Optimization of Communication Systems
Lecture 1: Introduction and Convexity

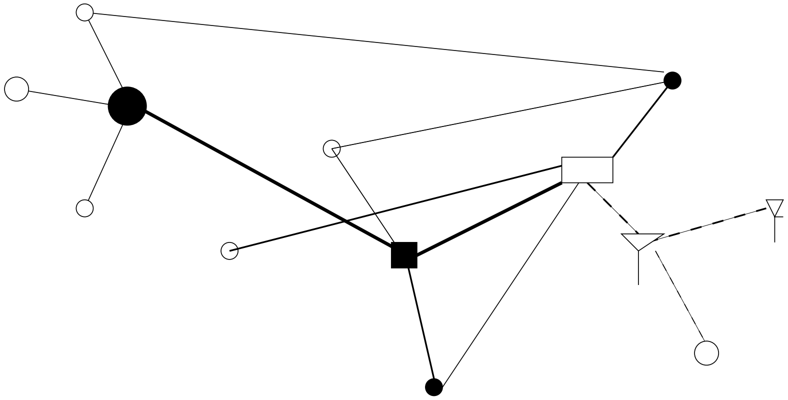
Professor M. Chiang
Electrical Engineering Department, Princeton University

Chinese University of Hong Kong
August 9, 2004

Lecture Outline

- Communication systems
- Optimization: theory, algorithm, mentality
- Convex sets
- Convex functions

Communication Systems



How to send information over a communication medium?

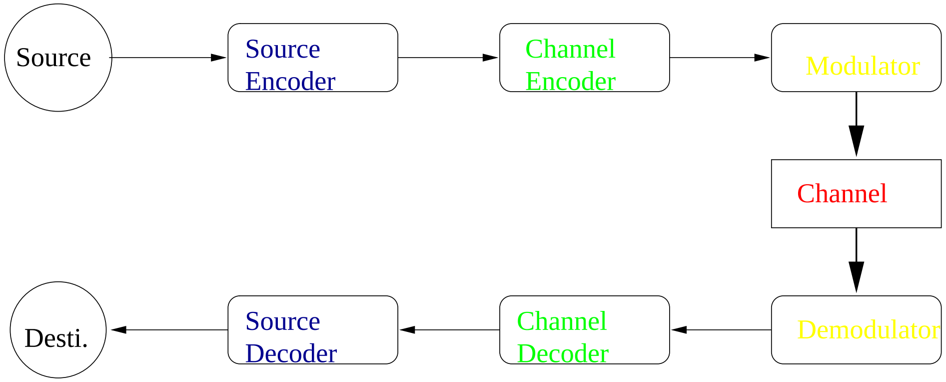
Divide and conquer: break the overall big problem into smaller ones with standardized interfaces

Each layer provides a service to upper layers and utilizes the services provided by lower layers

Questions

- How to meet the requirements from the applications of the information (like accuracy, throughput, latency, jittering, mobility support...)?
- How to represent and use the information?
- How to utilize the communication medium?
- How to connect users?
- How to reach one point from another?
- How to coordinate among the transmitters and receivers?
- How to regulate competition among users?
- How to make the system robust to failures, attacks, variations, growth across space and over time?

Point-to-Point Communication Channel


Source

Source Encoder

Channel Encoder

Modulator

Channel

Demodulator

Channel Decoder

Source Decoder

Desti.

Compress analog signals into digital data

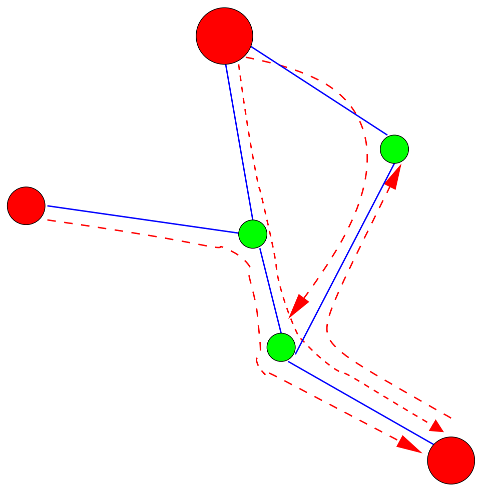
Add redundancy to protect against channel impairments

Map digital data onto physical waveforms suitable for the medium

Questions

- How to describe the channel and estimate its characteristics (twisted pair, coaxial cable, optic fiber, radio, acoustic, storage)?
- How fast can data be sent reliably?
- How to compress signals?
- How to add redundancy to compensate for noise (thermal noise, impulse noise ...), interference (from other users, from reflections, among symbols) ...
- How to use the communication resources (time, frequency, engineering design parameters) efficiently?
- What happens when multiple transmitters send data to multiple receivers?

Communication Networks



Not necessarily a direct link, but a networked communication system

Questions on last slide remain, plus more questions (and opportunities)

Questions

- Fixed or dynamic topology? Who are transceivers and who are relays?
- Direct link or switched architecture? Circuit switch or packet switch or something else?
- How to divide into (possibly different types of) subnetworks?
- End-to-end control or hop-by-hop control?
- How to get on the communication medium?
- How to get from one point to another?
- How to monitor and adjust overall state of the network?
- How to ensure accurate, secure, dependable, timely, and usable transfer of information across space among competing users?

Model, Analysis and Design

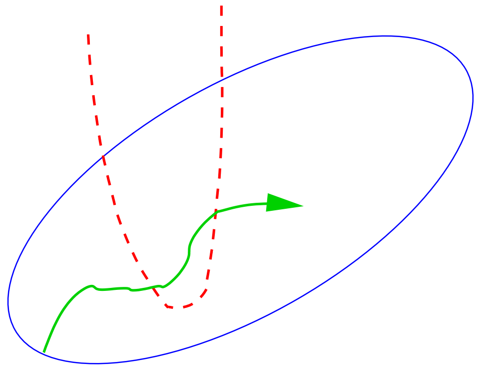
Empirical data from field trials

Computer simulations

Analytic tools

- Mathematical models of networks
- Information theory, coding theory, communication theory
- Digital signal processing algorithms
- Queuing theory and other probabilistic tools
- Systems control theory, graph theory, game theory, economics modelling, physics/biology modelling...
- Optimization theory

Optimization



$$\begin{array}{ll}\text{minimize} & f(x)\\ \text{subject to} & x \in C\end{array}$$

Optimization variables: x. Constant parameters describe objective function f and constraint set C

Questions

- How to describe the constraint set?
- Can the problem be solved globally and uniquely?
- What kind of properties does it have? How does it relate to another optimization problem?
- Can we numerically solve it in an efficient, robust, and distributed way?
- Can we optimize multiple objectives simultaneously?
- Can we optimize over a sequence of time instances?
- Can we find the problem for a given solution?

Applications

Theory and algorithms of optimization are extremely powerful:

- Communication systems
- Other information science areas: signal/image/video processing, systems control, algorithms, graphics, data analysis, theoretical computer science ...
- Other engineering disciplines: aerospace, mechanical, chemical, civil, transportation, computer architecture, analog circuit design ...
- Physics, chemistry, biology ...
- Economics, finance, management ...
- Analysis, probability, statistics, differential equations ...

Methodologies

Widely known: linear programming is powerful and easy to solve

Modified view: watershed between easy and hard optimization problems is not linearity, but convexity

- Local optimality is also global optimality
- Lagrange duality theory well developed
- Know a lot about the problem and solution structures
- Efficiently compute the solutions numerically

Need to know how to recognize and formulate convex optimization, and use the recently developed tools to solve your problem (an objective of this course)

Active research area with many exciting recent and ongoing developments, and other challenges (discrete optimization, nonconvex problems, robust and distributed algorithms...)

Optimization of Communication Systems

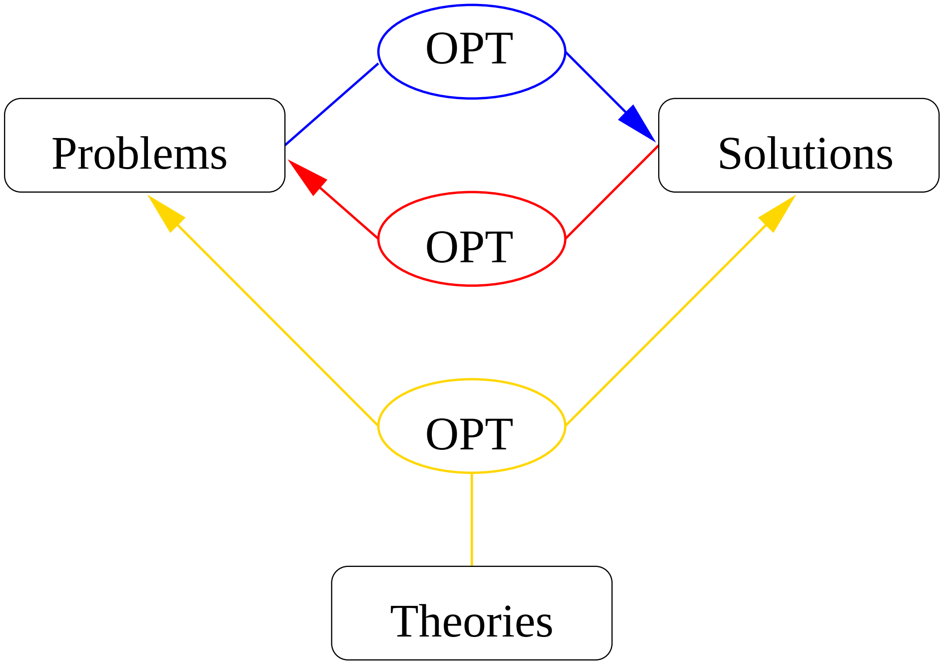
Three meanings of 'optimization of communication systems':

- Formulate the problem as an optimization problem
- Interpret a given solution as an optimizer/algorithm for an optimization problem
- Extend the underlying theory by optimization theoretic techniques

A remarkably powerful, versatile, widely applicable and not yet fully recognized viewpoint

Applications in communication systems also stimulate new developments in optimization theory and algorithms

Optimization of Communication Systems



What This Course Is About

How problems in communication systems can be formulated and solved as optimization problems

- Classic results (starting with 1940s)
- Current research (papers being published as we speak)

Applications topics:

Information theory problems, transmitter and receiver design, channel decoding, detection and estimation, multiple antenna beamforming, network resource allocation and utility maximization, optical network topology design, wireless power control and medium access, network flow problems, IP routing, TCP congestion control, cross layer design

Methodology topis:

Linear programming, convex optimization, quadratic programming, geometric programming, integer programming, robust optimization, Pareto optimization, dynamic programming, Lagrange duality, KKT optimality conditions, gradient methods, interior point methods, distributed algorithms

What This Course Is Not About

- Not a math course on convex analysis (not many rigorous proofs)
- Not an OR course on nonlinear optimization (only basic optimization/algorithm topics)
- Not an EE course on digital communication (cover only selected topics)
- Not a EE/CS course on networking (cover only selected topics)
- Not a CS course on algorithms (little computational complexity analysis)

‘Just enough’ background materials presented ‘just-in-time’

Acknowledgements

The first course devoted to systematic treatment of the subject

Course materials drawn from a variety of sources (many textbooks, a number of recent journal/conference papers, ongoing research projects...) and distilled into a common framework

Jointly developed with Professor Steven Low at Caltech (netlab.caltech.edu)

Many thanks to many people, particulary

- Professor Stephen Boyd (Stanford)
- Professor Tom Luo (U. Minnesota)
- Professor Wei Yu (U. Toronto)

Books and Papers

- M. Chiang, *ELE539 Lecture Notes 2004* Does not contain all the information, and complemented by a lot of discussion and graphs in class
- S. Boyd and L. Vandenberghe, *Convex Optimization* Cambridge University Press 2004. Free download from www.stanford.edu/~boyd/cvxbook.html
- Recent papers

Week 1

August 9

- Communication systems and optimization mentality
- Convex set and convex functions
- Convex optimization and Lagrange duality

August 11

- LP
- Network flow problems

August 13

- QP and GP
- Basic information theory and resource allocation problems

Week 2

August 18

- Network rate allocation and utility maximization
- TCP congestion control

August 19

- Advances in utility maximization: Internet
- Advances in utility maximization: wireless networks
- Layering as optimization decomposition

August 20

- SDP
- Detection and estimation problems

Week 3

August 23

- Numerical algorithms: gradient and Newton's methods
- Numerical algorithms: Interior point methods

August 25

- Wireless MIMO transceiver design
- DSL spectrum management and generalized waterfilling

August 27

- DP and applications
- Integer constrained, nonconvex optimization, and applications

Second Half of Lecture 1

Why Does Convexity Matter?

- The watershed between easily solvable problem and intractable ones is not ‘linearity’, but ‘convexity’
- So we’ll start with convex optimization framework, then specialize into different special cases (including linear programming)
- Only covers the very basic concepts and results in convex analysis without proofs

This and next lectures are primarily mathematical, but a wide range of applications will soon follow

Convex Set

Set C is a convex set if the line segment between any two points in C lies in C, ie, if for any $x_1, x_2 \in C$ and any $\theta \in [0, 1]$, we have

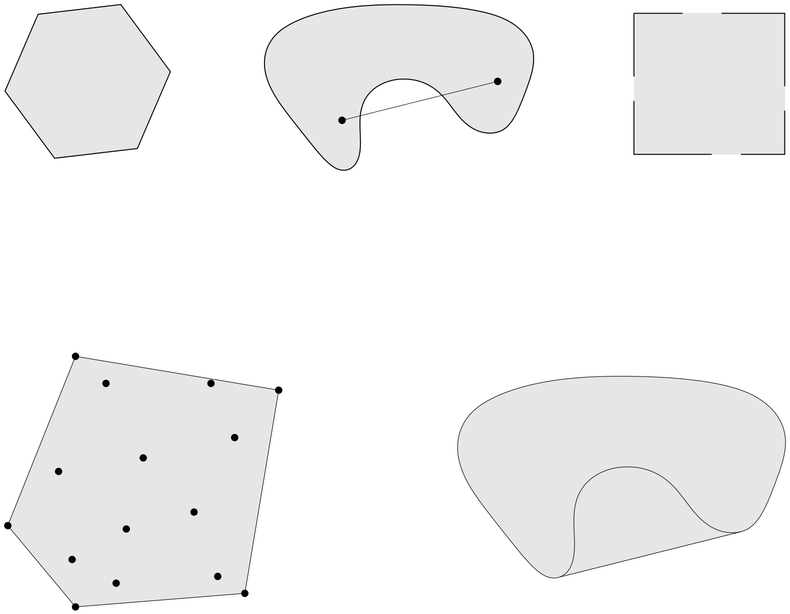
$$\theta x_1 + (1 - \theta) x_2 \in C$$

Convex hull of C is the set of all convex combinations of points in C:

$$\left\{ \sum_{i=1}^{k} \theta_i x_i | x_i \in C, \theta_i \geq 0, i = 1, 2, \ldots, k, \sum_{i=1}^{k} \theta_i = 1 \right\}$$

Can generalize to infinite sums and integrals

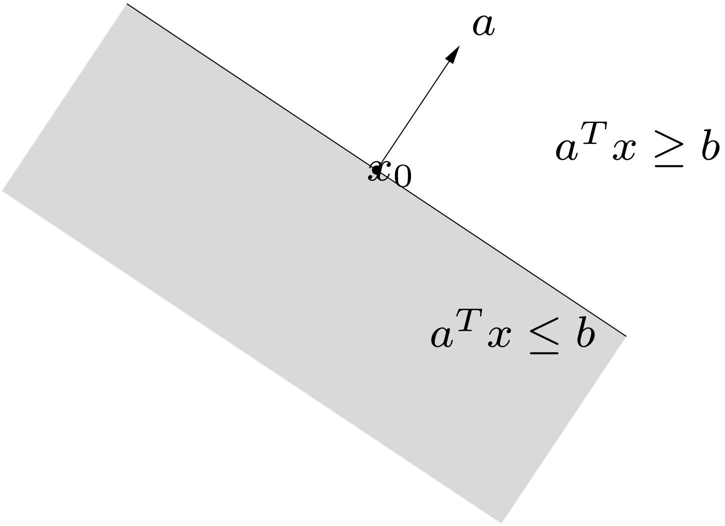
Examples



Examples of Convex Sets

- Hyperplane in $\mathbf{R}^n$ is a set: $\{x | a^T x = b\}$ where $a \in \mathbf{R}^n, a \neq 0, b \in \mathbf{R}$

Divides $\mathbf{R}^n$ into two halfspaces: eg, $\{x | a^T x \leq b\}$ and $\{x | a^T x > b\}$



- Polyhedron is the solution set of a finite number of linear equalities and inequalities (intersection of finite number of halfspaces and hyperplanes)

Examples of Convex Sets

- Euclidean ball in $\mathbf{R}^n$ with center x_c and radius r:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x_c + ru \mid \|u\|_2 \leq 1\}$$

Verify its convexity by triangle inequality

- Generalize to ellipsoids:

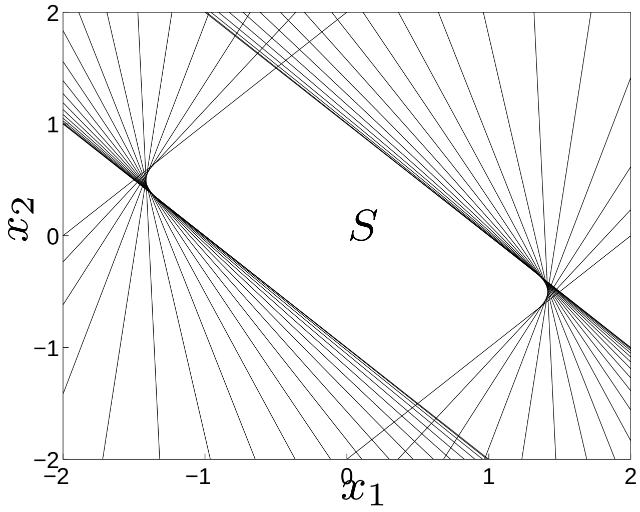
$$\mathcal{E}(x_c, P) = \left\{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\right\}$$

P: symmetric and positive definite. Lengths of semi-exes of $\mathcal{E}$ are $\sqrt{\lambda_i}$ where λ_i are eigenvalues of P

Convexity-Preserving Operations

- Intersection.
- Example: $S = \left\{x \in \mathbf{R}^m \,|\, |p(t)| \leq 1 \text{for} |t| \leq \frac{\pi}{3}\right\}$ where $p(t) = \sum_{k=1}^{m} x_k \cos kt$.

Since $S = \bigcap_{|t| \leq \frac{\pi}{3}} S_t$, where $S_t = \left\{x | -1 \leq (\cos t, \ldots, \cos mt)^T x \leq 1\right\}$, S is convex



Convexity-Preserving Operations

- Linear-fractional functions: $f : \mathbf{R}^n \to \mathbf{R}^m$:

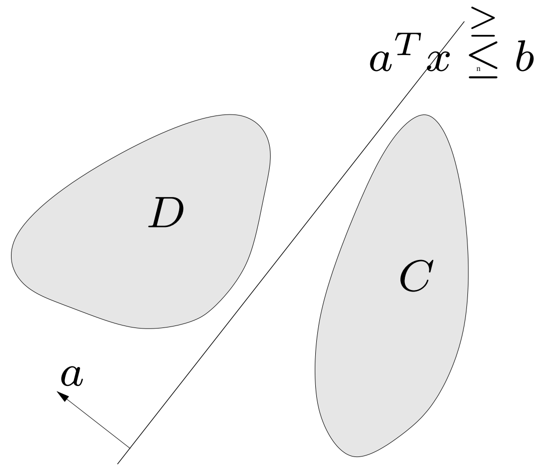
$$f(x) = \frac{Ax + b}{c^T x + d}, \quad \mathbf{dom}\, f = \{x | c^T x + d > 0\}$$

- If set C in $\mathbf{dom}\, f$ is convex, image $f(C)$ is also convex set
- Example: $p_{ij} = \mathbf{Prob}(X = i, Y = j), q_{ij} = \mathbf{Prob}(X = i | Y = j)$. Since

$$q_{ij} = \frac{p_{ij}}{\sum_k p_{kj}},$$

if C is a convex set of joint prob. for (X, Y), the resulting set of conditional prob. of X given Y is also convex

Separating Hyperplane Theorem



- C and D: non-intersecting convex sets, *i.e.*, $C \bigcap D = \phi$. Then there exist $a \neq 0$ and b such that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$.

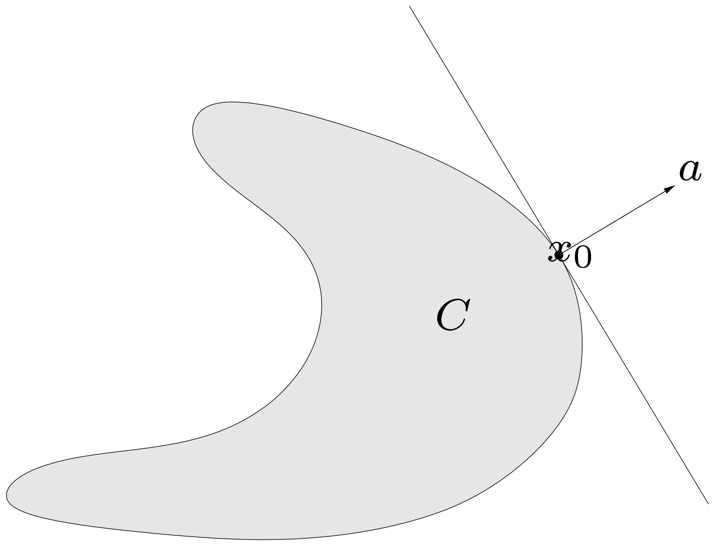
- Application: Theorem of alternatives for strict linear inequalities:

$$Ax \prec b$$

are infeasible if and only if there exists $\lambda \in \mathbf{R}^m$ such that

$$\lambda \neq 0, \;\; \lambda \succeq 0, \;\; A^T\lambda = 0, \;\; \lambda^T b \leq 0.$$

Supporting Hyperplane Theorem



- Given a set $C \in \mathbf{R}^n$ and a point x_0 on its boundary, if $a \neq 0$ satisfies $a^T x \leq a^T x_0$ for all $x \in C$, then $\{x | a^T x = a^T x_0\}$ is called a supporting hyperplane to C at x_0
- For any nonempty convex set C and any x_0 on boundary of C, there exists a supporting hyperplane to C at x_0

Convex Functions

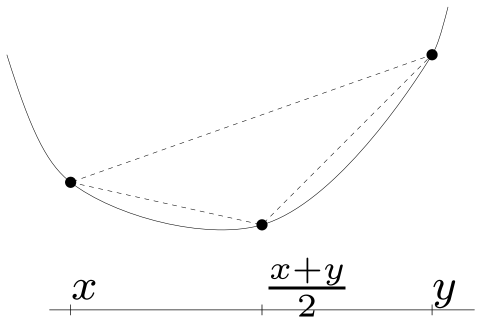
$f : \mathbf{R}^n \to \mathbf{R}$ is a convex function if $\mathbf{dom}\, f$ is a convex set and for all $x, y \in \mathbf{dom}\, f$ and $\theta \in [0, 1]$, we have

$$f(\theta x + (1 - \theta)y) \geq \theta f(x) + (1 - \theta)f(y)$$

f is strictly convex if strict inequality above for all $x \neq y$ and $0 < \theta < 1$

f is concave if $-f$ is convex

- Affine functions are convex and concave

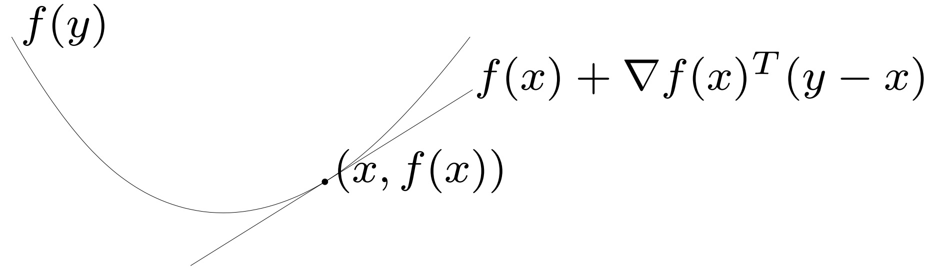


Conditions of Convex Functions

1. For differentiable functions, f is convex iff

$$f(y) - f(x) \geq \nabla f(x)^T (y - x)$$

for all $x, y \in \mathbf{dom}\, f$, and $\mathbf{dom}\, f$ is convex



- $f(y) \geq \tilde{f}(x)$ where $\tilde{f}(x)$ is first order Taylor expansion of f at x.
- Local information (first order Taylor approximation) about a convex function provides global information (global underestimator).
- If $\nabla f(x) = 0$, then $f(y) \geq f(x)$, $\forall y$, thus x is a global minimizer of f

Conditions for Convex Functions

2. For twice differentiable functions, f is convex iff

$$\nabla^2 f(x) \succeq 0$$

for all $x \in \mathbf{dom}\, f$ (upward slope) and $\mathbf{dom}\, f$ is convex

3. f is convex iff for all $x \in \mathbf{dom}\, f$ and all v,

$$g(t) = f(x + tv)$$

is convex on its domain $\{t \in \mathbf{R} | x + tv \in \mathbf{dom}\, f\}$

Examples of Convex or Concave Functions

- e^{ax} is convex on $\mathbf{R}$, for any $a \in \mathbf{R}$
- x^a is convex on $\mathbf{R}_{++}$ when $a \geq 1$ or $a \leq 0$, and concave for $0 \leq a \leq 1$
- $|x|^p$ is convex on $\mathbf{R}$ for $p \geq 1$
- $\log x$ is concave on $\mathbf{R}_{++}$
- $x \log x$ is strictly convex on $\mathbf{R}_{++}$
- Every norm on $\mathbf{R}^n$ is convex
- $f(x) = \max\{x_1, \ldots, x_n\}$ is convex on $\mathbf{R}^n$
- $f(x) = \log \sum_{i=1}^n e^{x_i}$ is convex on $\mathbf{R}^n$
- $f(x) = \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}$ is concave on $\mathbf{R}^n_{++}$

Convexity-Preserving Operations

- $f = \sum_{i=1}^{n} w_i f_i$ convex if f_i are all convex and $w_i \geq 0$

- $g(x) = f(Ax + b)$ is convex iff $f(x)$ is convex

- $f(x) = \max\{f_1(x), f_2(x)\}$ convex if f_i convex, *e.g.*, sum of r largest components is convex

- $f(x) = h(g(x))$, where $h : \mathbf{R}^k \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}^k$.

If $k = 1$: $f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$. So

f is convex if h is convex and nondecreasing and g is convex, or if h is convex and nonincreasing and g is concave ...

- $g(x) = \inf_{y \in C} f(x, y)$ is convex if f is convex and C is convex

- $g(x, t) = tf(x/t), x \in \mathbf{R}^n, t \in \mathbf{R}$ is convex if f is convex

Conjugate Function

Given $f : \mathbf{R}^n \to \mathbf{R}$, conjugate function $f^* : \mathbf{R}^n \to \mathbf{R}$ defined as:

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$

with domain consisting of $y \in \mathbf{R}^n$ for which the supremum if finite

- $f^*(y)$ always convex: it is the pointwise supremum of a family of affine functions of y
- Fenchel's inequality: $f(x) + f^*(y) \geq x^T y$ for all x, y (by definition)
- $f^{**} = f$ if f is convex and closed

Useful for Lagrange duality theory

Examples of Conjugate Functions

- $f(x) = ax + b,\ f^*(a) = -b$
- $f(x) = -\log x,\ f^*(y) = -\log(-y) - 1$ for $y < 0$
- $f(x) = e^x,\ f^*(y) = y \log y - y$
- $f(x) = x \log x,\ f^*(y) = e^{y-1}$
- $f(x) = \frac{1}{2} x^T Q x,\ f^*(y) = \frac{1}{2} y^T Q^{-1} y$ (Q is positive definite)
- $f(x) = \log \sum_{i=1}^{n} e^{x_i},\ f^*(y) = \sum_{i=1}^{n} y_i \log y_i$ if $y \succeq 0$ and $\sum_{i=1}^{n} y_i = 1$ ($f^*(y) = \infty$ otherwise)

Log-concave Functions

$f : \mathbf{R}^n \to \mathbf{R}$ is log-concave if $f(x) > 0$ and $\log f$ is concave

Many probability distributions are log-concave:

- Cumulative distribution function of Gaussian density
- Multivariate normal distribution
- Exponential distribution
- Uniform distribution
- Wishart distribution

Summary

- Definitions of convex sets and convex functions
- Convexity-preserving operations
- Global information from local characterization: Support Hyperplane Theorem
- Convexity is the watershed between 'easy' and 'hard' optimization problems. Recognize convexity. Utilize convexity.

Readings: Section 2.1-2.3, 2.5, and 3.1-3.3 in Boyd and Vandenberghe