EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
Lecture 17

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

May 29, 2012

Outline

Distance and Projection on a Set

The *distance* from a point $\mathbf{x}_0 \in \mathbb{R}^n$ to a closed set $\mathcal{C} \subseteq \mathbb{R}^n$, with respect to the norm $||\cdot||$, is defined as

$$\mathrm{dist}(\mathbf{x}_0, \mathcal{C}) \triangleq \inf\{||\mathbf{x} - \mathbf{x}_0|| : \mathbf{x} \in \mathcal{C}\} \; .$$

The *projection* of $\mathbf{x}_0$ on $\mathcal{C}$ is any point $\mathbf{z} \in \mathcal{C}$ for which $||\mathbf{z} - \mathbf{x}_0|| = \mathrm{dist}(\mathbf{x}_0, \mathcal{C})$. The notation $P_{\mathcal{C}} : \mathbb{R}^n \to \mathbb{R}^n$ is used to denote any function for which $P_{\mathcal{C}}(\mathbf{x}_0)$ is a projection of $\mathbf{x}_0$ onto $\mathcal{C}$, i.e., for all $\mathbf{x}_0$, we have

$$P_{\mathcal{C}}(\mathbf{x}_0) \in \mathcal{C} \, , \; ||P_{\mathcal{C}}(\mathbf{x}_0) - \mathbf{x}_0|| = \mathrm{dist}(\mathbf{x}_0, \mathcal{C}) \iff P_{\mathcal{C}}(\mathbf{x}_0) = \mathrm{argmin}\{||\mathbf{x} - \mathbf{x}_0|| : \mathbf{x} \in \mathcal{C}\} \; .$$

Properties:

- If $\mathcal{C}$ is closed and convex, and $||\cdot||$ is strictly convex, then for any $\mathbf{x}_0$, there is a *unique* projection $P_{\mathcal{C}}(\mathbf{x}_0)$ onto $\mathcal{C}$.
- If, for every $\mathbf{x}_0$, there is a unique *Euclidean* projection of $\mathbf{x}_0$ on $\mathcal{C}$, then $\mathcal{C}$ is closed and convex.

Example: *Projection onto rank-k matrices*
Suppose that we have

$$\mathbf{X}_0 \in \mathbb{R}^{m \times n} \, , \; \mathcal{C} = \left\{\mathbf{X} \in \mathbb{R}^{m \times n} : \mathrm{rank}(\mathbf{X}) \leq k\right\} \, , \; ||\cdot|| = ||\cdot||_2 = \text{ spectral norm} \, .$$

If $\mathbf{X}_0$ has an SVD given by $\mathbf{X}_0 = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$, where $r = \mathrm{rank}(\mathbf{X}_0)$, and we have $\sigma_1 \geq \cdots \geq \sigma_r > 0$, then $\mathbf{Y} = \sum_{i=1}^{\min\{k,r\}} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ is a projection of $\mathbf{X}_0$ on $\mathcal{C}$.

Projection Via Indicator and Support Functions

The problem of finding the distance from a point $\mathbf{x}_0$ to a set $\mathcal{C}$ (along with its dual) can be expressed compactly using the *indicator function* $I_{\mathcal{C}}$ and the *support function* $S_{\mathcal{C}}$:

$$I_{\mathcal{C}}(\mathbf{x}) \triangleq \begin{cases} 0\,, & \mathbf{x} \in \mathcal{C} \\ \infty\,, & \mathbf{x} \notin \mathcal{C} \end{cases}\,, \quad S_{\mathcal{C}}(\mathbf{x}) \triangleq \sup_{\mathbf{y} \in \mathcal{C}} \mathbf{x}^T \mathbf{y}\,.$$

The primal problem of finding the distance between $\mathbf{x}_0$ and $\mathcal{C}$ can then be expressed as

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x} - \mathbf{x}_0|| \\ \text{subject to} & I_{\mathcal{C}}(\mathbf{x}) \leq 0 \end{array} \iff \begin{array}{ll} \text{minimize} & ||\mathbf{y}|| \\ \text{subject to} & I_{\mathcal{C}}(\mathbf{x}) \leq 0 \\ & \mathbf{x}_0 - \mathbf{x} = \mathbf{y} \end{array}\,.$$

Note that the primal problem is convex if $\mathcal{C}$ is a *closed convex* set. The dual problem of the second form of the primal is given by

$$\begin{array}{ll} \text{maximize} & \mathbf{z}^T \mathbf{x}_0 - S_{\mathcal{C}}(\mathbf{z}) \\ \text{subject to} & ||\mathbf{z}||_{\circledast} \leq 1 \end{array}\,.$$

- If $\mathcal{C}$ is a closed convex set (i.e., the primal is convex), then if $\mathbf{z}$ is dual optimal with a positive objective value, then $\mathbf{z}^T \mathbf{x}_0 > \mathbf{z}^T \mathbf{x}$ for all $\mathbf{x} \in \mathcal{C}$, i.e., $\mathbf{z}$ defines a *separating hyperplane* between $\mathbf{x}_0$ and $\mathcal{C}$.

Projecting a Point on a Convex Set

If $\mathcal{C}$ is convex, then $P_{\mathcal{C}}(\mathbf{x}_0)$ and $\operatorname{dist}(\mathbf{x}_0, \mathcal{C})$ can be computed via convex optimization. Representing $\mathcal{C}$ by a set of convex inequalities $f_i(\mathbf{x}) \leq 0$ for $i = 1, \ldots, m$ and affine equalities $\mathbf{A}\mathbf{x} = \mathbf{b}$ ($\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{b} \in \mathbb{R}^p$), $\operatorname{dist}(\mathbf{x}_0, \mathcal{C})$ and $P_{\mathcal{C}}(\mathbf{x}_0)$ can be found by solving

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x} - \mathbf{x}_0|| \\ \text{subject to} & f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} \iff \begin{array}{ll} \text{minimize} & ||\mathbf{y}|| \\ \text{subject to} & f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x}_0 - \mathbf{x} = \mathbf{y} \end{array} .$$

The dual problem of the second form of the primal is given by the following.

$$\begin{array}{ll} \text{maximize} & \boldsymbol{\mu}^T \mathbf{x}_0 + \inf\limits_{x} \left\{ \sum\limits_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \boldsymbol{\nu}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) - \boldsymbol{\mu}^T \mathbf{x} \right\} \\ \text{subject to} & \boldsymbol{\lambda} \succeq \mathbf{0} \\ & ||\boldsymbol{\mu}||_{\circledast} \leq 1 \end{array} .$$

- The primal problem is feasible if and only if $\mathcal{C}$ is nonempty.
- When the primal is feasible, its optimal value is $\operatorname{dist}(\mathbf{x}_0, \mathcal{C})$, and any optimal point is a projection of $\mathbf{x}_0$ on $\mathcal{C}$.
- The dual problem leads to insights regarding separation of $\mathbf{x}_0$ and $\mathcal{C}$.

Euclidean Projection Examples

Euclidean projection on a polyhedron:

The projection of $\mathbf{x}_0$ on a polyhedron $\mathcal{P} \triangleq \{\mathbf{x} : \mathbf{A}\mathbf{x} \preceq \mathbf{b}\}$ can be computed by solving the QP

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x} - \mathbf{x}_0||_2^2 \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

- For the hyperplane $\mathcal{C}_{\text{hp}} = \{\mathbf{x} : \mathbf{a}^T\mathbf{x} = b\}$ and halfspace $\mathcal{C}_{\text{hs}} = \{\mathbf{x} : \mathbf{a}^T\mathbf{x} \leq b\}$, we have

$$P_{\mathcal{C}_{\text{hp}}}(\mathbf{x}_0) = \mathbf{x}_0 + \frac{\left(b - \mathbf{a}^T\mathbf{x}_0\right)}{||\mathbf{a}||_2^2}\mathbf{a}\,,\; P_{\mathcal{C}_{\text{hs}}}(\mathbf{x}_0) = \begin{cases} \mathbf{x}_0 + \left(b - \mathbf{a}^T\mathbf{x}_0\right)\mathbf{a}/\,||\mathbf{a}||_2^2\,, & \mathbf{a}^T\mathbf{x}_0 > b \\ \mathbf{x}_0\,, & \mathbf{a}^T\mathbf{x}_0 \leq b \end{cases} .$$

- For the box $\mathcal{C}_{\text{box}} = \{\mathbf{x} : \mathbf{l} \preceq \mathbf{x} \preceq \mathbf{u}\}$ (where $\mathbf{l} \prec \mathbf{u}$), we have $\left[P_{\mathcal{C}_{\text{box}}}(\mathbf{x}_0)\right]_k = l_k$ for $[\mathbf{x}_0]_k \leq l_k$, $[\mathbf{x}_0]_k$ for $l_k \leq [\mathbf{x}_0]_k \leq u_k$, and u_k for $[\mathbf{x}_0]_k \geq u_k$.

Euclidean projection on a proper cone:

The projection of $\mathbf{x}_0$ onto a proper cone $\mathcal{K}$ can be obtained by solving the problem

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x} - \mathbf{x}_0||_2^2 \\ \text{subject to} & \mathbf{x} \succeq_{\mathcal{K}} \mathbf{0} \end{array} .$$

The KKT conditions for this problem can be expressed in terms of a primal nonnegative and dual nonpositive decomposition as follows.

$$\mathbf{x}_0 = \mathbf{x}_+ + \mathbf{x}_-\,,\; \mathbf{x}_+ \succeq_{\mathcal{K}} \mathbf{0}\,,\; \mathbf{x}_- \preceq_{\mathcal{K}^\circledast} \mathbf{0}\,,\; \mathbf{x}_+^T\mathbf{x}_- = 0\,.$$

- For $\mathcal{K} = \mathbb{R}_+^n$, we have $\left[P_{\mathcal{K}}(\mathbf{x}_0)\right]_k = \max\left\{[\mathbf{x}_0]_k\,, 0\right\}$.
- For $\mathcal{K} = \mathbb{S}_+^n$ and the Frobenius norm $||\cdot||_F$, we have $P_{\mathcal{K}}(\mathbf{X}_0) = \sum_{i=1}^n \max\{\lambda_i, 0\}\,\mathbf{v}_i\mathbf{v}_i^T$, where $\mathbf{X}_0 = \sum_{i=1}^n \lambda_i\mathbf{v}_i\mathbf{v}_i^T$ is the eigenvalue decomposition of $\mathbf{X}_0 \in \mathbb{S}^n$.

Separating a Point from a Convex Set

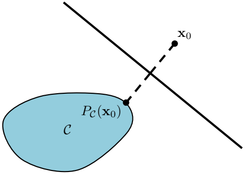
Suppose $\mathcal{C}$ is a closed convex set. Then, we have the following.

- If $\mathbf{x}_0 \in \mathcal{C}$, then $\mathrm{dist}(\mathbf{x}_0, \mathcal{C}) = 0$ with an optimal point $\mathbf{x}_0$.
- If $\mathbf{x}_0 \notin \mathcal{C}$, then $\mathrm{dist}(\mathbf{x}_0, \mathcal{C}) > 0$, and the optimal value of the distance problem is positive. Furthermore, a dual optimal point provides a separating hyperplane between $\mathbf{x}_0$ and $\mathcal{C}$.

For the Euclidean norm, if $P_{\mathcal{C}}(\mathbf{x}_0)$ denotes the Euclidean projection of $\mathbf{x}_0$ on $\mathcal{C}$, where $\mathbf{x}_0 \notin \mathcal{C}$, then the hyperplane characterized by

$$(P_{\mathcal{C}}(\mathbf{x}_0) - \mathbf{x}_0)^T (\mathbf{x} - (1/2)(\mathbf{x}_0 + P_{\mathcal{C}}(\mathbf{x}_0))) = 0,$$

strictly separates $\mathbf{x}_0$ from $\mathcal{C}$.



For a general norm $||\cdot||$, a separating hyperplane can be obtained via Lagrange duality. From the dual problem, if $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}$ are dual feasible with a positive objective value, then we have

$$\boldsymbol{\mu}^T \mathbf{x}_0 - \boldsymbol{\mu}^T \mathbf{x} + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \boldsymbol{\nu}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) > 0, \text{ for all } \mathbf{x},$$

and so $\boldsymbol{\mu}^T \mathbf{x}_0 > \boldsymbol{\mu}^T \mathbf{x}$ for $\mathbf{x} \in \mathcal{C}$. Thus, $\boldsymbol{\mu}$ defines a strictly separating hyperplane in this case.

Expressing the Distance Between Two Sets

The distance between two sets $\mathcal{C}$ and $\mathcal{D}$, with respect to the norm $||\cdot||$, is defined as

$$\operatorname{dist}(\mathcal{C}, \mathcal{D}) \triangleq \inf\{||\mathbf{x} - \mathbf{y}|| : \mathbf{x} \in \mathcal{C} \, , \; \mathbf{y} \in \mathcal{D}\} \; .$$

Properties:

- The sets $\mathcal{C}$ and $\mathcal{D}$ do not intersect if $\operatorname{dist}(\mathcal{C}, \mathcal{D}) > 0$.
- The sets $\mathcal{C}$ and $\mathcal{D}$ intersect if $\operatorname{dist}(\mathcal{C}, \mathcal{D}) = 0$ and the infimum is attained.
- The distance between sets can be expressed as the distance between a point and a set.

$$\operatorname{dist}(\mathcal{C}, \mathcal{D}) = \operatorname{dist}(\mathbf{0}, \mathcal{C} - \mathcal{D}) = \operatorname{dist}(\mathbf{0}, \mathcal{D} - \mathcal{C}) \; .$$

Distance and Separation Via Indicator and Support Functions:

The problem of finding the distance between sets can be posed in terms of indicator functions. The dual of this problem can be expressed using support functions.

Primal: ***Dual:***

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x} - \mathbf{y}|| \\ \text{subject to} & I_{\mathcal{C}}(\mathbf{x}) \le 0 \\ & I_{\mathcal{D}}(\mathbf{y}) \le 0 \end{array} \iff \begin{array}{ll} \text{minimize} & ||\mathbf{w}|| \\ \text{subject to} & I_{\mathcal{C}}(\mathbf{x}) \le 0 \\ & I_{\mathcal{D}}(\mathbf{y}) \le 0 \\ & \mathbf{x} - \mathbf{y} = \mathbf{w} \end{array} \quad \begin{array}{ll} \text{maximize} & -S_{\mathcal{C}}(-\mathbf{z}) - S_{\mathcal{D}}(\mathbf{z}) \\ \text{subject to} & ||\mathbf{z}||_{\circledast} \le 1 \end{array} \; .$$

If $\mathcal{C}$ and $\mathcal{D}$ are convex sets, then the primal problem is convex and (assuming strong duality holds) the dual problem can be used to find a separating hyperplane. Specifically, if $\mathbf{z}$ is dual feasible with a positive objective value, then $S_{\mathcal{D}}(\mathbf{z}) < -S_{\mathcal{C}}(-\mathbf{z})$, which means that

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

In other words, $\mathbf{z}$ defines a hyperplane that *strictly* separates $\mathcal{C}$ and $\mathcal{D}$.

Computing the Distance Between Convex Sets

Suppose $\mathcal{C}$ and $\mathcal{D}$ are described by two sets of convex inequalities

$$\mathcal{C} = \{\mathbf{x} : f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m\}\ ,\ \mathcal{D} = \{\mathbf{x} : g_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, p\}\ .$$

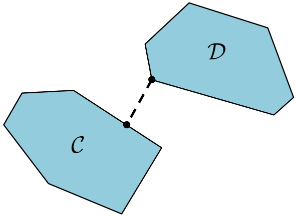
This can admit affine equalities by including two convex inequalities for each affine equality. Then, we can find $\mathrm{dist}(\mathcal{C}, \mathcal{D})$ by solving the convex optimization problem

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

Example: *Euclidean distance between polyhedra*

$$\mathcal{C} = \{\mathbf{x} : \mathbf{A}_1\mathbf{x} \preceq \mathbf{b}_1\}\ ,\ \mathcal{D} = \{\mathbf{x} : \mathbf{A}_1\mathbf{x} \preceq \mathbf{b}_1\}\ .$$

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x} - \mathbf{y}||_2 \\ \text{subject to} & \mathbf{A}_1\mathbf{x} \preceq \mathbf{b}_1 \\ & \mathbf{A}_2\mathbf{x} \preceq \mathbf{b}_2 \end{array}\ .$$



Squaring the objective, we obtain an equivalent QP.

Separating Convex Sets

The dual of the problem of finding the distance between convex sets has an interesting interpretation in terms of separating hyperplanes. Starting from the equivalent form of the primal problem

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

we obtain the dual problem

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

If $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are dual feasible with a positive objective value (i.e., $\text{dist}(\mathcal{C}, \mathcal{D}) > 0$), then

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

for all $\mathbf{x}$ and $\mathbf{y}$. In particular, for $\mathbf{x} \in \mathcal{C}$ and $\mathbf{y} \in \mathcal{D}$, we have $\mathbf{z}^T \mathbf{x} > \mathbf{z}^T \mathbf{y}$, and so $\mathbf{z}$ defines a hyperplane that *strictly* separates $\mathcal{C}$ and $\mathcal{D}$.

Configurations and the Gram Matrix

Consider a set of vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^n$ with known Euclidean lengths

$$l_1 = ||\mathbf{a}_1||_2 \ , \ \ldots \ , \ l_n = ||\mathbf{a}_n||_2 \ .$$

We refer to this set as a *configuration* in general and as a *basis* when they are linearly independent.

The lengths, distances, correlation coefficients, and angles between vectors of the configuration can be expressed in terms of the *Gram matrix* $\mathbf{G} \in \mathbb{R}^{n \times n}$ given by

$$\mathbf{G} \triangleq \mathbf{A}^T \mathbf{A} \, , \ \mathbf{A} \triangleq \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} .$$

- *Length:* (of $\mathbf{a}_i$)

$$l_i \triangleq ||\mathbf{a}_i||_2 = (G_{i,i})^{1/2} \ , \ G_{i,i} = l_i^2 \ , \ i = 1, \ldots, n \, .$$

- *Distance:* (between $\mathbf{a}_i$ and $\mathbf{a}_j$)

$$d_{i,j} \triangleq ||\mathbf{a}_i - \mathbf{a}_j||_2 = \left(l_i^2 + l_j^2 - 2G_{i,j} \right)^{1/2} \ , \ G_{i,j} = \frac{l_i^2 + l_j^2 - d_{i,j}^2}{2} \ , \ i, j = 1, \ldots, n \, .$$

- *Correlation coefficient:* (between $\mathbf{a}_i$ and $\mathbf{a}_j$)

$$\rho_{i,j} \triangleq \frac{\mathbf{a}_i^T \mathbf{a}_j}{||\mathbf{a}_i||_2 \, ||\mathbf{a}_j||_2} = \frac{G_{i,j}}{l_i l_j} \ , \ G_{i,j} = l_i l_j \rho_{i,j} \ , \ i, j = 1, \ldots, n \, .$$

- *Angle:* (between $\mathbf{a}_i$ and $\mathbf{a}_j$)

$$\theta_{i,j} \triangleq \arccos \rho_{i,j} = \arccos(G_{i,j} / \, (l_i l_j)) \ , \ G_{i,j} = l_i l_j \cos \theta_{i,j} \ , \ i, j = 1, \ldots, n \, .$$

These quantities are invariant under orthogonal transformations. Namely, if $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is orthogonal, then the configuration $\mathbf{Q}\mathbf{a}_1, \ldots, \mathbf{Q}\mathbf{a}_n$ has the same Gram matrix as the original one.

Realizability of the Gram Matrix

Euclidean distance problems can be posed in terms of the Gram matrix $\mathbf{G} = \mathbf{A}^T \mathbf{A}$.

- From linear algebra, a matrix $\mathbf{G} \in \mathbb{S}^n$ is the Gram matrix of a configuration $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ (i.e., $\mathbf{G} = \mathbf{A}^T \mathbf{A}$ with $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$) if and only if $\mathbf{G} \succeq \mathbf{0}$.
- When $\mathbf{G} \succeq \mathbf{0}$, we can construct a configuration with Gram matrix $\mathbf{G}$ by taking $\mathbf{A} = \mathbf{G}^{1/2}$, using an eigenvalue decomposition of $\mathbf{G}$.
- When $\mathbf{G} \succ \mathbf{0}$, a configuration can be obtained via the Cholesky decomposition $\mathbf{G} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is a lower triangular matrix, by taking $\mathbf{A} = \mathbf{L}^T$.
- *All* configurations with Gram matrix $\mathbf{G}$ can be constructed by orthogonal transformation, given any one solution $\mathbf{A}$. Namely, if $\widetilde{\mathbf{A}}^T \widetilde{\mathbf{A}} = \mathbf{G}$ is *any* solution, then $\widetilde{\mathbf{A}} = \mathbf{Q}\mathbf{A}$ for some orthogonal matrix $\mathbf{Q}$.

A set of lengths, distances, correlation coefficients, and/or angles is said to be *realizable*, i.e., those of some configuration, if and only if we have

$$\mathbf{G} \succeq \mathbf{0}\,,\ G_{i,i} = l_i\,,\ i = 1, \ldots, n\,.$$

In other words, $\mathbf{G}$ is realizable for a set of length, distance, correlation coefficient, and/or angle constraints if and only if the above condition is satisfied. This realizability constraint can be used to pose several geometric problems as convex optimization problems, with $\mathbf{G} \in \mathbb{S}^n$ as the optimization variable.

Configuration Constraints

- **Angle constraints:** A lower and upper bound on an angle, $\alpha \leq \theta_{i,j} \leq \beta$, is equivalent to the affine inequality constraints
$$l_i l_j \cos\alpha \geq G_{i,j} \geq l_i l_j \cos\beta \,.$$
- **Distance constraints:** A lower and upper bound on a distance, $d_{\min} \leq d_{i,j} \leq d_{\max}$, is equivalent to the affine inequality constraints
$$d_{\min}^2 \leq l_i^2 + l_j^2 - 2G_{i,j} \leq d_{\max}^2 \,.$$
- **Singular value constraints:** The singular values $\sigma_1 \geq \cdots \geq \sigma_n$ of $\mathbf{A}$ are the square roots of the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$ of $\mathbf{G}$. Thus, σ_1^2 is a *convex* function of $\mathbf{G}$, whereas σ_n^2 is a *concave* function of $\mathbf{G}$. So, the following constraints are convex.
$$\sigma_1^2(\mathbf{G}) \leq \sigma_{\max}^2 \,,\ \sigma_n^2(\mathbf{G}) \geq \sigma_{\min}^2 \,.$$
- **Condition number constraints:** The condition number $\kappa \triangleq \sigma_1/\sigma_n$ of $\mathbf{A}$ is a quasiconvex function of $\mathbf{G}$. Hence, the following constraint is convex.
$$\kappa(\mathbf{G}) \leq \kappa_{\max} \,.$$
Instead of this, we can minimize $\kappa(\mathbf{G})$ over all configurations satisfying other convex geometric constraints, by quasiconvex optimization.

Dual Basis and Volume Objectives

Dual Basis:

When $\mathbf{G} \succ \mathbf{0}$, the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ form a basis for $\mathbb{R}^n$. The associated *dual basis* $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is one that satisfies $\mathbf{b}_i^T \mathbf{a}_j = \delta_{i,j}$, where $\delta_{i,j}$ is the *Kronecker delta* function.

- The dual basis vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are simply the rows of $\mathbf{A}^{-1}$. As such, the Gram matrix associated with the dual basis is $\mathbf{G}^{-1}$.
- The squared lengths of the dual basis vectors are given by

$$||\mathbf{b}_i||_2^2 = \mathbf{e}_i^T \mathbf{G}^{-1} \mathbf{e}_i \ , \ i = 1, \ldots, n \, ,$$

  and are convex functions of $\mathbf{G}$. So, they can be minimized or upper bound constrained. Also, $\sum_{i=1}^{n} ||\mathbf{b}_i||_2^2 = \mathrm{tr}\left(\mathbf{G}^{-1}\right)$, is a measure of a well conditioned basis and is convex in $\mathbf{G}$.

Volume Under Linear Transformation:

The volume of a set $\mathcal{C} \subseteq \mathbb{R}^n$ under a linear transformation $\mathbf{A}$ (i.e., $\mathcal{C}_{\mathrm{lt}} \triangleq \{\mathbf{A}\mathbf{u} : \mathbf{u} \in \mathcal{C}\}$) is

$$\mathrm{vol}(\mathcal{C}_{\mathrm{lt}}) = \mathrm{vol}(\mathcal{C}) \left(\det\left(\mathbf{A}^T \mathbf{A}\right)\right)^{1/2} = \mathrm{vol}(\mathcal{C}) \left(\det \mathbf{G}\right)^{1/2} \, .$$

- **Example:** *Ellipsoid* ($\mathcal{E} \triangleq \left\{\mathbf{A}\mathbf{u} : ||\mathbf{u}||_2 \leq 1\right\}$)

$$\mathrm{vol}(\mathcal{E}) = \gamma_{\mathrm{ub}} \left(\det \mathbf{G}\right)^{1/2} \ , \ \gamma_{\mathrm{ub}} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} \, .$$

- **Example:** *Simplex* ($\mathcal{S} \triangleq \{\mathbf{A}\mathbf{u} : \mathbf{u} \in \mathrm{conv}\{\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_n\}\}$)

$$\mathrm{vol}(\mathcal{S}) = \gamma_{\mathrm{ub}} \left(\det \mathbf{G}\right)^{1/2} \ , \ \gamma_{\mathrm{us}} = 1/n! \, .$$

- The volume of the warped set $\mathcal{C}_{\mathrm{lt}}$ can be maximized by equivalently maximizing $\log\det \mathbf{G}$.

Bounding Correlations for Angle Only Problems

For problems which only involve angles (or equivalently correlation coefficients), many constraints on the Gram matrix $\mathbf{G}$ can be simplified as they will not depend on the actual lengths of the vectors of the configuration. This follows from the fact that

$$\mathbf{G} = \mathrm{diag}(\mathbf{l})\, \mathbf{C}\, \mathrm{diag}(\mathbf{l}) \ ,$$

where $\mathbf{l}$ is the vector of lengths given by

$$\mathbf{l} = \begin{bmatrix} l_1 & \cdots & l_n \end{bmatrix}^T \ ,$$

and $\mathbf{C}$ is the *correlation matrix* given by

$$C_{i,j} = \rho_{i,j} = \cos\theta_{i,j} \ , \ i,j = 1,\ldots,n \,.$$

- If $\mathbf{G} \succeq \mathbf{0}$ for any $\mathbf{l} \succ \mathbf{0}$, then $\mathbf{G} \succeq \mathbf{0}$ for all $\mathbf{l} \succ \mathbf{0}$, which occurs if and only if $\mathbf{C} \succeq \mathbf{0}$.
- A set of angles $\theta_{i,j} \in [0,\pi]$ for $i,j = 1,\ldots,n$ is realizable if and only if $\mathbf{C} \succeq \mathbf{0}$, which is an LMI in $\rho_{i,j}$ for $i,j = 1,\ldots,n$.
- As an example, given lower and upper bounds on some angles, we can find the minimum and maximum possible value of some other angle, over all configurations, by solving two SDPs.

Euclidean Distance Problems

In a *Euclidean distance problem*, the only concern is the set of distances between vectors $d_{i,j}$.

- Distances are invariant under orthogonal transformations, i.e., the configuration $\widetilde{\mathbf{A}} = \mathbf{QA}$ has the same distances as the original one $\mathbf{A}$, for any orthogonal matrix $\mathbf{Q}$.
- Distances are invariant under translation, i.e., the configuration $\widetilde{\mathbf{A}} = \mathbf{A} - \mathbf{b1}^T$ has the same distances as the original one $\mathbf{A}$, for any $\mathbf{b} \in \mathbb{R}^n$. If $\mathbf{b}$ is the *centroid* given by
$$\mathbf{b} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{a}_i = (1/n)\,\mathbf{A1}\,,$$
then we have $\widetilde{\mathbf{A}}\mathbf{1} = \mathbf{0}$. Hence, without loss of generality, we can assume that $\mathbf{A1} = \mathbf{0}$, i.e., $\mathbf{A}$ is orthogonal to $\mathbf{1}$ and lies solely in $\mathbf{1}^{\perp}$.
- Defining $\mathbf{z} \in \mathbb{R}^n$ as $z_i = l_i^2$, and $\mathbf{D} \in \mathbb{S}^n$ by $D_{i,j} = d_{i,j}^2$ (with $D_{i,i} = 0$), the condition $\mathbf{G} \succeq \mathbf{0}$ for some set of lengths is equivalent to
$$\mathbf{G} = \left(\mathbf{z1}^T + \mathbf{1z}^T - \mathbf{D}\right)/2 \succeq \mathbf{0} \text{ for some } \mathbf{z} \succeq \mathbf{0}\,.$$
A matrix $\mathbf{D} \in \mathbb{S}^n$ with nonnegative elements and zero diagonal that satisfies the above relation is called a *Euclidean distance matrix*.
- Assuming $\mathbf{A1} = \mathbf{0}$, we can show that $\mathbf{G} \succeq \mathbf{0}$ if and only if the following conditions on $\mathbf{D}$ hold:
$$D_{i,i} = 0\,,\ i = 1,\ldots,n\,,\ D_{i,j} \geq 0\,,\ i,j = 1,\ldots,n\,,$$
$$\left(\mathbf{I} - (1/n)\,\mathbf{11}^T\right)\mathbf{D}\left(\mathbf{I} - (1/n)\,\mathbf{11}^T\right) \preceq \mathbf{0}\,.$$
Therefore, any Euclidean distance problem that is convex in the squared distances can be expressed as convex problem with variable $\mathbf{D} \in \mathbb{S}^n$.

The Löwner-John Ellipsoid

Let $\mathcal{C} \in \mathbb{R}^n$ be some bounded set with nonempty interior. Suppose we wish to find the minimum volume ellipsoid that covers $\mathcal{C}$, which is called the *Löwner-John ellipsoid*. To characterize this ellipsoid, it will be convenient to use the following parametrization:

$$\mathcal{E} = \left\{ \mathbf{v} : ||\mathbf{A}\mathbf{v} + \mathbf{b}||_2 \leq 1 \right\} ,$$

where we can assume without loss of generality that $\mathbf{A} \in \mathbb{S}^n_{++}$. Then, the Löwner-John ellipsoid $\mathcal{E}_{\mathrm{lj}}$ can be found by solving the following problem.

Minimum Volume Covering Ellipsoid Problem

$$\begin{array}{ll} \text{minimize} & \log\det \mathbf{A}^{-1} \\ \text{subject to} & \sup\limits_{\mathbf{v} \in \mathcal{C}} ||\mathbf{A}\mathbf{v} + \mathbf{b}||_2 \leq 1 \end{array} ,$$

with variables $\mathbf{A} \in \mathbb{S}^n$ and $\mathbf{b} \in \mathbb{R}^n$, and there is an implicit constraint that $\mathbf{A} \succ \mathbf{0}$.

- This is a convex optimization problem as the objective and constraint functions are both convex in $\mathbf{A}$ and $\mathbf{b}$.
- Evaluating the constraint function involves solving a convex maximization problem and is only tractable in certain special cases.

Examples

Minimum volume ellipsoid covering a finite set:

Suppose $\mathcal{C} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$. The Löwner-John ellipsoid can be found by solving the problem:

$$\begin{array}{ll} \text{minimize} & \log\det \mathbf{A}^{-1} \\ \text{subject to} & ||\mathbf{A}\mathbf{x}_i + \mathbf{b}||_2 \leq 1\,,\ i = 1, \ldots, m \end{array} .$$

- Norm constraint $||\mathbf{A}\mathbf{x}_i + \mathbf{b}||_2 \leq 1$ equivalent to convex quadratic one $||\mathbf{A}\mathbf{x}_i + \mathbf{b}||_2^2 \leq 1$.
- As an ellipsoid covers any set $\mathcal{C}$ if and only if it covers its convex hull, this problem yields the Löwner-John ellipsoid for the polyhedron $\mathcal{P} = \mathrm{conv}\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$.

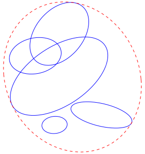
Minimum volume ellipsoid covering a union of ellipsoids:

Suppose $\mathcal{C} = \bigcup_{i=1}^{m} \mathcal{E}_i$, where $\mathcal{E}_1, \ldots, \mathcal{E}_m$ are ellipsoids of the following form.

$$\mathcal{E}_i = \left\{ \mathbf{x} : \mathbf{x}^T \mathbf{A}_i \mathbf{x} + 2\mathbf{b}_i^T \mathbf{x} + c_i \leq 0 \right\}\ ,\ \mathbf{A}_i \in \mathbb{S}_{++}^n\ ,\ i = 1, \ldots, m\,.$$

The Löwner-John ellipsoid can be found by solving the following convex problem.

$$\begin{array}{ll} \text{minimize} & \log\det \mathbf{A}^{-1} \\ \text{subject to} & \tau_1 \geq 0\,,\ \ldots\,,\ \tau_m \geq 0 \\ & \begin{bmatrix} \mathbf{A}^2 - \tau_i \mathbf{A}_i & \widetilde{\mathbf{b}} - \tau_i \mathbf{b}_i & \mathbf{0} \\ \left(\widetilde{\mathbf{b}} - \tau_i \mathbf{b}_i\right)^T & -1 - \tau_i c_i & \widetilde{\mathbf{b}}^T \\ \mathbf{0} & \widetilde{\mathbf{b}} & -\mathbf{A}^2 \end{bmatrix} \preceq \mathbf{0}\,,\ i = 1, \ldots, m \end{array} ,$$



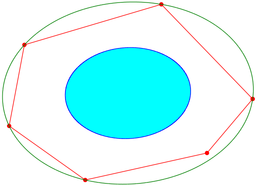
with variables $\mathbf{A}^2 \in \mathbb{S}^n$, $\widetilde{\mathbf{b}} = \mathbf{A}\mathbf{b} \in \mathbb{R}^n$, and $\tau_i \in \mathbb{R}$ for $i = 1, \ldots, m$.

Efficiency of Löwner-John Ellipsoidal Approximation

Let $\mathcal{E}_{\text{lj}}$ be the Löwner-John ellipsoid of a convex set $\mathcal{C} \subseteq \mathbb{R}^n$, which is bounded with nonempty interior, and whose center is $\mathbf{x}_0$. Then, it can be shown that the Löwner-John ellipsoid, shrunk by a factor of n, lies inside $\mathcal{C}$. Namely, we have

$$\mathbf{x}_0 + (1/n)\left(\mathcal{E}_{\text{lj}} - \mathbf{x}_0\right) \subseteq \mathcal{C} \subseteq \mathcal{E}_{\text{lj}} \,.$$

Example: *Polyhedron in* $\mathbb{R}^2$



Further Extensions:

- If $\mathcal{C}$ is *symmetric*, then the factor of n can be improved to $\sqrt{n}$. Namely, we have
$$\mathbf{x}_0 + \left(1/\sqrt{n}\right)\left(\mathcal{E}_{\text{lj}} - \mathbf{x}_0\right) \subseteq \mathcal{C} \subseteq \mathcal{E}_{\text{lj}} \,.$$
- This allows us to approximate any norm by a quadratic norm. In other words, if $||\cdot||$ is any norm, $\mathcal{C} \triangleq \{\mathbf{x} : ||\mathbf{x}|| \leq 1\}$ is its unit ball, and $\mathcal{E}_{\text{lj}} = \left\{\mathbf{x} : \mathbf{x}^T \mathbf{A} \mathbf{x} \leq 1\right\}$ (where $\mathbf{A} \in \mathbb{S}^n_{++}$) is the Löwner-John ellipsoid of $\mathcal{C}$, then we have
$$||\mathbf{z}||_{\text{lj}} \leq ||\mathbf{z}|| \leq \sqrt{n}\, ||\mathbf{z}||_{\text{lj}} \,, \text{ where } ||\mathbf{z}||_{\text{lj}} \triangleq \left(\mathbf{z}^T \mathbf{A} \mathbf{z}\right)^{1/2} \,.$$

The Maximum Volume Inscribed Ellipsoid

Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set that is bounded with nonempty interior. To find the ellipsoid of maximum volume that lies inside $\mathcal{C}$, we will use the parametrization

$$\mathcal{E} = \left\{ \mathbf{B}\mathbf{u} + \mathbf{d} : ||\mathbf{u}||_2 \leq 1 \right\} ,$$

where we can assume without loss of generality that $\mathbf{B} \in \mathbb{S}^n_{++}$. Then, the maximum volume inscribed ellipsoid $\mathcal{E}_{\text{mvie}}$ can be obtained by solving the following convex optimization problem.

Maximum Volume Inscribed Ellipsoid Problem

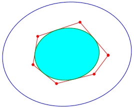
$$\begin{array}{ll} \text{maximize} & \log\det\mathbf{B} \\ \text{subject to} & \sup\limits_{||\mathbf{u}||_2 \leq 1} I_{\mathcal{C}}(\mathbf{B}\mathbf{u} + \mathbf{d}) \leq 0 \end{array} ,$$

with variables $\mathbf{B} \in \mathbb{S}^n$ and $\mathbf{d} \in \mathbb{R}^n$, and there is an implicit constraint that $\mathbf{B} \succ \mathbf{0}$.

Efficiency of Ellipsoidal Inner Approximations:

- The ellipsoid $\mathcal{E}_{\text{mvie}}$, expanded by a factor of n about $\mathbf{x}_0$, the center of $\mathcal{C}$, lies inside $\mathcal{C}$. Thus,
$$\mathcal{E}_{\text{mvie}} \subseteq \mathcal{C} \subseteq \mathbf{x}_0 + n\left(\mathcal{E}_{\text{mvie}} - \mathbf{x}_0\right) .$$
- If $\mathcal{C}$ is symmetric, the factor improves to $\sqrt{n}$. Hence, $\mathcal{E}_{\text{mvie}} \subseteq \mathcal{C} \subseteq \mathbf{x}_0 + \sqrt{n}\left(\mathcal{E}_{\text{mvie}} - \mathbf{x}_0\right)$.

Example: *Polyhedron in* $\mathbb{R}^2$



Examples

Maximum volume ellipsoid in a polyhedron:

Suppose $\mathcal{C} = \left\{\mathbf{x} : \mathbf{a}_i^T \mathbf{x} \leq b_i \, , \; i = 1, \ldots, m\right\}$. Then, we have

$$\sup_{||\mathbf{u}||_2 \leq 1} I_{\mathcal{C}}(\mathbf{B}\mathbf{u} + \mathbf{d}) \leq 0 \iff ||\mathbf{B}\mathbf{a}_i||_2 + \mathbf{a}_i^T \mathbf{d} \leq b_i \leq b_i \, , \; i = 1, \ldots, m \, .$$

The maximum volume ellipsoid inscribed in $\mathcal{C}$ can then be found by solving the convex problem

$$\begin{array}{ll} \text{minimize} & \log\det \mathbf{B}^{-1} \\ \text{subject to} & ||\mathbf{B}\mathbf{a}_i||_2 + \mathbf{a}_i^T \mathbf{d} \leq b_i \, , \; i = 1, \ldots, m \end{array} .$$

Maximum volume ellipsoid in an intersection of ellipsoids:

Suppose $\mathcal{C} = \bigcap_{i=1}^m \mathcal{E}_i$, where $\mathcal{E}_1, \ldots, \mathcal{E}_m$ are ellipsoids of the following form.

$$\mathcal{E}_i = \left\{\mathbf{x} : \mathbf{x}^T \mathbf{A}_i \mathbf{x} + 2\mathbf{b}_i^T \mathbf{x} + c_i \leq 0\right\} \, , \; \mathbf{A}_i \in \mathbb{S}_{++}^n \, , \; i = 1, \ldots, m \, .$$

The maximum volume inscribed ellipsoid can be found by solving the following convex problem.

$$\begin{array}{ll} \text{minimize} & \log\det \mathbf{B}^{-1} \\ \text{subject to} & \begin{bmatrix} -\lambda_i - c_i + \mathbf{b}_i^T \mathbf{A}_i^{-1} \mathbf{b}_i & \mathbf{0} & \left(\mathbf{d} + \mathbf{A}_i^{-1} \mathbf{b}_i\right)^T \\ \mathbf{0} & \lambda_i \mathbf{I} & \mathbf{B} \\ \mathbf{d} + \mathbf{A}_i^{-1} \mathbf{b}_i & \mathbf{B} & \mathbf{A}_i^{-1} \end{bmatrix} \succeq \mathbf{0} \, , \; i = 1, \ldots, m \end{array} ,$$

with variables $\mathbf{B} \in \mathbb{S}^n$, $\mathbf{d} \in \mathbb{R}^n$, and $\boldsymbol{\lambda} \in \mathbb{R}^m$.

Affine Invariance of Extremal Volume Ellipsoids

Affine Invariance Property of Extremal Volume Ellipsoids:

Let $\mathcal{C}, \mathcal{D} \subseteq \mathbb{R}^n$ each be a bounded set with nonempty interior, where $\mathcal{D}$ is additionally convex. Also, let $\mathbf{T} \in \mathbb{R}^{n \times n}$ be any nonsingular matrix. Then, we have the following.

- If $\mathcal{E}_{\text{lj}}$ is the Löwner-John ellipsoid of $\mathcal{C}$, then the Löwner-John ellipsoid of $\mathbf{T}\mathcal{C}$ is $\mathbf{T}\mathcal{E}_{\text{lj}}$.
- If $\mathcal{E}_{\text{mvie}}$ is the maximum volume inscribed ellipsoid of $\mathcal{D}$, then the maximum volume inscribed ellipsoid of $\mathbf{T}\mathcal{D}$ is $\mathbf{T}\mathcal{E}_{\text{mvie}}$.

Proof of Affine Invariance:

- Let $\mathcal{E}$ be any ellipsoid that covers $\mathcal{C}$ (or is inscribed in $\mathcal{D}$). Then the ellipsoid $\mathbf{T}\mathcal{E}$ covers $\mathbf{T}\mathcal{C}$ (or is inscribed in $\mathbf{T}\mathcal{D}$).
- Conversely, every ellipsoid that covers $\mathbf{T}\mathcal{C}$ (or is inscribed in $\mathbf{T}\mathcal{D}$) is of the form $\mathbf{T}\mathcal{E}$, where $\mathcal{E}$ is an ellipsoid that covers $\mathcal{C}$ (or is inscribed in $\mathcal{D}$).
- Therefore, the relation $\widetilde{\mathcal{E}} = \mathbf{T}\mathcal{E}$ gives a one-to-one correspondence between ellipsoids covering $\mathbf{T}\mathcal{C}$ (or inscribed in $\mathbf{T}\mathcal{D}$) and ellipsoids covering $\mathcal{C}$ (or inscribed in $\mathcal{D}$).
- Moreover, the volumes of the corresponding ellipsoids are related by the ratio $|\det \mathbf{T}|$.
- Hence, if $\mathcal{E}$ has minimum volume among ellipsoids covering $\mathcal{C}$, then $\mathbf{T}\mathcal{E}$ has minimum volume among ellipsoids covering $\mathbf{T}\mathcal{C}$. Similarly, if $\mathcal{E}$ has maximum volume among ellipsoids inscribed in $\mathcal{D}$, then $\mathbf{T}\mathcal{E}$ has maximum volume among ellipsoids inscribed in $\mathbf{T}\mathcal{D}$.

Depth and the Chebyshev Center

Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a bounded set with nonempty interior, and $\mathbf{x} \in \mathcal{C}$. The *depth* of $\mathbf{x}$ is

$$\mathrm{depth}(\mathbf{x}, \mathcal{C}) \triangleq \mathrm{dist}(\mathbf{x}, \mathbb{R}^n \setminus \mathcal{C}) \ .$$

Intuitively, the depth gives the radius of the largest ball (for the specific norm), centered at $\mathbf{x}$, that lies in $\mathcal{C}$. A *Chebyshev center* of $\mathcal{C}$ is any point of maximum depth in $\mathcal{C}$:

$$\mathbf{x}_{\mathrm{cheb}}(\mathcal{C}) = \operatorname*{argmax}_{\mathbf{x}} \mathrm{depth}(\mathbf{x}, \mathcal{C}) = \operatorname*{argmax}_{\mathbf{x}} \mathrm{dist}(\mathbf{x}, \mathbb{R}^n \setminus \mathcal{C}) \ .$$

Chebyshev Center of a Convex Set:

When $\mathcal{C}$ is convex, the depth is a concave function for $\mathbf{x} \in \mathcal{C}$, so computing the Chebyshev center is a convex optimization problem. More specifically, suppose $\mathcal{C}$ is defined by a set of convex inequalities:

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

Then, a Chebyshev center can be found by solving the following problem.

$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & g_i(\mathbf{x}, R) \leq 0 \, , \ i = 1, \ldots, m \end{array} , \text{ where } g_i(\mathbf{x}, R) \triangleq \sup_{||\mathbf{u}|| \leq 1} f_i(\mathbf{x} + R\mathbf{u}) \ .$$

- Each function g_i is convex, since it is the pointwise supremum of a family of convex functions of $\mathbf{x}$ and R.
- Evaluating g_i entails solving a convex maximization problem, which may be hard.

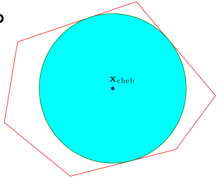
Examples

Chebyshev center of a polyhedron:

Suppose $\mathcal{C} = \left\{ \mathbf{x} : \mathbf{a}_i^T \mathbf{x} \leq b_i \, , \; i = 1, \ldots, m \right\}$. Then, we have

$$g_i(\mathbf{x}, R) = \sup_{||\mathbf{u}|| \leq 1} \mathbf{a}_i^T \left(\mathbf{x} + R\mathbf{u} \right) = \mathbf{a}_i^T \mathbf{x} + R \, ||\mathbf{a}_i||_{\circledast} - b_i \, ,$$

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



$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & \mathbf{a}_i^T \mathbf{x} + R \, ||\mathbf{a}_i||_{\circledast} \leq b_i \, , \; i = 1, \ldots, m \quad . \\ & R \geq 0 \end{array}$$

Euclidean Chebyshev center of intersection of ellipsoids:

Suppose $\mathcal{C}$ is an intersection of m ellipsoids, defined by quadratic inequalities,

$$\mathcal{C} = \left\{ \mathbf{x} : \mathbf{x}^T \mathbf{A}_i \mathbf{x} + 2\mathbf{b}_i^T \mathbf{x} + c_i \leq 0 \, , \; i = 1, \ldots, m \right\} \, , \text{ where } \mathbf{A}_i \in \mathbb{S}_{++}^n \, .$$

The Chebyshev centering problem can be shown to be the following.

$$\begin{array}{ll} \text{maximize} & R \\ \text{subject to} & \begin{bmatrix} -\lambda_i - c_i + \mathbf{b}_i^T \mathbf{A}_i^{-1} \mathbf{b}_i & \mathbf{0} & \left(\mathbf{x} + \mathbf{A}_i^{-1} \mathbf{b}_i \right)^T \\ \mathbf{0} & \lambda_i \mathbf{I} & R\mathbf{I} \\ \mathbf{x} + \mathbf{A}_i^{-1} \mathbf{b}_i & R\mathbf{I} & \mathbf{A}_i^{-1} \end{bmatrix} \succeq \mathbf{0} \, , \; i = 1, \ldots, m \quad , \end{array}$$

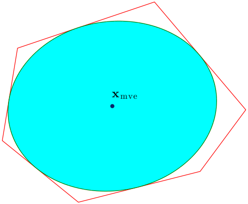
which is an SDP with variables $R \in \mathbb{R}$, $\boldsymbol{\lambda} \in \mathbb{R}^m$, and $\mathbf{x} \in \mathbb{R}^n$.

Maximum Volume Ellipsoid Center

- Recall that the Chebyshev center $\mathbf{x}_{\text{cheb}}$ of a set $\mathcal{C} \subseteq \mathbb{R}^n$ is the center of the largest ball that lies in $\mathcal{C}$.
- As an extension of this idea, we define the *maximum volume ellipsoid center* of $\mathcal{C}$, denoted $\mathbf{x}_{\text{mve}}$, as the center of the maximum volume ellipsoid that lies in $\mathcal{C}$.
- We can compute $\mathbf{x}_{\text{mve}}$ by finding the maximum volume inscribed ellipsoid for $\mathcal{C}$. To do this, we first solve the optimization problem

$$\begin{array}{ll} \text{maximize} & \log\det \mathbf{B} \\ \text{subject to} & \sup\limits_{||\mathbf{u}||_2 \leq 1} I_{\mathcal{C}}(\mathbf{B}\mathbf{u} + \mathbf{d}) \leq 0 \end{array},$$

with variables $\mathbf{B} \in \mathbb{S}^n$ and $\mathbf{d} \in \mathbb{R}^n$, and an implicit constraint that $\mathbf{B} \succ \mathbf{0}$. Then we simply set $\mathbf{x}_{\text{mve}} = \mathbf{d}^\star$.



Analytic Center of a Set of Inequalities

The *analytic center* $\mathbf{x}_{\rm ac}$ of a set of convex inequalities and affine equalities,

$$f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m\,,\ \mathbf{F}\mathbf{x} = \mathbf{g}\,,$$

is defined as an optimal point for the convex optimization problem

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

with variable $\mathbf{x} \in \mathbb{R}^n$ and implicit constraints $f_i(\mathbf{x}) < 0$ for $i = 1, \ldots, m$. The objective of this problem is called the *logarithmic barrier* associated with the set of inequalities.

- The analytic center $\mathbf{x}_{\rm ac}$ is the point that maximizes the product (or equivalently the geometric mean) of the slacks $-f_i(\mathbf{x})$, subject to $\mathbf{F}\mathbf{x} = \mathbf{g}$ and $f_i(\mathbf{x}) < 0$.
- This center is *not* a function of the feasible set $\mathcal{C} = \{\mathbf{x} : f_i(\mathbf{x}) < 0\ \ \forall\, i\,,\ \mathbf{F}\mathbf{x} = \mathbf{g}\}$: two sets of inequalities can describe the same set, but have different analytic centers.
- The analytic center $\mathbf{x}_{\rm ac}$ is independent of affine changes of coordinates.
- This center is also invariant under positive scalings of the inequality functions and reparametrization of the equality constraints. In other words, if $\alpha_1, \ldots, \alpha_m > 0$, and $\widetilde{\mathbf{F}}$ and $\widetilde{\mathbf{g}}$ are such that $\widetilde{\mathbf{F}}\mathbf{x} = \widetilde{\mathbf{g}}$ if and only if $\mathbf{F}\mathbf{x} = \mathbf{g}$, then the analytic center of

  $$\alpha_i f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m\,,\ \widetilde{\mathbf{F}}\mathbf{x} = \widetilde{\mathbf{g}}\,,$$

  is the same as the analytic center of

  $$f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m\,,\ \mathbf{F}\mathbf{x} = \mathbf{g}\,.$$

Ellipsoids from Analytic Center of Linear Inequalities

The analytic center $\mathbf{x}_{\mathrm{ac}}$ of a set of linear inequalities

$$\mathbf{a}_i^T \mathbf{x} \leq b_i \; , \; i = 1, \ldots, m \, ,$$

is the solution of the unconstrained minimization problem

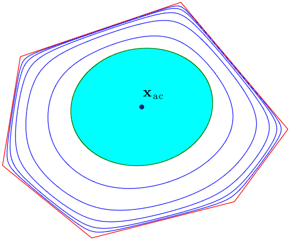
$$\text{minimize} \quad \phi(\mathbf{x}) \triangleq - \sum_{i=1}^{m} \log\left(b_i - \mathbf{a}_i^T \mathbf{x}\right) \; ,$$

with implicit constraint $b_i - \mathbf{a}_i^T \mathbf{x} > 0$, $i = 1, \ldots, m$. It can be shown that the analytic center here defines an inscribed and a covering ellipsoid, in terms of $\mathbf{H} \triangleq \nabla^2 \phi(\mathbf{x}_{\mathrm{ac}})$. Namely, we have

$$\mathcal{E}_{\mathrm{inner}} \subseteq \mathcal{P} \subseteq \mathcal{E}_{\mathrm{outer}} \, ,$$

where

$$\begin{aligned} \mathcal{P} &= \left\{ \mathbf{x} : \mathbf{a}_i^T \mathbf{x} \leq b_i \; , \; i = 1, \ldots, m \right\} \, , \\ \mathcal{E}_{\mathrm{inner}} &= \left\{ \mathbf{x} : (\mathbf{x} - \mathbf{x}_{\mathrm{ac}})^T \, \mathbf{H} \, (\mathbf{x} - \mathbf{x}_{\mathrm{ac}}) \leq 1 \right\} \, , \\ \mathcal{E}_{\mathrm{outer}} &= \left\{ \mathbf{x} : (\mathbf{x} - \mathbf{x}_{\mathrm{ac}})^T \, \mathbf{H} \, (\mathbf{x} - \mathbf{x}_{\mathrm{ac}}) \leq m \, (m-1) \right\} \, . \end{aligned}$$



Here, $\mathbf{H}$ is the Hessian of $\phi(\mathbf{x})$ at $\mathbf{x}_{\mathrm{ac}}$, given by

$$\mathbf{H} = \sum_{i=1}^{m} d_i^2 \mathbf{a}_i \mathbf{a}_i^T \; , \; d_i = \frac{1}{b_i - \mathbf{a}_i^T \mathbf{x}_{\mathrm{ac}}} \; , \; i = 1, \ldots, m \, .$$

Analytic Center of a Linear Matrix Inequality

The definition of an analytic center can be extended to sets described by generalized inequalities with respect to a cone $\mathcal{K}$, if we define a *generalized logarithm* for $\mathcal{K}$.

Generalized Logarithm for a Proper Cone:

Let $\mathcal{K} \subseteq \mathbb{R}^q$ is a proper cone. We say $\psi : \mathbb{R}^q \to \mathbb{R}$ is a generalized logarithm for $\mathcal{K}$ if

- The function ψ is concave, closed, twice continuously differentiable, $\operatorname{dom} \psi = \operatorname{int} \mathcal{K}$, and $\nabla^2 \psi(\mathbf{y}) \prec \mathbf{0}$ for $\mathbf{y} \in \operatorname{int} \mathcal{K}$.
- There is a constant $\theta > 0$ (the degree of ψ) such that for all $\mathbf{y} \succ_{\mathcal{K}} \mathbf{0}$ and all $s > 0$,

$$\psi(s\mathbf{y}) = \psi(\mathbf{y}) + \theta \log s \,.$$

In other words, ψ behaves like a logarithm along any ray in the cone $\mathcal{K}$.

For example, to find the analytic center of the LMI

$$x_1 \mathbf{A}_1 + \cdots x_n \mathbf{A}_n \preceq \mathbf{B} \,,$$

we can take $\psi(\mathbf{X}) = \log \det \mathbf{X}$ as a generalized logarithm for $\mathbb{S}_+^p$ with degree p, and solve the problem

$$\text{minimize} \quad -\log \det(\mathbf{B} - x_1 \mathbf{A}_1 - \cdots - x_n \mathbf{A}_n) \quad .$$