Convex Optimization II

Lecture 6: Convex Optimization Problems and Lagrange Duality

Hamed Shah-Mansouri

Department of Electrical Engineering

Sharif University of Technology

1400-2

Reference

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*, first edition, Cambridge University Press, 2004.

All materials and figures in this lecture are from [1].

Thanks to Prof. Vincent Wong and Prof. Stephen Boyd for all the slides used in this lecture.

Optimization Problem in Standard Form

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

- $\mathbf{x} \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, p$, are the equality constraint functions
- Domain $\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\ f_i \cap \bigcap_{i=1}^{p} \mathbf{dom}\ h_i$

Optimal value:

$$
p^\star = \inf\{f_0(\mathbf{x}) \mid f_i(\mathbf{x}) \leq 0, i = 1, \ldots, m, \ \ h_i(\mathbf{x}) = 0, i = 1, \ldots, p\}
$$

- $p^\star = \infty$ if problem is infeasible (no $\mathbf{x}$ satisfies the constraints)
- $p^\star = -\infty$ if problem is unbounded below

Optimal and Locally Optimal Points

- $\mathbf{x}$ is feasible if $\mathbf{x} \in \mathcal{D}$ and it satisfies the constraints
- A feasible $\mathbf{x}^\star$ is optimal if $f_0(\mathbf{x}^\star) = p^\star$; X_{opt} is the set of optimal points
- $\tilde{\mathbf{x}}$ is locally optimal if there is an $R > 0$ such that $\tilde{\mathbf{x}}$ is optimal for

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \quad h_i(\mathbf{x}) = 0, \quad i = 1, \dots, p \\ & ||\mathbf{x} - \tilde{\mathbf{x}}||_2 \leq R \end{array}$$

Examples: (with $n = 1$, $m = p = 0$)

- $f_0(x) = 1/x,\ \textbf{dom}\ f_0 = \mathbf{R}_{++}: \ \ p^\star = 0$, but the optimal value is not achieved
- $f_0(x) = -\log x,\ \textbf{dom}\ f_0 = \mathbf{R}_{++}: \ \ p^\star = -\infty$, this problem is unbounded below
- $f_0(x) = x \log x,\ \textbf{dom}\ f_0 = \mathbf{R}_{++}: \ \ p^\star = -1/e,\ x^\star = 1/e$
- $f_0(x) = x^3 - 3x,\ p^\star = -\infty$, local optimal at $x = 1$

Implicit Constraints

- The standard form optimization problem has an implicit constraint

$$\mathbf{x} \in \mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\ f_i \cap \bigcap_{i=1}^{p} \mathbf{dom}\ h_i$$

- The constraints $f_i(\mathbf{x}) \leq 0$ and $h_i(\mathbf{x}) = 0$ are the explicit constraints.
- A problem is unconstrained if it has no explicit constraints ($m = p = 0$).

  Example:

$$\text{minimize } f_0(\mathbf{x}) = -\sum_{i=1}^{k} \log(b_i - \mathbf{a}_i^T \mathbf{x})$$

  is an unconstrained problem with implicit constraints $\mathbf{a}_i^T \mathbf{x} < b_i$

Feasibility Problem

$$\begin{array}{ll} \text{find} & \mathbf{x} \\ \text{subject to} & f_i(\mathbf{x}) \le 0, \quad i = 1, \dots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \dots, p \end{array}$$

can be considered a special case of the general problem with $f_0(\mathbf{x}) = 0$.

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & f_i(\mathbf{x}) \le 0, \quad i = 1, \dots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \dots, p \end{array}$$

- $p^\star = 0$ if the feasible set is nonempty; any feasible $\mathbf{x}$ is optimal.
- $p^\star = \infty$ if the feasible set is empty.

Convex Optimization Problem

- Standard form convex optimization problem

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) \le 0, \quad i = 1, \dots, m \\ & a_i^T \mathbf{x} = b_i, \quad i = 1, \dots, p, \end{aligned}$$

  where $f_0, f_1, \dots f_m$ are convex; equality constraints are affine.

- Often written as

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) \le 0, \quad i = 1, \dots, m \\ & A\mathbf{x} = \mathbf{b}. \end{aligned}$$

- Important property: the feasible set of a convex problem is convex.

- If the objective function $g_0(\mathbf{x})$ is concave, then maximizing $g_0(\mathbf{x})$ is equivalent to minimizing $-g_0(\mathbf{x})$.

Local and Global Optima

Any locally optimal point of a convex problem is (globally) optimal.

Proof:

- Suppose x is locally optimal for a convex problem. That is, x is feasible and

$$f_0(x) = \inf\{f_0(z) \mid z \text{ feasible}, ||z - x||_2 \leq R\} \tag{2}$$

  for some $R > 0$.

- Suppose x is not globally optimal. There is a feasible y such that $f_0(y) < f_0(x)$. Clearly, $||y - x||_2 > R$.

- Consider the point $w = (1 - \theta)x + \theta y$, where $\theta = \frac{R}{2||y-x||_2}$.

- We have $||w - x||_2 = R/2 < R$. By convexity of the feasible set, w is feasible.

- By convexity of f_0, we have

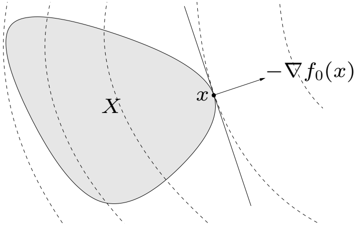
$$f_0(w) \leq (1 - \theta)f_0(x) + \theta f_0(y) < f_0(x),$$

  which contradicts (2). Hence, there exists no feasible y with $f_0(y) < f_0(x)$. That is, x is globally optimal.

Optimality Criterion for Differentiable f_0 (Convex Problem)

x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T(y - x) \geq 0 \quad \text{for all feasible } y$$



if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

Special Cases

- **unconstrained problem**

$$\nabla f_0(x) = 0$$

- **minimization over nonnegative orthant**

$$\text{minimize } f_0(x) \qquad \text{subject to } x \succeq 0$$

The optimality condition can be expressed as

$$x \succeq 0, \qquad \nabla f_0(x) \succeq 0, \qquad x_i(\nabla f_0(x))_i = 0, \quad i = 1, \ldots, n.$$

Equivalent Problems

We call two problems equivalent if from a solution of one, a solution of the other is readily found, and vice versa.

Example:

$$\begin{array}{ll} \text{minimize} & \alpha_0 f_0(\mathbf{x}) \\ \text{subject to} & \alpha_i f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \beta_i h_i(\mathbf{x}) = 0, \quad i = 1, \dots, p \end{array} \tag{3}$$

where $\alpha_i > 0, i = 0, 1, \dots, m$, and $\beta_i \neq 0, i = 1, \dots, p$.

Problems (1) and (3) are equivalent.

However, they are not the same unless α_i and β_i are all equal to one.

General transformations that can yield equivalent problems include:

- Change of variables
- Transformation of objective and constraint functions
- Optimizing over some variables

Change of Variables

- Suppose $\phi : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one, with image covering domain $\mathcal{D}$ in problem (1).
- We define functions

$$\tilde{f}_i(\mathbf{z}) = f_i(\phi(\mathbf{z})), \quad i = 0, \ldots, m, \qquad \tilde{h}_i(\mathbf{z}) = h_i(\phi(\mathbf{z})), \quad i = 1, \ldots, p.$$

- Consider the problem

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(\mathbf{z}) \\ \text{subject to} & \tilde{f}_i(\mathbf{z}) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_i(\mathbf{z}) = 0, \quad i = 1, \ldots, p \end{array} \tag{4}$$

- Problems (1) and (4) are related by change of variable $\mathbf{x} = \phi(\mathbf{z})$.
- Problems (1) and (4) are equivalent.
- If $\mathbf{x}$ solves problem (1), then $\mathbf{z} = \phi^{-1}(\mathbf{x})$ solves problem (4).
- If $\mathbf{z}$ solves problem (4), then $\mathbf{x} = \phi(\mathbf{z})$ solves problem (1).

Transformation of Objective and Constraint Functions

- Suppose function $\psi_0 : \mathbf{R} \to \mathbf{R}$ is monotone increasing. That is, for all $x, y \in \mathbf{dom}\ \psi_0$ such that $x \leq y$, we have $\psi_0(x) \leq \psi_0(y)$
- Functions $\psi_i, \ldots, \psi_m : \mathbf{R} \to \mathbf{R}$ satisfy $\psi_i(u) \leq 0$ if and only if $u \leq 0$
- Functions $\psi_{m+1}, \ldots, \psi_{m+p} : \mathbf{R} \to \mathbf{R}$ satisfy $\psi_i(u) = 0$ if and only if $u = 0$
- We define functions

$$\tilde{f}_i(\mathbf{x}) = \psi_i(f_i(\mathbf{x})), \quad i = 0, \ldots, m, \qquad \tilde{h}_i(\mathbf{x}) = \psi_{m+i}(h_i(\mathbf{x})), \quad i = 1, \ldots, p.$$

- Consider the problem

$$\begin{aligned} \text{minimize} \quad & \tilde{f}_0(\mathbf{x}) \\ \text{subject to} \quad & \tilde{f}_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \\ & \tilde{h}_i(\mathbf{x}) = 0, \quad i = 1, \ldots, p \end{aligned} \tag{5}$$

- Problems (1) and (5) are equivalent.
- The feasible sets are identical. The optimal points are also identical.

Example: Transformation of Objective Function

- Consider the unconstrained Euclidean norm minimization problem,

$$\text{minimize} \ \ ||A\mathbf{x} - \mathbf{b}||_2,$$

  with variable $\mathbf{x} \in \mathbf{R}^n$.

- Since the norm is non-negative, we can as well solve the problem

$$\text{minimize} \ \ ||A\mathbf{x} - \mathbf{b}||_2^2 = (A\mathbf{x} - \mathbf{b})^T(A\mathbf{x} - \mathbf{b}).$$

- These two problems are equivalent.
- Note that the objective function in the first problem is not differentiable at any $\mathbf{x}$ with $A\mathbf{x} - \mathbf{b} = 0$.
- The objective function in the second problem is differentiable for all $\mathbf{x}$.

Example: Transformation of Constraint Functions

$$
\begin{array}{ll}
\text{minimize} & f_0(x) = x_1^2 + x_2^2 \\
\text{subject to} & f_1(x) = x_1/(1 + x_2^2) \leq 0 \\
& h_1(x) = (x_1 + x_2)^2 = 0
\end{array}
$$

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (according to our definition): f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

$$
\begin{array}{ll}
\text{minimize} & x_1^2 + x_2^2 \\
\text{subject to} & x_1 \leq 0 \\
& x_1 + x_2 = 0
\end{array}
$$

Example: Geometric Programming Problem

- GP in standard form is not a convex optimization problem.

$$\begin{array}{lll} \underset{\mathbf{x} \succ \mathbf{0}}{\text{minimize}} & \sum_{k=1}^{K_0} d_{0k}\, x_1^{a_{0k}^{(1)}} x_2^{a_{0k}^{(2)}} \cdots\ x_n^{a_{0k}^{(n)}} & \\ \text{subject to} & \sum_{k=1}^{K_i} d_{ik}\, x_1^{a_{ik}^{(1)}} x_2^{a_{ik}^{(2)}} \cdots\ x_n^{a_{ik}^{(n)}} \leq 1, & i = 1, \ldots, m, \\ & d_l\, x_1^{a_l^{(1)}} x_2^{a_l^{(2)}} \cdots\ x_n^{a_l^{(n)}} = 1, & l = 1, \ldots, M. \end{array}$$

- The main idea to convert a GP into a convex problem is based on a logarithmic change of variables, and a logarithmic transformation of the objective and constraint functions.

- Consider a logarithmic change of all the variables and multiplicative constants:

$$y_i = \log x_i,\ \ (\text{so } x_i = e^{y_i}), \qquad i = 1, \ldots, n,$$

$$b_{ik} = \log d_{ik},\ \ (\text{so } d_{ik} = e^{b_{ik}}), \qquad k = 1, \ldots, K_i,\ i = 0, 1, \ldots, m,$$

$$b_l = \log d_l,\ \ (\text{so } d_l = e^{b_l}), \qquad l = 1, \ldots, M.$$

Example: Geometric Programming Problem

- After the logarithmic change of variables, the problem becomes (Why?)

$$
\begin{array}{lll}
\underset{\boldsymbol{y}}{\text{minimize}} & \displaystyle\sum_{k=1}^{K_0} \exp\left(\boldsymbol{a}_{0k}^T \boldsymbol{y} + b_{0k}\right) & \\
\text{subject to} & \displaystyle\sum_{k=1}^{K_i} \exp\left(\boldsymbol{a}_{ik}^T \boldsymbol{y} + b_{ik}\right) \le 1, & i = 1, \ldots, m, \\
& \exp\left(\boldsymbol{a}_l^T \boldsymbol{y} + b_l\right) = 1, & l = 1, \ldots, M,
\end{array}
$$

where

$$
\boldsymbol{a}_{ik}^T = \begin{bmatrix} a_{ik}^{(1)} & a_{ik}^{(2)} & \cdots & a_{ik}^{(n)} \end{bmatrix},
$$

$$
\boldsymbol{a}_l^T = \begin{bmatrix} a_l^{(1)} & a_l^{(2)} & \cdots & a_l^{(n)} \end{bmatrix},
$$

and variables $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$.

Example: Geometric Programming Problem

- After a logarithmic transformation of the objective and constraint functions, we have the following problem

$$
\begin{array}{lll}
\underset{\boldsymbol{y}}{\text{minimize}} & \log \sum_{k=1}^{K_0} \exp\left(\boldsymbol{a}_{0k}^T \boldsymbol{y} + b_{0k}\right) & \\
\text{subject to} & \log \sum_{k=1}^{K_i} \exp\left(\boldsymbol{a}_{ik}^T \boldsymbol{y} + b_{ik}\right) \leq 0, & i = 1, \dots, m, \\
& \boldsymbol{a}_l^T \boldsymbol{y} + b_l = 0, & l = 1, \dots, M.
\end{array}
$$

- Is the above problem convex? Why?

Optimizing over Some Variables

- We always have

$$\inf_{x,y} f(x,y) = \inf_{x} \tilde{f}(x)$$

  where $\tilde{f}(x) = \inf_y f(x,y)$.

- We can always minimize a function by first minimizing over some of the variables, and then minimizing over the remaining ones.

Optimizing over Some Variables: Example

- Consider a problem with strictly convex quadratic objective (i.e., $P_{11}, P_{22} \succeq 0$), with some of the variables unconstrained

$$\begin{array}{ll} \text{minimize} & \mathbf{x}_1^T P_{11}\mathbf{x}_1 + 2\mathbf{x}_1^T P_{12}\mathbf{x}_2 + \mathbf{x}_2^T P_{22}\mathbf{x}_2 \\ \text{subject to} & f_i(\mathbf{x}_1) \leq 0, \quad i = 1, \dots, m \end{array}$$

- We can first analytically minimize over $\mathbf{x}_2$:

$$\inf_{\mathbf{x}_2} \left(\mathbf{x}_1^T P_{11}\mathbf{x}_1 + 2\mathbf{x}_1^T P_{12}\mathbf{x}_2 + \mathbf{x}_2^T P_{22}\mathbf{x}_2\right) = \mathbf{x}_1^T \left(P_{11} - P_{12}P_{22}^{-1}P_{12}^T\right)\mathbf{x}_1$$

- The original problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{x}_1^T \left(P_{11} - P_{12}P_{22}^{-1}P_{12}^T\right)\mathbf{x}_1 \\ \text{subject to} & f_i(\mathbf{x}_1) \leq 0, \quad i = 1, \dots, m. \end{array}$$

Duality Mentality

- Bound or solve an optimization problem via a different optimization problem.

Lagrangian

standard form problem (not necessarily convex)

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

variable $x \in \mathbf{R}^n$, domain $\mathcal{D}$, optimal value $p^\star$

Lagrangian: $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with $\mathbf{dom}\, L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$,

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

- weighted sum of objective and constraint functions
- λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$
- ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange Dual Function

- Lagrange dual function: $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

$$\begin{aligned} g(\lambda,\nu) &= \inf_{x\in\mathcal{D}} L(x,\lambda,\nu) \\ &= \inf_{x\in\mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right) \end{aligned}$$

  g is concave, can be $-\infty$ for some λ, ν.

- Lower bound property: if $\lambda \succeq 0$, then $g(\lambda,\nu) \le p^\star$

  Proof: if $\tilde{x}$ is feasible and $\lambda \succeq 0$, we have

$$L(\tilde{x},\lambda,\nu) = f_0(\tilde{x}) + \sum_{i=1}^{m} \lambda_i f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x}) \le f_0(\tilde{x})$$

  Thus,

$$g(\lambda,\nu) = \inf_{x\in\mathcal{D}} L(x,\lambda,\nu) \le L(\tilde{x},\lambda,\nu) \le f_0(\tilde{x})$$

  Minimizing over all feasible $\tilde{x}$ gives $g(\lambda,\nu) \le p^\star$.

The Dual Problem

Lagrange dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

- finds best lower bound on $p^\star$, obtained from Lagrange dual function
- a convex optimization problem; optimal value denoted $d^\star$
- λ, ν are dual feasible if $\lambda \succeq 0$, $(\lambda, \nu) \in \mathbf{dom}\, g$
- often simplified by making implicit constraint $(\lambda, \nu) \in \mathbf{dom}\, g$ explicit

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array} \qquad \begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0 \end{array}$$

Weak and strong duality

Weak Duality: $d^\star \leq p^\star$

- always holds for convex and nonconvex problems
- can be used to find nontrivial lower bounds for difficult problems

Strong Duality: $d^\star = p^\star$

- does not hold in general
- (usually) holds for convex problems
- conditions that guarantee strong duality in convex problems are called constraint qualifications

Duality Gap: $p^\star - d^\star$

- gives the gap between the optimal value of the primal problem (i.e., problem (1)) and the best (i.e., greatest) lower bound obtained from dual problem
- always non-negative

Slater's Constraint Qualification

- strong duality holds for a convex problem

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

- constraint qualifications provide conditions on the problem, beyond convexity, under which strong duality holds.

- Slater's condition : There exist $\mathbf{x} \in \mathbf{int}\, \mathcal{D}$ such that

$$f_i(\mathbf{x}) < 0, \quad i = 1, \ldots, m, \quad A\mathbf{x} = \mathbf{b}.$$

- can be sharpened. If the first k constraint functions are affine, then strong duality holds provided the following condition holds: There exist $\mathbf{x} \in \mathbf{relint}\, \mathcal{D}$ (interior relative to affine hull) such that

$$f_i(\mathbf{x}) \leq 0, \;\; i = 1, \ldots, k, \quad f_i(\mathbf{x}) < 0, \;\; i = k+1, \ldots, m, \quad A\mathbf{x} = \mathbf{b}.$$

A Nonconvex Problem with Strong Duality

- Strong duality holds for any optimization problem with quadratic objective and one quadratic inequality constraint, provided Slater's condition holds

- Example:

$$\begin{array}{ll} \text{minimize} & x^T A x + 2 b^T x \\ \text{subject to} & x^T x \leq 1, \end{array}$$

where $A \in \mathbf{S}^n$ and $b \in \mathbf{R}^n$.

When $A \nsucceq 0$, this is not a convex problem.

However, it can be shown that we have zero duality gap for this problem.

Complementary Slackness

assume strong duality holds, $x^\star$ is primal optimal, $(\lambda^\star, \nu^\star)$ is dual optimal

$$\begin{aligned} f_0(x^\star) = g(\lambda^\star, \nu^\star) &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^\star f_i(x) + \sum_{i=1}^p \nu_i^\star h_i(x) \right) \\ &\leq f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star f_i(x^\star) + \sum_{i=1}^p \nu_i^\star h_i(x^\star) \\ &\leq f_0(x^\star) \end{aligned}$$

hence, the two inequalities hold with equality

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0$ for $i = 1, \ldots, m$ (known as complementary slackness):

$$\lambda_i^\star > 0 \Longrightarrow f_i(x^\star) = 0, \qquad f_i(x^\star) < 0 \Longrightarrow \lambda_i^\star = 0$$

Karush-Kuhn-Tucker (KKT) conditions for Non-Convex Problem

- Assume functions $f_0, \ldots, f_m, h_1, \ldots, h_p$ are differentiable. They may not be convex.
- If strong duality holds and $x^\star$, $\lambda^\star$, $\nu^\star$ are optimal points, then they must satisfy the following KKT conditions:

$$\begin{aligned} f_i(x^\star) &\leq 0, \quad i = 1, \ldots, m \\ h_i(x^\star) &= 0, \quad i = 1, \ldots, p \\ \lambda_i^\star &\geq 0, \quad i = 1, \ldots, m \\ \lambda_i^\star f_i(x^\star) &= 0, \quad i = 1, \ldots, m \\ \nabla f_0(x^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla f_i(x^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(x^\star) &= 0. \end{aligned}$$

- The last condition is due to the fact that $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$ over x, its gradient must vanish.

KKT Conditions for Convex Problem

- If f_i are convex, h_i are affine, $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ are points that satisfy the KKT conditions,

$$\begin{aligned} f_i(\tilde{x}) &\leq 0, \quad i = 1, \ldots, m \\ h_i(\tilde{x}) &= 0, \quad i = 1, \ldots, p \\ \tilde{\lambda}_i &\geq 0, \quad i = 1, \ldots, m \\ \tilde{\lambda}_i f_i(\tilde{x}) &= 0, \quad i = 1, \ldots, m \\ \nabla f_0(\tilde{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^{p} \tilde{\nu}_i \nabla h_i(\tilde{x}) &= 0, \end{aligned}$$

  then $\tilde{x}$ and $(\tilde{\lambda}, \tilde{\nu})$ are primal and dual optimal, with zero duality gap.
- Since $\tilde{\lambda}_i \geq 0$ and $L(x, \tilde{\lambda}, \tilde{\nu})$ is convex in x, the last condition implies $\tilde{x}$ minimizes $L(x, \tilde{\lambda}, \tilde{\nu})$. We have

$$g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) = f_0(\tilde{x}) + \sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{x}) + \sum_{i=1}^{p} \tilde{\nu}_i h_i(\tilde{x}) = f_0(\tilde{x})$$

- Thus, $\tilde{x}$ and $(\tilde{\lambda}, \tilde{\nu})$ have zero duality gap, and are primal and dual optimal.
- If a convex problem satisfies Slater's condition, then the KKT conditions provide necessary and sufficient conditions for optimality.

Example: Water-filling (with $\alpha_i > 0$)

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^n \log(x_i + \alpha_i) \\ \text{subject to} & x \succeq 0, \quad \mathbf{1}^T x = 1 \end{array}$$

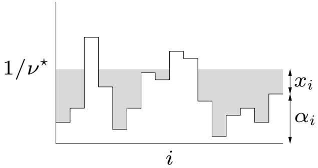
x is optimal iff $x \succeq 0$, $\mathbf{1}^T x = 1$, and there exist $\lambda \in \mathbf{R}^n$, $\nu \in \mathbf{R}$ such that

$$\lambda \succeq 0, \qquad \lambda_i x_i = 0, \qquad \frac{1}{x_i + \alpha_i} + \lambda_i = \nu$$

- if $\nu < 1/\alpha_i$: $\lambda_i = 0$ and $x_i = 1/\nu - \alpha_i$
- if $\nu \geq 1/\alpha_i$: $\lambda_i = \nu - 1/\alpha_i$ and $x_i = 0$
- determine ν from $\mathbf{1}^T x = \sum_{i=1}^n \max\{0, 1/\nu - \alpha_i\} = 1$

interpretation

- n patches; level of patch i is at height α_i
- flood area with unit amount of water
- resulting level is $1/\nu^\star$



How to Solve a Convex Optimization Problem?

- There are several efficient software to solve convex optimization problems
- Examples include CVX (http://cvxr.com/cvx), MOSEK (http://www.mosek.com/), CPLEX, and Matlab
- In MATLAB, you can use the optimization toolbox. The key function you need to use is *fmincon*:
- https://www.mathworks.com/help/optim/ug/fmincon.html
- [x, fval, exitflag] = fmincon(fun, x0, A, b, Aeq, beq, lb, ub, nonlcon, options)

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) \leq 0 \\ & ceq(x) = 0 \\ & A \cdot x \leq b \\ & Aeq \cdot x = beq \\ & lb \leq x \leq ub \end{array}$$

Summary

- Convexity mentality: Convex optimization is *nice* for several reasons:
  - local optimum is global optimum
  - zero duality gap (under technical conditions)
  - KKT optimality conditions are necessary and sufficient
- Duality mentality: Can always bound primal through dual, sometimes solve primal through dual
- Reading: Sections 4.2, 5.1, 5.2 and 5.5 in Boyd and Vandenberghe.

Example: Geometric Programming Problem

$$
\begin{aligned}
\underset{x_1, x_2}{\text{maximize}} \quad & \log x_1 \log x_2 + \log x_1 \\
\text{subject to} \quad & \log x_1 + \mathrm{e}^{\log x_1 \log x_2} \leq 1
\end{aligned}
$$

$$
\begin{aligned}
\underset{x_1, x_2}{\text{minimize}} \; & -x_1 + x_2 \log x_2 \\
\text{subject to} \; & x_1 + x_2 = 1 \\
& x_1, x_2 \geq 0
\end{aligned}
$$