Midterm Exam Solution

- Instructor：Kuangyou Bruce Cheng
- TAs: Le Ngoc Hanh & 林筠禎

Problem 1

- A cylindrical tank has its height = h and the radius of the top and bottom = r. We want to minimize the total surface area (including the top, bottom, and the side) of this cylindrical tank. We are also told that the volume of this tank has to be fixed (i.e. $\pi r2h = C$, where C is a constant).

Problem 1

(5pt) Minimize $f(h, r) = 2\pi r^2 + 2\pi rh$, subject to the constraint $\pi r^2 h = C$.

- Please formulate this problem as an optimization problem. (5%)

Problem 1

- Please use Lagrange multiplier method to solve the problem (find the optimal r and h, and the corresponding function value). (10%)

(3pt) $L(r,h,\lambda) = (2\pi r^2 + 2\pi rh) + \lambda(\pi r^2 h - C)$

(2pt) $\begin{cases} \dfrac{\partial L}{\partial r} = 4\pi r + 2\pi h + 2\lambda\pi rh \\ \dfrac{\partial L}{\partial h} = 2\pi r + \lambda\pi r^2 \\ \dfrac{\partial L}{\partial \lambda} = \pi r^2 h - C \end{cases}$

(2pt) $\begin{cases} 4\pi r + 2\pi h + 2\lambda\pi rh = 0 \\ 2\pi r + \lambda\pi r^2 = 0 \\ \pi r^2 h - C = 0 \end{cases}$

(3pt) $h = \sqrt[3]{\frac{4C}{\pi}}, r = \sqrt[3]{\frac{C}{2\pi}}, f\left(\sqrt[3]{\frac{4C}{\pi}}, \sqrt[3]{\frac{C}{2\pi}}\right) = 2\pi\left(\frac{C}{2\pi}\right)^{\frac{2}{3}} + 2\pi\left(\frac{C}{2\pi}\right)^{\frac{1}{3}}\left(\frac{4C}{\pi}\right)^{\frac{1}{3}}$

Problem 2

- Find the stationary points of the following functions. Also use the second order derivative or Hessian matrix to identify whether the point is a local minimum, maximum, or neither.

Problem 2

- $f(x) = x^2e^{-x}$ (12%)

(2pt) $f'(x) = 2xe^{-x} - x^2e^{-x}$

(2pt) $x(2 - x)e^{-x} = 0$

(2pt) $x = 0\ or\ 2$

(2pt) $f''(x) = (2e^{-x} - 2xe^{-x}) - (2xe^{-x} - x^2e^{-x}) = (x^2 - 4x + 2)e^{-x}$

(1pt) $f''(0) = 2 > 0$

(1pt) $f''(2) = -2e^{-2} < 0$

(2pt) $(0, 0)$ is a local minimum and $(2, 4e^{-2})$ is a local maximum.

Problem 2

- f(x, y) = x + 4/(xy) + 5y (12%)

(2pt) $\begin{cases} \frac{\partial f}{\partial x} = 1 - \frac{4}{x^2 y} \\ \frac{\partial f}{\partial y} = -\frac{4}{xy^2} + 5 \end{cases}$

(2pt) $\begin{cases} 1 - \frac{4}{x^2 y} = 0 \\ -\frac{4}{xy^2} + 5 = 0 \end{cases}$

(2pt) $x = \sqrt[3]{20}, y = \sqrt[3]{\frac{4}{25}}$

Problem 2

- f(x, y) = x + 4/(xy) + 5y (12%)

(2pt) $H = \begin{bmatrix} \frac{8}{x^3 y} & \frac{4}{x^2 y^2} \\ \frac{4}{x^2 y^2} & \frac{8}{x y^3} \end{bmatrix}$

(1pt) At $x = \sqrt[3]{20}$ and $y = \sqrt[3]{\frac{4}{25}}$, $H = \begin{bmatrix} \sqrt[3]{\frac{2}{3}} & \sqrt[3]{\frac{25}{4}} \\ \sqrt[3]{\frac{25}{4}} & \sqrt[3]{6250} \end{bmatrix} \approx \begin{bmatrix} 0.873 & 1.842 \\ 1.842 & 18.42 \end{bmatrix}$

(1pt) $\begin{vmatrix} 0.873 - \lambda & 1.842 \\ 1.842 & 18.42 - \lambda \end{vmatrix} = 0, \lambda \approx 0.68 \; or \; 18.61$

(2pt) Since Hessian matrix has positive eigenvalues, $(\sqrt[3]{20}, \sqrt[3]{\frac{4}{25}}, 2\sqrt[3]{20} + 5\sqrt[3]{\frac{4}{25}})$ is a local minimum.

Problem 3-a: 15pts

Use classical methods (e.g. KKT conditions or duality theorem) to minimize the function $f(x, y) = (x-1)^2 + (y-1)^2$, subject to the following constraints: $2x + y >= 4$, and $x + 2y >=4$. Also, identify the active constraint(s).

Problem 3-a

Use the KKT method:

We have:

$$g_1 = -2x - y + 4$$

$$g_2 = -x - 2y + 4$$

$$L = (x-1)^2 + (y-1)^2 + \lambda_1(-2x - y + 4) + \lambda_2(-x - 2y + 4)$$

KKT conditions are as followed:

(2pts)

$$\begin{cases} \frac{\partial L}{\partial x} = 2(x-1) - 2\lambda_1 - \lambda_2 = 0 & (1) \\ \frac{\partial L}{\partial y} = 2(y-1) - \lambda_1 - 2\lambda_2 & (2) \\ \lambda_1 g_1 = 0 & (3) \\ \lambda_2 g_2 = 0 & (4) \\ g_1, g_2 \leq 0 & (5) \\ \lambda_1, \lambda_2 \geq 0 & (6) \end{cases}$$

Problem 3-a

Solve from (3):

Case 1: $\lambda_1 = 0$:

(3pts) Case 1.1: $\lambda_2 = 0$:

x =1; y=1; (violate to (5))

(3pts) Case 1.2: $-x - 2y + 4 = 0$:

x =1.2; y=1.4; (violate to (5))

Case 2: $-2x - y + 4 = 0$:

(3pts) Case 2.1: $\lambda_2 = 0$:

x =1.4; y=1.2; (violate to (5))

(3pts) Case 2.2: $-x - 2y + 4 = 0$:

$x = \frac{4}{3}; y = \frac{4}{3}$ (both constraints are active)

Finally, we have $(\frac{4}{3}, \frac{4}{3})$ is a local minimum. At this point, both constraints are active. (1pts)

Problem 3-b: 10pts

Please formulate the same problem as unconstrained optimization using exterior penalty functions. That is, transform the constraints to be part of the objective function. You do NOT need to solve this new unconstrained problem.

Problem 3-b

We have:

$f(X) = (x-1)^2 + (y-1)^2;$

(1.5 pts) $g_1 = -2x - y + 4;$

(1.5 pts) $g_2 = -x - 2y + 4;$

The function is formulated as:

(6.5 pts)
$$\phi\left(\begin{bmatrix} x \\ y \end{bmatrix}, r_k\right) = (x-1)^2 + (y-1)^2 + r_k\left[\left(max(0, -2x - y + 4)\right)^2 + \left(max(0, -x - 2y + 4)\right)^2\right],$$

where:

(0.5 pts) r_k is a positive constant;

$\begin{bmatrix} x \\ y \end{bmatrix}$ is in feasible region.

Problem 4-a: 15pts

- Minimize the function $f(x, y) = x^2 + y^2 - 7x - y - xy$.

Use the method of Nelder and Mead (downhill simplex method). Start from the simplex of [1; 1], [1.5; 1], [1; 1.5], and use the reflection coefficient $\alpha = 1$, expansion coefficient $\beta = 2$, and contraction coefficient $\gamma = 0.5$. You just need to find the next **2** simplexes (each simplex is composed of 3 points).

Problem 4-a

i=1:

Denote that: $X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $X_2 = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$; $X_3 = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}$

$f(X_1) = -7$; (1pts) $f(X_2) = -9.75$; (1pts) $f(X_3) = -6.75$; (1pts)

$f_{min} = f(X_2)$; $f_{max} = f(X_3)$;

Center point: $X_a = \begin{bmatrix} 1.25 \\ 1 \end{bmatrix}$ (1pts)

$X_r = X_a + \alpha(X_a - X_{max}) = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$; (1pts)

$f(X_r) = -9.25 > f_{min}$. So, we replace X_3 by X_r. (1pts)

The new simplex is: $\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} \right)$ (0.5pts)

Problem 4-a

i=2:

We find the second simplex by starting from the simplex of $X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, X_2 = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}, X_3 = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$

$f(X_1) = -7;$ (1pts) $\quad f(X_2) = -9.75;$ (1pts) $\quad f(X_3) = -9.25;$ (1pts)

$f_{min} = f(X_2); \quad f_{max} = f(X_1);$

$X_a = \begin{bmatrix} 1.5 \\ 0.75 \end{bmatrix}$ (1pts)

$X_e = \begin{bmatrix} 2.5 \\ 0.25 \end{bmatrix}$ (1pts)

$X_r = \begin{bmatrix} 2 \\ 0.5 \end{bmatrix};$ (1pts)

$f(X_e) = -12.0625 < f(X_r)$. So, we replace X_1 by X_e. (1pts)

$f(X_r) = -11.25 < f_{min}$. (1pts)

The new simplex is $\left(\begin{bmatrix} 2.5 \\ 0.25 \end{bmatrix}, \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} \right)$ (0.5pts)

Problem 4-b: 15pts

- Minimize the function $f(x, y) = x^2 + y^2 - 7x - y - xy$.

Use DFP method with the initial point $\underline{x}_1 = [1; 1]$. Find the next two points and the corresponding function values. If the minimum is found at $\underline{x}_2$, then you don't need to find $\underline{x}_3$.

Problem 4-b

$$X_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} 2x - 7 - y \\ 2y - x - 1 \end{bmatrix}$$ (1pts)

$$\nabla f(X_1) = \begin{bmatrix} -6 \\ 0 \end{bmatrix}$$ (1pts)

$$S_1 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$ (1pts)

$$X_2 = X_1 + \lambda_1 S_1 = \begin{bmatrix} 1 + 6\lambda_1 \\ 1 \end{bmatrix}$$ (1pts)

Min. $f(X_2)$, we get $\lambda_1 = \frac{1}{2}$. (1pts)

$$X_2 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$ (1pts)

$$\nabla f(X_2) = \begin{bmatrix} 0 \\ -3 \end{bmatrix};$$ (1pts)

$$g_1 = \begin{bmatrix} 6 \\ -3 \end{bmatrix};$$ (1pts)

$$B_2 = \begin{bmatrix} 0.7 & 0.4 \\ 0.4 & 0.8 \end{bmatrix};$$ (3pts)

$$S_2 = \begin{bmatrix} 1.2 \\ 2.4 \end{bmatrix};$$ (1pts)

$$X_3 = \begin{bmatrix} 4 + 1.2\lambda_2 \\ 1 + 2.4\lambda_2 \end{bmatrix};$$ (1pts)

Min. $f(X_3)$, we get $\lambda_2 = \frac{5}{6}$ (1pts)

So, $X_3 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$; (1pts)

Problem 5-a

(a) If $j=1$

Let $\underline{x}_1$ is the starting point and let the first search direction be

$\underline{S}_1 = -\nabla f(\underline{x}_1) = -(\mathbf{A}\underline{x}_1 + \mathbf{B})$

then $\underline{x}_2 = \underline{x}_1 + \lambda_1^* \underline{S}_1 \Rightarrow \mathbf{A}\underline{x}_2 = \mathbf{A}\underline{x}_1 + \mathbf{A}\lambda_1^*\underline{S}_1 \Rightarrow \mathbf{A}\underline{x}_2 + \mathbf{B} = \mathbf{A}\underline{x}_1 + \mathbf{B} + \mathbf{A}\lambda_1^*\underline{S}_1$

$\Rightarrow \nabla f(\underline{x}_2) = \nabla f(\underline{x}_1) + \mathbf{A}\lambda_1^*\underline{S}_1 \Rightarrow \underline{\mathbf{A}\underline{S}_1 = \dfrac{\nabla f(\underline{x}_2) - \nabla f(\underline{x}_1)}{\lambda_1^*}} \ldots(a)$

$\because \lambda_1^*$ is the minimizing step length in the direction $\underline{S}_1$, so that $\underline{S}_1^T \nabla f(\underline{x}_2) = 0 \Rightarrow \underline{\nabla f(\underline{x}_1)^T \nabla f(\underline{x}_2) = 0} \ldots(b)$

Now express the second search direction as (the problem said)

$\underline{S}_2 = -\nabla f(\underline{x}_2) + \beta_1 \underline{S}_1$

then $\underline{S}_1^T \mathbf{A}\underline{S}_2 = (\mathbf{A}\underline{S}_1)^T \underline{S}_2 = \underbrace{\left(\dfrac{\nabla f(\underline{x}_2) - \nabla f(\underline{x}_1)}{\lambda_1^*}\right)^T}_{\ldots from(a)} \left(-\nabla f(\underline{x}_2) + \beta_1 \underline{S}_1\right)$

$= \left(\dfrac{\nabla f(\underline{x}_2) - \nabla f(\underline{x}_1)}{\lambda_1^*}\right)^T \left(-\nabla f(\underline{x}_2) + \beta_1(-\nabla f(\underline{x}_1))\right)$

$= \dfrac{1}{\lambda_1^*}(\nabla f(\underline{x}_2) - \nabla f(\underline{x}_1))^T(-\nabla f(\underline{x}_2) - \beta_1 \nabla f(\underline{x}_1))$

$= \dfrac{1}{\lambda_1^*}\left(-\nabla f(\underline{x}_2)^T \nabla f(\underline{x}_2) + \underbrace{\nabla f(\underline{x}_1)^T \nabla f(\underline{x}_2)_{=0}}_{\text{(b)}} - \underbrace{\beta_1 \nabla f(\underline{x}_2)^T \nabla f(\underline{x}_1)_{=0}}_{\text{(b)}} + \beta_1 \nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1)\right)$

$= \dfrac{1}{\lambda_1^*}\left(-\nabla f(\underline{x}_2)^T \nabla f(\underline{x}_2) + \beta_1 \nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1)\right)$

$\because \beta_1 = \dfrac{\nabla f(\underline{x}_2)^T \nabla f(\underline{x}_2)}{\nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1)} \Rightarrow \nabla f(\underline{x}_2)^T \nabla f(\underline{x}_2) = \beta_1 \nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1)$

$\therefore \underline{S}_1^T \mathbf{A}\underline{S}_2 = \dfrac{1}{\lambda_1^*}\left(-\nabla f(\underline{x}_2)^T \nabla f(\underline{x}_2) + \beta_1 \nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1)\right)$

$= \dfrac{1}{\lambda_1^*}\left(-\beta_1 \nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1) + \beta_1 \nabla f(\underline{x}_1)^T \nabla f(\underline{x}_1)\right)_{=0} = 0$

So, when $j=1$, $\underline{S}_1^T \mathbf{A}\underline{S}_2 = 0$

Problem 5-b

- The solution to 5b can be found on page 188-189 of the textbook by Snyman (2018 version)