Assessment:

- Three assignments (15% each) due at dates TBD spread over the course.
- Final written examination (55%).

Got question about course material?

- Ask your peers on Ed discussion
- or ask instructor/tutor during office hours and tutorials
- Questions are not to be asked by email.

Office hours :

Regular weeks (subject to change, changes announced on LMS):

- Tu 3.00-4.00: room 189, Peter Hall building, start week 2,
- Wed 3.15-4.15: room 189, PHB
- Fri 11.00-12.00: Zoom, meeting ID 81136149049.

Aug 4-11: see subject overview on LMS.

Other comments

- The tutorials will involve the software R.
- You need to attend the tutorial that was assigned to you as the lists have been designed based on the computer rooms capacity.
- The official slides of the course will be the ones posted AFTER the class.
- I will try to post a version of the slides BEFORE the class. This will NOT be the official version of the slides.

  I will often replace those slides after the class to post the official version of the slides.

1 INTRODUCTION

Goal of the course: Describe, understand and discover properties of data in $p > 1$ dimensions. 多维数据 p是维度

要分析 n个向量 每个向量是p维

- We are interested in analysing a sample of n random vectors $X_1, \ldots, X_n$, also denoted by X_i for $i = 1, \ldots, n$, where X_1, X_2, etc all belong to $\mathbb{R}^p$:

$$\begin{aligned}
X_1 &= (X_{11}, \ldots, X_{1p}) \in \mathbb{R}^p \\
X_2 &= (X_{21}, \ldots, X_{2p}) \in \mathbb{R}^p \\
&\vdots \\
X_n &= (X_{n1}, \ldots, X_{np}) \in \mathbb{R}^p .
\end{aligned}$$

- This means that we have a sample of n individuals, and that for each individual we observe p variables (sometimes called features).

- $X_i = (X_{i1}, \ldots, X_{ip})$ is a row vector

$$X_i^T = \begin{pmatrix} X_{i1} \\ \ldots \\ X_{ip} \end{pmatrix} \text{ is a column vector.}$$

Example:

- We consider a health study involving $n = 100$ patients.

- On each patient we measure $p = 4$ quantities: age, weight, body mass index, systolic blood pressure.

- For the ith individual, where $i = 1, \ldots, 100$ we observe the vector

$$X_i = (X_{i1}, X_{i2}, X_{i3}, X_{i4}) .$$

- X_{i1}=age of ith patient
  X_{i2}=weight of ith patient
  X_{i3}= body mass index of ith patient
  X_{i4}=systolic blood pressure of ith patient.

Often we gather all the observations into a single $n \times p$ matrix

$$\mathcal{X} = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ X_{21} & \dots & X_{2p} \\ & \vdots & \\ X_{n1} & \dots & X_{np} \end{pmatrix}$$

- ☛ Each X_{ij}, for $i = 1, \dots, n$ and $j = 1, \dots, p$ is a random variable.
- ☛ $\mathcal{X}$ has n rows and p columns.
- ☛ The ith row represents the p variables corresponding to the ith individual,
- ☛ the jth column represents the jth variable for all n individuals.

- We call the value taken by $X_i = (X_{i1}, \dots, X_{ip})$ the observed value or realisation.

- Often we use a lower case to denote the observed value: $x_i = (x_{i1}, \dots, x_{ip})$ is the observed value (or realisation) of the random vector X_i.

What we will do in the course:

- The technique learned in basic statistics courses are not focused on multivariate data.

- For ex, how would we do a scatterplot in more than $p = 2$ dimensions? How to represent graphically multivariate data?

- In this course we will learn techniques especially designed for multivariate data.

- This course will introduce various CORE statistical methods for analysing and describing multivariate data.

- Topics such as neural networks, support vector machine, graphical models, etc will NOT be covered in this course.

MATERIAL COVERED IN THE COURSE:

- Revision of basic matrix results and multivariate data needed in the course (tentative schedule: weeks 1 and 2).

  ☛ As this is just a revision, we won't spend much time on examples but the reference book has examples.

- Principal component analysis, factor analysis, partial least squares.

- Classification (linear and quadratic discriminant, regression-based, including trees and forest).

- Clustering (K-means, K-medoids, hierarchical).

- If time permits, other topics such as correspondence analysis or more techniques for above topics.

Main reference books:

- Härdle, W. and Simar, L (2015). Applied multivariate statistical analysis, 4th edition.

  Available for download at the unimelb library website.

- Hastie, T. Tibshirani, R. and Friedman, J. (2009). The elements of statistical learning, 2nd edition (available online):
  `https://hastie.su.domains/Papers/ESLII.pdf`

We might also use

- Koch, I. (2013). Analysis of Multivariate and High-Dimensional Data. Cambridge University Press.
- K. V. Mardia, J. T. Kent, & J. M. Bibby (1979). Multivariate Analysis. New York: Academic Press.
- Other resources as needed.

2 Review of matrix properties

Sections 2.1, 2.2, 2.3, 2.4 and 2.6 in Härdle and Simar (2015).

2.1 Elementary operations

- Let $A = (a_{ij}) = (a_{ij})_{1\leq i\leq n, 1\leq j\leq p}$ be an $n \times p$ matrix (matrix with n rows and p columns):

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ & \vdots & & \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}$$

- We use A^T to denote the transpose of the matrix A:

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ & \vdots & & \\ a_{1p} & a_{2p} & \dots & a_{np} \end{pmatrix}$$

(rows become columns and columns become rows).

- Product of two matrices A, an $n \times p$ matrix, and B, a $p \times m$ matrix:

$$AB = C\,,$$

where C is an $n \times m$ matrix whose (i, j)th element is given by:

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

- Scalar product of two p dimensional vectors $x = (x_1, \ldots, x_p)^T$ and $y = (y_1, \ldots, y_p)^T$:

$$x^T y = \sum_{j=1}^{p} x_j y_j\,.$$

- Norm of a p dimensional vector $x = (x_1, \ldots, x_p)^T$ or $x = (x_1, \ldots, x_p)$:

$$\|x\| = \sqrt{\sum_{j=1}^{p} x_j^2} = \begin{cases} \sqrt{x^T x} & \text{if } x = (x_1, \ldots, x_p)^T \\ \sqrt{x x^T} & \text{if } x = (x_1, \ldots, x_p) \end{cases}$$

Some properties:

$$A + B = B + A$$

$$A(B + C) = AB + AC$$

$$A(BC) = (AB)C$$

$$(A^T)^T = A$$

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

BUT : $AC \neq CA$ 乘法交换律、乘法消除律 没有

Special matrices (taken from Härdle and Simar):

Table 2.1 Special matrices and vectors

Name	Definition	Notation	Example
Scalar	$p = n = 1$	a	3
Column vector	$p = 1$	a	$\begin{pmatrix} 1 \\ 3 \end{pmatrix}$
Row vector	$n = 1$	$a^\top$	$\begin{pmatrix} 1 & 3 \end{pmatrix}$
Vector of ones	$(\underbrace{1, \ldots, 1}_{n})^\top$	1_n	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Vector of zeros	$(\underbrace{0, \ldots, 0}_{n})^\top$	0_n	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$
Square matrix	$n = p$	$\mathcal{A}(p \times p)$	$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$
Diagonal matrix	$a_{ij} = 0, i \neq j, n = p$	$\mathrm{diag}(a_{ii})$	$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

Identity matrix	$\text{diag}(\underbrace{1, \ldots, 1}_{p})$	$\mathcal{I}_p$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
Unit matrix	$a_{ij} = 1, n = p$	$1_n 1_n^\top$	$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$
Symmetric matrix	$a_{ij} = a_{ji}$		$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$
Null matrix	$a_{ij} = 0$	0	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$
Upper triangular matrix	$a_{ij} = 0, i < j$		$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$
Idempotent matrix	$\mathcal{A}\mathcal{A} = \mathcal{A}$		$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$
Orthogonal matrix	$\mathcal{A}^\top \mathcal{A} = \mathcal{I} = \mathcal{A}\mathcal{A}^\top$ $A^T = A^{-1}$		$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$

各列长为1

各列正交

点乘=0

- A set of k rows (or columns) $a_1, \ldots, a_k$ of A are said to be linearly independent if none of them can be expressed as a nontrivial linear combination of the other $k-1$ rows (or columns). It is not possible to write

$$a_j = \sum_{i \neq j} c_i a_i \, ,$$

对任意列(或行)都没法表达为其他列(或行)的线性组合.

where the c_i's are all different from zero, or equivalently

$$\sum_{i=1}^{k} c_i a_i = 0 \Rightarrow c_1, \ldots, c_k = 0 \, .$$

对一个矩阵 / 一组vector 线性无关的 / 列 使等式成立的唯一解是 $c_1 \cdots c_k = 0$

- Rank: The rank of an $n \times p$ matrix A, denoted by $\text{rank}(A)$ is defined as the maximum number of linearly independent rows (or columns).

最大线性无关的行列的数量.

We always have

$$\text{rank}(A) \leq \min(n, p) \, .$$

n, p的最小值

行列

- Trace: The trace of a square $p \times p$ matrix A, denoted by $\mathrm{tr}(A)$, is the sum of its diagonal elements:

$$\mathrm{tr}(A) = \sum_{i=1}^{p} a_{ii} \, .$$

- Determinant: The determinant of a square $p \times p$ matrix A, denoted by $\det(A)$ or $|A|$, is a number computed from the matrix and which plays an important role in all sorts of problems. For a 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

it is computed by

$$|A| = a_{11}a_{22} - a_{21}a_{12} \, .$$

For larger matrices, we compute this recursively. If you forgot how a determinant is computed, see

`https://en.wikipedia.org/wiki/Determinant` square matrix

- If rank of $p \times p$ matrix A not full $\boxed{(\text{rank}(A) < p)}$ then $\det(A) = 0$. 不满秩 $\det(A)=0$

- Inverse: If $|A| \neq 0$, the inverse of a square $p \times p$ matrix A exists. It is denoted by A^{-1} and is such that

full rank $\Rightarrow$ $|A| \neq 0$ $\det(A) \neq 0$ $\Rightarrow$ non singular.

$$AA^{-1} = A^{-1}A = I_p$$

($I_p = p \times p$ identity matrix).

See Härdle and Simar, page 57, for how to compute the inverse.

We have

$$\boxed{|A^{-1}| = 1/|A|.}$$

- A non invertible square matrix is also said to be singular.

$\|$

not full rank matrix

$\|$

$\det(A)=0$

- Eigenvalues and eigenvectors of a square $p \times p$ matrix A:

  The (non zero) $p \times 1$ vector v is an eigenvector of A with eigenvalue λ if it is such that

  $$Av = \lambda v\,.$$

  Note that λ is a real number (not a vector).

  ☛ If A is symmetric then there are p eigenvalues and eigenvectors.

  ☛ All eigenvalues satisfy

  $$\boxed{|A - \lambda I_p| = 0}$$ 解行列式

  (they are the p roots of the above polynomial of order p in λ).

  ☛ The eigenvalues are not necessarily all different from each other.

  eigenvalues 不一定相同

- ☛ Constant multiples of an eigen vector v with eigenvalue λ are also eigenvectors with eigenvalue λ. 特征向量的数量是无限的

- ☛ It is common to work with scaled version of eigenvectors that have norm 1:

$$\|v\| = \sqrt{v^T v} = 1\,.$$

$|V| = \sqrt{\|V\|}$ 开根

- ☛ In practice we often compute them with a software, e.g. R.

- Suppose the square (满秩) $p \times p$ matrix A has eigenvalues $\lambda_1, \ldots, \lambda_p$.

  Let Λ be the diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$ with the λ_i's on the diagonal and 0 everywhere else. A的对角矩阵 $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$

  Then we have

$$\det(A) = (|A| = |\Lambda|) = \prod_{i=1}^{p} \lambda_i \quad \text{(连乘)} = \det(A)\,.$$

product of. eigenvalues

and

满秩矩阵

$$\boxed{\mathrm{tr}(A) = \mathrm{tr}(\Lambda) = \sum_{i=1}^{p} \lambda_i\,.}$$

- If A and B are two invertible matrices of the same dimension, then

$$(A \cdot B)^{-1} = B^{-1}A^{-1}\,.$$

Some other properties of matrix characteristics (from Härdle and Simar):

$\mathcal{A}(n \times p), \mathcal{B}(p \times q), \mathcal{C}(q \times n)$

$$= \mathrm{tr}(\mathcal{A}\mathcal{B}\mathcal{C}) = \mathrm{tr}(\mathcal{C}\mathcal{A}\mathcal{B}) = \mathrm{tr}(\mathcal{B}\mathcal{C}\mathcal{A})$$

$\mathcal{A}(n \times n),\ \mathcal{B}(n \times n),\ c \in \mathbb{R}$

$$\mathrm{tr}(\mathcal{A} + \mathcal{B}) = \mathrm{tr}\,\mathcal{A} + \mathrm{tr}\,\mathcal{B}$$

$$\mathrm{tr}(c\mathcal{A}) = c\ \mathrm{tr}\,\mathcal{A}$$

$$\boxed{|c\mathcal{A}| = c^n|\mathcal{A}|}$$

$$|\mathcal{A}\mathcal{B}| = |\mathcal{B}\mathcal{A}| = |\mathcal{A}||\mathcal{B}|$$

$\mathcal{A}(n \times p),\ \mathcal{B}(p \times n)$

$$\operatorname{tr}(\mathcal{A}\cdot \mathcal{B}) = \operatorname{tr}(\mathcal{B}\cdot \mathcal{A})$$

$$\operatorname{rank}(\mathcal{A}) \leq \min(n, p)$$

$$\operatorname{rank}(\mathcal{A}) \geq 0$$

$$\operatorname{rank}(\mathcal{A}) = \operatorname{rank}(\mathcal{A}^\top)$$

$$\boxed{\operatorname{rank}(\mathcal{A}^\top \mathcal{A}) = \operatorname{rank}(\mathcal{A})}$$

有时候算 $A^T A$ might be easier.

$$\operatorname{rank}(\mathcal{A} + \mathcal{B}) \leq \operatorname{rank}(\mathcal{A}) + \operatorname{rank}(\mathcal{B})$$

$$\operatorname{rank}(\mathcal{A}\mathcal{B}) \leq \min\{\operatorname{rank}(\mathcal{A}), \operatorname{rank}(\mathcal{B})\}$$

Errata: The dimensions of $\mathcal{A}$ and $\mathcal{B}$ have to match for $\mathcal{A} + \mathcal{B}$ to be defined in

$$\operatorname{rank}(\mathcal{A} + \mathcal{B}) \leq \operatorname{rank}(\mathcal{A}) + \operatorname{rank}(\mathcal{B}),$$

Spectral decomposition

- Spectral decomposition: Suppose A is a square and symmetric $p \times p$ matrix and let

  解 eigenvalue

  $\lambda_1, \ldots, \lambda_p$ denote its p eigenvalues and 不一定全不同

  $v_1, \ldots, v_p$ denote the associated $p \times 1$ eigenvectors of norm 1 and orthogonal to each other (each v_k is the eigenvector associated to λ_k).

Note: two $p \times 1$ column vectors v and w are orthogonal if

$$v^T w = \sum_{i=1}^{p} v_i w_i = 0 .$$

列向量乘得0.

Then we can always express A in the following way, which is called spectral decomposition of A:

matrix

symmetric
square

有了特征值，特征根
我们都能将矩阵表示成这样

eigenvector

$$\underset{p \times p}{A} = \sum_{j=1}^{p} \underbrace{\lambda_j}_{\text{eigenvalue}} v_j v_j^T .$$

This can also be written in matrix form, if we let

$$\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$$

and

Gamma

eigenvectorset

$$\Gamma = (v_1, \ldots, v_p)$$

the $p \times p$ orthogonal matrix whose columns are the p eigenvectors standardised to be of norm 1: (标准特征向量) 每个维除以norm

$$A = \Gamma \Lambda \Gamma^T .$$

p

norm 1

- In the above notation, if $A = \Gamma\Lambda\Gamma^T$ then if we take an integer power of A, for example A^α, we find

$$A^\alpha = \Gamma\Lambda^\alpha\Gamma^T. \qquad A^\alpha = \Gamma\Lambda^\alpha\Gamma^T$$

This is because the v_j's are orthonormal (i.e. orthogonal and of norm 1).

For example

$$A^2 = \Gamma\underset{\text{特征值对角矩阵}}{\Lambda}\overset{I}{\overset{\shortparallel}{\Gamma^T\Gamma}}\Lambda\Gamma^T = \Gamma\Lambda^2\Gamma^T\,.$$

This also works for negative powers if A is invertible (which happens if and only if the eigenvalues are all nonzero). For example,

满秩矩阵

$$A^{-1} = \Gamma\Lambda^{-1}\Gamma^T$$

(A^{-1}: the inverse of the matrix A).

Singular value decomposition

More generally, a similar decomposition exists for matrices that are not especially square matrices. In particular, any $n \times p$ matrix A with rank r can be decomposed as

3 2

不同矩阵

rank ≤ min(3,2)

$$A = \Gamma \Lambda \Delta^T$$

$n \times p$ $\quad$ $n\times r \;\; r\times r \;\; r \times p$

rank : # of non-zero eigenvalues of matrix

where the $n \times r$ matrix Γ and the $p \times r$ matrix Δ are column orthonormal, i.e. their columns are orthonormal (i.e., orthogonal and of norm 1), that is

$$\Gamma^T \Gamma = \Delta^T \Delta = I_r$$

Γ、Δ都正交

and

$$\Lambda = \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_r^{1/2})$$

where each $\lambda_i > 0$.

非零特征根的开方

The λ_i's are the nonzero eigenvalues of the matrices AA^T or A^TA; the columns of Γ and Δ are the corresponding r eigenvectors of those matrices, respectively.

Δ: eigenvectors of A^TA

Γ: eigenvectors of AA^T

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}_{2\times 3}$$

rank(A)=2 because it's impossible to express $(1,1,0) = c \cdot (0,1,1)$

$$AA^T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}_{2\times 3} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}_{3\times 2} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$\rightarrow$ # of eigenvalues = 2

$\text{rank}(A^TA) = \text{rank}(A) \Rightarrow 2$

$$\det(AA^T - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0$$

$$\lambda^2 - 4\lambda + 4 - 1 = 0$$

$$(\lambda - 2)^2 = 1$$

$$\begin{cases} \lambda - 2 = 1 \Rightarrow \lambda_1 = 3 \\ \lambda - 2 = -1 \Rightarrow \lambda_2 = 1 \end{cases}$$

$V_1 = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}$ associated to $\lambda_1 = 1$

$$AA^T \cdot V_1 = \lambda_1 \cdot V_1$$

$$\Rightarrow \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = 1 \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix} = 3 \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix}$$

$2\times2\times2\times1$

$$\begin{pmatrix} 2v_{11}+v_{12} \\ v_{11}+2v_{12} \end{pmatrix} = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix}$$

$$\begin{cases} v_{12} = -v_{11} \\ v_{11} = -v_{12} \end{cases} \Rightarrow \begin{pmatrix} -v_{12} \\ v_{12} \end{pmatrix} \Rightarrow \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

e.ve$_1$

$$\|v_1\| = \sqrt{1+1} = \sqrt{2}$$

normalize: $v_1 = \begin{pmatrix} -\frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} \end{pmatrix}$

$$\begin{pmatrix} 2v_{21}+v_{22} \\ v_{21}+2v_{22} \end{pmatrix} = \begin{pmatrix} 3v_{21} \\ 3v_{22} \end{pmatrix}$$

$$\begin{cases} v_{22} = v_{21} \\ v_{21} = v_{22} \end{cases} \Rightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\|v_2\| = \sqrt{2}$$

normalize $v_2 = \begin{pmatrix} \frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} \end{pmatrix}$

$$\begin{array}{cc} \lambda_1 & \lambda_2 \end{array}$$

$$\Gamma = \begin{pmatrix} -\frac{1}{\sqrt2} & \frac{1}{\sqrt2} \\ \frac{1}{\sqrt2} & \frac{1}{\sqrt2} \end{pmatrix}$$ (orthogonal matrix)

$$A^T A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$3\times2\times2\times3$

rank 相同 $r(A^TA) = r(AA^T) \Rightarrow$ 非零 特征值个数相同

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{pmatrix} v_{11} \\ v_{12} \\ v_{13} \end{pmatrix} = \lambda \begin{pmatrix} v_{11} \\ v_{12} \\ v_{13} \end{pmatrix}$$

$v_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ $\quad v_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ $\quad v_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

$\lambda_1 = 3$ $\quad \lambda_2 = 1$ $\quad \lambda_3 = 0$

$1+4+1=6$ $\quad$ norm $= \frac{1}{\sqrt6}$ $\quad$ norm: $\frac{1}{\sqrt2}$ $\quad$ norm $= \frac{1}{\sqrt3}$

$$\Delta = \begin{pmatrix} \frac{1}{\sqrt6} & \frac{1}{\sqrt2} & \frac{1}{\sqrt3} \\ \frac{2}{\sqrt6} & 0 & -\frac{1}{\sqrt3} \\ \frac{1}{\sqrt6} & \frac{1}{\sqrt2} & \frac{1}{\sqrt3} \end{pmatrix}$$

$$\begin{array}{ccc} \lambda_1 & \lambda_2 & \lambda_3 \end{array}$$

AA^T的特征向量

$\Gamma \Lambda \Delta^T$

$$\Lambda = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \sqrt{\lambda_2} & \\ & & \sqrt{3} \\ & & & 0 \end{bmatrix}$$

$$A = \Gamma \Lambda \Delta^T$$