SF2822 Applied nonlinear optimization, final exam
Wednesday August 20 2015 8.00–13.00
Brief solutions

1. (a) As $g(x^*) = 0$, the constraint is active, and as $\nabla g(x^*)$ is nonzero, it holds that x^* is a regular point. Hence, for x^* to be a local minimizer to (NLP), the first-order necessary optimality conditions must hold. Hence, there must exist a nonnegative λ^* such that $\nabla f(x^*) = \nabla g(x^*)\lambda^*$, i.e.,

$$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \lambda^*.$$

There is no such λ^*. Hence, x^* is not a local minimizer to (NLP).

(b) The first-order optimality conditions

$$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \lambda^*,$$

are only violated in the last component if $\lambda^* = 2$. Hence, for this value of λ^*, we have

$$\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \lambda^* + \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix}.$$

Consequently, if we add a second constraint in the form of the bound-constraint $-x_3 \geq -x_3^*$ to (NLP), the first-order optimality conditions are satisfied for $\lambda_1^* = 2$, $\lambda_2^* = 2$.

In order to verify if x^* is a local minimizer, we now examine the second-order optimality conditions. We obtain

$$\begin{aligned} \nabla^2 \mathcal{L}(x^*, \lambda^*) &= \nabla^2 f(x^*) - \nabla^2 g(x^*)\lambda_1^* \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} - 2 \begin{pmatrix} -5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \end{aligned}$$

As we have strict complementarity, we now want to check the definiteness of the reduced Hessian of the Lagrangian with respect to the active constraint gradients, given by

$$A(x^*) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

However, $\nabla^2 \mathcal{L}(x^*, \lambda^*)$ is a diagonal matrix with positive diagonal elements, hence positive definite. Thus, the reduced Hessian is also positive definite.

We conclude that the second-order sufficient optimality conditions hold at x^* together with λ^*. Hence, x^* is a local minimizer to the problem where the bound-constraint $-x_3 \geq -x_3^*$ has been added.

2. No constraints are active at the initial point. Hence, the working set is empty, i.e., $\mathcal{W} = \emptyset$. Since $H = I$ and $c = 0$, we obtain $p^{(0)} = -(Hx^{(0)} + c) = -x^{(0)}$. The maximum steplength is given by

$$\alpha_{\max} = \min_{i: a_i^T p^{(0)} < 0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = \frac{1}{3},$$

where the minimum is attained for $i = 3$. Consequently, $\alpha^{(0)} = 1/3$ so that

$$x^{(1)} = x^{(0)} + \alpha^{(0)} p^{(0)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \end{pmatrix},$$

with $\mathcal{W} = \{3\}$. The solution to the corresponding equality-constrained quadratic program is given by

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ -\lambda_3^{(2)} \end{pmatrix} = - \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \\ 0 \end{pmatrix}$$

One way of solving this system of linear equations is to first express $p^{(1)}$ in $\lambda_3^{(2)}$ from the first two equations as

$$p_1^{(1)} = -\frac{2}{3} + \lambda_3^{(2)}, \quad p_2^{(1)} = -\frac{4}{3} + \lambda_3^{(2)}.$$

Insertion into the last equation gives $\lambda_3^{(2)} = 1$, so that

$$p^{(1)} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \end{pmatrix}^T.$$

The maximum steplength is given by

$$\alpha_{\max} = \min_{i: a_i^T p^{(0)} < 0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = 7,$$

which is attained for $i = 2$. Hence, $\alpha^{(1)} = 1$, so that

$$x^{(2)} = x^{(1)} + \alpha^{(1)} p^{(1)} = \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \end{pmatrix} + \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Since $\lambda_3^{(2)} \geq 0$, it follows that $x^{(2)}$ is the optimal solution.

3. Since $g(x^{(0)}) > 0$, it is not necessary to introduce slack variables for the constraints. If slack variables are not introduced, the Newton step Δx, $\Delta \lambda$ is given by

$$\begin{pmatrix} \nabla^2_{xx} \mathcal{L}(x, \lambda) & -A(x)^T \\ \Lambda A(x) & G(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} \nabla f(x) - A(x)^T \lambda \\ G(x) \lambda - \mu e \end{pmatrix},$$

where $G(x) = \operatorname{diag}(g(x))$, $\Lambda = \operatorname{diag}(\lambda)$ and e is the vector of ones.

In our case we get

$$\begin{pmatrix} 1+2\lambda & 0 & 2x_1 \\ 0 & 1+\lambda & x_2 \\ -2\lambda x_1 & -\lambda x_2 & 2-x_1^2-\frac{1}{2}x_2^2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} x_1 - 3 + 2\lambda x_1 \\ x_2 - 2 + \lambda x_2 \\ (2 - x_1^2 - \frac{1}{2}x_2^2)\lambda - \mu \end{pmatrix}.$$

Then, for the first iteration we obtain

$$\begin{pmatrix} 5 & 0 & 2 \\ 0 & 3 & 0 \\ -4 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x_1^{(0)} \\ \Delta x_2^{(0)} \\ \Delta\lambda^{(0)} \end{pmatrix} = -\begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}.$$

The next iterate is given by $x^{(1)} = x^{(0)} + \alpha^{(0)}\Delta x_1^{(0)}$, $\lambda^{(1)} = \lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)}$, where $\alpha^{(0)}$ is given by some approximate linesearch. The steplength $\alpha^{(0)}$ must be chosen such that $g(x^{(0)} + \alpha^{(0)}\Delta x^{(0)}) > 0$ and $\lambda^{(0)} + \alpha^{(0)}\Delta\lambda^{(0)} > 0$.

4. (See the course material.)

5. (a) The Lagrange multiplier $\lambda^{(1)}$ corresponds to an inequality constraint in the SQP subproblem. Hence, it must be nonnegative. This is not the case in the printout.

(b) We have

$$f(x) = \frac{1}{2}(x_1+1)^2 + \frac{1}{2}(x_2+2)^2, \quad g(x) = 3(x_1+x_2-2)^2 + (x_1-x_2)^2 - 6,$$
$$\nabla f(x) = \begin{pmatrix} x_1+1 \\ x_2+2 \end{pmatrix}, \quad \nabla g(x) = \begin{pmatrix} 8x_1+4x_2-12 \\ 4x_1+8x_2-12 \end{pmatrix},$$
$$\nabla^2 f(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \nabla^2 g(x) = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}.$$

Insertion of numerical values in the expressions above gives the first QP-problem according to

$$\begin{array}{ll} \min & \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2 + p_1 + 2p_2 \\ \text{subject to} & -12p_1 - 12p_2 \geq -6. \end{array}$$

This is a convex QP-problem with a globally optimal solution given by $p_1 = -1$, $p_2 = -2$ and $\lambda = 0$, so that $x^{(1)} = (-1 \quad -2)^T$, $\lambda^{(1)} = 0$.

(c) We can see that $g(x^{(1)}) = 3 \cdot 25 + 1 - 6 = 70 \geq 0$, so that $x^{(1)}$ is feasible to (NLP). In addition, since $f(x)$ is a strictly convex quadratic function and $\lambda^{(0)} = 0$, it follows that $x^{(1)}$ is a global minimizer to $f(x)$ over all $I\!R^2$. Hence, since $x^{(1)}$ is feasible to (NLP), it follows that it is a global minimizer to (NLP).