C6.2/B2. Continuous Optimization

Problem Sheet 4

Part C and OMMS: Please attempt Problems 2, 3, 6, 7, and 9 over the vacation, for the fourth intercollegiate class in Trinity Term.

MMSC MSc: Please hand-in Problems 3, 6 and 7. For the sixth class, please attempt Problem 9 as well, but hand-in is optional.

The other problems are optional for everyone.

1. The fundamental theorem of linear inequalities, also known as *Farkas' Lemma* states that: given any vectors $b \in \mathbb{R}^n$ and $a_i \in \mathbb{R}^n$, $i \in \{1, \ldots, m\}$, the set

$$\{s : \; b^T s < 0 \quad \text{and} \quad a_i^T s \geq 0, i \in \{1, \ldots, m\}\}$$

is empty if and only if

$$b \in C = \{\sum_{i=1}^{m} a_i y_i : y_i \geq 0, \, i \in \{1, \ldots, m\}\}.$$

(In other words, a vector b lies in the cone C generated by the vectors a_i if and only if it cannot be separated from the vectors a_i by a separating hyperplane generated by s.) Use this lemma in the next part of the problem (for appropriate choices of b, a_i and m).

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^p$ are $\mathcal{C}^1$ functions. Let x^* be a local minimizer of

$$\min_x f(x) \quad c(x) \geq 0.$$

Show that, provided a suitable first-order constraint qualification holds, there exists a vector $\lambda_* \in \mathbb{R}^p$ of Lagrange multipliers such that

$$\nabla f(x^*) = J(x^*)^T \lambda^*, \quad c(x^*) \geq 0, \quad \lambda^* \geq 0, \quad \lambda_i^* c_i(x^*) = 0, \, i \in \{1, \ldots, p\}.$$

(These are the KKT conditions for inequality-constrained problems. Use ideas and approaches from the proof of Theorem 16; note that we only need a first-order representation of the feasible path in the proof of Theorem 16. Recall that it is sufficient to consider the active constraints at x^*.)

2. Consider the following constrained optimization problem

$$\min_{(x_1,x_2) \in \mathbb{R}^2} \; \frac{1}{2} x_1^2 - x_2^2 \quad \text{subject to} \quad 1 - x_1^2 - x_2^2 \geq 0. \tag{1}$$

   (a) Write down the system of Karush-Kuhn-Tucker (KKT) conditions for problem (1). By solving this system as a function of x and the Lagrange multiplier, find (all) the KKT points $\hat{x}$ of problem (1).

   (b) State the Slater constraint qualification for problem (1); does it hold for this problem? Conclude from this if each (local) minimizer of (1) is a KKT point (in other words, if first-order necessary optimality conditions hold for this problem).

   (c) Using (a) and (b), or otherwise, find a global minimizer of (1). Argue also that a global minimizer of (1) exists.

   (d) Is the constraint of problem (1) active at each (local or global) minimizer of (1)? Justify your answer.

3. Consider the problem

$$\min_{(x_1,x_2) \in \mathbb{R}^2} f(x_1, x_2) = -x_1 + x_2 \quad \text{subject to} \quad \begin{cases} 0 \leq x_1 \leq a, \\ 0 \leq x_2 \leq 1, \\ x_2 \geq x_1^2, \end{cases} \tag{2}$$

where $a > 0$ is a fixed positive constant.

(a) By drawing a diagram of the feasible region and the contours of $f(x)$, or otherwise, determine the solution of problem (2).

(b) Show that the set of active constraints at the solution differs according to whether or not a is greater than a certain fixed value $\overline{a}$, and determine $\overline{a}$. Obtain the value of the Lagrange multipliers of the active constraints at the solution in both cases, and verify that the KKT conditions are satisfied.

4. Consider the following quadratic programming problem.

$$\begin{array}{llcrcrcrl}
\text{minimize} & f(x) & = & (x_1-1)^2 + (x_2-4)^2 & & & & & \\
\text{subject to} & & & -2x_1 & - & x_2 & = & -8 & (1) \\
 & & & x_1 & - & x_2 & \geq & -2 & (2) \\
 & & & x_1 & & & \geq & 0 & (3) \\
 & & & & & x_2 & \geq & 0 & (4)
\end{array}$$

For the points $x = (4\ 0)^T$ and $x = (2\ 4)^T$, compute the Lagrange multipliers and hence determine whether the KKT conditions are satisfied. If the KKT conditions cannot be satisfied then determine a feasible descent direction. If they are, determine whether the second order sufficient conditions are satisfied and deduce whether the vertex is a minimizer.

5. For the following problem,

$$\begin{array}{llcl}
\text{minimize} & f(x) & = & x_1^2 + x_2^2 \\
\text{subject to} & c(x) & = & (x_1-1)^3 - x_2^2 = 0
\end{array}$$

(i) Sketch contours of $f(x)$ and the points which satisfy $c(x) = 0$. Deduce the solution x^* of the problem.

(ii) Use the constraint to eliminate x_2 and show that the solution obtained in (i) is obtained. What happens if the problem is solved by eliminating x_1?

(iii) Attempt to solve the problem by the method of Lagrange multipliers. Show that either $x_2 = 0$ or $\lambda = -1$ and that both lead to a contradiction. Deduce that the stationary point condition for the Lagrangian function is not a necessary condition for a minimizer. What notable property does $\nabla c(x^*)$ have?

6. Consider the problem

$$\min_{x\in\mathbb{R}^2} -x_1 - x_2 \quad \text{subject to} \quad 1 - x_1^2 - x_2^2 = 0. \tag{3}$$

(a) Use the first-order necessary optimality (KKT) conditions to solve this problem.

(b) Let $x(\mu) = (x_1(\mu), x_2(\mu))$ be a local minimizer of the quadratic penalty function for (3). Show that $x_1(\mu) = x_2(\mu)$ and $2x_1(\mu)^3 - x_1(\mu) - \mu/2 = 0$.

(c) Among the two solutions for $x(\mu)$, pick the one for which $x_1(\mu) > 0$. Show that as $\mu \to 0$,

$$x_1(\mu) = \frac{1}{\sqrt{2}} + a\mu + O(\mu^2).$$

Find the constant a.

(d) Now consider the problem

$$\begin{aligned}
&\min -x_1 - x_2 \\
\text{s.t.}\quad & 1 - x_1^2 - x_2^2 = 0, \\
& x_2 - x_1^2 \geq 0.
\end{aligned}$$

Show how the penalty function may be modified to solve this problem. Show that there is a range of values of μ for which the minimisers of the two penalty functions agree.

7. Consider the problem

$$\begin{aligned}
&\min -x_1x_2x_3 \\
\text{s.t.}\quad & 72 - x_1 - 2x_2 - 2x_3 = 0.
\end{aligned} \tag{4}$$

(i) For $x^* = (24\ 12\ 12)^T$ verify that there exists a Lagrange multiplier λ^* such that (x^*, λ^*) is a KKT point.

(ii) Now let

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

where $Q(x,\mu)$ is the quadratic penalty function for (4). Verify that the explicit expression for $x(\mu)$ given by

$$x_1(\mu) = 2x_2(\mu), \quad x_2(\mu) = x_3(\mu) = \frac{24}{1+\sqrt{1-8\mu}}$$

satisfies $\nabla_x Q(x(\mu),\mu) = 0$, and verify that $x(\mu) \to x^*$ as $\mu \to 0$.

(iii) Let $\mu = 1/9$. Find $x(\mu)$ and verify that $\nabla^2_{xx} Q(x(\mu),\mu)$ is positive definite, so that $x(\mu)$ is a local minimizer of $Q(x,\mu)$.

(iv) Show that $-c(x(\mu))/\mu \to \lambda^*$, where c is the equality constraint function in (4).

8. (a) Show that the logarithmic barrier function for the problem of minimizing $1/(1+x^2)$ subject to $x \geq 1$ is unbounded from below for all μ.

   Comment: Thus the barrier function approach will not always work.

   (b) Find the minimizer $x(\mu)$, and its related Lagrange multiplier estimate $\lambda(\mu)$, of the logarithmic barrier function for the problem of minimizing $\frac{1}{2}x^2$ subject to $x \geq 2a$ where $a > 0$. What is the rate of convergence of $x(\mu)$ to x_* as a function of μ? And the rate of convergence of $\lambda(\mu)$ to λ_* as a function of μ?

   Comment: Problems with strictly complementary solutions (for which $\lambda_i^* > 0$ whenever $c_i(x^*) = 0$) generally have $x(\mu) - x_* = \mathcal{O}(\mu)$ and $\lambda(\mu) - \lambda_* = \mathcal{O}(\mu)$ as $\mu \to 0$.

   (c) Find the minimizer $x(\mu)$, and its related Lagrange multiplier estimate $\lambda(\mu)$, of the logarithmic barrier function for the problem of minimizing $\frac{1}{2}x^2$ subject to $x \geq 0$. How do the errors $x(\mu) - x_*$ and $\lambda(\mu) - \lambda_*$ behave as a function of μ?

   Comment: Without strict complementarity, the errors $x(\mu) - x_*$ and $\lambda(x(\mu)) - \lambda_*$ are generally larger than in the strictly complementary case.

9. Apply the augmented Lagrangian function to minimize

$$f(x) = 2x_1^2 - x_2^2 \quad \text{subject to} \quad c(x) = x_1 + x_2 - 1 = 0.$$

The estimate of the Lagrange multiplier of the constraint is revised by the formula

$$\lambda^{k+1} = \lambda^k - \frac{c(x(\lambda^k))}{\sigma}$$

where $x(\lambda^k)$ is a minimizer of the augmented Lagrangian function. Show that the sequence of values of λ^k converges if $\sigma > 0$ is sufficiently small. Find the value of σ such that each iteration reduces the difference between λ^k and the optimal multiplier λ^* by a factor of 10.

10. Suppose that an algorithm for unconstrained minimization fails if the ratio of the largest to the smallest eigenvalue of the Hessian matrix exceeds 10^{10} at the required solution. It is used to find an approximate solution of the problem

$$f(x) = x_1^2 + 2x_2^2 \quad \text{subject to} \quad x_1 + x_2 - 1 \geq 0$$

in two ways. Specifically, the functions

$$x_1^2 + 2x_2^2 + r(x_1+x_2-1)^2 \quad \text{and} \quad x_1^2 + 2x_2^2 - r\log(x_1+x_2-1)$$

are minimized over $\mathbb{R}^2$ using a large and a small value of r, respectively. Estimate the accuracy of the approximate solution in each case when r is close to a value that causes failure.