Week 2: Linear Transformation and Matrix Representation

Definition. A linear transformation is a function between vector spaces that preserves vector addition and scalar multiplication. If $T : V \to W$ is linear then for all $u, v \in V$ and all scalars c,

$$T(u+v) = T(u) + T(v), \qquad T(cu) = c\,T(u),$$

or equivalently for scalars c_1, c_2 and vectors v_1, v_2,

$$T(c_1 v_1 + c_2 v_2) = c_1 T(v_1) + c_2 T(v_2).$$

Matrix representation (2$\times 2 case$). Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Then the linear map $T(u) = Au$ gives

$$Au = \begin{bmatrix} a_{11}u_1 + a_{12}u_2 \\ a_{21}u_1 + a_{22}u_2 \end{bmatrix}.$$

The images of the standard basis vectors are

$$Ae_1 = A\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \qquad Ae_2 = A\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}.$$

Examples. 1. Identity transformation:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Example: $\begin{bmatrix} 2 \\ 3 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

2. Swap (exchange) operation:

$$S = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad S\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_2 \\ u_1 \end{bmatrix}.$$

Basis action: $Se_1 = e_2,\ Se_2 = e_1$.

3. Rotation by 90° (counterclockwise):

$$R_{90} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad R_{90}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} -u_2 \\ u_1 \end{bmatrix}.$$

Examples: $R_{90}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},\ R_{90}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$.

Linear Transformations

Consider a linear rule T from $\mathbb{R}^n$ to $\mathbb{R}^m$.

$$T : \mathbb{R}^n \to \mathbb{R}^m.$$

Let $u \in \mathbb{R}^n$. We write the vector u in coordinate form as

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m.$$

The transformation T is a rule that assigns to each vector $u \in \mathbb{R}^n$ a vector $T(u) \in \mathbb{R}^m$. We call $b = T(u)$ the *image* of u under T.

Domain and codomain:

$$\text{Domain}(T) = \mathbb{R}^n, \qquad \text{Codomain}(T) = \mathbb{R}^m.$$

Definition (Image). For $u \in \mathbb{R}^n$ the vector $T(u) \in \mathbb{R}^m$ is called the image of u under T.

Definition (Range). The set of all images

$$\{\, T(x) \mid x \in \mathbb{R}^n \,\}$$

is called the *range* of T, sometimes written Range(T) or Im(T).

Remarks on representation by columns. If $e_1, \ldots, e_n$ denote the standard basis vectors of $\mathbb{R}^n$, then the matrix of T (with respect to the standard bases) has as its j-th column the vector $T(e_j) \in \mathbb{R}^m$. For a general vector $x = \sum_{j=1}^n x_j e_j$,

$$T(x) = \sum_{j=1}^{n} x_j T(e_j).$$

Examples of notation used in transformations:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \mapsto T(u) = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

b = A x.

$$\{\, b : b = Ax \text{ for } x \in \mathbb{R}^n \,\} \;=\; \text{Range}(A).$$

If

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

then the matrix equation $Ax = b$ is equivalent to the system of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Write the columns of A as column vectors

$$\boldsymbol{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \; \boldsymbol{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \; \ldots, \; \boldsymbol{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Then

$$Ax = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n.$$

Thus b is in the range of A precisely when b can be expressed as a linear combination of the column vectors of A.

Definition: The column space (also called the column span or range) of the matrix A is the set of all linear combinations of the column vectors of A. It is denoted $\mathrm{Col}(A)$ or $\mathrm{Range}(A)$.

The dimension of the column space is called the rank of the matrix A ($\mathrm{rank} A$).

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation defined by $T(x) = Ax$, then

$$\mathrm{Range}(T) = \{\, T(x) : x \in \mathbb{R}^n \,\} = \mathrm{Range}(A) = \mathrm{Col}(A),$$

which is a subspace of $\mathbb{R}^m$. In particular, $\{T(x)\}$ (the set of all images under T) is a vector subspace of $\mathbb{R}^m$.

Let T be a transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ represented by a matrix A.

Example.

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ 4x_1 + 2x_2 \end{pmatrix}.$$

Observe the column structure:

$$A = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \begin{pmatrix} 1 & \tfrac{1}{2} \end{pmatrix} \quad \text{or} \quad A = \left[\begin{smallmatrix} 2 \\ 4 \end{smallmatrix} \;\middle|\; \begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right],$$

and the second column equals $\frac{1}{2}$ times the first column. Hence the column space (image) is one-dimensional:

$$\mathrm{Im}(T) = \mathrm{Col}(A) = \mathrm{span}\{(2,4)^\top\} = \mathrm{span}\{(1,2)^\top\}.$$

Kernel (null space). Solve $A\mathbf{x} = \mathbf{0}$:

$$\begin{cases} 2x_1 + x_2 = 0, \\ 4x_1 + 2x_2 = 0 \end{cases}$$

The two equations are dependent; from $2x_1 + x_2 = 0$ we get $x_2 = -2x_1$. Thus

$$\ker(T) = \ker(A) = \{\, (x_1, x_2)^\top : x_2 = -2x_1 \,\} = \mathrm{span}\{(1,-2)^\top\},$$

a one-dimensional subspace (a line through the origin in $\mathbb{R}^2$).

Definitions and remarks. - The kernel of T (also called the null space of A) is

$$\ker(T) = \{\, \mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0} \,\} = \{\, \mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0} \,\}.$$

- Solving $A\mathbf{x} = \mathbf{0}$ is solving a homogeneous linear system; its solution set is a subspace of $\mathbb{R}^n$. - The image (column space) of A is the subspace of $\mathbb{R}^m$ spanned by the columns of A.

Let V, W be two vector spaces. A transformation is a mapping $T : V \to W$ that obeys the following rules:

1. $T(u_1 + u_2) = T(u_1) + T(u_2)$. (Transform of a sum is the sum of transforms.)
2. $T(\alpha u) = \alpha\, T(u)$ for any scalar α. (Homogeneity / scaling.)

Suppose $B = \{v_1, v_2, \ldots, v_k\}$ is a basis of the vector space V, and let v be any vector in V. Write

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

for some scalars $\alpha_1, \ldots, \alpha_k$. If T is linear then

$$\begin{aligned} T(v) &= T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) \\ &= T(\alpha_1 v_1) + T(\alpha_2 v_2) + \cdots + T(\alpha_k v_k) \\ &= \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_k T(v_k). \end{aligned}$$

Definition (vector basis): a basis is a subset of a vector space consisting of vectors that are linearly independent and that span the space.

Remarks about linear maps:

- A linear map always sends the origin to the origin: $T(0) = 0$.
- A linear map maps any line passing through the origin to a line (or the origin) in the target space.

Matrix Representation of a Transformation

Proposition. Any linear transformation $T : V \to W$ is completely determined by what it does to the basis vectors of V.

Let V and W be vector spaces and let $T : V \to W$ be linear. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis for V. For $i = 1, \ldots, n$ define

$$w_i := T(v_i) \in W.$$

Then there exists a unique linear map $T : V \to W$ such that $T(v_i) = w_i$ for all i.

Proof. Let $v \in V$. Since $\{v_1, \ldots, v_n\}$ is a basis, v has a unique representation

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, \qquad \alpha_i \in \mathbb{F}.$$

Define

$$T(v) := \sum_{i=1}^{n} \alpha_i w_i.$$

If $v = v_j$ then the coefficients satisfy $\alpha_j = 1$ and $\alpha_i = 0$ for $i \neq j$, so $T(v_j) = w_j$ as required.

Claim: T is linear. Let $u, v \in V$ with

$$u = \sum_{i=1}^{n} \alpha_i v_i, \qquad v = \sum_{i=1}^{n} \beta_i v_i,$$

and let c be a scalar. Then

$$u + v = \sum_{i=1}^{n} (\alpha_i + \beta_i) v_i \quad \Rightarrow \quad T(u+v) = \sum_{i=1}^{n} (\alpha_i + \beta_i) w_i = \sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{n} \beta_i w_i = T(u) + T(v),$$

and

$$c\,v = \sum_{i=1}^{n} (c\alpha_i) v_i \quad \Rightarrow \quad T(cv) = \sum_{i=1}^{n} (c\alpha_i) w_i = c \sum_{i=1}^{n} \alpha_i w_i = c\,T(v).$$

Hence T is linear. Uniqueness follows because any linear map is determined by its action on a basis.

Remark (matrix representation). If $\{u_1, \ldots, u_m\}$ is a basis of W and each w_i is written in coordinates relative to that basis as

$$w_i = \sum_{j=1}^{m} a_{ji}\, u_j,$$

then the matrix of T (with respect to the chosen bases of V and W) has columns $(a_{1i}, a_{2i}, \ldots, a_{mi})^T$, i.e. $A = (a_{ji})$.

article

Linearity of a linear transformation

Let V be a vector space with basis $\{v_1, \ldots, v_k\}$. Suppose

$$v = \sum_{i=1}^{k} \alpha_i v_i, \qquad u = \sum_{i=1}^{k} \beta_i v_i.$$

Then

$$v + u = \sum_{i=1}^{k} (\alpha_i + \beta_i)\, v_i,$$

and for a linear transformation $T : V \to W$,

$$\begin{aligned} T(v+u) &= T\Big(\sum_{i=1}^{k} (\alpha_i + \beta_i)\, v_i\Big) = \sum_{i=1}^{k} (\alpha_i + \beta_i)\, T(v_i) \\ &= \sum_{i=1}^{k} \alpha_i T(v_i) + \sum_{i=1}^{k} \beta_i T(v_i) = T(v) + T(u). \end{aligned}$$

For a scalar c,

$$T(cv) = T\Big(\sum_{i=1}^{k} (c\alpha_i)\, v_i\Big) = \sum_{i=1}^{k} (c\alpha_i)\, T(v_i) = c \sum_{i=1}^{k} \alpha_i T(v_i) = c\, T(v).$$

Thus T is additive and homogeneous, i.e. linear.

How to get the matrix representation of a linear transformation

To obtain the matrix representation of a linear transformation, compute the images of the basis vectors and place those image vectors as the columns of the matrix (with respect to the chosen bases).

Example

Let $V = \mathbb{R}^2$, $W = \mathbb{R}^2$, and define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - y \end{pmatrix}.$$

Use the standard basis $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Compute the images of the basis vectors:

$$T(e_1) = T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad T(e_2) = T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The matrix A of T with respect to the standard bases is obtained by placing these images as columns:

$$A = \big[T(e_1) \quad T(e_2)\big] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Indeed, for any vector $\begin{pmatrix} x \\ y \end{pmatrix} = xe_1 + ye_2$,

$$T \begin{pmatrix} x \\ y \end{pmatrix} = x\, T(e_1) + y\, T(e_2) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Obtaining the standard matrix with a linear transformation T.

- General idea: For a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ the standard matrix A is given by the images of the standard basis vectors:

$$A = \begin{bmatrix} T(e_1) & T(e_2) \end{bmatrix}, \qquad e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

- Diagrammatic idea: every vector in the domain is mapped to some vector in the codomain; to find the matrix it suffices to compute $T(e_1)$ and $T(e_2)$.

Examples:

1) Zero transformation

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Hence

$$T(e_1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

2) Swap coordinates

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix}.$$

Thus

$$T(e_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

3) Reflection across the y-axis

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}.$$

Then

$$T(e_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

4) Reflection across the x-axis

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}.$$

So

$$T(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \qquad A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

- Note: when performing row/column operations on matrices (e.g. to obtain RREF) one may swap rows or columns as needed; however for the standard matrix of a linear map, compute the images of basis vectors and place them as columns.

Example. Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^2$. Define

$$T : \mathbb{R}^2 \to \mathbb{R}^2, \qquad T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ y \end{pmatrix}.$$

Then on the standard basis

$$T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

so the matrix of T relative to the standard bases is

$$m(T) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Example 5. Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. A general linear map $T : \mathbb{R}^2 \to \mathbb{R}^3$ can be written as

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \\ a_{31}x + a_{32}y \end{pmatrix}.$$

Hence

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \qquad T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix},$$

and the matrix of T (a 3×2 matrix) is

$$m(T) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

Question. Does T map the entire $\mathbb{R}^2$ onto entire $\mathbb{R}^3$?

$$\text{No: } \operatorname{rank}(m(T)) \le 2 < 3,$$

so the image of T is at most a 2-dimensional subspace of $\mathbb{R}^3$ and cannot equal all of $\mathbb{R}^3$.

Linearity reminders:

$$T(v + w) = T(v) + T(w), \qquad T(cv) = cT(v)$$

for all $v, w \in V$ and all scalars c.

eg1. $V = \mathbb{R}^3 \longrightarrow W = \mathbb{R}^2$.

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + x_2 \\ x_3 \end{pmatrix}.$$

Images of standard basis vectors:

$$T\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad T\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Matrix of T (with respect to the standard bases):

$$M = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad M\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_3 \end{pmatrix}.$$

Does T map entire $\mathbb{R}^3$?

Suppose U, V, W are three vector spaces. Let

$$T_1 : U \to V, \qquad T_2 : V \to W$$

be linear transformations and define $T = T_2 \circ T_1 : U \to W$. Then T is linear:

$$\begin{aligned} T(u+v) &= T_2\big(T_1(u+v)\big) = T_2\big(T_1(u) + T_1(v)\big) \\ &= T_2\big(T_1(u)\big) + T_2\big(T_1(v)\big) = T(u) + T(v), \end{aligned}$$

and for any scalar c,

$$T(cu) = T_2\big(T_1(cu)\big) = T_2\big(c\,T_1(u)\big) = c\,T_2\big(T_1(u)\big) = c\,T(u).$$

Therefore $T = T_2 \circ T_1$ is a linear transformation $U \to W$.

Eigen vectors & Eigen values

- If $T(\mathbf{u}) = k\,\mathbf{u}$ where T is a linear transformation and k is a scalar, we say (call) the vector $\mathbf{u}$ the *eigenvector* of T and the scalar k is called the *eigenvalue*.
- The eigenvector is a non-zero vector.

Question: How do we identify/obtain the eigenvectors for a linear map $T : \mathbb{R}^n \to \mathbb{R}^n$?
Suppose $A \in \mathbb{R}^{n\times n}$ is the matrix representation of T. If $T(\mathbf{v}) = k\mathbf{v}$, this can be expressed as

$$A\mathbf{v} = k\mathbf{v}.$$

Equivalently,

$$A\mathbf{v} = kI\mathbf{v} \quad\Longrightarrow\quad (A - kI)\mathbf{v} = \mathbf{0},$$

a homogeneous system of equations.

Here k is a scalar and I denotes the $n \times n$ identity matrix.

Since we seek a non-trivial solution $\mathbf{v} \neq \mathbf{0}$ to the homogeneous system $(A-kI)\mathbf{v} = \mathbf{0}$, the matrix $(A-kI)$ must be singular (non-invertible). Hence

$$\det(A - kI) = 0,$$

which is the characteristic equation of A.

$$\boxed{\det(A - kI) = 0}$$

Eigenvalues and eigenvectors (notes)

- The roots of the characteristic equation are scalars that scale the vector upon the action of T on a vector $\boldsymbol{v}$.

- The roots are called eigenvalues of A.

Example. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x + 2y \end{pmatrix}.$$

The matrix of T (with respect to the standard basis) is

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Characteristic equation: solve $\det(A - \lambda I) = 0$.

$$A - \lambda I = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}, \qquad \det(A - \lambda I) = (2-\lambda)^2 - 1.$$

Compute the polynomial:
$$(2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda-3)(\lambda-1).$$
Hence the eigenvalues are
$$\lambda = 3, \quad \lambda = 1.$$

Find eigenvectors.
- For $\lambda = 3$:
$$(A-3I)\boldsymbol{v} = 0 \quad \Longrightarrow \quad \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
This gives $-v_1 + v_2 = 0$, so $v_1 = v_2$. Thus the eigenspace is
$$\operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\},$$
and any eigenvector for $\lambda = 3$ is $t \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $t \neq 0$.
- For $\lambda = 1$:
$$(A-I)\boldsymbol{v} = 0 \quad \Longrightarrow \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$
This gives $v_1 + v_2 = 0$, so $v_2 = -v_1$. Thus the eigenspace is
$$\operatorname{span}\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\},$$
and any eigenvector for $\lambda = 1$ is $s \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $s \neq 0$.

Summary:
$$\text{Eigenvalues: } \{3, 1\}, \qquad \text{Eigenvectors: } \begin{cases} \lambda = 3: & \operatorname{span}\{(1,1)^\mathsf{T}\}, \\ \lambda = 1: & \operatorname{span}\{(1,-1)^\mathsf{T}\}. \end{cases}$$

Notes on invariant subspaces and eigenvalues
1. Examples of subspaces (lines through the origin)
$$\left\{ \begin{pmatrix} 3 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\} \quad \Rightarrow \quad \text{line passing through the origin} \Rightarrow \text{ subspace}$$
$$\left\{ \begin{pmatrix} 5 \\ -5 \end{pmatrix} \right\} \quad \Rightarrow \quad \text{line passing through the origin} \Rightarrow \text{ subspace}$$
2. Invariant subspace

- A subspace U is called an invariant subspace of a linear operator T if $T(U) \subseteq U$.
- (Note in the notes:) the eigenspace condition $T(\mathbf{v}) = \lambda \mathbf{v}$ for some scalar λ defines eigenvectors (and their span is an invariant subspace).

3. Characteristic polynomial and roots Let $A \in \mathbb{R}^{n \times n}$. The characteristic equation is
$$\det(A - \lambda I) = 0.$$
Properties:

- $\det(A - \lambda I)$ is a polynomial in λ of degree n.
- Counting multiplicity, there are n roots of this polynomial (in $\mathbb{C}$).

Questions to consider about the roots:

1. Are the n roots distinct?
2. Are there repeated roots?
3. Are the roots complex?
4. Different cases for the spectrum

- Case 1: distinct eigenvalues.
- Case 2: repeated eigenvalues.
- Case 3: complex eigenvalues.

5. Example 1 Let

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Compute the characteristic polynomial:

$$\det(A - \lambda I) = \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3.$$

Hence the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 1$, which are distinct.

Remark: Eigenvectors corresponding to distinct eigenvalues are linearly independent. (Proof: exercise / DIY.)

6. Example 2 (repeated eigenvalue) Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

The characteristic polynomial is

$$\det(A - \lambda I) = (1-\lambda)^2,$$

so $\lambda = 1$ is a repeated eigenvalue (algebraic multiplicity 2). A corresponding eigenvector is, for example, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

7. Algebraic multiplicity (AM) The algebraic multiplicity of an eigenvalue is the number of times that eigenvalue occurs as a root of the characteristic polynomial $\det(A - \lambda I)$.

Notes: Geometric Multiplicity, Rotation, and Diagonalization

Definition. Geometric multiplicity (GM) of an eigenvalue := number of linearly independent eigenvectors associated with that eigenvalue.

Example (informal).

- If the algebraic multiplicity (AM) of λ_1 is 3 and the algebraic multiplicity of λ_2 is 1, then each algebraic multiplicity satisfies AM ≥ 1.
- In general: $\text{GM}(\lambda) \leq \text{AM}(\lambda)$ for each λ.
- If GM of a specific eigenvalue is strictly less than its algebraic multiplicity AM, we say that the corresponding eigenvalue is *deficient*.

Example: Rotation matrix (90 degrees).

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{(rotation by } 90^\circ\text{)}.$$

Compute the characteristic polynomial:

$$\det(A - \lambda I) = \lambda^2 + 1 = 0,$$

so the eigenvalues are $\lambda = \pm i$ (complex) even though A is a real matrix.

Diagonalization of a matrix A (2-by-2 case)

Consider a linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ represented by a 2×2 matrix A.

Suppose A has two distinct eigenvalues $\lambda_1 \neq \lambda_2$. Let v_1 and v_2 be corresponding eigenvectors (so each eigenvalue has at least one eigenvector).

Define the matrix P whose columns are the eigenvectors of A:

$$P = [\, v_1 \; v_2 \,].$$

Then

$$AP = A[\, v_1 \; v_2 \,] = [\, Av_1 \; Av_2 \,] = [\, \lambda_1 v_1 \; \lambda_2 v_2 \,] = [\, v_1 \; v_2 \,] \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

If P is invertible (equivalently, v_1 and v_2 are linearly independent), we obtain the diagonalization

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Remarks:

- Distinct eigenvalues guarantee linearly independent eigenvectors in the 2×2 case.
- Diagonalization requires that the geometric multiplicity for each eigenvalue equals its algebraic multiplicity (so that there are enough independent eigenvectors to form P).

$$AP = PD, \quad \text{where } P = [u_1 \; u_2 \; \cdots], \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

where P=[u$_1$ u_2 $\cdots$], $\quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

$$(AP)P^{-1} = PDP^{-1} \qquad \Longrightarrow \qquad \boxed{A = PDP^{-1}}$$

(eigendecomposition of A).

From $A = PDP^{-1}$ we get

$$A^2 = PD^2P^{-1}, \qquad A^3 = PD^3P^{-1}, \qquad A^k = PD^kP^{-1}.$$

Thus $P^{-1}AP = D$ gives diagonalisation of A.

If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues such that some λ_i has algebraic multiplicity > 1 and $\mathrm{GM}(\lambda_i) <$ (its algebraic multiplicity), then the geometric multiplicities satisfy

$$\mathrm{GM}(\lambda_1) + \mathrm{GM}(\lambda_2) + \cdots + \mathrm{GM}(\lambda_n) < n,$$

so we cannot obtain n linearly independent eigenvectors and A is not diagonalizable.

Example: Consider

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Here $\lambda = 1$ with algebraic multiplicity 2 (AM = 2) but geometric multiplicity GM = 1. Any matrix P formed from eigenvectors is not invertible, so A cannot be written as PDP^{-1}. Hence A is not diagonalizable.

Remarks: - P matrix is not invertible $\implies$ A cannot be diagonalized. - D cannot be expressed as $P^{-1}AP$ if P is singular.

Question (from notes): Suppose A is 3×3 such that it has an eigenvalue with algebraic multiplicity 2. Discuss diagonalizability of A in the following cases: (i) $\mathrm{AM}(\lambda) = 2$ and $\cdots$ (notes list several subcases), (ii) $\cdots$ (iii) If it has 3 distinct eigenvalues.

What can we say about diagonalizability of A in each case?

CASE 3. Eigenvalues are distinct.

$\Rightarrow$ Eigenvectors corresponding to each eigenvalue are in different directions (linearly independent).

$$\Rightarrow \; P = [\mathrm{evec}_1 \; \mathrm{evec}_2 \; \mathrm{evec}_3]$$

Hence P is invertible and $A = PDP^{-1}$ exists (i.e. A is diagonalizable).

CASE 2. $\mathrm{AM}(\lambda_1) = 2, \; \mathrm{AM}(\lambda_2) = 1$.
(Here A is a 3×3 matrix.)
Consider the nullspace of $(A - \lambda_1 I)$. By the rank–nullity theorem

$$\mathrm{rank}(A - \lambda_1 I) + \mathrm{nullity}(A - \lambda_1 I) = \text{number of columns of } A.$$

If we can find two linearly independent eigenvectors for λ_1 (i.e. geometric multiplicity = 2) and one eigenvector for λ_2 (geometric multiplicity = 1), then we obtain three linearly independent eigenvectors in total. Thus P is invertible and A is diagonalizable.

CASE 1. A is 3×3 with a single eigenvalue repeated three times.
Example (upper triangular form shown in the notes):

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Properties:

$$\det(A) = 1, \qquad \mathrm{Tr}(A) = 3.$$

Recall:

- The product of the eigenvalues of A equals $\det(A)$.
- The sum of the eigenvalues of A equals $\mathrm{Tr}(A)$ (sum of diagonal entries).

To find eigenvectors solve

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

For the example matrix above the only eigenvalue is $\lambda = 1$ with algebraic multiplicity 3. The corresponding eigenspace (nullspace of $A - I$) has dimension 1 (only one independent eigenvector), so the geometric multiplicity is 1. Therefore A is not diagonalizable in this case.

$n_2 = 0,\ n_3 = 0.$

Therefore any eigenvector has the form

$$\begin{pmatrix} n_1 \\ 0 \\ 0 \end{pmatrix} = n_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Hence $\lambda = 1$ is deficient and A is not diagonalizable.

1. Eigenvectors corresponding to distinct eigenvalues are linearly independent. Hence these can be used as a basis for the column space (or for $\mathbb{R}^n$ when enough eigenvectors exist).

2. Let A be an $n \times n$ matrix and let λ be an eigenvalue with associated eigenvector v. Then

$$Av = \lambda v.$$

For any scalar c,

$$A(cv) = c(Av) = c\lambda v = \lambda(cv),$$

so cv is also an eigenvector (or the zero vector if $c = 0$). If u is another eigenvector corresponding to the same eigenvalue λ, then for scalars c_1, c_2,

$$A(c_1 v + c_2 u) = c_1 Av + c_2 Au = c_1 \lambda v + c_2 \lambda u = \lambda(c_1 v + c_2 u).$$

Thus the set of all vectors associated with the eigenvalue λ is closed under addition and scalar multiplication; i.e. the eigenspace

$$E_\lambda = \{x \in \mathbb{R}^n : Ax = \lambda x\}$$

is a vector subspace of $\mathbb{R}^n$, called the eigensubspace (or eigenspace) corresponding to λ.