KTH Mathematics

SF2822 Applied nonlinear optimization, final exam
Saturday June 5 2008 8.00–13.00
Brief solutions

---

1. We have

$$f(x) = \frac{1}{2}(x_1+1)^2 + \frac{1}{2}(x_2+2)^2, \qquad g(x) = 3(x_1+x_2-2)^2 + (x_1-x_2)^2 - 6,$$

$$\nabla f(x) = \begin{pmatrix} x_1+1 \\ x_2+2 \end{pmatrix}, \qquad \nabla g(x) = \begin{pmatrix} 8x_1+4x_2-12 \\ 4x_1+8x_2-12 \end{pmatrix},$$

$$\nabla^2 f(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \nabla^2 g(x) = \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}.$$

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

$$\begin{array}{ll} \min & \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2 + p_1 + 2p_2 \\ \text{då} & -12p_1 - 12p_2 = -6. \end{array}$$

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

$$\begin{aligned} p_1 + 12\lambda &= -1, \\ p_2 + 12\lambda &= -2, \\ -12p_1 - 12p_2 &= -6. \end{aligned}$$

The solution is given by $p_1 = 3/4$, $p_2 = -1/4$ and $\lambda = -7/48$, which agree with the printout from the SQP-solver.

(b) We can see that $\nabla^2 f(x)$ and $\nabla^2 g(x)$ are positive definite, 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

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

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

$$\begin{aligned} x_1(\mu) - \frac{\mu}{x_1(\mu)+x_2(\mu)} &= 0, \\ x_2(\mu) - \frac{\mu}{x_1(\mu)+x_2(\mu)} &= 0. \end{aligned}$$

These equations are symmetric in $x_1(\mu)$ and $x_2(\mu)$. Hence, $x_1(\mu) = x_2(\mu)$. This mean that $2x_1(\mu)^2 - \mu = 0$, from which it follows that $x_1(\mu)^2 = \mu/2$. If one includes $x_1(\mu) = x_2(\mu)$ in the implicit constraint, it follows that $x_1(\mu) = x_2(\mu) = \sqrt{\mu/2}$. 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 $\lambda_i(\mu) = \mu/g_i(x(\mu))$, $i = 1, \dots, m$. Here we only have one constraint, so

$$\lambda(\mu) = \frac{\mu}{\sqrt{\frac{\mu}{2}} + \sqrt{\frac{\mu}{2}}} = \sqrt{\frac{\mu}{2}}.$$

(b) As $\mu \to 0$ it follows that $x(\mu) \to (0\ 0)^T$ and $\lambda(\mu) \to 0$. Let $x^* = (0\ 0)^T$ and $\lambda^* = 0$. 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 $\|x(\mu) - x^*\|_2 = \sqrt{\mu}$. The square root comes from the fact that we do not have strict complementarity at the solution, i.e., the constraint is active with a zero multiplier.

If the constraint was given by $x_1 + x_2 \geq a$, for a given a, we obtain

$$x_1(\mu) = x_2(\mu) = \frac{a}{4} + \sqrt{\frac{a^2}{16} + \frac{\mu}{2}}.$$

Hence, for $a \neq 0$, we obtain $\|x(\mu) - x^*\|_2 = O(\mu)$. It is only for the degenerate case, $a = 0$, we obtain $\|x(\mu) - x^*\|_2 = O(\sqrt{\mu})$.

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

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

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

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

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

The condition $a^T\tilde{x} + 2 = 0$ gives $\tilde{\lambda}_2 = -1$. Hence, $a = (-1\ \ 2\ -2)^T$.
If $a = (-1\ \ 2\ -2)^T$, then $\tilde{x}$ fulfils the first order of necessary optimality conditions together with $\tilde{\lambda} = (0\ -1)^T$.

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

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

We also have that $A_A(\tilde{x}) = a^T$, where we can let $a^T = (B\ N)$ for $B = -1$ and $N = (2\quad -2)$. We then obtain a matrix whose columns form a basis for the null space of $A_A(\tilde{x})$ as

$$Z_A(\tilde{x}) = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

which gives

$$Z_A(\tilde{x})^T\nabla^2 f(\tilde{x})Z_A(\tilde{x}) = \begin{pmatrix} 10 & -8 \\ -8 & 6 \end{pmatrix}.$$

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

5. (a) The relaxed problem is a non-convex quadratic programming problem. To obtain a lower bound of the original problem we do need to calculate a global minimizer of this non-convex relaxed problem, which in general is not computationally tractable.

(b) If we let (SDP') be the problem arising as the constraint $Y = xx^T$ is added to (SDP) we can replace Y with xx^T, which by (i) gives

$$(SDP')\qquad \begin{array}{ll} \min & c^Tx + \frac{1}{2}x^THx \\ \text{då} & \begin{pmatrix} xx^T & x \\ x^T & 1 \end{pmatrix} \succeq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ & x_j^2 = x_j, \quad j = 1,\dots,n. \end{array}$$

By hint (ii) we can see that the constraint

$$\begin{pmatrix} xx^T & x \\ x^T & 1 \end{pmatrix} \succeq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

is always fulfilled, hence (SDP') may be written as

$$(SDP')\qquad \begin{array}{ll} \min & c^Tx + \frac{1}{2}x^THx \\ & x_j^2 = x_j, \quad j = 1,\dots,n. \end{array}$$

But $x_j^2 = x_j$ if and only if $x_j \in \{0,1\}$. Hence, (SDP') and (P) are equivalent.