Matrix Decomposition

Jo Ciucă
Australian National University
comp36706670@anu.edu.au

Outline

This lecture:

- The Determinant & the Trace
- Eigenvalues & Eigenvectors

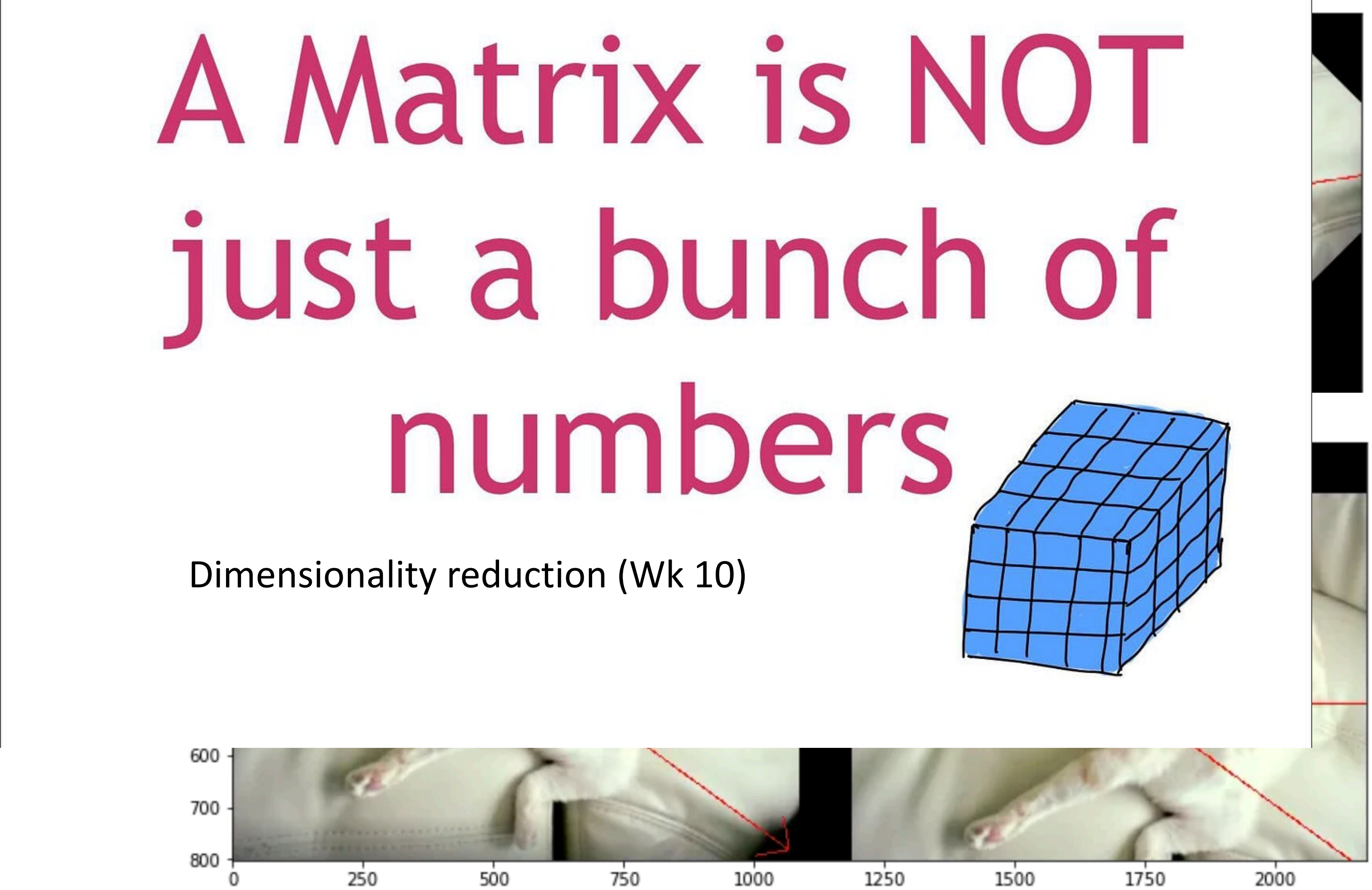
Next lecture (tomorrow):

- Diagonalization & Eigendecomposition
- Singular Value Decomposition

The essence of a Matrix

Cholesky decomposition: Symmetric, PD matrix =LL^T

e.g. covariance matrix of a multivariate Gaussian (Wk 6)



Dimensionality reduction (Wk 10)

The Determinant

- A number associated with a square matrix that essentially “packs” it.

- We write the determinant as $\det(\boldsymbol{A})$ or sometimes as $|\boldsymbol{A}|$ so that

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

- The determinant of a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is a function that maps $\boldsymbol{A}$ onto a real number.

Example 1: Testing for Matrix Invertibility

- If $\boldsymbol{A}$ is a 1×1 matrix, then $\boldsymbol{A} = a \Rightarrow \boldsymbol{A}^{-1} = \frac{1}{a}$. It holds if and only if $a \neq 0$.
- For 2×2 matrices, if $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, recall that the inverse of $\boldsymbol{A}$ is

$$\boldsymbol{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- Hence, $\boldsymbol{A}$ is invertible if and only if

$$a_{11}a_{22} - a_{12}a_{21} \neq 0$$

- This quantity is the determinant of $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$, i.e.,

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

- For a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ it holds that $\boldsymbol{A}$ is invertible if and only if $det(\boldsymbol{A}) \neq 0$.

- We have explicit (closed-form) expressions for determinants of small matrices in terms of the elements of the matrix. For $n = 1$,

$$\det(\boldsymbol{A}) = \det(a_{11}) = a_{11}$$

- For $n = 2$,

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

which we have observed in the preceding example.

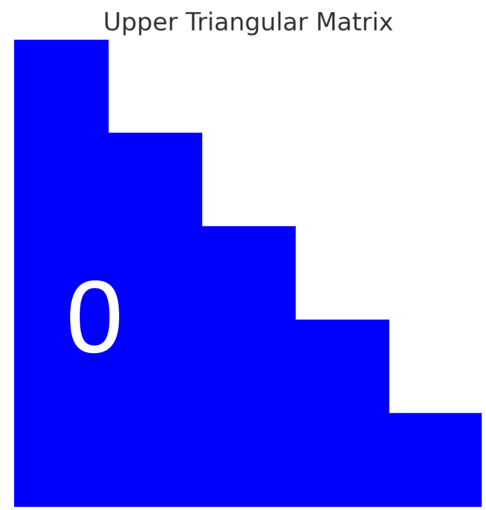
- For $n = 3$ (known as Sarrus' rule),

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$

$$- a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}$$

Triangular Matrices

- We call a square matrix T an upper-triangular matrix if T_{ij} for $i > j$, i.e., the matrix is zero below its diagonal.
- Analogously, we define a lower-triangular matrix as a matrix with zeros above its diagonal.



- For a triangular matrix $T \in \mathbb{R}^{n \times n}$, the determinant is the product of the diagonal elements, i.e.,

$$\det(T) = \prod_{i=1}^{n} T_{ii}$$

Properties of the determinant

1. $\det(I_n) = 1$

2. Exchanging two rows of a matrix reverses the sign of the determinant.

3. The determinant is a linear function.

$$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\begin{vmatrix} a + a' & b + b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$$

Let’s prove these

4. If two rows are equal, the determinant is 0.

5. Adding/subtracting a scaled row from another doesn’t change the determinant.

6. If we have a row of zeros, the determinant is 0.
7. For a triangular matrix, *T*:

$$\det(T) = \prod_{i=1}^{n} T_{ii}$$

We know that:

- Exchanging two rows/columns changes the sign of $\det(\boldsymbol{A})$. (Rule 2).
- Multiplication of a column/row with $\lambda \in \mathbb{R}$ scales $\det(\boldsymbol{A})$ by λ (Rule 3a). In particular, $\det(\lambda \boldsymbol{A}) = \lambda^n \det(\boldsymbol{A})$
- Adding a multiple of a column/row to another one does not change $\det(\boldsymbol{A})$. (Rule 5)

- We can use Gaussian elimination to compute $\det(\boldsymbol{A})$ by bringing $\boldsymbol{A}$ it into row-echelon form. We can stop Gaussian elimination when we have $\boldsymbol{A}$ in a triangular form where the elements below the diagonal are all 0.

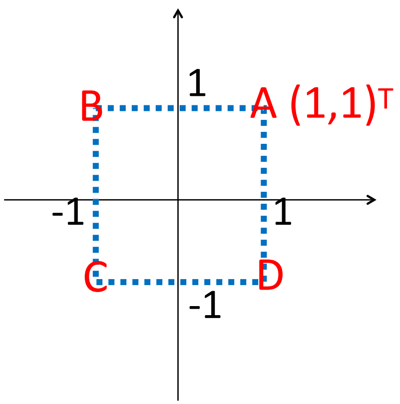
- Recall: the determinant of a triangular matrix is the product of the diagonal elements.

More properties:

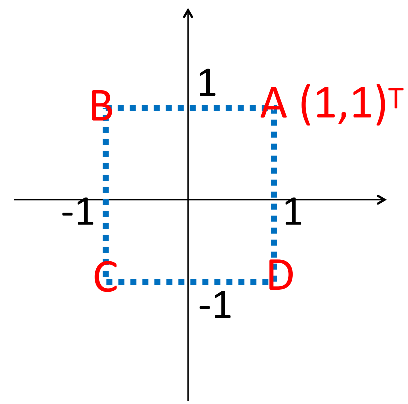
- For $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, we have the following properties:
- $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{A}) \det(\boldsymbol{B})$
- $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$
- If $\boldsymbol{A}$ is regular (invertible), then $\det(\boldsymbol{A}^{-1}) = \dfrac{1}{\det(\boldsymbol{A})}$

Understanding the determinant

- Matrices characterize linear transformations.



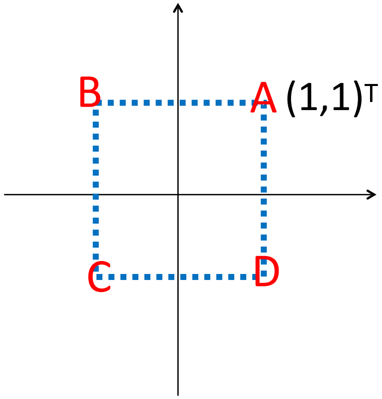
Left multiplied by a matrix



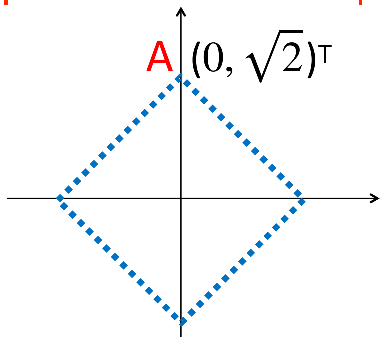
$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Determinant and invertibility

$$\begin{vmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{vmatrix} = 1$$



45° counterclock wise rotation



Theorem 1: Laplace expansion

- Consider a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$. Then, for all $j = 1, \ldots, n$:
- 1. Expansion along the column j

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det\left(\boldsymbol{A}_{\boldsymbol{k},\ \boldsymbol{j}}\right)$$

- *2.* Expansion along row j

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{jk} \det\left(\boldsymbol{A}_{\boldsymbol{j},\boldsymbol{k}}\right)$$

- Here $\boldsymbol{A}_{k,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the submatrix of $\boldsymbol{A}$ that we obtain when deleting row k and column j.

Example 2

2. Expansion along row j

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j}\, a_{jk} \det(\boldsymbol{A}_{j,k})$$

- Let us compute the determinant of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

- Using the **Laplace expansion along the first row**, yielding:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = (-1)^{1+1} \bullet 1 \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} + (-1)^{1+2} \bullet 2 \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix} + (-1)^{1+3} \bullet 3 \begin{vmatrix} 3 & 1 \\ 0 & 0 \end{vmatrix}$$

- We compute the determinants of all the 2×2 matrices and obtain

$$\det(\boldsymbol{A}) = 1(1-0) - 2(3-0) + 3(0-0) = -5$$

- For completeness, we can compare this result to computing the determinant using **Sarrus' rule**:

$$\det(\boldsymbol{A}) = 1 \bullet 1 \bullet 1 + 3 \bullet 0 \bullet 3 + 0 \bullet 2 \bullet 2 - 0 \bullet 1 \bullet 3 - 1 \bullet 0 \bullet 2 - 3 \bullet 2 \bullet 1 = 1 - 6 = -5.$$

Example 2

- Let us use **Gaussian elimination** in order to obtain the following determinant:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \; row2 - 3 \times row1$$

$$\rightsquigarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -7 \\ 0 & 0 & 1 \end{bmatrix}$$

- Now we have the upper triangular form (row-echelon form).

$$\det(\boldsymbol{A}) = 1 \times (-5) \times 1 = -5$$

- We can verify this result with the previous example.

The Trace and its properties

The trace of a square matrix $A \in \mathbb{R}^{n \times n}$ is the sum of the diagonal elements of A.

$$\mathrm{tr}(A) := \sum_{i=1}^{n} a_{ii}$$

Properties:

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$$

$$\mathrm{tr}(\alpha A) = \alpha \, \mathrm{tr}(A)$$

$$\mathrm{tr}(I_n) = n$$

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

Eigenvalues and Eigenvectors

- For $\lambda \in \mathbb{R}$ and a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$

$$p_{\boldsymbol{A}}(\lambda) := \det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right)$$

$$= c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n$$

$c_0, \cdots, c_{n-1} \in \mathbb{R}$, is the **characteristic polynomial** of $\boldsymbol{A}$.

<u>Example</u>:

For $\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$, we have:

$$p_{\boldsymbol{A}}(\lambda) = \det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right) = \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) - 8 = \left(\lambda - 5\right)(\lambda + 1)$$

- The characteristic polynomial $p_{\boldsymbol{A}}(\lambda) := \det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right)$ will allow us to compute eigenvalues and eigenvectors.

Theorem

$\lambda \in \mathbb{R}$ is eigenvalue of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ if and only if λ is a root of the characteristic polynomial $p_{\boldsymbol{A}}(\lambda)$ of $\boldsymbol{A}$.

Example:

- $\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$, we have,

$$p_{\boldsymbol{A}}(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) - 8 = (\lambda - 5)(\lambda + 1)$$

- Eigenvalues are $\lambda_1 = 5$, $\lambda_2 = -1$

Definition:

Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a square matrix. Then $\lambda \in \mathbb{R}$ is an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$ is the corresponding eigenvector of $\boldsymbol{A}$ if

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}.$$

We call this equation the **eigenvalue equation**.

Matrix we are finding the eigenvector/eigenvalue of

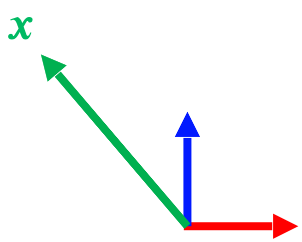
eigenvalue

$$Ax = \lambda x$$

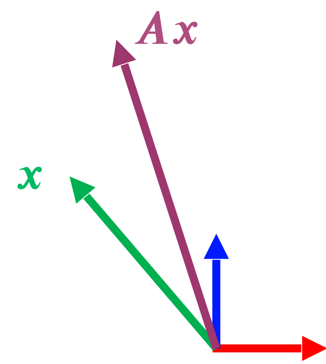
$$(A - \lambda I)x = 0$$

identity matrix

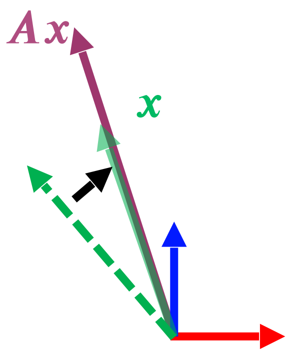
Eigenvalues and Eigenvectors



Left multiplied by a matrix $\boldsymbol{A}$

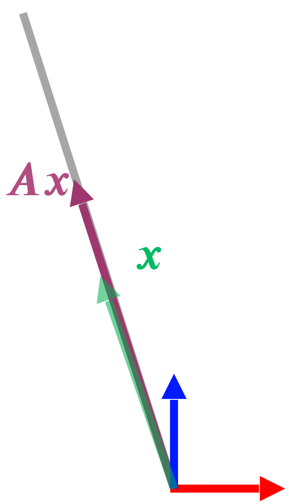


Everything is normal



Now, $\boldsymbol{x}$ and $\boldsymbol{Ax}$ are of the same line. The length of $\boldsymbol{Ax}$ is greater than $\boldsymbol{x}$

$$\boldsymbol{Ax} = \lambda \boldsymbol{x}$$



The grey line is the eigenspace of $\boldsymbol{A}$ with respect to λ
Every vector on this grey line is an eigenvector of $\boldsymbol{A}$, and they all correspond to the eigenvalue λ

The following statements are equivalent:

- λ is an eigenvalue of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$
- There exists an $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ with $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ or equivalently $(\boldsymbol{A} - \lambda\boldsymbol{I}_n)\boldsymbol{x} = \boldsymbol{0}$ can be solved non-trivially, i.e., $\boldsymbol{x} \neq \boldsymbol{0}$.
- $\mathrm{rk}(\boldsymbol{A} - \lambda\boldsymbol{I}_n) < n$
- $\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$

Non-uniqueness of eigenvectors

- If $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ associated with eigenvalue λ, then for any $c \in \mathbb{R}\backslash\{0\}$ it holds that $c\boldsymbol{x}$ is an eigenvector of A with the same eigenvalue since

$$\boldsymbol{A}(c\boldsymbol{x}) = c\boldsymbol{A}\boldsymbol{x} = c\lambda\boldsymbol{x} = \lambda(c\boldsymbol{x})$$

- Thus, all vectors that are collinear (point in the same or opposite direction) to $\boldsymbol{x}$ are also eigenvectors of $\boldsymbol{A}$.

Definition:

Let a square matrix $\boldsymbol{A}$ have an eigenvalue λ_i. The algebraic multiplicity of λ_i is the number of times the root appears in the characteristic polynomial.

<u>Example 3</u>:

- $\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$, we have,

$$p_{\boldsymbol{A}}(\lambda) = \det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right) = \begin{vmatrix} 1-\lambda & 4 \\ 2 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) - 8 = \left(\lambda - 5\right)(\lambda + 1)$$

- Eigenvalues are $\lambda_1 = 5,\ \lambda_2 = -1$
- Hence it has two distinct eigenvalues and each occurs only once, so the algebraic multiplicity of both eigenvalues is one.

Example 4:

- $\boldsymbol{B} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$, we have,

$$p_{\boldsymbol{B}}(\lambda) = \det(\boldsymbol{B} - \lambda \boldsymbol{I}) = \begin{vmatrix} 5-\lambda & 0 \\ 0 & 5-\lambda \end{vmatrix} = (5-\lambda)^2$$

- Eigenvalues are $\lambda_1 = \lambda_2 = 5$.
- The eigenvalue 5 has algebraic multiplicity of 2.

Definition:

For $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, the union of the $\mathbf{0}$ vector and the set of all eigenvectors of $\boldsymbol{A}$ associated with an eigenvalue λ is a subspace of $\mathbb{R}^n$, which is called the eigenspace of $\boldsymbol{A}$ with respect to λ and is denoted by E_λ.

The set of all eigenvalues of $\boldsymbol{A}$ is called the eigenspectrum, or just spectrum, of $\boldsymbol{A}$.

If λ is an eigenvalue of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, then the corresponding eigenspace E_λ is the solution space of the homogeneous system of linear equations $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \mathbf{0}$

Example 6: The case of the Identity Matrix

The identity matrix $\boldsymbol{I} \in \mathbb{R}^{n \times n}$ has characteristic polynomial $p_{\boldsymbol{I}}(\lambda) = \det(\boldsymbol{I} - \lambda \boldsymbol{I}) = (1 - \lambda)^n = 0$. It has only one eigenvalue $\lambda = 1$ that occurs n times.

- Moreover, $\boldsymbol{I}\boldsymbol{x} = \lambda \boldsymbol{x}$ holds for all vectors $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$.

- Therefore, the sole eigenspace E_1 of the identity matrix spans n dimensions, and all n standard basis vectors of $\mathbb{R}^n$ are eigenvectors of $\boldsymbol{I}$.

Useful properties:

- A matrix $\boldsymbol{A}$ and its transpose $\boldsymbol{A}^T$ possess the same eigenvalues, but not necessarily the same eigenvectors.
- Symmetric, positive definite matrices always have positive, real eigenvalues.

$$\forall \boldsymbol{x} \in V \setminus \{\boldsymbol{0}\} : \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > 0$$

$$\boxed{\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \longrightarrow \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} = \boldsymbol{x}^T \lambda \boldsymbol{x} > 0 \longrightarrow \lambda > 0}$$

Example 5 (Computing Eigenvalues, Eigenvectors, and Eigenspaces)

- Let us find the eigenvalues and eigenvectors of the 2×2 matrix

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

- **Step 1: Characteristic Polynomial**. We need to compute the roots of the characteristic polynomial $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$ to find the eigenvalues.

- **Step 2: Eigenvalues.** The characteristic polynomial is

$$p_A(\lambda) = \det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right) = \det\left(\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \begin{vmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} = (4-\lambda)(3-\lambda) - 2 \bullet 1$$

- We factorize the characteristic polynomial and obtain
- $p_A(\lambda) = (4-\lambda)(3-\lambda) - 2 \bullet 1 = 10 - 7\lambda + \lambda^2 = (2-\lambda)\left(5-\lambda\right)$

  giving the roots $\lambda_1 = 2$ and $\lambda_2 = 5$.

- **Step 3: Eigenvectors and Eigenspaces.** From our definition of the eigenvector $\boldsymbol{x} \neq \boldsymbol{0}$, there will be a vector such that $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, i.e., $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$.
- We find the eigenvectors that correspond to these eigenvalues by looking at vectors $\boldsymbol{x}$ such that

$$\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \boldsymbol{x} = \boldsymbol{0}$$

- For $\lambda = 5$ we obtain

$$\begin{bmatrix} 4-5 & 2 \\ 1 & 3-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \boldsymbol{0}$$

- We solve this homogeneous system and obtain a solution space

$$E_5 = \text{span}[\begin{bmatrix} 2 \\ 1 \end{bmatrix}]$$

- This eigenspace is one-dimensional as it possesses a single basis vector.
- Analogously, we find the eigenvector for $\lambda = 2$ by solving

$$\begin{bmatrix} 4-2 & 2 \\ 1 & 3-2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \boldsymbol{x} = \boldsymbol{0}$$

- The corresponding eigenspace is given as

$$E_2 = \text{span}[\begin{bmatrix} 1 \\ -1 \end{bmatrix}]$$

Definition:

Let λ_i be an eigenvalue of a square matrix $\boldsymbol{A}$. Then the geometric multiplicity of λ_i is the number of linearly independent eigenvectors associated with λ_i. In other words, it is the dimensionality of the eigenspace spanned by the eigenvectors associated with λ_i.

- In our previous example, the geometric multiplicity of $\lambda = 5$ and $\lambda = 2$ is 1.
- In another example, the matrix $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ has two repeated eigenvalues $\lambda_1 = \lambda_2 = 2$. The algebraic multiplicity of λ_1 and λ_2 is 2.
- The eigenvalue has only one distinct unit eigenvector $\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and thus geometric multiplicity is 1.

Theorem

The eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ of a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ are linearly independent.

- Eigenvectors of a matrix with n distinct eigenvalues form a basis of $\mathbb{R}^n$.

Definition:

A square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is defective if it possesses fewer than n linearly independent eigenvectors.

- Looking at the eigenspaces of a defective matrix, it follows that the sum of the dimensions of the eigenspaces is less than n.
- A defective matrix cannot have n distinct eigenvalues, as distinct eigenvalues have linearly independent eigenvectors.

Theorem

Given a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, we can always obtain a symmetric, positive semidefinite matrix $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ by defining

$$\boldsymbol{S} := \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}$$

Proof:

- **Symmetry**: $\boldsymbol{S} := \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\right)^{\mathrm{T}} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{\mathrm{T}} = \boldsymbol{S}^{\mathrm{T}}$
- **Positive semidefinite**: $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \left(\boldsymbol{A}\boldsymbol{x}\right)^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{0}$

- If $\mathrm{rk}(\boldsymbol{A}) = n$, then $\boldsymbol{S} := \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ is positive definite.

The Spectral Theorem

If $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is symmetric, there exists an orthonormal basis of the corresponding vector space V consisting of eigenvectors of A, and each eigenvalue is real.

Example 7

- Consider the matrix

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$$

- The characteristic polynomial of A is

$$p_A(\lambda) = (\lambda - 1)^2(\lambda - 7)$$

We obtain the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 7$, where λ_1 is a repeated eigenvalue. Following our standard procedure for computing eigenvectors, we obtain the eigenspaces:

$$E_1 = \mathrm{span}[\ \underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}_{=:\boldsymbol{x}_1},\ \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_2}\ ],\ E_7 = \mathrm{span}[\ \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_3}\ ]$$

- We see that $\boldsymbol{x}_3$ is orthogonal to both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. However, since $\boldsymbol{x}_1^{\mathrm{T}}\boldsymbol{x}_2 = 1 \neq 0$, they are not orthogonal. The spectral theorem states that there exists an orthogonal basis, but the one we have is not orthogonal.
- However, we can construct one.

- To construct such a basis, we exploit the fact that $\boldsymbol{x_1}, \boldsymbol{x_2}$ are eigenvectors associated with the same eigenvalue λ. Therefore, for any $\alpha,\ \beta \in \mathbb{R}$ it holds that

$$\boldsymbol{A}\left(\alpha \boldsymbol{x_1} + \beta \boldsymbol{x_2}\right) = \boldsymbol{A}\boldsymbol{x_1}\alpha + \boldsymbol{A}\boldsymbol{x_2}\beta = \lambda_1\left(\alpha \boldsymbol{x_1} + \beta \boldsymbol{x_2}\right)$$

- i.e., any linear combination of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ is also an eigenvector of $\boldsymbol{A}$ associated with λ_1. The Gram-Schmidt algorithm is a method for iteratively constructing an orthogonal/orthonormal basis from a set of basis vectors using such linear combinations.

- Therefore, even if $\boldsymbol{x_1}$and $\boldsymbol{x_2}$ are not orthogonal, we can apply the Gram-Schmidt algorithm and find eigenvectors associated with $\lambda_1 = 1$ that are orthogonal to each other (and to $\boldsymbol{x_3}$). In our example, we will obtain

$$\boldsymbol{x_1'} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x_2'} = \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

- which are orthogonal to each other, orthogonal to $\boldsymbol{x_3}$, and eigenvectors of $\boldsymbol{A}$ associated with $\lambda_1 = 1$.

Theorems

The determinant of a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is the product of its eigenvalues,

$$\det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i$$

where λ_i are (possibly repeated) eigenvalues of $\boldsymbol{A}$.

The trace of a matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is the sum of its eigenvalues:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} \lambda_i$$

Check your understanding

- The eigenvalues of a projection matrix are 0 and 1.
- The sum of eigenvalues of the permutation matrix $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is 1.

True or False:

- The eigenvalues of $A + 5I$ are the same as the eigenvalues of A . F
- The eigenvectors of $A + 5I$ are the same as the eigenvectors of A . T