Grundlegende Variationelle Methoden mit Anwendungen aus der Bildverarbeitung

Stephen Keeling
Institut für Mathematik
Universität Graz, Graz, Austria

Literatur:

P.E. Gill, W. Murray, M.H. Wright, *Practical Optimization.*

H. Amann, J. Escher, *Analysis II.*

I.M. Gelfand, S.V. Fomin, *Calculus of Variations.*

D.G. Luenberger, *Optimization by Vector Space Methods.*

Beispiel: Sei $f \in C^2(\boldsymbol{R}^2, \boldsymbol{R})$ und

$$\min_{\boldsymbol{x} \in \boldsymbol{R}^2} f(\boldsymbol{x}), \quad f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2 = x_1^2 + x_2^2.$$

Notwendige Optimalitätsbedingung:

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} f_{x_1} \\ f_{x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \qquad \partial f(\boldsymbol{x})[\boldsymbol{h}] = \nabla f(\boldsymbol{x}) \cdot \boldsymbol{h}$$

$$0 = \partial f(\boldsymbol{x}^\star)[\boldsymbol{h}] = 2\boldsymbol{x}^\star \cdot \boldsymbol{h}, \ \forall \boldsymbol{h} \in \boldsymbol{R}^2 \ \Rightarrow \ \boldsymbol{x}^\star = \boldsymbol{0}$$

Hinreichende Optimalitätsbedingung:

$$H_f(\boldsymbol{x}) = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \partial^2 f(\boldsymbol{x})[\boldsymbol{h}, \boldsymbol{k}] = \boldsymbol{h} \cdot H_f(\boldsymbol{x}) \cdot \boldsymbol{k}$$

$$\partial^2 f(\boldsymbol{x}^\star)[\boldsymbol{h}]^2 = \boldsymbol{h} \cdot [2I] \cdot \boldsymbol{h} = 2\|\boldsymbol{h}\|^2 > 0 \ \Rightarrow \ f(\boldsymbol{x}^\star) = \min.$$

Basis der Bedingungen und Lösungsmethoden.

Sei $J \in C^2(H, \boldsymbol{R})$ und

$$\min_{u \in H} J(u).$$

Nimm an: $J(u^\star) = \min_{u \in H} J(u)$. Taylorsatz: $v \in H$,

$$J(u^\star + v) = J(u^\star) + \partial J(u^\star)[v] + \tfrac{1}{2}\partial^2 J(u^\star)[v]^2 + \sigma \|v\|_H^2$$

wobei $\sigma \to 0$ für $\|v\|_H \to 0$.

Nimm an:

$$\partial J(u^\star)[-\hat{v}] < 0 < \partial J(u^\star)[\hat{v}].$$

Wenn $|\varepsilon|$ klein genug ist, gilt

$$\varepsilon^{-1}[J(u^\star + \varepsilon\hat{v}) - J(u^\star)] = \partial J(u^\star)[\hat{v}] + \sigma|\varepsilon| > 0.$$

Dann folgt Widerspruch:

$$\begin{array}{l} \varepsilon > 0 \ \Rightarrow \ J(u^\star + \varepsilon\hat{v}) > J(u^\star) \\ \varepsilon < 0 \ \Rightarrow \ J(u^\star + \varepsilon\hat{v}) < J(u^\star). \end{array}$$

Also notwendige (1.Ordnung) Optimalitätsbedingungen:

$$\partial J(u^\star)[v] = D_v J(u^\star) = 0, \ \forall v \in H.$$

Nimm an:

$$\partial^2 J(u^\star)[\hat{v}]^2 < 0.$$

Da $\partial J(u^\star)[v] = 0$ gilt, folgt:

$$J(u^\star + v) = J(u^\star) + \tfrac{1}{2}\partial^2 J(u^\star)[v]^2 + \sigma \|v\|_H^2.$$

Wenn $|\varepsilon|$ klein genug ist, gilt

$$2\varepsilon^{-2}[J(u^\star + \varepsilon\hat{v}) - J(u^\star)] = \partial^2 J(u^\star)[\hat{v}]^2 + \sigma|\varepsilon| < 0.$$

Dann folgt Widerspruch:

$$J(u^\star + \varepsilon \hat{v}) < J(u^\star), \ \forall \varepsilon \text{ klein genug.}$$

Also notwendige (2.Ordnung) Optimalitätsbedingungen:

$$\partial^2 J(u^\star)[v]^2 \geq 0, \ \forall v \in H.$$

Nimm an: (Koerzivität)

$$\exists \alpha > 0 \text{ s.d. } \partial^2 J(u^\star)[v]^2 \geq \alpha \|v\|_H^2, \ \forall v \in H.$$

Da $\partial J(u^\star)[v] = 0$ gilt, folgt:

$$J(u^\star + v) = J(u^\star) + \tfrac{1}{2}\partial^2 J(u^\star)[v]^2 + \sigma \|v\|_H^2$$

Nimm $\|v\|_H$ klein genug, s.d. $|\sigma| \leq \alpha/4$,

$$J(u^\star + v) - J(u^\star) \geq \tfrac{1}{2}\alpha \|v\|_H^2 - \tfrac{1}{4}\alpha \|v\|_H^2 > 0.$$

Also hinreichende (2.Ordnung) Optimalitätsbedingungen:

$$\partial^2 J(u^\star)[v]^2 \geq \alpha \|v\|_H^2, \ \forall v \in H.$$

Äquivalent in endlich dimensionalen Räumen:

$$\partial^2 J(u^\star)[v]^2 > 0, \ \forall v \in H.$$

Existenz und Eindeutigkeit.

Sei $J : Y \subset X \to \boldsymbol{R}$

- koerziv auf Y, oder $J(u) \to \infty$ für $\|u\|_Y \to \infty$, und
- von unten halbstetig auf X, oder $J(u) \leq \liminf\limits_{n\to\infty} J(u_n)$ wenn $\|u - u_n\|_X \to 0$.

Wenn Y kompakt eingebettet in X wird, gibt es ein Minimum $u^\star \in X$ für J. Wenn J streng konvex ist, ist das Minimum eindeutig.

Minimierende Folge $\{u_n\} \subset Y$.

- Y-Koerzivität $\Rightarrow \{u_n\}$ Y-beschränkt.
- $Y \overset{k}{\hookrightarrow} X \Rightarrow \exists \{u_{n'}\}$ und $u^\star \in X$, so daß $u_n \xrightarrow{X} u^\star$.
- Halbstetigkeit von unten $\Rightarrow J(u^\star) \le J(u_{n'})$.
- Strenge Konvexität $\Rightarrow J[(u_1^\star + u_2^\star)/2] < \frac{1}{2}[J(u_1^\star) + J(u_2^\star)] = J_{\min}$.

Zusammenfassung der Definitionen.

Richtungsableitung:

$$D_v J(u) = \lim_{\varepsilon \to 0} \frac{J(u + \varepsilon v) - J(u)}{\varepsilon}$$

Ableitung: $\partial J(u) \in \mathcal{L}(H, \boldsymbol{R})$ (linear, stetig/beschränkt)

$$\lim_{\|v\|_H \to 0} \frac{|J(u+v) - J(u) - \partial J(u)[v]|}{\|v\|_H} = 0.$$

Satz: J differenzierbar in $u^\star \Rightarrow D_v J(u^\star)$ existieren $\forall v \in H$ und $D_v J(u^\star) = \partial J(u^\star)[v]$.

Satz: $\partial J(u^\star) \in \mathcal{L}(H, \boldsymbol{R})$ (H Hilbert) $\Rightarrow \exists g \in H$ (Rieszsche Darstellung) s.d. $D_v J(u^\star) = \partial J(u^\star)[v] = \langle g, v \rangle_H$. (“$g = \nabla J$”)

Höhere Ableitungen:

$$\partial^m J(u) \in \mathcal{L}^m(H, \boldsymbol{R}), \qquad \mathcal{L}^m = \mathcal{L}(H, \mathcal{L}^{m-1}(H, \boldsymbol{R})),$$

$$\lim_{\|v\|_H \to 0} \frac{\|\partial^{m-1} J(u+v) - \partial^{m-1} J(u) - \partial^m J(u)[v]\|_{\mathcal{L}^{m-1}(H,\boldsymbol{R})}}{\|v\|_H} = 0.$$

$$\|A\|_{\mathcal{L}(B,C)} = \sup_{u \in B} \frac{\|Au\|_C}{\|u\|_B}$$

$$\partial^m J(u)[v_1, \ldots, v_m] = \partial(\cdots(\partial(\partial J(u)[v_1])[v_2]\cdots)[v_m]$$

Taylorsatz: $u, v \in H$,

$$J(u+v) = \sum_{k=0}^{m} \frac{1}{k!} \partial^k J(u)[v]^k + \sigma \|v\|_H^m$$

wobei $\sigma \to 0$ für $\|v\|_H \to 0$.

Beispiel. Finde die Kurve $u^\star$ mit kleinster Länge zwischen zwei Punkten.

Eingeschränktes Problem:

$$\min_{w \in C^1[0,1]} J(w) \quad \text{s.d.} \quad w(0) = 0, \; w(1) = 1.$$

wobei die Länge so gegeben wird:

$$J(w) = \int_0^1 \sqrt{1 + [w'(x)]^2} dx.$$

Die Menge $\{w \in C^1[0,1] : w(0) = 0, w(1) = 1\}$ ist kein Vektorraum. Randbedingungen sind Einschränkungen – später.

Nicht eingeschränktes Problem: $w = x + u$, $u(0) = 0 = u(1)$

$$\min_{u \in C_0^1[0,1]} J(u), \quad J(u) = \int_0^1 \sqrt{1 + [1 + u'(x)]^2} dx.$$

Die Menge $C_0^1[0,1] = \{u \in C^1[0,1] : u(0) = 0, u(1) = 0\}$ *ist* ein (Banach) Vektorraum.

Notwendige Optimalitätsbedingung: $D_v J(u^\star) = 0, \forall v \in C_0^1[0,1]$.

$$
\begin{aligned}
D_v J(u) &= \lim_{\varepsilon\to 0} \frac{J(u+\varepsilon v)-J(u)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \int_0^1 \frac{\sqrt{1+[1+u'(x)+\varepsilon v'(x)]^2}-\sqrt{1+[1+u'(x)]^2}}{\varepsilon} dx
\end{aligned}
$$

$$
= \cdots = \int_0^1 \frac{[1+u'(x)]v'(x)}{\sqrt{1+[1+u'(x)]^2}} dx.
$$

Wenn $u^\star$ genügend glatt ist, gilt

$$
\begin{aligned}
D_v J(u^\star) &= \frac{1+u^{\star\prime}(x)}{\sqrt{1+[1+u^{\star\prime}(x)]^2}} \; \underbrace{v(x)}_{=0} \Bigg|_{x=0}^{x=1} \\
&- \int_0^1 v(x) \left(\frac{1+u^{\star\prime}(x)}{\sqrt{1+[1+u^{\star\prime}(x)]^2}} \right)' dx.
\end{aligned}
$$

Wenn $v(x) \to \delta(x-\hat{x})$, $\hat{x} \in (0,1)$, $D_v J(u^\star) = 0 \Rightarrow$

$$
\left(\frac{1+u^{\star\prime}(x)}{\sqrt{1+[1+u^{\star\prime}(x)]^2}} \right)' = 0, \;\; x \in (0,1).
$$

Die allgemeine Lösung:

$$
\left(\frac{1+u^{\star\prime}(x)}{\sqrt{1+[1+u^{\star\prime}(x)]^2}} \right) = k_1 \;\; \Rightarrow \;\; u^{\star\prime}(x) = k_2 \;\; \Rightarrow \;\; u^\star(x) = k_2 x + k_3
$$

Die Randbedingungen $u \in C_0^1[0,1] \Rightarrow u^\star(x) = 0$.

Das heisst, $w^\star(x) = x + u^\star(x) = x$ gibt eine Gerade.

Auch notwendig: $\partial J(u^\star)[v] = 0$, $\forall v \in C_0^1[0,1]$.

J differenzierbar $\Rightarrow \partial J(u^\star)[v] = D_v J(u^\star) = 0$, $\forall v \in C_0^1[0,1]$.

J differenzierbar?

$$\lim_{\|v\|_{C_0^1} \to 0} \frac{|J(u+v) - J(u) - \partial J(u)[v]|}{\|v\|_{C_0^1}} = 0?$$

wobei:

$$\|v\|_{C_0^1} = \sup_{x\in[0,1]} |v'(x)|, \quad \|v\|_{C^1} = \sup_{x\in[0,1]} |v(x)| + \sup_{x\in[0,1]} |v'(x)|$$

die auf $C_0^1[0,1]$ äquivalent sind.

Aufgabe: Zeige ja, und daher gilt $\partial J(u^\star)[v] = D_v J(u^\star) = 0$, $\forall v \in C_0^1[0,1]$.

Hinweis: Wenn $J(u) = \int F(u')$ gilt, folgt

$$J(u+v) - J(u) = \int [F(u'+v') - F(u')] = \int [F'(u')v' + \sigma|v'|].$$

wobei $\sigma \to 0$ für $|v'| \to 0$.

Auch notwendig: $\partial^2 J(u^\star)[v]^2 \geq 0$, $\forall v \in C_0^1[0,1]$.

$$\begin{aligned} D_w D_v J(u) &= \lim_{\varepsilon\to 0} \frac{D_v J(u+\varepsilon w) - D_v J(u)}{\varepsilon} \\ &= \cdots = \int_0^1 \frac{v'(x)w'(x)}{[1+[1+u'(x)]^2]^{\frac{3}{2}}} dx \end{aligned}$$

$J(u)$ zweimal differenzierbar $\Rightarrow \partial^2 J(u)[v,w] = D_w D_v J(u)$, $\forall v, w \in C_0^1[0,1]$.

$J(u)$ zweimal differenzierbar?

$$\lim_{\|v\|_{C_0^1} \to 0} \frac{\|\partial J(u+v) - \partial J(u) - \partial^2 J(u)[v]\|_{\mathcal{L}(C_0^1, \boldsymbol{R})}}{\|v\|_{C_0^1}} = 0?$$

Aufgabe: Zeige ja, und daher gilt $\partial^2 J(u)[v,w] = D_w D_v J(u)$, $\forall v, w \in C_0^1[0,1]$.

Daher folgt

$$\begin{aligned}\partial^2 J(u^\star)[v]^2 = D_v^2 J(u^\star) &= \int_0^1 \frac{v'(x)^2}{[1+[1+u'(x)]^2]^{\frac{3}{2}}} dx \\ &= 2^{-\frac{3}{2}} \int_0^1 [v'(x)]^2 dx \geq 0\end{aligned}$$

Hinreichend: $\partial^2 J(u^\star)[v]^2 \geq \alpha \|v\|_{C_0^1}^2$, $\forall v \in C_0^1[0,1]$.

$$\partial^2 J(u^\star)[v]^2 = 2^{-\frac{3}{2}} \int_0^1 [v'(x)]^2 dx \geq \alpha \|v\|_{C_0^1}^2 = \alpha \sup_{x \in [0,1]} |v'(x)|?$$

Nein! Aber es gilt

$$\partial^2 J(u^\star)[v]^2 \geq \alpha \|v\|_{H_0^1}^2, \ \forall v \in H_0^1[0,1] \quad (\alpha = 2^{-\frac{3}{2}})$$

wobei

$$\|v\|_{H_0^1}^2 = \int_0^1 [v'(x)]^2 dx \qquad \|v\|_{H^1}^2 = \int_0^1 \{[v(x)]^2 + [v'(x)]^2\} dx$$

die äquivalent auf $H_0^1[0,1]$ sind.

Also natürliche Frage: Warum nicht mit $H_0^1[0,1]$ statt $C_0^1[0,1]$ arbeiten?

Aufgabe: Zeige ja, alle obigen Behauptungen gelten ebensogut mit $H_0^1[0,1]$, z.B. gilt

$$\lim_{\|v\|_{H_0^1} \to 0} \frac{|J(u+v) - J(u) - \partial J(u)[v]|}{\|v\|_{H_0^1}} = 0.$$

Zusammenfassung:

- Notwendige $D_v J(u^\star) = 0$ liefert einen Kandidaten $u^\star$,
- Notwendige $D_v J(u^\star) = 0$ leicht zu zeigen,
- Notwendige $\partial J(u^\star)[v] = 0$ folgt, wenn J in $u^\star$ glatt ist,
- Notwendige $\partial^2 J(u^\star)[v]^2 \geq 0$ schwieriger zu zeigen,
- Hinreichende $\partial^2 J(u^\star)[v]^2 \geq \alpha \|v\|_H^2$ noch schwieriger zu zeigen.

Beispiel. Sei $\tilde{u} \in L^2[0,1]$ verrauscht und finde $u \in H^1[0,1]$ eine entrauschte Abschätzung.

Formulierung:

$$\min_{u \in H^1} J(u), \quad J(u) = \int_0^1 [u(x) - \tilde{u}(x)]^2 dx + \mu \int_0^1 [u'(x)]^2 dx$$

$$\mu \to 0 \Rightarrow u^\star \to \tilde{u}. \qquad \mu \to \infty \Rightarrow u^\star \to \int_0^1 \tilde{u}(x) dx.$$

Notwendige Optimalitätsbedingung: $D_v J(u^\star) = 0$, $\forall v \in H^1[0,1]$.

$$\begin{aligned} D_v J(u) &= \lim_{\varepsilon \to 0} \frac{J(u + \varepsilon v) - J(u)}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^1 \{[u(x) + \varepsilon v(x) - \tilde{u}(x)]^2 - [u(x) - \tilde{u}(x)]^2\} dx \\ &\qquad + \frac{\mu}{\varepsilon} \int_0^1 \{[u'(x) + \varepsilon v'(x)]^2 - [u'(x)]^2\} dx \\ = \cdots &= 2 \int_0^1 [u(x) - \tilde{u}(x)] v(x) dx + 2\mu \int_0^1 u'(x) v'(x) dx \end{aligned}$$

Wenn $u^\star$ genügend glatt ist, gilt

$$\begin{aligned} D_v J(u^\star) &= 2 \int_0^1 [u^\star(x) - \tilde{u}(x)] v(x) dx \\ &\quad + 2\mu u^{\star\prime}(x) v(x) \Big|_{x=0}^{x=1} - 2\mu \int_0^1 v(x) u^\star(x)'' dx. \end{aligned}$$

Wenn $v \to 1$ am Rand und 0 sonst, oder $v(x) \to \delta(x - \hat{x})$, $\hat{x} \in (0,1)$, $D_v J(u^\star) = 0 \Rightarrow$

$$\begin{cases} -\mu u^{\star\prime\prime} + u^\star &= \tilde{u},\ 0 < x < 1 \\ u^{\star\prime} &= 0,\ \ x = 0,1 \end{cases}$$

Charakterisierung von $u^\star$.

$$\mu \to 0 \Rightarrow u^\star \to \tilde{u}. \qquad \mu \to \infty \Rightarrow u^\star \to \int_0^1 \tilde{u}(x)dx.$$

Auch notwendig: $\partial J(u^\star)[v] = 0$, $\forall v \in H^1[0,1]$.

J differenzierbar $\Rightarrow \partial J(u^\star)[v] = D_v J(u^\star) = 0$, $\forall v \in H^1[0,1]$.

J differenzierbar?

$$\lim_{\|v\|_{H^1} \to 0} \frac{|J(u+v) - J(u) - \partial J(u)[v]|}{\|v\|_{H^1}} = 0?$$

wobei:

$$\|v\|_{H^1}^2 = \int_0^1 \{[v(x)]^2 + [v'(x)]^2\} dx$$

Aufgabe: Zeige ja, und daher gilt $\partial J(u^\star)[v] = D_v J(u^\star) = 0$, $\forall v \in H^1[0,1]$.

Auch notwendig: $\partial^2 J(u^\star)[v]^2 \geq 0$, $\forall v \in H^1[0,1]$.

$$\begin{aligned} D_w D_v J(u) &= \lim_{\varepsilon \to 0} \frac{D_v J(u + \varepsilon w) - D_v J(u)}{\varepsilon} \\ &= \cdots = 2 \int_0^1 w(x) v(x) dx + 2\mu \int_0^1 w'(x) v'(x) dx \end{aligned}$$

$J(u)$ zweimal differenzierbar $\Rightarrow \partial^2 J(u)[v,w] = D_w D_v J(u)$, $\forall v, w \in H^1[0,1]$.

$J(u)$ zweimal differenzierbar?

$$\lim_{\|v\|_{H^1} \to 0} \frac{\|\partial J(u+v) - \partial J(u) - \partial^2 J(u)[v]\|_{\mathcal{L}(H^1, \boldsymbol{R})}}{\|v\|_{H^1}} = 0?$$

Aufgabe: Zeige ja, und daher gilt $\partial^2 J(u)[v,w] = D_w D_v J(u)$, $\forall v, w \in H^1[0,1]$.

Daher folgt

$$\begin{aligned} \partial^2 J(u^\star)[v]^2 = D_v^2 J(u^\star) &= 2\int_0^1 [v(x)]^2 dx + 2\mu \int_0^1 [v'(x)]^2 dx \\ &\geq 2\min\{1,\mu\}\|v\|_{H^1}^2 \end{aligned}$$

Hinreichend! $\partial^2 J(u^\star)[v]^2 \geq \alpha \|v\|_{H^1}^2$, $\forall v \in H^1[0,1]$.

Numerische Methoden.

Diskretisiere

$$0 = \tfrac{1}{2} D_v J(u) = B(u,v) - F(v), \quad \forall v \in H^1[0,1]$$

$$\begin{aligned} B(u,v) &= \int_0^1 [u(x)v(x) + \mu u'(x)v'(x)]dx \\ F(v) &= \int_0^1 \tilde{u}(x)v(x)dx \end{aligned}$$

mit finiten Elementen (S_h = span von Dachfunktionen),

$$B(u_h, \chi) = F(\chi) \qquad \forall \chi \in S_h \subset H^1[0,1]$$

oder mit finiten Differenzen für

$$\begin{cases} -\mu u'' + u = \tilde{u}, \ 0 < x < 1 \\ \qquad\quad u' = 0, \ x = 0, 1 \end{cases}$$

oder minimiere ($h = 1/n$, $x_i = ih$, $U_i \approx u(x_{i-\frac{1}{2}})$)

$$J_h(U) = h \sum_{i=1}^{n} [U_i - \tilde{U}_i]^2 + h\mu \sum_{i=1}^{n-1} \left[\frac{U_{i+1} - U_i}{h}\right]^2$$

Optimiere dann diskretisiere, oder diskretisiere dann optimiere?

Fast niemals die selbe Diskretisierung! Wichtige Unterschiede im allgemeinen, aber hier keine!

Aufgabe: Zeige $\nabla J_h(U^\star) = 0 \Rightarrow$

$$[-\mu\Delta_h + I]\, U^\star = \tilde{U}, \qquad -\Delta_h = \frac{1}{h^2}\begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix}$$

und $U^\star \to u^\star$ für $n \to \infty$.

Beispiel. Sei $\tilde{u} \in L^2[0,1]$ verrauscht und finde $u \in H^1[0,1]$ eine entrauschte Abschätzung mit scharfen Kanten.

Quadratische Strafe zu dissipativ. Neue Formulierung:

$$\min_{u \in H^1} J(u), \quad J(u) = \int_0^1 [u(x) - \tilde{u}(x)]^2 dx + \int_0^1 \phi(u'(x)^2) dx$$

wobei $s \mapsto \phi(s^2)$ konvex ist.

Notwendige Optimalitätsbedingung: $D_v J(u^\star) = 0$, $\forall v \in H^1[0,1]$.

$$D_v J(u) = \lim_{\varepsilon \to 0} \frac{J(u + \varepsilon v) - J(u)}{\varepsilon}$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_0^1 \{[u(x) + \varepsilon v(x) - \tilde{u}(x)]^2 - [u(x) - \tilde{u}(x)]^2\} dx$$

$$+ \frac{1}{\varepsilon} \int_0^1 \{\phi([u'(x) + \varepsilon v'(x)]^2) - \phi([u'(x)]^2)\} dx$$

$$= 2 \int_0^1 [u(x) - \tilde{u}(x)] v(x) dx + 2 \int_0^1 \phi'(u'(x)^2) u'(x) v'(x) dx$$

Wenn $u^\star$ genügend glatt ist, gilt

$$\begin{aligned} D_v J(u^\star) &= 2\int_0^1 [u^\star(x) - \tilde{u}(x)] v(x) dx \\ &+ 2\phi'(u^{\star\prime}(x)^2) u^{\star\prime}(x) v(x)\Big|_{x=0}^{x=1} \\ &- 2\mu \int_0^1 v(x) [\phi'(u^{\star\prime}(x)^2) u^{\star\prime}(x)]' dx. \end{aligned}$$

Wenn $v \to 1$ am Rand und 0 sonst, oder $v(x) \to \delta(x - \hat{x})$, $\hat{x} \in (0,1)$, $D_v J(u^\star) = 0 \Rightarrow$

$$\begin{cases} -[\phi'(u^{\star\prime 2}) u^{\star\prime}]' + u^\star = \tilde{u}, \ 0 < x < 1 \\ \qquad\qquad\qquad\quad u^{\star\prime} = 0, \ \ x = 0, 1 \end{cases}$$

Charakterisierung von $u^\star$.

Lösungsmethoden.

Lagged Diffusivity oder Fixpunkt Iteration.

$$\begin{cases} -[\phi'(u_l'^2) u_{l+1}']' + u_{l+1} = \tilde{u}, \ 0 < x < 1 \\ \qquad\qquad\qquad\qquad u_{l+1}' = 0, \ \ x = 0, 1 \end{cases}$$

Newtons Verfahren.

Hintergrund: löse $F(\boldsymbol{x}) = \boldsymbol{0}$ oder $\boldsymbol{h} \cdot F(\boldsymbol{x}) = 0$, $\forall \boldsymbol{h} \in \boldsymbol{R}^n$.
Newtons Iteration: $\nabla F(\boldsymbol{x}_k)[\boldsymbol{x}_{k+1} - \boldsymbol{x}_k] = -F(\boldsymbol{x}_k)$
oder $\boldsymbol{h} \cdot \nabla F(\boldsymbol{x}_k) d\boldsymbol{x}_k = -\boldsymbol{h} \cdot F(\boldsymbol{x}_k)$, $\forall \boldsymbol{h} \in \boldsymbol{R}^n$.

Gegenstück im Funktionenraum:

$$\partial J(u)[v] = 0, \qquad \forall v \in H$$

$$\partial^2 J(u_k)[v, du_k] = -\partial J(u_k)[v], \qquad \forall v \in H$$

Aufgabe: Zeige $J(u)$ zweimal differenzierbar und daher gilt $\partial^2 J(u)[v,w] = D_w D_v J(u)$, $\forall v, w \in H^1[0,1]$.

$$D_w D_v J(u) = \lim_{\varepsilon \to 0} \frac{D_v J(u+\varepsilon w) - D_v J(u)}{\varepsilon} = \cdots$$

$$= \int_0^1 \{[2vw] + \phi'(u'^2)[2v'w'] + \phi''(u'^2)[2u'v'][2u'w']\}dx$$

Newtons Verfahren $\partial^2 J(u_k)[v, du_k] = -\partial J(u_k)[v]$ wird:

$$\int_0^1 \{[2vdu_k] + \phi'(u_k'^2)[2v'du_k'] + \phi''(u_k'^2)[2u_k'v'][2u_k'du_k']\}dx$$

$$= -\int_0^1 \{2(u_k - \tilde{u})v + 2\phi'(u_k'^2)u_k'v'\}dx$$

Aufgabe: Wenn u_k und u_{k+1} genügend glatt sind, verwende partielle Integration, wähle v strategisch, und zeige

$$\left\{\begin{array}{rcll} -[2\phi''(u_k'^2)u_k'^2 u_{k+1}']' & + & [2\phi''(u_k'^2)u_k'^3]' & \\ -[\phi'(u_k'^2)u_{k+1}']' + u_{k+1} & = & \tilde{u}, & 0 < x < 1 \\ u_{k+1}' & = & 0, & x = 0, 1 \end{array}\right.$$

Fixpunkt Iteration ohne erste Zeile.

Mögliche numerischen Diskretisierungen wie vorher.

Wenn $s \mapsto \phi(s^2)$ nicht konvex, stationäre Lösungen nicht notwendigerweise minimierend oder eindeutig.

Minimierung unter linearen Einschränkungen.

Beispiel. Finde den kleinste Norm Punkt in einer Ebene.

Formulierung: mache L stationär

$$L(\boldsymbol{x}, \lambda) = J(\boldsymbol{x}) + \langle \lambda, E(\boldsymbol{x}) \rangle = \|\boldsymbol{x}\|^2 + \lambda(\boldsymbol{a} \cdot \boldsymbol{x} - b)$$

Notwendige Optimalitätsbedingung:

$$\left.\begin{array}{r} \nabla_{\boldsymbol{x}} L(\boldsymbol{x},\lambda) = 2\boldsymbol{x} + \lambda \boldsymbol{a} = 0 \\ \nabla_{\lambda} L(\boldsymbol{x},\lambda) = \boldsymbol{a}\cdot\boldsymbol{x} - b = 0 \end{array}\right\} \begin{array}{l} \lambda = -2b/\|\boldsymbol{a}\|^2 \\ \boldsymbol{x}^\star = b\boldsymbol{a}/\|\boldsymbol{a}\|^2 \end{array}$$

Hinreichende Optimalitätsbedingung:

$$\partial^2 J(\boldsymbol{x})[\boldsymbol{y}]^2 \geq \alpha \|\boldsymbol{y}\|^2, \qquad \forall \boldsymbol{y} \text{ s.d. } \boldsymbol{a}\cdot\boldsymbol{y} = b$$

$$\partial^2 J(\boldsymbol{x})[\boldsymbol{y}]^2 = 2\|\boldsymbol{y}\|^2 \quad \Rightarrow \quad \alpha = 2$$

Basis der Bedingungen und Lösungsmethoden.

Seien $J \in C^2(H, \boldsymbol{R})$, $A \in \mathcal{L}(H, \Lambda)$ und

$$\min_{u\in H} J(u) \text{ s.d. } Au = r$$

Nimm an, daß $Au = r$ keine eindeutige Lösung hat: $\mathcal{N}(A) \neq \{0\}$.
Nimm an, daß $A \in \mathcal{L}(H,\Lambda)$ surjektiv ist: $\mathcal{N}(A)^\perp = \mathcal{R}(A^*)$.

$$\langle \lambda, Av\rangle_\Lambda = \langle A^*\lambda, v\rangle_H, \qquad \forall \lambda \in \Lambda, \forall v \in H.$$

Störungen p, $u_2 = u_1 + p$, in $\{u : Au = r\}$ erfüllen $Ap = 0$.

Nimm an: $J(u^\star) = \min_{u\in H} J(u)$ s.d. $Au = r$. Taylorsatz: $p \in \mathcal{N}(A)$,

$$J(u^\star + p) = J(u^\star) + \partial J(u^\star)[p] + \tfrac{1}{2}\partial^2 J(u^\star)[p]^2 + \sigma \|p\|_H^2$$

wobei $\sigma \to 0$ für $\|p\|_H \to 0$.

Wie vorher, sind die folgenden Bedingungen notwendig:

$$\begin{array}{rcll} \partial J(u^\star)[p] & = & 0, & \forall p \in \mathcal{N}(A) \\ \partial^2 J(u^\star)[p]^2 & \geq & 0, & \forall p \in \mathcal{N}(A) \end{array}$$

und die folgende hinreichend:

$$\exists \alpha > 0 \text{ s.d. } \partial^2 J(u^\star)[p]^2 \geq \alpha \|p\|_H^2, \qquad \forall p \in \mathcal{N}(A)$$

Die Eigenschaft $\partial J(u^\star) \in \boldsymbol{\mathcal{L}}(H, \boldsymbol{R})$ gibt die Rieszsche Darstellung: $\exists g \in H$ (“$g = \nabla J$”) s.d.

$$\partial J(u^\star)[v] = \langle g, v \rangle_H, \qquad \forall v \in H$$

Also ist “der Gradient” der Einschränkungsmenge orthogonal:

$$0 = \partial J(u^\star)[p] = \langle g, p \rangle_H, \qquad \forall p \in \mathcal{N}(A)$$

und $g \in \mathcal{N}(A)^\perp = \mathcal{R}(A^*)$.

Lagrangescher Multiplikator Satz: $\exists \lambda^\star \in \Lambda$ s.d. $g = -A^* \lambda^\star$.

$$\partial J(u^\star)[v] = \langle g, v \rangle_H = -\langle A^* \lambda^\star, v \rangle_H = -\langle \lambda^\star, Av \rangle_\Lambda$$

Notwendige Optimalitätsbedingung: $(u^\star, \lambda^\star)$ stationär für die Lagrangesche Funktion

$$L(u, \lambda) = J(u) + \langle \lambda, Au - r \rangle_\Lambda$$

oder

$$\begin{array}{lll} \forall v \in H, & 0 = \partial_u L(u^\star, \lambda^\star)[v] & = \partial J(u^\star)[v] + \langle \lambda^\star, Av \rangle_\Lambda \\ \forall \kappa \in \Lambda, & 0 = \partial_\lambda L(u^\star, \lambda^\star)[\kappa] & = \langle \kappa, Av - r \rangle_\Lambda \end{array}$$

Die Räume hier sind Hilberträume - mit einem Innenprodukt ausgestattet. Es gibt ein ähnliches Argument, wenn die nur Banachräume sind - ohne ein Innenprodukt. Dann gelten $g \in H^*$ und $A \in \boldsymbol{\mathcal{L}}(H, \Lambda^*)$, wobei $X^* = \boldsymbol{\mathcal{L}}(X, \boldsymbol{R})$ definiert wird.

Beispiel: Finde die Kurve $u^\star$ mit kleinster Länge zwischen zwei Punkten.

Eingeschränktes Problem:

$$\min_{u \in C^1[0,1]} J(u) \quad \text{s.d.} \quad u(0) = 0, \; u(1) = 1.$$

wobei:

$$J(u) = \int_0^1 \sqrt{1+[u'(x)]^2}dx, \qquad Au = \begin{bmatrix} u(0) \\ u(1) \end{bmatrix}, \quad r = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$A \in \boldsymbol{\mathcal{L}}(C^1[0,1], \boldsymbol{R}^2)$

$$\|Au\|_{\boldsymbol{R}^2} \leq 2 \max_{x=0,1} |u(x)| \leq 2\|u\|_{C^1}$$

ist surjektiv:

$$\forall r \in \boldsymbol{R}^2, \quad \exists u \in C^1[0,1] \quad \text{s.d. } Au = r.$$

Notwendige Optimalitätsbedingung:

$$\begin{aligned} L(u,\lambda) &= J(u) + \langle \lambda, Au - r \rangle_{\boldsymbol{R}^2} \\ &= \int_0^1 \sqrt{1+[u'(x)]^2}dx + \lambda_1 u(0) + \lambda_2 (u(1) - 1) \end{aligned}$$

stationär, oder

$$\forall v \in C^1[0,1]$$

$$0 = \partial_u L(u^\star, \lambda^\star)[v] = \int_0^1 \frac{u^{\star\prime}(x)v'(x)}{\sqrt{1+[u^{\star\prime}(x)]^2}}dx + \lambda_1^\star v(0) + \lambda_2^\star v(1)$$

$$\forall \kappa \in \boldsymbol{R}^2$$

$$0 = \partial_\lambda L(u^\star, \lambda^\star)[\kappa] = \begin{bmatrix} \kappa_1 u^\star(0) \\ \kappa_2 (u^\star(1) - 1) \end{bmatrix}$$

Wenn $u^\star$ genügend glatt ist, gilt

$$\begin{aligned} \partial_u L(u^\star, \lambda^\star)[v] &= \left. \frac{u^{\star\prime}(x)}{\sqrt{1+[u^{\star\prime}(x)]^2}} v(x) \right|_{x=0}^{x=1} + \lambda_1^\star v(0) + \lambda_2^\star v(1) \\ &- \int_0^1 v(x) \left(\frac{u^{\star\prime}(x)}{\sqrt{1+[u^{\star\prime}(x)]^2}} \right)' dx \end{aligned}$$

Wenn $v(x) \to \delta(x - \hat{x})$, $\hat{x} \in (0,1)$, $\partial_u L(u^\star, \lambda^\star) = 0 \Rightarrow$

$$\left(\frac{u^{\star\prime}(x)}{\sqrt{1 + [u^{\star\prime}(x)]^2}} \right)' = 0, \;\; x \in (0,1).$$

Die allgemeine Lösung:

$$\left(\frac{u^{\star\prime}(x)}{\sqrt{1 + [u^{\star\prime}(x)]^2}} \right) = k_1 \;\; \Rightarrow \;\; u^{\star\prime}(x) = k_2 \;\; \Rightarrow \;\; u^\star(x) = k_2 x + k_3$$

Die Nebenbedingungen $\Rightarrow u^\star(x) = x$ und

$$\lambda_1 = \frac{u^{\star\prime}(0)}{\sqrt{1 + [u^{\star\prime}(0)]^2}} = \frac{1}{\sqrt{2}} \qquad \lambda_2 = -\frac{u^{\star\prime}(1)}{\sqrt{1 + [u^{\star\prime}(1)]^2}} = -\frac{1}{\sqrt{2}}.$$

$\lambda \neq \mathbf{0} \Rightarrow$ Einschränkungen *aktiv.*

Aufgabe: Etabliere den Rahmen und löse:

$$\min_{u \in H} J(u) = \int_0^1 \sqrt{1 + [u'(x)]^2} dx \;\;\text{ s.d. }\;\; u(x_i) = u_i, \;\; i = 1, n.$$

Hinweis: $H = H^1[0,1]$ oder $H = C^0[0,1] \cap_{i=1}^{n-1} C^0(x_i, x_{i+1})$.

Beispiel: Sei $\tilde{u} \in L^2[0,1]$ verrauscht und finde $u \in H^1[0,1]$ eine entrauschte Abschätzung mit $u(x_0) = u_0$, $x_0 \in (0,1)$.

Formulierung:

$$\min_{u \in H^1} J(u) \;\;\text{ s.d. }\; u(x_0) = u_0$$

wobei

$$J(u) = \int_0^1 [u(x) - \tilde{u}(x)]^2 dx + \mu \int_0^1 [u'(x)]^2 dx$$

$$Au = u(x_0), \quad r = u_0$$

$A \in \boldsymbol{\mathcal{L}}(H^1[0,1], \boldsymbol{R})$

$$|Au| = |u(x_0)| \leq \|u\|_{C^0} \leq \epsilon \|u\|_{H^1} \qquad (\epsilon \text{ von Einbettung})$$

ist surjektiv:

$$\forall r \in \boldsymbol{R}, \quad \exists u \in H^1[0,1] \quad \text{s.d. } Au = r.$$

Notwendige Optimalitätsbedingung:

$$\begin{aligned} L(u,\lambda) &= J(u) + \langle \lambda, Au - r \rangle_{\boldsymbol{R}} \\ &= \int_0^1 [u(x) - \tilde{u}(x)]^2 dx + \mu \int_0^1 [u'(x)]^2 dx + \lambda [u(x_0) - u_0] \end{aligned}$$

stationär, oder

$$\begin{aligned} &\forall v \in H^1[0,1] \\ 0 = \partial_u L(u^\star, \lambda^\star)[v] &= 2 \int_0^1 \{[u^\star(x) - \tilde{u}(x)] v(x) + \mu u^{\star\prime}(x) v'(x)\} dx \\ &\qquad + \lambda^\star v(x_0) \\ &\forall \kappa \in \boldsymbol{R} \\ 0 = \partial_\lambda L(u^\star, \lambda^\star)[\kappa] &= \kappa [u^\star(x_0) - u_0] \end{aligned}$$

Wenn $u^\star$ genügend glatt ist, gilt

$$\begin{aligned} \partial_u L(u^\star, \lambda^\star)[v] &= 2 \int_0^1 [u^\star(x) - \tilde{u}(x)] v(x) dx + \lambda^\star v(x_0) \\ &+ 2\mu u^{\star\prime}(x) v(x) \big|_{x=0}^{x=x_0} + 2\mu u^{\star\prime}(x) v(x) \big|_{x=x_0}^{x=1} \\ &- 2\mu \int_0^{x_0} v(x) u^{\star\prime\prime}(x) dx - 2\mu \int_{x_0}^1 v(x) u^{\star\prime\prime}(x) dx \end{aligned}$$

Wenn $v \to 1$, $x \in \{0, x_0, 1\}$, und sonst 0, oder $v(x) \to \delta(x - \hat{x})$, $\hat{x} \in (0, x_0) \cup (x_0, 1)$, $\partial_u L(u^\star, \lambda^\star) = 0 = \partial_\lambda L(u^\star, \lambda^\star) \Rightarrow$

$$\begin{cases} -\mu u^{\star\prime\prime} + u = \tilde{u}, & x \in (0, x_0) \cup (x_0, 1) \\ u^{\star\prime} = 0, & x = 0, 1 \\ -\mu [u^{\star\prime}] = \frac{1}{2}\lambda, & x = x_0 \\ u^\star = u_0, & x = x_0 \end{cases}$$

Nimm $\tilde{u} = 0$, $\mu = 1$.

$$u'' = u \;\Rightarrow\; u(x) = A_i \cosh(x) + B_i \sinh(x),$$
$$i = 1 \text{ in } [0, x_0],\; i = 2 \text{ in } [x_0, 1].$$

Die Nebenbedingungen $\Rightarrow$

$$u^\star(x) = \begin{cases} u_0 \dfrac{\cosh(x)}{\cosh(x_0)}, & x \in [0, x_0] \\[2ex] u_0 \dfrac{\cosh(x)}{\cosh(x_0)} \dfrac{[1 - \tanh(1)\tanh(x)]}{[1 - \tanh(1)\tanh(x_0)]}, & x \in [x_0, 1] \end{cases}$$

und

$$\lambda = \frac{2u_0}{\cosh(x_0)^2} \frac{\tanh(1)}{[1 - \tanh(1)\tanh(x_0)]}$$

$\lambda \neq 0 \Rightarrow$ Einschränkung aktiv.

Minimierung unter nicht linearen Einschränkungen.

Seien $J \in C^2(H, \boldsymbol{R})$, $E \in C^2(H, \Lambda)$ und

$$\min_{u \in H} J(u) \quad \text{s.d.} \quad E(u) = 0$$

Nimm an: $J(u^\star) = \min\limits_{u \in H} J(u)$ s.d. $E(u^\star) = 0$.

Bahn $u(t)$, $u(0) = u^\star$, $E(u(t)) = 0 \Rightarrow$

$$0 = D_t E(u(t)) = \partial E(u(t))[D_t u(t)]$$

Störungen $p = D_t u(0)$ im Tangentraum erfüllen

$$0 = \partial E(u^\star)[p]$$

Nimm an: $\mathcal{N}(\partial E(u^\star)) \neq \{0\}$.

Nimm an: $\partial E(u^\star) \in \mathcal{L}(H, \Lambda)$ surjektiv, so
$$\mathcal{N}(\partial E(u^\star))^\perp = \mathcal{R}(\partial E(u^\star)^*).$$

Taylorsatz: $p \in \mathcal{N}(\partial E(u^\star))$,
$$J(u^\star + p) = J(u^\star) + \partial J(u^\star)[p] + \tfrac{1}{2}\partial^2 J(u^\star)[p]^2 + \sigma \|p\|_H^2$$
wobei $\sigma \to 0$ für $\|p\|_H \to 0$.

Wie vorher, sind die folgenden Bedingungen notwendig:
$$\begin{aligned} \partial J(u^\star)[p] &= 0, \qquad \forall p \in \mathcal{N}(\partial E(u^\star)) \\ \partial^2 J(u^\star)[p]^2 &\ge 0, \qquad \forall p \in \mathcal{N}(\partial E(u^\star)) \end{aligned}$$
und die folgende hinreichend:
$$\exists \alpha > 0 \text{ s.d. } \partial^2 J(u^\star)[p]^2 \ge \alpha \|p\|_H^2, \qquad \forall p \in \mathcal{N}(\partial E(u^\star)).$$

Notwendige Optimalitätsbedingung: wie vorher ist $(u^\star, \lambda^\star)$ stationär für die Lagrangesche Funktion
$$L(u, \lambda) = J(u) + \langle \lambda, E(u) \rangle_\Lambda$$
oder
$$\begin{aligned} \forall v \in H, \quad 0 = \partial_u L(u^\star, \lambda^\star)[v] &= \partial J(u^\star)[v] + \langle \lambda^\star, \partial E(u^\star)[v] \rangle_\Lambda \\ \forall \kappa \in \Lambda, \quad 0 = \partial_\lambda L(u^\star, \lambda^\star)[\kappa] &= \langle \kappa, E(u^\star) \rangle_\Lambda \end{aligned}$$

Newtons Verfahren.
$$\begin{aligned} \partial_u \partial_u L(u, \lambda)[v, du] + \partial_u \partial_\lambda L(u, \lambda)[v, d\lambda] &= -\partial_u L(u, \lambda)[v] \\ \partial_\lambda \partial_u L(u, \lambda)[\kappa, du] + \partial_\lambda \partial_\lambda L(u, \lambda)[\kappa, d\lambda] &= -\partial_\lambda L(u, \lambda)[\kappa] \end{aligned}$$
oder
$$\begin{array}{rcl} \begin{array}{r} \partial^2 J(u)[v, du] + \\ \langle \lambda, \partial^2 E(u)[v, du] \rangle_\Lambda \end{array} + \langle \Delta\lambda, \partial E(u)[v] \rangle_\Lambda &=& \begin{array}{l} -\partial J(u)[v] \\ -\langle \lambda, \partial E(u)[v] \rangle_\Lambda \end{array} \\ \langle \kappa, \partial E(u)[du] \rangle_\Lambda &=& -\langle \kappa, E(u) \rangle_\Lambda \end{array}$$

oder

$$\begin{bmatrix} W & Z^* \\ Z & 0 \end{bmatrix} \begin{bmatrix} du \\ d\lambda \end{bmatrix} = - \begin{bmatrix} R \\ S \end{bmatrix}$$

Gaußsche Elimination,

$$\begin{aligned} (ZW^{-1}Z^*)d\lambda &= ZW^{-1}R - S \\ W du &= R - Z^* d\lambda \end{aligned}$$

Notwendig: $W[p]^2 = \partial_u^2 L(u^\star, \lambda^\star)[p]^2 = D_t^2 J(u(0)) \geq 0$, $\forall p \in \mathcal{N}(\partial E(u^\star))$.

Augmentierte Lagrangesche Funktion.

$$L_A(u, \lambda) = J(u) + \langle \lambda, E(u) \rangle_\Lambda + \rho \langle E(u), E(u) \rangle_\Lambda$$

Wenn $(u^\star, \lambda^\star)$ stationär für $L(u, \lambda)$ ist, dann auch für $L_A(u, \lambda)$,

$$\begin{aligned} \partial_u L_A(u^\star, \lambda^\star)[v] &= \underbrace{\partial J(u^\star)[v] + \langle \lambda^\star, \partial E(u^\star)[v] \rangle_\Lambda}_{=0} \\ &+ 2\rho \langle \underbrace{E(u^\star)}_{=0}, \partial E(u^\star)[v] \rangle_\Lambda \\ \partial_\lambda L_A(u^\star, \lambda^\star)[k] &= \langle \kappa, \underbrace{E(u^\star)}_{=0} \rangle_\Lambda \end{aligned}$$

aber $\partial_u^2 L_A(u^\star, \lambda^\star)[v, w]$ besser konditioniert mit $\rho > 0$:

$$\begin{aligned} \partial_u^2 L_A(u^\star, \lambda^\star)[v, w] &= \partial_u^2 L(u^\star, \lambda^\star)[v, w] \\ &+ 2\rho \langle \partial E(u^\star)[v], \partial E(u^\star)[w] \rangle_\Lambda \\ &+ 2\rho \langle \underbrace{E(u^\star)}_{=0}, \partial^2 E(u^\star)[v, w] \rangle_\Lambda \\ \\ &= \partial_u^2 L(u^\star, \lambda^\star)[v, w] \\ &+ 2\rho \langle \partial E(u^\star)[v], \partial E(u^\star)[w] \rangle_\Lambda \end{aligned}$$

oder

$$\partial_u^2 L(u^\star, \lambda^\star)[v]^2 = W[v]^2 + 2\rho \|\partial E(u^\star)[v]\|_\Lambda^2 \geq \beta \|v\|_H^2.$$

Wende Newtons Verfahren für L_A an,

$$\begin{bmatrix} W + \rho A & Z^* \\ Z & 0 \end{bmatrix} \begin{bmatrix} du \\ d\lambda \end{bmatrix} = - \begin{bmatrix} R \\ S \end{bmatrix}$$

oder implementiere Augmentierte Lagrangesche SQP Interation:

K. Ito, K. Kunisch, *Augmented Lagrangian SQP Methods in Hilbert Spaces and Application to Control in the Coefficient Problems*, SIAM J. on Optimization, Vol. 6, No. 1, pp. 96-125, February 1996.