Optimalitätsbedingungen

Einführung in das Optimierungsproblem

Diese Ausarbeitung meines Seminarthemas beschäftigt sich mit Optimalitätsbedingungen von Optimierungsproblemen. Dazu werden wir uns zuerst zwei Methoden anschauen, die uns später bei der Ausarbeitung der Optimalitätsbedingungen helfen. Zum Schluss werden wir sie an einem Beispiel veranschaulichen.

Im Allgemeinen sieht ein **Optimierungsproblem** wie folgt aus:

$$\begin{aligned} &\min_{(y,u)\in Y\times U} J(y,u) \\ \text{unter der Nebenbedingung} \quad & e(y,u)=0, \ (y,u)\in W_{ad}, \end{aligned} \tag{1}$$

wobei $J : Y \times U \to \mathbb{R}$ eine **Zielfunktion**, $e : Y \times U \to Z$ ein Operator und $U \ni u \mapsto y(u) \in Y$ ein **Lösungsoperator** sind. Des Weiteren sind Y, U, Z Banachräume und $W_{ad} \subset W := Y \times U$ eine nichtleere, abgeschlossene Teilmenge.

Wir nehmen an, dass sowohl die Zielfunktion J, als auch der Operator e stetig Fréchet-differenzierbar sind. Es existiere zu jedem $u \in U$ genau eine Lösung $y(u) \in Y$, sodass die **Zustandsgleichung** $e(y,u) = 0$ erfüllt ist. Zudem sei $e_y(y(u),u) \in \mathcal{L}(Y,Z)$ stetig invertierbar.

Mithilfe des Theorems über Implizite Funktionen [Thm. 1.41] wissen wir, dass $y(u)$ stetig differenzierbar ist. Setzen wir $y(u)$ in (1) ein, so erhalten wir das **Reduzierte Problem**

$$\begin{aligned} &\min_{u\in U} \hat{J}(u) := J(y(u),u) \\ \text{unter der Nebenbedingung} \quad & u \in \hat{U}_{ad} := \{u \in U : (y(u),u) \in W_{ad}\}. \end{aligned}$$

Im Weiteren wollen wir die Ableitung der reduzierten Zielfunktion berechnen. Dazu gibt es zwei Methoden, die wir nun kennenlernen.

Analysemethoden der reduzierten Zielfunktion

Sensitivitätsanalyse

Die Sensibilität $\delta_s y = dy(u,s)$, wobei $s \in U$ die Richtung angibt, ist gegeben durch die linearisierte Zustandsgleichung $e_y(y(u),u)\delta_s y = -e_u(y(u),u)s$. Die Richtungsableitung der reduzierten Zielfunktion berechnen wir in zwei Schritten.

1. Berechne die Sensibilität $\delta_s y = dy(u,s)$ durch Lösen der linearisierten Zustandsgleichung $e_y(y(u),u)\delta_s y = -e_u(y(u),u)s$.

2. Berechne die Richtungsableitung der reduzierten Zielfunktion $d\hat{J}'(u) = \langle \hat{J}'(u), s\rangle_{U^*,U}$ durch $d\hat{J}'(u) = \langle J_y(y(u),u), \delta_s y\rangle_{Y^*,Y} + \langle J_u(y(u),u), s\rangle_{U^*,U}$.

Um die allgemeine Ableitung der reduzierten Zielfunktion zu erhalten, müssen wir jede ihrer Richtungsableitungen berechnen. Da dies sehr viele sein können, wird diese Methode sehr teuer. Dagegen effizienter ist der nächste Ansatz.

Adjungierte Analyse

Wir definieren drei wichtige Begriffe, die wir im weiteren Verlauf noch häufig nutzen werden.

Adjungierter Zustand:	$p = p(u) \in Z^*$
Adjungierte Gleichung:	$e_y(y(u),u)^* p = -J_y(y(u),u)$
Adjungierte Gradientendarstellung:	$\hat{J}'(u) = e_u(y(u),u)^* p + J_u(y(u),u)$

Lagrange basierte Sicht auf die adjungierte Analyse

Wir haben gesehen, dass die reduzierte Zielfunktion abgeleitet die adjungierte Gradientendarstellung ergibt. Diese können wir auch auf eine andere Weise schreiben. Dazu definieren wir die Lagrangefunktion durch

$$L : Y \times U \times Z^* \to \mathbb{R} \text{ mit } L(y,u,p) := J(y,u) + \langle p, e(y,u)\rangle_{Z^*,Z}.$$

Durch Einsetzen von $y = y(u)$ und die Wahl von $p = p(u)$, sodass $L_y(y(u),u,p(u)) = 0$ gilt, erhalten wir die adjungierte Gradientendarstellung in folgender Form:

$$\hat{J}'(u) = L_u(y(u),u,p(u)) = J_u(y(u),u) + e_u(y(u),u)^* p.$$

Optimalitätsbedingungen für einfach bedingte Probleme

Wir betrachten das Optimierungsproblem

$$\min_{w \in W} J(w) \quad \text{unter } w \in C, \tag{2}$$

wobei W ein Banachraum, $C \subset W$ nichtleere, abgeschlossene, konvexe Teilmenge und die Zielfunktion $J : W \to \mathbb{R}$ Gâteaux-differenzierbar ist.

Definition 1 (Def. 1.30)**.** *$\bar{w} \in W$ ist **optimal** für (2), falls gilt:*

$$\bar{w} \in C \text{ und } J(\bar{w}) \leq J(w) \; \forall w \in C.$$

Theorem 2 (Thm. 1.46)**.** *Sei W ein Banachraum, $C \subset W$ eine nichtleere, konvexe Teilmenge und die Zielfunktion $J : V \to \mathbb{R}$ definiert auf einer offenen Umgebung V von C. Weiter sei $\bar{w}$ eine lokale Lösung von (2), wobei J Gâteaux-differenzierbar ist.*

*(i) Dann gilt die **Optimalitätsbedingung**:*

$$\bar{w} \in C, \; \langle J'(\bar{w}), w - \bar{w} \rangle_{W^*,W} \geq 0 \; \forall w \in C. \tag{3}$$

(ii) Falls J konvex auf C gilt, so ist (3) notwendig und hinreichend für ein globales Optimum.

(iii) Falls J strikt konvex auf C ist, dann existiert höchstens eine Lösung von (2).

(iv) Falls sowohl W reflexiv, C abgeschlossen und konvex, als auch J konvex und stetig mit $\lim\limits_{w \in C, \|w\|_W \to \infty} J(w) = \infty$ gelten, dann exisiert eine (globale=lokale) Lösung von (2).

Beweis:

(i) Sei $w \in C$ beliebig. Da C konvex ist, gilt $w(t) := \bar{w} + t(w - \bar{w}) \in C \; \forall t \in [0,1]$. Weiter gilt aufgrund der Definition 1.30 und der Konvexität von C:

$$0 \leq J(w(t)) - J(\bar{w}) = J(\bar{w} + t(w - \bar{w})) - J(\bar{w}) \; \forall t \in [0,1].$$

Mit der Definition der dualen Paarung und der Gâteau-Differenzierbarkeit von J folgt:

$$\langle J'(\bar{w}, w - \bar{w} \rangle_{W^*,W} = dJ(\bar{w}, w - \bar{w}) = \lim_{t \to 0^+} \frac{J(\bar{w} + t(w - \bar{w}) - J(\bar{w})}{t} \geq 0.$$

(ii) Sei J konvex und $w \in C$ beliebig. Dann gilt:

$$J(w) - J(\bar{w}) = \lim_{t \to 0^+} \frac{J(\bar{w}) + t(J(w) - J(\bar{w})) - J(\bar{w})}{t} \overset{J\ konvex}{\geq} \lim_{t \to 0^+} \frac{J(\bar{w} + t(w - \bar{w}) - J(\bar{w})}{t}$$

Da J Gâteau-differenzierbar ist, können wir mithilfe von Aussage (i) die Optimalität von $\bar{w}$ folgern:

$$\lim_{t \to 0^+} \frac{J(\bar{w} + t(w - \bar{w}) - J(\bar{w})}{t} = \langle J'(\bar{w}, w - \bar{w}\rangle_{W^*,W} \geq 0$$

(iii) Sei J strikt konvex. Wir nehmen an, dass w_1 und 2_2 mit $w_1 \neq w_2$ zwei globale Lösungen sind. Dann wäre aber $\frac{w_1+w_2}{2} \in C$ eine bessere Lösung. Dies ist ein Widerspruch und somit haben wir höchstens eine Lösung.

(iv) Sei $(w_k)_{k \in \mathbb{N}} \subset C$ eine minimierende Folge, d.h. sie konvergiert gegen ihr Minimum. Angenommen, die Folge $(w_k)_{k \in \mathbb{N}}$ wäre nicht beschränkt. Dann würde $\|w_k\|_W \overset{k \to \infty}{\to} \infty$ gelten und somit nach Voraussetzung $J(w_k) \overset{k \to \infty}{\to} \infty$. Demnach ist $(w_k)_{k \in \mathbb{N}}$ beschränkt. Da W ein reflexiver Banachraum und $C \in W$ abgeschlossen ist, ist auch C ein reflexiver Banachraum. Nun existiert eine Teilfolge $(w_{k_j})_{j \in \mathbb{N}}$ von $(w_k)_{k \in \mathbb{N}}$, die schwach gegen ein $\bar{w}$ konvergiert [Thm. 1.17]. Weiter liegt $\bar{w}$ in C, da C sowohl abgeschlossen, als auch konvex und somit schwach folgenabgeschlossen ist [Thm. 1.16]. Stetigkeit und Konvexität von J führen zu einer schwachen Unterhalbstetigkeit von J, d.h. $w_k \rightharpoonup \bar{w} \Rightarrow \liminf_{k \to \infty} J(w_k) \geq J(\bar{w})$ [Thm. 1.18]. Zudem gilt $J(\bar{w}) \leq \liminf_{k \to \infty} J(w_k) \leq \lim_{k \to \infty} J(w_k) = \inf_{w \in C} J(w) \leq J(w) \forall w \in C$. Demnach ist $\bar{w}$ eine optimale Lösung von (2) [Def. 1.30].

□

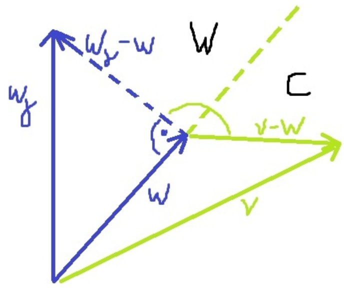
Bemerkung:
Die Optimalitätsbedingung (3) wird auch **Variationsungleichung** genannt.

Lemma 3 (Lem. 1.11)**.** *Sei W ein Hilbertraum, $C \subset W$ eine nichtleere, abgeschlossene und konvexe Teilmenge und $P : W \to C$ eine Projektion auf C.*
Dann sind für alle $y \in Y$ und $\gamma > 0$ äquivalent:

$$w \in C,\ (y, v - w)_W \geq 0\ \forall v \in C.$$
$$w - P(w - \gamma y) = 0.$$

Beweisidee: Setze $w_\gamma := w - \gamma y$. Dann ist die folgende Äquivalenz äquivalent zu der, die zu zeigen ist.

$$w \in C,\ (w_\gamma - w, v - w)_W \leq 0\ \forall v \in C. \quad P(w_\gamma) = w.$$



□

Korollar 4 (Kor. 1.2)**.** *Sei W ein Hilbertraum, $C \subset W$ eine nichtleere, abgeschlossene, konvexe Menge und die Zielfunktion $J : V \to \mathbb{R}$ sei definiert auf einer offenen Umgebung V von C. Weiter sei $\bar{w}$ lokale Lösung von (2), wobei J Gâteau-differenzierbar ist.*

Dann gilt die Optimalitätsbedingung:

$$\bar{w} = P(\bar{w} - \gamma \nabla J(\bar{w})), \tag{4}$$

wobei $\gamma > 0$ beliebig, aber fest ist und $\nabla J(w) \in W$ die Riesz-Darstellung von $J'(w) \in W^*$ bezeichnet.

Falls $\emptyset \neq C$ abgeschlossen und konvex ist, dann sind die Aussagen (3) und (4) äquivalent.

Optimalitätsbedingungen für steuerungsbedingte Probleme

In diesem Kapitel beschäftigen wir uns mit einem nichtlinearen Problem, welches wie folgt aussieht:

$$\begin{aligned} &\min_{(y,u) \in Y \times U} J(y,u) \\ &\text{unter der Nebenbedingung } e(y,u) = 0,\ u \in U_{ad}, \end{aligned} \tag{5}$$

wobei $U_{ad} := \{u \in U \mid a \leq u \leq b\}$.

Voraussetzung 5 (Vor. 1.47)**.**

1. *U_{ad} ist eine nichtleere, abgeschlossene und konvexe Teilmenge von U.*

2. *$J : Y \times U \to \mathbb{R}$, $e : Y \times U \to Z$ sind stetig Fréchet-differenzierbar, wobei U, Y, Z Banachräume sind.*

3. *Für alle u in einer Umgebung $V \subset U$ von U_{ad} existiert genau eine Lösung $y(u)$ aus Y, sodass die Zustandsgleichugn $e(y(u), u) = 0$ erfüllt ist.*

4. $e_y(y(u), u) \in \mathcal{L}(Y, Z)$ *hat eine beschränkte Inverse.*

Unter den gegebenen Voraussetzungen erhalten wir das reduzierte Problem von (5), welches nur von u und nicht mehr von y abhängt.

$$\min_{u \in U} \hat{J}(u) = J(y(u), u) \quad \text{unter } u \in U_{ad}. \tag{6}$$

Theorem 6 (Thm. 1.48)**.** *Seien die Voraussetzungen aus Vor. 1.47 erfüllt. Falls* $\bar{u}$ *eine lokale Lösung des reduzierten Problems (6) ist, so erfüllt* $\bar{u}$ *die Variationsungleichung, d.h.* $\bar{u} \in U_{ad},\ \langle \hat{J}'(\bar{u}), u - \bar{u} \rangle_{U^*,U} \geq 0\ \forall u \in U_{ad}$.

Beweis: Wende Theorem 1.46 an. □

Korollar 7 (Kor. 1.3)**.** *Seien die Voraussetzungen aus Vor. 1.47 erfüllt. Zudem sei* $(\bar{y}, \bar{u})$ *optimale Lösung des Problems (5).*

Dann existiert ein adjungierter Zustand $\bar{p} \in Z^*$*, sodass folgende Optimalitätsbedingung erfüllt ist:*

$$\begin{aligned} e(\bar{y}, \bar{u}) &= 0 && \textit{(Zustandsgleichung)} \\ e_y(\bar{y}, \bar{u})^*\ \bar{p} &= -J_y(\bar{y}, \bar{u}) && \textit{(Adjungierte Gleichung)} \\ \bar{u} \in U_{ad}, \langle J_u(\bar{y}, \bar{u}) + e_u(\bar{y}, \bar{u})^*\ \bar{p}, u - \bar{u} \rangle_{U^*,U} &\geq 0\ \forall u \in U_{ad} && \textit{(Adj. Gradientendarstellung)} \end{aligned}$$

oder mithilfe der Lagrange-Funktion in kompakter Form:

$$\begin{aligned} L_p(\bar{y}, \bar{u}, \bar{p}) &= e(\bar{y}, \bar{u}) = 0 \\ L_y(\bar{y}, \bar{u}, \bar{p}) &= 0 \\ \bar{u} \in U_{ad}, \langle L_u(\bar{y}, \bar{u}, \bar{p}), u - \bar{u} \rangle_{U^*,U} &\geq 0\ \forall u \in U_{ad} \end{aligned}$$

oder in Variationsform:

$$\begin{aligned} \langle e(\bar{y}, \bar{u}), p \rangle_{Z,Z^*} &= 0\ \forall p \in Z^* \\ \langle L_y(\bar{y}, \bar{u}, \bar{p}), v \rangle_{Y^*,Y} &= 0\ \forall v \in Y \\ \bar{u} \in U_{ad}, \langle L_u(\bar{y}, \bar{u}, \bar{p}), u - \bar{u} \rangle_{U^*,U} &\geq 0\ \forall u \in U_{ad} \end{aligned}$$

Lemma 8 (Lem. 1.12)**.** *Sei $a, b \in U = L^2(\Omega)$ mit $a \leq b$ und $U_{ad} := \{u \in L^2(\Omega) : a \leq u \leq b\}$. Dann sind äquivalent:*

(i)

$$\bar{u} \in U_{ad}, \ \ (\nabla \hat{J}(\bar{u}), u - \bar{u})_U \geq 0 \ \forall u \in U_{ad}.$$

(iii)

$$\begin{aligned} \exists \ \lambda_a, \lambda_b \in L^2(\Omega) : \quad & \nabla \hat{J}(\bar{u}) + \lambda_b - \lambda_a = 0, \\ & \bar{u} \geq a, \ \ \lambda_a \geq 0, \ \ \lambda_a(\bar{u} - a) = 0, \\ & \bar{u} \leq b, \ \ \lambda_b \geq 0, \ \ \lambda_b(b - \bar{u}) = 0. \end{aligned}$$

Beweis:

(iii) $\Rightarrow$ (i):

$$\begin{aligned} &(\nabla \hat{J}(\bar{u}), u - \bar{u})_{L^2(\Omega)} = \int_\Omega \nabla \hat{J}(\bar{u})(u - \bar{u}) \ dx = \int_\Omega -(\lambda_b - \lambda_a)(u - \bar{u}) \ dx \\ &= - \int_\Omega \lambda_b(u - \bar{u}) - \lambda_a(u - \bar{u}) \ dx = - \int_\Omega \lambda_b(u - \bar{u}) + \lambda_a(\bar{u} - u) \ dx \\ &\underset{u \ \in \ U_{ad}}{\geq} \ - \int_\Omega \underbrace{\lambda_b(b - \bar{u})}_{=0} - \underbrace{\lambda_a(\bar{u} - a)}_{=0} \ dx = 0 \end{aligned}$$

(i) $\Rightarrow$ (iii): Da $\bar{u} \in U_{ad}$, liegt $\bar{u}$ zwischen a und b. Definiere $\lambda_a := \max(\nabla \hat{J}(\bar{u}), 0)$ und $\lambda_b := \max(-\nabla \hat{J}(\bar{u}), 0)$. Nun erfüllen λ_a und λ_b schon $\nabla \hat{J}(\bar{u}) + \lambda_b - \lambda_a = 0$ und $\lambda_a \geq 0, \lambda_b \geq 0$.

Falls $\bar{u} = a$ oder $\lambda_a = 0$, dann gilt trivialerweise $\lambda_a(\bar{u} - a) = 0$. Analog gilt $\lambda_b(b - \bar{u}) = 0$, wenn $\bar{u} = b$ oder $\lambda_b = 0$.

Sei $\bar{u} > a$. Angenommen, es gelte $\nabla \hat{J}(\bar{u}) > 0$. Ohne Beschränkung der Allgemeinheit existiert eine Menge $M \subset I_b := \{x \mid a(x) < \bar{u}(x) \leq b(x)\}$ mit $\nabla \hat{J}(\bar{u}) > 0$ auf M. Wähle $u := \bar{u} + 1_M(a - \bar{u})$. Dann gilt $u \in U_{ad}, \ u - \bar{u} < 0$ auf M und sonst $u - \bar{u} = 0$. Somit gilt:

$$(\nabla \hat{J}(\bar{u}), u - \bar{u})_U = \int_{I_b} \nabla \hat{J}(\bar{u}) \ (\bar{u} + 1_M(a - \bar{u}) - \bar{u}) \ dx = \int_M \nabla \hat{J}(\bar{u}) \ (a - \bar{u}) \ dx < 0.$$

Dies steht im Widerspruch zu (i). Daher gilt $\nabla \hat{J}(\bar{u}) \leq 0$. Es folgt $\lambda_a = 0$ und somit auch $\lambda_a(\bar{u} - a) = 0$.

Sei $\bar{u} < b$. Analog zu oben gilt nun $\nabla \hat{J}(\bar{u}) \geq 0$ und es folgt $\lambda_b(b - \bar{u}) = 0$.

□

Im nächsten Theorem fassen wir unsere bisherigen Kenntnisse über Optimalitätsbedingungen zusammen.

Theorem 9. *Falls* $(\bar{y}, \bar{u})$ *eine optimale Lösung des Problems (5) ist, dann existiert ein* $\bar{p} \in Z^*, \lambda_a, \lambda_b \in U^*$, *sodass folgende* ***Optimalitätsbedingung im schwachen Sinn*** *erfüllt ist:*

$$\begin{aligned} e(\bar{y}, \bar{u}) &= 0, \\ e_y(\bar{y}, \bar{u})^* \ \bar{p} &= -J_y(\bar{y}, \bar{u}), \\ J_u(\bar{y}, \bar{u}) + e_u(\bar{y}, \bar{u})^* \ \bar{p} + \lambda_b - \lambda_a &= 0, \\ \bar{u} \geq a, \lambda_a \geq 0, \lambda_a(\bar{u} - a) &= 0, \\ \bar{u} \leq b, \lambda_b \geq 0, \lambda_b(b - \bar{u}) &= 0. \end{aligned}$$

Optimalitätsbedingungen für Probleme mit allgemeiner Nebenbedingung

Wir betrachten nun das allgemeine Problem:

$$\begin{aligned} &\min_{w \in W} J(w) \\ \text{unter der Nebenbedingung } &G(w) \in K_G, \ w \in C, \end{aligned} \tag{7}$$

wobei die Zielfunktion $J : W \to \mathbb{R}$ und $G : W \to V$ stetig Fréchet-differenzierbar sind und W, V Banachräume sind. C sei eine nichtleere, abgeschlossene, konvexe Teilmenge von V. Zudem sei K_G ein abgeschlossener, konvexer Kegel in V.

Definition 10. *K heißt **Kegel**, falls für alle $\lambda > 0$ gilt:* $\quad v \in K \Rightarrow \lambda v \in K$.

Weiter sind die Menge F_{ad} und der Operator $G(w)$ der Nebenbedingung wie folgt definiert:

$$\begin{aligned} F_{ad} &:= \{w \in W | G(w) \in K_G, w \in C\} \\ G(w) &:= (e(w), c(w)) \in \{0\} \times K =: K_G. \end{aligned}$$

Definition 11 (Def. 1.31)**.** *Sei $F_{ad} \subset W$ nichtleere Teilmenge. Der **Tangentialkegel** von F_{ad} in $w \in F_{ad}$ ist definiert durch*

$$T(F_{ad}, w) := \{s \in W \mid \exists\ \eta_k > 0, w_k \in F_{ad} : \lim_{k \to \infty} w_k = w, \lim_{k \to \infty} \eta_k (w_k - w) = s\}.$$

Definition 12. *Ein* ***linearisierter Kegel*** *in* $\bar{w} \in F_{ad}$ *ist definiert durch*

$$L(F_{ad}, G, K_G, C, \bar{w}) := \{\eta d \mid \eta > 0, d \in W : G(w) + G'(w)d \in K_G,\ \bar{w} + d \in C\}.$$

Definition 13. *Drei wichtige Bedingungen sind im Folgenden definiert.*

$$\begin{aligned} \textit{Constraint Qualification:} \quad & L(F_{ad}, G, K_G, C, \bar{w}) \subset T(F_{ad}, w) \\ \textit{Regularitätsbedingung von Robinson:} \quad & 0 \in int(G(\bar{w}) + G'(\bar{w})(C - \bar{w}) - K_G) \\ \textit{Slaters Bedingung:} \quad & \exists\ \tilde{w} \in W : G(\tilde{w}) \in int(K_G), \tilde{w} \in C \end{aligned}$$

Theorem 14 (Thm. 1.54)**.** *Die Regularitätsbedingung von Robinson impliziert Constraint Qualification.*

Theorem 15 (Thm. 1.55)**.** *Seien sowohl $C \subset W$, als auch $K_G \subset V$ abgeschlossen und konvex. Falls $G : W \to V$ konvex ist, dann folgt aus der Slaters Bedingung die Regularitätsbedingung von Robinson für alle $\bar{w} \in C$ mit $G(\bar{w}) \in K_G$.*

Beweis: Für genügend kleines $\epsilon > 0$ gilt $G(\tilde{w}) + B_V(\epsilon) \subset K_G$. Sei $\bar{w} \in C$ beliebig, aber fest mit $G(\bar{w}) \in K_G$ und $w(t) := \bar{w} + t(\tilde{w} - \bar{w})$ für $t \in [0, 1]$. Da G konvex ist, gilt für alle $t \in [0, 1]$:

$$K_G \ni \frac{G(w(t)) - (1-t)G(\bar{w}) - tG(\tilde{w})}{t} = \frac{G(w(t)) - G(\bar{w})}{t} + G(\bar{w}) - G(\tilde{w}).$$

Durch Anwendung des Differenzenquotienten erhalten wir die folgende Konvergenz:

$$\frac{G(w(t)) - G(\bar{w})}{t} = \frac{G(\bar{w} + t(\tilde{w} - \bar{w}) - G(\bar{w})}{t(\tilde{w} - \bar{w})}(\tilde{w} - \bar{w}) \ \underrightarrow{t \to 0}\ G'(\bar{w})(\tilde{w} - \bar{w}).$$

Wir wissen nun, dass ein $d \in K_G$ existiert mit $G(\bar{w})G'(\bar{w})(\tilde{w} - \bar{w}) - K_G = G(\tilde{w}) + d - K_G$. Da $d \in K_G$ ist und somit auch $d + K_G \subset K_G$, gilt weiter $G(\tilde{w}) + d - K_G \supset G(\tilde{w}) - K_G$. Mit $\tilde{w} \in C$ aus der Slaters Bedingung folgt

$$G(\bar{w})G'(\bar{w})(C - \bar{w}) - K_G \ \supset\ G(\bar{w})G'(\bar{w})(\tilde{w} - \bar{w}) - K_G \ \supset\ G(\tilde{w}) - K_G \ \supset\ B_V(\epsilon)$$

und daraus die Behauptung. □

Theorem 16 (Thm. 1.52)**.** *Sei die Zielfunktion $J : W \to \mathbb{R}$ stetig Fréchet-differenzierbar. Dann gilt für jede lokale Lösung $\bar{w}$ von (7):*

$$\bar{w} \in F_{ad},\ \langle J'(\bar{w}), s\rangle_{W^*,W} \geq 0\ \forall s \in T(F_{ad}, w).$$

Beweis: Da $\bar{w}$ eine Lösung von (7) ist, gilt schon $\bar{w} \in F_{ad}$.

Sei $s \in T(F_{ad}, \bar{w})$ beliebig. Wegen der Definition des Tangentialkegels wissen wir, dass eine Folge $(w_k)_{k\in\mathbb{N}} \subset F_{ad}$ und $\eta_k > 0$ existieren mit $w_k \to \bar{w}$ und $\eta_k(w_k - \bar{w}) \to s$ für $k \to \infty$. Da J stetig ist und $\bar{w}$ optimal gilt $0 \leq \eta_k(J(w_k) - J(\bar{w}))$ für alle $k \in \mathbb{N}$. Nun wenden wir die qualitative Taylorformel im Entwicklungspunkt $\bar{w}$ an und erhalten:

$$\eta_k(J(w_k) - J(\bar{w})) = \langle J'(\bar{w}), \eta_k(w_k - \bar{w})\rangle_{W^*,W} + o(\|w_k - \bar{w}\|_W) \xrightarrow{k\to\infty} \langle J'(\bar{w}), s\rangle_{W^*,W}.$$

□

Theorem 17 (Thm. 1.56)**.** *Seien $J : W \to \mathbb{R}$, $G : W \to V$ stetig Fréchet-differenziebrar, W, V Banachräume, $C \subset W$ nichtleere, abgeschlossene und konvexe Teilmenge und $K_G \subset V$ abgeschlossen und konvex. Zudem sei $\bar{w}$ eine lokale Lösung von (7), die die Regularitätsbedingung von Robinson erfüllt. Dann gilt für $\bar{w}$ die folgende Optimalitätsbedingung:*

Es existiert ein $\bar{q} \in V^*$ mit

$$\begin{aligned}
&G(\bar{w}) \in K_G,\\
&\bar{q} \in K_G^\circ := \{q \in V^* : \langle q, v\rangle_{V^*,V} \leq 0 \ \forall v \in K_G\},\\
&\langle \bar{q}, G(\bar{w})\rangle_{V^*,V} = 0,\\
&\bar{w} \in C, \langle J'(\bar{w}) + G'(\bar{w})^*\bar{q}, w - \bar{w}\rangle_{W^*,W} \geq 0 \forall w \in C.
\end{aligned}$$

Bemerkung: K_G° bezeichnet den dualen Kegel von K_G.

Anwendungsbeispiel

Wir untersuchen die optimalverteilte Kontrolle einer gleichmäßig verteilten Temperatur mit Randtemperatur Null. Dabei bezeichnen y die Temperatur und y_d die gewünschte Temperatur. Das zugehörige Optimierungsproblem ist das Folgende:

$$\begin{aligned}
\min_{(y,u)\in Y\times U} J(y,u) = \frac{1}{2}\|y - y_d\|_H^2 + \frac{\alpha}{2}\|u\|_U^2 \qquad\qquad&\\
\text{unter der Nebenbedingung} \quad -\Delta y = \gamma y \quad &\text{in } \Omega,\\
y = 0 \quad &\text{in } \partial\Omega,\\
a \leq u \leq b \quad &\text{in } \Omega,
\end{aligned} \tag{8}$$

wobei $\gamma \in L^\infty(\Omega) \setminus \{0\}$ mit $\gamma \geq 0$ und $a, b \in L^2(\Omega)$ mit $a \leq b$ gelten.

Wir wählen

$$\begin{aligned} U &= H = L^2(\Omega), \\ U_{ad} &= \{u \in L^2(\Omega) : a \leq u \leq b\}, \\ Y_{ad} &= Y = H_0^1(\Omega), \\ Z &= Y^* = H^{-1}(\Omega). \end{aligned}$$

Somit erhalten wir die Zielfunktion und den Operator der Nebenbedingung in folgender Form:

$$\begin{aligned} J &: H_0^1(\Omega) \times L^2(\Omega) \to \mathbb{R}, \\ e &: H_0^1(\Omega) \times L^2(\Omega) \to H^{-1}(\Omega). \end{aligned}$$

Wir wollen nun die schwache Formulierung von $\begin{aligned} -\Delta y &= \gamma y \text{ in } \Omega \\ y &= 0 \text{ in } \partial\Omega \end{aligned}$

herleiten. Sei $v \in H_0^1$ beliebig, aber fest. Dann gilt:

$$\int_\Omega -\Delta y \ v \ dx \overset{Gauß}{=} \int_\Omega \nabla y \ \nabla v \ dx - \underbrace{\int_{\partial\Omega} \nabla y \ v \ n \ dx}_{=0,\ da\ v \in H_0^1} = \int_\Omega \nabla y \ \nabla v \ dx \overset{Lemma\ 1.7}{=} a(y,v)$$

$$\int_\Omega -\Delta y \ v \ dx = \int_\Omega \gamma u \ v \ dx = (\gamma u, v)_{L^2(\Omega)}$$

Also können wir sie schreiben als $a(y,v) = (\gamma u, v)_{L^2(\Omega)}\ \forall v \in H_0^1(\Omega)$ oder $Ay + Bu = 0$, wobei $A \in \mathcal{L}(H_0^1(\Omega), H^{-1}(\Omega))$ mit $Ay = a(y,.)$ und $B \in \mathcal{L}(L^2(\Omega) H^{-1}(\Omega))$ mit $Bu = -(\gamma u, .)_{L^2(\Omega)}$. Wir erhalten die Zustandsgleichung in der Form

$$a(y,v) - (\gamma u, v)_{L^2(\Omega)} = 0 \ \text{ beziehungsweise } \ Ay + Bu = 0.$$

Zunächste berechnen wir die partiellen Ableitungen von J und e:

$$
\begin{aligned}
J_y(y,u) &= \frac{d}{dy}\Big(\frac{1}{2}\|y-y_d\|^2_{L^2(\Omega)}\Big) = \frac{1}{2}\int_\Omega \frac{d}{dy}(|y-y_d|^2)dx = \int_\Omega |y-y_d|\ dx = (y-y_d,.)_{L^2(\Omega)},\\
J_u(y,u) &= \frac{d}{dy}\Big(\frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}\Big) = \frac{\alpha}{2}\int_\Omega \frac{d}{dy}(|u|^2)dx = \alpha\int_\Omega |u|\ dx = (\alpha u,.)_{L^2(\Omega)},\\
e_y(y,u)(v) &= \frac{d}{dy}\Big(\int_\Omega \nabla y\ \nabla v\ dx\Big) = \frac{d}{dy}\Big(\int_\Omega \nabla y\ \nabla v\ dx - \underbrace{\int_{\partial\Omega} y\ \nabla v\ n\ dx}_{=0,\ da\ y=0\ \partial\Omega}\Big) \overset{Gauß}{=} \frac{d}{dy}\Big(\int_\Omega -y\ \Delta v\ dx\Big)\\
&= \int_\Omega -\Delta v\ dx = (-\Delta v,.)_{L^2(\Omega)} \quad \text{beziehungsweise}\\
e_y(y,u) &= \frac{d}{dy}(Ay+Bu) = A,\\
e_u(y,u)(v) &= \frac{d}{du}\Big(\int_\Omega -\gamma u\ v\ dx\Big) = \int_\Omega -\gamma\ v\ dx = (-\gamma v,\ .)_{L^2(\Omega)} \quad \text{beziehungsweise}\\
e_u(y,u) &= \frac{d}{dy}(Ay+Bu) = B.
\end{aligned}
$$

Da $H_0^1(\Omega)$ ein Hilbertraum ist, wissen wir aufgrund der Reflexivität, dass $H_0^1(\Omega)^{**} = H_0^1(\Omega)$ und somit $H^{-1}(\Omega)^* = H_0^1(\Omega)$ ist. Um nun zum einen das reduzierte Optimierungsproblem herzuleiten und zum anderen die zugehörigen Optimalitätsbedingungen zu erhalten, prüfen wir die Annahmen aus Voraussetzung 1.47.

1. $\emptyset \neq \{u \in L^2(\Omega) : a \leq u \leq b\} \subset L^2(\Omega)$ ist nichtleer, abgeschlossen und konvex.
2. J und e sind stetig Fréchet-differenzierbar, wobei $H_0^1(\Omega), H^{-1}(\Omega)$ und $L^2(\Omega)$ Banachräume sind.
3. Für jedes u in einer Umgebung $V \subset L^2(\Omega)$ von $\{u \in L^2(\Omega) : a \leq u \leq b\}$ existiert genau eine Lösung $y(u) \in H_0^1(\Omega)$, sodass die Zustandsgleichung $Ay + Bu = 0$ erfüllt ist.
4. $e_y(y(u), u) = A \in \mathcal{L}(H_0^1(\Omega), H^{-1}(\Omega))$ hat eine beschränkte Inverse.

Das reduzierte Problem von (8) unter den Voraussetzungen 1.47 lautet nun

$$
\begin{gathered}
\min_{u\in U} \hat{J}(u) = \frac{1}{2}\|y(u)-y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}\\
\text{unter der Nebenbedingung } u \in U_{ad}.
\end{gathered}
\tag{9}
$$

Da die Voraussetzungen 1.47 erfüllt sind, können wir Theorem 1.48 anwenden. Die adjungierte Gradientendarstellung liefert:

$$\hat{J}'(\bar{u}) = J_u(y(u), u) + e_u(y(u), u)^* \ p(\bar{u}).$$

Da $L^2(\Omega) = H = U$ ein Hilbertraum ist und somit $L^2(\Omega)^* = L^2(\Omega)$ gilt, folgt mit der Rieszschen Darstellung:

$$\begin{aligned}
J_u(y(u), u) &= (\alpha u, .)_{L^2(\Omega)} = \alpha u \\
e_u(y(u), u)(v) &= (-\gamma v,\ .)_{L^2(\Omega)} = -\gamma v \\
\langle e_u(y(u), u)^* \ p(u),\ v\rangle_{L^2(\Omega)^*, L^2(\Omega)} &= \langle p(u),\ e_u(y(u), u)(v)\rangle_{H_0^1(\Omega), H^{-1}(\Omega)} = \langle p(u),\ -\gamma v\rangle_{H_0^1(\Omega), H^{-1}(\Omega)} \\
&= (p(u),\ -\gamma v)_{L^2(\Omega)} = (-\gamma p(u),\ v)_{L^2(\Omega)} = \langle -\gamma p(u),\ v\rangle_{L^2(\Omega)^*, L^2(\Omega)} \\
e_u(y(u), u)^* \ p(u) &= -\gamma p(u) \\
&\Rightarrow \hat{J}'(\bar{u}) = \alpha \bar{u} - \gamma p(\bar{u}).
\end{aligned}$$

Demnach besagt das Theorem 1.48, dass jede lokale Lösung $\bar{u}$ von (9) die Variationsungleichung erfüllt, das heißt es gilt:

$$\bar{u} \in U_{ad},\ (\alpha \bar{u} - \gamma p(\bar{u}),\ u - \bar{u})_{L^2(\Omega)} \geq 0 \ \forall u \in U_{ad}.$$

Des Weiteren finden wir mithilfe des Korollars 1.3 für jede optimale Lösung $(\bar{y}, \bar{u})$ von (8) ein $\bar{p}$ aus $H_0^1(\Omega)$, sodass die Variationsform erfüllt ist.

$$\begin{aligned}
\langle e(\bar{y}, \bar{u}), v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} &= e(\bar{y}, \bar{u})(v) = a(\bar{y}, v) - (\gamma \bar{u}, v)_{L^2(\Omega)}, \\
\langle e_y(\bar{y}, \bar{u})^* \ \bar{p}, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} &= \langle \bar{p}, e_y(\bar{y}, \bar{u})(v)\rangle_{H_0^1(\Omega), H^{-1}(\Omega)} = \langle \bar{p}, -\Delta v\rangle_{H_0^1(\Omega), H^{-1}(\Omega)} \\
&= (\bar{p}, -\Delta v)_{L^2(\Omega)} = \int_\Omega -\bar{p} \ \Delta v \ dx \\
&\overset{Gauß}{=} \int_\Omega \nabla \bar{p} \ \nabla v \ dx - \underbrace{\int_{\partial\Omega} \bar{p} \ \nabla v \ n \ dx}_{=0,\ da\ \bar{p} \in H_0^1(\Omega)} = a(\bar{p}, v) \quad \text{beziehungsweise} \\
\langle e_y(\bar{y}, \bar{u})^* \ \bar{p}, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} &= \langle \bar{p}, e_y(\bar{y}, \bar{u}) \ v\rangle_{H_0^1(\Omega), H^{-1}(\Omega)} = \langle \bar{p}, Av\rangle_{H_0^1(\Omega), H^{-1}(\Omega)} \\
&= \langle Av, \bar{p}\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} = a(v, \bar{p}) = a(\bar{p}, v), \\
\langle L_y(\bar{y}, \bar{u}, \bar{p}), v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} &= \langle J_y(\bar{y}, \bar{u}), v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} + \langle e_y(\bar{y}, \bar{u})^* \ \bar{p}, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)} \\
&= (\bar{y} - y_d, v)_{L^2(\Omega)} + a(\bar{p}, v), \\
L_u(\bar{y}, \bar{u}, \bar{p}) &= J_u(\bar{y}, \bar{u}) + e_u(\bar{y}, \bar{u})^* \ \bar{p} = \alpha \bar{u} - \gamma \bar{p}, \\
\langle L_u(\bar{y}, \bar{u}, \bar{p}), u - \bar{u}\rangle_{L^2(\Omega), L^2(\Omega)} &= (\alpha \bar{u} - \gamma \bar{p}, u - \bar{u})_{L^2(\Omega)}.
\end{aligned}$$

Somit sieht die Optimalitätsbedingung in Variationsform wie folgt aus:

$$\begin{aligned} a(\bar{y}, v) - (\gamma \bar{u}, v)_{L^2(\Omega)} &= 0 \ \forall v \in H_0^1(\Omega), \\ (\bar{y} - y_d, v)_{L^2(\Omega)} + a(\bar{p}, v) &= 0 \ \forall v \in H_0^1(\Omega), \\ (\alpha \bar{u} - \gamma \bar{p}, u - \bar{u})_{L^2(\Omega)} &\geq 0 \ \forall u \in U_{ad}. \end{aligned}$$

Sei nun wieder $(\bar{y}, \bar{u})$ eine optimale Lösung von (8). Wir wollen das Theorem über die Optimalitätsbedingung im schwachen Sinn anwenden. Dazu berechnen wir zuerst:

$$\begin{aligned} e(\bar{y}, \bar{u}) &= -\Delta \bar{y} - \gamma \bar{u} = 0 \\ \langle e_y(\bar{y}, \bar{u})^* \ \bar{p}, v \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} &= a(\bar{p}, v) = \int_\Omega \nabla \bar{p} \ \nabla v \ dx - \underbrace{\int_{\partial\Omega} \nabla \bar{p} \ v \ n \ dx}_{=0, \ da \ v \in H_0^1(\Omega)} \overset{Gauß}{=} \int_\Omega -\Delta \bar{p} \ v \ dx \\ &= (-\Delta \bar{p}, v)_{L^2(\Omega)} = \langle -\Delta \bar{p}, v \rangle \\ &\Rightarrow e_y(\bar{y}, \bar{u})^* \ \bar{p} = -\Delta \bar{p} \\ J_y(\bar{y}, \bar{u}) + e_y(\bar{y}, \bar{u})^* \ \bar{p} &= (\bar{y} - y_d) - \Delta \bar{p} \\ J_y(\bar{y}, \bar{u}) + e_y(\bar{y}, \bar{u})^* \ \bar{p} &= \alpha \bar{u} - \gamma \bar{p} \end{aligned}$$

Schließlich erhalten wir:
$\exists \ \bar{p} \in H_0^1(\Omega), \ \lambda_a, \lambda_b \in L^2(\Omega):$

$$\begin{aligned} -\Delta \bar{y} = \gamma \bar{u}, \ \bar{y}|_{\partial\Omega} &= 0, \\ -\Delta \bar{p} = -(\bar{y} - y_d), \ \bar{p}|_{\partial\Omega} &= 0, \\ \alpha \bar{u} - \gamma \bar{p} + \lambda_b - \lambda_a &= 0, \\ \bar{u} \geq a, \ \lambda_a \geq 0, \ \lambda_a(\bar{u} - a) &= 0, \\ \bar{u} \leq b, \ \lambda_b \geq 0, \ \lambda_b(b - \bar{u}) &= 0. \end{aligned}$$