"Least Squares" Fitting Formulas to Data

- Most common application of Normal Equation.
- Find the "best fitting" formula of given form for some data.

MATLAB Example

\>> x = -20 : 0.1 : 20 ; (Generate vector of x-values)

\>> x = x' ; (Transpose x to be a column)

\>> y = 3 * x.^2 + x - 2 ; ($y = 3x^2 + x - 2$)

\>> y = y + (rand(size(y)) - 0.5) * 1000 ;

(Insert random error into y)

\>> scatter(x, y) (plot data with error)

(Matrix A in Normal Eqn)

\>> A = [x.^2 x x.^0] ;

\>> (A' * A) \ (A' * y) (Solve Normal Eqn $\hat{x} = \cancel{\frac{A^Tb}{A^TA}}$)

Answer should be approx. [3 1 -2]

\>> hold on

\>> plot(x, ans[1] * x.^2 + ans[2] * x + ans[3])

Goal:

Given data

x	y
x_0	y_0
x_1	y_1
x_2	y_2
⋮	⋮

and formula with unknown constants $y = \cdots$

Find value for constants so that the formula predicts values as close to the data as possible.

EX: (Best line through data points)

Find the best line through the points

$(-1,-2) \quad (0,1) \quad (1,2) \quad (2,5)$

Data:

x	y
-1	-2
0	1
1	2
2	5

plug into formula to get system of equations

Formula: $ax + b = y$

$$a(-1) + b = (-2)$$
$$a(0) + b = 1$$
$$a(1) + b = 2$$
$$a(2) + b = 5$$

Matrix Equation

$$\begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 5 \end{bmatrix}$$

$x \quad 1 \quad\quad y$

Normal Equation

$$\begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 14 \\ 6 \end{bmatrix}$$

Solution

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 4 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 14 \\ 6 \end{bmatrix} = \begin{bmatrix} 11/5 \\ 2/5 \end{bmatrix}$$

$$\boxed{\text{Best Line: } y = \tfrac{11}{5} x + \tfrac{2}{5}}$$

EX: (Best quadratic through data points)

Find the best quadratic $y = ax^2 + bx + c$ through

$(-1,-2) \quad (0,-1) \quad (1,0) \quad (2,0)$

Data:

x	y
-1	-2
0	-1
1	0
2	0

Formula:

$ax^2 + bx + c = y$

$a(-1)^2 + b(-1) + c = (-2)$

$a(0)^2 + b(0) + c = (-1)$

$a(1)^2 + b(1) + c = 0$

$a(2)^2 + b(2) + c = 0$

Matrix:

$$\begin{array}{cccc} x^2 & x & 1 & \\ \end{array} \qquad y$$

$$\begin{bmatrix} (-1)^2 & -1 & 1 \\ 0 & 0 & 1 \\ 1^2 & 1 & 1 \\ 2^2 & 2 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

Note: This will be easier to solve if we change order

$$\begin{array}{c} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \\ \begin{array}{ccc} 1 & x & x^2 \end{array} \end{array} \begin{bmatrix} c \\ b \\ a \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 0 \\ 0 \end{bmatrix} \qquad (c + bx + ax^2 = y)$$

Next Year: Change this example!

Normal Equation

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix}\begin{bmatrix} \hat{c} \\ \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \\ -2 \end{bmatrix}$$

Solve with **LU decomposition**:

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 & 6 \\ 0 & 5 & 5 \\ 0 & 5 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 1 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 & 6 \\ 0 & 5 & 5 \\ 0 & 0 & 4 \end{bmatrix}$$

(Divide by L)

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & 1 & 1 \end{bmatrix} \Big\backslash \begin{bmatrix} -3 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} -3 \\ 7/2 \\ -1 \end{bmatrix}$$

(Divide by U)

$$\begin{bmatrix} 4 & 2 & 6 \\ 0 & 5 & 5 \\ 0 & 0 & 4 \end{bmatrix} \Big\backslash \begin{bmatrix} -3 \\ 7/2 \\ -1 \end{bmatrix} = \begin{bmatrix} -17/20 \\ 19/20 \\ -1/4 \end{bmatrix} \begin{array}{l} \leftarrow \hat{c} \\ \leftarrow \hat{b} \\ \leftarrow \hat{a} \end{array}$$

$$\boxed{\text{Best Quadratic: } y = -\tfrac{1}{4}x^2 + \tfrac{19}{20}x - \tfrac{17}{20}}$$

(**Note:** When solving by hand, it is usually easier if you write low order terms first

EX $\quad y = a + bx + cx^2$
instead of
$y = ax^2 + bx + c$)

This method can even be used to fit more complicated formulas to data.

EX: Write the normal equation to compute the best fit for $f = a + bx + cy$

to the data

x	1	1	0	-1
y	0	1	1	1
f	0	1	2	-1

Data:

x	y	f
1	0	0
1	1	1
0	1	2
-1	1	-1

Formula:

$a + bx + cy = f$

$a + b\cdot 1 + c\cdot 0 = 0$

$a + b\cdot 1 + c\cdot 1 = 1$

$a + b\cdot 0 + c\cdot 1 = 2$

$a + b(-1) + c\cdot 1 = -1$

Matrix:

$$\begin{array}{ccc} & x & y \end{array} \qquad f$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix}$$

(EX continued)

Normal Equation:

$$\begin{bmatrix} 4 & 1 & 3 \\ 1 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}\begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$$

To compute $\hat{a}, \hat{b}, \hat{c}$ you would now use LU-Decomp...

EX: Write the normal equation to compute the best fit for

$$f = a + b\cos x + c\cos 2x$$

to the data

x	$-\pi/2$	$-\pi/4$	0	$\pi/4$	$\pi/2$
f	-2	-1	0	1	2

Formula

$a + b\cos x + c\cos 2x = f$

$a + b\cdot 0 + c\cdot(-1) = -2$

$a + b\,1/\sqrt{2} + c\cdot 0 = -1$

$a + b\cdot 1 + c\cdot 1 = 0$

$a + b\cdot 1/\sqrt{2} + c\cdot 0 = 1$

$a + b\cdot 0 + c\cdot(-1) = 2$

Matrix

$$\begin{array}{c} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1/\sqrt{2} & 0 \\ 1 & 1 & 1 \\ 1 & 1/\sqrt{2} & 0 \\ 1 & 0 & -1 \end{bmatrix} \\ \quad\ \ \uparrow\ \cos x \quad \uparrow\ \cos(2x) \end{array} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{array}{c} \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \\ \uparrow\ f \end{array}$$

This is called the "discrete cosine transform"

(EX continued)

Normal Equation:

$$\begin{bmatrix} 5 & 2+2/\sqrt{5} & 1 \\ 2+2/\sqrt{5} & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}\begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Ha ha ha...

Orthogonalization and QR-decomposition

Two vectors are called "orthogonal" if

$\mathbf{v}^T\mathbf{w} = 0$ ⇝ (i.e. $\mathbf{v}\cdot\mathbf{w} = 0$ in MAT 120; $\mathbf{v}\perp\mathbf{w}$ "perpendicular")

Matrices whose columns are all orthogonal to each other are very nice.

EX: Solve $\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ **Ans:** $\begin{bmatrix} 6/3 \\ 3/6 \\ -1/2 \end{bmatrix}$

Note: Columns are all orthogonal to each other:

$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = 0 \qquad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = 0 \qquad \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = 0$

The solution to $A\mathbf{x} = \mathbf{b}$ can be computed by solving the normal equation. Usually this does not make things simpler... but if the columns of A are all <u>orthogonal</u> then the normal equation becomes <u>simple</u>!

<u>Equation</u> $\rightsquigarrow$ <u>Normal Equation</u>

$$\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ -1 \end{bmatrix}$$

A^TA is <u>diagonal</u> because columns were <u>orthogonal</u>!

So
$$\hat{x} = 6/3 = 2$$
$$\hat{y} = 3/6 = 1/2$$
$$\hat{z} = -1/2$$

Since the original system has a solution, $\hat{x}, \hat{y}, \hat{z}$ must be the solution.

<u>Formula</u>: If the columns of A are all <u>orthogonal</u>, then the best approximate solution to $A\mathbf{x} = \mathbf{b}$ is

$$\hat{\mathbf{x}} = \begin{bmatrix} \frac{c_1 \cdot b}{c_1 \cdot c_1} \\ \frac{c_2 \cdot b}{c_2 \cdot c_2} \\ \vdots \end{bmatrix} \quad \text{where } A = \begin{bmatrix} | & | & \\ c_1 & c_2 & \cdots \\ | & | & \end{bmatrix}$$

<u>EX</u>: Find best approximate solution to

$$\begin{bmatrix} 1 & -1 & 4 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \\ 3 & 0 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 2 \end{bmatrix}$$

<u>Note</u>: Columns are orthogonal

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = 0 \qquad \begin{bmatrix} 1 \\ 1 \\ 0 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \\ 1 \\ -2 \end{bmatrix} = 0 \qquad \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 2 \\ 1 \\ -2 \end{bmatrix} = 0$$

<u>Best approx. solution</u>:
$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} \frac{1-1+0+6}{1+1+0+9} \\ \frac{-1-1+2+0}{1+1+4+0} \\ \frac{4-2+1-4}{16+4+1+4} \end{bmatrix} = \begin{bmatrix} 6/11 \\ 0 \\ -1/25 \end{bmatrix}$$

<u>EX</u>: Find solution to

$$\begin{bmatrix} 1 & -3 & 1 \\ 2 & 1 & -4 \\ 1 & 1 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

<u>Note</u>: Columns are orthogonal

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} = 0 \qquad \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -4 \\ 7 \end{bmatrix} = 0 \qquad \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -4 \\ 7 \end{bmatrix} = 0$$

Solution = Best Approx. Solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} 4/6 \\ -1/11 \\ 4/66 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -1/11 \\ 2/33 \end{bmatrix}$$

What about matrices whose columns are not all orthogonal to each other?

→ **Orthogonalization** (a.k.a. "Gram-Schmidt")

and **QR-Decomposition.**

Idea: Given a matrix A use column operations to change columns so that they are all orthogonal to each other.

→ New, orthogonal-column matrix $= Q$

→ Triangular matrix of column op.s $= R$

→ Since we are making the columns nicer anyway, we will also divide each column by its length

Def: A matrix is called "Orthogonal" if

(1) All columns are orthogonal to each other

(2) Each column has length $= 1$

Notation: Usually we name matrices $\mathbf{A}$ or $\mathbf{B}$; but orthogonal matrices are so special, we use $\mathbf{Q}$

Note: Being orthogonal means that each column

- has dot product with other columns $= 0$
- has dot product with itself $= 1$

EX: $\begin{bmatrix} 2/3 & 2/3 & 1/3 \\ 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \end{bmatrix}$ is orthogonal

$$c_1 \cdot c_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix} \cdot \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix} = 4/9 + 1/9 + 4/9 = 1$$

$$c_2 \cdot c_2 = \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix} \cdot \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix} = 4/9 + 4/9 + 1/9 = 1$$

$$c_3 \cdot c_3 = \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix} \cdot \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix} = 1/9 + 4/9 + 4/9 = 1$$

$$c_1 \cdot c_2 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix} \cdot \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix} = 4/9 - 2/9 - 2/9 = 0$$

$$c_1 \cdot c_3 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix} \cdot \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix} = 2/9 + 2/9 - 4/9 = 0$$

$$c_2 \cdot c_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix} \cdot \begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix} = 2/9 - 4/9 + 2/9 = 0$$

Note: For orthogonal matrices, $Q^TQ = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \\ \vdots & \vdots & & \end{bmatrix} = I$

so the normal equation is <u>really</u> simple.

Make columns orthogonal to each other using (Gram-Schmidt) <u>Orthogonalization</u>

- <u>Step 1</u> Make each column orthogonal to the ones before it
- <u>Step 2</u> Divide all columns by their length

Begin with a set of vectors (columns of A)

$c_1, c_2, c_3 \ldots$

→ Convert to orthogonal set

$q_1, q_2, q_3 \ldots$

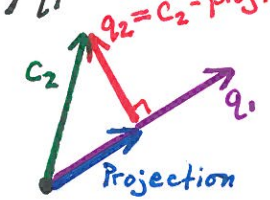
$c_1 \rightsquigarrow q_1 = c_1$ (c₁ doesn't have any columns before it, so it is fine.)

$$c_2 \rightsquigarrow q_2 = c_2 - \begin{pmatrix}\text{projection of} \\ c_2 \text{ onto } q_1\end{pmatrix}$$

$$= c_2 - \left(\frac{q_1 \cdot c_2}{q_1 \cdot q_1}\right) q_1$$

(q_2 = component of c_2 which is $\perp$ to q_1

$= c_2$ − projection)

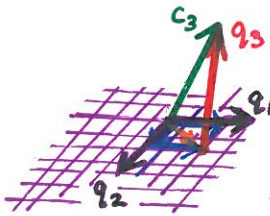


$$c_3 \rightsquigarrow q_3 = c_3 - \begin{pmatrix}\text{projection of} \\ c_3 \text{ onto } c_1 \,\&\, c_2\end{pmatrix}$$

$$= c_3 - \begin{pmatrix}\text{projection of} \\ c_3 \text{ onto } q_1 \,\&\, q_2\end{pmatrix} \quad (*)$$

$$= c_3 - \begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix} \begin{bmatrix} \frac{q_1 \cdot c_3}{q_1 \cdot q_1} \\ \frac{q_2 \cdot c_3}{q_2 \cdot q_2} \end{bmatrix}$$

$$= c_3 - \left(\left(\frac{q_1 \cdot c_3}{q_1 \cdot q_1}\right) q_1 + \left(\frac{q_2 \cdot c_3}{q_2 \cdot q_2}\right) q_2 \right)$$



(*) Projection of c_3 onto $\begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix}$ is easy to calculate because these columns are orthogonal — solution to normal equation for $\begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix} x = c_3$ is $\hat{x} = \begin{bmatrix} \frac{q_1 \cdot c_3}{q_1 \cdot q_1} \\ \frac{q_2 \cdot c_3}{q_2 \cdot q_2} \end{bmatrix}$

So projection is $\begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix} \hat{x}$ written above.

This pattern continues...

$$c_4 \rightsquigarrow q_4 = c_4 - \left(\left(\frac{q_1 \cdot c_4}{q_1 \cdot q_1}\right) q_1 + \left(\frac{q_2 \cdot c_4}{q_2 \cdot q_2}\right) q_2 + \left(\frac{q_3 \cdot c_4}{q_3 \cdot q_3}\right) q_3 \right)$$

Finally, divide all columns by their length.

$$q_1 / \sqrt{q_1 \cdot q_1} \qquad q_2 / \sqrt{q_2 \cdot q_2} \qquad q_3 / \sqrt{q_3 \cdot q_3} \qquad \text{etc}$$

(Gram-Schmidt) Orthogonalization

Begin with set of vectors $c_1, c_2, c_3, \dots$

Step 1: Orthogonalize

$$q_1 = c_1$$

$$q_2 = c_2 - \frac{q_1 \cdot c_2}{q_1 \cdot q_1} q_1$$

(alt: $(q_1 \cdot q_1)\, c_2 - (q_1 \cdot c_2)\, q_1$)

$$q_3 = c_3 - \frac{q_1 \cdot c_3}{q_1 \cdot q_1} q_1 - \frac{q_2 \cdot c_3}{q_2 \cdot q_2} q_2$$

$$q_4 = c_4 - \frac{q_1 \cdot c_4}{q_1 \cdot q_1} q_1 - \frac{q_2 \cdot c_4}{q_2 \cdot q_2} q_2 - \frac{q_3 \cdot c_4}{q_3 \cdot q_3} q_3$$

etc

Step 2: Normalize (divide by length)

$$\bar{q}_1 = q_1 / \sqrt{q_1 \cdot q_1} \qquad \bar{q}_2 = q_2 / \sqrt{q_2 \cdot q_2} \qquad \text{etc}$$

Note: Since we normalize at the end, only the direction of q_k matters

i.e $\begin{bmatrix} 2 \\ 4 \\ -8 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 2/3 \\ -4/3 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1 \\ -2 \end{bmatrix} \rightsquigarrow \bar{q} = \begin{bmatrix} 1/\sqrt{21} \\ 2/\sqrt{21} \\ -4/\sqrt{21} \end{bmatrix}$

So we can simplify computations by "throwing out" common factors from q & "multiplying away" fractions

EX: Orthogonalize (& Normalize)

$$\underset{c_1}{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}} \quad \underset{c_2}{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}} \quad \underset{c_3}{\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}}$$

$$q_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \qquad q_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$q_3 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} - \frac{4}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

ignore multiple

$$= \frac{1}{3} \left(\begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ -4 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix} \right) = \frac{1}{3} \begin{bmatrix} -4 \\ 2 \\ 4 \end{bmatrix} = \cancel{\tfrac{2}{3}} \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$$

$$\bar{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix} \qquad \bar{q}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix} \qquad \bar{q}_3 = \begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$$

(Note: $\sqrt{9} = 3$)

EX: Orthogonalize $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

$$q_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

$$q_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1+2+0+0}{1+4+0+1}\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{3}{6}\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} 2 \\ 2 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}\right) = \cancel{\frac{1}{2}}\begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}$$

$$q_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} - \frac{1+2+0+2}{1+4+0+1}\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} - \frac{1+0+2-2}{1+0+4+1}\begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \end{bmatrix} - \frac{5}{6}\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 0 \\ -4 \\ 4 \\ 8 \end{bmatrix} = \cancel{\frac{4}{6}}\begin{bmatrix} 0 \\ -1 \\ 1 \\ 2 \end{bmatrix}$$

Check that $\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ $\begin{bmatrix} 1 \\ 0 \\ 2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ -1 \\ 1 \\ 2 \end{bmatrix}$ are all orthog.

$$\bar{q}_1 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \end{bmatrix} \qquad \bar{q}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 0 \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} \qquad \bar{q}_3 = \begin{bmatrix} 0 \\ -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

$$Q = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 0 \\ 2/\sqrt{6} & 0 & -1/\sqrt{6} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \end{bmatrix}$$

What about **QR - decomposition**?

If $A = QR$ where Q is orthogonal

$$Q^TA = \cancel{Q^TQ}R$$

$$Q^TA = R$$

dot products of columns of A and columns of Q

$$R = \begin{bmatrix} q_1 \cdot c_1 & q_1 \cdot c_2 & q_1 \cdot c_3 & \cdots \\ \cancel{q_2 \cdot c_1}\ 0 & q_2 \cdot c_2 & q_2 \cdot c_3 & \cdots \\ \cancel{q_3 \cdot c_1}\ 0 & \cancel{q_3 \cdot c_2}\ 0 & q_3 \cdot c_3 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

Upper Δ matrix!

Note: If you know Q, then computing R is part of solving the normal equation:

$$(A^TA)\hat{x} = (A^Tb)$$

$$\Updownarrow$$

$$(Q^TA)\hat{x} = (Q^Tb)$$

This is how you use Q to solve the normal equation for A!

This is R = upper Δ matrix so dividing by it is easy!

A

Extra note about fitting polynomials to data:

When fitting polynomials to data using least squares, the normal equation has a nice form (which engineers sometimes memorize)...

Data:

x	y
x_1	y_1
x_2	y_2
x_3	y_3
$\vdots$	$\vdots$
x_n	y_n

Points: (x_1, y_1), (x_2, y_2), (x_3, y_3), $\vdots$, (x_n, y_n)

Linear. Equation $a + bx = y$

Matrix
$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$

Normal Equation
$$\begin{bmatrix} n & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \end{bmatrix}$$

dot products of columns of A

dot product of A with y

Quadratic. Equation $a + bx + cx^2 = y$

Matrix
$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$

Normal Equation
$$\begin{bmatrix} n & \sum_i x_i & \sum_i x_i^2 \\ \sum_i x_i & \sum_i x_i^2 & \sum_i x_i^3 \\ \sum_i x_i^2 & \sum_i x_i^3 & \sum_i x_i^4 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \\ \hat{c} \end{bmatrix} = \begin{bmatrix} \sum_i y_i \\ \sum_i x_i y_i \\ \sum_i x_i^2 y_i \end{bmatrix}$$

→ This is useful if you are solving by hand... However, if using a computer calculation engine, like MatLab, it will be *much* faster to create A and then use $(A^T A)$ and $(A^T b)$ because MatLab is programmed with special tricks to compute these *very* quickly.