1.

2. 矩阵的概念

   引入案例 (1)

   为了避免肥胖，提升员工健康状况，2018年初大数据部门组织月度跑步活动。规则如下：部门为参与者在月初定制月度目标，对完成目标者进行奖励，对未完成者进行惩罚，奖惩金额为：$x_j = t_j - e_j y_j = h_j y_j$，

   其中x_j为第j月总奖惩金额，t_j为总公里数，e_j为月度目标，h_j为实际距离与月度目标

   的差，y_j为每月对每公里的奖惩金额。活动影响良好，同时云部门也开展起来。以下数

   据为第一季度部分参与员工每月与月度目标差以及第一季度的总奖励值：

月份 / 姓名	h_1	h_2	h_3	w
小陈	10	8	12	20
小刘	4	4	2	8
小姚	2	-4	-2	-5

表1 大数据部门

月份 / 姓名	h_1	h_2	h_3	w
小李	2	4	5	10
小黄	4	2	2	6
小傅	-2	2	2	3

表2 云部门

   列方程求解，过掉

标量、向量、矩阵

- **向量**：一个向量是一列数。这些数是有序排列的。通过次序中的索引，我们可以确定每个单独的数。例如：

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- **矩阵**：由$m \times n$个数$a_{ij}(i=1, 2, \cdots, m; j=1, 2, \cdots, n)$排成$m$行$n$列的数表：

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix}$$

称为m行n列矩阵，简称$m \times n$矩阵。记作：

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

简记为$A = A_{m\times n} = (a_{ij})_{m\times n} = (a_{ij})$.特殊的，行数与列数都等于$n$的矩阵称为$n$阶矩阵或$n$阶方阵。

矩阵的加减乘与转置

```
# 加法
l=[[1,2,3],[4,5,6],[7,8,9]]
l1=[[9,8,7],[6,5,4],[3,2,1]]
l_result=[]
for index in range(len(l)):
    ll_result=[]
    l2=l[index]
    l3=l1[index]
    for i in range(len(l2)):
        ll_result.append(l2[i]+l3[i])
    l_result.append(ll_result)
for ll in l_result:
    print(ll)
```

```python
#减法
l=[[1,2,3],[4,5,6],[7,8,9]]
l1=[[1,2,3],[4,5,6],[7,8,9]]
l_result=[]
for index in range(len(l)):
    ll_result=[]
    l2=l[index]
    l3=l1[index]
    for i in range(len(l2)):
        ll_result.append(l2[i]-l3[i])
    l_result.append(ll_result)
for ll in l_result:
    print(ll)
```

```python
# 数字与矩阵相乘
l=[[1,2,3],[4,5,6],[7,8,9]]
index=2
l_result=[]
for l1 in l:
    ll_result=[]
    for i in l1:
       ll_result.append(i*index)
    l_result.append(ll_result)
for ll in l_result:
    print(ll)
```

```
#矩阵与矩阵相乘，要求第一个矩阵的列数必须等于第二个矩阵的行数
l1=[[1,2,3],[4,5,6]]#2X3的矩阵
l2=[[1,2],[3,4],[5,6]]#3X2的矩阵
index=0
l_result=[]
for i in range(len(l1)):
    ll_result = []
    for j in range(len(l2[0])):
        result = 0
        for k in range(len(l2)):
            result+=l1[i][k]*l2[k][j]
        ll_result.append(result)
    l_result.append(ll_result)
print(l_result)
```

```
#矩阵的转置
l1=[[1,2,3],[4,5,6]]
l_result=[]
for i in range(len(l1[0])):
    ll_result=[]
    for j in range(len(l1)):
        ll_result.append(l1[j][i])
    l_result.append(ll_result)
print(l_result)
```

主对角线：在一个n阶方阵中，从左上角到右下角这一斜线上的n 个元素的位置，叫做n阶方阵的主对角线

次对角线：在一个n阶方阵中，从右上角到左下角这一斜线上的n 个元素的位置，叫做n阶方阵的次对角线

对角矩阵：

对角矩阵

- **对角矩阵**：主对角线之外的元素皆为0的矩阵。常写为$diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

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

单位矩阵

单位矩阵：所有沿主对角线的元素都是1，而其他位置的所有元素都是0的矩阵。任意矩阵与单位矩阵相乘，都不会改变。

$$I_n=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}.$$

正交矩阵：设方阵$A=(a_{ij})_{n\times n}$，满足$AA^T=A^TA=I_n$，则称A为正交矩阵。即$A^{-1}=A^T$.

1. 概率

   为什么要使用概率？
   概率论使我们能够提出不确定的声明以及在不确定性存在的情况下进行推理，而信息论使我们能够量化概率分布中的不确定性总量

   随机试验

   满足以下三个特点的试验称为 随机试验：

   1. 可以在相同的条件下重复进行。

2. 每次试验的可能结果不止一个，并且能事先明确试验的所有可能结果。
3. 进行一次试验之前不能确定哪一个结果会出现。

举例：

1. 抛两枚硬币，出现正面I、反面U的情况
2. 抛一枚骰子，观察可能出现的点数情况

样本点、样本空间、随机事件
样本点（sample point）：）：随机试验的每一个可能的结果称为样本点，用f表示。
样本空间（sample space））：随机试验F的所有可能结果组成的集合，记作T，即T = f 1 ，f 2 ，…，f o .
随机事件（random variables events ））：样本空间T的任一子集B。属于事件B的样
本点出现，则称事件B发生。特别的，仅含一个样本点的随机事件，称为基本事件 基本事件。
举例：
随机试验E :抛一枚骰子，观察可能出现的点数情况。
样本空间为：S = 1，2，3，4，5，6 .
样本点为：e_i = 1，2，3，4，5，6.
随机事件A_1 ：“骰子出现的点数为5”，即A_1 ={x|x=5}

频率与概率
频率：在相同的条件下，进行n次试验，在这n次试验中，事件A发生的次数n_A 称为事件A发生的频数。

概率：比值 $\frac{n_A}{n}$ 称为事件A发生的概率，并记成P(A)
概率的特性：设E是随机试验，S是其样本空间。对于E的每一事件A赋予一个实数，记为P(A)，称为事件A的概率：
非负性：对于每一个事件A，有0 ≤ p(A) ≤ 1.
规范性：对于必然事件S，有P(S) = 1
可列可加性：设A_1 ， A_2 是两两互斥事件，则有P(A_1 + A_2)=P(A_1)+P(A_2) 加法定理

性质一：如果A属于B，则P(B-A)=P(B)-P(A) P(A)小于等于P(B)

性质二：如果A，B为随机事件，那么P(AUB)=P(A)+P(B)-P(AB) 广义加法定理

条件概率：

在很多情况下，我们感兴趣的是某个事件在给定其他事件发生时出现的概率

这种概率叫做条件概率P(B|A)，称作在事件A出现的情况下事件B的概率

乘法公式：

P(X|Y)=$\frac{P(YX)}{P(X)}$

设所有可能出现的试验结果为全体U={e1,e2.....en}，其中导致导致A出现的试验结果有m个，导致事件B出现的结果有k个，导致A与B同时出现的结果有r个，显然（r小于等于m，也小于等于k）事件A出现，也就是导致A事件的m个实验中有一个出现，在这个条件下，导致B出现的实验结果有且仅有r个

所以P(B|A)=r/m 同时除以n，得到p(AB)/P(A)

例子：一批零件 100个，次品率10% ，连续两次从这批零件中任取一个零件，第一次取出的零件不放回，求第二次才取到正品的概率

全概率公式：

设诸多事件A1，A2...An为两两互斥事件，且事件B为事件A1+A2+A3+...An的子事件

P(B)=P(A1)P(B|A1)+...P(An)P(B|An)

贝叶斯公式：

P(Ai|B)=P(Ai)P(B|Ai)/P(A1)P(B|A1)+...+P(An)P(B|An)

数学期望：

在试验中，每次可能结果的概率乘以其结果的总和，它反映随机变量平均取值的大小

方差：方差描述随机变量对于数学期望的偏离程度

平均数：$M=\frac{x_1+x_2+x_3+\cdots+x_n}{n}$（n表示这组数据个数，x1、x2、x3......xn表示这组数据具体数值）

方差公式：$s^2=\frac{(x_1-M)^2+(x_2-M)^2+(x_3-M)^2+\cdots+(x_n-M)^2}{n}$