University of Vienna
Faculty of Mathematics

Nonlinear Optimization Problems

Milutin Popovic

Contents

1 Sheet 7

1.1 Exercise 43

Consider the optimization problem

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

with $x = (x_1, x_2)^T$. For $\alpha > 0$, find the minimum $x^*(\alpha)$ of the penalty function

$$P(x;\alpha) := f(x) + \frac{\alpha}{2}\|g_+(x)\|^2 \tag{2}$$

and the limit points $x^* = \lim_{\alpha\to+\infty} x^*(\alpha)$ and $\lambda^* = \lim_{\alpha\to+\infty} \alpha g_+(x^*(\alpha))$. Find out if (x^*, λ^*) is a KKT point of the constrained optimization problem.
First we find the minimum of $P(x;\alpha)$.

$$\nabla P(x;\alpha) = \nabla f(x) + \frac{\alpha}{2}\left(\nabla\left(\max(0,-x_1)\right)^2 + \nabla\left(\max\left(0,-x_2\right)\right)^2\right) \tag{3}$$

since $\frac{\partial}{\partial x_i}\max(0,-x_i)^2$ is $2x_i$ for $x_i < 0$ and 0 otherwise for all $i = 1, 2$, so we have the equations

$$\nabla P(x;\alpha) = \begin{pmatrix} 2(x_1+1) \\ 2(x_2+2) \end{pmatrix} + \alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \tag{4}$$

which gives

$$x^*(\alpha) = \begin{pmatrix} -2(2+\alpha)^{-1} \\ -4(2+\alpha)^{-1} \end{pmatrix} \tag{5}$$

$$x^* = \lim_{\alpha\to+\infty} x^*(\alpha) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{6}$$

and

$$\lambda^* = \lim_{\alpha\to+\infty} \alpha g_+(x^*(\alpha)) \tag{7}$$

$$= \lim_{\alpha\to+\infty} \begin{pmatrix} \max(0, \frac{2\alpha}{(2+\alpha)} \\ \max(0, \frac{4\alpha}{(2+\alpha)} \end{pmatrix} \tag{8}$$

$$= \begin{pmatrix} 2 \\ 4 \end{pmatrix}. \tag{9}$$

All that is left is to show that (x^*, λ^*) is a KKT point by $\nabla_x L(x^*, \lambda^*) = 0$

$$\nabla f(x^*) + \lambda_1^* \nabla g_1(x^*) + \lambda_2^* \nabla g_2(x^*) = \tag{10}$$

$$= \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} -2 \\ -4 \end{pmatrix} \tag{11}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{12}$$

we conclude that $\left(x^* = (0,0)^T, \lambda^* = (2,4)^T\right)$ is a KKT point.

1.2 Exercise 44

Consider the optimization problem

$$\begin{aligned} \min \quad & f(x) := x^2, \\ \text{s.t.} \quad & g(x) := 1 - \ln(x) \leq 0 \end{aligned} \tag{13}$$

and the penalized optimization problem

$$\min_{x\in\mathbb{R}} P(x;\alpha) = f(x) + \alpha\phi\left(\frac{g(x)}{\alpha}\right). \tag{14}$$

with $\phi(t) = e^t - 1$ (*exponential penalty function*). For $\alpha > 0$ find the optimal solution $x^*(\alpha)$ of the penalized optimization problem and prove x^*, the limit of $x^*(\alpha)$ as $\alpha \downarrow 0$ is an optimal solution of the constrained optimization problem.
To find the minimum we differentiate $P(x;\alpha)$ w.r.t x

$$\frac{d}{dx}P(x;\alpha) = \frac{d}{dx}\left(x^2 + \alpha\left(\exp\left(\frac{1-\ln(x)}{\alpha}\right) - 1\right)\right) = \tag{15}$$

$$= 2x - e^{\frac{1}{\alpha}} x^{-\frac{\alpha+1}{\alpha}} \tag{16}$$

setting to 0 give the equation

$$x^{-\frac{\alpha+1}{\alpha}} = 2e^{-\frac{1}{\alpha}} x \tag{17}$$

$$x^*(\alpha) = \left(\frac{1}{2}e^{\frac{1}{\alpha}}\right)^{\frac{\alpha}{2\alpha+1}}. \tag{18}$$

Then

$$x^* = \lim_{\alpha\downarrow 0} x^*(\alpha) = e. \tag{19}$$

First of all $f(x) = x^2$ is strictly convex and the condition $1-\ln(x) \leq 0$ is equivalent the condition $x \geq e$. So there is no $y \in \{x \in \mathbb{R} : x > e\}$ such that $f(y) = y^2 < e^2$. We conclude that $x^* = e$ is the optimal solution for the constrained optimization problem.

1.3 Exercise 45

Consider the optimization problem

$$\begin{aligned} \min \quad & f(x) := x^2, \\ \text{s.t.} \quad & h(x) := x - 1 = 0 \end{aligned} \tag{20}$$

and its optimal solution $x^* = 1$. For $\overline{\alpha} > 0$ such that x^* is a minimum of the ℓ_1-penalty function $P_1(\cdot, \alpha)$ for all $\alpha \geq \overline{\alpha}$. We have that for $\overline{\alpha}$ and x^* and some $x \in \mathbb{R}$ that

$$P_1(x; \overline{\alpha}) < P_1(x^*, \overline{\alpha}) \tag{21}$$
$$x^2 + \overline{\alpha}|x-1| < 1 \qquad \Big| \frac{d}{dx}\cdot \tag{22}$$
$$2x + \overline{\alpha}\frac{x-1}{|x-1|} < 0 \tag{23}$$
$$\alpha < -\frac{2x|x-1|}{x-1} \longrightarrow 2 \quad \text{as } x \downarrow 1 \tag{24}$$

so $\alpha \geq 2$.

1.4 Exercise 46

Consider the optimization problem in Exercise 40

$$\begin{aligned} \min \quad & f(x) := \gamma + c^T x + \frac{1}{2}x^T Q x, \\ \text{s.t} \quad & h(x) := b^T x = 0, \end{aligned} \tag{25}$$

The penalized optimization problem

$$P(x; \alpha) := f(x) + \frac{\alpha}{2}\left(h(x)\right)^2 \tag{26}$$

with solution

$$x^*(\alpha) = \left(\frac{\alpha}{1 + \alpha b^T Q^{-1} b} Q^{-1} b b^T - I\right) Q^{-1} c \tag{27}$$

and solution to the constrained optimization problem

$$x^* = \lim_{\alpha \to \infty} x^*(\alpha) \tag{28}$$
$$= \left(\frac{Q^{-1} b b^T}{b^T Q^{-1} b} - I\right) Q^{-1} c. \tag{29}$$

1.4.1 Part a

Prove that

$$\mu^* := \lim_{\alpha \to +\infty} \alpha h(x^*(\alpha)) \tag{30}$$

is a Lagrange multiplier corresponding to the optimal solution x^*.

$$\alpha h(x^*(\alpha)) = \alpha b^T x^*(\alpha) \tag{31}$$
$$= \alpha\left(\frac{\alpha}{1 + \alpha b^T Q^{-1} b} b^T Q^{-1} b\, b^T - b^T\right) Q^{-1} c \tag{32}$$
$$= \left(\frac{\alpha^2}{1 + \alpha b^T Q^{-1} b} b^T Q^{-1} b - \alpha\right) b^T Q^{-1} c \tag{33}$$
$$= \left(\frac{\alpha^2 b^T Q^{-1} b - \alpha - \alpha^2 b^T Q^{-1} b}{1 + \alpha b^T Q^{-1} b}\right) b^T Q^{-1} c \tag{34}$$
$$= \left(\frac{-\alpha}{1 + \alpha b^T Q^{-1} b}\right) b^T Q^{-1} c \tag{35}$$
$$\tag{36}$$

then we let the $\alpha \to +\infty$ and we get

$$\mu^* = -\frac{b^T Q^{-1} c}{b^T Q^{-1} b}. \tag{37}$$

Now we check if μ^* is the Lagrange multiplier w.r.t x^*.

$$L(x,\mu) = f(x) + \mu h(x) \tag{38}$$
$$= \gamma + c^T x + \frac{1}{2}x^T Q x + \mu b^T x, \tag{39}$$

we need the condition $\nabla L(x^*,\mu^*) = 0$, which is satisfied if

$$\nabla L(x^*,\mu^*) = c + Qx^* + \mu^* b = 0 \qquad \Big| b^T Q^{-1} \tag{40}$$
$$-\mu^* b^T Q^{-1} b = b^T Q^{-1} c + b^T x^*. \tag{41}$$
$$\tag{42}$$

we know that $b^T x^* = 0$ is satisfied then

$$\mu^* = -\frac{b^T Q^{-1} c}{b^T Q^{-1} b}. \tag{43}$$

which is the same as taking the limit.

1.4.2 Part b

A bit confused here.

1.5 Exercise 47

Prove that the following functions are NCP-functions.

1. minimum function

$$\varphi(a,b) = \min\{a,b\} \tag{44}$$

2. Fischer-Burgmeister function

$$\varphi(a,b) = \sqrt{a^2+b^2} - a - b \tag{45}$$

3. penalized minimum function

$$\varphi(a,b) = 2\lambda \min\{a,b\} + (1-\lambda)a_+ b_+ \tag{46}$$

where $a_+ = \max\{0,a\}$, $b_+ = \max\{0,b\}$ and $\lambda \in (0,1)$

For 1. we have that $\min\{a,b\} = 0$ if

$$\Leftrightarrow a = 0 \quad \text{for} \quad b \geq 0 \quad \text{then} \quad ab = 0 \tag{47}$$
$$\Leftrightarrow b = 0 \quad \text{for} \quad a \geq 0 \quad \text{then} \quad ab = 0. \tag{48}$$

The minimum function is an NCP-function
For 2. we have

$$\varphi(a,b) = \sqrt{a^2+b^2} - a - b = 0 \tag{49}$$

then

$$a^2 + b^2 = (a+b)^2, \tag{50}$$

here we need $a \geq 0$ and $b \geq 0$ to preserve the root. Solving the above we get $2ab = 0$ or simply $ab = 0$, which means φ is an NCP-function
For 3. we have that

$$\varphi(a,b) = -2\lambda \min(a,b) + (1-\lambda)\max(0,a)\max(0,b) = 0 \tag{51}$$
$$-2\lambda \min(a,b) = (1-\lambda)\max(0,a)\max(0,b) = 0. \tag{52}$$

The solution is either $a = 0$ with $b \geq 0$ or $b = 0$ with $a \geq 0$ in the first case we get that $a \cdot b = 0$, which means this is an NCP function.

1.6 Exercise 48

Let $(x^*, \lambda^*, \mu^*) \in \mathbb{R}^{n+m+p}$ be a KKT point of the optimization problem.

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \leq 0, i = 1, \dots, m \qquad\qquad h_j(x) = 0, j = 1, \dots, p \end{aligned} \tag{53}$$

all functions are considered to be twice continuously differentiable. Additionally we have that

- $g_i(x^*) + \lambda_i^* \neq 0$ for all $i = 1, \dots, p$
- $\{\nabla g_i(x^*)\}_{i \in \mathcal{A}(x^*)}$ and $\{\nabla h_j(x^*)\}_{j=1,\dots,p}$ are linearly independent (LICQ)
- second order sufficient optimality condition is satisfied

Let $\Phi : \mathbb{R}^{n+m+p} \to \mathbb{R}^{n+m+p}$ be defined as

$$\Phi := \begin{pmatrix} \nabla_x L(x, \lambda, \mu) \\ h(x) \\ \phi(-g(x), \lambda) \end{pmatrix} \tag{54}$$

where

$$\phi(-g(x), \lambda) := \begin{pmatrix} \varphi(-g_1(x), \lambda_1) \\ \vdots \\ \varphi(-g_m(x), \lambda_1) \end{pmatrix} \in \mathbb{R}^m \tag{55}$$

and $\varphi : \mathbb{R}^2 \to \mathbb{R}$ with $\varphi(a, b) = \min\{a, b\}$. Show that the matrix $\nabla\Phi$ is well defined and regular. The matrix is well defined because first of all, the functions f, g_i, h_j are C^2 and $\min\{-g_i(x), \lambda_i\}$ is differentiable because of the strict complementarity condition, meaning that

$$\nabla\varphi(-g_i(x^*, \lambda_i^*) = \begin{cases} -\nabla g_i(x^*) & i \in \mathcal{A}(x^*) \\ 0 & i \notin \mathcal{A}(x^*) \end{cases} \tag{56}$$

Then we need to show that he matrix $\nabla\Phi(x^*, \lambda^*, \mu^*)$ is regular, first of all the matrix has the following form

$$\nabla\Phi = \begin{pmatrix} \nabla_x^2 L(x, \lambda, \mu) & \nabla h(x)^T & \nabla\phi(x)^T \\ \nabla h(x) & 0 & 0 \\ \nabla\phi(x) & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(n+m+p)\times(n+m+p)}. \tag{57}$$

To show that $\nabla\Phi(x^*, \lambda^*, \mu^*)$ is regular we show that $\ker(\nabla\Phi(x^*, \lambda^*, \mu^*)) = \emptyset$.
Let $q = (q^{(1)}, q^{(2)}, q^{(3)})^T \in \mathbb{R}^{n+m+p}$ then we need to find the solution of

$$\nabla\Phi(x^*, \lambda^*, \mu^*) \begin{pmatrix} q^{(1)} \\ q^{(2)} \\ q^{(3)} \end{pmatrix} = 0. \tag{58}$$

These are three equations

$$\nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} + \nabla h(x^*)^T q^{(2)} + \nabla\phi(x^*)^T q^{(3)} = 0 \tag{59}$$
$$\nabla h(x^*) q^{(1)} = 0 \tag{60}$$
$$\nabla\phi(x^*) q^{(1)} = 0. \tag{61}$$

By multiplying 59 with $(q^{(1)})^T$ we get that

$$(q^{(1)})^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} + (q^1)^T \nabla h(x^*)^T q^{(2)} + (q^1)^T \nabla\phi(x^*)^T q^{(3)} = \tag{62}$$

$$= (q^{(1)})^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} + \sum_{j=1}^{p} q_j^{(2)} \underbrace{(q^{(1)})^T \nabla h_j(x^*)}_{=0\ (60)} + \sum_{i=1}^{m} q_i^{(3)} \underbrace{(q^{(1)})^T \nabla\phi(-g_i(x^*), \lambda_i^*)}_{=0\ (61)} \tag{63}$$

$$= 0, \tag{64}$$

in summary

$$(q^{(1)})^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} = 0. \tag{65}$$

Since second order sufficient optimality condition is satisfied then $q^{(1)} \in T_2(x^*)$, and the only solution is $q^{(1)} = 0$. Equation 59 is left with

$$\nabla h(x^*)^T q^{(2)} + \nabla \phi(x^*) q^{(3)} = \tag{66}$$

$$= \sum_{j=1}^{p} q_j^{(2)} \nabla h_j(x^*) + \sum_{i \in \mathcal{A}(x^*)} q_i^{(3)} (-\nabla g_i(x^*)) = 0 \tag{67}$$

since LICQ is fulfilled these vectors are linearly independent and by definition of linear independence the only $q^{(2)}, q^{(3)}$ fulfilling the above condition are $q^{(2)} = 0$ and $q^{(3)} = 0$. Thereby $q = 0$ and $\ker(\nabla \Phi(x^*, \lambda^*, \mu^*)) = \emptyset$, so the matrix is regular.