EE 227A: Convex Optimization

Lecture 1: Optimization Models

Laurent El Ghaoui

EECS and IEOR Departments
UC Berkeley

Jan. 22, 2013
Outline

What is Optimization?
  Definition
  Examples
  Nomenclature
  Other standard forms
  Extensions

The Role of Convexity
  Global vs. local optima
  Convex problems
  Software
  Non-convex problems
  Non-convex problems

Linear Algebra & Optimization

References
Outline

What is Optimization?
- Definition
- Examples
- Nomenclature
- Other standard forms
- Extensions

The Role of Convexity
- Global vs. local optima
- Convex problems
- Software
- Non-convex problems
- Non-convex problems

Linear Algebra & Optimization

References
Optimization problem

A standard form

An optimization problem is a problem of the form

\[ p^* := \min_x f_0(x) \ \text{subject to} \ f_i(x) \leq 0, \ i = 1, \ldots, m, \]

where

- \( x \in \mathbb{R}^n \) is the decision variable;
- \( f_0 : \mathbb{R}^n \rightarrow \mathbb{R} \) is the objective (or, cost) function;
- \( f_i : \mathbb{R}^n \rightarrow \mathbb{R}, \ i = 1, \ldots, m \) represent the constraints;
- \( p^* \) is the optimal value.

Often the above is referred to as a “mathematical program” (for historical reasons).
Example
Least-squares regression

\[
\min_w \| X^T w - y \|_2
\]

where
- \( X = [x_1, \ldots, x_m] \) is a \( n \times m \) matrix of data points (\( x_i \in \mathbb{R}^n \));
- \( y \) is a response vector;
- \( \| \cdot \|_2 \) is the \( l_2 \) (i.e., Euclidean) norm.
- Many variants (with e.g., constraints) exist (more on this later).
- Perhaps the most popular / useful optimization problem.
Example
Linear classification

\[
\min_{w,b} \sum_{i=1}^{m} \max(0, 1 - y_i (w^T x_i + b))
\]

where

- \( X = [x_1, \ldots, x_m] \) is a \( n \times m \) matrix of data points \( (x_i \in \mathbb{R}^n) \);
- \( y \in \{-1, 1\} \) is a \textit{binary} response vector;
- Many variants (with \textit{e.g.}, constraints) exist (more on this later).
- Very useful for classifying data (\textit{e.g.}, text documents).
What is Optimization?

Definition
Examples
Nomenclature
Other standard forms
Extensions
Convexity
Global vs. local optima
Convex problems
Software
Non-convex problems
Non-convex problems
Linear Algebra & Optimization
References

Nomenclature
A toy optimization problem

\[
\begin{align*}
\min_{x} & \quad 0.9x_1^2 - 0.4x_1x_2 - 0.6x_2^2 - 6.4x_1 - 0.8x_2 \\
\text{s.t.} & \quad -1 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 3.
\end{align*}
\]

- **Feasible set** in light blue.
- 0.1- **suboptimal set** in darker blue.
- **Unconstrained minimizer**: \(x_0\); optimal point: \(x^*\).
- **Level sets** of objective function in red lines.
- A **sub-level set** in red fill.
Other standard forms

**Equality constraints.** We may single out equality constraints, if any:

\[
\min_{x} f_0(x) \text{ subject to } h_i(x) = 0, \ i = 1, \ldots, p,
\]
\[
f_i(x) \leq 0, \ i = 1, \ldots, m,
\]

where \( h_i \)'s are given. Of course, we may reduce the above problem to the standard form above, representing each equality constraint by a pair of inequalities.

**Abstract forms.** Sometimes, the constraints are described abstractly via a set condition, of the form \( x \in X \) for some subset \( X \) of \( \mathbb{R}^n \). The corresponding notation is

\[
\min_{x \in X} f_0(x).
\]
Minimization vs. maximization

Some problems come in the form of maximization problems. Such problems are readily cast in standard form via the expression

$$\max_{x \in \mathcal{X}} f_0(x) = - \min_{x \in \mathcal{X}} g_0(x),$$

where $g_0 := -f_0$.

- **Minimization** problems correspond to loss, cost or risk minimization.
- **Maximization** problems typically correspond to utility or return (e.g., on investment) maximization.
Penalization

A trade-off between two objectives is commonly accomplished via a *penalized* problem:

$$\max_x f(x) + \lambda g(x),$$

where $f$ and $g$ represent loss and risk functions, and $\lambda > 0$ is a risk-aversion parameter.

*Example:* penalized least-squares

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

Here, the risk term $\|w\|_2^2$ controls the variance associated with noise in $X$. 
Robust optimization

Definition

In many instances the problem data is not known exactly. Assume that the functions $f_i$ in the original problem also depend on an “uncertainty” vector $u$ that is unknown, but bounded: $u \in \mathcal{U}$, with the set $\mathcal{U}$ given.

Robust counterpart:

$$\min_x \max_{u \in \mathcal{U}} f_0(x, u)$$

subject to $\forall u \in \mathcal{U}, \ f_i(x, u) \leq 0, \ i = 1, \ldots, m$. 

- Robust counterparts are sometimes tractable.
- If not, systematic procedures exist to generate approximations.
Robust optimization

Geometry

Given $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, consider the constraint in $x \in \mathbb{R}^n$

$$(a + u)^T x \leq b,$$

with $u$'s components are only known within a given set $\mathcal{U}$. The robust counterpart is:

$$\forall u \in \mathcal{U} : (a + u)^T x \leq b.$$
Stochastic optimization

Definition

In stochastic programming, the uncertainty is described by a random variable, with known distribution.

Two-stage stochastic linear program with recourse:

\[
\min_{x \in \mathcal{X}} a^T x + f(x) : \quad f(x) = \mathbb{E}[\min_{w, y \in \mathcal{Y}(x, w)} c(w)^T y].
\]

- \(x\)-variables correspond to decisions taken now.
- \(y\)-variables correspond to decisions taken when uncertainty \(w\) is revealed.

- Stochastic problems are usually very hard.
- Most known approaches are very expensive to solve.
Outline

What is Optimization?
- Definition
- Examples
- Nomenclature
- Other standard forms
- Extensions

The Role of Convexity
- Global vs. local optima
- Convex problems
- Software
- Non-convex problems
- Non-convex problems

Linear Algebra & Optimization

References
Global vs. local minima
The curse of optimization

- Point in red is **globally** optimal (optimal for short).
- Point in green is only **locally** optimal.
- In many applications, we are interested in global minima.

Curse of optimization
Optimization algorithms for general problems can be trapped in local minima.
**Convex function**

**Definition**

A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is **convex** if it satisfies the condition

$$\forall x, y \in \mathbb{R}^n, \lambda \in [0, 1] : f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Geometrically, the graph of the function is “bowl-shaped”.

Convex function.

Non-convex function.
Convexity and local minima

When trying to minimize convex functions, specialized algorithms will always converge to a global minimum, irrespective of the starting point, provided some (weak) assumptions on the function hold.

The Newton algorithm.
Convex optimization

Definition

The problem in standard form

\[ p^* := \min_x f_0(x) \text{ subject to } f_i(x) \leq 0, \quad i = 1, \ldots, m, \]

is convex if the functions \( f_0, \ldots, f_m \) are all convex.

Examples:

- Linear programming \((f_0, \ldots, f_m \text{ affine})\).
- Quadratic programming \((f_0 \text{ convex quadratic}, f_1, \ldots, f_m \text{ affine})\).
- Second-order cone programming \((f_0 \text{ linear, } f_i \text{'s of the form } \|A_i x + b_i\|_2 + c_i^T x + d_i, \text{ for appropriate data } A_i, b_i, c_i, d_i)\).
Software for convex optimization

- Free (if you have matlab): CVX [3], Yalmip, Mosek’s student version [1].
- Commercial: Mosek, CPLEX, etc.
Non-convex problems
Examples

- **Boolean/integer optimization:** some variables are constrained to be Boolean or integers. Convex optimization can be used for getting (sometimes) good approximations.

- **Cardinality-constrained problems:** we seek to bound the number of non-zero elements in a vector variable. Convex optimization can be used for getting good approximations.

- **Non-linear programming:** usually non-convex problems with differentiable objective and functions. Algorithms provide only local minima.

Not all non-convex problems are hard! *e.g.*, low-rank approximation problem.
Outline

What is Optimization?
  Definition
  Examples
  Nomenclature
  Other standard forms
  Extensions

The Role of Convexity
  Global vs. local optima
  Convex problems
  Software
  Non-convex problems
  Non-convex problems

Linear Algebra & Optimization

References
Linear algebra and optimization

It is **very important** to master linear algebra:

- Scalar products, norms.
- Eigenvalues and singular values.

Why is it important?

- Some “simple” optimization problems can be solved via linear algebra.
- Conversely, most linear algebra algorithms actually solve some optimization problems.
- Most optimization algorithms use linear algebra inside the hood.

For background, consult the hyper-textbook:

Food for thought

We are given a set of points in $x_1, \ldots, x_m$ in $\mathbb{R}^n$.

- **Problem 1**: find a line in $\mathbb{R}^n$ such that the average squared distance from the line to the points is minimized.
- **Problem 2**: find a hyperplane in $\mathbb{R}^n$ such that the average squared distance from the hyperplane to the points is minimized.
Outline

What is Optimization?
  Definition
  Examples
  Nomenclature
  Other standard forms
  Extensions

The Role of Convexity
  Global vs. local optima
  Convex problems
  Software
  Non-convex problems
  Non-convex problems

Linear Algebra & Optimization

References
References


