Linear algebra: determinants and eigenvalues/eigenvectors

Statistical Natural Language Processing 1

Çağrı Çöltekin

University of Tübingen
Seminar für Sprachwissenschaft

Winter Semester 2024/2025

version: 7d87f40 @2024-11-28

Quick recap

So far we reviewed:

- Vectors, matrices
- Operations on vectors and matrices: scalar multiplication, addition, dot product, matrix multiplication
- Matrices as operators (linear functions / transformations)
- Linearity and linear combinations
- Solving systems of linear equations, elimination
- Finding matrix inverse
- Linear regression

Today's plan

- Determinant
- Eigenvalues and eigenvectors

Determinant

- The determinant of a square matrix is a number that provides a lot of information about the matrix
  - Whether the matrix has an inverse or not
  - Calculating eigenvalues and eigenvectors
  - Solving systems of linear equations
  - Determining the (signed) 'change of volume' caused by the linear transformation defined by the matrix

Calculating the determinant

- The determinant of a 2x2 matrix is

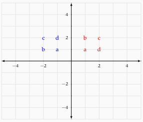
$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb$$

- The determinant of larger matrices are defined recursively
  - Choose a row or column
  - The determinant is the sum of the each element in the row (or column) multiplied by its *cofactor*
  - The cofactor of an element a_{ij} is the determinant of 'sub-matrix' (or *minor*) multiplied by -1^{i+j}
  - The minor of a_{ij} is the matrix obtained by removing row i and column j from the original matrix

Determinant

example geometric interpretation (1)

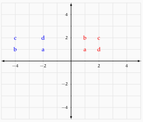
- $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
- $\det(A) = ?$



Determinant

example geometric interpretation (2)

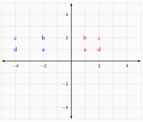
- $A = \begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}$
- $\det(A) = ?$



Determinant

example geometric interpretation (3)

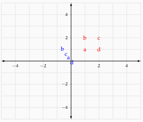
- $A = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$
- $\det(A) = ?$



Determinant

example geometric interpretation (3)

- $A = \begin{bmatrix} \cos 120 \\ \sin 120 \end{bmatrix} \times \begin{bmatrix} \cos 120 & \sin 120 \end{bmatrix} = \begin{bmatrix} 0.25 & -0.43 \\ -0.43 & 0.75 \end{bmatrix}$
- $\det(A) = ?$



Some properties of determinants

- $\det(\mathbf{I}) = 1$
- If two columns or rows are the same, the determinant is 0
- If we multiply a row $\mathbf{A}$ with a scalar c, determinant becomes $c \det A$
- $\begin{vmatrix} a + a' & b + b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$
- If we exchange two rows of $\mathbf{A}$, determinant becomes $-\det \mathbf{A}$
- Elementary row operations do not change the determinant (except permutations)
- $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$

Eigenvalues and eigenvectors

- We can view any linear transformation as a combination of scaling and rotation (and reflection)
- The linear transformation defined by a matrix does not change the directions of some vectors, vectors in these directions are called the *eigenvectors*
- The scaling factor in these directions is called *eigenvalues*
- More formally, if v is an eigenvector of $\mathbf{A}$ with corresponding eigenvalue λ,

$$\mathbf{A}v = \lambda v$$

- Independent eigenvectors of a symmetric are orthogonal

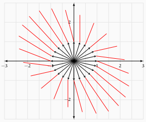
Eigenvalues and eigenvectors

visualization

- We start with the vectors (black arrows)
- The red lines trace the vector after transformation with

$$\begin{bmatrix} 2.3660 & -0.3660 \\ -0.6340 & 2.6340 \end{bmatrix}$$

- In some directions, the vector is only scaled



Finding eigenvalues and eigenvectors

- We can start from the definition

$$\mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v}$$

- Rearranging,

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

- This means the matrix $\mathbf{A} - \lambda\mathbf{I}$ should be singular for non-zero $\boldsymbol{v}$, and

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

- Now we can first solve the equation for λ, and knowing λs we can find the corresponding eigenvectors

Finding eigenvalues and eigenvectors

an example

$$\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$$

Solution:

$$\lambda_1 = 5$$
$$\lambda_2 = 3$$
$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Properties of eigenvalues and eigenvectors

- An $n \times n$ matrix $\mathbf{A}$ has n eigenvalues (which can be complex, or repeated)
- The sum of eigenvalues is the sum of the diagonal of $\mathbf{A}$ (the *trace* of $\mathbf{A}$)
- The product of the eigenvalues is the determinant
- $\mathbf{A}$ and $\mathbf{A}^T$ have the same eigenvalues
- For symmetric matrices, the eigenvectors can be chosen to be orthonormal
- If all eigenvalues of a symmetric are positive, it is called a *positive definite* matrix. More formally, if $\mathbf{A}$ is positive definite, then $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is positive for any $\mathbf{x}$
- If all eigenvalues of a symmetric are non-negative, it is called a *positive semi-definite* matrix

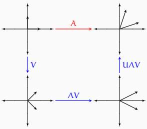
Diagonalization

(eigenvalue decomposition)

- An $n \times n$ with n independent eigenvalues can be *diagonalized* using eigenvalues and eigenvectors
- We take the matrix $\mathbf{S}$ whose columns are the eigenvalues of $\mathbf{A}$, and the diagonal matrix $\boldsymbol{\Lambda}$ with eigenvalues of $\mathbf{A}$, then

$$\mathbf{AS} = \mathbf{S}\boldsymbol{\Lambda}$$
$$\mathbf{A} = \mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}$$
$$\mathbf{S}^{-1}\mathbf{AS} = \boldsymbol{\Lambda}$$

The geometry of eigenvalue decomposition



Matrix powers and matrix inverse

- Matrix powers can be easily calculated with diagonalization

$$\mathbf{Ax} = \lambda\mathbf{x}$$
$$\mathbf{AAx} = \lambda\mathbf{Ax}$$
$$\mathbf{A}^2\mathbf{x} = \lambda^2\mathbf{x}$$

- In general,

$$\mathbf{A}^2 = \mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}\mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}$$
$$= \mathbf{S}\boldsymbol{\Lambda}^2\mathbf{S}^{-1}$$
$$\mathbf{A}^k = \mathbf{S}\boldsymbol{\Lambda}^k\mathbf{S}^{-1}$$

- Inverse is also easy to obtain after eigendecomposition

$$\mathbf{A}^{-1} = \mathbf{S}\boldsymbol{\Lambda}^{-1}\mathbf{S}^{-1}$$

Summary / next

- We reviewed eigenvalues and eigenvectors
- Eigenvalues and eigenvectors have many practical applications from image compression to clustering and dimensionality reduction

Next:

- SVD and pseudo inverse

Further reading

Any of the linear algebra references provided earlier.