Notes 06: Strong duality

1. Geometric interpretation [BV §5.3.1]

Consider the graph

$$\mathcal{G} = \{(f_1(x), \ldots, f_m(x), h_1, \ldots, h_p(x), f_0(x)) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}\},$$

the set of values taken on by the constraint and objective functions. The primal optimum value can be expressed as

$$p^\star = \inf\{t \mid (u, v, t) \in \mathcal{G}, u \leqslant 0, v = 0\}.$$

The dual function at (λ, ν) amounts to minimizing the affine function

$$(\lambda, \nu, 1)^\top (u, v, t) = \sum_{1 \leqslant i \leqslant m} \lambda_i u_i + \sum_{1 \leq i \leqslant p} \nu_i v_i + t$$

over $(u, v, t) \in \mathcal{G}$, so $g(\lambda, \nu) = \inf\{(\lambda, \nu, 1)^\top (u, v, t) \mid (u, v, t) \in \mathcal{G}\}$.

If the infimum is finite, then the inequality

$$(\lambda, \nu, 1)^\top (u, v, t) \geqslant g(\lambda, \nu)$$

defines a supporting hyperplane to $\mathcal{G}$. This hyperplane is "nonvertical" as the last coordinate is nonzero.

Weak duality means $p^\star \geqslant g(\lambda, \nu)$ for any $\lambda \geqslant 0$.

Consider the epigraph

$$\mathcal{A} = \{(u, v, t) \mid \exists x \in \mathbb{R}^n,\ f_i(x) \leqslant u_i,\ 1 \leqslant i \leqslant m; h_i(x) = v_i,\ 1 \leqslant i \leqslant p, f_0(x) \leqslant t\}$$

of set of values taken on by the constraint and objective functions. If the optimization program is convex, then $\mathcal{A}$ is a convex set.

As before, primal optimum $p^\star = \inf\{t \mid (0, 0, t) \in \mathcal{A}\}$ and weak duality means

$$p^\star = (\lambda, \nu, 1)^\top (0, 0, p^\star) \geqslant g(\lambda, \nu) \text{ for any } \lambda \geqslant 0.$$

Strong duality holds if and only if equality holds in the last inequality for some dual feasible (λ, ν), i.e. there is a nonvertical supporting hyperplane to $\mathcal{A}$ at the boundary point $(0, 0, p^\star)$.

2. Slater's condition [BV §5.3.2]

Strong duality always holds for linear programs.

It fails in general for semidefinite programs and many other convex programs.

But typical convex programs that arises in practice can be shown to satisfy a sufficient condition for strong duality, known as Slater's condition.

Slater's condition says that a convex program is strictly feasible. More precisely, there is some point $\tilde{x}$ such that $f_i(\tilde{x}) < 0$ for $1 \leqslant i \leqslant m$ and $h_i(\tilde{x}) = 0$ for $1 \leqslant i \leqslant p$. We will express the equality constraints as $Ax = b$ where A is p-by-n matrix.

Theorem 2.1. *If a convex program satisfies slater's condition, strong duality holds.*

Proof. We will assume A has full row-rank (no equality constraint is redundant), for otherwise we can consider only a subset of rows of A that has the same row-rank as A.

We also assume primal optimal $p^\star$ is finite, for otherwise $p^\star = d^\star = -\infty$ by weak duality.

Since the program is convex, the "epigraph" $\mathcal{A}$ is a convex set. Define another convex set

$$\mathcal{B} = \{(0, 0, s) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid s < p^\star\}$$

that is disjoint from $\mathcal{A}$.

A version of separating hyperplane theorem (that we did not prove; see [BV §2.5.1]) can separate these two convex sets nonstrictly: there is $(\lambda, \nu, \mu) \neq 0$ and $\alpha \in \mathbb{R}$ such that

$$(\lambda, \nu, \mu)^\top (u, v, t) \geqslant \alpha \quad \forall (u, v, t) \in \mathcal{A}$$

and

$$(\lambda, \nu, \mu)^\top (u, v, t) \leqslant \alpha \quad \forall (u, v, t) \in \mathcal{B}.$$

The inequality for $\mathcal{A}$ implies $\lambda \geqslant 0$ and $\mu \geqslant 0$. The inequality for $\mathcal{B}$ simply says $\mu t \leqslant \alpha$ for all $t < p^\star$, hence $\mu p^\star \leqslant \alpha$. Together with the inequality for $\mathcal{A}$, we get for any $x \in \mathbb{R}^n$,

$$(\lambda, \nu, \mu)^\top (f(x), h(x), f_0(x)) \geqslant \alpha \geqslant \mu p^\star,$$

where $f(x) = (f_1(x), \ldots, f_m(x)) \in \mathbb{R}^m$ and $h(x) = (h_1(x), \ldots, h_p(x)) \in \mathbb{R}^p$.

If $\mu > 0$, then the last inequality divided by μ gives

$$L(x, \lambda/\mu, \nu/\mu) \geqslant p^\star.$$

Taking infimum over $x \in \mathbb{R}^n$, we get $g(\lambda/\mu, \nu/\mu) \geqslant p^\star$, and strong duality holds.

The remaining case is $\mu = 0$. We will rule out this case using Slater's condition. We have for any $x \in \mathbb{R}^n$,

$$(\lambda, \nu, \mu)^\top (f(x), h(x), f_0(x)) = (\lambda, \nu)^\top (f(x), h(x)) \geqslant 0.$$

In particular, for the strictly feasible point $\tilde{x}$,

$$\sum_{1 \leqslant i \leqslant m} \lambda_i f_i(\tilde{x}) \geqslant 0.$$

Since $f_i(\tilde{x}) < 0$ and $\lambda_i \geqslant 0$ for $1 \leqslant i \leqslant m$, in fact $\lambda = 0$.

Since $(\lambda, \nu, \mu) \neq 0$ and $\mu = 0$ and $\lambda = 0$, we must have $\nu \neq 0$. Earlier we showed that $\nu^\top h(x) = (\lambda, \nu, \mu)^\top (f(x), h(x) f_0(x)) \geqslant 0$ for all $x \in \mathbb{R}^n$. Since $\nu^\top h(x) = \nu^\top (Ax - b)$ and A has full row-rank, $\nu^\top A \neq 0$. Further, the strictly feasible point satisfies $\nu^\top (A\tilde{x} - b) = 0$. It is impossible for all points in a small ball (in $\mathbb{R}^n$) centered at $\tilde{x}$ to all lie on the same side of the hyperplane $\nu^\top (Ax - b) \geq 0$, when $\tilde{x}$ itself lies on the hyperplane. This rules out the case $\mu = 0$. □