Homework Assignment 4

Abhay Gupta (Andrew Id: abhayg)

October 25, 2018

1 Question 1

(a) With y(2) = 1, given differential equation is

$$\frac{dy}{dx} = \frac{1}{3y^2}$$
$$3y^2 dy = dx$$
$$\int 3y^2 dy = \int dx$$
$$y^3 = x + C$$

Putting, $y(2) = 1$ in the above equation, we get,

$$1 = 2 + C$$
$$C = -1$$

Thus the equation is $y = \sqrt[3]{x-1}$.

(b) Taking $y(2) = 1$ and a step size of $h = 0.05$, we get, the values in Table 1 and Figure 1. To compute the numerical correctness, we calculate the mean squared error of the computed error $e = y(x_i) - y_i$ vector. The error for Euler's estimation algorithm is 0.007045.

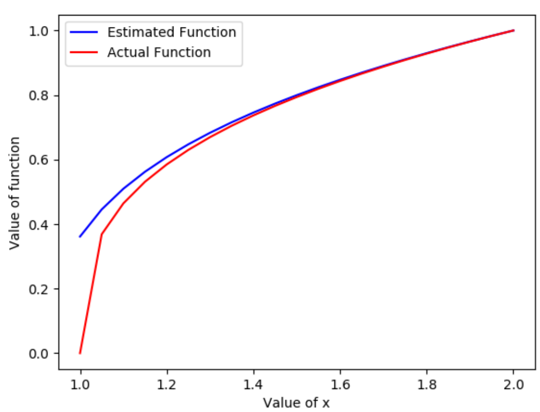

Figure 1: Euler's Estimation Graph

(c) Taking $y(2) = 1$ and a step size of $h = 0.05$, we get, the values in Table 2 and Figure 2. To compute the numerical correctness, we calculate the mean squared error of the computed error $e = y(x_i) - y_i$ vector. The error for RK4's estimation algorithm is 0.00053.

(d) Taking $y(2) = 1$ and a step size of $h = 0.05$, we get, the values in Table 3 and Figure 3. To compute the numerical correctness, we calculate the mean squared error of the computed error $e = y(x_i) - y_i$ vector. The error for RK4's estimation algorithm is 0.00244.

x	calculated $f(x)$	true $f(x)$	error
2.0	1.0	1.0	0.0
1.95	0.983	0.983	-0.00029
1.9	0.966	0.965	-0.00061
1.85	0.948	0.947	-0.00097
1.8	0.93	0.928	-0.00139
1.75	0.91	0.909	-0.00186
1.7	0.89	0.888	-0.00241
1.65	0.869	0.866	-0.00305
1.6	0.847	0.843	-0.0038
1.55	0.824	0.819	-0.00469
1.5	0.799	0.794	-0.00577
1.45	0.773	0.766	-0.00708
1.4	0.746	0.737	-0.00872
1.35	0.716	0.705	-0.01081
1.3	0.683	0.669	-0.01355
1.25	0.647	0.63	-0.0173
1.2	0.607	0.585	-0.02267
1.15	0.562	0.531	-0.03098
1.1	0.51	0.464	-0.04544
1.05	0.445	0.368	-0.07702
1.0	0.361	0.0	-0.36142

Table 1: Euler's Estimation Table

x	calculated $f(x)$	true $f(x)$	error
2.0	1.0	1.0	0.0
1.95	0.983	0.983	0.0
1.9	0.965	0.965	0.0
1.85	0.947	0.947	0.0
1.8	0.928	0.928	0.0
1.75	0.909	0.909	0.0
1.7	0.888	0.888	0.0
1.65	0.866	0.866	0.0
1.6	0.843	0.843	0.0
1.55	0.819	0.819	0.0
1.5	0.794	0.794	0.0
1.45	0.766	0.766	0.0
1.4	0.737	0.737	0.0
1.35	0.705	0.705	0.0
1.3	0.669	0.669	0.0
1.25	0.63	0.63	0.0
1.2	0.585	0.585	0.0
1.15	0.531	0.531	0.0
1.1	0.464	0.464	0.0
1.05	0.368	0.368	3e-05
1.0	0.103	0.0	-0.10346

Table 2: RK4 Estimation Table

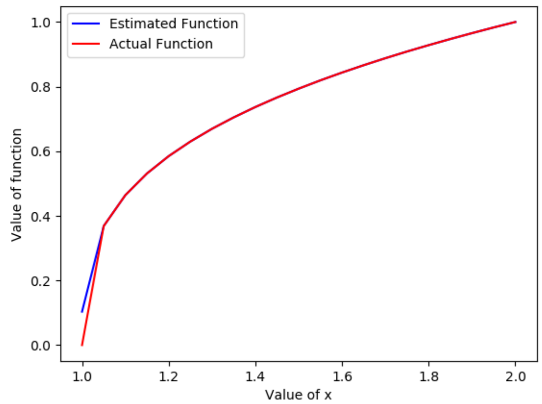

Figure 2: RK4 Estimation Graph

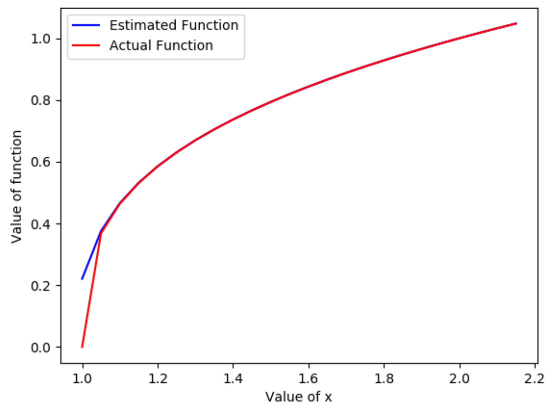

Figure 3: AB4 Estimation Graph

(e) In $\mathbb{R}$, the global minima of the function occurs at $x = 1$ and the function is monotonically increasing. The derivative of the function at $x = 1 \longrightarrow \infty$. However, when we compute the derivate at $1 + \Delta x$, we have a real-value for the derivative and this is the value we use to compute the estimated value for the function at $x = 1$, which causes the estimation to go off-track as all the methods assume there is somewhat of a continuity of values of the function being estimated and the difference of the derivatives of the function at two nearby points.

2 Question 2

$f(x, y) = x^3 + y^3 - 2x^2 + 3y^2 - 8$

(a)

$$\frac{\partial f}{\partial x} = 3x^2 - 4x$$
$$3x^2 - 4x = 0$$
$$x(3x - 4) = 0$$
$$x = 0, x = \frac{4}{3}$$

x	calculated $f(x)$	true $f(x)$	error
2.0	1.0	1.0	0.0
1.95	0.983	0.983	-0.0
1.9	0.965	0.965	-0.0
1.85	0.947	0.947	-0.0
1.8	0.928	0.928	-0.0
1.75	0.909	0.909	-0.0
1.7	0.888	0.888	-0.0
1.65	0.866	0.866	-1e-05
1.6	0.843	0.843	-1e-05
1.55	0.819	0.819	-1e-05
1.5	0.794	0.794	-2e-05
1.45	0.766	0.766	-3e-05
1.4	0.737	0.737	-4e-05
1.35	0.705	0.705	-6e-05
1.3	0.67	0.669	-0.0001
1.25	0.63	0.63	-0.00017
1.2	0.585	0.585	-0.00031
1.15	0.532	0.531	-0.00068
1.1	0.466	0.464	-0.00181
1.05	0.376	0.368	-0.00711
1.0	0.221	0.0	-0.22082

Table 3: AB4 Estimation Table

$$\frac{\partial f}{\partial y} = 3y^2 + 6y$$
$$3y^2 + 6y = 0$$
$$3y(y+2) = 0$$
$$y = 0, y = -2$$

Plotting the curves, we get, Figures 4 and 5

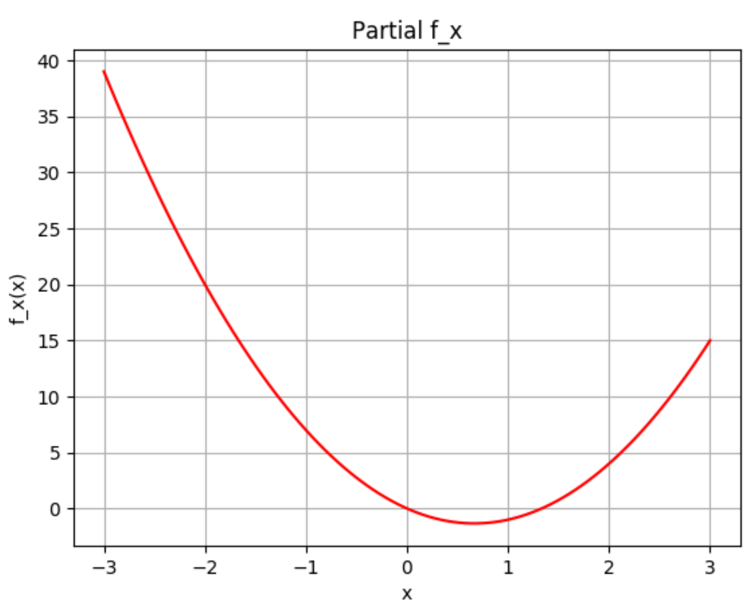

Figure 4: $f_x(x)$

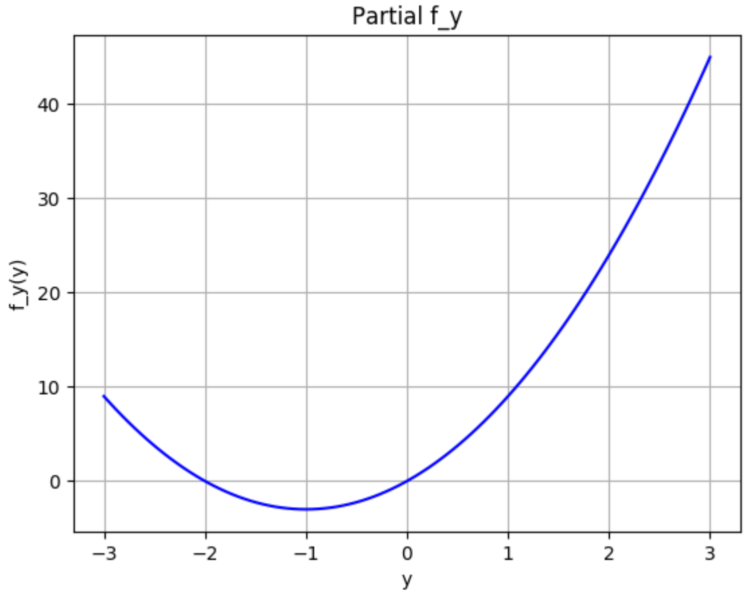

Figure 5: $f_y(y)$

Thus, the possible critical points are (0,0), (0,-2), (4/3,0) and (4/3,-2).

Calculating $D = f_{xx}f_{yy} - f_{xy}^2$, we get,

$$\begin{aligned} f_{xx} &= 6x - 4, f_{yy} = 6y + 6, f_{xy} = 0 \\ D &= (6x - 4)(6y + 6) - 0 \\ D &= (36xy + 36x - 24y - 24) \end{aligned}$$

Plugging in the points,

$$\begin{aligned} (0,0) &: D = -24, f_{xx} = -4 \implies \text{saddle point} \\ (0,-2) &: D = 24, f_{xx} = -4 \implies \text{maxima} \\ (4/3,0) &: D = 24, f_{xx} = 4 \implies \text{minima} \\ (4/3,-2) &: D = -24, f_{xx} = 4 \implies \text{saddle point} \end{aligned}$$

(b) The gradient is given by,

$$\nabla f(x,y) = \begin{pmatrix} 3x^2 - 4x \\ 3y^2 + 6y \end{pmatrix}$$

At $(x, y) = (1, -1)$, we get,

$$
\begin{aligned}
\nabla f(1,-1) = u &= \begin{pmatrix} -1 \\ -3 \end{pmatrix} \\
\text{Taking } x - tu, \text{ we get } x - tu &= \begin{pmatrix} 1+t \\ -1+3t \end{pmatrix} \\
f(x - tu) &= 28t^3 + t^2 - 10t - 7 \\
\text{Calculating minimum of } f(x - tu), \text{ such that } t > 0 \text{ we get } t &= \frac{1}{3} \\
\text{Now, } x - tu &= \begin{pmatrix} \frac{4}{3} \\ 0 \end{pmatrix} \\
\nabla f(4/3, 0) = u &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned}
$$

Visualizing the plot, we get, Figures 6, 7. The yellow arrow signifies the direction of descent.

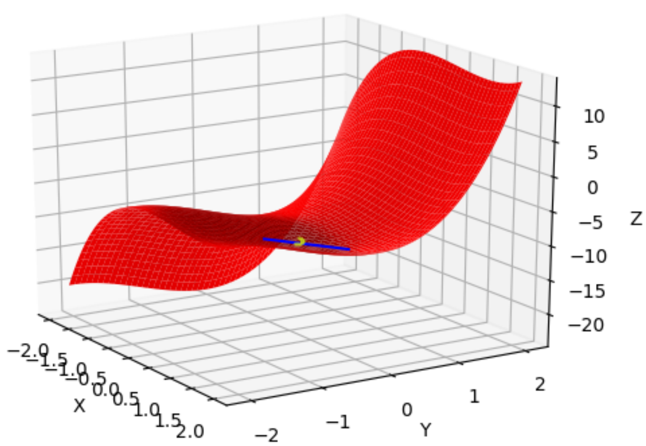

Figure 6: Steepest Descent Direction

We need only **1** step of the steepest descent to converge to an overall local minimum of f.

3 Question 3

(a) Consider the matrix Q which if n x n and is positive symmetric definite. For any eigenvectors v_i and v_j with distinct eigenvalues λ_i and λ_j, we get,

$$
\begin{aligned}
Qv_i &= \lambda_i v_i \\
v_j' Q v_i &= v_j' \lambda_i v_i \\
(Qv_j)' v_i &= \lambda_i v_j' v_i \\
\lambda_j v_j' v_i &= \lambda_i v_j' v_i \\
(\lambda_j - \lambda_i) v_j' v_i &= 0
\end{aligned}
$$

Since $\lambda_i \neq \lambda_j \implies v_j' v_i = 0$. Now substitute this in the equation

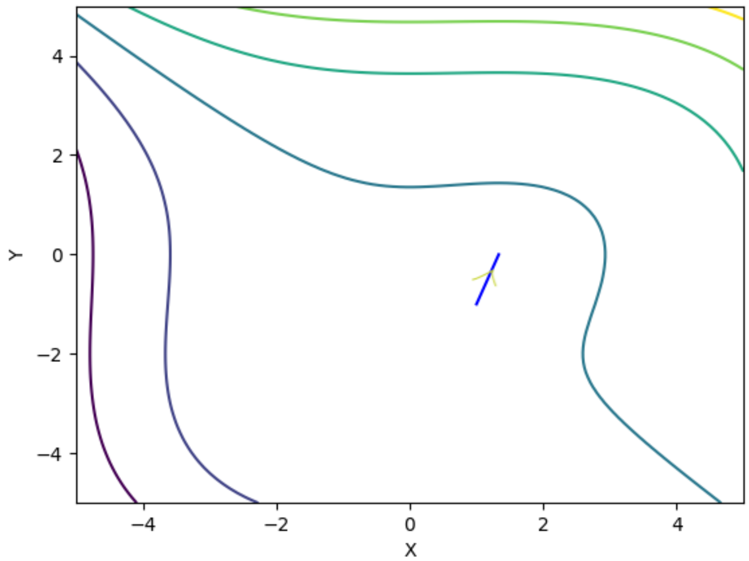

Figure 7: Steepest Descent

$$\begin{aligned} v_j' Q v_i &= v_j' \lambda_i v_i \\ &= \lambda_i v_j' v_i \\ &= 0 \end{aligned}$$

This implies that the eigen vectors v_j and v_i are Q-orthogonal

(b) From the eigen decomposition of a matrix, we have $Q = A\Lambda A^{-1}$ and since Q here is positive real symmetric definite, we have the decomposition as $Q = A\Lambda A^T$ (the eigenvectors are pairwise-orthogonal to each other and hence $A^{-1} = A^T$), where

$\Lambda = \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ 0 & ... & ... & 0 \\ 0 & 0 & ... & \lambda_n \end{pmatrix}$, where $\lambda_1, \lambda_2, ...\lambda_n$ are the eigen values of the Q matrix.

Now consider two eigen-basis vectors a_i and a_j of the Q matrix ($\implies a_i^T a_j = 0$), we get

$$\begin{aligned} a_i^T Q a_j &= a_i^T \lambda_j a_j \\ &= \lambda_j a_i^T a_j \\ &= 0 \end{aligned}$$

Hence any two basis vectors are Q-orthogonal.

4 Question 4

(a) For a purely quadratic polynomial, the directions d_k and d_{k-1} of the conjugate gradient descent algorithm are Q-orthogonal $\implies d_k^T Q d_{k-1} = 0$. Consider $d_k^T Q d_k$, we get,

$$
\begin{aligned}
d_k^T Q d_k &= d_k^T Q(-g_k + \beta_{k-1} d_{k-1}) \\
&= -d_k^T Q g_k + \beta_{k-1}(d_k^T Q d_{k-1}) \\
&= -d_k^T Q g_k
\end{aligned}
$$

(b) Consider y_k, we get,

$$
\begin{aligned}
y_k &= x_k - g_k \\
Q y_k &= Q x_k - Q g_k \\
Q y_k + b &= Q x_k + b - Q g_k \\
\nabla f(y_k) &= \nabla f(x_k) - Q g_k \\
p_k &= g_k - Q g_k \\
\implies Q g_k &= g_k - p_k
\end{aligned}
$$

(c) We can re-write the algorithm for conjugate gradient descent as

$$
\begin{aligned}
\text{Let } \ d_0 &= -g_0 = -\nabla f(x_0) \\
\text{for } \ k = 0, 1, ...n-1 \ \text{ do :}& \\
\alpha_k &= \frac{-g_k^T d_k}{d_k^T Q d_k} = \frac{g_k^T d_k}{d_k^T Q g_k} = \frac{g_k^T d_k}{d_k^T (g_k - p_k)} \\
x_{k+1} &= x_k + \alpha_k d_k \\
\beta_k &= \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k} = -\frac{(Q g_{k+1})^T d_k}{d_k^T Q g_k} = -\frac{(g_{k+1} - p_{k+1})^T d_k}{d_k^T (g_k - p_k)} \\
d_{k+1} &= -g_{k+1} + \beta_k d_k \\
\text{Return } \ x_n&
\end{aligned}
$$

5 Question 5

Consider a rectangle of perimeter of k, then the perimeter of the rectangle with length x and breadth y is given by $2x + 2y = k \implies 2x + 2y - k = 0$. We have to find the rectangle of maximum area $a(x, y) = xy$, which is equivalent to minimizing the area $a(x, y) = -xy$. Thus, applying the constraints, we get,

$$
\begin{aligned}
F(x, y, \lambda) &= -xy + \lambda(2x + 2y - k) \\
\nabla F(x, y, \lambda) &= \begin{pmatrix} -y + 2\lambda \\ -x + 2\lambda \\ 2x + 2y - k \end{pmatrix} \text{ taking derivative to compute maximum} \\
\nabla F(x, y, \lambda) &= 0 \\
\implies y &= 2\lambda \\
\implies x &= 2\lambda \\
\text{Putting this in } 2x + 2y = k, \text{ , we get} 8\lambda = k &\implies \lambda = \frac{k}{8} \\
\implies y &= \frac{k}{4}, x = \frac{k}{4}
\end{aligned}
$$

Thus the area is given by $a(x, y) = k^2/16$ and it is achieved at $x^* = \begin{pmatrix} k/4 & k/4 & k/8 \end{pmatrix}^T$

For the second order sufficiency conditions, we first need to compute the hessian of $F(x, y, \lambda)$. Doing this, we get,

$$\nabla F(x, y, \lambda) = \begin{pmatrix} -y + 2\lambda \\ -x + 2\lambda \\ 2x + 2y - k \end{pmatrix}$$
$$\nabla^2 F(x, y, \lambda) = \begin{pmatrix} 0 & -1 & 2 \\ -1 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$$

Now consider the vector $d = \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix}^T$ to be such that $\nabla F(x, y, \lambda) + d^T \nabla h(x, y, \lambda) = 0$ at $x = x^*$, where $h(x, y, \lambda) = 2x + 2y - k$.

$$\begin{aligned} \nabla F(x, y, \lambda) + d^T \nabla h(x, y, \lambda)\bigg|_{x=x^*} &= \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} \\ &= 2d_1 + 2d_2 = 0 \implies d_1 = -d_2, d_3 = 0 \end{aligned}$$

Now consider $\nabla^2 h(x, y, \lambda)$, we get,

$$\nabla^2 h(x, y, \lambda) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Now taking the inner product $d^T \left(\nabla^2 F(x, y, \lambda) + d^T \nabla^2 h(x, y, \lambda) \right) d$, and substituting the above results, we get,

$$\begin{aligned} d^T \left(\nabla^2 F(x, y, \lambda) + d^T \nabla^2 h(x, y, \lambda) \right) d \bigg|_{x=x^*} &= d^T \left(\nabla^2 F(x, y, \lambda) \right) d \bigg|_{x=x^*} \\ &= \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix} \begin{pmatrix} 0 & -1 & 2 \\ -1 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \\ &= \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix} \begin{pmatrix} -d_2 + 2d_3 \\ -d_1 + 2d_3 \\ 2d_1 + 2d_2 \end{pmatrix} \\ &= \begin{pmatrix} d_1 & d_2 & d_3 \end{pmatrix} \begin{pmatrix} -d_2 \\ -d_1 \\ 0 \end{pmatrix} \\ &= -2d_1 d_2 \\ &= 2d_1^2 \end{aligned}$$

Thus the quantity $d^T \left(\nabla^2 F(x, y, \lambda) + d^T \nabla^2 h(x, y, \lambda) \right) d$ is always greater than 0 implies it is a positive definite matrix, at the minima x^*. This proves the second-order sufficiency conditions.