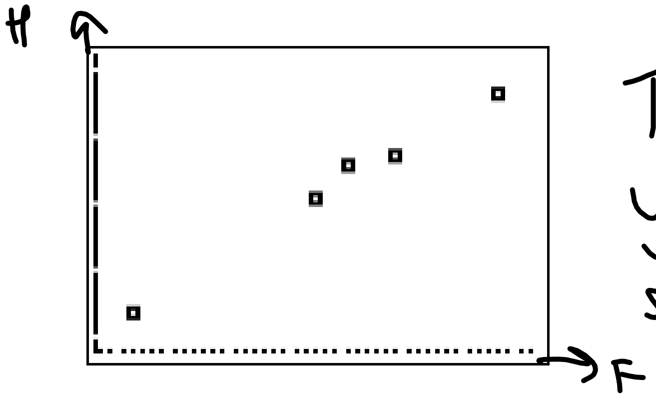
Least Squares: Orthogonal Projections

Femur:	38	56	59	64	74
Humerus:	41	63	70	72	84



There is no single line $y = mx + b$ that is satisfied by all the data points.

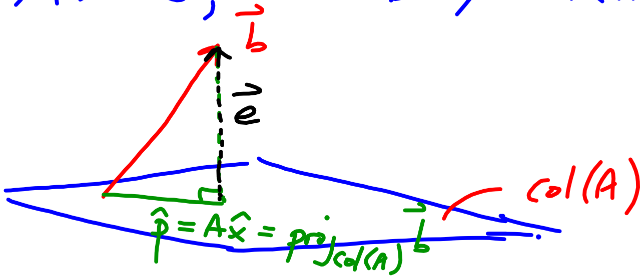
Make it fit: Assume $y = mx + b$ is the line that models this relationship. Then it must be true that:

$$\begin{cases} 38m + b = 41 \\ 56m + b = 63 \\ 59m + b = 70 \\ 64m + b = 72 \\ 74m + b = 84 \end{cases} \rightarrow \begin{bmatrix} 38 & 1 \\ 56 & 1 \\ 59 & 1 \\ 64 & 1 \\ 74 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 41 \\ 63 \\ 70 \\ 72 \\ 84 \end{bmatrix}$$

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

We know from the graph that this is an inconsistent system. Recall: If $A\vec{x} = \vec{b}$ has no solution $\vec{x}$, $\vec{b}$ is not in col(A).

$A\vec{x}=\vec{b}$, and $\vec{b} \notin col(A)$ ...



If $\vec{b}$ is not in $col(A)$, what's the closest we could get to $\vec{b}$ while in $col(A)$? The vector $\vec{p} = proj_{col(A)}\vec{b}$.

Also, since $\vec{p} \in col(A)$, we know $\vec{p} = A\hat{x}$ for some heretofore undetermined vector $\hat{x}$.

- $\vec{b} = \vec{e} + \vec{p} \implies \vec{e} = \vec{b} - \vec{p}$.
$$= \vec{b} - A\hat{x}$$
- $\vec{e} \perp col(A)$, so $\vec{e}$ is in $null(A^T)$.

So $A^T\vec{e} = \vec{0}$

$$A^T(\vec{b} - A\hat{x}) = \vec{0}$$

$$A^T\vec{b} - A^TA\hat{x} = \vec{0}$$

$$A^TA\hat{x} = A^T\vec{b}$$ "normal equations"

Generally speaking, A^T will be invertible only if A is square and has linearly independent columns. In a least-squares situation, A^T is usually not square, hence not invertible, i.e., we usually cannot split A^TA apart. However, A^TA is ALWAYS square and will always be invertible if A has linearly independent columns.

$$\hat{x} = (A^TA)^{-1}A^T\vec{b}$$ "least squares solution"

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

$$\underbrace{\begin{bmatrix} 38 & 1 \\ 56 & 1 \\ 59 & 1 \\ 64 & 1 \\ 74 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m \\ b \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} 41 \\ 63 \\ 70 \\ 72 \\ 84 \end{bmatrix}}_{\vec{b}}$$

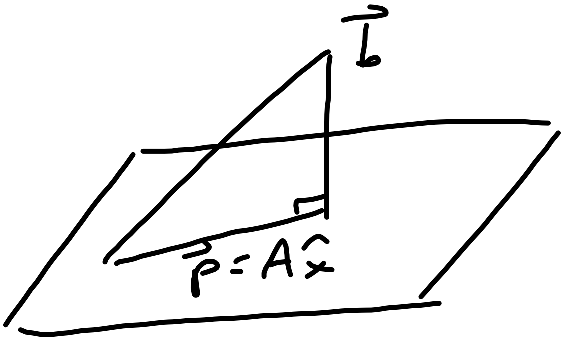
On Calc:

$$([A]^T[A])^{-1}[A]^T[B]$$

$$\hat{x} = \left(\begin{bmatrix} 38 & 56 & 59 & 64 & 74 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 38 & 1 \\ 56 & 1 \\ 59 & 1 \\ 64 & 1 \\ 74 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 38 & 56 & 59 & 64 & 74 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 41 \\ 63 \\ 70 \\ 72 \\ 84 \end{bmatrix}$$

$$= \begin{bmatrix} 17633 & 291 \\ 291 & 5 \end{bmatrix}^{-1} A^T \vec{b} = \begin{bmatrix} 1.196900\ldots \\ -3.659586\ldots \end{bmatrix}$$

So y = 1.1969x - 3.6596



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

$\text{proj}_{\text{col}(A)} \vec{b} = A\hat{x} = \underbrace{A(A^T A)^{-1} A^T}\vec{b}$.

$m \times m$ projection matrix for projecting <u>any</u> vector $\vec{b}$ onto col(A).

(3, −1, 2) onto the subspace V with basis $B = \{(1,1,2), (0,-1,3)\}$

$A = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 2 & 3 \end{bmatrix}$ So $V = \text{col}(A)$.

proj. (3, −1, 2) onto col(A).

$P = \underbrace{A(A^T A)^{-1} A^T}$ for projection $\vec{b}$ onto cols of A.

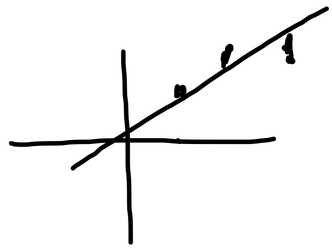
$P = \dfrac{aa^T}{a^T a}$ for projection $\vec{b}$ onto $\vec{a}$.

Calculus:

Suppose $a_1 m = b_1$

$a_2 m = b_2$

$a_3 m = b_3$



What should m be?

Least-Squares:

Minimize sum of squared errors:

$$\sum(E^2) = \sum(a_i m - b_i)^2$$

$$\sum(E^2) = (a_1 m - b_1)^2 + (a_2 m - b_2)^2 + (a_3 m - b_3)^2$$

$$\frac{d(\sum(E^2))}{dm} = 2(a_1 m - b_1)a_1 + 2(a_2 m - b_2)a_2 + 2(a_3 m - b_3)a_3$$

$$= 2a_1^2 m - 2a_1 b_1 + 2a_2^2 m - 2a_2 b_2 + 2a_3^2 m - 2a_3 b_3$$

Min. could only occur when $\frac{d(\sum(E^2))}{dm} = 0 \ldots$

$$2a_1^2 m + 2a_2^2 m + 2a_3^2 m = 2a_1 b_1 + 2a_2 b_2 + 2a_3 b_3$$

$$m = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{a_1^2 + a_2^2 + a_3^2}$$

$$m = \frac{a \cdot b}{a \cdot a} = \frac{a^T b}{a^T a}$$

Least-Squares: QR

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

Since $A = QR$,

$$A^T = (QR)^T = R^T Q^T$$

$$(R^T \underbrace{Q^T) Q}_{I} R\hat{x} = R^T Q^T \vec{b}$$

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

R^T is invertible, so

$$R\hat{x} = Q^T\vec{b}$$ (R is upper Δ).

solvable via back substitution