LECTURE 15

Eigenvalues

15.1 Invariant subspaces

Let $f \in \mathcal{L}(V)$ and let U denote a subspace of V. A natural question to ask is what set f will map U to. Is there any relationship that we can establish between $f(U)$ and U, e.g., $f(U) = U$ or $f(U) \subset U$? The following example shows that, in general, we cannot expect to be able to make a general statement in this respect.

Example 15.1 Let $V = \mathbb{R}^2$ and let

$$f(\mathbf{v}) = A\mathbf{v}, \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Define $U_1 = \operatorname{span}\{\mathbf{e}_1\}, U_2 = \operatorname{span}\{\mathbf{e}_2\}$. Then

$$f(\mathbf{e}_1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad f(\mathbf{e}_2) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Hence, $f(U_1) = U_1$, but $f(U_2) \neq U_2$; in fact, $f(U_2) \cap U_2 = \{\mathbf{0}\}$.

This example shows that the relationship $f(U) \subseteq U$ is special. This observation justifies the following definition.

Definition 15.1 — Invariant subspace. Let $V(\mathbb{F})$ be a non-trivial vector space and let $f \in \mathcal{L}(V)$. We say a subspace U is invariant under f, or f-invariant, if $f(U) \subseteq U$, i.e., $f(\mathbf{u}) \in U$ for all $\mathbf{u} \in U$.

An immediate consequence of the invariance property is that we can restrict our study of f to its action on invariant subspaces.

Proposition 15.1 Let $f \in \mathcal{L}(V)$ and let $U \leq V$ be an f-invariant subspace. Let $\tilde{f}$ be defined via $\tilde{f}(\mathbf{u}) = f(\mathbf{u})$ for all $\mathbf{u} \in U$. Then $\tilde{f} \in \mathcal{L}(U)$.

The linear map $\tilde{f}$ in the above proposition is called the restriction of f to U and is denoted by $\tilde{f} = f\,|_U$.

Some subspaces are invariant under any $f \in \mathcal{L}(V)$.

Exercise 15.1 Show that $U \leq V$ is invariant under $f \in \mathcal{L}(V)$ for the following choices of U:
1. $U = \{\mathbf{0}\}$;
2. $U = V$;
3. $U = \ker f$;
4. $U = \text{im } f$.

We would like to identify other invariant subspaces. Of particular interest will be one-dimensional invariant subspaces.

Motivated by Example 16.2, let us examine the invariance property for the following simple choice of subspace: let $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ and let $U = \text{span}\{\mathbf{v}\}$. Then any non-zero $\mathbf{u} \in U$ has the form $\mathbf{u} = a\mathbf{v}$ for some $a \in \mathbb{F} \setminus \{0\}$. Invariance would then require that $f(\mathbf{u}) \in U$, i.e., $f(a\mathbf{v}) = b\mathbf{v}$, for some $b \in \mathbb{F}$. Using the linearity of f we find

$$f(a\mathbf{v}) = b\mathbf{v} \iff af(\mathbf{v}) = b\mathbf{v} \iff f(\mathbf{v}) = \lambda\mathbf{v}, \qquad \lambda \in \mathbb{F}.$$

Hence, if there exists $\mathbf{v} \in V$ such that the above relation holds, then $U = \text{span}\{\mathbf{v}\}$ is f-invariant. This result justifies the following definition.

Definition 15.2 Let $f \in \mathcal{L}(V)$, where $V(\mathbb{F})$ is a finite-dimensional vector space. Let $\lambda \in \mathbb{F}, \mathbf{v} \in V \setminus \{\mathbf{0}\}$ satisfy the equation

$$f(\mathbf{v}) = \lambda\mathbf{v}. \tag{15.1}$$

Then
- λ is called an **eigenvalue of** f;
- $\mathbf{v}$ is called **an eigenvector of** f;
- the pair $(\lambda, \mathbf{v})$ will be referred to as an **eigenpair of** f or an **eigensolution of** 15.1;
- the set of eigenvalues is called the **spectrum of** f: $\text{sp} f := \{\lambda \in \mathbb{F} : f(\mathbf{v}) = \lambda\mathbf{v}, \mathbf{v} \in V\}$;
- equation 15.1 is called the **eigenvalue equation**.

We emphasise the following facts ensuing from, or stated in, the definition:
- eigenvectors are, by definition, non-zero vectors; without this restriction, any scalar $a \in \mathbb{F}$ would be an eigenvalue, since $f(\mathbf{0}) = \mathbf{0} = a \cdot \mathbf{0}$;
- eigenvectors are unique up to a multiplicative constant: if $\mathbf{v}$ is an eigenvector, then $a\mathbf{v}$ is also an eigenvector, for any $a \in \mathbb{F}$, since $f(a\mathbf{v}) = af(\mathbf{v}) = a(\lambda\mathbf{v}) = \lambda(a\mathbf{v})$;
- an eigenvector is associated with a *single* eigenvalue, since $A\mathbf{v} = \lambda\mathbf{v} = \lambda'\mathbf{v} \implies (\lambda - \lambda')\mathbf{v} = \mathbf{0} \implies \lambda = \lambda'$, given that $\mathbf{v} \neq \mathbf{0}$;
- the eigenvalue equation is nonlinear, since the unknowns λ and $\mathbf{v}$ are multiplied together on the right; this means that we expect $\lambda = \lambda(\mathbf{v})$ and $\mathbf{v} = \mathbf{v}(\lambda)$: this explains the term **eigenpair**.

Before we investigate further this type of f-invariance, we need to establish if the eigenvalue equation for f has any solutions. The answer will depend on the choice of field $\mathbb{F}$.

15.2 The eigenvalue problem: matrix formulations

Let us consider the matrix representation of 15.1. The term on the right-hand side can be written as $\lambda\mathbf{v} = (\lambda id_{\mathrm{V}})(\mathbf{v})$, i.e., a multiple of the identity. Then the matrix representation of 15.1 is

$$A\mathbf{x} = (\lambda I)\mathbf{x} = \lambda\mathbf{x}, \tag{15.2}$$

where $A \in \mathbb{F}^{n \times n}$ is the matrix representation of f and $\mathbf{x} = \varphi_V^{-1}(\mathbf{v}) \in \mathbb{F}^n$, with $\lambda \in \mathbb{F}$. We will study this representation, in order to establish results for the eigenvalue problem 15.1.

Let us re-write 15.2 as follows:

$$A\mathbf{x} = \lambda I\mathbf{x} \iff (\lambda I - A)\mathbf{x} = \mathbf{0}.$$

The above formulation suggests two distinct approaches:

1. Assuming λ is known, one can view this relation as a linear system of equations with coefficient matrix $\lambda I - A$, with zero right hand side. Hence, a non-trivial solution $\mathbf{x} \in \mathbb{F}^n \setminus \{\mathbf{0}\}$ exists if and only if $\det(\lambda I - A) = 0$. This is an eigenvalue-only characterisation! While it seems awkward, it allows us to analyse the existence and properties of eigenvalues independently of the corresponding eigenvectors.
2. Assuming λ is known, one can view eigenvectors as being elements in a certain kernel: $\mathbf{x} \in \ker(\lambda I - A)$. Therefore, the study of such subspaces is useful in providing a description of eigenvectors.

These observations indicate that the study (and computation) of eigenvalues should precede that of eigenvectors. Indeed, this is our approach below.

15.3 Eigenvalues

15.3.1 The characteristic polynomial

The choice of field in our discussion is either $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$.

Definition 15.3 Let $A \in \mathbb{F}^{n\times n}$. The characteristic polynomial of A is the polynomial $p_A \in \mathcal{P}_n(\mathbb{F})$ defined via

$$p_A(t) = \det(tI - A).$$

We note that, by the properties/definition of the determinant, the polynomial p_A is indeed a polyonimal of degree n, which is also **monic**, i.e., it leading coefficient equal to one; its general form is included below:

$$p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{n-1} t^{n-1} + t^n.$$

With this notation in place, the eigenvalue problem (15.2) is equivalent to the following problem:

$$\text{find } \lambda \in \mathbb{F} \text{ such that } p_A(\lambda) = 0.$$

Hence, the eigenvalues of A are the roots (or zeros) of the charateristic polynomial of A.

R If $\mathbb{F} = \mathbb{C}$, then $p_A \in \mathcal{P}_n(\mathbb{C})$, while if $\mathbb{F} = \mathbb{R}$, then $p_A \in \mathcal{P}_n(\mathbb{R})$. In other words, if A is a complex (real) matrix, then the coefficients of the characteristic polynomial p_A are complex (real). This distinction is important, as outlined in the discussion below.

15.3.2 Eigenvalues of complex matrices

Let $A \in \mathbb{C}^{n\times n}$. Then $p_A \in \mathcal{P}_n(\mathbb{C})$. Let us consider the existence of roots of p_A. We recall the following fundamental results for polynomials with complex coefficients.

Theorem 15.2 — Fundamental Theorem of Algebra. Every polynomial $p \in \mathcal{P}_n(\mathbb{C})$ has a zero in $\mathbb{C}$ for any $n \in \mathbb{N}$.

By this result, we can write any polynomial of generic degree n, say $p_n(t)$, in the form

$$p_n(t) = (t - z)p_{n-1}(t),$$

where $z \in \mathbb{C}$ and $p_{n-1}(t)$ is a polynomial of degree $n-1$. Applying the theorem again to $p_{n-1}(t)$, we obtain a further factorisation of $p_n(t)$. This observation leads to the following result.

Proposition 15.3 Let $p \in \mathcal{P}_n(\mathbb{C})$, with $n \in \mathbb{N}$. Then p can be expressed uniquely as the product of n monic linear polynomials $t - z_j$:

$$p(t) = a_n(t - z_1)(t - z_2)\cdots(t - z_n),$$

where $a_n, z_1, \dots, z_n \in \mathbb{C}$.

This result confirms that the polynomial equation $p(t) = 0$, has n zeros in $\mathbb{C}$. We immediately deduce the following result concerning eigenvalues.

Proposition 15.4 Let $A \in \mathbb{C}^{n\times n}$. Then A has n complex eigenvalues, given by the n roots of $p_A(t)$. In particular, the form of the characteristic polynomial is

$$p_A(t) = (t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n),$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$.

15.3.3 Eigenvalues of real matrices

Let $A \in \mathbb{R}^{n\times n}$. Then $p_A \in \mathcal{P}_n(\mathbb{R})$. In this case, a similar result to that of Theorem 15.2 does not hold: polynomials with real coefficients do not always have real roots: a simple counter-example is provided by the polynomial $p(t) = t^2 + 1$. In this case, a result that provides a description of the roots of p_A is provided by the factorisation of polynomials with real coefficients. First, we note the following type of quadratic polynomial with real coefficients.

Definition 15.4 A monic quadratic polynomial with real coefficients $p \in \mathcal{P}_2(\mathbb{R})$ is called irreducible if it has the form

$$p(t) = (t - z)(t - \bar{z}) = t - 2\mathrm{Re}z\ t + |z|^2,$$

with $z \in \mathbb{C}$ and $\mathrm{Im}(z) \neq 0$.

R The term *irreducible* refers to the fact that a monic polynomial $p \in \mathcal{P}_2(\mathbb{R})$ can only be factorised as the product of two monic polynomials in $\mathcal{P}_1(\mathbb{C})$ and not in $\mathcal{P}_1(\mathbb{R})$.

The main result regarding factorisation of polynomials over $\mathbb{R}$ is included below.

Proposition 15.5 Let $p \in \mathcal{P}_n(\mathbb{R})$, with $n \in \mathbb{N}$. Then p can be expressed uniquely as the product of monic linear and irreducible quadratic polynomials with real coefficients in the form

$$p(t) = a_n(t - x_1)(t - x_2)\cdots(t - x_m)(t^2 + b_1 t + c_1)(t^2 + b_2 t + c_2)\cdots(t^2 + b_k t + c_k),$$

where $a_n, x_i, b_j, c_j \in \mathbb{R}$, $i = 1, \ldots, m$ and $j = 1, \ldots, k$, with $n = m + 2k$.

Hence, a polynomial of degree n with real coefficients has
- $0 \leq m \leq n$ real roots;
- k pairs of complex-conjugate roots, with $0 \leq k \leq \frac{n-m}{2}$.

The above representation allows us to provide a description (albeit theoretical) of the real eigenvalues of a real matrix A.

Proposition 15.6 Let $A \in \mathbb{R}^{n\times n}$. Then A has a number $0 \leq m \leq n$ of real eigenvalues equal to the number of linear factors in the factorisation of its characteristic polynomial p_A.

R Note that the case real matrices $A \in \mathbb{R}^{n\times n}$ corresponds to endomorphisms f defined on real vector spaces. Since we are interested in the invariance of one-dimensional subspaces of a real vector space $V(\mathbb{R})$, we cannot take into account the complex roots of the characteristic polynomial of A, as the scalars arising in the invariance property (i.e., the eigenvalues) need to be real.

We summarise the above discussion as follows:
- We have at most n roots of p_A, hence at most n eigenvalues, possibly non-distinct.
- If $\mathbb{F} = \mathbb{C}$, then we will always have n eigenvalues in $\mathbb{C}$, some possibly real.

- If $\mathbb{F} = \mathbb{R}$, then we are not guaranteed to have n eigenvalues in $\mathbb{R}$.

This means that if we pose the eigenvalue problem over $\mathbb{C}$, then we are guaranteed to have n eigenvalues in $\mathbb{C}$, possibly with some in $\mathbb{R}$. For this reason, we reformulate the eigenvalue problem for $f \in \mathcal{L}(V)$ as follows:

$$\text{find } \lambda \in \mathbb{C} \text{ and } \mathbf{v} \in \mathbb{C}^n \text{ such that } f(\mathbf{v}) = \lambda \mathbf{v}.$$

We note that in the above formulation we necessarily have $V = V(\mathbb{C})$.

This formulation leads to the obvious question: how do we deal with the case of real linear maps? By the above results, we could have real maps that yield characteristic polynomials with

- no real roots (if n is even, this is a possibility); in this case, f has no eigensolutions over $\mathbb{R}$ and hence no invariant subspaces of dimension one;
- n real roots: in this case, we could pose the problem over $\mathbb{F} = \mathbb{R}$: this would be natural and also efficient, in practice. We know that such maps exist: an example is given by the identity map, which has n real eigenvalues, all equal to one. Later, we will identify other real linear maps with n real eigenvalues.

One way of reconciling the real case with the above complex formulation is to note that any real map induces a complex map: $f \leftrightarrow f + i \cdot o$, where o is the zero map over $\mathbb{C}$. We can then seek eigensolutions over $\mathbb{C}$: if all the eigenvalues turn out to be real, we can also choose the eigenvectors to be real and reformulate the eigenvalue problem over $\mathbb{R}$, *a posteriori*.

15.3.4 Algebraic multiplicity

If the characteristic polynomial has repeated roots, then its general form over $\mathbb{C}$ is given by

$$p(t) = a_n (t - \lambda_1)^{\alpha_1} (t - \lambda_2)^{\alpha_2} \cdots (t - \lambda_\ell)^{\alpha_\ell},$$

where $\alpha_1 + \alpha_2 + \cdots + \alpha_\ell = n$. Let us define the map $\alpha : \mathrm{sp} f \mapsto \{1, 2, \ldots, n\}$ via $\alpha(\lambda_k) = \alpha_k$.

Definition 15.5 Let V be an n-dimensional vector space and let $f \in \mathcal{L}(V)$. The algebraic multiplicity of an eigenvalue λ_k of f is $\alpha(\lambda_k) = \alpha_k$.

Note that we can indeed have $\alpha(\lambda) = n$: this is the case of the eigenvalues of the identity map, which are all equal (to one); in fact, any map $f = c \cdot id$ for some $c \in \mathbb{F}$ will also satisfy this property. Note also that if $\alpha(\lambda) = 1$ for all $\lambda \in \mathrm{sp} f$, then we have n distinct eigenvalues and therefore n distinct eigenvectors. If $\alpha(\lambda) > 1$ for some λ, i.e., at least one eigenvalue is a repeated eigenvalue, then there may be situations where the total number of eigenvectors is less than n: see Example 16.2. This is an important distinction which we will discuss in the next lecture.

LECTURE 16

Eigenvectors

We consider again the eigenvalue equation

$$f(\mathbf{v}) = \lambda\mathbf{v},$$

where $f \in \mathcal{L}(V)$, $\lambda \in \mathbb{F}$. Let us recall also its matrix formulation in the two equivalent forms considered previously:

$$A\mathbf{x} = \lambda\mathbf{x} \iff (\lambda I - A)\mathbf{x} = \mathbf{0}.$$

16.1 Eigenspaces

The first observation is that $\mathbf{v}$ belongs to a certain subspace of V.

Proposition 16.1 Let $(\lambda, \mathbf{v})$ be an eigenpair of f. Then $\mathbf{v} \in E_\lambda := \ker(f - \lambda id_{\mathrm{V}})$. Moreover, $E_\lambda \leq V$.

Proof. Since $\lambda\mathbf{v} = (\lambda id_{\mathrm{V}})\mathbf{v}$, we find that

$$f(\mathbf{v}) - (\lambda id_{\mathrm{V}})(\mathbf{v}) = \mathbf{0} \iff (f - \lambda id_{\mathrm{V}})(\mathbf{v}) = \mathbf{0} \iff \mathbf{v} \in \ker(f - \lambda id_{\mathrm{V}}).$$

Finally, since$f, id_{\mathrm{V}} \in \mathcal{L}(V)$, the map $f - \lambda id_{\mathrm{V}} \in V$, by closure in $\mathcal{L}(V)$. The result then follows, as the kernel of an endomorphism on V is a subspace of V. ∎

The subspace property in the previous result suggests the next definition.

Definition 16.1 — Eigenspace. The subspace E_λ is the eigenspace of f associated with eigenvalue λ.

Note that for any λ, E_λ is non-trivial, since it contains at least one non-zero vector: an eigenvector associated with λ. This means that $\dim E_\lambda \geq 1$. Let us derive further properties of eigenspaces.

Proposition 16.2 Let $(\lambda, \mathbf{v}), (\lambda', \mathbf{v}')$ denote two distinct eigenpairs. Then $E_\lambda \cap E_{\lambda'} = \{\mathbf{0}_{\mathrm{V}}\}$.

Proof. Assume, for a contradiction, that there exists a nonzero $\mathbf{u}$ such that $\mathbf{u} \in E_\lambda \cap E_{\lambda'}$. Then

$$f(\mathbf{u}) = \lambda\mathbf{u} = \lambda'\mathbf{u} \implies (\lambda - \lambda')\mathbf{u} = \mathbf{0}_{\mathrm{V}} \implies \lambda = \lambda',$$

which is the contradiction we sought. ∎

We immediately obtain the following corollary.