

## HOMEWORK PCA

1.1 设  $X = (X_1, \dots, X_m)^\top$  是  $m$  维随机变量, 均值为  $E(X) \stackrel{\text{def}}{=} \mu$ , 协方差矩阵为  $\text{cov}(X) \stackrel{\text{def}}{=} \Sigma$ . 设  $\Sigma$  的特征值为  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ , 特征值对应的单位特征向量为  $\alpha_1, \dots, \alpha_m$  则  $X$  的第  $k$  个主成分是  $Y_k = \alpha_k^\top X$ , 方差为  $\text{var}(Y_k) = \alpha_k^\top \Sigma \alpha_k$ .

证明以下性质:

$$\sum_k \rho^2(Y_k, X_i) = 1$$

其中,  $\rho(Y_k, X_i) = \frac{\sqrt{\lambda_k} \alpha_{ik}}{\sqrt{\sigma_{ii}}}$ ,  $\sigma_{ii} = \text{var}(X_i)$ ,  $\alpha_{ik} = e_i^\top \alpha_k$ ,  $\alpha_k$  为第  $k$  个特征向量,  $e_i$  为基本单位向量, 其第  $i$  个变量为 1, 其余为 0.

1.2 对以下样本数据进行主成分分析:

$$X = \begin{bmatrix} 2 & 2 \\ 4 & 3 \\ 5 & 3 \\ 5 & 4 \\ 6 & 5 \\ 8 & 7 \end{bmatrix}$$

以上题目请以PDF格式提交。

提交时间: 12月22日, 晚20:00之前。请预留一定的时间, 迟交作业扣3分, 作业抄

袭0分。