AI 비전공자를 위한 기초 수학 1: 선형 대수

Math for AI Beginner Part 1: Linear Algebra

Week 5: Determinant of Square Matrix and Eigenvalue Problem

Linear Algebra

What we will do during the first 4 weeks…

Review of matrices

Systems of linear algebraic equations AX=B

Finding solution(s) by row operations

Inverse of a square matrix A

Finding inverse by row operations

Determinant of a square matrix A

Calculating determinant by row operations

Bigger pictures: how are these 3 topics connected together with AI?

In between, we will look at vector space and ***linearly independent vectors.***

Matrix eigen-problem

Diagonalisation problem

AI Applications: Deep Learning & SVM

Example:

A and B are two $N \times N$ matrices which are invertible.

Show that: $(AB)^{-1} = B^{-1}A^{-1}$.

Note that $(AB)^{-1}$ denotes "the inverse of AB".

So, what we are asked to do is this:

Prove that the "inverse of AB" is $B^{-1}A^{-1}$.

$(AB)\ (B^{-1}A^{-1})$ *Is this equal to* I*?*

$$= A\left(B\ B^{-1}\right)A^{-1}$$

$$= A\,\mathrm{I}\,A^{-1} = A\,A^{-1} = \mathrm{I}$$

Thus:

$(AB)^{-1} = B^{-1}A^{-1}$

We will now look at the idea of the determinant of a square matrix.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & a_{N3} & \cdots & a_{NN} \end{pmatrix}$$

The elements a_{ij} of the matrix A can be used to compute a number called the determinant of A.

det(A)

|A|

For a starting point, our approach is to define det(A) in terms of determinants of smaller matrices.

The smallest square matrix one can find is of order 1×1.

Let A = (a) then we define det(A) = a.

We calculate det(A) using the *I*-th row of A.

I-th row of A

$$a_{I1} \quad a_{I2} \quad a_{I3} \quad \cdots a_{Ij} \quad \cdots \quad a_{IN}$$

$$\det(\mathbf{A}) = \sum_{j=1}^{N} (-1)^{I+j} a_{Ij} \det(\mathbf{A}_{Ij})$$

sum over all the columns in the row

(N–1)×(N–1) matrix obtained by deleting the *I*-th row and the *j*-th column of A

It can be shown that no matter which row we use to do the calculation the value of det(A) is the same.

We calculate det(A) using the J-th column of A.

J-th column of A

$$\begin{array}{c} a_{1J} \\ a_{2J} \\ \vdots \\ \boxed{a_{iJ}} \\ \vdots \\ a_{NJ} \end{array}$$

sum over all the rows in the column

$(N-1)\times(N-1)$ matrix obtained by deleting the i-th row and the J-th column of A

$$\det(\mathbf{A}) = \sum_{i=1}^{N} (-1)^{i+J}\, a_{iJ}\, \det(\mathbf{A}_{iJ})$$

It can be shown that no matter which column we use to do the calculation the value of det(A) is the same. The value is also the same as the one calculated using "row expansion".

Take $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Calculating det(A) using first row:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \det(A) = (-1)^{1+1}\, a \det(d) + (-1)^{1+2}\, b \det(c) = ad - bc$$

Calculating det(A) using second column:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \det(A) = (-1)^{2+1}\, b \det(c) + (-1)^{2+2}\, d \det(a) = -bc + ad$$

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = (-1)^{1+1} a \det\begin{pmatrix} e & f \\ h & i \end{pmatrix} + (-1)^{1+2} b \det\begin{pmatrix} d & f \\ g & i \end{pmatrix} + (-1)^{1+3} c \det\begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

$$= a(ei - fh) - b(di - fg) + c(dh - eg)$$

$$= aei + bfg + cdh - bdi - ceg - afh$$

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - bdi - afh - ceg$$

$$\begin{array}{ccccc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}$$

There is no such secret formula for the determinants of 4×4 or larger square matrices!

It is easy to compute the determinant of an upper triangular matrix.

$$\det\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} \\ 0 & u_{22} & u_{23} & u_{24} & u_{25} \\ 0 & 0 & u_{33} & u_{34} & u_{35} \\ 0 & 0 & 0 & u_{44} & u_{45} \\ 0 & 0 & 0 & 0 & u_{55} \end{pmatrix} = (-1)^{1+1} u_{11} \det\begin{pmatrix} u_{22} & u_{23} & u_{24} & u_{25} \\ 0 & u_{33} & u_{34} & u_{35} \\ 0 & 0 & u_{44} & u_{45} \\ 0 & 0 & 0 & u_{55} \end{pmatrix}$$

$$= u_{11}(-1)^{1+1} u_{22} \det\begin{pmatrix} u_{33} & u_{34} & u_{35} \\ 0 & u_{44} & u_{45} \\ 0 & 0 & u_{55} \end{pmatrix}$$

$$\vdots$$

$$= u_{11}u_{22}u_{33}u_{44}u_{55}$$

The determinant of an upper (or lower) triangular matrix is the product of all the diagonal elements of the matrix.

To calculate the determinant of a matrix A, can we first perform row operations on A to obtain an upper triangular matrix U, work out det(U) easily, and deduce what det(A) is?

If we perform a legitimate row operation on A to obtain B, there is a simple relation between det(A) and det(B).

$$\mathbf{A} \xrightarrow{R_i \leftrightarrow R_j \text{ (once)}} \mathbf{B} \qquad (-1)\det(\mathbf{A}) = \det(\mathbf{B})$$

$$\mathbf{A} \xrightarrow[\text{(once)}]{R_i \to \alpha R_i + \beta R_j} \mathbf{B} \qquad \alpha \det(\mathbf{A}) = \det(\mathbf{B})$$

Example:

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 3 & 4 & 5 \\ -1 & 2 & 1 & 1 & 2 \\ -1 & 1 & 2 & 2 & 3 \\ 0 & 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\Downarrow$$

$$\mathbf{U} = \begin{pmatrix} -1 & 2 & 1 & 1 & 2 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & -8 \end{pmatrix}$$

All the row operations used are:

$$R_2 \leftrightarrow R_1$$
$$R_4 \leftrightarrow R_3$$
$$R_4 \rightarrow R_4 - R_1$$
$$R_5 \rightarrow R_5 + R_1$$
$$R_4 \rightarrow 2R_4 + R_2$$

$$R_5 \rightarrow 2R_5 - 3R_2$$
$$R_4 \rightarrow R_4 - 5R_3$$

$$R_5 \rightarrow R_5 + 5R_3$$
$$R_5 \rightarrow R_5 + 3R_4$$

To find the determinant of A:

$$(-1)(-1)(1)(1)(2)(2)(1)(1)(1)\det(\mathbf{A})$$
$$= \det(\mathbf{U})$$
$$= (-1)(2)(1)(1)(-8)$$

AI 비전공자를 위한 기초 수학 1: 선형 대수학

Math for AI Beginner Part 1: Linear Algebra

Week 5: Determinant of Square Matrix and Eigenvalue Problem

The Big Picture

Solving AX = B	Finding inverse of A	Finding det(A)
$\mathbf{A} \mid \mathbf{B}$	$\mathbf{A} \mid \mathbf{I}$	$\mathbf{A}$
$\downarrow$	$\downarrow$	$\downarrow$
$\mathbf{U} \mid \mathbf{C}$	$\mathbf{U} \mid \mathbf{P}$	$\mathbf{U}$
	$\downarrow$	$c\det(\mathbf{A}) = \det(\mathbf{U})$
	$\mathbf{I} \mid \mathbf{V}$ (if possible)	

If all diagonal elements of U are not zero, then AX=B has a unique solution, A is invertible and det(A) is not zero.

If U has a zero diagonal element, then det(A)=0, A does not have an inverse and AX=B has either no solution or infinitely many solutions.

We can say something about det(A) and the solutions of the homogeneous system AX=0.

If det(A) is not zero then AX=0 has a unique solution given by X=0.

If det(A) is zero then AX=0 has infinitely many solutions.

If AX=0 has a unique solution X=0 then det(A) is not zero.

If AX=0 has infinitely many solutions then det(A)=0.

Example:

Are the vectors $\begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 7 \\ -1 \end{pmatrix}$ **linearly independent?**

Does the homogeneous system

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 2 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ 3 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c_4 \begin{pmatrix} 1 \\ 2 \\ 7 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

have a unique solution?

If the system has a unique solution then the vectors are linearly independent. Otherwise, they are linearly dependent.

$$c_1\begin{pmatrix}1\\2\\2\\0\end{pmatrix}+c_2\begin{pmatrix}1\\0\\3\\1\end{pmatrix}+c_3\begin{pmatrix}1\\0\\0\\0\end{pmatrix}+c_4\begin{pmatrix}1\\2\\7\\-1\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

$$\begin{aligned} c_1+c_2+c_3+c_4&=0\\ 2c_1+2c_4&=0\\ 2c_1+3c_2+7c_4&=0\\ c_2-c_4&=0 \end{aligned}$$

$$\begin{pmatrix}1&1&1&1\\2&0&0&2\\2&3&0&7\\0&1&0&-1\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\\c_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

$$\det\begin{pmatrix}1&1&1&1\\2&0&0&2\\2&3&0&7\\0&1&0&-1\end{pmatrix}$$

$$=(-1)^{3+1}(1)\det\begin{pmatrix}2&0&2\\2&3&7\\0&1&-1\end{pmatrix}$$

$$=-16$$

The system has a unique solution. Thus, the given vectors are linearly independent.

Linear Algebra

What we will do during the first 4 weeks…

Review of matrices

Systems of linear algebraic equations AX=B

Finding solution(s) by row operations

Inverse of a square matrix A

Finding inverse by row operations

Determinant of a square matrix A

Calculating determinant by row operations

Bigger pictures: how are these 3 topics connected together with AI?

In between, we will look at vector space and ***linearly independent vectors.***

Matrix eigen-problem

Diagonalisation problem

AI Applications: Deep Learning & SVM

We will now look at two related problems: matrix eigen-problem and the diagonalisation of a square matrix.

These two problems have many applications:

- vibration in complicated spring-mass systems
- finding principal strains and stresses
- finding principal axes
- modeling population growth
- Queue theory in Financial Engineering
- Machine Learning Algorithm (Deep Learning)

Review

The homogeneous system $AX=O$ has infinitely many solution if and only of $\det(A)=0$.

The matrix eigen-problem may be stated as follows:

Given an $N \times N$ matrix $\mathbf{A}$, can we find $N \times 1$ matrix $\mathbf{X}$ such that $\mathbf{AX} = \lambda\mathbf{X}$, where λ is a real or complex number?

Obviously $\mathbf{X} = \mathbf{0}$ satisfies $\mathbf{AX} = \lambda\mathbf{X}$ for any λ !

But we are more interested in finding $\mathbf{X} \neq \mathbf{0}$.

$N\times 1$ matrix $N\times N$ identity matrix

$$\mathbf{AX} = \lambda\mathbf{X} = \lambda\mathbf{IX}$$

$N\times N$ matrix number

$$\Rightarrow \quad \mathbf{AX} - \lambda\mathbf{IX} = \mathbf{0}$$

$$\Rightarrow \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$$

Regard this as a homogeneous system of linear algebraic equations

For more than just the trivial solution X=0, it is required that

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

We can find non-trivial $\mathbf{X}$ such that $\mathbf{AX} = \lambda\mathbf{X}$ only for λ which satisfies $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

To solve the eigen-matrix problem, we find λ first from:

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

characteristic (polynomial) equation of A

Values of λ satisfying the characteristic equation are called eigenvalues of $\mathbf{A}$.

If $\mathbf{A}$ is an *N*×*N* matrix, the characteristic equation is a polynomial equation of order *N*.

The matrix $\mathbf{A}$ can then have up to *N* distinct eigenvalues.

If λ is known then we may solve the homogeneous system

$$(A-\lambda I)\, X = 0$$

to find non-trivial X for the eigen-problem.

We refer to all the solutions X of the above homogeneous system as eigenvectors of A corresponding to the eigenvalue λ.

Example:

Find eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$.

$$\det[\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}] = 0$$

$$\Rightarrow \quad \det \begin{pmatrix} 3-\lambda & 2 \\ 1 & 2-\lambda \end{pmatrix} = 0$$

$$(3-\lambda)(2-\lambda) - 2 = 0$$

$$\Rightarrow (\lambda - 1)(\lambda - 4) = 0$$

The eigenvalues of A are λ=1 and λ=4.

To find the eigenvectors, we have to solve the homogeneous system

$$\begin{pmatrix} 3-\lambda & 2 \\ 1 & 2-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$ for each eigenvalue.

$\lambda = 1$

$$\begin{array}{cc|c} 2 & 2 & 0 \\ 1 & 1 & 0 \end{array}$$

Easy to see that if we let $y = t$ then $x = -t$.

All eigenvectors corresponding to the eigenvalue 1 are given by $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -t \\ t \end{pmatrix}$.

$\lambda = 4$

$$\begin{array}{cc|c} -1 & 2 & 0 \\ 1 & -2 & 0 \end{array}$$

Easy to see that if we let $y = s$ then $x = 2s$.

All eigenvectors corresponding to the eigenvalue 4 are given by $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2s \\ s \end{pmatrix}$.

Example:

Find eigenvalues and eigenvectors of

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

The eigenvalues are 1, –1 and 2.

The characteristic equation is

$$\det \begin{pmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 1 \\ 1 & 1 & 0-\lambda \end{pmatrix} = 0$$

$$\Rightarrow \quad \lambda(1-\lambda)^2 + 2(1-\lambda) = 0$$

$$\Rightarrow \quad (1-\lambda)(\lambda-2)(\lambda+1) = 0$$

To find eigenvectors, we have to solve the homogeneous system:

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

$$\lambda = 1$$

$$\left(\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{array}\right)$$

First 2 rows imply *z* = 0.

Let *y* = *t* . Last row implies that *x* = − *y* + *z* = − *t* .

All eigenvectors corresponding to eigenvalue 1 are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -t \\ t \\ 0 \end{pmatrix}.$$

Note that t represents an arbitrary real number.

To find eigenvectors, we have to solve the homogeneous system:

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

$$\lambda = -1$$

$$\left(\begin{array}{ccc|c} 2 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 2 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Let $z = p$.

2^{nd} row implies that $2y + z = 0$. Thus, $y = -p/2$.

1^{st} row implies that $2x + z = 0$. Thus, $x = -p/2$.

Eigenvectors corresponding to eigenvalue -1 are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -p/2 \\ -p/2 \\ p \end{pmatrix}$$

To find eigenvectors, we have to solve the homogeneous system:

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

$$\lambda = 2$$

$$\left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & -2 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Let $z = s$.

2nd row implies that $-y + z = 0$.
Thus, $y = s$.

1st row implies that $-x + z = 0$.
Thus, $x = s$.

All eigenvectors corresponding to eigenvalue 2 are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} s \\ s \\ s \end{pmatrix}$$

Example:

Find eigenvalues and eigenvectors of

$$\mathbf{A} = \begin{pmatrix} \mathbf{3} & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 3 \end{pmatrix}$$

This matrix has only one eigenvalue, that is, 3.

We say that the eigenvalue 3 has a multiplicity of 4.

From the characteristic equation, we may think of the matrix A as having 4 eigenvalues with the same value 3.

The characteristic equation is

$$\det \begin{pmatrix} 3-\lambda & 0 & 0 & 0 \\ 0 & 3-\lambda & 0 & 0 \\ 1 & 1 & 3-\lambda & 0 \\ 1 & 1 & 1 & 3-\lambda \end{pmatrix} = 0 \quad \Rightarrow \quad (3-\lambda)^4 = 0$$

To find the eigenvectors, we have to solve

$$\begin{pmatrix} 3-\lambda & 0 & 0 & 0 \\ 0 & 3-\lambda & 0 & 0 \\ 1 & 1 & 3-\lambda & 0 \\ 1 & 1 & 1 & 3-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The 4^{th} row gives $x + y + z = 0$.

The 3^{rd} row gives $x + y = 0$.

Thus, $z = 0$.

If we let $x = t$ then $y = -t$.

$\lambda = 3$

$$\left(\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{array}\right)$$

What about w?

It can be anything!

Let $w = s$.

The eigenvectors are given by $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} t \\ -t \\ 0 \\ s \end{pmatrix}$

Summary

Can we find $\mathbf{X} \neq 0$ such that $\mathbf{AX} = \lambda\mathbf{X}$?

Yes, provided that λ satisfies $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

Values of λ satisfying $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ are called *eigenvalues* of $\mathbf{A}$.

For each eigenvalue λ, we solve the homogeneous system

$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = \mathbf{0}$ to find eigenvectors of $\mathbf{A}$.