IMPERIAL COLLEGE LONDON

EE4-29
EE9-CS3-2

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2013

MSc and EEE/EIE PART IV: MEng and ACGI

Corrected Copy

OPTIMIZATION

Thursday, 2 May 10:00 am

Time allowed: 3:00 hours

There are FOUR questions on this paper.

Answer FOUR questions.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : A. Astolfi
Second Marker(s) : P.L. Dragotti

OPTIMISATION

1. Consider the function

$$f(x_1,x_2) = \frac{1}{2}x_1^2\left(\frac{1}{6}x_1^2+1\right)+x_2\arctan x_2-\frac{1}{2}\ln(x_2^2+1).$$

a) Compute the unique stationary point of the function.
(Hint: recall that $\frac{d\arctan x}{dx} = \frac{1}{1+x^2}$.)

[2 marks]

b) Using second order sufficient conditions of optimality show that the stationary point determined in part a) is a local minimizer.
Show, in addition, that the function is convex. Finally, show that the local minimizer is a global minimizer.
(Hint: convexity of a function f is implied by the condition $\nabla^2 f(x) > 0$ for all x.)

[4 marks]

c) Consider the problem of minimizing the function f using Newton's method.

i) Write Newton's iteration for the minimization of the function f.

[2 marks]

ii) Perform 4 steps of Newton's iteration with starting point

$$(x_1,x_2) = (1,2).$$

[4 marks]

d) Consider the function

$$f_2(x_2) = x_2\arctan x_2 - \frac{1}{2}\ln(x_2^2+1).$$

i) Using the iteration derived in part c.i) write Newton's iteration for the minimization of the function f_2. [1 mark]

ii) Write the Newton's iteration in part d.i) in the form

$$x_2(k+1) = \psi(x_2(k)).$$

Write explicitly the function ψ. [1 mark]

iii) Plot on the same graph the functions x_2 and $\psi(x_2)$. Exploiting the graph explain why Newton's iteration for the minimization of f_2 converges for initial conditions sufficiently close to zero, and diverges otherwise.
(Hint: use the graph to *execute* Newton's iteration graphically.)

[4 marks]

e) Exploiting the results in part d), and the fact that the function f is the sum of two functions of one variable each, determine (qualitatively) for which initial points the Newton's iteration for the minimization of f converges to the minimizer.

[2 marks]

2. Consider a set of 3 words w_1, w_2, and w_3. The words have to be coded using binary strings. Let s_i be the length of the binary string coding w_i. Clearly the variables s_i have to be non-negative integers.

Assume that the word i occurs with probability p_i, with $p_i \in (0,1)$. Recall that $p_1 + p_2 + p_3 = 1$.

The problem of minimizing the mean word length can be formulated as follows:

$$\min_{s_1,s_2,s_3} \sum_{i=1}^{3} p_i s_i,$$
$$\sum_{i=1}^{3} 2^{-s_i} \leq 1.$$

For simplicity, in this formulation we ignore the fact that the lengths s_i are integers and that $s_i \geq 0$ for all i, that is we do not define multipliers associated to these constraints.

a) Write first order necessary conditions of optimality for the problem.

[2 marks]

b) Using the conditions in part a) determine a candidate optimal solution.
(Hint: recall that

$$\frac{d}{dx} 2^{-x} = -2^{-x} \log 2.$$

[6 marks]

c) Using second order necessary conditions of optimality show that the solution determined in part b) is a local minimizer. [4 marks]

d) Evaluate the so-called source entropy, that is the function

$$E(p_1, p_2, p_3) = \sum_{i=1}^{m} p_i s_i^\star,$$

where the $s_i^\star$ denote the optimal solutions determined in part b).

[2 marks]

e) Using the method of constraints elimination, determine for which values of the probabilities p_1, p_2, and p_3 the source entropy E determined in part d) is minimized. Disregard the conditions $p_i \in (0,1)$ and recall that $p_1 + p_2 + p_3 = 1$.

[6 marks]

3. A gambler at a horserace has an amount $b > 0$ to bet. The gambler estimates that p_i is the probability that horse i will win, and knows that s_i has been bet by others on horse i, with $i = 1, 2$, that is there are only two horses on the race.

Let $x_i \geq 0$ denote the amount bet by the gambler on horse i.

The total amount bet on the race is shared out in proportion to the bets on the winning horse.

The gambler's optimal strategy is given by the solution of the optimization problem

$$\max_{x_1,x_2} \sum_{i=1}^{2} \frac{p_i x_i}{s_i + x_i}$$
$$\sum_{i=1}^{2} x_i = b$$
$$x_1 \geq 0, x_2 \geq 0.$$

Assume that

$$p_1 = p_2 = \frac{1}{2}, \qquad s_1 = 1, \qquad s_2 = 9.$$

a) Write first order necessary conditions of optimality for the problem.

[4 marks]

b) Exploiting the conditions in part a) show that a candidate optimal solution is given by

$$x_i = \sqrt{\frac{s_i p_i}{\lambda}} - s_i \qquad i = 1,2$$

in which $\lambda > 0$ is the optimal multiplier. Use this candidate optimal solution to determine the optimal multiplier as a function of b. Then, show that this candidate solution is valid only for $b \geq 2$. [12 marks]

c) Assume $b = 1$. Show that a candidate optimal solution is given by

$$x_1 = b = 1, \qquad x_2 = 0.$$

Show, in addition, that

$$x_1 = 0, \qquad x_2 = b = 1$$

is not a candidate optimal solution. [4 marks]

4. Consider the situation in which a certain quantity of water $R > 0$ is to be allocated to three different users. Let $x_i \geq 0$ be the quantity of water allocated to user i, with $i = 1,\ 2,\ 3$.

The goal is to determine the allocation such that the total benefit from all users is maximized.

The benefit resulting from an allocation of x_i to the user i is

$$B_i(x_i) = \alpha_i x_i - x_i^2,$$

with $\alpha_1 = 1$, $\alpha_2 = 2$, and $\alpha_3 = 3$.

Note, finally, that the allocations x_i have to be selected such that

$$x_1 + x_2 + x_3 = R.$$

a) Sketch the graphs of the three utility functions B_i, for $i = 1,\ 2,\ 3$. [2 marks]

b) Write the problem of maximizing the total benefit of all users as an optimization problem.
[2 marks]

c) Write first order necessary conditions of optimality for the problem formulated in part b).
[4 marks]

d) Using the conditions in part c), show that there exists a value $R_1 > 0$ such that, for all $R \in [0, R_1]$, the only candidate optimal solution is given by

$$x_1 = 0 \qquad x_2 = 0 \qquad x_3 > 0.$$

[2 marks]

e) Using the conditions in part c), show that there exists a value $R_2 > R_1$, with R_1 as in part d), such that, for all $R \in (R_1, R_2]$, the only candidate optimal solution is such that

$$x_1 = 0 \qquad x_2 > 0 \qquad x_3 > 0.$$

[4 marks]

f) Finally, using the conditions in part c), show that for all $R > R_2$, with R_2 as in part e), the only candidate optimal solution is such that

$$x_1 > 0 \qquad x_2 > 0 \qquad x_3 > 0.$$

[4 marks]

g) Exploiting the results in parts d), e) and f) sketch the graphs of the optimal allocations x_i as a function of R, for $R > 0$.
[2 marks]

Optimisation - Model answers 2013

(Note to external examiners: all questions involve mostly applications of standard methods and concepts to unseen examples.)

Question 1

a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{3}x_1(x_1^2+3) \\ \arctan x_2 \end{bmatrix}.$$

Hence, the point $(0,0)$ is the unique stationary point. [2 marks]

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} x_1^2+1 & 0 \\ 0 & \frac{1}{1+x_2^2} \end{bmatrix}.$$

Note that

$$\nabla^2 f(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is positive definite, hence the point $(0,0)$ is a local minimizer. In addition, $\nabla^2 f > 0$ for all (x_1, x_2), hence the function is convex. For (strictly) convex function, a stationary point is a global minimizer, hence $(0,0)$ is a global minimizer. [4 marks]

c) i) Newton's iteration, considering that the function f is the sum of a function of x_1 and of a function of x_2, gives two *decoupled* equations, namely

$$x_1^{k+1} = \frac{2}{3}\frac{x_1^3}{1+x_1^2} \qquad\qquad x_2^{k+1} = x_2 - (1+x_2^2)\arctan x_2.$$

[2 marks]

ii) The first five elements of the sequences $\{x_1^k\}$ and $\{x_2^k\}$ are

$$x_1^0 = 1,\ x_1^1 = 1/3,\ x_1^2 = 1/45,\ x_1^3 = 1/136755,\ x_1^4 = 1/3836373661058445 \approx 0,$$

and

$$x_2^0 = 2,\ x_2^1 = -3.5357,\ x_2^2 = 13.95095909,\ x_2^3 = -279.3440667,\ x_2^4 = 122016.9990.$$

[4 marks]

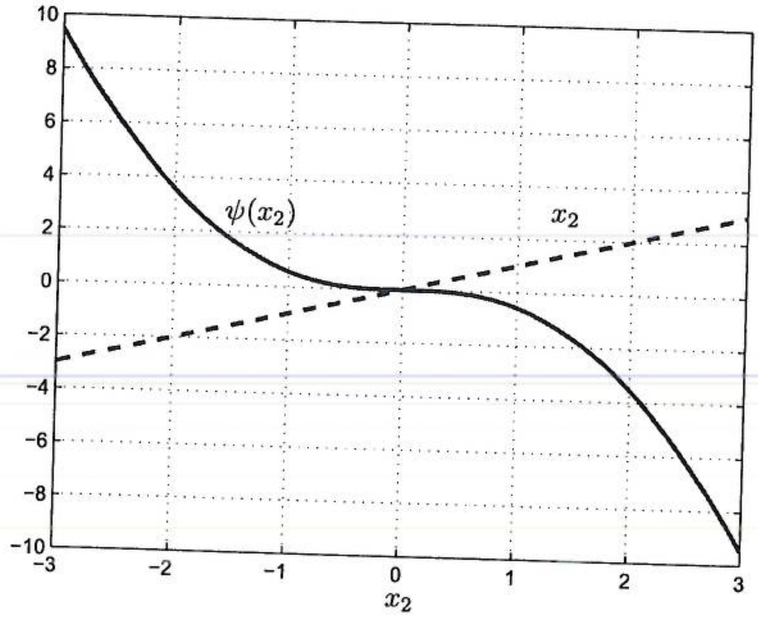
d) i) The iteration is the same as the "x_2" iteration in part c.i). [1 mark]

ii) The function ψ is given by

$$\psi(x_2) = x_2 - (1+x_2^2)\arctan(x_2).$$

[1 mark]

iii) The graphs are displayed in the following figure.



One can use the graph to show how Newton's iteration works. In fact, pick a point x_2^k on the x_2-axis, and *lift* it (up or down) on the graph of the function ψ. Then *move* the point horizontally on the graph of the function x_2, and then vertically on the x_2-axis. This is the point x_2^{k+1}. Iterating the procedure one can construct the sequence $\{x_2^k\}$. Using this approach, one concludes that if x_2^0 is sufficiently close to zero the iteration yields a sequence converging to $x_2 = 0$. If $|x_2|$ is large, then the sequence diverges.

[4 marks]

e) As shown in part c.i), Newton's iteration is composed of two decoupled iterations. The iteration for x_1 yields a globally converging sequence, whereas the iteration for x_2 converges only for $|x_2^0|$ sufficiently small (to be precise, for $|x_2^0| < 1.39...$). Hence, for all initial points (x_1^0, x_2^0) such that $|x_2^0| < 1.39...$, the iteration yields a sequence converging to the global minimizer.

[2 marks]

Question 2

a) The Lagrangian of the problem is

$$L(s_1, s_2, s_3, \rho) = p_1 s_1 + p_2 s_2 + p_3 s_3 + \rho(2^{-s_1} + 2^{-s_2} + 2^{-s_3} - 1).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial s_1} = p_1 - \rho\, 2^{-s_1} \log 2 \qquad 0 = \frac{\partial L}{\partial s_2} = p_2 - \rho\, 2^{-s_2} \log 2 \qquad 0 = \frac{\partial L}{\partial s_3} = p_3 - \rho\, 2^{-s_3} \log 2$$

$$\rho(2^{-s_1} + 2^{-s_2} + 2^{-s_3} - 1) = 0 \qquad \rho \geq 0 \qquad 2^{-s_1} + 2^{-s_2} + 2^{-s_3} - 1 \leq 0$$

[2 marks]

b) Using the complementarity condition one has two cases.

Case 1: $\rho = 0$. In this case, each of the conditions $0 = \frac{\partial L}{\partial s_i}$ reduces to $p_i = 0$, which is not possible.

Case 2: $\rho > 0$. In this case (recall that $p_1 + p_2 + p_3 = 1$)

$$\rho = \frac{1}{\log 2},$$

hence

$$2^{-s_i} = p_i$$

yielding the candidate optimal solution

$$s_i = -\log_2 p_i.$$

[6 marks]

c) The Hessian matrix of the Lagrangian, evaluated at the candidate optimal solution, is

$$\nabla^2 L^\star = \log 2 \;\; \mathrm{diag}(p_1, p_2, p_3),$$

which is positive definite. As a result, the candidate optimal solution is a (local) minimizer. [4 marks]

d) The source entropy is

$$E(p_1, p_2, p_3) = -(p_1 \log_2 p_1 + p_2 \log_2 p_2 + p_3 \log_2 p_3).$$

[2 marks]

e) Eliminating p_3, that is using the equation $p_3 = 1 - p_1 - p_2$, yields

$$\tilde{E}(p_1, p_2) = -(p_1 \log_2 p_1 + p_2 \log_2 p_2 + (1 - p_1 - p_2) \log_2(1 - p_1 - p_2)).$$

The stationary points of $\tilde{E}$ are the solutions of

$$0 = \frac{\partial \tilde{E}}{\partial p_1} = -\log_2 p_1 + \log_2(1 - p_1 - p_2) \qquad 0 = \frac{\partial \tilde{E}}{\partial p_2} = -\log_2 p_2 + \log_2(1 - p_1 - p_2).$$

These equations have the unique solution $p_1 = p_2 = 1/3$. At this point, the Hessian matrix of $\tilde{E}$ is negative definite, hence $p_1 = p_2 = p_3 = 1/3$ is a local (actually it is global) maximizer for the source entropy. [6 marks]

Question 3

a) The Lagrangian of the problem is (one has to change sign to the cost function to have a minimization problem)

$$L(x_1, x_2, \lambda, \rho_1, \rho_2) = -\frac{1}{2}\frac{x_1}{1+x_1} - \frac{1}{2}\frac{x_2}{9+x_2} + \lambda(x_1 + x_2 - b) - \rho_1 x_1 - \rho_2 x_2.$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = -\frac{1}{2}\frac{1}{1+x_1} + \frac{1}{2}\frac{x_1}{(1+x_1)^2} + \lambda - \rho_1 \qquad 0 = \frac{\partial L}{\partial x_2} = -\frac{1}{2}\frac{1}{9+x_2} + \frac{1}{2}\frac{x_2}{(9+x_2)^2} + \lambda - \rho_2$$

$$x_1 + x_2 - b = 0 \qquad -x_1 \leq 0 \qquad -x_2 \leq 0 \qquad \rho_1 \geq 0 \qquad \rho_2 \geq 0$$

$$\rho_1 x_1 = 0 \qquad \rho_2 x_2 = 0$$

[4 marks]

b) Replacing

$$x_1 = \sqrt{\frac{1}{2\lambda}} - 1 \qquad x_2 = 3\sqrt{\frac{1}{2\lambda}} - 9$$

in the necessary conditions gives

$$0 = -\rho_1 \qquad 0 = -\rho_2 \qquad 2\sqrt{\frac{2}{\lambda}} - 10 - b = 0$$

From this condition one has $\lambda = \dfrac{8}{(10+b)^2}$ hence

$$x_1 = \frac{3}{2} + \frac{1}{4}b \qquad x_2 = -\frac{3}{2} + \frac{3}{4}b.$$

This is a solution provided $x_1 \geq 0$ and $x_2 \geq 0$, that is provided $b \geq 2$.
Note that for $b \in (0, 2)$ this solution is not feasible, and the actual solution should be such that at least one of the inequality constraints is active. [12 marks]

c) Replacing $x_1 = b = 1$ and $x_2 = 0$ in the necessary conditions yields

$$0 = -\frac{1}{8} + \lambda - \rho_1 \qquad 0 = -\frac{1}{18} + \lambda - \rho_2 \qquad \rho_1 = 0$$

hence

$$x_1 = 1 \qquad x_2 = 0 \qquad \lambda = \frac{1}{8} \qquad \rho_1 = 0 \qquad \rho_2 = \frac{5}{72}$$

gives a candidate optimal solution.
Replacing $x_1 = 0$ and $x_2 = b = 1$ in the necessary conditions yields

$$0 = -\frac{1}{2} + \lambda - \rho_1 \qquad 0 = -\frac{9}{200} + \lambda - \rho_2 \qquad \rho_2 = 0$$

hence

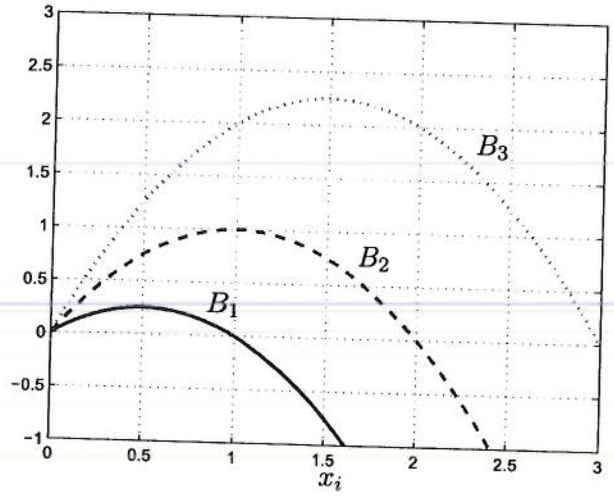
$$x_1 = 0 \qquad x_2 = 1 \qquad \lambda = \frac{9}{200} \qquad \rho_1 = -\frac{91}{200} \qquad \rho_2 = 0,$$

which is not admissible since $\rho_1 < 0$. [4 marks]

This result can be interpreted as follows. If the amount to be bet b is below a certain value (2 in this example) and if both horses have the same probability to win, then the optimal gambling strategy is to bet all the available amount on the horse which attracts the smallest amount of bets.

Question 4

a) The graphs of the functions B_i are given in the following figure.



[2 marks]

b) The problem can be written as (note the change in the sign of the objective function)

$$\min_{x_1,x_2,x_3} -(x_1 - x_1^2) - (2x_2 - x_2^2) - (3x_3 - x_3^2)$$

$$x_1 + x_2 + x_3 = R$$

$$-x_1 \leq 0 \qquad -x_2 \leq 0 \qquad -x_3 \leq 0$$

[2 marks]

c) The Lagrangian of the problem is

$$\begin{aligned} L(x_1, x_2, x_3, \lambda, \rho_1, \rho_2, \rho_3) = & -(x_1 - x_1^2) - (2x_2 - x_2^2) - (3x_3 - x_3^2) \\ & +\lambda(x_1 + x_2 + x_3 - R) - \rho_1 x_1 - \rho_2 x_2 - \rho_3 x_3. \end{aligned}$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = -1+2x_1+\lambda-\rho_1 \qquad 0 = \frac{\partial L}{\partial x_2} = -2+2x_2+\lambda-\rho_2 \qquad 0 = \frac{\partial L}{\partial x_3} = -3+2x_3+\lambda-\rho_3$$

$$x_1 + x_2 + x_3 - R = 0 \qquad -x_1 \leq 0 \qquad -x_2 \leq 0 \qquad -x_3 \leq 0$$

$$\rho_1 x_1 = 0 \qquad \rho_2 x_2 = 0 \qquad \rho_3 x_3 = 0 \qquad \rho_1 \geq 0 \qquad \rho_2 \geq 0 \qquad \rho_3 \geq 0$$

[4 marks]

d) The selection $x_1 = 0$, $x_2 = 0$ and $x_3 = R$ is an optimal solution only if

$$0 = -1 + \lambda - \rho_1 \qquad 0 = -2 + \lambda - \rho_2 \qquad 0 = -3 + 2R + \lambda \qquad \rho_1 \geq 0 \qquad \rho_2 \geq 0$$

The first three equations yield

$$\lambda = 3 - 2R \qquad \rho_1 = 2 - 2R \qquad \rho_2 = 1 - 2R.$$

As a result, this is a solution only if $R \in [0, 1/2]$, thus $R_1 = 1/2$. (For $R > R_1$ the condition $\rho_2 \geq 0$ is violated.)

[2 marks]

e) Using the selection $x_1 = 0$, $x_2 > 0$ and $x_3 > 0$ in the necessary conditions yields

$$0 = -1 + \lambda - \rho_1 \qquad 0 = -2 + 2x_2 + \lambda \qquad 0 = -3 + 2x_3 + \lambda \qquad x_2 + x_3 = R \qquad \rho_1 \geq 0$$

The first four equations yield

$$\lambda = \frac{5}{2} - R \qquad \rho_1 = \frac{3}{2} - R \qquad x_2 = \frac{R}{2} - \frac{1}{4} \qquad x_3 = \frac{R}{2} + \frac{1}{4}$$

As a result, this is a solution only if $R \in [1/2, 3/2]$, thus $R_2 = 3/2$. (For $R > R_2$ the condition $\rho_1 \geq 0$ is violated.)

[4 marks]

f) Using the selection $x_1 > 0$, $x_2 > 0$ and $x_3 > 0$ in the necessary conditions yields

$$0 = -1 + 2x_1 + \lambda \qquad 0 = -2 + 2x_2 + \lambda \qquad 0 = -3 + 2x_3 + \lambda \qquad x_1 + x_2 + x_3 = R.$$

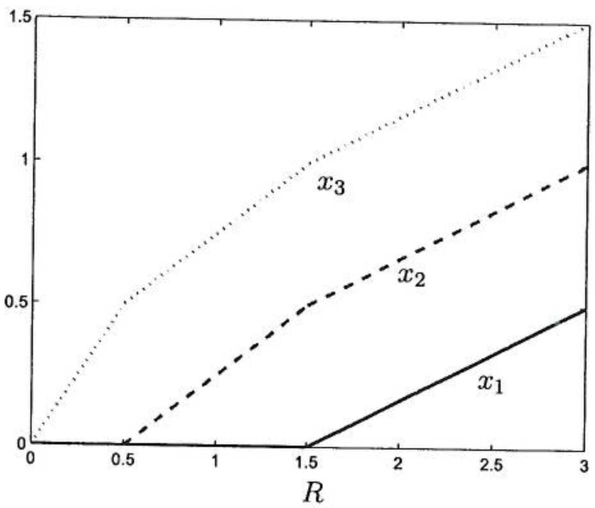
These equations yield

$$x_1 = -\frac{1}{2} + \frac{1}{3}R \qquad x_2 = \frac{1}{3}R \qquad x_3 = \frac{1}{2} + \frac{1}{3}R \qquad \lambda = 2 - \frac{2}{3}R$$

Note that all x_i's are positive for $R > 3/2 = R_2$.

[4 marks]

g) The optimal allocations x_i as a function of $R > 0$ are displayed in the following graph.



[2 marks]