The Karush–Kuhn–Tucker (KKT) conditions

1 Introduction

When refering to the KKT conditions it is important to notice that the course book by Nash and Sofer does not state them in their original form in Theorem 14.3, but the necessary conditions that can be stated if the functions involved are *twice* continuously differentiable. When comparing the KKT conditions with the optimality conditions in linear programming on the one hand, and the unconstrained case on the other, we find that it is wiser to state the KKT conditions in their original form, which only assumes that the functions are once continuously differentiable.

2 The unconstrained case

2.1 Case of C^1 functions

Let f be in C^1. Then:

(necessary local optimality conditions):

x^* is a local minimum of f over $\Re^n \Longrightarrow \nabla f(x^*) = 0^n$.

2.2 Case of C^2 functions

Let f be in C^2. Then:

(necessary local optimality conditions):

x^* is a local minimum of f over $\Re^n \Longrightarrow \{\ \nabla f(x^*) = 0^n$ *and* $p^{\mathrm{T}} \nabla^2 f(x^*) p \geq 0$ for all $p \in \Re^n\ \}$.

3 The constrained case

3.1 Case of C^1 functions

Consider the problem to

$$\begin{aligned} \text{minimize } & f(x), \\ \text{subject to } & g(x) \geq 0^m, \\ & h(x) = 0^\ell, \end{aligned}$$

where $f : \Re^n \mapsto \Re$, $g : \Re^n \mapsto \Re^m$, $h : \Re^n \mapsto \Re^\ell$ are all in C^1.

(necessary local optimality conditions):

$\{$ x^* is a local minimum of f over $X := \{ x \in \Re^n \mid g(x) \geq 0^m,\ h(x) = 0^\ell \}$ *and* x^* is regular $\}$ $\Longrightarrow$

1. $x^* \in X$ (primal feasibility), and there exist vectors $\lambda^* \in \Re^m$ and $\mu^* \in \Re^\ell$ such that
2. $(\lambda^*)^{\mathrm{T}} g(x^*) = 0$ (complementarity),
3. $\nabla_x L(x^*, \lambda^*, \mu^*) := \nabla f(x^*) - (\lambda^*)^{\mathrm{T}} \nabla g(x^*) - (\mu^*)^{\mathrm{T}} \nabla h(x^*) = 0^n$ (dual feasibility), and
4. $\lambda^* \geq 0^m$ (dual feasibility).

3.2 Case of C^2 functions

Let $f : \Re^n \mapsto \Re$, $g : \Re^n \mapsto \Re^m$, $h : \Re^n \mapsto \Re^\ell$ all be in C^2.

(necessary local optimality conditions):

$\{$ x^* is a local minimum of f over X *and* x^* is regular $\}$ $\Longrightarrow$

1. $x^* \in X$ (primal feasibility), and there exist vectors $\lambda^* \in \Re^m$ and $\mu^* \in \Re^\ell$ such that
2. $(\lambda^*)^{\mathrm{T}} g(x^*) = 0$ (complementarity),
3. $\nabla_x L(x^*, \lambda^*, \mu^*) = 0^n$ (dual feasibility),
4. $\lambda^* \geq 0^m$ (dual feasibility), *and*

$p^{\mathrm{T}} \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) p \geq 0$ for all $p \in \{ p \neq 0^n \mid \nabla g_i(x^*)^{\mathrm{T}} p \geq 0,\ i \in \mathcal{I}(x^*); \quad \nabla h_i(x^*)^{\mathrm{T}} p = 0,\ i = 1, \dots, \ell \}$, where $\mathcal{I}(x^*) := \{ i = 1, \dots, m \mid g_i(x^*) = 0 \}$.

4 The LP case

Consider the problem to

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

Let $g(x) = \begin{pmatrix} Ax - b \\ x \end{pmatrix} \in \Re^{m+n}$, and introduce the multiplier vector $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \in \Re^{m+n}$. With $L(x, \lambda) := c^{\mathrm{T}}x - \lambda_1^{\mathrm{T}}(Ax - b) - \lambda_2^{\mathrm{T}}x$, we obtain the KKT conditions that

1. $Ax^* \geq b$, $x^* \geq 0^n$ (primal feasibility),
2. $(\lambda_1^*)^{\mathrm{T}}(Ax^* - b) = 0$, $\lambda_2^{\mathrm{T}}x^* = 0$ (complementarity),
3. $c - A^{\mathrm{T}}\lambda_1^* - \lambda_2^* = 0^n$ (dual feasibility), and
4. $\lambda^* \geq 0^{m+n}$ (dual feasibility).

Eliminate $\lambda_2^* = c - A^{\mathrm{T}}\lambda_1^*$. Then, we obtain that $(\pi^* := \lambda_1^*)$

1. $Ax^* \geq b$, $x^* \geq 0^n$ (primal feasibility),
2. $(\pi^*)^{\mathrm{T}}(Ax^* - b) = 0$, $(x^*)^{\mathrm{T}}(c - A^{\mathrm{T}}\pi^*) = 0$ (complementarity),
3. $A^{\mathrm{T}}\pi^* \leq c$ (dual feasibility), and
4. $\pi^* \geq 0^m$ (dual feasibility).

This is the optimality conditions for LP. Notice that since LP is a problem with linear constraints, any feasible solution is regular, so the KKT conditions are then guaranteed to be necessary, and LP is also a convex problem, so KKT is at the same time sufficient for global optimality.