Homework # 6

1) Computer the Jacobian matrix for the following:

- (15pts)

$$F(x,y) = \begin{pmatrix} x^2 + sin(x) \\ x(y-2) \\ y^2 - 3xy \end{pmatrix} \neq xy - 2x$$

- (15pts)

$$G(x,y,z) = \begin{pmatrix} x^2 - y^2 \\ 3xyz - 5 \end{pmatrix}$$

Jacobian matrix   we know

$$\begin{matrix} \frac{df_1}{dx} & \frac{df_1}{dy} & \cdots \\ \frac{df_2}{dx} & \frac{df_2}{dy} \\ \frac{df_3}{dx} & \frac{df_3}{dy} \end{matrix}$$

1) $$F(x,y) = \begin{bmatrix} 2x + \cos(x) & 0 \\ y-2 & x \\ -3y & 2y-3x \end{bmatrix}$$

Course: Introduction to Machine Learning
Assignment 5
Important NOTE: In order to get full credit, for every question, you need to provide the details of your work on how to get to a solution or the end of a proof.
Instructor: Tan Bui-Thanh
TA: Cole Nockolds
Due date: 11:00 pm, Tuesday, November 19
All submission MUST be in pdf format, except codes. All code outputs should be reported in pdf format.

1 Question 1: (40 points) (Kernel SVM with scikit-learn)

Based on the given code below, generate a new training set of 20 samples. Using scikit-learn, plot the decision boundaries for the binary classification of the data with the following approaches:

1. Polynomial kernel (use the entire dataset): You will need to produce six plots for the six unique pairs of degree and γ (γ is an input to the scikit-learn funtion):
   - Degree = 2, 5.
   - $\gamma = 1, 2, 5$.

   Discuss the significance of γ.

2. Radial basic kernel:
$$K(\boldsymbol{x}, \boldsymbol{z}) = \exp\left(-\frac{||\boldsymbol{x} - \boldsymbol{z}||^2}{2\sigma^2}\right).$$
   You will need to produce six plots for the six unique pairs of σ and number of samples:
   - $\sigma = 0.2, 1.0, 2.0$
   - Training data = 10 samples, 20 samples.

2 (40 points) (Linear Regression with Kernel)

Given the model:
$$y = \boldsymbol{\theta}^T \boldsymbol{\phi}(x) + \epsilon.$$

1. Derive the following kernel form (including the form/equation for β_i) for the Ridge Regression approach.
$$y_{\text{test}} = \boldsymbol{\phi}(x_{\text{test}})^\top \boldsymbol{\theta} = \sum_{i=1}^{n} \beta_i \boldsymbol{\phi}(x_{\text{test}})^\top \boldsymbol{\phi}(x^{(i)}),$$

2. Starting from the given code below, write a short code using the derived kernel form, from part 1, with a polynomial kernel function (degree = 4). Plot the predicted function together with the true function and discuss the results.

3. Now, with the same derived kernel formula, change the basis function to the Gaussian basis function:

$$\phi_j(x) = \exp\left(-\frac{(x-u_j)^2}{2\sigma^2}\right), \quad j = 0, ..., 10$$

   Find the regression solution for the following case: $u_j = [0, 0.1, 0.2, ..., 1]$ and s = 1. Plot the predicted function together with the true function and discuss the results.

3 (20 points) (Kernel basis)

Let $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$, show that the Radial Basis Function (RBF; sometimes called Gaussian) kernel,

$$K(\boldsymbol{x}, \boldsymbol{z}) = \exp\left(-\frac{||\boldsymbol{x}-\boldsymbol{z}||_2^2}{2\sigma^2}\right),$$

is a Mercer kernel, where $||\cdot||_2$ denotes ℓ_2-norm (Euclidean norm) in $\mathbb{R}^n$.

2) $G(x,y,z) = \begin{bmatrix} 2x & -2y & 0 \\ 3yz & 3xz & 3xy \end{bmatrix}$

2) (15pts) Compute the Jacobian of the polar coordinates transformation

$$x = rcos\theta, y = rsin\theta$$

$$T(r,\theta) = \begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

Jacobian matrix: $\left| \frac{d(x,y)}{d(r,\theta)} \right| = \left| \begin{pmatrix} \frac{dx}{dr} & \frac{dx}{d\theta} \\ \frac{dy}{dr} & \frac{dy}{d\theta} \end{pmatrix} \right| = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$

Jacobian determinant: $\cos\theta(r\cos\theta) - (-r\sin\theta)(\sin\theta) \Rightarrow r\cos^2\theta + r\sin^2\theta \Rightarrow r(\cos^2 + \sin^2) = \boxed{r}$

3) (15pts) Derive the weak form of the 1D wave equation

$$\frac{\partial^2 u(x,t)}{\partial t^2} - c^2\,\nabla^2 u(x,t) = f(x,t)$$

where c is a scalar valued wave speed.

To get weak form, multiply by test function

Integrate by spatial derivative

IBP: $uv - \int v\,du$

We know:

$\frac{d^2 u(x,t)}{dt^2}$ is a temporal derivative because $\frac{du(x,t)}{dt}$ is in respect to time

$$\int v(x)\left(\frac{d^2 u(x,t)}{dt^2} - c^2\,\nabla^2 u(x,t)\right)dx = \int f(x,t)\,v(x)\,dx \Rightarrow$$

$$\int \frac{d^2 u(x,t)}{dt^2} v(x)\,dx - c^2 \int \nabla^2 u(x,t)\,v(x)\,dx = \int f(x,t)\,v(x)\,dx$$

$u = v(x)$ $\quad$ $v = \frac{du(x,t)}{dx}$ $\quad$ IBP $\Rightarrow$ $\quad$ $v(x)\cdot\frac{du(x,t)}{dx} - \int \frac{du(x,t)}{dx}\cdot\frac{dv(x)}{dx}\,dx$

$du = \frac{dv(x)}{dx}$ $\quad$ $dv = \frac{d^2u(x,t)}{dx^2}$

$\frac{d}{dx} = \nabla$

$$\int \frac{d^2u(x,t)}{dt^2}\, v(x)\,dx - c^2\left[v(x)\cdot \nabla u(x,t) - \int \nabla\cdot u(x,t)\cdot\nabla v(x)\,dx \right] = \int f(x,t)\, v(x)\,dx$$

$$\left[v(x)\,\nabla u(x,t) \right]_0^1 = 0$$

$$\int \frac{d^2u(x,t)}{dt^2}\, v(x)\,dx + c^2 \int \nabla\cdot u(x,t)\cdot\nabla v(x)\,dx = \int f(x,t)\, v(x)\,dx$$

This is after the Treyhan announcement

3) (20pts) Let $f(x) = e^x$ for $0 \leq x \leq 2$. Approximate $f(0.25)$ using Lagrangian linear interpolation with $x_0 = 0$ and $x_1 = 0.5$. Compare your answer to the real one.

We know that lagrangian polynomials have a $\frac{df^2}{dx} = 0$ @ desired points

$f(0) = e^{(0)} = 1 \qquad f(0.5) = e^{(0.5)} \approx 1.6487$

$$f(x) = f_1 + f_2 = y_0 \frac{(x - x_1)}{(x_0 - x_1)} + y_1 \frac{(x - x_0)}{(x_1 - x_0)} = 1 \frac{(x - 0.5)}{(0 - 0.5)} + 1.6487 \frac{(x - 0)}{(0.5 - 0)}$$

$$f(x) = -2x + 1 + 1.6487(2x) = 1.2974x + 1$$

$$f(0.25) = 1.2974(0.25) + 1 = 1.32435$$

$$\frac{1.32435 - 1.28402}{1.28402} = .0314 = 3.14\%$$

$$f(0.25) = e^{(0.25)} = 1.28402$$

$1.32435 - 1.28402 = 0.0433$ = 4 units above

3) (20pts) Let $P_3(x)$ by the Lagrangian interpolating polynomial for the data $(0,0)$, $(0.5,y)$, $(1,3)$ and $(2,2)$. Find y if the coefficient of the x^3 in $P_3(x)$ is 6.

$x_0 \quad x_1 \quad x_2 \quad x_3$

$$f(x) = \overset{f_1(x)}{y_0\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}} + \overset{f_2(x)}{y_1\frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}} + \overset{f_3(x)}{y_2\frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}}$$

$$+ \overset{f_3(x)}{y_3\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}}$$

$$f(x) = \cancelto{0}{\frac{0(x-0.5)(x-1)(x-2)}{(0-0.5)(0-1)(0-2)}} + y\frac{(x-0)(x-1)(x-2)}{(0.5-0)(0.5-3)(0.5-2)} + \frac{3(x-0)(x-1)(x-2)}{(1-0)(1-0.5)(1-2)}$$

$$+ \frac{2(x-0)(x-0.5)(x-1)}{(2-0)(2-0.5)(2-1)}$$

$$f(x) = \frac{1.5yx^3 - 3x^3 - 4.5yx^2 + 9.5625x^2 + 3yx - 6.5625x}{0.5625}$$

Coefficients of $x^3 = 6$, pull out x^3 solve for y

$$\frac{1.5yx^3 - 3x^3}{0.5625} = 6x^3 \Rightarrow \frac{1.5y - 3}{0.5625} = 6 \Rightarrow 1.5y - 3 = 3.375$$

$$\Rightarrow 1.5y = 6.375 \Rightarrow y = \frac{6.375}{1.5} = 4.25$$