Matrix manipulation

18 February 2026 10:53 PM

sample feature matrix:

$$X = \begin{bmatrix} 100 & 1 \\ 200 & 2 \\ 300 & 3 \end{bmatrix}$$

For,

```
X_bias = np.hstack ([np.ones((X.shape[0],1)), X])
```

np.ones((X.shape[0], 1))

creates a matrix of RXC = X.shape[0] X 1

for our sample: X.shape[0] = 3

$$\therefore \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

np.stack(...., X)

horizontally stacks the ones matrix with the X matrix

$$\begin{bmatrix} 1 & 100 & 1 \\ 1 & 200 & 2 \\ 1 & 300 & 3 \end{bmatrix}$$

```
# y_hat = np.add(X.dot(w[1:]), w[0])
```

↳ hypothesis function

used to compute the predicted values

used to compute the predicted values

loss = y_hat - y → R×1 matrix of prediction errors

sample loss: $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$

grad = x_bias . T . dot(loss)

X_bias.T → transpose the x bias matrix : ensures correct dimensions for multiplication of matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 100 & 200 & 300 \\ 1 & 2 & 3 \end{bmatrix}$$

x_bias.T dot(loss) → matrix multiplication of x_bias with loss matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 100 & 200 & 300 \\ 1 & 2 & 3 \end{bmatrix}_{3\times3} \times \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}_{3\times1}$$

resultant grad = $\begin{bmatrix} 2+3+5 \\ 200+600+1500 \\ 2+6+15 \end{bmatrix}_{3\times1}$ → $\frac{\partial}{\partial w_0}(E)$, $\frac{\partial}{\partial w_1}(E)$, $\frac{\partial}{\partial w_2}(E)$