Lecture 4, Outline

- **Orthogonality**
- **Projections onto subspaces**
- **Least squares method**
- **Orthogonal matrices and Gram-Schmidt orthogonalization**

SUGGESTED READING:
G. Strang, Linear algebra and its applications, Chapter 3

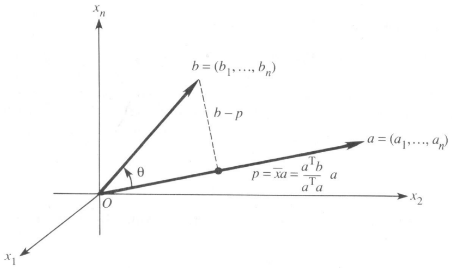
Orthogonality

- Inner product: $x^T x = x_1^2 + \cdots + x_n^2 = ||x||^2$ (definition of a norm)
- Orthogonality (in $\mathbf{R}^n$): $x \perp y \iff x^T y = 0$
- Orthogonal subspaces V, W: $v^T w = 0$ for all $v \in V$ and $w \in W$

Orthogonality of the 4 fundamental subspaces

The *row space is orthogonal to the nullspace* (in $\mathbf{R}^n$) and the *column space is orthogonal to the left nullspace* (in $\mathbf{R}^m$).

Distance between point b and line a?



$$a \perp (b - \bar{x}a) \Rightarrow a^T(b - \bar{x}a) = 0$$
$$\Rightarrow \bar{x} = \frac{a^T b}{a^T a}$$

Projections

- Projection onto a line carried out by a *projection matrix* P, $p = a\bar{x} = \underbrace{\frac{aa^T}{a^T a}}_{P} b$, where a is $(n \times 1)$.
- Instead of a line a, we are given a plane, or a subspace S of $\mathbf{R}^n$, and we look for the point p on that subspace closest to b.
- If $|a| = 1$, then $a^T a = 1 \Rightarrow P = aa^T$.

<u>Example</u>

- Projection onto a line carried out by a *projection matrix* P, $p = a\bar{x} = \underbrace{\frac{aa^T}{a^T a}}_{P} b$, where a is $(n \times 1)$.
- Instead of a line a, we are given a plane, or a subspace S of $\mathbf{R}^n$, and we look for the point p on that subspace closest to b.
- If $|a| = 1$, then $a^T a = 1 \Rightarrow P = aa^T$.

<u>EXAMPLE</u>

Line going through $a = (1, 1, 1)$. $P = ?$

$$P = \frac{aa^T}{a^T a} = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}.$$

Notice that P *is symmetric, rank is 1 and* $P^2 = P$.

Least squares method with one variable

- Despite their unsolvability, inconsistent equations arise in practice and have to be solved (imagine e.g. m observations, prone to error, and n unknowns, with $m > n$).
- *Strategy*: Choose x that minimizes the average (sum of squares) error in the equations.

EXAMPLE

Least squares method with one variable

- Despite their unsolvability, inconsistent equations arise in practice and have to be solved (imagine e.g. m observations, prone to error, and n unknowns, with $m > n$).
- *Strategy*: Choose x that minimizes the average (sum of squares) error in the equations.

EXAMPLE

$2x = b_1, 3x = b_2, 4x = b_3.$

$E^2 = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2 \Rightarrow$

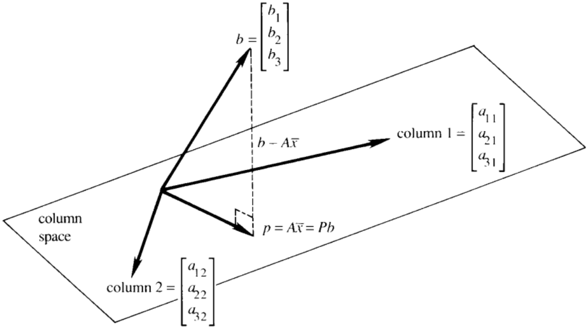
$\frac{dE^2}{dx} = 2[(2x - b_1)2 + (3x - b_2)3 + (4x - b_3)4] \overset{!}{=} 0$

$\bar{x} = \frac{2b_1 + 3b_2 + 4b_3}{2^2 + 3^2 + 4^2}.$

Least squares = projection of b onto a subspace

The least squares solution to $ax = b$ is $\bar{x} = \frac{a^T b}{a^T a}$.

Least squares method with several variables



Use

1. Geometry: Since p is the projection of b onto the column space of A, the error vector $b - A\bar{x}$ must be perpendicular to that space (i.e. must lie in the left nullspace): $A^T(b - A\bar{x}) = 0$, or $\boxed{A^T A\bar{x} = A^T b}$; or
2. Algebra: Setting partial derivatives of $E^2 = E^T E = (Ax - b)^T(Ax - b)$ to zero gives $2A^T Ax - 2A^T b = 0$.

Least squares method, summary and examples

Least squares, summary

The least squares solution to an inconsistent system $Ax = b$ satisfies $A^T A\bar{x} = A^T b$. If columns of A are lin. indep., than $A^T A$ is invertible and $\bar{x} = (A^T A)^{-1} A^T b$ (compare with $\bar{x} = \frac{1}{a^T a} a^T b$).

Note 1: We are effectively solving $A\bar{x} = p$ instead of $Ax = b$.
Note 2: When A is square and invertible, we can write $(A^T A)^{-1}$ as $A^{-1}(A^T)^{-1}$: $p = A\bar{x} = \underbrace{A(A^T A)^{-1} A^T}_{P} b = AA^{-1}(A^T)^{-1} A^T b = b$.

<u>EXAMPLE</u>

Least squares, summary

The least squares solution to an inconsistent system $Ax = b$ satisfies $A^T A\bar{x} = A^T b$. If columns of A are lin. indep., than $A^T A$ is invertible and $\bar{x} = (A^T A)^{-1} A^T b$ (compare with $\bar{x} = \frac{1}{a^T a} a^T b$).

Note 1: We are effectively solving $A\bar{x} = p$ instead of $Ax = b$.

Note 2: When A is square and invertible, we can write $(A^T A)^{-1}$ as $A^{-1}(A^T)^{-1}$: $p = A\bar{x} = \underbrace{A(A^T A)^{-1} A^T}_{P} b = AA^{-1}(A^T)^{-1} A^T b = b$.

EXAMPLE

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix},$$

$$b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

EXAMPLE

Least squares method, summary and examples

Least squares, summary

The least squares solution to an inconsistent system $Ax = b$ satisfies $A^T A\bar{x} = A^T b$. If columns of A are lin. indep., than $A^T A$ is invertible and $\bar{x} = (A^T A)^{-1} A^T b$ (compare with $\bar{x} = \frac{1}{a^T a} a^T b$).

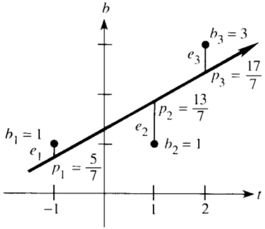
Note 1: We are effectively solving $A\bar{x} = p$ instead of $Ax = b$.

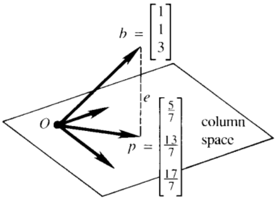
Note 2: When A is square and invertible, we can write $(A^T A)^{-1}$ as $A^{-1}(A^T)^{-1}$: $p = A\bar{x} = \underbrace{A(A^T A)^{-1} A^T}_{P} b = AA^{-1}(A^T)^{-1} A^T b = b$.

<u>EXAMPLE</u>

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

<u>EXAMPLE</u>





Orthogonal matrices

Vectors $q_1, \ldots, q_k$ *orthonormal* if

$$q_i^T q_j = \begin{cases} 0 & \text{if } i \neq j \quad \text{(orthogonality)} \\ 1 & \text{if } i = j \quad \text{(normalization)} \end{cases}$$

- Standard basis $\Rightarrow Q = I$
- *Orthogonal matrix* Q: a square matrix with orthonormal columns, $Q^T Q = I \Rightarrow \boxed{Q^T = Q^{-1}}$

EXAMPLE

Vectors $q_1, \dots, q_k$ *orthonormal* if

$$q_i^T q_j = \begin{cases} 0 & \text{if } i \neq j \quad \text{(orthogonality)} \\ 1 & \text{if } i = j \quad \text{(normalization)} \end{cases}$$

- Standard basis $\Rightarrow Q = I$
- *Orthogonal matrix* Q: a square matrix with orthonormal columns, $Q^T Q = I \Rightarrow \boxed{Q^T = Q^{-1}}$

EXAMPLE

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad Q^T = Q^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \textit{(rotation)}$$

Why are Qs important?

Orthogonal matrices

Vectors $q_1, \ldots, q_k$ *orthonormal* if

$$q_i^T q_j = \begin{cases} 0 & \text{if } i \neq j \quad \text{(orthogonality)} \\ 1 & \text{if } i = j \quad \text{(normalization)} \end{cases}$$

- Standard basis $\Rightarrow Q = I$
- *Orthogonal matrix Q*: a square matrix with orthonormal columns, $Q^T Q = I \Rightarrow \boxed{Q^T = Q^{-1}}$

EXAMPLE

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad Q^T = Q^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \textit{(rotation)}$$

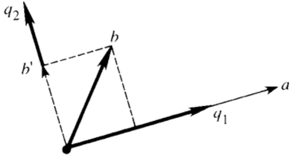
Why are Qs important?

- Any $b = x_1 q_1 + x_2 q_2 + \cdots + x_n q_n$, where $x = Q^T b$ (Fourier expansion).
- Least squares become easy:

$$Q^T Q\bar{x} = Q^T b \Rightarrow \underline{\bar{x} = Q^T b}, \; p = Q\bar{x} = \underbrace{QQ^T}_{P} b$$

Gram-Schmidt process and QR factorization

How to turn *independent*, non-orthogonal vectors into orthogonal ones?
Repeated projections!



$\mathbf{a}, \mathbf{b}, \mathbf{c} \to \mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}$:

$$a' = a, \quad q_1 = a'/\|a'\|$$

$$b' = b - (q_1^T b) q_1, \quad q_2 = b'/\|b'\|$$

$$c' = c - (q_1^T c) q_1 - (q_2^T c) q_2, \quad q_3 = c'/\|c'\|$$

$$\begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} q_1^T a & q_1^T b & q_1^T c \\ & q_2^T b & q_2^T c \\ & & q_3^T c \end{bmatrix} \Rightarrow \boxed{\text{A=QR}}$$

EXAMPLE

Gram-Schmidt process and QR factorization

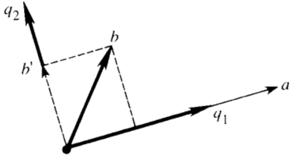
How to turn *independent*, non-orthogonal vectors into orthogonal ones?
Repeated projections!

$\mathbf{a},\mathbf{b},\mathbf{c} \rightarrow \mathbf{q_1},\mathbf{q_2},\mathbf{q_3}$:

$a' = a, \quad q_1 = a'/\|a'\|$

$b' = b - (q_1^T b)q_1, \quad q_2 = b'/\|b'\|$

$c' = c - (q_1^T c)q_1 - (q_2^T c)q_2, \quad q_3 = c'/\|c'\|$



$$\begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} q_1^T a & q_1^T b & q_1^T c \\ & q_2^T b & q_2^T c \\ & & q_3^T c \end{bmatrix} \Rightarrow \boxed{\text{A=QR}}$$

<u>EXAMPLE</u>

$$a = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}. \qquad q_1, q_2, q_3 = ?$$