2 Notwendige Bedingungen erster Ordnung für Optimalsteuerprobleme

2.1 Quadratische Optimierungsprobleme im Hilbertraum

Wir betrachten grundsätzlich das Problem

$$\min_{u \in U_{\mathrm{ad}}} \; j(u) = \frac{1}{2}\left\|Su - y_d\right\|_H^2 + \frac{\lambda}{2}\|u\|_U^2,$$

wobei U und H Hilberträume sind und U_{ad} eine konvexe und abgeschlossene Menge. Der Operator $S\colon U \to H$ sei linear und stetig. Die Existenz einer Lösung dieses Problems folgt direkt aus Satz 1.2.9.

derivative concept

optimality conditions

Für die Optimalitätsbedingungen benötigen wir Ableitungsbegriffe. Das kennt man schon aus der Schule. Sei dazu $F\colon U \to V$ gegeben.

directional derivative

- **Richtungsableitung**: Existiert der Grenzwert

$$\delta F(u,h) = \lim_{t \downarrow 0} \frac{F(u+th) - F(u)}{t} \qquad \text{in } V$$

  in V, so heißt F *richtungsdifferenzierbar* in u in Richtung h.

- **Erste Variation**: Existiert die Richtungsableitung in alle Richtungen h, so heißt die Abbildung $h \mapsto \delta F(u;h)$ auch *erste Variation*.

- **Gâteaux-Ableitung**: Ist die erste Variation in u linear und stetig, so heißt der entsprechende lineare Operator A *Gâteaux-Ableitung* in u; wir schreiben auch $A = F'(u)$.

- **Fréchet-Ableitung**: Ist F Gâteaux-differenzierbar und gilt

$$F(u+h) = F(u) + F'(u)h + \tau(h) \qquad \text{in } V$$

  mit einer Funktion $\tau\colon U \to V$, die

$$\frac{\|\tau(h)\|_V}{\|h\|_U} \to 0 \quad \text{für} \quad \|h\|_U \to 0,$$

erfüllt, so heißt A *Fréchet-Ableitung*. Wir schreiben auch $\tau(h) \in o(\|h\|_U)$.

Beispiele 2.1.1. Einige Beispiele Fréchet-differenzierbarer Funktionen:

(1) Sei $A \in \mathcal{L}(U, V)$ ein linearer stetiger Operator, $b \in V$ und $f(u) = Au + b$. Dann ist $f'(u)h = Ah$, also $f'(u) = A$.

(2) Sei H ein Hilbertraum und $f \colon H \to \mathbb{R}$ gegeben durch $f(u) = \|u\|_H^2 = (u, u)_H$. Dann ist $f'(u)h = 2(u, h)_H$.

(3) Sei $U = C[0, 1]$ und $V = \mathbb{R}$ und $f(u) = \sin(u(1))$. Dann ist $f'(u)h = \cos(u(1))h(1)$. Für ein solches Resultat braucht man eine Kettenregel.

chain rule

May 8, 2024

Lecture 4

Lemma 2.1.2 (Kettenregel)**.** *Es seien $F \colon U \to V$ und $G \colon V \to Z$ Fréchet-differenzierbar. Setze $E \coloneqq G \circ F$. Dann gilt*

$$E'(u)h = G'(F(u))\Big[F'(u)h\Big].$$

Beispiel 2.1.3. Es sei $A \in \mathcal{L}(U, H)$ und $b \in H$ mit $E(u) = \frac{1}{2}\|Au + b\|_H^2$, also $E = G \circ F \colon U \to \mathbb{R}$ mit $G \colon H \to \mathbb{R}$ gegeben durch $G(v) = \frac{1}{2}\|v\|_H^2$ und $F \colon U \to H$ gegeben durch $F(u) = Au + b$. Dann ist $G'(v) = (v, h)_H$ und $F'(u)h = Ah$, also nach der Kettenregel und mit den Riesz-Isomorphismen $\mathcal{R}_H$ und $\mathcal{R}_U$ von H und U:

$$\begin{aligned} E'(u)h = \big(Au + b, Ah\big)_H &= \big\langle \mathcal{R}_H^{-1}(Au + b), Ah\big\rangle_{H^\star, H} \\ &= \big\langle A^* \mathcal{R}_H^{-1}(Au + b), h\big\rangle_{U^\star, U} = \big(\mathcal{R}_U A^* \mathcal{R}_H^{-1}(Au + b), h\big)_U. \end{aligned}$$

Im Falle $U = H$ erhalten wir also

$$E'(u)h = \big(A^\star(Au + b), h\big)_U.$$

Wir können nun für unsere einfache Modellaufgabe die notwendige Optimalitätsbedingung herleiten. Sei dazu $\bar{u}$ ein lokales Minimum und $u \in U_{\text{ad}}$. Dann gilt $\bar{u} + t(u - \bar{u}) \in U_{\text{ad}}$ für $t \in [0, 1]$, da U_{ad} konvex angenommen war. Da zudem $\bar{u}$ lokales Minimum ist, gilt für hinreichend kleines t

$$j(\bar{u}) \leq j\big(\bar{u} + t(u - \bar{u})\big),$$

also

$$0 \leq \frac{j(\bar{u} + t(u - \bar{u})) - j(\bar{u})}{t}, \qquad \text{und mit } t \downarrow 0\text{:} \qquad j'(\bar{u})(u - \bar{u}) \geq 0.$$

Das heißt, die notwendige Bedingung erster Ordnung sieht genauso aus wie im $\mathbb{R}^n$, also wie in der nichtlinearen Optimierung, für eine konvexe Nebenbedingung.

Für das konkrete j ergibt sich mit Beispiel 2.1.3, wobei wir der Einfachheit halber $U = H$ annehmen,

$$j'(\bar{u})(u - \bar{u}) = \big(S^\star(S\bar{u} - y_d) + \lambda\bar{u}, u - \bar{u}\big)_U.$$

Damit lautet die notwendige Optimalitätsbedingung erster Ordnung

$$\big(S^\star(S\bar{u} - y_d) + \lambda\bar{u}, u - \bar{u}\big) \geq 0 \qquad \forall u \in U_{\text{ad}}. \tag{2.1.1}$$

Da wir ein konvexes Funktional auf einer konvexen Menge minimieren, ist diese Bedingung auch *hinreichend*.

2.2 Optimale Steuerung elliptischer Probleme

Wir wollen uns nun mit dem Minimierungsproblem aus dem ersten Kapitel beschäftigen, also mit

$$\min_{y,u \in L^2(\Omega)} J(y,u) := \frac{1}{2}\big\|y - y_d\big\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Omega)}$$

mit

$$\left.\begin{aligned} -\Delta y &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{auf } \Gamma \end{aligned}\right\} \tag{2.2.1}$$

und

$$u \in U_{\text{ad}} = \Big\{u \in L^2(\Omega) \colon u_a \leq u(x) \leq u_b \text{ fast überall auf } \Omega\Big\}.$$

in advance

Wir setzen $\lambda \geq 0$ voraus. Wir wollen dieses Problem auf die Struktur von Kapitel 2.1 bringen. Dazu setzen wir $H = U = L^2(\Omega)$. Der Operator S wird definiert über die schwache Lösung von (2.2.1), also $Su = y$, wenn

$$\int_\Omega \nabla y \cdot \nabla v \,\mathrm{d}x = \int_\Omega u \cdot v \,\mathrm{d}x \quad \forall v \in H_0^1(\Omega).$$

Das ist sinnvoll, denn Satz 1.2.15 (Lax-Milgram) liefert uns die Stetigkeit der Abbildung $S\colon H_0^1(\Omega)^\star \to H_0^1(\Omega)$. Diese ist tatsächlich auch stetig invertierbar. Da $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$, also, $H_0^1(\Omega) \subseteq L^2(\Omega)$ und die Identitätsabbildung ist stetig, faktorisieren wir

$$L^2(\Omega) \cong L^2(\Omega)^\star \hookrightarrow H_0^1(\Omega)^\star \quad \xrightarrow{S} \quad H_0^1(\Omega) \hookrightarrow L^2(\Omega)$$

und erhalten S auch als stetigen linearen Operator

$$S\colon L^2(\Omega) \to L^2(\Omega), \quad u \mapsto y \text{ schwache Lösung von (2.2.1)}.$$

Somit können wir im Zielfunktional y sinnvoll durch Su ersetzen. Die Existenz einer Lösung $\bar{u}$ des Optimierungsproblems folgt somit unmittelbar aus Satz 1.2.9. Die Lösung

immediately
close in space
sofort: close in zeit

ist sogar eindeutig, da unser Zielfunktional strikt konvex ist. Das sieht man im Fall $\lambda > 0$ sehr schnell. Im Fall $\lambda = 0$ benutzt man die Injektivität des Operators S.

Wir wollen nun die nowendige und hinreichende Optimalitätsbedingung aufstellen. Dafür benötigen wir den adjungierten Operator $S^\star$. Laut Definition muss für alle $u, v \in L^2(\Omega)$ gelten, dass

$$(Su, v)_{L^2(\Omega)} = (u, S^\star v)_{L^2(\Omega)}.$$

Sei $v \in L^2(\Omega)$ und sei $z \in H_0^1(\Omega)$ die eindeutige schwache Lösung des Problems

$$\begin{aligned} -\Delta z &= v \quad \text{in } \Omega \\ z &= 0 \quad \text{auf } \Gamma, \end{aligned}$$

sprich, es gilt

$$\int_\Omega \nabla z \cdot \nabla w \, \mathrm{d}x = \int_\Omega w \cdot v \, \mathrm{d}x \quad \forall w \in H_0^1(\Omega).$$

Sei $u \in L^2(\Omega)$ beliebig und schreiben wir $y = Su \in H_0^1(\Omega)$, so folgt mit dem vorigen

v is the distributed control

$$(Su, v)_{L^2(\Omega)} = \int_\Omega y \cdot v \, \mathrm{d}x = \int_\Omega \nabla z \cdot \nabla y \, \mathrm{d}x = \int_\Omega z \cdot u = (u, z)_{L^2(\Omega)}.$$

Es folgt, dass $S^\star v = z$. Also ist $S^\star$ wieder der Lösungsoperator von (2.2.1) und wir erhalten $S = S^\star$. Insbesondere ist S selbstadjungiert. selfadjoint

In den Optimalitätsbedingungen steht der Ausdruck $S^\star(S\bar{u} - y_d)$. Wir bezeichnen

$$\bar{p} = S^\star(S\bar{u} - y_d)$$

als *adjungierten Zustand*. Wegen $\bar{y} = S\bar{u}$ und der Gestalt von $S^\star$ ist $\bar{p}$ also durch adjoint state

$$\int_\Omega \nabla v \cdot \nabla \bar{p} \, \mathrm{d}x = \int_\Omega (\bar{y} - y_d) \cdot v \, \mathrm{d}x \quad \forall v \in H_0^1(\Omega)$$

charakterisiert, oder in klassischer Schreibweise

$$\begin{aligned} -\Delta \bar{p} &= \bar{y} - y_d \quad \text{in } \Omega \\ \bar{p} &= 0 \quad \text{auf } \Gamma. \end{aligned}$$

Der Ausdruck $\bar{y} - y_d$ beschreibt hierbei den "Defekt" des Zielfunktionals.

Die Optimalitätsbedingung lautet folglich

$$\left(\bar{p} + \lambda \bar{u}, u - \bar{u}\right)_{L^2(\Omega)} \geq 0 \quad \forall u \in U_{\mathrm{ad}}.$$

Das gesamte Optimalitätssystem hat demnach die Form:

$$\begin{aligned} -\Delta\bar y &= \bar u \quad \text{in } \Omega, & -\Delta\bar p &= \bar y - y_d \quad \text{in } \Omega,\\ \bar y &= 0 \quad \text{auf } \Gamma, & \bar p &= 0 \qquad\quad \text{auf } \Gamma,\end{aligned}$$

$$\big(\bar p + \lambda\bar u, u - \bar u\big)_{L^2(\Omega)} \geq 0 \quad \forall u \in U_{\text{ad}}.$$

May 14, 2024
Lecture 5 ⟶

Für viele Untersuchungen ist es sinnvoll, die Variationsungleichung

$$\big(\bar p + \lambda\bar u, u - \bar u\big)_{L^2(\Omega)} \geq 0 \quad \forall u \in U_{\text{ad}} \tag{2.2.2}$$

äquivalent umzuformulieren. Ausgeschrieben lautet diese Beziehung

$$\int_\Omega \big(\bar p(x) + \lambda\bar u(x)\big)\big(u(x) - \bar u(x)\big)\,\mathrm{d}x \geq 0 \quad \forall u \in U_{\text{ad}}.$$

Unser U_{ad} ist gegeben durch

$$U_{ad} = \Big\{u \in L^2(\Omega)\colon u_a \leq u(x) \leq u_b \text{ fast überall in } \Omega\Big\}.$$

Daher muss die Ungleichung tatsächlich äquivalent punktweise fast überall gelten, also für fast jedes $x \in \Omega$:

$$\big(\bar p(x) + \lambda\bar u(x)\big)\big(u(x) - \bar u(x)\big) \geq 0 \quad \forall u \in U_{\text{ad}}. \tag{2.2.3}$$

Das sieht man so: Angenommen, es gäbe ein $u \in U_{\text{ad}}$, so dass die umgekehrte Ungleichung auf einer Menge A mit Maß $|A| > 0$ gälte. Dann wäre aber die Funktion

$$\hat u(x) := \chi_A(x)u(x) + \chi_{\Omega\setminus A}\bar u(x) = \begin{cases} u(x) & \text{in } A,\\ \bar u(x) & \text{sonst}\end{cases}$$

wieder in U_{ad} und es folgte

$$\int_\Omega \big(\bar p(x) + \lambda\bar u(x)\big)\big(\hat u(x) - \bar u(x)\big)\,\mathrm{d}x = \int_A \big(\bar p(x) + \lambda\bar u(x)\big)\big(u(x) - \bar u(x)\big)\,\mathrm{d}x < 0.$$

Das kann nicht sein, also gilt (2.2.3) eben doch punktweise für fast jedes $x \in \Omega$. Damit kann man nun noch etwas mehr arbeiten: Es folgt direkt

$$\bar u(x) = u_a \qquad\qquad \text{für fast alle} \quad x \in \big[\bar p + \lambda\bar u > 0\big]$$

und umgekehrt reverse

$$\bar p(x) + \lambda\bar u(x) \geq 0 \qquad\qquad \text{für fast alle} \quad x \in \big[\bar u = u_a\big], \tag{2.2.4}$$

und analog für u_b. Es bleibt die Menge $\big[\bar p + \lambda\bar u = 0\big]$. Für x aus dieser Menge unterscheiden wir:

- $\lambda > 0$: Dann gilt
$$\bar{u}(x) = \frac{-\bar{p}(x)}{\lambda}.$$
Definiert man die konvexe Projektion
$$P_{[\alpha,\beta]}(z) = \begin{cases} z, & z \in [\alpha,\beta], \\ \alpha, & z < \alpha, \\ \beta, & z > \beta, \end{cases}$$
dann kann man die Variationsungleichung (2.2.2) bzw. (2.2.3) wiederum äquivalent schreiben als
$$\bar{u}(x) = P_{[u_a,u_b]}\Big(-\frac{\bar{p}(x)}{\lambda}\Big) \qquad \text{für fast alle } x \in \Omega. \tag{2.2.5}$$
Im Fall $\lambda > 0$ vererbt sich die Regularität $\bar{p} \in H^1(\Omega)$ auf die optimale Steuerung $\bar{u}$, es treten aber in der Regel Knickstellen auf.

- $\lambda = 0$: Hier erhalten wir im Fall $x \in [\bar{p} = 0]$ *keine* Information für $\bar{u}(x)$ aus dem Optimalitätssystem. Die Lösungen können nun sehr unterschiedliche Gestalt besitzen. Wir betrachten zwei Extremfälle.

  Extremfall 1: $\bar{y} \equiv y_d$
  Damit können wir die gewünschte Funktion y_d exakt ansteuern und der optimale Zielfunktionswert ist Null. Wir erhalten
$$\begin{aligned} -\Delta\bar{p} = \bar{y} - y_d = 0 \quad & \text{in } \Omega \\ \bar{p} = 0 \quad & \text{auf } \Gamma \end{aligned}$$
  und damit $\bar{p} \equiv 0$ auf ganz Ω. Die Optimalitätsbedingung enthält hier keine Information. Aus $\bar{y} \equiv y_d$ erhalten wir aber
$$\bar{u} = -\Delta\bar{y} = -\Delta y_d.$$

  Extremfall 2: $|[\bar{p} = 0]| = 0$.
  In diesem Extremfall nimmt die optimale Steuerung $\bar{u}$ fast überall in Ω die extremalen Werte u_a bzw. u_b an, d.h.
$$\bar{u}(x) \in \{u_a, u_b\} \quad \text{fast überall in } \Omega.$$
  Eine solche Struktur bezeichnet man als *Bang-Bang* und die Steuerung $\bar{u}$ entsprechend als *Bang-Bang-Steuerung*, siehe Abbildung 2.1.

  Wir müssen im Fall $\lambda = 0$ also sogar mit (einer Vielzahl von) Sprungstellen rechnen und die Regularität von $\bar{u}$ geht nicht über das generische $L^2(\Omega)$ hinaus. Das hat potenziell Auswirkungen auf die numerische Analysis.

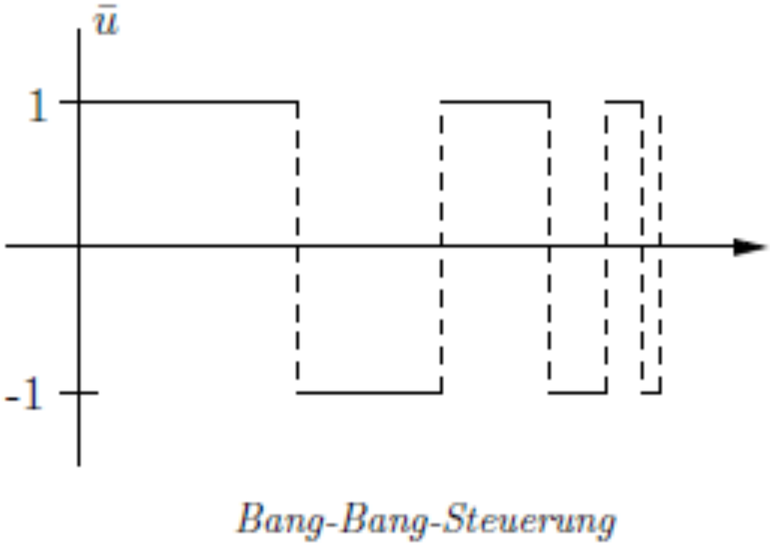

Abbildung 2.1: Bang-Bang Steuerung

Bemerkung 2.2.1. Bang-Bang-Steuerungen sind extremale Punkte der zulässigen Menge U_{ad}. Solche extremalen Punkte besitzen besondere Eigenschaften, die man ausnutzen kann.