TLDR

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

- A vector space is a nonempty set with a "zero vector" and two operations that can be thought of a "vector addition" and "scalar multiplication." The operations must obey several conditions.
- There are notions of subspaces, linear functions, linear combinations, spans, linear independence, and bases for vector spaces. The definitions are essentially the same as for \mathbb{R}^n , with one minor caveat when we are considering linear combinations and independence of infinite sets of vectors.
- Every vector space has a basis, and every basis for a given vector space has the same number of elements, which could be infinite. This number of elements is the *dimension* of the vector space.
- If X and Y are sets, then let $\operatorname{\mathsf{Fun}}(X,Y)$ be the set of functions $f:X\to Y$. The set $\operatorname{\mathsf{Fun}}(X,\mathbb{R})$ is naturally a vector space. If X is finite then $\overline{\dim\operatorname{\mathsf{Fun}}(X,\mathbb{R})}=|X|$.
- If U and V are vector spaces, then let $\mathsf{Lin}(U,V)$ be the set of linear functions $f:U\to V$. The set $\mathsf{Lin}(U,V)$ is naturally a vector space. If $\dim U<\infty$ then $\boxed{\dim \mathsf{Lin}(U,\mathbb{R})=\dim U}$. Moreover, if W is another vector space and $f\in \mathsf{Lin}(V,W)$ and $g\in \mathsf{Lin}(U,V)$, then $f\circ g\in \mathsf{Lin}(U,W)$.
- Suppose $f:U\to V$ is a linear function between vector spaces. Define $\operatorname{range}(f)=\{f(u):u\in U\}\subseteq V \text{ and } \operatorname{kernel}(f)=\{u\in U:f(u)=0\}\subseteq U.$ These sets are subspaces. If $\dim U<\infty$ then $\dim\operatorname{range}(f)+\dim\operatorname{kernel}(f)=\dim U$.
- Let A be an $n \times n$ matrix. Let λ be a number and suppose $0 \neq v \in \mathbb{R}^n$. If $Av = \lambda v$ then we say that v is an eigenvector for A and that λ is an eigenvalue for A. More specifically, v is an eigenvector with eigenvalue λ for A. This happens if and only if $0 \neq v \in \text{Nul}(A - \lambda I_n)$.

For example,
$$v = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$$
 is an eigenvector with eigenvalue $\lambda = -4$ for $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$ since
$$\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix}.$$

The zero vector is not allowed to be an eigenvector, but 0 can occur as an eigenvalue.

- The eigenvalues λ for A are the numbers such that $\det(A \lambda I_n) = 0$.
- The eigenvectors with eigenvalue λ for A are the nonzero elements of Nul $(A \lambda I_n)$.
- If A is a triangular matrix, then its eigenvalues are its diagonal entries.

For example, the eigenvalues of $\left[\begin{array}{ccc} 1 & 6 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{array}\right]$ are 0 and 1.

1 Last time: vector spaces

A (real) vector space V is a set containing a zero vector, denoted 0, with vector addition and scalar multiplication operations that let us produce new vectors $u + v \in V$ and $cv \in V$ from given elements $u, v \in V$ and $c \in \mathbb{R}$. Several conditions must be satisfied so that these operations behave exactly like vector addition and scalar multiplication for \mathbb{R}^n . Most importantly, we require that

- 1. u + v = v + u and (u + v) + w = u + (v + w).
- 2. v-v=0 where we define u-v=u+(-1)v.
- 3. v + 0 = v
- 4. cv = v if c = 1.

There are a few other more conditions to give the full definition (see the notes from last time).

By convention, we refer to elements of vector spaces as *vectors*.

Example. All subspace of \mathbb{R}^n are vector spaces, with the usual zero vector and vector operations.

The set of $m \times n$ matrices is a vector space, with the usual addition and scalar multiplication operations. The zero vector in this vector space is the $m \times n$ zero matrix.

Most vector spaces that we encounter are either subspaces of \mathbb{R}^n or subspaces of the following construction.

Proposition. Let X be a set and let V be a vector space.

Then the set $\operatorname{\mathsf{Fun}}(X,V)$ of all functions $f:X\to V$ is a vector space once we define

$$\begin{split} 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{split}$$

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

Definition. The definitions of a *subspace* of a vector space and of *linear transformations* between vector spaces are identical to the ones we have already seen for subspaces of \mathbb{R}^n :

- A subset $H \subseteq V$ is a subspace if $0 \in H$ and if $u + v \in H$ and $cv \in H$ for all $u, v \in H$ and $c \in \mathbb{R}$.
- A function $f: U \to V$ is linear if f(u+v) = f(u) + f(v) and f(cv) = cf(v) for all $u, v \in U$ and $c \in \mathbb{R}$.

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

Example. If U and V are vector spaces then let Lin(U,V) be the set of linear functions $f:U\to V$.

Then Lin(U, V) is a subspace of Fun(U, V).

Can you make sense of this statement? "Lin($\mathbb{R}^n, \mathbb{R}^m$) is the vector space of $m \times n$ matrices."

Example. A function $f: \mathbb{R} \to \mathbb{R}$ is a polynomial if it has the formula

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

for some nonnegative integer n and some coefficients $a_0, a_1, \ldots, a_n \in \mathbb{R}$.

The set of polynomial functions $\mathbb{R} \to \mathbb{R}$ is a subspace of $\mathsf{Fun}(\mathbb{R}, \mathbb{R})$.

Example. Suppose V is a vector space. Choose $v \in V$. Given a linear function $f: V \to \mathbb{R}$, define

$$v^*(f) = f(v).$$

Then v^* is a linear function $Lin(V, \mathbb{R}) \to \mathbb{R}$.

Let's go deeper: the function with the formula $v \mapsto v^*$ is a linear function $V \to \text{Lin}(\text{Lin}(V, \mathbb{R}), \mathbb{R})$.

If $V = \mathbb{R}^n$ then this function $V \to \text{Lin}(\text{Lin}(V, \mathbb{R}), \mathbb{R})$ is invertible.

Let V be a vector space. The definitions of linear combinations and linear independence for vectors in V are mostly the same as for vectors in \mathbb{R}^n , with one caveat.

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

$$c_1v_1 + c_2v_2 + \dots + c_kv_k$$

for some scalars $c_1, c_2, \ldots, c_k \in \mathbb{R}$.

We must be a little careful when defining linear combinations for infinite sets. Specifically: a *linear* combination of an infinite set of vectors is a linear combination of some **finite** subset of the vectors.

Definition. The *span* of a set of vectors is the set of all linear combinations that can be formed from the vectors. The span of a set of vectors in V is a subspace of V.

Example. The subspace of polynomials in $\operatorname{Fun}(\mathbb{R},\mathbb{R})$ is the span of the set of functions $1,x,x^2,x^3,\ldots$. The infinite sum $e^x=1+x+\frac{1}{2}x+\frac{1}{6}x^2+\frac{1}{24}x^3+\cdots+\frac{1}{n!}x^n+\ldots$ does not belong to this subspace.

Definition. A finite list of vectors $v_1, v_2, \ldots, v_k \in V$ is *linearly independent* if it is impossible to express $0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$ except when $c_1 = c_2 = \cdots = c_k = 0$.

An infinite list of vectors is defined to be *linearly independent* if every finite subset of the vectors 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, \ldots is a basis for V is the same as saying that for each $v \in V$, there a unique coefficients $x_1, x_2, x_3, \cdots \in \mathbb{R}$, all but finitely many of which are zero, such that $v = x_1b_1 + x_2b_2 + x_3b_3 + \ldots$

Theorem. Let V be a vector space.

- 1. V has at least one basis.
- 2. Every basis of V has the same number of elements (but this could be infinite).
- 3. If A is a subset of linearly independent vectors in V then V has a basis B with $A \subseteq B$.
- 4. If C is a subset of vectors in V whose span is V then V has a basis B with $B \subseteq C$.

Definition. The dimension of a vector space V is the number dim V of elements in any of its bases.

Example. If X is a finite set then dim $\operatorname{Fun}(X,\mathbb{R}) = |X|$ where |X| is the size of X.

2 More on dimension

If V is a finite-dimensional vector space then I claim that dim $Lin(V, \mathbb{R}) = \dim V$.

Suppose b_1, b_2, \ldots, b_n is a basis for V.

Then a basis for $Lin(V,\mathbb{R})$ is given by the linear functions $\phi_1,\phi_2,\ldots,\phi_n:V\to\mathbb{R}$ with the formulas

$$\phi_i(x_1b_1 + x_2b_2 + \dots x_nb_n) = x_i \quad \text{for } x_1, x_2, \dots, x_n \in \mathbb{R}.$$

The unique way to express a linear function $f:V\to\mathbb{R}$ as a linear combination of these functions is

$$f = f(b_1)\phi_1 + f(b_2)\phi_2 + \dots + f(b_n)\phi_n.$$

Assume $V = \mathbb{R}^n$. Then we can think of $\text{Lin}(\mathbb{R}^n, \mathbb{R})$ as the vector space of $1 \times n$ matrices.

If $b_1 = e_1, b_2 = e_2, ..., b_n = e_n$ is the standard basis, then $\phi_1 = e_1^T, \phi_2 = e_2^T, ..., \phi_n = e_n^T$.

Definition. Suppose U and V are vector spaces and $f: U \to V$ is a linear function.

Define $\operatorname{range}(f) = \{f(x) : x \in U\} \subseteq V \text{ and } \operatorname{kernel}(f) = \{x \in U : f(x) = 0\} \subseteq U.$

These sets are subspaces which generalize the column space and null space of a matrix.

We have a version of the rank-nullity theorem for arbitrary vector spaces:

Theorem (Rank-Nullity Theorem). If $\dim U < \infty$ then $\dim \mathsf{range}(f) + \dim \mathsf{kernel}(f) = \dim U$.

This specializes to our earlier statement about matrices when $U = \mathbb{R}^n$ and $V = \mathbb{R}^m$.

We can prove the theorem in a self-contained, completely abstract way, but it's a little involved.

Proof. If b_1, b_2, \ldots, b_n is a basis for U then the span of $f(b_1), f(b_2), \ldots, f(b_n)$ must be equal to $\mathsf{range}(f)$.

Therefore $\dim \operatorname{range}(f) \leq \dim U < \infty$. Since $\operatorname{kernel}(f) \subseteq U$, we also have $\dim \operatorname{kernel}(f) < \infty$.

Let $k = \dim \mathsf{range}(f)$ and $l = \dim \mathsf{kernel}(f)$.

Choose $u_1, u_2, \ldots, u_k \in U$ such that $f(u_1), f(u_2), \ldots, f(u_k)$ is a basis for range(f).

Choose a basis v_1, v_2, \ldots, v_l for kernel(f). We will check that $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_l$ is a basis for U.

To show linear independence, suppose $a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_l \in \mathbb{R}$ are such that

$$a_1u_1 + a_2u_2 + \dots + a_ku_k + b_1v_1 + b_2v_2 + \dots + b_lv_l = 0.$$

Applying f to both sides gives $a_1 f(u_1) + a_2 f(u_2) + \cdots + a_k f(u_k) = 0$, so $a_1 = a_2 = \cdots = a_k = 0$.

But this implies $b_1v_1 + b_2v_2 + \cdots + b_lv_l = 0$, so we also have $b_1 = b_2 = \cdots = b_l = 0$.

Our vectors $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_l$ are therefore linearly independent in U.

Now let $x \in U$. By assumption $f(x) = c_1 f(u_1) + c_2 f(u_2) + \dots + c_k f(u_k)$ for some $c_1, c_2, \dots, c_k \in \mathbb{R}$.

The vector $x - c_1u_1 - c_2u_2 - \cdots - c_ku_k$ is then in the span of v_1, v_2, \ldots, v_l since it belongs to kernel(f).

We conclude that x is a linear combination of $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_l$, so this is a basis for U.

3 Eigenvectors and eigenvalues

We return to the concrete setting of \mathbb{R}^n and its subspaces. Let A be a square $n \times n$ matrix.

Definition. An eigenvector of A is a nonzero vector $v \in \mathbb{R}^n$ such that

$$Av = \lambda v$$

for a number $\lambda \in \mathbb{R}$. (λ is the Greek letter "lambda.")

The number λ is called the *eigenvalue* of A for the eigenvector v.

We require eigenvectors to be nonzero because if v=0 then $Av=\lambda v=0$ for all numbers $\lambda\in\mathbb{R}$.

The number 0 is allowed to be an eigenvalue of A, however.

Example. If we are given A and v, it is easy to check whether v is an eigenvector: just compute Av.

For example, if
$$A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$$
 and $v = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ then $Av = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4v$.

Therefore v is an eigenvector of A with eigenvalue -4.

Example. What are the eigenvectors of the matrix $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$?

If $v \in \mathbb{R}^4$ were an eigenvector with eigenvalue λ then

$$Av = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} v_2 \\ v_3 \\ v_4 \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix}.$$

The last equation implies that $0 = \lambda v_4$ and $v_4 = \lambda v_3$ and $v_3 = \lambda v_2$ and $v_2 = \lambda v_1$. In other words,

$$0 = \lambda v_4 = \lambda^2 v_3 = \lambda^3 v_2 = \lambda^4 v_1.$$

If $\lambda \neq 0$ then this would mean that $v_1 = v_2 = v_3 = v_4 = 0$, but remember that v should be nonzero. Therefore the only possible eigenvalue of A is $\lambda = 0$. The eigenvectors of A with eigenvalue 0 are

$$v = \begin{bmatrix} v_1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{where } v_1 \neq 0.$$

To say that " λ is an eigenvalue of A" means that there exists a nonzero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$. Recall that I_n denotes the $n \times n$ identity matrix. We abbreviate by setting $I = I_n$.

Proposition. A number $\lambda \in \mathbb{R}$ is an eigenvalue of A if and only if $A - \lambda I$ is not invertible.

Proof. The equation $Ax = \lambda x$ has a nonzero solution $x \in \mathbb{R}^n$ if and only if $(A - \lambda I)x = 0$ has a nonzero solution, which occurs if and only if $\text{Nul}(A - \lambda I) \neq \{0\}$, or equivalently when $A - \lambda I$ is not invertible. \square

Example. If
$$A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$$
 then

$$A-7I = \left[\begin{array}{cc} 1 & 6 \\ 5 & 2 \end{array}\right] - \left[\begin{array}{cc} 7 & 0 \\ 0 & 7 \end{array}\right] = \left[\begin{array}{cc} -6 & 6 \\ 5 & -5 \end{array}\right] \sim \left[\begin{array}{cc} 1 & -1 \\ 1 & -1 \end{array}\right] \sim \left[\begin{array}{cc} 1 & -1 \\ 0 & 0 \end{array}\right] = \mathsf{RREF}(A-7I).$$

Since $RREF(A-7I) \neq I$, the matrix A-7I is not invertible so 7 is an eigenvalue of A.

Corollary. A number $\lambda \in \mathbb{R}$ is an eigenvalue of A if and only if $\det(A - \lambda I) = 0$.

Proof. Remember that $A - \lambda I$ is not invertible if and only if $\det(A - \lambda I) = 0$.

Another way of defining an eigenvector: the eigenvectors of A with eigenvalue λ are precisely the nonzero elements of the null space $\text{Nul}(A - \lambda I)$. Since we know how to construct a basis for the null space of any matrix, we also know how to find all eigenvectors of a matrix for any given eigenvalue.

Example. In the previous example, $\mathsf{RREF}(A-7I) = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ so Ax = 7x if and only if (A-7I)x = 0

if and only if $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where $x_1 - x_2 = 0$. In this linear system, x_2 is a free variable, and we can rewrite x as $x = \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This means $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a basis for Nul(A - 7I).

Therefore every eigenvector of A with eigenvalue 7 has the form $\begin{bmatrix} a \\ a \end{bmatrix}$ for some $a \in \mathbb{R}$.

One calls the set of all $v \in \mathbb{R}^n$ with $Av = \lambda v$ the eigenspace of A for λ . We also call this the λ -eigenspace of A. Note that this is just the null space of $A - \lambda I$. A number is an eigenvalue of A if and only if the corresponding eigenspace is nonzero (that is, contains a nonzero vector).

Example. Suppose we were told that $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$ has 2 as an eigenvalue.

To find a basis for the 2-eigenspace of A, we row reduce

$$A - 2I = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1/2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathsf{RREF}(A - 2I).$$

Thus Ax = 2x if and only if $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1 - \frac{1}{2}x_2 + 3x_3 = 0$, that is, if and only if

$$x = \begin{bmatrix} \frac{1}{2}x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

The vectors $\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$ are then a basis for the 2-eigenspance of A.

Recall that a matrix is *triangular* if its nonzero entries all appear on or above the main diagonal, or all appear on or below the main diagonal.

Theorem. The eigenvalues of a triangular square matrix A are its diagonal entries.

Proof. If A has diagonal entries d_1, d_2, \ldots, d_n then $A - \lambda I$ is triangular with diagonal entries $d_1 - \lambda, d_2 - \lambda, \ldots, d_n - \lambda$, so $\det(A - \lambda I) = (d_1 - \lambda)(d_2 - \lambda) \cdots (d_n - \lambda)$ which is zero if and only if $\lambda \in \{d_1, d_2, \ldots, d_n\}$. \square

Example. The eigenvalues of the matrix $\begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ are 3, 0, and 2.

4 Vocabulary

Keywords from today's lecture:

1. Subspace of a vector space.

A nonempty subset closed under linear combinations.

2. 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

$$c_1v_1 + c_2v_2 + \cdots + c_nv_n$$

where $c_1, c_2, \ldots, c_p \in \mathbb{R}$. 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.

3. 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 $\operatorname{\mathsf{Fun}}(\mathbb{R},\mathbb{R})$.

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

4. **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.

5. Linear functions.

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

$$f(u+v) = f(u) + f(v)$$
 and $f(cv) = cf(v)$

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

6. **Eigenvector** for an $n \times n$ matrix A.

A nonzero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$ for some real number $\lambda \in \mathbb{R}$.

The number λ is the **eigenvalue** of A for v.

$$\begin{bmatrix} 1\\1\\1 \end{bmatrix} \text{ is an eigenvector for } \begin{bmatrix} 0 & 2 & 0\\2 & 0 & 0\\0 & 0 & 2 \end{bmatrix} \text{ with eigenvalue 2 as } \begin{bmatrix} 0 & 2 & 0\\2 & 0 & 0\\0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1\\1\\1 \end{bmatrix} = \begin{bmatrix} 2\\2\\2 \end{bmatrix}.$$

7. λ -eigenspace for an $n \times n$ matrix A, where $\lambda \in \mathbb{R}$.

The subspace $\operatorname{Nul}(A - \lambda I) \subseteq \mathbb{R}^n$ where I is the $n \times n$ identity matrix.

If λ is not an eigenvalue of A, then this subspace is $\{0\}$.

But if λ is an eigenvalue of A, then the subspace is nonzero.