Basic Mathematics , Fall 2020

Karen Keryan,
ASDS, YSU

October 15, 2020

Definition

The trace of a square matrix $A \in \mathbb{R}^{n \times n}$ *is given by*

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

in other words, the trace is the sum of the diagonal elements of A.

Remark

For $A; B \in \mathbb{R}^{n \times n}$ *the trace satisfies the following properties:*

1. $tr(A + B) = tr(A) + tr(B)$
2. $tr(\alpha A) = \alpha tr(A)$
3. $tr(I_n) = n$
4. $tr(AB) = tr(BA)$

Corollary

- *In particular, for two vectors $x, y \in \mathbb{R}^n$*

$$tr(\mathbf{x}\mathbf{y}^T) = tr(\mathbf{y}^T\mathbf{x}) = \mathbf{y}^T\mathbf{x} \in \mathbb{R}.$$

- *If $A \sim B$ (i.e. $\exists\, Q$ s.t. $A = QBQ^{-1}$), then $tr A = tr B$*
- *All transformation matrices A_T of a linear mapping $T : V \to V$ have the same trace.*

Definition

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

$$p_A(\lambda) = \det(A - \lambda I) = c_0 + c_1\lambda + c_2\lambda^2 + \ldots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n$$

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

$$c_0 = \det(A), \quad c_{n-1} = (-1)^{n-1} tr(A).$$

Eigenvalues and Eigenvectors

Definition

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

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

Remark

If $\mathbf{x}$ is an eigenvector of A, then so is $c\mathbf{x}$, for any nonzero $c \in \mathbb{R}$.

Theorem

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

Definition

For $A \in \mathbb{R}^{n\times n}$ the set of all eigenvectors of A associated with an eigenvalue λ together with the zero vector is called **eigenspace of** A *with respect to λ and is denoted by E_λ. The set of all eigenvalues of A is called* **spectrum,** *of A.*

Example

Find the eigenvalues and eigenvectors of a 2×2 matrix

$$A = \begin{bmatrix} 5 & 8 \\ 2 & 5 \end{bmatrix}$$

Proposition

- *If $A \sim B$, then eigenvalues of A and B are the same.*
- *A linear mapping T has eigenvalues that are independent of the choice of basis of its transformation matrix.*

Proposition

- *A matrix A and its transpose A^T have the same eigenvalues, but not necessarily the same eigenvectors.*
- *If S is a symmetric matrix, then all its eigenvalues are real.*
- *The eigenvalues of a symmetric positive definite matrix are positive (and real).*
- *The eigenvectors of symmetric matrices are always orthogonal to each other.*

Theorem

For any matrix $A \in \mathbb{R}^{m \times n}$ the matrices $A^T A$ and AA^T are symmetric and positive semi-definite.

Theorem

Any symmetric matrix $A = A^T \in \mathbb{R}^{n \times n}$ has n independent eigenvectors that form an orthogonal basis for $\mathbb{R}^n$.

Definition

Let a square matrix A have an eigenvalue λ algebraic . The **algebraic multiplicity** *of λ is the number of times the root appears in the characteristic polynomial.*

Definition

Let a square matrix A have an eigenvalue λ. The **geometric multiplicity** *of λ is the total number of linearly independent eigenvectors associated with λ. In other words it is the dimension of the eigenspace E_λ spanned by the eigenvectors associated with λ.*

Example

Find the geometric and algebraic multiplicity of the eigenvalue(s) of the matrix $A = \begin{bmatrix} 3 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Theorem

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

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

where λ_i *are (possibly repeated) eigenvalues of* A*.*

Theorem

The trace of a matrix $A \in \mathbb{R}^{n\times n}$ *is the sum of its eigenvalues, i.e.,*

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

where λ_i *are (possibly repeated) eigenvalues of* A*.*

Cholesky Decomposition

A symmetric positive definite matrix A can be factorized into a product $A = LL^T$, where L is a lower triangular matrix with positive diagonal elements:

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ l_{n1} & \dots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & \dots & l_{n1} \\ \vdots & \ddots & \vdots \\ 0 & \dots & l_{nn} \end{bmatrix}$$

and L is called the Cholesky factor of A.

Example

Find Cholesky decomposition of a general 2×2 symmetric matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}.$$

Applications of Cholesky decompozition

Simulation. Generating samples of multivariate normal distribution with covariance matrix A is hard, as it has a lot of dependencies, but generating sample points of $N(0, I_n)$ is easy, and using the relationship

$$N(0, A) = N(0, LL^T) = LN(0, I_n)$$

one can generate sample points of Gaussian distribution $N(0, I_n)$ and then transform them into points of Gaussian distribution $N(0, A)$.

Simulation. Generating samples of multivariate normal distribution with covariance matrix A is hard, as it has a lot of dependencies, but generating sample points of $N(0, I_n)$ is easy, and using the relationship

$$N(0, A) = N(0, LL^T) = LN(0, I_n)$$

one can generate sample points of Gaussian distribution $N(0, I_n)$ and then transform them into points of Gaussian distribution $N(0, A)$.

Linear equations. The Cholesky decomposition is mainly used for the numerical solution of linear equations $\mathbf{Ax} = \mathbf{b}$. If A is symmetric and positive definite, then we can solve $\mathbf{Ax} = \mathbf{b}$ by first computing the Cholesky decomposition $\mathbf{A} = \mathbf{LL}^{\mathrm{T}}$, then solving $\mathbf{Ly} = \mathbf{b}$ for y by forward substitution, and finally solving $\mathbf{L^T x} = \mathbf{y}$ for x by back substitution.

Corollary

If $A = LL^T$, *then*

$$\det A =$$

Corollary

If $A = LL^T$*, then*

$$\det A = (\det L)^2 =$$

Corollary

If $A = LL^T$*, then*

$$\det A = (\det L)^2 = (l_{11} \cdot \ldots \cdot l_{nn})^2.$$

Moreover

$$A^{-1} = (L^{-1})^T (L^{-1}).$$

Definition

A matrix $A \in \mathbb{R}^{n \times n}$ is **diagonalizable** *if it is similar to a diagonal matrix, i.e. there exists a matrix $P \in \mathbb{R}^{n \times n}$ so that*

$$D = P^{-1}AP,$$

(equivalently $AP = PD$ or $A = PDP^{-1}$) where

$$D = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{bmatrix}$$

Proposition

If $D \in \mathbb{R}^{n\times n}$ is a diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$, then

$$AP = PD$$

if and only if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A and the $\mathbf{p}_i$ are the corresponding eigenvectors of A, where $P = [\mathbf{p}_1 \dots \mathbf{p}_n]$.

Theorem

A symmetric matrix $S = S^T \in \mathbb{R}^{n\times n}$ can be diagonalized into

$$S = PDP^T$$

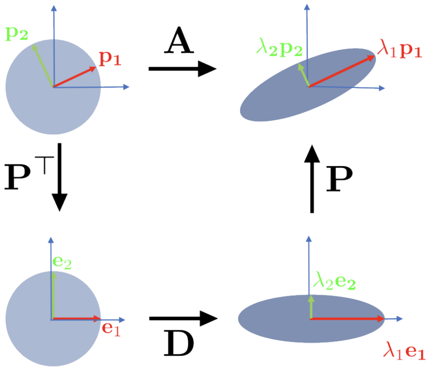
where P is matrix of n orthogonal eigenvectors,i.e. $P^T = P^{-1}$, and D is a diagonal matrix of its n eigenvalues.

Example

Compute the eigendecomposition of a (symmetric) matrix $A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$

Proposition

If $A = PDP^{-1}$, then $A^k = PD^kP^{-1}$ for any $k \in N$, and $D^k = \begin{bmatrix} \lambda_1^k & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n^k \end{bmatrix}$


$\mathbf{p_2}$
$\mathbf{p_1}$
$\mathbf{A}$
$\lambda_2\mathbf{p_2}$
$\lambda_1\mathbf{p_1}$
$\mathbf{P}^\top$
$\mathbf{P}$
$\mathbf{e}_2$
$\mathbf{e}_1$
$\mathbf{D}$
$\lambda_2\mathbf{e_2}$
$\lambda_1\mathbf{e_1}$