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

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

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 $1 - \ln(x) \le 0$ is equivalent the condition $x \ge e$. 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

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

so $\alpha \geq 2$.

1.4 Exercise 46

Consider the optimization problem in Exercise 40

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

The penalized optimization problem

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

with solution

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

and solution to the constrained optimization problem

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

1.4.1 Part a

Prove that

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

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

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

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

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

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

$$\begin{aligned} L(x,\mu) &= f(x) + \mu h(x) && (38)\\ &= \gamma + c^T x + \frac{1}{2}x^T Qx + \mu b^T x, && (39)\end{aligned}$$

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

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

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

$$\begin{aligned} &\Leftrightarrow a = 0 \quad \text{for} \quad b \geq 0 \quad \text{then} \quad ab = 0 && (47)\\ &\Leftrightarrow b = 0 \quad \text{for} \quad a \geq 0 \quad \text{then} \quad ab = 0. && (48)\end{aligned}$$

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

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

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.

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

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

- $g_i(x^*) + \lambda_i^* \neq 0$ for all $i = 1, \ldots, p$
- $\{\nabla g_i(x^*)\}_{i \in \mathcal{A}(x^*)}$ and $\{\nabla h_j(x^*)\}_{j=1,\ldots,p}$ 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

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

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

$$\nabla\varphi(a, b) = \begin{cases} (1, 0)^T & a < b \\ (0, 1)^T & a > b \end{cases} \tag{56}$$

and not differentiable for $a = b$ which is never the case since, per condition $g_i(x^*) + \lambda_i^* \neq 0$ for all i.

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

$$\nabla\Phi = \begin{pmatrix} \nabla^2_{xx} L(x^*, \lambda^*, \mu^*) & \nabla g(x^*)^T & \nabla h(x^*)^T \\ \nabla h(x^*) & 0 & 0 \\ \nabla\phi_1(-g(x^*), \lambda^*) & \nabla\phi_2(-g(x^*), \lambda^*) & 0 \end{pmatrix} \in \mathbb{R}^{(n+m+p)\times(n+m+p)}. \tag{57}$$

For convinience the matrix $\phi(-g(x^*), \lambda^*)$ was split into two parts because of the chain rule the that matrix has the following form

$$\phi_1(-g(x^*), \lambda^*) = \begin{pmatrix} \partial_{g_1}\varphi(-g_1(x^*), \lambda_1)(\partial_{x_1} g_1(x^*)) & \ldots & \partial_{g_1}\varphi(-g_1(x^*), \lambda_1)(\partial_{x_n} g_1(x^*)) \\ \vdots & \vdots & \vdots \\ \partial_{g_m}\varphi(-g_m(x^*), \lambda_m)(\partial_{x_1} g_m(x^*)) & \ldots & \partial_{g_m}\varphi(-g_m(x^*), \lambda_m)(\partial_{x_n} g_1(x^*)) \end{pmatrix} \in \mathbb{R}^{n\times m} \tag{58}$$

and

$$\phi_2(-g(x^*), -\lambda^*) = \begin{pmatrix} \partial_{\lambda_1}\varphi(-g_1(x^*), \lambda_1^*) & 0 & \ldots \\ & \ddots & \\ \ldots & 0 & \partial_{\lambda_m}\varphi(-g_m(x^*), \lambda_m^*) \end{pmatrix} \in \mathbb{R}^{m\times m} \tag{59}$$

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

$$\nabla\Phi(x^*, \lambda^*, \mu^*) \begin{pmatrix} q^{(1)} \\ q^{(2)} \\ q^{(3)} \end{pmatrix} = 0. \tag{60}$$

These are three equations

$$\nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} + \nabla g(x^*)^T q^{(2)} + \nabla h(x^*)^T q^{(3)} = 0 \tag{61}$$
$$\nabla h(x^*) q^{(1)} = 0 \tag{62}$$
$$\nabla \phi_1(-g(x^*), \lambda^*) q^{(1)} + \nabla \phi_2(-g(^*), \lambda^*) q^{(2)} = 0. \tag{63}$$

By multiplying 61 with $(q^{(1)})^T$ we get that

$$(q^{(1)})^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} \tag{64}$$
$$+ \sum_{i=1}^{m} q_i^{(3)} \underbrace{(q^{(1)})^T \nabla g_i(x^*)}_{=0 \ (63)} \tag{65}$$
$$+ \sum_{j=1}^{p} q_j^{(2)} \underbrace{(q^{(1)})^T \nabla h_j(x^*)}_{=0 \ (62)} \tag{66}$$
$$= 0.$$

It is not directly obvious why the term 65 in the above equation is directly zero. To see this we have to separate two cases, the first addresses what happens in 63 in the case $i \in \mathcal{A}(x^*)$. In this case $\nabla_{g_i} \varphi(-g_i(x^*), \lambda_i^*) = 1$ since $g_i(x^*) = 0$ with $\lambda_i^* > 0$ and thereby $-g_i(^*) - \lambda_i^* < 0$, so we have that the this specific entry is

$$(\nabla \phi_1(-g(x^*), \lambda^*))_i = -\nabla g_i(x^*) \tag{67}$$
$$(\nabla \phi_2(-g(x^*), \lambda^*))_i = 0, \tag{68}$$

evaluating the equation in 63 we get

$$-\nabla g_i(x^*)^T q^{(1)} = 0 \qquad \forall i \in \mathcal{A}(x^*). \tag{69}$$

In the other case $i \notin \mathcal{A}(x^*)$, $g_i(x^*) < 0$ with $\lambda^* = 0$ and thereby $-g_i(^*) - \lambda_i^* > 0$ so $\varphi(-g_i(x^*), \lambda^*) = \lambda_i^*$. The entries of the matrix are

$$(\nabla \phi_1(-g(x^*), \lambda^*))_i = 0 \tag{70}$$
$$(\nabla \phi_2(-g(x^*), \lambda^*))_i = (0, \dots, 0, \underbrace{1}_{i}, 0, \dots, 0)^T, \tag{71}$$

evaluating the equation in 63 we get

$$q_i^{(2)} = 0 \qquad \forall i \notin \mathcal{A}(x^*). \tag{72}$$

Both cases contribute to the fact that the sum evaluates to 0 in term 65. In summary we are left with

$$(q^{(1)})^T \nabla_x^2 L(x^*, \lambda^*, \mu^*) q^{(1)} = 0. \tag{73}$$

Since second order sufficient optimality condition is satisfied then $q^{(1)} \in T_2(x^*)$, and the only solution is $q^{(1)} = 0$. Equation 61 is left with

$$\nabla g(x^*)^T q^{(2)} + \nabla h(x^*) q^{(3)} = \tag{74}$$
$$= \sum_{i=1}^{m} q_i^{(2)} \nabla g_i(x^*) + \sum_{j=1}^{p} q_j^{(3)} \nabla h_j(x^*) = \tag{75}$$
$$= \sum_{i \in \mathcal{A}(x^*)} q_i^{(2)} \nabla g_i(x^*) + \sum_{j=1}^{p} q_j^{(3)} \nabla h_j(x^*) = 0 \tag{76}$$

in the last equation we remove all $q_i^{(2)} = 0$ which are exactly all $i \notin \mathcal{A}(x^*)$. 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.