Week 7 PYQs
Machine Learning Foundations
May 22 - Sep 25

Note

1. If no questions are listed in that term Quiz, it indicates that the corresponding quiz did not include questions from that week.

2. The answer key may contain minor inaccuracies, as answers were taken directly from the original quiz/exam papers. Some of these might have been corrected later.

3. For any issues or suggestions regarding this document, feel free to contact:
   `23f3004032@ds.study.iitm.ac.in`
   or reach out on Discourse: `@23f3004032`

Quiz 1 May 22

Quiz 2 May 22

1. Consider the data points x_1, x_2, x_3 to answer the given subquestions:

$$x_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

   (a) Calculate the mean vector of the data points x_1, x_2, x_3 and write the summation of all elements of the mean vector obtained. [Ans:3]

   (b) Let

$$C = \frac{1}{3} \sum_{i=1}^{3} (x_i - \bar{x})(x_i - \bar{x})^T$$

   for the data points x_1, x_2, x_3 is calculated. Find the trace of C. [Ans: 1.4]

   (c) Project data points x_1, x_2, x_3 onto a one dimensional space using PCA. Let z_1, z_2, z_3 denote the projection of x_1, x_2, x_3 respectively. Calculate the summation of all elements of z_2.

   [Ans: 0]

ET May 22

Quiz 1 Sep 22

Quiz 2 Sep 22

1. Consider the data points

$$x_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix},\ x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ x_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

If projected onto first principal component, what is projected variance?

(a) 0.34

(b) 1

(c) 1.34

(d) None of these

[Ans: 1.34]

2. Consider these data points to answer the following question:

$$x_1 = \begin{bmatrix} -4 \\ 4 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad x_4 = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$$

The projected (new) data points x_1', x_2', x_3', x_4' corresponding to the original points x_1, x_2, x_3, x_4 in one dimensional PCA is

(a)
$$x_1' = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad x_2' = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad x_3' = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \quad x_4' = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

(b)
$$x_1' = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad x_2' = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \quad x_3' = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \quad x_4' = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

(c)
$$x_1' = \begin{bmatrix} -4 \\ 4 \end{bmatrix}, \quad x_2' = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \quad x_3' = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad x_4' = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$$

(d)
$$x_1' = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \quad x_2' = \begin{bmatrix} -4 \\ 4 \end{bmatrix}, \quad x_3' = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad x_4' = \begin{bmatrix} 4 \\ -4 \end{bmatrix}$$

[Ans:c]

ET Sep 22

1. What are the properties of principal components in PCA?

   (a) All principal components are orthogonal to each other.

   (b) The number of principal components for n-dimensional data are atmost n-1.

   (c) The first principal component accounts for most of the possible variability of the original data i.e, maximum possible variance.

   (d) The first principal component is the eigenvector of the covariance matrix corresponding to the maximum eigenvalue.

   [Ans: 1, 3, 4]

Quiz 1 Jan 23

Quiz 2 Jan 23

1. For dataset $D = \{x_1, \ldots, x_n\}$, the matrix

$$C = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T$$

is called covariance matrix:

(a) always

(b) only when dataset is centered

(c) only when dataset has maximum variance

(d) Both when dataset is centered and when dataset has maximum variance

[Ans: D]

2. Which are limitations of PCA?

(a) PCA does work well for non-linearly correlated data.

(b) PCA always consider the low variance components in the data as noise and recommend us to throw away those components. But, sometimes those components play a major role in a supervised learning task.

(c) If the variables are correlated, PCA can achieve dimension reduction. If not, PCA just orders them according to their variances.

(d) PCA always finds orthogonal principal components. Sometimes, our data demands non-orthogonal principal components to represent the data.

[Ans: B, C, D]

3.

$$x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$C = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T, \quad \text{where } \bar{x} = \frac{x_1 + x_2 + x_3}{3}$$

What is the sum of the eigenvalues of the covariance matrix C? (correct to two decimals)

[Ans: 1.33]

ET Jan 23

1. Suppose you have a two-dimensional dataset $x_1, x_2, x_3, ...x_n$ with mean zero. Suppose the covariance matrix $C = \sum_{i=1}^{n} x_i x_i^T = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$. For projection using PCA onto a line, what is the reconstruction error? Enter the answer as integer. [Ans: 2]

Quiz 1 May 23

Quiz 2 May 23

1. Suppose you have a 3-dimensional dataset $\{x_1, x_2, \ldots, x_n\}$ with mean zero. Suppose the covariance matrix

$$C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

   For projection using PCA onto a line, what is the projected variance?

   [Ans: 3]

ET May 23

1. Which among the following is true for PCA?

   (a) The first principal component represents the direction along which the variance of the dataset is maximized.

   (b) The first principal component represents the direction along which the variance of the dataset is minimized.

   (c) The first principal component is the eigenvector corresponding to the smallest eigenvalue of the covariance matrix.

   (d) The first principal component is the eigenvector corresponding to the largest eigenvalue of the covariance matrix.

   [Ans: 1, 4]

Quiz 1 Sep 23

Quiz 2 Sep 23

1. Which among the following statements are true about principal component analysis (PCA)?

   (a) PCA is a supervised learning algorithm.

   (b) Principal components are the eigenvectors of the sample covariance matrix.

   (c) Principal components are eigenvectors of the centered data matrix.

   (d) If p_1 and p_2 are the principal component vectors, then p_1 is orthogonal to p_2.

   [Ans: B, D]

2. Consider dataset

$$\left\{ \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

   (a) Compute the sample covariance matrix.

   i. $C = \frac{1}{4} \begin{bmatrix} 46 & 30 \\ 30 & 36 \end{bmatrix}$

   ii. $C = \frac{1}{4} \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix}$

   iii. $C = \frac{1}{4} \begin{bmatrix} 46 & 30 \\ 30 & 46 \end{bmatrix}$

   iv. $C = \frac{1}{4} \begin{bmatrix} 36 & 30 \\ 30 & 36 \end{bmatrix}$

   [Ans: 2]

   (b) Which is the principal direction that is chosen for PCA?

   i. $\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

   ii. $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

   iii. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

   iv. $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

   [Ans: 2]

   (c) If $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ and $\tilde{x}_4$ are the projections of the data points onto the first principal component, then which among the following are true?

   i. $\tilde{x}_1 = (3, 0), \tilde{x}_2 = (3, 0), \tilde{x}_3 = (5, 0), \tilde{x}_4 = (1, 0)$

ii. $\tilde{x}_1 = (5/2, 5/2), \tilde{x}_2 = (5/2, 5/2), \tilde{x}_3 = (9/2, 9/2), \tilde{x}_4 = (1/2, 1/2)$

iii. $\tilde{x}_1 = (3, 0), \tilde{x}_2 = (3, 0), \tilde{x}_3 = (5, 4), \tilde{x}_4 = (1, 0)$

iv. $\tilde{x}_1 = (3, 2), \tilde{x}_2 = (3, 2), \tilde{x}_3 = (5, 4), \tilde{x}_4 = (1, 0)$

[Ans: Option 4]

ET Sep 23

1. Consider the following data points: $X_1 = \begin{bmatrix}0\\0\end{bmatrix}, X_2 = \begin{bmatrix}2\\1\end{bmatrix}, X_3 = \begin{bmatrix}1\\2\end{bmatrix}$. For the given data points, the goal is to find transformed data points for one dimensional PCA. Based on the information, answer the given subquestions. Which of the following represents the covariance matrix, C for the given data points?

   (a) $\frac{1}{3}\begin{bmatrix}2 & 2\\2 & 2\end{bmatrix}$

   (b) $\begin{bmatrix}1 & 2\\2 & 1\end{bmatrix}$

   (c) $\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$

   (d) $\frac{1}{3}\begin{bmatrix}2 & 1\\1 & 2\end{bmatrix}$

   [Ans: 1]

2. What is the value of the projected variance? [Ans: 1]

Quiz 1 Jan 24

Quiz 2 Jan 24

1. Consider dataset
$$\left\{ \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix} \right\}$$

   (a) Compute the sample covariance matrix.

   i. $C = \frac{1}{3} \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}$

   ii. $C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

   iii. $C = \frac{1}{3} \begin{bmatrix} 10 & 10 \\ 10 & 10 \end{bmatrix}$

   iv. $C = \frac{1}{3} \begin{bmatrix} 11 & 11 \\ 11 & 11 \end{bmatrix}$

   [Ans: 1]

   (b) Which is the principal direction for PCA?

   i. $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

   ii. $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

   iii. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

   iv. $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

   [Ans: 2]

   (c) If $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ are projections onto the first principal component, which is correct?

   i. $\tilde{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

   ii. $\tilde{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

   iii. $\tilde{x}_1 = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 3/2 \\ 1/2 \end{bmatrix}$

   iv. $\tilde{x}_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

   [Ans: 4]

ET Jan 24

1. Consider a data set $\{x_1, x_2, ..., x_n\}$, where $x_i \in \mathbb{R}^d, i = 1, 2, ...n$ and $d >> n$. Let $A = \begin{bmatrix} (x_1 - \overline{x})^T \\ (x_2 - \overline{x})^T \\ \vdots \\ (x_n - \overline{x})^T \end{bmatrix}$, and $C = \frac{1}{n} A^T A$, where $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$. Which of the following options is/are true?

   (a) $Rank(C) \geq n$

   (b) $Rank((x_i - \overline{x})(x_i - \overline{x})^T) = 1$ for all $i = 1, 2, ..., n$

   (c) For this problem, it is efficient to perform PCA using eigenvalues/eigenvectors of $\frac{1}{n} AA^T$

   (d) For this problem, it is efficient to perform PCA using eigenvalues/eigenvectors of C.

[Ans: 2, 3]

Quiz 1 May 24

Quiz 2 May 24

1. What is the primary goal of Principal Component Analysis (PCA)?

   (a) To find directions that capture the maximum variance in the data

   (b) To find directions that capture the minimum variance

   (c) To increase dimensionality

   (d) To reduce number of data points

[Ans: 1]

2. Consider dataset

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \right\}$$

   (a) Compute the sample covariance matrix.

   i. $C = \frac{1}{3} \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$

   ii. $C = \frac{1}{3} \begin{bmatrix} 2 & -2 & 2 \\ -2 & 2 & -2 \\ 2 & -2 & 2 \end{bmatrix}$

   iii. $C = \frac{1}{3} \begin{bmatrix} -2 & 2 & -2 \\ 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix}$

   iv. $C = \begin{bmatrix} -2 & 2 & -2 \\ 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix}$

[Ans: 2]

   (b) Which is the principal direction chosen for PCA?

   i. $\begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix}$

   ii. $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$

   iii. $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$

iv. $\begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

[Ans: 4]

(c) If $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ are projections onto the first principal component, which is correct?

i. $\tilde{x}_1 = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} -2 \\ 4 \\ -1 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 7 \\ -1 \\ 4 \end{bmatrix}$

ii. $\tilde{x}_1 = \begin{bmatrix} -4/3 \\ 2/3 \\ -1/3 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} 2/3 \\ -4/3 \\ 1/3 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} -7/3 \\ 1/3 \\ -4/3 \end{bmatrix}$

iii. $\tilde{x}_1 = \begin{bmatrix} 4/3 \\ 2/3 \\ 1/3 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} 2/3 \\ 4/3 \\ 1/3 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 7/3 \\ 1/3 \\ 4/3 \end{bmatrix}$

iv. $\tilde{x}_1 = \begin{bmatrix} 4/3 \\ -2/3 \\ 1/3 \end{bmatrix}$, $\tilde{x}_2 = \begin{bmatrix} -2/3 \\ 4/3 \\ -1/3 \end{bmatrix}$, $\tilde{x}_3 = \begin{bmatrix} 7/3 \\ -1/3 \\ 4/3 \end{bmatrix}$

[Ans: 4]

ET May 24

1. Which of the following is/are true about PCA?

   (a) PCA will transform the original data set onto a lower dimension subspace such that the variance of the project is maximized.

   (b) PCA calculates the mean of each data set to determine its significance.

   (c) PCA can be used to reduce the dimensionality of the dataset.

   (d) PCA will transform the original data set onto a lower dimension subspace such that the reconstruction error is maximized.

   [Ans: 1, 3]

Quiz 1 Sep 24

Quiz 2 Sep 24

1. Consider the following dataset:

$$\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \end{pmatrix} \right\}$$

Suppose we want to project the above dataset onto a 1 dimensional space using PCA.

(a) The sample covariance matrix C for the given dataset is

i. $C = \frac{1}{3} \begin{pmatrix} 8 & 8 \\ 8 & 8 \end{pmatrix}$

ii. $C = \frac{1}{3} \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}$

iii. $C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

iv. $C = \frac{1}{3} \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$

[Ans: 1]

(b) The eigenvalues of the covariance matrix are

i. $\lambda_1 = 16, \lambda_2 = 0$

ii. $\lambda_1 = 4, \lambda_2 = 0$

iii. $\lambda_1 = 1, \lambda_2 = 0$

iv. $\lambda_1 = \frac{16}{3}, \lambda_2 = 0$

[Ans: 4]

(c) If $\tilde{x}_1, \tilde{x}_2$ and $\tilde{x}_3$ are the projections of the data points onto the first principal component, then which among the following is correct?

i. $\tilde{x}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \tilde{x}_2 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \tilde{x}_3 = \begin{pmatrix} 6 \\ 5 \end{pmatrix}$

ii. $\tilde{x}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \tilde{x}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tilde{x}_3 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

iii. $\tilde{x}_1 = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}, \tilde{x}_2 = \begin{pmatrix} 2/2 \\ 2/2 \end{pmatrix}, \tilde{x}_3 = \begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}$

iv. $\tilde{x}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \tilde{x}_2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \tilde{x}_3 = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$

[Ans: 1]

ET Sep 24

Quiz 1 Jan 25

Quiz 2 Jan 25

1. Consider the following dataset of three points in $\mathbb{R}^2$: $\{\begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}3\\6\end{bmatrix}, \begin{bmatrix}5\\10\end{bmatrix}\}$. Compute the sample covariance matrix C for the given dataset.

   (a) $C = \frac{1}{3}\begin{bmatrix}6 & 12\\12 & 24\end{bmatrix}$

   (b) $C = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$

   (c) $C = \frac{1}{9}\begin{bmatrix}14 & 28\\28 & 56\end{bmatrix}$

   (d) $C = \frac{1}{9}\begin{bmatrix}10 & 20\\20 & 40\end{bmatrix}$

   [Ans: 1]

2. What is the principal direction chosen for PCA for the dataset $\{\begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}3\\6\end{bmatrix}, \begin{bmatrix}5\\10\end{bmatrix}\}$?

   (a) $\frac{1}{\sqrt{5}}\begin{bmatrix}2\\1\end{bmatrix}$

   (b) $\frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix}$

   (c) $\begin{bmatrix}1\\0\end{bmatrix}$

   (d) $\begin{bmatrix}0\\1\end{bmatrix}$

   [Ans: 2]

3. Compute the projections $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$ of the data points $\{\begin{bmatrix}2\\4\end{bmatrix}, \begin{bmatrix}3\\6\end{bmatrix}, \begin{bmatrix}5\\10\end{bmatrix}\}$ onto the first principal component. Which of the following is correct?

   (a) $\tilde{x}_1 = \begin{bmatrix}2\\4\end{bmatrix}, \tilde{x}_2 = \begin{bmatrix}3\\6\end{bmatrix}, \tilde{x}_3 = \begin{bmatrix}5\\10\end{bmatrix}$

   (b) $\tilde{x}_1 = \begin{bmatrix}1/\sqrt{5}\\2/\sqrt{5}\end{bmatrix}, \tilde{x}_2 = \begin{bmatrix}2/\sqrt{5}\\4/\sqrt{5}\end{bmatrix}, \tilde{x}_3 = \begin{bmatrix}4/\sqrt{5}\\8/\sqrt{5}\end{bmatrix}$

   (c) $\tilde{x}_1 = \begin{bmatrix}-2\\-4\end{bmatrix}, \tilde{x}_2 = \begin{bmatrix}-3\\-6\end{bmatrix}, \tilde{x}_3 = \begin{bmatrix}-5\\-10\end{bmatrix}$

(d) $\tilde{x}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 4/\sqrt{5} \end{bmatrix}, \tilde{x}_2 = \begin{bmatrix} 3/\sqrt{5} \\ 6/\sqrt{5} \end{bmatrix}, \tilde{x}_3 = \begin{bmatrix} 5/\sqrt{5} \\ 10/\sqrt{5} \end{bmatrix}$

[Ans: 1]

ET Jan 25

1. In the context of Linear Algebra, which of the following statements about Principal Component Analysis (PCA) is/are true?

   (a) The principal components of a matrix are given by the eigenvectors of its covariance matrix.

   (b) The principal components of a matrix are given by the singular values of its covariance matrix.

   (c) PCA finds the principal components by maximizing the variance along each component.

   (d) The principal components form a basis that is not necessarily orthogonal.

   [Ans: 1, 3]

Quiz 1 May 25

Quiz 2 May 25

1. Let $A \in \mathbb{R}^{n \times d}$ be a centered data matrix, where $d > n$. Let $u_i \in \mathbb{R}^n$ be an eigenvector of AA^T corresponding to the eigenvalue $\lambda_i > 0$, representing the i-th principal component in the sample space. Let $v_i \in \mathbb{R}^d$ denote the corresponding principal component direction in the feature space (i.e., the eigenvector of A^TA). Which of the following correctly computes v_i from u_i?

   (a) The i-th principal component direction is $v_i = u_i$.

   (b) The i-th principal component direction is $v_i = Au_i$.

   (c) The i-th principal component direction is $v_i = \frac{1}{\sqrt{\lambda_i}} A^T u_i$.

   (d) The i-th principal component direction is $v_i = \frac{1}{\lambda_i} A^T u_i$.

   [Ans: 3]

2. Consider the following dataset: $\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}5\\6\end{pmatrix}\}$. Suppose we want to project the above dataset onto a 1-dimensional space using PCA. The sample covariance matrix C for the given dataset is:

   (a) $C = \frac{1}{3}\begin{pmatrix}8 & 8\\8 & 8\end{pmatrix}$

   (b) $C = \frac{1}{3}\begin{pmatrix}4 & 4\\4 & 4\end{pmatrix}$

   (c) $C = \begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}$

   (d) $C = \frac{1}{3}\begin{pmatrix}2 & 2\\2 & 2\end{pmatrix}$

   [Ans: 1]

3. The eigenvalues of the covariance matrix for the dataset $\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}5\\6\end{pmatrix}\}$ are:

   (a) $\lambda_1 = 4, \lambda_2 = 0$

   (b) $\lambda_1 = \frac{8}{3}, \lambda_2 = 0$

   (c) $\lambda_1 = \frac{16}{3}, \lambda_2 = 0$

   (d) $\lambda_1 = 8, \lambda_2 = 0$

   [Ans: 3]

4. The first principal component for the dataset $\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}5\\6\end{pmatrix}\}$ is:

(a) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

(c) $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(d) $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

[Ans: 3]

ET May 25

Quiz 1 Sep 25

Quiz 2 Sep 25

1. Consider a dataset D = (-2, -2), (-1,-1), (0, 0), (1, 1), (2, 2). If the covariance matrix of the dataset is $C = \begin{bmatrix} a & a \\ a & a \end{bmatrix}$,

   (a) find a. [Ans: 1.9 to 2.1]

   (b) If the unit-norm eigenvector corresponding to the largest eigenvalue of C is v, find $v^T C v$. [Ans: 3.9 to 4.1]

   (c) The data-points (100, 100) and (-100, 100) are added to D to get D'. PCA is run on D'. Let the covariance matrix of D' be C' and the unit-norm eigenvector corresponding to the largest eigenvalue of C' be v'. Which of the following is true?

      i. $v' = v$ or $v' = -v$
      ii. v and v' are orthogonal to each other
      iii. The angle between v and v' is 45°
      iv. We can't comment on the angle between v and v' with the given information

   [Ans: 2]

ET Sep 25

1. In Principal Component Analysis (PCA), the first principal component is the direction that:

   (a) Minimizes the variance of the projected data.

   (b) Maximizes the variance of the projected data.

   (c) Is always the x-axis.

   (d) Is the eigenvector corresponding to the smallest eigenvalue of the covariance matrix.

   [Ans: 2]

2. Consider the following dataset: $X = \{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\3\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}\}$. The sample covariance matrix C for the given dataset is:

   (a) $\frac{1}{3}\begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}$

   (b) $\begin{pmatrix}2 & 1\\1 & 3\end{pmatrix}$

   (c) $\begin{pmatrix}3 & 1\\1 & 2\end{pmatrix}$

   (d) $\begin{pmatrix}10 & 5\\5 & 6\end{pmatrix}$

   [Ans: 1]