OPTIMISATION

1. Note that

$$\begin{aligned} f(x) &= (x-a_1)^2+(x-a_2)^2+(x-a_3)^2+(x-a_4)^2 \\ &= x^2-2a_1+a_1^2+x^2-2a_2+a_2^2+x^2-2a_3+a_3^2+x^2-2a_4+a_4^2 \\ &= 4x^2-2(a_1+a_2+a_3+a_4)x+a_1^2+a_2^2+a_3^2+a_4^2. \end{aligned}$$

a) The first order necessary condition of optimality is

$$0 = \nabla f = 8x-2(a_1+a_2+a_3+a_4),$$

which yields

$$x^* = \frac{a_1+a_2+a_3+a_4}{4}.$$

x^* is the average of a_1, a_2, a_3, a_4. [2 marks]

Typical mistakes include the incorrect solution of the equation defining the stationary points.

b) The second order sufficient condition of optimality is

$$\nabla^2 f = 8 > 0$$

Hence, x^* is a local minimizer. Note that f is strictly convex and this implies that x^* is global minimizer. Alternatively, note that

$$\lim_{|x|\to+\infty} f(x) = +\infty.$$

Hence, f is radially unbounded. [4 marks]

Typical mistakes include the incorrect computation of the second order derivative, the "non-recognition" that the function is convex and/or radially unbounded.

c) Note that $f(x) = 4x^2+c^2$, with $c^2 = a_1^2+a_2^2+a_3^2+a_4^2$.

i) The gradient is

$$\nabla f = 8x$$

and the gradient algorithm gives

$$x_{k+1} = x_k - 8\alpha_k x_k = (1-8\alpha_k)x_k.$$

Note that

$$f(x_{k+1}) = 4(1-16\alpha_k+64\alpha_k^2)x_k^2+c^2$$

and

$$f(x_k) = 4x_k^2+c^2$$

yields

$$f(x_{k+1})-f(x_k) = 4(64\alpha_k^2-16\alpha_k)x_k^2.$$

To find the exact line search parameter solve

$$0 = \frac{\partial[f(x_{k+1})-f(x_k)]}{\partial\alpha} = 4(128\alpha_k-16)x_k^2,$$

Answers

obtaining

$$\alpha^* = \frac{1}{8}.$$

[4 marks]

Typical mistakes include the incorrect computation of the exact line search parameter or the incorrect definition of the parameter itself.

ii) Note now that

$$x_{k+1} = x_k - 8\gamma\alpha^* x_k = (1-\gamma)x_k.$$

To have convergence we need

$$|1-\gamma| < 1.$$

Hence, $\gamma \in (0,2)$. For $\gamma = 0$ or $\gamma = 2$,

$$|x_{k+1}| = |x_k|,$$

and the sequence does not converge. For $\gamma \in (2,3]$

$$|x_{k+1}| > |x_k|$$

hence the sequence diverges. For $\gamma \in (0,2)$ the speed of convergence is linear, since

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} = \frac{|x_{k+1}|}{|x_k|} = |1-\gamma|.$$

[10 marks]

Typical mistakes include the incorrect definition of the iteration and then the inability to study the iteration as a discrete-time system to assess convergence properties. Some have also overlooked the absolute value in the condition $|1-\gamma| < 1$ and/or have not sudied correctly the cases in which $|1-\gamma| = 1$.

Answers

2. The Lagrangian function is

$$L(x,\lambda,\rho) = (x_1-6)^2 + (x_2-7)^2 + \rho(x_1+x_2-7).$$

a) The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 2(x_1-6)+\rho, \tag{2.1}$$
$$0 = \frac{\partial L}{\partial x_2} = 2(x_2-7)+\rho, \tag{2.2}$$
$$\rho \geq 0, \tag{2.3}$$
$$x_1+x_2-7 \leq 0, \tag{2.4}$$
$$\rho(x_1+x_2-7) = 0. \tag{2.5}$$

[2 marks]

Typical mistakes include writing only some of the necessary conditions, not writing the complementarity condition or writing the inequalities with incorrect "direction".

b) If $\rho = 0$, equations (2.1) and (2.2) yield $x_1 = 6$, $x_2 = 7$, which is not a feasible solution because the constraint (2.4) is violated.
If $\rho > 0$ then equation (2.5) yields

$$x_1+x_2-7 = 0,$$

while equations (2.1) and (2.2)

$$x_1-6 = x_2-7.$$

The solution of these two equations is $x_1 = 3$ and $x_2 = 4$ which yields, by equation (2.1), $\rho = 6$. The point $(x_1,x_2) = (3,4)$ is the closest point in the admissible set to the center of the circle $(x_1-6)^2+(x_2-7)^2 = r^2$, hence it is the solution of the problem, *i.e.* it gives the smallest circle which intersects the admissible set, see Figure 2.1. [8 marks]

Typical mistakes include the incorrect solution of the equations identifying the candidate optimal solutions and the lack of geometric understanding of the problem.

c) Consider the penalty function

$$P_c(x_1,x_2) = (x_1-6)^2 + (x_2-7)^2 + c(\max\{0,x_1+x_2-7\})^2.$$

i) If $\max\{0,x_1+x_2-7\} = 0$ then

$$P_c(x_1,x_2) = (x_1-6)^2 + (x_2-7)^2,$$

with the stationary point $(6,7)$. However, this point violates the hypothesis since $\max\{0,x_1+x_2-7\} = 6$. If $\max\{0,x_1+x_2-7\} = x_1+x_2-7$ then

$$P_c(x_1,x_2) = (x_1-6)^2 + (x_2-7)^2 + c(x_1+x_2-7)^2$$

and the stationary point is

$$x_c^\star = \left(6\frac{1+c}{1+2c}, \frac{7+8c}{1+2c}\right).$$

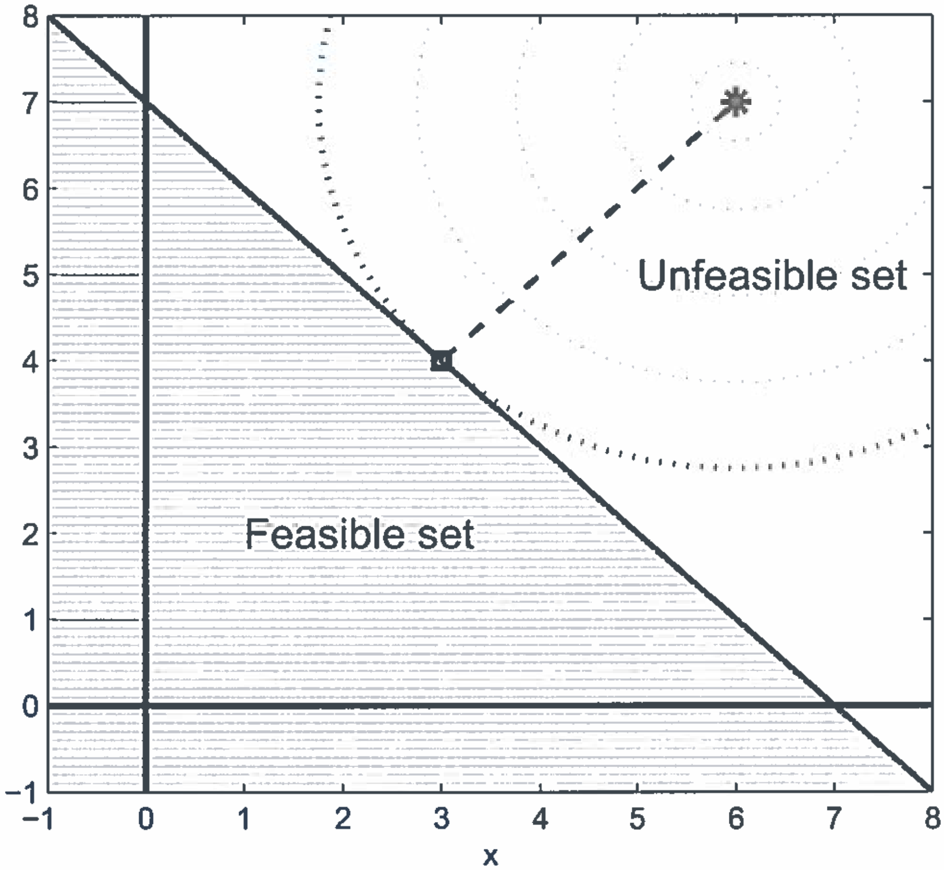

Figure 2.1 Feasible set, unfeasible set, unconstrained optimal (star) and constrained optimal (square).

[6 marks]

Some students have not recognized that the function has to be studied in two regions and that one of the solution has to be discarded since it is outside the set in which the function that has generated the solution is defined. In addition, they may have made mistakes in the computation of the stationary point.

ii) The function P_c is a sum of squares, hence it is strictly convex, radially unbounded and x_c^* is its global minimizer. [2 marks]

Some students have not recognized that the function is a sum of squares, it is stricly convex and radially unbounded.

iii) Note that

$$\lim_{c \to +\infty} x_c^* = (3,4),$$

which is the solution of the problem. [2 marks]

The typical mistake is the incorrect evaluation of the limit.

Answers

3. The Lagrangian function is

$$L(x,\lambda,\rho) = -xy^2 + \lambda(y - x + \varepsilon), \qquad \varepsilon > 0.$$

a) The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x} = -y^2 - \lambda, \tag{3.1}$$

$$0 = \frac{\partial L}{\partial y} = -2xy + \lambda, \tag{3.2}$$

$$y - x + \varepsilon = 0. \tag{3.3}$$

[2 marks]

Some students may have written the incorrect necessary conditions.

b) From equations (3.1) and (3.2) we obtain

$$-y^2 - 2xy = 0,$$

which has a solution if

$$y = 0,$$

or

$$y + 2x = 0.$$

i) If $y = 0$, equation (3.3) yields $x = \varepsilon$, equation (3.1) yields $\lambda = 0$ and the cost is $L = 0$.

ii) If $y + 2x = 0$, equation (3.3) yields $x = \frac{\varepsilon}{3}$, $y = -2\frac{\varepsilon}{3}$, equation (3.1) yields $\lambda = \frac{-4\varepsilon^2}{9}$ and the cost is $L = \frac{-4\varepsilon^3}{27}$.

[6 marks]

Some students may have solved incorrectly the equations definiting the candidate optimal solutions. In addition, some may have been misled by the fact that the problem does not have a global solution, although it clearly has a local solution.

c) The second order sufficient conditions of optimality are

$$s'\nabla^2 L s = s' \begin{bmatrix} 0 & -2y \\ -2y & -2x \end{bmatrix} s > 0 \tag{3.4}$$

and

$$0 = \nabla g s = \begin{bmatrix} -1 & 1 \end{bmatrix} s. \tag{3.5}$$

Condition (3.5) yields

$$s = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \alpha,$$

with $\alpha \neq 0$.

i) Consider

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

hence, condition (3.4) yields

$$s'\nabla^2 L s = -2\varepsilon\alpha^2 < 0,$$

which implies that the point $(\varepsilon, 0)$ is a local maximizer.

 ©Imperial College London

Answers

ii) Consider

$$\nabla^2 L = \begin{bmatrix} 0 & \frac{4\varepsilon}{3} \\ \frac{4\varepsilon}{3} & -2\frac{\varepsilon}{3} \end{bmatrix},$$

hence, condition (3.4) yields

$$s'\nabla^2 Ls = -\frac{10}{3}\varepsilon\alpha^2 < 0,$$

which implies that the point $\left(\frac{\varepsilon}{3}, \frac{-2\varepsilon}{3}\right)$ is a local minimizer.

[6 marks]

Typical mistakes include the incorrect computation of the Hessian or the erroneous use of the sufficient conditions.

d) The optimal cost is

$$J^*(\varepsilon) = \frac{-4\varepsilon^3}{27}$$

and

$$\frac{\partial J^*}{\partial \varepsilon} = -\frac{4}{9}\varepsilon^2 = \lambda^*(\varepsilon).$$

This demonstrates that the variation of the cost as a function of ε is equal to the optimal multiplier, *i.e.* since ε is a measure of constraint violation $\lambda^*(\varepsilon)$ gives a measure of the variation of the cost as a function of the "violation".

[6 marks]

This is essentially a straightforward question which requires the correct interpretation of the results of the exercise, something that some students have found hard.

Answers

4. a) When only the equality constraint is active

$$\frac{\partial g}{\partial x} = \begin{bmatrix} 1 & 1 \end{bmatrix} \neq 0,$$

which implies that all points are regular points. When both constraints are active

$$\begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial h}{\partial x} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}.$$

Since the column vectors are independent, all points are regular. [2 marks]

Some students have not applied the definition and have incorrectly interpreted the constrains.

b) Note that the Lagrangian function is

$$L(x,\lambda,\rho) = 2x_1^2 + 9x_2 + \lambda(x_1 + x_2 - 4) - \rho x_1.$$

The first order necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 4x_1 + \lambda - \rho, \tag{4.1}$$

$$0 = \frac{\partial L}{\partial x_2} = 9 + \lambda, \tag{4.2}$$

$$x_1 + x_2 - 4 = 0, \tag{4.3}$$

$$\rho \geq 0, \tag{4.4}$$

$$-x_1 \leq 0, \tag{4.5}$$

$$\rho x_1 = 0. \tag{4.6}$$

[2 marks]

Some students may have written the incorrect necessary conditions, or the inequalities in the incorrect "direction".

c) From equation (4.6), we have $\rho = 0$ or $x_1 = 0$.
If $\rho = 0$, equation (4.2) yields $\lambda = -9$, equation (4.1) yields $x_1 = \frac{9}{4}$ and equation (4.3) yields $x_2 = \frac{7}{4}$.
If $x_1 = 0$, equation (4.3) yields $x_2 = 4$, equation (4.2) yields $\lambda = -9$ and equation (4.1) yields $\rho = -9 \not\geq 0$.
The only candidate optimal solution is $x_1 = \frac{9}{4}$, $x_2 = \frac{7}{4}$, $\lambda = -9$ and $\rho = 0$.

[4 marks]

Some students may have computed the wrong candidate optimal solution, or may have found more than one candidate solution.

d) Consider

$$B_r(x_1,x_2) = 2x_1^2 + 9x_2 + r\log x_1 + \frac{1}{r}(x_1 + x_2 - 4)^2.$$

i) The first order optimality conditions are

$$0 = \frac{\partial B_r}{\partial x_1} = 4x_1 + \frac{r}{x_1} + \frac{2}{r}(x_1 + x_2 - 4), \tag{4.7}$$

$$0 = \frac{\partial B_r}{\partial x_2} = 9 + \frac{2}{r}(x_1 + x_2 - 4). \tag{4.8}$$

Answers

Hence, $4x_1 + \frac{r}{x_1} - 9 = 0$, which has the solutions

$$x_1^+(r) = \frac{9+\sqrt{81-16r}}{8}, \qquad x_1^-(r) = \frac{9-\sqrt{81-16r}}{8}.$$

x_2 is obtained from equation (4.8), *i.e.* $x_1 + x_2 - 4 + \frac{9r}{2} = 0$. [4 marks]

Some students may have not recognized that the equations could be solved directly for x_1 and hence may have obtained incorrect solutions. They may have also written incorrect necessary conditions of optimality for the barrier function.

ii) Computing

$$\lim_{r\to 0} x_1^+(r) = \lim_{r\to 0} \frac{9+\sqrt{81-16r}}{8} = \frac{9}{4}$$

yields $x_1 = \frac{9}{4}$ and $x_2 = \frac{7}{4}$, *i.e.* the optimal solution. [4 marks]

Incorrect limits!

iii) Computing

$$\lim_{r\to 0} x_1^-(r) = \lim_{r\to 0} \frac{9-\sqrt{81-16r}}{8} = 0$$

yields $x_1 = 0$ and $x_2 = 4$. This is a local maximizer. [4 marks]

Incorrect limits and/or characterization of the limit point.

Answers