TLDR

Quick summary of today's notes. Lecture starts on next page.

- The determinant has a geometric interpretation in terms of volumes. This is the reason why determinants appear when you do variable substitutions in multivariable integrals.

  The columns of a 2×2 matrix A are the sides of a unique parallelogram in $\mathbb{R}^2$.

  The absolute value of $\det A$ is the area of this parallelogram.

  This fact generalizes to n dimensions if we replace "parallelogram" by its n-dimensional analogue.

- We introduce the concept of a *vector space* to generalize the idea of a subspace of $\mathbb{R}^n$.

  Formally, an *(abstract) vector space* is a nonempty set with a "zero vector" and two operations that can be thought of a "vector addition" and "scalar multiplication."

  These operations are subject to several conditions.

  All subspaces of $\mathbb{R}^n$, including $\mathbb{R}^n$ itself, are examples of vector spaces.

  The set of polynomials in one variable is another example of a vector space.

- There are notions of linear combinations, span, linear independence, subspaces, bases, and dimension for abstract vector spaces. The definitions are all the same as for $\mathbb{R}^n$.

- If X and Y are sets, then let $\mathsf{Fun}(X, Y)$ be the set of functions $f : X \to Y$.

  The sets $\mathsf{Fun}(X, \mathbb{R})$ and $\mathsf{Fun}(X, \mathbb{R}^n)$ are naturally vector spaces.

  More generally, if V is a vector space, then $\mathsf{Fun}(X, V)$ is naturally a vector space.

  The corresponding vector operations and zero vector are

$$\begin{aligned} f + g &= (\text{ the function that maps } x \mapsto f(x) + g(x) \text{ for } x \in X\ ), \\ cf &= (\text{ the function that maps } x \mapsto c \cdot f(x) \text{ for } x \in X\ ), \\ 0 &= (\text{ the function that maps } x \mapsto 0 \in V \text{ for } x \in X\ ), \end{aligned}$$

  for $f, g \in \mathsf{Fun}(X, V)$ and $c \in \mathbb{R}$.

  Most abstract vector spaces of interest arise naturally as subspaces of $\mathsf{Fun}(X, V)$ for some V.

- If U and V are vector spaces then a function $f : U \to V$ is *linear* if

$$f(u + v) = f(u) + f(v) \qquad \text{and} \qquad f(cv) = c \cdot f(v)$$

  for all $u, v \in U$ and $c \in \mathbb{R}$.

1 Last time: determinants

Let n be a positive integer.

Theorem. The determinant is the unique function $\det : \{ n \times n \text{ matrices} \} \to \mathbb{R}$ such that

(1) $\det I_n = 1$.

(2) Switching two columns reverses the sign of the determinant.

(3) $\det A$ is linear as a function of a single column A if all other columns are fixed.

For 1×1 and 2×2 matrices, we have $\det \left[\begin{array}{c} a \end{array}\right] = a$ and $\det \left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = ad - bc$.

The *diagonal (positions)* of an $n \times n$ matrix are the positions $(1,1)$, $(2,2)$, …, (n,n).

The *diagonal entries* of a matrix are the entries in these positions.

A matrix is *upper triangular* if all of its nonzero entries are in positions on or above the diagonal.

A matrix is *lower triangular* if all of its nonzero entries are in positions on or below the diagonal.

A *triangular matrix* is a square matrix which is either upper or lower triangular.

A *diagonal matrix* is a matrix which is both upper and lower triangular: in other words, all of its nonzero entries appear in diagonal positions.

Theorem. If A is triangular square matrix then $\det A$ is the product of the diagonal entries of A.

Theorem. A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem. If A and B are $n \times n$ matrices then $\det(AB) = (\det A)(\det B)$ and $\det(A^T) = \det A$.

Algorithm to compute $\det A$.

Input: an $n \times n$ matrix A.

1. Start by setting $c = 1$.
2. Row reduce A to an echelon form E, while doing the following:
   (a) When you switch two rows, multiply c by -1.
   (b) When you rescale a row by a nonzero factor λ, divide c by λ.
   (c) When you add a multiple of a row to another row, don't do anything to c.

The determinant of A is then given by $\det A = c \cdot \det E =$ the product of c and all the diagonal entries of the echelon form E.

Another (usually more time consuming) way to compute $\det A$:

Theorem. Fix a matrix

$$A = \left[\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{array}\right].$$

Define $A^{(i,j)}$ as the submatrix formed by deleting row i and column j. Then

$$\boxed{\det A = a_{11} \det A^{(1,1)} - a_{12} \det A^{(1,2)} + a_{13} \det A^{(1,3)} - \cdots - (-1)^n a_{1n} \det A^{(1,n)}}$$

Each $A^{(1,j)}$ is a square matrix smaller than A, so $\det A^{(1,j)}$ can be computed by the same formula.

Example. $\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei-fh)-b(di-fg)+c(dh-eg) = a(ei-fh)-d(bi-ch)+g(bf-ce).$

The recursive formula for $\det A$ is most useful if A has many entries which are zero.

Example. If $A = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 3 & 4 & 5 \\ 1 & 6 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$ then $\det A = \det \begin{bmatrix} 3 & 4 & 5 \\ 6 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} - 0 + 2 \det \begin{bmatrix} 0 & 3 & 5 \\ 1 & 6 & 0 \\ 0 & 1 & 1 \end{bmatrix} - 0$ and

$$\det \begin{bmatrix} 3 & 4 & 5 \\ 6 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} 3 & 6 & 1 \\ 4 & 0 & 1 \\ 5 & 0 & 1 \end{bmatrix} = -\det \begin{bmatrix} 6 & 3 & 1 \\ 0 & 4 & 1 \\ 0 & 5 & 1 \end{bmatrix} = -\det \begin{bmatrix} 6 & 0 & 0 \\ 3 & 4 & 5 \\ 1 & 1 & 1 \end{bmatrix} = -6 \det \begin{bmatrix} 4 & 5 \\ 1 & 1 \end{bmatrix} = 6$$

since we taking transposes doesn't change the determinant, and switching columns reverses the sign of the determinant. Similarly, we have

$$\det \begin{bmatrix} 0 & 3 & 5 \\ 1 & 6 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} 0 & 1 & 0 \\ 3 & 6 & 1 \\ 5 & 0 & 1 \end{bmatrix} = -\det \begin{bmatrix} 3 & 1 \\ 5 & 1 \end{bmatrix} = -(3-5) = 2.$$

Therefore $\det A = 6 + 2 \cdot 2 = 10$.

2 Interpreting the determinant geometrically

The last thing we'll mention is how the determinant of a matrix A can be interpreted as measuring the volume of the region bounded by the columns of A. This easiest way to explain this is in 2 dimensions:

Proposition. If $u, v \in \mathbb{R}^2$ are two vectors and $A = \begin{bmatrix} u & v \end{bmatrix}$ then the area of the parallelogram with sides u and v is $|\det A|$.

Proof idea. Make things simple by putting u and v both in the first quadrant. Draw a picture of the parallelogram P inside the rectangle R whose diagonal is $u + v$ and whose sides are on the x- and y-axes. Then compute the area of P by subtracting the areas of an appropriate number of rectangular and triangular regions from R. One finds that this area is $ad - bc$ if $u = \begin{bmatrix} a \\ c \end{bmatrix}$ and $v = \begin{bmatrix} b \\ d \end{bmatrix}$. □

This property generalizes to 3 and higher dimensions, once we replace "area" by "volume" and then specify what the higher dimensional analogue of a parallelogram is. (In 3 dimensions, it's the *parallelepiped*).

Theorem. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation with standard matrix A. If S is a parallelogram in the $\mathbb{R}^2$ plane then the area of $T(S)$ is the area of S times $|\det A|$.

3 Vector spaces

This course focuses on $\mathbb{R}^n$ and its subspaces.

These objects are examples of *(real) vector spaces*.

There is also a notion of a *complex vector space* where our scalars can be complex numbers from $\mathbb{C}$ rather than just $\mathbb{R}$. Essentially all of the theory is the same, so for now we stick to real vector spaces which are more closely aligned with applications.

The definition of a general vector space is given as follows:

Definition. A *vector space* is a nonempty set V with two operations called *vector addition* and *scalar multiplication* satisfying several conditions. We refer to the elements of V as *vectors*.

The vector addition operation for V should be a rule that takes two input vectors $u, v \in V$ and produces an output vector $u + v \in V$ such that

(a) $u + v = v + u$.

(b) $(u + v) + w = u + (v + w)$.

(c) There exists a unique *zero vector* $0 \in V$ with the property that $0 + v = v$ for all $v \in V$.

The scalar multiplication operation for V should be a rule that takes a scalar input $c \in \mathbb{R}$ and an input vector $v \in V$ and produces an output vector $cv \in V$ such that

(a) If $c = -1$ then $v + (-1)v = 0$.

(b) $c(u + v) = cu + cv$.

(c) $(c + d)v = cv + dv$ for $c, d \in \mathbb{R}$.

(d) $c(dv) = (cd)v$ for $c, d \in \mathbb{R}$.

(e) If $c = 1$ then $1v = v$.

Notation: If V is a vector space and $v \in V$ then we define $-v = (-1)v$ and $u - v = u + (-v)$.

Example. $\mathbb{R}^n$ and any subspace of $\mathbb{R}^n$ is a vector space, with the usual operations of vector addition and scalar multiplication.

Example. Let $\mathbb{R}^\infty$ be the set of infinite sequences $a = (a_1, a_2, a_3, \dots)$ of real numbers $a_i \in \mathbb{R}$. Define

$$a + b = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \dots) \qquad \text{and} \qquad ca = (ca_1, ca_2, ca_3, \dots)$$

for $a, b \in \mathbb{R}^\infty$ and $c \in \mathbb{R}$.

These operations make $\mathbb{R}^\infty$ into a vector space.

The zero vector in this space is the sequence $0 = (0, 0, 0, \dots) \in \mathbb{R}^\infty$.

It is rarely necessary to check the axioms of a vector space in detail, and not too useful to memorise the abstract definition. If we have a set with operations that look like vector addition and scalar multiplication for $\mathbb{R}^n$, then we usually have a vector space. Moreover, it's usually easy to identify any vector space we encounter as a special case of a few general constructions like the following:

Example. Let X be any set. Define $\mathsf{Fun}(X, \mathbb{R})$ as the set of functions $f : X \to \mathbb{R}$.

Given $f, g \in \mathsf{Fun}(X, \mathbb{R})$ define $f + g$ as the function with the formula

$$(f + g)(x) = f(x) + g(x) \qquad \text{for } x \in X.$$

Given $c \in \mathbb{R}$ and $f \in \mathsf{Fun}(X, \mathbb{R})$, define cf as the function with the formula

$$(cf)(x) = cf(x) \qquad \text{for } x \in X.$$

The set $\mathsf{Fun}(X, \mathbb{R})$ is a vector space relative to these operations.

The corresponding zero vector in $\mathsf{Fun}(X, \mathbb{R})$ is the function $f(x) = 0$.

In a sense which can be made precise, we have

$\mathbb{R}^n = \mathsf{Fun}(\{1, 2, 3, \ldots, n\}, \mathbb{R})$.

$\mathbb{R}^\infty = \mathsf{Fun}(\{1, 2, 3, \ldots\}, \mathbb{R})$.

More generally, if V is any vector space then the set of functions $\mathsf{Fun}(X, V) = \{f : X \to V\}$ is a vector space for similar definitions of vector addition and scalar multiplication.

As an example of how one can use the axioms to prove properties of a general vector space, consider the following identities which are obvious for subspaces of $\mathbb{R}^n$.

Proposition. If V is a vector space then $0v = 0$ and $c0 = 0$ for all $c \in \mathbb{R}$ and $v \in V$.

Proof. We have $0v = (0 + 0)v = 0v + 0v$ so $0 = 0v - 0v = (0v + 0v) - 0v = 0v + (0v - 0v) = 0v + 0 = 0v$.

Similarly, $c0 = c(0 + 0) = c0 + c0$ so $0 = c0 - c0 = (c0 + c0) - c0 = c0 + (c0 - c0) = c0 + 0 = c0$. □

4 Subspaces, bases, and dimension

Notions of subspaces, bases, and dimension for abstract vector spaces are essentially the same as for $\mathbb{R}^n$.

Definition. A *subspace* of a vector space V is a subset H containing the zero vector of V, such that if $u, v \in H$ and $c \in \mathbb{R}$ then $u + v \in H$ and $cv \in H$.

If $H \subset V$ is a subspace then H is itself a vector space with the same operations of scalar multiplication and vector addition.

Example. V is a subspace of itself and $\{0\} \subset V$ is a subspace.

Example. $\mathbb{R}^2$ is technically not a subspace of $\mathbb{R}^3$ since $\mathbb{R}^2$ is not a subset of $\mathbb{R}^3$.

Example. Let X be any set. Let $Y \subset X$ be a subset. Define H as the subset of $\mathsf{Fun}(X, \mathbb{R})$ consists of the functions $f : X \to \mathbb{R}$ with $f(y) = 0$ for all $y \in Y$. Then H is a subspace.

Example. The set of all functions $\mathsf{Fun}(\mathbb{R}^n, \mathbb{R}^m)$ is a vector space since $\mathbb{R}^m$ is a vector space. The subset of linear functions $f : \mathbb{R}^n \to \mathbb{R}^m$ is a subspace of this vector space.

Let V be a vector space.

Definition. A *linear combination* of a finite list of vectors $v_1, v_2, \ldots, v_k \in V$ is a vector of the form

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k$$

for some scalars $c_1, c_2, \ldots, c_k \in \mathbb{R}$. A linear combination by definition only involves finitely many vectors.

Definition. The *span* of a set of vectors is the set of all linear combinations that can be formed from the vectors. It is important to note that each such linear combination can only involve finitely many vectors at a time. The span of a set of vectors in V is a subspace of V.

Example. Let $V = \mathsf{Fun}(\mathbb{R}, \mathbb{R})$. The span of the infinite set of functions $1, x, x^2, x^3, \dots \in V$ is the subspace of *polynomial functions*. Note that each polynomial function is a linear combination of a finite number of monomials $c_nx^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0$.

Definition. A finite list of vectors $v_1, v_2, \dots, v_k \in V$ is *linearly independent* if it is impossible to express $0 = c_1v_1 + c_2v_2 + \cdots + c_kv_k$ except when $c_1 = c_2 = \cdots = c_k = 0$. An infinite list of vectors is linearly independent if every finite subset is linearly independent.

Definition. A *basis* of a vector space V is a subset of linearly independent vectors whose span is V. Saying $b_1, b_2, b_3, \dots$ is a basis for V is the same as saying that each $v \in V$ can be expressed as a uniquely linear combination of basis elements.

Theorem. Let V be a vector space.

1. V has at least one basis.
2. Every basis of V has the same size.
3. If A is a subset of linearly independent vectors in V then V has a basis B with $A \subset B$.
4. If C is a subset of vectors in V whose span is V then V has a basis B with $B \subset C$.

When V has a basis that is finite in size, the proof of the previous theorem follows from the case when V is a subspace $\mathbb{R}^n$ (which was shown in earlier lectures). When V has no finite basis, the properties in the theorem still hold, but their proofs in general are slightly beyond the scope of this course.

Definition. As for subspaces of $\mathbb{R}^n$, we define the *dimension* of a vector space V to be the common size of any of its bases. Denote the dimension of V by $\dim V$.

Corollary. If $H \subset V$ is a subspace then $\dim H \leq \dim V$, and if $\dim H = \dim V$ then $H = V$.

Proof. This follows from the last two parts of the previous theorem. □

Example. If X is a finite set then $\dim \mathsf{Fun}(X, \mathbb{R}) = |X|$ where $|X|$ is the size of X. A basis is given by the functions $\delta_y : X \to \mathbb{R}$ for $y \in X$, defined by the formulas

$$\delta_y(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y. \end{cases}$$

The following is a more interesting example involving the space of solutions of a differential equation. The problem of of describing all solutions to a differential equation is an important motivation for the consideration of abstract vectors spaces (rather than just subspaces of $\mathbb{R}^n$) in the first place.

Example. Let V be the subset of $\mathsf{Fun}(\mathbb{R}, \mathbb{R})$ of twice-differentiable functions $f : \mathbb{R} \to \mathbb{R}$ with

$$f'' + f = 0.$$

Here f'' denotes the second derivative of f. The subset V is a subspace of $\mathsf{Fun}(\mathbb{R}, \mathbb{R})$ (check this!).

The vector space V contains the functions $\cos x$ and $\sin x$ since $(\cos x)' = -\sin x$ and $(\sin x)' = \cos x$.

These functions are linearly independent since if we could express

$$a \cos x + b \sin x = 0 \qquad \text{for all } x \in \mathbb{R}$$

then setting $x = 0$ would imply $a = 0$ and setting $x = \pi/2$ would imply $b = 0$.

We conclude that $\dim V \geq 2$. What is $\dim V$? Is it finite? We'll answer this question in a moment.

Suppose U and V are vector spaces.

Definition. A function $f : U \to V$ is *linear* if

$$f(u+v) = f(u) + f(v) \qquad \text{and} \qquad f(cv) = cf(v) \qquad \text{for all } c \in \mathbb{R} \text{ and } u, v \in U.$$

Define $\mathrm{range}(f) = \{f(x) : x \in U\}$ and $\mathrm{kernel}(f) = \{x \in U : f(x) = 0\}$.

Proposition. If $f : U \to V$ is linear then $\mathrm{range}(f)$ and $\mathrm{kernel}(f)$ are subspaces.

These subspaces are generalisations of the column space and null space of a matrix.

Proposition. If U, V, W are vector spaces and $f : V \to W$ and $g : U \to V$ are linear functions then $f \circ g : U \to V \to W$ is linear, where $f \circ g(x) = f(g(x))$.

Check this yourself!

If D is the subspace of twice-differentiable functions in $\mathsf{Fun}(\mathbb{R}, \mathbb{R})$ and $\mathcal{L} : D \to \mathsf{Fun}(\mathbb{R}, \mathbb{R})$ is the function $\mathcal{L}(f) = f'' + f$, then $\mathcal{L}$ is a linear map and the subspace

$$V = \{f \in D : f'' + f = 0\}$$

in our previous example is precisely $\mathrm{kernel}(\mathcal{L})$.

To compute the dimension of this subspace, some notation is useful. Define

$0! = 1$ and $n! = n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1$ for integers $n > 0$.

So that in general $n!$ (pronounced "n factorial") is the product of all positive integers at most n.

Now suppose we could write $f \in V$ as

$$f(x) = a_0/0! + a_1x/1! + a_2x^2/2! + a_3x^3/3! + a_4x^4/4! + \dots$$

for some real numbers $a_0, a_1, a_2, a_3, a_4, \dots \in \mathbb{R}$. Then

$$f'(x) = a_1/0! + a_2x/1! + a_3x^2/2! + a_4x^3/3! + a_5x^4/4! + \dots$$

and

$$f''(x) = a_2/0! + a_3x/1! + a_4x^2/2! + a_5x^3/3! + a_6x^4/4! + \dots$$

Since $f'' + f = 0$ we have

$$0 = (a_0 + a_2)/0! + (a_1 + a_3)x/1! + (a_2 + a_4)x^2/2! + (a_3 + a_5)x^3/3! + (a_4 + a_6)x^4/4! + \dots$$

this means

$$a_0 + a_2 = 0 \qquad \text{and} \qquad a_1 + a_3 = 0 \qquad \text{and} \qquad a_2 + a_4 = 0 \qquad \text{and} \qquad a_3 + a_5 = 0 \qquad \text{etc.}$$

Therefore $a_0 = -a_2 = a_4 = -a_6 = a_8 = \dots$ and $a_1 = -a_3 = a_5 = -a_7 = a_9 = \dots$ so

$$f(x) = a_0(1 - x^2/2! + x^4/4! - x^6/6! + \dots) + a_1(x/1! - x^3/3! + x^5/5! - x^7/7! + \dots).$$

Remembering our Taylor series from calculus, this shows that

$$f(x) = a_0 \cos x + a_1 \sin x.$$

Thus, the linearly independent functions $\cos x$ and $\sin x$ span the vector space V.

These functions are therefore a basis and $\dim V = 2$.

5 Vocabulary

Keywords from today's lecture:

1. **Vector spaces**.

   A vector space is a nonempty set V with two operations called *vector addition* and *scalar multiplication* that formally resemble the operations of vector addition and scalar multiplication for elements of $\mathbb{R}^n$. The precise definition involves a long list of axioms governing these operations, but in practice it's rarely necessary to remember the axioms.

   Example: Any subspace of $\mathbb{R}^n$.

   Example: Given a set X, the set $\mathsf{Fun}(X, \mathbb{R})$ of functions $f : X \to \mathbb{R}$, provided we define $f + g$ as the function with the formula

   $$(f + g)(x) = f(x) + g(x) \qquad \text{for } x \in X$$

   and define cf as the function with the formula

   $$(cf)(x) = cf(x) \qquad \text{for } x \in X$$

   whenever $f, g : X \to \mathbb{R}$ and $c \in \mathbb{R}$.

2. **Subspace** of a vector space.

   A nonempty subset closed under linear combinations.

3. **Linearly combination** and **span** of elements in a vector space.

   A linear combination of a finite set of vectors $v_1, v_2, \dots v_p \in V$ is a vector of the form

   $$c_1v_1 + c_2v_2 + \cdots + c_pv_p$$

   where $c_1, c_2, \dots, c_p \in \mathbb{R}$. A linear combination of an infinite set of vectors is a linear combination of some finite subset. The set of all linear combinations of a set of vectors is the span of the vectors.

4. **Linearly independent** elements in a vector space.

   A list of elements in a vector space is **linearly dependent** if one vector can be expressed as a linear combination of a finite subset of the other vectors. If this is impossible, then the vectors are linearly independent.

   Example: $\cos(x)$ and $\sin(x)$ are linearly independently in $\mathsf{Fun}(\mathbb{R}, \mathbb{R})$.

   Example: the infinite list of functions $1, x, x^2, x^3, x^4, \dots$ are linearly independent in $\mathsf{Fun}(\mathbb{R}, \mathbb{R})$.

5. **Basis** and **dimension** of a vector space.

   A set of linearly independent elements whose span is the entire vector space.

   Every basis in a vector space has the same number of elements. This number is defined to be the **dimension** of the vector space.

6. **Linear functions**.

   If U and V are vector spaces, then a function $f : U \to V$ is linear when

   $$f(u + v) = f(u) + f(v) \qquad \text{and} \qquad f(cv) = cf(v)$$

   for all $u, v \in U$ and $c \in \mathbb{R}$.