Introduction to Optimization

Vincent Leclère (CERMICS, ENPC)

Autumn 2026

Why should I bother to learn this stuff?

- Many useful engineering models are convex: least squares, energy minimization, resource allocation, economic dispatch, . . .
- Recognizing convexity tells you that a numerical solution can often be trusted as a global optimum, not merely a locally good point.
- Convex models can often be solved reliably even when they have many variables and constraints.
- Dual variables can quantify the value of additional resources or relaxed constraints.

Contents

Convex sets — recall

♡

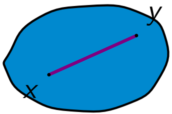
- C is convex if for any two points x and y in C the segment $[x, y] \subset C$, *i.e.*

$$\forall x, y \in C, \ \forall \theta \in [0, 1], \ \theta x + (1 - \theta) y \in C.$$

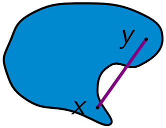
- The convex hull of C as the set of convex combination of elements of C, *i.e.*

$$\mathrm{conv}(C) := \Big\{ \sum_{i=1}^{K} \theta_i x_i \ | \ \forall x_i \in C, \\ \forall \theta_i \in [0, 1], \ \sum_{i=1}^{K} \theta_i = 1, \ \forall i \in [K], \ \forall K \in \mathbb{N} \Big\}$$

- $\mathrm{conv}(C)$ is the smallest convex set containing C.



convex



nonconvex

Examples of convex sets

- Linear or affine (sub)-space, half-space, hyperplane.
- Ball (for any norm) and ellipsoid.
- Polyhedron (finite intersection of half-spaces).
- The set of symmetric positive (semi-)definite matrices.

Exercise. Prove these results.

Some sets are not convex:

- The set of integers $\mathbb{Z}$.
- $\{x \in \mathbb{R}^n \mid \|x\|_2 \geq 1\}$.
- The set of orthogonal matrices.

Exercise. Prove these results.

Examples of convex sets

- Linear or affine (sub)-space, half-space, hyperplane.
- Ball (for any norm) and ellipsoid.
- Polyhedron (finite intersection of half-spaces).
- The set of symmetric positive (semi-)definite matrices.

Exercise. Prove these results.

Some sets are not convex:

- The set of integers $\mathbb{Z}$.
- $\{x \in \mathbb{R}^n \mid \|x\|_2 \geq 1\}$.
- The set of orthogonal matrices.

Exercise. Prove these results.

Operations preserving convexity

Assume that all sets denoted by C (indexed or not) are convex.

- $C_1 + C_2$ and $C_1 \times C_2$ are convex sets.
- For any arbitrary index set $\mathcal{I}$ the intersection $\bigcap_{i \in \mathcal{I}} C_i$ is convex.
- Let f be an affine function. Then $f(C)$ and $f^{-1}(C)$ are convex.
- In particular, $C + x_0$, and tC are convex. The projection of C on any affine space is convex.

Exercise. Prove these results.

Contents

Convex functions — recall

♡

- A function $f : X \to \bar{\mathbb{R}}$ is convex if its epigraph is convex.
- $f : X \to \mathbb{R} \cup \{+\infty\}$ is convex iff

$$\forall t \in [0,1],\ \forall x, y \in X,$$
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

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

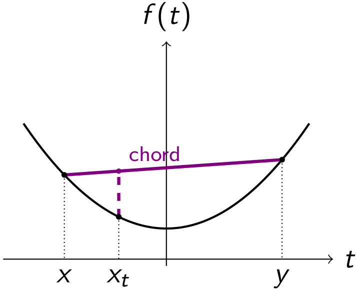


Figure: The graph of a convex function lies below its chords.

Convex function and tangent

Theorem

A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if it is always above its tangents:

$$f(x) \geq f(y) + \nabla f(y)^\top (x - y). \quad \forall x, \forall y$$

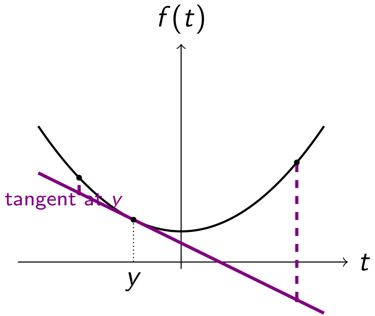


Figure: A differentiable convex function lies above every tangent.

Hessian matrix and semidefinite matrices

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. The Hessian matrix of f at x is the matrix of second derivatives

$$\nabla^2 f(x) := \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}.$$

- $\nabla^2 f(x)$ is symmetric.

- A symmetric matrix A is
  - positive semidefinite ($A \succeq 0$) if $\forall z \in \mathbb{R}^n, z^\top A z \geq 0$
  - positive definite ($A \succ 0$) if $\forall z \in \mathbb{R}^n \setminus \{0\}, z^\top A z > 0$
- If A is symmetric, then:
  - A is positive semidefinite iff all its eigenvalues are nonnegative (≥ 0),
  - A is positive definite iff all its eigenvalues are positive (> 0).
  - $A - \alpha I$ is positive semidefinite (also written $A \succeq \alpha I$) iff all its eigenvalues are $\geq \alpha$.

Hessian matrix and semidefinite matrices

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. The Hessian matrix of f at x is the matrix of second derivatives
$$\nabla^2 f(x) := \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}.$$
- $\nabla^2 f(x)$ is symmetric.
- A symmetric matrix A is
  - positive semidefinite ($A \succcurlyeq 0$) if $\forall z \in \mathbb{R}^n, z^\top A z \geq 0$
  - positive definite ($A \succ 0$) if $\forall z \in \mathbb{R}^n \setminus \{0\}, z^\top A z > 0$
- If A is symmetric, then:
  - A is positive semidefinite iff all its eigenvalues are nonnegative (≥ 0),
  - A is positive definite iff all its eigenvalues are positive (> 0).
  - $A - \alpha I$ is positive semidefinite (also written $A \succcurlyeq \alpha I$) iff all its eigenvalues are $\geq \alpha$.

Hessian matrix and semidefinite matrices

- Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. The Hessian matrix of f at x is the matrix of second derivatives

$$\nabla^2 f(x) := \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}.$$

- $\nabla^2 f(x)$ is symmetric.

- A symmetric matrix A is
  - positive semidefinite ($A \succeq 0$) if $\forall z \in \mathbb{R}^n, z^\top A z \geq 0$
  - positive definite ($A \succ 0$) if $\forall z \in \mathbb{R}^n \setminus \{0\}, z^\top A z > 0$
- If A is symmetric, then:
  - A is positive semidefinite iff all its eigenvalues are nonnegative (≥ 0),
  - A is positive definite iff all its eigenvalues are positive (> 0).
  - $A - \alpha I$ is positive semidefinite (also written $A \succeq \alpha I$) iff all its eigenvalues are $\geq \alpha$.

Convex functions: strict and strong convexity

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is strictly convex iff

$$\forall t \in]0,1[, \quad \forall x, y \in X, \qquad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-convex iff $\forall t \in]0,1[, \quad \forall x, y \in X,$

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x-y\|^2$$

- If $f \in C^1(\mathbb{R}^n)$
  - $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ iff f convex
  - if strict inequality holds, then f strictly convex
  - $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-convex iff $\forall x, y \in X$

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$$

- If $f \in C^2(\mathbb{R}^n)$,
  - $\nabla^2 f(x) \succeq 0$, for all x, iff f convex
  - if $\nabla^2 f(x) \succ 0$, for all x, then f strictly convex
  - if $\nabla^2 f(x) \succeq \alpha I$, for all x, then f is α-convex

Convex functions: strict and strong convexity

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is strictly convex iff
$$\forall t \in]0,1[, \quad \forall x, y \in X, \qquad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$
- $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-convex iff $\forall t \in]0,1[, \quad \forall x, y \in X,$
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x-y\|^2$$
- If $f \in C^1(\mathbb{R}^n)$
  - $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ iff f convex
  - if strict inequality holds, then f strictly convex
  - $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-convex iff $\forall x, y \in X$
$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$$
- If $f \in C^2(\mathbb{R}^n)$,
  - $\nabla^2 f(x) \succeq 0$, for all x, iff f convex
  - if $\nabla^2 f(x) \succ 0$, for all x, then f strictly convex
  - if $\nabla^2 f(x) \succeq \alpha I$, for all x, then f is α-convex

Convex functions: strict and strong convexity

- $f : X \to \mathbb{R} \cup \{+\infty\}$ is strictly convex iff
$$\forall t \in]0,1[, \quad \forall x, y \in X, \qquad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$
- $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-convex iff $\forall t \in]0,1[, \quad \forall x, y \in X,$
$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}\alpha t(1-t)\|x-y\|^2$$
- If $f \in C^1(\mathbb{R}^n)$
  - $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ iff f convex
  - if strict inequality holds, then f strictly convex
  - $f : X \to \mathbb{R} \cup \{+\infty\}$ is α-convex iff $\forall x, y \in X$
$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\alpha}{2}\|y - x\|^2$$
- If $f \in C^2(\mathbb{R}^n)$,
  - $\nabla^2 f(x) \succcurlyeq 0$, for all x, iff f convex
  - if $\nabla^2 f(x) \succ 0$, for all x, then f strictly convex
  - if $\nabla^2 f(x) \succcurlyeq \alpha I$, for all x, then f is α-convex

Important examples

- The indicator function of a set $C \subset X$,

$$\mathbb{I}_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \textit{otherwise} \end{cases}$$

  is convex iff C is convex.
- $x \mapsto e^{ax}$ is convex for any $a \in \mathbb{R}$
- $x \mapsto \|x\|^q$ is convex for $q \geq 1$ and any norm
- $x \mapsto \ln(x)$ is concave
- $x \mapsto x\ln(x)$ is convex

Contents

Basic properties

- If f, g convex, $t > 0$, then $tf + g$ is convex.
- ⚠ If f, g convex, then $f \circ g$ not necessarily convex, but:
  - If f convex non-decreasing, g convex, then $f \circ g$ convex.
  - If f convex and a affine, then $f \circ a$ is convex.
- If $(f_i)_{i \in I}$ is a family of convex functions, then $\sup_{i \in I} f_i$ is convex.
- The domain and the sublevel sets of a convex function are convex.

Exercise. Prove these results

Basic properties

- If f, g convex, $t > 0$, then $tf + g$ is convex.
- ⚠ If f, g convex, then $f \circ g$ not necessarily convex, but:
  - If f convex non-decreasing, g convex, then $f \circ g$ convex.
  - If f convex and a affine, then $f \circ a$ is convex.
- If $(f_i)_{i \in I}$ is a family of convex functions, then $\sup_{i \in I} f_i$ is convex.
- The domain and the sublevel sets of a convex function are convex.

Exercise. Prove these results

Basic properties

- If f, g convex, $t > 0$, then $tf + g$ is convex.
- ⚠ If f, g convex, then $f \circ g$ not necessarily convex, but:
  - If f convex non-decreasing, g convex, then $f \circ g$ convex.
  - If f convex and a affine, then $f \circ a$ is convex.
- If $(f_i)_{i \in I}$ is a family of convex functions, then $\sup_{i \in I} f_i$ is convex.
- The domain and the sublevel sets of a convex function are convex.

Exercise. Prove these results

Convex function: regularity

Consider a convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$.

- f is continuous (on $\mathbb{R}^n$) if and only if $\mathrm{dom}(f) = \mathbb{R}^n$ (i.e., if it is finite everywhere)
- f is continuous on the interior of its domain
- f is Lipschitz continuous on any compact subset of the interior of its domain

Convex functions: existence, uniqueness and characterization of minima ♡

- If f is convex, then any local minimum is a global minimum. Further, the set of global minimizers is convex.
- If f is strictly convex, then there exists at most one global minimizer.
- If $f : \mathbb{R}^n \to \mathbb{R}$ is α-convex, then it admits a unique global minimizer.

Exercise. Prove these results.

First-order optimality conditions are necessary and sufficient in the convex case:

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable. Then $x^\sharp$ is a global minimizer of f iff $\nabla f(x^\sharp) = 0$.

Convex functions: existence, uniqueness and characterization of minima ♡

- If f is convex, then any local minimum is a global minimum. Further, the set of global minimizers is convex.
- If f is strictly convex, then there exists at most one global minimizer.
- If $f : \mathbb{R}^n \to \mathbb{R}$ is α-convex, then it admits a unique global minimizer.

Exercise. Prove these results.

First-order optimality conditions are necessary and sufficient in the convex case:

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable. Then $x^\sharp$ is a global minimizer of f iff $\nabla f(x^\sharp) = 0$.

First-order optimality on a convex set

Let $X \subseteq \mathbb{R}^n$ be convex and let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable.

Theorem

A point $x^\sharp \in X$ is a global minimizer of

$$\operatorname*{Min}_{x \in X} f(x)$$

if and only if

$$\nabla f(x^\sharp)^\top d \geq 0 \qquad \forall d \in T_X(x^\sharp).$$

Thus, for convex problems, the constrained first-order necessary condition from Lecture 1 is also sufficient.

Contents

Abstract convex optimization problems

An abstract optimization problem of the form

$$\underset{x \in X}{\text{Min}} \quad f(x)$$

is convex if $f : X \to \overline{\mathbb{R}}$ is convex and X is a convex set.

A convex optimization problem has nice properties:

- any local minimum is a global minimum,
- if f is strictly convex, there exists at most one global minimizer,
- if $f : \mathbb{R}^n \to \mathbb{R}$ is α-convex and X is nonempty, closed and convex, there exists a unique global minimizer,
- first-order optimality conditions are necessary and sufficient.

Abstract convex optimization problems

An abstract optimization problem of the form

$$\underset{x \in X}{\mathrm{Min}} \quad f(x)$$

is convex if $f : X \to \overline{\mathbb{R}}$ is convex and X is a convex set.

A convex optimization problem has nice properties:

- any local minimum is a global minimum,
- if f is strictly convex, there exists at most one global minimizer,
- if $f : \mathbb{R}^n \to \mathbb{R}$ is α-convex and X is nonempty, closed and convex, there exists a unique global minimizer,
- first-order optimality conditions are necessary and sufficient.

Differentiable convex optimization problem

$$\begin{aligned}
\underset{x\in\mathbb{R}^n}{\text{Min}} \quad & f(x) && (P)\\
\text{s.t.} \quad & g_i(x) = 0 && \forall i \in [n_E]\\
& h_j(x) \leq 0 && \forall j \in [n_I]
\end{aligned}$$

This is called a convex differentiable optimization problem if f is convex, each g_i is affine, and each h_j is convex differentiable.

Exercise. Show that a convex differentiable optimization problem is a convex optimization problem.

Why bother with convex optimization?

- Many applications:
  - Portfolio optimization
  - Statistical regression
  - Model fitting
  - Optimal / Robust control
  - Signal processing
  - ...
- Strong theoretical results:
  - KKT are necessary and sufficient (no local optima or saddle point)
  - Strong duality theory
  - Complexity theory
- Efficient algorithms in theory and practice
  - ➥ Algorithms can be applied to other settings in a more or less heuristic way

Contents

KKT - convex case

Theorem (KKT - convex case)

Consider a convex differentiable optimization problem

$$\underset{\tilde{x}\in\mathbb{R}^n}{\mathrm{Min}} \left\{ f(\tilde{x}) \quad \middle| \quad g_i(\tilde{x}) = 0,\ \forall i \in [n_E], \quad h_j(\tilde{x}) \leq 0,\ \forall j \in [n_I] \right\}$$

If the constraints are qualified[a] *at x, then x is optimal if and only if there exists dual variables λ, μ such that*

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{n_E} \lambda_i \nabla g_i(x) + \sum_{j=1}^{n_I} \mu_j \nabla h_j(x) = 0 & \nabla_x \mathcal{L} = 0 \\ g(x) = 0, \quad h(x) \leq 0 & \textit{Primal feasibility} \\ \lambda \in \mathbb{R}^{n_E}, \quad \mu \in \mathbb{R}_+^{n_I} & \textit{dual feasibility} \\ \mu_j h_j(x) = 0 \qquad \forall j \in [n_I] & \textit{complementary slackness} \end{cases}$$

[a] Even without qualification, in the convex case, any primal–dual pair satisfying KKT is optimal.

Slater qualification condition

$$
\begin{aligned}
\underset{x\in\mathbb{R}^n}{\text{Min}} \quad & f(x) & & (P) \\
\text{s.t.} \quad & g_i(x) = 0 & & \forall i \in [n_E] \\
& h_j(x) \leq 0 & & \forall j \in [n_I]
\end{aligned}
$$

where f is convex, each g_i is affine, and each h_j is convex differentiable.

We say that $x_S \in \mathbb{R}^n$ is a Slater point if it strictly satisfies the inequality constraints, *i.e.*

- $g_i(x_S) = 0$ for all $i \in [n_E]$,
- $h_j(x_S) < 0$ for all $j \in [n_I]$.

Theorem

If there exists a Slater point x_S, then the constraints are qualified at every feasible point.

Slater qualification condition

$$
\begin{aligned}
\underset{x\in\mathbb{R}^n}{\text{Min}} \quad & f(x) && (P)\\
\text{s.t.} \quad & g_i(x) = 0 && \forall i \in [n_E]\\
& h_j(x) \leq 0 && \forall j \in [n_I]
\end{aligned}
$$

where f is convex, each g_i is affine, and each h_j is convex differentiable.

We say that $x_S \in \mathbb{R}^n$ is a Slater point if it strictly satisfies the inequality constraints, *i.e.*

- $g_i(x_S) = 0$ for all $i \in [n_E]$,
- $h_j(x_S) < 0$ for all $j \in [n_I]$.

Theorem

If there exists a Slater point x_S, then the constraints are qualified at every feasible point.

Contents

Lagrangian duality — recall

For the primal problem

$$
\begin{aligned}
(P) \quad & \underset{x \in \mathbb{R}^n}{\mathrm{Min}} \quad f(x) \\
& \text{s.t.} \quad g_i(x) = 0 \qquad \forall i \in [n_E], \\
& \qquad\quad h_j(x) \leq 0 \qquad \forall j \in [n_I],
\end{aligned}
$$

recall the Lagrangian and the dual function

$$\mathcal{L}(x; \lambda, \mu) := f(x) + \lambda^\top g(x) + \mu^\top h(x),$$

$$d(\lambda, \mu) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x; \lambda, \mu), \qquad (\lambda, \mu) \in \mathbb{R}^{n_E} \times \mathbb{R}_+^{n_I}.$$

The dual problem is

$$(D) \qquad \underset{\lambda \in \mathbb{R}^{n_E}, \mu \in \mathbb{R}_+^{n_I}}{\mathrm{Max}} \quad d(\lambda, \mu),$$

and weak duality gives

$$d(\lambda, \mu) \leq \mathrm{val}(D) \leq \mathrm{val}(P) \leq f(x)$$

for every primal-feasible x and dual-feasible (λ, μ).

Lagrangian duality — recall

For the primal problem

$$
\begin{aligned}
(P) \quad & \underset{x \in \mathbb{R}^n}{\mathrm{Min}} \quad f(x) & \\
& \text{s.t.} \quad g_i(x) = 0 & \forall i \in [n_E], \\
& \qquad\quad h_j(x) \leq 0 & \forall j \in [n_I],
\end{aligned}
$$

recall the Lagrangian and the dual function

$$\mathcal{L}(x; \lambda, \mu) := f(x) + \lambda^\top g(x) + \mu^\top h(x),$$

$$d(\lambda, \mu) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x; \lambda, \mu), \qquad (\lambda, \mu) \in \mathbb{R}^{n_E} \times \mathbb{R}_+^{n_I}.$$

The dual problem is

$$(D) \qquad \underset{\lambda \in \mathbb{R}^{n_E}, \mu \in \mathbb{R}_+^{n_I}}{\mathrm{Max}} d(\lambda, \mu),$$

and weak duality gives

$$d(\lambda, \mu) \leq \mathrm{val}(D) \leq \mathrm{val}(P) \leq f(x)$$

for every primal-feasible x and dual-feasible (λ, μ).

Strong duality and KKT conditions

If (P) is convex and there exists a Slater point, then

$$\mathrm{val}(P) = \mathrm{val}(D),$$

i.e.

$$\inf_{x} \sup_{\substack{\lambda \in \mathbb{R}^{n_E} \\ \mu \in \mathbb{R}_+^{n_I}}} \mathcal{L}(x; \lambda, \mu) = \sup_{\substack{\lambda \in \mathbb{R}^{n_E} \\ \mu \in \mathbb{R}_+^{n_I}}} \inf_{x} \mathcal{L}(x; \lambda, \mu).$$

Theorem

If (P) is convex and there exists a Slater point, then the following assertions are equivalent:

1. $x^\sharp$ *is an optimal solution of* (P),
2. ($\exists\, \lambda^\sharp, \mu^\sharp$ *such that*) $(x^\sharp; \lambda^\sharp, \mu^\sharp)$ *is a saddle point of* $\mathcal{L}$,
3. ($\exists\, \lambda^\sharp, \mu^\sharp$ *such that*) $(x^\sharp; \lambda^\sharp, \mu^\sharp)$ *satisfies the KKT conditions,*
4. $\lambda^\sharp, \mu^\sharp$ *is an optimal solution of* (D).

Strong duality and KKT conditions

If (P) is convex and there exists a Slater point, then

$$\mathrm{val}(P) = \mathrm{val}(D),$$

i.e.

$$\inf_{x} \sup_{\substack{\lambda \in \mathbb{R}^{n_E} \\ \mu \in \mathbb{R}_+^{n_I}}} \mathcal{L}(x;\lambda,\mu) = \sup_{\substack{\lambda \in \mathbb{R}^{n_E} \\ \mu \in \mathbb{R}_+^{n_I}}} \inf_{x} \mathcal{L}(x;\lambda,\mu).$$

Theorem

If (P) is convex and there exists a Slater point, then the following assertions are equivalent:

1. $x^\sharp$ *is an optimal solution of* (P),
2. $(\exists\, \lambda^\sharp, \mu^\sharp$ *such that*$)$ $(x^\sharp; \lambda^\sharp, \mu^\sharp)$ *is a saddle point of* $\mathcal{L}$,
3. $(\exists\, \lambda^\sharp, \mu^\sharp$ *such that*$)$ $(x^\sharp; \lambda^\sharp, \mu^\sharp)$ *satisfies the KKT conditions,*
4. $\lambda^\sharp, \mu^\sharp$ *is an optimal solution of* (D).

Contents

Marginal interpretation of the dual variables ♡

Consider the perturbed value function

$$\begin{aligned} v(q) = \underset{x\in\mathbb{R}^n}{\text{Min}} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \leq q. \end{aligned}$$

In particular, $v(0) = \mathrm{val}(P)$.

Assume that (P) is convex and satisfies Slater's condition, and let $(\lambda^\sharp, \mu^\sharp)$ be optimal dual variables at $q = 0$. Then

$$\boxed{v(q) \geq v(0) - (\mu^\sharp)^\top q.}$$

Thus $-\mu^\sharp$ gives the marginal sensitivity of the optimal value to the right-hand side. In particular, if v is differentiable at 0,

$$\boxed{\nabla_q v(0) = -\mu^\sharp.}$$

Exercise. Prove this result

Marginal interpretation of the dual variables

Consider the perturbed value function

$$
\begin{aligned}
v(q) = \underset{x\in\mathbb{R}^n}{\mathrm{Min}} \quad & f(x) \\
\text{s.t.} \quad & g(x) = 0, \\
& h(x) \leq q.
\end{aligned}
$$

In particular, $v(0) = \mathrm{val}(P)$.

Assume that (P) is convex and satisfies Slater's condition, and let $(\lambda^\sharp, \mu^\sharp)$ be optimal dual variables at $q = 0$. Then

$$\boxed{v(q) \geq v(0) - (\mu^\sharp)^\top q.}$$

Thus $-\mu^\sharp$ gives the marginal sensitivity of the optimal value to the right-hand side. In particular, if v is differentiable at 0,

$$\boxed{\nabla_q v(0) = -\mu^\sharp.}$$

Exercise. Prove this result

Marginal interpretation of the dual variables

Consider the perturbed value function

$$\begin{aligned} v(q) = \underset{x \in \mathbb{R}^n}{\mathrm{Min}} \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0, \\ & h(x) \leq q. \end{aligned}$$

In particular, $v(0) = \mathrm{val}(P)$.

Assume that (P) is convex and satisfies Slater's condition, and let $(\lambda^\sharp, \mu^\sharp)$ be optimal dual variables at $q = 0$. Then

$$\boxed{v(q) \geq v(0) - (\mu^\sharp)^\top q.}$$

Thus $-\mu^\sharp$ gives the marginal sensitivity of the optimal value to the right-hand side. In particular, if v is differentiable at 0,

$$\boxed{\nabla_q v(0) = -\mu^\sharp.}$$

Exercise. Prove this result

Exercise

Exercise. Consider the following problem, for $b \in \mathbb{R}$,

$$\begin{aligned} \underset{x\in\mathbb{R}}{\text{Min}} \quad & x^2 \\ \text{s.t.} \quad & x \leq b \end{aligned}$$

1. Do optimal dual variables exist?
2. Without solving the dual, give the optimal dual variable μ_b.

Recovering KKT conditions from Lagrangian duality ♡

$$
\begin{aligned}
(P) \quad \underset{x\in\mathbb{R}^n}{\text{Min}} \quad & f(x) \\
\text{s.t.} \quad & Ax = b \\
& h_j(x) \leq 0 \qquad\qquad \forall j \in [n_I]
\end{aligned}
$$

with associated Lagrangian

$$\mathcal{L}(x;\lambda,\mu) := f(x) + \lambda^\top(Ax-b) + \mu^\top h(x)$$

In the convex case, for $\mu \geq 0$, the map $x \mapsto \mathcal{L}(x;\lambda,\mu)$ is convex. The KKT conditions can therefore be seen as:

1. $\nabla_x \mathcal{L}(x;\lambda,\mu) = 0$ (Lagrangian minimized in x)
2. $Ax = b$, $h(x) \leq 0$ (x primal feasible; $Ax = b$ also follows from $\nabla_\lambda \mathcal{L} = 0$)
3. $\mu \geq 0$ ((λ,μ) dual feasible)
4. $\mu_j = 0$ or $h_j(x) = 0$, for all $j \in [n_I]$ (complementary slackness $\rightsquigarrow 2^{n_I}$ possibilities).

Recovering KKT conditions from Lagrangian duality ♡

$$
(P) \quad \underset{x\in\mathbb{R}^n}{\text{Min}} \quad f(x)
$$
$$
\text{s.t.} \quad Ax = b
$$
$$
h_j(x) \leq 0 \qquad \forall j \in [n_I]
$$

with associated Lagrangian

$$
\mathcal{L}(x;\lambda,\mu) := f(x) + \lambda^\top (Ax - b) + \mu^\top h(x)
$$

In the convex case, for $\mu \geq 0$, the map $x \mapsto \mathcal{L}(x;\lambda,\mu)$ is convex. The KKT conditions can therefore be seen as:

1. $\nabla_x \mathcal{L}(x;\lambda,\mu) = 0$ (Lagrangian minimized in x)
2. $Ax = b$, $h(x) \leq 0$ (x primal feasible; $Ax = b$ also follows from $\nabla_\lambda \mathcal{L} = 0$)
3. $\mu \geq 0$ ((λ,μ) dual feasible)
4. $\mu_j = 0$ or $h_j(x) = 0$, for all $j \in [n_I]$ (complementary slackness $\rightsquigarrow$ 2^{n_I} possibilities).

What you should know

- What a convex set and a convex function are, and the main operations preserving convexity.
- The distinction between convex, strictly convex and strongly convex functions.
- The consequences of convexity for global optimality, existence and uniqueness.
- For convex problems, KKT conditions become sufficient; under Slater's condition they characterize optimality and strong duality holds.
- The marginal interpretation of optimal dual variables.

What you should be able to do

- Show that a simple set or function is convex.
- Recognize whether an optimization problem is convex.
- Use strict or strong convexity to establish uniqueness or existence of a solution.
- Interpret a dual multiplier as a marginal value.