SF2822 Applied nonlinear optimization, final exam
Wednesday August 16 2023 8.00–13.00
Brief solutions

1. (a) The objective function is $f(x) = e^{x_1} + x_1x_2 + x_2^2 - 2x_2x_3 + x_3^2 - 3x_1 - x_2 - x_3$. Differentiation gives

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

In particular, $\nabla f(\widetilde{x}) = (-1\ -1\ -1)^T$. With $g_1(x) = -x_1^2 - x_2^2 - x_3^2 + 4$ we get $g_1(\widetilde{x}) = 2$, which mean that constraint 1 is inactive 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 $a^T\widetilde{x} + 3 = 0$ with $\tilde{\lambda}_2 \geq 0$.
The condition $\nabla f(\widetilde{x}) = a\tilde{\lambda}_2$ takes the form

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

and it can not be fulfilled with $\tilde{\lambda}_2 = 0$. Hence, $\tilde{\lambda}_2 > 0$, and we obtain $a_1 = a_2 = a_3 = -1/\tilde{\lambda}_2$. The condition $-3/\tilde{\lambda}_2 + 3 = 0$ gives $\tilde{\lambda}_2 = 1$. Hence, $a = (-1\ -1\ -1)^T$.
If $a = (-1\ -1\ -1)^T$, 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 $A_+(\widetilde{x}) = A_A(\widetilde{x}) = a^T$, where we can let $a^T = (B\ N)$ for $B = -1$ and $N = (-1\ \ -1)$. We then obtain a matrix whose columns form a basis for the null space of $A_+(\widetilde{x})$ as

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

which gives

$$Z_+(\widetilde{x})^T\nabla^2 f(\widetilde{x})Z_+(\widetilde{x}) = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix},$$

for which an LDL^T-factorization gives

$$\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

As $d_{22} = -1 < 0$, $Z_+(\widetilde{x})^T\nabla^2 f(\widetilde{x})Z_+(\widetilde{x})$ is not positive semidefinite. By strict complementarity, $A_+(\widetilde{x}) = A_A(\widetilde{x})$, so that $Z_+(\widetilde{x}) = Z_A(\widetilde{x})$. Therefore, $\widetilde{x}$ does not fulfil the second-order necessary optimality conditions and is therefore not a local minimizer.

2. 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

$$\begin{aligned} p_1 - \lambda &= 0, \\ p_2 - \lambda &= 0, \\ p_1 + p_2 &= -2. \end{aligned}$$

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).

3. We may make use of the fact that the problem has only simple bounds.

Constraint 1 and 2 are in the working set at the initial point, i.e., x_1 and x_2 are set to zero. The search direction is given by

$$h_{33}p_3^{(0)} = -(h_{33}x_3^{(0)} + c_3), \quad \text{i.e.} \quad 3p_3^{(0)} = -4,$$

so that $p^{(0)} = (0 \;\; 0 \;\; -4/3)^T$. The maximum steplength is given by $\alpha_{\max} = 3/4$, so that $\alpha^{(0)} = 3/4$ which gives $x^{(1)} = (0\ 0\ 0)^T$. All three constraints are active, so $p^{(1)} = 0$ and $x^{(2)} = x^{(1)}$. The multipliers are given by $\lambda^{(2)} = Hx^{(2)} + c = c$. Since $\lambda_1^{(2)} < 0$, constraint 1 is deleted from the working set. The search direction is given by

$$h_{11}p_1^{(2)} = -\lambda_1^{(2)}, \quad \text{i.e.} \quad 2p_1^{(2)} = 2,$$

so that $p^{(2)} = (1\ 0\ 0)^T$. The maximum steplength is infinite, so that $\alpha^{(2)} = 1$ which gives $x^{(3)} = (1\ 0\ 0)^T$. The multipliers are given by $\lambda^{(3)} = Hx^{(3)} + c = (0\ 1\ -1)^T$.

Since $\lambda_3^{(3)} < 0$, constraint 3 is deleted from the working set. The search direction is given by

$$\begin{pmatrix} h_{11} & h_{13} \\ h_{31} & h_{33} \end{pmatrix} \begin{pmatrix} p_1^{(3)} \\ p_3^{(3)} \end{pmatrix} = - \begin{pmatrix} \lambda_1^{(3)} \\ \lambda_3^{(3)} \end{pmatrix}, \quad \text{i.e.} \quad \begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} p_1^{(3)} \\ p_3^{(3)} \end{pmatrix} = - \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

so that $p^{(3)} = (1\ 0\ 1)^T$. The maximum steplength is infinite, so that $\alpha^{(3)} = 1$ which gives $x^{(4)} = (2\ 0\ 1)^T$. The multipliers are given by $\lambda^{(4)} = Hx^{(4)} + c = (0\ 2\ 0)^T$. Since $\lambda^{(4)} \geq 0$, an optimal solution has been found.

4. (See the course material.)

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 $\Lambda_{\mathcal{U}} = \text{diag}(\lambda_{\mathcal{U}})$, $Y_{\mathcal{U}} = \text{diag}(y_{\mathcal{U}})$, $\Lambda_{\mathcal{L}} = \text{diag}(\lambda_{\mathcal{L}})$, $Y_{\mathcal{L}} = \text{diag}(y_{\mathcal{L}})$, $X = \text{diag}(x)$ and $N = \text{diag}(\eta)$. For a positive barrier parameter μ, the perturbed first-order optimality conditions may be written

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