

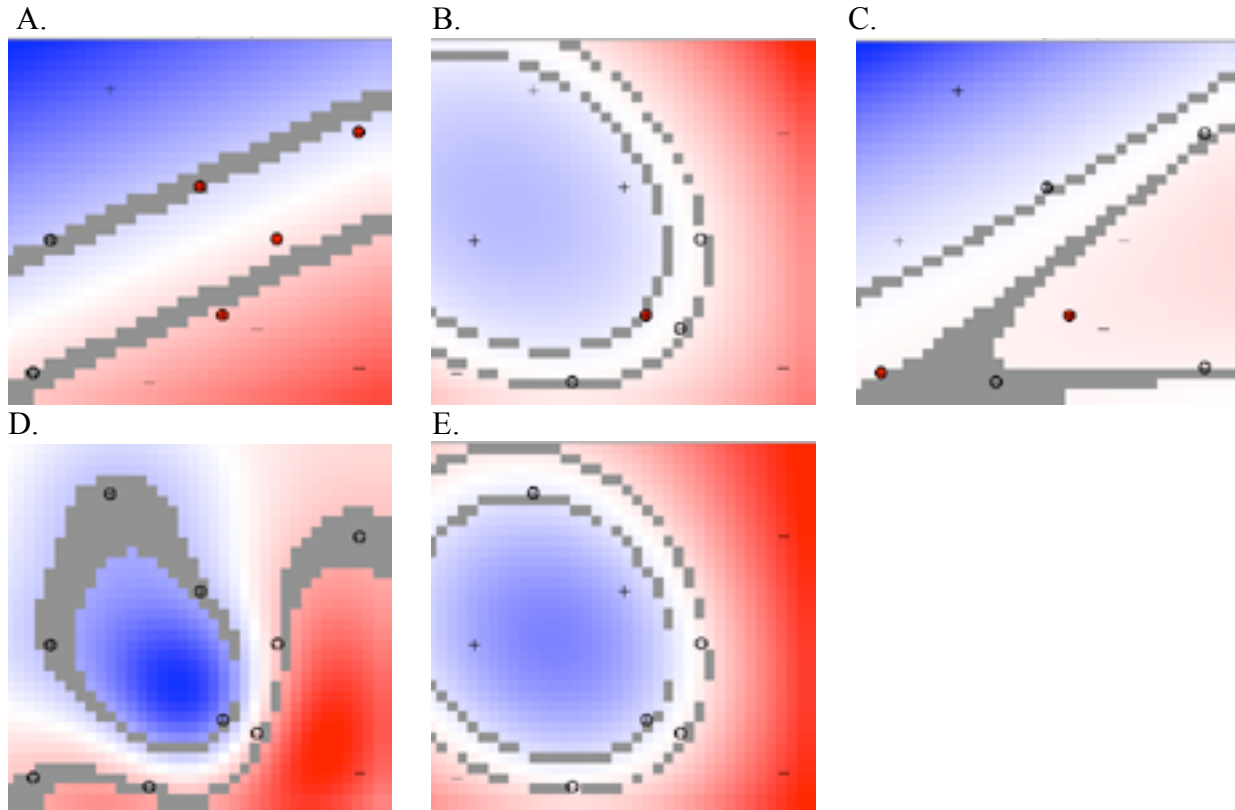
Final Exam 2002 Problem 6: Support Vector Machines (14 Points)

Part A: (2 Points)

The following diagrams represent graphs of support vector machines trained to separate pluses (+) from minuses (-) for the same data set. The origin is at the lower left corner in all diagrams. Which represents the best classifier for the training data? *See the separate color sheet for a clearer view of these diagrams.*

Indicate your choice here:

E gives proper values to all points. So does D, but at the risk of overfitting.



Part B: (5 Points)

Match the diagrams in Part 1 with the following kernels:

- Radial basis function, sigma .08
- Radial basis function, sigma .5
- Radial basis function, sigma 2.0
- Linear
- Second order polynomial

Part C: (3 Points)

Order the following diagrams from *smallest* support vector weights to *largest* support vector weights, assuming all diagrams are produced by the same mechanism using a linear kernel (that is, there is no transformation from the dot-product space).

The origin is at the lower left corner in all diagrams. Support vector weights are also referred to as α_i values or LaGrangian multipliers. *See the separate color sheet for a clearer view of these diagrams.*

Smallest

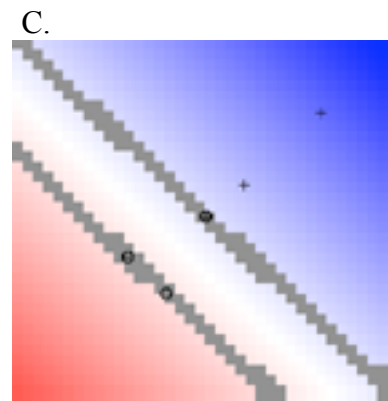
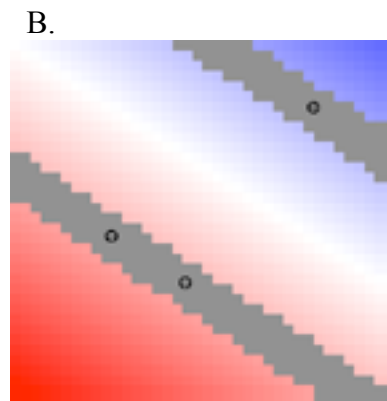
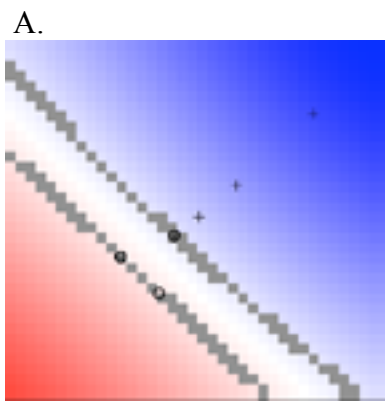
B

Medium

C

Largest

A



Part D (4 Points)

Suppose a support vector machine for separating pluses from minuses finds a plus support vector at the point $\mathbf{x}_1 = (1, 0)$, a minus support vector at $\mathbf{x}_2 = (0, 1)$.

You are to determine values for the classification vector \mathbf{w} and the threshold value b . Your expression for \mathbf{w} may contain \mathbf{x}_1 and \mathbf{x}_2 because those are vectors with known components, but you are not to include any α_i or y_i . Hint: think about the values produced by the decision rule for the support vectors, \mathbf{x}_1 and \mathbf{x}_2 .

\mathbf{w}

b

The decision boundary goes through the origin, so $b = 0$.

We know $\mathbf{w} = (a_1 \mathbf{x}_1 - a_2 \mathbf{x}_2)$, and \mathbf{x}_1 is a support vector so, $\mathbf{w} \cdot \mathbf{x}_1 = 1$. Substituting for \mathbf{w} , we get $\mathbf{w} \cdot \mathbf{x}_1 = (a_1 \mathbf{x}_1 - a_2 \mathbf{x}_2) \cdot \mathbf{x}_1 = a_1 \mathbf{x}_1 \cdot \mathbf{x}_1 - a_2 \mathbf{x}_2 \cdot \mathbf{x}_1 = a_1$ because \mathbf{x}_1 and \mathbf{x}_2 are orthogonal and \mathbf{x}_1 is a unit vector. Hence $\mathbf{w} \cdot \mathbf{x}_1 = a_1 = 1$. Same reasoning yields $a_2 = 1$. So $\mathbf{w} = (a_1 \mathbf{x}_1 - a_2 \mathbf{x}_2) = \mathbf{x}_1 - \mathbf{x}_2 = [1 -1]$.