Mathematics 327
Introduction to least squares

Open the Jupyter notebook on `CoCalc` called `February-08`. That notebook defines a function `QR` that will be helpful in this activity.

1. Find the QR factorization of the matrix $A = \begin{bmatrix} 1 & 2 \\ 1 & -2 \\ -2 & 1 \end{bmatrix}$. What are the matrices Q and R?

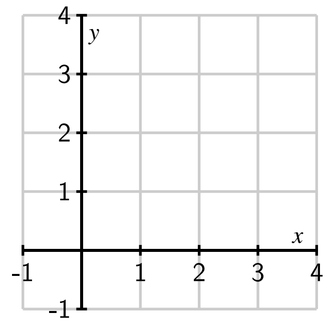
   Find Q^TQ and explain how this shows that the columns of Q are orthonormal.

   Find QQ^T and use it to find the orthogonal projection of $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$ onto $\mathrm{Col}(A)$.

   What is $\mathrm{rank}(QQ^T)$ and why does this make sense geometrically?

2. **Introduction to least squares** Suppose we have data points $(1, 1)$, $(2, 1)$, and $(3, 3)$ and that we would like to find a line passing through all three points.

   Plot the points below. Is there a line passing through the three points?



Let's write the equation of a line as $y = b + mx$. If a line passes through the point $(1, 1)$, we know that $y = 1$ when $x = 1$, which tells us that $b + m(1) = b + m = 1$. Write similar equations using the other two data points.

We now have a linear system with three equations in the unknowns b and m. Write this in the form $A\mathbf{x} = \mathbf{b}$ and note that the unknown vector $\mathbf{x} = \begin{bmatrix} b \\ m \end{bmatrix}$ describes a line.

Is your linear system consistent? What does this say about the existence of a line passing through the three points?

In spite of the fact that our linear system has no solution, we would like to find the vector $\mathbf{x}$ that is as close as possible to being a solution. That is, we would like to find $\mathbf{x}$ so that $A\mathbf{x}$ is as close to $\mathbf{b}$ as possible. For this reason, find the orthogonal projection $\widehat{\mathbf{b}}$ of $\mathbf{b}$ onto $\mathrm{Col}(A)$.

Since $\widehat{\mathbf{b}}$ is in Col(A), the equation $A\mathbf{x} = \widehat{\mathbf{b}}$ must be consistent. Find its solution $\mathbf{x} = \begin{bmatrix} b \\ m \end{bmatrix}$ and then sketch the line $y = b + mx$ along with the data points above.

3. Let's think about this process more generally. Suppose that $A\mathbf{x} = \mathbf{b}$ is inconsistent, but that we would like to find $\mathbf{x}$ that is as close as possible to $\mathbf{b}$. This means that $A\mathbf{x}$ should equal $\widehat{\mathbf{b}}$, the orthogonal projection of $\mathbf{b}$ onto Col(A).

   This means that $\mathbf{b} - A\mathbf{x}$ should be orthogonal to Col(A). Said another way, $\mathbf{b} - A\mathbf{x}$ is in Col$(A)^\perp$. Explain why
$$A^T(\mathbf{b} - A\mathbf{x}) = \mathbf{0}.$$

   Rearrange this condition to find the *normal equations*
$$A^TA\mathbf{x} = A^T\mathbf{b}.$$

   For the problem of finding the line through three points, write the normal equations.

   Find the solution to this system and check that it is the same that you found earlier.

   Suppose that you want to find the line through the points $(1, 2)$, $(3, 3)$, and $(5, 4)$. How does this change the problem?