Week_2_mlt_questions

Jan 2024 Quiz 1

1. Let $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a valid kernel. Is $\sin(x_1)\cos(x_2)k(x_1, x_2)$ a valid kernel? Here, $x_1, x_2 \in \mathbb{R}$?

(a) Yes

(b) No

[Ans: b]

2. Consider the following kernel:

$$k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

$$k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T\mathbf{y})^2 + 1$$

Which of the following transformation mapping ϕ may correspond to the kernel k?

(a) $\phi([x_1, x_2]^T) = [x_1^2, \sqrt{2}x_1x_2, x_2^2, 1]^T$

(b) $\phi([x_1, x_2]^T) = [x_1^2, x_1 + x_2, x_2^2, 1]^T$

(c) $\phi([x_1, x_2]^T) = [x_1, \sqrt{2}x_1^2x_2^2, x_2, 1]^T$

(d) $\phi([x_1, x_2]^T) = [x_1, x_1x_2, x_2, 1]^T$

[Ans: a]

3. Consider a dataset consisting of 1000 samples in a 20-dimensional space. You apply Kernel PCA with a polynomial kernel of degree 3 to reduce the dimensionality of the data. Which of the following statements regarding Kernel PCA is true?

(a) Kernel PCA always results in a lower-dimensional representation of the data compared to standard PCA.

(b) Kernel PCA can only be applied to datasets that are linearly separable.

(c) Choosing a higher degree polynomial kernel in Kernel PCA can lead to increased flexibility in capturing non-linear relationships.

(d) None of these.

[Ans: c]

May 2024 Quiz 1

4. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ be two vectors in $\mathbb{R}^2$. Let k_1 and k_2 be two functions defined from $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$:

$$k_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + x_1 x_2 y_1 y_2$$

$$k_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + x_1 x_2 y_1 y_2 + 1$$

Which of the following statements is true?

(a) Both k_1 and k_2 are valid kernels.

(b) k_1 is a valid kernel, but k_2 is not a valid kernel.

(c) k_2 is a valid kernel, but k_1 is not a valid kernel.

(d) Neither k_1 nor k_2 is a valid kernel.

[Ans: a]

5. Kernel PCA with a polynomial kernel of degree 3 is run on a dataset of 100 data-points. Let ϕ be the transformation corresponding to this kernel. If the largest eigenvalue of the centered kernel matrix is 10, find the largest eigenvalue of the covariance matrix of the transformed dataset, where ϕ is used to effect the transformation.

[Ans: 0.1]

6. $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is a polynomial kernel of degree 4. Find the value of $k\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)$.

[Ans: 16]

Sep 2024 Quiz 1

7. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ be vectors in $\mathbb{R}^2$. Define the functions k_1 and k_2 as:

$$k_1(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + (x_1 + x_2)(y_1 + y_2)$$

$$k_2(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + (x_1^2 + y_2^2) + 3$$

Which of the following statements is true?

(a) Both k_1 and k_2 are valid kernels.

(b) k_1 is a valid kernel, but k_2 is not a valid kernel.

(c) k_2 is a valid kernel, but k_1 is not a valid kernel.

(d) Neither k_1 nor k_2 is a valid kernel.

[Ans: b]

8. Consider the following kernel:

$$k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

$$k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y})^2 + 1$$

Which of the following transformation mapping ϕ may correspond to the kernel k?

(a) $\phi([x_1, x_2]^T) = [x_1^2, \sqrt{2}x_1 x_2, x_2^2, 1]^T$

(b) $\phi([x_1, x_2]^T) = [x_1^2, x_1 + x_2, x_2^2, 1]^T$

(c) $\phi([x_1, x_2]^T) = [x_1, \sqrt{2}x_1^2 x_2^2, x_2, 1]^T$

(d) $\phi([x_1, x_2]^T) = [x_1, x_1 x_2, x_2, 1]^T$

[Ans: a]

9. Given n data points in a d-dimensional space with a non-linear relationship, we apply kernel PCA to reduce the dimensionality and select the first k principal components. Is it possible for k to be greater than d?

(a) Yes

(b) No

[Ans: a]

Jan 2025 Quiz 1

10. Which of the following statements is/are true? (MSQ)

(a) $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ k(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1^T \mathbf{x}_2)^3$ is a valid kernel.

(b) $k : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ k(\mathbf{x}_1, \mathbf{x}_2) = e^{-\|\mathbf{x}_1 - \mathbf{x}_2\|^2}$ is a valid kernel.

(c) The trace of XX^T is equal to the trace of $X^T X$.

(d) Eigenvectors of XX^T are the same as eigenvectors of $X^T X$.

[Ans: a, b, c]

May 2025 Quiz 1

11. Let $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a valid kernel function such that $k(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^T \phi(\mathbf{y})$, where $\phi : \mathbb{R}^d \to \mathbb{R}^D$, where $d < D$. Define

$$k_1(\mathbf{x}, \mathbf{y}) = \frac{100k(\mathbf{x}, \mathbf{y})}{\sqrt{k(\mathbf{x}, \mathbf{x})k(\mathbf{y}, \mathbf{y})}}$$

Find $\phi_1 : \mathbb{R}^d \to \mathbb{R}^D$ such that $k_1(\mathbf{x}, \mathbf{y}) = \phi_1(\mathbf{x})^T \phi_1(\mathbf{y})$. (Note: $\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^T \mathbf{x}}$)

(a) $\phi_1(\mathbf{x}) = \frac{100\phi(\mathbf{x})}{\|\phi(\mathbf{x})\|_2}$

(b) $\phi_1(\mathbf{x}) = \frac{\phi(\mathbf{x})}{10\|\phi(\mathbf{x})\|_2}$

(c) $\phi_1(\mathbf{x}) = \frac{10\phi(\mathbf{x})}{\|\phi(\mathbf{x})\|_2}$

(d) Such a ϕ_1 does not exist, i.e., k_1 is not a valid kernel.

[Ans: c]

12. Kernel PCA with a polynomial kernel is applied on the following dataset with three features:

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

The k-th principal component is given by $\mathbf{w}_k = \phi(X)\boldsymbol{\alpha}_k$, where ϕ is the transformation mapping corresponding to the kernel $k(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i^T \mathbf{x}_j)^2$ for some $\boldsymbol{\alpha}_k \in \mathbb{R}^n$. Find the centered kernel matrix K^c.

(a) $\begin{bmatrix} -1.5 & 1.5 \\ 1.5 & -1.5 \end{bmatrix}$

(b) $\begin{bmatrix} 1.5 & -1.5 \\ -1.5 & 1.5 \end{bmatrix}$

(c) $\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & -1 \\ 4 & -4 \end{bmatrix}$

[Ans: b]

2024 Apr 28 ET

13. Which of the following statements accurately describes the characteristics of kernel functions? Assume the dataset to be mean-centered. (MSQ)

(a) Kernel PCA can reconstruct original PCA if the kernel function is $k(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + 1)^2$.

(b) The dimensionality of the transformed dataset $\phi(X)$, computed using the kernel function, is always smaller than the original feature space.

(c) The dimensionality of the transformed dataset $\phi(X)$, computed using the kernel function, can exceed the original feature space.

(d) The dimension of the transformed dataset $\phi(X)$, whose inner products the kernel function computes, can be infinite.

[Ans: c, d]

2024 Sep 01 ET

14. Consider $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, a polynomial kernel of degree p. If the kernel always outputs a non-negative value, which of the following are possible values for p? (MSQ)

(a) 2

(b) 4

(c) 3

(d) 5

[Ans: a, b]

2025 Apr 13 ET

15. A dataset of 1000 second-hand cars has four features: kilometers driven (x_1), mileage (x_2), the present price of the car (x_3), and the selling price (x_4). The selling price seems to have the following relationship (approximate) with the other three features.

$$x_4^2 = x_1 x_2 + x_1 x_2 x_3$$

If we want to project the dataset into a lower dimensional space, which of the following task would be most appropriate?

(a) Kernel PCA with a polynomial kernel of degree 2.

(b) Kernel PCA with a polynomial kernel of degree 3.

(c) Kernel PCA with a polynomial kernel of degree 4.

(d) Standard PCA

[Ans: b]

2025 Aug 31 ET

16. Consider the following kernel function: $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$

$$k([\mathbf{x}_1, \mathbf{x}_2]^T, [\mathbf{y}_1, \mathbf{y}_2]^T) = 1 + x_1y_1 + x_2y_2 + x_1^2y_1^2x_2y_2 + x_1y_1x_2^2y_2^2$$

Find the appropriate transformation mapping ϕ for the given kernel.

(a) $\phi([x_1, x_2]^T) = [1, x_1x_2, x_1^2x_2^2]^T$

(b) $\phi([x_1, x_2]^T) = [1, x_1, x_2, x_1^2x_2^2]^T$

(c) $\phi([x_1, x_2]^T) = [1, x_1, x_2, x_1^2x_2, x_1x_2^2]^T$

(d) $\phi([x_1, x_2]^T) = [1, x_1, x_2, x_1x_2, x_1^2x_2^2]^T$

[Ans: c]

2025 Dec 18 ET

17. For the polynomial kernel $K(\mathbf{f}, \mathbf{g}) = (\mathbf{f}^T\mathbf{g} + 1)^2$, the explicit feature map for the vector $\mathbf{f} = (1, 2, -1)$ is 10-dimensional. Which of the following options is the correct transformed feature vector $\phi(\mathbf{f})$?

(a) $(1, 1, 2, -1, 1, 4, 1, 2, -1, -2)$

(b) $(1, \sqrt{2}, 2\sqrt{2}, -\sqrt{2}, 1, 4, 1, 2\sqrt{2}, -\sqrt{2}, -2\sqrt{2})$

(c) $(1, 1, 4, 1, 2\sqrt{2}, -\sqrt{2}, -2\sqrt{2}, 0, 0, 0)$

(d) $(1, \sqrt{2}, \sqrt{2}, \sqrt{2}, 1, 1, 1, 1, 1, 1)$

[Ans: b]

Oct 2022 Quiz 1

18. A function k is defined as

$$k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \qquad k(x_1, x_2) = x_1 + x_1^2 x_2^2 + x_1 x_2 + 1.$$

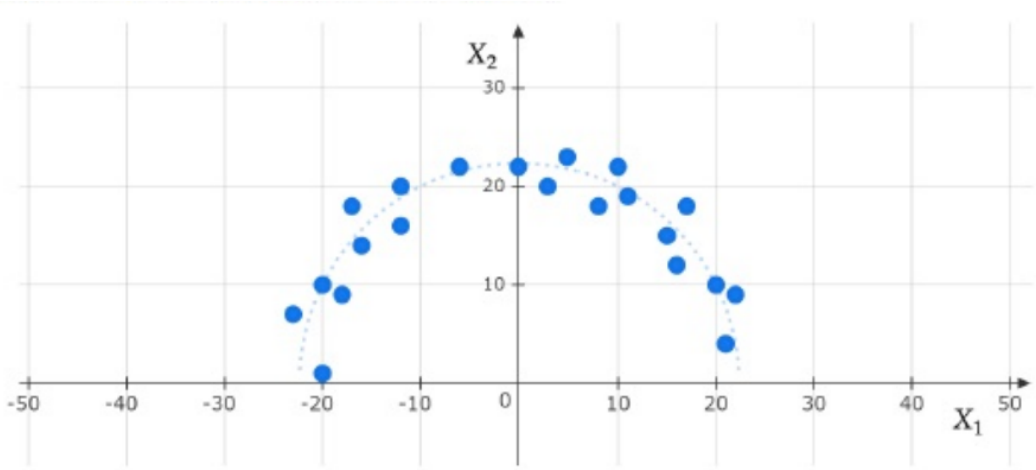
Is k a valid kernel?

(a) Yes

(b) No

[Ans: b]

19. Consider the dataset as shown in the figure. We want to project the dataset into another feature space so that it lives in a linear subspace of a higher-dimensional space. We apply kernel PCA with a polynomial kernel of the appropriate degree to achieve the same. What will be the dimension of the transformed feature space?



[Ans: 6]

Jul 2023 Quiz 1

20. Which of the following statements about kernel PCA is/are always true? (MSQ)

(a) The number of principal components corresponding to nonzero eigenvalues obtained using kernel PCA on a d-dimensional dataset cannot be more than d.

(b) For every transformation mapping $\phi : \mathbb{R}^d \to \mathbb{R}^D$, there exists a valid kernel function.

(c) For two valid kernels $k_1, k_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $ak_1 + bk_2$ is a valid kernel function. Here both a and b are positive real numbers.

(d) The output of a valid kernel function can never be a negative real number.

[Ans: b, c]

21. Which of the following statements is/are true? (MSQ)

(a) $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $k(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}_1^T \mathbf{x}_2)^3$ is a valid kernel.

(b) $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $k(\mathbf{x}_1, \mathbf{x}_2) = -(\mathbf{x}_1^T \mathbf{x}_2)^3$ is a valid kernel.

(c) Nonzero eigenvalues of XX^T and X^TX are the same.

(d) Eigenvectors of XX^T are the same as eigenvectors of X^TX.

[Ans: a, c]

Oct 2023 Quiz 1

22. **Comprehension:**
Suppose that we have a dataset (data matrix) $X \in \mathbb{R}^{d \times n}$, where d denotes the number of features and n denotes the number of samples. Assume that $d \gg n$ and the data is centered. The representative line is given by the unit vector $\mathbf{w}$, pointing in the direction of the line. In all the cases, we maximize the objective function given below:

$$f(\mathbf{w}) = \mathbf{w}^T C \mathbf{w}, \qquad C = \frac{1}{n} X X^T.$$

Here C is a covariance matrix. The kernel matrix is defined to be $K = X^T X$. Based on the above data, answer the given subquestions.

i. Which of the following statements are true? (MSQ)

   (a) The eigenvectors of C and K are related by a scaling factor $r \in \mathbb{R}$.

   (b) Computing eigenvectors using the matrix K reduces the number of computations significantly compared with using the matrix C.

   (c) Suppose the maximum eigenvalue of C is λ_1, then the maximum eigenvalue of XX^T is $n\lambda_1$.

   (d) The eigenvectors of C are in $\mathbb{R}^d$ and those of K are in $\mathbb{R}^d$.

   (e) The eigenvalues of both C and K are non-negative.

   [Ans: b, c, e]

ii. Which of the following matrices helps in transforming the eigenvectors of K to the eigenvectors of C?

   (a) X

   (b) X^T

   (c) C^{-1}

   (d) K^{-1}

   (e) X^{-1}

   [Ans: a]

iii. Suppose that 95% of the variance of the data is captured using $d = 1$ eigenvectors. This implies that:

   (a) The features are highly correlated (or dependent).

   (b) The features have low correlation (or are independent).

   [Ans: b]

Oct 2025 Quiz 1

23. Which of the following are valid kernel functions? Select all that apply. (MSQ)

(a) $K(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}\right)$, $\sigma > 0$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(b) $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + c)^d$, $c \geq 0$, $d \in \mathbb{N}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(c) $K(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(d) $K(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|}{\sigma}\right)$, $\sigma > 0$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

[Ans: a, b, d]

Mar 2026 Quiz 1

24. Consider a polynomial kernel of degree 3 on a feature space of dimension 4. How many features will the transformed space have?

[Ans: 35]

25. Consider the Gaussian kernel, $k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, with $\sigma = 1$. Compute the value of $k((0, 1), (1, 0))$ and enter your answer correct to two decimal places.

[Ans: 0.37]

26. If $k_1 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ and $k_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ are valid kernels, then $2k_1 + 5k_2$ is also a valid kernel.

(a) True

(b) False

[Ans: a]

27. If X is a data matrix of shape $d \times n$, $X^T X$ and XX^T have the same eigenvectors.

(a) True

(b) False

[Ans: b]

Jul 2026 Quiz 1

28. Consider the following kernel function defined for scalar inputs $x_1, x_2 \in \mathbb{R}$:

$$K(x_1, x_2) = 10x_1^2 x_2^2.$$

Which of the following feature maps $\phi : \mathbb{R} \to \mathbb{R}$ satisfies $K(x_1, x_2) = \phi(x_1)^T \phi(x_2)$ for all $x_1, x_2 \in \mathbb{R}$?

(a) $\phi(x) = 10x^2$

(b) $\phi(x) = x^2$

(c) $\phi(x) = \sqrt{10}\, x^2$

(d) $\phi(x) = \sqrt{10x^2}$

[Ans: d]

29. Assume that

$$K_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}, \qquad K_2(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2}\right)$$

are valid (positive semi-definite) kernels. Which of the following are also valid kernels? (MSQ)

(a) $\frac{K_2(\mathbf{x},\mathbf{y})}{\sqrt{K_2(\mathbf{x},\mathbf{x})}\sqrt{K_2(\mathbf{y},\mathbf{y})}} \cdot \mathbf{x}^T\mathbf{y}$

(b) $\mathbf{x}^T\mathbf{y} + \exp\left(-\frac{\|\mathbf{x}\|^2+\|\mathbf{y}\|^2}{2\sigma^2}\right) \exp\left(\frac{\mathbf{x}^T\mathbf{y}}{\sigma^2}\right)$

(c) $(K_1(\mathbf{x}, \mathbf{y}) + K_2(\mathbf{x}, \mathbf{y}))^2$

(d) $K_1(\mathbf{x}, \mathbf{y}) - K_2(\mathbf{x}, \mathbf{y})$

[Ans: a, b, c]

Dec 2022 End Term

30. Let $k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a valid kernel. Is $x_1 x_2 k(x_1, x_2)$ a valid kernel? Here $x_1, x_2 \in \mathbb{R}$.

(a) Yes

(b) No

[Ans: a]

Apr 2023 End Term

31. A kernel k is defined as

$$k : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},$$

$$k([x_1, x_2]^T, [y_1, y_2]^T) = 1 + x_1 y_1 + x_1^2 y_1^2 + x_2^2 y_2^2.$$

Which of the following transformation mappings corresponds to this kernel function?

(a) $\phi([x_1, x_2]^T) = [1, x_1^2 + x_2^2]^T$

(b) $\phi([x_1, x_2]^T) = [1, x_1 + x_1^2 + x_2^2]^T$

(c) $\phi([x_1, x_2]^T) = [1, x_1^2, x_2^2]^T$

(d) $\phi([x_1, x_2]^T) = [1, x_1, x_1^2, x_2^2]^T$

[Ans: d]

Sep 2023 End Term

32. Which of the following matrices can be expressed as $K = X^T X$ for some data matrix X? Choose the most appropriate answer. (MSQ)

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

(b) $\begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 2 \\ 2 & 4 \end{bmatrix}$

[Ans: a, b]

Dec 2023 End Term

33. A team was given a dataset $X \in \mathbb{R}^{d \times n}$ where d denotes the number of features and n denotes the number of samples. They found that there are 10 samples in the dataset and each sample contains 100 features. Assume that the data points $\mathbf{x}_4$ to $\mathbf{x}_{10}$ are all linear combinations of the linearly independent data points $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$.

Suppose the team applies kernel PCA by computing the kernel matrix $K = X^T X$ on the dataset. Choose all the correct statements. (MSQ)

(a) The principal directions given by kernel PCA are the same as those given by linear PCA.

(b) Kernel PCA takes fewer computations than linear PCA to find the principal directions.

(c) The eigenvectors of the kernel matrix K are pointing in the same direction as the eigenvectors of the covariance matrix C.

(d) Kernel PCA takes more computations than linear PCA to find the principal directions.

[Ans: a, b]