Least Squares

We want to find a line/curve that minimizes the sum of the square of the error caused due to deviation.

Least squares

The least squares for $Ax = b$ is $\hat{x}$ for which,

$$||b - A\hat{x}|| \leq ||b - Ax||$$

for all x

We can use the normal equation to solve for $\hat{x}$

Normal Equation

$$A^T A\hat{x} = A^T\vec{b}$$

Manipulating this we can get this,

$$\hat{x} = (A^T A)^{-1} A^T\vec{b}$$

Using QR Factorization

$$R\hat{x} = Q^T\vec{b}$$

Proof

$$\begin{aligned} A^T A\hat{x} &= A^T\vec{b} \\ (QR)^T QR\hat{x} &= (QR)^T\vec{b} \\ R^T Q^T QR\hat{x} &= R^T Q^T\vec{b} \\ R^T R\hat{x} &= R^T Q^T\vec{b} \\ R\hat{x} &= Q^T\vec{b} \end{aligned}$$

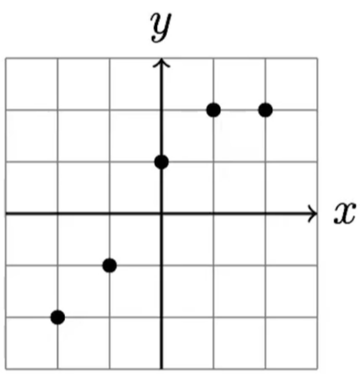
How to solve

1. Construct QR Factorization
2. Solve $R\hat{x} = Q^T\vec{b}$

Line

Say we want to find a line $y = mx + b$ that is the best fit for the following points:

x	-2	-1	1	1	2
y	-2	-1	1	2	2



We can create a list of linear equations using this:

$$\begin{gathered} m(-2) + b = -2 \\ m(-1) + b = -1 \\ m(1) + b = 1 \\ m(1) + b = 2 \\ m(2) + b = 2 \end{gathered}$$

We can turn this in to a matrix equation like so,

$$A\vec{x} = \vec{b}$$

$$\begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

Compute QR,

$$\begin{bmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{10}} \\ \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{5}} & 0 \\ 0 & \frac{10}{\sqrt{10}} \end{bmatrix}$$

Now compute $Q^T\vec{b}$,

$$Q^T\vec{b} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{11}{\sqrt{10}} \end{bmatrix}$$

Finally solve $R\hat{x} = Q^T\vec{b}$,

$$\begin{bmatrix} \frac{5}{\sqrt{5}} & 0 \\ 0 & \frac{10}{\sqrt{10}} \end{bmatrix} \hat{x} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{11}{\sqrt{10}} \end{bmatrix}$$

$$\hat{x} = \boxed{\begin{bmatrix} \frac{2}{5} \\ \frac{11}{10} \end{bmatrix}}$$

So we get $\boxed{y = \dfrac{2}{5} + \dfrac{11}{10}x}$