T6: KKT conditions

Pablo Arias Martínez - ENS Paris-Saclay, UPF

October 20, 2022

Optimization and inference techniques for Computer Vision

Previously on...

Convex functions and convex sets

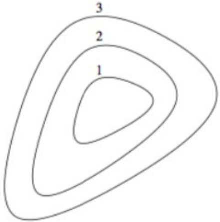
Draw some examples. . .

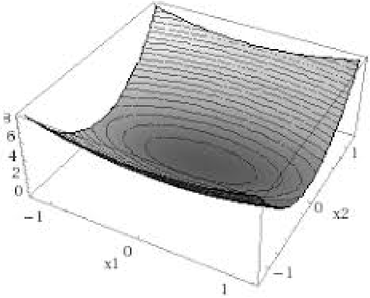
Sublevel sets of convex functions

α-**sublevel set** of $f : \mathbb{R}^n \to \mathbb{R}$:

$$C_\alpha = \{\mathbf{x} \in \mathrm{dom}(f) \mid f(\mathbf{x}) \leq \alpha\}.$$

- sublevel sets of convex functions are convex.
- functions with all their sub-level sets convex, are not necessary convex. These broarder class of functions are called **quasi-convex** (and are also easy to optimize).

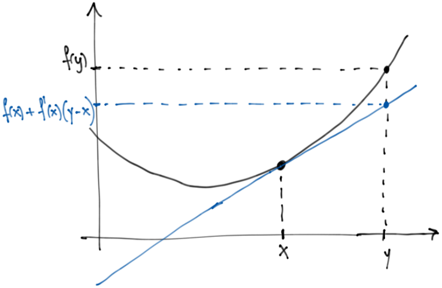




Gradient as a global underestimator

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a **differentiable** function and let C be a convex subset of $\mathbb{R}^n$. Then:

$$f \text{ is } \textbf{convex} \quad \Longleftrightarrow \quad f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in C.$$



(*i.e.* the first order Taylor approximation of f is a **global underestimator**.)

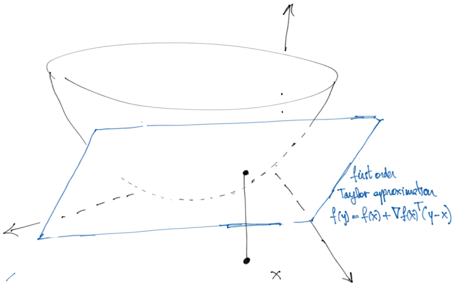
Moreover, f is **strictly convex** if and only if:

$$f(\mathbf{y}) > f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in C.$$

Gradient as a global underestimator

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a **differentiable** function and let C be a convex subset of $\mathbb{R}^n$. Then:

$$f \text{ is } \textbf{convex} \quad \iff \quad f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in C.$$



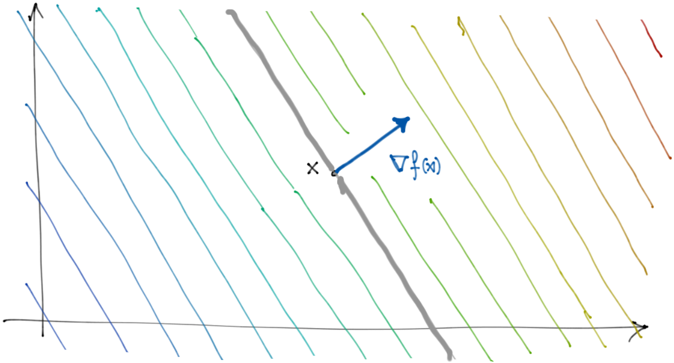
(*i.e.* the first order Taylor approximation of f is a **global underestimator**.)

Moreover, f is **strictly convex** if and only if:

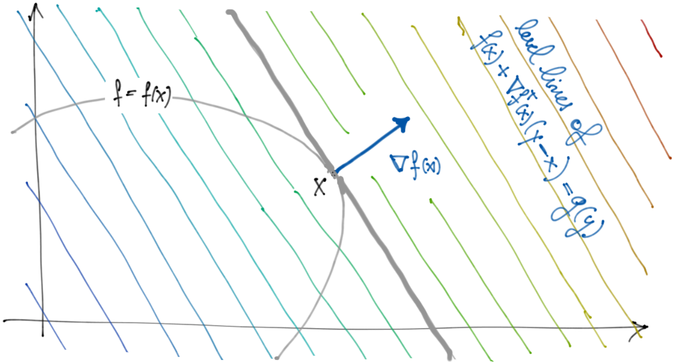
$$f(\mathbf{y}) > f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in C.$$

Gradient as a global underestimator

$f(\mathbf{y}) \geq f(\mathbf{x})$, for all $\mathbf{y}$ in the half-space $\{\mathbf{y} \in \mathbf{R}^n \,:\, \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \geq 0\}$

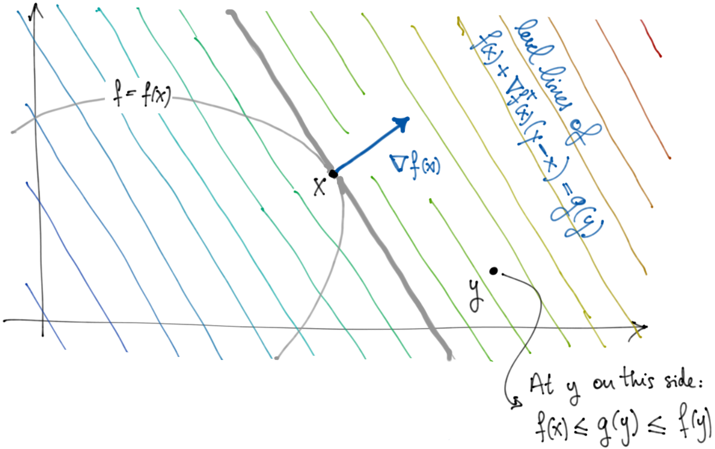


$f(\mathbf{y}) \geq f(\mathbf{x})$, for all $\mathbf{y}$ in the half-space $\{\mathbf{y} \in \mathbf{R}^n \;:\; \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \geq 0\}$



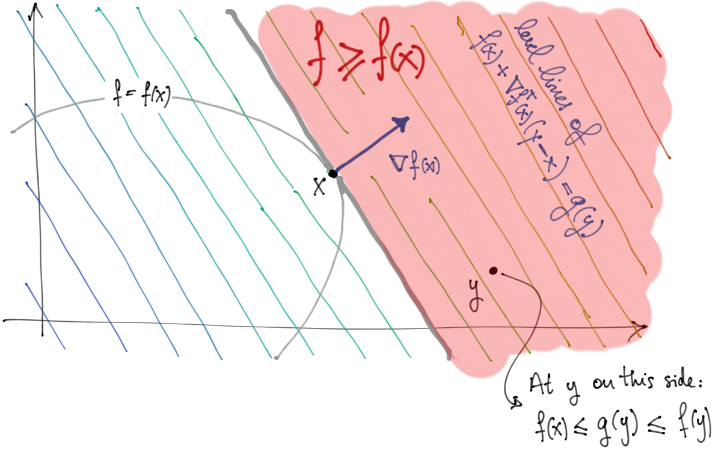
Gradient as a global underestimator

$f(\mathbf{y}) \geq f(\mathbf{x})$, for all $\mathbf{y}$ in the half-space $\{\mathbf{y} \in \mathbf{R}^n \, : \, \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \geq 0\}$



Gradient as a global underestimator

$f(\mathbf{y}) \geq f(\mathbf{x})$, for all $\mathbf{y}$ in the half-space $\{\mathbf{y} \in \mathbf{R}^n \ : \ \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \geq 0\}$



Optimality conditions for convex problems

Consider the **constrained minimization problem**

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in C \end{array}$$

Theorem
Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a **convex function** such that f is not identically $-\infty$ or $+\infty$, and C a convex set. Then

- Any local minimum of f over C is also a global minimum over C.
- Moreover, if f is **strictly convex**, then the global minimum in C is unique.

Optimality condition for unconstrained convex problems

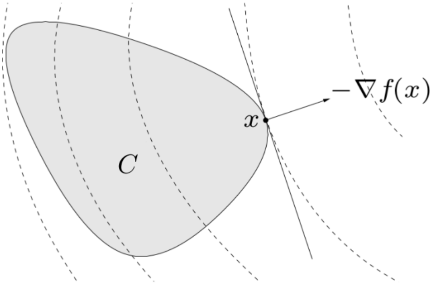
For **convex unconstrained minimization problems** with continuously differentiable f we have the following necessary and sufficient condition.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and convex function. Then:

$$\mathbf{x}^* \text{ is a global minimum of } f \quad \iff \quad \nabla f(\mathbf{x}^*) = 0.$$

Minimization of convex functions on convex sets

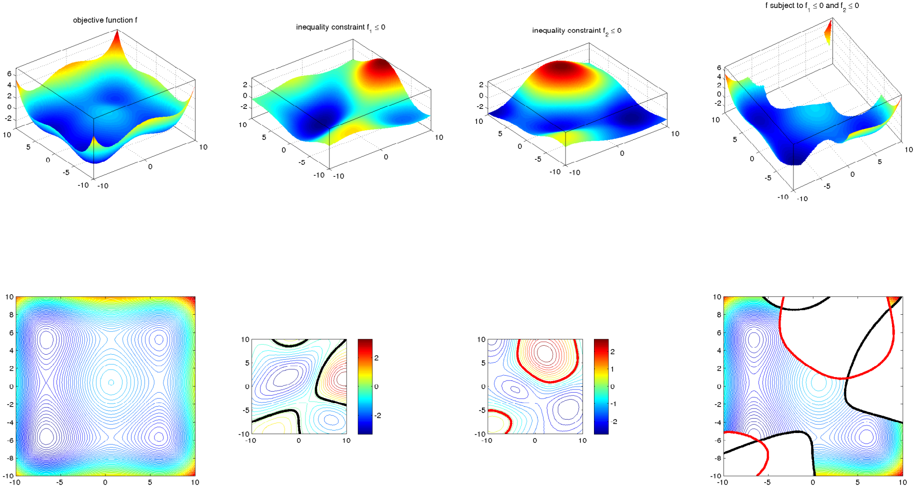
Consider the **constrained minimization problem** $\begin{cases} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in C \end{cases}$



Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and convex function and C a set. Then:

$$\mathbf{x}^* \text{ is a global minimum of } f \text{ in } C \quad \iff \quad \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \geq 0 \quad \forall \mathbf{x} \in C.$$

Explicit constraints: two constaints



Inequality and equality constraints

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $C \subset \mathbb{R}^n$:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & x \in C \end{array}$$

Inequality and equality constraints

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $C \subset \mathbb{R}^n$:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & x \in C \end{array}$$

Let $f, c_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$.

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m. \end{array}$$

The set C is given by $C = \{x \in \mathbb{R}^n \mid c_1(\mathbf{x}) \leq 0, \ldots, c_k(\mathbf{x}) \leq 0\}$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $C \subset \mathbb{R}^n$:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & x \in C \end{array}$$

Let $f, c_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$.

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m. \end{array}$$

The set C is given by $C = \{x \in \mathbb{R}^n \mid c_1(\mathbf{x}) \leq 0, \dots, c_k(\mathbf{x}) \leq 0\}$.

Equality constraints: Suppose $c_j = -c_i$ for some i, j. Then

$$c_j(\mathbf{x}) \leq 0 \Rightarrow c_i(\mathbf{x}) \geq 0 \text{ and } c_i(\mathbf{x}) \leq 0 \qquad \Rightarrow \qquad c_i(\mathbf{x}) = c_j(\mathbf{x}) = 0.$$

Constraint qualification

From now on, we are going to separate equality constraints from inequality constraints, and we are going to assume that the contraints are **qualified**.

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, p$.

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & c_i(\mathbf{x}) \le 0, \quad i = 1, \dots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, \quad j = 1, \dots, p. & \text{equality constraints} \end{array}$$

Constraint qualification

From now on, we are going to separate equality constraints from inequality constraints, and we are going to assume that the contraints are **qualified**.

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

$$\begin{array}{lllr} \text{minimize} & f(\mathbf{x}) & & \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, & i = 1, \ldots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, & j = 1, \ldots, p. & \text{equality constraints} \end{array}$$

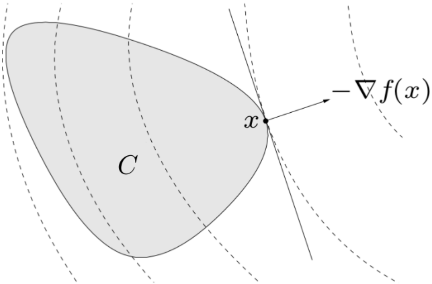
Qualified constraints: we assume that it exists a feasible $\mathbf{x}$ such that

$$c_i(\mathbf{x}) < 0 \text{ for all } i = 1, \ldots, m.$$

This excludes the possibility of pairs i, j with $c_i = -c_j$, among other things.

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

minimize	$f(\mathbf{x})$		
subject to	$c_i(\mathbf{x}) \leq 0,$	$i = 1, \ldots, m,$	inequality constraints
	$d_j(\mathbf{x}) = 0,$	$j = 1, \ldots, p.$	equality constraints



How to express this N&S optimality condition explicitly when C is defined as

$$C = \{\mathbf{x} \in \mathbb{R}^n \,:\, c_i(\mathbf{x}) \leq 0, i = 1, \ldots, m, \quad d_j(\mathbf{x}) = 0, j = 1, \ldots, p\}?$$

KKT conditions for convex differentiable problems

General form of a convex constrained optimization problem

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, \quad j = 1, \ldots, p. & \text{equality constraints} \end{array}$$

For a convex problem: we need a convex objective f and convex set C defined by the constraints:

- Inequality constraints: convex functions c_i
- Equality constraints: **affine functions** functions $d_j(x) = \mathbf{a}_j \mathbf{x} + b_j$

General form of a convex constrained optimization problem

Let $f, c_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $\mathbf{a}_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$, $j = 1, \ldots, p$.

$$
\begin{array}{lll}
\text{minimize} & f(\mathbf{x}) & \\
\text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, & \text{inequality constraints} \\
 & a_j\mathbf{x} - b_j = 0, \quad j = 1, \ldots, p. & \text{equality constraints}
\end{array}
$$

For a convex problem: we need a convex objective f and convex set C defined by the constraints:

- Inequality constraints: convex functions c_i
- Equality constraints: **affine functions** functions $d_j(x) = \mathbf{a}_j\mathbf{x} + b_j$

General form of a convex constrained optimization problem

Let $f, c_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{b} \in \mathbb{R}^p$.

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, & \text{inequality constraints} \\ & \mathbf{A}\mathbf{x} - \mathbf{b} = 0. & \text{equality constraints} \end{array}$$

For a convex problem: we need a convex objective f and convex set C defined by the constraints:

- Inequality constraints: convex functions c_i
- Equality constraints: **affine functions** functions $d_j(x) = \mathbf{a}_j \mathbf{x} + b_j$

In addition we will assume that f, c_i are continuosly differentiable functions.

Lagrangian function

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, \quad j = 1, \ldots, p. & \text{equality constraints} \end{array}$$

Lagrangian function

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, m$, $j = 1, \dots, p$.

$$
\begin{array}{lll}
\text{minimize} & f(\mathbf{x}) & \\
\text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, & \text{inequality constraints} \\
& d_j(\mathbf{x}) = 0, \quad j = 1, \dots, p. & \text{equality constraints}
\end{array}
$$

The **Lagrangian function** for the problem is defined as

$$
\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i c_i(\mathbf{x}) + \sum_{j=1}^{p} \nu_j d_j(\mathbf{x})
$$

where

$$
\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m), \quad \boldsymbol{\nu} = (\nu_1, \dots, \nu_p)
$$

are the **Lagrange multipliers** of the problem.

Karush-Kuhn-Tucker (KKT) optimality conditions

The **KKT conditions** for our optimization problem, at a point $\mathbf{x}$ are:

$$\mathrm{KKT}(\mathbf{x}) =$$

$$\begin{cases} \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla c_i(\mathbf{x}) + \sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x}) = 0 & \text{stationarity} \\ d_j(\mathbf{x}) = 0, j = 1, \ldots, p & \text{primal feasibility} \\ c_i(\mathbf{x}) \leq 0, i = 1, \ldots, m & \text{primal feasibility} \\ \lambda_i \geq 0, i = 1, \ldots, m & \text{dual feasibility} \\ \lambda_i c_i(\mathbf{x}) = 0, i = 1, \ldots, m & \text{complementary slackness} \end{cases}$$

Karush-Kuhn-Tucker (KKT) optimality conditions

The **KKT conditions** for our optimization problem, at a point $\mathbf{x}$ are:

$$\mathrm{KKT}(\mathbf{x}) =$$

$$\begin{cases} \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla c_i(\mathbf{x}) + \sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x}) = 0 & \text{stationarity} \\ d_j(\mathbf{x}) = 0, j = 1, \dots, p & \text{primal feasibility} \\ c_i(\mathbf{x}) \leq 0, i = 1, \dots, m & \text{primal feasibility} \\ \lambda_i \geq 0, i = 1, \dots, m & \text{dual feasibility} \\ \lambda_i c_i(\mathbf{x}) = 0, i = 1, \dots, m & \text{complementary slackness} \end{cases}$$

For a **convex optimization problem** with **qualified constraints** and **continuously differentiable** $f, c_i, i = 1, \dots, m$:

$$\mathbf{x}^* \text{ is a global minimum} \quad \Longleftrightarrow \quad \mathrm{KKT}(\mathbf{x}^*).$$

Karush-Kuhn-Tucker (KKT) optimality conditions

- The KKT conditions are a set of equations and inequalities that we can solve to find the solution to a convex optimization problem.

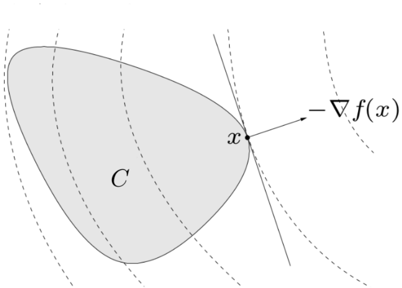
- The KKT conditions only depend on local quantities at a point $\mathbf{x}$. **Convexity** allows us to go from local conditions to a global property (global optimizer)

- In practice, it is often very hard to solve analitycally the KKT conditions. But they can be exploited to derive efficient optimization algorithms (such as interior point methods).

Understanding the KKT conditions

The KKT conditions might seem complicated, but they are nothing else that the conditions for which

$$\nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \geq 0, \forall \mathbf{y} \in C.$$



(Recall that this is a necessary and sufficient condition for optimality for convex differentiable problems)

Their complicated expression derives from the set C:

$$C = \{\mathbf{x} \in \mathbb{R}^n \mid c_i(\mathbf{x}) \leq 0, i = 1, \ldots, m, \quad d_j(\mathbf{x}) = 0, j = 1, \ldots, p\}.$$

Suppose we only have equality constraints. The KKT conditions simplify to:

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + \sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x}) = 0 \qquad \text{stationarity}$$

$$d_j(\mathbf{x}) = 0, j = 1, \ldots, p \qquad \text{primal feasibility}$$

Understanding the KKT conditions: I) equality constraints

Suppose we only have equality constraints. The KKT conditions simplify to:

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + \sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x}) = 0 \qquad \text{stationarity}$$

$$d_j(\mathbf{x}) = 0, j = 1, \ldots, p \qquad \text{primal feasibility}$$

$\nabla f(\mathbf{x})$ is a linear combination of the gradients to the constraint functions:

$$\nabla f(\mathbf{x}) = -\sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x})$$

Understanding the KKT conditions: I) equality constraints

Suppose we only have equality constraints. The KKT conditions simplify to:

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\nu}) = \nabla f(\mathbf{x}) + \sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x}) = 0 \qquad \text{stationarity}$$

$$d_j(\mathbf{x}) = 0, j = 1, \ldots, p \qquad \text{primal feasibility}$$

$\nabla f(\mathbf{x})$ is a linear combination of the gradients to the constraint functions:

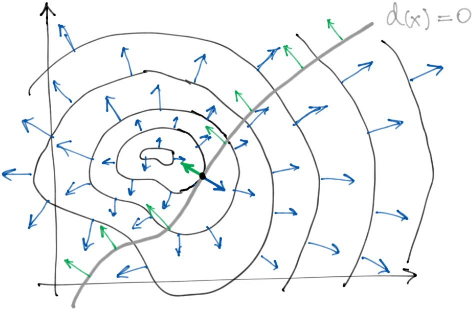
$$\nabla f(\mathbf{x}) = -\sum_{j=1}^{p} \nu_j \nabla d_j(\mathbf{x})$$

This is exactly the **method of Lagrange multipliers** you saw in Calculus!

Understanding the KKT conditions: I) equality constraints

For simplicity, we consider a single equality constraint.

$\nabla f(\mathbf{x})$ is colinear with the gradient of the constraint: $\nabla f(\mathbf{x}) = -\nu \nabla d(\mathbf{x})$.



Note: the drawings correspond to a nonconvex problem!! For a convex problem, the equality constraints have to be affine (a line).

Understanding the KKT conditions: II) only inequality constraints

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla c_i(\mathbf{x}) = 0 \qquad \text{stationarity}$$

$$c_i(\mathbf{x}) \leq 0, i = 1, \ldots, m \qquad \text{primal feasibility}$$

$$\lambda_i \geq 0, i = 1, \ldots, m \qquad \text{dual feasibility}$$

$$\lambda_i c_i(\mathbf{x}) = 0, i = 1, \ldots, m \qquad \text{complementary slackness}$$

$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla c_i(\mathbf{x}) = 0$	stationarity
$c_i(\mathbf{x}) \leq 0, i = 1, \ldots, m$	primal feasibility
$\lambda_i \geq 0, i = 1, \ldots, m$	dual feasibility
$\lambda_i c_i(\mathbf{x}) = 0, i = 1, \ldots, m$	complementary slackness

Complementary slackness: $\lambda_i c_i(\mathbf{x}) = 0$

if $\lambda_i > 0$	then $c_i(\mathbf{x}) = 0$ (**active constraint**)
if $c_i(\mathbf{x}) < 0$ (**inactive constraint**)	then $\lambda_i = 0$

Understanding the KKT conditions: II) only inequality constraints

$$
\begin{aligned}
&\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x},\boldsymbol{\lambda}) = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla c_i(\mathbf{x}) = 0 && \text{stationarity} \\
&c_i(\mathbf{x}) \leq 0, i = 1, \ldots, m && \text{primal feasibility} \\
&\lambda_i \geq 0, i = 1, \ldots, m && \text{dual feasibility} \\
&\lambda_i c_i(\mathbf{x}) = 0, i = 1, \ldots, m && \text{complementary slackness}
\end{aligned}
$$

Complementary slackness: $\lambda_i c_i(\mathbf{x}) = 0$

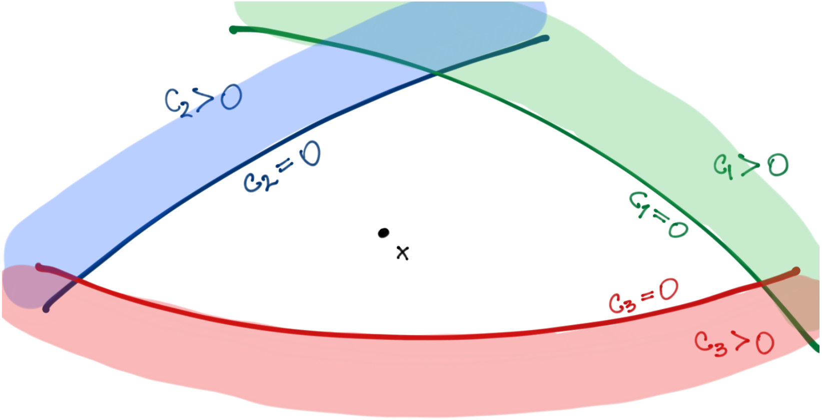
if $\lambda_i > 0$	then $c_i(\mathbf{x}) = 0$ (**active constraint**)
if $c_i(\mathbf{x}) < 0$ (**inactive constraint**)	then $\lambda_i = 0$

Let $A(x) = \{i = 1, \ldots, m \mid c_i(\mathbf{x}) = 0\}$, the set of active constraints. Then

$$
\nabla f(\mathbf{x}) = -\sum_{i \in A(\mathbf{x})} \lambda_i \nabla c_i(\mathbf{x}) = 0, \quad \text{with } \lambda_i \geq 0.
$$

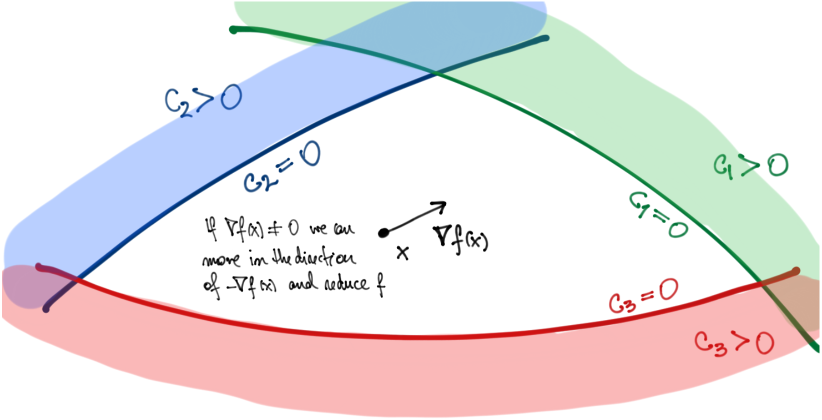
Example: All constraints inactive $c_i(\mathbf{x}) < 0, \forall i$

$$\nabla f(\mathbf{x}) = 0$$



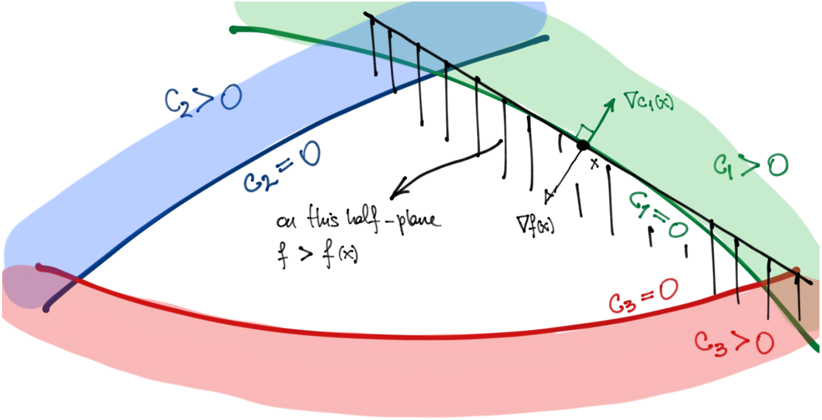
Example: All constraints inactive $c_i(\mathbf{x}) < 0, \forall i$

$$\nabla f(\mathbf{x}) = 0$$



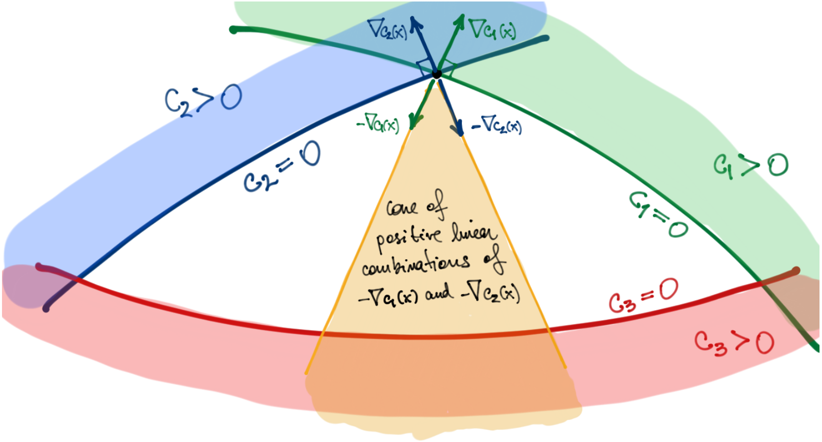
Example: Only one active constraint $c_1(\mathbf{x}) = 0, c_i(\mathbf{x}) < 0, i = 2, \dots, m$

$$\nabla f(\mathbf{x}) = -\lambda_1 \nabla c_1(\mathbf{x}), \quad \text{with } \lambda_1 \geq 0$$



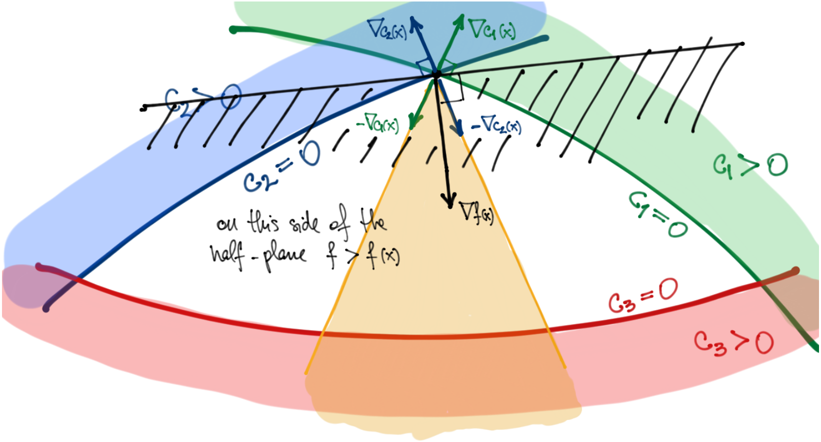
Example: Two active constraints $c_1(\mathbf{x}) = 0, c_2(\mathbf{x}) = 0, c_i(\mathbf{x}) < 0, i = 3, \ldots, m$

$$\nabla f(\mathbf{x}) = -\lambda_1 \nabla c_1(\mathbf{x}) - \lambda_2 \nabla c_2(\mathbf{x}) \quad \text{with } \lambda_1, \lambda_2 \geq 0.$$



Example: Two active constraints $c_1(\mathbf{x}) = 0$, $c_2(\mathbf{x}) = 0$, $c_i(\mathbf{x}) < 0$, $i = 3, \ldots, m$

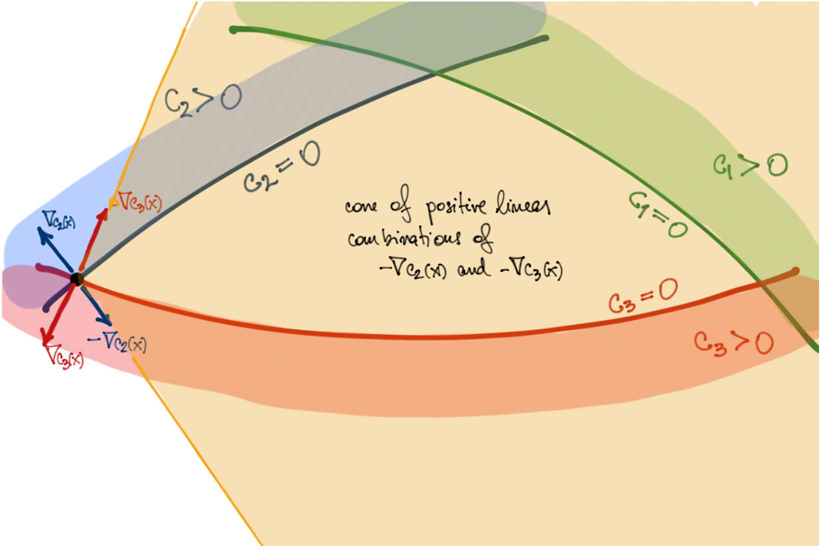
$$\nabla f(\mathbf{x}) = -\lambda_1 \nabla c_1(\mathbf{x}) - \lambda_2 \nabla c_2(\mathbf{x}) \quad \text{with } \lambda_1, \lambda_2 \geq 0.$$



Understanding the KKT conditions: II) only inequality constraints

Example: Two active constraints $c_1(\mathbf{x}) = 0, c_2(\mathbf{x}) = 0, c_i(\mathbf{x}) < 0, i = 3, \ldots, m$

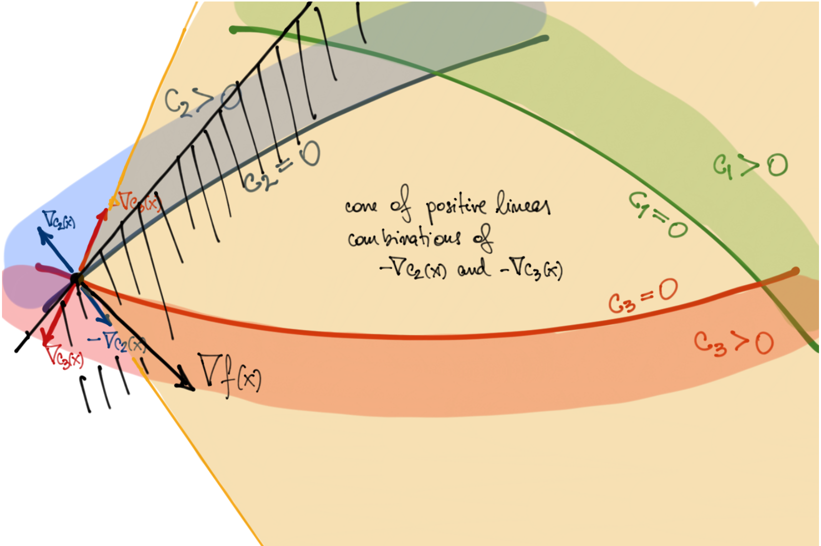
$$\nabla f(\mathbf{x}) = -\lambda_1 \nabla c_1(\mathbf{x}) - \lambda_2 \nabla c_2(\mathbf{x}) \quad \text{with } \lambda_1, \lambda_2 \geq 0.$$



Understanding the KKT conditions: II) only inequality constraints

Example: Two active constraints $c_1(\mathbf{x}) = 0, c_2(\mathbf{x}) = 0, c_i(\mathbf{x}) < 0, i = 3, \ldots, m$

$$\nabla f(\mathbf{x}) = -\lambda_1 \nabla c_1(\mathbf{x}) - \lambda_2 \nabla c_2(\mathbf{x}) \quad \text{with } \lambda_1, \lambda_2 \geq 0.$$



KKT conditions without convexity

Constrained differentiable optimization problem

Let $f, c_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, d_j, $j = 1, \ldots, p$.

$$(\mathcal{P}) \begin{cases} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & c_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m, \qquad \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, \quad j = 1, \ldots, p. \qquad \text{equality constraints} \end{cases}$$

For the **optimization problem** $(\mathcal{P})$ with **qualified constraints*** and **continuously differentiable** $f, c_i, i = 1, \ldots, m, d_j, j = 1, \ldots, p$:

$$\mathbf{x}^* \text{ is a global minimum} \quad \Longrightarrow \quad \text{KKT}(\mathbf{x}^*).$$

Constrained differentiable optimization problem

Let $f, c_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, d_j, $j = 1, \ldots, p$.

$$(\mathcal{P}) \begin{cases} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & c_i(\mathbf{x}) \le 0, \quad i = 1, \ldots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, \quad j = 1, \ldots, p. & \text{equality constraints} \end{cases}$$

$$\boxed{\begin{array}{c} \text{For the } \textbf{optimization problem } (\mathcal{P}) \text{ with } \textbf{qualified constraints}^* \text{ and} \\ \textbf{continuously differentiable } f, c_i, i = 1, \ldots, m, d_j, j = 1, \ldots, p: \\[1em] \mathbf{x}^* \text{ is a global minimum} \quad \Longrightarrow \quad \mathrm{KKT}(\mathbf{x}^*). \end{array}}$$

(*) For this statement we use a more restrictive **constraint qualification**:
The gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at $\mathbf{x}^*$.

Constrained differentiable optimization problem

- If we remove the convexity assumption, the KKT conditions are necessary conditions, but not sufficient.

- There is a sufficient **second-order** optimality condition based on the $\nabla^2_{\mathbf{xx}}\mathcal{L}$, the Hessian of the Lagrangian function. We won't cover it, but in summary it requires that the second directional derivatives in all admissible directions are strictly positive.

- A technical detail: the constraint qualification in the previous statement is more restrictive than that used in convex problems. There are other less-restrictive constraint qualifications, but we won't cover them.

What's next?

Does our problem have a solution?
(Existence) ✓

Does our problem have an unique solution?
(Uniqueness) ✓

Is it possible to find find the solution?
(Convexity) ✓

How to tell if a point $\mathbf{x}$ is a solution for a constrained problem?
(Optimality conditions - convex differentiable problems) ✓

What's next?

Does our problem have a solution?
(Existence) ✓

Does our problem have an unique solution?
(Uniqueness) ✓

Is it possible to find find the solution?
(Convexity) ✓

How to tell if a point $\mathbf{x}$ is a solution for a constrained problem?
(Optimality conditions - convex differentiable problems) ✓
(Optimality conditions - non-convex differentiable problems) ✓
(Optimality conditions - convex non-differentiable problems) ✗

Non-differentiable convex problems

Convex optimization does not require differentiability

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Let C be a closed convex subset of $\mathbb{R}^N$.

Consider the **constrained minimization problem**

$$\min_{\mathbf{x} \in C} f(\mathbf{x}).$$

Theorem
Assume that C is a **convex subset** of $\mathbb{R}^n$.

Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a **convex function** such that f is not identically $-\infty$ or $+\infty$. Then, **a local minimum of f over C is also a global minimum over C**.

Moreover, if f is **strictly convex**, then **any global minimum in C is unique** (there exists at most one global minimum over C).

Subgradient of a function

Recall that convex and differentiable functions f satisfy the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$$

What if f is not differentiable?

Subgradient of a function

Recall that convex and differentiable functions f satisfy the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$$

What if f is not differentiable?

We say that $\mathbf{g} \in \mathbb{R}^n$ is a **subgradient** of a function $f : \mathbb{R}^n \to \mathbb{R}$ (not necessarily convex) at $\mathbf{x}$ if satisfies the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}), \; \forall \mathbf{y} \in \mathrm{dom}(f)$$

Subgradient of a function

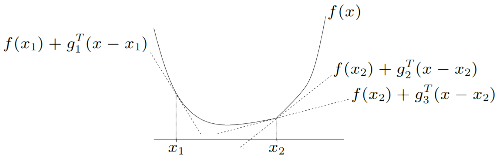
Recall that convex and differentiable functions f satisfy the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$$

What if f is not differentiable?

We say that $\mathbf{g} \in \mathbb{R}^n$ is a **subgradient** of a function $f : \mathbb{R}^n \to \mathbb{R}$ (not necessarily convex) at $\mathbf{x}$ if satisfies the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}), \ \forall \mathbf{y} \in \mathrm{dom}(f)$$



From the course slides *Convex Optimization II - Stanford*

Subgradient of a function

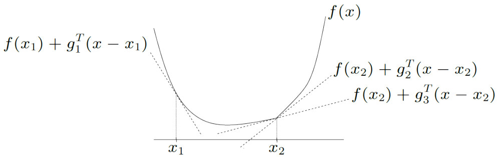
Recall that convex and differentiable functions f satisfy the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)$$

What if f is not differentiable?

We say that $\mathbf{g} \in \mathbb{R}^n$ is a **subgradient** of a function $f : \mathbb{R}^n \to \mathbb{R}$ (not necessarily convex) at $\mathbf{x}$ if satisfies the condition:

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}), \, \forall \mathbf{y} \in \mathrm{dom}(f)$$



From the course slides *Convex Optimization II - Stanford*

If f is convex and differentiable, $\nabla f(\mathbf{x})$ is a subgradient of f at $\mathbf{x}$.

Example: subgradients in $\mathbb{R}^2$

Subdifferential of a function

The **subdifferential** $\partial f(\mathbf{x})$ of $f : \mathbb{R}^n \to \mathbb{R}$ at a point $\mathbf{x}$ is the set of all subgradients:

$$\partial f(\mathbf{x}) = \{\mathbf{g} \in \mathbb{R}^n : \; f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}) \; \forall \mathbf{y} \in \mathbb{R}^n\}$$

Subdifferential of a function

The **subdifferential** $\partial f(\mathbf{x})$ of $f : \mathbb{R}^n \to \mathbb{R}$ at a point $\mathbf{x}$ is the set of all subgradients:

$$\partial f(\mathbf{x}) = \{\mathbf{g} \in \mathbb{R}^n : \; f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}) \; \forall \mathbf{y} \in \mathbb{R}^n\}$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is a **convex function**, then the subdifferential $\partial f(\mathbf{x})$ is **non-empty, convex and compact** for all $\mathbf{x} \in \mathbb{R}^n$. $\Rightarrow$**(It exists!)**

Subdifferential of a function

The **subdifferential** $\partial f(\mathbf{x})$ of $f : \mathbb{R}^n \to \mathbb{R}$ at a point $\mathbf{x}$ is the set of all subgradients:

$$\partial f(\mathbf{x}) = \{\mathbf{g} \in \mathbb{R}^n : \; f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}) \; \forall \mathbf{y} \in \mathbb{R}^n\}$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is a **convex function**, then the subdifferential $\partial f(\mathbf{x})$ is **non-empty, convex and compact** for all $\mathbf{x} \in \mathbb{R}^n$. $\Rightarrow$**(It exists!)**

If f is differentiable at the point $\mathbf{x}$, then $\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\}$.

Subdifferential of a function

The **subdifferential** $\partial f(\mathbf{x})$ of $f : \mathbb{R}^n \to \mathbb{R}$ at a point $\mathbf{x}$ is the set of all subgradients:

$$\partial f(\mathbf{x}) = \{\mathbf{g} \in \mathbb{R}^n : \; f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{y} - \mathbf{x}) \; \forall \mathbf{y} \in \mathbb{R}^n\}$$

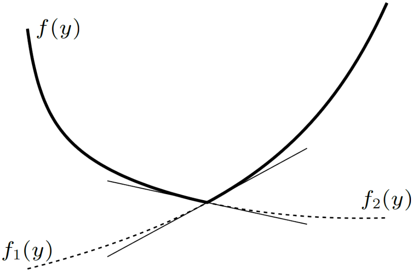
If $f : \mathbb{R}^n \to \mathbb{R}$ is a **convex function**, then the subdifferential $\partial f(\mathbf{x})$ is **non-empty, convex and compact** for all $\mathbf{x} \in \mathbb{R}^n$. $\Rightarrow$**(It exists!)**

If f is differentiable at the point $\mathbf{x}$, then $\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\}$.

Example: the subdifferential of the $f(\mathbf{x}) = |\mathbf{x}|$.

Subdifferential of a function

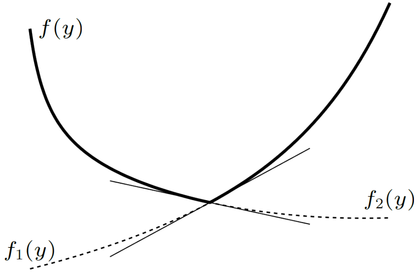
Let define $f = \max\{f_1, f_2\}$, with f_1, f_2 convex and differentiable functions:



From the course slides *Convex Optimization II - Stanford*

Subdifferential of a function

Let define $f = \max\{f_1, f_2\}$, with f_1, f_2 convex and differentiable functions:

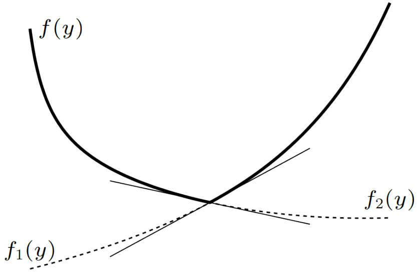


From the course slides *Convex Optimization II - Stanford*

- When $f_1(\mathbf{x}) > f_2(\mathbf{x})$, there is an unique subgradient $p = \nabla f_1(\mathbf{x})$.

Subdifferential of a function

Let define $f = \max\{f_1, f_2\}$, with f_1, f_2 convex and differentiable functions:

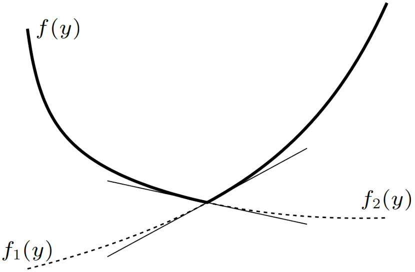


From the course slides *Convex Optimization II - Stanford*

- When $f_1(\mathbf{x}) > f_2(\mathbf{x})$, there is an unique subgradient $p = \nabla f_1(\mathbf{x})$.
- When $f_2(\mathbf{x}) > f_1(\mathbf{x})$, there is an unique subgradient $p = \nabla f_2(\mathbf{x})$.

Subdifferential of a function

Let define $f = \max\{f_1, f_2\}$, with f_1, f_2 convex and differentiable functions:



From the course slides *Convex Optimization II - Stanford*

- When $f_1(\mathbf{x}) > f_2(\mathbf{x})$, there is an unique subgradient $p = \nabla f_1(\mathbf{x})$.
- When $f_2(\mathbf{x}) > f_1(\mathbf{x})$, there is an unique subgradient $p = \nabla f_2(\mathbf{x})$.
- At $f_1(\mathbf{x}) = f_2(\mathbf{x})$, the subdifferential forms the line segment $[\nabla f_2(\mathbf{x}), \nabla f_1(\mathbf{x})]$.

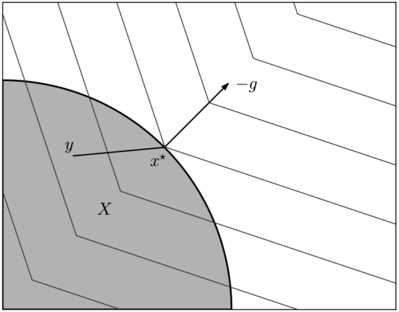
Example: subdifferntial in $\mathbb{R}^2$

Convex unconstrained minimization: optimality condition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be and convex function. Then:

$$\mathbf{x}^* \text{ is a global minimum of } f \text{ in } \mathrm{dom}(f) \quad \iff \quad \mathbf{0} \in \partial f(\mathbf{x}^*).$$

Convex constrained minimization: optimality condition



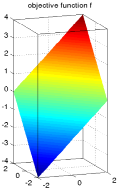
Let $f : \mathbb{R}^n \to \mathbb{R}$ be and convex function and C a set. Then:

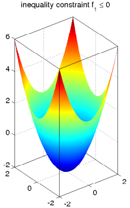
$\mathbf{x}^*$ is a global minimum of f in C $\iff$

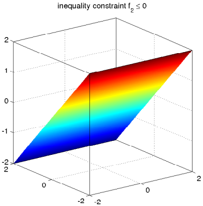
$$\exists \mathbf{g} \in \partial f(\mathbf{x}^*) \text{ such that } \mathbf{g}^T(\mathbf{x} - \mathbf{x}^*) \geq 0 \quad \forall \mathbf{x} \in C. \quad (1)$$

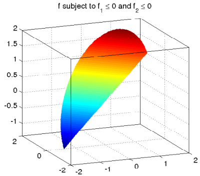
Explicit solution of the KKT conditions

Examples: solving the KKT conditions manually

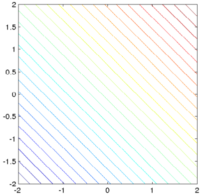


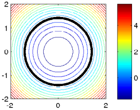


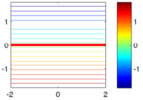


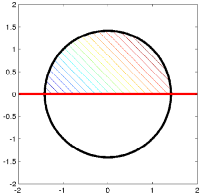


$$
\begin{array}{ll}
\text{minimize} & x_1 + x_2 \\
\text{subject to} & x_1^2 + x_2^2 - 2 \leq 0 \\
 & -x_2 \leq 0.
\end{array}
$$



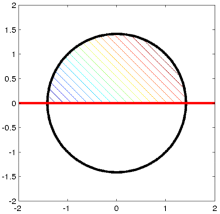






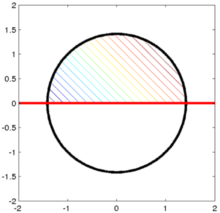
Examples: solving the KKT conditions manually

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) = x_1 + x_2 \\ \text{subject to} & c_1(\mathbf{x}) = x_1^2 + x_2^2 - 2 \leq 0 \\ & c_2(\mathbf{x}) = -x_2 \leq 0. \end{array}$$



Examples: solving the KKT conditions manually

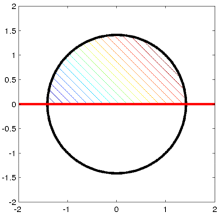
$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) = x_1 + x_2 \\ \text{subject to} & c_1(\mathbf{x}) = x_1^2 + x_2^2 - 2 \leq 0 \\ & c_2(\mathbf{x}) = -x_2 \leq 0. \end{array}$$



Lagrangian: $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = x_1 + x_2 + \lambda_1(x_1^2 + x^2 - 2) + \lambda_2(-x_2)$

Examples: solving the KKT conditions manually

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) = x_1 + x_2 \\ \text{subject to} & c_1(\mathbf{x}) = x_1^2 + x_2^2 - 2 \leq 0 \\ & c_2(\mathbf{x}) = -x_2 \leq 0. \end{array}$$



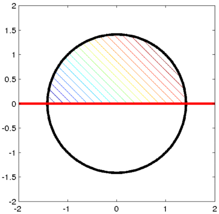
Lagrangian: $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = x_1 + x_2 + \lambda_1(x_1^2 + x^2 - 2) + \lambda_2(-x_2)$

Let's write the KKT conditions, KKT($\mathbf{x}$)

$$\begin{array}{lll} \text{KKT}(\mathbf{x}): & (S1) & \dfrac{\partial \mathcal{L}}{\partial x_1}(\mathbf{x}, \boldsymbol{\lambda}) = 1 + 2\lambda_1 x_1 = 0 \\ & (S2) & \dfrac{\partial \mathcal{L}}{\partial x_2}(\mathbf{x}, \boldsymbol{\lambda}) = 1 + 2\lambda_1 x_2 - \lambda_2 = 0 \\ & (DF) & \lambda_1, \lambda_2 \geq 0 \\ & (PF1) & x_1^2 + x_2^2 - 2 \leq 0 \\ & (PF2) & -x_2 \leq 0 \\ & (CS1) & \lambda_1(x_1^2 + x_2^2 - 2) = 0 \\ & (CS2) & \lambda_2(-x_2) = 0 \end{array}$$

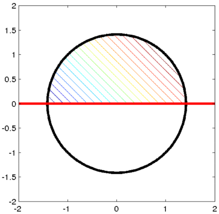
Examples: solving the KKT conditions manually

$$
\mathrm{KKT}(\mathbf{x}): \quad
\begin{array}{lr}
(S1) & 1 + 2\lambda_1 x_1 = 0 \\
(S2) & 1 + 2\lambda_1 x_2 - \lambda_2 = 0 \\
(DF) & \lambda_1, \lambda_2 \geq 0 \\
(PF1) & x_1^2 + x_2^2 - 2 \leq 0 \\
(PF2) & -x_2 \leq 0 \\
(CS1) & \lambda_1(x_1^2 + x_2^2 - 2) = 0 \\
(CS2) & \lambda_2(-x_2) = 0
\end{array}
$$



Examples: solving the KKT conditions manually

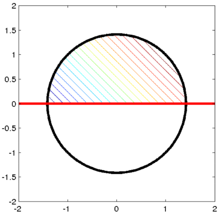
$$
\begin{aligned}
\mathrm{KKT}(\mathbf{x}):\quad & (S1) & 1 + 2\lambda_1 x_1 = 0 \\
& (S2) & 1 + 2\lambda_1 x_2 - \lambda_2 = 0 \\
& (DF) & \lambda_1, \lambda_2 \geq 0 \\
& (PF1) & x_1^2 + x_2^2 - 2 \leq 0 \\
& (PF2) & -x_2 \leq 0 \\
& (CS1) & \lambda_1(x_1^2 + x_2^2 - 2) = 0 \\
& (CS2) & \lambda_2(-x_2) = 0
\end{aligned}
$$



Let's start from $(S1)$: $\lambda_1 = -\frac{1}{2x_1}$.
From here we get that $\lambda_1 > 0$ (the constraint c_1 is **active**), and, using (DF), that $x_1 < 0$. Thus the solution is somewhere in the top left quarter of the circle.

Examples: solving the KKT conditions manually

$$
\mathrm{KKT}(\mathbf{x}): \quad
\begin{array}{lr}
(S1) & 1 + 2\lambda_1 x_1 = 0 \\
(S2) & 1 + 2\lambda_1 x_2 - \lambda_2 = 0 \\
(DF) & \lambda_1, \lambda_2 \geq 0 \\
(PF1) & x_1^2 + x_2^2 - 2 \leq 0 \\
(PF2) & -x_2 \leq 0 \\
(CS1) & \lambda_1(x_1^2 + x_2^2 - 2) = 0 \\
(CS2) & \lambda_2(-x_2) = 0
\end{array}
$$

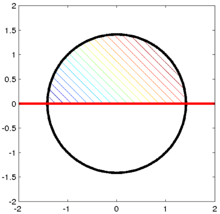


Let's start from $(S1)$: $\lambda_1 = -\frac{1}{2x_1}$.
From here we get that $\lambda_1 > 0$ (the constraint c_1 is **active**), and, using (DF), that $x_1 < 0$. Thus the solution is somewhere in the top left quarter of the circle.

Let's now see if c_2 is active.
$(CS2):\ \lambda_2 x_2 = 0$. Suppose $\lambda_2 = 0$, and substitute in $(S2):\ 1 + 2\lambda_1 x_2 = 0$.
We have $\lambda_1 > 0$ and $x_2 \geq 0$ due to $(PF2)$. Thus $1 + 2\lambda_1 x_2$ cannot be 0, and therefore $\lambda_2 > 0$, and c_2 is also active, which means that $x_2 = 0$.

Examples: solving the KKT conditions manually

$$
\mathrm{KKT}(\mathbf{x}): \quad
\begin{array}{lr}
(S1) & 1 + 2\lambda_1 x_1 = 0 \\
(S2) & 1 + 2\lambda_1 x_2 - \lambda_2 = 0 \\
(DF) & \lambda_1, \lambda_2 \geq 0 \\
(PF1) & x_1^2 + x_2^2 - 2 \leq 0 \\
(PF2) & -x_2 \leq 0 \\
(CS1) & \lambda_1(x_1^2 + x_2^2 - 2) = 0 \\
(CS2) & \lambda_2(-x_2) = 0
\end{array}
$$



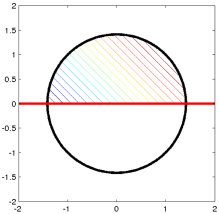
Let's start from $(S1)$: $\lambda_1 = -\frac{1}{2x_1}$.
From here we get that $\lambda_1 > 0$ (the constraint c_1 is **active**), and, using (DF), that $x_1 < 0$. Thus the solution is somewhere in the top left quarter of the circle.

Let's now see if c_2 is active.
$(CS2): \ \lambda_2 x_2 = 0$. Suppose $\lambda_2 = 0$, and substitute in $(S2): \ 1 + 2\lambda_1 x_2 = 0$.
We have $\lambda_1 > 0$ and $x_2 \geq 0$ due to $(PF2)$. Thus $1 + 2\lambda_1 x_2$ cannot be 0, and therefore $\lambda_2 > 0$, and c_2 is also active, which means that $x_2 = 0$.

Using $x_2 = 0$ in $(PF1): x_1^2 = 2 \Rightarrow x_1 = -\sqrt{2}$. Thus $\mathbf{x}^* = (-\sqrt{2}, 0)$.

Examples: solving the KKT conditions manually

$$
\begin{array}{lll}
\mathrm{KKT}(\mathbf{x}): & (S1) & 1 + 2\lambda_1 x_1 = 0 \\
& (S2) & 1 + 2\lambda_1 x_2 - \lambda_2 = 0 \\
& (DF) & \lambda_1, \lambda_2 \geq 0 \\
& (PF1) & x_1^2 + x_2^2 - 2 \leq 0 \\
& (PF2) & -x_2 \leq 0 \\
& (CS1) & \lambda_1(x_1^2 + x_2^2 - 2) = 0 \\
& (CS2) & \lambda_2(-x_2) = 0
\end{array}
$$



Let's start from $(S1)$: $\lambda_1 = -\frac{1}{2x_1}$.
From here we get that $\lambda_1 > 0$ (the constraint c_1 is **active**), and, using (DF), that $x_1 < 0$. Thus the solution is somewhere in the top left quarter of the circle.

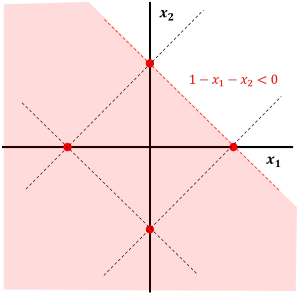
Let's now see if c_2 is active.
$(CS2): \lambda_2 x_2 = 0$. Suppose $\lambda_2 = 0$, and substitute in $(S2): 1 + 2\lambda_1 x_2 = 0$.
We have $\lambda_1 > 0$ and $x_2 \geq 0$ due to $(PF2)$. Thus $1 + 2\lambda_1 x_2$ cannot be 0, and therefore $\lambda_2 > 0$, and c_2 is also active, which means that $x_2 = 0$.

Using $x_2 = 0$ in $(PF1): x_1^2 = 2 \Rightarrow x_1 = -\sqrt{2}$. Thus $\mathbf{x}^* = (-\sqrt{2}, 0)$.

From $(S1)$ we get $\lambda_1 = \frac{1}{2\sqrt{2}}$, and from $(S2)$ we get $\lambda_2 = 1$.

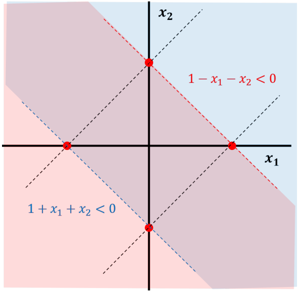
Examples: solving the KKT conditions manually

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) = \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{8}\right)^2 \\ \text{subject to:} & -c_1(\mathbf{x}) = 1 - x_1 - x_2 \geq 0 \\ & -c_2(\mathbf{x}) = 1 + x_1 + x_2 \geq 0 \\ & -c_3(\mathbf{x}) = 1 + x_1 - x_2 \geq 0 \\ & -c_4(\mathbf{x}) = 1 - x_1 + x_2 \geq 0. \end{array}$$



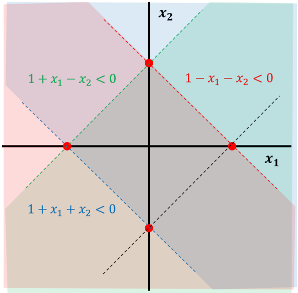
Examples: solving the KKT conditions manually

$$
\begin{array}{ll}
\text{minimize} & f(\mathbf{x}) = \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{8}\right)^2 \\
\text{subject to:} & -c_1(\mathbf{x}) = 1 - x_1 - x_2 \geq 0 \\
& -c_2(\mathbf{x}) = 1 + x_1 + x_2 \geq 0 \\
& -c_3(\mathbf{x}) = 1 + x_1 - x_2 \geq 0 \\
& -c_4(\mathbf{x}) = 1 - x_1 + x_2 \geq 0.
\end{array}
$$



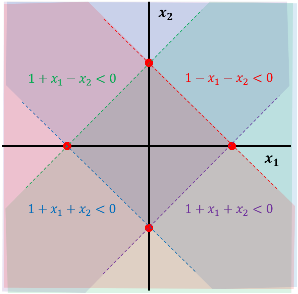
Examples: solving the KKT conditions manually

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) = \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{8}\right)^2 \\ \text{subject to:} & -c_1(\mathbf{x}) = 1 - x_1 - x_2 \geq 0 \\ & -c_2(\mathbf{x}) = 1 + x_1 + x_2 \geq 0 \\ & -c_3(\mathbf{x}) = 1 + x_1 - x_2 \geq 0 \\ & -c_4(\mathbf{x}) = 1 - x_1 + x_2 \geq 0. \end{array}$$



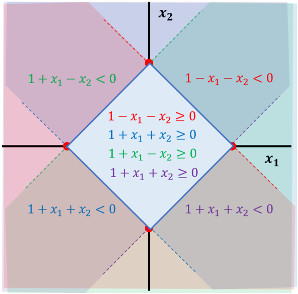
Examples: solving the KKT conditions manually

$$
\begin{array}{ll}
\text{minimize} & f(\mathbf{x}) = \left(x_1 - \frac{3}{2}\right)^2 + \left(x_2 - \frac{1}{8}\right)^2 \\
\text{subject to:} & -c_1(\mathbf{x}) = 1 - x_1 - x_2 \geq 0 \\
 & -c_2(\mathbf{x}) = 1 + x_1 + x_2 \geq 0 \\
 & -c_3(\mathbf{x}) = 1 + x_1 - x_2 \geq 0 \\
 & -c_4(\mathbf{x}) = 1 - x_1 + x_2 \geq 0.
\end{array}
$$



Examples: solving the KKT conditions manually

$$
\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) = \left(x_1 - \tfrac{3}{2}\right)^2 + \left(x_2 - \tfrac{1}{8}\right)^2 \\
\text{subject to:} \quad & -c_1(\mathbf{x}) = 1 - x_1 - x_2 \geq 0 \\
& -c_2(\mathbf{x}) = 1 + x_1 + x_2 \geq 0 \\
& -c_3(\mathbf{x}) = 1 + x_1 - x_2 \geq 0 \\
& -c_4(\mathbf{x}) = 1 - x_1 + x_2 \geq 0.
\end{aligned}
$$



Lagrangian function - different sign convention

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

$$\begin{array}{lllr} \text{minimize} & f(\mathbf{x}) & & \\ \text{subject to} & c_i(\mathbf{x}) \geq 0, & i = 1, \ldots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, & j = 1, \ldots, p. & \text{equality constraints} \end{array}$$

Sometimes we express the inequalities as upper-level sets $c_i(\mathbf{x}) \geq 0$. This formulation requires c_i to be concave (so that $-c_i$ is convex).

Lagrangian function - different sign convention

Let $f, c_i, d_j : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $j = 1, \ldots, p$.

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) & \\ \text{subject to} & c_i(\mathbf{x}) \geq 0, \quad i = 1, \ldots, m, & \text{inequality constraints} \\ & d_j(\mathbf{x}) = 0, \quad j = 1, \ldots, p. & \text{equality constraints} \end{array}$$

Sometimes we express the inequalities as upper-level sets $c_i(\mathbf{x}) \geq 0$. This formulation requires c_i to be concave (so that $-c_i$ is convex).

With this convention, the Lagrangian is as follows:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i c_i(\mathbf{x}) - \sum_{j=1}^{p} \nu_j d_j(\mathbf{x})$$

The KKT conditions **are the same.**