8. Geometric problems

• extremal volume ellipsoids

• centering

• classification

• placement and facility location
**Minimum volume ellipsoid around a set**

**Löwner-John ellipsoid** of a set $C$: minimum volume ellipsoid $E$ s.t. $C \subseteq E$

- parametrize $E$ as $E = \{v \mid \|Av + b\|_2 \leq 1\}$; w.l.o.g. assume $A \in S^{n}_{++}$
- $\text{vol } E$ is proportional to $\det A^{-1}$; to compute minimum volume ellipsoid,

  $$\text{minimize (over } A, b) \quad \log \det A^{-1}$$

  $$\text{subject to } \sup_{v \in C} \|Av + b\|_2 \leq 1$$

  convex, but evaluating the constraint can be hard (for general $C$)

**finite set** $C = \{x_1, \ldots, x_m\}$:

  $$\text{minimize (over } A, b) \quad \log \det A^{-1}$$

  $$\text{subject to } \|Ax_i + b\|_2 \leq 1, \quad i = 1, \ldots, m$$

also gives Löwner-John ellipsoid for polyhedron $\text{conv}\{x_1, \ldots, x_m\}$
Maximum volume inscribed ellipsoid

maximum volume ellipsoid $\mathcal{E}$ inside a convex set $C \subseteq \mathbb{R}^n$

- parametrize $\mathcal{E}$ as $\mathcal{E} = \{Bu + d \mid \|u\|_2 \leq 1\}$; w.l.o.g. assume $B \in \mathbb{S}_{++}^n$
- $\text{vol} \mathcal{E}$ is proportional to $\det B$; can compute $\mathcal{E}$ by solving

\[
\begin{align*}
\text{maximize} & \quad \log \det B \\
\text{subject to} & \quad \sup_{\|u\|_2 \leq 1} I_C(Bu + d) \leq 0
\end{align*}
\]

(where $I_C(x) = 0$ for $x \in C$ and $I_C(x) = \infty$ for $x \notin C$)

convex, but evaluating the constraint can be hard (for general $C$)

polyhedron $\{x \mid a^T_i x \leq b_i, \ i = 1, \ldots, m\}$:

\[
\begin{align*}
\text{maximize} & \quad \log \det B \\
\text{subject to} & \quad \|Ba_i\|_2 + a^T_i d \leq b_i, \quad i = 1, \ldots, m
\end{align*}
\]

(constraint follows from $\sup_{\|u\|_2 \leq 1} a^T_i (Bu + d) = \|Ba_i\|_2 + a^T_i d$)
Efficiency of ellipsoidal approximations

$C \subseteq \mathbb{R}^n$ convex, bounded, with nonempty interior

- Löwner-John ellipsoid, shrunk by a factor $n$, lies inside $C$
- maximum volume inscribed ellipsoid, expanded by a factor $n$, covers $C$

example (for two polyhedra in $\mathbb{R}^2$)

factor $n$ can be improved to $\sqrt{n}$ if $C$ is symmetric
Centering

some possible definitions of ‘center’ of a convex set $C$:

- center of largest inscribed ball (‘Chebyshev center’) for polyhedron, can be computed via linear programming (page 4–19)
- center of maximum volume inscribed ellipsoid (page 1–3)

MVE center is invariant under affine coordinate transformations
Analytic center of a set of inequalities

the analytic center of set of convex inequalities and linear equations

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

is defined as the optimal point of

\[
\begin{align*}
m\text{inimize} & \quad - \sum_{i=1}^{m} \log(-f_i(x)) \\
\text{subject to} & \quad Fx = g
\end{align*}
\]

• more easily computed than MVE or Chebyshev center (see later)
• not just a property of the feasible set: two sets of inequalities can describe the same set, but have different analytic centers
analytic center of linear inequalities $a^T_i x \leq b_i, \ i = 1, \ldots, m$

$x_{ac}$ is minimizer of

$$
\phi(x) = -\sum_{i=1}^{m} \log(b_i - a^T_i x)
$$

inner and outer ellipsoids from analytic center:

$$
E_{inner} \subseteq \{x \mid a^T_i x \leq b_i, \ i = 1, \ldots, m\} \subseteq E_{outer}
$$

where

$$
E_{inner} = \{x \mid (x - x_{ac})^T \nabla^2 \phi(x_{ac})(x - x_{ac}) \leq 1\}
$$

$$
E_{outer} = \{x \mid (x - x_{ac})^T \nabla^2 \phi(x_{ac})(x - x_{ac}) \leq m(m - 1)\}
$$
Linear discrimination

separate two sets of points \( \{x_1, \ldots, x_N\}, \{y_1, \ldots, y_M\} \) by a hyperplane:

\[
a^T x_i + b > 0, \quad i = 1, \ldots, N, \quad a^T y_i + b < 0, \quad i = 1, \ldots, M
\]

homogeneous in \( a, b \), hence equivalent to

\[
a^T x_i + b \geq 1, \quad i = 1, \ldots, N, \quad a^T y_i + b \leq -1, \quad i = 1, \ldots, M
\]

a set of linear inequalities in \( a, b \)
Robust linear discrimination

(Euclidean) distance between hyperplanes

\[ \mathcal{H}_1 = \{ z \mid a^T z + b = 1 \} \]
\[ \mathcal{H}_2 = \{ z \mid a^T z + b = -1 \} \]

is \( \text{dist}(\mathcal{H}_1, \mathcal{H}_2) = 2/\|a\|_2 \)

to separate two sets of points by maximum margin,

\[
\begin{aligned}
\text{minimize} & \quad (1/2)\|a\|_2 \\
\text{subject to} & \quad a^T x_i + b \geq 1, \quad i = 1, \ldots, N \\
& \quad a^T y_i + b \leq -1, \quad i = 1, \ldots, M
\end{aligned}
\]

(after squaring objective) a QP in \( a, b \)
Lagrange dual of maximum margin separation problem (1)

maximize \( 1^T \lambda + 1^T \mu \)
subject to \( 2 \left\| \sum_{i=1}^{N} \lambda_i x_i - \sum_{i=1}^{M} \mu_i y_i \right\|_2 \leq 1 \)
\( 1^T \lambda = 1^T \mu, \quad \lambda \succeq 0, \quad \mu \succeq 0 \) (2)

from duality, optimal value is inverse of maximum margin of separation interpretation

• change variables to \( \theta_i = \lambda_i / 1^T \lambda, \quad \gamma_i = \mu_i / 1^T \mu, \quad t = 1 / (1^T \lambda + 1^T \mu) \)
• invert objective to minimize \( 1 / (1^T \lambda + 1^T \mu) = t \)

minimize \( t \)
subject to \( \left\| \sum_{i=1}^{N} \theta_i x_i - \sum_{i=1}^{M} \gamma_i y_i \right\|_2 \leq t \)
\( \theta \succeq 0, \quad 1^T \theta = 1, \quad \gamma \succeq 0, \quad 1^T \gamma = 1 \)

optimal value is distance between convex hulls
Approximate linear separation of non-separable sets

\[
\begin{align*}
\text{minimize} & \quad 1^T u + 1^T v \\
\text{subject to} & \quad a^T x_i + b \geq 1 - u_i, \quad i = 1, \ldots, N \\
& \quad a^T y_i + b \leq -1 + v_i, \quad i = 1, \ldots, M \\
& \quad u \succeq 0, \quad v \succeq 0
\end{align*}
\]

• an LP in \(a, b, u, v\)

• at optimum, \(u_i = \max\{0, 1 - a^T x_i - b\}\), \(v_i = \max\{0, 1 + a^T y_i + b\}\)

• can be interpreted as a heuristic for minimizing \#misclassified points
Support vector classifier

minimize \[ \|a\|_2 + \gamma(1^T u + 1^T v) \]
subject to \[ a^T x_i + b \geq 1 - u_i, \quad i = 1, \ldots, N \]
\[ a^T y_i + b \leq -1 + v_i, \quad i = 1, \ldots, M \]
\[ u \succeq 0, \quad v \succeq 0 \]

produces point on trade-off curve between inverse of margin \(2/\|a\|_2\) and classification error, measured by total slack \(1^T u + 1^T v\)

same example as previous page, with \(\gamma = 0.1\):
Nonlinear discrimination

separate two sets of points by a nonlinear function:

\[ f(x_i) > 0, \quad i = 1, \ldots, N, \quad f(y_i) < 0, \quad i = 1, \ldots, M \]

• choose a linearly parametrized family of functions

\[ f(z) = \theta^T F(z) \]

\[ F = (F_1, \ldots, F_k) : \mathbb{R}^n \to \mathbb{R}^k \text{ are basis functions} \]

• solve a set of linear inequalities in \( \theta \):

\[ \theta^T F(x_i) \geq 1, \quad i = 1, \ldots, N, \quad \theta^T F(y_i) \leq -1, \quad i = 1, \ldots, M \]
**quadratic discrimination:** \[ f(z) = z^T P z + q^T z + r \]

\[ x_i^T P x_i + q^T x_i + r \geq 1, \quad y_i^T P y_i + q^T y_i + r \leq -1 \]

can add additional constraints (e.g., \( P \preceq -I \) to separate by an ellipsoid)

**polynomial discrimination:** \( F(z) \) are all monomials up to a given degree
Placement and facility location

• $N$ points with coordinates $x_i \in \mathbb{R}^2$ (or $\mathbb{R}^3$)
• some positions $x_i$ are given; the other $x_i$’s are variables
• for each pair of points, a cost function $f_{ij}(x_i, x_j)$

**placement problem**

\[
\text{minimize} \quad \sum_{i \neq j} f_{ij}(x_i, x_j)
\]

variables are positions of free points

**interpretations**

• points represent plants or warehouses; $f_{ij}$ is transportation cost between facilities $i$ and $j$
• points represent cells on an IC; $f_{ij}$ represents wirelength
**example:** minimize $\sum_{(i,j) \in A} h(\|x_i - x_j\|_2)$, with 6 free points, 27 links

optimal placement for $h(z) = z$, $h(z) = z^2$, $h(z) = z^4$

histograms of connection lengths $\|x_i - x_j\|_2$