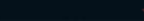


Lecture.7 Linear/Logistic Regression(2) - Linear Regression



Model and Tensors

$$J = \frac{1}{N} \sum_{i=1}^N J_0^{(i)}$$

$$\therefore \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}_0} \cdot \frac{\partial \mathbf{z}_0}{\partial \mathbf{z}_1} \cdot \frac{\partial \mathbf{z}_1}{\partial \mathbf{w}}$$

$$(1, 0) \cdot (0, 1) \cdot (1, \ln)$$

$$= (1, \ln)$$

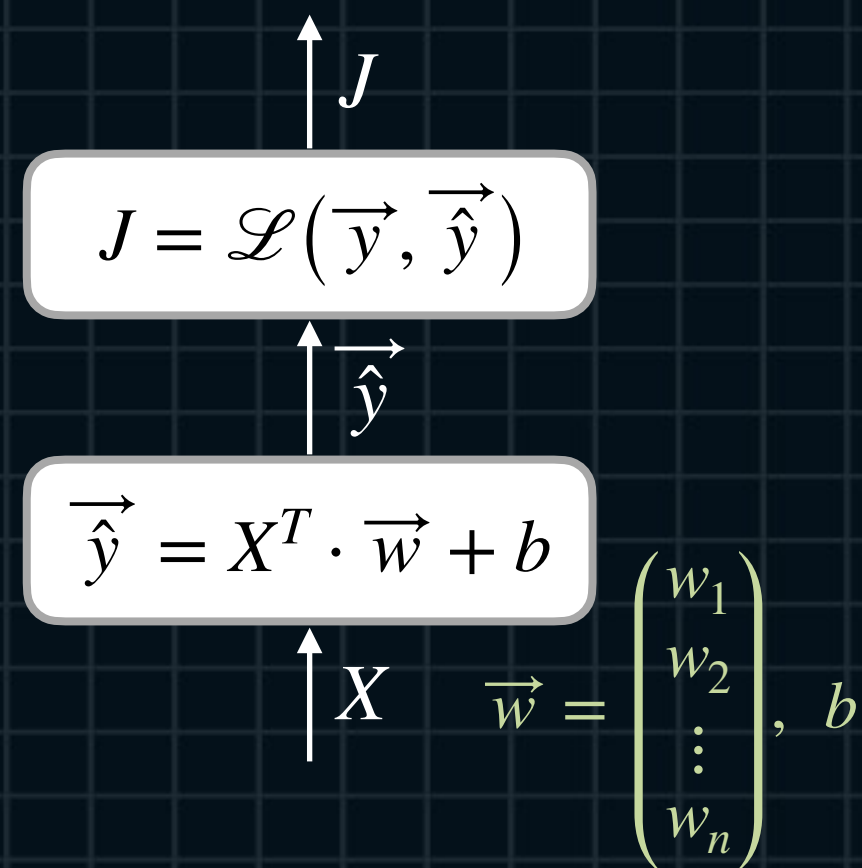
$$\cdot \frac{\partial J}{\partial b} = \frac{\partial J}{\partial x} \cdot \frac{\partial x}{\partial y} \cdot \frac{\partial y}{\partial b}$$

$$= (1, 1)$$

$$\vec{J}_0 = \begin{pmatrix} (y^{(1)} - \hat{y}^{(1)})^2 \\ (y^{(2)} - \hat{y}^{(2)})^2 \\ \vdots \\ (y^{(N)} - \hat{y}^{(N)})^2 \end{pmatrix} = \begin{pmatrix} J_0^{(1)} \\ J_0^{(2)} \\ \vdots \\ J_0^{(N)} \end{pmatrix}$$

$$\vec{\hat{y}} = X^T \cdot \vec{w} + b = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_{l_I}^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_{l_I}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(N)} & x_2^{(N)} & \dots & x_{l_I}^{(N)} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_{l_I} \end{pmatrix} + b = \begin{pmatrix} \hat{y}^{(1)} \\ \hat{y}^{(2)} \\ \vdots \\ \hat{y}^{(N)} \end{pmatrix}$$

$$X^T = \begin{pmatrix} \longleftarrow & (\overrightarrow{x^{(1)}})^T & \longrightarrow \\ \longleftarrow & (\overrightarrow{x^{(2)}})^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & (\overrightarrow{x^{(N)}})^T & \longrightarrow \end{pmatrix} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_{l_i}^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_{l_i}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(N)} & x_2^{(N)} & \dots & x_{l_i}^{(N)} \end{pmatrix}$$



$$\frac{\partial J}{\partial w} = -\frac{2}{N} \sum_{i=1}^N x_i^{(n)} (y_i^{(n)} - \hat{y}_i^{(n)})$$

• 1 sample의 Gradient를 N 번 더함
 • 1 sample의 Gradient를 더함
 Gradient를 update 시키는 방법
 1 batch를 더함 (Mini-batch를 사용함)

$$J = (y - \hat{y})^2 = (y - (x^T w + b))^2$$

$$\frac{\partial J}{\partial w} = -2 \cdot x^T \cdot (y - \hat{y})$$

$$\frac{\partial J}{\partial b} = \frac{\partial J}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial b} = -2 \cdot (y - \hat{y}) \cdot 1 = -2(y - \hat{y})$$