EE4-29
EE9CS3-2

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2011

MSc and EEE/ISE PART IV: MEng and ACGI

OPTIMIZATION

Friday, 6 May 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. Gelenbe

OPTIMISATION

1. Consider the function

$$f(x) = x_1^4 + x_1x_2 + \frac{1}{2}x_2^2.$$

a) Compute the stationary points of the function. [2 marks]

b) Using second order sufficient conditions classify the stationary points determined in part a), that is say which is a local minimum, or a local maximum, or a saddle point. [6 marks]

c) Sketch on the (x_1, x_2)-plane the level lines of the function f. [4 marks]

d) Consider the point $P_0 = (0,0)$.

i) Determine a direction d_0 which is a descent direction for f at P_0. [4 marks]

ii) Consider the problem of performing an exact line search along the direction d_0 starting from P_0. Determine a solution to such a problem. [4 marks]

2. Consider the function

$$f(x) = ax^2 + bx + c,$$

where $a < 0$, b and c are given constants, and the problem of maximizing f in the admissible set $x \in [0, 10]$. This problem can be solved with two approaches.

a) Consider the function f as a function of a scalar variable defined in a compact set. Compute the maximizer of f in the admissible set.
(Hint: Note that the value of the maximizer depends upon the constants a, b and c and that the maximizer may not be a stationary point of f.) [8 marks]

b) Study the problem as a constrained optimization problem.

i) Write first order necessary conditions of optimality. [4 marks]

ii) Using the first order necessary conditions of optimality in part b.i), compute candidate optimal solutions. [4 marks]

iii) Verify that the solution determined in part a) coincides with the solution determined in part b.ii). [4 marks]

3. Consider the optimisation problem

$$\begin{cases} \min\limits_{x_1,x_2} \; 1 - x_1^2 - x_2^2, \\ \\ x_1 \geq 0, \\ x_2 \geq 0, \\ x_1 + x_2 - 1 \leq 0. \end{cases}$$

a) Sketch in the (x_1, x_2)-plane the admissible set and the level lines of the objective function. Show that all admissible points are regular points for the constraints. Hence, using only graphical considerations determine the optimal solutions of the considered problem. [4 marks]

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

c) Using the conditions in part b) determine candidate optimal solutions for the considered optimisation problem. [6 marks]

d) Write the second order sufficient conditions of optimality for the candidate optimal solutions determined in part c). Discuss if the sufficient conditions allow to classify the candidate optimal solutions or otherwise. [6 marks]

4. Consider the function

$$\phi(b) = \min_{x_1,x_2,x_3} \; -(x_1 + x_2 + x_3),$$

where x_1, x_2 and x_3 are such that

$$x_1^2 + 2x_2^2 + 3x_3^2 \leq b^2, \qquad x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_3 \geq 0,$$

and b is a constant. The function $\phi(b)$ can be computed solving the given optimization problem.

a) Show, without solving the optimization problem, that $\phi(0) = 0$ and $\phi(b) = \phi(-b)$. [2 marks]

b) Write first order necessary conditions of optimality for the considered optimization problem. [6 marks]

c) Using the conditions derived in part b) compute candidate optimal solutions. [10 marks]

d) Write explicitly the function $\phi(b)$. Verify that, as established in part a), $\phi(b) = \phi(-b)$ and $\phi(0) = 0$. [2 marks]

5. Consider the optimisation problem

$$\begin{cases} \min\limits_{x_1,x_2,x_3} & -(x_1x_2+x_2x_3+x_1x_3), \\ & x_1+x_2+x_3=1. \end{cases}$$

a) Using the method of constraint elimination determine a solution for the considered problem. [4 marks]

b) Consider the use of an exact penalty function to solve the considered problem.

i) Write the exact penalty function for the problem. [4 marks]

ii) Determine the unique stationary point of the exact penalty function. [6 marks]

iii) Show that the exact penalty function has a global minimizer for any positive value of its parameter ε, and that this global minimizer coincides with the solution determined in part a). [6 marks]

6. Consider the optimisation problem

$$\begin{cases} \min\limits_{x_1,x_2} (x_1-2)^2+(x_1-2x_2)^2, \\ x_1-x_2 \le 0. \end{cases}$$

a) Write first order necessary conditions of optimality for the considered optimization problem. [2 marks]

b) Using the conditions derived in part a) compute candidate optimal solutions. [4 marks]

c) The considered problem can be solved minimizing the barrier function

$$B_\varepsilon(x_1,x_2)=(x_1-2)^2+(x_1-2x_2)^2-\frac{\varepsilon^2}{x_1-x_2},$$

with $\varepsilon > 0$ and sufficiently small.

i) Compute, approximately, the unique stationary point of the function B_ε.
(Hint: show that the stationary point can be approximated, for ε small, by $(x_1^\star, 1)$, where $x_1^\star = 1-0.5\varepsilon$.) [6 marks]

ii) Show that the stationary point determined in part c.i) is a minimizer of B_ε. [6 marks]

iii) Show that the approximate stationary point determined in part c.i) is admissible for all $\varepsilon > 0$ and that, as ε goes to zero, the stationary point approaches the solution determined in part b). [2 marks]

EE4-29
EE9 CS7-2

Optimisation - Model answers 2011

(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

$$0 = \nabla f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4x_1^3 + x_2 \\ x_1 + x_2 \end{bmatrix}.$$

Replacing the second equation in the first yields $x_1(4x_1^2 - 1) = 0$. Hence, the stationary points are

$$P_1 = (0,0) \qquad P_2 = \left(\frac{1}{2}, -\frac{1}{2}\right) \qquad P_2 = \left(-\frac{1}{2}, \frac{1}{2}\right).$$

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} 12x_1^2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Note that

$$\nabla^2 f(P_1) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

is indefinite and

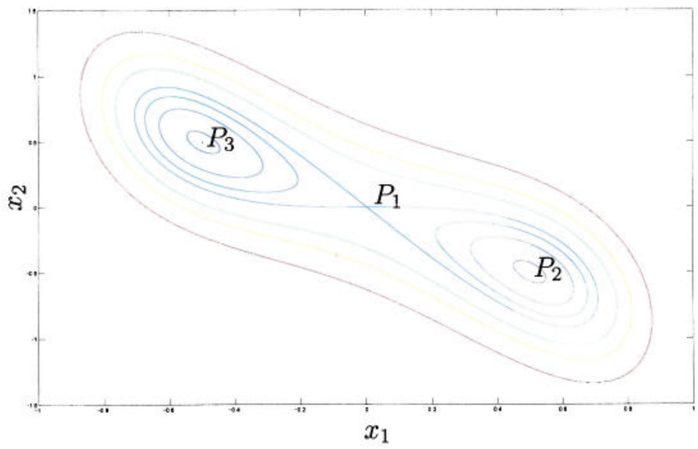
$$\nabla^2 f(P_2) = \nabla^2 f(P_3) = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$$

is positive definite. Hence P_2 and P_3 are local minimizers, and P_1 is a saddle point.

c) The level lines of f can be sketched using the following considerations.

- Around the minimizers the level lines are closed.
- The value of f at the saddle point P_1 is 0.
- The value of f at the local minimizers P_2 and P_3 is $-1/16$. There is a closed level line which goes through the saddle point and encircles both local minimizers.

A sketch of the level lines is in the figure in the next page .



d) Note that $\nabla f(P_0) = 0$, hence for any direction d the scalar product $\nabla' f\ d$ is zero, *i.e.* it is not possible to use first order sufficient conditions to establish if a direction is a descent direction.

i) Let, for example, $d_0 = [1, -1]'$ and consider the restriction of the function f along d_0, with *initial point* P_0, namely

$$f(P_0 + \alpha d_0) = \alpha^2 \left(-\frac{1}{2} + \alpha^2 \right).$$

For any $\alpha > 0$ and sufficiently small (namely $\alpha \in (0, 1/\sqrt{2})$)

$$f(P_0) > f(P_0 + \alpha d_0),$$

hence d_0 is a descent direction for f at P_0.

ii) To solve an exact line search problem along d_0 at P_0 one has to find the global minimizer, if it exists, of $f(P_0 + \alpha d_0)$. Note that the function

$$f(P_0 + \alpha d_0) = \alpha^2(-1/2 + \alpha^2)$$

is radially unbounded (and bounded from below), hence possesses a global minimizer, which is a stationary point. The stationary points of this function are $\alpha = 0$ (local maximizer) and $\alpha = \pm 1/2$ (local minimizer). Hence, an exact line search along d_0, starting at P_0, gives either the point P_2 or the point P_3.

Question 2

a) The maximum of f in the set $x \in [0, 10]$ is achieved either at a stationary point or at $x = 0$ and $x = 10$. The only stationary point of f is $x^\star = -\frac{b}{2a}$. Note now that

$$f(0) = c \qquad f(10) = 100a + 10b + c \qquad f(x^\star) = -\frac{b^2}{2a} + c.$$

Consider now the following three cases.

- $x^\star < 0$. In this case the maximizer is either $x = 0$ or $x = 10$. Note also that in this case a and b have the same sign. As a result, since $a < 0$, $b < 0$ and the maximizer is $x = 0$.
- $x^\star > 10$. In this case the maximizer is either $x = 0$ or $x = 10$. Note also that in this case a and b have opposite sign, hence $b > 0$. In addition, $b > -20a$, hence $5b + 100a > 0$. As a result $100a + 10b > 0$ and the maximizer is $x = 10$.
- $0 < x^\star < 10$. In this case the maximizer is the stationary point if the function is concave, *i.e.* if $a < 0$, which is the case.

b) The problem can be written as

$$\begin{aligned} &\max_x \;\; ax^2 + bx + c, \\ &-x \leq 0, \\ &x - 10 \leq 0. \end{aligned}$$

The Lagrangian of the problem is (note the change in sign of the objective function to transform the problem into a minimization problem)

$$L(x, \rho_1, \rho_2) = -(ax^2 + bx + c) - \rho_1 x + \rho_2(x - 10).$$

i) The necessary conditions of optimality are

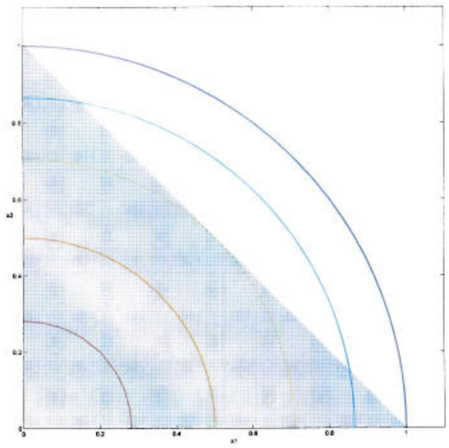
$$\begin{gathered} \frac{dL}{dx} = -2ax - b - \rho_1 + \rho_2 = 0 \\ -x \leq 0 \qquad x - 10 \leq 0 \\ \rho_1 \geq 0 \qquad \rho_2 \geq 0 \\ \rho_1 x = 0 \qquad \rho_2(x - 10) = 0. \end{gathered}$$

ii) Using the complementarity conditions we consider four cases.

- $\rho_1 = 0$, $\rho_2 = 0$. In this case the only candidate solution is $x = -\frac{b}{2a}$.
- $\rho_1 > 0$, $\rho_2 = 0$. In this case the only candidate solution is $x = 0$. Note that $\rho_1 > 0$ implies $b < 0$, hence $0 > -\frac{b}{2a}$.
- $\rho_1 = 0$, $\rho_2 > 0$. In this case the only candidate solution is $x = 10$. Note that $\rho_2 > 0$ implies $-\frac{b}{2a} > 10$.
- $\rho_1 > 0$, $\rho_2 > 0$. In this case there is no candidate solution.

iii) As in part a), there are three cases. If $0 < -\frac{b}{2a} < 10$ then the optimal solution is $x^\star = -\frac{b}{2a}$. If $0 > -\frac{b}{2a}$ then the optimal solution is $x^\star = 0$. If $-\frac{b}{2a} > 10$ then the optimal solution is $x^\star = 10$.

Question 3



a) The admissible set is the shaded area in the figure above. The level lines are the solid lines. Note that the gradient vectors of the constraints are independent at all points of the admissible set: hence all points are regular points. The function f decreases in the direction of the arrows. There are two (global) minimizers: $(x_1, x_2) = (1, 0)$ and $(x_1, x_2) = (0, 1)$. The value of the function at the optimal point is $f(1, 0) = f(0, 1) = 0$.

b) The Lagrangian of the problem is

$$L(x_1, x_2, \rho_1, \rho_2, \rho_3) = 1 - x_1^2 - x_2^2 - \rho_1 x_1 - \rho_2 x_2 + \rho_3(x_1 + x_2 - 1).$$

The first order necessary conditions of optimality are

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

$$x_1 \geq 0 \qquad x_2 \geq 0 \qquad x_1 + x_2 - 1 \leq 0 \qquad \rho_1 \geq 0 \qquad \rho_2 \geq 0 \qquad \rho_3 \geq 0$$

$$\rho_1 x_1 = 0 \qquad \rho_2 x_2 = 0 \qquad \rho_3(x_1 + x_2 - 1) = 0$$

c) To determine candidate optimal solutions exploiting the complementarity conditions, consider the following cases.

- $\rho_1 = 0,\ \rho_2 = 0,\ \rho_3 = 0$. The only candidate solution is $x_1 = x_2 = 0$.
- $\rho_1 = 0,\ \rho_2 = 0,\ \rho_3 > 0$. The only candidate solution is $x_1 = x_2 = 1/2$.
- $\rho_1 = 0,\ \rho_2 > 0,\ \rho_3 = 0$. No candidate solution.
- $\rho_1 = 0,\ \rho_2 > 0,\ \rho_3 > 0$. The only candidate sol ution is $x_1 = 1$, $x_2 = 0$.
- $\rho_1 > 0,\ \rho_2 = 0,\ \rho_3 = 0$. No candidate solution.
- $\rho_1 > 0,\ \rho_2 = 0,\ \rho_3 > 0$. The only candidate solution is $x_1 = 0$, $x_2 = 1$.
- $\rho_1 > 0,\ \rho_2 > 0,\ \rho_3 = 0$. No candidate solution.

- $\rho_1 > 0,\ \rho_2 > 0,\ \rho_3 > 0$. No candidate solution.

In summary there are four candidate optimal solutions

$$P_1 = (0,0) \qquad P_2 = (1,0) \qquad P_3 = (0,1) \qquad P_4 = (1/2, 1/2).$$

d) Note that

$$\nabla_{xx} L = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

consider now the four candidate points.

- P_1. The sufficient condition requires

$$s' \nabla_{xx} L s > 0$$

for all $s \neq 0$, such that

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} s = 0.$$

Note that the set of *admissible* s is empty, hence the sufficient conditions do not allow to classify the point P_0. Note also that the condition of strict complementarity does not hold at P_0.

- P_2. The sufficient condition requires

$$s' \nabla_{xx} L s > 0$$

for all $s \neq 0$, such that

$$\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} s = 0.$$

Note that the set of *admissible* s is empty, hence the sufficient conditions do not allow to classify the point P_2.

- P_3. The sufficient condition requires

$$s' \nabla_{xx} L s > 0$$

for all $s \neq 0$, such that

$$\begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} s = 0.$$

Note that the set of *admissible* s is empty, hence the sufficient conditions do not allow to classify the point P_3.

- P_4. The sufficient condition requires

$$s' \nabla_{xx} L s > 0$$

for all $s \neq 0$, such that

$$\begin{bmatrix} 1 & 1 \end{bmatrix} s = 0,$$

which is not the case.

In summary, the second order sufficient conditions rule out optimality of P_4, yet do not show that P_2 and P_3 are (global) minimizers, and that P_0 is a (global) maximizer.

Question 4

a) For $b = 0$ the admissible set is composed of the singleton $(x_1, x_2, x_3) = (0, 0, 0)$. At this point the objective function $-(x_1 + x_2 + x_3)$ is zero, hence $\phi(0) = 0$.
Replacing b with $-b$ does not change the formulation of the optimization problem, hence $\phi(b) = \phi(-b)$.

b) The Lagrangian of the problem is

$$L(x_1, x_2, x_3, \rho_1, \rho_2, \rho_3, \rho_4) = -(x_1+x_2+x_3)-\rho_1 x_1-\rho_2 x_2-\rho_3 x_3+\rho_4(x_1^2+2x_2^2+3x_3^2-b^2).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = -1-\rho_1+2\rho_4 x_1 \qquad 0 = \frac{\partial L}{\partial x_2} = -1-\rho_2+4\rho_4 x_2 \qquad 0 = \frac{\partial L}{\partial x_3} = -1-\rho_3+6\rho_4 x_3$$

$$x_1^2 + 2x_2^2 + 3x_3^2 - b^2 \leq 0 \qquad -x_1 \leq 0 \qquad -x_2 \leq 0 \qquad -x_3 \leq 0$$

$$\rho_1 \geq 0 \qquad \rho_2 \geq 0 \qquad \rho_3 \geq 0 \qquad \rho_4 \geq 0$$

$$\rho_1 x_1 = 0 \qquad \rho_2 x_2 = 0 \qquad \rho_3 x_3 = 0 \qquad \rho_4(x_1^2 + 2x_2^2 + 3x_3^2 - b^2) = 0.$$

c) Since ρ_1, ρ_2 and ρ_3 are non-negative, the conditions

$$0 = \frac{\partial L}{\partial x_1} = \frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial x_3}$$

imply $x_1 \neq 0$, $x_2 \neq 0$, $x_3 \neq 0$ and $\rho_4 > 0$. As a result, $\rho_1 = 0$, $\rho_2 = 0$, $\rho_3 = 0$ and $x_1^2 + 2x_2^2 + 3x_3^2 - b^2 = 0$. The candidate solutions are therefore described by

$$x_1 = \frac{1}{2\rho_4} \qquad x_2 = \frac{1}{4\rho_4} \qquad x_3 = \frac{1}{6\rho_4},$$

with ρ_4 such that

$$\frac{1}{\rho_4^2}\left(\frac{1}{4} + \frac{1}{8} + \frac{1}{12}\right) = b^2,$$

yielding the candidate optimal solution

$$x_1 = \frac{|b|}{2}\sqrt{\frac{24}{11}} \qquad x_2 = \frac{|b|}{4}\sqrt{\frac{24}{11}} \qquad x_3 = \frac{|b|}{6}\sqrt{\frac{24}{11}}$$

d) A direct substitution yields

$$\phi(b) = -|b|\sqrt{\frac{22}{12}}.$$

Hence, $\phi(0) = 0$, and $\phi(b) = \phi(-b)$.

Question 5

a) Using the constrains we obtain, for example,

$$x_3 = 1 - x_1 - x_2$$

which substituted in the objective function gives

$$x_1^2 + x_1x_2 + x_2^2 - x_2 - x_1.$$

This is a quadratic function, with a positive definite Hessian matrix, hence it is a strictly convex function. The (global) minimizer is therefore the stationary point of the function, which is the solution of the equations

$$0 = 2x_1 + x_2 - 1 \qquad 0 = x_1 + 2x_2 - 1,$$

namely $x_1 = 1/3$ and $x_2 = 1/3$. Using again the equation of the constrain we obtain the solution of the optimization problem: $(x_1, x_2, x_3) = (1/3, 1/3, 1/3)$.

b) An exact penalty function for a constraint optimization problem with equality constraints is

$$G(x) = f(x) - g'(x)(\frac{\partial g}{\partial x}\frac{\partial g'}{\partial x})^{-1}\frac{\partial g}{\partial x}\nabla f + \frac{1}{\epsilon}\|g(x)\|^2,$$

with $\epsilon > 0$, and sufficiently small.

i) For the considered problem we have

$$G(x_1, x_2, x_3) = -x_1x_2 - x_2x_3 - x_1x_3 + \frac{2}{3}(x_1+x_2+x_3-1)(x_1+x_2+x_3) + \frac{1}{\epsilon}(x_1+x_2+x_3-1)^2.$$

The function is well-defined for all (x_1, x_2, x_3), since $\frac{\partial g}{\partial x}\frac{\partial g'}{\partial x}$ is constant and invertible.

ii) The stationary points of the function G are the solutions of

$$0 = \frac{dG}{dx_1} = \frac{1}{3}\left(4x_1 + x_2 + x_3 - 2 + \frac{6}{\epsilon}(x_1 + x_2 + x_3 - 1)\right)$$
$$0 = \frac{dG}{dx_2} = \frac{1}{3}\left(x_1 + 4x_2 + x_3 - 2 + \frac{6}{\epsilon}(x_1 + x_2 + x_3 - 1)\right)$$
$$0 = \frac{dG}{dx_3} = \frac{1}{3}\left(x_1 + x_2 + 4x_3 - 2 + \frac{6}{\epsilon}(x_1 + x_2 + x_3 - 1)\right)$$

These equations have the unique solution $(x_1, x_2, x_3) = (1/3, 1/3, 1/3)$.

iii) The function $G(x_1, x_2, x_3)$ is a quadratic function with Hessian matrix

$$\nabla^2 G = \frac{1}{3}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 4 & 1 \end{bmatrix} + \frac{2}{\epsilon}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Note that, for any $\epsilon > 0$, $\nabla^2 G$ is positive definite, hence G is a (strictly) convex function. As a result, its unique stationary point, determined in part b.ii), is a global minimizer of G, which coincides with the optimal solution determined in part a).

Question 6

a) The Lagrangian of the problem is

$$L(x_1, x_2, \rho) = (x_1 - 2)^2 + (x_1 - 2x_2)^2 + \rho(x_1 - x_2).$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 4x_1 - 4x_2 - 4 + \rho \qquad 0 = \frac{\partial L}{\partial x_2} = -4x_1 + 8x_2 - \rho$$

$$x_1 - x_2 \leq 0 \qquad \rho \geq 0 \qquad \rho(x_1 - x_2) = 0$$

b) The only candidate solution is $x_1 = x_2 = 1$, with multiplier $\rho = 4$.

c) Note that, for $\epsilon > 0$, the barrier function is not defined for $x_1 = x_2$, *i.e.* it is not defined on the boundary of the admissible set.

i) The stationary points of the barrier function are the solutions of

$$0 = 4x_1 - 4x_2 - 4 + \frac{\epsilon^2}{(x_1 - x_2)^2} \qquad 0 = -4x_1 + 8x_2 - \frac{\epsilon^2}{(x_1 - x_2)^2}$$

Adding these two equations yields $-4 + 4x_2 = 0$, hence $x_2 = 1$. Replacing $x_2 = 1$ and $x_1 = 1 - \epsilon/2$, as suggested, in the first equation and performing a series expansion yields

$$0 = -2\epsilon + \ldots.$$

Hence, an approximation of the stationary point is

$$\left(1 - \frac{\epsilon}{2}, 1\right).$$

ii) The Hessian matrix of the function B_ϵ at the approximate stationary point is

$$\nabla^2 B_\epsilon = 4 \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} + \frac{16}{\epsilon} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Note that $\nabla^2 B_\epsilon$ is positive definite for $\epsilon > 0$. As a result, the stationary point of B_ϵ is an approximate local minimizer.

iii) The approximate stationary point is such that

$$x_1 - x_2 = \left(1 - \frac{\epsilon}{2}\right) - 1 = -\frac{\epsilon}{2} < 0$$

for all $\epsilon > 0$, *i.e.* it is admissible for all $\epsilon > 0$. In addition, as ϵ goes to zero, the approximate stationary point tends to the solution of the considered optimization problem.