Optimization Models

EECS 127 / EECS 227AT

Laurent El Ghaoui

EECS department
UC Berkeley

Fall 2018

LECTURE 21

Review: Convex Models

I have dual citizenship.

Simone Biles

Outline

Linear algebra

What we have seen in linear algebra:

- Matrix-vector, matrix-matrix product.
- Norms, projection on a line.
- Solving linear equations, least-squares.
- Spectral decomposition of symmetric matrices.
- Singular value decomposition of arbitrary matrices.
- Principal component analysis, low-rank approximation.

Linear algebra problems as optimization problems

- Solving linear systems of equations: (A is a matrix, b a vector)

$$\min_x \; 0 \; : \; Ax = b$$

- Ridge regression: (X is a data matrix, y a vector and $\lambda > 0$ a "regularization" parameter)

$$\min_w \; \|X^T w - y\|_2^2 + \lambda \|w\|_2^2$$

(includes projection on a line as special case!);

- Maximum-variance direction and PCA ($C = C^T \succeq 0$ is a covariance matrix):

$$\max_w \; w^T C w \; : \; w^T w = 1,$$

- Low-rank approximation:

$$\min_{x,y} \; \|X - xy^T\|_F$$

where x, y are vectors, and X is a given data matrix. Norm in objective can be the largest singular value norm, with same result.

Conic optimization

Linear programming

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^\top x \\ \text{s.t.:} \quad & Ax \le b, \quad Cx = d \end{aligned}$$

,

with A, b, c, C, d matrices or vectors of appropriate size.

Important applications examples:

- solving linear equations;
- linear regression problems based on l_1- and l_∞-norms;
- resource management;
- network flows.

Conic optimization

Quadratic programming

$$\begin{aligned} \text{minimize} \quad & x^\top Q x + c^\top x \\ \text{subject to:} \quad & Ax \leq b, \;\; Cx = d, \end{aligned}$$

with A, b, C, d, H matrices or vectors of appropriate size, and with Q PSD (symmetric and positive-semidefinite, also denoted $Q \succeq 0$). When $Q \succ 0$ (positive-definite), QP is a regularized version of LP, with a unique solution (if problem is feasible). The model includes LP as a special case.

Important applications examples:

- solving linear equations via least-squares;
- sparsity-constrained least-squares (LASSO);
- portfolio optimization, index tracking;
- linear-quadratic control.

Conic optimization

Quadratically constrained quadratic programming

The convex quadratic-constrained quadratic program (QCQP) model is

$$\begin{aligned} \min_{x} \quad & x^\top Q_0 x + a_0^\top x \\ \text{s.t.:} \quad & x^\top Q_i x + a_i^\top x \le b_i, \quad i = 1, \ldots, m, \end{aligned}$$

with a_i, b_i vectors and scalars, and PSD matrices Q_i PSD ($Q_i = Q_i^\top \succeq 0$), $i = 0, 1, \ldots, m$. The model includes LP, QP as a special case.

Important applications examples:

- minimization of the maximum of quadratic functions;
- geometric problems, such as finding a point in the intersection of ellipses.
- portfolio optimization with multiple risk (variance) constraints.

Conic optimization

Second-order cone programming

The SOCP model is

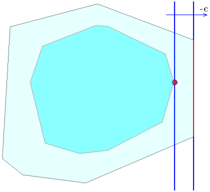
$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^\top x \\ \text{s.t.:} \quad & \|A_i x + b_i\|_2 \leq c_i^\top x + d_i, \quad i = 1, \ldots, m, \end{aligned}$$

with A_i, b_i, c_i, d_i, $i = 1, \ldots, m$ matrices of appropriate size. The model includes LP, QP and QCQP as a special case.

Important applications examples:

- linear regression problems involving sums-of-powers of variables;
- grasping problems in robotics; truss optimization; etc.
- robust linear programming with ellipsoidal uncertainty.

Robust optimization



Robust counterpart:

$$
\begin{array}{ll}
\min\limits_{x} & \max\limits_{u \in \mathcal{U}} f_0(x, u) \\
\text{s.t.} & \forall\, u \in \mathcal{U}, \;\; f_i(x, u) \leq 0, \\
 & i = 1, \ldots, m
\end{array}
$$

- functions f_i now depend on a second variable u, the "uncertainty", which is constrained to lie in given set $\mathcal{U}$;
- in the case of robust linear programming, and with sets $\mathcal{U}$ that are ellipsoids or boxes, robust counterpart is tractable (an SOCP).

Are there other conic problems?

The last "Russian doll"

Yes!

- It's called "semidefinite programming" (SDP), and it involves optimization with PSD matrices as variables.
- It includes SOCP as a special case.
- SDPs are beyond the scope of this class, but are extensively covered in EE 227BT.

Solving conic problems

Due to their structure, conic problems can be efficiently solved globally:

- conic optimization solvers can provide a (near) optimal point, or unambiguously determine that the problem is infeasible;
- in practical terms, general SOCPs with dense input data and tens of thousands of variables and constraints can be solved in minutes on an ordinary laptop;
- With sparse input data the reach is much higher.

This is in sharp contrast with solvers for general nonlinear programming:

- users must provide a “good” initial guess;
- a “wrong” one may lead to a very sub-optimal solution;
- in the case of constrained problems, the solver may fail to find a feasible (let alone optimal) point, resulting in total failure, with little in the way of diagnostics.

Convex sets

A subset $C \subseteq \mathbb{R}^n$ is said to be *convex* if it contains the line segment between any two points in it:

$$x_1, x_2 \in C,\ \lambda \in [0,\ 1] \quad \Rightarrow \quad \lambda x_1 + (1-\lambda) x_2 \in C.$$

- The intersection of convex sets is convex.
- The affine transformation of a convex set is convex.

Convex functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if $\operatorname{dom} f$ is a convex set, and for all $x, y \in \operatorname{dom} f$ and all $\lambda \in [0, 1]$ it holds that

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

- A function is convex if and only if its epigraph is.
- The pointwise maximum of a family of functions is convex.
- The composition of a convex function with an affine map is convex.
- The non-negatively weighted sum of convex functions is convex.
- A twice-differentiable function is convex if and only if its Hessian is PSD everywhere.

Convex problem

Standard form

$$p^* = \min_{x \in \mathbb{R}^n} \; f_0(x) \text{ subject to: } \quad f_i(x) \leq 0, \quad i = 1, \ldots, m,$$
$$Ax = b,$$

where

- $f_0, \ldots, f_m$ are convex functions;
- The equality constraints are affine, and represented via the matrix $A \in \mathbb{R}^{q \times n}$ and vector $b \in \mathbb{R}^q$.

- Covers many fitting problems in statistics, as well as engineering design problems.
- Includes conic optimization (LP, QP, QCQP and SOCP) as special cases.
- Can be used to approximate non-convex problems.