Nonlinear optimization problem without constraints NPP:

Objective function f(x) :

$$f(\mathbf{x}): R^n \longrightarrow R^1$$

The optimization problem consists of finding a vector of decision variables, belonging to the feasible set of solutions R^n

such, that

$$\forall \mathbf{x} \in R^n \qquad f\left(\hat{\mathbf{x}}\right) \le f(\mathbf{x})$$

It is denoted as:

$$\min_{\mathbf{x} \in R^n} f(\mathbf{x}) = f\left(\hat{\mathbf{x}}\right)$$

Local minimum and global minimum of the function *f(x)*

The vector $\hat{x}$ is local minimum of the function *f(x) in the space R^n, if there exists such open space* $E \subset R^n$ *of a point* $\hat{x}$ *, that*

$$\forall x \in E \quad f(\hat{x}) \le f(x)$$

Additionaly if $f(\hat{x}) < f(x)$ for $x \ne \hat{x}$ then there exists strict local minimum.

The vector $\hat{x}$ is global minimum of the function *f(x) in the space R^n, if there exists such open space* R^n *of a point* $\hat{x}$ *, that*

$$\forall x \in R^n \quad f(\hat{x}) \le f(x)$$

Additionaly if $f(\hat{x}) < f(x)$ for $x \ne \hat{x}$ then there exists the strict global minimum in this point .

Nonlinear optimization problem without constraints:

Definition.

A direction d in the space R^n we denote any n-dimentional column vector.

Let us assume, that there exists a point $\mathbf{x} \in R^n$ and scalar $\tau \in [0;+\infty)$.

Any point $\mathbf{y} \in R^n$ belonging to half stright line, which starts in point x, in direction $d \ne 0$ will be determined as follows:

$$\mathbf{y} = \mathbf{x} + \tau\mathbf{d}$$

Lemma.

Let $f: \mathbf{X} = R^n \to R^1$ be a differentiable function in the point $\mathbf{x^0} \in \mathbf{X}$

Let us assume, that there exist a direction d, for which:

$$\left\langle \nabla f(x^0), d \right\rangle < 0,$$

Then there exists such $\sigma > 0$, that for all $\tau \in (0,\sigma]$ the following condition is fulfilled

$$f(x^0 + \tau d) < f(x^0).$$

Proof: let us use the property of Gateaux differential, with the directional derivative definition.

Nonlinear optimization problem without constraints:

Theorem. Let $f: \mathbf{X} = R^n \to R^1$ will be differentiable function. If $\hat{x} \in X$ minimize function *f(x) in the form:*

$$f(\hat{\mathbf{x}}) \le f(\mathbf{x}), \qquad \forall \mathbf{x} \in \mathbf{X},$$

then

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

Proof: not directly.

The point $\hat{x}$ is named as a stationary point.

Theorem. Let $f: X = R^n \to R^1$ will be convex and differentiable function. For the point $\hat{\mathbf{x}} \in X$ function *f(x)* constitutes minimal value of the function:

$$f(\hat{x}) \le f(x),$$

For each $x \in X$ if and only if , when $\hat{x}$ fulfills the following condition:

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

It is the necessary condition for local extremum *f(x)* in the point. $\hat{x}$

Global minimum of the function f(x)

Theorem:

If the function $f: X = R^n \to R^1$ will be strictly convex and differentiable, then the vector $\hat{\mathbf{x}} \in X$, which fulfills the necessary condition

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

is the only one global minimum of the function *f(x).*

Sufficiant conditions for the nonlinear optimization problem without constraints:

Function f(x) is continuous and twice differentiable. The function f(x) has the matrix with second derivatives named as hesjan A

Matrix A has some main subdeterminants $|A|_i$

$$|A|_1 = \left| \frac{\partial^2 f}{\partial x_1^2}(x) \right|$$

$$|A|_i = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \ldots\ldots & \frac{\partial^2 f}{\partial x_1 x_i}(x) \\ \ldots\ldots & & \ldots\ldots \\ \frac{\partial^2 f}{\partial x_i x_1}(x) & \ldots\ldots & \frac{\partial^2 f}{\partial x_i^2}(x) \end{vmatrix}$$

$$|A|_n = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \ldots & \frac{\partial^2 f}{\partial x_1 x_i}(x) & \ldots & \frac{\partial^2 f}{\partial x_1 x_n}(x) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{\partial^2 f}{\partial x_i x_1}(x) & \ldots & \frac{\partial^2 f}{\partial x_i^2}(x) & \ldots & \frac{\partial^2 f}{\partial x_i x_n}(x) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \frac{\partial^2 f}{\partial x_n x_1}(x) & \ldots & \frac{\partial^2 f}{\partial x_n x_i}(x) & \ldots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{vmatrix}$$

The stationary conditions for nonlinear programming problem without constraints

Let us assume, that $\hat{x}$ is a stationary point for a function *f(x)*.

Then the following relations are fulfilled:

1. If the hesjan matrix A is positive defined : $\left|A(\hat{x})\right|_i > 0 \quad for \quad i = 1,...,n$ then a function f(x) has local minimum in this point.

2. If the hesjan matrix A is negative defined : $(-1)^i \left|A(\hat{x})\right|_i > 0 \quad for \quad i = 1,...,n$ then a function f(x) has local maximum in this point.

3. If the hesjan is half-positive defined: $\left|A(\hat{x})\right|_i \geq 0 \quad for \quad i = 1,...,n-1 \quad and \quad \left|A(\hat{x})\right|_n = 0$
   or the hesjan is half-negative defined:
   $$(-1)^i \left|A(\hat{x})\right|_i \geq 0 \quad for \quad i = 1,...,n-1 \quad and \quad \left|A(\hat{x})\right|_n = 0$$
   In this moment it is impossible to decide what type of an extremum of a function *f(x is* in this point.

4. If two conditions 1 and 2 are not fulfilled (then the hesjan A is not determined) then the function *f(x) has not extremum in the point* $\hat{x}$.

Theorem.

If the function *f(x) is twice differentiable, then in each local minimum for the nonlinear programming problem without constraints the following optimality conditions are fulfilled:*

$\nabla f(\hat{x}) = 0$ I order condition

$d^T A(\hat{x}) d > 0$ for $\forall d \neq 0$ II order condition

$A = \nabla^2 f(\hat{x})$ Matrix *A* has to be strictly positive

Nonlinear programming problem with constraints NPP :

Let us find $\hat{\mathbf{x}}$ such, that:

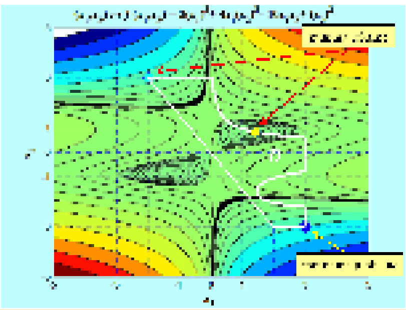
$$f(\hat{\mathbf{x}}) = \min_{\mathbf{x} \in X} f(\mathbf{x}),$$

where:

$$X = \{\mathbf{x} : g_i(\mathbf{x}) \leq 0, \ i = 1,...,m\},$$
$$f : X = R^n \to R^1$$
$$g_i : X = R^n \to R^1.$$

Example of NP problem

$$\min_{x \in X} f(x) = 2(x_1)^2 + 4x_1(x_2)^3 - 10x_2x_1 + (x_2)^2$$
$$X = \{(x_1, x_2) \mid x_1(x_2)^2(2.4 + x_2) \leq 3$$
$$\wedge\ 3/2x_1 + x_2 \geq 0$$
$$\wedge\ (-3 \leq x_i \leq 3, i = 1,2)\}.$$



Nonlinear programming problem with constraints NPP :

Definition.

The direction d starting in the point x is feasible, if there exists such $\sigma > 0$, that for any

$$\tau \in [0;\sigma], \mathbf{x} + \tau\mathbf{d} \in X \ ,$$

The set of all feasible directions:

$$D(x) = \{d : \exists \sigma > 0 \text{ such, that}$$
$$\tau \in [0;\sigma] \Rightarrow \mathbf{x} + \tau\mathbf{d} \in X \ \}.$$

Lagrange function

Definition.

The Lagrange function for nonlinear programming problem is a scalar function as follows:

$$L(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) + \langle \boldsymbol{\lambda}, g(\mathbf{x}) \rangle,$$

where $\boldsymbol{\lambda} \in R^m$ **is a vector of Lagrange multipliers.**

Active constraints:

1. $g_i(\hat{\mathbf{x}} + \tau\mathbf{d}) \le 0, \qquad \forall i \in A(\hat{\mathbf{x}})$

   where $\tau \in [0,\sigma]$

2. $g_i(\hat{\mathbf{x}} + \tau\mathbf{d}) = g_i(\hat{\mathbf{x}}) + \tau\langle \nabla g_i(\hat{\mathbf{x}}), \mathbf{d} \rangle + O(\tau) \le 0$

3. **The necessary condition for the feasibility of direction vector is:**

$$\langle \nabla g_i(\hat{\mathbf{x}}), \mathbf{d} \rangle \le 0, \qquad \forall i \in A(\hat{\mathbf{x}})$$

Farkas lemma

Let us define the set of n-dimensional vectors in R^n space

$\{b, a^i, i = 1,\ldots,m\}$.

The relation $\langle b, x \rangle \ge 0$

set up for each $x \in R^n$, **which fulfills**

$$\langle -a^i, x \rangle \ge 0,$$

If and only if, the vector λ **exists** $\lambda = [\lambda_1,\ldots,\lambda_m]^T \ge 0$

Such, that

$$b + \sum_{i=1}^{m} \lambda_1 a^i = 0.$$

The set of directions:

$$D_1(\hat{x}) = \left\{ d : < \nabla g_i(\hat{x}), d > \le 0, i \in A(\hat{x}), < \nabla f(\hat{x}), d > \ge 0 \right\}$$

$$D_2(\hat{x}) = \left\{ d : < \nabla g_i(\hat{x}), d > \le 0, i \in A(\hat{x}), < \nabla f(\hat{x}), d > < 0 \right\}$$

$$D_3(\hat{x}) = \left\{ d : < \nabla g_i(\hat{x}), d > > 0 \ \textit{for some} \ i \in A(\hat{x}) \right\}$$

$$D(\hat{x}) = D_1(\hat{x}) \cup D_2(\hat{x}) \cup D_3(\hat{x})$$

Kuhn-Tucker-Karusch necessary conditions:

Theorem: If

a)functions f i g_i **are differentiable;**

b) $\hat{x}$ **is a local minimum of NPP , there exists**

$\hat{\lambda} \ge 0, \dim \hat{\lambda} = m.$ **such, that**

$$\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^{m} \hat{\boldsymbol{\lambda}}_i \nabla g_i(\hat{\mathbf{x}}) = 0$$

$$\hat{\lambda}_i g_i(\hat{\mathbf{x}}) = 0, \quad \mathrm{i} = 1,\ldots,\mathrm{m},$$

if and only if

$$D_2(\hat{x}) = \phi$$

Proof: necessary conditions Kuhn-Tucker-Karusch

Proof. Let $b = \nabla f(\hat{x})$ **and**

$a^i = \nabla g_i(\hat{\mathbf{x}}), \qquad \forall i \in A(\hat{\mathbf{x}})$

Then according to the Farkas lemma, there exists $\hat{\lambda}_i \ge 0, \ i \in A(\hat{x})$

Such, that $\nabla f(\hat{\mathbf{x}}) = \sum_{i \in A(\hat{\mathbf{x}})} \hat{\lambda}_i (-\nabla g_i(\hat{\mathbf{x}}))$

when $\langle \nabla f(\hat{\mathbf{x}}), \mathbf{d} \rangle \ge 0, \quad \forall \mathbf{d} \in R^n,$

And the condition for the active constraints has to be fulfilled : $\langle \nabla g_i(\hat{\mathbf{x}}), \mathbf{d} \rangle \le 0, \qquad \forall i \in A(\hat{\mathbf{x}})$

if and only if $D_2(\hat{x}) = \phi$

Proof: necessary conditions Kuhn-Tucker-Karusch

For $i \notin A(\hat{\mathbf{x}})$ **let us assume** $\hat{\lambda}_i = 0$

So, two following equations are fulfilled:

$$\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^{m} \hat{\boldsymbol{\lambda}} \nabla g_i(\hat{\mathbf{x}}) = 0$$

$$\hat{\lambda}_i g_i(\hat{\mathbf{x}}) = 0, \quad \mathrm{i} = 1,\ldots,\mathrm{m},$$

Ckd

Kuhn-Tucker_Karush necessary conditions with the help of Langrangian function

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}).$$

Necessary conditions :

$$\nabla_x L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}) = 0,$$

$$\left\langle \hat{\boldsymbol{\lambda}}, \nabla_\lambda L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}) \right\rangle = 0$$

$$\nabla_\lambda L(\hat{\mathbf{x}}, \hat{\boldsymbol{\lambda}}) \le 0$$

$$\hat{\boldsymbol{\lambda}} \ge 0$$

Sufficient conditions defined as the regularity condition:

1. **All functions for the contraints $g_i(x)$ are linear– the Karlin regularity condition.**
2. **. All functions for the contraints $g_i(x)$ are convex and the feasible set of solutions is not empty - the Slater regularity condition.**
3. **The gradients of all active constraints,**

$$\nabla g_i(\hat{x}), \quad \textbf{for} \quad i \in A(\hat{x}),$$

are lineary independent–Fiacco and McCormick regularity condition.

Kuhn-Karusch-Tucker theorem – necessary and sufficient conditions for nonlinear programming problem with constraints (NPP)

Theorem. If

a)functions f and g_i **are differentiable;**

b) $\hat{x}$ **is local minimum of NPP,**

c) Regularity condition for constraints is fulfilled in the point $\hat{x}$

Then there exists $\hat{\boldsymbol{\lambda}} \ge 0$, dim $\hat{\boldsymbol{\lambda}} = m.$

Such, that in the point $\hat{\mathbf{x}}$ **the Kuhn-Karusch-Tucker necessary conditions are fulfilled:**

$$\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^{m} \hat{\boldsymbol{\lambda}}_i \nabla g_i(\hat{\mathbf{x}}) = 0$$

$$\hat{\lambda}_i g_i(\hat{\mathbf{x}}) = 0, \quad \mathrm{i} = 1,\ldots,\mathrm{m},$$

Example I

Taking under consideration the K-K-T conditions solve the following nonlinear optimization problem:

$$\min_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x}) = x_1^2 + x_1 * x_2 + 0.5 * x_2^2 - x_1 - x_2$$

$$X = \left\{ x : \begin{matrix} x_1 + x_2 \le 0 \\ x_2 \ge 2 \end{matrix} \right\}$$

The solutions of NPP problem:

$$\hat{\mathbf{x}} = [-2, 2]^T \qquad f(\hat{\mathbf{x}}) = 2$$

Nonlinear programming problem with constraints NPP with additional constraints for variables:

$$f(\hat{\mathbf{x}}) = \min_{\mathbf{x} \in X} f(\mathbf{x}),$$

where :

$$X = \{\mathbf{x} : g_i(\mathbf{x}) \le 0, \mathbf{x} \ge 0, \ i = 1,\ldots,m\},$$

$$f(x) : R^n \to R^1$$

and $g_i(x) : R^n \to R^1, i = 1,\ldots,m$

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}).$$

Lagrange conditions for NNP with x≥0

$$\nabla_x L\left(\hat{x},\hat{\lambda}\right) \ge 0 \qquad \nabla_\lambda L\left(\hat{x},\hat{\lambda}\right) \le 0$$

$$\left\langle \hat{\mathbf{x}}, \nabla_x L\left(\hat{x},\hat{\lambda}\right)\right\rangle = 0 \qquad \left\langle \hat{\lambda}, \nabla_\lambda L\left(\hat{x},\hat{\lambda}\right)\right\rangle = 0$$

$$\hat{\mathbf{x}} \ge 0 \qquad \hat{\lambda} \ge 0$$

Examples for Kuhn-Tucker-Karusch conditions

Example II

Taking under consideration the K-T-K conditions solve the following nonlinear optimization problem:

$$\min_{\mathbf{x}\in\mathbf{X}} f(\mathbf{x}) = 2*x_1^2 + x_1*x_2 + 0.5*x_2^2 - x_1 - x_2$$

$$X = \left\{ x : \begin{matrix} x_1 + x_2 \le 0 \\ x_2 \ge 0 \end{matrix} \right\}$$

The solutions of NPP problem:

$$\hat{\mathbf{x}} = [0, 0]^T \qquad f(\hat{\mathbf{x}}) = 0$$

Nonlinear programming problem with constraints NPP with additional equations as constraints :

$$f(\hat{x}) = \min_{x\in X} f(x)$$

where :

$$X = \{\mathbf{x} : g_i(\mathbf{x}) \le 0,\ i = 1,\ldots,p,\quad h_i(x) = 0, i = p+1,\ldots,m\},$$

Where:

$$f(x) : X = R^n \to R^1$$
$$g_i(x) : X = R^n \to R^1, i = 1,\ldots,p$$
$$h_i(x) : X = R^n \to R^1, i = p+1,\ldots,m$$

$$L(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}).$$

Theorem: If

a) Functions f and g_i **for *i=1,…,m* are differentiable;**
b) $\hat{x}$ **is local minimum of NPP,**

Then there exists: $\hat{\lambda}_i \ge 0,\ i = 1,\ldots,p$

and $\hat{\lambda}_i, i = p+1,\ldots,m$ without any constraints on variables,

such that in the point $\hat{x}$ the Kuhn-Karusch-Tucker necessary conditions are fulfilled:

$$\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^{m} \hat{\boldsymbol{\lambda}}_i \nabla g_i(\hat{\mathbf{x}}) = 0$$

$$\hat{\lambda}_i g_i(\hat{\mathbf{x}}) = 0, \quad \mathrm{i} = 1,\ldots, \mathrm{p}$$

Example III

Taking under consideration the K-K-T conditions solve the following nonlinear optimization problem:

$$\min_{\mathbf{x}\in\mathbf{X}} f(\mathbf{x}) = x_1^2 + x_1x_2 + x_2^2$$

$$X = \left\{ x : \begin{matrix} 2x_1 - x_2 \ge 4 \\ x_2 = 0 \end{matrix} \right\}$$

The solutions of NPP problem:

$$\hat{\mathbf{x}} = [2, 0]^T ; \hat{\lambda} = 2 \quad f(\hat{\mathbf{x}}) = 4$$