COMP6704 Lecture 5

Advanced Topics in Optimization

Convex Problems (2)

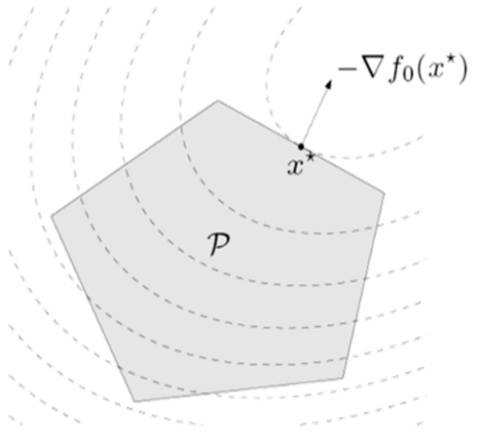
Fall, 2022

Instructor: WU, Xiao-Ming

Quadratic Programming

$$
\begin{array}{ll}
\underset{x}{\text{minimize}} & (1/2)\, x^T P x + q^T x + r \\
\text{subject to} & Gx \leq h \\
& Ax = b
\end{array}
$$

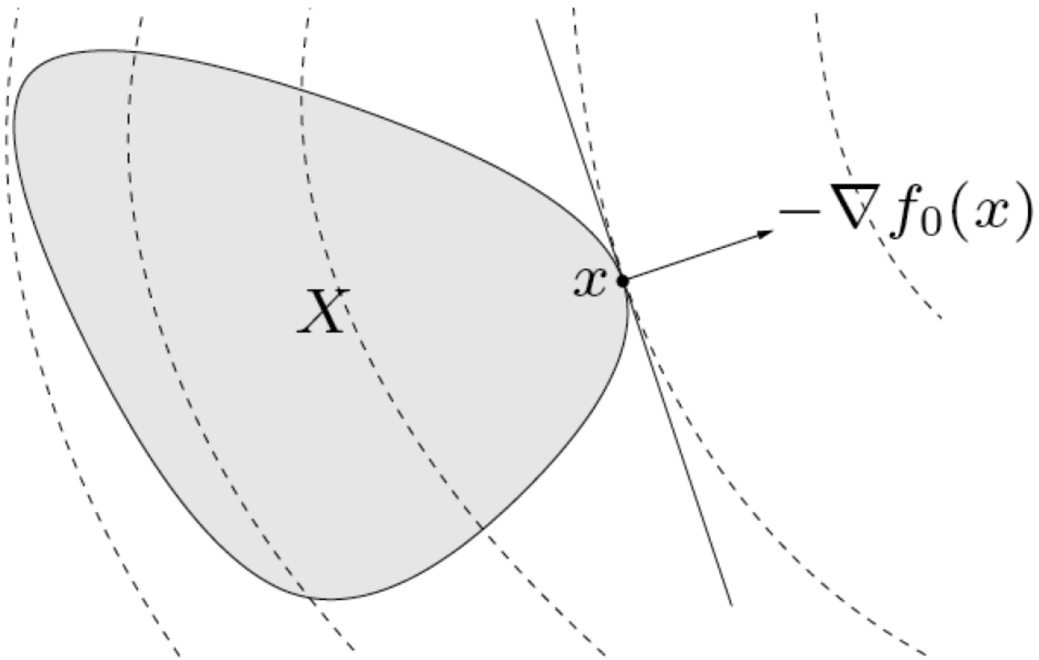
- Convex problem (assuming $P \in \mathbb{S}^n \succeq 0$): convex quadratic objective and affine constraint functions.
- Minimization of a convex quadratic function over a polyhedron:



Optimality Criterion for Differentiable f_0

Minimum Principle: A feasible point x is optimal if and only if

$$\nabla f_0(x)^T (y - x) \geq 0 \qquad \text{for all feasible } y$$



Exercise

Prove that $x^\star = (1, 1/2, -1)$ is optimal for the optimization problem

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & -1 \le x_i \le 1, \quad i = 1, 2, 3, \end{array}$$

where

$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \qquad q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13.0 \end{bmatrix}, \qquad r = 1.$$

Solution. We verify that $x^\star$ satisfies the optimality condition (4.21). The gradient of the objective function at $x^\star$ is

$$\nabla f_0(x^\star) = (-1, 0, 2).$$

Therefore the optimality condition is that

$$\nabla f_0(x^\star)^T (y - x) = -1(y_1 - 1) + 2(y_2 + 1) \ge 0$$

for all y satisfying $-1 \le y_i \le 1$, which is clearly true.

Least Squares in Matrix Form

$A \in \mathbf{R}^{k \times n}$ (with $k \geq n$), a_i^T are the rows of A, and the vector $x \in \mathbf{R}^n$

$$\text{minimize} \quad f_0(x) = \|Ax - b\|_2^2 = \textstyle\sum_{i=1}^k (a_i^T x - b_i)^2.$$

$$(A^T A)x = A^T b,$$

$$x = (A^T A)^{-1} A^T b$$

Quadratically Constrained QP (QCQP)

$$
\begin{array}{lll}
\underset{x}{\text{minimize}} & (1/2)\, x^T P_0 x + q_0^T x + r_0 & \\
\text{subject to} & (1/2)\, x^T P_i x + q_i^T x + r_i \leq 0 & i = 1, \ldots, m \\
& Ax = b &
\end{array}
$$

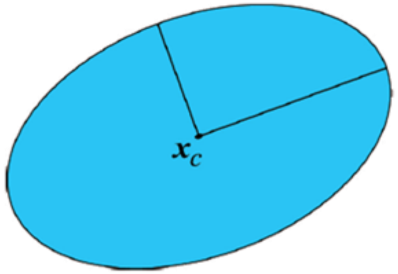
- Convex problem (assuming $P_i \in \mathbb{S}^n \succcurlyeq 0$): convex quadratic objective and constraint functions.

Ellipsoids

- Let A a square $(n \times n)$ matrix.
  - A is positive semidefinite when $x'Ax \geq 0$ for all $x \in \mathbb{R}^n$
  - A is positive definite when $x'Ax > 0$ for all $x \in \mathbb{R}^n, x \neq 0$.
- An ellipsoid is a set of the form

$$\varepsilon = \{x | (x - x_0)' P^{-1} (x - x_0) \leq 1\}$$

- where P is symmetric and positive definite



  - x_0 is the center of the ellipsoid ε
  - A ball $\{x | \, \|x - x_0\| \leq r\}$ is a special case of the ellipsoid $(P = r^2 I)$
  - Ellipsoids are convex

Exercise

4.23 *ℓ_4-norm approximation via QCQP.* Formulate the ℓ_4-norm approximation problem

$$\text{minimize} \quad \|Ax - b\|_4 = \left(\sum_{i=1}^m (a_i^T x - b_i)^4\right)^{1/4}$$

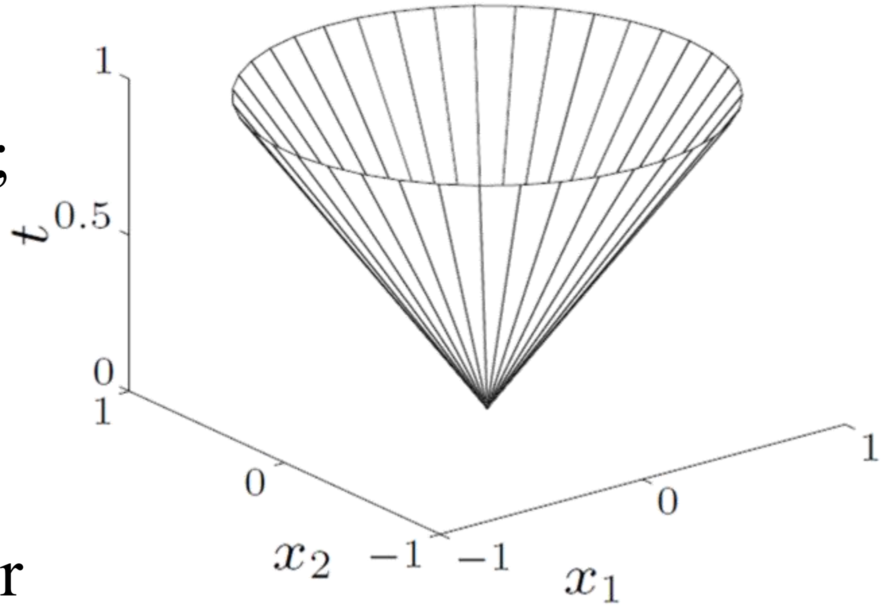
as a QCQP. The matrix $A \in \mathbf{R}^{m\times n}$ (with rows a_i^T) and the vector $b \in \mathbf{R}^m$ are given.

Solution.

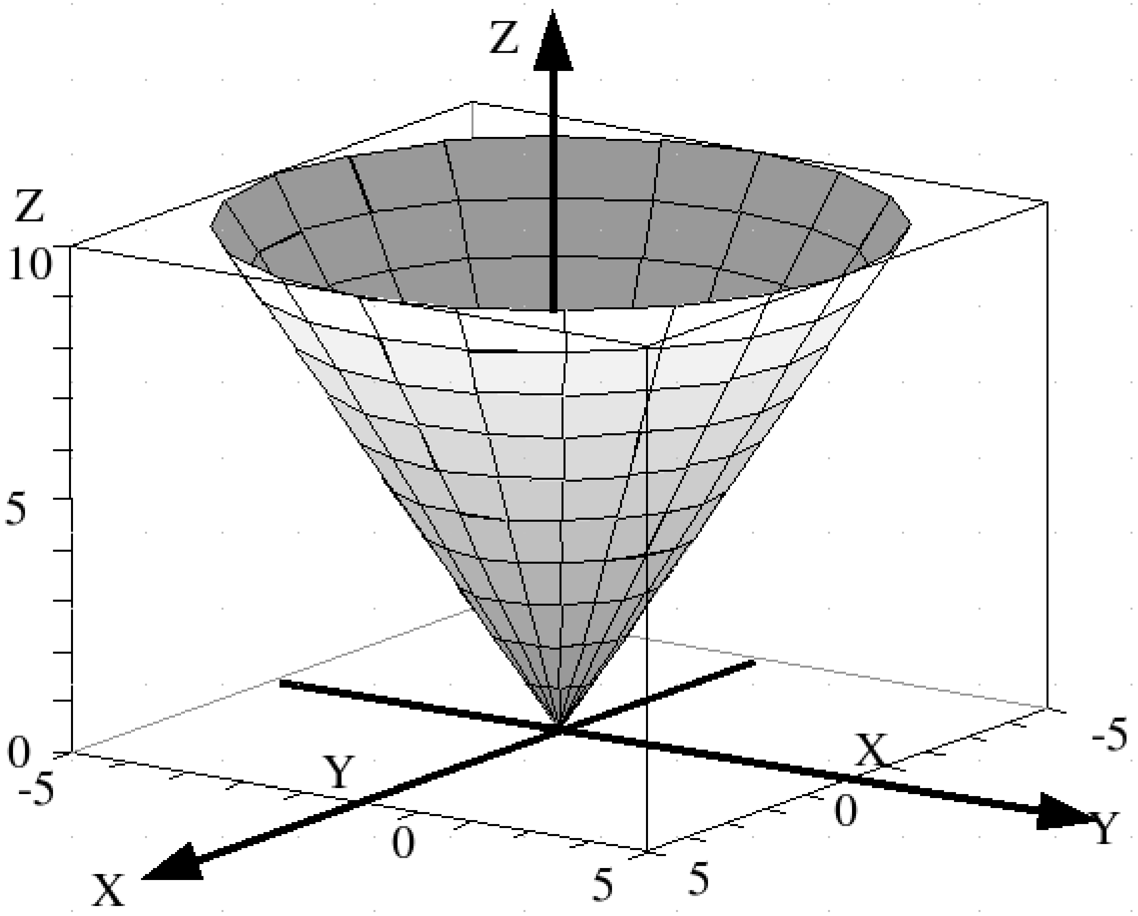
$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m z_i^2 \\ \text{subject to} & a_i^T x - b_i = y_i, \quad i = 1, \ldots, m \\ & y_i^2 \le z_i, \quad i = 1, \ldots, m \end{array}$$

Norm Balls and Norm Cones

- **norm**: a function $\|\cdot\|$ that satisfies
  - $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
  - $\|tx\| = |t|\|x\|$ for $t \in \mathbb{R}$
  - $\|x + y\| \leq \|x\| + \|y\|$
- notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{symb}$ is particular norm
- **norm ball** with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$
- **norm cone**: $\{(x, t) \mid \|x\| \leq t\}$
- Euclidean norm cone is called second-order cone
- Norm balls and cones are convex



$$x^2 + y^2 \leq z^2$$



Second Order Cone

- $\|u\| < t$
  - u-vector of dimension '$d - 1$'
  - t-scalar variable
  - Cone lies in 'd' dimensions

- Second Order Cone defines a convex set

Example: Second Order Cone in 3D

$$x^2 + y^2 \leq z^2$$

Second-Order Cone Programming (SOCP)

$$
\begin{array}{lll}
\underset{x}{\text{minimize}} & f^T x & \\
\text{subject to} & \|A_i x + b_i\| \leq c_i^T x + d_i & i = 1, \ldots, m \\
& Fx = g &
\end{array}
$$

- Convex problem: linear objective and second-order cone constraints
- For A_i=0, $b_i = 0$, it reduces to an LP.
- For $c_i = 0$, it reduces to a QCQP.
- More general than QCQP and LP.

Motivating Example

- Original:

$$mimimize \sqrt{(x+2)^2 + (y+1)^2} + \sqrt{(x+y)^2}$$

- Transformed:

$$\text{minimize } u + v$$

$$(x+2)^2 + (y+1)^2 \leq u^2$$

$$(x+y)^2 \leq v^2$$

$$u, v \geq 0$$

Motivating Example

- Original:

$$mimimize\ \sqrt{(x+2)^2+(y+1)^2}+\sqrt{(x+y)^2}$$

- Transformed:

$$\text{minimize } u+v$$

$$r^2+s^2\le u^2$$

$$t^2\le v^2$$

$$x+2=r$$

$$y+1=s$$

$$x+y=t$$

$$u,v\ge 0$$

Generally Accepted SOCP Form

- SOCP general form:

$$\text{minimize } f^T x$$

$$\text{subject to } \|A_i x + b_i\| \leq c_i^T x + d_i \ \forall i$$

where

- $x \in \mathbb{R}^n$ is the variable
- $f \in \mathbb{R}^n$
- $A_i \in \mathbb{R}^{m_i,n}$
- $b_i \in \mathbb{R}^{m_i}$
- $c_i \in \mathbb{R}^n$
- $d_i \in \mathbb{R}$

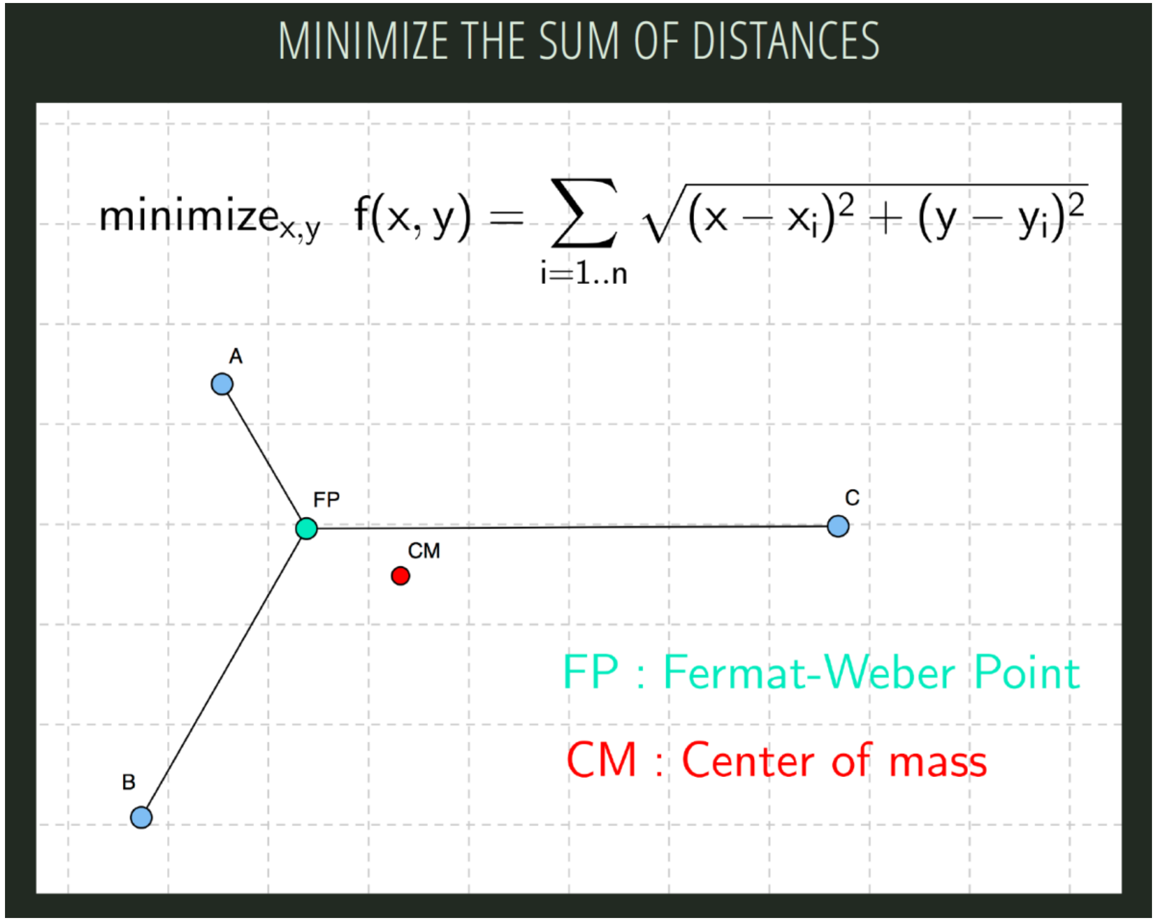
Sum of Norms

$$\text{minimize} \ \sum_{i=1}^{p} \|F_i x + g_i\|$$

$$\Longrightarrow$$

$$\begin{aligned}
\text{minimize} \ & \sum_{i=1}^{p} y_i \\
\text{subject to} \ & \sum_{j=1}^{q_i} u_{ij}^2 - y_i^2 \leq 0, \ \ i = 1..p \\
& (F_i x + g_i)_j - u_{ij} = 0, \ \ i = 1..p, \ \ j = 1..q_i \\
& y_i \geq 0, \ \ i = 1..p
\end{aligned}$$

The classical *Fermat-Weber* problem is a special case of the sum of norms problem. The Fermat-Weber problem considers where to place a facility so that the sum of the distances from this facility to a set of fixed locations is minimized. This problem is formulated as $\min_{\mathbf{x}} \sum_{i=1}^{k} \|\mathbf{d_i} - \mathbf{x}\|$, where the $\mathbf{d_i}$, $i = 1, \ldots, k$, are the fixed locations and $\mathbf{x}$ is the unknown facility location.



Max of Norms

$$\text{minimize} \quad \max_{i=1..p} \|F_i x + g_i\|$$

$$\Longrightarrow$$

$$\begin{aligned} & \text{minimize } y \\ & \text{subject to } \sum_{j=1}^{q_i} u_{ij}^2 - y^2 \leq 0, \;\; i = 1..p \\ & \qquad (F_i x + g_i)_j - u_{ij} = 0, \;\; i = 1..p, \;\; j = 1..q_i \\ & \qquad y_i \geq 0, \;\; i = 1..p \end{aligned}$$

Linear Regression

- Recall the second order cone:

$$K := \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \; : \; x \in \mathbb{R}^n, \; t \in \mathbb{R}, \; t \geq \|x\| \right\}.$$

- The least squares solution of a linear system of equations $Ax = b$ is the solution of

$$\min_x \|Ax - b\| = \min\{t \; : \; t \geq \|Ax - b\|\}.$$

This is a second order cone programming problem (why?)

Exercise

Hyperbolic constraints as SOC constraints. Verify that $x \in \mathbf{R}^n$, $y, z \in \mathbf{R}$ satisfy

$$x^T x \leq yz, \qquad y \geq 0, \qquad z \geq 0$$

if and only if

$$\left\| \begin{bmatrix} 2x \\ y - z \end{bmatrix} \right\|_2 \leq y + z, \qquad y \geq 0, \qquad z \geq 0.$$

Use this observation to cast the following problems as SOCPs.

(a) *Maximizing harmonic mean.*

$$\text{maximize} \quad \left(\sum_{i=1}^m 1/(a_i^T x - b_i) \right)^{-1},$$

with domain $\{x \mid Ax \succ b\}$, where a_i^T is the ith row of A.

Solution

(a) The problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T t \\ \text{subject to} & t_i(a_i^T x + b_i) \geq 1, \quad i = 1, \ldots, m \\ & t \succeq 0. \end{array}$$

Writing the hyperbolic constraints as SOC constraints yields an SOCP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T t \\ \text{subject to} & \left\| \begin{bmatrix} 2 \\ a_i^T x + b_i - t_i \end{bmatrix} \right\|_2 \leq a_i^T x + b_i + t_i, \quad i = 1, \ldots, m \\ & t_i \geq 0, \quad a_i^T x + b_i \geq 0, \quad i = 1, \ldots, m. \end{array}$$

Positive Semidefinite Cone

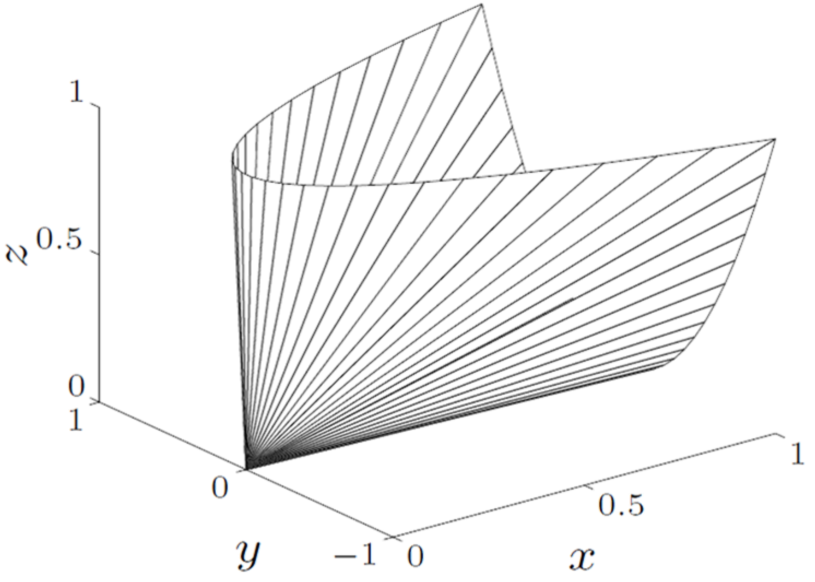
- **notation**:
  - $\mathbb{S}^n$ is set of symmetric $n \times n$ matrices
  - $\mathbb{S}_+^n = \{X \in \mathbb{S}^n | X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbb{S}_+^n \Leftrightarrow z^T X z \geq 0 \text{ for all } z$$

    $\mathbb{S}_+^n$ is a convex cone
  - $\mathbb{S}_{++}^n = \{X \in \mathbb{S}^n | X \succ 0\}$: positive definite $n \times n$ matrices

Example

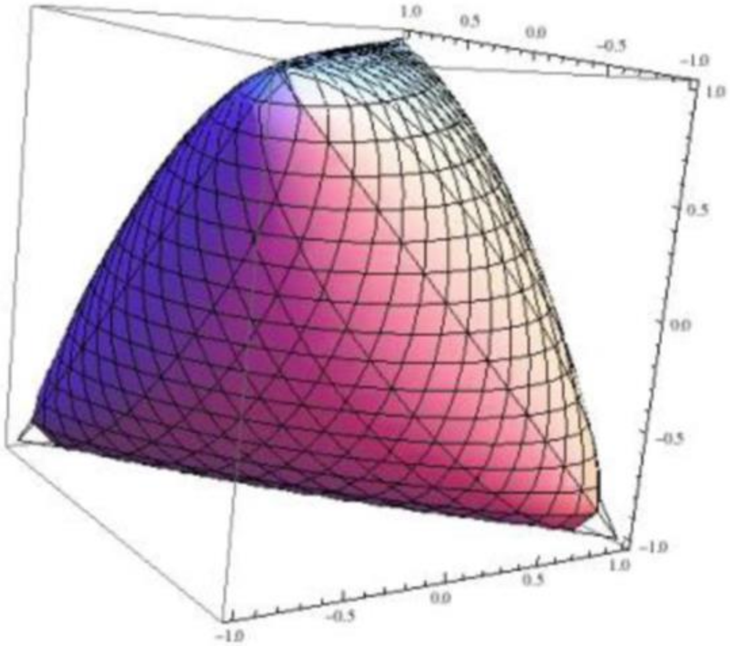
$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbb{S}_+^2$$



Spectrahedron

The feasible set of an SDP is called a spectrahedron. Every polyhedron is a spectrahedron.

$$\left\{ (x, y, z) \;\middle|\; \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0 \right\}$$



The so-called “elliptope."

http://www.princeton.edu/~amirali/Public/Teaching/ORF523/ORF523_Lec9.pdf

Positive Semidefinite Matrices

Theorem (Properties of p.s.d. matrices):

Let $X \in \mathbb{S}_+^{n\times n}$. The following are equivalent:

- $X \in \mathbb{S}_+^{n\times n}$ or $X \succeq 0$ (X is p.s.d.);
- $z^T X z \geq 0, \forall z \in \mathbb{R}^n$;
- $\lambda_{min}(X) \geq 0$;
- All principal minors of X are nonnegative;
- $X = LL^T$ for some $L \in \mathbb{R}^{n\times n}$.

- A nonsingular matrix $X \succeq 0$ is called positive definite ($X \succ 0$ or $X \in \mathbb{S}_{++}^{n\times n}$).

Semidefinite Programming

Let $X \in S^n$. We can think of X as a matrix, or equivalently, as an array of n^2 components of the form $(x_{11}, \ldots, x_{nn})$. We can also just think of X as an object (a vector) in the space S^n. All three different equivalent ways of looking at X will be useful.

What will a linear function of X look like? If $C(X)$ is a linear function of X, then $C(X)$ can be written as $C \bullet X$, where

$$C \bullet X := \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} X_{ij}.$$

Semidefinite Programming

$$
\begin{array}{llll}
SDP: & \text{minimize} & C \bullet X & \\
& \text{s.t.} & A_i \bullet X = b_i & , i = 1, \ldots, m, \\
& & X \succeq 0, &
\end{array}
$$

Notice that in an SDP that the variable is the matrix X, but it might be helpful to think of X as an array of n^2 numbers or simply as a vector in S^n. The objective function is the linear function $C \bullet X$ and there are m linear equations that X must satisfy, namely $A_i \bullet X = b_i \quad , i = 1, \ldots, m$. The variable X also must lie in the (closed convex) cone of positive semidefinite symmetric matrices S^n_+. Note that the data for SDP consists of the symmetric matrix C (which is the data for the objective function) and the m symmetric matrices $A_1, \ldots, A_m$, and the $m-$vector b, which form the m linear equations.

Let us see an example of an SDP for $n = 3$ and $m = 2$. Define the following matrices:

$$A_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 7 \\ 1 & 7 & 5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 2 & 8 \\ 2 & 6 & 0 \\ 8 & 0 & 4 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 9 & 0 \\ 3 & 0 & 7 \end{pmatrix},$$

and $b_1 = 11$ and $b_2 = 19$. Then the variable X will be the 3×3 symmetric matrix:

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix},$$

$$
\begin{aligned}
SDP: \quad \text{minimize} \quad & x_{11} + 4x_{12} + 6x_{13} + 9x_{22} + 0x_{23} + 7x_{33} \\
\text{s.t.} \quad & \\
& x_{11} + 0x_{12} + 2x_{13} + 3x_{22} + 14x_{23} + 5x_{33} = 11 \\
& 0x_{11} + 4x_{12} + 16x_{13} + 6x_{22} + 0x_{23} + 4x_{33} = 19 \\
& X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \succeq 0.
\end{aligned}
$$

Notice that SDP looks remarkably similar to a linear program. However, the standard LP constraint that x must lie in the nonnegative orthant is replaced by the constraint that the variable X must lie in the cone of positive semidefinite matrices. Just as "$x \geq 0$" states that each of the n components of x must be nonnegative, it may be helpful to think of "$X \succeq 0$" as stating that each of the n *eigenvalues* of X must be nonnegative.

QCQP to SDP

$$
\begin{array}{lll}
QCQP: & \underset{x}{\text{minimize}} & x^T Q_0 x + q_0^T x + c_0 \\
& \text{s.t.} & x^T Q_i x + q_i^T x + c_i \leq 0, \;\; i = 1, \dots, m,
\end{array}
$$

$$
\begin{array}{lll}
QCQP: & \underset{x,\theta}{\text{minimize}} & \theta \\
& \text{s.t.} & x^T Q_0 x + q_0^T x + c_0 - \theta \leq 0 \\
& & x^T Q_i x + q_i^T x + c_i \leq 0, \;\; i = 1, \dots, m.
\end{array}
$$

$$Q_i = M_i^T M_i$$

$$
\begin{pmatrix} I & M_i x \\ x^T M_i^T & -c_i - q_i^T x \end{pmatrix} \succeq 0 \quad \Longleftrightarrow \quad x^T Q_i x + q_i^T x + c_i \leq 0.
$$

QCQP to SDP

$$
\begin{array}{lll}
QCQP: & \underset{x,\theta}{\text{minimize}} & \theta \\
& \text{s.t.} & x^T Q_0 x + q_0^T x + c_0 - \theta \leq 0 \\
& & x^T Q_i x + q_i^T x + c_i \leq 0, \;\; i = 1, \ldots, m.
\end{array}
$$

$$
\begin{array}{lll}
QCQP: & \underset{x,\theta}{\text{minimize}} & \theta \\
& \text{s.t.} & \\
& & \begin{pmatrix} I & M_0 x \\ x^T M_0^T & -c_0 - q_0^T x + \theta \end{pmatrix} \succeq 0 \\
& & \begin{pmatrix} I & M_i x \\ x^T M_i^T & -c_i - q_i^T x \end{pmatrix} \succeq 0, \;\; i = 1, \ldots, m.
\end{array}
$$

SOCP to SDP

$$\text{SOCP:} \quad \begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & Ax = b \\ & \|Q_i x + d_i\| \le (g_i^T x + h_i) \ , \quad i = 1, \ldots, k. \end{array}$$

$$\|Qx + d\| \le (g^T x + h) \iff \begin{pmatrix} (g^T x + h)I & (Qx + d) \\ (Qx + d)^T & g^T x + h \end{pmatrix} \succeq 0.$$

$$M = \begin{pmatrix} P & v \\ v^T & d \end{pmatrix} \succeq 0 \iff d - v^T P^{-1} v \ge 0.$$

$$\text{SDPSOCP:} \quad \begin{array}{ll} \min_x & c^T x \\ \text{s.t.} & Ax = b \\ & \begin{pmatrix} (g_i^T x + h_i)I & (Q_i x + d_i) \\ (Q_i x + d_i)^T & g_i^T x + h_i \end{pmatrix} \succeq 0 \ , \quad i = 1, \ldots, k. \end{array}$$

Property: $M \succeq tI$ if and only if $\lambda_{\min}(M) \geq t$.

To see why this is true, let us consider the eigenvalue decomposition of $M = QDQ^T$, and consider the matrix R defined as:

$$R = M - tI = QDQ^T - tI = Q(D - tI)Q^T.$$

Then

$$M \succeq tI \iff R \succeq 0 \iff D - tI \succeq 0 \iff \lambda_{\min}(M) \geq t.$$

Property: $M \preceq tI$ if and only if $\lambda_{\max}(M) \leq t$.

$$S := B - \sum_{i=1}^{k} w_i A_i.$$

Now suppose that we wish to find weights w to minimize the difference between the largest and the smallest eigenvalues of S. This problem can be written down as:

$$\begin{array}{lll} EOP: & \underset{w,S}{\text{minimize}} & \lambda_{\max}(S) - \lambda_{\min}(S) \\ & \text{s.t.} & S = B - \sum_{i=1}^{k} w_i A_i \\ & & Gw \leq d. \end{array}$$

Then EOP can be written as:

$$
\begin{array}{llll}
EOP: & \underset{w, S, \mu, \lambda}{\text{minimize}} & \mu - \lambda \\
& \text{s.t.} & S = B - \sum_{i=1}^{k} w_i A_i \\
& & Gw \leq d \\
& & \lambda I \preceq S \preceq \mu I .
\end{array}
$$

This last problem is a semidefinite program.