Convex Optimization for Signal Processing and Communications: From Fundamentals to Applications

Chong-Yung Chi

Institute of Communications Engineering &
Department of Electrical Engineering
National Tsing Hua University, Taiwan 30013
E-mail: cychi@ee.nthu.edu.tw
Web: http://www.ee.nthu.edu.tw/cychi/

Invited tutorial talk at Aalborg University, Denmark, 2018/9/20

Acknowledgment: My post doctor, **Dr. Chia-Hsiang Lin** for preparing slides of Part II.
My Ph.D. student, **Yao-Rong Syu** for some slides preparation.

Outline

1. **Part I: Fundamentals of Convex Optimization**
2. **Part II: Application in Hyperspectral Image Analysis: (Big Data Analysis and Machine Learning)**
3. **Part III: Application in Wireless Communications (5G Systems)**
   - **Subsection I:** Outage Constrained Robust Transmit Optimization for Multiuser MISO Downlinks
   - **Subsection II:** Outage Constrained Robust Hybrid Coordinated Beamforming for Massive MIMO Enabled Heterogeneous Cellular Networks

Optimization problem

- **Optimization problem:**

$$\begin{array}{rl} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in \mathcal{C} \end{array} \quad (1)$$

where $f(\mathbf{x})$ is the objective function to be minimized and $\mathcal{C}$ is the feasible set from which we try to find an optimal solution. Let

$$\mathbf{x}^\star = \arg\min_{\mathbf{x}\in\mathcal{C}} f(\mathbf{x}) \quad \text{(optimal solution or global minimizer)} \quad (2)$$

- **Challenges in applications:**

• Local optima; large problem size; decision variable $\mathbf{x}$ involving real and/or complex vectors, matrices; feasible set $\mathcal{C}$ involving generalized inequalities, etc.

• Computational complexity: NP-hard; polynomial-time solvable.

• Performance analysis: Performance insights, properties, perspectives, proofs (e.g., identifiability and convergence), limits and bounds.

Convex sets and convex functions-1

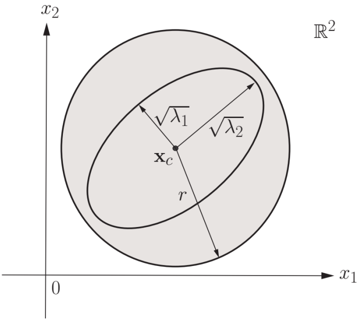
- **Affine (convex) combination:** Provided that C is a nonempty set,

$$\mathbf{x} = \theta_1 \mathbf{x}_1 + \cdots + \theta_K \mathbf{x}_K, \ \mathbf{x}_i \in C \ \forall i \tag{3}$$

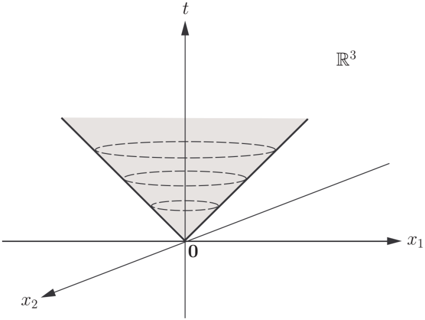
is called an *affine (a convex) combination* of $\mathbf{x}_1, \ldots, \mathbf{x}_K$ (K vectors or points of a set) if $\sum_{i=1}^{K} \theta_i = 1$, $\theta_i \in \mathbb{R}$ ($\theta_i \in \mathbb{R}_+$), $K \in \mathbb{Z}_{++}$.

- **Affine (convex) set:**
  - C is an *affine (a convex) set* if C is closed under the operation of *affine (convex) combination*;
  - an affine set is constructed by *lines*;
  - a convex set is constructed by *line segments*.

- **Conic set:**
  - If $\theta \mathbf{x} \in C$ for any $\theta \in \mathbb{R}_+$ and any $\mathbf{x} \in C$, then the set C is a *cone* and it is constructed by *rays starting from the origin*;
  - the linear combination (3) is called a *conic combination* if $\theta_i \geq 0 \ \forall i$;



Ellipsoid and Euclidean ball centered at $\mathbf{x}_c$



Second-order cone
$C = \left\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 \leq t\right\}$

- Left plot: An ellipsoid with semiaxes $\sqrt{\lambda_1}, \sqrt{\lambda_2}$, and an Euclidean ball with radius $r > \max\{\sqrt{\lambda_1}, \sqrt{\lambda_2}\}$ in $\mathbb{R}^2$; right plot: Second-order cone in $\mathbb{R}^3$.

Convex sets and convex functions-2

- Let $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$. Affine hull of $\mathcal{A}$ *(the smallest affine set containing $\mathcal{A}$)* is defined as

$$\textbf{aff } \mathcal{A} \triangleq \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \mid \sum_{i=1}^{N} \theta_i = 1,\ \theta_i \in \mathbb{R}\ \forall i \right\}$$
$$= \left\{ \mathbf{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \mid \boldsymbol{\alpha} \in \mathbb{R}^p \right\} \quad \text{(affine set representation)}$$

where $\mathbf{C} \in \mathbb{R}^{M \times p}$ is of full column rank, $\mathbf{d} \in \operatorname{aff} \mathcal{A}$, and

$$\operatorname{affdim} \mathcal{A} = p \leq \min\{N-1, M\}.$$

- $\mathcal{A}$ is *affinely independent* ($\mathcal{A}.\mathcal{I}.$) with $\operatorname{affdim} \mathcal{A} = N-1$ if the set $\{\mathbf{a} - \mathbf{a}_i \mid \mathbf{a} \in \mathcal{A}, \mathbf{a} \neq \mathbf{a}_i\}$ is linearly independent for any i; moreover,

$$\textbf{aff } \mathcal{A} = \{\mathbf{x} \mid \mathbf{b}^T\mathbf{x} = h\} \triangleq \mathcal{H}(\mathbf{b}, h) \quad \text{(when } M = N\text{)}$$

is a hyperplane, where $(\mathbf{b}, h)$ can be determined from $\mathcal{A}$ (closed-form expressions available).

Convex sets and convex functions-3

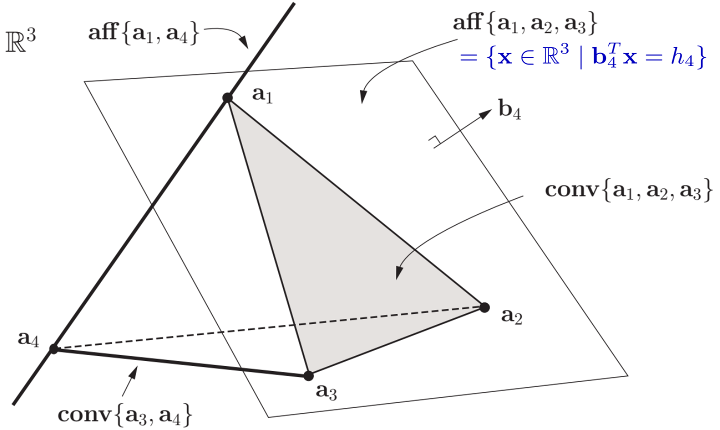


Figure 1: An illustration in $\mathbb{R}^3$, where $\mathbf{conv}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ is a simplex defined by the shaded triangle, and $\mathbf{conv}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$ is a simplex (and also a simplest simplex) defined by the tetrahedron with the four extreme points $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4\}$.

Convex sets and convex functions-4

- Let $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$. Convex hull of $\mathcal{A}$ *(the smallest convex set containing $\mathcal{A}$)* is defined as

$$\textbf{conv } \mathcal{A} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \mid \sum_{i=1}^{N} \theta_i = 1,\ \theta_i \in \mathbb{R}_+ \ \forall i \right\} \subset \text{aff } \mathcal{A}$$

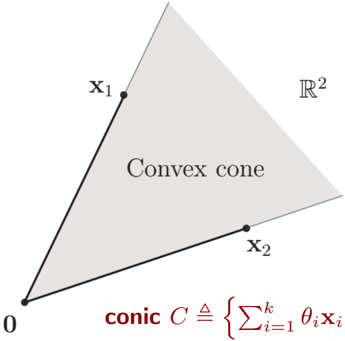
• $\text{conv}\, \mathcal{A}$ is called a *simplex* (a polytope with N vertices) if $\mathcal{A}$ is $\mathcal{A}.\mathcal{I}.$

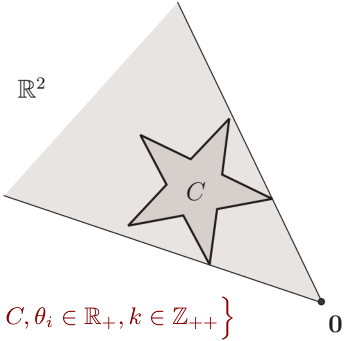
• When $\mathcal{A}$ is $\mathcal{A}.\mathcal{I}.$ and $M = N - 1$, $\text{conv}\, \mathcal{A}$ is called a *simplest simplex of N vertices* (i.e., $\mathbf{a}_1, \ldots, \mathbf{a}_N$);

$$\textbf{aff } (\mathcal{A} \setminus \{\mathbf{a}_i\}) = \mathcal{H}(\mathbf{b}_i, h_i),\ i \in \mathcal{I}_N = \{1, \ldots, N\}$$

is a *hyperplane*, where the N boundary hyperplane parameters $\{(\mathbf{b}_i, h_i), i = 1, \ldots, N\}$ can be uniquely determined from $\mathcal{A}$ (*closed-form expressions available*) and vice versa.

Convex Set Examples

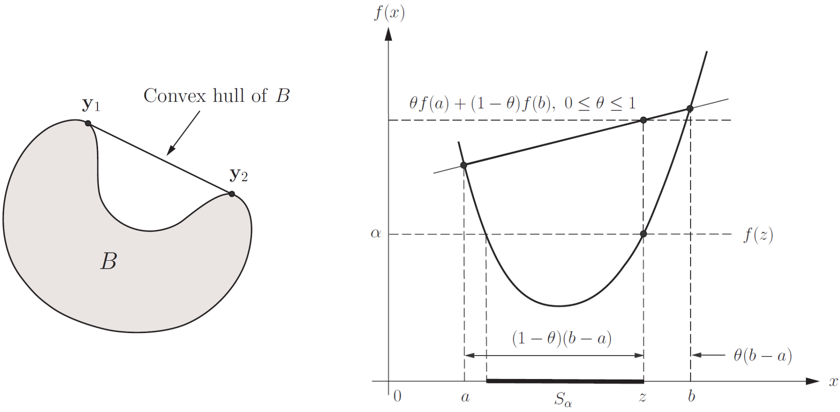




$$\textbf{conic } C \triangleq \left\{ \sum_{i=1}^{k} \theta_i \mathbf{x}_i \mid \mathbf{x}_i \in C, \theta_i \in \mathbb{R}_+, k \in \mathbb{Z}_{++} \right\}$$
$$= \{\theta \mathbf{x} \mid \mathbf{x} \in \textbf{conv } C, \theta \geq 0\}$$

- Left plot: **conic** C (called the *conic hull of* C) is a *convex cone* formed by $C = \{\mathbf{x}_1, \mathbf{x}_2\}$ via conic combinations, i.e., *the smallest conic set that contains* C; right plot: **conic** C formed by another set C (star).

Convex sets and convex functions-5



- Left plot: $\frac{\mathbf{y}_1+\mathbf{y}_2}{2} \notin B$, implying that B is not a convex set; right plot: $f(x)$ is a convex function (by (4)).

Convex sets and convex functions-6

- **Convex function:** f is convex if $\mathbf{dom}\ f$ (the domain of f) is a convex set, and for all $\mathbf{x}, \mathbf{y} \in \mathbf{dom}\ f$,

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta)f(\mathbf{y}),\ \forall\, 0 \leq \theta \leq 1. \quad (4)$$

  - f is *concave* if $-f$ is convex.

Some Examples of Convex Functions

- An *affine function* $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$ is both convex and concave on $\mathbb{R}^n$.
- $f(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x} + 2\mathbf{q}^T\mathbf{x} + r$, where $\mathbf{P} \in \mathbb{S}^n, \mathbf{q} \in \mathbb{R}^n$ and $r \in \mathbb{R}$ is convex if and only if $\mathbf{P} \in \mathbb{S}^n_+$.
- Every norm on $\mathbb{R}^n$ (e.g., $\|\cdot\|_p$ for $p \geq 1$) is convex.
- Linear function $f(\mathbf{X}) = \mathrm{Tr}(\mathbf{A}\mathbf{X})$ (where $\mathrm{Tr}(\mathbf{V})$ denotes the trace of a square matrix $\mathbf{V}$) is both convex and concave on $\mathbb{R}^{n\times n}$.
- $f(\mathbf{X}) = -\log\ \det(\mathbf{X})$ is convex on $\mathbb{S}^n_{++}$.

Ways Proving Convexity of a Function

First-order Condition

Suppose that f is differentiable. f is a convex function if and only if $\mathbf{dom}\ f$ is a convex set and

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbf{dom}\ f. \tag{5}$$

Second-order Condition

Suppose that f is twice differentiable. f is a convex function if and only if $\mathbf{dom}\ f$ is a convex set and

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}\ \text{(positive-semidefinite)}, \forall \mathbf{x} \in \mathbf{dom}\ f. \tag{6}$$

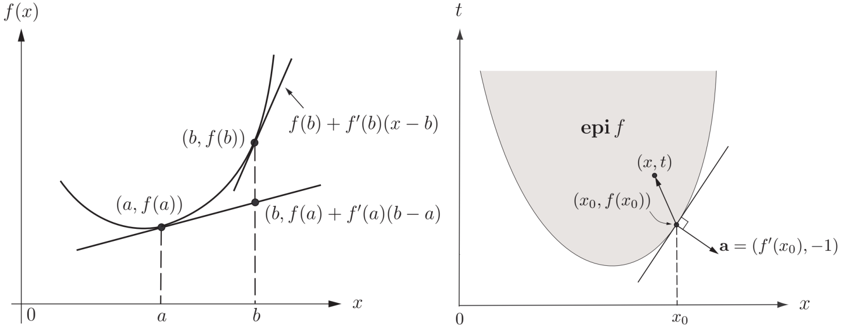
Epigraph

The *epigraph* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\mathbf{epi}\ f = \{(\mathbf{x}, t) \mid \mathbf{x} \in \mathbf{dom}\ f,\ f(\mathbf{x}) \leq t\} \subseteq \mathbb{R}^{n+1}. \tag{7}$$

A function f is convex if and only if $\mathbf{epi}\ f$ is a convex set.

First-order Condition and Epigraph

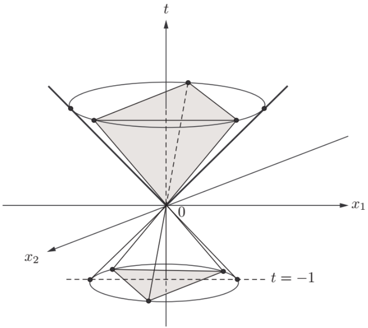


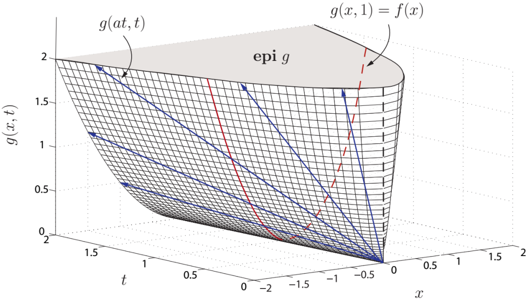
- Left plot: first-order condition for a convex function f for the one-dimensional case: $f(b) \geq f(a) + f'(a)(b - a)$, for all $a, b \in \mathbf{dom}\ f$; right plot: the epigraph of a convex function $f : \mathbb{R} \to \mathbb{R}$.

Convex Functions (Cont'd)

Convexity Preserving Operations

- *Intersection,* $\bigcap_i C_i$ *of convex sets* C_i *is a convex set;*
  Nonnegative weighted sum, $\sum_i \theta_i f_i$ *(where* $\theta_i \geq 0$*) of convex functions* f_i *is a convex function.*
- *Image,* $h(C) = \{h(\mathbf{x}) \mid \mathbf{x} \in C)\}$*, of a convex set* C *via affine mapping* $h(\mathbf{x}) \triangleq \mathbf{A}\mathbf{x} + \mathbf{b}$*, is a convex set;*
  Composition $f(h(\mathbf{x}))$ *of a convex function* f *with affine mapping* h *is a convex function;*
- *Image,* $p(C)$*, of a convex set* C *via perspective mapping* $p(\mathbf{x}, t) \triangleq \mathbf{x}/t$ *is a convex set;*
  Perspective, $g(\mathbf{x}, t) = tf(\mathbf{x}/t)$ *(where* $t > 0$*) of a convex function* f *is a convex function.*





- Left plot: pinhole camera interpretation of the perspective mapping $p(x, t) = x/t, \ t > 0$;
  right plot: epigraph of the perspective $g(x, t) = tf(x/t), \ t > 0$ of $f(x) = x^2$, where each ray is associated with $g(at, t) = a^2 t$ for a different value of a.

Convex optimization problem

- **Convex problem:**

$$\text{(CVXP)} \quad p^\star = \min_{\mathbf{x} \in \mathcal{C}} \ f(\mathbf{x}) \tag{8}$$

is a **convex problem** if the objective function $f(\cdot)$ is a **convex function** and $\mathcal{C}$ is a **convex set** (called the feasible set) in standard form as follows:

$$\mathcal{C} = \{\mathbf{x} \in \mathcal{D} \mid f_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0, i = 1, \ldots, m, j = 1, \ldots, p\},$$

where $f_i(\mathbf{x})$ is convex for all i and $h_j(\mathbf{x})$ is affine for all j and

$$\mathcal{D} = \textbf{dom}\, f \ \cap \ \left\{ \bigcap_{i=1}^{m} \textbf{dom}\ f_i \right\} \bigcap \left\{ \bigcap_{i=1}^{p} \textbf{dom}\ h_i \right\}$$

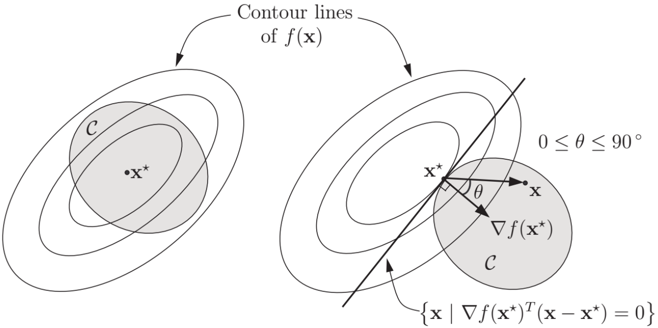
is called the problem domain.

- **Advantages:**
  - *Global optimality*: $\mathbf{x}^\star$ can be obtained by closed-form solution, analytically (algorithm), or numerically by convex solvers (e.g., CVX and SeDuMi).
  - *Computational complexity*: Polynomial-time solvable.
  - *Performance analysis*: KKT conditions are the backbone for analysis.

Global optimality and solution

- **An optimality criterion:** Any suboptimal solution to CVXP (8) is *globally optimal*. Assume that f is differentiable. Then a point $\mathbf{x}^\star \in \mathcal{C}$ is optimal *if and only if*

$$\nabla f(\mathbf{x}^\star)^T(\mathbf{x} - \mathbf{x}^\star) \geq 0, \ \forall \mathbf{x} \in \mathcal{C} \tag{9}$$



Case 1: $\nabla f(\mathbf{x}^\star) = \mathbf{0}_n$ Case 2: $\nabla f(\mathbf{x}^\star)^T(\mathbf{x} - \mathbf{x}^\star) \geq 0, \ \forall \mathbf{x} \in \mathcal{C}$

(where **int** $\mathcal{C} \neq \emptyset$ is assumed)

Global optimality and solution

- Besides the optimality criterion (9), a complementary approach for solving CVXP (8) is founded on the *"duality theory"*.

  - **Dual problem:**

$$
\begin{aligned}
&\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \triangleq f(\mathbf{x}) + \textstyle\sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \quad \text{(Lagrangian)} \\
&g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) > -\infty \quad \text{(dual function)} \\
&d^\star = \max \left\{ g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \mid \boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\nu} \in \mathbb{R}^p \right\} \quad \text{(dual problem)} \\
&\quad \leq p^\star = \min \left\{ f(\mathbf{x}) \mid \mathbf{x} \in \mathcal{C} \right\} \quad \text{(primal problem CVXP (8))}
\end{aligned}
\tag{10}
$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_p)$ are dual variables; " $\succeq$ " stands for an abbreviated *generalized inequality* defined by the *proper cone* $K = \mathbb{R}_+^m$, i.e., the first orthant, *(a closed convex solid and pointed cone)*, i.e., $\boldsymbol{\lambda} \succeq_K \mathbf{0} \Leftrightarrow \boldsymbol{\lambda} \in K$.

CVXP (8) and its dual can be solved simultaneously by solving the so-called *KKT conditions*.

Global optimality and solution

- **KKT conditions:**
  Suppose that $f, f_1, \ldots, f_m, h_1, \ldots, h_p$ are differentiable and $\mathbf{x}^\star$ is primal optimal and $(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ is dual optimal to CVXP (8). Under *strong duality*, i.e.,

$$p^\star = d^\star = \mathcal{L}(\mathbf{x}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$$

  (which holds true under *Slater's condition: a strictly feasible point exists, i.e.,* $\text{relint}\ \mathcal{C} \neq \emptyset$), the KKT conditions for solving $\mathbf{x}^\star$ *and* $(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ are as follows:

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^\star, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star) = \mathbf{0}, \tag{11a}$$

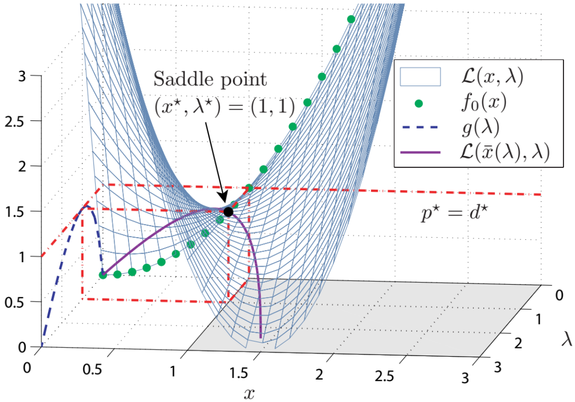
$$f_i(\mathbf{x}^\star) \leq 0,\ i = 1, \ldots, m, \tag{11b}$$

$$h_i(\mathbf{x}^\star) = 0,\ i = 1, \ldots, p, \tag{11c}$$

$$\lambda_i^\star \geq 0,\ i = 1, \ldots, m, \tag{11d}$$

$$\lambda_i^\star f_i(\mathbf{x}^\star) = 0,\ i = 1, \ldots, m. \ \textit{(complementary slackness)} \tag{11e}$$

  The above KKT conditions (11) and the optimality criterion (9) are equivalent under Slater's condition.



- Lagrangian $\mathcal{L}(x, \lambda)$, dual function $g(\lambda)$, and primal-dual optimal solution $(x^\star, \lambda^\star) = (1, 1)$ of the convex problem $\min\{f_0(x) = x^2 \mid (x-2)^2 \leq 1\}$ with strong duality. Note that $f_0(x^\star) = g(\lambda^\star) = \mathcal{L}(x^\star, \lambda^\star) = 1$.

Standard Convex Optimization Problems

Linear Programming (LP) - Inequality Form

$$\begin{aligned} \min \ & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \ & \mathbf{G}\mathbf{x} \preceq \mathbf{h}, \quad (\preceq \text{ stands for componentwise inequality}) \\ & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{12}$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{b} \in \mathbb{R}^p$, $\mathbf{G} \in \mathbb{R}^{m \times n}$, $\mathbf{h} \in \mathbb{R}^m$, and $\mathbf{x} \in \mathbb{R}^n$ is the unknown vector variable.

LP- Standard Form

$$\begin{aligned} \min \ & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \ & \mathbf{x} \succeq \mathbf{0}, \\ & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{13}$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{b} \in \mathbb{R}^p$, and $\mathbf{x} \in \mathbb{R}^n$ is the unknown vector variable.

Standard Convex Optimization Problems (Cont'd)

Quadratic Programming (QP): Convex if and only if $\mathbf{P} \succeq \mathbf{0}$ (i.e., $\mathbf{P}$ is positive semidefinite)

$$\begin{aligned} \min \quad & \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{G}\mathbf{x} \preceq \mathbf{h}, \end{aligned} \tag{14}$$

where $\mathbf{P} \in \mathbb{S}^n$, $\mathbf{G} \in \mathbb{R}^{m\times n}$, and $\mathbf{A} \in \mathbb{R}^{p\times n}$.

Quadratically constrained QP (QCQP): Convex if and only if $\mathbf{P}_i \succeq \mathbf{0}, \ \forall i$

$$\begin{aligned} \min \quad & \frac{1}{2}\mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0 \\ \text{s.t.} \quad & \frac{1}{2}\mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T\mathbf{x} + r_i \leq 0, \ i = 1, \ldots, m, \\ & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{15}$$

where $\mathbf{P}_i \in \mathbb{S}^n$, $i = 0, 1, \ldots, m$, and $\mathbf{A} \in \mathbb{R}^{p\times n}$.

Standard Convex Optimization Problems (Cont'd)

Second-order cone programming (SOCP)

$$\begin{aligned} \min \;& \mathbf{c}^T\mathbf{x} \\ \text{s.t. } & \|\mathbf{A}_i\mathbf{x}+\mathbf{b}_i\|_2 \le \mathbf{f}_i^T\mathbf{x}+d_i,\; i=1,\dots,m, \\ & \mathbf{F}\mathbf{x}=\mathbf{g}, \end{aligned} \tag{16}$$

where $\mathbf{A}_i \in \mathbb{R}^{n_i\times n}$, $\mathbf{b}_i \in \mathbb{R}^{n_i}$, $\mathbf{f}_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$, $\mathbf{F} \in \mathbb{R}^{p\times n}$, $\mathbf{g} \in \mathbb{R}^p$, $\mathbf{c} \in \mathbb{R}^n$, and $\mathbf{x} \in \mathbb{R}^n$ is the vector variable.

Semidefinite programming (SDP) - Standard Form

$$\begin{aligned} \min \;& \mathrm{Tr}(\mathbf{C}\mathbf{X}) \\ \text{s.t. } & \mathbf{X} \succeq \mathbf{0}, \\ & \mathrm{Tr}(\mathbf{A}_i\mathbf{X}) = b_i,\; i=1,\dots,p, \end{aligned} \tag{17}$$

with variable $\mathbf{X} \in \mathbb{S}^n$, where $\mathbf{A}_i \in \mathbb{S}^n$, $\mathbf{C} \in \mathbb{S}^n$, and $b_i \in \mathbb{R}$.

Alternating direction method of multipliers (ADMM)

- Consider the following convex optimization problem:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n,\mathbf{z}\in\mathbb{R}^m} \quad & f_1(\mathbf{x}) + f_2(\mathbf{z}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{S}_1, \ \ \mathbf{z} \in \mathcal{S}_2 \\ & \mathbf{z} = \mathbf{A}\mathbf{x} \end{aligned} \tag{18}$$

where $f_1 : \mathbb{R}^n \mapsto \mathbb{R}$ and $f_2 : \mathbb{R}^m \mapsto \mathbb{R}$ are convex functions, $\mathbf{A}$ is an $m \times n$ matrix, and $\mathcal{S}_1 \subset \mathbb{R}^n$ and $\mathcal{S}_2 \subset \mathbb{R}^m$ are nonempty convex sets.

- The considered dual problem of (18) is given by

$$\max_{\boldsymbol{\nu}\in\mathbb{R}^m} \min_{\mathbf{x}\in\mathcal{S}_1,\mathbf{z}\in\mathcal{S}_2} \left\{ f_1(\mathbf{x}) + f_2(\mathbf{z}) + \frac{c}{2}\|\mathbf{A}\mathbf{x} - \mathbf{z}\|_2^2 + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{z}) \right\}, \tag{19}$$

where c is a penalty parameter, and $\boldsymbol{\nu}$ is the dual variable associated with the equality constraint in (18).

ADMM (Cont'd)

- Inner minimization (convex problems):

$$\mathbf{z}(q+1) = \arg\min_{\mathbf{z}\in\mathcal{S}_2} \left\{ f_2(\mathbf{z}) - \boldsymbol{\nu}(q)^T\mathbf{z} + \frac{c}{2}\left\|\mathbf{A}\mathbf{x}(q) - \mathbf{z}\right\|_2^2 \right\}, \tag{20a}$$

$$\mathbf{x}(q+1) = \arg\min_{\mathbf{x}\in\mathcal{S}_1} \left\{ f_1(\mathbf{x}) + \boldsymbol{\nu}(q)^T\mathbf{A}\mathbf{x} + \frac{c}{2}\left\|\mathbf{A}\mathbf{x} - \mathbf{z}(q+1)\right\|_2^2 \right\}. \tag{20b}$$

ADMM Algorithm

1: Set $q = 0$, choose $c > 0$.
2: Initialize $\boldsymbol{\nu}(q)$ and $\mathbf{x}(q)$.
3: **repeat**
4: Solve (20a) and (20b) for $\mathbf{z}(q+1)$ and $\mathbf{x}(q+1)$ by two distributed equipments including *the information exchange of $\mathbf{z}(q+1)$ and $\mathbf{x}(q+1)$ between them;*
5: $\boldsymbol{\nu}(q+1) = \boldsymbol{\nu}(q) + c\,(\mathbf{A}\mathbf{x}(q+1) - \mathbf{z}(q+1))$;
6: $q := q+1$;
7: **until** the predefined stopping criterion is satisfied.

- When $\mathcal{S}_1$ is bounded or $\mathbf{A}^T\mathbf{A}$ is invertible, ADMM is guaranteed to converge and the obtained $\{\mathbf{x}(q), \mathbf{z}(q)\}$ is an optimal solution of problem (18).

Nonconvex problem

- **Reformulation into a convex problem:**
Equivalent representations (e.g. epigraph representations); function transformation; change of variables, etc.

- **Stationary-point solutions**: Provided that $\mathcal{C}$ is closed and convex but f is nonconvex, a point $\mathbf{x}^\star$ is a *stationary point* of the nonconvex problem (1) if

$$f'(\mathbf{x}^\star;\mathbf{v}) \triangleq \liminf_{\lambda\downarrow 0} \frac{f(\mathbf{x}^\star+\lambda\mathbf{v})-f(\mathbf{x}^\star)}{\lambda} \geq 0 \quad \forall \mathbf{x}^\star+\mathbf{v}\in\mathcal{C} \tag{21}$$

$$\Leftrightarrow \quad \nabla f(\mathbf{x}^\star)^T(\mathbf{x}-\mathbf{x}^\star)\geq 0 \quad \forall \mathbf{x}\in\mathcal{C} \quad \text{(when } f \text{ is differentiable)}$$

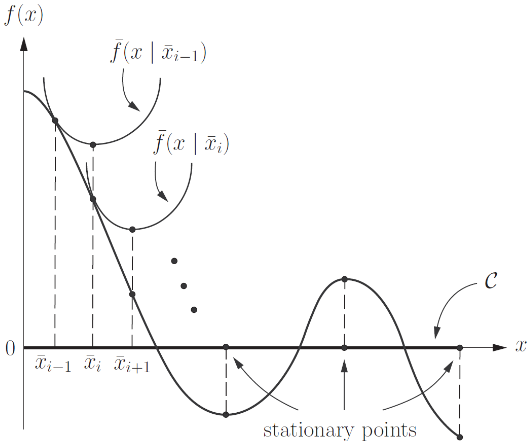
where $f'(\mathbf{x}^\star;\mathbf{v})$ is the *directional derivative* of f at a point $\mathbf{x}^\star$ in direction $\mathbf{v}$. Block successive upper bound minimization (BSUM) [Razaviyayn'13] guarantees a stationary-point solution under some convergence conditions.

• KKT points (i.e., solutions of KKT conditions) are also stationary points provided that the Slater's condition is satisfied.

[Razaviyayn'13] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optimiz.*, vol. 23, no. 2, pp. 11261153, 2013.

Stationary points and BSUM

- An illustration of stationary points of problem (1) for a nonconvex f and convex $\mathcal{C}$; convergence to a stationary point of (1) by BSUM.

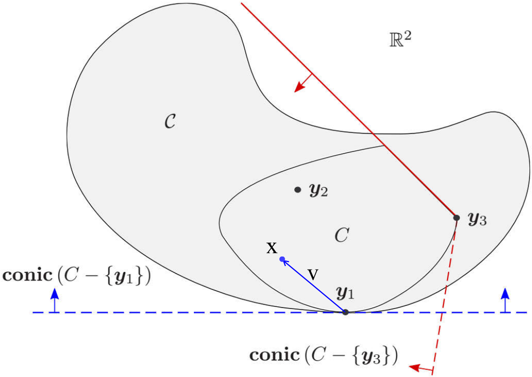


Stationary points for nonconvex feasible set

- An illustration of stationary points of problem (1) when both f and $\mathcal{C}$ are nonconvex. If $\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3$ are stationary points of $\min_{\mathbf{x}\in C} f(\mathbf{x})$ where $C \subset \mathcal{C}$ is convex, then **conic** $(C - \{\boldsymbol{y}_i\}) = \{\theta\mathbf{v} \mid \mathbf{v} \in C - \{\boldsymbol{y}_i\}, \theta \geq 0\}$ and

$$\mathcal{C} - \{\boldsymbol{y}_i\} \triangleq \{\mathbf{v} = \mathbf{x} - \{\boldsymbol{y}_i\} \mid \mathbf{x} \in \mathcal{C}\} \subset \textbf{conic}\ (C - \{\boldsymbol{y}_i\}),\ i = 1, 2$$

$$\Longrightarrow\ \boldsymbol{y}_1, \boldsymbol{y}_2 \text{ are also stationary points of (1).}$$



Nonconvex problem

- Approximate solutions to problem (1) when f is convex but $\mathcal{C}$ is nonconvex:
  - Convex *restriction* to $\mathcal{C}$: *Successive convex approximation (SCA)*

$$\boldsymbol{x}_i^\star = \arg\min_{\mathbf{x}\in C_i} f(\mathbf{x}) \in C_{i+1} \tag{22}$$

where $C_i \subset \mathcal{C}$ is convex for all i. Then

$$f(\boldsymbol{x}_{i+1}^\star) = \min_{\mathbf{x}\in C_{i+1}} f(\mathbf{x}) \leq f(\boldsymbol{x}_i^\star) \tag{23}$$

After convergence, an approximate solution $\boldsymbol{x}_i^\star$ is obtained.

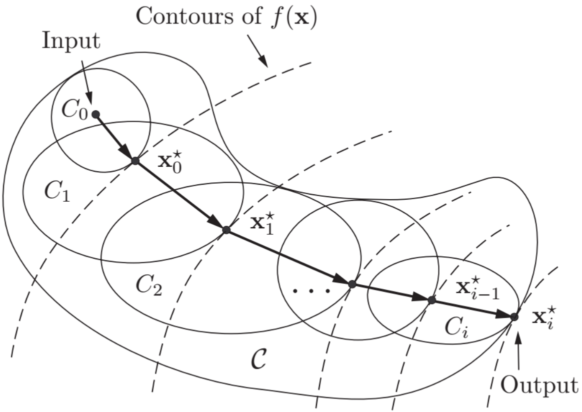
  - Convex *relaxation* to $\mathcal{C}$ (e.g., *semidefinite relaxation (SDR)*):

$$\begin{aligned} \mathcal{C}' &= \{\mathbf{X} \in \mathbb{S}_+^n \mid \operatorname{rank}(\mathbf{X}) = 1\} \subset \mathcal{C} \text{ relaxed to } \operatorname{conv} \mathcal{C}' = \mathbb{S}_+^n \text{ (SDR)}; \\ \mathcal{C}' &= \{-3, -1, +1, +3\} \subset \mathcal{C} \text{ relaxed to } \operatorname{conv} \mathcal{C}' = [-3, 3] \end{aligned} \tag{24}$$

The obtained $\boldsymbol{X}^\star$ or $\boldsymbol{x}^\star$ may not be feasible to problem (1); for SDR, a good approximate solution can be obtained from $\boldsymbol{X}^\star$ via *Gaussian randomization*.

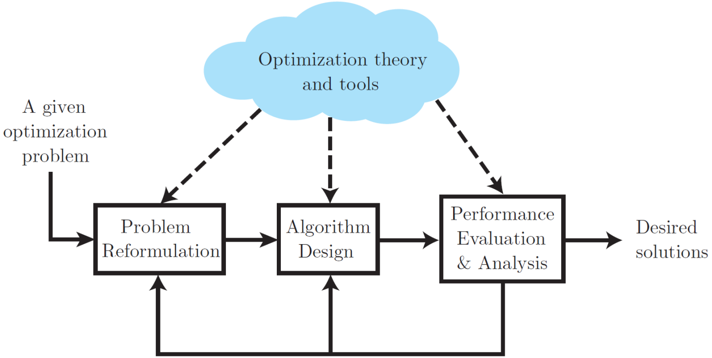
Successive Convex Approximation (SCA)

- Illustration of SCA for (1) when f is convex but $\mathcal{C}$ is nonconvex, yielding a solution $\mathbf{x}_i^\star$ (which is a stationary point under some mild condition).



Algorithm development

- **Foundamental theory and tools:** Calculus, linear algebra, matrix analysis and computations, convex sets, convex functions, convex problems (e.g., geometric program (GP), LP, QP, SOCP, SDP), duality, interior-point method; CVX and SeDuMi.



A new book

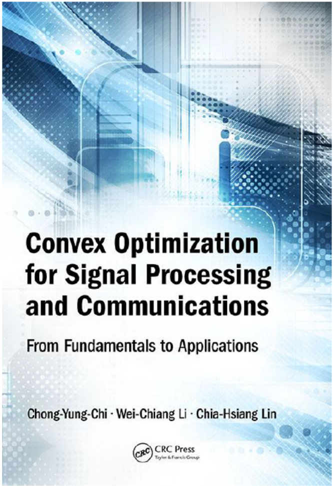
Convex Optimization for Signal Processing and Communications: From Fundamentals to Applications

Chong-Yung Chi, Wei-Chiang Li, Chia-Hsiang Lin

(Publisher: CRC Press, 2017, 432 pages, ISBN: 9781498776455)

Motivation: Most of mathematical books are hard to read for engineering students and professionals due to *lack of enough fundamental details and tangible linkage* between mathematical theory and applications.

- The book is written in a *causally sequential fashion*; namely, one can *review/peruse the related materials introduced in early chapters/sections again,* to overpass hurdles in reading.
- *Covers* convex optimization *from fundamentals to advanced applications*, while holding a *strong link from theory to applications*.
- *Provides comprehensive proofs and perspective interpretations*, many insightful figures, examples and remarks to illuminate core convex optimization theory.



- *Illustrates*, by *cutting-edge applications*, how to apply the convex optimization theory, like a *guided journey/exploration* rather than pure mathematics.
- Has been used for a *2-week short course* under the book title at 9 major universities *(Shandong Univ, Tsinghua Univ, Tianjin Univ, BJTU, Xiamen Univ., UESTC, SYSU, BUPT, SDNU)* in Mainland China more than 17 times since early 2010.

Thank you for your attention!

Acknowledgment: Financial support by NTHU; my students, visiting students and scholars, short-course participants for almost uncountable questions, interactions and comments over the last decade.