Convex problems

Nonlinear optimization
Instituto Superior Técnico and Carnegie Mellon University PhD course
João Xavier
TA: Hung Tuan

- special classes of convex problems
  - linear programming (LP)
  - quadratic programming (QP)
  - second-order cone programming (SOCP)
  - semi-definite programming (SDP)

- the classes are nested: LP $\subset$ QP $\subset$ SOCP $\subset$ SDP

- there exist efficient algorithms for each class (and available for free)

- their complexity increases along the chain: a SDP takes much more time to solve than a "comparable" LP

Quadratic programming (QP)

- minimization of a convex quadratic over finitely many affine constraints

- in $\mathbf{R}^n$, a QP is a problem of the form

$$\begin{array}{ll} \underset{x}{\text{minimize}} & x^T A x + b^T x + c \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & c_j^T x = d_j, \quad j = 1, \ldots, p \end{array}$$

  with $A \succeq 0$

- LP $\subset$ QP by taking $A = 0$

Example: portfolio management

- you have T euros to invest across n assets
- r_i is the random rate of return of the asset $i = 1, \ldots, n$
- x_i is the amount you invest in asset $i = 1, \ldots, n$
- you receive the random amount

$$r^T x = r_1 x_1 + r_2 x_2 + \cdots + r_n x_n$$

- the average and covariance of the random vector r are known:

$$\mu = \mathbf{E}(r) \qquad \Sigma = \textbf{cov}\,(r) = \mathbf{E}\left((r-\mu)(r-\mu)^T\right)$$

(this implies: $\mathbf{E}(r^T x) = \mu^T x$ and $\textbf{var}(r^T x) = x^T \Sigma x$)

- how much should you invest in each asset?

- risk-averse formulation:

$$\begin{array}{ll} \underset{x}{\text{maximize}} & \mu^T x - \beta x^T \Sigma x \\ \text{subject to} & 1^T x = T \\ & x \geq 0 \end{array}$$

- the constant $\beta > 0$ sets the tradeoff between the expected gain and risk of the portfolio $x = (x_1, \ldots, x_n)$

- a QP because $\Sigma \succeq 0$

Example: fire-station location

- original formulation:

$$\underset{x}{\text{minimize}} \quad \max\left\{\|x - p_1\|, \ldots, \|x - p_K\|\right\}$$

- not a QP
- reformulate as

$$\underset{x}{\text{minimize}} \quad \max\left\{\|x - p_1\|^2, \ldots, \|x - p_K\|^2\right\}$$

or

$$\underset{x}{\text{minimize}} \quad \|x\|^2 + \max\left\{-2p_1^T x + \|p_1\|^2, \ldots, -2p_K^T x + \|p_K\|^2\right\}$$

- introduce an epigraph variable

$$\begin{array}{ll} \underset{x,t}{\text{minimize}} & \|x\|^2 + t \\ \text{subject to} & -2p_k^T x + \|p_k\|^2 \leq t, \quad k = 1, \ldots, K. \end{array}$$

- a QP

Second-order cone programming (SOCP)

- in $\mathbf{R}^n$, a SOCP is a problem of the form

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & \|A_i x + b_i\| \leq c_i^T x + d_i, \quad i = 1, \ldots, m. \end{array}$$

- key-fact: for $w \in \mathbf{R}^n$ and $y, z \in \mathbf{R}$, there holds

$$\begin{cases} \|w\|^2 \leq yz \\ y \geq 0 \\ z \geq 0 \end{cases} \quad \Leftrightarrow \quad \left\| \begin{bmatrix} 2w \\ y - z \end{bmatrix} \right\| \leq y + z$$

- QP $\subset$ SOCP:

$$\begin{array}{ll}\underset{x}{\text{minimize}} & x^T A x + b^T x + c \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & c_j^T x = d_j, \quad j = 1, \ldots, p\end{array}$$

with $A \succeq 0$, can be reformulated as

$$\begin{array}{ll}\underset{x,s}{\text{minimize}} & s + b^T x + c \\ \text{subject to} & \left\| \begin{bmatrix} A^{1/2} x \\ s - 1 \end{bmatrix} \right\| \leq s + 1 \\ & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & c_j^T x = d_j, \quad j = 1, \ldots, p\end{array}$$

Example: input design

- a team of two vehicles move in the plane with dynamics

$$x_i(t) = A_i x_i(t-1) + B_i u_i(t), \quad t = 1, \ldots, T$$

- $x_i(t) = (p_i(t), v_i(t)) \in \mathbf{R}^4$ is state of vehicle i at time t ($p_i(t)$=position, $v_i(t)$=velocity)

- $u_i(t) \in \mathbf{R}^n$ is control that we apply to vehicle i at time t

- matrices $A_i \in \mathbf{R}^{4\times 4}, B_i \in \mathbf{R}^{4\times n}$ and initial states $x_i(0)$ are given

- we want to move the vehicles to given desired positions $q_i \in \mathbf{R}^2$ at time T and stop them there

- preferably, vehicles should stay within r distance units of each other at all times (otherwise, their wireless link starts deteriorating)

- goal: design a minimum-energy input sequence $u_i(t)$, $t = 1, \dots T$

- formulation:

$$\begin{array}{ll} \underset{p_i(t), v_i(t), u_i(t)}{\text{minimize}} & \sum_{t=1}^{T} \|u_1(t)\|^2 + \|u_2(t)\|^2 + \rho \left((\|p_1(t) - p_2(t)\| - r)_+ \right)^2 \\ \text{subject to} & \begin{bmatrix} p_i(t) \\ v_i(t) \end{bmatrix} = A_i \begin{bmatrix} p_i(t-1) \\ v_i(t-1) \end{bmatrix} + B_i u_i(t) \\ & p_i(T) = q_i \\ & v_i(T) = 0 \end{array}$$

The constraints run for $t = 1, \dots, T$ and $i = 1, 2$

- $\rho > 0$ is given trade-off parameter

- not a SOCP formulation: constraints are OK, but objective is not linear

- since $((\cdot)_+)^2$ is nondecreasing, we may reformulate as

$$\begin{array}{ll} \underset{p_i(t), v_i(t), u_i(t), s(t)}{\text{minimize}} & \sum_{t=1}^{T} \|u_1(t)\|^2 + \|u_2(t)\|^2 + \rho\,(s(t)_+)^2 \\ \text{subject to} & \|p_1(t) - p_2(t)\| - r \leq s(t) \\ & \begin{bmatrix} p_i(t) \\ v_i(t) \end{bmatrix} = A_i \begin{bmatrix} p_i(t-1) \\ v_i(t-1) \end{bmatrix} + B_i u_i(t) \\ & p_i(T) = q_i \\ & v_i(T) = 0 \end{array}$$

- introduce epigraph variables

$$
\begin{array}{ll}
\underset{p_i(t), v_i(t), u_i(t), s(t), \alpha(t), \beta(t)}{\text{minimize}} & \sum_{t=1}^{T} \alpha(t) + \rho\beta(t) \\
\text{subject to} & \|u_1(t)\|^2 + \|u_2(t)\|^2 \leq \alpha(t) \\
& (s(t)_+)^2 \leq \beta(t) \\
& \|p_1(t) - p_2(t)\| - r \leq s(t) \\
& \begin{bmatrix} p_i(t) \\ v_i(t) \end{bmatrix} = A_i \begin{bmatrix} p_i(t-1) \\ v_i(t-1) \end{bmatrix} + B_i u_i(t) \\
& p_i(T) = q_i \\
& v_i(T) = 0
\end{array}
$$

- objective is linear now, but constraints are not SOCP

- Fact: for $x, y \in \mathbf{R}$, there holds

$$(x_+)^2 \leq y \quad \Leftrightarrow \quad \exists_{z \in \mathbf{R}} : x_+ \leq z \text{ and } z^2 \leq y$$

$$\begin{array}{ll}
\underset{p_i(t), v_i(t), u_i(t), s(t), \alpha(t), \beta(t), z(t)}{\text{minimize}} & \sum_{t=1}^{T} \alpha(t) + \rho\beta(t) \\
\text{subject to} & \|u_1(t)\|^2 + \|u_2(t)\|^2 \leq \alpha(t) \\
& s(t)_+ \leq z(t) \\
& z(t)^2 \leq \beta(t) \\
& \|p_1(t) - p_2(t)\| - r \leq s(t) \\
& \begin{bmatrix} p_i(t) \\ v_i(t) \end{bmatrix} = A_i \begin{bmatrix} p_i(t-1) \\ v_i(t-1) \end{bmatrix} + B_i u_i(t) \\
& p_i(T) = q_i \\
& v_i(T) = 0
\end{array}$$

- a SOCP formulation:

$$
\begin{array}{ll}
\underset{p_i(t),v_i(t),u_i(t),s(t),\alpha(t),\beta(t),z(t)}{\text{minimize}} & \sum_{t=1}^{T} \alpha(t) + \rho\beta(t) \\
\text{subject to} & \left\| \begin{bmatrix} 2u_1(t) \\ 2u_2(t) \\ \alpha(t) - 1 \end{bmatrix} \right\| \leq \alpha(t) + 1 \\
& 0 \leq z(t), s(t) \leq z(t) \\
& \left\| \begin{bmatrix} 2z(t) \\ \beta(t) - 1 \end{bmatrix} \right\| \leq \beta(t) + 1 \\
& \|p_1(t) - p_2(t)\| \leq s(t) + r \\
& \begin{bmatrix} p_i(t) \\ v_i(t) \end{bmatrix} = A_i \begin{bmatrix} p_i(t-1) \\ v_i(t-1) \end{bmatrix} + B_i u_i(t) \\
& p_i(T) = q_i \\
& v_i(T) = 0
\end{array}
$$