Consider a training dataset with $m$ samples in $R^d$ space. Consider a kernel which projects any sample in $R^k$ space.

We would like to find out the Kernel SVM prediction on a test sample. What is the time complexity of this prediction

(A) $O(d)$

(B) $O(k)$

(C) $O(m)$

(D) **[Ans]** $O(md)$

(E) $O(mk)$

(F) $O(mdk)$

(G) None of these

Consider the primal form of soft margin SVM:

$$min_w \frac{1}{2} w^T w + C \sum_i \xi_i$$

subject to $y_i(w^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0 \forall i.$

Now find its dual form as a function of $\alpha, x, y$

(A) **[Ans]** $w = \sum_i \alpha_i x_i y_i$ similar to hard margin SVM

(B) **[Ans]** $\sum_i x_i y_i = 0$ similar to hard margin SVM

(C) **[Ans]** Dual form objective remains same as hard margin SVM

(D) Dual form constraints remain same as hard margin SVM

(E) None of these

What does the kernel $K(x, x') = \frac{x^T x'}{\|x\| \|x'\|}$ do?

(A) Project all points on the line $Y = X$

(B) Project all points on a line

(C) **[Ans]** Project all points on a unit circle

(D) Project all points on a non-unit circle

(E) None of these

Consider the dataset $\mathcal{D}(\mathcal{X}, \mathcal{Y})$:

$\mathcal{X} = \{[1, 0]^T, [0, 1]^T, [.5, .5]^T, [1, 1]^T\}$, $\mathcal{Y} = \{+1, +1, -1, -1\}$

We use the kernel $K(x, x') = \frac{x^T x'}{\|x\| \|x'\|}$ to peform kernel SVM (hard-margin)

(A) $\mathcal{D}$ is linearly separable in the original feature space

(B) **[Ans]** $\mathcal{D}$ is linearly separable in the new feature space

(C) We see 4 unique points in the new feature space

(D) If wee add the point $\{[2, 2]^T, 1\}$ to $\mathcal{D}$, then $\mathcal{D}$ is linearly separable in the new feature space

(E) **[Ans]** If wee add the point $\{[2, 2]^T, -1\}$ to $\mathcal{D}$, then $\mathcal{D}$ is linearly separable in the new feature space

(F) None of these

Consider the dataset $\mathcal{D}$:

$\mathcal{X} = \{[1,0]^T, [0,1]^T, [.5,.5]^T, [1,1]^T\}$, $\mathcal{Y} = \{+1,+1,-1,-1\}$

We use the kernel $K(x,x') = \frac{x^T x'}{\|x\|\|x'\|}$ to peform kernel SVM (hard-margin). Then the decision boundary in the new feature space can be written as $w_1 x_1 + w_2 x_2 + w_3 = 0$

(A) $w_1 = 1, w_2 = -1$

(B) **[Ans]** $w_1 = 1, w_2 = 1$

(C) $w_1 = 1, w_2 = 0$

(D) $w_1 = 0, w_2 = 1$

(E) $w_1 = 0, w_2 = 0$

(F) None of these