Abstract Linear Algebra M416 notes

Zhao Yang

Department of Mathematics
University of Illinois at Urbana-Champaign

This note will collect key ingredients taught in class. It cannot be used as a substitute of your textbook.

December 2, 2019

1.2 Vector spaces

Definition (Vector space (1 set+2 operations+8 axioms))

A vector space over a field F is a set V together with two operations that satisfy the eight axioms below.

Two operation:

- Vector addition "+": $V \times V \to V$, $(x, y) \to x + y$, $x, y \in V$
- Scalar multiplication "$\cdot$": $F \times V \to V$, $(a, x) \to a \cdot x$, $a \in F$, $x \in V$.

Eight axioms:

VS1 commutativity of addition: $x + y = y + x$, $\forall x, y \in V$;

VS2 associativity of addition: $(x + y) + z = x + (y + z)$, $\forall x, y, z \in V$;

VS3 identity element of addition: $\exists 0 \in V$ $s.t.$ $x + 0 = x$, $\forall x \in V$;

VS4 inverse element of addition: $\forall x \in V$, $\exists y \in V$ $s.t.$ $x + y = 0$;

VS5 identity element of scalar multiplication: $1 \cdot x = x$, where 1 denotes the multiplicative identity in F;

VS6 compatibility of scalar multiplication with field multiplication: $(a \cdot b) \cdot x = a \cdot (b \cdot x)$, $\forall a, b \in F$, $x \in V$;

VS7 distributivity of scalar multiplication w.r.t. vector addition: $a(x + y) = ax + ay$, $\forall a \in F$, $x, y \in V$;

VS8 distributivity of scalar multiplication w.r.t. field addition: $(a + b)x = ax + bx$, $\forall a, b \in F$, $x \in V$;

Remarks and examples

Remark

- no direction, no magnitude
- no ambiguity when writing $x + y + z$ without parentheses. more generally linear combination $a_1x_1 + a_2x_2 + \cdots + a_nx_n$
- To show a structure is a vector space
  1. describe the set V;
  2. define two operation: vector addition and scalar multiplication;
  3. verify the two operations satisfy the 8 axioms.

Examples

F^n, $M_{m\times n}$, $\mathcal{F}(S, F)$, $C[a, b]$, $P(F)$, $P_n(F)$, sequence, various functional spaces, etc.

Nonexample

polynomials of degree n; first quadrant; [FIS] page 15 ex 17;

Theorems and Corollary

Theorem (1.1 Cancellation law)

$\forall x, y, z \in V$, *if* $x + z = y + z$ *then* $x = y$. *In other words,* z *gets cancelled.*

Corollary (uniqueness of zero vector and inverse)

0 *vector is unique; the inverse of a vector is unique (homework).*

Theorem

The following statements are true.

(a) $0 \cdot x = 0$, $\forall x \in V$;

(b) $(-a) \cdot x = -(a \cdot x) = a(-x)$;

(c) $a \cdot 0 = 0$, $\forall a \in F$.

1.3 Subspaces

Definition

A subset W of a vector space V over a field F is called a subspace of V if W is a vector space over F with the operations of addition and scalar multiplication defined in V.

Theorem (1.3)

$W \subset V$ is a subspace of V if and only if

(a) $0 \in W$;

(b) *closed under vector addition:* $x + y \in W, \forall x, y \in W$;

(c) *closed under scalar multiplication:* $c \cdot x \in W, \forall c \in F, x \in W$

Examples and nonexamples

operations on a matrix: (1) transpose "t" (2) trace of square matrix.

Examples

1. symmetric matrices $\{A \in M_{n\times n} : A^t = A\}$;
2. skew-symmetric matrices $\{A \in M_{n\times n} : A^t = -A\}$;
3. null spaces of linear maps $\{(x, y, z) : x + y + 2z = 0\}$;
4. polynomials with degree less than or equal to n: $P_n(F) = \{p \in P(F) : deg(p) \leq n\}$;
5. $C[a, b] \subset \mathcal{F}([a, b], R)$;
6. diagonal matrices $\{A \in M_{n\times n} : A_{ij} = 0, i \neq j\}$;
7. trace 0 matrices $\{A \in M_{n\times n} : trace(A) = 0\}$.

...

Nonexample

1. $\{(x, y, z) : x + y + 2z = 1\}$;

Theorem

Theorem (1.4)

Let W_1, W_2 be two subspaces of V, then $W_1 \cap W_2$ is a subspace of W_1, W_2, and V.

Question: is the union set $W_1 \cup W_2$ necessarily a subspace of V?
Answer: [FIS page 21 ex 19].

1.4 Linear combinations and systems of linear equations

Definition (linear combination)

Let V be a vector space over field F, for $v_1, v_2, \cdots, v_n \in V$ and $a_1, a_2, \cdots, a_n \in F$, the vector $a_1 \cdot v_1 + a_2 \cdot v_2 + \cdots + a_n \cdot v_n$ is called a linear combination of vectors $v_1, \cdots, v_n$ with **coefficients** $a_1, \cdots, a_n$.

Definition (span(S))

Let S be a nonempty subset of a vector space V. The span of S is the set consisting all linear combinations of the vectors in S. Namely

$$span(S) = \{a_1 \cdot v_1 + a_2 \cdot v_2 + \cdots + a_n \cdot v_n : a_i \in F, v_i \in S, i = 1, \cdots, n\}$$

Theorem (1.5)

For vector spave V and $\emptyset \neq S \subset V$, $span(S)$ is a subspace of V. Moreover, for all subspace W of V, if $S \subset W$ then $span(S) \subset W$.

linear combinations v.s. linear systems

Write linear system in linear combination of vectors form a 2×2 example

$$\begin{cases} 10x + 5y = 6 \\ 5x + 2y = 2 \end{cases} \to \text{linear combination?}$$

linear combinations v.s. linear systems

Write linear system in linear combination of vectors form a 2×2 example

$$\begin{cases} 10x + 5y = 6 \\ 5x + 2y = 2 \end{cases} \rightarrow \text{linear combination?}$$

more generally, a system of m equations with n unknowns x_i and $(n+1) \times m$ known quantities a_{ij}, b_i.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = b_m \end{cases} \rightarrow \text{linear combination?}$$

continued, $span(S)$

Definition (generating set)

We say a subset S of a vector space V generates V if $span(S) = V$.

Find the generating sets of following subspaces

Examples

1. symmetric matrices $\{A \in M_{n\times n} : A^t = A\}$;
2. skew-symmetric matrices $\{A \in M_{n\times n} : A^t = -A\}$;
3. null spaces of linear maps $\{(x, y, z) : x + y + 2z = 0\}$;
4. polynomials with degree less than or equal to n: $P_n(F) = \{p \in P(F) : deg(p) \leq n\}$;
5. diagonal matrices $\{A \in M_{n\times n} : A_{ij} = 0, i \neq j\}$;
6. trace 0 matrices $\{A \in M_{n\times n} : trace(A) = 0\}$.

Solving a linear system

Reference: "http://linear.ups.edu/html/section-RREF.html"
A system of m equations with n unknowns x_i and $(n+1)\times m$ known quantities a_{ij}, b_i

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a_{mn}x_n = b_m \end{cases} \tag{1}$$

can be simply put as $Ax = b$ where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

Definition (augmented matrix)

The matrix $[A|b]$ is called the **augmented matrix** of the system of equations (1).

Row operations

Definition (Row operations)

The following three operations will transform an $m \times n$ matrix into a different matrix of the same size, and each is known as a row operation.

1. $R_i \leftrightarrow R_j$: Swap the location of rows i and j.
2. αR_i: Multiply row i by the nonzero scalar α.
3. $\alpha R_i + R_j$: Multiply row i by the scalar α and add to row j.

Definition (Row-Equivalent Matrices)

Two matrices, A and B, are row-equivalent if one can be obtained from the other by a sequence of row operations.

Theorem (Row-Equivalent Matrices represent Equivalent Systems)

Suppose that A and B are row-equivalent augmented matrices. Then the systems of linear equations that they represent are equivalent systems.

Echelon Form

Definition (Reduced Row-Echelon Form)

A matrix is in reduced row-echelon form if it meets all of the following conditions:

1. If there is a row where every entry is zero, then this row (called a **zero row**) lies below any other row that contains a nonzero entry;
2. The leftmost nonzero entry of a row is equal to 1. This 1 is called a **leading 1**. A column containing a leading 1 will be called a **pivot column**.
3. The leftmost nonzero entry of a row is the only nonzero entry in its column.
4. Consider any two different leftmost nonzero entries, one located in row i, column j and the other located in row s, column t. If $s > i$, then $t > j$. **In other words, leading 1s move to the right as you move down rows.**

The number of nonzero rows will be denoted by r (rank), which is also equal to the number of leading 1's and the number of pivot columns.
The set of column indices for the pivot columns will be denoted by $D = \{d_1, d_2, \cdots, d_r\}$, while the columns that are not pivot columns will be denoted as $F = \{f_1, f_2, \cdots, f_{n-r}\}$.

Gauss-Jordan Elimination

Gauss-Jordan Elimination Let M be an $m \times n$ matrix. Denote rows by R_i, $i = 1, \cdots, m$ and columns by C_j, $j = 1, \cdots, n$.

0. Set $r = 0$, $j = 0$;
1. If $j \geq n$ stop and return current matrix. Otherwise, set $j = j + 1$;
2. If C_j is 0 below row r, go to step 1;
3. Set $r = r + 1$. Arrange the j^{th} entry of R_r to be nonzero by swapping R_r with one $R_{r+1}, \cdots, R_m$ if necessary;
4. Scale R_r so that j^{th} entry is 1;
5. For i in $[1, \cdots, m]$, $i \neq r$ clear the j^{th} entry of R_i by setting $R_i = -cR_r + R_i$ where c is the j^{th} entry of R_i;
6. go to step 1.

Theorem

The returned matrix from Gauss-Jordan Elimination procedure is in reduced row-echelon form.

Gauss-Jordan Elimination, continued

Proof.

we will prove following claim by induction.
Claim: When we return to step 1, the $r \times j$ "upper matrix" is in row reduced echelon form and the $(m-r) \times j$ "lower matrix" below the "upper matrix" is a zero matrix.
Base step: The first time we return to step 1, the first column must be $\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^t$. Being [1] the "upper matrix" is certainly in row reduced echelon form. The "lower matrix" is zero matrix also follows.
Inductive step: (a) if we return to step 1 via step 2, then in the last round the upper matrix is augmented by an arbitrary column, hence in row reduced echelon form. And, the lower zero matrix is augmented by a zero column, hence still being zero. (b) if we return to step 1 via step 6, then in last round we have 1 more leading 1 in the right-bottom corner of the upper matrix and the new upper matrix is still in row reduced echelon form. For lower zero matrix, its number of rows decreases by 1 and its number of columns increases by 1. $\square$

Theorem

Theorem

Row-Equivalent Matrix in Echelon Form Suppose A is a matrix. Then there is a matrix B so that A and B are row-equivalent. B is in reduced row-echelon form.

Proof.

Perform Gauss-Jordan Elimination procedure on matrix A and call the returned matrix as matrix B. B is in reduced row-echelon form. $\square$

Examples

Example

1.
$$\begin{cases} -7x_1 - 6x_2 - 12x_3 = -33 \\ 5x_1 + 5x_2 + 7x_3 = 24 \\ x_1 + 4x_3 = 5 \end{cases}$$

2.
$$\begin{cases} x_1 - x_2 + 2x_3 = 1 \\ 2x_1 + x_2 + x_3 = 8 \\ x_1 + x_2 = 5 \end{cases}$$

3.
$$\begin{cases} 2x_1 + x_2 + 7x_3 - 7X_4 = 2 \\ -3x_1 + 4x_2 - 5x_3 - 6X_4 = 3 \\ x_1 + x_2 + 4x_3 - 5x_4 = 2 \end{cases}$$

Types of solution sets

Reference: "http://linear.ups.edu/html/section-TSS.html"

Definition (consistent system)

A system of linear equations is consistent if it has at least one solution. Otherwise, the system is called inconsistent.

Definition

Suppose A is the augmented matrix of a consistent system of linear equations and B is a row-equivalent matrix in reduced row-echelon form. Suppose j is the index of a pivot column of B. Then the variable x_j is **dependent**. A variable that is not dependent is called **independent** or **free**.

Theorem (inconsistency)

Suppose A is the augmented matrix of a system of linear equations with n variables. Suppose also that B is a row-equivalent matrix in reduced row-echelon form with r nonzero rows. Then the system of equations is inconsistent if and only if column $n+1$ of B is pivot column.

Types of solution sets, continued,

Theorem (consistent systems)

Suppose A is the augmented matrix of a consistent system of linear equations with n variables. Suppose also that B is a row-equivalent matrix in reduced row-echelon form with r pivot columns. Then $r \leq n$.

i. *If $r = n$, then the system has a unique solution, and if $r < n$;*
ii. *If $r < n$, then the system has infinitely many solutions. And the solution set can be described with $n - r$ free variable.*

Corollary

Suppose a consistent system of linear equations has m equations in n variables. If $n > m$, then the system has infinitely many solutions.

1.5 Linear dependence and linear independence

Definition (independence and dependence)

Let $v_1, \cdots, v_n$ be n vectors in a vector space V over F. $v_1, \cdots, v_n$ are linear independent if the equation $a_1 \cdot v_1 + \cdots a_n v_n = 0$ has only trivial solution $a_1 = \cdots = a_n = 0$ and linear dependent if the equation $a_1 \cdot v_1 + \cdots a_n v_n = 0$ has nontrivial solutions.

Theorem

Let $v_1, \cdots, v_n$ be vectors in vector space V over F, then $v_1, \cdots, v_n$ are linear dependent if and only if at least one vector from $v_1, \cdots, v_n$ can be represented as a linear combination of the other vectors.

Theorem

Let $v_1, \cdots, v_n$, and w be vectors in a vector space V over F. Assume w can be represented as a linear combination of $v_1, \cdots, v_n$, then the representation is unique if and only if $v_1, \cdots, v_n$ are linear independent.

More theorems

Definition (independent and dependent subsets)

Let V be a vector space over F. A subset S of V will be called a linear independent set if any finite number of distinct vectors in S are linear independent. A subset S will be called a linear dependent set if there exists a finite number of distinct vectors in S that are linear dependent.

Theorem (1.6)

Let V be a vector space and $S_1 \subset S_2 \subset V$, then

i. *S_1 is linear dependent $\Rightarrow$ S_2 is linear dependent.*

ii. *S_2 is linear independent $\Rightarrow$ S_1 is linear independent.*

Theorem (1.7)

Let S be a linearly independent subset of a vector space V, and let $v \in V, v \notin S$. Then $S \cup \{v\}$ is linearly dependent if and only if $v \in span(S)$.

1.6 Bases and dimension

Definition (Basis)

A basis β for a vector space V is a linear independent subset of V that generates V. If β is a basis for V, we also say that the vectors of β form a basis for V.

Remark

Check (1) If β is a linear independent set. (2) If $span(\beta) = V$, or simply $span(\beta) \supset V$.

Examples

1. $\emptyset$ is a basis for the zero vector space;
2. the standard basis $\{e_1, \cdots, e_n\}$ for F^n;
3. the standard basis $\{E^{ij} : i = 1, \cdots, m,\ j = 1, \cdots, n\}$ for $M_{m\times n}(F)$;
4. the standard basis $\{1, x, \cdots, x^n\}$ for $P_n(F)$;
5. the standard basis $\{1, x, x^2, \cdots\}$ for $P(F)$.

Theorems

Theorem (1.8)

Let V be a vector space and $\beta = \{u_1, u_2, \cdots, u_n\}$ be a subset of V. Then β is a basis for V if and only if each $v \in V$ can be uniquely expressed as a linear combination of vectors of β, that is, can be expressed in the form $v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$ for unique scalars $a_1, a_2, \cdots, a_n$.

Theorem (1.9)

If a vector space V is generated by a finite set S, then some subset of S is a basis for V. Hence V has a finite basis.

Theorem

Every vector space has a basis. (proof relies on Zorn's lemma.)

Theorem (Replacement theorem)

Let V be a vector space that is generated by a set G containing exactly n vectors, and let L be a linearly independent subset of V containing exactly m vectors, Then $m \leq n$ and there exists a subset of H of G containing exactly $n - m$ vectors such that $L \cup H$ generates V.

Dimension

Corollary

Suppose G is a finite generating set for V and $U \subset V$ is linear independent, then U is finite and the number of elements in U is less than or equal to the number of elements in G.

Corollary (1)

Let V be a vector space having a finite basis. Then every basis for V contains the same number of vectors.

Definition (Dimension)

A vector space is called finite-dimensional if it has a basis consisting of a finite number of vectors. The unique number of vectors in each basis for V is called the dimension of V and is denoted by $dim(V)$. A vector space that is not finite-dimensional is called infinite-dimensional.

Dimension for subspaces

Corollary (2)

Let V be a vector space with dimension n.

(a) *Any finite generating set for V contains at least n vectors, and a generating set for V that contains exactly n vectors is a basis for V.*

(b) *Any linearly independent subsets of V that contains exactly n vectors is a basis for V.*

(c) *Every linearly independent subset of V can be extended to a basis for V.*

Theorem (1.11)

Let W be a subspace of a finite-dimensional vector space V. Then W is finite-dimensional and $dim(W) \leq dim(V)$. Moreover, if $dim(W) = dim(V)$, then $V = W$.

Dimension for subspaces, continued

Example

For $A \in M_{m\times n}(R)$,

(i) show the solution space of a homogeneous linear system $\{x : Ax = 0\}$ $\big(Null(A)\big)$ is a subspace of R^n. What is the dimension of the solution space?

(ii) Find a basis for the row space of A $\big(\mathcal{R}(A)\big)$ by using Gauss-Jordan elimination.

Theorem

If W_1 and W_2 are finite-dimensional subspaces of a vector space V, then the subspace $W_1 + W_2$ is finite-dimensional and $dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2)$. Hence the sum is a **direct sum** *if and only if $dim(W_1 + W_2) = dim(W_1) + dim(W_2)$.*

Example

$A = (A + A^t)/2 + (A - A^t)/2$, spaces of symmetric matrices and skew-symmetric matrices. Dimension relation?

2.1 Linear transformations, null spaces, and ranges

Definition (Linear transformation)

Let V and W be vector spaces (over F). We call a function $T : V \to W$ a linear transformation from V to W if, for all $x, y \in V$ and $c \in F$, we have

(a) $T(x + y) = T(x) + T(y)$ and

(b) $T(cx) = cT(x)$.

Definition (Null space and range)

Suppose $T : V \to W$ is a linear transformation.

i. The null space (or kernel) of T is $\mathcal{N}(T) := \{v \in V : T(v) = 0_w\}$.

ii. The range (or image) of T is $\mathcal{R}(T) := \{T(v) : v \in V\}$.

Theorem (2.1)

The kernel of T is a subspace of V and the image of T is a subspace of W.

Examples

Examples

1. $T : R^2 \to R^3,\ T([x, y]) = [x + y, x - y, 2x]$
2. Projection $P : R^3 \to R^3,\ P([x, y, z]) = [x, y, 0]$
3. Rotation $S : R^2 \to R^2,\ S([x, y]) = [\cos(\theta)x - \sin(\theta)y, \sin(x) + \cos(\theta)y]$
4. Matrix transpose $T : M_{m\times n} \to M_{n\times m},\ T(A) = A^t$
5. Differentiation $D : P(R) \to P(R)\ D(p(x)) := \frac{dp(x)}{dx}$ or $D(x^n) = nx^{n-1}$
6. Integration $I : P(R) \to P(R)\ I(p(x)) := \int_0^x p(y)dy$ or $I(x^n) = \frac{1}{n+1}x^{n+1}$
7. many more (Fourier transform, Laplace transform)

Remark

Examples 1-3 are special cases of $T : R^n \to R^m,\ T(v) = Av$ for some $A \in M_{m\times n}$.

Nullity, Rank

Theorem (2.2)

Let V and W be vector spaces and T be a linear transformation from V to W. If $\beta = \{v_1, v_2, \cdots, v_n\}$ is a basis for V, then

$$\mathcal{R}(T) = span(\{T(v_1), T(v_2), \cdots, T(v_n)\}).$$

In other words, $T(\beta)$ is a generating set for the range of T.

Definition (Nullity and rank)

Let V and W be vector spaces and T be a linear transformation from V to W. If $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are finite-dimensional, then we define

$$nullity(T) := dim(\mathcal{N}(T)), \quad rank(T) := dim(\mathcal{R}(T)).$$

Remark

If V is finite-dimensional, then both $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are finite-dimensional. Moreover, we have following theorem relating $dim(V)$, $nullity(T)$, and $rank(T)$.

The dimension theorem

Theorem (2.3 Dimension theorem)

Let V and W be vector spaces and T be a linear transformation from V to W. If V is finite-dimensional, then

$$nullity(T) + rank(T) = dim(V).$$

Theorem (2.4)

Let $T : V \to W$ be a linear transformation. Then T is one to one $\Leftrightarrow \mathcal{N}(T) = \{0\}$.

Theorem (2.5)

Let V and W be vector space of equal finite dimension and T be a linear transformation from V to W. Then the following are equivalent.

(a) T *is one to one;*

(b) T *is onto;*

(c) $rank(T) = dim(V)$.

Defining a linear transformation

Theorem (2.6)

Let V and W be vector spaces and V be finite-dimensional with a basis $\{v_1, \cdots v_n\}$. For $\{u_1, \cdots, u_n\}$ in W, there **exists** *a* **unique** *linear transformation $T : V \to W$ such that $T(v_i) = u_i$ for $i = 1, \cdots, n$.*

Corollary

Let V and W be vector spaces and V be finite-dimensional with a basis $\{v_1, \cdots v_n\}$. If $T, U : V \to W$ are linear and $T(v_i) = U(v_i)$ for $i = 1, \cdots, n$, then $T = U$.

Remark

A linear transformation is completely determined by its action on a basis.

2.2 The matrix representation of a linear transformation

Definition (Ordered basis)

Let V be a finite-dimensional vector space. An ordered basis for V is a basis for V endowed with a specific order.

Examples (Standard ordered basis)

$\{e_1, e_2, \cdots, e_n\}$ for F^n and $\{1, x, \cdots, x^{n-1}\}$ for $P_{n-1}(F)$.

Definition (coordinate)

Let $\beta = \{v_1, v_2, \cdots, v_n\}$ be an ordered basis for a finite dimensional vector space V over F. For $x \in V$, let $c_1, c_2, \cdots, c_n \in F$ be the unique scalars such that $x = \sum_{i=1}^{n} c_i v_i$. We define the coordinate vector of x relative to β, denoted $[x]_\beta$, by

$$[x]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \in F^n.$$

Remark

(i) With this definition, we have $x = \beta\,[x]_\beta$. The map $V \to F^n : x \to [x]_\beta$ is a linear transformation, which is one to one and onto (hence an isomorphism);

(ii) Coordinate vectors relative to different order bases are different.

Ideas:

(i) Identify n-dimensional abstract vector space V over F with F^n;

(ii) Identify a linear transformation T from an n-dimensional abstract vector space V (β) to an m-dimensional abstract vector space W (γ) with an $m \times n$ matrix.

In particular, we have following diagram

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\cdot]_\beta}"] & W \arrow[d, "{[\cdot]_\gamma}"] \\ F^n \arrow[r, "{[T]_\beta^\gamma}"] & F^m \end{tikzcd} \qquad \begin{tikzcd} x \arrow[r, "T"] \arrow[d, "{[\cdot]_\beta}"] & T(x) \arrow[d, "{[\cdot]_\gamma}"] \\ {[x]_\beta} \arrow[r, "{[T]_\beta^\gamma}"] & {[T(x)]_\gamma = [T]_\beta^\gamma [x]_\beta} \end{tikzcd}$$

Proof of Thm 2.14

Assume V is an n-dimension vector space with an ordered basis $\beta = \{v_1, v_2, \cdots, v_n\}$ and W is an m-dimension vector space with an ordered basis $\gamma = \{w_1, w_2, \cdots, w_m\}$. T is a linear transformation from V to W with

$$T(v_i) = \sum_{j=1}^{m} a_{ji} w_j = \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{mi} \end{bmatrix} = \gamma \, [T(v_i)]_\gamma .$$

Now, for $x \in V$, $x = \sum_{i=1}^{n} c_i v_i$, direct computation yields

$$\begin{aligned} T(x) = \sum_{i=1}^{n} c_i T(v_i) &= \begin{bmatrix} T(v_1) & T(v_2) & \cdots & T(v_n) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \\ &= \begin{bmatrix} w_1 & w_2 & \cdots & w_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \\ &= \gamma \, A \, [x]_\beta = \gamma \, [T]_\beta^\gamma \, [x]_\beta . \end{aligned}$$

Moreover, we have $[T(x)]_\gamma = [T]_\beta^\gamma \, [x]_\beta$.

Matrix representation

Definition (Matrix representation)

We call the $m \times n$ matrix A the matrix representation of T in the ordered bases β and γ and write $A = [T]_\beta^\gamma$. If $V = W$ and $\beta = \gamma$, then we write $A = [T]_\beta$.

Let V and W be vector spaces over F. $\mathcal{F}(V, W)$ denotes set of **all functions** from V to W. For $T, U \in \mathcal{F}(V, W)$ and $a \in F$, define $T + U : V \to W$ and $aT : V \to W$ in the usual way. That is

$$(T + U)(x) = T(x) + U(x), \quad (aT)(x) = aT(x).$$

By previous results, $\mathcal{F}(V, W)$ is a vector space.

Definition (set of all linear transformations)

We denote the set of all linear transformations from V to W by $\mathcal{L}(V, W)$, that is

$$\mathcal{L}(V, W) := \{T \in \mathcal{F}(V, W) : T \text{ is linear}\}.$$

When $W = V$, we denote $\mathcal{L}(V, V)$ simply by $\mathcal{L}(V)$.

Remark

(i) $\mathcal{L}(V, W)$ is a subspace $\mathcal{F}(V, W)$ established in Thm 2.7 (omitted);

(ii) Will show a complete identification of $\mathcal{L}(V, W)$ with $M_{m\times n}(F)$ in Section 2.4 by using the map $[]_\beta^\gamma$.

Matrix representation, continued

Example (FIS P84 4 modified)

Define $T : M_{2\times 2}(R) \to P_2(R)$ by $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b) + (2d)x + bx^2$. Let $\beta = \{E_{11}+E_{12}, E_{12}-E_{22}, E_{21}, E_{22}\}$ and $\gamma = \{1, 2x, x^2\}$ be ordered bases of $M_{2\times 2}(R)$ and $P_2(R)$ respectively. Compute $[T]_\beta^\gamma$.

Theorem (2.8 $[\cdot]_\beta^\gamma$ as a linear transformation)

Let V and W be finite-dimensional vector spaces over F with ordered bases β and γ. Consider the map $\mathcal{L}(V,W) \to M_{m\times n}(F)$ that maps T to $[T]_\beta^\gamma$. The map is linear that is it satisfies for $T, U \in \mathcal{L}(V,W)$

(a) $[T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$;

(b) $[aT]_\beta^\gamma = a[T]_\beta^\gamma$.

Remark

Indeed the map is an isomorphism and hence $\mathcal{L}(V,W)$ and $M_{m\times n}(F)$ are isomorphic.

2.3 Composition of linear transformations and matrix multiplication

Theorem (2.9)

Let V, W, and Z be vector spaces over the same field F, and let $T : V \to W$ and $U : W \to Z$ be linear. Then $UT : V \to Z$ is linear.

Theorem (2.10)

Let V be a vector space. Let $T, U_1, U_2 \in \mathcal{L}(V)$. Then

(a) $T(U_1 + U_2) = TU_1 + TU_2$ *and* $(U_1 + U_2)T = U_1T + U_2T$;

(b) $T(U_1U_2) = (TU_1)U_2$;

(c) $TI = IT = T$;

(d) $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$.

Definition (Matrix multiplication)

$A \in M_{p\times m}$, $B \in M_{m\times n}$, the product of A and B is an $p \times n$ matrix with ij-entry $(AB)_{ij} = \sum_{k=1}^{m} A_{ik}B_{kj}$.

Continued,

Theorem (2.11)

Let V, W, and Z be finite-dimensional vector spaces with ordered bases $\alpha(n)$, $\beta(m)$, and $\gamma(p)$. Let $T : V \to W$ and $U : W \to Z$ be linear transformations. Then $[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$.

Corollary

Let $T, U \in \mathcal{L}(V)$. then $[UT]_\beta = [U]_\beta [T]_\beta$.

Definition

The kronecker delta $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The $n \times n$ identity matrix I_n is defined by $(I_n)_{ij} = \delta_{ij}$.

Properties of matrix multiplication

Direct calculation shows:

Theorem (2.12 Properties of matrix multiplication)

Let A be an $m \times n$ matrix, B and C be $n \times p$ matrices, and D and E be $q \times m$ matrices. Then

(a) $A(B + C) = AB + AC$ *and* $(D + E)A = DA + EA$;

(b) $a(AB) = (aA)B = A(aB)$;

(c) $I_m A = A = A I_n$;

(d) *If V is an n-dimensional vector space with an ordered basis β, then* $[I_V]_\beta = I_n$.

Theorem (2.13 Block matrix multiplication)

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Let u_j and v_j denote the j^{th} columns of AB and B, then

(a) $u_j = Av_j$, *that is*

$$AB = A\begin{bmatrix} v_1 & v_2 & \cdots & v_p \end{bmatrix} = \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_p \end{bmatrix};$$

(b) $v_j = Be_j$.

One more property and the main theorem

Theorem (2.16)

If $(AB)C$ is defined, then $(AB)C = A(BC) = ABC$.

Theorem (2.14)

Let V and W be finite-dimensional vector spaces having ordered bases β and γ and $T : V \to W$ be linear. Then, for $x \in V$

$$[T(x)]_\gamma = [T]_\beta^\gamma [x]_\beta.$$

Proof.

Done! $\square$

Define linear transformation from a matrix

Definition

Let $A \in M_{m\times n}(F)$, we denote byL_A the mapping $F^n \to F^m$ defined by $L_A(x) = Ax$ for column vector $x \in F^n$. We call L_A a left-multiplication transformation.

Theorem (2.15)

Let $A, B \in M_{m\times n}(F)$, then $L_A(L_B)$ is linear. Let β, γ be the standard ordered bases for F^n and F^m, we have following properties:

(a) $[L_A]_\beta^\gamma = A;$

(b) $L_A = L_B \Leftrightarrow A = B;$

(c) $L_{A+B} = L_A + L_B$ *and* $L_{aA} = aL_A;$

(d) *If* $T : F^n \to F^m$ *is linear, then there exists* $C \in M_{m\times n}(F)$ *such that* $T = L_C$ *in fact* $C = [T]_\beta^\gamma;$

(e) *If* $E \in M_{n\times p}$, *then* $L_{AE} = L_A L_E;$

(f) *If* $m = n$ *then* $L_{I_n} = I_{F^n}.$

2.4 Invertibility and isomorphism

Definition (homomorphism)

In algebra, a homomorphism is a structure-preserving map between two algebraic structures of the same type. Specializing to vector spaces, a homomorphism is a linear transformation.

Definition (Invertible linear transformation, isomorphism)

Let V and W be vector space let T be a homomorphism from V to W, if T is invertible, namely if there exists function $U : W \to V$ such that $TU = I_W$ $UT = I_V$, T is then called an invertible linear transformation or simply an isomorphism. U is called the inverse of T and is denoted by T^{-1}.

Remark

Assume $T : V \to W$, $U : W \to Z$ are invertible, then

1. $(TU)^{-1} = U^{-1}T^{-1}$;
2. $(T^{-1})^{-1} = T$;
3. Let V and W be finite dimensional vector spaces of equal dimension and $T : V \to W$ be linear. T is invertible $\Leftrightarrow$ $rank(T) = dim(V)$.

Continued,

Theorem (2.17)

The inverse function of an invertible linear transformation is linear.

Definition (isomorphic vector spaces)

Let V and W be vector spaces, we say that V is isomorphic to W if there exists an isomorphism T from V to W.

Remark

By theorem 2.17 and remark 2, W is also isomorphic to V provided that V is isomorphic to W. We may simply say V and W are isomorphic. Moreover "isomorphic" is an equivalence relation.

Examples

i. Let V be an n-dimensional space over F with an ordered basis β, the coordinate map $\phi_\beta : V \to F^n$ defined by $\phi_\beta(x) = [x]_\beta$ is an isomorphism. [Thm 2.21]
ii. Let V/W be an n/m-dimensional space over F with an ordered basis β/γ and $\mathcal{L}(V, W)$ be the space of all linear transformations from V to W, the matrix representation map $\Phi(T) = [T]_\beta^\gamma$ is an isomorphism. [Thm 2.20]

Theorems

Definition (Invertible matrix)

Let $A \in M_{n\times n}(F)$. A is invertible if there exists $B \in M_{n\times n}(F)$ such that $AB = BA = I$.

Remark Let $A, B \in M_{n\times n}(F)$, if $AB = Id$, then $BA = Id$. So definition can be changed to "such that $AB = Id$".

Lemma

Let T be an isomorphism from V to W. Then V is finite-dimensional if and only if W is finite dimensional. In this case, $dim(V) = dim(W)$.

Theorem (2.18)

Let V and W be finite-dimensional vector spaces with ordered bases β and γ. Let $T : V \to W$ be linear. Then T is invertible if and only if $[T]_\beta^\gamma$ is invertible.

Furthermore, $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$.

Theorems

Theorem (2.19)

Let V and W be finite dimensional vector spaces over F. Then V is isomorphic to W if and only if $dim(V) = dim(W)$.

Corollary

All n dimensional vector space over F are isomorphic and they are isomorphic to F^n.

Examples

i $\mathcal{L}(V, W)$, $M_{m\times n}(F)$, and F^{nm} are isomorphic;

ii $\mathcal{L}(\mathcal{L}(V, W), W)$ is isomorphic to F^{nmm}.

$$
\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "{[\cdot]_\beta}"] & W \arrow[d, "{[\cdot]_\gamma}"] \\
F^n \arrow[r, "{[T]_\beta^\gamma}"] & F^m
\end{tikzcd}
\quad \Rightarrow \quad
\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, leftrightarrow, "\phi_\beta"] & W \arrow[d, leftrightarrow, "\phi_\gamma"] \\
F^n \arrow[r, "{L_{[T]_\beta^\gamma}}"] & F^m
\end{tikzcd}
$$

2.5 The change of coordinate matrix

Question: β, β' are two bases of an n dimensional space V. Already know $[x]_\beta \neq [x]_{\beta'}$, but how are they related to each other? Answer: following theorem:

Theorem (2.22)

Let β and β' be two ordered bases for a finite-dimensional vector space V, and let $Q = [I_V]_{\beta'}^{\beta}$. Then

(a) *Q is invertible;*

(b) *For any $v \in V$, $[v]_\beta = Q[v]_{\beta'}$.*

Question: β, β' are two bases of an n dimensional space V. $T : V_\beta \to V_\beta$ corresponds the matrix representation $[T]_\beta$ and $T : V_{\beta'} \to V_{\beta'}$ corresponds the matrix representation $[T]_{\beta'}$. How T_β and $T_{\beta'}$ are related to each other? Answer: following theorem:

Theorem (2.23)

Let T be a linear operator on a finite-dimensional vector space V, and β and β' be ordered bases for V. Let $Q = [I_v]_{\beta'}^{\beta}$. Then $[T]_{\beta'} = Q^{-1}[T]_\beta Q$.

Similar matrices

Definition (Similar matrices)

$A, B \in M_{n\times n}(F)$. We say B is similar to A if there exists an invertible matrix T such that $B = T^{-1}AT$.

Remark

- i The relation of matrix similirity is an equivalence relation;
- ii Matrix representations of a linear transformation $T : V \to V$ in different order bases are similar;
- iii Can we find "simpler" matrix B that is similar to matrix A? (Chapter 5 and 7)

Theorem (FIS 118 10)

If A and B are similar, then $trace(A) = trace(B)$

Proof.

For matrices M_1, M_2, we have $trace(M_1M_2) = trace(M_2M_1)$ □

Selected facts from Chapter 3

Overview: Up to now, section 3.4 (down); section 3.1-3.3 (almost down)
Section 3.1

Definition (Elementary matrix)

An $n \times n$ elementary matrix is a matrix obtained by performing an elementary operation on I_n. The elementary matrix is said to be of type 1,2, or 3 according to whether the elementary operation performed on I_n is a type 1,2 or 3 operation.

Theorem (3.2)

Elementary matrices are invertible. Inverse matrix is of same type.

Observation:
Performing a row/column operation on a matrix is equivalent to multiplying the corresponding elementary matrix on the left/right.

Continued,

Section 3.2 Finding inverse matrix via Gauss-Jordan elimination.

Example

Find inverse of

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

Continued,

Section 3.2 Finding inverse matrix via Gauss-Jordan elimination.

Example

Find inverse of

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

Augmenting A by $[\ A \mid Id\ ]$ and performing a series of row operations on $[\ A \mid Id\ ]$ or equivalently multiplying $[\ A \mid Id\ ]$ on the left by a series of elementary matrices $E_m \cdots E_2 E_1$, we have

$$E_m \cdots E_2 E_1 [\ A \mid Id\ ] = [\ E_m \cdots E_2 E_1 A \mid E_m \cdots E_2 E_1\ ] = [\ Id \mid E_m \cdots E_2 E_1\ ]$$

Linear system recap

Section 3.3 Structure of solution space of linear system $Ax = b$

Definition

The linear system $Ax = b$ is said to be homogeneous if $b = 0$ and nonhomogeneous if $b \neq 0$.

Recall that $L_A(x) = Ax$, then the solution space of the homogeneous system $Ax = 0$ is then $Null(L_A)$, which is a subspace.
Now consider nonhomogeneous system $Ax = b$ and let p be a **specific solution**. Then we have following representation of the solution space

$$\{x : Ax = b\} = p + Null(L_A).$$

Equivalence

Let $A \in M_{n\times n}(F)$, the following statement are equivalent

i. $rank(A) = n$;

ii. $null(A) = \{0\}$;

iii. A is invertible;

iv. $det(A) \neq 0$;

v. L_A is onto;

vi. L_A is one to one;

vii. L_A is invertible;

viii. Columns of A are linear independent;

ix. Rows of A are linear independent;

x. $Ax = b$ has a unique solution for every b.

Chapter 4 Determinants: 4.1 Determinants of 2×2 matrices

Definition

For 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant of A is defined to be

$$\det(A) := ad - bc.$$

Remark

i. $\det(AB) = \det(A)\det(B) = \det(BA)$

ii. Not linear, but "linear on each row/column"

Theorem (4.1)

The function $\det : M_{2\times 2}(F) \to F$ *is a linear function of each row of a* 2×2 *matrix when the other row is held fixed. That is, if* $u, v,$ *and* w *are row vector in* F^2 *and* $k \in F$, *we have*

$$\det\begin{pmatrix} u + kv \\ w \end{pmatrix} = \det\begin{pmatrix} u \\ w \end{pmatrix} + k\det\begin{pmatrix} v \\ w \end{pmatrix}.$$

Inverse of 2×2 matrices, geometric meaning of determinant

Theorem (4.2)

For 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) \neq 0$ if and only if A is invertible. Inverse of A is

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Geometric meaning of determinant: Let $u, v \in F^2$ be two row vectors, then the area of the parallelogram having u and v as adjacent sides is $\left|\det \begin{pmatrix} u \\ v \end{pmatrix}\right|$

Definition (orientation)

If $\beta = \{u, v\}$ is an ordered basis for R^2, we define the orientation of β to be the real number

$$O\begin{pmatrix} u \\ v \end{pmatrix} := \operatorname{sgn}\left(\det \begin{pmatrix} u \\ v \end{pmatrix}\right) = \begin{cases} 1 & \text{right-handed} \\ -1 & \text{left-handed} \end{cases}.$$

4.2 Determinant of $n \times n$ matrices

Definition

Let $A \in M_{n\times n}(F)$. If $n = 1$, so that $A = (A_{11})$, we define $\det(A) = A_{11}$. For $n \geq 2$, we define $\det(A)$ recursively as

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j}), \tag{2}$$

where $\tilde{A}_{ij}$ denotes the $(n-1) \times (n-1)$ matrix obtained from deleting row i and column j of A. The scalar $c_{ij} := (-1)^{i+j} \det(\tilde{A}_{ij})$ is called the cofactor of the entry of A in row i column j. With this definition, (2) can be rewritten as

$$\det(A) = \sum_{j=1}^{n} A_{1j} c_{1j}. \quad \textbf{(cofactor expansion along the first row of } A\textbf{)} \tag{3}$$

Continued,

Theorem (4.3)

The determinant is a linear function of each row when the remaining rows are held fixed. That is

$$\det\begin{pmatrix} r_1 \\ \vdots \\ r_{i-1} \\ u+kv \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix} = \det\begin{pmatrix} r_1 \\ \vdots \\ r_{i-1} \\ u \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix} + k\det\begin{pmatrix} r_1 \\ \vdots \\ r_{i-1} \\ v \\ r_{i+1} \\ \vdots \\ r_n \end{pmatrix}.$$

Corollary

If A has a row consisting entirely of zeros, then $\det(A) = 0$.

Continued,

Lemma

If row i of A equals e_k for some k, then $\det(A) = (-1)^{i+k}\det(\tilde{A}_{ik})$.

Theorem (4.4)

The determinant can be evaluated by cofactor expansion along row i. That is $\det(A) = \sum_{j=1}^{n}(-1)^{i+j}A_{ij}\det(\tilde{A}_{ij})$.

Theorem (4.5 interchanging two rows)

Let $A \in M_{n\times n}(F)$ and B is obtained from A by interchanging any two rows of A, then $\det(B) = -\det(A)$.

Corollary

If A has two identical rows, then $\det(A) = 0$.

Theorem (4.6 adding a multiple of one row to another)

If B is obtained by adding a multiple of one row of A to another row of A. Then $\det(B) = \det(A)$.

Continued

Corollary

If $\text{rank}(A) < n$, *then* $\det(A) = 0$.

Example

Find determinants of following matrices.

$$\det\left(\begin{bmatrix} 3a_1 & 3a_2 & 3a_3 \\ 3b_1 & 3b_2 & 3b_3 \\ 3c_1 & 3c_2 & 3c_3 \end{bmatrix}\right)$$

$$\det\left(\begin{bmatrix} 2a_1 & 2a_2 & 2a_3 \\ 3b_1 + 5c_1 & 3b_2 + 5c_2 & 3b_3 + 5c_3 \\ 7c_1 & 7c_2 & 7c_3 \end{bmatrix}\right)$$

$$\det\left(\begin{bmatrix} 1 & 5 & 0 & 0 \\ 7 & 2 & -1 & 5 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 1 & 0 \end{bmatrix}\right)$$

4.3 Properties of determinants

Theorem (4.7)

For $A, B \in M_{n\times n}(F)$, $\det(AB) = \det(A)\det(B)$.

Corollary

A *is invertible if and only if* $\det(A) \neq 0$, *furthermore* $\det(A^{-1}) = \det(A)^{-1}$.

Theorem

For $A \in M_{n\times n}(F)$, $\det(A^t) = \det(A)$.

Theorem (4.9 Cramer's rule)

Let $Ax = b$ *be the matrix form of a system of* n *linear equations in* n *unknowns, if* $\det(A) \neq 0$, *then the system has a unique solution, and*

$$x_k = \frac{\det(M_k)}{\det(A)}$$

where M_k *is the* $n \times n$ *matrix obtained by replacing column* k *of* A *by* b.

Definition (Diagonalizable linear transformation)

A linear operator T on a finite dimensional vector space V is called diagonalizable if there is an ordered basis $\beta = \{v_1, v_2, \cdots, v_n\}$ such that $[T]_\beta$ is a diagonal matrix. A square matrix A is called diagonalizable if L_A is diagonalizable.

Suppose α, β are two ordered bases of V, by theorem 2.23,

$$diag(\lambda_1, \lambda_2, \cdots, \lambda_n) = [T]_\beta = \left([I_V]_\beta^\alpha\right)^{-1}[T]_\alpha [I_V]_\beta^\alpha = S^{-1}[T]_\alpha S \tag{4}$$

where $S = [I_V]_\beta^\alpha$. Therefore, T is diagonalizable if $[T]_\alpha$ is similar to a diagonal matrix. $T(v_i) = \lambda_i v_i$ for $i = 1, \cdots, n$ and v_i's are nonzero vector.

Definition (Eigenvector and eigenvalue)

Let T be a linear operator on a vector space V. A nonzero vector $v \in V$ is called an eigenvector of T if $T(v) = \lambda v$ for some scalar λ. The scalar λ is called the eigenvalue of T corresponding to the eigenvector v.
A nonzero vector $v \in F^n$ is called an eigenvector of A if it is an eigenvector of L_A.

Eigenvalues and eigenvectors

Observation: Denoting $A = [T]_\alpha$, equation (4) also reads

$$S\ diag(\lambda_1, \lambda_2, \cdots, \lambda_n) = AS$$

column i of the transition matrix S is an eigenvector corresponds to the eigenvalue λ_i.

Theorem (5.1)

T is diagonalizable $\Leftrightarrow$ there exists an ordered basis β consisting of eigenvectors of T. Furthermore, if T is diagonalizable, $\beta = \{v_1, v_2, \cdots, v_n\}$ is an ordered basis of eigenvectors of T, then $D := [T]_\beta$ is diagonal matrix and D_{jj} is the eigenvalue corresponding to v_j.

Remark

"$\Leftarrow$": To diagonalize a matrix or a linear operator is to find a basis consisting of eigenvectors and corresponding eigenvalues.

Question: how to find eigenvectors and eigenvalues of a square matrix A.

Finding eigenvalues and corresponding eigenvectors

Noting that

$$Av = \lambda v \Leftrightarrow Av - \lambda v = 0 \Leftrightarrow (A - \lambda Id)v = 0,$$

existence of nontrivial solution v implies $\det(A - \lambda Id) = 0$ and $v \in ker(A - \lambda Id)$.

Definition (Characteristic polynomial of a matrix)

The polynomial $p(\lambda) := det(A - \lambda Id)$ is called the characteristic polynomial of A.

Definition (Characteristic polynomial of a linear transformation)

Let T be a linear operator on an n-dimensional vector space V with ordered basis β. We define the characteristic polynomial of T to be the characteristic polynomial of $[T]_\beta$.

Remark

Definition does not depend on selection of basis.

Continued,

Theorem (5.2)

A scalar λ is an eigenvalue of A if and only if $\det(A - \lambda Id) = 0$, in other words, if and only if λ is a root of the characteristic polynomial.

Theorem (5.3')

Let $A \in M_{n\times n}(R)$. The characteristic polynomial of A is a polynomial of degree n. And it has exactly n roots in $\mathbb{C}$ **(Fundamental theorem of algebra)**.

Theorem (5.4)

Let λ be an eigenvalue of T. A vector v is an eigenvector corresponding to λ if and only if $v \neq 0$ and $v \in Null(T - \lambda I)$.

Definition (Eigenspace)

Let λ be an eigenvalue of T. The subspace $E_\lambda := Null(T - \lambda I)$ is called λ-eigenspace.

5.2 Diagonalizability

Theorem (5.5)

Let T be a linear operator on a vector space V and $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of T, Let $0 \neq v_i \in E_{\lambda_i}$ for $i = 1, 2, \cdots, k$, then $v_1, v_2, \cdots, v_k$ are linearly independent.

Corollary

Let T be a linear operator on an n-dimensional vector space V. If T has n distinct eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, then T is diagonalizable.

Definition

A polynomial $p(x) \in P(F)$ splits over F if there are scalars $c, r_1, r_2, \cdots, r_n$ in F such that $p(x) = c(x - r_1)(x - r_2) \ldots (x - r_n)$

Theorem (Fundamental theorem of algebra)

Every polynomial $p(x) \in P(\mathbb{C})$ splits over $\mathbb{C}$.

Continued,

Theorem (5.6)

T is diagonalizable $\Rightarrow$ The characteristic polynomial $\det(T - \lambda Id))$ splits.

Remark

The converse is false. Example $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$.

Definition (multiplicity)

Let λ_0 be an eigenvalue of a linear operator T with characteristic polynomial $f(t)$.
The **algebraic multiplicity** of λ_0 is the largest k for which $(\lambda - \lambda_0)^k$ is a factor of $\det(T - \lambda Id)$.
The **geometric multiplicity** of λ_0 is $dim(E_{\lambda_0})$.

Theorem (5.7)

$1 \leq$ *geometric multiplicity of* $\lambda_0 \leq$ *algebraic multiplicity of* λ_0.

continued

Lemma

Let T be a linear operator, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalue of T. For each $i = 1, 2, \cdots, k$, let $v_i \in E_{\lambda_i}$. If

$$v_1 + v_2 + \cdots + v_k = 0,$$

then $v_i = 0$ for all i.

Theorem (5.8)

Let T be a linear operator, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalue of T. For each $i = 1, 2, \cdots, k$, let S_i be a finite linearly independent subset of the eigenspace E_{λ_i}. Then $S = S_1 \cup S_2 \cup \ldots \cup S_k$ is a a linearly independent subset of V.

The main theorem

Theorem (5.8)

Let T be a linear operator, and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalue of T. For each $i = 1, 2, \cdots, k$, let S_i be a finite linearly independent subset of the eigenspace E_{λ_i}. Then $S = S_1 \cup S_2 \cup \ldots \cup S_k$ is a a linearly independent subset of V.

Theorem (5.9 main)

Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of T. Then

(a) *T is diagonalizable if and only if the algebraic multiplicity of λ_i is equal to geometric multiplicity for all i.*

(b) *If T is diagonalizable and β_i is an ordered basis for E_{λ_i} for each i, then $\beta = \cup_{i=1}^k \beta_i$ is an ordered basis for V consisting of eigenvectors of T.*

Eigen decomposition and eigen projections

Corollary (Eigen decomposition)

Let T be a diagonalizable linear operator on V. Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of T. For $v \in V$, there are unique vectors $v_i \in E_{\lambda_i}$ for $i = 1, \cdots, k$, such that $v = \sum_{i=1}^{k} v_i$. That is

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_k}.$$

Furthermore, there are eigenprojections P_{λ_i} such that $\mathcal{R}(P_{\lambda_i}) = E_{\lambda_i}$. Hence

$$V = \mathcal{R}(P_{\lambda_1}) \oplus \mathcal{R}(P_{\lambda_2}) \oplus \cdots \oplus \mathcal{R}(P_{\lambda_k}).$$

Example, continued

Example

$$\begin{bmatrix} 4 & 0 & -2 \\ 8 & 2 & -8 \\ 4 & 0 & -2 \end{bmatrix}, \begin{matrix} \lambda_1 = 2, 2 \\ \lambda_2 = 0 \end{matrix}, E_2 = span\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad E_0 = span\left\{ \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \right\}.$$

Hence,

$$\begin{bmatrix} 4 & 0 & -2 \\ 8 & 2 & -8 \\ 4 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix}^{-1}.$$

Q: $P_{\lambda_1}, P_{\lambda_2}$?

Example, continued

Example

$$\begin{bmatrix} 4 & 0 & -2 \\ 8 & 2 & -8 \\ 4 & 0 & -2 \end{bmatrix}, \begin{matrix} \lambda_1 = 2, 2 \\ \lambda_2 = 0 \end{matrix}, E_2 = span\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}, \quad E_0 = span\left\{ \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \right\}.$$

Hence,

$$\begin{bmatrix} 4 & 0 & -2 \\ 8 & 2 & -8 \\ 4 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix}^{-1}.$$

Q: $P_{\lambda_1}, P_{\lambda_2}$?Let

$$P_2 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & 0 & -1 \\ 4 & 1 & -4 \\ 2 & 0 & -1 \end{bmatrix},$$

$$P_0 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 4 \\ 0 & 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 0 & 1 \\ -4 & 0 & 4 \\ -2 & 0 & 2 \end{bmatrix}.$$

5.4 Invariant subspace and the Cayley-Hamilton Theorem

Later, before Chapter 7

Definition (Inner product spaces)

An inner product space V is a vector space over F (either $\mathbb{C}$ or $\mathbb{R}$) together with an inner product, i.e., with a map

$$\langle \cdot, \cdot \rangle : V \times V \to F$$

that satisfies for all $x, y, z \in V$ and $a \in F$ the following three properties:

(a) Conjugate symmetry:
$$\langle x, y \rangle = \overline{\langle y, x \rangle};$$

(b) Linearity in the first argument:
$$\langle ax, y \rangle = a\langle x, y \rangle, \quad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle;$$

(c) Positive-definite:
$$\langle x, x \rangle > 0, \quad \text{for } x \neq 0.$$

Typical inner product spaces

Example

1. $\mathbb{C}^n$ with $\langle x, y\rangle = y^* x$ where $y^* = \overline{y^t}$ is the conjugate transpose of column vector y; $\mathbb{R}^n$ with $\langle x, y\rangle = y^t x$;
2. $C([a,b];\mathbb{R})$ continuous real-valued functions on $[a,b]$ with
$$\langle f(x), g(x)\rangle = \int_a^b f(x)g(x)dx;$$
$C([a,b];\mathbb{C})$ continuous complex-valued functions on $[a,b]$ with
$$\langle f(x), g(x)\rangle = \int_a^b f(x)\overline{g(x)}dx;$$
3. $M_{n\times n}(\mathbb{C})$ with $\langle A, B\rangle = trace(B^* A)$.

Elementary properties of inner product

Theorem (6.1)

Let V be an inner product space. Then for $x, y, z \in V$ and $c \in F$, the following statements are true.

(a) *Conjugate linearity in the second argument:*

$$\langle x, ay\rangle = \bar{a}\langle x, y\rangle, \quad \langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle.$$

Inner product is a sesquilinear form.

(b) $\langle x, 0\rangle = \langle 0, x\rangle = 0$;

(c) $\langle x, x\rangle = 0$ *if and only if* $x = 0$;

(d) *If* $\langle x, y\rangle = \langle x, z\rangle$ *for all* $x \in V$, *then* $y = z$.

Definition (Induced norm)

An inner product naturally induces an associated norm defined by $||x|| = \sqrt{\langle x, x\rangle}$ on V.

Continued,

Theorem (6.2)

The following statements hold:

(a) $||cx|| = |c| \cdot ||x||$;

(b) $||x|| \geq 0$ *for all* $x \in V$ *and* $||x|| = 0$ *if and only if* $x = 0$;

(c) *Cauchy-Schwarz Inequality* $|\langle x, y\rangle| \leq ||x|| \cdot ||y||$;

(d) *Triangle Inequality* $||x + y|| \leq ||x|| + ||y||$.

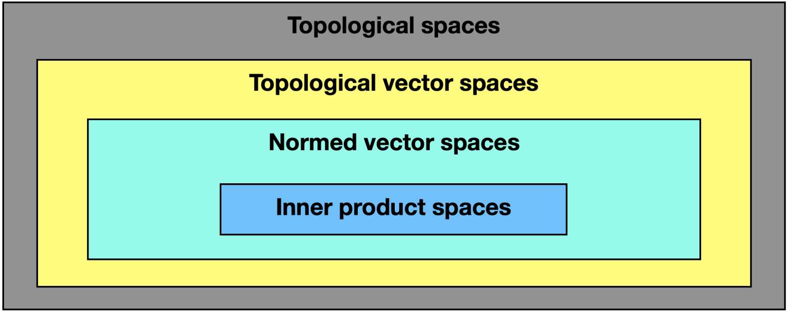
Remark

(a),(b),(d)$\Rightarrow$ $|| \cdot ||$ is a norm on V. Hence, inner product spaces are normed vector spaces.

Definition

Two vectors x, y are called orthogonal if $\langle x, y\rangle = 0$.
A vector x is called a unit vector if $||x|| = 1$.
A subset S of V is called orthogonal if any two distinct vectors in S are orthogonal and orthonormal if S is orthogonal and consists entirely of unit vectors.

Hierarchy of mathematical spaces



6.2 Gram-Schmidt orthogonalization process

Definition (Orthonormal Basis)

An ordered basis β of V is called an orthonormal basis if β is orthonormal.

Theorem (6.3)

Let V be an inner product space and $S = \{v_1, v_2, \cdots, v_k\}$ be an orthogonal subset of V consisting of nonzero vectors. If $y \in span(S)$, then $y = \sum_{i=1}^{k} \frac{\langle y, v_i \rangle}{||v_i||^2} v_i$.

Corollary (1)

If, in addition to the hypothese of Thm 6.3, S is orthonormal, then $y = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$.

Corollary (2)

The subset in Thm 6.3 is linearly independent.

Generating an orthogonal basis

Theorem (Gram-Schmidt process)

Let V be an inner product space and $S = \{w_1, w_2, \cdots, w_n\}$ be a linearly independent subset of V. Define $S' = \{v_1, v_2, \cdots, v_n\}$, where $v_1 = w_1$ and

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{||v_j||^2} v_j, \text{ for } 2 \le k \le n.$$

Then S' is an orthogonal set of nonzero vectors such that $span(S') = span(S)$.

Q: How to generate an orthonormal basis?

Generating an orthogonal basis

Theorem (Gram-Schmidt process)

Let V be an inner product space and $S = \{w_1, w_2, \cdots, w_n\}$ be a linearly independent subset of V. Define $S' = \{v_1, v_2, \cdots, v_n\}$, where $v_1 = w_1$ and

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{||v_j||^2} v_j, \text{ for } 2 \leq k \leq n.$$

Then S' is an orthogonal set of nonzero vectors such that $span(S') = span(S)$.

Q: How to generate an orthonormal basis?

Theorem (Gram-Schmidt process that generates an orthonormal basis)

Let $v_1 = w_1/||w_1||$ and

$$\tilde{v}_k = w_k - \sum_{j=1}^{k-1} \langle w_k, v_j \rangle v_j, \quad v_k = \tilde{v}_k/||\tilde{v}_k||, \quad \text{for } 2 \leq k \leq n.$$

Orthogonal complement space

Theorem (6.5)

Let V be a nonzero n-dimensional inner product space. Then V has an orthonormal basis $\beta = \{v_1, v_2, \cdots, v_n\}$. Furthermore, $x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$ for all $x \in V$. The scalars $\langle x, v_i \rangle$ $i = 1, 2, \cdots, n$ are called Fourier coefficients.

Corollary

Let T be a linear operator on V and $A = [T]_\beta$. Then $A_{ij} = \langle T(v_j), v_i \rangle$.

Definition (Orthogonal complement)

Let V be an inner product space and S be a subset of V. We define the Orthogonal complement of S as

$$S^\perp := \{x \in V : \langle x, y \rangle = 0, \quad \forall y \in S\}.$$

Remark

$S^\perp$ is a subspace of V.

Orthogonal decomposition

Theorem (6.6 Orthogonal decomposition)

Let W be a finite-dimensional subspace of an inner product space V. Then for $y \in V$, there exists unique vectors $u \in W$ and $z \in W^{\perp}$ such that $y = u + z$. Furthermore, if $\{v_1, v_2, \cdots, v_k\}$ is an orthonormal basis for W, then

$$u = \sum_{i=1}^{k} \langle y, v_i \rangle v_i.$$

Furthermore, u is the unique vector in W that is closest to y, that is $||y - x|| \geq ||y - u||$ for all $x \in W$ and equality holds if and only if $x = u$.

Remark

The above theorem implies $V = W \oplus W^{\perp}$.

6.3 The adjoint operator

Definition (Linear functional, Dual space)

Let V be a vector space over F. A linear transformation from V to F is called a linear functional. The vector space $\mathcal{L}(V, F)$ is called the dual space of V and denotes by V'.

Remark

For $y \in V$, let $T_y(x) = \langle x, y \rangle$, we have $T_y \in V'$. The following theorem implies every functional in V' is equal to a T_y for some $y \in V$. That is every functional is of the form T_y.

Theorem (6.8 Riesz representation theorem finite dimension version)

Let V be a finite dimensional inner product space and $g \in V'$, then there is a unique vector $y \in V$ such that $g(x) = \langle x, y \rangle$ for all $x \in V$.

Remark

By proof of theorem, $y = \sum_{i=1}^{n} \overline{g(v_i)} v_i$ where $\{v_1, v_2, \cdots, v_n\}$ is an orthonormal basis of V.

Adjoint operator

Theorem (6.9 Adjoint operator)

Let V be a finite dimensional inner product space. And T be a linear operator on V. Then there exists a unique function $T^* : V \to V$ such that $\langle T(x), y\rangle = \langle x, T^*(y)\rangle$ for all $x, y \in V$. Furthermore, T^* is linear.

Theorem (6.10)

$[T^*]_\beta = [T]_\beta^*$

Theorem (6.11)

(a) $(T+U)^* = T^* + U^*$; $(A+B)^* = A^* + B^*$;

(b) $(cT)^* = \bar{c}T^*$; $(cA)^* = \bar{c}A^*$;

(c) $(TU)^* = U^*T^*$; $(AB)^* = B^*A^*$;

(d) $T^{**} = T$; $A^{**} = A$;

5.4 Invariant subspaces preliminary

Definition (Invariant subspaces)

Let T be a linear operator on a vector space V. A subspace W of V is called a T-invariant subspace of V if $T(W) \subset W$.

Remark

(a) Invariant subspaces: $\{0\}$, V, $\mathcal{R}(T)$, $\mathcal{N}(T)$, E_λ;

(b) Can restrict T on W (T_W)and consider matrix representation, eigenvalues, eigenvectors, diagonalizability on W space.

Example

$T(a,b,c) = (-b+c, a+c, 3c)$. $W = span\{e_1, e_2\}$. $[T_W]$, eigenvalues ?

Theorem (5.21)

Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Then the characteristic polynomial of T_W divides the characteristic polynomial of T. That is $\det(T_W - \lambda Id) | \det(T - \lambda Id)$.

6.4 Normal and self-adjoint operators

Theorem (Schur)

Let T be a linear operator on a finite dimensional inner product space V. Suppose that the characteristic polynomial of T splits. Then there exists an orthonormal basis β such that the matrix $[T]_\beta$ is upper trianglur.

Definition

A linear operator is normal if $TT^* = T^*T$, self-adjoint if $T = T^*$, and unitary (Hermitian) if $TT^* = T^*T = Id$.

Remark

Self-adjoint and unitary operators are normal.

Continued

Theorem (6.15)

The following statements hold for normal operators:

(a) $||T(x)|| = ||T^*(x)||$.

(b) $T - cI$ *is also normal.*

(c) $T(x) = \lambda x \Leftrightarrow T^*(x) = \bar{\lambda}x$.

(d) *If* λ_1, λ_2 *are distinct eigenvalues of* T *with corresponding eigenvectors* x_1 *and* x_2, *then* x_1 *and* x_2 *are othogonal.*

Theorem (6.16)

Let T *be a linear operator on a finite-dimensional complex inner product space* V. *Then,*
T *is normal* $\Leftrightarrow$ *There exists orthonormal basis for* V *consisting of eigenvectors of* T.

Corollary (Thm 6.19)

Matrix A *is normal* $\Leftrightarrow$ A *is unitarily equivalent to a diagnonal matrix.*

6.5 Unitary and orthogonal operator

Theorem (6.17)

Let T be a linear operator on a finite-dimensional real inner product space V. Then, T is self-adjoint $\Leftrightarrow$ There exists orthonormal basis for V consisting of eigenvectors of T.

Corollary (Thm 6.20)

Real matrix A is symmetric $\Leftrightarrow$ A is orthogonally equivalent to a diagnonal matrix.

Theorem (6.18)

The following statements are equivalent.

(a) T is unitary $TT^* = T^*T = Id$.

(b) T preserves inner product: $\langle T(x), T(y)\rangle = \langle x, y\rangle$ for all $x, y \in V$.

(c) If β is an orthonormal basis for V, then $T(\beta)$ is an orthonormal basis for V.

(d) There exists an orthonormal basis β for V such that $T(\beta)$ is an orthonormal basis for V.

(e) T preserves norm (isometry): $||T(x)|| = ||x||$ for all $x \in V$.

Introduction of Jordan canonical form

Definition (Jordan block)

A **Jordan block** J is a matrix, having zeros everywhere except along the diagonal and superdiagonal, with each element of the diagonal consisting of a single number λ, and each element of the superdiagonal consisting of a 1.

$$J_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

Let V be a vector space with an ordered basis $\beta = \{v_1, \cdots, v_m\}$. If $[T]_\beta = J_\lambda$, what can we obtain from this?

Continued

Theorem (Jordan canonical form of a matrix)

Let $A \in M_{n\times n}(\mathbb{C})$. Then A is similar to a block diagonal matrix J:

$$J = \begin{bmatrix} J_{\lambda_1} & O & \cdots & O \\ O & J_{\lambda_2} & \cdots & O \\ \vdots & \vdots & & \vdots \\ O & O & \cdots & J_{\lambda_k} \end{bmatrix}$$

where each block matrix J_{λ_i} is a Jordan block. The matrix J (unique up to the order of the Jordan blocks) is called the **Jordan canonical form** *of matrix A*

Remark

i. λ_i are not necessarily distinct.

ii. A is diagonalizable if and only if every Jordan block J_{λ_i} is a one by one matrix.

Proof.

Generalized eigenspaces; Invariant subspaces; Cayley-Hamilton theorem. □