KTH Mathematics

SF2822 Applied nonlinear optimization, final exam
Saturday December 20 2008 8.00–13.00
Brief solutions

---

1. (a) The quadratic programming subproblems must have nonnegative values on λ. Since this is not the case in the prinout, the prinout cannot be correct.

(b) We have

$$
\begin{aligned}
f(x) &= e^{x_1} + \frac{1}{2}(x_1 + x_2 - 4)^2 + (x_1 - x_2)^2, \\
g(x) &= -(x_1 - 3)^2 - x_2^2 + 9, \\
\nabla f(x) &= \begin{pmatrix} e^{x_1} + 3x_1 - x_2 - 4 \\ -x_1 + 3x_2 - 4 \end{pmatrix}, \qquad \nabla g(x) = \begin{pmatrix} -2(x_1 - 3) \\ -2x_2 \end{pmatrix}, \\
\nabla^2 f(x) &= \begin{pmatrix} e^{x_1} + 3 & -1 \\ -1 & 3 \end{pmatrix}, \qquad \nabla^2 g(x) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.
\end{aligned}
$$

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

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}p^T H p + c^T p \\
\text{subject to} & Ap \geq b,
\end{array}
$$

with

$$
H = \begin{pmatrix} 4 & -1 \\ -1 & 3 \end{pmatrix}, \quad c = \begin{pmatrix} -3 \\ -4 \end{pmatrix}, \quad A = \begin{pmatrix} 6 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \end{pmatrix}.
$$

This is a convex quadratic program. If we guess that the constraint is inactive, we obtain

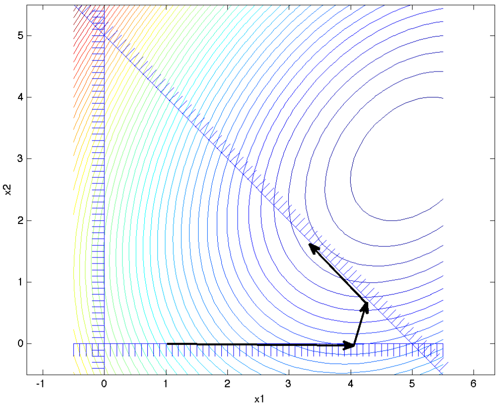
$$
p = -H^{-1}c = \begin{pmatrix} \frac{13}{11} \\ \frac{19}{11} \end{pmatrix}.
$$

For this p, it holds that $Ap \geq b$, and hence we have the optimal solution to the QP-problem, with $\lambda = 0$.

(c) The fact that the λ components from the prinout are negative suggests that the inequality constraint is incorrectly treated as an equality, i.e., the printout corresponds to

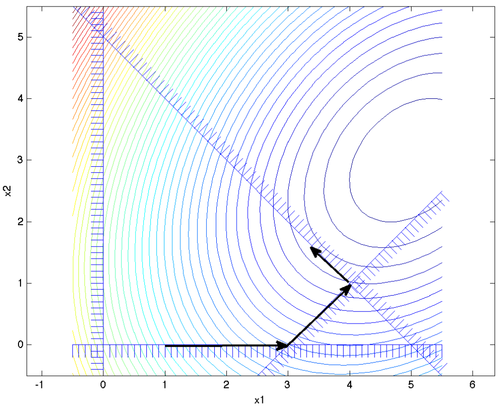
$$
\begin{array}{ll}
\text{minimize} & e^{x_1} + \frac{1}{2}(x_1 + x_2 - 4)^2 + (x_1 - x_2)^2 \\
\text{subject to} & -(x_1 - 3)^2 - x_2^2 + 9 = 0.
\end{array}
$$

2. (a) The iterations are illustrated in the figure below:



In the first iteration the search direction points at $(4\ 0)^T$, which is feasible. At this point, the multiplier of the constraint $x_2 \geq 0$ is negative, and the constraint is deleted from the active set. In the second iteration, the search direction points at $(5\ 3)^T$, but is limited by the constraint $-x_1 - x_2 \geq -5$, which is added. The search direction now points at $(7/2\ 3/2)^T$, which is feasible. The multiplier is positive, and the problem is thus solved.

(b) The iterations are illustrated in the figure below:



In the first iteration the search direction points at $(4\ 0)^T$, but the step is limited by the constraint $x_2 - x_1 \geq -3$, which is added. A zero step is taken, and the multiplier for the constraint $x_2 \geq 0$ is negative. This constraint is deleted. The new step is limited by the constraint $-x_1 - x_2 \geq -5$, which is added. A zero step is taken, and the multiplier for the constraint $x_2 - x_1 \geq -3$ is negative. This constraint is deleted, and the new step leads to the point $(7/2\ 3/2)^T$, which is feasible. The multiplier is positive, and the problem is thus solved.

3. (See the course material.)

4. The objective function is $f(x) = e^{x_1} + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 + x_3^2 - 2x_1 - x_2 - 3x_3$.

Differentiation gives

$$\nabla f(x) = \begin{pmatrix} e^{x_1} + x_1 + x_2 - 2 \\ x_1 + x_2 - 1 \\ 2x_3 - 3 \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} e^{x_1}+1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

(a) Insertion of numerical values gives $\nabla f(\widetilde{x}) = -(1\ 1\ 1)^T$. With $g_1(x) = x_1^2 + x_2^2 + x_3^2 - 1$ we get $g_1(\widetilde{x}) = 0$, which mean that constraint 1 is active in $\widetilde{x}$. With $g_2(x) = -x_1^2 - x_2^2 - x_3^2 + 2$ we get $g_2(\widetilde{x}) = 1$, which mean that constraint 2 is not active in $\widetilde{x}$.
The first-order of necessary optimality conditions require the existence of a $\tilde{\lambda}_1$ such that $\nabla f(\widetilde{x}) = \nabla g_1(\widetilde{x})\tilde{\lambda}_1$. This holds for $\tilde{\lambda}_1 = -1$, i.e., the first-order necessary optimality conditions hold at $\widetilde{x}$ with Lagrange multiplier vector $\tilde{\lambda} = (-1\ \ 0)^T$.

(b) We obtain

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \nabla^2 f(x) - \lambda_1 \nabla^2 g_1(x) = \begin{pmatrix} e^{x_1} + 1 - 2\lambda_1 & 1 & 0 \\ 1 & 1 - 2\lambda_1 & 0 \\ 0 & 0 & 2 - 2\lambda_1 \end{pmatrix}.$$

In particular,

$$\nabla^2_{xx}\mathcal{L}(\widetilde{x},\tilde{\lambda}) = \nabla^2 f(\widetilde{x}) - \tilde{\lambda}_1 \nabla^2 g_1(\widetilde{x}) = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

As $\nabla^2_{xx}\mathcal{L}(\widetilde{x},\tilde{\lambda}) \succ 0$, it must hold that $Z(\widetilde{x})^T \nabla^2_{xx}\mathcal{L}(\widetilde{x},\tilde{\lambda}) Z(\widetilde{x}) \succ 0$, and we need not compute $Z(\widetilde{x})$. Hence, the second-order sufficient optimality conditions hold.

(c) As $\tilde{\lambda}_1 \le 0$, it follows that $\widetilde{x}$ satisfies the first-order necessary optimality conditions for

$$(NLP') \quad \begin{array}{ll} \text{minimize} & e^{x_1} + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 + x_3^2 - 2x_1 - x_2 - 3x_3 \\ \text{subject to} & x_1^2 + x_2^2 + x_3^2 + x_1 + x_2 - x_3 \le 0. \end{array}$$

But (NLP') is a convex optimization problem, since $e^{x_1} + \frac{1}{2}x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 + x_3^2 - 2x_1 - x_2 - 3x_3$ and $x_1^2 + x_2^2 + x_3^2 + x_1 + x_2 - x_3$ are both convex functions on $I\!R^n$. Hence, since (NLP') is a relaxation of (NLP) for which the objective function is identical and the global minimizer $\widetilde{x}$ is feasible to (NLP), it follows that $\widetilde{x}$ is a global minimizer to (NLP).

5. Note that $X_{ii} = e_i^T X e_i$, where e_i is the ith unit vector. In addition, we may write $e_i^T X e_i = \text{trace}(e_i e_i^T X)$. Hence, we may rewrite (SDP_2) as

$$\begin{array}{ll} \underset{X \in \mathcal{S}^n}{\text{minimize}} & \text{trace}(WX) \\ \text{subject to} & \text{trace}(e_i e_i^T X) = 1, \quad i = 1, \dots, n, \\ & X \succeq 0. \end{array}$$

If we take the dual of this semidefinite program we obtain

$$\begin{array}{ll} \underset{y \in I\!R^n}{\text{maximize}} & \sum_{i=1}^n y_i \\ \text{subject to} & \sum_{i=1}^n e_i e_i^T y_i \preceq W. \end{array}$$

By noting that $\sum_{i=1}^n y_i = e^T y$, where e is the vector of ones, and that $\sum_{i=1}^n e_i e_i^T y_i = \text{diag}(y)$, an equivalent form is

$$\begin{array}{ll} \underset{y \in I\!R^n}{\text{maximize}} & e^T y \\ \text{subject to} & \text{diag}(y) \preceq W. \end{array}$$

But this problem is equivalent to (SDP_1), since we may change the sign of the objective function while changing maximization to minimization without altering the problem.

(As (SDP_1) and (SDP_2) both have nonempty relative interior, the duality gap is zero, and they provide the same bound on the optimal value of the original problem.)

This exercise is based on Exercise 5.39 in [1], where a more thorough discussion can be found.

References

[1] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge University Press, Cambridge, 2004.