A. Einführung in die Variationsrechnung

David Ebert

13. Oktober 2016

Inhaltsverzeichnis

1 Normen

Definition 1. *Hier einige Normen zur späteren Verwendung:*

$$\|x\|_0 = \max_{t\in[a,b]} |x(t)|$$

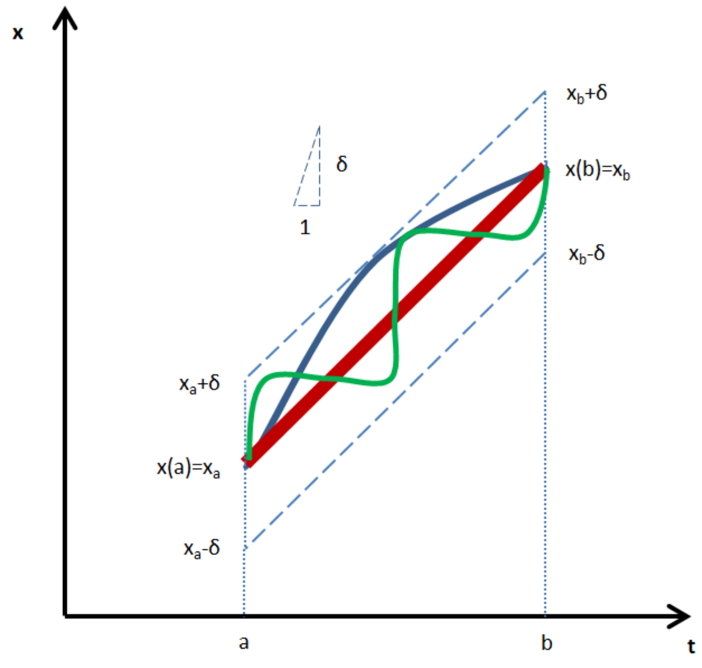
$$\|x\|_1 = \max_{t\in[a,b]} \{|x(t)|, |\dot{x}(t)|\}$$

Definition 2. *Die δ-Umgebung von x_0 ist*

$$U^\delta(x_0) = \{x | \|x - x_0\| < \delta\}$$

Beispiel 3. *Die nachfolgende Abbildung zeigt δ-Umgebungen zur Geraden x_0 zwischen x_a und x_b bezüglich der zwei Normen in $C^1[a,b]$. Die Gerade kann als das Minimum eines Funktionals verstanden werden, das die Länge des "Weges" der Funktion repräsentiert, da die Gerade der kürzeste Weg von x_a nach x_b ist. Die Funktion über der Gerade erfüllt die Umgebungsbedingung zu x_0 für $\|\cdot\|_0$ und $\|\cdot\|_1$, da sowohl die Differenz der Ableitung kleiner als δ bleibt,als auch die Abweichung im Funktionswert gegenüber x_0. Die andere Funktion um die Gerade erfüllt nur die Umgebungsbedingung für $\|\cdot\|_0$.*
Die Funktionsschar in der Umgebung bezüglich $\|\cdot\|_0$ umfasst also mehr Funktionen.

Abbildung 4.



Da wir unter verschiedenen Normen also verschiedenen Funktionsscharen betrachten, sind die Minima von Funktionalen auf den Funktionen der Funktionsscharen im Allgemeinen nicht identisch.

Definition 5. *Ein Minimum, das auf Funktionen aus der Umgebung bezüglich der $\|\cdot\|_0$-Norm gilt, heißt stark, wenn es nur bezüglich der $\|\cdot\|_1$-Norm gilt, schwach.*

Beispiel 6. *Wir betrachten*

$$J[t] = \int_0^1 (\dot{x}^2 + \dot{x}^3)\mathrm{dt} \to \mathrm{Min.}$$

auf $C^1[0,1]$ mit $\|\cdot\|_1$.
Randbedingungen: $x(0) = x(1) = 0$ Das Integral wird minimal für

$$2\dot{x} + 3\dot{x}^2 = \mathrm{const.}$$

oder

$$\dot{x} = \mathrm{const.}$$

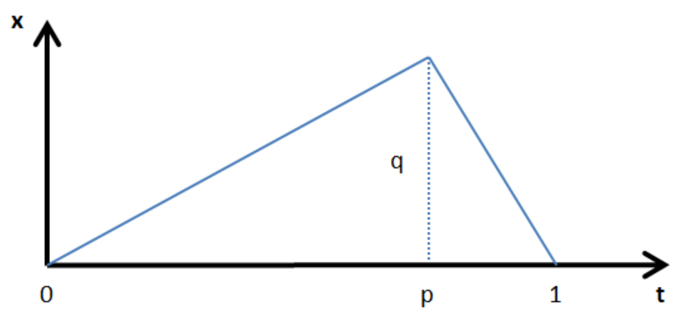
Man benötigt hierfür die Euler'sche Differentialgleichung, die in einem anderen Vortrag (Thema B) behandelt wird. Durch Einsetzen der Randbedingungen liefert die Extremale

$$x_0(t) = 0$$

Beschränkt man sich auf Verleichsfunktionen mit $|\dot{x}| < 1$ für alle $t \in [0,1]$, so ist der Integrand stets positiv, und man erkennt, dass die Extremale das Minimum $J[x_0] = 0$ erzeugt. 0 ist also das Minimum auf dem Vektorraum $C^1[0,1]$.

Jetzt betrachten wir $J[t]$ auf dem Vektorraum $C[0,1]$ mit der Norm $\|\cdot\|_0$. Es sei für das Integral der Weg $x = f(t)$ vorgegeben:

$$f(t) = \begin{cases} \frac{q}{p}t & \text{für } 0 \le t \le p \\ \frac{q}{p-1}(t-1) & \text{für } p < t \le 1 \end{cases}$$



Dann wird

$$J = \underbrace{\int_0^p (\frac{q^2}{p^2} + \frac{q^3}{p^3})\mathrm{d}t}_{=:J_1} + \underbrace{\int_p^1 (\frac{q^2}{(p-1)^2} + \frac{q^3}{(p-1)^3})\mathrm{d}t}_{=:J_2}$$

wobei wir hier zur Einfachheit ignorieren, dass die benötigte Ableitung in der Unstetigkeitsstelle nicht definiert ist.

$$\begin{aligned} J_1 &= \frac{q^2}{p} + \frac{q^3}{p^2} \\ &= \frac{q^2}{p}(1 + \frac{q}{p}) \end{aligned}$$

$$J_2 = \frac{q^2}{1-p}(1 - \frac{q}{1-p})$$

Lässt man p gegen 1 *gehen, so wird* $J_2 < 0$ *werden und* $|J_2| \to \infty$ *gehen, also wird* $J < 0$*, also kleiner als das Minimum* $J[x_0] = 0$*. Die Lösung diese scheinbaren Widerspruchs ist, dass das Minimum* $J[x_0] = 0$ *nur schwach ist.*

2 Ableitungen und Differentiale

2.1 Ableitungsbegriff für normierte Vektorräume

Dieses Kapitel beschäftigt sich damit, wie man in normierten Vektorräumen Ableitungen bilden kann.
Die bekannte Definition für Funktionen $f : \mathbb{R} \to \mathbb{R}$

$$f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} \tag{1}$$

versagt bei Funktionen $f : V \to \mathbb{R}$, weil $h \in V$ ein Vektor ist und man nicht durch Vektoren teilen kann.

Definition 7. *$f'(x_0)$ ist <u>Ableitung</u> von $y = f(x)$ in x_0 dann und nur dann, wenn ein $\delta > 0$ existiert, so dass*

$$f(x_0+h) - f(x_0) = f'(x_0)h + o[h] \tag{2}$$

für alle $h \in V$ mit $\|h\| < \delta$ ist,
wobei im Fall $h \in \mathbb{R}$ gilt

$$\lim_{h\to\infty} \frac{o[h]}{h} = 0$$

Satz 8. *Die Definition (2) und die von (1) sind äquivalent (, falls beide gültig sind).*

Beweis. <u>(2)⇒(1)</u>
Dividieren durch h:

$$\frac{f(x_0+h) - f(x_0)}{h} = f'(x_0) + \frac{o[h]}{h}$$

Grenzübergang:

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} = f'(x_0)$$

wegen

$$\lim_{h\to 0} \frac{o[h]}{h} = 0$$

<u>(1)⇒(2)</u>
Wegen

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} - f'(x_0) = 0$$

folgt für

$$\epsilon_1[h] = \frac{f(x_0+h) - f(x_0)}{h} - f'(x_0)$$

$$\lim_{h\to 0} \epsilon_1[h] = 0$$

Damit folgt 2, wenn man setzt:

$$o[h] = h\epsilon_1[h]$$

□

2.2 Differentiale

Definition 9. *Eine Teilmenge $U \subset V$ eines normierten Vektorraums V heißt offene Teilmenge, wenn es für jedes $x \in U$ ein $\delta > 0$ gibt, so dass $x + h \in U$ für alle $h \in V$ mit $\|h\| < \delta$ gilt.*

Definition 10. *Das Funktional $J[x]$ sei definiert auf einer offenen Teilmenge U des normierten Vektorraums V. Man nennt $\mathrm{d}J(x_0, h)$ Fréchet-Differential (bzw. normal differenzierbar) von $J[x]$ an der Stelle x_0, wenn es ein $\delta > 0$ gibt, so dass für alle $h \in V$ mit $\|h\| < \delta$ gilt*

$$J[x_0 + h] - J[x_0] = \mathrm{d}J[x_0, h] + o[\|h\|]$$

wobei $\mathrm{d}J(x_0, h)$ linear in h und

$$\lim_{\|h\|\to 0} \frac{o[\|h\|]}{\|h\|} = 0$$

ist.

Definition 11. *Das Funktional $J[x]$ sei definiert auf einer offenen Teilmenge U des normierten Vektorraums V. Man nennt $\mathrm{D}J(x_0, h)$ Gateaux-Differential von $J[x]$ an der Stelle x_0 (bzw. es existiert die Richtungsableitung in Richtung h), wenn für $\xi \in \mathbb{R}$*

$$\mathrm{D}J(x_0, h) = \frac{\mathrm{d}}{\mathrm{d}\xi} J[x_0 + \xi h]|_{\xi=0}$$

für alle $h \in V$ existiert.

Satz 12. *Existiert an der Stelle $x_0 \in U \subset V$ das Fréchet-Differential, so existiert auch das Gateaux-Differential, und die beiden sind einander gleich:*

$$\mathrm{d}J(x_0, h) = \mathrm{D}J(x_0, h)$$

Beweis. Nach Definition des Fréchet-Differentials gilt für alle $h \in V$ mit $\|h\| < \delta$:

$$J[x_0 + h] - J[x_0] = \mathrm{d}J(x_0, h) + o[\|h\|] \ , \lim_{\|h\|\to 0} \frac{o[\|h\|]}{\|h\|} = 0$$

Dann gilt für festes, jedoch beliebiges $h \in V$ und alle $\xi \in \mathbb{R}$ mit $\|\xi h\| < \delta$:

$$\frac{J[x_0 + \xi h] - J[x_0]}{\xi} = \frac{\mathrm{d}J(x_0, \xi h)}{\xi} + \frac{o[\xi\|h\|]}{\xi}$$

Wegen der Linearität des Fréchet-Differentials folgt daraus

$$\frac{J[x_0 + \xi h] - J[x_0]}{\xi} = \mathrm{d}J(x_0, h) + \frac{o[\xi\|h\|]}{\xi}$$

und nach Grenzübergang $\xi \to 0$

$$\lim_{\xi\to 0} \frac{J[x_0 + \xi h] - J[x_0]}{\xi} = \mathrm{d}J(x_0, h)$$

oder

$$\mathrm{D}J(x_0, h) = \mathrm{d}J(x_0, h)$$

□

Definition 13. *Das Funktional $J[x]$ sei definiert auf einer offenen Teilmenge U des normierten Vektorraums V.*
Man nennt $\mathrm{D}^2 J(x_0 + \xi h)$ <u>zweites Gateaux-Differential</u> von $J[x]$ an der Stelle x_0, wenn

$$\mathrm{D}^2 J(x_0, h) = \frac{\mathrm{d}^2}{\mathrm{d}^2\xi} J[x_0 + \xi h]|_{\xi=0}$$

für alle $h \in V$ existiert.

3 Extrema

Definition 14. *Sei U eine offene Teilmenge des Vektorraums V. Ein $x_0 \in U$ erzeugt ein relatives Minimum (auch: lokales Minimum) für $J[x]$ in U genau dann, wenn es ein $\delta > 0$ gibt, so dass gilt*

$$J[x_0 + h] - J[x_0] \geq 0$$

für alle $h \in V$ mit $\|h\| < \delta$.

Definition 15. *Sei W Teilraum eines Vektorraums V und U eine Teilmenge von V mit folgender Eigenschaft:*
Für jedes $x \in U$ gibt es ein $\delta \in 0$, so dass auch $x + h \in U$ ist für alle $h \in W$ mit $\|h\| < \delta$. Dann soll U bezüglich W offene oder kurz W-offene Teilmenge von V heißen.

Satz 16. *Sei $f : U \to \mathbb{R}$ eine auf einer offenen Menge zweimal Gateaux-differenzierbare reelle Funktion f. Dann ist das zweite Gateaux-Differenzial $f''(x_0)(x, x)$ bei allen lokalen Minima (Maxima) eine nicht negative (nicht positive) Bilinearform: $f''(x_0)(x, x) \geq 0$ bzw. ≤ 0 für alle $x \in V$.*
Sei $f : U \to \mathbb{R}$ sogar eine auf einer offenen Menge zweimal Fréchet-differenzierbare reelle Funktion f mit zweitem Fréchet-Differential $f''(x_0)(x, x)$. Gibt es einen kritischen Punkt $x_0 \in U$ und ein $\delta > 0$ mit

$$f''(x_0)(x, x) \geq \delta\|x\|^2 \text{ bzw. } f''(x_0)(x, x) \leq -\delta\|x\|^2 \text{ für alle } x \in U^\epsilon(x_0),$$

dann ist der kritische Punkt ein lokales Minimum bzw. Maximum.

Beweis. Wenn x_0 ein lokales Maximum bzw. Minimum von f ist, dann für alle $x \in V$ auch $t = 0$ von $t \mapsto f(x_0 + tx)$.
Deshalb folgt die erste Aussage aus den Eingenschaften isolierter kritischer Punkte im Eindimensionalen.
Zur 2. Aussage:
f ist bei x_0 zweimal Fréchet-differenzierbar. Für ein Minimum gilt also

$$f(x_0 + x) - f(x_0) = \underbrace{f'(x_0)(x)}_{=0} + \underbrace{f''(x_0)(x, x)}_{\geq\delta\|x\|^2} + o[\|x\|^2]$$

Wegen der Definition des $o[\ ]$-Symbols kann man $o[\|x\|^2]$ auf einer Umgebung $U^\epsilon(x_0)$ abschätzen mit $o[\|x\|^2] < \frac{\delta}{2}\|x\|^2$, wobei $\delta > 0$ für kleine ϵ-Umgebungen beliebig klein wird. Wählen wir jetzt, wie gefordert, $f''(x_0)(x, x) \geq \delta\|x\|^2$, so ist $f(x_0 + x) - f(x_0) \geq 0$ in einer Umgebung um $U^\epsilon(x_0)$. Dort ist x_0 also ein Minimum. Der Beweis für Maxima verläuft analog.

□

Bemerkung 17. *Auf endlichdimensionalen Räumen ist die Bedingung $f''(x_0)(x,x) \geq \delta\|x\|^2$ äquivalent zu $f''(x_0)(x,x) > 0$ für alle $x \in V \setminus \{0\}$. In unendlichdimensionalen Räumen nicht. Die zweite Bedingung ist dann auch nicht hinreichend für ein lokales Minimum.*

Um dies für $C[0,1]$, einem unendlichdimensionalen Raum, zu zeigen, hier ein Beispiel:

Beispiel 18. *Sei $f : C([0,1]) \to \mathbb{R},\ x \mapsto \int_0^1 x^2(t)(t - x(t))\mathrm{d}t$.*
Das Gateaux-Differential wird berechnet (Richtung $y(t) \in C[0,1]$):

$$f(x(t) + sy(t)) = \int_0^1 (x(t) + sy(t))^2 t - (x(t) + sy(t))^3 \ \mathrm{d}t$$

$$\frac{\mathrm{d}}{\mathrm{d}s} f(x(t) + sy(t)) = \int_0^1 2(x(t) + sy(t))y(t)t - 3(x(t) + sy(t))^2 y(t) \ \mathrm{d}t$$

$$\frac{\mathrm{d}}{\mathrm{d}^2 s} f(x(t) + sy(t)) = \int_0^1 2y(t)^2 t - 6(x(t) + sy(t))y(t)^2 \ \mathrm{d}t$$

jetzt setzten wir $s = 0$ wegen $\mathrm{D}^2 f(x(t) + sy(t)) = \frac{\mathrm{d}^2}{\mathrm{d}s^2} f[x(t) + sy(t)]|_{s=0}$

$$\mathrm{D}^2 f[x(t), y(t)] = \int_0^1 2y(t)^2 t - 6x(t)y(t)^2 \ \mathrm{d}t = 2\int_0^1 y(t)^2(t - 3x(t)) \ \mathrm{d}t$$

wir wollen die 2. Ableitung in Punkt $f(x = 0)$ betrachten:

$$\mathrm{D}^2 f[0, y(t)] = 2\int_0^1 y(t)^2 t \ \mathrm{d}t$$

Wir haben den Richtungsvektor y genannt, eigentlich interessierte uns aber im Beispiel die Richtung x, daher setzen wir $y = x$.
Dann ist $f''(0)(x,x) = 2\int_0^1 x^2(t)t \ \mathrm{d}t > 0$ für alle $x \in C([0,1]) \setminus \{0\}$. Sei

$$x_\epsilon(t) = \begin{cases} \epsilon - t & \text{für } 0 \leq t < \epsilon \\ 0 & \text{für } \epsilon \leq t \leq 1 \end{cases}$$

mit $\epsilon \in (0,1)$. Dann gilt $\|x\|_0 = \epsilon$ und

$$\begin{aligned} f(sx_\epsilon) &= s^2 \int_0^\epsilon (\epsilon - t)^2 t \mathrm{d}t - s^3 \int_0^\epsilon (\epsilon - t)^3 \mathrm{d}t \\ &= -\frac{s^2}{3}(\epsilon - t)^3 t]_0^\epsilon - (\frac{s^2}{12} - \frac{s^3}{4})(\epsilon - t)^4]_0^\epsilon \\ &= (\frac{s^2}{12} - \frac{s^3}{4})\epsilon^4 \end{aligned}$$

Setzt man jetzt $s = 1$ ein, so sieht man, dass $f(x_\epsilon) = -\frac{1}{6}\epsilon^4$. Das heißt es gibt ein $x_\epsilon \in U^\epsilon(0)$, so dass $f(x_\epsilon) < 0$ für jedes $\epsilon > 0$. Also ist $x = 0$ kein lokales Minimum.

4 Quellen/Literatur

1. Kapitel 1.3 aus E. Klingbeil: „Variationsrechnung“

2. Skript Ana I/II (2013) Uni Mannheim