University of Vienna
Faculty of Mathematics

Nonlinear Optimization Problems

Milutin Popovic

Contents

1 Sheet 3 **1**
1.1 Exercise 13 . . . 1
1.1.1 Part a . . . 1
1.1.2 Part b . . . 2
1.2 Exercise 14 . . . 2
1.3 Exercise 15 . . . 3
1.4 Exercise 16 . . . 4
1.4.1 Part a . . . 4
1.4.2 Part b . . . 5
1.5 Exercise 17 . . . 6
1.6 Exercise 18 . . . 7

1 Sheet 3

1.1 Exercise 13

1.1.1 Part a

Solve

$$\begin{aligned} \min \quad & -x_1 - 2x_2, \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 4 \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned} \tag{1}$$

rewriting it in to reduced notation

$$\begin{aligned} \min \quad & -x_1 - 2x_2, \\ \text{s.t.} \quad & g_1(x_1, x_2) = x_1^2 + x_2^2 - 4 \leq 0 \\ & g_2(x_1, x_2) = -x_1 \leq 0 \\ & g_3(x_1, x_2) = -x_2 \leq 0 \end{aligned} \tag{2}$$

We know that for a KKT point (x, λ) we have that the Lagrangian of the problem satisfies

$$\nabla_x L(x, \lambda) = 0 \tag{3}$$

$$\lambda \geqq 0, \quad \lambda^T g(x) = 0. \tag{4}$$

Then for $-\lambda_2 x_1 = 0$ and $-\lambda_3 x_2 = 0$ the only solution is for $\lambda_2, \lambda_2 = 0$. Now we have a system of three equations with three unknowns x_1, x_2, λ_1 that we can solve

$$\nabla f(x) + \nabla(\lambda^T g(x)) = 0 \tag{5}$$

$$\begin{pmatrix} -1 + 2x_1\lambda_1 \\ -2 + 2x_2\lambda_1 \\ -\lambda_1(x_1^2 + x_2^2 - 4) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{6}$$

Solving the first and second equation we get

$$x_1 = \frac{1}{2\lambda_1}, \quad x_2 = \frac{1}{\lambda_1} \tag{7}$$

$$x_2 = \frac{1}{2}x_1. \tag{8}$$

Plugging this into equation 3 and considering $x_1 \geq 0$, $x_2 \geq 0$ which tells us what root to take we get

$$x_1^2 + \frac{1}{4}x_1^2 - 4 = 0 \tag{9}$$

$$\Rightarrow x_1 = \frac{4}{\sqrt{5}}, \quad x_2 = \frac{2}{\sqrt{5}} \tag{10}$$

The solution the optimization problem is $x^* = (\frac{4}{\sqrt{5}}, \frac{2}{\sqrt{5}})^T$.

1.1.2 Part b

Verify if $x = (2,4)^T$ is an optimal solution of the optimization problem and determine a KKT point.

$$\begin{aligned} \min \quad & (x_1 - 4)^2 + (x_2 - 3)^2, \\ \text{s.t.} \quad & x_1^2 \leq x_2 \\ & x_2 \leq 4 \end{aligned} \tag{11}$$

i.e.

$$\begin{aligned} \min \quad & (x_1 - 4)^2 + (x_2 - 3)^2, \\ \text{s.t.} \quad & g_1(x) = x_1^2 - x_2 \leq 0 \\ & g_2(x) = x_2 - 4 \leq 0 \end{aligned} \tag{12}$$

We again use the KKT optimality conditions

$$\nabla f(x^*) + \nabla(\lambda^T g(x)) = 0 \tag{13}$$

$$\begin{pmatrix} 2(x_1 - 4) + 2\lambda_1 x_1 \\ 2(x_2 - 3) - \lambda_1 + \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{14}$$

substituting for $x = (2,4)^T$ gives

$$\begin{pmatrix} -4 + 4\lambda_1 \\ 2 - \lambda_1 + \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{15}$$

Which gives $\lambda_1 = 1, \lambda_2 = -1$ and tells us that $x = (2,4)^T$ is an optimal solution, and $\left(x^* = (2,4)^T, \lambda^* = (1,-1)^T\right)$ is a KKT point.

1.2 Exercise 14

Solve the following optimization problem

$$\begin{aligned} \min \quad & \Sigma_{i=1}^n (x_i - a_i)^2, \\ \text{s.t.} \quad & \Sigma_{i=1}^n x_i^2 \leq 1 \\ & \Sigma_{i=1}^n x_i = 0 \end{aligned} \tag{16}$$

To solve this we use the KKT optimality conditions for $g(x) = \Sigma_{i=1}^n x_i^2$ and $h(x) = \Sigma_{i=1}^n x_i = 0$.

$$\nabla f(x) + \lambda \nabla g(x) + \mu \nabla h(x) = 0 \tag{17}$$
$$\lambda \geq 0,\ g(x) \leq 0,\ \lambda g(x) = 0,\ h(x) = 0. \tag{18}$$

From the first equation we get

$$2x_i - 2a_i + 2\lambda x_i + \mu = 0 \tag{19}$$
$$2\left(1+\lambda\right)x_i + \mu - 2a_i = 0 \tag{20}$$
$$x_i = \frac{2a_i - \mu}{2(1+\lambda)} \qquad \forall i \in \{1, \ldots, n\}. \tag{21}$$

By substituting the derived expression for x_i into $h(x) = 0$ we get

$$\sum \frac{2a_i - \mu}{2(1+\lambda)} = 0 \tag{22}$$
$$\Rightarrow \mu = \frac{2}{n}\sum a_i \tag{23}$$

plugging this into $\lambda g(x) = 0$ we get

$$\sum \left(\frac{2a_i - \mu}{2\left(1+\lambda\right)}\right)^2 - 1 = 0 \tag{24}$$
$$\sum \left(2a_i - \mu\right)^2 = 4(1+\lambda)^2 \tag{25}$$
$$= 4\sum a_i^2 - 2\mu \sum a_i - n\mu^2 = 4\left(1+\lambda\right)^2 \tag{26}$$
$$\sum a_i^2 = (1+\lambda)^2. \tag{27}$$

Since $\lambda \geq 0$ then $(1+\lambda) \geq 1$ and the root is positive

$$\lambda = 1 - \sqrt{\sum a_i} \tag{28}$$

Then x_i becomes

$$x_i = \frac{2a_i - \mu}{2(1-\lambda)} \tag{29}$$
$$= \frac{2a_i - \frac{2}{n}\sum_j a_j}{2\sum_j a_j} \tag{30}$$
$$= \frac{a_i}{\sum_j a_j} - \frac{1}{n} \tag{31}$$
$$= \frac{1}{n}\left(\frac{a_i}{\langle a \rangle} - 1\right) \tag{32}$$

where $\langle a \rangle$ denotes the standard mean of $a = (a_1, \ldots, a_n)^T$.

1.3 Exercise 15

Consider the function

$$f:\ \mathbb{R}^2 \to \mathbb{R} \tag{33}$$
$$(x_1, x_2) \mapsto 3x_1^4 - 4x_1^2 x_2 + x_2^2 \tag{34}$$

Prove that the following statements for $x^* = (0,0)^T$ are true

1. x^* is a critical point of f
2. x^* is a strict local minimum of f along any line going through the origin

3. x^* is not a local minimum of f

For 1. we check if $\nabla f(x^*) = 0$ then x^* is a critical point

$$\nabla f(x) = \begin{pmatrix} 12x_1 + 8x_1x_2 \\ 4x_1^2 + 2x_2 \end{pmatrix} \tag{35}$$

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{36}$$

For 2. we need minimize $f(x)$ subjected to all lines through the origin. We start with lines $x_2 = mx_1$ for $m \neq 0$.

$$f(x_1, x_2 = mx_1) = 3x_1^4 - 4mx_1^3 + m^2x_1^2, \tag{37}$$

set $g(x) = f(x, mx)$. We need to check that $x = 0$ is local minimum of $g(x)$

$$g'(x) = 12x^3 - 12mx^2 + 2m^2x \tag{38}$$

$$g''(x) = 36x^2 - 24mx + 2m^2. \tag{39}$$

$$g'(0) = 0 \qquad g''(0) = 2m^2 > 0. \tag{40}$$

So f is a strict local min along lines $x_2 = mx_1$. Next we check along $x_1 = mx_2$

$$f(x_1 = mx_2, x_2) = 3m^4x_2^4 - tm^2x_2^3 + x_2^2 \tag{41}$$

set $g(x) = f(x_1 = mx_2, x_2)$

$$g'(x) = 12m^4x^3 - 12m^2x^2 + 2x \tag{42}$$

$$g''(x) = 36m^4x^2 - 24m^2x + 2 \tag{43}$$

$$g'(0) = 0 \qquad g''(0) = 2 > 0. \tag{44}$$

For 3. we need to show that $x^* = (0,0)^T$ is not a local minimum of f. Consider $x_2 = 2x_1^2$

$$f(x_1, 2x_1^2) = 3x_1^4 - 8x_1^4 + 4x_1^4 \tag{45}$$

$$= -x_1^4 < f(x^*) = 0 \qquad \forall x_1 \in \mathbb{R} \setminus \{0\} \tag{46}$$

We have found function values smaller than that of $f(x^*) = 0$.

1.4 Exercise 16

1.4.1 Part a

Formulate a statement concerning the solutions of the optimization problem

$$\begin{aligned} \max \quad & x_1, \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 1 \\ & (x_1 - 1)^2 + x_2^2 \geq 1 \\ & x_1 + x_2 \geq 1. \end{aligned} \tag{47}$$

using geometric arguments and verify this statement by means of arguments.
Let X be the optimization domain

$$X = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, (x_1 - 1)^2 + x_2^2 \geq 0, x_1 + x_2 \geq 0\right\} \tag{48}$$

Then X has the following graphical representation in the red domain of the plot below.

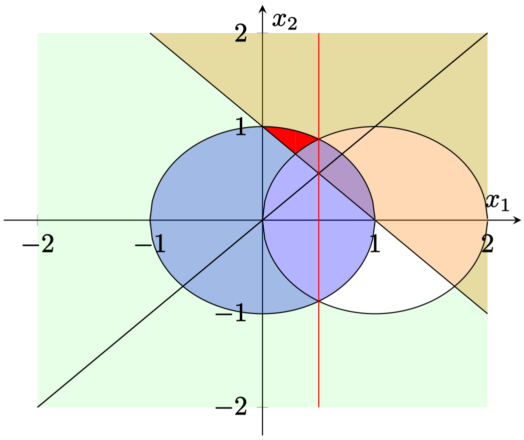

Figure 1: Area: Red: X, Blue: $x_1^2 + x_2^2 \leq 1$, Green: $(x_1 - 1)^2 + x_2^2 \geq 1$ and Orange: $x_1 + x_2 \geq 1$

Solutions are in the red area. But since $f(x_1, x_2) = x_1$ actually only depends on x_1 we can choose any x_2 then the maximum in the area is at $x_1 = \frac{1}{2}$. The analytical argumentation on the other hand follows KKT optimality condition, for this we transform the maximization problem into a minimization problem by multiplying the objective function f by -1.

$$\begin{aligned} \min \quad & -x_1, \\ \text{s.t.} \quad & g_1(x) = x_1^2 + x_2^2 - 1 \leq 0 \\ & g_2(x) = 1 - (x_1 - 1)^2 - x_2^2 \leq 0 \\ & g_3(x) = 1 - x_1 - x_2 \leq 0. \end{aligned} \tag{49}$$

then $\nabla L(x, \lambda) = 0$, $\lambda^T g(x) = 0$ and $\lambda^T \geqq 0$ will give us the optimal solution for the optimization problem

$$\begin{pmatrix} -1 + 2\lambda_1 x_1 - 2\lambda_2(x_1 - 1) - \lambda_3 \\ \lambda_1 x_2 - 2\lambda_2 x_2 - \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{50}$$

we set $x_2 = 0$ since x_2 is not dependent on objective then we get that $\lambda_3 = 0$ and

$$\lambda_1 = -\lambda_2 \frac{1 - (x_1 - 1)^2}{x_2^2 - 1}. \tag{51}$$

we are left with

$$-1 + 2\lambda_1 x_1 - 2\lambda_2(x_1 - 1) - \lambda_3. \tag{52}$$

Then $\lambda^T g(x)$ gives us

$$\lambda_1 = -\lambda_2 \frac{1 - (x_1 - 1)^2}{x_1^2 - 1} \tag{53}$$

substituting into 52 back and calculating we arrive at the equation

$$x_1^2 - x_1 + 1 = 0 \tag{54}$$

which gives $x_1 = \frac{1}{2}$.

1.4.2 Part b

Verify if $x^* = (1, 1)^T$ fulfills the constraint qualifications of LICQ, MFCQ, ABADIE-CQ.

$$\begin{aligned} \min \quad & x_1, \\ \text{s.t.} \quad & g_1(x) = x_1 + x_2 - 2 \leq 0 \\ & g_2(x) = 1 - x_1 x_2 \leq 0 \\ & g_3(x) = -x_1 \leq 0. \qquad\qquad g_4(x) = -x_2 \leq 0. \end{aligned} \tag{55}$$

Tangent cone are all tangent vectors $x_2 = \frac{1}{x_1} = 1$

$$T_X(x^*) = \left\{\begin{pmatrix} \lambda \\ -\lambda \end{pmatrix}, \begin{pmatrix} -\lambda \\ \lambda \end{pmatrix} : \lambda \geq 0\right\} \tag{56}$$

The linearized tangent cone is at

$$T_{\text{lin}}(x^*) = \left\{d \in \mathbb{R}^2 : \begin{pmatrix} 1 \\ 1 \end{pmatrix}^T d \leq 0, \begin{pmatrix} -1 \\ -1 \end{pmatrix}^T d \leq 0, \right\} \tag{57}$$

$$= \left\{\begin{pmatrix} \lambda \\ -\lambda \end{pmatrix}, \begin{pmatrix} -\lambda \\ \lambda \end{pmatrix} : \lambda \geq 0\right\}. \tag{58}$$

Which means x^* fulfills ABADIE-CQ. For MFCQ we need strict inequality $\nabla g_i(x^*)^T d < 0$ for $i \in \{1, 2\}$, which is not fulfilled for any $d \in T_{\text{lin}}(x)$. For LICQ we need that $\{\nabla g_i(x^*)\}_{i\in\{1,2\}}$ are linearly independent. But

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, (-1, -1) \tag{59}$$

are not linearly independent.

1.5 Exercise 17

Find out by using second order optimality conditions if $x^* = (0, 0)^T$ is a local minimum of

$$\begin{aligned} \min \quad & -x_1^2 + x_2, & (60) \\ \text{s.t.} \quad & g_1(x) = x_1^3 - x_2 \leq 0 \\ & g_2(x) = -mx_1 + x_2 \leq 0 \\ & & (61) \end{aligned}$$

where $m \geq 0$. We need to check that

$$d^T \nabla_x^2 L(x^*, \lambda^*) d > 0 \qquad \forall d \in T_2(x^*), \tag{62}$$

where

$$\begin{aligned} T_2(x^*) = \{d \in \mathbb{R}^2 : & \nabla g_i(x^*) d = 0 \;\; i \in \mathcal{A}_{\geq}(x^*), & (63) \\ & \nabla g_i(x^*) d \leq 0 \;\; i \in \mathcal{A}_0(x^*)\} & (64) \\ & & (65) \end{aligned}$$

and

$$\begin{aligned} \mathcal{A}_0(x^*) &= \{i \in \mathcal{A}(x^*) : \lambda_i^* = 0\} & (66) \\ \mathcal{A}_>(x^*) &= \{i \in \mathcal{A}(x^*) : \lambda_i^* > 0\} & (67) \\ & & (68) \end{aligned}$$

The gradients are

$$\begin{aligned} \nabla g_1(x)|_{x^*} &= (0, -1)^T & (69) \\ \nabla g_2(x)|_{x^*} &= (-m, 1)^T. & (70) \\ & & (71) \end{aligned}$$

Note that the KKT conditions $\lambda^T g(x) = 0$ and $\lambda \geqq 0$ are satisfied only if

$$\lambda^T g(x) = \lambda_1(-x_2 + x_1^3) + \lambda_2(-mx_1 + x_2) = 0 \tag{72}$$

$$\lambda_1 = \lambda_2 \frac{mx_1 - x_2}{x_1^3 - x_2}, \tag{73}$$

since $x_1^3 - x_2 \leq 0$ then the $\lambda^T \geqq 0$ is satisfied only if $\lambda 2 = 0$ which means $\lambda_1 = 0$. Which gives us $\mathcal{A}_> = \emptyset$ and $\mathcal{A}_0 = \{1, 2\}$ which means $T_2(x^*) = T_{\text{lin}}(x^*)$ and

$$T_{\text{lin}}(x^*) = \left\{ d \in \mathbb{R}^2 : \begin{pmatrix} 0 \\ -1 \end{pmatrix}^T d \leq 0, \begin{pmatrix} -m \\ -1 \end{pmatrix}^T d \leq 0, \right\} \tag{74}$$

$$= \left\{ \begin{pmatrix} \lambda \\ \lambda m \end{pmatrix} : \lambda \geq 0 \right\} \tag{75}$$

Now we calculate the hessian of the Lagrangian

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) \tag{76}$$

$$= f(x) \tag{77}$$

$$\nabla^2 L(x, \lambda) = \nabla^2 f(x) \tag{78}$$

$$= \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \tag{79}$$

Then

$$\begin{pmatrix} \lambda \\ \lambda m \end{pmatrix}^T \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \lambda m \end{pmatrix} = -2\lambda^2 \not> 0. \tag{80}$$

We conclude that $x^* = (0, 0)^T$ is not a local minimum of the optimization problem.

1.6 Exercise 18

Let $(t_k)_{k \geq 0} \subseteq \mathbb{R}$ be a monotonically decreasing sequence and t^* an accumulation point of it. Show that the sequence $(t_k)_{k \geq 0}$ converges to t^*.
We know that t^* is an accumulation point of $(t_k)_{k \geq 0}$ so

$$\forall U_\varepsilon(t^*) = [t^* - \varepsilon, t^* + \varepsilon], \varepsilon > 0 \;\; \forall N \in \mathbb{N} \;\; \exists n \geq N : t_n \in U_\varepsilon(t^*) \tag{81}$$

$$\text{i.e.} \quad |t_n - t^*| < \varepsilon \qquad \forall n \geq N \in \mathbb{N} \tag{82}$$

since $(t_k)_{k \geq 0}$ monotonically decreasing, $t_0 > t_1 > \ldots > t_k > \ldots$ we have that $\forall n \in N$

$$\varepsilon_n > |t_n - t^*| > |t_{n+1} - t^*| \tag{83}$$

so there exists a positive, strictly monotonically decreasing subsequence $(\varepsilon_k)_{k \geq 0}$ of $(t_k)_{k \geq 0}$ defined by $\varepsilon_n > |t_n - t^*|$ and $\varepsilon_n > \varepsilon_{n+1}$ that converges to 0 so $(t_k)_{k \geq 0}$ converges to t^*.