SF2822 Applied nonlinear optimization, final exam
Thursday June 3 2021 8.00–13.00
Brief solutions

---

1. (a) The objective value is $f(x) = e^{x_1} + x_1 x_2 + x_2^2 - 2x_2 x_3 + x_3^2$. To take the derivative gives

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

Especially we get $\nabla f(\tilde{x}) = (1 \ {-2} \ 2)^T$. With $g_1(x) = -x_1^2 - x_2^2 - x_3^2 + 10$ we get $g_1(\tilde{x}) = 9$, which mean that constraint 1 is not active in $\tilde{x}$. As $\nabla f(\tilde{x}) \neq 0$ constraint 2 has to be active with non-negative Lagrange multipliers for $\tilde{x}$ to be fulfil the first order of necessary optimality conditions. We get $\nabla f(\tilde{x}) = a\tilde{\lambda}_2$, where $\tilde{\lambda}_2 \geq 0$ and $a^T\tilde{x} = 2$.

The condition $\nabla f(\tilde{x}) = a\tilde{\lambda}_2$ can not be fulfilled as $\tilde{\lambda}_2 = 0$. With that we $a = \frac{1}{\lambda_2}\nabla f(\tilde{x})$. By combining with $a^T\tilde{x} = 2$, we have $\lambda_2 = 1$ and $a = \nabla f(\tilde{x}) = (1 \ {-2} \ 2)^T$.

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 $N = (-2 \ \ 2)$. With that we obtain a base

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

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. With that $\tilde{x}$ does not fulfil the second order of necessary optimality conditions and is therefore not a local min point.

2. (See the course material.)

3. We reformulate the constrains as $Ax \geq b$, where

$$A = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

Iteration 1: We have $\mathcal{W}^{(0)} = \{3,4\}$. The solution of the equality-constrained quadratic subproblem is given by

$$\begin{pmatrix} 2 & 1 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_1^{(0)} \\ d_2^{(0)} \\ d_3^{(0)} \\ -\lambda_3^{(1)} \\ -\lambda_4^{(1)} \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ -4 \\ 0 \\ 0 \end{pmatrix}.$$

To solve this system of equations gives $d^{(0)} = (0\ \frac{7}{3}\ 0\ )^T$, $\lambda_3^{(1)} = -\frac{5}{3}$ and $\lambda_4^{(1)} = 4$. Maximal step length is given by

$$\alpha_{\max}^{(0)} = \min_{i:a_i^T d^{(0)}<0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T d^{(0)}} = \frac{6}{7} < 1, \quad \text{for } i = 1.$$

Therefore we get $\alpha^{(0)} = \frac{6}{7}$, $\mathcal{W}^{(1)} = \{1,3,4\}$ and $x^{(1)} = x^{(0)} + \alpha^{(0)} d^{(0)} = (0\ 2\ 0)^T$.

Iteration 2: The solution of the equality-constrained quadratic subproblem is given by

$$\begin{pmatrix} 2 & 1 & 0 & -1 & 1 & 0 \\ 1 & 3 & 0 & -1 & 0 & 0 \\ 0 & 0 & 4 & -1 & 0 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_1^{(1)} \\ d_2^{(1)} \\ d_3^{(1)} \\ -\lambda_1^{(2)} \\ -\lambda_3^{(2)} \\ -\lambda_4^{(2)} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -4 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Solving this system of equations gives $d^{(1)} = (0\ 0\ 0)^T$, $\lambda_1^{(2)} = 1$, $\lambda_3^{(2)} = -1$ and $\lambda_4^{(2)} = 5$. As $d^{(1)} = (0\ 0\ 0)^T$ we get $x^{(2)} = x^{(1)} = (0\ 2\ 0)^T$.

Iteration 3: As $\lambda_3^{(2)} < 0$ we get $\mathcal{W}^{(2)} = \{1,4\}$. The solution of the equality-constrained quadratic subproblem is given by

$$\begin{pmatrix} 2 & 1 & 0 & -1 & 0 \\ 1 & 3 & 0 & -1 & 0 \\ 0 & 0 & 4 & -1 & 1 \\ -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_1^{(2)} \\ d_2^{(2)} \\ d_3^{(2)} \\ -\lambda_1^{(3)} \\ -\lambda_4^{(3)} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -4 \\ 0 \\ 0 \end{pmatrix}.$$

Solving this system of equations give $d^{(2)} = (\frac{1}{3}\ -\frac{1}{3}\ 0\ )^T$, $\lambda_1^{(3)} = \frac{5}{3}$ and $\lambda_4^{(3)} = \frac{17}{3}$. Maximal step length is given by

$$\alpha_{\max}^{(2)} = \min_{i:a_i^T d^{(2)}<0} \frac{a_i^T x^{(2)} - b_i}{-a_i^T d^{(2)}} = 3 > 1, \quad \text{for } i = 2.$$

Therefore we get $\alpha^{(2)} = 1$ and $x^{(3)} = x^{(2)} + \alpha^{(2)} d^{(2)} = (\frac{1}{3}\ \frac{5}{3}\ 0)^T$.

Iteration 3: As $\lambda^{(3)} \geq 0$, we complete the active-set method.

The optimal solution is $x = (\frac{1}{3}\ \frac{5}{3}\ 0)^T$ and the corresponding Lagrange multipliers is $\lambda = (\frac{5}{3}\ 0\ 0\ \frac{17}{3})^T$.

4. We have

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

$$g(x) = 1 - \frac{1}{2}x_1^2 - \frac{1}{2}x_2^2, \quad \nabla g(x) = \begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix}, \quad \nabla^2 g(x) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

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

$$\begin{array}{ll} \min & \frac{1}{2}\begin{pmatrix} p_1 & p_2 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\ \text{subject to} & -p_2 \geq -\frac{1}{2}. \end{array}$$

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

$$\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} -2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}\lambda \tag{1}$$

$$-p_2 \geq -\frac{1}{2} \tag{2}$$

$$\lambda(-p_2 + \frac{1}{2}) = 0, \lambda \geq 0. \tag{3}$$

If (2) is not active. $\lambda = 0$ and $-p_2 > -\frac{1}{2}$. Then, from (1), we have $p_1 = \frac{2}{3}, p_2 = \frac{2}{3}$, which is in contradiction with $-p_2 \geq -\frac{1}{2}$.
Then, (2) is active and $\lambda \geq 0$. We have $p_1 = \frac{2}{3}, p_2 = \frac{1}{2}$ and $\lambda = \frac{1}{2}$. Hence, $x^{(1)} = x^{(0)} + p = (\frac{2}{3}\ \frac{3}{2})^T, \lambda^{(1)} = \frac{1}{2}$.

(b) The Newton step Δx, $\Delta\lambda$ is given by

$$\begin{pmatrix} \nabla^2_{xx}\mathcal{L}(x,\lambda) & A(x)^T \\ \Lambda A(x) & -G(x) \end{pmatrix}\begin{pmatrix} \Delta x \\ -\Delta\lambda \end{pmatrix} = -\begin{pmatrix} \nabla f(x) - A(x)^T\lambda \\ G(x)\lambda - \mu e \end{pmatrix},$$

where $\Lambda = \text{diag}(\lambda)$ and $G(x) = \text{diag}(g(x))$.
Which gives the following system of linear equations

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -2 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ -\Delta\lambda \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is given by $\Delta x = (\frac{2}{3}\ 0)^T$, $\Delta\lambda = 0$.
Step-size α should be calculated such that $g(x + \alpha\Delta x) > 0$ and $\lambda + \alpha\Delta\lambda > 0$. This is valid for $\alpha = 1$. If we ignore the demand of the merit function and choose $\alpha = 1$, we obtain that $x^{(1)} = (\frac{2}{3}\ 1)^T$ and $\lambda^{(1)} = 2$.

5. (a) The dual problem can for example be written on the form

$$(DSDP) \quad \begin{array}{ll} \max & y \\ \text{subject to} & Iy \preceq M. \end{array}$$

(b) Let $\eta_i(M)$, $i = 1, \ldots, n$, denote the eigenvalue of M. If we add a multiple of the unit matrix of M the eigenvalues are shifted with that multiple. With that the matrix $M - Iy$ obtain the eigenvalues $\eta_i(M) - y$. Therefore y become feasible to $(DSDP)$ if and only if $y \leq \eta_{\min}(M)$, where $\eta_{\min}(M)$ denote the smallest eigenvalue of M. With that the optimal y become the smallest eigenvalue of M, which therefore is an optimal value of $(DSDP)$.

(c) If we restrict X to have the form xx^T in $(PSDP)$ we obtain the following problem

$$(P) \quad \begin{array}{ll} \min & \text{trace}(Mxx^T) \\ \text{subject to} & \text{trace}(xx^T) = 1. \end{array}$$

As $\text{trace}(Axx^T) = x^T Ax$ for a symmetric $n \times n$-matrix A then (P) can equivalently be written as

$$(P) \quad \begin{array}{ll} \min & x^T Mx \\ \text{subject to} & x^T x = 1. \end{array}$$

The optimal value of (P) is the smallest eigenvalue of M and the optimal solution x^* is an eigenvector of the norm one corresponding to this eigenvalue. As (P) is a restrification of $PSDP$), the optimal value of (P) is at least as large as the optimal value of $(DSDP)$. Our choice of x^* give the same objective function value in (P) as the optimal value of $(DSDP)$. With that x^* is the optimal solution of (P) which implies that $x^* x^{*T}$ is an optimal solution of $(PSDP)$.