16-811 Homework 4

Changsheng Shen (Bobby)
changshs@andrew.cmu.edu

10/24/2019

Q1:

(a).

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

Since $y(2) = \sqrt{2}$

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

Then we have:

$$y^2 = 2x - 2$$
$$y = \pm\sqrt{2x - 2}$$

Since $y(2) = \sqrt{2}$

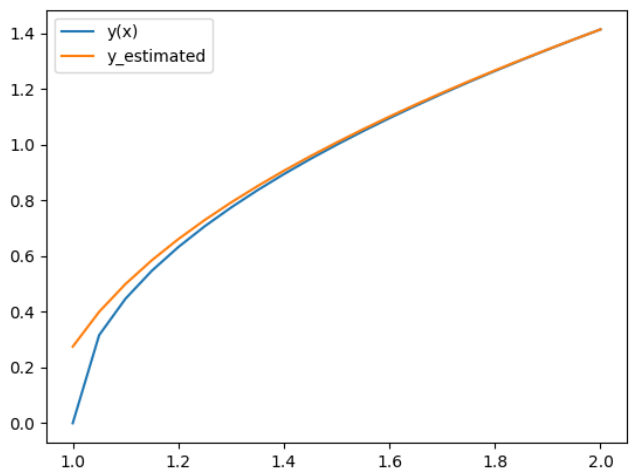
Then

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

(b). See q1.py

Euler's method:

x	$y(x_i)$	$y(x_i) - y_i$
2.0	1.4142135623730951	0.0
1.95	1.3788582233137678	-0.00045334810474573217
1.9	1.342596337099073	-0.000955550599199162
1.85	1.305355061655082	-0.00151458061455223
1.8	1.2670513070148661	-0.002140242947514359
1.75	1.2275896058799702	-0.0028447344883812953
1.7	1.1868593815765813	-0.003643424956658148
1.65	1.144731391710097	-0.004555966610959139
1.6	1.101053022826279	-0.005607907815946556
1.55	1.0556419490964637	-0.0068331009263120635
1.5	1.008277404666607	-0.00827740466660698
1.45	0.9586878772522415	-0.010004579201727704
1.4	0.9065332592723956	-0.012106068272479842
1.35	0.8513780848884671	-0.014718058354391439
1.3	0.7926497703035902	-0.018053101062107002
1.25	0.7295702087201059	-0.022463427533558344
1.2	0.6610367085821061	-0.028581176548430287
1.15	0.5853979439706101	-0.037675386465444105
1.1	0.4999859596700464	-0.05277236417008824
1.05	0.3999831515252011	-0.0837553855083637
1.0	0.2749778861550428	-0.2749778861550428

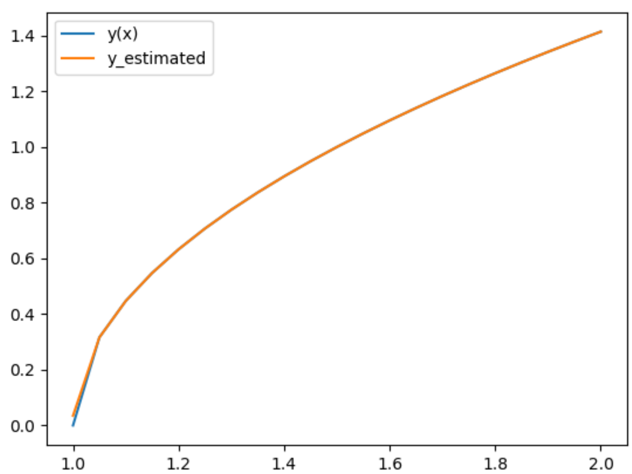


(c).

See q1.py

4-th order Runge-Kutta method:

x	$y(x_i)$	$y(x_i) - y_i$
2.0	1.4142135623730951	0.0
1.95	1.378404874894564	3.1445801518259486e-10
1.9	1.34164078577881	7.210638752752629e-10
1.85	1.303840479788252	1.2522778369827847e-09
1.8	1.2649110621126183	1.9547334861869103e-09
1.75	1.2247448684947508	2.8968381027283385e-09
1.7	1.183215952438591	4.181332169395091e-09
1.65	1.1401754191324296	5.9667082386027914e-09
1.6	1.0954451065052435	8.50508885541501e-09
1.55	1.048808835958106	1.221204559698208e-08
1.5	0.999999982198333	1.7801667029360146e-08
1.45	0.9486832714874119	2.6563101829246705e-08
1.4	0.8944271500334284	4.096648731355401e-08
1.35	0.836659960429315	6.61047606653753e-08
1.3	0.7745965557456568	1.1349582640374223e-07
1.25	0.7071065686624443	2.1252410331573657e-07
1.2	0.6324550805080248	4.515256509796117e-07
1.15	0.5477213900557294	1.167449436567658e-06
1.1	0.44720935257977523	4.242920182917498e-06
1.05	0.3161961398337223	3.162618311508547e-05
1.0	0.0349558741707901	-0.0349558741707901

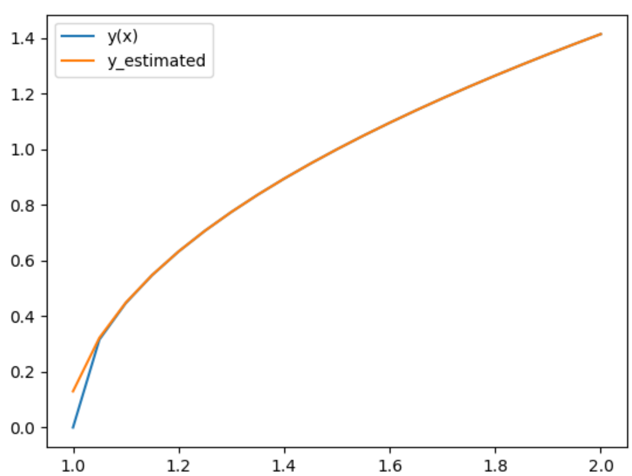


(d).

See q1.py

4-th order Adams-Bashforth method:

x	$y(x_i)$	$y(x_i) - y_i$
2.0	1.4142135623730951	0.0
1.95	1.3784052798425974	-4.046335753127295e-07
1.9	1.3416417196624542	-9.331625803188359e-07
1.85	1.3038420826781232	-1.6016375934402305e-06
1.8	1.2649135369653146	-2.472897962846332e-06
1.75	1.2247484905646338	-3.619173044810964e-06
1.7	1.1832211042929865	-5.147673063365943e-06
1.65	1.1401826446878964	-7.219588758600182e-06
1.6	1.0954551969301782	-1.0081919845861265e-05
1.55	1.0488229731708116	-1.412500066000355e-05
1.5	1.0000199876061941	-1.998760619414952e-05
1.45	0.9487120566219236	-2.875857140982596e-05
1.4	0.8944695770123618	-4.238601244610507e-05
1.35	0.8367245920233066	-6.456548923094996e-05
1.3	0.7746995014286151	-0.00010283218713191822
1.25	0.7072808032084978	-0.00017402202195027083
1.2	0.6327760881776687	-0.00032055614399295607
1.15	0.5483912182068935	-0.0006686607017275525
1.1	0.4489140295052684	-0.0017004340053102251
1.05	0.3225011673412169	-0.006273401324379535
1.0	0.1300828836872516	-0.1300828836872516



Q2:

$$\frac{\partial f}{\partial x} = 3x^2 - 4x$$

$$\frac{\partial f}{\partial y} = 3y^2 + 6y$$

Setting $\frac{\partial f}{\partial x} = 0$ gives:

$$x = 0 \text{ or } x = \frac{4}{3}$$

Setting $\frac{\partial f}{\partial y} = 0$ gives:

$$y = 0 \text{ or } y = -2$$

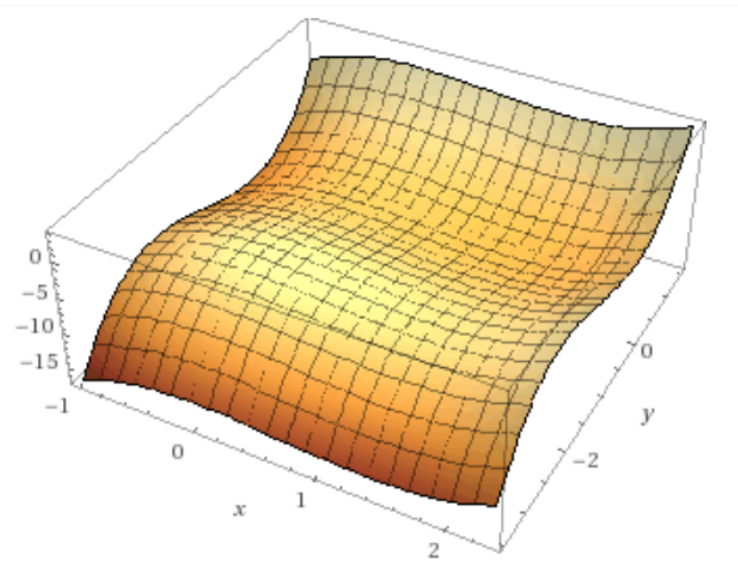


Figure 1: Sketch of $f(x, y)$

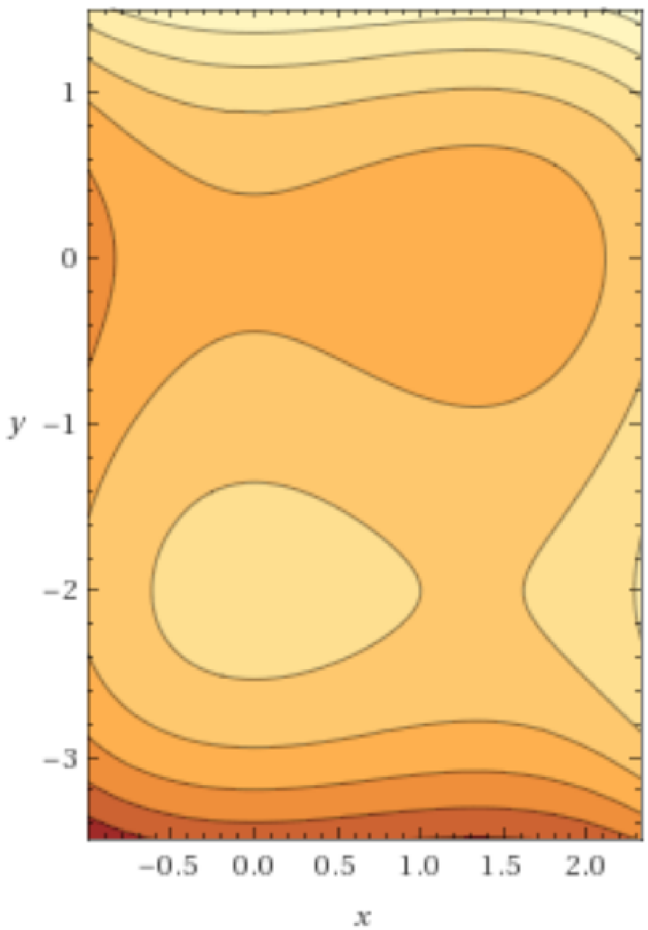


Figure 2: Sketch of the contours

Then we have four critical points:

$$(0,0),\ (\frac{4}{3},0),\ (0,-2),\ \text{and}\ (\frac{4}{3},-2)$$

To evaluate these critical points, compute the hessian matrix H

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 6x-4 & 0 \\ 0 & 6y+6 \end{bmatrix}$$

For the point (0, 0):

$$H = \begin{bmatrix} -4 & 0 \\ 0 & 6 \end{bmatrix}$$

So it is a saddle point.

For the point $(\frac{4}{3}, 0)$:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$$

So it is a local minima.

For the point $(0, -2)$:

$$H = \begin{bmatrix} -4 & 0 \\ 0 & -6 \end{bmatrix}$$

So it is a local maxima.

For the point $(\frac{4}{3}, -2)$:

$$H = \begin{bmatrix} 2 & 0 \\ 0 & -6 \end{bmatrix}$$

So it is a saddle point.

(b).

$$\nabla f(x, y) = (3x^2 - 4x, 3y^2 + 6y)$$

For the initial point (1, -1):

$$\nabla f(1, -1) = (-1, 3)$$

Minimizing

$$g(t) = f(x - t\nabla f(x, y)) = f(1 + t, -1 + 3t)$$

Setting $g(t) = 0$, solve for t with the constraint $t > 0$, we get:

$$t = \frac{1}{3}$$

Then the next point is:

$$(1 + \frac{1}{3}, -1 + 3 \cdot \frac{1}{3}) = (\frac{4}{3}, 0)$$

which is the local minima point, analyzed in (a).

Therefore, only one steepest descent step is needed.

Q3:

(a).

Suppose v_1 and v_2 are two eigenvectors of Q with distinct eigenvalues λ_1 and λ_2 ($\lambda_1 \neq \lambda_2$).

Then we have:

$$Qv_1 = \lambda_1 v_1$$
$$Qv_2 = \lambda_2 v_2$$

$$v_2^T Q v_1 = \lambda_1 v_2^T v_1$$
$$v_1^T Q v_2 = \lambda_2 v_1^T v_2$$

For the first equation, take the transpose of both sides, note that the right hand side of the equation is a scalar, so its transpose equals to its own:

$$(v_2^T Q v_1)^T = (\lambda_1 v_2^T v_1)^T = \lambda_1 v_2^T v_1$$

$$v_1^T Q^T v_2 = \lambda_1 v_2^T v_1$$

Since Q is a symmetric matrix, $Q^T = Q$

So:

$$v_1^T Q v_2 = \lambda_1 v_2^T v_1$$

Therefore:

$$\lambda_2 v_1^T v_2 = \lambda_1 v_2^T v_1$$

Also since $v_1^T v_2 = v_2^T v_1$:

Then

$$(\lambda_2 - \lambda_1) v_1^T v_2 = 0$$

Since we know that $\lambda_1 \neq \lambda_2$, so $v_1^T v_2 = 0$.

Therefore:

$$v_1^T Q v_2 = \lambda_2 v_1^T v_2 = 0$$

So we proved that v_1 and v_2 are Q-orthogonal.

(b).

For any two orthogonal basis vectors (of eigenvectors of Q) v_1 and v_2, we have the inner product (dot product) $v_1 \cdot v_2 = v_1^T v_2 = 0$.

Therefore:

$$v_1^T Q v_2 = v_1^T \lambda_2 v_2 = \lambda_2 v_1^T v_2 = \lambda_2 \cdot 0 = 0$$

So v_1 and v_2 are Q-orthogonal.

Q4:

(a).

$$\beta_k = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k}$$

$$d_{k+1} = -g_{k+1} + \beta_k d_k$$

$$d_k = -g_k + \beta_{k-1} d_{k-1}$$

Therefore

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

With this:

$$x_{k+1} = x_k + \alpha_k d_k = x_k - \frac{d_k^T g_k}{d_k^T Q d_k} d_k = x_k + \frac{d_k^T g_k}{d_k^T Q g_k} d_k$$

Therefore it is necessary to use Q only to evaluate g_k and Q_{g_k}

(b).

In the purely quadratic form: $\nabla f(y_k) = g_k = Q y_k + b$ where $y_k = x_k - g_k$

Therefore:

$$g_k - p_k = g_k - \nabla f(y_k) = g_k - Q(x_k - g_k) - b = g_k - (Q x_k + b) - Q g_k = Q g_k$$

(c).

From (a) and (b) we know that:

$$x_{k+1} = x_k + \frac{d_k^T g_k}{d_k^T Q g_k} d_k = x_k + \frac{d_k^T g_k}{d_k^T (g_k - p_k)} d_k$$

We can get $d_k^T g_k$ with:

$$d_k^T g_k = d_k^T Q x_0 + \alpha_0 d_k^T Q d_0 + \alpha_1 d_k^T Q d_1 + \ldots + \alpha_{k-1} d_k^T Q d_{k-1} + d_k^T b = d_k^T (Q x_0 + b) = d_k^T g_0$$

which does not require knowledge of the Hessian or f or a line search.

Q5:

Denote the given perimeter as C, then the problem can be state as:

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

subject to the constraints:

$$\begin{aligned} x &> 0 \\ y &> 0 \\ 2(x+y) &= C \end{aligned}$$

Since

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

is equivalent to

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

Denote

$$h(x) = 2(x+y) - C$$

and

$$F(x,y,\lambda) = -xy + \lambda h(x,y)$$

$$\nabla F = \begin{bmatrix} \frac{dF}{dx} \\ \frac{dF}{dy} \\ \frac{dF}{d\lambda} \end{bmatrix} = \begin{bmatrix} -y+2\lambda \\ -x+2\lambda \\ 2x+2y-C \end{bmatrix}$$

Setting $\Delta F = 0$ and solve it:

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0.25C \\ 0.25C \\ 0.125C \end{bmatrix}$$

Therefore when $x = y = \frac{C}{4}$, the rectangle (square) has the greatest area.

For the second-order sufficiency conditions:

$$[\nabla^2 f] = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

$$[\nabla^2 h] = \begin{bmatrix} \frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 h}{\partial x \partial y} \\ \frac{\partial^2 h}{\partial y \partial x} & \frac{\partial^2 h}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Therefore we have

$$L(x^*, y^*) = [\nabla^2 f] + \lambda[\nabla^2 h] - \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

For all vector v such that $v^T \nabla h = 0$, since $\nabla h = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

So

$$v = \begin{bmatrix} c \\ -c \end{bmatrix} = c \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where $c \in R$

Then we have:

$$v^T L(x^*, y^*) = 2c^2 \geq 0$$

which verifies the second-order sufficiency conditions.