SF2822 Applied nonlinear optimization, final exam
Thursday June 1 2023 8.00–13.00
Brief solutions

1. (a) GAMS has terminated successfully with model status ``Locally Optimal'', meaning that a point $x^*$ and Lagrange multiplier vector $\lambda^*$ that together satisfy the first-order necessary optimality conditions for $(NLP)$ have been computed. From ``LEVEL'' of ``VAR x'', we obtain the solution

$$x^* \approx \begin{pmatrix} 0.433 & 0.433 & 0.567 \end{pmatrix}^T.$$

Analogously, the Lagrange multipliers of the constraints are given by ``MARGIN'' of ``EQU cons1'', ``EQU cons2'', and ``VAR x'', as

$$\lambda^* \approx \begin{pmatrix} 0.000 & 1.134 & 0.000 & 0.000 & 0.000 \end{pmatrix}^T.$$

(b) We have $f(x) = 2e^{(x_1-1)} + (x_2 - x_1)^2 + x_3^2 = f_1(x_1 - 1) + f_2(x_2 - x_1) + f_2(x_3)$, for $f_1(y) = 2e^y$ and $f_2(y) = y^2$. Then, $f_1''(y) = 2e^y \geq 0$ and $f_2''(y) = 2 \geq 0$, so that f_1 and f_2 are convex functions on $I\!R$. As linear tranformations preserve convexity, we obtain f as a sum of three convex functions, hence convex. In addition, $g_2(x) = x_1 + x_3 - 1$, which is linear. Hence, x^* and λ_2^* satisfy the first order necessary optimality conditions for

$$(NLP') \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_2(x) \geq 0, \end{array}$$

which is a convex optimization problem. Hence, x^* is a global minimizer to (NLP'). As (NLP') is a relaxation of (NLP) created by omitting constraints that are satisfied at x^*, x^* is a global minimizer to (NLP) as well.

(c) The expected change in the objective function is given by the Lagrange multiplier, up to first order, hence $1.456 + 1.134t$.

2. We may make use of the fact that the problem has only simple bounds. The solutions below are stated for the general form $Ax - b$. We first note that H is diagonally dominant, hence positive definite, so that (QP) is a convex optimization problem.

At iteration k, search direction $p^{(k)}$ and Lagrange multipliers $\lambda_{\mathcal{W}^{(k)}}^{(k+1)}$ are given by

$$\begin{pmatrix} H & A_{\mathcal{W}^{(k)}}^T \\ A_{\mathcal{W}^{(k)}} & 0 \end{pmatrix} \begin{pmatrix} p^{(k)} \\ -\lambda_{\mathcal{W}^{(k)}}^{(k+1)} \end{pmatrix} = - \begin{pmatrix} Hx^{(k)} + c \\ 0 \end{pmatrix}.$$

We have $x^{(0)} = (-1 \ 1 \ 0)^T$. Constraints 1 and 5 are active at $x^{(0)}$, so that $\mathcal{W}^{(0)} = \{1, 5\}$. We obtain

$$\begin{pmatrix} 2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(0)} \\ p_2^{(0)} \\ p_3^{(0)} \\ -\lambda_1^{(1)} \\ -\lambda_5^{(1)} \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

which gives

$$p^{(0)} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} \lambda_1^{(1)} \\ \lambda_5^{(1)} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{5}{2} \end{pmatrix}.$$

The maximum steplength $\alpha_{\max}$ is given by the maximum α such that $A(x^{(0)} + \alpha p^{(0)}) \geq b$, which gives $\alpha_{\max} = 2$, so that $\alpha^{(0)} = 1$ which gives $x^{(1)} = (-1\ 1\ 1/2)^T$. As $\lambda_5^{(1)} < 0$, we let $\mathcal{W}^{(1)} = \mathcal{W}^{(0)} \backslash \{5\} = \{1\}$.

We obtain

$$\begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ p_3^{(1)} \\ -\lambda_1^{(2)} \end{pmatrix} = \begin{pmatrix} -1 \\ -\frac{5}{2} \\ 0 \\ 0 \end{pmatrix},$$

which gives

$$p^{(1)} = \begin{pmatrix} 0 \\ -\frac{5}{3} \\ \frac{5}{6} \end{pmatrix}, \quad \lambda_1^{(1)} = -\frac{2}{3}.$$

The maximum steplength $\alpha_{\max}$ is given by the maximum α such that $A(x^{(1)} + \alpha p^{(1)}) \geq b$, which gives $\alpha_{\max} = 3/5$, where the maximum is attained for constraint 6, so that $\alpha^{(1)} = 3/5$ which gives $x^{(2)} = (-1\ 0\ 1)^T$ and $\mathcal{W}^{(2)} = \mathcal{W}^{(1)} \cup \{6\} = \{1, 6\}$.

We obtain

$$\begin{pmatrix} 2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(2)} \\ p_2^{(2)} \\ p_3^{(2)} \\ -\lambda_1^{(3)} \\ -\lambda_6^{(3)} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

which gives

$$p^{(2)} = \begin{pmatrix} 0 \\ -\frac{1}{2} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \lambda_1^{(3)} \\ \lambda_6^{(3)} \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

The maximum steplength $\alpha_{\max}$ is given by the maximum α such that $A(x^{(2)} + \alpha p^{(2)}) \geq b$, which gives $\alpha_{\max} = 2$, so that $\alpha^{(2)} = 1$ which gives $x^{(1)} = (-1\ -1/2\ 1)^T$. As $\lambda_1^{(3)} < 0$, we let $\mathcal{W}^{(3)} = \mathcal{W}^{(2)} \backslash \{1\} = \{6\}$.

We obtain

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(3)} \\ p_2^{(3)} \\ p_3^{(3)} \\ -\lambda_6^{(4)} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{pmatrix},$$

which gives

$$p^{(3)} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{6} \\ 0 \end{pmatrix}, \quad \lambda_6^{(4)} = \frac{2}{3}.$$

The maximum steplength $\alpha_{\max}$ is given by the maximum α such that $A(x^{(3)} + \alpha p^{(3)}) \ge b$, which gives $\alpha_{\max} = 2$, so that $\alpha^{(3)} = 1$ which gives $x^{(4)} = (-2/3 \ -2/3 \ 1)^T$. In addition, we have $\lambda^{(4)}_{\mathcal{W}^{(3)}} \ge 0$, so that $x^{(4)}$ is optimal.

3. (a) In this case, $x^{(0)} = 0$, so that $Ax^{(0)} = 0 > b$, since b has all components -1. Therefore, if s is introduced as $s = Ax - b$, we may let $s^{(0)} = Ax^{(0)} - b$, and then $Ax^{(k)} - s^{(k)} = b$ will be maintained at all Newton iterations, since this is a linear constraint. Therefore, we may use $Ax - b$ directly.

(b) The linear system of equations takes the form

$$\begin{pmatrix} H & -A^T \\ \operatorname{diag}(\lambda^{(0)})A & \operatorname{diag}(Ax^{(0)} - b) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} Hx^{(0)} + c - A^T\lambda^{(0)} \\ \operatorname{diag}(Ax^{(0)} - b)\operatorname{diag}(\lambda^{(0)})e - \mu^{(0)}e \end{pmatrix},$$

where e is the vector of ones. Insertion of numerical values gives

$$\begin{pmatrix} 2 & 1 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & -1 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta\lambda_3 \\ \Delta\lambda_4 \\ \Delta\lambda_5 \\ \Delta\lambda_6 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

(c) The unit step is accepted only if $-e < x^{(0)} + \Delta x < e$ and $\lambda^{(0)} + \Delta\lambda > 0$. Since $\lambda^{(0)} + \Delta\lambda \not> 0$, the unit step is not accepted. We may for example let $\alpha^{(0)} = 0.99\alpha_{\max}$, where $\alpha_{\max}$ is the maximum step, to maintain $x^{(0)} + \alpha\Delta x \ge 0$ and $\lambda^{(0)} + \alpha\Delta\lambda \ge 0$, i.e., $\alpha_{\max} = -\lambda_3^{(0)}/(\Delta\lambda_3)$. Then $x^{(1)} = x^{(0)} + \alpha^{(0)}\Delta x$ and $\lambda^{(1)} = \lambda^{(0)} + \alpha^{(0)}\Delta\lambda$.

4. (See the course material.)

5. (a) The QP subproblem becomes

$$(QP) \quad \begin{array}{ll} \text{minimize} & \frac{1}{2}p^T\nabla^2_{xx}\mathcal{L}(x^{(k)}, \lambda^{(k)})p + \nabla f(x^{(k)})^T p \\ \text{subject to} & A(x^{(k)})p \ge -g(x^{(k)}). \end{array}$$

Assume that there exists an x such that $g_i(x) \ge 0$. Then, if g_i is concave, we have

$$0 \ge -g_i(x) \ge -g_i(x^{(k)}) - \nabla g_i(x^{(k)})^T(x - x^{(k)}).$$

Repeating for $i = 1, \dots, m$ gives

$$0 \ge -g(x) \ge -g(x^{(k)}) - A(x^{(k)})(x - x^{(k)}),$$

which implies

$$A(x^{(k)})(x - x^{(k)}) \ge -g(x^{(k)}),$$

so that $x - x^{(k)}$ is feasible to the SQP subproblem.

(b) Linearization of the objective function in (NLP') gives

$$f(x^{(k)} + p) + Me^T(u^{(k)} + q) \approx f(x^{(k)}) + \nabla f(x^{(k)})^T p + Me^T u^{(k)} + Me^T q,$$

where e is the vector of ones. Since u only appears linearly in (NLP'), the quadratic part of the objective function in (QP') is identical to that in the objective function of the QP subproblem associated with (NLP). The objective function therefore becomes

$$\tfrac{1}{2} p^T \nabla^2_{xx} \mathcal{L}(x^{(k)}, \lambda^{(k)}) p + \nabla f(x^{(k)})^T p + Me^T q.$$

Setting a linearization of the constraints feasible in (NLP') gives

$$\begin{aligned} g(x^{(k)} + p) + u^{(k)} + q \approx g(x^{(k)}) + A(x^{(k)})p + u^{(k)} + q &\ge 0, \\ u^{(k)} + q &\ge 0. \end{aligned}$$

The QP subproblem associated with (NLP') at iteration k may consequently be written as

$$(QP') \quad \begin{array}{ll} \text{minimize} & \tfrac{1}{2} p^T \nabla^2_{xx} \mathcal{L}(x^{(k)}, \lambda^{(k)}) p + \nabla f(x^{(k)})^T p + Me^T q \\ \text{subject to} & A(x^{(k)})p + q \ge -g(x^{(k)}) - u^{(k)}, \\ & q \ge -u^{(k)}. \end{array}$$

If in addition $u^{(k)} = 0$, (QP') takes the form

$$(QP') \quad \begin{array}{ll} \text{minimize} & \tfrac{1}{2} p^T \nabla^2_{xx} \mathcal{L}(x^{(k)}, \lambda^{(k)}) p + \nabla f(x^{(k)})^T p + Me^T q \\ \text{subject to} & A(x^{(k)})p + q \ge -g(x^{(k)}), \\ & q \ge 0. \end{array}$$

The first-order necessary optimality conditions for (QP') then become

$$\begin{aligned} \nabla^2_{xx} \mathcal{L}(x^{(k)}, \lambda^{(k)}) p + \nabla f(x^{(k)}) &= A(x^{(k)})^T \lambda, \\ Me &= \lambda + \eta, \\ A(x^{(k)})p + q &\ge -g(x^{(k)}), \\ \lambda &\ge 0, \\ (A(x^{(k)})p + q + g(x^{(k)}))^T \lambda &= 0, \\ q &\ge 0, \\ \eta &\ge 0, \\ q^T \eta &= 0. \end{aligned}$$

Now assume that $Me - \lambda > 0$. Then, $\eta > 0$, since $\eta = Me - \lambda$. But then, the complementarity condition $q^T \eta = 0$ and nonnegativity of q gives $q = 0$. Then, the optimality conditions take the form

$$\begin{aligned} \nabla^2_{xx} \mathcal{L}(x^{(k)}, \lambda^{(k)}) p + \nabla f(x^{(k)}) &= A(x^{(k)})^T \lambda, \\ A(x^{(k)})p &\ge -g(x^{(k)}), \\ \lambda &\ge 0, \\ (A(x^{(k)})p + g(x^{(k)}))^T \lambda &= 0, \end{aligned}$$

which are exactly the optimality conditions of the QP subproblem associated with (NLP). Therefore, based on the optimality conditions we conclude that $q^{(k)} = 0$ in addition to $p^{(k)}$ and $\lambda^{(k+1)}$ being optimal solution and Lagrange multipliers respectively of the corresponding QP subproblem associated with (NLP).