C6.2/B2. Continuous Optimization

Problem Sheet 6

Please attempt Problems 1 and 3; time permitting, please also attempt Problem 6 (it will be marked if you attempt it, but it will not count towards your Sheet 6 mark). The other problems are optional.

1. Consider the problem

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

   (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 (1). 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.

2. Let $\|\cdot\|$ be the Euclidean norm. Consider the *quartic* penalty function

$$\Phi(x,\mu) = f(x) + \frac{1}{4\mu}\|c(x)\|^4$$

for the equality-constrained minimization problem

$$\min_{x\in\mathbb{R}^n} \quad f(x) \quad \text{subject to} \quad c(x) = 0, \tag{2}$$

where $f,\, c \in \mathcal{C}^2$. Suppose that

$$y_i^k = -\frac{\|c(x^k)\|^2 c_i(x^k)}{\mu_k},$$

that

$$\|\nabla_x \Phi(x^k, \mu_k)\| \leq \epsilon_k,$$

where ϵ_k converges to zero as $k \to \infty$, and that x^k converges to x^* for which the Jacobian $J(x^*)$ of the constraints c is full rank. Show that x^* satisfies the first-order necessary optimality conditions for the problem (2) and $\{y^k\}$ converges to the associated Lagrange multipliers y^*. (*hint:* use the proof of global convergence of the quadratic penalty method.)

3. Consider the problem

$$\begin{aligned} &\min -x_1x_2x_3 \\ &\text{s.t.} \quad 72 - x_1 - 2x_2 - 2x_3 = 0. \end{aligned} \tag{3}$$

(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 (3). 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 (3).

4. (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.

5. Consider the linear programming problem

$$\max_{(y_1,y_2)\in\mathbb{R}^2} y_1 + \alpha y_2 \quad \text{subject to} \quad \begin{cases} y_1 + y_2 \leq 1, \ 2y_1 - y_2 \leq 2, \\ y_1 \geq -1, \ y_2 \geq -1, \end{cases} \tag{4}$$

where $\alpha \in [0, 1]$. Graphically or otherwise, find the solution set of problem (4) as a function of $\alpha \in [0, 1]$.

- Possibly by re-writing (4) or directly, write down the KKT conditions for (4) and its dual. Then write down the perturbed (primal-dual) system of optimality conditions of (4) (which are also the equations of the primal-dual central path).

- Now let $\alpha := 1$ in (4). Show that as $\mu \to 0$, the points $(y_1(\mu), y_2(\mu))$ of the central path converge to the solution of the following optimization problem

$$\min_{(y_1,y_2)\in\mathbb{R}^2} -\log(2 - 2y_1 + y_2) - \log(1 + y_1) - \log(1 + y_2) \quad \text{subject to} \quad y_1 + y_2 = 1. \tag{5}$$

(*hint*: you would have to solve (5) and justify it has a unique solution; then use the central path equations to show the limit).

6. 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.

7. 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.