# SF2822 Applied nonlinear optimization, final exam
**Friday August 23 2013 8.00–13.00**
**Brief solutions**

---

**1.** (a) We have

$$\nabla f(x) = \begin{pmatrix} 2e^{(x_1-1)} + 2x_1 - 2x_2 \\ 2x_2 - 2x_1 \\ 4x_3 \end{pmatrix} \quad \text{and in particular} \quad \nabla f(x^*) = \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}.$$

Since $\nabla f(x^*) \neq 0$, the point $x^*$ does not satisfy the first-order optimality conditions for an unconstrained problem. Hence, at least one constraint must be active. The point $x^*$ is feasible, and the only potentially active constraint is constraint 2 for $c = 3$. Since

$$\nabla g_2(x) = \begin{pmatrix} 1 & 0 & 2 \end{pmatrix}^T,$$

it follows that for $c = 3$, the first-order necessary optimality conditions require a $\lambda_2 \geq 0$ such that

$$\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \lambda_2,$$

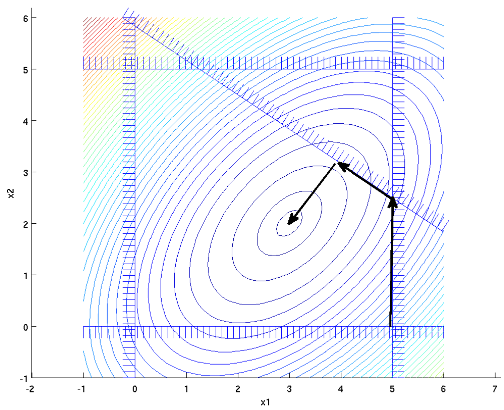
which holds for $\lambda_2 = 2$. Hence, for $c = 3$, it holds that $x^*$ satisfies the first-order necessary optimality conditions.

(b) The objective function is convex, and the only active constraint is linear. Hence, $x^*$ is a global minimizer to

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

But since $x^*$ is feasible to $(NLP)$ as well, and the only difference between $(NLP)$ and $(NLP')$ is that we have omitted the constraints that are not active at $x^*$, it follows that $x^*$ is globally optimal to $(NLP)$ as well.

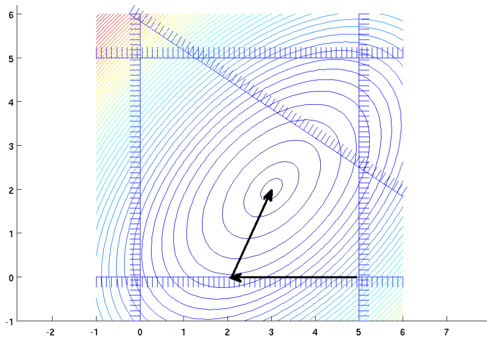
**2.** (a) The iterates are illustrated in the figure below:




At the first iteration constraint 3 is in the working set. The direction points at $(5\ 3)^T$, which is infeasible. The maximum step gives the new point $(5\ \frac{5}{2})^T$.

Constraint 5 is added, which gives a vertex and hence a zero step. Constraint 3 has a negative multiplier, and it is hence deleted. The direction points at $(\frac{305}{76}\ \frac{60}{19})^T$, which is feasible. Hence, the step there is taken. At this point, constraint 5 has a negative multiplier, and it is hence deleted. The direction points at (3 2) which is feasible. Hence, the step there is taken. No constraints are active, and we have found the optimal solution.

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




At the first iteration constraint 2 is in the working set. The direction points at $(2\ 0)^T$, which is feasible. Hence, the step there is taken. Constraint 2 has a negative multiplier, and it is hence deleted. The direction points at (3 2) which is feasible. Hence, the step there is taken. No constraints are active, and we have found the optimal solution.

**3.** We have

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

(a) The QP subproblem becomes

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}p^T\nabla^2_{xx}\mathcal{L}(x^{(0)}, \lambda^{(0)})p + \nabla f(x^{(0)})^T p \\ \text{subject to} & \nabla g_1(x^{(0)})^T p = -g_1(x^{(0)}), \\ & \nabla g_2(x^{(0)})^T p \geq -g_2(x^{(0)}), \\ & \nabla g_3(x^{(0)})^T p \geq -g_3(x^{(0)}). \end{array}$$

Insertion of numerical values gives

$$\begin{array}{ll} \text{minimize} & p_1^2 + p_2^2 - p_1 - 3p_2 \\ \text{subject to} & -p_1 + p_2 = 0, \\ & p_2 \geq -2, \\ & -p_1 \geq -4. \end{array}$$

The first constraint gives $p_1 = p_2$, so that we obtain

$$\begin{array}{ll} \text{minimize} & 2p_1^2 - 4p_1 \\ \text{subject to} & p_1 \geq -2, \\ & -p_1 \geq -4. \end{array}$$

The optimal solution is given by $p_1 = p_2 = 1$, and the Lagrange multiplier vector is given by $\lambda_1 = -1$, $\lambda_2 = 0$, $\lambda_3 = 0$.
Hence, we obtain $x^{(1)} = x^{(0)} + p = (1 \;\; 1)^T$ and $\lambda^{(1)} = \lambda = (-1 \;\; 0 \;\; 0)^T$.

(b) The change in the quadratic programming subproblem is that the constraint $-p_1 + p_2 = 0$ is replaced by $-p_1 + p_2 \geq 0$, i.e.,

$$\begin{array}{ll} \text{minimize} & p_1^2 + p_2^2 - p_1 - 3p_2 \\ \text{subject to} & -p_1 + p_2 \geq 0, \\ & p_2 \geq -2, \\ & -p_1 \geq -4. \end{array}$$

In the solution computed, it holds that $\lambda_1 < 0$. Hence, the solution does not remain optimal to the new problem. By temporarily ignoring the constraint $-p_1+p_2 \geq 0$, we obtain $p_1 = 1/2$, $p_2 = 3/2$, which is feasible and hence optimal. Thus, in this situation we obtain $x^{(1)} = x^{(0)} + p = (1/2 \;\; 3/2)^T$ and $\lambda^{(1)} = \lambda = (0 \;\; 0 \;\; 0)^T$.

(c) If the constraint is linear, the linearization of the constraint is exact, i.e.,

$$g_1(x^{(1)}) = g_1(x^{(0)}) + \nabla g_1(x^{(1)})^T p^{(0)}.$$

Since the constraints in the subproblem are given by the linearization, they will be satisfied in both cases.

**4.** (See the course material.)

**5.** (a) Let $f_\mu(x)$ denote the objective function of $(DQP_\mu)$, i.e.,

$$f_\mu(x) = \frac{1}{2} x^T H x + c^T x + \frac{1}{\mu} \sum_{i=1}^{n} x_i(1 - x_i).$$

Then we obtain

$$\nabla f_\mu(x) = Hx + c + \frac{1}{\mu} e - \frac{2}{\mu} x, \quad \nabla^2 f_\mu(x) = H - \frac{2}{\mu} I,$$

where $e$ is the vector of ones and $I$ is the identity matrix.
We obtain

$$\lambda_{\min}(\nabla^2 f_\mu(x)) = \lambda_{\min}(H) - \frac{2}{\mu},$$

since adding a multiple of the identity matrix just shifts the eigenvalues. Hence, $\lambda_{\min}(\nabla^2 f_\mu(x)) < 0$ if

$$\mu < \frac{2}{\lambda_{\min}(H)}.$$

Hence, let $\bar{\mu} = 2/\lambda_{\min}(H)$, where we let $\bar{\mu} = \infty$ if $\lambda_{\min}(H) = 0$.

(b) For the one-dimensional problem, we obtain

$$f'_\mu(x) = Hx + c + \frac{1}{\mu} - \frac{2}{\mu} x,$$

so that

$$f'_\mu(0) = c + \frac{1}{\mu}, \quad f'_\mu(1) = H + c - \frac{1}{\mu},$$

Hence, $f'_\mu(0) > 0$ if

$$\mu < \frac{1}{\max\{-c, 0\}}$$

and $f'_\mu(1) < 0$ if

$$\mu < \frac{1}{\max\{H + c, 0\}}.$$

Here, division by zero should be interpreted as infinity. Hence, by letting

$$\hat{\mu} = \min\left\{\frac{1}{\max\{-c, 0\}}, \frac{1}{\max\{H + c, 0\}}\right\}$$

we obtain $f'_\mu(0) > 0$ and $f'_\mu(1) < 0$ for $0 < \mu < \hat{\mu}$. This means that $x = 0$ and $x = 1$ are both local minimizers to $(DQP_\mu)$. If, in addition, $\mu < \bar{\mu}$, these are the only two local minimizers. The global minimizer is found by comparing the objective values of $f_\mu(0)$ and $f_\mu(1)$.

(c) The one-dimensional problem has two local minimizers for $\mu$ sufficiently small. One may expect that the $n$-dimensional problem has $2^n$ local minimizers for $\mu$ sufficiently small. Small values of $\mu$ also create high degree of nonconvexity. Hence, finding the global minimizer of such a problem is probably not a viable approach.

(d) If $H$ is diagonal, the problem decomposes into $n$ separate one-dimensional problems

$$(DQP^i) \qquad \begin{array}{ll} \underset{x_i \in I\!R}{\text{minimize}} & \frac{1}{2}H_{ii}x_i^2 + c_i x_i \\ \text{subject to} & x_i \in \{0, 1\}. \end{array}$$

Such a problem can easily be solved, as we have only two feasible points.

If $H_{ii}/2 + c_i < 0$, then $x_i^* = 1$, otherwise $x_i^* = 0$. Repeating this argument for $i = 1, \ldots, n$ gives the optimal solution $x^*$.