Group theory
(and some applications in Computer Science)

Mai Hoàng Biên

Faculty of Mathematics and Computer Science, University of Science, VNUHCM

Chapter 3. Some classes of Groups

1. 3.1. Dihedral groups (in computer visions)
2. 3.2. Braid group Cryptography
3. 3.3. Linear groups (in Neural Networks)
4. 3.4. Groups of points (of Elliptic Curve Cryptography)

Linear groups. Somes applications of linear groups

Data Analysis:

In neural networks, employing linear transformations aids in analyzing and extracting features from data. Linear transformations, such as matrix operations, not only reduce data dimensionality but also enhance classification capabilities and deepen understanding of data structure.

Linear groups. Somes applications of linear groups

Data Analysis:

In neural networks, employing linear transformations aids in analyzing and extracting features from data. Linear transformations, such as matrix operations, not only reduce data dimensionality but also enhance classification capabilities and deepen understanding of data structure.

Machine Learning and Neural Networks:

In the realm of machine learning and neural networks, linear transformations play a pivotal role in constructing data-driven models. Linear layers, such as fully connected layers in neural networks, are integral components in the architectures of these models.

Somes applications of linear groups

Information Analysis:

In natural language processing and textual data analysis, linear methods like principal component analysis (PCA) support data dimensionality reduction and extraction of crucial information.

Somes applications of linear groups

Information Analysis:

In natural language processing and textual data analysis, linear methods like principal component analysis (PCA) support data dimensionality reduction and extraction of crucial information.

Image Processing:

In image processing, linear transformations are utilized to filter and extract features from images. Applying linear matrices allows for transformations such as convolutions, playing a crucial role in image processing.

Linear groups

3.1. Definition

Let $n > 1$ be a integer. The group $\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathrm{M}_n(\mathbb{R}) \mid |A| \neq 0\}$ is called the general linear group of degree n over $\mathbb{R}$.

Linear groups

3.1. Definition

Let $n > 1$ be a integer. The group $\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathrm{M}_n(\mathbb{R}) \mid |A| \neq 0\}$ is called the general linear group of degree n over $\mathbb{R}$.

1. Every subgroup of $\mathrm{GL}_n(\mathbb{R})$ is called a linear group of degree n over $\mathbb{R}$.

Linear groups

3.1. Definition

Let $n > 1$ be a integer. The group $\mathrm{GL}_n(\mathbb{R}) = \{A \in \mathrm{M}_n(\mathbb{R}) \mid |A| \neq 0\}$ is called the general linear group of degree n over $\mathbb{R}$.

1. Every subgroup of $\mathrm{GL}_n(\mathbb{R})$ is called a linear group of degree n over $\mathbb{R}$.
2. $\mathrm{SL}_n(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) : |A| = 1\}$ is called special linear groups.

3.2. Some special matrices

Let $A \in \mathrm{M}_n(\mathbb{R})$.

1. A is called *symmetric* if $A^t = A$. Here A^t is the transpose of A.
2. A is called orthogonal if $A^{-1} = A^t$, that is, $AA^t = I_n$.

Orthogonal groups

3.3. Proposition.

If $A \in GL_n(\mathbb{R})$ is orthogonal, then $|A| = \pm 1$.

Proof.

Orthogonal groups

3.3. Proposition.

If $A \in GL_n(\mathbb{R})$ is orthogonal, then $|A| = \pm 1$.

Proof.

3.4. Theorem

The set of orthogonal matrices in $GL_n(\mathbb{R})$ is a linear group (of degree n) over $\mathbb{R}$.

Proof.

finite linear groups

Symmetries in neural networks: a linear group action approach

R Folk†, A Kartashov‡, P Lisoněk§ and P Paule§
† Institute for Theoretical Physics, Linz University, Austria
‡ Neural Networks Laboratory, Moscow Institute for Radiotechnics, Electronics & Automation, Moscow, Russia
§ Research Institute for Symbolic Computation, Linz University, Austria

Received 13 July 1992, in final form 15 March 1993

Abstract. Neural networks storing Hadamard patterns have been completely classified with respect to permutation symmetry. The symmetry group of the Hadamard patterns is found to be isomorphic to $GL(n, F_2)$, and the symmetry groups of the networks are explicitly constructed for the most important classes. The volumes of different equivalence classes have been calculated.

The fields $\mathbb{Z}_p$ for p a prime number

Recall that, for a prime number p, the set $\mathbb{Z}_p = \{\overline{0}, \overline{1}, \cdots, \overline{p-1}\}$ includes of all residue classes modulo p with two operators $\overline{a} + \overline{b} = \overline{a+b}$ and $\overline{a} \cdot \overline{b} = \overline{ab}$ for every $\overline{a}, \overline{b} \in \mathbb{Z}_p$.

The fields $\mathbb{Z}_p$ for p a prime number

Recall that, for a prime number p, the set $\mathbb{Z}_p = \{\overline{0}, \overline{1}, \cdots, \overline{p-1}\}$ includes of all residue classes modulo p with two operators $\overline{a} + \overline{b} = \overline{a+b}$ and $\overline{a} \cdot \overline{b} = \overline{ab}$ for every $\overline{a}, \overline{b} \in \mathbb{Z}_p$. Put

$$\mathbb{Z}_p^* = \{\overline{1}, \cdots, \overline{p-1}\}.$$

3.5. Proposition

$(\mathbb{Z}_p, +)$ and $(\mathbb{Z}_p^*, \cdot)$ are abelian groups.

Proof.

Linear groups over $\mathbb{Z}_p$

Let $GL_n(\mathbb{Z}_p)$ be the group of invertible matrices of degree n whose entries belongs to $\mathbb{Z}_p$. Then,

3.6. Theorem

1. $GL_n(\mathbb{Z}_p) = \{A \in M_n(\mathbb{Z}_p) \mid |A| \neq \overline{0}\}$ and $GL_n(\mathbb{Z}_p)$ has $(p^n - 1)(p^n - p) \cdots (p^n - p^{n-1})$ matrices.

Linear groups over $\mathbb{Z}_p$

Let $GL_n(\mathbb{Z}_p)$ be the group of invertible matrices of degree n whose entries belongs to $\mathbb{Z}_p$. Then,

3.6. Theorem

1. $GL_n(\mathbb{Z}_p) = \{A \in M_n(\mathbb{Z}_p) \mid |A| \neq \overline{0}\}$ and $GL_n(\mathbb{Z}_p)$ has $(p^n - 1)(p^n - p) \cdots (p^n - p^{n-1})$ matrices.
2. $SL_n(\mathbb{Z}_p) = \{A \in M_n(\mathbb{Z}_p) \mid |A| = 1\}$ is a normal subgroup of $GL_n(\mathbb{Z}_p)$ and $SL_n(\mathbb{Z}_p)$ has $\frac{1}{p-1}(p^n - 1)(p^n - p) \cdots (p^n - p^{n-1})$ matrices.

Example

Find the number of matrices in $GL_3(\mathbb{Z}_2)$, $GL_2(\mathbb{Z}_3)$, $GL_3(\mathbb{Z}_3)$, $SL_4(\mathbb{Z}_3)$.

Singular value decompositions and its applications

Singular value decomposition (SVD)

Open the book, Chapter 13, Page 237.

Singular value decompositions and its applications

Singular value decomposition (SVD)

Open the book, Chapter 13, Page 237.
For convenience, in this course, all matrices we want to focus of **singular value decompositions of symmetric matrices over** $\mathbb{R}$.
However, many results work on any case.

SVD of real symmetric matrices

3.7. Theorem

Let $A \in M_n(\mathbb{R})$. If A is symmetric, that is, $A^t = A$, there exists an orthogonal matrix $Q \in GL_n(\mathbb{R})$ such that $A = QDQ^t$ where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

SVD of real symmetric matrices

3.7. Theorem

Let $A \in M_n(\mathbb{R})$. If A is symmetric, that is, $A^t = A$, there exists an orthogonal matrix $Q \in GL_n(\mathbb{R})$ such that $A = QDQ^t$ where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

In this lecture, we will find this decomposition step by step as follows:

1. Find $\lambda_1, \lambda_2, \cdots, \lambda_n$ of D which are called eigenvalues of A.

SVD of real symmetric matrices

3.7. Theorem

Let $A \in M_n(\mathbb{R})$. If A is symmetric, that is, $A^t = A$, there exists an orthogonal matrix $Q \in GL_n(\mathbb{R})$ such that $A = QDQ^t$ where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

In this lecture, we will find this decomposition step by step as follows:

1. Find $\lambda_1, \lambda_2, \cdots, \lambda_n$ of D which are called eigenvalues of A.
2. Find $P \in GL_n(\mathbb{R})$ such that $A = PDP^{-1}$.

SVD of real symmetric matrices

3.7. Theorem

Let $A \in M_n(\mathbb{R})$. If A is symmetric, that is, $A^t = A$, there exists an orthogonal matrix $Q \in GL_n(\mathbb{R})$ such that $A = QDQ^t$ where

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

In this lecture, we will find this decomposition step by step as follows:

1. Find $\lambda_1, \lambda_2, \cdots, \lambda_n$ of D which are called eigenvalues of A.
2. Find $P \in GL_n(\mathbb{R})$ such that $A = PDP^{-1}$.
3. "build" an orthogonal matrix Q, that is, $Q^t = Q$, from P such that $A = QDQ^t$.

Notation

In this lecture,

1. Matrices are square matrices of degree n whose entries are in $\mathbb{R}$.
2. The Vector space $\mathbb{R}^n$ is $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R}\}$. The zero vector is denoted by $0 = (0, 0, \cdots, 0)$.

3.8. Definition

Let $A \in M_n(\mathbb{R})$ and $c \in \mathbb{R}$.

Eigenvalues and eigenvectors of matrices

3.8. Definition

Let $A \in M_n(\mathbb{R})$ and $c \in \mathbb{R}$. The value c is called an eigenvalue of A if there exists $0 \neq v \in \mathbb{R}^n$ such that $Av^t = cv^t$. In this case, v is called an eigenvector of A with respect to c.

Eigenvalues and eigenvectors of matrices

3.8. Definition

Let $A \in M_n(\mathbb{R})$ and $c \in \mathbb{R}$. The value c is called an eigenvalue of A if there exists $0 \neq v \in \mathbb{R}^n$ such that $Av^t = cv^t$. In this case, v is called an eigenvector of A with respect to c.

Remark

c is an eigenvalues of A if and only if the system of linear equations with coefficient matrix $A - cI_n$ has a non-trivial solution.

3.8. Definition

Let $A \in M_n(\mathbb{R})$ and $c \in \mathbb{R}$. The value c is called an eigenvalue of A if there exists $0 \neq v \in \mathbb{R}^n$ such that $Av^t = cv^t$. In this case, v is called an eigenvector of A with respect to c.

Remark

c is an eigenvalues of A if and only if the system of linear equations with coefficient matrix $A - cI_n$ has a non-trivial solution.

example

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Then $c = 3$ is an eigenvalue of A but 2 is not an eigenvalue. Show....(Students need to attend the class to listen to the teacher's explanations.)

3.9. Theorem

If c is an eigenvalue of A, then the set $E_c = \{v \in \mathbb{R}^n \mid Av^t = cv^t\}$ is a nonzero vector subspace of $\mathbb{R}^n$. We call E_c the eigenspace with respect to c

Proof. In fact, E_c is the subspace of $\mathbb{R}^n$ consists all solutions of the system of linear equations with coefficient matrix $A - cI_n$ or $cI_n - A$. (Students need to attend the class to listen to the teacher's explanations.)

Vector subspace of eigenvectors

3.9. Theorem

If c is an eigenvalue of A, then the set $E_c = \{v \in \mathbb{R}^n \mid Av^t = cv^t\}$ is a nonzero vector subspace of $\mathbb{R}^n$. We call E_c the eigenspace with respect to c

Proof. In fact, E_c is the subspace of $\mathbb{R}^n$ consists all solutions of the system of linear equations with coefficient matrix $A - cI_n$ or $cI_n - A$. (Students need to attend the class to listen to the teacher's explanations.)

Example

Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$. Show that 3 is an eigenvalue of A and find E_3.

3.10. Definition

For a matrix A, the characteristic polynomial $p_A(t)$ in indeterminate t of A is defined as the determinant

$$p_A(t) = |tI_n - A|.$$

Characteristic polynomials of matrices

3.10. Definition

For a matrix A, the characteristic polynomial $p_A(t)$ in indeterminate t of A is defined as the determinant

$$p_A(t) = |tI_n - A|.$$

Example

Find the characteristic polynomials of $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$,

$B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & -3 \end{pmatrix}$ and $C = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Characteristic polynomials of matrices

3.11. Proposition

For a matrix A, the characteristic polynomial $p_A(t)$ of A is monic of degree n, that is, it is of the form

$$p_A(t) = |tI_n - A| = t^n + b_{n-1}t^{n-1} + \cdots + b_1 t + b_0.$$

Characteristic polynomials of matrices

3.11. Proposition

For a matrix A, the characteristic polynomial $p_A(t)$ of A is monic of degree n, that is, it is of the form

$$p_A(t) = |tI_n - A| = t^n + b_{n-1}t^{n-1} + \cdots + b_1 t + b_0.$$

Additionally, if $A^t = A$, then $p_A(t)$ has the form

$$p_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$$

in which $n_1 + n_2 + \cdots + n_r = n$.

Characteristic polynomials of matrices

3.11. Proposition

For a matrix A, the characteristic polynomial $p_A(t)$ of A is monic of degree n, that is, it is of the form

$$p_A(t) = |tI_n - A| = t^n + b_{n-1}t^{n-1} + \cdots + b_1 t + b_0.$$

Additionally, if $A^t = A$, then $p_A(t)$ has the form

$$p_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$$

in which $n_1 + n_2 + \cdots + n_r = n$.

Example

look at back the last example.

3.12. Theorem

c is an eigenvalue of A if and only if c is a root of $p_A(t) = 0$.

3.12. Theorem

c is an eigenvalue of A if and only if c is a root of $p_A(t) = 0$.

Example

Find all eigenvalues of $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & -3 \end{pmatrix}$ and

$C = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Diagonal entries in SVD

Now we have the firs result on SVD of matrices.

3.13. Theorem

Let $A \in M_n(\mathbb{R})$. Assume that $A = QDQ^t$ in which Q is orthogonal and $D = \mathrm{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})$ is a diagonal. Then, λ_i is an eigenvalue of A, that is, λ_i is a root of the characteristic polynomial $p_A(t) = |tI_n - A|$ of A. Moreover, if

$$p_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r},$$

then λ_i appears n_i times in D.

Diagonal entries in SVD

Now we have the firs result on SVD of matrices.

3.13. Theorem

Let $A \in M_n(\mathbb{R})$. Assume that $A = QDQ^t$ in which Q is orthogonal and $D = \mathrm{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})$ is a diagonal. Then, λ_i is an eigenvalue of A, that is, λ_i is a root of the characteristic polynomial $p_A(t) = |tI_n - A|$ of A. Moreover, if

$$p_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r},$$

then λ_i appears n_i times in D.

Example. Find the diagonal matrix in the SVD of

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & -3 \end{pmatrix} \text{ and } C = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Dimensions of eigenspaces

3.14. Theorem

Assume that $A \in M_n(\mathbb{R})$ is symmetric with characteristic polynomial
$p_A(t) = |tI_n - A| = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$.

1. For every $1 \leq i \leq r$, the dimension of the eigenspace E_{λ_i} is n_i.

Dimensions of eigenspaces

3.14. Theorem

Assume that $A \in M_n(\mathbb{R})$ is symmetric with characteristic polynomial
$p_A(t) = |tI_n - A| = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$.

1. For every $1 \leq i \leq r$, the dimension of the eigenspace E_{λ_i} is n_i.
2. Moreover, if $\{u_{i1}, u_{i2}, \cdots, u_{in_i}\}$ is a basis of E_{λ_i} for every $1 \leq i \leq r$, and

$$P = (u_{11}^t u_{12}^t \cdots u_{1n_1}^t \cdots u_{r1}^t u_{r2}^t \cdots u_{rn_r}^t),$$

then $A = PDP^{-1}$.

Algorithm to find invertible matrix P and diagonal matrix D such that $A = PDP^{-1}$

1. Find the characteristic polynomial $p_A(t) = |tI_n - A|$.
2. Find eigenvalues λ_i of A.
3. Find the basis of the eigenspace E_{λ_i}.
4. The matrix P is the matrix whose columns are from the previous step.

Example

1. Find an invertible matrix P and diagonal matrix D such that $A = PDP^{-1}$. $A = \begin{pmatrix} 5 & 4 & 6 \\ 4 & 5 & 6 \\ -4 & -4 & -5 \end{pmatrix}$ and $\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$.
2. Find the formula of Fibonacci sequence: 1,1,2,3,5,8,....
3. Find SVD of $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$ and apply to find A^n for every $n \geq 1$.

Exercises

Scalar product

In this lecture, for two vectors $u = (u_1, u_2, \dots, u_n)$ and $v = (v_1, v_2, \dots, v_n)$ in $\mathbb{R}^n$, we consider the scalar product of u and v as follows:

$$uv = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

3.15. Remarks and definitions

1. The product $u^2 = uu = u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$. The product $u^2 = 0$ if and only if $u = 0$. Put $||u|| = \sqrt{u^2}$ and we call it the *length* or *norm* of the vector u. A vector u is called an unit if $||u|| = 1$.
2. $|||u|| - ||v|| \leq ||u + v|| \leq ||u|| + ||v||$.

The angle of two vectors

We can show that for every $u, v \in V$, $|uv| \leq ||u|| \cdot ||v||$ which implies that

$$-1 \leq \frac{uv}{||u|| \cdot ||v||} \leq 1.$$

3.16. Definition

The angle of two vectors u and v is angle $\alpha \in [0, 180^o]$ such that

$$\cos\alpha = \frac{uv}{||u|| \cdot ||v||}.$$

The angle of two vectors

We can show that for every $u, v \in V$, $|uv| \leq ||u|| \cdot ||v||$ which implies that

$$-1 \leq \frac{uv}{||u|| \cdot ||v||} \leq 1.$$

3.16. Definition

The angle of two vectors u and v is angle $\alpha \in [0, 180^o]$ such that

$$\cos\alpha = \frac{uv}{||u|| \cdot ||v||}.$$

We say that two vectors u, v are *orthogonal* or *perpendicular* if $\alpha = 90^o$, that is $uv = 0$. In this case, we write $u \perp v$.

Orthonormal basis

Let W be a subspace of $\mathbb{R}^n$ with basis $B = \{u_1, u_2, \ldots, u_m\}$.

3.17. Definition

The basis B is called an *orthonormal basis* of W if $u_i \perp u_j$ and $||u_i|| = 1$ for every $1 \leq i \neq j \leq m$.

Orthonormal basis

Let W be a subspace of $\mathbb{R}^n$ with basis $B = \{u_1, u_2, \ldots, u_m\}$.

3.17. Definition

The basis B is called an *orthonormal basis* of W if $u_i \perp u_j$ and $||u_i|| = 1$ for every $1 \leq i \neq j \leq m$.

Example

See the whitebroad.

Orthorgonal matrix and orthonormal basis

3.18. Theorem

For a matrix $A \in \mathrm{M}_n(\mathbb{R})$, the following statements are equivalent.

1. A is orthorgonal, that is, $A^{-1} = A^t$.
2. The set of columns of A are orthonormal.
3. The set of lines of A are orthonormal.

Orthorgonal matrix and orthonormal basis

3.18. Theorem

For a matrix $A \in \mathrm{M}_n(\mathbb{R})$, the following statements are equivalent.

1. A is orthorgonal, that is, $A^{-1} = A^t$.
2. The set of columns of A are orthonormal.
3. The set of lines of A are orthonormal.

Example

See the whitebroad.

Orthorgonal matrix and orthonormal basis

3.18. Theorem

For a matrix $A \in \mathrm{M}_n(\mathbb{R})$, the following statements are equivalent.

1. A is orthorgonal, that is, $A^{-1} = A^t$.
2. The set of columns of A are orthonormal.
3. The set of lines of A are orthonormal.

Example

See the whitebroad.

The existence of orthonormal bases

Every subspace W of $\mathbb{R}^n$ contains an orthonormal basis.

Gram-Schmidt algorithm

Let W be a subspace of $\mathbb{R}^n$ with basis $A = \{u_1, u_2, \ldots, u_m\}$. We can find an orthornormal basis of W which is defined from A as follows:

1. Put $v_1 = u_1$.

Gram-Schmidt algorithm

Let W be a subspace of $\mathbb{R}^n$ with basis $A = \{u_1, u_2, \ldots, u_m\}$. We can find an orthornormal basis of W which is defined from A as follows:

1. Put $v_1 = u_1$.
   $v_2 = u_2 - \frac{u_2 v_1}{v_1 v_1} v_1$

Gram-Schmidt algorithm

Let W be a subspace of $\mathbb{R}^n$ with basis $A = \{u_1, u_2, \ldots, u_m\}$. We can find an orthornormal basis of W which is defined from A as follows:

1. Put $v_1 = u_1$.
   $v_2 = u_2 - \frac{u_2 v_1}{v_1 v_1} v_1$
   $v_3 = u_3 - \frac{u_3 v_1}{v_1 v_1} v_1 - \frac{u_3 v_2}{v_2 v_2} v_2$

Gram-Schmidt algorithm

Let W be a subspace of $\mathbb{R}^n$ with basis $A = \{u_1, u_2, \ldots, u_m\}$. We can find an orthornormal basis of W which is defined from A as follows:

1. Put $v_1 = u_1$.
   $v_2 = u_2 - \frac{u_2 v_1}{v_1 v_1} v_1$
   $v_3 = u_3 - \frac{u_3 v_1}{v_1 v_1} v_1 - \frac{u_3 v_2}{v_2 v_2} v_2$
   $\vdots$
   $v_m = u_m - \frac{u_m v_1}{v_1 v_1} v_1 - \frac{u_m v_2}{v_2 v_2} v_2 - \cdots - \frac{u_m v_{m-1}}{v_{m-1} v_{m-1}} v_{m-1}.$

Gram-Schmidt algorithm

Let W be a subspace of $\mathbb{R}^n$ with basis $A = \{u_1, u_2, \ldots, u_m\}$. We can find an orthornormal basis of W which is defined from A as follows:

1. Put $v_1 = u_1$.
   $v_2 = u_2 - \frac{u_2 v_1}{v_1 v_1} v_1$
   $v_3 = u_3 - \frac{u_3 v_1}{v_1 v_1} v_1 - \frac{u_3 v_2}{v_2 v_2} v_2$
   $\vdots$
   $v_m = u_m - \frac{u_m v_1}{v_1 v_1} v_1 - \frac{u_m v_2}{v_2 v_2} v_2 - \cdots - \frac{u_m v_{m-1}}{v_{m-1} v_{m-1}} v_{m-1}$. We can choose αv_i in stead of v_i for some nonzero element α.
2. Put $w_1 = \frac{v_1}{||v_1||}, w_2 = \frac{v_2}{||v_2||}, \cdots, w_m = \frac{v_m}{||v_m||}$.

Gram-Schmidt algorithm

Let W be a subspace of $\mathbb{R}^n$ with basis $A = \{u_1, u_2, \ldots, u_m\}$. We can find an orthornormal basis of W which is defined from A as follows:

1. Put $v_1 = u_1$.
   $v_2 = u_2 - \frac{u_2 v_1}{v_1 v_1} v_1$
   $v_3 = u_3 - \frac{u_3 v_1}{v_1 v_1} v_1 - \frac{u_3 v_2}{v_2 v_2} v_2$
   $\vdots$
   $v_m = u_m - \frac{u_m v_1}{v_1 v_1} v_1 - \frac{u_m v_2}{v_2 v_2} v_2 - \cdots - \frac{u_m v_{m-1}}{v_{m-1} v_{m-1}} v_{m-1}$. We can choose αv_i in stead of v_i for some nonzero element α.
2. Put $w_1 = \frac{v_1}{||v_1||}, w_2 = \frac{v_2}{||v_2||}, \cdots, w_m = \frac{v_m}{||v_m||}$.

Then, $B = \{w_1, w_2, \ldots, w_m\}$ is an orthonormal basis of W.

Examples

Find an orthonormal basis of the subspace W of $\mathbb{R}^4$ with basis
$A = \{u_1 = (1, 1, 0, 0), u_2 = (1, 0, -1, 1), u_3 = (0, 1, 1, 1)\}$.
See the whitebroad.

Orthonormal bases of eigenspaces

3.19. Theorem

Let $A \in \mathrm{M}_n(\mathbb{R})$. Assume that A has $\lambda_1, \lambda_2, \ldots, \lambda_r$ eigenvalues. If B_i is the orthonormal basis of eigenspace of E_{λ_i}, then $B_1 \cup B_2 \cup \cdots \cup B_r$ are an orthonormal set.

Algorithm to find an orthogonal Q such that $A = QDQ^t$

Let $A \in \mathrm{M}_n(\mathbb{R})$ be a symmetric matrix.

1. **Find the characteristic polynomial $p_A(t)$: $p_A(t) = |tI_n - A|$.**

Algorithm to find an orthogonal Q such that $A = QDQ^t$

Let $A \in \mathrm{M}_n(\mathbb{R})$ be a symmetric matrix.

1. **Find the characteristic polynomial $p_A(t)$:** $p_A(t) = |tI_n - A|$.
2. **Find the diagonal matrix D:** assume that $p_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$. Then, $D = \mathrm{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})$ is a diagonal in which λ_i appears n_i times in D.

Algorithm to find an orthogonal Q such that $A = QDQ^t$

Let $A \in \mathrm{M}_n(\mathbb{R})$ be a symmetric matrix.

1. **Find the characteristic polynomial $p_A(t)$:** $p_A(t) = |tI_n - A|$.
2. **Find the diagonal matrix D:** assume that $p_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$. Then, $D = \mathrm{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})$ is a diagonal in which λ_i appears n_i times in D.
3. **Find a basis of eigenspace E_{λ_i}:** for each $1 \leq i \leq i$, find a basis A_i of eigenspaces E_{λ_i}. Here E_{λ_i} is the solution space of the system of linear equations with coefficient matrix $A - \lambda_i I_n$.

Algorithm to find an orthogonal Q such that $A = QDQ^t$

Let $A \in \mathrm{M}_n(\mathbb{R})$ be a symmetric matrix.

1. **Find the characteristic polynomial $p_A(t)$:** $p_A(t) = |tI_n - A|$.
2. **Find the diagonal matrix D:** assume that $p_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$. Then, $D = \mathrm{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})$ is a diagonal in which λ_i appears n_i times in D.
3. **Find a basis of eigenspace E_{λ_i}:** for each $1 \leq i \leq i$, find a basis A_i of eigenspaces E_{λ_i}. Here E_{λ_i} is the solution space of the system of linear equations with coefficient matrix $A - \lambda_i I_n$.
4. **Gram–Schmidt orthonormalization of A_i:** find an orthonormal basis B_i from A_i by the Gram-Schmidt algorithm.

Algorithm to find an orthogonal Q such that $A = QDQ^t$

Let $A \in \mathrm{M}_n(\mathbb{R})$ be a symmetric matrix.

1. **Find the characteristic polynomial $p_A(t)$:** $p_A(t) = |tI_n - A|$.
2. **Find the diagonal matrix D:** assume that $p_A(t) = (t-\lambda_1)^{n_1}(t-\lambda_2)^{n_2}\cdots(t-\lambda_r)^{n_r}$. Then, $D = \mathrm{diag}(\lambda_{i_1}, \lambda_{i_2}, \ldots, \lambda_{i_n})$ is a diagonal in which λ_i appears n_i times in D.
3. **Find a basis of eigenspace E_{λ_i}:** for each $1 \leq i \leq i$, find a basis A_i of eigenspaces E_{λ_i}. Here E_{λ_i} is the solution space of the system of linear equations with coefficient matrix $A - \lambda_i I_n$.
4. **Gram–Schmidt orthonormalization of A_i:** find an orthonormal basis B_i from A_i by the Gram-Schmidt algorithm.

The matrix Q is the matrix whose columns are from the vectors from $B_1, B_2, \ldots, B_r$.

Example

Find the SVD of A.

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

Exercises

Find an orthogonal matrix Q and diagonal matrix D such that $A = QDQ^t$ with

1. $A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}$.
2. $A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -1 \\ 2 & -1 & 0 \end{pmatrix}$.
3. $A = \begin{pmatrix} 1 & -3 & -1 \\ -3 & 1 & 1 \\ -1 & 1 & 5 \end{pmatrix}$.
4. $A = \begin{pmatrix} 6 & 2 & 2 \\ 2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}$.