Assignment 2

1. Lagrange Multiplier Problem

a) $\min_{x} x_2 - x_1$

$g(x_1, x_2) = -x_1 + 4x_2 \leq 0$

$h(x_1, x_2) = \frac{1}{10}x_1^2 - x_2 - 3 = 0$

① Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda, \mu) = (x_2 - x_1) + \mu(-x_1 + 4x_2) + \lambda(\tfrac{1}{10}x_1^2 - x_2 - 3)$$

② KKT-conditions

$$\nabla_{x_1}\mathcal{L} = -1 - \mu + \tfrac{1}{5}\lambda x_1 = 0$$

$$\nabla_{x_2}\mathcal{L} = 1 + 4\mu - \lambda = 0$$

$\mu \geq 0$

$\mu g(x_1, x_2) = 0 \qquad h(x_1, x_2) = 0$

③ case 1 $\mu = 0$

$$\begin{cases} -1 + \frac{1}{5}\lambda x_1 = 0 \\ 1 - \lambda = 0 \\ \frac{1}{10}x_1^2 - x_2 - 3 = 0 \end{cases} \Rightarrow \begin{cases} \lambda^* = 1 \\ x_1^* = 5 \\ x_2^* = -\frac{1}{2} \end{cases}$$

$\nabla g(x_1^*, x_2^*) = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \quad \nabla h(x_1^*, x_2^*) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ regular point !

③ case 2 $\mu > 0$, $g(x_1, x_2) = 0$

$$\begin{cases} -1 - \mu + \frac{1}{5}\lambda x_1 = 0 \\ 1 + 4\mu - \lambda = 0 \\ -x_1 + 4x_2 = 0 \\ \frac{1}{10}x_1^2 - x_2 - 3 = 0 \end{cases} \Rightarrow \begin{cases} -1 - \mu + \frac{1}{5}\lambda x_1 = 0 \\ \lambda = 1 + 4\mu \\ x_1 = 4x_2 \\ \frac{1}{10}x_1^2 - x_2 - 3 = 0 \end{cases} \Rightarrow \begin{cases} (\frac{4}{5}x_1 - 1)\mu = 1 - \frac{1}{5}x_1 \\ \frac{16}{10}x_2^2 - x_2 - 3 = 0 \end{cases}$$

$$\Rightarrow \begin{cases} x_1^* = \frac{5 - \sqrt{505}}{4} \\ x_2^* = \frac{5 - \sqrt{505}}{16} \\ \lambda^* = \frac{-3\sqrt{505}}{101} \\ \mu^* = \frac{-3\sqrt{505}}{404} - \frac{1}{4} < 0 \ \lightning \end{cases} \quad \text{or} \quad \begin{cases} x_1^* = \frac{5 + \sqrt{505}}{4} \\ x_2^* = \frac{5 - \sqrt{505}}{16} \\ \lambda^* = \frac{3\sqrt{505}}{101} \\ \mu^* = \frac{3\sqrt{505}}{404} - \frac{1}{4} < 0 \ \lightning \end{cases}$$

Not KKT point Not KKT point

④ The optimal solution is $x^* = (5, -\frac{1}{2})$ with optimal value $-\frac{11}{2}$

(b) $\min_{x_1, x_2} \; x_1^2 + x_2^2$

$g_1(x_1, x_2) = -x_1 - x_2 + 3 \leq 0$

$g_2(x_2) = -x_2 + 2 \leq 0$

It is linearly constrained, so regularity is not required.

① Lagrangian

$$\mathcal{L}(x_1, x_2, \mu_1, \mu_2) = x_1^2 + x_2^2 + \mu_1(-x_1 - x_2 + 3) + \mu_2(-x_2 + 2)$$

② KKT condition

$$\nabla_{x_1} \mathcal{L} = 2x_1 - \mu_1 = 0$$

$$\nabla_{x_2} \mathcal{L} = 2x_2 - \mu_1 - \mu_2 = 0$$

$$\mu_1 \geq 0 \ , \ \mu_2 \geq 0$$

$$\mu_1 g_1(x_1, x_2) = 0 \quad \mu_2 g_2(x_2) = 0$$

③ case 1 $\mu_1 = 0 \ , \ \mu_2 = 0$

$$\Rightarrow x_1^* = 0, \ x_2^* = 0$$

$g_1 = 0 + 0 + 3 \not\leq 0$ ↯ Not feasible

case 2 $\mu_1 > 0, \ \mu_2 = 0 \ , g(x_1, x_2) = 0$

$$\begin{cases} -x_1 - x_2 + 3 = 0 \\ 2x_1 - \mu_1 = 0 \\ 2x_2 - \mu_1 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = \frac{3}{2} \\ x_2^* = \frac{3}{2} \\ \mu_1 = 3 \end{cases}$$

$g_2 = -\frac{3}{2} + 2 \not\leq 0$ ↯ Not feasible

case 3 $\mu_1 = 0, \ \mu_2 > 0 \ , g(x_2) = 0$

$$\begin{cases} x_2 = 2 \\ 2x_1 = 0 \\ 2x_2 - \mu_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = 0 \\ x_2^* = 2 \\ \mu_2 = 4 \end{cases}$$

$g_1 = 0 - 2 + 3 \not\leq 0$ ↯ Not feasible

case 4 $\mu_1 > 0 \ , \mu_2 > 0 \ , g(x_1, x_2) = 0 \ , g(x_2) = 0$

$$\begin{cases} x_2 = 2 \\ -x_1 - x_2 + 3 = 0 \\ 2x_1 - \mu_1 = 0 \\ 2x_2 - \mu_1 - \mu_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = 1 \\ x_2^* = 2 \\ \mu_1 = 2 \\ \mu_2 = 2 \end{cases}$$

④ The optimal solution is $[1 \quad 2]^T$ with optimal value 5.

(c) ① Lagrangian

$$\mathcal{L}(x_1, x_2, \mu) = (x_1 - 1)^2 + x_1 x_2^2 - 2 + \mu(x_1^2 + x_2^2 - 4)$$

② KKT condition

$$\nabla_{x_1}\mathcal{L} = 2(x_1 - 1) + x_2^2 + 2\mu x_1$$

$$\nabla_{x_2}\mathcal{L} = 2x_1 x_2 + 2\mu x_2$$

$$\mu \geq 0, \quad \mu g(x_1, x_2) = 0$$

③
Case 1 $\mu = 0$

$$\begin{cases} 2(x_1 - 1) + x_2^2 = 0 \\ 2x_1 x_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = 0 \\ x_2^* = \pm\sqrt{2} \end{cases} \text{ or } \begin{cases} x_1^* = 1 \\ x_2^* = 0 \end{cases}$$

Case 2 $\mu > 0, \ g(x_1, x_2) = 0$

$$\Rightarrow \begin{cases} x_1^2 + x_2^2 = 4 & ① \\ 2(x_1 - 1) + x_2^2 + 2\mu x_1 = 0 & ② \\ 2x_1 x_2 + 2\mu x_2 = 0 & ③ \end{cases}$$

③ $\Rightarrow \quad 2x_2(x_1 + \mu) = 0 \Rightarrow$

a) $x_2 = 0, \ x_1 = \pm 2$

If $x_1 = 2, \quad \mu = -\frac{1}{2}$ ↯

If $x_1 = -2, \quad \mu = -\frac{3}{2}$ ↯

b) $x_1 = -\mu$

①, ② $\Rightarrow -3x_1^2 + 2x_1 + 2 = 0$

$$\Rightarrow \begin{cases} x_1^* = \dfrac{1-\sqrt{7}}{3} \\ u = \dfrac{-1+\sqrt{7}}{3} \\ x_2^* = \pm 1.9233 \end{cases}$$

$\left(x_1 = \frac{1+\sqrt{7}}{3} \text{ is not KKT point as } u < 0 \right)$ ↯

④ Optimal solution :

$x^* = [1, 0]^T \quad f = -2$

$x^* = [0, \pm\sqrt{2}]^T \quad f = -1$

$x^* = [\frac{1-\sqrt{7}}{3}, \pm 1.9233]^T \quad f = -1.6311$

The optimal solution is $x^* = [1, 0]^T$
with optimal value -2.

2 Lagrange Augmentation

1° Lagrangian equation:

$$\mathcal{L}(x_1, x_2, \lambda) = (x_1 - 1)^2 - x_1 x_2 + \lambda(-x_1 - x_2 + 4)$$

KKT condition:

$$\nabla_{x_1}\mathcal{L} = 2(x_1 - 1) - x_2 - \lambda = 0 \quad ①$$

$$\nabla_{x_2}\mathcal{L} = -x_1 - \lambda = 0 \quad ②$$

$$-x_1 - x_2 + 4 = 0 \quad ③$$

② $\Rightarrow$ $x_1 = -\lambda$

③ $\Rightarrow$ $x_2 = -x_1 + 4 = \lambda + 4$

① $\Rightarrow$ $2(-\lambda - 1) - \lambda - 4 - \lambda = 0$

$\Rightarrow$ $\lambda = -\frac{3}{2}$, $x_1^* = \frac{3}{2}$, $x_2^* = \frac{5}{2}$

The optimal solution is $x^* = [\frac{3}{2}, \frac{5}{2}]^T$

with optimal value $-\frac{7}{2}$

2^{0} $g(\lambda_k) \doteq f(x_k) + \lambda_k^T h(x_k) - \frac{1}{2\partial}\|\lambda_k - \lambda_{k-1}\|_2^2$

$$\max_{\lambda_k} g(\lambda_k)$$

$$g(\lambda_k) = f(x_k) + \lambda_k^T h(x_k) - \frac{1}{2\partial}(\|\lambda_k\|_2^2 - 2\lambda_k^T \lambda_{k-1} + \|\lambda_{k-1}\|_2^2)$$

$$\nabla g(\lambda_k) = h(x_k) - \frac{1}{2\partial}(2\lambda_k - 2\lambda_{k-1})$$

$$= h(x_k) - \frac{1}{\partial}(\lambda_k - \lambda_{k-1}) = 0$$

Stationary point:

$$\Rightarrow \lambda_k^* = \partial h(x_k) + \lambda_{k-1}$$

$\nabla^2 g(\lambda_k) = -\frac{1}{\partial} I \prec 0 \Rightarrow \lambda_k^*$ is a strict global maximum point of $g(\lambda_k)$

$$\max_{\lambda_k} g(\lambda_k) = g(\lambda_k^*)$$

$$= f(x_k) + (\partial h(x_k) + \lambda_{k-1})^T h(x_k) - \frac{1}{2\partial}\|\partial h(x_k)\|_2^2$$

$$= f(x_k) + \lambda_{k-1}^T h(x_k) + \partial\|h(x_k)\|_2^2 - \frac{\partial}{2}\|h(x_k)\|_2^2$$

$$= f(x_k) + \lambda_{k-1}^T h(x_k) + \frac{\partial}{2}\|h(x_k)\|_2^2$$

There, eq.(3) leads to eq.(4), prouf done!

3° $\min_{x_k} f(x_k) + \lambda_{k-1}^T h(x_k) + \frac{\alpha}{2}\|h(x_k)\|_2^2$

$f(x) = (x_1-1)^2 - x_1x_2$ s.t. $x_1+x_2=4$

Without losing generality, the subscript k can be removed.

Let $g(x_1,x_2) = (x_1-1)^2 - x_1x_2 + \lambda(x_1+x_2-4) + \frac{\alpha}{2}(x_1+x_2-4)^2$

$\frac{\partial g(x_1,x_2)}{\partial x_1} = 2(x_1-1) - x_2 + \lambda + \alpha(x_1+x_2-4) = 0$ ①

$\frac{\partial g(x_1,x_2)}{\partial x_2} = -x_1 + \lambda + \alpha(x_1+x_2-4) = 0$ ②

①−② ⇒ $3x_1 - x_2 - 2 = 0 \Rightarrow x_2 = 3x_1 - 1$ ③

②,③ ⇒ $x_1^* = \frac{6\alpha-\lambda}{4\alpha-1}$, $x_2^* = \frac{10\alpha-3\lambda+2}{4\alpha-1}$ $(\alpha\neq\frac{1}{4})$!

$\nabla^2 g(x_1,x_2) = \begin{bmatrix} \alpha+2 & \alpha-1 \\ \alpha-1 & \alpha \end{bmatrix}$, To ensure that $\nabla^2 g(x_1,x_2) \succeq 0$ and x^* is a global minimum point

$\lambda_1+\lambda_2 = 2+2\alpha \geq 0$
$\lambda_1\lambda_2 = 4\alpha-1 \geq 0$
$\Rightarrow \alpha \geq \frac{1}{4}$, finally $\alpha > \frac{1}{4}$ as $\alpha \neq \frac{1}{4}$

Update rule: $x_{1k} = \frac{6\alpha-\lambda_{k-1}}{4\alpha-1}$, $x_{2k} = \frac{10\alpha-3\lambda_{k-1}+2}{4\alpha-1}$

9° What is the effect of choosing a larger value for α

Larger value of α makes smaller value change of x_k after each iteration. As the picture shows, larger value of α leads to x_k closer to optimal solution x^* after same number of iterations; therefore x_k converges to x^* more quickly.

3° Least square fitting

① By looking at the data, $d=2$

② Formulate a linear system of the form $Ax \simeq b$

$$A=\begin{bmatrix} x_1^0y_1^0 & x_1^0y_1^1 & x_1^0y_1^2 & x_1^1y_1^0 & x_1^1y_1^1 & x_1^1y_1^2 & x_1^2y_1^0 & x_1^2y_1^1 & x_1^2y_1^2 \\ \vdots & & & & & & & & \vdots \\ x_i^0y_i^0 & x_i^0y_i^1 & x_i^0y_i^2 & x_i^1y_i^0 & x_i^1y_i^1 & x_i^1y_i^2 & x_i^2y_i^0 & x_i^2y_i^1 & x_i^2y_i^2 \\ \vdots & & & & & & & & \vdots \\ x_n^0y_n^0 & x_n^0y_n^1 & x_n^0y_n^2 & x_n^1y_n^0 & x_n^1y_n^1 & x_n^1y_n^2 & x_n^2y_n^0 & x_n^2y_n^1 & x_n^2y_n^2 \end{bmatrix}_{n\times 9}$$

$$x=\begin{bmatrix} w_{00} & w_{01} & w_{02} & w_{10} & w_{11} & w_{12} & w_{20} & w_{21} & w_{22} \end{bmatrix}^T$$

$$b=[z_1, z_2 \cdots, z_n]^T$$

③ $x=\begin{bmatrix} 0.911 & 0.941 & 0.552 & -4.302 & -0.031 & 0.771 & 0.489 & -0.051 & 0 \end{bmatrix}$