Eigenvalue decomposition

17.1 Canonical forms

In Lecture 14 we introduced the concepts of similarity and diagonalisation. Recall that a square matrix A is said to be diagonalisable if it is similar to a diagonal matrix D:

$$A = MDM^{-1} \iff D = M^{-1}AM,$$

for some invertible M. In this case, D represents the canonical form of the equivalence class of matrices similar to A. We say that the associated endomorphisms are diagonalisable. Results on diagonalisability hold in general when $\mathbb{F} = \mathbb{C}$. Hence, we consider this case first; we will discuss the choice $\mathbb{F} = \mathbb{R}$ at the end of the lecture.

The following result provides a characterisation of diagonalisation of endomorphisms $f \in \mathcal{L}(V(\mathbb{C}))$, and therefore equivalently of square matrices $A \in \mathbb{C}^{n \times n}$.

> **Proposition 17.1** Let $V(\mathbb{C})$ be an n-dimensional vector space with basis $B_V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. An endomorphism $f : V \to V$ has a diagonal matrix representation relative to B_V if and only $\mathbf{v}_j$ is an eigenvector for f for all $j = 1, \ldots, n$.

Proof. $\boxed{\Longrightarrow}$ Assume that f has a diagonal matrix representation $D \in \mathbb{C}^{n \times n}$ with diagonal entries λ_i. Equivalently, we have $[D]_{ij} = \lambda_i \delta_{ij}$ for $i, j = 1, \ldots, n$. Consider

$$f(\mathbf{v}_j) = \sum_{i=1}^{n} y_i \mathbf{v}_i.$$

By Proposition 10.1, $\mathbf{y} = D\mathbf{x}$, where $\mathbf{x} = \mathbf{e}_j$, given that we are only evaluating the jth vector in the basis B_V. We find

$$\mathbf{y} = D\mathbf{e}_j = \mathbf{c}_j(D) = \lambda_j \mathbf{e}_j \implies y_i = \lambda_j \delta_{ij}.$$

Therefore,

$$f(\mathbf{v}_j) = \sum_{i=1}^{n} y_i \mathbf{v}_i = \sum_{i=1}^{n} \lambda_j \delta_{ij} \mathbf{v}_i = \lambda_j \mathbf{v}_j,$$

which indicates that $\mathbf{v}_j$ is an eigenvector of f, with corresponding eigenvalue λ_j.

$\boxed{\Longleftarrow}$ Assume now that A is the matrix representation relative to a basis of eigenvectors B_V. Noting that $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$, by definition of matrix representation, we get

$$f(\mathbf{v}_i) = \sum_{j=1}^{n} a_{ji} \mathbf{v}_j.$$

On the other hand,

$$f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i = \sum_{j=1}^{n} \lambda_i \delta_{ji} \mathbf{v}_j.$$

By uniqueness of representation of $f(\mathbf{v}_i)$ in the basis B_V, we find $a_{ji} = \lambda_i \delta_{ji}$ for all j, which implies that A is a diagonal matrix with the eigenvalues as diagonal entries. $\blacksquare$

(R) Note that f does not have to be invertible in order for its matrix representation to have a diagonal canonical form.

This result implies that the only situation where we have a canonical form for an endomorphism defined on an n-dimensional vector spaces corresponds to the case where we have n eigenvectors. In turn, this implies that we require

$$\sum_{k=1}^{n} \gamma(\lambda_k) = n.$$

This observation justifies the following terminology.

> **Definition 17.1 — Defective matrix.** We say a matrix $A \in \mathbb{C}^{n \times n}$ is defective if it does not have n linearly independent eigenvectors.

It is important to realise that defective matrices arise 'naturally' in applications and the lack of a complete set of eigenvectors cannot be avoided by a change of basis. In fact, using this very concept, we can see that the matrix in Example 16.2 can be viewed as a canonical form for an equivalence class of 2×2 matrices which have a single eigenvector associated with the repeated eigenvalue $\lambda = 1$.

The following results are obvious corollaries.

> **Corollary 17.2** Defective matrices are not diagonalisable.

> **Corollary 17.3** If the eigenvalues of an endomorphism f are distinct, then f is diagonalisable over $\mathbb{C}$.

We end this subsection with the general statement of diagonalisability of square matrices, which can be seen as a corollary of the above results.

> **Theorem 17.4 — Eigenvalue decomposition.** Let $A \in \mathbb{C}^{n \times n}$ be a square, non-defective matrix. Then A is diagonalisable with canonical form $D = X^{-1}AX$, where
>
> $$D_{ii} = \lambda_i \in \mathbb{C}, \quad \mathbf{c}_i(X) = \mathbf{x}_i \in \mathbb{C}^n, \quad i = 1, 2, \ldots, n,$$
>
> where $\mathbf{x}_i$ are the eigenvectors of A.

(R) Note that for general real matrices we are not guaranteed to have a diagonal canonical form over $\mathbb{R}$ (i.e., a diagonal matrix with real entries), even if the eigenvalues are distinct. This is due to the fact that the characteristic polynomial of a general real matrix cannot be factorised over $\mathbb{R}$, i.e., it cannot be written as a product of real linear polynomials only (see Proposition 15.5).

17.2 Invariant subspaces

We started our discussion of eigenvalues and eigenvectors motivated by the concept of subspace invariance for the case of one-dimensional subspaces. It is therefore fitting to round up our discussion by revisiting this concept and considering the case of subspaces of dimension greater than one.

What is the consequence of invariance of a k-dimensional subspace with regard to matrix representations? To answer this, let us consider the simple case of a two-dimensional f-invariant subspace $U = \mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of an n-dimensional vector space V. Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ be a basis for V. Then

$$
\begin{cases}
f(\mathbf{u}_1) & = a_{11}\mathbf{u}_1 + a_{21}\mathbf{u}_2 \\
f(\mathbf{u}_2) & = a_{12}\mathbf{u}_1 + a_{22}\mathbf{u}_2 \\
f(\mathbf{v}_3) & = b_{13}\mathbf{u}_1 + b_{23}\mathbf{u}_2 + b_{33}\mathbf{v}_3 + \cdots + b_{n3}\mathbf{v}_n \\
& \cdots \\
f(\mathbf{v}_n) & = b_{1n}\mathbf{u}_1 + b_{2n}\mathbf{u}_2 + b_{3n}\mathbf{v}_3 + \cdots + b_{nn}\mathbf{v}_n,
\end{cases}
$$

so that the matrix representation in the basis B is

$$
A = \begin{bmatrix}
a_{11} & a_{12} & b_{13} & \cdots & \cdots & b_{1n} \\
a_{21} & a_{22} & b_{23} & \cdots & \cdots & b_{2n} \\
0 & 0 & b_{33} & \cdots & \cdots & b_{3n} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & b_{n3} & \cdots & \cdots & b_{nn}
\end{bmatrix} = \begin{bmatrix} A_U & B \\ O & C \end{bmatrix},
$$

where A_U is the matrix representation of $\tilde{f} = f\,|_U$ and O is the $(n-2) \times 2$ zero matrix. Thus, the matrix representation of a linear map with a two-dimensional invariant subspace has a block upper-triangular form. This is just a special case of a more general result on invariance which we state below.

Proposition 17.5 Let $f \in \mathcal{L}(V)$ and let $U \leq V$, with $\dim U = k, \dim V = n$. Then U is f-invariant if and only if it has a block triangular matrix representation of the form

$$
A = \begin{bmatrix} A_U & B \\ O & C \end{bmatrix},
$$

where $A_U \in \mathbb{C}^{k \times k}, O \in \mathbb{C}^{(n-k) \times k}, B \in \mathbb{C}^{k \times (n-k)}, C \in \mathbb{C}^{(n-k) \times (n-k)}$.

The following exercise is a corollary of the above result.

Exercise 17.1 Let $f \in \mathcal{L}(V)$ and assume $V = U \oplus W$, with U, W f-invariant subspaces of V. Show that f has a block-diagonal matrix representation:

$$
A = \begin{bmatrix} A_U & O \\ O & A_W \end{bmatrix},
$$

where A_U, A_W are square matrices of dimensions summing up to $\dim V$.

(R) Block diagonal matrices are commonly written as **matrix direct sums** $A = A_U \oplus A_W$.

We can use Proposition 17.5 to construct a canonical form for an endomorphism $f \in \mathcal{L}(V)$ over $\mathbb{R}$. Let us assume first that f has a non-defective matrix representation A. For ease of argument, let us assume that the eigenvalues are distinct. By Proposition 15.5, we have in general m real eigenvalues, with $k = (n-m)/2$ complex-conjugate roots. We first need the following result.

Lemma 17.6 Let $(\lambda, \mathbf{v})$ be a complex eigenpair of A, where

$$\lambda = a + ib, \quad \mathbf{v} = \mathbf{u}_1 + i\mathbf{u}_2,$$

with $a, b \in \mathbb{R}$ with $b \neq 0$ and $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a linearly independent set in $\mathbb{R}^n$. Moreover,

$$A[\mathbf{u}_1 \ \mathbf{u}_2] = [\mathbf{u}_1 \ \mathbf{u}_2] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

Proof. First, note that if $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a linearly dependent set in $\mathbb{R}^n$, then $\mathbf{u}_1 = c\mathbf{u}_2$, for some $c \in \mathbb{R}$; since $\mathbf{v} = \mathbf{u}_1 + i\mathbf{u}_2 = (1 + ic)\mathbf{u}_1$ is an eigenvector, the vector $\tilde{\mathbf{v}} = \mathbf{u}_1 \in \mathbb{R}^n$ is also one. This results in $A\mathbf{u}_1 = (a + ib)\mathbf{u}_1$; comparing imaginary parts, we obtain $b = 0$, a contradiction. Hence, $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a linearly independent set in $\mathbb{R}^n$.

To derive the second statement, we note that

$$A\mathbf{v} = \lambda\mathbf{v} \iff A(\mathbf{u}_1 + i\mathbf{u}_2) = (a + ib)(\mathbf{u}_1 + i\mathbf{u}_2) \iff \begin{cases} A\mathbf{u}_1 = & a\mathbf{u}_1 - b\mathbf{u}_2, \\ A\mathbf{u}_2 = & a\mathbf{u}_2 + b\mathbf{u}_1, \end{cases}$$

after identifying real and imaginary parts. The result follows by writting these identities in matrix form. $\blacksquare$

Corollary 17.7 Let $A \in \mathbb{R}^{n \times n}$. Let $(\lambda, \mathbf{v})$ be a complex eigenpair of A, where

$$\lambda = a + ib, \quad \mathbf{v} = \mathbf{u}_1 + i\mathbf{u}_2,$$

with $a, b \in \mathbb{R}$, $b \neq 0$ and $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then $U := \text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ is an A-invariant subspace of $\mathbb{R}^n$.

Proof. Let $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ be a basis for $\mathbb{R}^n$. Let $\mathbf{u} \in U$. Then, using Lemma (17.6),

$$A\mathbf{u} = A(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) = \alpha A\mathbf{u}_1 + \beta A\mathbf{u}_2 = \alpha(a\mathbf{u}_1 - b\mathbf{u}_2) + \beta(a\mathbf{u}_2 + b\mathbf{u}_1) = (a\alpha + b\beta)\mathbf{u}_1 + (a\beta - b\alpha)\mathbf{u}_2 \in U.$$

$\blacksquare$

The results of the lemma and the corollary lead to the following important observation: given a real matrix A, the two complex one-dimensional A-invariant subspaces of $\mathbb{C}^n$ associated with a complex eigenvalue and its complex-conjugate induce a real two-dimensional A-invariant subspace of $\mathbb{R}^n$. This means that we can replace the complex diagonal canonical form of A with a real canonical form that involves a block diagonal matrix, with real blocks of size 1×1, if the corresponding eigenvalue is real, or of size 2×2 for corresponding complex eigenvalues $\lambda, \bar{\lambda}$.

Proposition 17.8 Let $A \in \mathbb{R}^{n \times n}$ be a non-defective matrix with
- m real eigenvalues: $\lambda_1, \dots, \lambda_m$, $i = 1, \dots, m$;
- $k = (n - m)/2$ complex-conjugate eigenvalues: $\lambda_j = a_j \pm ib_j$, $j = 1, \dots, k$.

Then $A = MDM^{-1}$, with $D, M \in \mathbb{R}^{n \times n}$ and with D a block-diagonal matrix with
- m blocks of size 1×1: $\lambda_1, \dots, \lambda_m$;
- k blocks of size 2×2 of the form

$$\begin{bmatrix} a_j & b_j \\ -b_j & a_j \end{bmatrix}, \quad j = 1, \dots, k.$$

The block-diagonal matrix D is the real canonical form of a square real matrix A.

We briefly summarise our findings below: any non-defective matrix can be diagonalised as follows:
- $A \in \mathbb{C}^{n \times n}$ is similar to a complex diagonal matrix, with eigenvalues on the diagonal;
- $A \in \mathbb{R}^{n \times n}$ is similar to a block diagonal matrix, with diagonal entries as described in Prop. 17.8.