For Support Vector Machines:

(A) A hard margin linear SVM yields a valid (constraint satisfying) solution for any data.

(B) If the data is linearly separable, linear SVM and linear perceptron will give the same solution.

(C) maximization of $\mathbf{w}^T\mathbf{w}$ with the associated constraints is leading to the maximization of the margin.

(D) **[Ans]** minimization of $\mathbf{w}^T\mathbf{w}$ with the associated constraints is leading to the maximization of the margin.

(E) None of the above.

Kernel SVMs are:

(A) Always hard margin

(B) Always soft margin

(C) **[Ans]** Can be either hard margin or soft margin

(D) If there is a hard margin feasible, soft-margin K-SVMs eventually finds this.

(E) None of the above

In the context of Softmargin Linear SVMs:

(A) If there is a hard margin feasible, soft-margin linear SVMs always finds this.

(B) **[Ans]** If there is a hard margin feasible, soft-margin linear SVMs can finds this with certain 'C'.

(C) If the data is linearly separable, we should not use soft margin SVM.

(D) **[Ans]** The larger the C, the formulation become closer and closer to hard margin SVM.

(E) None of the above.

Number of Support Vectors:

(A) is $2 \times d$

(B) **[Ans]** depends on the kernel we use.

(C) **[Ans]** can be as large as N

(D) can be as small as 1

(E) None of the above.

In the context of K-SVM:

(A) Formulation change with kernel.

(B) **[Ans]** Performance of the solution change with kernel.

(C) We can use different Kernels during training and testing.

(D) **[Ans]** If the kernel $\kappa(x, y) = x^T y$, then K-SVM is equivalent to Linear SVM

(E) All the above four are true.