SECTION C: CONTINUOUS OPTIMISATION
PROBLEM SET 7: SOLUTIONS

HONOUR SCHOOL OF MATHEMATICS, OXFORD UNIVERSITY
HILARY TERM 2006, DR RAPHAEL HAUSER

Solution to Problem 1. (i) The KKT conditions of are the following,

$$-x_2x_3 + \lambda = 0 \tag{0.1}$$
$$-x_1x_3 + 2\lambda = 0 \tag{0.2}$$
$$-x_2x_3 + 2\lambda = 0 \tag{0.3}$$
$$72 - x_1 - 2x_2 - 2x_3 = 0. \tag{0.4}$$

Clearly, x^* satisfies (0.4). Moreover, it follows from (0.1),(0.2),(0.3) that

$$\lambda^* = x_2^*x_3^* = \frac{1}{2}x_1^*x_3^* = \frac{1}{2}x_1^*x_2^* = 144.$$

all of which are satisfied.

(ii) We have

$$Q(x,\mu) = -x_1x_2x_3 + \frac{1}{2\mu}\big(72 - x_1 - 2x_2 - 2x_3\big)^2,$$
$$\nabla_x Q(x,\mu) = \begin{bmatrix} -x_2x_3 - \frac{1}{\mu}(72 - x_1 - 2x_2 - 2x_3) \\ -x_3x_1 - \frac{2}{\mu}(72 - x_1 - 2x_2 - 2x_3) \\ -x_1x_2 - \frac{2}{\mu}(72 - x_1 - 2x_2 - 2x_3) \end{bmatrix}. \tag{0.5}$$

Substituting $x_1 = 2x_2$ and $x_3 = x_2$ into (0.5), all three entries of $\nabla_x Q(x,\mu)$ become

$$-x_2^2 - \frac{1}{\mu}(72 - 6x_2).$$

Setting this expression to zero, we find x_2 by solving a quadratic expression:

$$x_2 = \frac{3}{\mu}\big(1 \pm \sqrt{1-8\mu}\big) = \frac{3\big(1 \pm \sqrt{1-8\mu}\big)\big(1 \mp \sqrt{1-8\mu}\big)}{\mu\big(1 \mp \sqrt{1-8\mu}\big)} = \frac{24}{1 \mp \sqrt{1-8\mu}}.$$

Thus, the expression given for $x(\mu)$ satisfies the claim. Moreover, clearly,

$$x_2(\mu) \to 12,$$

and then $x_1(\mu) \to 24$, $x_3(\mu) \to 12$.

(iii) $x_2(\mu) = x_3(\mu) = 24/(1 + 1/3) = 18$, $x_1(\mu) = 36$. Moreover,

$$D^2_{xx}Q(x,\mu) = \begin{bmatrix} \frac{1}{\mu} & -x_3 + \frac{2}{\mu} & -x_2 + \frac{2}{\mu} \\ -x_3 + \frac{2}{\mu} & \frac{4}{\mu} & -x_1 + \frac{4}{\mu} \\ -x_2 + \frac{2}{\mu} & -x_1 + \frac{4}{\mu} & \frac{4}{\mu} \end{bmatrix},$$

so that

$$D^2_{xx}Q(x(1/9),1/9) = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 36 \end{bmatrix} \succ 0.$$

(iv) We have

$$-\frac{g(x(\mu))}{\mu} = -\frac{72 - 6\frac{24}{1+\sqrt{1-8\mu}}}{\mu} = \frac{72}{1+\sqrt{1-8\mu}} \times \frac{1-\sqrt{1-8\mu}}{\mu}.$$

Clearly,

$$\lim_{\mu\to 0} \frac{72}{1+\sqrt{1-8\mu}} = 36.$$

Moreover, the rule of Bernoulli-Hôpital implies

$$\lim_{\mu\to 0} \frac{1-\sqrt{1-8\mu}}{\mu} = \lim_{\mu\to 0} \frac{\frac{8}{2\sqrt{1-8\mu}}}{1} = 4.$$

Hence,

$$\lim_{\mu\to 0} -\frac{g(x(\mu))}{\mu} = 36 \times 4 = \lambda^*.$$

Solution to Problem 2. (i) The KKT conditions are

$$-1 + 2\lambda x_1 = 0 \tag{0.6}$$
$$-1 + 2\lambda x_2 = 0 \tag{0.7}$$
$$1 - x_1^2 - x_2^2 = 0. \tag{0.8}$$

If $\lambda = 0$ then (0.6) and (0.7) are violated, so there are no solutions corresponding to this case. If $\lambda \neq 0$ then $x_1 = x_2 = 1/(2\lambda)$, thus (0.8) implies that the KKT points are (x^*,λ^*) and $(-x^*,-\lambda^*)$, where $\lambda^* = x_1^* = x_2^* = 1/\sqrt{2}$.

(ii) We have

$$Q(x,\mu) = -x_1 - x_2 + \frac{1}{2\mu}(1 - x_1^2 - x_2^2)^2.$$

The stationary points of the problem

$$\min_{x\in\mathbb{R}^2} Q(x,\mu)$$

satisfy

$$\nabla_x Q(x,\lambda) = \begin{bmatrix} -1 - \frac{2x_1}{\mu}(1 - x_1^2 - x_2^2) \\ -1 - \frac{2x_2}{\mu}(1 - x_1^2 - x_2^2) \end{bmatrix} = 0,$$

which implies

$$\mu = -2x_1(1 - x_1^2 - x_2^2) = -2x_2(1 - x_1^2 - x_2^2). \tag{0.9}$$

Since $\mu > 0$, we have $1-x_1^2-x_2^2 \neq 0$, and hence (0.9) shows that $x_1 = x_2$. Substituting this back into (0.9), we find $2x_1^3(\mu) - x_1(\mu) - \mu/2 = 0$.

(iii) We have

$$x_1(1 - 2x_1^2) + \mu/2 = 0. \tag{0.10}$$

So, as $\mu \to 0$, it must be the case that $x_1(1 - 2x_1^2) \to 0$. Since $x_1 \to 0$ would imply that $x_2 \to 0$, and hence the penalty term would blow up, this shows that

$$x_1(\mu) \overset{\mu\to 0}{\longrightarrow} \frac{1}{\sqrt{2}} = x_1^*.$$

Using the ansatz $x_1(\mu) = 1/\sqrt{2} + a\mu + O(\mu^2)$ where g is some function, (0.10) implies

$$\frac{1 - 2\big(1/\sqrt{2} + a\mu + O(\mu^2)\big)^2}{\mu} = -\frac{1}{2x_1(\mu)} \overset{\mu\to 0}{\longrightarrow} -\frac{1}{\sqrt{2}}.$$

Expanding the left hand side, we find

$$-\frac{1}{\sqrt{2}} = \lim_{\mu\to 0} \frac{-2\sqrt{2}a\mu + O(\mu^2)}{\mu} = -2\sqrt{2}a,$$

which shows that $a = 1/4$.

(iv) The new penalty function is

$$\tilde{Q}(x, \mu) = Q(x, \mu) + \frac{1}{2\mu}\tilde{g}^2(x),$$

where

$$\tilde{g}(x) = \begin{cases} x_2 - x_1^2 & \text{if } x_1^2 > x_2, \\ 0 & \text{otherwise.} \end{cases}$$

To find the minimisers of $\tilde{Q}$ we solve the first order conditions,

$$\nabla_x \tilde{Q}(x, \mu) = \nabla_x Q(x, \mu) + \frac{\tilde{g}(x)}{\mu} \begin{bmatrix} -2x_1 \\ 1 \end{bmatrix} = 0.$$

Since we want x to be a local minimiser of $Q(x, \mu)$ at the same time, we also need $\nabla_x(Q, \mu) = 0$, so that the criterion becomes

$$\begin{aligned} -2x_1\tilde{g}(x) &= 0 \\ \tilde{g}(x) &= 0, \end{aligned}$$

which is satisfied if and only if the second equation is satisfied. That is, we are looking for the values of μ for which $x_2(\mu) - x_1^2(\mu) \geq 0$, and since $x_2(\mu) = x_1(\mu)$, this translates into

$$\frac{1}{\sqrt{2}} + \frac{\mu}{4} + O(\mu^2) - \Big(\frac{1}{\sqrt{2}} + \frac{\mu}{4} + O(\mu^2)\Big)^2 \geq 0,$$

which is equivalent to

$$\frac{\sqrt{2}-1}{2}\Big(1-\frac{\mu}{2}-O(\mu^2)\Big)\geq 0.$$

But this equation is clearly satisfied for all μ small enough, thus, there exists a value $\bar{\mu}>0$ such that the minimisers of $Q(x,\mu)$ and $\tilde{Q}(x,\mu)$ coincide for all $\mu\in(0,\bar{\mu}]$.

Solution to Problem 3. (i) Looking at the last line, we find

$$\begin{aligned}
&x_1, x_2 \approx \frac{1}{\sqrt{2}} = 0.7071\\
&x_3 \approx 0.
\end{aligned}$$

Thus, we are led to suspect that $x^* = (1/\sqrt{2}, 1/\sqrt{2}, 0)$. Moreover, we have

$$\begin{aligned}
g_1(x) = x_3 \geq 0 &\qquad \text{active,}\\
g_2(x) = 1 - x_3 \geq 0 &\qquad \text{inactive,}\\
g_3(x) = 1 - x_1^3 - x_3 \geq 0 &\qquad \text{inactive,}\\
g_4(x) = 1 - x_1^2 - x_2^2 - x_3^2 \geq 0 &\qquad \text{active.}
\end{aligned}$$

This is consistent with our guess for x^*, since $\mathcal{A}(x^*) = \{1,4\}$. Furthermore, the theory says

$$\begin{aligned}
&\lambda_1^* \approx -\frac{\tilde{g}_1(x)}{\mu_4} = 1.017 \approx 1\\
&\lambda_2^* \approx -\frac{\tilde{g}_2(x)}{\mu_4} = 0\\
&\lambda_3^* \approx -\frac{\tilde{g}_3(x)}{\mu_4} = 0\\
&\lambda_4^* \approx -\frac{\tilde{g}_4(x)}{\mu_4} = 0.7061 \approx \frac{1}{\sqrt{2}}.
\end{aligned}$$

Thus, we suspect that $\lambda^* = (1,0,0,1/\sqrt{2})$.

(ii) The KKT conditions are

$$\begin{aligned}
-1 + 3\lambda_3 x_1^2 + 2\lambda_4 x_1 &= 0, \\
-1 - 2\lambda_4 x_2 &= 0, \\
1 - \lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4 x_3 &= 0, \\
x_3 &\geq 0, \\
1 - x_3 &\geq 0, \\
1 - x_1^3 - x_3 &\geq 0, \\
1 - x_1^2 - x_2^2 - x_3^2 &\geq 0, \\
\lambda_1(x_3) &= 0, \\
\lambda_2(1 - x_3) &= 0, \\
\lambda_3(1 - x_1^3 - x_3) &= 0, \\
\lambda_4(1 - x_1^2 - x_2^2 - x_3^2) &= 0, \\
\lambda_1, \ldots, \lambda_4 &\geq 0.
\end{aligned}$$

It is easily checked that (x^*, λ^*) satisfies all these equations and inequalities.

(iii) The theory says

$$|g_1(x)| = O(\mu_4), \tag{0.11}$$

$$|g_4(x)| = O(\mu_4). \tag{0.12}$$

It follows from (0.11) that $x_3 = 0$ to about 3 digits of accuracy, which is consistent with the value from the table. In other words, our hunch that $x_3^* = 0$ is not contradicted by the data.

Furthermore, let $\delta_1 = x_1(\mu_4) - x_1^*$ and $\delta_2 = x_2(\mu_4) - x_2^*$. If our hunch that $x_1^* = x_2^* = 1/\sqrt{2}$ is true, then $\delta_1, \delta_2 > 0$, and it follows from (0.12) that the following should hold,

$$O(\mu_4) = 1 - \left(\frac{1}{\sqrt{2}} + \delta_1\right)^2 - \left(\frac{1}{\sqrt{2}} + \delta_2\right)^2 - O(\mu_4^2) = -\sqrt{2}(\delta_1 + \delta_2) - \delta_1^2 - \delta_2^2 - O(\mu_4^2).$$

Thus,

$$\delta_1 + \delta_2 = O(\mu_4) = O(10^{-3}),$$

and since δ_1 and δ_2 are of the same sign, this implies that it should be the case that

$$\delta_1, \delta_2 = O(\mu_4) = 10^{-3}.$$

Indeed, this requirement is consistent with the data. Thus, we do not reject the hunch. This is further confirmed by part (ii), where we showed that (x^*, λ^*) is a KKT point.