ &= \left\|\begin{bmatrix}\Lambda_n(1\!:\!\rho\,,1\!:\!\rho) & \mathbf{0}^{\mathrm T}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}-\Lambda_n\right\|_{\mathrm F}^2 = \sum_{i=\rho+1}^{\eta}\lambda(Y^{\mathrm T}Y)_i^2\end{aligned} \tag{1511}$$

projection of list Y on a subspace solves projection of Gram matrix $Y^{\mathrm T}Y$ [*sic*] on a positive semidefinite cone (1509); quite a remarkable interpretation.[7.15] [292, §14.4] [334, §2]

Yet there is a more plain interpretation:

$$X^\star X^{\star\mathrm T}=Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}YY^{\mathrm T}=Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}YY^{\mathrm T}Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}\in\mathbb{S}_+^n \tag{1512}$$

is the orthogonal projection of $YY^{\mathrm T}$ on the closest $\rho(\rho+1)/2$-dimensional subspace

$$Q_n(:\,,1\!:\!\rho)\mathbb{S}^\rho Q_n(:\,,1\!:\!\rho)^{\mathrm T} = Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}\mathbb{S}^nQ_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T} \tag{1513}$$

of a rotated Cartesian coordinate system $Q_n\mathbb{S}^nQ_n^{\mathrm T}$ in isomorphic $\mathbb{R}^{n(n+1)/2}$. That it is the closest subspace, comes from §2.13.8.1.1. That (1513) is the smallest subspace containing the smallest face (that contains $X^\star X^{\star\mathrm T}$) of PSD cone $\mathbb{S}_+^n$, is a result from §2.9.2.4.

[7.14] Reconstruction of $X^\star$, with dimension ρ instead of n , is disclosed in §5.12.2.

[7.15] This might imply existence of an isomorphism (§2.2.1.0.1) relating vector space $\mathbb{R}^{nN}$ (containing vectorized list X) to vector space $\mathbb{R}^{N(N+1)/2}$ (containing vectorized cone $\mathbb{S}_+^N$); but there is none. Such an isomorphism might be an isometry (2.2.1.1.1) were $\|X^\star X^{\star\mathrm T}-YY^{\mathrm T}\|_{\mathrm F}^2$ equal to

$$\begin{array}{ll}\underset{X\in\mathbb{R}^{n\times N}}{\text{minimize}} & \|X-Y\|_{\mathrm F}^2\\ \text{subject to} & X\mathbf{1}=\mathbf{0}\\ & \operatorname{rank}X\leq\rho\end{array} = \|Q_n(:\,,1\!:\!\rho)Q_n(:\,,1\!:\!\rho)^{\mathrm T}Y-Y\|_{\mathrm F}^2 = \sum_{i=\rho+1}^{\eta}\lambda(Y^{\mathrm T}Y)_i$$

but the square of eigenvalues is absent with respect to (1511); numerically verifiable by means of problem transformation in §4.9 and a few convex iterations (§4.5.1).

7.1.5 Problem 1 in spectral norm, convex case

When instead we pose the matrix 2-norm (spectral norm) in Problem 1 (1476) for the convex case $\rho = N-1$, then the new problem

$$\begin{array}{ll}\underset{D}{\text{minimize}} & \|-V_{\mathcal{N}}^{\mathrm{T}}(D-H)V_{\mathcal{N}}\|_2 \\ \text{subject to} & D \in \mathbb{EDM}^N\end{array} \tag{1514}$$

is convex although its solution is not necessarily unique;[7.16] giving rise to nonorthogonal projection (§E.1) on positive semidefinite cone $\mathbb{S}_+^{N-1}$. Indeed, its solution set includes the Frobenius solution (1478) for the convex case whenever $-V_{\mathcal{N}}^{\mathrm{T}}HV_{\mathcal{N}}$ is a normal matrix. [215, §1] [206] [66, §8.1.1] Proximity problem (1514) is equivalent to

$$\begin{array}{ll}\underset{\mu\,,\,D}{\text{minimize}} & \mu \\ \text{subject to} & -\mu I \preceq -V_{\mathcal{N}}^{\mathrm{T}}(D-H)V_{\mathcal{N}} \preceq \mu I \\ & D \in \mathbb{EDM}^N\end{array} \tag{1515}$$

by (1896) where

$$\mu^\star = \max_i \left\{ \left|\lambda\big(-V_{\mathcal{N}}^{\mathrm{T}}(D^\star - H)V_{\mathcal{N}}\big)_i\right| \,,\quad i=1\ldots N-1 \right\} \in \mathbb{R}_+ \tag{1516}$$

is the minimized largest absolute eigenvalue (due to matrix symmetry).

For lack of unique solution here, we prefer the Frobenius rather than spectral norm.

7.2 Second prevalent problem: Projection on EDM cone in $\sqrt{d_{ij}}$

Let

$$\sqrt[\circ]{D} \triangleq [\sqrt{d_{ij}}\,] \in \mathcal{K} = \mathbb{S}_h^N \cap \mathbb{R}_+^{N\times N} \tag{1517}$$

be an unknown matrix of absolute distance; *id est*,

$$D = [d_{ij}] \triangleq \sqrt[\circ]{D} \circ \sqrt[\circ]{D} \in \mathbb{EDM}^N \tag{1518}$$

where $\circ$ denotes Hadamard product. The second prevalent proximity problem is a Euclidean projection (in the natural coordinates $\sqrt{d_{ij}}$) of matrix H on a nonconvex subset of the boundary of the nonconvex cone of Euclidean absolute-distance matrices $\operatorname{rel}\partial\sqrt{\mathbb{EDM}^N}$: (§6.3, *confer* Figure **165**b)

$$\left.\begin{array}{ll}\underset{\sqrt[\circ]{D}}{\text{minimize}} & \|\sqrt[\circ]{D} - H\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \leq \rho \\ & \sqrt[\circ]{D} \in \sqrt{\mathbb{EDM}^N}\end{array}\right\} \quad \text{Problem 2} \tag{1519}$$

where

$$\sqrt{\mathbb{EDM}^N} = \{\sqrt[\circ]{D} \mid D \in \mathbb{EDM}^N\} \qquad (1344)$$

This statement of the second proximity problem is considered difficult to solve because of the constraint on desired affine dimension ρ (§5.7.2) and because the objective function

$$\|\sqrt[\circ]{D} - H\|_{\mathrm{F}}^2 = \sum_{i,j}(\sqrt{d_{ij}} - h_{ij})^2 \tag{1520}$$

[7.16] For each and every $|t| \leq 2$, for example, $\begin{bmatrix} 2 & 0 \\ 0 & t \end{bmatrix}$ has the same spectral-norm value.

is expressed in the natural coordinates; projection on a doubly nonconvex set.

Our solution to this second problem prevalent in the literature requires measurement matrix H to be nonnegative;

$$H = [h_{ij}] \in \mathbb{R}_+^{N\times N} \tag{1521}$$

If the H matrix given has negative entries, then the technique of solution presented here becomes invalid. As explained in §7.0.1, projection of H on $\mathcal{K} = \mathbb{S}_h^N \cap \mathbb{R}_+^{N\times N}$ (1463) prior to application of this proposed solution is incorrect.

7.2.1 Convex case

When $\rho = N-1$, the rank constraint vanishes and a convex problem that is equivalent to (1470) emerges:[7.17]

$$\begin{array}{cl}\underset{\sqrt[\circ]{D}}{\text{minimize}} & \|\sqrt[\circ]{D} - H\|_{\mathrm F}^2\\ \text{subject to} & \sqrt[\circ]{D} \in \sqrt{\mathbb{EDM}^N}\end{array} \quad\Leftrightarrow\quad \begin{array}{cl}\underset{D}{\text{minimize}} & \sum\limits_{i,j} d_{ij} - 2h_{ij}\sqrt{d_{ij}} + h_{ij}^2\\ \text{subject to} & D \in \mathbb{EDM}^N\end{array} \tag{1522}$$

For any fixed i and j, the argument of summation is a convex function of d_{ij} because (for nonnegative constant h_{ij}) the negative square root is convex in nonnegative d_{ij} and because $d_{ij} + h_{ij}^2$ is affine (convex). Because the sum of any number of convex functions in D remains convex [66, §3.2.1] and because the feasible set is convex in D, we have a convex optimization problem:

$$\begin{array}{cl}\underset{D}{\text{minimize}} & \mathbf{1}^{\mathrm T}(D - 2H\circ\sqrt[\circ]{D}\,)\mathbf{1} + \|H\|_{\mathrm F}^2\\ \text{subject to} & D \in \mathbb{EDM}^N\end{array} \tag{1523}$$

The objective function being a sum of strictly convex functions is, moreover, strictly convex in D on the nonnegative orthant. Existence of a unique solution $D^\star$ for this second prevalent problem depends upon nonnegativity of H and a convex feasible set (§3.1.1).[7.18]

7.2.1.1 Equivalent semidefinite program, Problem 2, convex case

Convex problem (1522) is numerically solvable for its global minimum using an interior-point method [461] [327] [315] [452] [12] [169]. We translate (1522) to an equivalent semidefinite program (SDP) for a pedagogical reason made clear in §7.2.2.2 and because there exist readily available computer programs for numerical solution [195] [454] [455] [406] [36] [453] [395] [379].

Substituting a new matrix variable $Y \triangleq [y_{ij}] \in \mathbb{R}_+^{N\times N}$

$$h_{ij}\sqrt{d_{ij}} \leftarrow y_{ij} \tag{1524}$$

Boyd proposes: problem (1522) is equivalent to the semidefinite program

$$\begin{array}{cl}\underset{D\,,\,Y}{\text{minimize}} & \sum\limits_{i,j} d_{ij} - 2y_{ij} + h_{ij}^2\\ \text{subject to} & \begin{bmatrix} d_{ij} & y_{ij}\\ y_{ij} & h_{ij}^2\end{bmatrix} \succeq 0\,, \quad i,j=1\ldots N\\ & D \in \mathbb{EDM}^N\end{array} \tag{1525}$$

[7.17] still thought to be a nonconvex problem as late as 1997 [401] even though discovered convex by de Leeuw in 1993. [116] [57, §13.6] Yet using methods from §3, it can be easily ascertained: $\|\sqrt[\circ]{D} - H\|_{\mathrm F}$ is not convex in D.

[7.18] The transformed problem in variable D no longer describes Euclidean projection on an EDM cone. Otherwise we might erroneously conclude $\sqrt{\mathbb{EDM}^N}$ were a convex body by the *Bunt-Motzkin theorem* (§E.9.0.0.1).

To see that, recall: $d_{ij} \geq 0$ is implicit to $D \in \mathbb{EDM}^N$ (§5.8.1, (1068)). So when $H \in \mathbb{R}_+^{N \times N}$ is nonnegative, as assumed,

$$\begin{bmatrix} d_{ij} & y_{ij} \\ y_{ij} & h_{ij}^2 \end{bmatrix} \succeq 0 \;\Leftrightarrow\; h_{ij}\sqrt{d_{ij}} \geq \sqrt{y_{ij}^2} \tag{1526}$$

by Theorem A.3.1.0.4. Minimization of the objective function implies maximization of y_{ij} that is bounded above. Hence nonnegativity of y_{ij} is implicit to (1525) and, as desired, $y_{ij} \rightarrow h_{ij}\sqrt{d_{ij}}$ as optimization proceeds. ♦

If the given matrix H is now assumed symmetric and nonnegative,

$$H = [h_{ij}] \in \mathbb{S}^N \cap \mathbb{R}_+^{N \times N} \tag{1527}$$

then $Y = H \circ \sqrt[\circ]{D}$ must belong to $\mathcal{K} = \mathbb{S}_h^N \cap \mathbb{R}_+^{N \times N}$ (1463). Because $Y \in \mathbb{S}_h^N$ (§B.4.2 *no.*20), then

$$\|\sqrt[\circ]{D} - H\|_{\mathrm{F}}^2 \;=\; \sum_{i,j} d_{ij} - 2y_{ij} + h_{ij}^2 \;=\; -N \operatorname{tr}\big(V(D - 2\,Y)V\big) + \|H\|_{\mathrm{F}}^2 \tag{1528}$$

So convex problem (1525) is equivalent to the semidefinite program

$$\begin{array}{cl} \underset{D\,,\,Y}{\text{minimize}} & -\operatorname{tr}\big(V(D - 2\,Y)V\big) \\ \text{subject to} & \begin{bmatrix} d_{ij} & y_{ij} \\ y_{ij} & h_{ij}^2 \end{bmatrix} \succeq 0\,, \quad N \geq j > i = 1 \ldots N-1 \\ & Y \in \mathbb{S}_h^N \\ & D \in \mathbb{EDM}^N \end{array} \tag{1529}$$

where the constants h_{ij}^2 and N have been dropped arbitrarily from the objective.

7.2.1.2 Gram-form semidefinite program, Problem 2, convex case

There is great advantage to expressing problem statement (1529) in Gram-form because Gram matrix G is a bidirectional bridge between point list X and distance matrix D ; *e.g*, §5.4.2.2.8, §6.7.0.0.1. This way, problem convexity can be maintained while simultaneously constraining point list X , Gram matrix G , and distance matrix D at our discretion.

Convex problem (1529) may be equivalently written via linear bijective (§5.6.1) EDM operator $\mathbf{D}(G)$ (1061);

$$\begin{array}{cl} \underset{G \in \mathbb{S}_c^N\,,\; Y \in \mathbb{S}_h^N}{\text{minimize}} & -\operatorname{tr}\big(V(\mathbf{D}(G) - 2\,Y)V\big) \\ \text{subject to} & \begin{bmatrix} \langle \Phi_{ij}\,,\, G\rangle & y_{ij} \\ y_{ij} & h_{ij}^2 \end{bmatrix} \succeq 0\,, \quad N \geq j > i = 1 \ldots N-1 \\ & G \succeq 0 \end{array} \tag{1530}$$

where distance-square $D = [d_{ij}] \in \mathbb{S}_h^N$ (1045) is related to $G = [g_{ij}] \in \mathbb{S}_c^N \cap \mathbb{S}_+^N$ Gram matrix entries by

$$\begin{aligned} d_{ij} &= g_{ii} + g_{jj} - 2g_{ij} \\ &= \langle \Phi_{ij}\,,\, G\rangle \end{aligned} \tag{1060}$$

where

$$\Phi_{ij} = (e_i - e_j)(e_i - e_j)^{\mathrm{T}} \in \mathbb{S}_+^N \tag{1047}$$

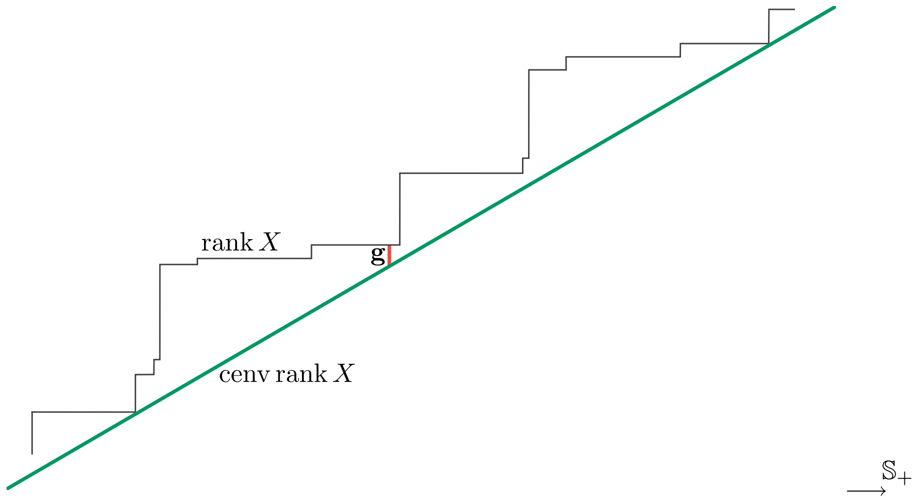

Figure 181: Abstraction of convex envelope of rank function. Rank is a quasiconcave monotonic function on a positive semidefinite cone $\mathbb{S}_+$, but its convex envelope is the largest convex function whose epigraph contains it. Vertical bar labelled $\mathbf{g}$ measures a trace−rank gap; *id est*, rank found always exceeds estimate; large decline in trace required here for only a small decrease in rank.

Confinement of G to the geometric center subspace provides numerical stability and no loss of generality (*confer* (1408)); implicit constraint $G\mathbf{1}=\mathbf{0}$ is otherwise unnecessary.

To include constraints on the list $X\in\mathbb{R}^{n\times N}$, we would first rewrite (1530)

$$\begin{array}{cl}\underset{G\in\mathbb{S}_c^N\,,\; Y\in\mathbb{S}_h^N\,,\; X\in\mathbb{R}^{n\times N}}{\text{minimize}} & -\operatorname{tr}(V(\mathbf{D}(G)-2Y)V)\\ \text{subject to} & \begin{bmatrix}\langle\Phi_{ij}\,,G\rangle & y_{ij}\\ y_{ij} & h_{ij}^2\end{bmatrix}\succeq 0\,,\quad N\geq j>i=1\ldots N-1\\ & \begin{bmatrix} I & X\\ X^{\mathrm T} & G\end{bmatrix}\succeq 0\\ & X\in\mathcal{C}\end{array}\tag{1531}$$

and then introduce the constraints, realized here in abstract membership to some convex set $\mathcal{C}$. This problem realization includes a convex relaxation of the nonconvex constraint $G=X^{\mathrm T}X$. If desired, more constraints on G could be introduced. These techniques are discussed in §5.4.2.2.8.

7.2.2 Minimization of affine dimension in Problem 2

When desired affine dimension ρ is diminished, the rank function becomes reinserted into problem (1525) that is then rendered difficult to solve because feasible set $\{D\,,Y\}$ loses convexity in $\mathbb{S}_h^N\times\mathbb{R}^{N\times N}$. Indeed, the rank function is quasiconcave (§3.14) on a positive semidefinite cone; (§2.9.2.9.2) *id est*, its sublevel sets are not convex.

7.2.2.1 Rank minimization heuristic

A remedy developed in [298] [156] [157] [155] introduces convex envelope of the quasiconcave rank function: (Figure **181**)

7.2.2.1.1 Definition. *Convex envelope.* [229]
Convex envelope $\operatorname{cenv} f$ of a function $f:\mathcal{C}\to\mathbb{R}$ is defined to be the largest convex function g such that $g\leq f$ on convex domain $\mathcal{C}\subseteq\mathbb{R}^n$. [7.19] △

- [156] [155] Convex envelope of rank function: for σ_i a singular value, (1737)

$$\operatorname{cenv}(\operatorname{rank} A) \text{ on } \{A\in\mathbb{R}^{m\times n} \mid \|A\|_2\leq\kappa\} \;=\; \frac{1}{\kappa}\mathbf{1}^{\mathrm{T}}\sigma(A) \;=\; \frac{1}{\kappa}\operatorname{tr}\sqrt{A^{\mathrm{T}}A} \tag{1532}$$

$$\operatorname{cenv}(\operatorname{rank} A) \text{ on } \{A \text{ normal} \mid \|A\|_2\leq\kappa\} \;=\; \frac{1}{\kappa}\|\lambda(A)\|_1 \;=\; \frac{1}{\kappa}\operatorname{tr}\sqrt{A^{\mathrm{T}}A} \tag{1533}$$

$$\operatorname{cenv}(\operatorname{rank} A) \text{ on } \{A\in\mathbb{S}_+^n \mid \|A\|_2\leq\kappa\} \;=\; \frac{1}{\kappa}\mathbf{1}^{\mathrm{T}}\lambda(A) \;=\; \frac{1}{\kappa}\operatorname{tr}(A) \tag{1534}$$

A properly scaled trace thus represents the best convex lower bound on rank for positive semidefinite matrices. The idea, then, is to substitute convex envelope for rank of some variable $A\in\mathbb{S}_+^M$ (§A.6.2.2)

$$\operatorname{rank} A \;\leftarrow\; \operatorname{cenv}(\operatorname{rank} A) \;\propto\; \operatorname{tr} A \;=\; \sum_i \sigma(A)_i \;=\; \sum_i \lambda(A)_i \tag{1535}$$

which is equivalent to the sum of all eigenvalues or singular values.

- [155] Convex envelope of the cardinality function is proportional to the 1-norm:

$$\operatorname{cenv}(\operatorname{card} x) \text{ on } \{x\in\mathbb{R}^n \mid \|x\|_\infty\leq\kappa\} \;=\; \frac{1}{\kappa}\|x\|_1 \tag{1536}$$

$$\operatorname{cenv}(\operatorname{card} x) \text{ on } \{x\in\mathbb{R}_+^n \mid \|x\|_\infty\leq\kappa\} \;=\; \frac{1}{\kappa}\mathbf{1}^{\mathrm{T}}x \tag{1537}$$

7.2.2.2 Applying trace rank-heuristic to Problem 2

Substituting rank envelope for rank function in Problem 2, for $D\in\mathbb{EDM}^N$ (*confer* (1203))

$$\operatorname{cenv}\operatorname{rank}(-V_{\mathcal{N}}^{\mathrm{T}}DV_{\mathcal{N}}) = \operatorname{cenv}\operatorname{rank}(-VDV) \;\propto\; -\operatorname{tr}(VDV) \tag{1538}$$

and for desired affine dimension $\rho\leq N-1$ and nonnegative H [*sic*] we get a convex optimization problem

$$\begin{array}{cl}\underset{D}{\text{minimize}} & \|\sqrt[\circ]{D}-H\|_{\mathrm{F}}^2\\ \text{subject to} & -\operatorname{tr}(VDV)\leq\kappa\,\rho\\ & D\in\mathbb{EDM}^N\end{array} \tag{1539}$$

where $\kappa\in\mathbb{R}_+$ is a constant determined by cut-and-try. The equivalent semidefinite program makes κ variable: for nonnegative and symmetric H

$$\begin{array}{cl}\underset{D\,,\,Y\,,\,\kappa}{\text{minimize}} & \kappa\,\rho+2\operatorname{tr}(VYV)\\ \text{subject to} & \begin{bmatrix} d_{ij} & y_{ij}\\ y_{ij} & h_{ij}^2\end{bmatrix}\succeq 0\,,\quad N\geq j>i=1\ldots N-1\\ & -\operatorname{tr}(VDV)\leq\kappa\,\rho\\ & Y\in\mathbb{S}_h^N\\ & D\in\mathbb{EDM}^N\end{array} \tag{1540}$$

[7.19] Provided $f\not\equiv+\infty$ and there exists an affine function $h\leq f$ on $\mathbb{R}^n$, then the convex envelope is equal to the convex conjugate (the *Legendre-Fenchel transform*) of the convex conjugate of f; *id est*, the conjugate-conjugate function f^{**}. [230, §E.1] which is the same as (1529), the problem with no explicit constraint on affine dimension. As the present problem is stated, the desired affine dimension ρ yields to the variable scale factor κ ; ρ is effectively ignored.

Yet this result is an illuminant for problem (1529) and it equivalents (all the way back to (1522)): When the given measurement matrix H is nonnegative and symmetric, finding the closest EDM D (as in problem (1522), (1525), or (1529)) implicitly entails minimization of affine dimension (*confer* §5.8.4, §5.14.4). Those non-rank-constrained problems are each inherently equivalent to cenv(rank)-minimization problem (1540), in other words, and their optimal solutions are unique because of the strictly convex objective function in (1522).

7.2.2.3 Rank-heuristic insight

Minimization of affine dimension by use of this trace rank-heuristic (1538) tends to find a list configuration of least energy; rather, it tends to optimize compaction of the reconstruction by minimizing total distance. (1073) It is best used where some physical equilibrium implies such an energy minimization; *e.g*, [399, §5].

For this Problem 2, the trace rank-heuristic arose naturally in the objective in terms of V. We observe: V (in contrast to $V_{\mathcal{N}}^{\mathrm{T}}$) spreads energy over all available distances (§B.4.2 *no.*20, contrast *no.*22) although the rank function itself is insensitive to choice of auxiliary matrix.

Trace rank-heuristic (1534) is useless when a main diagonal is constrained to be constant. Such would be the case were optimization over an elliptope (§5.4.2.2.1), or when the diagonal represents a Boolean vector; *e.g*, §4.2.3.1.1, §4.7.0.0.9.

7.2.2.4 Rank minimization heuristic beyond convex envelope

Fazel, Hindi, & Boyd [157] [457] [158] propose a rank heuristic more potent than trace (1535) for problems of rank minimization;

$$\operatorname{rank} Y \;\leftarrow\; \log\det(Y+\varepsilon I) \tag{1541}$$

the concave surrogate function $\log\det$ in place of quasiconcave $\operatorname{rank} Y$ (§2.9.2.9.2) when $Y\in\mathbb{S}_+^n$ is variable and where ε is a small positive constant. They propose minimization of the surrogate by substituting a sequence comprising infima of a linearized surrogate about the current estimate Y_i ; *id est*, from the first-order Taylor series expansion about Y_i on some open interval of $\|Y\|_2$ (§D.1.7)

$$\log\det(Y+\varepsilon I) \approx \log\det(Y_i+\varepsilon I) + \operatorname{tr}\big((Y_i+\varepsilon I)^{-1}(Y-Y_i)\big) \tag{1542}$$

we make the surrogate sequence of infima over bounded convex feasible set $\mathcal{C}$

$$\arg\inf_{Y\in\mathcal{C}} \operatorname{rank} Y \;\leftarrow\; \lim_{i\to\infty} Y_{i+1} \tag{1543}$$

where, for $i=0\ldots$

$$Y_{i+1} = \arg\inf_{Y\in\mathcal{C}} \operatorname{tr}\big((Y_i+\varepsilon I)^{-1}Y\big) \tag{1544}$$

a matrix analogue to the reweighting scheme disclosed in [239, §4.11.3]. Choosing $Y_0=I$, the first step becomes equivalent to finding the infimum of $\operatorname{tr} Y$; the trace rank-heuristic (1535). The intuition underlying (1544) is the new term in the argument of trace; specifically, $(Y_i+\varepsilon I)^{-1}$ weights Y so that relatively small eigenvalues of Y found by the infimum are made even smaller.

To see that, substitute the nonincreasingly ordered diagonalizations

$$\begin{aligned} Y_i + \varepsilon I &\triangleq Q(\Lambda + \varepsilon I)Q^{\mathrm{T}} && \text{(a)} \\ Y &\triangleq U \Upsilon U^{\mathrm{T}} && \text{(b)} \end{aligned} \tag{1545}$$

into (1544). Then from (1893) we have,

$$\begin{aligned} \inf_{\Upsilon \in U^{\star\mathrm{T}} \mathcal{C} U^{\star}} \delta\big((\Lambda + \varepsilon I)^{-1}\big)^{\mathrm{T}} \delta(\Upsilon) &= \inf_{\Upsilon \in U^{\mathrm{T}} \mathcal{C} U} \; \inf_{R^{\mathrm{T}} = R^{-1}} \operatorname{tr}\big((\Lambda + \varepsilon I)^{-1} R^{\mathrm{T}} \Upsilon R\big) \\ &\leq \inf_{Y \in \mathcal{C}} \operatorname{tr}\big((Y_i + \varepsilon I)^{-1} Y\big) \end{aligned} \tag{1546}$$

where $R \triangleq Q^{\mathrm{T}} U$ in U on the set of orthogonal matrices is a bijection. The role of ε is, therefore, to limit maximum weight; the smallest entry on the main diagonal of Υ gets the largest weight. ♦

7.2.2.5 Applying $\log\det$ rank-heuristic to Problem 2

When the $\log\det$ rank-heuristic is inserted into Problem 2, problem (1540) becomes the problem sequence in i

$$\begin{array}{cl} \underset{D,\,Y,\,\kappa}{\text{minimize}} & \kappa\,\rho + 2\operatorname{tr}(VYV) \\ \text{subject to} & \begin{bmatrix} d_{jl} & y_{jl} \\ y_{jl} & h_{jl}^2 \end{bmatrix} \succeq 0\,, \quad l > j = 1 \ldots N-1 \\ & -\operatorname{tr}\big((-VD_iV + \varepsilon I)^{-1} VDV\big) \leq \kappa\,\rho \\ & Y \in \mathbb{S}_h^N \\ & D \in \mathbb{EDM}^N \end{array} \tag{1547}$$

where $D_{i+1} \triangleq D^\star \in \mathbb{EDM}^N$ and $D_0 \triangleq \mathbf{1}\mathbf{1}^{\mathrm{T}} - I$.

7.2.2.6 Tightening this $\log\det$ rank-heuristic

Like the trace method, this $\log\det$ technique for constraining rank offers no provision for meeting a predetermined upper bound ρ. Yet since eigenvalues are simply determined, $\lambda(Y_i + \varepsilon I) = \delta(\Lambda + \varepsilon I)$, we may certainly force selected weights to ε^{-1} by manipulating diagonalization (1545a). Empirically we find this sometimes leads to better results, although affine dimension of a solution cannot be guaranteed.

7.2.2.7 Cumulative summary of rank heuristics

We have studied a perturbation method of rank reduction in §4.3 as well as the trace heuristic (convex envelope method §7.2.2.1.1) and $\log\det$ heuristic in §7.2.2.4. There is another good contemporary method called LMIRank [323] based on alternating projection (§E.10).[7.20]

7.2.2.7.1 Example. *Unidimensional scaling.*
We apply the convex iteration method from §4.5.1 to numerically solve an instance of Problem 2; a method empirically superior to the foregoing convex envelope and $\log\det$ heuristics for rank regularization and enforcing affine dimension.

Unidimensional scaling, [118] a historically practical application of multidimensional scaling (§5.12), entails solution of an optimization problem having local minima whose

[7.20] that does not solve the *ball packing* problem presented in §5.4.2.2.6.

multiplicity varies as the factorial of point-list cardinality; geometrically, it means reconstructing a list constrained to lie in one affine dimension. Given nonnegative symmetric matrix $H = [h_{ij}] \in \mathbb{S}^N \cap \mathbb{R}_+^{N\times N}$ (1527) whose entries h_{ij} are all known, the nonconvex problem in terms of point list is

$$\underset{\{x_i\in\mathbb{R}\}}{\text{minimize}} \sum_{i\,,\,j=1}^{N} (|x_i - x_j| - h_{ij})^2 \tag{1470}$$

called a *raw stress* problem [57, p.34] which has an implicit constraint on dimensional embedding of points $\{x_i \in \mathbb{R}\,,\ i=1\ldots N\}$. This problem has proven NP-hard; *e.g*, [80].

As always, we first transform variables to distance-square $D\in\mathbb{S}_h^N$; so begin with convex problem (1529) on page 472

$$\begin{array}{cl}\underset{D\,,\,Y}{\text{minimize}} & -\operatorname{tr}\big(V(D-2\,Y)V\big)\\ \text{subject to} & \begin{bmatrix} d_{ij} & y_{ij}\\ y_{ij} & h_{ij}^2\end{bmatrix} \succeq 0\,, \quad N\geq j>i=1\ldots N-1\\ & Y\in\mathbb{S}_h^N\\ & D\in\mathbb{EDM}^N\\ & \operatorname{rank} V_{\mathcal{N}}^{\mathrm T} D V_{\mathcal{N}} = 1\end{array} \tag{1548}$$

that becomes equivalent to (1470) by making explicit the constraint on affine dimension via rank. The iteration is formed by moving the dimensional constraint to the objective:

$$\begin{array}{cl}\underset{D\,,\,Y}{\text{minimize}} & -\langle V(D-2\,Y)V\,,\,I\rangle - w\langle V_{\mathcal{N}}^{\mathrm T} D V_{\mathcal{N}}\,,\,W\rangle\\ \text{subject to} & \begin{bmatrix} d_{ij} & y_{ij}\\ y_{ij} & h_{ij}^2\end{bmatrix} \succeq 0\,, \quad N\geq j>i=1\ldots N-1\\ & Y\in\mathbb{S}_h^N\\ & D\in\mathbb{EDM}^N\end{array} \tag{1549}$$

where w (≈ 10) is a positive scalar just large enough to make $\langle V_{\mathcal{N}}^{\mathrm T} D V_{\mathcal{N}}\,,\,W\rangle$ vanish to within some numerical precision, and where direction matrix W is an optimal solution to semidefinite program (1892a)

$$\begin{array}{cl}\underset{W}{\text{minimize}} & -\langle V_{\mathcal{N}}^{\mathrm T} D^\star V_{\mathcal{N}}\,,\,W\rangle\\ \text{subject to} & 0\preceq W\preceq I\\ & \operatorname{tr} W = N-1\end{array} \tag{1550}$$

one of which is known in closed form. Semidefinite programs (1549) and (1550) are iterated until convergence in the sense defined on page 250. This iteration is not a projection method. (§4.5.1.1) Convex problem (1549) is neither a relaxation of unidimensional scaling problem (1548); instead, problem (1549) is a convex equivalent to (1548) at convergence of the iteration.

Jan de Leeuw provided us with some test data

$$H = \begin{bmatrix} 0.000000 & 5.235301 & 5.499274 & 6.404294 & 6.486829 & 6.263265\\ 5.235301 & 0.000000 & 3.208028 & 5.840931 & 3.559010 & 5.353489\\ 5.499274 & 3.208028 & 0.000000 & 5.679550 & 4.020339 & 5.239842\\ 6.404294 & 5.840931 & 5.679550 & 0.000000 & 4.862884 & 4.543120\\ 6.486829 & 3.559010 & 4.020339 & 4.862884 & 0.000000 & 4.618718\\ 6.263265 & 5.353489 & 5.239842 & 4.543120 & 4.618718 & 0.000000 \end{bmatrix} \tag{1551}$$

and a globally optimal solution

$$\begin{aligned} X^\star &= [\,-4.981494 \quad -2.121026 \quad -1.038738 \quad 4.555130 \quad 0.764096 \quad 2.822032\,] \\ &= [\quad x_1^\star \qquad x_2^\star \qquad x_3^\star \qquad x_4^\star \qquad x_5^\star \qquad x_6^\star \quad] \end{aligned} \tag{1552}$$

found by searching $6!$ local minima of (1470) [118]. By iterating convex problems (1549) and (1550) about twenty times (initial $W=\mathbf{0}$) we find global infimum 98.12812 to stress problem (1470), and by (1293) we find a corresponding one-dimensional point list that is a rigid transformation in $\mathbb{R}$ of $X^\star$.

Here we found the infimum to accuracy of the given data, but that ceases to hold as problem size increases. Because of machine numerical precision and an interior-point method of solution, we speculate, accuracy degrades quickly as problem size increases beyond this. □

7.3 Third prevalent problem: Projection on EDM cone in d_{ij}

In summary, we find that the solution to problem [(1472.3) p.461] *is difficult and depends on the dimension of the space as the geometry of the cone of EDMs becomes more complex.*

−Hayden, Wells, Liu, & Tarazaga, 1991 [216, §3]

Reformulating Problem 2 (p.470), in terms of EDM D, changes it considerably:

$$\left.\begin{array}{cl} \underset{D}{\text{minimize}} & \|D-H\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \leq \rho \\ & D \in \mathbb{EDM}^N \end{array}\right\} \quad \text{Problem 3} \tag{1553}$$

This third prevalent proximity problem is a Euclidean projection of given matrix H on a generally nonconvex subset ($\rho < N-1$) of $\partial\mathbb{EDM}^N$ the boundary of the convex cone of Euclidean distance matrices relative to subspace $\mathbb{S}_h^N$ (Figure **165**d). Because coordinates of projection are distance-square and H now presumably holds distance-square measurements, numerical solution to Problem 3 is generally different than that of Problem 2.

For the moment, we need make no assumptions regarding measurement matrix H.

7.3.1 Convex case

$$\begin{array}{cl} \underset{D}{\text{minimize}} & \|D-H\|_{\mathrm{F}}^2 \\ \text{subject to} & D \in \mathbb{EDM}^N \end{array} \tag{1554}$$

When the rank constraint disappears (for $\rho = N-1$), this third problem becomes obviously convex because the feasible set is then the entire EDM cone and because the objective function

$$\|D-H\|_{\mathrm{F}}^2 = \sum_{i,j}(d_{ij}-h_{ij})^2 \tag{1555}$$

is a strictly convex quadratic in D;[7.21]

$$\begin{array}{ll}\underset{D}{\text{minimize}} & \sum_{i,j} d_{ij}^2 - 2h_{ij}\,d_{ij} \,+\, h_{ij}^2\\ \text{subject to} & D \in \mathbb{EDM}^N\end{array} \tag{1556}$$

Optimal solution $D^\star$ is therefore unique, as expected, for this simple projection on the EDM cone equivalent to (1471).

7.3.1.1 Equivalent semidefinite program, Problem 3, convex case

In the past, this convex problem was solved numerically by means of alternating projection. (Example 7.3.1.1.1) [180] [172] [216, §1] We translate (1556) to an equivalent semidefinite program because we have a good solver:

Assume the given measurement matrix H to be nonnegative and symmetric;[7.22]

$$H = [h_{ij}] \in \mathbb{S}^N \cap \mathbb{R}_+^{N\times N} \qquad (1527)$$

We then propose: Problem (1556) is equivalent to the semidefinite program, for

$$\partial \triangleq [d_{ij}^2] = D\circ D \tag{1557}$$

a matrix of distance-square squared,

$$\begin{array}{ll}\underset{\partial\;,\;D}{\text{minimize}} & -\operatorname{tr}\big(V(\partial - 2\,H\circ D)V\big)\\ \text{subject to} & \begin{bmatrix}\partial_{ij} & d_{ij}\\ d_{ij} & 1\end{bmatrix}\succeq 0\,, \quad N\geq j > i = 1\ldots N-1\\ & D\in\mathbb{EDM}^N\\ & \partial\in\mathbb{S}_h^N\end{array} \tag{1558}$$

where

$$\begin{bmatrix}\partial_{ij} & d_{ij}\\ d_{ij} & 1\end{bmatrix}\succeq 0 \;\Leftrightarrow\; \partial_{ij}\geq d_{ij}^2 \tag{1559}$$

Symmetry of input H facilitates trace in the objective (§B.4.2 *no.*20), while its nonnegativity causes $\partial_{ij}\rightarrow d_{ij}^2$ as optimization proceeds.

7.3.1.1.1 Example. *Alternating projection on nearest EDM.*
By solving (1558) we confirm the result from an example given by Glunt, Hayden, Hong, & Wells [180, §6] who found analytical solution to convex optimization problem (1554) for particular cardinality $N=3$ by using the alternating projection method of von Neumann (§E.10):

$$H = \begin{bmatrix}0&1&1\\1&0&9\\1&9&0\end{bmatrix}, \qquad D^\star = \begin{bmatrix}0&\frac{19}{9}&\frac{19}{9}\\ \frac{19}{9}&0&\frac{76}{9}\\ \frac{19}{9}&\frac{76}{9}&0\end{bmatrix} \tag{1560}$$

[7.21] For nonzero $Y\in\mathbb{S}_h^N$ and some open interval of $t\in\mathbb{R}$ (§3.13.0.0.2, §D.2.3)

$$\frac{d^2}{dt^2}\|(D+t\,Y)-H\|_{\mathrm F}^2 \;=\; 2\operatorname{tr}Y^{\mathrm T}Y > 0 \qquad ♦$$

[7.22] If that H given has negative entries, then the technique of solution presented here becomes invalid. Projection of H on $\mathcal{K}$ (1463) prior to application of this proposed technique, as explained in §7.0.1, is incorrect.

The problem (1554), of projecting H on the EDM cone, is transformed to an equivalent iterative sequence of projections on the two convex cones (1412) from §6.8.1.1. Utilizing projector (1415) in an ordinary alternating projection, input H goes to $D^\star$ with an accuracy of four decimal places in about 17 iterations. Affine dimension corresponding to this optimal solution is $r=1$.

Obviation of semidefinite programming's computational expense is the principal advantage of this alternating projection technique. □

7.3.1.2 Schur-form semidefinite program, Problem 3 convex case

Semidefinite program (1558) can be reformulated by moving the objective function in

$$\begin{array}{cl}\underset{D}{\text{minimize}} & \|D-H\|_{\mathrm F}^2\\ \text{subject to} & D\in\mathbb{EDM}^N\end{array}\qquad(1554)$$

to the constraints. This makes an equivalent epigraph form of the problem: for any measurement matrix H

$$\begin{array}{cl}\underset{t\in\mathbb{R}\;,\;D}{\text{minimize}} & t\\ \text{subject to} & \|D-H\|_{\mathrm F}^2\leq t\\ & D\in\mathbb{EDM}^N\end{array}\qquad(1561)$$

We can transform this problem to an equivalent Schur-form semidefinite program; (§3.5.3)

$$\begin{array}{cl}\underset{t\in\mathbb{R}\;,\;D}{\text{minimize}} & t\\ \text{subject to} & \left[\begin{array}{cc}tI & \operatorname{vec}(D-H)\\ \operatorname{vec}(D-H)^{\mathrm T} & 1\end{array}\right]\succeq 0\\ & D\in\mathbb{EDM}^N\end{array}\qquad(1562)$$

characterized by great sparsity and structure. The advantage of this SDP is lack of conditions on input H; *e.g*, negative entries would invalidate any solution provided by (1558). (§7.0.1.2)

7.3.1.3 Gram-form semidefinite program, Problem 3 convex case

Further, this problem statement may be equivalently written in terms of a Gram matrix via linear bijective (§5.6.1) EDM operator $\mathbf{D}(G)$ (1061);

$$\begin{array}{cl}\underset{G\in\mathbb{S}_c^N\;,\;t\in\mathbb{R}}{\text{minimize}} & t\\ \text{subject to} & \left[\begin{array}{cc}tI & \operatorname{vec}(\mathbf{D}(G)-H)\\ \operatorname{vec}(\mathbf{D}(G)-H)^{\mathrm T} & 1\end{array}\right]\succeq 0\\ & G\succeq 0\end{array}\qquad(1563)$$

To include constraints on the list $X\in\mathbb{R}^{n\times N}$, we would rewrite this:

$$\begin{array}{cl}\underset{G\in\mathbb{S}_c^N\;,\;t\in\mathbb{R}\;,\;X\in\mathbb{R}^{n\times N}}{\text{minimize}} & t\\ \text{subject to} & \left[\begin{array}{cc}tI & \operatorname{vec}(\mathbf{D}(G)-H)\\ \operatorname{vec}(\mathbf{D}(G)-H)^{\mathrm T} & 1\end{array}\right]\succeq 0\\ & \left[\begin{array}{cc}I & X\\ X^{\mathrm T} & G\end{array}\right]\succeq 0\\ & X\in\mathcal{C}\end{array}\qquad(1564)$$

where $\mathcal{C}$ is some abstract convex set. This technique is discussed in §5.4.2.2.8.

7.3.1.4 Dual interpretation, projection on EDM cone

From §E.9.1.2 we learn that projection on a convex cone has dual form. In the circumstance that $\mathcal{K}$ is a convex cone and point x exists exterior to the cone or on its boundary, distance to the nearest point Px in $\mathcal{K}$ is found as the optimal value of the objective

$$\begin{array}{rll}\|x-Px\| = \underset{a}{\text{maximize}} & a^{\mathrm T}x \\ \text{subject to} & \|a\|\leq 1 \\ & a\in\mathcal{K}^{\circ}\end{array} \tag{2235}$$

where $\mathcal{K}^{\circ}$ is the polar cone.

Applying this result to (1554), we get a convex optimization for any given symmetric matrix H exterior to or on the EDM cone boundary:

$$\begin{array}{ll}\underset{D}{\text{minimize}} & \|D-H\|_{\mathrm F}^2 \\ \text{subject to} & D\in\mathbb{EDM}^N\end{array} \quad\equiv\quad \begin{array}{ll}\underset{A^{\circ}}{\text{maximize}} & \langle A^{\circ},\, H\rangle \\ \text{subject to} & \|A^{\circ}\|_{\mathrm F}\leq 1 \\ & A^{\circ}\in\mathbb{EDM}^{N^{\circ}}\end{array} \tag{1565}$$

Then, from (2237), projection of H on cone $\mathbb{EDM}^N$ is

$$D^{\star} = H - A^{\circ\star}\langle A^{\circ\star},\, H\rangle \tag{1566}$$

Critchley proposed, instead, projection on the polar EDM cone in his 1980 thesis [98, p.113]: In that circumstance, by projection on the algebraic complement (§E.9.2.2.1),

$$D^{\star} = A^{\star}\langle A^{\star},\, H\rangle \tag{1567}$$

which is equal to (1566) when $A^{\star}$ solves

$$\begin{array}{ll}\underset{A}{\text{maximize}} & \langle A\,,\, H\rangle \\ \text{subject to} & \|A\|_{\mathrm F}=1 \\ & A\in\mathbb{EDM}^N\end{array} \tag{1568}$$

This projection of symmetric H on polar cone $\mathbb{EDM}^{N^{\circ}}$ can be made a convex problem, of course, by relaxing the equality constraint $(\|A\|_{\mathrm F}\leq 1)$.

7.3.2 Minimization of affine dimension in Problem 3

When desired affine dimension ρ is diminished, Problem 3 (1553) is difficult to solve [216, §3] because the feasible set in $\mathbb{R}^{N(N-1)/2}$ loses convexity. By substituting rank envelope (1538) into Problem 3, then for any given H we get a convex problem

$$\begin{array}{ll}\underset{D}{\text{minimize}} & \|D-H\|_{\mathrm F}^2 \\ \text{subject to} & -\operatorname{tr}(VDV)\leq\kappa\,\rho \\ & D\in\mathbb{EDM}^N\end{array} \tag{1569}$$

where $\kappa\in\mathbb{R}_+$ is a constant determined by cut-and-try. Given κ, problem (1569) is a convex optimization having unique solution in any desired affine dimension ρ; an approximation to Euclidean projection on that nonconvex subset of the EDM cone containing EDMs with corresponding affine dimension no greater than ρ.

The SDP equivalent to (1569) does not move κ into the variables as on page 474: for nonnegative symmetric input H and distance-square squared variable ∂ as in (1557),

$$\begin{array}{cl}\underset{\partial\,,\,D}{\text{minimize}} & -\operatorname{tr}\big(V(\partial-2\,H\circ D)V\big)\\ \text{subject to} & \begin{bmatrix}\partial_{ij} & d_{ij}\\ d_{ij} & 1\end{bmatrix}\succeq 0\,,\quad N\geq j>i=1\ldots N-1\\ & -\operatorname{tr}(V D\, V)\leq\kappa\,\rho\\ & D\in\mathbb{EDM}^N\\ & \partial\in\mathbb{S}_h^N\end{array}\tag{1570}$$

That means we will not see equivalence of this cenv(rank)-minimization problem to the non-rank-constrained problems (1556) and (1558) like we saw for its counterpart (1540) in Problem 2.

Another approach to affine dimension minimization is to project instead on the polar EDM cone; discussed in §6.8.1.5.

7.3.3 Constrained affine dimension, Problem 3

When one desires affine dimension diminished further below what can be achieved via cenv(rank)-minimization as in (1570), spectral projection can be considered a natural means in light of its successful application to projection on a rank ρ subset of a positive semidefinite cone in §7.1.4.

Yet it is wrong here to zero eigenvalues of $-VDV$ or $-VGV$ or a variant to reduce affine dimension, because that particular method comes from projection on a positive semidefinite cone (1499); zeroing those eigenvalues here in Problem 3 would place an elbow in the projection path (Figure **180**) thereby producing a result that is necessarily suboptimal. Problem 3 is instead a projection on the EDM cone whose associated spectral cone is considerably different. (§5.11.2.3) Proper choice of spectral cone is demanded by diagonalization of that variable argument to the objective:

7.3.3.1 Cayley-Menger form

We use Cayley-Menger composition of the Euclidean distance matrix to solve a problem that is the same as Problem 3 (1553): (§5.7.3.0.1)

$$\begin{array}{cl}\underset{D}{\text{minimize}} & \left\|\begin{bmatrix}0 & \mathbf{1}^{\mathrm T}\\ \mathbf{1} & -D\end{bmatrix}-\begin{bmatrix}0 & \mathbf{1}^{\mathrm T}\\ \mathbf{1} & -H\end{bmatrix}\right\|_{\mathrm F}^2\\ \text{subject to} & \operatorname{rank}\begin{bmatrix}0 & \mathbf{1}^{\mathrm T}\\ \mathbf{1} & -D\end{bmatrix}\leq\,\rho+2\\ & D\in\mathbb{EDM}^N\end{array}\tag{1571}$$

a projection of H on a generally nonconvex subset (when $\rho<N-1$) of the Euclidean distance matrix cone boundary $\operatorname{rel}\partial\mathbb{EDM}^N$; *id est*, projection from the EDM cone interior or exterior on a subset of its relative boundary (§6.5, (1341)).

Rank of an optimal solution is intrinsically bounded above and below;

$$2\,\leq\,\operatorname{rank}\begin{bmatrix}0 & \mathbf{1}^{\mathrm T}\\ \mathbf{1} & -D^\star\end{bmatrix}\leq\,\rho+2\,\leq\,N+1\tag{1572}$$

Our proposed strategy for low-rank solution is projection on that subset of a spectral cone $\lambda\left(\begin{bmatrix}0 & \mathbf{1}^{\mathrm T}\\ \mathbf{1} & -\mathbb{EDM}^N\end{bmatrix}\right)$ (§5.11.2.3) corresponding to affine dimension not in excess of that ρ

desired; *id est*, spectral projection on

$$\begin{bmatrix} \mathbb{R}_+^{\rho+1} \\ \mathbf{0} \\ \mathbb{R}_- \end{bmatrix} \cap \partial\mathcal{H} \subset \mathbb{R}^{N+1} \tag{1573}$$

where

$$\partial\mathcal{H} = \{\lambda \in \mathbb{R}^{N+1} \mid \mathbf{1}^{\mathrm{T}}\lambda = 0\} \qquad (1273)$$

is a hyperplane through the origin. This pointed polyhedral cone (1573), to which membership subsumes the rank constraint, is not full-dimensional.

Given desired affine dimension $0 \le \rho \le N-1$ and diagonalization (§A.5) of unknown EDM D

$$\begin{bmatrix} 0 & \mathbf{1}^{\mathrm{T}} \\ \mathbf{1} & -D \end{bmatrix} \triangleq U\Upsilon U^{\mathrm{T}} \in \mathbb{S}_h^{N+1} \tag{1574}$$

and given symmetric H in diagonalization

$$\begin{bmatrix} 0 & \mathbf{1}^{\mathrm{T}} \\ \mathbf{1} & -H \end{bmatrix} \triangleq Q\Lambda Q^{\mathrm{T}} \in \mathbb{S}^{N+1} \tag{1575}$$

having eigenvalues arranged in nonincreasing order, then by (1286) problem (1571) is equivalent to

$$\begin{array}{cl} \underset{\Upsilon,\,R}{\text{minimize}} & \left\|\delta(\Upsilon) - \pi\big(\delta(R^{\mathrm{T}}\Lambda R)\big)\right\|^2 \\ \text{subject to} & \delta(\Upsilon) \in \begin{bmatrix} \mathbb{R}_+^{\rho+1} \\ \mathbf{0} \\ \mathbb{R}_- \end{bmatrix} \cap \partial\mathcal{H} \\ & \delta(QR\Upsilon R^{\mathrm{T}}Q^{\mathrm{T}}) = \mathbf{0} \\ & R^{-1} = R^{\mathrm{T}} \end{array} \tag{1576}$$

where π is the permutation operator from §7.1.3 arranging its vector argument in nonincreasing order,[7.23] where

$$R \triangleq Q^{\mathrm{T}}U \in \mathbb{R}^{N+1\times N+1} \tag{1577}$$

in U on the set of orthogonal matrices is a bijection, and where $\partial\mathcal{H}$ insures one negative eigenvalue. Hollowness constraint $\delta(QR\Upsilon R^{\mathrm{T}}Q^{\mathrm{T}}) = \mathbf{0}$ makes problem (1576) difficult by making the two variables dependent.

Our plan is to instead divide problem (1576) into two and then alternate their solution:

$$\begin{array}{cll} \underset{\Upsilon}{\text{minimize}} & \left\|\delta(\Upsilon) - \pi\big(\delta(R^{\mathrm{T}}\Lambda R)\big)\right\|^2 & \\ \text{subject to} & \delta(\Upsilon) \in \begin{bmatrix} \mathbb{R}_+^{\rho+1} \\ \mathbf{0} \\ \mathbb{R}_- \end{bmatrix} \cap \partial\mathcal{H} & \text{(a)} \\ & & \\ \underset{R}{\text{minimize}} & \|R\Upsilon^\star R^{\mathrm{T}} - \Lambda\|_{\mathrm{F}}^2 & \\ \text{subject to} & \delta(QR\Upsilon^\star R^{\mathrm{T}}Q^{\mathrm{T}}) = \mathbf{0} & \text{(b)} \\ & R^{-1} = R^{\mathrm{T}} & \end{array} \tag{1578}$$

[7.23] Recall, any permutation matrix is an orthogonal matrix.

Proof. We justify disappearance of the hollowness constraint in convex optimization problem (1578a): From arguments in §7.1.3 with regard to permutation operator π, cone membership constraint $\delta(\Upsilon)\in\begin{bmatrix}\mathbb{R}_+^{\rho+1}\\ \mathbf{0}\\ \mathbb{R}_-\end{bmatrix}\cap\partial\mathcal{H}$ from (1578a) is equivalent to

$$\delta(\Upsilon)\in\begin{bmatrix}\mathbb{R}_+^{\rho+1}\\ \mathbf{0}\\ \mathbb{R}_-\end{bmatrix}\cap\,\partial\mathcal{H}\,\cap\,\mathcal{K}_{\mathcal{M}} \tag{1579}$$

where $\mathcal{K}_{\mathcal{M}}$ is the monotone cone (§2.13.10.4.3). Membership of $\delta(\Upsilon)$ to the *polyhedral cone of majorization* (Theorem A.1.2.0.1)

$$\mathcal{K}_{\lambda\delta}^* = \partial\mathcal{H}\cap\mathcal{K}_{\mathcal{M}+}^* \qquad (1596)$$

where $\mathcal{K}_{\mathcal{M}+}^*$ is the dual monotone nonnegative cone (§2.13.10.4.2), is a condition (in absence of a hollowness constraint) that would insure existence of a symmetric hollow matrix $\begin{bmatrix}0&\mathbf{1}^{\rm T}\\ \mathbf{1}&-D\end{bmatrix}$. Curiously, intersection of this feasible superset $\begin{bmatrix}\mathbb{R}_+^{\rho+1}\\ \mathbf{0}\\ \mathbb{R}_-\end{bmatrix}\cap\partial\mathcal{H}\cap\mathcal{K}_{\mathcal{M}}$ from (1579) with the cone of majorization $\mathcal{K}_{\lambda\delta}^*$ is a benign operation; *id est*,

$$\partial\mathcal{H}\,\cap\,\mathcal{K}_{\mathcal{M}+}^*\cap\,\mathcal{K}_{\mathcal{M}} \;=\; \partial\mathcal{H}\cap\mathcal{K}_{\mathcal{M}} \tag{1580}$$

verifiable by observing conic dependencies (§2.10.3) among the aggregate of halfspace-description normals. The cone membership constraint in (1578a) therefore inherently insures existence of a symmetric hollow matrix. ♦

Optimization (1578b) would be a Procrustes problem (§C.4) were it not for the hollowness constraint. It is, instead, a minimization over the intersection of the nonconvex manifold of orthogonal matrices with another nonconvex set in variable R specified by the hollowness constraint. We solve problem (1578b) by a method introduced in §4.7.0.0.2: Define $R=[\,r_1\cdots r_{N+1}\,]\in\mathbb{R}^{N+1\times N+1}$ and make the assignment

$$\begin{aligned}G &= \begin{bmatrix}r_1\\ \vdots\\ r_{N+1}\\ 1\end{bmatrix}[\,r_1^{\rm T}\cdots\; r_{N+1}^{\rm T}\;\;1\,] \qquad\qquad \in\mathbb{S}^{(N+1)^2+1}\\ &= \begin{bmatrix}R_{11}&\cdots&R_{1,N+1}&r_1\\ \vdots&\ddots&&\vdots\\ R_{1,N+1}^{\rm T}&&R_{N+1,N+1}&r_{N+1}\\ r_1^{\rm T}&\cdots&r_{N+1}^{\rm T}&1\end{bmatrix}\triangleq\begin{bmatrix}r_1r_1^{\rm T}&\cdots&r_1r_{N+1}^{\rm T}&r_1\\ \vdots&\ddots&&\vdots\\ r_{N+1}r_1^{\rm T}&&r_{N+1}r_{N+1}^{\rm T}&r_{N+1}\\ r_1^{\rm T}&\cdots&r_{N+1}^{\rm T}&1\end{bmatrix}\end{aligned} \tag{1581}$$

where $R_{ij}\triangleq r_ir_j^{\rm T}\in\mathbb{R}^{N+1\times N+1}$ and $\Upsilon_{ii}^\star\in\mathbb{R}$. Since $R\,\Upsilon^\star R^{\rm T}=\sum\limits_{i=1}^{N+1}\Upsilon_{ii}^\star R_{ii}$, then problem

(1578b) is equivalently expressed:

$$\begin{array}{cll}\underset{R_{ii}\in\mathbb{S}\,,\;R_{ij}\,,\;r_i}{\text{minimize}} & \left\|\sum_{i=1}^{N+1}\Upsilon_{ii}^\star R_{ii}-\Lambda\right\|_{\rm F}^2 & \\ \text{subject to} & \operatorname{tr}R_{ii}=1\,, & i=1\ldots N+1\\ & \operatorname{tr}R_{ij}=0\,, & i<j=2\ldots N+1\\ & G=\begin{bmatrix}R_{11} & \cdots & R_{1,N+1} & r_1\\ \vdots & \ddots & & \vdots\\ R_{1,N+1}^{\rm T} & & R_{N+1,N+1} & r_{N+1}\\ r_1^{\rm T} & \cdots & r_{N+1}^{\rm T} & 1\end{bmatrix}(\succeq 0) & \\ & \delta\left(Q\sum_{i=1}^{N+1}\Upsilon_{ii}^\star R_{ii}\,Q^{\rm T}\right)=\mathbf{0} & \\ & \operatorname{rank}G=1 & \end{array} \tag{1582}$$

The rank constraint is regularized by method of convex iteration developed in §4.5. Problem (1582) is partitioned into two convex problems:

$$\begin{array}{cll}\underset{R_{ij}\,,\;r_i}{\text{minimize}} & \left\|\sum_{i=1}^{N+1}\Upsilon_{ii}^\star R_{ii}-\Lambda\right\|_{\rm F}^2+\langle G\,,W\rangle & \\ \text{subject to} & \operatorname{tr}R_{ii}=1\,, & i=1\ldots N+1\\ & \operatorname{tr}R_{ij}=0\,, & i<j=2\ldots N+1\\ & G=\begin{bmatrix}R_{11} & \cdots & R_{1,N+1} & r_1\\ \vdots & \ddots & & \vdots\\ R_{1,N+1}^{\rm T} & & R_{N+1,N+1} & r_{N+1}\\ r_1^{\rm T} & \cdots & r_{N+1}^{\rm T} & 1\end{bmatrix}\succeq 0 & \\ & \delta\left(Q\sum_{i=1}^{N+1}\Upsilon_{ii}^\star R_{ii}\,Q^{\rm T}\right)=\mathbf{0} & \end{array} \tag{1583}$$

and

$$\begin{array}{cl}\underset{W\in\,\mathbb{S}^{(N+1)^2+1}}{\text{minimize}} & \langle G^\star,\,W\rangle\\ \text{subject to} & 0\preceq W\preceq I\\ & \operatorname{tr}W=(N+1)^2\end{array} \tag{1584}$$

then alternated with convex problem (1578a) until a rank-1 G matrix is found and the objective of (1578a) is minimized.[7.24] An optimal solution to (1584) is known in closed form (p.539).

7.4 Conclusion

The importance and application of solving rank- or cardinality-constrained problems are enormous, a conclusion generally accepted *gratis* by the mathematics and engineering communities. Rank-constrained semidefinite programs arise in many vital feedback and control problems [201], optics [82] (Figure **138**), and communications [165] [289] (Figure **116**). *For example, one might be interested in the minimal order dynamic output feedback which stabilizes a given linear time invariant plant* (*this problem is considered to be among the most important open problems in control*). −[299] Rank and cardinality constraints also arise naturally in combinatorial optimization (§4.7.0.0.11, Figure **127**), and find application to facial recognition (Figure **6**), cartography (Figure **161**), latent

[7.24] The hollowness constraint in (1583) may cause numerical instability; in that case, it may be moved to the objective within an added weighted norm. Conditions for termination of the iteration would then comprise a vanishing norm of hollowness.

semantic indexing [264], sparse or low-rank matrix completion for preference models and collaborative filtering, multidimensional scaling or principal component analysis (§5.12), medical imaging (Figure **122**), digital filter design with time domain constraints [427], molecular conformation (Figure **154**), sensor-network localization and wireless location (Figure **99**), *etcetera.*

There has been little progress in spectral projection since the discovery by Eckart & Young in 1936 [149] leading to a formula for projection on a rank ρ subset of a positive semidefinite cone (§2.9.2.1). [170] The only closed-form spectral method presently available for solving proximity problems, having a constraint on rank, is based on their discovery (Problem 1, §7.1, §5.13).

- One popular recourse is intentional misapplication of Eckart & Young's result by introducing spectral projection on a positive semidefinite cone into Problem 3 via $\mathbf{D}(G)$ (1061), for example. [80] Since Problem 3 instead demands spectral projection on the EDM cone, any solution acquired that way is necessarily suboptimal.

- A second recourse is problem redesign: A presupposition to all proximity problems in this chapter is that matrix H is given. We considered H having various properties such as nonnegativity, symmetry, hollowness, or lack thereof. It was assumed that if H did not already belong to the EDM cone, then we wanted an EDM closest to H in some sense; *id est*, input-matrix H was assumed corrupted somehow. For practical problems, it withstands reason that such a proximity problem could instead be reformulated so that some or all entries of H were unknown but bounded above and below by known limits; the norm objective is thereby eliminated as in the development beginning on page 261. That particular redesign (*the art,* p.8), in terms of the Gram-matrix bridge between point list X and EDM D, at once encompasses proximity and completion problems.

- A third recourse is to apply the method of convex iteration as we did in §7.2.2.7.1. This technique is applicable to any semidefinite problem requiring a rank constraint; it places a regularization term in the objective that enforces the rank constraint.