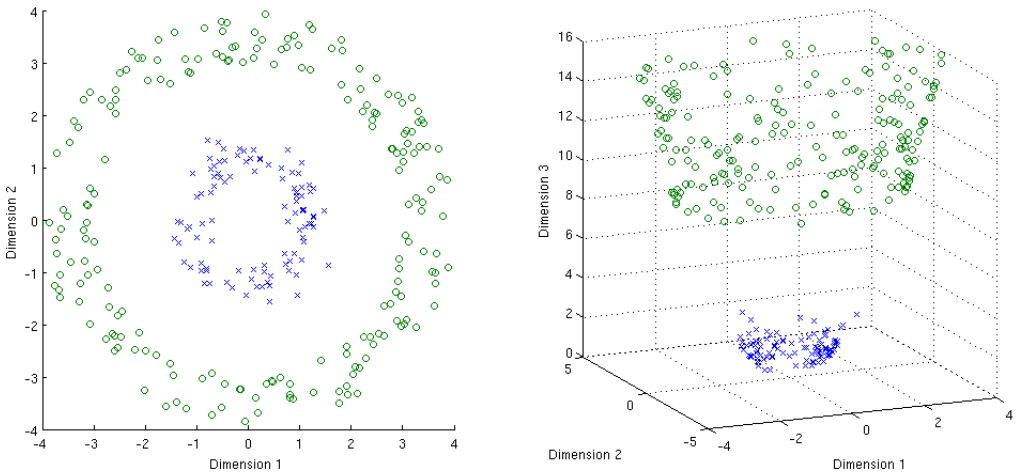
4. Kernels-II

In this question, we will practice some specific kernel methods.

4. (a)

2/2 points (graded)

In the figure below, a set of points in 2-D is shown on the left. On the right, the same points are shown mapped to a 3-D space via some transform $\phi(x)$, where x denotes a point in the 2-D space. Notice that $\phi(x)_1 = x_1$ and $\phi(x)_2 = x_2$, or in other words, the first and second coordinates are unchanged by the transformation.



Which of the following functions could have been used to compute the value of the 3rd coordinate, $\phi(x)_3$ for each point?

- $\phi(x)_3 = x_1 + x_2$
- $\phi(x)_3 = x_1^2 + x_2^2$ ✔
- $\phi(x)_3 = x_1 x_2$
- $\phi(x)_3 = x_1^2 - x_2^2$

Think about how a linear decision boundary in the 3 dimensional space ($\{\phi \in \mathbb{R}^3 : \theta \cdot \phi + \theta_0 = 0\}$) might appear in the original 2 dimensional space.

For example, suppose the decision boundary in the 3 dimensional space is $z = 4$.

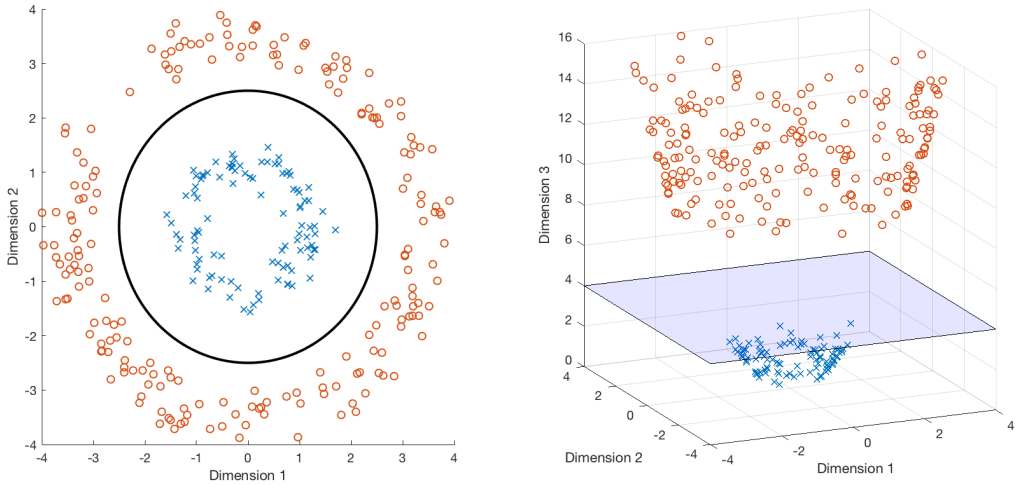
Provide an equation $f(x_1, x_2) = 0$ in the 2 dimensional space such that all the points (x_1, x_2) with $f(x_1, x_2) > 0$ correspond to $z > 4$ in the 3 dimensional space.

$f(x_1, x_2) = 0 =$ x_1^2+x_2^2-4 ✔ **Answer:** x_1^2 + x_2^2 - 4

$x_1^2 + x_2^2 - 4$

Solution:

- With $x = [x_1; x_2]$, one mapping which could satisfy the mapping is $\phi(x)_3 = x_1^2 + x_2^2$. The decision boundary is shown below.
- As a result, the decision boundary at $z = 4$ corresponds to $x_1^2 + x_2^2 = 4$



Submit You have used 1 of 3 attempts

Answers are displayed within the problem

4. (b)

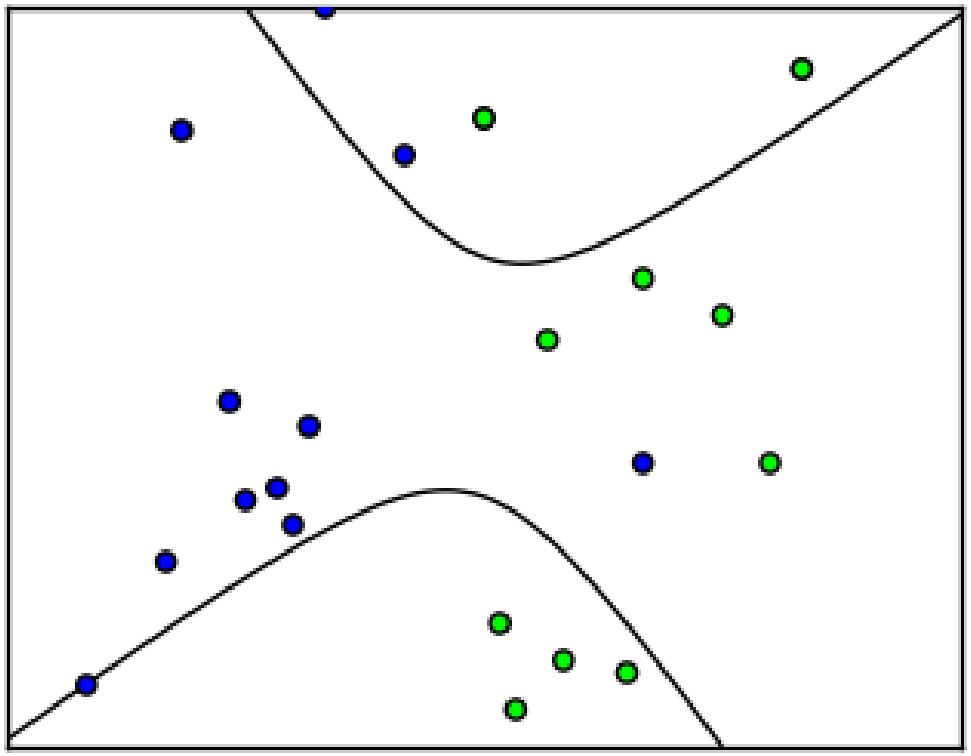
5/5 points (graded)

Consider fitting a kernelized SVM to a dataset $(x^{(i)}, y^{(i)})$ where $x^{(i)} \in \mathbb{R}^2$ and $y^{(i)} \in \{1, -1\}$ for all $i = 1, \ldots, n$. To fit the parameters of this model, one computes θ and θ_0 to minimize the following objective:

$$L(\theta, \theta_0) = \frac{1}{n} \sum_{i=1}^{n} \text{Loss}_h \left(y^{(i)} \left(\theta \cdot \phi \left(x^{(i)} \right) + \theta_0 \right) \right) + \frac{\lambda}{2} \|\theta\|^2$$

where ϕ is the feature vector associated with the kernel function. Note that, in a kernel method, the optimization problem for training would be typically expressed solely in terms of the kernel function $K(x, x')$ (dual) rather than using the associated feature vectors $\phi(x)$ (primal). We use the primal only to highlight the classification problem solved.

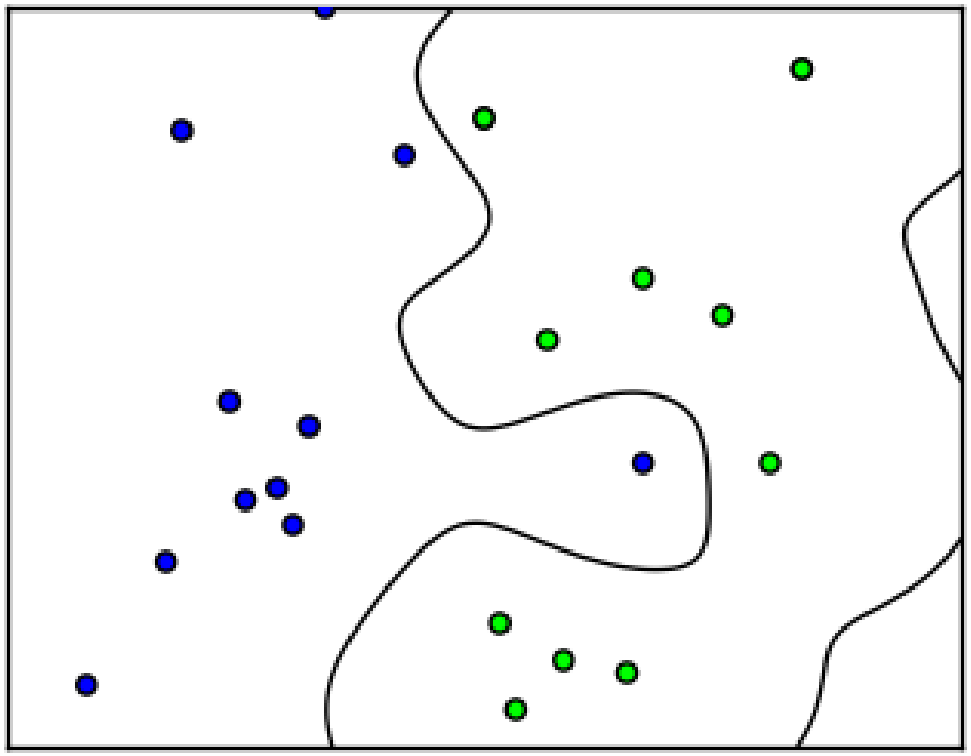
The plots below show 4 different kernelized SVM models estimated from the same 11 data points. We used a different kernel to obtain each plot but got confused about which plot corresponds to which kernel. Help us out by assigning each plot to one of the following models: linear kernel, quadratic kernel, order 3 kernel, and RBF kernel.



Which kernel is used in the above model?

- linear kernel

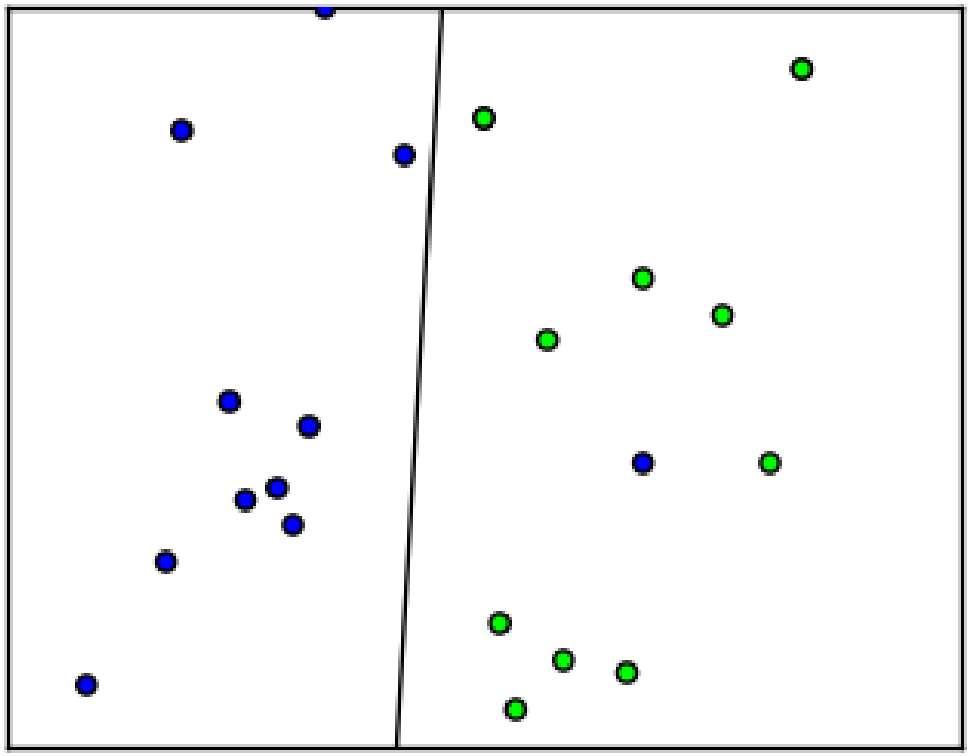
- (•) quadratic kernel ✔
- ( ) order 3 kernel
- ( ) RBF kernel



Which kernel is used in the above model?

- linear kernel

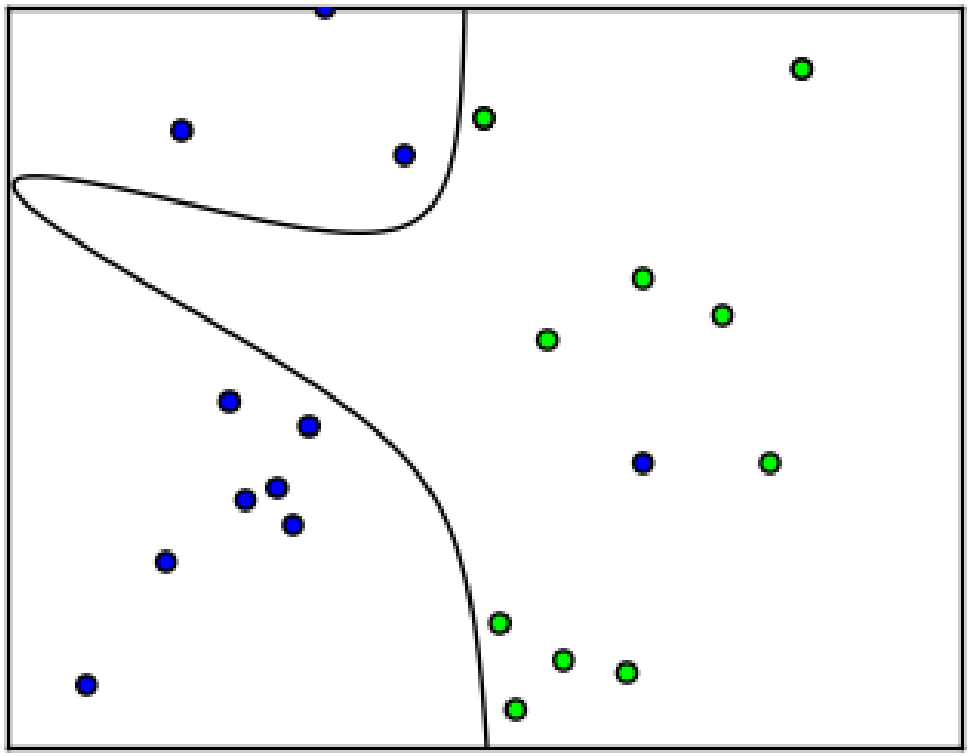
- quadratic kernel
- order 3 kernel
- RBF kernel ✔



Which kernel is used in the above model?

- linear kernel ✔

- quadratic kernel
- order 3 kernel
- RBF kernel



Which kernel is used in the above model?

- linear kernel

- quadratic kernel
- order 3 kernel ✔
- RBF kernel

How would you describe qualitatively how the resulting classifiers vary with the value of λ? If the value of λ is increased, the fitting of model would be

- better fit on training data (sharper decision boundary)
- worse fit on training data (flatter decision boundary) ✔

Solution:

- From examining the number of bends in the decision boundaries:
- 3rd plot corresponds to the linear kernel.
- 1st plot corresponds to the quadratic kernel.
- 4th plot corresponds to the 3rd-order kernel.

- 2nd plot corresponds to the Gaussian RBF kernel.
- Large λ penalty on θ results in flatter/ less “squiggly" lines.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Homework 3 / 4. Kernels-II

Add a Post

Show all posts ▼ by recent activity ▼

- [Staff] Misunderstood the deadline for HM3 — 3
  Hello everyone, I thought that the deadline was till 00:00 today and I just noticed that I can't submit anymore. I would really appreciate it if is pos...
- Plane in 3D vs. circle in 2D — 4
  Hi all, I am slightly confused about the answer given to the second part of 4(a). As it is worded, it seems a little like it is saying that a circle in 2D w...
- Explanations for the graphs of the quadratic kernel and cubic Kernel? — 3
- 4a - 2nd part — 11
  I thought that was as simple as taking the equation from the first part (which I got right) and add a number so that every x1,x2 point would be ab...
- [STAFF] 4b Datasets

I found the right answers by guessing but I have no understanding of how it works. Can you provide us the numerical datasets of these example... 2

Understanding RBF 2

Very disappointing that people are trying to cheat at this graduate level class from MITx
how will they do in their job even if they successfully cheat here? very very immature thinking... 2

4b: Matching the graphs. I guessed them correctly. But I don't really understand why they are right .
I'm mostly confused about the quadratic kernel, since it doesn't look like a "normal" quadratic graph. I still don't understand why the graph matc... 4

How can I understand order 3 kernel
I checked the property of quadratic, RBF and linear but I cannot understand how to identify the order 3. Is it because the boundary is pushed aw... 3

further explanation needed on f(x_1, x_2)=0 7

[STAFF] [URGENT] It's rather disappointing that someone tried to cheat their way through :/
So as I search for the word "order 3" and then go to google images to have a visual image of what order 3 is. I stumbled upon an image of the ex... 9

Finished Question but No Green Check
Title says it all - I answered everything but still don't have a green check at the top for this question. Could someone from the staff please confir... 2

[staff] Some wording / expression cleanup for 4b
Community TA 2

4(b)