~~Exam~~ Master Copy (Aug. 04)

E4.29
C1.1

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2004

MSc and EEE PART IV: MEng and ACGI

OPTIMIZATION

Wednesday, 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) :	J.C. Allwright

[4.29]

Special instructions for invigilators: None

Information for candidates:

- All functions are sufficiently smooth.
- ∇f denotes the gradient of the function f. Note that ∇f is a column vector.
- $\nabla^2 f$ denotes the Hessian matrix of the function f. Note that this is a symmetric square matrix.

[4.29]

1. Consider the problem of minimizing the function

$$f(x) = x_1^2 + x_2^2 - x_1$$

(a) Compute the unique stationary point of the function, and show that the function is radially unbounded. [4]

(b) Using second order sufficient conditions show that the stationary point determined in part (a) is a local minimum. Also show that the point is a global minimum. [6]

(c) Consider the minimization of the function f using the gradient algorithm. Express analytically the form of the generic iteration, *i.e.*

$$p_{k+1} = p_k - \alpha \nabla f \qquad (\star)$$

(where $p_i = [x_1^i, x_2^i]^T$). [2]

(d) Consider the initial point $p_0 = [1, 1]$ and apply one step of the gradient algorithm $(\star)$ with exact line search. Verify that p_1 coincides with the stationary point determined in part (a). [4]

(e) It is known that for quadratic functions, such as the function f above, the gradient algorithm is globally convergent, however the speed of convergence may be very slow. Discuss why, for the function f, the gradient algorithm with exact line search converges in one step. [4]

2. Consider the problem of minimizing the function

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

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

(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. [6]

(c) Consider the minimization of the function f using Newton's algorithm. Express analytically the form of the generic iteration, *i.e.*

$$p_{k+1} = p_k - [\nabla^2 f]^{-1} \nabla f \qquad (\star\star)$$

(where $p_i = [x_1^i, x_2^i]^T$). [2]

(d) Equation $(\star\star)$ defines a nonlinear discrete time system with equilibria coinciding with the stationary points of the function f.

Consider the linear approximation of system $(\star\star)$ around the equilibrium corresponding to the local minimum of the function f with $x_1 > 0$, and compute the eigenvalues associated with the linear approximation.

Interpret the result obtained in terms of convergence properties of sequences generated by Newton's algorithm and initialized close to a local minimum. [4]

(e) Consider the initial point

$$p_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and, using the results in part (c), apply four steps of Newton's algorithm to generate the points p_1, p_2, p_3, p_4. Comment on the speed of convergence of the sequence. [4]

[4.29]

3. Consider the minimization problem

$$\begin{cases} \min_x 1 - x_1^2 - x_2^2 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x_1 + x_2 - 1 \leq 0. \end{cases}$$

(a) Show that all points in the admissible set are regular points for the constraints. [2]

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

(c) Using the conditions derived in part (b), compute candidate optimal solutions. [8]

(d) Show that the admissible set is compact.

Hence deduce the existence of a global minimum for the optimization problem.

Determine the global minimum of the problem.

Is this minimum unique? [6]

[4.29]

4. Consider the optimization problem

$$\begin{cases} \min_{x_1,x_2} -x_1 - x_2 \\ x_1^2 + x_2^2 - 1 = 0 \end{cases}$$

(a) Transform this minimization problem into an unconstrained minimization problem using the method of sequential penalty functions. [4]

(b) State the necessary conditions of optimality for the unconstrained problem of part (a). Hence compute approximate candidate optimal solutions for the unconstrained optimization problem. Discuss the feasibility of those candidate optimal solutions.
(Hint: you may show that optimal points of the unconstrained problem are such that $x_1^\star = x_2^\star$. Moreover, use the fact that the solutions of $1 + 4x\frac{1-2x^2}{\epsilon} = 0$, for ϵ positive and small, are

$$\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \qquad -\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \qquad -\frac{1}{4}\epsilon.)$$

[10]

(c) Consider the stationary points of the sequential penalty function of part (b). Consider the limit for $\epsilon \to 0$ of these stationary points and thus determine candidate optimal solutions for the original constrained optimization problem. [6]

[4.29]

5. Consider the discrete time system

$$x_{k+1} = a x_k$$

with $x_k \in I\!R$, and output $y_k = x_k$. Consider also the auxiliary discrete time system

$$\xi_{k+1} = \alpha \xi_k$$

with $\xi_k \in I\!R$, output $\eta_k = \xi_k$, and such that $\xi_0 = x_0 \neq 0$.

Consider now the problem of determining the constant α such that the cost

$$J(\alpha) = \frac{1}{2}\left(e_1^2 + e_2^2 + \cdots e_N^2\right)$$

is minimized, where $e_i = y_i - \eta_i$ and $N \geq 1$. (This can be regarded as a classical parameter identification problem.)

(a) Pose the above problem as an unconstrained optimization problem in the decision variable α, parameterized by a and x_0. [4]

(b) Assume $N = 1$. Show that $J(a) = 0$ and $J(\alpha) > 0$ for all $\alpha \neq a$. Hence show that the function $J(\alpha)$ has a unique local minimum which is also a global minimum. [4]

(c) Suppose $N = 2$. Compute the stationary points of $J(\alpha)$ as a function of a. Note that the number of stationary points is a function of the value of a. Hence, determine the local minima and the local maxima of the function $J(\alpha)$. [8]

(d) For $N = 2$ and $a = 3/2$, the function $J(\alpha)$ is as shown in Figure 5.1. Let $L = 12$ be the Lipschitz constant of $J(\alpha)$ for $\alpha \in [-2, 2]$. Apply four steps of the Schubert-Mladineo algorithm for the minimization of the function $J(\alpha)$ assuming that a global minimum is in the set $I_1 = \{\alpha \in I\!R \mid -2 \leq \alpha \leq 2\}$ and that the starting point of the algorithm is selected to be $\alpha = 2$. [4]

[4.29]

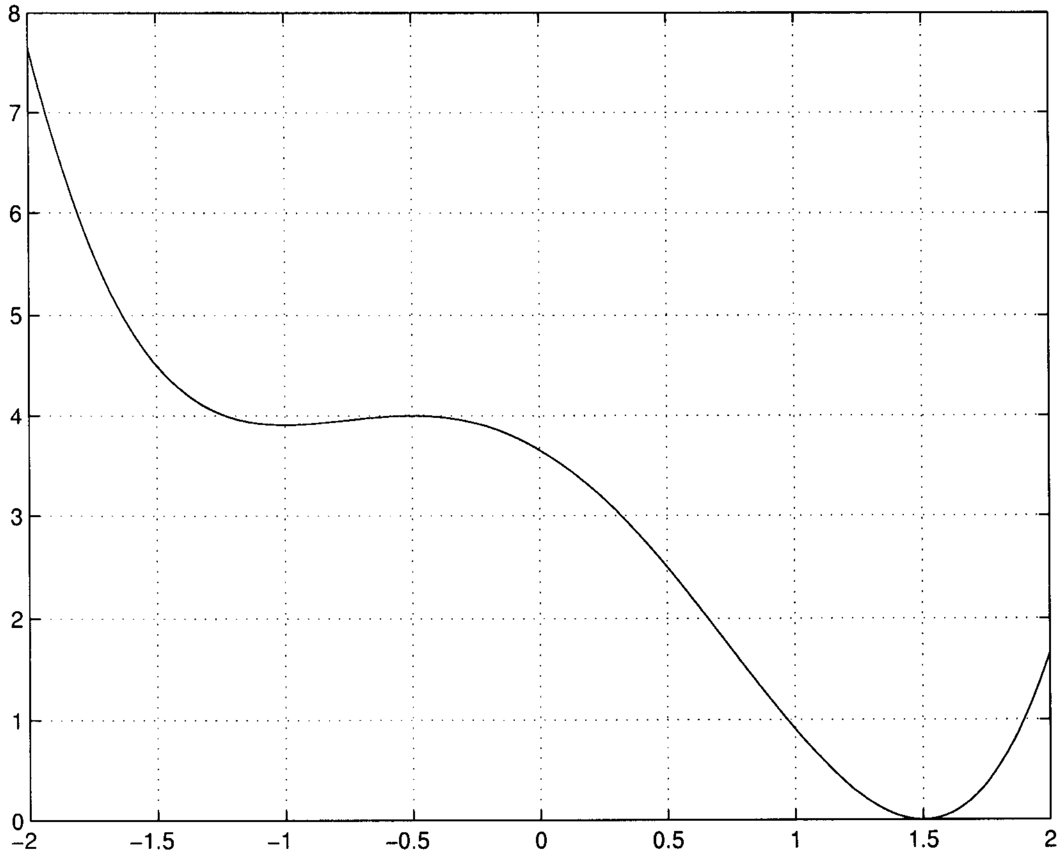

Figure 5.1: The function $J(\alpha)$.

[4.29]

6. Consider the optimization problem

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

(a) Sketch in the (x_1, x_2)-plane the level surfaces of the function to be minimized and the admissible set. (Hint: plot the level surfaces corresponding to $x_1^2 - x_2^2 = 0$ and $x_1^2 - x_2^2 = \pm 4$.) [4]

(b) Using first order necessary conditions, compute candidate optimal solutions. Use second order sufficient conditions to decide which of the candidate points is a local minimum or a local maximum. [8]

(c) Compute an exact penalty function for the minimization problem and verify that the candidate optimal solutions determined in part (b) are stationary points of the exact penalty function. [8]

Optimisation - Model answers 2004

Question 1

(a) The stationary point of the function f are computed solving the equation

$$0 = \nabla f = \begin{bmatrix} 2x_1 - 1 \\ 2x_2 \end{bmatrix},$$

yielding $x_1^\star = 1/2$ and $x_2^\star = 0$. The function f is of the form $x'Qx + c'x$ with $Q = \text{diag}(1,1) > 0$, hence it is radially unbounded.

(b) Note that $\nabla^2 f = \text{diag}(2,2) > 0$, hence $x^\star$ is a local minimum. It is also a global minimum for the following reason: the function f is C^1 and radially unbounded, therefore the global minimum is a stationary point.

(c) The generic iteration of the gradient algorithm for the considered function f is

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

(d) Let $x_1^0 = x_2^0 = 1$. Hence

$$x_1^1 = 1 - \alpha \qquad x_2^1 = 1 - 2\alpha.$$

Note now that $f(x_1^1, x_2^1) = 1 - 5\alpha + 5\alpha^2$ and this is minimized by $\alpha^\star = 1/2$, yielding

$$x_1^1 = 1 - \alpha^\star = 1/2 = x_1^\star \qquad x_2^1 = 1 - 2\alpha^\star = 0 = x_2^\star.$$

(e) For the considered function the gradient algorithm with exact line search converges in one step (from any initial point) because the function is quadratic and the minimum and maximum eigenvalues of $\nabla^2 f$ coincide.

Question 2

(a) The stationary point of the function f are computed solving the equation

$$0 = \nabla f = \begin{bmatrix} 4x_1^3 + x_2 \\ x_1 + x_2 \end{bmatrix},$$

yielding the stationary points

$$p^\star = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \tilde{p}^\star = \begin{bmatrix} -1/2 \\ 1/2 \end{bmatrix} \qquad \hat{p}^\star = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}$$

.

(b) Note that

$$\nabla^2 f(p^\star) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \not\geq 0 \qquad \nabla^2 f(\tilde{p}^\star) = \nabla^2 f(\hat{p}^\star) = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} > 0.$$

Hence, $p^\star$ is a saddle point, and $\tilde{p}^\star$ and $\hat{p}^\star$ are local minima.

(c) The generic iteration of Newton's algorithm for the considered function f is

$$x_1^{k+1} = \frac{8(x_1^k)^3}{12(x_1^k)^2 - 1} \qquad x_2^{k+1} = -\frac{8(x_1^k)^3}{12(x_1^k)^2 - 1}.$$

(d) The linear approximation of the above nonlinear discrete time system around the point $\hat{p}^\star$ is the system

$$p_{k+1} = Ap_k = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} p_k.$$

This linear system is such that $p_1 = 0$ for any p_0, and this explains the local 'fast' speed of convergence of Newton's iteration.

(e) A simple computation yields the sequence

$$p_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad p_1 = \begin{bmatrix} 0.7272727273 \\ -0.7272727273 \end{bmatrix} \qquad p_2 = \begin{bmatrix} 0.57552339460 \\ -0.57552339460 \end{bmatrix}$$

$$p_3 = \begin{bmatrix} 0.51266296461 \\ -0.51266296461 \end{bmatrix} \qquad p_4 = \begin{bmatrix} 0.5004542259 \\ -0.5004542259 \end{bmatrix}$$

and this shows the fast convergence (approximately two exact digits for each iteration) of Newton's algorithm.

Question 3

(a) The admissible set is the shaded area in Figure 1. The arrows denote the gradient of the constraints on the boundary of the admissible set. As can be seen, these vectors are always independent, therefore all points are regular points for the constraints.

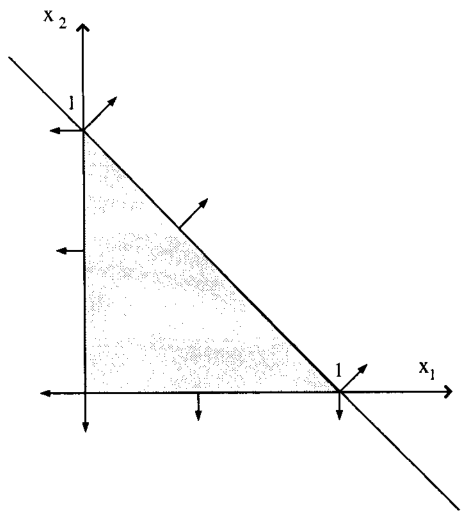

Figure 1: Admissible set

(b) Define the Lagrangian

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

The first order necessary conditions of optimality are

$$\begin{array}{l} -2x_1 - \rho_1 + \rho_3 = 0 \\ -2x_2 - \rho_2 + \rho_3 = 0 \\ -x_1 \leq 0 \quad -x_2 \leq 0 \quad x_1 + x_2 - 1 \leq 0 \\ \rho_1 \geq 0 \quad \rho_2 \geq 0 \quad \rho_3 \geq 0 \\ -\rho_1 x_1 = 0 \quad -\rho_2 x_2 = 0 \quad \rho_3(x_1 + x_2 - 1) = 0. \end{array}$$

(c) To compute candidate optimal solutions, note that from the last line of the necessary conditions we have the following possibilities:

P1 $\rho_1 = 0,\ \rho_2 = 0,\ \rho_3 = 0$;

P2 $\rho_1 = 0,\ \rho_2 = 0,\ \rho_3 > 0$;

P3 $\rho_1 = 0,\ \rho_2 > 0,\ \rho_3 = 0$;

P4 $\rho_1 = 0,\ \rho_2 > 0,\ \rho_3 > 0$;

P5 $\rho_1 > 0,\ \rho_2 = 0,\ \rho_3 = 0$;

P6 $\rho_1 > 0,\ \rho_2 = 0,\ \rho_3 > 0$;

P7 $\rho_1 > 0,\ \rho_2 > 0,\ \rho_3 = 0$;

P8 $\rho_1 > 0,\ \rho_2 > 0,\ \rho_3 > 0$.

This yields the following candidate points

P1 $x_1 = x_2 = 0$ hence $f = 1$;

P2 $x_1 = x_2 = 1/2$ hence $f = 1/2$;

P3 impossible;

P4 $x_1 = 1$, $x_2 = 0$, hence $f = 0$;

P5 impossible;

P6 $x_1 = 0$, $x_2 = 1$, hence $f = 0$;

P7 impossible;

P8 impossible.

As a result, we have only four candidate points:

$$P_1 = (0,0) \quad P_2 = (1/2, 1/2) \quad P_4 = (1,0) \quad P_6 = (0,1).$$

(d) The admissible set is closed (the constraints include the equality sign) and bounded (see Figure 1), hence compact. By Weierstrass theorem the function f has a global minimum in such a set. The function f attains its global minimum at P_4 and P_6, which are therefore both global minima. (This can be also shown noting that the problem is symmetric, i.e. changing x_1 into x_2 and x_2 into x_1 yields the same problem.)

Question 4

(a) A sequential penalty function for the constrained problem is

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

(b) The necessary conditions of optimality for F_ϵ are

$$0 = \nabla F_\epsilon = \begin{bmatrix} -1 + \frac{4x_1}{\epsilon}(x_1^2 + x_2^2 - 1) \\ -1 + \frac{4x_2}{\epsilon}(x_1^2 + x_2^2 - 1) \end{bmatrix}.$$

As a result,

$$\frac{1}{x_1} = \frac{4}{\epsilon}(x_1^2 + x_2^2 - 1)$$
$$\frac{1}{x_2} = \frac{4}{\epsilon}(x_1^2 + x_2^2 - 1)$$

yielding $x_1 = x_2$. Let $x_1 = x_2 = x$. From the first equation we have

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

This equation has approximate solutions

$$\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \qquad -\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \qquad -\frac{1}{4}\epsilon.$$

As a result, F_ϵ has three stationary points:

$$P_1 \approx (\frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon, \frac{\sqrt{2}}{2} + \frac{1}{8}\epsilon) \qquad P_2 \approx (\frac{-\sqrt{2}}{2} + \frac{1}{8}\epsilon, \frac{-\sqrt{2}}{2} + \frac{1}{8}\epsilon) \qquad P_3 \approx (-\frac{1}{4}\epsilon, -\frac{1}{4}\epsilon).$$

Note that none of the above points is feasible, for any $\epsilon > 0$.

(c) The stationary points of F_ϵ are such that

$$\lim_{\epsilon \to 0} P_1 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \qquad \lim_{\epsilon \to 0} P_2 = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}) \qquad \lim_{\epsilon \to 0} P_3 = (0, 0).$$

Hence, P_1 and P_2 converge to the admissible set. P_1 is a (local) solution of the optimization problem considered.

Question 5

(a) The problem can be formulated as

$$\min_{\alpha \in I\!R} J(\alpha) = \min_{\alpha \in I\!R} \frac{1}{2}(a-\alpha)^2 x_0^2 + (a^2-\alpha^2)^2 x_0^2 + \cdots + (a^N - \alpha^N)^2 x_0^2,$$

i.e. as an unconstrained optimization problem in the decision variable α and parameterized by a and x_0.

(b) For $N = 1$ one has

$$J(\alpha) = \frac{1}{2}(a-\alpha)^2 x_0^2.$$

Hence, $J(a) = 0$ and $J(\alpha) > 0$ for all $\alpha \neq a$. This shows (using the very definition of global minimum) that $\alpha = a$ is a global minimum.

(c) If $N = 2$ one has

$$J(\alpha) = \frac{1}{2} x_0^2 \left((a-\alpha)^2 + (a^2-\alpha^2)^2 \right).$$

Hence,

$$\frac{dJ(\alpha)}{d\alpha} = -x_0^2 (a-\alpha)(2\alpha^2 + 2\alpha a + 1).$$

Therefore, the stationary points are

$$P_1 = a \qquad P_2 = -\frac{a}{2} + \frac{\sqrt{a^2-2}}{2} \qquad P_3 = -\frac{a}{2} - \frac{\sqrt{a^2-2}}{2}.$$

We conclude that, if $|a| < \sqrt{2}$ there is only one stationary point, whereas if $|a| \geq \sqrt{2}$ there are three stationary points. Computing second derivatives we have that P_1 is always a local minimum, and, for $|a| \geq \sqrt{2}$, P_2 is a local maximum and P_3 is a local minimum.

(d) A sketch of the application of the Schubert-Mladineo algorithm is shown in Figure 2.

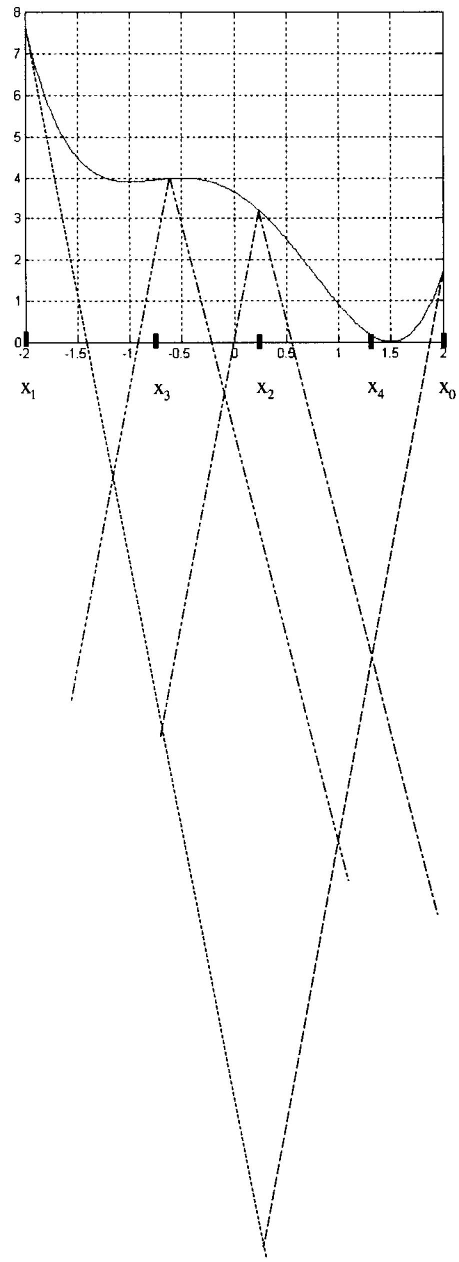

Figure 2: Application of the Schubert-Mladineo algorithm

Question 6

(a) The level sets and the admissible set are depicted in Figure 3.

(b) Let

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

The first order necessary conditions are

$$\begin{aligned} 2x_1 + \lambda &= 0 \\ -2x_2 - 2\lambda x_2 &= 0 \\ x_1 - x_2^2 &= 0 \end{aligned}$$

and these yield the candidate optimal points

$$P_1 = (0,0) \qquad P_2 = (1/2, \sqrt{2}/2) \qquad P_3 = (1/2, -\sqrt{2}/2),$$

with corresponding multipliers

$$\lambda_1 = 0 \qquad \lambda_2 = -1 \qquad \lambda_3 = -1.$$

The second order sufficient conditions are

$$s' \nabla^2_{xx} L(x^\star, \lambda^\star) s > 0$$

for $s \neq 0$ such that

$$[1, -2x_2^\star] s = 0.$$

For P_1 one has $s = [0,1]'$ and $s'(\nabla^2_{xx} L)s < 0$, hence P_1 is a local maximum. For P_2 one has $s = [\sqrt{2}, 1]'$ and $s'(\nabla^2_{xx} L)s = 4$, and for P_3 one has $s = [\sqrt{2}, -1]'$ and $s'(\nabla^2_{xx})Ls = 4$. Hence, P_2 and P_3 are local minima.

(c) An exact penalty function for the considered problem is

$$L_a(x_1,x_2) = x_1^2 - x_2^2 - \frac{(2x_1 + 4x_2^2)(x_1 - x_2^2)}{1 + 4x_2^2} + \frac{(x_1 - x_2^2)^2}{\epsilon},$$

with $\epsilon > 0$. Its stationary points are the solutions of

$$0 = \nabla L_a(x_1,x_2) = \begin{bmatrix} 2\dfrac{-x_1\epsilon + 4x_1\epsilon x_2^2 - \epsilon x_2^2 + x_1 + 4x_1 x_2^2 - x_2^2 - 4x_2^4}{(1+4x_2^2)\epsilon} \\ 2x_2 \dfrac{-\epsilon + 8\epsilon x_1^2 - 2x_1\epsilon - 2x_1 - 16x_1x_2^2 - 32x_1x_2^4 + 2x_2^2 + 16x_2^4 + 32x_2^6}{(1+4x_2^2)^2\epsilon} \end{bmatrix}.$$

By direct substitution we verify that, for any $\epsilon > 0$, the points P_1, P_2 and P_3 are stationary points of L_a.

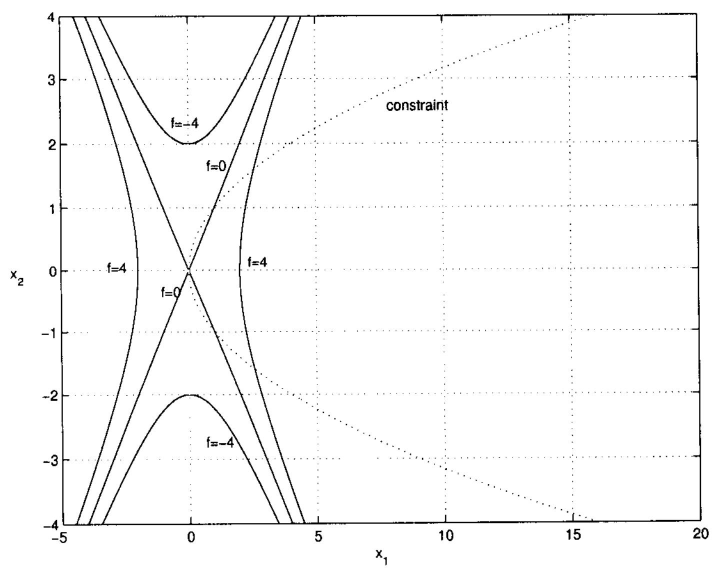

Figure 3: Level sets and constraint for Question 6.