Convex Optimization

Stephen Boyd   Lieven Vandenberghe

Revised slides by Stephen Boyd, Lieven Vandenberghe, and Parth Nobel

8. Geometric problems

Outline

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 $\mathcal{E}$ with $C \subseteq \mathcal{E}$
- parametrize $\mathcal{E}$ as $\mathcal{E} = \{v \mid \|Av + b\|_2 \leq 1\}$; can assume $A \in \mathbf{S}^n_{++}$
- $\mathbf{vol}\,\mathcal{E}$ is proportional to $\det A^{-1}$; to find Löwner-John ellipsoid, solve problem

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

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

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

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

  also gives Löwner-John ellipsoid for polyhedron $\mathbf{conv}\{x_1, \ldots, x_m\}$

Maximum volume inscribed ellipsoid

- maximum volume ellipsoid $\mathcal{E}$ with $\mathcal{E} \subseteq C$, $C \subseteq \mathbf{R}^n$ convex
- parametrize $\mathcal{E}$ as $\mathcal{E} = \{Bu + d \mid \|u\|_2 \leq 1\}$; can assume $B \in \mathbf{S}^n_{++}$
- $\mathbf{vol}\, \mathcal{E}$ is proportional to $\det B$; can find $\mathcal{E}$ by solving

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

  (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)

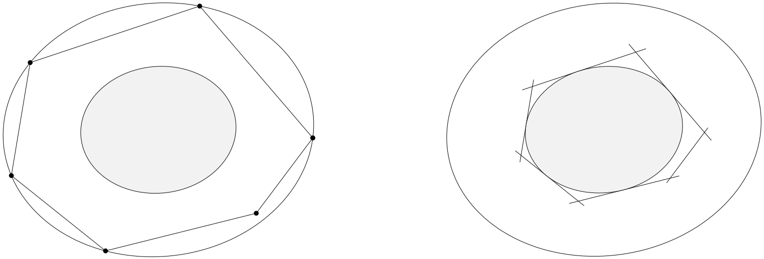
- **polyhedron** $\{x \mid a_i^T x \leq b_i,\ i = 1, \ldots, m\}$:

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

  (constraint follows from $\sup_{\|u\|_2 \leq 1} a_i^T (Bu + d) = \|Ba_i\|_2 + a_i^T d$)

Efficiency of ellipsoidal approximations

- $C \subseteq \mathbf{R}^n$ convex, bounded, with nonempty interior
- Löwner-John ellipsoid, shrunk by a factor n (around its center), lies inside C
- maximum volume inscribed ellipsoid, expanded by a factor n (around its center) covers C
- **example** (for polyhedra in $\mathbf{R}^2$)



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

Outline

Extremal volume ellipsoids

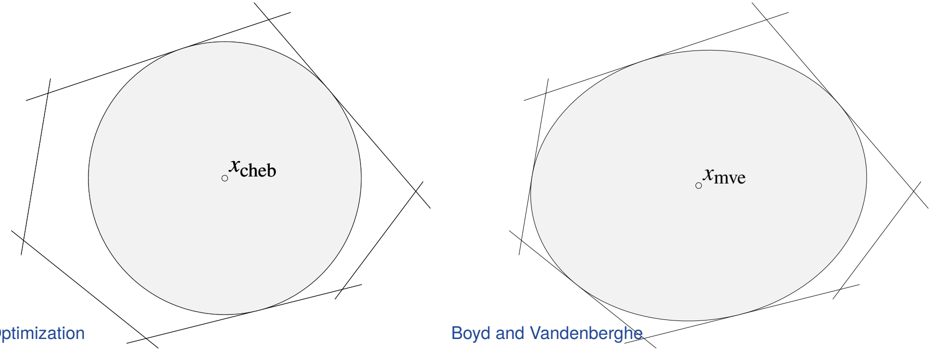
Centering

Classification

Placement and facility location

Centering

- many possible definitions of 'center' of a convex set C
- Chebyshev center: center of largest inscribed ball
  - for polyhedron, can be found via linear programming
- center of maximum volume inscribed ellipsoid
  - invariant under affine coordinate transformations



Analytic center of a set of inequalities

- the **analytic center** of set of convex inequalities and linear equations

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

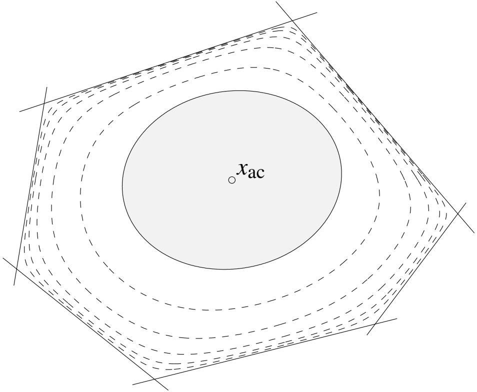
  is defined as solution of

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

- objective is called the **log-barrier** for the inequalities
- (we'll see later) analytic center more easily computed than MVE or Chebyshev center
- two sets of inequalities can describe the same set, but have different analytic centers

Analytic center of linear inequalities

- $a_i^T x \leq b_i$, $i = 1, \ldots, m$
- x_{ac} minimizes $\phi(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$
- dashed lines are level curves of ϕ



Inner and outer ellipsoids from analytic center

- we have

$$\mathcal{E}_{\text{inner}} \subseteq \{x \mid a_i^T x \leq b_i,\ i = 1, \ldots, m\} \subseteq \mathcal{E}_{\text{outer}}$$

  where

$$\begin{aligned} \mathcal{E}_{\text{inner}} &= \{x \mid (x - x_{\text{ac}})^T \nabla^2 \phi(x_{\text{ac}})(x - x_{\text{ac}}) \leq 1\} \\ \mathcal{E}_{\text{outer}} &= \{x \mid (x - x_{\text{ac}})^T \nabla^2 \phi(x_{\text{ac}})(x - x_{\text{ac}}) \leq m(m-1)\} \end{aligned}$$

- ellipsoid expansion/shrinkage factor is $\sqrt{m(m-1)}$
  (cf. n for Löwner-John or max volume inscribed ellpsoids)

Outline

Linear discrimination

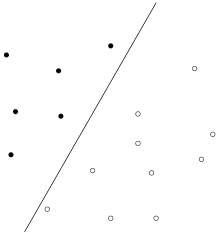
- separate two sets of points $\{x_1, \ldots, x_N\}$, $\{y_1, \ldots, y_M\}$ by a hyperplane
- *i.e.*, find $a \in \mathbf{R}^n$, $b \in \mathbf{R}$ with

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

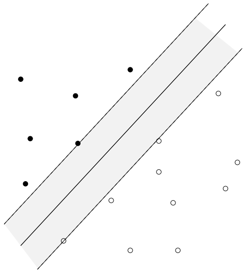
- homogeneous in a, b, hence equivalent to

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

a set of linear inequalities in a, b, *i.e.*, an LP feasibility problem



Robust linear discrimination



(Euclidean) distance between hyperplanes

$$
\begin{aligned}
\mathcal{H}_1 &= \{z \mid a^T z + b = 1\} \\
\mathcal{H}_2 &= \{z \mid a^T z + b = -1\}
\end{aligned}
$$

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

to separate two sets of points by maximum margin,

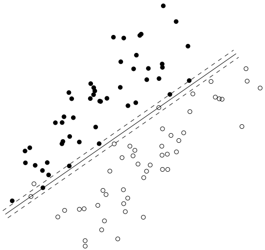
$$
\begin{array}{ll}
\text{minimize} & (1/2)\|a\|_2^2 \\
\text{subject to} & a^T x_i + b \geq 1, \quad i = 1, \ldots, N \\
& a^T y_i + b \leq -1, \quad i = 1, \ldots, M
\end{array} \qquad (2)
$$

a QP in a, b

Approximate linear separation of non-separable sets

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

- an LP in a, b, u, v
- at optimum, $u_i = \max\{0, 1 - a^T x_i - b\}$, $v_i = \max\{0, 1 + a^T y_i + b\}$
- equivalent to minimizing the sum of violations of the original inequalities

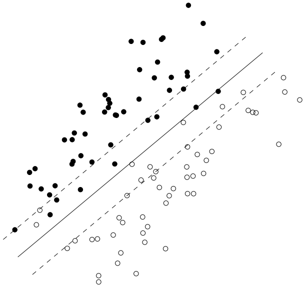


Support vector classifier

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

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

example on previous slide, with $\gamma = 0.1$:



Nonlinear discrimination

- separate two sets of points by a nonlinear function f: find $f : \mathbf{R}^n \to \mathbf{R}$ with

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

- choose a linearly parametrized family of functions $f(z) = \theta^T F(z)$
  - $\theta \in \mathbf{R}^k$ is parameter
  - $F = (F_1, \ldots, F_k) : \mathbf{R}^n \to \mathbf{R}^k$ are basis functions

- solve a set of linear inequalities in θ:

$$\theta^T F(x_i) \geq 1, \quad i = 1, \ldots, N, \qquad \theta^T F(y_i) \leq -1, \quad i = 1, \ldots, M$$

Examples

- **quadratic discrimination**: $f(z) = z^T P z + q^T z + r$, $\theta = (P, q, r)$
- solve LP feasibility problem with variables $P \in \mathbf{S}^n$, $q \in \mathbf{R}^n$, $r \in \mathbf{R}$

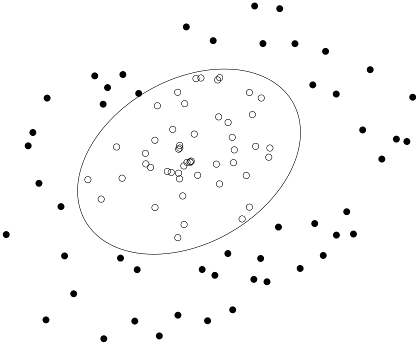
$$x_i^T P x_i + q^T x_i + r \geq 1, \qquad y_i^T P y_i + q^T y_i + r \leq -1$$

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

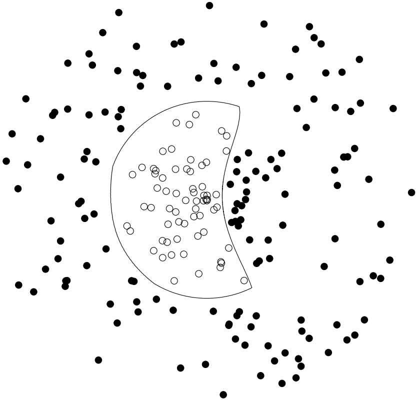
- **polynomial discrimination**: $F(z)$ are all monomials up to a given degree d
- *e.g.*, for $n = 2$, $d = 3$

$$F(z) = (1,\ z_1,\ z_2,\ z_1^2,\ z_1 z_2,\ z_2^2,\ z_1^3,\ z_1^2 z_2,\ z_1 z_2^2,\ z_2^3)$$

Example



separation by ellipsoid



separation by 4th degree polynomial

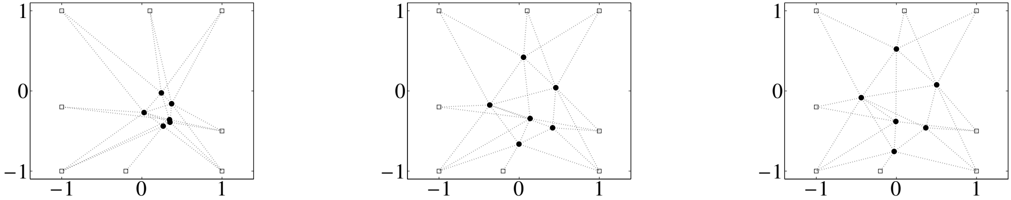
Outline

Placement and facility location

- N points with coordinates $x_i \in \mathbf{R}^2$ (or $\mathbf{R}^3$)
- 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**: minimize $\sum_{i \neq j} f_{ij}(x_i, x_j)$
- **interpretations**
  - points are locations of plants or warehouses; f_{ij} is transportation cost between facilities i and j
  - points are locations of cells in an integrated circuit; f_{ij} represents wirelength

Example

- minimize $\sum_{(i,j)\in\mathcal{E}} h(\|x_i - x_j\|_2)$, with 6 free points, 27 edges
- optimal placements for $h(z) = z$, $h(z) = z^2$, $h(z) = z^4$



- histograms of edge lengths $\|x_i - x_j\|_2$, $(i,) \in \mathcal{E}$