SF2822 Applied nonlinear optimization, final exam
Thursday August 18 2016 8.00–13.00
Brief solutions

1. 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

$$\alpha_{\max} = \min_{i: a_i^T p^{(0)} < 0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = \frac{1}{2},$$

where the minimium is attained for $i = 2$. Consequently, $\alpha^{(0)} = 1/2$ so that

$$x^{(1)} = x^{(0)} + \alpha^{(0)} p^{(0)} = \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 \\ -3 \\ -6 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{3}{2} \\ 3 \end{pmatrix},$$

with $\mathcal{W} = \{2\}$. The solution to the corresponding equality-constrained quadratic progam is given by

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ p_3^{(1)} \\ -\lambda_2^{(2)} \end{pmatrix} = - \begin{pmatrix} 0 \\ \frac{3}{2} \\ 3 \\ 0 \end{pmatrix}$$

One way of solving this system of linear equations is to first express $p^{(1)}$ in $\lambda_2^{(2)}$ from the first three equations as

$$p_1^{(1)} = \lambda_2^{(2)}, \quad p_2^{(1)} = -\frac{3}{2} + 2\lambda_2^{(2)}, \quad p_3^{(1)} = -3 + \lambda_2^{(2)}.$$

Insertion into the last equation gives $\lambda_2^{(2)} = 1$, so that

$$p^{(1)} = \begin{pmatrix} 1 & \frac{1}{2} & -2 \end{pmatrix}^T.$$

The maximum steplength is given by

$$\alpha_{\max} = \min_{i: a_i^T p^{(0)} < 0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = \frac{9}{5}.$$

Hence, $\alpha^{(1)} = 1$, so that

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

Since $\lambda_2^{(2)} \geq 0$, it follows that $x^{(2)}$ is the optimal solution.

2. (a) Since $Ax^{0)} > b$, there is no need to introduce s. We may let $s^{(0)} = Ax^{0)} - b = (1\ 1\ 1)^T$. Then, as $Ax - s = b$ is a linear equation, we will have $s^{(k)} = Ax^{(k)} - b$ throughout. Consequently, $s^{(k)}$ is just a notation for $Ax^{(k)} - b$ in this situation.

(b) The linear system of equations takes the form

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

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

$$\begin{pmatrix} 1 & 0 & 0 & -2 & -1 & -1 \\ 0 & 1 & 0 & -1 & -2 & -1 \\ 0 & 0 & 1 & -1 & -1 & -2 \\ 2 & 1 & 1 & 6 & 0 & 0 \\ 2 & 4 & 2 & 0 & 6 & 0 \\ 3 & 3 & 6 & 0 & 0 & 12 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta\lambda_3 \end{pmatrix} = -\begin{pmatrix} -7 \\ -5 \\ -3 \\ 5 \\ 11 \\ 35 \end{pmatrix}.$$

(c) The unit step is accepted only if $A(x^{(0)}+\Delta x)-b>0$ and $\lambda^{(0)}+\Delta\lambda>0$. This is the case here, and we may let $x^{(1)}=x^{(0)}+\Delta x$ and $\lambda^{(1)}=\lambda^{(0)}+\Delta\lambda$.

3. (See the course material.)

4. (a) The objective function is $f(x)=e^{x_1}-x_1x_2+x_2^2-2x_2x_3+x_3^2$. Differentiation gives

$$\nabla f(x) = \begin{pmatrix} e^{x_1}-x_2 \\ -x_1+2x_2-2x_3 \\ -2x_2+2x_3 \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} e^{x_1} & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -2 & 2 \end{pmatrix}.$$

In particular, $\nabla f(\widetilde{x}) = (0\;2\,{-2})^T$. With $g_1(x)=-x_1^2-x_2^2-x_3^2+5$ we get $g_1(\widetilde{x})=4$, which mean that constraint 1 is not active in $\widetilde{x}$. The first-order necessary optimality conditions require the existence of a $\tilde{\lambda}_2$ such that $\nabla f(\widetilde{x})=a\tilde{\lambda}_2$ and $a^T\widetilde{x}+1=0$.

The condition $\nabla f(\widetilde{x})=a\tilde{\lambda}_2$ can not be fulfilled with $\tilde{\lambda}_2=0$. Hence, $\tilde{\lambda}_2\neq 0$, and we obtain

$$a=\frac{1}{\tilde{\lambda}_2}\nabla f(\widetilde{x}) = \frac{1}{\tilde{\lambda}_2}\begin{pmatrix} 0 & 2 & -2 \end{pmatrix}^T.$$

The condition $a^T\widetilde{x}+1=0$ gives $\tilde{\lambda}_2=-2$. Hence, $a=(0\;\;-1\;\;1)^T$.

If $a=(0\;\;-1\;\;1)^T$, then $\widetilde{x}$ fulfils the first order of necessary optimality conditions together with $\tilde{\lambda}=(0\;\;-2)^T$.

(b) As we only have one active linear constraint in $\widetilde{x}$ we obtain

$$\nabla^2_{xx}\mathcal{L}(\widetilde{x},\tilde{\lambda}) = \nabla^2 f(\widetilde{x}) = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -2 & 2 \end{pmatrix}.$$

We also have that $A_A(\widetilde{x})=a^T$, where we can let $a^T=(N\;B)$ for $B=1$ and $N=(0\;\;-1)$. We then obtain a matrix whose columns form a basis for the null space of $A_A(\widetilde{x})$ as

$$Z_A(\widetilde{x}) = \begin{pmatrix} I \\ -B^{-1}N \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

which gives

$$Z_A(\widetilde{x})^T \nabla^2 f(\widetilde{x}) Z_A(\widetilde{x}) = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}.$$

But $Z_A(\widetilde{x})^T \nabla^2 f(\widetilde{x}) Z_A(\widetilde{x}) \not\succeq 0$ since $Z_A(\widetilde{x})^T \nabla^2 f(\widetilde{x}) Z_A(\widetilde{x})$ is a 2×2-matrix with negative determinant. Hence, $\widetilde{x}$ does not fulfil the second-order necessary optimality conditions and is therefore not a local minimizer.

5. (a) The matrix D has three negative eigenvalues. Since the problem has two constraints, at most two constraints can be active at any feasble point. At least three constraints would have to be active for the reduced Hessian of the objective function with respect to the active constraints to be positive semidefinite. This is a necessary condition for a local minimizer. Hence, no local minimizer can exist.

(b) As outlined in Question 5a, since D is diagonal with $D_{ii} < 0$, $i = 1, 3, 5$, there must be at least three active constraints at any local minimizer. Hence, it must hold that $Ax^* = b$, and in addition at least one more constraint must be active, which is linearly independent from the rows of A. Since $Ae_1 = 0$ and $e_1^T D e_1 < 0$, it follows that it is necessary to add a bound constraint of the form either $-x_1 \geq -1$ or $x_1 \geq 1$. If such a constraint is added, a matrix whose columns gives a basis for the resulting nullspace is given by A matrix whose columns form a basis for the nullspace of A is given by

$$Z = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix},$$

so that

$$Z^T D Z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which is a positive definite matrix. We assume that the bound-constraint $x_1 \geq 1$ is added, and try to establish the existence of Lagrange multipliers. The requirement becomes $Dx^* + c = A^T\lambda + e_1\lambda_3$, where $\lambda = (\lambda_1 \ \lambda_2)^T$. Thus,

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix},$$

which has a solution $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 1$. Since $\lambda_2 < 0$ we conclude that the first-order necessary conditions are not fulfilled and x^* is not a local minimizer. Hence, we cannot add a bound constraint so that x^* becomes a local minimizer.