

$$\theta_j^{(0)} = \theta_j^0 \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^m (h_0(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

$$\left[1 - \alpha \frac{\lambda}{m} < 1\right] \quad \hookrightarrow 0.99 \quad \text{Just little bit less than 1.}$$

$$\theta_j^{(0)} = \underbrace{\theta_j^0 \left(1 - \alpha \frac{\lambda}{m}\right)}_{0.99 \approx 1} - \underbrace{\alpha \frac{1}{m} \sum_{i=1}^m (h_0(x^{(i)}) - y^{(i)}) x_j^{(i)}}_{\text{Original grad-desc. formula without regularization}}$$

$$0.99 \approx 1$$

$$0.99 \theta_j^0$$

Original grad-desc.
formula without
regularization

Normal Eqⁿ (with Regularization)

$$X = \begin{bmatrix} (x^{(1)})^T \\ \vdots \\ (x^{(m)})^T \end{bmatrix}_{m \times (n+1)} \quad y = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix}_{m \times 1} \Rightarrow \mathbb{R}^m$$

$$\min_{\theta} J(\theta)$$

$$\theta = (X^T X + \lambda \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix})^{-1} X^T y$$

$$(n+1)(n+1)$$

$$\text{eg) } n=2 \Rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$