Technische Universität Wien

Seminararbeit

Wahrscheinlichkeitsmaße mit vorgegebenen Marginalmaßen

Autor:
Daniel TONEIAN

Betreuer:
Ao.Univ.Prof. Dipl.-Ing. Dr.techn.
Harald WORACEK

19. Juni 2019

Inhaltsverzeichnis

1 Einleitung

Seien $(X, \mathcal{F})$ und $(Y, \mathcal{G})$ Mengen mit σ-Algebren und $\mathcal{F} \times \mathcal{G}$ die Produkt-σ-Algebra auf $X \times Y$. Für ein Maß λ auf $(X \times Y, \mathcal{F} \times \mathcal{G})$ sind Maße μ und ν auf $(X, \mathcal{F})$ bzw. $(Y, \mathcal{G})$ durch $\mu(A) := \lambda(A \times Y)$ bzw. $\nu(B) := \lambda(X \times B)$ für $A \in \mathcal{F}, B \in \mathcal{G}$. Dies ist äquivalent dazu, dass $\mu = \lambda \circ \mathrm{proj}_X^{-1}$ bzw. $\nu = \lambda \circ \mathrm{proj}_Y^{-1}$ ist, wobei proj die entsprechende Projektion bezeichnet. Diese Maße μ, ν heißen die Marginalmaße von λ.
Hat man umgekehrt Maße μ und ν auf $(X, \mathcal{F})$ bzw. $(Y, \mathcal{G})$, stellt sich die Frage, ob es ein Maß λ auf $(X \times Y, \mathcal{F} \times \mathcal{G})$ gibt, welches μ und ν als Marginalmaße hat. Wie aus der Maßtheorie bekannt ist, ist die Antwort stets ja, nämlich hat das Produktmaß $\mu \times \nu$ die gewünschte Eigenschaft.
Für gegebene μ, ν ist jedoch das Produktmaß typischerweise nicht das einzige Maß λ mit den Marginalmaßen μ und ν. Man kann sich deswegen fragen, welche zusätzlichen Anforderungen man an λ stellen kann. Zwei solche mögliche Arten von Anforderungen sollen hier für den Fall von Wahrscheinlichkeitsmaßen vorgestellt werden.
Im Wesentlichen folgen die Struktur und der Inhalt dieser Arbeit Teilen einer Publikation von V. Strassen (Strassen, 1965). Im 3. Abschnitt geben wir einen Einschub zu analytischen Mengen, der eine Auswahl von Resultaten ist, die in (Kechris, 2012) gefunden werden können.

2 Wahrscheinlichkeitsmaße mit vorgegebenen Marginalmaßen

Definition 2.1. Sei X ein vollständiger, separabler, metrischer Raum. Ein Gewicht $\hat{x}$ sei eine positive, aber nicht notwendigerweise nach oben beschränkte, stetige Funktionen auf X, für die es also eine positive reelle Zahl m gibt, sodass m eine untere Schranke für $\hat{x}$ ist. Dann sei $C_b(X, \hat{x})$ der Raum aller stetigen, reellwertigen Funktionen x auf X, für die die Funktion $||.||$ definiert durch

$$||x|| = \sup\{|x(s)|/\hat{x}(s) : s \in X\},$$

endlich ist.

Bemerkung 2.2. Für einen vollständigen, separablen, metrischen Raum X mit Gewicht $\hat{x}$ ist $C_b(X, \mathbb{R}, \hat{x})$ ein Banachraum, da die Division durch $\hat{x}$ ein Isomorphismus auf den Raum aller beschränkten Funktionen auf X mit der Supremumsnorm ist.

Bemerkung 2.3. Zu Räumen $X_1, X_2, \ldots, X_n$ mit Gewichten $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$ kann man den Raum $C_b(\Pi_{i=1}^n X_i, \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i})$ definieren. Klarerweise ist $x \in C_b(X_i, \hat{x}_i)$ genau dann, wenn $x \circ \mathrm{proj}_{X_i} \in C_b(\Pi_{i=1}^n X_i, \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i})$.

Bevor wir das erste Hauptresultat vorgestellen, müssen wir noch an zwei aus der Stochastik bekannte Sätze erinnern. Der Beweis dieser Sätze würde den Umfang dieser Arbeit sprengen. Wir verweisen den Leser auf (Kechris, 2012) Seite 112 sowie (Bogachev, 2007) Seite 202.

Satz 2.4. *Sei X ein vollständiger, separabler metrischer Raum. Dann ist die Menge aller endlichen Maße auf X mit der Topologie der schwachen Konvergenz so metrisierbar, dass sie zu einem vollständigen, separablen metrischen Raum wird.*

Satz 2.5. *Für eine Menge M von Maßen auf einem separablen, vollständigen metrischen Raum X sind äquivalent:*

1. *Der Abschluss von M bezüglich der schwach-$*$-Topologie ist kompakt.*

2. *$\sup_{\mu \in M} \mu(X) < \infty$ und M ist straff, das heißt, für jedes $\epsilon > 0$ gibt es eine kompakte Menge $K^\epsilon \subset X$, sodass für alle $\mu \in M$ folgendes gilt: $\mu(X \backslash K^\epsilon) < \epsilon$.*

Satz 2.6. *Seien vollständige, separable, metrische Räume $X_1, X_2, \ldots, X_n$, mit Gewichten $\hat{x}_1$, $\hat{x}_2, \ldots, \hat{x}_n$ gegeben. Sei Π die Menge aller Wahrscheinlichkeitsmaße π auf $\Pi_{i=1}^n X_i$ sodass $\Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i}$ π-integrierbar ist, versehen mit der Topologie $\mathcal{J}$, die durch die Funktionale $\pi \mapsto \int x\, d\pi$ für $x \in C_b(\Pi_{i=1}^n X_i, \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i})$ induziert wird. Das ist die relative schwach-$*$-Topologie, wenn Π als Teilmenge von $C_b(\Pi_{i=1}^n X_i, \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i})^*$ aufgefasst wird. Sei Λ eine nicht-leere, $\mathcal{J}$-abgeschlossene und konvexe Teilmenge von Π und sei für $1 \leq i \leq n$ ein Borelwahrscheinlichkeitsmaß μ_i in X_i gegeben, sodass $\hat{x}_i$ bezüglich μ_i-integrierbar ist. Eine notwendige und hinreichende Bedingung für die Existenz eines Wahrscheinlichkeitsmaßes $\lambda \in \Lambda$ sodass $\mu_i = \lambda \circ \mathrm{proj}_{X_i}^{-1}$ für $1 \leq i \leq n$ ist, dass*

$$\sum_{i=1}^{n} \int x_i\, d\mu_i \leq \sup\{\int \sum_{i=1}^{n} x_i \circ \mathrm{proj}_{X_i}\, d\gamma : \gamma \in \Lambda\} \tag{1}$$

für $x_i \in C_b(X_i, \mathbb{R}, \hat{x}_i), 1 \leq i \leq n$.

Beweis. Eine kleine Rechnung

$$\sum_{i=1}^{n} \int x_i\, d\mu_i = \int \sum_{i=1}^{n} x_i \circ \mathrm{proj}_{X_i}\, d\lambda \leq \sup\{\int \sum_{i=1}^{n} x_i \circ \mathrm{proj}_{X_i}\, d\gamma : \gamma \in \Lambda\}$$

zeigt, dass die Ungleichung tatsächlich eine notwendige Bedingung ist. Deswegen sei (1) angenommen. Sei M die Menge aller Tupel $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, in denen für jedes $1 \leq i \leq n$ α_i ein Wahrscheinlichkeitsmaß auf X_i ist und sodass $\hat{x}_i$ α_i-integrierbar ist. Dann ist offensichtlich $(\mu_1, \mu_2, \ldots, \mu_n) \in M$. M kann durch die Zuweisung

$$(\alpha_1, \alpha_2, \ldots, \alpha_n)(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \int x_i\, d\alpha_i$$

als Teilmenge von $(\Pi_{i=1}^n X_i)^*$ aufgefasst werden. Die relative schwach-$*$-Topologie auf M sei $\mathcal{J}_1$. $\mathcal{J}_1$ ist metrisierbar: Es genügt dies anstatt für das Produkt $\Pi_{i=1}^n X_i$ nur für einen Raum X_i zu zeigen, denn das Produkt metrisierbarer Räume ist wieder metrisierbar. Hierfür sei Π_{X_i} die Menge aller Wahrscheinlichkeitsmaße α in X_i, für die $\hat{x}_i$ α_i-integrierbar ist. Wir müssen nun zeigen, dass Π_{X_i} metrisierbar als Teilmenge von X_i^* mit der schwach-$*$-Topologie ist. Sei für jede Borelmenge $E \subset X_i$

$$(\phi_{X_i}\alpha)(E) = \int_E \hat{x}_i\, d\alpha_i,$$

dann ist ϕ_{X_i} ein Homöomorphismus von Π_{X_i} auf die abgeschlossene Teilmenge $\{m : \int (\frac{1}{\hat{x}_i})\, dm = 1\}$ der Menge aller endlichen positiven Maße in X_i mit der schwach-$*$-Topologie. Die Homöomorphie ist sofort ersichtlich, wenn man sich vor Augen hält, dass für jede stetige und beschränkte Funktion f auf X_i gilt, dass $f\hat{x}_i \in X_i$ und umgekehrt für jedes $x \in X_i$ die Funktion $\frac{x}{\hat{x}_i}$ eine stetige und beschränkte Funktion auf X_i ist. Π_{X_i} ist also metrisierbar. Sei M_Λ die Menge aller $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in M$, sodass es ein $\gamma \in \Lambda$ gibt, welches die Randmaße α_i für $1 \leq i \leq n$ hat. M_Λ ist klarerweise konvex und $(\mu_1, \mu_2, \ldots, \mu_n)$ ist im $\mathcal{J}_1$-Abschluss von M_Λ, da sonst $(\mu_1, \mu_2, \ldots, \mu_n)$ auch nicht im schwach-$*$-Abschluss von M_Λ (in $(\Pi_{i=1}^n X_i)^*$) wäre, es also nach dem Satz von Hahn-Banach ein $(x_1, x_2, \ldots, x_n) \in C_b(\Pi_{i=1}^n X_i, \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i})$ geben müsste, für das

$$(\mu_1, \mu_2, \ldots, \mu_n)(x_1, x_2, \ldots, x_n) > \sup\{(\alpha_1, \alpha_2, \ldots, \alpha_n)(x_1, x_2, \ldots, x_n) : (\alpha_1, \alpha_2, \ldots, \alpha_n) \in M_\Lambda\},$$

was ein Widerspruch zur angenommenen Ungleichung (1) ist. Sei $(\alpha_1^k, \alpha_2^k, \ldots, \alpha_n^k)_{k \geq 1}$ eine Folge von Elementen in M_Λ, die gegen $(\mu_1, \mu_2, \ldots, \mu_n)$ $\mathcal{J}_1$-konvergiert und sei $\lambda_k \in \Lambda$ so, dass es die

Randmaße $(\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k)$ hat. Dann gilt:

$$\lim_{k\to\infty} \int \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i}\ d\lambda_k = \lim_{k\to\infty} \sum_{i=1}^n \int \hat{x}_i\, d\alpha_i^k = \sum_{i=1}^n \int \hat{x}_i\, d\mu_i,$$

also $\sup_{n\in\mathbb{N}} \int \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i}\ d\lambda_n < \infty$. Da die Mengen $\{\alpha_i^k : k \in \mathbb{N}\}$ für $1 \leq i \leq n$ präkompakt sind, sind sie auch straff, und da die Funktionen ϕ_{X_i} Homöomorphismen sind, gibt es für jedes $\epsilon > 0$ und $1 \leq i \leq n$ Mengen $K_{X_i}^\epsilon \subset X_i$, sodass

$$\sup_{k\in\mathbb{N}} \int_{X_i \setminus K_{X_i}^\epsilon} \hat{x}_i\, d\alpha_i^k < \epsilon.$$

Deswegen gilt für $K^\epsilon := \Pi_{i=1}^n K_{X_i}^\epsilon$

$$\sup_{k\in\mathbb{N}} \int_{(\Pi_{i=1}^n X_i)\setminus K^\epsilon} \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i}\ d\lambda_k < n\epsilon.$$

Ist nun $\phi\gamma$ durch $(\phi\gamma)(E) := \int_E \Sigma_{i=1}^n \hat{x}_i \circ \mathrm{proj}_{X_i}\ d\gamma$ für jede Borelmenge $E \in \Pi_{i=1}^n X_i$ definiert, so folgt, dass $(\phi\lambda_k)_{k\in\mathbb{N}}$ straff und in weiterer Folge auch relativ kompakt ist. Das selbe gilt auch für $(\lambda_k)_{k\in\mathbb{N}}$ bezüglich der $\mathcal{J}$-Topologie. Dann gilt für jeden Häufungspunkt λ von $(\lambda_k)_{k\in\mathbb{N}}$, dass $\lambda \in \Lambda$, da Λ abgeschlossen ist. Die Beziehung (1) gilt, da Projektionen stetig sind. □

3 Zwischenspiel: Analytische Mengen

Für unser zweites Hauptresultat wird aus technischen Gründen ein Einschub nötig. In jenem Theorem werden die Maße von Projektionen von Borelmengen betrachtet. Unangenehmer Weise sind diese Mengen im Allgemeinen nicht mehr borel, deswegen ist apriori nicht klar, dass das Maß überhaupt auf dieser Menge definiert ist. Abhilfe schafft der Begriff der analytischen Menge, der hier eingeführt wird. Da diese Begriffsbildung sehr nützlich, aber nicht aus unseren Vorlesungen bekannt ist, wird hier eine etwas längere Einführung gegeben. Diese ist dem Buch (Kechris, 2012) entnommen, wo der Leser auch alle Beweis findet.

Definition 3.1. Ein topologischer Raum $(X, \mathcal{T})$ heißt vollständig metrisierbar, falls es eine mit der Topologie $\mathcal{T}$ verträgliche Metrik d gibt, sodass (X, d) ein vollständiger metrischer Raum ist.

Definition 3.2. Ein separabler, vollständig metrisierbarer topologischer Raum wird als polnischer Raum bezeichnet.

Beispiel 3.3. Beispiele polnischer Räume sind $\mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathbb{C}^n, \mathbb{R}^\mathbb{N}, \mathbb{C}^\mathbb{N}$, das Einheitsintervall $\mathbb{I} = [0, 1]$, der Cantorraum $\mathcal{C} = 2^\mathbb{N}$ und der Baireraum $\mathcal{N} = \mathbb{N}^\mathbb{N}$

Polnische Räume habe eine interessante Struktur:

Satz 3.4. *Jeder separable metrisierbare Raum ist zu einem Unterraum des Hilbertwürfels $\mathbb{I}^\mathbb{N}$ homöomorph. Die polnischen Räume sind genau die homöomorphen Bilder der G_δ Unterräume des Hilbertwürfels.*

Wir kommen zum Begriff der analytischen Menge.

Definition 3.5. Sei X ein polnischer Raum. Eine Menge $A \subseteq X$ heißt analytisch, falls es einen polnischen Raum Y und eine stetige Funktion $f : Y \to X$ gibt, für die $F(Y) = A$ gilt. Die Menge aller analytischen Mengen eines Raumes X wird als $\Sigma_1^1(X)$ notiert.

Es gibt typischerweise mehr analytische Mengen als Borelmengen, was im nächsten Satz klar zum Ausdruck kommt.

Satz 3.6. *Sei X ein überabzählbar großer polnischer Raum. Dann gilt: $\mathcal{B}(X) \subsetneq \Sigma_1^1(X)$.*

Definition 3.7. Sei X ein polnischer Raum und $(P_s)_{s\in\mathbb{N}^{<\mathbb{N}}}$ eine Familie von Teilmengen, genannt Suslin-Schema, die durch $\mathbb{N}^{<\mathbb{N}}$, der Menge aller endlichen Folgen in $\mathbb{N}$, indiziert ist. Für $s \in \mathbb{N}^{\mathbb{N}}$ setzen wir

$$\mathcal{A}_s P_s = \bigcup_{x\in\mathcal{N}} \bigcap_n P_{x_{|n}}.$$

Ist nun Γ eine beliebige Menge von Teilmengen von X, so bezeichnet $\mathcal{A}\Gamma$ die Menge aller $\mathcal{A}_s P_s$, für die $P_s \subset X$ in Γ ist. Der Operator $\mathcal{A}$ heißt Suslin-Operator.

Es gibt einige äquivalente Definitionen von analytischen Mengen:

Satz 3.8. *Sei X ein polnischer Raum und $A \subseteq X$. Dann sind äquivalent:*

1. *A ist analytisch.*
2. *Es gibt einen polnischen Raum Y und eine Borelmenge $B \subseteq X \times Y$ mit $A = \mathrm{proj}_X(B)$.*
3. *Es gibt ein abgeschlossenes $F \subseteq X \times \mathcal{N}$ mit $A = \mathrm{proj}_X(F)$.*
4. *Es gibt ein Suslin-Schema bestehend aus abgeschlossenen Mengen mit $A = \mathcal{A}_s F_s$*

Analytische Mengen sind nützlicherweise stabil unter vielen Operationen:

Proposition 3.9. *Seien X und Y polnische Räume, $A_n \subseteq X$ und auch $A \subseteq X$, analytisch, $B \subseteq Y$ analytisch und $f : X \to Y$ Borel, so sind auch $\cup_{n\in\mathbb{N}} A_n, \cap_{n\in\mathbb{N}} A_n, f(A)$ und $f^{-1}(B)$ analytisch.*

Analytische Mengen sind aber nicht unter Komplementen abgeschlossen. Sie haben schöne Trennungseigenschaften:

Satz 3.10. *(Lusin'scher Trennungssatz) Seien A ein polnischer Raum und $A, B \subseteq X$ disjunkte analytische Teilmengen, so gibt es eine Borelmenge $C \subseteq X$, die A und B trennt, also $A \subseteq C$ und $B \cap C = \emptyset$*

Man kann auch gewisse Aussagen über die Struktur analytischer Mengen treffen.

Satz 3.11. *Sei X ein polnischer Raum und $A \subset X$ analytisch. Dann ist A entweder abzählbar oder enthält eine Kopie der Cantormenge.*

Definition 3.12. Sei X ein polnischer Raum und sei $A \subseteq X$. Eine Menge A heißt co-analytisch, falls ihr Komplement analytisch ist. Die Menge der co-analytischen Mengen wird mit Π_1^1 bezeichnet. Mengen, die sowohl analytisch als auch co-analytisch sind, werden bi-analytisch genannt.

Analytischen Mengen haben eine enge Beziehung zu Borelmengen.

Satz 3.13. *Sei X ein polnischer Raum. Dann sind die bi-analytischen Mengen genau die Borelmengen.*

Satz 3.14. *Seien X, Y polnische Räume, und $f : X \to Y$. Dann sind folgende Eigenschaften äquivalent:*

1. *f ist borel.*

2. graph(f) *ist borel.*

3. graph(f) *ist analytisch.*

Insbesondere sind Bijektionen, die borel sind, Borelisomorphismen.

Definition 3.15. Sei X ein Hausdorffraum. Eine Kapazität auf X ist eine Abbildung $\gamma : \mathcal{P}(X) \to [0, \infty]$, sodass:

1. $A \subseteq B \implies \gamma(A) \leq \gamma(B)$;
2. $A_0 \subseteq A_1 \subseteq \cdots \implies (\gamma(A_n) \to \gamma(\cup_{n\in\mathbb{N}} A_n))$;
3. für jedes kompakte $K \subseteq X$ gilt $\gamma(K) \leq \infty$; falls weiters $\gamma(K) < r$ für ein $r \in \mathbb{R}$, dann gibt es eine offene Menge U, sodass $U \supseteq K, \gamma(U) < r$.

Beispiel 3.16. Sei X ein polnischer Raum und μ ein endliches Borelmaß auf X. Dann ist das durch μ induzierte äußere Maß μ^* eine Kapazität.

Definition 3.17. Sei γ eine Kapazität auf einem Hausdorffraum X. $A \subseteq X$ heißt $\boldsymbol{\gamma}$ -kapazitierbar, falls $\gamma(A) = \sup\{\gamma(K) : K \text{ kompakt}, K \subseteq A\}$. A heißt universell kapazitierbar, falls es für jede Kapazität γ kapazitierbar ist.

Proposition 3.18. *Sei X ein polnischer Raum, $A \subseteq X$, μ ein endliches Borelmaß auf X und $\gamma = \mu^*$. Dann ist die Menge A genau dann γ -kapazitierbar, falls sie μ-messbar ist.*

Satz 3.19. *(Choquetscher Kapazitierbarkeitssatz) Sei X ein polnischer Raum. Dann ist jede analytische Teilmenge von X universell kapazitierbar.*

4 Ein weiterer Satz über Wahrscheinlichkeitsmaße mit vorgegebenen Marginalmaßen

Lemma 4.1. *Sei $u \in L^p(\mathbb{R}^n, \mu), p \in [1, \infty)$ und $u \geq 0$, dann gilt:*

$$\int_{\mathbb{R}^n} |u|^p \, d\mu = p \int_0^\infty t^{p-1} \mu(\{|u| > t\}) \, dt.$$

Beweis. Der Satz von Fubini liefert

$$\begin{aligned} p \int_0^\infty t^{p-1} \mu(\{|u| > t\}) \, dt &= \int_0^\infty \int_{\mathbb{R}^n} p \cdot t^{p-1} \mathbb{1}_{(0,|u(x)|)}(t) \, d\mu(x) \, dt \\ &= \int_{\mathbb{R}^n} \int_0^\infty p \cdot t^{p-1} \mathbb{1}_{(0,|u(x)|)}(t) \, dt \, d\mu(x) \\ &= \int_{\mathbb{R}^n} \int_0^{|u(x)|} p \cdot t^{p-1} \, dt \, d\mu(x) \\ &= \int_{\mathbb{R}^n} |u|^p \, d\mu \end{aligned}$$

□

Lemma 4.2. *Sei S ein vollständiger, separabler metrischer Raum, $z : S \mapsto \mathbb{R}$ eine beschränkte, nicht-negative, messbare Funktion und μ ein Wahrscheinlichkeitsmaß auf S, so gilt:*

$$\int_0^{\sup z(S)} \min\{\mu(\{z(s) > r\}), 1-\epsilon\}\, dr = \sup\{\int_0^{\sup z(S)} \hat{\mu}(\{z > t\})\, dt : \\ \hat{\mu} \textit{ ist ein endliches Maß auf } S, \hat{\mu}(S) = 1-\epsilon, \hat{\mu} \leq \mu\} \tag{2}$$

Beweis. Offensichtlich ist die rechte Seite kleiner als die linke, da für jedes $\hat{\mu}$ der Integrand punktweise kleiner als auf der linken Seite ist. Sei $t^* = \inf\{t : \mu(\{z > t\}) < 1-\epsilon\}$. Definiere nun

$$\hat{\mu}_n := \mu|_{\{z>t^*+\frac{1}{n}\}} + \mu|_{\{z>t^*\}\setminus\{z>t^*+\frac{1}{n}\}} \cdot \frac{1-\epsilon-\mu(\{z > t^*+\frac{1}{n}\})}{\mu(\{z>t^*\}\setminus\{z>t^*+\frac{1}{n}\})}.$$

Dieses Maß hat nun Masse $1-\epsilon$ und es gilt $\hat{\mu}_n \leq \mu$. Die Rechnung

$$\Big|\int_0^{\sup z(S)} \min\{\mu(\{z(s) > r\}), 1-\epsilon\}\, dr - \int_0^{\sup z(S)} \hat{\mu}(\{z > t\})\, dt\Big| \\ \leq \int_{t^*}^{t^*+\frac{1}{n}} |\mu(\{z > t\}) - \hat{\mu}(\{z > t\})|\, dt \leq \frac{2}{n}$$

zeigt, dass das Supremum in (2) tatsächlich den Wert der linken Seite der Gleichung erreicht. □

Satz 4.3. *Seien X, Y vollständige, separable metrische Räume, ω eine nichtleere, abgeschlossene Teilmenge von $X \times Y$, μ ein Borelwahrscheinlichkeitsmaß auf X und ν ein Borelwahrscheinlichkeitsmaß auf Y, $\epsilon \geq 0$. Dann ist die Existenz eines Borelwahrscheinlichkeitsmaßes λ auf $X \times Y$ mit Randmaßen μ und ν und $\lambda(\omega) \geq 1-\epsilon$, äquivalent dazu, dass für alle offenen $U \subset X$*

$$\nu(U) \leq \mu(\mathrm{proj}_X(\omega \cap (X \times U))) + \epsilon. \tag{3}$$

Beweis. Zuerst wird die Notwendigkeit der Bedingung überprüft:

$$\begin{aligned} \nu(U) &\leq \mu(\mathrm{proj}_X(\omega \cap (X \times U))) + \epsilon \\ \nu(U) - \epsilon &\leq \mu(\mathrm{proj}_X(\omega \cap (X \times U))) \\ \lambda(X \times U) - \epsilon &\leq \lambda(\mathrm{proj}_X(\omega \cap (X \times U)) \times Y) \\ 1 - \lambda(X \times U^c) - \epsilon &\leq \lambda(\mathrm{proj}_X(\omega \cap (X \times U)) \times Y) \\ 1 - \epsilon &\leq \lambda(\mathrm{proj}_X(\omega \cap (X \times U)) \times Y) + \lambda(X \times U^c) \end{aligned}$$

Hier sieht man sofort, dass ω einerseits Teilmenge der Vereinigung der beiden Mengen auf der rechten Seite ist, andererseits ist aber nach Annahme $1-\epsilon \leq \lambda(\omega)$, also die Ungleichung erfüllt. Um zu zeigen, dass (3) hinreichend für die Existenz des Maßes ist, verwenden wir Satz 2.6, setzen $\hat{x}_1 = \hat{x}_2 = 1, \Lambda = \{\lambda : \lambda \in \Pi, \lambda(\omega) \leq 1-\epsilon\}$. Dabei ist die Konvexität von Λ trivial und die schwache Abgeschlossenheit gilt, da bekanntermaßen für jede abgeschlossene Menge A und jede Folge von Maßen $(\alpha_n)_{n\in\mathbb{N}}$, die schwach gegen ein Maß α konvergiert, $\limsup_{n\to\infty} \alpha_n(A) \leq \alpha(A)$. Wir müssen also zeigen, dass (1) erfüllt ist. Dafür dürfen wir annehmen, dass sowohl x als auch y positiv sind, da sie beschränkt sind und die Addition von Konstanten nichts an (1) ändert. Weiters können wir annehmen, dass $\mathrm{proj}_X(\omega) = X$, indem wir einen Punkt $\bar{y}$ zu Y hinzufügen und ω um $X \times \{\bar{y}\}$ erweitern. Indem wir (3) und Lemma 4.1 benutzen, errechnen wir

$$\int y\, d\nu = \int_0^{\sup y(Y)} \nu(\{t : y(t) > r\})\, dr \leq \int_0^{\sup y(Y)} \mu(\mathrm{proj}_X(\omega \cap (X \times \{t : y(t) > r\}))) \\ \leq \int_0^{\sup y(Y)} \min\{\mu(\{s : \sup y(\omega_s) > r\}) + \epsilon, 1\}\, dr,$$

wobei $\omega_s = \{t : (s,t) \in \omega\}$ ist. Sei $y_0(s) = \sup y(\omega_x)$, dann ist y_0 beschränkt, nichtnegativ und messbar und es gilt, unter Benutzung von Lemma 4.2,

$$\begin{aligned}
&\int y\, d\nu \leq \int_0^{\sup y(Y)} \min\{\mu(\{y_0 > r\}), 1-\epsilon\}\, dr + \epsilon \sup y(Y) \\
&= \int_0^{\sup y_0(X)} \min\{\mu(\{y_0 > r\}), 1-\epsilon\}\, dr + \epsilon \sup y(Y) \\
&= \sup\{\int_0^{\sup y_0(X)} \hat{\mu}(\{y_0 > t\})\, dt : \hat{\mu} \text{ ist ein endliches Maß auf } X, \hat{\mu}(X) = 1-\epsilon, \hat{\mu} \leq \mu\} + \epsilon \sup y(Y) \\
&= \sup\{\int y_0\, d\hat{\mu} : \hat{\mu} \text{ ist ein endliches Maß auf } X, \hat{\mu}(X) = 1-\epsilon, \hat{\mu} \leq \mu\} + \epsilon \sup y(Y).
\end{aligned}$$

Weiters gilt für jedes $\hat{\mu} \in \{\nu : \nu \text{ ist ein endliches Maß auf } X, \nu(X) = 1-\epsilon, \nu \leq \mu\}$

$$\int x\, d\mu = \int x\, d(\hat{\mu} + (\mu - \hat{\mu})) = \int yx\, d\hat{\mu} + \int x\, d(\mu - \hat{\mu}) \leq \int x\, d\hat{\mu} + \epsilon \sup x(X)$$

und somit gemeinsam

$$\begin{aligned}
&\int x\, d\mu + \int y\, d\nu \\
&\leq \sup\{\int (x + y_0\, d\hat{\mu} : \hat{\mu} \text{ ist ein endliches Maß auf } X, \hat{\mu}(X) = 1-\epsilon, \hat{\mu} \leq \mu\} + \epsilon(\sup x(X) + \sup y(Y)) \\
&\leq (1-\epsilon) \sup(x + y_0)(X) + \epsilon \sup(x(X) + y(Y)) \\
&= (1-\epsilon) \sup\{x(s) + y(t) : (s,t) \in \omega\} + \epsilon \sup(x(X) + y(Y)) \\
&= \sup\{\int (x \circ \mathrm{proj}_X + y \circ \mathrm{proj}_Y)\, d\gamma : \gamma \in \Lambda\}.
\end{aligned}$$

Das ist aber genau die Ungleichung, die zu zeigen war. □

Bemerkung 4.4. Wir konnten Satz 2.6 für endlich viele Räume formulieren. Für Satz 4.3 ist es auf dem ersten Blick nicht klar, wie man (3) überhaupt auf mehr als zwei Räume formulieren soll. Es stellt sich heraus, dass (3) zu

$$\mu(U) + \nu(V) \leq \sup\{\gamma(V \times Y) + \gamma(X \times U) : \gamma \text{ ist Wahrscheinlichkeitsmaß auf } X \times Y, \gamma(\omega) \geq 1-\epsilon\} \tag{4}$$

äquivalent ist, wobei dies für alle offenen $U \in X$ und $V \in Y$ gelten muss. Für (4) ist nun klar, wie man eine Formulierung für mehr als zwei Räume finden kann. Jedoch ist das keine hinreichende Bedingung mehr, um die Existenz eines Maßes mit vorgegebenen Marginalmaßen sicherzustellen, selbst wenn $\epsilon = 0$, wie folgendes Beispiel zeigt:

$$S = T = \{1, 2\}, \; R = \{1, 2, 3\},$$

$$\omega = \{(1,1,1), (1,2,2), (212), (2,2,3)\},$$

$$\mu(\{1\}) = \mu(\{2\}) = \nu(\{1\}) = \nu(\{2\}) = \frac{1}{2}, \quad \alpha(\{1\}) = \frac{1}{3}, \quad \alpha(\{2\}) = \frac{1}{2}, \quad \alpha(\{3\}) = \frac{1}{6},$$

dann gilt: $\mu(V) + \nu(U) + \alpha(W) \leq \sup\{\gamma(V \times T \times R) + \gamma(S \times U \times R) + \gamma(S \times T \times W) : \gamma(\omega) = 1\}$ für alle $V \subset S, U \subset T, W \subset R$, aber es gibt kein Wahrscheinlichkeitsmaß γ auf $S \times T \times R$ mit den Randmaßen μ, ν und α, sodass $\gamma(\omega) = 1$.

Literatur

Bogachev, V. I. (2007). *Measure theory* (Bd. 2). Springer Science & Business Media.

Kechris, A. (2012). *Classical descriptive set theory* (Bd. 156). Springer Science & Business Media.

Strassen, V. (1965). The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, *36* (2), 423–439.