Final copy

E4.29
C1.1
ISE4.55

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2006

MSc and EEE PART IV: MEng and ACGI

# OPTIMIZATION

Friday, 28 April 10:00 am

Time allowed: 3:00 hours

**There are SIX 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) : E.C. Kerrigan

# OPTIMISATION

1. a) An electrical engineer wants to maximize the current $I$ between two points A and B of a complex network by adjusting the values $x_1$ and $x_2$ of two variable resistors. The engineer does not have a model of the network and decides to opt for this procedure.

- Keep the value $x_2$ fixed and adjust $x_1$ to maximize $I$.
- Keep the value $x_1$ fixed and adjust $x_2$ to maximize $I$.
- Repeat the above steps until no further improvement can be obtained.

Explain if this approach has sound theoretical basis: *i.e.* discuss under what assumptions the above procedure determines a stationary point of the function $I$. [ 6 marks ]

b) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Suppose that $x_\star$ is a local minimum of $f$ along every line that passes through $x_\star$, *i.e.* the function

$$g(\alpha) = f(x_\star + \alpha d)$$

is minimized at $\alpha = 0$ for all $d \in \mathbb{R}^n$.

i) Show that $\nabla f(x_\star) = 0$. [ 4 marks ]

ii) Is $x_\star$ a local minimum of $f$? [ 2 marks ]

iii) Consider the function

$$f(x_1, x_2) = (x_2 - x_1^2)(x_2 - 2x_1^2).$$

Show that the point $(0,0)$ is a local minimum of $f$ along every line that passes through (0,0). Show that the point $(0,0)$ is not a local minimum of $f$. (Hint: consider the values of $f$ for $x_1 = y$ and $x_2 = my^2$ and $m \in \mathbb{R}$.) [ 8 marks ]

2. Consider the problem of minimizing the function

$$f(x_1, x_2) = \frac{1}{3}x_1^2 - \alpha x_1^4 + \frac{1}{4}x_1^6 + x_1 x_2 + x_2^2,$$

where $\alpha$ is a constant.

a) Compute all stationary points of the function. [ 6 marks ]

b) Let $\alpha = 5/12$. Using second order sufficient conditions *classify* the stationary points determined in part a), *i.e.* say which is a local minimum, or a local maximum, or a saddle point. [ 10 marks ]

c) Let $\alpha = 5/12$. Show that the function $f$ is radially unbounded and hence compute the global minimum of $f$. Is the global minimizer unique? [ 4 marks ]

3. Consider the problem of minimizing the function

$$f(x) = x - \log x,$$

with $x > 0$.

a) Compute analytically the minimum of $f$. [ 2 marks ]

b) Write Newton's iteration for the proposed algorithm. [ 2 marks ]

c) Consider the Newton's iteration in part b) with initial point $x_0 = 1.99$. Compute ten steps of the Newton's iteration. Argue that the resulting sequence converges to the minimum of $f$. Show that the sequence converges to the minimum of $f$ with quadratic speed of convergence. [ 6 marks ]

d) Consider the Newton's iteration in part b) with initial point $x_0 = 2.01$. Compute five steps of the Newton's iteration. Argue that the resulting sequence diverges. [ 4 marks ]

e) Consider the Newton's iteration in part b). Show that

i) if the initial point $x_0 = 2$ then $x_k = 0$, for all $k \geq 1$;

ii) if the initial point $x_0 = 0$ then $x_k = 0$, for all $k \geq 1$;

iii) if the initial point $x_0 > 2$ then $x_k < 0$, for all $k \geq 1$ and the sequence does not converge;

iv) if the initial point $x_0 \in (0,2)$ then $x_k \in (0,2)$, for all $k \geq 1$ and the sequence converges to the minimum determined in part a).

[ 6 marks ]

4. Let $Q \in \mathbb{R}^{n\times n}$ with $Q = Q' > 0$, $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ $b \in \mathbb{R}^m$ and $y \in \mathbb{R}^m$. Consider the minimization problem

$$P : \begin{cases} \min_x \frac{1}{2} x' Q x \\ Ax - b \leq 0 \end{cases}$$

and the so-called dual problem

$$D : \begin{cases} \min_y \frac{1}{2} y' A Q^{-1} A' y + b' y \\ -y \leq 0. \end{cases}$$

a) Write first order necessary conditions of optimality for the problem P. (Denote the multiplier with $\rho$.) [ 2 marks ]

b) Write first order necessary conditions of optimality for the problem D. (Denote the multiplier with $\sigma$.) [ 2 marks ]

c) Let $y_\star$ and $\sigma_\star$ be such that the optimality conditions in part b) hold. Show that

$$x_\star = -Q^{-1} A' y_\star \qquad \rho_\star = y_\star$$

are such that the optimality conditions in part a) hold. [ 8 marks ]

d) Consider the minimization problem

$$P_1 : \begin{cases} \min_x \frac{1}{2} x' x \\ x_1 + 1 \leq 0 \end{cases}$$

with $x \in \mathbb{R}^n$ and $x = [x_1, x_2, \cdots, x_n]'$. Exploiting the results above solve this problem. (Hint: write the dual $D_1$ of problem $P_1$, solve problem $D_1$, and then obtain a solution to problem $P_1$ exploiting the results in part c).) [ 8 marks ]

5. Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,\, x_2} x_1^2 + x_2 \\ x_1^2 + (x_2 - 1)^2 = 4. \end{cases}$$

a) Sketch in the $(x_1, x_2)$-plane the level surfaces of the function to be minimized and the admissible set. Hence show that all points in the admissible set are regular points for the constraints. [ 4 marks ]

b) Using only graphical considerations, determine the solution of the considered problem. [ 2 marks ]

c) Show that the considered problem can be solved by eliminating the variable $x_1$ and obtaining the optimization problem

$$\begin{cases} \min\limits_{x_2} 4 - (x_2 - 1)^2 + x_2 \\ -1 \leq x_2 \leq 3. \end{cases}$$

[ 4 marks ]

d) Solve the optimization problem in part c) and hence obtain a solution for the considered optimization problem. [ 4 marks ]

e) Suppose that one wants to solve the considered optimization problem using recursive quadratic programming methods. Write the quadratic programming problem associated with the considered optimization problem. [ 6 marks ]

6. Consider the optimization problem

$$\begin{cases} \max\limits_{x_1,\, x_2,\, x_3} x_1^\alpha x_2^\alpha x_3^\alpha \\ x_1 + x_2 + x_3 - 1 = 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0, \end{cases}$$

with $\alpha > 1$.

a) State first order necessary conditions of optimality for this constrained optimization problem. [ 4 marks ]

b) Using the conditions derived in part a), compute candidate optimal solutions. Show that there is only one candidate solution such that $x_1 \neq 0$, $x_2 \neq 0$ and $x_3 \neq 0$. [ 8 marks ]

c) Consider the candidate optimal solution with $x_1 \neq 0$, $x_2 \neq 0$ and $x_3 \neq 0$ determined in part b). Show using second order sufficient conditions that such a candidate optimal point is a local maximum. [ 8 marks ]

Master Copy E4.29/ISE 4.55/C1.1

# Optimisation - Model answers 2006

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

## Question 1

a) The engineer is applying the so-called coordinate directions method with an exact line search (without derivatives), as described in Section 2.9 of the Lecture Notes, for the minimization of the function $-I = -I(x_1, x_2)$. This approach provides a sequence of points converging to a stationary point of the function $I$ provided that the initial point is selected inside a compact level set of $-I(x_1, x_2)$.

b) i) Note that, by assumption, the function

$$\frac{dg}{d\alpha} = \nabla f(x_\star + \alpha d)'d$$

is zero for $\alpha = 0$ and for every $d$. This means that

$$\nabla f(x_\star)'d = 0$$

for every $d$, and this implies that

$$\nabla f(x_\star) = 0.$$

ii) Without further information on $f$ it is not possible to draw any conclusion on $x_\star$, *i.e.* $x_\star$ is a stationary point of $f$, but it may be a local minimum, a local maximum or a saddle point.

iii) Consider a line that goes through zero, namely $x_2 = \gamma x_1$, and note that

$$f(x_1, \gamma x_1) = (\gamma x_1 - x_1^2)(\gamma x_1 - 2x_1^2) = \gamma^2 x_1^2 - 3\gamma x_1^3 + 2x_1^4$$

and this shows that for all $\gamma$ the point $x_1 = 0$ is a minimum of $f(x_1, \gamma x_1)$. For completeness we have also to consider the line $x_1 = 0$ (which corresponds formally to $\gamma = \infty$). Note that

$$f(0, x_2) = x_2^2$$

hence the point $x_2 = 0$ is a minimum of $f(0, x_2)$.
To show that $(0,0)$ is not a local minimum of $f$ note first that $f(0,0) = 0$ and then let $x_1 = y$ and $x_2 = my^2$. Note that

$$f(y, my^2) = y^4(m-1)(m-2).$$

Pick $m \in (1,2)$ and note that for such values of $m$

$$f(y, my^2) = y^4(m-1)(m-2) < 0$$

for all $y \neq 0$. This shows that close to the point $(0,0)$, where the function is zero, there are points in which the function takes negative values. Hence, $(0,0)$ is not a local minimum of $f$.

## Question 2

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

$$0 = \nabla f = \begin{bmatrix} 2/3x_1 - 4\alpha x_1^3 + 3/2x_1^5 + x_2 \\ x_1 + 2x_2 \end{bmatrix}.$$

Solving the second equation for $x_2$ yields $x_2 = -1/2x_1$, and upon replacement in the first equation we obtain

$$\frac{1}{6}x_1 - 4\alpha x_1^3 + \frac{3}{2}x_1^5 = 0,$$

yielding

$$x_{1a} = 0 \qquad x_{1b} = \frac{1}{3}\sqrt{12\alpha + 3\sqrt{16\alpha^2 - 1}} \qquad x_{1c} = -\frac{1}{3}\sqrt{12\alpha + 3\sqrt{16\alpha^2 - 1}}$$

$$x_{1d} = \frac{1}{3}\sqrt{12\alpha - 3\sqrt{16\alpha^2 - 1}} \qquad x_{1e} = -\frac{1}{3}\sqrt{12\alpha - 3\sqrt{16\alpha^2 - 1}}.$$

b) For $\alpha = 5/12$ we obtain the following stationary points

$$P_a = (0,0) \qquad P_b = (1,-1/2) \qquad P_c = (-1,1/2)$$

$$P_d = (1/3,-1/6) \qquad P_e = (-1/3,1/6).$$

Note now that, for $\alpha = 5/12$,

$$\nabla^2 f = \begin{bmatrix} 2/3 - 5x_1^2 + 15/2x_1^4 & 1 \\ 1 & 2 \end{bmatrix}$$

and that

$$\nabla^2 f(P_a) = \begin{bmatrix} 2/3 & 1 \\ 1 & 2 \end{bmatrix} > 0$$

$$\nabla^2 f(P_b) = \nabla^2 f(P_c) = \begin{bmatrix} 19/6 & 1 \\ 1 & 2 \end{bmatrix} > 0$$

$$\nabla^2 f(P_d) = \nabla^2 f(P_e) = \begin{bmatrix} 11/54 & 1 \\ 1 & 2 \end{bmatrix} \not\geq 0.$$

As a result, $P_a$, $P_b$ and $P_c$ are local minima, and $P_d$ and $P_e$ are saddle points.

c) Note that

$$-\frac{5}{12}x_1^4 + \frac{1}{4}x_1^6 = x_1^4\left(\frac{1}{4}x_1^2 - \frac{5}{12}\right)$$

is radially unbounded. Hence

$$f(x_1,x_2) = (\frac{1}{3}x_1^2 + x_1x_2 + x_2^2) + x_1^4\left(\frac{1}{4}x_1^2 - \frac{5}{12}\right),$$

is also radially unbounded. The global minimum of $f$ is also a local minimum of $f$. Note that

$$f(P_a) = 0 \qquad f(P_b) = f(P_c) = -0.833\cdots.$$

Hence, $P_b$ and $P_c$ are both global minimizers, therefore the global minimum is not unique.

## Question 3

a) The minimum of $f$ is obtained solving $\nabla f = 1 - 1/x = 0$, yielding $x = 1$. Note that $x = 1$ is indeed a minimum (a global one), because the function $f$ is convex for all $x > 0$.

b) The Newton's iteration is $x_{k+1} = x_k - \frac{1}{\nabla^2 f(x_k)} \nabla f(x_k) = x_k - x_k^2(1 - \frac{1}{x_k}) = (2 - x_k)x_k$.

c) Let $x_0 = 1.99$ then

$$\begin{aligned} x_1 &= 0.01990 \\ x_2 &= 0.03940399 \\ x_3 &= 0.07725530557208 \\ x_4 &= 0.14854222890512 \\ x_5 &= 0.27501966404215 \\ x_6 &= 0.47440351247444 \\ x_7 &= 0.72374833230079 \\ x_8 &= 0.92368501609341 \\ x_9 &= 0.99417602323134 \\ x_{10} &= 0.99996608129460. \end{aligned}$$

The sequence is converging to $x = 1$, i.e. to the local minimum of $f$.
To establish quadratic speed of convergence note that

$$\frac{\mathcal{E}_{k+1}}{\mathcal{E}_k^2} = \frac{|x_{k+1} - 1|}{|x_k - 1|^2} = \frac{|(2 - x_k)x_k - 1|}{(x_k - 1)^2} = 1.$$

d) Let $x_0 = 2.01$ then

$$\begin{aligned} x_1 &= -0.02010 \\ x_2 &= -0.04060401 \\ x_3 &= -0.08285670562808 \\ x_4 &= -0.17257864492369 \\ x_5 &= -0.37494067853109. \end{aligned}$$

We then infer that the sequence will decrease and $\lim_{k\to\infty} x_k = -\infty$.

e) The first two points are trivial noting that

$$x_{k+1} = (2 - x_k)x_k$$

and the r.h.s. is zero for $x_k = 0$ or $x_k = 2$.
If $x_0 > 2$ then $x_1 < 0$. Moreover if $x_k < 0$ then

$$x_{k+1} = (2 - x_k)x_k < x_k,$$

which proves the third claim.
Finally, if $x_k \in (0, 2)$ then it is easy to verify that

$$0 < x_{k+1} = (2 - x_k)x_k < 2.$$

Moreover, if $x_k = 1$ then $x_{k+1} = 1$, hence $x = 1$ is an equilibrium of the discrete-time system $x_{k+1} = (2 - x_k)x_k$. Finally, if $x_k \in (1, 2)$

$$0 < x_{k+1} < 1,$$

and if $x_k \in (0, 1)$

$$x_k < x_{k+1} < 1,$$

which shows convergence of the sequence to $x = 1$.

## Question 4

a) Let $L_P = \frac{1}{2}x'Qx + \rho'(Ax-b)$ be the Lagrangian for problem $P$. The first order necessary conditions of optimality for problem $P$ are

$$\begin{aligned} &Qx_\star + A'\rho_\star = 0 \\ &Ax_\star - b \leq 0 \\ &\rho_\star \geq 0 \\ &\rho_\star'(Ax_\star - b) = 0. \end{aligned}$$

b) Let $L_D = \frac{1}{2}y'AQ^{-1}A'y + b'y + \sigma'(-y)$ be the Lagrangian for problem $D$. The first order necessary conditions of optimality for problem $D$ are

$$\begin{aligned} &AQ^{-1}A'y_\star + b - \sigma_\star = 0 \\ &-y_\star \leq 0 \\ &\sigma_\star \geq 0 \\ &\sigma_\star'(-y_\star) = 0. \end{aligned}$$

c) Replacing $x_\star = -Q^{-1}A'y_\star$ and $\rho_\star = y_\star$ in the equations in part a) yields

$$\begin{array}{rcl} Q(-Q^{-1}A'y_\star) + A'y_\star & = & 0 \\ A(-Q^{-1}A'y_\star) - b & \leq & 0 \\ y_\star & \geq & 0 \\ y_\star'(A(-Q^{-1}A'y_\star) - b) & = & 0. \end{array}$$

The first of the above equations holds trivially.
For the second one note that

$$A(-Q^{-1}A'y_\star) - b = -\sigma_\star \leq 0$$

by the third of the equations in b).
The third equation holds by the second of the equations in b).
The fourth equation holds exploiting the first and the fourth of the equations in b), hence we conclude the claim.

d) Problem $P_1$ is of the form of problem $P$ with $Q = I$, $A = [1, 0, \cdots, 0]$ and $b = -1$. Hence, the dual $D_1$ is

$$D_1 : \begin{cases} \min\limits_y \frac{1}{2}y^2 - y \\ -y \leq 0 \end{cases}$$

with $y \in I\!R$. The problem $D_1$ has the solution $y_\star = 1$ and $\sigma_\star = 0$. Hence, the solution to problem $P_1$ is

$$x_\star = -[1, 0, \cdots, 0]' \qquad \rho_\star = 1.$$

## Question 5

a) The level sets and the admissible set are depicted in Figure 1. Note that the constraint is always active, and that the gradient of the constraint is never zero, hence all points are regular points.

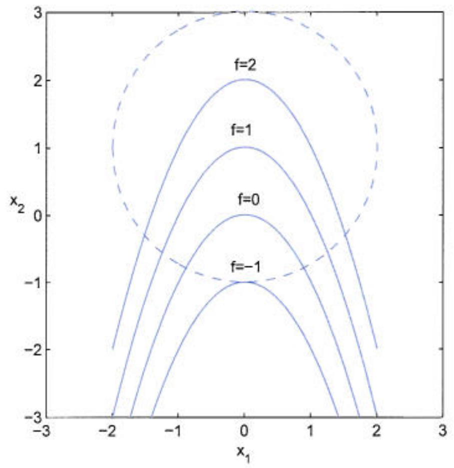


Figure 1: Level sets and constraint for Question 5.

b) The solution of the problem is obtained when the level set of $f$ is tangent to the admissible set in its lower point. As a result, the optimal point is $(x_1, x_2) = (0, -1)$.

c) We can solve the constraint yielding

$$x_1^2 = 4 - (x_2 - 1)^2.$$

Replacing in $f$ we obtain the function to minimize

$$\tilde{f} = 4 - (x_2 - 1)^2 + x_2.$$

Note that $x_2$ is not *free*. In fact, from the constraint

$$(x_2 - 1)^2 = 4 - x_1^2 \leq 4$$

we obtain

$$-1 \leq x_2 \leq 3.$$

This shows that eliminating the variable $x_1$ yields the constrained scalar problem given in part b).

d) Note that a solution to the minimization problem in part c) is obtained at a stationary point of $\tilde{f}$ or at the boundary of the admissible set. The function $\tilde{f}$ has a stationary point (a local maximum) for $x_2 = 3/2$. Note now that

$$\tilde{f}(-1) = -1 \qquad \tilde{f}(3/2) = 21/4 \qquad \tilde{f}(3) = 3.$$

Therefore the function $\tilde{f}$ attains its minimum for $x_2 = -1$. Replacing this into the constraint yields

$$x_1 = 0$$

and this coincides with the optimal solution obtained in part b).

c) Consider the optimization problem $\min_x f(x)$ subject to $g(x) = 0$. Using recursive quadratic programming methods to solve this problem one obtains the quadratic programming problem

$$PQ_1^{k+1} \quad \begin{cases} \min\limits_{\delta} f(x_k) + \nabla f(x_k)'\delta + \frac{1}{2}\delta' \nabla^2_{xx} L(x_k, \lambda_k)\delta \\ \frac{\partial g(x_k)}{\partial x}\delta = 0, \end{cases}$$

where $L = f + \lambda' g$, $\delta = x - x_k$, and $x_k$ and $\lambda_k$ are the current estimates of the solution and of the multiplier.

For the specific example it is enough to replace the function $f$, $g$ in the above expression.

## Question 6

a) Let

$$L = -x_1^\alpha x_2^\alpha x_3^\alpha + \lambda(x_1 + x_2 + x_3 - 1) + \mu_1(-x_1) + \mu_2(-x_2) + \mu_3(-x_3).$$

The first order necessary conditions of optimality for the problem are

$$\begin{aligned}
&-\alpha x_1^{\alpha-1} x_2^\alpha x_3^\alpha + \lambda - \mu_1 = 0\\
&-\alpha x_1^\alpha x_2^{\alpha-1} x_3^\alpha + \lambda - \mu_2 = 0\\
&-\alpha x_1^\alpha x_1^\alpha x_3^{\alpha-1} + \lambda - \mu_3 = 0\\
&x_1 + x_2 + x_3 - 1 = 0\\
&-x_1 \le 0\\
&-x_2 \le 0\\
&-x_3 \le 0\\
&\mu_1 \ge 0\\
&\mu_2 \ge 0\\
&\mu_3 \ge 0\\
&x_1\mu_1 = 0\\
&x_2\mu_2 = 0\\
&x_3\mu_3 = 0.
\end{aligned}$$

b) Consider the condition $x_1\mu_1 = 0$. This implies $\mu_1 = 0$ or $x_1 = 0$.
If $x_1 = 0$ then, because $\alpha > 1$,

$$\lambda = \mu_1 = \mu_2 = \mu_3 = \kappa \ge 0,$$

for some constant $\kappa$. If $\kappa > 0$ then $x_2 = 0$ and $x_3 = 0$ which is not feasible. If $\kappa = 0$ then any $x_2$ and $x_3$ such that

$$x_2 + x_3 - 1 = 0 \qquad x_2 \ge 0 \qquad x_3 \ge 0$$

satisfies necessary conditions of optimality.
We obtain similar conclusions from the conditions $x_2\mu_2 = 0$ and $x_3\mu_3 = 0$.
To obtain other candidate solution we have to consider the case $x_1 \ne 0$, $x_2 \ne 0$ and $x_3 \ne 0$. In this case, $\mu_1 = \mu_2 = \mu_3 = 0$, and

$$\begin{aligned}
&\alpha x_1^{\alpha-1} x_2^\alpha x_3^\alpha - \lambda = 0\\
&\alpha x_1^\alpha x_2^{\alpha-1} x_3^\alpha - \lambda = 0\\
&\alpha x_1^\alpha x_1^\alpha x_3^{\alpha-1} - \lambda = 0.
\end{aligned}$$

The above equations imply

$$x_1 = x_2 = x_3$$

which, together with the constraint $x_1 + x_2 + x_3 - 1 = 0$, yields the candidate optimal solution $(x_1, x_2, x_3) = (1/3, 1/3, 1/3)$.
In summary, all candidate optimal solutions are

$$\begin{aligned}
&x_1 = x_2 = x_3 = 1/3\\
&x_1 = 0,\ x_2 + x_3 = 1,\ x_2 \ge 0,\ x_3 \ge 0\\
&x_2 = 0,\ x_1 + x_3 = 1,\ x_1 \ge 0,\ x_3 \ge 0\\
&x_3 = 0,\ x_1 + x_2 = 1,\ x_1 \ge 0,\ x_2 \ge 0.
\end{aligned}$$

c) At the point $(x_1, x_2, x_3) = (1/3, 1/3, 1/3)$ the only active constraint is the equality constraint. Hence, the second order sufficient conditions of optimality are

$$s'\nabla^2 L s > 0$$

for all $s \neq 0$ such that

$$[1,\ 1,\ 1]s = 0.$$

Note now that at the considered point

$$\nabla^2 L = -\left(\frac{1}{3}\right)^{3\alpha-2} \begin{bmatrix} \alpha(\alpha-1) & \alpha^2 & \alpha^2 \\ \alpha^2 & \alpha(\alpha-1) & \alpha^2 \\ \alpha^2 & \alpha^2 & \alpha(\alpha-1) \end{bmatrix}$$

and the admissible $s$ can be parameterized as

$$s_a = [\beta, 0, -\beta]' \qquad s_b = [\gamma, -\gamma, 0]'.$$

As a result

$$s_a'\nabla^2 L s_a = 2\left(\frac{1}{3}\right)^{3\alpha-2}\beta^2\alpha > 0 \qquad s_b'\nabla^2 L s_b = 2\left(\frac{1}{3}\right)^{3\alpha-2}\gamma^2\alpha > 0$$

which show that the considered point is a local minimizer.