Math 405 : Learning From Data

Week 2 : Dot Product, Orthogonality, and Projection

Dr. Maher Boudabra

Lecture 1: Dot Product and Its Properties

Lecture 1: Outline

Lecture 1: Dot Product and Its Properties

Lecture 2: Orthogonality and Orthonormal Bases

Lecture 3: Projections and Least Squares

Motivation: Why Inner Products?

- Measure angles and lengths in vector spaces
- Define orthogonality (perpendicularity)
- Fundamental for:
  - Geometry and distance measurements
  - Least squares approximations
  - Fourier analysis
  - Signal processing
  - Quantum mechanics
- Generalization of the familiar dot product from $\mathbb{R}^2$ and $\mathbb{R}^3$

Definition of Dot Product

For vectors $\mathbf{x} = (x_1, \ldots, x_n)^T$, $\mathbf{y} = (y_1, \ldots, y_n)^T$ in $\mathbb{R}^n$:

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i$$

Example:

$$\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = (1)(3) + (2)(0) + (-1)(4) = 3 + 0 - 4 = -1$$

Algebraic Properties of Dot Product

For vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ and scalar $c \in \mathbb{R}$:

1. **Commutativity:** $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. **Linearity:** $\mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$
3. **Distributivity:** $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
4. **Positive Definiteness:** $\mathbf{v} \cdot \mathbf{v} \geq 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ iff $\mathbf{v} = \mathbf{0}$

These properties make the dot product an example of an **inner product**.

Geometric Interpretation: Length

The **length** or **norm** of a vector $\mathbf{v} \in \mathbb{R}^n$ is defined as:

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

Example:

$$\left\| \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\| = \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5$$

A vector with $\|\mathbf{v}\| = 1$ is called a **unit vector**.

Geometric Interpretation: Angle

For nonzero vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the angle θ between them satisfies:

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}, \quad 0 \leq \theta \leq \pi$$

Example: Angle between $\mathbf{u} = (1,1)^T$ and $\mathbf{v} = (1,0)^T$:

$$\cos\theta = \frac{(1)(1) + (1)(0)}{\sqrt{1^2 + 1^2}\sqrt{1^2 + 0^2}} = \frac{1}{\sqrt{2}} \Rightarrow \theta = 45°$$

Special Case: Orthogonality

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** (perpendicular) if:

$$\mathbf{u} \cdot \mathbf{v} = 0$$

Examples:

- $(1, 0)^T$ and $(0, 1)^T$ are orthogonal
- $(1, 2, -1)^T$ and $(2, -1, 0)^T$ are orthogonal since: $1 \cdot 2 + 2 \cdot (-1) + (-1) \cdot 0 = 2 - 2 + 0 = 0$
- The zero vector is orthogonal to every vector

Cauchy-Schwarz Inequality

For any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$:

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

Equality holds if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

Proof sketch: Consider the quadratic function:

$$f(t) = \|\mathbf{u} + t\mathbf{v}\|^2 = \|\mathbf{v}\|^2 t^2 + 2(\mathbf{u} \cdot \mathbf{v})t + \|\mathbf{u}\|^2 \geq 0$$

Discriminant must be ≤ 0: $4(\mathbf{u} \cdot \mathbf{v})^2 - 4\|\mathbf{u}\|^2\|\mathbf{v}\|^2 \leq 0$

Triangle Inequality

For any vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$:

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Proof:

$$\begin{aligned}
\|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\
&= \|\mathbf{u}\|^2 + 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2|\mathbf{u} \cdot \mathbf{v}| + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \quad \text{(by Cauchy-Schwarz)} \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

General Inner Products

An **inner product** on a real vector space V is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying:

1. **Symmetry:** $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. **Bilinearity:** $\langle a\mathbf{u} + b\mathbf{w}, \mathbf{v} \rangle = a\langle \mathbf{u}, \mathbf{v} \rangle + b\langle \mathbf{w}, \mathbf{v} \rangle$
3. **Positive Definiteness:** $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ for $\mathbf{v} \neq \mathbf{0}$

The dot product is the **standard inner product** on $\mathbb{R}^n$, but many others exist.

Examples of Different Inner Products

1. **Weighted dot product** on $\mathbb{R}^n$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1 x_1 y_1 + w_2 x_2 y_2 + \cdots + w_n x_n y_n$$

   where $w_i > 0$ are fixed weights.

2. **Integral inner product** on functions:

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

   on the space of continuous functions on $[a, b]$.

Application: Distance Between Vectors

The **distance** between vectors $\mathbf{u}$ and $\mathbf{v}$ is defined as:

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \sqrt{(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})}$$

Properties:

- $d(\mathbf{u},\mathbf{v}) \geq 0$, and $= 0$ iff $\mathbf{u}=\mathbf{v}$
- $d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$ (symmetry)
- $d(\mathbf{u},\mathbf{w}) \leq d(\mathbf{u},\mathbf{v}) + d(\mathbf{v},\mathbf{w})$ (triangle inequality)

Lecture 1 Summary

- Dot product: $\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i$
- Length: $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$
- Angle: $\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$
- Orthogonality: $\mathbf{u} \cdot \mathbf{v} = 0$
- Key inequalities: Cauchy-Schwarz and Triangle
- General inner products satisfy symmetry, bilinearity, positive definiteness

Lecture 2: Orthogonality and Orthonormal Bases

Orthogonal Sets

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is called **orthogonal** if:

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{for all } i \neq j$$

Examples:

- Standard basis in $\mathbb{R}^n$: $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$
- $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \right\}$ in $\mathbb{R}^3$

Orthonormal Sets

An orthogonal set where each vector has unit length is called **orthonormal**:

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Examples:

- Standard basis in $\mathbb{R}^n$
- $\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ in $\mathbb{R}^2$

Why Orthogonal Sets Are Useful

Theorem

An orthogonal set of nonzero vectors is linearly independent.

Proof: Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is orthogonal and

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Take dot product with $\mathbf{v}_i$:

$$c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) = \mathbf{0} \cdot \mathbf{v}_i = 0$$

But $\mathbf{v}_j \cdot \mathbf{v}_i = 0$ for $j \neq i$, so $c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = 0$ Since $\mathbf{v}_i \neq \mathbf{0}$, $\mathbf{v}_i \cdot \mathbf{v}_i > 0$, so $c_i = 0$ for all i.

Orthogonal Bases

An **orthogonal basis** for a subspace $W \subseteq \mathbb{R}^n$ is a basis that is also an orthogonal set.

An **orthonormal basis** is an orthogonal basis where each vector has unit length.

Advantages:

- Easy to compute coordinates
- Numerically stable
- Simplify many calculations

Coordinates Relative to Orthogonal Bases

If $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthogonal basis for W, then any $\mathbf{w} \in W$ can be written as:

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

where the coefficients are given by:

$$c_i = \frac{\mathbf{w} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}$$

If the basis is orthonormal, this simplifies to:

$$c_i = \mathbf{w} \cdot \mathbf{v}_i$$

Example: Coordinates in Orthogonal Basis

Let $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$

This is an orthogonal basis for $\mathbb{R}^3$.

Express $\mathbf{w} = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}$ in this basis:

$$c_1 = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} = \frac{4+2+0}{1+1+0} = \frac{6}{2} = 3$$

$$c_2 = \frac{\mathbf{w} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} = \frac{4-2+1}{1+1+1} = \frac{3}{3} = 1$$

$$c_3 = \frac{\mathbf{w} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} = \frac{4-2-2}{1+1+4} = \frac{0}{6} = 0$$

So $\mathbf{w} = 3\mathbf{v}_1 + \mathbf{v}_2 + 0\mathbf{v}_3$

Gram-Schmidt Orthogonalization Process

Algorithm to convert a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ into an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$:

1. $\mathbf{v}_1 = \mathbf{u}_1$
2. $\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$
3. $\mathbf{v}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$
4. Continue similarly for $\mathbf{v}_4, \ldots, \mathbf{v}_k$

To get an orthonormal basis, normalize each $\mathbf{v}_i$: $\mathbf{q}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$

Gram-Schmidt Example

Convert $\mathbf{u_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{u_2} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{u_3} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ to orthogonal basis.

Step 1: $\mathbf{v_1} = \mathbf{u_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

Step 2: $\mathbf{v_2} = \mathbf{u_2} - \frac{\mathbf{u_2} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}} \mathbf{v_1} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$

Step 3: $\mathbf{v_3} = \mathbf{u_3} - \frac{\mathbf{u_3} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}} \mathbf{v_1} - \frac{\mathbf{u_3} \cdot \mathbf{v_2}}{\mathbf{v_2} \cdot \mathbf{v_2}} \mathbf{v_2}$

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

QR Factorization

Any $m \times n$ matrix A with linearly independent columns can be factored as:

$$A = QR$$

where:

- Q is an $m \times n$ matrix with orthonormal columns
- R is an $n \times n$ upper triangular matrix with positive diagonal entries

This is the matrix formulation of the Gram-Schmidt process.

Orthogonal Matrices

A square matrix Q is **orthogonal** if its columns form an orthonormal basis for $\mathbb{R}^n$.

Properties:

- $Q^TQ = I$ (and $QQ^T = I$ if Q is square)
- $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$ (preserves length)
- $(Q\mathbf{x}) \cdot (Q\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ (preserves dot product)
- $Q^{-1} = Q^T$ (easy inverse)

Orthogonal matrices represent rotations and reflections.

Examples of Orthogonal Matrices

1. Rotation matrix in $\mathbb{R}^2$:

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

2. Reflection matrix:

$$Q = I - 2\mathbf{v}\mathbf{v}^T \quad \text{where } \|\mathbf{v}\| = 1$$

3. Permutation matrices (reorder coordinates)

Orthogonal Complements

For a subspace $W \subseteq \mathbb{R}^n$, its **orthogonal complement** is:

$$W^\perp = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{v} \cdot \mathbf{w} = 0 \text{ for all } \mathbf{w} \in W\}$$

Properties:

- $W^\perp$ is a subspace of $\mathbb{R}^n$
- $(W^\perp)^\perp = W$
- $\dim W + \dim W^\perp = n$
- $\mathbb{R}^n = W \oplus W^\perp$ (direct sum)

Lecture 2 Summary

- Orthogonal/orthonormal sets and bases
- Coordinates in orthogonal bases are easy to compute
- Gram-Schmidt process converts any basis to orthogonal basis
- QR factorization: $A = QR$
- Orthogonal matrices preserve lengths and angles
- Orthogonal complements: $W^\perp$

Lecture 3: Projections and Least Squares

Lecture 3: Outline

Orthogonal Projection onto a Line

The **orthogonal projection** of vector $\mathbf{v}$ onto a nonzero vector $\mathbf{u}$ is:

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$$

Properties:

- $\text{proj}_{\mathbf{u}}(\mathbf{v})$ is parallel to $\mathbf{u}$
- $\mathbf{v} - \text{proj}_{\mathbf{u}}(\mathbf{v})$ is orthogonal to $\mathbf{u}$
- $\text{proj}_{\mathbf{u}}(\mathbf{v})$ is the closest point to $\mathbf{v}$ on the line spanned by $\mathbf{u}$

Geometric Interpretation of Projection

$$\text{proj}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{u}\|^2}\right) \mathbf{u}$$

The scalar $\frac{\mathbf{v} \cdot \mathbf{u}}{\|\mathbf{u}\|^2}$ is called the **component** of $\mathbf{v}$ in the direction of $\mathbf{u}$.

Orthogonal Projection onto a Subspace

For a subspace $W \subseteq \mathbb{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, the orthogonal projection of $\mathbf{v}$ onto W is:

$$\text{proj}_W(\mathbf{v}) = \frac{\mathbf{v} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{v} \cdot \mathbf{u}_k}{\mathbf{u}_k \cdot \mathbf{u}_k}\mathbf{u}_k$$

If the basis is orthonormal $\{\mathbf{q}_1, \dots, \mathbf{q}_k\}$, this simplifies to:

$$\text{proj}_W(\mathbf{v}) = (\mathbf{v} \cdot \mathbf{q}_1)\mathbf{q}_1 + \cdots + (\mathbf{v} \cdot \mathbf{q}_k)\mathbf{q}_k$$

Properties of Orthogonal Projection

For any $\mathbf{v} \in \mathbb{R}^n$ and subspace $W \subseteq \mathbb{R}^n$:

1. $\text{proj}_W(\mathbf{v}) \in W$
2. $\mathbf{v} - \text{proj}_W(\mathbf{v}) \in W^\perp$
3. $\text{proj}_W(\mathbf{v})$ is the unique vector in W closest to $\mathbf{v}$
4. $\|\mathbf{v}\|^2 = \|\text{proj}_W(\mathbf{v})\|^2 + \|\mathbf{v} - \text{proj}_W(\mathbf{v})\|^2$ (Pythagorean theorem)

Projection Matrix

The orthogonal projection onto a subspace W can be represented by a **projection matrix** P such that:

$$\text{proj}_W(\mathbf{v}) = P\mathbf{v}$$

If $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ is an orthonormal basis for W, then:

$$P = \mathbf{q}_1\mathbf{q}_1^T + \mathbf{q}_2\mathbf{q}_2^T + \cdots + \mathbf{q}_k\mathbf{q}_k^T$$

Properties of projection matrices:

- $P^2 = P$ (idempotent)
- $P^T = P$ (symmetric)
- Eigenvalues are 0 or 1

Least Squares Problems

Often we encounter inconsistent linear systems:

$$A\mathbf{x} = \mathbf{b}$$

where $\mathbf{b}$ is not in the column space of A.

The **least squares solution** minimizes:

$$\|A\mathbf{x} - \mathbf{b}\|^2$$

This occurs when $A\mathbf{x}$ is the projection of $\mathbf{b}$ onto the column space of A.

Normal Equations

The least squares solution to $A\mathbf{x} = \mathbf{b}$ satisfies:

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

These are called the **normal equations**.

Why? The error $\mathbf{b} - A\mathbf{x}$ is orthogonal to the column space of A, so:

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

Solving Least Squares Problems

1. Compute $A^T A$ and $A^T \mathbf{b}$
2. Solve the normal equations $A^T A\mathbf{x} = A^T \mathbf{b}$
3. Alternatively, use QR factorization: if $A = QR$, then

$$A^T A\mathbf{x} = A^T \mathbf{b} \Rightarrow R^T Q^T QR\mathbf{x} = R^T Q^T \mathbf{b} \Rightarrow R^T R\mathbf{x} = R^T Q^T \mathbf{b}$$

Since R is invertible, $R\mathbf{x} = Q^T \mathbf{b}$ (easier to solve)

Example: Least Squares Fit

Find the line $y = ax + b$ that best fits the points: (1,1), (2,3), (3,3), (4,4)

Set up system:

$$\begin{aligned} a(1) + b &= 1 \\ a(2) + b &= 3 \\ a(3) + b &= 3 \\ a(4) + b &= 4 \end{aligned} \Rightarrow A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \\ 3 \\ 4 \end{pmatrix}$$

Normal equations:

$$A^T A = \begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix}, \quad A^T \mathbf{b} = \begin{pmatrix} 35 \\ 11 \end{pmatrix}$$

Solve: $a = 0.8$, $b = 0.75$

Applications of Least Squares

- **Data fitting:** Regression analysis, curve fitting
- **Signal processing:** Filtering, noise reduction
- **Statistics:** Parameter estimation
- **Computer graphics:** Approximation of surfaces
- **Control theory:** System identification
- **Econometrics:** Modeling economic relationships

Orthogonal Distance

The distance from a point $\mathbf{b}$ to a subspace W is:

$$d(\mathbf{b}, W) = \|\mathbf{b} - \text{proj}_W(\mathbf{b})\|$$

For the least squares problem $A\mathbf{x} = \mathbf{b}$, the minimum error is:

$$\min \|A\mathbf{x} - \mathbf{b}\| = d(\mathbf{b}, \text{col}(A))$$

where col(A) is the column space of A.

Gram Matrix

For vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, the **Gram matrix** is:

$$G = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 & \cdots & \mathbf{v}_1 \cdot \mathbf{v}_k \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 & \cdots & \mathbf{v}_2 \cdot \mathbf{v}_k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_k \cdot \mathbf{v}_1 & \mathbf{v}_k \cdot \mathbf{v}_2 & \cdots & \mathbf{v}_k \cdot \mathbf{v}_k \end{pmatrix}$$

The Gram matrix is:

- Always positive semidefinite
- Positive definite if and only if the vectors are linearly independent
- Used to compute volumes of parallelepipeds

Lecture 3 Summary

- Orthogonal projection onto lines and subspaces
- Projection matrices: $P^2 = P$, $P^T = P$
- Least squares: minimize $\|A\mathbf{x} - \mathbf{b}\|$
- Normal equations: $A^T A\mathbf{x} = A^T\mathbf{b}$
- Applications in data fitting, statistics, signal processing
- Distance to subspace: $\|\mathbf{b} - \mathrm{proj}_W(\mathbf{b})\|$
- Gram matrix and its properties