Convex Optimization

Kontonis Vasilis

17 February, 2017

Corelab, NTUA

Contents

Quadratic Programming

QP & QCQP

Basic **QP** Problem

$$\begin{aligned} \text{minimize} \quad & (1/2)x^T P x + q^T x + r \\ \text{subject to} \quad & Gx \leqslant h \\ & Ax = b \end{aligned}$$

If we allow **quadratic** inequality constraints we have a **QCQP** problem

$$\begin{aligned} \text{minimize} \quad & (1/2)x^T P x + q^T x + r \\ \text{subject to} \quad & (1/2)x^T P_i x + q_i^T x + r_i \leqslant 0, \; i = 1, \dots, m \\ & Ax = b \end{aligned}$$

- The feasible region is the intersection of ellipsoids.
- Generalizes QP and LP.

Bounded Least Squares

The well-known least squares problem

$$\text{minimize} \quad \|Ax - b\|_2^2 = x^T(A^TAx) - 2b^TAx + b^Tb$$

In the unconstraint case we can obtain the **normal** equations

$$A^TAx = A^Tb.$$

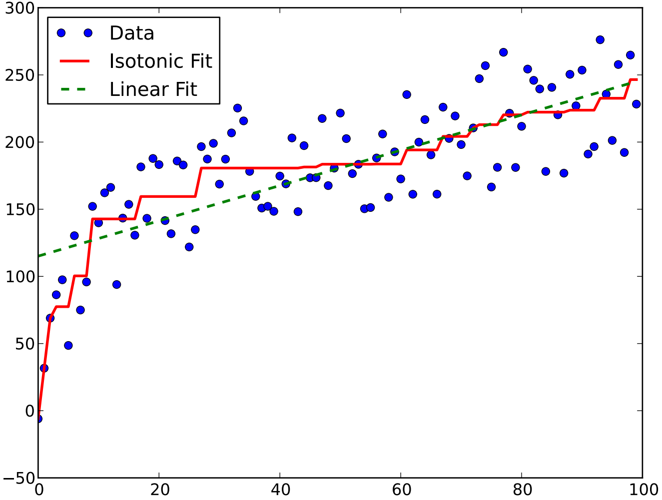
The QP for constraint Least Squares:

$$\begin{aligned} &\text{minimize} && \|Ax - b\|_2^2 \\ &\text{subject to} && l_i \leqslant x_i \leqslant u_i,\ i = 1, \ldots, n \end{aligned}$$

Examples:

- Estimation of non-negative parameters.
- Isotonic (or Monotonic) Regression, $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$.

Isotonic Regression



Polyhedra Distance

Let $P_1 = \{x \mid A_1 x \leqslant b_1\}$ and $P_2 = \{x \mid A_2 x \leqslant b_2\}$ be two polyhedra in $\mathbb{R}^n$.

$$\text{dist}(P_1, P_2) = \inf\{\|x_1 - x_2\|_2 \mid x_1 \in P_1,\ x_2 \in P_2\}$$

The QP:

$$\begin{aligned} \text{minimize} \quad & \|x_1 - x_2\|_2^2 \\ \text{subject to} \quad & A_1 x_1 \leqslant b_1,\ A_2 x_2 \leqslant b_2 \end{aligned}$$

Bounding Variance

We want to bound the variance of a function f of the RV of Chebyshev Inequalities problem.

$$\mathsf{Var}[f(X)] = \mathbb{E}[f^2(X)] - (\mathbb{E}[f(X)])^2 = \sum f_i^2 p_i - \left(\sum f_i p_i\right)^2$$

QP:

$$\begin{array}{ll} \text{maximize} & \mathsf{Var}[f(X)] \\ \text{subject to} & \alpha_i \leqslant a_i^T p \leqslant \beta_i,\ i = 1, \ldots, m \\ & p \geqslant 0,\ \mathbf{1}^T p = 1 \end{array}$$

Linear Program with Random Cost

Let $c \in \mathbb{R}^n$ be a Random Vector, with mean $\bar{c}$ and covariance $\mathbb{E}(c - \bar{c})(c - \bar{c})^T = \Sigma$.
Basic LP:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Gx \leqslant h \\ & Ax = b \end{aligned}$$

- Trade-off between small expected cost and small cost variance.
- Define the risk-sensitive cost $\mathbb{E}[c^T x] + \gamma \text{Var}(c^T x)$, where γ is the **risk-aversion** parameter. Is the covariance matrix PSD?

QP:

$$\begin{aligned} \text{minimize} \quad & \bar{c}^T x + \gamma x^T \Sigma x \\ \text{subject to} \quad & Gx \leqslant h \\ & Ax = b \end{aligned}$$

Markowitz portfolio Optimization

- n assets held over a period of time.
- x_i(dollars) amount of asset i held throughout the period.
- p_i relative change in the price of asset i over the period, $r = p^T x$ return of the portfolio.
- We do not allow "shorting" assets, $x \geqslant 0$.
- Total budget is assumed to be 1, $\mathbf{1}^T x = 1$.

We assume p to be a Random Vector with mean $\bar{p}$ and covariance Σ.
QP:

$$\begin{aligned} \text{minimize} \quad & x^T \Sigma x \\ \text{subject to} \quad & \bar{p}^T x \geqslant r_{min} \\ & \mathbf{1}^T x = 1, \; x \geqslant 0 \end{aligned}$$

Markowitz portfolio Optimization

Extensions:

- To allow short positions $x_i < 0$ we introduce x_{long}, x_{short} s.t.

  $x_{long} \geqslant 0,\ x_{short} \geqslant 0,\ x = x_{long} - x_{short},\ \mathbf{1}^T x_{short} \geqslant \eta \mathbf{1}^T x_{long}$

- Include linear transaction costs to go from an initial portfolio x_{init} to a desired portfolio x, which then is held over the period.

$$x = x_{init} + u_{buy} - u_{sell},$$

$$u_{buy} \geqslant 0, u_{sell} \geqslant 0.$$

  Initial buying and selling involves zero net cach:

$$(1 - f_{sell})\mathbf{1}^T u_{sell} = (1 + f_{buy})\mathbf{1}^T u_{buy}$$

$$f_{buy}, f_{sell} > 0.$$

Second-Order Cone Programming

Dual Spaces

Linear Maps
Let $X,\ Y$ be two normed spaces.

- A map $T : X \to Y$ s.t $T(\lambda x_1 + \mu x_2) = \lambda T(x_1) + \mu T(x_2)$ is a **linear map**.
- T is bounded if there is a constant c s.t. $\|Tx\|_Y \leqslant c\|x\|_X$.
  $\|T\| = \min\{c \geqslant 0 : \forall x \in X,\ \|Tx\| \leqslant c\|x\|\}$.
- Operator Norm $\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|$.
- $\|F\|_2 = \sup\{\|Fx\|_2 \mid \|u\|_2 \leqslant 1\} = \sqrt{\lambda_{max}(F^T F)}$

Linear Functional
A Linear functional is a Linear Map $F : X \to \mathbb{R}$.

Dual Space
Let X be a normed space. The space X^* of the bounded linear functionals $F : X \to \mathbb{R}$ is the dual space of X.

Dual Norms

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. Its **dual** norm is defined

$$\|z\|_* = \sup\{z^\mathsf{T}x \mid \|x\| \leqslant 1\} = \sup\{|z^\mathsf{T}x| \mid \|x\| \leqslant 1\}$$

- $\|x\|_{**} = \|x\|$. Does not hold in infinite-dimensional vector spaces.
- The ℓ_2 norm is self-dual.
- The dual of ℓ_∞-norm is the ℓ_1-norm.

Definition

Norm Cone: $C = \{(x,t) \mid \|x\| \leqslant t\} \subseteq \mathbb{R}^{n+1}$.
SOCP Definition:

$$\begin{aligned} \text{minimize} \quad & f^T x \\ \text{subject to} \quad & \|A_i x + b_i\|_2 \leqslant c_i^T x + d_i,\ i = 1, \ldots, m \\ & Fx = g \end{aligned}$$

- SOCP is a generalization of LP and QCQP.
  $x^T P_0 x + 2q_0^T x + r_0 = \|P_0^{1/2} x + P_0^{-1/2} q_0\|^2 + r_0 - q_0^T P_0^{-1} q_0$
  The optimal values of the QCQP and the SOCP are equal up to a square root and a constant.
- The second-order cone constraint requires that the affine function $(Ax + b, c^T x + d)$ lies in the second-order cone in $\mathbb{R}^{k+1}$

Robust Linear Programming

Often we only know approximations of the coefficients the usual LP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leqslant b_i,\ i = 1, \dots, m \end{array}$$

Assume that c, b_i are known exactly but a_i are known to lie in ellipsoids $\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leqslant 1\}$.

Robust SOCP:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i{}^T x + \|P_i^T x\|_2 \leqslant b_i,\ i = 1, \dots, m \end{array}$$

Linear Programming with Random Constraints

Statistical framework for the robust LP.
Each constraint a_i is a Gaussian Random Vector with mean $\bar{a}_i$ and covariance Σ_i and the constraints must hold with confidence at least $\eta \geqslant 1/2$ (Why?)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \Pr[a_i^T x \leqslant b_i] \geqslant \eta \end{array}$$

Equivalent **SOCP**:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta) \|\Sigma^{1/2} x\|_2 \leqslant b_i, \; i = 1, \ldots, m \end{array}$$

Questions?

References i

S. Boyd and L. Vandenberghe.
Convex Optimization.
Cambridge University Press, Cambridge, UK ; New York, Mar. 2004.

M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret.
Applications of second-order cone programming.
Linear Algebra and its Applications, 284(13):193–228, Nov. 1998.