Duality

$$\begin{aligned} &\text{minimize } f(x) \\ &\text{s.t. } x \in \mathcal{X} \\ &\quad h_i(x) = 0 \text{ for } i = 1, \dots, m \\ &\quad g_j(x) \le 0 \text{ for } j = 1, \dots, r \end{aligned}$$

$$L(x, \lambda, \mu) = f(x) + \underbrace{\sum_{i=1}^{m} \lambda_i h_i(x)}_{\lambda^T h(x)} + \underbrace{\sum_{j=1}^{r} \mu_j g_j(x)}_{\mu^T g(x)}$$

Dual Function

$$D(\lambda, \mu) = \min_{x \in \mathcal{X}} L(x, \lambda, \mu)$$

on convex set

$$\mathcal{C} = \{(\lambda, \mu) : \lambda \in \mathbb{R}^m, \mu_j \ge 0, j = 1, \dots, r\}$$

Result $D(\lambda, \mu)$ is concave on $\mathcal{C}$

Proof Let (λ, μ) and $(\tilde{\lambda}, \tilde{\mu}) \in \mathcal{C}$.

For $\alpha \in [0, 1]$,

$$D(\alpha\lambda + (1-\alpha)\tilde{\lambda}, \alpha\mu + (1-\alpha)\tilde{\mu})$$

$$= \min_{x \in \mathcal{X}} f(x) + (\alpha\lambda + (1-\alpha)\tilde{\lambda})^T h(x) + (\alpha\mu + (1-\alpha)\tilde{\mu})^T g(x)$$

$$= \min_{x \in \mathcal{X}} \alpha[f(x) + \lambda^T h(x) + \mu^T g(x)] + (1-\alpha)[f(x) + \tilde{\lambda}^T h(x) + \tilde{\mu}^T g(x)]$$

$$\ge \min_{x \in \mathcal{X}} \alpha[f(x) + \lambda^T h(x) + \mu^T g(x)] + \min_{x \in \mathcal{X}} (1-\alpha)[f(x) + \tilde{\lambda}^T h(x) + \tilde{\mu}^T g(x)]$$

Thus,

$$D(\alpha\lambda + (1-\alpha)\tilde{\lambda}, \alpha\mu + (1-\alpha)\tilde{\mu})$$
$$\geq \alpha D(\lambda, \mu) + (1-\alpha) D(\tilde{\lambda}, \tilde{\mu})$$ $\boxtimes$

<u>Weak Duality</u>

Define the feasibility set

$$\mathcal{F} = \{x : x \in \mathcal{S}, h(x) = 0, g(x) \leq 0\}$$

$$\max_{(\lambda,\mu) \in \mathcal{C}} D(\lambda, \mu) \leq \min_{x \in \mathcal{F}} f(x)$$

<u>Proof</u> For $(\lambda, \mu) \in \mathcal{C}$, $x \in \mathcal{F}$,

$$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x) \leq f(x)$$

$$\Rightarrow \min_{x \in \mathcal{F}} L(x, \lambda, \mu) \leq f(x) \text{ for all } x \in \mathcal{F}$$
$$\leq \min_{x \in \mathcal{F}} f(x) = f^*$$

Since $\mathcal{F} \subseteq \mathcal{S}$,

$$\min_{x \in \mathcal{S}} L(x, \lambda, \mu) \leq f^*$$

i.e. $D(\lambda, \mu) \leq f^* \quad \forall (\lambda, \mu) \in \mathcal{C}$

$$\Rightarrow \max_{(\lambda,\mu) \in \mathcal{C}} D(\lambda, \mu) \leq f^*.$$

<u>Strong Duality</u> Under some conditions, equality holds, i.e.,

$$\underbrace{\max_{(\lambda,\mu)\in \mathcal{C}} D(\lambda,\mu)}_{\text{dual problem}} = \underbrace{\min_{x\in \mathcal{F}} f(x)}_{\text{primal problem}}$$

<u>Result</u> Suppose f is convex, h_i's are affine, g_j's are convex, and $\mathcal{S} = \mathbb{R}^n$. If x^* is an optimal solution for primal problem, x^* is regular, and (λ^*, μ^*) are corresponding Lagrange multipliers, then strong duality holds and (λ^*, μ^*) maximize $D(\lambda,\mu)$.

<u>Proof</u> Under regularity assumption, using first-order KKT necessary conditions,

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0 \quad \text{— (1)}$$

and

$$\mu^{*T} g(x^*) = 0 \quad \text{— (2)}$$

Since f is convex, h_i's are affine, g_j's are convex, $L(x, \lambda^*, \mu^*)$ is convex in x. Thus, by (1),

$$L(x^*, \lambda^*, \mu^*) = \min_{x\in\mathbb{R}^n} L(x, \lambda^*, \mu^*)$$

$$\leq \max_{(\lambda,\mu)\in\mathcal{C}} \min_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \quad \text{— (3)}$$

Furthermore,

$$L(x^*, \lambda^*, \mu^*) = f(x^*) + \lambda^{*T} \underbrace{h(x^*)}_{=0} + \underbrace{\mu^{*T} g(x^*)}_{=0}$$

$$= f(x^*)$$

$$\geq f(x^*) + \lambda^T h(x^*) + \mu^T g(x^*)$$

for all $(\lambda, \mu) \in \mathcal{C}$

$$= L(x^*, \lambda, \mu), \quad \forall (\lambda, \mu) \in \mathcal{C}$$

$$\Rightarrow \quad L(x^*, \lambda^*, \mu^*) \geq \max_{(\lambda, \mu) \in \mathcal{C}} L(x^*, \lambda, \mu)$$

$$\geq \min_{x \in \mathbb{R}^n} \max_{(\lambda, \mu) \in \mathcal{C}} L(x, \lambda, \mu) \quad -(4)$$

From (3) and (4),

$$\min_{x \in \mathbb{R}^n} \max_{(\lambda, \mu) \in \mathcal{C}} L(x, \lambda, \mu) \leq L(x^*, \lambda^*, \mu^*) \leq \max_{(\lambda, \mu) \in \mathcal{C}} \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \quad -(5)$$

<u>Lemma</u> Consider function $g(y, z)$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$.

$$\max_{z \in \mathcal{Z}} \min_{y \in \mathcal{Y}} g(y, z) \leq \min_{y \in \mathcal{Y}} \max_{z \in \mathcal{Z}} g(y, z)$$

<u>Proof</u> HW 5

By Lemma,

$$\max_{(\lambda,\mu)\in\mathcal{C}} \min_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \leq \min_{x\in\mathbb{R}^n} \max_{(\lambda,\mu)\in\mathcal{C}} L(x,\lambda,\mu) \quad -(6)$$

From (5), (6) we get

$$\max_{(\lambda,\mu)\in\mathcal{C}} \min_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \min_{x\in\mathbb{R}^n} \max_{(\lambda,\mu)\in\mathcal{C}} L(x,\lambda,\mu) = L(x^*,\lambda^*,\mu^*)$$

$$\max_{(\lambda,\mu)\in\mathcal{C}} L(x,\lambda,\mu) = \max_{(\lambda,\mu)\in\mathcal{C}} f(x) + \lambda^T h(x) + \mu^T g(x)$$

$$= \begin{cases} \infty \text{ if } h(x)\neq 0 \text{ or } g(x)\not\leq 0, \text{ i.e. } x\notin\mathcal{F} \\ f(x) \text{ if } x\in\mathcal{F} \end{cases}$$

$$\Rightarrow \min_{x\in\mathbb{R}^n} \max_{(\lambda,\mu)\in\mathcal{C}} L(x,\lambda,\mu) = \min_{x\in\mathcal{F}} f(x)$$

Also $\max_{(\lambda,\mu)\in\mathcal{C}} \min_{x\in\mathbb{R}^n} L(x,\lambda,\mu) = \max_{(\lambda,\mu)\in\mathcal{C}} D(\lambda,\mu)$

Therefore $\max_{(\lambda,\mu)\in\mathcal{C}} D(\lambda,\mu) = \min_{x\in\mathcal{F}} f(x)$

Furthermore,

$$\max_{(\lambda,\mu)\in\mathcal{C}} D(\lambda,\mu) = L(x^*,\lambda^*,\mu^*) = D(\lambda^*,\mu^*).$$

i.e., (λ^*,μ^*) solve dual problem.

If the optimization problem is a linear program and it is feasible, Then strong duality holds (always).

Why?

Two ways to prove:

1) Simplex method
2) Farkas Lemma

Strong Duality result can be generalized to $\mathcal{S}$ being a convex subset of $\mathbb{R}^n$, but in this case, we need the existence of a point x in the relative interior of $\mathcal{S}$ s.t. $g_j(x) < 0 \ \forall j$ and $h_i(x) = 0 \ \forall i$.

This is called Slater's condition

Example

$$\text{minimize } x_1^2 + x_2^2 - 4x_1 - 2x_2 + 2$$

$$\text{s.t.} \quad x_1 + x_2 \le 2, \quad x_1 + 2x_3 \le 3$$

We showed that $x^* = \left(\frac{3}{2}, \frac{1}{2}\right)$ is the global min with $\mu^* = (1, 0)$.

$$f^* = f(x^*) = -\frac{5}{2}.$$

What is the dual of this convex program?

$$L(x, \mu) = x_1^2 + x_2^2 - 4x_1 - 2x_2 + 2 + \mu_1(x_1 + x_2 - 2) + \mu_2(x_1 + 2x_2 - 3).$$

$$= x_1^2 + x_2^2 + (\mu_1 + \mu_2 - 4)x_1 + (\mu_1 + 2\mu_2 - 2)x_2 + 2 - 2\mu_1 - 3\mu_2$$

$$D(\mu) = \min_{x \in \mathbb{R}^2} L(x, \mu) \leftarrow \text{convex}$$

min satisfies $\nabla_x L(x, \mu) = 0$

i.e.
$$2x_1 + (\mu_1 + \mu_2 - 4) = 0$$
$$2x_2 + (\mu_1 + 2\mu_2 - 2) = 0$$

i.e. $x_1 = \dfrac{-(\mu_1 + \mu_2 - 4)}{2}$, $x_2 = \dfrac{-(\mu_1 + 2\mu_2 - 2)}{2}$

$$D(\mu) = -\left(\frac{\mu_1+\mu_2-4}{2}\right)^2 - \left(\frac{\mu_1+2\mu_2-2}{2}\right)^2$$
$$+ 2 - 2\mu_1 - 3\mu_2$$

<u>Dual Problem</u>

$$\underset{\mu_1 \geq 0,\ \mu_2 \geq 0}{\text{maximize}} \quad D(\mu) = \underset{\mu_1 \geq 0,\ \mu_2 \geq 0}{\text{minimize}} \quad -D(\mu)$$

Note: $\nabla D(\mu) = \begin{bmatrix} -\mu_1 - \frac{3}{2}\mu_2 + 1 \\ -\frac{3}{2}\mu_1 - \frac{5}{2}\mu_2 + 1 \end{bmatrix}$

$$\nabla^2 D(\mu) = \begin{bmatrix} -1 & -\frac{3}{2} \\ -\frac{3}{2} & -\frac{5}{2} \end{bmatrix} \prec 0$$

$\Rightarrow$ $D(\mu)$ is strictly concave.

Can show that the maximizer of $D(\mu)$ under the constraint $\mu_1 \geq 0, \mu_2 \geq 0$ is

$$\mu_1^* = 1, \quad \mu_2^* = 0$$

$$D(\mu^*) = -\frac{9}{4} - \frac{1}{4} + 2 - 2 + 0$$
$$= \frac{-10}{4} = -\frac{5}{2} = f(x^*).$$