1. Introduction

- mathematical optimization
- least-squares and linear programming
- convex optimization
- example
- course goals and topics
- nonlinear optimization
- brief history of convex optimization

In layman's terms, the mathematical science of Optimization is the study of how to make a good choice when confronted with conflicting requirements. The qualifier *convex* means: when an optimal solution is found, then it is guaranteed to be a best solution; there is no better choice.

Mathematical optimization

(mathematical) optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le b_i, \quad i = 1, \ldots, m \end{array}$$

- $x = (x_1, \ldots, x_n)$: optimization variables
- $f_0 : \mathbf{R}^n \to \mathbf{R}$: objective function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$: constraint functions

optimal solution $x^\star$ has smallest value of f_0 among all vectors that satisfy the constraints

Examples

portfolio optimization

- variables: amounts invested in different assets
- constraints: budget, max./min. investment per asset, minimum return
- objective: overall risk or return variance

device sizing in electronic circuits

- variables: device widths and lengths
- constraints: manufacturing limits, timing requirements, maximum area
- objective: power consumption

data fitting

- variables: model parameters
- constraints: prior information, parameter limits
- objective: measure of misfit or prediction error

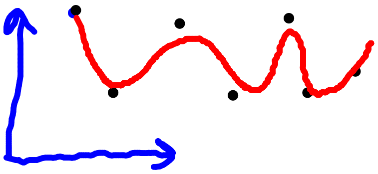
Solving optimization problems

general optimization problem

- very difficult to solve
- methods involve some compromise, *e.g.*, very long computation time, or not always finding the solution

exceptions: certain problem classes can be solved efficiently and reliably

- least-squares problems
- linear programming problems
- convex optimization problems



• coordinates of pts = rows of A

• x = coefficients of the polynomial ($f(a) = \sum_i x_i a^i$)

Least-squares

$$\text{minimize} \quad \|Ax - b\|_2^2$$

$x \in C$ → suppose: I need $\sum_i x_i^2 \le d$

solving least-squares problems

- analytical solution: $x^\star = (A^T A)^{-1} A^T b$ → example of closed form solution
- reliable and efficient algorithms and software
- computation time proportional to $n^2 k$ ($A \in \mathbf{R}^{k \times n}$); less if structured
- a mature technology

using least-squares

- least-squares problems are easy to recognize
- a few standard techniques increase flexibility (*e.g.*, including weights, adding regularization terms)

Linear programming

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

solving linear programs

- no analytical formula for solution
- reliable and efficient algorithms and software
- computation time proportional to $n^2 m$ if $m \ge n$; less with structure
- a mature technology

using linear programming

- not as easy to recognize as least-squares problems
- a few standard tricks used to convert problems into linear programs (*e.g.*, problems involving ℓ_1- or ℓ_∞-norms, piecewise-linear functions)

Convex optimization problem

For LP $A^T x \le b$

$$\left\{\begin{array}{l} f_1(x) \le b_1 \\ f_2(x) \le b_2 \\ \vdots \\ f_m(x) \le b_m \end{array}\right.$$

$$\begin{array}{ll} \text{minimize} & \boxed{f_0(x)} \\ \text{subject to} & \boxed{f_i(x) \le b_i,} \quad i = 1, \dots, m \end{array}$$

$\|Ax - b\|^2$ for Least squares

$c^T x$ for Linear program (LP)

$A: \begin{bmatrix} a_1 \\ a_2 \\ a_m \end{bmatrix}$

$b: \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

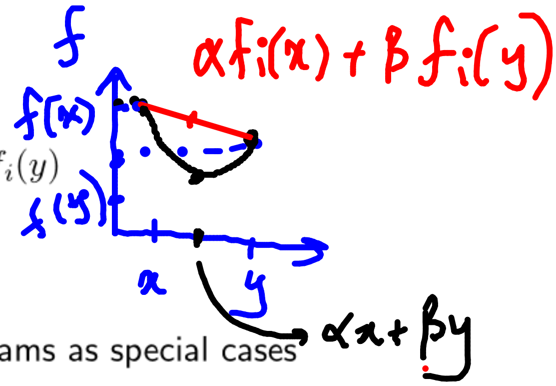
$a_i^T x \le b_i$

- objective and constraint functions are convex:

convex combination

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y)$$

if $\alpha + \beta = 1$, $\alpha \ge 0$, $\beta \ge 0$ $(\alpha, \beta \in [0,1])$



- includes least-squares problems and linear programs as special cases

what we missed: $f_i(\cdot)$ should be defined at each convex combination $\alpha x + \beta y$ ... Domain must have some property which we will call convexity

solving convex optimization problems

- no analytical solution
- reliable and efficient algorithms
- computation time (roughly) proportional to $\max\{n^3, n^2m, F\}$, where F is cost of evaluating f_i's and their first and second derivatives
- almost a technology

using convex optimization

- often difficult to recognize
- many tricks for transforming problems into convex form
- surprisingly many problems can be solved via convex optimization

Brief history of convex optimization

theory (convex analysis): ca1900–1970

algorithms

- 1947: simplex algorithm for linear programming (Dantzig)
- 1960s: early interior-point methods (Fiacco & McCormick, Dikin, . . .)
- 1970s: ellipsoid method and other subgradient methods
- 1980s: polynomial-time interior-point methods for linear programming (Karmarkar 1984)
- late 1980s–now: polynomial-time interior-point methods for nonlinear convex optimization (Nesterov & Nemirovski 1994)

applications

- before 1990: mostly in operations research; few in engineering
- since 1990: many new applications in engineering (control, signal processing, communications, circuit design, . . .); new problem classes (semidefinite and second-order cone programming, robust optimization)

Convex optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & \underbrace{f_i(x) \le b_i}_{x \in C}, \quad i = 1, \ldots, m \end{array}$$

- objective and constraint functions are convex:

Initial part of course

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y)$$

if $\alpha + \beta = 1$, $\alpha \ge 0$, $\beta \ge 0$

- includes least-squares problems and linear programs as special cases

Convex analysis: Calculus of inequalities

Convex geometry is easiest of geometries

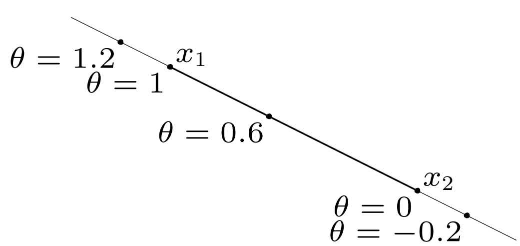
2. Convex sets

- affine and convex sets
- some important examples
- operations that preserve convexity
- generalized inequalities
- separating and supporting hyperplanes
- dual cones and generalized inequalities

Affine set

line through x_1, x_2: all points

$$x = \theta x_1 + (1 - \theta) x_2 \qquad (\theta \in \mathbf{R})$$



affine set: contains the line through any two distinct points in the set

example: solution set of linear equations $\{x \mid Ax = b\}$ insight from linear algebra on geometry etc?

(conversely, every affine set can be expressed as solution set of system of linear equations)

For answer: pages 145 to 181 of

http://www.cse.iitb.ac.in/~cs709/notes/LinearAlgebra.pdf

Rank of A

r=m=n	r=m<n	r=n<m
R=I	R=[I F]	R=[I 0]^T
Unique solution	Infinitely many solutions	0 or 1 solution

Figure 3.3: Summary of the properties of the solutions to the system of equations $Ax = b$.

$A_{m \times n}$ matrix

$$A\mathbf{x}_{complete} = A(\mathbf{x}_{particular} + \mathbf{x}_{nullspace}) = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Example:

$$A\mathbf{x} = \begin{bmatrix} 1 & 2 & 2 & 2 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 8 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b$$

(Gauss Elimination)

$$[A\ \mathbf{b}] = \begin{bmatrix} 1 & 2 & 2 & 2 & b_1 \\ 2 & 4 & 6 & 8 & b_2 \\ 3 & 6 & 8 & 10 & b_3 \end{bmatrix} \overset{E_{2,1},E3,1}{\Longrightarrow} \begin{bmatrix} [1] & 2 & 2 & 2 & b_1 \\ 0 & 0 & [2] & 4 & b_2 - 2b_1 \\ 0 & 0 & 2 & 4 & b_3 - 3b_1 \end{bmatrix}$$

$$\overset{E3,2}{\Longrightarrow} \begin{bmatrix} [1] & 2 & 2 & 2 & b_1 \\ 0 & 0 & [2] & 4 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 - b_1 - b_2 \end{bmatrix}$$

condition for solvability: $b_3 - b_1 - b_2 = 0$

Thus:

The system of equations $A\mathbf{x} = \mathbf{b}$ is solvable when $\mathbf{b}$ is in the column space $C(A)$.

Another way of describing solvability is:

The system of equations $A\mathbf{x} = \mathbf{b}$ is solvable if a combination of the rows of A produces a zero row, the requirement on $\mathbf{b}$ is that the same combination of the components of $\mathbf{b}$ has to yield zero.

Steps to find $x_{particular}$:

1. $\mathbf{x}_{particular}$[2]: Set all free variables (corresponding to columns with no pivots) to 0. In the example above, we should set $x_2 = 0$ and $x_4 = 0$.
2. Solve $A\mathbf{x} = \mathbf{b}$ for pivot variables.

In this example:

$$x_1 + 2x_3 = b_1 \quad \& \quad 2x_3 = b_2 - 2b_1$$

$$\Rightarrow \quad \mathbf{x}_{particular} = \begin{bmatrix} b_2 + 3b_1 \\ 0 \\ \frac{b_2 - 2b_1}{2} \\ 0 \end{bmatrix}$$

Now: $\mathbf{x}_{complete} = \mathbf{x}_{particular} + \mathbf{x}_{nullspace}$ since $x_{nullspace}$ is s.t. $Ax_{nullspace} = 0$

$$\boxed{A\mathbf{x}_{complete} = A(\mathbf{x}_{particular} + \mathbf{x}_{nullspace}) = \mathbf{b} + \mathbf{0} = \mathbf{b}}$$

Eg: if we choose $\mathbf{b} = [5\ 1\ 6]^T$, we get $\mathbf{x}_{particular} = \begin{bmatrix} -2 \\ 0 \\ \frac{3}{2} \\ 0 \end{bmatrix}$ &

$$x_{complete} = \underbrace{\begin{bmatrix} -2 \\ 0 \\ \frac{3}{2} \\ 0 \end{bmatrix}}_{x_{particular}} + \underbrace{\begin{bmatrix} -2 & 2 \\ 1 & 0 \\ 0 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}_{x_{nullspace}} \quad (3.36)$$

Show that $x_{complete} = \theta x_1 + (1-\theta) x_2$ for some $x_1, x_2 \in R^4$ & $\theta \in \mathbb{R}$

$\xRightarrow[\text{that}]{\text{Proves}}$ $\{x \mid Ax = b\}$ is an affine set

Q: What is a more generalised definition of affine sets?