

Lecture 2c: Introduction to Convex Programming

.pull-left[

@luk036

2023-09-27

] .pull-right[

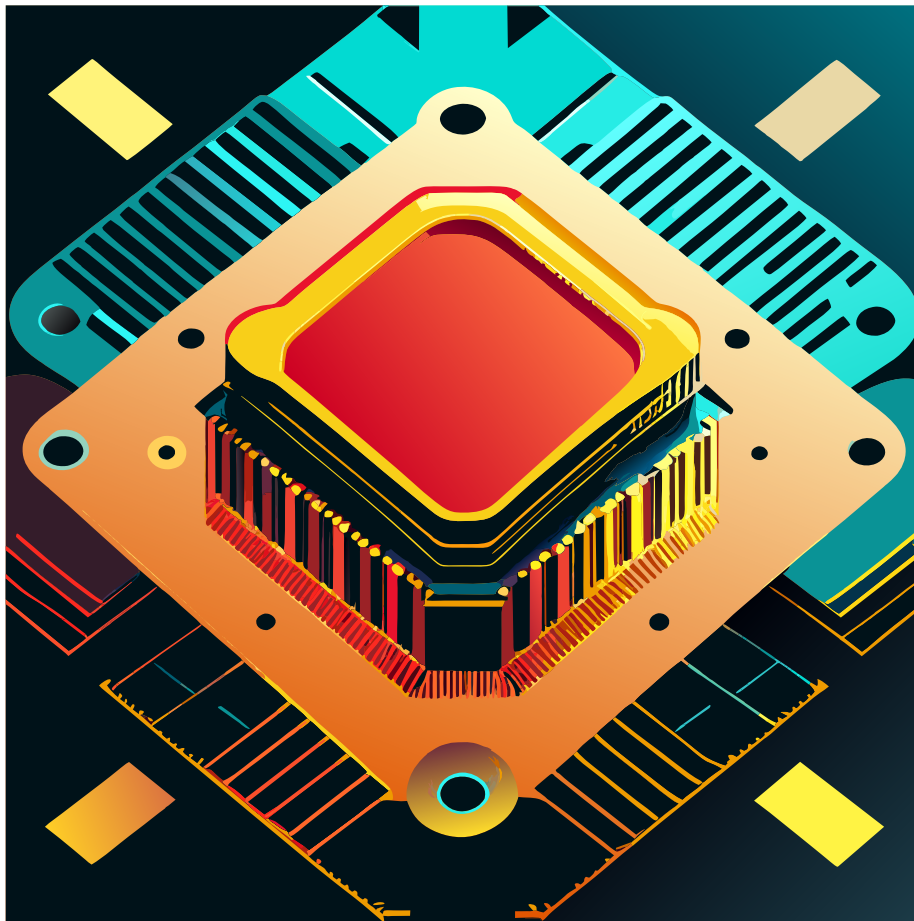


Figure 1: image

]

Abstract

This lecture provides an introduction to the convex programming and covers various aspects of optimization. The lecture begins with an overview of optimization, including linear and nonlinear programming, duality and convexity, and approximation techniques. It then delves into more specific topics within continuous optimization, such as linear programming problems and their standard form, transformations to standard form, and the duality of linear programming problems. The lecture also touches on nonlinear programming, discussing the standard form of an NLPP (nonlinear programming problem) and the necessary conditions of optimality known as the Karush-Kuhn-Tucker (KKT) conditions. Convexity is another important concept explored in the document, with explanations on the definition of convex functions and their properties. The lecture also discusses the duality of convex optimization problems and their usefulness in computation. Finally, the document briefly mentions various unconstrained optimization techniques, descent methods, and approximation methods under constraints.

Overview

- Introduction
 - Linear programming
 - Nonlinear programming
 - Duality and Convexity
 - Approximation techniques
 - Convex Optimization
 - Books and online resources.
-

Classification of Optimizations

- Continuous
 - Linear vs Non-linear
 - Convex vs Non-convex
 - Discrete
 - Polynomial time Solvable
 - NP-hard
 - * Approximatable
 - * Non-approximatable
 - Mixed
-

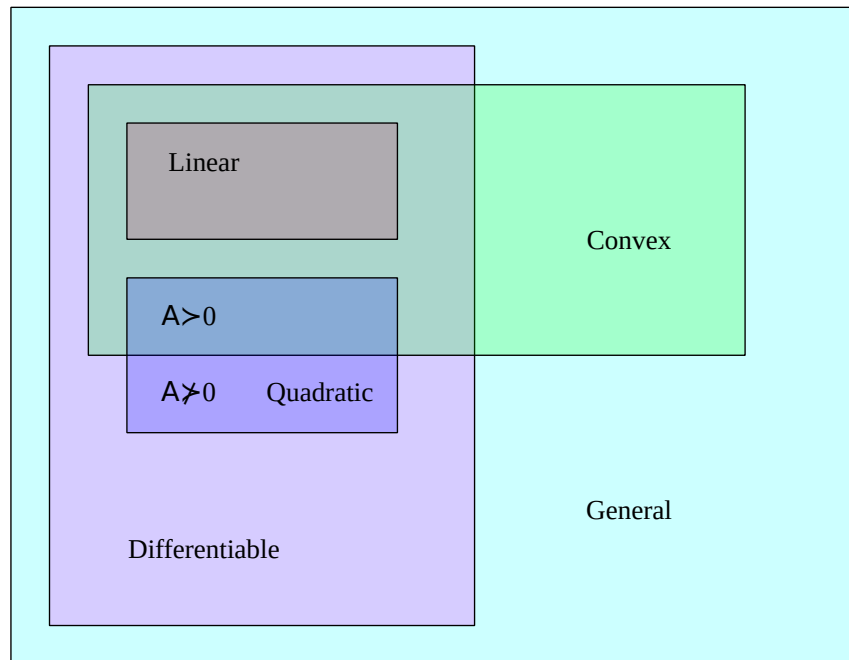


Figure 2: classification

Continuous Optimization

Linear Programming Problem

- An LPP in standard form is:

$$\min\{c^T x \mid Ax = b, x \geq 0\}.$$

- The ingredients of LPP are:
 - An $m \times n$ matrix A , with $n > m$
 - A vector $b \in \mathbb{R}^m$
 - A vector $c \in \mathbb{R}^n$
-

Example

$$\begin{aligned} \text{minimize} \quad & 0.4x_1 + 3.4x_2 - 3.4x_3 \\ \text{subject to} \quad & 0.5x_1 + 0.5x_2 = 3.5 \\ & 0.3x_1 - 0.8x_2 + 8.4x_3 = 4.5 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Transformations to Standard Form

- Theorem: Any LPP can be transformed into the standard form.
 - Variables not restricted in sign:
 - Decompose x to two new variables $x = x_1 - x_2, x_1, x_2 \geq 0$
 - Transforming inequalities into equalities:
 - By putting slack variable $y = b - Ax \geq 0$
 - Set $x' = (x, y), A' = (A, 1)$
 - Transforming a max into a min
 - $\max(\text{expression}) = \min(-\text{expression})$;
-

Duality of LPP

- If the primal problem of the LPP: $\min\{c^T x \mid Ax \geq b, x \geq 0\}$.
 - Its dual is: $\max\{y^T b \mid A^T y \leq c, y \geq 0\}$.
 - If the primal problem is: $\min\{c^T x \mid Ax = b, x \geq 0\}$.
 - Its dual is: $\max\{y^T b \mid A^T y \leq c\}$.
-

Nonlinear Programming

- The standard form of an NLPP is

$$\min\{f(x) \mid g(x) \leq 0, h(x) = 0\}.$$

- Necessary conditions of optimality, Karush- Kuhn-Tucker (KKT) conditions:
 - $\nabla f(x) + \mu \nabla g(x) + \lambda \nabla h(x) = 0$,
 - $\mu g(x) = 0$,
 - $\mu \geq 0, g(x) \leq 0, h(x) = 0$
-

Convexity

- A function $f: K \subseteq \mathbb{R}^n \mapsto \mathbb{R}$ is convex if K is a convex set and $f(y) \geq f(x) + \nabla f(x)(y - x), y, x \in K$.
- **Theorem:** Assume that f and g are convex differentiable functions. If the pair (x, m) satisfies the KKT conditions above, x is an optimal solution of the problem. If in addition, f is strictly convex, x is the only solution of the problem.

(Local minimum = global minimum)

Duality and Convexity

- Dual is the NLPP:

$$\max\{\theta(\mu, \lambda) \mid \mu \geq 0\},$$

where $\theta(\mu, \lambda) = \inf_x [f(x) + \mu g(x) + \lambda h(x)]$

- Dual problem is always convex.
 - Useful for computing the lower/upper bound.
-

Applications

- Statistics
 - Filter design
 - Power control
 - Machine learning
 - SVM classifier
 - logistic regression
-

class: nord-light, middle, center

Convexify the non-convex's

Change of curvature: square

Transform:

$$0.3 \leq \sqrt{x} \leq 0.4$$

into:

$$0.09 \leq x \leq 0.16.$$

Note that $\sqrt{\cdot}$ are **monotonic concave** functions in $(0, +\infty)$.

Generalization: - Consider $|H(\omega)|^2$ (power) instead of $|H(\omega)|$ (magnitude). - square root -> Spectral factorization

Change of curvature: square

Transform:

$$x^2 + y^2 \geq 0.16, \quad (\text{non-convex})$$

into:

$$x' + y' \geq 0.16, \quad x', y' \geq 0$$

Then:

$$x_{\text{opt}} = \pm \sqrt{x'_{\text{opt}}}, \quad y_{\text{opt}} = \pm \sqrt{y'_{\text{opt}}}.$$

Change of curvature: sine

Transform:

$$\sin^2 x \leq 0.4, \quad 0 \leq x \leq \pi/2$$

into:

$$y \leq 0.4, \quad 0 \leq y \leq 1$$

Then:

$$x_{\text{opt}} = \sin^{-1}(\sqrt{y_{\text{opt}}}).$$

Note that $\sin(\cdot)$ are monotonic concave functions in $(0, \pi/2)$.

Change of curvature: log

Transform:

$$\pi \leq x/y \leq \phi$$

into:

$$\pi' \leq x' - y' \leq \phi'$$

where $z' = \log(z)$.

Then:

$$z_{\text{opt}} = \exp(z'_{\text{opt}}).$$

Generalization: - Geometric programming

Change of curvature: inverse

Transform:

$$\log(x) + \frac{c}{x} \leq 0.3, \quad x > 0$$

into:

$$-\log(y) + c \cdot y \leq 0.3, \quad y > 0.$$

Then:

$$x_{\text{opt}} = y_{\text{opt}}^{-1}.$$

Note that $\sqrt{\cdot}$, $\log(\cdot)$, and $(\cdot)^{-1}$ are monotonic functions.

Generalize to matrix inequalities

Transform:

$$\log(\det X) + \text{Tr}(X^{-1}C) \leq 0.3, X \succ 0$$

into:

$$-\log(\det Y) + \text{Tr}(Y \cdot C) \leq 0.3, Y \succ 0$$

Then:

$$X_{\text{opt}} = Y_{\text{opt}}^{-1}.$$

Change of variables

Transform:

$$(a + b \cdot y)x \leq 0, x > 0$$

into:

$$a \cdot x + b \cdot z \leq 0, x > 0$$

where $z = yx$.

Then:

$$y_{\text{opt}} = z_{\text{opt}}x_{\text{opt}}^{-1}$$

Generalize to matrix inequalities

Transform:

$$(A + B\mathbf{Y})X + X(A + B\mathbf{Y})^T \prec 0, X \succ 0$$

into:

$$AX + XA^T + BZ + Z^TB^T \prec 0, X \succ 0$$

where $Z = YX$.

Then:

$$Y_{\text{opt}} = Z_{\text{opt}}X_{\text{opt}}^{-1}$$

Other thoughts

- Minimizing any quasi-convex function subject to convex constraints can easily be transformed into a convex programming.
 - Replace a non-convex constraint with a sufficient condition (such as its lower bound). Less optimal.
 - Relaxation + heuristic
 - Decomposition
-

Unconstraint Techniques

- Line search methods
 - Fixed or variable step size
 - Interpolation
 - Golden section method
 - Fibonacci's method
 - Gradient methods
 - Steepest descent
 - Quasi-Newton methods
 - Conjugate Gradient methods
-

General Descent Method

1. **Input:** a starting point $x \in \text{dom } f$
 2. **Output:** x^*
 3. **repeat**
 1. Determine a descent direction p .
 2. Line search. Choose a step size $\alpha > 0$.
 3. Update. $x := x + \alpha p$
 4. **until** stopping criterion satisfied.
-

Some Common Descent Directions

- Gradient descent: $p = -\nabla f(x)^\top$
 - Steepest descent:
 - $\Delta x_{nsd} = \text{argmin}\{\nabla f(x)^\top v \mid \|v\| = 1\}$
 - $\Delta x = \|\nabla f(x)\| \Delta x_{nsd}$ (un-normalized)
 - Newton's method:
 - $p = -\nabla^2 f(x)^{-1} \nabla f(x)$
 - Conjugate gradient method:
 - p is “orthogonal” to all previous p 's
 - Stochastic subgradient method:
 - p is calculated from a set of sample data (instead of using all data)
 - Network flow problems:
 - p is given by a “negative cycle” (or “negative cut”).
-

Approximation Under Constraints

- Penalization and barriers
- Dual method
- Interior Point method

- Augmented Lagrangian method

Books and Online Resources

- Pablo Pedregal. Introduction to Optimization, Springer. 2003 (O224 P371)
- Stephen Boyd and Lieven Vandenberghe, Convex Optimization, Dec. 2002
- Mittlemann, H. D. and Spellucci, P. Decision Tree for Optimization Software, World Wide Web, <http://plato.la.asu.edu/guide.html>, 2003

count: false class: nord-dark, middle, center

.pull-left[

Q & A

] .pull-right[

]



Figure 3: image