Vv214 Linear Algebra

RC5

Li Yuzhou

SJTU-UM Joint Institute
Shanghai Jiao-Tong University

July 3, 2019

Table of Contents

Table of Contents

Orthogonal transformation and orthogonal matrices

Definition

A linear transformation T from R^n to R^n is called orthogonal if it preserves the length of vectors:

$$\|T(\vec{x})\| = \|\vec{x}\|, \quad \text{for all } \vec{x} \text{ in } \mathbb{R}^n$$

If $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, we say that A is an orthogonal matrix.

Remark

To find whether a transformation is orthogonal, you can just illustrate whether the length is preserved.

Orthogonal transformation and orthogonal matrices

Example

The rotation

$$T(\vec{x}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \vec{x}$$

is an orthogonal transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$, and

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is an orthogonal matrix, for all angles θ.

Orthogonal transformation and orthogonal matrices

Exercise

Consider an orthogonal transformation T from R^n to R^n. If the vectors $\vec{(v)}$ and $\vec{(w)}$ in R^n are orthogonal, prove that $T(\vec{(v)})$ and $T(\vec{(w)})$ are also orthogonal.

Hint

We can use the theorem of Pythagoras to prove two vectors are orthogonal.

$$\|\vec{x}+\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

Orthogonal transformations and orthonormal bases

$\|T(\vec{x})\|^2 = \|x_1 T(\vec{e}_1) + \cdots + x_n T(\vec{e}_n)\|$
$= x_1^2 + x_2^2 + \cdots + x_n^2$
$= \|x\|^2$

If A is an ortho- mat

$\|\vec{x}\| = \|AA^{-1}\vec{x}\| = \|A^{-1}\vec{x}\|$

Theorem

1. A linear transformation T from R^n to R^n is orthogonal if and only if vectors $T(\vec{e}_1), T(\vec{e}_2), \ldots, T(\vec{e}_n)$ form an orthonormal basis of R^n.
2. An $n \times n$ matrix A is orthogonal if and only if its columns form an orthonormal basis of R^n.

Remark

1. You can try to prove this using the statement we just proved and the linearity.
2. Checking whether the columns form an orthonormal basis is the fastest way to check a matrix to be orthogonal or not.

Product and Inverse

Theorem

1. The product AB of two orthogonal $n \times n$ matrices A and B is orthogonal.
2. The inverse A^{-1} of an orthogonal $n \times n$ matrix A is orthogonal.

Remark

1. You can prove this directly using the definition of orthogonal transformation.

Matrix Transpose

Property of transpose

1. We say that a square matrix A is symmetric if $A^T = A$, and A is called skew-symmetric if $A^T = -A$.
2. An $n \times n$ matrix A is orthogonal if and only if $A^T A = I_n$, or equivalently $A^{-1} = A^T$.

$$(AB)^T = B^T A^T.$$

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

$$\text{rank}(A) = \text{rank}\left(A^T\right).$$

Remark

1. You can prove statement 2 using $\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}$.

The matrix of an orthogonal projection

Consider a subspace V of $\mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$. The matrix of the orthogonal projection onto V is

$$QQ^T, \quad \text{where} \quad Q = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_m \\ | & | & & | \end{bmatrix}.$$

Pay attention to the order of the factors (QQ^T as opposed to Q^TQ). ∎

Remark

Constructing the matrix column by column also works fine to find this matrix.

Exercise

Find the matrix of the orthogonal projection onto the subspace of $\mathbb{R}^4$ spanned by

$$\vec{u}_1 = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad \vec{u}_2 = \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}.$$

$$\text{proj}_V \vec{x} = (\vec{x}\cdot\vec{u}_1)\vec{u}_1 + (\vec{x}\cdot\vec{u}_2)\vec{u}_2$$

Remark
Try both ways!

$$\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

Summary

Consider an $n \times n$ matrix A. Then the following statements are equivalent:

i. A is an orthogonal matrix.

ii. The transformation $L(\vec{x}) = A\vec{x}$ preserves length, that is, $\|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x}$ in $\mathbb{R}^n$.

iii. The columns of A form an orthonormal basis of $\mathbb{R}^n$.

iv. $A^T A = I_n$.

v. $A^{-1} = A^T$.

Table of Contents

Gram-Schmidt process

Definition

It gives the way to construct an orthonormal bases given an arbitrary basis of the space.
The construction process is as follow:

1.

$$\vec{v}_j = \vec{v}_j^{\parallel} + \vec{v}_j^{\perp}, \qquad \text{with respect to span}\left(\vec{v}_1, \ldots, \vec{v}_{j-1}\right).$$

$$\vec{v}_j^{\perp} = \vec{v}_j - \vec{v}_j^{\parallel} = \vec{v}_j - (\vec{u}_1 \cdot \vec{v}_j)\,\vec{u}_1 - \cdots - (\vec{u}_{j-1} \cdot \vec{v}_j)\,\vec{u}_{j-1}$$

2.

$$\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1, \quad \vec{u}_2 = \frac{1}{\|\vec{v_2}^{\perp}\|}\vec{v}_2^{\perp}, \ldots, \overline{u}_m = \frac{1}{\|\vec{v_m}^{\perp}\|}\vec{v}_m^{\perp}$$

The calculation will not be more than 4 dimension.

QR factorization

Definition

Consider an $n \times m$ matrix M with linearly independent columns $\vec{v}_1, \ldots, \vec{v}_m$. Then there exists an $n \times m$ matrix Q whose columns $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal and an upper triangular matrix R with positive diagonal entries such that

$$M = QR.$$

This representation is unique. Furthermore, $r_{11} = \|\vec{v}_1\|$, $r_{jj} = \left\|\vec{v}_j^{\perp}\right\|$ (for $j = 2, \ldots, m$), and $r_{ij} = \vec{u}_i \cdot \vec{v}_j$ (for $i < j$). ■

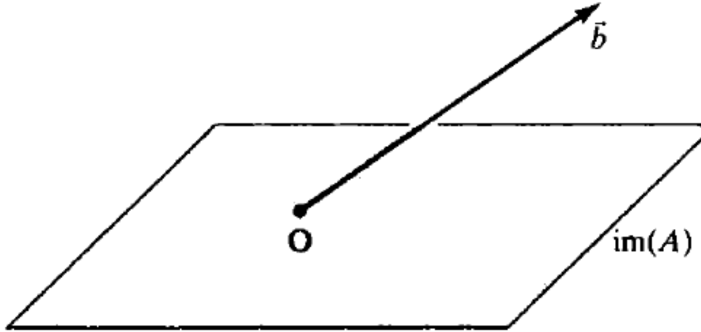
Just perform the Gram-Schmidt process for the set of columns of M and all the information about matrix Q and R will be shown in the process.

Table of Contents

Least squares approximation

$A\vec{x} = \vec{b}$

Definition



We are trying to find the vector that are closest to the vector $\vec{b}$ in the image of matrix A.

Least squares approximation

Theorem

1. $$(\mathrm{im} A)^{\perp} = \ker\left(A^T\right).$$

2. $$\ker(A) = \ker\left(A^T A\right).$$

3. If $\ker(A)=0$, then $A^T A$ is invertible.

Try to prove these statements!

1. $(\mathrm{im}\, A)^{\perp} = \ker(A^T)$ $\quad V^{\perp}$ $\quad \{\vec{x} : \vec{x}\cdot\vec{v} = 0\}$

$A = [\vec{v}_1 \ \cdots \ \vec{v}_m]$ $\Rightarrow \vec{x}\cdot\vec{v}_i = 0.$

$\Updownarrow$

$\vec{v}_i^{\,T}\vec{x} = 0.$

2. $A^T A\vec{x} = 0$

$\Rightarrow A\vec{x} = 0.$

Suppose: $A\vec{x} \neq 0$, $A^T \cdot A\vec{x} = 0$

$\Rightarrow A\vec{x} \in \ker(A^T) = \mathrm{im}(A)^{\perp}$

$\mathrm{im}(A)$

Least squares approximation

Method

The least-squares solutions of the system

$$A\vec{x} = \vec{b}$$

are the exact solutions of the (consistent) system

$$A^T A\vec{x} = A^T\vec{b}.$$

The system $A^T A\vec{x} = A^T\vec{b}$ is called the *normal equation* of $A\vec{x} = \vec{b}$.

If $\ker(A) = \{\vec{0}\}$, then the linear system

$$A\vec{x} = \vec{b}$$

has the unique least-squares solution

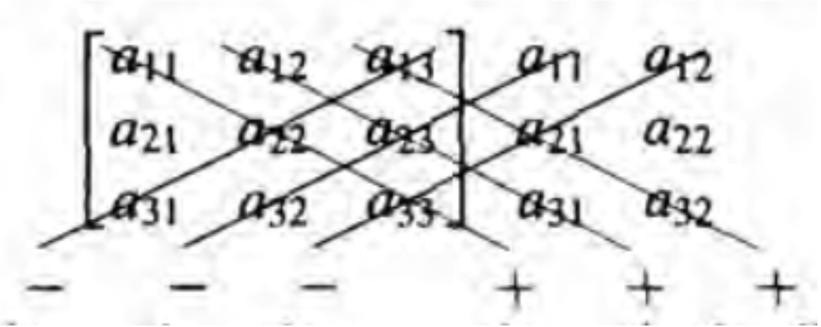
$$\vec{x}^* = (A^T A)^{-1} A^T\vec{b}.$$

Note that $\vec{x}^*$ is not the orthogonal projection of $\vec{b}$, $A\vec{x}^*$ is.

Table of Contents

Finding determinant for 3×3 matrix

Sarrus's rule



Finding determinant for 3×3 matrix

Typical Type for Sarrus's rule

Find the determinant of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}.$$

Sarrus's rule is for and only for 3×3 matrix!

Properties for finding matrix

1. $\det B = -\det A$ if B is obtained by switching any two columns or any rows of a matrix A.
2. $$\det \begin{bmatrix} - & \vec{v}_1 & - \\ & \vec{v}_2 & - \\ - & \overline{x} + y & - \end{bmatrix}$$
3. $$\det \begin{bmatrix} - & \vec{v}_1 & - \\ & \vec{v}_2 & - \\ - & k\overline{x} & - \end{bmatrix}$$

From these properties, we can have some other conclusions...

4×4 matrix

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

$\det A = (-1)^s k_1 k_2 \cdots k_r$

Patterns and Inversions

Patterns, inversions, and determinants[4]

A *pattern* in an $n \times n$ matrix A is a way to choose n entries of the matrix so that there is one chosen entry in each row and in each column of A.

With a pattern P we associate the product of all its entries, denoted prod P.

Two entries in a pattern are said to be *inverted* if one of them is located to the right and above the other in the matrix.

The *signature* of a pattern P is defined as $\operatorname{sgn} P = (-1)^{(\text{number of inversions in } P)}$.

The determinant of A is defined as

$$\det A = \sum (\operatorname{sgn} P)(\operatorname{prod} P),$$

where the sum is taken over all $n!$ patterns P in the matrix A. Thus we are summing up the products associated with all patterns with an even number of inversions, and we are subtracting the products associated with the patterns with an odd number of inversions.

Patterns and Inversions typical problem

$$A = \begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Patterns and Inversions typical problem

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 & 5 \\ 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

only 1 case.

Patterns and Inversions typical problem

$$A = \begin{bmatrix} 6 & 0 & 1 & 0 & 0 \\ 9 & 3 & 2 & 3 & 7 \\ 8 & 0 & 3 & 2 & 9 \\ 0 & 0 & 4 & 0 & 0 \\ 5 & 0 & 5 & 0 & 1 \end{bmatrix}.$$

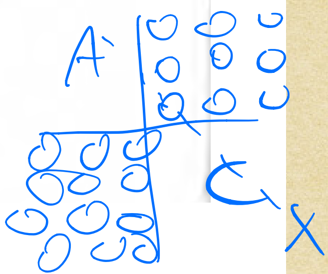
Determinant of Block Matrix

If $M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, where A and C are square matrices (not necessarily of the same size), then

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = (\det A)(\det C).$$

Likewise,

$$\det \begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = (\det A)(\det C).$$



More properties

1. $$\det\left(A^T\right) = \det A.$$

2. $$\det(AB) = (\det A)(\det B)$$

3. $$A \text{ is similar to } B, \text{ then } \det A = \det B$$

4. $$\det\left(A^{-1}\right) = \frac{1}{\det A} = (\det A)^{-1}.$$

Laplace expansion

We can compute the determinant of an $n \times n$ matrix A by Laplace expansion down any columns or along any row.

1. Expansion down the *jth* column:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det\left(A_{ij}\right).$$

2. Expansion down the *ith* row:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det\left(A_{ij}\right)$$

Exercise for Laplace Expansion

$$A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 9 & 1 & 3 & 0 \\ 9 & 2 & 2 & 0 \\ 5 & 0 & 0 & 3 \end{bmatrix}.$$