Notation and Math

Prediction for a linear regression model

$$\mathbf{x} = \begin{bmatrix} 0.04 \\ 0.05 \\ 0.06 \\ 0.02 \\ -0.04 \\ -0.03 \\ 0.04 \\ 0.00 \\ 0.02 \\ -0.02 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 152 \\ -16 \\ -254 \\ 560 \\ 278 \\ -393 \\ 97 \\ -19 \\ 179 \\ 630 \\ 114 \end{bmatrix} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \\ w_6 \\ w_7 \\ w_8 \\ w_9 \\ w_{10} \end{bmatrix}$$

- Predict the value of $\mathbf{x}$:
  - Straightforward calculation:
    $\hat{y} = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6 + w_7x_7 + w_8x_8 + w_9x_9 + w_{10}x_{10}$
  - Inner product:
    - augment $\mathbf{x}$ with a 1, then compute $\hat{y} = \mathbf{w}^T\mathbf{x}$

$$\begin{bmatrix} 152 & -16 & -254 & 560 & 278 & -393 & 97 & -19 & 179 & 630 & 114 \end{bmatrix} \begin{bmatrix} 1 \\ 0.04 \\ 0.05 \\ 0.06 \\ 0.02 \\ -0.04 \\ -0.03 \\ 0.04 \\ 0.00 \\ 0.02 \\ -0.02 \end{bmatrix}$$

    - Remove w_0 from $\mathbf{w}$, then compute $\hat{y} = w_0 + \mathbf{w}'^T\mathbf{x}$

$$152+ \quad \begin{bmatrix} -16 & -254 & 560 & 278 & -393 & 97 & -19 & 179 & 630 & 114 \end{bmatrix} \begin{bmatrix} 0.04 \\ 0.05 \\ 0.06 \\ 0.02 \\ -0.04 \\ -0.03 \\ 0.04 \\ 0.00 \\ 0.02 \\ -0.02 \end{bmatrix}$$

Feature vector, data matrix augmented with a 1
Merge bias/intercept into coefficient/weight vector

$x_0 = 1$

$$\mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_1^{(1)} & \dots & x_d^{(1)} \\ 1 & x_1^{(2)} & \dots & x_d^{(2)} \\ \vdots & \vdots & & \vdots \\ 1 & x_1^{(N)} & \dots & x_d^{(N)} \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_d \end{bmatrix}$$

Design matrix: $X \in \mathbb{R}^{N\times(d+1)}$
Weight vector: $\mathbf{w} \in \mathbb{R}^{d+1}$
Feature vector: $\mathbf{x} \in \mathbb{R}^{d+1}$

$$\hat{y}^{(i)} = w_0 + w_1 \cdot x_1^{(i)} + w_2 \cdot x_2^{(i)} + \cdots + w_d \cdot x_d^{(i)} = [w_0 \quad w_1 \quad w_2 \quad \dots \quad w_d] \begin{bmatrix} x_0^{(i)} \\ x_1^{(i)} \\ x_2^{(i)} \\ \vdots \\ x_d^{(i)} \end{bmatrix} = \mathbf{w}^T \mathbf{x}^{(i)}$$

$$\hat{\mathbf{y}} = X \cdot \mathbf{w}$$

Inner Product vs Outer Product

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$$

- Inner product

$$\mathbf{w}^T\mathbf{x} = [3 \quad 4 \quad 5 \quad 6] \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = 3 \cdot 1 + 4 \cdot 2 + 5 \cdot 3 + 6 \cdot 4$$

- Outer product

$$\mathbf{w}\mathbf{x}^T = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix} [1 \quad 2 \quad 3 \quad 4] = \begin{bmatrix} 3 \cdot 1 & 3 \cdot 2 & 3 \cdot 3 & 3 \cdot 4 \\ 4 \cdot 1 & 4 \cdot 2 & 4 \cdot 3 & 4 \cdot 4 \\ 5 \cdot 1 & 5 \cdot 2 & 5 \cdot 3 & 5 \cdot 4 \\ 6 \cdot 1 & 6 \cdot 2 & 6 \cdot 3 & 6 \cdot 4 \end{bmatrix}$$