Curs 9

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

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

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

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

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

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

$$f(x,y) = \underbrace{100x^2}_{\geq 0} + \underbrace{y^2}_{\geq 0} \geq 0 = f(0,0)$$

Minimul lui f este atins în $\underline{(0,0)}$

Care este alg. gradient pt opt f.

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

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

$$= \begin{pmatrix} x_n \\ y_n \end{pmatrix} - t \begin{pmatrix} 200x_n \\ 2y_n \end{pmatrix}$$

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

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

soluție

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

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

Convergență: $\Leftrightarrow \begin{cases} |1-2\alpha t| < 1 \\ |1-2t| < 1 \end{cases}$

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

$\begin{cases} t < \frac{1}{100} \\ t < 1 \end{cases}$ trebuie ales $t \in (0, \frac{1}{100})$

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

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

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

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

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

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

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

Pași pt care "Grad ~~D~~escent" converge sunt $t \in \left(0, \frac{2}{\lambda_n}\right)$

<u>Slow convergența?</u>

<u>Thm</u>: pt $x^T A x - b^T x$ avem convergență <u>GARANTAT</u> pt

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

Garanția de convergență $\Rightarrow$ <u>Alg util?</u>

Nr de iterații pt care este prea <u>mare</u>

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

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

→ O componentă a lui ∇f car este >> ca și celelalte.

→ GD utilizează aceleași pas în ambele direcții; nu va reuși să avanseze în direcția $\boxed{y}$

Ce putem face pentru a remedia situația?

$$\nabla f(x,y) = \begin{pmatrix} 2000x \\ 2y \end{pmatrix}$$

→ Să modificăm factorii lui $\boxed{x \text{ și } y}$

$$\nabla f \rightsquigarrow \begin{pmatrix} \cdots \\ \cdots \\ \cdots \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Matriz $\bar{A}$: $\quad \bar{A}\,\nabla f = \begin{pmatrix} x \\ y \end{pmatrix}$

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

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

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

met inverso

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

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

Newton 1D:

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

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

Metoda lui Newton în 2D: