HW10

xw-zeng

2022-12-20

Problem 1

1、 $\sum_k \rho^2(X_k, X_i) = \sum_k \frac{\lambda_k a_{ik}^2}{\sigma_{ii}} = (\sqrt{\lambda_k} a_k^T) \Sigma^{-1} (\sqrt{\lambda_k} a_k) = \lambda_k (a_k^T \Sigma^{-1} a_k).$

问题转化为证明 $a_k^T \Sigma^{-1} a_k = \lambda_k^{-1}$.

a_k 为 Σ 的特征值 λ_k 对应的单位特征向量.

$\lambda_k a_k = \Sigma a_k$ ← 特征值定义.

$\lambda_k \Sigma^{-1} a_k = \Sigma^{-1} \Sigma a_k$ (左乘 Σ^{-1}).

$\lambda_k a_k^T \Sigma^{-1} a_k = a_k^T a_k = \|a_k\|^2 = 1$ (左乘 a_k, a_k为单位特征向量, 模长为 1).

$\therefore a_k^T \Sigma^{-1} a_k = \lambda_k^{-1}$. $\Rightarrow \sum_k \rho^2(Y_k, X_i) = 1$ 得证.

Problem 2

导入样本点。

```
X <- data.frame(matrix(c(2,4,5,5,6,8,2,3,3,4,5,7),6,2)); X
```

```
##   X1 X2
## 1  2  2
## 2  4  3
## 3  5  3
## 4  5  4
## 5  6  5
## 6  8  7
```

计算样本协方差矩阵。

```r
covX <- cov(X); covX
```

```
##     X1  X2
## X1 4.0 3.4
## X2 3.4 3.2
```

1. 基于样本协方差矩阵

对样本协方差矩阵进行特征值分解，得到特征值与相应的特征向量。

```r
lambda <- eigen(covX)$values; lambda
```

```
## [1] 7.0234486 0.1765514
```

```r
a <- eigen(covX)$vectors; a
```

```
##            [,1]       [,2]
## [1,] -0.7472755  0.6645144
## [2,] -0.6645144 -0.7472755
```

计算得到主成分。

```r
Y <- t(a) %*% t(as.matrix(X))
t(Y)
```

```
##            [,1]        [,2]
## [1,]  -2.823580 -0.16552215
## [2,]  -4.982645  0.41623116
## [3,]  -5.729921  1.08074555
## [4,]  -6.394435  0.33347009
## [5,]  -7.806225  0.25070901
## [6,] -10.629804  0.08518686
```

2. 基于样本相关系数矩阵

计算样本相关系数矩阵。

```r
corX <- cor(X)
```

对样本相关系数矩阵进行特征值分解，得到特征值与相应的特征向量。

```r
lambda <- eigen(corX)$values; lambda
```

```
## [1] 1.95032889 0.04967111
```

```r
a <- eigen(corX)$vectors; a
```

```
##           [,1]       [,2]
## [1,] 0.7071068 -0.7071068
## [2,] 0.7071068  0.7071068
```

计算得到主成分。

```r
Y <- t(a) %*% t(as.matrix(X))
t(Y)
```

```
##           [,1]       [,2]
## [1,]  2.828427  0.0000000
## [2,]  4.949747 -0.7071068
## [3,]  5.656854 -1.4142136
## [4,]  6.363961 -0.7071068
## [5,]  7.778175 -0.7071068
## [6,] 10.606602 -0.7071068
```