Module 8: Linear Algebra II

Operational math bootcamp

Statistical Sciences
UNIVERSITY OF TORONTO

Emma Kroell

University of Toronto

July 22, 2022

Outline

Last time:

- Linear independence and bases
- Linear maps, null space, range, inverses
- Matrices as linear maps

Today:

- Determinants
- Inner product spaces

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

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Notation:

$$\det(A) = |A|$$

Statistical Sciences
UNIVERSITY OF TORONTO

Example: 3×3 matrix

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

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

$$\det(A) = aei + bfg + cdh - gec - hfa - idb$$

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.

$$\begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 & \\ & & \alpha_3 \end{bmatrix} \qquad \begin{bmatrix} * & * & * \\ 0 & * & * \\ & & * \end{bmatrix}$$

Sketch of proof

$$\begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix} \qquad \alpha_1 (-1)^{1+1} M_{1,1}$$

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 $Re(z) = \frac{z+\overline{z}}{2}$.

$x \in \mathbb{R}$, $x = \overline{x}$

Definition

Let V be an $\mathbb{F}$-vector space. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ 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 $\overline{a}$ 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*.

$$\langle x, \alpha y + \beta z \rangle = \overline{\langle \alpha y + \beta z, x \rangle} = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$$

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.

($\Rightarrow$) Suppose $x = 0$. Then since $0x = 0$.
Therefore $\langle x, y \rangle = \langle 0x, y \rangle = 0 \langle x, y \rangle = 0$.
for any $y \in V$.

($\Leftarrow$) Suppose $\langle x, y \rangle = 0 \ \forall y \in V$.
Then it holds for $y = x \Rightarrow \langle x, x \rangle = 0$ □

$\Rightarrow x = 0$

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.

Take $t \in \mathbb{F}$, $x, y \in V$

$$0 \le \langle x - ty, x - ty \rangle = \langle x, x \rangle - t\langle y, x \rangle - \bar{t}\langle x, y \rangle + |t|^2 \langle y, y \rangle$$

Choose $t = \dfrac{\langle x, y \rangle}{\langle y, y \rangle}$

$$0 \le \langle x, x \rangle - 2\frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle^2}\langle y, y \rangle$$

$$= \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}$$

$$\therefore |\langle x, y \rangle|^2 \le \langle x, x \rangle \langle y, y \rangle$$

□

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

Norm: $\|x\| \geq 0 \ \forall x$, $\|x\| = 0 \iff x = 0$

IP: $\langle x, x \rangle \geq 0$ $\langle x, x \rangle = 0 \iff x = 0$

WTS $\|\alpha x\| = |\alpha| \, \|x\|$

IP: $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{|\alpha|^2 \langle x, x \rangle}$

$\alpha \in \mathbb{C}: \alpha \bar{\alpha} = |\alpha|^2$ $\Rightarrow |\alpha| \sqrt{\langle x, x \rangle}$

Proof continued

$x, y \in V$

$$\|x+y\|^2 = \langle x+y, x+y\rangle$$
$$= \langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle$$
$$= \|x\|^2 + 2\,\mathrm{Re}\langle x,y\rangle + \|y\|^2$$
$$\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2$$
$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$
$$\leq (\|x\| + \|y\|)^2$$

Adjoint

Definition

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

To show uniqueness, suppose $T: V \to U$ s.t.

$$\langle Su, v\rangle = \langle u, Tv\rangle \quad \forall u \in U, \forall v \in V$$

$$\langle Su, v\rangle = \langle u, S^* v\rangle$$

Using conjugate symmetry, $S^* = T$

Statistical Sciences
UNIVERSITY OF TORONTO

Proof continued

$$\langle u, S^*(\alpha v + w)\rangle = \langle Su, \alpha v + w\rangle$$
$$= \bar{\alpha}\langle Su, v\rangle + \langle Su, w\rangle$$
$$= \langle u, \alpha S^* v\rangle + \langle u, S^* w\rangle$$
$$= \langle u, \alpha S^* v + S^* w\rangle$$

Example

Define $S : \mathbb{R}^3 \to \mathbb{R}^2$ 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

$$\langle Sx, y \rangle_{\mathbb{R}^2} = \langle (2x_1 + x_3, -x_2), (y_1, y_2) \rangle_{\mathbb{R}^2}$$
$$= 2x_1 y_1 + x_3 y_1 - x_2 y_2$$
$$= \langle (x_1, x_2, x_3), (2y_1, -y_2, y_1) \rangle_{\mathbb{R}^3}$$
$$:= \langle x, S^* y \rangle$$

$$\therefore S^* y = (2y_1, -y_2, y_1)$$

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 $i = 1, \dots, n$ and $j = 1, \dots, m$.

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.

□

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. $U^* U = UU^* = I$.

Statistical Sciences
UNIVERSITY OF TORONTO

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:
$$U = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

$$U^* = \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix}$$

$$UU^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

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*.

$A^T = A$: symmetric

$A^* = A$: Hermitian

Orthogonality and Gram-Schmidt

Definition

Two vectors $\mathbf{x}, \mathbf{y} \in V$ are called *orthogonal* if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, denoted $\mathbf{x} \perp \mathbf{y}$. We call them *orthonormal* if additionally the vectors are normalized, i.e. $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$. A basis $\mathbf{x}_1, \ldots, \mathbf{x}_n$ of V is called *orthonormal basis (ONB)*, if the vectors are pairwise orthogonal and normalized.

Proposition

Let $\mathbf{x}_1, \ldots, \mathbf{x}_k \in V$ be orthonormal. Then the system of vectors is linearly independent.

Proof.

Suppose $0 = \sum_{i=1}^{k} \alpha_i x_i$

$$\Rightarrow 0 = \left\| \sum_{i=1}^{k} \alpha_i x_i \right\|^2$$

$$= \left\langle \sum_{i=1}^{k} \alpha_i x_i , \sum_{i=1}^{k} \alpha_i x_i \right\rangle$$

$$= \sum_{i,j=1}^{k} \langle \alpha_i x_i , \alpha_j x_j \rangle$$

$$= \sum_{i=1} |\alpha_i|^2 \|x_i\|^2 + \sum_{i,j=1 \; i \neq j}^{n} \alpha_i \overline{\alpha_j} \langle x_i , x_j \rangle$$

□

$$= \sum_{i=1}^{n} |\alpha_i|^2 \Rightarrow \alpha_i = 0$$

Proposition (Orthogonal Decomposition)

Let $\mathbf{x}, \mathbf{y} \in V$ with $\mathbf{y} \neq 0$. Then, there exist $c \in F$ and $\mathbf{z} \in V$ such that $\mathbf{x} = c\mathbf{y} + \mathbf{z}$ with $\mathbf{y} \perp \mathbf{z}$.

Given a basis we can obtain an ONB from it using the Gram-Schmidt algorithm by reiterating the orthogonal decomposition from above.

Proposition (Gram-Schmidt Algorithm)

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \in V$ be a system of linearly independent vectors. Define $\mathbf{y}_1 = \mathbf{x}_1 / \|\mathbf{x}_1\|$. For $i = 2, \ldots, n$ define $\mathbf{y}_j$ inductively by

$$\mathbf{y}_i = \frac{\mathbf{x}_i - \sum_{k=1}^{i-1} \langle \mathbf{x}_i, \mathbf{y}_k \rangle \mathbf{y}_k}{\|\mathbf{x}_i - \sum_{k=1}^{i-1} \langle \mathbf{x}_i, \mathbf{y}_k \rangle \mathbf{y}_k\|}.$$

Then the $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are orthonormal and

$$\mathrm{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} = \mathrm{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_n\}.$$

The proof is omitted but can be found in the book.

Next time

- Eigenvalues and eigenvectors
- Matrix decompositions

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