Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4180 Optimization Theory Spring 2015

Exercise set 7 solution

1 We drop the asterisk notation and write x and λ for x^* and λ^*. If x is a KKT point, then there exist Lagrange parameters λ such that:

$$\nabla f(x) - \sum_{i=1}^{N} \lambda_i \nabla c_i(x) = 0. \tag{1}$$

To show that $\lambda = [\lambda_1 ... \lambda_N]^T$ is unique if the LICQ holds, we assume the opposite; that there exists another set of Lagrange parameters $\tilde{\lambda} = [\tilde{\lambda}_1, ..., \tilde{\lambda}_N]^T$ such that

$$\nabla f(x) - \sum_{i=1}^{N} \tilde{\lambda}_i \nabla c_i(x) = 0. \tag{2}$$

Subtracting (2) from (1), we see that

$$\sum_{i=1}^{N} (\tilde{\lambda}_i - \lambda_i) \nabla c_i(x) = 0,$$

and especially, since $\lambda_i = \tilde{\lambda}_i = 0$ for $i \notin \mathcal{A}(x)$,

$$\sum_{i \in \mathcal{A}(x)} (\tilde{\lambda}_i - \lambda_i) \nabla c_i(x) = 0. \tag{3}$$

But the LICQ condition states that all $\nabla c_i(x), i \in \mathcal{A}(x)$ are linearily independent, and so 3 holds if and only if $\tilde{\lambda}_i - \lambda_i = 0$, i.e. $\tilde{\lambda}_i = \lambda_i$.

If the LICQ fails, then one can find a_i, of which at least one is nonzero, such that

$$\sum_{i \in \mathcal{A}(x)} a_i \nabla c_i(x) = 0. \tag{4}$$

By choosing $\tilde{\lambda}_i = \lambda_i + a_i$, we find that (3) holds. If, in addition, $\tilde{\lambda}_i \geq 0$, the KKT conditions are satisfied and $\tilde{\lambda}$ is an alternative set of Lagrange parameters, meaning the Lagrange multipliers are non-unique.

2 We begin by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x} \in \mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c(\mathbf{x}) = 0,$$

where

$$f(\mathbf{x}) = xy,$$
$$c(\mathbf{x}) = x^2 + y^2 - 1,$$

We then find the Lagrangian function and its gradient:

$$\mathcal{L}(\mathbf{x}, \lambda) = xy - \lambda(x^2 + y^2 - 1)$$
$$\nabla_x \mathcal{L}(\mathbf{x}, \lambda) = \begin{bmatrix} y - 2\lambda x \\ x - 2\lambda y \end{bmatrix}.$$

The KKT conditions can now be stated as:

$$y - 2\lambda x = 0 \tag{5a}$$
$$x - 2\lambda y = 0 \tag{5b}$$
$$x^2 + y^2 - 1 = 0. \tag{5c}$$

Equations (5a) and (5b) give $\lambda = \pm\frac{1}{2}$, and $y = \pm x$. Plugging this into (5c), we find four candidates for KKT points:

$$\mathbf{x}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \mathbf{x}_2 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \quad \mathbf{x}_3 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \mathbf{x}_4 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

The points $\mathbf{x}_1$ and $\mathbf{x}_2$ have Lagrange parameter $\lambda = \frac{1}{2}$, while $\mathbf{x}_3$ and $\mathbf{x}_4$ have $\lambda = -\frac{1}{2}$. Since $\nabla c(x) = [2x, 2y]^T \neq 0$ at all these points and thereby the LICQ holds, these points are all KKT points.

To check whether the points are minimizers, we consider the second order sufficient conditions. At $\mathbf{x}_1$ and $\mathbf{x}_2$, the critical cone is $\mathcal{C}(\mathbf{x}) = \text{span}\{[-1, 1]^T\}$, while at the points $\mathbf{x}_3$ and $\mathbf{x}_4$ we have $\mathcal{C}(\mathbf{x}) = \text{span}\{[1, 1]^T\}$. We find the Hessian of the Lagrange function:

$$\nabla^2 \mathcal{L}(\mathbf{x}, \lambda) = \begin{bmatrix} -2\lambda & 1 \\ 1 & -2\lambda \end{bmatrix},$$

and find that for $\mathbf{x}_1$ and $\mathbf{x}_2$,

$$\mathbf{w}^T \nabla^2 \mathcal{L}(\mathbf{x}_{1,2}, \frac{1}{2})\mathbf{w} < 0 \quad \forall \mathbf{w} \in \mathcal{C}(\mathbf{x}_{1,2}), \mathbf{w} \neq 0,$$

thereby satisfying the sufficient condition for maximizers. On the other hand,

$$\mathbf{w}^T \nabla^2 \mathcal{L}(\mathbf{x}_{3,4}, \frac{1}{2})\mathbf{w} > 0 \quad \forall \mathbf{w} \in \mathcal{C}(\mathbf{x}_{3,4}), \mathbf{w} \neq 0,$$

such that $\mathbf{x}_3$ and $\mathbf{x}_4$ are in fact local minimizers. Since $f(\mathbf{x}_3) = f(\mathbf{x}_4)$, both $\mathbf{x}_3$ and $\mathbf{x}_4$ are global solutions.

$\boxed{3}$ **a)** We begin by stating the problem in standard form, writing $\mathbf{x} = [x, y]^T$:

$$\min_{\mathbf{x} \in \mathbb{R}^2} f(\mathbf{x}) \quad \text{s.t.} \quad c_i(\mathbf{x}) \geq 0, \quad i = 1, 2$$

where

$$\begin{aligned} f(\mathbf{x}) &= -x^2 - (y-1)^2, \\ c_1(\mathbf{x}) &= y - Cx^2, \\ c_2(\mathbf{x}) &= 2 - y. \end{aligned}$$

We then find the Lagrangian function and its gradient:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \lambda) &= -x^2 - (y-1)^2 - \lambda_1(y - Cx^2) - \lambda_2(2-y) \\ \nabla_x \mathcal{L}(\mathbf{x}, \lambda) &= \begin{bmatrix} -2x + 2Cx\lambda_1 \\ -2y + 2 - \lambda_1 + \lambda_2. \end{bmatrix}. \end{aligned}$$

The KKT conditions can now be stated in full as:

$$-2x + 2Cx\lambda_1 = 0 \tag{6a}$$
$$-2y + 2 - \lambda_1 + \lambda_2 = 0 \tag{6b}$$
$$y - Cx^2 \geq 0 \tag{6c}$$
$$2 - y \geq 0 \tag{6d}$$
$$\lambda_i \geq 0, \quad i = 1, 2 \tag{6e}$$
$$\lambda_1(y - Cx^2) = 0 \tag{6f}$$
$$\lambda_2(2 - y) = 0 \tag{6g}$$

We can now investigate whether (0,0) is a KKT point. We see that (6a), (6c), (6d) and (6f) are satisfied, and that (6g) is satisfied if $\lambda_2 = 0$. This means that (6b) is satisfied when $\lambda_1 = 2$, and thus also (6d) is satisfied. Thereby, (0,0) is a KKT point, regardless of the choice of C.

The only active constraint at (0,0) is c_1, with nonzero gradient

$$\nabla c_1(0,0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Thus, the LICQ holds.

b) Since $\lambda_1 > 0$, and c_1 is the only active constraint at (0,0), the critical cone is given as

$$\mathcal{C}(0,0) = \{w \in \mathbb{R}^2 : w^T \nabla c_1(0,0) = 0\} = \{w \in \mathbb{R}^2 : w = [w_1, 0]^T\}.$$

Also, the Hessian of the Lagrangian function is

$$\nabla^2 \mathcal{L}(0,0,2) = \begin{bmatrix} -2 + 4C & 0 \\ 0 & -4 \end{bmatrix}.$$

The second order necessary condition is that

$$w^T \nabla^2 \mathcal{L}(0,0,2) w \geq 0 \quad \forall w \in \mathcal{C}(0,0),$$

i.e.

$$(-2 + 4C) w_1^2 \geq 0 \quad \forall w_1 \in \mathbb{R}.$$

The second order sufficient condition is that

$$w^T \nabla^2 \mathcal{L}(0,0,2) w > 0 \quad \forall w \in \mathcal{C}(0,0), w \neq 0$$

i.e.

$$(-2+4C)w_1^2 > 0 \quad \forall w_1 \in \mathbb{R}\backslash\{0\}.$$

These conditions are clearly satisfied for $C > \frac{1}{2}$. When $C < \frac{1}{2}$, the necessary condition is not satisfied and (0,0) is not a minimum. When $C = \frac{1}{2}$, (0,0) is not a local minimum as can be seen by approaching along the line $y = \frac{1}{2}x^2$; here, we have

$$\begin{aligned}
\phi(x) &= f(x, \frac{1}{2}x^2) = x^2 - (\frac{1}{2}x^2 - 1)^2 = -\frac{1}{4}x^4 - 1, \\
\phi'(x) &= -x^3, \\
\phi''(x) &= -3x^2,
\end{aligned}$$

meaning (0,0) is a local maximum of the function along this curve. On the other hand, approaching along the line $x = 0$, we have

$$\begin{aligned}
\phi(y) &= f(0,y) = -(y-1)^2, \\
\phi'(y) &= 2 - 2y,
\end{aligned}$$

meaning f is increasing in positive y direction, such that (0,0) is a local minimum along this line. Hence, (0,0) is neither a local minimum nor a local maximum in this case.

$\boxed{4}$ **a)** From exercise set 6, we have

$$\begin{aligned}
\mathcal{L}(x,y,\lambda_1,\lambda_2) &= xy - \lambda_1(y-x) - \lambda_2(x^3 - y^4) \\
\nabla^2 \mathcal{L}(0,0,\lambda_1,\lambda_2) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\
\nabla c_1(0,0) &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
\nabla c_2(0,0) &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
\end{aligned}$$

We also saw that at (0,0), $\lambda_1 = \lambda_2 = 0$. From (12.53) in N&W, we see that

$$\mathcal{C}(0,0) = \{d \in \mathbb{R}^2 : \nabla c_1(0,0)^T d \geq 0\} = \{d \in \mathbb{R}^2 : d_2 \geq d_1\}.$$

From this, we see that

$$d^T \nabla^2 \mathcal{L}(0,0,\lambda_1,\lambda_2) d = 2d_1 d_2,$$

and by choosing e.g. $d_1 = -1$ and $d_2 = 1$, we get

$$d^T \nabla^2 \mathcal{L}(0,0,\lambda_1,\lambda_2) d = -2 < 0.$$

b) From exercise set 6, we had that the tangent cone at (0,0) was given by

$$\mathcal{T}(0,0) = \{d \in \mathbb{R}^2 : d_1 \geq 0 \text{ and } d_2 \geq d_1\}.$$

Now, we still have

$$d^T \nabla^2 \mathcal{L}(0,0,\lambda_1,\lambda_2) d = 2d_1 d_2,$$

but since $d_1 \geq 0$ and thus $d_2 \geq 0$, we have $2d_1 d_2 \geq 0$, and thus

$$d^T \nabla^2 \mathcal{L}(0,0,\lambda_1,\lambda_2) d \geq 0 \quad \forall d \in \mathcal{T}(0,0).$$