Optimization methods. Seminar 11. Intro to duality.

Alexandr Katrutsa

Moscow Institute of Physics and Technology
Department of Control and Applied Mathematics

June 27, 2018

Reminder

- Existence solution of the optimization problem
- Optimality conditions for
  - general optimization problem
  - unconstrained optimization problem
  - equality constrained optimization problem
  - equality and inequality constrained optimization problem

Notations

Problem

$$\min f(x) = p^*$$
$$\text{s.t. } g_i(x) = 0,\ i = 1, \ldots, m$$
$$h_j(x) \leq 0,\ j = 1, \ldots, p$$

Lagrangian

$$L(x, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x)$$

Dual variables

Vectors $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are called *dual variables*.

Dual function

A function $g(\boldsymbol{\mu}, \boldsymbol{\lambda}) = \inf\limits_x L(x, \boldsymbol{\lambda}, \boldsymbol{\mu})$ is called *dual function*.

Dual function properties

Concavity

The dual function is concave as infimum of affine functions of $(\boldsymbol{\mu}, \boldsymbol{\lambda})$ independently of convexity of the primal problem.

Lower bound

For all $\boldsymbol{\lambda}$ and for $\boldsymbol{\mu} \geq 0$ the following holds $g(\boldsymbol{\mu}, \boldsymbol{\lambda}) \leq p^*$.

Dual problem

$$\max g(\boldsymbol{\mu}, \boldsymbol{\lambda}) = d^*$$
$$\text{s.t. } \boldsymbol{\mu} \geq 0$$

What for?

- Dual problem is concave independently on convexity of primal problem
- Lower bound **can be** tight

Conjugate function again

Consider the problem

$$\begin{aligned} &\min f_0(x) \\ \text{s.t. } &\mathbf{A}\mathbf{x} \leq \mathbf{b} \\ &\mathbf{C}\mathbf{x} = \mathbf{d} \end{aligned}$$

Then

$$\begin{aligned} g(\boldsymbol{\lambda}, \boldsymbol{\mu}) &= \inf_{\mathbf{x}} (f_0(\mathbf{x}) + \boldsymbol{\lambda}^\intercal(\mathbf{A}\mathbf{x} - \mathbf{b}) + \boldsymbol{\mu}^\intercal(\mathbf{C}\mathbf{x} - \mathbf{d})) = \\ &- \mathbf{b}^\intercal\boldsymbol{\lambda} - \boldsymbol{\mu}^\intercal\mathbf{d} + \inf_{\mathbf{x}}(f_0(\mathbf{x}) + (\mathbf{A}^\intercal\boldsymbol{\lambda} + \mathbf{C}^\intercal\boldsymbol{\mu})^\intercal\mathbf{x}) = \\ &- \mathbf{b}^\intercal\boldsymbol{\lambda} - \boldsymbol{\mu}^\intercal\mathbf{d} - f_0^*(-\mathbf{A}^\intercal\boldsymbol{\lambda} - \mathbf{C}^\intercal\boldsymbol{\mu}) \end{aligned}$$

Domains of dual and conjugate functions are related:

$$\text{dom } g = \{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mid -\mathbf{A}^\intercal\boldsymbol{\lambda} - \mathbf{C}^\intercal\boldsymbol{\mu} \in \text{dom } f_0^*\}$$

Examples

Find dual function:

- Minimal norm solution of linear system

$$\begin{aligned} & \min \|\mathbf{x}\|_2^2 \\ \text{s.t. } & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

- Linear programming

$$\begin{aligned} & \min \mathbf{c}^\intercal \mathbf{x} \\ \text{s.t. } & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

- Partitioning problem

$$\begin{aligned} & \min \mathbf{x}^\intercal \mathbf{W} \mathbf{x} \\ \text{s.t. } & x_i^2 = 1, \; i = 1, \ldots, n \end{aligned}$$

Weak and strong duality

Definition

Optimal values of the primal objective and dual objective are related as

$$d^* \leq p^*.$$

If $d^* < p^*$, then *weak* duality holds.
If $d^* = p^*$, then *strong* duality holds.

Remark

Weak duality always holds by construction of the dual problem

Questions

- When the strong duality is hold?
- How to use duality to test optimality?

Suboptimality criterion

By construction $p^* \geq g(\boldsymbol{\lambda}, \boldsymbol{\mu})$, therefore
$f_0(x) - p^* \leq f_0(x) - g(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \varepsilon$.

Definition

Difference $f_0(x) - g(\boldsymbol{\lambda}, \boldsymbol{\mu})$ is called *duality gap* and gives upper bound for difference between current function value and optimal one.

How to use:

- stopping criterion in iterative process
- theoretical estimate of convergence speed
- check optimality of given point

Slater's condition

Theorem

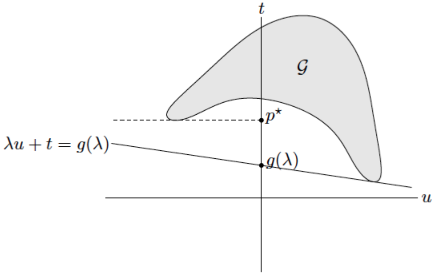
If a problem is convex and there exists x inside the interior of the feasible set, i.e. inequality constraints hold with strict inequalities, then the strong duality holds.

- Solution of linear system with minimum norm
- Linear programming
- Quadratically Constrained Quadratic Program (QCQP)

Geometric interpretation

$$\min_x f_0(x), \text{ where } f_1(x) \le 0.$$

$$g(\lambda) = \inf_{(u,t)\in\mathcal{G}} (t + \lambda u) \qquad \mathcal{G} = \{(f_1(x), f_0(x)) \mid x \in \mathcal{D}\}$$



- $\lambda = 0$
- λ^* — optimal value
- $\lambda > \lambda^*$

Complementary slackness condition

Let $\mathbf{x}^*$ and $(\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ be solutions of the primal and dual problems, therefore

$$
\begin{aligned}
& f(\mathbf{x}^*) = g(\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \\
& f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* g_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j^* h_j(\mathbf{x}^*) \leq \\
& f(\mathbf{x}^*), \qquad \boldsymbol{\mu} \geq 0
\end{aligned}
$$

Complementary slackness condition

$$
\mu_j^* h_j(\mathbf{x}^*) = 0, \qquad j = 1, \ldots, p
$$

For every inequality constraint:

- Lagrange multiplier is zero
- inequality is active

KKT conditions

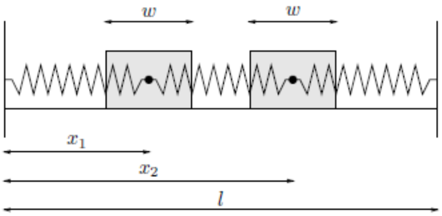
From the last seminar we know KKT conditions:

1. $g_i(x^*) = 0$ — primal feasibility
2. $h_j(x^*) \leq 0$ — primal feasibility
3. $\mu_j^* \geq 0$ — dual feasibility
4. $\mu_j^* h_j(x^*) = 0$ — complementary slackness
5. $\nabla_x L(x^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = 0$ — stationariness of Lagrangian in primal variables

Example ($\mathbf{P} \in \mathbb{S}^n_+$)

$$
\begin{aligned}
&\min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2}\mathbf{x}^\intercal \mathbf{P}\mathbf{x} + \mathbf{q}^\intercal \mathbf{x} + r \\
\text{s.t. } &\mathbf{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

Mechanical interpretation



Search equilibrium state of the system:

$$
\begin{aligned}
&\min_{\mathbf{x}\in\mathbb{R}^3} \frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2(x_2 - x_1)^2 + \frac{1}{2}k_3(l - x_2)^2\\
\text{s.t. } & \frac{w}{2} - x_1 \leq 0\\
& w + x_1 - x_2 \leq 0\\
& \frac{w}{2} - l + x_2 \leq 0
\end{aligned}
$$

Examples

- Negative entropy with linear constraints

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^n} & \sum_{i=1}^n x_i \log x_i \\
\text{s.t. } & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\
& \mathbf{1}^\mathsf{T}\mathbf{x} = 1
\end{aligned}
$$

- State dual problem, solve it and recover solution of the primal problem from the dual solution:

$$
\begin{aligned}
& \min \frac{1}{2}x^2 + 2y^2 + \frac{1}{2}z^2 + x + y + 2z \\
\text{s.t. } & x + 2y + z = 4
\end{aligned}
$$

- Lagrange relaxation for the binary linear programming problem:

$$
\begin{aligned}
& \min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^\mathsf{T}\mathbf{x} \\
\text{s.t. } & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\
& x_i \in \{0, 1\}, \quad i = 1, \dots, n
\end{aligned}
$$

Recap

- Dual problem: what and why?
- Weak and strong duality
- Slater's condition
- Geometrical and mechanical interpretations