

Curs 9

- Alg. optimizării "Gradient descent"
- aplicat pe o funcție pătratică:

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$f(x, y) = 100x^2 + y^2$$

$$\frac{\partial f}{\partial x}(x, y) = 200x + 0 = 200x$$

$$\frac{\partial f}{\partial y}(x, y) = 0 + 2y = 2y$$

$$\nabla f(x, y) = \begin{pmatrix} 200x \\ 2y \end{pmatrix}$$

$$f(x, y) = 100x^2 + y^2 \geq 0 = f(0, 0)$$

$\geq 0 \quad \geq 0$

Minimul lui f este atins în (0, 0)

Constr alg. gradient pt opt f.

$$\underline{\text{Init}} : \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \mathbb{R}^2$$

$$\underline{\text{Iter}} : \begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - t \overset{\text{fix}}{\nabla f(x_n, y_n)}$$

$$= \begin{pmatrix} \tilde{x}_n \\ \tilde{y}_n \end{pmatrix} - t \begin{pmatrix} 200\tilde{x}_n \\ 2\tilde{y}_n \end{pmatrix}$$

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n (1 - 200t) \\ y_n (1 - 2t) \end{pmatrix}$$

$$= \dots = \begin{pmatrix} x_0 (1 - 200t)^{n+1} \\ y_0 (1 - 2t)^{n+1} \end{pmatrix} \xrightarrow{?} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\underline{x_0 (1 - 200t)^n} \xrightarrow{n} 0$$

$$\underline{y_0 (1 - 2t)^n} \xrightarrow{n} 0$$

converge

Convergenz: $\Leftrightarrow \begin{cases} |1-200t| < 1 \\ |1-2t| < 1 \end{cases}$

$$\begin{cases} -1 < 1-200t < 1 \\ -1 < 1-2t < 1 \end{cases}$$

$$\begin{cases} t < \frac{1}{100} \\ t < 1 \end{cases} \quad \text{früher als } t \in (0, \frac{1}{100})$$

$$\frac{1}{2} (x \ y) \begin{pmatrix} \underline{200} & 0 \\ 0 & \underline{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 100x^2 + y^2$$

$$f(x,y) = \frac{1}{2} (x \ y) A \begin{pmatrix} x \\ y \end{pmatrix} - b \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\nabla f = A \begin{pmatrix} x \\ y \end{pmatrix} - b$$

$$A \cdot (\text{vect}) - b$$

Pentru $\frac{1}{2} x^T A x - b^T x$

A are val proprii $\lambda_1 \leq \dots \leq \lambda_n$

$$\leadsto f(x) \sim \underline{\lambda_1} x_1^2 + \dots + \underline{\lambda_n} x_n^2$$

Pasii pt cu "Grad Descent"
converge sunt $t \in (0, \frac{2}{\lambda_n})$

Slow convergence?

Thm: pt $x^T A x - b^T x$ avem
convergent GARANȚAT pt

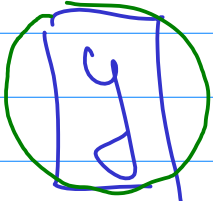
$$t < \frac{2}{\lambda_n}$$

Garantia de convergent \Rightarrow Alg util?
Nr de iterații pt conv este pro mare

$$f(x,y) = 10^4 x^2 + y^2$$

$$\nabla f(x,y) = \begin{pmatrix} 20000x \\ 2y \end{pmatrix}$$

→ O componentă a lui ∇f care este >> ca și cealaltă.

→ GD utilizează același pas în ambele direcții; nu va reuși să avanseze în direcția 

Ce putem face pentru a remedia situația?

$$\nabla f(x,y) = \begin{pmatrix} 20000x \\ 2y \end{pmatrix}$$

→ Să modificăm factorii lui $\begin{bmatrix} x & y \\ 1 & 1 \end{bmatrix}$

$$\nabla f \leadsto \left(\begin{array}{c|c} \dots & x \\ \hline \dots & y \end{array} \right)$$

Matrix \bar{A} : $\bar{A} \nabla f = \begin{pmatrix} x \\ y \end{pmatrix}$

$$\underbrace{\begin{pmatrix} \frac{1}{20000} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}}_{\bar{A}} \begin{pmatrix} 20000x \\ 2y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$A = \begin{pmatrix} 20000 & 0 \\ 0 & 2 \end{pmatrix} \quad \nabla f \rightarrow f = x^T A x - b^T x$$

$$\text{Hess } f = D^2 f = A$$

met
inverse

$$A^{-1} = \begin{pmatrix} \frac{1}{20000} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

$$x_{n+1} = x_n - \cancel{\nabla^2 f(x_n)^{-1} \nabla f(x_n)}$$

Newton 1D:

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

$$x_{n+1} = x_n - \nabla^2 f(x_n)^{-1} \nabla f(x_n)$$

Methode bei Newton im 2D: