



SF2822 Applied nonlinear optimization, final exam  
Wednesday August 15 2018 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 a particular nonlinear programming problem

$$\begin{aligned} & \text{minimize} && f(x) \\ (NLP) & \text{subject to} && g_i(x) \geq 0, \quad i = 1, \dots, 24, \\ & && x \in \mathbb{R}^9, \end{aligned}$$

where  $f : \mathbb{R}^9 \rightarrow \mathbb{R}$  and  $g : \mathbb{R}^9 \rightarrow \mathbb{R}^{24}$  are twice continuously differentiable. It is known that  $f$  and  $-g_i$ ,  $i = 1, \dots, 5$ , are convex on  $\mathbb{R}^9$ . It is also known that  $-g_i$ ,  $i = 6, 7, \dots, 24$ , are not convex on  $\mathbb{R}^9$ .

A highly reliable friend of yours named AR has written a Matlab file “`prob`” with syntax “[`f,gradf,g,A,HessL`] = `prob(x,lambda)`”. For given “`x`” and “`lambda`” ( $x$  and  $\lambda$ ), the function returns

$$\begin{aligned} \mathbf{f} & f(x), \\ \mathbf{gradf} & \nabla f(x), \\ \mathbf{g} & g(x), \\ \mathbf{A} & A(x), \quad \text{i.e.,} \quad \begin{pmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_{24}(x)^T \end{pmatrix}, \\ \mathbf{HessL} & \nabla_{xx}^2 \mathcal{L}(x, \lambda), \quad \text{i.e.,} \quad \nabla^2 f(x) - \sum_{i=1}^{24} \lambda_i \nabla^2 g_i(x), \end{aligned}$$

for this particular problem (*NLP*). You do not have access to “`prob`”, but we assume that the function is correct.

A not so reliable friend of yours named AF claims that he has solved the problem by making use of “`prob`”, finding a solution  $x^*$  and corresponding Lagrange multiplier vector  $\lambda^*$ . It is not clear what AF means by “solved”. You have access to a printout of some calculations that he has done in Matlab to verify the solution “`xstar`” and corresponding Lagrange multiplier vector “`lambdastar`”. Since these calculations only involve Matlab built-in functions and “`prob`”, we assume that they are correct in spite of AF not being so reliable. The Matlab printout can be found at the end of the exam.

AF has made a number of claims. Based on the Matlab printout, comment each of AF’s claims. Motivate each of your comments.

- (a) Claim by AF: “The point  $x^*$  is a regular point to (NLP).” .....(2p)
- (b) Claim by AF: “The point  $x^*$  satisfies, together with  $\lambda^*$ , the first-order necessary optimality conditions to (NLP).” .....(2p)
- (c) Claim by AF: “The point  $x^*$  satisfies, together with  $\lambda^*$ , the second-order necessary optimality conditions to (NLP).” .....(2p)
- (d) Claim by AF: “The point  $x^*$  satisfies, together with  $\lambda^*$ , the second-order sufficient optimality conditions to (NLP).” .....(2p)
- (e) Claim by AF: “The point  $x^*$  is a global minimizer to (NLP).” .....(2p)

2. Consider the nonlinear programming problem (NLP) defined by

$$\begin{aligned}
 & \text{minimize} && \frac{1}{2}(x_1 + x_2)^2 + \frac{3}{2}x_1 - \frac{9}{2}x_2 \\
 \text{(NLP)} & \text{subject to} && x_1 \cdot x_2 - 1 \geq 0. \\
 & && x_1 \geq 0, \\
 & && x_2 \geq 0.
 \end{aligned}$$

We want to solve (NLP) by sequential quadratic programming. Let  $x^{(0)} = (2 \ \frac{1}{2})^T$ ,  $\lambda^{(0)} = (1 \ 0 \ 0)^T$  and perform one iteration, i.e., calculate  $x^{(1)}$  and  $\lambda^{(1)}$ . You may solve the subproblem in an arbitrary way that need not be systematic, e.g. graphically, and you do not need to perform any linesearch. ....(10p)

Note: According to the convention of the textbook we define the Lagrangian  $\mathcal{L}(x, \lambda)$  as  $\mathcal{L}(x, \lambda) = f(x) - \lambda^T g(x)$ , where  $f(x)$  is the objective function and  $g(x)$  is the constraint function, with the inequality constraints written as  $g(x) \geq 0$ .

3. Consider the quadratic program (QP) given by

$$\begin{aligned}
 \text{(QP)} & \text{minimize} && \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \\
 & \text{subject to} && x_1 - 1 \geq 0.
 \end{aligned}$$

- (a) For a given positive barrier parameter  $\mu$ , find the corresponding optimal solution  $x(\mu)$  and the corresponding multiplier estimate  $\lambda(\mu)$  to the barrier-transformed problem. It is possible to obtain an analytic expression for this small problem. ....(5p)
- (b) Show that  $x(\mu)$  and  $\lambda(\mu)$  which you obtained in Question 3a converge to the optimal solution  $x^*$  and Lagrange multiplier  $\lambda^*$  respectively of (QP). ... (3p)
- (c) For  $x(\mu)$  and  $\lambda(\mu)$  which you obtained in Question 3a, how does  $\|x(\mu) - x^*\|_2$  and  $\|\lambda(\mu) - \lambda^*\|_2$  behave when  $\mu$  is small and positive? Is this as expected? Comment on the result. ....(2p)

4. Consider the quadratic program (QP) defined by

$$\begin{aligned}
 \text{(QP)} & \text{minimize} && \frac{1}{2}x^T H x + c^T x \\
 & \text{subject to} && A x = b,
 \end{aligned}$$

where

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad c = \begin{pmatrix} -4 \\ -5 \\ -6 \\ -7 \\ -8 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

The optimal solution to  $(QP)$  is given by  $x^* = (5 \ 4 \ 3 \ 2 \ 1)^T$ .

- (a) Determine a matrix  $Z$  whose columns form a basis for the nullspace of  $A$ . (2p)
- (b) It turns out that  $c_1$  was not correctly given in the original problem. It should have been  $c_1 = 16$ . Call this new problem  $(QP_2)$ . Solve  $(QP_2)$  making use of  $x^*$  and  $Z$ . ..... (4p)
- (c) It turns out that in addition to  $c_1 = 16$ , there should have been constraints  $x \geq 0$ . Call this new problem  $(QP_3)$ . Solve  $(QP_3)$  making use of your calculations in Question 4b. .... (4p)

5. 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  $A_j, j = 1, \dots, n$ , are symmetric  $m \times m$ -matrices. The corresponding dual problem is given by

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

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

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

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

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

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

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

without proof.

*Good luck!*



```
>> [g lambdastar]

ans =

    0    0.4330
  0.2500    0
    0    0.2165
  0.4330    0
  0.5000    0
  0.4330    0
  0.2500    0
  0.2500    0
  0.4330    0
  0.5000    0
  0.4330    0
  0.7500    0
  0.2500    0
  0.2500    0
  0.7500    0
  0.7500    0
  0.2500    0
  0.2500    0
  0.7500    0
  0.2500    0
  0.2500    0
  0.2500    0
  0.2500    0
  0.7500    0
  0.2500    0

>> norm(gradf-A'*lambdastar)

ans =

  8.7771e-17

>> eps

ans =

  2.2204e-16

>> active=find(g<sqrt(eps))

active =

    1
    3
```

(Matlab printout continues on next page.)

```
>> rank(A(active,:))
```

```
ans =
```

```
2
```

```
>> Z=null(A(active,:))
```

```
Z =
```

```
0.2222 -0.2750 0.3849 0.4763 -0.3849 -0.4763 0
0.2750 0.2222 0.4763 -0.3849 -0.4763 0.3849 0
0 0 0 0 0 0 1.0000
0.8889 -0.0393 -0.1925 0.0680 0.1925 -0.0680 0
0.0393 0.8889 0.0680 0.1925 -0.0680 -0.1925 0
-0.1925 -0.0680 0.6667 0.1179 0.3333 -0.1179 0
-0.0680 0.1925 -0.1179 0.6667 0.1179 0.3333 0
0.1925 0.0680 0.3333 -0.1179 0.6667 0.1179 0
0.0680 -0.1925 0.1179 0.3333 -0.1179 0.6667 0
```

```
>> eig(Z'*HessL*Z)
```

```
ans =
```

```
1.2990
0.6330
0.5000
0.5000
0.5000
0.5000
0
```

(End of Matlab printout.)