Introduction to Convex Optimization

Chee Wei Tan

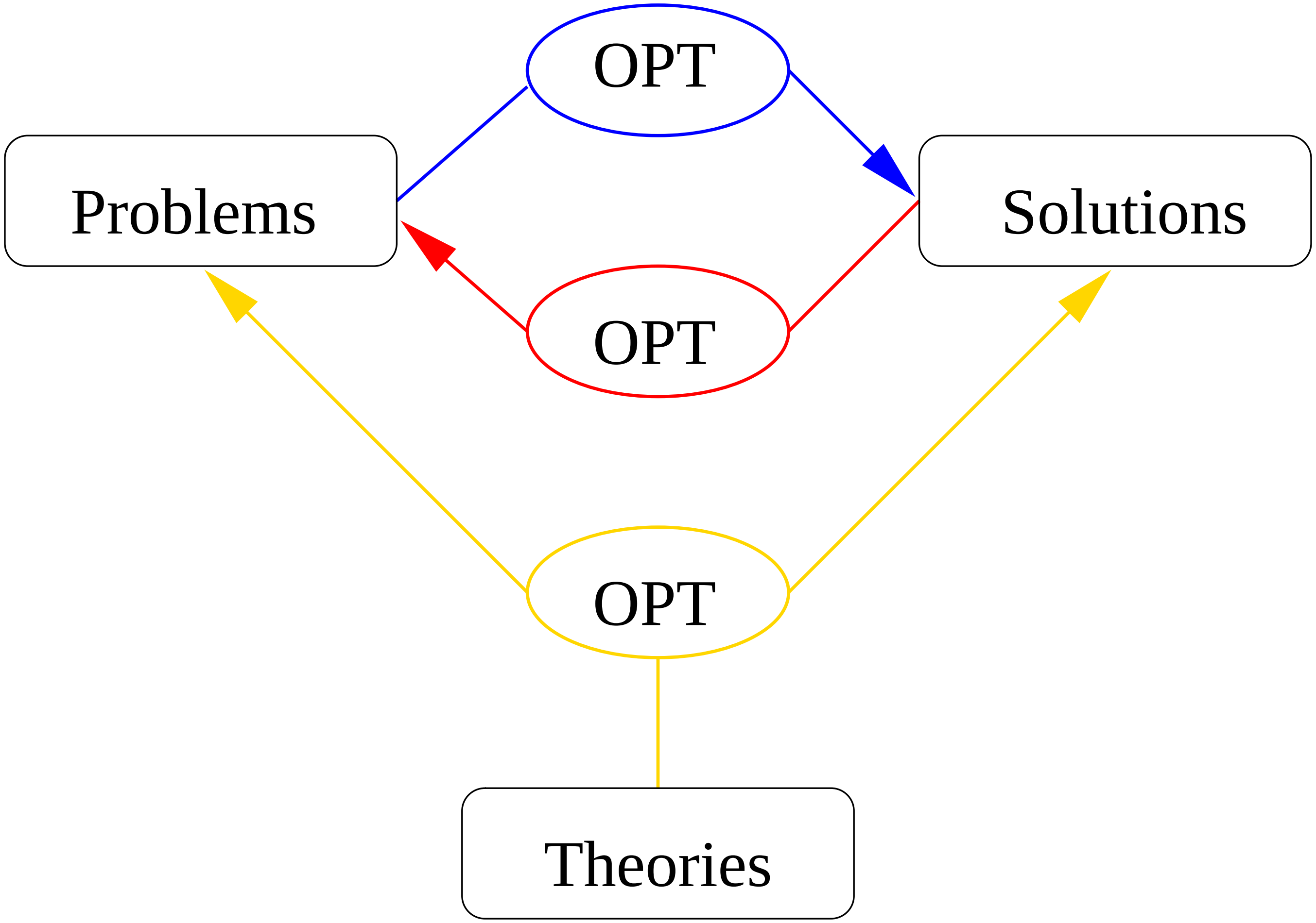
Convex Optimization and its Applications to Computer Science

Outline

- Optimization
- Examples
- Overview of Syllabus
- An Example: The Illumination Problem
- References

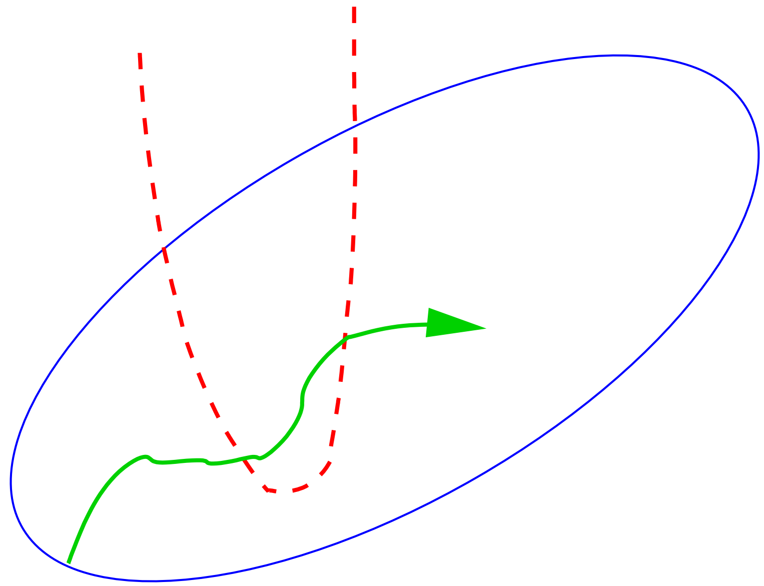
Acknowledgement: Grateful thanks to Mung Chiang (Princeton) and Stephen Boyd (Stanford) for course materials in this class.

Optimization Mentality



Optimization as a mathematical language, as a rule for algorithm and system design, as a unifying principle

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 formulate an optimization problem and describe its constraint set? What properties does it have?

- Is the problem feasible or solvable?

- Can this problem possibly relate to other optimization problems?

- Can we find an algorithm to solve the problem in an efficient and scalable way?

- If we can't solve the original problem, can we reformulate it, approximate it or relax it to an easier one? Can we bound it?

- Can we find an optimization problem for a given solution?

Solving Optimization Problems

- **general nonlinear optimization problem**
  - often difficult to solve, e.g., nonconvex problems
  - methods involve some compromise, e.g., very long computational time
- **local optimization**
  - find a suboptimal solution
  - computationally fast but dependent on initial point and parameter tuning
- **global optimization**
  - find a global optimal solution
  - computationally slow

Convex optimization is both art and technology. You need to understand the art of problem formulation and exploit problem structure to design algorithms to solve the problem. Over time, this process becomes a technology, e.g., least-squares.

Perspective

Widely known: linear programming is powerful and easy to solve

Modified view: "`... the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity.`"

– R. Rockafellar, SIAM Review 1993.

- 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

We'll start with introducing convex optimization theory and algorithms, and then specialize into various applications involving convex and even nonconvex problems

Examples

- Least-Squares:

$$\text{minimize}_x \;\; \|Ax - b\|^2$$

Analytical solution: $x^\star = (A^TA)^{-1}A^Tb$; reliable and efficient algorithms and software; computational time proportional to n^2k ($A \in \mathcal{R}^{k\times n}$); less with structure; a mature technology

- Linear Programming:

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

No analytical formula for solution; reliable and efficient algorithms and software; computational time proportional to n^2m if $m \geq n$; less with structure; a mature technology
Not as easy to recognize as least squares

- Convex Optimization:

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

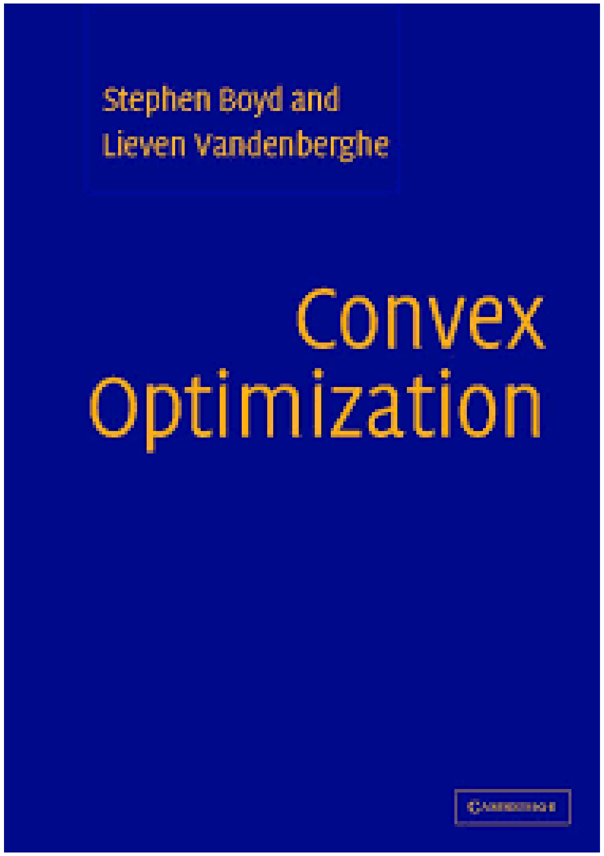
Typically, no analytical formula for solution; reliable and efficient algorithms and software; computational time (roughly) proportional to $\max\{n^3, n^2m, F\}$ where F is the cost of evaluating the objective and constraint functions f_i and their first and second order derivatives (if they exist); *almost a technology*

- Applications of convex optimization

Often difficult to recognize, many tricks for transforming problems, surprisingly many problems can be solved (or be tackled elegantly) via convex optimization

Once the formulation is done, solving the problem is, like least-squares or linear programming, (almost) technology. Learn tricks from examples and applications. Convex optimization software are important as they determine how soon convex optimization becomes a technology.

How to Learn Convex Optimization?

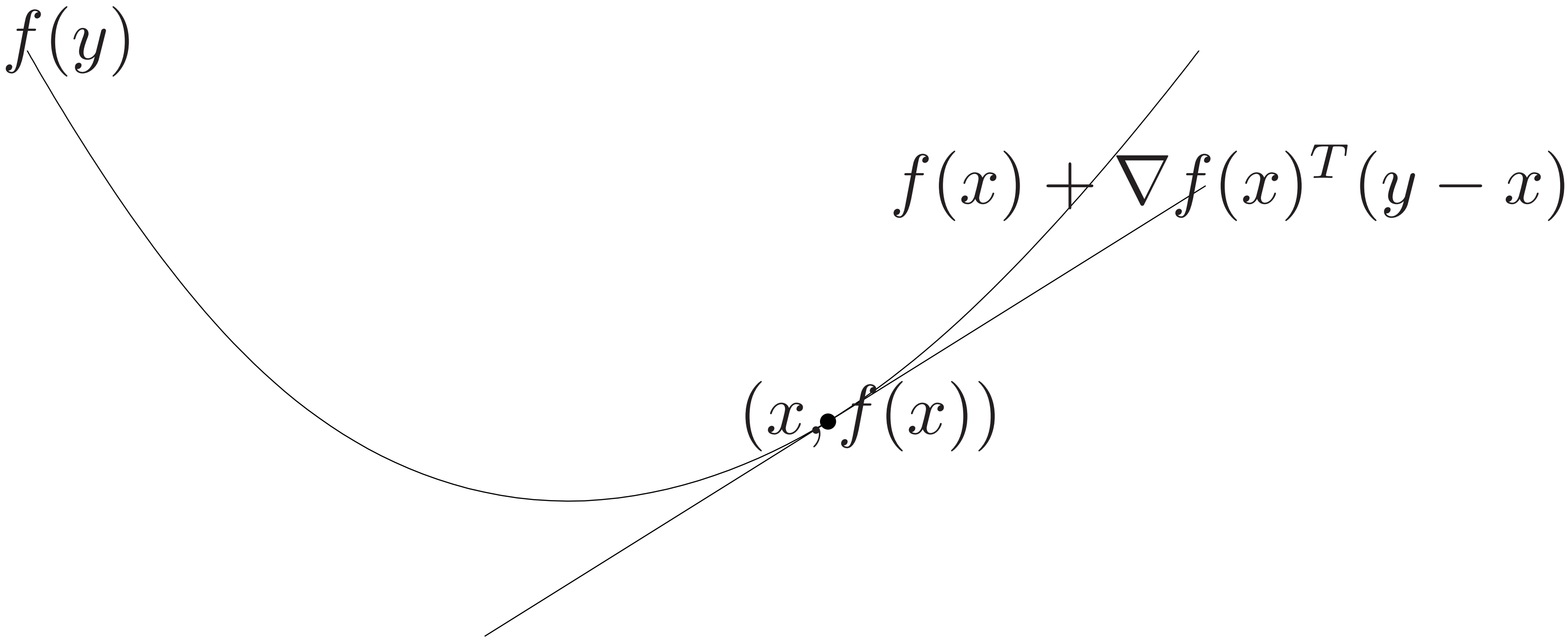


不憤不啓，
不悱不發。
舉一隅不以
三隅反，
則不復也。
《論語·述而
第七》

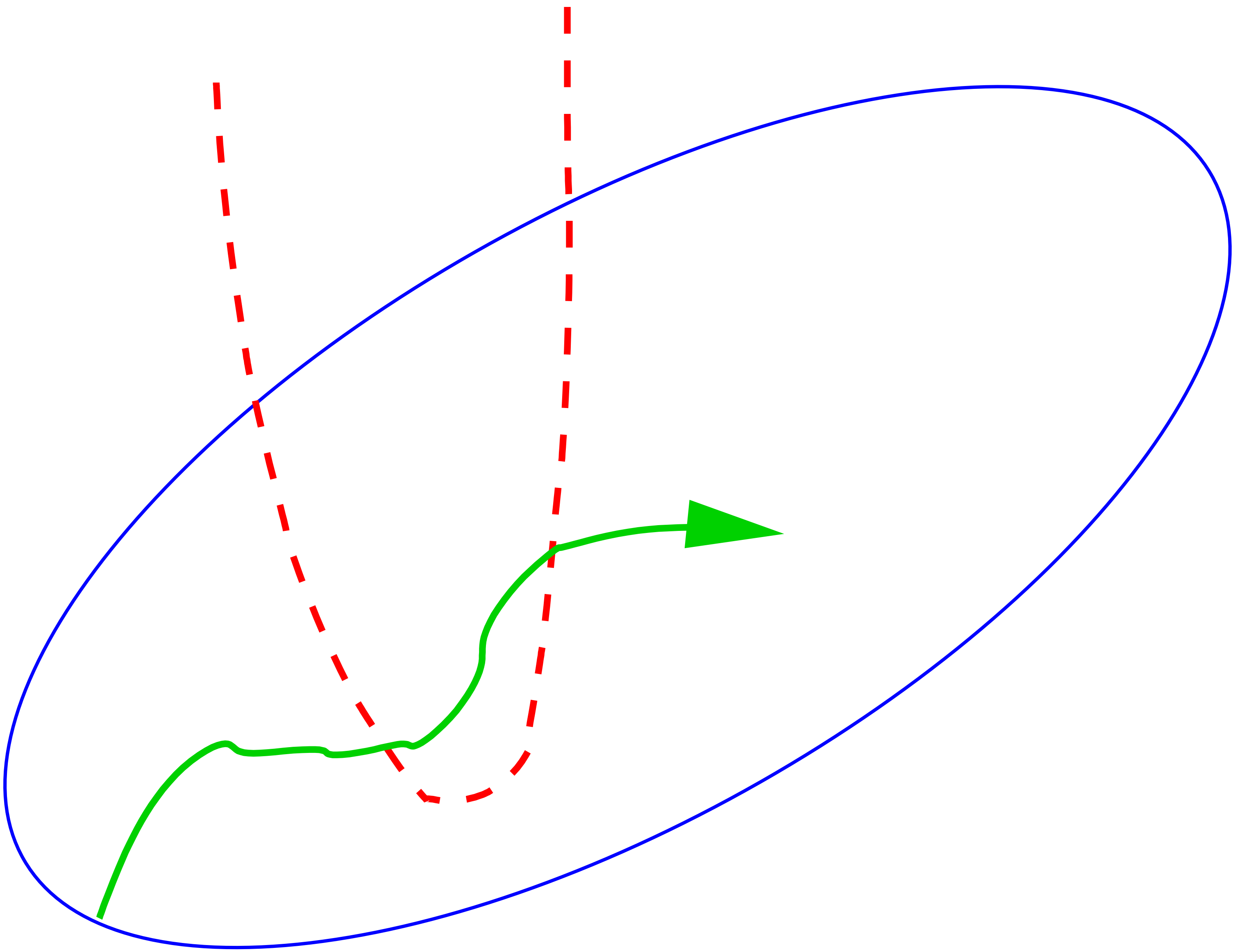
Find new ways to interpret a mathematical statement given one way. Draw inferences about other possible applications from one instance

Learn how to recognize and formulate convex optimization, and use optimization tools to creatively solve problems (an objective of this course)

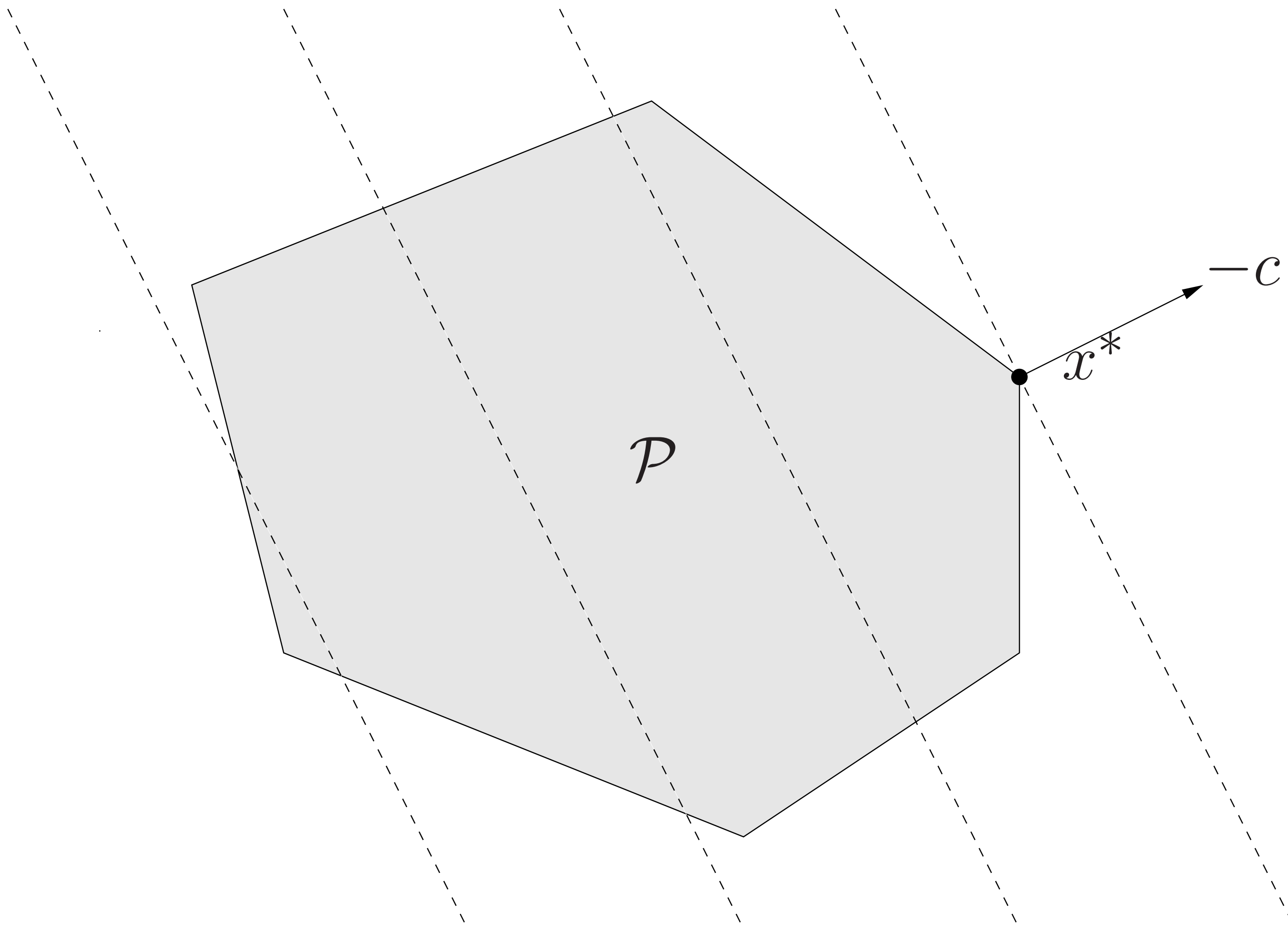
Convex Sets and Convex Functions



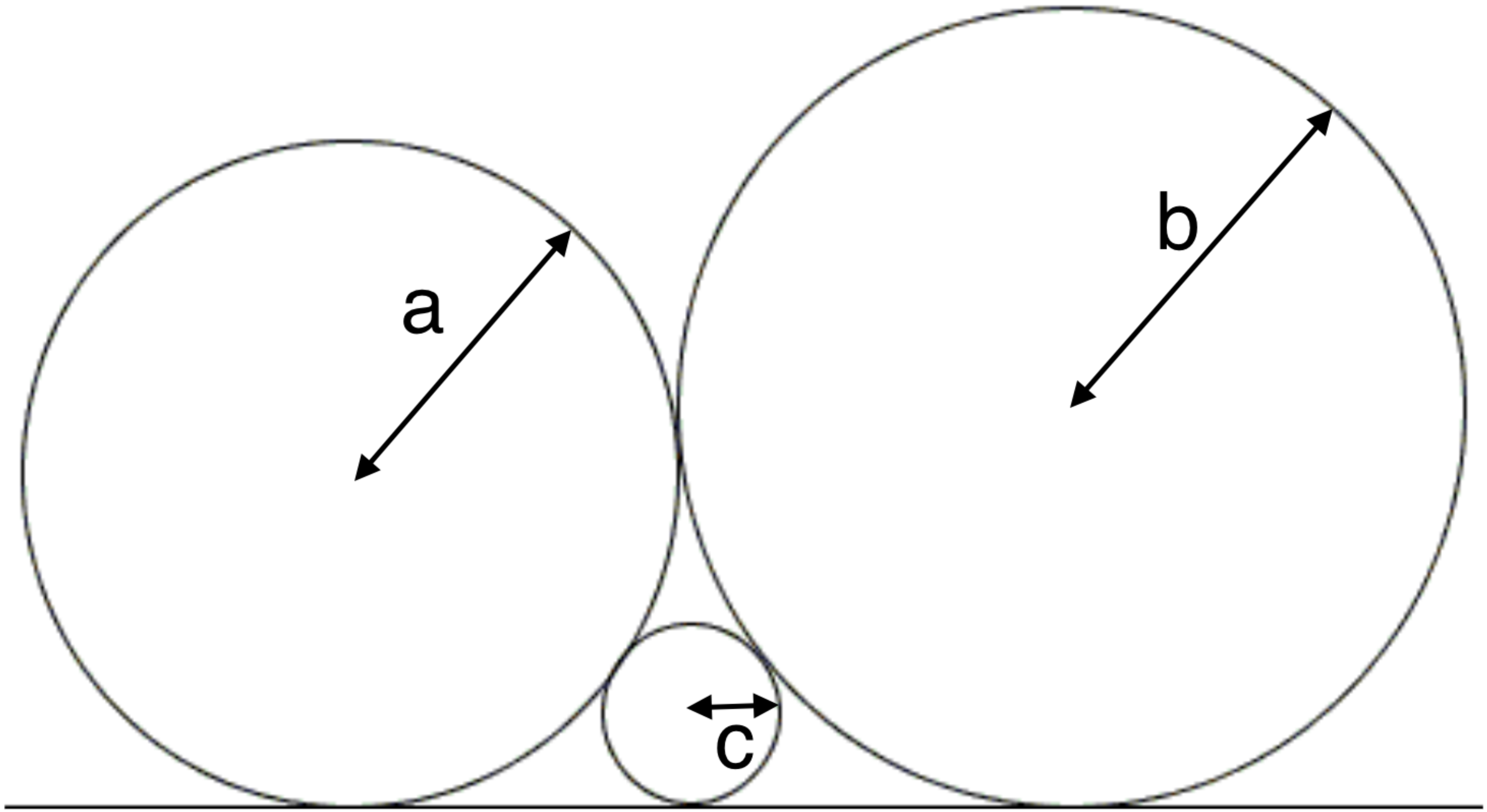
Convex Optimization and Lagrange Duality



Linear Programming and Quadratic Programming

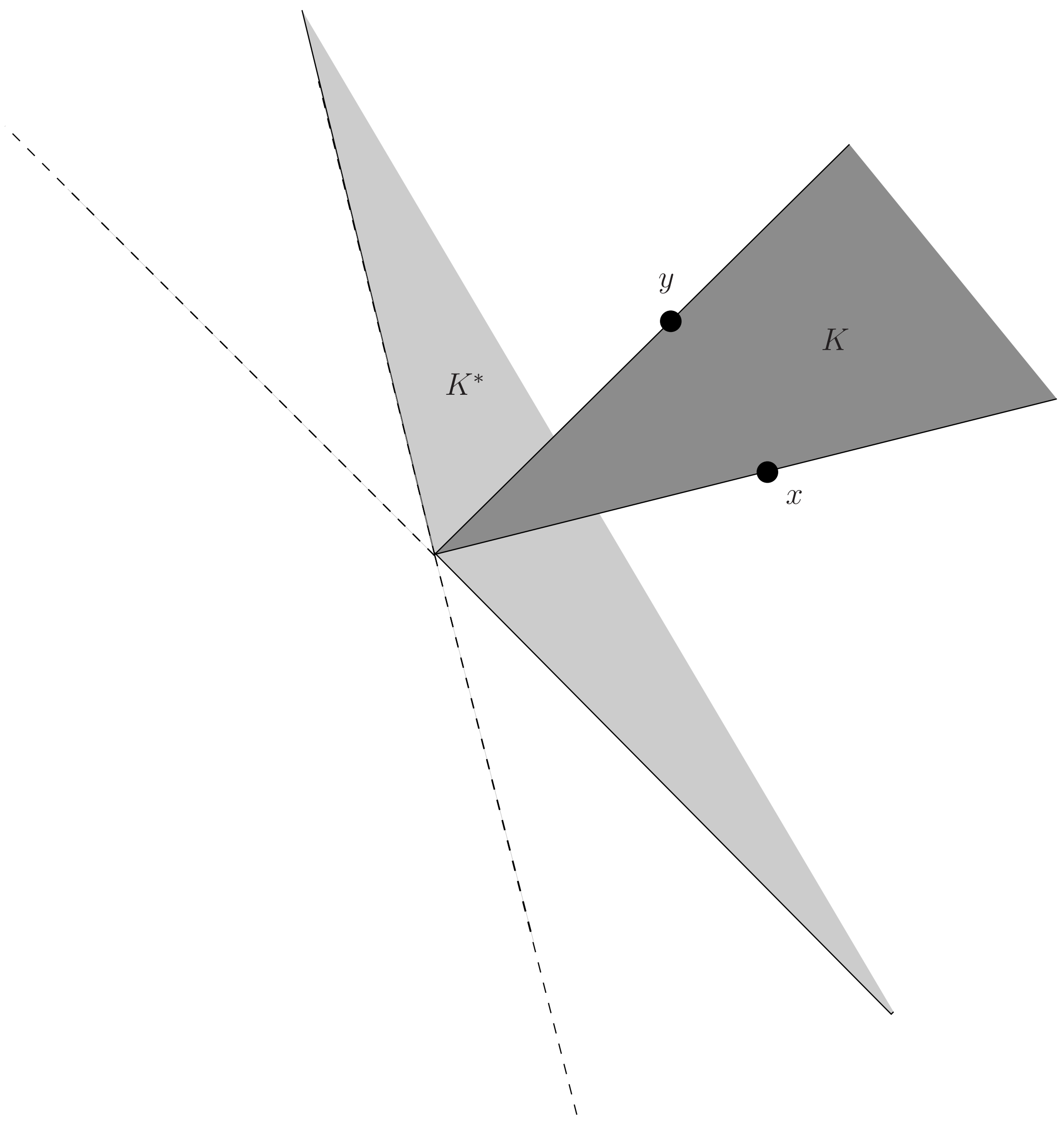


Geometric Programming

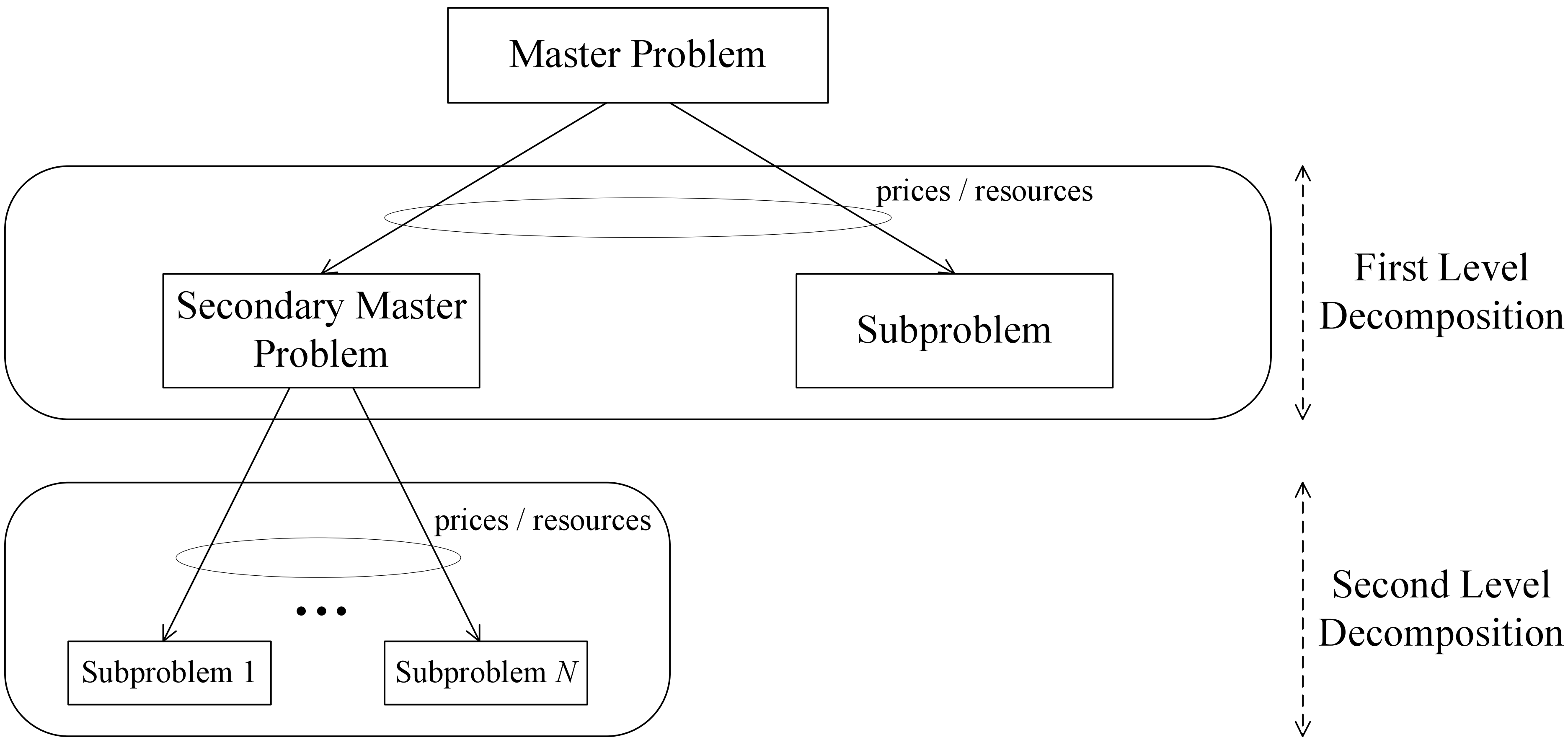


$$a + b \geq 2\sqrt{ab}$$

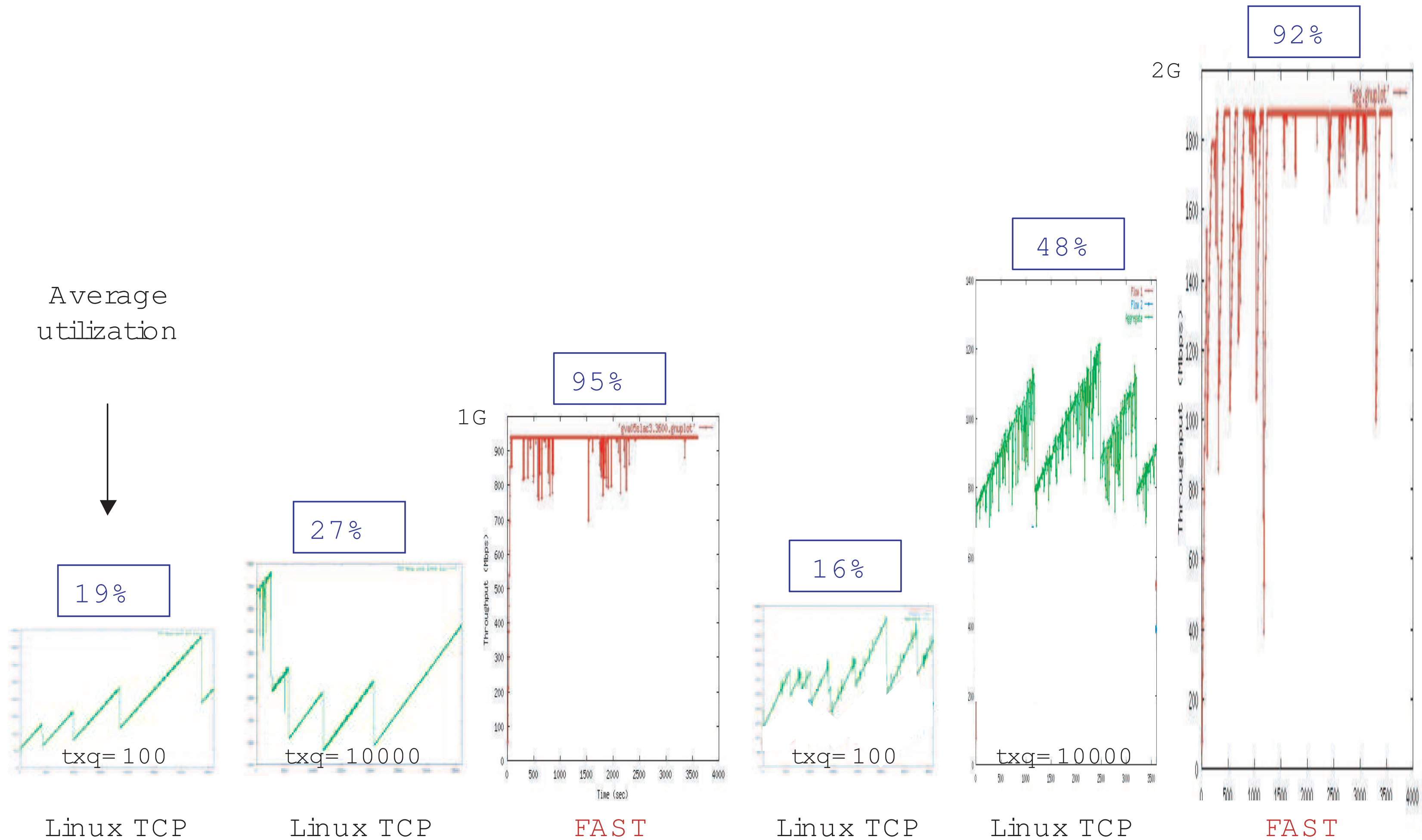
Semidefinite Programming



Decomposition

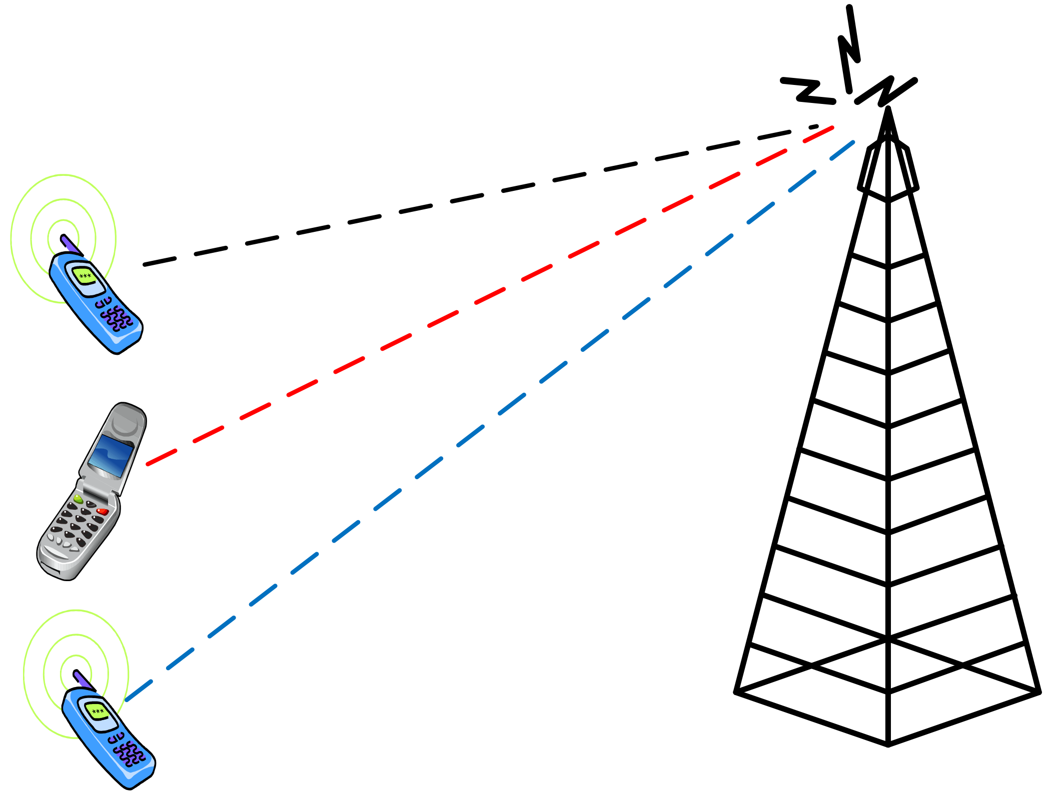


Network Utility Maximization in Internet



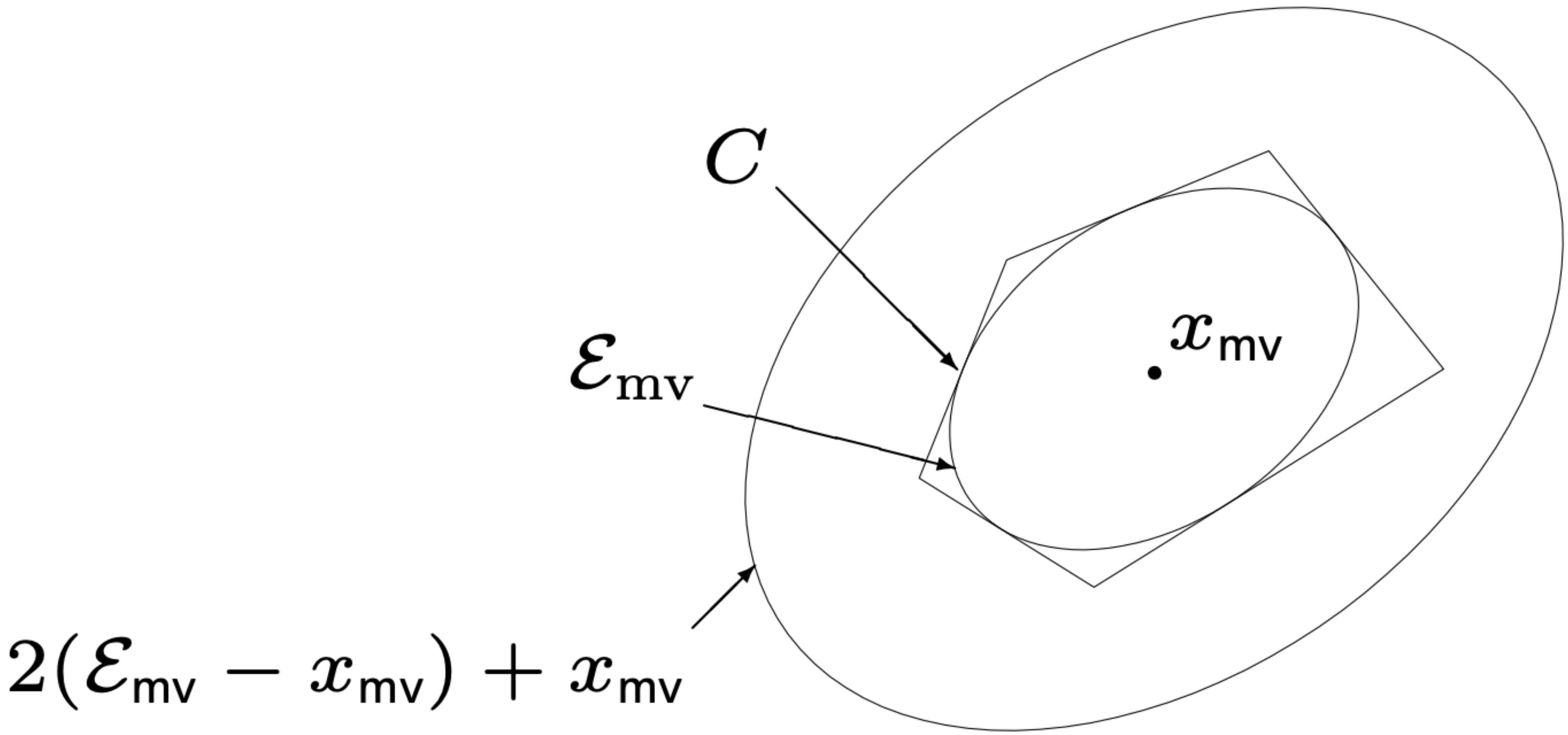
Web Intelligence: Ranking, Statistical Inference

Wireless Network Optimization



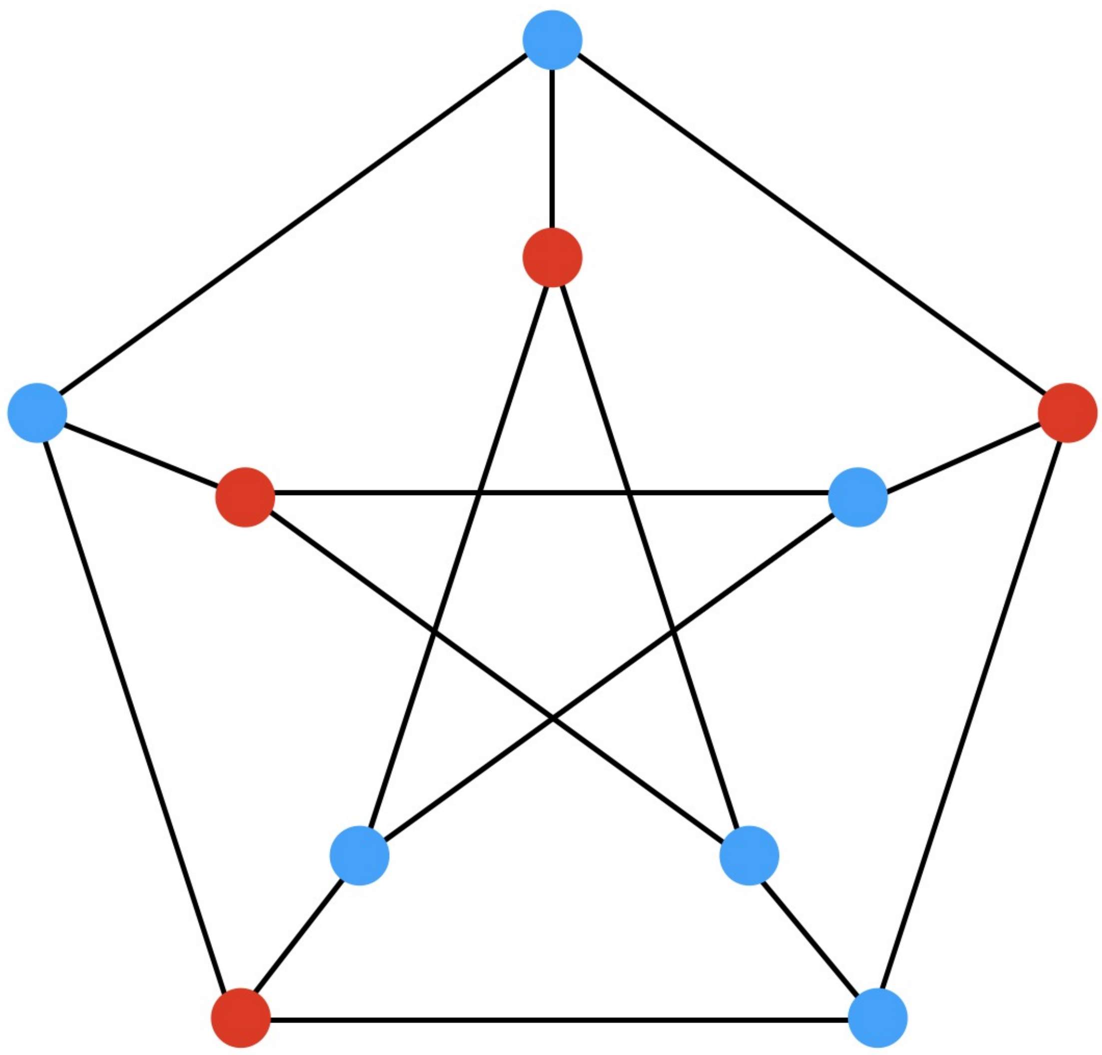
$$\max_{p_1+p_2\le 1} \min\left\{\frac{p_1}{p_2+1}, \frac{p_2}{p_1+1}\right\} = \rho\left(\left[\begin{array}{cc} 1 & 2 \\ 2 & 1 \end{array}\right]\right)^{-1}$$

Computational Geometry



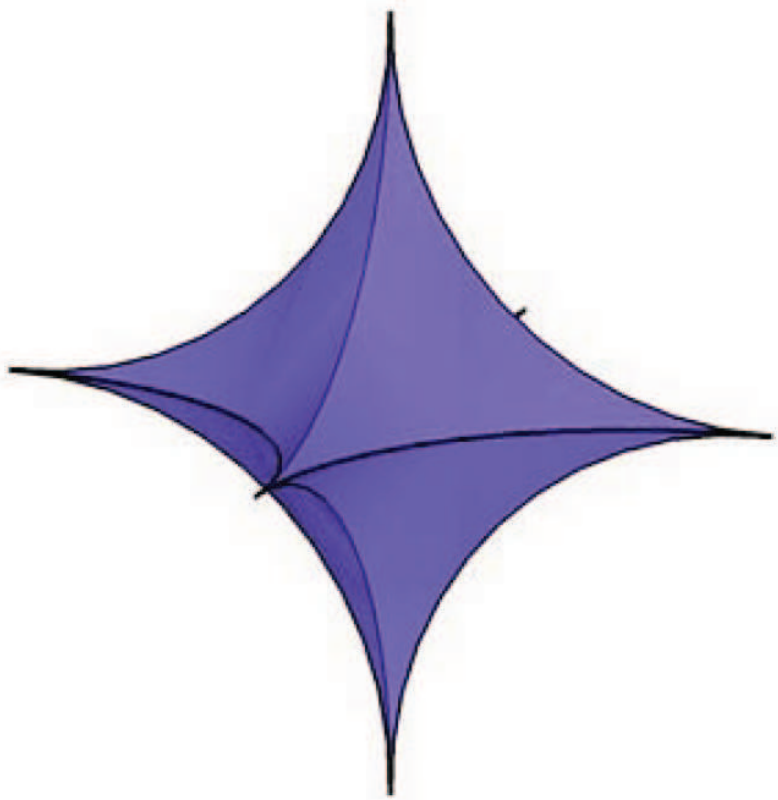
$\varepsilon_{\text{mv}} \subseteq C \subseteq 2(\varepsilon_{\text{mv}} - x_{\text{mv}}) + x_{\text{mv}}$ where $x_{\text{mv}} = \text{center}(\varepsilon_{\text{mv}})$

Combinatorial Graph Problems



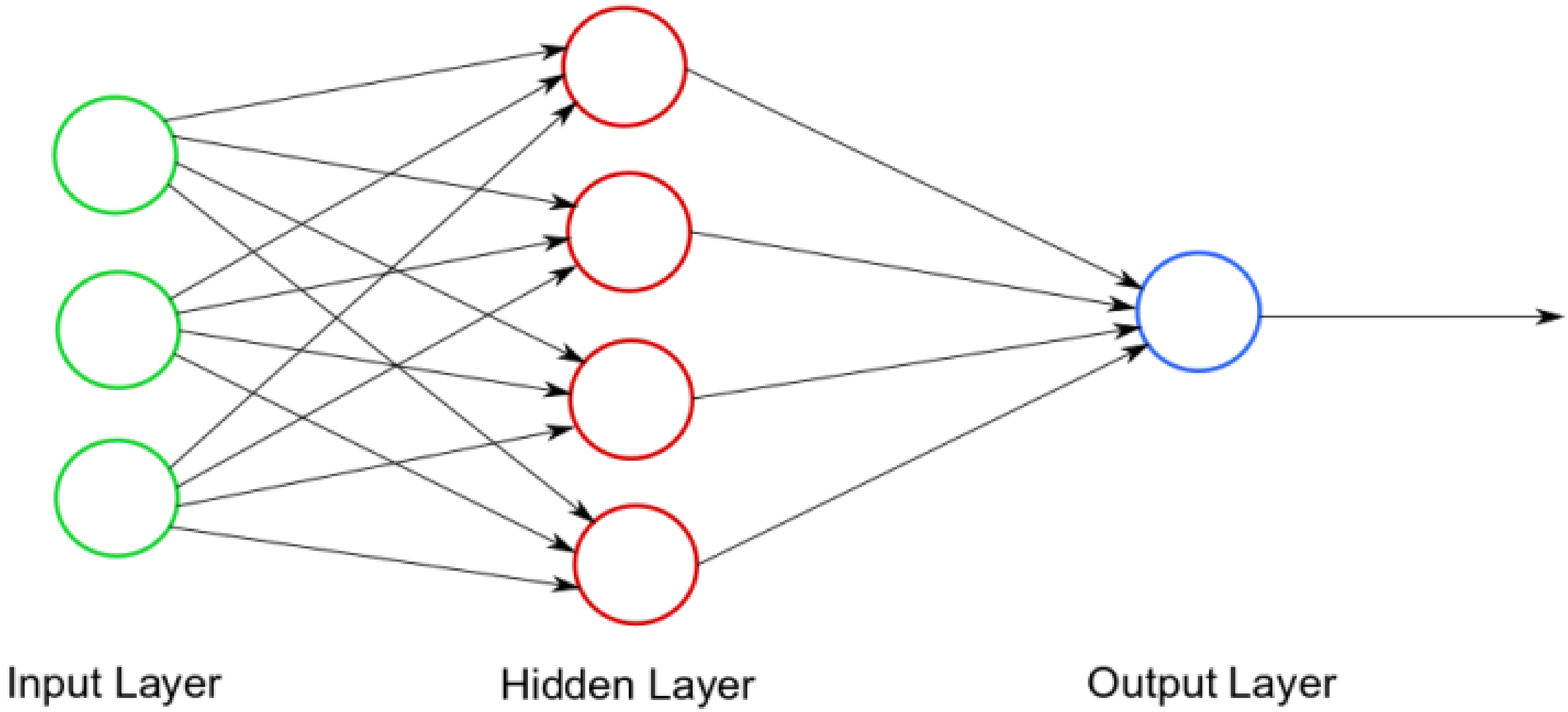
Find a bipartite subgraph with the maximum number of edges

Non-Convex Problems: Convex-Cardinality, SOS Proofs, Reasoning in AI



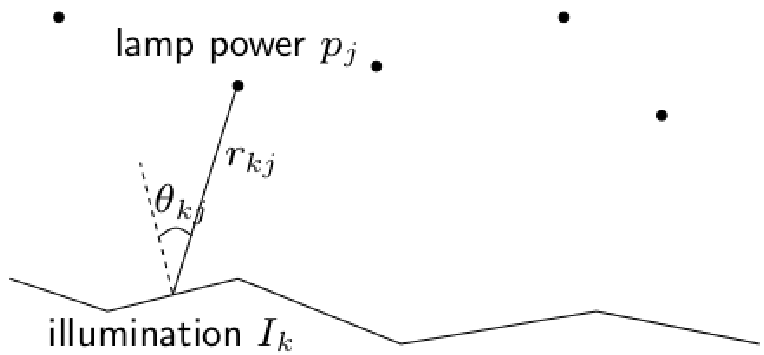
Sparsity, Compressed Sensing, Shortest-Path Routing, Shortest Proof in AI, Dynamic Programming

Optimization Methods in Machine Learning



Illumination Problem

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

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

How to Solve Illumination Problem

1. use uniform power: $p_j = p$, vary p

2. use least-squares:

$$\text{minimize} \quad \sum_{k=1}^{n} (I_k - I_{\mathsf{des}})^2$$

   round p_j if $p_j > p_{\max}$ or $p_j < 0$

3. use weighted least-squares:

$$\text{minimize} \quad \sum_{k=1}^{n} (I_k - I_{\mathsf{des}})^2 + \sum_{j=1}^{m} w_j (p_j - p_{\max})^2$$

   iteratively adjust weights w_j until $0 \le p_j \le p_{\max}$

4. use linear programming:

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

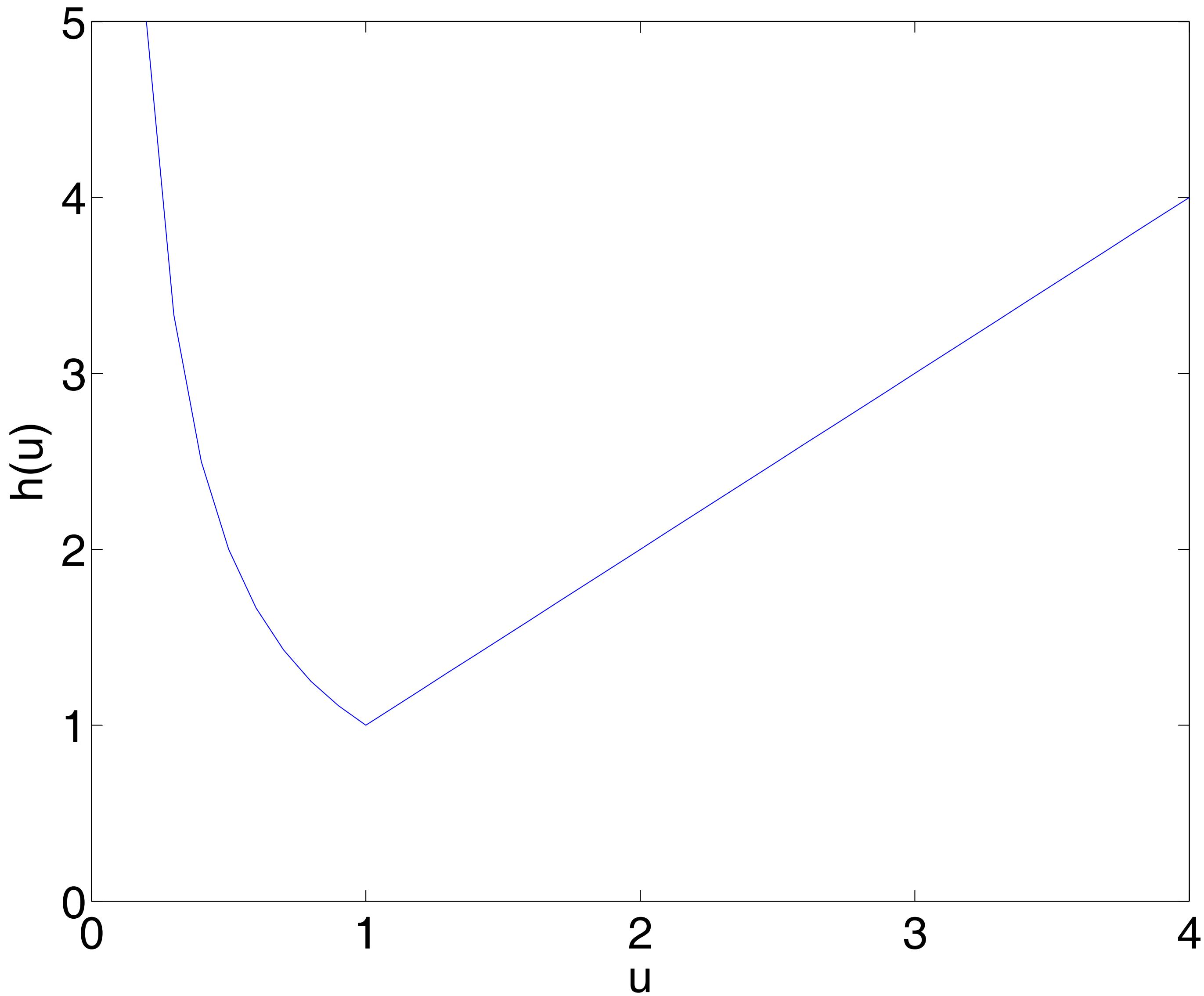
which can be solved via linear programming

of course these are suboptimal 'solutions' (to the original problem)

5. use convex optimization: problem is equivalent to

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

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


5
4
3
2
1
0
h(u)
0
1
2
3
4
u

$f_0(p)$ is convex because maximum of convex functions is convex

exact solution obtained with effort $\approx$ modest factor $\times$ least-squares effort

additional constraints: does adding (a) or (b) below complicate the problem?

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

- answer: with (a), still easy to solve; with (b), extremely difficult
- moral: (untrained) intuition doesn't always work; without the proper background very easy problems can appear quite similar to very difficult problems

6. Can you think of yet another way to solve the Illumination Problem exactly?

References

- Primary Textbook: Convex Optimization by Stephen Boyd and Lieven Vandenberghe, Cambridge University Press, 2004
- Introduction to Matrix Analysis, R. Bellman, SIAM, 1997
- Nonlinear Programming, D. Bertsekas, Athena Scientific, 1999
- Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering, by A. Ben-Tal and A. Nemirovski, SIAM, 2001
- Parallel and Distributed Computation: Numerical Methods by D. Bertsekas and J. N. Tsitsiklis, 1st Edition, Prentice Hall, 1989