Prof. Dr. Matthias Gerdts
Alberto De Marchi
LRT-1 UniBw
Wintertrimester 2021

Universität der Bundeswehr München
Professur für Ingenieurmathematik

Optimierung ME BSc · Übung 6

1. Betrachten Sie die Optimierungsprobleme

$$\min \quad x_1^2 - \frac{1}{2}x_2^2 \qquad \text{u.d.N.} \quad x_1 + x_2 - 1 = 0, \tag{1}$$

$$\min \quad x_1^2 - \frac{1}{2}x_2^2 + \frac{\eta}{2}(x_1 + x_2 - 1)^2 \qquad \text{u.d.N.} \quad x_1 + x_2 - 1 = 0. \tag{2}$$

(a) Berechnen Sie den KKT-Punkt $(\bar{x}, \bar{\mu})$ des Problems (1).

(b) Finden Sie ein $\bar{\eta} > 0$, so dass für alle $\eta > \bar{\eta}$ der in (a) berechnete KKT-Punkt globales unrestringiertes Minimum der Lagrangefunktion

$$L_a(\cdot, \bar{\mu}, \eta) := x_1^2 - \frac{1}{2}x_2^2 + \frac{\eta}{2}(x_1 + x_2 - 1)^2 + \bar{\mu}(x_1 + x_2 - 1)$$

des Problems (2) ist, wobei der Multiplikator $\bar{\mu}$ fix ist.

Lösung:

(a) Die Lagrangefunktion von (1) ist

$$L(x, \mu) = x_1^2 - \frac{1}{2}x_2^2 + \mu(x_1 + x_2 - 1).$$

Damit lauten die KKT-Bedingungen

$$\begin{aligned} 0 &= 2x_1 + \mu \\ 0 &= -x_2 + \mu \\ 0 &= x_1 + x_2 - 1 \end{aligned}$$

Hieraus ergibt sich die eindeutige Lösung $\bar{x} = (-1, 2)$ und $\bar{\mu} = 2$.
Der KKT-Punkt $\bar{x} = (-1, 2)$, $\bar{\mu} = 2$, ist ein Minimum, weil die Hessematrix der Lagrangefunktion auf dem kritische Kegel positive definit ist.

$$\begin{aligned} \nabla_{xx} L(x, \mu) &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \\ T_K(x) &= \{d \in \mathbb{R}^2 \,|\, \nabla h(x)^\top d = 0\} \\ &= \{d \in \mathbb{R}^2 \,|\, d_1 + d_2 = 0\} \\ &= \{(d, -d) \,|\, d \in \mathbb{R}\} \end{aligned}$$

$$\begin{pmatrix} d \\ -d \end{pmatrix}^\top \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} d \\ -d \end{pmatrix} = d^2 > 0 \quad \forall d \neq 0$$

(b) Setzen wir $\bar{\mu} = 2$ in die Lagrangefunktion des Problems (2) ein so erhalten wir die Funktion

$$L_a(x, 2, \eta) := x_1^2 - \frac{1}{2}x_2^2 + \frac{\eta}{2}(x_1 + x_2 - 1)^2 + 2(x_1 + x_2 - 1)$$

Im KKT-Punkt verschwindet der Gradient, denn es gilt

$$\nabla_x L_a(\bar{x}, 2, \eta) = \begin{pmatrix} 2\bar{x}_1 + \eta(\bar{x}_1 + \bar{x}_2 - 1) + 2 \\ -\bar{x}_2 + \eta(\bar{x}_1 + \bar{x}_2 - 1) + 2 \end{pmatrix} = \begin{pmatrix} 2(-1) + \eta \cdot 0 + 2 \\ -2 + \eta \cdot 0 + 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Für die Hessematrix erhalten wir

$$\nabla^2_{xx} L_a(x, 2, \eta) = \begin{bmatrix} 2+\eta & \eta \\ \eta & -1+\eta \end{bmatrix}.$$

- Für diese symmetrische Matrix berechnen wir die Eigenwerte mittels

  $$0 = (2 + \eta - \lambda)(-1 + \eta - \lambda) - \eta^2 = \lambda^2 - \lambda(2\eta + 1) + (\eta - 2),$$

  also

  $$\lambda_{1,2} = \frac{1}{2}(2\eta + 1 \pm \sqrt{(2\eta+1)^2 - 4(\eta - 2)}) = \frac{1}{2}(2\eta + 1 \pm \sqrt{4\eta^2 + 9}).$$

  Wir erkennen, dass für $\eta > 2$ stets

  $$(2\eta + 1)^2 = 4\eta^2 + 4\eta + 1 > 4\eta^2 + 9$$

  gilt, und somit sind die Eigenwerte der Hessematrix für alle $\eta > 2$ positiv und damit die Matrix positiv definit.
- Sylvester's criterion:

  $$\begin{aligned} \det \begin{bmatrix} 2+\eta \end{bmatrix} &= 2 + \eta > 0 & &\Rightarrow \eta > -2 \\ \det \begin{bmatrix} 2+\eta & \eta \\ \eta & -1+\eta \end{bmatrix} &= (2+\eta)(-1+\eta) - \eta^2 & & \\ &= \eta - 2 > 0 & &\Rightarrow \eta > 2 \\ \Rightarrow \quad & \eta > 2 & & \end{aligned}$$

Damit ist $L_a(x, 2, \eta)$ nach Satz 3.2.3 für alle $\eta > 2$ strikt konvex und folglich ist nach Satz 3.2.2 der stationäre Punkt eindeutiges globales Minimum.

Bemerkung: Die Existenz eines $\bar{\eta} > 0$ mit der Eigenschaft, dass KKT-Punkte von (1) strikte lokale Minima der erweiterten Lagrangefunktion $L_a(x, \bar{\mu}, \eta)$ sind, folgt aus Hilfssatz 5.6.1. Dieser sagt jedoch nichts über die Größe von $\bar{\eta}$ aus.

2. Gegeben sei das restringierte Optimierungsproblem

$$\min \quad x_1 + x_2 \qquad \text{u.d.N.} \quad x_1^2 - x_2 = 0.$$

Berechnen Sie mit

(a) dem Multiplikator-Penalty-Verfahren

(b) dem Lagrange-Newton-Verfahren

für $x^{[0]} = 0$, $\mu^{[0]} = \alpha$ und $\eta^{[0]} = 1$ die Lösung des Problems.

Welchen Einfluss hat der Startwert $\mu^{[0]}$? Und $\eta^{[0]}$?

Lösung:

$f(x) := x_1 + x_2$, $h(x) = x_1^2 - x_2$. Mit der Lagrangefunktion

$$L(x, \mu) = f(x) + \mu^\top h(x) = x_1 + x_2 + \mu(x_1^2 - x_2)$$

und mit der erweiterte Lagrangefunktion

$$L_a(x, \mu, \eta) = f(x) + \frac{\eta}{2}\|h(x)\|^2 + \mu^\top h(x)$$

erhalten wir

$$\nabla f(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \nabla h(x) = \begin{pmatrix} 2x_1 \\ -1 \end{pmatrix},$$

$$\nabla_x L(x, \mu) = \begin{pmatrix} 1 + 2\mu x_1 \\ 1 - \mu \end{pmatrix}, \qquad \nabla^2_{xx} L(x, \mu) = \begin{bmatrix} 2\mu & 0 \\ 0 & 0 \end{bmatrix},$$

$$\nabla_x L_a(x, \mu, \eta) = \begin{pmatrix} 1 + 2x_1\mu + 2x_1\eta(x_1^2 - x_2) \\ 1 - \mu - \eta(x_1^2 - x_2) \end{pmatrix}.$$

(a) Damit gehen wir in das Multiplikator-Penalty-Verfahren.

$$\begin{aligned} x^{[k+1]} &\leftarrow \arg\min_x L_a(x, \mu^{[k]}, \eta^{[k]}) \\ &\Downarrow \\ 0 &= \nabla_x L_a(x^{[k+1]}, \mu^{[k]}, \eta^{[k]}) \\ &\Downarrow \\ \begin{pmatrix} x_1^{[k+1]} \\ x_2^{[k+1]} \end{pmatrix} &= \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{4} + \frac{\mu^{[k]} - 1}{\eta^{[k]}} \end{pmatrix} \\ \mu^{[k+1]} &= \mu^{[k]} + \eta^{[k]} h(x^{[k+1]}) \end{aligned}$$

(k=0) $x_1 = 0$, $x_2 = 0$, $h(x) = 0$, $\mu = \alpha$, $\eta = 1$.
(k=1) $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{4} + \alpha - 1$, $h(x) = 1 - \alpha$, $\mu = 1$, $\eta = 10$.
(k=2) $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{4}$, $h(x) = 0$, $\mu = 1$, $\eta = 10$.
(k=3) $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{4}$, $h(x) = 0$, $\mu = 1$, $\eta = 10$. Dies ist bereits die Lösung!

(b) Damit gehen wir in das Lagrange-Newton-Verfahren.

$$0 = \begin{pmatrix} \nabla_x L(x, \mu) \\ h(x) \end{pmatrix} = \begin{pmatrix} 1 + 2\mu x_1 \\ 1 - \mu \\ x_1^2 - x_2 \end{pmatrix} \qquad \Rightarrow \begin{pmatrix} x_1^\star \\ x_2^\star \\ \mu^\star \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 1 \end{pmatrix}$$

Newton-Verfahren: Lineares Gleichungssystem, Lösung, und Neue Iterierte.

$$\begin{bmatrix} \nabla^2_{xx} L(x^{[k]}, \mu^{[k]}) & \nabla h(x^{[k]})^\top \\ \nabla h(x^{[k]}) & 0 \end{bmatrix} \begin{pmatrix} d^{[k]} \\ v^{[k]} \end{pmatrix} = -\begin{pmatrix} \nabla_x L(x^{[k]}, \mu^{[k]}) \\ h(x^{[k]}) \end{pmatrix}$$

Wir haben

$$\underbrace{\begin{bmatrix} 2\mu^{[k]} & 0 & 2x_1^{[k]} \\ 0 & 0 & -1 \\ 2x_1^{[k]} & -1 & 0 \end{bmatrix}}_{=:K^{[k]}} \begin{pmatrix} d_1^{[k]} \\ d_2^{[k]} \\ v^{[k]} \end{pmatrix} = - \begin{pmatrix} 1 + 2\mu^{[k]} x_1^{[k]} \\ 1 - \mu^{[k]} \\ (x_1^{[k]})^2 - x_2^{[k]} \end{pmatrix}$$

Determinant $\det K^{[k]} = -2\mu^{[k]}$. $\mu^{[k]} \neq 0$ gilt $K^{[k]}$ invertierbar und

$$\begin{pmatrix} d_1^{[k]} \\ d_2^{[k]} \\ v^{[k]} \end{pmatrix} = \begin{pmatrix} -\frac{2x_1^{[k]}+1}{2\mu^{[k]}} \\ x_2^{[k]} - (x_1^{[k]})^2 - \frac{2x_1^{[k]}+1}{\mu^{[k]}} x_1^{[k]} \\ 1 - \mu[k] \end{pmatrix}$$

(k=0) $x_1 = 0$, $x_2 = 0$, $\mu = \alpha \neq 0$. $d_1 = -\frac{1}{2\alpha}$, $d_2 = 0$, $v = 1 - \alpha$.
(k=1) $x_1 = -\frac{1}{2\alpha}$, $x_2 = 0$, $\mu = 1$. $d_1 = -\frac{1}{2} + \frac{1}{2\alpha}$, $d_2 = \frac{1}{2\alpha} - \frac{1}{4\alpha^2}$, $v = 0$.
(k=2) $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{2\alpha} - \frac{1}{4\alpha^2}$, $\mu = 1$. $d_1 = 0$, $d_2 = \frac{1}{4} - \frac{1}{2\alpha} + \frac{1}{4\alpha^2}$, $v = 0$.
(k=3) $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{4}$, $\mu = 1$. $d_1 = 0$, $d_2 = 0$, $v = 0$. Dies ist bereits die Lösung!

3. Gegeben sei das restringierte Optimierungsproblem

$$\min \quad (2x_1 - x_2 - 1)^2 + x_1 x_2 \qquad \text{u.d.N.} \quad x_1 + x_2 = 3.$$

Berechnen Sie (4 Iterierte) mit

(a) dem Multiplikator-Penalty-Verfahren
(b) dem Lagrange-Newton-Verfahren

für $x^{[0]} = 0$, $\mu^{[0]} = 0$ und $\eta^{[0]} = 1$ die Lösung des Problems.

Lösung:

$f(x) := (2x_1 - x_2 - 1)^2 + x_1 x_2$, $h(x) = 3 - x_1 - x_2$. Mit der Lagrangefunktion

$$L(x, \mu) = f(x) + \mu^\top h(x) = (2x_1 - x_2 - 1)^2 + x_1 x_2 + \mu(3 - x_1 - x_2)$$

und mit der erweiterte Lagrangefunktion

$$L_a(x, \mu, \eta) = f(x) + \frac{\eta}{2} \|h(x)\|^2 + \mu^\top h(x)$$

erhalten wir

$$\nabla f(x) = \begin{pmatrix} 8x_1 - 3x_2 - 4 \\ -3x_1 + 2x_2 + 2 \end{pmatrix}, \qquad \nabla h(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix},$$

$$\nabla_x L(x, \mu) = \begin{pmatrix} 8x_1 - 3x_2 - 4 - \mu \\ -3x_1 + 2x_2 + 2 - \mu \end{pmatrix}, \qquad \nabla^2_{xx} L(x, \mu) = \begin{bmatrix} 8 & -3 \\ -3 & 2 \end{bmatrix},$$

$$\nabla_x L_a(x, \mu, \eta) = \begin{pmatrix} 8x_1 - 3x_2 - 4 + \eta(x_1 + x_2 - 3) - \mu \\ -3x_1 + 2x_2 + 2 + \eta(x_1 + x_2 - 3) - \mu \end{pmatrix}.$$

(a) Damit gehen wir in das Multiplikator-Penalty-Verfahren.

$$
\begin{aligned}
x^{[k+1]} &\leftarrow \arg\min_x L_a(x, \mu^{[k]}, \eta^{[k]}) \\
&\Downarrow \\
0 &= \nabla_x L_a(x^{[k+1]}, \mu^{[k]}, \eta^{[k]}) \\
&= \begin{bmatrix} 8+\eta^{[k]} & \eta^{[k]}-3 \\ \eta^{[k]}-3 & 2+\eta^{[k]} \end{bmatrix} x^{[k+1]} - \begin{pmatrix} 3\eta^{[k]}+4+\mu^{[k]} \\ 3\eta^{[k]}-2+\mu^{[k]} \end{pmatrix} \\
&\Downarrow \langle \eta^{[k]} \geq 0 \rangle \\
\begin{pmatrix} x_1^{[k+1]} \\ x_2^{[k+1]} \end{pmatrix} &= \frac{1}{16\eta^{[k]}+7} \begin{bmatrix} 2+\eta^{[k]} & 3-\eta^{[k]} \\ 3-\eta^{[k]} & 8+\eta^{[k]} \end{bmatrix} \begin{pmatrix} 3\eta^{[k]}+4+\mu^{[k]} \\ 3\eta^{[k]}-2+\mu^{[k]} \end{pmatrix}
\end{aligned}
$$

(k=0) $x_1 = 0$, $x_2 = 0$, $h(x) = 3$, $\mu = 0$, $\eta = 1$.

(k=1) $x_1 = 1$, $x_2 = 1$, $h(x) = 1$, $\mu = 1$, $\eta = 10$.

(k=2) $x_1 \approx 1.29940$, $x_2 \approx 1.65868$, $h(x) \approx 0.041916$, $\mu \approx 1.4192$, $\eta = 10$.

(k=3) $x_1 \approx 1.31195$, $x_2 \approx 1.68629$, $h(x) \approx 0.001757$, $\mu \approx 1.4367$, $\eta = 10$.

(k=4) $x_1 \approx 1.31248$, $x_2 \approx 1.68745$, $h(x) \approx 0.000073$, $\mu \approx 1.4375$, $\eta = 10$.

(k=5) $x_1 \approx 1.31250$, $x_2 \approx 1.68750$, $h(x) \approx 0.000003$, $\mu \approx 1.4375$, $\eta = 10$.

(b) Damit gehen wir in das Lagrange-Newton-Verfahren.

$$
0 = \begin{pmatrix} \nabla_x L(x,\mu) \\ h(x) \end{pmatrix} = \begin{pmatrix} 8x_1 - 3x_2 - 4 - \mu \\ -3x_1 + 2x_2 + 2 - \mu \\ 3 - x_1 - x_2 \end{pmatrix} = \begin{bmatrix} 8 & -3 & -1 \\ -3 & 2 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \mu \end{pmatrix} - \begin{pmatrix} 4 \\ -2 \\ -3 \end{pmatrix}
$$

$$\Downarrow$$

$$
\begin{bmatrix} 8 & -3 & -1 \\ -3 & 2 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \mu \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \\ -3 \end{pmatrix}
$$

$$\Downarrow$$

$$
\begin{pmatrix} x_1^\star \\ x_2^\star \\ \mu^\star \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 21 \\ 27 \\ 23 \end{pmatrix} \approx \begin{pmatrix} 1.3125 \\ 1.6875 \\ 1.4375 \end{pmatrix}
$$