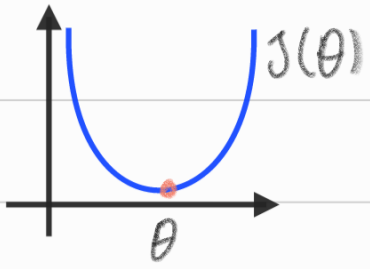
Normal Equation 정규 방정식

- Method to solve θ analytically



$J(\theta) = a\theta^2 + b\theta + c \quad (\theta \in \mathbb{R})$

미분값이 0 일때, θ 도출

$J(\theta_0, \theta_1, \cdots, \theta_m) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 \quad (\theta \in \mathbb{R}^{n+1})$

편미분 활용 $\frac{\partial}{\partial \theta_j} J(\theta) = \cdots = 0$ (for every j)

solve for $\theta_0, \theta_1, \cdots \theta_n$

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

Gradient Descent	Normal Equation
- Need to choose α	- No need to choose α
- Needs many iter.	- Don't need to iterate
- Works well even when n is large	- Need to compute $(X^T X)^{-1}$ - Slow if n is very large

∴ feature가 많지 않으면 NE도 좋은 방법 (1000↓)

Normal Equation Noninvertibility

$\theta = (X^TX)^{-1}X^Ty$ 에서 X^TX 가 역행렬이 없다면?

A. 역행렬이 없는 경우는 매우 드물지만, 그렇더라도

pinv() <psendo-inv> 사용하면 계산 (in octave)

원인 ① Redundant features

② Too many features ($m<n$) 로 인해 발생