Lec. 15 (11/10/2022)

Example: Dual of LP

Primal problem: $p^* = \min_{\underline{x} \in \mathbb{R}^n} \langle \underline{c}, \underline{x} \rangle$

s.t. $G\underline{x} \preceq \underline{h} \in \mathbb{R}^m$

$A\underline{x} = \underline{b} \in \mathbb{R}^p$

Step 1: Write the Lagrangian:

$$\underbrace{L(\underline{x}, \underline{\lambda}, \underline{\nu})}_{\text{is affine in primal variable } \underline{x}} = \langle \underline{c}, \underline{x} \rangle + \langle \underline{\lambda}, G\underline{x} - \underline{h} \rangle + \langle \underline{\nu}, A\underline{x} - \underline{b} \rangle$$

$$= (\underline{c}^T + \underline{\lambda}^T G + \underline{\nu}^T A)\,\underline{x} - \underline{\lambda}^T \underline{h} - \underline{\nu}^T \underline{b}$$

Step 2 : derive Lagrange dual function :

$$g(\underline{\lambda}, \underline{\nu}) := \inf_{\underline{x} \in \mathbb{R}^n} L(\underline{x}, \underline{\lambda}, \underline{\nu})$$

$$= \begin{cases} -\underline{\lambda}^T \underline{h} - \underline{\nu}^T \underline{b} & \text{if } \underline{c} + G^T \underline{\lambda} + A^T \underline{\nu} = \underline{0} \\ -\infty & \text{else} \end{cases}$$

Step 3 : derive the Lagrange dual problem

$$\sup_{\substack{\underline{\lambda} \in \mathbb{R}^m_{\geq 0} \\ \underline{\nu} \in \mathbb{R}^p}} g(\underline{\lambda}, \underline{\nu}) \iff \begin{aligned} &\min_{\begin{pmatrix} \underline{\lambda} \\ \underline{\nu} \end{pmatrix}} \begin{pmatrix} \underline{h} \\ \underline{b} \end{pmatrix}^T \begin{pmatrix} \underline{\lambda} \\ \underline{\nu} \end{pmatrix} \\ &\text{s.t. } \underline{\lambda} \succeq \underline{0}, \ \underline{c} + G^T \underline{\lambda} + A^T \underline{\nu} = \underline{0} \end{aligned}$$

A different LP

Example: Nonconvex but strong duality holds

This problem has a name: "Trust region problem"

$$\left.\begin{array}{ll} \min_{\underline{x} \in \mathbb{R}^n} & \underline{x}^T A \underline{x} + 2\underline{b}^T \underline{x} \\ \text{s.t.} & \underline{x}^T \underline{x} \le 1 \end{array}\right\} \quad A \in S^n$$

possibly sign indefinite

Step 1: Lagrangian

$$L(\underline{x}, \lambda) = \underline{x}^T A \underline{x} + 2\underline{b}^T \underline{x} + \lambda(\underline{x}^T \underline{x} - 1)$$

next pg.

Step 2 : The dual function.

$$g(\lambda) = \inf_{\underline{x} \in \mathbb{R}^n} L(\underline{x}, \lambda)$$

pseudo-inverse

Verify yourself :

$$g(\lambda) = \begin{cases} -\underline{b}^T (A+\lambda I)^{+} \underline{b} - \lambda & \text{if } A+\lambda I \succcurlyeq 0, \ \underline{b} \in \text{range}(A+\lambda I) \\ -\infty & \text{else} \end{cases}$$

Step 3: Dual problem:

$$\begin{cases} \max & -\underline{b}^T (A+\lambda I)^{\dagger}\, \underline{b} - \lambda \\ \text{s.t.} & A + \lambda I \succeq 0 \\ \text{and} & \underline{b} \in \text{range}(A+\lambda I) \end{cases}$$

Convex problem ← the text p. 229 checks this

(has to be from duality theory)

The text also shows: $\boxed{d^* = p^*}$ ← Strong duality holds!!

Application of duality:

(Algorithm to solve convex problem)

$$\underbrace{\underline{x}^{(k)}}_{\text{feasible primal sequence}} \qquad \underbrace{(\underline{\lambda}^{(k)}, \underline{\nu}^{(k)})}_{\text{feasible dual sequence}}$$

numerical tolerance

Stopping condition:

$$f_0(\underline{x}^{(k)}) - g(\underline{\lambda}^{(k)}, \underline{\nu}^{(k)}) \leq \text{numerical tolerance}$$

Complimentary slackness :

Suppose strong duality holds

$$\Longleftrightarrow \quad p^* = d^*$$

$$\Longleftrightarrow \quad f_0(\underline{x}^*) = \underbrace{g(\underline{\lambda}^*, \underline{\nu}^*)}_{\shortparallel}$$

$\underline{x}^*$ is primal optimizer

$(\underline{\lambda}^*, \underline{\nu}^*)$ are dual optimizers

$$= \inf_{\underline{x} \in \mathbb{R}^n} L(\underline{x}, \underline{\lambda}^*, \underline{\nu}^*)$$

$$\leq L(\underline{x}, \underline{\lambda}^*, \underline{\nu}^*)\Big|_{\underline{x} = \underline{x}^*}$$

$$\Leftrightarrow \cancel{f_0(\underline{x}^*)} \le \cancel{f_0(\underline{x}^*)} + \sum_{i=1}^{m} \lambda_i^* f_i(\underline{x}^*) + \boxed{\sum_{i=1}^{p} \nu_i^* \underbrace{h_i(\underline{x}^*)}_{\to 0}}$$

0 (because $\underline{x}^*$ is feasible)

$$\Leftrightarrow \boxed{0 \le \sum_{i=1}^{m} \lambda_i^* f_i(\underline{x}^*)}$$

On the other hand, $\lambda_i^* \ge 0$ and $f_i(\underline{x}^*) \le 0$

$$\boxed{\therefore \sum_{i=1}^{m} \lambda_i^* f_i(\underline{x}^*) \le 0}$$

$\therefore$ We must have :

$$\sum_{i=1}^{m} \underbrace{\lambda_i^*}_{\geq 0} \underbrace{f_i(\underline{x}^*)}_{\leq 0} = 0$$

$$\Longleftrightarrow \boxed{\lambda_i^* f_i(\underline{x}^*) = 0 \quad \forall\, i = 1, \ldots, m}$$

$$\therefore \left.\begin{aligned} \lambda_i^* > 0 &\Longleftrightarrow f_i(\underline{x}^*) = 0 \\ \lambda_i^* = 0 &\Longleftrightarrow f_i(\underline{x}^*) < 0 \end{aligned}\right\} \text{complimentary slackness}$$

Relation between Lagrange duality and Legendre-Fenchel conjugate duality

We know that the Legendre-Fenchel conjugate of $f(\cdot)$ is

$$\underbrace{f^*(\underline{y})}_{\text{convex in } \underline{y}} = \sup_{\underline{x} \in \mathrm{dom}(f)} \left\{ \langle \underline{y}, \underline{x} \rangle - \underbrace{f(\underline{x})}_{\text{possibly nonconvex in } \underline{x}} \right\}$$

Consider the primal problem:

may be complicated

$$\min_{\underline{x} \in \mathbb{R}^n} f_0(\underline{x})$$
$$\text{s.t.} \quad A\underline{x} \leq \underline{b}$$
$$C\underline{x} = \underline{d}$$

linear constraint

Can directly derive Lagrange dual function $g(\underline{\lambda}, \underline{\nu})$:

$$g(\underline{\lambda}, \underline{\nu}) = \inf_{\underline{x} \in \mathbb{R}^n} L(\underline{x}, \underline{\lambda}, \underline{\nu})$$
$$= \inf_{\underline{x} \in \mathbb{R}^n} \left\{ f_0(\underline{x}) + \underline{\lambda}^T (A\underline{x} - \underline{b}) + \underline{\nu}^T (C\underline{x} - \underline{d}) \right\}$$

$$= -\underline{b}^T\underline{\lambda} - \underline{d}^T\underline{\nu} + \inf_{\underline{x}\in\mathbb{R}^n}\left\{ f_0(\underline{x}) + \left(A^T\underline{\lambda} + C^T\underline{\nu}\right)^T\underline{x} \right\}$$

$$= -\underline{b}^T\underline{\lambda} - \underline{d}^T\underline{\nu} \underbrace{- f_0^*}\left(-A^T\underline{\lambda} - \underline{C}^T\underline{\nu}\right)$$

we are using

$$\inf_{\underline{x}} (\ldots) = -\sup_{\underline{x}}\{-(\ldots)\}$$

ditto

Moral: If we encounter nonlinear optimization problem with linear constraints, then we can directly write down Lagrange dual function & hence the Lagrange dual problem using Fenchel conjugates.

Example:

$$\max_{\underline{x} \in \Delta^{n-1}} f_0(\underline{x}) = -\sum_{i=1}^{n} x_i \log x_i$$

$$\text{s.t.} \quad A\underline{x} \preceq \underline{b} \in \mathbb{R}^m$$

(m linear inequality constraints)

But in HW 4, Prob. 2, we had:

$$f(u) = u \log u \Rightarrow f^*(v) = e^{v-1}$$

$$\therefore \; f_0^*(\underline{y}) = \sum_{i=1}^{n} e^{y_i - 1}, \quad \mathrm{dom}(f_0^*) = \mathbb{R}^n$$

Then,

$$g(\underline{\lambda}, \underline{\nu}) = -\underline{b}^T \underline{\lambda} - \nu - \sum_{i=1}^{n} \exp(-\underline{a}_i^T \underline{\lambda} - \underline{\nu} - 1)$$

$$= -\underline{b}^T \underline{\lambda} - \nu - \exp(-\nu - 1) \sum_{i=1}^{n} e^{-\underline{a}_i^T \underline{\lambda}}$$

where $\underline{a}_i$ is i^{th} column of matrix A

∴ Lagrange Dual problem :

$$d^* = \min_{\substack{\underline{\lambda} \in \mathbb{R}^m_{\geq 0} \\ \nu \in \mathbb{R}}} \left\{ \underline{b}^T \underline{\lambda} + \nu + \exp(-\nu - 1) \sum_{i=1}^{n} \exp(-\underline{a}_i^T \underline{\lambda}) \right\}$$

We can further simplify by minimizing over $\nu \in \mathbb{R}$ while holding $\underline{\lambda} \in \mathbb{R}^m_{\geq 0}$ fixed

This returns optimal ν :

$$\nu^* = \log\left(\sum_{i=1}^{n} \exp(-\underline{a}_i^T \underline{\lambda}) \right) - 1$$

next pg.

$$\therefore \; d^* = \min_{\underline{\lambda} \in \mathbb{R}^n_{\geq 0}} \left\{ \underline{b}^T \underline{\lambda} + \log\left(\sum_{i=1}^{n} \exp\left(-\underline{a}_i^T \underline{\lambda}\right)\right)\right\}$$

(sub. ν^* back)

This is a GP

End of example

See Lec. 13, p. 6-13.