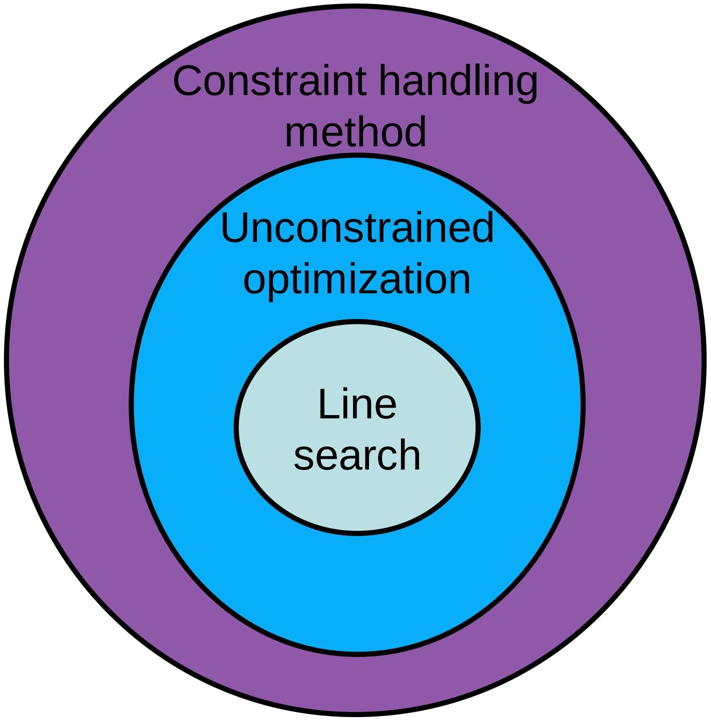
optimality conditions

Slides by
Jussi Hakanen
Post-doctoral researcher
jussi.hakanen@jyu.fi
Presented by
Mohammad Tabatabaei
Post-doctoral researcher
mohammad.tabatabaei@jyu.fi

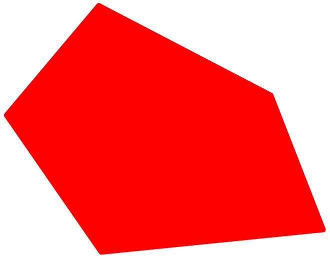
Structure of optimization methods

- Typically
  - Constraint handling converts the problem to (a series of) unconstrained problems
  - In unconstrained optimization a search direction is determined at each iteration
  - The best solution in the search direction is found with line search

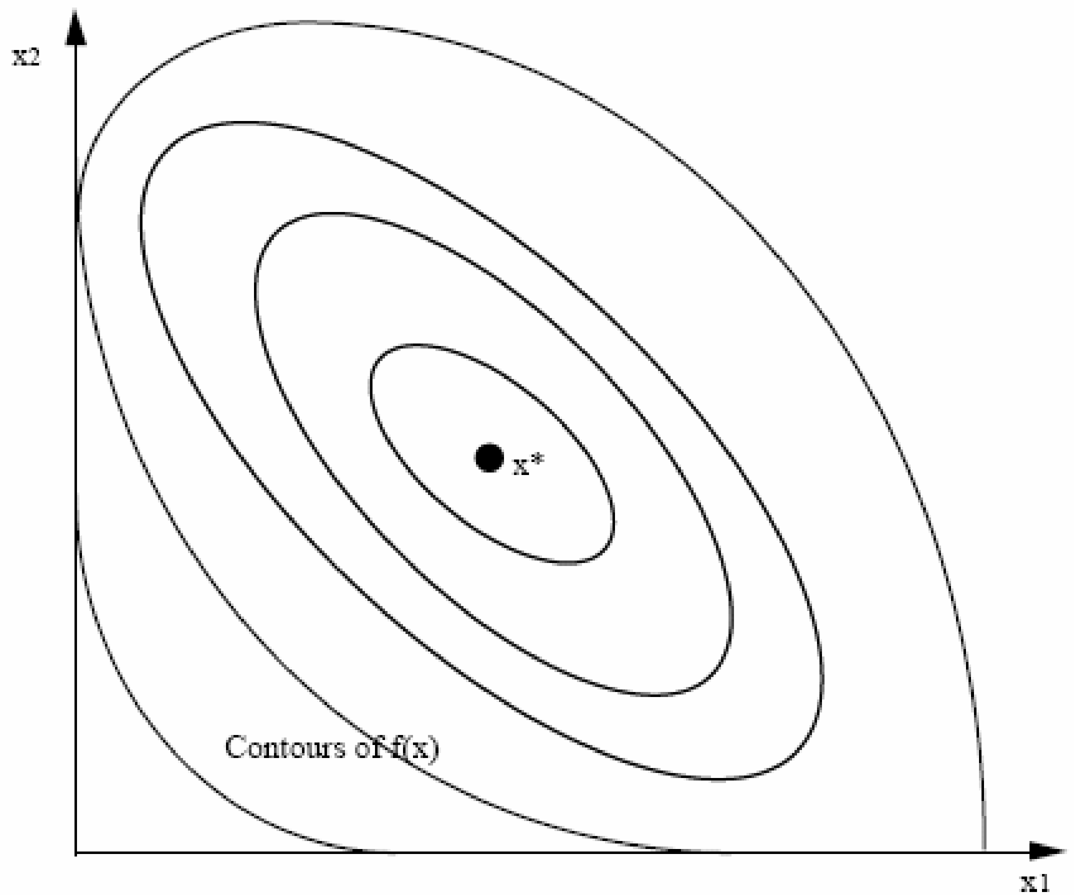


Constrained problem

- $\min f(x),\ s.t.\ x \in S \subset R^n$
- S is the *feasible region*, consists of all the points that satisfy the constraints
- Constraints can be e.g.
  - Box constraints: $x_i^l \leq x_i \leq x_i^u,\ i = 1, \ldots, n$
  - Inequality constraints: $g_j(x) \leq 0,\ j = 1, \ldots, m$
  - Equality constraints: $h_j(x) = 0,\ j = 1, \ldots, l$
- Special case: linear constraints
  - $Ax \leq b$ (i.e., $g(x) = Ax - b$)
  - $Ax = b$ (i.e., $h(x) = Ax - b$)

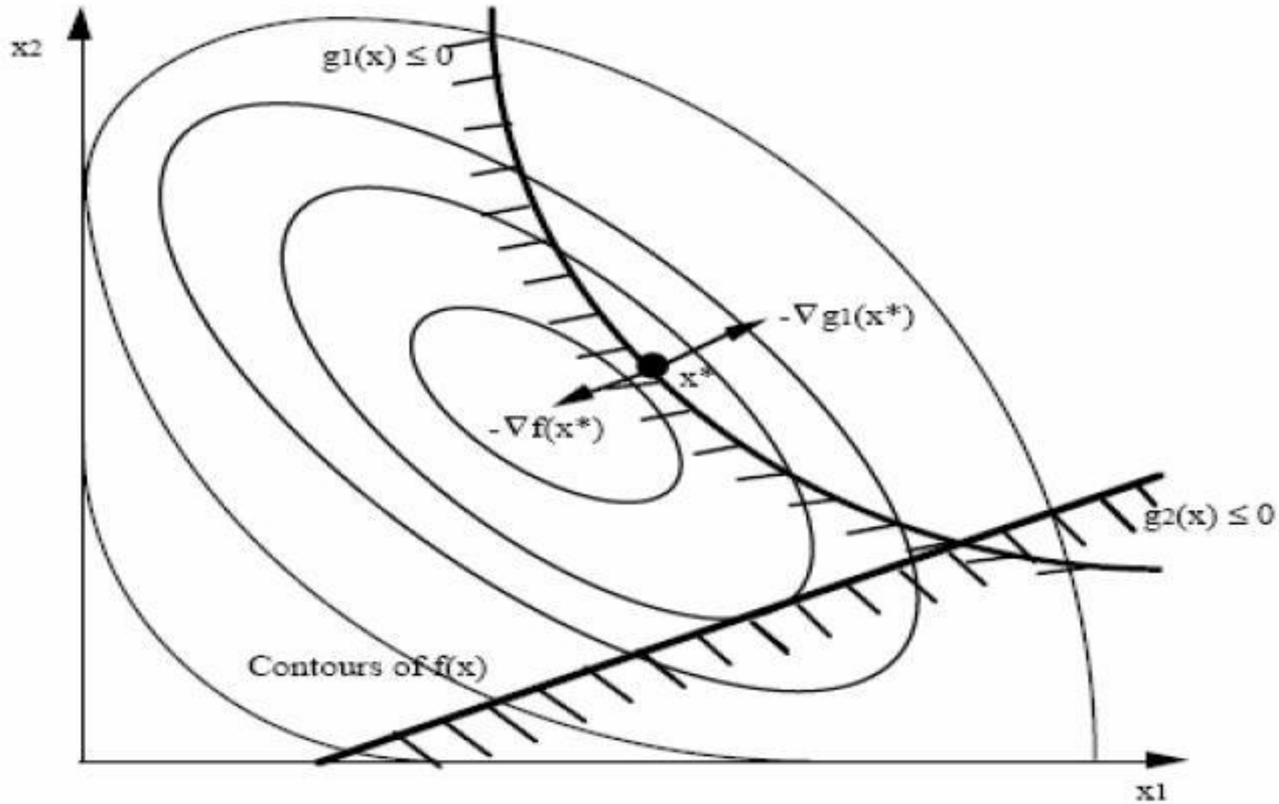


Unconstrained problem



Adopted from Prof. L.T. Biegler (Carnegie Mellon University)

Inequality constraints



Min $f(x)$

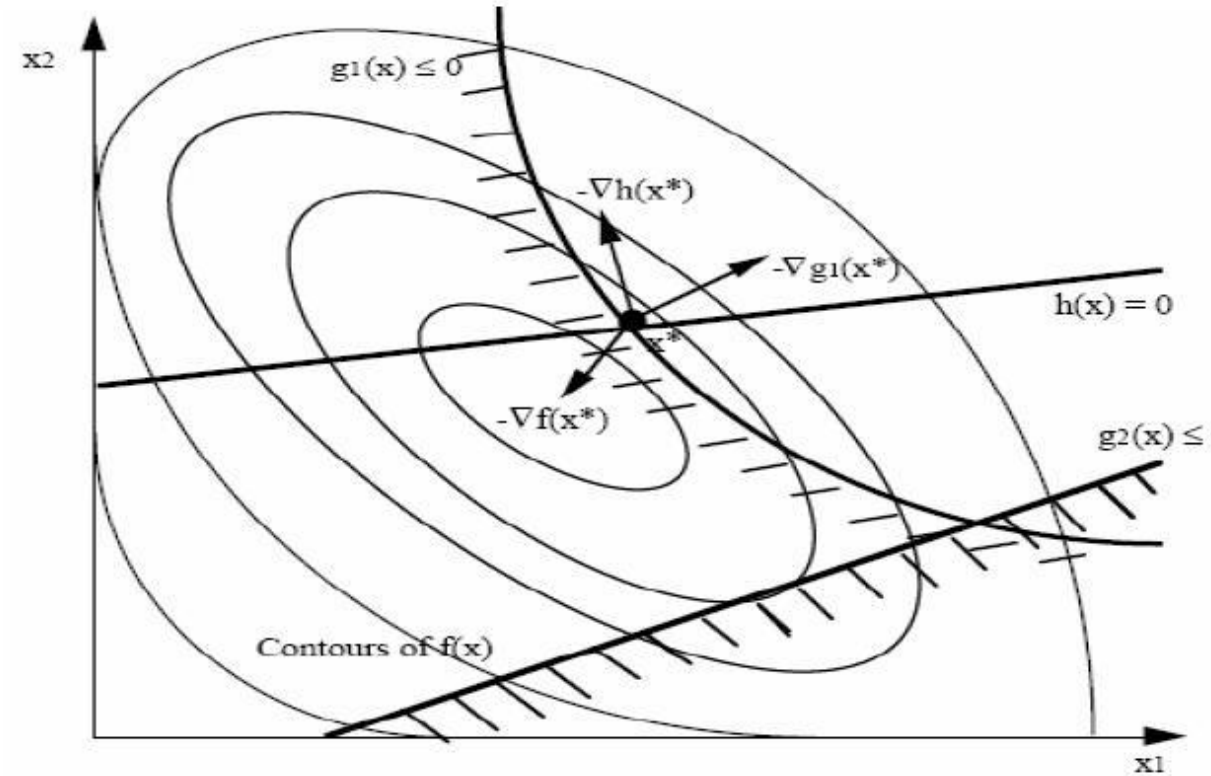
s.t. $g(x) \leq 0$

Analogy: Ball rolling down valley pinned by fence

Note: Balance of forces (∇f, ∇g_1)

Adopted from Prof. L.T. Biegler (Carnegie Mellon University)

Inequality and equality constraints



Problem: Min $f(x)$

s.t. $g(x) \le 0$

$h(x) = 0$

<u>Analogy</u>: Ball rolling on rail pinned by fences

Balance of forces: ∇f, ∇g_1, ∇h

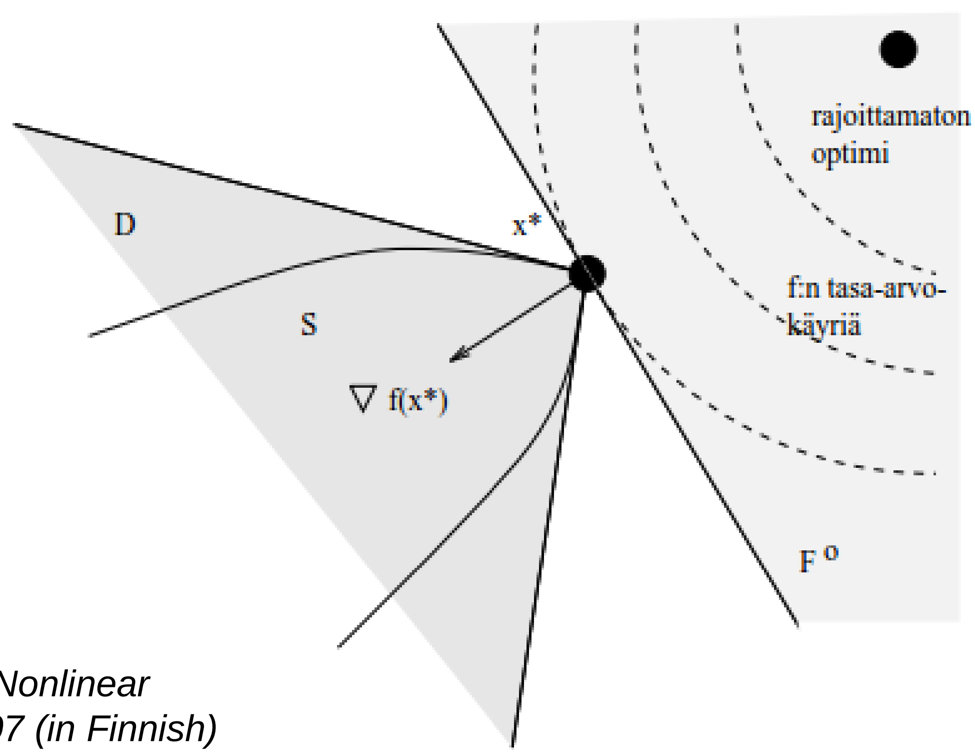
Adopted from Prof. L.T. Biegler (Carnegie Mellon University)

Feasible descent directions

- **Definition**: Let $S \subset R^n, S \neq \emptyset$ and $x^* \in cl\, S$. Then
$$D = \{d \in R^n \mid d \neq 0 \;\&\; x^* + \alpha d \in S \;\forall \alpha \in (0, \delta)\}$$
for some $\delta > 0$ is the cone of feasible directions of S in x^*.
  - Each $d \in D, d \neq 0$ is a feasible direction
- **Definition**: $F = \{d \in$

A necessary condition

- Let us consider $\min f(x),\ s.t.\ x \in S \subset R^n$ where $S \neq \emptyset$
- **Theorem:** Let f be differentiable in $x^* \in S$. If x^* is a local minimizer, then $F^\circ \cap D = \emptyset$ where $F^\circ = \{d \in$



From *Miettinen: Nonlinear optimization, 2007 (in Finnish)*

Problem with only inequality constraints

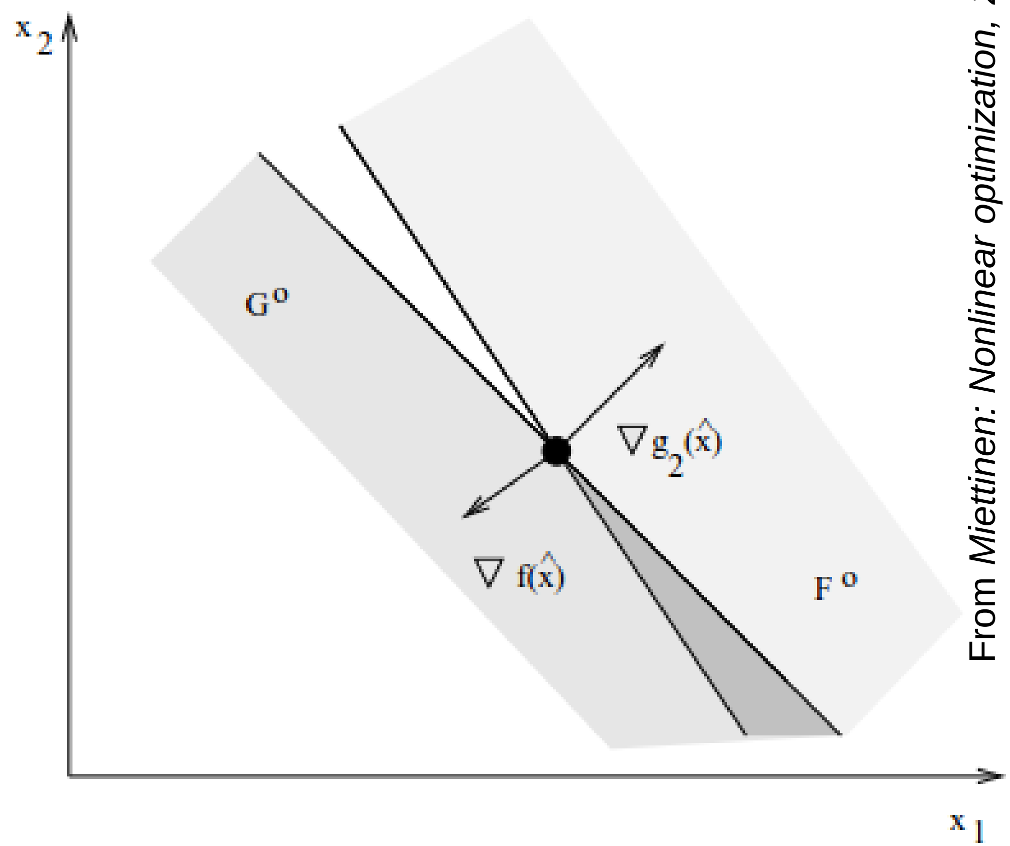
- Let us consider
  $\min f(x),\ s.t.\ g_i(x) \leq 0, i = 1, \dots, m$
- $S = \{x \in R^n \,|g_i(x) \leq 0\ \forall\ i = 1, \dots, m\}$ where $g_i : R^n \to R\ \forall i$
- Let I denote the set of active constraints in x^* i.e. $I = \{i\ |g_i(x^*) = 0\}$.

A necessary condition (for inequality constraints only)

- Define $G^{\circ} = \{d \in R^n \,|\nabla g_i(x^*)^T d < 0 \;\forall i \in I\}$
- **Theorem:** Let f and g_i be continuously differentiable in $x^* \in S$. If x^* is a local minimizer, then $F^{\circ} \cap G^{\circ} = \emptyset$.

Example

- $\min f(x) = (x_1 - 3)^2 + (x_2 - 2)^2$
- $s.t. \; x_1^2 + x_2^2 \leq 5, \; x_1 + x_2 \leq 3, \; x_1, x_2 \geq 0$
- $\hat{x} = \left(\frac{9}{5}, \frac{6}{5}\right)^T, I = \{2\}$
- $F^\circ \cap G^\circ \neq \emptyset$
- $\hat{x}$ is not a local minimizer!



From Miettinen: Nonlinear optimization, 2007 (in Finnish)

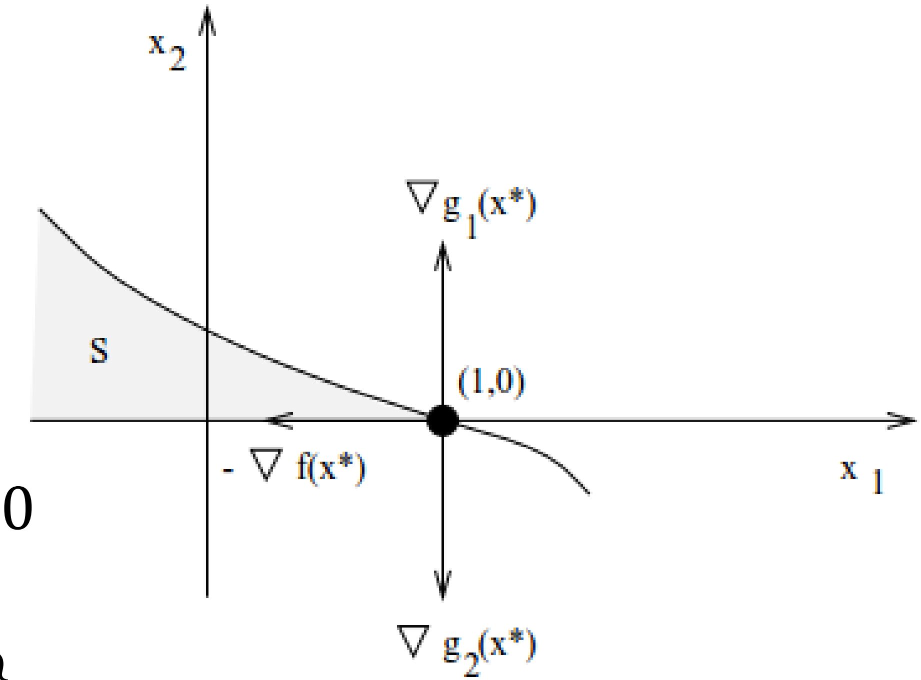
Notes

- If $\nabla f(x^*) = 0$, then $F^{\circ} = \emptyset$ and $F^{\circ} \cap G^{\circ} = \emptyset$
  - All critical points satisfy necessary conditions
- Any $x^* \in S$ that satisfies $g_i(x^*) = 0$ for some $i \in I$ satisfies also the necessary conditions

Fritz John conditions (for inequality constraints only)

- **Necessary conditions:** Let f and g_i be continuously differentiable in $x^* \in S$. If x^* is a local minimizer, then there exist multipliers $\lambda \geq 0$ and $\mu_i \geq 0,\ i = 1, \dots, m$ such that $(\lambda, \mu) \neq 0$ and
  1) $\lambda \nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0$
  2) $\mu_i g_i(x^*) = 0$ for all $i = 1, \dots, m$.
- Multipliers are usually called as *Lagrange multipliers*
  – $\min L(x, \mu) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x), s.t.\ x \in R^n$
- *Conditions 2) are called complementarity conditions*

Example



From Miettinen: *Nonlinear optimization, 2007 (in Finnish)*

- $\min f(x) = x_1$
- $s.t.\ \begin{aligned} x_2 - (1 - x_1)^3 &\le 0 \\ -x_2 &\le 0 \end{aligned}$
- $x^* = (1{,}0)^T,\ I = \{1{,}2\}$
- $\nabla f(x^*) = (1{,}0)^T, \nabla g_1(x^*) = (0{,}1)^T, \nabla g_2(x^*) = (0, -1)^T$
- $\lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ only if $\lambda = 0$!
  - $\mu_1 = \mu_2 = \alpha (> 0)$ will do $\rightarrow$ necessary conditions are satisfied even though x^* is not a local minimizer

KKT conditions

- Assume that $\lambda > 0$ i.e. the gradient of the objective function is taken into account every time.
- Since $\lambda > 0$ we can divide by it and can assume that $\lambda = 1$
- In order to satisfy KKT conditions, some regularity must be assumed from the constraints
  - Constraint qualifications
- **Definition:** A point $x^* \in S$ is *regular* if the set of gradients of the active constraints $\nabla g_i(x^*)$ $(i \in I)$ is linearly independent.

KKT conditions (for inequality constraints)

- **Necessary conditions:** Let f and g_i be continuously differentiable in a regular $x^* \in S$. If x^* is a local minimizer, then there exist multipliers $\mu_i \geq 0,\ i = 1, \ldots, m$ such that
  1) $\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0$
  2) $\mu_i g_i(x^*) = 0$ for all $i = 1, \ldots, m$.
- Note: If $g_i(x) \geq 0$, *then* $\mu_i \leq 0$

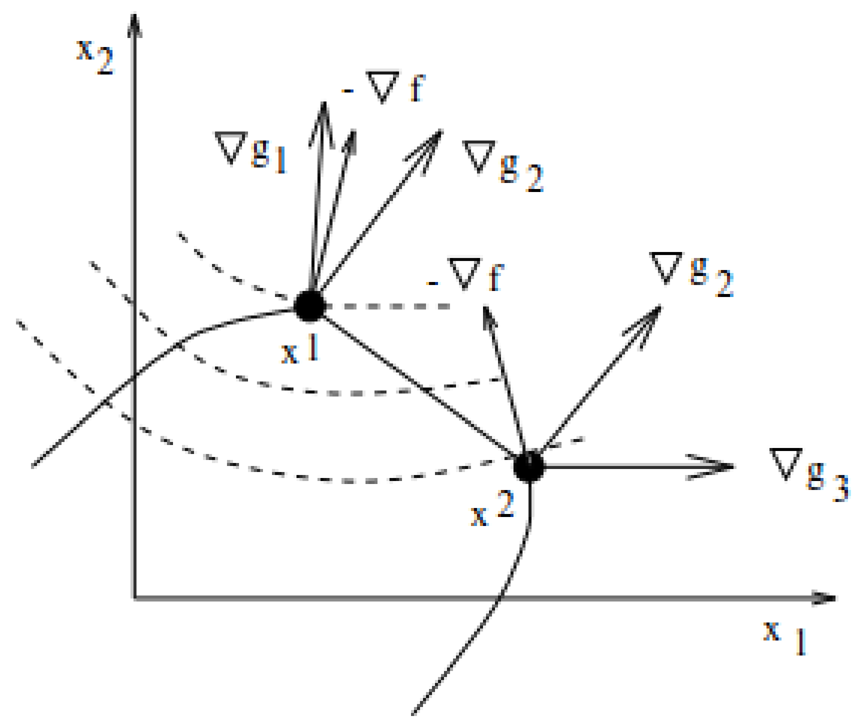
KKT conditions (for inequality constraints)

- **Sufficient conditions:** Let f and g_i be continuously differentiable and convex. Consider $x^* \in S$. If there exist multipliers $\mu_i \geq 0,\ i = 1, \ldots, m$ such that
  1) $\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) = 0$
  2) $\mu_i g_i(x^*) = 0$ for all $i = 1, \ldots, m$,

  then x^* is a global minimizer.

Note

- Alternatively, one can write
$$-\nabla f(x^*) = \sum_{i=1}^{m} \mu_i \nabla g_i(x^*)$$
- That is, $-\nabla f(x^*)$ belongs to the cone defined by the gradients of
the active constraints



Problem with both inequality and equality constraints

- Let us consider $\min f(x),\ s.t.\ g_i(x) \leq 0, i = 1, \dots, m$ and $h_j(x) = 0, j = 1, \dots, l$
- $S = \{x \in R^n \mid g_i(x) \leq 0\ \forall\ i = 1, \dots, m \,\&\, h_j(x) = 0\ \forall\ j = 1, \dots, l\}$ where $g_i: R^n \to R\ \forall i$ and $h_j: R^n \to R\ \forall j$

KKT conditions

- **Definition:** A point $x^* \in S$ is *regular* if the set of gradients of the active inequality constraints $\nabla g_i(x^*)$ $(i \in I)$ and equality constraints $\nabla h_i(x^*)$ $(i = 1, \ldots, l)$ are linearly independent.
- **Necessary conditions:** Let f, g_i and h_i be continuously differentiable in a regular $x^* \in S$. If x^* is a local minimizer, then there exist multipliers $\mu_i \geq 0,\ i = 1, \ldots, m$ and $\nu_i, i = 1, \ldots, l$ such that
  1) $\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{l} \nu_i \nabla h_i(x^*) = 0$
  2) $\mu_i g_i(x^*) = 0$ for all $i = 1, \ldots, m$.

KKT conditions

- **Sufficient conditions:** Let f, g_i and h_i be continuously differentiable and convex. Consider $x^* \in S$. If there exist multipliers $\mu_i \geq 0,\ i = 1, \dots, m$ and $\nu_i, i = 1, \dots, l$ such that
  1) $\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{i=1}^{l} \nu_i \nabla h_i(x^*) = 0$
  2) $\mu_i g_i(x^*) = 0$ for all $i = 1, \dots, m$,

  then, x^* is a global minimizer.

Example

- $\min f(x) = (x_1 - 3)^2 + (x_2 - 2)^2$
  $x_1^2 + x_2^2 \le 5,$
- s.t. $x_1 + 2x_2 = 4,$
  $x_1, x_2 \ge 0$
- **Formulate the KKT conditions**
- $$\begin{gathered}
\mathbf{2x_1 - 6 + 2\mu_1 x_1 - \mu_2 + \nu_1 = 0} \\
\mathbf{2x_2 - 4 + 2\mu_1 x_2 - \mu_3 + 2\nu_1 = 0} \\
\mathbf{\mu_1\left(x_1^2 + x_2^2 - 5\right) = 0} \\
\mathbf{-\mu_2 x_1 = 0} \\
\mathbf{-\mu_3 x_2 = 0} \\
\mathbf{x_1 + 2x_2 - 4 = 0}
\end{gathered}$$
- $\boldsymbol{\mu_1, \mu_2, \mu_3 \ge 0}$