Constraints (equalities)

$$f(\mathbf{x}) \to \text{min, s.t. } g_1(\mathbf{x}) = 0, \ldots, g_m(\mathbf{x}) = 0$$

All functions are continuously differentiable.

A necessary condition for $\mathbf{x}^*$ to be a local extremum is given, if there exists multipliers $\lambda_1, \ldots, \lambda_{m+1}$ with at least one $\lambda_i \neq 0$ for $i = 1, \ldots, m+1$, such that:

$$\lambda_1 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_{i+1} \nabla g_i(\mathbf{x}^*) = \mathbf{0}$$

.

The Langrange multipliers λ_i are named after Lagrange (1736-1813), who discovered this theorem, but could not prove it. It took 100 years before the proof was found.

Show that this yields m+n equations with m+n+1 unknowns.

A standard approach is to try $\lambda_1 = 0$ and $\lambda_1 = 1$ (Lagrange multiplier rule).

Constraints (inequalities)

$f(\mathbf{x}) \to \min$, s.t. $g_1(\mathbf{x}) \leq 0, \ldots, g_m(\mathbf{x}) \leq 0$, all functions are continuously differentiable.

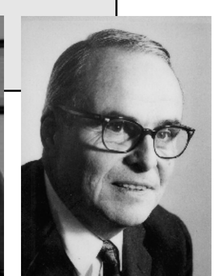
The Karush Kuhn Tucker conditions are said to hold for $\mathbf{x}^*$, if there exists multipliers $\lambda_1 \geq 0$, ..., $\lambda_{m+1} \geq 0$ and at least one $\lambda_i > 0$ for $i = 1, \ldots, m+1$, such that:

(1) $\lambda_1 \nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_{i+1} \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.
(2) $\lambda_{i+1} g_i(\mathbf{x}^*) = 0, i = 1, \ldots, m$

KKT Theorem - Neccessary conditions for smooth, convex programming: Assume the objective and all constraint functions are convex in some ϵ-neighborhood of $\mathbf{x}^*$, if $\mathbf{x}^*$ is a local minimum, then there exists $\lambda_1, \ldots, \lambda_{m+1}$ such that KKT conditions are fulfilled.

Harold W. Kuhn
US-American
Mathematician
1924-2014

Albert William Tucker
Canadian
Mathematician,
1905-1995



Constraint (inequality)

As in the case of Lagrange multiplier, we get $m+n$ non-linear equations, the solution of which results in candidate solutions.

The KKT conditions are sufficient for optimiality, provided $\lambda_1 = 1$. In this case $\mathbf{x}^*$ is a local minimum.

Note that if $\mathbf{x}^*$ is in the interior of the feasible region (a Slater point), all $g_i(\mathbf{x}) < 0$ and thus $\lambda_1 > 0$.

[Brinkhuis, Tikhomirov, 2005]

Multiobjective Optimization [cf. Miettinnen '99]

Fritz John neccessary conditions

A neccessary condition for $\mathbf{x}^*$ to be a locally efficient point is that there exists vectors $\lambda_1, \ldots, \lambda_k$ and $\upsilon_1, \ldots, \upsilon_m$ such that

(0) $\lambda \succ \mathbf{0}, \upsilon \succ \mathbf{0}$
(1) $\sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) - \sum_{i=1}^{m} \upsilon_i \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.
(2) $\upsilon_i g_i(\mathbf{x}^*) = 0, i = 1, \ldots, m$

Karush Kuhn Tucker sufficient conditions for a solution to be Pareto optimal: Let $\mathbf{x}^*$ be a feasible point. Assume that all objective functions are locally convex and all constraint functions are locally concave, and the Fritz John conditions hold in $\mathbf{x}^*$, then $\mathbf{x}^*$ is a local efficient point.

Unconstrained Multiobjective Optimization

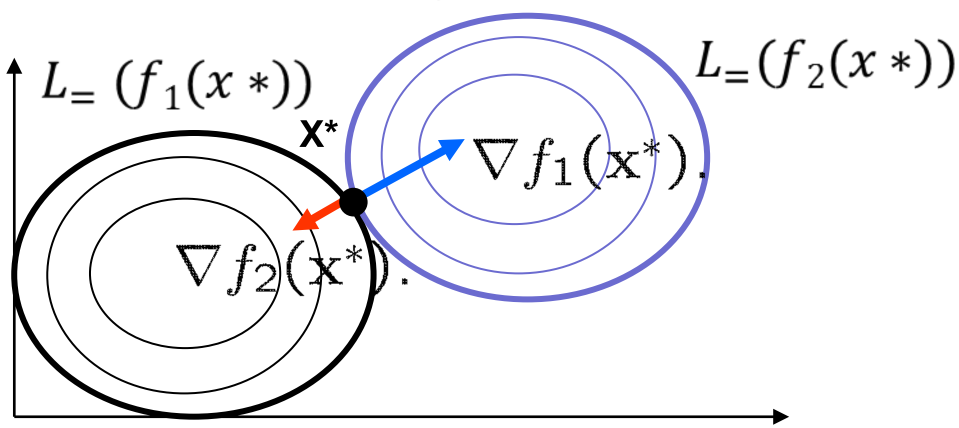
In the unconstrained case Fritz John neccessary conditions reduce to

There exist numbers $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, such that

(1) $\lambda \succ \mathbf{0}$

(2) $\sum_{i=1}^{k} \lambda_i \nabla f_i(\mathbf{x}^*) = \mathbf{0}$.

$\mathbf{x}^*$ is optimum for some linear scalarization with some weights $\lambda_1, \ldots, \lambda_k$.



In 2-dimensional spaces this criterion reduces to the observation, that either one of the objectives has a zero gradient (neccesary condition for ideal points) or the gradients are parallel.