

Gradient descent.

$$W^{(t+1)} = W^t + \eta \frac{dW}{dt} \quad \text{for hill } \text{---} \text{ for valley}$$

move than 1 variable. $W^{(t+1)} = W^t + \eta \nabla Q(W)$.

$$\text{minimize } RSS = \min [\sum (y_i - (w_0 + w_1 x_i))^2]$$

multi regression: Treat Power Term as feature.

$$\text{break sine component. } \sin(at-b) = \sin(at)\cos b - \cos(at)\sin b$$

seasonality $\xrightarrow{\text{new constant}}$ new constant

$$y_i = w_0 h_0(x_i) + w_1 h_1(x_i) + \dots + w_n h_n(x_i) + \epsilon$$

$$\begin{aligned} \text{multi input. } Y &= W_0 \hat{x}_0 + W_1 \hat{x}_1 + \dots + W_n \hat{x}_n \\ &= [\hat{x}] \cdot W + \epsilon \end{aligned}$$

$$RSS = (Y - HW)^T (Y - HW)$$

$$\nabla RSS = -2Y^T (Y - HW).$$

$$-2H^T Y + 2H^T H W = 0 \quad H^T \text{ can be not invertible}$$

$$W = (H^T H)^{-1} H^T Y \rightarrow O(n^3) !!!$$

least square: column space:

$$A = \begin{bmatrix} \vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \end{bmatrix} \quad \text{span}(\vec{v}_1 \dots \vec{v}_n) \text{ linear combination.}$$

Gradient descent $W^{(t)} + \eta (-2H^T)(Y - HW)$.

$$RSS(W) = \sum (y_i - h^T W)^2$$

$$= \sum (y_i - w_0 h_0 - w_1 h_1 - w_2 h_2 - \dots - w_n h_n)^2$$

$$\frac{dRSS}{dW_i} = 2(y_i - w_0 h_0 - w_1 h_1 - \dots - w_n h_n) \cdot (-h_i)$$

$$= 2 \sum (y_i - h^T W) \cdot (-h_i)$$

Review LA.

$$W^{t+1} = W^t + \eta \underline{2h_i} (\sum y_i - h^T W)$$

proof. "feature same"

"uncertain"

$$\xrightarrow{\text{weight}} \# \text{ features} \quad , \quad \xrightarrow{W} \# \text{ features}$$

$$H = \# \text{ trials} \quad \left[\quad \right] \quad \left[\quad \right] \quad \left[\quad \right] \quad \xrightarrow{W} \# \text{ features}$$





































































































































































































