Module 8: Linear Algebra II

Operational math bootcamp

Statistical Sciences
UNIVERSITY OF TORONTO

Emma Kroell

University of Toronto

July 25, 2023

Outline

Last time:

- Vector spaces and subspaces
- Linear independence and bases
- Linear maps, null space, range

Today:

- Inverses of linear maps
- Matrices as linear maps
- Determinants
- Inner product spaces

Definition (Product of linear maps)

Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. We define the product $ST \in \mathcal{L}(U, W)$ for $\mathbf{u} \in U$ as $ST(\mathbf{u}) = S(T(\mathbf{u}))$.

Definition

A linear map $T : U \to V$ is *invertible* if there exists a linear map $S : V \to U$ such that ST is the identity map on U and TS is the identity map on V. Such a map S is called the *inverse* of T.

If T is invertible, we denote the inverse by T^{-1}. This is justified by the fact that the inverse is unique:

Proposition

Any invertible linear map has a unique inverse.

Proof.

$\square$

Theorem

A linear map is invertible if and only if it is injective and surjective.

See proof in the book.

Definition

An invertible linear map is called an *isomorphism*. If there exists an isomorphism from one vector space to another, we say that the vector spaces are *isomorphic*.

Theorem

Two finite-dimensional vector spaces over $\mathbb{F}$ are isomorphic if and only if they have the same dimension.

Proof.

$(\Rightarrow)$

$\square$

Proof continued

$(\Leftarrow)$

Linear maps and matrices

Example

Let $A \in M_{m \times n}$ be a fixed matrix. Then, we can define a linear map $T_A \colon \mathbb{F}^n \to \mathbb{F}^m$ via $T_A(\mathbf{v}) = A\mathbf{v}$, where we recall matrix vector multiplication $(A\mathbf{v})_i = \sum_{k=1}^n A_{ik} v_k$ for $i = 1, \ldots, m$.

Next we will see that we can use matrices to represent linear maps between finite dimensional vector spaces.

Definition

Let $T \in \mathcal{L}(U, V)$ where U and V are vector spaces. Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be bases for U and V respectively. The matrix of T with respect to these bases is the $m \times n$ matrix $\mathcal{M}(T)$ with entries A_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ defined by

$$T\mathbf{u}_k = A_{1k}\mathbf{v}_1 + \cdots + A_{mk}\mathbf{v}_m$$

i.e. the kth column of A is the scalars needed to write $T\mathbf{u}_k$ as a linear combination of the basis of V:

$$T\mathbf{u}_k = \sum_{i=1}^{m} A_{ik}\mathbf{v}_i$$

Note that since a linear map $T \in \mathcal{L}(U, V)$ is uniquely determined by its image on a basis of U, we see that once we pick basis of U and V its matrix representation is uniquely determined.

Example

Let $D \in \mathcal{L}(\mathbb{P}_4(\mathbb{R}), \mathbb{P}_3(\mathbb{R}))$ be the differentiation map, $Dp = p'$. Find the matrix of D with respect to the standard bases of $\mathbb{P}_3(\mathbb{R})$ and $\mathbb{P}_4(\mathbb{R})$.

Standard basis: $1, x, x^2, x^3, (x^4)$

$T(u_1)$

$T(u_2)$

$T(u_3)$

$T(u_4)$

$T(u_5)$

The matrix is:

- Observe that if we choose bases $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ for U, V and represent $T \in \mathcal{L}(U, V)$ as a matrix $\mathcal{M}(T)$, then the corresponding map can be obtained by just working with the coordinates of vectors in U, V with respect to the chosen basis
- If $\mathbf{u} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$, then the coordinates of $T(\mathbf{u})$ with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_m$ can be obtained by the matrix vector multiplication $\mathcal{M}(T)\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is the $n \times 1$ matrix with entries α_i

Example

If we want to find the derivative of $p = x^4 + 12x^3 - 5x^2 + 7$ with respect to the standard monomial basis of $\mathbb{P}_4(\mathbb{R})$, we use $\mathcal{M}(D)$ from the previous example to obtain

$$\mathcal{M}(D)\boldsymbol{\alpha} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 7 \\ 0 \\ -5 \\ 12 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -10 \\ 36 \\ 4 \end{pmatrix}.$$

Thus, translating back into the monomial basis of $\mathbb{P}_3(\mathbb{R})$ gives $D(p) = -10x + 36x^2 + 4x^3$.

Other points

- Looking at matrices as representations of linear maps gives us an intuitive explanation for why we do matrix multiplication the way we do! In fact, we want matrix multiplication to represent composition of linear maps
- We can use matrices to solve linear systems.

Determinants

Determinant

- The determinant is a function from $M_{n \times n} \to \mathbb{F}$, i.e. it is a function from the entries of a square matrix to a real or complex number.
- The determinant has applications in solving linear systems, computing eigenvalues, etc

Example: 2×2 matrix

The determinant of a 2×2 matrix is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} =$$

Example: 3×3 matrix

There is a **trick** for finding the determinant of a 3 by 3 matrix:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

Cofactor expansion

For other $n \times n$ matrices, one can compute the determinant using **cofactor expansion**.

Definition (Cofactor expansion)

Let $A = \{a_{j,k}\}_{j,k=1}^{n}$ be a $n \times n$ matrix. Let $M_{j,k}$ denote the determinant of the $(n-1) \times (n-1)$ matrix obtained by removing the j^{th} row and the k^{th} column of A.

For each row $j = 1, \ldots, n$

$$|A| = \sum_{k=1}^{n} a_{j,k}(-1)^{j+k} M_{j,k}.$$

Similarly, for each column $k = 1, \ldots, n$

$$|A| = \sum_{j=1}^{n} a_{j,k}(-1)^{j+k} M_{j,k}.$$

The numbers $C_{j,k} = (-1)^{j+k} M_{j,k}$ are called *cofactors*.

Proposition

The determinant of a diagonal matrix or triangular matrix is the product of the entries on the diagonal.

Sketch of proof

Inverse of a matrix

Theorem

Let A be an $n \times n$ invertible matrix and let $C = \{C_{j,k}\}_{j,k=1}^{n}$ be its cofactor matrix. Then

$$A^{-1} = \frac{1}{|A|} C^T$$

Connection to last lecture: The matrix A is invertible if and only if the linear map represented by the matrix is an isomorphism.

Cramer's rule

Corollary

Suppose A is an $n \times n$ invertible matrix. The linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution given by

$$x_i = \frac{|A_i|}{|A|}, \quad i, \ldots, n,$$

where A_i is the matrix obtained by replacing the i^{th} column of A with $\mathbf{b}$.

Transpose of a matrix

Definition

The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix, denoted A^T, defined entry-wise as $\{A^T_{j,k}\} = \{A_{k,j}\}$ for $j = 1, \ldots, m$ and $k = 1, \ldots n$ (i.e. the rows of A are the columns of A^T and the columns of A are the rows of A^T)

Properties of the determinant

Proposition

$|A| \neq 0$ if and only if A is invertible

Proposition

Let A be an $n \times n$ real matrix.

1. If A has a zero column, then $|A| = 0$.
2. If A has two equal columns, then $|A| = 0$.
3. If one column of A is a multiple of another, then $|A| = 0$.
4. $|AB| = |A||B|$
5. $|\alpha A| = \alpha^n |A|$ for $\alpha \in \mathbb{F}$
6. $|A^T| = |A|$

Inner product spaces

Complex numbers

Recall that for a complex number $z = a + ib$, we define the following:

- Real part: $Re(z) = a$,
- Imaginary part: $Im(z) = b$,
- Complex conjugate: $\overline{z} = a - ib$,
- Modulus: $|z| = \sqrt{Re(z)^2 + Im(z)^2} = \sqrt{a^2 + b^2}$

We have $|z|^2 = z\overline{z}$ and $Re(z) = \frac{z+\overline{z}}{2}$.

Definition

Let V be an $\mathbb{F}$-vector space. A function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{F}$ is called *inner product* on V if the following holds:

1. (Conjugate) symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ for all $\mathbf{x}, \mathbf{y} \in V$, where $\bar{a}$ denotes the complex conjugate for $a \in \mathbb{C}$
2. Linearity in the first argument: $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha, \beta \in \mathbb{F}$
3. Positive definiteness: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ if and only if $\mathbf{x} = \mathbf{0}$

A vector space equipped with an inner product is called an *inner product space*.

Example

- Standard inner product on $\mathbb{R}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$
- Standard inner product on $\mathbb{C}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i \overline{y}_i$ for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$
- On the space of polynomials $\mathbb{P}_n(\mathbb{R})$: $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x$ for $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{P}_n(\mathbb{R})$

Proposition

Let V be an inner product space. Then $\mathbf{x} = \mathbf{0}$ if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for all $\mathbf{y} \in V$.

Proof.

$\square$

Cauchy-Schwarz Inequality

Proposition

Let V be an inner product space. Then

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}\sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}$$

for all $\mathbf{x}, \mathbf{y} \in V$.

Proof.

□

Proposition

Let V be an inner product space. Then $\langle \cdot, \cdot \rangle$ induces a norm on V via $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ for all $\mathbf{x} \in V$.

Proof

Proof continued

Adjoint

Definition

Let U, V be inner product spaces and $S \colon U \to V$ be a linear map. The *adjoint* S^* of S is the linear map $S^* \colon V \to U$ defined such that

$$\langle S\mathbf{u}, \mathbf{v} \rangle_V = \langle \mathbf{u}, S^*\mathbf{v} \rangle_U \qquad \text{for all } \mathbf{u} \in U, \mathbf{v} \in V.$$

Proposition

Let U, V be inner product spaces and $S\colon U \to V$ be a linear map. Then S^* is unique and linear.

Proof

Proof continued

Example

Define $S\colon \mathbb{R}^3 \to \mathbb{R}^2$ by $S\mathbf{x} = (2x_1 + x_3, -x_2)$. Then, for all $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$ the defining equation for the adjoint operator leads to

Proposition

Let $A \in M_{m\times n}(\mathbb{F})$ be a matrix and $T_A \colon \mathbb{F}^n \to F^m \colon \mathbf{x} \mapsto A\mathbf{x}$. Then, $T_A^*(\mathbf{x}) = A^*\mathbf{x}$, where $A^* \in M_{n\times m}(\mathbb{F})$ with $(A^*)_{ij} = \overline{A_{ji}}$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

In particular, if $\mathbb{F} = \mathbb{R}$, the adjoint of the matrix is given by its transpose,denoted A^T, and if $\mathbb{F} = \mathbb{C}$, it is given by its conjugate transpose, denoted A^*.

Proof.

$\square$

Definition

A matrix $O \in M_n(\mathbb{R})$ is called *orthogonal* if its inverse is given by its transpose, i.e. $O^T O = OO^T = I$.

A matrix $U \in M_n(\mathbb{C})$ is called *unitary* if the inverse is given by the conjugate transpose, i.e. $U^* U = UU^* = I$.

Example

- Let $\varphi \in [0, 2\pi]$. Then
$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$
is an orthogonal matrix. What does it describe geometrically?
- The following is a unitary matrix:
$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Definition

Let $A \in M_n(\mathbb{F})$. We call A *self-adjoint* if $A^* = A$. In the case $\mathbb{F} = \mathbb{R}$, such an A is called *symmetric* and if $\mathbb{F} = \mathbb{C}$, such an A is called *Hermitian*.

References

Axler S. *Linear Algebra Done Right*. 3rd ed. Undergraduate Texts in Mathematics. Springer, 2015. Available from: https://link.springer.com/book/10.1007/978-3-319-11080-6

Treil S. *Linear Algebra Done Wrong*. 2017. Available from: https://www.math.brown.edu/streil/papers/LADW/LADW.html