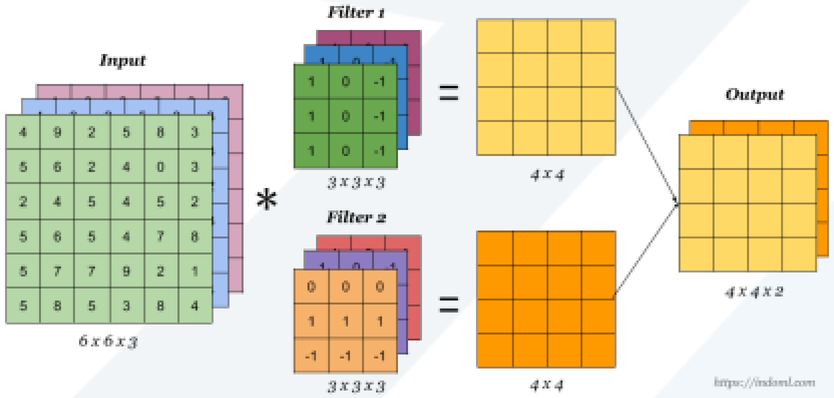
CECC122: Linear Algebra and Matrix Theory

Lecture Notes 6: Inner Product Spaces



Ramez Koudsieh, Ph.D.

Faculty of Engineering

Department of Informatics

Manara University

Chapter 5

Inner Product Spaces

1. General Inner Product
2. Orthonormal Bases: Gram-Schmidt Process
3. Mathematical Models and Least Square Analysis

1. General Inner Product

- **Definition:** Let $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ be vectors in a real vector space V, and let c be any scalar. An inner product on V is a function that associates a real number $<\boldsymbol{u}, \boldsymbol{v}>$ with each pair of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and satisfies the following axioms:

  (1) $<\boldsymbol{u}, \boldsymbol{v}> = <\boldsymbol{v}, \boldsymbol{u}>$

  (2) $<\boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w}> = <\boldsymbol{u}, \boldsymbol{v}> + <\boldsymbol{u}, \boldsymbol{w}>$

  (3) $c <\boldsymbol{u}, \boldsymbol{v}> = <c\boldsymbol{u}, \boldsymbol{v}>$

  (4) $<\boldsymbol{v}, \boldsymbol{v}> \geq 0$ and $<\boldsymbol{v}, \boldsymbol{v}> = 0$ if and only if $\boldsymbol{v} = \mathbf{0}$

- **Notes:**

  (1) $\boldsymbol{u} . \boldsymbol{v}$ = dot product (Euclidean inner product for R^n)
  (2) $<\boldsymbol{u}, \boldsymbol{v}>$ = general inner product for vector space V

- Note: A vector space V with an inner product is called an inner product space.
  Vector space: $(V, +, .)$
  Inner product space: $(V, +, ., <, >)$

- Example 1: (Euclidean inner product for R^n)
  The dot product in R^n satisfies the four axioms of an inner product.
  $$\boldsymbol{u} = (u_1, u_2, \ldots, u_n), \boldsymbol{v} = (v_1, v_2, \ldots, v_n)$$
  $$< \boldsymbol{u}, \boldsymbol{v} > = \boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

- Example 2: (A different inner product for R^n)
  Show that the function defines an inner product on R^2, where $\boldsymbol{u} = (u_1, u_2)$ and $\boldsymbol{v} = (v_1, v_2)$: $<\boldsymbol{u}, \boldsymbol{v}> = u_1v_1 + 2u_2v_2$.
  (1) $<\boldsymbol{u}, \boldsymbol{v}> = u_1v_1 + 2u_2v_2 = v_1u_1 + 2v_2u_2 = <\boldsymbol{v}, \boldsymbol{u}>$

(2) $\boldsymbol{w} = (w_1, w_2)$

$$< \boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w} > = u_1(v_1 + w_1) + 2u_2(v_2 + w_2)$$
$$= u_1v_1 + u_1w_1 + 2u_2v_2 + 2u_2w_2$$
$$= (u_1v_1 + 2u_2v_2) + (u_1w_1 + 2u_2w_2)$$
$$= < \boldsymbol{u}, \boldsymbol{v} > + < \boldsymbol{u}, \boldsymbol{w} >$$

(3) $c <\boldsymbol{u}, \boldsymbol{v}> = c\ (u_1v_1 + 2u_2v_2) = (cu_1)v_1 + 2(cu_2)v_2 = <c\boldsymbol{u}, \boldsymbol{v}>$

(4) $< \boldsymbol{v}, \boldsymbol{v} > = {v_1}^2 + 2{v_2}^2 \geq 0$

$$< \boldsymbol{v}, \boldsymbol{v} > = 0 \Rightarrow {v_1}^2 + 2{v_2}^2 = 0 \Rightarrow v_1 = v_2 = 0 \quad (\boldsymbol{v} = \boldsymbol{0})$$

- Note: (An inner product on R^n)

$< \boldsymbol{u}, \boldsymbol{v} > = c_1u_1v_1 + c_2u_2v_2 + \cdots + c_nu_nv_n, \qquad c_i > 0$ (weights)

- **Example 3: (A function that is not an inner product)**

  Show that the following function is not an inner product on R^3

$$< \boldsymbol{u},\ \boldsymbol{v} > = u_1v_1 - 2u_2v_2 + u_3v_3$$

  Let $v = (1, 2, 1)$, then $<v, v> = (1)(1) - 2(2)(2) + (1)(1) = -6 < 0$

  Axiom 4 is not satisfied. Thus this function is not an inner product on R^3

- **Theorem 1: (Properties of inner products)**

  Let $\boldsymbol{u}$, v and w be vectors in an inner product space V, and let c be any real number.

  (1) $<\mathbf{0},\ \boldsymbol{v}> = <\boldsymbol{v},\ \mathbf{0}> = \mathbf{0}$

  (2) $<\boldsymbol{u} + \boldsymbol{v},\ \boldsymbol{w}> = <\boldsymbol{u},\ \boldsymbol{w}> + <\boldsymbol{v},\ \boldsymbol{w}>$

  (3) $<\boldsymbol{u},\ c\boldsymbol{v}> = c<\boldsymbol{u},\ \boldsymbol{v}>$

- Example 4: (The Standard Inner Product on $M_n(R)$)

$$A, B \in M_n(R), \quad \langle A, B\rangle = \operatorname{tr}(AB^T)$$

for the 2×2 matrices $A = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$

$$\langle A, B\rangle = \operatorname{tr}(AB^T) = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4$$

$$\|A\| = \sqrt{\langle A, A\rangle} = \sqrt{a_1^2 + a_2^2 + a_3^2 + a_4^2}$$

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$$

$$\langle A, B\rangle = \operatorname{tr}(AB^T) = 1(-1) + 2(0) + 3(3) + 4(2) = 16$$

$$\|A\| = \sqrt{\langle A, A\rangle} = \sqrt{1^2 + 2^2 + 3^2 + 4^2} = \sqrt{30}, \quad \|B\| = \sqrt{\langle B, B\rangle} = \sqrt{14}$$

- Example 5: (The Standard Inner Product on P_n)

$p, q \in P_n \qquad p = a_0 + a_1 x + \cdots + a_n x^n \qquad q = b_0 + b_1 x + \cdots + b_n x^n$

$\langle p, q \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n \qquad \|p\| = \sqrt{\langle p, p \rangle} = \sqrt{a_0^2 + a_1^2 + \cdots + a_n^2}$

- Norm (length) of $\boldsymbol{u}$: $\|\boldsymbol{u}\| = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}$
- Note: $\|\boldsymbol{u}\|^2 = \langle \boldsymbol{u}, \boldsymbol{u} \rangle$
- Distance between $\boldsymbol{u}$ and $\boldsymbol{v}$: $d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\| = \sqrt{\langle \boldsymbol{u} - \boldsymbol{v}, \boldsymbol{u} - \boldsymbol{v} \rangle}$
- Angle between two nonzero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$: $\cos\theta = \dfrac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|},\ 0 \le \theta \le \pi$
- Orthogonal: $(\boldsymbol{u} \perp \boldsymbol{v})$ $\qquad$ $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if $<\boldsymbol{u}, \boldsymbol{v}> = 0$

- Notes:

  (1) If $\|\boldsymbol{v}\| = 1$, then $\boldsymbol{v}$ is called a unit vector

  (2) $\|\boldsymbol{v}\| \neq 1$, $\boldsymbol{v} \neq \mathbf{0}$ $\xrightarrow{\text{Normalizing}}$ $\dfrac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$ (the unit vector in the direction of $\boldsymbol{v}$)

- Properties of norm:

  (1) $\|\boldsymbol{u}\| \geq 0$ (2) $\|\boldsymbol{u}\| = 0$ if and only if $\boldsymbol{u} = \mathbf{0}$ (3) $\|c\boldsymbol{u}\| = |c|\|\boldsymbol{u}\|$

- (Properties of distance)

  (1) $d(\boldsymbol{u}, \boldsymbol{v}) \geq 0$ (2) $d(\boldsymbol{u}, \boldsymbol{v}) = 0$ if and only if $\boldsymbol{u} = \boldsymbol{v}$ (3) $d(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{v}, \boldsymbol{u})$

- Note: Norm, Distance and Orthogonality depend on the inner product being used.

- Example 6: $\boldsymbol{u}=$ (1, 0) and $\boldsymbol{v}=$ (0, 1) in R^2

  Euclidean inner product:

  $$\|\boldsymbol{u}\|=\sqrt{1^2+0^2}=1,\quad \|\boldsymbol{v}\|=\sqrt{0^2+1^2}=1,\quad d(\boldsymbol{u},\boldsymbol{v})=\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{1^2+(-1)^2}=\sqrt{2}$$

  Weighted Euclidean inner product: $\langle\boldsymbol{u},\boldsymbol{v}\rangle=3u_1v_1+2u_2v_2$

  $$\|\boldsymbol{u}\|=\sqrt{3(1)^2+2(0)^2}=\sqrt{3},\quad \|\boldsymbol{v}\|=\sqrt{3(0)^2+2(1)^2}=\sqrt{2}$$

  $$d(\boldsymbol{u},\boldsymbol{v})=\|\boldsymbol{u}-\boldsymbol{v}\|=\sqrt{3(1)^2+2(-1)^2}=\sqrt{5}$$

- Example 7: $\boldsymbol{u}=$ (1, 1) and $\boldsymbol{v}=$ (1, –1) in R^2

  Euclidean inner product: $\boldsymbol{u}.\boldsymbol{v}=1(1)+(-1)(1)=0\quad\Rightarrow\quad\boldsymbol{u}\perp\boldsymbol{v}$

  Weighted Euclidean inner product: $\langle\boldsymbol{u},\boldsymbol{v}\rangle=3u_1v_1+2u_2v_2$

  $$\langle\boldsymbol{u},\boldsymbol{v}\rangle=3u_1v_1+2u_2v_2=3(1)(1)+2(-1)(1)=1\neq 0$$

- Theorem 2: Let u and v be vectors in an inner product space V.

  (1) Cauchy-Schwarz inequality: $|\langle \boldsymbol{u},\boldsymbol{v}\rangle| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$

  (2) Triangle inequality: $\|\boldsymbol{u}+\boldsymbol{v}\| \le \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$

  (3) Pythagorean theorem: u and v are orthogonal iff $\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$

Orthogonal Complements

- Definition: If W is a subspace of a real inner product V, then the set of all vectors in V that are orthogonal to every vector in W is called the orthogonal complement of W and is denoted by the symbol $W^\perp$.

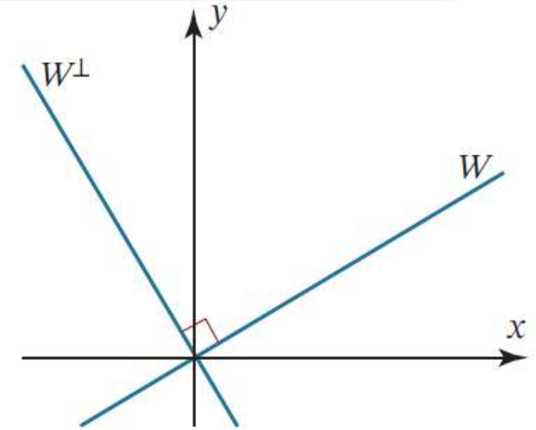
- Theorem 3: (Properties of Orthogonal Complements)

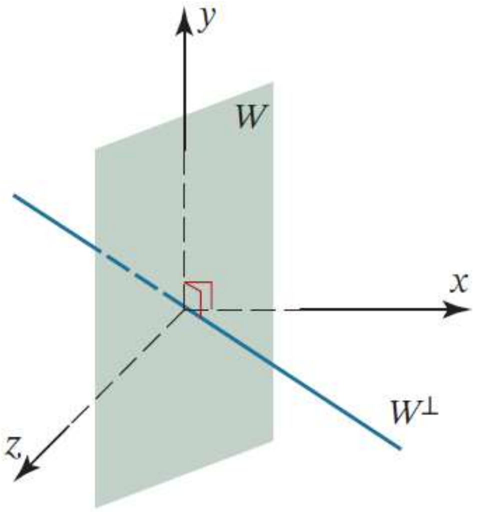
  If W is a subspace of a real inner product V, then:

  (a) $W^\perp$ is a subspace of V (b) $W^\perp \cap W = \{\mathbf{0}\}$

- Example 8: (Orthogonal Complements)

in R^2





in R^3

2. Orthonormal Bases: Gram-Schmidt Process

Definition: A set S of vectors in an inner product space V is called an orthogonal set if every pair of vectors in the set is orthogonal.

$$S = \{v_1, v_2, \cdots, v_n\} \subseteq V \Rightarrow \langle v_i, v_j \rangle = 0, \quad i \neq j$$

An orthogonal set in which each vector is a unit vector is called orthonormal.

$$S=\{v_1, v_2, \cdots, v_n\} \subseteq V \Rightarrow \langle v_i, v_j \rangle = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

- Note: If S is a basis, then it is called an orthogonal/orthonormal basis.
- Example 9: (A nonstandard orthonormal basis for R^3)

  Show that the following set is an orthonormal basis.

$$S=\{(v_1=\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (v_2=-\frac{\sqrt{2}}{6}, \frac{\sqrt{2}}{6}, \frac{2\sqrt{2}}{3}), (v_3=\frac{2}{3}, -\frac{2}{3}, \frac{1}{3})\}$$

$$v_1 . v_2 = -\frac{1}{6}+\frac{1}{6}+0=0 \qquad \|v_1\|=\sqrt{v_1 . v_1}=\sqrt{\frac{1}{2}+\frac{1}{2}+0}=1$$

$$v_1 . v_3 = \frac{2}{3\sqrt{2}}-\frac{2}{3\sqrt{2}}+0=0 \qquad \|v_2\|=\sqrt{v_2 . v_2}=\sqrt{\frac{2}{36}+\frac{2}{36}+\frac{8}{9}}=1$$

$$v_2 . v_3 = -\frac{\sqrt{2}}{9}-\frac{\sqrt{2}}{9}+\frac{2\sqrt{2}}{9}=0 \qquad \|v_3\|=\sqrt{v_3 . v_3}=\sqrt{\frac{4}{9}+\frac{4}{9}+\frac{1}{9}}=1$$

Thus S is an orthonormal set

- Example 10: (An orthonormal basis for P_3)
  with the inner product $\langle p, q\rangle = a_0b_0 + a_1b_1 + a_2b_2 + a_3b_3$
  the standard basis $B = \{1, x, x^2, x^3\}$ is orthonormal
- Theorem 4: (Orthogonal sets are linearly independent)
  If $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.
- Theorem 5: (Coordinates relative to an orthonormal basis)
  If $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is an orthogonal/orthonormal basis for an inner product space V, and if $\boldsymbol{u}$ is any vector in V, then

$$\boldsymbol{u} = \frac{\langle \boldsymbol{u}, \boldsymbol{v}_1\rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{u}, \boldsymbol{v}_2\rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{u}, \boldsymbol{v}_n\rangle}{\|\boldsymbol{v}_n\|^2}\boldsymbol{v}_n \qquad S \text{ orthogonal}$$

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{u}, \mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{u}, \mathbf{v}_n\rangle \mathbf{v}_n$$ S orthonormal

- Note: If $S = \{v_1, v_2, \ldots, v_n\}$ is an orthogonal/orthonormal basis for an inner product space V and $w \in V$, then the corresponding coordinate matrix of w relative to B is

$$[\mathbf{u}]_S = \left(\frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}, \frac{\langle \mathbf{u}, \mathbf{v}_2\rangle}{\|\mathbf{v}_2\|^2}, \cdots, \frac{\langle \mathbf{u}, \mathbf{v}_n\rangle}{\|\mathbf{v}_n\|^2} \right)^T$$

S orthogonal

$$[\mathbf{w}]_S = \left(\langle \mathbf{u}, \mathbf{v}_1\rangle, \langle \mathbf{u}, \mathbf{v}_2\rangle, \cdots, \langle \mathbf{u}, \mathbf{v}_n\rangle \right)^T$$

S orthonormal

- Example 11: (Representing vectors relative to an orthonormal basis)

Find the coordinates of vector $w = (5, -5, 2)$ relative to the following orthonormal basis for R^3 $S = \{(\frac{3}{5}, \frac{4}{5}, 0), (-\frac{4}{5}, \frac{3}{5}, 0), (0, 0, 1)\}$

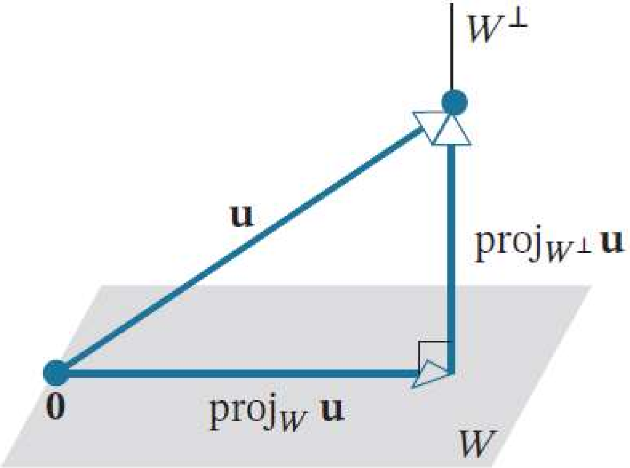
$$< \boldsymbol{w},\boldsymbol{v}_1 > = \boldsymbol{w}.\boldsymbol{v}_1 = (5, -5, 2) \,.\, (\tfrac{3}{5}, \tfrac{4}{5}, 0) = -1$$
$$< \boldsymbol{w},\boldsymbol{v}_2 > = \boldsymbol{w}.\boldsymbol{v}_2 = (5, -5, 2) \,.\, (-\tfrac{4}{5}, \tfrac{3}{5}, 0) = -7 \quad \Rightarrow [\boldsymbol{w}]_S = \begin{bmatrix} -1 \\ -7 \\ 2 \end{bmatrix}$$
$$< \boldsymbol{w},\boldsymbol{v}_3 > = \boldsymbol{w}.\boldsymbol{v}_3 = (5, -5, 2) \,.\, (0, 0, 1) = 2$$

Orthogonal Projections

- **Theorem 6: (Projection Theorem)**

  If W is a finite-dimensional subspace of an inner product space V, then every vector $\boldsymbol{u}$ in V can be expressed in exactly one way as $\boldsymbol{u} = \boldsymbol{w}_1 + \boldsymbol{w}_2$, where $\boldsymbol{w}_1$ is in W and $\boldsymbol{w}_2$ is in $W^{\perp}$.

$$\boldsymbol{u} = \text{proj}_W \boldsymbol{u} + \text{proj}_{W^\perp} \boldsymbol{u} = \text{proj}_W \boldsymbol{u} + (\boldsymbol{u} - \text{proj}_W \boldsymbol{u})$$



- Theorem 7: (formulas for calculating orthogonal projection)

  Let W be a finite-dimensional subspace of an inner product space V. If $S = \{v_1, v_2, \ldots, v_r\}$ is an orthogonal/orthonormal basis for W, then

$$\operatorname{proj}_W \boldsymbol{u} = \frac{\langle \boldsymbol{u}, \boldsymbol{v}_1\rangle}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 + \frac{\langle \boldsymbol{u}, \boldsymbol{v}_2\rangle}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 + \cdots + \frac{\langle \boldsymbol{u}, \boldsymbol{v}_r\rangle}{\|\boldsymbol{v}_r\|^2}\boldsymbol{v}_r \qquad S \text{ orthogonal}$$

$$\operatorname{proj}_W \boldsymbol{u} = \langle \boldsymbol{u}, \boldsymbol{v}_1\rangle\boldsymbol{v}_1 + \langle \boldsymbol{u}, \boldsymbol{v}_2\rangle\boldsymbol{v}_2 + \cdots + \langle \boldsymbol{u}, \boldsymbol{v}_r\rangle\boldsymbol{v}_r \qquad S \text{ orthonormal}$$

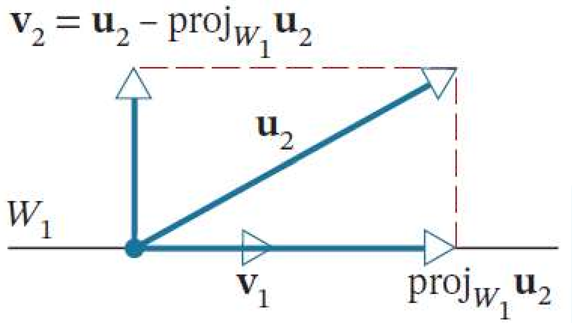
The Gram-Schmidt Process

- Theorem 8: (Projection Theorem)

  Every nonzero finite-dimensional inner product space has an orthonormal basis.

Proof (Gram-Schmidt orthonormalization construction)

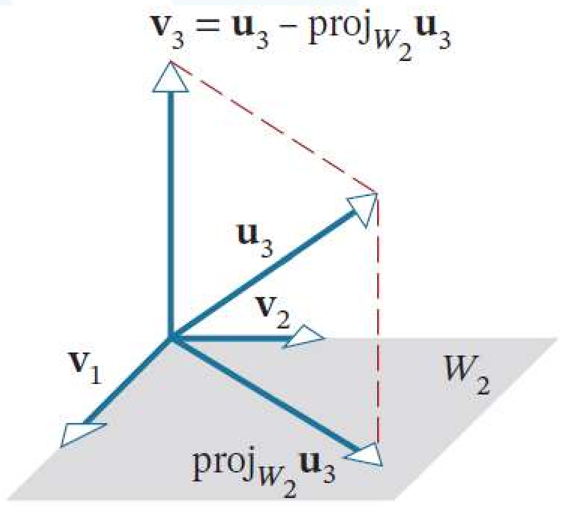
Let W be any nonzero finite-dimensional subspace of an inner product space, and suppose that $\{u_1, u_2, \ldots, u_r\}$ is any basis for W.



Step 1: Let $v_1 = u_1$

Step 2: $v_2 = u_2 - \text{proj}_{W_1} u_2 = u_2 - \dfrac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1$

$W_1 = \text{span}(v_1)$ and $v_2 \perp v_1$, $v_2 \neq \mathbf{0}$



Step 3: $v_3 = u_3 - \text{proj}_{W_2} u_3 = u_3 - \dfrac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \dfrac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2$

$W_2 = \text{span}(v_1, v_2)$ and $v_3 \perp W_2$, $v_3 \neq \mathbf{0}$

Continuing in this way we will produce after r steps an orthogonal set of nonzero vectors $\{v_1, v_2, \ldots, v_r\}$.

By normalizing these basis vectors we can obtain an orthonormal basis.

- Theorem 9: (Gram-Schmidt orthonormalization process)

  (1) Let $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ is a basis for an inner product space V

  (2) Let $B' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, where

$$\boldsymbol{v}_1 = \boldsymbol{u}_1$$

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\langle \boldsymbol{u}_2, \boldsymbol{v}_1 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle}\boldsymbol{v}_1$$

$$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \frac{\langle \boldsymbol{u}_3, \boldsymbol{v}_1 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle}\boldsymbol{v}_1 - \frac{\langle \boldsymbol{u}_3, \boldsymbol{v}_2 \rangle}{\langle \boldsymbol{v}_2, \boldsymbol{v}_2 \rangle}\boldsymbol{v}_2$$

$$\vdots$$

$$\boldsymbol{v}_n = \boldsymbol{u}_n - \sum_{i=1}^{n-1} \frac{\langle \boldsymbol{u}_n, \boldsymbol{v}_i \rangle}{\langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle}\boldsymbol{v}_i$$

Then B' is an orthogonal basis for V

(3) Let $w_i = \dfrac{v_i}{\|v_i\|}$

Then $B'' = \{w_1, w_2, \ldots, w_n\}$ is an orthonormal basis for V

Also, span$\{u_1, u_2, \ldots, u_n\} =$ span$\{w_1, w_2, \ldots, w_n\}$ for $k = 1, 2, ..., n$

- Example 12: (Applying the Gram-Schmidt orthonormalization process)

  Apply the Gram-Schmidt orthonormalization process to the basis B for R^2

$$B = \{u_1 = (1, 1),\ u_2 = (0, 1)\}$$

$$v_1 = u_1 = (1, 1)$$

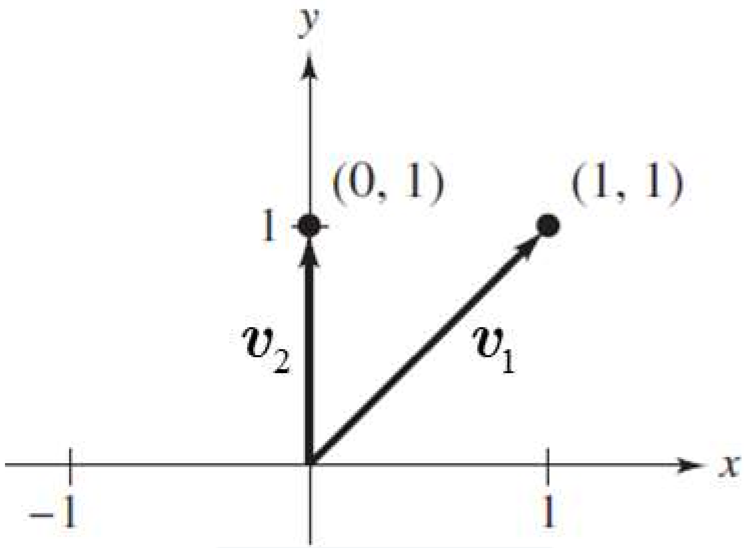
$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (0, 1) - \frac{1}{2}(1, 1) = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

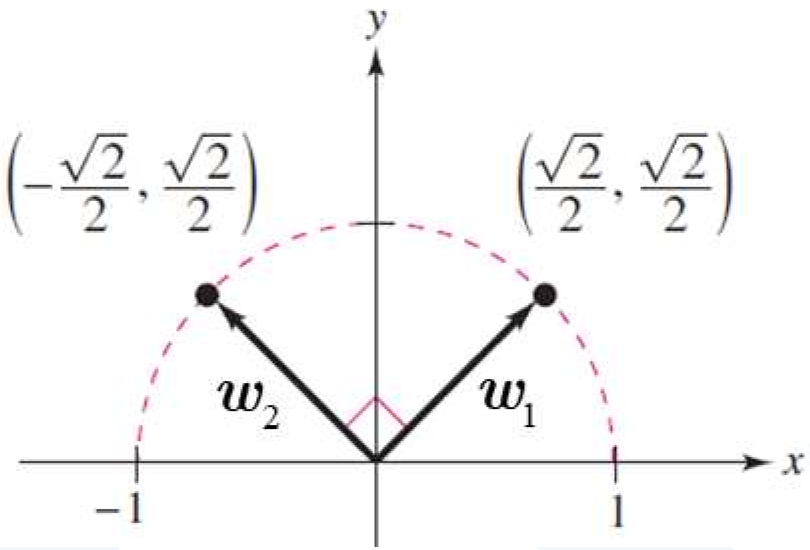
The set $B' = \{v_1, v_2\}$ is an orthogonal basis for R^2

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(1, 1) = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$$

$$w_2 = \frac{v_2}{\|v_2\|} = \frac{1}{1/\sqrt{2}}(-\frac{1}{2}, \frac{1}{2}) = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$$

The set $B'' = \{w_1, w_2\}$ is an orthonormal basis for R^2





- Example 13: (Applying the Gram-Schmidt orthonormalization process)

Apply the Gram-Schmidt orthonormalization process to the basis B for R^3

$$B = \{u_1 = (1, 1, 0),\ u_2 = (1, 2, 0),\ u_3 = (0, 1, 2)\}$$

$$v_1 = u_1 = (1, 1, 0)$$

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (1, 2, 0) - \frac{3}{2}(1, 1, 0) = (-\frac{1}{2}, \frac{1}{2}, 0)$$

$$v_3 = u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2 = (1, 2, 0) - \frac{1}{2}(1, 1, 0) - \frac{1/2}{1/2}(-\frac{1}{2}, \frac{1}{2}, 0) = (0, 0, 2)$$

The set $B' = \{v_1, v_2, v_3\}$ is an orthogonal basis for R^3

$$w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}(1, 1, 0) = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$$

$$w_2 = \frac{v_2}{\|v_2\|} = \frac{1}{1/\sqrt{2}}\left(-\frac{1}{2}, \frac{1}{2}, 0\right) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)$$

$$w_3 = \frac{v_3}{\|v_3\|} = \frac{1}{2}(0, 0, 2) = (0, 0, 1)$$

The set $B'' = \{w_1, w_2, w_3\}$ is an orthonormal basis for R^3

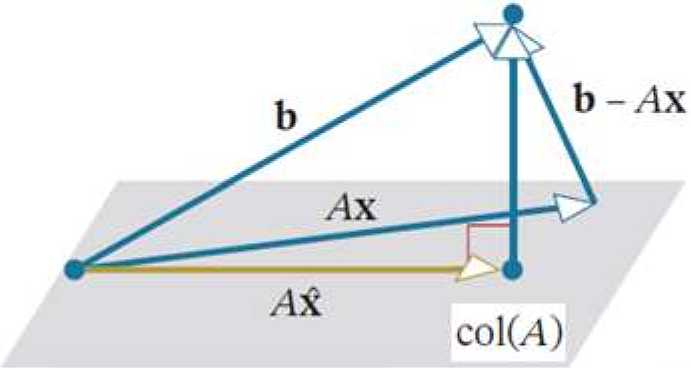
3. Mathematical Models and Least Square Analysis

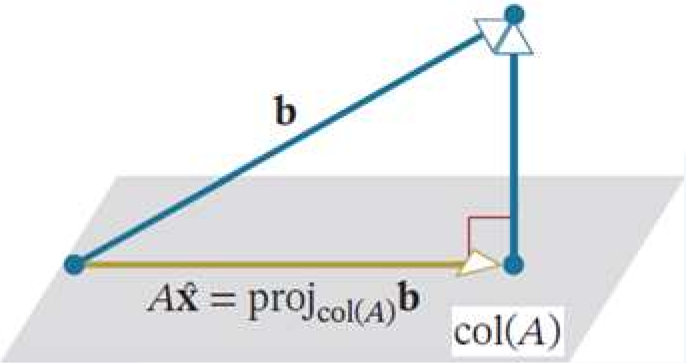
Best Approximation; Least Squares

- **Least Squares Problem:** Given $Ax = b$ of m equations in n unknowns, find x in R^n that minimizes $\|b - Ax\|$ with respect to the Euclidean inner product on R^m. We call x, if it exists, a least squares solution of $Ax = b$, $b - Ax$ the least squares error vector, and $\|b - Ax\|$ the least squares error.

$$\boldsymbol{b} - A\boldsymbol{x} = \begin{bmatrix} e_1 \\ e_1 \\ \vdots \\ e_m \end{bmatrix} \Rightarrow \|\boldsymbol{b} - A\boldsymbol{x}\|^2 = e_1^2 + e_2^2 + \cdots + e_m^2$$

- Note: For every vector $\boldsymbol{x}$ in R^n, the product $A\boldsymbol{x}$ is in the column space of A because it is a linear combination of the column vectors of A. Find a least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ is equivalent to find a vector $A\widehat{\boldsymbol{x}}$ in the col(A) that is closest to $\boldsymbol{b}$ (it minimizes the length of the vector $\boldsymbol{b} - A\boldsymbol{x}$) $\Rightarrow A\widehat{\boldsymbol{x}} = \text{proj}_{\text{col}(A)}\boldsymbol{b}$.





- Theorem 10: (Best Approximation Theorem)

  If W is a finite-dimensional subspace of an inner product space V, and if $\boldsymbol{b}$ is a vector in V, then $\text{proj}_W \boldsymbol{b}$ is the best approximation to $\boldsymbol{b}$ from W in the sense that $\|\boldsymbol{b} - \text{proj}_W \boldsymbol{b}\| < \|\boldsymbol{b} - \boldsymbol{w}\|$ for every vector $\boldsymbol{w}$ in W that is different from $\text{proj}_W \boldsymbol{b}$.

- If $V = R^n$ and $W = \text{col}(A)$, then the best approximation to $\boldsymbol{b}$ from $\text{col}(A)$ is $\text{proj}_{\text{col}(A)} \boldsymbol{b}$.

- Finding Least Squares Solutions: $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$

  This is called the normal equations associated with $A\boldsymbol{x} = \boldsymbol{b}$.

- Example 14: (Finding Least Squares Solutions)

  Find the Least Squares Solution, the least squares error vector, and the least squares error of the linear system:

$$\begin{array}{rcrcl} x & - & y & = & 4 \\ 3x & + & 2y & = & 1 \\ -2x & + & 4y & = & 3 \end{array}$$

$$A^T A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

$$A^T \boldsymbol{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

$$A^T A\boldsymbol{x} = A^T \boldsymbol{b} \Rightarrow \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 17/95 \\ 143/285 \end{bmatrix}$$

$$\boldsymbol{b} - A\boldsymbol{x} = \begin{bmatrix} 1232/285 \\ -154/285 \\ 77/57 \end{bmatrix}, \quad \text{and} \quad \|\boldsymbol{b} - A\boldsymbol{x}\| \approx 4.556$$

- Theorem 11: (a unique least squares solution)

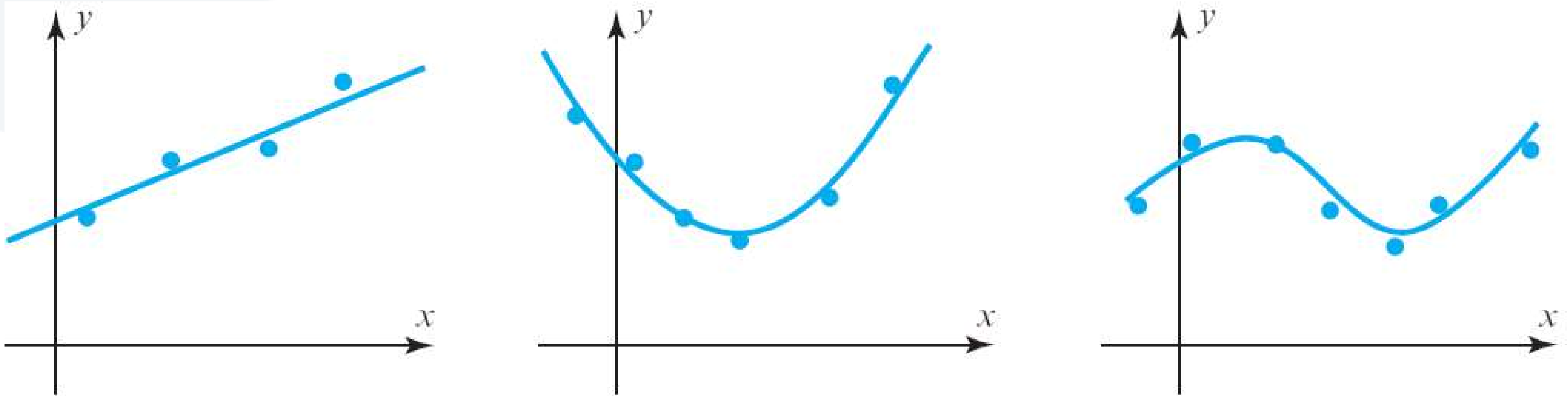
  If A is an $m\text{x}n$ matrix with linearly independent column vectors, then for every $m\text{x}1$ matrix $\boldsymbol{b}$, the linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least squares solution. This solution is given by: $\boldsymbol{x} = (A^T A)^{-1} A^T \boldsymbol{b}$.

  Moreover, if W is the column space of A, then the orthogonal projection of $\boldsymbol{b}$ on W is: $\text{proj}_W \boldsymbol{b} = A\boldsymbol{x} = A(A^T A)^{-1} A^T \boldsymbol{b}$.

Mathematical Modeling Using Least Squares

- Fitting a Curve to Data

  A common problem in experimental work is to find a mathematical relationship $y = f(x)$ between two variables x and y by "fitting" a curve to points in the plane $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.



(a) $y = a + bx$ (b) $y = a + bx + cx^2$ (c) $y = a + bx + cx^2 + dx^3$

mathematical model

Least Squares Fit of a Straight Line $y = a + bx$

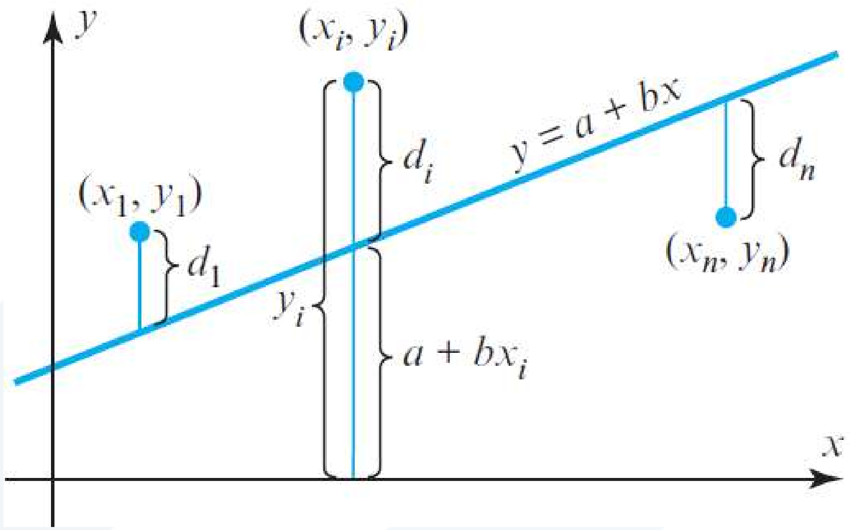
$$\begin{aligned} y_1 &= a + bx_1 \\ y_2 &= a + bx_2 \\ &\vdots \\ y_n &= a + bx_n \end{aligned} \quad \Rightarrow \quad M\boldsymbol{v} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \boldsymbol{y}$$

$$M\boldsymbol{v}=\boldsymbol{y} \quad \Rightarrow \quad M^T M\boldsymbol{v}=M^T\boldsymbol{y} \quad \Rightarrow \quad \boldsymbol{v}^*=\begin{bmatrix} a^* \\ b^* \end{bmatrix}=(M^T M)^{-1}M^T\boldsymbol{y}$$

$y=a^*+b^*x$ Least squares line of best fit or the regression line

It minimizes $\|\boldsymbol{y}-M\boldsymbol{v}\|^2=[y_1-(a+bx_1)]^2+[y_2-(a+bx_2)]^2+\cdots+[y_n-(a+bx_n)]^2$

$d_1=|y_1-(a+bx_1)|, d_2=|y_2-(a+bx_2)|, \cdots, d_n=|y_n-(a+bx_n)|$ residuals

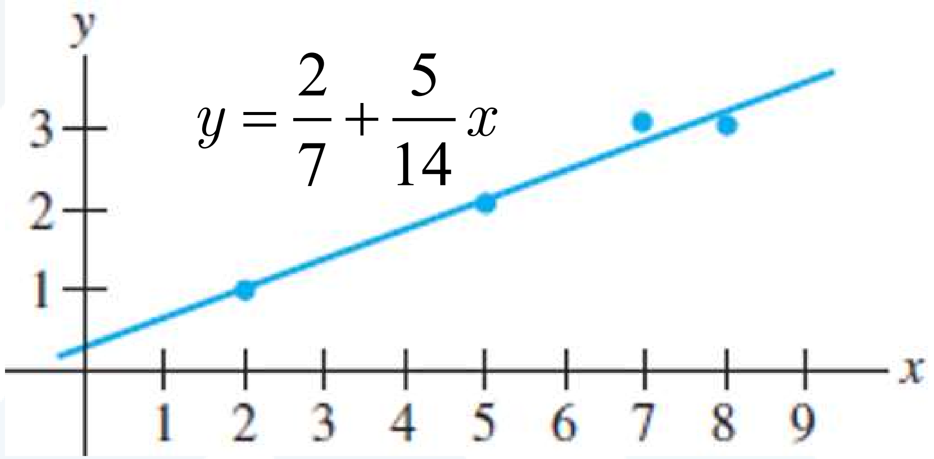


- Example 15: (Least Squares Straight Line Fit)

Find the least squares straight line fit to the points (2, 1), (5, 2), (7, 3), and (8, 3)

$$M^T M = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$M^T \boldsymbol{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$



$$\boldsymbol{v}^* = (M^T M)^{-1} M^T \boldsymbol{y} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ \frac{5}{14} \end{bmatrix}$$

Least Squares Fit of a Polynomial $y = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

$$\begin{aligned}
a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_m x_1^m &= y_1 \\
a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_m x_2^m &= y_2 \\
&\vdots \\
a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_m x_n^m &= y_n
\end{aligned}$$

$$M\boldsymbol{v} = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^m \\ 1 & x_2 & x_2^2 & \ldots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \boldsymbol{y}$$

$$M\boldsymbol{v} = \boldsymbol{y} \quad \Rightarrow \quad M^T M\boldsymbol{v} = M^T \boldsymbol{y} \quad \Rightarrow \quad \boldsymbol{v}^* = (M^T M)^{-1} M^T \boldsymbol{y}$$

- Example 16: (Fitting a Quadratic Curve to Data)

Newton's second law of motion $s = s_0 + v_0 t + \frac{1}{2} g t^2$

Laboratory experiment

Time t (sec)	.1	.2	.3	.4	.5
Displacement s (ft)	-0.18	0.31	1.03	2.48	3.73

Approximate g

Let $s = a_0 + a_1 t + a_2 t^2$

(0.1,–0.18), (0.2, 0.31), (0.3, 1.03), (0.4, 2.48), (0.5, 3.73)

$$M=\begin{bmatrix}1 & t_1 & t_1^2\\ 1 & t_2 & t_2^2\\ 1 & t_3 & t_3^2\\ 1 & t_4 & t_4^2\\ 1 & t_5 & t_5^2\end{bmatrix}=\begin{bmatrix}1 & 0.1 & 0.01\\ 1 & 0.2 & 0.04\\ 1 & 0.3 & 0.09\\ 1 & 0.4 & 0.16\\ 1 & 0.5 & 0.25\end{bmatrix},\quad \boldsymbol{y}=\begin{bmatrix}s_1\\ s_2\\ s_3\\ s_4\\ s_5\end{bmatrix}=\begin{bmatrix}-0.18\\ 0.31\\ 1.03\\ 2.48\\ 3.73\end{bmatrix}$$

$$\boldsymbol{v}^*=\begin{pmatrix}a_0^*\\ a_1^*\\ a_2^*\end{pmatrix}=(M^TM)^{-1}M^T\boldsymbol{y}=\begin{pmatrix}-0.4\\ 0.35\\ 16.1\end{pmatrix}$$

$$g=2a_2^*=2(16.1)=32.2\text{ feet}/s^2$$

$$s_0=a_0^*=-0.4\text{ feet}\qquad v_0=a_1^*=0.35\text{ feet/s}$$