KTH Mathematics

SF2822 Applied nonlinear optimization, final exam
Thursday December 17 2009 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

$$\alpha_{\max} = \min_{i:a_i^T p^{(0)}<0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = \frac{1}{5},$$

where the minimum is attained for $i = 3$. Consequently, $\alpha^{(0)} = 1/5$ so that

$$x^{(1)} = x^{(0)} + \alpha^{(0)} p^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \frac{1}{5}\begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{4}{5} \\ \frac{8}{5} \end{pmatrix},$$

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

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \end{pmatrix}\begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ p_3^{(1)} \\ -\lambda_3^{(2)} \end{pmatrix} = -\begin{pmatrix} 0 \\ \frac{4}{5} \\ \frac{8}{5} \\ 0 \end{pmatrix}$$

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

$$p_1^{(1)} = \lambda_3^{(2)}, \quad p_2^{(1)} = -\frac{4}{5} + \lambda_3^{(2)}, \quad p_3^{(1)} = -\frac{8}{5} + 2\lambda_3^{(2)}.$$

Insertion into the last equation gives $\lambda_3^{(2)} = 2/3$, so that

$$p^{(1)} = \begin{pmatrix} \frac{2}{3} & -\frac{2}{15} & -\frac{4}{15} \end{pmatrix}^T.$$

The maximum steplength is given by

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

as $Ap \geq 0$. Hence, $\alpha^{(1)} = 1$, so that

$$x^{(2)} = x^{(1)} + \alpha^{(1)} p^{(1)} = \begin{pmatrix} 0 \\ \frac{4}{5} \\ \frac{8}{5} \end{pmatrix} + \begin{pmatrix} \frac{2}{3} \\ -\frac{2}{15} \\ -\frac{4}{15} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{4}{3} \end{pmatrix}.$$

Since $\lambda_3^{(2)} \geq 0$, 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

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

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

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

(c) The unit step is accepted only if $A(x^{(0)} + \Delta x) - b > 0$ and $\lambda^{(0)} + \Delta\lambda > 0$. Since $A\Delta x \geq 0$, there is no restriction on the step for the x-variables, but since $\lambda_1^{(0)} + \Delta\lambda_1 \not> 0$, the unit step is not accepted for the λ-variables. 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 $\lambda^{(1)} = \lambda^{(0)} + \alpha^{(0)}\Delta\lambda$.

3. (See the course material.)

4. (a) Since (NLP') is formed by perturbing the first constraint of (NLP) from $h(x) \geq 0$ to $h(x) \geq 1/2$, sensitivity analysis gives the estimate

$$f(\tilde{x}) + \frac{1}{2}\tilde{\lambda}_1 = f(\tilde{x}) + 1 = 6.$$

(b) The QP-subproblem takes the form

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}p^T\nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)})p + \nabla f(x^{(0)})^T p \\ \text{subject to} & \nabla g(x^{(0)})\ p \geq -g(x^{(0)}). \end{array}$$

We obtain

$$\begin{aligned} \nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)}) &= \nabla^2 f(x^{(0)}) - \lambda_1^{(0)}\nabla^2 h(x^{(0)}) \\ &= \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} - 2\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Insertion of numerical values gives

$$\begin{array}{ll} \text{minimize} & 3p_1^2 + 2p_2^2 + 4p_1 \\ \text{subject to} & 2p_1 \geq \frac{1}{2}, \\ & p_1 \geq -5, \\ & p_2 \geq -4. \end{array}$$

This is a separable problem, so that minimization can be done with respect to p_1 and p_2 independently. We obtain $p_1 = 1/4$ and $p_2 = 0$ with Lagrange multipliers $\lambda_1 = 11/4$, $\lambda_2 = 0$ and $\lambda_3 = 0$. Consequently,

$$x^{(1)} = \begin{pmatrix} \frac{21}{4} \\ 4 \end{pmatrix}, \quad \lambda^{(1)} = \begin{pmatrix} \frac{11}{4} \\ 0 \\ 0 \end{pmatrix}.$$

5. The second-order sufficient optimality conditions for (NLP_1) imply that
(i) $g(x^*) \geq 0$,
(ii) $\nabla f(x^*) = A(x^*)^T \lambda^*$ for some $\lambda^* \geq 0$,
(iii) $\lambda_i^* g_i(x^*) = 0$, $i = 1, \ldots, m$, and
(iv) $Z_+(x^*)^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) Z_+(x^*) \succ 0$,
where $A_+(x^*)$ contains the rows of $A(x^*)$ for which λ^* has positive components, and $Z_+(x^*)$ is a matrix whose columns form a basis for $\text{null}(A_+(x^*))$.

We may write (NLP_2) as

$$(NLP_2) \qquad \begin{array}{ll} \text{minimize} & \widetilde{f}(z,x) \\ \text{subject to} & \widetilde{g}(z,x) \geq 0, \end{array}$$

with

$$\widetilde{f}(z,x) = z, \quad \widetilde{g}(z,x) = \begin{pmatrix} z - f(x) \\ g(x) \end{pmatrix}.$$

Associated with (NLP_2), we may define the Lagrangian function

$$\tilde{\mathcal{L}}(z, x, \mu, \eta) = z - \mu(z - f(x)) - \eta^T g(x),$$

where μ is the Lagrange multiplier associated with $z - f(x) \geq 0$ and η are the Lagrange multipliers associated with $g(x) \geq 0$.

We now want to find z^*, μ^* and η^* so that the second-order sufficient optimality conditions (i)–(iv) hold, but associated with (NLP_2). This means that we want to find z^*, μ^* and η^* such that
(i') $\begin{pmatrix} z^* - f(x^*) \\ g(x^*) \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$,
(ii') $\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\nabla f(x^*) & A(x^*)^T \end{pmatrix} \begin{pmatrix} \mu^* \\ \eta^* \end{pmatrix}$ for some $\mu^* \geq 0$ and $\eta^* \geq 0$,
(iii') $\mu^*(z^* - f(x^*)) = 0$, $\eta_i^* g_i(x^*) = 0$, $i = 1, \ldots, m$, and
(iv') $\tilde{Z}_+(z^*, x^*)^T \nabla^2_{z,x} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \tilde{Z}_+(z^*, x^*) \succ 0$,
where $\tilde{Z}_+(z^*, x^*)$ is a matrix whose columns form a basis for $\text{null}(\tilde{A}_+(z^*, x^*))$, with $\tilde{A}_+(z^*, x^*)$ defined as the matrix comprising the rows of

$$\begin{pmatrix} 1 & -\nabla f(x^*)^T \\ 0 & A(x^*) \end{pmatrix}$$

for which the associated components of the multipliers μ^* and η^* of (ii') are positive.

We now verify these conditions. For (i') to hold, we must have $z^* \geq f(x^*)$, since $g(x^*) \geq 0$ holds by (i).

For (ii'), the first equation reads $1 = \mu^*$. Hence, $\mu^* = 1$ must hold. With $\mu^* = 1$, the second block of equations reads

$$0 = -\nabla f(x^*) + A(x^*)^T \eta^*,$$

which holds for $\eta^* = \lambda^*$ by (ii). Since $\mu^* = 1 > 0$ and $\lambda^* \geq 0$ by (ii), (ii') holds.

Since $\mu^* > 0$, (iii') holds if $z^* = f(x^*)$, since (iii) implies that $\eta_i^* g_i(x^*) = 0$, $i = 1, \ldots, m$, if $\eta^* = \lambda^*$. In addition, since $z^* = f(x^*)$, (i') holds.

Finally, to verify (iv'), taking the derivatives gives

$$\nabla^2_{z,x} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) = \begin{pmatrix} \nabla^2_{xx} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) & \nabla^2_{xz} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \\ \nabla^2_{zx} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) & \nabla^2_{zz} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) \end{pmatrix}.$$

Since $\mu^* > 0$ and $\eta^* = \lambda^*$, we obtain

$$\tilde{A}_+(z^*, x^*) = \begin{pmatrix} 1 & -\nabla f(x^*)^T \\ 0 & A_+(x^*) \end{pmatrix} = \begin{pmatrix} 1 & -\lambda_+^{*\,T} A_+(x^*) \\ 0 & A_+(x^*) \end{pmatrix}.$$

Note that $\operatorname{rank}(\tilde{A}_+(z^*, x^*)) = \operatorname{rank}(A_+(x^*)) + 1$, since the first row of $\tilde{A}_+(z^*, x^*))$ is not linearly dependent on the other rows. Hence, $\operatorname{null}(A_+(z^*, x^*))$ and $\operatorname{null}(A_+(x^*))$ have the same dimension. Since

$$\begin{pmatrix} 1 & -\lambda_+^{*\,T} A_+(x^*) \\ 0 & A_+(x^*) \end{pmatrix} \begin{pmatrix} 0 \\ Z_+(x^*) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{we may let} \quad \tilde{Z}_+(z^*, x^*) = \begin{pmatrix} 0 \\ Z_+(x^*) \end{pmatrix}.$$

Then,

$$\begin{aligned} & \tilde{Z}_+(z^*, x^*)^T \nabla^2_{z,x} \tilde{\mathcal{L}}(z^*, x^*, \mu^*, \eta^*) \tilde{Z}_+(z^*, x^*) \\ &= \begin{pmatrix} 0 & Z_+(x^*)^T \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) \end{pmatrix} \begin{pmatrix} 0 \\ Z_+(x^*) \end{pmatrix} \\ &= Z_+(x^*)^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) Z_+(x^*) \succ 0, \end{aligned}$$

as required, where (iv) has been used in the last step. This means that the second-order sufficient optimality conditions hold for (NLP_2) with $z^* = f(x^*)$, $\mu^* = 1$ and $\eta^* = \lambda^*$.