KTH Mathematics

SF2822 Applied nonlinear optimization, final exam
Wednesday June 10 2009 8.00–13.00
Brief solutions

---

1. The objective function is $f(x) = e^{x_1} - x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 + 2x_3^2 - x_1 + x_2 - 2x_3$ and the constraint functions are $g_1(x) = x_1^2 + x_2^2 + x_3^2 - 1$, $g_2(x) = -x_1^2 - x_2^2 - x_3^2 + 2$. $g_3(x) = x_2$. Differentiation gives

$$\nabla f(x) = \begin{pmatrix} e^{x_1} - 2x_1 + x_2 - 1 \\ x_1 + x_2 + 1 \\ 4x_3 - 2 \end{pmatrix}, \quad \nabla g(x)^T = \begin{pmatrix} 2x_1 & 2x_2 & 2x_3 \\ -2x_1 & -2x_2 & -2x_3 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \begin{pmatrix} e^{x_1} - 2 - 2\lambda_1 + 2\lambda_2 & 1 & 0 \\ 1 & 1 - 2\lambda_1 + 2\lambda_2 & 0 \\ 0 & 0 & 4 - 2\lambda_1 + 2\lambda_2 \end{pmatrix}.$$

(a) Insertion of numerical values gives $\nabla f(\widetilde{x}) = (0\ 1\ 2)^T$, $g_1(\widetilde{x}) = 0$, $g_2(\widetilde{x}) = 1$ and $g_3(\widetilde{x}) = 0$. Hence, $\widetilde{x}$ is feasible with constraints 1 and 3 active.
As $\nabla g_1(\widetilde{x}) = (0\ 0\ 1)^T$ and $\nabla g_3(\widetilde{x}) = (0\ 1\ 0)^T$, it follows that $\widetilde{x}$ is a regular point. Hence, for $\widetilde{x}$ to be a local minimizer, the first-order of necessary optimality conditions must hold. They require the existence of a $\tilde{\lambda}_1$ and $\tilde{\lambda}_3$, with $\tilde{\lambda}_3 \geq 0$, such that $\nabla f(\widetilde{x}) = \nabla g_1(\widetilde{x})\tilde{\lambda}_1 + \nabla g_3(\widetilde{x})\tilde{\lambda}_3$, i.e.,

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_3 \end{pmatrix},$$

which is satisfied for $\tilde{\lambda}_1 = 1$, $\tilde{\lambda}_3 = 1$, i.e., the first-order necessary optimality conditions hold at $\widetilde{x}$ with Lagrange multiplier vector $\tilde{\lambda} = (1\ \ 0\ \ 1)^T$.
We obtain

$$\nabla^2_{xx}\mathcal{L}(\widetilde{x},\tilde{\lambda}) = \nabla^2 f(\widetilde{x}) - \tilde{\lambda}_1 \nabla^2 g_1(\widetilde{x}) = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

We obtain $Z(\widetilde{x}) = (1\ \ 0\ \ 0)^T$, so that $Z(\widetilde{x})^T \nabla^2_{xx}\mathcal{L}(\widetilde{x},\tilde{\lambda}) Z(\widetilde{x}) < 0$. Hence, the second-order necessary optimality conditions do not hold, i.e., $\widetilde{x}$ is not a local minimizer to (NLP).
As for $\widehat{x}$, insertion of numerical values gives $\nabla f(\widehat{x}) = (e - 3\ \ 2\ -2)^T$, $g_1(\widehat{x}) = 0$, $g_2(\widehat{x}) = 1$ and $g_3(\widehat{x}) = 0$. Hence, $\widehat{x}$ is feasible with constraints 1 and 3 active.
As $\nabla g_1(\widehat{x}) = (2\ 0\ 0)^T$ and $\nabla g_3(\widehat{x}) = (0\ 1\ 0)^T$, it follows that $\widehat{x}$ is a regular point. Hence, for $\widehat{x}$ to be a local minimizer, the first-order of necessary optimality

conditions must hold. They require the existence of a $\hat{\lambda}_1$ and $\hat{\lambda}_3$, with $\hat{\lambda}_3 \geq 0$, such that $\nabla f(\widehat{x}) = \nabla g_1(\widehat{x})\hat{\lambda}_1 + \nabla g_3(\widehat{x})\hat{\lambda}_3$, i.e.,

$$\begin{pmatrix} e-3 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_3 \end{pmatrix},$$

which has no solution. Hence, as $\widehat{x}$ is a regular point, $\widehat{x}$ is not a local minimizer to (NLP).

We conclude that neither $\widetilde{x}$ nor $\widehat{x}$ are local minimizers to (NLP).

2. If the problem is put on the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \geq 0, \quad x \in I\!R^2, \end{array}$$

we obtain

$$\nabla f(x)^T = \begin{pmatrix} x_1 + x_2 + \frac{5}{2} & x_1 + x_2 - \frac{1}{2} \end{pmatrix}, \quad \nabla g(x)^T = \begin{pmatrix} x_2 & x_1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

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

With $x^{(0)} = (2\ \frac{1}{2})^T$ and $\lambda^{(0)} = (1\ 0\ 0)^T$, the first QP-problem becomes

$$\begin{array}{ll} \text{minimize} & \frac{1}{2} \begin{pmatrix} p_1 & p_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} 5 & 2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\ \text{subject to} & \begin{pmatrix} \frac{1}{2} & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ -2 \\ -\frac{1}{2} \end{pmatrix}. \end{array}$$

The optimal solution of the QP-problem is given by the feasible point which is closest, in 2-norm, to $(-5\ \ -2)^T$, i.e., $p^{(0)} = (-2\ \frac{1}{2})^T$ with Lagrange multipliers $\lambda^{(1)} = (\frac{5}{4}\ \frac{19}{8}\ 0)^T$. Thus, we have $\lambda^{(1)}$, and $x^{(1)}$ is given by $x^{(1)} = x^{(0)} + p^{(0)} = (0\ 1)^T$.

3. (See the course material.)

4. (a) We may write $A = (I\ a)$, with $a = (1\ \ -1\ \ 1\ \ -1)^T$. Then, a matrix whose columns form a basis for the nullspace of A is given by $Z = (-a^T\ 1)^T = (-1\ \ 1\ \ -1\ \ 1\ \ 1)^T$.

(b) As the new cost may be written as $c - 27e_1$, the step to the minimizer of the new problem can be written as $p = Zp_Z$, where

$$Z^T HZp_Z = -Z^T(Hx^* + c - 27e_1).$$

As x^* is optimal to the original problem we have $Z^T(Hx^* + c) = 0$, so that $Z^T HZp_Z = 27Z^T e_1$. Insertion of numerical values gives $15p_z = -27$, i.e., $p_Z = -27/15 = -9/5$. Hence, if the optimal solution to the new problem is denoted by $\bar{x}$, we obtain

$$\bar{x} = x^* - \frac{9}{5}\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6.8000 \\ 2.2000 \\ 4.8000 \\ 0.2000 \\ -0.8000 \end{pmatrix}.$$

(c) As $\bar{x}_5 < 0$, $\bar{x}$ is not feasible to the third problem. In the previous exercise, we computed p as the first step in an active-set method for solving the third problem. The maximum steplength is given by the maximum α such that $x^* + \alpha p \geq 0$. We obtain $\alpha = 5/9$. The new point, $\widehat{x}$, becomes $\widehat{x} = x^* + 5/9p = (6\ 3\ 4\ 1\ 0)^T$. This point is in fact optimal, as the Lagrange multiplier of an added constraint will become positive. If the constraint $x_5 \geq 0$ is added as a fifth constraint, this can be verified algebraically by solving

$$H\widehat{x} + c = \begin{pmatrix} -24 \\ -1 \\ 2 \\ -6 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \\ \hat{\lambda}_4 \\ \hat{\lambda}_5 \end{pmatrix},$$

to obtain the Lagrange multipliers. We obtain $\hat{\lambda}_1 = -24$, $\hat{\lambda}_2 = -1$, $\hat{\lambda}_3 = 2$, $\hat{\lambda}_4 = -6$, $\hat{\lambda}_5 = 12$. As $\hat{\lambda}_5 \geq 0$, the solution is optimal.

5. (a) The function $f(y) = y_+^2$ has derivative $f'(y) = 0$ for $y < 0$ and $f'(y) = 2y$ for $y > 0$. Hence, $f'(y)$ is continuous with $f'(0) = 0$. The second derivative is given by $f''(y) = 0$ for $y < 0$ and $f''(y) = 1$ for $y > 0$. Hence, f'' is discontinuous at $y = 0$. As a consequence, the objective function has discontinuous Hessian at points where $p_i^T x = u_i$ for some $i \in \mathcal{U}$ or $p_i^T x = l_i$ for some $i \in \mathcal{L}$.

(b) Consider a fixed x and minimize over y in (QP). We want to show that $y_i = (p_i^T x - u_i)_+$, $i \in \mathcal{U}$, and $y_i = (l_i - p_i^T x)_+$, $i \in \mathcal{L}$. Assume that $p_i^T x - u_i < 0$ for some $i \in \mathcal{U}$. Then, $y_i = 0$, since $y_i = 0$ is the the minimizer of y_i^2. Similarly, if $p_i^T x - u_i \geq 0$, the optimal choice of y_i is $y_i = p_i^T x - u_i$, as y_i^2 is a strictly increasing function for $y_i > 0$. Hence, $y_i = (p_i^T x - u_i)_+$, $i \in \mathcal{U}$, as required. The argument for $i \in \mathcal{L}$ is analogous.

(c) We may write the Lagrangian function as

$$l(x, y, \lambda, \eta) = \tfrac{1}{2}\sum_{i\in\mathcal{U}} y_i^2 + \tfrac{1}{2}\sum_{i\in\mathcal{L}} y_i^2 - \sum_{i\in\mathcal{U}} \lambda_i(y_i - p_i^T x + u_i) - \sum_{i\in\mathcal{L}} \lambda_i(y_i + p_i^T x - l_i) - x^T\eta,$$

for Lagrange multipliers $\lambda_i \geq 0$, $i \in \mathcal{U} \cup \mathcal{L}$, and $\eta \geq 0$. Let $P_{\mathcal{U}}$ be the matrix whose rows comprise p_i^T, $i \in \mathcal{I}$, and analogously for $P_{\mathcal{U}}$. Let subscripts $''\mathcal{U}''$ and $''\mathcal{L}''$ respectively denote the vectors with components in the two sets. Also, let $\Lambda_{\mathcal{U}} = \mathrm{diag}(\lambda_{\mathcal{U}})$, $Y_{\mathcal{U}} = \mathrm{diag}(y_{\mathcal{U}})$, $\Lambda_{\mathcal{L}} = \mathrm{diag}(\lambda_{\mathcal{L}})$, $Y_{\mathcal{L}} = \mathrm{diag}(y_{\mathcal{L}})$, $X = \mathrm{diag}(x)$ and $N = \mathrm{diag}(\eta)$. For a positive barrier parameter μ, the perturbed first-order optimality conditions may be written

$$\begin{aligned}
P_{\mathcal{U}}^T \lambda_{\mathcal{U}} - P_{\mathcal{L}}^T \lambda_{\mathcal{L}} - \eta &= 0, \\
y_{\mathcal{U}} - \lambda_{\mathcal{U}} &= 0, \\
y_{\mathcal{L}} - \lambda_{\mathcal{L}} &= 0, \\
\Lambda_{\mathcal{U}}(y_{\mathcal{U}} - P_{\mathcal{U}} x + u_{\mathcal{U}}) &= \mu e, \\
\Lambda_{\mathcal{L}}(y_{\mathcal{L}} + P_{\mathcal{L}} x - l_{\mathcal{L}}) &= \mu e, \\
N x &= \mu e.
\end{aligned}$$