Precalculus advanced

Chapter 4:

Matrix transformations
Eigenvalues & Eigenvectors

Foundation Year - 2022/2023
Dr. Grace Younes

Disclaimer: These notes closely follow the material in the textbooks cited in the bibliography, with some additions by the author. These notes should not be distributed or used for commercial purposes.

Outline

Outline

- Informally, a function is a rule that accepts inputs and produces outputs. For instance, $f(x) = x^2$ is a function that accepts one number x as its input, and outputs the square of that number: $f(2) = 4$. In this section, we interpret matrices as functions.

- Hereafter, when there is no ambiguity, we may simply denote a vector like any variable, i.e., without accenting it by a right arrow, that is simply v instead of $\vec{v}$.
- Let A be a matrix with m rows and n columns. Consider the matrix equation $b = Ax$ (we write it this way instead of $Ax = b$ to be reminded of the notation $y = f(x)$). If we vary x, then b will also vary; in this way, we think of A as a function with independent variable x and dependent variable b.
    - The independent variable (the input) is x, which is a vector in $\mathbb{R}^n$.
    - The dependent variable (the output) is b, which is a vector in $\mathbb{R}^m$.
- The set of all possible output vectors are the vectors b such that $Ax = b$ has some solution;
  this is the same as the column space of A

Example (Projection onto the xy-plane)

Let

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

Describe the function $b = Ax$ geometrically.

Solution: In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbb{R}^3$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$$

Multiplication by A simply sets the z-coordinate equal to zero: it projects vertically onto the xy-plane.

Example (Reflection)

Let

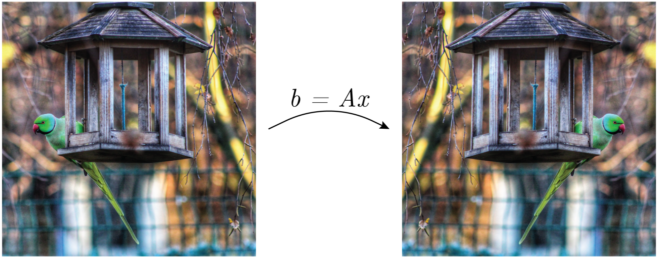
$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution: In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

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

Multiplication by A negates the x-coordinate: it reflects over the y-axis.



Example (Dilation)

Let

$$A = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution: In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1.5x \\ 1.5y \end{pmatrix} = 1.5\begin{pmatrix} x \\ y \end{pmatrix}.$$

Multiplication by A is the same as scalar multiplication by 1.5: it scales or dilates the plane by a factor of 1.5.

$$b = Ax$$

Example (Identity)

Let

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

Describe the function $b = Ax$ geometrically.

Solution: In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

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

Multiplication by A does not change the input vector at all: it is the identity transformation which does nothing.

$b = Ax$

Example (Rotation)

Let

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

Describe the function $b = Ax$ geometrically.

Solution: In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

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

We substitute a few test points in order to understand the geometry of the transformation:
Multiplication by A is *counterclockwise rotation* by 90°.

$b = Ax$

Example (Shear)

Let

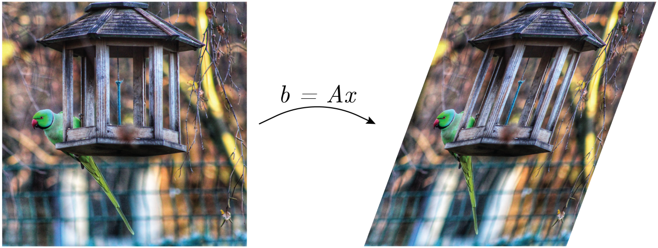
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution: In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

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

Multiplication by A adds the y-coordinate to the x-coordinate; this is called a shear in the x-direction.



DEFINITION

A transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule T that assigns to each vector x in $\mathbb{R}^n$ a vector $T(x)$ in $\mathbb{R}^m$.

- $\mathbb{R}^n$ is called the **domain** of T.
- $\mathbb{R}^m$ is called the **codomain** of T.
- For x in $\mathbb{R}^n$, the vector $T(x)$ in $\mathbb{R}^m$ is the **image** of x under T.
- The set of all images $\{T(x) \mid x \text{ in } \mathbb{R}^n\}$ is the **range** of T.

The notation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ means "T is a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$"

- Now we specialize the general notions and vocabulary from the previous frame to the functions defined by matrices that we considered in the first frame.

Definition

Let A be an $m \times n$ matrix. The matrix transformation associated to A is the transformation

$$T : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad \text{defined by} \quad T(x) = Ax.$$

This is the transformation that takes a vector x in $\mathbb{R}^n$ to the vector Ax in $\mathbb{R}^m$.

- If A has n columns, then it only makes sense to multiply A by vectors with n entries. This is why the domain of $T(x) = Ax$ is $\mathbb{R}^n$.
- If A has n rows, then Ax has m entries for any vector x in $\mathbb{R}^n$; this is why the codomain of $T(x) = Ax$ is $\mathbb{R}^m$.

The definition of a matrix transformation T tells us how to evaluate T on any given vector: we multiply the input vector by a matrix. For instance, let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

and let $T(x) = Ax$ be the associated matrix transformation. Then

$$T\begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix} = A\begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix} = \begin{pmatrix} -14 \\ -32 \end{pmatrix}.$$

- Suppose that A has columns $v_1, v_2, \ldots, v_n$. If we multiply A by a general vector x, we get

$$Ax = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n.$$

This is just a general linear combination of $v_1, v_2, \ldots, v_n$. Therefore, the outputs of $T(x) = Ax$ are exactly the linear combinations of the columns of A :

The range of T is the column space of A.

Let A be an $m \times n$ matrix, and let $T(x) = Ax$ be the associated matrix transformation.

- The domain of T is $\mathbb{R}^n$, where n is the number of columns of A.
- The codomain of T is $\mathbb{R}^m$, where m is the number of rows of A.
- The range of T is the column space of A.

Example

Let

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix}$$

and define $T(x) = Ax$. The domain of T is $\mathbb{R}^3$, and the codomain is $\mathbb{R}^2$. The range of T is the column space; since all three columns are collinear, the range is a line in $\mathbb{R}^2$.

Example

Let

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

and define $T(x) = Ax$. The domain of T is $\mathbb{R}^2$, and the codomain is $\mathbb{R}^3$. The range of T is the column space; since A has two columns which are not collinear, the range is a plane in $\mathbb{R}^3$.

Example (Projection onto the xy-plane)

Let

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

and let $T(x) = Ax$. What are the domain, the codomain, and the range of T ?

Solution: Geometrically, the transformation T projects a vector directly "down" onto the xy-plane in $\mathbb{R}^3$.
The inputs and outputs have three entries, so the domain and codomain are both $\mathbb{R}^3$. The possible outputs all lie on the xy-plane, and every point on the xy plane is an output of T (with itself as the input), so the range of T is the xy-plane.

Be careful not to confuse the codomain with the range here. The range is a plane, but it is a plane in $\mathbb{R}^3$, so the codomain is still $\mathbb{R}^3$. The outputs of T all have three entries; the last entry is simply always zero.

Outline

DEFINITION (ONE-TO-ONE TRANSFORMATIONS)

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if, for every vector b in $\mathbb{R}^m$, the equation $T(x) = b$ has at most one solution x in $\mathbb{R}^n$.

Remark: Another word for one-to-one is injective.

Here are some equivalent ways of saying that T is one-to-one:

- For every vector b in $\mathbb{R}^m$, the equation $T(x) = b$ has zero or one solution x in $\mathbb{R}^n$.
- Different inputs of T have different outputs.
- If $T(u) = T(v)$ then $u = v$.

Here are some equivalent ways of saying that T is not one-to-one:

- There exists some vector b in $\mathbb{R}^m$ such that the equation $T(x) = b$ has more than one solution x in $\mathbb{R}^n$.
- There are two different inputs of T with the same output.
- There exist vectors u, v such that $u \neq v$ but $T(u) = T(v)$.

Theorem (One-to-one matrix transformations)

Let A be an $m \times n$ matrix, and let $T(x) = Ax$ be the associated matrix transformation. The following statements are equivalent:

1. *T is one-to-one.*
2. *For every b in $\mathbb{R}^m$, the equation $T(x) = b$ has at most one solution.*
3. *For every b in $\mathbb{R}^m$, the equation $Ax = b$ has a unique solution or is inconsistent.*
4. *$Ax = 0$ has only the trivial solution.*
5. *The columns of A are linearly independent.*
6. *A has a pivot in every column.*

Example

Let A be the matrix

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

and define $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(x) = Ax$. Is T one-to-one?

Solution: The reduced row echelon form of A is

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

Hence A has a pivot in every column, so T is one-to-one.

EXAMPLE

Let

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

and define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(x) = Ax$. Is T one-to-one? If not, find two different vectors u, v such that $T(u) = T(v)$.

Solution: The matrix A is already in reduced row echelon form. It does not have a pivot in every column, so T is not one-to-one. Therefore, we know from the theorem that $Ax = 0$ has nontrivial solutions. If v is a nontrivial (i.e., nonzero) solution of $Av = 0$, then $T(v) = Av = 0 = A0 = T(0)$, so 0 and v are different vectors with the same output. For instance,

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

Geometrically, T is projection onto the xy-plane. Any two vectors that lie on the same vertical line will have the same projection. For b on the xy-plane, the solution set of $T(x) = b$ is the entire vertical line containing b. In particular, $T(x) = b$ has infinitely many solutions.

EXAMPLE

Let A be the matrix

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

and define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by $T(x) = Ax$. Is T one-to-one? If not, find two different vectors u, v such that $T(u) = T(v)$.

Solution: The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

There is not a pivot in every column, so T is not one-to-one. Therefore, we know from the theorem that $Ax = 0$ has nontrivial solutions. If v is a nontrivial (i.e., nonzero) solution of $Av = 0$, then $T(v) = Av = 0 = A0 = T(0)$, so 0 and v are different vectors with the same output. In order to find a nontrivial solution, we find the parametric form of the solutions of $Ax = 0$ using the reduced matrix above:

$$\begin{cases} x - z &= 0 \\ y + z &= 0 \end{cases} \Longrightarrow \begin{cases} x = z \\ y = -z \end{cases}$$

The free variable is z. Taking $z = 1$ gives the nontrivial solution

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

Wide matrices do not have one-to-one transformations.

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a one-to-one matrix transformation, what can we say about the relative sizes of n and m ?

1. The matrix associated to T has n columns and m rows. Each row and each column can only contain one pivot, so in order for A to have a pivot in every column, it must have at least as many rows as columns: $n \leq m$.
2. This says that, for instance, $\mathbb{R}^3$ is "too big" to admit a one-to-one linear transformation into $\mathbb{R}^2$.

Note that there exist tall matrices that are not one-to-one: for example,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

does not have a pivot in every column.

Definition (Onto transformations)

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is onto if, for every vector b in $\mathbb{R}^m$, the equation $T(x) = b$ has at least one solution x in $\mathbb{R}^n$.

Remark: Another word for onto is surjective.

Here are some equivalent ways of saying that T is onto:

- The range of T is equal to the codomain of T.
- Every vector in the codomain is the output of some input vector.

Here are some equivalent ways of saying that T is not onto:

- The range of T is smaller than the codomain of T.
- There exists a vector b in $\mathbb{R}^m$ such that the equation $T(x) = b$ does not have a solution.
- There is a vector in the codomain that is not the output of any input vector.

Theorem (Onto matrix transformations)

Let A be an $m \times n$ matrix, and let $T(x) = Ax$ be the associated matrix transformation. The following statements are equivalent:

1. *T is onto.*
2. *$T(x) = b$ has at least one solution for every b in $\mathbb{R}^m$.*
3. *$Ax = b$ is consistent for every b in $\mathbb{R}^m$.*
4. *The columns of A span $\mathbb{R}^m$.*
5. *A has a pivot in every row.*

EXAMPLE

Let A be the matrix

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

and define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by $T(x) = Ax$. Is T onto?

Solution: The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Hence A has a pivot in every row, so T is onto.

Example

Let A be the matrix

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

and define $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(x) = Ax$. Is T onto? If not, find a vector b in $\mathbb{R}^3$ such that $T(x) = b$ has no solution.

Solution: The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Hence A does not have a pivot in every row, so T is not onto. In fact, since

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ x \end{pmatrix},$$

we see that for every output vector of T, the third entry is equal to the first. Therefore,

$$b = (1, 2, 3)$$

is not in the range of T.

Tall matrices do not have onto transformations.

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is an onto matrix transformation, what can we say about the relative sizes of n and m ?

1. The matrix associated to T has n columns and m rows. Each row and each column can only contain one pivot, so in order for A to have a pivot in every row, it must have at least as many columns as rows: $m \leq n$.
2. This says that, for instance, $\mathbb{R}^2$ is "too small" to admit an onto linear transformation to $\mathbb{R}^3$.

Note that there exist wide matrices that are not onto: for example,

$$\begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix}$$

does not have a pivot in every row.

Comparison

Let A be an $m \times n$ matrix, and $T : \mathbb{R}^n \to \mathbb{R}^m$ is the matrix transformation $T(x) = Ax$.

T is **one-to-one**	T is **onto**
$T(x) = b$ has at most one solution for every b	$T(x) = b$ has at least one solution for every b
The columns of A are linearly independent	The columns of A span $\mathbb{R}^m$.
A has a pivot in every column.	A has a pivot in every row.

One-to-one is the same as onto for square matrices.

- a square matrix has a pivot in every row if and only if it has a pivot in every column.
- Therefore, a matrix transformation T from $\mathbb{R}^n$ to itself is one-to-one if and only if it is onto: in this case, the two notions are equivalent.
- Conversely, by this note and this note, if a matrix transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is both one-to-one and onto, then $m = n$.

Note that in general, a transformation T is both one-to-one and onto if and only if $T(x) = b$ has exactly one solution for all b in $\mathbb{R}^m$.

Definition

Let A be an $n \times n$ matrix.

1. An eigenvector of A is a nonzero vector v in $\mathbb{R}^n$ such that $Av = \lambda v$, for some scalar λ
2. An eigenvalue of A is a scalar λ such that the equation $Av = \lambda v$ has a nontrivial solution.

If $Av = \lambda v$ for $v \neq 0$, we say that λ is the eigenvalue for v, and that v is an eigenvector for λ.

Remarks:

- The German prefix "eigen" roughly translates to "self" or "own". An eigenvector of A is a vector that is taken to a multiple of itself by the matrix transformation $T(x) = Ax$, which perhaps explains the terminology. On the other hand, "eigen" is often translated as "characteristic"; we may think of an eigenvector as describing an intrinsic, or characteristic, property of A.
- Eigenvalues and eigenvectors are only for square matrices.
- Eigenvectors are by *definition nonzero*. Eigenvalues may be equal to zero. In fact, we do not consider the zero vector to be an eigenvector: since $A0 = 0 = \lambda 0$ for *every* scalar λ, the associated eigenvalue would be undefined.

Example

Consider the matrix

$$A = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \quad \text{and vectors} \quad v = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Which are eigenvectors? What are their eigenvalues?

Solution:

- We have

$$Av = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4v.$$

  Hence, v is an eigenvector of A, with eigenvalue $\lambda = 4$.

- On the other hand,

$$Aw = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}.$$

  which is not a scalar multiple of w. Hence, w is not an eigenvector of A.

Use this link to view the online demo

Example

Consider the matrix

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \quad \text{and vectors} \quad v = \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix} \quad w = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

Which are eigenvectors? What are their eigenvalues?

Solution:

- We have

$$Av = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 32 \\ 8 \\ 2 \end{pmatrix} = 2v.$$

  Hence, v is an eigenvector of A, with eigenvalue $\lambda = 2$.

- On the other hand,

$$Aw = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 28 \\ 1 \\ 1 \end{pmatrix},$$

  which is not a scalar multiple of w. Hence, w is not an eigenvector of A.

Example

Let

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \quad v = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Is v an eigenvector of A ? If so, what is its eigenvalue?

Solution: The product is

$$Av = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0v.$$

Hence, v is an eigenvector with eigenvalue zero.
As noted above, an eigenvalue is allowed to be zero, but an eigenvector is not.

Use this link to view the online demo

Fact: Eigenvectors with distinct eigenvalues are linearly independent

- Let $v_1, v_2, \ldots, v_k$ be eigenvectors of a matrix A, and suppose that the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct (all different from each other).
- Then $\{v_1, v_2, \ldots, v_k\}$ is linearly independent.

- Let A be an $n \times n$ matrix, and let λ be a scalar. The eigenvectors with eigenvalue λ, if any, are the nonzero solutions of the equation $Av = \lambda v$. We can rewrite this equation as follows:

$$\begin{aligned} & Av = \lambda v \\ \iff & Av - \lambda v = 0 \\ \iff & Av - \lambda I_n v = 0 \\ \iff & (A - \lambda I_n)\, v = 0. \end{aligned}$$

- Therefore, the eigenvectors of A with eigenvalue λ, if any, are the nontrivial solutions of the matrix equation $(A - \lambda I_n)\, v = 0$, i.e., the nonzero vectors in $\mathrm{Nul}\,(A - \lambda I_n)$.
- If the equation $(A - \lambda I_n)\, v = 0$ has no nontrivial solutions, then λ is not an eigenvalue of A.

- In other words,

$$\lambda \text{ is an eigenvalue of } A$$

if and only if

$$(A - \lambda I_n)\, v = 0 \text{ has non trivial solutions}$$

if and only if

$$\det (A - \lambda I_n) = 0$$

Definition

Let A be an $n \times n$ matrix, and let λ be an eigenvalue of A. The λ-**eigenspace** of A is the solution set of vectors such that

$$(A - \lambda I_n)\, v = 0$$

i.e., the subspace $\mathrm{Nul}\,(A - \lambda I_n)$.

Outline

Definition

Let A be an $n \times n$ matrix. The characteristic polynomial of A is the function $f(\lambda)$ given by

$$f(\lambda) = \det(A - \lambda I_n).$$

EXAMPLE

Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution: We have

$$
\begin{aligned}
f(\lambda) = \det(A - \lambda I_2) &= \det\left(\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) \\
&= \det\begin{pmatrix} 5-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\
&= (5-\lambda)(1-\lambda) - 2\cdot 2 = \lambda^2 - 6\lambda + 1.
\end{aligned}
$$

EXAMPLE

Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Solution: We compute the determinant by expanding cofactors along the third column:

$$f(\lambda) = \det(A - \lambda I_3) = \det\begin{pmatrix} -\lambda & 6 & 8 \\ \frac{1}{2} & -\lambda & 0 \\ 0 & \frac{1}{2} & -\lambda \end{pmatrix}$$
$$= 8\left(\frac{1}{4} - 0 \cdot -\lambda\right) - \lambda\left(\lambda^2 - 6 \cdot \frac{1}{2}\right)$$
$$= -\lambda^3 + 3\lambda + 2.$$

The Characteristic Polynomial

Theorem (Eigenvalues are roots of the characteristic polynomial)

Let A be an $n \times n$ matrix, and let $f(\lambda) = \det(A - \lambda I_n)$ be its characteristic polynomial. Then a number λ_0 is an eigenvalue of A if and only if $f(\lambda_0) = 0$.

Proof. The matrix equation $(A - \lambda_0 I_n)\,x = 0$ has a nontrivial solution if and only if $\det(A - \lambda_0 I_n) = 0$. Therefore,

$$\begin{aligned}
\lambda_0 \text{ is an eigenvalue of } A &\iff Ax = \lambda_0 x \text{ has a nontrivial solution} \\
&\iff (A - \lambda_0 I_n)\,x = 0 \text{ has a nontrivial solution} \\
&\iff A - \lambda_0 I_n \text{ is not invertible} \\
&\iff \det(A - \lambda_0 I_n) = 0 \\
&\iff f(\lambda_0) = 0
\end{aligned}$$

Example

Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution: In the above example we computed the characteristic polynomial of A to be $f(\lambda) = \lambda^2 - 6\lambda + 1$. We can solve the equation $\lambda^2 - 6\lambda + 1 = 0$ using the quadratic formula:

$$\lambda = \frac{6 \pm \sqrt{36 - 4}}{2} = 3 \pm 2\sqrt{2}.$$

Therefore, the eigenvalues are $3 + 2\sqrt{2}$ and $3 - 2\sqrt{2}$.
To compute the eigenvectors, we solve the homogeneous system of equations $(A - \lambda I_2)x = 0$ for each eigenvalue λ. When $\lambda = 3 + 2\sqrt{2}$, we have

Solution:

$$A-(3+\sqrt{2})I_2=\begin{pmatrix} 2-2\sqrt{2} & 2 \\ 2 & -2-2\sqrt{2} \end{pmatrix}$$

$$\xrightarrow{R_1=R_1\times(2+2\sqrt{2})}\begin{pmatrix} -4 & 4+4\sqrt{2} \\ 2 & -2-2\sqrt{2} \end{pmatrix}$$

$$\xrightarrow{R_2=R_2+R_1/2}\begin{pmatrix} -4 & 4+4\sqrt{2} \\ 0 & 0 \end{pmatrix}$$

$$\xrightarrow{R_1=R_1/-4}\begin{pmatrix} 1 & -1-\sqrt{2} \\ 0 & 0 \end{pmatrix}$$

The parametric form of the general solution is $x=(1+\sqrt{2})y$, so the $(3+2\sqrt{2})$ eigenspace is the line spanned by $\begin{pmatrix} 1+\sqrt{2} \\ 1 \end{pmatrix}$. We compute in the same way that the $(3-2\sqrt{2})$-eigenspace is the line spanned by $\begin{pmatrix} 1-\sqrt{2} \\ 1 \end{pmatrix}$.

DEFINITION

The trace of a square matrix A is the number $\mathrm{Tr}(A)$ obtained by summing the diagonal entries of A :

$$\mathrm{Tr}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{pmatrix} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Theorem

Let A be an $n \times n$ matrix, and let $f(\lambda) = \det(A - \lambda I_n)$ be its characteristic polynomial. Then $f(\lambda)$ is a polynomial of degree n. Moreover, $f(\lambda)$ has the form

$$f(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \operatorname{Tr}(A) \lambda^{n-1} + \cdots + \det(A)$$

In other words, the coefficient of λ^{n-1} is $\pm \operatorname{Tr}(A)$, and the constant term is $\det(A)$ (the other coefficients are just numbers without names).

Proof.

First we notice that

$$f(0) = \det(A - 0I_n) = \det(A),$$

so that the constant term is always $\det(A)$.
We will prove the rest of the theorem only for 2×2 matrices; We can write a 2×2 matrix as $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$; then

$$f(\lambda) = \det(A - \lambda I_2) = \det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)(d-\lambda) - bc$$
$$= \lambda^2 - (a+d)\lambda + (ad - bc) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$$

□

Recipe: The characteristic polynomial of a 2×2 matrix. When $n = 2$, the previous theorem tells us all of the coefficients of the characteristic polynomial:

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$$

This is generally the fastest way to compute the characteristic polynomial of a 2×2 matrix

EXAMPLE

Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$$

Solution: We have

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - (5+1)\lambda + (5\cdot 1 - 2\cdot 2) = \lambda^2 - 6\lambda + 1,$$

as in the above example.

Remark: By the above theorem, the characteristic polynomial of an $n \times n$ matrix is a polynomial of degree n. Since a polynomial of degree n has at most n roots, this gives a proof of the fact that an $n \times n$ matrix has at most n eigenvalues.

Eigenvalues of a triangular matrix It is easy to compute the determinant of an upper- or lower-triangular matrix; this makes it easy to find its eigenvalues as well.

Corollary

If A is an upper- or lower-triangular matrix, then the eigenvalues of A are its diagonal entries.

Proof.

Suppose for simplicity that A is a 3×3 upper-triangular matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

Its characteristic polynomial is

$$f(\lambda) = \det(A - \lambda I_3) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{pmatrix}$$

This is also an upper-triangular matrix, so the determinant is the product of the diagonal entries:

$$f(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda).$$

The zeros of this polynomial are exactly a_{11}, a_{22}, a_{33}. □

EXAMPLE

Find the eigenvalues of the matrix

$$A = \begin{pmatrix} 1 & 7 & 2 & 4 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & \pi & 101 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Solution: The eigenvalues are the diagonal entries $1, \pi, 0$. (The eigenvalue 1 occurs twice, so it is said to be of multiplicity 2)

Kuttler, K.,
Elementary linear algebra. The Saylor Foundation, 2012.

Strang, G.,
Linear algebra and its applications. Belmont, CA: Thomson, Brooks/Cole, 2006.

OpenStax,
Calculus Volumes 1, 2, and 3

Margalit, D., Rabinoff, J. and Rolen, L.,
"Interactive linear algebra." Georgia Institute of Technology (2017).