SECTION C: CONTINUOUS OPTIMISATION
PROBLEM SET 4

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

***Problem 1.** Let B_k be a positive definite symmetric matrix. Prove the claims of Lemma 3.1, Lecture 7:

(i) Show that the model function m_k is strictly decreasing along the path $y(\tau)$.

(ii) Show that $\|y(\tau)-x_k\|$ is strictly increasing along the path $y(\tau)$. Hint: consider the space $2D = \text{span}\{-\nabla f(x_k), y_k^{qn} - y_k^u\}$ and the strictly convex quadratic function

$$\begin{aligned}\phi : 2D &\to \mathbb{R}\\ z &\mapsto m_k(x_k + z).\end{aligned}$$

Show that if the conjugate gradient algorithm is applied to the minimisation of $\phi(z)$ starting from $z_0 = 0$, then $d_0 = -\nabla f(x_k)$ and $d_1 = \alpha^{-1}(y_k^{qn} - y_k^u)$ for some $\alpha > 0$. Then use the fact that in the proof of Theorem 3.4 of Lecture 7 we showed that $d_0^{\mathrm{T}} d_1 > 0$.

(iii) Show that if $\Delta \geq \|B_k^{-1}\nabla f(x_k)\|$ then $y(\Delta) = y_k^{qn}$.

(iv) Prove that if $\Delta \leq \|B_k^{-1}\nabla f(x_k)\|$ then $\|y(\Delta) - x_k\| = \Delta$.

(v) Derive the following limits:

$$\begin{aligned}\lim_{\Delta\to 0+} \frac{x(\Delta) - x_k}{\Delta} &= -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}\\ \lim_{\tau\to 0+} \frac{y(\tau) - y(0)}{\tau} &= \frac{-\|\nabla f(x_k)\|^2}{\nabla f(x_k)^{\mathrm{T}} B_k \nabla f(x_k)} \nabla f(x_k)\end{aligned}$$

Hint: you may use the fact that a necessary condition for $x = x(\Delta)$ is that there exists $\lambda \geq 0$ such that

$$-\nabla m_k(x) = \lambda(x - x_k). \tag{0.1}$$

***Problem 2.**

(i) Farkas' lemma is the following result: let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a matrix and let $\mathbf{b} \in \mathbb{R}^m$ be a vector. Then exactly one of the following two situations occurs:

(I) $\exists\, x \in \mathbb{R}^n$ such that $Ax \leq b$.
(II) $\exists\, y \in \mathbb{R}^m_+$ such that $A^{\mathrm{T}} y = 0$ and $y^{\mathrm{T}} b < 0$.

Prove this result using the fundamental theorem of linear inequalities. Hint: note that if A' is the matrix $\begin{bmatrix} I & A & -A \end{bmatrix}$ then $Ax \leq b$ has a solution if and only if $A'x' = b$ has a solution $x' \geq 0$. Use the fundamental theorem of linear inequalities.

(ii) Sometimes the Farkas lemma is formulated in an equivalent form which says that the following conditions are equivalent:

(I) $\exists\, x \in \mathbb{R}^n_+$ such that $Ax = b$.
(II) $A^{\mathrm{T}} y \geq 0$ implies that $b^{\mathrm{T}} y \geq 0$.

Prove this result.

(iii) Let $a_0, \dots, a_m \in \mathbb{R}^n$. Using part (ii), show that the condition

$$\left\{x \in \mathbb{R}^n : a_i^{\mathrm{T}} x \leq 0,\ (i = 0, \dots, m)\right\} = \left\{x \in \mathbb{R}^n : a_i^{\mathrm{T}} x \leq 0,\ (i = 1, \dots, m)\right\}$$

(in other words: the inequality $a_0^{\mathrm{T}} x \leq 0$ is redundant) holds if and only if a_0 lies in the cone $\mathrm{cone}(a_1, \dots, a_m)$ generated by the a_i $(i \neq 0)$.

***Problem 3.** Let $a_1, \dots, a_m \in \mathbb{R}^n$ be linearly independent vectors and consider the linear programme

$$\begin{aligned} \text{(P)} \qquad & \max_{x \in \mathbb{R}^n} c^{\mathrm{T}} x \\ & \text{s.t.} \quad Ax \leq b, \end{aligned}$$

where $A = \begin{bmatrix} a_1 & \cdots & a_m \end{bmatrix}$, and where $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$ are given vectors. Consider also the dual programme

$$\begin{aligned} \text{(D)} \qquad & \min_{y \in \mathbb{R}^m} b^{\mathrm{T}} y \\ & \text{s.t.} \quad A^{\mathrm{T}} y = c, \\ & \qquad\quad y \geq 0. \end{aligned}$$

We will be interested in the following set of equations:

$$A^{\mathrm{T}} y = c, \quad y \geq 0 \tag{0.2}$$

$$Ax \leq b \tag{0.3}$$

$$c^{\mathrm{T}} x - b^{\mathrm{T}} y = 0. \tag{0.4}$$

Points y that satisfy (0.2) are called *dual feasible*, whereas points x that satisfy (0.3) are called *primal feasible*.

(i) Show that if x is primal feasible and y is dual feasible then

$$c^{\mathrm{T}} x \leq b^{\mathrm{T}} y.$$

This property is called *weak LP duality*.

(ii) Using part (i), show that if (x^*, y^*) satisfies (0.2)–(0.4) then x^* is an optimal solution for (P) and y^* is an optimal solution for (D).

(iii) Now let x^* be a maximiser of (P) and consider the set of indices $J = \{i : a_i^{\mathrm{T}} x = b_i\}$. Apply the fundamental theorem of linear inequalities to the vectors $\{a_i : i \in J\}$ and c (playing the the role of b in the statement of the theorem) and show that Alternative (I) must hold.

(iv) Conclude that there exists a vector y^* representing a dual optimal solution such that x^*, y^* satisfy (0.2)–(0.4).

(v) From part (iv) conclude that *strong LP duality* holds: (P) has an optimal solution x^* if and only if (D) has an optimal solution y^*, and whenever this is the case, the duality gap at x^* and y^* is zero. Hint: you may assume that the roles of (P) and (D) can be exchanged, since the bidual is the primal.