Robotics 811 - Homework 4

Xiang Zhi Tan

October 28, 2015

1 Q1

1(a)

Fristly, we notice that the equation is a separative form of differentiation and can be separate into

$$\frac{dy}{dx} = \frac{2}{x^2}(1-y)$$
$$(1-y)dy = \frac{2}{x^2}dx$$

By applying integration to both side, we get

$$\int (1-y)dy = \int \frac{2}{x^2}dx$$
$$y - \frac{1}{2}y^2 + c = -2x^{-1} + c$$

By observing the equation, we notice it has the form of a quadratic equation, we then rearrange it and put it into the form of the quadratic formula.

$$y - \frac{1}{2}y^2 + c = -2x^{-1} + c$$
$$-\frac{1}{2}y^2 + y + (2x^{-1} - c) = 0$$
$$y^2 + (-2y) + (-4x^{-1} + c) = 0$$
$$y = \frac{2 \pm \sqrt{4 - 4(-4x^{-1} + c)}}{2}$$
$$y = 1 \pm \frac{\sqrt{4 + 16x^{-1} - 4c}}{2}$$
$$y = 1 \pm \frac{\sqrt{4}\sqrt{1 + 4x^{-1} - c}}{2}$$
$$y = 1 \pm \sqrt{1 + 4x^{-1} - c}$$
$$y = 1 \pm \sqrt{4x^{-1} - c}$$

Because we know the initial condition of $y(1) = -1$, we can plug it back into the previous equation and solve for c

$$-1 = 1 \pm \sqrt{\frac{4}{1} - c}$$

Because the square root will be positive, the only way to have a negative on the lefthand side would be that the $\pm$ is a negative sign.

$$-1 = 1 - \sqrt{4 - c}$$
$$2 = \sqrt{4 - c}$$
$$c = 0$$

Now, we kno that the analytic solution would be the following:

$$y(x) = 1 - 2\sqrt{x^{-1}}$$

1(b)

We impleted and used the Euler Method, following is the table with the found value and real answer from the expected value

n	x	y	real y
0	1.000000	-1.000000	-1.000000
1	0.950000	-1.050000	-1.051957
2	0.900000	-1.104050	-1.108185
3	0.850000	-1.162726	-1.169305
4	0.800000	-1.226723	-1.236068
5	0.750000	-1.296894	-1.309401
6	0.700000	-1.374293	-1.390457
7	0.650000	-1.460248	-1.480695
8	0.600000	-1.556452	-1.581989
9	0.550000	-1.665109	-1.696799
10	0.500000	-1.789149	-1.828427
11	0.450000	-1.932562	-1.981424
12	0.400000	-2.100956	-2.162278
13	0.350000	-2.302507	-2.380617
14	0.300000	-2.549691	-2.651484
15	0.250000	-2.862707	-3.000000
16	0.200000	-3.276924	-3.472136
17	0.150000	-3.861457	-4.163978
18	0.100000	-4.775677	-5.324555
19	0.050000	-6.507076	-7.944272
20	0.000000	-11.835382	-Inf

By observing the values, we notice this method is not too accurate and there are deviation in the end values.

1(c)

We implemented and used the fourth-order Runge-Kutta method. Following is the table containing the result

n	x	y	real y
0	1.000000	-1.000000	-1.000000
1	0.950000	-1.051957	-1.051957
2	0.900000	-1.108185	-1.108185
3	0.850000	-1.169305	-1.169305
4	0.800000	-1.236068	-1.236068
5	0.750000	-1.309401	-1.309401
6	0.700000	-1.390457	-1.390457
7	0.650000	-1.480695	-1.480695
8	0.600000	-1.581990	-1.581989
9	0.550000	-1.696800	-1.696799
10	0.500000	-1.828429	-1.828427
11	0.450000	-1.981426	-1.981424
12	0.400000	-2.162282	-2.162278
13	0.350000	-2.380624	-2.380617
14	0.300000	-2.651498	-2.651484
15	0.250000	-3.000033	-3.000000
16	0.200000	-3.472225	-3.472136
17	0.150000	-4.164292	-4.163978
18	0.100000	-5.326343	-5.324555
19	0.050000	-7.973392	-7.944272
20	0.000000	-183129202253624666683263156224.000000	-Inf

By observing the result, we observe that the values don't deviate until $n = 8$. Compared to Euler methods, fourth-order Runge Kutta is more accurate.

1(d)

We implemented and used the fourth-order Adams-Bashforth. Following is the table containing the result.

0	1.000000	-1.000000	-1.000000
1	0.950000	-1.051952	-1.051957
2	0.900000	-1.108174	-1.108185
3	0.850000	-1.169285	-1.169305
4	0.800000	-1.236038	-1.236068
5	0.750000	-1.309357	-1.309401
6	0.700000	-1.390393	-1.390457
7	0.650000	-1.480601	-1.480695
8	0.600000	-1.581854	-1.581989
9	0.550000	-1.696602	-1.696799
10	0.500000	-1.828133	-1.828427
11	0.450000	-1.980974	-1.981424
12	0.400000	-2.161562	-2.162278
13	0.350000	-2.379424	-2.380617
14	0.300000	-2.649367	-2.651484
15	0.250000	-2.995918	-3.000000
16	0.200000	-3.463310	-3.472136
17	0.150000	-4.141388	-4.163978
18	0.100000	-5.248560	-5.324555
19	0.050000	-7.504348	-7.944272
20	0.000000	-15.471548	-Inf

By observing the result, we see that the fourth-order Adams-Bashforth is more accurate than Euler's method but less accurate then fourth-order Runge-Kutta method.

2 Q2

2(a)

Firstly, we plotted the function in matlab using the *contour* function on the range, $y = [-4, 2]$ and $x = [-2, 4]$. Following is the result of the contour plot

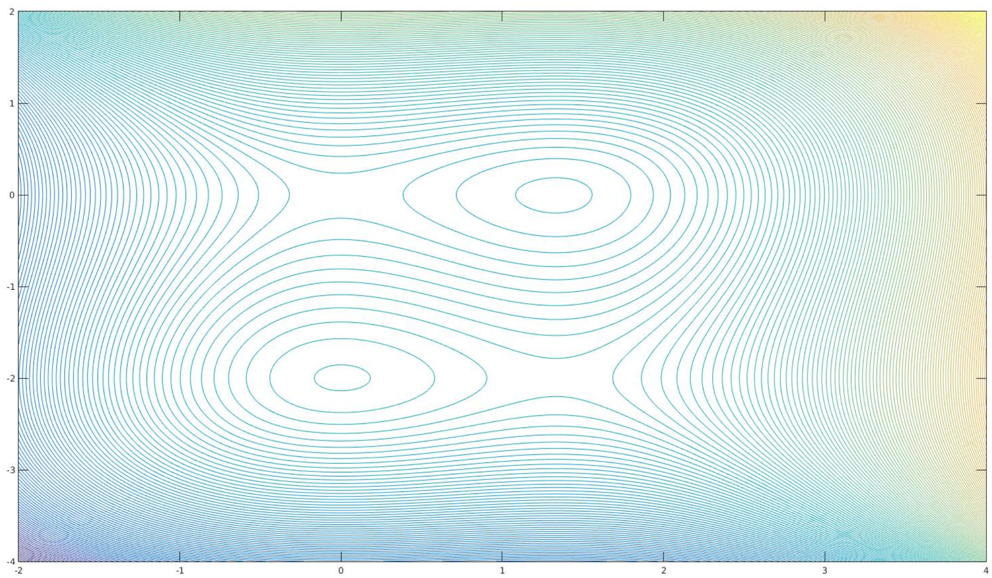

Figure 1: plot of the function $f(x, y) = x^3 + y^3 - 2x^2 + 3y^2 - 8$

By examining the sketch created, we notice there are four critical points, $(0, -2), (0, 0), (\frac{4}{3}, -2), (\frac{4}{3}, 0)$. We then measure the gradients(in the domain space) of the nearby points to determine the type of critical point. Following are the finding for each point:

1. Point$(0, -2)$
   This is a local maxima. By measuring the gradient at 8 points around the point with a step size of 0.1 relative to this point. we found them all to have a position gradient which means they are moving upwards towards $(0, -2)$ which means this is a local maximum

2. Point$(0, 0)$
   This is a saddle points. When measuring the gradient at 8 points around the point with a step size of 0.1 relative to this point, we measured a positive gradient at the two points of $(0.1, 0)$ and $(-0.1, 0)$ to have a negative gradient where as the remaining 6 points to have a negative gradient. This fits the definition of a saddle point.

3. Point$(\frac{4}{3}, 0)$
   This is the local minima. When measuring the gradient at the 8 points around the point with the step size of 0.1 relative to this point, all gradients are negative, this means that the direction of all these points are towards this point meaning that this point must serve as the local minima.

4. Point$(\frac{4}{3}, -2)$
   This is a saddle points. When measuring the gradient, we found a negative gradient at point $(1.23333, -2)$ and $(1.43333, -2)$ and positive gradient at all remaining points. This shows that this must be a saddle point.

2(b)

The gradient(partial derivative) of x and y can be found with the following equation:

$$\frac{\partial}{\partial x} = x(3x-4)$$
$$\frac{\partial}{\partial y} = y(3y+6)$$

Using the steepest descent algorithm, we first calculate the gradient at point $(1,-1)$. Which gives us the gradient at x, $\frac{\partial}{\partial x} = -1$ and gradient at y, $\frac{\partial}{\partial y} = -3$. Since the gradient are both non zero, we try to minimize the function $g(t) = f(x + t\vec{u})$ with u being the gradient. Minimizing the function, we found $t = \frac{1}{3}$. We update the points using the following algorithm $x^{n+1} = x^n - t\vec{u}$. After updating the points, we found the points to be $(\frac{4}{3}, 0)$ which is one of the critical points(also the local minima). The gradient at that point is also 0, stopping the algorithm. Therefore, we need only one step to converge to the local minimum.

3 Q3

3(a)

First, we know that eigenvectors of Q which is a real symmetric positive definite matrix are orthogonal. This means that given two eigenvectors, v_1 and v_2 come from two distinct eigenvalues, λ_1 and λ_2. Their inner product would be 0. By grinding through the equations, we could derive the Q-orthogonal definition.

$$\begin{aligned}
< v_1, v_2 > &= 0 \\
< \lambda_1 v_1, v2 > &= 0 \ \lambda \text{ is a scalar and will not change the result of the inner product} \\
< Qv_1, v_2 > &= 0 \\
(Qv_1)^T v_2 &= 0 \\
v_1^T Q^T v_2 &= 0 \\
v_1^T Q v_2 &= 0 \text{ the definition of Q-orthogonal}
\end{aligned}$$

3(b)

The solution on the previous section would be sufficient as it is based on the fact that eigenvectors of Q are orthogonal to each other, which is now explicitly given.

4 Q4

5 Q5

6 Q6

As we are finding the maximum area with a given parameter,p, we know that the objective function must be

$$f(x,y) = xy$$

with the constraint of

$$2y + 2x - p = 0$$

First we construct the Lagrangian

$$L(x,y,\alpha) = xy + \alpha(2y + 2x - p)$$

We then compute the gradient at set it to 0

$$\Delta L(x, y, \alpha) = \begin{pmatrix} y + 2\alpha \\ x + 2\alpha \\ 2x + 2y - p \end{pmatrix} = \vec{0}$$

We are now given 3 equations

$$y = 2\alpha \tag{1}$$

$$x = 2\alpha \tag{2}$$

$$\begin{aligned} 2x &= -2y + p \\ x &= \frac{-2y + p}{2} \end{aligned} \tag{3}$$

Insert 3 into 2, we get

$$2\alpha = \frac{-2y + p}{2} \tag{4}$$

Then, insert 4 into 1, we can solve y

$$\begin{aligned} y &= \frac{-2y + p}{2} \\ 2y &= -2y + p \\ y &= \frac{p}{4} \end{aligned} \tag{5}$$

We then insert 5 back into 2 and we solve x

$$x = \frac{p}{4} \tag{6}$$

We have now found the critical points in the equation, where $x = \frac{p}{4}$ and $y = \frac{p}{4}$. Notice that $x = y$. To show that we are achieving the maximum, we need to verify the second order sufficiency where we need to show the following matrix

$$L(x^*) = \Delta^2 f(x*) + \alpha^T \Delta^2 h(x^*) \tag{7}$$

is negative semi-definate on m which is $m = y|\Delta h(x^*)y = 0$. First we solve the partial derivatives for $f(x)$ and $h(x)$

$$\begin{aligned} \frac{\partial}{\partial^2 x} &= 0 \\ \frac{\partial}{\partial^2 y} &= 0 \\ \frac{\partial}{\partial xy} &= 1 \\ \frac{\partial}{\partial yx} &= 1 \\ \Delta^2 h &= 0 \end{aligned}$$

We have solved $L(x^*)$

$$L(x^*) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

We then check where the matrix is positive semi-definite or negative semi-definite by choosing a vector from the subspace M and apply $y^T Ly$.

$$\begin{aligned} \Delta h(x) &= [22] \\ [22]y &= 0 \\ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} &= \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned} \tag{8}$$

$$\begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -2 \tag{9}$$

This shows the matrix is negative semi-definite and therefore, x^* must be the maximum of f.

7 Q7

7(a)

We can first rearrange the constraints 2 and 3 into the following form by moving elements on the right handside to the left and multiplying constraints 2 by -1.

$$-w^T x_i - b + 1 - \xi_i \leq 0$$
$$\text{if } y_i = 1 w^T x_i + b + 1 - \xi_i \leq 0$$
$$\text{if } y_i = -1$$

Through observing the inequalities, we notice that the only difference is the element $w^T x_i$ which has a different sign that could be change by the values of y_i. This allows us to combine both inequalities into the following constraint.

$$-y_i(w^T x_i + b) + 1 - \xi_i \leq= 0$$

7(b)

The Lagrangian for our optimization problem is as following:

$$L([w\ 0\ \xi]^T, \alpha, \beta) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\alpha_i g_i([w\ b\ \xi]^T) - \sum_{i=1}^{l}\beta_i\xi_i$$

7(c)

Following is the steps to minimize the Lagrangian which is ther primal form of the SVM. Here's the original Lagrangian.

$$L([w\ 0\ \xi]^T, \alpha, \beta) = \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\alpha_i(-y_i(w^T x_i + b) + 1 - \xi_i) - \sum_{i=1}^{l}\beta_i\xi_i \tag{10}$$

First we, find the partial of w,b and ξ_i. Following are the partials follow by setting them to 0.
Finding partial of w

$$\begin{aligned} \frac{\partial}{\partial w} &= \frac{2}{2}w - \sum_{i=1}^{l}(\alpha_i y_i x_i) = 0 \\ w &= \sum_{i=1}^{l}(\alpha_i y_i x_i) \end{aligned} \tag{11}$$

Finding the partial of b.

$$\begin{aligned} \frac{\partial}{\partial b} &= -\sum_{i=1}^{l}(\alpha_i y_i) = 0 \\ \sum_{i=1}^{l}(\alpha_i y_i) &= 0 \end{aligned} \tag{12}$$

Finding the partial of xi_i.

$$\begin{aligned} \frac{\partial}{\partial \xi_i} &= C - \sum_{i=1}^{l}(\alpha_i) - \sum_{i=1}^{l}(\beta_i) = 0 \\ C &= \sum_{i=1}^{l}(\alpha_i) + \sum_{i=1}^{l}(\beta_i) \end{aligned} \tag{13}$$

Now we insert equations 11, 13 in equation 10.

$$
\begin{aligned}
L &= \frac{1}{2}(\sum_{i=1}^{l}(\alpha_i y_i x_i))^2 + (\sum_{i=1}^{l}(\alpha_i) + \sum_{i=1}^{l}(\beta_i))\sum_{i=1}^{l}\xi_i + \sum_{i=1}^{l}\alpha_i(-y_i((\sum_{j=1}^{l}(\alpha_j y_j x_j))x_i + b) + 1 - \xi_i) - \sum_{i=1}^{l}\beta_i\xi_i \\
L &= \frac{1}{2}(\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i\alpha_j y_i y_j x_j^T x_i)) + \sum_{i=1}^{l}(\alpha_i\xi_i) + \sum_{i=1}^{l}(\beta_i\xi_i) + \sum_{i=1}^{l}(\alpha_i\alpha_j y_i y_j x_j^T x_i)) \\
&- b\sum_{i=1}^{l}(\alpha_i y_i) + \sum_{i=1}^{l}(\alpha_i) - \sum_{i=1}^{l}(\alpha_i\xi_i) - \sum_{i=1}^{l}\beta_i\xi_i \\
L &= \sum_{i=1}^{l}(\alpha_i) - \frac{1}{2}(\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i\alpha_j y_i y_j x_j^T x_i)) - b\sum_{i=1}^{l}(\alpha_i y_i)
\end{aligned}
\tag{14}
$$

By applying equation 12 in the previous equation, we can derive the final form

$$
L^*(\alpha) = \sum_{i=1}^{l}(\alpha_i) - \frac{1}{2}(\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i\alpha_j y_i y_j x_j^T x_i)) \tag{15}
$$

7(d)

As we want to write the equation in terms of H and f such that $L^*(\alpha) = \frac{1}{2}\alpha^T H\alpha + f^T\alpha$. By looking at equation 15, we see a similar structure between them. For the first part $\frac{1}{2}\alpha^T H\alpha$

$$
\begin{aligned}
\frac{1}{2}\alpha^T H\alpha &= \frac{1}{2}(\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i\alpha_j y_i y_j x_j^T x_i)) \\
H &= (\sum_{i=1}^{l}\sum_{j=1}^{l}(y_i y_j x_j^T x_i))
\end{aligned}
$$

For second part $f^T\alpha$

$$
\begin{aligned}
f^T\alpha &= \sum_{i=1}^{l}(\alpha_i) \\
f^T &= [1, 1, ...1]
\end{aligned}
$$