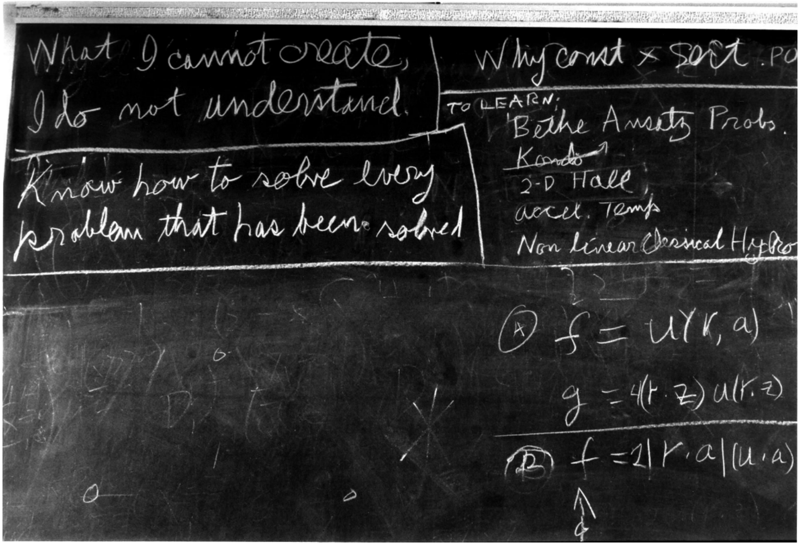
Orthogonality and Linear Transformations

Wang Xiao

Shanghai University of Finance and Economics

April 8, 2026



What I cannot create,
I do not understand.

Why const × sort.PO

TO LEARN:
Bethe Ansatz Probs.
Kondo
2-D Hall
accel. Temp
Non Linear Classical Hydro

Know how to solve every
problem that has been solved

$$\textcircled{A}\quad f = u(r, a)$$

$$g = 4(r \cdot z)\, u(r \cdot z)$$

$$\textcircled{B}\quad f = 2|r \cdot a|(u \cdot a)$$

Inconsistent Linear Systems

Key Problem

For $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$, the system $A\mathbf{x} = \mathbf{b}$ is *inconsistent* if no solution exists (intersection of hyperplanes is empty).

Goal: Find the "best approximation" to a solution.

New perspective on solving a linear system

Finding a set of coefficients $\mathbf{x} = (x_1, ..., x_n)^T$ satisfying the system $A\mathbf{x} = \mathbf{b}$ is equivalent to finding a linear combination of the column vectors of A, i.e., $x_1 A^1 + ... + x_n A^n$, so that

$$x_1 A^1 + ... + x_n A^n = \mathbf{b}.$$

Therefore, if the linear system $A\mathbf{x} = \mathbf{b}$ cannot be solved exactly, the best we can do is to find a set of coefficients $\mathbf{x} = (x_1, ..., x_n)^\top$ so that it "*almost*" solves the system, and this can be geometrically translated into the following:

The distance between the linear combination $x_1 A^1 + ... + x_n A^n$ *and* $\mathbf{b}$ *is minimized*

Formally, this leads to the concept of *least square solution.*

Definition (Least Squares Solution)

A vector $\hat{\mathbf{x}} \in \mathbb{R}^n$ is a *least squares solution* of $A\mathbf{x} = \mathbf{b}$ if:

$$\|A\hat{\mathbf{x}} - \mathbf{b}\| \leq \|A\mathbf{x} - \mathbf{b}\| \quad \forall \mathbf{x} \in \mathbb{R}^n$$

(Minimizes the distance from $\mathbf{b}$ to the column space of A.)

Projection onto a Subspace

Definition (Projection of b onto V)

Let $V \subset \mathbb{R}^m$ be a finite-dimensional subspace. The *projection* of $\mathbf{b} \in \mathbb{R}^m$ onto V, denoted $\text{proj}_V \mathbf{b}$, is the unique vector in V such that:

$$\mathbf{b} - \text{proj}_V \mathbf{b} \perp V \quad (\text{i.e., } (\mathbf{b} - \text{proj}_V \mathbf{b}) \cdot \mathbf{v} = 0 \forall \mathbf{v} \in V)$$

Proposition

The least squares solution $\hat{\mathbf{x}}$ of $A\mathbf{x} = \mathbf{b}$ satisfies $A\hat{\mathbf{x}} = \text{proj}_{C(A)} \mathbf{b}$, where $C(A)$ is the column space of A.

Normal Equations for Least Squares

Theorem (Normal Equations)

A vector $\hat{\mathbf{x}}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if it satisfies:

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

Sketch.

$A\hat{\mathbf{x}} = \mathrm{proj}_{C(A)}\mathbf{b} \iff \mathbf{b} - A\hat{\mathbf{x}} \perp C(A)$
$\iff (\mathbf{b} - A\hat{\mathbf{x}}) \cdot A\mathbf{e}_j = 0$ for all standard basis vectors $\mathbf{e}_j$
$\iff A^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0 \iff A^T A\hat{\mathbf{x}} = A^T\mathbf{b}.$ □

Corollary

If $A^T A$ is invertible (columns of A are linearly independent), there is a unique least squares solution:

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

Example: Fitting a Line to Data

Example

Find the least squares line $y = mx + c$ for data points $(1,2), (2,3), (3,5)$.

Step 1: Set up the system $A\mathbf{x} = \mathbf{b}$

For $\mathbf{x} = \begin{bmatrix} m \\ c \end{bmatrix}$, we have:

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$

$A\mathbf{x} = \mathbf{b}$ is inconsistent (no line passes through all points).

Step 2: Compute normal equations

$$A^T A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}, \quad A^T \mathbf{b} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 23 \\ 10 \end{bmatrix}$$

Solve $\begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 23 \\ 10 \end{bmatrix}$:

$$m = \frac{3}{2}, \quad c = \frac{1}{3} \implies y = \frac{3}{2}x + \frac{1}{3}$$

Orthogonal and Orthonormal Sets

Definition (Orthogonal Set)

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subset \mathbb{R}^n$ is *orthogonal* if:

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \forall i \neq j$$

It is *orthonormal* if additionally $\|\mathbf{v}_i\| = 1$ for all i.

Proposition

An orthogonal set of *nonzero* vectors is linearly independent.

Proof.

Suppose $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0$. Dot both sides with $\mathbf{v}_i$:

$$c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = 0 \implies c_i = 0 \quad (\text{since } \mathbf{v}_i \neq 0 \implies \mathbf{v}_i \cdot \mathbf{v}_i > 0)$$

Corollary

An orthogonal set of k nonzero vectors in $\mathbb{R}^n$ is a basis for its span (dimension k).

Projection Formula for Orthogonal Bases

Theorem

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal basis for $V \subset \mathbb{R}^m$. For any $\mathbf{b} \in \mathbb{R}^m$:

$$\text{proj}_V \mathbf{b} = \sum_{i=1}^{k} \frac{\mathbf{b} \cdot \mathbf{v}_i}{\|\mathbf{v}_i\|^2} \mathbf{v}_i$$

If the basis is orthonormal ($\|\mathbf{v}_i\| = 1$), this simplifies to:

$$\text{proj}_V \mathbf{b} = \sum_{i=1}^{k} (\mathbf{b} \cdot \mathbf{v}_i) \mathbf{v}_i$$

Idea.
Write $\mathrm{proj}_V \mathbf{b} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$. Dot with $\mathbf{v}_j$:

$$(\mathrm{proj}_V \mathbf{b}) \cdot \mathbf{v}_j = c_j(\mathbf{v}_j \cdot \mathbf{v}_j) \implies c_j = \frac{\mathbf{b} \cdot \mathbf{v}_j}{\|\mathbf{v}_j\|^2}$$

(Since $\mathbf{b} - \mathrm{proj}_V \mathbf{b} \perp \mathbf{v}_j$, so $\mathbf{b} \cdot \mathbf{v}_j = (\mathrm{proj}_V \mathbf{b}) \cdot \mathbf{v}_j$.) $\square$

Theorem (Gram-Schmidt process)

Given a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ *for an inner product space* V*, we obtain an orthogonal basis* $\{\mathbf{w}_1, \dots, \mathbf{w}_k\}$ *for* V *as follows:*

$$\begin{aligned} \mathbf{w}_1 &= \mathbf{v}_1 \\ \mathbf{w}_2 &= \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 \\ &\vdots \end{aligned}$$

and, assuming $\mathbf{w}_1, \dots, \mathbf{w}_j$ *have been defined,*

$$\begin{aligned} \mathbf{w}_{j+1} &= \mathbf{v}_{j+1} - \frac{\mathbf{v}_{j+1} \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\mathbf{v}_{j+1} \cdot \mathbf{w}_2}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \dots - \frac{\mathbf{v}_{j+1} \cdot \mathbf{w}_j}{\|\mathbf{w}_j\|^2}\mathbf{w}_j \\ &\vdots \\ \mathbf{w}_k &= \mathbf{v}_k - \frac{\mathbf{v}_k \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\mathbf{v}_k \cdot \mathbf{w}_2}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 - \dots - \frac{\mathbf{v}_k \cdot \mathbf{w}_{k-1}}{\|\mathbf{w}_{k-1}\|^2}\mathbf{w}_{k-1}. \end{aligned}$$

Example

Let $\mathbf{v}_1 = (1,1,1,1)^\top, \mathbf{v}_2 = (3,1,-1,1)^\top$ and $\mathbf{v}_3 = (1,1,3,3)$. Use Gram-Schmit process to give an orthogonal basis for $V = \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \subset \mathbb{R}^4$. We can take

$$\mathbf{w}_1 = \mathbf{v}_1 = (1,1,1,1);$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 = (2,0,-2,0)$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_1}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 = (0,-1,0,1)$$

We can furthermore make them a orthonormal basis by normalization with respect to norms.

Example: Projection with Orthogonal Basis

Example

Let $V = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ where $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (1, -1, 2)$ (orthogonal). Compute $\text{proj}_V\mathbf{b}$ for $\mathbf{b} = (3, 1, 4)$.

Step 1: Verify orthogonality

$\mathbf{v}_1 \cdot \mathbf{v}_2 = (1)(1) + (1)(-1) + (0)(2) = 0$ (orthogonal).

Step 2: Compute coefficients

$$\|\mathbf{v}_1\|^2 = 1^2 + 1^2 + 0^2 = 2, \quad \mathbf{b} \cdot \mathbf{v}_1 = 3(1) + 1(1) + 4(0) = 4$$

$$\|\mathbf{v}_2\|^2 = 1^2 + (-1)^2 + 2^2 = 6, \quad \mathbf{b} \cdot \mathbf{v}_2 = 3(1) + 1(-1) + 4(2) = 10$$

Step 3: Apply projection formula

$$\text{proj}_V\mathbf{b} = \frac{4}{2}\mathbf{v}_1 + \frac{10}{6}\mathbf{v}_2 = 2(1, 1, 0) + \frac{5}{3}(1, -1, 2) = \left(\frac{11}{3}, \frac{1}{3}, \frac{10}{3}\right)$$

Orthogonal Complements and Projections

Definition (Orthogonal Complement)

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

$$V^\perp = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{v} = 0 \forall \mathbf{v} \in V\}$$

Theorem (Direct Sum Decomposition)

For any subspace $V \subset \mathbb{R}^n$:

$$\mathbb{R}^n = V \oplus V^\perp$$

Every vector $\mathbf{x} \in \mathbb{R}^n$ has a unique decomposition:

$$\mathbf{x} = \text{proj}_V \mathbf{x} + \text{proj}_{V^\perp} \mathbf{x}$$

where $\text{proj}_V \mathbf{x} \in V$ and $\text{proj}_{V^\perp} \mathbf{x} = \mathbf{x} - \text{proj}_V \mathbf{x} \in V^\perp$.

Corollary

$(V^\perp)^\perp = V$ and $\dim V + \dim V^\perp = n$.

Orthogonality of Fundamental Subspaces

Recall the four fundamental subspaces of $A \in \mathbb{R}^{m \times n}$: 1. Column space $C(A) \subset \mathbb{R}^m$ 2. Nullspace $N(A) \subset \mathbb{R}^n$ 3. Row space $R(A) = C(A^T) \subset \mathbb{R}^n$ 4. Left nullspace $N(A^T) \subset \mathbb{R}^m$

Theorem (Orthogonality Relations)

$$N(A) = R(A)^{\perp} \quad \text{and} \quad N(A^T) = C(A)^{\perp}$$
$$C(A) = N(A^T)^{\perp} \quad \text{and} \quad R(A) = N(A)^{\perp}$$

Sketch for $N(A) = R(A)^{\perp}$.

$\mathbf{x} \in N(A) \iff A\mathbf{x} = 0 \iff$ each row of A is orthogonal to $\mathbf{x} \iff \mathbf{x} \perp R(A)$. $\square$

Linear Transformations: Recap

Definition (Linear Transformation)

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is *linear* if: 1. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
2. $T(c\mathbf{x}) = cT(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $c \in \mathbb{R}$

Proposition

Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be written as $T(\mathbf{x}) = A\mathbf{x}$ for a unique $m \times n$ matrix A (called the *standard matrix* of T).

Standard Matrix Construction

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{R}^n$. Then:

$$A = \begin{bmatrix} | & | & \cdots & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & \cdots & | \end{bmatrix}$$

For any $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, $T(x) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) = A\mathbf{x}$.

Example: Standard Matrix of Rotation

Example

Find the standard matrix of $T : \mathbb{R}^2 \to \mathbb{R}^2$, rotation by θ counterclockwise.

Step 1: Compute $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$

- Standard basis: $\mathbf{e}_1 = (1, 0)$, $\mathbf{e}_2 = (0, 1)$ - Rotation of $\mathbf{e}_1$: $T(\mathbf{e}_1) = (\cos\theta, \sin\theta)$ - Rotation of $\mathbf{e}_2$: $T(\mathbf{e}_2) = (-\sin\theta, \cos\theta)$

Step 2: Construct the standard matrix

$$[T]_E = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Verify: For $x = (x_1, x_2)$,

$$[T]_E x = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (x_1 \cos\theta - x_2 \sin\theta, x_1 \sin\theta + x_2 \cos\theta)$$

(Matches geometric rotation formula.)

Matrix of T with Respect to a Basis B

Definition (Matrix of $T : V \to V$ w.r.t. B)

Let V be finite-dimensional, $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ an ordered basis for V, and $T : V \to V$ linear. Define a_{ij} by:

$$T(\mathbf{v}_j) = a_{1j}\mathbf{v}_1 + a_{2j}\mathbf{v}_2 + \cdots + a_{nj}\mathbf{v}_n$$

The *matrix of* T *with respect to* B is:

$$[T]_B = \begin{bmatrix} | & | & \cdots & | \\ [T(\mathbf{v}_1)]_B & [T(\mathbf{v}_2)]_B & \cdots & [T(\mathbf{v}_n)]_B \\ | & | & \cdots & | \end{bmatrix}$$

where $[w]_B$ is the coordinate vector of w w.r.t. B.

Proposition (Coordinate Transformation)

For any $v \in V$:

$$[T(v)]_B = [T]_B [v]_B$$

Sketch.

Write $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. Then

$$T(\mathbf{v}) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j) = \sum_{j=1}^{n} c_j \sum_{i=1}^{n} a_{ij}\mathbf{v}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} c_j \right) \mathbf{v}_i.$$

Thus, $[T(\mathbf{v})]_B = [T]_B[\mathbf{v}]_B$. □

Example: Matrix of Projection w.r.t. Orthogonal Basis

Example

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be projection onto the line $y = x$. Let $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ where $\mathbf{v}_1 = (1, 1)$ (along line), $\mathbf{v}_2 = (1, -1)$ (perpendicular). Find $[T]_B$.

Step 1: Compute $T(\mathbf{v}_1)$ and $T(\mathbf{v}_2)$

- $\mathbf{v}_1$ is on the line: $T(\mathbf{v}_1) = \mathbf{v}_1 = 1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2$ - $\mathbf{v}_2$ is perpendicular to the line: $T(\mathbf{v}_2) = 0 = 0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2$

Step 2: Find coordinate vectors

$$[T(\mathbf{v}_1)]_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad [T(\mathbf{v}_2)]_B = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Step 3: Construct $[T]_B$

$$[T]_B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

(Diagonal! Orthogonal bases simplify matrix representations.)

Change of Basis w.r.t. standard basis

Proposition

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with standard matrix $[T]_E$. Let $B = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ be an ordered basis for $\mathbb{R}^n$ and let $[T]_B$ be the matrix for T with respect to B. Let P be the $n \times n$ matrix whose columns are given by the vectors $\mathbf{v}_1, ..., \mathbf{v}_n$. Then we have

$$[T]_E P = P[T]_B.$$

We call P the *change of basis matrix from the standard basis to the basis* B.

Proof

Recall the matrix of T w.r.t. B is

$$\begin{aligned} T(\mathbf{v}_1) &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + ... + a_{n1}\mathbf{v}_n = [\mathbf{v}_1, ..., \mathbf{v}_n]\,[T]_B^1 \\ &\vdots \\ T(\mathbf{v}_n) &= a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + ... + a_{nn}\mathbf{v}_n = [\mathbf{v}_1, ..., \mathbf{v}_n]\,[T]_B^n \end{aligned}$$

On the other hand, each $\mathbf{v}_1, ..., \mathbf{v}_n$ can be expressed as a linear combination of the standard basis,

$$\begin{aligned} \mathbf{v}_1 &= [\mathbf{e}_1, ..., \mathbf{e}_n]\, P^1 \\ &\vdots \\ \mathbf{v}_n &= [\mathbf{e}_1, ..., \mathbf{e}_n]\, P^n \end{aligned}$$

Thus, the linear transformation T acting on B has another expression:

$$\begin{aligned} T(\mathbf{v}_1) &= T(p_{11}\mathbf{e}_1 + ... + p_{n1}\mathbf{e}_n) = p_{11}T(\mathbf{e}_1) + ... + p_{n1}T(\mathbf{e}_n) = [T(\mathbf{e}_1), ..., T(\mathbf{e}_n))]\, P^1 \\ &\vdots \\ T(\mathbf{v}_n) &= T(p_{1n}\mathbf{e}_1 + ... + p_{nn}\mathbf{e}_n) = p_{1n}T(\mathbf{e}_1) + ... + p_{nn}T(\mathbf{e}_n) = [T(\mathbf{e}_1), ..., T(\mathbf{e}_n))]\, P^n \end{aligned}$$

Then we complete the proof:

$$P[T]_B^i = [\mathbf{v}_1, ..., \mathbf{v}_n]\, [T]_B^i = [T(\mathbf{e}_1), ..., T(\mathbf{e}_n)]\, P^i = [T]_E P^i \quad \text{for all } i = 1, ..., n$$

Change-of-Basis Matrix

Definition (Change-of-Basis Matrix)

Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ (old basis) and $B' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\}$ (new basis) for V. The *change-of-basis matrix* $P = [I]_B^{B'}$ (from B to B') has columns $[\mathbf{v}'_j]_B$:

$$P = \begin{bmatrix} | & | & \cdots & | \\ [\mathbf{v}'_1]_B & [\mathbf{v}'_2]_B & \cdots & [\mathbf{v}'_n]_B \\ | & | & \cdots & | \end{bmatrix}$$

Proposition (Coordinate Transformation)

For any $\mathbf{v} \in V$:

$$[\mathbf{v}]_{B'} = P^{-1}[\mathbf{v}]_B$$

(P is invertible because B' is a basis.)

Corollary (Standard Basis Special Case)

If $B = E$ (standard basis), then $P = \begin{bmatrix} \mathbf{v}'_1 & \mathbf{v}'_2 & \dots & \mathbf{v}'_n \end{bmatrix}$ (columns are B' vectors). For $\mathbf{v} \in \mathbb{R}^n$:

$$[\mathbf{v}]_{B'} = P^{-1}\mathbf{v}$$

Change-of-Basis Formula for T

Theorem (General Change-of-Basis Formula)

Let $T : V \to V$ be linear, B, B' ordered bases for V, and $P = [I]_B^{B'}$ (change-of-basis from B to B'). Then:

$$[T]_{B'} = P^{-1}[T]_B P$$

Proof.

For any $\mathbf{v} \in V$:

$$[T(\mathbf{v})]_{B'} = P^{-1}[T(\mathbf{v})]_B = P^{-1}[T]_B[\mathbf{v}]_B = P^{-1}[T]_B P[\mathbf{v}]_{B'}$$

Since this holds for all $\mathbf{v}$, we have $[T]_{B'} = P^{-1}[T]_B P$. ☐

Definition (Similar Matrices)

Two $n \times n$ matrices A and B are *similar* if there exists an invertible P such that $B = P^{-1}AP$.

Corollary

Matrices representing the same linear transformation (w.r.t. different bases) are similar.

Example (Change-of-basis for rotation)

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation by 90° (standard matrix $[T]_E = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$). Let $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ where $\mathbf{v}_1 = (1, 1)$, $\mathbf{v}_2 = (1, -1)$. Find $[T]_B$.

Step 1: Construct change-of-basis matrix P (from E to B)

$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

Step 2: Compute P^{-1}

$$P^{-1} = \frac{1}{\det P} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{-2} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

Step 3: Apply change-of-basis formula

$$[T]_B = P^{-1}[T]_E P = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Diagonalization: Motivating Application

Definition (Diagonalizable Matrix)

A matrix A is *diagonalizable* if it is similar to a diagonal matrix D:

$$D = P^{-1}AP \quad \text{for some invertible } P$$

Proposition

$T : V \to V$ is *diagonalizable* (has a diagonal matrix representation) if and only if V has a basis of *eigenvectors* of T (vectors $\mathbf{v} \neq \mathbf{0}$ with $T(\mathbf{v}) = \lambda\mathbf{v}$ for some scalar λ).

Why Diagonalization Matters

- Simplifies matrix powers: $D^k = P^{-1}A^kP \implies A^k = PD^kP^{-1}$ - Reveals geometric behavior of T (eigenvectors are invariant under T) - Key for applications (differential equations, data analysis, etc.)

Abstract Vector Spaces

Definition (Abstract Vector Space)

A set V over $\mathbb{R}$ is a *vector space* if it has two operations: 1. Addition: $V \times V \to V$ (commutativity, associativity, identity, inverses) 2. Scalar multiplication: $\mathbb{R} \times V \to V$ (distributivity, compatibility, identity)

Examples

- P_k: Polynomials of degree $\leq k$ (e.g., $P_2 = \{a + bx + cx^2 \mid a, b, c \in \mathbb{R}\}$) - $M_{m\times n}$: $m \times n$ real matrices - $C^0(I)$: Continuous functions on interval I - $C^\infty(I)$: Infinitely differentiable functions on I

Definition (Subspace)

A subset $W \subset V$ is a *subspace* if it is closed under addition and scalar multiplication.

Linear Transformations on Abstract Vector Spaces

Definition (Linear Transformation $T : V \to W$)

Let V, W be vector spaces over $\mathbb{R}$. A function $T : V \to W$ is *linear* if: 1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ 2. $T(c\mathbf{v}) = cT(\mathbf{v})$ for all $\mathbf{v} \in V$ and $c \in \mathbb{R}$

Examples

1. $D : P_k \to P_{k-1}$, differentiation: $D(a_0 + a_1x + \cdots + a_kx^k) = a_1 + 2a_2x + \cdots + ka_kx^{k-1}$ 2. $S : P_k \to P_{k+1}$, integration: $S(p) = \int_0^x p(t)dt$ 3. $T : M_{2\times 2} \to \mathbb{R}$, trace: $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d$

Proposition

T is linear $\iff$ $T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$ for all $c_1, c_2 \in \mathbb{R}$ and $\mathbf{v}_1, \mathbf{v}_2 \in V$.

Kernel and Image of a Linear Transformation

Definition (Kernel (Nullspace) of T)

$$\ker(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}_W\} \quad (\mathbf{0}_W \text{ is the zero vector of } W)$$

Definition (Image (Range) of T)

$$\text{image}(T) = \{\mathbf{w} \in W \mid \exists \mathbf{v} \in V \text{ such that } T(\mathbf{v}) = \mathbf{w}\}$$

Proposition

$\ker(T)$ is a subspace of V, and $\text{image}(T)$ is a subspace of W.

Examples

1. $D : P_k \to P_{k-1}$: $\ker(D) = \{a_0 \mid a_0 \in \mathbb{R}\}$ (constant polynomials), $\text{image}(D) = P_{k-1}$ 2. $T : M_{2\times 2} \to \mathbb{R}$ (trace): $\ker(T) = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \right\}$, $\text{image}(T) = \mathbb{R}$

Rank-Nullity Theorem

Theorem (Rank-Nullity Theorem)

Let V be finite-dimensional, and $T : V \to W$ linear. Then:

$$\dim \ker(T) + \dim \operatorname{image}(T) = \dim V$$

where:

- $\dim \ker(T)$ = *nullity* of T
- $\dim \operatorname{image}(T)$ = *rank* of T

Sketch.

Let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a basis for $\ker(T)$, extend to a basis $B' = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ of V. Show $C = \{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\}$ is a basis for image(T). Then:

$$\dim \ker(T) = k, \quad \dim \operatorname{image}(T) = n - k \implies k + (n - k) = n = \dim V$$

□

Corollary

If $\dim V = \dim W$ (finite-dimensional), then T is bijective $\iff$ $\ker(T) = \{0\}$ $\iff$ image$(T) = W$.

Example: Rank-Nullity Theorem

Example

Let $T : P_3 \to P_2$ be differentiation: $T(p) = p'$. Verify Rank-Nullity.

Step 1: Compute $\dim V$

$P_3 = \{a + bx + cx^2 + dx^3\}$ has basis $\{1, x, x^2, x^3\}$, so $\dim V = 4$.

Step 2: Compute $\dim \ker(T)$

$\ker(T) = \{p \in P_3 \mid p' = 0\} = \{a \mid a \in \mathbb{R}\}$ (constants). Basis $\{1\}$, so $\dim \ker(T) = 1$.

Step 3: Compute $\dim \mathrm{image}(T)$

$\mathrm{image}(T) = \{p' \mid p \in P_3\} = \{b + 2cx + 3dx^2\} = P_2$. Basis $\{1, x, x^2\}$, so $\dim \mathrm{image}(T) = 3$.

Step 4: Verify Rank-Nullity

$$\dim \ker(T) + \dim \mathrm{image}(T) = 1 + 3 = 4 = \dim V$$

Matrix of $T : V \to W$ (Abstract Spaces)

Definition (Matrix of T w.r.t. Bases B and C)

Let V (dim n) have basis $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, W (dim m) have basis $C = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$, and $T : V \to W$ linear. For each $\mathbf{v}_j$, write:

$$T(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \cdots + a_{mj}\mathbf{w}_m$$

The *matrix of T with respect to B and C* is:

$$[T]_{B,C} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{mj} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Proposition

For any $\mathbf{v} \in V$:

$$[T(\mathbf{v})]_C = [T]_{B,C}[\mathbf{v}]_B$$

Corollary

If $V = W$ and $B = C$, this reduces to the earlier definition $[T(\mathbf{v})]_B = [T]_B[\mathbf{v}]_B$.

Example: Matrix of Differentiation $D : P_2 \to P_1$

Example

Let $D(p) = p'$, $V = P_2$ (basis $B = \{1, x, x^2\}$), $W = P_1$ (basis $C = \{1, x\}$). Find $[D]_{B,C}$.

Step 1: Compute $T(v_j)$ for B

- $D(1) = 0 = 0 \cdot 1 + 0 \cdot x$ - $D(x) = 1 = 1 \cdot 1 + 0 \cdot x$ - $D(x^2) = 2x = 0 \cdot 1 + 2 \cdot x$

Step 2: Find $[T(v_j)]_C$

$$[D(1)]_C = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad [D(x)]_C = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad [D(x^2)]_C = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Step 3: Construct $[D]_{B,C}$

$$[D]_{B,C} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Verify: For $p = 2 + 3x + 4x^2$, $[p]_B = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$

$$[D(p)]_C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix} \implies D(p) = 3 + 8x \quad \text{(correct!)}$$

Key Concepts Summary

1. **Least Squares Solutions**: Minimize $\|A\mathbf{x} - \mathbf{b}\|$, solve $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$
2. **Orthogonal Bases**: Simplify projections ($\text{proj}_V \mathbf{b} = \sum \frac{\mathbf{b}\cdot\mathbf{v}_i}{\|\mathbf{v}_i\|^2}\mathbf{v}_i$)
3. **Matrix of T w.r.t. a Basis**: $[T]_B$ encodes T via coordinate vectors
4. **Change-of-Basis Formula**: $[T]_{B'} = P^{-1}[T]_B P$ (similar matrices)
5. **Abstract Linear Transformations**: Extend to $P_k, M_{m\times n}, C^0(I)$, etc.
6. **Rank-Nullity Theorem**: $\dim\ker(T) + \dim\text{image}(T) = \dim V$

Unifying Theme: Linear transformations are best understood via their matrix representations, which simplify with well-chosen bases (e.g., orthogonal, eigenbases).