Stochastische Optimierung

Mitschrift[1] zur Vorlesung von Prof. Römisch
Humboldt Universität zu Berlin

Wintersemester 2002/2003

[1]Gefundene Fehler bitte an Stefan Vigerske (vigerske@mathematik.hu-berlin.de).

Inhaltsverzeichnis

Kapitel 1

Einleitung

Gegenstand: Optimierungsprobleme mit stochastischen Parametern, unter besonderer Berücksichtigung dynamischer Modelle

Historie: ≈1950, Dantzig, Wets, Birge, Rockafellar, ...; Kall, Ziemba, Prekopa, Dupačova, ...

1.1 Mathematisches Modell

Gegeben ist ein stochastischer (Daten-)Prozeß

$$\xi := \{\xi_t\}_{t=1}^T$$

über $(\Omega, \mathcal{F}, \mathbb{P})$. Dabei ist ξ_1 determiniert mit bekannter Wahrscheinlichkeitsverteilung.

Gesucht ist ein stochastischer (Entscheidungs-)Prozeß

$$x := \{x_t\}_{t=1}^T$$

über $(\Omega, \mathcal{F}, \mathbb{P})$. Dabei ist x_1 determiniert und x_t nicht antizipativ, d.h. x_t hängt nur von $\xi_1, \ldots, \xi_t$ ab, bzw. x_t ist $\mathcal{F}_t$-meßbar, wobei

$$\mathcal{F}_t := \sigma\{\xi_1, \ldots, \xi_t\}.$$

Dann ist

$$\{\emptyset, \Omega\} = \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_{t-1} \subseteq \mathcal{F}_t \subseteq \ldots \subseteq \mathcal{F}_T = \mathcal{F}$$

Minimiere

$$\mathbb{E}[f(x,\xi) = f(x_1, \ldots, x_T, \xi_1, \ldots, \xi_T)]$$

$$\text{bzgl. } x_t \in X_t, \quad g_t(x_1, \ldots, x_t, \xi_t) \leq 0 \quad (\mathbb{P}\text{-fast-überall}), \quad x_t \text{ ist } \mathcal{F}_t\text{-meßbar.}$$

Oder eine allgemeinere Zielfunktion $\mathbb{F}[f(x,\xi)]$, wobei $\mathbb{F}$ ein Risiko-Funktional auf einem linearen Raum von reellen Zufallsvariablen ist.

Eine spezielle Form ist

$$f(x,\xi) = \sum_{t=1}^T f_t(x_t, \xi_t),$$

speziell wenn $f_t(\cdot, \xi_t)$ konvex ist. Im linearen Fall ist

$$\begin{aligned} f(x,\xi) &= \sum_{t=1}^T \langle c_t(\xi_t), x_t \rangle \\ g_t(x_1, \ldots, x_t, \xi_t) &= A_{tt}(\xi_t)\, x_t + \sum_{\tau=1}^{t-1} A_{t\tau}(\xi_\tau)\, x_\tau - b_t(\xi_t) \end{aligned}$$

1.2 Anwendungsbeispiele

Beispiel 1.1 (Blumen-Junge).

- a_t Preis für den Kauf einer Lilie am Tag t, $c_t > a_t$ Verkaufspreis
- $d_t = \xi_t$ zufälliger Bedarf an Lilien am Tag t
- x_t Anzahl der gekauften Lilien am Tag t
- z_t Anzahl der vernichteten Lilien am Tag t, s_{t-1} Anzahl der vom Vortag übrigen Lilien

Ziel: Gewinnmaximierung:

$$\max \left\{ \begin{array}{l} (c_1 - a_1)(x_1 - s_1) \\ +\mathbb{E}\left[\sum_{t=2}^{T} (c_t - a_t)(x_t + s_{t-1} - s_t) - c_t z_t \right] \end{array} \middle| \begin{array}{ll} x_t \geq 0, z_t \geq 0, s_t \geq 0, & t = 1, \ldots, T \\ z_t \geq s_{t-1} - d_t, & t = 2, \ldots, T \\ s_t + z_t \geq x_t + s_{t-1} - d_t, & t = 2, \ldots, T \\ (x_t, z_t, s_t)\ \mathcal{F}_t\text{-meßbar} & t = 1, \ldots, T \\ s_0 = 0, \quad z_1 = 0 & \end{array} \right\}$$

Beispiel 1.2 (stochastisches Strom-Management).

- I Menge von Erzeugereinheiten und Verträge
- T Anzahl von Zeitperioden
- $\{\xi_t\}_{t=1}^{T}$ Prozeß bestehend aus stochastischer Last/Preise/Zuflüsse
- $S_i(\xi)$ Restriktionsmenge für Einheit i
- $p_{i,t}$ Entscheidung für i und t
- $r_i(\xi_t, p_{i,t})$ Gewinn der Entscheidung $p_{i,t}$ in der Einheit i in t
- $C(\xi)$ Gleichgewichtsrestriktionen, Lastdeckung, Reserve

$$\max \left\{ \mathbb{E}\left[\sum_{i \in I} \sum_{t=1}^{T} r_i(\xi_t, p_{i,t}) \right] \middle| \begin{array}{ccc} p_i & \in & S_i(\xi) \\ p & \in & C(\xi) \end{array} \right\}$$

Beispiel 1.3 (OD-Optimierung von Airlines).

- I origins i, J destinations j
- $L \subseteq I \times J$ legs l (Flug von i nach j)
- $l \in L$, $k_l \in \mathbb{N}$ Buchungsklassen
- $x_{i,j,k}$ Buchungsgrenze (maximale Anzahl der Sitze in Klasse k von i nach j)
- c_l Kapazität von leg l von i nach j

$$x_{i,j,k} \geq 0, \qquad \sum_{k=1}^{k_l} x_{i,j,k} \leq c_l, \quad l = l(i,j)$$

- $f_{i,j,k}$ Buchungspreise für Klasse k von i nach j
- $d_{i,j,k}$ Passagieraufkommen für Klasse k von i nach j

$$\max \left\{ \mathbb{E}\left[\sum_{i,j,k} f_{i,j,k} x_{i,j,k} \right] \middle| 0 \leq x_{i,j,k} \leq d_{i,j,k}, \quad \sum_{k=1}^{k_l} x_{i,j,k} \leq c_l \right\}$$

Ziel: dynamische Modelle, Buchungssteuerung, dynamische Preise $f_{i,j,k,t}$, Grenzen $x_{i,j,k,t}$, Passagieraufkommen $d_{i,j,k,t}$

Kapitel 2

Optimalität, Dualität und Stabilität

2.1 Meßbarkeit

Ziel. Theorie meßbarer mengenwertiger Abbildungen und normaler Integranden.

$(\Omega, \mathcal{F})$ bezeichne einen meßbaren Raum, manchmal auch $(\Omega, \mathcal{F}, \mathbb{P})$ mit dem Wahrscheinlichkeitsmaß $\mathbb{P}$. $\mathcal{B}(\mathbb{R}^m)$ sei die Borel-σ-Algebra. Sei

$$\mathbb{B}(x, r) := \{y \in \mathbb{R}^m \,|\, \|y - x\| \le r\}$$

die abgeschlossene Kugel um $x \in \mathbb{R}^m$ mit Radius r (bzgl. einer Norm $\|\cdot\|$) und bezeichne

$$\mathbb{B} := \mathbb{B}(0, 1) = \{y \in \mathbb{R}^m \,|\, \|y\| \le 1\}$$

die abgeschlossene Einheitskugel im $\mathbb{R}^m$. Definiere den Abstand eines Punktes $x \in \mathbb{R}^m$ zu einer Menge $\emptyset \ne B \subseteq \mathbb{R}^m$ mittels

$$d(x, B) := \inf_{y \in B} \|x - y\|$$

und $d(x, \emptyset) := +\infty$. Dann gilt $|d(x, B) - d(\tilde{x}, B)| \le \|x - \tilde{x}\|$ für alle $x, \tilde{x} \in \mathbb{R}^m$ und $\emptyset \ne B \subseteq \mathbb{R}^m$.

Für eine mengenwertige Abbildung $\Gamma : \Omega \rightrightarrows \mathbb{R}^m$ von Ω in den $\mathbb{R}^m$, d.h. $\Gamma(\omega) \subseteq \mathbb{R}^m$, $\omega \in \Omega$, sei

$$\begin{aligned} \Gamma^{-1}(x) &:= \{\omega \in \Omega \,|\, x \in \Gamma(\omega)\}, \qquad x \in \mathbb{R}^m, \\ \Gamma^{-1}(B) &:= \bigcup_{x \in B} \Gamma^{-1}(x), \qquad B \subseteq \mathbb{R}^m. \end{aligned}$$

Dann ist

$$\Gamma^{-1}(B) = \{\omega \in \Omega \,|\, \Gamma(\omega) \cap B \ne \emptyset\}.$$

Sei weiterhin

$$\mathrm{dom}\Gamma := \Gamma^{-1}(\mathbb{R}^m) = \{\omega \in \Omega \,|\, \Gamma(\omega) \ne \emptyset\}$$

und

$$\mathrm{gph}\Gamma := \{(\omega, x) \in \Omega \times \mathbb{R}^m \,|\, x \in \Gamma(\omega)\}.$$

Satz 2.1. *Es sei $\Gamma : \Omega \rightrightarrows \mathbb{R}^m$ abgeschlossen-wertig. Dann sind die folgenden Aussagen äquivalent:*

1. *Γ ist meßbar, d.h. $\Gamma^{-1}(O) \in \mathcal{F}$ für alle $O \subseteq \mathbb{R}^m$ offen.*
2. *$\Gamma^{-1}(D) \in \mathcal{F}$ für alle $D \subseteq \mathbb{R}^m$ abgeschlossen.*

3. *Die Funktion* $d(x, \Gamma(\cdot)) : \Omega \to \bar{\mathbb{R}}$ *ist meßbar für alle* $x \in \mathbb{R}^m$.

4. $\mathrm{dom}\Gamma \in \mathcal{F}$ *und es existiert eine abzählbare Familie* $\{x_n\}_{n\in\mathbb{N}}$ *meßbarer Abbildungen von* $\mathrm{dom}\Gamma$ *in* $\mathbb{R}^m$, *so daß* (Castaing-Selektion)

$$\Gamma(\omega) = c\ell\{x_n(\omega) \,|n \in \mathbb{N}\} \quad \forall \omega \in \mathrm{dom}\Gamma.$$

Beweis. 1.⇒2. Sei $D \subset \mathbb{R}^m$, $D \neq \emptyset$, abgeschlossen. Dann definiere die offenen Mengen

$$D_n := \left\{ x \in \mathbb{R}^m \,\middle|\, d(x, D) < \frac{1}{n} \right\} \qquad \forall n \in \mathbb{N}.$$

Offensichtlich ist $D \subseteq D_n$, $n \in \mathbb{N}$, und $D = \bigcap_{n\in\mathbb{N}} D_n$. Nach Voraussetzung ist $\Gamma^{-1}(D_n) \in \mathcal{F}$, also folgt

$$\Gamma^{-1}(D) = \bigcap_{n\in\mathbb{N}} \Gamma^{-1}(D_n) \in \mathcal{F}.$$

2.⇒1. Sei $O \subseteq \mathbb{R}^m$, $O \neq \emptyset$, offen. Dann läßt sich O als abzählbare Vereinigung $O = \bigcup_{n\in\mathbb{N}} \mathbb{B}(x_n, \varepsilon_n)$ darstellen. Wegen $\Gamma^{-1}(\mathbb{B}(x_n, \varepsilon_n)) \in \mathcal{F}$ folgt dann

$$\Gamma^{-1}(O) = \bigcup_{n\in\mathbb{N}} \Gamma^{-1}(\mathbb{B}(x_n, \varepsilon_n)) \in \mathcal{F}.$$

2.⇔3. [RW98, Theorem 14.3] Für $x \in \mathbb{R}^m$ und $\alpha \in [0, +\infty]$ ist $d(x, \Gamma(\omega)) \leq \alpha$, gdw. $\Gamma(\omega) \cap \mathbb{B}(x, \alpha) \neq \emptyset$. Also ist

$$\{\omega \in \Omega \,|d(x, \Gamma(\omega)) \leq \alpha\} = \Gamma^{-1}(\mathbb{B}(x, \alpha)).$$

2.⇒3. Da $\mathbb{B}(x, \alpha)$ abgeschlossen ist, gilt $\Gamma^{-1}(\mathbb{B}(x, \alpha)) \in \mathcal{F}$, also $\{\omega \in \Omega \,|d(x, \Gamma(\omega)) \leq \alpha\} \in \mathcal{F}$, d.h. $d(x, \Gamma(\cdot))$ ist meßbar.

3.⇒2. Sei $D \subseteq \mathbb{R}^n$ abgeschlossen. Dann existieren $x_n \in D$, $\varepsilon_n \in \mathbb{R}_+$, $n \in \mathbb{N}$, so daß $D = \bigcup_{n\in\mathbb{N}} \mathbb{B}(x_n, \varepsilon_n)$. Also ist

$$\Gamma^{-1}(D) = \bigcup_{n\in\mathbb{N}} \Gamma^{-1}(\mathbb{B}(x_n, \varepsilon_n)) = \bigcup_{n\in\mathbb{N}} \{\omega \in \Omega \,|d(x_n, \Gamma(\omega)) \leq \varepsilon_n\} \in \mathcal{F}.$$

4.⇒1. Sei $O \subseteq \mathbb{R}^m$ offen. Dann ist

$$\Gamma^{-1}(O) = \bigcup_{n\in\mathbb{N}} \{\omega \in \mathrm{dom}\Gamma \,|x_n(\omega) \in O\} = \left(\bigcup_{n\in\mathbb{N}} x_n^{-1}(O) \right) \cap \mathrm{dom}\Gamma \in \mathcal{F},$$

da die x_n, $n \in \mathbb{N}$, meßbar sind.

1.⇒4. [RW98, Theorem 14.5] □

Satz 2.2. *Sei* $\Gamma : \Omega \rightrightarrows \mathbb{R}^m$ *abgeschlossen-wertig und stetig. Ist* Γ *meßbar, so gilt*

$$\mathrm{gph}\Gamma \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m) := \sigma\{A \times B \,|A \in \mathcal{F}, B \in \mathcal{B}(\mathbb{R}^m)\}.$$

Die Umkehrung gilt, falls $\{\Omega, \mathcal{F}, \mathbb{P}\}$ *vollständig*[1] *ist.*

Beweis. (⇐) Sei $O \subseteq \mathbb{R}^m$ offen. Dann ist $G := \mathrm{gph}\Gamma \cap \{\Omega \times O\} \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$ und es ist

$$\Gamma^{-1}(O) = \{\omega \in \Omega \,|\exists x \in \mathbb{R}^m \text{ mit } (\omega, x) \in G\} = \mathrm{Pr}_\Omega(G).$$

Wegen der Vollständigkeit von $(\Omega, \mathcal{F}, \mathbb{P})$ folgt $\mathrm{Pr}_\Omega(G) \in \mathcal{F}$ (Debreu '67, Sainte-Beuve '74).

[1] $\{\Omega, \mathcal{F}, \mathbb{P}\}$ heißt vollständig, falls für $A \in \mathcal{F}$ mit $\mathbb{P}(A) = 0$ gilt: $\tilde{A} \subseteq A \Rightarrow \tilde{A} \in \mathcal{F}$.

($\Rightarrow$) Sei Γ meßbar. Es gilt $(\omega, x) \in \mathrm{gph}\Gamma \Leftrightarrow \omega \in \Omega$ und $x \in \Gamma(\omega) \Leftrightarrow \omega \in \Omega$ und $\forall \rho \in \mathbb{Q}_+ \exists a \in \mathbb{Q}^m$ mit $x \in \mathbb{B}(a, \rho)$ und $\Gamma(\omega) \cap \mathbb{B}(a, \rho) \neq \emptyset$. Dann ist

$$\mathrm{gph}\Gamma = \bigcap_{\rho \in \mathbb{Q}_+} \bigcup_{a \in \mathbb{Q}^m} \underbrace{\Gamma^{-1}(\mathbb{B}(a, \rho))}_{\in \mathcal{F}, \text{ wg. Satz 2.1}} \times \underbrace{\mathbb{B}(a, \rho)}_{\in \mathcal{B}(\mathbb{R}^m)} \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$$

□

Satz 2.3. *Es seien $\Gamma_j : \Omega \rightrightarrows \mathbb{R}^m$, $j \in J$, abgeschlossen-wertig und meßbar, J sei höchstens abzählbar. Dann sind die mengenwertigen Abbildungen*

$$\omega \mapsto \bigcup_{j \in J} \Gamma_j(\omega), \qquad \omega \mapsto \bigcap_{j \in J} \Gamma_j(\omega)$$

meßbar.

Beweis. Sei

$$\Gamma(\omega) := \bigcup_{j \in J} \Gamma_j(\omega) \qquad \forall \omega \in \Omega,$$

und $D \subseteq \mathbb{R}^m$ abgeschlossen. Dann ist

$$\Gamma^{-1}(D) = \{\omega \in \Omega \,|\, \Gamma(\omega) \cap D \neq \emptyset\} = \bigcup_{j \in J} \underbrace{\Gamma_j^{-1}(D)}_{\in \mathcal{F}} \in \mathcal{F}.$$

Sei nun

$$\Gamma(\omega) := \bigcap_{j \in J} \Gamma_j(\omega) \qquad \forall \omega \in \Omega.$$

Wir beweisen die Meßbarkeit für den Fall, daß $(\Omega, \mathcal{F}, \mathbb{P})$ vollständig ist (ansonsten siehe [RW98, Theorem 14.11]). Es ist

$$\mathrm{gph}\Gamma = \left\{ (\omega, x) \in \Omega \times \mathbb{R}^m \,\middle|\, x \in \Gamma(\omega) = \bigcap_{j \in J} \Gamma_j(\omega) \right\} = \bigcap_{j \in J} \mathrm{gph}\Gamma_j \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$$

und $\mathrm{gph}\Gamma_j \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$ nach Satz 2.2. Mit Satz 2.2 folgt, daß Γ meßbar ist. □

Wir vereinbaren

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

Definition (normaler Integrand). Eine Funktion $f : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ heißt *normaler Integrand*, falls die mengenwertige Abbildung $E_f : \Omega \rightrightarrows \mathbb{R}^m \times \mathbb{R}$

$$\omega \mapsto E_f(\omega) := \mathrm{epi} f(\omega, \cdot) := \{(x, \alpha) \in \mathbb{R}^m \times \mathbb{R} \,|\, f(\omega, x) \leq \alpha\}$$

abgeschlossen-wertig und meßbar ist.

Satz 2.4. *Es sei $f : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ ein normaler Integrand. Dann gilt*

1. *$f(\omega, \cdot) : \mathbb{R}^m \to \bar{\mathbb{R}}$ ist unterhalbstetig für alle $\omega \in \Omega$, d.h. $\liminf_{x \to \bar{x}} f(\omega, x) \geq f(\omega, \bar{x})$ $\forall \bar{x} \in \mathbb{R}^m$.*
2. *Ist $x : \Omega \to \mathbb{R}^m$ meßbar, so ist $f(\cdot, x(\cdot)) : \Omega \to \bar{\mathbb{R}}$ meßbar.*
   $f(\cdot, x) : \Omega \to \bar{\mathbb{R}}$ ist meßbar für alle $x \in \mathbb{R}^m$.
3. *f ist $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$-meßbar.*

Ist $(\Omega, \mathcal{F}, \mathbb{P})$ vollständig, f $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$-meßbar und $f(\omega, \cdot)$ unterhalbstetig, so ist f ein normaler Integrand.

Beweis. 1. Sei $\bar{x} \in \mathbb{R}^m$, $\bar{\alpha} := \liminf\limits_{x \to \bar{x}} f(\omega, x)$ und $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^m$ mit $x_n \to \bar{x}$ und $f(\omega, x_n) \to \bar{\alpha}$. Wegen $(x_n, f(\omega, x_n)) \in E_f(\omega)$ folgt aus der Abgeschlossenheit von $E_f(\omega)$, daß $(\bar{x}, \bar{\alpha}) = \lim\limits_{n\to\infty} (x_n, f(\omega, x_n)) \in E_f(\omega)$, also $f(\omega, \bar{x}) \leq \bar{\alpha}$.

2. Sei $x : \Omega \to \mathbb{R}^m$ meßbar und $R_\alpha : \Omega \rightrightarrows \mathbb{R}^m \times \mathbb{R}$,

$$R_\alpha(\omega) := \{x(\omega)\} \times (-\infty, \alpha] \qquad \forall \omega \in \Omega, \alpha \in \mathbb{R}.$$

Dann gilt für beliebige $\alpha \in \mathbb{R}$:

$$\{\omega \in \Omega \,|\, f(\omega, x(\omega)) \leq \alpha\} = \{\omega \in \Omega \,|\, E_f(\omega) \cap R_\alpha(\omega) \neq \emptyset\} = \mathrm{dom}\,(E_f(\cdot) \cap R_\alpha(\cdot)) \in \mathcal{F},$$

falls $E_f(\cdot) \cap R_\alpha(\cdot)$ meßbar ist. Dies gilt, da E_f und R_α meßbar sind und wegen Satz 2.3.
Sei $x(\omega) \equiv x \in \mathbb{R}^m$. Dann ist $f(\cdot, x)$ meßbar.

3. Sei $\alpha : \Omega \to \bar{\mathbb{R}}$ meßbar und wir betrachten $L : \Omega \rightrightarrows \mathbb{R}^m$.

$$L(\omega) := \{x \in \mathbb{R}^m \,|\, f(\omega, x) \leq \alpha(\omega)\}, \quad \forall \omega \in \Omega$$

Dann ist L abgeschlossen-wertig.
Wir zeigen: L ist meßbar.
Sei $D \subseteq \mathbb{R}^m$ abgeschlossen und wir betrachten $R : \Omega \rightrightarrows \mathbb{R}^m \times \mathbb{R}$,

$$R(\omega) := D \times (-\infty, \alpha(\omega)].$$

R ist meßbar und daher gilt

$$\begin{aligned} L^{-1}(D) &= \{\omega \in \Omega \,|\, L(\omega) \cap D \neq \emptyset\} \\ &= \{\omega \in \Omega \,|\, E_f(\omega) \cap R(\omega) \neq \emptyset\} \\ &= \mathrm{dom}\,(E_f(\cdot) \cap R(\cdot)) \in \mathcal{F}, \end{aligned}$$

d.h. L ist meßbar (Satz 2.1). Für $\alpha(\omega) \equiv \alpha \in \mathbb{R}$ gilt nach Satz 2.2 speziell:

$$\{(\omega, x) \in \Omega \times \mathbb{R}^m \,|\, f(\omega, x) \leq \alpha\} = \mathrm{gph} L \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m).$$

Dann ist f $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m)$-meßbar.

Der Rest folgt analog zu Satz 2.2. □

Beispiel 2.5. Es sei $\Gamma : \Omega \rightrightarrows \mathbb{R}^m$ abgeschlossen-wertig und meßbar.
Wir betrachten die Indikator-Funktion $\delta_\Gamma : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$,

$$\delta_\Gamma(\omega, x) := \begin{cases} 0, & x \in \Gamma(\omega) \\ \infty, & x \notin \Gamma(\omega) \end{cases}.$$

$\delta_\Gamma(\omega, \cdot)$ ist unterhalbstetig.

δ_Γ ist ein normaler Integrand gdw. Γ abgeschlossenwertig und meßbar ist.

$$\left(\mathrm{epi}\delta_\Gamma(\omega, \cdot) = \{(x, \alpha) \in \mathbb{R}^m \times \mathbb{R} \,|\, \delta_\Gamma(\omega, x) \leq \alpha\} = \begin{cases} \emptyset, & \omega \notin \mathrm{dom}\Gamma \\ \Gamma(\omega) \times [0, \infty), & \omega \in \mathrm{dom}\Gamma \end{cases} \right)$$

Wir betrachten die Optimierungsaufgabe:

$$\min \{f_0(\omega, x) \,|\, x \in \Gamma(\omega)\}$$

Äquivalent dazu ist

$$\min \{f_0(\omega, x) + \delta_\Gamma(\omega, x) \,|\, x \in \mathbb{R}^m\}$$

Falls $f_0 : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ ein normaler Integrand ist, so ist auch $f_0 + \delta_\Gamma$ ein normaler Integrand.

Folgerung. 1. Es sei $f : \Omega \times \mathbb{R}^m \to \mathbb{R}$ eine Carathéodory-Funktion, d.h. $f(\cdot, x)$ ist meßbar für alle $x \in \mathbb{R}^m$, und $f(\omega, \cdot)$ ist stetig für alle $\omega \in \Omega$.

Dann ist f ein normaler Integrand.

2. Es seien $f_j : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ normale Integranden und $\alpha_j : \Omega \to \bar{\mathbb{R}}$ meßbar für alle $j \in J$, wobei J höchstens abzählbar ist. Dann ist

$$\omega \mapsto \Gamma(\omega) := \{x \in \mathbb{R}^m \,|\, f_j(\omega, x) \leq \alpha_j(\omega) \quad \forall j \in J\}$$

abgeschlossen-wertig und meßbar.

Beweis. 1. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ vollständig (allgemeiner Fall: [RW98, 14.29]). Für $\alpha \in \mathbb{R}$ gilt

$$\{(\omega, x) \,|\, f(\omega, x) \leq \alpha\} = \bigcap_{\rho \in \mathbb{Q}_+} \bigcup_{a \in \mathbb{Q}^m} \underbrace{\{\omega \in \Omega \,|\, f(\omega, a) \leq \alpha + \rho\}}_{\in \mathcal{F}} \times \mathbb{B}(a, \rho) \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^m).$$

Mit Satz 2.4 folgt dann, daß f ein normaler Integrand ist.

2. Aus dem Beweis von Satz 2.4, 3., folgt, daß die Funktionen $L_j : \Omega \rightrightarrows \mathbb{R}^m$, $L_j(\omega) := \{x \in \mathbb{R}^m \,|\, f_j(\omega, x) \leq \alpha_j(\omega)\}$, $j \in J$, meßbar und abgeschlossen-wertig sind. Nach Satz 2.3 ist dann auch $\omega \mapsto \Gamma(\omega) = \bigcap_{j \in J} L_j(\omega)$ meßbar und abgeschlossen-wertig. □

Satz 2.6. *Es sei $f : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ ein normaler Integrand und wir betrachten*

$$\begin{aligned} v(\omega) &:= \inf_{x \in \mathbb{R}^m} f(\omega, x), \qquad \forall \omega \in \Omega, \\ S(\omega) &:= \operatorname*{argmin}_{x \in \mathbb{R}^m} f(\omega, x) = \{x \in \mathbb{R}^m \,|\, f(\omega, x) \leq v(\omega)\}, \qquad \forall \omega \in \Omega. \end{aligned}$$

Dann ist $v : \Omega \to \bar{\mathbb{R}}$ meßbar und $S : \Omega \rightrightarrows \mathbb{R}^m$ abgeschlossen-wertig und meßbar.

Beweis. Für $\alpha \in \mathbb{R}$ ist

$$\begin{aligned} \{\omega \in \Omega \,|\, v(\omega) < \alpha\} &= \{\omega \in \Omega \,|\, \exists x \in \mathbb{R}^m : f(\omega, x) < \alpha\} \\ &= \{\omega \in \Omega \,|\, E_f(\omega) \cap (\mathbb{R}^m \times (-\infty, \alpha)) \neq \emptyset\} \in \mathcal{F}, \end{aligned}$$

also ist v meßbar. Die Meßbarkeit von S folgt aus der letzten Folgerung. S ist abgeschlossen-wertig, da $f(\omega, \cdot)$ unterhalbstetig ist. □

Es sei $\mathcal{G} \subset \mathcal{F}$ eine weitere σ-Algebra.

$$\begin{aligned} \mathcal{N}(\mathcal{G}, \mathbb{R}^m) &:= \left\{f : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}} \,|\, f \text{ ist normaler Integrand bzgl. } (\Omega, \mathcal{G})\right\} \\ \mathcal{N}_c(\mathcal{G}, \mathbb{R}^m) &:= \{f \in \mathcal{N}(\mathcal{G}, \mathbb{R}^m) \,|\, \{x \in \mathbb{R}^m : f(\omega, x) \leq \alpha\} \text{ ist kompakt } \forall \omega \in \Omega, \alpha \in \mathbb{R}\} \end{aligned}$$

Satz 2.7 (inf-**Projektion normaler Integranden).** *Es sei $f \in \mathcal{N}_c\left(\mathcal{G}, \mathbb{R}^m \times \mathbb{R}^k\right)$ und $\bar{f} : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$,*

$$\bar{f}(\omega, x) := \inf_{y \in \mathbb{R}^k} f(\omega, x, y) \qquad \forall (\omega, x) \in \Omega \times \mathbb{R}^m$$

Dann gilt: $\bar{f} \in \mathcal{N}_c(\mathcal{G}, \mathbb{R}^m)$.

Beweis. 1. Wir zeigen, daß das Minimum angenommen wird (falls es existiert), d.h. für alle $(\omega, \bar{x}) \in \Omega \times \mathbb{R}^m$ existiert ein $\bar{y} \in \mathbb{R}^k$, so daß $\bar{f}(\omega, \bar{x}) = f(\omega, \bar{x}, \bar{y})$.

Sei $(\omega, \bar{x}) \in \Omega \times \mathbb{R}^m$. Falls $\bar{f}(\omega, \bar{x}) = +\infty$, so ist $f(\omega, \bar{x}, y) = +\infty$ für alle $y \in \mathbb{R}^k$ (trivial).

Sei nun $\bar{f}(\omega, \bar{x}) < +\infty$. Dann existiert eine Folge $\{y_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^k$ mit

$$\bar{f}(\omega, \bar{x}) = \lim_{n \to \infty} f(\omega, \bar{x}, y_n).$$

Also existiert ein $n_0 \in \mathbb{N}$ und $\alpha \in \mathbb{R}$, so daß $f(\omega, \bar{x}, y_n) \leq \alpha$ für alle $n \geq n_0$. Die Menge $\left\{(\bar{x}, y_n) \in \mathbb{R}^n \times \mathbb{R}^k \,|\, f(\omega, \bar{x}, y_n) \leq \alpha, n \geq n_0\right\}$ ist nach Voraussetzung beschränkt, also existiert eine konvergente Teilfolge $\{y_{n_l}\}_{l \in \mathbb{N}} \subseteq \{y_n\}_{n \in \mathbb{N}}$, $y_{n_l} \to \bar{y} \in \mathbb{R}^k$. Aus der Unterhalbstetigkeit von f folgt nun

$$\bar{f}(\omega, \bar{x}) = \lim_{l \to \infty} f(\omega, \bar{x}, y_{n_l}) \geq f(\omega, \bar{x}, \bar{y}) \geq \bar{f}(\omega, \bar{x}),$$

also $\bar{f}(\omega, \bar{x}) = \bar{f}(\omega, \bar{x}, \bar{y})$.

2. Wir zeigen, daß $\bar{f}(\omega, \cdot)$ unterhalbstetig ist für alle $\omega \in \Omega$.

   Sei $\omega \in \Omega$ und $\{x_n\}_{n \in \mathbb{N}}$ eine Folge in $\mathbb{R}^m$ mit $x_n \to \bar{x}$. Nach 1. existieren $y_n \in \mathbb{R}^k$ mit

   $$\bar{f}(\omega, x_n) = f(\omega, x_n, y_n) \quad \forall n \in \mathbb{N}.$$

   Es ist zu zeigen: $\liminf_{n \to \infty} \bar{f}(\omega, x_n) \geq \bar{f}(\omega, \bar{x})$.

   O.B.d.A. gelte $\liminf_{n \to \infty} \bar{f}(\omega, x_n) = \lim_{n \to \infty} \bar{f}(\omega, x_n) < \infty$.

   Dann existiert ein $\alpha \in \mathbb{R}$ und $n_0 \in \mathbb{N}$, so daß $f(\omega, x_n, y_n) \leq \alpha$, $n \geq n_0$, also ist $\{(x_n, y_n)\}_{n \geq n_0}$ nach Voraussetzung an f beschränkt. Dann existiert eine konvergente Teilfolge $\{y_{n_l}\}_{l \in \mathbb{N}} \subseteq \{y_n\}_{n \in \mathbb{N}}$, $(x_{n_l}, y_{n_l}) \to (\bar{x}, \bar{y})$, und es folgt

   $$\liminf_{n \to \infty} \bar{f}(\omega, x_n) = \lim_{l \to \infty} \bar{f}(\omega, x_{n_l}) = \lim_{l \to \infty} f(\omega, x_{n_l}, y_{n_l}) \geq f(\omega, \bar{x}, \bar{y}) \geq \bar{f}(\omega, \bar{x}).$$

   Aus der Unterhalbstetigkeit von $\bar{f}(\omega, \cdot)$ für alle $\omega \in \Omega$ folgt die Abgeschlossenheit von $\operatorname{epi}\bar{f}(\omega, \cdot)$, $\omega \in \Omega$.

3. $\mathcal{G}$-Meßbarkeit von $\omega \mapsto \operatorname{epi}\bar{f}(\omega, \cdot)$. Es ist

   $$\begin{aligned} E_{\bar{f}}(\omega) &= \left\{(x, \alpha) \in \mathbb{R}^m \times \mathbb{R} \,\middle|\, \bar{f}(\omega, x) \leq \alpha\right\} \\ &= \left\{(x, \alpha) \in \mathbb{R}^m \times \mathbb{R} \,\middle|\, \exists \bar{y} \in \mathbb{R}^k : f(\omega, x, \bar{y}) \leq \alpha\right\} \\ &= P \circ E_f(\omega), \end{aligned}$$

   wobei $P : \mathbb{R}^m \times \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}^m \times \mathbb{R}$ mittels $P(x, y, \alpha) := (x, \alpha)$ definiert ist. P ist meßbar.

   Sei $O \subseteq \mathbb{R}^m \times \mathbb{R}$ offen. Dann ist

   $$\begin{aligned} E_{\bar{f}}^{-1}(O) = \left\{\omega \in \Omega \,\middle|\, E_{\bar{f}}(\omega) \cap O \neq \emptyset\right\} &= \{\omega \in \Omega \,|\, P \circ E_f(\omega) \cap O \neq \emptyset\} \\ &= \left\{\omega \in \Omega \,\middle|\, E_f(\omega) \cap P^{-1}(O) \neq \emptyset\right\} \\ &= E_f^{-1}\left(P^{-1}(O)\right) \in \mathcal{G}. \end{aligned}$$

   Aus 2. und 3. folgt nun $\bar{f} \in \mathcal{N}(\mathcal{G}, \mathbb{R}^m)$.

4. Sei $\omega \in \Omega$ und $\alpha \in \mathbb{R}$. Zu zeigen: $\left\{x \in \mathbb{R}^m \,\middle|\, \bar{f}(\omega, x) \leq \alpha\right\}$ ist kompakt.

   Sei $\{x_n\}_{n \in \mathbb{N}}$ eine Folge mit $\bar{f}(\omega, x_n) \leq \alpha$ $\forall n \in \mathbb{N}$. Dann existiert eine Folge $\{y_n\}_{n \in \mathbb{N}} \subset \mathbb{R}^k$ mit $\bar{f}(\omega, x_n) = f(\omega, x_n, y_n) \leq \alpha$. Aufgrund der Voraussetzungen an f existiert dann wieder eine konvergente Teilfolge von $\{(x_n, y_n)\}_{n \in \mathbb{N}} \subset \mathbb{R}^m \times \mathbb{R}^k$, also auch von $\{x_n\}_{n \in \mathbb{N}}$. □

Es sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, $\mathcal{G} \subset \mathcal{F}$ eine σ-Algebra.

$$\mathcal{N}_i(\mathcal{G}, \mathbb{R}^m) := \left\{ f : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}} \;\middle|\; \begin{array}{rcl} \exists \bar{f} \in \mathcal{N}(\mathcal{G}, \mathbb{R}^m),\ q : \Omega \to \bar{\mathbb{R}} & : & \\ \int_\Omega |q(\omega)| \, \mathbb{P}(d\omega) & < & \infty \\ \forall x \in \mathbb{R}^m : \quad f(\omega, x) & \geq & q(\omega) \ \mathbb{P}\text{-f.ü.} \\ \mathbb{P}\left(f(\omega, x) = \bar{f}(\omega, x) \, \forall x \in \mathbb{R}^m\right) & = & 1 \end{array} \right\}$$

$$\mathcal{N}_{ci}(\mathcal{G}, \mathbb{R}^m) := \left\{ f : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}} \;\middle|\; \begin{array}{rcl} \exists \bar{f} \in \mathcal{N}_c(\mathcal{G}, \mathbb{R}^m),\ q : \Omega \to \bar{\mathbb{R}} & : & \\ \int_\Omega |q(\omega)| \, \mathbb{P}(d\omega) & < & \infty \\ \forall x \in \mathbb{R}^m : \quad f(\omega, x) & \geq & q(\omega) \ \mathbb{P}\text{-f.ü.} \\ \mathbb{P}\left(f(\omega, x) = \bar{f}(\omega, x) \, \forall x \in \mathbb{R}^m\right) & = & 1 \end{array} \right\}$$

Definition (bedingte Erwartung). Für $z : \Omega \to \bar{\mathbb{R}}$ $\mathcal{F}$-meßbar mit $z(\omega) \geq q(\omega)$ $\mathbb{P}$-f.ü., $\int_\Omega |q(\omega)| \, \mathbb{P}(d\omega) < \infty$, heißt

$$y =: \mathbb{E}[z \,|\mathcal{G}] : \Omega \to \bar{\mathbb{R}}$$

bedingte Erwartung von z bzgl. $\mathcal{G}$, falls y $\mathcal{G}$-meßbar ist und

$$\int_A y(\omega) \, \mathbb{P}(d\omega) = \int_A z(\omega) \, \mathbb{P}(d\omega) \quad \forall A \in \mathcal{G}.$$

Die bedingte Erwartung von z bzgl. $\mathcal{G}$ existiert und ist $\mathbb{P}$-f.ü. eindeutig. ([Sh96, Kap. II.7])

Frage: Gegeben ist $f \in \mathcal{N}_i(\mathcal{F}, \mathbb{R}^m)$, $x : \Omega \to \mathbb{R}^m$ $\mathcal{G}$-meßbar. Dann ist $f(\cdot, x(\cdot))$ nach Satz 2.4, 2., $\mathcal{F}$-meßbar. Ist die Abbildung $\mathbb{E}[f(\cdot, x(\cdot)) \,|\mathcal{G}] : \Omega \to \bar{\mathbb{R}}$, die nach Definition $\mathcal{G}$-meßbar ist, darstellbar in der Form $g(\cdot, x(\cdot))$ mit $g \in \mathcal{N}_i(\mathcal{G}, \mathbb{R}^m)$?

Satz 2.8 (Dynkin/Evstigneev '76 [DE76, 325-328]). *Für jede σ-Algebra $\mathcal{G} \subset \mathcal{F}$ und jedes $f \in \mathcal{N}_i(\mathcal{F}, \mathbb{R}^m)$ existiert ein $g \in \mathcal{N}_i(\mathcal{G}, \mathbb{R}^m)$, so daß für alle $\mathcal{G}$-meßbaren $x : \Omega \to \mathbb{R}^m$*

$$\mathbb{E}[f(\cdot, x(\cdot)) \,|\mathcal{G}](\omega) = g(\omega, x(\omega)) \qquad \mathbb{P}\text{-f.ü.}$$

g heißt regulär bedingte Erwartung *von f bzgl. $\mathcal{G}$. Notation: $\mathbb{E}^r[f \,|\mathcal{G}] := g$.*

Satz 2.9. *1. Ist $f \in \mathcal{N}_{ci}\left(\mathcal{G}, \mathbb{R}^m \times \mathbb{R}^k\right)$ und $\varphi : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ definiert durch*

$$\varphi(\omega, x) := \inf_{y \in \mathbb{R}^k} f(\omega, x, y) \qquad \forall (\omega, x) \in \Omega \times \mathbb{R}^m,$$

so gilt $\varphi \in \mathcal{N}_{ci}(\mathcal{G}, \mathbb{R}^m)$

2. Ist $f \in \mathcal{N}_{ci}(\mathcal{F}, \mathbb{R}^m)$, so gilt $\mathbb{E}^r[f \,|\mathcal{G}] \in \mathcal{N}_{ci}(\mathcal{G}, \mathbb{R}^m)$.

Beweis. 1. Sei $f \in \mathcal{N}_{ci}\left(\mathcal{G}, \mathbb{R}^m \times \mathbb{R}^k\right)$. Dann existiert ein $\bar{f} \in \mathcal{N}_c\left(\mathcal{G}, \mathbb{R}^m \times \mathbb{R}^k\right)$ mit

$$\mathbb{P}\left(\left\{\omega \,\middle|\, f(\omega, x, y) = \bar{f}(\omega, x, y) \, \forall (x, y) \in \mathbb{R}^m \times \mathbb{R}^k\right\}\right) = 1$$

Sei

$$\bar{\varphi}(\omega, x) := \inf_{y \in \mathbb{R}^k} \bar{f}(\omega, x, y) \qquad \forall (\omega, x) \in \Omega \times \mathbb{R}^m.$$

Nach Satz 2.7 ist dann $\bar{\varphi} \in \mathcal{N}_c(\mathcal{G}, \mathbb{R}^m)$ und

$$\mathbb{P}(\{\omega \,|\varphi(\omega, x) = \bar{\varphi}(\omega, x) \ \forall x \in \mathbb{R}^m\}) = 1$$

Da $q : \Omega \to \bar{\mathbb{R}}$ existiert mit $\int_\Omega |q(\omega)| \, \mathbb{P}(d\omega) < \infty$ und $f(\omega, x, y) \geq q(\omega)$ $\mathbb{P}$-f.ü. für alle $(x, y) \in \mathbb{R}^m \times \mathbb{R}^k$, folgt $\bar{\varphi}(\omega, x) \geq q(\omega)$. Also ist $\varphi(\omega, x) \geq q(\omega)$ $\mathbb{P}$-f.ü. für alle $x \in \mathbb{R}^m$ und daher $\varphi \in \mathcal{N}_{ci}(\mathcal{G}, \mathbb{R}^m)$.

2. Nach Satz 2.8 gilt: $\mathbb{E}^r[f \,|\mathcal{G}] \in \mathcal{N}_i(\mathcal{G}, \mathbb{R}^m)$.

Zu zeigen: Es existiert ein $\bar{g} : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ mit $\mathbb{P}[\{\omega \,|\bar{g}(\omega, x) = \mathbb{E}^r[f \,|\mathcal{G}](\omega, x),\ \forall x \in \mathbb{R}^n\}] = 1$ und $\mathbb{P}[\{\omega \,|\{x \in \mathbb{R}^m \,|\bar{g}(\omega, x) \leq \alpha\}$ ist kompakt für alle $\alpha \in \mathbb{R}\}] = 1$.

Beweis: vgl. [Ev76, Theorem 5, 267-272] □

2.2 Dynamische Optimierung und determinierte äquivalente stochastische Programme

Wir betrachten das stochastische Optimierungsproblem

$$\min \left\{ \mathbb{E}[f(x_1,\ldots,x_T,\xi(\omega))] \;\middle|\; \begin{array}{c} x_t \in X_t \\ g_t(x_1,\ldots,x_t,\xi_t) \le 0 \quad \mathbb{P}-\text{f.ü.} \qquad t=1,\ldots,T \\ x_t \text{ ist } \mathcal{F}_t-\text{meßbar} \end{array} \right\} \tag{SP}$$

wobei $\xi = \{\xi_t\}_{t=1}^T$ ein stochastischer Prozeß über $(\Omega,\mathcal{F},\mathbb{P})$ ist, $\mathcal{F}_t := \sigma(\xi_1,\ldots,\xi_t)$, $t=1,\ldots,T$, $\mathcal{F}_1 = \{\emptyset,\Omega\}$, $\mathcal{F}_T = \mathcal{F}$, $X_t \subseteq \mathbb{R}^{m_t}$,

$$\begin{aligned} f &: \underbrace{\bigtimes_{t=1}^{T} \mathbb{R}^{m_t}}_{\mathbb{R}^m} \times \underbrace{\bigtimes_{t=1}^{T} \mathbb{R}^{s_t}}_{\mathbb{R}^s} \to \bar{\mathbb{R}}, \\ g_t &: \bigtimes_{\tau=1}^{t} \mathbb{R}^{m_\tau} \times \mathbb{R}^{s_t} \to \mathbb{R}^{d_t}. \end{aligned}$$

Wir definieren die Funktion $f_T : \Omega \times \mathbb{R}^m \to \bar{\mathbb{R}}$ durch

$$f_T(\omega,x_1,\ldots,x_T) := \begin{cases} f(x_1,\ldots,x_T,\xi(\omega)) & x_t \in X_t, g_t(x_1,\ldots,x_t,\xi_t) \le 0,\ t=1,\ldots,T \\ +\infty & \text{sonst} \end{cases}$$

für alle $(\omega,x_1,\ldots,x_T) \in \Omega \times \mathbb{R}^m = \Omega \times \bigtimes_{t=1}^T \mathbb{R}^{m_t}$.

Voraussetzung (A): f, X_t und g_t, $t=1,\ldots,T$, haben die Eigenschaft, daß $f_T \in \mathcal{N}_{ci}(\mathcal{F},\mathbb{R}^m)$.

(Kriterien dafür sind Beispiel 2.5 bzw. die darauffolgende Folgerung für die Zugehörigkeit zu $\mathcal{N}(\mathcal{F},\mathbb{R}^m)$)

Wir definieren rekursiv die Funktionen („*Bellmann-Rekursion*“)

$$\begin{aligned} \varphi_t(\omega,x_1(\omega),\ldots,x_t(\omega)) &:= \mathbb{E}^r[f_t(\cdot,x_1(\cdot),\ldots,x_t(\cdot))|\mathcal{F}_t](\omega) \quad \mathbb{P}\text{-f.ü., } t=T,\ldots,1 \quad (2.1) \\ f_{t-1}(\omega,x_1,\ldots,x_{t-1}) &:= \inf_{y\in\mathbb{R}^{m_t}} \varphi_t(\omega,x_1,\ldots,x_{t-1},y) \quad \mathbb{P}\text{-f.ü., } t=T,\ldots,2 \quad (2.2) \end{aligned}$$

Satz 2.10. *Es seien die folgenden Voraussetzungen erfüllt:* $f_T \in \mathcal{N}_{ci}(\mathcal{F},\mathbb{R}^m)$, *d.h.* (A) *ist erfüllt, und es existiere ein* $\tilde{x} : \Omega \to \mathbb{R}^m$ *mit* $\tilde{x}_t$ $\mathcal{F}_t$*-meßbar und* $\mathbb{E}[f_T(\omega,\tilde{x}_1(\omega),\ldots,\tilde{x}_T(\omega))] < \infty$.

Dann ist eine Abbildung $\bar{x} : \Omega \to \mathbb{R}^m$ *optimal für* (SP) *gdw. die Bellmann-Gleichungen*

$$\varphi_t(\omega,\bar{x}_1(\omega),\ldots,\bar{x}_t(\omega)) = f_{t-1}(\omega,\bar{x}_1(\omega),\ldots,\bar{x}_{t-1}(\omega)) \quad \mathbb{P}\text{-f.ü.}, t=1,\ldots,T,$$

erfüllt sind. Überdies existiert eine optimale Lösung von (SP).

Beweisschritte (vgl. [Ev76, Theorem 1 und 2]). Nach Satz 2.9 ist $\varphi_t \in \mathcal{N}_{ci}\left(\mathcal{F}_t, \bigtimes_{\tau=1}^t \mathbb{R}^{m_\tau}\right)$, $f_{t-1} \in \mathcal{N}_{ci}\left(\mathcal{F}_t, \bigtimes_{\tau=1}^{t-1} \mathbb{R}^{m_\tau}\right)$, $t=1,\ldots,T$.

($\Rightarrow$) (SP) ist äquivalent zu

$$\min\{\mathbb{E}[f_T(\omega,x_1(\omega),\ldots,x_T(\omega))]\,|x_t \text{ ist } \mathcal{F}_t\text{-meßbar}, t=1,\ldots,T\}$$

und dies besitzt nach Voraussetzung eine zulässige Lösung mit endlichem Zielfunktionswert.

Wegen $\varphi_t \in \mathcal{N}_{ci}$ werden die Infima angenommen und die Lösungen sind meßbar (Satz 2.6). Dann sind die Bellmann-Gleichungen erfüllbar.

($\Leftarrow$) Es sei $\bar{x} : \Omega \to \mathbb{R}^m$ sukzessive durch (2.2) so definiert, daß $\bar{x}_t$ $\mathcal{F}_t$-meßbar ist. Dies ist möglich nach Satz 2.6 und Auswahl einer meßbaren Selektion aus der Lösungsabbildung.

Zu zeigen:

$$\mathbb{E}[f_T(\omega, \bar{x}_1(\omega), \ldots, \bar{x}_T(\omega))] \leq \mathbb{E}[f_T(\omega, x_1(\omega), \ldots, x_T(\omega))]$$

für jede zulässige Lösung $x : \Omega \to \mathbb{R}^m$ von (SP).

Wir zeigen induktiv:

$$\mathbb{E}[f_t(\omega, \bar{x}_1(\omega), \ldots, \bar{x}_t(\omega))] \leq \mathbb{E}[f_t(\omega, x_1(\omega), \ldots, x_t(\omega))], \quad t = 0, \ldots, T$$

$t = 0$: trivial

$t - 1 \to t$: Es gelte

$$\mathbb{E}[f_{t-1}(\omega, \bar{x}_1(\omega), \ldots, \bar{x}_{t-1}(\omega))] \leq \mathbb{E}[f_{t-1}(\omega, x_1(\omega), \ldots, x_{t-1}(\omega))]$$

für alle zulässigen x. Es gilt

$$\begin{aligned}
\mathbb{E}[f_t(\omega, \bar{x}_1(\omega), \ldots, \bar{x}_t(\omega))] &\overset{(2.1)}{=} \mathbb{E}[\varphi_t(\omega, \bar{x}_1(\omega), \ldots, \bar{x}_t(\omega))] \\
&\overset{(2.2)}{=} \mathbb{E}[f_{t-1}(\omega, \bar{x}_1(\omega), \ldots, \bar{x}_{t-1}(\omega))] \\
&\overset{\text{Ind.vor.}}{\leq} \mathbb{E}[f_{t-1}(\omega, x_1(\omega), \ldots, x_{t-1}(\omega))] \\
&\overset{(2.2)}{\leq} \mathbb{E}[\varphi_t(\omega, x_1(\omega), \ldots, x_{t-1}(\omega), x_t(\omega))] \\
&\overset{(2.1)}{=} \mathbb{E}[f_t(\omega, x_1(\omega), \ldots, x_t(\omega))]
\end{aligned}$$

Existenz der optimalen Lösung: Wir betrachten rekursiv die Minimierungsprobleme

$$\inf_{y \in \mathbb{R}^{m_t}} \varphi_t(\omega, x_1(\omega), \ldots, x_{t-1}(\omega), y)$$

für beliebige $(\omega, x_1, \ldots, x_{t-1}) \in \Omega \times \bigtimes_{\tau=1}^{t-1} \mathbb{R}^{m_\tau}$ und $x_1, \ldots, x_{t-1}$ zulässig.

$\varphi_t : \Omega \times \bigtimes_{\tau=1}^{t-1} \mathbb{R}^{m_\tau} \times \mathbb{R}^{m_t} \to \bar{\mathbb{R}}$ ist $\mathcal{F}_t \otimes \mathcal{B}\left(\bigtimes_{\tau=1}^{t-1} \mathbb{R}^{m_\tau}\right)$-meßbar als Funktion von $y \in \mathbb{R}^{m_t}$.

Nach Satz 2.7 existiert eine meßbare Selektion $\eta_t(\omega, x_1, \ldots, x_{t-1})$ der Lösungsmengen-Abbildung des Minimierungsproblems, d.h. es gilt

$$\varphi_t(\omega, x_1, \ldots, x_{t-1}, \eta_t(\omega, x_1, \ldots, x_{t-1})) = f_{t-1}(\omega, x_1, \ldots, x_{t-1})$$

Induktives Argument: Seien die $\bar{x}_1, \ldots, \bar{x}_{t-1}$ bereits bestimmt und $\mathcal{F}_1$-$, \ldots, \mathcal{F}_{t-1}$-meßbar. Dann ist $\bar{x}_t := \eta_t(\cdot, \bar{x}_1(\cdot), \ldots, \bar{x}_{t-1}(\cdot)) : \Omega \to \mathbb{R}^{m_t}$ $\mathcal{F}_t$-meßbar.

Die so konstruierten $\bar{x}_1, \ldots, \bar{x}_T$ erfüllen die Bellmann-Gleichungen, also ist $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_T)$ eine optimale Lösung. □

Bemerkung 2.11 (determiniertes Äquivalent). Für $t = 1$ gilt die Rekursionsbeziehung des determinierten Optimierungsproblems

$$\inf_{x_1 \in \mathbb{R}^{m_1}} \mathbb{E}[f_1(\omega, x_1)] \tag{DA}$$

zur Bestimmung der optimalen Erststufen-Entscheidung. Satz 2.10 gibt Bedingungen, wann (DA) lösbar ist und zu einer Lösung $\bar{x}_1$ "ausgebaut" werden kann.

In diesem Sinne sind (SP) und (DA) äquivalent. (SP) ist i.a. unendlichdimensional und (DA) stets endlichdimensional. "Erkauft" wird dies mit der i.a. sehr komplizierten Struktur von f_1, selbst dann, wenn das Ausgangsproblem (SP) "einfach" ist.

Bemerkung 2.12 (lineare stochastische Programme). Für affin-lineare Funktionen $c_t(\cdot)$, $b_t(\cdot)$, $A_{t\tau}(\cdot)$, $\tau = 1, \ldots, t-1$, $t = 1, \ldots, T$, seien

$$\begin{aligned} f(x_1, \ldots, x_T, \xi) &:= \sum_{t=1}^{T} c_t(\xi_t) x_t, \\ g_t(x_1, \ldots, x_t; \xi_t) &:= A_{tt} x_t + \sum_{\tau=1}^{t-1} A_{t\tau}(\xi_t) x_\tau - b_t(\xi_t) \quad \text{Gleichheitsrestriktionen,} \end{aligned}$$

und seien die $X_t \subseteq \mathbb{R}^{m_t}$ konvex und polyedrisch. Dann sind die Voraussetzungen von Satz 2.10 erfüllt, falls die folgenden Bedingungen gültig sind:

(A1) "relativ vollständige Kompensation" ("relatively complete recourse")

Für jedes $t = 1, \ldots, T$ gilt $\mathbb{P}$-f.ü.

$$\mathcal{M}_t(x_1, \ldots, x_{t-1}, \xi_t) := \left\{ x_t \in X_t \,\middle|\, A_{tt} x_t = b_t(\xi_t) - \sum_{\tau=1}^{t-1} A_{t\tau}(\xi_t) x_\tau \right\} \neq \emptyset$$

für alle $x_\tau \in \mathcal{M}_\tau(x_1, \ldots, x_{\tau-1}, \xi_\tau(\omega))$, $\tau = 1, \ldots, t-1$.

(A2) "duale Zulässigkeit"

Für jedes $t = 1, \ldots, T$ ist $\mathbb{P}$-f.ü. das zu

$$\min \left\{ c_t(\xi_t(\omega)) x_t \,\middle|\, x_t \in X_t, A_{tt} x_t = b_t(\xi_t(\omega)) - \sum_{\tau=1}^{t-1} A_{t\tau}(\xi_t(\omega)) x_\tau \right\}$$

duale Optimierungsproblem zulässig $\forall x_\tau \in \mathcal{M}_\tau(x_1, \ldots, x_{\tau-1}, \xi_\tau(\omega))$, $\tau = 1, \ldots, t-1$.

(A3) $\min \{ c_1(\xi_1) x_1 \,|\, x_1 \in X_1, A_{11} x_1 = b_1(\xi_1) \}$ ist inf-kompakt.

(A4) $\mathbb{E}[\|\xi(\omega)\|] = \int_\Omega \|\xi(\omega)\| \, \mathbb{P}(d\omega) < \infty$

Bemerkung 2.13 (konvexe stochastische Programme). Es seien $f(\cdot, \xi)$ konvex, $g_t(\cdot, \xi_t)$ konvex, X_t konvex und abgeschlossen für alle $t = 1, \ldots, T$. Falls $f_t \in \mathcal{N}_{ci}(\mathcal{F}, \mathbb{R}^m)$, so sind sämtliche Funktionen φ_t und f_t, $t = 1, \ldots, T$, konvexe normale Integranden. Insbesondere ist (DA) ein konvexes Optimierungsproblem.

Beispiel 2.14 (zweistufige lineare stochastische Programme). Sei $T := 2$. Seien

$$\begin{aligned} f(x_1, x_2; \xi) &= c_1 x_1 + c_2(\xi) x_2 \\ g_2(x_1, x_2; \xi) &= A_{22} x_2 + A_{21}(\xi) x_1 - b_2(\xi) \end{aligned}$$

für affin-lineare Funktionen $c_2(\cdot)$, $b_2(\cdot)$ und $A_{21}(\cdot)$. Seien X_1, X_2 Polyeder und die affin lineare Restriktion $g_1(x_1; "\xi_1") = 0$ in X_1 integriert. Dann hat ein zweistufiges lineares stochastisches Programm die Form

$$\min \{ c_1 x_1 + \mathbb{E}[c_2(\xi) x_2] \,|\, x_1 \in X_1, x_2 \in X_2; A_{22} x_2 = b_2(\xi) - A_{21}(\xi) x_1 \quad \mathbb{P}\text{-f.ü.} \}$$

Es seien die Bedingungen (A1)-(A4) aus Bemerkung 2.12 erfüllt.

Die Bellmann-Gleichungen sind

$$\begin{aligned} f_2(\omega, x_1, x_2) &:= f(x, \xi(\omega)) = c_1 x_1 + c_2(\xi(\omega)) x_2 \quad \rightsquigarrow \quad f_2 \text{ ist normaler Integrand} \\ f_1(\omega, x_1) &:= \inf \{ c_1 x_1 + c_2(\xi(\omega)) x_2 \,|\, x_2 \in X_2, A_{22} x_2 = b_2(\xi(\omega)) - A_{21}(\xi(\omega)) x_1 \} \\ &= c_1 x_1 + \inf \{ c_2(\xi(\omega)) x_2 \,|\, x_2 \in X_2, A_{22} x_2 = b_2(\xi(\omega)) - A_{21}(\xi(\omega)) x_1 \} \\ &= c_1 x_1 + \Phi(c_2(\xi(\omega)), b_2(\xi(\omega)) - A_{21}(\xi(\omega)) x_1) \\ \text{wobei } \Phi(u, t) &:= \inf \{ uy \,|\, y \in X_2, A_{22} y = t \} \end{aligned}$$

Die Menge

$$\mathrm{dom}\Phi = \{(u,t) \,|{-\infty} < \Phi(u,t) < \infty\}$$

ist ein Polyeder und es existieren endlich viele polyedrische Mengen D_j, $j = 1, \ldots, l$, so daß

$$\bigcup_{j=1}^{\infty} D_j = \mathrm{dom}\Phi,$$

$D_i \cap D_j$ ist gemeinsame Seite von D_i und D_j oder leer ($D_i \cap D_j$ hat Lebesgue-Maß 0), $i \neq j$, und

$$\Phi(u,t) = \langle C_j u, t\rangle \qquad \forall (u,t) \in D_j, j = 1, \ldots, l,$$

wobei C_j eine Matrix geeigneter Dimension ist, d.h. Φ ist eine stückweise quadratische Funktion von (u,t). Überdies ist Φ auf domΦ stetig und die Funktion $\Phi(\cdot,t)$ ist stückweise linear konkav ($\forall t$) und $\Phi(u,\cdot)$ ist stückweise linear konvex ($\forall u$). (Hauptresultat der sogenannten linearen parametrischen Optimierung; [WW69, RW98])

Konsequenz: $f_1(\omega,\cdot)$ ist stückweise linear konvex.

Deterministisches Äquivalent zum zweistufigen stochastischen Programm:

$$\min\{\mathbb{E}[f_1(\omega,x_1)]|\, x_1 \in X_1\} = \min\left\{\int_\Omega f_1(\omega,x_1)\,\mathbb{P}(d\omega)\middle|\, x_1 \in X_1\right\}$$

2.3 Dualität

Wir betrachten das stochastische Optimierungsproblem

$$\min\left\{\mathbb{E}[f(x_1,\ldots,x_T;\xi)]\;\middle|\;\begin{array}{lll} x_t \in X_t & & \\ h_t(x_t,\xi_t) \le 0 & \mathbb{P}\text{-f.ü.} & \\ g_t(x_1,\ldots,x_t;\xi_t) \le 0 & \mathbb{P}\text{-f.ü.} & t = 1,\ldots,T \\ x_t \quad \mathcal{F}_t\text{-meßbar} & & \end{array}\right\} \tag{SP}$$

für $f : \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}$, $h_t : \mathbb{R}^{m_t} \times \mathbb{R}^{s_t} \to \mathbb{R}^{k_t}$, $g_t : \left(\bigtimes_{\tau=1}^t \mathbb{R}^{m_\tau}\right) \times \mathbb{R}^{s_t} \to \mathbb{R}^{d_t}$, wobei durch die explizite Formulierung der Restriktionen $h_t(x_t,\xi_t) \le 0$ Kopplungen der Komponenten von x_t ausgedrückt werden sollen, die vorher in die dynamischen Restriktionen integriert waren. Wir nehmen der Einfachheit halber an, daß

$$\xi \in \bigtimes_{t=1}^T L_\infty(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{s_t})$$

(d.h. ξ ist wesentlich beschränkt) und wir suchen auch die Entscheidung

$$x \in X := \bigtimes_{t=1}^T L_\infty(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{m_t}).$$

Wir formulieren (SP) als Optimierungsproblem im Banachraum X:

$$\min\left\{F(x)\;\middle|\;\begin{array}{lll} x \in X, \quad x_t \in X_t & & \\ h_t(x_t,\xi_t) \le 0 & \mathbb{P}\text{-f.ü.} & \\ g_t(x_1,\ldots,x_t;\xi_t \le 0) & \mathbb{P}\text{-f.ü.} & t = 1,\ldots,T \\ H_t(x_t) = 0 & & \end{array}\right\}$$

mit

$$\begin{aligned} F(x) &:= \mathbb{E}[f(x_1,\ldots,x_T;\xi)] \\ H_t(x_t) &:= x_t - \mathbb{E}[x_t\,|\mathcal{F}_t], \quad t = 1,\ldots,T. \end{aligned}$$

Voraussetzung (V): $\xi \in \bigtimes_{t=1}^{T} L_\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{s_t})$ und X_t kompakt, $t = 1, \ldots, T$.

Wir definieren folgende Mengen von *Lagrange-Multiplikatoren*:

$$\begin{aligned}
\Lambda &:= \Lambda_1 \times \Lambda_2 \times \Lambda_3 \\
\Lambda_1 &:= \bigtimes_{t=1}^{T} L_1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{m_t}) \\
\Lambda_2 &:= \left\{ \lambda_2 \in \bigtimes_{t=1}^{T} L_1\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{k_t}\right) | \lambda_2(\omega) \geq 0 \quad \mathbb{P}\text{-f.ü.} \right\} \\
\Lambda_3 &:= \left\{ \lambda_3 \in \bigtimes_{t=1}^{T} L_1\left(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{d_t}\right) | \lambda_3(\omega) \geq 0 \quad \mathbb{P}\text{-f.ü.} \right\}
\end{aligned}$$

Beachte: $\{x \in L_p(\Omega, \mathcal{F}, \mathbb{P}) | x(\omega) \geq 0 \quad \mathbb{P} - \text{f.ü.}\}$ hat leeres Inneres für $p \in [1, \infty)$!

Die *Lagrange-Funktion* $L : X \times \Lambda \to \mathbb{R}$,

$$L(x, \lambda) := F(x) + \mathbb{E}\left[\sum_{t=1}^{T} \lambda_{1t} \cdot H_t(x_t) + \sum_{t=1}^{T} \lambda_{2t} \cdot h_t(x_t; \xi_t) + \sum_{t=1}^{T} \lambda_{3t} \cdot g_t(x_1, \ldots, x_t; \xi_t) \right],$$

ist für alle $(x, \lambda) \in X \times \Lambda$ wohl-definiert. Die *duale Funktion* $D : \Lambda \to \bar{\mathbb{R}}$ ist definiert mittels

$$D(\lambda) := \inf \{ L(x, \lambda) | x \in X,\ x_t \in X_t \quad \mathbb{P}\text{-f.ü.},\ t = 1, \ldots, T \}, \quad \lambda \in \Lambda,$$

und das *duale Problem* ist

$$\max \{ D(\lambda) | \lambda \in \Lambda \} \tag{D}$$

Unter der Voraussetzung (V) ist das duale Problem wohldefiniert.

Lemma 2.15. *Es sei* (V) *erfüllt. Dann ist die duale Funktion auf dem konvexen Kegel* Λ *konkav und es gilt*

$$\sup_{\lambda \in \Lambda} D(\lambda) \leq \inf_{x \in X, x \text{ zulässig}} F(x) \qquad \textit{(schwache Dualität)}$$

Beweis. Für alle $x \in X$ ist $L(x, \cdot) : \Lambda \to \mathbb{R}$ affin linear und Λ ist als kartesisches Produkt konvexer Kegel ebenfalls ein konvexer Kegel. Es seien $\lambda, \tilde{\lambda} \in \Lambda$ und $\alpha \in [0, 1]$ gegeben. Dann ist

$$\begin{aligned}
D\left(\alpha\lambda + (1 - \alpha)\tilde{\lambda}\right) &= \inf \left\{ L\left(x, \alpha\lambda + (1 - \alpha)\tilde{\lambda}\right) \middle| x \in X,\ x_t \in X_t \quad \mathbb{P}\text{-f.ü.},\ t = 1, \ldots, T \right\} \\
&= \inf \left\{ \alpha L(x, \lambda) + (1 - \alpha) L\left(x, \tilde{\lambda}\right) \middle| x \in X,\ x_t \in X_t \quad \mathbb{P}\text{-f.ü.},\ t \in 1, \ldots, T \right\} \\
&\geq \alpha \inf \{ L(x, \lambda) | x \in X,\ x_t \in X_t \quad \mathbb{P}\text{-f.ü.},\ t \in 1, \ldots, T \} \\
&\quad + (1 - \alpha) \inf \left\{ L\left(x, \tilde{\lambda}\right) \middle| x \in X,\ x_t \in X_t \quad \mathbb{P}\text{-f.ü.},\ t \in 1, \ldots, T \right\} \\
&= \alpha D(\lambda) + (1 - \alpha) D\left(\tilde{\lambda}\right)
\end{aligned}$$

Gilt $\inf F(x) = +\infty$, so ist die schwache Dualität trivial. Im Fall $\inf F(x) < +\infty$ gilt für jedes zulässige $\bar{x}$ mit $F(\bar{x}) < +\infty$ und alle $\lambda \in \Lambda$

$$D(\lambda) = \inf \{ L(x, \lambda) | x \in X,\ x_t \in X_t \quad \mathbb{P}\text{-f.ü.},\ t \in 1, \ldots, T \} \leq L(\bar{x}, \lambda) \leq F(\bar{x}).$$

Maximierung über $\lambda \in \Lambda$ und Minimierung über zulässige $x \in X$ liefert die schwache Dualität. □

Im allgemeinen heißt

$$\inf_{x \in X, x \text{ zulässig}} F(x) - \sup_{\lambda \in \Lambda} D(\lambda)$$

Dualitätslücke von (SP). Die Größe der Dualitätslücke ist häufig ein Gradmesser für die Kompliziertheit von (SP).

Satz 2.16 (Rockafellar / Wets '78 [SI78, 16-36]). *Es sei die Voraussetzung* (V) *erfüllt. Die Mengen X_t und Funktionen $f(\cdot;\xi)$, $h_t(\cdot;\xi_t)$, $g_t(\cdot,\ldots,\cdot;\xi_t)$, $t=1,\ldots,T$ seien konvex. Das Programm* (SP) *besitze relativ vollständige Kompensation im Sinne von* (A1) *in Bemerkung 2.12 und es sei strikt zulässig, d.h. es existieren ein $\varepsilon > 0$ und ein $\bar{x} \in X$, so daß $\bar{x} = \mathbb{E}[\bar{x}_t \,|\mathcal{F}_t]$, $t = 1,\ldots,T$, $\mathbb{B}(\bar{x}_t,\varepsilon) \subseteq X_t$, $h_t(\bar{x}_t,\xi_t) \le -\varepsilon$, $g_t(\bar{x}_1,\ldots,\bar{x}_t;\xi_t) \le -\varepsilon$ $\mathbb{P}$-f.ü., $t=1,\ldots,T$.*

Dann existieren optimale Lösungen $\lambda^* \in \Lambda$ von (D) *und $x^* \in X$ von* (SP)*, so daß*

$$D(\lambda^*) = F(x^*).$$

Insbesondere gilt starke Dualität*:*

$$\max_{\lambda\in\Lambda} D(\lambda) = \min_{x\in X, x \text{ zulässig}} F(x).$$

Bemerkung 2.17 (partielle Dualisierung). Eine Möglichkeit ist die Dualisierung nur einer Gruppe von Restriktionen (Nichtantizipativitätsbedingungen (NA-Bedingungen), verkoppelnde Restriktionen, dynamische Restriktionen). Lemma 2.15 und Satz 2.16 bleiben dabei sinngemäß erhalten. Wir führen dazu folgende Lagrange-Funktionen

$$\begin{aligned}
&L_1 : X\times\Lambda_1 \to \mathbb{R}, \quad L_1(x,\lambda_1) := F(x) + \mathbb{E}\left[\sum_{t=1}^T \lambda_{1t} H_t(x_t)\right] && \forall (x,\lambda_1) \in X\times\Lambda_1,\\
&L_2 : X\times\Lambda_2 \to \mathbb{R}, \quad L_2(x,\lambda_2) := F(x) + \mathbb{E}\left[\sum_{t=1}^T \lambda_{2t} h_t(x_t;\xi_t)\right] && \forall (x,\lambda_2) \in X\times\Lambda_2,\\
&L_3 : X\times\Lambda_3 \to \mathbb{R}, \quad L_3(x,\lambda_3) := F(x) + \mathbb{E}\left[\sum_{t=1}^T \lambda_{3t} g_t(x_1,\ldots,x_t;\xi_t)\right] && \forall (x,\lambda_3) \in X\times\Lambda_3,
\end{aligned}$$

und zugehörige duale Funktionen $D_i : \Lambda_i \to \mathbb{R}$ und duale Probleme (D_i) ein, $i = 1,2,3$:

$$D_i(\lambda_i) := \inf\{L_i(x_i,\lambda_i)\,|x \in X,\ x \text{ eindeutig für verbleibende Restriktionen}\}, \quad \forall \lambda_i \in \Lambda_i,$$

$$\max\{D_i(\lambda_i)\,|\lambda_i \in \Lambda_i\} \qquad (D_i)$$

Motivation: Man sollte möglichst wenige Restriktionen dualisieren, da i.a. mit wachsender Anzahl von dualisierten Restriktionen die Dualitätslücke wächst (vgl. [LR01]). Einen Vergleich von Dualitätslücken für die 3 Dualitätsansätze findet man in [DR02], z.B.

$$\sup\{D_3(\lambda_3)\,|\lambda_3 \in \Lambda_3\} \le \sup\{D_1(\lambda_1)\,|\lambda_1\in\Lambda_1\} \quad \text{(i.a. strikt).}$$

Algorithmische Effekte der partiellen Dualisierung: Zur Vereinfachung nehmen wir an

$$F(x) = \mathbb{E}\left[\sum_{t=1}^T f_t(x_t;\xi_t)\right] \qquad \text{(Separabilität).}$$

Szenario-Dekomposition D_1: Dekomposition des stochastischen Programms in "realisierungsweise" Optimierungsprobleme: Es gilt

$$\begin{aligned}
\mathbb{E}[\lambda_{1t}H_t(x_t)] = \mathbb{E}[\lambda_{1t}x_t] - \mathbb{E}[\lambda_{1t}\mathbb{E}[x_t\,|\mathcal{F}_t]] &= \mathbb{E}[\lambda_{1t}x_t] - \mathbb{E}[\mathbb{E}[\lambda_{1t}\mathbb{E}[x_t\,|\mathcal{F}_t]]]\\
&= \mathbb{E}[\lambda_{1t}x_t] - \mathbb{E}[\mathbb{E}[\lambda_{1t}\,|\mathcal{F}_t]\,\mathbb{E}[x_t\,|\mathcal{F}_t]]\\
&= \mathbb{E}[\lambda_{1t}x_t] - \mathbb{E}[\mathbb{E}[\mathbb{E}[\lambda_{1t}\,|\mathcal{F}_t]\,x_t\,|\mathcal{F}_t]]\\
&= \mathbb{E}[\lambda_{1t}x_t] - \mathbb{E}[x_t\mathbb{E}[\lambda_{1t}\,|\mathcal{F}_t]] = \mathbb{E}[x_tH_t(\lambda_{1t})],
\end{aligned}$$

so daß $L_1(x,\lambda_1) = \mathbb{E}\left[\sum_{t=1}^T \{f_t(x_t,\xi_t) + \lambda_{1t}H_t(x_t)\}\right] = \mathbb{E}\left[\sum_{t=1}^T \{f_t(x_t,\xi_t) + H_t(\lambda_{1t})\,x_t\}\right].$

Dann ist (mit [RW98, Theorem 14.60])

$$\begin{aligned} D_1(\lambda_1) &= \inf\left\{ \mathbb{E}\left[\sum_{t=1}^{T} f_t(x_t;\xi_t) + H_t(\lambda_{1t})\, x_t \;\middle|\; \begin{array}{rl} h_t(x_t;\xi_t) \le 0 & \mathbb{P}\text{-f.ü.}, \\ g_t(x_1,\dots,x_t;\xi_t) \le 0 & \mathbb{P}\text{-f.ü.}, \\ x_t \in X_t & t=1,\dots,T \end{array} \right] \right\} \\ &= \mathbb{E}\left[\inf\left\{ \sum_{t=1}^{T} \left[f_t(x_t;\xi_t) + H_t(\lambda_{1t})\, x_t \right] \;\middle|\; \begin{array}{rl} h_t(x_t;\xi_t) \le 0 & \mathbb{P}\text{-f.ü.}, \\ g_t(x_1,\dots,x_t;\xi_t) \le 0 & \mathbb{P}\text{-f.ü.}, \\ x_t \in X_t & t=1,\dots,T \end{array} \right\} \right] \end{aligned}$$

geographische Dekomposition D_2 Dekomposition des stochastischen Programms in stochastische Programme für jede Komponente des Entscheidungsvektor (falls die Restriktionen "$g_t(x_1,\dots,x_t;\xi_t) \le 0$" und "$x_t \in X_t$", $t = 1,\dots,T$, ausschließlich komponentenweise gegeben sind).

nodale Dekomposition D_3 Dekomposition des stochastischen Programms in stochastische Programme für x_t, $t = 1,\dots,T$.

2.4 Stabilität

Motivation: Die Wahrscheinlichkeitsverteilung $\mathbb{P}$ von $\xi = \{\xi_t\}_{t=1}^{T}$ ist in praktischen Anwendungen meist *nicht* oder *nicht vollständig* bekannt. Falls $\mathbb{P}$ bekannt ist, so erfordert die Berechnung der Zielfunktion $\mathbb{E}[f(x;\xi)]$ die Auswertung mehrdimensionaler Integrale (Monte-Carlo Methoden, Approximation von $\mathbb{P}$ durch einfachere Verteilungen). Deshalb ist eine *Störungstheorie* stochastischer Programme bzgl. der Verteilung P von ξ notwendig.

Ist es möglich, allgemeine Resultate aus der Störungstheorien für allgemeine Optimierungsprobleme anzuwenden?

- Bank / Guddat / Klatte / Kummer / Tammer: Nonlinear parametric optimization, 1982
- Rockafellar / Wets 1998 ([RW98])
- Robinson 70er/80er
- Bonnans / Shapiro, Springer 2000

Ausgangspunkt: Die unendlich-dimensionale Formulierung von (SP) führt zu methodischen Problemen. Dies führt zur Verwendung der determinierten Äquivalente. Dafür wird eine Störungstheorie benötigt, die ohne starke analytische Voraussetzungen auskommt.

Wir betrachten stochastische Optimierungsprobleme der Form

$$\min\left\{ \int_{\Omega} f_0(x,\xi(\omega))\, \mathbb{P}(d\omega) \;\middle|\; x \in X \right\},$$

wobei $f_0 : \mathbb{R}^m \times \mathbb{R}^s \to \bar{\mathbb{R}}$ ein normaler Integrand ist, $X \subseteq \mathbb{R}^m$ abgeschlossen, $\xi : \Omega \to \mathbb{R}^s$ ein Zufallsvektor über $(\Omega, \mathcal{F}, \mathbb{P})$. Dies ist äquivalent zu

$$\min\left\{ \int_{\Xi} f_0(x,\xi)\, P(d\xi) \;\middle|\; x \in X \right\},$$

wobei $\Xi \subseteq \mathbb{R}^s$ abgeschlossen ist und P eine Wahrscheinlichkeitsverteilung von ξ definiert auf Ξ $\left(P := \mathbb{P} \circ \xi^{-1}\right)$. Wir betrachten

$$\begin{aligned} v(P) &:= \inf\left\{ \int_{\Xi} f_0(x,\xi)\, P(d\xi) \;\middle|\; x \in X \right\} && \text{(Optimalwert)} \\ S(P) &:= \left\{ x \in X \;\middle|\; \int_{\Xi} f_0(x,\xi)\, P(d\xi) = v(P) \right\} && \text{(Lösungsmenge)} \end{aligned}$$

Mathematisches Problem: Stetigkeitseigenschaften der Abbildung $v(\cdot)$ und $S(\cdot)$ in der "Original"-Verteilung P, wobei im Raum der Wahrscheinlichkeitsmaße eine geeignete Topologie zu definieren ist.

Es sei $U \subseteq \mathbb{R}^m$ eine beliebige nichtleere, offene, beschränkte Menge. Wir betrachten

$$\mathcal{F}_U := \{f_0(x,\cdot)\,|x \in X \cap c\ell U\} \qquad \big(\text{Menge von meßbaren Funktionen von } \Xi \text{ in } \bar{\mathbb{R}}\big)$$

$$\mathcal{P}_{\mathcal{F}_U}(\Xi) := \left\{ Q \text{ Wahrscheinlichkeitsmaß auf } \Xi \;\middle|\; \begin{array}{l} \int_\Xi \inf_{x \in X \cap \mathbb{B}(0,r)} f_0(x,\xi)\,Q(d\xi) > -\infty \quad r > 0 \\ \sup_{x \in X \cap c\ell U} \int_\Xi f_0(x,\xi)\,Q(d\xi) < +\infty \end{array} \right\}$$

$$\begin{aligned} d_{\mathcal{F}_U}(P,Q) &:= \sup_{x \in X \cap c\ell U} \left| \int_\Xi f_0(x,\xi)\,P(d\xi) - \int_\Xi f_0(x,\xi)\,Q(d\xi) \right| \\ &= \sup_{x \in X \cap c\ell U} \left| \int_\Xi f_0(x,\xi)\,(P-Q)(d\xi) \right| \qquad \forall P, Q \in \mathcal{P}_{\mathcal{F}_U}(\Xi) \end{aligned}$$

$d_{\mathcal{F}_U} : \mathcal{P}_{\mathcal{F}_U} \times \mathcal{P}_{\mathcal{F}_U} \to \mathbb{R}_+$ ist symmetrisch und erfüllt die Dreiecksungleichung. $d_{\mathcal{F}_U}$ ist eine *Pseudo-Metrik* ([Du89, Kap. 2.1]), denn i.a. gilt <u>nicht</u>: $d_{\mathcal{F}_U}(P,Q) = 0 \Rightarrow P = Q$ (Ursache: $\mathcal{F}_U$ ist nicht reichhaltig genug).

$\mathcal{P}_{\mathcal{F}_U}$ wird ein topologischer Raum mit der Basis, die durch alle Kugeln gebildet wird. Er ist i.a. nicht Hausdorff (er ist genau dann Hausdorff, wenn $d_{\mathcal{F}_U}$ eine Metrik ist).

Eine Folge $\{P_n\}_{n\in\mathbb{N}}$ konvergiert gegen P in $\mathcal{P}_{\mathcal{F}_U} \Leftrightarrow d_{\mathcal{F}_U}(P_n, P) \xrightarrow{n\to\infty} 0$.

Lemma 2.18. *Unter den obigen Voraussetzungen gilt: Die Abbildung*

$$(x,Q) \mapsto \int_\Xi f_0(x,\xi)\,Q(d\xi)$$

von $\mathbb{R}^m \times \mathcal{P}_{\mathcal{F}_U}$ *in* $\bar{\mathbb{R}}$ *ist auf* $(X \cap c\ell U) \times \mathcal{P}_{\mathcal{F}_U}$ *sequentiell unterhalbstetig.*

Beweis. Zu zeigen: Für $x_n \to x$ in $X \cap c\ell U$ und $Q_n \to Q$ in $\mathcal{P}_{\mathcal{F}_U}$ gilt

$$\int_\Xi f_0(x,\xi)\,Q(d\xi) \le \liminf_{n\to\infty} \int_\Xi f_0(x_n,\xi)\,Q_n(d\xi).$$

Es ist

$$\begin{aligned} \int_\Xi f_0(x,\xi)\,Q(d\xi) &\le \liminf_{n\to\infty} \int_\Xi f_0(x_n,\xi)\,Q(d\xi) \qquad \text{(Fatou-Lemma)} \\ &= \liminf_{n\to\infty} \left(\int_\Xi f_0(x_n,\xi)\,Q_n(d\xi) + \int_\Xi f_0(x_n,\xi)\,(Q-Q_n)(d\xi) \right) \\ &\le \liminf_{n\to\infty} \left(\int_\Xi f_0(x_n,\xi)\,Q_n(d\xi) + d_{\mathcal{F}_U}(Q,Q_n) \right) \\ &\le \liminf_{n\to\infty} \left(\int_\Xi f_0(x_n,\xi)\,Q_n(d\xi) \right) \end{aligned}$$

□

Um Lösungsmengen nichtkonvexer Optimierungsprobleme behandeln zu können, werden zunächst "Lokalisierungen" von v und S benötigt.

$$\begin{aligned} v_U(P) &:= \inf \left\{ \int_\Xi f_0(x,\xi)\,P(d\xi) \;\middle|\; x \in X \cap c\ell U \right\} \\ S_U(P) &:= \left\{ x \in X \cap c\ell U \;\middle|\; \int_\Xi f_0(x,\xi)\,P(d\xi) = v_U(P) \right\} \end{aligned}$$

Definition (vollständige Menge lokaler Minimalpunkte). Eine nichtleere Menge $S \subset \mathbb{R}^m$ heißt *vollständige Menge lokaler Minimalpunkte* (CLM Menge) von (SP), falls eine offene, beschränkte Menge $U \subset \mathbb{R}^m$ existiert, so daß $S = S_U(P) \subset U$.

Es gilt: Falls (SP) beschränkt ist, so ist $S(P)$ eine CLM Menge. CLM Mengen sind nach Definition in der Tat Mengen lokaler Minimalpunkte.

Satz 2.19 (Stabilität). *Es sei $P \in \mathcal{P}_{\mathcal{F}_U}(\Xi)$, $S(P)$ sei nichtleer und $U \subset \mathbb{R}^m$ sei eine offene, beschränkte Menge mit $S(P) \subset U$. Dann gilt für alle $Q \in \mathcal{P}_{\mathcal{F}_U}(\Xi)$:*

$$|v(P) - v_U(Q)| \leq d_{\mathcal{F}_U}(P, Q)$$

und

$$\begin{aligned} \emptyset \neq S_U(Q) &\subseteq S(P) + \mathbb{B}\left(0, \psi_P^{-1}(2d_{\mathcal{F}_U}(P,Q))\right) \\ \textit{bzw.} \quad \sup_{x \in S_U(Q)} & d(x, S(P)) \leq \psi_P^{-1}(2d_{\mathcal{F}_U}(P,Q)), \end{aligned}$$

wobei die Wachstumsfunktion $\psi_P : \mathbb{R}_+ \to \mathbb{R}_+$ von (SP) *definiert ist durch*

$$\begin{aligned} \psi_P(\tau) &:= \min\left\{ \int_\Xi f_0(x,\xi)\, P(d\xi) - v(P) \,\middle|\, d(x, S(P)) \geq \tau,\ x \in X \cap c\ell U \right\} \quad \forall \tau \in \mathbb{R}_+, \\ \psi_P^{-1}(r) &:= \sup\{\tau \in \mathbb{R}_+ \,|\, \psi_P(\tau) \leq r\} \quad \forall r \in \mathbb{R}_+. \end{aligned}$$

(ψ_P ist nichtfallend und unterhalbstetig. Dann ist ψ_P^{-1} nichtfallend und $\lim_{r \searrow 0} \psi_P^{-1}(r) \geq 0 = \psi_P^{-1}(0)$.)

Die Menge $S_U(Q)$ ist eine CLM Menge, falls $d_{\mathcal{F}_U}(P,Q)$ hinreichend klein ist.

Beweis. Nach Lemma 2.18 ist die Abbildung $x \mapsto \int_\Xi f_0(x,\xi)\, Q(d\xi)$ unterhalbstetig auf $X \cap c\ell U$ für jedes $Q \in \mathcal{P}_{\mathcal{F}_U}(\Xi)$. Da $X \cap c\ell U$ kompakt ist, gilt $S_U(Q) \neq \emptyset$.

Sei nun $Q \in \mathcal{P}_{\mathcal{F}_U}(\Xi)$ beliebig.

1. Sei $\bar{x} \in S(P)$ und $\tilde{x} \in S_U(Q)$. Dann ist

$$\begin{aligned} v(P) - v_U(Q) &\leq \int_\Xi f_0(\tilde{x},\xi)\, P(d\xi) - \int_\Xi f_0(\tilde{x},\xi)\, Q(d\xi) \leq d_{\mathcal{F}_U}(P,Q) \\ v_U(Q) - v(P) &\leq \int_\Xi f_0(\bar{x},\xi)\, Q(d\xi) - \int_\Xi f_0(\bar{x},\xi)\, P(d\xi) \leq d_{\mathcal{F}_U}(P,Q). \end{aligned}$$

2. Nach Definition der Wachstumsfunktion $\psi_P(\cdot)$ gilt:

$$\int_\Xi f_0(x,\xi)\, P(d\xi) - v(P) \geq \psi_P(d(x, S(P))) \qquad \forall x \in X \cap c\ell U.$$

Sei $\tilde{x} \in S_U(Q)$. Dann ist

$$\begin{aligned} 2d_{\mathcal{F}_U}(P,Q) &\geq \int_\Xi f_0(\tilde{x},\xi)\, P(d\xi) - \underbrace{\int_\Xi f_0(\tilde{x},\xi)\, Q(d\xi)}_{=v_U(Q)} + v_U(Q) - v(P) \\ &= \int_\Xi f_0(\tilde{x},\xi)\, P(d\xi) - v(P) \\ &\geq \psi_P(d(\tilde{x}, S(P))). \end{aligned}$$

Also folgt $d(\tilde{x}, S(P)) \leq \psi_P^{-1}(2d_{\mathcal{F}_U}(P,Q))$.

3. Noch zu zeigen: $S_U(Q)$ ist CLM Menge bzgl. U, falls $d_{\mathcal{F}_U}(P,Q)$ "klein" ist.
   Angenommen, es existiert eine Folge $\{P_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}_{\mathcal{F}_U}(\Xi)$ mit

$$d_{\mathcal{F}_U}(P,P_n) \xrightarrow{n\to\infty} 0 \quad \text{und} \quad S_U(P_n) \not\subset U \;\forall n\in\mathbb{N}.$$

   Dann existiert für alle $n\in\mathbb{N}$ ein

$$x_n \in S_U(P_n)\setminus U \subseteq X\cap c\ell U\setminus U.$$

   Aufgrund der Kompaktheit von $X\cap c\ell U\setminus U$ existiert dann eine in $X\cap c\ell U\setminus U$ konvergente Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$, $x_{n_k} \xrightarrow{k\to\infty} x\in X\cap c\ell U\setminus U$, und mit Lemma 2.18 folgt

$$\int_\Xi f_0(x,\xi)\,P(d\xi) \le \liminf_{k\to\infty}\int_\Xi f_0(x_{n_k},\xi)\,P_{n_k}(d\xi) = \liminf_{k\to\infty} v_U(P_{n_k}) = v(P),$$

   also ist $x\in S(P)\subset U$. Widerspruch zu $x\in X\cap c\ell U\setminus U$. □

Bemerkung 2.20 (Wachstum). Falls die Wachstumsfunktion $\psi_P(\cdot)$ die Eigenschaft $\psi_P(\tau)\ge \gamma\cdot\tau^k$, für alle kleinen $\tau\in\mathbb{R}_+$ und gewisse $\gamma>0$, $k\ge 1$, hat, so sagt man, daß (SP) ein *Wachstum der Ordnung* k besitzt ($k=1$ und $k=2$: lineares bzw. quadratisches Wachstum), d.h.

$$\int_\Xi f_0(x,\xi)\,P(d\xi) \ge v(P)+\gamma d(x,S(P))^k \qquad \forall x\in X\cap c\ell U,\quad d(x,S(P)) \text{ klein}$$

In diesem Fall gilt: $\psi_P^{-1}(r)\le\left(\frac{r}{\gamma}\right)^{\frac{1}{k}}$ für "kleines" $r\in\mathbb{R}_+$.

Bemerkung 2.21 (Globalisierung). Die lokalisierten Konzepte $v_U(\cdot)$ und $S_U(\cdot)$ in Satz 2.19 können in den folgenden beiden Fällen durch $v(\cdot)$ bzw. $S(\cdot)$ ersetzt werden:

1. Die Probleme

$$\min\left\{\int_\Xi f_0(x,\xi)\,Q(d\xi)\,\middle|\, x\in X\right\}$$

   sind konvex und $S_U(Q)$ ist CLM Menge. Dann ist $S_U(Q)=S(Q)\subset U$ und $v_U(Q)=v(Q)$.
2. X ist kompakt (wir wählen U so, daß $X\subseteq U$).

Bemerkung 2.22 (Vergrößerung von $\mathcal{F}_U$). Wählt man eine Klasse $\mathcal{F}$ von meßbaren Funktionen von Ξ in $\bar{\mathbb{R}}$ und $\mathcal{F}_U\subseteq\mathcal{F}$, so vergrößert sich $d_{\mathcal{F}_U}$ zu $d_{\mathcal{F}}$ und die Aussage von Satz 2.19 bleibt richtig (dabei "verkleinert" sich auch $\mathcal{P}_{\mathcal{F}_U}$).

Dies ist von Bedeutung, um die schwer handhabbaren Pseudo-Metriken für die Anwendung nutzbar zu machen, insbesondere für deren Berechnung bzw. Abschätzung und für die Charakterisierung der Konvergenz bzgl. $d_{\mathcal{F}_U}$.

Genauer: Man sucht eine Klasse $\mathcal{F}$, so daß eine Normalisierungskonstante $C=C(\mathcal{F})>0$ existiert, so daß $Cf_0(x,\cdot)\in\mathcal{F}$, $\forall x\in X\cap c\ell(U)$, und daß $\mathcal{F}$ "nicht zu groß" ist.

Beispiel (zweistufige lineare stochastische Programme, vgl. Bsp. 2.14). f_0 hat die Form

$$f_0(x,\xi)=\langle c_1,x\rangle+\Phi(c_2(\xi),A_{21}(\xi)x-b_2(\xi)) \qquad \forall(x,\xi)\in X_1\times\Xi$$

wobei

$$\Phi(u,t):=\inf\{\langle u,y\rangle|\,y\in X_2, A_{22}y=t\}$$

für Polyeder X_1, X_2 und affin lineare Funktionen $c_2(\cdot)$, $b_2(\cdot)$ und $A_{21}(\cdot)$.

Unter den Voraussetzungen (A1) und (A2) (vgl. 2.12) existiert ein $L>0$, so daß

$$\left|f_0(x,\xi)-f_0\left(x,\tilde{\xi}\right)\right| \le L\|x\|\max\left\{1,\|\xi\|,\left\|\tilde{\xi}\right\|\right\}\left\|\xi-\tilde{\xi}\right\|, \qquad \forall\xi,\tilde{\xi}\in\Xi \quad \forall x\in X_1\cap c\ell U.$$

Offene Frage: Wie ist die Struktur von f_0 im Fall $T > 2$?

Wir wählen $\mathcal{F} = \mathcal{F}_2(\Xi)$, wobei für $p \geq 1$

$$\begin{aligned}
\mathcal{F}_p(\Xi) &:= \left\{ f : \Xi \to \mathbb{R} \,\middle|\, \left| f(\xi) - f\left(\tilde{\xi}\right) \right| \leq \max\left\{1, \|\xi\|^{p-1}, \left\|\tilde{\xi}\right\|^{p-1}\right\} \left\|\xi - \tilde{\xi}\right\| \; \forall \xi, \tilde{\xi} \in \Xi \right\}, \\
\zeta_p &:= d_{\mathcal{F}_p(\Xi)}, \\
\mathcal{P}_p(\Xi) &:= \left\{ Q \,\middle|\, \int_\Xi \|\xi\|^p \, Q(d\xi) < \infty \right\}.
\end{aligned}$$

ζ_p ist definiert auf $\mathcal{P}_p(\Xi)$ (Fortet-Mourier-Metriken ('53; [Ra91])).

Es gilt: Eine Folge $\{P_n\}_{n\in\mathbb{N}}$ konvergiert gegen P im Sinne von $\zeta_p \Leftrightarrow \{P_n\}_{n\in\mathbb{N}}$ konvergiert schwach gegen P und

$$\int_\Xi \|\xi\|^p \, P_n(d\xi) \to \int_\Xi \|\xi\|^p \, P(d\xi).$$

Ferner existieren Dualitätssätze und Abschätzungen für ζ_p (vgl. [Ra91]).

Bemerkung 2.23 (Beziehungen zur schwachen Konvergenz).

Definition (schwache Konvergenz ([Bi68, Du89])). Eine Folge $\{P_n\}_{n\in\mathbb{N}}$ in $\mathcal{P}(\Xi)$ *konvergiert schwach* gegen P, gdw. für jede (Lipschitz-)stetige, beschränkte Funktion $f : \Xi \to \mathbb{R}$:

$$\lim_{n\to\infty} \int_\Xi f(\xi) \, P_n(d\xi) = \int_\Xi f(\xi) \, P(d\xi)$$

Notation. $P_n \xrightarrow{w} P$ (w ... weak)

Definition (P-Uniformitätsklasse). Konvergiert $d_\mathcal{F}(P_n, P)$ gegen 0 für jede Folge $\{P_n\}_{n\in\mathbb{N}}$ mit $P_n \xrightarrow{w} P$, so heißt $\mathcal{F}$ *P-Uniformitätsklasse*.

Proposition (Topsøe '67). *$\mathcal{F}$ ist eine P-Uniformitätsklasse, gdw. $\mathcal{F}$ gleichmäßig beschränkt ist und $\mathcal{F}$ ist gleichgradig stetig P-f.ü. in Ξ.*

Falls $\mathcal{F}$ nicht gleichmäßig beschränkt ist, kann $d_\mathcal{F}(P_n, P) \to 0$ nur für Folgen $P_n \xrightarrow{w} P$ erwartet werden, für die $\mathcal{F}$ gleichmäßig integrierbar bzgl. $\{P_n\}_{n\in\mathbb{N}}$ ist, d.h.

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \sup_{f\in\mathcal{F}} \int_{f(\xi)>R} |f(\xi)| \, P_n(d\xi) = 0.$$

Bemerkung (diskrete Wahrscheinlichkeitsmaße). Diskrete Wahrscheinlichkeitsmaße

$$P = \sum_{j=1}^{n} p_j \delta_{\xi_j},$$

$p_j > 0$, $\xi_j \in \Xi$, $j = 1, \ldots, n$, $\sum_{j=1}^n p_j = 1$, $n \in \mathbb{N}$, δ_ξ das Dirac-Maß, $\delta_\xi(B) := \begin{cases} 1 & \xi \in B \\ 0 & \text{sonst} \end{cases}$, $B \subseteq \Xi$, liegen in vielen Fällen in den Mengen $\mathcal{P}_{\mathcal{F}_U}$.

Satz 2.24 ([Bi68]). *In $\mathcal{P}(\Xi)$ ist die Menge aller diskreten Wahrscheinlichkeitsmaße bzgl. der schwachen Konvergenz dicht.*

Kapitel 3

Stochastische Programme mit diskreten Wahrscheinlichkeitsmaßen

Gegeben sei ein diskretes Wahrscheinlichkeitsmaß

$$\tilde{P} := \sum_{j=1}^{N} p_j \delta_{\xi_j},$$

so daß $d_{\mathcal{F}_U}\left(P, \tilde{P}\right)$ "klein" ist. Dann betrachten wir (SP) mit $P := \tilde{P}$.

Konzeptionell wird das so modelliert, daß ξ auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit

$$\Omega = \{\omega_1, \ldots, \omega_N\}, \qquad p_j := \mathbb{P}(\{\omega_j\}) > 0, \quad j = 1, \ldots, N,$$

definiert ist. Dann ist $\sum_{j=1}^{N} p_j = 1$. Sei

$$\xi_t^j := \xi_t(\omega_j), \quad x_t^j := x_t(\omega_j), \qquad j = 1, \ldots, N, \ t = 1, \ldots, T.$$

Es galt:

$$\{\emptyset, \Omega\} = \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots \subseteq \mathcal{F}_t \subseteq \ldots \subseteq \mathcal{F}_T = \mathcal{F} = \text{Potenzmenge von } \Omega.$$

Sei $\mathcal{C}_t$ ein endlicher *Erzeuger* von $\mathcal{F}_t$, d.h.

$$\mathcal{F}_t = \sigma(\mathcal{C}_t) \quad \text{für} \quad \mathcal{C}_t = \{E_{ti}\}_{i \in I_t}$$

mit $\emptyset \notin \mathcal{C}_t$, $\bigcup_{i \in I_t} E_{ti} = \Omega$, $E_{ti} \cap E_{tk} = \emptyset$, $i \neq k$.

Für $E_{ti} \in \sigma(\mathcal{C}_{t+1})$ gilt wegen $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$, daß E_{ti} Vereinigung gewisser Mengen aus $\mathcal{C}_{t+1}$ ist, und

$$1 \leq \text{card} I_t \leq \text{card} I_{t+1} \leq N, \quad t = 1, \ldots, T-1.$$

Lemma 3.1. *Für jeden Zufallsvektor $\{\xi_t\}_{t=1}^{T}$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ gilt: ξ_t ist bzgl. $\mathcal{F}_t$ meßbar, gdw.*

$$\xi_t(\omega_k) = \mathbb{E}[\xi_t \,|E_{tj}](\omega_k) = \frac{\sum\limits_{i:\,\omega_i \in E_{tj}} p_i \xi_t^i}{\sum\limits_{i:\,\omega_i \in E_{tj}} p_i} \qquad \forall\, \omega_k \in E_{tj},\ j \in I_t,\ t = 1, \ldots, T.$$

Beweis. ξ_t ist bzgl. $\mathcal{F}_t$ meßbar, gdw.

$$\xi_t = \mathbb{E}[\xi_t \,|\mathcal{F}_t] = \sum_{j \in I_t} \mathbb{E}[\xi_t \,|E_{tj}]\, \chi_{E_{tj}} = \sum_{j \in I_t} \left(\frac{1}{\mathbb{P}(E_{tj})} \int_{E_{tj}} \xi_t(\omega)\, \mathbb{P}(d\omega) \right) \chi_{E_{tj}} = \sum_{j \in I_t} \frac{\sum\limits_{i:\, \omega_i \in E_{tj}} p_i \xi_t^i}{\sum\limits_{i:\, \omega_i \in E_{tj}} p_i} \chi_{E_{tj}}$$

Dann ist ξ_t konstant auf E_{tj} und es gilt $\xi_t(\omega) = \mathbb{E}[\xi_t \,|E_{tj}]$, $\forall \omega \in E_{tj}$, $j \in I_t$, $t = 1, \dots, T$. □

Bemerkung 3.2. Aus Lemma 3.1 ergeben sich für die Nichtantizipativitätsbedingungen (NA) $x_t = \mathbb{E}[x_t \,|\mathcal{F}_t]$, $t = 1, \dots, T$, in (SP) die folgenden Gleichungen:

$$x_t^k = \sum_{j \in I_t} \chi_{E_{tj}}(\omega_k) \frac{\sum\limits_{i:\, \omega_i \in E_{tj}} p_i x_t^i}{\sum\limits_{i:\, \omega_i \in E_{tj}} p_i}, \qquad k = 1, \dots, N, \quad t = 1, \dots, T-1. \tag{NA}$$

Dies sind $(T-1)\,N$ lineare (vektorielle) Gleichungsrestriktionen.

Für $t = 1$ ist $I_1 = \{1\}$ und $E_{11} = \Omega$. Dann ergibt sich

$$x_1^k = \underbrace{\left(\sum_{i=1}^N p_i \right)^{-1}}_{=1} \sum_{i=1}^N p_i x_1^i = \sum_{i=1}^N p_i x_1^i, \qquad k = 1, \dots, N,$$

also $x_1^1 = \ldots = x_1^N$.

Für $t = T$ ist $I_T = \{1, \dots, N\}$ und $\mathcal{C}_T = \{\{\omega_1\}, \dots, \{\omega_N\}\}$. Dann ergibt sich

$$x_T^k = \frac{1}{p_k} p_k x_T^k = x_T^k,$$

also sind die Gleichungen redundant.

Szenario-Schreibweise von (SP):

$$\min \left\{ \sum_{i=1}^N p_i f\left(x_1^i, \dots, x_T^i; \xi_1^i, \dots, \xi_T^i\right) \,\middle|\, x_t^i \in X_t,\ g_t\left(x_1^i, \dots, x_t^i; \xi_t^i\right) \le 0,\ \begin{array}{l} i = 1, \dots, N, \\ t = 1, \dots, T \end{array} \text{ und (NA)} \right\} \tag{SP$_s$}$$

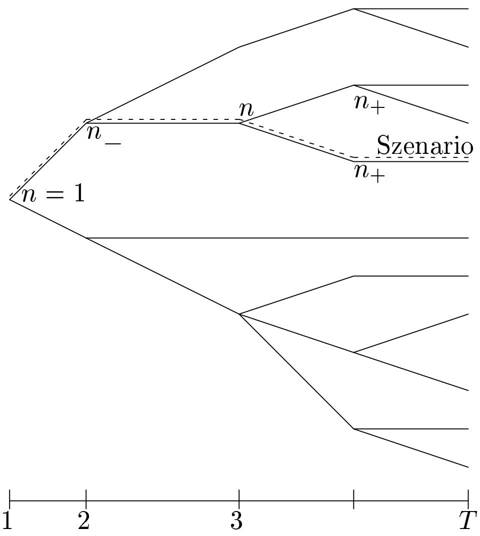
Dies ist ein Optimierungsproblem mit NT (vektoriellen) Variablen und $N(2T-1)$ Restriktionen (ohne "$x_t^i \in X_t$").

Idee: Dimensionsreduzierung durch explizite Berücksichtigung der (NA)-Bedingungen. Dabei sei $N_t = \mathrm{card}(\mathcal{C}_t) = \mathrm{card}(I_t)$ und es gilt $E_{tj} \in \sigma(\mathcal{C}_{t+1})$, $j = 1, \dots, N_t$, $t = 1, \dots, T$, d.h. E_{tj} ist eine Vereinigung gewisser Mengen $E_{t+1,k}$, $k \in I_{t+1}$. Konstruiere nun einen Graph, dessen Knotenmenge den Elementen E_{tj} von $\mathcal{C}_t$, $t = 1, \dots, T$, bzw. den Szenarien von x_t, $t = 1, \dots, T$, entsprechen und dessen Kanten den obigen "Vereinigungen" entsprechen, d.h. es existiert eine Kante zwischen den zu E_{tj} und $E_{t+1,k}$ gehörigen Knoten, gdw. $E_{t+1,k} \subseteq E_{tj}$, $j \in J_t$, $k \in J_{t+1}$, $t = 1, \dots, T-1$. Dieser Graph besitzt eine Baumstruktur mit der Wurzel in $E_{11} = \Omega$ und den Blättern $\{\omega_1\}, \dots, \{\omega_N\}$.

$\mathcal{N} \subset \mathbb{N}$ sei eine Numerierung der Knotenmenge des Graphen, so daß $n = 1$ der Wurzelknoten ist.

Notation. Zu $n \in \mathcal{N}$ bezeichne

- n_- den Vorgängerknoten von n,
- $n_{-k} := \left(n_{-(k-1)}\right)_-$, wobei $n_{-1} := n_-$,
- $\mathcal{N}_+(n)$ die Menge der Nachfolger von n,
- $t(n)$ ist so definiert, daß $n_{-t(n)+1} = 1$,
- $\mathcal{N}_t := \{n \in \mathcal{N} \,|\, t(n) = t\}$
  $\mathcal{N}_T$ ist die Menge der Blätter ($\mathrm{card}\mathcal{N}_T = N$).
- $i(n)$ der Szenario-Index aus $\{1, \ldots, N\}$, der zum Blatt n gehört



Wir definieren:

$$\begin{aligned}\pi_n &:= p_{i(n)} = \mathbb{P}\left(\left\{\omega_{i(n)}\right\}\right) && \forall n \in \mathcal{N}_T,\\ \pi_n &:= \sum_{n_+ \in \mathcal{N}_+(n)} \pi_{n_+} && \forall n \in \mathcal{N} \setminus \mathcal{N}_T.\end{aligned}$$

Dann ist $\sum_{n \in \mathcal{N}_t} \pi_n = 1$ für alle $t = 1, \ldots, T$.
Wir identifizieren nun

$$\left\{\xi_t^i\right\}_{t=1}^T \Leftrightarrow \{\xi^n\}_{n \in \mathcal{N}} \qquad \left\{x_t^i\right\}_{t=1}^T \Leftrightarrow \{x^n\}_{n \in \mathcal{N}}$$

und erhalten die **Szenariobaum-Schreibweise von (SP)**:

$$\min \left\{ \sum_{n \in \mathcal{N}_T} \pi_n f\left(x^1, \ldots, x^{n_-}, x^n; \xi^1, \ldots, \xi^{n_-}, \xi^n\right) \,\middle|\, \begin{array}{l} x^n \in X_{t(n)} \\ g_{t(n)}\left(x^1, \ldots, x^n; \xi^n\right) \leq 0 \end{array} \quad \forall n \in \mathcal{N} \right\} \quad (\mathrm{SP}_B)$$

Dies ist ein Optimierungsproblem mit $\mathrm{card}(\mathcal{N})$ (vektoriellen) Variablen und $\mathrm{card}(\mathcal{N})$ (vektoriellen) Restriktionen (ohne "$x^n \in X_{t(n)}$").

Bemerkung 3.3. Haben die Funktion f und g_t die spezielle *separable* Struktur

$$\begin{aligned} f(x_1, \ldots, x_T; \xi_1, \ldots, \xi_T) &= \sum_{t=1}^T f_t(x_t, \xi_t),\\ g_t(x_1, \ldots, x_t; \xi_t) &= A_{tt} x_t + A_{t,t-1}(\xi_t) x_{t-1} - b_t(\xi_t), \end{aligned}$$

so entstehen in der Baum-Formulierung folgende Strukturen:

$$\begin{aligned} \sum_{n \in \mathcal{N}_T} \pi_n f\left(x^1, \ldots, x^n; \xi^1, \ldots, \xi^n\right) &= \sum_{n \in \mathcal{N}} \pi_n f_{t(n)}\left(x^n; \xi^n\right)\\ g_{t(n)}\left(x^1, \ldots, x^n; \xi^n\right) &= A_{t(n),t(n)} x^n + A_{t(n),t(n)-1}\left(\xi^n\right) x^{n_-} - b_{t(n)}\left(\xi^n\right) \end{aligned}$$

Die Restriktionsmatrix ist dabei strukturiert und schwach besetzt.

Bemerkung 3.4 (Generierung von Szenariobäumen). Generelle Methodik:

1. Erstellung eines stochastischen Modells für den Prozeß $\{\xi_t\}_{t=1}^T$ (Varianten: Historische Verläufe nach erfolgter Datenanalyse, Anpassung von Zeitreihen- oder Regressions-Modellen) und Simulation von Szenarien.
2. Untersuchung der Ähnlichkeiten von Szenarien in jedem t mit "Bündelung" ähnlicher Szenarien (Cluster-Analyse).

Für eine konkrete Anwendung müssen nur das Modell in 1. und das Ähnlichkeitsmaß in 2. geeignet gewählt werden.

Kapitel 4

Dekompositionsmethoden

Es existieren **primale** und **duale** Zugänge zur Dekomposition, d.h. sukzessive Zerlegung der großen, strukturierten Optimierungsprobleme (SP_s) bzw. (SP_B). Literatur: [BL97, RS03]

4.1 Primal

Idee: Sukzessive Einführung von "Schnitten" ("cuts"), d.h. lineare Restriktionen zur Verkleinerung der zulässigen Menge und der Verbesserung der Zielfunktion.

Methodik: nested Benders - Dekomposition (u.a. Gassmann '90)

4.2 Dual

Die Dualisierung von Nichtantizipativitätsbedingungen führt zur Szenario-Dekomposition. Die Dualisierung verkoppelnder Restriktionen führt zur geographischen Dekomposition. Es entsteht dann zum primalen Problem (SP_s) bzw. (SP_B) ein duales Problem.

Danach wird im linearen Fall (bei Vorliegen von Dualität) ein sogenanntes primal-duales Lösungsverfahren angewendet, das gleichzeitig das primale und duale Problem löst (z.B. augmented Lagrangian methods, splitting methods, interior point methods). Im nichtkonvexen Fall wird das duale Problem gelöst und damit eine untere Schranke für (SP) erhalten. Dies ist Startpunkt für Globalisierungstechniken (branch-and-bound, Heuristiken, etc.)

4.2.1 Szenario-Dekomposition

Das duale Problem ist

$$\max_{\lambda_1} D_1(\lambda_1)$$

wobei die duale Funktion $D_1(\cdot)$ definiert ist mittels

$$\begin{aligned} D_1(\lambda_1) &:= \inf \left\{ \sum_{i=1}^{n} p_i \left[f\left(x_1^i, \ldots, x_T^i; \xi_1^i, \ldots, \xi_T^i\right) + \sum_{t=1}^{T} x_t^i H_t^i(\lambda_{1,t}) \right] \;\middle|\; \begin{array}{r} x_t^i \in X_t \\ h_t\left(x_t^i, \xi_t^i\right) \leq 0 \\ g_t\left(x_1^i, \ldots, x_t^i; \xi_t^i\right) \leq 0 \\ t = 1, \ldots, T \\ i = 1, \ldots, N \end{array} \right\} \\ &= \sum_{i=1}^{n} p_i D_{1i}(\lambda_1) \end{aligned}$$

wobei für $i = 1, \ldots, N$

$$D_{1i}(\lambda_1) := \inf \left\{ f\left(x_1^i, \ldots, x_T^i; \xi_1^i, \ldots, \xi_T^i\right) + \sum_{t=1}^{T} x_t^i H_t^i(\lambda_{1t}) \;\middle|\; \begin{array}{r} x_t^i \in X_t \\ h_t\left(x_t^i, \xi_t^i\right) \leq 0 \\ g_t\left(x_1^i, \ldots, x_t^i; \xi_t^i\right) \leq 0 \\ t = 1, \ldots, T \end{array} \right\}.$$

Dies entspricht N Optimierungsproblemen, eines für jedes Szenario.

4.2.2 Geographische Dekomposition

Ausgangspunkt sei das folgende primale Problem in Szenariobaumform ($m \ll N(T-1)$):

$$\min \left\{ \sum_{n \in \mathcal{N}} \pi_n \sum_{j=1}^{m} f_{t(n)}^j\left(x_j^n; \xi^n\right) \;\middle|\; \begin{array}{rl} x_j^n \in X_{t(n)}^j & j = 1, \ldots, m \\ g_{t(n)}^j\left(x_j^{n-}, x_j^n; \xi^n\right) \leq 0 & j = 1, \ldots, m \\ \sum_{j=1}^{m} h_{t(n)}^j\left(x_j^n; \xi^n\right) \leq 0 & n \in \mathcal{N} \end{array} \right\}$$

Dann zerfällt die duale Funktion $D_2(\cdot)$ folgendermaßen

$$D_2(\lambda_2) = \sum_{j=1}^{m} D_{2j}(\lambda_2)$$

wobei für $j = 1, \ldots, m$

$$D_{2j}(\lambda_2) := \inf \left\{ \sum_{n \in \mathcal{N}} \left[\pi_n f_{t(n)}^j\left(x_j^n; \xi^n\right) + \lambda_2^n h_{t(n)}^j\left(x_j^n; \xi^n\right)\right] \;\middle|\; \begin{array}{rl} x_j^n \in X_{t(n)}^j & n \in \mathcal{N} \\ g_{t(n)}^j\left(x_j^{n-}, x_j^n; \xi^n\right) \leq 0 & \end{array} \right\}$$

Dies sind m Optimierungsprobleme, jeweils für $\left\{x_j^n\right\}_{n \in \mathcal{N}}$, $j = 1, \ldots, m$.

Literaturverzeichnis

[Bi68] Billingsley: Convergence of probability measures, Academic Press, 1968

[BL97] Birge, Louveaux: Introduction to stochastic programming, Springer 1997
Bibliothek: SK870 B618

[DR02] D. Dentcheva, W. Römisch: Duality gaps in nonconvex stochastic optimization, Mathematical Programming Ser. A 101 (2004), 515-535

[DE76] Dynkin, Evstigneev: Regular conditional expectation of correspondences, Theory of Probability and its Applications 21, 1976

[Du89] Dudley: Real Analysis and Probability, 1989

[Ev76] I.V. Evstigneev: Measurable selections and dynamic programming, Math. of Operations Research, 1, 1976

[KW94] Kall, Wallace: Stochastic Programming, Wiley, 1994
Bibliothek: SK870 K14

[LR01] Lemarechal, Renaud: Math. Progr. 2001

[Pr95] Prékopa: Stochastic Programming, Kluwer, 1995, 0-7923-3482-5
Bibliothek: Wp1807a Ls

[Ra91] Rachev: Probability Metrics and the Stability of Stochastic Models, Wiley, 1991
Bibliothek: Wr1032 Ls

[RS03] Rieszczynski, Shapiro (Eds.): Handbook of stachastic programming, Elsevier, 2003

[RW98] Rockafellar, Wets: Variational Analysis, Springer, Berlin 1998 (Chapter 14)
Bibliothek: SK660 R682

[Sh96] Shiryaev: Probability, Springer, 1996

[SI78] SIAM Journal Control Optimization 16, 1978

[WW69] Walkup, Wets: Pacific Journal of Mathematics 28 (1969), 465-475

Index