Convex Optimization with Engineering Applications

Mikael Johansson
EE/Automatic Control
mikaelj@ee.kth.se

Anders Forsgren
Math/OptSyst
andersf@kth.se

Mathematical optimization

Mathematical *optimization problem*

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \end{array}$$

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

optimal solution $x^\star$ has smallest value of $f_0(x)$ satisfying constraints

Examples

Portfolio optimization

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

Device sizing in electronic circuits

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

…and many many more

Our focus will be on problems in communications and control…

Solving optimization problems

general optimization problems

- very difficult to solve (unless special structure)
- "long computation times or not sure to find optimal solution"

exceptions: certain problem classes can be solved efficiently

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

Least-squares

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

solving least-squares problems

- analytical solution $x^\star = (A^TA)^{-1}A^Tb$
- reliable and efficient algorithms and software
- computation time proportional to n^2k ($A \in \mathbb{R}^{k \times n}$)
- a mature technology

using least-squares

- least-squares problem easy to recognize
- a few standard techniques increase flexibility (weights, regularization)

Linear programming

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

solving linear programming problems

- no analytical formula for solution
- reliable and efficient algorithms and software
- computation time proportional to nm^2 (if $m \geq n$)
- 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

Convex optimization problems

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

- objective and constraint functions are convex

$$f_i(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y) \qquad \text{for all } \theta \in [0,1]$$

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

Convex optimization

solving convex optimization problems

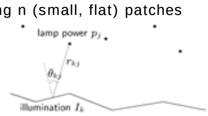
- no analytical solution
- reliable and efficient algorithms
- computation time roughly $\max\{n^3, n^2 m, F\}$
- almost a technology

using convex optimization

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

Example

m lamps illuminating n (small, flat) patches



intensity I_k at patch k depends linearly on lamp powers p_j

$$I_k = \sum_{j=1}^{m} a_{kj} p_j, \quad a_{kj} = r_{kj}^{-2} \max\{\cos\theta_{kj}, 0\}$$

problem: achieve desired illumination I_{des} with bounded lamp powers

$$\begin{array}{ll}\text{minimize} & \max_{k=1,\dots,n} |\log I_k - \log I_{\text{des}}|\\ \text{subject to} & 0 \le p_j \le p_{\max}, \quad j = 1, \dots, m\end{array}$$

Solution approaches

1. Use uniform power $p_j = p$ vary p
2. Use least-squares:
   $$\text{minimize} \quad \textstyle\sum_{k=1}^{n} (I_k - I_{\text{des}})^2$$
   round p_j if $p_j > p_{max}$ or $p_j < 0$
3. Use weighted least-squares
   $$\text{minimize} \quad \textstyle\sum_{k=1}^{n} (I_k - I_{\text{des}})^2 + w_k (P_k - P_{\max}/2)^2$$
   iteratively adjust weights w_j until $0 \le p_j \le p_{\max}$
4. Use linear programming to solve
   $$\begin{array}{ll}\text{minimize} & \max_{k=1,\dots,n} |I_k - I_{\text{des}}|\\ \text{subject to} & 0 \le p_j \le p_{\max} \quad j = 1, \dots, m\end{array}$$

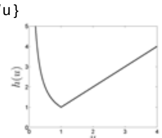
of course, these are all approximate (suboptimal) 'solutions'

Solution approaches (cont'd)

5. Use convex optimization. Problem is equivalent to

$$\begin{array}{ll}\text{minimize} & f_0(p) = \max_{k=1,\dots,n} h(I_k / I_{\text{des}})\\ \text{subject to} & 0 \le p_j \le p_{\max}, \quad j = 1, \dots, m\end{array}$$

with $h(u) = \max\{u, 1/u\}$



f_0 is convex, since maximum of convex functions is convex

Exact solution obtained with effort ≈ modest factor × least-squares effort

Extensions: simple or hard?

Does constraint 1. or 2. complicate the problem?

1. no more than half of total power is in any 10 lamps
2. no more than half of the lamps are on ($p_j > 0$)

Answer: with (1), still easy to solve; with (2), extremely difficult.

Moral: (untrained) intuition doesn't always work; without proper background, very easy problem can appear quite similar to very difficult problems.

Key message I

"The great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity"

R. Rockafellar, SIAM Review 1993.

Recognize and utilize convexity!

Key message II

A wide range of design *problems in engineering* can be formulated and solved as *convex optimization* problems.

Examples:

- routing in data networks
- Internet congestion control
- model-predictive control
- analysis and design of linear control systems
- filter design
- multi-antenna beamforming
- power allocation in cellular systems
- cross-layer optimization of wireless networks

⋮

Example: power allocation in wireless

Wireless system with L transmitter-receiver pairs (communication links)

Reliable communication if signal-to-interference and noise ratios

$$\gamma_l(P) = \frac{G_{ll}P_l}{\sigma_l + \sum_{j\neq l} G_{lj}P_j}$$

exceed target values γ_{tgt}.

Problem: Find minimum power allocation that meets SINR constraints

$$\begin{array}{lll} \text{minimize} & \sum_{l=1}^{L} P_l & \\ \text{subject to} & \gamma_l(P) \geq \gamma_{\text{tgt}} & l = 1,\dots,L \\ & 0 \leq P_l \leq P_{\max} & l = 1,\dots,L \end{array}$$

Problem can be formulated, solved and analyzed as a linear program.

Example: model predictive control (MPC)

Discrete-time linear system

$$x(t+1) = Ax(t) + Bu(t) \qquad x(0) = x_0$$

find control sequence $U=(u_0,u_1, \dots, u_K)$ that minimizes loss function

$$J(x,u) = \sum_{k=1}^{K} x_k^T Q x_k + u_k^T R u_k$$

while obeying control and state constraints

$$Cx_k + Du_k \leq \eta \qquad k = 1,\dots,K$$

MPC via quadratic programming

Using the fact that

$$x(t) = A^t x_0 + A^{t-1}Bu_0 + \cdots + ABu_{t-2} + Bu_{t-1} = H_t \begin{bmatrix} x_0 \\ u \end{bmatrix}$$

we can eliminate state variables and re-write problem as

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} x_0 \\ u \end{bmatrix}^T \begin{bmatrix} Q_e & S_e^T \\ S_e & R_e \end{bmatrix} \begin{bmatrix} x_0 \\ u \end{bmatrix} \\ \text{subject to} & C_e x_0 + D_e u \leq \eta_e \end{array}$$

Convex *quadratic programming* problem in u – efficiently solved!

Example: explicit MPC

By convex duality (Lecture 5), u is optimal if and only if (u,λ) satisfies

$$\begin{array}{rl} R_e u + S_e x_0 + D_e^T \lambda = 0 & \\ \lambda^{(i)}(C_e^{(i)} x_0 + D_e^{(i)} u - \eta_e^{(i)}) = 0 & i = 1,\dots,M \\ \lambda \geq 0 & \\ C_e x_0 + D_e u - \eta_e \leq 0 & \end{array}$$

Optimal control is locally affine in x_0,

$$u = -R_e^{-1}(S_e x_0 + D_e^T \lambda)$$

within a polyhedron defined by the other optimality constraints.

MPC-equivalent feedback is piecewise linear, can be computed off-line!

Example: L_2 gain analysis

Closed loop linear system with disturbance w and output y

$$\dot{x} = Ax + Bw \qquad y = Cx \qquad x(0) = 0$$

has L_2 gain γ if, for all (square integrable) w and all $t \geq 0$,

$$\int_{\tau=0}^{t} y(\tau)^T y(\tau)\, d\tau \leq \gamma^2 \int_{\tau=0}^{t} w(\tau)^T w(\tau)\, d\tau$$

Alternative characterization: non-negative storage function $V(x) \geq 0$

$$\frac{\partial V}{\partial x}(Ax + Bw) \leq \gamma^2 w^T w - y^T y$$

How to find storage function? How to find the best one (smallest γ)?

L_2 gain analysis via convex optimization

Fact: for linear systems, sufficient to consider quadratic V(x)

$$V(x) = x^T P x$$

Gain conditions translate to *linear matrix inequalities* (in P, γ^2)

$$P \succeq 0$$

$$\begin{bmatrix} A^T P + PA + C^T C & PB \\ BP & -\gamma^2 I \end{bmatrix} \preceq 0$$

Convex conditions: can find optimal P and smallest γ efficiently!

Many extensions: controller design, other performances, …

Example: filter design

Finite impulse response (FIR) filter

$$y(t) = \sum_{\tau=0}^{n-1} h_\tau u(t-\tau) \; = \; h_0 u(t) + h_1 u(t-1) + \cdots + h_{n-1} u(t-n+1)$$

Filter frequency response

$$H(\omega) = h_0 + h_1 e^{-i\omega} + \cdots + h_{n-1} e^{-i(n-1)\omega} =$$
$$= \sum_{t=0}^{n-1} h_t \cos t\omega + i \sum_{t=0}^{n-1} h_t \sin t\omega$$

Filter design problem: choose h to optimize specifications!

Chebyshev design via convex optimization

Chebyshev design: minimize deviation from desired freq. response

$$\text{minimize} \quad \max_{\omega \in [0, 2\pi]} |H(\omega) - H_{\text{des}}(\omega)|$$

After discretizing frequency, we can re-write the problem as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & |H(\omega_k) - H_{\text{des}}(\omega_k)| \leq t \qquad k = 1, \ldots, K \end{array}$$

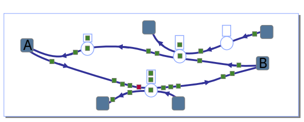
which results in the *second-order cone program*

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|A_k h - b_k\| \leq t \qquad k = 1, \ldots, K \end{array}$$

Convex, hence readily solved!

Example: routing in data networks

Find best routes for forwarding traffic from source to destination



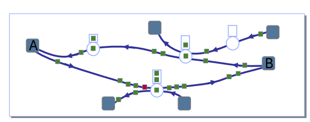
If we view data traffic as a continuous flow, optimal routing solves

$$\begin{array}{lll} \text{minimize} & \sum_{l=1}^{L} f_l(t_l) & \\ \text{subject to} & Ax^{(d)} = s^{(d)} & \text{flow conservation at nodes} \\ & t_l = \sum_{d=1}^{N} x_l^{(d)} & \text{total traffic across links} \\ & t_l \leq c_l & \text{capacity constraints} \\ & x^{(d)} \succeq 0 & \text{traffic is non-negative} \end{array}$$

Convex optimization problem if loss functions $f_l(\cdot)$ are convex

Internet congestion control

Users contend for limited network resources



Congestion control: helps share resources fairly and efficiently

- When router buffer fills up, packets are marked (or dropped)
- Users react to congestion signals by reducing transmission rate

Interplay between control mechanisms in routers and end nodes

Congestion control as utility maximization

Optimal network operation from solving utility maximization problem

$$\begin{array}{lll} \text{maximize} & \sum_{p=1}^{P} u_p(s_p) & \\ \text{subject to} & r_l^T s \le c_l & l = 1, \dots, L \\ & s_p \ge 0 & p = 1, \dots, P \end{array}$$

where u_p is a strictly concave *utility function*

Key observation: the dual problem admits decentralized solution

- links update congestion measures (Lagrange multipliers)
- sources update source rates, in reaction to congestion signals

Gives insight into current (and future) congestion control schemes!

Optimization of engineering systems

Three meanings of "optimization of engineering systems"

- formulate design as an optimization problem
- interpret solution as algorithm for an underlying optimization problem
- extend fundamental theory using optimization-theoretic techniques

A remarkably powerful, versatile and widely applicable viewpoint.

After this course…

…you will be able to

1. Recognize/formulate problems as convex optimization problems
2. Develop numerical codes for problems of moderate size
3. Characterize optimal solution, give limits of performance, …

What we will cover…

Theoretical tools

- convex sets, convex functions, convex optimization, duality

Important classes of convex optimization problems

- linear programming, quadratic programming, second-order cone progamming, geometric programming, semidefinite programming

Numerical algorithms

- Simple algorithms, basic ideas behind interior-point methods, hands-on experience using state-of-the art optimization packages.

Applications

- wireless power control, network flow problems, TCP congestion control, IP routing, cross-layer design, transceiver design, …

…and what we will not

- mathematical proofs, advanced concepts from convex analysis
- non-convex optimization (e.g., integer programming)
- detailed accounts on digital communications, networking, control, …

Administrative issues

Course content and schedule

1. Introduction
2. Convex sets and convex functions
3. Linear programming, Lagrangian relaxation and duality
4. Linear programming, Lagrangian relaxation and duality
5. Convex programming and semidefinite programming
6. Geomtric programming and second-order cone programming
7. Sensitivity and multiobjective optimization
8. Smooth unconstrained minimization
9. Interior methods
10. Decomposition and large-scale optimization
11. Applications in communications and control
12. Applications in communications and control

First part theory and exercises, second part applications and research!

Philosophy of the course

- Can only cover a small sample from a vast research field
  - view the course as a guided tour
- You can only learn by working with the material
  - be active in class, read the book, and do the homeworks!
- Motivation for theory comes from the need to understand
  - use course to help you work through a research problem in your field!

Course variants

Course offered in two basic variants

A four-credit version
- A basic understanding of convex optimization and its applications
- Requires completion of hand-ins and a research paper presentation

An eight-credit version
- In-depth working knowledge of convex optimization
- Additional requirements: take-home exam and short research project
- Only for very well-motivated PhD students

Hand-ins

Four (4) mandatory hand-ins during the course.

Topic	Deadline
Convexity	November 7^{nd}
Convex optimization problems	November 14^{th}
Duality	November 21^{th}
Numerical algorithms	November 30^{rd}

Aim: help you to get started working with the material.

Late homework solutions are *not* accepted.

Research paper presentations

Read and present a research paper, individually or in groups

In 10-15 minutes, try to address the following:

- What engineering problem is being solved? Why is it important?
- How can the problem be formulated as an optimization problem
- Is the problem convex? Use what you learned in course to explain!
- How is the problem solved? Explain any new techniques!
- What type of results are obtained?
- Are there any obvious flaws, or extensions?

Try to teach us what you have learned!

Preliminary presentation date: December 7^{th}, 2006.

Take-home exam

Individually (no cooperation), but with the help of all course material
- Course book, lecture notes, exercises, library, Internet, …

Eight hours should be enough.

Preliminary exam dates: December 8^{th}-15^{th}, 2006.

Research project

The most fun and useful part of the course!

A short optimization project related to your own research
- formulate and solve a research problem using convex optimization
- implement a numerical solver that you could use in your work
- a comparative study of (optimization-based) solutions to a problem

Documented by well-written 5-15 page report (see separate instructions)

Project proposals due: November 23rd, 2006
Final project deadline: January 14th, 2007 (sharp)

Interact (both proposals and project work) with us!

Course book

The course will cover most of the material in

"Convex optimization" by S. Boyd and L. Vandenberghe
Cambridge University Press, 2004
ISBN 0 521 83378 7

The book is available as a (huge) PDF file via
http://www.stanford.edu/~boyd/cvxbook.html
but we recommend you to buy the printed book (it's only £45)

Course software

This year's course will make heavy use of **cvx** under Matlab

cvx: a parser/solver for (a subset of) convex optimization problems

```
m = 20; n = 10; p = 4;
A = randn(m,n); b = randn(m,1);
C = randn(p,n); d = randn(p,1);
cvx_begin
    variable x(n)
    minimize( norm( A * x - b, 2 ) )
    subject to
        C * x == d;
        norm( x, Inf ) <= 0.4;
cvx_end
```

Download from http://www.stanford.edu/~boyd/cvx/

Acknowledgements

Slides borrow heavily from lectures by Stephen Boyd (Stanford)
and Lieven Vandenberghe (UCLA)

http://www.ucla.edu/ee236b/

Some graduate level courses on optimization in communications

- "Optimization of Communications Systems", Princeton
  http://www.princeton.edu/~chiangm/class.html
- "Convex Optimization" at HKUST
  http://www.ece.ust.hk/~palomar/courses/ELEC692Q/
- "Convex optimization of Communication Systems" (short course)
  http://www.s3.kth.se/~mikaelj/