SF2822 Applied nonlinear optimization, final exam
Thursday August 21 2014 14.00–19.00
Brief solutions

1. We have

$$f(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2, \qquad g(x) = x_1 + x_2 + x_2^2 + 2,$$

$$\nabla f(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad \nabla g(x) = \begin{pmatrix} 1 \\ 1 + 2x_2 \end{pmatrix},$$

$$\nabla^2 f(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \nabla^2 g(x) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

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

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

This is a convex QP-problem with a globally optimal solution given by

$$p_1 - \lambda = 0,$$
$$p_2 - \lambda = 0,$$
$$p_1 + p_2 = -2.$$

The solution is given by $p_1 = -1$, $p_2 = -1$ and $\lambda = -1$, which agrees with the printout from the SQP-solver.

(b) We can see that $\nabla^2 f(x)$ is positive definite and $\nabla^2 g(x)$ is positive semidefinite, independently of x. Moreover λ is non-positive in all iterations. This implies that the solution to each QP subproblem is optimal also for the case when the equality constraint is changed to a less than or equal constraint. Hence, the iterates would not change at all if the constraint was changed as suggested.

(c) The inequality-constrained problem is a convex problem, and in addition a relaxation of the original problem. Hence we get convergence towards a global minimizer of this problem, which is also a global minimizer of (NLP).

2. (a) The problem (QP) is a convex quadratic program. The primal part of the trajectory is obtained as minimizer to the barrier-transformed problem

$$(P_\mu) \qquad \min \quad \tfrac{1}{2}x_1^2 + \tfrac{1}{2}x_2^2 - \mu \ln(x_1 + x_2 - 2)$$

under the implicit condition that $x_1 + x_2 - 2 > 0$. The first-order optimality conditions of (P_μ) gives

$$x_1(\mu) - \frac{\mu}{x_1(\mu) + x_2(\mu) - 2} = 0,$$
$$x_2(\mu) - \frac{\mu}{x_1(\mu) + x_2(\mu) - 2} = 0.$$

These equations are symmetric in $x_1(\mu)$ and $x_2(\mu)$. Hence, $x_1(\mu) = x_2(\mu)$. This means that $2x_1(\mu)^2 - 2x_1(\mu) - \mu = 0$, from which it follows that

$$x_1(\mu) = x_2(\mu) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{\mu}{2}} = \frac{1}{2} + \frac{1}{2}\sqrt{1 + 2\mu}.$$

where the plus sign has been chosen for the square root to enforce $x_1(\mu) + x_2(\mu) - 2 > 0$. Since (P_μ) is a convex problem, this is a global minimizer.

The dual part of the trajectory, i.e. $\lambda(\mu)$, is normally given by $\lambda_i(\mu) = \mu/g_i(x(\mu))$, $i = 1, \ldots, m$. Here we only have one constraint, so

$$\lambda(\mu) = \frac{\mu}{x_1(\mu) + x_2(\mu) - 2} = \frac{\mu}{\sqrt{1+2\mu} - 1} = \frac{1}{2} + \frac{1}{2}\sqrt{1+2\mu}.$$

(b) As $\mu \to 0$ it follows that $x(\mu) \to (1\ 1)^T$ and $\lambda(\mu) \to 1$. Let $x^* = (1\ 1)^T$ and $\lambda^* = 1$. Then x^* and λ^* satisfy the first-order optimality conditions of (QP). Since (QP) is a convex problem, this is sufficient for global optimality of (QP).

(c) We have

$$x_1(\mu) - x_1^* = x_2(\mu) - x_2^* = -\frac{1}{2} + \frac{1}{2}\sqrt{1+2\mu} = \frac{1}{2}\mu + o(\mu).$$

This is as expected. We would expect $\|x(\mu) - x^*\|_2$ to be of the order μ near an optimal solution where regularity holds.

3. (See the course material.)

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

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

In particular, $\nabla f(\widetilde{x}) = (0\ {-1}\ {-1})^T$. With $g_1(x) = -x_1^2 - x_2^2 - x_3^2 + 5$ we get $g_1(\widetilde{x}) = 3$, which mean that constraint 1 is not active at $\widetilde{x}$. Since $\nabla f(\widetilde{x}) \neq 0$, constraint 2 must be active for $\widetilde{x}$ to possibly satisfy the first-order necessary optimality conditions. These conditions require the existence of a $\tilde{\lambda}_2$ such that $\nabla f(\widetilde{x}) = a\tilde{\lambda}_2$ and $a^T\widetilde{x} + 2 = 0$ with $\tilde{\lambda}_2 \geq 0$.

The condition $\nabla f(\widetilde{x}) = a\tilde{\lambda}_2$ takes the form

$$\begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \tilde{\lambda}_2.$$

and it can not be fulfilled with $\tilde{\lambda}_2 = 0$. Hence, $\tilde{\lambda}_2 > 0$, and we obtain $a_1 = 0$, $a_2 = a_3 = -1/\tilde{\lambda}_2$. The condition $-2/\tilde{\lambda}_2 + 2 = 0$ so that $\tilde{\lambda}_2 = 1$. Hence, $a = (0\ {-1}\ {-1})^T$.

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

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

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

Since $\tilde{\lambda}_2 > 0$, we also have that $A_+(\tilde{x}) = a^T$, where we can let $a^T = (N \; B)$ for $B = -1$ and $N = (0 \quad -1)$. We then obtain a matrix whose columns form a basis for the null space of $A_+(\tilde{x})$ as

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

which gives

$$Z_+(\tilde{x})^T \nabla^2 f(\tilde{x}) Z_+(\tilde{x}) = \begin{pmatrix} 1 & 1 \\ 1 & 8 \end{pmatrix},$$

which is a positive definite matrix. Hence, $\tilde{x}$ fulfils the second-order sufficient optimality conditions and is therefore a local minimizer.

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 $f''(y) = 2$ 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 $P_{\mathcal{L}}$. Let subscripts $''\mathcal{U}''$ and $''\mathcal{L}''$ respectively denote the vectors with components in the two sets. Also, let $\Lambda_{\mathcal{U}} = \operatorname{diag}(\lambda_{\mathcal{U}})$, $Y_{\mathcal{U}} = \operatorname{diag}(y_{\mathcal{U}})$, $\Lambda_{\mathcal{L}} = \operatorname{diag}(\lambda_{\mathcal{L}})$, $Y_{\mathcal{L}} = \operatorname{diag}(y_{\mathcal{L}})$, $X = \operatorname{diag}(x)$ and $N = \operatorname{diag}(\eta)$. For a positive barrier parameter μ, the perturbed first-order optimality conditions may be written

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