KKT OPTIMALITY CONDITIONS

CHRISTOPHER J. HILLAR

1. Complementary Slackness

Recall the primal and dual problems:

$$\begin{gathered} \text{minimize } f_0(x) \\ \text{subject to } f_i(x) \le 0 \\ h_i(x) = 0, \end{gathered}$$

$$\begin{gathered} \text{maximize } g(\lambda, \nu) \\ \text{subject to } \lambda \ge 0, \end{gathered}$$

where

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x),$$

and

$$g(\lambda, \nu) := \inf_{x \in \mathcal{D}} L(x, \lambda, \nu).$$

We assume now that the primal and dual optimality values are attained and equal (so that strong duality holds). Let x^* be a primal optimal and (λ^*, ν^*) be a dual optimal point. We have:

$$\begin{aligned} f_0(x^*) &= g(\lambda^*, \nu^*) \\ &= \min_{x \in \mathcal{D}} L(x, \lambda^*, \nu^*) \\ &\le L(x^*, \lambda^*, \nu^*) \\ &= f_0(x^*) + \sum \lambda^* f_i(x^*) + \sum v_i^* h_i(x^*) \\ &\le f_0(x^*). \end{aligned} \tag{1.1}$$

Thus, all the inequalities are actually equalities. In particular,

$$L(x^*, \lambda^*, \nu^*) = \min_{x \in \mathcal{D}} L(x, \lambda^*, \nu^*). \tag{1.2}$$

But more importantly, *complementary slackness* holds:

$$\sum \lambda_i^* f_i(x^*) = 0 \quad \Longrightarrow \quad \lambda_i^* f_i(x^*) = 0, \quad i = 1, \ldots, m.$$

An interesting interpretation of all this is that (x^*, λ^*, ν^*) is a *saddle point* for the Lagrangian $L(x, \lambda, \nu)$ (for $\lambda \geq 0$):

$$
\begin{aligned}
L(x^*, \lambda^*, \nu^*) &= f_0(x^*) + \sum \lambda_i^* f_i(x^*) + \sum \nu_i^* h_i(x^*) \\
&\geq f_0(x^*) + \sum \lambda_i f_i(x^*) + \sum \nu_i h_i(x^*) \\
&= L(x^*, \lambda, \nu).
\end{aligned} \tag{1.3}
$$

while $L(x^*, \lambda^*, \nu^*) \leq L(x, \lambda^*, \nu^*)$ by (1.2).

2. KKT

We now assume additionally that f_i and h_i are differentiable (but general otherwise). By (1.2), x^* minimizes $L(x, \lambda^*, \nu^*)$ over x. Thus, its gradient must vanish at x^*:

$$\nabla f_0(x^*) + \sum \lambda_i^* \nabla f_i(x^*) + \sum \nu_i^* \nabla h_i(x^*) = 0.$$

Thus, we have:

$$
\begin{aligned}
f_i(x^*) &\leq 0 \\
h_i(x^*) &= 0 \\
\lambda_i^* &\geq 0 \\
\lambda_i^* f_i(x^*) &= 0 \\
\nabla f_0(x^*) + \sum \lambda_i^* \nabla f_i(x^*) + \sum \nu_i^* \nabla h_i(x^*) &= 0.
\end{aligned} \tag{2.1}
$$

This system is called the *Karush-Kuhn-Tucker (KKT)* conditions.

References

[1] S. Boyd, L. Vandenberghe, *Convex Optimization*.