Chapter 1

The Characteristic equation of a matrix

1.1 Eigenvalues and eigenvectors

Definition 1.1.0.1. *: An element $\lambda \in \mathbb{F}$ is an eigenvalue of a matrix $A \in \mathbb{F}^{n*n}$ if there exists a nonzero vector $x \in \mathbb{F}^n$ such that $Ax = \lambda x$. The vector x is said to be an eigenvector of A corresponding to the eigenvalue λ.*
A nonzero row vector y is a left eigenvector of A, corresponding to the eigenvalue λ, if $yA = \lambda y$.

For $A \in \mathbb{F}^{n*n}$, the characteristic polynomial of A is given by $p_A(x) = det(xI - A)$.
Eigenvalues were initial used by Leonhard Euler in 1743 in connection with the solution to an order linear differential equation with constant coefficients.
Geometrically, the equation implies that the n-vectors $\boldsymbol{Ax}$ and $\boldsymbol{x}$ are parallel.

1.1.1 The characterstic polynomial

The polynomial $det(A - \lambda I_{n\times n})$ is called the **characteristic polynomial** of A and is often denoted by $ch_A(\lambda)$. The equation

$$det(A - \lambda I_{n\times n}) = 0$$

is called the **characteristic equation** of A. Hence the eigenvalues of A are the roots of the characteristic polynomial of A.

- The algebraic multiplicity, $\alpha(\lambda)$, of $\lambda \in \sigma(A)$ is the number of times the eigenvalue occurs as a root in the characteristic polynomial of A.

- The spectrum of $A \in \mathbb{F}^{n*n}$, $\sigma(A)$, is the multi-set of all eigenvalues of A, with eigenvalue λ appearing $\alpha(\lambda)$ times in $\sigma(A)$.

- The spectral radius of $A \in \mathbb{F}^{n*n}$ is

  $$\rho(A) = max\{|\lambda| : \lambda \in \sigma(A)\}.$$

  Let $p(x) = c_n x^n + c_{n-1}x^{n-1} + ... + c_2x^2 + c_1x + c_0$ be a polynomial with coefficients in $\mathbb{F}$. Then $p(A) = c_nA^n + c_{n-1}A^{n-1} + ... + c_2A^2 + c_1A + c_0I$.

- For $A \in \mathbb{F}^{n*n}$, the minimal polynomial of A, $q_A(x)$, is the unique monic polynomial of least degree for which $q_A(A) = 0$.

- The vector space $ker(A - \lambda I)$, for $\lambda \in (A)$, is the eigenspace of $A \in \mathbb{F}^{n*n}$ corresponding to λ, and is denoted by $E_\lambda(A)$.
- The geometric multiplicity, $\gamma(\lambda)$, of an eigenvalue λ is the dimension of the eigenspace $E_\lambda(A)$.
- An eigenvalue λ is simple if $\alpha(\lambda) = 1$.

An eigenvalue λ is semi-simple if $\alpha(\lambda) = \gamma(\lambda)$.

Theorem 1.1.1.1. *If A is an $n \times n$ matrix and λ is a eigenvalue of A, then the set of all eigenvectors of λ, together with the zero vector, forms a subspace of $\mathbb{R}^n$.*

$$E(\lambda) = \{0\} \bigcup \{x : x \textit{ is an eigenvector corresponding to } \lambda\}.$$

(of all eigenvalues corresponding to λ, together with 0) is a subspace of $\mathbb{R}^n$. This subspace $E(\lambda)$ is called the ***eigenspace*** *of λ.*

Proof:
Since $0 \in E(\lambda)$, we have $E(\lambda)$ is nonempty.
Next let us check $E(\lambda)$ is closed under addition and scalar multiplication or not.
Suppose $x, y \in E(\lambda)$ and c be a scalar. Then,

$$Ax = \lambda x \textit{ and } Ay = \lambda y.$$

So,

$$A(x + y) = Ax + Ay = \lambda x + \lambda y = \lambda(x + y).$$

So, $x + y$ is an eigenvector corresponding to λ or zero. So, $x + y \in E(\lambda)$ and $E(\lambda)$ is closed under addition. Also,

$$A(cx) = c(Ax) = c(\lambda x) = \lambda(cx).$$

So, $cx \in E(\lambda)$ and $E(\lambda)$ is closed under scalar multiplication.
Therefore, $E(\lambda)$ is a subspace of $\mathbb{R}^n$. The proof is complete.

$$\begin{aligned} Ax &= \lambda x \quad (A_{n\times n} \\ Ax - \lambda x &= 0 \\ \underbrace{(A - \lambda I)}_{matrix} \underbrace{x}_{vector} &= \underbrace{0}_{vector} \end{aligned}$$

Need $(A - \lambda I)$ to not be invertible, because if it was, we would only have the trivial solution $x = 0$. Set

$$\Rightarrow \textit{ Set } det(A - \lambda I) = 0.$$

The roots of the characteristic equation are the eigenvalues λ.
For each eigenvalue λ, find its eigenvector by solving $(Ax - \lambda I)x = 0$.

Shortcut method to find the characteristic equation of a matrix *:*
For 2×2,

$$\lambda^2 - \underbrace{(trace(A)}_{sum\ of\ diagonal\ entries} \lambda + det(A) = 0.$$

For 3×3,

$$\lambda^3 - (trace(A)\lambda^2 + \underbrace{(C_{11} + C_{22} + C_{33})}_{sum\ of\ the\ diagonal\ cofactors} \lambda - det(A) = 0.$$

Example 1.1.1.1. *Find the eigenvalues and eigenvectors of* $\begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 3 & -1 \end{pmatrix}$.

Solution: $det(A - \lambda I) = 0$

$$\Rightarrow \left|\begin{pmatrix} -1-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 3 & -1-\lambda \end{pmatrix}\right| = (1-\lambda)\left|\begin{pmatrix} 2-\lambda & 1 \\ 3 & -1-\lambda \end{pmatrix}\right| + (-1)\left|\begin{pmatrix} 1 & 1 \\ 0 & -1-\lambda \end{pmatrix}\right|$$

$$\begin{aligned} &= (-1-\lambda)[(2-\lambda)(-1-\lambda) - 3] - (-1-\lambda) = 0 \\ &= (-1-\lambda)[(-2-\lambda+\lambda^2) - 3] - (-1-\lambda) = 0 \\ &\Rightarrow (-1-\lambda)[(\lambda^2-\lambda-5)] - (-1-\lambda) = 0 \\ &\Rightarrow (-1-\lambda)((\lambda^2-\lambda-5) - 1) = 0 \\ &\Rightarrow (-1-\lambda)(\lambda^2-\lambda-6) = 0 \\ &\Rightarrow (-1-\lambda)(\lambda-3)(\lambda+2) = 0 \end{aligned}$$

$\lambda_1 = -1,\ \lambda_2 = -2,\ \lambda_3 = 3$ *are the eigenvalues of* A.
For $\lambda_1 = -1$,

$$(A - \lambda I)\vec{x} = 0 \Rightarrow (A+I)\vec{v_1} = 0 \begin{pmatrix} 0 & 1 & 0 & : & 0 \\ 1 & 3 & 1 & : & 0 \\ 0 & 3 & 0 & : & 0 \end{pmatrix} R_1 \leftrightarrow R_2 \begin{pmatrix} 1 & 3 & 1 & : & 0 \\ 0 & 1 & 0 & : & 0 \\ 0 & 3 & 0 & : & 0 \end{pmatrix} R_3 \leftrightarrow R_3 - 3R_1 \begin{pmatrix} 1 & 3 & 1 & : & 0 \\ 0 & 1 & 0 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{pmatrix}$$

$$\begin{cases} x + 3y + z = 0 \Rightarrow x = -z \\ \Rightarrow y = 0 \end{cases} \quad z \text{ is free, let } z = 1 \Rightarrow \vec{v_1} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

For $\lambda_2 = -2$,

$$(A - \lambda I)\vec{x} = 0 \Rightarrow (A + 2I)\vec{v_2} = 0$$

$$\Rightarrow \begin{pmatrix} 1 & 1 & 0 & : & 0 \\ 1 & 4 & 1 & : & 0 \\ 0 & 3 & 1 & : & 0 \end{pmatrix} R_1 \leftrightarrow R_2 \begin{pmatrix} 1 & 1 & 0 & : & 0 \\ 0 & 3 & 1 & : & 0 \\ 0 & 3 & 1 & : & 0 \end{pmatrix} R_3 \leftrightarrow R_3 - R_2 \begin{pmatrix} 1 & 1 & 0 & : & 0 \\ 0 & 3 & 1 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{pmatrix}$$

which is the REF of the matrix.

$$\begin{cases} x + y = 0 \Rightarrow x = -y \Rightarrow x = \frac{1}{3}z \\ 3y + z = 0 \Rightarrow y = -\frac{1}{3}z \end{cases}$$

Z *is free. Let* $z = 3$. *and* $\vec{v_2} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$.

For $\lambda_3 = 3$,

$$(A - \lambda I)x = 0 \Rightarrow (A - 3I)\vec{v_3} = 0.$$

$$\begin{pmatrix} -4 & 1 & 0 & : & 0 \\ 1 & -1 & 1 & : & 0 \\ 0 & 3 & -4 & : & 0 \end{pmatrix} R_1 \leftrightarrow R_2 \begin{pmatrix} 1 & -1 & 1 & : & 0 \\ 0 & -3 & 4 & : & 0 \\ 0 & 3 & -4 & : & 0 \end{pmatrix} R_3 \leftrightarrow R_3 + R_2 \begin{pmatrix} 1 & -1 & 1 & : & 0 \\ 0 & -3 & 4 & : & 0 \\ 0 & 0 & 0 & : & 0 \end{pmatrix}$$

$$\begin{cases} x - y + z = 0 \Rightarrow x = y - z \\ 3y + 4z = 0 \Rightarrow y = -\frac{4}{3}z \end{cases}.$$

z is free, let $z = 3$. $\vec{v_3} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$ *are eigenvectors of* A.

Properties of eigenvalues and eigenvectors *Let A be an $n \times n$ invertible matrix. The following are true:*

1. *If A is triangular, then the diagonal elements of A are the eigenvalues of A.*
2. *If λ is an eigenvalue of A with eigenvector $\vec{x}$, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with eigenvector $\vec{x}$.*
3. *If λ is an eigenvalue of A then λ is an eigenvalue of A^T.*
4. *The sum of the eigenvalues of A is equal to $tr(A)$, the trace of A.*
5. *The product of the eigenvalues of A is the equal to $det(A)$, the determinant of A.*

1.2 Similarity of matrices

Definition 1.2.0.1. *: Let A and B be an $n \times n$ matrices. We say that A is similar to B and write $A \backsim B$ if there is an invertible matrix S such that $A = SBS^{-1}$. Similarity of matrices is an equivalence relation meaning that*

1. *$A \backsim A$ because $A = IAI^{-1}$*
2. *If $A \backsim B$, then $B \backsim A$ because $A = SBS^{-1}$ implies $S^{-1}AS = S^{-1}(SBS^{-1})S = (S^{-1}S)B(S^{-1}S) = IBI = B$;*
3. *If $A \backsim B$ and $B \backsim C$, then $A \backsim C$ because $A = SBS^{-1}$ and $B = TCT^{-1}$ implies $A = SBS^{-1} = S(TCT^{-1})S^{-1} = (ST)C(T^{-1}S^{-1}) = (ST)C(ST)^{-1}$.*

Thus, the word "similar" behaves as it does in its everyday use:

1. *Any thing is similar to itself.*
2. *If one thing is similar to another thing, then the other thing is similar to the first thing.*
3. *If a thing is similar to a second thing and the second thing is similar to the third thing, then the first thing is similar to the third thing.*

Theorem 1.2.0.2. *: Similar matrices have the same*

1. *determinant;*
2. *characteristic polynomial;*
3. *eigenvalue.*

Proof. *Suppose A and B are similar. Then $A = SBS^{-1}$ for some invertible matrix S.*

1. *Since $det(S^{-1}) = (detS)^{-1}$, we have $det(A) = (detS)(detB)(detS^{-1}) = (detS)(detB)(detS)^{-1} = (detS)(detS^{-1})(detB) = detB$.*
   We used the fact that the determinant of a matrix is a number.
2. *Since $S(tI)S^{-1} = ti, A - tI = S(B - tI)s^{-1}$; i.e., $A - tI$ **and** $B - tI$ are similar. (2) now follows from (1) because the characteristic polynomial of A is $det(A - tI$.*

3. *The eigenvalues of a matrix are the roots of its characteristic polynomial. Since A and B have the same characteristic polynomial they have the same eigenvalues.*

Warning: *Although they have the same eigenvalues similar matrices do not usually have the same eigenvectors or eigenspaces. Nevertheless there is a precise relationship between the eigenspaces of a similar matrices.*

Proposition 1.2.0.1. *: Suppose $A = SBS^{-1}$. Let $E_\lambda(A)$ be the $\lambda-$ eigenspace for A and $E_\lambda(B)$ the $\lambda-$ eigenspace for B. Then $E_\lambda(B) = S^{-1}E_\lambda(A)$, i.e., $E_\lambda(B) = \{S^{-1}X : X \in E_\lambda(A)\}$*
In particular, the dimension of the $\lambda-$eigenspaces for A and B are the same.

Proof. If $X \in E_\lambda(A)$, then $\lambda X = AX = SBS^{-1}X$ so $BS^{-1}X = S^{-1}\lambda X = \lambda(S^{-1}X)$; i.e., $S^{-1}X$ is a $\lambda - eigenvector$ for B or, equivalently, $S^{-1}E_\lambda(A) \subseteq E_\lambda(B)$.

Starting from the fact that $B = S^{-1}AS$, the same sort of argument shows that $SE_\lambda(B) \subseteq E_\lambda(A)$.

Therefore

$E_\lambda(B) = I.E_\lambda(B) = S^{-1}S.E_\lambda(B) \subseteq S^{-1}.E_\lambda(A) \subseteq E_\lambda(B)$.

In particular, $E_\lambda(B) \subseteq S^{-1}.E_\lambda(A) \subseteq E_\lambda(B)$ so these three sets are equal, i.e., $E_\lambda(B) =\subseteq S^{-1}.E_\lambda(A) = \{S^{-1}X : X \in E_\lambda(A)\}$.
That $dimE_\lambda(A) = dimE_\lambda(B)$ is proved. □

If A and B are similar and $A^r = 0$ and $B^r = 0$.

Corollary 1.2.0.1. *: Similar matrices have the same rank and nullity.*

Proof. Suppose A and B are similar matrices. By definition, the nullity of A is the dimension of its null space. But $N(A) = E_0(A)$, the 0-eigenspace of A. By the above proposition $E_0(A)$ and $E_0(B)$ have the same dimension. Hence A and B have the same nullity.

Since rank+nullity= n, A and B also have the same rank. □

Example 1.2.0.2. *Show that the matrices $B = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$ and $A = \begin{pmatrix} 8 & -5 \\ 3 & -2 \end{pmatrix}$ are similar.*

Solution

Take $S = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, then $SBS^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}(\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix})^{-1} = \begin{pmatrix} 8 & 11 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -5 \\ 3 & -2 \end{pmatrix} = A$. Therefore $A \backsim B$.

1.3 The spectral radius of a matrix

Definition 1.3.0.2. *: For an $n \times n$ matrix A define*

1. *$\sigma(A) = \{\lambda : Ax = \lambda x$ has a solution for a nonzero vector $x\}$. $\sigma(A)$ is called the **spectrum** of A.*

2. *$\rho(A) = max\{|\lambda| : \lambda \in \sigma(A)\}$. $\rho(A)$ is called the **spectral radius** of A.*

Example 1.3.0.3. *: Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Then $\lambda = 1$ is an eigenvalue of A with eigenvector $\boldsymbol{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Also $\lambda = 3$ is an eigenvalue of A with eigenvector $\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The spectrum of A is $\sigma(A) = \{1, 3\}$ and the spectral radius of A is $\rho(A) = 3$.*

Example 1.3.0.4. *: The* 3×3 *matrix* $B = \begin{pmatrix} -3 & 0 & 6 \\ -12 & 9 & 26 \\ 4 & -4 & -9 \end{pmatrix}$ *has eigenvalues:* $-1, -3, 1$. *Pertaining to the eigenvalues are the eigenvectors* $\{\begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}\} \leftrightarrow 1$, $\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\} \leftrightarrow -3$, $\{\begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}\} \leftrightarrow -1$

For an $n \times n$ *matrix* A, $\parallel A \parallel = max\{|A_{ij}| : 1 \leq i, j \leq n\}$.

$$\lim_{n \to \inf} \parallel A^n \parallel = \begin{cases} 0 \; if \; \rho(A) < 1, \\ ?????????, \; if \; \rho(A) > 1. \end{cases}$$

1.4 Diagonilization

Recall, a matrix, D*, is diagonal if it is square and the only non-zero entries are on the diagonal. This is equivalent to* $D\vec{e_i} = \lambda_i \vec{e_i}$ *where here* $\vec{e_i}$ *are the standard vector and the* λ_i *are the diagonal entries.*
And Square matrices A *and* M *are similar if there is an invertible matrix* P *such that* $P^{-1}AP = M$.
A linear transformation, $T : \mathbb{R}^n \to \mathbb{R}^n$*, is diagonalizable if there is a basis* B *of* $\mathbb{R}^n$ *so that* $[T]_B$ *is diagonal. This means* $[T]$ *is similar to the diagonal matrix* $[T]_B$*. Similarly, a matrix* $A \in \mathbb{R}^{n \times n}$ *is diagonalizable if it is similar to some diagonal matrix* D*.*

Steps of Diagonalizing an $n \times n$ **matrix and/or Linear mapping:**

1. *Compute the matrix representation of the linear map with respect the bases of the vector spaces for which the map is defined say* A*.*

2. *Compute* $\det(A - \lambda I)$*.*

3. *Find all the eigenvalues of* A*.*

4. *Find* n *linearly independent eigenvectors* $p_1, p_2, ..., p_n$ *for* A *with corresponding eigenvalues* $\lambda_1, \lambda_2, ..., \lambda_n$*. If* n *independent eigenvectors do not exists, then* A *is not diagonalizable.*

5. *If* A *has* n *linearly independent eigenvectors as above, write*

$$P = [p_1 \; p_2 \; .... \; p_n]$$

   and

6. *Then* $D = P^{-1}AP$ *is a diagonal matrix with diagonal elements are eigenvalues of* $A/[T]$*. That is*

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Theorem 1.4.0.3. *An* $n \times n$ *matrix* A *is diagonalizable if and only if it has* n *linearly independent eigenvectors.*
Proof*:*

(By induction). The statement is trivially true if $n = 1$.
Suppose it is true for $n = 1, 2, \cdots, k-1$; i.e. if $\lambda_1, \cdots, \lambda_{k-1}$ are distinct and $v_1, \cdots, v_{k-1}$ are corresponding eigenvectors, then they are LI.
Suppose λ_k is an eigenvalue distinct from $\lambda_1, \cdots, \lambda_{k-1}$ and v_k is a corresponding eigenvector. For some scalars $c_1, ..., c_k$ suppose that

$$c_1 v_1 + ... + c_k v_k = 0.$$

So, $A(c_1 v_1 + ... + c_k v_k) = A0 = 0$.
By linearity of A we get -

$$c_1 A v_1 + ... + c_k A v_k = 0 \Rightarrow c_1 \lambda_1 v_1 + ... + c_k \lambda_k v_k = 0.$$

$$\Rightarrow c_1(\lambda_1 - \lambda_k) v_1 + ... + c_{k-1}(\lambda_{k-1} - \lambda_k) v_{k-1} = 0.$$

By LI of

$$\{v_1, ..., v_k\},$$

$$c_i(\lambda_i - \lambda_k) = 0, \ \forall i = 1 : k-1.$$

Since $\lambda_i \neq \lambda_k$, $c_i = 0$, $\forall i = 1 : k$, we get $c_k = 0$ also, so that $\{v_1, .., v_k\}$ are LI.

Theorem 1.4.0.4. *An $n \times n$ matrix A is diagonalizable if and only if there is a basis of $\mathbb{R}^n$ consisting only of eigenvectors of A.*

Geometrical Interpretation Applications :

- *Let's look at A as the matrix representing a linear transformation $T = TA$ in standard coordinates, ie, $T(x) = Ax$.*
- *let's assume A has a set of linearly independent vectors $B = \{v_1, v_2, \cdots, v_n\}$ corresponding to the eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then B is a basis of $\mathbb{R}^n$.*
- *what is the matrix representing T wrt the basis B?*

$$A_{[B,B]} = P^{-1} A P$$

  where $P = [v_1 \ v_2 \ ... \ v_n]$
- *hence, the matrices A and $A_{[B,B]}$ are similar, they represent the same linear transformation:*
- *A in the standard basis*
- *$A_{[B,B]}$ in the basis B of eigenvectors of A*
- *$A_{[B,B]} = [[T(v_1)]_B \ [T(v_2)]_B \ \cdots \ [T(v_n)]_B]$*
  for those vectors in particular $T(v_i) = Av_i = \lambda_i v_i$ hence diagonal matrix $A_{[B,B]} = D$.

Example 1.4.0.5. $A = \begin{pmatrix} -3 & -1 & -2 \\ 1 & -1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$
Eigenvalue $\lambda_1 = -1$ *has multiplicity* 2; $\lambda_2 = -2$.

$$(A+I) = \begin{pmatrix} -2 & -1 & -2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \overset{RREF}{\longleftrightarrow} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The rank is 2.
The null space $(A+I)$ *therefore has dimension* 1 *(rank-nullity theorem). We find only one linearly independent vector:* $x = [-1, 0, 1]^T$.
Hence the matrix A *cannot be diagonalized.*

(multiplicity)

Definition 1.4.0.3. *An eigenvalue* λ_0 *of a matrix* A *has*
algebraic multiplicity k *if* k *is the largest integer such that* $(\lambda - \lambda_0)^k$ *is a factor of the characteristic polynomial.*
geometric multiplicity k *if* k *is the dimension of the eigenspace of* λ_0*, ie,* $dim(N(A - \lambda_0 I))$.

Theorem 1.4.0.5. *For any eigenvalue of a square matrix, the geometric multiplicity is no more than the algebraic multiplicity.*

Theorem 1.4.0.6. *A matrix is diagonalizable if and only if all its eigenvalues are real numbers and, for each eigenvalue, its geometric multiplicity equals the algebraic multiplicity.*

Complete Set of Eigenvectors

A complete set of eigenvectors for an matrix A *is a set of* n *linearly independent set of eigenvectors for* A.
A matrix that does not have a complete set of eigenvectors is said to be deficient.
Summary:
n *distinct eigenvalues* $\rightarrow$ *complete set* $\rightarrow$ *diagonalizability.*

Orthogonal Matrix:

Definition 1.4.0.4. *A square matrix* Q *is said to be an orthogonal matrix if*

$$QQ^T = Q^TQ = I.$$

If Q *is nonsingular, then we have a wonderful simplicity for the inverse of the matrix* Q*, in fact,*

$$QQ^T = I \text{ implies } Q = Q^{-1}.$$

Note: $QQ^T = I$ *if and only if the columns of* Q *are such that:*

$$q^{(i)T} q^j = \delta_{ij} := \begin{cases} 1, & if\ i = j \\ 0, & if\ i \neq j. \end{cases}$$

1.5 Block Decomposable Matrices

Definition 1.5.0.5. *: A partitioned matrix also called a block matrix is a partition of a matrix into rectangular smaller matrices called blocks. Equivalently, using a system of horizontal and vertical (dashed) lines, we can partition a matrix A into sub-matrices called blocks (cells) of A. Clearly a given matrix may be divided into blocks in different ways. The convenience of the partition of matrices, say A and B, into blocks is that the result of operations on A and B can be obtained by carrying out the computation with the blocks, just as if they were the actual elements of the matrices. This is illustrated below, where the notation $A = (A_{ij})$ will be used for a block matrix A with blocks A_{ij}.*

Suppose that $A = (A_{ij})$ and $B = (B_{ij})$ are block matrices with the same numbers of row and column blocks and suppose that corresponding blocks have the same size. Then adding the corresponding blocks of A and B also adds corresponding elements of A and B, and multiplying each block of A by a scalar k multiplies each element of A by k.

$$\text{Thus, } A + B = \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1n}+B_{1n} \\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2n}+B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & \cdots & A_{mn}+B_{mn} \end{pmatrix} = ((A_{ij}) + (B_{ij}))_{m\times n}.$$

Example 1.5.0.6. *P can be partitioned into 4 2×2 blocks*

$$P = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 3 & 3 & 4 & 4 \\ 3 & 3 & 4 & 4 \end{bmatrix}$$

$$P_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},\ P_{12} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix},\ P_{21} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix},\ P_{22} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

Then we can write the partitioned matrix like this

$$P = P_{partitioned} = \left(\begin{array}{c|c} P_{11} & P_{12} \\ \hline P_{21} & P_{22} \end{array}\right)$$

Note:

When we are partitioning any matrix, the diagonal blocks must be square matrix.

1.6 Cayley-Hamilton theorem and Minimal polynomial

Note: The minimum polynomial is often also called the "minimal polynomial".

Definition 1.6.0.6. *Let $T : V \to V$ be a linear transformation of a finite-dimensional vector space over the field $\mathbb{F}$. The minimum polynomial $m_T(x)$ is the monic polynomial over F of smallest degree such that $m_T(T) = 0$.*
To say that a polynomial is monic is to say that its leading coefficient is 1. Our definition of the characteristic polynomial ensures that $c_T(x)$ is also always a monic polynomial.
The definition of the minimum polynomial (in particular, the assumption that it is monic) ensures that it is unique.

Theorem 1.6.0.7. ***(Cayley–Hamilton Theorem)*** *Let $T : V \to V$ be a linear transformation of a finite-dimensional vector space V. If $c_T(x)$ is the characteristic polynomial of T, then $c_T(T) =$*

0.
It is tempting to give the following "proof":

$$P_A(A) = det(A - AI) = det\boldsymbol{0} = 0.$$

This is wrong. In the first place, notice that the output of determinant is a scalar, while $P_A(A)$ is a matrix-indeed, writing $P_A(t) = a_0 + a_1 t + ... + a_{n-1}t^{n-1} + t^n$, we have

$$P_A(A) = a_0 I + a_2 A^2 + ... + a_{n-1}A^{n-1} + A^n.$$

This is an $n \times n$ matrix, not a determinant of anything.
We first observe that if A happens to be diagonallizable (in particular, if A has n distinct eigenvalues), then writing any $v \in \mathbb{C}^n$ as a linear combination of eigenvectors $v_1,\ v_2, ...,\ v_n$, we get

$$P_A(A)v = P_A(A)\sum_{j=1}^{n} c_j v_j = \sum_{j=1}^{n} c_j P_A(A)v_j = \sum_{j=1}^{n} c_j P_A(\lambda_j)v_j = 0,$$

showing that $P_A(A) = 0$. However, if A is diagonalizable then we need to use a different argument.

Facts about polynomials *:*
Let $\mathbb{F}$ be a field and recall $\mathbb{F}[x]$ denotes the set of polynomials with coefficients from $\mathbb{F}$:

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

(where $a_i \in \mathbb{F}$).
Then $\mathbb{F}[x]$ is an example of what is known as a Euclidean domain. A summary of its main properties are:

- *We can add, multiply and subtract polynomials;*
- *Euclidean Algorithm: if we attempt to divide $f(x)$ by $g(x)$ (where $g(x) \neq 0$), we obtain*

  $$f(x) = g(x)q(x) + r(x)$$

  where either $r(x) = 0$ or the degree of $r(x)$ satisfies deg $r(x) < degg(x)$ (i.e., we can perform long-division with polynomials).
- *When the remainder is 0, that is, when $f(x) = g(x)q(x)$ for some polynomial $q(x)$, we say that $g(x)$ divides $f(x)$.*
- *If $f(x)$ and $g(x)$ are non-zero polynomials, their greatest common divisor is the polynomial $d(x)$ of largest degree dividing them both. It is uniquely determined up to multiplying by a scalar and can be expressed as*

  $$d(x) = a(x)f(x) + b(x)g(x)$$

  for some polynomials $a(x),\ b(x)$.
  Those familiar with divisibility in the integers $\mathbb{Z}$ will recognize these facts as being standard properties of $\mathbb{Z}$.

Proposition 1.6.0.2. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear transformation. If $f(x)$ is any polynomial (over $\mathbb{F}$) such that $f(T) = 0$, then the minimum polynomial $m_T(x)$ divides $f(x)$.*

Corollary 1.6.0.2. *Suppose that $T : V \to V$ is a linear transformation of a finite-dimensional vector space V. Then the minimum polynomial $m_T(x)$ divides the characteristic poly- nomial $c_T(x)$.*

Theorem 1.6.0.8. *Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear transformation of V. Then the roots of the minimum polynomial $m_T(x)$ and the roots of the characteristic polynomial $c_T(x)$ coincide.*

Theorem 1.6.0.9. *Let V be a finite-dimensional vector space over the field $\mathbb{F}$ and let $T : V \to V$ be a linear transformation. Then T is diagonalisable if and only if the minimum polynomial $m_T(x)$ is a product of distinct linear factors.*

Lemma 1.6.0.1. *Let $T : V \to V$ be a linear transformation of a vector space over the field $\mathbb{F}$ and let $f(x)$ and $g(x)$ be coprime polynomials over $\mathbb{F}$. Then*

$$ker f(T)g(T) = ker f(T) \oplus ker g(T).$$

1.7 Application of eigenvalues and eigenvectors

1. ***Communication systems:***
   Eigenvalues were used by Claude Shannon to determine the theoretical limit to how much information can be transmitted through a communication medium like your telephone line or through the air. This is done by calculating the eigenvectors and eigenvalues of the communication channel (expressed a matrix), and then waterfilling on the eigenvalues. The eigenvalues are then, in essence, the gains of the fundamental modes of the channel, which themselves are captured by the eigenvectors.

2. ***Designing bridges:***
   The natural frequency of the bridge is the eigenvalue of smallest magnitude of a system that models the bridge. The engineers exploit this knowledge to ensure the stability of their constructions. [Watch the video on the collapse of the Tacoma Narrow Bridge which was built in 1940]

3. ***Designing car stereo system:***
   Eigenvalue analysis is also used in the design of the car stereo systems, where it helps to reproduce the vibration of the car due to the music.

4. ***Electrical Engineering:***
   The application of eigenvalues and eigenvectors is useful for decoupling three-phase systems through symmetrical component transformation.

5. ***Mechanical Engineering:***
   Eigenvalues and eigenvectors allow us to "reduce" a linear operation to separate, simpler, problems. For example, if a stress is applied to a "plastic" solid, the deformation can be dissected into "principle directions"- those directions in which the deformation is greatest. Vectors in the principle directions are the eigenvectors and the percentage deformation in each principle direction is the corresponding eigenvalue.
   Oil companies frequently use eigenvalue analysis to explore land for oil. Oil, dirt, and other substances all give rise to linear systems which have different eigenvalues, so eigenvalue analysis can give a good indication of where oil reserves are located. Oil companies place

probes around a site to pick up the waves that result from a huge truck used to vibrate the ground. The waves are changed as they pass through the different substances in the ground. The analysis of these waves directs the oil companies to possible drilling sites.

Exercise 1.1. *:*

1. *Find the eigenvalues, eigenvectors, spectral radius and minimal polynomial of the following matrix.*

   a. $\begin{pmatrix} 1 & 6 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ *b.* $\begin{pmatrix} 0 & -2 & -1 \\ 1 & 5 & 3 \\ -1 & -2 & 0 \end{pmatrix}$

2. *Determine whether the matrix* $\begin{pmatrix} -1 & 2 & -1 \\ -4 & 5 & -2 \\ -4 & 3 & 0 \end{pmatrix}$ *is diagonalizable or not.*