Trigonometric basis

7. (14 points) We can define a non-linear feature transformation using trigonometric functions. In this question, we will consider only the case in which our original $x^{(i)}$ are in $\mathbb{R}$ (that is, the input dimension $d = 1$.)

   Define the kth order trigonometric basis feature transformation to be

   $$\phi(x) = (1, \sin(x), \cos(x), \sin(2x), \cos(2x), \sin(x/2), \cos(x/2), \ldots, \sin(kx), \cos(kx), \sin(x/k), \cos(x/k)) \quad .$$

   (a) Sindy thinks that this basis is missing an important aspect, and suggests that it would be useful to add to the feature vector components of the form

   $$j \sin(x), j \cos(x)$$

   for values of j from 2 to k.

   Cosima thinks that Sindy's suggestion won't add any expressive power (that is, that any function that could be represented using Sindy's basis can also be represented using the original one.)

   Who is right? √ **Cosima** ◯ Sindy

   Let h be a hypothesis written in terms of Sindy's basis of order 2:

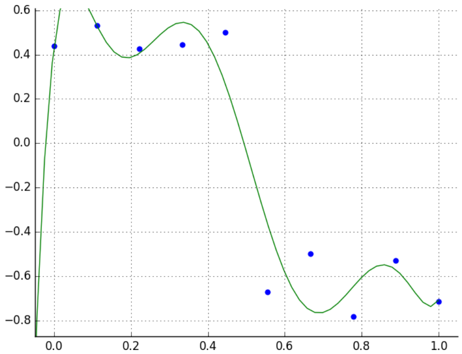
   $$\begin{aligned} h(x) = &\theta_0 + \theta_1 \sin(x) + \theta_2 \cos(x) + \\ &\theta_3 \sin(2x) + \theta_4 \cos(2x) + \theta_5 \sin(x/2) + \theta_6 \cos(x/2) + \theta_7 2 \sin(x) + \theta_8 2 \cos(x) \quad . \end{aligned}$$

   If Cosima is right, show how to describe this hypothesis in terms of the original trigonometric basis of order 2, using parameters $\theta_0, \ldots, \theta_8$.

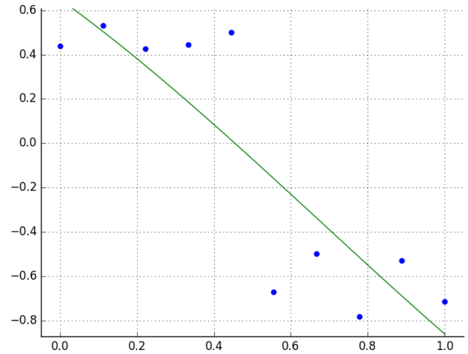
   Solution:

   $$\begin{aligned} h(x) = &\theta_0 + (\theta_1 + 2\theta_7) \sin(x) + (\theta_2 + 2\theta_8) \cos(x) + \\ &\theta_3 \sin(2x) + \theta_4 \cos(2x) + \theta_5 \sin(x/2) + \theta_6 \cos(x/2) \end{aligned}$$

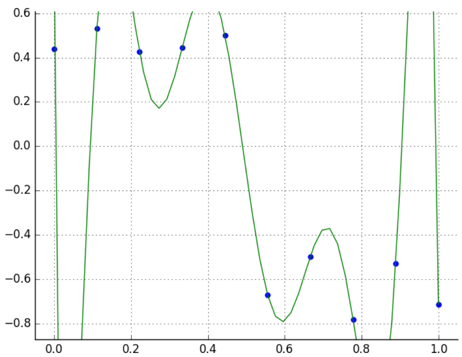
   (b) We used the trigonometric basis, for several different values of k to transform the input to a new feature space, performed linear regression, and obtained the following plots of $h(x)$ versus x. For each plot, provide the correct k value, chosen from the set $0, 1, 2, 3, 7$.



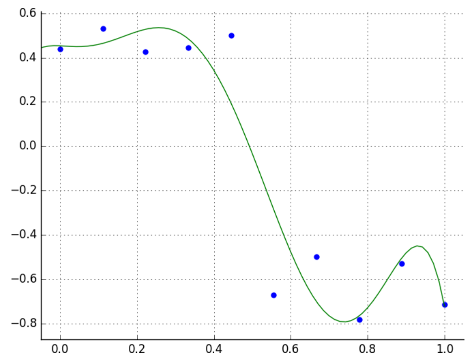
(A) k: **3**



(B) k: **1**



(C) k: **7**



(D) k: **2**

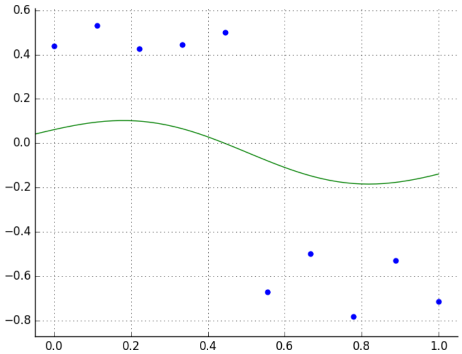
(c) Which hypotheses from the plots above do you think would lead to the best test-set performance on the range of x values between 0 and 1 (the range that is plotted)?

$\surd$ **A** ◯ B ◯ C $\surd$ **D**

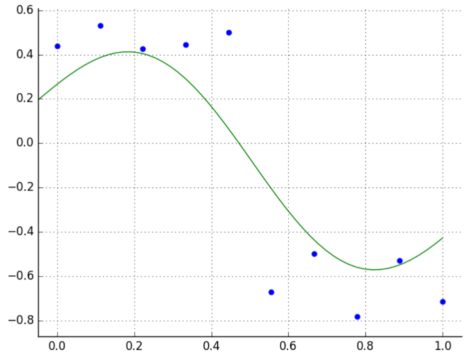
We accepted A, D, or both.

(d) We also used the trigonometric basis on the same data with a large fixed value of k, but performed ridge regression with various values of regularization parameter λ. For each plot, provide the correct λ value, chosen from the set $(0.0, 10^{-7}, 1e1, 1e2)$

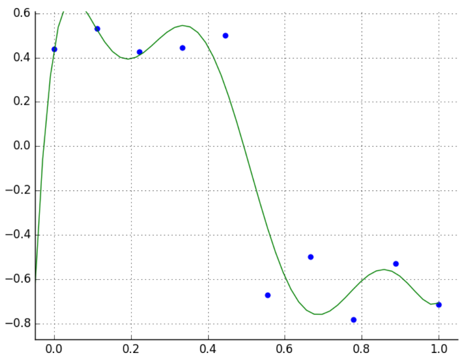
Name: ____________________



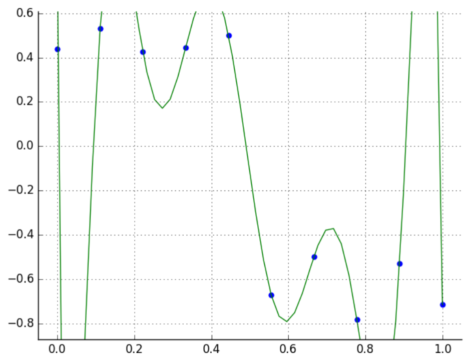
λ: **1e2**



λ: **1e1**



λ: **1e-7**



λ: **0.0**