11/10 More KKT ~~[illegible]~~

Recap: converting to standard form: two basic tools:

1) introduce slack variables

2) split unconstrained variables

Slack variables: if we have some inequality constraint $Ax \leq b$, introduce $s \geq 0$ and replace $Ax \leq b$ with $Ax + s = b$ and $s \geq 0$. Of course, now x in unconstrained.

Split unconstrained variables: if a variable is unconstrained, introduce two new variables, $x_+ \geq 0$ and $x_- \geq 0$ (called the positive part of x and the negative part of x) and write $x = x_+ - x_-$. Notice that if x is positive, $x_+ > 0$ and $x_- = 0$. Likewise, if x is negative, $x_- > 0$ and $x_+ = 0$. We can compute x_+ and x_- from x using:

$$x_+ = \max(0, x), \qquad x_- = \max(-x, 0).$$

Once we make this split, we effectively replace the variable x with the variables x_+ and x_-. So, in the previous example, we replace:

"$Ax + s = b, \quad s \geq 0, \quad x \in \mathbb{R}^n$"

with:

"$Ax + s = b, \quad s \geq 0, \quad x = x_+ - x_-, \quad x_+ \geq 0, \quad x_- \geq 0$"

To actually write this correctly in standard form, we need to eliminate "$x = x_+ - x_-$", giving:

"$A(x_+ - x_-) + s = b, \quad s \geq 0, \quad x_+ \geq 0, \quad x_- \geq 0$".

Now for more KKT theory. Problem of interest:

$$(*) \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) = 0, \quad i \in \mathcal{E} \\ & g_i(x) \leq 0, \quad i \in \mathcal{I}. \end{array} \Bigg\} \Rightarrow \mathcal{X} = \{x \in \mathbb{R}^n : g_i(x) = 0, \ i \in \mathcal{E}, \ g_i(x) \leq 0, \ i \in \mathcal{I}\}$$

<u>Def</u>: <u>active index set</u>: $\mathcal{A}(x) = \{i \in \mathcal{I} : g_i(x) = 0\} \cup \mathcal{E}$.

This is just the set of constraint indices for which the corresponding constraint $g_i(x)$ is zero.

We also need to define a slightly more technical condition:

<u>Def</u>: if x is feasible (i.e. $x \in \mathcal{X}$), and the set:

$$\{\nabla g_i(x)\}_{i \in \mathcal{A}(x)}$$

is linearly independent, then we say that the <u>linear independence constraint qualification</u> holds.

Theorem: Let f, g_i be $C^1(X)$ for all $i \in \mathcal{E} \cup \mathcal{I}$.

Let x^* be a constrained local minimum of $(*)$. Then if the linear independence constraint qualification holds at x^*, there exists a vector of Lagrange multipliers $\lambda^* \geq 0$ such that:

$$\nabla f(x^*) + \sum_{i \in \mathcal{E}} \lambda_i^* \nabla g_i(x^*) + \sum_{i \in \mathcal{I}} \lambda_i^* \nabla g_i(x^*) = 0 \quad \text{[stationarity]}$$

$$g_i(x^*) = 0, \ i \in \mathcal{E}, \qquad g_i(x^*) \leq 0, \ i \in \mathcal{I} \quad \text{[primal feasibility]}$$

$$\lambda_i^* \geq 0, \ i \in \mathcal{I} \quad \text{[dual feasibility]}$$

$$\lambda_i^* g_i^*(x^*) = 0, \ i \in \mathcal{E} \cup \mathcal{I} \quad \text{[complementary slackness]}.$$

Proof: too complex for this class. See Nocedal & Wright.

We'll now use the KKT conditions to derive optimality conditions for the standard form LP:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

Note: the KKT conditions are first-order necessary conditions. There are also second-order conditions for constrained problems, but we do not need them yet.

Let's introduce some Lagrange multipliers. We need a separate LM for each component of each inequality ~~of~~ or equality constraint. Write:

$$g_i(x) = (Ax-b)_i = \sum_{j=1}^{n} a_{ij} x_j - b_i \leq 0, \quad i=1,\dots,m$$

$$g_i(x) = -x_i, \quad i = m+1, \dots, m+n,$$

so that $\mathcal{I} = \{1,\dots,m\}$ and $\mathcal{E} = \{m+1,\dots,m+n\}$. We'll let $\lambda_1,\dots,\lambda_m$ be the LMs corresponding to the equality constraints $g_1(x) = 0, \dots, g_m(x) = 0$ and $s_1,\dots,s_n$ be the LMs corresponding to $g_{m+1}(x) \leq 0, \dots, g_{m+n}(x) \leq 0$.

The Lagrangian for this problem is:

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

The KKT conditions are now:

$$0 = \frac{\partial L}{\partial x} = c + A^T\lambda - s \quad \text{[stationarity]}$$

$$Ax=b, \ x \geq 0 \quad \text{[primal feasibility]}$$

$$s \geq 0 \quad \text{[dual feasibility]}$$

$$x_i s_i = 0 \ \forall i \quad \text{[complementary slackness]}$$

<u>Note</u>: since $x \geq 0$ and $s \geq 0$, "$x_i s_i = 0 \ \forall i$" is equivalent to $x^T s = 0$.

OK, let's assume ~~the~~ (x^*, λ^*, s^*) satisfy these KKT conditions. What can we conclude? Recall the optimal cost is $f(x^*) = c^T x^*$. Then:

$$c^T x^* = (s^* - A^T\lambda^*)^T x^*$$
$$= s^{*T} x^* - \lambda^{*T} A x^* = 0 - \lambda^{*T} A x^* = -\lambda^{*T} b.$$

The function $-\lambda^{*T} b$ is important... will come back to it. It's called the dual function, and this argument says the optimal primal and dual values are equal.

We can say more: the KKT conditions in this case are actually sufficient, and a feasible point x is optimal iff $x^T s^* = 0$. Why? Let x be primal feasible. Then:

$$c^T x = (s^* - A^T\lambda^*)^T x = \underbrace{s^{*T} x}_{s^* \geq 0,\ x \geq 0} - b^T \lambda^* \geq -b^T \lambda^* = c^T x^*.$$

Hence, $c^T x \geq c^T x^*$ for all x feasible. Furthermore, the inequality above becomes ~~an~~ an equality exactly when $s^{*T} x = 0$, or when complementary slackness holds. Hence, if (x^*, λ^*, s^*) satisfy the KKT conditions, x^* is optimum for the LP.