OPTIMISATION

1\. a) The minimization problem can be re-cast as the problem of finding the stationary points of the function q, that is as the problem of solving the scalar equation

$$k(x) = x^3 + 4x^2 - 10 = 0.$$

As noted in the exam paper, this equation has a solution at $x = 1.365230013$, which is actually the only solution. Note also that the second derivative of f at $x = 1.365230013$ is positive, hence the point is a local minimizer.

Typical mistakes include re-casting the problem as the problem of finding the roots of the function q or of the function $q = q^\star$.

[2 marks]

b) Newton's iteration for the solution of the equation $k(x) = 0$ is

$$x_{k+1} = x_k - \frac{x_k^3 + 4x_k^2 - 10}{x_k(3x_k + 8)} = \frac{2x_k^3 + 4x_k^2 + 10}{x_k(3x_k + 8)}.$$

Typical errors: incorrect expression of Newton's iteration, mistakes in simplifying the expression, use of incorrect functions.

[2 marks]

c) Setting $x_0 = 3$ yields

$x_1 = 1.960784314$	$RE_1 = 0.1363174326$
$x_2 = 1.486238507$	$RE_2 = 0.5728643018$
$x_3 = 1.371823522$	$RE_3 = 3.721080242$
$x_4 = 1.365251224$	$RE_4 = 73.99652652$

Since the sequence of the relative errors diverges the speed of convergence of the method is not of order three (note the *cube* in the denominator of the definition of the relative error). This is not un-expected, since under the given conditions one can only claim quadratic speed of convergence of Newton's method.

Numerical mistakes leading to non-sensical values of the relative errors.

[6 marks]

d) The Levenberg-Marquardt iteration is given by the two equations

$$x_{k+1} = x_k - 2\frac{x_k^3 + 4x_k^2 - 10}{(x_k(3x_k + 8)) + (\bar{x}(3\bar{x} + 8))}, \qquad \bar{x} = x_k - \frac{x_k^3 + 4x_k^2 - 10}{x_k(3x_k + 8)}.$$

Typical mistakes: use of incorrect functions, errors in simplifications.

[2 marks]

e) Setting $x_0 = 3$ yields

$$\begin{aligned}
x_1 &= 1.644853060 & RE_1 &= 0.06400339279,\\
x_2 &= 1.369582249 & RE_2 &= 0.1990643754,\\
x_3 &= 1.365230035 & RE_3 &= 0.2668611603,\\
x_4 &= 1.365230013 & RE_4 &\approx 0.
\end{aligned}$$

Since the sequence of the relative errors converges to zero, the speed of convergence of the Levenberg-Marquardt iteration is at least of order three, definitely faster than Newton's iteration.

Numerical mistakes leading to non-sensical values of the relative errors.

[8 marks]

2. a) The Lagrangian for the problem is

$$L(x_1,x_2,\rho) = x_1^2 + x_2^2 - 3 + \rho(x_1 - 3x_2 + 2).$$

The necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 2x_1 + \rho, \qquad 0 = \frac{\partial L}{\partial x_2} = 2x_2 - 3\rho,$$

$$x_1 - 3x_2 + 2 \le 0, \qquad \rho \ge 0, \qquad \rho(x_1 - 3x_2 + 2) = 0.$$

Using the last condition, that is the complementarity condition, we have two cases.

i) $\rho = 0$: does not gives any candidate optimal solution;

ii) $\rho > 0$: gives the candidate optimal solution $x_1 = -1/5$ and $x_2 = 3/5$, with optimal multiplier $\rho = 2/5$.

Incorrect necessary conditions, in particular the complementarity condition. Incorrect use of the complementarity condition to determine candidate solutions. Too many/few candidate solutions.

[4 marks]

b) Note that

$$\nabla^2 L = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} > 0,$$

hence the second order sufficient condition of optimality is trivially satisfied at the candidate optimal solution, which is therefore a strict minimizer. (Note that the minimizer is a global one.)

Incorrect use of the second order sufficient conditions of optimality. Several students failed to observe that, since $\nabla^2 L$ is strictly positive, there is no need to evaluate the gradient of the active constraint.

[4 marks]

c) The Lagrangian for the problem with the slack variable is

$$L_s(x_1,x_2,s,\lambda,\rho) = x_1^2 + x_2^2 - 3 + \lambda(x_1 - 3x_2 + 2 + s) + \rho(-s).$$

This is a problem with three decision variables x_1, x_2 and s and two multipliers, λ and ρ. The necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 2x_1 + \lambda, \qquad 0 = \frac{\partial L}{\partial x_2} = 2x_2 - 3\lambda, \qquad 0 = \frac{\partial L}{\partial s} = \lambda - \rho,$$

$$x_1 - 3x_2 + 2 + s = 0, \qquad -s \le 0, \qquad \rho \ge 0, \qquad \rho s = 0.$$

Incorrect Lagrangian function (which is defined in an extended space) and incorrect set of necesary conditions of optimality.

[2 marks]

d) The Lagrangian for the log penalty function problem is

$$L_{log}(x_1,x_2,s,\lambda) = x_1^2 + x_2^2 - 3 - \varepsilon \log s + \lambda(x_1 - 3x_2 + 2 + s).$$

This is a problem with three decision variables, x_1, x_2 and s, and one multiplier, λ.

i) The necessary conditions of optimality for this problem are

$$0 = \frac{\partial L}{\partial x_1} = 2x_1 + \lambda, \qquad 0 = \frac{\partial L}{\partial x_2} = 2x_2 - 3\lambda,$$

$$0 = \frac{\partial L}{\partial s} = -\frac{\varepsilon}{s} + \lambda, \qquad x_1 - 3x_2 + 2 + s = 0.$$

Incorrect Lagrangian function (which is defined in a new extended space) and incorrect set of necesary conditions of optimality.

[2 marks]

ii) The condition $0 = -\frac{\varepsilon}{s} + \lambda$ can be rewritten as $s\lambda = \varepsilon$ which is clearly an approximation of the complementarity condition $\rho s = 0$ determined in part c).

Most students failed to indentify the relation between the complementarity condition in part c) and the condition $\frac{\partial L}{\partial s}$ *in part d.i), hence made incorrect claims.*

[4 marks]

iii) Solving the necessary conditions of optimality yields $x_1 = -1/5$, $x_2 = 3/5$, $s = 5/2\ \varepsilon$, $\lambda = 2/5$. As ε goes to zero we recover the optimal solution of the given problem.

Incorrect solution of the four linear equations expressing the necessary conditions of optimality and/or incorrect evaluation of the limits for $\varepsilon \to 0$.

[4 marks]

3. a) The Lagrangian for the problem is

$$L(x,\rho) = x^2 + \rho(1-x).$$

The necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x} = 2x - \rho, \qquad 1 - x \leq 0 \qquad \rho \geq 0, \quad \rho(1-x) = 0.$$

Incorrect necessary conditions of optimality (in particular incorrect directions of the inequalities).

[2 marks]

b) Using the complementarity conditions gives two cases: $\rho = 0$ and $x = 1$. The former does not give a feasible solution and the latter gives the solution $x = 1$ and $\rho = 2$, which is the only candidate optimal solution. The optimal cost is $f^\star = 1$.

Incorrect use of the necessary conditions to find candidate optimal points and the optimal solution of the problem.

[2 marks]

c) The Lagrange relaxation of the problem is given by

$$\min_x x^2 + \mu(1-x), \qquad \mu \geq 0.$$

Minimizing with respect to x yields $x^\star = \mu/2$, hence $q(\mu) = \mu - \frac{\mu^2}{4}$.

Incorrect minimization in the Lagrange relaxation problem and incorrect evaluation of the function q. Note that the minimization has to be performed, in this case, as a function of x and does not involve any constraint.

[6 marks]

d) The dual problem is

$$\max_\mu \mu - \frac{\mu^2}{4},$$
$$\mu \geq 0.$$

This problem can be solved considering the Lagrangian (one has to transform the problem into a minimization problem and adjust the inequality constraint)

$$L_d(\mu,\rho) = \frac{\mu^2}{4} - \mu + \rho(-\mu)$$

and writing the necessary conditions of optimality

$$0 = \frac{\partial L_d}{\partial \mu} = \mu/2 - 1 - \rho, \qquad -\mu \leq 0 \qquad \rho \geq 0, \quad \rho\mu = 0.$$

These give the only candidate solution $\mu = 2$ with $\rho = 0$. The optimal cost is $q^\star = q(2) = 1$, which coincides with $f^\star$, hence the duality gap is zero. Since the duality gap is zero, the optimal solution of the original problem solves the equality $f(x) = q^\star$.

Incorrect set of necessary conditions of optimality and/or incorrect evaluation of the optimal solution. Some students also failed to recongnize that the duality gap is zero and were not able to use this fact as requested.

[6 marks]

e) The sequential penalty function for the considered problem is

$$F_\varepsilon(x) = x^2 + \frac{1}{\varepsilon}\max(0, 1-x)^2.$$

This function has a unique stationary point which is given by

$$x_\varepsilon = \frac{1}{1+\varepsilon}.$$

Note that this is a global minimizer, it is not feasible for any $\varepsilon > 0$, and it is such that

$$\lim_{\varepsilon\to 0} x_\varepsilon = 1,$$

that is the global minimizer of the sequential penalty function converges, as ε tends to zero, to the solution of the considered optimization problem giving the optimal cost

$$\lim_{\varepsilon\to 0} F_\varepsilon\left(\frac{1}{1+\varepsilon}\right) = f^\star = 1.$$

Some students wrote an incorrect sequential penalty function and/or performed its unconstrained minimization incorrectly. Some also made mistakes in the evaluation of the limits involved in the solution of the problem.

[4 marks]

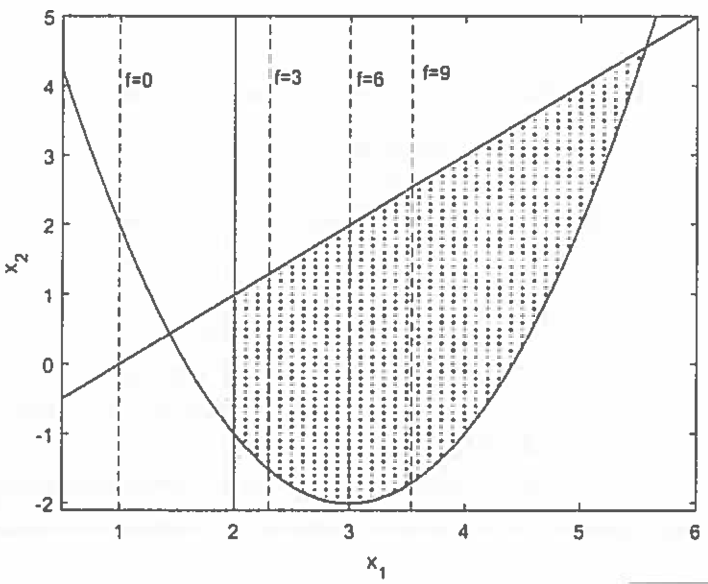

Figure 4.1 Admissible set (dotted set) and level lines of the objective function (dashed lines).

4. a) The admissible set and the level lines of the objective function are as in Figure 4.1. Note that all points are regular points for the constraints.

The incorrect admissible set was drawn and some students failed to study regularity of all admissible points.

[4 marks]

b) From Figure 4.1 we conclude that all points (x_1, x_2) such that $x_1 = 2$ and $x_2 \in [-1, 1]$ are optimal solutions for which the value of the objective function is $f^* = 2$.

Most students did not observe that the (global) minimizer is not unique and/or gave an incorrect set of optimal points (most notably, overlooked the fact that x_2 is constrained by feasibility.

[2 marks]

c) The Lagrangian of the problem is

$$L(x_1, x_2, \rho_1, \rho_2, \rho_3) = x_1^2 - x_1 + \rho_1(2 - x_1) + \rho_2((x_1 - 3)^2 - x_2 - 2) + \rho_3(1 - x_1 + x_2).$$

The necessary conditions of optimality are

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

$$2 - x_1 \leq 0, \qquad (x_1 - 3)^2 - x_2 - 2 \leq 0, \qquad 1 - x_1 + x_2 \leq 0,$$

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

$$\rho_1(2 - x_1) = 0, \qquad \rho_2((x_1 - 3)^2 - x_2 - 2) = 0, \qquad \rho_3(1 - x_1 + x_2) = 0.$$

Incorrect Lagrangian function and/or set of necessary conditions of optimality.

[4 marks]

d) Using the complementarity conditions one has eight cases.

i) All constraints are inactive: no candidate optimal point.

ii) $2-x_1=0$: all points such that $x_1=2$ and $x_2 \in (-1,1)$ are candidate optimal solutions. The strict complementarity condition holds, that is $\rho_1 = 3$.

iii) $(x_1-3)^2-x_2-2=0$: no candidate optimal point.

iv) $1-x_1+x_2=0$: no candidate optimal point.

v) $2-x_1=0$ and $(x_1-3)^2-x_2-2=0$: one candidate optimal point $(x_1,x_2)=(2,-1)$. For this point $\rho_2=0$, hence the strict complementarity condition does not hold.

vi) $2-x_1=0$ and $1-x_1+x_2=0$: one candidate optimal point $(x_1,x_2)=(2,1)$. For this point $\rho_3=0$, hence the strict complementarity condition does not hold.

vii) $(x_1-3)^2-x_2-2=0$ and $1-x_1+x_2=0$: no candidate optimal point.

viii) All constraints are active: no candidate optimal point.

Incorrect use of the complementarity condition to identify candidate optimal solutions. This is a somewhat tedious exercise which is much simplified by the availability of a correct answer to part a). Most students failed to connect the result of this part to the outcome of parts a) and b)!

[8 marks]

e) There are three types of candidate optimal points.

- $x_1=2$ and $x_2 \in (-1,1)$: none of these points is a strict local minimizer, hence the second order sufficient conditions of optimality do not hold (these are for strict minimizers). As a matter of fact, at these optimal points
$$\nabla^2 L = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix},$$
and
$$s'\nabla^2 L s = 0,$$
for all s such that $\dfrac{\partial(2-x_1)}{\partial x} s = 0.$
- $x_1=2$ and $x_2=1$. At this point the strict complementarity condition does not hold, hence the second order sufficient conditions of optimality cannot be used to classify the point.
- $x_1=2$ and $x_2=-1$. Similarly to the previous case, at this point the strict complementarity condition does not hold, hence the second order sufficient conditions of optimality cannot be used to classify the point.

Almost all students did not observe that, since the minimizer is non-strict, the sufficient conditions cannot be used. They also failed to observe that the strict complementarity condition does not always hold.

[4 marks]