ENTSCHEIDUNGSTHEORETISCHE GRUNDLAGEN

DER

STOCHASTISCHEN OPTIMIERUNG

von

Kurt Marti

Nr.2 (1970)

Einleitung

Es seien X,Z (separierte) topologische Vektorräume und X_+ ein Kegel in X. Unser Ausgangspunkt ist die "deterministische" Optimierungsaufgabe

Bestimme

$$\inf \{cx \mid Ax=b, x \geq \theta\},$$

wobei $A \in L(X,Z)$ $b \in Z$ und $c \in X^*_+$. Sind nun die "Daten" (A,b,c) stochastische Grössen, d.h. ist (A,b,c) eine messbare Abbildung

$$(A,b,c) : \Omega \longrightarrow L(X,Z)\times Z\times X^*$$

wobei Ω Träger eines Wahrscheinlichkeits(W-)raumes $(\Omega, \mathfrak{A}, P)$ und $c \in X^*_+$ P-fast sicher ist, so heisst (A,b,c) ein stochastisches Programm und $(A(\omega), b(\omega), c(\omega))$, $\omega \in \Omega$ eine Realisierung dieses Programms.

Eine Entscheidung, d.h. die Wahl eines Elements x aus X oder aus einer Teilmenge X_0 von X führt bei einer darauf folgenden Realisierung dieses Programms zum Ergebnis

$$e = e(\omega,x) = (c(\omega)x, A(\omega)x-b(\omega)).$$

Damit stehen wir vor dem Problem der Wahl einer Entscheidungsregel, wozu wir zumindest eine Ordnung auf der Menge

$$\mathbb{E} = \{e(\omega,x) \mid \omega \in \Omega, x \in X_0\}$$

der Ergebnisse benötigen. Da wir später im wesentlichen die Bayes-Regel verwenden werden, stellt sich das weitere Problem der Existenz einer (zumindest) messbaren Nutzenfunktion der verwendeten Ordnung auf $\mathbb{E}$.

Die Erläuterung der in der vorliegenden Arbeit verwendeten Zeichen und Begriffe befindet sich am Schluss der jeweiligen Seite oder im Anhang.

I. Entscheidungstheoretische Grundlagen

§ 1. Ordnungen und Skalen

Unter einer Ordnung (geordneten Menge) $(X,\sim,<)$ verstehen wir eine Menge X mit zwei binären Relationen "$\sim,<$", die den Axiomen

O1. Für alle $(x,y) \in X\times X$ gilt genau eine der Relationen $x<y$, $y<x$, $x\sim y$

O2. "$\sim$" ist eine Äquivalenzrelation

O3. "$<$" ist transitiv

genügen.

Von grosser Bedeutung sind die Ordnungen für die ein Homomorphismus von $(X,\sim,<)$ in $(\mathbb{R},=,<)$ existiert:

Def. 1.1. Eine Skala einer Ordnung $(X,\sim,<)$ ist eine Abbildung $m: X \longrightarrow R$ mit den Eigenschaften

$$x<y \implies m(x)<m(y)$$
$$x\sim y \implies m(x)=m(y)$$

Bemerkung. Ist m Skala der Ordnung $(X,\sim,<)$, dann gelten die Implikationen

$$x<y \iff m(x)<m(y)$$
$$x\sim y \iff m(x)=m(y)$$

Zur Definition von Entscheidungsregeln benötigt man Skalen, die bis auf gewisse lineare Transformationen eindeutig bestimmt sind:

Def. 1.2. Es sei S die Menge der Skalen einer Ordnung $(X,\sim,<)$ und $\mathcal{E}$ eine Teilmenge von S ($\mathcal{E}$ repräsentiert eine gewisse für $m \in S$ definierte (zusätzliche) Eigenschaft. Eine Skala m dieser Ordnung mit $m \in \mathcal{E}$ heisst $\mathcal{E}$ - Nutzenfunktion, falls sich ein beliebiges $m' \in \mathcal{E}$ von m höchstens um ein Element aus $\mathbf{T} = \{\tau \mid \tau(t)=\alpha t+\beta,\ t \in \mathbb{R};\ \alpha,\beta \in \mathbb{R},\ \alpha>0\}$ unterscheidet.

$\mathbb{R}$ = Menge der reellen Zahlen; Statt $x<y$ schreiben wir auch $y>x$.

Im Zusammenhang mit dem Problem der Existenz einer Nutzenfunktion brauchen wir die folgenden Begriffe und Aussagen (siehe [8]):

Def. 1.3. Eine Ordnung $(X,=,<)$ heißt ordnungsvollständig, falls jede nichtleere, nach oben beschränkte Teilmenge von X ein Supremum besitzt.

Def. 1.4. Ein abgeschlossenes Intervall $[a,b]=\{x \mid a \leqslant x \leqslant b\}$ einer Ordnung $(X,=,<)$ heißt Lücke, falls $a<b$ und das offene Intervall $(a,b)=\{x \mid a<x<b\}=\emptyset$ ist.

Def. 1.5. Die Intervalltopologie einer Ordnung $(X,=,<)$ ist die von den Intervallen $\{x \mid x>a\}, \{x \mid x<a\}, a \in X$, erzeugte Topologie (auf X).

Def. 1.6. Eine Operation $o: X \times X \to X$ auf einer geordneten Menge $(X,=,<)$ heißt metrisch, falls die Axiome

M1. $(aob)o(cod)=(aoc)o(bod)$ für beliebige $a,b,c,d \in X$
M2. $x \mapsto xoa, x \in X$ und $x \mapsto aox$, $x \in X$ sind für jedes $a \in X$ eineindeutige und bezüglich der Intervalltopologie stetige Operatoren.

erfüllt sind.

Def. 1.7. Eine Operation $\oplus: X \times X \to X$ auf einer geordneten Menge $(X,=,<)$ heißt additiv, falls die Axiome

A1. $\oplus$ ist kommutativ und assoziativ
A2. $x \mapsto x \oplus a, x \in X$ ist für jedes $a \in X$ ein eineindeutiger und bezüglich der Intervalltopologie stetiger Operator.

gelten.

Offenbar ist jede additive Operation metrisch.

$\emptyset$ = leere Menge

<u>Satz 1.1.</u> Eine geordnete Menge $(X,=,\sim)$ ist genau dann zusammenhängend (in bezug auf die Intervalltopologie), falls $(X,=,\sim)$ ordnungsvollständig ist und keine Lücken hat.

Beweis: Pfanzagel, Theory of Measurement.

<u>Satz 1.2.</u> Es sei $(X,=,<)$ eine geordnete, zusammenhängende Menge, die mindestens zwei Elemente hat. "o" sei eine Operation $o: X\times X \rightarrow X$. Es existiert genau dann eine (bezüglich der Intervalltopologie) stetige Skala m der Ordnung $(X,=,<)$ mit der Eigenschaft $m(aob)=\alpha m(a)+\beta m(b)+\gamma$, wobei α,β,γ gewisse reelle Konstante sind, wenn "o" eine metrische Operation ist. Die Skala m ist eindeutig bestimmt bis auf positive lineare Transformationen $\tau \in \mathbb{T}$. Die Konstanten α,β sind eindeutig bestimmt; es gilt $\alpha=\beta=1$ genau dann, wenn "o" assoziativ ist.

Beweis: Pfanzagel, Theroy of Measurement.

§ 2. Darstellung von Ordnungen

In dem in der Einleitung beschriebenen Modell ist $\theta=\theta_{\mathbb{R}\times Z}$ das bestmögliche Ergebnis (keine Kosten, keine Verletzung der Restriktion). Man wird daher von zwei Ergebnissen e_1, e_2 dasjenige vorziehen, das in einem im folgenden zu präzisierenden Sinn "näher" bei θ liegt. Es sei dazu X eine Menge, θ ein festes Element von X und σ eine Kette (in bezug auf die mengentheoretische Inklusion) in X mit $\theta \in \bigcap\sigma$. Durch

$$x<y \iff \text{es gibt ein } U\in\sigma \text{ mit } x\in U \text{ und } y\notin U$$

$$x\sim y \iff x\nless y \text{ und } x\ngtr y ,$$

wobei $x, y\in X$, werden zwei Relationen auf X definiert. Wie man leicht zeigt, gilt der

<u>Satz 2.1.</u> $(X,\sim,<)$ ist eine Ordnung mit $\theta \lesssim x$ für alle $x\in X$.

Wir nennen "$\sim,<$" die durch σ erzeugte Ordnung.

§ 3. Existenz von Skalen

Es sei nun X ein linearer Raum, $\theta=\theta_X$, σ eine Kette in X mit $\theta\in\bigcap\sigma$ und e ein weiteres (festes) Element von X. Durch

$$q(S) = \sup\{\lambda\geq o \mid \lambda e \in S\},\ S\in\sigma \qquad (1)$$

und

$$m(x) = \begin{cases} \inf\{q(S) \mid x\in S\}, & \{S \mid x\in S\}\neq\emptyset \\ +\infty & , \{S \mid x\in S\}=\emptyset \end{cases},\ x\in X \qquad (2)$$

definieren wir zwei Funktionale $q : \sigma \longrightarrow \bar{\mathbb{R}}$ und $m : X \longrightarrow \bar{\mathbb{R}}$. Zur Untersuchung des Zusammenhangs zwischen m und der durch σ erzeugten Ordnung sei zunächst $x\sim y$. Wegen $\{U\in\sigma \mid x\in U\} = \{V\in\sigma \mid y\in V\}$ gilt $x\sim y \Longrightarrow m(x)=m(y)$. Es sei nun $x<y$. Demnach existiert ein $U_o\in\sigma$ mit $x\in U_o$ und $y\notin U_o$. Ist $V\in\sigma$ mit $y\in V$, dann gilt $U_o\subset V$ und daher

$$m(x) \leqslant q(U_o) \leqslant q(V).$$

Daraus folgt im Fall $\{V\in\sigma \mid y\in V\}\neq\emptyset$ die Ungleichung $m(x)\leqslant m(y)$. Da für $\{V\in\sigma \mid y\in V\}=\emptyset$ nach Definition $m(y)=+\infty$ ist, ergibt sich die Implikation $x<y \Longrightarrow m(x)\leqslant m(y)$. Damit m eine Skala der Ordnung $(X,\sim,<)$ ist, benötigen wir aber die Implikation $x<y \Longrightarrow m(x)<m(y)$. Wie das folgende Beispiel 1 zeigt, folgt diese nicht aus den bisherigen Voraussetzungen:
Es sei dazu $X = \mathbb{R}^2$, $\sigma=\{S(\theta,1),\ S((-1,0),2)\}$ ($S(x,\varepsilon) = \{y\in\mathbb{R}^2 \mid \|y-x\| < \varepsilon\}$) und $e=(1,0)$. Für $x=\theta$ und $y = (-2,0)$ gilt dann $x<y$ und $m(x)=1=m(y)$.

Wir benötigen daher die zusätzliche Bedingung

(M) q ist streng monoton (bezüglich "$\subsetneqq$")

Leider ist, wie das folgende Beispiel 2 zeigt, (M) noch nicht hinreichend für die oben verlangte Implikation.

$\bar{\mathbb{R}} = \mathbb{R}\cup\{+\infty,-\infty\}$

Es sei $X=\mathbb{R}^2$, $\sigma=\{S_n|n=0,1,2,\dots\}$ mit $S_0=\{x|\ |x_1|+|x_2|<1\}$ und $S_n = [-1,1+\frac{1}{n}]\times[-1,1]$, $n=1,2,\dots$ sowie $e=(1,0)$. Wegen $q(S_0)=1$ und $q(S_n)=1+\frac{1}{n}$, $n\geq 1$ ist (M) erfüllt. Für $x=(0,0)$ und $y=(0.6,0.6)$ gilt dann $x\in S_0$, $y\notin S_0$ und $y\in S_n$, $n\geq 1$, also $x<y$ und $m(x)=1=m(y)$.
Das obige Beispiel legt nun aber die folgende Bedingung

(T) Gilt $x<y$, dann enthält $\{S\in\sigma|x\in S, y\notin S\}$ mindestens zwei Elemente.

nahe.
Wie man nun leicht sieht, ergeben (M) und (T) zusammen eine hinreichende Bedingung für die gewünschte Implikation. Ist nämlich $x<y$, dann existieren nach (T) zwei Elemente $U_1,U_2\in\sigma$ mit $x\in U_i$ und $y\notin U_i$, $i=1,2$. Es sei $U_1\subset U_2$. Daraus folgt dann

$$\begin{aligned} m(x) &\leq q(U_1) \\ &< q(U_2) \qquad \text{(M)} \\ &\leq q(V) \end{aligned}$$

für alle $V\in\sigma$ mit $y\in V$ und daher $m(x)<m(y)$.
Es gilt nun der

<u>Satz 3.1.</u> Es sei X ein linearer Raum und σ eine Kette in X mit $\theta_X\in\cap\sigma$. Ferner sei e ein festes Element von X. Gelten die zusätzlichen Bedingungen (M) und (T), dann ist das in (2) definierte Funktional m eine Skala der durch σ erzeugten Ordnung. Ist X zusätzlich ein messbarer Raum $(X,\mathfrak{A})$ und gilt $\sigma\subset\mathfrak{A}$, dann ist m messbar.

Beweis. Es ist nur noch die Messbarkeit von m zu zeigen. Wir betrachten dazu $M(\alpha)=\{x|m(x)<\alpha\}$, $\alpha\in\mathbb{R}$. Es gilt

$$\begin{aligned} M(\alpha) &= \{x|\inf\{q(S)|x\in S\}<\alpha\} \\ &= \{x|\text{es gibt ein } S\in\sigma \text{ mit } x\in S \text{ und } q(S)<\alpha\} = \\ &= \bigcup\{S|q(S)<\alpha\}. \end{aligned}$$

Im Fall $M(\alpha)\neq\emptyset$ sei $\alpha_0=\sup\{q(S)|q(S)<\alpha\}$ und $\{S_n\}$ eine Folge mit $q(S_n)<\alpha$ und $q(S_n)\to\alpha_0$ (existiert ein S_0,

$q(S_o)<\alpha$ mit $q(S_o)=\alpha_o$, dann sei $S_n \equiv S_o$).Nun gilt $\bigcup_{n=1}^{\infty} S_n \subset M(\alpha)$, ferner existiert zu einem S, $q(S)<\alpha$ ein S_n mit $q(S)\leqslant q(S_n)$. Wegen der Monotonie von q folgt dann $S \subset S_n$, also $M(\alpha)= =\bigcup_{n=1}^{\infty} S_n$. Daher ist $M(\alpha)$ für jedes $\alpha \in \mathbb{R}$ messbar, was die Messbarkeit von m impliziert.

Im Hinblick auf das Bestimmen optimaler Entscheidungen sind die Konvexitätseigenschaften von m von Interesse.Wir zeigen deshalb das

<u>Lemma 3.1.</u> Gelten die Voraussetzungen des obigen Satzes 3.1. und ist σ zusätzlich ein System konvexer Mengen, dann ist m schwach konvex.

Beweis. Es sei dazu $x<y$ und $z=\lambda x+(1-\lambda)y$, $0\leqslant\lambda\leqslant 1$. Für ein $S \in \sigma$ mit $y \in S$ gilt dann $x \in S$ und daher $z \in S$, was $z \lesssim y$ impliziert. Daraus folgt $m(z)\leqslant m(y)$.

Von grosser Bedeutung ist, wie sich später herausstellen wird, die folgende Klasse von Ketten in X: Es sei

$$\sigma = \{\lambda K \mid \lambda>0\};$$

dabei ist K eine konvexe Teilmenge von X mit $\theta \in K$. Wie das folgende Beispiel zeigt, besitzt nicht jede der durch eine dieser Ketten definierten Ordnungen eine Skala in der Form (2):

Es sei wieder $X=\mathbb{R}^2$, $K=\{x \mid \|x\|\leqslant 1\}\setminus\{(0,1)\}$ und $e=(1,0)$.Für $x=(1,0)$, $y=(0,1)$ gilt $x<y$; aus $q(\lambda K)=\lambda$ folgt jedoch $m(x)=1=m(y)$.

Ein positives Resultat ergibt sich, wenn die obige Klasse geeignet eingeschränkt wird:

<u>Satz 3.2.</u> Es sei X ein linearer Raum und K, $K \neq X$, eine konvexe, algebraisch offene oder abgeschlossene Teilmenge von X, die θ_X als algebraisch inneren Punkt enthält. Dann ist das Minkowskifunktional Q_K von K eine Skala der durch $\sigma=\{\lambda K \mid \lambda>0\}$ erzeugten Ordnung. Ist X zusätzlich ein topologischer Vektorraum und gilt $K \in \mathfrak{B}_X$, dann ist Q_K messbar.

Beweis. Da $K \neq X$, existiert ein $x \in X$ mit $Q_K(x)>0$ und damit ein $e \in X$ mit $Q_K(e)=1$. Nach Voraussetzung ist K al-

---

$A\setminus B = \{x \mid x \in A, x \notin B\}$; Es sei M eine konvexe Menge, eine Abbildung $f:M \to \mathbb{R}$ heisst schwach konvex, falls für alle $x,y \in M$ und $0\leqslant\lambda\leqslant 1$ $f(\lambda x+(1-\lambda)y)\leqslant\max\{f(x),f(y)\}$ gilt.

gebraisch offen oder abgeschlossen. Deshalb gilt $K=\{x|Q_K(x)<1\}$ bzw. $K=\{x|Q_K(x)\leqslant 1\}$. Ist jetzt $S=\alpha K$, $\alpha>0$, dann gilt für (1)

$$\begin{aligned} q(S) &= \sup\{\lambda\geqslant 0|\lambda e\in\alpha K\} \\ &= \sup\{\lambda\geqslant 0|Q_K(\tfrac{\lambda}{\alpha}e)<(\leqslant)1\} \\ &= \sup\{\lambda\geqslant 0|\tfrac{\lambda}{\alpha}<(\leqslant)1\} = \alpha. \end{aligned}$$

Daraus folgt für ein $x\in X$

$$\begin{aligned} m(x)&=\inf\{q(S)|x\in S\}=\inf\{q(\alpha K)|\alpha>0,\ x\in\alpha K\} \\ &=\inf\{\alpha>0|x\in\alpha K\}=Q_K(x)\ . \end{aligned}$$

Von der Betrachtung im Anschluss an die Definition von (2) wissen wir, dass $x\sim y \Longrightarrow m(x)=m(y)$ gilt. Es sei nun $x<y$. Es gibt also ein $\alpha_o>0$, so dass $x\in\alpha_o K$ und $y\notin\alpha_o K$. Daraus folgt $Q_K(x)\leqslant\alpha_o\leqslant Q_K(y)$. Ist K algebraisch offen, so gilt ferner $Q_K(\frac{x}{\alpha_o})<1$ und damit $Q_K(x)<Q_K(y)$. Im andern Fall ist $\alpha_o<Q_K(y)$, denn aus $\alpha_o=Q_K(y)$ folgt $y\in\alpha_o K$, was nach Voraussetzung ausgeschlossen ist. Da $m=Q_K$, folgt $m(x)<m(y)$, d.h. $m=Q_K$ ist eine Skala der durch σ erzeugten Ordnung. Die Messbarkeit von m folgt nun aus Satz 3.1. zusammen mit dem folgenden

Lemma 3.2. Es sei X ein topologischer Vektorraum. Ist B eine Borelmenge von X, dann ist auch λB, $\lambda\in\mathbb{R}$ messbar.

Beweis. Für $\lambda=0$ ist $\lambda B = \emptyset$ oder $=\{\theta\}$, also eine Borelmenge. Es sei nun $\lambda\neq 0$, $\mathfrak{O}=\mathfrak{O}_X$ und $\mathfrak{B}=\mathfrak{B}_X$. Dann gilt $\lambda\mathfrak{B}\supset\lambda\mathfrak{O}=\mathfrak{O}$ und daher $\lambda\mathfrak{B}\supset\mathfrak{B}$, da $\lambda\mathfrak{B}$ eine σ-Algebra ist. Daraus folgt $\frac{1}{\lambda}\mathfrak{B}\supset\mathfrak{B}$ und damit $\lambda\mathfrak{B}\subset\mathfrak{B}$.

Wir erwähnen noch eine Bedingung dafür, dass Q_K stetig ist:

Lemma 3.3. Es sei K eine konvexe Teilmenge eines topologischen Vektorraumes X, die θ als algebraisch inneren Punkt enthält. Eine notwendige und hinreichende Bedingung dafür, dass das Minkowskifunktional Q_K von K stetig ist, ist, dass K mindestens einen (topologisch) inneren Punkt enthält. In diesem Fall stimmen algebraisch Inneres und Inneres sowie die algebraische Hülle

und die abgeschlossene Hülle von K überein.
Beweis: Köthe, Topologische lineare Räume.
Abschliessend bemerken wir noch, dass Q_K konvex ist.

§ 4. Existenz von Nutzenfunktionen

Ist X ein linearer Raum und σ eine Kette in X, die die Voraussetzungen von Satz 3.1 erfüllt, dann besitzt die durch σ erzeugte Ordnung $(X,\sim,<)$ eine Skala s. Nun sind durch Bayes-Kriterien definierte Entscheidungsmodelle nur bedingt sinnvoll, wenn nicht sichergestellt ist, dass die dazu verwendete Skala (Verlustfunktion) bis auf positive lineare Transformationen $\tau \in \mathbb{T}$ eindeutig bestimmt ist. Zur Untersuchung dieses Problems betrachten wir die Ordnung $(X/_\sim,=,<)$; dabei ist $X/_\sim = \{[x]\}$ der Quotient von X modulo "$\sim$", die Relation "=" versteht sich dabei im mengentheoretischen Sinn und "<" ist durch $[x]<[y] \Longleftrightarrow x<y$ definiert. Ist s eine Skala von $(X,\sim,<)$ so ist $\tilde{s}$, $\tilde{s}([x]) = s(x)$ eine von $(X/_\sim,=,<)$ und umgekehrt.
Im weitern beschränken wir uns auf den Fall $\sigma=\{\lambda K|\lambda>o\}$, wobei wir voraussetzen, dass K die in Satz 3.2 aufgestellten Bedingungen erfüllt. Das Minkowskifunktional $m = Q_K$ ist dann nach Satz 3.2 eine Skala der durch σ erzeugten Ordnung $(X,\sim,<)$.

Wir betrachten nun ein beliebiges Element e von X mit $m(e)=1$. Für die von $(X,\sim,<)$ auf $\vec{e} = \{\lambda e|\lambda>o\}$ induzierte Ordnung gilt

$$\lambda e<\mu e \Longleftrightarrow m(\lambda e)<m(\mu e) \Longleftrightarrow \lambda<\mu$$
$$\lambda e\sim\mu e \Longleftrightarrow m(\lambda e)=m(\mu e) \Longleftrightarrow \lambda=\mu ,$$

d.h. sie ist vom Typ $(\vec{e},=,<)$, ferner ist $\lambda e \to \lambda$ ein Isomorphismus von $(\vec{e},=,<)$ auf $(\mathbb{R}_+,=,<)$.
Daher gilt das folgende

Lemma 4.1. Die durch $x \oplus y = x + y$ (Addition in X) definierte Operation $\oplus$ auf $\vec{e}$ ist additiv im Sinne von Definition 1.7.

$\mathbb{R}_+$ = Menge der nichtnegativen reellen Zahlen sowie das

Lemma 4.2. Die Ordnung $(\vec{e},=,<)$ ist zusammenhängend (siehe Satz 1.1).

Es sei nun S die Menge der Skalen der Ordnung $(X,\sim,<)$, s_e die Restriktion von $s \in S$ auf $\vec{e}$ und $\mathcal{E} = \{s \mid s \in S,\ s_e$ ist stetig (bez. der Intervalltopologie), $s_e(x \oplus y) = s_e(x) + s_e(y)$ für alle $x,y \in \vec{e}\}$.

Nun gilt der

Satz 4.1. Gelten für $\sigma = \{\lambda K \mid \lambda > o\}$ dieselben Voraussetzungen wie in Satz 3.2, dann ist $m = Q_K$ eine $\mathcal{E}$-Nutzenfunktion der durch σ auf X erzeugten Ordnung $(X,\sim,<)$.

Beweis. Offensichtlich ist m_e stetig und es gilt $m_e(x \oplus y) = m_e(x) + m_e(y)$ für alle $x,y \in \vec{e}$. Nach Lemma 4.1 ist die Operation $\oplus$ auf $\vec{e}$ additiv, also metrisch (Def. 1.6). Da $(\vec{e},=,<)$ nach Lemma 4.2 noch zusammenhängend ist, folgt aus Satz 1.2, dass ein s_e mit $s \in \mathcal{E}$ bis auf eine positive lineare Transformation eindeutig bestimmt ist, es gilt also für ein s_e mit $s \in \mathcal{E}$ die Gleichung $s_e = \alpha m_e + \beta$, $\alpha,\beta \in \mathbb{R}$, $\alpha > o$. Es sei nun x ein beliebiges Element von X und $\lambda e \in \vec{e}$ mit $x \sim \lambda e$. Für ein $s \in \mathcal{E}$ gilt dann $s(x) = s(\lambda e) = s_e(\lambda e) = \alpha m_e(\lambda e) + \beta = \alpha m(x) + \beta$, also $s = \alpha m + \beta$ mit $\alpha,\beta \in \mathbb{R}$ und $\alpha > o$.

Bemerkungen. Durch die zusätzliche Bedingung der Stetigkeit und Additivität für eine Skala $s \in S$ auf einem Strahl $\vec{e}$ von X ist diese bis auf eine positive lineare Transformation eindeutig festgelegt.

Wird der Nullpunkt der Skalen $s = \alpha m + \beta$ in $\mathcal{E}$ durch $s(\Theta) = o$ fixiert (was in der Folge geschehen soll), dann ist m als Element von $\mathcal{E}$ bis auf einen positiven Faktor eindeutig bestimmt.

§ 5. Entscheidungsregeln

Die Möglichkeit eine Entscheidung zu fällen, d.h. aus einer Menge X_0 von verfügbaren Aktionen eine in einem gewissen Sinn optimale auszuwählen setzt voraus, dass der Entscheidende von zwei verschiedenen Aktionen weiss, ob er die eine der andern vorzieht oder ob sie für ihn gleichrangig sind. Dies bedeutet aber, dass X_0 eine geordnete Menge $(X_0,\sim,<)$ ist.

Wir sagen

<u>Def. 5.1.</u> Eine Aktion $x_0 \in X_0$ heisst ($\lesssim$-)optimal, wenn $x_0 \lesssim x$ für alle $x \in X_0$.

In der uns hier interessierenden Ungewissheitssituation ist einer Aktion x durch

$$\omega \to e(\omega,x) = (c(\omega)x, A(\omega)x - \bar{b}(\omega))$$

eine (messbare) Abbildung $x(.) \in \mathfrak{F}(\Omega,\mathbb{R}\times Z)$ zugeordnet. Wir identifizieren durch

$$x \approx y \iff x(.) = y(.)$$

diejenigen Aktionen x deren x(.) übereinstimmen. Die Menge $X_0/\approx$ bezeichnen wir wieder mit X_0, ihre Elemente heissen wieder Aktionen. Dadurch wird eine Aktion x durch das

---

$\mathfrak{F}(A,B) = \{f \mid f: A \longrightarrow B\}$

ihr zugeordnete $x(.)$ charakterisiert, was wir durch $x = x(.)$ ausdrücken. Wie in der Einleitung bemerkt wurde, ist die Menge $\mathbb{E}$ der Ergebnisse eine geordnete Menge. Da die Ordnung in $\mathbb{E}$ die Rangfolge in der die einzelnen Ergebnisse für den Entscheidenden stehen wiederspiegelt, ist es zwingend, dass die Ordnung $(X_o,\sim,<)$ mit der durch

$$(u,v)\in\rho \iff u(\omega)\leqslant v(\omega) \text{ für alle } \omega\in\Omega$$

definierten Halbordnung ρ auf $\mathfrak{F}(\Omega,\mathbb{R}\times Z)$ verträglich ist; d.h. wir verlangen die Implikation

$$(x(.),y(.))\in\rho \implies x\lesssim y. \tag{1}$$

Besitzt nun die oben erwähnte Ordnung auf $\mathbb{E}$ eine bis auf einen positiven Faktor eindeutig bestimmte Skala m (siehe z.B. Satz 4.1), dann wird eine Aktion x auch durch die Teilmenge

$$\{\lambda M(.,x)\,|\,M(\omega,x) = m(e(\omega,x)),\lambda>0\}$$

von $\mathfrak{F}(\Omega,\mathbb{R})$ charakterisiert (wir identifizieren wieder diejenigen x deren $M(.,x)$ übereinstimmen), was wir durch $x(.) = M(.,x)$ ausdrücken. Daher erzeugt jede Ordnung "$\sim,<$" auf $\mathfrak{F}(\Omega,\mathbb{R})$ (oder auf einer $\{\lambda M(.,x)\,|\,\lambda>0, x\in X_o\}$ enthaltenden Teilmenge von $\mathfrak{F}(\Omega,\mathbb{R})$) mit

$$u\lesssim v \iff \lambda u\lesssim\lambda v \quad (\lambda>0)$$

eine Ordnung auf X_o. Wir geben zwei wichtige Beispiele, wobei wir noch voraussetzen, dass $m\geqslant 0$ und messbar ist. Es sei

$$x\lesssim y \iff \begin{cases} \text{A.} & \|x(.)\|_\infty \leqslant \|y(.)\|_\infty \\ \text{B.} & \sum_{p=1}^{n}\kappa_p\|x(.)\|_p \leqslant \sum_{p=1}^{n}\kappa_p\|y(.)\|_p \end{cases},$$

wobei $\|x(.)\|_\infty = \text{ess.sup}\{|x(\omega)| \mid \omega\in\Omega\}$, $\|x(.)\|_p = (\int|x(\omega)|^p dP(\omega))^{\frac{1}{p}}$ und $\kappa_1,\dots,\kappa_n$ nichtnegative Konstante sind. Offenbar erfüllen diese beiden Ordnungen die obige Verträglichkeitsbedingung (1). Eine Aktion heisst Minimax- bzw. Bayes-optimal, falls sie A-bzw. B-optimal ist. Ist $\|u\| = \text{ess.sup}\{|u(\omega)| \mid \omega \in \Omega\}$ oder $\|u\| = \sum_{p=1}^{n} \kappa_p \|u\|_p$, $u \in \mathfrak{M}(\Omega,\mathbb{R}) = \{u \in \mathfrak{F}(\Omega,\mathbb{R}) \mid u \text{ messbar}\}$, dann besteht das Auffinden einer in diesem Sinn optimalen Aktion in dem folgenden Minimum-Norm-Problem

$$\begin{aligned} &\text{Bestimme } \min \|u\| \\ &\qquad \text{bez. } u \in X_0' = \{M(.,x) \mid x \in X_0\} \end{aligned} \tag{2}$$

in $\mathfrak{M}(\Omega,\mathbb{R})$.

Wir behandeln nun etwas eingehender die Konsequenzen des in Satz 4.1 betrachteten Falls. Es seien dazu X und Z normierte Räume, $K \neq \mathbb{R}\times Z$ eine konvexe, algebraisch offene oder abgeschlossene Teilmenge von $\mathbb{R}\times Z$ (normiert durch $\|(r,z)\| = |r| + \|z\|$), die $\theta_{\mathbb{R}\times Z}$ als inneren Punkt enthält. Nach Satz 4.1 und der Bemerkung danach ist dann das Minkowskifunktional von K eine bis auf einen positiven Faktor eindeutig bestimmte Skala der durch $\sigma = \{\lambda K \mid \lambda > 0\}$ erzeugten Ordnung auf $\mathbb{R}\times Z$. Ferner ist m stetig und konvex, und es gibt ein $\mu > 0$ mit $m((r,z)) \leq \mu\|(r,z)\|$.

Es zeigt sich nun, dass das Optimierungsproblem (2) im wesentlichen konvex ist, denn es ist, falls X_0 konvex ist, wegen der Konvexität von $x \mapsto M(.,x)$ äquivalent zum Problem

$$\begin{aligned} &\text{Bestimme } \min \|u\| \\ &\qquad \text{bez. } u \in X_1 = \{u \in \mathfrak{M}(\Omega,\mathbb{R}) \mid u \geq u_0 \in X_0'\}. \end{aligned} \tag{3}$$

Über die Endlichkeit der rechten Seiten der Implikationen in (A) und (B) gilt das

<u>Lemma 5.1.</u> Sind die Zufallsgrössen $\|A\|, \|b\|, \|c\|$ beschränkt im wesentlichen bzw. p-fach integrierbar ($p \geq 1$), dann ist $\|x(.)\|_\infty$ bzw. $\|x(.)\|_p \in \mathbb{R}$ für alle $x \in X_0$.

Beweis. Zunächst impliziert die Stetigkeit von m die Messbarkeit von x(.). Die Behauptung folgt dann aus der Ungleichung

$$\begin{aligned} 0 \leq x(\omega) &= m((c(\omega)x, A(\omega)x - b(\omega)) \leq \\ &\leq \mu\{\|x\|(\|A(\omega)\| + \|c(\omega)\|) + \|b(\omega)\|\}. \end{aligned}$$

Wir geben nun eine hinreichende Bedingung für die Existenz einer Lösung von (3), wozu die Betrachtung von (2) genügt.

Satz 5.1. Ist $X_o \subset X$ eine kompakte Menge und sind $\|A\|, \|b\|, \|c\|$ beschränkt im wesentlichen bzw. p-fach integrierbar ($p \geq 1$), dann existiert eine A-bzw. B-optimale Aktion.

Beweis. Nach Lemma 5.1 ist X_o' eine Teilmenge des normierten Raumes $\mathcal{L} = \mathcal{L}^\infty = \{u \in \mathfrak{M}(\Omega,\mathbb{R}) \mid u \text{ beschränkt im wesentlichen}\}$ bzw. $= \mathcal{L}^p$. Ferner folgt aus

$$\begin{aligned} |M(\omega,x)-M(\omega,y)| &= |m(e(\omega,x))-m(e(\omega,y))| \\ &\leq \mu \|(c(\omega)(x-y), A(\omega)(x-y)\| \\ &\leq \mu \; \|x-y\| \, (\|c(\omega)\| + \|A(\omega)\|), \end{aligned}$$

dass $x \mapsto M(.,x)$ eine stetige Abbildung von X in $\mathcal{L}$ ist. Damit ist dann X_o' als stetiges Bild einer kompakten Menge ebenfalls kompakt. Da das Zielfunktional von (2) stetig ist, hat es auf X_o' ein Minimum.

Eine optimale Aktion ist von zweifelhaftem Wert, wenn sie nicht zulässig ist:

Def. 5.2. Eine Aktion x_o heisst zulässig, wenn kein $x \in X_o$ existiert, derart dass $(x(.),x_o(.)) \in \rho$ und $x(\omega)<x_o(\omega)$ für mindestens ein $\omega \in \Omega$.

Wir zeigen dazu den

Satz 5.2. Ist $P_{(A,b,c)}(Q)>0$ für jede nichtleere offene Teilmenge Q von $L(X,Z) \times Z \times X^*$ und ist x_o eine Bayes($\kappa_1>0$)-optimale Aktion mit $\|x_o(.)\| = \sum_{p=1}^{n} \kappa_p \|x_o(.)\|_p \in \mathbb{R}$ (siehe Lemma 5.1), dann ist x_o zulässig.

Beweis. Für ein beliebiges $x \in X_o$ und mit $\omega=(A,b,c) \in L(X,Z)\times Z\times X^*$ ist $M(.,x) : \omega \mapsto m((cx,Ax-b))$ eine stetige Abbildung von $L(X,Z)\times Z\times X^*$ in $\mathbb{R}$. Ist nun x_o nicht zulässig, dann existiert ein $x \in X_o$ mit $(M(.,x),M(.,x_o)) \in \rho$ und $M(\omega,x)< M(\omega,x_o)$ für ein $\omega_o \in L(X,Z)\times Z\times X^*$. Daher ist $\|M(.,x)\| \in \mathbb{R}$ und es gibt ein $\varepsilon>0$, so dass $M(\omega,x)<M(\omega,x_o)- \frac{\eta}{2}$, $\eta = M(\omega_o,x_o)-M(\omega_o,x)$, für alle ω mit $\|\omega-\omega_o\|<\varepsilon$. Daraus folgt dann mit $\Pi = P_{(A,b,c)}$

$$\|x_o(.)\| - \|x(.)\| =$$
$$= \kappa_1 \int (M(\omega,x_o)-M(\omega,x))d\Pi(\omega) + \sum_{p=2}^{n} \kappa_p(\|M(.,x_o)\|_p - \|M(.,x)\|_p)$$
$$\geq \kappa_1 \frac{\eta}{2}\, \Pi\{\omega \mid \|\omega-\omega_o\| < \varepsilon\} > 0,$$

also $x_o > x$, was ein Widerspruch zur Optimalität von x_o ist.

§ 6. Randomisierte Aktionen

Steht man wiederholt vor derselben Entscheidungssituation, so kann man sich fragen, ob es nicht geeigneter wäre, statt stets dieselbe (optimale) Aktion $x \in X_o$ zu verwenden, einen Zufallsmechanismus (Strategie)

$$\phi \in \Phi = \{\phi \mid \phi \text{ ist ein W-Mass auf } \mathcal{B}_X \text{ mit } \phi(X_o)=1\}$$

(X_o wird als messbar vorausgesetzt) zu benützen. Wir werden im folgenden zeigen, dass in einem wichtigen Fall dazu kein Grund besteht, was vor allem im Hinblick auf die Bestimmung optimaler Strategien von grosser Bedeutung ist. Zur Behandlung dieses Problems benötigen wir zunächst die folgenden Hilfssätze:

<u>Lemma 6.1.</u> Es seien X und Z normierte Räume. Werden die Produktalgebren $\mathcal{B}_{X^*} \otimes \mathcal{B}_X$ und $\mathcal{B}_X \otimes \mathcal{B}_{L(X,Z)} \otimes \mathcal{B}_Z$ von der (Norm)Topologie des entsprechenden Raumes erzeugt, dann ist (falls m wie in § 5 definiert ist)

$$M : (\omega,x) \mapsto m((c(\omega)x, A(\omega)x-b(\omega))), \quad (\omega,x) \in \Omega \times X$$

messbar.

Beweis. Zunächst sind die Abbildungen

$$\left.\begin{array}{l} S_1 : (\omega,x) \mapsto (x,c(\omega)) \\ S_2 : (\omega,x) \mapsto (x,A(\omega),b(\omega)) \end{array}\right\} \quad (\omega,x) \in \Omega \times X$$

messbar. Ferner zeigt eine leichte Rechnung, dass die Operatoren

$$\begin{array}{lll} T_1 : (x,c) \mapsto cx & & (x,c) \in X \times X^* \\ T_2 : (x,A,b) \mapsto Ax-b & & (x,A,b) \in X \times L(X,Z) \times Z \end{array}$$

stetig und daher (wegen der obigen Voraussetzung) messbar sind. Die Behauptung folgt jetzt aus der Darstellung

$$M = m \circ (T_1 \circ S_1, T_2 \circ S_2).$$

Der folgende Hilfssatz gibt eine hinreichende Bedingung für das Bestehen der im obigen Lemma benötigten Voraussetzung.

Lemma 6.2. Sind $X, X^*, L(X,Z), Z$ separable normierte Räume, dann wird die Produktalgebra $\mathcal{B}_X \otimes \mathcal{B}_{X^*}$ bzw. $\mathcal{B}_X \otimes \mathcal{B}_{L(X,Z)} \otimes \mathcal{B}_Z$ vom System der offenen Mengen in $X \times X^*$ bzw. in $X \times L(X,Z) \times Z$ erzeugt.

Beweis. Wir führen den Beweis für $\mathcal{B} = \mathcal{B}_X \otimes \mathcal{B}_{X^*}$. Nach Definition der Produktalgebra wird $\mathcal{B}$ von der Basis $\mathbb{Q} = \{Q_1 \times Q_2 \mid Q_1 \in \mathcal{O}_X, Q_2 \in \mathcal{O}_{X^*}\}$ der Produkttopologie in $R = X \times X^*$ (normiert durch $\|(x,x^*)\| = \max\{\|x\|, \|x^*\|\}$) erzeugt. Ist $U \subset X$ bzw. $U^* \subset X^*$ eine (nach Voraussetzung existierende) abzählbare und in X bzw. in X^* dichte Menge, dann ist $U \times U^*$ eine solche in R. Damit ist dann $\mathbb{Q}' = \{S((u,u^*),\varepsilon) \mid (u,u^*) \in U \times U^*, \varepsilon \text{ rational}\}$, $S((u,u^*),\varepsilon) = \{(x,x^*) \in X \times X^* \mid \|(x-u, x^*-u^*)\| < \varepsilon\}$ eine abzählbare Basis der Topologie in R, woraus $\mathfrak{A}(\mathbb{Q}') \supset \mathcal{O}_R$ folgt. Anderseits gilt wegen $S((u,u^*),\varepsilon) = S(u,\varepsilon) \times S(u^*,\varepsilon)$ die Inklusion $\mathbb{Q}' \subset \mathbb{Q}$. Die Behauptung folgt jetzt aus $\mathfrak{A}(\mathcal{O}_R) \subset \mathfrak{A}(\mathbb{Q}') \subset \mathcal{B} = \mathfrak{A}(\mathbb{Q}) \subset \mathfrak{A}(\mathcal{O}_R)$.

Bemerkung. Sind X und Z endlichdimensionale normierte Räume, dann ist die Voraussetzung des obigen Lemmas erfüllt.

Wir kommen nun zur Behandlung der zu Beginn dieses Abschnitts gestellten Frage. Es seien dazu $X, X^*, L(X,Z), Z$ separable normierte Räume. Damit ist dann wegen der Hilfssätze 6.2 und 6.1 die oben definierte Abbildung M zusammen mit ihren Potenzen M^p messbar. Wir definieren jetzt die Ordnung auf Φ durch

$$\phi \lesssim \psi \iff r(\phi) \leq r(\psi),$$

wobei $r(\phi) = \sum_{p=1}^{n} \kappa_p (\int M^p d(P \otimes \phi))^{\frac{1}{p}}$.

Ist ε_x das Einpunktmass in x, dann gilt der

Satz 6.1. Ist x_o Lösung des Problems $\min\{\|x(.)\| = \sum_{p=1}^{n} \kappa_p \|x(.)\|_p \mid x \in X_o\}$, dann gilt $\varepsilon_{x_o} \lesssim \phi$ für alle $\phi \in \Phi$.

Beweis. Zunächst gilt nach Fubini

$$r(\phi) = \sum_{p=1}^{n} \kappa_p \left(\int_X \int_\Omega M^p(\omega,x)\,dP(\omega)\,d\phi(x) \right)^{\frac{1}{p}} ,$$

ferner folgt aus der Ungleichung von Jensen

$$\kappa_p \left(\int_X \left\{ \int_\Omega M^p(\omega,x)\,dP(\omega) \right\} d\phi(x) \right)^{\frac{1}{p}} \geq \int_X \kappa_p \left(\int_\Omega M^p(\omega,x)\,dP(\omega) \right)^{\frac{1}{p}} d\phi(x).$$

Damit ist dann wegen $\phi(X_o) = 1$

$$r(\phi) \geq \int_X \sum_{p=1}^{n} \kappa_p \|x(.)\|_p \,d\phi(x)$$

$$= \int_{X_o} \|x(.)\|\,d\phi(x) \geq \|x_o(.)\| = r(\varepsilon_{x_o}).$$

II. Erzeugung von Ordnungen durch lineare Programme

§ 1. Einführung

Unter einem linearen Programm (LP) verstehen wir ein Quintupel (Y,Z,Y_+,M,s), wobei Y und Z lineare Räume, Y_+ ein Kegel in Y, $M : Y \longrightarrow Z$ ein linearer Operator und $s : Y \longrightarrow \mathbb{R}$ ein lineares Funktional sind. Die Abbildung $m : Z \longrightarrow \bar{\mathbb{R}}$,

$$m(z) = m_{(Y,Z,Y_+,M,s)}(z) = \begin{cases} \inf\{sy \mid My=z, y \in Y_+\}, & z \in MY_+ \\ +\infty & , z \notin MY_+ \end{cases}$$

heisse Wert des Programms (Y,Z,Y_+,M,s).

Es seien nun X und Z topologische Vektorräume, X_+ ein Kegel in X und

$$(A,b,c) : (\Omega, \mathfrak{A}, P) \longrightarrow L(X,Z)\times Z \times X^*$$

($c \in X^*_+$ P-fast sicher) ein stochastisches Programm. Im zweistufigen Modell der stochastischen Optimierung wird die Ordnung in $X_0 \subset X_+$ durch

$$x \underset{\sim}{<} y \iff \int m(e(\omega,x))dP(\omega) \leq \int m(e(\omega,y))dP(\omega)$$

definiert, wobei $m((r,z)) = |r|+q(z)$, $(r,z) \in \mathbb{R}\times Z$ und q der Wert eines LP (Y,Z,Y_+,M,s) ist. Setzt man $Y_1=\mathbb{R}^2\times Z$, $Y_{1+} = \mathbb{R}^2_+ \times Y_+$, $M_1(t_1,t_2,y) = (t_1,t_2,My)$ und $s_1(t_1,t_2,y) = t_1+t_2+sy$, $(t_1,t_2,y) \in Y_1$, so ist m offenbar der Wert von $(Y_1,\mathbb{R}\times Z, Y_{1+}, M_1, s_1)$.

Wir untersuchen daher im folgenden den Wert m eines LP $(Y,\mathbb{R}\times Z,Y_+,M,s)$; dabei wird zunächst gezeigt, dass m unter gewissen Bedingungen die Nutzenfunktion einer Ordnung in $\mathbb{R}\times Z$ ist.

§ 2. Der Wert eines linearen Programms als Nutzenfunktion

Es sei (Y,W,Y_+,M,s), $W=\mathbb{R}\times Z$ ein LP mit der Eigenschaft

(V) $\qquad MY_+ = W$, s ist positiv

Der Wert m dieses Programms hat dann die Eigenschaften

Lemma 2.1. a) $0 \leq m < \infty$, b) $m(\theta)=0$, c) $m(\lambda w)=\lambda m(w)$ für alle $w \in W$ und $\lambda \geq 0$, d) $m(w_1+w_2) \leq m(w_1)+m(w_2)$ für alle $w_1, w_2 \in W$.

Beweis: Klar.

Wir betrachten nun die Menge

$$K = \{w \in W \mid m(w) < 1\} .$$

Es gilt das

Lemma 2.2. K ist eine konvexe, algebraisch offene Menge mit $\theta \in K$.

Beweis. Aus Lemma 2.1 folgt sofort die Konvexität von K und $\theta \in K$. Es sei nun $p \in K$ und $w \in W$. Für $\mu=\max\{1, m(w), m(-w)\}$ und $\varepsilon = \frac{1-m(p)}{\mu}$, ist dann $p+tw \in K$ für alle $|t|<\varepsilon$, d.h. $p \in K^i$ und damit $K^i = K$.

Die Bedeutung der Menge K liegt im folgenden

Lemma 2.3. m ist das Minkowskifunktional Q_K von K.

Beweis. Da nach Lemma 2.2. $\theta \in K^i$, existiert das Minkowskifunktional von K und es gilt für ein $w \in W$

$$Q_K(w) = \inf\{\lambda>0 \mid \tfrac{w}{\lambda} \in K\} = \inf\{\lambda>0 \mid m(\tfrac{w}{\lambda}) < 1\} =$$
$$= \inf\{\lambda>0 \mid m(w)<\lambda\} = m(w),$$

d.h. $m = Q_K$.

Es folgt jetzt der

Satz 2.2. Ist $m \neq 0$ der Wert eines Programms (Y,W,Y_+,M,s) mit der Eigenschaft (V), dann ist m Nutzenfunktion der durch $\sigma=\{\lambda K \mid \lambda>0\}$, $K=\{w \mid m(w)<1\}$ erzeugten Ordnung in W.

Beweis. Nach Lemma 2.2 und 2.3 ist m das Minkowskifunktional der konvexen und algebraisch offenen Menge K. Da noch $\theta \in K$ und (nach Voraussetzung) $K \neq W$, folgt die Behauptung aus I, Satz 4.1.

Im obigen Satz wird die Voraussetzung (V) gemacht; im folgenden Abschnitt geben wir dafür notwendige und hinreichende Bedingungen.

§ 3. Lineare Programme mit der Eigenschaft (V)

Wir zitieren zunächst zwei Aussagen über die Existenz positiver linearer Funktionale:

<u>Satz 3.1.</u> Ist Y ein lokalkonvexer topologischer Vektorraum mit einem Kegel Y_+, dann existiert ein $y^* \in Y^*_+$, $y^* \neq \theta^*$ genau dann, wenn Y_+ nicht dicht in Y ist.
Beweis: Peressini. Ordered Topological Vector Spaces.

<u>Satz 3.2.</u> Ist Y ein normierter Raum mit einem abgeschlossenen Kegel Y_+, dann existiert für jedes $y \in Y_+, y \neq \theta$ ein $y^* \in Y^*_+$ mit $y^*(y)>0$. Ist Y zusätzlich separabel, dann gibt es ein $s \in Y^*_+$ mit $sy>0$ für alle $y \in Y_+$, $y \neq \theta$.

Beweis: Krasnoselskii, Positive solutions of operator equations.

Wir machen nun einige Bemerkungen zum ersten Teil von (V):

<u>Satz 3.3.</u> Es seien Y und Z lineare Räume, Y_+ ein Kegel in Y und e eine Ordnungseinheit in Y. Ist $M : Y \longrightarrow Z$ ein surjektiver linearer Operator mit $e \in \mathrm{Kern}(M)$, dann gilt $MY_+ = Z$.
Beweis. Es sei $z \in Z$. Nach Voraussetzung existiert ein $y \in Y$ mit $My = z$. Da Y_+ eine Ordnungseinheit enthält, hat man die Darstellung $y=y_1 - y_2$ mit $y_1,y_2 \in Y_+$, ferner existiert ein $\alpha \geq 0$ mit $\alpha e \geq y_2$. Für $\tilde{y} = y + \alpha e$ ist dann $M\tilde{y} = z$ und $\tilde{y} \geq \theta$.

Bemerkung. Sind Y,Z endlichdimensionale Räume und ist Y_+ der positive Orthant von Y, dann ergeben sich aus dem obigen Satz im wesentlichen die Sätze 2 und 3 von § 5 in [11].

Ist Z_+ ein Kegel in einem linearen Raum Z mit $Z_+ - Z_+ = Z$, so lässt sich leicht ein LP mit der Eigenschaft (V) konstruieren:
<u>Satz 3.4.</u> Ist Z ein linearer Raum, Z_+ ein Kegel in Z mit $Z_+ - Z_+ = Z$ und sind s_1, s_2 zwei positive lineare Funktionale auf Z, dann ist $(Z\times Z, Z, Z_+\times Z_+, M, s)$ mit $M(z_1,z_2)=z_1-z_2$, $s(z_1,z_2)=s_1z_1 + s_2z_2$, $(z_1,z_2)\in Y=Z\times Z$ ein LP mit der Eigen-

schaft (V). Ist Z ein Vektorverband, dann gilt für den Wert m dieses Programms

$$m(z) = s_1 z^+ + s_2 z^- , \quad z \in Z.$$

Beweis. Der erste Teil des Satzes ist klar. Es sei nun $z \in Z$. Ist $y = (z_1, z_2)$ eine beliebige Lösung für z, d.h. ist $z = z_1 - z_2$ mit $z_1, z_2 \in Z_+$, dann gilt wegen $z_1 \geqslant z^+$ und $z_2 \geqslant z^-$ (minimale Dekomposition von z in einem Vektorverband)

$$\begin{aligned} sy &= s_1 z_1 + s_2 z_2 \\ &\geqslant s_1 z^+ + s_2 z^- , \end{aligned}$$

woraus die Behauptung folgt.

Wir suchen jetzt eine Charakterisierung für LP mit der Eigenschaft (V). Dazu betrachten wir zunächst den Fall endlichdimensionaler Räume.

Es sei also (Y, Z, Y_+, M, s) ein LP mit $Y = \mathbb{R}^n$, $Z = \mathbb{R}^m$ und $Y_+ = \mathbb{R}^m_+$ (positiver Orthant von $\mathbb{R}^m$), das (V) erfüllt. Ist wie in § 2 m der Wert des obigen Programms und $K = \{z \in Z \mid m(z) < 1\}$, dann gilt (zunächst noch allgemein)

$$\begin{aligned} K &= \{z \in Z \mid \inf\{sy \mid My = z,\ y \geqslant \theta\} < 1\} \\ &= \{z \mid \text{es existiert ein } y \geqslant \theta \text{ mit } My = z \text{ und } sy < 1\} \\ &= \{My \mid y \geqslant \theta,\ sy < 1\}. \end{aligned}$$

Es sei nun $\{e_1, \dots, e_n\}$ die durch $e_{ij} = \delta_{ij}$ (Kroneckersymbol) definierte natürliche Basis von Y; wir setzen noch $se_i > 0$, $i = 1, \dots, n$ voraus. Mit $k_i = \frac{1}{se_i} Me_i$, $i = 1, \dots, n$ und $k_0 = \theta_Z$ folgt dann

$$\begin{aligned} K &= \{\sum_{i=1}^{n} \alpha_i k_i \mid \alpha_i \geqslant 0,\ i = 1, \dots, n,\ \sum_{i=1}^{n} \alpha_i < 1\} \\ &\subset \{\sum_{i=1}^{n} \alpha_i k_i \mid \alpha_i \geqslant 0,\ i = 1, \dots, n,\ \sum_{i=1}^{n} \alpha_i \leqslant 1\} = \operatorname{conv}\{k_0, k_1, \dots, k_n\}. \end{aligned}$$

Anderseits gilt für ein beliebiges μ, $0 < \mu < 1$

---

conv(M) = konvexe Hülle der Menge M

$$K \supset \{\sum_{i=1}^{n} \alpha_i k_i \mid \alpha_i \geq 0,\ i=1,\dots,n,\ \sum_{i=1}^{n} \alpha_i \leq \mu\}$$
$$= \mu \operatorname{conv}\{k_0,k_1,\dots,k_n\}\ .$$

Nun ist $P = \operatorname{conv}\{k_0,k_1,\dots,k_n\}$ eine konvexe Menge und wegen Lemma 2.2 und den obigen Inklusionen gilt $\theta_Z \in P^i$. Daher existiert das Minkowskifunktional von P (und von μP), und es ist

$$Q_{\mu P} \geq Q_K \geq Q_P\ ,\ 0<\mu<1.$$

Wegen $m = Q_K$ (Lemma 2.3) und $Q_{\mu P} = \frac{1}{\mu} Q_P$ folgt daraus

$$\frac{1}{\mu} Q_P \geq m \geq Q_P$$

und damit

$$|m-Q_P| = m-Q_P \leq (\frac{1}{\mu} - 1)Q_P$$

für alle μ, $0<\mu<1$, was $m=Q_P$ impliziert. Der Wert eines LP $(\mathbb{R}^n,\mathbb{R}^m,\mathbb{R}^n_+,M,s)$ mit der Eigenschaft (V) ist demnach das Minkowskifunktional eines konvexen Polyeders P in $\mathbb{R}^m$ mit $\theta \in P^i$. Wir zeigen die Umkehrung:

Es sei dazu $P = \operatorname{conv}\{k_1,\dots,k_n\}$ ein konvexes Polyeder in $Z = \mathbb{R}^m$ mit $\theta \in P^i$, ferner $Y = \mathbb{R}^n$ und $Y_+ = \mathbb{R}^n_+$. Ist dann $\{e_1,\dots,e_n\}$ wieder die natürliche Basis von Y, so wird durch $Me_i = k_i$ und $se_i = 1$, $i=1,\dots,n$ ein linearer Operator $M : Y \longrightarrow Z$ mit $MY_+ = Z$ und ein positives lineares Funktional $s : Y \longrightarrow \mathbb{R}$ definiert; denn zunächst hat jedes $z \in Z$ eine Darstellung $z = \lambda \sum_{i=1}^{n} \alpha_i k_i$ mit $\lambda>0$ und $\alpha_i \geq 0$, $\sum_{i=1}^{n} \alpha_i = 1$. Daraus folgt aber $z = My$ mit $y = (\lambda\alpha_1,\dots,\lambda\alpha_n) \in Y_+$. Dass s positiv ist, ist klar. Wie oben zeigt man dann wieder $m = Q_P$. Zusammenfassend gilt der

<u>Satz 3.5.</u> Ist Y, Z je ein endlichdimensionaler Raum und Y_+ der positive Orthant von Y in bezug auf eine Basis in Y, dann ist jedem Programm (Y,Z,Y_+,M,s) mit der Eigenschaft (V) und $sy>0$, $y \in Y_+$, $y \neq \theta$ ein konvexes Polyeder P in Z mit $\theta_Z \in P^i$ zugeordnet und umgekehrt, derart dass der Wert m dieses Programms das Minkowskifunktional von P ist.

Bevor wir uns abschliessend mit dem allgemeinen Fall befassen, zeigen wir im folgenden Satz, dass Minkowskifunktionale gewisser konvexer Mengen in Z durch Werte linearer Programme approximiert werden können:

Satz 3.6. Ist $Z = E^n$ und R eine kompakte konvexe Menge in Z mit $\theta \in \overset{o}{R}$, dann gilt $Q_R = \lim_{n\to\infty} \pi_n$, wobei Q_R das Minkowskifunktional von R und π_n der Wert eines LP $(Y_n, Z, Y_{n+}, M_n, s_n)$ mit der Eigenschaft (V) ist.

Beweis. Nach Valentine, Konvexe Mengen existiert zu jedem λ, $0<\lambda<1$, ein konvexes Polyeder P_λ mit $\lambda R \subset P_\lambda \subset R$. Da $\theta \in \overset{o}{P}_\lambda$, gilt

$$\frac{1}{\lambda} Q_R \geqslant Q_{P_\lambda} \geqslant Q_R$$

woraus

$$|Q_{P_\lambda} - Q_R| \leqslant \left(\frac{1}{\lambda} - 1\right) Q_R$$

folgt. Ist deshalb $\{\lambda_n\}$ eine Folge mit $0<\lambda_n<1$ und $\lambda_n \to 1$, dann gilt

$$Q_R = \lim_{n\to\infty} Q_{P_{\lambda_n}} .$$

Die Behauptung folgt nun aus Satz 3.5.

Wir behandeln jetzt die oben angekündigte Verallgemeinerung von Satz 3.5 :

Es sei dazu Z ein beliebiger linearer Raum und P eine konvexe Teilmenge von Z mit $\theta_Z \in P^i$, ferner B eine Teilmenge von P mit $\mathrm{conv}(B) = P$. Da $\theta \in P$, gilt

$P = \{\sum_{i=1}^{n} \alpha_i b_i \,|\, \alpha_i \geqslant 0, \sum_{i=1}^{n} \alpha_i \leqslant 1,\ b_i \in B,\ n \in \mathbb{N}\}$. Nun existiert ein linearer Raum Y mit $\dim Y = \mathrm{card}(B)$, es sei $\mathcal{B} = \{e_i \,|\, i \in I\}$ eine Basis von Y. Wir indizieren die Elemente von B derart, dass $B = \{b_i \,|\, i \in I\}$. Durch $Me_i = b_i$, $se_i = 1$ werden zwei lineare Abbildungen $M : Y \to Z$ und $s : Y \to \mathbb{R}$ definiert. Mit $Y_+ = \{\sum_{j=1}^{n} \alpha_j e_{i_j} \,|\, \alpha_j \geqslant 0,\ i_j \in I,\ n \in \mathbb{N}\}$ ist s positiv und es gilt $MY_+ = Z$. Ist m wieder der Wert des so definierten Programms (Y, Z, Y_+, M, s) und $K = \{z \in Z \,|\, m(z) < 1\}$, dann gilt nach Lemma 2.3 $m = Q_K$. Ferner erhält man

$$\begin{aligned} K &= \{My \,|\, y \in Y_+,\ sy<1\} \\ &\subset \{My \,|\, y \in Y_+\ \ sy \leqslant 1\} = \end{aligned}$$

$\overset{o}{R}$ = Inneres von R

$= \{\sum_{i=1}^{n} \alpha_i b_i \,|\, \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i \leq 1,\ b_i \in B,\ n \in \mathbb{N}\} = P$

und für μ, $0<\mu<1$ wieder $\mu P \subset K$, woraus dann wie oben $m=Q_P$ folgt. Damit gilt der

Satz 3.7. Jedem LP (Y,Z,Y_+,M,s) mit der Eigenschaft (V) ist eine konvexe Menge $P \subset Z$ mit $\theta_Z \in P^i$ zugeordnet und umgekehrt, derart dass der Wert dieses Programms das Minkowskifunktional von P ist.

Nach Satz 2.2 ist der Wert eines LP mit der Eigenschaft (V) Nutzenfunktion einer gewissen Ordnung. Da in Kapitel I zumindest messbare Nutzenfunktionen benötigt werden, behandeln wir im folgenden

§ 4. Stetigkeitsaussagen über Werte linearer Programme

Es seien Y und Z normierte Räume und Y_+ ein abgeschlossener Kegel in Y. Ferner sei $M \in L(Y,Z)$ mit $MY_+ = Z$ und $s \in Y^*_+$, $s \neq \theta^*$ (siehe Satz 3.2). Ist m der Wert des Programms (Y,Z,Y_+,M,s), dann definieren wir wieder $K = \{z \in Z \,|\, m(z)<1\}$. Da m das Minkowskifunktional von K ist, liegt der Schlüssel unseres Problems im

Lemma 4.1. Es sei Z ein normierter Raum und K eine konvexe Teilmenge von Z. Ist $\theta_Z \in \overset{o}{K}$, dann ist das Minkowskifunktional Q_K von K stetig; ferner ist $Q_K(z) \leq \frac{1}{\rho}\|z\|$, $z \in Z$, wobei ρ eine Zahl mit $\rho>0$ und $S(\theta,\rho)=\{z \,|\, \|z\|<\rho\} \subset K$ ist.

Beweis: Köthe, Topologische lineare Räume.

Nach § 3 gilt

$$K = \{My \,|\, y \geq \theta,\ sy<1\}$$
$$\supset \{My \,|\, y \geq \theta,\ \|y\| < \frac{1}{\|s\|}\}.$$

Für ein r mit $0<r<\frac{1}{\|s\|}$ folgt daraus

$$K \supset \{My \,|\, y \geq \theta,\ \|y\| \leq r\} = r\,\{My \,|\, y \geq \theta,\ \|y\| \leq 1\}.$$

Sind Y und Z Banachräume, dann ist M wegen $MY = MY_+ = Z$ und $M \in L(Y,Z)$ eine offene Abbildung (open mapping theorem). Nun ist aber $\{y \,|\, y \geq \theta,\ y < \frac{1}{\|s\|}\}$ i.a. keine offene Menge (in Y), so dass dieses Theorem hier i.a. keine Aussage liefert. Eine Aussage lässt sich dagegen für den Fall, dass $\{My \mid y \geq \theta, \|y\| \leq 1\}$ abgeschlossen ist machen. Es gilt nämlich der folgende

<u>Satz 4.1.</u> Es sei Y ein normierter - Z ein Banachraum, Y_+ ein abgeschlossener Kegel in Y und (Y,Z,Y_+,M,s) ein Programm mit der Eigenschaft (V) und $M \in L(Y,Z)$, $s \in Y^*$. Ist $\{My \mid y \geq \theta, \|y\| \leq 1\}$ eine abgeschlossene Teilmenge von Z, dann ist der Wert m des Programms stetig.

Beweis. Nach Lemma 4.1 haben wir nur zu zeigen, dass θ_Z innerer Punkt von $K = \{z \mid m(z) < 1\}$ ist. Wir betrachten dazu die abgeschlossene Menge $B = \{y \mid y \geq \theta, \|y\| \leq 1\}$. Aus $\bigcup_{n=1}^{\infty} nB = Y_+$ folgt $Z = MY_+ = \bigcup_{n=1}^{\infty} nMB$. Da MB nach Voraussetzung eine abgeschlossene Teilmenge eines Banachraumes Z ist, lässt sich nun der folgende Hilfssatz anwenden:

<u>Lemma 4.2.</u> Ist Z ein vollständiger metrischer Raum und $\{A_n\}$ eine Folge abgeschlossener Teilmengen von Z mit $Z = \bigcup_{n=1}^{\infty} A_n$, so existiert ein n_0, derart dass A_{n_0} eine offene Teilmenge (von Z) enthält.

Beweis: Bonic, Linear Functional Analysis.

Wäre nun MB eine absolutkonvexe Menge, so erhielten wir aus dem obigen Satz die Existenz einer Kugelumgebung von θ_Z in einem n_0MB und damit auch in MB. Da dies i.a. nicht der Fall ist, suchen wir eine geeignete absolutkonvexe Teilmenge von MB. Dazu stellen wir zunächst einige Eigenschaften von A = MB zusammen: A ist konvex und (nach Voraussetzung) abgeschlossen, enthält θ_Z und es gilt $\bigcup_{n=1}^{\infty} nA = Z$; ferner ist $\lambda z \in A$ für jedes $z \in A$ und $|\lambda| \leq \rho$ mit einem gewissen ρ, $0 < \rho \leq 1$. Die ersten vier Eigenschaften sind klar. Wir zeigen die letzte. Da A konvex ist und θ_Z enthält, gilt $\lambda z \in A$ für jedes $z \in A$ und λ mit $0 \leq \lambda \leq 1$. Zu $-z \neq \theta_Z$ existiert wegen (V) ein $y \geq \theta$, $y \neq \theta$ mit $My = -z$. Ist $\|y\| \leq 1$, dann erfüllt $\rho = \|y\|$, im andern Fall $\rho = \frac{1}{\|y\|}$ die obige Bedingung. Nach diesen Vorbereitungen betrachten wir die

Eigenschaften von

$$A_1 = \{z \in A \mid \lambda z \in A \text{ für alle } \lambda \text{ mit } |\lambda| \leqslant 1\}.$$

Zunächst liegt θ_Z in A_1 und A_1 ist konvex, denn sind z_1, $z_2 \in A_1$ und $\alpha,\beta \geqslant 0$, $\alpha+\beta=1$, dann folgt wegen der Konvexität von A

$$\lambda(\alpha z_1 + \beta z_2) = \alpha\lambda z_1 + \beta\lambda z_2 \in A, \quad |\lambda| \leqslant 1,$$

also $\alpha z_1 + \beta z_2 \in A_1$. A_1 ist kreisförmig: Es sei dazu $z \in A_1$, ferner μ eine feste Zahl mit $|\mu| \leqslant 1$; für $|\lambda| \leqslant 1$ gilt dann

$$\lambda(\mu z) = (\lambda\mu)z \in A,$$

also $\mu z \in A_1$. Da die Menge A_1 konvex und kreisförmig ist, ist sie absolutkonvex. A_1 ist abgeschlossen: Es sei $z \in \bar{A}_1$. Es gibt also eine Folge $\{z_n\}$ mit $z_n \in A_1$ und $z_n \to z$. Für ein festes λ mit $|\lambda| \leqslant 1$ gilt dann $\lambda z_n \in A$ und $\lambda z_n \to \lambda z$, woraus $\lambda z \in A$ und damit $z \in A_1$ folgt, da A abgeschlossen ist. Schliesslich ist $\bigcup_{n=1}^{\infty} nA_1 = Z$. Es sei dazu $z \in Z$; wegen $\bigcup_{n=1}^{\infty} nA = Z$ existiert ein $n \in \mathbb{N}$ und ein $a \in A$ mit $z = na$. Nach dem Obigen gibt es nun ein r, $0<r\leqslant 1$ mit $\mu a \in A$ für alle μ, $|\mu| \leqslant r$. Für $a_1 = ra$ ergibt sich $\lambda a_1 = (\lambda r)a \in A$ für alle $|\lambda| \leqslant 1$, also $a_1 \in A_1$. Damit gilt dann $z = na = \frac{n}{r}a_1 \in n'A_1$ mit $n' \in \mathbb{N}$ und $n' \geqslant \frac{n}{r}$. Da Z vollständig ist, existiert nun nach Lemma 4.2 ein n_o, so dass $n_o A_1$ eine offene Menge von Z enthält, womit auch A_1 eine offene Menge von Z, also eine gewisse Kugel $S(a',\rho)$, $a' \in A_1, \rho>0$ enthält. Da A_1 absolutkonvex ist, gilt $S(-a',\rho) \subset A_1$ und daher $S(\theta_Z,\rho) \subset A_1$. Daraus folgt aber, dass θ_Z innerer Punkt von A und damit von K ist.

Für endlichdimensionale Räume liefert Satz 4.1 das folgende

Korollar 4.1. Es seien Y und Z endlichdimensionale Räume und Y_+ ein beliebiger abgeschlossener Kegel in Y. Erfüllt das Programm (Y,Z,Y_+,M,s) die Bedingung (V), dann ist sein Wert stetig.

Beweis. Zunächst ist $B = \{y \mid y \geqslant \theta, \|y\| \leqslant 1\}$ eine kompakte Menge in Y. Als stetiges Bild von B ist auch

---

$\bar{M}$ = abgeschlossene Hülle von M

$\{My \mid y \geq \theta, \|y\| \leq 1\}$ kompakt, also abgeschlossen in Z. Die Behauptung folgt jetzt aus Satz 4.1.

Abschliessend zeigen wir noch den

Satz 4.2. Ist Z ein normierter Vektorverband (d.h. aus $\theta \leq x \leq y$ folg $\|x\| \leq \|y\|$) und sind d_1 und d_2 zwei Elemente aus Z^*, dann ist (siehe Satz 3.4) $m(z) = d_1 z^+ + d_2 z^-$ ein stetiges Funktional und es gilt $|m(z)| \leq (\|d_1\| + \|d_2\|)\|z\|$.

Beweis. Da Z ein normierter Vektorverband ist, sind

$$z \to z^+ \text{ und } z \to z^-$$

stetige Operartoren, was die Stetigkeit von m impliziert. Aus $\|z^+\| \leq \|z\|$ und $\|z^-\| \leq \|z\|$ folgt dann der zweite Teil der Behauptung.

III. Die Differenzierbarkeit der Zielfunktion bei Bayes-Kriterien

§ 1. Einführung

Es seien X und Z normierte Räume, X_0 eine Teilmenge von X und $(A,b,c) : (\Omega,\mathfrak{A},P) \longrightarrow L(X,Z)\times Z\times X^*$ $(\|(A,b,c)\| = \|A\|+\|b\|+\|c\|)$ ein normintegrierbares stochastisches Programm. Ferner sei die Nutzenfunktion der Ordnung in $\mathbb{E} = \{(c(\omega)x, A(\omega)x-b(\omega)) \mid \omega\in\Omega, x\in X_0\}$ das Minkowskifunktional $Q=Q_K$ einer konvexen Teilmenge K von $W = \mathbb{R}\times Z$ $(\|(r,z)\| = |r|+\|z\|)$ mit $\theta_W \in \overset{\circ}{K}$. Orientiert sich der Entscheidende am Bayeskriterium, dann steht man nach I, § 5 (im einfachsten Fall) vor dem Optimierungsproblem

Bestimme

$$\begin{aligned} &\min f(x) = EQ((cx, Ax-b)) \\ &\text{bez. } x\in X_0. \end{aligned} \tag{1}$$

Die Lösung von (1) wird in vielen Fällen erleichert, wenn man weiss, dass f zumindest Gateaux(G)-differenzierbar ist:

<u>Def. 1.1.</u> Es seien X und Z lineare Räume. Der Operator $T : X \longrightarrow Z$ heisst G-differenzierbar im Punkt $x\in X$, falls für jedes $h\in X$ der Grenzwert

$$\delta T(x,h) = \lim_{t\to 0} \frac{1}{t}\{T(x+th)-Tx\}$$

existiert.

§ 2. Die Gateaux-Differenzierbarkeit der Zielfunktion f

Da nach unseren Voraussetzungen $\bar{K} = \{w\in W \mid Q(w)\leq 1\}$, können wir wegen $Q_K = Q_{\bar{K}}$ die Menge K als abgeschlossen voraussetzen. Ist nun x ein festes und y ein beliebiges Element von X, dann gilt

$$\frac{1}{t}\{f(x+ty) - f(x)\} =$$

$$= \int \frac{1}{t}\{Q((c(\omega)x+tc(\omega)y, A(\omega)x-b(\omega)+tA(\omega)y))-Q((c(\omega)x, A(\omega)x-b(\omega))\}dP(\omega).$$

$E\chi = \int\chi(\omega)dP(\omega)$

Wir müssen daher Q auf die G-Differenzierbarkeit an der Stelle $(c(\omega)x, A(\omega)x-b(\omega))$ untersuchen. Aus $\theta_W \in \overset{o}{K}$ folgt das

Lemma 2.1. Für beliebige $w,h \in W$ existiert der Grenzwert

$$\tau(w,h) = \lim_{\substack{t\to 0\\ t>0}} \frac{1}{t}\{Q(w+th) - Q(w)\}$$

und es gilt

a) $\tau(w,h) \leq Q(h)$

b) $\tau(w,h_1 + h_2) \leq \tau(w,h_1) + \tau(w,h_2)$

c) $\tau(w,\alpha h) = \alpha\tau(w,h)$ für $\alpha \geq 0$

d) $-\tau(w,-h) \leq \tau(w,h)$

e) $\tau(w,\alpha w) = \alpha Q(w)$ für $\alpha \in \mathbb{R}$

Beweis: Dunford-Schwartz, Linear Operators I.

Mit Hilfe von τ ergibt sich nun das folgende Kriterium

Lemma 2.2. Q ist in w genau dann G-differenzierbar, falls für jedes $h \in W$ die Gleichung $-\tau(w,-h) = \tau(w,h)$ gilt.

Beweis: Dunford-Schwartz, Linear Operators I.

Die folgende Definition ermöglicht eine Verbindung zwischen der Struktur des Randes ∂K von K und der G-Differenzierbarkeit von $Q = Q_K$:

Def. 2.1. Es sei W ein topologischer Vektorraum, A eine Teilmenge von W. Ein $w^* \in W^*$ heisst tangential zu A in $a \in \partial K$, wenn ein $c \in \mathbb{R}$ existiert, so dass $w^*(a) = c$ und $w^*(w) \leq c$ für alle $w \in A$.

Es gilt nun das

Lemma 2.3. Es sei K eine abgeschlossene und konvexe Teilmenge des lokalkonvexen topologischen Vektorraumes W mit $\theta_W \in \overset{o}{K}$. Das Minkowskifunktional Q_K von K ist dann und nur dann in $w \in \partial K$ G-differenzierbar, wenn in w genau ein zu K tangentiales $w^* \in W^*$ existiert.

Beweis: Dunford-Schwartz, Linear Operators I.

Aus Lemma 2.1. und 2.2 folgt nun das

Korollar 2.1. Ist Q_K in w G-differenzierbar, dann ist $\delta Q_K(w,.) \in W^*$.

Beweis. Die Additivität und Homogenität von $\delta Q(w,.)$ ergibt sich wegen $\delta Q(w,h) = \tau(w,h)$ (Lemma 2.2) aus (b) und (c) von Lemma 2.1; die Beschränktheit folgt aus $|\delta Q(w,h)| = |\tau(w,h)| = \tau(w,\operatorname{sgn}\tau(w,h)h) \leq Q(\operatorname{sgn}\tau(w,h)h) \leq \frac{1}{\rho}\|h\|$, wobei $\rho > 0$ und $S(\theta_W,\rho) \subset K$.

Wir setzen in diesem Fall $\delta Q(w,.) = \operatorname{grad} Q(w) \equiv \nabla Q(w)$. Offenbar ist für $w \in \partial K$ $(Q(w) = 1)$ $\nabla Q(w)$ tangential zu K.

Nach diesen Vorbereitungen benötigen wir nun eine Abschätzung für den Differentialquotienten $\frac{1}{t}\{Q(w+th) - Q(w)\}$. Da Q subadditiv ist und $S(\theta_W,\rho) \subset K, \rho>0$, gilt für $w_1, w_2 \in W$

$$|Q(w_1) - Q(w_2)| \leq \frac{1}{\rho}\|w_1 - w_2\|,$$

woraus

$$\left|\frac{1}{t}\{Q(w+th) - Q(w)\}\right| \leq \frac{1}{|t|}\,\frac{1}{\rho}\|th\| = \frac{1}{\rho}\|h\|$$

folgt. Mit $w = (c(\omega)x, A(\omega)x-b(\omega))$ und $h = (c(\omega)y, A(\omega)y)$ gilt daher

$$\left|\frac{1}{t}\{Q((c(\omega)x+tc(\omega)y, A(\omega)x-b(\omega)+tA(\omega)y)) - Q((c(\omega)x, A(\omega)x-b(\omega)))\}\right| \leq \frac{1}{\rho}\|(c(\omega)y, A(\omega)y)\| \leq \frac{1}{\rho}\|y\|(\|c(\omega)\| + \|A(\omega)\|).$$

Da der Betrag des Differentialquotienten demnach durch eine integrierbare Zufallsvariable majorisiert wird, gilt der

Satz 2.1. Ist $(A,b,c) : (\Omega, \mathfrak{A}, P) \longrightarrow L(X,Z) \times Z \times X^*$ ein norm-integrierbares stochastisches Programm und $N_{x_0} = \{\omega \in \Omega \mid Q$ ist in $(c(\omega)x_0, A(\omega)x_0-b(\omega))$ nicht G-differenzierbar$\}$ eine P-Nullmenge, dann ist $f(x) = EQ((cx, Ax-b))$, $x \in X$ in x_0 G-differenzierbar.

Beweis. Es sei $\{t_n\}$ eine beliebige Folge mit $t_n \neq 0$ und $t_n \to 0$. Für jedes feste $\omega \notin N_{x_0}$ und $y \in X$ gilt dann

$$g_n(\omega) \equiv \frac{1}{t_n}\{Q((c(\omega)x_0 + t_n c(\omega)y, A(\omega)x_0 - b(\omega) + t_n A(\omega)y)) - Q((c(\omega)x_0, A(\omega)x_0 - b(\omega)))\} \to$$
$$\to \nabla Q((c(\omega)x_0, A(\omega)x_0 - b(\omega)))(c(\omega)y, A(\omega)y).$$

Da ferner $|g_n| \leq \frac{1}{\rho}\|y\|(\|A\| + \|c\|)$ für alle $n \in \mathbb{N}$ und nach Voraussetzung $P(N_{x_0}) = 0$, ergibt sich aus dem Konvergenzsatz von Lebesgue die Existenz einer integrierbaren Zufallsgrösse g mit $g(\omega) = \nabla Q((c(\omega)x_0, A(\omega)x_0 - b(\omega)))(c(\omega)y, A(\omega)y)$ f.s.

und $\int |g_n - g| dP \to 0$. Wegen der Ungleichung

$$\left|\frac{1}{t_n}\{f(x_o+t_n y) - f(x_o)\} - Eg\right| = |Eg_n - Eg| \leq E|g_n - g|$$

ist f in x_o G-differenzierbar, und es gilt

$$\delta f(x_o,y) = \int \nabla Q((c(\omega)x_o, A(\omega)x_o - b(\omega)))(c(\omega)y, A(\omega)y)dP(\omega).$$

Im Anschluss an Satz 2.1 erhalten wir noch das

Korollar 2.2. $\delta f(x_o,.)$ ist ein Element von X^*.

Beweis. Zunächst ist dieses Funktional sicher additiv und homogen; die Stetigkeit folgt mit $w=(c(\omega)x_o, A(\omega)x_o-b(\omega))$ und $h=(c(\omega)y, A(\omega)y)$ aus

$$|\nabla Q((c(\omega)x_o, A(\omega)x_o-b(\omega)))(c(\omega)y, A(\omega)y)| = |\tau(w,h)| \leq$$
$$\leq \frac{1}{\rho}\|y\|(\|c(\omega)\| + \|A(\omega)\|).$$

Abschliessend untersuchen wir die in Satz 2.1 definierte Menge N_x :

§ 3. Eigenschaften der Menge $N_x = \{\omega \in \Omega | Q$ ist in $(c(\omega)x, A(\omega)x-b(\omega))$ nicht G-differenzierbar$\}$

Zur Untersuchung dieser Menge (für festes $x \in X$) betrachten wir

$$\Delta = \{w \in W | Q \text{ ist in } w \text{ nicht G-differenzierbar}\} .$$

Offenbar gilt $N_x = (cx, Ax-b)^{-1}(\Delta)$ und damit $P(N_x) = P_{(cx,Ax-b)}(\Delta)$, falls Δ eine messbare Teilmenge von W ist. Es gilt nun zunächst das

Lemma 3.1. Ist Z ein separabler Raum, dann ist Δ messbar.

Beweis. Nach § 2 liegt ein $w \in W$ genau dann in Δ , falls mindestens ein $h \in W$ existiert, so dass $\tau(w,h) \neq -\tau(w,-h)$. Wir betrachten deshalb die durch $s : (w,h) \mapsto \tau(w,h)$, $(w,h) \in W \times W$ definierte Abbildung. Es sei $\{t_n\}$ eine Folge mit $t_n > 0$ und $t_n \to 0$. Mit $s_n((w,h)) = \frac{1}{t_n}\{Q(w+t_n h) - Q(w)\}$, $(w,h) \in W \times W$ gilt dann $s = \lim_{n\to\infty} s_n$ (punktweise). Wir zeigen, dass s messbar ist: Nach Lemma 6.2 von Kapitel I wird die Produktalgebra $\mathcal{B}_W \otimes \mathcal{B}_W$ wegen der Separabilität von W

(R und Z sind separabel) vom System der offenen Mengen in $W \times W$ erzeugt; daher folgt aus der Stetigkeit der s_n deren Messbarkeit und damit auch die Messbarkeit von s. Wird $\sigma: W\times W \longrightarrow \mathbb{R}$ durch $\sigma((w,h)) = \tau(w,h) - \tau(w,-h)$ definiert, dann ist $A = \sigma^{-1}\{0\}$ eine messbare Teilmenge von $W\times W$, woraus für $h \in W$ die Messbarkeit des "h-Schnittes" $\{w \in W \mid (w,h) \in A\}$ von A folgt. Es sei nun $\{w_n\}$ eine (wegen der Separabilität von W existierende) Folge mit $\overline{\{w_n\}} = W$. Da die Gleichung $\tau(w,h) = -\tau(w,-h)$ genau dann für alle $h \in W$ gilt, wenn sie für alle w_n gilt, ergibt sich

$$\begin{aligned} \mathrm{Kompl}(\Delta) &= \bigcap_{h\in W}\{w \in W \mid \sigma((w,h)) = 0\} = \\ &= \bigcap_{n=1}^{\infty}\{w \in W \mid \sigma((w,w_n)) = 0\} = \\ &= \bigcap_{n=1}^{\infty}\{w \in W \mid (w,w_n) \in A\} = \bigcap_{n=1}^{\infty} A^{w_n} \quad (w_n\text{-Schnitt}) \end{aligned}$$

oder

$$\Delta = \bigcup_{n=1}^{\infty} \mathrm{Kompl}(A^{w_n}),$$

woraus die Behauptung folgt.

Es gilt nun der

<u>Satz 3.1.</u> Es sei $(A,b,c) : (\Omega, \mathfrak{A}, P) \longrightarrow L(X,Z)\times Z\times X^*$ ein normintegrierbares stochastisches Programm, wobei Z ein separabler Raum ist. Ist jede Hyperebene in $L(X,Z)\times Z\times X^*$ eine $P_{(A,b,c)}$-Nullmenge und liegt Δ in der Vereinigung von abzählbar vielen Hyperebenen von $W = \mathbb{R}\times Z$, dann ist N_x für jedes $x \in X$ eine P-Nullmenge.

Beweis. Nach dem obigen Lemma ist Δ messbar, ferner gilt $\Delta \subset \bigcup_{n=1}^{\infty} H_n$ mit $H_n = \{w \in W \mid w_n^*(w) = c_n\}$, wobei $w_n^* \in W^*$ und $c_n \in \mathbb{R}$ ist. Wir definieren nun für ein festes $x \in X$ den Operator $T_x : L(X,Z)\times Z\times X^* \longrightarrow W$ durch $T_x(A,b,c) = (cx, Ax-b)$. Da T_x linear und stetig ist, ist $(w_n^* \circ T_x)^{-1}\{c_n\}$ für jedes $n \in \mathbb{N}$ eine Hyperebene in $L(X,Z)\times Z\times X^*$ und somit nach Voraussetzung eine $P_{(A,b,c)}$-Nullmenge. Daher gilt

$$\begin{aligned} P(N) = P_{(cx,Ax-b)}(\Delta) &\leq \sum_{n=1}^{\infty} P_{(cx,Ax-b)}(H_n) = \\ &= \sum_{n=1}^{\infty} P_{(A,b,c)}((w_n^* \circ T_x)^{-1}\{c_n\}) = 0. \end{aligned}$$

Im Fall endlichdimensionaler euklidscher Räume folgt aus dem obigen Satz der

<u>Satz 3.2.</u> Es sei $X=E^n$, $Z=E^m$ und $(A,b,c) : (\Omega, \mathfrak{A}, P) \rightarrow L(X,Z)\times \times Z\times X^*$ ein normintegrierbares stochastisches Programm mit einer in bezug auf das Lebesgue-Mass in $L(X,Z)\times Z\times X^*$ absolut stetigen Verteilung $P_{(A,b,c)}$. Ist die Nutzenfunktion der Ordnung in $\mathbb{E} = \{(c(\omega)x, A(\omega)x-b(\omega)) \mid \omega \in \Omega, x \in X_0 \subset X\}$ der Wert m eines LP $(E^\mu, W, E^\mu_+, M, s)$ mit der Eigenschaft (V), dann ist die Zielfunktion $f(x) = Em((cx, Ax-b))$ des Optimierungsproblems (1) von § 1 G-differenzierbar.

Beweis. Nach II, § 3 und § 4 ist m das Minkowskifunktional eines abgeschlossenen und konvexen Polyeders $\mathbb{P} = \mathrm{conv}\{k_0, k_1, \ldots, k_\mu\}$ in W mit $\theta_W \in \mathring{\mathbb{P}}$. Da $\mathbb{P}$ beschränkt ist, gilt $m(w) = 0$ genau dann, wenn $w = \theta$ ist. Daher liegt ein $w \neq \theta$ genau dann in Δ, falls $\frac{1}{m(w)} w \in \Delta_1 = \{w \in \partial\mathbb{P} \mid Q$ ist in w nicht G-differenzierbar$\}$. Nun liegen diejenigen Randpunkte von $\mathbb{P}$ in denen Q nicht G-differenzierbar ist, in denen also nach Lemma 2.3 mehr als eine zu $\mathbb{P}$ tangentiale Hyperebene existiert, in der Vereinigung aller (endlich vielen) untereinander möglichen Durchschnitte D einer gewissen endlichen Anzahl von Hyperebenen in W. Nach dem Obigen liegt daher $w \in \Delta$ in einer der endlich vielen θ_W und einen der Durchschnitte D enthaltenden Hyperebene von W. Die Behauptung folgt jetzt aus Satz 3.1.

Bezeichnungen

Es sei X ein linearer Raum; mit $\theta=\theta_X$ bezeichnen wir das Nullelement von X. Ist M eine Teilmenge von X, dann heisst $p \in M$ ein algebraisch innerer Punkt von M, falls zu jedem $x \in X$ ein $\varepsilon>0$ existiert, so dass $p+tx \in M$ für alle $|t|\leqslant\varepsilon$. Ein Punkt $p \in X$ heisst algebraischer Randpunkt von M, falls p weder algebraisch innerer Punkt von M noch vom Komplement von M ist. Das algebraisch Innere M^i von M ist die Menge aller algebraisch inneren Punkte von M, M heisst algebraisch offen, wenn $M = M^i$. Die algebraische Hülle M^a von M ist die Vereinigung von M^i mit der Menge der algebraischen Randpunkte von M, M heisst algebraisch abgeschlossen, wenn $M = M^a$. Ist K eine konvexe Teilmenge von X mit $\theta_X \in K^i$, dann nennen wir das Funktional $Q_K(x) = \inf\{\lambda>0 \mid \frac{x}{\lambda} \in K\}$, $x \in X$ das Minkowskifunktional von K. Es sei $X_+ \subset X$ ein Kegel in X, d.h. eine Teilmenge von X mit $X_+ + X_+ = \{x+y \mid x,y \in X_+\} \subset X_+$, $\alpha X_+ = \{\alpha x \mid x \in X_+\} \subset X_+$ für alle $\alpha\geqslant 0$ und $X_+ \cap (-X_+) = \{\theta_X\}$. Existiert bezüglich der durch $x \geqslant y \Longleftrightarrow x-y \in X_+$ definierten Halbordnung in X für jedes Paar $(x,y) \in X\times X$ die Elemente $\sup\{x,y\}$ und $\inf\{x,y\}$, dann heisst X ein Vektorverband, wir setzen dann $x^+ = \sup\{x,\theta_X\}$ und $x^- = (-x)^+$. Eine Ordnungseinheit von X ist ein Element $e \in X_+$, derart dass zu jedem $x \in X$ ein $\alpha\geqslant 0$ mit $\alpha e \geqslant x$ existiert. Ein Funktional $f : X \longrightarrow \mathbb{R}$ heisst positiv, falls $f(x)\geqslant 0$ für alle $x\geqslant\theta_X$. Es sei X und Z je ein topologischer Vektorraum; mit X^* bzw. $L(X,Z)$ bezeichnen wir die Menge aller linearen stetigen Funktionale $x^* : X \longrightarrow \mathbb{R}$ bzw. Operatoren $A : X \longrightarrow Z$; sind X,Z normiert, dann ist $\|A\| = \sup\{\|Ax\| \mid \|x\|=1\}$. Ist X_+ ein Kegel in X, dann ist X^*_+ die Menge aller positiven Elemente von X^*. Mit $\mathcal{O}_X$ bzw. $\mathcal{B}_X$ bezeichnen wir das System der offenen Mengen in X bzw. die von $\mathcal{O}_X$ in X erzeugte σ-Algebra der Borelmengen. Die Messbarkeit einer Abbildung auf bzw. in einem(n) topologischen Raum bezieht sich immer auf die entsprechende σ-Algebra der Borelmengen.

Literatur

[1] Bauer, Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie, de Gruyter, Berlin 1968

[2] Bonic, Linear Functional Analysis, Gordon and Breach, New York 1969

[3] Dunford-Schwartz, Linear Operators I, Wiley, New York 1967

[4] Ferguson, Mathematical Statistics, Academic Press, New York 1967

[5] Köthe, Topologische lineare Räume, Springer, Berlin 1966

[6] Krasnoselskii, Positive Solutions of Operator Equations, Noordhoff, Groningen 1964

[7] Peressini, Ordered Topological Vector Spaces, Harper and Row, New York 1967

[8] Pfanzagel, Theory of Measurement, Physica, Würzburg 1968

[9] Schneeweiss, Entscheidungstheorie bei Risiko, Springer, Berlin 1967

[10] Valentine, Konvexe Mengen, Bibliographisches Institut Mannheim 1968

[11] Kall, Qualitative Aussagen zu einigen Problemen der stochastischen Programmierung, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, Band 6, S.246-272 (1966)

[12] Wessels, Stochastic Programming, Statistica Neerlandica 21, 1967

[13] Wets, Stochastic Programs with Recourse: A survey I, Boeing 1969