Continuous Optimisation: Chpt 3 Exercises

October 10, 2016

1. For a set $\mathcal{F} \subseteq \mathbb{R}^n$, a function $f \in \mathrm{C}^1$, $f : \mathcal{F} \to \mathbb{R}$ and a point $\mathbf{x}_0 \in \mathcal{F}$ consider the following statements:

   1. $\mathbf{x}_0$ is a global minimum.
   2. $\mathbf{x}_0$ is a local minimum.
   3. $\nexists \mathbf{h} \in \mathbb{R}^n$ such that $\nabla f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} < 0$ and $\mathbf{x}_0 + \varepsilon\mathbf{h} \in \mathcal{F}$ for all $\varepsilon > 0$ small enough.
   4. $\nabla f(\mathbf{x}_0)^\mathsf{T}(\mathbf{x} - \mathbf{x}_0) \geq 0$ for all $\mathbf{x} \in \mathcal{F}$.

   Show that in general (1) $\Rightarrow$ (2) $\Rightarrow$ (3) and (3) $\Leftarrow$ (4).

   Also show that if $\mathcal{F}$ is a convex set and f a convex function then (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4).

Solution: We split the proof into the following parts

- <u>(1) $\Rightarrow$ (2):</u> This follows trivially from the definitions.

- <u>(2) $\Rightarrow$ (3):</u> If (3) is not true then there exists $\mathbf{h} \in \mathbb{R}^n$ such that $\nabla f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} < 0$ and $\mathbf{x}_0 + \varepsilon\mathbf{h} \in \mathcal{F}$ for all $\varepsilon > 0$ small enough. We have $f(\mathbf{x}_0 + \varepsilon\mathbf{h}) = f(\mathbf{x}_0) + \varepsilon\nabla f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} + o(\varepsilon)$, and as $\nabla f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} < 0$ this implies that $f(\mathbf{x}_0 + \varepsilon\mathbf{h}) < f(\mathbf{x}_0)$ for all $\varepsilon > 0$ small enough. Therefore $\mathbf{x}_0$ is not a local minimiser. In other words, $\neg$(3) $\Rightarrow$ $\neg$(2), or equivalently (2) $\Rightarrow$ (3). This is true in general and does not use convexity.

- <u>(4) $\Rightarrow$ (3):</u> If (3) is not true then there exists $\mathbf{h} \in \mathbb{R}^n$ such that $\nabla f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} < 0$ and $\mathbf{x}_0 + \varepsilon\mathbf{h} \in \mathcal{F}$ for all $\varepsilon > 0$ small enough. For such a small $\varepsilon > 0$, letting $\mathbf{x} = \mathbf{x}_0 + \varepsilon\mathbf{h} \in \mathcal{F}$ we have $\nabla f(\mathbf{x}_0)^\mathsf{T}(\mathbf{x} - \mathbf{x}_0) = \varepsilon f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} < 0$, and thus (4) is also not true. In other words, $\neg$(3) $\Rightarrow$ $\neg$(4), or equivalently (4) $\Rightarrow$ (3). This is true in general and does not use convexity.

- <u>If convex then (3) $\Rightarrow$ (4):</u> If (4) is not true then there exists $\mathbf{x} \in \mathcal{F}$ such that $\nabla f(\mathbf{x}_0)^\mathsf{T}(\mathbf{x} - \mathbf{x}_0) < 0$. Letting $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$, we have $\nabla f(\mathbf{x}_0)^\mathsf{T}\mathbf{h} = \nabla f(\mathbf{x}_0)^\mathsf{T}(\mathbf{x} -$

$\mathbf{x}_0) < 0$ and $\mathbf{x}_0 + \varepsilon\mathbf{h} = \varepsilon\mathbf{x} + (1-\varepsilon)\mathbf{x}_0 \in \mathcal{F}$ for all $\varepsilon \in (0,1]$. In other words $\neg(4) \Rightarrow \neg(3)$, or equivalently $(3) \Rightarrow (4)$. This is true for general $f \in \mathrm{C}^1$, $f : \mathcal{F} \to \mathbb{R}$ with $\mathcal{F}$ convex.

- If convex then (4) $\Rightarrow$ (1): If (4) is true then f being convex on $\mathcal{F}$ implies that for all $\mathbf{x} \in \mathcal{F}$ we have $f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\mathsf{T}(\mathbf{x} - \mathbf{x}_0) \geq f(\mathbf{x}_0)$, implying that $\mathbf{x}_0$ is a global minimiser and (1) holds. In other words (4) $\Rightarrow$ (1) for convex functions.

2. Show that if (C) is convex problem then $\mathcal{F}$ is convex set.

Solution: If (C) is a convex problem then considering arbitrary $\mathbf{x}, \mathbf{y} \in \mathcal{F}$ and $\theta \in [0,1]$, for all $j = 1, \ldots, m$ we have

$$g_j(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta g_j(\mathbf{x}) + (1-\theta) g_j(\mathbf{y}) \leq 0.$$

Therefore $(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \in \mathcal{F}$, and $\mathcal{F}$ is convex.

As an alternative proof: For all j we have that g_j is a convex function, and thus also a quasiconvex function. Therefore the set $\{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \leq 0\}$ is a convex set. As the intersection of convex sets is convex, this implies that $\mathcal{F} = \bigcap_{j=1}^m \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \leq 0\}$ is a convex set.

3. Show that if $g_1, \ldots, g_m$ are convex functions and $\mathbf{x}, \mathbf{y} \in \mathcal{F}$ then $\nabla g_i(\mathbf{x})^\mathsf{T}(\mathbf{y} - \mathbf{x}) \leq 0$ for all $i \in \mathcal{J}_\mathbf{x}$.

Solution: Consider an arbitrary $i \in \mathcal{J}_\mathbf{x}$. We have $g_i(\mathbf{x}) = 0$ and $g_i(\mathbf{y}) \leq 0$. As g_i is convex, we have $0 \geq g_i(\mathbf{y}) \geq g_i(\mathbf{x}) + \nabla g_i(\mathbf{x})^\mathsf{T}(\mathbf{y} - \mathbf{x}) = \nabla g_i(\mathbf{x})^\mathsf{T}(\mathbf{y} - \mathbf{x})$, completing the proof.

4. For each of the following problems, answer the following questions:

   (a) Is the problem convex?

   (b) Does Slater's condition hold?

   (c) What are the KKT points for this problem?

   (d) What is the global minimiser to this problem?

   1. $\min_\mathbf{x}\{x_1 - x_2 : x_1^2 + x_2^2 \leq 4,\ x_2 \leq 1\}$;
   2. $\min_\mathbf{x}\{x_1 : x_1^2 \leq x_2,\ x_2 \leq 0\}$;
   3. $\min_\mathbf{x}\{x_1 : x_1 + x_2^2 \geq 1,\ x_1^3 \geq 0\}$;

4. $\min_{\mathbf{x}}\{\mathbf{c}^{\mathsf{T}}\mathbf{x} : (\mathbf{x}-\mathbf{a})^{\mathsf{T}}Q(\mathbf{x}-\mathbf{a}) \le 1\}$, where $Q \succ 0$ and $\mathbf{c} \neq \mathbf{0}$.

Solution:

1. (a) We have

$$\begin{aligned} f(\mathbf{x}) &= x_1 - x_2, & \nabla f(\mathbf{x}) &= \begin{pmatrix} 1 \\ -1 \end{pmatrix}, & \nabla^2 f(\mathbf{x}) &= 0, \\ g_1(\mathbf{x}) &= x_1^2 + x_2^2 - 4, & \nabla g_1(\mathbf{x}) &= \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}, & \nabla^2 g_1(\mathbf{x}) &= 2I \succ 0, \\ g_2(\mathbf{x}) &= x_2 - 1, & \nabla g_2(\mathbf{x}) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, & \nabla g_2(\mathbf{x}) &= 0. \end{aligned}$$

Therefore f, g_1, g_2 are convex functions on $\mathbb{R}^2$, and the problem is a convex problem.

(b) Considering $\mathbf{z} = \mathbf{0}$ we have $g_1(\mathbf{z}) = -4 < 0$ and $g_2(\mathbf{z}) = -1 < 0$. Therefore $\mathbf{z}$ is a Slater point and Slater's condition holds.

(c) To satisfy the KKT conditions, we require $\mathbf{x}^*, \boldsymbol{\lambda} \in \mathbb{R}^2$ such that

$$g_1(\mathbf{x}^*) \le 0, \qquad g_2(\mathbf{x}^*) \le 0, \tag{1}$$
$$\lambda_1, \lambda_2 \ge 0, \tag{2}$$
$$\lambda_1 g_1(\mathbf{x}^*) = 0, \qquad \lambda_2 g_2(\mathbf{x}^*) = 0, \tag{3}$$
$$\nabla f(\mathbf{x}^*) = -\lambda_1 \nabla g_1(\mathbf{x}^*) - \lambda_2 \nabla g_2(\mathbf{x}^*). \tag{4}$$

The constraint (4) is equivalent to

$$1 = -2\lambda_1 x_1^*, \qquad -1 = -2\lambda_1 x_2^* - \lambda_2.$$

Therefore $\lambda_1 > 0$, $x_1^* < 0$ and $g_1(\mathbf{x}^*) = 0$.
We now consider two possible cases:

i. $\underline{\lambda_2 = 0}$: Then we have

$$x_1^* = \frac{-1}{2\lambda_1} < 0, \qquad x_2^* = \frac{1}{2\lambda_1} = -x_1^*, \qquad 0 = g_1(\mathbf{x}^*) = 2(x_1^*)^2 - 4.$$

Therefore $\mathbf{x}^* = \begin{pmatrix} -\sqrt{2} & \sqrt{2} \end{pmatrix}^{\mathsf{T}}$, however we then get the contradiction $g_1(\mathbf{x}^*) = \sqrt{2} - 1 > 0$, therefore this can not be a KKT point.

ii. $\underline{\lambda_2 > 0}$: Then we have $0 = g_2(\mathbf{x}^*) = x_2^* - 1$ and thus $x_2^* = 1$ and $0 = g_1(\mathbf{x}^*) = (x_1^*)^2 - 3$. Therefore $\mathbf{x}^* = \begin{pmatrix} -\sqrt{3} & 1 \end{pmatrix}^{\mathsf{T}}$ and $\lambda_1 = \frac{1}{-2\mathbf{x}_1^*} = \frac{1}{2\sqrt{3}}$ and $\lambda_2 = 1 - 2\lambda_1^* x_2^* = 1 + \frac{1}{2\sqrt{3}}$. We can then check that equations (1)–(4) hold for this.

Therefore $\mathbf{x}^* = \begin{pmatrix}-\sqrt{3} & 1\end{pmatrix}^\mathsf{T}$ is the unique KKT point, with corresponding multipliers $\boldsymbol{\lambda} = \begin{pmatrix}\frac{1}{2\sqrt{3}}, & 1 + \frac{1}{2\sqrt{3}}\end{pmatrix}^\mathsf{T}$.

(d) As we have a convex problem with Slater's condition holding, $\mathbf{x}^* = \begin{pmatrix}-\sqrt{3} & 1\end{pmatrix}^\mathsf{T}$ is the unique global minimiser.

2. (a) We have

$$\begin{aligned}
f(\mathbf{x}) &= x_1, & \nabla f(\mathbf{x}) &= \begin{pmatrix}1\\0\end{pmatrix}, & \nabla^2 f(\mathbf{x}) &= 0,\\
g_1(\mathbf{x}) &= x_1^2 - x_2, & \nabla g_1(\mathbf{x}) &= \begin{pmatrix}2x_1\\-1\end{pmatrix}, & \nabla^2 g_1(\mathbf{x}) &= \begin{pmatrix}2 & 0\\0 & 0\end{pmatrix} \succeq 0,\\
g_2(\mathbf{x}) &= x_2, & \nabla g_2(\mathbf{x}) &= \begin{pmatrix}0\\1\end{pmatrix}, & \nabla g_2(\mathbf{x}) &= 0.
\end{aligned}$$

Therefore f, g_1, g_2 are convex functions on $\mathbb{R}^2$, and the problem is a convex problem.

(b) It can be observed that $\mathbf{x} = \mathbf{0}$ is the only point for which both inequalities hold, and this is not a Slater point. Therefore Slater's condition does not hold.

(c) To satisfy the KKT conditions, we require $\mathbf{x}^*, \boldsymbol{\lambda} \in \mathbb{R}^2$ such that

$$g_1(\mathbf{x}^*) \leq 0, \qquad g_2(\mathbf{x}^*) \leq 0, \tag{5}$$
$$\lambda_1, \lambda_2 \geq 0, \tag{6}$$
$$\lambda_1 g_1(\mathbf{x}^*) = 0, \qquad \lambda_2 g_2(\mathbf{x}^*) = 0, \tag{7}$$
$$\nabla f(\mathbf{x}^*) = -\lambda_1 \nabla g_1(\mathbf{x}^*) - \lambda_2 \nabla g_2(\mathbf{x}^*). \tag{8}$$

The constraints (5) imply that $\mathbf{x}^* = \mathbf{0}$, and the constraint (8) is then equivalent to

$$1 = -2\lambda_1 x_1^* = 0, \qquad 0 = -2\lambda_1 - \lambda_2.$$

This is impossible, therefore there are no KKT points.

(d) There is only one feasible point for this problem, and thus it is also the global minimiser, i.e. the global minimiser is $\mathbf{x}^* = \mathbf{0}$.

3. (a) We have

$$\begin{aligned}
f(\mathbf{x}) &= x_1, & \nabla f(\mathbf{x}) &= \begin{pmatrix}1\\0\end{pmatrix}, & \nabla^2 f(\mathbf{x}) &= 0,\\
g_1(\mathbf{x}) &= 1 - x_1 - x_2^2, & \nabla g_1(\mathbf{x}) &= \begin{pmatrix}-1\\-2x_2\end{pmatrix}, & \nabla^2 g_1(\mathbf{x}) &= \begin{pmatrix}0 & 0\\0 & -2\end{pmatrix} \not\succeq 0,\\
g_2(\mathbf{x}) &= -x_1^3, & \nabla g_2(\mathbf{x}) &= \begin{pmatrix}-3x_1^2\\0\end{pmatrix}, & \nabla g_2(\mathbf{x}) &= \begin{pmatrix}-6x_1 & 0\\0 & 0\end{pmatrix}.
\end{aligned}$$

Therefore g_1 is not a convex function on $\mathbb{R}^2$, and the problem is a non-convex problem.

(b) The point $\mathbf{x} = \begin{pmatrix}1 & 1\end{pmatrix}^\mathsf{T}$ is a Slater point.

(c) To satisfy the KKT conditions, we require $\mathbf{x}^*, \boldsymbol{\lambda} \in \mathbb{R}^2$ such that

$$g_1(\mathbf{x}^*) \le 0, \qquad g_2(\mathbf{x}^*) \le 0, \tag{9}$$
$$\lambda_1, \lambda_2 \ge 0, \tag{10}$$
$$\lambda_1 g_1(\mathbf{x}^*) = 0, \qquad \lambda_2 g_2(\mathbf{x}^*) = 0, \tag{11}$$
$$\nabla f(\mathbf{x}^*) = -\lambda_1 \nabla g_1(\mathbf{x}^*) - \lambda_2 \nabla g_2(\mathbf{x}^*). \tag{12}$$

The constraint (12) is equivalent to

$$1 = \lambda_1 + 3\lambda_2 (x_1^*)^2, \qquad 0 = 2\lambda_1 x_2^*.$$

We now consider two possible cases:

i. $\underline{\lambda_1 = 0}$: Then $1 = 3\lambda_2(x_1^*)$, implying that $\lambda_2 > 0$ and $0 = g_2(\mathbf{x}^*) = -x_1^*$. Therefore $x_1^* = 0$. However this then gives the contradiction $1 = \lambda_1 + 3\lambda_2(x_1^*)^2 = 0$.

ii. $\underline{\lambda_1 > 0}$: Then $x_2^* = 0$ and $0 = g_1(\mathbf{x}^*) = 1 - x_1^*$. Therefore $\mathbf{x}^* = \begin{pmatrix}1 & 0\end{pmatrix}^\mathsf{T}$. We then have $g_2(\mathbf{x}^*) = -1 < 0$ and $\lambda_2 = 0$. This then implies that $\boldsymbol{\lambda} = \begin{pmatrix}1 & 0\end{pmatrix}^\mathsf{T}$. We can then check that equations (9)–(12) hold for this

Therefore $\mathbf{x}^* = \begin{pmatrix}1 & 0\end{pmatrix}^\mathsf{T}$ is the unique KKT point, with corresponding multipliers $\boldsymbol{\lambda} = \begin{pmatrix}1 & 0\end{pmatrix}^\mathsf{T}$.

(d) Considering a graph of this problem it can be seen that the set of global minimisers is the set $\{\mathbf{x} \in \mathbb{R}^2 : x_1 = 0,\ |x_2| \ge 1\}$, giving an optimal value of 0. In constrast, for the KKT point we have $f(1,0) = 1 > 0$, and thus the KKT point is not a global minimiser.
N.B. In fact the KKT point is not even a local minimiser, as for $t \in [0,1]$, letting $\mathbf{x}_t = \begin{pmatrix}1 - t^2, & t\end{pmatrix}^\mathsf{T}$, then for all $\varepsilon > 0$ there exists $t \in (0,1]$ such that $\mathbf{x}_t$ is feasible for the problem, $\left\| \mathbf{x}_t - \begin{pmatrix}1 & 0\end{pmatrix}^\mathsf{T} \right\|_2 \le \varepsilon$ and $f(\mathbf{x}_t) = 1 - t^2 < 1 = f(1,0)$.

4. (a) We have

$$f(\mathbf{x}) = \mathbf{c}^\mathsf{T}\mathbf{x}, \qquad \nabla f(\mathbf{x}) = \mathbf{c}, \qquad \nabla^2 f(\mathbf{x}) = 0,$$
$$g(\mathbf{x}) = (\mathbf{x} - \mathbf{a})^\mathsf{T} Q (\mathbf{x} - \mathbf{a}) - 1, \quad \nabla g(\mathbf{x}) = 2Q(\mathbf{x} - \mathbf{a}), \quad \nabla^2 g(\mathbf{x}) = 2Q \succ 0.$$

Therefore f, g are convex functions on $\mathbb{R}^2$, and the problem is a convex problem.

(b) Considering $\mathbf{z} = \mathbf{a}$ we have $g(\mathbf{z}) = -1 < 0$. Therefore $\mathbf{z}$ is a Slater point and Slater's condition holds.

(c) To satisfy the KKT conditions, we require $\mathbf{x}^* \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$ such that

$$
\begin{aligned}
&g(\mathbf{x}^*) \leq 0, && (13)\\
&\lambda \geq 0, && (14)\\
&\lambda g(\mathbf{x}^*) = 0, && (15)\\
&\nabla f(\mathbf{x}^*) = -\lambda \nabla g(\mathbf{x}^*). && (16)
\end{aligned}
$$

The constraint (16) is equivalent to

$$\mathbf{c} = -2\lambda Q(\mathbf{x}^* - \mathbf{a}).$$

Therefore $\lambda > 0$ and $(\mathbf{x}^* - \mathbf{a}) = \frac{-1}{2\lambda} Q^{-1}\mathbf{c}$ and

$$0 = g(\mathbf{x}^*) = \frac{1}{4\lambda^2}\mathbf{c}^\mathsf{T} Q^{-1} Q Q^{-1}\mathbf{c} - 1 = \frac{1}{4\lambda^2}\mathbf{c}^\mathsf{T} Q^{-1}\mathbf{c} - 1.$$

This implies that $\lambda = \frac{1}{2}\sqrt{\mathbf{c}^\mathsf{T} Q^{-1}\mathbf{c}}$ and $\mathbf{x}^* = \mathbf{a} - \frac{1}{\sqrt{\mathbf{c}^\mathsf{T} Q^{-1}\mathbf{c}}} Q^{-1}\mathbf{c}$.

We can then check that equations (13)–(16) hold for this.

Therefore $\mathbf{x}^* = \mathbf{a} - \frac{1}{\sqrt{\mathbf{c}^\mathsf{T} Q^{-1}\mathbf{c}}} Q^{-1}\mathbf{c}$ is the unique KKT point, with corresponding multiplier $\lambda = \frac{1}{2}\sqrt{\mathbf{c}^\mathsf{T} Q^{-1}\mathbf{c}}$.

(d) As we have a convex problem with Slater's condition holding, $\mathbf{x}^* = \mathbf{a} - \frac{1}{\sqrt{\mathbf{c}^\mathsf{T} Q^{-1}\mathbf{c}}} Q^{-1}\mathbf{c}$ is the unique global minimiser.