Chapter 1

Overview

Convex Optimization
Euclidean Distance Geometry 2ε

People are so afraid of convex analysis.

–Claude Lemaréchal, 2003

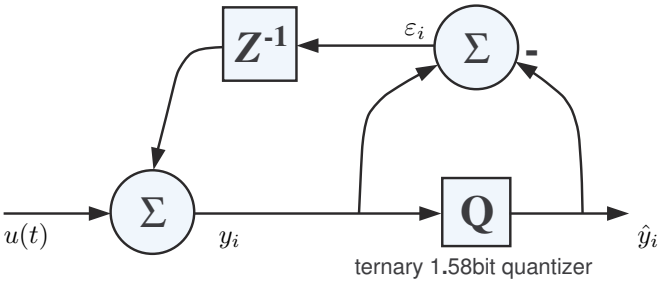
In layman's terms, the mathematical science of Optimization is a study of how to make good choices when confronted with conflicting requirements and demands. Optimization is a relatively new wisdom, historically, that can represent balance of real things. The qualifier *convex* means: when an optimal solution is found, then it is guaranteed to be a best solution; there is no better choice.

Any convex optimization problem has geometric interpretation. If a given optimization problem can be transformed to a convex equivalent, then this interpretive benefit is acquired. That is a powerful attraction: the ability to visualize geometry of an optimization problem. Conversely, recent advances in geometry and in graph theory hold convex optimization within their proofs' core. [465] [362]

This book is about convex optimization, convex geometry (with particular attention to distance geometry), and nonconvex, combinatorial, and geometrical problems that can be relaxed or transformed into convexity. A virtual flood of new applications follows by epiphany that many problems, presumed nonconvex, can be so transformed: [11] [12] [36, §4.3, p.316-322] [64] [103] [173] [176] [315] [340] [348] [408] [409] [461] [465] *e.g*, sigma delta analog-to-digital audio converter (A/D) antialiasing (Figure **1**).

Euclidean distance geometry is, fundamentally, a determination of point conformation (configuration, relative position or location) by inference from interpoint distance information. By *inference* we mean: *e.g*, given only distance information, determine whether there corresponds a *realizable* conformation of points; a *list* of points in some dimension that attains the given interpoint distances. Each point may represent simply location or, abstractly, any entity expressible as a vector in finite-dimensional Euclidean space; *e.g*, distance geometry of music [121].

It is a common misconception to presume that some desired point conformation cannot be recovered in absence of complete interpoint distance information. We might, for example, want to realize a constellation given only interstellar distance (or, equivalently, parsecs from our Sun and relative angular measurement; the Sun as vertex to two distant stars); called *stellar cartography,* an application evoked by Figure **3**. At first it may seem



$$\begin{array}{rl} \text{find} & u \\ \text{subject to} & R(\hat{y}-u) \preceq \frac{1}{3}\mathbf{1} \\ & R(u-\hat{y}) \preceq \frac{1}{3}\mathbf{1} \end{array}$$

Figure 1: Multibit sigma delta quantization is predominant technology for analog to digital audio signal conversion. [2, p.6] Input signal $u(t)$ is continuous. Delay z^{-1} here is analog, perhaps implemented by sample/hold circuit at MHz rate of $\hat{y}_i$ samples. Observing vector $\hat{y}$, signal u can be reconstructed by finding a point feasible to the set of linear inequalities representing this coarse quantizer recursion. R is a lower triangular matrix of ones. [111]

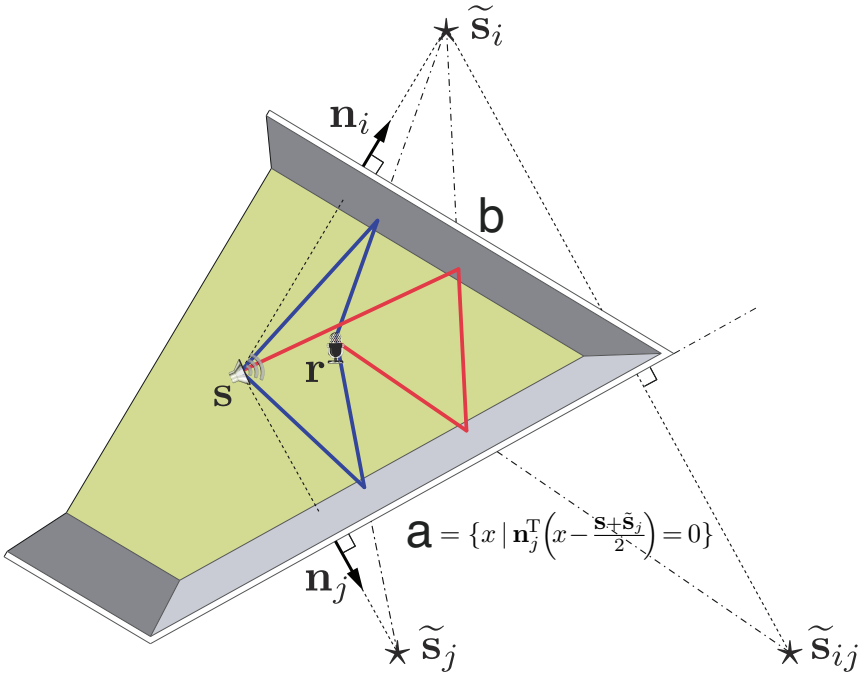


Figure 2: [133] [328] [130] Dokmanić & Parhizkar *et alii* discover an audio signal processing application of Euclidean distance matrices to room geometry estimation by discerning first acoustic reflections of stationary sound source $\mathbf{s}$. Locations of source and phantom $\star$ sources $\tilde{\mathbf{s}}_i$ and $\tilde{\mathbf{s}}_j$ are ascertained by measuring arrival times of first echoes (blue) at multiple microphone receivers. (Only one receiver $\mathbf{r}$ is illustrated. Second reflection (red) phantom $\tilde{\mathbf{s}}_{ij}$ ignored.) Phantom location is invariant to receiver position. All interpoint distances among receivers are known. Once source and phantoms are localized, normals $\mathbf{n}_j$ and $\mathbf{n}_i$ respectively identify truncated hyperplanes (walls) $\mathbf{a}$ and $\mathbf{b}$ bisecting perpendicular line segment connecting source $\mathbf{s}$ to a phantom.

Figure 3: *Orion nebula*. (Astrophotography by Massimo Robberto.)

that $\mathrm{O}(N^2)$ data is required, yet there are many circumstances where this can be reduced to $\mathrm{O}(N)$.

If we agree that a set of points may have a shape (three points can form a triangle and its interior, for example, four points a tetrahedron), then we can ascribe *shape* of a set of points to their convex hull. It should be apparent: from distance, these shapes can be determined only to within a *rigid transformation* (rotation, reflection, translation).

Absolute position information is generally lost, given only distance information, but we can determine the smallest possible dimension in which an unknown list of points can exist; that attribute is their *affine dimension* (a triangle in any ambient space has affine dimension 2, for example). In circumstances where stationary reference points are also provided, it becomes possible to determine absolute position or location; *e.g*, Figure **4**.

Geometric problems involving distance between points can sometimes be reduced to convex optimization problems. Mathematics of this combined study of geometry and optimization is rich and deep. Its application has already proven invaluable discerning organic *molecular conformation* by measuring interatomic distance along covalent bonds; *e.g*, Figure **5**. [97] [398] [160] [50] Many disciplines have already benefitted and simplified consequent to this theory; *e.g*, distance based *pattern recognition* (Figure **6**), *localization* in wireless sensor networks [51] [459] [49] by measurement of intersensor distance along channels of communication, *wireless location* of a radio-signal source such as cell phone by multiple measurements of signal strength, the *global positioning system* (GPS), *multidimensional scaling* (§5.12) which is a numerical representation of qualitative data by finding a low-dimensional scale, and audio signal processing: ultrasound tomography, room geometry estimation (Figure **2**), and perhaps dereverberation by localization of phantom sound sources [131] [130] [133]. [132]

Euclidean distance geometry provides some foundation for *artificial intelligence*. Together with convex optimization, distance geometry has found application to:

- *machine learning* by discerning naturally occurring manifolds in:
  - Euclidean bodies (Figure **7**, §6.7.0.0.1)
  - Fourier spectra of kindred utterances [244]
  - photographic image sequences [443]

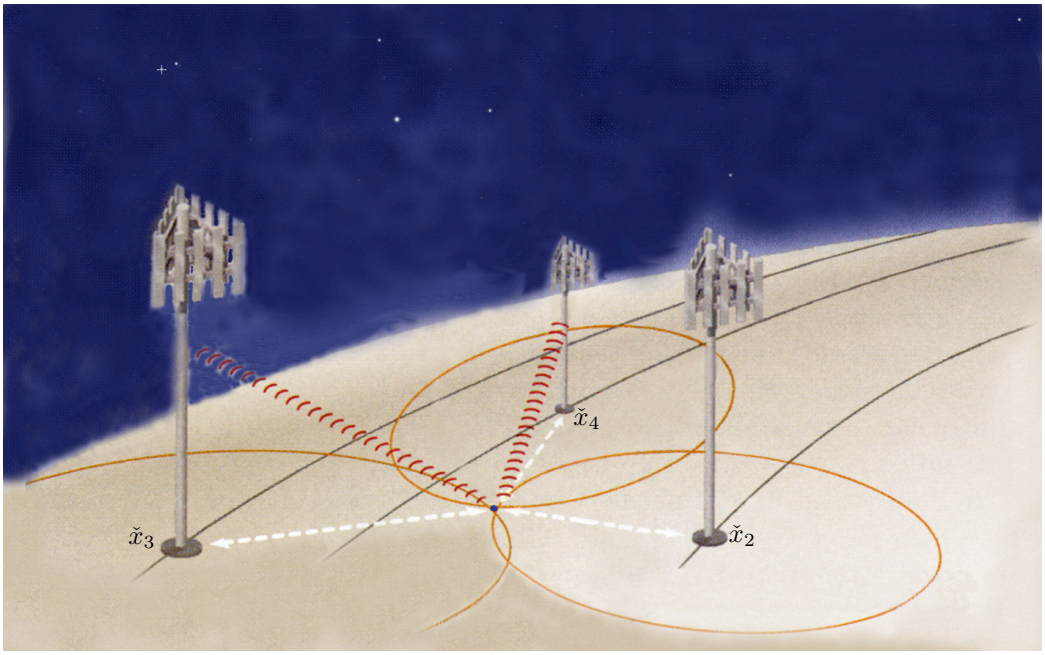


Figure 4: Application of trilateration (§5.4.2.2.8) is localization (determining position) of a radio signal source in 2 dimensions; more commonly known by radio engineers as the process "triangulation". In this scenario, anchors $\check{x}_2\,, \check{x}_3\,, \check{x}_4$ are illustrated as fixed antennae. [240] The radio signal source (a sensor $\bullet$ x_1) anywhere in affine hull of three antenna bases can be uniquely localized by measuring distance to each (dashed white arrowed line segments). Ambiguity of lone distance measurement to sensor is represented by circle about each antenna. Trilateration is expressible as a semidefinite program; hence, a convex optimization problem. [363]

- *robotics*; *e.g*, automated manufacturing, and autonomous navigation of vehicles maneuvering in formation (Figure **10**).

by chapter

We study the many manifestations and representations of pervasive convex Euclidean bodies. In particular, we make convex polyhedra, cones, and dual cones visceral through illustration in **chapter 2 Convex geometry** where geometric relation of polyhedral cones to nonorthogonal bases (biorthogonal expansion) is examined. It is shown that coordinates are unique in any conic system whose basis cardinality equals or exceeds spatial dimension; for high cardinality, a new definition of *conic coordinate* is provided in Theorem 2.13.13.0.1. The conic analogue to linear independence, called *conic independence,* is introduced as a tool for study, analysis, and manipulation of cones; a natural extension and next logical step in progression: linear, affine, conic. We explain conversion between halfspace- and vertex-description of a convex cone, we motivate the dual cone and provide formulae for finding it, and we show how first-order optimality conditions or alternative systems of linear inequality or *linear matrix inequality* can be explained by *dual generalized inequalities* with respect to convex cones. Arcane theorems of alternative generalized inequality are, in fact, simply derived from cone *membership relations*; generalizations of algebraic *Farkas' lemma* translated to geometry of convex cones.

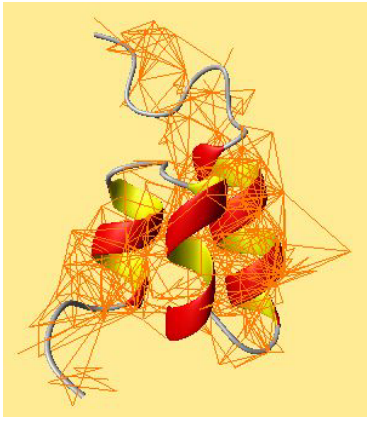


Figure 5: [219] [135] Distance data collected via nuclear magnetic resonance (NMR) helped render this three-dimensional depiction of a protein molecule. *At the beginning of the 1980s, Kurt Wüthrich* [Nobel laureate] *developed an idea about how NMR could be extended to cover biological molecules such as proteins. He invented a systematic method of pairing each NMR signal with the right hydrogen nucleus (proton) in the macromolecule. The method is called* sequential assignment *and is today a cornerstone of all NMR structural investigations. He also showed how it was subsequently possible to determine pairwise distances between a large number of hydrogen nuclei and use this information with a mathematical method based on distance-geometry to calculate a three-dimensional structure for the molecule.* [448] [214] −[319]

Any convex optimization problem can be visualized geometrically. Desire to visualize in high dimension [Sagan, *Cosmos − The Edge of Forever,* 22:55′] is deeply embedded in the mathematical psyche. [1] Chapter 2 provides tools to make visualization easier, and we teach how to visualize in high dimension. The concepts of face, extreme point, and extreme direction of a convex Euclidean body are explained here; crucial to understanding convex optimization. How to find the smallest face of any closed convex cone, containing convex set $\mathcal{C}$, is divulged; later shown to have practical application to presolving convex programs. The convex cone of positive semidefinite matrices, in particular, is studied in depth:

- We interpret, for example, inverse image of the positive semidefinite cone under affine transformation. (Example 2.9.1.0.2)
- Subsets of the positive semidefinite cone, discriminated by rank exceeding some lower bound, are convex. In other words, high-rank subsets of the positive semidefinite cone boundary united with its interior are convex. (Theorem 2.9.2.9.3) There is a closed form for projection on those convex subsets.
- The positive semidefinite cone is a circular cone in low dimension; *Geršgorin discs* specify inscription of a polyhedral cone into it. (Figure **51**)

Chapter 3 Geometry of convex functions observes Fenchel's analogy between convex sets and functions: We explain, for example, how the real affine function relates to convex functions as the hyperplane relates to convex sets. Partly a toolbox of practical useful convex functions and a cookbook for optimization problems, methods are drawn from the appendices about matrix calculus for determining convexity and discerning geometry.

Chapter 4. Semidefinite programming *has recently emerged to prominence because it admits a new problem type previously unsolvable by convex optimization techniques*

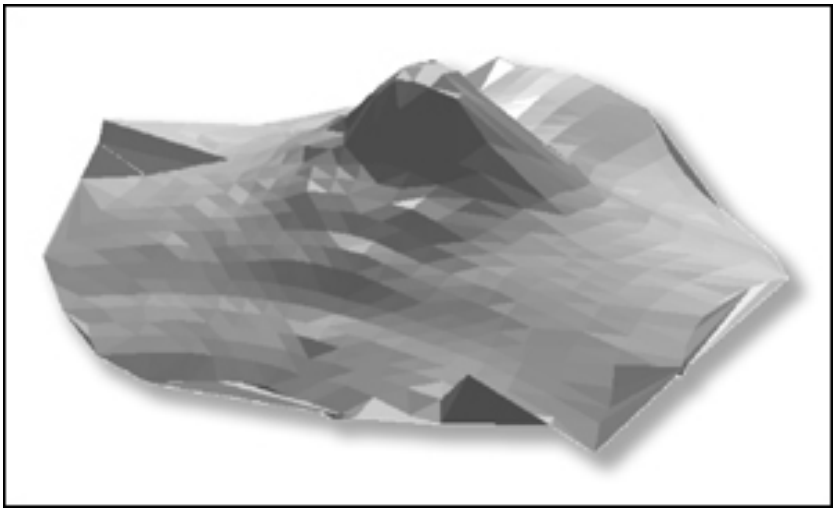


Figure 6: This coarsely discretized triangulated algorithmically flattened human face (made by Kimmel & the Bronsteins [258]) represents a stage in machine recognition of human identity; called *facial recognition.* Distance geometry is applied to determine discriminating-features.

and because it theoretically subsumes other convex types: linear programming, quadratic programming, second-order cone programming. −p.225 Semidefinite programming is reviewed with particular attention to optimality conditions for prototypical primal and dual problems, their interplay, and a perturbation method for rank reduction of optimal solutions (extant but not well known). *Positive definite Farkas' lemma* is derived, and we also show how to determine if a feasible set belongs exclusively to a positive semidefinite cone boundary. An arguably good three-dimensional polyhedral analogue to the positive semidefinite cone of 3×3 symmetric matrices is introduced: a new tool for visualizing coexistence of low- and high-rank optimal solutions in six isomorphic dimensions and a mnemonic aid for understanding semidefinite programs. We find a minimal cardinality Boolean solution to an instance of $Ax=b$:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x\|_0 \\ \text{subject to} & Ax=b \\ & x_i\in\{0,1\}\ , \quad i=1\ldots n \end{array} \tag{740}$$

The *sensor-network localization* problem is solved in any dimension in this chapter. We introduce a method of *convex iteration* for constraining rank in the form $\operatorname{rank} G\leq\rho$ and cardinality in the form $\operatorname{card} x\leq k$. Cardinality minimization is applied to a discrete image-gradient of the Shepp-Logan phantom, from Magnetic Resonance Imaging (MRI) in the field of medical imaging, for which we find a new lower bound of 1.9% cardinality. We show how to handle polynomial constraints, and how to transform a rank-constrained problem to a rank-1 problem.

The EDM is studied in **chapter 5 Euclidean Distance Matrix**; its properties and relationship to both positive semidefinite and Gram matrices. We relate the EDM to the four classical properties of Euclidean metric; thereby, observing existence of an infinity of properties of the Euclidean metric beyond triangle inequality. We proceed by deriving the fifth Euclidean metric property and then explain why furthering this endeavor is inefficient because the ensuing criteria (while describing polyhedra in angle or area, volume, content, and so on *ad infinitum*) grow linearly in complexity and number with problem size.

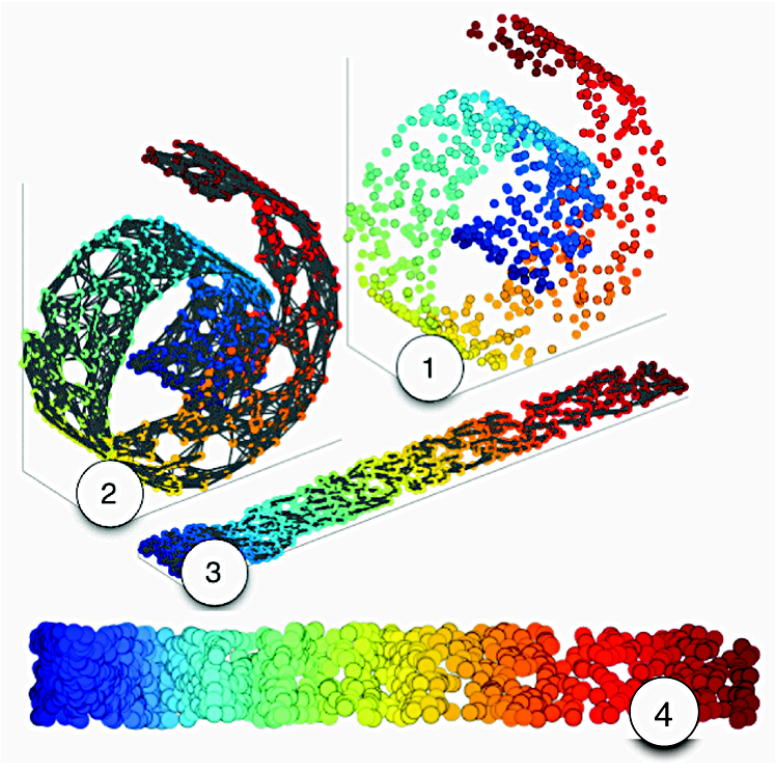


Figure 7: *Swiss roll,* Weinberger & Saul [443]. *The problem of manifold learning, illustrated for $N = 800$ data points sampled from a "Swiss roll" ①. A discretized manifold is revealed by connecting each data point and its $k = 6$ nearest neighbors ②. An unsupervised learning algorithm unfolds the Swiss roll while preserving the local geometry of nearby data points ③. Finally, the data points are projected onto the two-dimensional subspace that maximizes their variance, yielding a faithful embedding of the original manifold ④.*

Reconstruction methods are explained and applied to a map of the United States; *e.g*, Figure **8**. We also experimentally test a conjecture of Borg & Groenen by reconstructing a distorted but recognizable isotonic map of the USA using only ordinal (comparative) distance data: Figure **161**e-f. We demonstrate an elegant method for including dihedral (or *torsion*) angle constraints into a molecular conformation problem. We explain why *trilateration* (`a.k.a` *localization*) is a convex optimization problem. We show how to recover relative position given incomplete interpoint distance information, and how to pose EDM problems or transform geometrical problems to convex optimizations; *e.g*, *kissing number* of packed spheres about a central sphere (solved in $\mathbb{R}^{\mathbf{3}}$ by Isaac Newton).

The set of all Euclidean distance matrices forms a pointed closed convex cone called the *EDM cone*: $\mathbb{EDM}^N$. We offer a new proof of Schoenberg's seminal characterization of EDMs:

$$D \in \mathbb{EDM}^N \;\Leftrightarrow\; \begin{cases} -V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \succeq 0 \\ D \in \mathbb{S}_h^N \end{cases} \tag{1068}$$

Our proof relies on fundamental geometry; assuming, any EDM must correspond to a list of points contained in some polyhedron (possibly at its vertices) and *vice versa.* It is known, but not obvious, this *Schoenberg criterion* implies nonnegativity of the EDM entries; proved herein.

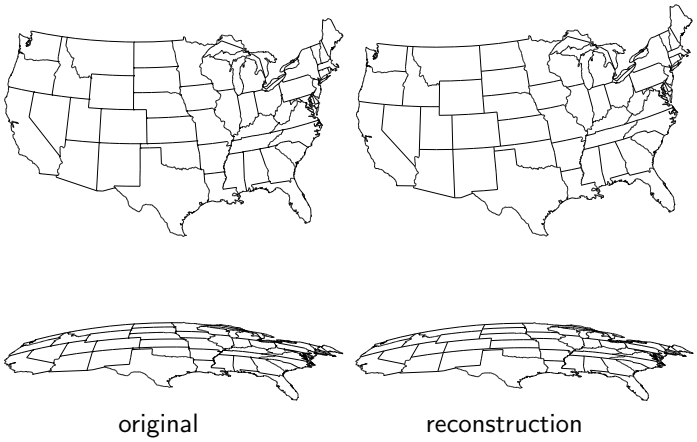


Figure 8: (*confer* Figure **161**) About five thousand points along borders constituting United States were used to create an exhaustive matrix of interpoint distance for each and every pair of points in an ordered set (a *list*); called *Euclidean distance matrix.* From that noiseless distance information, it is easy to reconstruct this nonconvex map exactly via Schoenberg criterion (1068). (§5.13.1.0.1) Map reconstruction is exact (to within a rigid transformation) given any number of interpoint distances; the greater the number of distances, the greater the detail (as it is for all conventional map preparation).

We characterize eigenvalue spectrum of an EDM, then devise a polyhedral spectral cone for determining membership of a given matrix (in Cayley-Menger form) to the convex cone of Euclidean distance matrices; *id est*, a matrix is an EDM if and only if its nonincreasingly ordered vector of eigenvalues belongs to a polyhedral spectral cone for $\mathbb{EDM}^N$

$$D \in \mathbb{EDM}^N \;\Leftrightarrow\; \begin{cases} \lambda\left(\begin{bmatrix} 0 & \mathbf{1}^{\mathrm{T}} \\ \mathbf{1} & -D \end{bmatrix}\right) \in \begin{bmatrix} \mathbb{R}^N_+ \\ \mathbb{R}_- \end{bmatrix} \cap \partial\mathcal{H} \\ D \in \mathbb{S}^N_h \end{cases} \tag{1286}$$

We will see: spectral cones are not unique.

In **chapter 6 Cone of distance matrices** we explain a geometric relationship between the cone of Euclidean distance matrices, two positive semidefinite cones, and the elliptope. We illustrate geometric requirements, in particular, for projection of a given matrix on a positive semidefinite cone that establish its membership to the EDM cone. The faces of the EDM cone are described, but still open is the question whether all its faces are exposed as they are for the positive semidefinite cone.

The *Schoenberg criterion,*

$$D \in \mathbb{EDM}^N \;\Leftrightarrow\; \begin{cases} -V_{\mathcal{N}}^{\mathrm{T}} D V_{\mathcal{N}} \in \mathbb{S}^{N-1}_+ \\ D \in \mathbb{S}^N_h \end{cases} \tag{1068}$$

for identifying a Euclidean distance matrix, is revealed to be a discretized *membership relation* (*dual generalized inequalities,* a new Farkas'-like lemma) between the EDM cone and its ordinary dual: $\mathbb{EDM}^{N\,*}$. A matrix criterion for membership to the dual EDM cone is derived that is simpler than the Schoenberg criterion:

$$D^* \in \mathbb{EDM}^{N\,*} \;\Leftrightarrow\; \delta(D^*\,\mathbf{1}) - D^* \succeq 0 \tag{1436}$$

There is a concise equality, relating the convex cone of Euclidean distance matrices to the positive semidefinite cone, apparently overlooked in the literature; an equality between

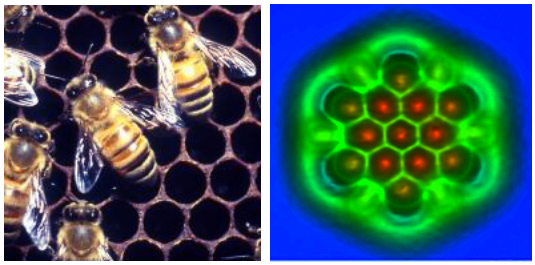
(a)  (b)

Figure 9: **(a)** These bees construct a honeycomb by solving a convex optimization problem (§5.4.2.2.6). The most dense packing of identical spheres about a central sphere in 2 dimensions is 6. Sphere centers describe a regular lattice. **(b)** *A hexabenzocoronene molecule (diameter: 1.4nm) imaged by noncontact atomic force microscopy using a microscope tip terminated with a single carbon monoxide molecule. The carbon-carbon bonds in the imaged molecule appear with different contrast and apparent lengths. Based on these disparities, the bond orders and lengths of the individual bonds can be distinguished.* (Image by Leo Gross.)

two large convex Euclidean bodies:

$$\mathbb{EDM}^N = \mathbb{S}_h^N \cap \left(\mathbb{S}_c^{N\perp} - \mathbb{S}_+^N\right) \qquad (1430)$$

Seemingly innocuous problems in terms of point position $x_i \in \mathbb{R}^n$ like

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

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

are difficult to solve. So, in **chapter 7 Proximity problems**, we instead explore methods of their solution by transformation to a few fundamental and prevalent Euclidean distance matrix proximity problems; the problem of finding that distance matrix closest, in some sense, to a given matrix $H = [h_{ij}]$:

$$\begin{array}{ll} \underset{D}{\text{minimize}} & \|-V(D-H)V\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \le \rho \\ & D \in \mathbb{EDM}^N \end{array} \qquad \begin{array}{ll} \underset{\sqrt[\circ]{D}}{\text{minimize}} & \|\sqrt[\circ]{D} - H\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \le \rho \\ & \sqrt[\circ]{D} \in \sqrt{\mathbb{EDM}^N} \end{array} \qquad (1472)$$

$$\begin{array}{ll} \underset{D}{\text{minimize}} & \|D-H\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \le \rho \\ & D \in \mathbb{EDM}^N \end{array} \qquad \begin{array}{ll} \underset{\sqrt[\circ]{D}}{\text{minimize}} & \|-V(\sqrt[\circ]{D} - H)V\|_{\mathrm{F}}^2 \\ \text{subject to} & \operatorname{rank} VDV \le \rho \\ & \sqrt[\circ]{D} \in \sqrt{\mathbb{EDM}^N} \end{array}$$

We apply a convex iteration method for constraining rank. Known heuristics for rank minimization are also explained. We offer new geometrical proof, in §7.1.4.0.1, of a famous discovery by Eckart & Young in 1936 [149]: Euclidean projection on that generally nonconvex subset of the positive semidefinite cone boundary comprising all semidefinite matrices having rank not exceeding a prescribed bound ρ. We explain how this problem is transformed to a convex optimization for any rank ρ.

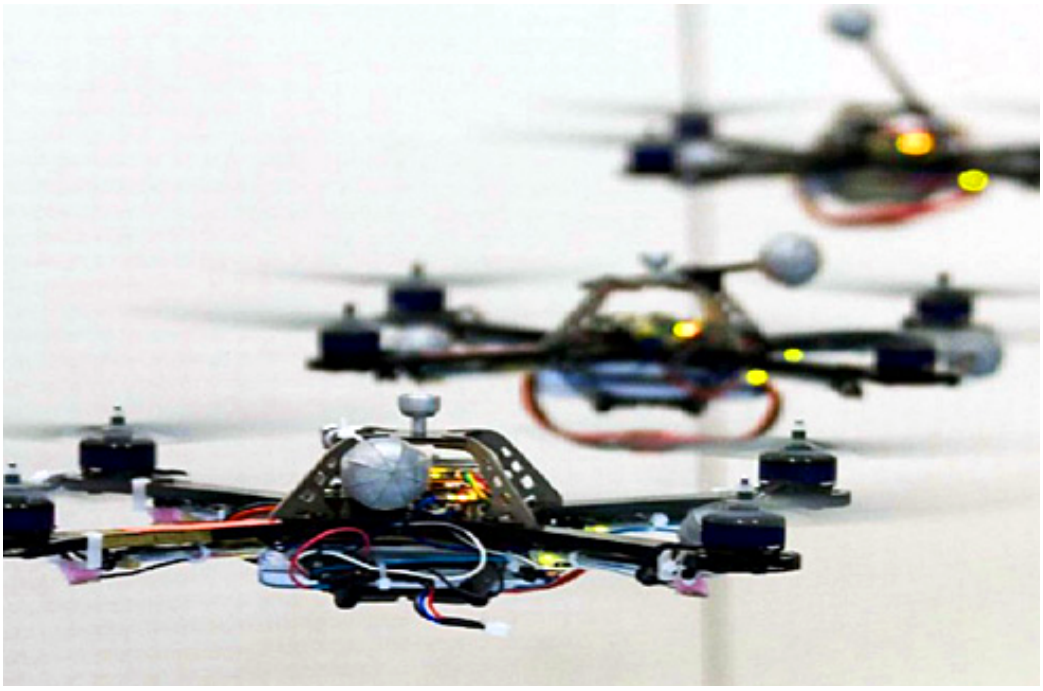


Figure 10: Nanocopter swarm. Robotic vehicles in concert can move larger objects or localize a plume of gas, liquid, or radio waves. [159]

appendices

We presume a reader already comfortable with elementary vector operations; [15, §3] formally known as *analytic geometry*. [450] Toolboxes are provided, in the form of appendices and code, so as to be more self-contained:

- linear algebra (**appendix A** is primarily concerned with proper statements of semidefiniteness for square matrices),
- simple matrices (dyad, doublet, elementary, Householder, Schoenberg, orthogonal, *etcetera*, in **appendix B**),
- collection of known analytical solutions to some important optimization problems (**appendix C**),
- matrix calculus remains somewhat unsystematized when compared to ordinary calculus (**appendix D** concerns matrix-valued functions, matrix differentiation and directional derivatives, Taylor series, and tables of first- and second-order gradients and matrix derivatives),
- an elaborate exposition offering insight into orthogonal and nonorthogonal projection on convex sets (the connection between projection and positive semidefiniteness, for example, or between projection and a linear objective function in **appendix E**),
- MATLAB code on *Wikimization* [431] to discriminate EDMs, to determine conic independence, to reduce or constrain rank of an optimal solution to a semidefinite program, compressed sensing (compressive sampling) for digital image and audio signal processing, and two distinct methods of reconstructing a map of the United States: one given only distance data, the other given only comparative distance.

Figure 11: Three-dimensional reconstruction of David from distance data.

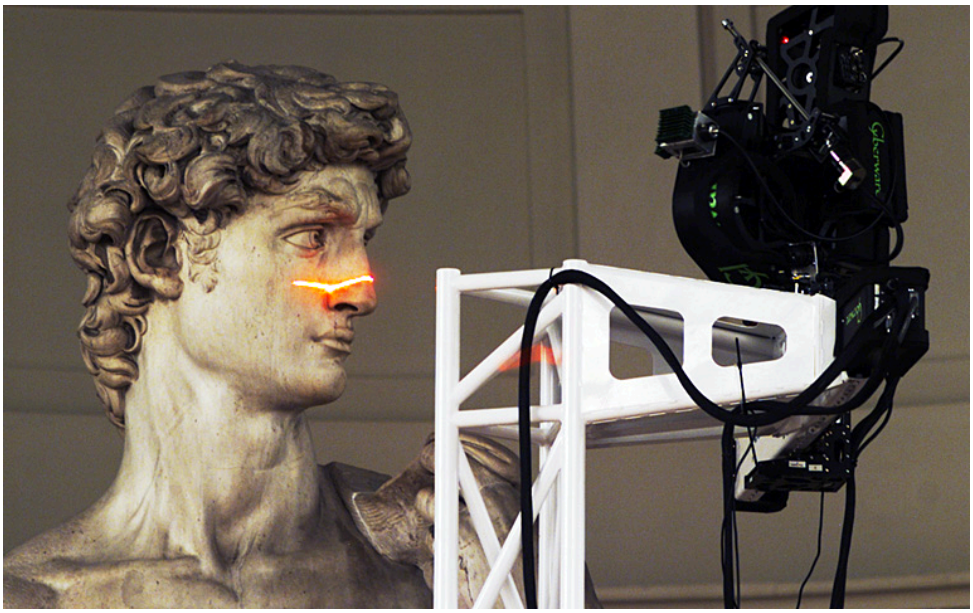


Figure 12: *Digital Michelangelo Project*, Stanford University. Measuring distance to David by laser rangefinder. (Spatial resolution: 0.29mm.) *Crystalix* commercialized a 3D image rendering laser by refining a stunning technique for interior engraving of cubic *photocrystal*.