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

*Examiner:* Anders Forsgren, tel. 08-790 71 27.

*Allowed tools:* Pen/pencil, ruler and eraser. *Note!* Calculator is not allowed.

*Solution methods:* Unless otherwise stated in the text, the problems should be solved by systematic methods, which do not become unrealistic for large problems. Motivate your conclusions carefully. If you use methods other than what have been taught in the course, you must explain thoroughly.

*Note!* Personal number must be written on the title page. Write only one question per sheet. Number the pages and write your name on each page.

22 points are sufficient for a passing grade. For 20-21 points, a completion to a passing grade may be made within three weeks from the date when the results of the exam are announced.

---

**1.** Consider the nonlinear programming problem

$$
(NLP) \qquad \begin{array}{ll}
\text{minimize} & 2e^{(x_1-1)} + (x_2 - x_1)^2 + x_3^2 \\
\text{subject to} & -x_1x_2x_3 \geq -2, \\
& x_1 + x_3 \geq 1, \\
& x_j \geq 0, \quad j = 1, 2, 3.
\end{array}
$$

A GAMS model of the problem has been created. The GAMS input file can be found at the end of the exam, and a partial GAMS output file reads:

```
               S O L V E      S U M M A R Y

     MODEL   nlpmodel              OBJECTIVE  obj
     TYPE    NLP                   DIRECTION  MINIMIZE
     SOLVER  SNOPT                 FROM LINE  21

**** SOLVER STATUS     1 Normal Completion
**** MODEL STATUS      2 Locally Optimal
**** OBJECTIVE VALUE                1.4559

 RESOURCE USAGE, LIMIT          0.043      1000.000
 ITERATION COUNT, LIMIT        14     2000000000
 EVALUATION ERRORS              0             0

SNOPT           24.8.5 r61358 Released May 10, 2017 DEG x86 64bit/Mac OS X

    GAMS/SNOPT, Large Scale Nonlinear SQP Solver
    S N O P T  7.2-12.1 (Jun 2013)
    P. E. Gill, UC San Diego
    W. Murray and M. A. Saunders, Stanford University
 Work space estimate computed by solver --     0.20 MB
 EXIT - Optimal Solution found, objective:        1.455938


                       LOWER     LEVEL     UPPER    MARGINAL

---- EQU objfun         -INF       .         .       -1.000
---- EQU cons1         -2.000    -0.106     +INF       .
---- EQU cons2          1.000     1.000     +INF      1.134
```

```
                                LOWER     LEVEL     UPPER    MARGINAL

---- VAR obj                    -INF      1.456     +INF        .

---- VAR x

          LOWER     LEVEL     UPPER    MARGINAL

j1          .        0.433     +INF  3.1022E-7
j2          .        0.433     +INF  -4.504E-8
j3          .        0.567     +INF        .
```

(a) Use the GAMS output file to give a point $x^*$ and Lagrange multiplier vector $\lambda^*$ that together satisfy the first-order necessary optimality conditions for $(NLP)$. ..................(4p)

(b) Is $x^*$ a global minimizer to $(NLP)$? .................. (4p)

(c) Assume that the second constraint is changed to $x_1 + x_3 \geq 1 + t$, where $t$ is a parameter. For $t$ near zero, give a prediction of the optimal value of the corresponding nonlinear program as a function of $t$. ..................(2p)

**2.** Consider the quadratic program $(QP)$ defined by

$$(QP) \quad \begin{array}{ll} \text{minimize} & \frac{1}{2}x^THx + c^Tx \\ \text{subject to} & \begin{pmatrix} I \\ -I \end{pmatrix} x \geq -\begin{pmatrix} e \\ e \end{pmatrix}, \end{array}$$

where $I$ is the identity matrix,

$$H = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad c = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \quad \text{and} \quad e = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

(This means that the constraints are given by $-1 \leq x_j \leq 1$, $j = 1, 2, 3$.)

Solve $(QP)$ by an active-set method, with the initial point $x^{(0)}$ given by $x^{(0)} = (-1\ 1\ 0)^T$ and the initial working set given by the constraints that are active at $x^{(0)}$. .................. (10p)

**3.** Consider the same quadratic program $(QP)$ as in Question 2.

Assume that we want to solve $(QP)$ with a primal-dual interior point method. Also assume that we initially choose $x^{(0)} = (0\ 0\ 0)^T$, $\lambda^{(0)} = (2\ 2\ 2\ 2\ 2\ 2)^T$, and $\mu = 2$.

(a) When the constraints are in the form $Ax \geq b$, one may introduce slack variables $s$ and rewrite the constraints as $Ax - s = b$, $s \geq 0$, when applying the interior method. Explain why this is not necessary for the given initial value $x^{(0)}$. .(2p)

(b) Formulate the linear system of equations to be solved in the first iteration of the primal-dual interior point method for the given initial values. Formulate the general form and then introduce explicit numerical values into the system of equations. ........................................................(5p)

(c) If the linear system of equations of Question 3b are solved, and the steps in the $x$-direction and the $\lambda$-direction are denoted by $\Delta x$ and $\Delta\lambda$ respectively, one obtains

$$\Delta x \approx \begin{pmatrix} -0.3039 \\ -0.1765 \\ 0.3627 \end{pmatrix}, \quad \Delta\lambda \approx \begin{pmatrix} -1.1922 \\ -1.4471 \\ -2.5255 \\ -2.4078 \\ -2.1529 \\ -1.0745 \end{pmatrix}.$$

Explain why it is not suitable to use the unit step, i.e, why it is not suitable to let $x^{(1)} = x^{(0)} + \Delta x$ and $\lambda^{(1)} = \lambda^{(0)} + \Delta\lambda$. Also explain how you would choose $x^{(1)}$ and $\lambda^{(1)}$. You need not give precise numerical values of $x^{(1)}$ and $\lambda^{(1)}$, but you should explain the principle. ........................................(3p)

**4.** Consider the semidefinite programming problem $(P)$ defined as

$$(P) \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & G(x) \succeq 0, \end{array}$$

where $G(x) = \sum_{j=1}^{n} A_j x_j - B$ for $B$ and $A_j$, $j = 1, \ldots, n$, are symmetric $m \times m$-matrices. The corresponding dual problem is given by

$$(D) \quad \begin{array}{lll} \text{maximize} & \operatorname{trace}(BY) & \\ \text{subject to} & \operatorname{trace}(A_j Y) = c_j, & j = 1, \ldots, n, \\ & Y = Y^T \succeq 0. & \end{array}$$

A barrier transformation of $(P)$ for a fixed positive barrier parameter $\mu$ gives the problem

$$(P_\mu) \qquad \text{minimize} \quad c^T x - \mu \ln(\det(G(x))).$$

(a) Show that the first-order necessary optimality conditions for $(P_\mu)$ are equivalent to the system of nonlinear equations

$$\begin{aligned} c_j - \operatorname{trace}(A_j Y) &= 0, \quad j = 1, \ldots, n, \\ G(x)Y - \mu I &= 0, \end{aligned}$$

assuming that $G(x) \succ 0$ and $Y \succ 0$ are kept implicitly. ..................(5p)

(b) Show that a solution $x(\mu)$ and $Y(\mu)$ to the system of nonlinear equations, such that $G(x(\mu)) \succ 0$ and $Y(\mu) \succ 0$, is feasible to $(P)$ and $(D)$ respectively with duality gap $m\mu$. ........................................................(3p)

(c) In linear programming, when $G(x)$ and $Y$ are diagonal, it is not an issue how the equation $G(x)Y - \mu I = 0$ is written. The linearizations of $G(x)Y - \mu I = 0$ and $YG(x) - \mu I = 0$ are then identical. Explain why this is in general not the case for semidefinite programming and how it can be handled. .......... (2p)

*Remark:* For a symmetric matrix $M$ we above use $M \succ 0$ and $M \succeq 0$ to denote that $M$ is positive definite and positive semidefinite respectively. You may use the relations

$$\frac{\partial \ln(\det(G(x)))}{\partial x_j} = \text{trace}(A_j G(x)^{-1}) \quad \text{for} \quad j = 1, \dots, n,$$

without proof.

**5.** Consider the nonlinear optimization problem $(NLP)$ given by

$$(NLP) \qquad \begin{array}{ll} \displaystyle\mathop{\text{minimize}}_{x \in I\!R^n} & f(x) \\ \text{subject to} & g_i(x) \geq 0, \quad i = 1, \dots, m, \end{array}$$

where $f$ and $g_i$, $i = 1, \dots, m$, are twice continuously differentiable.

(a) We want to solve $(NLP)$ by sequential quadratic programming. For a given point $x^{(k)}$, with $x^{(k)} \in I\!R^n$, and Lagrange multiplier vector estimate $\lambda^{(k)}$, with $\lambda^{(k)} \in I\!R^m$, formulate the SQP subproblem, i.e., the quadratic programming problem to be solved in the SQP method. In addition, show that if $(NLP)$ is feasible and $g_i$, $i = 1, \dots, m$, are concave functions on $I\!R^n$, then the SQP subproblem is feasible. .................................................. (5p)
*Hint:* If $\varphi : I\!R^n \to I\!R$ is convex and differentiable on $I\!R^n$, then $\varphi(u) \geq \varphi(v) + \nabla\varphi(v)^T(u - v)$ for all $u$, $v$ in $I\!R^n$.

(b) In order to create regularity and bound the multipliers, one may consider a reformulation of $(NLP)$ according to

$$(NLP') \qquad \begin{array}{ll} \displaystyle\mathop{\text{minimize}}_{x \in I\!R^n, u \in I\!R^m} & f(x) + M \sum_{i=1}^m u_i \\ \text{subject to} & g_i(x) + u_i \geq 0, \quad i = 1, \dots, m, \\ & u_i \geq 0, \quad i = 1, \dots, m, \end{array}$$

where $M$ is a large positive (fixed) number and $u$ are so-called elastic variables. Denote the Lagrange multipliers of the constraints $g(x) + u \geq 0$ by $\lambda$ and denote the Lagrange multipliers of the constraints $u \geq 0$ by $\eta$. Assume that we want to solve $(NLP')$ by sequential quadratic programming. Formulate the QP subproblem $(QP')$ for given $x^{(k)}$, $u^{(k)}$, $\lambda^{(k)}$ and $\eta^{(k)}$. Let $p$ and $q$ denote the variables corresponding to change in $x$ and $u$ respectively.
Assume that $\nabla^2_{xx}\mathcal{L}(x^{(k)}, \lambda^{(k)})$ is positive semidefinite. Let $p^{(k)}$, $q^{(k)}$ denote the optimal solution to $(SQP')$ and let $\lambda^{(k+1)}, \eta^{(k+1)}$ denote the corresponding Lagrange multipliers. Show that if $u^{(k)} = 0$ and $M > \max_{i=1,\dots,m}\{\lambda_i^{(k+1)}\}$, then $q^{(k)} = 0$ in addition to $p^{(k)}$ and $\lambda^{(k+1)}$ being optimal solution and Lagrange multipliers respectively of the corresponding QP subproblem associated with $(NLP)$. ............................................................ (5p)

*Note:* According to the convention of the book we define the Lagrangian $\mathcal{L}(x, \lambda)$ as $\mathcal{L}(x, \lambda) = f(x) - \lambda^T g(x)$, where $f(x)$ the objective function and $g(x)$ is the constraint function.

*Good luck!*

GAMS file for Question 1:

---

```
sets
        j / j1*j3 /;

variables
        obj
        x(j) ;

equations
        objfun
        cons1
        cons2 ;

objfun .. 2*exp(x("j1")-1)+power(x("j2")-x("j1"),2)+power(x("j3"),2) =l= obj;
cons1  .. -prod(j,x(j)) =g= -2;
cons2  .. x("j1")+x("j3") =g= 1;

x.lo(j) = 0;

model nlpmodel / all /;

solve nlpmodel using nlp minimizing obj;
```

---