Vorlesung Variationsrechnung und optimale Steuerung

7. Foliensatz, Version vom 8.7.2014

Michael Karow

Das Steuerungsproblem

Minimiere

$$J(x, u, t_*) = g(x(t_*)) + \int_{t_0}^{t_*} f^0(t, x(t), u(t))\, dt$$

unter den Nebenbedingungen

(a) $\dot{x}(t) = f(t, x(t), u(t))$, $\qquad$ (x stetig, DGL gilt für fast alle $t \in [t_0, t_*]$)

(b) $x(t_0) = x_0$, $\psi(x(t_*)) = 0$.

(c) $u \in L^\infty([t_0, t_*], \Omega)$ $\qquad$ $\emptyset \neq \Omega \subseteq \mathbb{R}^m$ beliebig

Die Funktionen $f^0, g : \mathbb{R}^n \to \mathbb{R}$, $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\psi : \mathbb{R}^n \to \mathbb{R}^d$ sind stetig diff'bar.
Wenn $\psi \equiv 0$, dann ist $x(t_*)$ frei.
Wenn $\psi(x) = x - x_1$, dann ist $x(t_*) = x_1$ fest.
Die Endzeit t_* kann fest oder variabel sein.

Dem Problem wird eine **Hamilton-Funktion** zugeordnet

$$H(t, x, \lambda, u) = \lambda_0\, f^0(t, x, u) + \lambda^\top f(t, x, u), \qquad \lambda_0 \in \{0, 1\},\ \ \lambda, x \in \mathbb{R}^n,\ \ u \in \mathbb{R}^m$$

λ heisst adjungierter Zustand (Kozustand).

'Normalfall': $\lambda_0 = 1$. Gilt immer bei freiem Endzustand.

Das Pontryaginsche Minimumprinzip

Sei $(x_*, u_*) : [t_0, t_*] \to \mathbb{R}^n \times \Omega$ eine Lösung der Minimierungsaufgabe auf der vorherigen Seite. Dann existieren $\lambda_0 \in \{0, 1\}$, $\nu \in \mathbb{R}^n$ und eine stetige und fast überall differenzierbare Funktion $\lambda : [t_0, t_*] \to \mathbb{R}^n$ mit $(\lambda_0, \lambda(t), \nu) \neq 0$, so dass folgende Aussagen gelten.

(i) Fast überall in $[t_0, t_*]$ (insbesondere an allen Stetigkeitspunkten von u_*) gilt die **Minimumbedingung**

$$H(t, x_*(t), \lambda(t), u_*(t)) = \min_{u \in \Omega} H(t, x_*(t), \lambda(t), u).$$

(ii) λ erfüllt die **adjungierte DGL**

$$\begin{aligned}\dot{\lambda}(t) &= -\nabla_x H(t, x_*(t), \lambda(t), u_*(t)) \\ &= -[f_x(t, x_*(t), u_*(t))]^\top \lambda(t) - \lambda_0 \nabla_x f^0(t, x_*(t), u_*(t)).\end{aligned}$$

(iii) Zur Endzeit gilt die **Transversalitätsbedingung**

$$\lambda(t_*) = \lambda_0 \nabla g(x_*(t_*)) + \psi_x(x_*(t_*))^\top \nu.$$

(iv) Die Funktion $t \mapsto H(t, x_*(t), \lambda(t), u_*(t))$ ist stetig und für fast alle t diff'bar mit

$$\frac{d}{dt} H(t, x_*(t), \lambda(t), u_*(t)) = H_t(t, x_*(t), \lambda(t), u_*(t)) \qquad \text{für fast alle } t \in [0, t_*].$$

(v) Im Fall freier Endzeit ist $\quad H(t_*, x_*(t_*), \lambda(t_*), u_*(t_*)) = 0$

Zur Notation: f_x und ψ_x bezeichnen Jacobi-Matrizen bzgl. der Variablen x.
∇_x ist der Gradient bzgl. x (Spaltenvektor). H_t ist partielle Ableitung von H nach t.

Bemerkungen zum Minimumprinzip I

(1) Wählt man $\lambda_0 = 0$ oder -1 statt $\lambda_0 = 0$ oder 1, dann wird die Minimumbedingung (i) zur Maximumbedingung

$$H(t, x_*(t), \lambda(t), u_*(t)) = \max_{u \in \Omega} H(t, x_*(t), \lambda(t), u).$$

Diese Formulierung findet man häufig in der Literatur.

(2) Die adjungierte DGL ist von der Form

$$\dot{\lambda}(t) = -A(t)^\top \lambda(t) - b(t)$$

mit

$$A(t) = f_x(t, x_*(t), u_*(t)), \qquad b(t) = \lambda_0 \nabla_x f^0(t, x_*(t), u_*(t)).$$

Es handelt sich also um eine inhomogene lineare DGL.

(3) Oft ist es bequemer, die adjungierte DGL in transponierter Form zu schreiben:

$$\dot{\lambda}(t)^\top = -\lambda(t)^\top f_x(t, x_*(t), u_*(t)) - \lambda_0 f_x^0(t, x_*(t), u_*(t)).$$

Dies gibt auch die Bedeutung von λ korrekt wieder, denn $\lambda(t)$ ist eigentlich eine Linearform (aus dem Dualraum von $\mathbb{R}^n$).

Bemerkungen zum Minimumprinzip II

(4) Man hat $\nabla_\lambda H = f(t,x,u)$. Damit lassen sich die DGL für x und λ in der fast symmetrischen Form

$$\dot{x}_*(t) = \nabla_\lambda H(t, x_*(t), \lambda(t), u_*(t))$$

$$\dot{\lambda}(t) = -\nabla_x H(t, x_*(t), \lambda(t), u_*(t))$$

schreiben. Dies sind die Hamiltonschen DGL.

(5) Wenn $u_*(t)$ innerer Punkt von Ω ist, dann folgt aus der Minimumbedingung, dass

$$H_u(t, x_*(t), \lambda(t), u_*(t)) = 0.$$

(5) Wenn $u_*(t)$ innerer Punkt von Ω ist, und u_*, x_*, λ an der Stelle t differenzierbar sind, dann folgt die Ableitungsgleichung in Aussage (iv) des Minimumprinzips so:

$$\begin{aligned} \frac{d}{dt} H &= H_t + (\nabla_x H)^\top \dot{x}_* + (\nabla_\lambda H)^\top \dot{\lambda} + \underbrace{H_u}_{0} \dot{u}_* \\ &= H_t - \dot{\lambda}^\top \dot{x}_* + \dot{x}_*^\top \dot{\lambda} = H_t \end{aligned}$$

Anwendungsbeispiel 1

Aufgabe: Minimiere ($\gamma \in \mathbb{R}$, $\alpha, \beta, t_* > 0$ fest)

$$J(x,u) = \int_0^{t_*} u(t) - \gamma\, x(t)\, dt, \qquad \dot{x} = -\beta\, x + \alpha\, u, \quad x(0) = x_0 > 0, \quad u \in [0,1]$$

Hamilton-Funktion:

$$H(x,\lambda,u) \;=\; u - \gamma\, x + \lambda\,(-\beta\, x + \alpha\, u) \;=\; (1 + \lambda\,\alpha)u - (\lambda\,\beta + \gamma)\, x$$

Minimumbedingung:

$$H(x,\lambda,u_*) = \min_{u\in[0,1]} H(x,\lambda,u) \quad \Rightarrow \quad u_* = \begin{cases} 0 & \text{wenn } \lambda > -1/\alpha, \\ \text{unbestimmt} & \text{wenn } \lambda = -1/\alpha, \\ 1 & \text{wenn } \lambda < -1/\alpha. \end{cases} \qquad (*)$$

Adjungierte DGL:

$$\dot{\lambda} = -H_x(x_*,\lambda,u_*) = -(\beta\,\lambda + \gamma)$$

Allgemeine Lösung: $\lambda(t) = c\, e^{-\beta\, t} - \frac{\gamma}{\beta}$

Endwert: $\lambda(t_*) = 0$ (weil $x(t_*)$ frei) $\quad \Rightarrow \quad c = \frac{\gamma}{\beta}\, e^{\beta\, t_*}$

Somit: $\lambda(t) = \frac{\gamma}{\beta}(e^{\beta\,(t_* - t)} - 1)$

Nun ergibt sich $u_*(t)$ aus $(*)$, und das Problem ist gelöst.

Anwendungsbeispiel 2

Aufgabe: Maximiere ($\alpha > 0$ fest)

$$-J(x,u) = \int_0^{t_*} (1-u(t))\,x(t)\,dt, \qquad \dot{x} = \alpha\, u\, x, \quad x(0) = x_0 > 0, \quad u \in [0,1]$$

Interpretation der Aufgabe:

x=Menge eines Gutes (Geld, Holz, Fisch), welches zu seiner eigenen Vermehrung gebraucht wird. $-J(x,u) =$ Menge die insgesamt konsumiert wird.

$1-u(t)$ Anteil der konsumierten Menge an der Gesamtmenge $x(t)$.

$u(t)x(t) =$ Menge, die zur Vermehrung zur Verfügung steht.

Weil stets $\alpha\, u \geq 0$ folgt aus $x_0 > 0$, dass $x(t) > 0$.

Hamilton-Funktion: $\quad H(x,\lambda,u) = (u-1)\,x + \lambda\,\alpha\,ux = (1+\lambda\alpha)x\,u - x$

Minimumbedingung:

$$H(x_*,\lambda,u_*) = \min_{u\in[0,1]} H(x_*,\lambda,u) \quad \Rightarrow \quad u_* = \begin{cases} 0 & \text{wenn } \lambda > -1/\alpha, \\ \text{unbestimmt} & \text{wenn } \lambda = -1/\alpha, \\ 1 & \text{wenn } \lambda < -1/\alpha. \end{cases}$$

Anwendungsbeispiel 2, Fortsetzung

Hamilton-Funktion:

$$H(x,\lambda,u) = (1+\lambda\alpha)x\,u - x.$$

Adjungierte DGL:

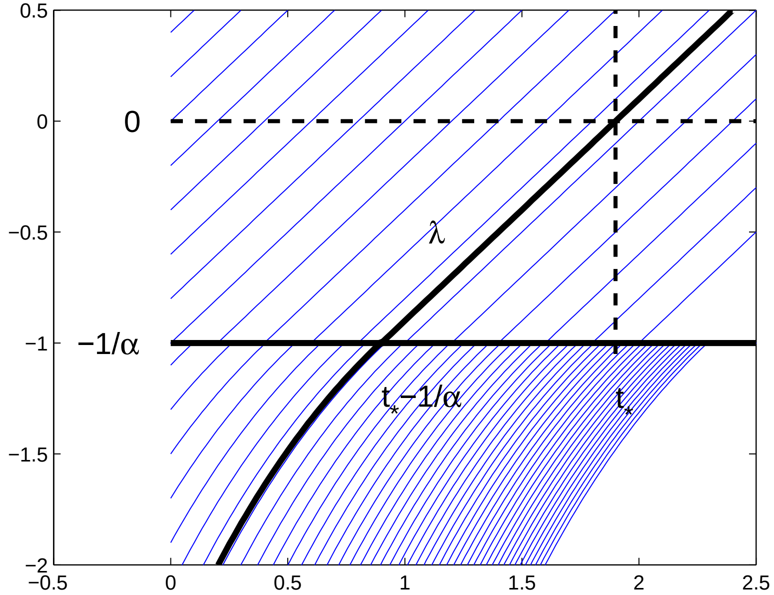
$$\dot\lambda = -H_x(x,\lambda,u_*) = -(1+\lambda\alpha)\,u_* + 1 = \begin{cases} 1 & \text{wenn } \lambda > -1/\alpha \;\; (\Rightarrow \;\; u_* = 0), \\ 1 & \text{wenn } \lambda = -1/\alpha \;\; (\Rightarrow \;\; u_* \text{ unbestimmt}), \\ -\lambda\alpha & \text{wenn } \lambda < -1/\alpha \;\; (\Rightarrow \;\; u_* = 1). \end{cases}$$

Lösung:

$$\lambda(t) = \begin{cases} t + d & \text{wenn } \lambda(t) \geq -1/\alpha, \;\; (\Rightarrow \;\; u_* = 0), \\ c\,e^{-\alpha t} & \text{wenn } \lambda(t) < -1/\alpha \;\; (\Rightarrow \;\; u_* = 1). \end{cases}$$

Endbedingung: $\lambda(t_*) = 0$ (weil $x(t_*)$ frei).

Wir suchen eine stetige Funktion $\lambda(t)$,
welche alle diese Bedingungen erfüllt.
Man findet sie leicht graphisch.
Resultat: $\lambda(t)$ ist linear für $t \geq t_* - 1/\alpha$,
$\lambda(t)$ ist Exponentialfunktion für $t < t_* - 1/\alpha$.
Umschalten von $u_* = 1$ auf $u_* = 0$ bei $t = t_* - 1/\alpha$



Minimumprinzip und Variationsrechnung I

Grundaufgabe der Variationsrechnung: Minimiere

$$J(x,u) = \int_{t_0}^{t_*} f(t, x(t), u(t))\, dt, \qquad \dot{x} = u, \qquad x(a) = x_a, \quad x(b) = x_b.$$

Hamilton-Funktion: $H(t, x, \lambda, u) = f(t, x, u) + \lambda^\top u$

Taylor-Entwicklung von f bezüglich u um $u_*(t)$:

$$f(t, x_*(t), u) = f(t, x_*(t), u_*(t)) + f_u(t, x_*(t), u_*(t))\,(u - u_*(t)) + E,$$

wobei

$$E = E(t, x_*(t), u_*(t), u) \qquad \text{und} \qquad \lim_{u \to u_*(t)} \frac{E(t, x_*(t), u_*(t), u)}{\|u - u_*(t)\|_2} = 0.$$

Minimumbedingung: Für $u \in \Omega$, und fast alle $t \in [t_0, t_*]$,

$$\begin{aligned}
0 &\le H(t, x_*(t), \lambda, u) - H(t, x_*(t), \lambda, u_*(t)) \\
&= (f(t, x_*(t), u) + \lambda^\top u) - (f(t, x_*(t), u_*(t)) + \lambda^\top u_*(t)) \\
&= (f(t, x_*(t), u) - (f(t, x_*(t), u_*(t)) + \lambda^\top (u - u_*(t)) \\
&= f_u(t, x_*(t), u_*(t))(u - u_*(t)) + E + \lambda^\top (u - u_*(t)) \\
&= (f_u(t, x_*(t), u_*(t)) + \lambda^\top)(u - u_*(t)) + E
\end{aligned}$$

Falls $u_*(t)$ innerer Punkt von Ω ist, folgt

$$(a) \quad \lambda^\top = -f_u(t, x_*(t), u_*(t)), \qquad\qquad (b) \quad 0 \le E(t, x_*(t), u_*(t), u).$$

Minimumprinzip und Variationsrechnung II

Zwischenergebnis. Falls $u_*(t)$ innerer Punkt von Ω ist:

$$(a) \quad \lambda^\top = -f_u(t, x_*(t), u_*(t)), \qquad (b) \quad 0 \leq E(t, x_*(t), u_*(t), u).$$

Adjungierte Gleichung:

$$\dot{\lambda}^\top = -H_x = -\frac{\partial}{\partial x}\left[f(t, x_*(t), u_*(t)) + \lambda^\top u_*(t)\right] = -f_x(t, x_*(t), u_*(t))$$

Kombination von (a), (b) und $u_*(t) = \dot{x}_*(t)$ ergibt

(a') $\frac{d}{dt} f_u(t, x_*(t), \dot{x}_*(t)) = f_x(t.x_*(t), \dot{x}_*(t))$ **(1. Euler-Lagrange Gleichung)**

(b') $0 \leq E(t, x_*(t), \dot{x}_*(t), u)$ **(Weierstraß-Kriterium)**.

Aussage (iv) im Minimumprinzip: $\frac{d}{dt} H(t, x_*(t), \lambda(t), u_*(t)) = H_t(t, x_*(t), \lambda(t), u_*(t))$.

Dies wird zu

$$\frac{d}{dt}\left[f(t, x_*(t), \dot{x}_*(t)) - f_u(t, x_*(t), \dot{x}_*(t))\, \dot{x}_*(t)\right] = f_t(t, x_*(t), \dot{x}_*(t)).$$

(2. Euler-Lagrange Gleichnung)

Minimumprinzip und zeitoptimale 0-Steuerung linearer Systeme

Problem: Minimiere

$$J(x,u,t_*) = \int_0^{t_*} 1\,dt = t_*, \qquad \dot{x} = Ax + Bu, \quad x(0) = x_0, \quad x(t_*) = 0, \quad t_* \text{ variabel},$$

und

$$u \in L^\infty([0,t_*],\Omega), \qquad \Omega \text{ konvex und kompakt.}$$

Hamilton-Funktion: $H(t,x,\lambda,u) = 1 + \lambda^\top(Ax + Bu)$.

Minimumbedingung:

$$H(x_*(t),\lambda(t),u_*(t)) = \min_{u\in\Omega} H(x_*(t),\lambda(t),u)$$

$$\Leftrightarrow$$

$$1 + \lambda(t)^\top(Ax_*(t) + Bu_*(t)) = \min_{u\in\Omega}(1 + \lambda(t)^\top(Ax_*(t) + Bu))$$

$$\Leftrightarrow$$

$$\lambda(t)^\top Bu_*(t) = \min_{u\in\Omega} \lambda(t)^\top Bu.$$

Adjungierte DGL:

$$\dot{\lambda}(t) = -\nabla_x H(t,x_*(t),\lambda(t),u_*(t)) = -A^\top\lambda(t)$$

Endbedingung: $\lambda(t_*) = \psi_x(x_*)\nu = \nu$, weil $\psi(x) = x$, also $\psi_x =$ Einheitsmatrix.

Ableitung des Zustands nach der Steuerung

Notation: $x^u(t)$ ist die Lösung des Anfangswertproblems:

$$\dot{x}^u(t) = f(t, x^u(t), u(t)), \qquad x^u(t_0) = x_0$$

Satz: Sei f stetig differenzierbar, $u + \epsilon\, v \in L^\infty([t_0, t_*], \Omega)$. Dann ist

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} x^{u+\epsilon\, v}(t) = \xi^v(t) \qquad (*)$$

Dabei ist ξ^v die Lösung des linearen Anfangswertproblems

$$\dot{\xi}^v(t) = A(t)\, \xi^v(t) + B(t)\, v, \qquad \xi^v(t_0) = 0.$$

mit

$$A(t) = f_x(t, x(t), u(t)), \qquad B(t) = f_u(t, x(t), u(t)).$$

Beweisidee: Nehme $(*)$ als Definition von ξ^v. Unter der Annahme der Differenzierbarkeit von $\epsilon \mapsto x^{u+\epsilon\, v}$ und der Vertauschbarkeit der Ableitungen nach ϵ und t hat man

$$\begin{aligned}
\frac{d}{dt} \underbrace{\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} x^{u+\epsilon\, v}(t)}_{\xi^v(t)} = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \dot{x}^{u+\epsilon\, v}(t) &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} f(t, x^{u+\epsilon\, v}(t), u(t) + \epsilon\, v(t)) \\
&= f_x(t, x^u(t), u(t)) \underbrace{\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} x^{u+\epsilon\, v}(t)}_{\xi^v(t)} + f_u(t, x^u(t), u(t))\, v(t).
\end{aligned}$$

Bemerkung: Die Abbildung $v \mapsto \xi^v$ ist linear.

Der Gradient einer Funktion

Sei $\mathcal{U}$ ein reeller Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$.
$g \in \mathcal{U}$ heisst **Gradient** der Funktion $J : \mathcal{U} \supseteq \mathcal{D} \to \mathbb{R}$ an der Stelle u, wenn

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} J(u+\epsilon\, v) = \langle g, v\rangle$$

für alle v, so dass $u+\epsilon\, v \in \mathcal{D}$ für hinreichend kleine $\epsilon > 0$. Für $\mathcal{U} = L^\infty([t_0,t_*],\mathbb{R}^m)$,

$$\langle g, v\rangle = \int_{t_0}^{t_*} g(t)^\top v(t)\, dt.$$

Klar:

$$\langle g, v\rangle < 0 \qquad \Rightarrow \qquad J(u+\epsilon\, v) < J(u) \quad \text{für hineichend kleine } \epsilon > 0.$$

Folgerung:
Sei u lokaler Minimierer von J. Dann

(a) $\langle g, v\rangle \geq 0$ für alle v mit $u+\epsilon\, v \in \mathcal{D}$ für hinreichend kleine $\epsilon > 0$;
(b) $g = 0$, falls g innerer Punkt von $\mathcal{D}$ (zum Beweis wähle $v = -g$).

Anwendung in der Numerik: Abstiegsverfahren.
Wenn $\langle g, v\rangle < 0$ dann bestimme $\epsilon > 0$ so dass $u_1 := u+\epsilon\, v \in \mathcal{D}$ und $J(u_1) < J(u)$.
Wiederhole das Verfahren mit u_1 anstatt u, dann mit $u_2 = u_1 + \epsilon\, v_1$, ..., und hoffe, dass die so definierte Folge $u, u_1, u_2, \ldots$ gegen einen Minimierer von J konvergiert.
Steilster Abstieg: $v = -g$. Diese Wahl von v ist nicht immer optimal für schnelle Konvergenz (siehe Konjugiertes Gradientenverfahren bei linearen Gleichungssystemen)

Der Gradient des Kostenfunktionals I

Wir schreiben das Kostenfunktional als reine Funktion der Steuerung:

$$J(u) = g(x^u(t_*)) + \int_{t_0}^{t_*} f^0(t, x^u(t), u(t))\, dt, \qquad u \in L^\infty([t_0, t_*], \Omega).$$

wobei

$$\dot{x}^u(t) = f(t, x^u(t), u(t)), \qquad x^u(t_0) = x_0.$$

Hamilton-Funktion: $\quad H(t, x, \lambda, u) = f^0(t, x, u) + \lambda^\top f(t, x, u).$

Satz: Sei λ die Lösung des linearen Endwertwertproblems

$$\dot{\lambda}(t) = \underbrace{-f_x(t, x^u(t), u(t))^\top \lambda(t) - \nabla_x f^0(t, x^u(t), u(t))}_{= -\nabla_x H(t, x^u(t), \lambda(t), u(t))}, \qquad \lambda(t_*) = \nabla g(x(t_*)).$$

Dann gilt

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} J(u + \epsilon\, v) \;=\; \int_{t_0}^{t_*} \big[\underbrace{f_u^0(t, x^u(t), u(t)) + \lambda(t)^\top f_u(t, x^u(t), u(t))}_{= H_u(t, x^u(t), \lambda(t), u(t))}\big]\, v(t)\, dt.$$

Die Funktion

$$t \mapsto H_u(t, x^u(t), \lambda(t), u(t))^\top = \nabla_u H(t, x^u(t), \lambda(t), u(t))$$

ist also der Gradient von J an der Stelle u.

Der Gradient des Kostenfunktionals II

Beweis. Zur Übersichtlichkeit sind für den Beweis unnötige Funktionsargumente weggelassen. Sei ξ die Ableitung von $u \mapsto x^u$ in Richtung v. Dann ist

$$\dot{\xi} = f_x \, \xi + f_u \, v, \qquad \xi(t_0) = 0.$$

Die Ableitung des Kostenfunktionals

$$J(u) = g(x^u(t_*)) + \int_{t_0}^{t_*} f^0(t, x^u, u) \, dt,$$

in Richtung v ist

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} J(u + \epsilon \, v) = \nabla g(x^u(t_*))^\top \xi(t_*) + \int_{t_0}^{t_*} f_u^0 \, v + f_x^0 \, \xi \, dt. \qquad (*)$$

Ziel ist es nun, ξ aus dieser Formel zu entfernen. Es ist

$$\frac{d}{dt}(\lambda^\top \xi) = \dot{\lambda}^\top \xi + \lambda^\top \dot{\xi} = (-\lambda^\top f_x - f_x^0)\xi + \lambda^\top (f_x \, \xi + f_u \, v) = -f_x^0 \, \xi + \lambda^\top f_u \, v.$$

Integrieren dieser Identität ergibt

$$\int_{t_0}^{t_*} f_x^0 \, \xi \, dt = - \left. \lambda^\top \xi \right|_{t=t_0}^{t=t_*} + \int_{t_0}^{t_*} \lambda^\top f_u \, v \, dt = -\lambda(t_*)^\top \xi(t_*) + \int_{t_0}^{t_*} \lambda^\top f_u \, v \, dt.$$

Einsetzen in $(*)$:

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} J(u + \epsilon \, v) \;\; = \;\; (\nabla g(x(t_*)) - \lambda(t_*))^\top \xi(t_*) + \int_{t_0}^{t_*} f_u^0 \, v + \lambda^\top f_u \, v \, dt.$$

Der erste Summand verschwindet, wenn man $\lambda(t_*) = \nabla g(x(t_*))$ setzt $\square$

Diskretisierung und Numerik

Idee zur numerischen Berechnung einer Optimalsteuerung:
Löse die Hamiltonschen DGL für x und λ bei gegebenen u numerisch, und bestimme so eine Näherung des Gradienten. Wende dann ein Abstiegsverfahren an.

Wir diskutieren hier das explizite Eulerverfahren.

Äquidistante Stützpunkte im Zeitintervall:

$$t_i = t_0 + i\,h, \qquad h = (t_* - t_0)/N \qquad i = 0, \ldots, N.$$

x_i, λ_i	Näherung an $x(t_i)$, $\lambda(t_i)$, $i = 0, \ldots, N$
u_i	Näherung an u im Intervall $[t_{i-1}, t_i]$, $i = 1, \ldots, N$.

Explizites Eulerverfahren für das Hamiltonsche Randwertproblem:

$$\begin{aligned} x_0 &= x(t_0), & x_{i+1} &= x_i + h\,f(t_i, x_i, u_i), \\ \lambda_N &= \nabla g(x_N), & \lambda_{i-1} &= \lambda_i + h\,[f_x(x_i, \lambda_i, u_i)^\top \lambda_i + \nabla_x f^0(x_i, \lambda_i, u_i)]. \end{aligned}$$

Diskretisiertes Kostenfunktional:

$$J_h(u_i) := g(x_N) + h \sum_{i=1}^{N} f^0(t_{i-1}, x_{i-1}, u_i),$$

Diskretisierter Gradient:

$$\nabla J_h(u_i) = \begin{bmatrix} \nabla_u f^0(t_0, x_0, u_1) + f_u(t_0, x_0, u_1)^\top \lambda_1 \\ \vdots \\ \nabla_u f^0(t_{N-1}, x_{N-1}, u_N) + f_u(t_{N-1}, x_{N-1}, u_N)^\top \lambda_N \end{bmatrix}.$$