SF2822 Applied nonlinear optimization, final exam
Tuesday June 3 2013 8.00–13.00
Brief solutions

1. (a) Both constraints are active at x^*. The first-order necessary optimality conditions then require the existence of λ_1^* and λ_2^*, with $\lambda_2^* \geq 0$, such that

$$\begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \lambda_1^* + \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \lambda_2^*.$$

There is a unique solution with $\lambda_1^* = -3$ and $\lambda_2^* = 1$, so that x^* satisfies the first-order necessary optimality conditions together with λ^*.

(b) Both Lagrange multipliers are nonzero, so that strict complementarity holds. A matrix $Z_+(x^*)$ whose columns form a basis for the nullspace of the matrix formed of the constraint gradients of the constraints with nonzero Lagrange multipliers, evaluated at x^*, is given by $Z_+(x^*) = (0\ 0\ 1)^T$. In addition to the first-order necessary optimality conditions, the second-order sufficient optimality conditions require

$$Z_+(x^*)^T \left(\nabla^2 f(x^*) - \lambda_2^* \nabla^2 g(x^*)\right) Z_+(x^*) \succ 0,$$

which gives

$$2 - \nabla^2 g(x^*)_{33} > 0.$$

Hence, x^* is a local minimizer if $\nabla^2 g(x^*)_{33} < 2$.

(c) Since conditions on f are only known at x^*, it is not sufficient to put any conditions on $\nabla^2 g(x)$ to ensure global minimality.

2. We may make use of the fact that the problem has only simple bounds.

Constraint 1 and 2 are in the working set at the initial point, i.e., x_1 and x_2 are set to zero. The search direction is given by

$$h_{33} p_3^{(0)} = -(h_{33} x_3^{(0)} + c_3), \quad \text{i.e.} \quad 3 p_3^{(0)} = -4,$$

so that $p^{(0)} = (0 \quad 0 \quad -4/3)^T$. The maximum steplength is given by $\alpha_{\max} = 3/4$, so that $\alpha^{(0)} = 3/4$ which gives $x^{(1)} = (0\ 0\ 0)^T$. All three constraints are active, so $p^{(1)} = 0$ and $x^{(2)} = x^{(1)}$. The multipliers are given by $\lambda^{(2)} = Hx^{(2)} + c = c$. Since $\lambda_1^{(2)} < 0$, constraint 1 is deleted from the working set. The search direction is given by

$$h_{11} p_1^{(2)} = -\lambda_1^{(2)}, \quad \text{i.e.} \quad 2 p_1^{(2)} = 2,$$

so that $p^{(2)} = (1\ 0\ 0)^T$. The maximum steplength is infinite, so that $\alpha^{(2)} = 1$ which gives $x^{(3)} = (1\ 0\ 0)^T$. The multipliers are given by $\lambda^{(3)} = Hx^{(3)} + c = (0\ 1\ -1)^T$. Since $\lambda_3^{(3)} < 0$, constraint 3 is deleted from the working set. The search direction is given by

$$\begin{pmatrix} h_{11} & h_{13} \\ h_{31} & h_{33} \end{pmatrix} \begin{pmatrix} p_1^{(3)} \\ p_3^{(3)} \end{pmatrix} = -\begin{pmatrix} \lambda_1^{(3)} \\ \lambda_3^{(3)} \end{pmatrix}, \quad \text{i.e.} \quad \begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} p_1^{(3)} \\ p_3^{(3)} \end{pmatrix} = -\begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

so that $p^{(3)} = (1\ 0\ 1)^T$. The maximum steplength is infinite, so that $\alpha^{(3)} = 1$ which gives $x^{(4)} = (2\ 0\ 1)^T$. The multipliers are given by $\lambda^{(4)} = Hx^{(4)} + c = (0\ 2\ 0)^T$. Since $\lambda^{(4)} \geq 0$, an optimal solution has been found.

3. (a) In this case $A = I$ and $b = 0$. Hence, since $Ax^{0)} = x^{(0)} > b = 0$, the initial point $x^{(0)}$ is strictly feasible and there is no need to introduce s. We may let $s^{(0)} = x^{0)} = (2\ 1\ 2)^T$. Then, as $x - s = 0$ is a linear equation, we will have $s^{(k)} = x^{(k)}$ throughout.

(b) The linear system of equations takes the form

$$\begin{pmatrix} H & -I \\ \text{diag}(\lambda^{(0)}) & \text{diag}(x^{(0)}) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} Hx^{(0)} + c - \lambda^{(0)} \\ \text{diag}(x^{(0)})\,\text{diag}(\lambda^{(0)})e - \mu^{(0)}e \end{pmatrix},$$

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

$$\begin{pmatrix} 2 & 0 & -2 & -1 & 0 & 0 \\ 0 & 2 & 1 & 0 & -1 & 0 \\ -2 & 1 & 3 & 0 & 0 & -1 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \\ \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta\lambda_3 \end{pmatrix} = -\begin{pmatrix} -3 \\ 3 \\ 3 \\ 1.8 \\ 1.8 \\ 1.8 \end{pmatrix}.$$

(c) The unit step is accepted only if $x^{(0)} + \Delta x > 0$ and $\lambda^{(0)} + \Delta\lambda > 0$. Since $x_2^{(0)} + \Delta x_s \not> 0$ and $\lambda_1^{(0)} + \Delta\lambda_1 \not> 0$, the unit step is not accepted. We may for example let $\alpha^{(0)} = 0.99\alpha_{\max}$, where $\alpha_{\max}$ is the maximum step, i.e., $\alpha_{\max} = -\lambda_1^{(0)}/(\Delta\lambda_1)$. 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) By adding an additional variable z, we may rewrite (P) as the nonlinear program

$$(NLP) \qquad \begin{array}{ll} \underset{x\in I\!R^n, z\in I\!R}{\text{minimize}} & z \\ \text{subject to} & z - f_i(x) \geq 0, \quad i = 1,\dots,m. \end{array}$$

As f_i, $i = 1,\dots,n$, are convex on $I\!R^n$, the functions $z - f_i(x)$ are concave on $I\!R^n \times I\!R$. Hence, (NLP) has a convex feasible region. In addition, it has a linear objective function, and is therefore a convex problem. Consequently, a local minimizer to (NLP) is also a global minimizer.

(b) The Lagrangian function associated with (NLP) is given by $\mathcal{L}(x,z,\lambda) = z - \sum_{i=1}^m \lambda_i(z - f_i(x))$. For a given (x,z,λ), the quadratic programming subproblem is given by

$$(QP) \qquad \begin{array}{ll} \underset{\Delta x\in I\!R^n, \Delta z\in I\!R}{\text{minimize}} & \frac{1}{2}\begin{pmatrix} \Delta z^T & \Delta x^T \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) \end{pmatrix} \begin{pmatrix} \Delta z \\ \Delta x \end{pmatrix} \\ & + \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \Delta z \\ \Delta x \end{pmatrix} \\ \text{subject to} & \begin{pmatrix} 1 & -\nabla f_i(x)^T \end{pmatrix} \begin{pmatrix} \Delta z \\ \Delta x \end{pmatrix} \geq -(z - f_i(x)), \quad i = 1,\dots,m. \end{array}$$

Simplification gives

$$(QP) \quad \begin{array}{ll} \underset{\Delta x \in \mathbb{R}^n, \Delta z \in \mathbb{R}}{\text{minimize}} & \frac{1}{2}\Delta x^T \left(\sum_{i=1}^m \lambda_i \nabla^2 f_i(x) \right) \Delta x + \Delta z \\ \text{subject to} & \Delta z - \nabla f_i(x)^T \Delta x \geq -(z - f_i(x)), \quad i = 1, \ldots, m. \end{array}$$

(c) The only concern regarding convexity of the quadratic programming subproblem is whether the Hessian is positive semidefinite. We know that the Lagrange multipliers of (NLP) are nonnegative, so it is natural to initially let $\lambda^{(0)} \geq 0$. The Hessian of the quadratic program is then given by $\sum_{i=1}^m \lambda_i^{(0)} \nabla^2 f_i(x^{(0)})$, which is positive semidefinite since $\lambda_i^{(0)} \geq 0$ and $\nabla^2 f_i(x^{(0)}) \succeq 0$ for $i = 1, \ldots, m$, due to the convexity of f_i, $i = 1, \ldots, m$. The Lagrange multipliers of the quadratic program give $\lambda^{(1)}$. They will be nonnegative, since (QP) has inequality constraints. We may now give this argument for the quadratic programming subproblem at iteration k and $k+1$, so convexity holds by induction if we initially let $\lambda^{(0)} \geq 0$. (In fact, we will have $\lambda^{(k)} \geq 0$, $\sum_{i=1}^m \lambda_i^{(k)} = 1$ for $k \geq 1$.)