Probleme (NLP) generale

- Reamintim problema (NLP) generala:

$$(\mathrm{NLP}): \quad \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.l.: } g(x) \leq 0, \;\; h(x) = 0$$

- Functiile $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$ si $h : \mathbb{R}^n \to \mathbb{R}^p$ sunt functii diferentiabile de doua ori.

- Multimea fezabila a problemei (NLP):

$$X = \{x \in \mathbb{R}^n \;:\; g(x) \leq 0, \;\; h(x) = 0\}$$

- Deci problema se rescrie ca:

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

Exemplu (NLP)

- Fie urmatoarea problema de optimizare (NLP):

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & x_1^4 + x_2^4 - 4x_1x_2 \\ \text{s.l:} \quad & x \geq 0, x_1 + x_2 = 5 \end{aligned}$$

- **Observam**:
  - functia obiectiv: $f(x) = x_1^4 + x_2^4 - 4x_1x_2$, $f : \mathbb{R}^2 \to \mathbb{R}$, $n = 2$
  - constrangeri de inegalitate $g(x) = -x$, $g : \mathbb{R}^2 \to \mathbb{R}^2$, $m = 2$
  - constrangeri de egalitate $h(x) = x_1 + x_2 - 5$, $h : \mathbb{R}^2 \to \mathbb{R}$, $p = 1$
- Multimea fezabila in acest caz

$$X = \{x \in \mathbb{R}^2 : \; x \geq 0, \; x_1 + x_2 = 5\}$$

este convexa!

Conditii de ordinul I pentru (NLP) - constrangeri convexe

Teorema: Fie X o multime convexa si $f \in \mathcal{C}^1$. Pentru problema de optimizare constransa

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

avem urmatoarele conditii de optimalitate:

- daca x^* este minim local atunci:

$$\nabla f(x^*)^T (x - x^*) \geq 0 \quad \forall x \in X$$

- daca in plus f este functie convexa, atunci x^* este punct de minim daca si numai daca:

$$\nabla f(x^*)^T (x - x^*) \geq 0 \quad \forall x \in X$$

Punctele x^* ce satisfac inegalitatea anterioara $\Rightarrow$ puncte stationare pentru (NLP) avand constrangeri convexe.

Constrangeri (multimi) active/inactive

Definitie: O constrangere de inegalitate $g_i(x) \leq 0$ se numeste activa in punctul fezabil $x \in X$ daca si numai daca $g_i(x) = 0$, altfel se numeste inactiva. Orice constrangere de egalitate $h_i(x) = 0$ este activa intr-un punct fezabil.

Definitie: Multimea de indecsi $\mathcal{A}(x) \subseteq \{1, \dots, m\}$ corespunzatoare constrangerilor active se numeste multime activa in punctul $x \in X$.

Exemplu: Pentru multimea fezabila:

$$X = \{x \in \mathbb{R}^2 : \ x \geq 0, \ x_1 + x_2 = 5\}$$

multimea activa in punctul $x = [0\ 1]^T$ este $\mathcal{A}([0\ 1]^T) = \{1\}$.

Conditii de ordinul I pt. (NLP): constrangeri de egalitate

Consideram intai probleme (NLP) ce au doar constrangeri de egalitate:

$$\text{(NLPe)}: \quad \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.l.}: h(x) = 0$$

Pentru problema (NLPe) toate constrangerile sunt considerate constrangeri active.

Vom studia conditiile de optimalitate pentru clasa de probleme (NLPe), pentru ca apoi sa extindem rezultatele obtinute la clasa generala (NLP).

Notiuni teoretice

- O curba pe o suprafata $S \Rightarrow$ o multime de puncte $x(t) \in S$ continuu parametrizate in t, pentru $a \leq t \leq b$.
- O curba $x(t)$ este diferentiabila daca exista $\dot{x} = \frac{\mathrm{d}x(t)}{\mathrm{d}t}$ si de doua ori diferentiabila daca $\ddot{x}(t)$ exista.
- O curba $x(t)$ trece prin punctul x^* daca exista $t^* \in [a, b]$ a.i. $x^* = x(t^*)$.
- Planul tangent in $x^* \in S \Rightarrow$ multimea tuturor vectorilor $\dot{x}(t^*)$ definiti de $x(t) \in S$ avand proprietatea ca $x(t^*) = x^*$.
- Pentru o functie $h : \mathbb{R}^n \to \mathbb{R}^p$, cu $h(x) = [h_1(x) \dots h_p(x)]^T$ notam Jacobianul prin $\nabla h(x)$. Reamintim ca $\nabla h(x)$ este o matrice $p \times n$,avand elementul de pe pozitia (i, j) egal cu $\frac{\partial h_i(x)}{\partial x_j}$.

Puncte regulate

Definitie: Un punct x^* ce satisface constrangerea $h(x^*) = 0$ se numeste punct regulat daca gradientii $\nabla h_1(x^*), \ldots, \nabla h_p(x^*)$ sunt liniar independenti (Jocabianul are rangul p).

Teorema: Intr-un punct regulat x^* al suprafetei S definita de constrangerile de egalitate $h(x) = 0$, planul tangent este egal cu:

$$\begin{aligned} M &= \{d \in \mathbb{R}^n \ : \ \nabla h(x^*) d = 0\} \\ &= \{d \in \mathbb{R}^n \ : \ \nabla h_i(x^*)^T d = 0 \ \forall i = 1, \ldots, p\} \end{aligned}$$

Lema: Fie x^* un punct regulat al constrangerilor $h(x) = 0$ si punct de extrem local (minim sau maxim local) al problemei de optimizare (NLPe). Atunci, orice $d \in \mathbb{R}^n$ ce satisface:

$$\nabla h(x^*) d = 0$$

trebuie sa satisfaca si:

$$\nabla f(x^*)^T d = 0.$$

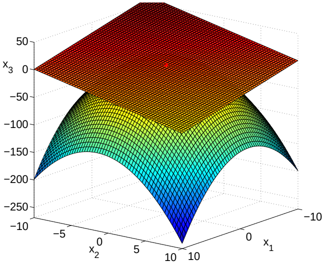
Exemplu

- Fie constrangerea data de functia $h : \mathbb{R}^3 \to \mathbb{R}$, $h(x) = x_1^2 + x_2^2 + 3x_1 + 3x_2 + x_3 - 1$. Consideram punctul $x^* = [0\ 0\ 1]^T$ pentru care $h(x^*) = 0$.
- Jacobianul lui $h(x)$ va fi:

$$\nabla h(x) = [2x_1 + 3\ \ 2x_2 + 3\ \ 1]$$

iar $\nabla h(x^*) = [3\ 3\ 1]$ are rang $1 \Rightarrow x^*$ punct regulat.

- Din teorema precedenta pentru caracterizarea planului tangent, rezulta ca orice directie tangenta $d = [d_1\ \ d_2\ \ d_3]^T$ va satisface $\nabla h(x^*) d = 0$, deci $3d_1 + 3d_2 + d_3 = 0$ (un plan)



Conditii necesare de ordinul I pentru (NLPe)

Teorema: Fie x^* un punct de extrem al functiei obiectiv f supusa la constrangerile $h(x) = 0$, i.e. al problemei de optimizare (NLPe), si presupunem ca x^* este un punct regulat pentru aceste constrangeri. Atunci, exista un multiplicator Lagrange $\mu^* \in \mathbb{R}^p$ a.i.:

$$(\textbf{KKT - NLPe}): \ \nabla f(x^*) + \nabla h(x^*)^T \mu^* = 0 \ \text{ si } \ h(x^*) = 0.$$

Punctele x^* pentru care exista μ^* a.i. conditiile (KKT-NLPe) sunt satisfacute $\Rightarrow$ puncte stationare pentru problema (NLPe) (puncte de minim, puncte de maxim sau puncte sa).

Conditiile (KKT-NLPe) nu reprezinta altceva decat conditiile de optimalitate de ordinul I pentru minimizarea Lagrangianului neconstrans $\mathcal{L}(x, \mu) = f(x) + \mu^T h(x)$:

$$\nabla_x \mathcal{L}(x, \mu) = 0 \ \text{ si } \ \nabla_\mu \mathcal{L}(x, \mu) = 0 \Longleftrightarrow \nabla \mathcal{L}(x, \mu) = 0$$

Exemplul 1

Fie problema (NLPe):

$$\min_{x\in\mathbb{R}^2} x_1 + x_2$$
$$\text{s.l.}: \; h(x) = x_1^2 + x_2^2 - 2 = 0$$

Putem observa ca orice punct fezabil este regulat. Drept urmare, orice punct de minim local satisface sistemul $\nabla\mathcal{L}(x,\mu) = 0$:

$$2\mu x_1 = -1$$
$$2\mu x_2 = -1$$
$$x_1^2 + x_2^2 = 2.$$

Solutiile (x_1^*, x_2^*, μ^*) ale sistemului sunt punctele stationare $(-1,-1,1/2)$ si $(1,1,-1/2)$. Sunt puncte de extrem cele doua solutii sau puncte sa

Exemplul 2

Consideram urmatoarea problema de optimizare:

$$\min_{x \in \mathbb{R}^3} 4x_1 - 2x_2 + x_3^2$$

$$\text{s.l. :} \;\; h_1(x) = x_1 - x_2 + x_3 = 0, \;\; h_2(x) = -x_1 + 2x_2 - x_3^2 = 0.$$

Jacobianul constrangerilor va fi de forma:

$$\nabla h(x) = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 2 & -2x_3 \end{bmatrix},$$

ce are rangul minim 2 $\Rightarrow$ orice punct fezabil este regulat.
Pentru a gasi punctele de extrem rezolvam sistemul dat de conditiile (KKT-NLPe):

$$4 + \mu_1 - \mu_2 = 0, \;\; -2 - \mu_1 + 2\mu_2 = 0, \;\; 2x_3 + \mu_1 - 2x_3\mu_2 = 0$$
$$x_1 - x_2 + x_3 = 0$$
$$-x_1 + 2x_2 - x_3^2 = 0.$$

Obtinem ca solutia unica a sistemului este data de $x^* = [-1 \; 0 \; 1]^T$ si $\mu_1^* = -6$, $\mu_2^* = -2$. Care este natura punctului x^*?

Exemplul 3

Consideram problema:

$$\min_{x\in\mathbb{R}^2} -x_1$$
$$\text{s.l.}: \; h_1(x) = (1-x_1)^3 + x_2 = 0, \; h_2(x) = (1-x_1)^3 - x_2 = 0.$$

Putem observa ca problema are un singur punct fezabil $x^* = [1 \; 0]^T \Rightarrow$ minim global.

De asemenea, avem $\nabla f(x^*) = [-1 \; 0]^T$, $\nabla h_1(x^*) = [0 \; 1]^T$ si $\nabla h_2(x^*) = [0 \; -1]^T \Rightarrow x^*$ NU este un punct regulat!

In acest caz, conditiile de optimalitate de ordinul I nu pot fi satisfacute $\Rightarrow \nexists \mu_1$ si μ_2 a.i.:

$$\mu_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \mu_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Deci conditia de regularitate este esentiala!

Conditii necesare de ordinul II pentru (NLPe)

Teorema: Presupunem ca x^* este un punct de minim local al problemei (NLPe) si un punct regulat pentru constrangerile aferente. Atunci exista $\mu^* \in \mathbb{R}^p$ astfel incat

$$\nabla f(x^*) + \nabla h(x^*)^T \mu^* = 0 \text{ si } h(x^*) = 0.$$

In plus, daca notam planul tangent in x^* prin $M = \{d \in \mathbb{R}^n \ : \ \nabla h(x^*) d = 0\}$, atunci matricea Hessiana a Lagrangianului in raport cu x

$$\nabla_x^2 \mathcal{L}(x^*, \mu^*) = \nabla^2 f(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla^2 h_i(x^*)$$

este pozitiv semidefinita pe planul tangent M, adica

$$d^T \nabla_x^2 \mathcal{L}(x^*, \mu^*) d \geq 0 \quad \forall d \in M.$$

Exemplu

Consideram din nou problema:

$$\min_{x\in\mathbb{R}^2:\ h(x)=x_1^2+x_2^2-2=0} x_1 + x_2,$$

pentru care matricea Hessiana a Lagrangianului in variabila x este:

$$\nabla_x^2 \mathcal{L}(x,\mu) = \nabla^2 f(x) + \mu \nabla^2 h(x) = \mu \left[\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array}\right].$$

Tinand cont ca planul tangent la un punct $x \neq 0$ este de forma $M = \{d:\ [2x_1\ 2x_2]d = 0\}$, pentru care putem alege o baza $D(x) = [-x_2\ x_1]^T$ avem:

$$D(x)^T \nabla_x^2 \mathcal{L}(x,\mu) D(x) = 2\mu(x_1^2 + x_2^2).$$

Pentru prima solutie a sistemului $\nabla \mathcal{L}(x,\mu) = 0$ avem $d_1^T \nabla_x^2 \mathcal{L}(-1,-1,1/2) d_1 = 2 > 0$, unde $d_1 = D(-1,-1) \Rightarrow$ posibil punct de minim local!
Pentru cea de-a doua solutie $d_2^T \nabla_x^2 \mathcal{L}(1,1,-1/2) d_2 = -2 < 0$, cu $d_2 = D(1,1) \Rightarrow$ acesata solutie nu este minim local!

Conditii suficiente de ordinul II pentru (NLPe)

Teorema: Presupunem un punct $x^* \in \mathbb{R}^n$ si un $\mu^* \in \mathbb{R}^p$ a.i.:

$$\nabla f(x^*) + \nabla h(x^*)^T \mu^* = 0 \text{ si } h(x^*) = 0.$$

Presupunem de asemenea ca matricea data de Hessiana Lagrangianului $\nabla_x^2 \mathcal{L}(x^*, \mu^*) = \nabla^2 f(x^*) + \sum_{i=1}^p \mu_i^* \nabla^2 h_i(x^*)$ este pozitiv definita pe planul tangent $M = \{d : \nabla h(x^*) d = 0\}$, adica

$$d^T \nabla_x^2 \mathcal{L}(x^*, \mu^*) d > 0 \quad \forall d \in M \setminus 0.$$

Atunci x^* este punct de minim local strict al problemei (NLPe).

Exemplu

Fie problema:

$$\min_{x \in \mathbb{R}^3 :\ x_1+x_2+x_3=3} -x_1x_2 - x_1x_3 - x_2x_3,$$

pentru care conditiile de optimalitate de ordinul I conduc la urmatorul sistem liniar:

$$\begin{aligned} -x_2 - x_3 + \mu &= 0 \\ -x_1 - x_3 + \mu &= 0 \\ -x_1 - x_2 + \mu &= 0 \\ x_1 + x_2 + x_3 &= 3, \end{aligned}$$

avand solutia unica $x_1^* = x_2^* = x_3^* = 1$ si $\mu^* = 2$.
Hessiana Lagrangianului va avea urmatoarea forma:

$$\nabla_x^2 \mathcal{L}(x, \mu) = \nabla^2 f(x) = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

Exemplu - continuare

Planul tangent intr-un punct x fezabil va fi de forma $M = \{d : \ [1\ 1\ 1]d = 0\} \Rightarrow d_1 = -d_2 - d_3 \Rightarrow$ alegand $[d_1\ d_2] = [1\ \ -1]$ si $[d_1\ d_2] = [-1\ \ -1] \Rightarrow$ o baza a planului tangent la suprafata definita de constrangerea $h(x) = 0$ va fi:

$$D(x) = \begin{bmatrix} 0 & 2 \\ 1 & -1 \\ -1 & -1 \end{bmatrix}$$

Drept urmare, obtinem ca x^* este minim strict local intrucat:

$$D(x^*)^T \nabla_x^2 \mathcal{L}(x^*, \mu^*) D(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \succ 0$$

Este Hessiana functiei obiectiv evaluata in x^* pozitiv definita?

Conditii de ordinul I pentru (NLP) generale

In continuare, vom extinde conditiile de optimalitate derivate anterior la cazul general:

$$\begin{aligned}(\text{NLP}): \min_{x \in \mathbb{R}^n} & \; f(x) \\ \text{s.l.}: & \; g(x) \leq 0, \; h(x) = 0.\end{aligned}$$

Definitie: Fie un punct x^* ce satisface constrangerile problemei (NLP), adica $h(x^*) = 0$, $g(x^*) \leq 0$ si $\mathcal{A}(x^*)$ multimea constrangerilor active. Numim punctul x^* punct regulat, daca gradientii functiilor de constrangere, $\nabla h_i(x^*)$ pentru $i = 1, \ldots, p$ si $\nabla g_j(x^*)$ pentru $j \in \mathcal{A}(x^*)$ sunt liniari independenti.

Planul tangent in punctul regulat x^* pentru problema generala (NLP) este planul tangent corespunzator pentru constrangerile active:

$$M = \{d : \nabla g_j(x^*)^T d = 0 \, \forall j \in \mathcal{A}(x^*), \; \nabla h_i(x^*)^T d = 0 \, \forall i = 1, \ldots, p\}.$$

Conditii necesare de ordinul I pentru (NLP)

Teorema: Fie x^* un punct de minim local pentru problema (NLP) generala si presupunem ca x^* este si regulat. Atunci, exista un vector $\lambda^* \in \mathbb{R}^m$ si un vector $\mu^* \in \mathbb{R}^p$ a.i. conditiile Karush-Kuhn-Tucker (KKT) au loc:

$$
\begin{aligned}
(KKT): \quad \nabla f(x^*) + \nabla h(x^*)^T \mu^* + \nabla g(x^*)^T \lambda^* &= 0 \\
g(x^*)^T \lambda^* &= 0 \\
g(x^*) \leq 0, \;\; h(x^*) &= 0 \\
\mu^* \in \mathbb{R}^p, \;\; \lambda^* &\geq 0
\end{aligned}
$$

Conditiile (KKT) au fost formulate pentru prima data de catre W. Karush in teza sa de master (1939) si publicate abia 12 ani mai tarziu in cartea publicata de catre W. Kuhn si W. Tucker!

Demonstratie

Intrucat $\lambda^* \geq 0$ si $g(x^*) \leq 0$, cea de-a doua relatie implica faptul ca λ_i^* poate fi nenula doar daca $g_i(x^*)$ este activa. Drept urmare, $g_i(x^*) < 0$ implica $\lambda_i^* = 0$, iar $\lambda_i^* > 0$ implica $g_i(x^*) = 0$.

Mai departe, tinand cont ca x^* este un punct de minim local pentru problema (NLP). Notand cu X multimea constrangerilor problemei (NLP), atunci x^* este un punct de minim local si pentru problema avand multimea de constrangeri o submultime a lui X si definita prin setarea constrangerilor active la zero.

Astfel, pentru problema (NLPe) ce ar rezulta, definita pentru o vecinatate a lui x^*, exista multiplicatori Lagrange si deci prima relatie este satisfacuta pentru $\lambda_i^* = 0$ daca $g_i(x^*) \neq 0$ si drept urmare si cea de-a doua relatie este satisfacuta.

Demonstratie - continuare

Mai trebuie aratat ca $\lambda_i^* \geq 0$ pentru constrangerile active $g_i(x^*) = 0$. Fie o componenta $\lambda_k^* < 0$ si S_k, respectiv M_k suprafata si planul tangent in x^* definit de toate constrangerile active, mai putin $g_k(x^*) = 0$.

Tinand cont ca x^* este un punct regulat, atunci $\exists d \in M_k$ a.i. $\nabla g_k(x^*) d < 0$.

Luand in continuare o curba $x(t) \in S_k$, trecand prin x^* la $t = 0$ si avand $\dot{x}(0) = d$, atunci, pentru un $t \geq 0$ suficient de mic $x(t)$ este fezabil iar din prima relatie a conditiilor (KKT) obtinem:

$$\frac{\mathrm{d}f(x(t))}{\mathrm{d}t}\Big|_{t=0} = \nabla f(x^*) d < 0$$

ceea ce contrazice ipoteza de minimalitate a lui x^*.

Interpretarea conditiilor (KKT)

- Solutiile x^* ce satisfac conditiile (KKT) se numesc puncte stationare (pot fi puncte de minim, puncte de maxim sau puncte sa) pentru problema (NLP) generala.
- Prima relatie a conditiilor (KKT) semnifica faptul ca x^* este punct stationar pentru functia Lagrange $\min_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda^*, \mu^*)$, reprezentand conditiile de optimalitate de ordinul I, adica $\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0$.
- Cea de-a doua relatie reprezinta conditiile de complementaritate: $g(x^*)^T \lambda^* = 0 \Longleftrightarrow g_i(x^*)\lambda_i^* = 0 \quad \forall i$.
- Ultimele doua relatii exprima fezabilitatea primala $(g(x^*) \leq 0,\ h(x^*) = 0)$, respectiv duala $(\lambda^* \geq 0)$.

Exemplul 1

Consideram problema:

$$\min_{x \in \mathbb{R}^3} 0.5\left(x_1^2 + x_2^2 + x_3^2\right)$$
$$\text{s.l.} g_1(x) = x_1 + x_2 + x_3 + 3 \leq 0,\ \ g_2(x) = x_1 \leq 0.$$

Observam ca orice punct fezabil este si punct regulat, iar din prima relatie a conditiilor (KKT) obtinem:

$$\begin{aligned} x_1^* + \lambda_1^* + \lambda_2^* &= 0 \\ x_2^* + \lambda_1^* &= 0 \\ x_3^* + \lambda_1^* &= 0. \end{aligned}$$

Pentru a calcula punctele (KKT), analizam conditia de complementaritate pentru care distingem patru cazuri:

- 1. g_1 si g_2 sunt ambele inactive;
- 2. g_1 inactiva si g_2 activa;
- 3. g_1 activa si g_2 inactiva;
- 4. g_1 si g_2 sunt ambele active.

Exemplul 1 - continuare

Analiza cazurilor:

- 1. $\Rightarrow x_1^* + x_2^* + x_3^* < -3$ si $x_1^* < 0 \Rightarrow \lambda_1^* = \lambda_2^* = 0 \Rightarrow$ $x_1^* = x_2^* = x_3^* = 0 \Rightarrow$ contradictie cu ipoteza;
- 2. $\Rightarrow x_1^* + x_2^* + x_3^* < -3$, $x_1^* = 0$, $\lambda_1^* = 0$ si $\lambda_2^* \geq 0 \Rightarrow$ $\lambda_1^* = -\lambda_2^* = 0 \Rightarrow x_2^* = x_3^* = 0 \Rightarrow$ contradictie cu ipoteza;
- 3. $\Rightarrow x_1^* + x_2^* + x_3^* = -3$, $x_1^* < 0$, $\lambda_1^* \geq 0$ si $\lambda_2^* = 0 \Rightarrow$ putem alege $x_1^* = x_2^* = x_3^* = -1$ si $\lambda_1^* = 1$ ce satisfac conditiile (KKT) $\Rightarrow$ aceasta solutia este un punct (KKT);
- 4. $\Rightarrow x_1^* + x_2^* + x_3^* = -3$, $x_1^* = 0$ si $\lambda_1^*, \lambda_2^* \geq 0 \Rightarrow$ $x_2^* = x_3^* = -3/2$ si $\lambda_1^* = -\lambda_2^* = 3/2 \Rightarrow$ contrazice conditia $\lambda_2^* \geq 0$.

Exemplul 2

Consideram problema:

$$\min_{x \in \mathbb{R}^2} 2x_1^2 + 2x_1 x_2 + x_2^2 - 10x_1 - 10x_2$$
$$\text{s.l.}: \; x_1^2 + x_2^2 \leq 5, \; 3x_1 + x_2 \leq 6.$$

Pentru acesata problema conditiile (KKT) sunt:

$$4x_1 + 2x_2 - 10 + 2\lambda_1 x_1 + 3\lambda_2 = 0, \; 2x_1 + 2x_2 - 10 + 2\lambda_1 x_2 + \lambda_2 = 0$$
$$\lambda_1(x_1^2 + x_2^2 - 5) = 0, \; \lambda_2(3x_1 + x_2 - 6) = 0$$
$$x_1^2 + x_2^2 \leq 5, \; 3x_1 + x_2 \leq 6, \; \lambda_1, \lambda_2 \geq 0.$$

Pentru acest exemplu putem considera doua constrangeri active, una sau niciuna. Presupunem prima constrangere activa iar cea de-a doua inactiva, rezultand un sistem de trei ecuatii corespunzator, avand solutia $x_1^* = 1, x_2^* = 2, \lambda_1^* = 1$ si $\lambda_2^* = 0$. Solutia verifica $3x_1 + x_2 \leq 6$ si $\lambda_1 \geq 1 \Rightarrow$ satisface conditiile (KKT).

Conditii necesare de ordinul II pentru (NLP)

Planul tangent in punctul regulat x^* pentru problema generala (NLP) este planul tangent corespunzator pentru constrangerile active:

$$M = \{d : \nabla g_j(x^*)^T d = 0 \, \forall j \in \mathcal{A}(x^*), \ \nabla h_i(x^*)^T d = 0 \, \forall i = 1, \ldots, p\}.$$

Teorema: Fie f, g si h functii continuu diferentiabile de doua ori si x^* un punct regulat pentru constrangerile din problema (NLP) generala. Daca x^* este un punct de minim local pentru problema (NLP), atunci exista $\lambda^* \in \mathbb{R}^m$ si $\mu^* \in \mathbb{R}^p$ a.i. conditiile (KKT) sunt satisfacute si in plus Hessiana Lagrangianului in raport cu x:

$$\nabla_x^2 \mathcal{L}(x^*, \lambda^*, \mu^*) = \nabla^2 f(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla^2 h_i(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla^2 g_i(x^*)$$

este pozitiv semidefinita pe subspatiul tangent al constrangerilor active in x^*, adica:

$$d^T \nabla_x^2 \mathcal{L}(x^*, \lambda^*, \mu^*) d \geq 0 \quad \forall d \in M.$$

Conditii suficiente de ordinul II pentru (NLP)

Teorema: Fie f, g si h functii continuu diferentiabile de doua ori. Fie de asemenea:

- un punct regulat $x^* \in \mathbb{R}^n$
- variabilele duale $\lambda^* \in \mathbb{R}^m$ si $\mu^* \in \mathbb{R}^p$ pentru care conditiile (KKT) sunt satisfacute
- nu avem constrangeri de inegalitate degenerate, adica $\lambda_j^* > 0$ pentru orice $j \in \mathcal{A}(x^*)$
- Hessiana Lagrangianului $\nabla_x^2 \mathcal{L}(x^*, \lambda^*, \mu^*)$ este pozitiv definita pe planul tangent M

Atunci x^* este un punct de minim local strict pentru problema (NLP) generala.

Exemplu

Fie problema:

$$\min_{x \in \mathbb{R}^2:\ x_1^2 + x_2^2 - 1 \leq 0} x_2.$$

Observam ca punctul de minim global al acestei probleme este $x^* = [0\ \ -1]^T$. Aratam in continuare ca acesta este si minim strict. Prima conditie (KKT) are forma:

$$2\lambda x_1 = 0,\ 1 + 2\lambda x_2 = 0,$$

din care obtinem $\lambda > 0 \Rightarrow$ constrangerea este activa.
Adaugand aceasta conditie la cele doua anterioare, obtinem solutia $x^* = [0\ \ -1]^T$ si $\lambda^* = 1/2$.
Planul tangent in x^* este dat de $\{d :\ [0\ 2]d = 0\} = \{d :\ d_2 = 0\}$.
De asemenea, Hessiana Lagrangianului $\nabla_x^2 \mathcal{L}(x^*, \lambda^*) = 2\lambda^* I_2$ este pozitiv definita $\Rightarrow x^*$ punct de minim strict.

Conditii suficiente de ordinul I pentru probleme convexe

Teorema: Fie o problema convexa (CP) de forma:

$$(\mathrm{CP}): \quad f^* = \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.l.}: g(x) \leq 0, \; Ax = b,$$

in care functiile f si $g_1, \dots, g_m$ sunt functii convexe.
Daca urmatoarele conditii (KKT) sunt satisfacute:

$$(KKT - CP): \quad \nabla f(x^*) + \nabla g(x^*)^T \lambda^* + A^T \mu^* = 0$$
$$g(x^*)^T \lambda^* = 0$$
$$g(x^*) \leq 0, \; Ax^* = b$$
$$\mu^* \in \mathbb{R}^p, \; \lambda^* \geq 0,$$

atunci x^* este punct de minim global pentru problema convexa (CP), (λ^*, μ^*) este punct de maxim global pentru problema duala si dualitatea puternica are loc, adica $f^* = q^*$.

Exemplul 1

Consideram problema proiectiei originii $x_0 = 0$ pe subspatiul $X = \{x \in \mathbb{R}^n : \ Ax = b\}$, unde $A \in \mathbb{R}^{p \times n}$ de rang $p < n$.

Problema de optimizare convexa (CP):

$$\min_{x:\ Ax=b} \|x\|^2,$$

pentru care avem conditiile (KKT-CP):

$$x^* + A^T \mu^* = 0,\ Ax^* = b.$$

Obtinem deci $-AA^T\mu^* = b$ si intrucat AA^T este inversabila $\Rightarrow$ $x^* = A^T(AA^T)^{-1}b = A^+b$, unde A^+ este pseudoinversa lui A.

Exemplul 1 - cazul general

Extindem problema anterioara la cazul functiilor patratice generale:

$$\min_{x:\ Ax=b} \frac{1}{2}x^T Qx + q^T x,$$

unde $Q \succ 0$.

Conditiile (KKT-CP):

$$Qx^* + q + A^T\mu^* = 0,\ Ax^* = b.$$

Obtinem astfel ca solutiile primale si duale au urmatoarele forme:

$$\mu^* = -(AQ^{-1}A^T)^{-1}[AQ^{-1}q + b]$$
$$x^* = -Q^{-1}A^T\mu^* - Q^{-1}q.$$

Exemplul 2

Consideram urmatoarea problema convexa:

$$\min_{x \in \mathbb{R}^2} (x_1 - 5)^2 + (x_2 - 5)^2$$
$$\text{s.l.}: \; g_1(x) = x_1^2 + x_2^2 - 5 \leq 0, \; g_2(x) = x_1 + 2x_2 - 4 \leq 0$$
$$g_3(x) = -x_1 \leq 0, \; g_4(x) = -x_4 \leq 0.$$

Presupunem ca doar $g_1(x) \leq 0$ si $g_2(x) \leq 0$ sunt active $\Rightarrow$ $\lambda_3^* = \lambda_4^* = 0 \Rightarrow$ conditiile (KKT-CP) se reduc la urmatorul sistem:

$$\begin{bmatrix} 2(x_1^* - 5) \\ 2(x_2^* - 5) \end{bmatrix} + \begin{bmatrix} 2x_1^* & \frac{1}{2} \\ 2x_2^* & 1 \end{bmatrix} \begin{bmatrix} \lambda_1^* \\ \lambda_2^* \end{bmatrix} = 0$$
$$(x_1^*)^2 + (x_2^*)^2 - 5 = 0, \; x_1^* + 2x_2^* - 4 = 0.$$

Obtinem solutia $x^* = [2 \; 1]^T \Rightarrow x^*$ punct de minim global.