数学建模

线性代数模型

背景

- 线性代数作为一门重要的数学工具，不仅在数学各领域有重要的理论意义，不在物理、生物、化学、医学、生产管理等各方面有着广泛而重要的应用。
- 作为教学建模来说，掌握基本的线性代数方法也是必要的。
- 本次课程主要结合一些实际问题说明如何建立线性代数模型，并利用软件工具进行辅助分析求解。
- 首先以简单的案例介绍线性代数模型，其主要目的是通过这些案例加深对线性代数基本概念、相论和方法的理解，培养数学建模的意识。
- 以矩阵分解方法在推荐系统中的应用为例，介绍线性代数在信息专业中的几个典型应用。

线性代数模型

1. Matlab线性代数工具简介

2. 线性方程模型

3. 矩阵分解与应用

Matlab线性代数工具简介

- 符号行列式的计算
- 线性方程组的求解
- 特征值和特征向量的计算

符号行列式的计算

- **例1：**
- **计算** $$\begin{vmatrix} x-y & 1 & 1 & 1 \\ 1 & x-y & 1 & 1 \\ 1 & 1 & x+y & 1 \\ 1 & 1 & 1 & x+y \end{vmatrix}$$

```
%% 符号行列式
syms x y
A=[x-y 1 1 1;1 x-y 1 1;1 1 x+y 1;1 1 1 x+y];
det(A)
```

解方程

- **例2：**

$$\begin{vmatrix} 3 & 2 & 1 & 1 \\ 3 & 2 & 2-x^2 & 1 \\ 5 & 1 & 3 & 2 \\ 7-x^2 & 1 & 3 & 2 \end{vmatrix} = 0$$

```
%% 解方程
clear all
clc
syms x
A=[3 2 1 1;3 2 2-x^2 1;5 1 3 2;7-x^2 1 3 2];
D=det(A)
f=factor(D)
X=solve(D)
```

线性方程组求解

- **求矩阵A的最简行阶梯矩阵**

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

```
%% 线性方程组求解
clear all
clc
A=[2 1 -1 3 1; 4 2 -5 1 2 ;2 1 -1 -1 1];
rref(A)  %给出A的最简行阶梯形矩阵
% linesolve(A,y, options) %求解Ax=y
% null(A,'r') %求齐次方程组Ax=0的基础解系
```

```
ans =

    1.0000    0.5000         0         0    0.5000
         0         0    1.0000         0         0
         0         0         0    1.0000         0
```

特征值和特征向量的计算

- **方阵的特征值与特征向量在矩阵对角化和微分方程组求解等问题中有着广泛的应用，它可以用来分析矩阵的对角化问题，二次型化标准形以及二次型的正定性等问题.在MATLAB中与本实验相关的命令。**

- **例：求矩阵A的特征值和特征向量**

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

```
clear all
clc
A=[-1 -2 0;2 3 0;2 1 3];
P=poly(A) %求A的特征多项式
roots(P)%求多项式P的零点
orth(A)%求出矩阵A的列向量构成空间的一个规范正交基
[V,U]=eig(A)
%返回方阵A的特征值和特征向量，其中U的对角元素为A的特征值，V
的列向量为特征值对应的特征向量
```

线性代数模型

1. Matlab线性代数工具简介

2. 线性方程模型

3. 矩阵分解与应用

例1. 投入产出模型

- 某地区有3个重要产业， 即煤矿、发电厂和地方铁路。经成本核算，开采一元的煤，煤矿要支付0.3元的电费及0.25元的运输费。生产1元的电力，发电厂要支付0.65元的煤作燃料，为了运行电厂的辅助设备需消耗本身0.15元的电，还需要0.1元的运输费:作为铁路局，每创收1元的运输费， 铁路要支付0.6元的煤费及0.15元的电费。

- 在某一周内，煤矿接到金额为8万元的外地订货，发电厂接到金额为6.5万元的外地需求，外界对地方铁路没有需求，问3个企业在这一周内，总产值各多少才能满足自身及外界的需求? 3个企业间相互支付多少金额?3个企业各创造多少新价值?

模型求解

- **模型假设:为了简化问题，不考虑价格变动等其他因素的影响.**
- **模型建立:设煤矿、发电厂和铁路本周内的总产值依次为x1，x2，x3元，则根据需求有如下分配平衡方程组**

$$\begin{cases} x_1-(0\times x_1+0.65x_2+0.6x_3)=80000 \\ x_2-(0.3x_1+0.15x_2+0.15x_3)=65000 \\ x_3-(0.25x_1+0.1x_2+0\times x_3)=0 \end{cases}$$

$$X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, A=\begin{pmatrix} 0 & 0.65 & 0.6 \\ 0.3 & 0.15 & 0.15 \\ 0.25 & 0.1 & 0 \end{pmatrix}, Y=\begin{pmatrix} 80000 \\ 65000 \\ 0 \end{pmatrix}$$

- **求解**

$$(E-A)X=Y$$

投入产出分析表

$$\boldsymbol{B}=(\boldsymbol{E}-\boldsymbol{A})^{-1}-\boldsymbol{E},\quad \boldsymbol{C}=\boldsymbol{A}\begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix},\quad \boldsymbol{D}=(1,1,1)\boldsymbol{C}$$

- **矩阵B称为完全消耗矩阵，它与矩阵A一起在各个部门之间的投入产出中起平衡作用。矩阵C可以称为投入产出矩阵，它的元素表示煤矿、发电厂、铁路之间的投入产出关系，其中每一行给出了每一个企业分别用于企业内部和其他企业的消耗。向量D称为总投入向量，它的元素是矩阵C的对应列元素之和， 分别表示煤矿、发电厂、铁路得到的总投入。**

- **由矩阵C,向量Y, X和D,可得投入产出分析表。**

- **Matlab求解如下：**

```
%% 投入产出模型
A=[0 0.65 0.6;0.3 0.15 0.15;0.25 0.1 0];
Y=[80000 65000 0]';
E=eye(3);
X=(E-A)\Y
B=inv(E-A)-E
C=A*diag (X) %计算投入产出矩阵
D=ones(1,3)*C %计算总投入向量
F=X-D' %计算新创造价值向量
```

投入产出计算结果

	煤矿	发电厂	铁路	外接需求	总产出
煤矿	0	113720	46540	80000	240260
发电厂	72080	26240	11630	65000	174950
铁路	60060	17500	0	0	77560
新创造价值	108110	17500	19390		
总投入	132140	157460	58170		

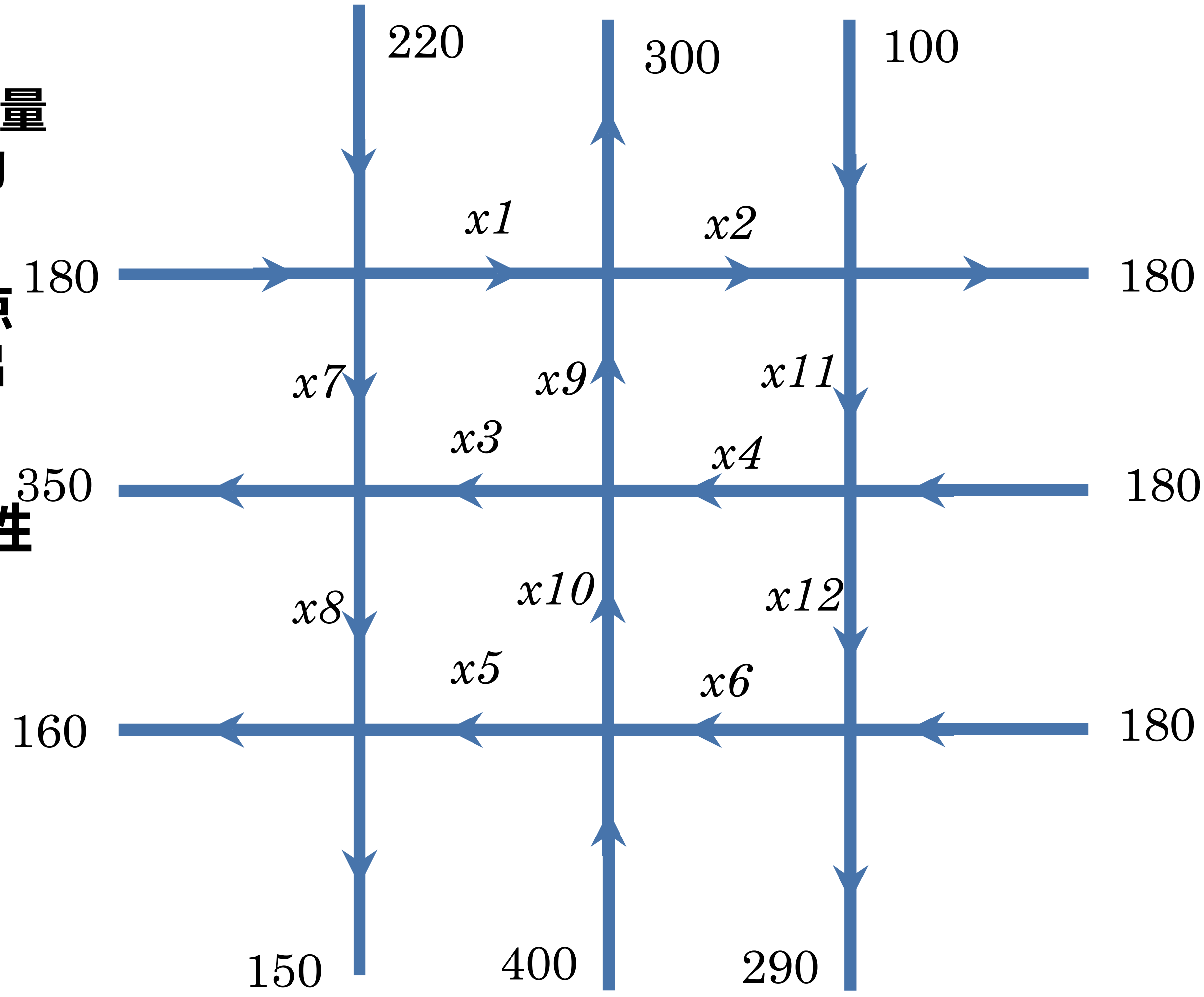
例2. 交通流量模型

- 问题背景：城市道路网中每条道路、每个交叉路口的车流量调查，是分析、评价及改善城市交通状况的基础.根据实际车流量信息可以设计流量控制方案，必要时设置单行线，以免大量车辆长时间拥堵.下面考虑一 简化的交通流量统计问题.

- 问题描述:某城市的交通图，如图所示，每条道路都是单行线，需要调查每条道路每小时的车流量.图中的数字表示该条路段的车流数.如果每个交叉路口进入和离开的车数相等，整个图中进入和离开的车数相等. 试解决以下问题:
  - 1)建立确定每条道路流量的线性方程组.
  - 2)分析哪些流量数据是多余的.
  - 3)为了唯一确定未知流量，需要增添哪几条道路的流量统计

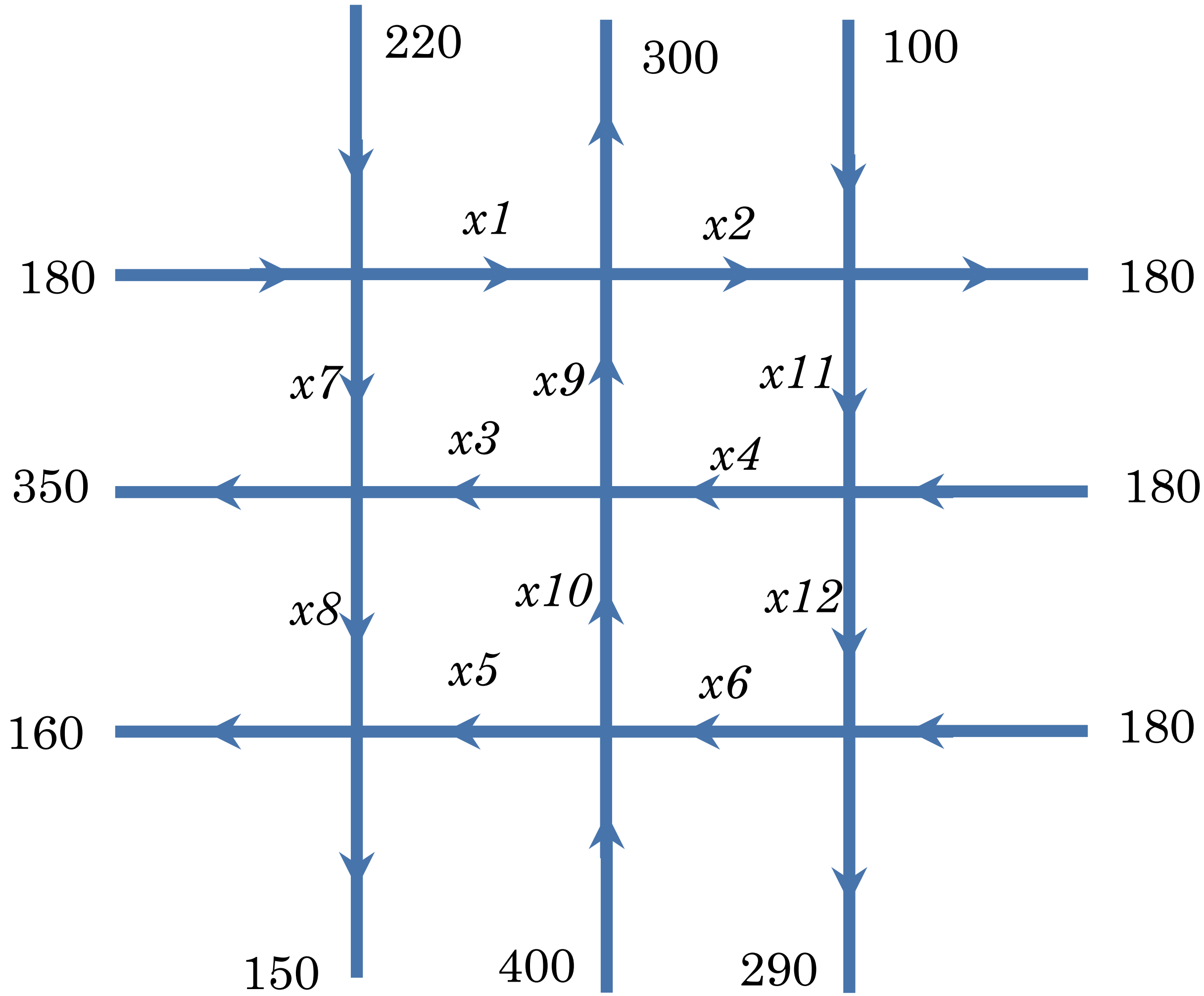
交通流量模型

- **模型假设：**
  - **1.全部流入网络的流量等于全部流出网络的流量**
  - **2. 全部流入每个节点的流量等于全部流出此节点的流量**
- **由流量假设，建立线性方程组模型：**



交通流量模型

$$
\begin{cases}
180 + 220 = x_1 + x_7 \\
x_1 + x_9 = 300 + x_2 \\
x_2 + 100 = 300 + x_{11} \\
x_3 + x_7 = 350 + x_3 \\
x_4 + x_{10} = x_3 + x_9 \\
x_{11} + 500 = x_4 + x_{12} \\
x_5 + x_8 = 160 + 150 \\
x_6 + 400 = x_5 + x_{10} \\
x_{12} + 150 = x_6 + 290
\end{cases}
$$



交通流量模型

- **模型求解**

$$\begin{cases} 180+220=x_1+x_7 \\ x_1+x_9=300+x_2 \\ x_2+100=300+x_{11} \\ x_3+x_7=350+x_3 \\ x_4+x_{10}=x_3+x_9 \\ x_{11}+500=x_4+x_{12} \\ x_5+x_8=160+150 \\ x_6+400=x_5+x_{10} \\ x_{12}+150=x_6+290 \end{cases}$$

$$\begin{cases} x_1+x_7=400 \\ x_1-x_2+x_9=300 \\ x_2-x_{11}=200 \\ x_3+x_7-x_8=350 \\ x_3-x_4+x_9-x_{10}=0 \\ x_4-x_{11}+x_{12}=500 \\ x_5+x_8=310 \\ x_5-x_6+x_{10}=400 \\ -x_6+x_{12}=140 \end{cases}$$

$$A=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

交通流量模型

- **结果分析**

$$B=\begin{pmatrix}1&0&0&0&0&0&0&0&1&0&-1&0&500\\0&1&0&0&0&0&0&0&0&0&-1&0&200\\0&0&1&0&0&0&0&0&1&-1&-1&1&500\\0&0&0&1&0&0&0&0&0&0&-1&1&500\\0&0&0&0&1&0&0&0&0&1&0&-1&260\\0&0&0&0&0&1&0&0&0&0&0&-1&-140\\0&0&0&0&0&0&1&0&-1&0&1&0&-100\\0&0&0&0&0&0&0&1&0&-1&0&1&50\\0&0&0&0&0&0&0&0&0&0&0&0&0\end{pmatrix}$$

$$\begin{cases}x_1=-x_9+x_{11}+500\\x_2=x_{11}+200\\x_3=-x_9+x_{10}+x_{11}-x_{12}+500\\x_4=x_{11}-x_{12}+500\\x_5=-x_{10}+x_{12}+260\\x_6=x_{12}-140\\x_7=x_9-x_{11}-100\\x_8=x_{10}-x_{12}+50\end{cases}$$

取(x9, x10, x11, x12)为自由未知量，分别赋4组值为(1,0,0,0)，(0, 1,0, 0)，(0,0,1,0)，(0,0,0, 1)，则对应齐次方程组基础解系中4个解向量为

$\eta_1=(-1,0,-1,0,0,0,1,0,1,0,0,0)^{\mathrm{T}}$, $\eta_2=(0,0,1,0,-1,0,0,1,0,1,0,0)^{\mathrm{T}}$

$\eta_3=(1,1,1,1,0,0,-1,0,0,0,1,0)^{\mathrm{T}}$, $\eta_4=(0,0,-1,-1,1,1,0,-1,0,0,0,1)^{\mathrm{T}}$

$$\boldsymbol{\eta}^*=(500,200,500,500,260,-140,-100,50,0,0,0,0)^{\mathrm{T}}$$

通解为：$x=k_1\eta_1+k_2\eta_2+k_3\eta_3+k_4\eta_4+\eta^*$

交通流量模型

$$\boldsymbol{\eta}^* = (500, 200, 500, 500, 260, -140, -100, 50, 0, 0, 0, 0)^{\mathrm{T}}$$

通解为：$x = k_1\eta_1 + k_2\eta_2 + k_3\eta_3 + k_4\eta_4 + \eta^*$

- **结果分析:**
  - **从增广矩阵的阶梯形行最简形式可以看出，式中最后一个方程是多余的，即最后一个方程中的流量数据150 和290不用统计。而如果要唯一确定未知流量，可以增添x9, x10, x11, x12 (x2, x9，x10, x12等) 这几条道路的流量统计. 在特解中，x6 和x7为负，这表明对应的两条单行线应该改变方向才合理.**

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 500 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 200 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 & 500 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 500 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 260 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & -140 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & -100 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 1 & 50 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

线性代数模型

1. Matlab线性代数工具简介

2. 线性方程模型

3. 矩阵分解与应用

矩阵分解在数据处理中的应用

- **Principal Component Analysis (PCA)**
  - **-Draw a plane closest to data points (Pearson, 1901)**
  - **-Retain most variance of the data (Hotelling, 1933)**
- **Singular Value Decomposition (SVD)**
  - **-Low-rank approximation (Eckart-Young, 1936)**
  - **-Practical applications or Efficient Computations (Golub-Kahan, 1965)**
- **Matrix Factorization In Recommender Systems**
  - **-“Factorization Meets the Neighborhood: a Multifaceted Collaborative Filtering Model”，by Yehuda Koren, ACM KDD 2008**

基于矩阵分解的方法

- **主成分分解 PCA**
- **奇异值分解 SVD**
- **低秩矩阵分解 low rank MF**

主成分分解(PCA)

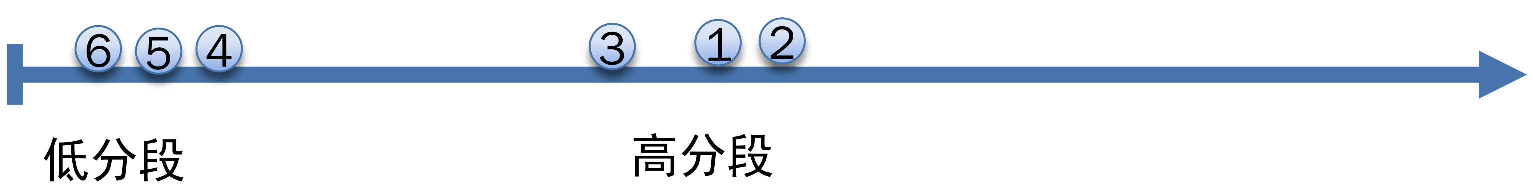
- **例：评估以下两门课程C1，C2的6位学生S1-S2的成绩…**

	S1	S2	S3	S4	S5	S6
C1	9	10	7	3	2	1
C2	6	4	5	3	2.8	1

主成分分解(PCA)

- **首先看一门课程**

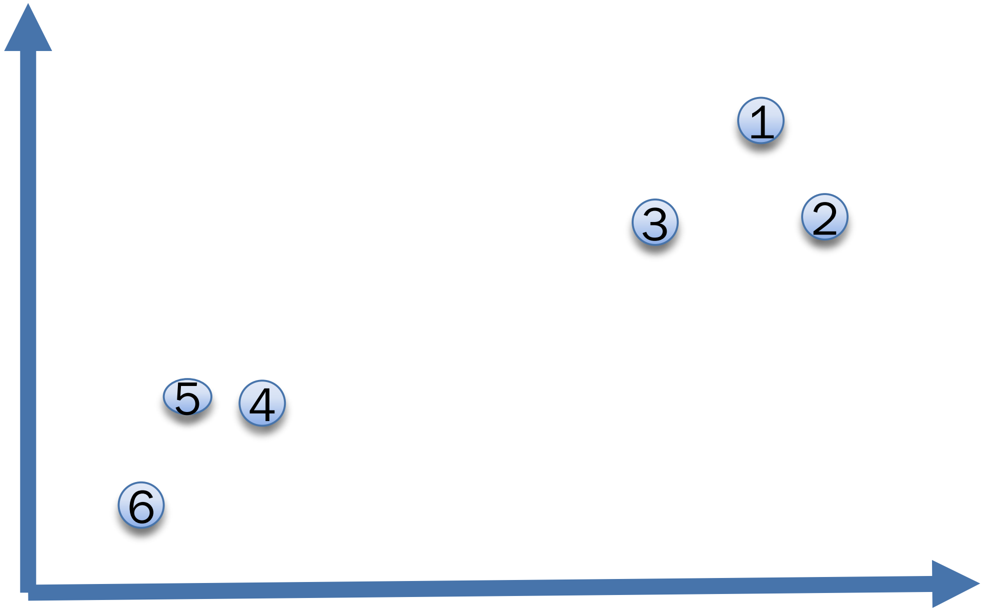
	S1	S2	S3	S4	S5	S6
C1	9	10	7	3	2	1



主成分分解(PCA)

- **分别用两个坐标轴反映两门课程的成绩**

	S1	S2	S3	S4	S5	S6
C1	9	10	7	3	2	1
C2	6	4	5	3	2.8	1



主成分分解(PCA)

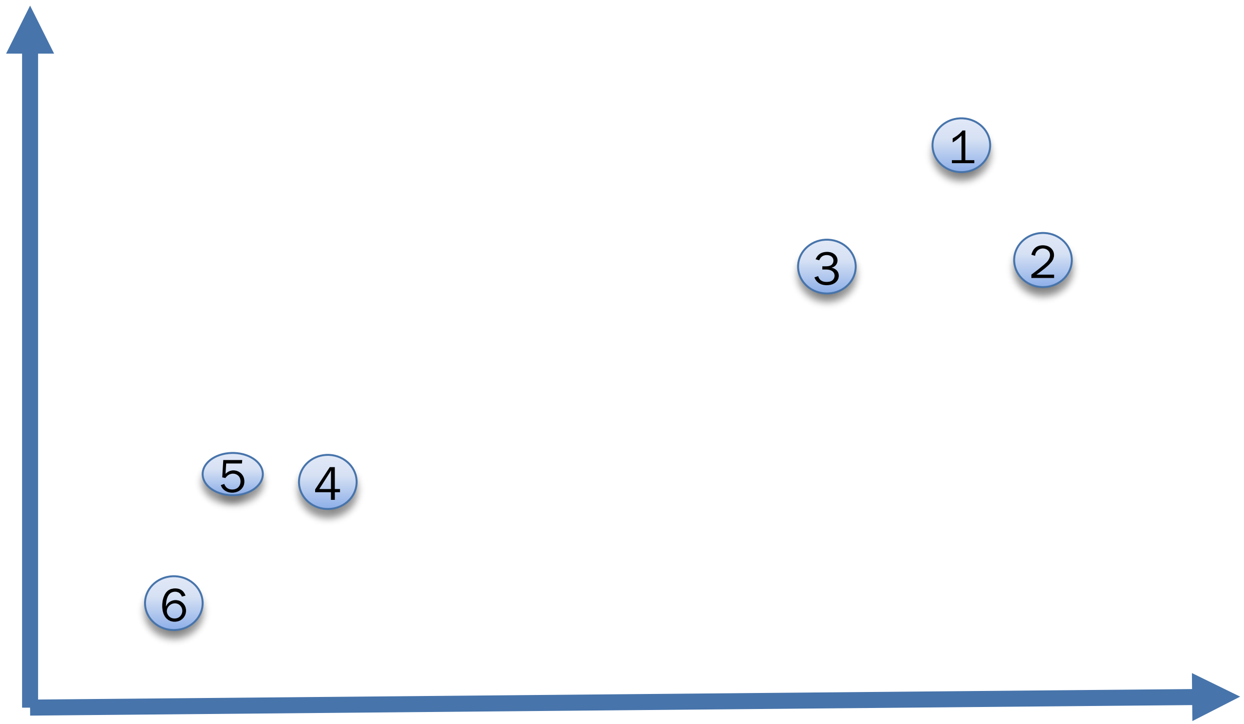
- **如果用更多的课程?**

	S1	S2	S3	S4	S5	S6
C1	9	10	7	3	2	1
C2	6	4	5	3	2.8	1
C3	10	8	7	2.5	1.3	2
C4	5	7	6	2	4	7

因此希望通过数据降维，在低维空间（如二维空间）呈现样本的主成分特征。

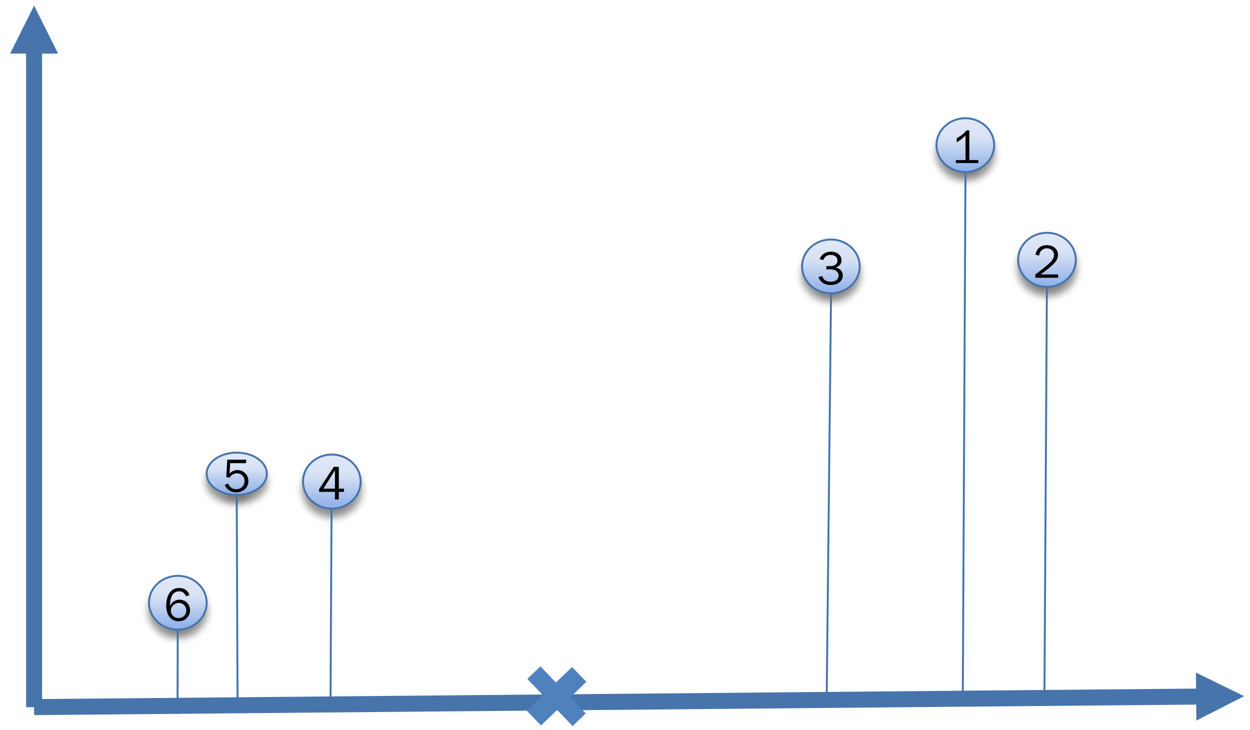
主成分分解(PCA)

以二维特征样本数据为例，分析降维方法。



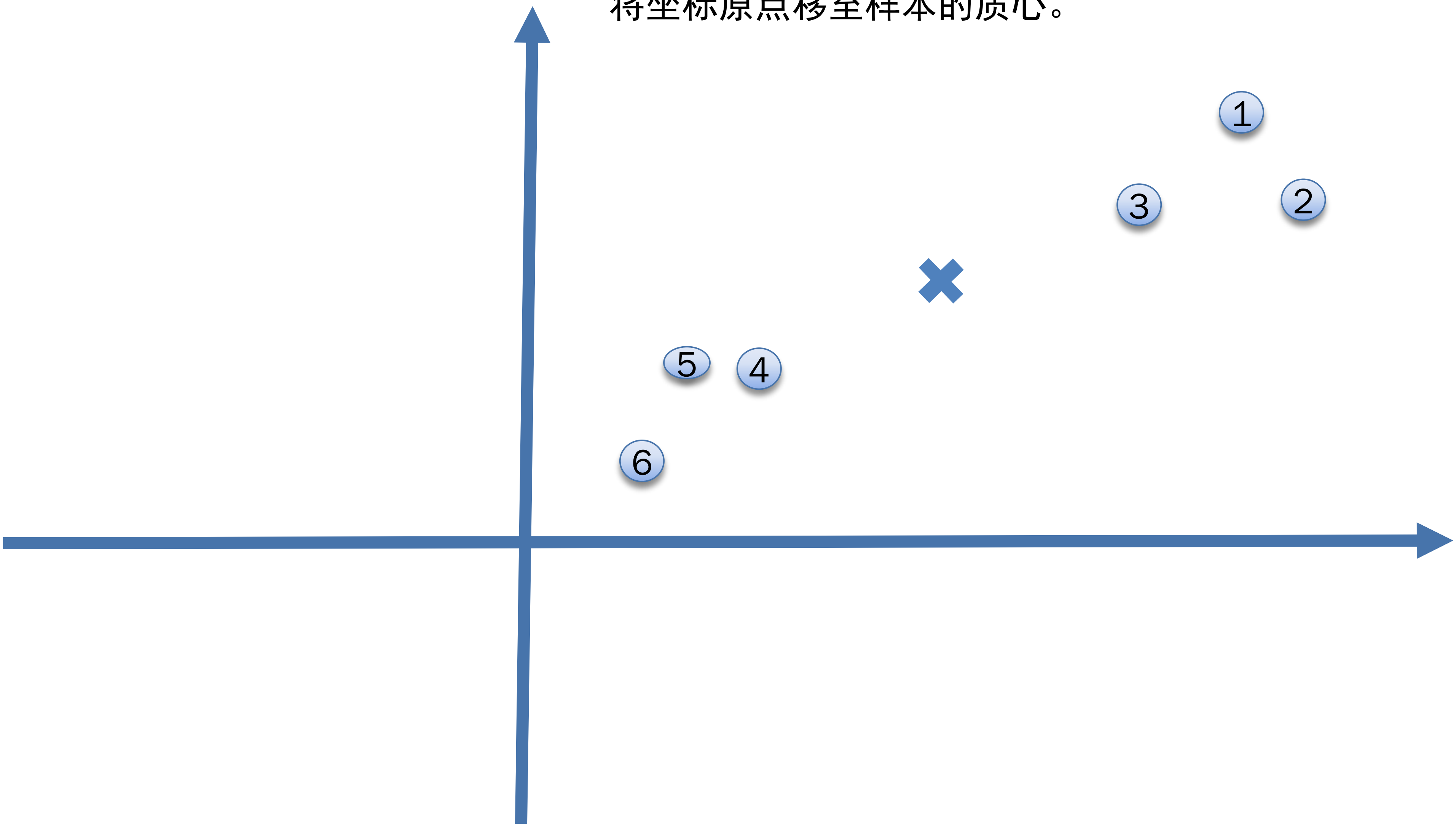
主成分分解(PCA)

计算每个维度特征的均值，获得样本质心在每个维度上的坐标



主成分分解(PCA)

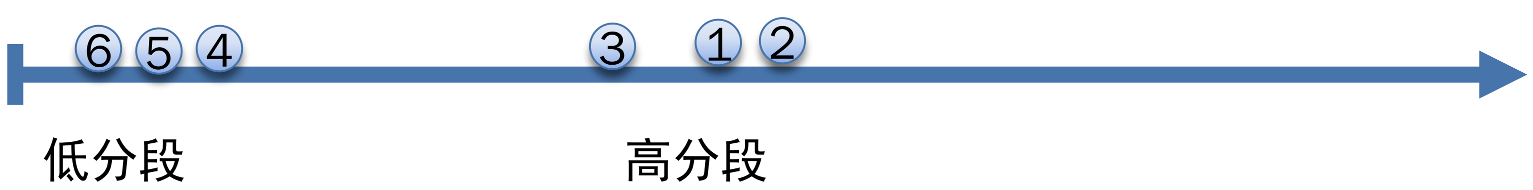
将坐标原点移至样本的质心。



主成分分解(PCA)

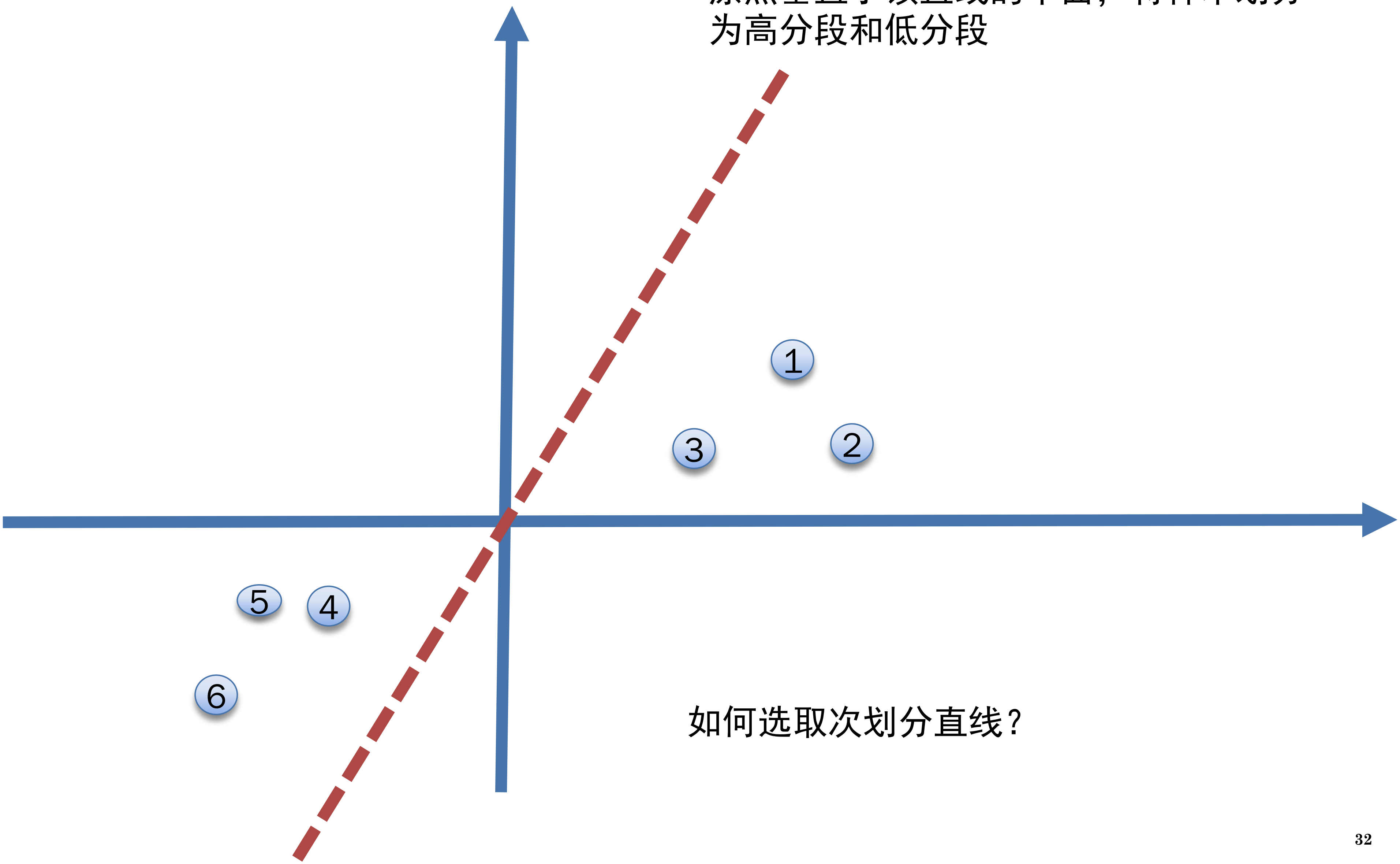
- **考虑对高维数据，做如下高低分段的划分**

	S1	S2	S3	S4	S5	S6
C1	9	10	7	3	2	1



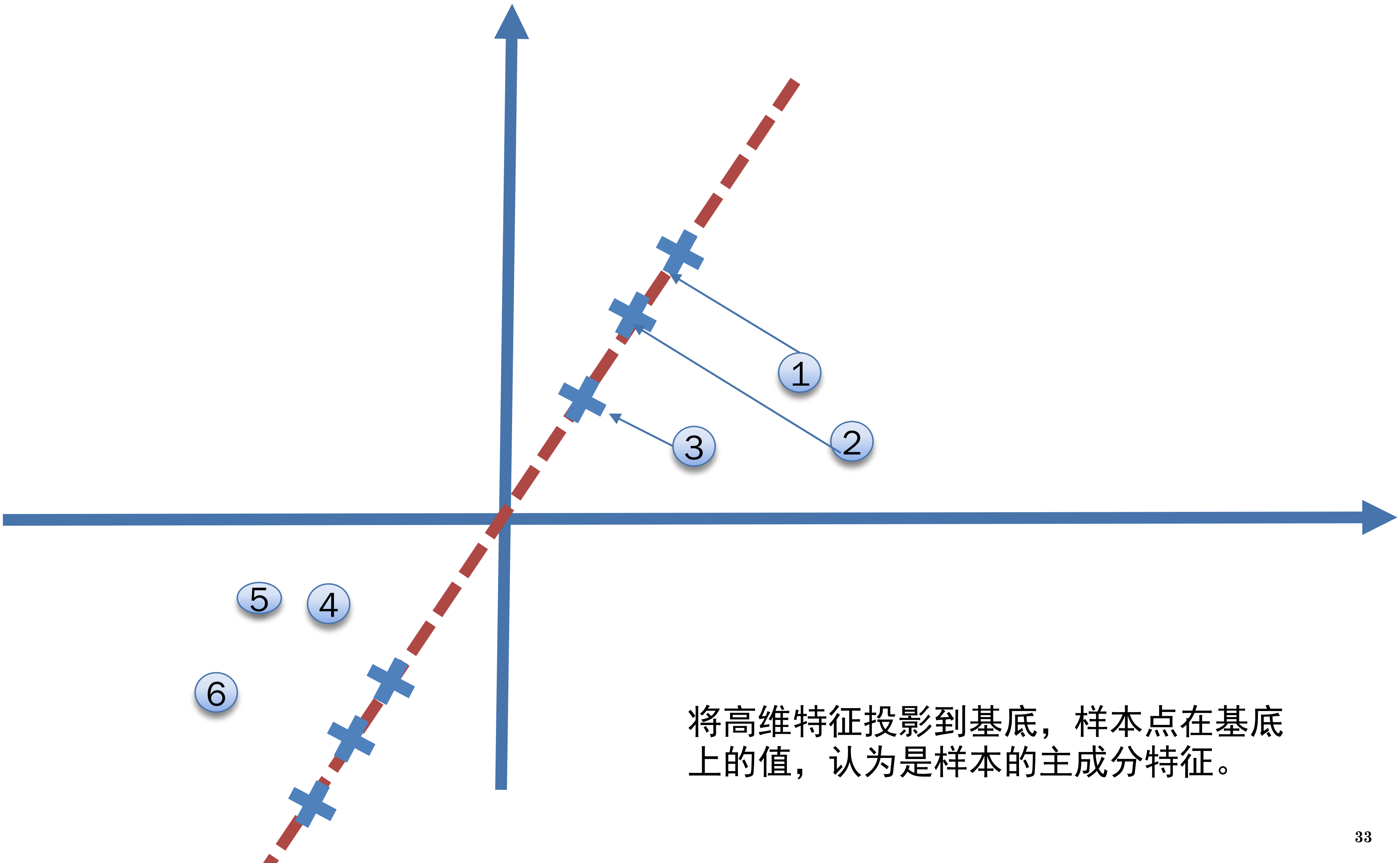
主成分分解(PCA)

假设某过原点的直线（坐标基底），过原点垂直于该直线的平面，将样本划分为高分段和低分段



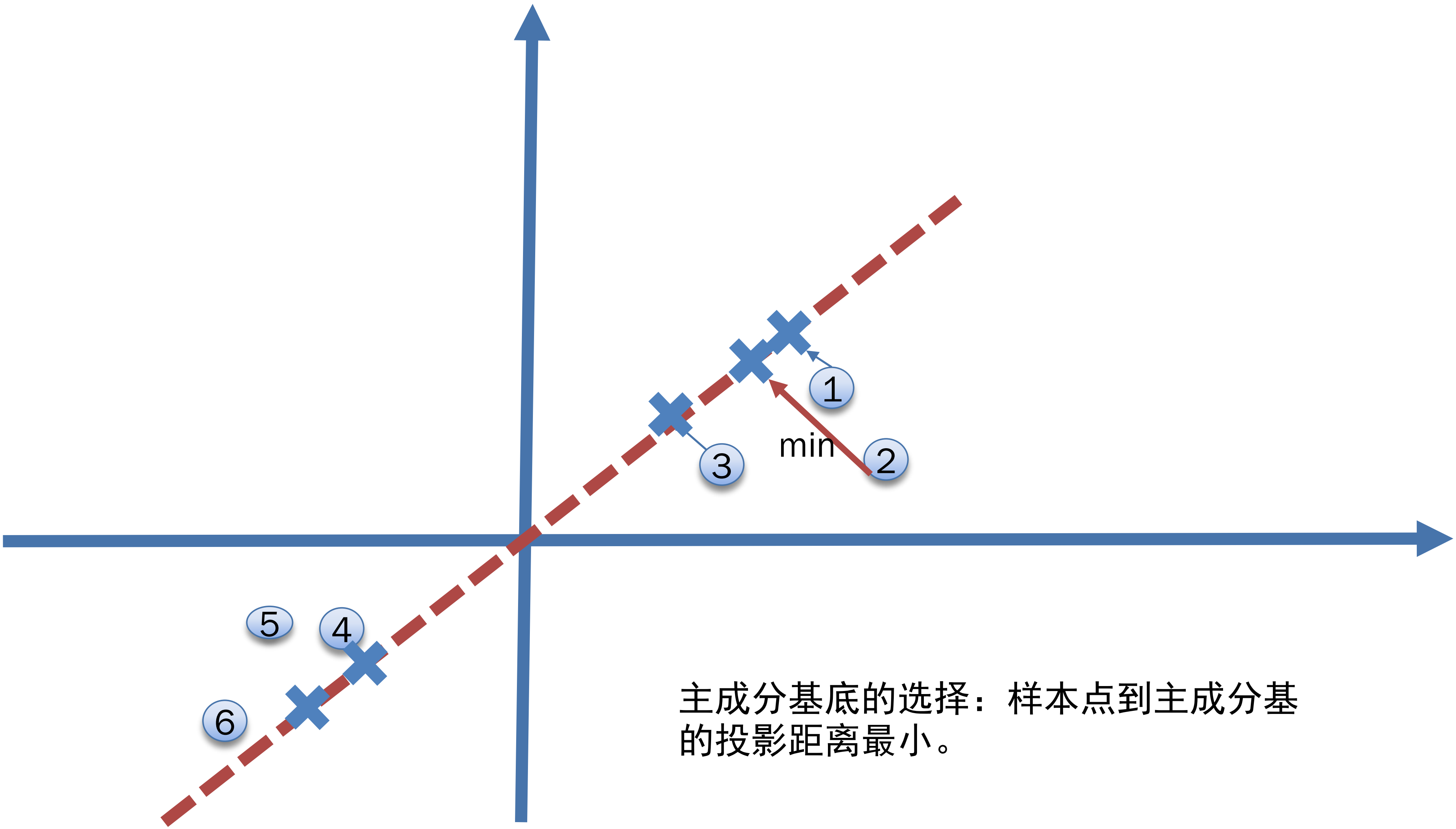
如何选取次划分直线？

主成分分解(PCA)



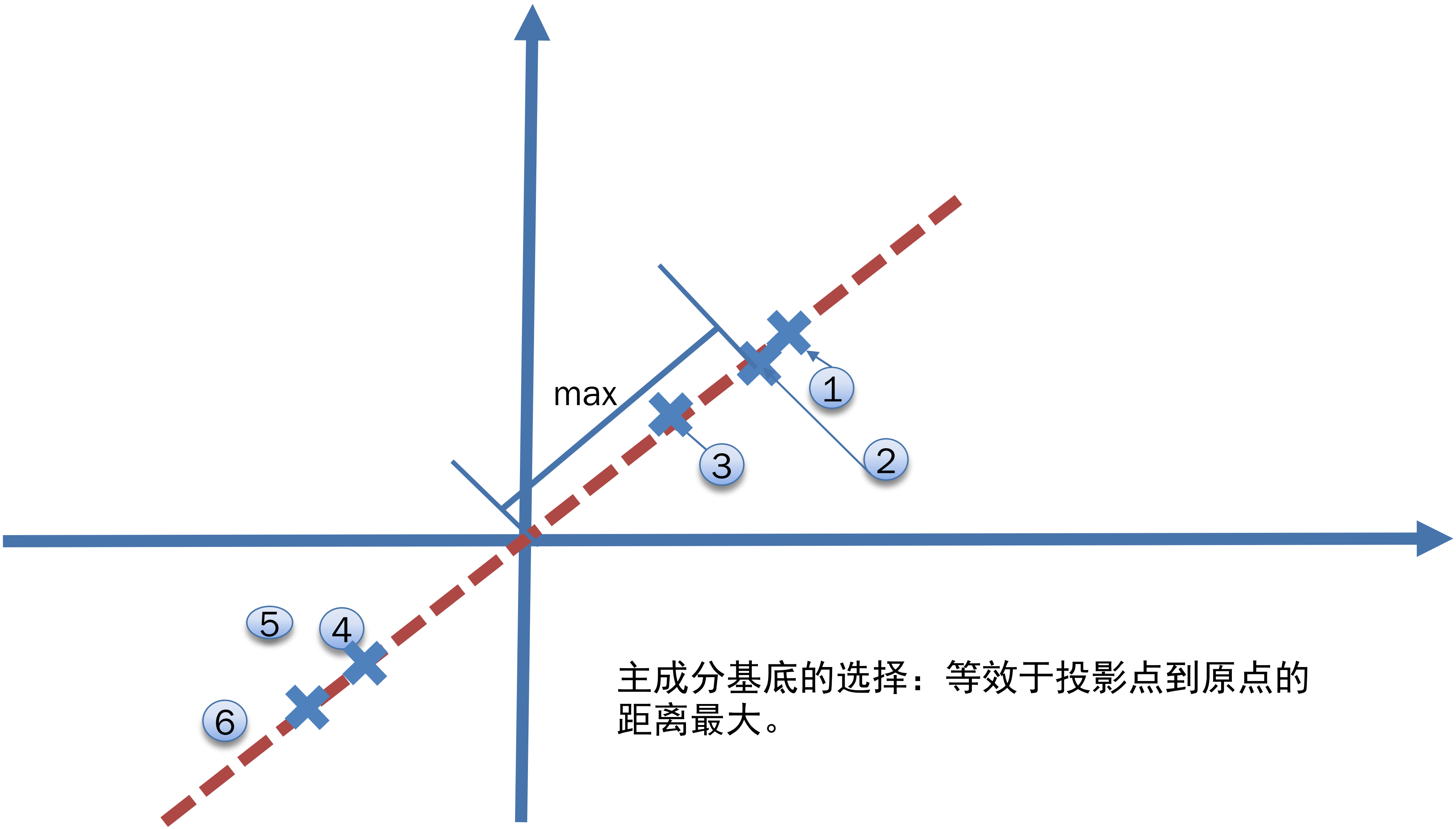
将高维特征投影到基底，样本点在基底上的值，认为是样本的主成分特征。

主成分分解(PCA)



主成分基底的选择：样本点到主成分基的投影距离最小。

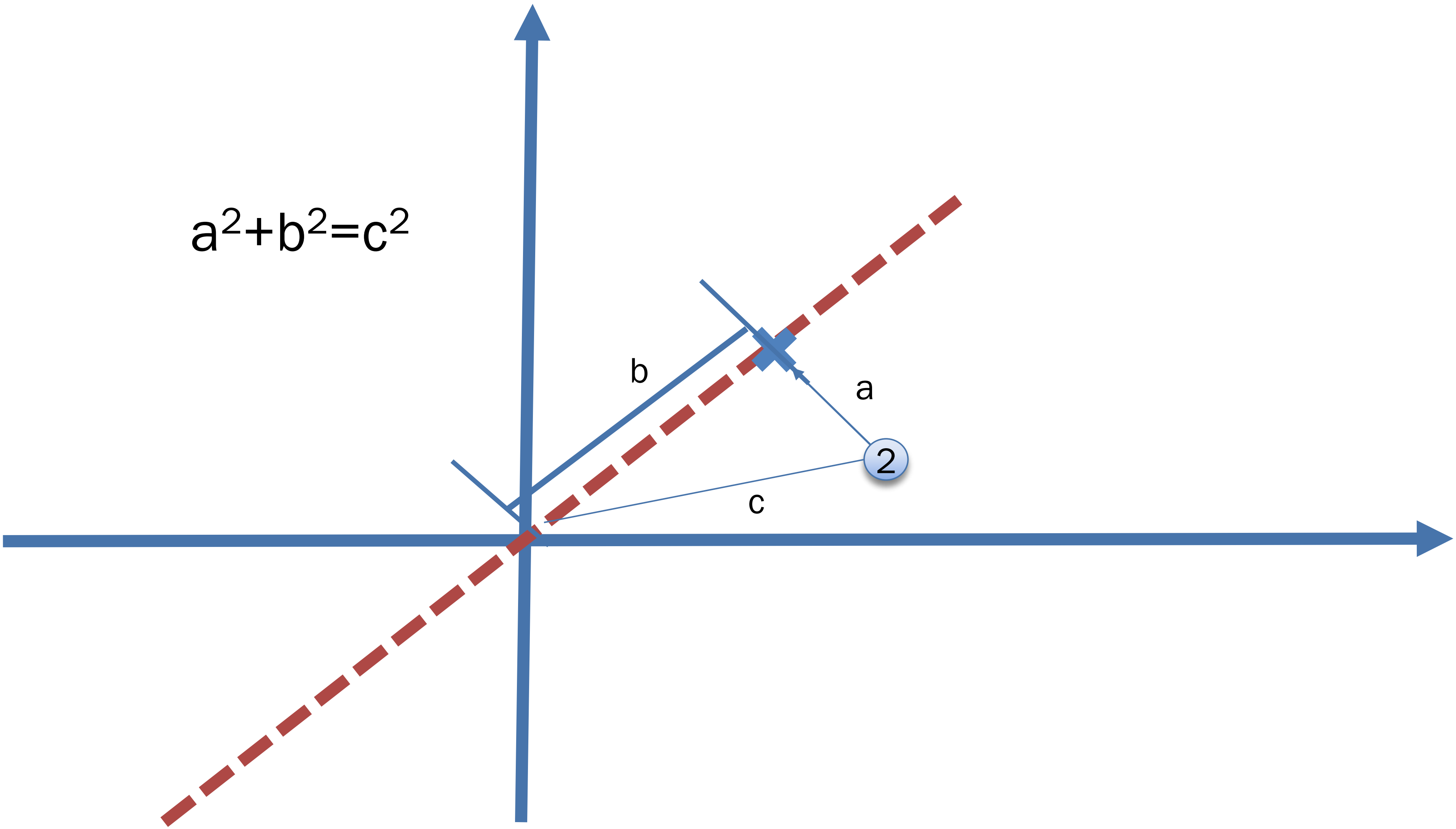
主成分分解(PCA)



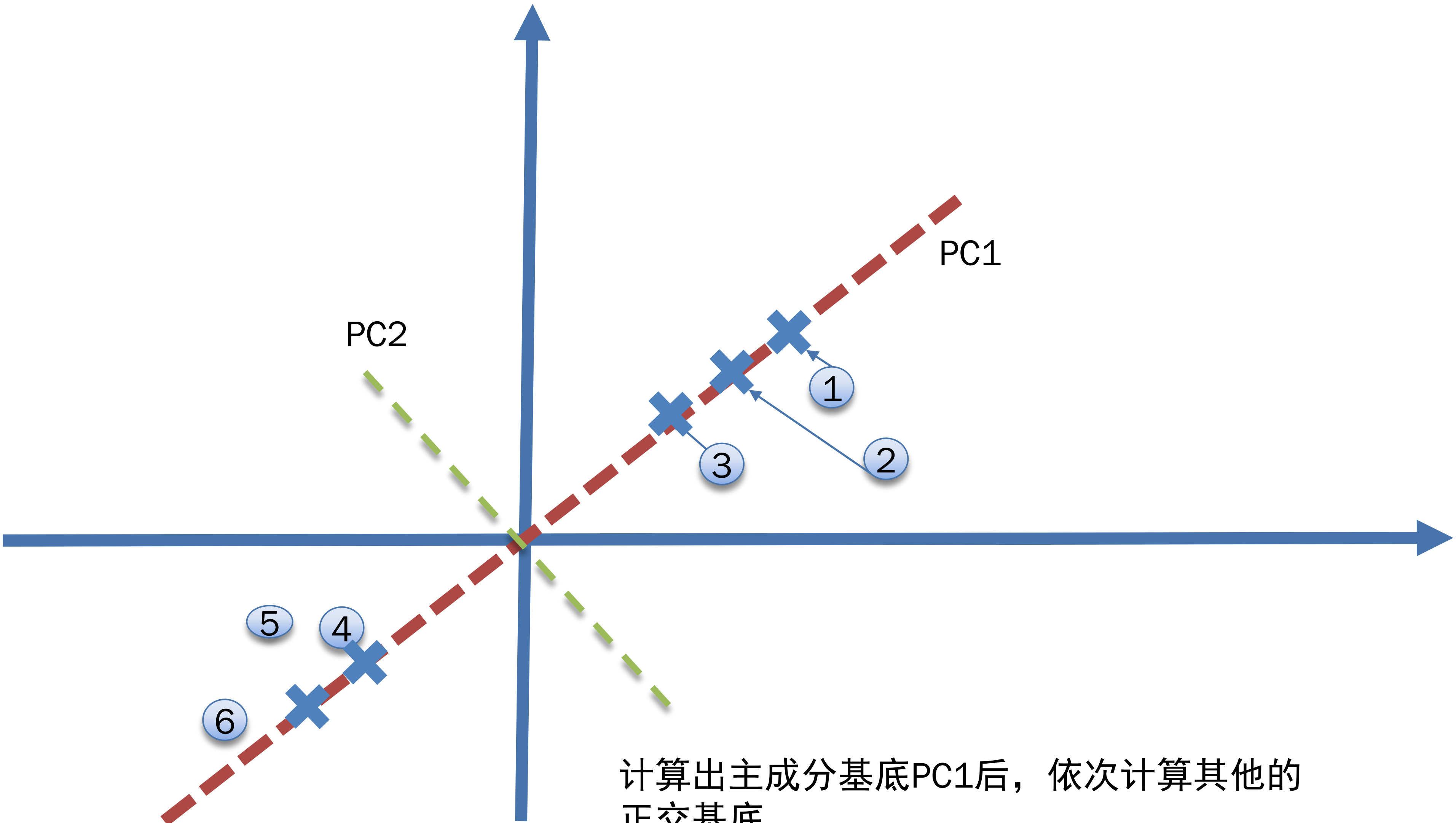
主成分基底的选择：等效于投影点到原点的距离最大。

主成分分解(PCA)

$a^2+b^2=c^2$

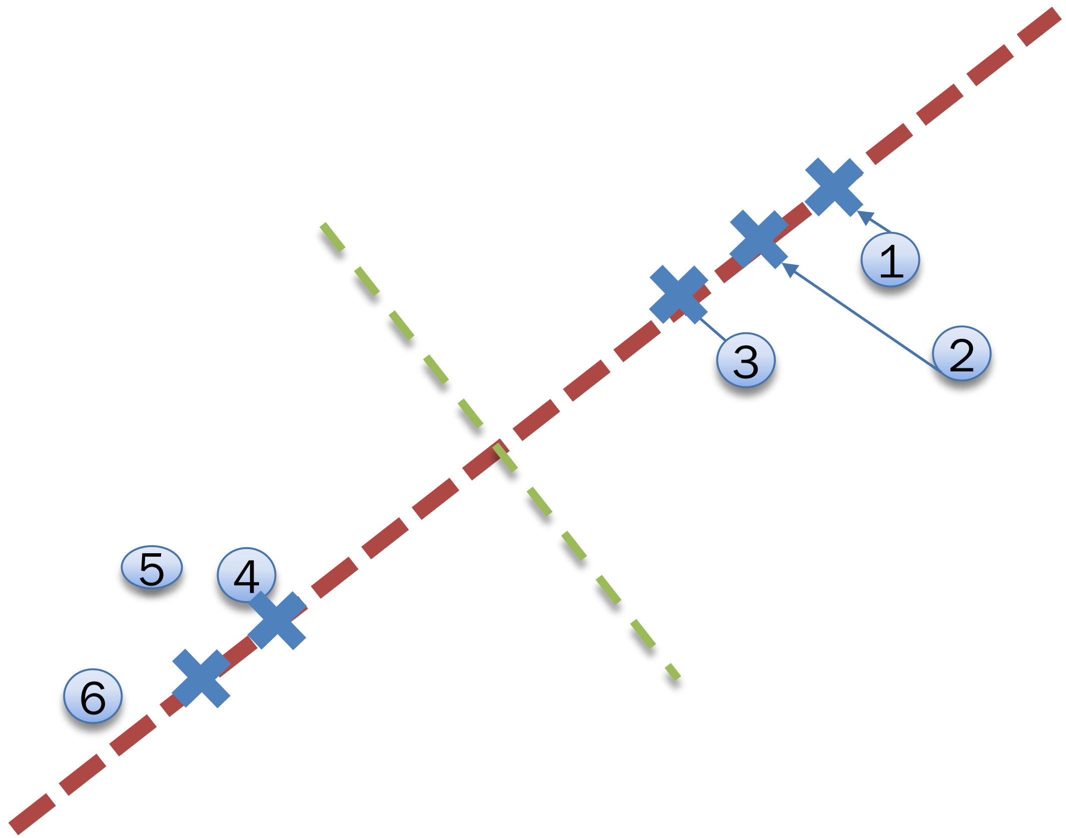


主成分分解(PCA)

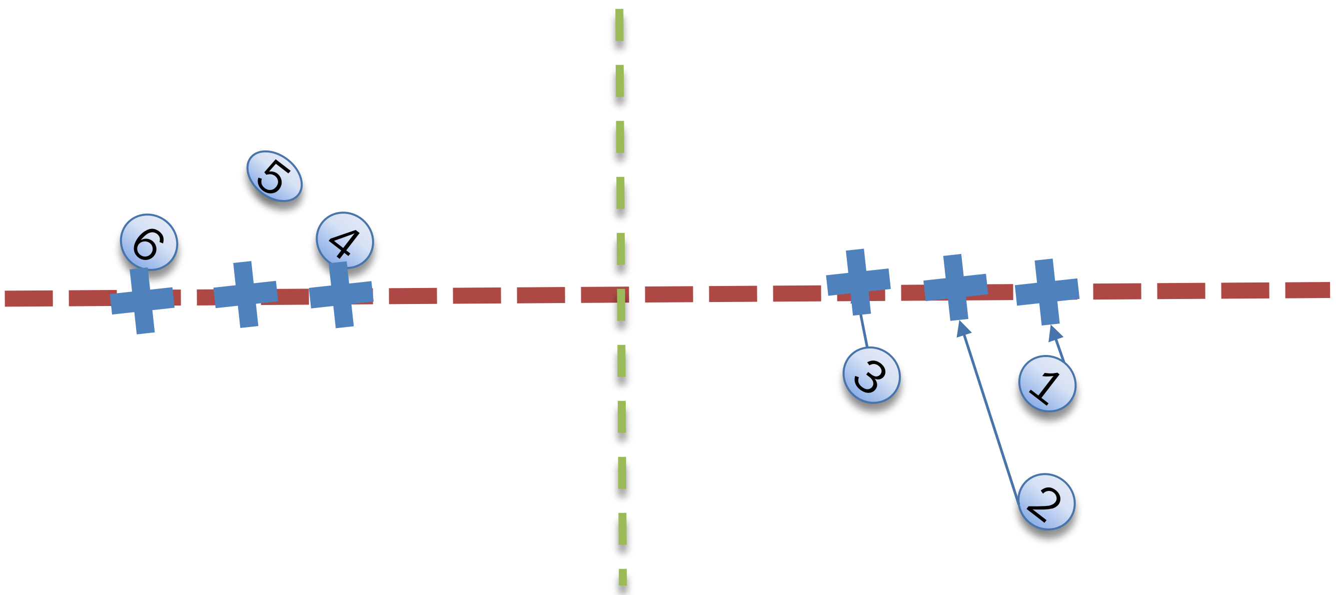


计算出主成分基底PC1后，依次计算其他的正交基底。

主成分分解(PCA)



主成分分解(PCA)



主成分分解(PCA)

- 对于高维特征数据样本降维：
- 1). 每一个成分可以理解为线性无关的特征的线性组合。
- 2). 基于数据样本在各个成分上的差异大小排序，描述成分的重要性。
- 3). 依次寻找在某基底上差异最大的成分，差异第二大的成分…
- 4). 当需要对数据进行降维时，忽略样本在次要成分量上的值，保留样本在几个主成分基底上的值。

主成分分解(PCA)

- 如何获得各个成分基底?
- 主成分分解（PCA）算法的原理与假设:
  - 每个样本的每个特征（Feature）取值构成样本数据矩阵。
  - 样本数据矩阵的协方差矩阵的每个特征向量，即为数据的每个主成分
    - 对应特征值最大的特征向量的方向，即为数据差异最大的主成分基底
- 每个特征向量对应一个特征值;
- 特征向量反映了数据差异的方向;
- 特征值反映了数据在该主成分方向上的差异大小。

如何计算协方差矩阵以及特征值、特征向量？

主成分分解(PCA)

- **以一组三维数据为例：**
- **Step1: 通过计算每个维度的均值，将每个样本在各个维度上的值，转换为相对质心的相对值**

原始数据

X	Y	Z
2.5	9.4	6.5
0.2	5.6	4.2
6	3.2	0.3
4.2	3.9	6.1
2.3	5	5.2
11	7	0.56
2.6	0.3	0.9
3.4	0.02	1.81
3.3	6.5	2.13
8	2	4.2

Mean(X) = 4.35
Mean (Y) = 4.292
Mean (Z) = 3.19

For example,
X11 = 2.5
Update:
X11 = 2.5-4.35 = -1.85

变换数据

X	Y	Z
-1.85	5.108	3.31
-4.15	1.308	1.01
1.65	-1.092	-2.89
-0.15	-0.392	2.91
-2.05	0.708	2.01
6.65	2.708	-2.63
-1.75	-3.992	-2.29
-0.95	-4.272	-1.38
-1.05	2.208	-1.06
3.65	-2.292	1.01

主成分分解(PCA)

- Step2: 使用变换后的矩阵，计算三个维度的协方差矩阵。

X	Y	Z
-1.85	5.108	3.31
-4.15	1.308	1.01
1.65	-1.092	-2.89
-0.15	-0.392	2.91
-2.05	0.708	2.01
6.65	2.708	-2.63
-1.75	-3.992	-2.29
-0.95	-4.272	-1.38
-1.05	2.208	-1.06
3.65	-2.292	1.01

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix}$$

```
>> C=cov(M)

C =

   10.0228    0.0329   -3.0013
    0.0329    9.0331    2.7696
   -3.0013    2.7696    5.4495
```

Matlab 函数: cov(Matrix)

主成分分解(PCA)

- **Step3: 计算特征值和特征向量**

```
>> C=cov(M)

C =

   10.0228    0.0329   -3.0013
    0.0329    9.0331    2.7696
   -3.0013    2.7696    5.4495
```

```
>> [V,D]=eig(C)

V =

    0.3603   -0.5838    0.7275
   -0.3863   -0.8033   -0.4533
    0.8491   -0.1177   -0.5150


D =

    2.9155         0         0
         0    9.4629         0
         0         0   12.1270
```

Matlab function:
[V,D] = eig(Covariance Matrix)
V = eigenvectors (each column)
D = eigenvalues (in the diagonal line)

此例子中，特征值2.9155 对应的特征向量即为 <0.3603, -0.3863, 0.8491>
每一个特征向量即可认为一个主成分基底的方向。

主成分分解(PCA)

- **Step4: 将特征值从大到小排序，对应的特征向量也相应重排序。我们可以选择其中最重要的K个成分，以K+2为例。**

```
>> [V,D]=eig(C)

V =

    0.3603   -0.5838    0.7275
   -0.3863   -0.8033   -0.4533
    0.8491   -0.1177   -0.5150


D =

    2.9155         0         0
         0    9.4629         0
         0         0   12.1270
```

K=2，即将原3维数据，降维为2维。本例中，选取特征向量矩阵中的后两列。

```
VM =

   -0.5838    0.7275
   -0.8033   -0.4533
   -0.1177   -0.5150
```

主成分分解(PCA)

- Step5: 将原始数据，投影到降维后的特征向量上，获得新的数据矩阵。

Original data, D, 10x3

X	Y	Z
2.5	9.4	6.5
0.2	5.6	4.2
6	3.2	0.3
4.2	3.9	6.1
2.3	5	5.2
11	7	0.56
2.6	0.3	0.9
3.4	0.02	1.81
3.3	6.5	2.13
8	2	4.2

Transformed data, TD, 10x3

X	Y	Z
-1.85	5.108	3.31
-4.15	1.308	1.01
1.65	-1.092	-2.89
-0.15	-0.392	2.91
-2.05	0.708	2.01
6.65	2.708	-2.63
-1.75	-3.992	-2.29
-0.95	-4.272	-1.38
-1.05	2.208	-1.06
3.65	-2.292	1.01

EigenMatrix, EM, 3X2

VM =

-0.5838	0.7275
-0.8033	-0.4533
-0.1177	-0.5150

FinalData (10xk) = TD (10x3) x EM (3xk), here k = 2

主成分分解(PCA)

- Step5: 将原始数据，投影到降维后的特征向量上，获得新的数据矩阵。

Original data, D, 10x3

X	Y	Z
2.5	9.4	6.5
0.2	5.6	4.2
6	3.2	0.3
4.2	3.9	6.1
2.3	5	5.2
11	7	0.56
2.6	0.3	0.9
3.4	0.02	1.81
3.3	6.5	2.13
8	2	4.2

After PCA

Final Data, FD, 10x2

FD =

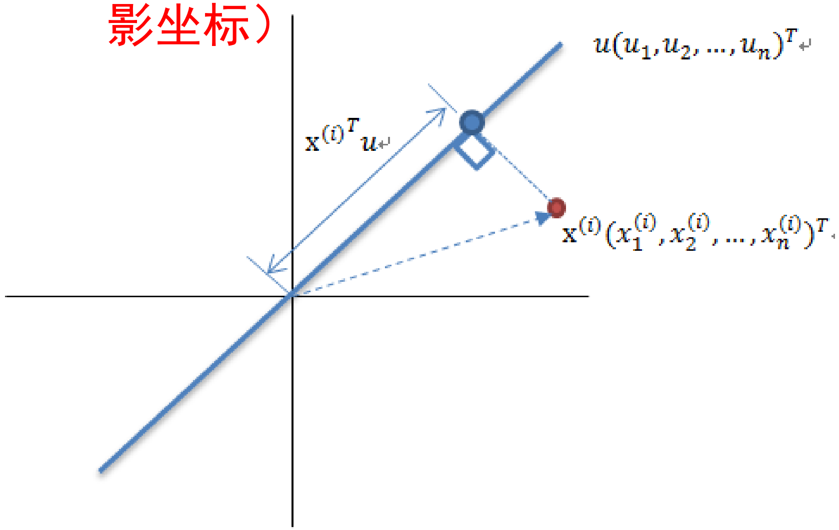
-3.4128	-5.3660
1.2532	-4.1322
0.2541	3.1837
0.0600	-1.4301
0.3915	-2.8475
-5.7481	4.9648
4.4980	1.7158
4.1487	1.9561
-1.0359	-1.2189
-0.4086	3.1742

FinalData (10xk) = TD (10x3) x EM (3xk), here k = 2

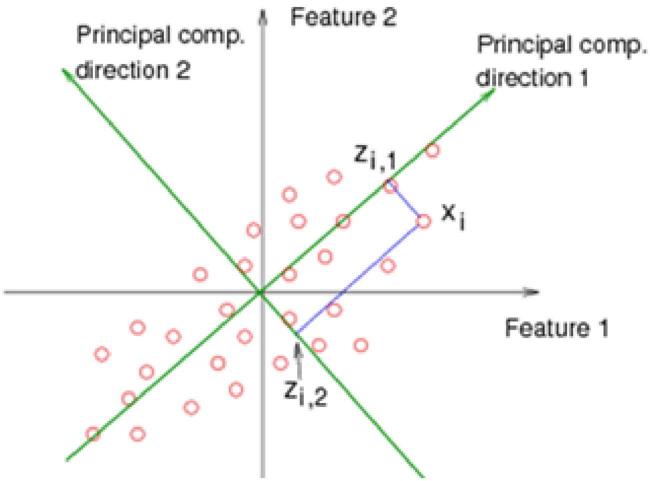
主成分分解(PCA)

- **Step5: 将原始数据，投影到降维后的特征向量上，获得新的数据矩阵。**

投影计算（每个高维样本的3维坐标与特征向量的点积即为样本在主成分基底上的投影坐标）



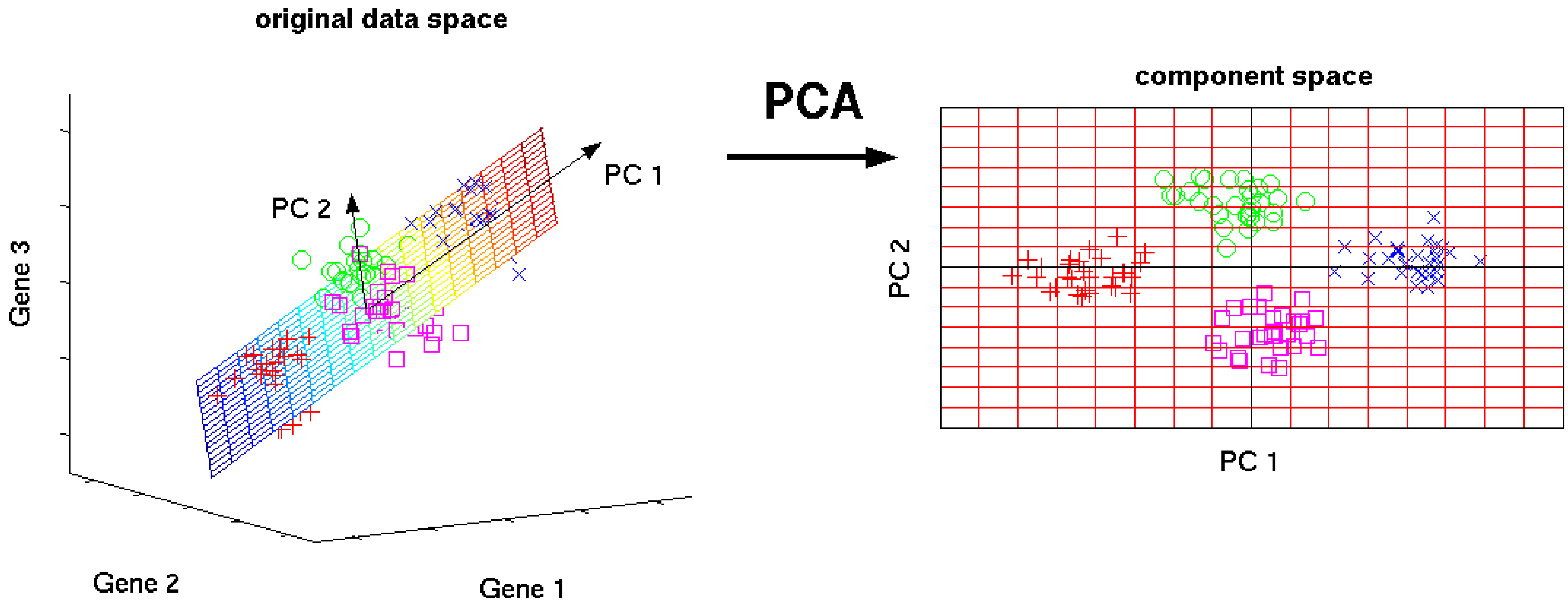
PCA 可视化



PCA实现高维数据到低维空间的线性投影

主成分分解(PCA)

- PCA 通过降维，可以方便高维数据可视化，同时尽可能的保留了原始数据中的差异信息。



基于矩阵分解的方法

- **主成分分解 PCA**
- **奇异值分解 SVD**
- **低秩矩阵分解 low rank MF**

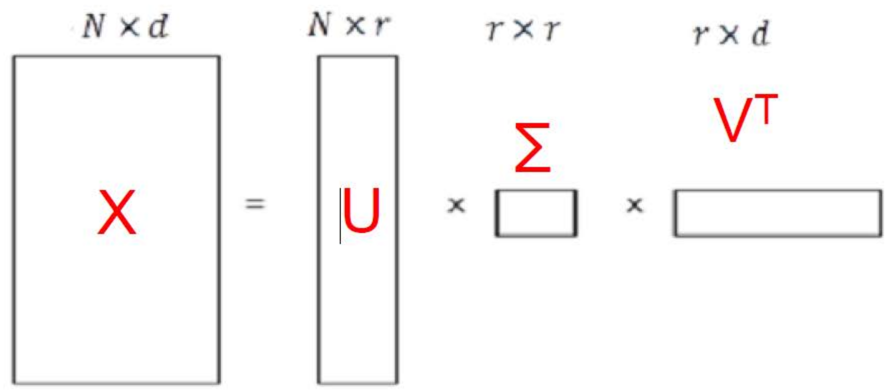
奇异值分解
——Singular Value Decomposition (SVD)

- 奇异值分解也是一种数据降维的方法
- 主要用于矩阵分解
- 奇异值分解也属于一种PCA
- 成名于2009年Netflix的推荐系统大赛 Netflix Prize Contest

奇异值分解(SVD)

- **任意一个 N × d 矩阵 X 可以按如下形式分解:**

$$\mathbf{X} = \mathbf{U} \times \Sigma \times \mathbf{V}^{\mathbf{T}}$$

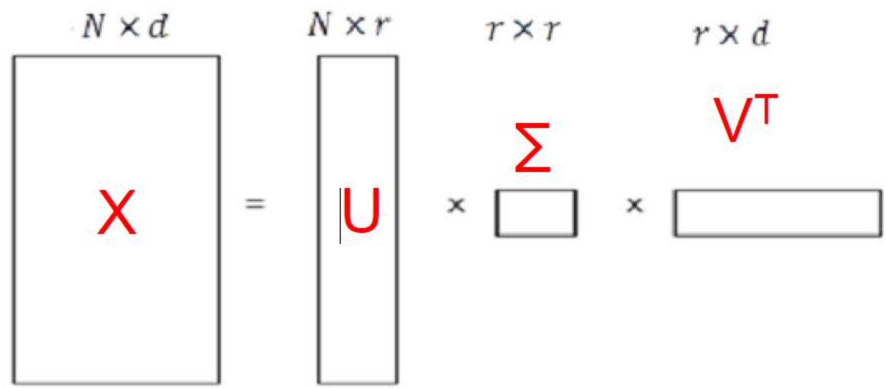


- **r 为矩阵X的秩，即矩阵X最大的线性无关的行数或列数**
- **U 为一个列正交的N × r 的矩阵;**
- **V 为一个列正交的 d × r的矩阵;**
- **Σ 为一个r × r的对角阵，其中对角线元素为矩阵X的奇异值，按降序排列。**

奇异值分解(SVD)

- **SVD与PCA的关系**

$$\mathbf{X} = \mathbf{U} \times \mathbf{\Sigma} \times \mathbf{V}^{\mathbf{T}}$$



$$C = \frac{1}{N-1} X^T X$$ 为 $d \times d$ 的协方差矩阵

C的特征向量等于X的右奇异向量，即V^T中包含了C 的特征向量。且$\sum$ 中的奇异值与C中的特征值的关系为：

$$\lambda_i = \frac{s_i^2}{N-1}$$

奇异值分解(SVD)

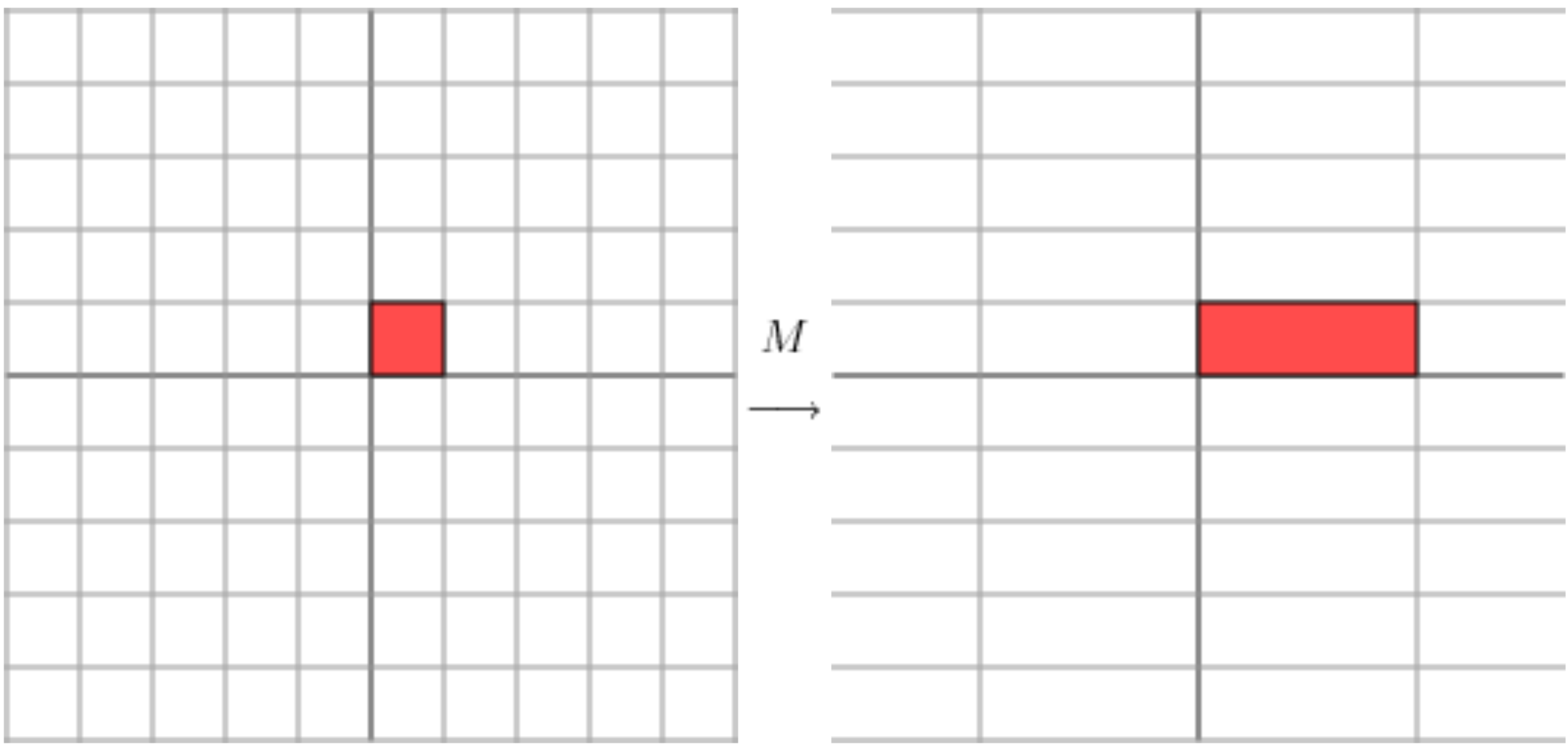
- **Matlab奇异值分解函数**

```
X = [1 0 1; -1 -2 0; 0 1 -1];
s = svd(X);%满秩矩阵的奇异值
X = [1 2; 3 4; 5 6; 7 8];
%[U,S,V] = svd(X) 奇异值分解
[U,S,V] = svd(X,'econ')
C = X'*X/3
[v,d] = eig(C)
```

奇异值分解的几何解释

矩阵乘法的几何含义

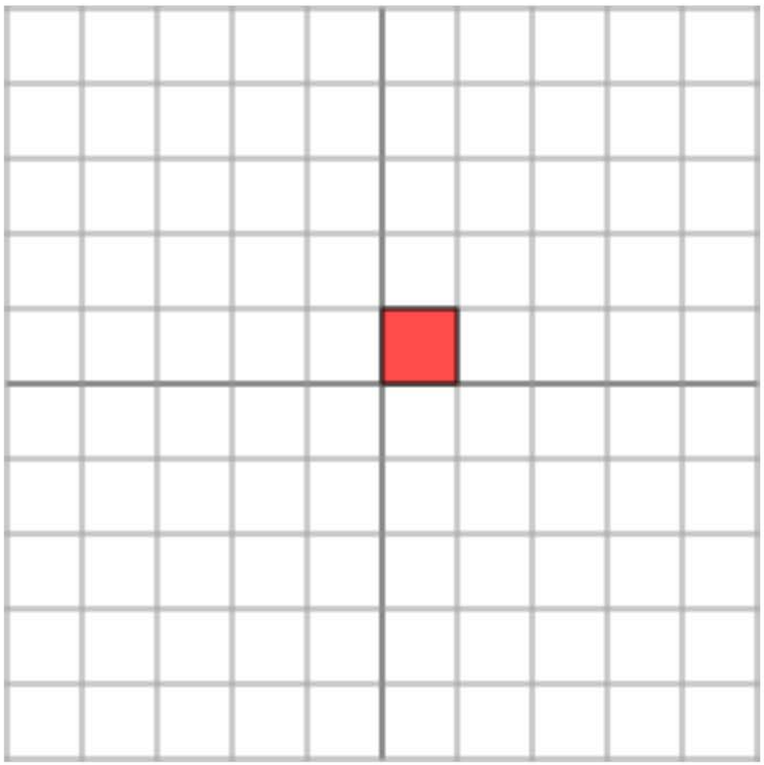
$$M = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ y \end{bmatrix}.$$



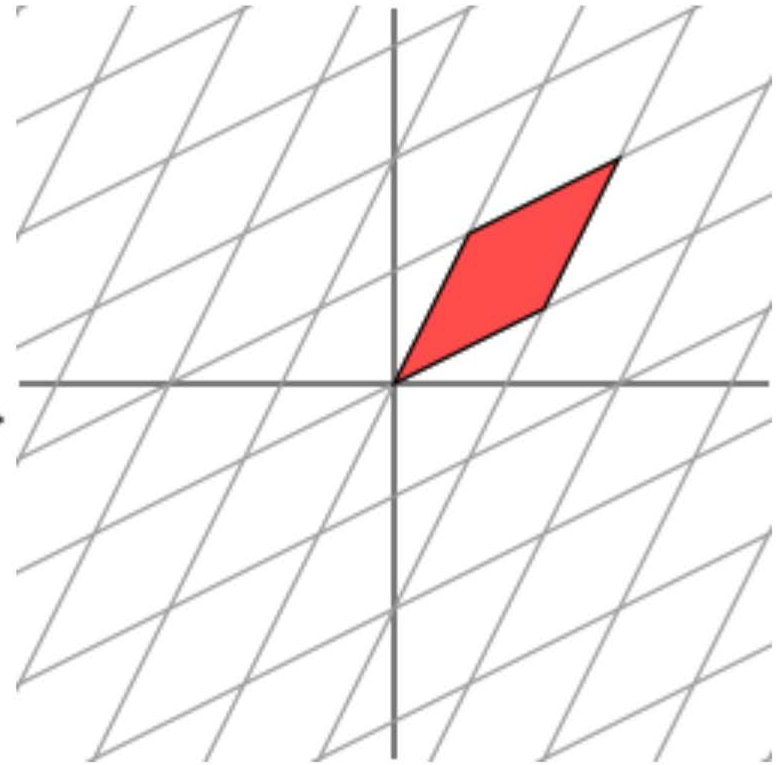
奇异值分解的几何解释

矩阵乘法的几何含义

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$



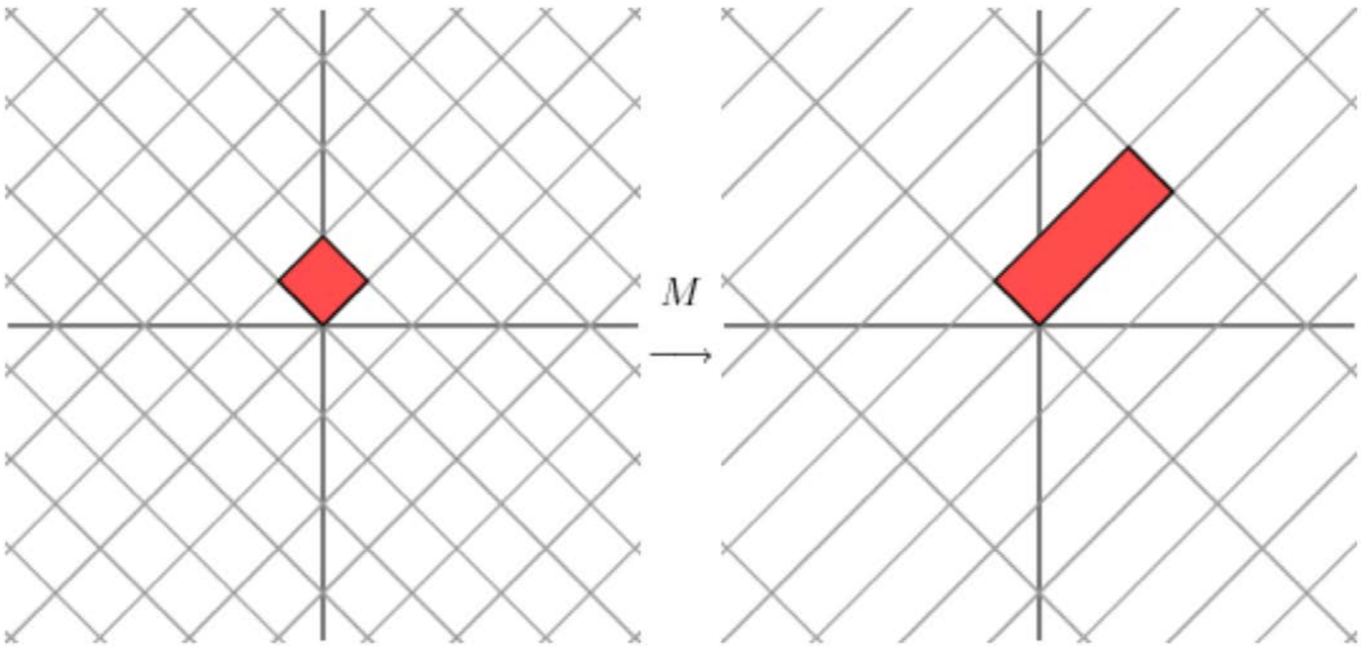
$\xrightarrow{M}$



矩阵左乘向量，相对于将空间中的向量旋转并拉伸

奇异值分解的几何解释

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$



如果空间中的向量表示为某一种特殊的基底的线性组合，该坐标基底在左乘M后，各基底向量方向保持不变（形态类似与对角矩阵相乘），只是在某些方向上，有伸缩。

奇异值分解的几何解释

- 给定一个**对称矩阵M**, 我们可以找到一组正交向量 $\mathbf{v_i}$ ， 使得 $M\mathbf{v_i}$ 等于$\mathbf{v_i}$与一个标量的乘积，即

$$Mv_i = \lambda_i v_i$$

其中 λ_i 是一个标量.

- 几何上，即表示向量 $\mathbf{v_i}$ 与 M 相乘等效于将向量拉伸缩放或180°翻转，但没有其他旋转。

- 正因为此特性, 向量 $\mathbf{v_i}$ 被称为矩阵 M 的特征向量，标量 λ_i 被称为特征值.

- 问题:

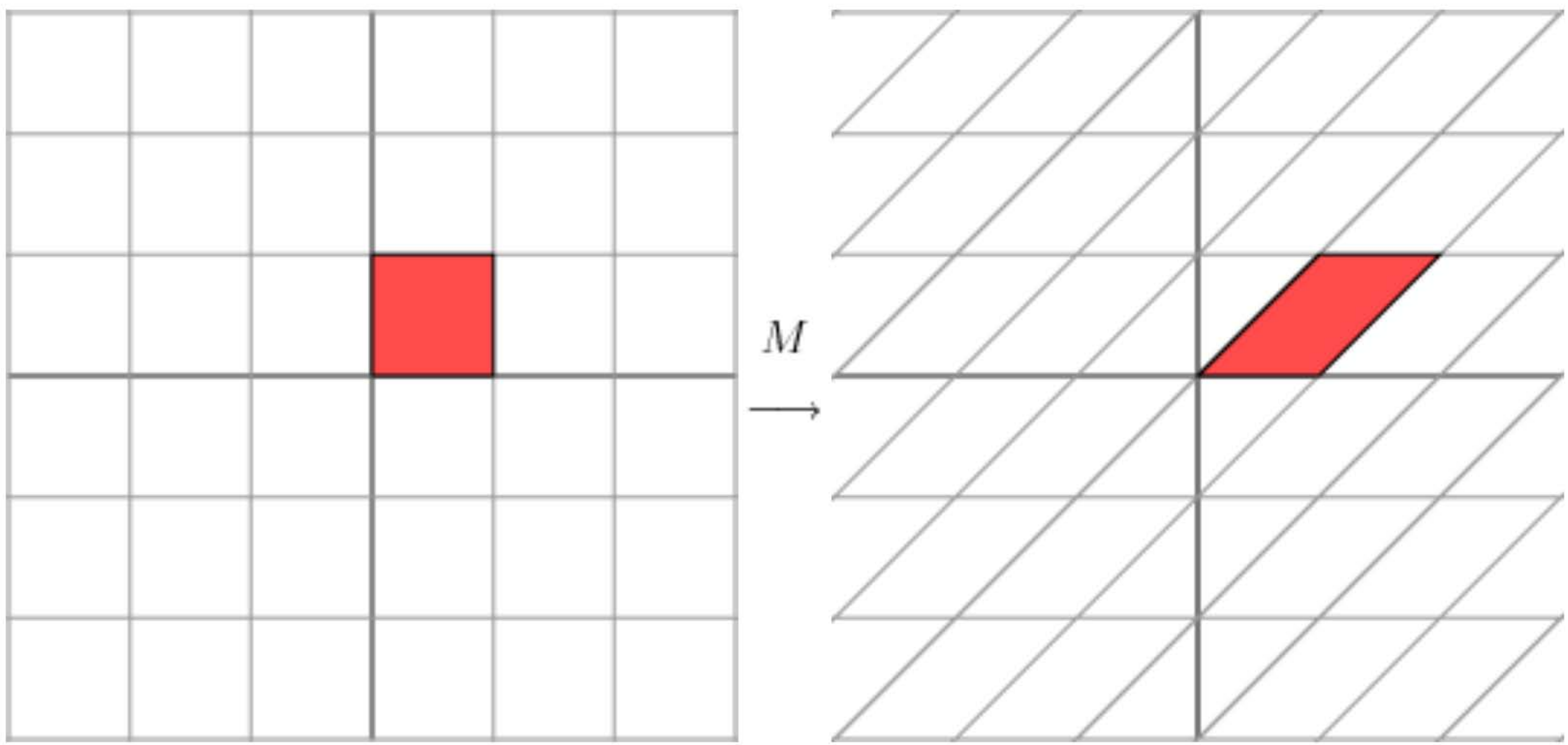
- 我们可以将一个正交的坐标基底，变换为另一个坐标基底么?

奇异值分解的几何解释

- 考虑一般的非对称矩阵M:

$$M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

- 乘以该矩阵会产生如下形变

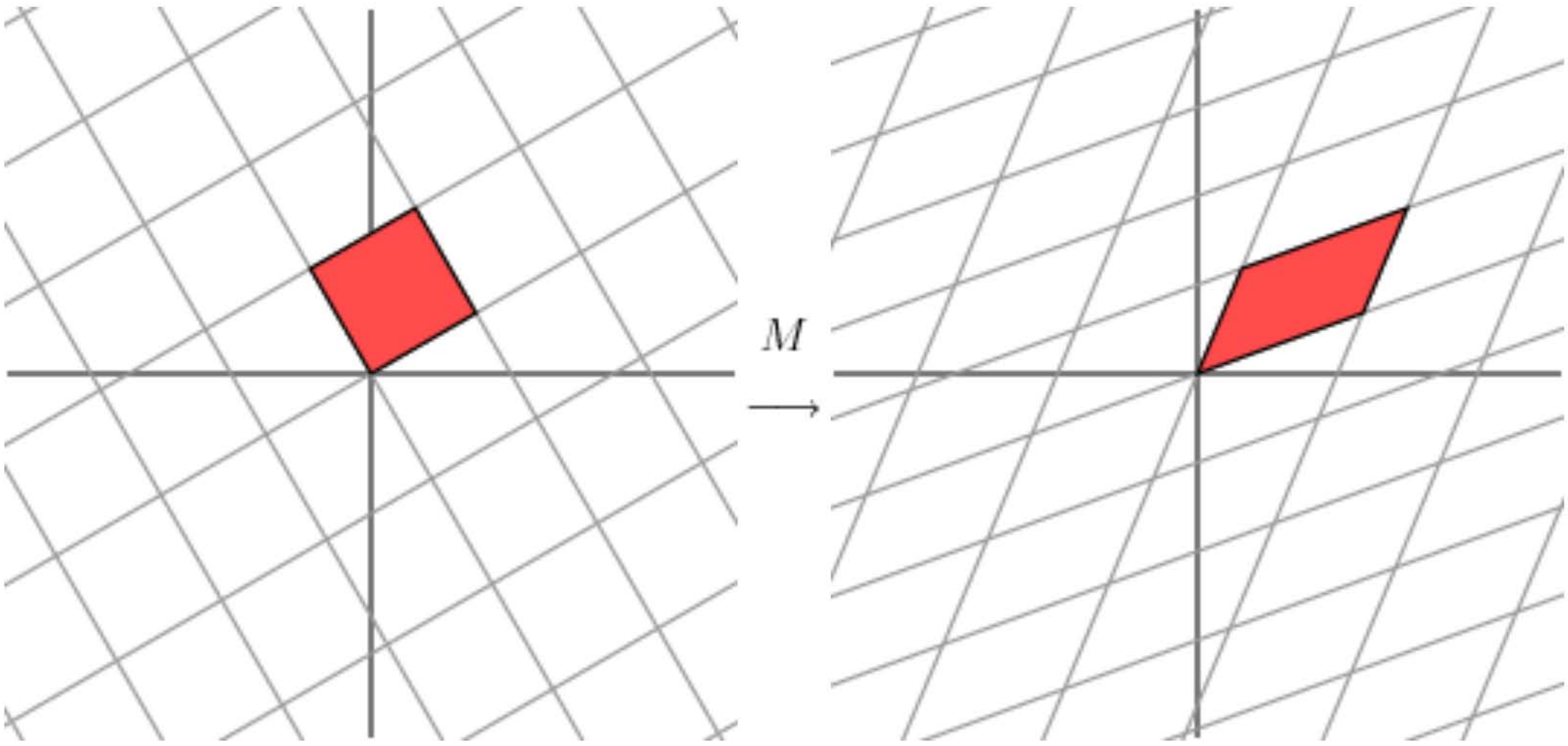


奇异值分解的几何解释

- 乘以某一般的非对称矩阵M，会该产生如下形变

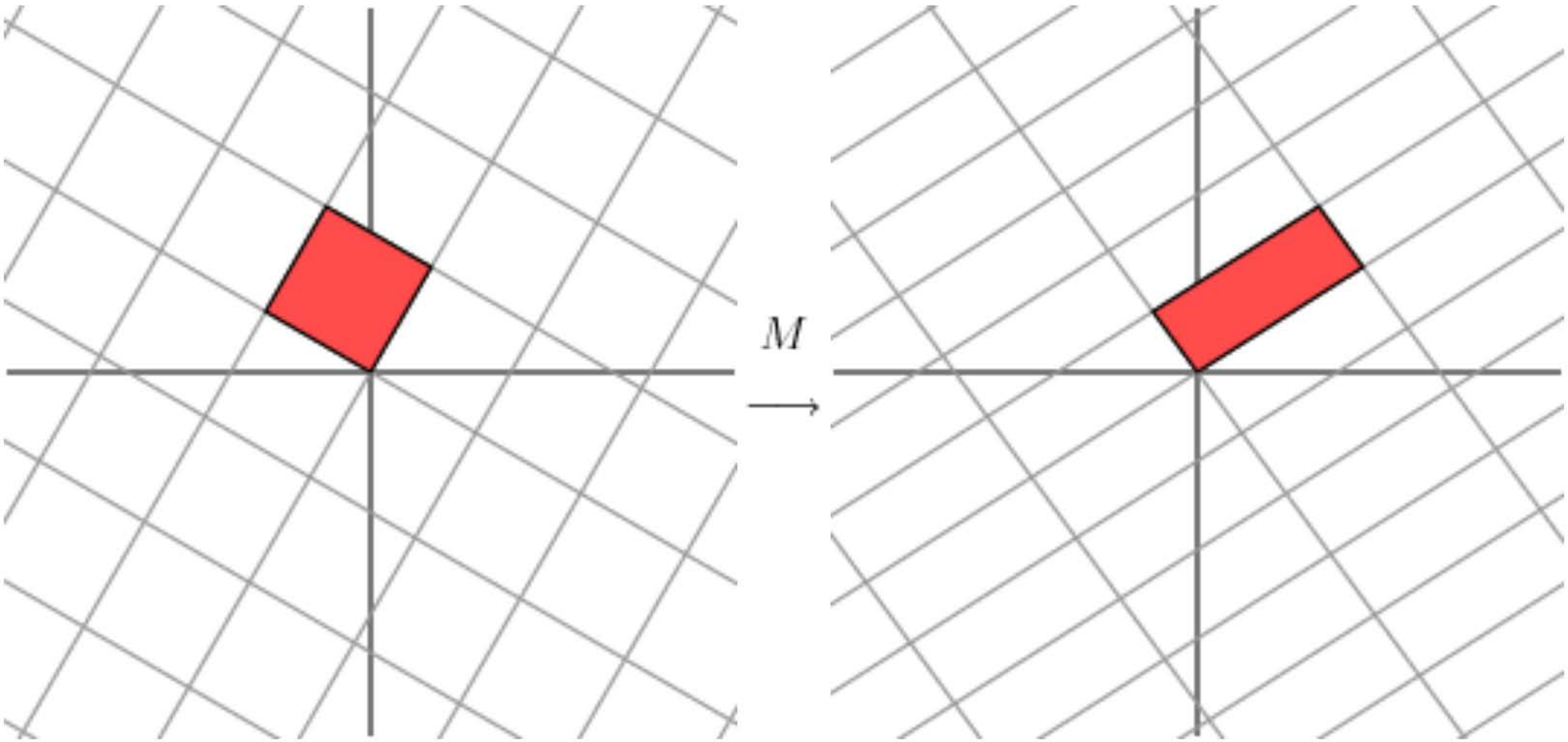
$$M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

- 如果旋转原空间中的坐标基底，变换空间中的方格会随之旋转和形变。



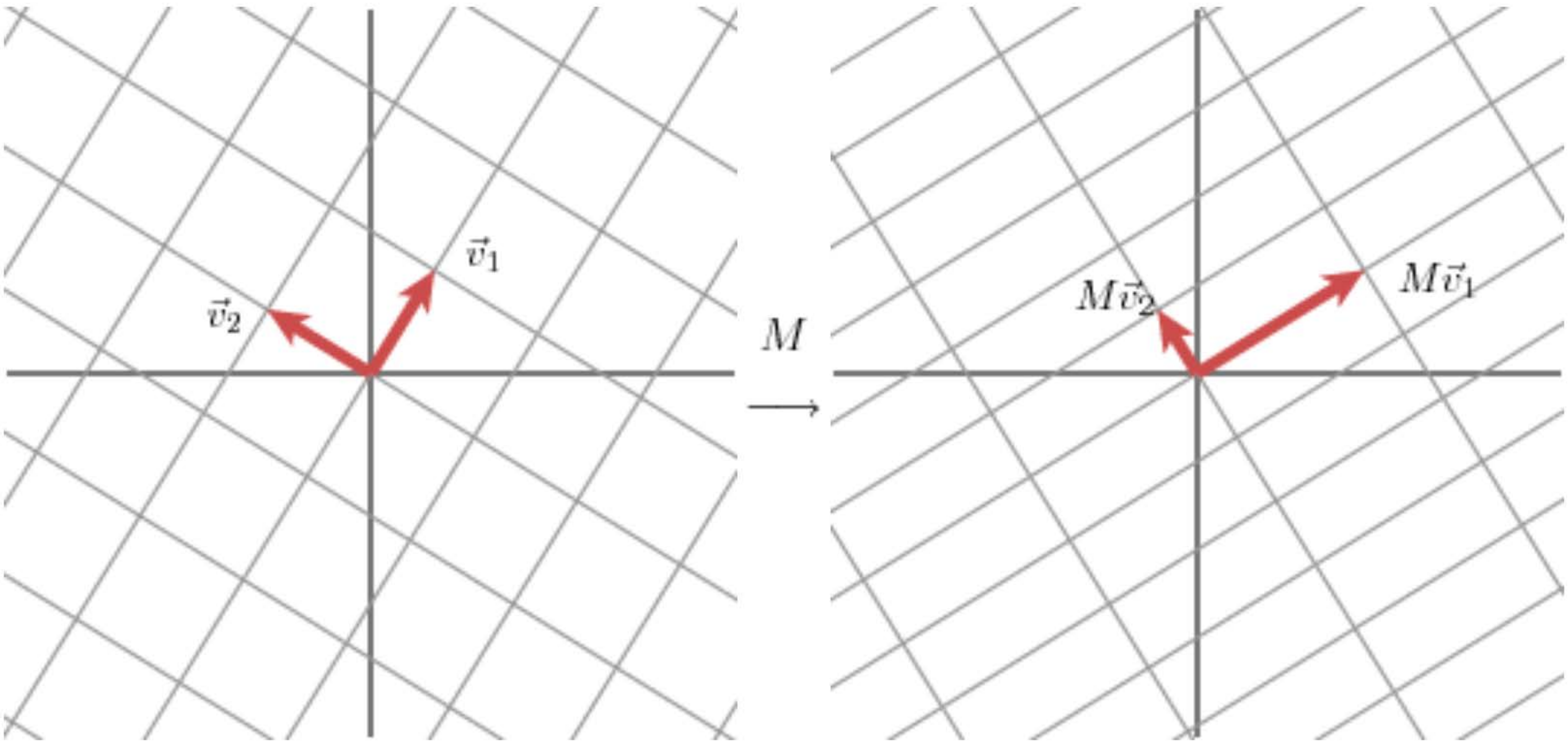
奇异值分解的几何解释

- 当旋转至某一特定角度时，如本例中$58.28°$，原空间中的正交基底单元格，变成了另一组正交向量构成的方格



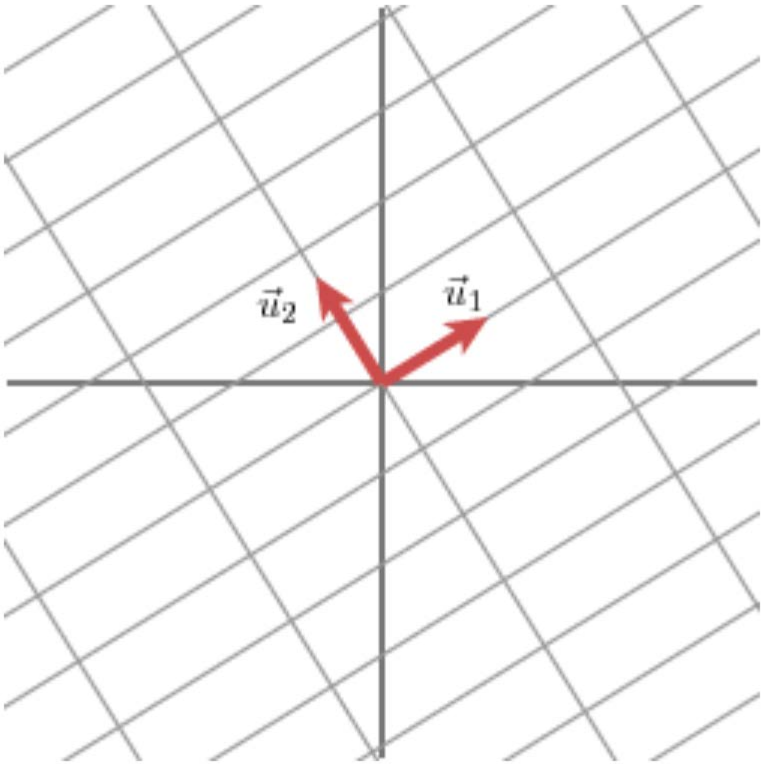
奇异值分解的几何解释

- 使用向量描述，即如果针对矩阵M 选择合适的正交基底，基底单位向量分别为 $\mathbf{v}_1$ 和 $\mathbf{v}_2$, 该单位向量与M相乘后得到向量 $M\mathbf{v}_1$ 和 $M\mathbf{v}_2$ 仍然保持正交。



奇异值分解的几何解释

- 我们分别用向量 $\mathbf{u_1}$ 和 $\mathbf{u_2}$ 表示 $M\mathbf{v_1}$ 和 $M\mathbf{v_2}$ 方向上的单位向量。
- 向量 $M\mathbf{v_1}$ 和 $M\mathbf{v_2}$的长度分别记作 σ_1 与 σ_2



- $M\mathbf{v_1} = \sigma_1\mathbf{u_1}$
  $M\mathbf{v_2} = \sigma_2\mathbf{u_2}$

奇异值分解的几何解释

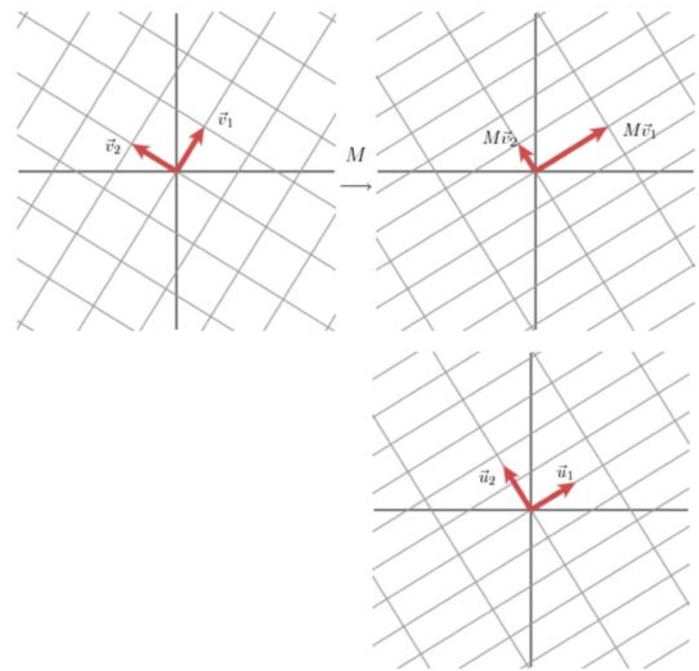
- 如此，我们可以对“矩阵M乘以一个向量”给与一种新的描述。
- 由于向量 $\mathbf{v}_1$ 和 $\mathbf{v}_2$ 为正交单位向量，因此任意向量x可以表示为：

$$x = (v_1 \cdot x)v_1 + (v_2 \cdot x)v_2$$

- 因此：

$$Mx = (v_1 \cdot x)Mv_1 + (v_2 \cdot x)Mv_2$$

$$Mx = (v_1 \cdot x)\sigma_1 u_1 + (v_2 \cdot x)\sigma_2 u_2$$



- 其中点积可以用向量转置相乘表示，即：

$$v \cdot x = v^T x$$

- 最终：

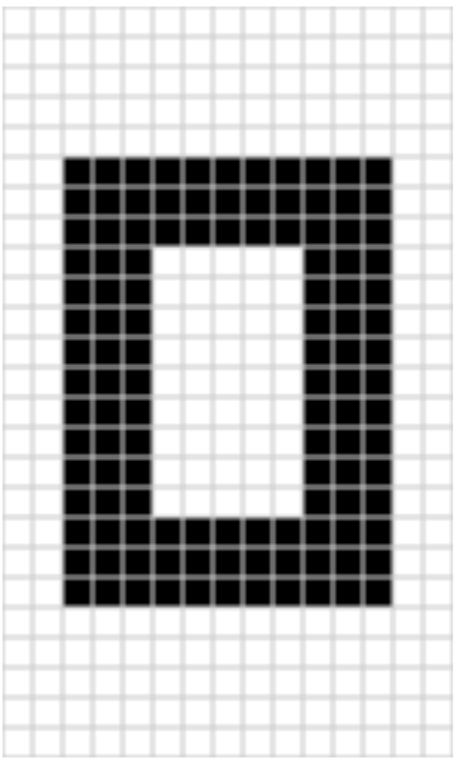
$$Mx = u_1\sigma_1 v_1^T x + u_2\sigma_2 v_2^T x$$

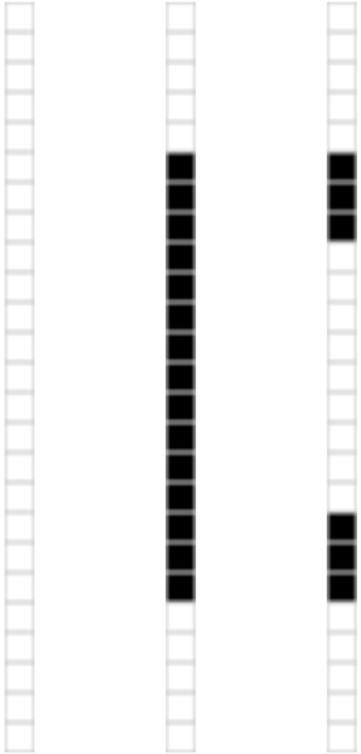
$$M = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T$$

- $M = U\Sigma V^T$

奇异值分解应用案例

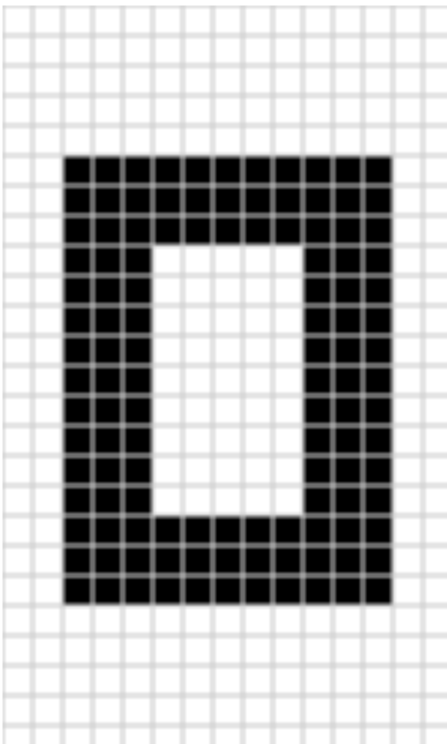
- **数据压缩**
- **一个15x25 黑白（01）像素阵列.**





奇异值分解应用

- **数据压缩**



$$M=\begin{bmatrix}
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1
\end{bmatrix}$$

奇异值分解应用

- **数据压缩**

$$
M=\begin{bmatrix}
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&1&1&1&1&1&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&0&0&0&0&0&0&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1
\end{bmatrix}
$$

- $\sigma_1 = 14.72$
- $\sigma_2 = 5.22$
- $\sigma_3 = 3.31$

$$M=\mathbf{u}_1\sigma_1\ \mathbf{v}_1^{\mathrm{T}}+\mathbf{u}_2\sigma_2\ \mathbf{v}_2^{\mathrm{T}}+\mathbf{u}_3\sigma_3\ \mathbf{v}_3^{\mathrm{T}}$$

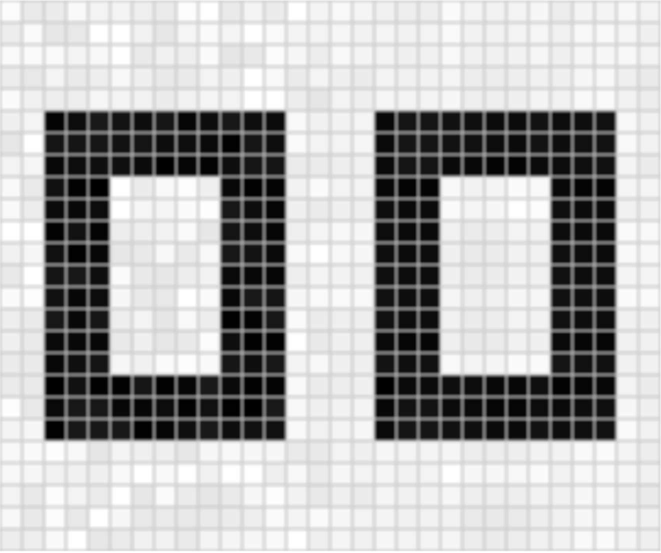
仅有三个奇异值不为0。对应三个奇异向量，意味着375个元素的矩阵，可以使用（25+15）*3+3 = 123个数值表示

奇异值分解应用

- **噪声消除**

- $\sigma_1 = 14.15$
  $\sigma_2 = 4.67$
  $\sigma_3 = 3.00$
  $\sigma_4 = 0.21$
  $\sigma_5 = 0.19$
  …
  $\sigma_{15} = 0.05$

$$M \approx \mathbf{u}_1\sigma_1\,\mathbf{v}_1^{\mathrm{T}} + \mathbf{u}_2\sigma_2\,\mathbf{v}_2^{\mathrm{T}} + \mathbf{u}_3\sigma_3\,\mathbf{v}_3^{\mathrm{T}}$$



基于矩阵分解的方法

- **主成分分解 PCA**
- **奇异值分解 SVD**
- **低秩矩阵分解 low rank MF**

以推荐系统为例

- 推荐系统（Recommender System， RS）
  - 为用户提供选项或建议

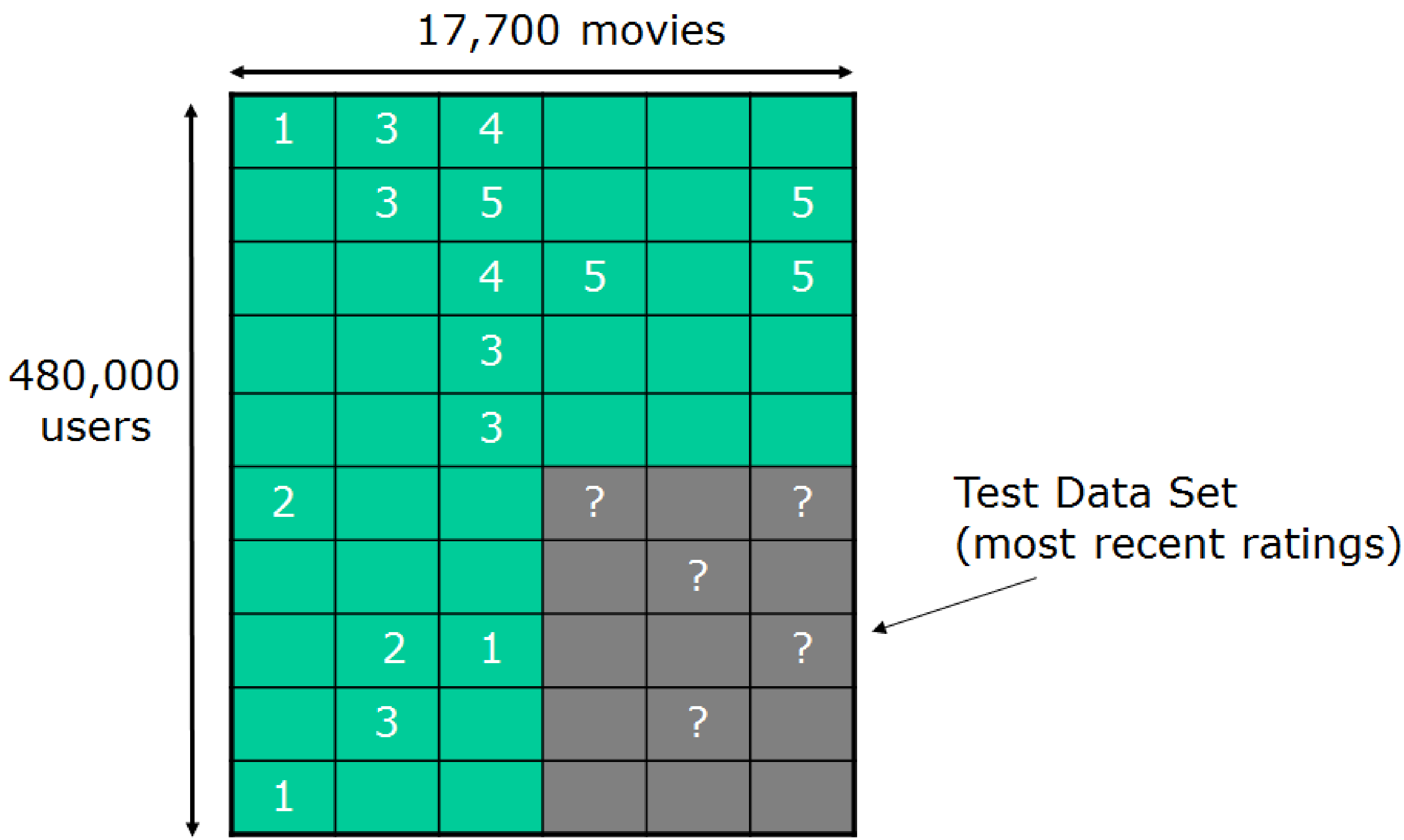
推荐系统的任务

- 1: 评分预测

User	Harry Potter	Batman	Spiderman
U1	5	3	4
U2	?	2	4
U3	4	2	?

- 2: Top-N 推荐
  - 清单推荐;
  - 例：热门新闻、热门微博

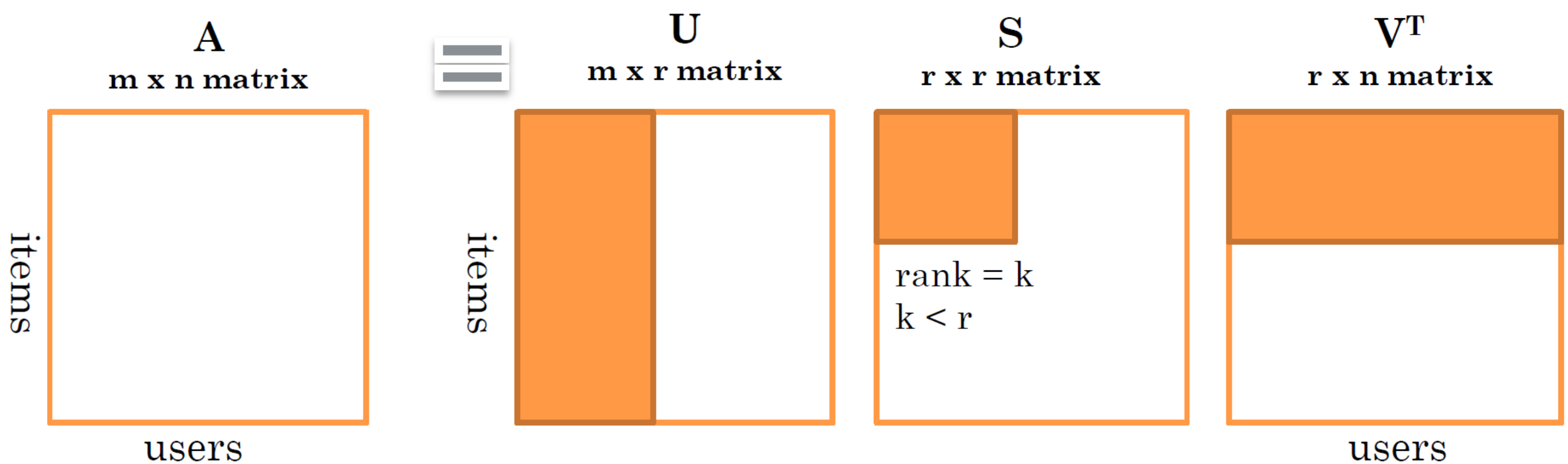
Netflix Contest



从 SVD 到 MF

- **SVD可以作为一种优化算法. 假设原数据矩阵 X, SVD 操作可以找到X的一种低秩（eg. rank=k）近似矩阵，使得近似矩阵D与原矩阵X之差的矩阵范数最小。min $||X - D||_F$**

$$A = U \times S \times V^T$$



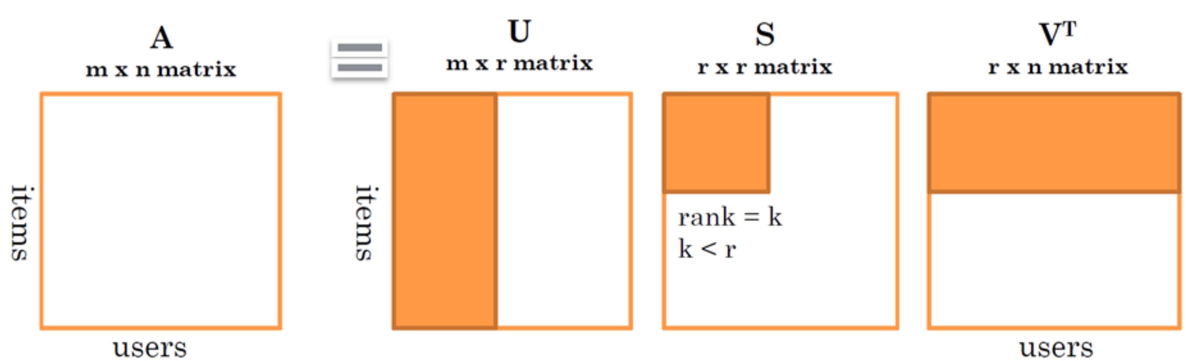
$$A_k = U_k \times S_k \times V_k^T$$

推荐系统中的矩阵分解

- 任务: 评分预测

User	Harry Potter	Batman	Spiderman
U1	5	3	4
U2	?	2	4
U3	4	2	?

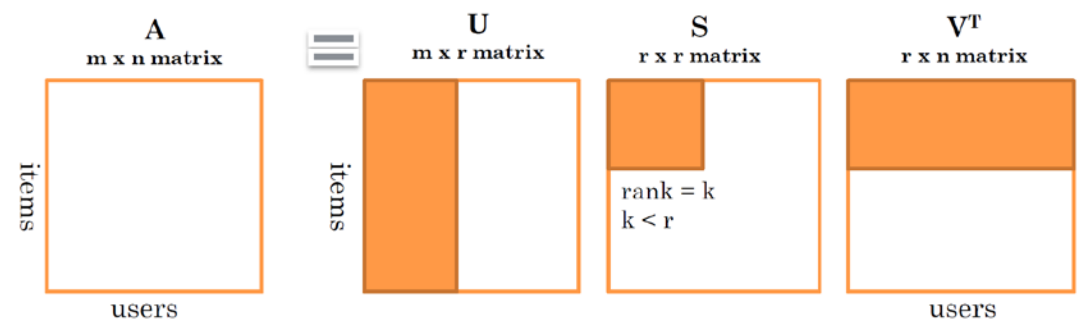
$$A = U \times S \times V^T$$



$$A_k = U_k \times S_k \times V_k^T$$

从SVD 到 MF

$$A = U \times S \times V^T$$



$$A_k = U_k \times S_k \times V_k^T$$

- 基于SVD的商品评分预测

$$C_{P_{pred}} = \overline{C} + U_K \cdot \sqrt{S_k}'(c) \cdot \sqrt{S_k} . V_k'(P).$$

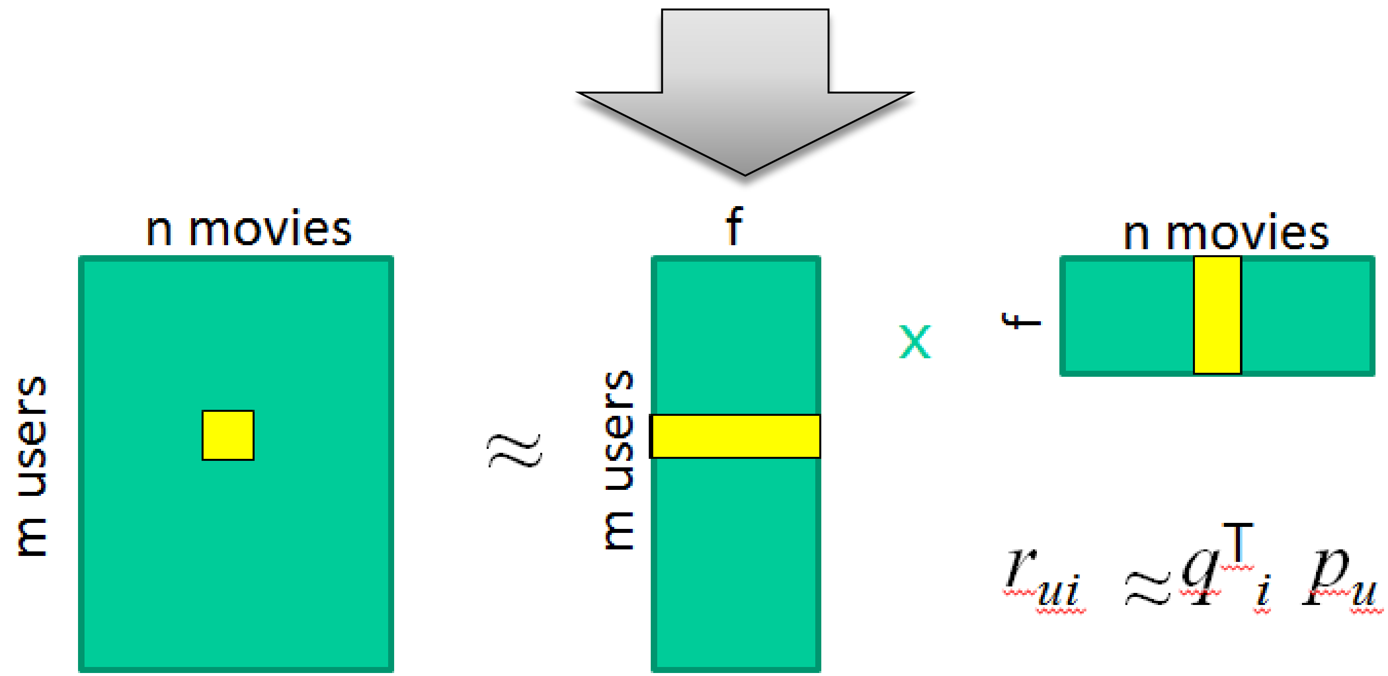
  - C 表示用户, P 表示商品(e.g. 电影)
  - 构建两个新的矩阵 $U_K \cdot \sqrt{S_k}'$ 和 $\sqrt{S_k} . V_k'$
  - 两个矩阵可以认为表示用户特征和商品特征

- 如果通过低秩分解，使得已知的评分和预测矩阵中的评分足够接近，则认为SVD分解相乘后得到的矩阵，可以预测未知的商品评分。

推荐系统中的矩阵分解

- 任务: 评分预测

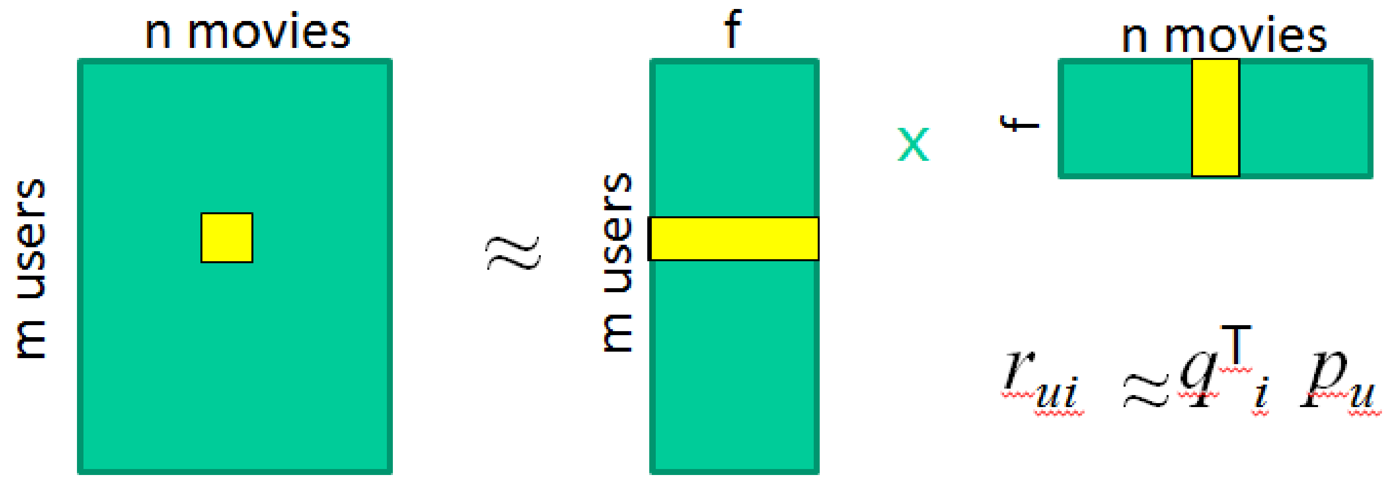
User	Harry Potter	Batman	Spiderman
U1	5	3	4
U2	?	2	4
U3	4	2	?



$$r_{ui} \approx q^{T}_{i}\, p_u$$

推荐系统中的矩阵分解

- **基本矩阵分解思想**



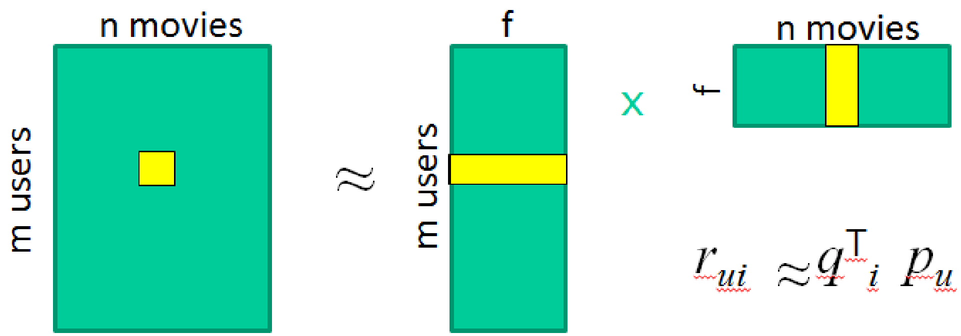
$$r_{ui} \approx q^{\mathsf{T}}_i \, p_u$$

- R = 评分矩阵, m 用户, n 商品;
- P = 用户矩阵, m 用户, f 潜在因素/特征;（反映用户喜好）
- Q = 商品矩阵, n 商品, f 潜在特征;（如电影类型）

- 评分r_{ui} 可以通过用户向量 p_u 和商品向量q_i 的点积来预测

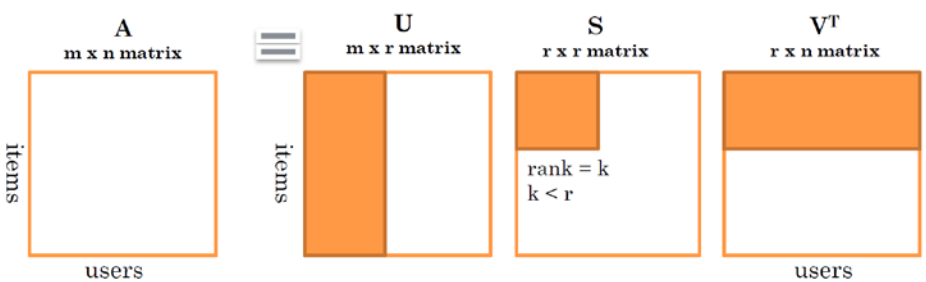
推荐系统中的矩阵分解

- **基本矩阵分解思想**



$$r_{ui} \approx q_i^T p_u$$

$$A = U \times S \times V^T$$



$$A_k = U_k \times S_k \times V_k^T$$

- SVD和MF的转换
  - P = 用户矩阵
  - Q = 商品矩阵
  - $U_K \cdot \sqrt{S_k}'$ = 用户矩阵
  - $\sqrt{S_k} \cdot V_k'$ = 商品矩阵

推荐系统中的矩阵分解

- **基本矩阵分解思想**
- 优化学习方法: 基于已知评分，训练学习用户矩阵和商品矩阵

$$\min_{q^* \cdot p^*} \sum_{(u,i) \in k} (r_{ui} - x_{ui})^2$$

- 其中
  - r_{ui} 为真实已知的用户u对商品i的评分
  - x_{ui} 为预测的用户u对商品i的评分
- X_{ui} 为用户特征p和商品特征q的点积

推荐系统中的矩阵分解

- **基本矩阵分解思想**
- 进一步优化：

$$\min_{q^* \cdot p^*} \sum_{(u,i) \in k} (r_{ui} - x_{ui})^2$$

$$\min_{q^* \cdot p^*} \sum_{(u,i) \in k} (r_{ui} - x_{ui})^2 + \lambda(|q_i|^2 + |p_u|^2)$$

- 综合优化目标代价函数：
  - 最小化误差
  - 正则化防止过拟合

推荐系统中的矩阵分解

User	Harry Potter	Batman	Spiderman
U1	5	3	4
U2	?	2	4
U3	4	2	?

	F1	F2	F3	F4	F5
U1	0.2763678	-0.37747	-1.26192	-1.54754	0.4738
U2	0.3999	-0.52747	-0.28946	-1.51597	0.73743
U3	0.2252336	-0.29125	-1.06624	-1.22463	0.37353

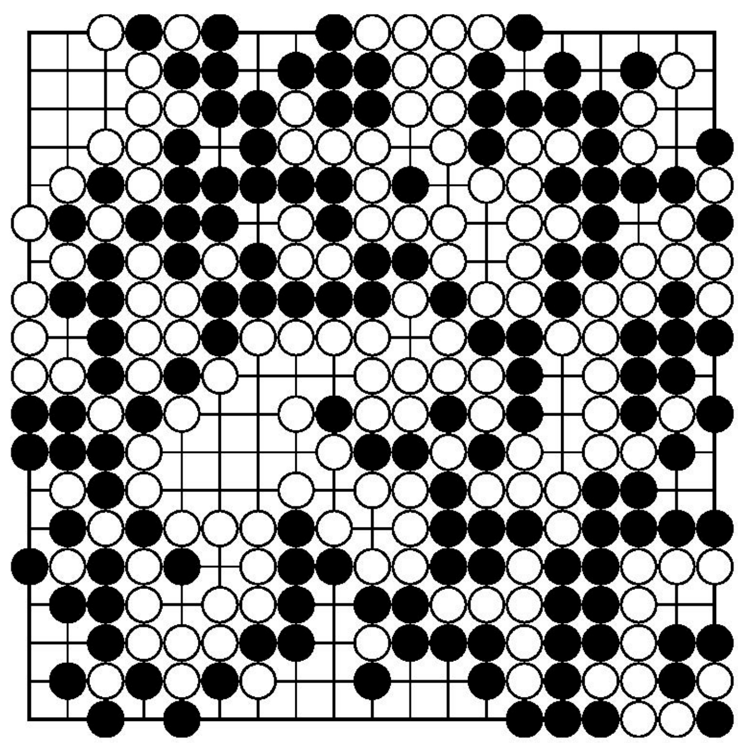
	HarryPotter	Batman	Spiderman
F1	0.301758	0.14297	0.43409
F2	-0.405407	-0.20245	-0.57514
F3	-1.399694	-0.9232	-0.46807
F4	-1.677619	-0.94013	-1.72021
F5	0.5118556	0.23861	0.79576

- 预测评分(U3, Spiderman) =
  - 黄色高亮向量点积 = 3.16822

实验练习

实验练习1

- 使用SVD，压缩所附棋谱二值图像，观测SVD压缩恢复图像效果。
- 自行添加随机噪声，观测SVD压缩恢复图像效果。



扩展练习2

- 基于已知观影评分数据，尝试使用矩阵分解，预测用户对影片的评分。
- 提交代码和简要分析报告，评价指标：
  - 1. 已知影片评分的预测误差
  - 2. 已知用户的top5/top10 最爱影片的预测误差
- 课程结束周，自愿提交。可作为课程综合成绩附加分。

```
#用户ID::影片ID::评分::时间戳
1::1193::5::978300760
1::661::3::978302109
1::914::3::978301968
1::3408::4::978300275
1::2355::5::978824291
1::1197::3::978302268
1::1287::5::978302039
1::2804::5::978300719
1::594::4::978302268
1::919::4::978301368
1::595::5::978824268
1::938::4::978301752
1::2398::4::978302281
1::2918::4::978302124
1::1035::5::978301753
```

End