See the Sec 3.3 (equations)

Why is $a \in R^+$ written there?

- (A) It could have been $a \in R^-$
- (B) [Ans] a negative does not lead to K being PSD
- (C) It is a typo.

See the Sec 3.3. Assume $\alpha_i \in R^+$; $\beta_i \in R^+$; $\kappa_i(\mathbf{p}, \mathbf{q})$ being a valid kernel. Also Kand L are some positive integers.

Then a new kernel $\kappa(\cdot, \cdot) =$

(A) [Ans]
$$\sum_{i=1}^{K} \kappa_i(\mathbf{p}, \mathbf{q})$$
 is a valid kernel.

(B) [Ans]
$$\prod_{i=1}^{L} \kappa_i(\mathbf{p}, \mathbf{q})$$
 is a valid kernel.

(C) [Ans]
$$\sum_{i=1}^{K} \alpha_i \kappa_i(\mathbf{p}, \mathbf{q})$$
 is a valid kernel.

(C) [Ans]
$$\sum_{i=1}^{n} \alpha_i \kappa_i(\mathbf{p}, \mathbf{q})$$
 is a valid kernel.

(D) [Ans] $\prod_{i=1}^{L} \beta_i \kappa_i(\mathbf{p}, \mathbf{q})$ is a valid kernel.

) [Ans]
$$\prod_{i=1}^{K} \beta_i \kappa_i(\mathbf{p}, \mathbf{q})$$
 is a valid kernel.
) [Ans] $\sum_{i=1}^{K} \alpha_i \kappa_i(\mathbf{p}, \mathbf{q}) + \prod_{i=1}^{L} \beta_i \kappa_i(\mathbf{p}, \mathbf{q})$ is a valid kernel.

(E) [Ans] $\sum_{i=1}^{K} \alpha_i \kappa_i(\mathbf{p}, \mathbf{q}) + \prod_{i=1}^{L} \beta_i \kappa_i(\mathbf{p}, \mathbf{q})$ is a valid kernel

E) **[Ans]**
$$\sum_{i=1}^K \alpha_i \kappa_i(\mathbf{p}, \mathbf{q}) + \prod_{i=1}^L \beta_i \kappa_i(\mathbf{p}, \mathbf{q})$$
 is a valid kerne

 $(\mathbf{x}^T\mathbf{y})^2$ (A) The initialization $\alpha_i = 0$ is a must. With no other initialization, this algorithm will not work (say will not converge)

See the pseudo-code for Kernel Perceptron (Algorithm 3). Assime the kernel to be

- (B) Step of computing Kernel Matrix (step 2) should have been in side the loop (repeat structure).
- (C) Since this is now Kernelized, with any data (irrespective of whether the data is linearly separable or not), this algorithm will converge.
 - (D) For data that is linearly separable, this algorithm will give you a linear decision boundary.
 - boundary.
 (E) [Ans] None of the above.

Look at the equation (94) related to the objective function:

(A) [Ans] This is an L1 softmargin SVM

(B) This is an L2 softmargin SVM

- (C) There is a typo. ξ_i should be replaced as ξ_i^2
- (D) There is a typo. LHS will have to be (E) None of the above.

Consider the decision making rule. "one side of a line (in 2D) is +ve class and other side of a line is -ve class" Figure 7 shows that VC dimension of a class of functions (lines) in 2D is 3.

What is the VC dimension in 1D for a function class. If $x > \theta$, positive, else negative.

Write your answer in the space provided.

(Sample answer (possibly incorrect): 1)