Algorithms for Optimal Decisions
Tutorial 8
Answers

Exercise 1 *Solve the following problem using the interior point method:*

$$\begin{aligned} \min_x \quad f(x) &= x_2 \\ s.t. g_1(x) &= x_2 - \sin(x_1) - \frac{x_1}{2} \geq 0 \\ & x_1, x_2 \geq 0. \end{aligned} \tag{1}$$

Solution : The sequence of unconstrained problems, using the logarithmic barrier function, is:

$$\min_x f(x) - \eta_k \log(x_2 - \sin(x_1) - \frac{x_1}{2}). \tag{2}$$

Since the objective functions of problems (2) for $k = 1, 2, \dots$ are twice differentiable we use first derivatives to determine possible local minima of (2) for different values of η_k. Assume that η_k is fixed.

$$\begin{array}{l} \frac{\partial}{\partial x_1}(x_2 - \eta_k \log(x_2 - \sin(x_1) - \frac{x_1}{2})) = 0 \\ \frac{\partial}{\partial x_2}(x_2 - \eta_k \log(x_2 - \sin(x_1) - \frac{x_1}{2})) = 0 \end{array} \quad \Rightarrow \tag{3}$$

$$\Rightarrow \begin{array}{l} \frac{\eta_k(\cos(x_1) + \frac{1}{2})}{x_2 - \sin(x_1) - \frac{x_1}{2}} = 0 \\ 1 - \frac{\eta_k}{x_2 - \sin(x_1) - \frac{x_1}{2}} = 0 \end{array} \tag{4}$$

There are two sets of possible solutions of the system (4):

- Set 1

$$\begin{aligned} x_1^{(k)} &= \frac{2\pi}{3} \pm 2m\pi, \\ x_2^{(k)} &= \sin(\frac{2\pi}{3} \pm 2m\pi) + \frac{\pi}{3} \pm m\pi + \eta_k, \quad m = 0, 1, 2, \dots \end{aligned}$$

- Set 2

$$\begin{aligned} x_1^{(k)} &= \frac{4\pi}{3} \pm 2m\pi, \\ x_2^{(k)} &= \sin(\frac{4\pi}{3} \pm 2m\pi) + \frac{2\pi}{3} \pm m\pi + \eta_k, \quad m = 0, 1, 2, \dots \end{aligned}$$

The Hessian matrix of the unconstrained problem (2) is

$$H^{(k)} = \begin{bmatrix} \sin(x_1^{(k)}) & 0 \\ 0 & \frac{1}{\eta_k} \end{bmatrix}. \tag{5}$$

For $x_1^{(k)} = \frac{2\pi}{3} \pm 2m\pi$, matrix $H^{(k)}$ is not positive definite (check it !!!). However, for $x_1^{(k)} = \frac{4\pi}{3} \pm 2m\pi$ it is positive definite and therefore satisfies the sufficient conditions for a local unconstrained minimum.

Thus there is an infinite number of trajectories, one for each local minimum of the constrained problem (1). Table 2 shows the data for two different trajectories corresponding to $m = 0$ and $m = -1$.

Table 1:

		$m = 0$		$m = -1$	
k	η_k	$x_1^{(k)}$	$x_2^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$
1	2	$\frac{4\pi}{3}$	1.027π	$-\frac{2\pi}{3}$	0.028π
2	1	$\frac{4\pi}{3}$	0.709π	$-\frac{2\pi}{3}$	-0.290π
3	$\frac{1}{2}$	$\frac{4\pi}{3}$	0.550π	$-\frac{2\pi}{3}$	-0.449π
4	$\frac{1}{10}$	$\frac{4\pi}{3}$	0.423π	$-\frac{2\pi}{3}$	-0.576π

Exercise 2 *Solve the following system of nonlinear equations:*

$$\mathcal{W}(y) = \begin{bmatrix} y_1 + y_2 - 3 \\ y_1^2 + y_2^2 - 9 \end{bmatrix} = 0, \tag{6}$$

starting the algorithm with the initial estimate $y_0 = (1, 5)$.

Solution : First we compute $\nabla\mathcal{W}(y)$:

$$\nabla\mathcal{W}(y) = \begin{bmatrix} 1 & 1 \\ 2y_1 & 2y_2 \end{bmatrix}, \tag{7}$$

and then $(\nabla\mathcal{W}(y))^{-1}$

$$(\nabla\mathcal{W}(y))^{-1} = \begin{bmatrix} -\frac{y_2}{y_1-y_2} & \frac{1}{2(y_1-y_2)} \\ \frac{y_1}{y_1-y_2} & -\frac{1}{2(y_1-y_2)} \end{bmatrix} \tag{8}$$

1. $j = 0$, $y_0 = (1, 5)$.
   - $$\mathcal{W}_0 = \begin{bmatrix} 3 \\ 17 \end{bmatrix}, \quad \|\mathcal{W}_0\|_2 = \sqrt{3^2 + 17^2} > \epsilon.$$
   - $$d_0 = -\begin{bmatrix} \frac{5}{4} & -\frac{1}{8} \\ -\frac{1}{4} & \frac{1}{8} \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 17 \end{bmatrix} = -\begin{bmatrix} \frac{15}{4} - \frac{17}{8} \\ -\frac{3}{4} + \frac{17}{8} \end{bmatrix} = -\begin{bmatrix} \frac{13}{8} \\ \frac{11}{8} \end{bmatrix}$$
   - We assumed that $\tau = 1$ so y_1 is computed:
     $$y_1 = y_0 + d_0 = \begin{bmatrix} 1 \\ 5 \end{bmatrix} - \begin{bmatrix} \frac{13}{8} \\ \frac{11}{8} \end{bmatrix} = \begin{bmatrix} -\frac{5}{8} \\ \frac{29}{8} \end{bmatrix}$$
   - $$(\nabla\mathcal{W}_1)^{-1} = \begin{bmatrix} \frac{29}{34} & -\frac{2}{17} \\ \frac{5}{34} & \frac{2}{17} \end{bmatrix}, \quad j = j + 1.$$
2. $j = 1$, $y_1 = (-\frac{5}{8}, \frac{11}{8})$.
   - $$\mathcal{W}_1 = \begin{bmatrix} 0 \\ \frac{145}{32} \end{bmatrix}, \quad \|\mathcal{W}_1\|_2 = \sqrt{0^2 + (\frac{290}{64})^2} > \epsilon.$$
   - $$d_1 = -\begin{bmatrix} \frac{29}{34} & -\frac{2}{17} \\ \frac{5}{34} & \frac{2}{17} \end{bmatrix} \cdot \begin{bmatrix} 0 \\ \frac{145}{32} \end{bmatrix} = -\begin{bmatrix} -\frac{2}{17} \cdot \frac{145}{32} \\ \frac{2}{17} \cdot \frac{145}{32} \end{bmatrix} = \begin{bmatrix} \frac{145}{272} \\ -\frac{145}{272} \end{bmatrix}$$
   - We assumed that $\tau = 1$ so y_2 is computed:
     $$y_2 = y_1 + d_1 = \begin{bmatrix} -\frac{5}{8} \\ \frac{29}{8} \end{bmatrix} + \begin{bmatrix} \frac{145}{272} \\ -\frac{145}{272} \end{bmatrix} = \begin{bmatrix} -\frac{25}{272} \\ \frac{841}{272} \end{bmatrix}$$
   - $$(\nabla\mathcal{W}_2)^{-1} = \begin{bmatrix} 0.97113 & -0.15704 \\ 0.02887 & 0.15704 \end{bmatrix}, \quad j = j + 1.$$

3. $j = 2$, $y_2 = (-\frac{25}{272}, \frac{841}{272})$.

- $$\mathcal{W}_2 = \begin{bmatrix} 0 \\ 0.56837 \end{bmatrix}, \quad \|\mathcal{W}_2\|_2 = 0.56837 > \epsilon.$$

- $$d_2 = -\begin{bmatrix} 0.97113 & -0.15704 \\ 0.02887 & 0.15704 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0.56837 \end{bmatrix} = \begin{bmatrix} 0.08926 \\ -0.08926 \end{bmatrix}$$

- We assumed that $\tau = 1$ so y_3 is computed:
$$y_3 = y_2 + d_2 = \begin{bmatrix} -\frac{15}{272} \\ \frac{841}{272} \end{bmatrix} + \begin{bmatrix} 0.08926 \\ -0.08926 \end{bmatrix} = \begin{bmatrix} -0.00265 \\ 3.00265 \end{bmatrix}$$

- $$(\nabla\mathcal{W}_3)^{-1} = \begin{bmatrix} 0.99912 & -0.16637 \\ 0.00088 & 0.16637 \end{bmatrix}, \quad j = j + 1.$$

4. $j = 3$, $y_3 = (-0.00265, 3.00265)$.

- $$\mathcal{W}_3 = \begin{bmatrix} 0 \\ 0.01593 \end{bmatrix}, \quad \|\mathcal{W}_3\|_2 = 0.01593 > \epsilon.$$

- $$d_3 = -\begin{bmatrix} 0.99912 & -0.16637 \\ 0.00088 & 0.16637 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0.01593 \end{bmatrix} = \begin{bmatrix} 0.00265 \\ -0.00265 \end{bmatrix}$$

- We assumed that $\tau = 1$ so y_4 is computed:
$$y_4 = y_3 + d_3 = \begin{bmatrix} -0.00265 \\ 3.00265 \end{bmatrix} + \begin{bmatrix} 0.00265 \\ -0.00265 \end{bmatrix} = \begin{bmatrix} -0.000002 \\ 3.000002 \end{bmatrix}$$

- $$(\nabla\mathcal{W}_4)^{-1} = \begin{bmatrix} 0.999999 & -0.166666 \\ 0 & 0.166666 \end{bmatrix}, \quad j = j + 1.$$

5. $j = 4$, $y_4 = (0.000002, 3.000002)$.

- $$\mathcal{W}_4 = \begin{bmatrix} 0 \\ 0.000014 \end{bmatrix}, \quad \|\mathcal{W}_4\|_2 = 1.41 \cdot 10^{-5} > \epsilon.$$

- $$d_4 = -\begin{bmatrix} 0.999999 & -0.166666 \\ 0 & 0.166666 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0.000014 \end{bmatrix} = \begin{bmatrix} 2 \cdot 10^{-06} \\ -2 \cdot 10^{-06} \end{bmatrix}$$

- We assumed that $\tau = 1$ so y_4 is computed:

$$y_5 = y_4 + d_4 = \begin{bmatrix} -0.000002 \\ 3.000002 \end{bmatrix} + \begin{bmatrix} 2 \cdot 10^{-06} \\ -2 \cdot 10^{-06} \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

- $$(\nabla \mathcal{W}_5)^{-1} = \begin{bmatrix} 1 & -\frac{1}{6} \\ 0 & \frac{1}{6} \end{bmatrix}, \quad j = j + 1.$$

6. $j = 5$, $y_4 = (0, 3)$.

- $$\mathcal{W}_5 = \begin{bmatrix} 0 \\ 1.0978 \cdot 10^{-11} \end{bmatrix}, \quad \|\mathcal{W}_5\|_2 = 1.0978 \cdot 10^{-11} < \epsilon.$$

Results obtained with BFGS update are shown in the following table:

Table 2:

k	y_k	$\mathcal{W}_k$	$\|\mathcal{W}_k\|_2$
	1.00000000	3.00000000	
0	5.00000000	17.00000000	17.2626765016
	-0.62500000	0.00000000	
1	3.62500000	4.53125000	4.5312500000
	-0.07575758	0.00000000	
2	3.07575758	0.46602388	0.4660238751
	-0.01279427	0.00000000	
3	3.01279427	0.07709300	0.0770929956
	-0.00031382	0.00000000	
4	3.00031382	0.00188314	0.0018831430
	-0.00000133	0.00000000	
5	3.00000133	0.00000800	0.0000079954
	0.00000000	0.00000000	
6	3.00000000	0.00000000	0.0000000008

Exercise 3 *Find the appropriate stepsize parameter for the following system of equations:*

$$\mathcal{W}(y) = \begin{bmatrix} y_1^2 + y_2^2 - 2 \\ e^{y_1 - 1} + y_2^3 - 2 \end{bmatrix} = 0, \tag{9}$$

choosing $y_0 = (2, \frac{1}{2})$ *as a starting point.*

Solution :

$$\mathcal{W}_0 = \begin{bmatrix} 2.25 \\ e - \frac{15}{8} \end{bmatrix}.$$

$$d_0 = -(\nabla\mathcal{W}_0)^{-1} \cdot \mathcal{W}_0 = - \begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix}.$$

We start with $\tau = 1$ so y_1 is computed:

$$y_1 = y_0 + d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} -0.996676 \\ 10.236705 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 578736.27157 > 2.886235$$

- $\tau = \frac{1}{2}$:

$$y_1 = y_0 + \frac{1}{2}d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} 0.501662 \\ 5.368353 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 12119.807835 > 2.886235$$

- $\tau = \frac{1}{4}$:

$$y_1 = y_0 + \frac{1}{4}d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} 1.250831 \\ 2.934176 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 334.673756 > 2.886235$$

- $\tau = \frac{1}{8}$:

$$y_1 = y_0 + \frac{1}{8}d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \frac{1}{8}\begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} 1.625415 \\ 1.717088 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 18.606057 > 2.886235$$

- $\tau = \frac{1}{16}$:

$$y_1 = y_0 + \frac{1}{16}d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \frac{1}{16}\begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} 1.812708 \\ 1.108544 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 4.468205 > 2.886235$$

- $\tau = \frac{1}{32}$:

$$y_1 = y_0 + \frac{1}{32}d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \frac{1}{32}\begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} 1.906354 \\ 0.804272 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 3.097106 > 2.886235$$

- $\tau = \frac{1}{128}$:

$$y_1 = y_0 + \frac{1}{128}d_0 = \begin{bmatrix} 2 \\ .5 \end{bmatrix} - \frac{1}{128}\begin{bmatrix} -2.996676 \\ 9.736705 \end{bmatrix} = \begin{bmatrix} 1.976588 \\ 0.576068 \end{bmatrix}$$

$$\frac{1}{2}\|\mathcal{W}_1\|_2^2 = 2.864341 > 2.886235$$

Had we chosen to multiply τ by $\frac{1}{10}$ at each iteration it would have converged in the third iteration, for $\tau = 0.01$.