Linear Regression: Analytic Solution *(find and work with a partner)*

In the case of *multiple* linear regression, each example has p features. We still have we an associated response (output) variable y for each example. We typically add a "fake 1" to the features of each example, so that $\boldsymbol{x} = [1\ x_1\ x_2 \cdots x_p]^T$. Our model is now the dot product between the weight vector $\boldsymbol{w}$ and the features:

$$h_{\boldsymbol{w}}(\boldsymbol{x}) = w_0 + w_1 x_1 + \cdots w_p x_p = \boldsymbol{w} \cdot \boldsymbol{x}$$

If we consider $\boldsymbol{X}$ to be the entire matrix of features (with dimensions $n \times (p+1)$) then we can reframe the linear regression problem in terms of vectors and matrices. In this case the analytic solution for the weight vector is:

$$\hat{\boldsymbol{w}} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y}$$

1. *Warmup.* Given matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ below, compute both $\boldsymbol{AB}$ and $\boldsymbol{BA}$. Is matrix multiplication commutative?

$$\boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{bmatrix} 0 & 2 \\ -4 & 1 \end{bmatrix}$$

2. Going back to our small example from the last handout, we will use this matrix/vector form to verify our solution for $\boldsymbol{w}$ (even though $p = 1$ we can still use the multiple linear regression derivation). We again have the data: $(x_1, y_1) = (1, 0)$ and $(x_2, y_2) = (0, 1)$. First, rewrite $\boldsymbol{X}$ with the added column of 1's. Then write $\vec{y}$ as a vector.

3. Now use the matrix/vector analytic solution to verify the weight vector $\boldsymbol{w}$.