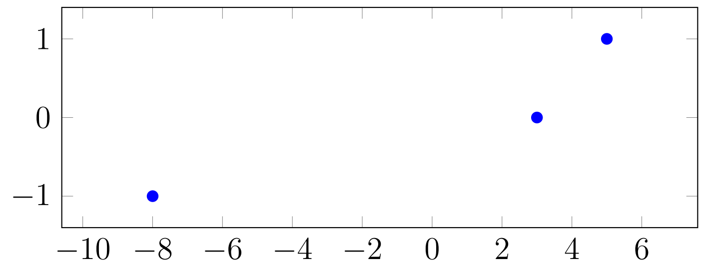
Least Squares Regression **Math 422 - Workshop**

Suppose we have three data points with x and y coordinates given by the following table:

x	y
3	0
5	1
-8	-1



1. Use the formula $r = \left(\frac{x-\bar{x}}{\|x-\bar{x}\|}\right)^T \left(\frac{y-\bar{y}}{\|y-\bar{y}\|}\right)$ to show that the correlation between x and y is $\frac{13}{14}$. Hint: To make your life a little easier, $\bar{x} = \bar{y} = 0$ here.

2. Consider a line $\hat{y} = a + bx$. Using the vector x, write $\hat{y}$ as vector with three entries. Each entry should be a function of a and b.

3. Find a formula for the sum of squared residuals $\|\hat{y} - y\|^2$. Your formula should be a function of a and b.

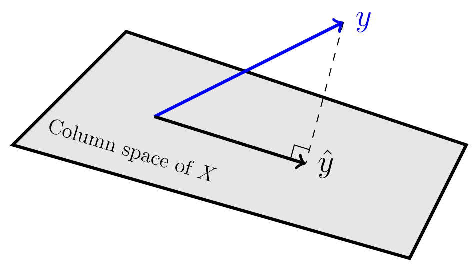
4. Differentiate the sum of squared residuals with respect to a and b to find the best fit regression line.

There is a different way to approach least squares regression using linear algebra instead of calculus. The vector $\hat{y}$ is a linear transformation of the two variables a and b, so you can write this as a matrix equation:

$$\hat{y} = \begin{bmatrix} 1 & 3 \\ 1 & 5 \\ 1 & -8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

The goal is to find the value of $\hat{y}$ that is the closest to y. But $\hat{y}$ is stuck in the range (i.e., column space) of the matrix

$$X = \begin{bmatrix} | & | \\ 1 & x \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 1 & 5 \\ 1 & -8 \end{bmatrix}.$$



To find $\begin{bmatrix} a \\ b \end{bmatrix}$, we need find values of a and b that make the vector $\hat{y} - y$ orthogonal to the column space of X. This is the same as requiring

$$X^T(\hat{y} - y) = 0,$$

because that just means that $\hat{y} - y$ is orthogonal to every column of X. Rewriting this equation, we get the **normal equation** for linear regression:

$$X^T X \begin{bmatrix} a \\ b \end{bmatrix} = X^T y.$$

5. Calculate $X^T X$ and $X^T y$. Remember that y is the vector $\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

6. Use linear algebra to solve the normal equation and find $\begin{bmatrix} a \\ b \end{bmatrix}$.