Einführung in die Variationsrechnung

Ricardo Peña Hoepner
Universität Mannheim

29. April 2012

-Begleitendes Skript zum Seminarvortrag-

Inhaltsverzeichnis

1 Einleitung

Schwerpunkt dieser Arbeit soll die Frage sein, inwiefern sich die in Analysis I und II kennengelernten Stetigkeits- und Differentiationsbegriffe von Funktionen sowie das Vorgehen bei der Extremwertsuche auf den unendlichdimensionalen Raum übertragen lassen. Anhand expliziter Berechnungen sollen die theoretischen, modifizierten Konzepte veranschaulicht werden.

Zunächst möchte ich zur Verdeutlichung dieser grundlegenden Problematik meines Vortrages folgendes Schaubild vorlegen:

(Sei X ein möglicherweise unendlichdimensionaler $\mathbb{R}$-Vektorraum)

Kapitel	**Gegenstand dieses Vortrages**	**Bekanntes Pendant**
1/2	$J : X \to \mathbb{R}$ differenzierbares Funktional	$f : \mathbb{R}^n \to \mathbb{R}$ differenzierbare Funktion
3/4	$\delta J[x_0] \in \mathcal{L}(X, \mathbb{R})$ *Linearität* $\not\Rightarrow$ *Stetigkeit*	$\nabla f(x_0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ Linearität $\Rightarrow$ Stetigkeit
4	x_0 stationärer Punkt: $\delta J[x_0](h) = 0 \quad \forall h$	x_0 kritischer Punkt: $\nabla f(x_0) = 0$
5	$\mathcal{L}(X, \mathbb{R}) = ?$	$\mathcal{L}(\mathbb{R}^n, \mathbb{R}) \simeq \mathbb{R}^n$

Definition 1.1.
Sei X ein $\mathbb{K}$-Vektorraum. Eine Abbildung $J : X \to \mathbb{K}$ heißt <u>Funktional auf X</u>.
J ist *linear*, wenn gilt:

- $J[\lambda x] = \lambda J[x] \qquad \forall x \in X, \lambda \in \mathbb{K}$ *(Homogenität)*
- $J[x + y] = J[x] + J[y] \qquad \forall x, y \in X$ *(Additivität)*

Bemerkung 1.2.

1. X ist typischerweise ein Funktionenraum. Ein Funktional ist also heuristisch formuliert eine „Funktion auf Funktionen“.
2. Im Folgenden wird $\mathbb{K} = \mathbb{R}$ angenommen.

2 Stetigkeit

Satz 2.1.
Sei X ein normierter $\mathbb{R}$-Vektorraum und $J : X \to \mathbb{R}$ ein lineares Funktional. J ist genau dann stetig, wenn es beschränkt ist, also wenn es eine Konstante $C \in \mathbb{R}$ gibt, so dass $\forall x \in X$ gilt:

$$|J[x]| \leq C\|x\|_X$$

Beweis.

> Einschub (Wiederholung Satz 9.30, Analysis II)
>
> Für eine Abbildung $f : X \to Y, x \mapsto f(x)$ zwischen metrischen Räumen (X, d) und (Y, e) ist folgendes äquivalent:
>
> - f ist stetig in x.
> - Für jede Folge $(x_k)_{k\in\mathbb{N}}$ in (X, d), die gegen x konvergiert, konvergiert auch die Folge $(f(x_k))_{k\in\mathbb{N}}$ in (Y, e) gegen $f(x)$.

„$\Rightarrow$"
Sei J stetig.
Annahme: J ist nicht beschränkt.
Dann gibt es für jedes $k \in \mathbb{N}$ ein $z_k \in X$ sodass $|J[z_k]| > k\|z_k\|_X$ gilt. Man erhält eine Folge $(x_k)_{k\in\mathbb{N}}$ in X mit $x_k := \frac{z_k}{\|z_k\|_X}$ und folgenden Eigenschaften:

$$\|x_k\|_X = 1 \text{ und } |J[x_k]| \geq k \quad \forall k \in \mathbb{N} \tag{2.1}$$

Man betrachte nun die Folge $(y_k)_{k\in\mathbb{N}}$ mit $y_k := \frac{x_k}{k}$.
Es gilt:

$$|J[y_k]| = |J[\frac{x_k}{k}]| \overset{\text{J linear}}{=} \frac{1}{k}|J[x_k]| \overset{(2.1)}{\geq} \frac{1}{k}k = 1 \quad \forall k \in \mathbb{N} \tag{2.2}$$

Gleichzeitig gilt aber:

$$y_k \to 0 \text{ für } k \to \infty,$$

denn $\forall \epsilon > 0 \; \exists N_\epsilon \in \mathbb{N}$, so dass $\forall k > N_\epsilon$ gilt:

$$\|y_k - 0\| = \|\frac{x_k}{k}\| = \frac{1}{k}\|x_k\| \overset{(2.1)}{=} \frac{1}{k} < \epsilon.$$

Nach Satz 9.30 (Einschub) folgt aus der Stetigkeit von J:

$$J[y_k] \overset{k\to\infty}{\longrightarrow} J[0] \overset{\text{J linear}}{=} 0.$$

Das ist aber ein Widerspruch zu (2.2). Also muss die Annahme falsch und somit J beschränkt sein.

„$\Leftarrow$“
Sei J beschränkt und $(x_k)_{k\in\mathbb{N}}$ eine in X konvergente Folge mit Grenzwert $x \in X$. Dann gilt:

$$|J[x_k] - J[x]| \overset{\text{J linear}}{=} |J[x_k - x]| \overset{\text{J beschränkt}}{\leq} C\|x_k - x\|_X \text{ mit } C \in \mathbb{R} \text{ konstant}$$

Wegen

$$0 \leq |J[x_k] - J[x]| \leq C\|x_k - x\| \overset{k\to\infty}{\longrightarrow} 0$$

folgt

$$J[x_k] \overset{k\to\infty}{\longrightarrow} J[x].$$

Aus Satz 9.30 (Einschub) folgt die Stetigkeit von J.

q.e.d.

Beispiel 2.2.
Seien $X = C[a,b]$ und $a, b \in \mathbb{R}$ mit $a < b$.
Außerdem sei $\|x\|_X = \|x\|_\infty := \sup\limits_{t\in[a,b]} |x(t)|$.
$(X, \|\cdot\|_\infty)$ ist ein normierter Vektorraum über $\mathbb{R}$, denn:

(i) $\|x\|_\infty = \sup\limits_{t\in[a,b]} \underbrace{|x(t)|}_{\geq 0} \geq 0$ und $\|x\|_\infty = 0 \Leftrightarrow x(t) = 0 \quad \forall t \in [a,b]$

(ii) $\|\lambda x\|_\infty = \sup\limits_{t\in[a,b]} |\lambda x(t)| = \sup\limits_{t\in[a,b]} |\lambda||x(t)| = |\lambda| \sup\limits_{t\in[a,b]} |x(t)| = |\lambda|\|x\|_\infty$

(iii) $$\begin{aligned}\|x+y\|_\infty &= \sup_{t\in[a,b]} |x(t) + y(t)| \leq \sup_{t\in[a,b]} |x(t)| + |y(t)| \\ &\leq \sup_{t\in[a,b]} |x(t)| + \sup_{t\in[a,b]} |y(t)| = \|x\|_\infty + \|y\|_\infty\end{aligned}$$

Betrachte jetzt das Funktional

$$J : X \to \mathbb{R} \text{ mit } J[x] := \int_a^b \alpha(t)x(t)\,dt$$

mit einer festen Funktion $\alpha \in X$. Wir wollen jetzt J auf Linearität und Stetigkeit untersuchen:

1. *Linearität*

   Seien $\lambda \in \mathbb{R}$ und $x, y \in X$ beliebig.

   $$\text{(i) } J[\lambda x] = \int_a^b \alpha(t)\lambda x(t)\,dt = \lambda \int_a^b \alpha(t)x(t)\,dt = \lambda J[x]$$

   $$\begin{aligned}\text{(ii) } J[x+y] &= \int_a^b \alpha(t)(x(t)+y(t))\,dt \\ &= \int_a^b (\alpha(t)x(t)+\alpha(t)y(t))\,dt \\ &= \int_a^b \alpha(t)x(t)\,dt + \int_a^b \alpha(t)y(t)\,dt \\ &= J[x]+J[y]\end{aligned}$$

2. *Stetigkeit*

   Nach Satz 2.1 reicht es aus, die Beschränktheit der Abbildung zu zeigen:
   zz.: $\exists C = \text{const.} \in \mathbb{R}$ mit $|J[x]| \leq C\|x\|_\infty \quad \forall x \in X$

   $$\begin{aligned}|J[x]| &= |\int_a^b \alpha(t)x(t)\,dt| \leq \int_a^b |\alpha(t)x(t)|\,dt = \int_a^b |\alpha(t)||x(t)|\,dt \leq \int_a^b |\alpha(t)| \sup_{s\in[a,b]} |x(s)|\,dt \\ &= \int_a^b |\alpha(t)|\|x\|_\infty\,dt = \|x\|_\infty \underbrace{\int_a^b |\alpha(t)|\,dt}_{:=C} = C\|x\|_\infty\end{aligned}$$

   C existiert, da α eine stetige Funktion auf dem kompakten Intervall $[a,b]$ ist.

$\Rightarrow J$ ist eine lineare und stetige Funktion.

Bemerkung 2.3.
Wie sich im Folgenden zeigen wird, ist die separate Untersuchung von Linearität und Stetigkeit gerechtfertigt. Wir werden später feststellen, dass die Implikation Linearität $\Rightarrow$ Stetigkeit im Unendlichdimensionalen nicht gegeben ist.

3 Die erste Variation

Definition 3.1. [Analogon zur Ableitungsdefinition]
Sei X ein normierter Vektorraum und $U \subset X$ eine offene Teilmenge. Das Funktional $J : U \longrightarrow \mathbb{R}$ heißt in $x_0 \in X$ differenzierbar, wenn es ein $A \in \mathcal{L}(X, \mathbb{R})$ gibt, sodass folgende Abbildung in x_0 stetig ist:

$$U \longrightarrow \mathbb{R}, x \mapsto \begin{cases} \frac{|J[x]-J[x_0]-A[x-x_0]|}{\|x-x_0\|_X} & \text{für } x \neq x_0 \\ 0 & \text{für } x = x_0 \end{cases} \tag{3.1}$$

A heißt <u>erste Variation von J</u> in x_0 und wird mit $\delta J[x_0]$ bezeichnet.

Satz 3.2.
Existiert ein solches A aus Definition 3.1, so ist es eindeutig.

<u>Beweis</u>.
Wenn A′ ebenfalls diese Bedingung erfüllt, dann gibt es für jedes $\frac{\varepsilon}{2} > 0$ ein $\delta^* := min\{\delta, \delta'\} > 0$, sodass gilt:

$$\begin{aligned} \frac{|(A-A')[x-x_0]|}{\|x-x_0\|_X} &\leq \frac{|J[x]-J[x_0]-A[x-x_0]|}{\|x-x_0\|_X} + \frac{|J[x]-J[x_0]-A'[x-x_0]|}{\|x-x_0\|_X} \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall x \in B(x_0, \delta^*) \end{aligned}$$

Daraus folgt $\forall \varepsilon > 0$:

$$\|A - A'\| = \sup_{x-x_0 \in X} \frac{|(A-A')[x-x_0]|}{\|x-x_0\|_X} = \sup_{x-x_0 \in X} \frac{|(A-A')[\frac{(x-x_0)\delta^*}{2\|x-x_0\|_X}]|}{\|\frac{(x-x_0)\delta^*}{2\|x-x_0\|_X}\|_X} < \varepsilon,$$

da $\frac{(x-x_0)\delta^*}{2\|x-x_0\|_X} \in B(0, \delta^*)$. Also gilt $\|A - A'\| = 0$. Da die Operatornorm eine Norm ist, folgt:

$$\|A - A'\| = 0 \Leftrightarrow A - A' = 0 \Leftrightarrow A = A'$$

q.e.d.

Proposition 3.3.
Ist $J \in \mathcal{L}(X, \mathbb{R})$ so gilt

$$\delta J[x_0] = J \quad \forall x_0 \in X$$

Beweis.
Sei $x_0 \in X$ beliebig. Es gilt:

$$\frac{|J[x] - J[x_0] - J[x - x_0]|}{\|x - x_0\|_X} \overset{J \in \mathcal{L}(X,\mathbb{R})}{=} \frac{|J[x] - J[x_0] - J[x] + J[x_0]|}{\|x - x_0\|_X} = 0.$$

Offensichtlich ist (3.1) damit in x_0 stetig. Wegen der Eindeutigkeit der Variation folgt die Aussage. **q.e.d.**

Bemerkung 3.4. [Richtungsableitung]
Sei X ein normierter Vektorraum und $U \subset X$ eine offene Teilmenge. Ist J in $x_0 \in U$ differenzierbar, so gilt für $h \in X$

$$\delta J[x_0](h) = \frac{d}{ds} J[x_0 + sh]\bigg|_{s=0} = \lim_{s \to 0} \frac{J[x_0 + sh] - J[x_0]}{s}.$$

$\delta J[x_0](h)$ heißt die erste Variation von J an der Stelle x_0 in Richtung h.

Beispiel 3.5 (Fortsetzung von Beispiel 2.2)**.**
Sei $X = C[a, b]$ mit $a, b \in \mathbb{R}$ und $a < b$.
Außerdem sei $\|x\|_\infty := \sup\limits_{t \in [a,b]} |x(t)|$ und

$$J : X \to \mathbb{R} \ \text{ mit } \ J[x] := \int\limits_a^b \alpha(t) x(t)\, dt$$

Nun wollen wir die erste Variation explizit berechnen:

$$\begin{aligned}
\delta J[x_0](h) &= \frac{d}{ds} J[x_0 + sh]\bigg|_{s=0} \\
&= \frac{d}{ds} \int\limits_a^b \alpha(t)(x_0(t) + sh(t))\, dt\bigg|_{s=0} \\
&= \frac{d}{ds} \int\limits_a^b \alpha(t) x_0(t) + \alpha(t) s h(t)\, dt\bigg|_{s=0} \\
&= \frac{d}{ds} \int\limits_a^b \alpha(t) x_0(t)\, dt\bigg|_{s=0} + \frac{d}{ds} s \int\limits_a^b \alpha(t) h(t)\, dt\bigg|_{s=0} \\
&= \int\limits_a^b \alpha(t) h(t)\, dt
\end{aligned}$$

Dieses Ergebnis überrascht nicht. Aus Beispiel 2.2 wissen wir bereits, dass $J \in \mathcal{L}(X, \mathbb{R})$ und aus Proposition 3.3 folgt, dass $\delta J[x_0] = J \;\; \forall x_0 \in X$ gelten muss.

Beispiel 3.6.
Gegeben sei $X = C[0,1]$ und das nichtlineare Funktional

$$J : X \to \mathbb{R} \;\text{ mit }\; J[x] := \int_0^1 x(t)^2(t - x(t))\, dt.$$

Wir berechnen die erste Variation:

$$\begin{aligned}
\delta J[x_0](h) &= \frac{d}{ds} J[x_0 + sh]\Big|_{s=0} \\
&= \frac{d}{ds} \int_0^1 (x_0(t) + sh(t))^2 (t - x_0(t) - sh(t))\, dt\Big|_{s=0} \\
&= \frac{d}{ds} \int_0^1 (x_0(t)^2 + 2x_0(t)sh(t) + s^2h(t)^2)(t - x_0(t) - sh(t))\, dt\Big|_{s=0} \\
&= \frac{d}{ds} \int_0^1 x_0(t)^2 t - x_0(t)^3 - sx_0(t)^2h(t) + 2sth(t)x_0(t) - 2sh(t)x_0(t)^2 \\
&- 2s^2h(t)^2x_0(t) + s^2h(t)^2t - s^2h(t)^2x_0(t) - s^3h(t)^3\, dt\Big|_{s=0} \\
&= \int_0^1 -x_0(t)^2h(t) + 2th(t)x_0(t) - 2h(t)x_0(t)^2 \\
&- 4sh(t)^2x_0(t) + 2sh(t)^2t - 2sh(t)^2x_0(t) - 3s^2h(t)^3\, dt\Big|_{s=0} \\
&= \int_0^1 2th(t)x_0(t) - 3h(t)x_0(t)^2\, dt \\
&= \int_0^1 (-3x_0(t)^2 + 2tx_0(t))h(t))\, dt
\end{aligned}$$

4 Notwendige Bedingung von Extrema

Im Folgenden sei $J : U \subset X \to \mathbb{R}$ ein differenzierbares Funktional.

Definition 4.1.
Ein Element $u \in U$ heißt stationärer Punkt von J, falls gilt:

$$\delta J[u](h) = 0 \quad \forall h$$

Definition 4.2.
J besitzt ein lokales Maximum *(bzw. Minimum)* in x_0, wenn es ein $\epsilon > 0$ gibt, sodass für alle $x \in B(x_0, \epsilon) \subset X$ gilt:

$$J[x] \leq J[x_0] \qquad (\textit{bzw. } J[x] \geq J[x_0]) \tag{4.1}$$

Ein lokales Maximum *(bzw. Minimum)* ist ein globales, wenn (4.1) $\forall x \in X$ gilt.

Satz 4.3.
Wenn J ein lokales Extremum in x_0 besitzt, dann ist x_0 ein stationärer Punkt von J.

Beweis.
Angenommen, J nimmt in x_0 ein Maximum an. Dann gilt:

$$\left.\begin{aligned} \lim_{s \downarrow 0} \frac{J[x_0 + sh] - J[x_0]}{s} \leq 0 \\ \lim_{s \uparrow 0} \frac{J[x_0 + sh] - J[x_0]}{s} \geq 0 \end{aligned}\right\} \Rightarrow \lim_{s \to 0} \frac{J[x_0 + sh] - J[x_0]}{s} = \delta J[x_0](h) = 0$$

Analoge Beweisführung bei einem Minimum in x_0. **q.e.d.**

5 Die Problematik unendlichdimensionaler Räume

Definition 5.1.
Sei X ein normierter $\mathbb{R}$-Vektorraum. Der Vektorraum der linearen Funktionale auf dem Vektorraum X heißt (algebraischer) Dualraum und wird mit

$$X^* := \{F : X \to \mathbb{R} \mid F \text{ ist linear}\}$$

bezeichnet. Der Vektorraum der linearen und stetigen Funktionale auf dem Vektorraum X heißt topologischer Dualraum und wird mit

$$X' := \{G : X \to \mathbb{R} \mid G \text{ ist linear und stetig}\} = \mathcal{L}(X, \mathbb{R})$$

bezeichnet. X' ist zusammen mit der Operatornorm $\|x'\|_{X'} := \sup\limits_{\|x\|_X=1} |x'(x)|$ ein Banachraum.

Bemerkung 5.2.
Für endlichdimensionale Räume gilt $X^* = X'$, da aus Linearität Stetigkeit folgt.
Im Unendlichdimensionalen gilt fast immer $X' \subsetneq X^*$, da i.A. Linearität $\not\Rightarrow$ Stetigkeit (siehe folgendes Beispiel).

Beispiel 5.3.
Betrachte

$$J : X \to \mathbb{R}, x \mapsto x'(0) \text{ mit } X = C^1[0,1] \text{ und der Norm } \|\cdot\|_\infty.$$

Da der Differentialoperator linear ist (vgl. u.a. Analysis II), ist J eine lineare Abbildung. Wir wollen sie nun auf Stetigkeit prüfen und konstruieren uns dazu folgende Folge$(x_k)_{k\in\mathbb{N}}$ auf X:

$$x_k(t) := \sin(kt + 2\pi) \in X \quad \forall k \in \mathbb{N}$$

Offensichtlich gilt

$$\|x_k\|_\infty = 1 \quad \forall k \in \mathbb{N}$$

Allerdings stellt man fest, dass

$$|J[x_k]| = |x_k'(0)| = |k \underbrace{\cos(2\pi)}_{=1}| = k \quad \forall k \in \mathbb{N}$$

Daraus folgt

$$\Rightarrow \forall C = const. \in \mathbb{R} \quad \exists N \in \mathbb{N} \quad \forall k \geq N : |J[x_k]| = k \geq C = C\|x_k\|_\infty$$

J ist also unbeschränkt und damit nach Satz 2.1 nicht stetig.

$$\Rightarrow J \in X^* \text{ aber } J \notin X'.$$

Satz 5.4.
Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum.
Dann gilt

$$\mathcal{L}(V,\mathbb{R}) \simeq V.$$

Beweis.
Man nehme zunächst an, dass $V = \mathbb{R}^n$ gilt. Dann kann man einen Isomorphismus über das Skalarprodukt definieren:

$$I : \mathbb{R}^n \to (\mathbb{R}^n)', w \mapsto I(w) \text{ mit } I(w) : \mathbb{R}^n \to \mathbb{R}, v \mapsto (I(w))(v) := \langle w, v\rangle$$

Wir wollen jetzt die Bijektivität der Abbildung I zeigen.
Injektivität:
Seien v und w aus $\mathbb{R}^n$.

$$\begin{aligned} I(v) = I(w) &\Rightarrow I(v)(x) = I(w)(x) \quad \forall x \in \mathbb{R}^n \\ &\Rightarrow \langle v, x\rangle = \langle w, x\rangle \quad \forall x \in \mathbb{R}^n \\ &\Rightarrow \langle v - w, x\rangle = 0 \qquad \forall x \in \mathbb{R}^n \\ &\Rightarrow v - w = 0 \\ &\Rightarrow v = w \end{aligned}$$

Surjektivität:
Sei $\phi \in (\mathbb{R}^n)' = \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ beliebig, d.h. von der Gestalt

$$\phi(x) = a_1 x_1 + \ldots + a_n x_n = \langle a, x\rangle$$

mit einem eindeutigen $a \in \mathbb{R}^n$.
Dann gilt:

$$I(a) = \phi$$

Da das Skalarprodukt eine Bilinearform ist, ist $I(w)$ für gegebenes w eine lineare Abbildung und aufgrund der Endlichdimensionalität auch stetig.

Ist V allgemein endlichdimensional, so lässt sich V über einen anderen Isomorphismus mit $\mathbb{R}^n$ identifizieren. **q.e.d.**

Bemerkung 5.5.
Im unendlichdimensionalen Fall gilt Satz 5.4 im Allgemeinen nicht. Es gibt jedoch Fälle, in denen man auch hier mithilfe eines Skalarproduktes einen Isomorphismus analog zum Beweis von Satz 5.4 konzipieren kann:

Ist V ein Hilbertraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$, so ist folgende Abbildung ein Isomorphismus:

$$j : V \to V', w \mapsto j(w) \text{ mit } j(w)(v) := \langle w, v\rangle$$

Betrachtet man beispielsweise den Hilbertraum L^2, so wird über sein Skalarprodukt

$$(f,g) \mapsto \int fg \, d\mu$$

ein Isomorphismus definiert:

Erinnerung: Satz 12.57 Analysis II

Für alle messbaren Mengen $A \subset \mathbb{R}^d$ und alle $1 \leq p,q \leq \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$ ist folgende Abbildung injektiv:

$$j : L^q(A) \to \mathcal{L}(L^p(A), \mathbb{K}), g \mapsto j(g) \;\text{ mit }\; j(g) : L^p(A) \to \mathbb{K}, f \mapsto \int_A fg d\mu$$

Fakt: Für $1 < q \leq \infty$ ist j sogar ein Isomorphismus von Banachräumen.
Wähle $p = q = 2$, dann folgt also

$$\mathcal{L}(L^2(A), \mathbb{K}) \simeq L^2(A).$$

Literatur

[1] Prof. Martin U. Schmidt - Skript zur Vorlesung Analysis II. Universität Mannheim. HS 2010/FS 2011

[2] Jean Dieudonne - Grundzüge der modernen Analysis 2. Braunschweig: Vieweg. 1975