SECTION C: CONTINUOUS OPTIMISATION
PROBLEM SET 4: SOLUTIONS

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

Solution to Problem 1: (i) Since $B_k \succ 0$, m_k is strictly convex. Therefore, $m_k|_{\{x_k - t\nabla f(x_k):\, t\in\mathbb{R}\}}$ is strictly convex, and since y_k^u is the global minimiser of this restricted function, m_k is strictly decreasing along the first section of the dogleg path. Likewise, the restriction

$$m_k|_{\{y_k^u+\tau(y_k^{qn}-y_k^u):\tau\in\mathbb{R}\}}$$

is strictly convex, and since y_k^{qn} is the global minimiser of m_k, it must also be the global minimiser of this restricted function. Therefore, m_k is strictly decreasing along the second section of the dogleg path.

(ii) It suffices to show that $y_k^u - x_k$ and $y_k^{qn} - y_k^u$ form an acute angle. Since $y_k^u - x_k \sim -\nabla f(x_k)$ up to a positive scalar, this is the same as showing that

$$\langle -\nabla f(x_k), y_k^{qn} - y_k^u\rangle > 0. \tag{0.1}$$

Now, using the notation of the hint, note that since

$$-\nabla f(x_k) = -\nabla m_k(x_k) \in 2D$$

and ϕ is the restriction of m_k to $2D + x_k$, it must be the case that $-\nabla f(x_k) = -\nabla\phi(0) = d_0$. Moreover, since the conjugate gradient algorithm takes exact line-search steps, the next update is $z_0 = y_k^u - x_k$. At that stage the remaining search space is one-dimensional and the conjugate gradient algorithm moves to the exact minimiser $z^* = z_1 + \alpha d_1$ of ϕ, where α is a positive number. But since the global minimiser y_k^{qn} is a member of $2D + x_k$, it must be the case that $y_k^{qn} - x_k = z^*$. Therefore,

$$\alpha d_1 = z^* - z_1 = (y_k^{qn} - x_k) - (y_k^u - x_k),$$

which shows that $y_k^{qn} - y_k^u = \alpha^{-1} d_1$. The relation (0.1) now follows from our proof of Theorem 3.4, Lecture 7.

(iii) This is trivial, because in that case the global minimiser y_k^{qn} lies in the trust region.

(iv) This is a trivial consequence of parts (i) and (ii).

(v) The second formula is trivial. Using the hint to prove the first formula, note that $x = y_k^{qn}$ is the unique point where the condition $\nabla m_k(x) = 0$ is satisfied. Therefore, denoting the number λ obtained in Equation (0.1) of the problem statement by $\lambda(\Delta)$, we must have $\lambda(\Delta) > 0$ for $\Delta < \|y_k^{qn} - x_k\|$. Moreover, since ∇m_k is continuous and R_k shrinks down to x_k, we have

$$\lim_{\Delta\to 0} \nabla m_k(x(\Delta)) = \lim_{x\to x_k} \nabla m_k(x) = \nabla m_k(x_k), \tag{0.2}$$

and hence,

$$\lim_{\Delta\to 0+} \Delta \times \lambda(\Delta) \overset{(iv)}{=} \lim_{\Delta\to 0+} \|x(\Delta) - x_k\| \times \lambda(\Delta) = \lim_{\Delta\to 0+} \|\nabla m_k(x(\Delta))\| = \|\nabla m_k(x_k)\|.$$

Together with Equation (0.1) from the problem statement and (0.2) this shows

$$\lim_{\Delta\to 0+} \frac{x(\Delta) - x_k}{\Delta} = \lim_{\Delta\to 0+} \frac{\lambda(\Delta)(x(\Delta) - x_k)}{\Delta \times \lambda(\Delta)} = \frac{-\nabla m_k(x_k)}{\|\nabla m_k(x_k)\|} = \frac{-\nabla f(x_k)}{\|\nabla f(x_k)\|},$$

as claimed.

Solution to Problem 2: (i) Using the hint, we find that $Ax \leq b$ has a solution if and only if b lies in the cone generated by the columns of A'. The fundamental theorem of linear inequalities says that this occurs if and only if there does not exist $y \in \mathbb{R}^m$ such that $A'^{\mathrm{T}} y \geq 0$ and $b^{\mathrm{T}} y < 0$. But

$$A'^{\mathrm{T}} y \geq 0 \Leftrightarrow (y \geq 0) \wedge (A^{\mathrm{T}} y = 0).$$

Thus, we have shown that $Ax \leq b$ has a solution if and only if $A^{\mathrm{T}} y = 0$, $y \geq 0$, $b^{\mathrm{T}} y < 0$ has no solution.

(ii) Clearly, if $x \geq 0$, $Ax = b$ and $A^{\mathrm{T}} y \geq 0$ then

$$b^{\mathrm{T}} y = (Ax)^{\mathrm{T}} y = x^{\mathrm{T}}(A^{\mathrm{T}} y) \geq 0.$$

This shows that (I)$\Rightarrow$(II).

On the other hand, suppose that (I) fails to hold, that is, $\nexists x \geq 0$ such that $Ax = b$. This means that $b \notin \mathrm{cone}(a_1, \dots, a_m)$, where a_i is the i-th column of A. The fundamental theorem of linear inequalities says that in this case $\exists\, y \in \mathbb{R}^m$ such that $A^{\mathrm{T}} y \geq 0$ and $b^{\mathrm{T}} y < 0$. This shows that $\neg$(I)$\Rightarrow \neg$(II).

(iii) This follows immediately from (ii) applied to the matrix $A = -\left[\begin{smallmatrix} a_1 & \cdots & a_m \end{smallmatrix}\right]^{\mathrm{T}}$ and $b = -a_0$, and where the roles of n and m are exchanged.

Solution to Problem 3: (i) We have $c^{\mathrm{T}} x = y^{\mathrm{T}} A x \leq y^{\mathrm{T}} b$, where the last inequality follows from Equation (0.3) of the problem statement and the fact that $y \geq 0$.

(ii) Equation (0.2) of the problem statement shows that y^* is dual feasible, and Equation (0.3) implies that x^* is primal feasible. If x^* is not primal optimal, then there exists a primal feasible point x such that

$$c^{\mathrm{T}} x^* < c^{\mathrm{T}} x \overset{(i)}{\leq} b^{\mathrm{T}} y^* \overset{(0.4)}{=} c^{\mathrm{T}} x^*.$$

Since this is a contradiction, x^* is primal optimal after all. The proof that y^* is dual optimal is analogous.

(iii) The fundamental theorem of linear inequalities shows that exactly one of the following two alternatives occur:

(I) $\exists\, \{y_i \geq 0 : i \in J\}$ such that $c = \sum_{i \in J} y_i a_i$.

(II) $\exists\, d \in \mathbb{R}^n$ such that $d^{\mathrm{T}} a_i \geq 0$ $(i \in J)$ and $d^{\mathrm{T}} c < 0$.

We claim that (I) holds. Suppose to the contrary that (II) holds. In this case,

$$\begin{aligned}
&a_i^{\mathrm{T}}(x^* - td) \leq b_i - t a_i^{\mathrm{T}} d \leq b_i, && (0 \leq t,\ i \in J),\\
&a_i^{\mathrm{T}}(x^* - td) \leq b_i, && \Big(0 \leq t < \min_{\{i \notin J : a_i^{\mathrm{T}} d < 0\}} -\frac{b_i - a_i^{\mathrm{T}} x^*}{a_i^{\mathrm{T}} d},\ i \notin J\Big),
\end{aligned}$$

that is, $x^* - td$ is feasible for small positive t, and then

$$c^{\mathrm{T}}(x^* - td) = c^{\mathrm{T}} x^* - t c^{\mathrm{T}} d > c^{\mathrm{T}} x^*$$

contradicts the assumption that x^* is a maximiser of (P).

(iv) For $i \neq J$ set $y_i^* = 0$ and for $i \in J$ set $y_i^* = y_i$. Then (I) shows that $A^{\mathrm{T}} y^* = c$. Since $y^* \geq 0$, this shows that y^* is dual feasible. Moreover, we have

$$\begin{aligned}
c^{\mathrm{T}} x^* &= \sum_{i=1}^{m} y_i^* (a_i^{\mathrm{T}} x^*) = \sum_{i \in J} y_i (a_i^{\mathrm{T}} x^*) + \sum_{i \notin J} 0 \times (a_i^{\mathrm{T}} x^*)\\
&= \sum_{i \in J} y_i b_i + \sum_{i \notin J} 0 \times b_i = \sum_{i=1}^{m} y_i^* b_i.
\end{aligned}$$

This shows that Equation (0.4) of the problem statement holds, and the remaining claims follow from part (ii).

(v) Part (iv) shows that the existence of a primal optimal solution x^* implies the existence of a dual optimal solution y^*. On the other hand, exchanging the roles of (P) and (D) we obtain the inverse implication. Since in this case, part (iv) also shows that (0.4) holds: the duality gap at the optimal solutions is zero.