In what follows P and D stand for the primal and the dual problems, respectively.

6.1 Weak duality

We have:

$$\begin{aligned} f(x) &\geq f(x) + u^T h(x) + v^T l(x) := L(x,u,v) \\ &\geq \min_x L(x,u,v) := g(u,v) \end{aligned} \tag{6.1}$$

and so:

$$f^* \geq g(u,v) \Longrightarrow f^* \geq g^*(u,v) \tag{6.2}$$

This is called the weak duality.

Note: The dual is always convex (even if P is not).

proof:

$$\begin{aligned} g(u,v) &= \min_x L(x,u,v) = \min_x \{f(x) + u^T h(x) + v^T l(x)\} \\ &= -\max_x \{f(x) + u^T h(x) + v^T l(x)\} \end{aligned} \tag{6.3}$$

This is a pointwise maximization of convex functions of u, v, and so it is convex in u, v.

6.2 Strong duality

$$f^* = g^* \tag{6.4}$$

When does this hold?

Slater's conditions: If P is convex and there exists at least one strictly feasible x, i.e., $h_i(x) < 0$, then we have strong duality.

An important extension: We only need this condition for the non-affine $h_i(x)$.

6.2.1 Strong duality in LP

In LP we have:

- Dual of dual is P
- Strong duality if P is feasible
- Strong duality if D is feasible
- The previous two points imply that we have strong duality unless both P and D are infeasible.

Example 6.1 (SVM)**.**

$$\begin{aligned} &\min_{\zeta,\beta,\beta_0} \frac{1}{2}||\beta||_2^2 + C\sum_i \zeta_i \\ &s.t\ \ \zeta_i \geq 0\,,\ y_i(x_i^T\beta + \beta_0) \geq 1 - \zeta_i \end{aligned} \tag{6.5}$$

The dual of this is:

$$\begin{aligned} &\max_w -\frac{1}{2}w^T\tilde{X}^T\tilde{X}w + 1^Tw \\ &s.t\ \ 0 \leq w \leq C1\ \ ,\ \ w^Ty = 0 \end{aligned} \tag{6.6}$$

where $\tilde{X} = diag(y)X$.

Clearly, $w = 0$ is dual feasible and so is primal feasible.

6.3 KKT conditions

1. Stationarity: $0 \in \partial(f(x) + u^Th(x) + v^Tl(x))$

For some pair (u, v), x minimizes the Lagrangian.

2. Complementary slackness: $u_ih_i(x) = 0\ \ ,\forall i$

3. Primal feasibility: $h_i(x) \leq 0\ ,\ \ l_i(x) = 0$

4. Dual feasibility: $u \geq 0$

Theorem 6.2 (Necessity)**.** *If x^* and (u^*, v^*) are optimal and $f^* = g^*$, then they satisfy KKT conditions.*

Proof: (i)

$$\begin{aligned} f(x^*) = g(u^*, v^*) &\leq \min_x f(x) + u^Th(x) + v^Tl(x) \\ &\leq f(x^*) + {u^*}^Th(x^*) + {v^*}^Tl(x^*) \\ &\leq f(x^*) \end{aligned} \tag{6.7}$$

Now replace $\leq$ with $=$. So from the second (in)equality we see that X^* is the minimizer of the Lagrangian. Also, from the last (in)equality we see that ${u^*}^Th(x^*) = 0$ and we have $u^* \geq 0$ and so we get the complementary slackness.

Theorem 6.3 (Sufficiency)**.** *If x^* and (u^*, v^*) satisfy KKT conditions then they are P and D optimal and $f^* = g^*$.*

Example 6.4 (SVM)**.** *KKT conditions:*

1. *Stationarity:* $w^T y = 0$, $\beta = w^T \tilde{X}$, $w = C1 - v$
2. *CS:* $v_i \zeta_i = 0$, $w_i(1 - \zeta_i - y_i(x_i^T \beta + \beta_0)) = 0$

Example 6.5 (constrained and Lagrangian forms)**.** *When are the two following forms equivalent?*

constrained form (C):

$$\begin{aligned} \min f(x) \\ s.t.\ h(x) \leq t \end{aligned} \tag{6.8}$$

Lagrangian form (L):

$$\min f(x) + \lambda h(x) \tag{6.9}$$

When C is strictly feasible, strong duality holds. So there exists λ such that for each x that solves C those x minimize L.

Now, if x^* solves L, then KKT condition for C hold by taking $t = h(x^*)$ and so x^* is a solution of C.

References