Chapter 8

Geometric problems

8.1 Projection on a set

The *distance* of a point $x_0 \in \mathbf{R}^n$ to a closed set $C \subseteq \mathbf{R}^n$, in the norm $\|\cdot\|$, is defined as

$$\mathbf{dist}(x_0, C) = \inf\{\|x_0 - x\| \mid x \in C\}.$$

The infimum here is always achieved. We refer to any point $z \in C$ which is closest to x_0, *i.e.*, satisfies $\|z - x_0\| = \mathbf{dist}(x_0, C)$, as a *projection* of x_0 on C. In general there can be more than one projection of x_0 on C, *i.e.*, several points in C closest to x_0.

In some special cases we can establish that the projection of a point on a set is unique. For example, if C is closed and convex, and the norm is strictly convex (*e.g.*, the Euclidean norm), then for any x_0 there is always exactly one $z \in C$ which is closest to x_0. As an interesting converse, we have the following result: If for every x_0 there is a unique Euclidean projection of x_0 on C, then C is closed and convex (see exercise 8.2).

We use the notation $P_C : \mathbf{R}^n \to \mathbf{R}^n$ to denote any function for which $P_C(x_0)$ is a projection of x_0 on C, *i.e.*, for all x_0,

$$P_C(x_0) \in C, \qquad \|x_0 - P_C(x_0)\| = \mathbf{dist}(x_0, C).$$

In other words, we have

$$P_C(x_0) = \operatorname{argmin}\{\|x - x_0\| \mid x \in C\}.$$

We refer to P_C as *projection on* C.

Example 8.1 *Projection on the unit square in* $\mathbf{R}^2$. Consider the (boundary of the) unit square in $\mathbf{R}^2$, *i.e.*, $C = \{x \in \mathbf{R}^2 \mid \|x\|_\infty = 1\}$. We take $x_0 = 0$.

In the ℓ_1-norm, the four points $(1, 0)$, $(0, -1)$, $(-1, 0)$, and $(0, 1)$ are closest to $x_0 = 0$, with distance 1, so we have $\mathbf{dist}(x_0, C) = 1$ in the ℓ_1-norm. The same statement holds for the ℓ_2-norm.

In the ℓ_∞-norm, all points in C lie at a distance 1 from x_0, and $\mathbf{dist}(x_0, C) = 1$.

Example 8.2 *Projection onto rank-k matrices.* Consider the set of $m \times n$ matrices with rank less than or equal to k,

$$C = \{X \in \mathbf{R}^{m\times n} \mid \mathbf{rank}\, X \leq k\},$$

with $k \leq \min\{m, n\}$, and let $X_0 \in \mathbf{R}^{m\times n}$. We can find a projection of X_0 on C, in the (spectral or maximum singular value) norm $\|\cdot\|_2$, via the singular value decomposition. Let

$$X_0 = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

be the singular value decomposition of X_0, where $r = \mathbf{rank}\, X_0$. Then the matrix $Y = \sum_{i=1}^{\min\{k,r\}} \sigma_i u_i v_i^T$ is a projection of X_0 on C.

8.1.1 Projecting a point on a convex set

If C is convex, then we can compute the projection $P_C(x_0)$ and the distance $\mathbf{dist}(x_0, C)$ by solving a convex optimization problem. We represent the set C by a set of linear equalities and convex inequalities

$$Ax = b, \qquad f_i(x) \leq 0, \quad i = 1, \ldots, m, \tag{8.1}$$

and find the projection of x_0 on C by solving the problem

$$\begin{array}{ll} \text{minimize} & \|x - x_0\| \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b, \end{array} \tag{8.2}$$

with variable x. This problem is feasible if and only if C is nonempty; when it is feasible, its optimal value is $\mathbf{dist}(x_0, C)$, and any optimal point is a projection of x_0 on C.

Euclidean projection on a polyhedron

The projection of x_0 on a polyhedron described by linear inequalities $Ax \preceq b$ can be computed by solving the QP

$$\begin{array}{ll} \text{minimize} & \|x - x_0\|_2^2 \\ \text{subject to} & Ax \preceq b. \end{array}$$

Some special cases have simple analytical solutions.

- The Euclidean projection of x_0 on a hyperplane $C = \{x \mid a^T x = b\}$ is given by
$$P_C(x_0) = x_0 + (b - a^T x_0)a/\|a\|_2^2.$$
- The Euclidean projection of x_0 on a halfspace $C = \{x \mid a^T x \leq b\}$ is given by
$$P_C(x_0) = \begin{cases} x_0 + (b - a^T x_0)a/\|a\|_2^2 & a^T x_0 > b \\ x_0 & a^T x_0 \leq b. \end{cases}$$

- The Euclidean projection of x_0 on a rectangle $C = \{x \mid l \preceq x \preceq u\}$ (where $l \prec u$) is given by

$$P_C(x_0)_k = \begin{cases} l_k & x_{0k} \leq l_k \\ x_{0k} & l_k \leq x_{0k} \leq u_k \\ u_k & x_{0k} \geq u_k. \end{cases}$$

Euclidean projection on a proper cone

Let $x = P_K(x_0)$ denote the Euclidean projection of a point x_0 on a proper cone K. The KKT conditions of

$$\begin{array}{ll} \text{minimize} & \|x - x_0\|_2^2 \\ \text{subject to} & x \succeq_K 0 \end{array}$$

are given by

$$x \succeq_K 0, \qquad x - x_0 = z, \qquad z \succeq_{K^*} 0, \qquad z^T x = 0.$$

Introducing the notation $x_+ = x$ and $x_- = z$, we can express these conditions as

$$x_0 = x_+ - x_-, \qquad x_+ \succeq_K 0, \qquad x_- \succeq_{K^*} 0, \qquad x_+^T x_- = 0.$$

In other words, by projecting x_0 on the cone K, we decompose it into the difference of two orthogonal elements: one nonnegative with respect to K (and which is the projection of x_0 on K), and the other nonnegative with respect to K^*.

Some specific examples:

- For $K = \mathbf{R}^n_+$, we have $P_K(x_0)_k = \max\{x_{0k}, 0\}$. The Euclidean projection of a vector onto the nonnegative orthant is found by replacing each negative component with 0.

- For $K = \mathbf{S}^n_+$, and the Euclidean (or Frobenius) norm $\|\cdot\|_F$, we have $P_K(X_0) = \sum_{i=1}^n \max\{0, \lambda_i\} v_i v_i^T$, where $X_0 = \sum_{i=1}^n \lambda_i v_i v_i^T$ is the eigenvalue decomposition of X_0. To project a symmetric matrix onto the positive semidefinite cone, we form its eigenvalue expansion and drop terms associated with negative eigenvalues. This matrix is also the projection onto the positive semidefinite cone in the ℓ_2-, or spectral norm.

8.1.2 Separating a point and a convex set

Suppose C is a closed convex set described by the equalities and inequalities (8.1). If $x_0 \in C$, then $\mathbf{dist}(x_0, C) = 0$, and the optimal point for the problem (8.2) is x_0. If $x_0 \notin C$ then $\mathbf{dist}(x_0, C) > 0$, and the optimal value of the problem (8.2) is positive. In this case we will see that any dual optimal point provides a separating hyperplane between the point x_0 and the set C.

The link between projecting a point on a convex set and finding a hyperplane that separates them (when the point is not in the set) should not be surprising. Indeed, our proof of the separating hyperplane theorem, given in §2.5.1, relies on

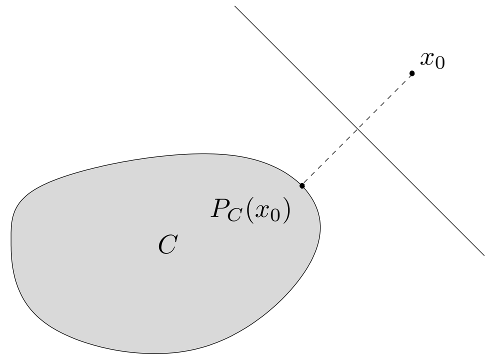


Figure 8.1 A point x_0 and its Euclidean projection $P_C(x_0)$ on a convex set C. The hyperplane midway between the two, with normal vector $P_C(x_0) - x_0$, strictly separates the point and the set. This property does not hold for general norms; see exercise 8.4.

finding the Euclidean distance between the sets. If $P_C(x_0)$ denotes the Euclidean projection of x_0 on C, where $x_0 \notin C$, then the hyperplane

$$(P_C(x_0) - x_0)^T (x - (1/2)(x_0 + P_C(x_0))) = 0$$

(strictly) separates x_0 from C, as illustrated in figure 8.1. In other norms, however, the clearest link between the projection problem and the separating hyperplane problem is via Lagrange duality.

We first express (8.2) as

$$\begin{array}{ll} \text{minimize} & \|y\| \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \\ & x_0 - x = y \end{array}$$

with variables x and y. The Lagrangian of this problem is

$$L(x, y, \lambda, \mu, \nu) = \|y\| + \sum_{i=1}^m \lambda_i f_i(x) + \nu^T (Ax - b) + \mu^T (x_0 - x - y)$$

and the dual function is

$$g(\lambda, \mu, \nu) = \begin{cases} \inf_x \left(\sum_{i=1}^m \lambda_i f_i(x) + \nu^T (Ax - b) + \mu^T (x_0 - x) \right) & \|\mu\|_* \leq 1 \\ -\infty & \text{otherwise,} \end{cases}$$

so we obtain the dual problem

$$\begin{array}{ll} \text{maximize} & \mu^T x_0 + \inf_x \left(\sum_{i=1}^m \lambda_i f_i(x) + \nu^T (Ax - b) - \mu^T x \right) \\ \text{subject to} & \lambda \succeq 0 \\ & \|\mu\|_* \leq 1, \end{array}$$

with variables λ, μ, ν. We can interpret the dual problem as follows. Suppose λ, μ, ν are dual feasible with a positive dual objective value, *i.e.*, $\lambda \succeq 0$, $\|\mu\|_* \leq 1$, and

$$\mu^T x_0 - \mu^T x + \sum_{i=1}^{m} \lambda_i f_i(x) + \nu^T (Ax - b) > 0$$

for all x. This implies that $\mu^T x_0 > \mu^T x$ for $x \in C$, and therefore μ defines a strictly separating hyperplane. In particular, suppose (8.2) is strictly feasible, so strong duality holds. If $x_0 \notin C$, the optimal value is positive, and any dual optimal solution defines a strictly separating hyperplane.

Note that this construction of a separating hyperplane, via duality, works for any norm. In contrast, the simple construction described above only works for the Euclidean norm.

Separating a point from a polyhedron

The dual problem of

$$\begin{array}{ll} \text{minimize} & \|y\| \\ \text{subject to} & Ax \preceq b \\ & x_0 - x = y \end{array}$$

is

$$\begin{array}{ll} \text{maximize} & \mu^T x_0 - b^T \lambda \\ \text{subject to} & A^T \lambda = \mu \\ & \|\mu\|_* \leq 1 \\ & \lambda \succeq 0 \end{array}$$

which can be further simplified as

$$\begin{array}{ll} \text{maximize} & (Ax_0 - b)^T \lambda \\ \text{subject to} & \|A^T \lambda\|_* \leq 1 \\ & \lambda \succeq 0. \end{array}$$

It is easily verified that if the dual objective is positive, then $A^T \lambda$ is the normal vector to a separating hyperplane: If $Ax \preceq b$, then

$$(A^T \lambda)^T x = \lambda^T (Ax) \leq \lambda^T b < \lambda^T A x_0,$$

so $\mu = A^T \lambda$ defines a separating hyperplane.

8.1.3 Projection and separation via indicator and support functions

The ideas described above in §8.1.1 and §8.1.2 can be expressed in a compact form in terms of the indicator function I_C and the support function S_C of the set C, defined as

$$S_C(x) = \sup_{y \in C} x^T y, \qquad I_C(x) = \begin{cases} 0 & x \in C \\ +\infty & x \notin C. \end{cases}$$

The problem of projecting x_0 on a closed convex set C can be expressed compactly as

$$\begin{array}{ll} \text{minimize} & \|x - x_0\| \\ \text{subject to} & I_C(x) \leq 0, \end{array}$$

or, equivalently, as

$$\begin{array}{ll} \text{minimize} & \|y\| \\ \text{subject to} & I_C(x) \leq 0 \\ & x_0 - x = y \end{array}$$

where the variables are x and y. The dual function of this problem is

$$\begin{array}{rcl} g(z,\lambda) & = & \inf_{x,y} \left(\|y\| + \lambda I_C(x) + z^T(x_0 - x - y) \right) \\ & = & \left\{ \begin{array}{ll} z^T x_0 + \inf_x \left(-z^T x + I_C(x) \right) & \|z\|_* \leq 1, \quad \lambda \geq 0 \\ -\infty & \text{otherwise} \end{array} \right. \\ & = & \left\{ \begin{array}{ll} z^T x_0 - S_C(z) & \|z\|_* \leq 1, \quad \lambda \geq 0 \\ -\infty & \text{otherwise} \end{array} \right. \end{array}$$

so we obtain the dual problem

$$\begin{array}{ll} \text{maximize} & z^T x_0 - S_C(z) \\ \text{subject to} & \|z\|_* \leq 1. \end{array}$$

If z is dual optimal with a positive objective value, then $z^T x_0 > z^T x$ for all $x \in C$, *i.e.*, z defines a separating hyperplane.

8.2 Distance between sets

The distance between two sets C and D, in a norm $\|\cdot\|$, is defined as

$$\mathbf{dist}(C,D) = \inf\{\|x - y\| \mid x \in C,\ y \in D\}.$$

The two sets C and D do not intersect if $\mathbf{dist}(C,D) > 0$. They intersect if $\mathbf{dist}(C,D) = 0$ and the infimum in the definition is attained (which is the case, for example, if the sets are closed and one of the sets is bounded).

The distance between sets can be expressed in terms of the distance between a point and a set,

$$\mathbf{dist}(C,D) = \mathbf{dist}(0, D - C),$$

so the results of the previous section can be applied. In this section, however, we derive results specifically for problems involving distance between sets. This allows us to exploit the structure of the set $C - D$, and makes the interpretation easier.

8.2.1 Computing the distance between convex sets

Suppose C and D are described by two sets of convex inequalities

$$C = \{x \mid f_i(x) \leq 0,\ i = 1,\dots,m\}, \qquad D = \{x \mid g_i(x) \leq 0,\ i = 1,\dots,p\}.$$

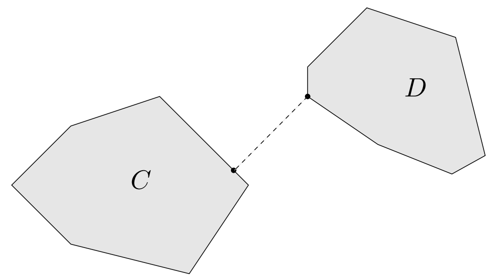


Figure 8.2 Euclidean distance between polyhedra C and D. The dashed line connects the two points in C and D, respectively, that are closest to each other in Euclidean norm. These points can be found by solving a QP.

(We can include linear equalities, but exclude them here for simplicity.) We can find $\mathbf{dist}(C, D)$ by solving the convex optimization problem

$$\begin{array}{ll} \text{minimize} & \|x - y\| \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & g_i(y) \leq 0, \quad i = 1, \ldots, p. \end{array} \tag{8.3}$$

Euclidean distance between polyhedra

Let C and D be two polyhedra described by the sets of linear inequalities $A_1 x \preceq b_1$ and $A_2 x \preceq b_2$, respectively. The distance between C and D is the distance between the closest pair of points, one in C and the other in D, as illustrated in figure 8.2. The distance between them is the optimal value of the problem

$$\begin{array}{ll} \text{minimize} & \|x - y\|_2 \\ \text{subject to} & A_1 x \preceq b_1 \\ & A_2 y \preceq b_2. \end{array} \tag{8.4}$$

We can square the objective to obtain an equivalent QP.

8.2.2 Separating convex sets

The dual of the problem (8.3) of finding the distance between two convex sets has an interesting geometric interpretation in terms of separating hyperplanes between the sets. We first express the problem in the following equivalent form:

$$\begin{array}{ll} \text{minimize} & \|w\| \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & g_i(y) \leq 0, \quad i = 1, \ldots, p \\ & x - y = w. \end{array} \tag{8.5}$$

The dual function is

$$g(\lambda, z, \mu) = \inf_{x,y,w} \left(\|w\| + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i g_i(y) + z^T(x - y - w) \right)$$

$$= \begin{cases} \inf_x \left(\sum_{i=1}^m \lambda_i f_i(x) + z^T x\right) + \inf_y \left(\sum_{i=1}^p \mu_i g_i(y) - z^T y\right) & \|z\|_* \leq 1 \\ -\infty & \text{otherwise,} \end{cases}$$

which results in the dual problem

$$\begin{array}{ll} \text{maximize} & \inf_x \left(\sum_{i=1}^m \lambda_i f_i(x) + z^T x\right) + \inf_y \left(\sum_{i=1}^p \mu_i g_i(y) - z^T y\right) \\ \text{subject to} & \|z\|_* \leq 1 \\ & \lambda \succeq 0, \quad \mu \succeq 0. \end{array} \tag{8.6}$$

We can interpret this geometrically as follows. If λ, μ are dual feasible with a positive objective value, then

$$\sum_{i=1}^m \lambda_i f_i(x) + z^T x + \sum_{i=1}^p \mu_i g_i(y) - z^T y > 0$$

for all x and y. In particular, for $x \in C$ and $y \in D$, we have $z^T x - z^T y > 0$, so we see that z defines a hyperplane that strictly separates C and D.

Therefore, if strong duality holds between the two problems (8.5) and (8.6) (which is the case when (8.5) is strictly feasible), we can make the following conclusion. If the distance between the two sets is positive, then they can be strictly separated by a hyperplane.

Separating polyhedra

Applying these duality results to sets defined by linear inequalities $A_1 x \preceq b_1$ and $A_2 x \preceq b_2$, we find the dual problem

$$\begin{array}{ll} \text{maximize} & -b_1^T \lambda - b_2^T \mu \\ \text{subject to} & A_1^T \lambda + z = 0 \\ & A_2^T \mu - z = 0 \\ & \|z\|_* \leq 1 \\ & \lambda \succeq 0, \quad \mu \succeq 0. \end{array}$$

If λ, μ, and z are dual feasible, then for all $x \in C$, $y \in D$,

$$z^T x = -\lambda^T A_1 x \geq -\lambda^T b_1, \qquad z^T y = \mu^T A_2 x \leq \mu^T b_2,$$

and, if the dual objective value is positive,

$$z^T x - z^T y \geq -\lambda^T b_1 - \mu^T b_2 > 0,$$

i.e., z defines a separating hyperplane.

8.2.3 Distance and separation via indicator and support functions

The ideas described above in §8.2.1 and §8.2.2 can be expressed in a compact form using indicator and support functions. The problem of finding the distance between two convex sets can be posed as the convex problem

$$\begin{array}{ll} \text{minimize} & \|x - y\| \\ \text{subject to} & I_C(x) \leq 0 \\ & I_D(y) \leq 0, \end{array}$$

which is equivalent to

$$\begin{array}{ll} \text{minimize} & \|w\| \\ \text{subject to} & I_C(x) \leq 0 \\ & I_D(y) \leq 0 \\ & x - y = w. \end{array}$$

The dual of this problem is

$$\begin{array}{ll} \text{maximize} & -S_C(-z) - S_D(z) \\ \text{subject to} & \|z\|_* \leq 1. \end{array}$$

If z is dual feasible with a positive objective value, then $S_D(z) < -S_C(-z)$, *i.e.*,

$$\sup_{x \in D} z^T x < \inf_{x \in C} z^T x.$$

In other words, z defines a hyperplane that strictly separates C and D.

8.3 Euclidean distance and angle problems

Suppose $a_1, \ldots, a_n$ is a set of vectors in $\mathbf{R}^n$, which we assume (for now) have known Euclidean lengths

$$l_1 = \|a_1\|_2, \quad \ldots, \quad l_n = \|a_n\|_2.$$

We will refer to the set of vectors as a *configuration*, or, when they are independent, a *basis*. In this section we consider optimization problems involving various geometric properties of the configuration, such as the Euclidean distances between pairs of the vectors, the angles between pairs of the vectors, and various geometric measures of the conditioning of the basis.

8.3.1 Gram matrix and realizability

The lengths, distances, and angles can be expressed in terms of the *Gram matrix* associated with the vectors $a_1, \ldots, a_n$, given by

$$G = A^T A, \qquad A = \left[\begin{array}{ccc} a_1 & \cdots & a_n \end{array} \right],$$

so that $G_{ij} = a_i^T a_j$. The diagonal entries of G are given by

$$G_{ii} = l_i^2, \quad i = 1, \ldots, n,$$

which (for now) we assume are known and fixed. The distance d_{ij} between a_i and a_j is

$$\begin{array}{rcl} d_{ij} & = & \|a_i - a_j\|_2 \\ & = & (l_i^2 + l_j^2 - 2a_i^T a_j)^{1/2} \\ & = & (l_i^2 + l_j^2 - 2G_{ij})^{1/2}. \end{array}$$

Conversely, we can express G_{ij} in terms of d_{ij} as

$$G_{ij} = \frac{l_i^2 + l_j^2 - d_{ij}^2}{2},$$

which we note, for future reference, is an affine function of d_{ij}^2.

The *correlation coefficient* ρ_{ij} between (nonzero) a_i and a_j is given by

$$\rho_{ij} = \frac{a_i^T a_j}{\|a_i\|_2 \|a_j\|_2} = \frac{G_{ij}}{l_i l_j},$$

so that $G_{ij} = l_i l_j \rho_{ij}$ is a linear function of ρ_{ij}. The angle θ_{ij} between (nonzero) a_i and a_j is given by

$$\theta_{ij} = \cos^{-1} \rho_{ij} = \cos^{-1}(G_{ij}/(l_i l_j)),$$

where we take $\cos^{-1} \rho \in [0, \pi]$. Thus, we have $G_{ij} = l_i l_j \cos \theta_{ij}$.

The lengths, distances, and angles are invariant under orthogonal transformations: If $Q \in \mathbf{R}^{n \times n}$ is orthogonal, then the set of vectors $Qa_i, \ldots, Qa_n$ has the same Gram matrix, and therefore the same lengths, distances, and angles.

Realizability

The Gram matrix $G = A^T A$ is, of course, symmetric and positive semidefinite. The converse is a basic result of linear algebra: A matrix $G \in \mathbf{S}^n$ is the Gram matrix of a set of vectors $a_1, \ldots, a_n$ if and only if $G \succeq 0$. When $G \succeq 0$, we can construct a configuration with Gram matrix G by finding a matrix A with $A^T A = G$. One solution of this equation is the symmetric squareroot $A = G^{1/2}$. When $G \succ 0$, we can find a solution via the Cholesky factorization of G: If $LL^T = G$, then we can take $A = L^T$. Moreover, we can construct *all* configurations with the given Gram matrix G, given any one solution A, by orthogonal transformation: If $\tilde{A}^T \tilde{A} = G$ is *any* solution, then $\tilde{A} = QA$ for some orthogonal matrix Q.

Thus, a set of lengths, distances, and angles (or correlation coefficients) is *realizable*, *i.e.*, those of some configuration, if and only if the associated Gram matrix G is positive semidefinite, and has diagonal elements $l_1^2, \ldots, l_n^2$.

We can use this fact to express several geometric problems as convex optimization problems, with $G \in \mathbf{S}^n$ as the optimization variable. Realizability imposes the constraint $G \succeq 0$ and $G_{ii} = l_i^2$, $i = 1, \ldots, n$; we list below several other convex constraints and objectives.

Angle and distance constraints

We can fix an angle to have a certain value, $\theta_{ij} = \alpha$, via the linear equality constraint $G_{ij} = l_i l_j \cos \alpha$. More generally, we can impose a lower and upper bound on an angle, $\alpha \leq \theta_{ij} \leq \beta$, by the constraint

$$l_i l_j \cos \alpha \geq G_{ij} \geq l_i l_j \cos \beta,$$

which is a pair of linear inequalities on G. (Here we use the fact that $\cos^{-1}$ is monotone decreasing.) We can maximize or minimize a particular angle θ_{ij}, by minimizing or maximizing G_{ij} (again using monotonicity of $\cos^{-1}$).

In a similar way we can impose constraints on the distances. To require that d_{ij} lies in an interval, we use

$$\begin{aligned} d_{\min} \leq d_{ij} \leq d_{\max} &\iff d_{\min}^2 \leq d_{ij}^2 \leq d_{\max}^2 \\ &\iff d_{\min}^2 \leq l_i^2 + l_j^2 - 2G_{ij} \leq d_{\max}^2, \end{aligned}$$

which is a pair of linear inequalities on G. We can minimize or maximize a distance, by minimizing or maximizing its square, which is an affine function of G.

As a simple example, suppose we are given ranges (*i.e.*, an interval of possible values) for some of the angles and some of the distances. We can then find the minimum and maximum possible value of some other angle, or some other distance, over all configurations, by solving two SDPs. We can reconstruct the two extreme configurations by factoring the resulting optimal Gram matrices.

Singular value and condition number constraints

The singular values of A, $\sigma_1 \geq \cdots \geq \sigma_n$, are the squareroots of the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of G. Therefore σ_1^2 is a convex function of G, and σ_n^2 is a concave function of G. Thus we can impose an upper bound on the maximum singular value of A, or minimize it; we can impose a lower bound on the minimum singular value, or maximize it. The condition number of A, σ_1/σ_n, is a quasiconvex function of G, so we can impose a maximum allowable value, or minimize it over all configurations that satisfy the other geometric constraints, by quasiconvex optimization.

Roughly speaking, the constraints we can impose as convex constraints on G are those that require $a_1, \ldots, a_n$ to be a well conditioned basis.

Dual basis

When $G \succ 0$, $a_1, \ldots, a_n$ form a basis for $\mathbf{R}^n$. The associated *dual basis* is $b_1, \ldots, b_n$, where

$$b_i^T a_j = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

The dual basis vectors $b_1, \ldots, b_n$ are simply the rows of the matrix A^{-1}. As a result, the Gram matrix associated with the dual basis is G^{-1}.

We can express several geometric conditions on the dual basis as convex constraints on G. The (squared) lengths of the dual basis vectors,

$$\|b_i\|_2^2 = e_i^T G^{-1} e_i,$$

are convex functions of G, and so can be minimized. The trace of G^{-1}, another convex function of G, gives the sum of the squares of the lengths of the dual basis vectors (and is another measure of a well conditioned basis).

Ellipsoid and simplex volume

The volume of the ellipsoid $\{Au \mid \|u\|_2 \leq 1\}$, which gives another measure of how well conditioned the basis is, is given by

$$\gamma(\det(A^T A))^{1/2} = \gamma(\det G)^{1/2},$$

where γ is the volume of the unit ball in $\mathbf{R}^n$. The log volume is therefore $\log \gamma + (1/2) \log \det G$, which is a concave function of G. We can therefore maximize the volume of the image ellipsoid, over a convex set of configurations, by maximizing $\log \det G$.

The same holds for any set in $\mathbf{R}^n$. The volume of the image under A is its volume, multiplied by the factor $(\det G)^{1/2}$. For example, consider the image under A of the unit simplex $\mathbf{conv}\{0, e_1, \ldots, e_n\}$, *i.e.*, the simplex $\mathbf{conv}\{0, a_1, \ldots, a_n\}$. The volume of this simplex is given by $\overline{\gamma}(\det G)^{1/2}$, where $\overline{\gamma}$ is the volume of the unit simplex in $\mathbf{R}^n$. We can maximize the volume of this simplex by maximizing $\log \det G$.

8.3.2 Problems involving angles only

Suppose we only care about the angles (or correlation coefficients) between the vectors, and do not specify the lengths or distances between them. In this case it is intuitively clear that we can simply assume the vectors a_i have length $l_i = 1$. This is easily verified: The Gram matrix has the form $G = \mathbf{diag}(l) C \,\mathbf{diag}(l)$, where l is the vector of lengths, and C is the correlation matrix, *i.e.*, $C_{ij} = \cos \theta_{ij}$. It follows that if $G \succeq 0$ for any set of positive lengths, then $G \succeq 0$ for *all* sets of positive lengths, and in particular, this occurs if and only if $C \succeq 0$ (which is the same as assuming that all lengths are one). Thus, a set of angles $\theta_{ij} \in [0, \pi]$, $i, j = 1, \ldots, n$ is realizable if and only if $C \succeq 0$, which is a linear matrix inequality in the correlation coefficients.

As an example, suppose we are given lower and upper bounds on some of the angles (which is equivalent to imposing lower and upper bounds on the correlation coefficients). We can then find the minimum and maximum possible value of some other angle, over all configurations, by solving two SDPs.

Example 8.3 *Bounding correlation coefficients.* We consider an example in $\mathbf{R}^4$, where we are given

$$\begin{array}{ll} 0.6 \leq \rho_{12} \leq 0.9, & 0.8 \leq \rho_{13} \leq 0.9, \\ 0.5 \leq \rho_{24} \leq 0.7, & -0.8 \leq \rho_{34} \leq -0.4. \end{array} \tag{8.7}$$

To find the minimum and maximum possible values of ρ_{14}, we solve the two SDPs

$$\begin{array}{ll} \text{minimize/maximize} & \rho_{14} \\ \text{subject to} & (8.7) \\ & \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{bmatrix} \succeq 0, \end{array}$$

with variables $\rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}, \rho_{24}, \rho_{34}$. The minimum and maximum values (to two significant digits) are -0.39 and 0.23, with corresponding correlation matrices

$$\begin{bmatrix} 1.00 & 0.60 & 0.87 & -0.39 \\ 0.60 & 1.00 & 0.33 & 0.50 \\ 0.87 & 0.33 & 1.00 & -0.55 \\ -0.39 & 0.50 & -0.55 & 1.00 \end{bmatrix}, \qquad \begin{bmatrix} 1.00 & 0.71 & 0.80 & 0.23 \\ 0.71 & 1.00 & 0.31 & 0.59 \\ 0.80 & 0.31 & 1.00 & -0.40 \\ 0.23 & 0.59 & -0.40 & 1.00 \end{bmatrix}.$$

8.3.3 Euclidean distance problems

In a *Euclidean distance problem*, we are concerned *only* with the distances between the vectors, d_{ij}, and do not care about the lengths of the vectors, or about the angles between them. These distances, of course, are invariant not only under orthogonal transformations, but also translation: The configuration $\tilde{a}_1 = a_1+b, \ldots, \tilde{a}_n = a_n+b$ has the same distances as the original configuration, for any $b \in \mathbf{R}^n$. In particular, for the choice

$$b = -(1/n)\sum_{i=1}^{n} a_i = -(1/n)A\mathbf{1},$$

we see that $\tilde{a}_i$ have the same distances as the original configuration, and also satisfy $\sum_{i=1}^{n} \tilde{a}_i = 0$. It follows that in a Euclidean distance problem, we can assume, without any loss of generality, that the average of the vectors $a_1, \ldots, a_n$ is zero, *i.e.*, $A\mathbf{1} = 0$.

We can solve Euclidean distance problems by considering the lengths (which cannot occur in the objective or constraints of a Euclidean distance problem) as free variables in the optimization problem. Here we rely on the fact that there is a configuration with distances $d_{ij} \geq 0$ if and only if there are lengths $l_1, \ldots, l_n$ for which $G \succeq 0$, where $G_{ij} = (l_i^2 + l_j^2 - d_{ij}^2)/2$.

We define $z \in \mathbf{R}^n$ as $z_i = l_i^2$, and $D \in \mathbf{S}^n$ by $D_{ij} = d_{ij}^2$ (with, of course, $D_{ii} = 0$). The condition that $G \succeq 0$ for some choice of lengths can be expressed as

$$G = (z\mathbf{1}^T + \mathbf{1}z^T - D)/2 \succeq 0 \text{ for some } z \succeq 0, \qquad (8.8)$$

which is an LMI in D and z. A matrix $D \in \mathbf{S}^n$, with nonnegative elements, zero diagonal, and which satisfies (8.8), is called a *Euclidean distance matrix.* A matrix is a Euclidean distance matrix if and only if its entries are the squares of the Euclidean distances between the vectors of some configuration. (Given a Euclidean distance matrix D and the associated length squared vector z, we can reconstruct one, or all, configurations with the given pairwise distances using the method described above.)

The condition (8.8) turns out to be equivalent to the simpler condition that D is negative semidefinite on $\mathbf{1}^\perp$, *i.e.*,

$$\begin{aligned} (8.8) \quad &\Longleftrightarrow \quad u^T D u \leq 0 \text{ for all } u \text{ with } \mathbf{1}^T u = 0 \\ &\Longleftrightarrow \quad (I - (1/n)\mathbf{1}\mathbf{1}^T)D(I - (1/n)\mathbf{1}\mathbf{1}^T) \preceq 0. \end{aligned}$$

This simple matrix inequality, along with $D_{ij} \geq 0$, $D_{ii} = 0$, is the classical characterization of a Euclidean distance matrix. To see the equivalence, recall that we can assume $A\mathbf{1} = 0$, which implies that $\mathbf{1}^T G\mathbf{1} = \mathbf{1}^T A^T A\mathbf{1} = 0$. It follows that $G \succeq 0$ if and only if G is positive semidefinite on $\mathbf{1}^\perp$, *i.e.*,

$$\begin{aligned} 0 &\preceq (I - (1/n)\mathbf{1}\mathbf{1}^T)G(I - (1/n)\mathbf{1}\mathbf{1}^T) \\ &= (1/2)(I - (1/n)\mathbf{1}\mathbf{1}^T)(z\mathbf{1}^T + \mathbf{1}z^T - D)(I - (1/n)\mathbf{1}\mathbf{1}^T) \\ &= -(1/2)(I - (1/n)\mathbf{1}\mathbf{1}^T)D(I - (1/n)\mathbf{1}\mathbf{1}^T), \end{aligned}$$

which is the simplified condition.

In summary, a matrix $D \in \mathbf{S}^n$ is a Euclidean distance matrix, *i.e.*, gives the squared distances between a set of n vectors in $\mathbf{R}^n$, if and only if

$$D_{ii} = 0, \quad i = 1, \ldots, n, \qquad D_{ij} \geq 0, \quad i, j = 1, \ldots, n,$$

$$(I - (1/n)\mathbf{1}\mathbf{1}^T)D(I - (1/n)\mathbf{1}\mathbf{1}^T) \preceq 0,$$

which is a set of linear equalities, linear inequalities, and a matrix inequality in D. Therefore we can express any Euclidean distance problem that is convex in the squared distances as a convex problem with variable $D \in \mathbf{S}^n$.

8.4 Extremal volume ellipsoids

Suppose $C \subseteq \mathbf{R}^n$ is bounded and has nonempty interior. In this section we consider the problems of finding the maximum volume ellipsoid that lies inside C, and the minimum volume ellipsoid that covers C. Both problems can be formulated as convex programming problems, but are tractable only in special cases.

8.4.1 The Löwner-John ellipsoid

The minimum volume ellipsoid that contains a set C is called the *Löwner-John ellipsoid* of the set C, and is denoted $\mathcal{E}_{\rm lj}$. To characterize $\mathcal{E}_{\rm lj}$, it will be convenient to parametrize a general ellipsoid as

$$\mathcal{E} = \{v \mid \|Av + b\|_2 \leq 1\}, \tag{8.9}$$

i.e., the inverse image of the Euclidean unit ball under an affine mapping. We can assume without loss of generality that $A \in \mathbf{S}^n_{++}$, in which case the volume of $\mathcal{E}$ is proportional to $\det A^{-1}$. The problem of computing the minimum volume ellipsoid containing C can be expressed as

$$\begin{array}{ll} \text{minimize} & \log\det A^{-1} \\ \text{subject to} & \sup_{v \in C} \|Av + b\|_2 \leq 1, \end{array} \tag{8.10}$$

where the variables are $A \in \mathbf{S}^n$ and $b \in \mathbf{R}^n$, and there is an implicit constraint $A \succ 0$. The objective and constraint functions are both convex in A and b, so the problem (8.10) is convex. Evaluating the constraint function in (8.10), however, involves solving a convex maximization problem, and is tractable only in certain special cases.

Minimum volume ellipsoid covering a finite set

We consider the problem of finding the minimum volume ellipsoid that contains the finite set $C = \{x_1, \ldots, x_m\} \subseteq \mathbf{R}^n$. An ellipsoid covers C if and only if it covers its convex hull, so finding the minimum volume ellipsoid that covers C

is the same as finding the minimum volume ellipsoid containing the polyhedron $\mathbf{conv}\{x_1, \ldots, x_m\}$. Applying (8.10), we can write this problem as

$$\begin{array}{ll} \text{minimize} & \log\det A^{-1} \\ \text{subject to} & \|Ax_i + b\|_2 \leq 1, \quad i = 1, \ldots, m \end{array} \tag{8.11}$$

where the variables are $A \in \mathbf{S}^n$ and $b \in \mathbf{R}^n$, and we have the implicit constraint $A \succ 0$. The norm constraints $\|Ax_i + b\|_2 \leq 1$, $i = 1, \ldots, m$, are convex inequalities in the variables A and b. They can be replaced with the squared versions, $\|Ax_i + b\|_2^2 \leq 1$, which are convex quadratic inequalities in A and b.

Minimum volume ellipsoid covering union of ellipsoids

Minimum volume covering ellipsoids can also be computed efficiently for certain sets C that are defined by quadratic inequalities. In particular, it is possible to compute the Löwner-John ellipsoid for a union or sum of ellipsoids.

As an example, consider the problem of finding the minimum volume ellipsoid $\mathcal{E}_{\text{lj}}$, that contains the ellipsoids $\mathcal{E}_1, \ldots, \mathcal{E}_m$ (and therefore, the convex hull of their union). The ellipsoids $\mathcal{E}_1, \ldots, \mathcal{E}_m$ will be described by (convex) quadratic inequalities:

$$\mathcal{E}_i = \{x \mid x^T A_i x + 2b_i^T x + c_i \leq 0\}, \quad i = 1, \ldots, m,$$

where $A_i \in \mathbf{S}^n_{++}$. We parametrize the ellipsoid $\mathcal{E}_{\text{lj}}$ as

$$\begin{array}{rcl} \mathcal{E}_{\text{lj}} & = & \{x \mid \|Ax + b\|_2 \leq 1\} \\ & = & \{x \mid x^T A^T A x + 2(A^T b)^T x + b^T b - 1 \leq 0\} \end{array}$$

where $A \in \mathbf{S}^n$ and $b \in \mathbf{R}^n$. Now we use a result from §B.2, that $\mathcal{E}_i \subseteq \mathcal{E}_{\text{lj}}$ if and only if there exists a $\tau \geq 0$ such that

$$\left[\begin{array}{cc} A^2 - \tau A_i & Ab - \tau b_i \\ (Ab - \tau b_i)^T & b^T b - 1 - \tau c_i \end{array}\right] \preceq 0.$$

The volume of $\mathcal{E}_{\text{lj}}$ is proportional to $\det A^{-1}$, so we can find the minimum volume ellipsoid that contains $\mathcal{E}_1, \ldots, \mathcal{E}_m$ by solving

$$\begin{array}{ll} \text{minimize} & \log\det A^{-1} \\ \text{subject to} & \tau_1 \geq 0, \ldots, \tau_m \geq 0 \\ & \left[\begin{array}{cc} A^2 - \tau_i A_i & Ab - \tau_i b_i \\ (Ab - \tau_i b_i)^T & b^T b - 1 - \tau_i c_i \end{array}\right] \preceq 0, \quad i = 1, \ldots, m, \end{array}$$

or, replacing the variable b by $\tilde{b} = Ab$,

$$\begin{array}{ll} \text{minimize} & \log\det A^{-1} \\ \text{subject to} & \tau_1 \geq 0, \ldots, \tau_m \geq 0 \\ & \left[\begin{array}{ccc} A^2 - \tau_i A_i & \tilde{b} - \tau_i b_i & 0 \\ (\tilde{b} - \tau_i b_i)^T & -1 - \tau_i c_i & \tilde{b}^T \\ 0 & \tilde{b} & -A^2 \end{array}\right] \preceq 0, \quad i = 1, \ldots, m, \end{array}$$

which is convex in the variables $A^2 \in \mathbf{S}^n$, $\tilde{b}$, $\tau_1, \ldots, \tau_m$.

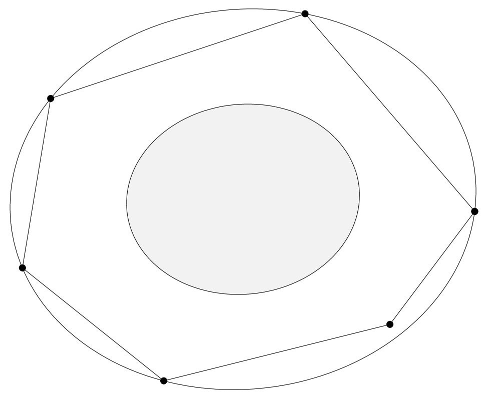


Figure 8.3 The outer ellipse is the boundary of the Löwner-John ellipsoid, *i.e.*, the minimum volume ellipsoid that encloses the points $x_1, \ldots, x_6$ (shown as dots), and therefore the polyhedron $\mathcal{P} = \mathbf{conv}\{x_1, \ldots, x_6\}$. The smaller ellipse is the boundary of the Löwner-John ellipsoid, shrunk by a factor of $n = 2$ about its center. This ellipsoid is guaranteed to lie *inside* $\mathcal{P}$.

Efficiency of Löwner-John ellipsoidal approximation

Let $\mathcal{E}_{\rm lj}$ be the Löwner-John ellipsoid of the convex set $C \subseteq \mathbf{R}^n$, which is bounded and has nonempty interior, and let x_0 be its center. If we shrink the Löwner-John ellipsoid by a factor of n, about its center, we obtain an ellipsoid that lies inside the set C:

$$x_0 + (1/n)(\mathcal{E}_{\rm lj} - x_0) \subseteq C \subseteq \mathcal{E}_{\rm lj}.$$

In other words, the Löwner-John ellipsoid approximates an arbitrary convex set, within a factor that depends only on the dimension n. Figure 8.3 shows a simple example.

The factor $1/n$ cannot be improved without additional assumptions on C. Any simplex in $\mathbf{R}^n$, for example, has the property that its Löwner-John ellipsoid must be shrunk by a factor n to fit inside it (see exercise 8.13).

We will prove this efficiency result for the special case $C = \mathbf{conv}\{x_1, \ldots, x_m\}$. We square the norm constraints in (8.11) and introduce variables $\tilde{A} = A^2$ and $\tilde{b} = Ab$, to obtain the problem

$$\begin{array}{ll} \text{minimize} & \log\det \tilde{A}^{-1} \\ \text{subject to} & x_i^T \tilde{A} x_i - 2\tilde{b}^T x_i + \tilde{b}^T \tilde{A}^{-1} \tilde{b} \le 1, \quad i = 1, \ldots, m. \end{array} \qquad (8.12)$$

The KKT conditions for this problem are

$$\begin{array}{c} \sum_{i=1}^m \lambda_i (x_i x_i^T - \tilde{A}^{-1}\tilde{b}\tilde{b}^T\tilde{A}^{-1}) = \tilde{A}^{-1}, \qquad \sum_{i=1}^m \lambda_i (x_i - \tilde{A}^{-1}\tilde{b}) = 0, \\ \lambda_i \ge 0, \qquad x_i^T \tilde{A} x_i - 2\tilde{b}^T x_i + \tilde{b}^T \tilde{A}^{-1} \tilde{b} \le 1, \qquad i = 1, \ldots, m, \\ \lambda_i (1 - x_i^T \tilde{A} x_i + 2\tilde{b}^T x_i - \tilde{b}^T \tilde{A}^{-1} \tilde{b}) = 0, \quad i = 1, \ldots, m. \end{array}$$

By a suitable affine change of coordinates, we can assume that $\tilde{A} = I$ and $\tilde{b} = 0$, *i.e.*, the minimum volume ellipsoid is the unit ball centered at the origin. The KKT conditions then simplify to

$$\sum_{i=1}^{m} \lambda_i x_i x_i^T = I, \qquad \sum_{i=1}^{m} \lambda_i x_i = 0, \qquad \lambda_i(1 - x_i^T x_i) = 0, \quad i = 1, \dots, m,$$

plus the feasibility conditions $\|x_i\|_2 \leq 1$ and $\lambda_i \geq 0$. By taking the trace of both sides of the first equation, and using complementary slackness, we also have $\sum_{i=1}^m \lambda_i = n$.

In the new coordinates the shrunk ellipsoid is a ball with radius $1/n$, centered at the origin. We need to show that

$$\|x\|_2 \leq 1/n \implies x \in C = \mathbf{conv}\{x_1, \dots, x_m\}.$$

Suppose $\|x\|_2 \leq 1/n$. From the KKT conditions, we see that

$$x = \sum_{i=1}^{m} \lambda_i (x^T x_i) x_i = \sum_{i=1}^{m} \lambda_i (x^T x_i + 1/n) x_i = \sum_{i=1}^{m} \mu_i x_i, \tag{8.13}$$

where $\mu_i = \lambda_i(x^T x_i + 1/n)$. From the Cauchy-Schwartz inequality, we note that

$$\mu_i = \lambda_i(x^T x_i + 1/n) \geq \lambda_i(-\|x\|_2 \|x_i\|_2 + 1/n) \geq \lambda_i(-1/n + 1/n) = 0.$$

Furthermore

$$\sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} \lambda_i (x^T x_i + 1/n) = \sum_{i=1}^{m} \lambda_i / n = 1.$$

This, along with (8.13), shows that x is a convex combination of $x_1, \dots, x_m$, hence $x \in C$.

Efficiency of Löwner-John ellipsoidal approximation for symmetric sets

If the set C is symmetric about a point x_0, then the factor $1/n$ can be tightened to $1/\sqrt{n}$:

$$x_0 + (1/\sqrt{n})(\mathcal{E}_{\text{lj}} - x_0) \subseteq C \subseteq \mathcal{E}_{\text{lj}}.$$

Again, the factor $1/\sqrt{n}$ is tight. The Löwner-John ellipsoid of the cube

$$C = \{x \in \mathbf{R}^n \mid -\mathbf{1} \preceq x \preceq \mathbf{1}\}$$

is the ball with radius $\sqrt{n}$. Scaling down by $1/\sqrt{n}$ yields a ball enclosed in C, and touching the boundary at $x = \pm e_i$.

Approximating a norm by a quadratic norm

Let $\|\cdot\|$ be any norm on $\mathbf{R}^n$, and let $C = \{x \mid \|x\| \leq 1\}$ be its unit ball. Let $\mathcal{E}_{\text{lj}} = \{x \mid x^T A x \leq 1\}$, with $A \in \mathbf{S}^n_{++}$, be the Löwner-John ellipsoid of C. Since C is symmetric about the origin, the result above tells us that $(1/\sqrt{n})\mathcal{E}_{\text{lj}} \subseteq C \subseteq \mathcal{E}_{\text{lj}}$. Let $\|\cdot\|_{\text{lj}}$ denote the quadratic norm

$$\|z\|_{\text{lj}} = (z^T A z)^{1/2},$$

whose unit ball is $\mathcal{E}_{\rm lj}$. The inclusions $(1/\sqrt{n})\mathcal{E}_{\rm lj} \subseteq C \subseteq \mathcal{E}_{\rm lj}$ are equivalent to the inequalities

$$\|z\|_{\rm lj} \leq \|z\| \leq \sqrt{n}\|z\|_{\rm lj}$$

for all $z \in \mathbf{R}^n$. In other words, the quadratic norm $\|\cdot\|_{\rm lj}$ approximates the norm $\|\cdot\|$ within a factor of $\sqrt{n}$. In particular, we see that any norm on $\mathbf{R}^n$ can be approximated within a factor of $\sqrt{n}$ by a quadratic norm.

8.4.2 Maximum volume inscribed ellipsoid

We now consider the problem of finding the ellipsoid of maximum volume that lies inside a convex set C, which we assume is bounded and has nonempty interior. To formulate this problem, we parametrize the ellipsoid as the image of the unit ball under an affine transformation, *i.e.*, as

$$\mathcal{E} = \{Bu + d \mid \|u\|_2 \leq 1\}.$$

Again it can be assumed that $B \in \mathbf{S}^n_{++}$, so the volume is proportional to $\det B$. We can find the maximum volume ellipsoid inside C by solving the convex optimization problem

$$\begin{array}{ll} \text{maximize} & \log\det B \\ \text{subject to} & \sup_{\|u\|_2 \leq 1} I_C(Bu + d) \leq 0 \end{array} \tag{8.14}$$

in the variables $B \in \mathbf{S}^n$ and $d \in \mathbf{R}^n$, with implicit constraint $B \succ 0$.

Maximum volume ellipsoid in a polyhedron

We consider the case where C is a polyhedron described by a set of linear inequalities:

$$C = \{x \mid a_i^T x \leq b_i,\ i = 1, \ldots, m\}.$$

To apply (8.14) we first express the constraint in a more convenient form:

$$\begin{aligned} \sup_{\|u\|_2 \leq 1} I_C(Bu + d) \leq 0 \quad &\Longleftrightarrow \quad \sup_{\|u\|_2 \leq 1} a_i^T(Bu + d) \leq b_i, \quad i = 1, \ldots, m \\ &\Longleftrightarrow \quad \|Ba_i\|_2 + a_i^T d \leq b_i, \quad i = 1, \ldots, m. \end{aligned}$$

We can therefore formulate (8.14) as a convex optimization problem in the variables B and d:

$$\begin{array}{ll} \text{minimize} & \log\det B^{-1} \\ \text{subject to} & \|Ba_i\|_2 + a_i^T d \leq b_i, \quad i = 1, \ldots, m. \end{array} \tag{8.15}$$

Maximum volume ellipsoid in an intersection of ellipsoids

We can also find the maximum volume ellipsoid $\mathcal{E}$ that lies in the intersection of m ellipsoids $\mathcal{E}_1, \ldots, \mathcal{E}_m$. We will describe $\mathcal{E}$ as $\mathcal{E} = \{Bu + d \mid \|u\|_2 \leq 1\}$ with $B \in \mathbf{S}^n_{++}$, and the other ellipsoids via convex quadratic inequalities,

$$\mathcal{E}_i = \{x \mid x^T A_i x + 2b_i^T x + c_i \leq 0\}, \quad i = 1, \ldots, m,$$

where $A_i \in \mathbf{S}^n_{++}$. We first work out the condition under which $\mathcal{E} \subseteq \mathcal{E}_i$. This occurs if and only if

$$\begin{aligned}
&\sup_{\|u\|_2 \le 1} \left((d+Bu)^T A_i (d+Bu) + 2b_i^T (d+Bu) + c_i \right) \\
&\quad = \; d^T A_i d + 2b_i^T d + c_i + \sup_{\|u\|_2 \le 1} \left(u^T B A_i B u + 2(A_i d + b_i)^T B u \right) \\
&\quad \le \; 0.
\end{aligned}$$

From §B.1,

$$\sup_{\|u\|_2 \le 1} \left(u^T B A_i B u + 2(A_i d + b_i)^T B u \right) \le -(d^T A_i d + 2b_i^T d + c_i)$$

if and only if there exists a $\lambda_i \ge 0$ such that

$$\begin{bmatrix} -\lambda_i - d^T A_i d - 2b_i^T d - c_i & (A_i d + b_i)^T B \\ B(A_i d + b_i) & \lambda_i I - B A_i B \end{bmatrix} \succeq 0.$$

The maximum volume ellipsoid contained in $\mathcal{E}_1, \ldots, \mathcal{E}_m$ can therefore be found by solving the problem

$$\begin{array}{ll}
\text{minimize} & \log\det B^{-1} \\
\text{subject to} & \begin{bmatrix} -\lambda_i - d^T A_i d - 2b_i^T d - c_i & (A_i d + b_i)^T B \\ B(A_i d + b_i) & \lambda_i I - B A_i B \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m,
\end{array}$$

with variables $B \in \mathbf{S}^n$, $d \in \mathbf{R}^n$, and $\lambda \in \mathbf{R}^m$, or, equivalently,

$$\begin{array}{ll}
\text{minimize} & \log\det B^{-1} \\
\text{subject to} & \begin{bmatrix} -\lambda_i - c_i + b_i^T A_i^{-1} b_i & 0 & (d + A_i^{-1} b_i)^T \\ 0 & \lambda_i I & B \\ d + A_i^{-1} b_i & B & A_i^{-1} \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m.
\end{array}$$

Efficiency of ellipsoidal inner approximations

Approximation efficiency results, similar to the ones for the Löwner-John ellipsoid, hold for the maximum volume inscribed ellipsoid. If $C \subseteq \mathbf{R}^n$ is convex, bounded, with nonempty interior, then the maximum volume inscribed ellipsoid, expanded by a factor of n about its center, covers the set C. The factor n can be tightened to $\sqrt{n}$ if the set C is symmetric about a point. An example is shown in figure 8.4.

8.4.3 Affine invariance of extremal volume ellipsoids

The Löwner-John ellipsoid and the maximum volume inscribed ellipsoid are both affinely invariant. If $\mathcal{E}_{\rm lj}$ is the Löwner-John ellipsoid of C, and $T \in \mathbf{R}^{n \times n}$ is nonsingular, then the Löwner-John ellipsoid of TC is $T\mathcal{E}_{\rm lj}$. A similar result holds for the maximum volume inscribed ellipsoid.

To establish this result, let $\mathcal{E}$ be any ellipsoid that covers C. Then the ellipsoid $T\mathcal{E}$ covers TC. The converse is also true: Every ellipsoid that covers TC has

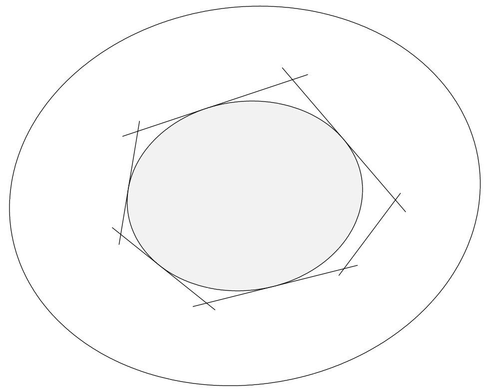


Figure 8.4 The maximum volume ellipsoid (shown shaded) inscribed in a polyhedron $\mathcal{P}$. The outer ellipse is the boundary of the inner ellipsoid, expanded by a factor $n = 2$ about its center. The expanded ellipsoid is guaranteed to cover $\mathcal{P}$.

the form $T\mathcal{E}$, where $\mathcal{E}$ is an ellipsoid that covers C. In other words, the relation $\tilde{\mathcal{E}} = T\mathcal{E}$ gives a one-to-one correspondence between the ellipsoids covering TC and the ellipsoids covering C. Moreover, the volumes of the corresponding ellipsoids are all related by the ratio $|\det T|$, so in particular, if $\mathcal{E}$ has minimum volume among ellipsoids covering C, then $T\mathcal{E}$ has minimum volume among ellipsoids covering TC.

8.5 Centering

8.5.1 Chebyshev center

Let $C \subseteq \mathbf{R}^n$ be bounded and have nonempty interior, and $x \in C$. The *depth* of a point $x \in C$ is defined as

$$\mathbf{depth}(x, C) = \mathbf{dist}(x, \mathbf{R}^n \setminus C),$$

i.e., the distance to the closest point in the exterior of C. The depth gives the radius of the largest ball, centered at x, that lies in C. A *Chebyshev center* of the set C is defined as any point of maximum depth in C:

$$x_{\text{cheb}}(C) = \operatorname{argmax} \mathbf{depth}(x, C) = \operatorname{argmax} \mathbf{dist}(x, \mathbf{R}^n \setminus C).$$

A Chebyshev center is a point inside C that is farthest from the exterior of C; it is also the center of the largest ball that lies inside C. Figure 8.5 shows an example, in which C is a polyhedron, and the norm is Euclidean.

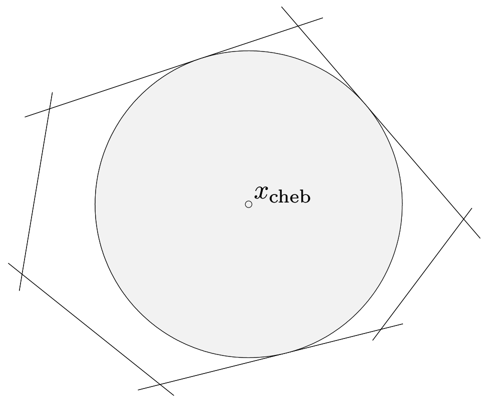


Figure 8.5 Chebyshev center of a polyhedron C, in the Euclidean norm. The center $x_{\rm cheb}$ is the deepest point inside C, in the sense that it is farthest from the exterior, or complement, of C. The center $x_{\rm cheb}$ is also the center of the largest Euclidean ball (shown lightly shaded) that lies inside C.

Chebyshev center of a convex set

When the set C is convex, the depth is a concave function for $x \in C$, so computing the Chebyshev center is a convex optimization problem (see exercise 8.5). More specifically, suppose $C \subseteq \mathbf{R}^n$ is defined by a set of convex inequalities:

$$C = \{x \mid f_1(x) \leq 0, \ldots, f_m(x) \leq 0\}.$$

We can find a Chebyshev center by solving the problem

$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & g_i(x, R) \leq 0, \quad i = 1, \ldots, m, \end{array} \tag{8.16}$$

where g_i is defined as

$$g_i(x, R) = \sup_{\|u\| \leq 1} f_i(x + Ru).$$

Problem (8.16) is a convex optimization problem, since each function g_i is the pointwise maximum of a family of convex functions of x and R, hence convex. However, evaluating g_i involves solving a convex *maximization* problem (either numerically or analytically), which may be very hard. In practice, we can find the Chebyshev center only in cases where the functions g_i are easy to evaluate.

Chebyshev center of a polyhedron

Suppose C is defined by a set of linear inequalities $a_i^T x \leq b_i$, $i = 1, \ldots, m$. We have

$$g_i(x, R) = \sup_{\|u\| \leq 1} a_i^T (x + Ru) - b_i = a_i^T x + R\|a_i\|_* - b_i$$

if $R \geq 0$, so the Chebyshev center can be found by solving the LP

$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & a_i^T x + R\|a_i\|_* \leq b_i, \quad i = 1, \ldots, m \\ & R \geq 0 \end{array}$$

with variables x and R.

Euclidean Chebyshev center of intersection of ellipsoids

Let C be an intersection of m ellipsoids, defined by quadratic inequalities,

$$C = \{x \mid x^T A_i x + 2b_i^T x + c_i \leq 0, \ i = 1, \ldots, m\},$$

where $A_i \in \mathbf{S}^n_{++}$. We have

$$\begin{array}{rcl} g_i(x, R) & = & \sup_{\|u\|_2 \leq 1} \left((x + Ru)^T A_i (x + Ru) + 2b_i^T (x + Ru) + c_i \right) \\ & = & x^T A_i x + 2b_i^T x + c_i + \sup_{\|u\|_2 \leq 1} \left(R^2 u^T A_i u + 2R(A_i x + b_i)^T u \right). \end{array}$$

From §B.1, $g_i(x, R) \leq 0$ if and only if there exists a λ_i such that the matrix inequality

$$\left[\begin{array}{cc} -x^T A_i x_i - 2b_i^T x - c_i - \lambda_i & R(A_i x + b_i)^T \\ R(A_i x + b_i) & \lambda_i I - R^2 A_i \end{array} \right] \succeq 0 \qquad (8.17)$$

holds. Using this result, we can express the Chebyshev centering problem as

$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & \left[\begin{array}{ccc} -\lambda_i - c_i + b_i^T A_i^{-1} b_i & 0 & (x + A_i^{-1} b_i)^T \\ 0 & \lambda_i I & RI \\ x + A_i^{-1} b_i & RI & A_i^{-1} \end{array} \right] \succeq 0, \quad i = 1, \ldots, m, \end{array}$$

which is an SDP with variables R, λ, and x. Note that the Schur complement of A_i^{-1} in the LMI constraint is equal to the lefthand side of (8.17).

8.5.2 Maximum volume ellipsoid center

The Chebyshev center x_{cheb} of a set $C \subseteq \mathbf{R}^n$ is the center of the largest ball that lies in C. As an extension of this idea, we define the *maximum volume ellipsoid center* of C, denoted x_{mve}, as the center of the maximum volume ellipsoid that lies in C. Figure 8.6 shows an example, where C is a polyhedron.

The maximum volume ellipsoid center is readily computed when C is defined by a set of linear inequalities, by solving the problem (8.15). (The optimal value of the variable $d \in \mathbf{R}^n$ is x_{mve}.) Since the maximum volume ellipsoid inside C is affine invariant, so is the maximum volume ellipsoid center.

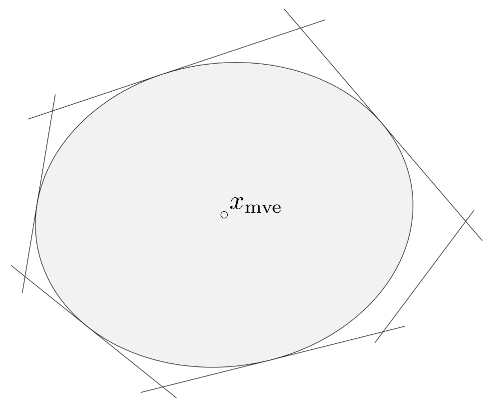


Figure 8.6 The lightly shaded ellipsoid shows the maximum volume ellipsoid contained in the set C, which is the same polyhedron as in figure 8.5. Its center $x_{\rm mve}$ is the maximum volume ellipsoid center of C.

8.5.3 Analytic center of a set of inequalities

The *analytic center* $x_{\rm ac}$ of a set of convex inequalities and linear equalities,

$$f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad Fx = g$$

is defined as an optimal point for the (convex) problem

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^m \log(-f_i(x)) \\ \text{subject to} & Fx = g, \end{array} \tag{8.18}$$

with variable $x \in \mathbf{R}^n$ and implicit constraints $f_i(x) < 0$, $i = 1, \ldots, m$. The objective in (8.18) is called the *logarithmic barrier* associated with the set of inequalities. We assume here that the domain of the logarithmic barrier intersects the affine set defined by the equalities, *i.e.*, the strict inequality system

$$f_i(x) < 0, \quad i = 1, \ldots, m, \qquad Fx = g$$

is feasible. The logarithmic barrier is bounded below on the feasible set

$$C = \{x \mid f_i(x) < 0, \ i = 1, \ldots, m, \ Fx = g\},$$

if C is bounded.

When x is strictly feasible, *i.e.*, $Fx = g$ and $f_i(x) < 0$ for $i = 1, \ldots, m$, we can interpret $-f_i(x)$ as the margin or slack in the ith inequality. The analytic center $x_{\rm ac}$ is the point that maximizes the product (or geometric mean) of these slacks or margins, subject to the equality constraints $Fx = g$, and the implicit constraints $f_i(x) < 0$.

The analytic center is *not* a function of the set C described by the inequalities and equalities; two sets of inequalities and equalities can define the same set, but have different analytic centers. Still, it is not uncommon to informally use the term 'analytic center of a set C' to mean the analytic center of a particular set of equalities and inequalities that define it.

The analytic center is, however, independent of affine changes of coordinates. It is also invariant under (positive) scalings of the inequality functions, and any reparametrization of the equality constraints. In other words, if $\tilde{F}$ and $\tilde{g}$ are such that $\tilde{F}x = \tilde{g}$ if and only if $Fx = g$, and $\alpha_1, \ldots, \alpha_m > 0$, then the analytic center of

$$\alpha_i f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad \tilde{F}x = \tilde{g},$$

is the same as the analytic center of

$$f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad Fx = g$$

(see exercise 8.17).

Analytic center of a set of linear inequalities

The analytic center of a set of linear inequalities

$$a_i^T x \leq b_i, \quad i = 1, \ldots, m,$$

is the solution of the unconstrained minimization problem

$$\text{minimize} \quad -\textstyle\sum_{i=1}^m \log(b_i - a_i^T x), \tag{8.19}$$

with implicit constraint $b_i - a_i^T x > 0$, $i = 1, \ldots, m$. If the polyhedron defined by the linear inequalities is bounded, then the logarithmic barrier is bounded below and strictly convex, so the analytic center in unique. (See exercise 4.2.)

We can give a geometric interpretation of the analytic center of a set of linear inequalities. Since the analytic center is independent of positive scaling of the constraint functions, we can assume without loss of generality that $\|a_i\|_2 = 1$. In this case, the slack $b_i - a_i^T x$ is the distance to the hyperplane $\mathcal{H}_i = \{x \mid a_i^T x = b_i\}$. Therefore the analytic center $x_{\rm ac}$ is the point that maximizes the product of distances to the defining hyperplanes.

Inner and outer ellipsoids from analytic center of linear inequalities

The analytic center of a set of linear inequalities implicitly defines an inscribed and a covering ellipsoid, defined by the Hessian of the logarithmic barrier function

$$-\sum_{i=1}^m \log(b_i - a_i^T x),$$

evaluated at the analytic center, *i.e.*,

$$H = \sum_{i=1}^m d_i^2 a_i a_i^T, \qquad d_i = \frac{1}{b_i - a_i^T x_{\rm ac}}, \quad i = 1, \ldots, m.$$

We have $\mathcal{E}_{\rm inner} \subseteq \mathcal{P} \subseteq \mathcal{E}_{\rm outer}$, where

$$\begin{array}{rcl} \mathcal{P} & = & \{x \mid a_i^T x \leq b_i,\ i = 1, \ldots, m\}, \\ \mathcal{E}_{\rm inner} & = & \{x \mid (x - x_{\rm ac})^T H (x - x_{\rm ac}) \leq 1\}, \\ \mathcal{E}_{\rm outer} & = & \{x \mid x - x_{\rm ac})^T H (x - x_{\rm ac}) \leq m(m-1)\}. \end{array}$$

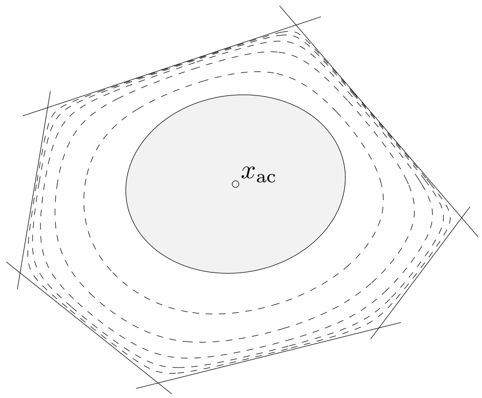


Figure 8.7 The dashed lines show five level curves of the logarithmic barrier function for the inequalities defining the polyhedron C in figure 8.5. The minimizer of the logarithmic barrier function, labeled $x_{\rm ac}$, is the analytic center of the inequalities. The inner ellipsoid $\mathcal{E}_{\rm inner} = \{x \mid (x - x_{\rm ac})H(x - x_{\rm ac}) \leq 1\}$, where H is the Hessian of the logarithmic barrier function at $x_{\rm ac}$, is shaded.

This is a weaker result than the one for the maximum volume inscribed ellipsoid, which when scaled up by a factor of n covers the polyhedron. The inner and outer ellipsoids defined by the Hessian of the logarithmic barrier, in contrast, are related by the scale factor $(m(m-1))^{1/2}$, which is always at least n.

To show that $\mathcal{E}_{\rm inner} \subseteq \mathcal{P}$, suppose $x \in \mathcal{E}_{\rm inner}$, *i.e.*,

$$(x - x_{\rm ac})^T H(x - x_{\rm ac}) = \sum_{i=1}^m (d_i a_i^T(x - x_{\rm ac}))^2 \leq 1.$$

This implies that

$$a_i^T(x - x_{\rm ac}) \leq 1/d_i = b_i - a_i^T x_{\rm ac}, \quad i = 1, \ldots, m,$$

and therefore $a_i^T x \leq b_i$ for $i = 1, \ldots, m$. (We have not used the fact that $x_{\rm ac}$ is the analytic center, so this result is valid if we replace $x_{\rm ac}$ with any strictly feasible point.)

To establish that $\mathcal{P} \subseteq \mathcal{E}_{\rm outer}$, we will need the fact that $x_{\rm ac}$ is the analytic center, and therefore the gradient of the logarithmic barrier vanishes:

$$\sum_{i=1}^m d_i a_i = 0.$$

Now assume $x \in \mathcal{P}$. Then

$$\begin{array}{rl} & (x - x_{\rm ac})^T H(x - x_{\rm ac}) \\ = & \displaystyle\sum_{i=1}^m (d_i a_i^T(x - x_{\rm ac}))^2 \end{array}$$

$$\begin{aligned}
&= \sum_{i=1}^m d_i^2(1/d_i - a_i^T(x - x_{\rm ac}))^2 - m \\
&= \sum_{i=1}^m d_i^2(b_i - a_i^T x)^2 - m \\
&\leq \left(\sum_{i=1}^m d_i(b_i - a_i^T x)\right)^2 - m \\
&= \left(\sum_{i=1}^m d_i(b_i - a_i^T x_{\rm ac}) + \sum_{i=1}^m d_i a_i^T(x_{\rm ac} - x)\right)^2 - m \\
&= m^2 - m,
\end{aligned}$$

which shows that $x \in \mathcal{E}_{\rm outer}$. (The second equality follows from the fact that $\sum_{i=1}^m d_i a_i = 0$. The inequality follows from $\sum_{i=1}^m y_i^2 \leq \left(\sum_{i=1}^m y_i\right)^2$ for $y \succeq 0$. The last equality follows from $\sum_{i=1}^m d_i a_i = 0$, and the definition of d_i.)

Analytic center of a linear matrix inequality

The definition of analytic center can be extended to sets described by generalized inequalities with respect to a cone K, if we define a logarithm on K. For example, the analytic center of a linear matrix inequality

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n \preceq B$$

is defined as the solution of

$$\text{minimize} \quad -\log\det(B - x_1 A_1 - \cdots - x_n A_n).$$

8.6 Classification

In pattern recognition and classification problems we are given two sets of points in $\mathbf{R}^n$, $\{x_1, \ldots, x_N\}$ and $\{y_1, \ldots, y_M\}$, and wish to find a function $f : \mathbf{R}^n \to \mathbf{R}$ (within a given family of functions) that is positive on the first set and negative on the second, *i.e.*,

$$f(x_i) > 0, \quad i = 1, \ldots, N, \qquad f(y_i) < 0, \quad i = 1, \ldots, M.$$

If these inequalities hold, we say that f, or its 0-level set $\{x \mid f(x) = 0\}$, *separates*, *classifies*, or *discriminates* the two sets of points. We sometimes also consider *weak separation*, in which the weak versions of the inequalities hold.

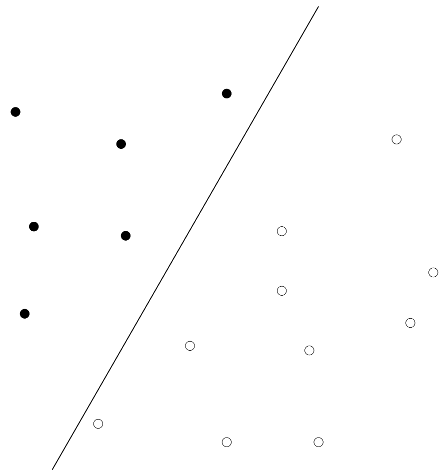


Figure 8.8 The points $x_1, \ldots, x_N$ are shown as open circles, and the points $y_1, \ldots, y_M$ are shown as filled circles. These two sets are classified by an affine function f, whose 0-level set (a line) separates them.

8.6.1 Linear discrimination

In linear discrimination, we seek an affine function $f(x) = a^T x - b$ that classifies the points, *i.e.*,

$$a^T x_i - b > 0, \quad i = 1, \ldots, N, \qquad a^T y_i - b < 0, \quad i = 1, \ldots, M. \tag{8.20}$$

Geometrically, we seek a hyperplane that separates the two sets of points. Since the strict inequalities (8.20) are homogeneous in a and b, they are feasible if and only if the set of nonstrict linear inequalities

$$a^T x_i - b \geq 1, \quad i = 1, \ldots, N, \qquad a^T y_i - b \leq -1, \quad i = 1, \ldots, M \tag{8.21}$$

(in the variables a, b) is feasible. Figure 8.8 shows a simple example of two sets of points and a linear discriminating function.

Linear discrimination alternative

The strong alternative of the set of strict inequalities (8.20) is the existence of λ, $\tilde{\lambda}$ such that

$$\lambda \succeq 0, \qquad \tilde{\lambda} \succeq 0, \qquad (\lambda, \tilde{\lambda}) \neq 0, \qquad \sum_{i=1}^{N} \lambda_i x_i = \sum_{i=1}^{M} \tilde{\lambda}_i y_i, \qquad \mathbf{1}^T \lambda = \mathbf{1}^T \tilde{\lambda} \tag{8.22}$$

(see §5.8.3). Using the third and last conditions, we can express these alternative conditions as

$$\lambda \succeq 0, \qquad \mathbf{1}^T \lambda = 1, \qquad \tilde{\lambda} \succeq 0, \qquad \mathbf{1}^T \tilde{\lambda} = 1, \qquad \sum_{i=1}^{N} \lambda_i x_i = \sum_{i=1}^{M} \tilde{\lambda}_i y_i$$

(by dividing by $\mathbf{1}^T\lambda$, which is positive, and using the same symbols for the normalized λ and $\tilde{\lambda}$). These conditions have a simple geometric interpretation: They state that there is a point in the convex hull of both $\{x_1, \ldots, x_N\}$ and $\{y_1, \ldots, y_M\}$. In other words: the two sets of points can be linearly discriminated (*i.e.*, discriminated by an affine function) if and only if their convex hulls do not intersect. We have seen this result several times before.

Robust linear discrimination

The existence of an affine classifying function $f(x) = a^T x - b$ is equivalent to a set of linear inequalities in the variables a and b that define f. If the two sets can be linearly discriminated, then there is a polyhedron of affine functions that discriminate them, and we can choose one that optimizes some measure of robustness. We might, for example, seek the function that gives the maximum possible 'gap' between the (positive) values at the points x_i and the (negative) values at the points y_i. To do this we have to normalize a and b, since otherwise we can scale a and b by a positive constant and make the gap in the values arbitrarily large. This leads to the problem

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & a^T x_i - b \geq t, \quad i = 1, \ldots, N \\ & a^T y_i - b \leq -t, \quad i = 1, \ldots, M \\ & \|a\|_2 \leq 1, \end{array} \tag{8.23}$$

with variables a, b, and t. The optimal value $t^\star$ of this convex problem (with linear objective, linear inequalities, and one quadratic inequality) is positive if and only if the two sets of points can be linearly discriminated. In this case the inequality $\|a\|_2 \leq 1$ is always tight at the optimum, *i.e.*, we have $\|a^\star\|_2 = 1$. (See exercise 8.23.)

We can give a simple geometric interpretation of the robust linear discrimination problem (8.23). If $\|a\|_2 = 1$ (as is the case at any optimal point), $a^T x_i - b$ is the Euclidean distance from the point x_i to the separating hyperplane $\mathcal{H} = \{z \mid a^T z = b\}$. Similarly, $b - a^T y_i$ is the distance from the point y_i to the hyperplane. Therefore the problem (8.23) finds the hyperplane that separates the two sets of points, and has maximal distance to the sets. In other words, it finds the thickest *slab* that separates the two sets.

As suggested by the example shown in figure 8.9, the optimal value $t^\star$ (which is half the slab thickness) turns out to be half the distance between the convex hulls of the two sets of points. This can be seen clearly from the dual of the robust linear discrimination problem (8.23). The Lagrangian (for the problem of minimizing $-t$) is

$$-t + \sum_{i=1}^N u_i(t + b - a^T x_i) + \sum_{i=1}^M v_i(t - b + a^T y_i) + \lambda(\|a\|_2 - 1).$$

Minimizing over b and t yields the conditions $\mathbf{1}^T u = 1/2$, $\mathbf{1}^T v = 1/2$. When these hold, we have

$$g(u, v, \lambda) = \inf_a \left(a^T (\sum_{i=1}^M v_i y_i - \sum_{i=1}^N u_i x_i) + \lambda \|a\|_2 - \lambda \right)$$

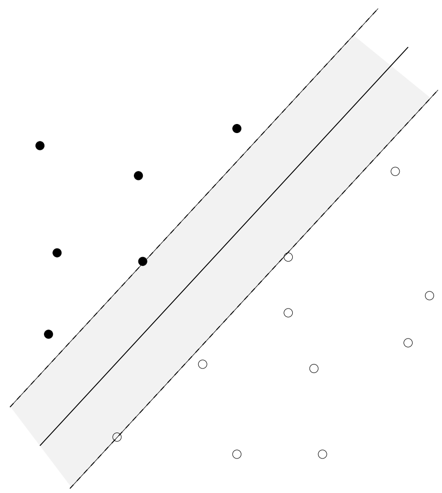


Figure 8.9 By solving the robust linear discrimination problem (8.23) we find an affine function that gives the largest gap in values between the two sets (with a normalization bound on the linear part of the function). Geometrically, we are finding the thickest slab that separates the two sets of points.

$$= \begin{cases} -\lambda & \left\| \sum_{i=1}^M v_i y_i - \sum_{i=1}^N u_i x_i \right\|_2 \leq \lambda \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem can then be written as

$$\begin{array}{ll} \text{maximize} & -\left\| \sum_{i=1}^M v_i y_i - \sum_{i=1}^N u_i x_i \right\|_2 \\ \text{subject to} & u \succeq 0, \quad \mathbf{1}^T u = 1/2 \\ & v \succeq 0, \quad \mathbf{1}^T v = 1/2. \end{array}$$

We can interpret $2\sum_{i=1}^N u_i x_i$ as a point in the convex hull of $\{x_1, \ldots, x_N\}$ and $2\sum_{i=1}^M v_i y_i$ as a point in the convex hull of $\{y_1, \ldots, y_M\}$. The dual objective is to minimize (half) the distance between these two points, *i.e.*, find (half) the distance between the convex hulls of the two sets.

Support vector classifier

When the two sets of points cannot be linearly separated, we might seek an affine function that approximately classifies the points, for example, one that minimizes the number of points misclassified. Unfortunately, this is in general a difficult combinatorial optimization problem. One heuristic for approximate linear discrimination is based on *support vector classifiers*, which we describe in this section.

We start with the feasibility problem (8.21). We first relax the constraints by introducing nonnegative variables $u_1, \ldots, u_N$ and $v_1, \ldots, v_M$, and forming the inequalities

$$a^T x_i - b \geq 1 - u_i, \quad i = 1, \ldots, N, \qquad a^T y_i - b \leq -(1 - v_i), \quad i = 1, \ldots, M. \quad (8.24)$$

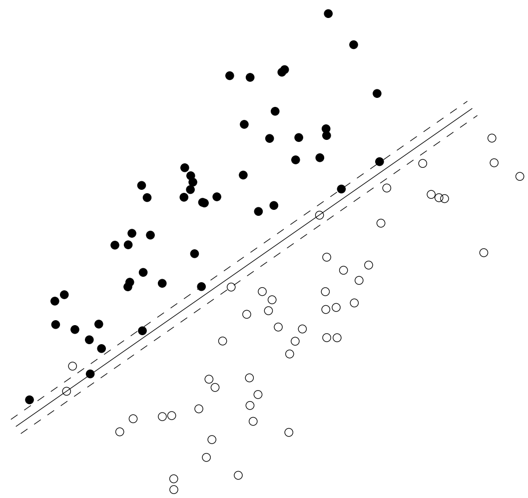


Figure 8.10 Approximate linear discrimination via linear programming. The points $x_1, \ldots, x_{50}$, shown as open circles, cannot be linearly separated from the points $y_1, \ldots, y_{50}$, shown as filled circles. The classifier shown as a solid line was obtained by solving the LP (8.25). This classifier misclassifies one point. The dashed lines are the hyperplanes $a^T z - b = \pm 1$. Four points are correctly classified, but lie in the slab defined by the dashed lines.

When $u = v = 0$, we recover the original constraints; by making u and v large enough, these inequalities can always be made feasible. We can think of u_i as a measure of how much the constraint $a^T x_i - b \geq 1$ is violated, and similarly for v_i. Our goal is to find a, b, and *sparse* nonnegative u and v that satisfy the inequalities (8.24). As a heuristic for this, we can minimize the sum of the variables u_i and v_i, by solving the LP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T u + \mathbf{1}^T v \\ \text{subject to} & a^T x_i - b \geq 1 - u_i, \quad i = 1, \ldots, N \\ & a^T y_i - b \leq -(1 - v_i), \quad i = 1, \ldots, M \\ & u \succeq 0, \quad v \succeq 0. \end{array} \tag{8.25}$$

Figure 8.10 shows an example. In this example, the affine function $a^T z + b$ misclassifies 1 out of 100 points. Note however that when $0 < u_i < 1$, the point x_i is correctly classified by the affine function $a^T z + b$, but violates the inequality $a^T x_i - b \geq 1$, and similarly for y_i. The objective function in the LP (8.25) can be interpreted as a relaxation of the number of points x_i that violate $a^T x_i - b \geq 1$ plus the number of points y_i that violate $a^T y_i - b \leq -1$. In other words, it is a relaxation of the number of points misclassified by the function $a^T z - b$, plus the number of points that are correctly classified but lie in the slab defined by $-1 < a^T z - b < 1$.

More generally, we can consider the trade-off between the number of misclassified points, and the width of the slab $\{z \mid -1 \leq a^T z - b \leq 1\}$, which is given by $2/\|a\|_2$. The standard *support vector classifier* for the sets $\{x_1, \ldots, x_N\}$,

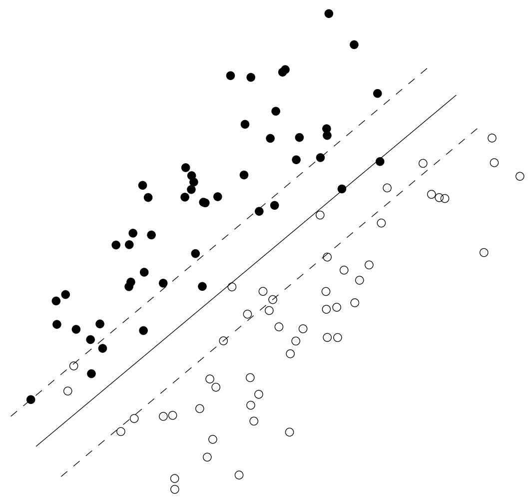


Figure 8.11 Approximate linear discrimination via support vector classifier, with $\gamma = 0.1$. The support vector classifier, shown as the solid line, misclassifies three points. Fifteen points are correctly classified but lie in the slab defined by $-1 < a^T z - b < 1$, bounded by the dashed lines.

$\{y_1, \ldots, y_M\}$ is defined as the solution of

$$\begin{array}{ll} \text{minimize} & \|a\|_2 + \gamma(\mathbf{1}^T u + \mathbf{1}^T v) \\ \text{subject to} & a^T x_i - b \geq 1 - u_i, \quad i = 1, \ldots, N \\ & a^T y_i - b \leq -(1 - v_i), \quad i = 1, \ldots, M \\ & u \succeq 0, \quad v \succeq 0, \end{array}$$

The first term is the proportional to the inverse of the width of the slab defined by $-1 \leq a^T z - b \leq 1$. The second term has the same interpretation as above, *i.e.*, it is a convex relaxation for the number of misclassified points (including the points in the slab). The parameter γ, which is positive, gives the relative weight of the number of misclassified points (which we want to minimize), compared to the width of the slab (which we want to maximize). Figure 8.11 shows an example.

Approximate linear discrimination via logistic modeling

Another approach to finding an affine function that approximately classifies two sets of points that cannot be linearly separated is based on the logistic model described in §7.1.1. We start by fitting the two sets of points with a logistic model. Suppose z is a random variable with values 0 or 1, with a distribution that depends on some (deterministic) explanatory variable $u \in \mathbf{R}^n$, via a logistic model of the form

$$\begin{array}{l} \mathbf{prob}(z = 1) = (\exp(a^T u - b))/(1 + \exp(a^T u - b)) \\ \mathbf{prob}(z = 0) = 1/(1 + \exp(a^T u - b)). \end{array} \qquad (8.26)$$

Now we assume that the given sets of points, $\{x_1, \ldots, x_N\}$ and $\{y_1, \ldots, y_M\}$, arise as samples from the logistic model. Specifically, $\{x_1, \ldots, x_N\}$ are the values of u for the N samples for which $z = 1$, and $\{y_1, \dots, y_M\}$ are the values of u for the M samples for which $z = 0$. (This allows us to have $x_i = y_j$, which would rule out discrimination between the two sets. In a logistic model, it simply means that we have two samples, with the same value of explanatory variable but different outcomes.)

We can determine a and b by maximum likelihood estimation from the observed samples, by solving the convex optimization problem

$$\text{minimize} \quad -l(a, b) \tag{8.27}$$

with variables a, b, where l is the log-likelihood function

$$\begin{aligned} l(a, b) = &\textstyle\sum_{i=1}^N (a^T x_i - b) \\ &- \textstyle\sum_{i=1}^N \log(1 + \exp(a^T x_i - b)) - \sum_{i=1}^M \log(1 + \exp(a^T y_i - b)) \end{aligned}$$

(see §7.1.1). If the two sets of points can be linearly separated, *i.e.*, if there exist a, b with $a^T x_i > b$ and $a^T y_i < b$, then the optimization problem (8.27) is unbounded below.

Once we find the maximum likelihood values of a and b, we can form a linear classifier $f(x) = a^T x - b$ for the two sets of points. This classifier has the following property: Assuming the data points are in fact generated from a logistic model with parameters a and b, it has the smallest probability of misclassification, over all linear classifiers. The hyperplane $a^T u = b$ corresponds to the points where $\mathbf{prob}(z = 1) = 1/2$, *i.e.*, the two outcomes are equally likely. An example is shown in figure 8.12.

Remark 8.1 *Bayesian interpretation.* Let x and z be two random variables, taking values in $\mathbf{R}^n$ and in $\{0, 1\}$, respectively. We assume that

$$\mathbf{prob}(z = 1) = \mathbf{prob}(z = 0) = 1/2,$$

and we denote by $p_0(x)$ and $p_1(x)$ the conditional probability densities of x, given $z = 0$ and given $z = 1$, respectively. We assume that p_0 and p_1 satisfy

$$\frac{p_1(x)}{p_0(x)} = e^{a^T x - b}$$

for some a and b. Many common distributions satisfy this property. For example, p_0 and p_1 could be two normal densities on $\mathbf{R}^n$ with equal covariance matrices and different means, or they could be two exponential densities on $\mathbf{R}^n_+$.

It follows from Bayes' rule that

$$\begin{aligned} \mathbf{prob}(z = 1 \mid x = u) &= \frac{p_1(u)}{p_1(u) + p_0(u)} \\ \mathbf{prob}(z = 0 \mid x = u) &= \frac{p_0(u)}{p_1(u) + p_0(u)}, \end{aligned}$$

from which we obtain

$$\begin{aligned} \mathbf{prob}(z = 1 \mid x = u) &= \frac{\exp(a^T u - b)}{1 + \exp(a^T u - b)} \\ \mathbf{prob}(z = 0 \mid x = u) &= \frac{1}{1 + \exp(a^T u - b)}. \end{aligned}$$

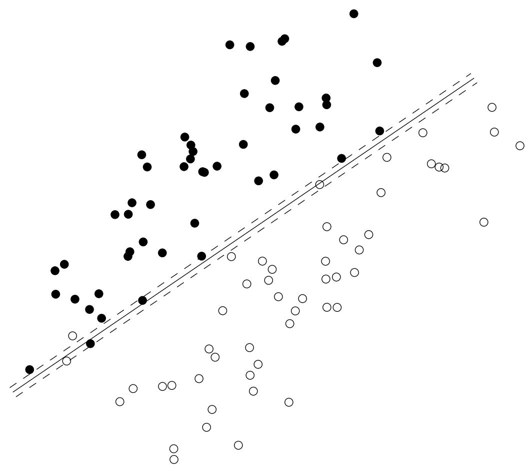


Figure 8.12 Approximate linear discrimination via logistic modeling. The points $x_1, \ldots, x_{50}$, shown as circles, cannot be linearly separated from the points $y_1, \ldots, y_{50}$, shown as squares. The maximum likelihood logistic model yields the hyperplane shown as a dark line, which misclassifies only two points. The two dashed lines show the lines defined by $a^T u - b = \pm 1$, where the probability of each outcome, according to the logistic model, is 73%. Three points are correctly classified, but lie in between the solid lines.

The logistic model (8.26) can therefore be interpreted as the posterior distribution of z, given that $x = u$.

8.6.2 Nonlinear discrimination

We can just as well seek a nonlinear function f, from a given subspace of functions, that is positive on one set and negative on another:

$$f(x_i) > 0, \quad i = 1, \ldots, N, \qquad f(y_i) < 0, \quad i = 1, \ldots, M.$$

Provided f is linear (or affine) in the parameters that define it, these inequalities can be solved in exactly the same way as in linear discrimination. In this section we examine some interesting special cases.

Quadratic discrimination

Suppose we take f to be quadratic: $f(x) = x^T P x + q^T x + r$. The parameters $P \in \mathbf{S}^n$, $q \in \mathbf{R}^n$, $r \in \mathbf{R}$ must satisfy the inequalities

$$\begin{array}{ll} x_i^T P x_i + q^T x_i + r > 0, & i = 1, \ldots, N \\ y_i^T P y_i + q^T y_i + r < 0, & i = 1, \ldots, M, \end{array}$$

which is a set of strict linear inequalities in the variables P, q, r. As in linear discrimination, we note that f is homogeneous in P, q, and r, so we can find a solution to the strict inequalities by solving the nonstrict feasibility problem

$$\begin{array}{ll} x_i^T P x_i + q^T x_i + r \geq 1, & i = 1, \ldots, N \\ y_i^T P y_i + q^T y_i + r \leq -1, & i = 1, \ldots, M. \end{array}$$

The separating surface $\{z \mid z^T P z + q^T z + r = 0\}$ is a quadratic surface, and the two classification regions

$$\{z \mid z^T P z + q^T z + r \leq 0\}, \qquad \{z \mid z^T P z + q^T z + r \geq 0\},$$

are defined by quadratic inequalities. Solving the quadratic discrimination problem, then, is the same as determining whether the two sets of points can be separated by a quadratic surface.

We can impose conditions on the shape of the separating surface or classification regions by adding constraints on P, q, and r. For example, we can require that $P \prec 0$, which means the separating surface is ellipsoidal. More specifically, it means that we seek an ellipsoid that contains all the points $x_1, \ldots, x_N$, but none of the points $y_1, \ldots, y_M$. This quadratic discrimination problem can be solved as an SDP feasibility problem

$$\begin{array}{lll} \text{find} & P,\ q,\ r & \\ \text{subject to} & x_i^T P x_i + q^T x_i + r \geq 1, & i = 1, \ldots, N \\ & y_i^T P y_i + q^T y_i + r \leq -1, & i = 1, \ldots, M \\ & P \preceq -I, & \end{array}$$

with variables $P \in \mathbf{S}^n$, $q \in \mathbf{R}^n$, and $r \in \mathbf{R}$. (Here we use homogeneity in P, q, r to express the constraint $P \prec 0$ as $P \preceq -I$.) Figure 8.13 shows an example.

Polynomial discrimination

We consider the set of polynomials on $\mathbf{R}^n$ with degree less than or equal to d:

$$f(x) = \sum_{i_1 + \cdots + i_n \leq d} a_{i_1 \cdots i_d} x_1^{i_1} \cdots x_n^{i_n}.$$

We can determine whether or not two sets $\{x_1, \ldots, x_N\}$ and $\{y_1, \ldots, y_M\}$ can be separated by such a polynomial by solving a set of linear inequalities in the variables $a_{i_1 \cdots i_d}$. Geometrically, we are checking whether the two sets can be separated by an algebraic surface (defined by a polynomial of degree less than or equal to d).

As an extension, the problem of determining the minimum degree polynomial on $\mathbf{R}^n$ that separates two sets of points can be solved via quasiconvex programming, since the degree of a polynomial is a quasiconvex function of the coefficients. This can be carried out by bisection on d, solving a feasibility linear program at each step. An example is shown in figure 8.14.

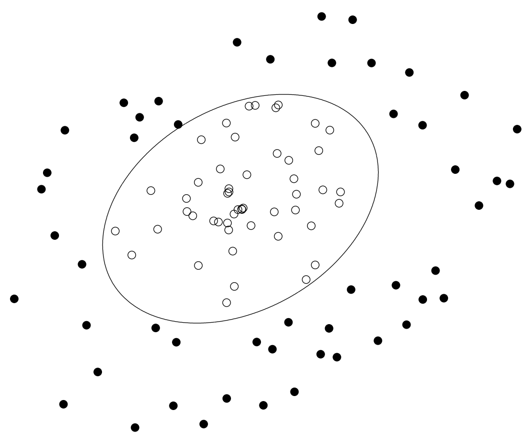


Figure 8.13 Quadratic discrimination, with the condition that $P \prec 0$. This means that we seek an ellipsoid containing all of x_i (shown as open circles) and none of the y_i (shown as filled circles). This can be solved as an SDP feasibility problem.

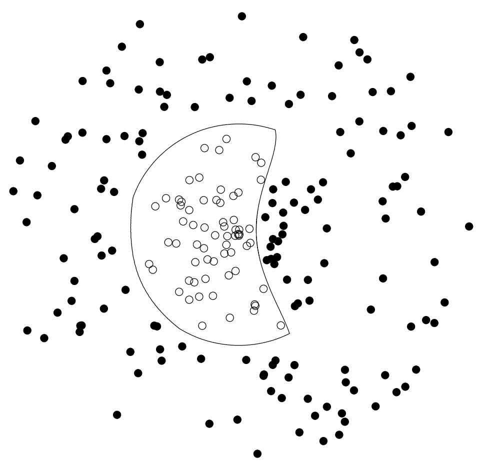


Figure 8.14 Minimum degree polynomial discrimination in $\mathbf{R}^2$. In this example, there exists no cubic polynomial that separates the points $x_1, \ldots x_N$ (shown as open circles) from the points $y_1, \ldots y_M$ (shown as filled circles), but they can be separated by fourth-degree polynomial, the zero level set of which is shown.

8.7 Placement and location

In this section we discuss a few variations on the following problem. We have N points in $\mathbf{R}^2$ or $\mathbf{R}^3$, and a list of pairs of points that must be connected by links. The positions of some of the N points are fixed; our task is to determine the positions of the remaining points, *i.e.*, to *place* the remaining points. The objective is to place the points so that some measure of the total interconnection length of the links is minimized, subject to some additional constraints on the positions. As an example application, we can think of the points as locations of plants or warehouses of a company, and the links as the routes over which goods must be shipped. The goal is to find locations that minimize the total transportation cost. In another application, the points represent the position of modules or cells on an integrated circuit, and the links represent wires that connect pairs of cells. Here the goal might be to place the cells in such a way that the total length of wire used to interconnect the cells is minimized.

The problem can be described in terms of an undirected graph with N nodes, representing the N points. With each node we associate a variable $x_i \in \mathbf{R}^k$, where $k = 2$ or $k = 3$, which represents its location or position. The problem is to minimize

$$\sum_{(i,j)\in\mathcal{A}} f_{ij}(x_i, x_j)$$

where $\mathcal{A}$ is the set of all links in the graph, and $f_{ij} : \mathbf{R}^k \times \mathbf{R}^k \to \mathbf{R}$ is a cost function associated with arc (i, j). (Alternatively, we can sum over all i and j, or over $i < j$, and simply set $f_{ij} = 0$ when links i and j are not connected.) Some of the coordinate vectors x_i are given. The optimization variables are the remaining coordinates. Provided the functions f_{ij} are convex, this is a convex optimization problem.

8.7.1 Linear facility location problems

In the simplest version of the problem the cost associated with arc (i, j) is the distance between nodes i and j: $f_{ij}(x_i, x_j) = \|x_i - x_j\|$, *i.e.*, we minimize

$$\sum_{(i,j)\in\mathcal{A}} \|x_i - x_j\|.$$

We can use any norm, but the most common applications involve the Euclidean norm or the ℓ_1-norm. For example, in circuit design it is common to route the wires between cells along piecewise-linear paths, with each segment either horizontal or vertical. (This is called *Manhattan routing*, since paths along the streets in a city with a rectangular grid are also piecewise-linear, with each street aligned with one of two orthogonal axes.) In this case, the length of wire required to connect cell i and cell j is given by $\|x_i - x_j\|_1$.

We can include nonnegative weights that reflect differences in the cost per unit distance along different arcs:

$$\sum_{(i,j)\in\mathcal{A}} w_{ij}\|x_i - x_j\|.$$

By assigning a weight $w_{ij} = 0$ to pairs of nodes that are not connected, we can express this problem more simply using the objective

$$\sum_{i<j} w_{ij}\|x_i - x_j\|. \tag{8.28}$$

This placement problem is convex.

Example 8.4 *One free point.* Consider the case where only one point $(u, v) \in \mathbf{R}^2$ is free, and we minimize the sum of the distances to fixed points $(u_1, v_1), \dots, (u_K, v_K)$.

- *ℓ_1-norm.* We can find a point that minimizes

$$\sum_{i=1}^{K} (|u - u_i| + |v - v_i|)$$

  analytically. An optimal point is any *median* of the fixed points. In other words, u can be taken to be any median of the points $\{u_1, \dots, u_K\}$, and v can be taken to be any median of the points $\{v_1, \dots, v_K\}$. (If K is odd, the minimizer is unique; if K is even, there can be a rectangle of optimal points.)

- *Euclidean norm.* The point (u, v) that minimizes the sum of the Euclidean distances,

$$\sum_{i=1}^{K} \left((u - u_i)^2 + (v - v_i)^2\right)^{1/2},$$

  is called the *Weber point* of the given fixed points.

8.7.2 Placement constraints

We now list some interesting constraints that can be added to the basic placement problem, preserving convexity. We can require some positions x_i to lie in a specified convex set, *e.g.*, a particular line, interval, square, or ellipsoid. We can constrain the relative position of one point with respect to one or more other points, for example, by limiting the distance between a pair of points. We can impose relative position constraints, *e.g.*, that one point must lie to the left of another point.

The *bounding box* of a group of points is the smallest rectangle that contains the points. We can impose a constraint that limits the points $x_1, \dots, x_p$ (say) to lie in a bounding box with perimeter not exceeding $P_{\max}$, by adding the constraints

$$u \preceq x_i \preceq v, \quad i = 1, \dots, p, \qquad 2\mathbf{1}^T(v - u) \le P_{\max},$$

where u, v are additional variables.

8.7.3 Nonlinear facility location problems

More generally, we can associate a cost with each arc that is a nonlinear increasing function of the length, *i.e.*,

$$\text{minimize} \quad \textstyle\sum_{i<j} w_{ij} h(\|x_i - x_j\|)$$

where h is an increasing (on $\mathbf{R}_+$) and convex function, and $w_{ij} \geq 0$. We call this a *nonlinear placement* or *nonlinear facility location problem.*

One common example uses the Euclidean norm, and the function $h(z) = z^2$, *i.e.*, we minimize

$$\sum_{i<j} w_{ij} \|x_i - x_j\|_2^2.$$

This is called a *quadratic placement problem.* The quadratic placement problem can be solved analytically when the only constraints are linear equalities; it can be solved as a QP if the constraints are linear equalities and inequalities.

Example 8.5 *One free point.* Consider the case where only one point x is free, and we minimize the sum of the squares of the Euclidean distances to fixed points $x_1, \ldots, x_K$,

$$\|x - x_1\|_2^2 + \|x - x_2\|_2^2 + \cdots + \|x - x_K\|_2^2.$$

Taking derivatives, we see that the optimal x is given by

$$\frac{1}{K}(x_1 + x_2 + \cdots + x_K),$$

i.e., the average of the fixed points.

Some other interesting possibilities are the 'deadzone' function h with deadzone width 2γ, defined as

$$h(z) = \begin{cases} 0 & |z| \leq \gamma \\ \gamma & |z| \geq \gamma, \end{cases}$$

and the 'quadratic-linear' function h, defined as

$$h(z) = \begin{cases} z^2 & |z| \leq \gamma \\ 2\gamma|z| - \gamma^2 & |z| \leq \gamma. \end{cases}$$

Example 8.6 We consider a placement problem in $\mathbf{R}^2$ with 6 free points, 8 fixed points, and 27 links. Figures 8.15–8.17 show the optimal solutions for the criteria

$$\sum_{(i,j)\in\mathcal{A}} \|x_i - x_j\|_2, \qquad \sum_{(i,j)\in\mathcal{A}} \|x_i - x_j\|_2^2, \qquad \sum_{(i,j)\in\mathcal{A}} \|x_i - x_j\|_2^4,$$

i.e., using the penalty functions $h(z) = z$, $h(z) = z^2$, and $h(z) = z^4$. The figures also show the resulting distributions of the link lengths.

Comparing the results, we see that the linear placement concentrates the free points in a small area, while the quadratic and fourth-order placements spread the points over larger areas. The linear placement includes many very short links, and a few very long ones (3 lengths under 0.2 and 2 lengths above 1.5.). The quadratic penalty function

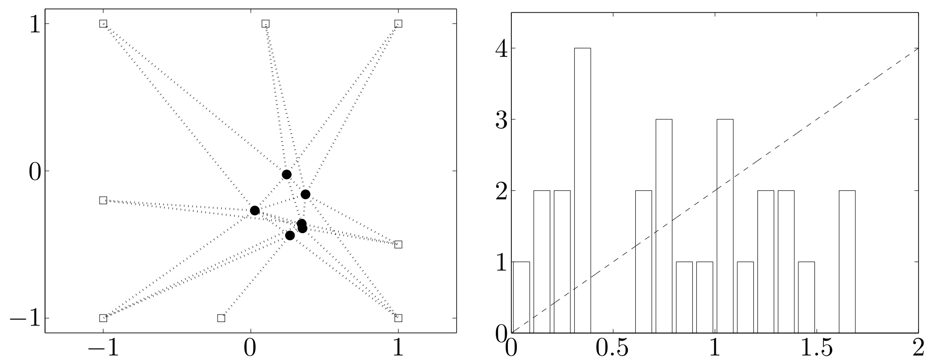


Figure 8.15 *Linear placement.* Placement problem with 6 free points (shown as dots), 8 fixed points (shown as squares), and 27 links. The coordinates of the free points minimize the sum of the Euclidean lengths of the links. The right plot is the distribution of the 27 link lengths. The dashed curve is the (scaled) penalty function $h(z) = z$.

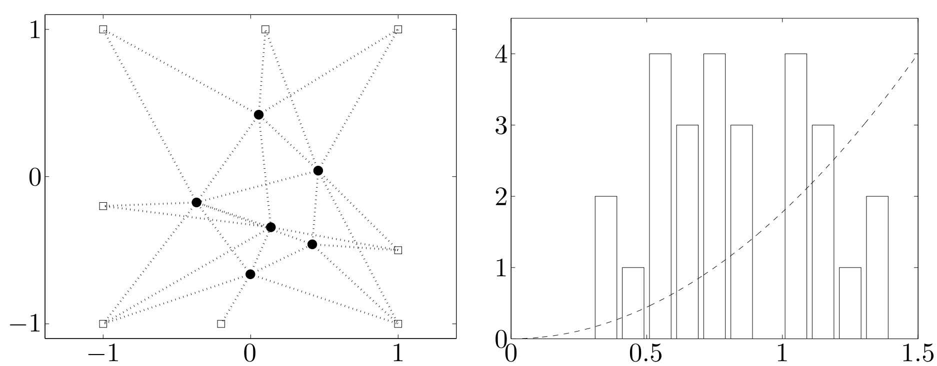


Figure 8.16 *Quadratic placement.* Placement that minimizes the sum of squares of the Euclidean lengths of the links, for the same data as in figure 8.15. The dashed curve is the (scaled) penalty function $h(z) = z^2$.

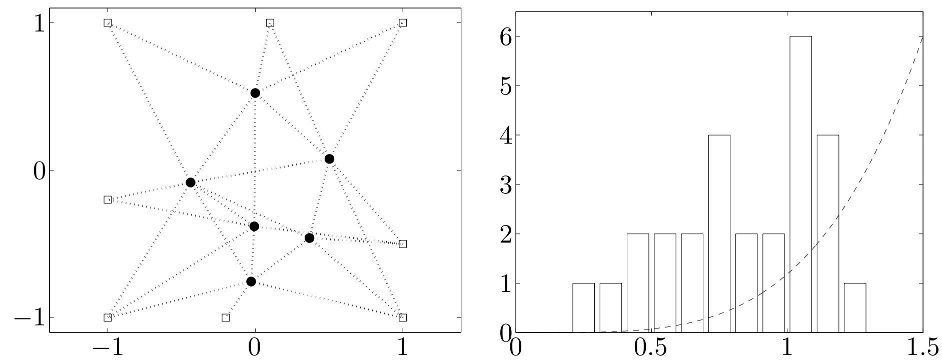


Figure 8.17 *Fourth-order placement.* Placement that minimizes the sum of the fourth powers of the Euclidean lengths of the links. The dashed curve is the (scaled) penalty function $h(z) = z^4$.

puts a higher penalty on long lengths relative to short lengths, and for lengths under 0.1, the penalty is almost negligible. As a result, the maximum length is shorter (less than 1.4), but we also have fewer short links. The fourth-order function puts an even higher penalty on long lengths, and has a wider interval (between zero and about 0.4) where it is negligible. As a result, the maximum length is shorter than for the quadratic placement, but we also have more lengths close to the maximum.

8.7.4 Location problems with path constraints

Path constraints

A p-link *path* along the points $x_1, \ldots, x_N$ is described by a sequence of nodes, $i_0, \ldots, i_p \in \{1, \ldots, N\}$. The length of the path is given by

$$\|x_{i_1} - x_{i_0}\| + \|x_{i_2} - x_{i_1}\| + \cdots + \|x_{i_p} - x_{i_{p-1}}\|,$$

which is a convex function of $x_1, \ldots, x_N$, so imposing an upper bound on the length of a path is a convex constraint. Several interesting placement problems involve path constraints, or have an objective based on path lengths. We describe one typical example, in which the objective is based on a maximum path length over a set of paths.

Minimax delay placement

We consider a directed acyclic graph with nodes $1, \ldots, N$, and arcs or links represented by a set $\mathcal{A}$ of ordered pairs: $(i, j) \in \mathcal{A}$ if and only if an arc points from i to j. We say node i is a *source node* if no arc $\mathcal{A}$ points to it; it is a *sink node* or *destination node* if no arc in $\mathcal{A}$ leaves from it. We will be interested in the maximal paths in the graph, which begin at a source node and end at a sink node.

The arcs of the graph are meant to model some kind of flow, say of goods or information, in a network with nodes at positions $x_1, \ldots, x_N$. The flow starts at a source node, then moves along a path from node to node, ending at a sink or destination node. We use the distance between successive nodes to model propagation time, or shipment time, of the goods between nodes; the total delay or propagation time of a path is (proportional to) the sum of the distances between successive nodes.

Now we can describe the minimax delay placement problem. Some of the node locations are fixed, and the others are free, *i.e.*, optimization variables. The goal is to choose the free node locations in order to minimize the maximum total delay, for any path from a source node to a sink node. Evidently this is a convex problem, since the objective

$$T_{\max} = \max\{\|x_{i_1} - x_{i_0}\| + \cdots + \|x_{i_p} - x_{i_{p-1}}\| \mid i_0, \ldots, i_p \text{ is a source-sink path}\} \tag{8.29}$$

is a convex function of the locations $x_1, \ldots, x_N$.

While the problem of minimizing (8.29) is convex, the number of source-sink paths can be very large, exponential in the number of nodes or arcs. There is a useful reformulation of the problem, which avoids enumerating all sink-source paths.

We first explain how we can evaluate the maximum delay $T_{\max}$ far more efficiently than by evaluating the delay for every source-sink path, and taking the maximum. Let τ_k be the maximum total delay of any path from node k to a sink node. Clearly we have $\tau_k = 0$ when k is a sink node. Consider a node k, which has outgoing arcs to nodes $j_1, \ldots, j_p$. For a path starting at node k and ending at a sink node, its first arc must lead to one of the nodes $j_1, \ldots, j_p$. If such a path first takes the arc leading to j_i, and then takes the longest path from there to a sink node, the total length is

$$\|x_{j_i} - x_k\| + \tau_{j_i},$$

i.e., the length of the arc to j_i, plus the total length of the longest path from j_i to a sink node. It follows that the maximum delay of a path starting at node k and leading to a sink node satisfies

$$\tau_k = \max\{\|x_{j_1} - x_k\| + \tau_{j_1}, \ldots, \|x_{j_p} - x_k\| + \tau_{j_p}\}. \tag{8.30}$$

(This is a simple *dynamic programming* argument.)

The equations (8.30) give a recursion for finding the maximum delay from any node: we start at the sink nodes (which have maximum delay zero), and then work backward using the equations (8.30), until we reach all source nodes. The maximum delay over any such path is then the maximum of all the τ_k, which will occur at one of the source nodes. This dynamic programming recursion shows how the maximum delay along any source-sink path can be computed recursively, without enumerating all the paths. The number of arithmetic operations required for this recursion is approximately the number of links.

Now we show how the recursion based on (8.30) can be used to formulate the minimax delay placement problem. We can express the problem as

$$\begin{array}{ll} \text{minimize} & \max\{\tau_k \mid k \text{ a source node}\} \\ \text{subject to} & \tau_k = 0, \quad k \text{ a sink node} \\ & \tau_k = \max\{\|x_j - x_k\| + \tau_j \mid \text{there is an arc from } k \text{ to } j\}, \end{array}$$

with variables $\tau_1, \ldots, \tau_N$ and the free positions. This problem is not convex, but we can express it in an equivalent form that is convex, by replacing the equality constraints with inequalities. We introduce new variables $T_1, \ldots, T_N$, which will be upper bounds on $\tau_1, \ldots, \tau_N$, respectively. We will take $T_k = 0$ for all sink nodes, and in place of (8.30) we take the inequalities

$$T_k \geq \max\{\|x_{j_1} - x_k\| + T_{j_1}, \ldots, \|x_{j_p} - x_k\| + T_{j_p}\}.$$

If these inequalities are satisfied, then $T_k \geq \tau_k$. Now we form the problem

$$\begin{array}{ll} \text{minimize} & \max\{T_k \mid k \text{ a source node}\} \\ \text{subject to} & T_k = 0, \quad k \text{ a sink node} \\ & T_k \geq \max\{\|x_j - x_k\| + T_j \mid \text{there is an arc from } k \text{ to } j\}. \end{array}$$

This problem, with variables $T_1, \ldots, T_N$ and the free locations, is convex, and solves the minimax delay location problem.

8.8 Floor planning

In placement problems, the variables represent the coordinates of a number of points that are to be optimally placed. A *floor planning problem* can be considered an extension of a placement problem in two ways:

- The objects to be placed are rectangles or boxes aligned with the axes (as opposed to points), and must not overlap.
- Each rectangle or box to be placed can be reconfigured, within some limits. For example we might fix the area of each rectangle, but not the length and height separately.

The objective is usually to minimize the size (*e.g.*, area, volume, perimeter) of the *bounding box*, which is the smallest box that contains the boxes to be configured and placed.

The non-overlap constraints make the general floor planning problem a complicated combinatorial optimization problem or rectangle packing problem. However, if the *relative positioning* of the boxes is specified, several types of floor planning problems can be formulated as convex optimization problems. We explore some of these in this section. We consider the two-dimensional case, and make a few comments on extensions to higher dimensions (when they are not obvious).

We have N cells or modules $C_1, \ldots, C_N$ that are to be configured and placed in a rectangle with width W and height H, and lower left corner at the position $(0, 0)$. The geometry and position of the ith cell is specified by its width w_i and height h_i, and the coordinates (x_i, y_i) of its lower left corner. This is illustrated in figure 8.18.

The variables in the problem are x_i, y_i, w_i, h_i for $i = 1, \ldots, N$, and the width W and height H of the bounding rectangle. In all floor planning problems, we require that the cells lie inside the bounding rectangle, *i.e.*,

$$x_i \geq 0, \qquad y_i \geq 0, \qquad x_i + w_i \leq W, \qquad y_i + h_i \leq H, \qquad i = 1, \ldots, N. \qquad (8.31)$$

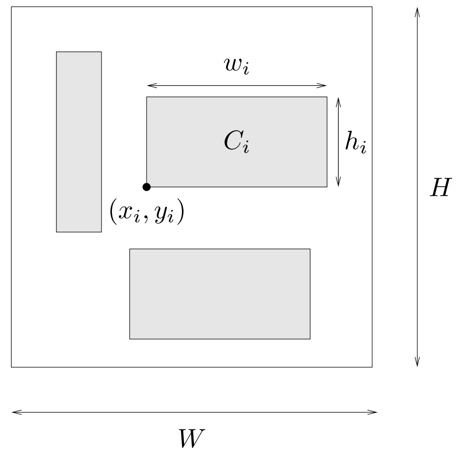


Figure 8.18 Floor planning problem. Non-overlapping rectangular cells are placed in a rectangle with width W, height H, and lower left corner at $(0, 0)$. The ith cell is specified by its width w_i, height h_i, and the coordinates of its lower left corner, (x_i, y_i).

We also require that the cells do not overlap, except possibly on their boundaries:

$$\mathbf{int}\,(C_i \cap C_j) = \emptyset \quad \text{for } i \neq j.$$

(It is also possible to require a positive minimum clearance between the cells.) The non-overlap constraint $\mathbf{int}(C_i \cap C_j) = \emptyset$ holds if and only if for $i \neq j$,

$$C_i \text{ is left of } C_j, \text{ or } C_i \text{ is right of } C_j, \text{ or } C_i \text{ is below } C_j, \text{ or } C_i \text{ is above } C_j.$$

These four geometric conditions correspond to the inequalities

$$x_i + w_i \leq x_j, \text{ or } x_j + w_j \leq x_i, \text{ or } y_i + h_j \leq y_j, \text{ or } y_j + h_i \leq y_i, \qquad (8.32)$$

at least one of which must hold for each $i \neq j$. Note the combinatorial nature of these constraints: for each pair $i \neq j$, at least one of the four inequalities above must hold.

8.8.1 Relative positioning constraints

The idea of relative positioning constraints is to specify, for each pair of cells, one of the four possible relative positioning conditions, *i.e.*, left, right, above, or below. One simple method to specify these constraints is to give two relations on $\{1, \ldots, N\}$: $\mathcal{L}$ (meaning 'left of') and $\mathcal{B}$ (meaning 'below'). We then impose the constraint that C_i is to the left of C_j if $(i, j) \in \mathcal{L}$, and C_i is below C_j if $(i, j) \in \mathcal{B}$. This yields the constraints

$$x_i + w_i \leq x_j \text{ for } (i, j) \in \mathcal{L}, \qquad y_i + h_i \leq y_j \text{ for } (i, j) \in \mathcal{B}, \qquad (8.33)$$

for $i, j = 1, \dots, N$. To ensure that the relations $\mathcal{L}$ and $\mathcal{B}$ specify the relative positioning of each pair of cells, we require that for each (i, j) with $i \neq j$, one of the following holds:

$$(i, j) \in \mathcal{L}, \qquad (j, i) \in \mathcal{L}, \qquad (i, j) \in \mathcal{B}, \qquad (j, i) \in \mathcal{B},$$

and that $(i, i) \notin \mathcal{L}$, $(i, i) \notin \mathcal{B}$. The inequalities (8.33) are a set of $N(N-1)/2$ linear inequalities in the variables. These inequalities imply the non-overlap inequalities (8.32), which are a set of $N(N-1)/2$ *disjunctions* of four linear inequalities.

We can assume that the relations $\mathcal{L}$ and $\mathcal{B}$ are anti-symmetric (*i.e.*, $(i, j) \in \mathcal{L} \Rightarrow (j, i) \notin \mathcal{L}$) and transitive (*i.e.*, $(i, j) \in \mathcal{L}$, $(j, k) \in \mathcal{L} \Rightarrow (i, k) \in \mathcal{L}$). (If this were not the case, the relative positioning constraints would clearly be infeasible.) Transitivity corresponds to the obvious condition that if cell C_i is to the left of cell C_j, which is to the left of cell C_k, then cell C_i must be to the left of cell C_k. In this case the inequality corresponding to $(i, k) \in \mathcal{L}$ is redundant; it is implied by the other two. By exploiting transitivity of the relations $\mathcal{L}$ and $\mathcal{B}$ we can remove redundant constraints, and obtain a compact set of relative positioning inequalities.

A minimal set of relative positioning constraints is conveniently described using two directed acyclic graphs $\mathcal{H}$ and $\mathcal{V}$ (for horizontal and vertical). Both graphs have N nodes, corresponding to the N cells in the floor planning problem. The graph $\mathcal{H}$ generates the relation $\mathcal{L}$ as follows: we have $(i, j) \in \mathcal{L}$ if and only if there is a (directed) path in $\mathcal{H}$ from i to j. Similarly, the graph $\mathcal{V}$ generates the relation $\mathcal{B}$: $(i, j) \in \mathcal{B}$ if and only if there is a (directed) path in $\mathcal{H}$ from i to j. To ensure that a relative positioning constraint is given for every pair of cells, we require that for every pair of cells, there is a directed path from one to the other in one of the graphs.

Evidently, we only need to impose the inequalities that correspond to the edges of the graphs $\mathcal{H}$ and $\mathcal{V}$; the others follow from transitivity. We arrive at the set of inequalities

$$x_i + w_i \leq x_j \text{ for } (i, j) \in \mathcal{H}, \qquad y_i + h_i \leq y_j \text{ for } (i, j) \in \mathcal{V}, \tag{8.34}$$

which is a set of linear inequalities, one for each edge in $\mathcal{H}$ and $\mathcal{V}$. The set of inequalities (8.34) is a subset of the set of inequalities (8.33), and equivalent.

In a similar way, the $4N$ inequalities (8.31) can be reduced to a minimal, equivalent set. The constraint $x_i \geq 0$ only needs to be imposed on the left-most cells, *i.e.*, for i that are minimal in the relation $\mathcal{L}$. These correspond to the sources in the graph $\mathcal{H}$, *i.e.*, those nodes that have no edges pointing to them. Similarly, the inequalities $x_i + w_i \leq W$ only need to be imposed for the right-most cells. In the same way the vertical bounding box inequalities can be pruned to a minimal set. This yields the minimal equivalent set of bounding box inequalities

$$\begin{array}{ll} x_i \geq 0 \text{ for } i \ \mathcal{L} \text{ minimal}, & x_i + w_i \leq W \text{ for } i \ \mathcal{L} \text{ maximal}, \\ y_i \geq 0 \text{ for } i \ \mathcal{B} \text{ minimal}, & y_i + h_i \leq H \text{ for } i \ \mathcal{B} \text{ maximal}. \end{array} \tag{8.35}$$

A simple example is shown in figure 8.19. In this example, the $\mathcal{L}$ minimal or left-most cells are C_1, C_2, and C_4, and the only right-most cell is C_5. The minimal set of inequalities specifying the horizontal relative positioning is given by

$$\begin{array}{c} x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_4 \geq 0, \qquad x_5 + w_5 \leq W, \qquad x_1 + w_1 \leq x_3, \\ x_2 + w_2 \leq x_3, \qquad x_3 + w_3 \leq x_5, \qquad x_4 + w_4 \leq x_5. \end{array}$$

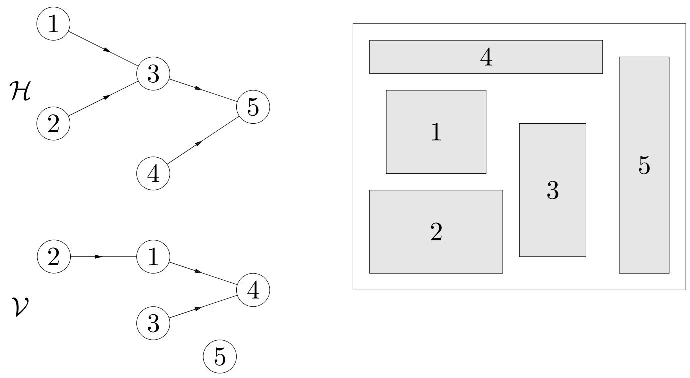


Figure 8.19 Example illustrating the horizontal and vertical graphs $\mathcal{H}$ and $\mathcal{V}$ that specify the relative positioning of the cells. If there is a path from node i to node j in $\mathcal{H}$, then cell i must be placed to the left of cell j. If there is a path from node i to node j in $\mathcal{V}$, then cell i must be placed below cell j. The floorplan shown at right satisfies the relative positioning specified by the two graphs.

The minimal set of inequalities specifying the vertical relative positioning is given by

$$y_2 \geq 0, \qquad y_3 \geq 0, \qquad y_5 \geq 0, \qquad y_4 + h_4 \leq H, \qquad y_5 + h_5 \leq H,$$
$$y_2 + h_2 \leq y_1, \qquad y_1 + h_1 \leq y_4, \qquad y_3 + h_3 \leq y_4.$$

8.8.2 Floor planning via convex optimization

In this formulation, the variables are the bounding box width and height W and H, and the cell widths, heights, and positions: w_i, h_i, x_i, and w_i, for $i = 1, \ldots, N$. We impose the bounding box constraints (8.35) and the relative positioning constraints (8.34), which are linear inequalities. As objective, we take the perimeter of the bounding box, *i.e.*, $2(W + H)$, which is a linear function of the variables. We now list some of the constraints that can be expressed as convex inequalities or linear equalities in the variables.

Minimum spacing

We can impose a minimum spacing $\rho > 0$ between cells by changing the relative position constraints from $x_i + w_i \leq x_j$ for $(i, j) \in \mathcal{H}$, to $x_i + w_i + \rho \leq x_j$ for $(i, j) \in \mathcal{H}$, and similarly for the vertical graph. We can have a different minimum spacing associated with each edge in $\mathcal{H}$ and $\mathcal{V}$. Another possibility is to fix W and H, and maximize the minimum spacing ρ as objective.

Minimum cell area

For each cell we specify a minimum area, *i.e.*, we require that $w_ih_i \geq A_i$, where $A_i > 0$. These minimum cell area constraints can be expressed as convex inequalities in several ways, *e.g.*, $w_i \geq A_i/h_i$, $(w_ih_i)^{1/2} \geq A_i^{1/2}$, or $\log w_i + \log h_i \geq \log A_i$.

Aspect ratio constraints

We can impose upper and lower bounds on the *aspect ratio* of each cell, *i.e.*,

$$l_i \leq h_i/w_i \leq u_i.$$

Multiplying through by w_i transforms these constraints into linear inequalities. We can also fix the aspect ratio of a cell, which results in a linear equality constraint.

Alignment constraints

We can impose the constraint that two edges, or a center line, of two cells are aligned. For example, the horizontal center line of cell i aligns with the top of cell j when

$$y_i + w_i/2 = y_j + w_j.$$

These are linear equality constraints. In a similar way we can require that a cell is flushed against the bounding box boundary.

Symmetry constraints

We can require pairs of cells to be symmetric about a vertical or horizontal axis, that can be fixed or floating (*i.e.*, whose position is fixed or not). For example, to specify that the pair of cells i and j are symmetric about the vertical axis $x = x_{\text{axis}}$, we impose the linear equality constraint

$$x_{\text{axis}} - (x_i + w_i/2) = x_j + w_j/2 - x_{\text{axis}}.$$

We can require that several pairs of cells be symmetric about an unspecified vertical axis by imposing these equality constraints, and introducing x_{axis} as a new variable.

Similarity constraints

We can require that cell i be an a-scaled translate of cell j by the equality constraints $w_i = aw_j$, $h_i = ah_j$. Here the scaling factor a must be fixed. By imposing only one of these constraints, we require that the width (or height) of one cell be a given factor times the width or height) of the other cell.

Containment constraints

We can require that a particular cell contains a given point, which imposes two linear inequalities. We can require that a particular cell lie inside a given polyhedron, again by imposing linear inequalities.

Distance constraints

We can impose a variety of constraints that limit the distance between pairs of cells. In the simplest case, we can limit the distance between the center points of cell i and j (or any other fixed points on the cells, such as lower left corners). For example, to limit the distance between the centers of cells i and j, we use the (convex) inequality

$$\|(x_i + w_i/2, y_i + h_i/2) - (x_j + w_j/2, y_j + h_j/2)\| \le D_{ij}.$$

As in placement problems, we can limit sums of distances, or use sums of distances as the objective.

We can also limit the distance $\mathbf{dist}(C_i, C_j)$ between cell i and cell j, *i.e.*, the minimum distance between a point in cell i and a point in cell j. In the general case this can be done as follows. To limit the distance between cells i and j in the norm $\|\cdot\|$, we can introduce four new variables u_i, v_i, u_j, v_j. The pair (u_i, v_i) will represent a point in C_i, and the pair (u_j, v_j) will represent a point in C_j. To ensure this we impose the linear inequalities

$$x_i \le u_i \le x_i + w_i, \qquad y_i \le v_i \le y_i + h_i,$$

and similarly for cell j. Finally, to limit $\mathbf{dist}(C_i, C_j)$, we add the convex inequality

$$\|(u_i, v_i) - (u_j, v_j)\| \le D_{ij}.$$

In many specific cases we can express these distance constraints more efficiently, by exploiting the relative positioning constraints or deriving a more explicit formulation. As an example consider the ℓ_∞-norm, and suppose cell i lies to the left of cell j (by a relative positioning constraint). The horizontal displacement between the two cells is $x_j - (x_i + w_i)$ Then we have $\mathbf{dist}(C_i, C_j) \le D_{ij}$ if and only if

$$x_j - (x_i + w_i) \le D_{ij}, \qquad y_j - (y_i + h_i) \le D_{ij}, \qquad y_i - (y_j + h_j) \le D_{ij}.$$

The first inequality states that the horizontal displacement between the right edge of cell i and the left edge of cell j does not exceed D_{ij}. The second inequality requires that the bottom of cell j is no more than D_{ij} above the top of cell i, and the third inequality requires that the bottom of cell i is no more than D_{ij} above the top of cell j. These three inequalities together are equivalent to $\mathbf{dist}(C_i, C_j) \le D_{ij}$. In this case, we do not need to introduce any new variables.

We can limit the ℓ_1- (or ℓ_2-) distance between two cells in a similar way. Here we introduce one new variable d_v, which will serve as a bound on the vertical displacement between the cells. To limit the ℓ_1-distance, we add the constraints

$$y_j - (y_i + h_i) \le d_v, \qquad y_i - (y_j + h_j) \le d_v, \qquad d_v \ge 0$$

and the constraints

$$x_j - (x_i + w_i) + d_v \le D_{ij}.$$

(The first term is the horizontal displacement and the second is an upper bound on the vertical displacement.) To limit the Euclidean distance between the cells, we replace this last constraint with

$$(x_j - (x_i + w_i)^2 + d_v^2) \le D_{ij}^2.$$

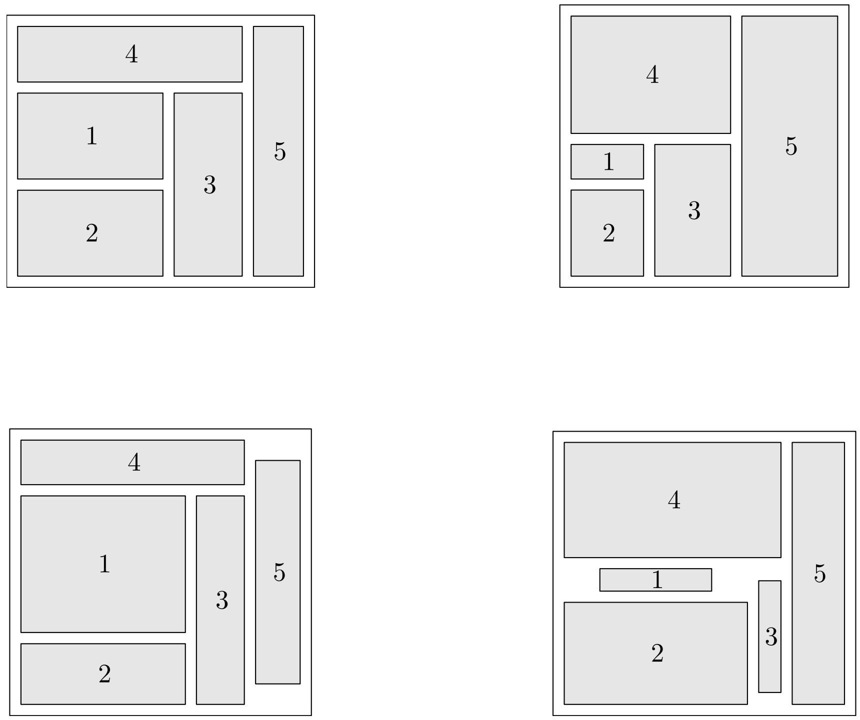


Figure 8.20 Four instances of an optimal floor plan, using the relative positioning constraints shown in figure 8.19. In each case the objective is to minimize the perimeter, and the same minimum spacing constraint between cells is imposed. We also require the aspect ratios to lie between 1/5 and 5. The four cases differ in the minimum areas required for each cell. The sum of the minimum areas is the same for each case.

Example 8.7 Figure 8.20 shows an example with 5 cells, using the ordering constraints of figure 8.19, and four different sets of constraints. In each case we impose the same minimum required spacing constraint, and the same aspect ratio constraint $1/5 \leq w_i/h_i \leq 5$. The four cases differ in the minimum required cell areas A_i. The values of A_i are chosen so that the total minimum required area $\sum_{i=1}^{5} A_i$ is the same for each case.

8.8.3 Floor planning via geometric programming

The floor planning problem can also be formulated as a *geometric program* in the variables x_i, y_i, w_i, h_i, W, H. The objectives and constraints that can be handled in this formulation are a bit different from those that can be expressed in the convex formulation.

First we note that the bounding box constraints (8.35) and the relative positioning constraints (8.34) are posynomial inequalities, since the lefthand sides are sums of variables, and the righthand sides are single variables, hence monomials. Dividing these inequalities by the righthand side yields standard posynomial inequalities.

In the geometric programming formulation we can minimize the bounding box area, since WH is a monomial, hence posynomial. We can also exactly specify the area of each cell, since $w_i h_i = A_i$ is a monomial equality constraint. On the other hand alignment, symmetry, and distance constraints cannot be handled in the geometric programming formulation. Similarity, however, can be; indeed it is possible to require that one cell be similar to another, without specifying the scaling ratio (which can be treated as just another variable).

Bibliography

The characterization of Euclidean distance matrices in §8.3.3 appears in Schoenberg [Sch35]; see also Gower [Gow85].

Our use of the term Löwner-John ellipsoid follows Grötschel, Lovász, and Schrijver [GLS88, page 69]. The efficiency results for ellipsoidal approximations in §8.4 were proved by John [Joh85]. Boyd, El Ghaoui, Feron, and Balakrishnan [BEFB94, §3.7] give convex formulations of several ellipsoidal approximation problems involving sets defined as unions, intersections or sums of ellipsoids.

The different centers defined in §8.5 have applications in design centering (see, for example, Seifi, Ponnambalan, and Vlach [SPV99]), and cutting-plane methods (Elzinga and Moore [EM75], Tarasov, Khachiyan, and Èrlikh [TKE88], and Ye [Ye97, chapter 8]). The inner ellipsoid defined by the Hessian of the logarithmic barrier function (page 420) is sometimes called the *Dikin ellipsoid*, and is the basis of Dikin's algorithm for linear and quadratic programming [Dik67]. The expression for the outer ellipsoid at the analytic center was given by Sonnevend [Son86]. For extensions to nonpolyhedral convex sets, see Boyd and El Ghaoui [BE93], Jarre [Jar94], and Nesterov and Nemirovski [NN94, page 34].

Convex optimization has been applied to linear and nonlinear discrimination problems since the 1960s; see Mangasarian [Man65] and Rosen [Ros65]. Standard texts that discuss pattern classification include Duda, Hart, and Stork [DHS99] and Hastie, Tibshirani, and Friedman [HTF01]. For a detailed discussion of support vector classifiers, see Vapnik [Vap00] or Schölkopf and Smola [SS01].

The Weber point defined in example 8.4 is named after Weber [Web71]. Linear and quadratic placement is used in circuit design (Kleinhaus, Sigl, Johannes, and Antreich [KSJA91, SDJ91]). Sherwani [She99] is a recent overview of algorithms for placement, layout, floor planning, and other geometric optimization problems in VLSI circuit design.

Exercises

Projection on a set

8.1 *Uniqueness of projection.* Show that if $C \subseteq \mathbf{R}^n$ is nonempty, closed and convex, and the norm $\|\cdot\|$ is strictly convex, then for every x_0 there is exactly one $x \in C$ closest to x_0. In other words the projection of x_0 on C is unique.

8.2 [Web94, Val64] *Chebyshev characterization of convexity.* A set $C \in \mathbf{R}^n$ is called a *Chebyshev set* if for every $x_0 \in \mathbf{R}^n$, there is a unique point in C closest (in Euclidean norm) to x_0. From the result in exercise 8.1, every nonempty, closed, convex set is a Chebyshev set. In this problem we show the converse, which is known as *Motzkin's theorem.*

Let $C \in \mathbf{R}^n$ be a Chebyshev set.

(a) Show that C is nonempty and closed.

(b) Show that P_C, the Euclidean projection on C, is continuous.

(c) Suppose $x_0 \notin C$. Show that $P_C(x) = P_C(x_0)$ for all $x = \theta x_0 + (1-\theta)P_C(x_0)$ with $0 \leq \theta \leq 1$.

(d) Suppose $x_0 \notin C$. Show that $P_C(x) = P_C(x_0)$ for all $x = \theta x_0 + (1-\theta)P_C(x_0)$ with $\theta \geq 1$.

(e) Combining parts (c) and (d), we can conclude that all points on the ray with base $P_C(x_0)$ and direction $x_0 - P_C(x_0)$ have projection $P_C(x_0)$. Show that this implies that C is convex.

8.3 *Euclidean projection on proper cones.*

(a) *Nonnegative orthant.* Show that Euclidean projection onto the nonnegative orthant is given by the expression on page 399.

(b) *Positive semidefinite cone.* Show that Euclidean projection onto the positive semidefinite cone is given by the expression on page 399.

(c) *Second-order cone.* Show that the Euclidean projection of (x_0, t_0) on the second-order cone

$$K = \{(x,t) \in \mathbf{R}^{n+1} \mid \|x\|_2 \leq t\}$$

is given by

$$P_K(x_0, t_0) = \begin{cases} 0 & \|x_0\|_2 \leq -t_0 \\ (x_0, t_0) & \|x_0\|_2 \leq t_0 \\ (1/2)(1 + t_0/\|x_0\|_2)(x_0, \|x_0\|_2) & \|x_0\|_2 \geq |t_0|. \end{cases}$$

8.4 The Euclidean projection of a point on a convex set yields a simple separating hyperplane

$$(P_C(x_0) - x_0)^T (x - (1/2)(x_0 + P_C(x_0))) = 0.$$

Find a counterexample that shows that this construction does not work for general norms.

8.5 [HUL93, volume 1, page 154] *Depth function and signed distance to boundary.* Let $C \subseteq \mathbf{R}^n$ be a nonempty convex set, and let $\mathbf{dist}(x, C)$ be the distance of x to C in some norm. We already know that $\mathbf{dist}(x, C)$ is a convex function of x.

(a) Show that the depth function,

$$\mathbf{depth}(x, C) = \mathbf{dist}(x, \mathbf{R}^n \setminus C),$$

is concave for $x \in C$.

(b) The *signed distance* to the boundary of C is defined as

$$s(x) = \begin{cases} \mathbf{dist}(x, C) & x \notin C \\ -\mathbf{depth}(x, C) & x \in C. \end{cases}$$

Thus, $s(x)$ is positive outside C, zero on its boundary, and negative on its interior. Show that s is a convex function.

Distance between sets

8.6 Let C, D be convex sets.

(a) Show that $\mathbf{dist}(C, x + D)$ is a convex function of x.

(b) Show that $\mathbf{dist}(tC, x + tD)$ is a convex function of (x, t) for $t > 0$.

8.7 *Separation of ellipsoids.* Let $\mathcal{E}_1$ and $\mathcal{E}_2$ be two ellipsoids defined as

$$\mathcal{E}_1 = \{x \mid (x - x_1)^T P_1^{-1}(x - x_1) \leq 1\}, \qquad \mathcal{E}_2 = \{x \mid (x - x_2)^T P_2^{-1}(x - x_2) \leq 1\},$$

where $P_1,\ P_2 \in \mathbf{S}^n_{++}$. Show that $\mathcal{E}_1 \cap \mathcal{E}_2 = \emptyset$ if and only if there exists an $a \in \mathbf{R}^n$ with

$$\|P_2^{1/2}a\|_2 + \|P_1^{1/2}a\|_2 < a^T(x_1 - x_2).$$

8.8 *Intersection and containment of polyhedra.* Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two polyhedra defined as

$$\mathcal{P}_1 = \{x \mid Ax \preceq b\}, \qquad \mathcal{P}_2 = \{x \mid Fx \preceq g\},$$

with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, $F \in \mathbf{R}^{p\times n}$, $g \in \mathbf{R}^p$. Formulate each of the following problems as an LP feasibility problem, or a set of LP feasibility problems.

(a) Find a point in the intersection $\mathcal{P}_1 \cap \mathcal{P}_2$.

(b) Determine whether $\mathcal{P}_1 \subseteq P_2$.

For each problem, derive a set of linear inequalities and equalities that forms a strong alternative, and give a geometric interpretation of the alternative.

Repeat the question for two polyhedra defined as

$$\mathcal{P}_1 = \mathbf{conv}\{v_1, \ldots, v_K\}, \qquad \mathcal{P}_2 = \mathbf{conv}\{w_1, \ldots, w_L\}.$$

Euclidean distance and angle problems

8.9 *Closest Euclidean distance matrix to given data.* We are given data $\hat{d}_{ij}$, for $i, j = 1, \ldots, n$, which are corrupted measurements of the Euclidean distances between vectors in $\mathbf{R}^k$:

$$\hat{d}_{ij} = \|x_i - x_j\|_2 + v_{ij}, \quad i, j = 1, \ldots, n,$$

where v_{ij} is some noise or error. These data satisfy $\hat{d}_{ij} \geq 0$ and $\hat{d}_{ij} = \hat{d}_{ji}$, for all i, j. The dimension k is not specified.

Show how to solve the following problem using convex optimization. Find a dimension k and $x_1, \ldots, x_n \in \mathbf{R}^k$ so that $\sum_{i,j=1}^n (d_{ij} - \hat{d}_{ij})^2$ is minimized, where $d_{ij} = \|x_i - x_j\|_2$, $i, j = 1, \ldots, n$. In other words, given some data that are approximate Euclidean distances, you are to find the closest set of actual Euclidean distances, in the least-squares sense.

8.10 *Minimax angle fitting.* Suppose that $y_1, \ldots, y_m \in \mathbf{R}^k$ are affine functions of a variable $x \in \mathbf{R}^n$:

$$y_i = A_i x + b_i, \quad i = 1, \ldots, m,$$

and $z_1, \ldots, z_m \in \mathbf{R}^k$ are given nonzero vectors. We want to choose the variable x, subject to some convex constraints, (*e.g.*, linear inequalities) to minimize the maximum angle between y_i and z_i,

$$\max\{\angle(y_1, z_1), \ldots, \angle(y_m, z_m)\}.$$

The angle between nonzero vectors is defined as usual:

$$\angle(u, v) = \cos^{-1}\left(\frac{u^T v}{\|u\|_2 \|v\|_2}\right),$$

where we take $\cos^{-1}(a) \in [0, \pi]$. We are only interested in the case when the optimal objective value does not exceed $\pi/2$.

Formulate this problem as a convex or quasiconvex optimization problem. When the constraints on x are linear inequalities, what kind of problem (or problems) do you have to solve?

8.11 *Smallest Euclidean cone containing given points.* In $\mathbf{R}^n$, we define a *Euclidean cone*, with center direction $c \neq 0$, and angular radius θ, with $0 \leq \theta \leq \pi/2$, as the set

$$\{x \in \mathbf{R}^n \mid \angle(c, x) \leq \theta\}.$$

(A Euclidean cone is a second-order cone, *i.e.*, it can be represented as the image of the second-order cone under a nonsingular linear mapping.)

Let $a_1, \ldots, a_m \in \mathbf{R}$. How would you find the Euclidean cone, of smallest angular radius, that contains $a_1, \ldots, a_m$? (In particular, you should explain how to solve the feasibility problem, *i.e.*, how to determine whether there is a Euclidean cone which contains the points.)

Extremal volume ellipsoids

8.12 Show that the maximum volume ellipsoid enclosed in a set is unique. Show that the Löwner-John ellipsoid of a set is unique.

8.13 *Löwner-John ellipsoid of a simplex.* In this exercise we show that the Löwner-John ellipsoid of a simplex in $\mathbf{R}^n$ must be shrunk by a factor n to fit inside the simplex. Since the Löwner-John ellipsoid is affinely invariant, it is sufficient to show the result for one particular simplex.

Derive the Löwner-John ellipsoid $\mathcal{E}_{\text{lj}}$ for the simplex $C = \mathbf{conv}\{0, e_1, \ldots, e_n\}$. Show that $\mathcal{E}_{\text{lj}}$ must be shrunk by a factor $1/n$ to fit inside the simplex.

8.14 *Efficiency of ellipsoidal inner approximation.* Let C be a polyhedron in $\mathbf{R}^n$ described as $C = \{x \mid Ax \preceq b\}$, and suppose that $\{x \mid Ax \prec b\}$ is nonempty.

(a) Show that the maximum volume ellipsoid enclosed in C, expanded by a factor n about its center, is an ellipsoid that contains C.

(b) Show that if C is symmetric about the origin, *i.e.*, of the form $C = \{x \mid -\mathbf{1} \preceq Ax \preceq \mathbf{1}\}$, then expanding the maximum volume inscribed ellipsoid by a factor $\sqrt{n}$ gives an ellipsoid that contains C.

8.15 *Minimum volume ellipsoid covering union of ellipsoids.* Formulate the following problem as a convex optimization problem. Find the minimum volume ellipsoid $\mathcal{E} = \{x \mid (x - x_0)^T A^{-1}(x - x_0) \leq 1\}$ that contains K given ellipsoids

$$\mathcal{E}_i = \{x \mid x^T A_i x + 2b_i^T x + c_i \leq 0\}, \quad i = 1, \ldots, K.$$

Hint. See appendix B.

8.16 *Maximum volume rectangle inside a polyhedron.* Formulate the following problem as a convex optimization problem. Find the rectangle

$$\mathcal{R} = \{x \in \mathbf{R}^n \mid l \preceq x \preceq u\}$$

of maximum volume, enclosed in a polyhedron $\mathcal{P} = \{x \mid Ax \preceq b\}$. The variables are $l, u \in \mathbf{R}^n$. Your formulation should not involve an exponential number of constraints.

Centering

8.17 *Affine invariance of analytic center.* Show that the analytic center of a set of inequalities is affine invariant. Show that it is invariant with respect to positive scaling of the inequalities.

8.18 *Analytic center and redundant inequalities.* Two sets of linear inequalities that describe the same polyhedron can have different analytic centers. Show that by adding redundant inequalities, we can make *any* interior point x_0 of a polyhedron

$$\mathcal{P} = \{x \in \mathbf{R}^n \mid Ax \preceq b\}$$

the analytic center. More specifically, suppose $A \in \mathbf{R}^{m\times n}$ and $Ax_0 \prec b$. Show that there exist $c \in \mathbf{R}^n$, $\gamma \in \mathbf{R}$, and a positive integer q, such that $\mathcal{P}$ is the solution set of the $m+q$ inequalities

$$Ax \preceq b, \qquad c^T x \le \gamma, \qquad c^T x \le \gamma, \quad \ldots, \quad c^T x \le \gamma \tag{8.36}$$

(where the inequality $c^T x \le \gamma$ is added q times), and x_0 is the analytic center of (8.36).

8.19 Let $x_{\rm ac}$ be the analytic center of a set of linear inequalities

$$a_i^T x \le b_i, \quad i = 1, \ldots, m,$$

and define H as the Hessian of the logarithmic barrier function at $x_{\rm ac}$:

$$H = \sum_{i=1}^m \frac{1}{(b_i - a_i^T x_{\rm ac})^2} a_i a_i^T.$$

Show that the kth inequality is redundant (*i.e.*, it can be deleted without changing the feasible set) if

$$b_k - a_k^T x_{\rm ac} \ge m(a_k^T H^{-1} a_k)^{1/2}.$$

8.20 *Ellipsoidal approximation from analytic center of linear matrix inequality.* Let C be the solution set of the LMI

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n \preceq B,$$

where $A_i, B \in \mathbf{S}^m$, and let $x_{\rm ac}$ be its analytic center. Show that

$$\mathcal{E}_{\rm inner} \subseteq C \subseteq \mathcal{E}_{\rm outer},$$

where

$$\begin{aligned} \mathcal{E}_{\rm inner} &= \{x \mid (x - x_{\rm ac})^T H (x - x_{\rm ac}) \le 1\}, \\ \mathcal{E}_{\rm outer} &= \{x \mid (x - x_{\rm ac})^T H (x - x_{\rm ac}) \le m(m-1)\}, \end{aligned}$$

and H is the Hessian of the logarithmic barrier function

$$-\log\det(B - x_1 A_1 - x_2 A_2 - \cdots - x_n A_n)$$

evaluated at $x_{\rm ac}$.

8.21 [BYT99] *Maximum likelihood interpretation of analytic center.* We use the linear measurement model of page 352,

$$y = Ax + v,$$

where $A \in \mathbf{R}^{m\times n}$. We assume the noise components v_i are IID with support $[-1, 1]$. The set of parameters x consistent with the measurements $y \in \mathbf{R}^m$ is the polyhedron defined by the linear inequalities

$$-\mathbf{1} + y \preceq Ax \preceq \mathbf{1} + y. \tag{8.37}$$

Suppose the probability density function of v_i has the form

$$p(v) = \begin{cases} \alpha_r (1 - v^2)^r & -1 \le v \le 1 \\ 0 & \text{otherwise,} \end{cases}$$

where $r \ge 1$ and $\alpha_r > 0$. Show that the maximum likelihood estimate of x is the analytic center of (8.37).

8.22 *Center of gravity.* The center of gravity of a set $C \subseteq \mathbf{R}^n$ with nonempty interior is defined as

$$x_{\rm cg} = \frac{\int_C u \, du}{\int_C 1 \, du}.$$

The center of gravity is affine invariant, and (clearly) a function of the set C, and not its particular description. Unlike the centers described in the chapter, however, it is very difficult to compute the center of gravity, except in simple cases (*e.g.*, ellipsoids, balls, simplexes).

Show that the center of gravity $x_{\rm cg}$ is the minimizer of the convex function

$$f(x) = \int_C \|u - x\|_2^2 \, du.$$

Classification

8.23 *Robust linear discrimination.* Consider the robust linear discrimination problem given in (8.23).

(a) Show that the optimal value $t^\star$ is positive if and only if the two sets of points can be linearly separated. When the two sets of points can be linearly separated, show that the inequality $\|a\|_2 \leq 1$ is tight, *i.e.*, we have $\|a^\star\|_2 = 1$, for the optimal $a^\star$.

(b) Using the change of variables $\tilde{a} = a/t$, $\tilde{b} = b/t$, prove that the problem (8.23) is equivalent to the QP

$$\begin{array}{ll} \text{minimize} & \|\tilde{a}\|_2 \\ \text{subject to} & \tilde{a}^T x_i - \tilde{b} \geq 1, \quad i = 1, \ldots, N \\ & \tilde{a}^T y_i - \tilde{b} \leq -1, \quad i = 1, \ldots, M. \end{array}$$

8.24 *Linear discrimination maximally robust to weight errors.* Suppose we are given two sets of points $\{x_1, \ldots, x_N\}$ and and $\{y_1, \ldots, y_M\}$ in $\mathbf{R}^n$ that can be linearly separated. In §8.6.1 we showed how to find the affine function that discriminates the sets, and gives the largest gap in function values. We can also consider robustness with respect to changes in the vector a, which is sometimes called the *weight vector*. For a given a and b for which $f(x) = a^T x - b$ separates the two sets, we define the *weight error margin* as the norm of the smallest $u \in \mathbf{R}^n$ such that the affine function $(a+u)^T x - b$ no longer separates the two sets of points. In other words, the weight error margin is the maximum ρ such that

$$(a+u)^T x_i \geq b, \quad i = 1, \ldots, N, \qquad (a+u)^T y_j \leq b, \quad i = 1, \ldots, M,$$

holds for all u with $\|u\|_2 \leq \rho$.

Show how to find a and b that maximize the weight error margin, subject to the normalization constraint $\|a\|_2 \leq 1$.

8.25 *Most spherical separating ellipsoid.* We are given two sets of vectors $x_1, \ldots, x_N \in \mathbf{R}^n$, and $y_1, \ldots, y_M \in \mathbf{R}^n$, and wish to find the ellipsoid with minimum eccentricity (*i.e.*, minimum condition number of the defining matrix) that contains the points $x_1, \ldots, x_N$, but not the points $y_1, \ldots, y_M$. Formulate this as a convex optimization problem.

Placement and floor planning

8.26 *Quadratic placement.* We consider a placement problem in $\mathbf{R}^2$, defined by an undirected graph $\mathcal{A}$ with N nodes, and with quadratic costs:

$$\text{minimize} \quad \textstyle\sum_{(i,j)\in\mathcal{A}} \|x_i - x_j\|_2^2.$$

The variables are the positions $x_i \in \mathbf{R}^2$, $i = 1, \ldots, M$. The positions x_i, $i = M+1, \ldots, N$ are given. We define two vectors $u, v \in \mathbf{R}^M$ by

$$u = (x_{11}, x_{21}, \ldots, x_{M1}), \qquad v = (x_{12}, x_{22}, \ldots, x_{M2}),$$

containing the first and second components, respectively, of the free nodes.

Show that u and v can be found by solving two sets of linear equations,

$$Cu = d_1, \qquad Cv = d_2,$$

where $C \in \mathbf{S}^M$. Give a simple expression for the coefficients of C in terms of the graph $\mathcal{A}$.

8.27 *Problems with minimum distance constraints.* We consider a problem with variables $x_1, \ldots, x_N \in \mathbf{R}^k$. The objective, $f_0(x_1, \ldots, x_N)$, is convex, and the constraints

$$f_i(x_1, \ldots, x_N) \leq 0, \quad i = 1, \ldots, m,$$

are convex (*i.e.*, the functions $f_i : \mathbf{R}^{Nk} \to \mathbf{R}$ are convex). In addition, we have the *minimum distance constraints*

$$\|x_i - x_j\|_2 \geq D_{\min}, \quad i \neq j, \quad i, j = 1, \ldots, N.$$

In general, this is a hard nonconvex problem.

Following the approach taken in floorplanning, we can form a *convex restriction* of the problem, *i.e.*, a problem which is convex, but has a smaller feasible set. (Solving the restricted problem is therefore easy, and any solution is guaranteed to be feasible for the nonconvex problem.) Let $a_{ij} \in \mathbf{R}^k$, for $i < j$, $i, j = 1, \ldots, N$, satisfy $\|a_{ij}\|_2 = 1$.

Show that the restricted problem

$$\begin{array}{ll} \text{minimize} & f_0(x_1, \ldots, x_N) \\ \text{subject to} & f_i(x_1, \ldots, x_N) \leq 0, \quad i = 1, \ldots, m \\ & a_{ij}^T(x_i - x_j) \geq D_{\min}, \quad i < j, \ i, j = 1, \ldots, N, \end{array}$$

is convex, and that every feasible point satisfies the minimum distance constraint.

Remark. There are many good heuristics for choosing the directions a_{ij}. One simple one starts with an approximate solution $\hat{x}_1, \ldots, \hat{x}_N$ (that need not satisfy the minimum distance constraints). We then set $a_{ij} = (\hat{x}_i - \hat{x}_j)/\|\hat{x}_i - \hat{x}_j\|_2$.

Miscellaneous problems

8.28 Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be two polyhedra described as

$$\mathcal{P}_1 = \{x \mid Ax \preceq b\}, \qquad \mathcal{P}_2 = \{x \mid -\mathbf{1} \preceq Cx \preceq \mathbf{1}\},$$

where $A \in \mathbf{R}^{m \times n}$, $C \in \mathbf{R}^{p \times n}$, and $b \in \mathbf{R}^m$. The polyhedron $\mathcal{P}_2$ is symmetric about the origin. For $t \geq 0$ and $x_c \in \mathbf{R}^n$, we use the notation $t\mathcal{P}_2 + x_c$ to denote the polyhedron

$$t\mathcal{P}_2 + x_c = \{tx + x_c \mid x \in \mathcal{P}_2\},$$

which is obtained by first scaling $\mathcal{P}_2$ by a factor t about the origin, and then translating its center to x_c.

Show how to solve the following two problems, via an LP, or a set of LPs.

(a) Find the largest polyhedron $t\mathcal{P}_2 + x_c$ enclosed in $\mathcal{P}_1$, *i.e.*,

$$\begin{array}{ll} \text{maximize} & t \\ \text{subject to} & t\mathcal{P}_2 + x_c \subseteq \mathcal{P}_1 \\ & t \geq 0. \end{array}$$

(b) Find the smallest polyhedron $t\mathcal{P}_2 + x_c$ containing $\mathcal{P}_1$, *i.e.*,

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \mathcal{P}_1 \subseteq t\mathcal{P}_2 + x_c \\ & t \geq 0. \end{array}$$

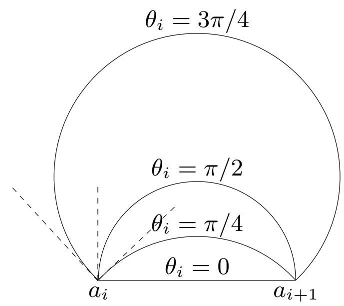
In both problems the variables are $t \in \mathbf{R}$ and $x_c \in \mathbf{R}^n$.

8.29 *Outer polyhedral approximations.* Let $\mathcal{P} = \{x \in \mathbf{R}^n \mid Ax \preceq b\}$ be a polyhedron, and $C \subseteq \mathbf{R}^n$ a given set (not necessarily convex). Use the support function S_C to formulate the following problem as an LP:

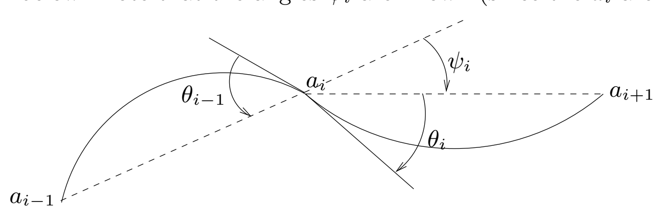
$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & C \subseteq t\mathcal{P} + x \\ & t \geq 0. \end{array}$$

Here $t\mathcal{P}+x = \{tu+x \mid u \in \mathcal{P}\}$, the polyhedron $\mathcal{P}$ scaled by a factor of t about the origin, and translated by x. The variables are $t \in \mathbf{R}$ and $x \in \mathbf{R}^n$.

8.30 *Interpolation with piecewise-arc curve.* A sequence of points $a_1, \ldots, a_n \in \mathbf{R}^2$ is given. We construct a curve that passes through these points, in order, and is an arc (*i.e.*, part of a circle) or line segment (which we think of as an arc of infinite radius) between consecutive points. Many arcs connect a_i and a_{i+1}; we parameterize these arcs by giving the angle $\theta_i \in (-\pi, \pi)$ between its tangent at a_i and the line segment $[a_i, a_{i+1}]$. Thus, $\theta_i = 0$ means the arc between a_i and a_{i+1} is in fact the line segment $[a_i, a_{i+1}]$; $\theta_i = \pi/2$ means the arc between a_i and a_{i+1} is a half-circle (above the linear segment $[a_1, a_2]$); $\theta_i = -\pi/2$ means the arc between a_i and a_{i+1} is a half-circle (below the linear segment $[a_1, a_2]$). This is illustrated below.



Our curve is completely specified by the angles $\theta_1, \ldots, \theta_n$, which can be chosen in the interval $(-\pi, \pi)$. The choice of θ_i affects several properties of the curve, for example, its *total arc length* L, or the *joint angle discontinuities*, which can be described as follows. At each point a_i, $i = 2, \ldots, n-1$, two arcs meet, one coming from the previous point and one going to the next point. If the tangents to these arcs exactly oppose each other, so the curve is differentiable at a_i, we say there is no joint angle discontinuity at a_i. In general, we define the joint angle discontinuity at a_i as $|\theta_{i-1}+\theta_i+\psi_i|$, where ψ_i is the angle between the line segment $[a_i, a_{i+1}]$ and the line segment $[a_{i-1}, a_i]$, *i.e.*, $\psi_i = \angle(a_i - a_{i+1}, a_{i-1} - a_i)$. This is shown below. Note that the angles ψ_i are known (since the a_i are known).



We define the *total joint angle discontinuity* as

$$D = \sum_{i=2}^{n} |\theta_{i-1} + \theta_i + \psi_i|.$$

Formulate the problem of minimizing total arc length length L, and total joint angle discontinuity D, as a bi-criterion convex optimization problem. Explain how you would find the extreme points on the optimal trade-off curve.