EE4-29

OPTIMIZATION – MODEL ANSWERS

1. a) The stationary points of the function f are computed by solving the equations

$$0 = \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2(x_1-2)(2(x_1-2)^2+x_2^2) \\ 2x_1^2x_2 - 8x_1x_2 + 10x_2 + 2 \end{bmatrix}.$$

As a result, the point $x_\star = (2,-1)$ is the unique stationary point. [2 marks]

b) The Hessian matrix of the function f is

$$\nabla^2 f(x) = \begin{bmatrix} 12(x_1-2)^2 + 2x_2^2 & 4(x_1-2)x_2 \\ 4(x_1-2)x_2 & 2(x_1-2)^2+2 \end{bmatrix},$$

hence

$$\nabla^2 f(x_\star) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Since $\nabla^2 f(x_\star) > 0$, $x_\star$ is a minimizer of f. Note now that

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

and the function $(x_1-2)^4+(x_2+1)^2$ takes non-negative values and it is radially unbounded (it is the sum of two squares, one involving x_1 and one involving x_2). Hence, f is radially unbounded, and since $f(x_\star) = 0$, $x_\star$ is the global minimizer of f. [4 marks]

c) The modified Newton's iteration is given by

$$x_{k+1} = x_k - \frac{1}{2}\nabla f(x_k) = \begin{bmatrix} x_{1,k} - (x_{1,k}-2)(2(x_{1,k}-2)^2 + x_{2,k}^2) \\ -4x_{2,k} - x_{2,k}x_{1,k}^2 + 4x_{2,k}x_{1,k} - 1 \end{bmatrix}.$$

[4 marks]

d) The points generated by the modified Newton's iteration from the starting point $x_0 = (3/2, 0)$ are

$$x_1 = (1.75, -1), \quad x_2 = (2.03125, -0.9375), \quad x_3 = (2.003723145, -.9990844731),$$

$$x_4 = (2.000006711, -.9999861507), \qquad x_5 = (2.000000000, -1.000000000).$$

[4 marks]

e) The points generated by the modified Newton's iteration from the starting point $x_0 = (1,0)$ are

$$x_1 = (3,-1), \qquad x_2 = (0,0), \qquad x_3 = (16,-1), \qquad x_4 = (-5486, 195).$$

[2 marks]

f) The research direction used in the modified Newton's iteration is $-1/2\nabla f(x_k)$, which is nothing else than the direction of the anti-gradient, hence it is a descent direction satisfying the condition of angle. The reason why the method is not globally convergent is that the line search parameter is *fixed* to $\alpha = 1$, and this may not yield a descent algorithm at each step. [4 marks]

2. a) Note that

$$\nabla f = \begin{bmatrix} x_1 \\ m x_2 \end{bmatrix},$$

hence the gradient algorithm is described by the iteration

$$x_{1,k+1} = x_{1,k} - \alpha x_{1,k}, \qquad x_{2,k+1} = x_{2,k} - \alpha m\, x_{2,k}.$$

Replacing x_{k+1} in f yields

$$f(x_{k+1}) = \frac{1}{2}\left(x_{1,k}^2 + m\, x_{2,k}^2\right) - \alpha\left(x_{1,k}^2 + m^2 x_{2,k}^2\right) + \frac{1}{2}\left(x_{1,k}^2 + m^3\, x_{2,k}^2\right)\alpha^2.$$

To obtain the exact linear search parameter one has to compute the stationary point of $f(x_{k+1})$ as a function of α (since $f(x_{k+1})$ is convex in α), that is

$$\alpha_\star = \frac{x_{1,k}^2 + m^2 x_{2,k}^2}{x_{1,k}^2 + m^3\, x_{2,k}^2}.$$

As a result, the gradient algorithm with exact line search is given by

$$x_{k+1} = x_k - \alpha_\star \nabla f(x_k),$$

as given in the exam question. [6 marks]

b) As indicated in the exam question, for the considered initial condition and value of m the value of $\alpha_\star$ is constant, namely

$$\alpha_\star = \frac{x_{1,0}^2 + m^2 x_{2,0}^2}{x_{1,0}^2 + m^3\, x_{2,0}^2} = 1/5.$$

As a result, the gradient iteration is given by

$$x_{1,k+1} = \frac{4}{5} x_{1,k}, \qquad x_{2,k+1} = -\frac{4}{5} x_{2,k}.$$

This yields

$$x_{1,k} = x_{1,0}\left(\frac{4}{5}\right)^k = 9\left(\frac{4}{5}\right)^k, \qquad x_{2,k} = x_{2,0}\left(-\frac{4}{5}\right)^k = (-1)^k\left(\frac{4}{5}\right)^k,$$

as indicated in the exam paper. [8 marks]

c) Note that $x_\star = 0$, hence

$$\|x_{k+1}\|^2 = \left(9\left(\frac{4}{5}\right)^{k+1}\right)^2 + \left((-1)^{k+1}\left(\frac{4}{5}\right)^{k+1}\right)^2 = 82\left(\frac{4}{5}\right)^{2(k+1)},$$

$$\|x_k\|^2 = 82\left(\frac{4}{5}\right)^{2k},$$

thus

$$\frac{\|x_{k+1} - x_\star\|}{\|x_k - x_\star\|} = \frac{4}{5}.$$

The sequence thus converges with linear speed of convergence. [6 marks]

3. a) To begin with, rewrite the inequality constraint as $4 - x_1 - x_2 \leq 0$. The Lagrangian of the problem is

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

The necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 4x_1 - \rho, \qquad 0 = \frac{\partial L}{\partial x_2} = 9 - \rho,$$

$$\rho \geq 0, \qquad (4 - x_1 - x_2) \leq 0, \qquad \rho(4 - x_1 - x_2) = 0.$$

[4 marks]

b) Note that $\rho = 9$, hence the constrain has to be satisfied with the equality sign, thus yielding the only candidate solution $x_{1,\star} = 9/4$, $x_{2,\star} = 7/4$, $\rho = 9$.

[4 marks]

c) The stationary points of B_r are obtained solving the equations

$$0 = \nabla B_r = \begin{bmatrix} 4x_1 - \dfrac{r}{(4 - x_1 - x_2)^2} \\ 9 - \dfrac{r}{(4 - x_1 - x_2)^2} \end{bmatrix}.$$

The second equation yields $\dfrac{r}{(4 - x_1 - x_2)^2} = 9$, which substituted in the first equation yields $x_1 = 9/4$. Replacing $x_1 = 9/4$ in

$$(4 - x_1 - x_2)^2 = \frac{r}{9}$$

and solving for x_2 yields two stationary points

$$P_1 = \left(\frac{9}{4}, \frac{7}{4} + \frac{1}{3}\sqrt{r}\right), \qquad P_2 = \left(\frac{9}{4}, \frac{7}{4} - \frac{1}{3}\sqrt{r}\right).$$

The point P_1 is admissible for all $r > 0$, whereas P_2 is outside the admissible set. Note also that

$$\lim_{r \to 0} P_1 = \left(\frac{9}{4}, \frac{7}{4}\right) = P_{1,\star},$$

which coincides with the candidate optimal solution determined in part b).

[6 marks]

d) The necessary conditions of optimality for the constrained problem are

$$0 = \frac{\partial L}{\partial x_1}, \qquad 0 = \frac{\partial L}{\partial x_2},$$

and those for B_r are

$$0 = \frac{\partial B_r}{\partial x_1}, \qquad 0 = \frac{\partial B_r}{\partial x_2}.$$

Comparing the conditions yields

$$\frac{\partial L}{\partial x_1} = \frac{\partial B_r}{\partial x_1}, \qquad \frac{\partial L}{\partial x_2} = \frac{\partial B_r}{\partial x_2}.$$

As a result

$$\rho = \frac{r}{(4 - x_1 - x_2)^2},$$

which evaluated at P_1 yields $\rho = 9$, consistently with part b). [6 marks]

4. a) The Lagrangian of the problem is

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

The necessary conditions of optimality are

$$0 = \frac{\partial L}{\partial x_1} = 4x_1 - 1 + 2\lambda x_1, \qquad 0 = \frac{\partial L}{\partial x_2} = 4x_2 + 2\lambda x_2, \qquad 0 = x_1^2+x_2^2-1.$$

[2 marks]

b) The condition $0 = \frac{\partial L}{\partial x_2}$ yields either $x_2 = 0$ or $\lambda = -2$. The latter, replaced in $0 = \frac{\partial L}{\partial x_1}$ yields $0 = -1$, hence no candidate solution. The former yields two candidate solutions

$$S_1 = (x_1,x_2,\lambda) = \left(1,0,-\frac{3}{2}\right), \qquad S_2 = (x_1,x_2,\lambda) = \left(-1,0,-\frac{5}{2}\right).$$

[4 marks]

c) Note that

$$\nabla^2 L(S_1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \nabla^2 L(S_2) = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

hence S_1 yields a local minimizer, that is the solution of the problem. Note that S_2 yields a local maximizer. [4 marks]

d) The point $x_k = (\cos\theta_k, \sin\theta_k)$ is admissible.

i) The linearization of the constraint is given by the (linear) constraint

$$2(\cos\theta_k x_1 + \sin\theta_k x_2 - 1) = 0.$$

[2 marks]

ii) The Lagrangian of the problem is

$$L_l(x_1,x_2,\lambda) = 2(x_1^2+x_2^2-1) - x_1 + 2\lambda(\cos\theta_k x_1 + \sin\theta_k x_2 - 1).$$

The necessary conditions of optimality are

$$0 = \frac{\partial L_l}{\partial x_1} = 4x_1 - 1 + 2\lambda\cos\theta_k, \qquad 0 = \frac{\partial L_l}{\partial x_2} = 4x_2 + 2\lambda\sin\theta_k,$$

$$0 = \cos\theta_k x_1 + \sin\theta_k x_2 - 1.$$

These are linear equations in x_1, x_2 and λ with the unique solution

$$x_{1,\star} = \cos\theta_k + \frac{1}{4}\sin^2\theta_k, \qquad x_{2,\star} = \sin\theta_k - \frac{1}{4}\sin\theta_k \cos\theta_k,$$

with multiplier $\lambda = -2 + 1/2\cos\theta_k$. [4 marks]

iii) The update law is given by

$$\theta_{k+1} = \frac{\sin\theta_k(4-\cos\theta_k)}{4\cos\theta_k + \sin^2\theta_k}.$$

The sequence generated from the initial value $\theta_0 = 0.1$ is

$$\theta_1 = 0.07518803294, \quad \theta_2 = 0.05647085060, \quad \theta_3 = 0.04238693407,$$

$$\theta_4 = 0.03180448680, \qquad \theta_5 = 0.02385939890.$$

From the above one could conclude that the sequence converges to $\theta = 0$, however it takes more than five iterations to increase by one the number of correct digits. This slow convergence is typical of algorithms with linear speed of convergence. [4 marks]

