Exemplu subiect:

1. Fie vectorii $a = [2\ \ 2]^T, b = [-1\ \ -1]^T$ şi funcţia $f : \mathbb{R}^2 \to \mathbb{R}$ definită de

$$f(x) = \log\left(e^{a^Tx} + e^{b^Tx}\right).$$

a. Este funcţia f convexă ?

b. Determinaţi punctele de optim şi natura lor corespunzătoare funcţiei f?

2. Fie funcţia $f : \mathbb{R}^2 \to \mathbb{R}$ definită de

$$f(x) = x_1^4 + x_1x_2 + (1 + x_2)^2.$$

a. Calculaţi prima iteraţie z_1 a metodei gradient cu pas constant $\alpha = 0.1$ pornind din $z_0 = [1\ \ 1]^T$.

b. Iteraţia z_1 de la punctul a) face o descreştere mai mare decât prima iteraţie a metodei Newton cu pas $\alpha = 1$?

c. Considerand multimea fezabila $X = \{x \in \mathbb{R}^2 | 0 \le x_1 \le 1, 1 \le x_2 \le 2\}$, aplicati primul pas al metodei gradient proiectat pe problema

$$\begin{aligned} \min_{x\in\mathbb{R}^2}\ & f(x) \\ \text{s.l.}\ & x \in X, \end{aligned}$$

pornind din $z_0 = [1\ \ 0]^T$ cu pas $\alpha = 0.1$.

3. Fie problema de optimizare constrânsă:

$$\begin{aligned} \min_{x\in\mathbb{R}^2}\ & -2x_1 + x_2 \\ \text{s.l.}\ & x_1^2 - x_2 \le 0, \\ & x_2 - 4 \le 0 \end{aligned}$$

a. Determinaţi problema duală

b. Determinaţi punctele KKT şi natura acestora

1 Rezolvari

1. a. Notand $c = b - a$, avem:

$$\begin{aligned} f(x) = \log\left(e^{a^Tx} + e^{b^Tx}\right) &= \log\left(e^{a^Tx}\right) + \log\left(1 + e^{(b-a)^Tx}\right) \\ &= a^Tx + \log\left(1 + e^{c^Tx}\right) \end{aligned}$$

Deoarece primul termen este liniar, investigam convexitatea celui de al doilea termen din suma. Deoarece functia f este diferentiabila, folosim criteriul de convexitate de ordinul II:

$$\begin{aligned} \nabla^2 f(x) &= \frac{e^{c^Tx} \cdot c \cdot c^T(1 + e^{c^Tx}) - e^{2c^Tx}cc^T}{(1 + e^{c^Tx})^2} \\ &= \left[\frac{e^{c^Tx}}{1 + e^{c^Tx}} - \left(\frac{e^{c^Tx}}{1 + e^{c^Tx}}\right)^2\right] cc^T \succeq 0 \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

Concluzionam ca $\nabla^2 f(x) \succeq 0 \quad \forall x \in \mathbb{R}^n$ ce implica ca f este convexa pe intreg domeniul.

b. Gradientul functiei f este dat de:

$$\nabla f(x) = a + \nabla \left[\log\left(1 + e^{c^T x}\right)\right] = a + \frac{e^{c^T x}}{1 + e^{c^T x}} c$$

Pentru determinarea punctelor de extrem aplicam conditiile suficiente de ordin I:

$$a + \frac{e^{c^T x}}{1 + e^{c^T x}} c = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \frac{e^{-3x_1 - 3x_2}}{1 + e^{-3x_1 - 3x_2}} \begin{bmatrix} -3 \\ -3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Sistemul se reduce la:

$$2 - \frac{3e^{-3x_1 - 3x_2}}{1 + e^{-3x_1 - 3x_2}} = 0 \quad \Leftrightarrow \quad e^{-3x_1 - 3x_2} = 2 \quad \Leftrightarrow \quad 3x_1 + 3x_2 = -\ln(2)$$

Concluzie: multimea optima $X^* = \{x | 3x_1 + 3x_2 = -\ln(2)\}$; deoarece f este convexa, punctele de extrem din X^* sunt puncte de minim global.

2. a. Gradientul functiei f este:

$$\nabla f(x) = \begin{bmatrix} 4x_1^3 + x_2 \\ x_1 + 2(1 + x_2). \end{bmatrix}$$

Din expresia Metodei Gradient cu pas constant avem:

$$z_1 = z_0 - \alpha \cdot \nabla f(z_0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 0.1 \begin{bmatrix} 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

b. Din expresia metodei Newton:

$$z_1^N = z_0 - \alpha \cdot (\nabla^2 f(z_0))^{-1} \nabla f(z_0)$$

Deducem usor ca $\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 & 1 \\ 1 & 2 \end{bmatrix}$ si mai mult $\nabla^2 f(z_0) = \begin{bmatrix} 12 & 1 \\ 1 & 2 \end{bmatrix}$. De aceea inversa Hessianei rezulta $(\nabla^2 f(z_0))^{-1} = \begin{bmatrix} 2/23 & -1/23 \\ -1/23 & 12/23 \end{bmatrix}$. In final prima iteratie Metodei Newton este:

$$z_1^N = z_0 - \alpha \cdot (\nabla^2 f(z_0))^{-1} \nabla f(z_0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2/23 & -1/23 \\ -1/23 & 12/23 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 18/23 \\ -32/23 \end{bmatrix}$$

In final, observam ca $f(z_1) = 2.56 < f(z_1^N) = 5.0046$.

c. Pentru iteratie de gradient proiectat sunt necesare 2 etape: (i) calculul pasului de gradient; (ii) calculul proiectiei pe multimea fezabila. Pentru prima etapa avem:

$$y = z_0 - \alpha \cdot \nabla f(z_0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 0.1 \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 0.6 \\ -0.3 \end{bmatrix}.$$

Pentru a doua etapa, calculul proiectiei se rezuma la rezolvarea problemei:

$$\min_{z \in \mathbb{R}^2} \|z - y\|_2^2 = (z_1 - y_1)^2 + (z_2 - y_2)^2$$
$$\text{s.l.} \;\; 0 \leq z_1 \leq 1, \quad 1 \leq z_2 \leq 2.$$

Se observa ca problema este separabila in doua probleme scalare independente: $\min\limits_{z_1 \in \mathbb{R}} (z_1 - y_1)^2$ s.l. $0 \leq z_1 \leq 1$ si $\min\limits_{z_2 \in \mathbb{R}} (z_2 - y_2)^2$ s.l. $1 \leq z_2 \leq 2$. Solutiile problemelor scalare sunt date de proiectia lui y_1, respectiv y_2 pe intervalul $[0, 1]$, respectiv $[1, 2]$. De aceea rezulta $z_1^* = 0.6$ si $z_2^* = 1$. Concluzionam:

$$[y]_X = [\, z_0 - \alpha \cdot \nabla f(z_0) \,]_X = \begin{bmatrix} 0.6 \\ 1 \end{bmatrix}.$$

3.a. Determinam functia Lagrangian asociata problemei:

$$\mathcal{L}(x,\lambda) = -2x_1 + x_2 + \lambda_1(x_1^2 - x_2) + \lambda_2(x_2 - 4),$$

mentinand $\lambda_1 \geq 0, \lambda_2 \geq 0$. Din definitia functiei duale avem: $q(\lambda) = \min_{x\in\mathbb{R}^2} \quad \mathcal{L}(x,\lambda) = \min\limits_{x\in\mathbb{R}^2} \ \lambda_1 x_1^2 - 2x_1 + x_2(1 - \lambda_1 + \lambda_2) - 4\lambda_2$. Deoarece functia Lagrangian este compusa dintr-o functie patratica scalara (in coordonata x_1) si respectiv o functie liniara (in coordonata x_2), rezolvarea problemei de minimizare presupune utilizarea conditiilor de optimalitate de ordin I pentru coordonata x_1 si respectiv, asigurarea egalitatii $1 - \lambda_1 + \lambda_2 = 0$ pentru coordonata x_2 (pas tipic pentru optimizarea liniara). De aceea, din conditiile de ord. I rezulta:

$$2\lambda_1 x_1(\lambda) - 2 = 0 \Rightarrow x_1(\lambda) = \frac{1}{\lambda_1}.$$

Utilizand constrangerea impusa variabilelor duale: $1 - \lambda_1 + \lambda_2 = 0$, in final rezulta problema duala:

$$\begin{aligned}
&\max_{\lambda\in\mathbb{R}^2} \ -\frac{1}{\lambda_1} - 4\lambda_2\\
&\text{s.l. } \ 1 - \lambda_1 + \lambda_2 = 0, \lambda_1 \geq 0, \lambda_2 \geq 0.
\end{aligned}$$

b. In cazul curent, sistemul KKT este dat de:

$$\begin{cases}
\begin{bmatrix} -2 \\ 1 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2x_1 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\\
\lambda_1(x_1^2 - x_2) = 0\\
\lambda_2(x_2 - 4) = 0\\
x_1^2 \leq x_2, x_2 \leq 4, \lambda \geq 0
\end{cases}$$

Din rezolvarea egalitatilor de optimalitate rezulta: $x_1 = \frac{1}{\lambda_1}$ si respectiv, $\lambda_1 = 1 + \lambda_2$. Ultima egalitate implica $\lambda_1 > 0$ is deci $x_1^2 = x_2 = \frac{1}{\lambda_1^2}$.

Raman 2 cazuri de verificat: $\lambda_2 = 0$ si $\lambda_2 > 0$. Observam ca din $\lambda_2 > 0$ rezulta $x_2 = 4$, care implica $\lambda_1 = \frac{1}{2}$ (contradictie cu egalitatea $\lambda_1 = 1 + \lambda_2$!). In final, ramane $\lambda_2 = 0$ care implica $\lambda_1 = 1$ si deci $x_1 = x_2 = 1$. Punct KKT: $(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix})$.

Pentru a verifica natura acestuia, determinam $\nabla_x^2 \mathcal{L}(x,\lambda) = \begin{bmatrix} 2\lambda_1 & 0 \\ 0 & 0 \end{bmatrix}$ care implica $\nabla_x^2 \mathcal{L}(x^*,\lambda^*) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$.

Separat, observam ca pentru punctul KKT calculat anterior, constrangerea $x_1^2 \leq x_2$ este activa (adica $(x_1^*)^2 = x_2^*$), iar constrangerea $x_2 \leq 4$ este inactiva (adica $x_2^* < 4$). De aceea, planul tagent la suprafata solutiilor este dat de $M(x^*) = \left\{d \in \mathbb{R}^2 | \ [\, 2x_1^* \ \ -1 \,] \cdot d = 0\right\} = \left\{d \in \mathbb{R}^2 \mid 2d_1 = d_2\right\}$. Se observa imediat ca $d^T \nabla_x^2 \mathcal{L}(x^*,\lambda^*) d = 2d_1^2 > 0$ pentru orice $d \in M(x^*)$ nenul. Din ultima relatie tragem concluzia ca punctul KKT calculat este cel putin punct de minim local. Mai departe, observand ca problema de optimizare primala este convexa, se deduce ca punctul KKT contine, de fapt, punct de optim (minim) global.