TEHNICI DE OPTIMIZARE
Curs 1

Andrei Pătraşcu

Departament Informatică
Universitatea din Bucureşti

Orar:

- În săptămânile impare se face seminar
- În săptămânile pare se face laborator

Punctaj:

- Examen final scris - 50%
- Teme laborator - 40% (**Prezenţă obligatorie!**)
- Activitate (şi prezenţă) seminar - 10%
- **Promovare**: nota minim 5 la examen final, nota minim 5 la temele de laborator

- B. Polyak, Introduction to Optimization, Optimization Software Inc., New York, 1987
- D. Bertsekas, Nonlinear Programming, Third Edition. Athena Scientic, 2016.
- Y. Nesterov, Introductory Lectures on Convex Optimization, Kluwer, 2004.
- Hiriart-Urruty, Jean-Baptiste, and Claude Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals.* Vol. 305. Springer science & business media, 1996.

$\approx$ **850** î. Hr. Problema reginei Dido (a Cartaginei): Care este aria maximă a unei suprafeţe de pământ delimitate cu o sfoară de lungime fixă?

outschool.com

$\approx$ **850** î. Hr. Problema reginei Dido (a Cartaginei): Care este aria maximă a unei suprafeţe de pământ delimitate cu o sfoară de lungime fixă?

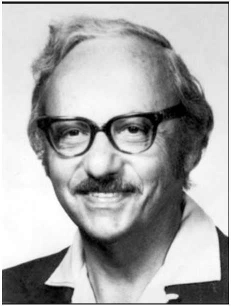
$$\max_{curba} \textbf{Arie}(curba) \ \text{ supus la } \ \textbf{Lungime}(curba) = ct.$$

- în plan se cunoaşte: $4\pi A \leq L^2$ pentru toate curbele închise
- egalitate doar pentru cerc.

sec. XVII-XVIII Bernoulli, Leibniz, Newton, L'Hopital publică soluţii ale câtorva probleme de calcul modelate prin căutarea de maxime/minime:

- forma optimă a unei nave cu rezistenţă minimă (Newton)
- curba celei mai rapide coborâri (Bernoulli)
- calcul variaţional (Euler)
- sisteme liniare de ecuaţii (Gauss)

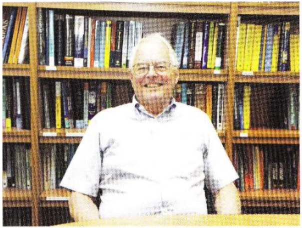
sec. XIX-XX Se conturează domeniul programării liniare: G. Dantzig investighează soluţii ale sistemelor liniare de inegalităţi pentru probleme logistice ale Pentagon-ului.



lclub-mathematicians.blogspot.com

’60-’70 se dezvoltă programarea neliniară

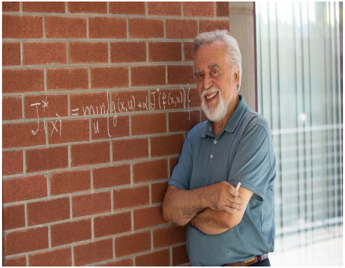
• R.T. Rockafeller: ”The original idea of ”programming” was synonimous with optimization. The only way programming turned out to be computer programming as we know it now is that programming was connected with running or managing a government program on a computer. Early examples, like food distribution programs, very much involved optimization in the sense that finding the best way to do a job, and computer were essential for that. It was then called computer programming.”



https://sites.math.washington.edu/ rtr/mypage.html

O fărâmă de istorie

70'- prezent - se dezvoltă Programarea (Optimizarea) Convexă



eurasip.org, coe.gatech.edu, fullcircle.asu.edu

Modele (probleme) de optimizare:

- **Discretă**: căutare în domeniu discret

$$
\begin{aligned}
\min\ & (x_1 + x_2)^2 \\
\text{s.t. } & x_1 \in \{0, 1\} \\
& x_2 \in \{0, 1\}
\end{aligned}
$$

Modele (probleme) de optimizare:

- **Discretă**: căutare în domeniu discret

$$\begin{aligned} \min\ & (x_1 + x_2)^2 \\ \text{s.t. } & x_1 \in \{0, 1\} \\ & x_2 \in \{0, 1\} \end{aligned}$$

- Continuă: căutare într-un domeniu continuu

$$\begin{aligned} \min\ & (x_1 + x_2)^2 \\ \text{s.t. } & 0 \le x_1 \le 1 \\ & 0 \le x_2 \le 1 \end{aligned}$$

$$\begin{aligned} \min\ & f(x) := (x_1 + x_2)^2 \\ \text{s.t. } & x_1 \in \{0, 1\} \\ & x_2 \in \{0, 1\} \end{aligned}$$

- Căutarea optimului se realizează (în general) prin verificarea unui număr finit de puncte ale mulţimii fezabile

$$\begin{aligned} \min\ & f(x) := (x_1 + x_2)^2 \\ \text{s.t. } & x_1 \in \{0, 1\} \\ & x_2 \in \{0, 1\} \end{aligned}$$

- Căutarea optimului se realizează (în general) prin verificarea unui număr finit de puncte ale mulţimii fezabile
- Mulţimea fezabilă include 4 puncte: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

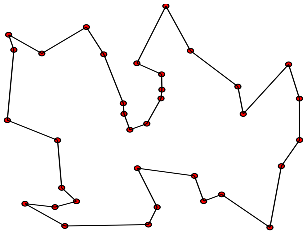
$$\begin{aligned} \min\ & f(x) := (x_1 + x_2)^2 \\ \text{s.t. } & x_1 \in \{0,1\} \\ & x_2 \in \{0,1\} \end{aligned}$$

- Căutarea optimului se realizează (în general) prin verificarea unui număr finit de puncte ale mulţimii fezabile
- Mulţimea fezabilă include 4 puncte: $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix}$
- $f\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = 4,\ f\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = 1,\ f\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = 1,\ f\left(\begin{bmatrix}0\\0\end{bmatrix}\right) = 0$

$$\begin{aligned} \min\ & f(x) := (x_1 + x_2)^2 \\ \text{s.t. } & x_1 \in \{0, 1\} \\ & x_2 \in \{0, 1\} \end{aligned}$$

- Căutarea optimului se realizează (în general) prin verificarea unui număr finit de puncte ale mulţimii fezabile
- Mulţimea fezabilă include 4 puncte: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- În general, numărul de puncte ale mulţimii fezabile creşte exponenţial cu dimensiunea problemei

Problema comis-voiajorului: Pentru o listă dată de oraşe şi distanţe între acestea, care este cea mai scurtă rută posibilă ce trece prin fiecare oraş exact o singură dată şi se întoarce la origine?



wikipedia.com

Distanţa între i şi j notăm d_{ij}, variabila x_{ij} este indicator pentru includerea (i, j) în ruta optimă:

$$\min_{x} \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} d_{ij} x_{ij}$$

$$\text{s.t. } x_{ij} \in \{0, 1\} \quad \forall i, j$$

$$\sum_{i, i \neq j} x_{ij} = 1 \quad \forall i$$

$$\sum_{j, i \neq j} x_{ij} = 1 \quad \forall j$$

- căutare ”brute-force”: $\mathcal{O}(n!)$ (complexitate exponenţială)

Distanţa între i şi j notăm d_{ij}, variabila x_{ij} este indicator pentru includerea (i, j) în ruta optimă:

$$\min_{x} \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} d_{ij} x_{ij}$$

$$\text{s.t. } x_{ij} \in \{0, 1\} \quad \forall i, j$$

$$\sum_{i, i \neq j} x_{ij} = 1 \quad \forall i$$

$$\sum_{j, i \neq j} x_{ij} = 1 \quad \forall j$$

- căutare ”brute-force”: $\mathcal{O}(n!)$ (complexitate exponenţială)
- programare dinamică: $\mathcal{O}\left(n^2 2^n\right)$

Distanţa între i şi j notăm d_{ij}, variabila x_{ij} este indicator pentru includerea (i, j) în ruta optimă:

$$\begin{aligned} \min_{x} \; & \sum_{i=1}^{n} \sum_{i\neq j, j=1}^{n} d_{ij} x_{ij} \\ \text{s.t. } & x_{ij} \in \{0, 1\} \quad \forall i, j \\ & \sum_{i, i\neq j} x_{ij} = 1 \quad \forall i \\ & \sum_{j, i\neq j} x_{ij} = 1 \quad \forall j \end{aligned}$$

- căutare ”brute-force”: $\mathcal{O}(n!)$ (complexitate exponenţială)
- programare dinamică: $\mathcal{O}\left(n^2 2^n\right)$
- TSPLIB (1991), cea mai mare problemă rezolvată $n = 85,900$

Distanţa între i şi j notăm d_{ij}, variabila x_{ij} este indicator pentru includerea (i, j) în ruta optimă:

$$\begin{aligned} \min_x \sum_{i=1}^{n} \sum_{i\neq j, j=1}^{n} d_{ij} x_{ij} \\ \text{s.t. } x_{ij} \in \{0, 1\} \quad \forall i, j \\ \sum_{i, i\neq j} x_{ij} = 1 \quad \forall i \\ \sum_{j, i\neq j} x_{ij} = 1 \quad \forall j \end{aligned}$$

- căutare ”brute-force”: $\mathcal{O}(n!)$ (complexitate exponenţială)
- programare dinamică: $\mathcal{O}(n^2 2^n)$
- TSPLIB (1991), cea mai mare problemă rezolvată $n = 85,900$
- $n = \mathcal{O}(10^6)$ suboptimalitate 3%

Modele (probleme) de optimizare:

- Discretă: căutare în domeniu discret

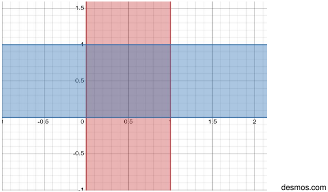
$$\begin{aligned} \min\ & (x_1 + x_2)^2 \\ \text{s.t. } & x_1 \in \{0, 1\} \\ & x_2 \in \{0, 1\} \end{aligned}$$

- **Continuă**: căutare într-un domeniu continuu

$$\begin{aligned} \min\ & (x_1 + x_2)^2 \\ \text{s.t. } & 0 \leq x_1 \leq 1 \\ & 0 \leq x_2 \leq 1 \end{aligned}$$

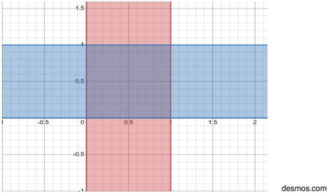
$$
\begin{aligned}
\min \ f(x) := & (x_1 + x_2)^2 \\
\text{s.t. } 0 \le & \ x_1 \le 1 \\
0 \le & \ x_2 \le 1
\end{aligned}
$$

- Mulţimea fezabilă este un poliedru convex:



$$\begin{aligned} \min \; f(x) :=& (x_1 + x_2)^2 \\ \text{s.t. } 0 \leq & \; x_1 \leq 1 \\ 0 \leq & \; x_2 \leq 1 \end{aligned}$$

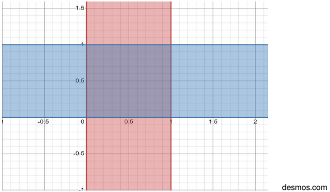
- Mulţimea fezabilă este un poliedru convex:



- Aproximarea optimului se realizează prin execuţia unui algoritm iterativ

$$\begin{aligned} \min \ f(x) :=& (x_1 + x_2)^2 \\ \text{s.t. } 0 \leq& \ x_1 \leq 1 \\ 0 \leq& \ x_2 \leq 1 \end{aligned}$$

- Mulţimea fezabilă este un poliedru convex:



- Aproximarea optimului se realizează prin execuţia unui algoritm iterativ
- $x^* = \pi_{[0,1]^2}(x^* - \nabla f(x^*)) \Leftrightarrow \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} \pi_{[0,1]}(-x_2^*) \\ \pi_{[0,1]}(-x_1^*) \end{bmatrix} \Leftrightarrow x^* = 0$

Problema comis-voiajorului: Pentru o listă dată de oraşe şi distanţe între acestea, care este cea mai scurtă rută posibilă ce trece prin fiecare oraş exact o singură dată şi se întoarce la origine?

$$\min_{x} \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} d_{ij} x_{ij}$$

$$\text{s.t. } x_{ij} \in [0, 1] \quad \forall i, j$$

$$\sum_{i, i \neq j} x_{ij} = 1 \quad \forall i$$

$$\sum_{j, i \neq j} x_{ij} = 1 \quad \forall j$$

Problema comis-voiajorului: Pentru o listă dată de oraşe şi distanţe între acestea, care este cea mai scurtă rută posibilă ce trece prin fiecare oraş exact o singură dată şi se întoarce la origine?

$$\min_{x} \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} d_{ij} x_{ij}$$

$$\text{s.t. } x_{ij} \in [0, 1] \quad \forall i, j$$

$$\sum_{i, i \neq j} x_{ij} = 1 \quad \forall i$$

$$\sum_{j, i \neq j} x_{ij} = 1 \quad \forall j$$

- relaxarea convexă a TSP: soluţia nu coincide cu soluţia TSP!

Problema comis-voiajorului: Pentru o listă dată de oraşe şi distanţe între acestea, care este cea mai scurtă rută posibilă ce trece prin fiecare oraş exact o singură dată şi se întoarce la origine?

$$\min_{x} \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} d_{ij} x_{ij}$$

$$\text{s.t. } x_{ij} \in [0, 1] \quad \forall i, j$$

$$\sum_{i, i \neq j} x_{ij} = 1 \quad \forall i$$

$$\sum_{j, i \neq j} x_{ij} = 1 \quad \forall j$$

- relaxarea convexă a TSP: soluţia nu coincide cu soluţia TSP!
- problemă de *programare liniară* (convexă)

Problema comis-voiajorului: Pentru o listă dată de oraşe şi distanţe între acestea, care este cea mai scurtă rută posibilă ce trece prin fiecare oraş exact o singură dată şi se întoarce la origine?

$$\begin{aligned} \min_{x} \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} & d_{ij} x_{ij} \\ \text{s.t. } x_{ij} \in [0,1] & \quad \forall i,j \\ \sum_{i, i \neq j} x_{ij} = 1 & \quad \forall i \\ \sum_{j, i \neq j} x_{ij} = 1 & \quad \forall j \end{aligned}$$

- relaxarea convexă a TSP: soluţia nu coincide cu soluţia TSP!
- problemă de *programare liniară* (convexă)
- $\mathcal{O}\left(n^3\right)$ (1987), $\mathcal{O}\left(n^{2.5}\right)$ (1989)

Problema comis-voiajorului: Pentru o listă dată de oraşe şi distanţe între acestea, care este cea mai scurtă rută posibilă ce trece prin fiecare oraş exact o singură dată şi se întoarce la origine?

$$\begin{aligned}
\min_{x} \ & \sum_{i=1}^{n} \sum_{i \neq j, j=1}^{n} d_{ij} x_{ij} \\
\text{s.t. } & x_{ij} \in [0,1] \quad \forall i,j \\
& \sum_{i, i \neq j} x_{ij} = 1 \quad \forall i \\
& \sum_{j, i \neq j} x_{ij} = 1 \quad \forall j
\end{aligned}$$

- relaxarea convexă a TSP: soluţia nu coincide cu soluţia TSP!
- problemă de *programare liniară* (convexă)
- $\mathcal{O}\left(n^3\right)$ (1987), $\mathcal{O}\left(n^{2.5}\right)$ (1989)
- $\mathcal{O}\left(n^{2+\frac{1}{6}}\right)$ (Cohen et al., 2019), $\mathcal{O}\left(n^{2+\frac{1}{18}}\right)$ (Jiang et al., 2020)

Model de optimizare cu constrângeri generale:

$$f^* := \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \;\; x \in Q.$$

- f se numeşte funcţie obiectiv (cost)

Model de optimizare cu constrângeri generale:

$$f^* := \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } \; x \in Q.$$

- f se numeşte funcţie obiectiv (cost)
- x variabila de căutare de dimensiune n

Model de optimizare cu constrângeri generale:

$$f^* := \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } x \in Q.$$

- f se numeşte funcţie obiectiv (cost)
- x variabila de căutare de dimensiune n
- Q se numeşte mulţime fezabilă. Defineşte mulţimea din $\mathbb{R}^n$ în care se caută optimul.

Model de optimizare cu constrângeri generale:

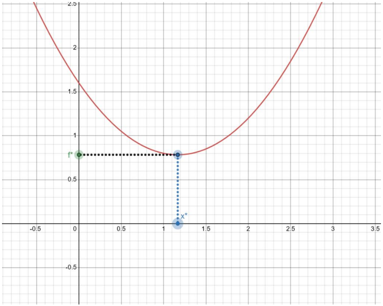
$$f^* := \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } x \in Q.$$

- f se numeşte funcţie obiectiv (cost)
- x variabila de căutare de dimensiune n
- Q se numeşte mulţime fezabilă. Defineşte mulţimea din $\mathbb{R}^n$ în care se caută optimul.
- Format explicit: $Q := \{x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0\}$

Model de optimizare cu constrângeri generale:

$$f^* := \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \;\; x \in Q.$$

Puncte de optim global:

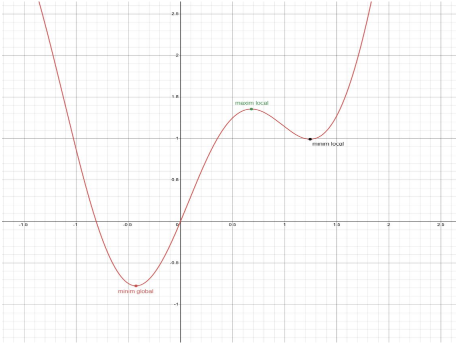


desmos.com

Model de optimizare cu constrângeri generale:

$$f^* := \min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } x \in Q.$$

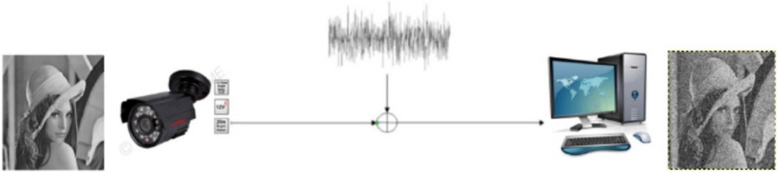
Puncte de optim local:



desmos.com

Surse de zgomot în achiziţii de semnal:

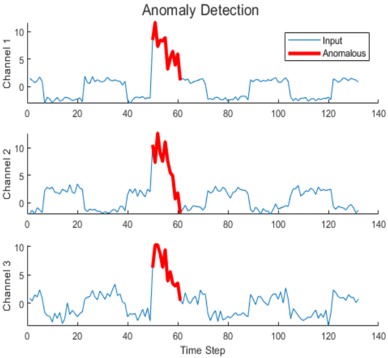
- perturbaţii senzori: factori fizici de mediu (e.g. temperatură)
- pierderi de date (comunicaţie)
- alterări de memorie



Filtre

$$\min_{X \in \mathbb{R}^{m \times n}} \frac{1}{2}\|X - Y\|_F^2 + \lambda \mathcal{L}(X)$$

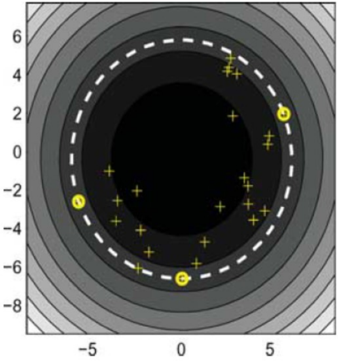
- $\mathcal{L}(\cdot)$ funcţie de penalitate de filtrare (zgomot gaussian)
- $\mathcal{L}(X) = \sum_{i,j} (x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2$
- $\mathcal{L}(X) = \sum_{i,j} |x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}|$



https://www.mathworks.com

Model bazat pe centre: $\min_{c \in \mathbb{R}^n} \sum_{i=1}^{m} \|c - y^i\|_2^2$

Model SVDD: $\min_{c \in \mathbb{R}^n, R > 0} R^2$ s.t. $\|c - y^i\|_2 \leq R$



Alte aplicaţii

- prelucrarea semnalelor: e.g. filtre
- statistică şi învăţare automată: e.g. detecţie de anomalii
- inginerie: e.g. control automat şi distribuit
- logistică şi operaţiuni: e.g. probleme de transport
- motoare de căutare: e.g. problema Google

- **O scurtă recapitulare**
- Convexitate: Mulţimi convexe
- Diferenţiabilitate
- Convexitate: Funcţii convexe

Vezi appendix!

- O scurtă recapitulare
- **Convexitate: Mulţimi convexe**
- Diferenţiabilitate
- Convexitate: Funcţii convexe

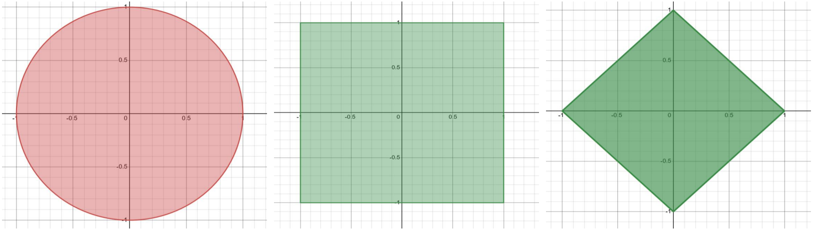
Definiţie

Mulţimea Q este ***convexă*** *dacă şi numai dacă* $\forall x, y \in Q$

$$\alpha x + (1 - \alpha) y \in Q \qquad \forall \alpha \in [0, 1].$$

Mai mult, Q este ***închisă*** *dacă îşi conţine toate punctele limită (frontiera).*

Pe scurt: segmentul aflat între oricare două puncte ale mulţimii se află, de asemenea, în mulţime.

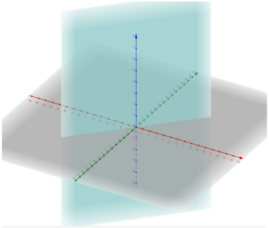


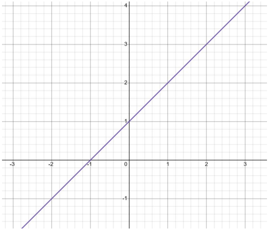
www.desmos.com

- punct e.g. $\{0\}$

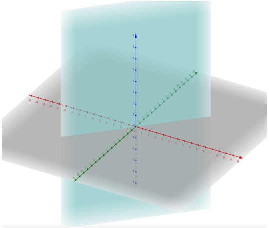
- punct e.g. $\{0\}$
- dreapta/segment e.g. $\mathbb{R}^1 : [0, \infty)$, $\mathbb{R}^2 : \{x \in \mathbb{R}^2 : x_1 + x_2 = 0\}$

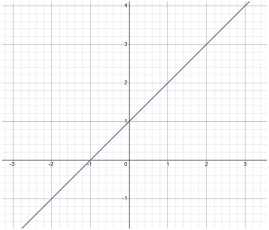
- punct e.g. $\{0\}$
- dreapta/segment e.g. $\mathbb{R}^1 : [0, \infty)$, $\mathbb{R}^2 : \{x \in \mathbb{R}^2 : x_1 + x_2 = 0\}$
- hiperplan $\mathbb{R}^n : \{x \in \mathbb{R}^n : a^T x = b\}$





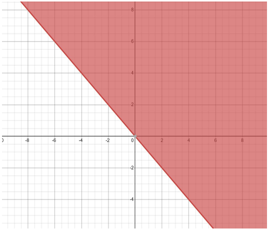
- punct e.g. $\{0\}$
- dreapta/segment e.g. $\mathbb{R}^1 : [0, \infty)$, $\mathbb{R}^2 : \{x \in \mathbb{R}^2 : x_1 + x_2 = 0\}$
- hiperplan $\mathbb{R}^n : \{x \in \mathbb{R}^n : a^T x = b\}$



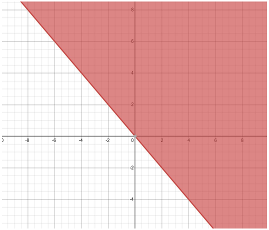


- varietate liniară $\mathbb{R}^n : \{x \in \mathbb{R}^n : Ax = b\}$

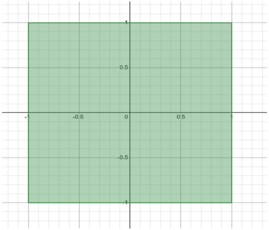
- semispaţiu $\quad S(a,b) := \{x \in \mathbb{R}^n : a^T x \leq b\}$

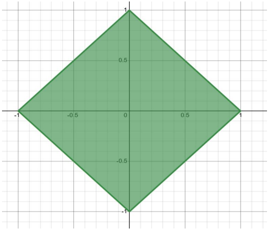


- semispaţiu $S(a, b) := \{x \in \mathbb{R}^n : a^T x \leq b\}$

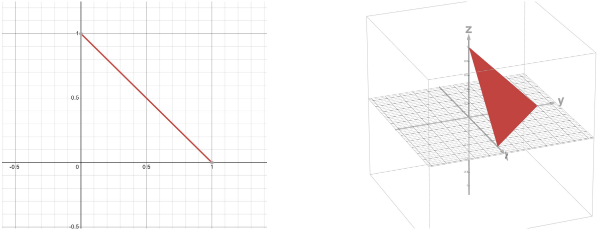


- poliedru $C := \{x \in \mathbb{R}^n : Ax \leq b\}$



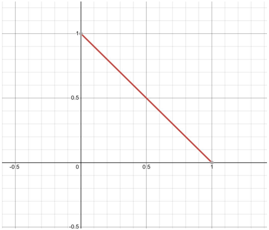


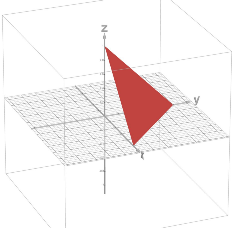
- simplex $\Delta_n := \{x \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 1, x \geq 0\}$



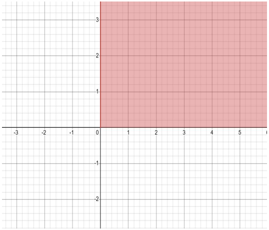
Mulţimi convexe - exemple

- simplex $\quad \Delta_n := \{x \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 1, x \geq 0\}$

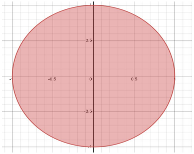


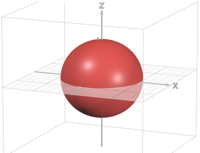


- con $\quad$ e.g. $\quad$ e.g. $\mathbb{R}_+ := [0, \infty), \mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \geq 0\}$



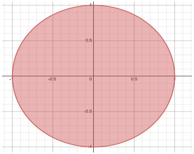
- bila $B(c,r) := \{x \in \mathbb{R}^n : \|x - c\| \leq r\}$, e.g. $c = 0, r = 1$

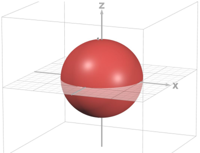




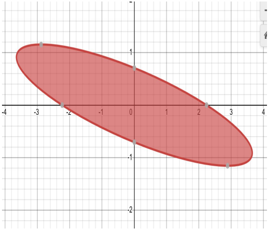
Mulţimi convexe - exemple non-afine

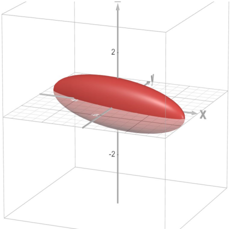
- bila $\quad B(c,r) := \{x \in \mathbb{R}^n : \|x - c\| \leq r\}$, e.g. $c = 0, r = 1$



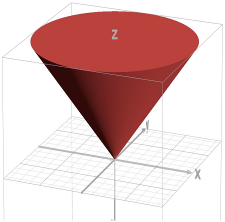


- elipsoid $\quad E(c,r,P) := \{x \in \mathbb{R}^n : (x-c)^T P(x-c) \leq r\}$, e.g.
  $c = 0, r = 1, (\text{st.}) P = \begin{bmatrix} 0.2 & 0.5 \\ 0.5 & 2 \end{bmatrix}; (\text{dr.}) P = \begin{bmatrix} 0.2 & 0.5 & 0 \\ 0.5 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$





- con non-afin   e.g. con Lorentz $K_L := \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq x_3\}$



Fie $C_1, C_2, \cdots, C_m \subset \mathbb{R}^n$ mulţimi convexe închise.

- *Intersecţia*: mulţimea $C = \bigcap_{i=1}^{m} C_i$ este închisă şi convexă

Fie $C_1, C_2, \cdots, C_m \subset \mathbb{R}^n$ mulţimi convexe închise.

- *Intersecţia*: mulţimea $C = \bigcap_{i=1}^{m} C_i$ este închisă şi convexă
- *Produs cartezian*: mulţimea $C = C_1 \times C_2 \times \cdots \times C_m$ este închisă şi convexă.

Fie $C_1, C_2, \cdots, C_m \subset \mathbb{R}^n$ mulţimi convexe închise.

- *Intersecţia*: mulţimea $C = \bigcap_{i=1}^{m} C_i$ este închisă şi convexă
- *Produs cartezian*: mulţimea $C = C_1 \times C_2 \times \cdots \times C_m$ este închisă şi convexă.
- *Compunerea cu un operator afin*: fie $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, mulţimea $C_i' = \{z \in C_i : Az - b\}$ este închisă şi convexă.

- Un hiperplan (afin) reprezintă intersecţia dintre două semi-spaţii

$$H(a,b) = \{x \in \mathbb{R}^n : a^T x \leq b, a^T x \geq b\} = S(a,b) \cap S(-a,-b)$$

- Un hiperplan (afin) reprezintă intersecţia dintre două semi-spaţii

$$H(a,b) = \{x \in \mathbb{R}^n : a^T x \leq b, a^T x \geq b\} = S(a,b) \cap S(-a,-b)$$

- O varietate liniară reprezintă intersecţia unui număr finit de hiperplane

$$V(A,b) = H(A_1,b_1) \cap H(A_2,b_2) \cap \cdots \cap H(A_m,b_m)$$

- Un hiperplan (afin) reprezintă intersecţia dintre două semi-spaţii

$$H(a,b) = \{x \in \mathbb{R}^n : a^T x \leq b, a^T x \geq b\} = S(a,b) \cap S(-a,-b)$$

- O varietate liniară reprezintă intersecţia unui număr finit de hiperplane

$$V(A,b) = H(A_1, b_1) \cap H(A_2, b_2) \cap \cdots \cap H(A_m, b_m)$$

- Mulţimea simplex (unitate) este intersecţia dintre un hiperplan şi un con convex

$$\Delta_n := \{x \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 1, x \geq 0\} = H(e,1) \cap \mathbb{R}^n_+$$

- Un hiperplan (afin) reprezintă intersecţia dintre două semi-spaţii

$$H(a,b) = \{x \in \mathbb{R}^n : a^T x \leq b, a^T x \geq b\} = S(a,b) \cap S(-a,-b)$$

- O varietate liniară reprezintă intersecţia unui număr finit de hiperplane

$$V(A,b) = H(A_1,b_1) \cap H(A_2,b_2) \cap \cdots \cap H(A_m,b_m)$$

- Mulţimea simplex (unitate) este intersecţia dintre un hiperplan şi un con convex

$$\Delta_n := \{x \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 1, x \geq 0\} = H(e,1) \cap \mathbb{R}^n_+$$

- Un poliedru reprezintă intersecţia unui număr finit de semi-spaţii

$$\{x \in \mathbb{R}^n : Ax \leq b\} = S(A_1,b_1) \cap S(A_2,b_2) \cap \cdots \cap S(A_m,b_m)$$

Translaţiile, intersecţiile şi semi-spaţiile sunt obiectele de bază în analiza convexă.

- $\{x \in \mathbb{R}^n : |a^T x + b| \leq r\}$

Translaţiile, intersecţiile şi semi-spaţiile sunt obiectele de bază în analiza convexă.

- $\{x \in \mathbb{R}^n : |a^T x + b| \leq r\}$
- Hipercub: $\{x \in \mathbb{R}^n : -1 \leq x \leq 1\}$

Translaţiile, intersecţiile şi semi-spaţiile sunt obiectele de bază în analiza convexă.

- $\{x \in \mathbb{R}^n : |a^T x + b| \leq r\}$
- Hipercub: $\{x \in \mathbb{R}^n : -1 \leq x \leq 1\}$
- $\{x \in \mathbb{R}^n : \|Ax\|_\infty \leq 1\}$

Translaţiile, intersecţiile şi semi-spaţiile sunt obiectele de bază în analiza convexă.

- $\{x \in \mathbb{R}^n : |a^T x + b| \leq r\}$
- Hipercub: $\{x \in \mathbb{R}^n : -1 \leq x \leq 1\}$
- $\{x \in \mathbb{R}^n : \|Ax\|_\infty \leq 1\}$
- $\{x \in \mathbb{R}^n : \|Ax\|_\infty \leq 1, \|x\|_2 \leq 1\}$

Definiţie

Fie $Z = \{z^1, z^2, \cdots, z^n\} \subset \mathbb{R}^n$, *pentru* $\alpha \in \Delta_n$ *combinaţia* $\sum_i \alpha_i z^i$ *se numeşte combinaţie convexă a punctelor din* Z.

Definiţie

Fie $Z = \{z^1, z^2, \cdots, z^n\} \subset \mathbb{R}^n$, *pentru* $\alpha \in \Delta_n$ *combinaţia* $\sum_i \alpha_i z^i$ *se numeşte combinaţie convexă a punctelor din* Z.

Definiţie (Reformulare)

Mulţimea $Q \subset \mathbb{R}^n$ *se numeşte convexă dacă pentru oricare set de puncte* $\{x^1, x^2, \cdots, x^n\} \subset Q$ *orice combinaţie convexă a punctelor din set rămâne în mulţimea* Q.

Definiţie

Fie $Z = \{z^1, z^2, \cdots, z^n\} \subset \mathbb{R}^n$, pentru $\alpha \in \Delta_n$ combinaţia $\sum_i \alpha_i z^i$ se numeşte combinaţie convexă a punctelor din Z.

Definiţie (Reformulare)

Mulţimea $Q \subset \mathbb{R}^n$ se numeşte convexă dacă pentru oricare set de puncte $\{x^1, x^2, \cdots, x^n\} \subset Q$ orice combinaţie convexă a punctelor din set rămâne în mulţimea Q.

Definiţie

Fie $Q \subset \mathbb{R}^n$ (nu neapărat convexă), acoperire convexă Q_h reprezintă cea mai "mică" mulţime convexă care include pe Q.

- O scurtă recapitulare
- Convexitate: Mulţimi convexe
- **Diferenţiabilitate**
- Convexitate: Funcţii convexe

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte diferenţiabilă în punctul x dacă $\exists a \in \mathbb{R}^n$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

$$f(x+y) = f(x) + a^T y + o(y)$$

- $a =: \nabla f(x)$ este *gradientul* funcţiei f în x

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte diferenţiabilă în punctul x dacă $\exists a \in \mathbb{R}^n$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

$$f(x+y) = f(x) + a^T y + o(y)$$

- $a =: \nabla f(x)$ este *gradientul* funcţiei f în x
- $\lim\limits_{||y|| \to 0} \frac{o(y)}{y} = 0$

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte diferenţiabilă în punctul x dacă $\exists a \in \mathbb{R}^n$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

$$f(x+y) = f(x) + a^T y + o(y)$$

- $a =: \nabla f(x)$ este *gradientul* funcţiei f în x
- $\lim_{||y|| \to 0} \frac{o(y)}{y} = 0$
- Pe scurt: funcţia este diferenţiabilă în x dacă putem găsi o funcţie liniară $g(y) := f(x) + \nabla f(x)^T y$ astfel încât $f(x+y) - g(y) = o(y)$

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte diferenţiabilă în punctul x dacă $\exists a \in \mathbb{R}^n$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

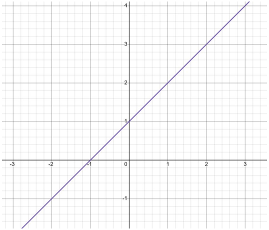
$$f(x+y) = f(x) + a^T y + o(y)$$

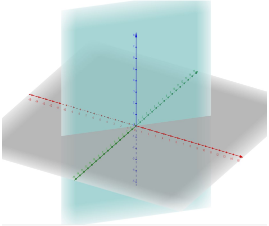
- $a =: \nabla f(x)$ este *gradientul* funcţiei f în x
- $\lim_{\|y\| \to 0} \frac{o(y)}{y} = 0$
- Pe scurt: funcţia este diferenţiabilă în x dacă putem găsi o funcţie liniară $g(y) := f(x) + \nabla f(x)^T y$ astfel încât $f(x+y) - g(y) = o(y)$
- $\nabla f(\cdot)$ este unic determinat şi definit de: $\nabla f(x) := \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$

Exemple:

- funcţia liniară: $f(x) = b^T x, \quad b \in \mathbb{R}^n$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$



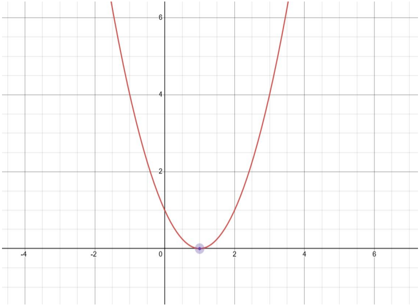


www.desmos.com

Exemple:

- funcţia pătratică: $f(x) = \frac{1}{2}x^T Ax - b^T x$, unde $A \in \mathbb{R}^{n\times n}, A = A^T$

$$\nabla f(x) = Ax - b$$



www.desmos.com

Exemple:

- funcţia pătratică: $f(x) = \frac{1}{2}x^T Ax - b^T x$, unde $A \in \mathbb{R}^{n\times n}, A = A^T$

$$\nabla f(x) = Ax - b$$

$$\begin{aligned} f(x+y) &= \frac{1}{2}(x+y)^T A(x+y) - b^T(x+y) \\ &= \frac{1}{2}x^T Ax - b^T x + (Ax-b)^T y + \frac{1}{2}y^T Ay \\ &= f(x) + (Ax-b)^T y + \frac{1}{2}y^T Ay, \end{aligned}$$

Observaţi: $\frac{1}{2}y^T Ay = o(y)$.

Definiţie

O funcţie f este continuă Lipschitz cu constanta L dacă satisface:

$$|f(x) - f(y)| \leq L\|x - y\| \quad \forall x, y$$

Echivalent orice funcţie Lipschitz satisface $\|\nabla f(x)\| \leq L \;\; \forall x$.

- $|x|$

Definiţie

O funcţie f este continuă Lipschitz cu constanta L dacă satisface:

$$|f(x) - f(y)| \leq L\|x - y\| \quad \forall x, y$$

Echivalent orice funcţie Lipschitz satisface $\|\nabla f(x)\| \leq L \;\; \forall x$.

- $|x|$
- $\|x\|$

Definiţie

O funcţie f este continuă Lipschitz cu constanta L dacă satisface:

$$|f(x) - f(y)| \leq L\|x - y\| \quad \forall x, y$$

Echivalent orice funcţie Lipschitz satisface $\|\nabla f(x)\| \leq L \ \ \forall x$.

- $|x|$
- $\|x\|$
- $\|Ax\|$

Definiţie

Prima derivată a funcţiei f este continuă Lipschitz cu constanta L dacă:

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y$$

Echivalent, se exprimă:

$$\|\nabla^2 f(x)\| \leq L \quad \forall x$$

$$|f(x) - f(y) - \nabla f(y)^T (x - y)| \leq \frac{L}{2}\|x - y\|^2 \quad \forall x, y$$

- $\|x\|^2$

Definiţie

Prima derivată a funcţiei f este continuă Lipschitz cu constanta L dacă:

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \quad \forall x, y$$

Echivalent, se exprimă:

$$\|\nabla^2 f(x)\| \leq L \quad \forall x$$

$$|f(x) - f(y) - \nabla f(y)^T(x - y)| \leq \frac{L}{2}\|x - y\|^2 \quad \forall x, y$$

- $\|x\|^2$
- $\|Ax\|^2$

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte *dublu diferenţiabilă în punctul* x dacă există matricea simetrică $H \in \mathbb{R}^{n\times n}$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

$$f(x+y) = f(x) + \langle \nabla f(x), y\rangle + \frac{1}{2}\langle Hy, y\rangle + o(\|y\|^2)$$

- $H =: \nabla^2 f(x)$ este *matricea derivatelor de ordin II* sau *Hessiana* funcţiei f în x

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte *dublu diferenţiabilă în punctul* x dacă există matricea simetrică $H \in \mathbb{R}^{n \times n}$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

$$f(x+y) = f(x) + \langle \nabla f(x), y \rangle + \frac{1}{2}\langle Hy, y \rangle + o(\|y\|^2)$$

- $H =: \nabla^2 f(x)$ este *matricea derivatelor de ordin II* sau *Hessiana* funcţiei f în x
- Pe scurt: funcţia este dublu diferenţiabilă în x dacă admite o aproximare patratică de ordin II în vecinătatea lui x, i.e. $g(y) := f(x) + \nabla f(x)^T y + \frac{1}{2} y^T H y$ astfel încât $|f(x+y) - g(y)| = o(\|y\|^2)$

Funcţii scalare: $f : \mathbb{R}^n \to \mathbb{R}$ se numeşte *dublu diferenţiabilă în punctul* x dacă există matricea simetrică $H \in \mathbb{R}^{n\times n}$ astfel încât, pentru orice $y \in \mathbb{R}^n$:

$$f(x+y) = f(x) + \langle \nabla f(x), y\rangle + \frac{1}{2}\langle Hy, y\rangle + o(\|y\|^2)$$

- $H =: \nabla^2 f(x)$ este *matricea derivatelor de ordin II* sau *Hessiana* funcţiei f în x
- Pe scurt: funcţia este dublu diferenţiabilă în x dacă admite o aproximare patratică de ordin II în vecinătatea lui x, i.e. $g(y) := f(x) + \nabla f(x)^T y + \frac{1}{2} y^T H y$ astfel încât $|f(x+y) - g(y)| = o(\|y\|^2)$
- $$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

Exemple:

- funcţia pătratică: $f(x) = \frac{1}{2}x^T Ax - b^T x$, unde $A \in \mathbb{R}^{n\times n}, A = A^T$

$$\nabla^2 f(x) = A$$

$$\begin{aligned} f(x+y) &= \frac{1}{2}(x+y)^T A(x+y) - b^T(x+y) \\ &= \frac{1}{2}x^T Ax - b^T x + (Ax-b)^T y + \frac{1}{2}y^T Ay \\ &= f(x) + \nabla f(x)^T y + \frac{1}{2}y^T Ay, \end{aligned}$$

Observaţi: $H = A$.

Definiţie

Derivată a doua a funcţiei (dublu diferenţiabile) f este continuă Lipschitz cu constanta L dacă:

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x - y\| \quad \forall x, y.$$

Echivalent, se arată:

$$\|\nabla f(x) - \nabla f(y) - \nabla^2 f(y)(x - y)\| \leq \frac{L}{2}\|x - y\|^2 \quad \forall x, y$$

$$|f(x) - f(y) - \nabla f(y)^T(x - y) - \frac{1}{2}(x - y)^T\nabla^2 f(y)(x - y)| \leq \frac{L}{6}\|x - y\|^3 \quad \forall x, y$$

- x^3

Definiţie

Derivată a doua a funcţiei (dublu diferenţiabile) f este continuă Lipschitz cu constanta L dacă:

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x - y\| \quad \forall x, y.$$

Echivalent, se arată:

$$\|\nabla f(x) - \nabla f(y) - \nabla^2 f(y)(x - y)\| \leq \frac{L}{2}\|x - y\|^2 \quad \forall x, y$$

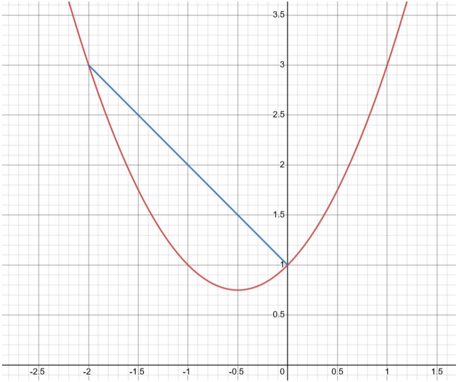
$$|f(x) - f(y) - \nabla f(y)^T(x - y) - \frac{1}{2}(x - y)^T\nabla^2 f(y)(x - y)| \leq \frac{L}{6}\|x - y\|^3 \quad \forall x, y$$

- x^3
- $\|x\|^3$

- O scurtă recapitulare
- Convexitate: Mulţimi convexe
- Diferenţiabilitate
- **Convexitate: Funcţii convexe**

Funcţia scalară $f : \mathbb{R}^n \to (-\infty, +\infty]$ este convexă dacă:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) \quad \forall \alpha \in [0, 1]$$



www.desmos.com

Funcţia scalară $f : \mathbb{R}^n \to (-\infty, +\infty]$ este convexă dacă:

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \quad \forall \alpha \in [0,1]$$

$$f(x) := a^T x + b$$

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= a^T[\alpha x + (1-\alpha)y] + b \\ &= \alpha a^T x + (1-\alpha)a^T y + b \\ &= \alpha(a^T x + b) + (1-\alpha)(a^T y + b). \end{aligned}$$

Funcţia scalară $f : \mathbb{R}^n \to (-\infty, +\infty]$ este convexă dacă:

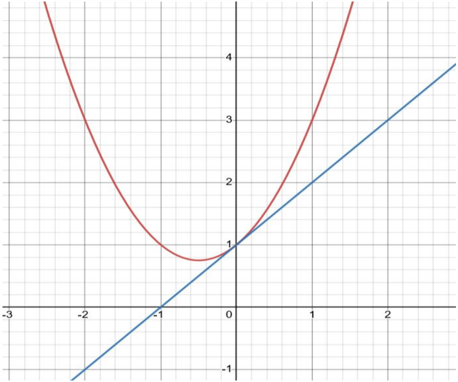
$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) \quad \forall \alpha \in [0,1]$$

$$f(x) := |a^T x + b|$$

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= |a^T[\alpha x + (1-\alpha)y] + b| \\ &= |\alpha(a^T x + b) + (1-\alpha)(a^T y + b)| \\ &\overset{I.T.}{\leq} |\alpha(a^T x + b)| + |(1-\alpha)(a^T y + b)| \\ &= \alpha|a^T x + b| + (1-\alpha)|a^T y + b|. \end{aligned}$$

Funcţia $f : \mathbb{R}^n \to (-\infty, +\infty]$ diferenţiabilă este convexă dacă:

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \qquad \forall x, y$$



www.desmos.com

Funcţia $f : \mathbb{R}^n \to (-\infty, +\infty]$ diferenţiabilă este convexă dacă:

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \qquad \forall x, y$$

Dacă f este dublu diferenţiabilă atunci definiţia de mai sus este echivalentă cu:

$$\nabla^2 f(x) \succeq 0 \qquad x \in \mathrm{dom}(f)$$

Dacă f este dublu diferenţiabilă atunci definiţia de mai sus este echivalentă cu:

$$\nabla^2 f(x) \succeq 0 \qquad x \in \mathrm{dom}(f)$$

Exemplu: $f : \mathbb{R} \to \mathbb{R}, f(x) = x^4$

$$\nabla^2 f(x) = 12x^2 \geq 0 \qquad \forall x \in \mathbb{R}.$$

$f : \mathbb{R}^2 \to \mathbb{R}, f(x) = \|x\|^4$

$$\nabla f(x) = 4\|x\|^2 x$$
$$\nabla^2 f(x) = 4\|x\| I_2 + 8xx^T \succeq 0 \qquad \forall x \in \mathbb{R}^2.$$

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă
- *Maximum*: funcţia $f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}$ este închisă şi convexă.

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă
- *Maximum*: funcţia $f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}$ este închisă şi convexă.
- *Compunerea interioară cu un operator afin*: fie $A \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$, funcţia $g(x) = f(Ax - b)$ este închisă şi convexă.

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă
- *Maximum*: funcţia $f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}$ este închisă şi convexă.
- *Compunerea interioară cu un operator afin*: fie $A \in \mathbb{R}^{p\times n}$, $b \in \mathbb{R}^p$, funcţia $g(x) = f(Ax - b)$ este închisă şi convexă.
- *Compunerea exterioară cu o funcţie crescătoare*: fie $\phi : \mathbb{R} \to \mathbb{R}$ o funcţie crescătoare, funcţia $g(x) = \phi(f(x))$ este închisă şi convexă.

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă

$$e.g.\ \|Ax - b\|_2^2 + \|x\|_1$$

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă

$$e.g.\ \|Ax - b\|_2^2 + \|x\|_1$$

- *Maximum*: funcţia $f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}$ este închisă şi convexă.

$$e.g.\ |a^T x + b|, ||Ax + b||_1$$

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă

$$e.g. \; \|Ax - b\|_2^2 + \|x\|_1$$

- *Maximum*: funcţia $f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}$ este închisă şi convexă.

$$e.g. \; |a^T x + b|, ||Ax + b||_1$$

- *Compunerea interioară cu un operator afin*: fie $A \in \mathbb{R}^{p \times n}, b \in \mathbb{R}^p$, funcţia $g(x) = f(Ax - b)$ este închisă şi convexă.

$$e.g. \; ||Ax + b||_p, \; p \geq 1$$

Fie $f_1, f_2, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ funcţii convexe închise.

- *Suma*: pentru $\alpha \in \mathbb{R}^m_+$, funcţia $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ este închisă şi convexă

$$e.g.\ \|Ax - b\|_2^2 + \|x\|_1$$

- *Maximum*: funcţia $f(x) = \max\{f_1(x), f_2(x), \cdots, f_m(x)\}$ este închisă şi convexă.

$$e.g.\ |a^T x + b|, ||Ax + b||_1$$

- *Compunerea interioară cu un operator afin*: fie $A \in \mathbb{R}^{p \times n}, b \in \mathbb{R}^p$, funcţia $g(x) = f(Ax - b)$ este închisă şi convexă.

$$e.g.\ ||Ax + b||_p,\ p \geq 1$$

- *Compunerea exterioară cu o funcţie crescătoare*: fie $\phi : \mathbb{R} \to \mathbb{R}$ o funcţie crescătoare, funcţia $g(x) = \phi(f(x))$ este închisă şi convexă.

$$e.g.\ \max\{0, (a^T x)^2 - b\}$$

Notaţii

$\mathbb{R}^n$	spaţiul real Euclidian n-dimensional
x_i	componenta i a vectorului $x \in \mathbb{R}^n$
$\mathbb{R}^n_+$	Cadranul $n-$dimensional ne-negativ din R^n.$R^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0\}$.
$\langle \cdot, \cdot \rangle$	produsul scalar: $\langle x, y \rangle = x^T y = x_1 y_1 + \cdots x_n y_n$
$\nabla f(x), f'(x)$	gradientul/derivata funcţiei scalare $f(x)$
$\nabla^2 f(x), f''(x)$	matricea Hessiană a funcţiei scalare $f(x)$
$o(h(x))$	pentru $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^s$, dacă $\frac{\\|g(x)\\|}{\\|h(x)\\|} \to 0$ când $\\|x\\| \to 0$, atunci $g(x) = o(h(x))$
$\mathcal{O}(h(x))$	pentru $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^s$, dacă există $\epsilon, \alpha > 0$ astfel încât $\\|g(x)\\| \leq \alpha \\|h(x)\\|$ pentru $\\|x\\| \leq \epsilon$, atunci $g(x) = \mathcal{O}(h(x))$
$[x]_+$	$\max\{0, x\}$
I_n	matricea identitate de ordin n
A^T	matricea transpusă asociată matricii A
$\Lambda(A)$	spectrul (mulţimea valorilor proprii) matricii simetrice A
$\\|A\\|$	norma matricii $A \in \mathbb{R}^{n \times n}$: $\\|A\\| = \max_{\\|x\\|=1} \\|Ax\\| = \max_{1 \leq i \leq n} \lvert\lambda_i(A)\rvert$
$A \succeq 0$	matricea simetrică A este positiv semidefinită: $x^T A x \geq 0$ pentru $\forall x \in \mathbb{R}^n$. Echivalent $\Lambda(A) \subset \mathbb{R}^n_+$.
$A \succeq B$	matricile A, B sunt simetrice şi $A - B \succeq 0$

- B. Polyak, Introduction to Optimization, Optimization Software Inc., New York, 1987
- D. Bertsekas, Nonlinear Programming, Third Edition. Athena Scientic, 2016.
- Y. Nesterov, Introductory Lectures on Convex Optimization, Kluwer, 2004.
- Hiriart-Urruty, Jean-Baptiste, and Claude Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals.* Vol. 305. Springer science & business media, 1996.