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,
>
> $$\hat{x} = (A^T A)^{-1} A^T\vec{b}$$
>
> > 🔥 **Proof** ⌄
> >
> > $\hat{x}$ is a least squares solution of $Ax = b$
> >
> > $\Longleftrightarrow A\hat{x} - b$ is as small as possible
> >
> > $\Longleftrightarrow A\hat{x} - b$ is orthogonal to $\mathrm{Col}(A)$
> >
> > $\overset{\text{FTLA}}{\Longleftrightarrow} A\hat{x} - b$ is in $\mathrm{Nul}(A^T)$
> >
> > $\Longleftrightarrow A^T(A\hat{x} - b) = 0$
> >
> > $\Longleftrightarrow A^T A\hat{x} = A^T b$

> ✏️ **Using QR Factorization** ⌄

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

How to solve

Method 1

1. Construct [QR Factorization](#)
2. Solve $R\hat{x} = Q^T\vec{b}$
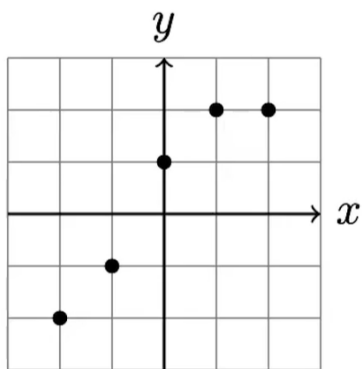
Method 2

1. Solve the normal equation for $\hat{x}$.

Method 3

1. [Use Mean Deviation](#)

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:

$$m(-2) + b = -2$$
$$m(-1) + b = -1$$
$$m(1) + b = 1$$
$$m(1) + b = 2$$
$$m(2) + b = 2$$

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

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

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

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

Curves

We can use this method to fit data to a curve using the function,

$$y = c_0 + c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x)$$

Lets take an example,
Say we want to modal:

x	-1	0	0	1
y	2	1	0	6

Using $y = c_1 x + c_2 x^2$

$$-c_1 + c_2 = 2$$
$$0c_1 + 0c_2 = 1$$
$$0c_1 + 0c_2 = 0$$
$$c_1 + c_2 = 6$$

Now we can use [one of the methods](#) to solve for c_1 and c_2