The dual LP (11/15)

Recall LP in standard form:

$$\begin{aligned} &\text{minimize} \quad f(x) = c^T x \\ &\text{subject to} \quad Ax = b \qquad (*) \\ &\qquad\qquad\quad\;\; x \geq 0 \end{aligned}$$

Let's write down the KKT conditions for this problem: Let $a_i \in \mathbb{R}^n$ be the i^{th} row of A. Then we can rewrite the constraints as:

$$g_i(x) = b_i - a_i^T x = 0, \quad i = 1, \dots, m,$$

$$g_{m+i}(x) = -x_i \leq 0, \quad i = 1, \dots, n.$$

So there are $m+n$ constraints total. The Lagrangian for $(*)$ is:

$$\mathcal{L}(x, \lambda, s) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{i=1}^{n} s_i g_{m+i}(x).$$

Recall the KKT conditions:

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x} = 0 & \text{[stationarity]} \\ g_i(x) = 0, \quad 1 \leq i \leq m & \\ g_i(x) \leq 0, \quad m < i \leq m+n & \text{[primal feasibility]} \\ s_i \geq 0, \quad 1 \leq i \leq n & \text{[dual feasibility]} \\ s_i g_i(x), \quad 1 \leq i \leq n & \text{[complementary slackness]} \end{cases}$$

Let's specialize this to $(*)$. First:

$$L(x,\lambda,s) = c^T x + \sum_{i=1}^{m} (b_i - a_i^T x)\lambda_i - \sum_{i=1}^{n} s_i x_i,$$

so that:

$$\frac{\partial L}{\partial x_j} = c_j - \sum_{i=1}^{m} (a_i^T e_j)\lambda_i - \sum_{i=1}^{n} s_i \delta_{i-j}$$

$$= c_j - \sum_{i=1}^{m} a_{ij}\lambda_i - s_j .$$

Hence:

$$\frac{\partial L}{\partial x} = \left(\frac{\partial L}{\partial x_1}, \ldots, \frac{\partial L}{\partial x_n} \right) = c - A^T\lambda - s.$$

So, the KKT conditions for $(*)$ with (x^*, λ^*, s^*) optimal are:

$A^T\lambda^* + s^* = c$ [stationarity]

$Ax^* = b$, $x^* \geq 0$ } [primal feasibility]

$s^* \geq 0$ [dual feasibility]

$x_i^* s_i^* = 0, \forall i$ [complementary slackness]

Using these conditions, we can derive sufficient conditions for optimality for $(*)$. How? First, observe the following:

$$c^T x^* = (s^* + A^T\lambda^*)^T x^* = s^{*T} x^* + \lambda^{*T} A x^* = \lambda^{*T} b.$$

In the last step, we used the fact that $s_i^* x_i^*$ $\forall i$, which implies $s^{*T} x^* = \sum_i s_i^* x_i^* = 0$.

Next, let x be another feasible point. Then:

$$c^T x = (s^* + A^T\lambda^*)^T x = s^{*T} x + \lambda^{*T} A x \geq \lambda^{*T} b = c^T x^*.$$

This follows since $s^* \geq 0$ and $x \geq 0$ together imply $s^{*T} x \geq 0$.

Note that equality is obtained above exactly when $x = x^*$, since then $s^{*T} x^* = 0$ (instead of $s^{*T} x \geq 0$). This essentially proves that x is optimal if and only if $s^{*T} x = 0$.

If we consider a fixed x, we could see how large we could make the value $b^T\lambda$, reasoning that since for optimal (x^*, λ^*, s^*) we have $\lambda^{*T} b = c^T x^*$, for suboptimal λ, we should have $c^T x^* \geq b^T \lambda$. If x is fixed, the remaining KKT conditions lead to the minimization problem:

$$\begin{aligned} &\text{minimize} \quad -b^T\lambda \\ &\text{subject to} \quad A^T\lambda + s = c \\ &\qquad\qquad\qquad s \geq 0 \end{aligned}$$

Let's define $s = c - A^T\lambda$ and rewrite this:

$$\begin{aligned} &\text{minimize} \quad -b^T\lambda \qquad\qquad (**) \\ &\text{subject to} \quad A^T\lambda \leq c. \end{aligned}$$

Let's denote the Lagrange multipliers for the inequality constraints suggestively by $x = (x_1, \ldots, x_n)$ so that the Lagrangian is:

$$L(\lambda, x) = -b^T\lambda + x^T(A^T\lambda - c).$$

Then, the KKT conditions are:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -b + Ax^* = 0 \quad \text{[stationarity]}$$

$$A^T \lambda^* \leq c \quad \text{[primal feasibility]}$$

$$x^* \geq 0 \quad \text{[dual feasibility]}$$

$$0 = (c - A^T\lambda^*)_i \, x_i^*, \quad 1 \leq i \leq n \quad \text{[complementary slackness]}$$

Using $s^* = c - A^T\lambda^*$, this can be rewritten:

$$Ax^* = b$$
$$A^T\lambda^* + s^* = c$$
$$s^* \geq 0$$
$$x^* \geq 0$$
$$s_i^* x_i^* = 0, \quad 1 \leq i \leq n.$$

But these are exactly the KKT conditions for (*)!

Hence, if (x^*, λ^*, s^*) solves the first set of KKT conditions, then (λ^*, x^*) must solve these, and vice versa!

<u>Note</u>: the LP (**) is the <u>dual</u> of (*). <u>The dual of (**) is (*)!</u>

You can check this...

If λ is feasible for (**) we call it <u>dual feasible</u>.

We can show that the KKT conditions for (**) also give sufficient conditions for optimality. If λ is some other dual feasible point, then:

$$b^T\lambda = x^{*T}A^T\lambda = x^{*T}A^T\lambda - c^Tx^* + c^Tx^*$$
$$= \underbrace{x^{*T}(A^T\lambda - c)}_{\leq 0} + c^Tx^* \leq c^Tx^* = b^T\lambda^*.$$

<u>Note</u>: should be clear by now that $c^Tx \geq b^T\lambda$ for primal feasible x and dual feasible λ. This is called <u>weak duality</u>. For LPs, we have <u>strong duality</u>, encapsulated in the following theorem:

<u>Theorem</u>: (Strong duality for LPs)

i) If (*) or (**) have a finite solution, so does the other, and $c^Tx^* = b^T\lambda^*$.

ii) If (*) or (**) is unbounded, then the other is infeasible.

<u>Proof</u>:

i) Follows almost immediately from the equivalence of the KKT conditions of (*) and (**) ⊠

ii) WLOG, assume $(*)$ is unbounded. Then, there exists a sequence $\{x_k\}$ such that each x_k is primal feasible, and:

$$\lim_{k \to \infty} c^T x_k = -\infty.$$

Assume that λ is dual feasible: i.e. $A^T \lambda \le c$. Clearly, $\lambda^T b$ is finite. Then, since $x_k \geq 0$, we have:

$$\lambda^T A x_k \le c^T x_k \quad \text{for all } k.$$

Hence:

$$\lambda^T b = \lambda^T A x_k \le c^T x_k \to -\infty,$$

which is a contradiction $\blacksquare$