FOUNDATIONS OF OPTIMIZATION: IE6001 Optimality Conditions

Napat Rujeerapaiboon

Semester I, AY2022/2023

Complementary Slackness

Assume that strong duality holds and that x^* is optimal in P, while (λ^*, μ^*) is optimal in D. Then,

$$f_0(x^*) = g(\lambda^*, \mu^*)$$
 (strong duality)
$$= \inf_{x} f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \mu_i^* h_i(x)$$
 (definition of g)
$$\leq f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{p} \mu_i^* h_i(x^*)$$

$$\leq f_0(x^*)$$
 ($\lambda^* \geq 0, f(x^*) \leq 0, h(x^*) = 0$)

and thus all inequalities hold as equalities. We conclude:

- x^* minimizes $L(x, \lambda^*, \mu^*)$
- $\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0$, which implies complementary slackness, i.e.,

$$\lambda_i^{\star} > 0 \implies f_i(x^{\star}) = 0$$
 and $f_i(x^{\star}) < 0 \implies \lambda_i^{\star} = 0$

Karush-Kuhn-Tucker (KKT) Conditions

Theorem: Assume that $f_0, \ldots, f_m, h_1, \ldots, h_p$ are differentiable, x^* is optimal in P, (λ^*, μ^*) is optimal in D and strong duality holds (but P may be nonconvex). Then, (x^*, λ^*, μ^*) satisfies:

$$f(x^{\star}) \leq 0, \ h(x^{\star}) = 0$$
 : primal feasibility $\lambda^{\star} \geq 0$: dual feasibility

$$\lambda^{\star} \geq 0$$
 : dual feasibility

$$\lambda_i^* f_i(x^*) = 0 \quad \forall i = 1, \dots, m$$
 : complementary slackness

$$\lambda_i^{\star} f_i(x^{\star}) = 0 \quad \forall i = 1, \dots, m$$
 : complement $\nabla f_0(x^{\star}) + \sum_{i=1}^m \lambda_i^{\star} \nabla f_i(x^{\star}) + \sum_{i=1}^p \mu_i^{\star} \nabla h_i(x^{\star}) = 0$: stationarity

Proof: Primal and dual feasibility follow immediately from optimality. Complementary slackness was derived on the previous slide. Stationary holds because x^* minimizes $L(x, \lambda^*, \mu^*)$.

KKT Conditions for Convex Problems

Theorem (Sufficiency): If P is convex and (x^*, λ^*, μ^*) satisfies the KKT conditions, then, x^* solves P and (λ^*, μ^*) solves D.

Proof: The KKT conditions imply

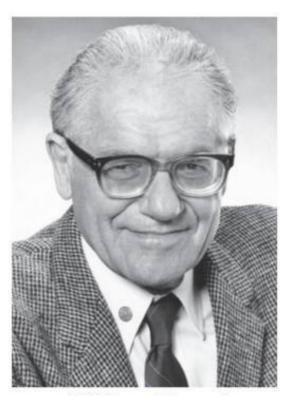
$$f_0(x^*) = L(x^*, \lambda^*, \mu^*)$$
 (complementary slackness)
= $\inf_x L(x, \lambda^*, \mu^*)$ (stationarity and convexity of L in x)
= $g(\lambda^*, \mu^*)$.

Thus, x^* and (λ^*, μ^*) are primal and dual feasible with the same objective value. Thus, they are both optimal.

Theorem (Necessity): Assume that a convex P satisfies Slater's condition. If x^* solves P, there is (λ^*, μ^*) such that (x^*, λ^*, μ^*) satisfies the KKT conditions.

Proof: Slater implies strong duality and solvability of D. The claim then follows from the KKT theorem (see previous slide).

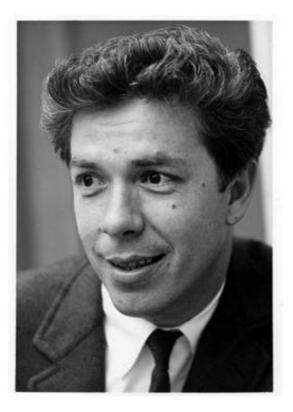
Who invented it?







Harold W. Kuhn



Albert W. Tucker

The KKT conditions were first named after Kuhn and Tucker, two famous professors from Princeton, who published them in 1951.

Later it was discovered that the conditions had been stated in the MSc thesis of the student William Karush in 1939.

Relevance of the KKT Conditions

The KKT conditions play an important role in optimization.

- Many algorithms for convex optimization are conceived as methods for solving the KKT conditions.
- Sometimes it is possible to solve the KKT conditions (and thus, the optimization problem) analytically.

Example: minimize
$$\frac{1}{2}x^{T}Px + q^{T}x + r$$
 subject to $Ax = b$ $(P \in \mathbb{S}_{+}^{n})$

KKT conditions:
$$Ax^* = b$$
, $Px^* + q + A^T \mu^* = 0$

$$\iff \begin{pmatrix} P & A^{\top} \\ A & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x}^{\star} \\ \mu^{\star} \end{pmatrix} = \begin{pmatrix} -q \\ b \end{pmatrix}$$

Separable Problems

Problem with a separable objective and a single constraint:

minimize
$$f_0(x) = \sum_{i=1}^n f_i(x_i)$$

subject to $a^\top x = b$

The Lagrangian is also separable:

$$L(x,\mu) = \sum_{i=1}^{n} f_i(x_i) + \mu(b - a^{\top}x) = b\mu + \sum_{i=1}^{n} (f_i(x_i) - \mu a_i x_i)$$

The dual objective is also separable:

$$g(\mu) = b\mu + \inf_{x} \sum_{i=1}^{n} (f_{i}(x_{i}) - \mu a_{i}x_{i})$$

$$= b\mu + \sum_{i=1}^{n} \inf_{x_{i}} (f_{i}(x_{i}) - \mu a_{i}x_{i})$$

$$= b\mu - \sum_{i=1}^{n} f_{i}^{*}(\mu a_{i})$$

D is a scalar optimization problem and thus simple. Given μ^* , we solve the problems $\inf_{x_i} (f_i(x_i) - \mu^* a_i x_i)$ separately to find x_i^* .

Main Take-Away Points

- KKT conditions: primal feasibility; dual feasibility; complementary slackness; stationarity
- **KKT theorems:** if Slater's condition holds, then (x^*, λ^*, μ^*) satisfies the KKT conditions iff x^* and (λ^*, μ^*) are primal and dual optimal, respectively