Solutions to Part A of Problem Sheet 6

Solution (6.1) The claim is that the neighbourhood

$$\mathcal{N}_{-\infty}(1) = \{(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \in \mathcal{F}^{\circ} : x_i s_i \geq \mu\}$$

coincides with the central path $\mathcal{C}$. Clearly, since μ is the *average* of the $x_i s_i$, $x_i s_i \geq \mu$ for all i must imply $x_i s_i = \mu$ for all i (we can't all be better or equal than average, unless we are all equal). But then, such a vector is clearly on the central path. Conversely, if $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \in \mathcal{C}$, then there exists a $\tau > 0$ such that $x_i s_i = \tau$ for all i. But then, $\mu = \frac{1}{n}\sum_{i=1}^n x_i s_i = \frac{1}{n}\sum_{i=1}^n \tau = \tau = x_i s_i$ for all i, so that $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) \in \mathcal{N}_{-\infty}(1)$.

Solution (6.2) The problem is of the form

$$\text{minimize } \langle \boldsymbol{c}, \boldsymbol{x} \rangle \quad \text{subject to } \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0},$$

with $\boldsymbol{A} = (1, -1)$, $\boldsymbol{b} = 2$, $\boldsymbol{c} = (0, 1)^{\top}$. The optimality conditions are

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} &= \boldsymbol{0} \\ \boldsymbol{A}^{\top}\boldsymbol{y} + \boldsymbol{s} - \boldsymbol{c} &= \boldsymbol{0} \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} &= \boldsymbol{0} \\ \boldsymbol{x} &\geq \boldsymbol{0} \\ \boldsymbol{s} &\geq \boldsymbol{0}. \end{aligned}$$

In our specific case, this translates to

$$\begin{aligned} x_1 - x_2 - 2 &= 0 \\ y + s_1 &= 0 \\ -y + s_2 - 1 &= 0 \\ x_1 s_1 &= 0 \\ x_2 s_2 &= 0 \\ x_1, x_2, s_1, s_2 &\geq 0. \end{aligned} \tag{1}$$

(2 marks) The central path is the set $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s})$ that arises when we change the conditions $x_i s_i = 0$ in (1) to $x_i s_i = \tau$, for $\tau > 0$. To compute the central path, we first eliminate x_2 and s_2 from (1):

$$x_2 = x_1 - 2, \quad s_2 = y + 1 = 1 - s_1.$$

The conditions $x_i s_i = \tau$ then become

$$x_1 s_1 = \tau, \quad (x_1 - 2)(1 - s_1) = \tau.$$

(2 marks) Multiplying out the second of these and replacing $x_1 s_1$ with τ and s_1 with τ / x_1 in the resulting expression gives

$$-\tau - 2 + 2\frac{\tau}{x_1} + x_1 = \tau \iff x_1^2 - 2(\tau + 1) + 2\tau = 0.$$

Solving this quadratic equation gives the solution

$$x_1 = \tau + 1 \pm \sqrt{\tau^2 + 1}.$$

From this we get the expression for x_2,

$$x_2 = x_1 - 2 = \tau - 1 \pm \sqrt{\tau^2 + 1}. \tag{1}$$

Since $(\tau - 1)^2 = \tau^2 - 2\tau + 1 < \tau^2 + 1$ for $\tau > 0$, we get that only one (the one with $+$) of the solutions in (1) is positive, and thus valid. (2 marks)

The central path is therefore defined by

$$\begin{aligned}
x_1 &= \tau + 1 + \sqrt{\tau^2 + 1} \\
x_2 &= \tau - 1 + \sqrt{\tau^2 + 1} \\
s_1 &= \frac{\tau}{x_1} \\
s_2 &= \frac{\tau}{x_2} \\
y &= s_2 - 1.
\end{aligned}$$

In particular, every $\tau > 0$ determines a unique point on the central path.

Solution (6.3) The central path is defined as the set of solutions of

$$\begin{aligned}
a_1 y + s_1 - c_1 &= 0 \\
a_2 y + s_2 - c_2 &= 0 \\
a_1 x_1 + a_2 x_2 - b &= 0 \\
x_1 s_1 &= \tau \\
x_2 s_2 &= \tau \\
x_1, x_2, s_1, s_2 &> 0
\end{aligned}$$

for $\tau > 0$. We want to derive conditions under which there is exactly one solution to this system. In a first step, we eliminate x_1 and x_2 by setting

$$x_1 = \frac{\tau}{s_1}, \quad x_2 = \frac{\tau}{s_2}.$$

Note that this can be done, as we assumed $s_1 > 0$ and $s_2 > 0$. We can therefore eliminate x_1 and x_2 from the third equations,

$$a_1 \frac{\tau}{s_1} + a_2 \frac{\tau}{s_2} - b = 0 \Leftrightarrow a_1 \tau s_2 + a_2 \tau s_1 - b s_1 s_2 = 0, \tag{0.1}$$

where we multiplied the whole thing by $s_1 s_2$, and consider the first three equations in s_1 and s_2. We can now use the first two equations to express s_1 in terms of s_2. Assuming that $a_2 \neq 0$, we get

$$y = \frac{c_2 - s_2}{a_2}$$

from the second equation, and plugging into the first one,

$$a_1\left(\frac{c_2-s_2}{a_2}\right)+s_1-c_1=0 \Rightarrow s_1=c_1-\frac{a_1}{a_2}(c_2-s_2).$$

Replacing s_1 in (0.1) with the above and simplifying things slightly, we arrive at a quadratic equation in s_2,

$$\frac{ba_1}{a_2}s_2^2+\left(b\left(c_1-\frac{a_1}{a_2}c_2\right)-2\tau a_1\right)s_2+\tau a_2\left(\frac{a_1}{a_2}c_2-c_1\right)=0.$$

This can be simplified as $As_2^2+Bs_2+C=0$, with

$$\begin{aligned} A &= \frac{ba_1}{a_2} \\ B &= \left(b\left(c_1-\frac{a_1}{a_2}c_2\right)-2\tau a_1\right) \\ C &= -\tau a_2\left(c_1-\frac{a_1}{a_2}c_2\right). \end{aligned}$$

Using the formula for the solution of a quadratic equation, we get two solutions

$$\begin{aligned} s_2^+ &= \frac{-B+\sqrt{B^2-4AC}}{2A} \\ s_2^- &= \frac{-B-\sqrt{B^2-4AC}}{2A} \end{aligned}$$

The term under the square root is always positive in our case, so we always have two real solutions. Moreover, we see that if $4AC<0$, then exactly one of the solutions is positive and one is negative (as the term under the square root is then always bigger than $|B|$). The condition $4AC<0$ translates to

$$ba_1\left(c_1-\frac{a_1c_2}{a_2}\right)>0,$$

where we removed the factor 4τ because this is always positive, which is satisfied if $a_1,a_2,b>0$ and $a_2c_1>c_2a_1$. What remains to be done is to check if also $s_1>0$ under these conditions. This, however, is the case since

$$s_1=c_1-\frac{a_1}{a_2}c_2+\frac{a_1}{a_2}s_2,$$

and $c_1-\frac{a_1}{a_2}c_2>0$ by assumption. Clearly, given $\tau>0$, we then also have $x_1>0$ and $x_2>0$. This shows that there is a unique point on the central path, given by the positive solution to a quadratic equation.

Solution (6.4) If $\boldsymbol{B}$ is the matrix with the $-y_i\boldsymbol{x}_i$ as rows, and $\boldsymbol{y}$ denotes the vector with the y_i as entries, define $\boldsymbol{A}=[\boldsymbol{B},-\boldsymbol{y}]$. Let $\boldsymbol{z}=(\boldsymbol{w},b)\in\mathbb{R}^{n+1}$. Then any solution of the feasibility problem

$$\boldsymbol{A}\boldsymbol{z}<0$$

has the property that $\langle \boldsymbol{w}, \boldsymbol{x}_i \rangle + b < 0$ if $y_i = -1$ and $\langle \boldsymbol{w}, \boldsymbol{x}_j \rangle + b > 0$ if $y_i = 1$, so we get the separating hyperplane from $\boldsymbol{w}, b$. If the convex hulls of the $\boldsymbol{x}_i$ with $y_i = 1$ and of the $\boldsymbol{x}_j$ with $y_j = -1$ are disjoint, we have two disjoint, bounded, closed convex sets, and these can be separated by a hyperplane. The existence of a separating hyperplane is equivalent to the linear programming feasibility problem having a solution.