definim $X = \begin{bmatrix} \ldots (x^1) \ldots \\ \ldots (x^2) \ldots \\ \vdots \\ \ldots (x^m) \ldots \end{bmatrix}$ $\Theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$

$\vec{y} = \begin{bmatrix} y^1 \\ \vdots \\ y^m \end{bmatrix}$

definim fct de cost $J(\Theta) = \frac{1}{2} \sum (h_\theta(x^i) - y^i)^2$

$\sum_i^m z_i^2 = z^T z$ $\Rightarrow$

$\Rightarrow J = \frac{1}{2} (X\Theta - y)^T (X\Theta - y)$

$\nabla_\Theta J(\Theta) = \vec{0} \Rightarrow$

$\nabla_\Theta \frac{1}{2} (X\Theta - y)^T (X\Theta - y) =$

$\nabla_\Theta \frac{1}{2} (\Theta^T X^T - y^T)(X\Theta - y) =$

$\nabla_\Theta \frac{1}{2} (\Theta^T X^T X\Theta - \Theta^T X^T y - y^T X\Theta + y^T y) =$

nu este fct de Θ

=

$tr A = ~~tr~~ A$ p/t $A \in \mathbb{R}$
nr real

$$\Rightarrow \nabla_\theta J = \frac{1}{2} \nabla_\theta tr(\theta^T X^T X \theta - \theta^T X^T y - y^T X \theta) =$$

$$~~tr~~ A = tr A^T$$

$$\Rightarrow \nabla_\theta J = \frac{1}{2} \nabla_\theta (tr \theta^T X^T X \theta - 2 tr y^T X \theta) =$$

$$= \frac{1}{2} (2 X^T X \theta - 2 X^T y)$$

$\nabla_{A^T} tr A B A^T C =$
$= B^T A^T C^T + B A^T C$

$$= X^T X \theta - X^T y = 0 \Rightarrow$$

$$\Rightarrow \theta = (X^T X)^{-1} X^T y$$

în cazul meu $\theta = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$

$h = b_0 + b_1 x$

$x_0 = 1$ intercept