Orthogonality

CS 556

Erisa Terolli

Orthogonality

Orthogonality

Orthogonal Vectors

Two vectors v and w are orthogonal when their dot product is 0.

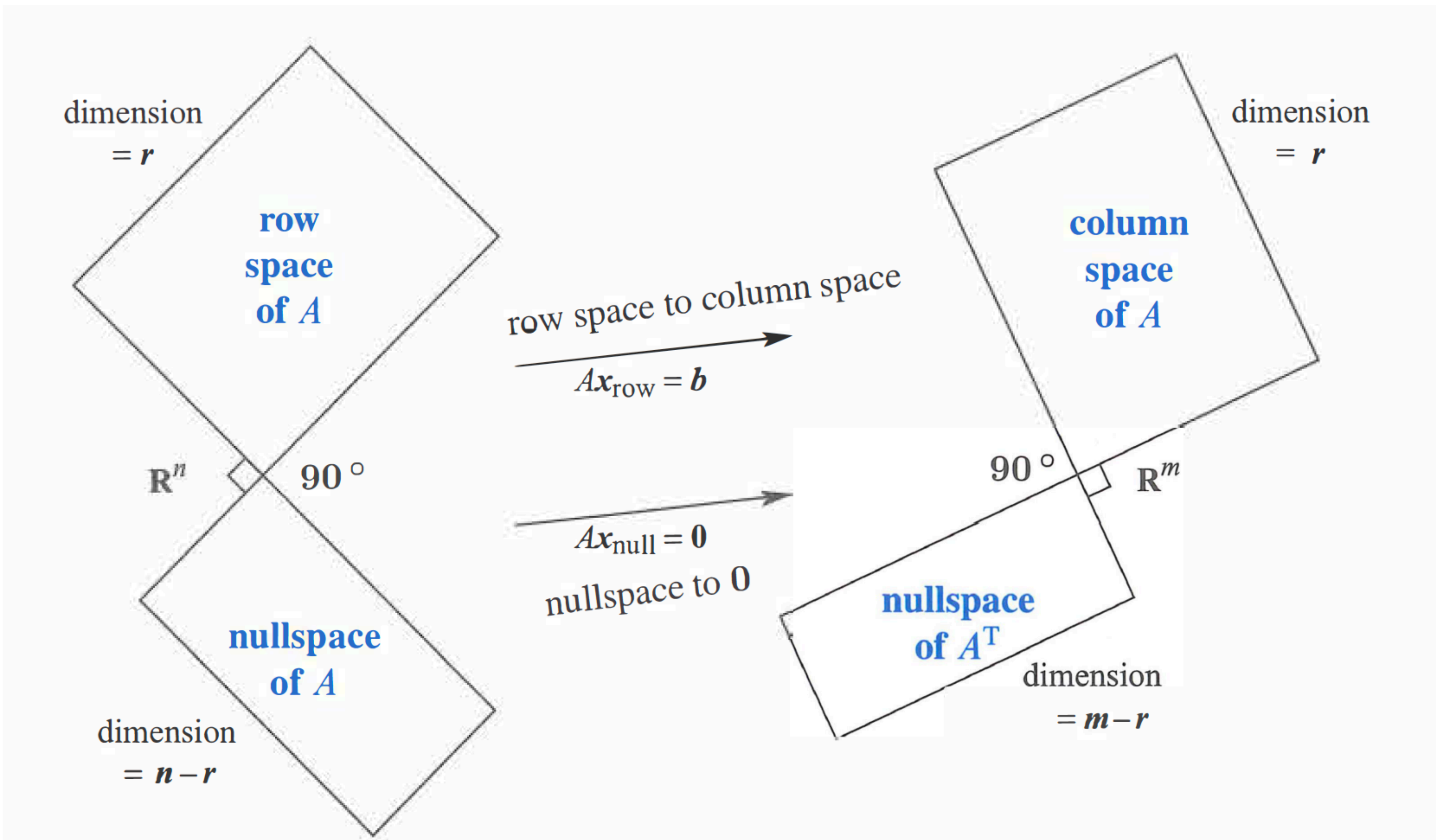
$$v^T w = 0, \qquad ||v||^2 + ||w||^2 = ||v + w||^2$$

Orthogonal Subspaces

Two subspaces V and W of a vector space are orthogonal if every vector $v \in V$ is perpendicular to every vector $w \in W$.

$$v^T w = 0 \; for \; all \; v \in V \; and \; w \in W$$

Orthogonality of the Four Spaces



Orthogonality of the Four Spaces

The null space $N(A)$ and the row space $C(A^T)$ are orthogonal subspaces of R^n.

Proof:

$$A\mathbf{x} = \begin{bmatrix} row\ 1 \\ \vdots \\ row\ m \end{bmatrix} [\mathbf{x}] = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \rightarrow \begin{matrix} (row\ 1) \cdot \mathbf{x} = 0 \\ \vdots \\ (row\ m) \cdot \mathbf{x} = 0 \end{matrix}$$

The left null space $N(A^T)$ and the column space $C(A)$ are orthogonal in R^m.

Projections

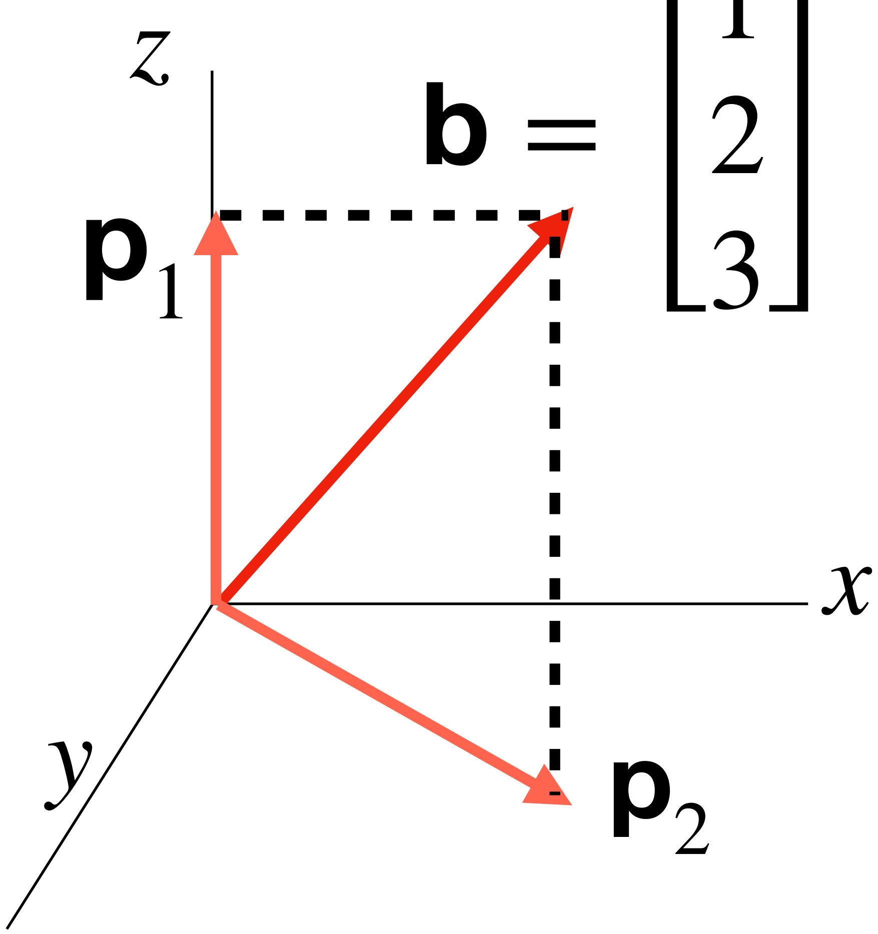
Projections

The projection of b onto a subspace S is the closest vector p in S. For example, when a vector b is projected onto a line, its projection p is the part of b along that line. When b is projected into a plane, p is the part in that plane.

A projection matrix P is a symmetric matrix with $P^2 = P$. The projection of b is Pb.

Projections

What is the projection of vector b onto the *z-axis* line and *xy* plane?



$$\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \quad \mathbf{P}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mathbf{1} \end{bmatrix}$$

$$\mathbf{p}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \mathbf{P}_2 = \begin{bmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

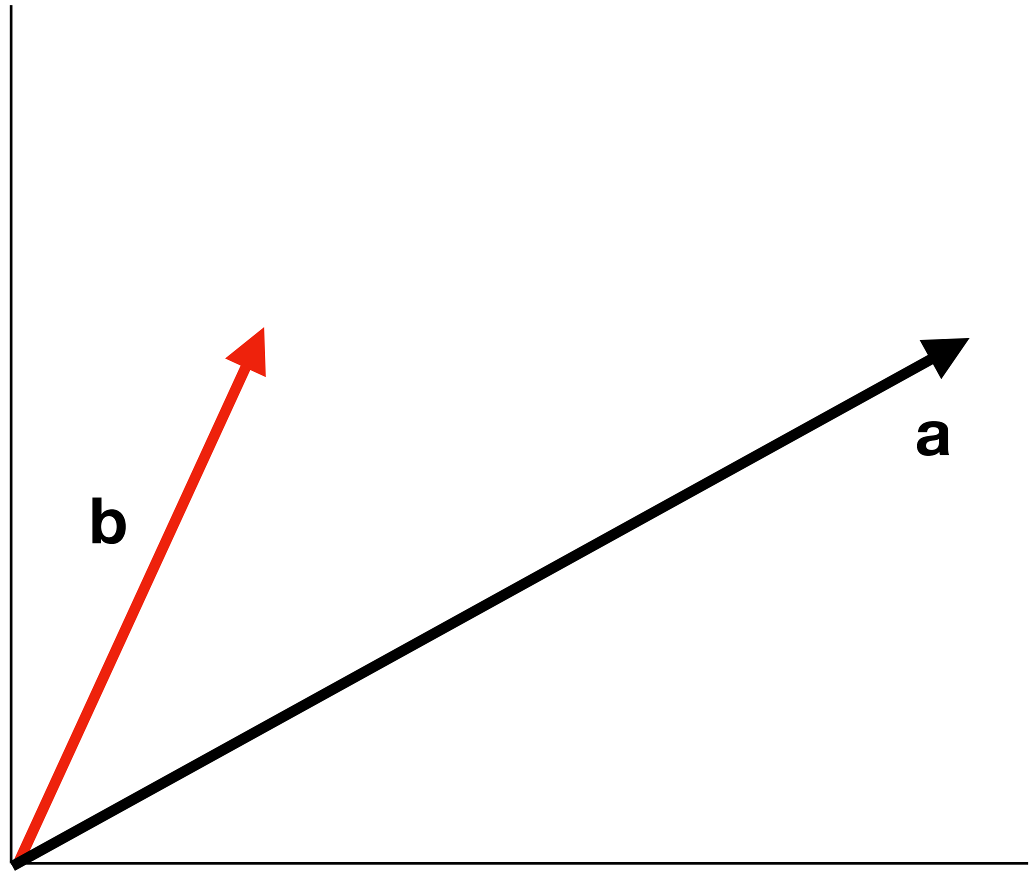
$$\mathbf{p}_1 + \mathbf{p}_2 = b,$$
$$\mathbf{P}_1 + \mathbf{P}_2 = I$$
the line and plane are orthogonal complements.

Why do we need projections?

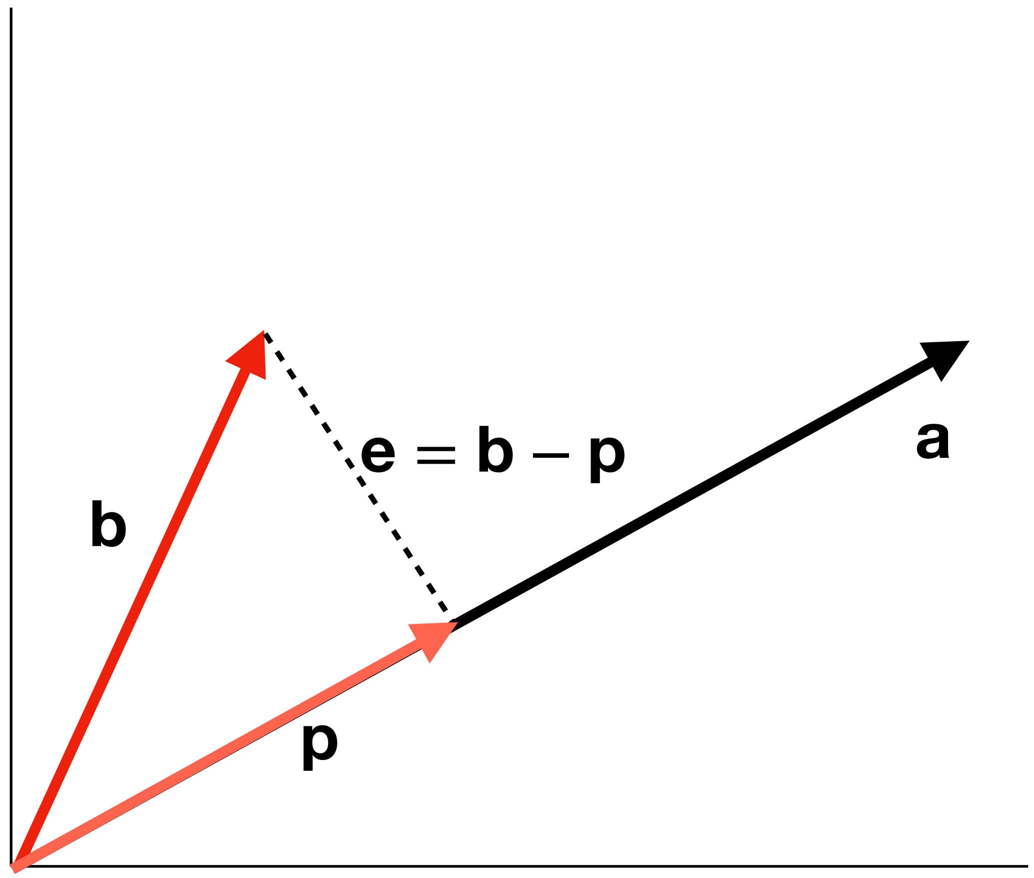
We need projections to cover cases when $A\mathbf{x} = \mathbf{b}$ does not have any solutions. In these cases we can solve the closest problem $A\hat{\mathbf{x}} = \mathbf{p}$.

If $A\mathbf{x} = \mathbf{b}$ does not have any solutions, $\mathbf{b}$ is not in the column space of A. We can solve $A\hat{\mathbf{x}} = \mathbf{p}$ instead where $\mathbf{p}$ is the projection of $\mathbf{b}$ onto the column space of A.

Projection onto a Line



Projection onto a Line



$$\mathbf{p} = \hat{x}\mathbf{a}, \ \ \mathbf{a} \perp (\mathbf{b} - \mathbf{p})$$

$$\mathbf{a} \cdot (\mathbf{b} - \hat{x}\mathbf{a}) = 0$$

$$\mathbf{a} \cdot \mathbf{b} - \hat{x}\mathbf{a} \cdot \mathbf{a} = 0$$

$$\mathbf{a}^T\mathbf{b} - \hat{x}\mathbf{a}^T\mathbf{a} = 0$$

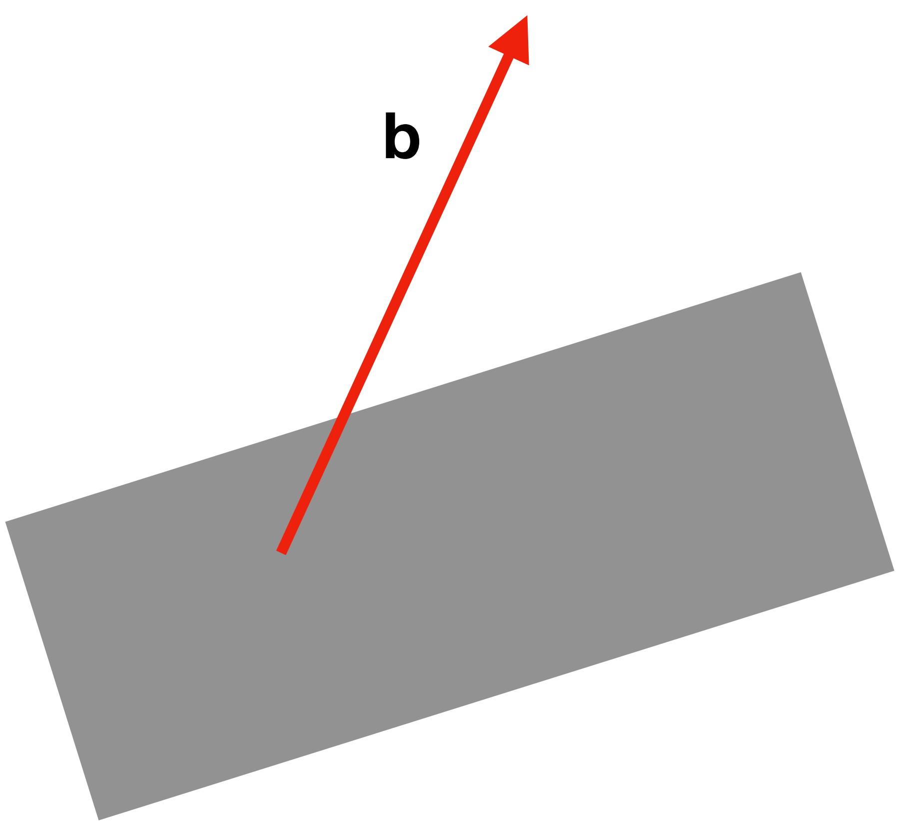
$$\hat{x} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$$

$$\mathbf{p} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}\mathbf{a}$$

$$\mathbf{P} = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$$

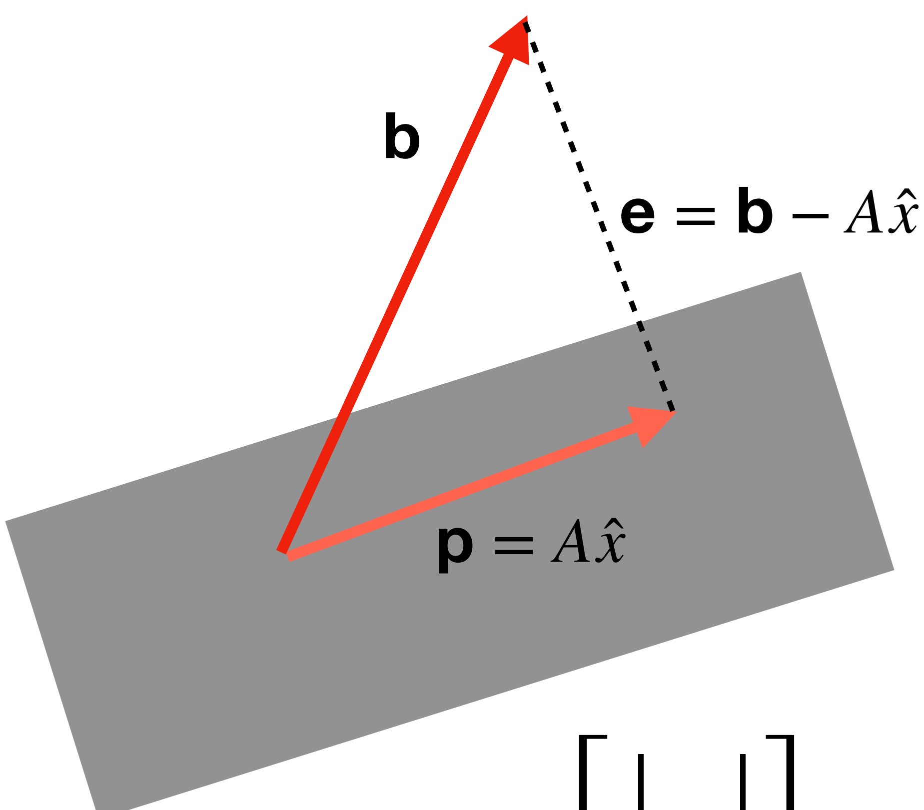
Projection onto a Subspace

Find the projection of $\mathbf{b}$ in $\mathbb{R}^m$ onto the subspace spanned by columns of A.



Projection onto a Subspace

Find the projection of **b** in $\mathbb{R}^m$ onto the subspace spanned by columns of A.



$$\mathbf{A} = \begin{bmatrix} | & | \\ a_1 & a_2 \\ | & | \end{bmatrix}$$

$a_1, and\ a_2\ are\ basis\ for\ plane$

$$\mathbf{p} = A\hat{x}$$

$$\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} (\mathbf{b} - A\hat{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

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

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

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

$$\mathbf{p} = A(A^T A)^{-1} A^T \mathbf{b}$$

$$\mathbf{P} = A(A^T A)^{-1} A^T$$

$A^T A$ Invertibility

Theorem

If A has linearly independent columns then $A^T A$ is invertible.

Proof:

Show that the the null space of $A^T A$ is only the zero vector.

$$A^T A x = 0$$

$$x^T A^T A x = 0$$

$$(Ax)^T A x = 0$$

$$Ax = 0$$

$$x = 0$$

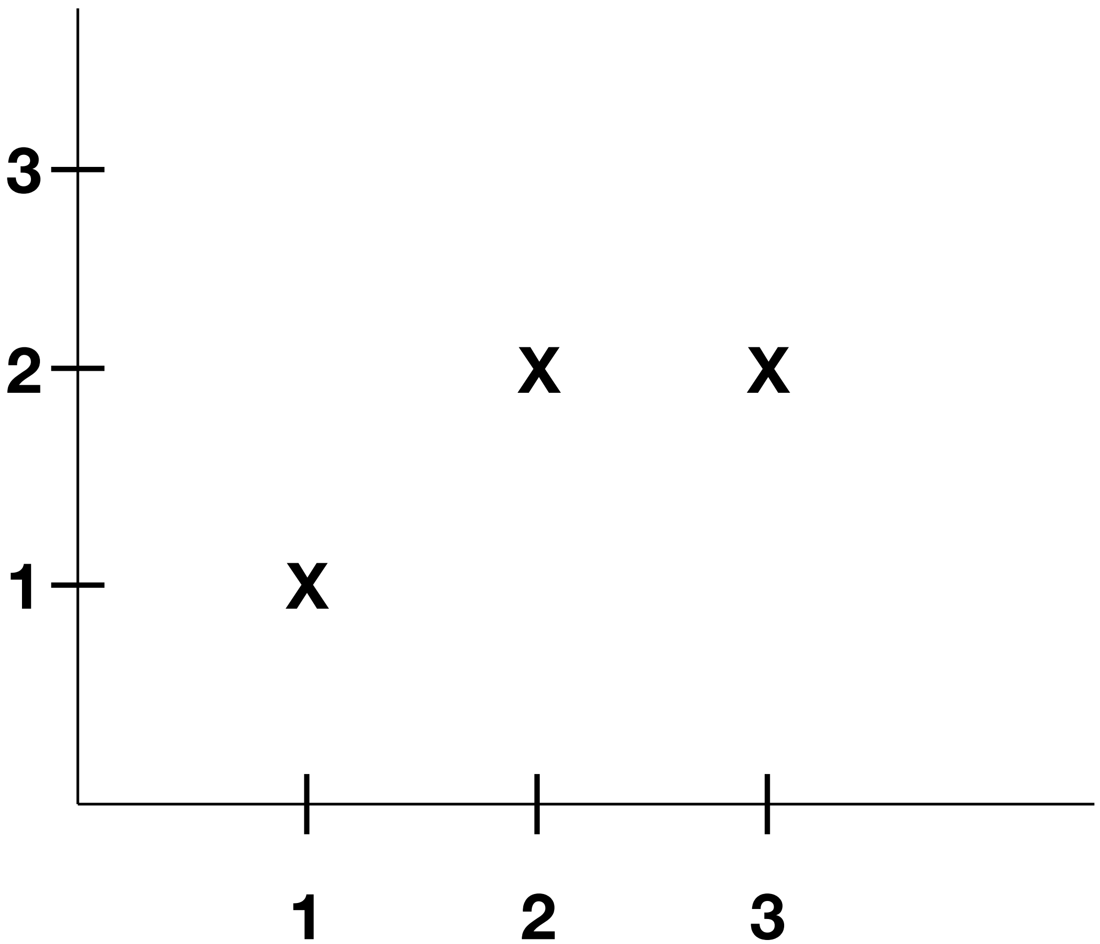
Since columns of A are independent

$\blacksquare$

Least Squares Approximation

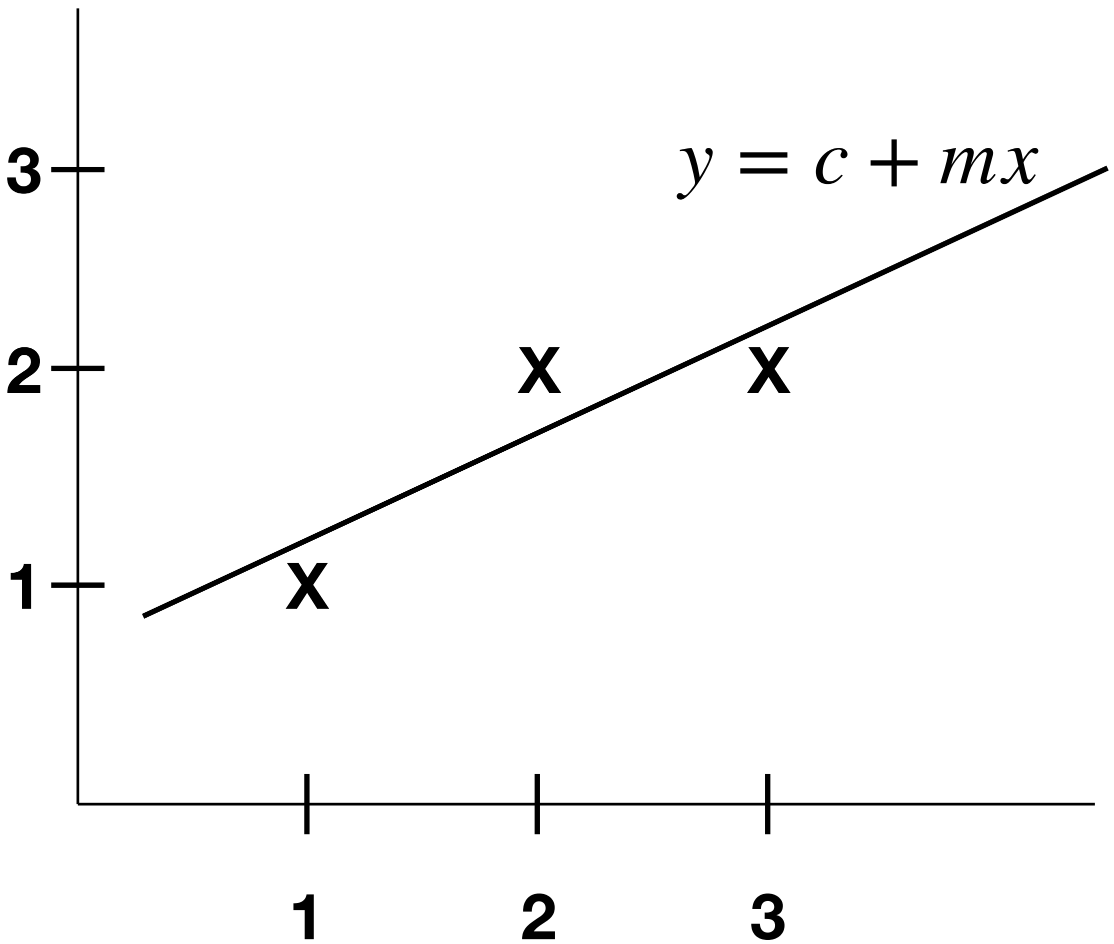
Least Squares Approximation

Find the closest line to the points (1,1), (2,2) and (3,2).



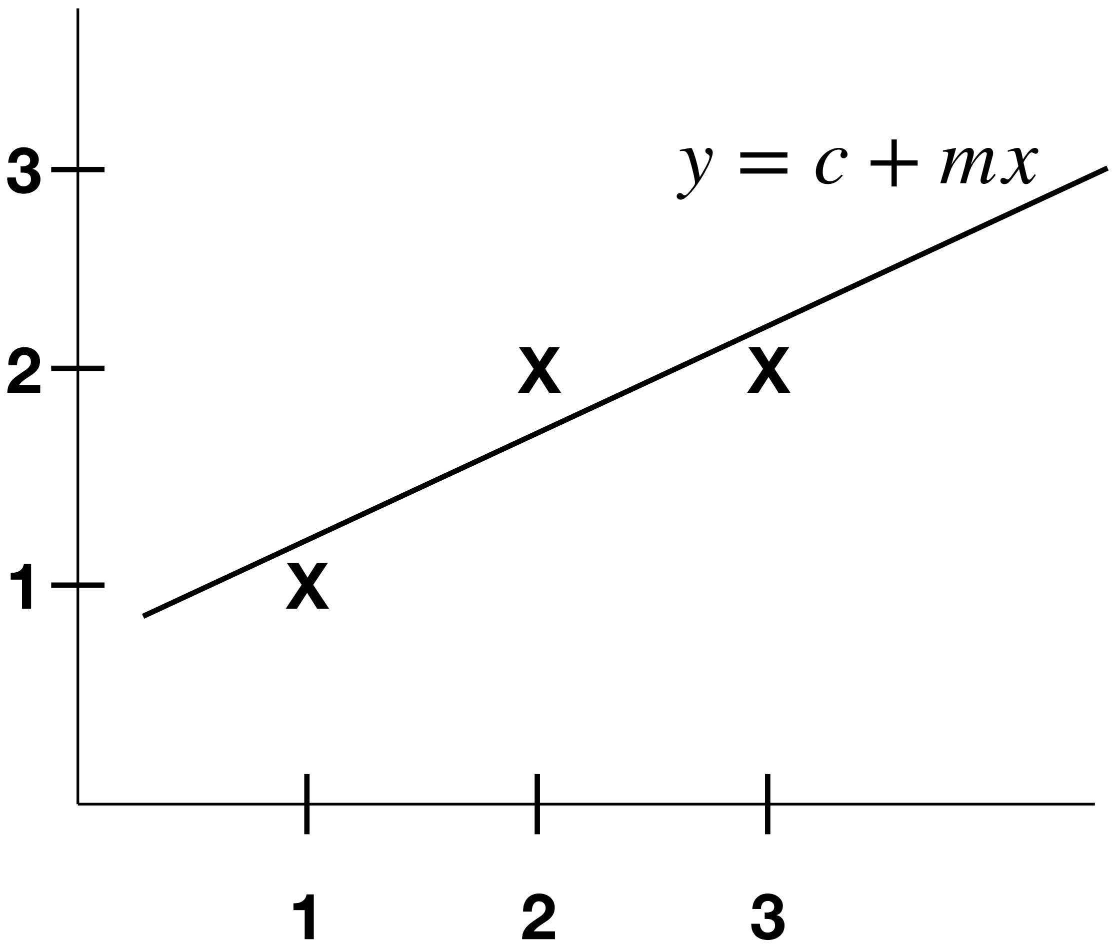
Least Squares Approximation

Find the closest line to the points (1,1), (2,2) and (3,2).



Least Squares Approximation

Find the closest line to the points (1,1), (2,2) and (3,2).



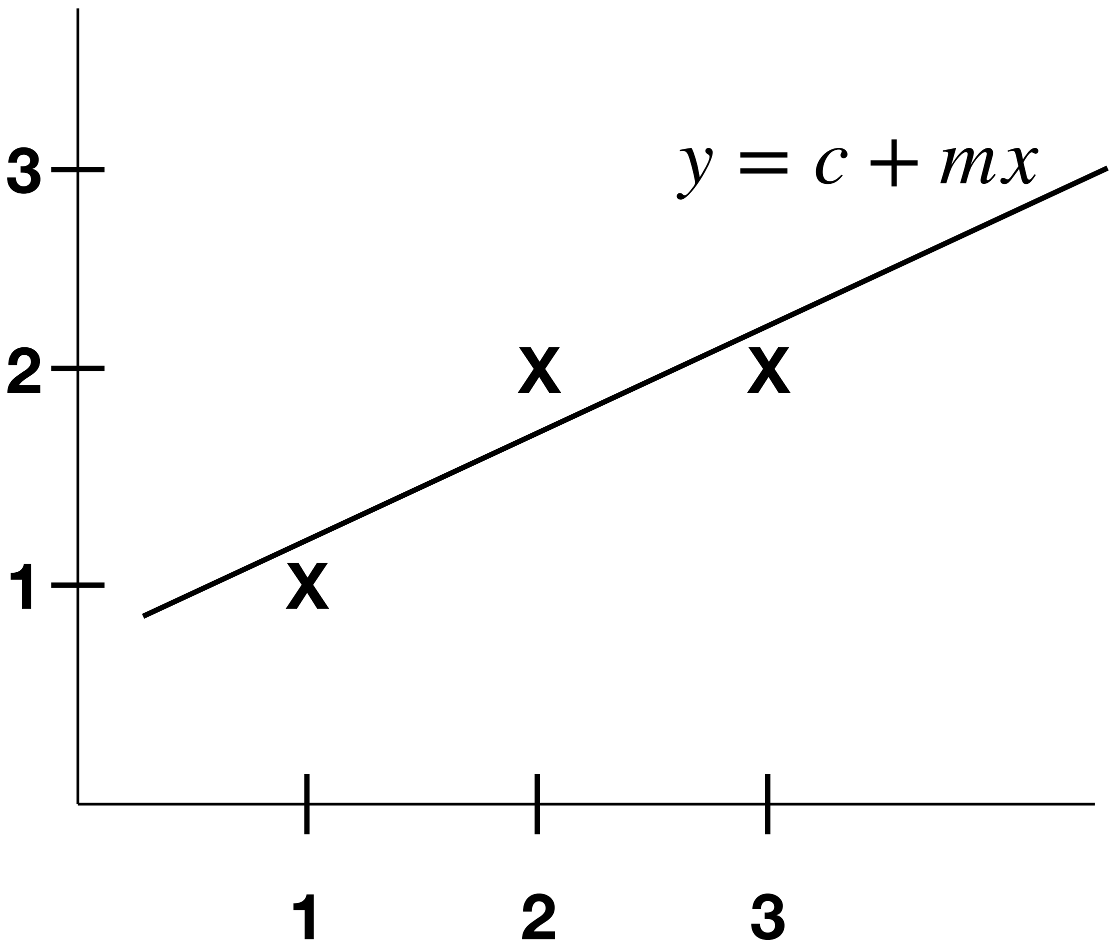
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

3 equation,
2 unknown
No solution

$$Ax = b$$

Least Squares Approximation

Find the closest line to the points (1,1), (2,2) and (3,2).



Minimize

$$e^2 = ||Ax - \mathbf{b}||^2$$

Find $\hat{x} = \begin{bmatrix} \hat{c} \\ \hat{m} \end{bmatrix}$

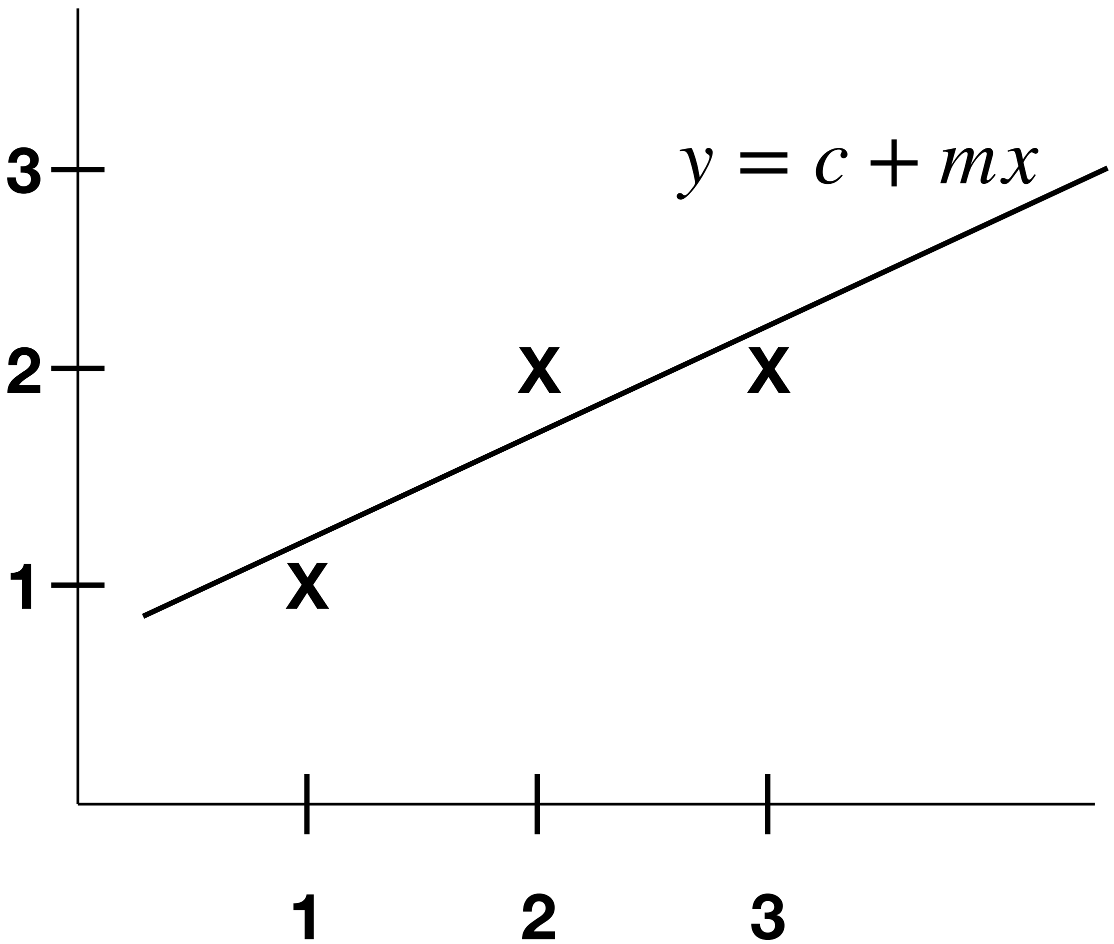
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

3 equation,
2 unknown
No solution

$$Ax = b$$

Least Squares Approximation

Find the closest line to the points (1,1), (2,2) and (3,2).



Minimize

$$e^2 = ||Ax - \mathbf{b}||^2$$

Find $\hat{x} = \begin{bmatrix} \hat{c} \\ \hat{m} \end{bmatrix}$

$$A\hat{x} = p$$

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

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 2 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 6 & 5 \\ 6 & 14 & 11 \end{array}\right]$$

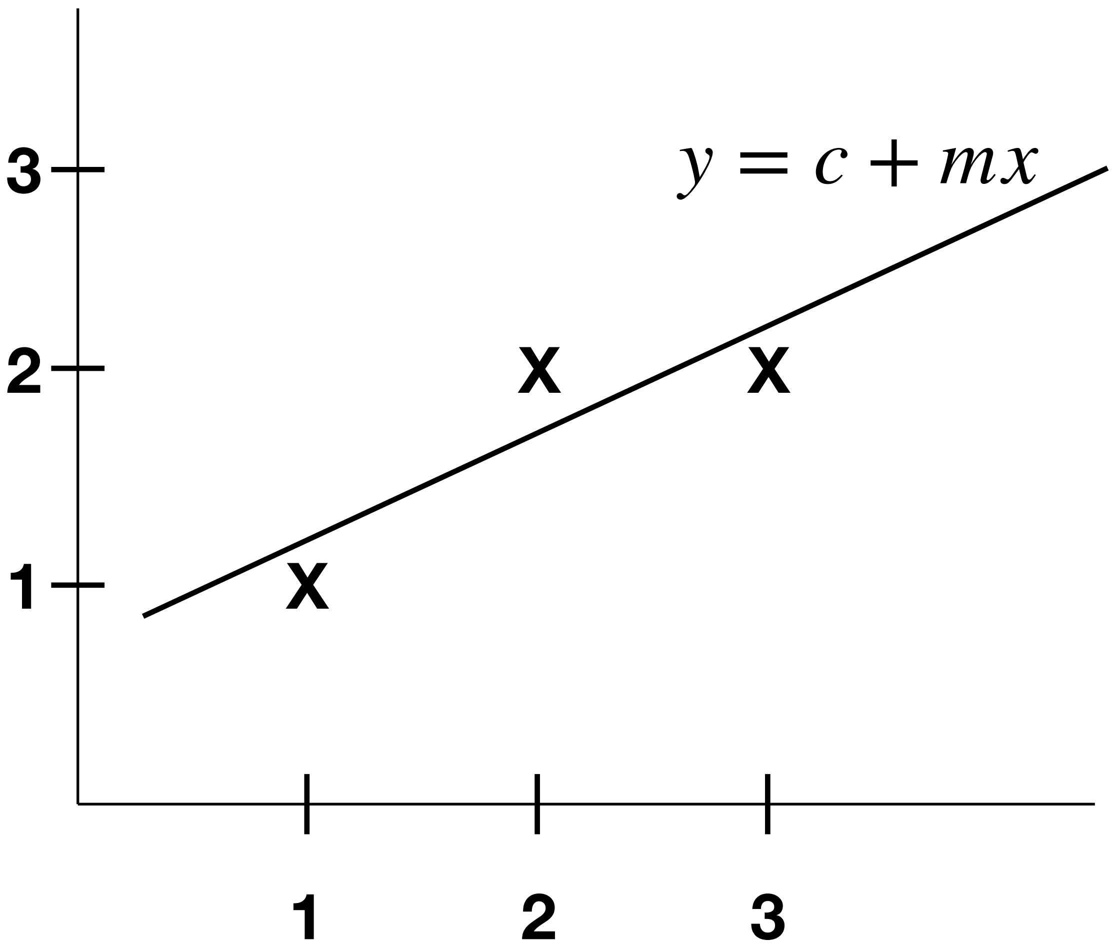
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

3 equation,
2 unknown
No solution

$$Ax = b$$

Least Squares Approximation

Find the closest line to the points (1,1), (2,2) and (3,2).



$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c \\ m \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

3 equation,
2 unknown
No solution

Minimize

$$e^2 = ||Ax - \mathbf{b}||^2$$

Find $\hat{x} = \begin{bmatrix} \hat{c} \\ \hat{m} \end{bmatrix}$

$$A\hat{x} = p$$

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

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 2 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 6 & 5 \\ 6 & 14 & 11 \end{array}\right]$$

$$\begin{cases} 3\hat{c} + 6\hat{m} = 5 \\ 6\hat{c} + 14\hat{m} = 11 \end{cases} \rightarrow \quad \hat{c} = \frac{2}{3} \quad \hat{m} = \frac{1}{2}$$

Orthonormal Vectors

Orthonormal Vectors

Vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ are **orthonormal** if:

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 0 & \textbf{when } i \neq j \ (\textbf{orthogonal vectors}\ ) \\ 1 & \textbf{when } i = j \ (\textbf{unit vectors: } ||\mathbf{q}_i|| = 1) \end{cases}$$

A matrix Q with orthonormal columns satisfies $Q^T Q = I$:

$$Q^T Q = \begin{bmatrix} - & \mathbf{q}_1^T & - \\ - & \mathbf{q}_2^T & - \\ - & \mathbf{q}_n^T & - \end{bmatrix} \begin{bmatrix} | & | & | \\ q_1 & q_2 & q_n \\ | & | & | \end{bmatrix} = I$$

Examples:

$$Q_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Permutation

$$Q_2 = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

Rotation

Projections Using Orthonormal Bases

$$\hat{x} = (A^T A)^{-1} A^T \mathbf{b}, \quad \mathbf{P} = A(A^T A)^{-1} A^T$$

Assume that A has orthonormal columns.

$$\hat{x} = (Q^T Q)^{-1} Q^T \mathbf{b}$$

$$\hat{x} = Q^T \mathbf{b}$$

$$\mathbf{P} = Q(Q^T Q)^{-1} Q^T$$

$$\mathbf{P} = Q I^{-1} Q^T$$

$$\mathbf{P} = Q Q^T$$

P satisfies both $P^T = P$ and $P^2 = P$.

Gram-Schmidt Process

Gram-Schmidt Process

For independent vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$, Gram-Schmidt process constructs orthonormal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

- Start with three independent vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$
- Construct three orthogonal vectors $\mathbf{A}, \mathbf{B}, \mathbf{C}$ as follows:
  - Choose $\mathbf{A} = \mathbf{a}$
  - To construct $\mathbf{B}$, start with $\mathbf{y}$ and subtract its projection along $\mathbf{A}$.
    $$\mathbf{B} = \mathbf{y} - \frac{\mathbf{A}^T\mathbf{y}}{\mathbf{A}^T\mathbf{A}}\mathbf{A}$$
  - To get $\mathbf{C}$, subtract its component in directions $\mathbf{A}$ and $\mathbf{B}$.
    $$\mathbf{C} = \mathbf{z} - \frac{\mathbf{A}^T\mathbf{z}}{\mathbf{A}^T\mathbf{A}}\mathbf{A} - \frac{\mathbf{B}^T\mathbf{z}}{\mathbf{B}^T\mathbf{B}}\mathbf{B}$$
- Produce three orthonormal vectors: $\mathbf{q}_1 = \mathbf{A}/||\mathbf{A}||$, $\mathbf{q}_2 = \mathbf{B}/||\mathbf{B}||$, $\mathbf{q}_3 = \mathbf{C}/||\mathbf{C}||$

Gram-Schmidt Example

$$x = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}, y = \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix}, z = \begin{bmatrix} 6 \\ -6 \\ 6 \end{bmatrix}$$

$$A = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$$

$$B = y - \frac{A^T y}{A^T A} A = \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix} - \frac{\begin{bmatrix} 2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix}}{\begin{bmatrix} 2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}} \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ -4 \end{bmatrix} - \frac{8}{8} \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix}$$

$$C = z - \frac{A^T z}{A^T A} A - \frac{B^T z}{B^T B} B = \begin{bmatrix} 6 \\ -6 \\ 6 \end{bmatrix} - \frac{24}{8} \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} + \frac{24}{24} \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$$q_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, q_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, q_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Thank you!