IMPERIAL COLLEGE LONDON

master

E4.29
C1.1
ISE4.55

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2007

MSc and EEE PART IV: MEng and ACGI

OPTIMIZATION

Wednesday, 9 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) : M.M. Draief

Optimisation

1. Consider the problem of minimizing the function

$$f(x_1, x_2) = \frac{1}{2n+2} x_1^{2n+2} - x_1 x_2 + \frac{1}{2} x_2^2,$$

where n is a positive integer.

a) Compute all stationary points of the function. [4 marks]

b) 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. [8 marks]

c) Show that the function f is radially unbounded and hence compute the global minimum of f. Is the global minimizer unique? [4 marks]

d) Consider the point $P_0 = (0, 0)$ and the direction

$$d = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Show that the direction d is a descent direction for f at P_0. [4 marks]

2. Consider the problem of approximating a matrix $Q \in \mathbb{R}^{n \times n}$ with a matrix of the form $A = \rho I$, with I the identity matrix of dimension $n \times n$ and $\rho \geq 0$.

As a measure of the distance between the two matrices we could use either the square of the Frobenius norm or the infinity norm. The Frobenius norm of a matrix $L \in \mathbb{R}^{n \times n}$ is defined as

$$\|L\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} L_{ij}^2},$$

where the L_{ij}'s denote the entry of the matrix L. The infinity norm of a matrix $L \in \mathbb{R}^{n \times n}$ is defined as

$$\|L\|_\infty = \max_i \sum_{j=1}^{n} |L_{ij}|.$$

The optimal approximation problem is thus the problem of determining the nonnegative constant ρ which minimizes

$$\|Q - \rho I\|_F^2$$

or

$$\|Q - \rho I\|_\infty.$$

a) Show that the considered optimal approximation problems can be written as constrained minimization problems with one inequality constraint. [2 marks]

b) Consider the Frobenius norm. Solve the problem derived in part a). Show that if $\text{trace}(Q) > 0$ then the optimal ρ is positive, and if $\text{trace}(Q) \leq 0$ then the optimal ρ is zero.
(The trace of a matrix is the sum of its diagonal elements.) [6 marks]

c) Consider the infinity norm and assume that $n = 2$, hence

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix},$$

that $0 < Q_{11} < Q_{22}$ and that $|Q_{12}| = |Q_{21}|$.

i) Sketch the graph of the function to be minimized. [4 marks]

ii) Argue that the optimal solution $\rho_\star$ is such that

$$0 < Q_{11} < \rho_\star < Q_{22}.$$

[4 marks]

iii) Compute the optimal solution $\rho_\star$. [4 marks]

3. Newton's method for the minimization of a function $f : \mathbb{R} \to \mathbb{R}$ is based on a quadratic approximation of the function at a given point. An alternative way to construct a quadratic approximation that does not require the computation of the second derivative is to consider an approximation based on the knowledge of two points x_k and x_{k-1} and of the values $f(x_k)$, $\frac{df(x_k)}{dx}$ and $\frac{df(x_{k-1})}{dx}$. Such an approximation is given by

$$q(x) = f(x_k) + \frac{df(x_k)}{dx}(x - x_k) + \frac{\dfrac{df(x_{k-1})}{dx} - \dfrac{df(x_k)}{dx}}{x_{k-1} - x_k} \frac{(x - x_k)^2}{2}.$$

a) Show that the function $q(x)$ is such that

$$q(x_k) = f(x_k), \qquad \frac{dq(x_k)}{dx} = \frac{df(x_k)}{dx}, \qquad \frac{dq(x_{k-1})}{dx} = \frac{df(x_{k-1})}{dx}.$$

[4 marks]

b) Compute the stationary point $x_\star$ of $q(x)$. [2 marks]

c) Consider the algorithm, known as the method of the false position, obtained by setting $x_{k+1} = x_\star$, with $x_\star$ as in part b), and argue that this algorithm provides an approximation of Newton's method that does not require the computation of the second derivative of f. [2 marks]

d) Show that the method of the false position applied to the minimization of a quadratic function $f = ax^2 + bx + c$, with $a > 0$, coincides with Newton's method. [4 marks]

e) Consider the function $f = \dfrac{x^4}{4} + x$. This function has a global minimizer at $x = -1$.

i) Show that the method of the false position yields the iteration

$$x_{k+1} = x_k - (x_k^3 + 1)\frac{1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2}.$$

[2 marks]

ii) Evaluate

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

and show that if $\lim_{k\to\infty} x_k = -1$ then

$$\lim_{k\to\infty} \frac{|\mathscr{E}_{k+1}|}{\mathscr{E}_k^2} = 1.$$

Hence, quantify the speed of convergence of the method. [6 marks]

4. Consider the optimization problem

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

a) Sketch in the (x_1, x_2)-plane the admissible set and show that there is a point which is not a regular point for the constraints. [4 marks]

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

c) Find candidate optimal solutions for the considered problem. [8 marks]

d) Prove that the non-regular point for the constraints is the global minimizer for the considered problem. [4 marks]

5. Consider the optimization problem

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

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. [4 marks]

c) This constrained optimization problem can be transformed into an unconstrained optimization problem by defining the so-called mixed penalty-barrier function

$$F_\varepsilon(x_1, x_2) = x_1^2 + x_2^2 + \frac{1}{\varepsilon}(x_2 - x_1 - 1)^2 + \frac{\varepsilon}{x_1},$$

with $\varepsilon > 0$ and considering the unconstrained minimization of $F_\varepsilon(x_1, x_2)$. Determine the stationary points of $F_\varepsilon(x_1, x_2)$. (Hint: solve $\nabla_{x_2} F_\varepsilon(x_1, x_2) = 0$ for x_2, and replace the obtained solution in the equation $\nabla_{x_1} F_\varepsilon(x_1, x_2) = 0$. Solve this last equation assuming that $x_1 = \alpha \varepsilon^{1/2}$, for some $\alpha > 0$ to be determined, and neglecting all terms ε^k, for $k \geq 1/2$.) [10 marks]

d) Show that the stationary point of $F_\varepsilon(x_1, x_2)$ computed in part c) tends, as ε tends to zero, to the optimal solution determined in part b). [2 marks]

6. Consider the optimization problem

$$\begin{cases} \min\limits_{x_1,\, x_2} x_1 x_2, \\ \frac{1}{2}x_1^2 + 2x_2^2 = 1. \end{cases}$$

a) State first order necessary conditions of optimality for such a constrained optimization problem. [2 marks]

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

c) Transform the minimization problem into an unconstrained minimization problem using the method of the exact augmented Lagrangian functions and write explicitly the exact augmented Lagrangian functions for the considered problem. [4 marks]

d) Show that the candidate optimal solutions determined in part b) are stationary points of the exact augmented Lagrangian function. [4 marks]

e) Find the global minimum for the considered problem. Is the global minimizer unique? [4 marks]

E4.29
C1.1
ISE 4.55

Optimisation - Model answers 2007

(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 equation

$$0 = \nabla f = \begin{bmatrix} x_1^{2n+1} - x_2 \\ -x_1 + x_2 \end{bmatrix}.$$

The second equation yields $x_2 = x_1$, hence the first equation becomes

$$0 = x_1^{2n+1} - x_1 = x_1(x_1^{2n} - 1).$$

The (real) solutions of this equation are $x_1 = 0$, $x_1 = 1$ and $x_1 = -1$. In summary, the function f has three stationary points

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

b) Note that (recall that n is a positive integer)

$$\nabla^2 f = \begin{bmatrix} (2n+1)x_1^{2n} & -1 \\ -1 & 1 \end{bmatrix}.$$

Hence

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

which is an indefinite matrix, and

$$\nabla^2 f(P_b) = \nabla^2 f(P_c) = \begin{bmatrix} 2n+1 & -1 \\ -1 & 1 \end{bmatrix} > 0.$$

As a result P_a is a saddle point, and P_b and P_c are local minimizers.

c) Note that

$$f = \frac{1}{2n+2}x_1^{2n+2} - x_1x_2 + \frac{1}{2}x_2^2 = \frac{1}{2n+2}x_1^{2n+2} - x_1^2 + \left(x_1^2 - x_1x_2 + \frac{1}{2}x_2^2\right).$$

The function

$$\frac{1}{2n+2}x_1^{2n+2} - x_1^2 = x_1^2\left(\frac{1}{2n+2}x_1^{2n} - 1\right)$$

is radially unbounded, as a function of x_1 alone, and the function $x_1^2 - x_1x_2 + \frac{1}{2}x_2^2$ is radially unbouded as a function of x_1 and x_2. As a result the global minimum of f is also a local minimum. Note that (recall again that n is a positive integer)

$$f(P_b) = f(P_c) = -\frac{1}{2}\frac{n}{n+1} < 0,$$

hence both P_b and P_c are global minimizers.

d) The point P_0 coincides with the saddle point P_a. The function f along the direction d is given by

$$\phi(\alpha) = f(\alpha, \alpha) = \frac{1}{2n+2}\alpha^{2n+2} - \frac{1}{2}\alpha^2.$$

Note that $\phi(0) = 0$ and that $\phi(\alpha) < 0$ for $\alpha > 0$ and sufficiently small (namely for all $\alpha \in \left(0, (n+1)^{\frac{1}{2n}}\right)$, hence d is a descent direction for f at P_0.

(Note that $\phi(\alpha)$ is negative also for $\alpha \in \left(-(n+1)^{\frac{1}{2n}}, 0\right)$, i.e. $-d$ is also a descent direction for f at P_0, but this is not requested.)

Question 2

a) The optimal approximation problems can be written as

$$P_f : \begin{cases} \min_{\rho} \|Q - \rho I\|_F^2 \\ \rho \geq 0 \end{cases} \qquad \text{or as} \qquad P_\infty : \begin{cases} \min_{\rho} \|Q - \rho I\|_\infty \\ \rho \geq 0. \end{cases}$$

b) Note that

$$\begin{aligned} \|Q - \rho I\|_F^2 = \; & (Q_{11} - \rho)^2 + Q_{12}^2 + \cdots + Q_{1n}^2 + \\ & Q_{21}^2 + (Q_{22} - \rho)^2 + Q_{23}^2 + \cdots + Q_{2n}^2 + \cdots + Q_{n1}^2 + \cdots + Q_{2,n-1}^2 + (Q_{nn} - \rho)^2 \end{aligned}$$

hence

$$\|Q - \rho I\|_F^2 = n\rho^2 - 2\rho \overbrace{(Q_{11} + Q_{22} + \cdots + Q_{nn})}^{\text{trace}(Q)} + \text{constant terms.}$$

If $\text{trace}(Q) > 0$ the function $\|Q - \rho I\|_F^2$, which is convex, has a global minimum for $\rho = \frac{\text{trace}(Q)}{n}$. If $\text{trace}(Q) \leq 0$ the function $\|Q - \rho I\|_F^2$ is monotonically increasing for $\rho \geq 0$, hence it achieves its minimum, in the set $\rho \geq 0$, for $\rho = 0$.

c) The optimal approximation problem is now

$$\tilde{P}_\infty : \begin{cases} \min_{\rho} \Big(\max\left(|Q_{11} - \rho| + |Q_{12}|, |Q_{21}| + |Q_{22} - \rho|\right) \Big) \\ \rho \geq 0. \end{cases}$$

A sketch of the function to be minimized is in the figure. From this, it is clear that $0 < Q_{11} < \rho_\star < Q_{22}$. Note that $\rho_\star$ is such that

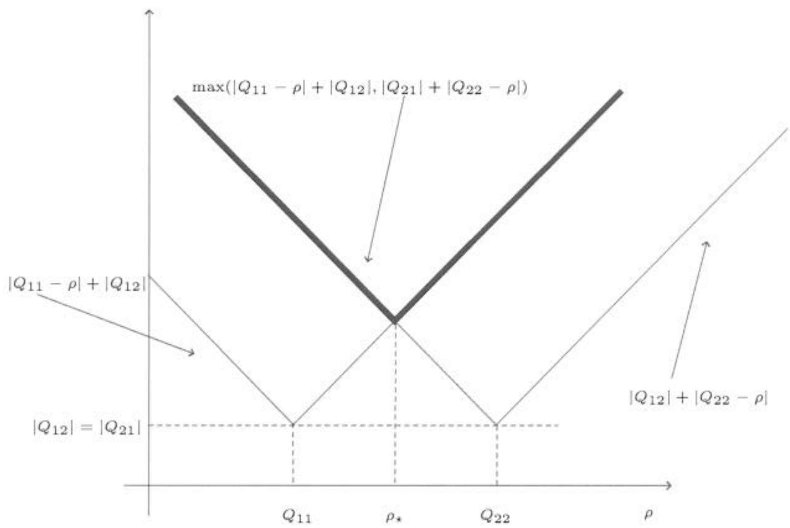
$$|Q_{11} - \rho_\star| + |Q_{12}| = |Q_{21}| + |Q_{22} - \rho_\star|.$$

However, because $0 < Q_{11} < \rho_\star < Q_{22}$ this can be rewritten as

$$\rho_\star - |Q_{11}| + |Q_{12}| = |Q_{21}| + |Q_{22}| - \rho_\star.$$

As a result (recall that $Q_{11} > 0$, $Q_{22} > 0$ and $|Q_{12}| = |Q_{21}|$)

$$\rho_\star = \frac{Q_{11} + Q_{22}}{2}.$$



max(|Q11 − ρ| + |Q12|, |Q21| + |Q22 − ρ|)
|Q11 − ρ| + |Q12|
|Q12| + |Q22 − ρ|
|Q12| = |Q21|
Q11
ρ⋆
Q22
ρ

Question 3

a) Setting $x = x_k$ in $q(x)$ yields $q(x_k) = f(x_k)$. Note that

$$\frac{dq(x)}{dx} = \frac{df(x_k)}{dx} + \frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}(x - x_k)$$

hence, setting $x = x_k$ and $x = x_{k-1}$ yields

$$\frac{dq(x_k)}{dx} = \frac{df(x_k)}{dx} \qquad \frac{dq(x_{k-1})}{dx} = \frac{df(x_{k-1})}{dx}.$$

b) The stationary point $x_\star$ of $q(x)$ is obtained solving the equation

$$\frac{dq(x)}{dx} = 0,$$

which yields

$$x_\star = x_k - \left(\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}\right)^{-1} \frac{df(x_k)}{dx}.$$

c) The method of the false position is therefore given by

$$x_{k+1} = x_k - \left(\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}\right)^{-1} \frac{df(x_k)}{dx}.$$

This algorithm is an approximation of Newton's method because the quantity

$$\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k}$$

is an approximation of $\frac{d^2 f(x)}{dx^2}$ at $x = x_k$. Note however that, unlike Newton's method, the method of the false position does not need the computation of the second derivative: it uses an approximation.

d) For quadratic functions one has

$$\frac{d^2 f(x)}{dx^2} = 2a$$

and

$$\frac{\frac{df(x_{k-1})}{dx} - \frac{df(x_k)}{dx}}{x_{k-1} - x_k} = \frac{(2ax_{k-1} + b) - (2ax_k + b)}{x_{k-1} - x_k} = 2a,$$

hence, for such functions, Newton's method and the method of the false position coincide.

e) If $f = \frac{x^4}{4} + x$ then $\frac{df(x)}{dx} = x^3 + 1$, and replacing in the expression of the considered method yields

$$x_{k+1} = x_k - \frac{x_{k-1} - x_k}{(x_{k-1}^3 + 1) - (x_k^3 + 1)}(x_k^3 + 1) = x_k - \frac{x_{k-1} - x_k}{x_{k-1}^3 - x_k^3}(x_k^3 + 1),$$

hence, noting that

$$x_{k-1}^3 - x_k^3 = (x_{k-1} - x_k)(x_{k-1}^2 + x_{k-1}x_k + x_k^2)$$

yields

$$x_{k+1} = x_k - (x_k^3 + 1)\frac{1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2}.$$

Note that

$$\begin{aligned} x_{k+1} + 1 = & \ x_k + 1 - (x_k^3 + 1)\frac{1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2} \\ = & \ (x_k + 1)(x_{k-1} + 1)\frac{x_k + x_{k-1} - 1}{x_{k-1}^2 + x_{k-1}x_k + x_k^2}, \end{aligned}$$

hence

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

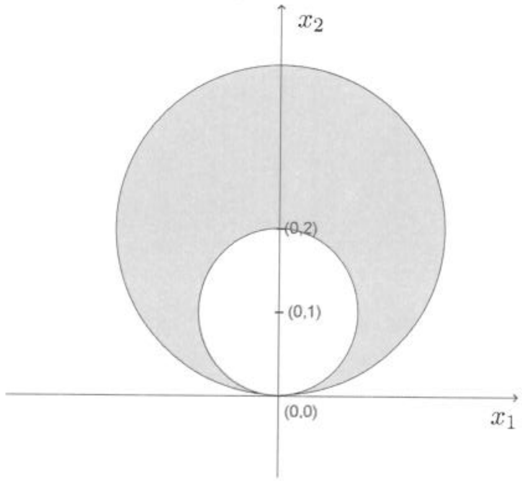
If $x_k \to -1$ then also $x_{k-1} \to -1$, hence $\frac{|\mathcal{E}_{k+1}|}{\mathcal{E}_k^2} = 1$, which shows that the algorithm has quadratic speed of convergence (if it converges).

Question 4

a) The admissible set is the set outside a circle of radius one and centered at $(0,1)$ and inside a circle of radius two and centered at $(0,2)$, which is the shaded region in the figure. The point $(0,0)$ is not a regular point for the constraints because at this point both constraints are active and their gradients, namely

$$\begin{bmatrix} 2x_1 \\ 2(x_2-1) \end{bmatrix} \qquad \begin{bmatrix} 2x_1 \\ 2(x_2-2) \end{bmatrix},$$

evaluated at the point, are linearly dependent.



b) To write necessary conditions of optimality rewrite first the constraints as

$$1 - x_1^2 - (x_2-1)^2 \le 0 \qquad x_1^2 + (x_2-2)^2 - 4 \le 0$$

and define the Lagrangian function

$$L(x_1, x_2, \mu_1, \mu_2) = x_1^2 + x_2 + \mu_1(1 - x_1^2 - (x_2-1)^2) + \mu_2(x_1^2 + (x_2-2)^2 - 4).$$

The necessary conditions of optimality are

$$\frac{dL}{dx_1} = 2x_1 - 2\mu_1 x_1 + 2\mu_2 x_1 = 0 \qquad \frac{dL}{dx_2} = 1 - 2\mu_1(x_2-1) + 2\mu_2(x_2-2) = 0$$

$$1 - x_1^2 - (x_2-1)^2 \le 0 \qquad x_1^2 + (x_2-2)^2 - 4 \le 0$$

$$\mu_1 \ge 0 \qquad \mu_2 \ge 0$$

$$\mu_1(1 - x_1^2 - (x_2-1)^2) = 0 \qquad \mu_2(x_1^2 + (x_2-2)^2 - 4) = 0.$$

c) To find candidate optimal solutions we exploit the complementarity conditions, hence we have four possibilities.

- $\mu_1 = 0$ and $\mu_2 = 0$.
  This selection yields $0 = \frac{dL}{dx_2} = 1$, hence no candidate optimal solution.

- $\mu_1 = 0$ and $x_1^2 + (x_2 - 2)^2 - 4 = 0$.
  This selection yields, from $0 = \frac{dL}{dx_1}$, either $x_1 = 0$ or $\mu_2 = -1$. The first option yields $x_2 = 0$ or $x_2 = 4$, whereas the second option violates the positivity of μ_2. Moreover, the selection $x_1 = 0$ and $x_2 = 4$ yields, from $0 = \frac{dL}{dx_2}$, $\mu_2 < 0$, hence it is not a candidate solution.
- $1 - x_1^2 - (x_2 - 1)^2 = 0$ and $\mu_2 = 0$.
  This selection yields, from $0 = \frac{dL}{dx_1}$, $x_1 = 0$ or $\mu_1 = 1$. The first option yields $x_2 = 0$ or $x_2 = 2$. The second option yields, from $0 = \frac{dL}{dx_2}$, $x_2 = 3/2$, hence, from $1 - x_1^2 - (x_2 - 1)^2 = 0$, $x_1 = \pm\frac{\sqrt{3}}{2}$.
- $1 - x_1^2 - (x_2 - 1)^2 = 0$ and $x_1^2 + (x_2 - 2)^2 - 4 = 0$.
  The only point consistent with these conditions is $(0, 0)$.

In summary the candidate solutions obtained so far are as follows.

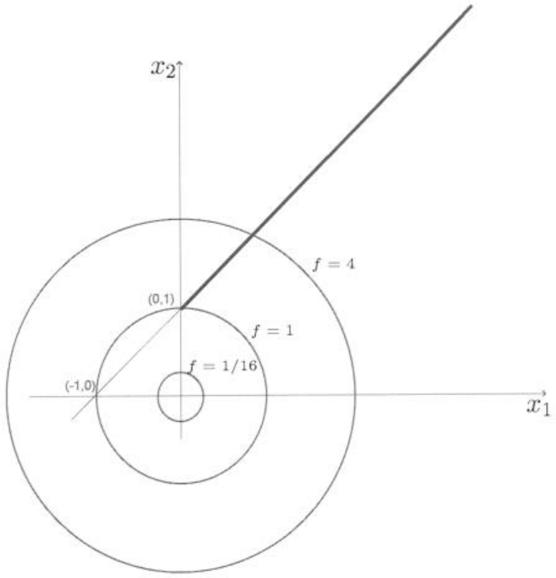
- $(0, 0)$.
- $(0, 2)$.
- $(\pm\frac{\sqrt{3}}{2}, \frac{3}{2})$.

Hence there are four candidate optimal solutions.

d) The nonregular point $(0, 0)$ is such that $x_1^2 + x_2 = 0$. Note now that the function $x_1^2 + x_2$ is always nonnegative in the admissible set and it is zero, in the admissible set, if and only if $x_1 = x_2 = 0$. Hence the nonregular point is a global minimum for the considered problem. Note that it is not possible to associate, in a unique way, a pair of optimal multipliers to this optimal point.

Question 5

a) The admissible set, and the level surfaces of the function to be minimized are as in the figure. There are two constraints active at the point $(0, 1)$ and their gradients, at this point, are independent. At any other admissible point there is only one active constraint, the equality constraint, and its gradient is always nonzero (it is a constant vector). Thus all points are regular points for the constraints.



b) The optimal solution is obtained considering the smallest circle centered at the origin intersecting the admissible set. Hence, the optimal solution is the point $(0, 1)$.

c) The stationary points of the mixed penalty-barrier function are the solutions of

$$0 = \nabla F_\epsilon = \begin{bmatrix} 2x_1 - \frac{2}{\epsilon}(x_2 - x_1 - 1) - \frac{\epsilon}{x_1^2} \\ 2x_2 + \frac{2}{\epsilon}(x_2 - x_1 - 1) \end{bmatrix}.$$

Solving the second equation yields

$$x_2 = \frac{x_1 + 1}{\epsilon + 1},$$

and replacing this in the first equation yields

$$0 = \frac{x_1^3(2\epsilon + 4) + 2x_1^2 - \epsilon(1 + \epsilon)}{(\epsilon + 1)x_1^2}.$$

Setting $x_1 = \alpha\sqrt{\epsilon}$ and neglecting all terms ϵ^k, with $k \geq 1/2$, yields $0 = (2\alpha^2 - 1)$, hence (recall that $\alpha > 0$) $x_1 = \sqrt{\epsilon/2}$, and $x_2 = \frac{\sqrt{\epsilon/2}+1}{\epsilon+1}$.

d) As $\epsilon \to 0$, the stationary point of the mixed penalty-barrier function tends to $(0, 1)$, which is the optimal solution of the considered problem.

Question 6

a) Define the Lagrangian

$$L(x_1, x_2, \lambda) = x_1 x_2 + \lambda(\frac{1}{2}x_1^2 + 2x_2^2 - 1).$$

The first order necessary conditions of optimality are

$$0 = \frac{dL}{dx_1} = x_2 + \lambda x_1 \qquad 0 = \frac{dL}{dx_2} = x_1 + 4\lambda x_2 \qquad \frac{1}{2}x_1^2 + 2x_2^2 - 1 = 0.$$

b) The conditions $\frac{dL}{dx_1} = \frac{dL}{dx_2} = 0$ can be rewritten as

$$\begin{bmatrix} \lambda & 1 \\ 1 & 4\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

If $4\lambda^2 - 1 \neq 0$ the above equation implies $x_1 = x_2 = 0$, which is not an admissible point. If $4\lambda^2 - 1 = 0$, or $\lambda = \pm\frac{1}{2}$. then $x_2 = \mp\frac{1}{2}x_1$, and replacing in the constrains yields the candidate solutions with the corresponding multipliers

$$(x_1, x_2, \lambda) = \left(1, -\frac{1}{2}, \frac{1}{2}\right) \qquad (x_1, x_2, \lambda) = \left(-1, \frac{1}{2}, \frac{1}{2}\right)$$

$$(x_1, x_2, \lambda) = \left(1, \frac{1}{2}, -\frac{1}{2}\right) \qquad (x_1, x_2, \lambda) = \left(-1, -\frac{1}{2}, -\frac{1}{2}\right).$$

c) The exact augmented Lagrangian function for a constraint optimization problem with equality constraints is

$$S(x, \lambda) = f(x) + \lambda' g(x) + \frac{1}{\epsilon}\|g(x)\|^2 + \eta\|\frac{\partial g(x)}{\partial x}\nabla_x L(x, \lambda)\|^2,$$

with $\epsilon > 0$ and $\eta > 0$. Hence, for the considered problem, we have

$$S(x_1, x_2, \lambda) = x_1 x_2 + \lambda(\frac{1}{2}x_1^2 + 2x_2^2 - 1) + \frac{1}{\epsilon}(\frac{1}{2}x_1^2 + 2x_2^2 - 1)^2 + \eta\left(\begin{bmatrix} x_1 & 4x_2 \end{bmatrix}\begin{bmatrix} x_2 + \lambda x_1 \\ x_1 + 4\lambda x_2 \end{bmatrix}\right)^2.$$

d) The stationary points of the function $S(x_1, x_2, \lambda)$ are the solutions of the equations

$$0 = \frac{dS}{dx_1} = x_2 + \lambda x_1 + \frac{2x_1}{\epsilon}(\frac{1}{2}x_1^2 + 2x_2^2 - 1) + 2\eta(5x_1 x_2 + \lambda x_1^2 + 16\lambda x_2^2)(5x_2 + 2\lambda x_1)$$

$$0 = \frac{dS}{dx_2} = x_1 + 4\lambda x_2 + \frac{8x_2}{\epsilon}(\frac{1}{2}x_1^2 + 2x_2^2 - 1) + 2\eta(5x_1 x_2 + \lambda x_1^2 + 16\lambda x_2^2)(5x_1 + 32\lambda x_2)$$

$$0 = \frac{dS}{d\lambda} = \frac{1}{2}x_1^2 + 2x_2^2 - 1 + 2\eta(5x_1 x_2 + \lambda x_1^2 + 16\lambda x_2^2)(x_1^2 + 16x_2^2).$$

Replacing the candidate points obtained in part b) shows that indeed they are stationary points for the augmented Lagrangian function. (Note that this is true for any ϵ and η.)

e) To find the global minimum we evaluate the function to be minimized at the candidate optimal solutions:

$$(x_1 x_2)_{x_1=1, x_2=-1/2} = -\frac{1}{2} \qquad (x_1 x_2)_{x_1=-1, x_2=1/2} = -\frac{1}{2}$$

$$(x_1 x_2)_{x_1=1, x_2=1/2} = \frac{1}{2} \qquad (x_1 x_2)_{x_1=-1, x_2=-1/2} = \frac{1}{2}.$$

Hence, the points $(1, -1/2)$ and $(-1, 1/2)$ are both global minimizers. (Note that the points $(1, 1/2)$ and $(-1, -1/2)$ are both global maximizers.)