

# Assignment 2

## 1. Lagrange Multiplier Problem

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

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

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

### ① Lagrangian

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

### ② KKT-conditions

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

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

$$\mu \geq 0$$

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

### ③ case 1 $\mu = 0$

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

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

③ case 2  $\lambda > 0$ ,  $g(x_1, x_2) = 0$

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

$$\Rightarrow \begin{cases} x_1^* = \frac{5 - \sqrt{505}}{4} \\ x_2^* = \frac{5 - \sqrt{505}}{16}^{2.2} \\ \lambda^* = \frac{-3\sqrt{505}}{101} \\ \mu^* = \frac{-3\sqrt{505}}{404} - \frac{1}{4} < 0 \end{cases} \text{ Not KKT point } \checkmark$$

or

$$\begin{cases} x_1^* = \frac{5 + \sqrt{505}}{4} \\ x_2^* = \frac{5 + \sqrt{505}}{16}^{2.1} \\ \lambda^* = \frac{3\sqrt{505}}{101} \\ \mu^* = \frac{3\sqrt{505}}{404} - \frac{1}{4} < 0 \end{cases} \text{ Not KKT point } \checkmark$$

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

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

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

$$\mu_1 > 0, \mu_2 > 0 \quad \checkmark$$

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

### ③ case 1 $\mu_1 = 0, \mu_2 = 0$

$$\Rightarrow x_1^* = 0, x_2^* = 0 \quad [3.1]$$

$$g_1 = 0 + 0 + 3 \neq 0 \quad \text{Not feasible} \quad \checkmark$$

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

$$\begin{cases} -x_1 + x_2 + 3 = 0 \\ 2x_1 - \mu_1 = 0 \\ 2x_2 - \mu_1 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = \frac{3}{2} \\ x_2^* = \frac{3}{2} \\ \mu_1 = 3 \end{cases} \quad [3.3]$$

$$g_2 = -\frac{3}{2} + 2 \neq 0 \quad \text{Not feasible} \quad \checkmark$$

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

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

$$g_1 = 0 - 2 + 3 \neq 0 \quad \text{Not feasible} \quad \checkmark$$

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

$$\begin{cases} x_2 = 2 \\ -x_1 + x_2 + 3 = 0 \\ 2x_1 - \mu_1 = 0 \\ 2x_2 - \mu_1 - \mu_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = 1 \\ x_2^* = 2 \\ \mu_1 = 2 \\ \mu_2 = 2 \end{cases} \quad [3.4] \quad \times$$

④ The optimal solution is  $[1 \ 2]^T$  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

$$\nabla_{x_1} \mathcal{L} = 2(x_1 - 1) + x_2^2 + 2\mu x_1, \quad 4.1$$

$$\nabla_{x_2} \mathcal{L} = 2x_1 x_2 + 2\mu x_2, \quad 4.2$$

$$\mu > 0 \quad \mu g(x_1, x_2) = 0 \quad 4.3$$

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

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

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

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

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

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

$$b) x_1 = -\mu$$

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

$$\Rightarrow \begin{cases} x_1^* = \frac{1-\sqrt{7}}{3} & \left( x_1 = \frac{1+\sqrt{7}}{3} \text{ is not } kKT \text{ point as } u < 0 \right) \\ u = \frac{-1+\sqrt{7}}{3} \\ x_2^* = \pm 1.9233 \end{cases}$$

5.1

④ Optimal solution :

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

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

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

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

## 2 Lagrange Augmentation

7<sup>o</sup> 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  $-\frac{7}{2}$

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

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

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

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

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

Stationary point.

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

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

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

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

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

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

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

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

$$f(x) = (x_1 - 1)^2 - x_1 x_2 \quad \text{s.t. } x_1 + x_2 = 4$$

Without loss of generality, the subscript  $k$ <sup>8.1</sup> can be removed. X

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

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

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

$$(1) - (2) \Rightarrow 3x_1 - x_2 - 2 = 0 \Rightarrow x_2 = 3x_1 - 1 \quad (3)$$

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

$$\nabla^2 g(x_1, x_2) = \begin{bmatrix} 2 & \alpha - 1 \\ \alpha - 1 & \alpha \end{bmatrix}$$

To ensure that  $x^*$  is a global minimum point,

$$\nabla^2 g(x_1, x_2) \geq 0$$

$$\Rightarrow \begin{cases} \text{Tr}(\nabla^2 g(x_1, x_2)) = 2\alpha + 2 \geq 0 \\ \det(\nabla^2 g(x_1, x_2)) = 4\alpha - 1 \geq 0 \end{cases} \Rightarrow \alpha \geq \frac{1}{4} \quad \checkmark$$

As  $\alpha \neq \frac{1}{4}$ , Finally  $\alpha > \frac{1}{4}$

$$\text{Update rule: } x_k = \left[ \frac{6\alpha - \lambda_{k-1}}{4\alpha - 1} \overset{8.2}{,} \frac{10\alpha - 3\lambda_{k-1} + 2}{4\alpha - 1} \right]^\top \quad \checkmark$$

Question 4 - 8 please see the code

9° What is effect of choosing a larger value for  $\lambda$

A larger value of  $\lambda$  makes the algorithm converge to the optimal solution more quickly.



### 3° Least square fitting

① By looking at the data,  $d=2$  X

10.1

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

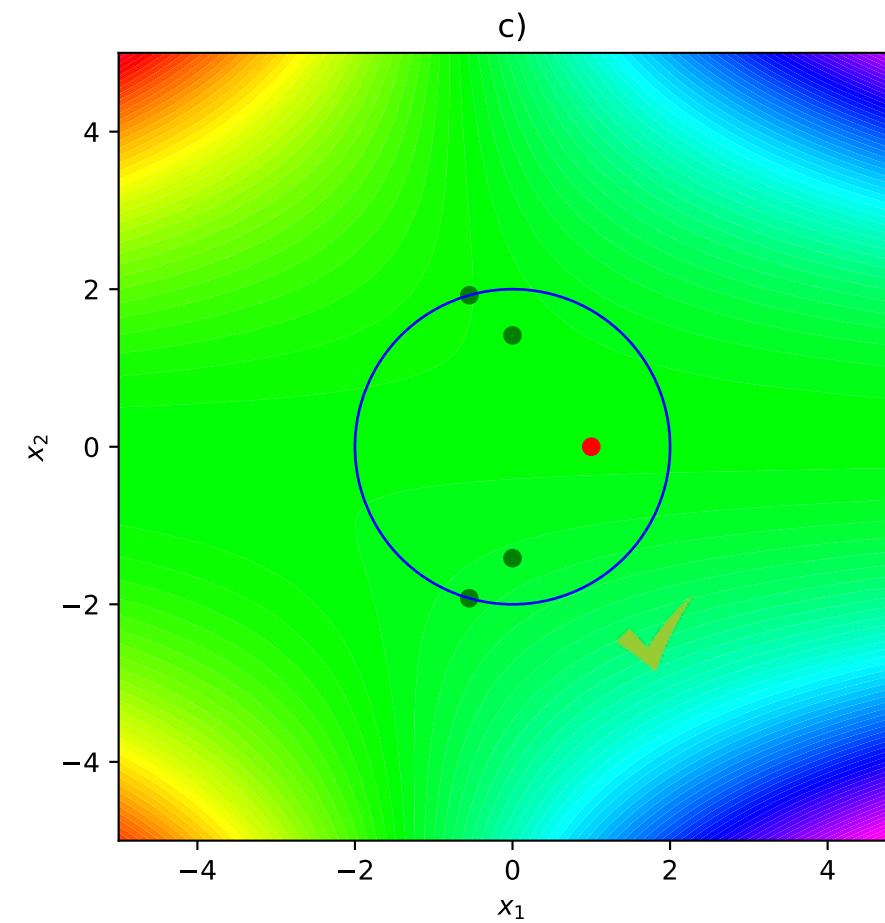
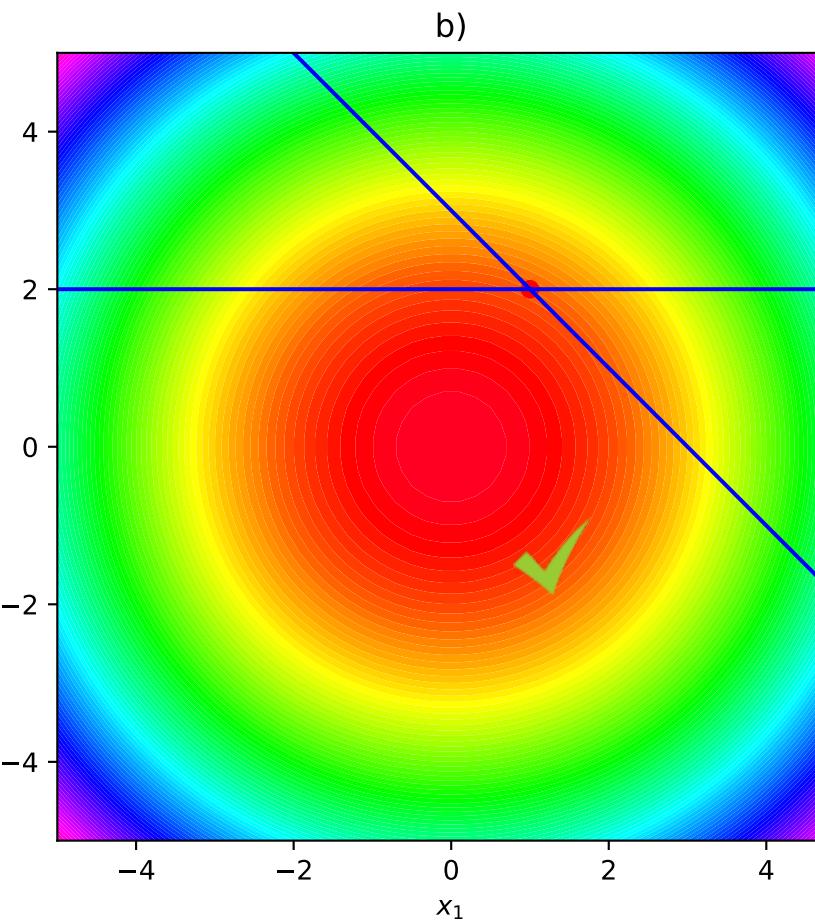
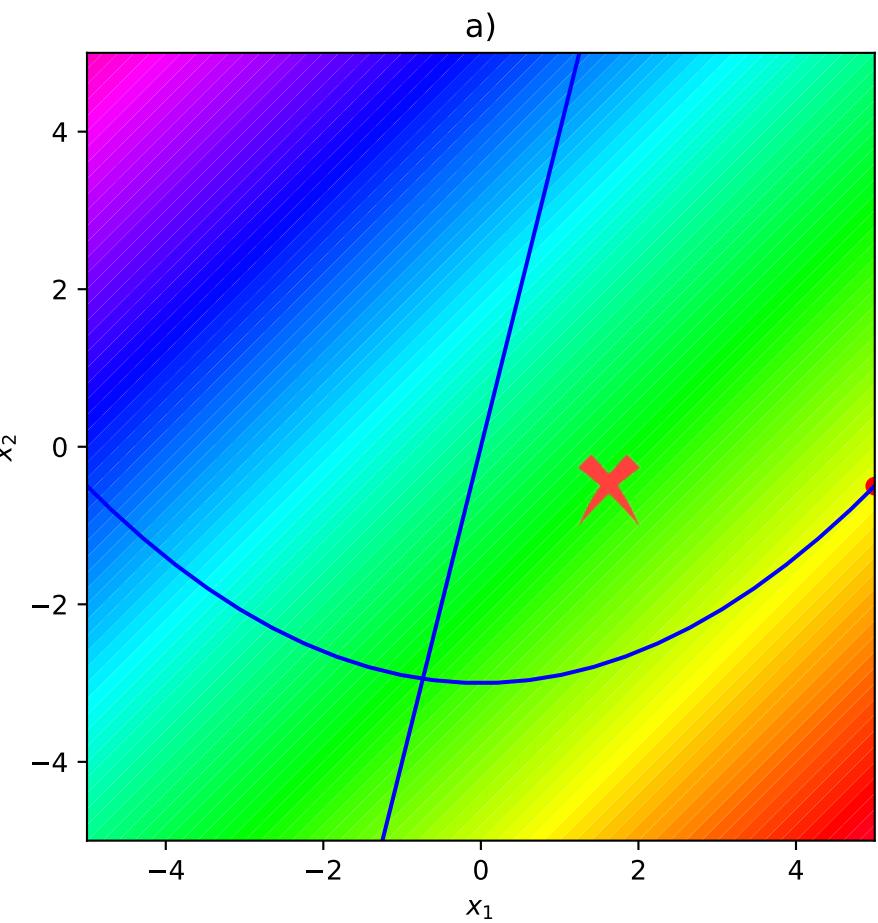
$$A = \begin{bmatrix} x_i^0 y_i^0 & x_i^0 y_i^1 & x_i^0 y_i^2 & x_i^1 y_i^0 & x_i^1 y_i^1 & x_i^1 y_i^2 & x_i^2 y_i^0 & x_i^2 y_i^1 & x_i^2 y_i^2 \\ \vdots & \vdots \\ x_n^0 y_n^0 & x_n^0 y_n^1 & x_n^0 y_n^2 & x_n^1 y_n^0 & x_n^1 y_n^1 & x_n^1 y_n^2 & x_n^2 y_n^0 & x_n^2 y_n^1 & x_n^2 y_n^2 \end{bmatrix}_{n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

$$x = [w_{00} \ w_{01} \ w_{02} \ w_{10} \ w_{11} \ w_{12} \ w_{20} \ w_{21} \ w_{22}]^T$$

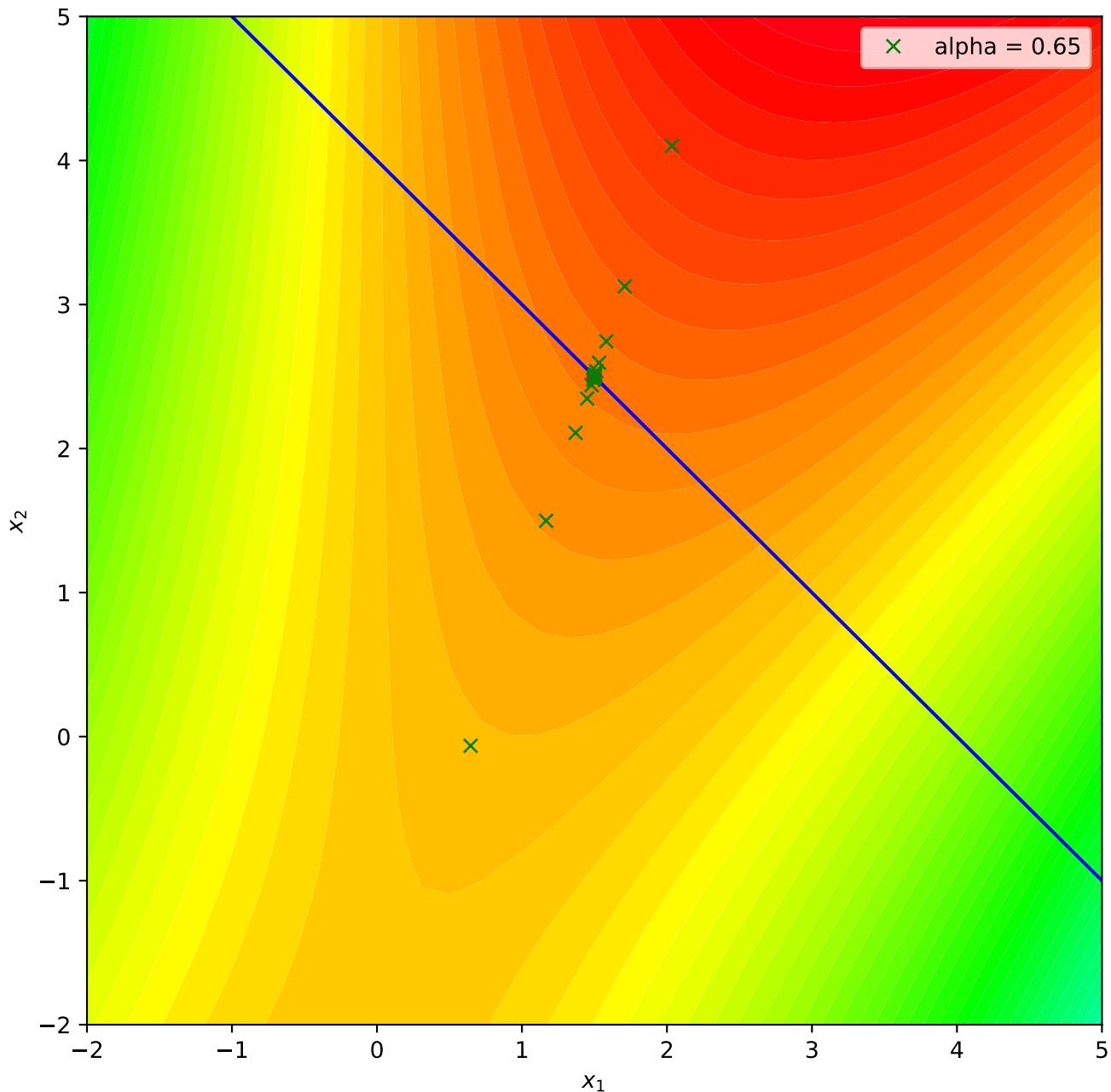
$$b = [z_1, z_2, \dots, z_n]^T$$

③  $x = [0.911 \ 0.941 \ 0.552 \ -4.302 \ -0.031 \ 0.771 \ 0.489 \ -0.051 \ 0]^T$

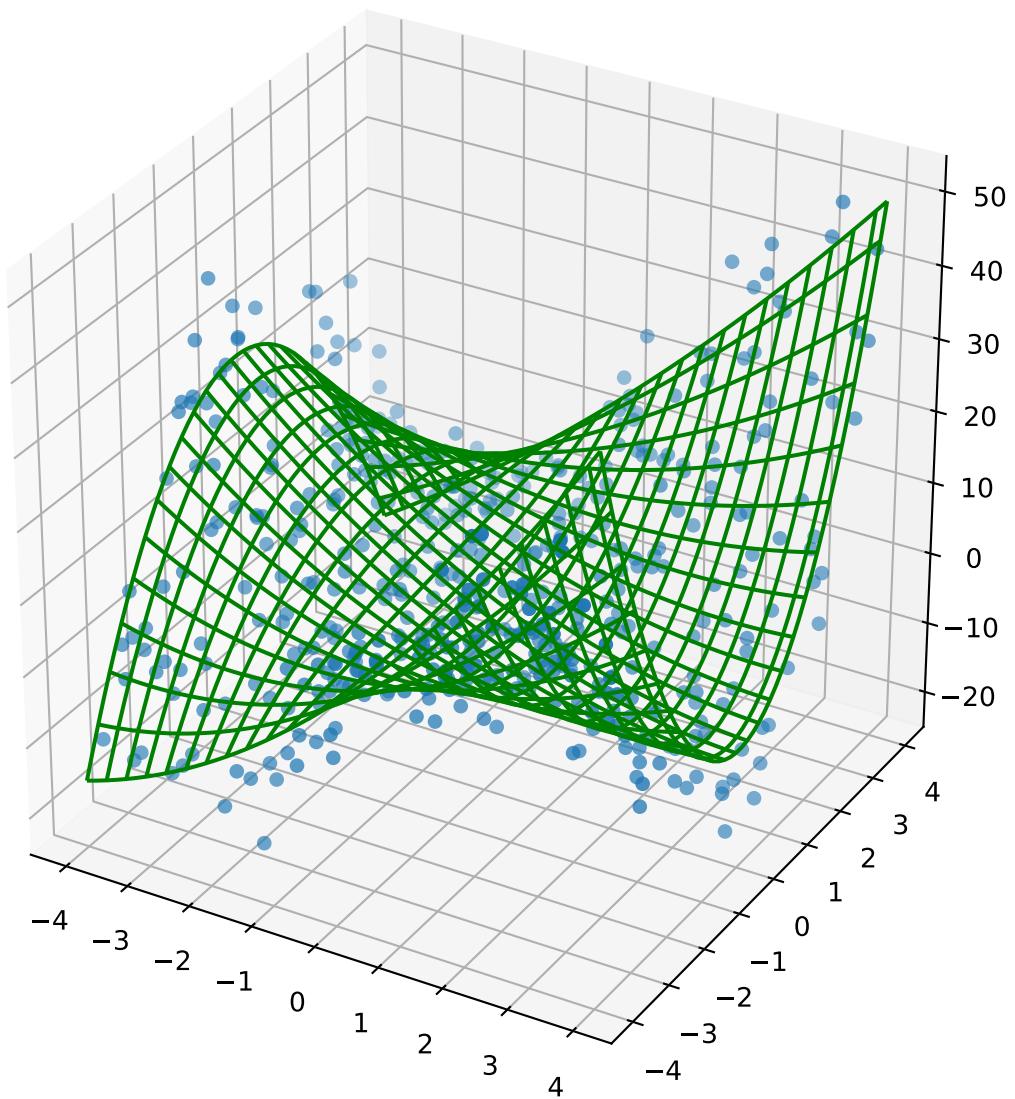
# Task 1 - Contour plots + Constraints



## Task 2 - Contour plots + Constraints + Iterations over k



### Task 3 - Data points vs. LS solution



# Index der Kommentare

---

- 2.1 +  
i guess this is a typo
- 2.2 looks like a 10 i hope you mean 16
- 3.1 steps?
- 3.2 please define when you use what equation, and what simplification you make
- 3.3 steps?
- 3.4 did not check if  $\mu > 0$
- 4.1 missing " $= 0$ "
- 4.2 what is g here? not defined
- 4.3 does it fulfill g?
- 5.1 please write down all points neatly and not only like this. this is rather confusing
- 8.1 well you can substitute it but you cannot remove it!
- 8.2 a little unclear since you removed k and added it back
- 10.1 should have at least used  $d = 3$  the MSE is rather large with  $d = 2$