Linear Regression with Matrices

Create original matrices

$$[x] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix} \qquad [y] = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$

independent variable

dependent variable

weights

Hypothesis Function

$$h(x_i) = \theta_0 + \theta_1 x_i$$

Equivalent

Create a new x matrix

$$[x] = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ & \vdots \\ 1 & x_m \end{bmatrix}$$

Multiply $[x]$ by $[\theta]$ ← Dot product

$$[x][\theta] = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} \theta_0 + \theta_1 x_1 \\ \theta_0 + \theta_1 x_2 \\ \theta_0 + \theta_1 x_3 \\ \vdots \\ \theta_0 + \theta_1 x_m \end{bmatrix}$$

Cost Function

$$\frac{1}{2m}\sum_{i=1}^{m}\Big(\underbrace{h(x_i)-y}_{\text{error}}\Big)^2$$

Do this calculation in several steps:

Error $h(x_i)-y$

$$[X][\theta]-[y]=\begin{bmatrix}\theta_0+\theta_1 X_1-y_1\\ \theta_0+\theta_1 x_2-y_2\\ \theta_0+\theta_1 x_3-y_3\\ \vdots\\ \theta_0+\theta_1 X_m-y_m\end{bmatrix}=\begin{bmatrix}e_1\\ e_2\\ e_3\\ \vdots\\ e_m\end{bmatrix}=[e]$$

Sum of Error^2 $\sum_{i=1}^{m} e^2$

Dot product

$$[e]^2=\underbrace{[e_1\ e_2\ e_3\cdots e_m]}_{\text{transpose of }[e]}\begin{bmatrix}e_1\\ e_2\\ e_3\\ \vdots\\ e_m\end{bmatrix}$$

equivalent

$$=e_1^2+e_2^2+e_3^2+\cdots+e_m^2$$

Cost $\frac{1}{2m}\sum e^2=\boxed{\frac{1}{2m}[e]^2}$

Gradient Descent $\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum (h(x_i)-y_i)(x_{0i})$

$h(x_i)-y_i = [e] \leftarrow$ error matrix from above

$x_{0i} \Leftarrow$ first colum from x matrix

dot product

$$\sum (h(x_i) - y_i)(x_{0i}) = [e_1\ e_2\ e_3 \cdots e_m] \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

transpose of error matrix

Equivalent

$$= e_1 + e_2 + e_3 + \cdots + e_m$$

Gradient Descent $\theta_1 = \theta_1 - \alpha \frac{1}{m} \sum ((h(x_i) - y_i)(x_{1i})$

$h(x_i)-y_i = [e] \leftarrow$ error matrix from above

$x_{1i} \Leftarrow$ second column from x matrix

dot product

$$\sum (h(x_i) - y_i)(x_{1i}) = [e_1\ e_2 e_3 \cdots e_m] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix}$$

transpose of error matrix

Equivalent

$$= e_1 x_1 + e_2 x_2 + \cdots + e_m x_m$$