Allgemeine Abweichungsmaße in der Risiko-Analysis

von

Philipp Kreins

Bachelorarbeit in Mathematik

vorgelegt der

Fakultät für Mathematik, Informatik und Naturwissenschaften

der

Rheinisch-Westfälischen Technischen Hochschule Aachen

im

September 2018

angefertigt am

Institut für Mathematik

Erstgutachter: Prof. Dr. Stanislaus Maier-Paape

Zweitgutachter: Prof. Dr. Erhard Cramer

Inhaltsverzeichnis

Einleitung

Wenn wir uns Modellierungen in der Finanzmathematik, wobei besonders die Portfolio-Theorie zu nennen ist, im Ingenieurwesen oder in der Statistik betrachten, werden in diesem Kontext häufig Zufallsvariablen verwendet und deren Erwartungswerte herangezogen, um diese grob einschätzen zu können. Jedoch können wir uns leicht vorstellen, dass deren Ausagekraft mit der Unsicherheit in den jeweiligen Zufallsvariablen, also mit „dem Maß, inwieweit sie von ihrem Erwartungswert abweichen", steht und fällt. Diese Unsicherheit wird typischerweise mit der Standardabweichung gemessen, obwohl auch gelegentlich die sogenannte mittlere absolute Abweichung Anwendung findet. Dies hat aber auch gute Gründe, denn z.B. in der Portfolio-Theorie vereinfachen sich bei Verwendung der Standardabweichung so manche Rechnungen deutlich.
Trotzdem ist sie vom Ansatz her nicht in jeder Situation bestens geeignet; denn wir können uns z.B. bei der Einschätzung von Risiken vorstellen, dass man eher daran interessiert ist, inwieweit eine Zufallsvariable unter- bzw. oberhalb des Erwartungswertes bleibt. Die Standardabweichung ermöglicht uns aber einen solchen Fokus nicht, da sie „symmetrisch" ist, d.h. Abweichungen nach oben und nach unten gleichermaßen gewichtet.
Dieses Problem haben u.A. Rockafellar et al. schon früh erkannt und haben deshalb in [12] mit Axiomen festgelegt, wann ein Funktional in diesem Kontext ein sogenanntes „Abweichungsmaß" für sie darstellt. Dieser allgemeine Ansatz hat sich als sehr sinnvoll herausgestellt; denn einerseits erfüllen neben der Standardabweichung sogar mehrere ganze Klassen an Funktionalen diese Axiome. Andererseits ist der Ansatz nicht zu allgemein gewählt, da sich unter einer wenig einschränkenden Bedingung, und zwar der sogenannten „Unterhalbstetigkeit", sehr interessante Dualitätsaussagen formulieren lassen. Daher wollen wir den Ansatz von Rockafellar et al. in dieser Arbeit vorstellen sowie die von Ihnen entwickelte Theorie genaustens ausarbeiten und an einigen Beispielen illustrieren.

Vorab stellen wir kurz die Inhalte vor: Thematisch ist diese Arbeit in zwei Kapitel unterteilt, wobei wir im ersten Kapitel hauptsächlich die Theorie entwickeln bzw. vorstellen und im zweiten Kapitel eher den Stellenwert dieser Theorie in Anwendungen beleuchten. Zusätzlich sind im Anhang einige bekannte bzw. nebensächliche Definitionen sowie nützliche Hilfsaussagen zusammengestellt.

Jedenfalls orientiert sich das **erste Kapitel** in weiten Teilen an [12], wobei wir zum Teil Schwerpunkte setzen, d.h. wir arbeiten die Beispiele detaillierter aus und verallgemeinern manche Aussagen bzw. ergänzen sie. Hingegen verzichten wir auch auf einige wenige Resultate bzw. zitieren sie nur.
Zu Beginn führen wir in **Sektion 1.1** axiomatisch die sogenannten „Abweichungsmaße" ein und diskutieren direkt einige Beispiele. Neben einem hinreichenden Kriterium dafür, wann ein „Abweichungsmaß" stetig ist, definieren wir noch die Eigenschaft der „lower range dominance", der nur manche „Abweichungsmaße" genügen, und nennen einige Rechengesetze für „Abweichungsmaße".
Anschließend kommen wir in **Sektion 1.2** zu dem zentralen Resultat von [12] und ebenfalls dieser Arbeit, und zwar der dualen Charakterisierung von „unterhalbstetigen Abweichungsmaßen", in der eine Eins-zu-Eins-Korrespondenz zwischen diesen und sogenannten „risk envelopes", die wir auch axiomatisch beschreiben, hergestellt wird. Für den Beweis verwenden wir einen bekannten Dualitätssatz über Hilberträumen. Zum besseren Verständnis präsentieren wir, wie man diese „risk envelopes" (geometrisch) interpretieren kann, und berechnen sie für einige Beispiele. Des Weiteren beschreiben wir, wie sich die „risk envelopes" unter Transformationen und Kombinationen von „unterhalbstetigen Abweichungsmaßen" verhalten.
In **Sektion 1.3** führen wir dann sogenannte „strictly expectation bounded" und „kohärente" Risikomaße ein, wozu wir natürlich wieder Beispiele anführen. Zwischen der ersteren Art dieser Risikomaße und den Abweichungsmaßen können wir über eine erstaunlich einfache Transformation eine Eins-zu-Eins-Beziehung herstellen. Obendrein erkennen wir in dieser Situation, dass sich die „Kohärenz" des Risikomaßes dual zur „lower range dominance" des korrespondierenden Abweichungsmaßes verhält. Schließlich benutzen wir diese Erkenntnis, um duale Charakterisierungen für diese Risikomaße zu erhalten, sofern sie „unterhalbstetig" sind.
Abschließend dient **Sektion 1.4** dazu, unter Benutzung von „Fehlerfunktionalen" auf zwei Arten jeweils große Klassen an Abweichungsmaßen zu erschließen. Dabei treffen wir ebenfalls Aussagen darüber, ob sie „lower range dominated" und/oder „(unterhalb-)stetig" sind. Jedoch zitieren wir dort jegliche Resultate nur ohne Beweis aus [12].

Im **zweiten Kapitel** wollen wir anhand zweier Beispiele, denen wir je einen Abschnitt widmen, festmachen, wie man die Resultate aus dem ersten Kapitel in praxisnäheren Situationen anwenden kann.
In der **Sektion 2.1** konstruieren wir ein Abweichungsmaß, das aktuell besonders in der Finanzmathematik eingesetzt wird. Als Ausgangspunkt dient der „Value-at-Risk", der aber nur unzureichende Charakteristika aufweist. Deshalb wurde der sogenannte „Conditional Value-at-Risk" entwickelt, aus dem wir mithilfe von Aussagen aus dem ersten Kapitel dieses gewünschte Abweichungsmaß mit schönsten Eigenschaften generieren können. Anschließend stellen wir auf Grundlage von [1] dann noch mit viel Arbeit einen Zusammenhang zwischen dem „Value-at-Risk" und

dem „Conditional Value-at-Risk“ her.
Schließlich stellen wir in **Sektion 2.2** unter einschränkenden Annahmen an den zugrunde liegenden Wahrscheinlichkeitsraum vor, wie aus allgemeinen Abweichungsmaßen „Risikomaße“ für die Portfolio-Theorie, also im Sinne von [9], generiert werden. Jedoch handelt es sich dabei um andere Risikomaße als diejenigen aus Abschnitt 1.3. Darüber hinaus wenden wir manche dieser „Risikomaße“ in einer Charakterisierung bzgl. der Existenz von sogenannten „master funds“ - ebenfalls aus [9] - an, indem wir die dort aufgestellten Bedingungen an diesen „Risikomaßen“ verifizieren.

Zum Einstieg in die Mathematik setzen wir Folgendes fest: Es sei von nun an immer **ein Wahrscheinlichkeitsraum**

$$(\Omega, \mathcal{M}, \mathcal{P})$$

gegeben, wobei Ω eine beliebige nicht-leere Menge, $\mathcal{M}$ als Teilmenge der Potenzmenge $\mathfrak{P}(\Omega)$ eine σ-Algebra und $\mathcal{P} : \mathcal{M} \to [0,1]$ ein Wahrscheinlichkeitsmaß ist. Damit können je nach Anwendungsgebiet zwar ganz unterschiedliche Sachverhalte beschrieben sein, aber im Allgemeinen stehen die Elemente in Ω für gewisse Szenarien in der Zukunft und in $\mathcal{M}$ sind einzelne dieser Szenarien zu Ereignissen zusammengefasst, deren Eintrittswahrscheinlichkeit $\mathcal{P}$ festlegt.
Zum näheren Verständnis der nachfolgenden Inhalte benötigt man zusätzlich zu grundlegenden Ergebnissen aus der Stochastik, Maß- und Integrationstheorie sowie der Funktionalanalysis Kenntnisse aus der Theorie der

(i) L^p-Räume über $(\Omega, \mathcal{M}, \mathcal{P})$, wozu wir eine zum Teil detaillierte Zusammenfassung in Sektion A.1 des Anhangs erstellt haben,

(ii) Unter- und Oberhalbstetigkeit, dem die Sektion A.2 des Anhangs gewidmet ist, sowie

(iii) zur Fenchel-Konjugierten, wozu wir in Sektion A.3 des Anhangs einige elementare Resultate zusammengestellt haben.

Hierbei wählen wir $L^2(\Omega)$ stets als Grundraum, auf dem wir sämtliche nachfolgendenden Funktionale definieren wollen. Da es sich bei den Elementen aus $L^2(\Omega)$ um Restklassen und nicht mehr um klassische Zufallsvariablen handelt, beschreiben wir in Bemerkung (A.1.13) und in deren näheren Umgebung, wie wir hier mit diesen Elementen umgehen.
Warum wir $L^2(\Omega)$ für unsere Zwecke bevorzugen, liegt auf der einen Seite daran, dass dieser als **Hilbertraum** besonders schöne Eigenschaften besitzt, worunter z.B. fällt, dass wir einen einfachen Zugang zu den dualen Charakterisierungen in Abschnitt 1.2 erhalten. Auf der anderen Seite ist es - wie eingangs angedeutet - unser Ziel eine ganze Klasse von „Abweichungsmaßen“ zu erschließen, die sich ähnlich wie die bekannte **Standardabweichung**

$$\sigma(X) := \|X - EX\|_2 \tag{0.1}$$

für eine integrierbare Zufallsvariable $X : \Omega \to \mathbb{R}$ verhalten. Dabei suggeriert die Definition (0.1) gerade, dass $L^2(\Omega)$ der i.A. größte L^p-Raum ist, auf dem die Standardabweichung noch ein endliches Funktional definiert. (Dazu verweisen wir zum Einen auf Lemma (A.1.16) für $p \geq 2$ sowie die anschließende Bemerkung (A.1.17) und zum Anderen auf die Bemerkung (A.1.22) für $1 \leq p < 2 =: q$.)

Um das Lesen dieser Arbeit angenehmer zu gestalten, bietet es sich an dieser Stelle an, vorab einen Überblick darüber zu geben, wo in diesem Werk welche Axiome zu finden sind; denn - wie schon in der obigen Themenvorstellung angedeutet - führen wir hier nicht nur einige Funktionale axiomatisch ein, sondern charakterisieren auch die sogenannten „risk envelopes“ in der dualen Charakterisierung von „unterhalbstetigen Abweichungsmaßen“ u.A. durch Axiome und behandeln auch manche Annahmen wie Axiome. Daher hoffen wir, dass besonders in den wenigen Fällen, wo in einem Beweis o. Ä. keine Referenz bezüglich der Herkunft der verwendeten Axiome vorliegt, die nachfolgende Tabelle hilfreich ist:

Axiome	**Referenz**
$(\mathcal{D}1) - (\mathcal{D}4)$	Definition (1.1.1)
$(\mathcal{D}1'), (\mathcal{D}2')$	Lemma (1.1.4)
$(\mathcal{D}5)$	Definition (1.1.8)
$(\mathcal{Q}1) - (\mathcal{Q}4)$	Satz (1.2.3)
$(\mathcal{R}1) - (\mathcal{R}5)$	Definition (1.3.1)
$(\mathcal{R}1'), (\mathcal{R}2')$	Lemma (1.3.4)
$(r1) - (r3s)$	Definition (2.2.3)
$(A1), (A2)$	Satz (2.2.5)

1 Theorie der Abweichungsmaße

1.1 Definition und erste Eigenschaften

Wir beginnen direkt mit der grundlegenden Definition von allgemeinen „Abweichungsmaßen“. Dabei wählen wir den axiomatischen Ansatz aus [12, Definition 1] sowie den zugehörigen Ergänzungen. Es sei jedoch bemerkt, dass wir in dieser Arbeit stets die Festlegungen bzgl. des Rechnens mit $\pm\infty$ aus Bemerkung (A.1.7) verwenden.

(1.1.1) Definition (Allgemeine Abweichungsmaße)
Ein Funktional $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ heißt Abweichungsmaß, *falls es folgende Eigenschaften erfüllt:*

($\mathcal{D}1$) $\mathcal{D}(X + C) = \mathcal{D}(X)$ *für alle* $X \in L^2(\Omega)$ *und* $C \in \mathbb{R}$,

($\mathcal{D}2$) $\mathcal{D}(0) = 0$ *und* $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$ *für alle* $X \in L^2(\Omega)$ *und* $\lambda > 0$,

($\mathcal{D}3$) $\mathcal{D}(X + Y) \leq \mathcal{D}(X) + \mathcal{D}(Y)$ *für alle* $X, Y \in L^2(\Omega)$,

($\mathcal{D}4$) $\mathcal{D}(X) > 0$ *für alle nicht-konstanten* $X \in L^2(\Omega)$.

Falls $\mathcal{D}$ zusätzlich endlich ist, d.h. $\mathcal{D}\big(L^2(\Omega)\big) \subseteq [0, \infty)$ gilt, nennt man $\mathcal{D}$ entsprechend ein endliches Abweichungsmaß. ◇

(1.1.2) Bemerkung

a) Wir bezeichnen die Axiome in

→ ($\mathcal{D}1$) als **Translationsinvarianz**,

→ ($\mathcal{D}2$) als **positive Homogenität**,

→ ($\mathcal{D}3$) als **Subadditivität** und

→ ($\mathcal{D}2$) und ($\mathcal{D}3$) zusammen als **Sublinearität**.

b) In obiger Situation ist $\mathcal{D}$ insbesondere konvex; denn zu $X, Y \in L^2(\Omega)$ und $\lambda \in [0, 1]$ ist

$$\mathcal{D}\left(\lambda X + (1-\lambda)Y\right) \underset{(\mathcal{D}3)}{\leq} \mathcal{D}(\lambda X) + \mathcal{D}\left((1-\lambda)Y\right) \underset{(\mathcal{D}2)}{=} \lambda \mathcal{D}(X) + (1-\lambda)\mathcal{D}(Y).$$

c) Außerdem hängt $\mathcal{D}(X)$ nach $(\mathcal{D}1)$ nur von $X - EX$ für jedes $X \in L^2(\Omega)$ ab, d.h. es ist $\mathcal{D}(X) = \mathcal{D}(X - EX)$. Darüber hinaus ist genau dann $\mathcal{D}(X) = 0$, wenn X konstant ist, das sich unmittelbar aus einer Kombination von $(\mathcal{D}1)$, $(\mathcal{D}2)$ und $(\mathcal{D}4)$ ergibt. Dies passt natürlich auch zu unserer Idee, dass $\mathcal{D}$ ein Maß für die Unsicherheit einer Zufallsvariablen sein soll, denn eine konstante Zufallsvariable $X \in L^2(\Omega)$ stimmt schließlich in fast allen $\omega \in \Omega$ mit ihrem Erwartungswert EX überein. Jedenfalls erhalten wir insgesamt

$$\begin{aligned} 0 = \mathcal{D}(X) = \mathcal{D}(X - EX) \quad &\Leftrightarrow \quad X - EX \text{ ist konstant} \\ &\Leftrightarrow \quad X \text{ ist konstant} \\ &\Leftrightarrow \quad X = EX \\ &\Leftrightarrow \quad X - EX = 0. \end{aligned}$$

Zusammen mit $(\mathcal{D}3)$ kann man dies so interpretieren, dass $\mathcal{D}$ auf dem Unterraum

$$\{X - EX \mid X \in L^2(\Omega)\} = \{X \in L^2(\Omega) \mid EX = 0\}$$

evtl. bis auf die Eigenschaft der Symmetrie eine Art „Norm“ darstellt, wobei wir mit **Symmetrie** stets die Beziehung $\mathcal{D}(X) = \mathcal{D}(-X)$ für alle $X \in L^2(\Omega)$ bezeichnen. In der Literatur wird dieser Unterraum auch gelegentlich als der Unterraum der „reinen Unsicherheit“ von $L^2(\Omega)$ bezeichnet.

d) Tatsächlich sind eine ganze Reihe von Abweichungsmaßen nicht symmetrisch, wie wir später sehen werden. Anschaulich können wir die Symmetrie eines Abweichungsmaßes $\mathcal{D}$ unter Verwendung der Identität

$$\mathcal{D}(X) = \mathcal{D}(X - EX)$$

für jedes $X \in L^2(\Omega)$ so interpretieren, dass $\mathcal{D}$ die Schwankungen von X nach oben und unten im Vergleich zum Erwartungswert gewissermaßen gleich gewichtet.

e) Aufgrund der i.A. nicht vorhandenen Symmetrie definiert

$$\tilde{\mathcal{D}} : L^2(\Omega) \to [0, \infty], \quad \tilde{\mathcal{D}}(X) := \mathcal{D}(-X) \tag{1.1}$$

ein Funktional, das zwar offensichtlich immer noch ein Abweichungsmaß ist, aber i.A. nicht mit $\mathcal{D}$ übereinstimmt. Ferner besitzen $\mathcal{D}$ und $\tilde{\mathcal{D}}$ dieselben Charakteristika in Bezug auf Endlichkeit, (Unterhalb-)Stetigkeit und Symmetrie. Dabei wird $\tilde{\mathcal{D}}$ gelegentlich als die **Reflexion** von $\mathcal{D}$ bezeichnet. $\diamond$

Zur Illustration betrachten wir ein paar erste Beispiele.

(1.1.3) Beispiele

a) Zuallererst wollen wir uns davon überzeugen, dass die **Standardabweichung**

$$\sigma : L^2(\Omega) \to [0,\infty), \quad \sigma(X) = \|X - EX\|_2$$

auch im Sinne von Definition (1.1.1) ein Abweichungsmaß ist, wobei die Endlichkeit und Nichtnegativität schon bekannt sind.
Die vier Axiome sind aber ebenfalls offensichtlich: Denn $(\mathcal{D}1)$ ergibt sich unmittelbar aus

$$(X + C) - E[X + C] = X - EX + C - \underbrace{E[C]}_{= C} = X - EX$$

für alle $X \in L^2(\Omega)$ und $C \in \mathbb{R}$, während $(\mathcal{D}2)$ und $(\mathcal{D}3)$ aus der Linearität von $X \mapsto X - EX$ und der positiven Homogenität bzw. Subadditivität von $\|\cdot\|_2$ folgen. Schließlich sei $X \in L^2(\Omega)$ mit $\sigma(X) = 0$, das aufgrund der positiven Definitheit von $\|\cdot\|_2$ nur für $X - EX = 0$ bzw. $X = EX$ möglich ist, d.h. X ist insbesondere konstant. Daher trifft auch $(\mathcal{D}4)$ zu.
Darüber hinaus ist σ offenbar symmetrisch und stetig. Dabei ist letzteres deshalb erfüllt, weil sowohl die Zuordnung $X \mapsto X - EX$ (siehe dazu z.B. Bemerkung (A.1.17)) als auch $\|\cdot\|_2$ auf dem Raum $L^2(\Omega)$ stetig sind.

b) Im Zusammenhang mit der Standardabweichung sind auch die **Standard-Semi-Abweichungen** $\sigma_+, \sigma_- : L^2(\Omega) \to [0,\infty)$ gegeben durch

$$\sigma_+(X) := \|[X - EX]_+\|_2 \qquad \text{und} \qquad \sigma_-(X) := \|[X - EX]_-\|_2$$

von besonderer Bedeutung, wobei wir die Notationen aus Definition (A.1.23) verwenden. Aufgrund der Beziehung $[Y]_- = [-Y]_+$ für alle Zufallsvariablen $Y : \Omega \to \mathbb{R}$ ist σ_- gerade die Reflexion von σ_+ aus Bemerkung (1.1.2) e). Deshalb gelten alle folgenden Eigenschaften von σ_+ auch für σ_-:
Zunächst ist mit $X - EX \in L^2(\Omega)$ nach Lemma (A.1.24) (iv) auch $[X - EX]_+ \in L^2(\Omega)$ für alle $X \in L^2(\Omega)$, sodass σ_+ endlich ist.
Obendrein ist σ_+ ein Abweichungsmaß: $(\mathcal{D}1)$ lässt sich nämlich mit demselben Argument wie in a) zeigen und $(\mathcal{D}2)$ ist ohnehin klar. Für den Nachweis von $(\mathcal{D}3)$ verwende man die Ungleichungen aus (A.1.24) (iii) sowie die Monotonie-Eigenschaften und die Subadditivität von $\|\cdot\|_2$. Es bleibt noch $(\mathcal{D}4)$ zu zeigen: Dazu sei $X \in L^2(\Omega)$, sodass

$$0 = \sigma_+(X) = \|[X - EX]_+\|_2$$

gilt, d.h. es ist $[X - EX]_+ = 0$ bzw. $X \leq EX$. Dann muss $\{X < EX\}$ aber eine Nullmenge sein; denn sonst existiert ein $C \in \mathbb{R}$ mit $C < EX$, sodass $\mathcal{P}(\{X \leq$

$C\}) > 0$ ist. Dies liefert dann mit

$$\begin{aligned} EX = \int_{\Omega} X \,\mathrm{d}\mathcal{P} &= \int_{\{X>C\}} X \,\mathrm{d}\mathcal{P} + \int_{\{X\leq C\}} X \,\mathrm{d}\mathcal{P} \\ &\leq \int_{\{X>C\}} EX \,\mathrm{d}\mathcal{P} + \int_{\{X\leq C\}} C \,\mathrm{d}\mathcal{P} \\ &= \mathcal{P}(\{X > C\}) \cdot EX + \mathcal{P}(\{X \leq C\}) \cdot C \\ &< \mathcal{P}(\{X > C\}) \cdot EX + \mathcal{P}(\{X \leq C\}) \cdot EX = EX \end{aligned}$$

einen Widerspruch. Da $\{X < EX\}$ somit eine Nullmenge sein muss, gilt $X \geq EX$ und daher insgesamt $X = EX$, sodass X konstant ist. Damit erfüllt σ_+ auch $(\mathcal{D}4)$. Außerdem ist σ_+ wie in a) als Verknüpfung stetiger Funktionen stetig, wenn man zusätzlich Lemma (A.1.24) (iv) heranzieht.
Jedoch ist σ_+ in den meisten Fällen nicht symmetrisch. Dazu nehmen wir an, dass eine Menge $M \in \mathcal{M}$ mit $\mathcal{P}(M) \in (0, \frac{1}{2})$ existiert, das in den meisten nicht-trivialen Wahrscheinlichkeitsräumen, wie z.B. in esssentiell unendlichen Wahrscheinlichkeitsräumen, siehe Definition (A.1.18), zutrifft. Dann betrachten wir die Zufallsvariable

$$X : \Omega \to \mathbb{R}, \quad X(\omega) := \frac{1}{\mathcal{P}(M)} \cdot \mathbb{1}_M(\omega), \tag{1.2}$$

wobei $\mathbb{1}_M$ die charakteristische Funktion der Menge M sei. So ist $EX = 1$ und somit

$$[X - EX]_+(\omega) = \left(\frac{1}{\mathcal{P}(M)} - 1\right) \cdot \mathbb{1}_M(\omega)$$

bzw.

$$\big[(-X) - E[-X]\big]_+(\omega) = \mathbb{1}_{\Omega\setminus M}(\omega)$$

für alle $\omega \in \Omega$. Daraus folgt

$$\sigma_+^2(X) = \mathcal{P}(M)\left(\frac{1}{\mathcal{P}(M)} - 1\right)^2 = \frac{1-\mathcal{P}(M)}{\mathcal{P}(M)}(1 - \mathcal{P}(M)) \tag{1.3}$$

sowie

$$\sigma_+^2(-X) = \mathcal{P}(\Omega \setminus M) = 1 - \mathcal{P}(M),$$

sodass wir

$$\sigma_+^2(X) \neq \sigma_+^2(-X)$$

erhalten. Denn zum Einen ist $\mathcal{P}(M) \neq 1$ und zum Anderen ist

$$\frac{1-\mathcal{P}(M)}{\mathcal{P}(M)} = 1 \qquad \Leftrightarrow \qquad \mathcal{P}(M) = \frac{1}{2},$$

was von Vornherein ausgeschlossen wurde. Die i.A. fehlende Symmetrie ist aber auch schon unmittelbar anhand der Definition von σ_+ erkennbar, da der Schwerpunkt offensichtlich auf Schwankungen nach oben gelegt wird, während er bei σ_- auf Schwankungen nach unten liegt.

c) Oft von Interesse ist auch, wie groß die maximale Abweichung vom Erwartungswert nach oben oder unten ist. Dabei widmen wir uns zunächst der Abweichung nach unten, welche das Funktional

$$\psi_- : L^2(\Omega) \to [0, \infty], \quad \psi_-(X) := EX - \text{ess inf } X = \text{ess sup } \big(EX - X\big) \tag{1.4}$$

bestimmt, wobei das essentielle Infimum und Supremum in Defintion (A.1.4) erklärt sind. Für die obige Umformung von $\psi_-(X)$ haben wir Lemma (A.1.6) (ii) genutzt. Als erstes müssen wir die Nicht-Negativität von ψ_- begründen: Sei $X \in L^2(\Omega)$, so ist $\{X < \text{ess inf } X\}$ nach Lemma (A.1.6) (iv) eine Nullmenge und daher ist $X \geq \text{ess inf } X$. Während $\psi_-(X) = EX - \text{ess inf } X \geq 0$ für $\text{ess inf } X = -\infty$ direkt erfüllt ist, ist es für $\text{ess inf } X \in \mathbb{R}$ wegen

$$EX = \int_\Omega X \, \mathrm{d}\mathcal{P} \geq \int_\Omega \text{ess inf } X \, \mathrm{d}\mathcal{P} = \text{ess inf } X$$

auch wahr.
Hingegen ist ψ_- i.A. nicht endlich: Dafür wählen wir z.B. in der Situation von Bemerkung (A.1.22) die Zufallsvariable $X := -X_{2,q} \in L^2(\Omega)$ für $q > 2$. Diese genügt der Eigenschaft, dass $\text{ess inf } X = -\infty$ und daher $\psi_-(X) = \infty$ ist. Falls wir jedoch einschränkend annehmen, dass $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich ist, dann ist ψ_- mithilfe von Lemma (A.2.5) insbesondere endlich.
Wie zu erwarten war, ist ψ_- ebenfalls ein **Abweichungsmaß**: Dazu seien $X \in L^2(\Omega)$ und $C \in \mathbb{R}$. Dann ist nach Definition des essentiellen Infimums

$$\begin{aligned} \psi_-(X + C) = E[X + C] - \text{ess inf } (X + C) &= E[X] + C - (\text{ess inf } X + C) \\ &= EX - \text{ess inf } X = \psi_-(X), \end{aligned}$$

d.h. $(\mathcal{D}1)$ ist erfüllt. Des Weiteren ist $(\mathcal{D}2)$ klar und $(\mathcal{D}3)$ ergibt sich aus der Linearität des Erwartungswertes auf dem Raum $L^2(\Omega)$ zusammen mit der Ungleichung aus Lemma (A.1.6) (vi). Zum Nachweis von $(\mathcal{D}4)$ sei $X \in L^2(\Omega)$ mit

$$0 = \psi_-(X) = EX - \text{ess inf } X,$$

also $EX = \text{ess inf } X$. Da wir schon wissen, dass $X \geq \text{ess inf } X$ gilt, ist also insgesamt

$$X \geq \text{ess inf } X = EX.$$

Ab hier erhalten wir analog zu b), dass X konstant sein muss. Dies entspricht gerade $(\mathcal{D}4)$.
Zu den Stetigkeitseigeschaften von ψ_- lässt sich folgendes sagen: Wenn $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich ist, dann ist das Funktional $\mathcal{H}_2$ aus Lemma (A.2.5) stetig, sodass dies zusammen mit der Stetigkeit der Abbildung $X \mapsto EX$ auf dem Raum $L^2(\Omega)$ impliziert, dass auch ψ_- stetig ist. Andernfalls verwenden wir die rechte Seite der Definition von ψ_- in (1.4), um zu sehen, dass ψ_- als Komposition von

der stetigen Zuordnung $X \mapsto EX - X$ und dem unterhalbstetigen Funktional $\mathcal{H}^2$ aus Lemma (A.2.5) auch wieder unterhalbstetig ist. Dabei betonen wir aber, dass ψ_- hier in der Tat nicht stetig ist, da das Funktional $\mathcal{H}_2$ wiederum in der anderen Darstellung von ψ_- nach Lemma (A.2.5) unstetig ist.
Darüber hinaus ist ψ_- ein weiteres Beispiel für ein Abweichungsmaß, das meistens nicht symmetrisch ist. Denn unter denselben (leicht einschränkenden) Annahmen wie in b) wählen wir die Zufallsvariable X aus (1.2) als Beispiel. Dann ist einerseits ess inf $X = 0$ und andererseits

$$\text{ess inf }(-X) = -\frac{1}{\mathcal{P}(M)} < -2,$$

da $\mathcal{P}(M) \in (0, \frac{1}{2})$ angenommen wurde. Daraus folgt

$$\begin{aligned}\psi_-(X) = EX - \text{ess inf } X = 1 < (-1) - \left(-\frac{1}{\mathcal{P}(M)}\right) &= E[-X] - \text{ess inf }(-X) \\ &= \psi_-(-X).\end{aligned}$$

Wenn wir jetzt die zu ψ_- gehörige Reflexion

$$\tilde{\psi}_-(X) = \psi_-(-X) = \text{ess sup }\big(E[-X] - (-X)\big) = \text{ess sup } X - EX \tag{1.5}$$

betrachten, fällt uns sofort ins Auge, dass diese gerade die maximale Abweichung einer Zufallsvariablen $X \in L^2(\Omega)$ von ihrem Erwartungswert nach oben misst. Daher bezeichnen wir dieses Funktional mit ψ_+. Wegen dieses Zusammenhangs von ψ_- und ψ_+ können wir wiederum folgern, dass sich alle Eigenschaften von ψ_- auf ψ_+ übertragen. ◇

Bemerkenswert ist an Definition (1.1.1), dass Rockafellar et al. Abweichungsmaße in [12, Proposition 1] durch zum Teil scheinbar schwächere Axiome wie folgt charakterisieren konnten.

(1.1.4) Lemma
Ein Funktional $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ ist genau dann ein Abweichungsmaß, wenn $\mathcal{D}$ die Axiome

($\mathcal{D}1'$) $\mathcal{D}(C) = 0$ *für alle* $C \in \mathbb{R}$,
($\mathcal{D}2'$) $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$ *für alle* $X \in L^2(\Omega)$ *und* $\lambda > 0$

sowie ($\mathcal{D}3$) und ($\mathcal{D}4$) erfüllt. ◇

Beweis
„ $\Rightarrow$ “ Während ($\mathcal{D}2'$) unmittelbar aus ($\mathcal{D}2$) folgt, wählen wir für den Nachweis von ($\mathcal{D}1'$) ein beliebiges $C \in \mathbb{R}$, sodass

$$\mathcal{D}(C) = \mathcal{D}(0 + C) \underset{(\mathcal{D}1)}{=} \mathcal{D}(0) \underset{(\mathcal{D}2)}{=} 0$$

folgt.

„$\Leftarrow$“ Zunächst erhalten wir $(\mathcal{D}2)$ aus $(\mathcal{D}1')$ mit $C = 0$ und $(\mathcal{D}2')$. Ferner seien $X \in L^2(\Omega)$ und $C \in \mathbb{R}$. Dann ergibt sich

$$\begin{aligned}\mathcal{D}(X) = \mathcal{D}(X + C - C) \underset{(\mathcal{D}3)}{\leq} \mathcal{D}(X + C) + \mathcal{D}(-C) &\underset{(\mathcal{D}1')}{=} \mathcal{D}(X + C)\\ &\underset{(\mathcal{D}3)}{\leq} \mathcal{D}(X) + \mathcal{D}(C)\\ &\underset{(\mathcal{D}1')}{=} \mathcal{D}(X),\end{aligned}$$

also $\mathcal{D}(X + C) = \mathcal{D}(X)$. Somit gilt auch $(\mathcal{D}1)$. □

Nachfolgend präsentieren wir ein interessantes hinreichendes Kriterium dafür, wann ein Abweichungsmaß stetig ist, das auf [12, Proposition 2] basiert. Im Beweis verwenden wir den Begriff der „Domäne“, der in Definition (A.2.10) erklärt ist.

(1.1.5) Satz (Stetigkeit von Abweichungsmaßen)
Sei $\mathcal{D} : L^2(\Omega) \to [0, \infty)$ ein endliches und unterhalbstetiges Abweichungsmaß. Dann ist $\mathcal{D}$ sogar stetig. ◇

(1.1.6) Bemerkung
Daraus können wir schließen, dass die Abweichungsmaße ψ_- und ψ_+ in Beispiel (1.1.3) c) nicht endlich sein können, falls der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell unendlich ist; denn wir haben bereits festgestellt, dass sie in dieser Situation zwar unterhalbstetig, aber nicht stetig sind, sodass sie nach Satz (1.1.5) nicht endlich sein können.

Im Zusammenhang mit Bemerkung (A.1.17) bedeutet dies also, dass wir je ein $X \in L^2(\Omega)$ mit $\operatorname{ess\,inf} X = -\infty$ oder $\operatorname{ess\,sup} X = \infty$ finden. In beiden Fällen ist dann $|X| \in L^2(\Omega)$ und $\|X\|_\infty = \operatorname{ess\,sup} |X| = \infty$, d.h. $L^2(\Omega) \subsetneq L^\infty(\Omega)$. ◇

Beweis von Satz (1.1.5)
Nach Bemerkung (1.1.2) b) ist $\mathcal{D}$ insbesondere konvex. Zudem bedeutet die Endlichkeit von $\mathcal{D}$ gerade, dass $\operatorname{dom}(\mathcal{D}) = L^2(\Omega)$ gilt. Zusammen mit der Unterhalbstetigkeit von $\mathcal{D}$ liefert dann der nachfolgende Satz (1.1.7), dass $\mathcal{D}$ stetig ist. □

Der anschließende Satz, auf den wir uns gerade bzw. auf den sich schon Rockafellar et al. in ihrer Proposition bezogen haben, ist ein weitaus allgemeineres Resultat aus der konvexen Analysis, das zwar von Idee her grundsätzlich [3, Satz (5.22)] entspricht, aber sogar eine Verallgemeinerung davon darstellt.

(1.1.7) Satz
Sei $(\mathcal{X}, d)$ ein vollständiger metrischer linearer Raum und $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ ein unterhalbstetiges konvexes Funktional. Dann ist $\mathcal{F}$ stetig auf $\operatorname{int}(\operatorname{dom} \mathcal{F})$. ◇

Beweis
Sei $x_0 \in \operatorname{int}(\operatorname{dom} \mathcal{F})$. Somit existiert zum Einen ein $C \in \mathbb{R}$ mit $\mathcal{F}(x_0) < C$ und zum Anderen ein $r > 0$ mit

$$B_r(x_0) \subseteq \operatorname{dom} \mathcal{F}. \tag{1.6}$$

Unser Ziel ist es nun eine Umgebung U von x_0 zu finden, sodass $\mathcal{F}(x) \leq C$ für alle $x \in U$ ist. Dann liefert Satz (A.2.13) nämlich die Stetigkeit von $\mathcal{F}$ in x_0.
Dazu nehmen wir Œ an, dass $x_0 = 0$ ist; denn ansonsten betrachten wir das Funktional

$$\mathcal{G} : \mathcal{X} \to (-\infty, \infty], \quad \mathcal{G}(x) := \mathcal{F}(x + x_0),$$

das offenbar genauso wie $\mathcal{F}$ unterhalbstetig und konvex ist. Obendrein gilt

$$\mathcal{G}(0) = \mathcal{F}(x_0) < C$$

sowie $B_r(0) \subseteq \operatorname{dom} \mathcal{G}$ aufgrund der Translationsinvarianz von d und $\mathcal{G}$ ist genau dann in Null stetig, wenn $\mathcal{F}$ in x_0 stetig ist.
Unter dieser vereinfachenden Annahme definieren wir

$$O := \mathcal{V}_C(\mathcal{F}) = \{x \in \mathcal{X} \mid \mathcal{F}(x) \leq C\}$$

aus Lemma (A.2.3). So ist nach Annahme einerseits $0 \in O$ und andererseits ist O aufgrund der Unterhalbstetigkeit von $\mathcal{F}$ und Lemma (A.2.3) abgeschlossen. Darüber hinaus erhalten wir aus der Konvexität von $\mathcal{F}$, dass auch O konvex ist. Auf dieser Grundlage untersuchen wir jetzt die Menge

$$V := O \cap (-O) = \{x \in \mathcal{X} \mid \mathcal{F}(x) \leq C \text{ und } \mathcal{F}(-x) \leq C\}.$$

Da mit O auch $-O$ abgeschlossen und konvex ist sowie die Null enthält, treffen diese drei Aussagen ebenfalls auf ihre Schnittmenge V zu. Ferner ist offensichtlich $-x \in V$ für alle $x \in V$, d.h. wir erhalten

$$V = -V. \tag{1.7}$$

Zwischenbehauptung: Zu jedem $x \in \mathcal{X}$ existiert ein $\delta = \delta(x) > 0$, sodass $tx \in V$ für alle $t \in \mathbb{R}$ mit $|t| \leq \delta$ ist.

Dazu: Wir betrachten die Funktion

$$\mathcal{H} : \mathbb{R} \to (-\infty, \infty], \quad \mathcal{H}(t) := \mathcal{F}(tx),$$

welche deshalb konvex ist, da $\mathcal{F}$ konvex ist. Aufgrund der Stetigkeit der Skalarmultiplikation auf $\mathcal{X}$ existiert ein $t_0 = t_0(x) > 0$, sodass mit der Beziehung (1.6) gerade

$$tx \in B_r(0) \subseteq \operatorname{dom} \mathcal{F} \tag{1.8}$$

für alle $t \in \mathbb{R}$ mit $|t| \leq t_0$ ist. Dann können wir zu jedem $t \in (-t_0, t_0)$ den Parameter $\lambda := \frac{1}{2} - \frac{t}{2t_0} \in (0,1)$ definieren und damit

$$t = \left(\frac{1}{2} - \frac{t}{2t_0}\right)(-t_0) + \left(\frac{1}{2} + \frac{t}{2t_0}\right)t_0 = \lambda(-t_0) + (1-\lambda)t_0$$

schreiben, wobei die rechte Seite dann natürlich eine Konvexkombination darstellt. Zusammen mit der Konvexität von $\mathcal{H}$ erhalten wir deshalb

$$\begin{aligned}\mathcal{H}(t) &= \mathcal{H}\big(\lambda(-t_0) + (1-\lambda)t_0\big) \\ &\leq \lambda\mathcal{H}(-t_0) + (1-\lambda)\mathcal{H}(t_0) \\ &\leq \lambda \cdot \max\{\mathcal{H}(-t_0), \mathcal{H}(t_0)\} + (1-\lambda) \cdot \max\{\mathcal{H}(-t_0), \mathcal{H}(t_0)\} \\ &= \max\{\mathcal{H}(-t_0), \mathcal{H}(t_0)\} =: K.\end{aligned}$$

Dabei ist $K < \infty$ aufgrund der Beziehung (1.8). Insgesamt ist $\mathcal{H}(t) \leq K$ für alle $t \in (-t_0, t_0)$, sodass $\mathcal{H}$ nach Satz (A.2.13) stetig in Null ist. In Kombination mit $\mathcal{H}(0) = \mathcal{F}(0) < C$ liefert dies die Existenz eines $\delta = \delta(x) > 0$, sodass

$$\mathcal{F}(tx) = \mathcal{H}(t) < C$$

für alle $t \in \mathbb{R}$ mit $|t| \leq \delta$ ist, das gerade der Zwischenbehauptung entspricht.

Die Aussage aus der Zwischenbehauptung können wir auch wie folgt formulieren: Zu jedem $x \in \mathcal{X}$ existiert ein $\theta = \theta(x) > 0$, sodass $x \in sV$ für alle $s \in \mathbb{R}$ mit $|s| \geq \theta$ gilt. Dies impliziert

$$\mathcal{X} = \bigcup_{n=1}^{\infty} nV,$$

wobei mit V natürlich auch nV für alle $n \in \mathbb{N}$ abgeschlossen und konvex ist sowie die Null enthält. Aufgrund der Vollständigkeit von $(\mathcal{X}, d)$ ist der Kategoriensatz von Baire, siehe [10, Satz (2.3.7)], anwendbar. Aus diesem folgt, dass wir ein $n_0 \in \mathbb{N}$ finden, sodass $n_0 V$ nicht nirgends dicht ist, d.h. es ist

$$\emptyset \neq \operatorname{int}(\overline{n_0 V}) = \operatorname{int}(n_0 V),$$

wobei wir für die letzte Gleichheit die Abgeschlossenheit von $n_0 V$ benutzt haben. Somit existiert ein $y \in \operatorname{int}(n_0 V)$, also ist $B_R(y) \subseteq n_0 V$ bzw. $\frac{1}{n_0} B_R(y) \subseteq V$ für ein geeignetes $R > 0$. So ergibt sich erneut mit der Stetigkeit der Skalarmultiplikation auf $\mathcal{X}$, dass $\frac{1}{n_0} B_R(y)$ offen und somit $z := \frac{y}{n_0} \in \operatorname{int}(V)$ ist. Wegen der Gleichheit (1.7) ist somit $z \in \operatorname{int}(-V)$ und mit derselben Argumentation wie oben ist dies äquivalent zu $-z \in \operatorname{int}(V)$. Insgesamt ist daher

$$\pm z \in \operatorname{int}(V). \tag{1.9}$$

Daraus können wir dann

$$0 \underset{(+)}{\in} \operatorname{int}\left(\frac{1}{2}V + \frac{1}{2}V\right) = \operatorname{int}(V) \tag{1.10}$$

folgern: Dabei erhalten wir die letzte Gleichheit direkt aus der Konvexität von V, sodass nur noch $(+)$ zu zeigen bleibt. Dazu wählen wir ein geeignetes $\mu > 0$ mit

$$B_\mu(z) \subseteq V \quad \text{und} \quad B_\mu(-z) \subseteq V,$$

das nach (1.9) existiert, und ein beliebiges $w \in B_\mu(0)$. So ist offenbar $w \pm z \in B_\mu(\pm z)$ aufgrund der Translationsinvarianz von d und damit

$$w = \frac{1}{2}(w+z) + \frac{1}{2}(w-z) \in \frac{1}{2}B_\mu(z) + \frac{1}{2}B_\mu(-z) \subseteq \frac{1}{2}V + \frac{1}{2}V.$$

Da w beliebig gewählt war, ist daher $B_\mu(0) \subseteq \frac{1}{2}V + \frac{1}{2}V$ bzw. $0 \in \operatorname{int}\left(\frac{1}{2}V + \frac{1}{2}V\right)$.

Nun existiert nach (1.10) eine Umgebung U von $x_0 = 0$ mit $U \subseteq V$, das nach Definition von V insbesondere bedeutet, dass $\mathcal{F}(x) \leq C$ für alle $x \in U$ ist. □

Motiviert von den Dualitätsbeziehungen der Abweichungsmaße, die wir in Abschnitt 1.2 genau untersuchen werden, erweiterten Rockafellar et al., die Definition (1.1.1) um ein spezielles Axiom, das bei weitem nicht alle Abweichungsmaße erfüllen, siehe dazu [12, Definition 2].

(1.1.8) Definition (Lower range dominance)
Ein Abweichungsmaß $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ heißt lower range dominated, *falls es der Eigenschaft*

$(\mathcal{D}5)$ $\mathcal{D}(X) \leq EX - \operatorname{ess\,inf} X$ *für alle* $X \in L^2(\Omega)$

genügt. ◇

(1.1.9) Bemerkung
Offensichtlich bedeutet das Axiom $(\mathcal{D}5)$, dass ein Abweichungsmaß $\mathcal{D}$, für das dieses zutrifft, in besonderer Weise von der Abweichung der Zufallsvariablen $X \in L^2(\Omega)$ nach unten im Verhältnis zu ihrem Erwartungswert abhängt. Besonders nützlich wird diese Eigenschaft für das Verständnis der sogenannten „risk envelopes“, die wir in Abschnitt 1.2 einführen werden. ◇

Nun wollen wir die bisher vorgestellten Beispiele erneut aufgreifen und herausarbeiten, ob sie lower range dominated sind.

(1.1.10) Beispiele

a) Die Abweichungsmaße σ_- aus Beispiel (1.1.3) b) und ψ_- aus Beispiel (1.1.3) c) sind stets lower range dominated: Während dies für ψ_- offensichtlich ist, müssen wir für σ_- unterscheiden, ob das essentielle Infimum einer gegebenen Zufallsvariable $X \in L^2(\Omega)$ endlich ist: Für $\operatorname{ess\,inf} X = -\infty$ ist die in $(\mathcal{D}5)$ geforderte Ungleichung offenbar erfüllt; falls hingegen $\operatorname{ess\,inf} X \in \mathbb{R}$ ist, folgt aus

$$0 \leq [X - EX]_- = \max\{0, EX - X\} \leq EX - \operatorname{ess\,inf} X$$

und den Monotonie-Eigenschaften von $\|\cdot\|_2$, dass

$$\sigma_-(X) = \|[X-EX]_-\|_2 \leq \|EX - \text{ess inf } X\|_2 = EX - \text{ess inf } X$$

und somit $(\mathcal{D}5)$ gilt.

b) Die verbleibenden Beispiele σ, σ_+ und ψ_+ aus Beispiel (1.1.3) sind allesamt i.A. nicht lower range dominated. Dazu nehmen wir erneut an, dass eine Menge $M \in \mathcal{M}$ mit $\mathcal{P}(M) \in (0, \frac{1}{2})$ existiert. (In den übrigen Fällen kann man leicht zeigen, dass sie lower range dominated sind, aber solche Wahrscheinlichkeitsräume sind für uns nicht so interessant.) Dann wissen wir bereits über die Zufallsvariable

$$X : \Omega \to \mathbb{R}, \quad X(\omega) := \frac{1}{\mathcal{P}(M)} \cdot \mathbb{1}_M(\omega),$$

aus (1.2), dass $EX = 1$ und ess inf $X = 0$ gilt. Außerdem ist

$$\text{ess sup } X = \frac{1}{\mathcal{P}(M)} > 2,$$

sodass wir

$$\psi_+(X) = \frac{1}{\mathcal{P}(M)} - 1 > 1 = EX - \text{ess inf } X$$

erhalten, d.h. ψ_+ ist in dieser Situation nicht lower range dominated. Um dies auch für σ_+ und σ zu zeigen, müssen wir zusätzlich von $(\Omega, \mathcal{M}, \mathcal{P})$ fordern, dass wir $\mathcal{P}(M)$ für eine Menge $M \in \mathcal{M}$ so klein wählen können, dass sogar

$$\frac{\big(1-\mathcal{P}(M)\big)^2}{\mathcal{P}(M)} > 1 \tag{1.11}$$

oder

$$\big(1-\mathcal{P}(M)\big)^2 - \mathcal{P}(M) > 0 \qquad \text{bzw.} \qquad 0 < \mathcal{P}(M) < \frac{3-\sqrt{5}}{2}$$

ist. Der Wahrscheinlichkeitsraum genügt dieser Anforderung, falls er z.B. essentiell unendlich ist; dazu benutze man Satz (A.1.20). Jedenfalls folgt daraus für die obige Zufallsvariable X zusammen mit Gleichung (1.3), dass

$$\sigma_+(X) = \sqrt{\frac{\big(1-\mathcal{P}(M)\big)^2}{\mathcal{P}(M)}} \underset{(1.11)}{>} 1 = EX - \text{ess inf } X$$

ist. Ferner ergibt sich wegen

$$0 \leq [X-EX]_+ \leq |X-EX|$$

und der Monotonie-Eigenschaften von $\|\cdot\|_2$, dass

$$\sigma(X) = \|X-EX\|_2 = \||X-EX|\|_2 \geq \|[X-EX]_+\|_2 = \sigma_+(X)$$

gilt, sodass sowohl σ als auch σ_+ nicht lower range dominated sind. ◇

Schließlich stellen wir noch zwei für uns wichtige Kombinationsmöglichkeiten von Abweichungsmaßen vor, die ROCKAFELLAR ET AL. jedoch in einem anderen Kontext in [12, Proposition 4] nebenbei behandelten. Dabei ist die Gültigkeit der jeweiligen Axiome eine elementare Konsequenz aus den Definitionen.

(1.1.11) Lemma
Seien $\mathcal{D}_1, \ldots, \mathcal{D}_n : L^2(\Omega) \to [0, \infty]$ Abweichungsmaße. Dann gilt:

(i) Das Funktional

$$\sum_{i=1}^{n} \lambda_i \mathcal{D}_i : L^2(\Omega) \to [0, \infty]$$

ist für $\lambda_1, \ldots, \lambda_n > 0$ ebenfalls ein Abweichungsmaß, das lower range dominated ist, falls es alle $\mathcal{D}_1, \ldots, \mathcal{D}_n$ sind und obendrein $\sum_{i=1}^{n} \lambda_i \leq 1$ gilt.

(ii) Das Funktional

$$max\{\mathcal{D}_1, \ldots, \mathcal{D}_n\} : L^2(\Omega) \to [0, \infty]$$

ist ein Abweichungsmaß, das genau dann lower range dominated ist, falls es alle $\mathcal{D}_1, \ldots, \mathcal{D}_n$ sind. ◇

(1.1.12) Bemerkung
Zu einem Abweichungsmaß $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ können wir neben der schon bekannten Reflexion $\tilde{\mathcal{D}}$, die wir in (1.1) definiert haben, die **Symmetrisierung**

$$\hat{\mathcal{D}} : L^2(\Omega) \to [0, \infty], \quad \hat{\mathcal{D}}(X) := \frac{1}{2}\big(\mathcal{D}(X) + \tilde{\mathcal{D}}(X)\big)$$

von $\mathcal{D}$ definieren, die ROCKAFELLAR ET AL. ebenfalls im Anschluss an [12, Definition 1] einführten. Nach Lemma (1.1.11) ist diese wiederum ein Abweichungsmaß, das - wie der Name schon andeutet - symmetrisch ist; das kann man nämlich unmittelbar anhand der Definition erkennen. ◇

Eine ähnliche Verknüpfung diskutieren wir in folgendem

(1.1.13) Beispiel
Im Zusammenhang mit ψ_- und ψ_+ aus Beispiel (1.1.3) c), die die maximalen Abweichungen vom Erwatungswert nach unten bzw. oben bestimmen, liegt der Gedanke nahe, auch die maximale Spannweite einer Zufallsvariablen zu messen, das das Funktional

$$\psi : L^2(\Omega) \to [0, \infty], \quad \psi(X) := \text{ess sup } X - \text{ess inf } X$$

realisiert. Sofort fällt uns ins Auge, dass wir

$$\begin{aligned} \psi(X) = \text{ess sup } X - \text{ess inf } X &= (\text{ess sup } X - EX) + (EX - \text{ess inf } X) \\ &= \psi_+(X) + \psi_-(X) \end{aligned} \tag{1.12}$$

für alle $X \in L^2(\Omega)$ schreiben können. Daher ist ψ nach Lemma (1.1.11) ebenfalls ein Abweichungsmaß, für das wiederum dieselben Aussagen bzgl. der Endlichkeit wie in Beispiel (1.1.3) c) zutreffen.
Jedoch ist ψ i.A. nicht lower range dominated; dazu reicht es anzunehmen, dass eine Menge $M \in \mathcal{M}$ mit $\mathcal{P}(M) \in (0,1)$ existiert. So erhält man unter Verwendung Zufallsvariable

$$X : \Omega \to \mathbb{R}, \quad X(\omega) := \frac{1}{\mathcal{P}(M)} \cdot \mathbb{1}_M(\omega)$$

wegen $EX = 1$, ess inf $X = 0$ und ess sup $X = \frac{1}{\mathcal{P}(M)}$ offensichtlich die Ungleichung

$$\psi(X) = \text{ess sup } X - \text{ess inf } X = \frac{1}{\mathcal{P}(M)} > 1 = EX - \text{ess inf } X,$$

sodass ψ in dieser Situation nicht $(\mathcal{D}5)$ erfüllt.
Um Aussagen über (Unterhalb-)Stetigkeit zu treffen, unterscheiden wir erneut Fälle bzgl. des Wahrscheinlichkeitsraumes: Falls $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich ist, ist ψ als Summe von ψ_+ und ψ_-, die bekanntlich in diesem Fall stetig sind, auch stetig. Ansonsten erhalten wir aus der Unterhalbstetigkeit von ψ_+ und ψ_-, dass nach Lemma (A.2.9) (i) genauso ψ unterhalbstetig ist. Wir müssen aber wieder feststellen, dass ψ in diesem Fall tatsächlich unstetig ist; dazu verweisen wir auf die gewählte Beispielfolge $\left(X_n^{(2,\infty)}\right)_{n\in\mathbb{N}}$ aus (A.7), die wir auch im Beweis des Zusatzes von Lemma (A.2.5) verwendet haben.
Im Gegensatz zu ψ_+ und ψ_- ist ψ symmetrisch; denn man erinnere sich an die Identität (1.12) und daran, dass ψ_+ gerade die Reflexion von ψ_- ist, das wir auch in Gleichung (1.5) nachgerechnet haben. $\diamond$

1.2 Dualität von Abweichungsmaßen

In dieser Sektion ist es unser wesentliches Ziel, eine bzw. die duale Charakterisierung von Abweichungsmaßen von Rockafellar et al. zu erarbeiten, die uns ein tieferes Verständnis von unterhalbstetigen Abweichungsmaßen ermöglicht. Als Grundlage wird der folgende zentrale Dualitätssatz benötigt. Als Inspiration für dessen Beweis dienten neben [12, Theorem 1] zudem die verschiedenen Ergebnisse aus [11, S. 112 - 114]. Dabei stellte sich heraus, dass die Verwendung der sogenannten „Fenchel-Konjugierten“ über Hilberträumen äußerst nützlich ist; die zugehörige Theorie ist in Abschnitt A.3 des Anhangs zu finden.

(1.2.1) Satz
Sei $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ ein Hilbertraum. Dann ist ein Funktional $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ genau dann unterhalbstetig, konvex und positiv homogen mit dom $\mathcal{F} \neq \emptyset$, falls eine nicht-leere, konvexe und abgeschlossene Menge $\mathcal{Y} \subseteq \mathcal{X}$ existiert, sodass

$$\mathcal{F}(x) = \mathop{sup}_{y\in\mathcal{Y}} \langle x, y\rangle$$

für alle $x \in \mathcal{X}$ geschrieben werden kann. In dieser Situation ist $\mathcal{Y}$ eindeutig und es gilt

$$\mathcal{Y} = \{y \in \mathcal{X} \mid \mathcal{F}(x) \geq \langle x, y \rangle \quad \textit{für alle } x \in \mathcal{X}\}. \tag{1.13}$$

Zudem ist $\mathcal{F}$ genau dann endlich (und daher stetig), falls $\mathcal{Y}$ beschränkt ist. ◇

Beweis
Wir zeigen zuerst die behauptete Äquivalenz, wobei sich dann herausstellen wird, dass wir in dieser Situation $\mathcal{Y}$ wie in (1.13) wählen können. Anschließend widmen wir uns der Eindeutigkeit sowie dem Zusatz.

Äquivalenz: „$\Rightarrow$" Sei $\mathcal{F}$ also unterhalbstetig, konvex und positiv homogen mit dom $\mathcal{F} \neq \emptyset$. Dann betrachten wir die zugehörige Fenchel-Konjugierte

$$\mathcal{F}^* : \mathcal{X} \to [-\infty, \infty], \quad \mathcal{F}^*(x) := \sup_{y \in \mathcal{X}} \{\langle x, y \rangle - \mathcal{F}(y)\}$$

von $\mathcal{F}$ aus Definition (A.3.1). Nach Satz (A.3.4) gilt somit:

(i) Es ist $\mathcal{F}^{**} = \mathcal{F}$.

(ii) Die Fenchel-Konjugierte $\mathcal{F}^*$ besitzt die Eigenschaft, dass $\mathcal{F}^* > -\infty$ und dom $\mathcal{F}^*$ nicht-leer ist. Darüber hinaus ist sie konvex und unterhalbstetig.

Insbesondere ist $\mathcal{F}^{**} = \mathcal{F}$ damit positiv homogen, sodass wir für $\lambda \in (0,1)$ und $x \in \mathcal{X}$ die Gleichungskette

$$\begin{aligned} \mathcal{F}(x) = \mathcal{F}^{**}(x) = \lambda \mathcal{F}^{**}(\lambda^{-1} x) &= \lambda \cdot \sup_{y \in \mathcal{X}} \{\langle \lambda^{-1} x, y \rangle - \mathcal{F}^*(y)\} \\ &= \sup_{y \in \mathcal{X}} \{\langle x, y \rangle - \lambda \mathcal{F}^*(y)\} \\ &= (\lambda \mathcal{F}^*)^*(x), \end{aligned} \tag{1.14}$$

d.h. $\mathcal{F} = (\lambda \mathcal{F}^*)^*$, wobei wir mehrmals die Definition der Fenchel- und Bikonjugierten verwendet haben. Da die Eigenschaften der Fenchel-Konjugierten $\mathcal{F}^*$ aus (ii) auch für $\lambda \mathcal{F}^*$ bestehen bleiben, können wir wiederum aus Satz (A.3.4) (i)

$$(\lambda \mathcal{F}^*)^{**} = \lambda \mathcal{F}^*$$

folgern. In Kombination mit der Identität (1.14) ergibt dies

$$\mathcal{F}^*(x) = \left((\lambda \mathcal{F}^*)^*\right)^*(x) = (\lambda \mathcal{F}^*)^{**}(x) = \lambda \mathcal{F}^*(x)$$

für alle $x \in \mathcal{X}$. Wegen $\mathcal{F}^*(\mathcal{X}) \subseteq (-\infty, \infty]$ nach (ii) und der Wahl von λ ist dies nur möglich, falls

$$\mathcal{F}^*(\mathcal{X}) \subseteq \{0, \infty\} \tag{1.15}$$

ist. Genauer gesagt erhalten wir

$$\mathcal{F}^*(x) = \begin{cases} 0 & \text{für } x \in \mathcal{Y}, \\ \infty & \text{für } x \notin \mathcal{Y} \end{cases} \tag{1.16}$$

mit einer geeignet zu wählenden Menge $\mathcal{Y} \subseteq \mathcal{X}$. Für unsere Zwecke ist die Wahl

$$\mathcal{Y} := \{y \in \mathcal{X} \mid \mathcal{F}^*(y) \leq 0\}$$

sinnvoll, da die Charakteristika von $\mathcal{F}^*$ in (ii) Folgendes für $\mathcal{Y}$ implizieren:

a) $\mathcal{Y}$ ist nicht-leer, weil dom $\mathcal{F}^*$ nicht-leer ist und daher nach (1.15) ein $y \in \mathcal{X}$ mit $\mathcal{F}^*(y) = 0$ existieren muss.
b) $\mathcal{Y}$ ist konvex, weil $\mathcal{F}^*$ konvex ist.
c) $\mathcal{Y}$ ist abgeschlossen, das sich aus der Unterhalbstetigkeit von $\mathcal{F}^*$ sowie Lemma (A.2.3) ergibt.

Nun ist für jedes $y \in \mathcal{X}$ genau dann $y \in \mathcal{Y}$, falls

$$0 \geq \mathcal{F}^*(y) = \sup_{x \in \mathcal{X}} \{\langle x, y \rangle - \mathcal{F}(x)\}$$

ist, das jedoch nichts anderes als

$$0 \geq \langle x, y \rangle - \mathcal{F}(x) \qquad \text{bzw.} \qquad \mathcal{F}(x) \geq \langle x, y \rangle$$

für alle $x \in \mathcal{X}$ bedeutet. Daher können wir

$$\mathcal{Y} = \{y \in \mathcal{X} \mid \mathcal{F}(x) \geq \langle x, y \rangle \quad \text{für alle } x \in \mathcal{X}\}$$

schreiben, das gerade mit der Darstellung (1.13) übereinstimmt. Ferner stellen wir fest, dass

$$\mathcal{F}(x) = \mathcal{F}^{**}(x) = \sup_{y \in \mathcal{X}} \{\langle x, y \rangle - \mathcal{F}^*(y)\} \underset{(1.16)}{=} \sup_{y \in \mathcal{Y}} \{\langle x, y \rangle - \underbrace{\mathcal{F}^*(y)}_{=\,0}\} = \sup_{y \in \mathcal{Y}} \langle x, y \rangle$$

für alle $x \in \mathcal{X}$ erfüllt ist, sodass wir die erste Implikation gezeigt haben.

„ $\Leftarrow$ " Es sei eine nicht-leere, konvexe und abgeschlossene Teilmenge $\mathcal{Y} \subseteq \mathcal{X}$ gegeben, wozu wir das Funktional

$$\mathcal{F} : \mathcal{X} \to (-\infty, \infty], \quad \mathcal{F}(x) := \sup_{y \in \mathcal{Y}} \langle x, y \rangle$$

definieren. Nun müsen wir die behaupteten Eigenschaften von $\mathcal{F}$ zeigen. Dabei folgen sowohl die Konvexität als auch die positive Homogenität aus der Bilinearität des Skalarproduktes und den Supremumseigenschaften. Obendrein ist dom $\mathcal{F} \neq \emptyset$, da z.B. $\mathcal{F}(0) = 0 < \infty$ ist.

Schließlich bleibt die Unterhalbstetigkeit von $\mathcal{F}$ zu zeigen: Dazu sei $C \in \mathbb{R}$, $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{V}_C(\mathcal{F})$ aus Lemma (A.2.3) und $x \in \mathcal{X}$ mit $x_n \xrightarrow[n\to\infty]{} x$. Nach Annahme ist zum Einen

$$\langle x_n, y\rangle \leq \mathcal{F}(x_n) \leq C$$

für alle $n \in \mathbb{N}$ und $y \in \mathcal{Y}$ und zum Anderen aufgrund der Cauchy-Schwarz-Ungleichung

$$\langle x_n, y\rangle \xrightarrow[n\to\infty]{} \langle x, y\rangle$$

für alle $y \in \mathcal{Y}$. Daher ist auch $\langle x, y\rangle \leq C$ für alle $y \in \mathcal{Y}$ und somit $\mathcal{F}(x) \leq C$ bzw. $x \in \mathcal{V}_C(\mathcal{F})$, d.h. $\mathcal{V}_C(\mathcal{F})$ ist abgeschlossen. Damit liefert Lemma (A.2.3) die Unterhalbstetigkeit von $\mathcal{F}$.

Eindeutigkeit: Seien $\mathcal{Y}_1 \subseteq \mathcal{X}$ und $\mathcal{Y}_2 \subseteq \mathcal{X}$ nicht-leere, konvexe und abgeschlossene Mengen, sodass

$$\sup_{y\in\mathcal{Y}_1} \langle x, y\rangle = \sup_{y\in\mathcal{Y}_2} \langle x, y\rangle \tag{1.17}$$

für alle $x \in \mathcal{X}$ ist. Dann muss schon $\mathcal{Y}_1 = \mathcal{Y}_2$ gelten; denn andernfalls sei Œ $\mathcal{Y}_1 \subsetneq \mathcal{Y}_2$, d.h. ex existiert ein $x_0 \in \mathcal{Y}_2 \setminus \mathcal{Y}_1$. So existiert nach dem Satz von Mazur über Hilberträumen, siehe Satz (A.3.7), ein $z \in \mathcal{X}$ und ein $C \in \mathbb{R}$ mit

$$\langle z, x_0\rangle > C \quad \text{und} \quad \langle z, y\rangle \leq C \quad \text{für alle } y \in \mathcal{Y}_1.$$

Damit ergibt sich

$$\sup_{y\in\mathcal{Y}_2} \langle z, y\rangle \geq \langle z, x_0\rangle > C \geq \sup_{y\in\mathcal{Y}_1} \langle z, y\rangle,$$

also ein Widerspruch zur Gleichung (1.17).

Zusatz: Zunächst bemerken wir, dass in der Situation der Äquivalenz aus der Endlichkeit von $\mathcal{F}$ sogar direkt dessen Stetigkeit folgt, da der Satz (1.1.7) anwendbar ist.

„ $\Rightarrow$ “ Sei $\mathcal{F}$ also zusätzlich endlich. Dann betrachten wir zu jedem $y \in \mathcal{Y}$ das lineare stetige Funktional

$$\mathcal{J}_y : \mathcal{X} \to \mathbb{R}, \quad \mathcal{J}_y(x) := \langle x, y\rangle.$$

Nach Definition ist

$$\mathcal{F}(x) = \sup_{y\in\mathcal{Y}} \langle x, y\rangle = \sup_{y\in\mathcal{Y}} \mathcal{J}_y(x) \tag{1.18}$$

für alle $x \in \mathcal{X}$. Zudem gilt

$$\left|\inf_{y\in\mathcal{Y}} \mathcal{J}_y(x)\right| = \left|\inf_{y\in\mathcal{Y}} \langle x, y\rangle\right| = \left|-\inf_{y\in\mathcal{Y}} \langle x, y\rangle\right| = \left|\sup_{y\in\mathcal{Y}} \langle -x, y\rangle\right| = |\mathcal{F}(-x)|$$

für alle $x \in \mathcal{X}$, das wegen der Endlichkeit von $\mathcal{F}$ zusammen mit der Identität (1.18) gerade

$$\sup_{y\in\mathcal{Y}} |\mathcal{J}_y(x)| = \max\left\{\left|\sup_{y\in\mathcal{Y}} \mathcal{J}_y(x)\right|, \left|\inf_{y\in\mathcal{Y}} \mathcal{J}_y(x)\right|\right\} = \max\{|\mathcal{F}(x)|, |\mathcal{F}(-x)|\} < \infty$$

für alle $x \in \mathcal{X}$ liefert. Da X obendrein ein Hilbertraum ist, folgt mit dem Satz von der gleichmäßigen Beschränktheit, siehe [10, Satz (7.1.1)], dass wir ein $C > 0$ finden, sodass

$$C \geq \|\mathcal{J}_y\|_{\mathcal{X}'} = \|y\|_{\mathcal{X}}$$

für alle $y \in \mathcal{Y}$ erfüllt ist, wobei wir die letzte Gleichheit aufgrund des isometrischen Isomorphismus zwischen $\mathcal{X}$ und seinem Dualraum $\mathcal{X}'$ erhalten. Demnach ist $\mathcal{Y}$ beschränkt.

„ $\Leftarrow$ “ Sei die Menge $\mathcal{Y}$ nun außerdem beschränkt, d.h. es existiert ein $C > 0$ mit $\|y\|_{\mathcal{X}} \leq C$ für alle $y \in \mathcal{Y}$. In Kombination mit der Cauchy-Schwarz-Ungleichung impliziert dies

$$\mathcal{F}(x) = \sup_{y\in\mathcal{Y}} \langle x, y\rangle \leq \sup_{y\in\mathcal{Y}} \|x\|_{\mathcal{X}} \cdot \|y\|_{\mathcal{X}} \leq \|x\|_{\mathcal{X}} \cdot C < \infty$$

für jedes $x \in \mathcal{X}$, das gerade der Endlichkeit von $\mathcal{F}$ entspricht. □

Um diesen Satz erstmals anzuwenden, präsentieren wir jetzt ein ziemlich bekanntes Resultat, das Rockafellar et al. jedoch ohne Nachweis für den Beweis von [12, Proposition 7] heranzogen. Diese Proposition werden wir auch in Abschnitt 1.4 notieren.

(1.2.2) Korollar
Sei $q \in [1, \infty]$. Dann gilt für alle $X \in L^2(\Omega)$ die Identität

$$\|X\|_q = \mathop{sup}_{Y\in\mathcal{Y}_q} E[XY], \tag{1.19}$$

wobei

$$\mathcal{Y}_q := \{Y \in L^2(\Omega) \mid \|Y\|_p \leq 1\} \tag{1.20}$$

für $p \in [1, \infty]$ mit $\frac{1}{q} + \frac{1}{p} = 1$ sei. ◇

Beweis
Wir greifen das Funktional

$$\mathcal{G}_{2,q} : L^2(\Omega) \to [0, \infty], \quad \mathcal{G}_{2,q}(X) := \|X\|_q$$

aus Lemma (A.2.4) auf. Dort haben wir bereits gesehen, dass $\mathcal{G}_{2,q}$ zumindest immer unterhalbstetig ist. Ferner ist einerseits $\mathcal{G}_{2,q}$ konvex und positiv homogen und andererseits ist dom $\mathcal{G}_{2,q} \neq \emptyset$, weil z.B. $\mathcal{G}_{2,q}(0) = 0 < \infty$ ist. Daher liefert Satz (1.2.1), dass eine nicht-leere, konvexe und abgeschlossene Menge $\mathcal{Y}_q \subseteq L^2(\Omega)$ existiert, sodass

$$\|X\|_q = \mathcal{G}_{2,q}(X) = \sup_{Y \in \mathcal{Y}_q} E[XY]$$

für alle $X \in L^2(\Omega)$ gilt. Dabei ist $\mathcal{Y}_q$ nach (1.13) eindeutig durch

$$\mathcal{Y}_q = \{Y \in L^2(\Omega) \mid \|X\|_q = \mathcal{G}_{2,q}(X) \geq E[XY] \quad \text{für alle } X \in L^2(\Omega)\}$$

bestimmt. Somit müssen wir nur noch zeigen, dass wir

$$\mathcal{Y}_q = \{Y \in L^2(\Omega) \mid \|Y\|_p \leq 1\}$$

für $p \in [1, \infty]$ mit $\frac{1}{q} + \frac{1}{p} = 1$ schreiben können.

Dazu: „ $\supseteq$ “ Sei $Y \in L^2(\Omega)$ mit $\|Y\|_p \leq 1$. Dann ist für jedes $X \in L^2(\Omega)$ aufgrund der Hölder-Ungleichung

$$E[XY] \leq |E[XY]| \leq E[|XY|] \leq \|X\|_q \cdot \|Y\|_p \leq \|X\|_q,$$

also $Y \in \mathcal{Y}_q$.

„ $\subseteq$ “ Sei $Y \in \mathcal{Y}_q$. Wir nehmen

$$\|Y\|_p > 1 \tag{1.21}$$

an und führen dies zu einem Widerspruch. Zur Vereinfachung unterscheiden wir drei Fälle bzgl. q:

1. Fall: $q \in (1, \infty)$
Demnach ist auch $p \in (1, \infty)$. So definieren wir die Zufallsvariablen

$$Y_n : \Omega \to \mathbb{R}, \quad Y_n(\omega) := \operatorname{sign}\big(Y(\omega)\big) \cdot \min\{|Y(\omega)|, n\}$$

und

$$Z_n : \Omega \to \mathbb{R}, \quad Z_n(\omega) := \operatorname{sign}\big(Y(\omega)\big) \cdot |Y_n(\omega)|^{p-1}$$

für $n \in \mathbb{N}$. Nach Definition ist also $|Y_n| \leq n$ und daher $|Z_n| \leq n^{p-1}$ für alle $n \in \mathbb{N}$, das $\|Y_n\|_2 \leq n$ und $\|Z_n\|_2 \leq n^{p-1}$ für alle $n \in \mathbb{N}$ impliziert, d.h. wir erhalten $(Y_n)_{n\in\mathbb{N}} \subseteq L^2(\Omega)$ und $(Z_n)_{n\in\mathbb{N}} \subseteq L^2(\Omega)$. Darüber hinaus gilt offenbar

$$|Y_n| = \min\{|Y|, n\} \xrightarrow[n\to\infty]{pktw.} |Y|$$

und somit

$$|Z_n| = |Y_n|^{p-1} \xrightarrow[n\to\infty]{pktw.} |Y|^{p-1}.$$

Daraus folgt zum Einen

$$|Z_n|^q \xrightarrow[n\to\infty]{pktw.} |Y|^{q(p-1)} = |Y|^p,$$

da $q = \frac{p}{p-1}$ ist, und zum Anderen

$$Y_n \cdot Z_n = |Y_n| \cdot |Z_n| \xrightarrow[n\to\infty]{pktw.} |Y| \cdot |Y|^{p-1} = |Y|^p,$$

wobei sich die erste Gleichheit daraus ergibt, dass $Y_n(\omega)$ und $Z_n(\omega)$ für alle $\omega \in \Omega$ und $n \in \mathbb{N}$ nach Konstruktion dasselbe Vorzeichen besitzen. Genauso impliziert

$$0 \leq |Y_n| = \min\{|Y|, n\} \leq \min\{|Y|, n+1\} = |Y_{n+1}|$$

für alle $n \in \mathbb{N}$ die Ungleichungen

$$0 \leq |Z_n|^q \leq |Z_{n+1}|^q$$

sowie

$$0 \leq Y_n \cdot Z_n = |Y_n| \cdot |Z_n| \leq |Y_{n+1}| \cdot |Z_{n+1}| = Y_{n+1} \cdot Z_{n+1}$$

für alle $n \in \mathbb{N}$. Dann liefert der Satz von der monotonen Konvergenz

$$\|Z_n\|_q^q = E[|Z_n|^q] \xrightarrow[n\to\infty]{} E[|Y|^p],$$

sodass wir ebenfalls

$$\|Z_n\|_q \xrightarrow[n\to\infty]{} \left(E[|Y|^p]\right)^{\frac{1}{q}}, \tag{1.22}$$

erhalten, sowie

$$E[Y_n Z_n] \xrightarrow[n\to\infty]{} E[|Y|^p]. \tag{1.23}$$

Aufgrund unserer ursprünglichen Annahme (1.21) ist nun sowohl

$$E[|Y|^p] > 1 \qquad \text{als auch} \qquad \left(E[|Y|^p]\right)^{\frac{1}{q}} > 1,$$

sodass uns die Konvergenzen (1.22) und (1.23) die Wahl eines $N \in \mathbb{N}$ mit

$$\infty > \|Z_N\|_q > 1 \qquad \text{und} \qquad \infty > E[Y_N Z_N] > 1 \tag{1.24}$$

erlauben, wobei die beiden Terme deshalb endlich sind, da Y_N und Z_N bekanntermaßen beschränkt sind. Jedoch gilt

$$\|Z_N\|_q \geq E[Z_N Y]$$

wegen $Y \in \mathcal{Y}_q$ und $Z_N \in L^2(\Omega)$, woraus sich

$$\|Z_N\|_q \geq E[Z_N Y] = E[|Z_N| \cdot |Y|] \geq E[|Z_N| \cdot |Y_N|] = E[Y_N Z_N].$$

ergibt, da Y und Z_N ebenfalls punktweise dasselbe Vorzeichen besitzen. Mittels erneuter Ausnutzung der Definitionen von Y_N und Z_N erhalten wir schließlich

$$\big(E[|Y_N|^p]\big)^{\frac{1}{q}} = \|Z_N\|_q \geq E[Y_N Z_N] = E[|Y_N|^p],$$

das uns zusammen mit (1.24) den gewünschten Widerspruch beschert; denn wegen $q \in (1, \infty)$ ist $\frac{1}{q} \in (0, 1)$, sodass

$$x^{\frac{1}{q}} \geq x \quad \text{für alle } x \in [0,1] \qquad \text{und} \qquad x^{\frac{1}{q}} < x \quad \text{für alle } x > 1$$

gilt.

<u>*2. Fall:*</u> $q = 1$
Demnach ist $p = \infty$. Nach der Annahme (1.21) ist also

$$1 < \|Y\|_\infty = \operatorname{ess\,sup} |Y|.$$

Nun sei $C \in \mathbb{R}$ mit $1 < C < \operatorname{ess\,sup} |Y|$, sodass nach Definition des essentiellen Supremums $\mathcal{P}(\{|Y| \geq C\}) > 0$ sein muss. So betrachten wir die Zufallsvariable

$$X : \Omega \to \mathbb{R}, \quad X(\omega) := \operatorname{sign}\big(Y(\omega)\big) \cdot \mathbb{1}_{\{|Y| \geq C\}}(\omega).$$

Da offenbar $X \in L^2(\Omega)$ sowie $Y \in \mathcal{Y}_q$ gilt, folgt wegen

$$\mathcal{P}(\{|Y| \geq C\}) = \|X\|_1 \geq E[XY] = \int\limits_{\{|Y| \geq C\}} |Y| \, \mathrm{d}\mathcal{P} \geq C \cdot \mathcal{P}(\{|Y| \geq C\}),$$

also $C \leq 1$, ein Widerspruch; denn wir haben $C > 1$ gewählt.

<u>*3. Fall:*</u> $q = \infty$
Demnach ist $p = 1$. Nach der Annahme (1.21) gilt wiederum $\|Y\|_1 > 1$. Daraus folgt jedoch, dass auch

$$\|Y\|_\infty = \operatorname{ess\,sup} |Y| > 1$$

gelten muss; denn andernfalls erhalten wir wegen $\mathcal{P}(\{|Y| > \operatorname{ess\,sup} |Y|\}) = 0$ gerade

$$\|Y\|_1 = E[|Y|] \leq E[\operatorname{ess\,sup} |Y|] = \operatorname{ess\,sup} |Y| \leq 1,$$

dessen Gegenteil wir aber angenommen haben.
Jedenfalls betrachten wir jetzt die Zufallsvariable

$$X : \Omega \to \mathbb{R}, \quad X(\omega) := \operatorname{sign}\big(Y(\omega)\big).$$

Es gilt offensichtlich $X \in L^2(\Omega)$ und wegen $\|Y\|_\infty > 1 > 0$ muss außerdem $\|X\|_\infty = 1$ gelten. Dann ergibt sich mit erneuter Ausnutzung von $Y \in \mathcal{Y}_q$, dass

$$1 = \|X\|_\infty \geq E[XY] = E[|Y|] = \|Y\|_1 > 1$$

ist, das aber nicht möglich ist.

Damit ist das Korollar bewiesen. □

Wie angekündigt, können wir mithilfe von Satz (1.2.1) unterhalbstetige Abweichungsmaße charakterisieren. Dies beruht hauptsächlich auf [12, Theorem 1].

(1.2.3) Satz (Duale Charakterisierung von Abweichungsmaßen)
Ein Funktional $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ ist genau dann ein unterhalbstetiges Abweichungsmaß, wenn eine Menge $\mathcal{Q} \subseteq L^2(\Omega)$ existiert, sodass

$$\mathcal{D}(X) = EX - \inf_{Q \in \mathcal{Q}} E[XQ] \tag{1.25}$$

für alle $X \in L^2(\Omega)$ geschrieben werden kann und $\mathcal{Q}$ folgenden Eigenschaften genügt:

- $(\mathcal{Q}1)$ *$\mathcal{Q}$ ist nicht-leer, konvex und abgeschlossen.*
- $(\mathcal{Q}2)$ *Für jede nicht-konstante Zufallsvariable $X \in L^2(\Omega)$ existiert eine Zufallsvariable $Q \in \mathcal{Q}$ mit $EX > E[XQ]$.*
- $(\mathcal{Q}3)$ *Für jedes $Q \in \mathcal{Q}$ ist $EQ = 1$.*

In dieser Situation ist $\mathcal{Q}$ bereits durch $(\mathcal{Q}1)$ sowie der Identität (1.25) eindeutig bestimmt und es gilt

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid \mathcal{D}(X) \geq EX - E[XQ] \quad \text{für alle } X \in L^2(\Omega)\}. \tag{1.26}$$

Zudem gilt:

(i) $\mathcal{D}$ ist genau dann lower range dominated, falls $\mathcal{Q}$ die Anforderung

$(\mathcal{Q}4)$ $Q \geq 0 \quad$ *für alle $Q \in \mathcal{Q}$*

erfüllt.

(ii) $\mathcal{D}$ ist genau dann endlich (und daher stetig), falls $\mathcal{Q}$ beschränkt ist.

(iii) $\mathcal{D}$ ist genau dann symmetrisch, falls $1 - \mathcal{Q}$ kreisförmig ist. ◇

Beweis
Zuallererst stellen wir fest, dass ein Funktional $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ genau dann positiv homogen und subadditiv ist, falls es positiv homogen und konvex ist; während die Hinrichtung nämlich direkt wie in Bemerkung (1.1.2) b) folgt, bleibt für die Rückrichtung nur die Subadditivität zu zeigen: Seien $X, Y \in L^2(\Omega)$, so erhalten wir

$$\mathcal{D}(X+Y) = \mathcal{D}\big(\frac{1}{2}(2X+2Y)\big) \leq \frac{1}{2}\mathcal{D}(2X) + \frac{1}{2}\mathcal{D}(2Y) = \mathcal{D}(X) + \mathcal{D}(Y),$$

d.h. $\mathcal{D}$ ist subadditiv.

Zusammen mit Satz (1.2.1) ist ein Funktional $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ demnach genau dann unterhalbstetig und erfüllt $(\mathcal{D}2)$ und $(\mathcal{D}3)$, wenn eine nicht-leere, konvexe und abgeschlossene Menge $\mathcal{Y} \subseteq L^2(\Omega)$ existiert, sodass $\mathcal{D}$ eine Darstellung der Form

$$\mathcal{D}(X) = \sup_{Y \in \mathcal{Y}} E[XY] \tag{1.27}$$

für alle $X \in L^2(\Omega)$ besitzt. Dabei ist $\mathcal{Y}$ eindeutig durch

$$\mathcal{Y} = \{Y \in L^2(\Omega) \mid \mathcal{D}(X) \geq E[XY] \quad \text{für alle } X \in L^2(\Omega)\} \tag{1.28}$$

bestimmt. Dazu muss man beachten, dass $\operatorname{dom} \mathcal{D} \neq \emptyset$ gilt, da $\mathcal{D}(0) = 0 < \infty$ wegen $(\mathcal{D}2)$ erfüllt ist, und umgekehrt $\mathcal{D}(0) = 0$ aus der Darstellung (1.27) folgt.
In dieser Situation können wir aber auch

$$\begin{aligned} \mathcal{D}(X) = \sup_{Y\in\mathcal{Y}} E[XY] = \sup_{Y\in\mathcal{Y}} E\big[X\big(1-(1-Y)\big)\big] &= \sup_{Y\in\mathcal{Y}} \big\{EX - E[X(1-Y)]\big\} \\ &= \sup_{Q\in 1-\mathcal{Y}} \big\{EX - E[XQ]\big\} \\ &= EX - \inf_{Q\in 1-\mathcal{Y}} E[XQ] \end{aligned} \tag{1.29}$$

für alle $X \in L^2(\Omega)$ schreiben, wobei wir dann

$$\begin{aligned} \mathcal{Q} &:= 1 - \mathcal{Y} \\ &= \{1 - Y \mid Y \in L^2(\Omega) \quad \text{mit} \quad \mathcal{D}(X) \geq E[XY] \quad \text{für alle } X \in L^2(\Omega)\} \\ &= \{1 - Y \mid Y \in L^2(\Omega) \quad \text{mit} \quad \mathcal{D}(X) \geq E[X] - E[X(1-Y)] \quad \text{für alle } X \in L^2(\Omega)\} \\ &= \{Q \in L^2(\Omega) \mid \mathcal{D}(X) \geq E[X] - E[XQ] \quad \text{für alle } X \in L^2(\Omega)\} \end{aligned} \tag{1.30}$$

definieren. Hierbei ist $\mathcal{Q}$ nicht-leer, konvex und abgeschlossen, da dies schon auf $\mathcal{Y}$ zutrifft; die Umkehrung gilt aufgrund der Beziehung $\mathcal{Y} = 1 - \mathcal{Q}$ aber offensichtlich auch. Daher ist $\mathcal{Q}$ durch die Eigenschaft $(\mathcal{Q}1)$ sowie die Darstellung (1.29) eindeutig bestimmt und kann wie in (1.30) konkret angegeben werden.

Nun arbeiten wir (stets in obiger Situation) die zusätzlichen Merkmale von $\mathcal{Q}$ heraus, die dadurch zustande kommen, dass $\mathcal{D}$ obendrein $(\mathcal{D}1)$ oder $(\mathcal{D}4)$ erfüllt:

(1) *Behauptung:* $\mathcal{D}$ erfüllt genau dann $(\mathcal{D}1)$, falls $\mathcal{Q}$ der Eigenschaft $(\mathcal{Q}3)$ genügt.

Dazu: „$\Rightarrow$“ $\mathcal{D}$ erfülle also zusätzlich $(\mathcal{D}1)$. Da auch hier offenbar $(\mathcal{D}1')$ aus Lemma (1.1.4) gilt, folgt

$$0 = \mathcal{D}(C) = E[C] - \inf_{Q\in\mathcal{Q}} E[CQ] = C - \inf_{Q\in\mathcal{Q}} C \cdot EQ$$

für alle $C \in \mathbb{R}$. Dann liefert die Wahl von $C = \pm 1$ die Gleichungen

$$0 = 1 - \inf_{Q\in\mathcal{Q}} EQ \qquad \text{sowie} \qquad 0 = -1 - \inf_{Q\in\mathcal{Q}} (-EQ) = -\Big(1 - \sup_{Q\in\mathcal{Q}} EQ\Big),$$

wobei letztere äquivalent zu

$$0 = 1 - \sup_{Q\in\mathcal{Q}} EQ$$

ist. Somit erhalten wir

$$0 = 1 - \sup_{Q\in\mathcal{Q}} EQ \leq 1 - EV \leq 1 - \inf_{Q\in\mathcal{Q}} EQ = 0$$

und daher $EV = 1$ für alle $V \in \mathcal{Q}$, das $(\mathcal{Q}3)$ entspricht.

„ $\Leftarrow$ “ Jetzt erfülle $\mathcal{Q}$ zusätzlich $(\mathcal{Q}3)$. So folgt für beliebige $C \in \mathbb{R}$ und $X \in L^2(\Omega)$ die Gleichungskette

$$\begin{aligned}\mathcal{D}(X+C) = E[X+C] - \inf_{Q\in\mathcal{Q}} E[(X+C)Q] &= EX + C - \inf_{Q\in\mathcal{Q}} (E[XQ] + C \cdot \underbrace{EQ}_{=1}) \\ &= EX - \inf_{Q\in\mathcal{Q}} E[XQ] = \mathcal{D}(X),\end{aligned}$$

sodass $\mathcal{D}$ das Axiom $(\mathcal{D}1)$ erfüllt.

(2) *Behauptung:* $\mathcal{D}$ erfüllt genau dann $(\mathcal{D}4)$, falls $\mathcal{Q}$ der Eigenschaft $(\mathcal{Q}2)$ genügt.

Dazu: Sei $X \in L^2(\Omega)$ nicht-konstant. Dann gilt genau dann

$$0 < \mathcal{D}(X) = EX - \inf_{Q\in\mathcal{Q}} E[XQ],$$

falls ein $Q \in \mathcal{Q}$ mit $EX > E[XQ]$ existiert. Aber dies entspricht gerade der Behauptung.

Insgesamt können wir damit festhalten, dass $\mathcal{D}$ genau dann ein unterhalbstetiges Abweichungsmaß ist, falls eine Menge $\mathcal{Q} \subseteq L^2(\Omega)$ mit den Eigenschaften $(\mathcal{Q}1)$ bis $(\mathcal{Q}3)$ existiert, sodass $\mathcal{D}(X)$ die Darstellung (1.25) für alle $X \in L^2(\Omega)$ besitzt. Dabei ist $\mathcal{Q}$ aber schon durch $(\mathcal{Q}1)$ sowie der Identität (1.25) eindeutig bestimmt und kann wie in (1.26) angegeben werden.

Schließlich widmen wir uns den Zusätzen:
(i) Zunächst beobachten wir, dass $\mathcal{D}$ genau dann lower range dominated ist, falls

$$EX - \inf_{Q\in\mathcal{Q}} E[XQ] = \mathcal{D}(X) \leq EX - \text{ess inf } X$$

für alle $X \in L^2(\Omega)$ ist, was aber nichts anderes als

$$\inf_{Q\in\mathcal{Q}} E[XQ] \geq \text{ess inf } X$$

für alle $X \in L^2(\Omega)$ bedeutet. Nun unterscheiden wir zwei Fälle:

<u>1. Fall:</u> *$\mathcal{Q}$ erfüllt $(\mathcal{Q}4)$.*
Seien $X \in L^2(\Omega)$ und $Q \in \mathcal{Q}$ beliebig gegeben. Falls $\text{ess inf } X = -\infty$ gilt, so ist offenbar

$$E[XQ] > -\infty = \text{ess inf } X$$

aufgrund der Hölder-Ungleichung; andernfalls ist $\text{ess inf } X \in \mathbb{R}$, sodass wir ebenfalls

$$E[XQ] = \int_\Omega XQ \, d\mathcal{P} \underset{(\mathcal{Q}4)}{\geq} \int_\Omega (\text{ess inf } X) \cdot Q \, d\mathcal{P} = (\text{ess inf } X) \cdot \underbrace{EQ}_{\underset{(\mathcal{Q}3)}{=} 1} = \text{ess inf } X$$

erhalten, wobei wir für die Ungleichung wieder $\mathcal{P}(\{X < \text{ess inf } X\}) = 0$ benutzt haben. Da Q beliebig gewählt war, folgt somit

$$\inf_{Q \in \mathcal{Q}} E[XQ] \geq \text{ess inf } X$$

und, weil genauso X beliebig gewählt war, ist $\mathcal{D}$ aufgrund unserer obigen Beobachtung lower range dominated.

<u>2. Fall:</u> *$\mathcal{Q}$ erfüllt ($\mathcal{Q}4$) nicht.*
Demnach muss ein $Q_0 \in \mathcal{Q}$ mit $\text{ess inf } Q_0 < 0$ existieren. Daher können wir ein $C \in \mathbb{R}$ mit $\text{ess inf } Q_0 < C < 0$ wählen, für das $\mathcal{P}(\{Q_0 \leq C\}) > 0$ gilt. Dann definieren wir die Zufallsvariable

$$X_0 : \Omega \to \mathbb{R}, \quad X_0(\omega) := -\min\{Q_0(\omega), 0\},$$

für die offenbar $X_0 \in L^2(\Omega)$ gilt. Nach Konstruktion ist $X_0 \geq 0$, also insbesondere $\text{ess inf } X_0 \geq 0$, und somit ergibt sich

$$\begin{aligned} E[X_0 Q_0] = \int_\Omega X_0 Q_0 \, \mathrm{d}\mathcal{P} &= \int_{\{Q_0 \leq 0\}} X_0 Q_0 \, \mathrm{d}\mathcal{P} = \int_{\{Q_0 \leq 0\}} -(Q_0)^2 \, \mathrm{d}\mathcal{P} \\ &\leq \int_{\{Q_0 \leq C\}} -(Q_0)^2 \, \mathrm{d}\mathcal{P} \leq \int_{\{Q_0 \leq C\}} -C^2 \, \mathrm{d}\mathcal{P} \\ &= -C^2 \cdot \mathcal{P}(\{Q_0 \leq C\}) < 0 \leq \text{ess inf } X_0. \end{aligned}$$

Daraus folgt

$$\inf_{Q \in \mathcal{Q}} E[X_0 Q] \leq E[X_0 Q_0] < \text{ess inf } X_0.$$

Aufgrund obiger Beobachtung kann $\mathcal{D}$ dann nicht lower range dominated sein.

(ii) Dies ergibt sich unmittelbar aus dem Zusatz von Satz (1.2.1) sowie der Tatsache, dass $\mathcal{Q}$ genau dann beschränkt ist, wenn $\mathcal{Y}$ beschränkt ist; um letzteres zu zeigen, benutze man, dass $\mathcal{Q} = 1 - \mathcal{Y}$ und $\|1\|_2 = 1$ gilt.

(iii) „$\Rightarrow$“ Sei $\mathcal{D}$ also darüber hinaus symmetrisch. Da $1 - \mathcal{Q} = \mathcal{Y}$ gilt, müssen wir zeigen, dass $\mathcal{Y}$ mit der konkreten Darstellung (1.28) kreisförmig ist. Als erstes erkennen wir, dass

$$\begin{aligned} \mathcal{Y} &= \{Y \in L^2(\Omega) \mid \mathcal{D}(X) \geq E[XY] \quad \text{für alle } X \in L^2(\Omega)\} \\ &= \{Y \in L^2(\Omega) \mid \mathcal{D}(-X) \geq E[(-X)Y] \quad \text{für alle } X \in L^2(\Omega)\} \\ &= \{Y \in L^2(\Omega) \mid \mathcal{D}(X) \geq E[X(-Y)] \quad \text{für alle } X \in L^2(\Omega)\} \\ &= \{-Z \mid Z \in L^2(\Omega) \quad \text{mit} \quad \mathcal{D}(X) \geq E[XZ] \quad \text{für alle } X \in L^2(\Omega)\} \\ &= -\mathcal{Y} \end{aligned} \tag{1.31}$$

ist, wobei die Symmetrie von $\mathcal{D}$ bei der dritten Gleichheit eingeht.
Sei nun $\lambda \in \mathbb{R}$ mit $|\lambda| \leq 1$ und $Y \in \mathcal{Y}$ beliebig. Da aufgrund der Darstellung (1.28) offenbar $0 \in \mathcal{Y}$ ist, erhalten wir

$$\lambda Y = \operatorname{sign}(\lambda) \cdot \big(|\lambda| Y + (1 - |\lambda|) \cdot 0\big) \in \mathcal{Y}$$

zusammen mit der Identität (1.31) und der Konvexität von $\mathcal{Y}$. Daher ist $\mathcal{Y}$ kreisförmig.

„ $\Leftarrow$ “ Sei $1 - \mathcal{Q} = \mathcal{Y}$ nun zusätzlich kreisförmig. Dann ist insbesondere $\mathcal{Y} = -\mathcal{Y}$ und somit gilt mit der Identität (1.27) gerade

$$\mathcal{D}(X) = \sup_{Y \in \mathcal{Y}} E[XY] = \sup_{Y \in -\mathcal{Y}} E[XY] = \sup_{Y \in \mathcal{Y}} E[X(-Y)] = \sup_{Y \in \mathcal{Y}} E[(-X)Y] = \mathcal{D}(-X)$$

für alle $X \in L^2(\Omega)$, d.h. $\mathcal{D}$ ist symmetrisch. □

Basierend auf Satz (1.2.3) führten Rockafellar et al. folgende Bezeichnung ein, siehe dazu [12, Definition 3] sowie die zugehörigen Bemerkungen.

(1.2.4) Definition (Risk envelopes)
Die Menge $\mathcal{Q} \subseteq L^2(\Omega)$, die in Satz (1.2.3) für ein unterhalbstetiges Abweichungsmaß $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ durch die dort beschriebenen Charakteristika eindeutig bestimmt ist, heißt risk envelope von $\mathcal{D}$. ◇

(1.2.5) Bemerkung (Interpretation von risk envelopes)
Sei $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ ein unterhalbstetiges Abweichungsmaß. Zur Vereinfachung der nachfolgenden Interpretation sei dieses obendrein lower range dominated. Dann erfüllt das risk envelope $\mathcal{Q}$ von $\mathcal{D}$ nach Satz (1.2.3) die Axiome $(\mathcal{Q}1)$ bis $(\mathcal{Q}4)$. Daher ist jedes $Q \in \mathcal{Q}$ aufgrund von $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$ eine Dichtefunktion, sodass wir damit ein weiteres Wahrscheinlichkeitsmaß $\mathcal{P}'_Q : \mathcal{M} \to [0, 1]$ wie in Satz (A.1.25) definieren können. Für jede Zufallsvariable $X \in L^2(\Omega)$ ist dann XQ nach der Hölder-Ungleichung integrierbar, sodass wir ebenfalls die Darstellung

$$E_{\mathcal{P}'_Q}[X] = E[XQ]$$

aus Satz (A.1.25) verwenden können. In den Anwendungen sind solche alternativen Wahrscheinlichkeitsmaße manchmal besser als $\mathcal{P}$ geeignet.
Jedenfalls erhalten wir aus Satz (1.2.3), dass

$$\mathcal{D}(X) = EX - \inf_{Q \in \mathcal{Q}} E[XQ] = \sup_{Q \in \mathcal{Q}} \Big\{ EX - E[XQ] \Big\} = \sup_{Q \in \mathcal{Q}} \Big\{ EX - E_{\mathcal{P}'_Q}[X] \Big\} \qquad (1.32)$$

gilt, d.h. $\mathcal{D}(X)$ entspricht der maximalen unteren Abweichung aller Erwartungswerte bzgl. dieser bestimmten alternativen Wahrscheinlichkeitsmaße im Verhältnis zu EX und $\mathcal{Q}$ enthält genau die dafür relevanten Dichtefunktionen.

Demnach ist $\mathcal{D}(X)$ genau dann zeimlich klein, wenn alle Erwartungswerte bzgl. dieser ausgewählten alternativen Wahrscheinlichkeitsmaße allesamt nicht viel kleiner bzw. sogar größer als EX sind; denn aufgrund von $(\mathcal{Q}2)$ sind nur diejenigen $Q \in \mathcal{Q}$ für die Supremumsbildung in (1.32) relevant, die $E_{\mathcal{P}'_Q}[X] \leq E[X]$ erfüllen. Diese Beobachtung passt auch dazu, wie wir uns ein Abweichungsmaß vorstellen würden; ein Erwartungswert erscheint uns nämlich dann ziemlich sicher, wenn die Erwartungswerte von derselben Zufallsvariablen unter Verwendung einiger unterschiedlicher Wahrscheinlichkeitsmaße kaum vom Ursprünglichen abweichen. Dennoch drängt sich die Frage auf, ob die $Q \in \mathcal{Q}$ mit $E_{\mathcal{P}'_Q}[X] > E[X]$ i.A. vernachlässigt werden. Dem wäre auf jeden Fall nicht so, falls $\mathcal{D}$ beispielsweise symmetrisch wäre; denn in dieser Situation würde dies $\mathcal{D}(X) = \mathcal{D}(-X)$ bzw.

$$\sup_{Q\in\mathcal{Q}} \left\{EX - E_{\mathcal{P}'_Q}[X]\right\} = \sup_{Q\in\mathcal{Q}} \left\{E_{\mathcal{P}'_Q}[X] - EX\right\}$$

nach (1.32) bedeuten, sodass die maximale Abweichung der alternativen Erwartungswerte nach oben und nach unten in Bezug zu EX also identisch wäre. Tatsächlich gibt es ein unterhalbstetiges (sogar stetiges!) Abweichungsmaß, das symmetrisch und lower range dominated ist, wie wir in Abschnitt 2.1 des zweiten Kapitels sehen werden. $\diamond$

Die risk envelopes können in vielen Fällen konkret bestimmt werden. Dies zeigen wir vorerst nur anhand von zwei aus Sektion 1.1 bekannten Beispielen.

(1.2.6) Beispiele

a) Die Standardabweichung σ aus Beispiel (1.1.3) a) ist bekanntlich ein endliches und stetiges, also insbesondere unterhalbstetiges, Abweichungsmaß, sodass das das risk envelope $\mathcal{Q}$ von σ existiert. Nun behaupten wir, dass

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid \sigma(1-Q) \leq 1 \quad \text{und} \quad EQ = 1\}$$

gilt.

Dazu: Zunächst zeigen wir, dass für die eindeutig bestimmte Menge $\mathcal{Y}$ aus dem Beweis von Satz (1.2.3) die Identität

$$\mathcal{Y} = \{Y \in L^2(\Omega) \mid \sigma(Y) \leq 1 \quad \text{und} \quad EY = 0\}$$

erfüllt ist, sodass die Aussage bzgl. $\mathcal{Q}$ direkt aus der Beziehung $\mathcal{Q} = 1 - \mathcal{Y}$ folgt.

„ $\supseteq$ “ Sei $Y \in L^2(\Omega)$ mit $\sigma(Y) \leq 1$ und $EY = 0$. Dies liefert aber gerade

$$1 \geq \sigma(Y) = \|Y - EY\|_2 = \|Y\|_2,$$

sodass wir unter Benutzung der Hölder-Ungleichung

$$\begin{aligned} E[XY] = E[XY] - EX \cdot EY = E[(X - EX)Y] &\leq E[|(X - EX)Y|] \\ &\leq \|X - EX\|_2 \cdot \|Y\|_2 \\ &\leq \sigma(X) \end{aligned}$$

für alle $X \in L^2(\Omega)$ erhalten, d.h. es ist $Y \in \mathcal{Y}$.

„ $\subseteq$ " Sei $Y \in \mathcal{Y}$. Aus $(\mathcal{Q}3)$ und der Identität $\mathcal{Y} = 1 - \mathcal{Q}$ folgt dann schon $EY = 0$, das wie oben $\sigma(Y) = \|Y\|_2$ impliziert. Somit ergibt sich zusammen mit der Darstellung (1.28) von $\mathcal{Y}$ und der Wahl von $X := Y$ die Ungleichung

$$\|Y\|_2 = \sigma(Y) \geq E[Y^2] = \|Y\|_2^2,$$

was nur möglich ist, falls $\sigma(Y) = \|Y\|_2 \leq 1$ ist; denn es gilt

$$x^2 \leq x \quad \text{für alle } x \in [0,1] \qquad \text{und} \qquad x^2 > x \quad \text{für alle } x > 1.$$

Damit ist Behauptung gezeigt. Zusätzlich erhalten wir aus dem Satz (1.2.3) die Beschränktheit von $\mathcal{Q}$, da σ endlich ist; dies können wir aber auch direkt anhand der Gestalt von $\mathcal{Q}$ sehen. Jedoch können wir die Gültigkeit von $(\mathcal{Q}4)$ wegen der i.A. fehlenden Lower range dominance meistens nicht erwarten, während $1 - \mathcal{Q}$ aufgrund der Symmetrie von σ immerhin kreisförmig ist.

b) Das Abweichungsmaß

$$\psi_- : L^2(\Omega) \to [0, \infty], \quad \psi_-(X) = EX - \text{ess inf } X$$

aus Beispiel (1.1.3) c) ist ebenfalls unterhalbstetig und sogar lower range dominated. Deshalb existiert das risk envelope $\mathcal{Q}$ von ψ_- und es erfüllt $(\mathcal{Q}1)$ bis $(\mathcal{Q}4)$. Da ψ_- offenbar für jedes $X \in L^2(\Omega)$ den größtmöglichen Wert annimmt, sodass ψ_- noch $(\mathcal{D}5)$ erfüllt, wird $\mathcal{Q}$ vermutlich die größte Menge sein, die noch allen Axiomen $(\mathcal{Q}1)$ bis $(\mathcal{Q}4)$ genügt. Daher liegt der Gedanke nahe, dass

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid Q \geq 0 \quad \text{und} \quad EQ = 1\} \tag{1.33}$$

gilt.

Dazu: „ $\subseteq$ " Dies folgt direkt aus der Gültigkeit von $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$.

„ $\supseteq$ " Sei $Q \in L^2(\Omega)$ mit $Q \geq 0$ und $EQ = 1$. Wie im Beweis von Satz (1.2.3), Zusatz (i), 1.Fall können wir

$$E[XQ] \geq \text{ess inf } X$$

für alle $X \in L^2(\Omega)$ schließen, das aber offensichtlich dasselbe wie

$$EX - E[XQ] \leq EX - \text{ess inf } X = \psi_-(X)$$

für alle $X \in L^2(\Omega)$ bedeutet, d.h. es ist $Q \in \mathcal{Q}$.

Damit ist unsere Vermutung bestätigt. Da ψ_- i.A. nicht endlich bzw. auch nicht symmetrisch ist, kann $\mathcal{Q}$ in diesen Fällen nicht beschränkt bzw. $1 - \mathcal{Q}$ nicht kreisförmig sein.
Schließlich bemerken wir noch, dass in der Situation von Bemerkung (1.2.5) sogar alle alternativen Wahrscheinlichkeitsmaße $\mathcal{P}'_Q$, die durch eine Dichtefunktion $Q \in L^2(\Omega)$ wie in Satz (A.1.25) definiert sind, herangezogen werden, da $\mathcal{Q}$ aus (1.33) all diese Dichtefunktionen enthält. Aufgrund der Darstellung (1.25) können wir nun

$$\text{ess inf } X = \inf_{Q \in \mathcal{Q}} E[XQ] = \inf_{Q \in \mathcal{Q}} E_{\mathcal{P}'_Q}[X] \tag{1.34}$$

für alle $X \in L^2(\Omega)$ schreiben. Das bedeutet aber gerade, dass wir zu jeder Zufallsvariablen $X \in L^2(\Omega)$ eine Folge von Wahrscheinlichkeitsmaßen mit zugehörigen Dichtefunktionen aus $L^2(\Omega)$ finden, bzgl. derer die Erwartungswerte von X von oben gegen ess inf X konvergieren, das durchaus erstaunlich ist. $\diamond$

Die risk envelopes von σ_+ und σ_- werden wir in einer viel allgemeineren Situation in Abschnitt 1.4 konkret bestimmen, während jene von ψ_+ und ψ mithilfe der Rechenregeln für risk envelopes am Ende dieses Abschnitts angegeben werden können.
Vorher widmen wir uns aber einer eher geometrischen Betrachtung von risk envelopes: Dazu stellen wir zunächst fest, dass jedes risk envelope $\mathcal{Q}$, das zu einem unterhalbstetigen Abweichungsmaß $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ korrespondiert, die Zufallsvariable $Q \equiv 1$ enthält; dies erkennen wir direkt anhand der Darstellung (1.26) von $\mathcal{Q}$. Nachfolgend zeigen wir, dass $Q \equiv 1$ geometrisch gesehen eine besondere Rolle spielt. Um diese zu verstehen, machen wir einen kurzen Exkurs zu Hyperebenen.

(1.2.7) Definition
Sei $\mathcal{X}$ ein $\mathbb{R}$-Vektorraum und $\mathcal{H}$ ein affiner Untervektorraum von $\mathcal{X}$, d.h. es existiert ein Untervektorraum $\mathcal{H}_0$ von $\mathcal{X}$ sowie ein ein $x \in \mathcal{X}$ mit $\mathcal{H} = \mathcal{H}_0 + x$. Dann nennt man $\mathcal{H}$ eine Hyperebene in *$\mathcal{X}$, falls*

$$codim\ \mathcal{H}_0 := dim\ {}^{\mathcal{X}}\!/_{\mathcal{H}_0} = 1$$

gilt. $\diamond$

Ohne Beweis notieren wir folgendes Resultat aus der Funktionalanalysis.

(1.2.8) Lemma
Sei $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ ein Hilbertraum. Dann ist $\mathcal{H}$ genau dann eine abgeschlossene Hyperebene in $\mathcal{X}$, falls ein $x_0 \in \mathcal{X} \setminus \{0\}$ und ein $C \in \mathbb{R}$ existieren, sodass

$$\mathcal{H} = \{x \in \mathcal{X} \mid \langle x, x_0 \rangle = C\}.$$

gilt. $\diamond$

Nun kommen wir zu den angekündigten geometrischen Eigenschaften, die auf [12, Proposition 3] basieren.

(1.2.9) Lemma
Sei $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ ein unterhalbstetiges Abweichungsmaß und $\mathcal{Q}$ das zugehörige risk envelope. Dann gilt:

(i) $\mathcal{Q}$ ist eine Teilmenge der abgeschlossenen Hyperebene

$$\mathcal{H} := \{X \in L^2(\Omega) \mid EX = 1\}$$

in $L^2(\Omega)$.

(ii) Sei $\mathcal{G}$ eine weitere abgeschlossene Hyperebene in $L^2(\Omega)$, also

$$\mathcal{G} = \{X \in L^2(\Omega) \mid E[XY] = C\}$$

für ein gewisses $Y \in L^2(\Omega) \setminus \{0\}$ und ein $C \in \mathbb{R}$, mit $1 \in \mathcal{G}$ und $\mathcal{H} \neq \mathcal{G}$, so existieren je ein $Q_1 \in \mathcal{Q}$ und ein $Q_2 \in \mathcal{Q}$, die in den zu $\mathcal{G}$ zugehörigen offenen Halbräumen liegen, d.h.

$$Q_1 \in \{X \in L^2(\Omega) \mid E[XY] < C\} \quad \textit{und} \quad Q_2 \in \{X \in L^2(\Omega) \mid E[XY] > C\}.$$

$\diamond$

Beweis
(i) Zum Einen ist $\mathcal{H}$ wegen $EX = E[1 \cdot X]$ für alle $X \in L^2(\Omega)$ nach Lemma (1.2.8) eine abgeschlossene Hyperebene und zum Anderen folgt die Inklusionsbeziehung unmittelbar aus $(\mathcal{Q}3)$.

(ii) Zunächst folgt aus $1 \in \mathcal{G}$, dass

$$C = E[1 \cdot Y] = EY \tag{1.35}$$

ist. Darüber hinaus kann Y nicht konstant sein; denn falls ein $K \in \mathbb{R}$ mit $Y \equiv K$ existiert, ist einerseits $K \neq 0$, da $Y \in L^2(\Omega) \setminus \{0\}$ ist, und andererseits $C = K$ nach Gleichung (1.35). Daher gilt für jedes $X \in L^2(\Omega)$ genau dann $X \in \mathcal{G}$, falls

$$E[X \cdot K] = K \quad \text{bzw.} \quad EX = 1$$

ist, sodass $\mathcal{G} = \mathcal{H}$ folgt, das aber ausgeschlossen war.
Jedenfalls ist mit Y dann auch $-Y$ nicht-konstant, sodass nach Axiom $(\mathcal{Q}2)$ je ein $Q_1 \in \mathcal{Q}$ und ein $Q_2 \in \mathcal{Q}$ mit

$$E[YQ_1] < EY \quad \text{bzw.} \quad E[(-Y)Q_2] < E[-Y]$$

oder äquivalenterweise

$$E[YQ_1] < EY \underset{(1.35)}{=} C \quad \text{bzw.} \quad E[YQ_2] > EY \underset{(1.35)}{=} C$$

existieren. Dies entspricht gerade der Behauptung. $\square$

Schließlich stellen wir noch zwei Lemmata mit verschiedenen Rechenregeln für risk envelopes vor. Das Erste beruht auf einer Beobachtung in den Beweisen der Sätze (1.2.1) und (1.2.3), die ROCKAFELLAR ET AL. nur nebenbei in [12, S. 59] bemerkten.

(1.2.10) Lemma
Sei $\mathcal{U}$ eine nicht-leere Teilmenge von $L^2(\Omega)$, die $(\mathcal{Q}2)$ und $(\mathcal{Q}3)$ erfüllt. Dann ist die Abbildung

$$\mathcal{D} : L^2(\Omega) \to [0, \infty], \quad \mathcal{D}(X) := EX - \inf_{Q \in \mathcal{U}} E[XQ]$$

ein unterhalbstetiges Abweichungsmaß mit zugehörigem risk envelope

$$\mathcal{Q} = cl\big(conv(\mathcal{U})\big),$$

also dem konvexen Abschluss von $\mathcal{U}$. ◇

Beweis
Zuerst schreiben wir

$$\mathcal{D}(X) = EX - \inf_{Q \in \mathcal{U}} E[XQ] = \sup_{Q \in \mathcal{U}} \big\{EX - E[XQ]\big\} = \sup_{Q \in \mathcal{U}} E[X(1-Q)]$$

für alle $X \in L^2(\Omega)$, sodass wir ähnlich wie im Beweis von Satz (1.2.1), die Rückrichtung der Äquivalenz, zeigen können, dass $\mathcal{D}$ unterhalbstetig ist und den Axiomen $(\mathcal{D}2)$ und $(\mathcal{D}3)$ genügt, wobei $\mathcal{D}(0) = 0$ ohnehin klar ist; denn für den Nachweis dieser Eigenschaften wird die Konvexität und Abgeschlossenheit der zugrundeliegenden Menge nicht benötigt.
Darüber hinaus folgen $(\mathcal{D}1)$ und $(\mathcal{D}4)$ direkt aus $(\mathcal{Q}2)$ und $(\mathcal{Q}3)$ wie im Beweis von Satz (1.2.3), (1) und (2), da auch dafür die Konvexität und Abgeschlossenheit der jeweiligen Menge nicht erforderlich ist. Somit ist $\mathcal{D}$ ein unterhalbstetiges Abweichungsmaß.

Deshalb bleibt nur noch das risk envelope $\mathcal{Q}$ zu bestimmen: Sei dazu $\hat{\mathcal{Q}} := \mathrm{cl}\big(\mathrm{conv}(\mathcal{U})\big)$. Zunächst beobachten wir, dass $\hat{\mathcal{Q}}$ offensichtlich nicht-leer und abgeschlossen ist. Zudem überträgt sich die Konvexität von $\mathrm{conv}(\mathcal{U})$ auf dessen Abschluss, sodass $\hat{\mathcal{Q}}$ das Axiom $(\mathcal{Q}1)$ erfüllt. Daher reicht es jetzt nachzurechnen, dass die Identität

$$\mathcal{D}(X) = EX - \inf_{Q \in \hat{\mathcal{Q}}} E[XQ] \tag{1.36}$$

für alle $X \in L^2(\Omega)$ gilt; so folgt nämlich $\mathcal{Q} = \hat{\mathcal{Q}}$ aufgrund der Eindeutigkeit des risk envelopes in Satz (1.2.3). Nachfolgend zeigen wir die zu (1.36) äquivalente Beziehung

$$\inf_{Q \in \hat{\mathcal{Q}}} E[XQ] = \inf_{Q \in \mathcal{U}} E[XQ]$$

für alle $X \in L^2(\Omega)$.

Dazu: Sei $X \in L^2(\Omega)$.

„ $\leq$ “ Dies ergibt sich direkt wegen $\mathcal{U} \subseteq \hat{\mathcal{Q}}$.

„ $\geq$ “ Sei dazu vorerst $Q \in \operatorname{conv}(\mathcal{U})$ beliebig, d.h. es existieren $\lambda_1, \ldots, \lambda_n > 0$ mit $\sum_{i=1}^n \lambda_i = 1$ und $Q_1, \ldots, Q_n \in \mathcal{U}$, sodass

$$Q = \sum_{i=1}^n \lambda_i Q_i$$

gilt. Dann erhalten wir

$$E[XQ] = \sum_{i=1}^n \lambda_i \cdot E[XQ_i] \geq \sum_{i=1}^n \lambda_i \cdot \inf_{V \in \mathcal{U}} E[XV] = \inf_{V \in \mathcal{U}} E[XV]. \tag{1.37}$$

Nun sei $Y \in \hat{\mathcal{Q}} = \operatorname{cl}\big(\operatorname{conv}(\mathcal{U})\big)$ beliebig, d.h. es existiert eine Folge $(Y_k)_{k \in \mathbb{N}} \subseteq \operatorname{conv}(\mathcal{U})$ mit $Y_k \xrightarrow[k \to \infty]{} Y$ in $L^2(\Omega)$. Aufgrund der Ungleichung (1.37) gilt zum Einen

$$E[XY_k] \geq \inf_{V \in \mathcal{U}} E[XV]$$

für alle $k \in \mathbb{N}$ und zum Anderen

$$E[XY_k] \xrightarrow[k \to \infty]{} E[XY]$$

aufgrund der Hölder-Ungleichung. Zusammen ergibt sich daraus

$$E[XY] \geq \inf_{V \in \mathcal{U}} E[XV],$$

sodass wiederum die beliebige Wahl von Y die Ungleichung

$$\inf_{Y \in \hat{\mathcal{Q}}} E[XY] \geq \inf_{V \in \mathcal{U}} E[XV]$$

impliziert.

Damit ist die behauptete Gleichung gezeigt. □

Im nächsten Lemma beschreiben wir die risk envelopes von einigen kombinierten bzw. variierten unterhalbstetigen Abweichungsmaßen. Diese Beziehungen stammen größtenteils aus [12, Proposition 4], das wir bereits in Abschnitt 1.1 im Zusammenhang mit Lemma (1.1.11) zitiert haben.

(1.2.11) Lemma
Seien $\mathcal{D}, \mathcal{D}_1, \ldots, \mathcal{D}_n : L^2(\Omega) \to [0, \infty]$ unterhalbstetige Abweichungsmaße mit zugehörigen risk envelopes $\mathcal{Q}, \mathcal{Q}_1, \ldots, \mathcal{Q}_n$. Dann gilt:

(i) Die Reflexion $\tilde{\mathcal{D}}$ von $\mathcal{D}$, die wir in (1.1) definiert haben, ist ein unterhalbstetiges Abweichungsmaß und für das risk envelope $\tilde{\mathcal{Q}}$ gilt die Identität

$$\tilde{\mathcal{Q}} = 2 - \mathcal{Q}.$$

(ii) Das Funktional

$$\sum_{i=1}^{n} \lambda_i \mathcal{D}_i : L^2(\Omega) \to [0, \infty]$$

ist für $\lambda_1, \dots, \lambda_n > 0$ ein unterhalbstetiges Abweichungsmaß und für das risk envelope $\mathcal{U}$ gilt die Identität

$$\mathcal{U} = cl\left(\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i\right) + 1 - \sum_{i=1}^{n} \lambda_i,$$

wobei die Menge $\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$ bereits abgeschlossen ist, falls alle (bis auf vielleicht eines der) $\mathcal{D}_1, \dots, \mathcal{D}_n$ endlich sind.

(iii) Das Funktional

$$max\{\mathcal{D}_1, \dots, \mathcal{D}_n\} : L^2(\Omega) \to [0, \infty]$$

ist ein unterhalbstetiges Abweichungsmaß und für das risk envelope $\mathcal{V}$ gilt die Identität

$$\mathcal{V} = cl\left(conv\left(\bigcup_{i=1}^{n} \mathcal{Q}_i\right)\right).$$

◇

Beweis

Zunächst wissen wir aus Bemerkung (1.1.2) e), dass $\tilde{\mathcal{D}}$ ein unterhalbstetiges Abweichungsmaß ist. Dass dies aber auch auf $\sum_{i=1}^{n} \lambda_i \mathcal{D}_i$ und $\max\{\mathcal{D}_1, \dots, \mathcal{D}_n\}$ zutrifft, folgt aus den Lemmata (1.1.11) und (A.2.9). Demnach müssen wir jeweils nur noch die risk envelopes bestimmen.

(i) Für das risk envelope $\tilde{\mathcal{Q}}$ ergibt sich unter Verwendung der jeweiligen Darstellungen der risk envelopes $\tilde{\mathcal{Q}}$ und $\mathcal{Q}$ aus (1.26), dass

$$\begin{aligned}
\tilde{\mathcal{Q}} &= \{Q \in L^2(\Omega) \mid \tilde{\mathcal{D}}(X) \geq EX - E[XQ] \quad \text{für alle } X \in L^2(\Omega)\} \\
&= \{Q \in L^2(\Omega) \mid \mathcal{D}(-X) \geq EX - E[XQ] \quad \text{für alle } X \in L^2(\Omega)\} \\
&= \{Q \in L^2(\Omega) \mid \mathcal{D}(Y) \geq E[-Y] - E[(-Y)Q] \quad \text{für alle } Y \in L^2(\Omega)\} \\
&= \{Q \in L^2(\Omega) \mid \mathcal{D}(Y) \geq E[Y] - E[Y(2-Q)] \quad \text{für alle } Y \in L^2(\Omega)\} \\
&= \{2 - U \mid U \in L^2(\Omega) \quad \text{mit} \quad \mathcal{D}(Y) \geq E[Y] - E[YU] \quad \text{für alle } Y \in L^2(\Omega)\} \\
&= 2 - \mathcal{Q}
\end{aligned}$$

gilt.

(ii) Sei $\hat{\mathcal{U}} := \text{cl}\,(\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i) + 1 - \sum_{i=1}^{n} \lambda_i$. So ist $\hat{\mathcal{U}}$ offensichtlich nicht-leer und als Translation der abgeschlossenen Menge $\text{cl}\,(\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i)$ wiederum abgeschlossen. Obendrein ist $\hat{\mathcal{U}}$ konvex; denn seien $\lambda \in [0,1]$ und zunächst $U, V \in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$ beliebig, d.h. es existieren $U_i, V_i \in \mathcal{Q}_i$ für $i = 1, \dots, n$ mit

$$U = \sum_{i=1}^{n} \lambda_i U_i \qquad \text{und} \qquad V = \sum_{i=1}^{n} \lambda_i V_i.$$

Dann ist

$$\lambda U + (1-\lambda)V = \lambda \sum_{i=1}^{n} \lambda_i U_i + (1-\lambda) \sum_{i=1}^{n} \lambda_i V_i = \sum_{i=1}^{n} \lambda_i \big(\underbrace{\lambda U_i + (1-\lambda) V_i}_{\in\, \mathcal{Q}_i} \big)$$
$$\in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$$

unter Benutzung der Konvexität von $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$, sodass $\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$ aufgrund der beliebigen Wahl von λ und U, V konvex ist. Damit sind aber auch dessen Abschluss sowie Translationen des Abschlusses konvex, d.h. insbesondere $\hat{\mathcal{U}}$ ist konvex. Insgesamt erfüllt $\hat{\mathcal{U}}$ also $(\mathcal{Q}1)$, sodass wir nur noch nachweisen müssen, dass

$$\sum_{i=1}^{n} \lambda_i \mathcal{D}_i(X) = EX - \inf_{Q \in \hat{\mathcal{U}}} E[XQ] \tag{1.38}$$

für alle $X \in L^2(\Omega)$ gilt; dann folgt nämlich $\mathcal{U} = \hat{\mathcal{U}}$ wegen der Eindeutigkeit des risk envelopes in Satz (1.2.3).

Dazu: Zunächst zeigen wir für $\mathcal{U}' := \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i + 1 - \sum_{i=1}^{n} \lambda_i$ die Identität

$$\sum_{i=1}^{n} \lambda_i \mathcal{D}_i(X) = EX - \inf_{Q \in \mathcal{U}'} E[XQ] \tag{1.39}$$

für alle $X \in L^2(\Omega)$. Für alle $X \in L^2(\Omega)$ erhalten wir unter Verwendung von

$$\mathcal{D}_i(X) = EX - \inf_{Q \in \mathcal{Q}_i} E[XQ]$$

für $i = 1, \ldots, n$, dass

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i \mathcal{D}_i(X) &= \sum_{i=1}^{n} \lambda_i \left(EX - \inf_{Q \in \mathcal{Q}_i} E[XQ] \right) \\
&= \sum_{i=1}^{n} \lambda_i EX - \sum_{i=1}^{n} \lambda_i \inf_{Q \in \mathcal{Q}_i} E[XQ] \\
&= \sum_{i=1}^{n} \lambda_i EX - \inf_{Q_1 \in \mathcal{Q}_1, \ldots, Q_n \in \mathcal{Q}_n} \sum_{i=1}^{n} \lambda_i E[XQ_i] \\
&= \sum_{i=1}^{n} \lambda_i EX - \inf_{Q_1 \in \mathcal{Q}_1, \ldots, Q_n \in \mathcal{Q}_n} E\left[X \left(\sum_{i=1}^{n} \lambda_i Q_i \right) \right] \\
&= \sum_{i=1}^{n} \lambda_i EX - \inf_{U \in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i} E[XU]
\end{aligned}$$

$$
\begin{aligned}
&= EX + \left(\sum_{i=1}^{n} \lambda_i - 1\right) EX - \inf_{U \in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i} E[XU] \\
&= EX - \inf_{U \in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i} \left(E[XU] + \left(1 - \sum_{i=1}^{n} \lambda_i\right) EX \right) \\
&= EX - \inf_{U \in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i} E\left[X \left(U + 1 - \sum_{i=1}^{n} \lambda_i \right) \right] \\
&= EX - \inf_{V \in \mathcal{U}'} E[XV]
\end{aligned}
$$

erfüllt ist. Somit ist die Identität (1.39) gezeigt. Jetzt können wir ähnlich wie im Beweis von Lemma (1.2.10) nachrechnen, dass

$$\inf_{Q \in \mathcal{U}'} E[XQ] = \inf_{Q \in \mathrm{cl}(\mathcal{U}')} E[XQ]$$

bzw.

$$\sum_{i=1}^{n} \lambda_i \mathcal{D}_i(X) = EX - \inf_{Q \in \mathcal{U}'} E[XQ] = EX - \inf_{Q \in \mathrm{cl}(\mathcal{U}')} E[XQ]$$

für alle $X \in L^2(\Omega)$ gilt. Wenn wir dann noch

$$\mathrm{cl}(\mathcal{U}') = \mathrm{cl}\left(\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i + 1 - \sum_{i=1}^{n} \lambda_i \right) = \mathrm{cl}\left(\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i \right) + 1 - \sum_{i=1}^{n} \lambda_i = \mathcal{U}$$

benutzen, folgt unmittelbar die behauptete Darstellung (1.38).

Schließlich kommen wir zu dem Zusatz: Seien alle (bis auf vielleicht eines der) $\mathcal{D}_1, \ldots, \mathcal{D}_n$ endlich. Œ nehmen wir an, dass $\mathcal{D}_1, \ldots, \mathcal{D}_{n-1}$ endlich sind und nur $\mathcal{D}_n$ evtl. nicht endlich ist. Da $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$ allesamt dem Axiom $(\mathcal{Q}1)$ genügen, also nicht-leer, konvex und abgeschlossen sind, sind sie nach einer Variante des Satzes von Mazur auch schwach abgeschlossen. Daher sind aber auch $\lambda_1 \mathcal{Q}_1, \ldots, \lambda_n \mathcal{Q}_n$ schwach abgeschlossen, da die Skalarmultiplikation bzgl. der schwachen Topologie auf dem Raum $L^2(\Omega)$ stetig ist. Ferner sind $\mathcal{Q}_1, \ldots, \mathcal{Q}_{n-1}$ und somit ebenfalls $\lambda_1 \mathcal{Q}_1, \ldots, \lambda_{n-1} \mathcal{Q}_{n-1}$ nach Satz (1.2.3) beschränkt, da $\mathcal{D}_1, \ldots, \mathcal{D}_{n-1}$ nach Annahme endlich sind. Insgesamt ergibt sich daraus, dass $\lambda_1 \mathcal{Q}_1, \ldots, \lambda_{n-1} \mathcal{Q}_{n-1}$ schwach folgenkompakt sind.

Dazu: Wir beweisen die Aussage beispielhaft für $M := \lambda_1 \mathcal{Q}_1$.
Sei $(X_m)_{m \in \mathbb{N}} \subseteq M$ beliebig. Einerseits impliziert die Beschränktheit von M, dass auch $(X_m)_{m \in \mathbb{N}}$ beschränkt ist, und andererseits ist $L^2(\Omega)$ als Hilbertraum reflexiv. Daher können wir mit [10, Satz (6.2.3)] und einem einfachen Skalierungsargument schließen, dass eine Teilfolge $(X_{m_k})_{k \in \mathbb{N}} \subseteq (X_m)_{m \in \mathbb{N}}$ und ein $X \in L^2(\Omega)$ mit $X_{m_k} \rightharpoonup X$ für $k \to \infty$ in $L^2(\Omega)$ existiert. Aufgrund der oben schon begründeten schwachen Abgeschlossenheit von M folgt damit $X \in M$, sodass M wegen der beliebigen Wahl von $(X_m)_{m \in \mathbb{N}}$ schwach folgenkompakt ist.

Jedenfalls können wir daraus ableiten, dass $\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$ schwach folgenabgeschlossen ist.

Dazu: Seien $(X_m)_{m\in\mathbb{N}} \subseteq \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$ und $X \in L^2(\Omega)$ so, dass $X_m \rightharpoonup X$ für $m \to \infty$ in $L^2(\Omega)$. Nun finden wir zu jedem $m \in \mathbb{N}$ gewisse $X_m^{(i)} \in \lambda_i \mathcal{Q}_i$ für $i = 1, \dots, n$ mit

$$X_m = \sum_{i=1}^{n} X_m^{(i)}.$$

Da $\lambda_1 \mathcal{Q}_1$ schwach folgenkompakt ist und $\left(X_m^{(1)}\right)_{m\in\mathbb{N}} \subseteq \lambda_1 \mathcal{Q}_1$ gilt, existiert eine Teilfolge $\left(m_1(k)\right)_{k\in\mathbb{N}} \subseteq (m)_{m\in\mathbb{N}}$ und ein $X^{(1)} \in \lambda_1 \mathcal{Q}_1$ mit

$$X_{m_1(k)}^{(1)} \rightharpoonup X^{(1)} \quad \text{für } k \to \infty \text{ in } L^2(\Omega).$$

Falls $2 \leq n-1$ ist, können wir erneut für $\left(X_{m_1(k)}^{(2)}\right)_{k\in\mathbb{N}} \subseteq \lambda_2 \mathcal{Q}_2$ unter Verwendung der schwachen Folgenkompaktheit von $\lambda_2 \mathcal{Q}_2$ folgern, dass eine Teilfolge $\left(m_2(k)\right)_{k\in\mathbb{N}} \subseteq \left(m_1(k)\right)_{k\in\mathbb{N}}$ und ein $X^{(2)} \in \lambda_2 \mathcal{Q}_2$ mit

$$X_{m_2(k)}^{(2)} \rightharpoonup X^{(2)} \quad \text{für } k \to \infty \text{ in } L^2(\Omega)$$

existiert. Natürlich gilt dann obendrein $X_{m_2(k)}^{(1)} \rightharpoonup X^{(1)}$ für $k \to \infty$ in $L^2(\Omega)$. Mittels Iteration dieses Argumentes erhalten wir also eine Teilfolge $\left(m_{n-1}(k)\right)_{k\in\mathbb{N}} \subseteq (m)_{m\in\mathbb{N}}$ sowie $X^{(i)} \in \lambda_i \mathcal{Q}_i$ mit

$$X_{m_{n-1}(k)}^{(i)} \rightharpoonup X^{(i)} \quad \text{für } k \to \infty \text{ in } L^2(\Omega)$$

für $i = 1, \dots, n-1$. Daraus resultiert dann

$$X_{m_{n-1}(k)}^{(n)} = X_{m_{n-1}(k)} - \sum_{i=1}^{n-1} X_{m_{n-1}(k)}^{(i)} \rightharpoonup X - \sum_{i=1}^{n-1} X^{(i)} =: X^{(n)} \quad \text{für } k \to \infty \text{ in } L^2(\Omega),$$

wobei $X^{(n)} \in \lambda_n \mathcal{Q}_n$ ist, da $\lambda_n \mathcal{Q}_n$ schwach folgenabgeschlossen ist. Insgesamt folgt daher

$$X = \sum_{i=1}^{n} \underbrace{X^{(i)}}_{\in\, \lambda_i \mathcal{Q}_i} \in \sum_{i=1}^{n} \lambda_i \mathcal{Q}_i,$$

d.h. $\sum_{i=1}^{n} \lambda_i \mathcal{Q}_i$ ist schwach folgenabgeschlossen.

Da aus der schwachen Folgenabgeschlossenheit die (starke) (Folgen-)Abgeschlossenheit folgt, ist der Zusatz also gezeigt.

(iii) Sei $\hat{\mathcal{V}} := \mathrm{cl}\left(\mathrm{conv}(\bigcup_{i=1}^{n} \mathcal{Q}_i)\right)$. Dann ist $\hat{\mathcal{V}}$ zum Einen offenbar nicht-leer und abgeschlossen und zum Anderen ist mit $\mathrm{conv}\left(\bigcup_{i=1}^{n} \mathcal{Q}_i\right)$ auch dessen Abschluss, also

$\hat{\mathcal{V}}$, konvex, sodass $\hat{\mathcal{V}}$ der Eigenschaft $(\mathcal{Q}1)$ genügt. Wie in (ii) müssen wir deshalb aufgrund der Eindeutigkeit des risk envelopes $\mathcal{V}$ nur zeigen, dass

$$\max\{\mathcal{D}_1(X), \dots, \mathcal{D}_n(X)\} = EX - \inf_{Q \in \hat{\mathcal{V}}} E[XQ]$$

für alle $X \in L^2(\Omega)$ gilt.

Dazu: Für alle $X \in L^2(\Omega)$ ergibt sich wiederum mit der Darstellung

$$\mathcal{D}_i(X) = EX - \inf_{Q \in \mathcal{Q}_i} E[XQ]$$

für $i = 1, \dots, n$, dass

$$\begin{aligned}\max\{\mathcal{D}_1(X), \dots, \mathcal{D}_n(X)\} &= \max\left\{EX - \inf_{Q \in \mathcal{Q}_1} E[XQ], \dots, EX - \inf_{Q \in \mathcal{Q}_n} E[XQ]\right\} \\ &= EX - \min\left\{\inf_{Q \in \mathcal{Q}_1} E[XQ], \dots, \inf_{Q \in \mathcal{Q}_n} E[XQ]\right\}\end{aligned}$$

erfüllt ist. Daher ist es ausreichend zu zeigen, dass

$$\min\left\{\inf_{Q \in \mathcal{Q}_1} E[XQ], \dots, \inf_{Q \in \mathcal{Q}_n} E[XQ]\right\} = \inf_{Q \in \hat{\mathcal{V}}} E[XQ]$$

für alle $X \in L^2(\Omega)$ gilt. Zur Vereinfachung rechnen wir aber vorher für $\mathcal{V}' := \bigcup_{i=1}^n \mathcal{Q}_i$ die Identität

$$\min\left\{\inf_{Q \in \mathcal{Q}_1} E[XQ], \dots, \inf_{Q \in \mathcal{Q}_n} E[XQ]\right\} = \inf_{Q \in \mathcal{V}'} E[XQ] \tag{1.40}$$

für alle $X \in L^2(\Omega)$ nach. Dafür sei eine Zufallsvariable $X \in L^2(\Omega)$ gegeben.

„ $\geq$ “ Offenbar folgt aus der Inklusion $\mathcal{Q}_i \subseteq \mathcal{V}'$ die Ungleichung

$$\inf_{Q \in \mathcal{Q}_i} E[XQ] \geq \inf_{Q \in \mathcal{V}'} E[XQ]$$

für alle $i = 1, \dots, n$, das aber gerade

$$\min\left\{\inf_{Q \in \mathcal{Q}_1} E[XQ], \dots, \inf_{Q \in \mathcal{Q}_n} E[XQ]\right\} \geq \inf_{Q \in \mathcal{V}'} E[XQ]$$

impliziert.

„ $\leq$ “ Sei $Q \in \mathcal{V}' = \bigcup_{i=1}^n \mathcal{Q}_i$ beliebig, d.h. es existiert ein $i \in \{1, \dots, n\}$ mit $Q \in \mathcal{Q}_i$. Demnach ist

$$E[XQ] \geq \inf_{U \in \mathcal{Q}_i} E[XU] \geq \min\left\{\inf_{U \in \mathcal{Q}_1} E[XU], \dots, \inf_{U \in \mathcal{Q}_n} E[XU]\right\}$$

und, da Q beliebig gewählt war, ergibt sich

$$\inf_{U\in\mathcal{V}'} E[XU] \geq \min\left\{\inf_{U\in\mathcal{Q}_1} E[XU],\ldots,\inf_{U\in\mathcal{Q}_n} E[XU]\right\}.$$

Somit ist die Identität (1.40) gezeigt.

Wie im Beweis von Lemma (1.2.10) folgt dann

$$\inf_{Q\in\mathcal{V}'} E[XQ] = \inf_{Q\in\hat{\mathcal{V}}} E[XQ]$$

bzw.

$$\max\{\mathcal{D}_1(X),\ldots,\mathcal{D}_n(X)\} = EX - \inf_{Q\in\mathcal{V}'} E[XQ] = EX - \inf_{Q\in\hat{\mathcal{V}}} E[XQ]$$

für alle $X \in L^2(\Omega)$, was zu zeigen war. □

Damit können wir jetzt also auch die risk envelopes von ψ_+ und ψ berechnen.

(1.2.12) Beispiele

a) Aus Beispiel (1.1.3) c) wissen wir bereits, dass ψ_+ die Reflexion des unterhalbstetigen Abweichungsmaßes ψ_- ist. Ferner besitzt das risk envelope $\mathcal{Q}$ von ψ_- aus (1.33) die Darstellung

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid Q \geq 0 \quad \text{und} \quad EQ = 1\}.$$

Nach Lemma (1.2.11) (i) gilt somit für das risk envelope $\tilde{\mathcal{Q}}$ von ψ_+ die Identität

$$\begin{aligned}\tilde{\mathcal{Q}} = 2 - \mathcal{Q} &= \{2 - Q \mid Q \in L^2(\Omega) \quad \text{mit} \quad Q \geq 0 \quad \text{und} \quad EQ = 1\}\\ &= \{U \in L^2(\Omega) \mid 2 - U \geq 0 \quad \text{und} \quad E[2-U] = 1\}\\ &= \{U \in L^2(\Omega) \mid U \leq 2 \quad \text{und} \quad EU = 1\}.\end{aligned}$$

Daran sehen wir auch gut, dass $\tilde{\mathcal{Q}}$ das Axiom $(\mathcal{Q}4)$ in der Regel nicht erfüllt, sodass ψ_+ nach Satz (1.2.3) i.A. nicht lower range dominated ist. Darüber hinaus ist $\tilde{\mathcal{Q}}$ in einigen Fällen nicht beschränkt sowie $1 - \tilde{\mathcal{Q}}$ nicht kreisförmig, da ψ_+ bekanntlich häufig nicht endlich bzw. nicht symmetrisch ist.

b) In Gleichung (1.12) haben wir nachgerechnet, dass ψ die Beziehung

$$\psi(X) = \text{ess sup } X - \text{ess inf } X = \psi_+(X) + \psi_-(X)$$

für alle $X \in L^2(\Omega)$ erfüllt. Nach Lemma (1.2.11) (ii) ergibt sich daher für das risk envelope $\mathcal{U}$ von ψ unter Benutzung der Bezeichnungen aus a), dass

$$\mathcal{U} = \text{cl}(\mathcal{Q} + \tilde{\mathcal{Q}}) + 1 - 2 = \text{cl}(\mathcal{Q} + \tilde{\mathcal{Q}}) - 1$$

gilt. Da sowohl ψ_- als auch ψ_+ i.A. nicht endlich sind, können wir in dieser Darstellung von $\mathcal{U}$ nicht ohne weitere Begründungen den Abschluss von $\mathcal{Q} + \tilde{\mathcal{Q}}$

vernachlässigen. Trotz dieser ziemlich abstrakten Gestalt von $\mathcal{U}$ können wir mithilfe von Satz (1.2.3) festhalten, dass $\mathcal{U}$ i.A. weder der Eigenschaft $(\mathcal{Q}4)$ genügt und noch beschränkt ist, da ψ i.A. nicht lower range dominated und nicht endlich ist. Jedoch ist $1 - \mathcal{U}$ immerhin kreisförmig, da ψ immer symmetrisch ist. $\diamond$

1.3 Beziehung zu Risikomaßen

Neben den bis jetzt ausführlich untersuchten Abweichungsmaßen gibt es noch sogenannte „Risikomaße". Während in diesem Gebiet eine große Vielzahl an verschiedenen Ansätzen verfolgt werden, beschränken wir uns in dieser Arbeit u.A. auf die zwei Konzepte von „Risikomaßen" aus [12, Definition 4] und [12, Definition 5], die Rockafellar et al. verwendeten. Das Ziel dieser Sektion ist es im Wesentlichen, den Zusammenhang zwischen diesen „Risikomaßen" und Abweichungsmaßen, den Rockafellar et al. erarbeiteten, zu erklären.
Vorher wollen wir uns aber für diese Theorie motivieren, wobei daraus im Endeffekt nur eines dieser beiden Konzepte, und zwar der sogenannten „kohärenten Risikomaße", resultierte: In vielen Situationen bzw. Modellen, besonders in der Finanzmathematik, kommt es vor, dass eine Vielzahl an Zufallsvariablen bzw. deren Werte gewisse Schlüsse zulassen, d.h. für eine solche Zufallsvariable $X : \Omega \to \mathbb{R}$ und einem Ereignis $\omega \in \Omega$ impliziert der Fall $X(\omega) < 0$ eher einen Verlust oder einen Rückgang, während der Fall $X(\omega) > 0$ zumeist einen Gewinn oder einen Zuwachs andeutet. Obendrein ist das Ausmaß des Verlustes bzw. des Gewinnes umso größer, je größer $|X(\omega)|$ im jeweiligen Fall ist. In dieser Situation interessiert man sich dafür, wie groß denn „das Gesamtausmaß an möglichen Verlusten" im Zusammenhang mit X ist. Jedoch wollen wir hier auch das Gewinnpotenzial berücksichtigen und genau dies sollen die „Risikomaße" bei uns durch einen Zahlenwert bewerten. Da wir Null als Bezugspunkt gewählt haben, ist der Gedanke naheliegend, dass sie vermutlich nicht translationsinvariant sein werden; das Gegenteil wird sogar der Fall sein, wie wir gleich sehen werden.

Jedenfalls kommen wir jetzt zur Definition.

(1.3.1) Definition (Risikomaße)
Ein Funktional $\mathcal{R} : L^2(\Omega) \to (-\infty, \infty]$ heißt kohärentes Risikomaß, *falls es folgende Eigenschaften aufweist:*

$(\mathcal{R}1)$ $\mathcal{R}(X + C) = \mathcal{R}(X) - C$ *für alle* $X \in L^2(\Omega)$ *und* $C \in \mathbb{R}$,
$(\mathcal{R}2)$ $\mathcal{R}(0) = 0$ *und* $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ *für alle* $X \in L^2(\Omega)$ *und* $\lambda > 0$,
$(\mathcal{R}3)$ $\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$ *für alle* $X, Y \in L^2(\Omega)$,
$(\mathcal{R}4)$ $\mathcal{R}(X) \leq \mathcal{R}(Y)$ *für alle* $X, Y \in L^2(\Omega)$ *mit* $X \geq Y$.

Hingegen bezeichnet man $\mathcal{R}$ als ein strictly expectation bounded Risikomaß, *falls es den Axiomen* $(\mathcal{R}1)$, $(\mathcal{R}2)$, $(\mathcal{R}3)$ *sowie*

$(\mathcal{R}5)$ $\mathcal{R}(X) > E[-X]$ *für alle nicht-konstanten* $X \in L^2(\Omega)$

genügt, d.h. auf das Axiom $(\mathcal{R}4)$ *wird i.A. verzichtet.*

In den jeweiligen Situationen nennt man $\mathcal{R}$ *ein* endliches Risikomaß, *falls* $\mathcal{R}$ *ein endliches Funktional darstellt, also falls* $\mathcal{R}\big(L^2(\Omega)\big) \subseteq \mathbb{R}$ *gilt.* ◇

(1.3.2) Bemerkung

a) Neben den schon aus Bemerkung 1.1.2) a) bekannten Bezeichnungen für $(\mathcal{R}2)$ und $(\mathcal{R}3)$ nennen wir das Charakteristikum in

$\rightarrow$ $(\mathcal{R}4)$ **Monotonie**.

Darüber hinaus heißt $\mathcal{R}$ natürlich kohärentes und strictly expectation bounded Risikomaß, falls es allen Eigenschaften $(\mathcal{R}1)$ bis $(\mathcal{R}5)$ genügt.

b) In beiden Fällen folgt analog zu Bemerkung (1.1.2) b) aus $(\mathcal{R}2)$ und $(\mathcal{R}3)$, dass $\mathcal{R}$ konvex ist.

c) Falls $\mathcal{R}$ ein strictly expectation bounded Risikomaß ist, können wir zudem beobachten, dass für alle konstanten $X \in L^2(\Omega)$ die Beziehung

$$\mathcal{R}(X) = \mathcal{R}(0 + X) \underset{(\mathcal{R}1)}{=} \mathcal{R}(0) - X \underset{(\mathcal{R}2)}{=} -X = E[-X]$$

gilt, d.h. es ist $\mathcal{R}(X) \geq E[-X]$ für alle $X \in L^2(\Omega)$, wobei genau dann Gleichheit besteht, falls X konstant ist. Offensichtlich kommt der Name „**strictly** expectation bounded" daher, dass die Ungleichung für nicht-konstante $X \in L^2(\Omega)$ strikt ist.

d) Für die weitere Interpretation beschränken wir uns auf kohärente Risikomaße: Denn in dieser Situation folgt aus den Axiomen $(\mathcal{R}2)$ und $(\mathcal{R}4)$, dass zu $X \in L^2(\Omega)$ die Ungleichungen

$$\mathcal{R}(X) \geq 0 \quad \text{für } X \leq 0 \qquad \text{sowie} \qquad \mathcal{R}(X) \leq 0 \quad \text{für } X \geq 0$$

gelten. Dies passt auch zu dem oben vorgestellten Hintergrund, da $\mathcal{R}$ für $X \leq 0$, also Verlust in jedem Fall, zumindest einen nicht-negativen Wert ausgeben sollte; dazu sei bemerkt, dass wir das Verlustpotenzial positiv gewichten, da wir uns schließlich darauf fokussieren wollen. Dementsprechend sind auch die negativen Werte $\mathcal{R}(X)$ für $X \geq 0$ sinnvoll. In allen anderen Fällen bzgl. X lässt sich keine Aussage über $\mathcal{R}(X)$ treffen, da $\mathcal{R}$ durchaus unterschiedliche Schwerpunkte setzen kann.
Im Vergleich dazu können wir die strictly expectation bounded Risikomaße zwar nicht wirklich interpretieren, aber diese werden später wichtig, sobald wir die Beziehungen zu Abweichungsmaßen herstellen werden.

e) Das Konzept der kohärenten Risikomaße haben ROCKAFELLAR ET AL. größtenteils aus [2] von ARTZNER ET AL. übernommen, wobei $(\mathcal{R}1)$ leicht verändert und der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ hier allgemein gelassen wurde. Die Eigenschaft der strictly expectation boundedness findet aber keine Entsprechung in [2]. Weitere Erweiterungen zu [2] finden sich z.B. in [4, 5]. $\diamond$

Zur Veranschaulichung diskutieren wir zwei einfache Beispiele.

(1.3.3) Beispiele

a) Ein sehr intuitives Beispiel eines kohärenten Risikomaßes ist durch

$$\mathcal{R} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{R}(X) := E[-X]$$

gegeben, wobei sich die Endlichkeit von $\mathcal{R}$ mithilfe der Hölder-Ungleichung ergibt.
Zunächst gilt $(\mathcal{R}1)$ wegen

$$\mathcal{R}(X + C) = E[-X] + E[-C] = E[-X] - C = \mathcal{R}(X) - C$$

für alle $X \in L^2(\Omega)$ und $C \in \mathbb{R}$. Des Weiteren ist die Gültigkeit von $(\mathcal{R}2)$ und $(\mathcal{R}3)$ offensichtlich, da $\mathcal{R}$ sogar linear ist. Schließlich seien $X, Y \in L^2(\Omega)$ mit $X \geq Y$ bzw. $-X \leq -Y$, sodass dies in Kombination mit den Monotonie-Eigenschaften des Integrals

$$\mathcal{R}(X) = E[-X] \leq E[-Y] = \mathcal{R}(Y)$$

liefert, d.h. $\mathcal{R}$ erfüllt auch $(\mathcal{R}4)$. Hingegen genügt es offenbar nie $(\mathcal{R}5)$.
Ferner ist $\mathcal{R}$ stetig, das erneut mittels der Hölder-Ungleichung nachgerechnet werden kann.

b) Sei nun

$$\mathcal{R} : L^2(\Omega) \to (-\infty, \infty], \quad \mathcal{R}(X) := -\text{ess inf } X.$$

Hierbei ist $\mathcal{R}$ nach Lemma (A.2.5) und Bemerkung (1.1.6) genau dann endlich, falls der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich ist.
Im Folgenden rechnen wir nach, dass $\mathcal{R}$ ein kohärentes und strictly expectation bounded Riskomaß ist: Zuallererst folgt $(\mathcal{R}1)$ aus

$$\begin{aligned}\mathcal{R}(X + C) = -\text{ess inf } (X + C) = -\big((\text{ess inf } X) + C\big) &= -(\text{ess inf } X) - C \\ &= \mathcal{R}(X) - C\end{aligned}$$

für alle $X \in L^2(\Omega)$ und $C \in \mathbb{R}$. Während $\mathcal{R}$ offensichtlich $(\mathcal{R}2)$ erfüllt, ergibt sich zum Einen $(\mathcal{R}3)$ aus Lemma (A.1.6) (vi) und zum Anderen $(\mathcal{R}4)$ aus Lemma (A.1.6) (viii). Ferner sei $X \in L^2(\Omega)$ nicht-konstant, so erhalten wir unter Benutzung von ψ_- aus Beispiel (1.1.3) c) die Ungleichung

$$\begin{aligned}\mathcal{R}(X) = -\text{ess inf } X = (EX - \text{ess inf } X) - EX = \psi_-(X) - EX \underset{(\mathcal{D}4)}{>} &-EX \\ &= E[-X],\end{aligned}$$

d.h. $\mathcal{R}$ genügt der Eigenschaft ($\mathcal{R}5$).
Schließlich impliziert Lemma (A.2.5) zusammen mit Bemerkung (A.2.2) b) noch, dass $\mathcal{R}$ stets unterhalbstetig ist und sogar stetig sein kann. Diese Verschärfung tritt aber genau dann ein, falls $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich ist. ◇

Wir werden später sehen, dass es auch eine ganze Reihe an strictly expectation bounded Risikomaßen gibt, die aber nicht kohärent sind.
Jedenfalls erkannten ROCKAFELLAR ET AL. in [12, S. 60], dass man analog zu Lemma (1.1.4) Funktionale, die ($\mathcal{R}1$) bis ($\mathcal{R}3$) erfüllen, durch scheinbar schwächere Axiome charakterisieren kann.

(1.3.4) Lemma
Ein Funktional $\mathcal{R} : L^2(\Omega) \to (-\infty, \infty]$ genügt genau dann den Charakteristika ($\mathcal{R}1$) bis ($\mathcal{R}3$), wenn R die Axiome

$$(\mathcal{R}1') \ \ \mathcal{R}(C) = -C \quad \textit{für alle } C \in \mathbb{R},$$
$$(\mathcal{R}2') \ \ \mathcal{R}(\lambda X) = \lambda \mathcal{R}(X) \quad \textit{für alle } X \in L^2(\Omega) \textit{ und } \lambda > 0$$

sowie ($\mathcal{R}3$) erfüllt. ◇

Obendrein können wir aus dem Satz (1.1.7) genauso wie für Abweichungsmaße ein hinreichendes Kriterium für Stetigkeit ableiten, wenn wir uns an Bemerkung (1.3.2) b) erinnern.

(1.3.5) Satz (Stetigkeit von Risikomaßen)
Sei $\mathcal{R} : L^2(\Omega) \to \mathbb{R}$ ein endliches, unterhalbstetiges und kohärentes bzw. strictly expectation bounded Risikomaß. Dann ist $\mathcal{R}$ sogar stetig. ◇

Nun präsentieren wir den entscheidenden Satz von ROCKAFELLAR ET AL., siehe dazu [12, Theorem 2], in dem u.A. der Zusammenhang von strictly expectation bounded Risikomaßen und Abweichungsmaßen hergestellt wird.

(1.3.6) Satz (Beziehung von Risiko- und Abweichungsmaßen)
Seien $\mathscr{D}$ die Menge aller Abweichungsmaße und $\mathscr{R}$ die Menge aller strictly expectation bounded Risikomaße. Dazu definieren wir die Abbildungen

$$\mathscr{F} : \mathscr{R} \to \mathscr{D}, \quad \mathcal{R} \mapsto \Big[\mathscr{F}[\mathcal{R}] : L^2(\Omega) \to [0, \infty],\ \mathscr{F}[\mathcal{R}](X) := \mathcal{R}(X - EX)\Big]$$

und

$$\mathscr{G} : \mathscr{D} \to \mathscr{R}, \quad \mathcal{D} \mapsto \Big[\mathscr{G}[\mathcal{D}] : L^2(\Omega) \to (-\infty, \infty],\ \mathscr{G}[\mathcal{D}](X) := \mathcal{D}(X) + E[-X]\Big].$$

Dann sind $\mathscr{F}$ und $\mathscr{G}$ wohldefiniert und invers zueinander, also insbesondere bijektiv.

Zusätzlich gilt für alle $\mathcal{R} \in \mathscr{R}$:

(i) $\mathcal{R}$ *ist genau dann kohärent, falls* $\mathscr{F}[\mathcal{R}]$ *lower range dominated ist.*

(ii) $\mathcal{R}$ *ist genau dann endlich bzw. (unterhalb-)stetig, falls* $\mathscr{F}[\mathcal{R}]$ *endlich bzw. (unterhalb-)stetig ist.* ◇

Beweis

Damit $\mathscr{F}$ und $\mathscr{G}$ wohldefiniert sind, müssen wir zeigen, dass $\mathscr{F}[\mathcal{R}] \in \mathscr{D}$ für alle $\mathcal{R} \in \mathscr{R}$ sowie $\mathscr{G}[\mathcal{D}] \in \mathscr{R}$ für alle $\mathcal{D} \in \mathscr{D}$ ist. Zur Abkürzung verwenden wir die Lemmata (1.1.4) und (1.3.4).

Zu $\mathscr{F}$*:* Sei dazu $\mathcal{R} \in \mathscr{R}$. Zunächst beobachten wir, dass die Eigenschaft ($\mathcal{R}1$) die Darstellung

$$\mathscr{F}[\mathcal{R}](X) = \mathcal{R}(X - EX) = \mathcal{R}(X) + EX \tag{1.41}$$

für alle $X \in L^2(\Omega)$ impliziert. Dies wollen wir wie folgt ausnutzen: Denn wir haben bereits in Bemerkung (1.3.2) c) festgestellt, dass $\mathcal{R}(X) \geq E[-X]$ für alle $X \in L^2(\Omega)$ gilt, sodass wir insgesamt die Ungleichung

$$\mathscr{F}[\mathcal{R}](X) = \mathcal{R}(X) + EX \geq E[-X] + EX = 0$$

für alle $X \in L^2(\Omega)$ erhalten, d.h. $\mathscr{F}[\mathcal{R}]$ ist nicht-negativ. Darüber hinaus erfüllt $\mathscr{F}[\mathcal{R}]$ das Axiom ($\mathcal{D}1'$) in Lemma (1.1.4) wegen

$$\mathscr{F}[\mathcal{R}](C) \underset{(1.41)}{=} \mathcal{R}(C) + E[C] \underset{(\mathcal{R}1')}{=} -C + C = 0$$

für alle $C \in \mathbb{R}$. Die Charakteristika ($\mathcal{D}2'$) und ($\mathcal{D}3$) ergeben sich unmittelbar aus ($\mathcal{R}2'$) und ($\mathcal{R}3$), wenn wir zusätzlich die Linearität von $X \mapsto E[X]$ auf dem Raum $L^2(\Omega)$ sowie die Identität (1.41) benutzen. Ferner ist

$$\mathscr{F}[\mathcal{R}](X) \underset{(1.41)}{=} \mathcal{R}(X) + EX \underset{(\mathcal{R}5)}{>} E[-X] + EX = 0$$

für alle nicht-konstanten $X \in L^2(\Omega)$, sodass wir auch ($\mathcal{D}4$) nachgewiesen haben. Daher liefert das Lemma (1.1.4), dass $\mathscr{F}[\mathcal{R}]$ ein Abweichungsmaß ist.

Zu $\mathscr{G}$*:* Sei nun $\mathcal{D} \in \mathscr{D}$. Zuerst ergibt sich wegen

$$\mathscr{G}[\mathcal{D}](C) = \mathcal{D}(C) + E[-C] \underset{(\mathcal{D}1')}{=} E[-C] = -C$$

für alle $C \in \mathbb{R}$, dass $\mathscr{G}[\mathcal{D}]$ der Eigenschaft ($\mathcal{R}1'$) genügt. Die Axiome ($\mathcal{R}2'$) und ($\mathcal{R}3$) sind eine direkte Konsequenz aus ($\mathcal{D}2'$) und ($\mathcal{D}3$) sowie der Linearität von $X \mapsto E[-X]$ auf dem Raum $L^2(\Omega)$. Schließlich erhalten wir für jedes nicht-konstante $X \in L^2(\Omega)$ die Ungleichung

$$\mathscr{G}[\mathcal{D}](X) = \mathcal{D}(X) + E[-X] \underset{(\mathcal{D}4)}{>} E[-X],$$

d.h. $(\mathcal{R}5)$ ist ebenfalls erfüllt. Nach Lemma (1.3.4) können wir dann schließen, dass $\mathscr{G}[\mathcal{D}]$ ein strictly expectation bounded Risikomaß ist.
Insgesamt haben wir also nachgerechnet, dass $\mathscr{F}$ und $\mathscr{G}$ wohldefiniert sind.

Dass sich $\mathscr{F}$ und $\mathscr{G}$ obendrein invers zueinander verhalten, können wir gut anhand der Definition von $\mathscr{G}$ sowie der Darstellung (1.41) erkennen.

Abschließend widmen wir uns den Zusätzen:
(i) Sei $\mathcal{R} \in \mathscr{R}$ und $\mathcal{D} := \mathscr{F}[\mathcal{R}]$. Da $\mathscr{G}$ die Umkehrabbildung von $\mathscr{F}$ ist, folgt demnach $\mathcal{R} = \mathscr{G}[\mathcal{D}]$.

„ $\Rightarrow$ “ Sei $\mathcal{R}$ zusätzlich kohärent. Außerdem sei ein beliebiges $X \in L^2(\Omega)$ gegeben. Falls dann $\text{ess inf}\, X = -\infty$ gilt, folgt die triviale Abschätzung

$$\mathcal{D}(X) \leq \infty = EX - \text{ess inf}\, X.$$

Andernfalls ist $\text{ess inf}\, X \in \mathbb{R}$ und, da wir in Lemma (A.1.6) (iv) bereits gezeigt haben, dass $X \geq \text{ess inf}\, X$ gilt, ist also $X - \text{ess inf}\, X \geq 0$. Dies impliziert in Kombination mit $(\mathcal{R}2)$ und $(\mathcal{R}4)$, dass

$$\mathcal{R}(X - \text{ess inf}\, X) \leq 0 \tag{1.42}$$

ist, sodass wir schließlich

$$\begin{aligned}\mathcal{D}(X) \underset{(\mathcal{D}1)}{=} \mathcal{D}(X - \text{ess inf}\, X) &\underset{(1.41)}{=} \mathcal{R}(X - \text{ess inf}\, X) + E[X - \text{ess inf}\, X] \\ &\underset{(1.42)}{\leq} E[X - \text{ess inf}\, X] = EX - \text{ess inf}\, X\end{aligned}$$

erhalten. Aufgrund der beliebigen Wahl von X folgt die Lower range dominance von $\mathcal{D}$.

„ $\Leftarrow$ “ Sei $\mathcal{D}$ zusätzlich lower range dominated. Ferner seien $X, Y \in L^2(\Omega)$ mit $X \geq Y$ bzw. $X - Y \geq 0$. Nach Lemma (A.1.6) (viii) ist daher $\text{ess inf}\,(X - Y) \geq 0$. Unter Verwendung dieser Tatsache ergibt sich

$$\mathcal{D}(X - Y) \underset{(\mathcal{D}5)}{\leq} E[X - Y] - \text{ess inf}\,(X - Y) \leq E[X - Y]$$

und somit

$$\mathcal{R}(X - Y) = \mathcal{D}(X - Y) + E[-(X - Y)] \leq E[X - Y] + E[-(X - Y)] = 0. \tag{1.43}$$

Insgesamt schließen wir

$$\mathcal{R}(X) = \mathcal{R}\big(Y + (X - Y)\big) \underset{(\mathcal{R}3)}{\leq} \mathcal{R}(Y) + \mathcal{R}(X - Y) \underset{(1.43)}{\leq} \mathcal{R}(Y).$$

Da X, Y beliebig gewählt waren, haben wir gezeigt, dass $\mathcal{R}$ kohärent ist.

(ii) Dies folgt unmittelbar aus der Definition von $\mathscr{F}$ und $\mathscr{G}$ sowie der Darstellung (1.41) und der Tatsache, dass es sich bei $X \mapsto E[\pm X]$ auf dem Raum $L^2(\Omega)$ um stetige, lineare Funktionale handelt. Darüber hinaus ist Lemma (A.2.9) (i) im Falle der Unterhalbstetigkeit hilfreich. □

(1.3.7) Bemerkung
Der Satz (1.3.6) liefert uns also eine ganz einfache Möglichkeit, aus einem Abweichungsmaß durch Anwenden der Transformation $\mathscr{G}$ ein strictly expectation bounded Risikomaß zu generieren. Zum besseren Verständnis dieser Transformation empfehlen wir an dieser Stelle, diese konkret anhand der bis jetzt diskutierten Beispiele an Abweichungsmaßen aus Abschnitt 1.1 „auszurechnen“.
In Abschnitt 1.4 folgt dann noch eine Vielzahl weiterer Abweichungsmaße, die man ebenso zu Risikomaßen transformieren kann. ◇

Stattdessen wollen wir hier nur das bis jetzt einzige konkret betrachtete strictly expectation bounded Risikomaß aus Beispiel (1.3.3) b) aufgreifen.

(1.3.8) Beispiel
Das Funktional

$$\mathcal{R} : L^2(\Omega) \to (-\infty, \infty], \quad \mathcal{R}(X) = -\text{ess inf } X.$$

aus Beispiel (1.3.3) b) ist bekanntlich ein strictly expectation bounded und kohärentes Risikomaß, das immer unterhalbstetig, aber i.A. nicht stetig ist. Dann ist nach Satz (1.3.6)

$$\mathscr{F}[\mathcal{R}] : L^2(\Omega) \to [0, \infty], \quad \mathscr{F}[\mathcal{R}](X) = -\text{ess inf } (X - EX) = EX - \text{ess inf } X$$

ein lower range dominated Abweichungsmaß, das ebenfalls unterhalbstetig, aber i.A. nicht stetig ist.
Dies wussten wir aber auch schon vorher, da $\mathscr{F}[\mathcal{R}]$ dem Abweichungsmaß ψ_- aus Beispiel (1.1.3) c) entspricht. ◇

Wenn wir jetzt die Sätze (1.3.6) und (1.2.3) kombinieren, erhalten wir ohne zusätzliche Arbeit eine duale Charakterisierung von unterhalbstetigen, strictly expectation bounded Risikomaßen. Dabei geht diese Idee ebenfalls auf Rockafellar et al. zurück, die aber nur ganz kurz in [12, Theorem 2] darauf eingegangen sind.

(1.3.9) Korollar (Duale Charakterisierung von Risikomaßen)
Ein Funktional $\mathcal{R} : L^2(\Omega) \to (-\infty, \infty]$ ist genau dann ein unterhalbstetiges, strictly expectation bounded Risikomaß, wenn eine Menge $\mathcal{Q} \subseteq L^2(\Omega)$ existiert, sodass

$$\mathcal{R}(X) = -\inf_{Q \in \mathcal{Q}} E[XQ] \tag{1.44}$$

für alle $X \in L^2(\Omega)$ geschrieben werden kann und $\mathcal{Q}$ den Axiomen $(\mathcal{Q}1)$ bis $(\mathcal{Q}3)$ aus Satz (1.2.3) genügt. In dieser Situation ist $\mathcal{Q}$ bereits durch $(\mathcal{Q}1)$ sowie der Identität (1.44) eindeutig bestimmt und es gilt

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid \mathcal{R}(X) \geq -E[XQ] \quad \text{für alle } X \in L^2(\Omega)\}. \tag{1.45}$$

Zudem gilt:

(i) $\mathcal{R}$ ist genau dann kohärent, falls $\mathcal{Q}$ der Anforderung $(\mathcal{Q}4)$ genügt.

(ii) $\mathcal{R}$ ist genau dann endlich (und daher stetig), falls $\mathcal{Q}$ beschränkt ist. ◇

(1.3.10) Bemerkung
In der Situation von Satz (1.3.6) und Korollar (1.3.9) ist die zu einem unterhalbstetigen $\mathcal{R} \in \mathscr{R}$ eindeutig bestimmte Menge $\mathcal{Q}$ offensichtlich das risk envelope des korrespondierenden unterhalbstetigen Abweichungsmaßes $\mathscr{F}[\mathcal{R}]$. ◇

Dies veranschaulichen wir in folgendem

(1.3.11) Beispiel
In der Situation von Beispiel (1.3.8) bedeutet dies, dass das Korollar (1.3.9) auf $\mathcal{R}$ anwendbar ist, und die eindeutig bestimmte Menge $\mathcal{Q}$ dem risk envelope von $\mathscr{F}[\mathcal{R}] = \psi_-$ entspricht, das wir in Beispiel (1.2.6) b) berechnet haben, d.h. es ist

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid Q \geq 0 \quad \text{und} \quad EQ = 1\}.$$

Dort haben wir auch die genauen Eigenschaften von $\mathcal{Q}$ diskutiert. ◇

Vor Abschluss dieses Paragraphen stellen wir fest, dass wir den ausschließlich kohärenten Risikomaßen bis jetzt nicht sonderlich viel Beachtung geschenkt haben. Dies liegt hauptsächlich daran, dass wir nicht ohne Weiteres wie in Satz (1.3.6) eine Beziehung zu Abweichungsmaßen herstellen können, sodass wir jene auch nicht ohne zusätzliche Arbeit wie in Korollar (1.3.9) dual charkterisieren können. Aber genau diesen letzten Punkt wollen wir angehen, wobei ROCKAFELLAR ET AL. den Anlass dazu in [12, S.62] gegeben haben.

(1.3.12) Satz (Duale Charakterisierung von Risikomaßen II)
Ein Funktional $\mathcal{R} : L^2(\Omega) \to (-\infty, \infty]$ ist genau dann ein unterhalbstetiges, kohärentes Risikomaß, wenn eine Menge $\mathcal{Q} \subseteq L^2(\Omega)$ existiert, sodass

$$\mathcal{R}(X) = -\inf_{Q \in \mathcal{Q}} E[XQ] \tag{1.46}$$

für alle $X \in L^2(\Omega)$ geschrieben werden kann und $\mathcal{Q}$ den Axiomen $(\mathcal{Q}1)$, $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$ genügt. In dieser Situation ist $\mathcal{Q}$ bereits durch $(\mathcal{Q}1)$ sowie der Identität (1.46) eindeutig bestimmt und es gilt

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid \mathcal{R}(X) \geq -E[XQ] \quad \text{für alle } X \in L^2(\Omega)\}.$$

Darüber hinaus ist $\mathcal{R}$ in dieser Situation genau dann endlich (und daher stetig), falls $\mathcal{Q}$ beschränkt ist. ◇

Beweis
Für den Beweis halten wir zunächst folgende Beobachtungen fest:

a) Wir bezeichnen mit $\tilde{\mathscr{D}}$ die Menge aller Funktionale $\mathcal{D} : L^2(\Omega) \to (-\infty, \infty]$, die die Charakteristika $(\mathcal{D}1)$, $(\mathcal{D}2)$, $(\mathcal{D}3)$ und $(\mathcal{D}5)$ erfüllen, und mit $\tilde{\mathscr{R}}$ die Menge aller kohärenten Risikomaße und definieren dazu die Abbildungen

$$\tilde{\mathscr{F}} : \tilde{\mathscr{R}} \to \tilde{\mathscr{D}}, \quad \mathcal{R} \mapsto \left[\tilde{\mathscr{F}}[\mathcal{R}] : L^2(\Omega) \to (-\infty, \infty],\ \tilde{\mathscr{F}}[\mathcal{R}](X) := \mathcal{R}(X - EX)\right]$$

sowie

$$\tilde{\mathscr{G}} : \tilde{\mathscr{D}} \to \tilde{\mathscr{R}}, \quad \mathcal{D} \mapsto \left[\tilde{\mathscr{G}}[\mathcal{D}] : L^2(\Omega) \to (-\infty, \infty],\ \tilde{\mathscr{G}}[\mathcal{D}](X) := \mathcal{D}(X) + E[-X]\right].$$

Dann lässt sich wie im Beweis von Satz (1.3.6) zeigen, dass $\tilde{\mathscr{F}}$ und $\tilde{\mathscr{G}}$ wohldefiniert und invers zueinander sind.
Dabei betonen wir besonders, dass jedes $\mathcal{R} \in \tilde{\mathscr{R}}$ i.A. nicht mehr dem Axiom $(\mathcal{R}5)$ genügt und wir deshalb nicht mehr garantieren können, dass $\tilde{\mathscr{F}}[\mathcal{R}] \geq 0$ gilt; diese Eigenschaft ist hier aber auch gar nicht nötig.
Jedenfalls seien $\mathcal{D} \in \tilde{\mathscr{D}}$ und $\mathcal{R} \in \tilde{\mathscr{R}}$. So lässt sich immerhin noch (analog wie dort) zeigen, dass $\tilde{\mathscr{F}}[\mathcal{R}]$ den Eigenschaften $(\mathcal{D}1)$, $(\mathcal{D}2)$, $(\mathcal{D}3)$ und $(\mathcal{D}5)$ genügt, auch wenn $\mathcal{R}$ evtl. nicht $(\mathcal{R}5)$ erfüllt. Zudem können wir genau wie dort nachrechnen, dass $\tilde{\mathscr{G}}[\mathcal{D}]$ den Eigenschaften $(\mathcal{R}1)$ bis $(\mathcal{R}4)$ genügt, auch wenn $\mathcal{D}$ evtl. nicht $(\mathcal{D}4)$ erfüllt. Entsprechend überträgt sich hier ebenfalls die (Unterhalb-) Stetigkeit.

b) Bei Betrachtung des Beweises von Satz (1.2.3) lässt sich folgendes Resultat analog zeigen:

Ein Funktional $\mathcal{D} : L^2(\Omega) \to (-\infty, \infty]$ ist genau dann unterhalbstetig und erfüllt die Eigenschaften $(\mathcal{D}1)$, $(\mathcal{D}2)$, $(\mathcal{D}3)$ und $(\mathcal{D}5)$, wenn eine Menge $\mathcal{Q} \subseteq L^2(\Omega)$ existiert, sodass

$$\mathcal{D}(X) = EX - \inf_{Q \in \mathcal{Q}} E[XQ] \tag{1.47}$$

für alle $X \in L^2(\Omega)$ geschrieben werden kann und $\mathcal{Q}$ den Axiomen $(\mathcal{Q}1)$, $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$ genügt. In dieser Situation ist $\mathcal{Q}$ bereits durch $(\mathcal{Q}1)$ sowie der Identität (1.47) eindeutig bestimmt und es gilt

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid \mathcal{D}(X) \geq EX - E[XQ] \quad \text{für alle } X \in L^2(\Omega)\}.$$

Zudem ist $\mathcal{D}$ in dieser Situation genau dann endlich (und daher stetig), falls $\mathcal{Q}$ beschränkt ist.

Dazu lassen wir in jenem Beweis einfach Teil (2), d.h. die „Äquivalenz der Gültigkeiten von $(\mathcal{D}4)$ und $(\mathcal{Q}2)$", sowie den Zusatz (iii) weg. Da in den verbleibenden Beweisschritten die Nicht-Negativität von $\mathcal{D}$ nicht benötigt wird, erhalten wir unmittelbar den Nachweis des obigen Resultates.

Die Kombination von a) und b) ergibt direkt die Behauptung. □

(1.3.13) Bemerkung
Der Satz (1.3.12) ermöglicht es uns nun, ohne große Mühe unterhalbstetige kohärente Risikomaße zu konstruieren. Wir finden nämlich vergleichsweise schnell Mengen $\mathcal{Q} \subseteq L^2(\Omega)$, die die Axiome $(\mathcal{Q}1)$, $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$ erfüllen, wie wir im nachfolgenden Beispiel erkennen werden. Dazu sei noch erwähnt, dass das Axiom $(\mathcal{Q}2)$, die für $\mathcal{Q}$ in diesem Zusammenhang nicht zwingend zutreffen muss, diese Wahl aber offenbar deutlich einschränkt.
Jedoch enthält $\mathcal{Q}$ hier i.A. nicht mehr $Q \equiv 1$, da die strictly expectation boundedness von $\mathcal{R}$ dafür erforderlich ist; dies sieht man beispielsweise an der Darstellung von $\mathcal{Q}$ in (1.45), wenn man zusätzlich Bemerkung (1.3.2) c) benutzt. ◇

Schließlich präsentieren wir das angekündigte

(1.3.14) Beispiel
Sei $Q \in L^2(\Omega)$ mit $Q \geq 0$ sowie $EQ = 1$. Dazu definieren wir

$$\mathcal{R} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{R}(X) := -E[XQ],$$

sodass wir mit $\mathcal{Q} := \{Q\}$ sehen, dass die Identität

$$\mathcal{R}(X) = -E[XQ] = -\inf_{U \in \mathcal{Q}} E[XU]$$

für alle $X \in L^2(\Omega)$ gilt. Nach Satz (1.3.12) ist $\mathcal{R}$ demnach ein endliches, stetiges und kohärentes Risikomaß, da $\mathcal{Q}$ offensichtlich $(\mathcal{Q}1)$, $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$ erfüllt und beschränkt ist.
Wenn wir jetzt z.B. $Q \equiv 1$ wählen, erhalten wir das Beispiel (1.3.3) a).

Wir können aber sogar noch einen Schritt weitergehen und zu $Q_1, \ldots, Q_n \in L^2(\Omega)$ mit $Q_i \geq 0$ und $EQ_i = 1$ für $i = 1, \ldots n$ die Menge

$$\mathcal{Q} := \operatorname{conv}\{Q_1, \ldots, Q_n\}$$

heranziehen; denn mittels einfacher Rechnungen erhält man, dass dieses $\mathcal{Q}$ ebenfalls die Axiome $(\mathcal{Q}1)$, $(\mathcal{Q}3)$ und $(\mathcal{Q}4)$ erfüllt (und obendrein beschränkt ist).

Jedenfalls ist in beiden Situationen vermutlich $(\mathcal{Q}2)$ i.A. nicht erfüllt. ◇

1.4 Fehlerfunktionale

Die Relevanz von erarbeiteten theoretischen Aussagen steht und fällt mit der Fülle von Beispielen, auf die sie angewendet werden kann. Jedoch haben wir bis jetzt nur wenige bzw. sehr verschiedene Bespiele von Abweichungsmaßen diskutiert. Daher ist es unser Ziel in diesem Paragraphen, eine (oder mehrere) große Klassen von Abweichungsmaßen (und damit automatisch auch Risikomaßen) zu erschließen. Dabei werden wir auch einige Zusammenhänge zwischen den bisherigen Beispielen beobachten. Als Schlüssel-Instrument dienen dabei die „Fehlerfunktionale", die ROCKAFELLAR ET AL. in [12, S. 66] definierten und anschließend untersuchten.

(1.4.1) Definition (Fehlerfunktionale)
Seien $a, b \geq 0$ und $p \in [1, \infty]$. Dann heißt

$$\mathcal{E}_{a,b,p} : L^2(\Omega) \to [0, \infty], \quad \mathcal{E}_{a,b,p}(X) := \|aX_+ + bX_-\|_p$$

Fehlerfunktional mit Parametern a, b und p. ◇

(1.4.2) Bemerkung

a) Zu $X \in L^2(\Omega)$ ist nach Lemma (A.1.24) (iv) ebenfalls $aX_+ + bX_- \in L^2(\Omega)$.

b) Für $p \in [1, \infty]$ ergibt sich im Spezialfall $a = b \geq 0$ nach Lemma (A.1.24) (ii), dass

$$\mathcal{E}_{a,b,p}(X) = a \cdot \|X_+ + X_-\|_p = a \cdot \||X|\|_p = a \cdot \|X\|_p$$

für alle $X \in L^2(\Omega)$ gilt.

c) Der Name Fehlerfunktional kommt daher, dass $\mathcal{E}_{a,b,p}$ das Ausmaß, inwieweit sich eine Zufallsvariable $X \in L^2(\Omega)$ von Null unterscheidet, bewertet. Je mehr sie also in speziellen Gesichtspunkten von Null abweicht, desto größer wird der Wert $\mathcal{E}_{a,b,p}(X)$, der auch gelegentlich als „Fehler" bezeichnet wird.
Durch Variation von a, b und p können die Werte von X auf manchen Bereichen, wie $\{X \leq -1\}$ und $\{-1 < X \leq 0\}$ sowie $\{0 < X < 1\}$ und $\{X \geq 1\}$, z.T. beliebig gewichtet werden. ◇

Mithilfe dieser konstruierten ROCKAFELLAR ET AL. auf zwei verschiedene Arten Abweichungsmaße. Wir beschränken uns hier aber lediglich darauf, einzusehen, wie sie konstruiert werden, und zu wissen, welche Charakteristika sie im Einzelnen aufweisen. Für die Beweise, die größtenteils ziemlich technisch sind, verweisen wir jeweils auf die Originalliteratur.

Zuallererst stellen wir nun einige Eigenschaften der Fehlerfunktionale zusammen. Dabei beginnen wir mit ganz elementaren Eigenschaften - beruhend auf [12, Proposition 6].

(1.4.3) Lemma
Seien $a, b \geq 0$ und $p \in [1, \infty]$. Dann gilt:

(i) $\mathcal{E}_{a,b,p}$ erfüllt $(\mathcal{D}2)$ und $(\mathcal{D}3)$.
(ii) $\mathcal{E}_{a,b,p}$ ist unterhalbstetig.

(iii) Falls sogar $a, b > 0$ sind, ist $\mathcal{E}_{a,b,p}(X) > 0$ für alle $X \in L^2(\Omega) \setminus \{0\}$. ◇

(1.4.4) Bemerkung

a) Der Beweis dieses Lemmas ist (als Einziges in diesem Abschnitt) nicht schwer; insbesondere sind die Punkte (i) und (iii) beinahe trivial. Darüber hinaus könnte man (ii) aus Lemma (A.1.24) (iv) und Lemma (A.2.4) folgern.

b) Aufgrund dieses Lemmas verhält sich $\mathcal{E}_{a,b,p}$ für $a, b > 0$ beinahe wie eine Norm. Jedoch ist es i.A. nicht symmetrisch, wobei sich die Symmetrie im Spezialfall $a = b$ ergibt.
Außerdem ist sogar $\mathcal{E}_{a,b,p}(X) = \infty$ für $p > 2$ und $X \in L^2(\Omega)$ möglich; dies lässt sich aber ausschließen, falls der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich ist und man sich an Lemma (A.1.21) erinnert. Alternativ ist $\mathcal{E}_{a,b,p}$ bei Vorliegen eines allgemeinen Wahrscheinlichkeitsraumes auch endlich, falls wir $p \in [1, 2]$ wählen; dies folgt einfach aus der Inklusionskette (A.2) und Lemma (A.1.24) (iv).

c) Falls hingegen $a = 0$ oder $b = 0$ ist, ist $\mathcal{E}_{a,b,p}(X) = 0$ für $X \in L^2(\Omega) \setminus \{0\}$ möglich; denn wenn wir z.B. $a = 0$ annehmen, folgt für $X \equiv 1$, dass zwar $X \in L^2(\Omega) \setminus \{0\}$, aber dennoch $\mathcal{E}_{a,b,p}(X) = 0$ gilt.
Nichtsdestotrotz wird $\mathcal{E}_{a,b,p}$ auch in diesem Fall seinem Namen gerecht, da wir in unserer Vorstellung zulassen, dass ein Fehlerfunktional unter ganz verschiedenen Gesichtspunkten den Fehler einer Zufallsvariablen $X \in L^2(\Omega)$ zu Null bewertet. ◇

Aufgrund der Punkte (i) und (ii) ist insbesondere $\operatorname{dom}(\mathcal{E}_{a,b,p}) \neq \emptyset$, sodass wir den Satz (1.2.1) auf die Fehlerfunktionale anwenden können. In Kombination mit den Identitäten (1.19) und (1.20) in Korollar (1.2.2) zeigten Rockafellar et al. in [12, Proposition 7] folgendes Resultat.

(1.4.5) Lemma (Dualität von Fehlerfunktionalen)
Zu allen $a, b \geq 0$ und $p \in [1, \infty]$ existiert eine nicht-leere, abgeschlossene und konvexe Menge $\mathcal{B}_{a,b,p} \subseteq L^2(\Omega)$, sodass

$$\mathcal{E}_{a,b,p}(X) = \sup_{Y \in \mathcal{B}_{a,b,p}} E[XY]$$

für alle $X \in L^2(\Omega)$ geschrieben werden kann. Durch diese Eigenschaften ist $\mathcal{B}_{a,b,p}$ eindeutig bestimmt. Sei nun $q \in [1, \infty]$ so gewählt, dass $\frac{1}{p} + \frac{1}{q} = 1$ gilt. Dann besitzt $\mathcal{B}_{a,b,p}$ die folgenden Darstellungen:

(i) Falls $a, b > 0$ sind, ist

$$\mathcal{B}_{a,b,p} = \{Y \in L^2(\Omega) \mid \|a^{-1}Y_+ + b^{-1}Y_-\|_q \leq 1\}.$$

(ii) *Falls $a > 0$ und $b = 0$ ist, ist*

$$\mathcal{B}_{a,b,p} = \{Y \in L^2(\Omega) \mid Y \geq 0 \quad \textit{und} \quad \|Y\|_q \leq a\}.$$

(iii) *Falls $a = 0$ und $b > 0$ ist, ist*

$$\mathcal{B}_{a,b,p} = \{Y \in L^2(\Omega) \mid Y \leq 0 \quad \textit{und} \quad \|Y\|_q \leq b\}.$$

◇

(1.4.6) Bemerkung
Im obigen Lemma wird eine konkrete Darstellung von $\mathcal{B}_{a,b,p}$ für $a = b = 0$ unterschlagen. Aber wir sehen sofort, dass $\mathcal{B}_{a,b,p} = \{0\}$ gilt; denn $\{0\}$ ist nicht-leer, abgeschlosen und konvex und obendrein gilt

$$\mathcal{E}_{a,b,p}(X) = 0 = \sup_{Y \in \{0\}} E[XY]$$

für alle $X \in L^2(\Omega)$. Da die Existenz- und Eindeutigkeitsaussage aufgrund von Satz (1.2.1) ebenso für diesen Spezialfall zutrifft, gilt also die behauptete Gleichheit. ◇

Unter Benutzung dieser Hilfsmittel können wir uns jetzt der Konstruktion der Abweichungsmaße widmen. Die erste Variante dazu ist mit Kenntnis von Lemma (1.4.3) sogar ziemlich naheliegend. Die ursprüngliche Formulierung des anschließenden Satzes sowie ein möglicher Beweis findet sich in [12, Theorem 3].

(1.4.7) Satz
Seien $p \in [1, \infty]$ und $a, b \geq 0$ mit $a + b > 0$. Dann ist

$$\mathcal{D} : L^2(\Omega) \to [0, \infty], \quad \mathcal{D}(X) := \mathcal{E}_{a,b,p}(X - EX)$$

ein unterhalbstetiges Abweichungsmaß. Das zugehörige risk envelope $\mathcal{Q}$ ist gegeben durch

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid EQ = 1 \quad \textit{und} \quad C - Q \in \mathcal{B}_{a,b,p} \quad \textit{für ein } C \in \mathbb{R}\}.$$

Zudem gilt:

(i) *$\mathcal{D}$ ist genau dann endlich (und daher stetig), falls $\mathcal{E}_{a,b,p}$ endlich ist.*

(ii) *$\mathcal{D}$ ist lower range dominated, falls $a = 0$ und $b \leq 1$ ist bzw. falls $p = 1$ und $a + b \leq 1$ ist. Die Aussagen sind sogar äquivalent, falls der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell unendlich ist.* ◇

Zur Illustration betrachten wir einige Beispiele. Dabei werden erste Zusammenhänge zu den schon bekannten Beispielen erkennbar.

(1.4.8) Beispiele

a) Sei $p \in [1,\infty]$. Dazu definieren wir

$$\mathcal{D}^p : L^2(\Omega) \to [0,\infty], \quad \mathcal{D}^p(X) := \|X - EX\|_p.$$

Unter Verwendung von Bemerkung (1.4.2) b) sehen wir, dass $\mathcal{D}^p(X) = \mathcal{E}_{1,1,p}(X - EX)$ für alle $X \in L^2(\Omega)$ gilt. Nach Satz (1.4.7) ist $\mathcal{D}^p$ demnach ein unterhalbstetiges Abweichungsmaß mit zugehörigem risk envelope

$$\mathcal{Q}_p = \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad \|C - Q\|_q \leq 1 \quad \text{für ein } C \in \mathbb{R}\},$$

wobei $q \in [1,\infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$ ist. Dabei sei bemerkt, dass wir für die obige Darstellung von $\mathcal{Q}$ zusätzlich die Darstellung von $\mathcal{B}_{1,1,p}$ aus Lemma (1.4.5) (i) benutzt haben.
Bzgl. der Endlichkeit und Stetigkeit können wir folgendes aussagen: Bekanntlich ist die Abbildung $X \mapsto X - EX$ stetig auf dem Raum $L^2(\Omega)$. Zudem ist $\|\cdot\|_p$ ebenfalls stetig auf dem Raum $L^2(\Omega)$, falls $p \in [1,2]$ bzw. falls $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich und p beliebig gewählt ist, wie wir in Lemma (A.2.4) gesehen haben. Deshalb ist $\mathcal{D}^p$ genau in diesen Fällen endlich und stetig. Andernfalls ist $\|\cdot\|_p$ nach Lemma (A.2.4) unterhalbstetig, aber nicht stetig auf dem Raum $L^2(\Omega)$. Dann lässt sich (unter Benutzung der Beispielfolge aus (A.7) im Beweis des Zusatzes dieses Lemmas) zeigen, dass auch $\mathcal{D}^p$ nur unterhalbstetig und nicht stetig ist. Somit ist $\mathcal{D}^p$ nach Satz (1.1.5) dann aber auch nicht endlich.
Schließlich ist $\mathcal{D}^p$ nach Satz (1.4.7) über essentiell unendlichen Wahrscheinlichkeitsräumen für kein p lower range dominated.

Im Spezialfall $p = 2$ entspricht $\mathcal{D}^p$ der Standardabweichung σ. Zu dieser konnten wir aber - besonders im Hinblick auf das risk envelope - in Beispiel (1.2.6) a) genauere Aussagen treffen.
Außerdem wird $\mathcal{D}^1$ auch gelegentlich als **mittlere absolute Abweichung** bezeichnet.

b) Sei wieder $p \in [1,\infty]$. So definieren wir

$$\mathcal{D}^p_+ : L^2(\Omega) \to [0,\infty], \quad \mathcal{D}^p_+(X) := \|[X - EX]_+\|_p = \mathcal{E}_{1,0,p}(X - EX)$$

und

$$\mathcal{D}^p_- : L^2(\Omega) \to [0,\infty], \quad \mathcal{D}^p_-(X) := \|[X - EX]_-\|_p = \mathcal{E}_{0,1,p}(X - EX).$$

Somit sind nach Satz (1.4.7) sowohl $\mathcal{D}^p_+$ als auch $\mathcal{D}^p_-$ unterhalbstetige Abweichungsmaße. Außerdem erhalten wir für die risk envelopes die Identitäten

$$\mathcal{Q}^p_+ = \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad C - Q \geq 0,\ \|C - Q\|_q \leq 1 \text{ für ein } C \in \mathbb{R}\}$$

sowie

$$\mathcal{Q}^p_- = \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad C - Q \leq 0,\ \|C - Q\|_q \leq 1 \text{ für ein } C \in \mathbb{R}\},$$

wobei wir für $q \in [1,\infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$ erneut die Darstellung von $\mathcal{B}_{a,b,p}$ für $(a,b) = (1,0)$ bzw. $(a,b) = (0,1)$ aus Lemma (1.4.5) benutzt haben.

Es lässt sich sogar zeigen, dass wir

$$\mathcal{Q}_+^p = \{Q \in L^2(\Omega) \mid EQ = 1,\ \text{ess sup}\ Q < \infty,\ \|\text{ess sup}\ Q - Q\|_q \leq 1\}$$

und

$$\mathcal{Q}_-^p = \{Q \in L^2(\Omega) \mid EQ = 1,\ \text{ess inf}\ Q > -\infty,\ \|Q - \text{ess inf}\ Q\|_q \leq 1\}$$

schreiben können.

Dazu: Wir zeigen die Aussage nur für $\mathcal{Q}_+^p$; denn der Nachweis für $\mathcal{Q}_-^p$ verläuft analog.

„ $\supseteq$ “ Sei $Q \in L^2(\Omega)$ mit $EQ = 1, \text{ess sup}\ Q < \infty$ und $\|\text{ess sup}\ Q - Q\|_q \leq 1$. Da $\text{ess sup}\ Q - Q \geq 0$ nach Lemma (A.1.6) (iii) gilt, können wir $C := \text{ess sup}\ Q$ wählen und erhalten $Q \in \mathcal{Q}_+^p$.

„ $\subseteq$ “ Sei $Q \in \mathcal{Q}_+^p$, d.h. es existiert ein $C \in \mathbb{R}$ mit $C - Q \geq 0$ und $\|C - Q\|_q \leq 1$. Dies impliziert

$$\text{ess sup}\ Q \leq C < \infty,$$

das wiederum

$$0 \leq \text{ess sup}\ Q - Q \leq C - Q$$

liefert. Aufgrund der Monotonie von $\|\cdot\|_q$ folgt daher

$$\|\text{ess sup}\ Q - Q\|_q \leq \|C - Q\|_q \leq 1.$$

Damit ist die Gleichheit gezeigt.

Des Weiteren sind $\mathcal{D}_-^p$ und $\mathcal{D}_p^+$ endlich und stetig, falls $p \in [1,2]$ bzw. falls $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich und p beliebig gewählt ist; denn wie schon in a) begründet, ist $\|\cdot\|_p$ in genau diesen Fällen endlich und stetig auf dem Raum $L^2(\Omega)$, während $X \mapsto [X - EX]_-$ und $X \mapsto [X - EX]_+$ auf dem Raum $L^2(\Omega)$ u.A. nach Lemma (A.1.24) (iv) ohnehin stetig sind. In den übrigen Fällen ist $\|\cdot\|_p$ auf dem Raum $L^2(\Omega)$ nach Lemma (A.2.4) wieder nur unterhalbstetig und nicht stetig, sodass sich (erneut mithilfe der Beispielfolge aus (A.7) im Beweis des Zusatzes von Lemma (A.2.4) unter Verwendung von Lemma (A.1.24) (iii)) nachweisen lässt, dass die beiden Abweichungsmaße auch nur unterhalbstetig und nicht stetig sind. Daher sind sie nach Satz (1.1.5) aber auch nicht endlich.
Darüber hinaus ist $\mathcal{D}_-^p$ immer lower range dominated und, falls $p = 1$ ist, ist es auch $\mathcal{D}_+^p$. Unter der Einschränkung, dass $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell unendlich ist, liegt die Lower range dominance sogar nur in diesen Fällen vor.

Schließlich betrachten wir noch zwei Spezialfälle: Zunächst entsprechen $\mathcal{D}_+^2$ und $\mathcal{D}_-^2$ den Standard-Semi-Abweichungen σ_+ und σ_- aus Beispiel (1.1.3) b), deren risk envelopes bis jetzt noch zu bestimmen waren. Zudem erkennen wir, dass

$$\begin{aligned}\mathcal{D}_+^\infty(X) = \text{ess sup } |[X - EX]_+| = \text{ess sup } [X - EX]_+ &= \text{ess sup } X - EX \\ &= \psi_+(X)\end{aligned}$$

sowie

$$\begin{aligned}\mathcal{D}_-^\infty(X) = \text{ess sup } |[X - EX]_-| = \text{ess sup } [X - EX]_- &= EX - \text{ess inf } X \\ &= \psi_-(X)\end{aligned}$$

für alle $X \in L^2(\Omega)$ gilt. $\diamond$

Abschließend zitieren wir die zweite, weniger intuitive Variante, Abweichungsmaße aus Fehlerfunktionalen zu generieren, indem wir - salopp gesagt - „den Abstand einer Zufallsvariablen zur Konstantheit“ bestimmen. Diese basiert auf [12, Theorem 4].

(1.4.9) Satz
Seien $p \in [1, \infty]$ und $a, b > 0$. Dann ist

$$\mathcal{D} : L^2(\Omega) \to [0, \infty], \quad \mathcal{D}(X) := \inf_{C \in \mathbb{R}} \mathcal{E}_{a,b,p}(X - C)$$

ein unterhalbstetiges Abweichungsmaß und für jedes $X \in L^2(\Omega)$ wird obiges Infimum angenommen, d.h. für jedes $X \in L^2(\Omega)$ existiert ein $C_0 = C_0(X) \in \mathbb{R}$ mit

$$\mathcal{D}(X) = \mathcal{E}_{a,b,p}(X - C_0).$$

Das zugehörige risk envelope $\mathcal{Q}$ ist gegeben durch

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid EQ = 1 \quad \textit{und} \quad 1 - Q \in \mathcal{B}_{a,b,p}\}. \tag{1.48}$$

Zudem gilt:

(i) $\mathcal{D}$ ist genau dann endlich (und daher stetig), falls $\mathcal{E}_{a,b,p}$ endlich ist.

(ii) $\mathcal{D}$ ist lower range dominated, falls $p = 1$ und $a \leq 1$ ist. Die Aussagen sind sogar äquivalent, falls der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell unendlich ist. $\diamond$

Als erste Anwendung dient folgendes

(1.4.10) Beispiel
Gegeben sei

$$\mathcal{D} : L^2(\Omega) \to [0, \infty], \quad \mathcal{D}(X) := \inf_{C \in \mathbb{R}} \mathcal{E}_{1,1,2}(X - C) = \inf_{C \in \mathbb{R}} \|X - C\|_2.$$

Nach Satz (1.4.9) ist $\mathcal{D}$ demnach ein unterhalbstetiges Abweichungsmaß. Da $\mathcal{E}_{1,1,2}$ offensichtlich endlich ist, folgt daraus sogar die Endlich- und Stetigkeit von $\mathcal{D}$. Obendrein ist $\mathcal{D}$ über essentiell unendlichen Wahrscheinlichkeitsräumen nicht lower range dominated. Jedenfalls zeigen wir nun, dass die Identität

$$\mathcal{D}(X) = \|X - EX\|_2 = \sigma(X)$$

für alle $X \in L^2(\Omega)$ zutrifft.

Dazu: Zu festem $X \in L^2(\Omega)$ sei

$$\phi : \mathbb{R} \to [0, \infty), \quad \phi(C) := \|X - C\|_2.$$

Um einen Minimierer von ϕ bzw. ϕ^2 zu finden, schreiben wir

$$\begin{aligned}\phi^2(C) = E[(X - C)^2] = E[X^2 - 2CX + C^2] &= C^2 - 2C \cdot EX + E[X^2] \\ &= (C - EX)^2 - (EX)^2 + E[X^2].\end{aligned}$$

Somit ist $C = EX$ sogar der eindeutige Minimierer von ϕ^2 und daher auch von ϕ. Daraus folgt die Behauptung.

Mit dem obigen Satz erhalten wir also für das risk envelope $\mathcal{Q}$ von σ aus (1.48) die Identität

$$\begin{aligned}\mathcal{Q} &= \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad 1 - Q \in \mathcal{B}_{1,1,2}\} \\ &= \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad \|1 - Q\|_2 \leq 1\} \\ &= \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad \sigma(1 - Q) \leq 1\},\end{aligned}$$

wobei wir für die zweite Gleichheit wiederum die Darstellung von $\mathcal{B}_{1,1,2}$ aus Lemma (1.4.5) sowie Bemerkung (1.4.2) b) verwendet haben. Diese ist konkreter als in Beispiel (1.4.8) a) und haben wir auch schon durch direktes Nachrechnen in Beispiel (1.2.6) a) gefunden. $\diamond$

An dieser Stelle haben wir jetzt alle theoretischen Aussagen soweit zusammen gestellt und wollen jetzt zum eher anwendungsorientierten Teil übergehen. Dabei wird dieser Satz ziemlich am Anfang, also im Zusammenhang mit dem „Conditional Value-at-Risk“, noch eine zentrale Rolle einnehmen.

2 Die Rolle von Abweichungsmaßen in der Finanzmathematik

Wie eingangs beschrieben, fokussieren wir uns in diesem Kapitel eher darauf aufzuzeigen, inwieweit die Theorie aus dem ersten Kapitel in der Praxis nützlich ist.

2.1 Der Conditional Value-at-Risk

Zu Anfang stellen wir fest, dass sich seit geraumer Zeit in der Risiko-Analysis im Aktienhandel, zwei Funktionale, zum Einen der sogenannte „Value-at-Risk“ und zum Anderen der „Conditional Value-at-Risk“, etabliert haben. Da besonders das Letztere in das Konzept des vorigen Kapitels passt und gleichzeitig in engem Verhältnis zum Ersteren steht, wollen wir die beiden in diesem Abschnitt näher untersuchen.

Dazu benutzen wir hauptsächlich die Konzepte von Acerbi und Tasche in [1]; denn in der Literatur gibt es diesbezüglich leicht verschiedene Definitionen. Wir beginnen mit der formalen Definition des „Value-at-Risks“, die gerade [1, Definition 2.2] entspricht. Da wir insbesondere darin mit Verteilungsfunktionen arbeiten, findet sich in Definition (A.1.2) u.A. unsere Notation für Verteilungsfunktionen.

(2.1.1) Definition (Value-at-Risk)
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $\alpha \in (0,1)$. Dann heißt

$$VaR_\alpha(X) := -inf\,\{z \in \mathbb{R} \mid F^X(z) > \alpha\}$$

der Value-at-Risk von X zum Niveau α. ◇

(2.1.2) Bemerkung

a) Bekanntlich gilt für die Verteilungsfunktion F^X, dass

$$\lim_{z\to-\infty} F^X(z) = 0 \qquad \text{und} \qquad \lim_{z\to\infty} F^X(z) = 1$$

ist, sodass $\mathrm{VaR}_\alpha(X) \in \mathbb{R}$ für jedes $\alpha \in (0,1)$ ist.

b) Anhand der Definition des Value-at-Risks sehen wir auch, dass dieser sogar nur von der Verteilungsfunktion von X abhängig ist. $\diamond$

An dieser Stelle verzichten wir auf eine präzise finanzmathematische Erklärung, wofür der Value-at-Risk genau verwendet wird. Stattdessen versuchen wir uns selbst an einer Interpretation. Daher verschaffen wir uns mit den beiden folgenden Lemmata einen ersten Eindruck.

(2.1.3) Lemma

Sei mit $X : \Omega \to \mathbb{R}$ eine Zufallsvariable gegeben, sodass die Verteilungsfunktion F^X stetig und streng monoton wachsend ist. Dann ist auch F^{-X} stetig und streng monoton wachsend, sodass sie eine Umkehrfunktion besitzt. In dieser Situation gilt

$$VaR_\alpha(X) = \left(F^{-X}\right)^{-1}(1-\alpha) = -\left(F^X\right)^{-1}(\alpha)$$

für jedes $\alpha \in (0,1)$. $\diamond$

Beweis

Zunächst ist

$$\mathcal{P}(\{X = z\}) = \mathcal{P}^X(\{z\}) = 0$$

für alle $z \in \mathbb{R}$, da F^X stetig ist. Daher ergibt sich

$$\begin{aligned} F^{-X}(z) = \mathcal{P}(\{-X \leq z\}) = \mathcal{P}(\{X \geq -z\}) &= 1 - \mathcal{P}(\{X < -z\}) \\ &= 1 - \mathcal{P}(\{X \leq -z\}) \\ &= 1 - F^X(-z) \end{aligned} \tag{2.1}$$

für alle $z \in \mathbb{R}$. Daraus folgt, dass auch F^{-X} streng monoton wachsend und stetig ist, da dies auch auf F^X zutrifft, sodass F^{-X} eine Umkehrfunktion $\left(F^{-X}\right)^{-1} : (0,1) \to \mathbb{R}$ besitzt.

Sei nun $\alpha \in (0,1)$. Dann erhalten wir mit (2.1) die Identität

$$\begin{aligned} \mathrm{VaR}_\alpha(X) = -\inf\{z \in \mathbb{R} \mid F^X(z) > \alpha\} &= \sup\{-z \in \mathbb{R} \mid F^X(z) > \alpha\} \\ &= \sup\{y \in \mathbb{R} \mid F^X(-y) > \alpha\} \\ &= \sup\{y \in \mathbb{R} \mid 1 - F^{-X}(y) > \alpha\} \\ &= \sup\{y \in \mathbb{R} \mid F^{-X}(y) < 1 - \alpha\} \end{aligned}$$

gilt. Da F^{-X} streng monoton wachsend und stetig ist sowie $1 - \alpha \in (0,1)$ ist, überlegt man sich nun leicht, dass

$$\sup\{y \in \mathbb{R} \mid F^{-X}(y) < 1 - \alpha\} = \left(F^{-X}\right)^{-1}(1-\alpha)$$

ist. Mit demselben Argument folgt

$$\inf\{z \in \mathbb{R} \mid F^X(z) > \alpha\} = \left(F^X\right)^{-1}(\alpha),$$

womit wir die Behauptung nachgewiesen haben. $\square$

(2.1.4) Lemma
Sei $\alpha \in (0,1)$. Dann gilt:

(i) $VaR_\alpha(0) = 0$,

(ii) $VaR_\alpha(X) \geq VaR_\alpha(Y)$ *für alle Zufallsvariablen* $X, Y : \Omega \to \mathbb{R}$ *mit* $X \leq Y$. $\diamond$

Beweis
(i) Die Zufallsvariable $X \equiv 0$ besitzt die Verteilungsfunktion

$$F^X : \mathbb{R} \to [0,1], \quad F^X(z) = \begin{cases} 0 & \text{für } z < 0, \\ 1 & \text{für } z \geq 0. \end{cases}$$

Da $\alpha > 0$ ist, folgt

$$-\mathrm{VaR}_\alpha(X) = \inf\left\{z \in \mathbb{R} \mid F^X(z) > \alpha\right\} = 0$$

und somit $\mathrm{VaR}_\alpha(X) = 0$.

(ii) Seien $X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen mit $X \leq Y$. Somit ist

$$\{Y \leq z\} \subseteq \{X \leq z\}$$

und deshalb $F^Y(z) \leq F^X(z)$ für alle $z \in \mathbb{R}$. Dies impliziert

$$\left\{z \in \mathbb{R} \mid F^Y(z) > \alpha\right\} \subseteq \left\{z \in \mathbb{R} \mid F^X(z) > \alpha\right\}$$

und damit ergibt sich wiederum

$$-\mathrm{VaR}_\alpha(Y) = \inf\left\{z \in \mathbb{R} \mid F^Y(z) > \alpha\right\} \geq \inf\left\{z \in \mathbb{R} \mid F^X(z) > \alpha\right\} = -\mathrm{VaR}_\alpha(X),$$

das die behauptete Ungleichung liefert. $\square$

(2.1.5) Bemerkung (Interpretation des Value-at-Risks)
Sei nun $X : \Omega \to \mathbb{R}$ eine beliebige Zufallsvariable, die wir wie in der Einführung von Abschnitt 1.3 interpretieren wollen; in dieser Situation wird X auch gelegentlich als „Gewinnfunktion“ bezeichnet, während sie entsprechend bei umgekehrter Interpretation ihrer Werte „Verlustfunktion“ genannt wird. Jedenfalls wollen wir ihr zu einem vorgegebenem Niveau $\alpha \in (0,1)$ mithilfe des Value-at-Risks einen gewissen Risikowert zuordnen. Aufgrund von Lemma (2.1.4) können wir ähnlich wie in Bemerkung (1.3.2) d) eine Aussage über diesen Risikowert treffen.
Unter den einschränkenden Annahmen an die Verteilungsfunktion wie in Lemma (2.1.3) entspricht $\mathrm{VaR}_\alpha(X)$ dann einfach der Stelle in $\mathbb{R}$, an der die Verteilungsfunktion der korrespondierenden Verlustfunktion $-X$ das Niveau $1-\alpha$ annimmt. $\diamond$

Obendrein geben uns die Charakteristika des Value-at-Risks aus Lemma (2.1.4) Grund zur Hoffnung, dass es sich bei dem Funktional

$$\mathcal{R} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{R}(X) := \mathrm{VaR}_\alpha(X)$$

um ein kohärentes Risikomaß im Sinne von Definition (1.3.1) handeln könnte. Dies ist aber i.A. nicht so, wie wir im anschließenden Beispiel feststellen werden; denn die Eigenschaft $(\mathcal{R}3)$ aus Definition (1.3.1) ist i.A. nicht erfüllt.

(2.1.6) Beispiel
Sei $\alpha \in \left[\frac{1}{2}, 1\right)$. Zusätzlich nehmen wir an, dass eine Menge $M \in \mathcal{M}$ mit

$$1 - \alpha \leq \mathcal{P}(M) \leq \alpha \tag{2.2}$$

existiert. Dazu definieren wir die Zufallsvariablen

$$X : \Omega \to \mathbb{R}, \quad X(\omega) := \mathbb{1}_M \qquad \text{sowie} \qquad Y := -X,$$

wobei $\mathbb{1}_M$ die charakteristische Funktion der Menge M sei. Dann erhalten wir einerseits

$$\mathrm{VaR}_\alpha(X+Y) = \mathrm{VaR}_\alpha(0) = 0$$

nach Lemma (2.1.4) (i). Andererseits erhalten wir die Verteilungsfunktionen $F^X, F^Y : \mathbb{R} \to [0,1]$ gegeben durch

$$F^X(z) = \begin{cases} 0 & \text{für } z < 0, \\ \mathcal{P}(M^c) & \text{für } 0 \leq z < 1, \\ 1 & \text{für } z \geq 1 \end{cases} \quad \text{bzw.} \quad F^Y(z) = \begin{cases} 0 & \text{für } z < -1, \\ \mathcal{P}(M) & \text{für } -1 \leq z < 0, \\ 1 & \text{für } z \geq 0. \end{cases}$$

Aufgrund der Annahme (2.2) ist neben $\mathcal{P}(M) \leq \alpha$ ebenso

$$\mathcal{P}(M^c) = 1 - \mathcal{P}(M) \leq 1 - (1-\alpha) = \alpha,$$

sodass direkt

$$\inf\{z \in \mathbb{R} \mid F^X(z) > \alpha\} = 1 \qquad \text{und} \qquad \inf\{z \in \mathbb{R} \mid F^Y(z) > \alpha\} = 0$$

folgt. Insgesamt ergibt sich

$$\mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y) = -1 < 0 = \mathrm{VaR}_\alpha(X+Y),$$

d.h. der Value-at-risk ist unter den obigen Einschränkungen nicht subadditiv. ◇

In der Praxis führt dieses Phänomen jedoch zu Problemen bzw. zu Risiko-Bewertungen, die nicht logisch erscheinen; ein Überblick über einige dies bezügliche Kritikpunkte findet sich in [6].

Jedenfalls gab es daher (oder vielleicht schon vorher bzw. parallel) einige Bemühungen, ein „Risikomaß“ mit schöneren Eigenschaften zu konstruieren, das aber dennoch gewisse Ähnlichkeiten zum Value-at-Risk vorweist. Aus diesen resultierte insbesondere der „Conditional Value-at-Risk“. Als Definition wählen wir die Darstellung ähnlich wie in [1, Definition 2.5], die auf den ersten Blick grundsätzlich verschieden zum Value-at-Risk zu sein scheint.

(2.1.7) Definition
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $\alpha \in (0,1)$. Dann heißt

$$CVaR_\alpha(X) := \inf_{C \in \mathbb{R}} \left\{ C + \alpha^{-1} E\big[[X+C]_-\big] \right\}$$

der Conditional Value-at-Risk *von X zum Niveau α.* ◇

(2.1.8) Bemerkung
Für jede Zufallsvariable $X : \Omega \to \mathbb{R}$ und $C \in \mathbb{R}$ ist $[X+C]_- \geq 0$ und somit quasi-integrierbar, d.h. $E\big[[X+C]_-\big]$ ist wohldefiniert und es gilt $E\big[[X+C]_-\big] \in [0,\infty]$. ◇

Bevor wir den Zusammenhang zum Value-at-Risk herstellen, wollen wir nun - als Motivation für diese ganze Sektion - aufzeigen, wie man aus dem Conditional Value-at-Risk ein Abweichungsmaß konstruieren kann. Mithilfe der in Kapitel 1 entwickelten Theorie lässt sich der Nachweis sehr elegant führen, wobei die Idee dazu von Rockafellar et al. stammt, siehe [12, S. 63, 72 - 73].

(2.1.9) Satz
Sei $\alpha \in (0,1)$. Dann ist das Funktional

$$\mathcal{D} : L^2(\Omega) \to [0,\infty), \quad \mathcal{D}(X) := CVaR_\alpha(X - EX)$$

ein endliches, stetiges und lower range dominated Abweichungsmaß. Das zugehörige risk envelope $\mathcal{Q}$ besitzt die Gestalt

$$\mathcal{Q} = \{ Q \in L^2(\Omega) \mid EQ = 1 \quad und \quad 0 \leq Q \leq \alpha^{-1} \}.$$

Zudem ist $\mathcal{D}$ sogar symmetrisch, falls $\alpha = \frac{1}{2}$ gewählt wird.

Das nach Satz (1.3.6) zu $\mathcal{D}$ korrespondierende Risikomaß ist gegeben durch

$$\mathcal{R} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{R}(X) = CVaR_\alpha(X)$$

und ist endlich, stetig, kohärent und strictly expectation bounded. ◇

Beweis
Sei $X \in L^2(\Omega)$ beliebig aber fest. Wegen $\alpha \in (0,1)$ ist $\alpha^{-1} - 1 > 0$, sodass wir dann

$$\begin{aligned}
\mathcal{D}(X) = \mathrm{CVaR}_\alpha(X - EX) &= \inf_{C \in \mathbb{R}} \big\{ C + \alpha^{-1} E\big[[(X - EX) + C]_- \big] \big\} \\
&= \inf_{C \in \mathbb{R}} \big\{ EX - (\underbrace{EX - C}_{=:\, C'}) + \alpha^{-1} E\big[[X - (\underbrace{EX - C}_{=\, C'})]_- \big] \big\} \\
&= \inf_{C' \in \mathbb{R}} \big\{ \underbrace{E[X - C']}_{= E\left[[X - C']_+\right] - E\left[[X - C']_-\right]} + \alpha^{-1} E\big[[X - C']_- \big] \big\} \\
&= \inf_{C' \in \mathbb{R}} \big\{ E\big[[X - C']_+ \big] + (\alpha^{-1} - 1) \cdot E\big[[X - C']_- \big] \big\} \\
&= \inf_{C' \in \mathbb{R}} E\big[[X - C']_+ + (\alpha^{-1} - 1) \cdot [X - C']_- \big] \\
&= \inf_{C' \in \mathbb{R}} \| [X - C']_+ + (\alpha^{-1} - 1) \cdot [X - C']_- \|_1 \\
&= \inf_{C' \in \mathbb{R}} \mathcal{E}_{1, \alpha^{-1} - 1, 1}(X - C')
\end{aligned}$$

schreiben können, wobei wir das Funktional $\mathcal{E}_{a,b,p}$ mit

$$a := 1 > 0, \qquad b := \alpha^{-1} - 1 > 0 \qquad \text{und} \qquad p := 1$$

in Definition (1.4.1) definiert haben. Somit ist Satz (1.4.9) anwendbar. Dieser liefert zunächst, dass $\mathcal{D}$ ein unterhalbstetiges Abweichungsmaß ist. Aufgrund von $p < 2$ impliziert die Hölder-Ungleichung, dass $\mathcal{E}_{a,b,p}$ ein endliches Funktional darstellt, sodass ebenfalls $\mathcal{D}$ endlich und deshalb stetig ist. Des Weiteren folgt aus $p = 1$ und $a = 1$, dass $\mathcal{D}$ sogar lower range dominated ist.

Zudem besitzt das risk envelope $\mathcal{Q}$ die Darstellung

$$\begin{aligned}
\mathcal{Q} &= \{ Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad 1 - Q \in \mathcal{B}_{a,b,p} \} \\
&= \{ Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad \| [1 - Q]_+ + (\alpha^{-1} - 1)^{-1} \cdot [1 - Q]_- \|_\infty \leq 1 \},
\end{aligned}$$

wobei wir für die zweite Gleichheit die Darstellung von $\mathcal{B}_{a,b,p}$ aus Lemma (1.4.5) (i) verwendet haben. Nachfolgend zeigen wir, dass

$$\mathcal{Q} = \{ Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad 0 \leq Q \leq \alpha^{-1} \}$$

gilt.

Dazu: „ $\subseteq$ “ Sei $Q \in \mathcal{Q}$, d.h. $EQ = 1$ und

$$\begin{aligned}
1 &\geq \| [1 - Q]_+ + (\alpha^{-1} - 1)^{-1} \cdot [1 - Q]_- \|_\infty \\
&= \max\big\{ \text{ess sup } \big([1 - Q]_+\big), (\alpha^{-1} - 1)^{-1} \cdot \text{ess sup } \big([1 - Q]_-\big) \big\} \\
&\geq (\alpha^{-1} - 1)^{-1} \cdot \text{ess sup } \big([1 - Q]_-\big) \\
&= (\alpha^{-1} - 1)^{-1} \cdot \max\big\{ 0, -\text{ess inf } (1 - Q) \big\} \\
&\geq (\alpha^{-1} - 1)^{-1} \cdot \big(-\text{ess inf } (1 - Q) \big),
\end{aligned}$$

das äquivalent zu

$$\alpha^{-1} - 1 \geq -\text{ess inf } (1 - Q) = -1 + \text{ess sup } Q \tag{2.3}$$

ist. Daraus folgt

$$\alpha^{-1} \geq \text{ess sup } Q \geq Q$$

und aufgrund der lower range dominance von $\mathcal{D}$ genügt $\mathcal{Q}$ dem Axiom $(\mathcal{Q}4)$ aus Satz (1.2.3), sodass auch $Q \geq 0$ ist. Insgesamt erhalten wir also

$$Q \in \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad 0 \leq Q \leq \alpha^{-1}\}.$$

„ $\supseteq$ " Sei $Q \in L^2(\Omega)$ mit $EQ = 1$ und $0 \leq Q \leq \alpha^{-1}$. Erneut können wir

$$\begin{aligned} &\|[1-Q]_+ + (\alpha^{-1}-1)^{-1} \cdot [1-Q]_-\|_\infty \\ &= \max\{\text{ess sup } \big([1-Q]_+\big), (\alpha^{-1}-1)^{-1} \cdot \text{ess sup } \big([1-Q]_-\big)\} \end{aligned} \tag{2.4}$$

schreiben. Wegen $0 \leq Q \leq \alpha^{-1}$ ist $1 - \alpha^{-1} \leq 1 - Q \leq 1$ und somit

$$[1-Q]_+ \leq 1 \qquad \text{bzw.} \qquad \text{ess sup } \big([1-Q]_+\big) \leq 1.$$

sowie

$$[1-Q]_- \leq \alpha^{-1} - 1 \qquad \text{bzw.} \qquad \text{ess sup } \big([1-Q]_-\big) \leq \alpha^{-1} - 1.$$

Dies impliziert zusammen mit der Identität (2.4), dass

$$\|[1-Q]_+ + (\alpha^{-1}-1)^{-1} \cdot [1-Q]_-\|_\infty \leq \max\{1, (\alpha^{-1}-1)^{-1} \cdot (\alpha^{-1}-1)\} = 1$$

gilt, d.h. es ist $Q \in \mathcal{Q}$.

Im Spezialfall $\alpha = \frac{1}{2}$ besitzt $\mathcal{D}$ das risk envelope

$$\mathcal{Q} = \{Q \in L^2(\Omega) \mid EQ = 1 \quad \text{und} \quad 0 \leq Q \leq 2\},$$

sodass

$$\begin{aligned} 1 - \mathcal{Q} &= \{Y \in L^2(\Omega) \mid E[1-Y] = 1 \quad \text{und} \quad 0 \leq 1 - Y \leq 2\} \\ &= \{Y \in L^2(\Omega) \mid EY = 0 \quad \text{und} \quad -1 \leq Y \leq 1\} \end{aligned}$$

gilt. Daher ist $1 - \mathcal{Q}$ offenbar kreisförmig, sodass $\mathcal{D}$ nach Satz (1.2.3) (iii) symmetrisch ist.

Sei nun wieder $\alpha \in (0,1)$ beliebig. Schließlich liefert Satz (1.3.6), dass das zu $\mathcal{D}$ korrespondierende Risikomaß

$$\mathcal{R} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{R}(X) = \text{CVaR}_\alpha(X - EX) - EX$$

endlich, stetig, kohärent und strictly expectation bounded ist. Es bleibt lediglich zu zeigen, dass

$$\text{CVaR}_\alpha(X + C) = \text{CVaR}_\alpha(X) - C$$

für alle Zufallsvariablen $X : \Omega \to \mathbb{R}$ und $C \in \mathbb{R}$ gilt; denn dann folgt

$$\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X)$$

für alle $X \in L^2(\Omega)$.
Sei also eine Zufallsvariable $X : \Omega \to \mathbb{R}$ und ein $C \in \mathbb{R}$ gegeben, dann ist

$$\begin{aligned}\mathrm{CVaR}_\alpha(X+C) &= \inf_{C' \in \mathbb{R}} \left\{C' + \alpha^{-1} E\left[[(X+C)+C']_-\right]\right\} \\ &= \inf_{C' \in \mathbb{R}} \left\{(\underbrace{C'+C}_{=:\,C''}) - C + \alpha^{-1} E\left[[X + (\underbrace{C+C'}_{=\,C''})]_-\right]\right\} \\ &= \inf_{C'' \in \mathbb{R}} \left\{C'' + \alpha^{-1} E\left[[X+C'']_-\right]\right\} - C \\ &= \mathrm{CVaR}_\alpha(X) - C.\end{aligned}$$

Damit haben wir alles gezeigt. □

Aufgrund dieses Satzes ist es nachvollziehbar, warum wir den Conditional Value-at-Risk hier näher untersuchen wollen. Bis jetzt fehlt aber die Verbindung zum Value-at-Risk. Deshalb streben wir als letztes Ziel dieses Abschnitts den Beweis folgenden Satzes an, den Acerbi und Tasche größtenteils in [1] gezeigt haben.

(2.1.10) Satz (Zusammenhang von CVaR und VaR)
Sei $\alpha \in (0,1)$ und $X \in L^1(\Omega)$. Dann gilt

$$CVaR_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha VaR_\beta(X)\,\mathrm{d}\beta. \tag{2.5}$$

◇

(2.1.11) Bemerkung (Interpretation des Conditional Value-at-Risks)
Anhand der Identität (2.5) sehen wir also, dass zu einer gegebenen Gewinnfunktion $X \in L^1(\Omega)$ und einem gegebenen Niveau $\alpha \in (0,1)$ der Wert $\mathrm{CVaR}_\alpha(X)$ gerade dem Mittelwert über alle Risikowerte $\mathrm{VaR}_\beta(X)$ zu den Niveaus $\beta \in (0,\alpha]$ entspricht. Der Vorteil ist offenbar, dass man dadurch evtl. eine bessere Einschätzung des Risikos, das sich hinter X verbirgt, erhält; denn der Value-at-Risk bestimmt bekanntlich jeweils **nur zu einem** gegebenen Niveau das Risiko. ◇

Den Beweis des Satzes (2.1.10) unterteilen wir in einzelne Schritte. In diesen werden wir ständig sogenannte „Quantilfunktionen" verwenden, die Acerbi und Tasche in [1, Definition 2.1] einführten.

(2.1.12) Definition (Quantilfunktionen)
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $\alpha \in (0,1)$. Dann wird

$$q_\alpha(X) := inf\,\{z \in \mathbb{R} \mid F^X(z) \geq \alpha\}$$

als das untere α-Quantil von X *und*

$$q^{\alpha}(X) := inf\,\{z \in \mathbb{R} \mid F^X(z) > \alpha\}$$

als das obere α-Quantil von X *bezeichnet.* ◇

(2.1.13) Bemerkung

a) Mit demselben Argument wie in Bemerkung (2.1.2) a) folgt, dass

$$q_{\alpha}(X) \in \mathbb{R} \qquad \text{und} \qquad q^{\alpha}(X) \in \mathbb{R}$$

für alle $\alpha \in (0,1)$ gilt.

b) Offensichtlich ist

$$\mathrm{VaR}_{\alpha}(X) = -q^{\alpha}(X).$$

c) Anhand obiger Definitionen lässt sich beobachten, dass

$$q_{\alpha}(X) = q^{\alpha}(X)$$

für fast alle $\alpha \in (0,1)$ gilt.

Dazu: Offenbar erhalten wir für alle $\alpha \in (0,1)$ die Ungleichung

$$q_{\alpha}(X) \leq q^{\alpha}(X),$$

sodass wir uns fragen, wann diese Ungleichung denn strikt ist. Dabei liegt folgende Vermutung nahe:

Zwischenbehauptung: Für $\alpha \in (0,1)$ ist die Ungleichung genau dann strikt, falls $y, z \in \mathbb{R}$ mit $y \neq z$ und

$$F^X(y) = F^X(z) = \alpha$$

existieren. Dies weisen wir nun nach:

„ $\Leftarrow$ “ Sei $\alpha \in (0,1)$, sodass solche $y, z \in \mathbb{R}$ wie oben beschrieben existieren; dabei sei Œ $y < z$. Die Monotonie von F^X liefert dann

$$F^X(y) = F^X(w) = F^X(z) = \alpha$$

für alle $w \in [y,z]$, das gerade

$$q_{\alpha}(X) \leq y < z \leq q^{\alpha}(X)$$

impliziert.

„ $\Rightarrow$ “ Sei jetzt $\alpha \in (0,1)$ mit $q_{\alpha}(X) < q^{\alpha}(X)$. So nehmen wir an, dass höchstens ein $y \in \mathbb{R}$ mit $F^X(y) = \alpha$ existiert. Dies führt zu zwei Fällen:

<u>1. Fall:</u> *Es existiert genau ein $y \in \mathbb{R}$ mit $F^X(y) = \alpha$.*
Mit der Monotonie von F^X erhalten wir dann

$$F^X(z) < \alpha \quad \text{für alle } z < y \qquad \text{und} \qquad F^X(z) > \alpha \quad \text{für alle } z > y,$$

sodass aber

$$q_\alpha(X) = y = q^\alpha(X)$$

gelten müsste.

<u>2. Fall:</u> *Es existiert kein $y \in \mathbb{R}$ mit $F^X(y) = \alpha$.*
In dieser Situation folgt aus der Monotonie sowie der rechtsseitigen Stetigkeit von F^X, dass ein $y \in \mathbb{R}$ mit

$$\lim_{z \nearrow y} F^X(z) \leq \alpha < F^X(y) \qquad \text{und} \qquad F^X(z) < \alpha \quad \text{für alle } z < y$$

existiert, d.h. α liegt in einer Sprungstelle von F^X. Zwangsläufig ergibt sich daraus wieder

$$q_\alpha(X) = y = q^\alpha(X).$$

In beiden Fällen erhalten wir somit einen Widerspruch, womit die Zwischenbehauptung nachgewiesen ist.

Jedenfalls sei jetzt

$$M := \{\alpha \in (0,1) \mid q_\alpha(X) < q^\alpha(X)\}.$$

So können wir wegen der Monotonie von F^X in Kombination mit der Zwischenbehauptung zu jedem $\alpha \in M$ ein maximales nicht-triviales Intervall $I_\alpha \subseteq \mathbb{R}$, also mit $\mathscr{L}(I_\alpha) > 0$ bzgl. des Lebesgue-Maßes $\mathscr{L}$, finden, sodass $F^X(z) = \alpha$ für alle und nur genau diese $z \in I_\alpha$ gilt. Dann wählen wir für alle $\alpha \in M$ genau ein $x_\alpha \in \mathbb{Q} \cap I_\alpha$, sodass offenbar für alle $\alpha, \beta \in M$ mit $\alpha < \beta$ auch $x_\alpha < x_\beta$ gilt. Somit haben wir eine Bijektion zwischen M und einer Teilmenge von $\mathbb{Q}$ erstellt, d.h. M ist abzählbar und somit eine Nullmenge bzgl. des Lebesgue-Maßes $\mathscr{L}$. Dies liefert die Behauptung. ◇

Darüber hinaus definierten Acerbi und Tasche in [1, Definition 2.6] den „Expected Shortfall“, den wir als ein zentrales Hilfsmittel benutzen werden.

(2.1.14) Definition (Expected Shortfall)
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $\alpha \in (0,1)$. Dann wird

$$ES_\alpha(X) := -\frac{1}{\alpha}\Big(E[X \cdot \mathbb{1}_{\{X \leq q_\alpha(X)\}}] + q_\alpha(X)\big(\alpha - \mathcal{P}(\{X \leq q_\alpha(X)\})\big)\Big)$$

als Expected Shortfall von X zum Niveau α *bezeichnet.* ◇

(2.1.15) Bemerkung
Da $X \cdot \mathbb{1}_{\{X \le q_\alpha(X)\}}$ offenbar durch $\max\{q_\alpha(X), 0\}$ nach oben beschränkt ist, ist

$$E\big[[X \cdot \mathbb{1}_{\{X \le q_\alpha(X)\}}]_+\big] \in [0, \infty).$$

Daraus folgt, dass $E[X \cdot \mathbb{1}_{\{X \le q_\alpha(X)\}}]$ wohldefiniert und

$$E[X \cdot \mathbb{1}_{\{X \le q_\alpha(X)\}}] \in [-\infty, \infty)$$

ist. ◇

Damit können wir auf das erste Resultat, das im Wesentlichen auf [1, Proposition 3.2] basiert, eingehen.

(2.1.16) Satz (Zusammenhang von Expected Shortfall und Value-at-Risk)
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable und $\alpha \in (0,1)$. Dann gilt

$$ES_\alpha(X) = \frac{1}{\alpha} \int_0^\alpha VaR_\beta(X)\, \mathrm{d}\beta.$$

◇

Beweis
Es genügt zu zeigen, dass

$$\mathrm{ES}_\alpha(X) = -\frac{1}{\alpha} \int_0^\alpha q_\beta(X)\, \mathrm{d}\beta \tag{2.6}$$

gilt; denn mit Bemerkung (2.1.13) b) und c) folgt dann

$$-\int_0^\alpha q_\beta(X)\, \mathrm{d}\beta = -\int_0^\alpha q^\beta(X)\, \mathrm{d}\beta = \int_0^\alpha \mathrm{VaR}_\beta(X)\, \mathrm{d}\beta$$

und damit die behauptete Identität.

Für den Nachweis von Gleichung (2.6) betrachten wir zusätzlich zu dem ohnehin vorliegenden Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ den Wahrscheinlichkeitsraum $(\Omega', \mathcal{M}', \mathcal{P}')$ gegeben durch

$$\Omega' := (0,1), \qquad \mathcal{M}' := \mathcal{B} \cap \mathfrak{P}\big((0,1)\big) \qquad \text{und} \qquad \mathcal{P}' := \mathscr{L},$$

wobei $\mathscr{L}$ das Lebesgue-Maß, $\mathcal{B}$ die Borel-Algebra auf $\mathbb{R}$ und $\mathfrak{P}\big((0,1)\big)$ die Potenzmenge von $(0,1)$ beschreiben. Dann definieren wir

$$Z : \Omega' \to \mathbb{R}, \quad Z(\beta) := q_\beta(X).$$

Da Z offenbar monoton wachsend ist, ist sie damit $\mathcal{P}'$-messbar, also eine Zufallsvariable.

1. Zwischenbehauptung: Die Verteilungen $(\mathcal{P}')^Z$ und $\mathcal{P}^X$ sind identisch.

Dazu: Aufgrund der Eigenschaften der Borel-Algebra auf $\mathbb{R}$ reicht es zu zeigen, dass die jeweiligen Verteilungsfunktionen übereinstimmen. Sei also $y \in \mathbb{R}$. Nachfolgend rechnen wir nach, dass

$$\{Z \leq y\} = \begin{cases} \emptyset & \text{für } F^X(y) = 0, \\ (0, F^X(y)] & \text{für } F^X(y) \in (0,1), \\ (0,1) = \Omega' & \text{für } F^X(y) = 1 \end{cases} \tag{2.7}$$

gilt.

<u>*1. Fall:*</u> $F^X(y) = 0$
Wir nehmen an, dass ein $\beta \in \Omega' = (0,1)$ mit $Z(\beta) \leq y$ existiert. Dann erhalten wir einerseits

$$F^X\big(Z(\beta)\big) \leq F^X(y) = 0$$

aufgrund der Monotonie von F^X und andererseits

$$0 < \beta \leq F^X\big(Z(\beta)\big)$$

wegen der rechtsseitigen Stetigkeit von F^X, also einen Widerspruch. Daher ist $\{Z \leq y\} = \emptyset$.

<u>*2. Fall:*</u> $F^X(y) \in (0,1)$
„ $\supseteq$ “ Wegen $F^X(y) \geq F^X(y)$ ist

$$Z\big(F^X(y)\big) = \inf\{z \in \mathbb{R} \mid F^X(z) \geq F^X(y)\} \leq y$$

und zusammen der Monotonie von Z ergibt sich $(0, F^X(y)] \subseteq \{Z \leq y\}$.

„ $\subseteq$ “ Dazu nehmen wir an, dass ein $\beta \in \big(F^X(y), 1\big)$ mit $Z(\beta) \leq y$ existiert. Wie im ersten Fall folgt dann aus der Monotonie sowie der rechtsseitigen Stetigkeit von F^X, dass

$$F^X(y) < \beta \leq F^X\big(Z(\beta)\big) \leq F^X(y)$$

gilt, das offenbar ein Widerspruch ist.
Daher ist $\{Z \leq y\} = (0, F^X(y)]$.

<u>*3. Fall:*</u> $F^X(y) = 1$
Sei $t \in \Omega' = (0,1)$. Dann ist

$$F^X(y) = 1 \geq t,$$

das $Z(t) \leq y$ und somit $t \in \{Z \leq y\}$ liefert. Aufgrund der beliebigen Wahl von t ist demnach $\{Z \leq y\} = (0,1) = \Omega'$.

Schließlich ergibt sich aus der Identität (2.7), dass in allen Fällen

$$F^Z(y) = \mathscr{L}(\{Z \leq y\}) = F^X(y)$$

gilt. Damit ist die 1. Zwischenbehauptung gezeigt.

2. Zwischenbehauptung: Die Verteilungen $(\mathcal{P}')^{Z \cdot \mathbb{1}_{\{Z \leq q_\alpha(X)\}}}$ und $\mathcal{P}^{X \cdot \mathbb{1}_{\{X \leq q_\alpha(X)\}}}$ sind identisch.

Dazu: Auch hier beschränken wir uns nur auf die Verteilungsfunktionen. Sei $y \in \mathbb{R}$. So unterscheiden wir zwei Fälle:

1. Fall: $y < 0$
Dann ist zum Einen

$$\{Z \cdot \mathbb{1}_{\{Z \leq q_\alpha(X)\}} \leq y\} = \{Z \leq q_\alpha(X)\} \cap \{Z \leq y\} = \{Z \leq \min\{q_\alpha(X), y\}\}$$

und zum Anderen

$$\{X \cdot \mathbb{1}_{\{X \leq q_\alpha(X)\}} \leq y\} = \{X \leq q_\alpha(X)\} \cap \{X \leq y\} = \{X \leq \min\{q_\alpha(X), y\}\}.$$

2. Fall: $y \geq 0$
In diesem Fall ist

$$\begin{aligned}\{Z \cdot \mathbb{1}_{\{Z \leq q_\alpha(X)\}} \leq y\} &= \big[\{Z \leq q_\alpha(X)\} \cap \{Z \leq y\}\big] \cup \{Z > q_\alpha(X)\} \\ &= \{Z \leq \min\{q_\alpha(X), y\}\} \cup \{Z \leq q_\alpha(X)\}^c\end{aligned}$$

und

$$\begin{aligned}\{X \cdot \mathbb{1}_{\{X \leq q_\alpha(X)\}} \leq y\} &= \big[\{X \leq q_\alpha(X)\} \cap \{X \leq y\}\big] \cup \{X > q_\alpha(X)\} \\ &= \{X \leq \min\{q_\alpha(X), y\}\} \cup \{X \leq q_\alpha(X)\}^c.\end{aligned}$$

In beiden Fällen sieht man direkt, dass

$$F^{Z \cdot \mathbb{1}_{\{Z \leq q_\alpha(X)\}}}(y) = F^{X \cdot \mathbb{1}_{\{X \leq q_\alpha(X)\}}}(y)$$

gilt, wenn man die 1. Zwischenbehauptung benutzt. Damit haben wir auch die 2. Zwischenbehauptung gezeigt.

Schließlich kommen wir zum wesentlichen Teil: Zunächst ist Z auf dem Intervall $(0, \alpha]$ für $\alpha \in (0, 1)$ quasi-integrierbar, da Z darauf nach oben beschränkt und monoton wachsend ist. Daher können wir

$$\int_0^\alpha q_\beta(X)\,\mathrm{d}\beta = \int_0^\alpha Z(\beta)\,\mathrm{d}\beta = \int_{\Omega'} Z(\beta) \cdot \mathbb{1}_{(0,\alpha]}(\beta)\,\mathrm{d}\beta = E[Z \cdot \mathbb{1}_{(0,\alpha]}] \tag{2.8}$$

schreiben. Wegen $Z(\alpha) = q_\alpha(X)$ und der Monotonie von Z ergibt sich

$$(0, \alpha] \subseteq \{Z \leq q_\alpha(X)\} \quad \text{und} \quad \{Z \leq q_\alpha(X)\} \cap (\alpha, 1) \subseteq \{Z = q_\alpha(X)\}. \tag{2.9}$$

Aus der ersten Inklusion folgt wiederum

$$(0,\alpha] = \{Z \leq q_\alpha(X)\} \setminus \big[\{Z \leq q_\alpha(X)\} \cap (\alpha,1)\big]$$

bzw.

$$\mathbb{1}_{(0,\alpha]} = \mathbb{1}_{\{Z\leq q_\alpha(X)\}} - \mathbb{1}_{\{Z\leq q_\alpha(X)\}\cap(\alpha,1)}. \tag{2.10}$$

Da $Z \cdot \mathbb{1}_{\{Z\leq q_\alpha(X)\}}$ wie in Bemerkung (2.1.15) quasi-integrierbar ist, leitet sich aus den Gleichungen (2.8), (2.9) und (2.10) die Identität

$$\begin{aligned}\int\limits_0^\alpha q_\beta(X)\,\mathrm{d}\beta &= E[Z\cdot\mathbb{1}_{\{Z\leq q_\alpha(X)\}}] - E[Z\cdot\mathbb{1}_{(\alpha,1)\cap\{Z\leq q_\alpha(X)\}}] \\ &= E[Z\cdot\mathbb{1}_{\{Z\leq q_\alpha(X)\}}] - q_\alpha(X)\cdot\mathscr{L}\big(\{Z\leq q_\alpha(X)\}\cap(\alpha,1)\big)\end{aligned} \tag{2.11}$$

ab. Zusammen mit der Gleichung (2.7) angewendet auf $y = q_\alpha(X)$ sowie der schon oft benutzten Ungleichung

$$F^X\big(q_\alpha(X)\big) = F^X\big(Z(\alpha)\big) \geq \alpha > 0$$

folgt dann, dass

$$\{Z \leq q_\alpha(X)\} \cap (\alpha,1) = \begin{cases} \big(\alpha, F^X\big(q_\alpha(X)\big)\big] & \text{für } F^X\big(q_\alpha(X)\big) \in (0,1), \\ (\alpha,1) & \text{für } F^X\big(q_\alpha(X)\big) = 1 \end{cases}$$

und somit in beiden Fällen

$$\mathscr{L}\big(\{Z\leq q_\alpha(X)\}\cap(\alpha,1)\big) = F^X\big(q_\alpha(X)\big) - \alpha = \mathcal{P}(\{X\leq q_\alpha(X)\}) - \alpha \tag{2.12}$$

ist. Zudem können wir aufgrund der Tatsache, dass $E[Z\cdot\mathbb{1}_{\{Z\leq q_\alpha(X)\}}]$ existiert, aus [8, Korollar 14.7]

$$E[Z\cdot\mathbb{1}_{\{Z\leq q_\alpha(X)\}}] = \int\limits_{\mathbb{R}} x\,\mathrm{d}(\mathcal{P}')^{Z\cdot\mathbb{1}_{\{Z\leq q_\alpha(X)\}}}(x) = \int\limits_{\mathbb{R}} x\,\mathrm{d}\mathcal{P}^{X\cdot\mathbb{1}_{\{X\leq q_\alpha(X)\}}}(x) \tag{2.13}$$

schließen, wobei wir für die zweite Gleichheit die 2. Zwischenbehauptung verwendet haben. Da jedoch auch $E[X\cdot\mathbb{1}_{\{X\leq q_\alpha(X)\}}]$ existiert, erhalten wir mittels erneuter Anwendung von [8, Korollar 14.7] zusammen mit Gleichung (2.13) sogar

$$E[Z\cdot\mathbb{1}_{\{Z\leq q_\alpha(X)\}}] = \int\limits_{\mathbb{R}} x\,\mathrm{d}\mathcal{P}^{X\cdot\mathbb{1}_{\{X\leq q_\alpha(X)\}}}(x) = E[X\cdot\mathbb{1}_{\{X\leq q_\alpha(X)\}}].$$

Dies liefert zusammen mit den Gleichungen (2.11) und (2.12) somit

$$\begin{aligned}\int\limits_0^\alpha q_\beta(X)\,\mathrm{d}\beta &= E[X\cdot\mathbb{1}_{\{X\leq q_\alpha(X)\}}] - q_\alpha(X)\cdot\big(\mathcal{P}(\{X\leq q_\alpha(X)\}) - \alpha\big) \\ &= E[X\cdot\mathbb{1}_{\{X\leq q_\alpha(X)\}}] + q_\alpha(X)\cdot\big(\alpha - \mathcal{P}(\{X\leq q_\alpha(X)\})\big).\end{aligned}$$

Wenn wir jetzt noch beide Seiten mit $-\frac{1}{\alpha}$ multiplizieren, sind wir fertig. $\square$

Nun müssen wir nur noch zeigen, dass der Expected Shortfall unter einschränkenden Bedingungen mit dem Conditional Value-at-Risk übereinstimmt. Dafür benötigen wir ein in der Wahrscheinlichkeitstheorie bekanntes Lemma. Die ursprüngliche Formulierung findet sich z.B. in [1, Proposition 4.2].

(2.1.17) Lemma
Sei $X \in L^1(\Omega)$ und $\alpha \in (0,1)$. Dann ist das Funktional

$$\mathcal{T}_\alpha : \mathbb{R} \to [0,\infty), \quad \mathcal{T}_\alpha(C) := \alpha \cdot E\big[[X-C]_+\big] + (1-\alpha) \cdot E\big[[X-C]_-\big]$$

konvex, endlich und stetig und es gilt

$$\lim_{|C|\to\infty} \mathcal{T}_\alpha(C) = \infty.$$

Daher nimmt $\mathcal{T}_\alpha$ sein Infimum an und die Menge der Minimierer ist gegeben durch

$$\big[q_\alpha(X), q^\alpha(X)\big].$$

◇

Beweis
Zuallererst erkennen wir, dass

$$\begin{aligned}\mathcal{T}_\alpha(C) &= \alpha \cdot E\big[[X-C]_+\big] + (1-\alpha) \cdot E\big[[X-C]_-\big] \\ &= E\big[\alpha \cdot [X-C]_+ + (1-\alpha) \cdot [X-C]_-\big] \\ &= \|\alpha \cdot [X-C]_+ + (1-\alpha) \cdot [X-C]_-\|_1 = \mathcal{E}_{\alpha,1-\alpha,1}(X-C)\end{aligned}$$

für alle $C \in \mathbb{R}$ gilt, sodass das Lemma (1.4.3) mit

$$a := \alpha, \qquad b := 1-\alpha \qquad \text{und} \qquad p := 1$$

wegen $\alpha \in (0,1)$ anwendbar ist. Aus diesem folgt unmittelbar, dass $\mathcal{T}_\alpha$ konvex und unterhalbstetig ist. Zudem ist es wegen $p < 2$ in Kombination mit der Hölder-Ungleichung endlich und nach Satz (1.1.7) also stetig. Darüber hinaus erhalten wir

$$\begin{aligned}\mathcal{T}_\alpha(C) &= \|\alpha \cdot [X-C]_+ + (1-\alpha) \cdot [X-C]_-\|_1 \\ &\geq \min\{\alpha, 1-\alpha\} \cdot \|[X-C]_+ + [X-C]_-\|_1 \\ &= \min\{\alpha, 1-\alpha\} \cdot \|X-C\|_1 \\ &\geq \underbrace{\min\{\alpha, 1-\alpha\}}_{>0} \cdot (|C| - \|X\|_1) \xrightarrow[|C|\to\infty]{} \infty\end{aligned}$$

unter Benutzung der Monotonie von $\|\cdot\|_1$ sowie der unteren Dreiecksungleichung.

Dann impliziert die Stetigkeit zusammen mit dem Verhalten von $\mathcal{T}_\alpha$ für $|C| \to \infty$, dass $\mathcal{T}_\alpha$ sein Infimum annimmt. Somit müssen wir nur noch die Minimierer bestimmen: Sei dazu $C \in \mathbb{R}$. Dann stellen wir zunächst fest, dass mit $X \in L^1(\Omega)$ auch

$X - C \in L^1(\Omega)$ bzw. $[X-C]_\pm \in L^1(\Omega)$ sind. Da zudem $[X-C]_\pm \geq 0$ sind, können wir bekanntlich

$$E\big[[X-C]_\pm\big] = \int_0^\infty 1 - F^{[X-C]_\pm}(z)\,\mathrm{d}z \tag{2.14}$$

schreiben; dies ist eine elementare Folgerung aus dem Satz von Tonelli, s. [8, Satz 15.11], die üblicherweise in der Vorlesung Stochastik II behandelt wird, sodass wir dies hier als bekannt voraussetzen.
Obendrein gilt für $z \geq 0$ die Identität

$$\begin{aligned} F^{[X-C]_-}(z) &= \mathcal{P}(\{[X-C]_- \leq z\}) \\ &= \mathcal{P}(\{\max\{0, C-X\} \leq z\}) \\ &= \mathcal{P}\Big(\{C-X<0\} \cup \big[\{C-X \geq 0\} \cap \{C-X \leq z\}\big]\Big) \\ &= \mathcal{P}(\{X>C\}) + \mathcal{P}(\{C-z \leq X \leq C\}) \\ &= 1 - F^X(C) + F^X(C) - \mathcal{P}(\{X < C-z\}) \\ &= 1 - \mathcal{P}(\{X < C-z\}). \end{aligned} \tag{2.15}$$

Da jedoch zum Einen

$$\mathcal{P}(\{X<y\}) = \lim_{t \nearrow y} F^X(t)$$

für alle $y \in \mathbb{R}$ gilt und zum Anderen F^X als monoton wachsende Funktion nur höchstens abzählbar viele Unstetigkeitsstellen besitzt, ist

$$\mathcal{P}(\{X<y\}) = F^X(y) \tag{2.16}$$

für fast alle $y \in \mathbb{R}$. Zusammen mit Gleichung (2.15) bedeutet dies, dass

$$F^{[X-C]_-}(z) = 1 - F^X(C-z)$$

für fast alle $z \geq 0$ gilt. Dies liefert wiederum in Kombination mit (2.14) die Identität

$$E\big[[X-C]_-\big] = \int_0^\infty 1 - \big(1 - F^X(C-z)\big)\,\mathrm{d}z = \int_0^\infty F^X(C-z)\,\mathrm{d}z.$$

Mit der Substitution $y = C - z$ und $\mathrm{d}y = -\mathrm{d}z$ erhalten wir dann

$$E\big[[X-C]_-\big] = -\int_C^{-\infty} F^X(y)\,\mathrm{d}y = \int_{-\infty}^C F^X(y)\,\mathrm{d}y. \tag{2.17}$$

Aus

$$[X-C]_+ = [C-X]_- = [(-X)-(-C)]_-$$

folgt dann unter Verwendung der Identität (2.17) mit $-X$ statt X bzw. $-C$ statt C die Gleichung

$$E\big[[X-C]_+\big] = E\big[[(-X)-(-C)]_-\big] = \int_{-\infty}^{-C} F^{-X}(y)\,\mathrm{d}y. \tag{2.18}$$

Da nun

$$F^{-X}(y) = \mathcal{P}(\{-X \le y\}) = \mathcal{P}(\{X \ge -y\}) = 1 - \mathcal{P}(\{X < -y)\})$$

für alle $y \in \mathbb{R}$ gilt, ergibt sich zusammen mit der für fast alle $y \in \mathbb{R}$ geltenden Identität (2.16), dass

$$F^{-X}(y) = 1 - F^X(-y)$$

für fast alle $y \in \mathbb{R}$ gilt, sodass wir in Kombination mit der Gleichung (2.18)

$$E\big[[X-C]_+\big] = \int_{-\infty}^{-C} 1 - F^X(-y)\,\mathrm{d}y$$

erhalten. Mit der Substitution $w = -y$ und $\mathrm{d}w = -\mathrm{d}y$ folgt dann

$$E\big[[X-C]_+\big] = -\int_{\infty}^{C} 1 - F^X(w)\,\mathrm{d}w = \int_{C}^{\infty} 1 - F^X(w)\,\mathrm{d}w. \tag{2.19}$$

Insgesamt ergibt sich aus den Gleichungen (2.17) und (2.19), dass

$$\mathcal{T}_\alpha(C) = \alpha \int_{C}^{\infty} 1 - F^X(z)\,\mathrm{d}z + (1-\alpha) \int_{-\infty}^{C} F^X(z)\,\mathrm{d}z \tag{2.20}$$

für jedes $C \in \mathbb{R}$ gilt.

Es können nun drei Fälle bzgl. des vorgegebenen α auftreten, die wir zum Teil schon in Bemerkung (2.1.13) c) diskutiert haben:

<u>1. Fall:</u> *Es existiert genau ein $y \in \mathbb{R}$ mit $F^X(y) = \alpha$.*
Dann gilt

$$F^X(z) < \alpha \quad \text{für alle } z < y \qquad \text{und} \qquad F^X(z) > \alpha \quad \text{für alle } z > y$$

und

$$q_\alpha(X) = y = q^\alpha(X).$$

<u>2. Fall:</u> *Es existiert kein $y \in \mathbb{R}$ mit $F^X(y) = \alpha$.*
So existiert ein $y \in \mathbb{R}$ mit

$$\lim_{z \nearrow y} F^X(z) \le \alpha < F^X(y) \qquad \text{und} \qquad F^X(z) < \alpha \quad \text{für alle } z < y,$$

sodass aufgrund der Monotonie von F^X natürlich auch

$$F^X(z) > \alpha \quad \text{für alle } z > y$$

ist, und es gilt

$$q_\alpha(X) = y = q^\alpha(X).$$

<u>3. Fall:</u> *Es existieren $y, z \in \mathbb{R}$ mit $y < z$ und $F^X(y) = F^X(z) = \alpha$.*
Dann gilt bekanntlich

$$q_\alpha(X) \leq y < z \leq q^\alpha(X). \tag{2.21}$$

Darüber hinaus folgt $F^X\big(q_\alpha(X)\big) \geq \alpha$ aus der rechtsseitigen Stetigkeit von F^X, sodass wir zusammen mit der Ungleichungskette (2.21) und der Monotonie von F^X

$$\alpha \leq F^X\big(q_\alpha(X)\big) \leq F^X(y) = \alpha,$$

also $F^X\big(q_\alpha(X)\big) = \alpha$, erhalten. Dies impliziert zusammen mit der Definition von $q^\alpha(X)$, dass

$$F^X(w) = \alpha \quad \text{für alle } w \in \big[q_\alpha(X), q^\alpha(X)\big)$$

gilt. Genauso ergibt sich unmittelbar aus der Definition von $q_\alpha(X)$, dass

$$F^X(w) < \alpha \quad \text{für alle } w < q_\alpha(X)$$

ist. Des Weiteren ist es zwar möglich, dass $F^X\big(q^\alpha(X)\big) = \alpha$ ist, aber wir können aus der Definition von $q^\alpha(X)$ definitiv schließen, dass

$$F^X(w) > \alpha \quad \text{für alle } w > q^\alpha(X)$$

gilt.

Unter Beachtung dieser drei Fälle können wir Folgendes bzgl. der Minimierer von $\mathcal{T}_\alpha$ beobachten:

Sei $C \in [q_\alpha(X), q^\alpha(X)]$. Mithilfe der Darstellung (2.20) von $\mathcal{T}_\alpha$ erkennen wir dann, dass

$$\mathcal{T}_\alpha(C) = \alpha \int_C^\infty 1 - F^X(z)\,\mathrm{d}z + (1-\alpha) \int_{-\infty}^C F^X(z)\,\mathrm{d}z$$

$$
\begin{aligned}
&= \alpha \int_{q^\alpha(X)}^{\infty} 1 - F^X(z)\,\mathrm{d}z + \alpha \int_{C}^{q^\alpha(X)} 1 - \underbrace{F^X(z)}_{=\,\alpha \text{ f.ü.}}\,\mathrm{d}z + (1-\alpha) \int_{-\infty}^{q_\alpha(X)} F^X(z)\,\mathrm{d}z \\
&\quad + (1-\alpha) \int_{q_\alpha(X)}^{C} \underbrace{F^X(z)}_{=\,\alpha \text{ f.ü.}}\,\mathrm{d}z \\
&= \alpha \int_{q^\alpha(X)}^{\infty} 1 - F^X(z)\,\mathrm{d}z + (1-\alpha) \int_{-\infty}^{q_\alpha(X)} F^X(z)\,\mathrm{d}z + \alpha(1-\alpha)\big(q^\alpha(X) - C\big) \\
&\quad + (1-\alpha)\alpha\big(C - q_\alpha(X)\big) \\
&= \alpha \underbrace{\int_{q^\alpha(X)}^{\infty} 1 - F^X(z)\,\mathrm{d}z}_{\leq\, E\left[[X-C]_+\right] \,\in\, [0,\infty)} + (1-\alpha) \underbrace{\int_{-\infty}^{q_\alpha(X)} F^X(z)\,\mathrm{d}z}_{\leq E\left[[X-C]_-\right] \,\in\, [0,\infty)} + \alpha(1-\alpha)\big(q^\alpha(X) - q_\alpha(X)\big) \\
&=: C_\alpha,
\end{aligned}
$$

also konstant, ist.

Sei nun $C \in \mathbb{R} \setminus [q_\alpha(X), q^\alpha(X)]$. Dabei nehmen wir Œ $C < q_\alpha(X)$ an; denn für $C > q^\alpha(X)$ verläuft die nachfolgende Rechnung analog. Jedenfalls erhalten wir wieder mithilfe der Darstellung (2.20) von $\mathcal{T}_\alpha$, dass

$$
\begin{aligned}
\mathcal{T}_\alpha(C) &= \alpha \int_{C}^{\infty} 1 - F^X(z)\,\mathrm{d}z + (1-\alpha) \int_{-\infty}^{C} F^X(z)\,\mathrm{d}z \\
&= \alpha \int_{q^\alpha(X)}^{\infty} 1 - F^X(z)\,\mathrm{d}z + \alpha \int_{q_\alpha(X)}^{q^\alpha(X)} 1 - F^X(z)\,\mathrm{d}z + \alpha \int_{C}^{q_\alpha(X)} 1 - F^X(z)\,\mathrm{d}z \\
&\quad + (1-\alpha) \int_{-\infty}^{C} F^X(z)\,\mathrm{d}z
\end{aligned}
\tag{2.22}
$$

ist. Jetzt wählen wir ein $C' \in \mathbb{R}$ mit

$$
C < C' < q_\alpha(X),
$$

so ergibt sich die Abschätzung

$$
\alpha \int_{C}^{q_\alpha(X)} 1 - F^X(z)\,\mathrm{d}z = \alpha \int_{C}^{C'} \underbrace{1 - F^X(z)}_{\geq\, 1 - F^X(C')}\,\mathrm{d}z + \alpha \int_{C'}^{q_\alpha(X)} \underbrace{1 - F^X(z)}_{\geq\, 1-\alpha \text{ f.ü.}}\,\mathrm{d}z
$$

$$
\begin{aligned}
&\geq \alpha\big(\underbrace{1-F^X(C')}_{>\,1-\alpha}\big)(C'-C)+\alpha(1-\alpha)\big(q_\alpha(X)-C'\big)\\
&> \alpha(1-\alpha)(C'-C)+\alpha(1-\alpha)\big(q_\alpha(X)-C'\big)\\
&= \alpha(1-\alpha)\big(q_\alpha(X)-C\big)\\
&= (1-\alpha)\int\limits_C^{q_\alpha(X)} \underbrace{\alpha}_{\geq\, F^X(z)\ \text{f.ü.}}\,\mathrm{d}z\\
&\geq (1-\alpha)\int\limits_C^{q_\alpha(X)} F^X(z)\,\mathrm{d}z.
\end{aligned}
$$

Zusammen mit Gleichung (2.22) impliziert dies

$$
\begin{aligned}
\mathcal{T}_\alpha(C) &> \alpha\int\limits_{q^\alpha(X)}^{\infty} 1-F^X(z)\,\mathrm{d}z+\alpha\int\limits_{q_\alpha(X)}^{q^\alpha(X)} \underbrace{1-F^X(z)}_{=\,1-\alpha\ \text{f.ü.}}\,\mathrm{d}z+(1-\alpha)\int\limits_C^{q_\alpha(X)} F^X(z)\,\mathrm{d}z\\
&\quad+(1-\alpha)\int\limits_{-\infty}^{C} F^X(z)\,\mathrm{d}z\\
&= \alpha\int\limits_{q^\alpha(X)}^{\infty} 1-F^X(z)\,\mathrm{d}z+\alpha(1-\alpha)\big(q^\alpha(X)-q_\alpha(X)\big)+(1-\alpha)\int\limits_{-\infty}^{q_\alpha(X)} F^X(z)\,\mathrm{d}z\\
&= C_\alpha.
\end{aligned}
$$

Demnach entspricht $\big[q_\alpha(X), q^\alpha(X)\big]$ wirklich der Menge der Minimierer von $\mathcal{T}_\alpha$. $\square$

Mithilfe dieser Hilfsaussagen können wir uns dem Beweis von Satz (2.1.10) widmen. Dabei beruht der letzte und somit entscheidende Schritt auf [1, Corollary 4.3].

Beweis von Satz (2.1.10)
Wegen $X \in L^1(\Omega)$ sind alle Voraussetzungen von Lemma (2.1.17) erfüllt, sodass wir für alle $C \in \mathbb{R}$ dann

$$
\begin{aligned}
\mathcal{T}_\alpha(C) &= \alpha\cdot E\big[[X-C]_+\big]+(1-\alpha)\cdot E\big[[X-C]_-\big]\\
&= E\big[[X-C]_-\big]+\alpha\big(E\big[[X-C]_+\big]-E\big[[X-C]_-\big]\big)\\
&= E\big[[X-C]_-\big]+\alpha E[X-C] \qquad (2.23)\\
&= E\big[[X-C]_-\big]-\alpha C+\alpha EX\\
&= \alpha\big(\alpha^{-1}\cdot E\big[[X-C]_-\big]-C\big)+\alpha EX
\end{aligned}
$$

schreiben können. Außerdem ist

$$\begin{aligned} E\big[[X-C]_-\big] &= \int_\Omega (C-X)\cdot \mathbb{1}_{\{C-X\geq 0\}}\,\mathrm{d}\mathcal{P} \\ &= \int_\Omega C\cdot \mathbb{1}_{\{X\leq C\}}\,\mathrm{d}\mathcal{P} - \int_\Omega X\cdot \mathbb{1}_{\{X\leq C\}}\,\mathrm{d}\mathcal{P} \\ &= C\cdot \mathcal{P}(\{X\leq C\}) - E[X\cdot \mathbb{1}_{\{X\leq C\}}], \end{aligned}$$

sodass sich zusammen mit der Identität (2.23)

$$\begin{aligned} \mathcal{T}_\alpha(C) &= \alpha\big(\alpha^{-1}\cdot C\cdot \mathcal{P}(\{X\leq C\}) - \alpha^{-1}\cdot E[X\cdot \mathbb{1}_{\{X\leq C\}}] - C\big) + \alpha EX \\ &= -\alpha\Big(\alpha^{-1}\cdot E[X\cdot \mathbb{1}_{\{X\leq C\}}] + \alpha^{-1}\cdot C\cdot \big(\alpha - \mathcal{P}(\{X\leq C\})\big)\Big) + \alpha EX \end{aligned} \tag{2.24}$$

ergibt. Da die Aussagen in Lemma (2.1.17) offenbar genauso für das Funktional

$$\mathcal{S}_\alpha := \alpha^{-1}\cdot(\mathcal{T}_\alpha - \alpha EX)$$

gelten, können wir also insbesondere festhalten, dass $\mathcal{S}_\alpha$ nach unten beschränkt ist und sein Infimum genau für $C\in\big[q_\alpha(X), q^\alpha(X)\big]$ annimmt, d.h.

$$\begin{aligned} \inf_{C'\in\mathbb{R}} \mathcal{S}_\alpha(C') &= \mathcal{S}_\alpha(C) \\ &= -\alpha^{-1}\Big(E[X\cdot \mathbb{1}_{\{X\leq C\}}] + C\cdot\big(\alpha - \mathcal{P}(\{X\leq C\})\big)\Big) \end{aligned} \tag{2.25}$$

für alle $C\in\big[q_\alpha(X), q^\alpha(X)\big]$, wobei wir für die zweite Gleichheit die Identität (2.24) verwendet haben. Wenn wir jetzt $C = q_\alpha(X)$ wählen, erhalten wir

$$\inf_{C'\in\mathbb{R}} \mathcal{S}_\alpha(C') = -\alpha^{-1}\Big(E[X\cdot \mathbb{1}_{\{X\leq q_\alpha(X)\}}] + q_\alpha(X)\big(\alpha - \mathcal{P}(\{X\leq q_\alpha(X)\})\big)\Big) = \mathrm{ES}_\alpha(X).$$

nach Definition (2.1.14). Darüber hinaus erkennen wir unter Verwendung der Gleichung (2.23), dass

$$\begin{aligned} \inf_{C'\in\mathbb{R}} \mathcal{S}_\alpha(C') &= \inf_{C'\in\mathbb{R}} \big\{\alpha^{-1}\cdot E\big[[X-C']_-\big] - C'\big\} \\ &= \inf_{C'\in\mathbb{R}} \big\{(\underbrace{-C'}_{=:\,C''}) + \alpha^{-1}\cdot E\big[[X+(\underbrace{-C'}_{=\,C''})]_-\big]\big\} \\ &= \inf_{C''\in\mathbb{R}} \big\{C'' + \alpha^{-1}\cdot E\big[[X+C'']_-\big]\big\} = \mathrm{CVaR}_\alpha(X) \end{aligned} \tag{2.26}$$

gilt. Schließlich folgt

$$\mathrm{CVaR}_\alpha(X) = \mathrm{ES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha VaR_\beta(X)\,\mathrm{d}\beta$$

mit Satz (2.1.16). □

Als Nebenresultat aus dem Beweis von Satz (2.1.10) erhalten wir noch folgendes

(2.1.18) Korollar (Vereinfachte Darstellung des CVaR)
Sei $X \in L^1(\Omega)$ und $\alpha \in (0,1)$. Falls die Verteilungsfunktion F^X stetig in $-VaR_\alpha(X)$ ist, gilt

$$CVaR_\alpha(X) = -\alpha^{-1} \cdot E[X \cdot \mathbb{1}_{\{X \leq -VaR_\alpha(X)\}}].$$

◇

Beweis
Zunächst wissen wir aufgrund der Gleichungen (2.25) und (2.26), dass

$$\mathrm{CVaR}_\alpha(X) = -\alpha^{-1}\Big(E[X \cdot \mathbb{1}_{\{X \leq C\}}] + C \cdot \big(\alpha - \mathcal{P}(\{X \leq C\})\big)\Big)$$

für alle $C \in \big[q_\alpha(X), q^\alpha(X)\big]$ gilt. Wenn wir jetzt

$$C = q^\alpha(X) = -\mathrm{VaR}_\alpha(X)$$

wählen, erhalten wir also

$$\begin{aligned}\mathrm{CVaR}_\alpha(X) = -\alpha^{-1}\Big(&E[X \cdot \mathbb{1}_{\{X \leq -\mathrm{VaR}_\alpha(X)\}}] \\ &- \mathrm{VaR}_\alpha(X) \cdot \big(\alpha - \mathcal{P}(\{X \leq -\mathrm{VaR}_\alpha(X)\})\big)\Big).\end{aligned} \tag{2.27}$$

Nun erinnern wir uns, dass zu dem vorgegebenem α einer der drei Fälle aus dem Beweis von Lemma (2.1.17) eintritt.
Dabei ist der 2. Fall jedoch nicht möglich; denn in dieser Situation ist F^X in

$$y = q^\alpha(X) = -\mathrm{VaR}_\alpha(X)$$

unstetig, das aber in der Voraussetzung dieses Korollars ausgeschlossen ist.
Demnach tritt entweder der 1. oder 3. Fall ein, wobei der erste gerade

$$\mathcal{P}(\{X \leq -\mathrm{VaR}_\alpha(X)\}) = F^X\big(q^\alpha(X)\big) = F^X(y) = \alpha$$

liefert. Falls hingegen der 3. Fall eintritt, gilt bekanntlich insbesondere

$$F^X(w) = \alpha \quad \text{für alle } w \in \big[q_\alpha(X), q^\alpha(X)\big) \qquad \text{und} \qquad F^X(w) > \alpha \quad \text{für alle } w > q^\alpha(X).$$

Wegen der Stetigkeit von F^X in $q^\alpha(X) = -\mathrm{VaR}_\alpha(X)$ muss demnach ebenfalls

$$\mathcal{P}(\{X \leq -\mathrm{VaR}_\alpha(X)\}) = F^X\big(q^\alpha(X)\big) = \alpha$$

gelten. Insgesamt ist also nur

$$\mathcal{P}(\{X \leq -\mathrm{VaR}_\alpha(X)\}) = \alpha$$

möglich, sodass dies zusammen mit der Identität (2.27) gerade

$$\mathrm{CVaR}_\alpha(X) = -\alpha^{-1} \cdot E[X \cdot \mathbb{1}_{\{X \leq -\mathrm{VaR}_\alpha(X)\}}]$$

ergibt. □

Zum Abschluss dieses Abschnitts diskutieren wir noch ein Beispiel, in dem wir mithilfe der gefundenen Zusammenhänge leicht den Conditional Value-at-Risk berechnen können.

(2.1.19) Beispiel
Sei $X : \Omega \to \mathbb{R}$ eine Treppenfunktion, d.h. es existieren ein $n \in \mathbb{N}$, $M_1, \ldots, M_n \in \mathcal{M}$ mit

$$\mathcal{P}\left(\bigcup_{i=1}^{n} M_i\right) = 1, \qquad \mathcal{P}(M_i) > 0 \quad \text{für alle } i = 1, \ldots, n$$

und

$$\mathcal{P}(M_i \cap M_j) = 0 \quad \text{für alle } i, j = 1, \ldots, n \text{ mit } i \neq j$$

sowie $a_1, \ldots, a_n \in \mathbb{R}$ mit

$$a_1 < a_2 < \cdots < a_n,$$

sodass

$$X = \sum_{i=1}^{n} a_i \cdot \mathbb{1}_{M_i}$$

gilt. Dabei sei mit $\mathbb{1}_{M_i}$ die charakteristische Funktion der Menge M_i bezeichnet. Darüber hinaus definieren wir $p_i := \mathcal{P}(M_i)$ für $i = 1, \ldots, n$. Als Verteilungsfunktion erhalten wir

$$F^X(z) = \begin{cases} 0 & \text{für } z < a_1, \\ \sum\limits_{j=1}^{i} p_j & \text{für } a_i \leq z < a_{i+1} \text{ mit } i = 1, \ldots, n-1, \\ 1 & \text{für } z \geq a_n, \end{cases}$$

sodass sich für den Value-at-Risk die Darstellung

$$\mathrm{VaR}_\alpha(X) = \begin{cases} -a_1 & \text{für } 0 < \alpha < p_1, \\ -a_i & \text{für } \sum\limits_{j=1}^{i-1} p_j \leq \alpha < \sum\limits_{j=1}^{i} p_j \text{ mit } i = 2, \ldots, n-1, \\ -a_n & \text{für } \sum\limits_{j=1}^{n-1} p_j \leq \alpha < 1 \end{cases}$$

ergibt. Da nun

$$-\mathrm{VaR}_\alpha(X) \in \{a_1, \ldots, a_n\}$$

für alle $\alpha \in (0, 1)$ gilt und $\{a_1, \ldots, a_n\}$ gerade der Menge der Unstetigkeitsstellen von F^X entspricht, ist Korollar (2.1.18) in diesem Beispiel für kein $\alpha \in (0, 1)$ anwendbar. Daher müssen wir auf Satz (2.1.10) zurückgreifen; dieser ist anwendbar, da offenbar

$X \in L^1(\Omega)$ gilt. Demnach ist

$$\begin{aligned}
&\mathrm{CVaR}_\alpha(X)\\
&= \frac{1}{\alpha}\int_0^\alpha VaR_\beta(X)\,\mathrm{d}\beta\\
&= \begin{cases}
\frac{1}{\alpha}\cdot\alpha\cdot(-a_1) = -a_1 & \text{für } 0 < \alpha < p_1,\\
\frac{1}{\alpha}\left(\sum\limits_{j=1}^{i-1} p_j(-a_j) + \left(\alpha - \sum\limits_{j=1}^{i-1} p_j\right)(-a_i)\right) & \text{für } \sum\limits_{j=1}^{i-1} p_j \le \alpha < \sum\limits_{j=1}^{i} p_j \text{ mit}\\
 & i = 2, \ldots, n-1,\\
\frac{1}{\alpha}\left(\sum\limits_{j=1}^{n-1} p_j(-a_j) + \left(\alpha - \sum\limits_{j=1}^{n-1} p_j\right)(-a_n)\right) & \text{für } \sum\limits_{j=1}^{n-1} p_j \le \alpha < 1.
\end{cases}
\end{aligned}$$

Dabei sei bemerkt, dass die Berechnung nur anhand der Definition des Conditional Value-at-Risks zwar weniger in diesem Beispiel, aber i.A. vermutlich ein großes bis kaum lösbares Problem darstellt. ◇

2.2 Anwendung in der Portfolio-Theorie

Bis zu diesem Zeitpunkt haben wir neben einer Vielzahl theoretischer Aussagen über Abweichungsmaße auch eine ganze Reihe an Beipielen kennengelernt. Besonders im Paragraphen 1.4 haben wir zwei große Klassen an Abweichungsmaßen erschlossen. Nun ist es aber an der Zeit, auf zumindest ein Anwendungsgebiet der Abweichungsmaße einzugehen. Dabei ist die Portfolio-Theorie als eines der Hauptanwendungsgebiete zu nennen. Als Hintergrundliteratur, in denen die Theorie auch schon weit vorangetrieben wurde, empfehlen wir [9] von Maier-Paape und Zhu. Dieses wählen wir auch als Ausgangssituation für diesen Abschnitt.
Zuallererst stellen wir kurz die Voraussetzungen in [9, S. 4] vor. Dort finden sich auch weitere Details sowie jegliche Interpretationen, auf die wir hier verzichten wollen.

Den bis jetzt allgemein gewählten Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ schränken wir wie folgt ein: Die Grundmenge Ω sei endlich, etwa $\Omega = \{\omega_1, \ldots, \omega_N\}$ für $N \in \mathbb{N}$, und als σ-Algebra $\mathcal{M}$ wählen wir die Potenzmenge $\mathfrak{P}(\Omega)$. Für unsere Zwecke ist es außerdem nützlich, Œ anzunehmen, dass ein $N_0 \in \{1, \ldots, N\}$ existiert, sodass

$$\mathcal{P}(\omega_i) > 0 \quad \text{für alle } i = 1, \ldots, N_0 \qquad \text{und} \qquad \mathcal{P}(\omega_i) = 0 \quad \text{für alle } i = N_0 + 1, \ldots, N$$

gilt. Nach Beispiel (A.1.19) a) ist $(\Omega, \mathcal{M}, \mathcal{P})$ demnach insbesondere essentiell endlich.

Weitere Definitionen, die größtenteils auf [9, Definition 1 und 2] basieren, fassen wir folgendermaßen kurz zusammen:

(2.2.1) Definition (Finanzmarkt, (zulässige) Portfolios)
Sei $M \in \mathbb{N}$.

(i) Sei $S_t = (S_t^0, \ldots, S_t^M)^T$ für $t = 0, 1$, wobei $S_0 \in \mathbb{R}_+^{M+1}$ mit $S_0^0 = 1$ sowie $S_1^i : \Omega \to \mathbb{R}$ für $i = 1, \ldots, M$ nicht-negative Zufallsvariablen und $S_1^0 = R > 0$ seien. Dann bezeichnen wir S_t für $t = 0, 1$ als Finanzmarkt mit einer Zeitperiode. *Dabei entsprechen S_0^0, S_1^0 dem risikofreien bond und die restlichen Komponenten S_t^i für $i = 1, \ldots, M$ repräsentieren den Preis der i-ten risikobehafteten Kapitalanlage zur Zeit t. Wir schreiben außerdem $\hat{S}_t = (S_t^1, \ldots, S_t^M)^T$.*

(ii) Ein $x = (x_0, \ldots, x_M)^T \in \mathbb{R}^{M+1}$ bezeichnen wir als Portfolio. *Wir schreiben genauso $\hat{x} = (x_1, \ldots, x_M)^T$.*

(iii) Eine nicht-leere, konvexe und abgeschlossene Menge $A \subseteq \mathbb{R}^{M+1}$ nennen wir eine Menge an zulässigen Portfolios. *Darüber hinaus bezeichnen wir eine nicht-leere, konvexe und abgeschlossene Menge A, die eine Teilmenge von*

$$\{x \in \mathbb{R}^{M+1} \mid S_0^T x = 1\}$$

ist, als Menge an zulässigen Portfolios mit Anfangspreis Eins. ◇

(2.2.2) Bemerkung
Für einen gegebenen Finanzmarkt S_t mit $t = 0, 1$, also mit einer Zeitperiode, gilt aufgrund der endlichen Kardinalität von Ω offenbar, dass $S_1^i \in L^2(\Omega)$ für alle $i = 1, \ldots, M$ ist. ◇

Zudem werden besonders hier „Risikomaße" für die Abschätzung von Risiken im Zusammenhang mit möglichen Investitionen benötigt. Diese sind aber grundlegend von den uns bekannten Risikomaßen aus Definition (1.3.1) verschieden. Die anschließende zugehörige Definition entspricht [9, Assumption 2].

(2.2.3) Definition (Risikomaße in der Portfolio-Theorie)
Sei A eine Menge an zulässigen Portfolios und $\mathfrak{r} : A \to [0, \infty)$ eine stetige Funktion. Dann nennen wir $\mathfrak{r}$ ein Risikomaß.

Außerdem seien folgende Axiome gegeben, die die Funktion $\mathfrak{r}$ später teilweise auch erfüllen soll:

(r1) Für jedes Portfolio $x \in A$ hängt die Funktion $\mathfrak{r}$ nur von $\hat{x}$ ab, d.h. es existiert eine Funktion

$$\hat{\mathfrak{r}} : \hat{A} \to [0, \infty),$$

wobei

$$\hat{A} := \{\hat{x} \in \mathbb{R}^M \mid (x_0, \hat{x}^T)^T \in A \text{ für ein } x_0 \in \mathbb{R}\}$$

sei, mit $\mathfrak{r}(x) = \hat{\mathfrak{r}}(\hat{x})$ für alle $x \in A$.

($r1n$) *Für $x \in A$ gilt genau dann $\mathfrak{r}(x) = 0$, falls $\hat{x} = 0$ gilt.*

($r2$) *Die Funktion $\mathfrak{r}$ ist konvex.*

($r2s$) *Die Funktion $\hat{\mathfrak{r}}$ aus ($r1$) ist strikt konvex.*

($r3$) *Die Funktion $\hat{\mathfrak{r}}$ aus ($r1$) ist positiv homogen, d.h. es gilt $\hat{\mathfrak{r}}(t\hat{x}) = t\hat{\mathfrak{r}}(\hat{x})$ für alle $t > 0$ und $\hat{x} \in \hat{A}$, für die auch $t\hat{x} \in \hat{A}$ ist.*

($r3s$) *Die Funktion $\hat{\mathfrak{r}}$ aus ($r1$) erfüllt ($r3$) und für alle $\hat{x}, \hat{y} \in \hat{A}$ mit $\hat{x} \neq \hat{y}$ und $\hat{\mathfrak{r}}(\hat{x}) = \hat{\mathfrak{r}}(\hat{y}) = 1$ sowie $\alpha \in (0,1)$ gilt*

$$\hat{\mathfrak{r}}\big(\alpha\hat{x} + (1-\alpha)\hat{y}\big) < \alpha\hat{\mathfrak{r}}(\hat{x}) + (1-\alpha)\hat{\mathfrak{r}}(\hat{y}) = 1.$$

◇

(2.2.4) Bemerkung
Mit A ist offenbar auch $\hat{A}$ konvex, sodass die Axiome ($r2s$) und ($r3s$) wohldefiniert sind. ◇

Jedenfalls wollen wir aus manchen Abweichungsmaßen Risikomaße im Sinne der Portfolio-Theorie konstruieren. Dabei werden wir sogar feststellen, dass diese Risikomaße auch einige der oben genannten Axiome erfüllen. Die grundsätzliche Idee dazu stammt z.B. aus [13, Abschnitt 3]. Zwar wurde dort ein beliebiger Wahrscheinlichkeitsraum zugrunde gelegt, aber durch entsprechende einschränkende Voraussetzungen an den Finanzmarkt ist der Nachweis analog durchführbar.

(2.2.5) Satz
Sei ein Finanzmarkt S_t mit $t = 0, 1$, also mit einer Zeitperiode, und ein unterhalbstetiges Abweichungsmaß $\mathcal{D} : L^2(\Omega) \to [0, \infty]$ gegeben, sodass sie folgende zwei Axiome erfüllen:

($A1$) *$S_1^0, \ldots, S_1^M$ sind linear unabhängig in $L^2(\Omega)$.*

($A2$) *Es ist $\mathcal{D}(S_1^i) < \infty$ und $\mathcal{D}(-S_1^i) < \infty$ für alle $i = 1, \ldots, M$.*

Dann ist

$$\mathfrak{r} : \mathbb{R}^{M+1} \to [0, \infty), \quad \mathfrak{r}(x) := \mathcal{D}(\hat{x}^T \hat{S}_1)$$

ein Risikomaß im Sinne von Definition (2.2.3), das die Axiome ($r1$) mit

$$\hat{\mathfrak{r}} : \mathbb{R}^{M} \to [0, \infty), \quad \hat{\mathfrak{r}}(\hat{x}) := \mathcal{D}(\hat{x}^T \hat{S}_1)$$

sowie ($r1n$), ($r2$) und ($r3$) erfüllt. ◇

(2.2.6) Bemerkung
Wenn $\mathfrak{r}$ auf $\mathbb{R}^{M+1}$ stetig ist und die Axiome ($r1$), ($r1n$), ($r2$) und ($r3$) auf $\mathbb{R}^{M+1}$ erfüllt, dann ist $\mathfrak{r}$ natürlich auch auf einer Menge A an zulässigen Portfolios stetig und erfüllt genauso darauf die genannten Axiome. Dasselbe gilt offenbar auch für die übrigen Axiome aus Definition (2.2.3). ◇

Beweis von Satz (2.2.5)
Zunächst zeigen wir, dass $\mathfrak{r}$ eine endliche und unterhalbstetige Funktion ist: Sei dazu $x \in \mathbb{R}^{M+1}$, so ist

$$0 \leq \mathfrak{r}(x) = \mathcal{D}(\hat{x}^T \hat{S}_1) \underset{(\mathcal{D}2),(\mathcal{D}3)}{\leq} \sum_{i=1}^{M} |x_i| \cdot \mathcal{D}\big(\operatorname{sign}(x_i) S_1^i\big) \underset{(A2)}{<} \infty,$$

also ist $\mathfrak{r}$ endlich.
Darüber hinaus sei $(x^{(n)})_{n \in \mathbb{N}} \subseteq \mathbb{R}^{M+1}$ mit $x^{(n)} \xrightarrow[n \to \infty]{} x$ in $\mathbb{R}^{M+1}$, dann folgt

$$\|\big(\hat{x}^{(n)}\big)^T \hat{S}_1 - \hat{x}^T \hat{S}_1\|_2 \leq \sum_{i=1}^{M} |\big(x^{(n)}\big)_i - x_i| \cdot \|S_1^i\|_2 \leq \|x^{(n)} - x\|_{\infty, \mathbb{R}^{M+1}} \sum_{i=1}^{M} \|S_1^i\|_2 \xrightarrow[n \to \infty]{} 0,$$

d.h. es gilt $\big(\hat{x}^{(n)}\big)^T \hat{S}_1 \xrightarrow[n \to \infty]{} \hat{x}^T \hat{S}_1$ in $L^2(\Omega)$. Damit liefert die Unterhalbstetigkeit von $\mathcal{D}$, dass

$$\mathfrak{r}(x) = \mathcal{D}(\hat{x}^T \hat{S}_1) \leq \liminf_{n \to \infty} \mathcal{D}\big(\big(\hat{x}^{(n)}\big)^T \hat{S}_1\big) = \liminf_{n \to \infty} \mathfrak{r}\big(x^{(n)}\big)$$

gilt, d.h. $\mathfrak{r}$ ist unterhalbstetig.

Des Weiteren erfüllt die Funktion $\mathfrak{r}$ offenbar $(r1)$ mit

$$\hat{\mathfrak{r}} : \mathbb{R}^M \to [0, \infty), \quad \hat{\mathfrak{r}}(\hat{x}) := \mathcal{D}(\hat{x}^T \hat{S}_1)$$

sowie $(r2)$ und $(r3)$, da $\mathcal{D}$ den Axiomen $(\mathcal{D}2)$ und $(\mathcal{D}3)$ genügt und daher insbesondere selbst konvex ist. Obendrein liefert Satz (1.1.7) dann direkt, dass $\mathfrak{r}$ sogar stetig ist.

Somit bleibt nur noch nachzuweisen, dass $\mathfrak{r}$ die Eigenschaft $(r1n)$ besitzt: Zunächst ist offenbar $\mathfrak{r}(0) = 0$ und, da $\mathcal{D}$ das Axiom $(\mathcal{D}4)$ erfüllt, genügt es zu zeigen, dass $\hat{x}^T \hat{S}_1$ als Element von $L^2(\Omega)$ für jedes $x \in \mathbb{R}^{M+1}$ mit $\hat{x} \neq 0$ nicht-konstant ist.
Dazu nehmen wir an, dass ein $x \in \mathbb{R}^{M+1}$ mit $\hat{x} \neq 0$ und ein $C \in \mathbb{R}$ existieren, sodass

$$\hat{x}^T \hat{S}_1 = C \qquad \text{in } L^2(\Omega)$$

gilt. Dies ist aber äquivalent zu

$$0 = \hat{x}^T \hat{S}_1 - C = \hat{x}^T \hat{S}_1 - \frac{C}{R} \cdot R = \hat{x}^T \hat{S}_1 - \frac{C}{R} \cdot S_1^0,$$

d.h. $S_1^0, \ldots, S_1^M$ sind linear abhängig in $L^2(\Omega)$, das aber $(A1)$ widerspricht. Demnach muss $\mathfrak{r}$ auch das Axiom $(r1n)$ erfüllen. □

Aufgrund der Endlichkeit der Grundmenge Ω ist dieser Satz auf viele Abweichungsmaße anwendbar.

(2.2.7) Beispiele
Sei ein Finanzmarkt S_t mit $t = 0, 1$, also mit einer Zeitperiode, der der Eigenschaft $(A1)$ genügt, gegeben. Da die Grundmenge Ω als endlich angenommen wurde, ist der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ bekanntlich essentiell endlich.
Daher sind sowohl die Abweichungsmaße $\mathcal{D}^p$ aus Beispiel (1.4.8) a) als auch $\mathcal{D}^p_\pm$ aus Beispiel (1.4.8) b) endlich und stetig, wie wir schon dort festgestellt haben, sodass sie insbesondere unterhalbstetig sind und $(A2)$ erfüllen. Ferner ist auch das Abweichungsmaß aus Satz (2.1.9), das auf dem Conditional Value-at-Risk basiert, endlich und stetig, sodass ebenfalls dieses die Voraussetzungen des obigen Satzes erfüllt.
Damit haben wir aber auch schon praktisch alle Abweichungsmaße, die konkret als Beispiel diskutiert wurden, behandelt. $\diamond$

Somit steht uns eine Vielzahl an Risikomaßen im Sinne von Definition (2.2.3), die auch schon einige der dort genannten Axiome erfüllen, zur Verfügung. Diese wollen wir jetzt in einem konkreten Resultat der Portfolio-Theorie anwenden, siehe [9, Theorem 11], das wir hier aber nur zitieren wollen. Darin taucht der Begriff „master fund" auf, dessen Erklärung z.B. in [13, Abschnitt 4] zu finden ist.

(2.2.8) Satz (Existenz eines master funds)
Sei ein Finanzmarkt S_t mit $t = 0, 1$, also mit einer Zeitperiode, und mit

$$E[\hat{S}_1] - R\hat{S}_0 \neq \hat{0} \tag{2.28}$$

sowie $A = \{x \in \mathbb{R}^{M+1} \mid S_0^T x = 1\}$ gegeben. Darüber hinaus sei mit $\mathfrak{r} : A \to [0, \infty)$ ein Risikomaß gegeben, das die Eigenschaften $(r1)$, $(r1n)$, $(r2)$ und $(r3s)$ besitzt und für das die Subniveaumengen

$$\{x \in A \mid r \geq \mathfrak{r}(x)\}$$

für alle $r \in \mathbb{R}$ kompakt seien. Es gilt $\hat{A} = \mathbb{R}^M$, sodass die Funktion

$$f : \mathbb{R}^M \to [0, \infty), \quad f(\hat{x}) := \frac{1}{2}\hat{\mathfrak{r}}(\hat{x})^2$$

wohldefinert ist. Außerdem sei $f^* : \mathbb{R}^M \to \mathbb{R}$ die zu f gehörige Fenchel-Konjugierte, die in diesem Fall endlich (und somit stetig) ist. Obendrein sei f^* in $E[\hat{S}_1] - R\hat{S}_0$ differenzierbar.

Dann existiert ein master fund genau dann, falls

$$\hat{S}_0^T \nabla f^*\big(E[\hat{S}_1] - R\hat{S}_0\big) > 0 \tag{2.29}$$

ist. $\diamond$

(2.2.9) Bemerkung

a) In obigem Satz müssen wir noch eine Notation erklären: Zu einem Zufallsvektor $X : \Omega \to \mathbb{R}^n$ definieren wir

$$E[X] := (E[X_1], \ldots, E[X_n])^T.$$

b) In [9, Theorem 11] wurde nicht begründet, warum notwendigerweise $\hat{A} = \mathbb{R}^M$ gelten muss.

Dazu: Sei $\hat{x} \in \mathbb{R}^M$, so definieren wir $x_0 := 1 - \hat{S}_0^T \hat{x}$. Wegen $S_0^0 = 1$ erhalten wir für $x := (x_0, \hat{x}^T)^T \in \mathbb{R}^{M+1}$ die Identität

$$S_0^T x = S_0^0 x_0 + \hat{S}_0^T \hat{x} = 1,$$

also $x \in A$ bzw. $\hat{x} \in \hat{A}$. Daraus folgt $\hat{A} = \mathbb{R}^M$.

Zudem wurde dort ohne Beweis verwendet, dass f^* endlich, also dass $f^*(\mathbb{R}^M) \subseteq \mathbb{R}$ gilt, (und somit stetig) ist.

Dazu: Wegen der Endlichkeit von f folgt zunächst aus Lemma (A.3.3), dass $f^* > -\infty$ sowie f^* konvex und unterhalbstetig ist.
Darüber hinaus ergibt sich aus der Stetigkeit von $\mathfrak{r}$ zusammen mit der speziellen Gestalt von A, dass $\hat{\mathfrak{r}}$ stetig auf $\hat{A} = \mathbb{R}^M$ ist. Außerdem können wir aus $(r1)$ und $(r1n)$ schließen, dass für $\hat{x} \in \mathbb{R}^M$ genau dann $\hat{\mathfrak{r}}(\hat{x}) = 0$ gilt, falls $\hat{x} = \hat{0}$ gilt. Diese beiden Eigenschaften implizieren für

$$B_1(\hat{0}) \equiv B_1^{\|\cdot\|_{2,\mathbb{R}^M}}(\hat{0}) \subseteq \mathbb{R}^M$$

zusammen mit der Kompaktheit von $\partial B_1(\hat{0})$ und wegen $\hat{\mathfrak{r}} \geq 0$, dass die Ungleichung

$$C := \min_{\hat{x} \in \partial B_1(\hat{0})} \hat{\mathfrak{r}}(\hat{x}) > 0$$

gilt. Dann erhalten wir für $\hat{x} \neq \hat{0}$ wegen $(r3)$ die Abschätzung

$$\hat{\mathfrak{r}}(\hat{x}) = \|\hat{x}\|_{2,\mathbb{R}^M} \cdot \hat{\mathfrak{r}}\Bigg(\underbrace{\frac{\hat{x}}{\|\hat{x}\|_{2,\mathbb{R}^M}}}_{\in\, \partial B_1(\hat{0})}\Bigg) \geq C\|\hat{x}\|_{2,\mathbb{R}^M},$$

das wiederum

$$f(\hat{x}) = \frac{1}{2}\hat{\mathfrak{r}}(\hat{x})^2 \geq \frac{1}{2}C^2\|\hat{x}\|^2_{2,\mathbb{R}^M}$$

liefert, wobei die letzte Ungleichung letztendlich auch für $\hat{x} = \hat{0}$ gilt. In Kombination mit der Cauchy-Schwarz-Ungleichung impliziert dies

$$\hat{x}^T\hat{y} - f(\hat{y}) \leq \|\hat{x}\|_{2,\mathbb{R}^M} \cdot \|\hat{y}\|_{2,\mathbb{R}^M} - \frac{1}{2}C^2\|\hat{y}\|^2_{2,\mathbb{R}^M}$$

für alle $\hat{x}, \hat{y} \in \mathbb{R}^M$, sodass für jedes feste $\hat{x} \in \mathbb{R}^M$ offensichtlich

$$f^*(\hat{x}) = \sup_{\hat{y} \in \mathbb{R}^M} \{\hat{x}^T\hat{y} - f(\hat{y})\} \leq \sup_{\hat{y} \in \mathbb{R}^M} \left\{\|\hat{x}\|_{2,\mathbb{R}^M} \cdot \|\hat{y}\|_{2,\mathbb{R}^M} - \frac{1}{2}C^2\|\hat{y}\|^2_{2,\mathbb{R}^M}\right\} < \infty$$

gilt, d.h. wegen $f^* > -\infty$ erhalten wir $\mathrm{dom}(f^*) = \mathbb{R}^M$. Aufgrund der Unterhalbstetigkeit und Konvexität von f^* liefert der Satz (1.1.7) schließlich, dass f^* sogar endlich ist.

c) Diese Charakterisierung wurde erstmals vor noch nicht allzu langer Zeit von MAIER-PAAPE und ZHU entwickelt und ist demnach eine Weiterentwicklung der Existenzbedingung von ROCKAFELLAR ET AL. in [13]. ◇

Unsere Hoffnung besteht nun darin, aus Satz (2.2.5) Risikomaße zu erhalten, die zum Einen überhaupt die Voraussetzungen des obigen Satzes erfüllen und für die wir zum Anderen den Term aus (2.29) konkret berechnen können. Dazu bieten sich die Abweichungsmaße

$$\mathcal{D}^p : L^2(\Omega) \to [0,\infty], \quad \mathcal{D}^p(X) := \|X - EX\|_p$$

für $p \in [1,\infty]$ aus Beispiel (1.4.8) a) am ehesten an.

(2.2.10) Lemma
Sei ein Finanzmarkt S_t mit $t = 0,1$, also mit einer Zeitperiode, der der Eigenschaft $(A1)$ genügt, gegeben. Dann ist

$$\mathfrak{r}_p : \mathbb{R}^{M+1} \to [0,\infty), \quad \mathfrak{r}_p(x) := \mathcal{D}^p(\hat{x}^T \hat{S}_1)$$

für alle $p \in [1,\infty]$ ein Risikomaß im Sinne von Definition (2.2.3), das die Axiome $(r1)$, $(r1n)$, $(r2)$ und $(r3)$ erfüllt. Zudem gilt:

(i) Die Funktion $\hat{\mathfrak{r}}_p$ aus $(r1)$ ist eine Norm auf dem $\mathbb{R}^M$ und sie besitzt für alle $\hat{x} \in \mathbb{R}^M$ die Gestalt

$$\hat{\mathfrak{r}}_p(\hat{x}) = \|B_p \hat{x}\|_{p,\mathbb{R}^{N_0}} \tag{2.30}$$

mit $B_p = G_p \cdot H \in \mathbb{R}^{N_0 \times M}$, wobei $G_p := diag\Big(\mathcal{P}(\omega_1)^{\frac{1}{p}}, \dots, \mathcal{P}(\omega_{N_0})^{\frac{1}{p}}\Big) \in \mathbb{R}^{N_0 \times N_0}$ (Erinnerung: $\frac{1}{\infty} := 0$) und

$$H = \begin{pmatrix} S_1^1(\omega_1) - E[S_1^1] & S_1^2(\omega_1) - E[S_1^2] & \cdots & S_1^M(\omega_1) - E[S_1^M] \\ S_1^1(\omega_2) - E[S_1^1] & S_1^2(\omega_2) - E[S_1^2] & \cdots & S_1^M(\omega_2) - E[S_1^M] \\ \vdots & \vdots & \vdots & \vdots \\ S_1^1(\omega_{N_0}) - E[S_1^1] & S_1^2(\omega_{N_0}) - E[S_1^2] & \cdots & S_1^M(\omega_{N_0}) - E[S_1^M] \end{pmatrix} \in \mathbb{R}^{N_0 \times M}$$

seien. Dabei gilt notwendigerweise $N_0 > M$ und die Matrix B_p besitzt Rang M.

(ii) Die Subniveaumengen

$$\{x \in A \mid r \geq \mathfrak{r}_p(x)\}$$

mit $A = \{x \in \mathbb{R}^{M+1} \mid S_0^T x = 1\}$ sind für alle $r \in \mathbb{R}$ kompakt.

(iii) Für $p \in (1,\infty)$ erfüllt $\mathfrak{r}_p$ sogar das Axiom $(r3s)$. ◇

Beweis
Die ersten Behauptungen folgen bereits aus der Diskussion in den Beispielen (2.2.7), da für alle $p \in [1,\infty]$ der Satz (2.2.5) auf $\mathcal{D}^p$ anwendbar ist; denn das Axiom $(A2)$ ist automatisch erfüllt. Daher bleiben nur noch die Zusätze zu zeigen:

(i) Da $\mathcal{D}^p$ nach Definition offenbar symmetrisch ist, gilt dies auch für $\mathfrak{r}_p$ und $\hat{\mathfrak{r}}_p$. Daraus folgt, dass

$$\hat{\mathfrak{r}}_p(t\hat{x}) \underset{(r3)}{=} |t| \cdot \hat{\mathfrak{r}}_p\big(\operatorname{sign}(t)\hat{x}\big) = |t| \cdot \hat{\mathfrak{r}}_p(\hat{x})$$

für alle $t \in \mathbb{R}$ und $\hat{x} \in \mathbb{R}^M$ gilt. Da $\mathfrak{r}_p$ wegen $(r2)$ zudem konvex ist, ist es aufgrund von $(r1)$ ebenfalls $\hat{\mathfrak{r}}_p$, sodass wir

$$\hat{\mathfrak{r}}_p(\hat{x}_1 + \hat{x}_2) = \hat{\mathfrak{r}}_p\left(\frac{1}{2} \cdot (2\hat{x}_1 + 2\hat{x}_2)\right) \le \frac{1}{2} \cdot \big(\hat{\mathfrak{r}}_p(2\hat{x}_1) + \hat{\mathfrak{r}}_p(2\hat{x}_2)\big) \underset{(r3)}{=} \hat{\mathfrak{r}}_p(\hat{x}_1) + \hat{\mathfrak{r}}_p(\hat{x}_2)$$

für alle $\hat{x}_1, \hat{x}_2 \in \mathbb{R}^M$ erhalten. Zusammen mit $(r1)$ und $(r1n)$ ergibt sich dann, dass es sich bei $\hat{\mathfrak{r}}_p$ um eine Norm auf dem $\mathbb{R}^M$ handelt.

Wenn wir uns an die Definition von N_0 erinnern, erhalten wir obendrein für $p \in [1,\infty)$ die Identität

$$\begin{aligned}
\hat{\mathfrak{r}}_p(\hat{x}) = \left\|\hat{x}^T\big(\hat{S}_1 - E[\hat{S}_1]\big)\right\|_p &= \left(\sum_{i=1}^{N} \mathcal{P}(\omega_i) \cdot \left|\sum_{j=1}^{M} x_j \cdot \big(S_1^j(\omega_i) - E[S_1^j]\big)\right|^p\right)^{\frac{1}{p}} \\
&= \left(\sum_{i=1}^{N_0} \mathcal{P}(\omega_i) \cdot \big|(H\hat{x})_i\big|^p\right)^{\frac{1}{p}} \\
&= \left(\sum_{i=1}^{N_0} \left|\mathcal{P}(\omega_i)^{\frac{1}{p}} (H\hat{x})_i\right|^p\right)^{\frac{1}{p}} \\
&= \left(\sum_{i=1}^{N_0} \big|(\underbrace{G_p H}_{=\,B_p}\,\hat{x})_i\big|^p\right)^{\frac{1}{p}} = \|B_p\hat{x}\|_{p,\mathbb{R}^{N_0}}
\end{aligned}$$

für alle $\hat{x} \in \mathbb{R}^M$. Für $p = \infty$ können wir hingegen

$$\begin{aligned}
\hat{\mathfrak{r}}_\infty(\hat{x}) = \operatorname{ess\,sup} \left|\hat{x}^T\big(\hat{S}_1 - E[\hat{S}_1]\big)\right| &= \max_{i=1,\dots,N_0} \left|\sum_{j=1}^{M} x_j \cdot \big(S_1^j(\omega_i) - E[S_1^j]\big)\right| \\
&= \max_{i=1,\dots,N_0} \big|(H\hat{x})_i\big| = \|B_\infty\hat{x}\|_{\infty,\mathbb{R}^{N_0}}
\end{aligned}$$

für alle $\hat{x} \in \mathbb{R}^M$ schreiben, da $G_\infty = I_{N_0}$ und somit $B_\infty = H$ gilt, wobei wir mit I_{N_0} die Einheitsmatrix in $\mathbb{R}^{N_0 \times N_0}$ bezeichnen.

Sei nun wieder $p \in [1, \infty]$ beliebig. Des Weiteren sei $\hat{x} \in \mathbb{R}^M$ mit $B_p\hat{x} = 0$, das äquivalent zu

$$0 = \|B_p\hat{x}\|_{p,\mathbb{R}^{N_0}} = \hat{\mathfrak{r}}_p(\hat{x})$$

ist. Dies ist aber nach obigen Überlegungen genau dann der Fall, falls $\hat{x} = \hat{0}$ gilt. Dies impliziert $N_0 \geq M$ und $\mathrm{Rang}(B_p) = M$. Dabei ist aber $M = N_0$ ausgeschlossen; denn es ist

$$\begin{aligned}\left(\mathcal{P}(\omega_1), \ldots, \mathcal{P}(\omega_{N_0})\right) \cdot H &= \left(\sum_{i=1}^{N_0} \mathcal{P}(\omega_i) S_1^1(\omega_i), \ldots, \sum_{i=1}^{N_0} \mathcal{P}(\omega_i) S_1^M(\omega_i)\right) - E[\hat{S}_1]^T \\ &= E[\hat{S}_1]^T - E[\hat{S}_1]^T = 0,\end{aligned}$$

das $\mathrm{Rang}(H) < N_0$ liefert. Da ferner $\mathrm{Rang}(B_p) = \mathrm{Rang}(H)$ gilt, weil G_p nach Wahl von N_0 invertierbar ist, folgt

$$M = \mathrm{Rang}(B_p) = \mathrm{Rang}(H) < N_0.$$

(ii) Sei $r \in \mathbb{R}$ vorgegeben. Falls $r < 0$ ist, ist wegen $\mathfrak{r} \geq 0$ offenbar

$$\{x \in A \mid r \geq \mathfrak{r}_p(x)\} = \emptyset$$

und daher insbesondere kompakt. Andernfalls erhalten wir zunächst, dass

$$M_r := \{x \in A \mid r \geq \mathfrak{r}_p(x)\} = A \cap \{x \in \mathbb{R}^{M+1} \mid r \geq \mathfrak{r}_p(x)\}$$

abgeschlossen ist, da sowohl A als auch $\{x \in \mathbb{R}^{M+1} \mid r \geq \mathfrak{r}_p(x)\}$ abgeschlossen sind; bei A ist es offensichtlich und bei der letzteren Menge folgt dies aus der Stetigkeit von $\mathfrak{r}_p$.
Somit müssen wir nur noch die Beschränktheit nachweisen: Zuerst sei gesagt, dass $\mu, \rho > 0$ existieren, sodass

$$\mu\|\hat{x}\|_{\infty,\mathbb{R}^M} \leq \hat{\mathfrak{r}}_p(\hat{x}) \leq \rho\|\hat{x}\|_{\infty,\mathbb{R}^M}$$

für alle $\hat{x} \in \mathbb{R}^M$ gilt, da $\hat{\mathfrak{r}}_p$ nach (i) eine Norm auf dem $\mathbb{R}^M$ darstellt und alle Normen auf diesem Raum äquivalent sind. Dann ergibt sich für jedes $x \in M_r$ wegen *(r1)*, dass

$$r \geq \mathfrak{r}_p(x) = \hat{\mathfrak{r}}_p(\hat{x}) \geq \mu\|\hat{x}\|_{\infty,\mathbb{R}^M},$$

also $\|\hat{x}\|_{\infty,\mathbb{R}^M} \leq \frac{r}{\mu}$ ist. Da insbesondere $x \in A$, also $S_0^T x = 1$, gilt, können wir

$$|x_0| = \left|\frac{1 - \hat{x}^T\hat{S}_0}{S_0^0}\right| \leq 1 + |\hat{x}^T\hat{S}_0| \leq 1 + \|\hat{S}_0\|_{1,\mathbb{R}^M} \cdot \|\hat{x}\|_{\infty,\mathbb{R}^M} \leq 1 + \|\hat{S}_0\|_{1,\mathbb{R}^M} \cdot \frac{r}{\mu}$$

wegen $S_0^0 = 1$ schließen, d.h. es ist

$$\|x\|_{\infty,\mathbb{R}^{M+1}} \leq \max\left\{\frac{r}{\mu}, 1 + \|\hat{S}_0\|_{1,\mathbb{R}^M} \cdot \frac{r}{\mu}\right\}.$$

Daher ist M_r beschränkt und somit kompakt.

(iii) Sei $p \in (1,\infty)$. Wir wissen schon, dass $\mathfrak{r}_p$ das Axiom $(r3)$ erfüllt. Jetzt seien $\hat{x},\hat{y} \in \mathbb{R}^M$ mit $\hat{x} \neq \hat{y}$ und $\hat{\mathfrak{r}}_p(\hat{x}) = \hat{\mathfrak{r}}_p(\hat{y}) = 1$ sowie $\alpha \in (0,1)$.
Ohne Beweis verwenden wir an dieser Stelle, dass die Einheitskugel bzgl. $\|\cdot\|_{p,\mathbb{R}^n}$ für alle $n \in \mathbb{N}$ strikt konvex ist, d.h. für alle $w,z \in \mathbb{R}^n$ mit $w \neq z$ und $\|w\|_{p,\mathbb{R}^n} = \|z\|_{p,\mathbb{R}^n} = 1$ ist

$$\|\lambda w + (1-\lambda)z\|_{p,\mathbb{R}^n} < 1 \qquad \text{für alle } \lambda \in (0,1). \tag{2.31}$$

Es sei erwähnt, dass dies für $p \in \{1,\infty\}$ aber nicht richtig ist, wie man sich leicht bildlich vorstellen kann.
Jedenfalls folgt aus den Erkenntnissen aus (i) bzgl. B_p bereits, dass mit $\hat{x} \neq \hat{y}$ auch $B_p\hat{x} \neq B_p\hat{y}$ ist. Zudem übersetzen sich die Annahmen unter Benutzung der Identität (2.30) zu

$$\|B_p\hat{x}\|_{p,\mathbb{R}^{N_0}} = \hat{\mathfrak{r}}_p(\hat{x}) = 1 \qquad \text{und} \qquad \|B_p\hat{y}\|_{p,\mathbb{R}^{N_0}} = \hat{\mathfrak{r}}_p(\hat{y}) = 1.$$

Dann liefert die Tatsache (2.31) für $n = N_0$, dass

$$\hat{\mathfrak{r}}_p(\alpha\hat{x} + (1-\alpha)\hat{y}) = \|\alpha B_p\hat{x} + (1-\alpha)B_p\hat{y}\|_{p,\mathbb{R}^{N_0}} < 1$$

ist, also genügt $\mathfrak{r}_p$ sogar dem Axiom $(r3s)$. □

Damit haben wir gesehen, dass das Risikomaß $\mathfrak{r}_p$ für $p \in (1,\infty)$ unter geeigneten Annahmen an den Finanzmarkt, also $(A1)$ und der Ungleichung (2.28), praktisch alle Voraussetzungen von Satz (2.2.8) erfüllt. Wir müssen jetzt lediglich noch die Fenchel-Konjugierte f_p^* von

$$f_p : \mathbb{R}^M \to [0,\infty), \quad f_p(\hat{x}) := \frac{1}{2}\left[\hat{\mathfrak{r}}_p(\hat{x})\right]^2 = \frac{1}{2}\|B_p\hat{x}\|^2_{p,\mathbb{R}^{N_0}} \tag{2.32}$$

bestimmen und auf Differenzierbarkeit untersuchen. Dies stellt aber ein nicht-triviales Problem dar:

Wir wissen schon aus Bemerkung (2.2.9) b), dass die Fenchel-Konjugierte f_p^* endlich und stetig ist. Um sie aber explizit zu bestimmen, verfolgen wir erstmal den naiven Ansatz der Extremwertberechnung. Dazu sei

$$g_p : \mathbb{R}^{N_0} \to [0,\infty), \quad g_p(z) := \frac{1}{2}\|z\|^2_{p,\mathbb{R}^{N_0}}. \tag{2.33}$$

Diese Funktion ist offenbar in $z \neq 0$ stetig differenzierbar mit

$$\nabla g_p(z) = \frac{1}{\|z\|^{p-2}_{p,\mathbb{R}^{N_0}}}\left(z_1|z_1|^{p-2},\ldots,z_{N_0}|z_{N_0}|^{p-2}\right)^T, \tag{2.34}$$

wobei wir dann sehen, dass der Gradient in $z = 0$ durch $\nabla g_p(0) := 0$ stetig fortsetzbar ist, d.h. es gilt $g_p \in C^1(\mathbb{R}^{N_0})$. Jedenfalls folgt dann mit der Kettenregel, dass auch $f_p \in C^1(\mathbb{R}^M)$ mit

$$\begin{aligned}\nabla f_p(\hat{x}) &= \frac{1}{\|B_p\hat{x}\|_{p,\mathbb{R}^{N_0}}^{p-2}} B_p^T\big((B_p\hat{x})_1|(B_p\hat{x})_1|^{p-2}, \ldots, (B_p\hat{x})_{N_0}|(B_p\hat{x})_{N_0}|^{p-2}\big)^T \\ &\in \mathbb{R}^M\end{aligned} \tag{2.35}$$

für $\hat{x} \neq \hat{0}$ und $\nabla f_p(\hat{0}) = 0$ ist, wobei wir dabei wieder benutzt haben, dass $B_p\hat{x} = 0$ nur für $\hat{x} = \hat{0}$ erfüllt ist. An dieser Stelle sei bemerkt, dass ∇g_p in $z = 0$ nicht differenzierbar ist, sodass dies auch für ∇f_p in $\hat{x} = \hat{0}$ nicht zu erwarten ist.

Falls wir nun zu festem $\hat{x} \in \mathbb{R}^M$ die Funktion

$$h_p : \mathbb{R}^M \to [0, \infty), \quad h_p(\hat{y}) := \hat{x}^T\hat{y} - f_p(\hat{y})$$

definieren, ist ebenfalls $h_p \in C^1(\mathbb{R}^M)$ mit

$$\nabla h_p(\hat{y}) = \hat{x} - \nabla f_p(\hat{y})$$

für alle $\hat{y} \in \mathbb{R}^M$. Unser Ansatz besteht nun darin, das Supremum von h_p mittels Extremwertberechnung zu bestimmen, wobei wir dafür noch begründen müssten, warum das Supremum tatsächlich angenommen wird. Jedenfalls unter Annahme, dass es angenommen wird, müssten wir zuallererst die Gleichung $\nabla h_p(\hat{y}) = 0$ lösen: Für $p = 2$ vereinfacht sich diese zu

$$\hat{x} = B_2^T\big((B_2\hat{y})_1, \ldots, (B_2\hat{y})_{N_0}\big)^T = (B_2^T B_2)\hat{y},$$

die wir sogar lösen können. Auf diesen Spezialfall kommen wir am Ende noch einmal zurück.
Für $p \neq 2$ sehen wir hingegen keinen Weg, wie wir die Gleichung $\nabla h_p(\hat{y}) = 0$ für beliebiges $\hat{x}$ exakt lösen können - sofern sie überhaupt lösbar ist - , sodass dieser naive Ansatz i.A. nicht wirklich zielführend ist. Daher verfolgen wir jetzt einen anderen Ansatz.

Als Ausgangspunkt dient der nachfolgende in der Theorie der Fenchel-Konjugierten bekannte Satz, wobei wir bewusst eine sehr allgemeine Ausgangssituation wählen; als Inspiration für den Beweis diente [9, Remark 10].

(2.2.11) Satz
Sei $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ ein Hilbertraum und $\|\cdot\|$ sei die zu $\langle\cdot,\cdot\rangle$ gehörige Norm. Zudem sei mit $\|\cdot\|_0$ eine weitere Norm auf $\mathcal{X}$ gegeben, welche äquivalent zu $\|\cdot\|$ sei. Dazu sei

$$\mathcal{F} : \mathcal{X} \to [0, \infty), \quad \mathcal{F}(x) := \frac{1}{2}\|x\|_0^2.$$

Die Fenchel-Konjugierte $\mathcal{F}^* : \mathcal{X} \to [0, \infty)$ von $\mathcal{F}$ ist dann gegeben durch

$$\mathcal{F}^*(x) = \frac{1}{2} \|x\|_{0^*}^2$$

für alle $x \in \mathcal{X}$, wobei

$$\| \cdot \|_{0^*} : \mathcal{X} \to [0, \infty), \quad \|x\|_{0^*} := \sup_{\|y\|_0 \leq 1} \langle x, y \rangle$$

ebenfalls eine Norm auf $\mathcal{X}$ darstellt. ◇

(2.2.12) Bemerkung
Die Norm $\| \cdot \|_{0^*}$ wird auch gelegentlich als die zu $\| \cdot \|_0$ duale Norm bezeichnet. ◇

Beweis von Satz (2.2.11)
Zuallererst beobachten wir, dass es sich bei $(\mathcal{X}, \| \cdot \|_0)$ um einen reflexiven Banachraum handelt.

Dazu: Die Äquivalenz der Normen impliziert unmittelbar, dass $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ und $(\mathcal{X}, \| \cdot \|_0)$ algebraisch und topologisch übereinstimmen. Daher ist $(\mathcal{X}, \| \cdot \|_0)$ einerseits vollständig, da dies schon auf $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ zutrifft, und andererseits stimmen damit die jeweiligen Dualräume algebraisch überein. Diese entsprechen sich aber ebenfalls topologisch, da man leicht zeigt, dass auch die Normen der jeweiligen Dualräume äquivalent sind. Mit demselben Argument stimmen dann auch die jeweiligen Bidualräume algebraisch und topologisch überein, sodass sich die Reflexivität von $(\mathcal{X}, \langle \cdot, \cdot \rangle)$ als Hilbertraum auf $(\mathcal{X}, \| \cdot \|_0)$ überträgt.

Außerdem müssen wir begründen, dass $\| \cdot \|_{0^*}$ eine Norm auf $\mathcal{X}$ darstellt.

Dazu: Offenbar ist die Menge $\mathcal{Y} := \{y \in \mathcal{X} \mid \|y\|_0 \leq 1\}$ nicht-leer, konvex und kreisförmig. Wegen der Äquivalenz der Normen ist sie ebenfalls abgeschlossen und beschränkt. Dann liefert Satz (1.2.1), dass $\| \cdot \|_{0^*}$ ein endliches, stetiges, konvexes und positiv homogenes Funktional darstellt. Wie mehrmals in dieser Arbeit gezeigt, folgt aus der Konvexität und positiven Homogenität die Subadditivität. Genauso impliziert die Kreisförmigkeit von $\mathcal{Y}$, dass $\| \cdot \|_{0^*}$ symmetrisch ist, was wiederum zusammen mit der positiven Homogenität

$$\|\lambda x\|_{0^*} = |\lambda| \|x\|_{0^*}$$

für alle $x \in \mathcal{X}$ und $\lambda \in \mathbb{R}$ liefert.
Schließlich sei $x \in \mathcal{X}$, so existiert aufgrund der Stetigkeit von $\| \cdot \|_0$ auf $\mathcal{X}$ ein $\lambda = \lambda(x) > 0$ mit $\|\lambda x\|_0 \leq 1$, sodass wir

$$\|x\|_{0^*} = \sup_{\|y\|_0 \leq 1} \langle x, y \rangle \geq \langle x, \lambda x \rangle = \lambda \|x\|^2$$

erhalten, d.h. es ist $\|x\|_{0^*} > 0$ für $x \neq 0$. Da offenbar $\|0\|_{0^*} = 0$ gilt, handelt es sich bei $\| \cdot \|_{0^*}$ also tatsächlich um eine Norm auf $\mathcal{X}$.

Dies führt uns zu folgender

Zwischenbehauptung: Zu jedem $x \in \mathcal{X}$ existiert ein $z = z(x) \in \mathcal{X}$ mit $\|z\|_0 = 1$ und

$$\|x\|_{0^*} = \langle x, z \rangle.$$

Dazu: Sei $x \in \mathcal{X}$. Œ sei dabei $x \neq 0$; denn für $x = 0$ können wir wegen $\|x\|_{0^*} = 0$ ein beliebiges $z \in \mathcal{X}$ mit $\|z\|_0 = 1$ wählen.
Jedenfalls existiert dann eine Folge $(y_n)_{n\in\mathbb{N}} \subseteq \overline{B_1^{\|\cdot\|_0}(0)}$ mit

$$\langle x, y_n \rangle \xrightarrow[n\to\infty]{} \sup_{\|y\|_0 \leq 1} \langle x, y \rangle = \|x\|_{0^*}. \tag{2.36}$$

Da $(\mathcal{X}, \|\cdot\|_0)$ ein reflexiver Banachraum ist, ist $\overline{B_1^{\|\cdot\|_0}(0)}$ nach [10, Satz 6.2.3] schwach folgenkompakt, d.h. es existiert eine Teilfolge $(y_{n_k})_{k\in\mathbb{N}} \subseteq (y_n)_{n\in\mathbb{N}}$ und ein $z \in \overline{B_1^{\|\cdot\|_0}(0)}$ mit

$$y_{n_k} \rightharpoonup z \quad \text{für } k \to \infty \text{ in } (\mathcal{X}, \|\cdot\|_0). \tag{2.37}$$

Da jedoch $(\mathcal{X}, \|\cdot\|_0)$ algebraisch und topologisch mit $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ übereinstimmt, impliziert die schwache Konvergenz in (2.37), dass insbesondere

$$\langle x, y_{n_k} \rangle \xrightarrow[k\to\infty]{} \langle x, z \rangle$$

gilt. Somit erhalten wir mit der Konvergenz in (2.36)

$$\|x\|_{0^*} \geq \langle x, z \rangle = \lim_{k\to\infty} \langle x, y_{n_k} \rangle = \|x\|_{0^*},$$

also $\|x\|_{0^*} = \langle x, z \rangle$.
Jetzt bleibt nur noch zu zeigen, dass $\|z\|_0 = 1$ gilt: Da $x \neq 0$ ist, wissen wir, dass

$$\langle x, z \rangle = \|x\|_{0^*} > 0$$

gilt, das insbesondere $z \neq 0$ und wegen $z \in \overline{B_1^{\|\cdot\|_0}(0)}$ damit $0 < \|z\|_0 \leq 1$ liefert. Weil ebenfalls

$$\left\| \frac{z}{\|z\|_0} \right\|_0 \leq 1$$

ist, folgt

$$\|x\|_{0^*} \geq \left\langle x, \frac{z}{\|z\|_0} \right\rangle = \frac{1}{\|z\|_0} \langle x, z \rangle = \underbrace{\frac{1}{\|z\|_0}}_{\geq 1} \underbrace{\|x\|_{0^*}}_{> 0},$$

das nur für $\|z\|_0 = 1$ möglich ist. Somit ist die Zwischenbehauptung bewiesen.

Damit können wir uns dem eigentlichen Beweis widmen, d.h. wir zeigen, dass

$$\mathcal{F}^*(x) = \sup_{y\in\mathcal{X}} \left\{ \langle x, y \rangle - \frac{1}{2}\|y\|_0^2 \right\} = \frac{1}{2}\|x\|_{0^*}^2$$

für alle $x \in \mathcal{X}$ gilt.

Sei dazu stets $x \in \mathcal{X}$ gegeben.

„ $\leq$ “ Wie üblich, können wir mithilfe der Zwischenbehauptung

$$\|x\|_{0^*} = \sup_{\|y\|_0 \leq 1} \langle x, y \rangle = \sup_{y \neq 0} \frac{\langle x, y \rangle}{\|y\|_0}$$

schreiben, sodass wir

$$\langle x, y \rangle = \frac{\langle x, y \rangle}{\|y\|_0} \cdot \|y\|_0 \leq \sup_{z \neq 0} \frac{\langle x, z \rangle}{\|z\|_0} \cdot \|y\|_0 = \|x\|_{0^*} \cdot \|y\|_0 \leq \frac{1}{2}\big(\|x\|_{0^*}^2 + \|y\|_0^2\big) \quad (2.38)$$

für alle $y \in \mathcal{X}$ mit $y \neq 0$ erhalten, wobei wir für die letzte Ungleichung die bekannte Ungleichung

$$ab \leq \frac{1}{2}\big(a^2 + b^2\big) \qquad \text{für alle } a, b \in \mathbb{R}$$

verwendet haben; für $y = 0$ gilt die Ungleichung (2.38) letztendlich aber auch. Durch Umstellen übersetzt sich die Ungleichung (2.38) dann zu

$$\langle x, y \rangle - \mathcal{F}(y) = \langle x, y \rangle - \frac{1}{2}\|y\|_0^2 \leq \frac{1}{2}\|x\|_{0^*}^2$$

für alle $y \in \mathcal{X}$, d.h. es gilt

$$\mathcal{F}^*(x) = \sup_{y \in \mathcal{X}} \left\{ \langle x, y \rangle - \frac{1}{2}\|y\|_0^2 \right\} \leq \frac{1}{2}\|x\|_{0^*}^2.$$

„ $\geq$ “ Aufgrund der Zwischenbehauptung finden wir ein $z = z(x) \in \mathcal{X}$ mit $\|z\|_0 = 1$ und $\|x\|_{0^*} = \langle x, z \rangle$. Dann definieren wir die Funktion

$$\mathcal{G} : \mathbb{R} \to \mathbb{R}, \quad \mathcal{G}(t) := \langle x, tz \rangle - \mathcal{F}(tz).$$

Für alle $t \in \mathbb{R}$ können wir

$$\mathcal{G}(t) = \langle x, tz \rangle - \frac{1}{2}\|tz\|_0^2 = t\|x\|_{0^*} - \frac{1}{2}t^2$$

aufgrund der Wahl von z schreiben, sodass $\mathcal{G}$ ihr Supremum für $t_{\max} = \|x\|_{0^*}$ annimmt. Einsetzen liefert

$$\mathcal{G}(t_{\max}) = \|x\|_{0^*}^2 - \frac{1}{2}\|x\|_{0^*}^2 = \frac{1}{2}\|x\|_{0^*}^2$$

und somit erhalten wir

$$\mathcal{F}^*(x) = \sup_{y \in \mathcal{X}} \{\langle x, y \rangle - \mathcal{F}(y)\} \geq \sup_{t \in \mathbb{R}} \{\langle x, tz \rangle - \mathcal{F}(tz)\} = \sup_{t \in \mathbb{R}} \mathcal{G}(t) = \frac{1}{2}\|x\|_{0^*}^2,$$

womit die Behauptung bewiesen ist. $\square$

Zur Veranschaulichung diskutieren wir ein

(2.2.13) Beispiel

Sei ein essentiell endlicher Wahrscheinlichkeitsraum $(\Omega', \mathcal{M}', \mathcal{P}')$ gegeben. Dann folgt aus Lemma (A.1.21), dass $L^p(\Omega) = L^q(\Omega)$ für alle $p, q \in [1, \infty]$ gilt und die einzelnen Normen äquivalent sind. Zudem handelt es sich bei $(L^2(\Omega), \|\cdot\|_2)$ um einen Hilbertraum. Somit ist der Satz (2.2.11) auf

$$\mathcal{F}_p : L^2(\Omega) \to [0, \infty), \quad \mathcal{F}_p(X) := \frac{1}{2}\|X\|_p^2$$

für $p \in [1, \infty]$ anwendbar. Obendrein haben wir in Korollar (1.2.2) gezeigt, dass die zu $\|\cdot\|_p$ duale Norm gerade $\|\cdot\|_q$ mit $q \in [1, \infty]$ und $\frac{1}{p} + \frac{1}{q} = 1$ ist. Demnach besitzt die Fenchel-Konjugierte $\mathcal{F}_p^* : L^2(\Omega) \to [0, \infty)$ von $\mathcal{F}_p$ die Gestalt

$$\mathcal{F}_p^*(X) = \frac{1}{2}\|X\|_q^2$$

für alle $X \in L^2(\Omega)$. ◇

Wenn wir in Satz (2.2.11) den $\mathbb{R}^n$ mit dem Standardskalarprodukt als Hilbertraum wählen und benutzen, dass alle Normen auf diesem Raum äquivalent sind, ergibt sich direkt dieses

(2.2.14) Korollar

Sei $n \in \mathbb{N}$ und $\|\cdot\|$ eine Norm auf dem $\mathbb{R}^n$. Dazu sei

$$\mathcal{F} : \mathbb{R}^n \to [0, \infty), \quad \mathcal{F}(x) := \frac{1}{2}\|x\|^2.$$

Die Fenchel-Konjugierte $\mathcal{F}^* : \mathbb{R}^n \to [0, \infty)$ von $\mathcal{F}$ ist dann gegeben durch

$$\mathcal{F}^*(x) = \frac{1}{2}\|x\|_*^2$$

für alle $x \in \mathbb{R}^n$, wobei $\|\cdot\|_*$ die zu $\|\cdot\|$ duale Norm sei. ◇

Das bekannteste Beispiel dazu ist Folgendes.

(2.2.15) Beispiel

Sei $p \in [1, \infty]$. So kann man zeigen, dass die zu $\|\cdot\|_{p,\mathbb{R}^n}$ duale Norm gerade $\|\cdot\|_{q,\mathbb{R}^n}$ mit $q \in [1, \infty]$ und $\frac{1}{p} + \frac{1}{q} = 1$ ist; dazu gehe man wie im Beweis des nachfolgenden Lemmas (2.2.16) vor, d.h. zu einem beliebigen $x \in \mathbb{R}^n$ verwende man einerseits die Hölder-Ungleichung und andererseits wähle man ein Element $w \in \mathbb{R}^n$ mit $\|w\|_{p,\mathbb{R}^n} \leq 1$ und $x^T w = \|x\|_{q,\mathbb{R}^n}$. Dabei muss man jedoch für die Wahl dieses w die Fälle $p = 1, p \in (1, \infty)$ und $p = \infty$ separat behandeln.
Nach Korollar (2.2.14) erhalten wir dann jedenfalls als Fenchel-Konjugierte von

$$\mathcal{G}_p : \mathbb{R}^n \to [0, \infty), \quad \mathcal{G}_p(x) := \frac{1}{2}\|x\|_p^2$$

die Funktion

$$\mathcal{G}_p^* : \mathbb{R}^n \to [0, \infty), \quad \mathcal{G}_p^*(x) := \frac{1}{2}\|x\|_q^2.$$ ◇

Jedenfalls wollen wir das Korollar (2.2.14) auch auf f_p in (2.32), also

$$f_p : \mathbb{R}^M \to [0,\infty), \quad f_p(\hat{x}) = \frac{1}{2}\|B_p\hat{x}\|^2_{p,\mathbb{R}^{N_0}},$$

unter den unmittelbar vorher genannten Annahmen an den Finanzmarkt anwenden. Vorab wissen wir aus Lemma (2.2.10), dass $B_p \in \mathbb{R}^{N_0\times M}$ mit $N_0 > M$ sowie $\mathrm{Rang}(B_p) = M$ insbesondere für alle $p \in (1,\infty)$ gelten. Aus der linearen Algebra folgt dann, dass B_p unendlich viele Linksinversen $B_p^\dagger \in \mathbb{R}^{M\times N_0}$ besitzt, wovon aber keine eine Rechtsinverse ist, d.h. für jede Linksinverse $B_p^\dagger$ gilt

$$B_p^\dagger B_p = I_M, \qquad \text{aber} \qquad B_pB_p^\dagger \neq I_{N_0}, \tag{2.39}$$

wobei für alle $n \in \mathbb{N}$ mit I_n die Einheitsmatrix in $\mathbb{R}^{n\times n}$ bezeichnet sei. Aufgrund des Phänomens in (2.39) sind wir im nachfolgenden Lemma gezwungen, eine Bedingung an die Linksinversen zu stellen. Dazu verwenden wir Operatornormen für Matrizen: Zu einer Matrix $C \in \mathbb{R}^{n\times n}$ sei

$$\|C\|_{p,\mathbb{R}^{n\times n}} := \sup_{x\neq 0} \frac{\|Cx\|_{p,\mathbb{R}^n}}{\|x\|_{p,\mathbb{R}^n}}.$$

Eine Auflistung an Eigenschaften sowie konkrete Darstellungen dieser Operatornormen für manche p finden sich in [7, Abschnitt 4].

(2.2.16) Lemma
Sei ein Finanzmarkt S_t mit $t = 0,1$, also mit einer Zeitperiode, der der Eigenschaft $(A1)$ genügt, sowie $p \in (1,\infty)$ gegeben. Sei $B_p^\dagger \in \mathbb{R}^{M\times N_0}$ eine Linksinverse zu der Matrix $B_p \in \mathbb{R}^{N_0\times M}$ aus Lemma (2.2.10) mit

$$\|B_pB_p^\dagger\|_{p,\mathbb{R}^{N_0\times N_0}} \leq 1, \tag{2.40}$$

dann ist die zu $\mathfrak{r}_p$ duale Norm gegeben durch

$$\mathbb{R}^M \to [0,\infty), \quad \hat{x} \mapsto \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}$$

mit $q \in (1,\infty)$ und $\frac{1}{p} + \frac{1}{q} = 1$. ◇

(2.2.17) Bemerkung

a) Die Bedingung (2.40) ist sogar äquivalent zu $\|B_pB_p^\dagger\|_{p,\mathbb{R}^{N_0\times N_0}} = 1$.

   Dazu: Wir müssen nur zeigen, dass für jede Linksinverse $B_p^\dagger$ von B_p, die der Ungleichung (2.40) genügt, sogar $\|B_pB_p^\dagger\|_{p,\mathbb{R}^{N_0\times N_0}} = 1$ gilt.
   Sei also eine solche Linksinverse $B_p^\dagger$ gegeben. Wegen

$$(B_pB_p^\dagger)^2 = (B_pB_p^\dagger)(B_pB_p^\dagger) = B_p(B_p^\dagger B_p)B_p^\dagger = B_pB_p^\dagger \tag{2.41}$$

ist $B_pB_p^\dagger$ eine Projektionsmatrix. Da $B_p^\dagger$ eine Linksinverse von B_p ist, kann $B_p^\dagger$ nicht die Nullmatrix in $\mathbb{R}^{M\times N_0}$ sein. Wegen $\mathrm{Rang}(B_p) = M$ ist demnach ebenfalls $B_pB_p^\dagger$ nicht die Nullmatrix in $\mathbb{R}^{N_0\times N_0}$. Somit existiert ein $y \in \mathbb{R}^{N_0}$ mit

$$x := (B_pB_p^\dagger)y \neq 0.$$

Zusammen mit der Projektionseigenschaft (2.41) erhalten wir dann

$$(B_pB_p^\dagger)x = (B_pB_p^\dagger)^2 y = (B_pB_p^\dagger)y = x$$

und aufgrund von $x \neq 0$ ergibt sich

$$\|B_pB_p^\dagger\|_{p,\mathbb{R}^{N_0\times N_0}} = \sup_{z\neq 0} \frac{\|(B_pB_p^\dagger)z\|_{p,\mathbb{R}^{N_0}}}{\|z\|_{p,\mathbb{R}^{N_0}}} \geq \frac{\|(B_pB_p^\dagger)x\|_{p,\mathbb{R}^{N_0}}}{\|x\|_{p,\mathbb{R}^{N_0}}} = 1.$$

In Kombination mit der Ungleichung (2.40) liefert dies schließlich

$$\|B_pB_p^\dagger\|_{p,\mathbb{R}^{N_0\times N_0}} = 1.$$

b) Die Aussage bzgl. der dualen Norm in Lemma (2.2.16) gilt genauso für $p \in \{1,\infty\}$, aber da dies für unsere Zwecke nicht relevant ist, verzichten wir hier auf die damit verbundenen Fallunterscheidungen. $\diamond$

Beweis von Lemma (2.2.16)
Wir müssen also nachweisen, dass

$$\sup_{\|B_p\hat{y}\|_{p,\mathbb{R}^{N_0}}\leq 1} \hat{x}^T\hat{y} = \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}$$

für alle $\hat{x} \in \mathbb{R}^M$ gilt. Dazu sei stets $\hat{x} \in \mathbb{R}^M$ gegeben. Œ sei dabei $\hat{x} \neq \hat{0}$; denn für $\hat{x} = \hat{0}$ trifft diese Gleichheit offenbar zu.

„ $\leq$ “ Für alle $\hat{y} \in \mathbb{R}^M$ mit $\|B_p\hat{y}\|_{p,\mathbb{R}^{N_0}} \leq 1$ ergibt sich aufgrund der Hölder-Ungleichung

$$\hat{x}^T\hat{y} = (\hat{x}^TB_p^\dagger)(B_p\hat{y}) \leq \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}} \cdot \underbrace{\|B_p\hat{y}\|_{p,\mathbb{R}^{N_0}}}_{\leq 1} \leq \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}},$$

sodass

$$\sup_{\|B_p\hat{y}\|_{p,\mathbb{R}^{N_0}}\leq 1} \hat{x}^T\hat{y} \leq \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}$$

folgt.

„ $\geq$ “ Da ebenfalls

$$\mathrm{Rang}\big((B_p^\dagger)^T\big) = \mathrm{Rang}(B_p^\dagger) = M$$

gilt, ist $(B_p^\dagger)^T\hat{x} \neq 0$ aufgrund der Annahme $\hat{x} \neq \hat{0}$. Damit ist

$$\hat{w} := \frac{1}{\left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}}} \cdot B_p^\dagger \underbrace{\begin{pmatrix} \left|\left[(B_p^\dagger)^T\hat{x}\right]_1\right|^{q-1} \operatorname{sign}\left(\left[(B_p^\dagger)^T\hat{x}\right]_1\right) \\ \vdots \\ \left|\left[(B_p^\dagger)^T\hat{x}\right]_{N_0}\right|^{q-1} \operatorname{sign}\left(\left[(B_p^\dagger)^T\hat{x}\right]_{N_0}\right) \end{pmatrix}}_{=:\, z}$$

wohldefiniert und es gilt

$$\begin{aligned} \|B_p\hat{w}\|_{p,\mathbb{R}^{N_0}} &= \frac{1}{\left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}}} \cdot \left\|B_pB_p^\dagger z\right\|_{p,\mathbb{R}^{N_0}} \\ &\leq \frac{1}{\left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}}} \cdot \underbrace{\|B_pB_p^\dagger\|_{p,\mathbb{R}^{N_0\times N_0}}}_{\leq 1} \cdot \|z\|_{p,\mathbb{R}^{N_0}} \\ &\leq \frac{1}{\left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}}} \cdot \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}} = 1, \end{aligned}$$

da wegen $\frac{1}{p} + \frac{1}{q} = 1$ gerade $p = \frac{q}{q-1}$ und somit $p(q-1) = q$ gilt. Nach Definition von $\hat{w}$ erhalten wir außerdem

$$\begin{aligned} \hat{x}^T\hat{w} &= \frac{1}{\left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}}} \cdot \underbrace{\hat{x}^TB_p^\dagger}_{=\left((B_p^\dagger)^T\hat{x}\right)^T} \cdot z = \frac{1}{\left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{\frac{q}{p}}} \cdot \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{q} \\ &= \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}^{q-\frac{q}{p}} = \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}}, \end{aligned}$$

da $q - \frac{q}{p} = 1$ gilt. Dies impliziert

$$\sup_{\|B_p\hat{y}\|_{p,\mathbb{R}^{N_0}} \leq 1} \hat{x}^T\hat{y} \geq \hat{x}^T\hat{w} = \left\|(B_p^\dagger)^T\hat{x}\right\|_{q,\mathbb{R}^{N_0}},$$

womit wir die Identität gezeigt haben. □

Nun drängt sich die Frage auf, ob die Bedingung (2.40) in obigem Lemma überhaupt erfüllbar ist; denn - wie oben schon erwähnt - ist keine der Linksinversen eine Rechtsinverse. Immerhin können wir dies für $p = 2$ garantieren, indem wir sogar eine entsprechende Linksinverse angeben.

(2.2.18) Lemma

Sei ein Finanzmarkt S_t mit $t = 0, 1$, also mit einer Zeitperiode, der der Eigenschaft $(A1)$ genügt, gegeben. Zudem sei $B_2 \in \mathbb{R}^{N_0\times M}$ die Matrix aus Lemma (2.2.10) für $p = 2$. Dann *ist $B_2^T B_2 \in \mathbb{R}^{M\times M}$ invertierbar und $B_2^\dagger := (B_2^T B_2)^{-1} B_2^T \in \mathbb{R}^{M\times N_0}$ ist eine Linksinverse zu B_2 mit*

$$\|B_2 B_2^\dagger\|_{2,\mathbb{R}^{N_0\times N_0}} \leq 1,$$

sodass die zu $\hat{\mathfrak{r}}_2$ duale Norm durch

$$\mathbb{R}^M \to [0,\infty), \quad \hat{x} \mapsto \left\|(B_2^\dagger)^T \hat{x}\right\|_{2,\mathbb{R}^{N_0}}$$

gegeben ist. ◇

Beweis
Zunächst ist $B_2^T B_2 \in \mathbb{R}^{M\times M}$ offenbar symmetrisch und wegen $\mathrm{Rang}(B_2) = M$ ist $B_2^T B_2$ sogar positiv definit, d.h. $B_2^T B_2$ ist insbesondere invertierbar. Darüber hinaus folgt unmittelbar

$$B_2^\dagger B_2 = (B_2^T B_2)^{-1}(B_2^T B_2) = I_M,$$

also ist $B_2^\dagger$ eine Linksinverse zu B_2.
Nun gilt nach [7, Satz 4.10], dass für jede Matrix $C \in \mathbb{R}^{n\times n}$ die Beziehung

$$\|C\|_{2,\mathbb{R}^{n\times n}} = \sqrt{\lambda_{\max}(C^T C)}$$

gilt, wobei $\lambda_{\max}(C^T C)$ der größte Eigenwert von $C^T C$ sei; dazu sei bemerkt, dass $C^T C$ symmetrisch und positiv semi-definit und somit diagonalisierbar ist sowie nur nicht-negative Eigenwerte besitzt.
Jedenfalls sei jetzt $n = N_0$ und $C := B_2 B_2^\dagger \in \mathbb{R}^{N_0\times N_0}$, so erhalten wir

$$C^T = (B_2^T)^T \big((B_2^T B_2)^{-1}\big)^T B_2^T = B_2\big((B_2^T B_2)^T\big)^{-1} B_2^T = B_2(B_2^T B_2)^{-1} B_2^T = C,$$

d.h. C ist symmetrisch. Daher ist $C^T C = C^2$ und obendrein ist

$$\begin{aligned} C^2 = \big(B_2(B_2^T B_2)^{-1} B_2^T\big)\big(B_2(B_2^T B_2)^{-1} B_2^T\big) &= B_2(B_2^T B_2)^{-1}(B_2^T B_2)(B_2^T B_2)^{-1} B_2^T \\ &= B_2(B_2^T B_2)^{-1} B_2^T = C. \end{aligned}$$

Dies liefert zum Einen

$$\|C\|_{2,\mathbb{R}^{N_0\times N_0}} = \sqrt{\lambda_{\max}(C^T C)} = \sqrt{\lambda_{\max}(C)}$$

und zum Anderen, dass C eine symmetrische Projektionsmatrix ist. Dann ist aus der linearen Algebra bekannt, dass C nur die Eigenwerte 0 und/oder 1 besitzt, das insbesondere $\lambda_{\max}(C) \in \{0,1\}$ und somit $\|C\|_{2,\mathbb{R}^{N_0\times N_0}} \leq 1$ impliziert. Dann ergibt sich der Rest direkt aus Lemma (2.2.16). □

Aufgrund dieses Lemmas haben wir Grund zur Hoffnung, dass auch für allgemeine $p \in (1,\infty)$ Linksinversen existieren, die die Bedingung (2.40) erfüllen. Dies müsste aber in der Praxis erprobt werden. Als Schwierigkeit kommt dabei noch hinzu, dass wir die Operatornorm für Matrizen für allgemeine p nicht so leicht wie für $p = 2$ auswerten können.

Zusammenfassend können wir also festhalten: Es sei ein Finanzmarkt S_t mit $t = 0, 1$, also mit einer Zeitperiode, der das Axiom $(A1)$ und die Ungleichung (2.28) erfüllt, sowie $p \in (1, \infty)$ gegeben. Falls nun die Matrix B_p aus Lemma (2.2.10) eine Linksinverse $B_p^\dagger$ besitzt, die der Bedingung (2.40) genügt - das nach Lemma (2.2.18) zumindest für $p = 2$ garantiert möglich ist -, so sind mit dem Risikomaß $\mathfrak{r}_p$ aus Lemma (2.2.10) alle Voraussetzungen des Satzes (2.2.8) bis auf die Ungleichung (2.29) erfüllt und die Fenchel-Konjugierte von f_p aus (2.32) ist wegen Lemma (2.2.16) und Korollar (2.2.14) gegeben durch

$$f_p^* : \mathbb{R}^M \to [0, \infty), \quad f_p^*(\hat{x}) = \frac{1}{2} \left\| (B_p^\dagger)^T \hat{x} \right\|_{q, \mathbb{R}^{N_0}}^2$$

mit $q \in (1, \infty)$ und $\frac{1}{p} + \frac{1}{q} = 1$. Aufgrund derselben Argumente wie für f_p, die im Zusammenhang mit (2.33), (2.34) und (2.35) angeführt wurden, ist $f_p^* \in C^1(\mathbb{R}^M)$ mit

$$\begin{aligned} \nabla f_p^*(\hat{x}) = & \frac{1}{\left\| (B_p^\dagger)^T \hat{x} \right\|_{q, \mathbb{R}^{N_0}}^{q-2}} B_p^\dagger \Big(\big[(B_p^\dagger)^T \hat{x} \big]_1 \left| \big[(B_p^\dagger)^T \hat{x} \big]_1 \right|^{q-2}, \dots, \\ & \big[(B_p^\dagger)^T \hat{x} \big]_{N_0} \left| \big[(B_p^\dagger)^T \hat{x} \big]_{N_0} \right|^{q-2} \Big)^T \end{aligned} \tag{2.42}$$

für $\hat{x} \neq \hat{0}$ und $\nabla f_p^*(\hat{0}) = 0$, sodass wir dies schließlich nur noch in die charakterisierende Bedingung (2.29) einsetzen müssen.

Abschließend gehen wir erneut auf den schönen Spezialfall $p = 2$ ein: Dazu wählen wir die Linksinverse $B_2^\dagger = (B_2^T B_2)^{-1} B_2^T$ aus Lemma (2.2.18). Dann vereinfacht sich der Gradient in (2.42) für $\hat{x} \neq \hat{0}$ zu

$$\nabla f_2^*(\hat{x}) = B_2^\dagger \Big(\big[(B_2^\dagger)^T \hat{x} \big]_1, \dots, \big[(B_2^\dagger)^T \hat{x} \big]_{N_0} \Big)^T = B_2^\dagger (B_2^\dagger)^T \hat{x}, \tag{2.43}$$

wobei

$$\begin{aligned} B_2^\dagger (B_2^\dagger)^T &= (B_2^T B_2)^{-1} B_2^T (B_2^T)^T \big((B_2^T B_2)^{-1} \big)^T \\ &= (B_2^T B_2)^{-1} (B_2^T B_2) (B_2^T B_2)^{-1} \\ &= (B_2^T B_2)^{-1} \end{aligned} \tag{2.44}$$

gilt; für $\hat{x} = \hat{0}$ ist die Identität (2.43) aber offenbar auch richtig.
Jedenfalls kann man jetzt leicht nachrechnen, dass für alle $i, j = 1, \dots, M$ die Beziehung

$$(B_2^T B_2)_{i,j} = \begin{cases} \operatorname{Var}(S_1^i) & \text{für } i = j, \\ \operatorname{Kov}(S_1^i, S_1^j) & \text{für } i \neq j \end{cases}$$

gilt, das aber wirklich nur für $p = 2$ stimmt. Somit entspricht $B_2^T B_2$ der sogenannten **Kovarianzmatrix** von S_1 und wir schreiben

$$\mathrm{Kov}(S_1) := B_2^T B_2.$$

Unter Benutzung dieser Notation erhalten wir also zusammen mit den Gleichungen (2.43) und (2.44) die Identität

$$\nabla f_2^*(\hat{x}) = \big(\mathrm{Kov}(S_1)\big)^{-1}\hat{x}$$

für alle $\hat{x} \in \mathbb{R}^M$, sodass sich die Bedingung (2.29) zu

$$\hat{S}_0^T \big(\mathrm{Kov}(S_1)\big)^{-1}\big(E[\hat{S}_1] - R\hat{S}_0\big) > 0$$

übersetzt. Diese Bedingung wurde aber bereits in [9, Abschnitt 4] entdeckt, indem Maier-Paape und Zhu auch dort das auf der Standardabweichung basierende Risikomaß $\mathfrak{r}_2$ verwendeten.

A Anhang

A.1 Wahrscheinlichkeitsräume und Zufallsvariablen

Wir betrachten einen **allgemeinen Wahrscheinlichkeitsraum**

$$(\Omega, \mathcal{M}, \mathcal{P}),$$

wobei Ω eine beliebige nicht-leere Menge, $\mathcal{M}$ als Teilmenge der Potenzmenge $\mathfrak{P}(\Omega)$ eine σ-Algebra und $\mathcal{P} : \mathcal{M} \to [0,1]$ ein Wahrscheinlichkeitsmaß ist. Aus der Maß- und Intergrationstheorie sind somit für diesen Raum die Konzepte der Messbarkeit und (Quasi-)Integrierbarkeit bekannt. Darauf aufbauend werden wir einige Begriffe wiederholen bzw. einführen und wichtige Ergebnisse vorstellen.

(A.1.1) Definition (Zufallsvariablen)
Eine messbare Abbildung $X : (\Omega, \mathcal{M}) \to (\mathbb{R}, \mathcal{B})$, wobei $\mathcal{B}$ die Borel-Algebra auf $\mathbb{R}$ bezeichnet, heißt Zufallsvariable*.* $\diamond$

Für eine Zufallsvariable $X : (\Omega, \mathcal{M}) \to (\mathbb{R}, \mathcal{B})$ schreiben wir dann nur noch $X : \Omega \to \mathbb{R}$, da immer klar sein wird, welche σ-Algebren gemeint sind.

(A.1.2) Definition (Verteilungsfunktion)
Zu einer Zufallsvariablen $X : \Omega \to \mathbb{R}$ bezeichnet man mit

$$\mathcal{P}^X : \mathcal{B} \to [0,1], \quad \mathcal{P}^X(A) := \mathcal{P}\big(X^{-1}(A)\big)$$

die Verteilung von X unter $\mathcal{P}$ *und mit*

$$F^X : \mathbb{R} \to [0,1], \quad F^X(x) := \mathcal{P}^X\big((-\infty, x]\big)$$

die Verteilungsfunktion zu X*.* $\diamond$

(A.1.3) Definition (Erwartungswerte)
Zu einer (quasi-)integrierbaren Zufallsvariablen $X : \Omega \to \mathbb{R}$ definiert man den Erwartungswert von X *durch*

$$EX := E[X] := \int_\Omega X \,\mathrm{d}\mathcal{P} := \int_\Omega X(\omega) \,\mathrm{d}\mathcal{P}(\omega).$$

$\diamond$

(A.1.4) Definition (Essentielles Supremum und Infimum)
Für eine gegebene Zufallsvariable $X : \Omega \to \mathbb{R}$ *nennt man*

$$\textit{ess sup}\, X := \inf_{\substack{M \in \mathcal{M}, \\ \mathcal{P}(M)=0}} \sup_{\omega \in \Omega \setminus M} X(\omega) \in (-\infty, \infty]$$

das essentielle Supremum von X *und*

$$\textit{ess inf}\, X := \sup_{\substack{M \in \mathcal{M}, \\ \mathcal{P}(M)=0}} \inf_{\omega \in \Omega \setminus M} X(\omega) \in [-\infty, \infty)$$

das essentielle Infimum von X. ◇

Für das essentielle Supremum bzw. Infimum bleiben einige Ungleichungen, die für das gewöhnliche Supremum bzw. Infimum schon bekannt sind, erhalten. Notwendig für dessen Formulierung ist folgende Sprechweise, die besonders im Zusammenhang mit den „L^p-Räumen“ nützlich ist.

(A.1.5) Definition ($\mathcal{P}$-fast sicher)
Man sagt, dass eine Beziehung für $\mathcal{P}$-fast alle *oder nur* fast alle $\omega \in \Omega$ *- sofern das zugrunde liegende Wahrscheinlichkeitsmaß bekannt ist - erfüllt ist, falls eine* $\mathcal{P}$*-Nullmenge* $M \in \mathcal{M}$ *existiert, sodass diese Beziehung für alle* $\omega \in \Omega \setminus M$ *gilt. Zur Abkürzung sagt man auch, dass die jeweilige Beziehung* $\mathcal{P}$-fast sicher ($\mathcal{P}$-f.s.) *bzw.* fast sicher (f.s.) *gilt.* ◇

(A.1.6) Lemma
Seien die Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ *gegeben. Dann gelten die Ungleichungen*

(i) $\textit{ess inf}\, X \leq \textit{ess sup}\, X$,

(ii) $\textit{ess inf}\, X = -\, \textit{ess sup}\, (-X)$,

(iii) $X \leq \textit{ess sup}\, X$ *f.s.*,

(iv) $\textit{ess inf}\, X \leq X$ *f.s.*,

(v) $\textit{ess sup}\, (X + Y) \leq \textit{ess sup}\, X + \textit{ess sup}\, Y$,

(vi) $\textit{ess inf}\, (X + Y) \geq \textit{ess inf}\, X + \textit{ess inf}\, Y$,

sowie für $X \leq Y$ *f.s.*

(vii) $\textit{ess sup}\, X \leq \textit{ess sup}\, Y$,

(viii) $\textit{ess inf}\, X \leq \textit{ess inf}\, Y$. ◇

(A.1.7) Bemerkung
In diesem Lemma sowie in diesem gesamten Werk benutzen wir folgende zum Teil willkürliche Festlegungen: Zu $a \in [-\infty, \infty]$ sei

$$a \cdot \pm\infty := \begin{cases} \pm\infty & \text{für } a > 0, \\ 0 & \text{für } a = 0, \\ \mp\infty & \text{für } a < 0 \end{cases}$$

und für $b \in (-\infty, \infty]$ und $c \in [-\infty, \infty)$ sei

$$b + \infty := \infty \qquad \text{und} \qquad c - \infty := -\infty.$$

Ferner definieren wir

$$\infty - \infty := 0, \qquad \frac{1}{\infty} := 0 \qquad \text{sowie} \qquad \infty^p := \infty.$$

für $p > 0$. ◇

Beweis von Lemma (A.1.6)
(i) Sei $\epsilon > 0$. Dann finden wir eine $\mathcal{P}$-Nullmenge $M \in \mathcal{M}$ mit

$$\text{ess sup } X + \frac{\epsilon}{2} \geq \sup_{\omega \in \Omega \setminus M} X(\omega) \qquad \text{und} \qquad \text{ess inf } X \leq \inf_{\omega \in \Omega \setminus M} X(\omega) + \frac{\epsilon}{2}.$$

Somit ist

$$\begin{aligned} \text{ess inf } X \leq \inf_{\omega \in \Omega \setminus M} X(\omega) + \frac{\epsilon}{2} \leq \sup_{\omega \in \Omega \setminus M} X(\omega) + \frac{\epsilon}{2} &\leq (\text{ess sup } X + \frac{\epsilon}{2}) + \frac{\epsilon}{2} \\ &= \text{ess sup } X + \epsilon. \end{aligned}$$

Weil ϵ beliebig gewählt war, ist die Ungleichung gezeigt.

(ii) Dies erhalten wir aus

$$\begin{aligned} \text{ess inf } X = \text{ess inf } -(-X) &= \sup_{\substack{M \in \mathcal{M}, \\ \mathcal{P}(M)=0}} \inf_{\omega \in \Omega \setminus M} -(-X(\omega)) \\ &= -\inf_{\substack{M \in \mathcal{M}, \\ \mathcal{P}(M)=0}} \sup_{\omega \in \Omega \setminus M} -X(\omega) \\ &= -\text{ ess sup } (-X). \end{aligned}$$

(iii) Nach Definition existiert eine Folge $(M_n)_{n \in \mathbb{N}} \subseteq \mathcal{M}$ von $\mathcal{P}$-Nullmengen mit

$$\sup_{\omega \in \Omega \setminus M_n} X(\omega) \xrightarrow[n \to \infty]{} \text{ess sup } X.$$

Dann sei $M := \bigcup_{i=1}^{\infty} M_i$, so ist M offenbar auch eine $\mathcal{P}$-Nullmenge und es gilt

$$\text{ess sup } X \leq \sup_{\omega \in \Omega \setminus M} X(\omega) \leq \sup_{\omega \in \Omega \setminus M_n} X(\omega) \xrightarrow[n \to \infty]{} \text{ess sup } X,$$

sodass wir

$$\text{ess sup } X = \sup_{\omega \in \Omega \setminus M} X(\omega)$$

erhalten. Da nun

$$X(\omega) \leq \sup_{\omega \in \Omega \setminus M} X(\omega) \quad \text{für alle } \omega \in \Omega \setminus M$$

gilt und M eine $\mathcal{P}$-Nullmenge ist, folgt $X \leq \text{ess sup } X$ f.s..

(iv) Dies ergibt sich unmittelbar aus (ii) und (iii).

(v) Sei wiederum $\epsilon > 0$, so existiert eine $\mathcal{P}$-Nullmenge $M \in \mathcal{M}$ mit

$$\text{ess sup } X + \frac{\epsilon}{2} \geq \sup_{\omega \in \Omega \setminus M} X(\omega) \quad \text{und} \quad \text{ess sup } Y + \frac{\epsilon}{2} \geq \sup_{\omega \in \Omega \setminus M} Y(\omega).$$

Damit erhalten wir die Abschätzungen

$$\begin{aligned} \text{ess sup } (X+Y) \leq \sup_{\omega \in \Omega \setminus M} (X+Y)(\omega) &\leq \sup_{\omega \in \Omega \setminus M} X(\omega) + \sup_{\omega \in \Omega \setminus M} Y(\omega) \\ &\leq (\text{ess sup } X + \frac{\epsilon}{2}) + (\text{ess sup } Y + \frac{\epsilon}{2}) \\ &= \text{ess sup } X + \text{ess sup } Y + \epsilon. \end{aligned}$$

Da dies für jedes $\epsilon > 0$ erfüllt ist, folgt die behauptete Ungleichung.

(vi) Dies folgt direkt aus der Kombination von (ii) und (v).

(vii) Nach Annahme existiert also eine $\mathcal{P}$-Nullmenge $N \in \mathcal{M}$ mit $X(\omega) \leq Y(\omega)$ für alle $\omega \in \Omega \setminus N$. Dann erhalten wir für jede weitere $\mathcal{P}$-Nullmenge $M \in \mathcal{M}$ die Ungleichungskette

$$\text{ess sup } X \leq \sup_{\omega \in \Omega \setminus (M \cup N)} X(\omega) \leq \sup_{\omega \in \Omega \setminus (M \cup N)} Y(\omega) \leq \sup_{\omega \in \Omega \setminus M} Y(\omega)$$

und somit die Behauptung.

(viii) Dies ergibt sich erneut aus (ii) und (vii). $\square$

(A.1.8) Definition

Zu einer Zufallsvariablen $X : \Omega \to \mathbb{R}$ und einem $p \in [1, \infty]$ definiert man

$$\|X\|_p := \begin{cases} \left(E[|X|^p]\right)^{\frac{1}{p}} & \textit{für } p \in [1, \infty), \\ \textit{ess sup } |X| & \textit{für } p = \infty \end{cases} \quad \in [0, \infty],$$

wobei $|X|^p \geq 0$ für alle $p \in [1,\infty)$ gilt und $|X|^p$ somit quasi-integrierbar ist.

Dazu definiert man

$$\mathcal{L}^p(\Omega) := \mathcal{L}^p(\Omega, \mathcal{M}, \mathcal{P}) := \{X : \Omega \to \mathbb{R} \text{ Zufallsvariable} \mid \|X\|_p \in [0,\infty)\}.$$

◇

(A.1.9) Bemerkung
Eine interessante Eigenschaft weist hier $\mathcal{L}^\infty(\Omega)$ auf: Sei dazu $X \in \mathcal{L}^\infty(\Omega)$, d.h. es existiert ein $C \in \mathbb{R}$ mit $\|X\|_\infty \leq C$. Dann finden wir eine $\mathcal{P}$-Nullmenge $M \in \mathcal{M}$, sodass $|X(\omega)| \leq C$ für alle $\omega \in \Omega \setminus M$ ist.

Denn sei $n \in \mathbb{N}$, so ist $\|X\|_\infty < C + \frac{1}{n}$. Nach Definition des essentiellen Supremums finden wir eine $\mathcal{P}$-Nullmenge $M_n \in \mathcal{M}$ mit

$$\sup_{\omega \in \Omega \setminus M_n} |X(\omega)| < C + \frac{1}{n}.$$

Dann ist aber auch $M := \bigcup_{n \in \mathbb{N}} M_n$ eine $\mathcal{P}$-Nullmenge und wir erhalten

$$\sup_{\omega \in \Omega \setminus M} |X(\omega)| \leq \sup_{\omega \in \Omega \setminus M_n} |X(\omega)| < C + \frac{1}{n}$$

für alle $n \in \mathbb{N}$, d.h. es muss $\sup\limits_{\omega \in \Omega \setminus M} |X(\omega)| \leq C$ gelten. ◇

Ohne Beweis zitieren wir folgenden zentralen

(A.1.10) Satz (Hölder-Ungleichung)
Für alle Zufallsvariablen $X, Y : \Omega \to \mathbb{R}$ gilt die Beziehung

$$E[|XY|] \leq \|X\|_p \cdot \|Y\|_q,$$

wobei $p, q \in [1,\infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$ sind. ◇

(A.1.11) Bemerkung

a) Mit der Hölder-Ungleichung kann gezeigt werden, dass $\|\cdot\|_p$ für $p \in [1,\infty]$ auf der Menge der Zufallsvariablen subadditiv ist. Somit können wir leicht einsehen, dass $\mathcal{L}^p(\Omega)$ ein Vektorraum ist und dass es sich bei $\|\cdot\|_p$ um eine Semi-Norm auf $\mathcal{L}^p(\Omega)$ handelt.

b) Im Spezialfall $p = 2$ ist durch

$$\langle \cdot, \cdot \rangle_2 : \mathcal{L}^2(\Omega) \times \mathcal{L}^2(\Omega) \to \mathbb{R}, \quad \langle X, Y \rangle_2 := E[XY]$$

eine positiv semi-definite symmetrische Bilinearform definiert. Dies ergibt sich direkt aus den Integraleigenschaften. ◇

Wie für das Lebesgue-Maß können wir daraus einen normierten Raum bzw. für $p = 2$ einen Prä-Hilbertraum konstruieren, indem wir den Unterraum der Zufallsvariablen, die der Zufallsvariablen $X \equiv 0$ für fast alle $\omega \in \Omega$ gleichen, „herausfaktorisieren".

(A.1.12) Lemma (L^p-Räume)
Zu $p \in [1, \infty]$ sei

$$\mathcal{N}_p(\Omega) := \{X : \Omega \to \mathbb{R} \ \textit{Zufallsvariable} \mid X \equiv 0 \ \textit{f.s.}\}$$

und

$$L^p(\Omega) := {}^{\mathcal{L}^p(\Omega)}\!\big/_{\mathcal{N}_p(\Omega)}.$$

Dann ist $(L^p(\Omega), \|\cdot\|_p)$ ein normierter Raum. Für $p = 2$ ist $(L^2(\Omega), \langle\cdot,\cdot\rangle_2)$ sogar ein Prä-Hilbertraum. ◇

(A.1.13) Bemerkung

a) In Lemma (A.1.12) haben wir unterschlagen, wie die Normen genau auf den L^p-Räumen definiert sind: Nach Konstruktion gilt für eine Restklasse $[X] \in L^p(\Omega)$, dass für jeden weiteren Repräsentanten Y dieser Restklasse $X = Y$ f.s. erfüllt ist. Somit ist $\|X\|_p = \|Y\|_p$. Daher ist es legitim festzulegen, dass

$$\|[X]\|_p := \|X\|_p$$

sei.

b) Wegen der schönen Eigenschaft, dass $\mathcal{P}(\Omega) = 1$ ist, sind L^p-Räume auch immer nicht-trivial; denn die Zufallsvariablen $X : \Omega \to \mathbb{R}$, für die ein $C \in \mathbb{R}$ mit $X \equiv C$ f.s. existiert, erfüllen $\|X\|_p = |C|$ für beliebiges $p \in [1, \infty]$, sodass $[X] \in L^p(\Omega)$ gilt.

c) Nun können wir auch ein Gleichheitszeichen bzw. Ungleichungszeichen zwischen den Elementen eines L^p-Raums definieren, das wir hier aber nur skizzieren wollen: Zu $p \in [1, \infty]$ sowie $* \in \{>, \geq, <, \leq\}$ schreiben wir für zwei Restklassen $[X], [Y] \in L^p(\Omega)$ genau dann $[X] * [Y]$, falls $X * Y$ f.s. gilt. Dann kann man leicht nachrechnen, dass diese Definition vertreterunabhängig ist, sowie die Relationen die meisten üblichen Eigenschaften besitzen.

d) Der Einfachheit halber schreiben wir hier immer X statt $[X]$, wobei dann im Kontext immer klar ist, wann wir Repräsentanten benutzen (müssen); dies äußert sich z. B. bei der Verwendung von „f.s.". Außerdem arbeiten wir regelmäßig bei gegebenen Restklassen $X, Y, Z \in L^p(\Omega)$ mit Mengen wie $\{X * Y\}$ mit $*$ wie in c) sowie mit Integralen mit Z als Integranden; in diesen Situationen handelt es sich dann natürlich immer um Repräsentanten. Dies ist auch insofern sinnvoll, da sich bei unterschiedlicher Wahl von Repräsentanten solche Mengen nur um Nullmengen unterscheiden und solche Integrale sogar gar nicht variieren.

Allgemein greifen wir immer auf Vetreter zurück, wenn wir konkret deren Eigenschaften als Abbildung benötigen.
Obendrein nennen wir X dann auch „Zufallsvariable", das nach Definition (A.1.1) natürlich nicht richtig ist. Aber es handelt sich hierbei auch nur um eine vereinfachende Sprechweise, mit der wir vorsichtig umgehen müssen. Beispielsweise sprechen wir dann in der Situation von b) von einer konstanten Zufallsvariable. Hingegen sagen wir, dass eine Zufallsvariable $X \in L^p(\Omega)$ nicht-konstant ist, falls für kein $C \in \mathbb{R}$ die Beziehung $X = C$ gilt, wobei wir hier erstmals die neue Operation „=" verwendet haben. ◇

Für diese Räume stellen wir nun ein erstes zentrales Resultat vor, das wir aber nur ohne Beweis nennen wollen.

(A.1.14) Satz (Fischer-Riesz)
Für alle $p \in [1, \infty]$ sind $(L^p(\Omega), \|\cdot\|_p)$ vollständig. ◇

(A.1.15) Bemerkung
Im Beweis des Satzes von Fischer-Riesz erhält man ein Nebenresultat, das gelegentlich in Beweisen Anwendung findet:

Gegeben sei ein $p \in [1, \infty]$, eine Folge $(X_n)_{n\in\mathbb{N}} \subseteq L^p(\Omega)$ und $X \in L^p(\Omega)$ mit $X_n \xrightarrow[n\to\infty]{} X$ in $L^p(\Omega)$. Dann finden wir eine Teilfolge $(X_{n_k})_{k\in\mathbb{N}} \subseteq (X_n)_{n\in\mathbb{N}}$, sodass

$$X_{n_k}(\omega) \xrightarrow[k\to\infty]{} X(\omega)$$

für fast alle $\omega \in \Omega$ zutrifft. Dies kürzen wir künftig durch

$$X_{n_k} \xrightarrow[k\to\infty]{pktw.} X \quad \text{f.s.}$$

ab. Dabei erkennen wir sofort, dass diese Eigenschaft vertreterunabhängig ist. ◇

Mit der Hölder-Ungleichung können wir wegen $\mathcal{P}(\Omega) = 1$ auch Beziehungen zwischen den einzelnen L^p-Räumen herstellen.

(A.1.16) Lemma
Für eine Zufallsvariable $X : \Omega \to \mathbb{R}$ und $p, q \in (1, \infty)$ mit $p \leq q$ ist

$$\|X\|_1 \leq \|X\|_p \leq \|X\|_q \leq \|X\|_\infty \tag{A.1}$$

und somit

$$L^\infty(\Omega) \subseteq L^q(\Omega) \subseteq L^p(\Omega) \subseteq L^1(\Omega). \tag{A.2}$$

◇

(A.1.17) Bemerkung
Daher ist das Funktional

$$L^p(\Omega) \to \mathbb{R}, \quad X \mapsto EX$$

für jedes $p \in [1, \infty]$ wohldefiniert, also endlich, und stetig; dazu benutze man die aus Lemma (A.1.16) resultierende Abschätzung

$$\left|EX\right| \leq E[|X|] = \|X\|_1 \leq \|X\|_p \tag{A.3}$$

für alle $X \in L^p(\Omega)$ sowie die Linearität des Erwartungswertes auf $L^p(\Omega)$. Dabei kann Letzteres u.A. deshalb angenommen werden, da jede Zufallsvariable $X \in L^p(\Omega)$ aufgrund der Ungleichung (A.3) insbesondere integrierbar ist. $\diamond$

Über die Inklusionsbeziehung in (A.2) können wir sogar noch konkretere Aussagen machen, wenn wir mehr Anforderungen an den zugrunde liegenden Wahrscheinlichkeitsraum stellen. Dafür müssen wir aber noch neue Bezeichnungen einführen, die Rockafellar et al. in [12, S. 53] definiert haben.

(A.1.18) Definition (Essentiell (un-)endliche Wahrscheinlichkeitsräume)
Der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ heißt essentiell endlich, *falls $\mathcal{P} : \mathcal{M} \to [0,1]$ nur endlich viele verschiedene Werte annimmt. Ansonsten wird er* essentiell unendlich *genannt.* $\diamond$

(A.1.19) Beispiele

a) Der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ ist offenbar essentiell endlich, falls Ω oder $\mathcal{M}$ endliche Mengen sind.

b) Sei $\Omega := [0,1]$, $\mathcal{P} := \mathscr{L}$ das Lebesgue-Maß und $\mathcal{M} := \mathcal{B} \cap \mathfrak{P}([0,1])$, wobei $\mathcal{B}$ die Borel-Algebra auf $\mathbb{R}$ und $\mathfrak{P}([0,1])$ die Potenzmenge von $[0,1]$ beschreiben. Dann ist $(\Omega, \mathcal{M}, \mathcal{P})$ ein Beispiel für einen essentiell unendlichen Wahrscheinlichkeitsraum. $\diamond$

Eine nützliche Charakterisierung von essentiell unendlichen Wahrscheinlichkeitsräumen, die Rockafellar et al. in [12, S. 68] unbewiesen und ohne Verweis behauptet haben, formulieren wir in

(A.1.20) Satz
Der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ ist genau dann essentiell unendlich, falls eine Mengenfolge $(M_n)_{n \in \mathbb{N}} \subseteq \mathcal{M}$ mit

$$\mathcal{P}(M_n) \xrightarrow[n \to \infty]{} 0 \qquad \textit{und} \qquad \mathcal{P}(M_n) > 0 \quad \textit{für alle } n \in \mathbb{N}$$

existiert. $\diamond$

Beweis
„ $\Leftarrow$ “ Wir nehmen an, dass eine wie oben beschriebene Mengenfolge $(M_n)_{n\in\mathbb{N}}$ existiere. Dann finden wir aufgrund der Annahme eine Teilfoge $(M_{n_k})_{k\in\mathbb{N}} \subseteq (M_n)_{n\in\mathbb{N}}$ mit

$$\mathcal{P}(M_{n_k}) > \mathcal{P}(M_{n_{k+1}}) > 0 \quad \text{für alle } k \in \mathbb{N},$$

sodass $\big(\mathcal{P}(M_{n_k})\big)_{k\in\mathbb{N}}$ eine Folge mit paarweise verschiedenen Folgengliedern darstellt, d.h. der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ ist essentiell unendlich.

„ $\Rightarrow$ “ Nun sei $(\Omega, \mathcal{M}, \mathcal{P})$ ein essentiell unendlicher Wahrscheinlichkeitsraum. Nach Definition (A.1.18) bedeutet dies, dass eine Mengenfolge $(D_n)_{n\in\mathbb{N}} \subseteq \mathcal{M}$ existiert, sodass die Glieder der Folge $\big(\mathcal{P}(D_n)\big)_{n\in\mathbb{N}}$ paarweise verschieden sind. Zusätzlich definieren wir $\mathcal{X} := \times_{i=1}^{\infty}\{0,1\}$, also die Menge aller Folgen, die nur die Werte 0 und 1 annehmen, welche offenbar abzählbar ist. Für jede Menge $M \in \mathcal{M}$ sei zudem

$$M^0 := M^c \in \mathcal{M} \qquad \text{und} \qquad M^1 := M \in \mathcal{M}.$$

Unter Verwendung dieser Bezeichnungen können wir dann zu jedem $\mu = (\mu_n)_{n\in\mathbb{N}} \in \mathcal{X}$ die Menge

$$N_\mu := \bigcap_{i=1}^{\infty} D_i^{\mu_i} \in \mathcal{M}$$

definieren. Damit kommen wir zur

1. Zwischenbehauptung: Es gilt:

(i) Für alle $\mu, \nu \in \mathcal{X}$ mit $\mu \neq \nu$ sind N_μ und N_ν disjunkt.
(ii) Für alle $n \in \mathbb{N}$ beobachten wir die Beziehung

$$D_n = \bigcup_{\substack{\mu\in\mathcal{X},\\ \mu_n=1}} N_\mu.$$

Dazu: (i) Seien $\mu, \nu \in \mathcal{X}$ mit $\mu \neq \nu$, d.h. wir finden ein $n \in \mathbb{N}$ mit $\mu_n \neq \nu_n$; Œ sei dabei $\mu_n = 0$ und $\nu_n = 1$. Damit ergibt sich

$$N_\mu = \bigcap_{i=1}^{\infty} D_i^{\mu_i} \subseteq D_n^c \qquad \text{und} \qquad N_\nu = \bigcap_{i=1}^{\infty} D_i^{\nu_i} \subseteq D_n,$$

das wiederum

$$N_\mu \cap N_\nu \subseteq D_n^c \cap D_n = \emptyset,$$

also $N_\mu \cap N_\nu = \emptyset$, liefert. Daher sind die Mengen N_μ und N_ν disjunkt.

(ii) „ $\subseteq$ “ Sei $x \in D_n$. Dann definieren wir die Folge $\nu = (\nu_i)_{i\in\mathbb{N}}$ durch

$$\nu_i = \begin{cases} 1, & \text{falls } x \in D_i, \\ 0, & \text{falls } x \in D_i^c, \end{cases}$$

also ist $\nu \in \mathcal{X}$. Damit erhalten wir offenbar $\nu_n = 1$ und

$$x \in \bigcap_{i=1}^{\infty} D_i^{\nu_i} = N_\nu \subseteq \bigcup_{\substack{\mu \in \mathcal{X}, \\ \mu_n = 1}} N_\mu,$$

womit

$$D_n \subseteq \bigcup_{\substack{\mu \in \mathcal{X}, \\ \mu_n = 1}} N_\mu$$

folgt.

„ $\supseteq$ “ Sei $\mu \in \mathcal{X}$ mit $\mu_n = 1$, so ergibt sich direkt

$$N_\mu = \bigcap_{i=1}^{\infty} D_i^{\mu_i} \subseteq D_n,$$

das

$$\bigcup_{\substack{\mu \in \mathcal{X}, \\ \mu_n = 1}} N_\mu \subseteq D_n$$

impliziert.
Damit ist die 1. Zwischenbehauptung bewiesen.

Zwar ist jetzt durchaus $\mathcal{P}(N_\mu) = 0$ für manche $\mu \in \mathcal{X}$ möglich, aber immerhin gilt die

2. Zwischenbehauptung: Es existieren unendlich viele, paarweise verschiedene $\mu \in \mathcal{X}$ mit $\mathcal{P}(N_\mu) > 0$.

Dazu: Wir nehmen das Gegenteil an, d.h. es existiere eine endliche Teilmenge $\mathcal{U} \subseteq \mathcal{X}$, etwa $|U| = N$ für ein $N \in \mathbb{N}$, sodass

$$\mathcal{P}(N_\mu) > 0 \quad \text{für alle } \mu \in \mathcal{U} \qquad \text{und} \qquad \mathcal{P}(N_\mu) = 0 \quad \text{für alle } \mathcal{X} \setminus \mathcal{U}$$

gilt. Jetzt definieren wir zu jeder Teilmenge $\mathcal{O} \subseteq \mathcal{X}$ die Menge

$$K_{\mathcal{O}} := \bigcup_{\mu \in \mathcal{O}} N_\mu,$$

wobei $K_{\mathcal{O}} \in \mathcal{M}$ ist, da mit $\mathcal{X}$ auch $\mathcal{O}$ abzählbar ist. Jedenfalls erhalten wir dann für jede Teilmenge $\mathcal{O} \subseteq \mathcal{X}$ die Identität

$$\mathcal{P}(K_{\mathcal{O}}) = \sum_{\mu \in \mathcal{O}} \mathcal{P}(N_\mu) = \sum_{\mu \in \mathcal{O} \cap \mathcal{U}} \mathcal{P}(N_\mu),$$

wobei wir für die erste Gleicheit die paarweise Disjunktheit der N_μ und für die zweite Gleichheit die Annahme ausgenutzt haben. Dies impliziert

$$|\{\mathcal{P}(K_{\mathcal{O}}) \mid \mathcal{O} \subseteq \mathcal{X}\}| \leq |\mathfrak{P}(\mathcal{U})| = 2^N, \tag{A.4}$$

wobei mit $\mathfrak{P}(\mathcal{U})$ die Potenzmenge von $\mathcal{U}$ bezeichnet sei. Jedoch gilt nach der 1. Zwischenbehauptung, (ii)

$$D_n = K_{\mathcal{O}_n} \qquad \text{mit} \qquad \mathcal{O}_n := \{\mu \in \mathcal{X} \mid \mu_n = 1\} \subseteq \mathcal{X}$$

für alle $n \in \mathbb{N}$ und $\big(\mathcal{P}(D_n)\big)_{n\in\mathbb{N}}$ ist nach Definition eine Folge mit paarweise verschiedenen Folgengliedern, d.h. es folgt

$$|\{\mathcal{P}(K_{\mathcal{O}}) \mid \mathcal{O} \subseteq \mathcal{X}\}| \geq |\{\mathcal{P}(D_n) \mid n \in \mathbb{N}\}| = \infty,$$

das aber der Ungleichung (A.4) widerspricht. Daher kann $\mathcal{U}$ nicht endlich sein und die 2. Zwischenbehauptung ist somit gezeigt.

Aufgrund der 2. Zwischenbehauptung finden wir also eine Folge $(\mu^{(n)})_{n\in\mathbb{N}} \subseteq \mathcal{X}$ mit paarweise verschiedenen Folgengliedern und $\mathcal{P}\big(N_{\mu^{(n)}}\big) > 0$ für alle $n \in \mathbb{N}$. Dann muss diese aber eine Teilfolge $(\mu^{(n_k)})_{k\in\mathbb{N}} \subseteq (\mu^{(n)})_{n\in\mathbb{N}}$ besitzen, sodass

$$\mathcal{P}\big(N_{\mu^{(n_k)}}\big) \xrightarrow[k\to\infty]{} 0$$

gilt; denn andernfalls existiert ein $\epsilon > 0$ und ein $N \in \mathbb{N}$ mit $\mathcal{P}\big(N_{\mu^{(n)}}\big) \geq \epsilon$ für alle $n \geq N$. Aufgrund der paarweisen Disjunktheit der $N_{\mu^{(n)}}$ ergibt sich dann

$$1 \geq \mathcal{P}\left(\bigcup_{i=1}^{n} N_{\mu^{(N+i)}}\right) = \sum_{i=1}^{n} \mathcal{P}\big(N_{\mu^{(N+i)}}\big) \geq n\epsilon$$

für alle $n \in \mathbb{N}$, das aber offensichtlich nicht möglich ist.
Daher muss eine solche Teilfolge $(\mu^{(n_k)})_{k\in\mathbb{N}}$ existieren. Mit $M_k := N_{\mu^{(n_k)}}$ für $k \in \mathbb{N}$ ist schließlich auch die Hinrichtung der Äquivalenz in diesem Satz gezeigt. □

Wegen des folgenden Lemmas müssen wir über essentiell endlichen Wahrscheinlichkeitsräumen nicht mehr zwischen den einzelnen L^p-Räumen unterscheiden.

(A.1.21) Lemma
Sei der Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ essentiell endlich. Dann sind alle L^p-Räume alegebraisch und topologisch identisch, d.h. für jedes $p, q \in (1, \infty)$ mit $p \leq q$ ist

$$L^\infty(\Omega) = L^q(\Omega) = L^p(\Omega) = L^1(\Omega)$$

und die einzelnen Normen sind äquivalent. ◇

Beweis
Aufgrund der Abschätzung (A.1) und der Inklusionskette (A.2) reicht es zu zeigen, dass eine Konstante $C > 0$ mit $\|X\|_\infty \leq C \cdot \|X\|_1$ für alle $X \in L^1(\Omega)$ existiert. Sei daher $X \in L^1(\Omega)$ und $R \in \mathbb{R}$ mit $R < \|X\|_\infty$ beliebig. Dann ergibt sich, dass $\mathcal{P}(\{|X| \geq R\}) > 0$ sein muss; denn die Annahme des Gegenteils liefert wegen

$$\|X\|_\infty = \operatorname{ess\,sup} |X| \leq \sup_{\omega\in\Omega\setminus\{|X|\geq R\}} |X(\omega)| \leq R < \|X\|_\infty$$

einen Widerspruch.
Da der Wahrscheinlichkeitsraum essentiell endlich ist, ergibt sich

$$\nu := \inf_{\substack{M \in \mathcal{M}, \\ \mathcal{P}(M) > 0}} \mathcal{P}(M) > 0$$

direkt aus der Definition (A.1.18). Daraus können wir schließen, dass $\mathcal{P}(\{|X| \geq R\}) \geq \nu$ und schließlich

$$\|X\|_1 = \int_\Omega |X(\omega)| \, \mathrm{d}\mathcal{P}(\omega) \geq \int_{\{|X| \geq R\}} |X(\omega)| \, \mathrm{d}\mathcal{P}(\omega) \geq \mathcal{P}(\{|X| \geq R\}) \cdot R \geq \nu \cdot R$$

ist. Äquivalent dazu ist $R \leq \frac{1}{\nu} \cdot \|X\|_1$ und, da R mit $R < \|X\|_\infty$ beliebig gewählt war, muss daher auch $\|X\|_\infty \leq \frac{1}{\nu} \cdot \|X\|_1$ gelten. Somit können wir $C := \frac{1}{\nu}$ wählen. □

(A.1.22) Bemerkung
Die Aussage aus Lemma (A.1.21) trifft über essentiell unendlichen Wahrscheinlichkeitsräumen i.A. nicht zu; dazu greifen wir das Beispiel (A.1.19) b) auf. Seien $p, q \in [1, \infty)$ mit $p < q$ und dazu definieren wir die Zufallsvariable

$$X_{p,q}(\omega) := \begin{cases} \frac{1}{\omega^{r(p,q)}} & \text{für } \omega \in (0,1], \\ 0 & \text{für } \omega = 0 \end{cases},$$

wobei $r(p,q) := \frac{p+q}{2pq}$ sei. Nach Definition ist $X_{p,q} \geq 0$ und auf $(0,1]$ stetig. Sei $a \in (0,1)$, so ergibt sich, dass

$$\begin{aligned} \left(\int_a^1 |X_{p,q}^p(\omega)| \, \mathrm{d}\omega \right)^{\frac{1}{p}} &= \left(\int_a^1 \frac{1}{\omega^{p \cdot r(p,q)}} \, \mathrm{d}\omega \right)^{\frac{1}{p}} = \left(\left[\frac{2q}{q-p} \cdot x^{\frac{q-p}{2q}} \right]_a^1 \right)^{\frac{1}{p}} \\ &= \left(\frac{2q}{q-p} \cdot \left[1 - a^{\frac{q-p}{2q}} \right] \right)^{\frac{1}{p}} \xrightarrow[a \downarrow 0]{} \left(\frac{2q}{q-p} \right)^{\frac{1}{p}}, \end{aligned}$$

da $\frac{q-p}{2q} > 0$ ist, und somit dass

$$\|X_{p,q}\|_p = \left(\frac{2q}{q-p} \right)^{\frac{1}{p}} < \infty$$

bzw. $X_{p,q} \in L^p(\Omega)$ ist. Hingegen erhält man mit einer analogen Rechnung

$$\left(\int_a^1 |X_{p,q}^q(\omega)| \, \mathrm{d}\omega \right)^{\frac{1}{q}} = \left(\frac{2p}{p-q} \cdot \left[1 - a^{\frac{p-q}{2p}} \right] \right)^{\frac{1}{q}} \xrightarrow[a \downarrow 0]{} \infty,$$

da $\frac{p-q}{2p} < 0$ ist, d.h. es ist $X_{p,q} \notin L^q(\Omega)$. Zudem ist offensichtlich $\|X_{p,q}\|_\infty = \infty$, also $X_{p,q} \notin L^\infty(\Omega)$. ◇

Nun entfernen wir uns etwas von Theorie der L^p-Räume und widmen uns zunächst gewissen „Zerlegungen" von Zufallsvariablen.

(A.1.23) Definition
Zu einer Zufallsvariablen $X : \Omega \to \mathbb{R}$ definiert man

$$X_+ : \Omega \to \mathbb{R}, \quad X_+(\omega) := max\{X(\omega), 0\}$$

und

$$X_- : \Omega \to \mathbb{R}, \quad X_-(\omega) := max\{-X(\omega), 0\}.$$

◇

Nachfolgend präsentieren wir einige zentrale Eigenschaften.

(A.1.24) Lemma
Sei $X : \Omega \to \mathbb{R}$ eine Zufallsvariable. Dann gilt:

(i) *X_+ und X_- sind ebenfalls Zufallsvariablen.*

(ii) *Es sind* $X = X_+ - X_-$ *und* $|X| = X_+ + X_-$.

(iii) *Für jede weitere Zufallsvariable $Y : \Omega \to \mathbb{R}$ ist*

$$0 \leq [X+Y]_+ \leq X_+ + Y_+ \quad \text{und} \quad 0 \leq [X+Y]_- \leq X_- + Y_-.$$

(iv) *Für alle $p \in [1, \infty]$ sind die Abbildungen*

$$L^p(\Omega) \to L^p(\Omega),\ Z \mapsto Z_+ \quad \text{und} \quad L^p(\Omega) \to L^p(\Omega),\ Z \mapsto Z_-$$

wohldefiniert und stetig. ◇

Beweis
(i) Dies ist ein elementares Ergebnis aus der Maß- und Integrationstheorie.

(ii) Dies folgt direkt aus der Definition (A.1.23).

(iii) Wir zeigen zunächst die erste Ungleichungskette: Da in dieser in die erste Ungleichung offensichtlich ist, bleibt nur die zweite zu zeigen. Sei dazu $\omega \in \Omega$.

1. Fall: $X(\omega) + Y(\omega) \geq 0$
Somit ist $X(\omega) \geq 0$ oder $Y(\omega) \geq 0$, Œ sei $X(\omega) \geq 0$. Falls $Y(\omega) \geq 0$ ist, erhalten wir

$$[X+Y]_+(\omega) = X(\omega) + Y(\omega) = X_+(\omega) + Y_+(\omega).$$

Andernfalls ist

$$[X+Y]_+(\omega) = X(\omega) + Y(\omega) < X(\omega) = X_+(\omega) + Y_+(\omega).$$

2. Fall: $X(\omega) + Y(\omega) < 0$
Somit ist
$$[X+Y]_+(\omega) = 0 \leq X_+(\omega) + Y_+(\omega).$$
Insgesamt ist die Ungleichung also gezeigt.

Für die zweite Ungleichungskette benutzt man die Beziehung $X_- = [-X]_+$ bzw. $Y_- = [-Y]_+$ und die schon gezeigte Ungleichung.

(iv) Wir zeigen die Wohldefiniertheit und Stetigkeit nur von der ersten Abbildung, da daraus wie in (iii) die Aussagen über die zweite Abbildung folgen. Seien also $p \in [1, \infty]$ und $X \in L^p(\Omega)$. Dann ist zunächst nach (ii)
$$0 \leq X_+ \leq X_+ + X_- = |X|$$
und schließlich folgt mit den Monotonie-Eigenschaften des Integrals bzw. des essentiellen Supremums aus Lemma (A.1.6) (vii)
$$0 \leq \|X_+\|_p \leq \||X|\|_p = \|X\|_p < \infty,$$
d.h. die Abbildung ist wohldefiniert.
Sei nun zusätzlich $(X_n)_{n\in\mathbb{N}} \subseteq L^p(\Omega)$ mit $X_n \xrightarrow[n\to\infty]{} X$ in $L^p(\Omega)$. Für beliebiges $\omega \in \Omega$ und $n \in \mathbb{N}$ erhalten wir mittels Fallunterscheidungen ähnlich wie in (iii) die Abschätzung
$$0 \leq |[X_n]_+(\omega) - X_+(\omega)| \leq |X_n(\omega) - X(\omega)|.$$
Mit der Monotonie liefert dies wie oben
$$0 \leq \|[X_n]_+ - X_+\|_p \leq \|X_n - X\|_p \xrightarrow[n\to\infty]{} 0,$$
das gerade der Stetigkeit der Abbildung entspricht. □

Zum Abschluss dieser Sektion stellen wir einen Satz vor, der es uns erlaubt, in einem vorliegenden Wahrscheinlichkeitsraum weitere Wahrscheinlichkeitsmaße zu konstruieren. Für den Beweis sei auf [8, Lemma (14.3)] verwiesen.

(A.1.25) Satz (Alternative Wahrscheinlichkeitsmaße)
Sei $X : \Omega \to \mathbb{R}$ eine Dichtefunktion, d.h. eine nicht-negative Zufallsvariable mit $EX = 1$. Dann definiert
$$\mathcal{P}' : \mathcal{M} \to [0,1], \quad \mathcal{P}'(M) := \int_M X(\omega)\,\mathrm{d}\mathcal{P}(\omega)$$
ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{M})$. Zudem ist für jede weitere Zufallsvariable $Y : \Omega \to \mathbb{R}$, für die XY $\mathcal{P}$-integrierbar ist, der Erwartungswert von Y bzgl. $\mathcal{P}'$ gegeben durch
$$E_{\mathcal{P}'}[Y] := \int_\Omega Y(\omega)\,\mathrm{d}\mathcal{P}'(\omega) = \int_\Omega X(\omega)Y(\omega)\,\mathrm{d}\mathcal{P}(\omega).$$
◇

A.2 Unter- und Oberhalbstetigkeit

In einem **metrischen Raum** $(\mathcal{X}, d)$ wollen wir für ein Funktional $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ abgeschwächte Arten von Stetigkeit einführen, da ein klassischer Stetigkeitsbegriff wegen der nicht vorausgesetzten Endlichkeit von $\mathcal{F}$ nicht ohne Weiteres definierbar ist. Dabei sollen diese Definitionen insofern schwächer sein, dass sie unter der Einschränkung, dass $\mathcal{F}$ endlich ist, also $\mathcal{F}(\mathcal{X}) \subseteq \mathbb{R}$ gilt, von der bekannten Stetigkeit impliziert werden.

(A.2.1) Definition (Unter- und Oberhalbstetigkeit)
Sei $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ ein Funktional und $x_0 \in \mathcal{X}$.

(i) *$\mathcal{F}$ heißt* unterhalbstetig in *x_0, falls für jede Folge $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{X}$ mit $x_n \xrightarrow[n\to\infty]{} x_0$ die Ungleichung*

$$\mathcal{F}(x_0) \leq \liminf_{n\to\infty} \mathcal{F}(x_n)$$

gilt. Darüber hinaus nennt man $\mathcal{F}$ unterhalbstetig, *falls $\mathcal{F}$ in jedem $x \in \mathcal{X}$ unterhalbstetig ist.*

(ii) *$\mathcal{F}$ heißt* oberhalbstetig in *x_0, falls für jede Folge $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{X}$ mit $x_n \xrightarrow[n\to\infty]{} x_0$ die Ungleichung*

$$\mathcal{F}(x_0) \geq \limsup_{n\to\infty} \mathcal{F}(x_n)$$

gilt. Darüber hinaus nennt man $\mathcal{F}$ oberhalbstetig, *falls $\mathcal{F}$ in jedem $x \in \mathcal{X}$ oberhalbstetig ist.* ◇

(A.2.2) Bemerkung

a) Aus der Definition (A.2.1) folgt direkt für endliches $\mathcal{F}$, also mit $\mathcal{F}(\mathcal{X}) \subseteq \mathbb{R}$, dass es genau dann stetig ist, falls es unter- **und** oberhalbstetig ist.

b) Aufgrund der Eigenschaften des Limes inferiors bzw. superiors können wir beobachten, dass $\mathcal{F}$ genau dann unterhalbstetig bzw. oberhalbstetig ist, falls $-\mathcal{F}$ oberhalbstetig bzw. unterhalbstetig ist. ◇

Zum praktischen Arbeiten mit unter- oder oberhalbstetigen Funktionen benötigen wir noch eine äußerst nützliche Charakterisierung.

(A.2.3) Lemma
Ein Funktional $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ ist genau dann unterhalbstetig bzw. oberhalbstetig, falls

$$\mathcal{V}_C(\mathcal{F}) := \{x \in \mathcal{X} \mid \mathcal{F}(x) \leq C\} \qquad \textit{bzw.} \qquad \mathcal{V}^C(\mathcal{F}) := \{x \in \mathcal{X} \mid \mathcal{F}(x) \geq C\}$$

für jedes $C \in \mathbb{R}$ abgeschlossen ist, wobei man $\mathcal{V}_C(\mathcal{F})$ als Subniveaumenge bzgl. *$\mathcal{F}$ und $\mathcal{V}^C(\mathcal{F})$ als* Superniveaumenge bzgl. *$\mathcal{F}$ bezeichnet.* ◇

Beweis
Wir zeigen die Äquivalenz nur für unterhalbstetige Funktionale, da sich die Charakterisierung für oberhalbstetige Funktionale daraus zusammen mit Bemerkung (A.2.2) b) ergibt.

„ $\Rightarrow$ “ Sei $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ unterhalbstetig und $C \in \mathbb{R}$. Für alle Folgen $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{V}_C(\mathcal{F})$ und $x \in \mathcal{X}$ mit $x_n \xrightarrow[n\to\infty]{} x$ folgt aus der Definition (A.2.1)

$$\mathcal{F}(x) \leq \liminf_{n\to\infty} \mathcal{F}(x_n) \leq \liminf_{n\to\infty} C = C,$$

also $x \in \mathcal{V}_C(\mathcal{F})$.

„ $\Leftarrow$ “ Sei $\mathcal{V}_C(\mathcal{F})$ für jedes $C \in \mathbb{R}$ abgeschlossen. Darüber hinaus sei $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{X}$ und $x \in \mathcal{X}$ mit $x_n \xrightarrow[n\to\infty]{} x$.

1. Fall: $\liminf\limits_{n\to\infty} \mathcal{F}(x_n) = \infty$
Dann erhalten wir die triviale Abschätzung

$$\mathcal{F}(x) \leq \infty = \liminf_{n\to\infty} \mathcal{F}(x_n).$$

2. Fall: $\liminf\limits_{n\to\infty} \mathcal{F}(x_n) < \infty$
Wir wählen ein $C \in \mathbb{R}$ mit $C > \liminf\limits_{n\to\infty} \mathcal{F}(x_n)$ beliebig. Dann existiert eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$, sodass $\mathcal{F}(x_{n_k}) < C$, also $x_{n_k} \in \mathcal{V}_C(\mathcal{F})$, für alle $k \in \mathbb{N}$ gilt. Als Teilfolge haben wir immer noch $x_{n_k} \xrightarrow[k\to\infty]{} x$. Aufgrund der Abgeschlossenheit von $\mathcal{V}_C(\mathcal{F})$ folgt somit auch $x \in \mathcal{V}_C(\mathcal{F})$, also $\mathcal{F}(x) \leq C$. Da C beliebig gewählt war, muss schließlich auch $\mathcal{F}(x) \leq \liminf\limits_{n\to\infty} \mathcal{F}(x_n)$ erfüllt sein. □

Zur Illustration untersuchen wir drei Familien von Funktionalen mit Bezug zu Abschnitt (A.1) auf (Oberhalb-/Unterhalb-)Stetigkeit.

(A.2.4) Lemma
Zu einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ und $p, q \in [1, \infty]$ sei das Funktional

$$\mathcal{G}_{p,q} : L^p(\Omega) \to [0, \infty], \quad \mathcal{G}_{p,q}(X) := \|X\|_q$$

gegeben. Dann ist $\mathcal{G}_{p,q}$

(i) endlich und stetig, falls $q \leq p$ ist, und

(ii) unterhalbstetig, falls $q > p$ ist.

Falls $(\Omega, \mathcal{M}, \mathcal{P})$ sogar essentiell endlich ist, ist $\mathcal{G}_{p,q}$ endlich und stetig für alle $p, q \in [1, \infty]$. Andernfalls ist $\mathcal{G}_{p,q}$ für $q > p$ nicht stetig. ◇

Beweis
(i) Dies folgt aus der Abschätzung (A.1).

(ii) Sei $(X_n)_{n\in\mathbb{N}} \subseteq L^p(\Omega)$ und $X \in L^p(\Omega)$ mit $X_n \xrightarrow[n\to\infty]{} X$ in $L^p(\Omega)$. Jetzt müssen wir zwei Fälle bzgl. q unterscheiden:

<u>1. Fall:</u> $q \in [1,\infty)$
Dann existiert eine Teilfolge $(X_{n_k})_{k\in\mathbb{N}} \subseteq (X_n)_{n\in\mathbb{N}}$, sodass

$$\mathcal{G}_{p,q}(X_{n_k}) \xrightarrow[k\to\infty]{} \liminf_{m\to\infty} \mathcal{G}_{p,q}(X_m) =: L \in [0,\infty] \tag{A.5}$$

erfüllt ist. Zudem existiert nach Bemerkung (A.1.15) eine weitere Teilfolge $(X_{n_{k_l}})_{l\in\mathbb{N}} \subseteq (X_{n_k})_{k\in\mathbb{N}}$ mit

$$X_{n_{k_l}} \xrightarrow[l\to\infty]{pktw.} X \quad \text{f.s..}$$

Daher erhalten wir auch

$$|X_{n_{k_l}}|^q \xrightarrow[l\to\infty]{pktw.} |X|^q \quad \text{f.s.}$$

und somit liefert des Lemma von Fatou

$$\begin{aligned}\mathcal{G}_{p,q}(X)^q = \int_\Omega |X|^q \,\mathrm{d}\mathcal{P} \le \liminf_{l\to\infty} \int_\Omega |X_{n_{k_l}}|^q \,\mathrm{d}\mathcal{P} &= \liminf_{l\to\infty} \mathcal{G}_{p,q}(X_{n_{k_l}})^q \\ &= \lim_{k\to\infty} \mathcal{G}_{p,q}(X_{n_k})^q \\ &= L^q,\end{aligned}$$

wobei wir im Falle von $L = \infty$ noch einmal explizit auf die Festlegungen aus Bemerkung (A.1.7) verweisen. Da $\mathcal{G}_{p,q} \ge 0$ und $L \ge 0$ ist, folgt

$$\mathcal{G}_{p,q}(X) \le L = \liminf_{m\to\infty} \mathcal{G}_{p,q}(X_m).$$

<u>2. Fall:</u> $q = \infty$
Hier wollen wir erstmals das Lemma (A.2.3) anwenden. Dazu sei $C \in \mathbb{R}$ beliebig und unsere Folge $(X_n)_{n\in\mathbb{N}} \subseteq L^p(\Omega)$ mit $X_n \xrightarrow[n\to\infty]{} X$ in $L^p(\Omega)$ erfülle zusätzlich die Bedingung, dass $\|X_n\|_\infty = \mathcal{G}_{p,\infty}(X_n) \le C$ für alle $n \in \mathbb{N}$ gilt. Anders als im 1. Fall wählen wir nun direkt eine Teilfolge $(X_{n_k})_{k\in\mathbb{N}} \subseteq (X_n)_{n\in\mathbb{N}}$ mit

$$X_{n_k} \xrightarrow[k\to\infty]{pktw.} X \quad \text{f.s..} \tag{A.6}$$

Nach Bemerkung (A.1.9) existiert dann eine Nullmenge $M \in \mathcal{M}$ mit $|X_{n_k}(\omega)| \le C$ für alle $\omega \in \Omega \setminus M$. Zusammen mit der Konvergenz in (A.6) finden wir eine weitere Nullmenge $N \in \mathcal{M}$ mit $M \subseteq N$, sodass

$$|X_{n_k}(\omega)| \xrightarrow[k\to\infty]{} |X(\omega)| \quad \text{und} \quad |X_{n_k}(\omega)| \le C \quad \text{für alle } \omega \in \Omega \setminus N \text{ und } k \in \mathbb{N}$$

zutrifft. Dies impliziert aber gerade $|X(\omega)| \leq C$ für alle $\omega \in \Omega \setminus N$ und schließlich

$$\mathcal{G}_{p,\infty}(X) = \|X\|_\infty \leq \sup_{\omega \in \Omega \setminus N} |X(\omega)| \leq C.$$

Da $C \in \mathbb{R}$ beliebig gewählt war, ist $\mathcal{V}_C(\mathcal{G}_{p,\infty})$ für jedes $C \in \mathbb{R}$ abgeschlossen, was nach Lemma (A.2.3) die Unterhalbstetigkeit von $\mathcal{G}_{p,\infty}$ liefert.

Jetzt widmen wir uns den Zusätzen: Sei $(\Omega, \mathcal{M}, \mathcal{P})$ zunächst essentiell endlich, sodass nach Lemma (A.1.21) nicht nur $L^p(\Omega) = L^q(\Omega)$ gilt, sondern auch die zugehörigen Normen äquivalent sind. Daraus leitet sich sofort die Endlichkeit und Stetigkeit von $\mathcal{G}_{p,q}$ ab.
Wenn $(\Omega, \mathcal{M}, \mathcal{P})$ jedoch essentiell unendlich ist, existiert nach Satz (A.1.20) eine Mengenfolge $(M_n)_{n \in \mathbb{N}} \subseteq \mathcal{M}$ mit $\mathcal{P}(M_n) \xrightarrow[n \to \infty]{} 0$ und $\mathcal{P}(M_n) > 0$ für alle $n \in \mathbb{N}$. Zu $q > p$, also insbesondere $p \in [1, \infty)$, sei

$$r(p,q) := \begin{cases} \frac{p+q}{2pq} & \text{für } q \in [1, \infty), \\ \frac{1}{2p} & \text{für } q = \infty, \end{cases}$$

und für $n \in \mathbb{N}$ betrachten wir dann die Zufallsvariablen

$$X_n^{(p,q)} : \Omega \to \mathbb{R}, \quad X_n^{(p,q)}(\omega) := \frac{1}{\mathcal{P}(M_n)^{r(p,q)}} \cdot \mathbb{1}_{M_n}(\omega), \tag{A.7}$$

wobei wir mit $\mathbb{1}_{M_n}$ die charakteristische Funktion der Menge M_n bezeichnen.

1. Fall: $q \in [1, \infty)$
Nach Konstruktion erhalten wir

$$\|X_n^{(p,q)}\|_p = \left(\mathcal{P}(M_n) \cdot \frac{1}{\mathcal{P}(M_n)^{p \cdot r(p,q)}} \right)^{\frac{1}{p}} = \mathcal{P}(M_n)^{\frac{q-p}{2pq}} \xrightarrow[n \to \infty]{} 0,$$

da $\frac{q-p}{2pq} > 0$ ist, aber

$$\mathcal{G}_{p,q}(X_n^{(p,q)}) = \|X_n^{(p,q)}\|_q = \left(\mathcal{P}(M_n) \cdot \frac{1}{\mathcal{P}(M_n)^{q \cdot r(p,q)}} \right)^{\frac{1}{q}} = \mathcal{P}(M_n)^{\frac{p-q}{2pq}} \xrightarrow[n \to \infty]{} \infty,$$

da $\frac{p-q}{2pq} < 0$ ist.

2. Fall: $q = \infty$
So ergibt sich analog $\|X_n^{(p,q)}\|_p \xrightarrow[n \to \infty]{} 0$, aber

$$\mathcal{G}_{p,\infty}(X_n^{(p,\infty)}) = \|X_n^{(p,\infty)}\|_\infty = \frac{1}{\mathcal{P}(M_n)^{\frac{1}{2p}}} \xrightarrow[n \to \infty]{} \infty. \tag{A.8}$$

In beiden Fällen ist $\mathcal{G}_{p,q}$ demnach nicht stetig. □

(A.2.5) Lemma
Zu einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{M}, \mathcal{P})$ und $p \in [1, \infty]$ seien die Funktionale

$$\mathcal{H}^p : L^p(\Omega) \to (-\infty, \infty], \quad \mathcal{H}^p(X) := \textit{ess sup } X$$

und

$$\mathcal{H}_p : L^p(\Omega) \to [-\infty, \infty), \quad \mathcal{H}_p(X) := \textit{ess inf } X$$

gegeben. Dann ist $\mathcal{H}^p$ bzw. $\mathcal{H}_p$

(i) endlich und stetig, falls $p = \infty$ ist, und

(ii) unterhalbstetig bzw. oberhalbstetig, falls $p < \infty$ ist.

Falls $(\Omega, \mathcal{M}, \mathcal{P})$ sogar essentiell endlich ist, sind $\mathcal{H}^p$ und $\mathcal{H}_p$ endlich und stetig für alle $p \in [1, \infty]$. Andernfalls sind sie für $p < \infty$ nicht stetig. ◇

Beweis
Wir zeigen die jeweiligen Aussagen nur für $\mathcal{H}^p$; denn für $\mathcal{H}_p$ folgen die entsprechenden Behauptungen aus Lemma (A.1.6) (ii) und der Bemerkung (A.2.2) b).

(i) Sei $X \in L^\infty(\Omega)$. Dann ist wegen $-|X| \leq X \leq |X|$ nach Lemma (A.1.6) (ii), (vii) und (viii)

$$\underbrace{-\text{ ess inf } |X|}_{\leq 0} = \text{ess sup } (-|X|) \leq \text{ess sup } X \leq \underbrace{\text{ess sup } |X|}_{\geq 0}$$

und daher

$$\begin{aligned} |\mathcal{H}^\infty(X)| = |\text{ess sup } X| &\leq \max\{\big|-\text{ess inf } |X|\big|, \text{ ess sup } |X|\} \\ &= \max\{\underbrace{\text{ess inf } |X|}_{\leq \text{ ess sup } |X|}, \text{ ess sup } |X|\} \\ &= \text{ess sup } |X| = \|X\|_\infty < \infty, \end{aligned} \tag{A.9}$$

also ist $\mathcal{H}^\infty$ endlich. Sei nun zusätzlich $(X_n)_{n\in\mathbb{N}} \subseteq L^\infty(\Omega)$ mit $X_n \xrightarrow[n\to\infty]{} X$ in $L^\infty(\Omega)$. Mit Lemma (A.1.6) (v) erhalten wir die Ungleichungen

$$\text{ess sup } X \leq \text{ess sup } X_n + \text{ess sup } (X - X_n)$$

und

$$\text{ess sup } X_n \leq \text{ess sup } X + \text{ess sup } (X_n - X),$$

sodass

$$\begin{aligned} |\mathcal{H}^\infty(X) - \mathcal{H}^\infty(X_n)| &= \big|\text{ess sup } X - \text{ess sup } X_n\big| \\ &\leq \max\{\text{ess sup } (X - X_n), \text{ ess sup } (X_n - X)\} \\ &\leq \text{ess sup } |X - X_n| = \|X - X_n\|_\infty \xrightarrow[n\to\infty]{} 0 \end{aligned} \tag{A.10}$$

folgt. Daher ist $\mathcal{H}^\infty$ auch stetig.

(ii) Dies wird völlig analog wie im Beweis von Lemma (A.2.4) (ii), 2. Fall, bewiesen.

Schließlich kommen wir zu den Zusätzen: Sei $(\Omega, \mathcal{M}, \mathcal{P})$ zunächst essentiell endlich, sodass nach Lemma (A.1.21) die Räume $L^p(\Omega)$ und $L^\infty(\Omega)$ sowohl algebraisch als auch topologisch übereinstimmen. Wenn wir diese Aussage mit den hergeleiteten Ungleichungen in (A.9) und (A.10) verknüpfen, ergibt sich die Endlichkeit und Stetigkeit von $\mathcal{H}^p$.
Falls $(\Omega, \mathcal{M}, \mathcal{P})$ aber essentiell unendlich ist, so betrachten wir für $p < \infty$ in der Situation vom Beweis von Lemma (A.2.4) die in (A.7) definierte Folge $(X_n^{(p,\infty)})_{n\in\mathbb{N}} \subseteq L^p(\Omega)$, die bekanntlich $X_n^{(p,\infty)} \xrightarrow[n\to\infty]{} 0$ in $L^p(\Omega)$ erfüllt. Da $X_n^{(p,\infty)} \geq 0$ für alle $n \in \mathbb{N}$ ist, folgt zusammen mit der Divergenz in (A.8)

$$\mathcal{H}^p(X_n^{(p,\infty)}) = \mathcal{G}_{p,\infty}(X_n^{(p,\infty)}) \xrightarrow[n\to\infty]{} \infty.$$

Somit kann $\mathcal{H}^p$ nicht stetig sein. □

Darüber hinaus gibt es eine weitere Charakterisierung von Unterhalbstetigkeit, die aber eher von theoretischer Natur ist. Dazu benötigen wir aber auch eine neue Bezeichnung.

(A.2.6) Definition (Der Epigraph)
Zu einem Funktional $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ heißt die Menge

$$epi\ \mathcal{F} := \{(x,z) \in \mathcal{X} \times \mathbb{R} \mid \mathcal{F}(x) \leq z\}$$

der Epigraph von *$\mathcal{F}$.* ◇

(A.2.7) Bemerkung

a) Das Wort „epi“ stammt aus dem Griechischen und bedeutet „über“. Daher ist die Bezeichnung auch sinnvoll; denn wir können uns z.B. den Epigraphen einer Funktion $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ als die Fläche oberhalb des Graphen von $\mathcal{F}$ im zweidimensionalen Koordinatensystem, wobei der Graph selbst mit eingeschlossen ist, vorstellen.

b) Die folgende Abbildung

$$d_{\mathcal{X}\times\mathbb{R}} : (\mathcal{X} \times \mathbb{R}) \times (\mathcal{X} \times \mathbb{R}) \to \mathbb{R}, \quad d_{\mathcal{X}\times\mathbb{R}}\big((x,z),(y,w)\big) := d(x,y) + |z-w|$$

definiert eine Metrik auf $\mathcal{X} \times \mathbb{R}$, sodass $(\mathcal{X} \times \mathbb{R}, d_{\mathcal{X}\times\mathbb{R}})$ ein metrischer Raum ist.
◇

(A.2.8) Lemma
Ein Funktional $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ ist genau dann unterhalbstetig, wenn epi $\mathcal{F}$ abgeschlosen ist. ◇

Beweis
„ $\Rightarrow$ “ Sei $\mathcal{F}$ unterhalbstetig. Für jede Folge $\big((x_n, z_n)\big)_{n\in\mathbb{N}} \subseteq \operatorname{epi} \mathcal{F}$ und $(x, z) \in \mathcal{X} \times \mathbb{R}$ mit $(x_n, z_n) \xrightarrow[n\to\infty]{} (x, z)$ in $\mathcal{X} \times \mathbb{R}$ gilt insbesondere $x_n \xrightarrow[n\to\infty]{} x$ in $\mathcal{X}$ und $z_n \xrightarrow[n\to\infty]{} z$ in $\mathbb{R}$. Da $\mathcal{F}(x_n) \leq z_n$ für alle $n \in \mathbb{N}$ ist, ergibt sich aufgrund der Unterhalbstetigkeit von $\mathcal{F}$ die Ungleichung

$$\mathcal{F}(x) \leq \liminf_{n\to\infty} \mathcal{F}(x_n) \leq \liminf_{n\to\infty} z_n = z$$

und damit $(x, z) \in \operatorname{epi} \mathcal{F}$.

„ $\Leftarrow$ “ Sei $\operatorname{epi} \mathcal{F}$ abgeschlossen und $C \in \mathbb{R}$. Darüber hinaus seien $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{V}_C(\mathcal{F})$ aus Lemma (A.2.3) und $x \in \mathcal{X}$ mit $x_n \xrightarrow[n\to\infty]{} x$. Anders formuliert bedeutet dies gerade $(x_n, C) \in \operatorname{epi} \mathcal{F}$ für alle $n \in \mathbb{N}$. Da offensichtlich $(x_n, C) \xrightarrow[n\to\infty]{} (x, C)$ erfüllt ist, folgt $(x, C) \in \operatorname{epi} \mathcal{F}$ wegen der Abgeschlossenheit von $\operatorname{epi} \mathcal{F}$ und somit $\mathcal{F}(x) \leq C$ bzw. $x \in \mathcal{V}_C(\mathcal{F})$. Nach Lemma (A.2.3) ist $\mathcal{F}$ also unterhalbstetig. □

Natürlich können wir unter- bzw. oberhalbstetige Funktionale so kombinieren, dass wir wiederum solche erhalten. Dabei beschränken wir uns aber auf die für uns relevanten Verknüpfungen und benötigen dafür erneut unsere Festlegungen bzgl. des Umgangs mit $\pm\infty$ in Bemerkung (A.1.7).

(A.2.9) Lemma
Seien die Funktionale $\mathcal{F}_1, \ldots, \mathcal{F}_n : \mathcal{X} \to [-\infty, \infty]$ gegeben. Dann gilt:

(i) Falls alle $\mathcal{F}_1, \ldots, \mathcal{F}_n$ unter- bzw. oberhalbstetig sind, so ist auch

$$\sum_{i=1}^{n} \lambda_i \mathcal{F}_i : \mathcal{X} \to [-\infty, \infty]$$

für $\lambda_1, \ldots, \lambda_n > 0$ unter- bzw. oberhalbstetig.

(ii) Falls alle $\mathcal{F}_1, \ldots, \mathcal{F}_n$ unterhalbstetig sind, so ist auch

$$max\{\mathcal{F}_1, \ldots, \mathcal{F}_n\} : \mathcal{X} \to [-\infty, \infty]$$

unterhalbstetig und, falls alle $\mathcal{F}_1, \ldots, \mathcal{F}_n$ oberhalbstetig sind, so ist auch

$$min\{\mathcal{F}_1, \ldots, \mathcal{F}_n\} : \mathcal{X} \to [-\infty, \infty]$$

oberhalbstetig. ◇

Beweis
(i) Wir zeigen die Aussage nur für den Fall, dass alle $\mathcal{F}_1, \ldots, \mathcal{F}_n$ unterhalbstetig sind, da der Nachweis im anderen Fall analog verläuft.

Seien $x \in \mathcal{X}$ und $(x_k)_{k\in\mathbb{N}} \subseteq \mathcal{X}$ mit $x_k \xrightarrow[k\to\infty]{} x$. Dann gilt

$$\begin{aligned}\sum_{i=1}^n \lambda_i \mathcal{F}_i(x) \leq \sum_{i=1}^n \lambda_i \cdot \liminf_{k\to\infty} \mathcal{F}_i(x_k) &= \sum_{i=1}^n \liminf_{k\to\infty} \big(\lambda_i \mathcal{F}_i(x_k)\big) \\ &\leq \liminf_{k\to\infty} \left(\sum_{i=1}^n \lambda_i \mathcal{F}_i(x_k)\right)\end{aligned}$$

aufgrund der Unterhalbstetigkeit von $\mathcal{F}_1, \ldots, \mathcal{F}_n$ und den Eigenschaften des Limes inferiors. Daher ist ebenfalls $\sum_{i=1}^n \lambda_i \mathcal{F}_i$ unterhalbstetig.

(ii) Zusammen mit Bemerkung (A.2.2) b) genügt es erneut, nur die erste Behauptung zu beweisen.
Seien $C \in \mathbb{R}$, $(x_k)_{k\in\mathbb{N}} \subseteq \mathcal{V}_C(\max\{\mathcal{F}_1, \ldots, \mathcal{F}_n\})$ aus Lemma (A.2.3) und $x \in \mathcal{X}$ mit $x_k \xrightarrow[k\to\infty]{} x$. Also ist

$$\mathcal{F}_i(x_k) \leq \max\{\mathcal{F}_1(x_k), \ldots, \mathcal{F}_n(x_k)\} \leq C$$

für alle $i \in \{1, \ldots, n\}$ und $k \in \mathbb{N}$. Nach Lemma (A.2.3) ist somit $\mathcal{F}_i(x) \leq C$ für alle $i \in \{1, \ldots, n\}$ und daraus folgt

$$\max\{\mathcal{F}_1(x), \ldots, \mathcal{F}_n(x)\} \leq C$$

bzw. $x \in \mathcal{V}_C(\max\{\mathcal{F}_1, \ldots, \mathcal{F}_n\})$, sodass Lemma (A.2.3) die Unterhalbstetigkeit von $\max\{\mathcal{F}_1, \ldots, \mathcal{F}_n\}$ liefert. $\square$

Abschließend präsentieren wir eine hinreichende Bedingung für Stetigkeit, also nach Bemerkung (A.2.2) a) insbesondere für Unter- und Oberhalbstetigkeit. Vorher müssen wir aber eine in der Literatur gebräuchliche Bezeichnung einführen.

(A.2.10) Definition
Zu einem Funktional $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ heißt die Menge

$$dom\ \mathcal{F} := \{x \in \mathcal{X} \mid \mathcal{F}(x) < \infty\}$$

die Domäne von $\mathcal{F}$. $\diamond$

Zudem müssen wir den Konvexitätsbegriff verallgemeinern.

(A.2.11) Definition
Sei $(\mathcal{X}, d)$ sogar ein metrischer linearer Raum. Dann wird ein Funktional $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ als konvex *bezeichnet, falls für alle $x, y \in dom(\mathcal{F})$ und $\lambda \in [0, 1]$ die Ungleichung*

$$\mathcal{F}\big(\lambda x + (1-\lambda)y\big) \leq \lambda \mathcal{F}(x) + (1-\lambda)\mathcal{F}(y)\ (< \infty)$$

gilt, also insbesondere $dom(\mathcal{F})$ konvex ist. $\diamond$

(A.2.12) Bemerkung
Aufgrund von Bemerkung (A.1.7) gilt für ein konvexes Funktional $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ sogar

$$\mathcal{F}\big(\lambda x + (1-\lambda)y\big) \leq \lambda \mathcal{F}(x) + (1-\lambda)\mathcal{F}(y)$$

für alle $x, y \in \mathcal{X}$ und $\lambda \in [0,1]$. ◇

Nun kommen wir zu dem angekündigten Satz, welcher in einer allgemeineren Variante in [3, Satz (5.20)] zu finden ist.

(A.2.13) Satz
Sei $(\mathcal{X}, d)$ sogar ein metrischer linearer Raum. Zudem sei $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ ein konvexes Funktional und $x_0 \in \mathcal{X}$. Falls eine Umgebung $U \subseteq \mathcal{X}$ von x_0 und eine Konstante $C > 0$ existiert, sodass $\mathcal{F}(x) \leq C$ für alle $x \in U$ ist, dann ist $\mathcal{F}$ stetig in x_0. ◇

Beweis
Œ sei $x_0 = 0$ sowie $\mathcal{F}(0) = 0$; denn andernfalls betrachten wir das modifizierte Funktional

$$\mathcal{G} : X \to (-\infty, \infty], \quad \mathcal{G}(x) := \mathcal{F}(x + x_0) - \mathcal{F}(x_0).$$

Offensichtlich ist $\mathcal{G}(0) = 0$, $\mathcal{G}$ konvex und

$$\mathcal{G}(x) \leq C - \mathcal{F}(x_0) \leq \max\{1,\ C - \mathcal{F}(x_0)\} =: K$$

für alle $x \in V := U - x_0$. Dabei ist nach Definition $K > 0$. Darüber hinaus ist $\mathcal{F}$ genau dann in x_0 stetig, wenn $\mathcal{G}$ in Null stetig ist.

Also können wir uns jetzt dem eigentlichen Beweis unter Annahme $x_0 = 0$ und $\mathcal{F}(0) = 0$ widmen: Sei $\epsilon > 0$ und $\epsilon_0 := \min\{\frac{\epsilon}{C}, \frac{1}{2}\} \in (0,1)$. Da U eine Umgebung von $x_0 = 0$ ist, ist aus der Funktionalanalysis bekannt, dass eine kreisförmige Umgebung O von Null mit $O \subseteq U$ existiert. Daher ist $\epsilon_0 O \subseteq O$ eine weitere Null-Umgebung und für jedes $x \in \epsilon_0 O$ können wir

$$x = (1 - \epsilon_0) \cdot 0 + \epsilon_0 \cdot \frac{x}{\epsilon_0}$$

schreiben. Dann ist

$$\mathcal{F}(x) \leq (1 - \epsilon_0) \cdot \underbrace{\mathcal{F}(0)}_{=0} + \epsilon_0 \cdot \mathcal{F}\left(\frac{x}{\epsilon_0}\right) \leq \epsilon_0 C, \tag{A.11}$$

da $\mathcal{F}$ konvex und $\frac{x}{\epsilon_0} \in O \subseteq U$ ist. Andererseits ist die rechte Seite von

$$0 = \frac{1}{1+\epsilon_0} \cdot x + \frac{\epsilon_0}{1+\epsilon_0} \cdot \left(-\frac{x}{\epsilon_0}\right)$$

ebenfalls eine Konvexkombination. Da O kreisförmig ist, ist auch $-\frac{x}{\epsilon_0} \in O \subseteq U$, sodass erneut u.A. mit der Konvexität von $\mathcal{F}$

$$\begin{aligned} 0 = \mathcal{F}(0) \leq \frac{1}{1+\epsilon_0} \cdot \mathcal{F}(x) + \frac{\epsilon_0}{1+\epsilon_0} \cdot \mathcal{F}\left(-\frac{x}{\epsilon_0}\right) &\leq \frac{1}{1+\epsilon_0} \cdot \mathcal{F}(x) + \frac{\epsilon_0 C}{1+\epsilon_0} \\ &= \frac{1}{1+\epsilon_0} \cdot \left(\mathcal{F}(x) + \epsilon_0 C\right) \end{aligned}$$

folgt. Diese Ungleichung ist äquivalent zu $\mathcal{F}(x) \geq -\epsilon_0 C$ und zusammen mit der Ungleichung in (A.11) gilt die Abschätzung

$$|\mathcal{F}(x)| \leq \epsilon_0 C \leq \frac{\epsilon}{C} C = \epsilon$$

für alle $x \in \epsilon_0 O$. Schließlich wählen wir ein geeignetes $\delta > 0$ mit $B_\delta(0) \subseteq \epsilon_0 O$ und somit ist

$$|\mathcal{F}(0) - \mathcal{F}(x)| = |\mathcal{F}(x)| \leq \epsilon$$

für alle $x \in B_\delta(0)$ erfüllt, d.h. $\mathcal{F}$ ist stetig in $x_0 = 0$. □

Dieser Satz wird äußerst hilfreich für den Beweis von Satz (1.1.7) sein, in dem wir eine hinreichende Bedingung dafür vorstellen, wann eine unterhalbstetige Funktion sogar stetig ist.

A.3 Die Fenchel-Konjugierte

Als wichtiger Baustein für die zentralen Dualitätssätze dieser Arbeit beschäftigen wir uns in dieser Sektion mit der „Fenchel-Konjugierten". Da diese Theorie lediglich als Hilfsmittel dient, werden wir sie nur in Grundzügen und unter vereinfachten Bedingungen, d.h. über einem **Prä-Hilbertraum** $(\mathcal{X}, \langle\cdot,\cdot\rangle)$, herleiten. Die wesentlichen Ideen stammen dabei aus [3, Abschnitt 6]. An dieser Stelle erinnern wir noch einmal an unsere Festlegungen bzgl. des Umgangs mit $\pm\infty$ in Bemerkung (A.1.7), um die Wohldefiniertheit der anschließenden, allseits bekannten Definition zu garantieren.

(A.3.1) Definition (Die Fenchel-Konjugierte)
Sei $\mathcal{F} : \mathcal{X} \to [-\infty, \infty]$ ein Funktional. Dann nennt man das Funktional

$$\mathcal{F}^* : \mathcal{X} \to [-\infty, \infty], \quad \mathcal{F}^*(x) := \sup_{y \in \mathcal{X}} \left\{\langle x, y\rangle - \mathcal{F}(y)\right\}$$

die Fenchel-Konjugierte zu $\mathcal{F}$. *Des Weiteren bezeichnet man das Funktional $\mathcal{F}^{**} := (\mathcal{F}^*)^*$ als die* Bikonjugierte zu $\mathcal{F}$. ◇

Bevor wir die Fenchel-Konjugierte auf ihre Eigenschaften untersuchen, stellen wir eine bekannte Charakterisierung von Konvexität vor, die wir später benötigen.

(A.3.2) Lemma
Sei $(\mathcal{Y}, d)$ ein metrischer linearer Raum. Ein Funktional $\mathcal{F} : \mathcal{Y} \to (-\infty, \infty]$ ist genau dann konvex, falls epi $\mathcal{F} \subseteq \mathcal{Y} \times \mathbb{R}$ konvex ist. $\diamond$

Beweis
Seien $\lambda \in [0,1]$ und $x, y \in \mathrm{dom}(\mathcal{F})$. Dann ist die Ungleichung

$$\mathcal{F}\big(\lambda x + (1-\lambda)y\big) \leq \lambda \mathcal{F}(x) + (1-\lambda)\mathcal{F}(y) \tag{A.12}$$

äquivalent zu

$$\big(\lambda x + (1-\lambda)y, \lambda \mathcal{F}(x) + (1-\lambda)\mathcal{F}(y)\big) \in \text{epi } \mathcal{F}.$$

Nun seien in dieser Situation $z, w \in \mathbb{R}$ mit $(x,z), (y,w) \in$ epi $\mathcal{F}$, also $\mathcal{F}(x) \leq z$ und $\mathcal{F}(y) \leq w$, beliebig, d.h. insbesondere sind die Wahlen $z = \mathcal{F}(x)$ und $w = \mathcal{F}(y)$ möglich. Jedenfalls gilt somit

$$\mathcal{F}\big(\lambda x + (1-\lambda)y\big) \leq \lambda \mathcal{F}(x) + (1-\lambda)\mathcal{F}(y) \leq \lambda z + (1-\lambda)w,$$

das wiederum nichts anderes als

$$\lambda(x,z) + (1-\lambda)\big(y,w\big) = \big(\lambda x + (1-\lambda)y, \lambda z + (1-\lambda)w\big) \in \text{epi } \mathcal{F}$$

bedeutet. Daher trifft die Ungleichung (A.12) genau dann für alle $\lambda \in [0,1]$ und $x, y \in \mathrm{dom}(\mathcal{F})$ zu, wenn epi $\mathcal{F}$ konvex ist, sodass die Behauptung aufgrund der Definition (A.2.11) gezeigt ist. $\square$

Wie angekündigt, erarbeiten wir jetzt die wesentlichen Aussagen über die Fenchel- und Bikonjugierte. Als Erstes präsentieren wir eine spezielle Variante von [3, Lemma (6.4)].

(A.3.3) Lemma
Sei ein Funktional $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ mit dom $\mathcal{F} \neq \emptyset$ gegeben. Dann gilt für die Fenchel-Konjugierte $\mathcal{F}^*$, dass $\mathcal{F}^* > -\infty$ sowie $\mathcal{F}^*$ konvex und unterhalbstetig ist. $\diamond$

Beweis
Zuallererst finden wir wegen dom $\mathcal{F} \neq \emptyset$ ein $z \in \mathcal{X}$ mit $\mathcal{F}(z) < \infty$, sodass

$$\mathcal{F}^*(x) = \sup_{y \in \mathcal{X}} \{\langle x, y\rangle - \mathcal{F}(y)\} \geq \langle x, z\rangle - \mathcal{F}(z) > -\infty$$

für alle $x \in \mathcal{X}$ gilt, d.h. es ist $\mathcal{F}^* > -\infty$. Damit können wir nun die weiteren Behauptungen zeigen:

Konvexität: Seien $x, y \in \operatorname{dom}(\mathcal{F}^*)$ sowie $\lambda \in [0,1]$. So erhalten wir

$$\begin{aligned}
&\mathcal{F}^*\big(\lambda x + (1-\lambda)y\big) \\
&= \sup_{a \in \mathcal{X}} \left\{ \langle \lambda x + (1-\lambda)y, a\rangle - \mathcal{F}(a) \right\} \\
&= \sup_{a \in \mathcal{X}} \left\{ \big[\lambda \langle x, a\rangle - \lambda \mathcal{F}(a)\big] + \big[(1-\lambda)\langle y, a\rangle - (1-\lambda)\mathcal{F}(a)\big] \right\} \\
&\leq \sup_{a \in \mathcal{X}} \left\{ \big[\lambda \langle x, a\rangle - \lambda \mathcal{F}(a)\big] \right\} + \sup_{a \in \mathcal{X}} \left\{ \big[(1-\lambda)\langle y, a\rangle - (1-\lambda)\mathcal{F}(a)\big] \right\} \\
&= \lambda \cdot \sup_{a \in \mathcal{X}} \left\{ \big[\langle x, a\rangle - \mathcal{F}(a)\big] \right\} + (1-\lambda) \cdot \sup_{a \in \mathcal{X}} \left\{ \big[\langle y, a\rangle - \mathcal{F}(a)\big] \right\} \\
&= \lambda \mathcal{F}^*(x) + (1-\lambda)\mathcal{F}^*(y),
\end{aligned}$$

d.h. wegen der beliebigen Wahl von x, y und λ folgt die Konvexität von $\mathcal{F}^*$.

Unterhalbstetigkeit: Seien $C \in \mathbb{R}$, $(x_n)_{n \in \mathbb{N}} \subseteq \mathcal{V}_C(\mathcal{F}^*)$ aus Lemma (A.2.3) und $x \in \mathcal{X}$ mit $x_n \xrightarrow[n \to \infty]{} x$. Nach Annahme gilt daher

$$\langle x_n, y\rangle - \mathcal{F}(y) \leq \mathcal{F}^*(x_n) \leq C \tag{A.13}$$

für alle $n \in \mathbb{N}$ und $y \in \mathcal{X}$. Da die Cauchy-Schwarz-Ungleichung nun

$$\langle x_n, y\rangle \xrightarrow[n \to \infty]{} \langle x, y\rangle$$

für alle $y \in \mathcal{X}$ liefert, folgt zusammen mit der Abschätzung (A.13), dass

$$\langle x, y\rangle - \mathcal{F}(y) \leq C$$

für alle $y \in \mathcal{X}$ ist. Damit ist auch $\mathcal{F}^*(x) \leq C$, also $x \in \mathcal{V}_C(\mathcal{F}^*)$, sodass sich mit Lemma (A.2.3) die Unterhalbstetigkeit von $\mathcal{F}^*$ ergibt. □

Schon kommen wir zu dem für uns wichtigsten Ergebnis in dieser Theorie, das im Wesentlichen auf [3, Folgerung (6.8)] sowie einigen vorhergehenden Resultaten beruht.

(A.3.4) Satz
Sei $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ sogar ein Hilbertraum. Zudem sei $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ ein konvexes und unterhalbstetiges Funktional mit $\operatorname{dom} \mathcal{F} \neq \emptyset$. Dann gilt:

(i) Es ist $\mathcal{F}^{**} = \mathcal{F}$.

(ii) Die Fenchel-Konjugierte $\mathcal{F}^*$ besitzt die Eigenschaft, dass $\mathcal{F}^* > -\infty$ und $\operatorname{dom} \mathcal{F}^*$ nicht-leer ist. Darüber hinaus ist sie konvex und unterhalbstetig. ◇

Als Hilfsaussage für den Beweis dieses Satzes benutzen wir ein Resultat aus der konvexen Analysis, das unter schwächeren Bedingungen in [3, Satz (5.24)] bewiesen wurde. Die Beweisidee haben wir aber von dort übernommen.

(A.3.5) Lemma
Sei $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ sogar ein Hilbertraum. Zudem sei $\mathcal{F} : \mathcal{X} \to (-\infty, \infty]$ ein konvexes und unterhalbstetiges Funktional und es gelte dom $\mathcal{F} \neq \emptyset$. Dann ist

$$\mathcal{F} = \sup \{\mathcal{G} \mid \mathcal{G} : \mathcal{X} \to \mathbb{R} \text{ affin und stetig mit } \mathcal{G} \leq \mathcal{F}\},$$

wobei man ein Funktional $\mathcal{H} : \mathcal{X} \to \mathbb{R}$ genau dann affin *nennt, falls $\mathcal{H} - \mathcal{H}(0)$ linear ist.*◇

Für den Nachweis dieser Aussage benötigen wir eine geeignete Topologie auf $(\mathcal{X} \times \mathbb{R})$.

(A.3.6) Bemerkung
Wie in Bemerkung (A.2.7) b) erhält man auf kanonische Art und Weise eine Topologie auf $\mathcal{X} \times \mathbb{R}$: Denn die Abbildung

$$\langle\cdot,\cdot\rangle_{\mathcal{X}\times\mathbb{R}} : (\mathcal{X} \times \mathbb{R}) \times (\mathcal{X} \times \mathbb{R}) \to \mathbb{R}, \quad \langle(x,z),(y,w)\rangle_{\mathcal{X}\times\mathbb{R}} := \langle x,y\rangle + zw$$

definiert ein Skalarprodukt, sodass $(\mathcal{X} \times \mathbb{R}, \langle\cdot,\cdot\rangle_{\mathcal{X}\times\mathbb{R}})$ ein Prä-Hilbertraum ist bzw. sogar ein Hilbertraum, da sich aufgrund dieser Definition von $\langle\cdot,\cdot\rangle_{\mathcal{X}\times\mathbb{R}}$ die Vollständigkeit von $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ überträgt. ◇

Beweis von Lemma (A.3.5)
„$\geq$“ Klar.

„$\leq$“ Dabei reicht es zu zeigen, dass für alle $(x,a) \in \mathcal{X} \times \mathbb{R}$ mit $\mathcal{F}(x) > a$ ein affines, stetiges Funktional $\mathcal{H} : \mathcal{X} \to \mathbb{R}$ mit $\mathcal{H} \leq \mathcal{F}$ existiert, sodass $\mathcal{H}(x) \geq a$ gilt; denn nehmen wir an, dass wir ein $y \in \mathcal{X}$ mit

$$\mathcal{F}(y) > \sup \{\mathcal{G}(y) \mid \mathcal{G} : \mathcal{X} \to \mathbb{R} \text{ affin und stetig mit } \mathcal{G} \leq \mathcal{F}\} =: \mathcal{J}(y)$$

finden. Dann können wir ein $a \in \mathbb{R}$ mit $\mathcal{F}(y) > a > \mathcal{J}(y)$ auswählen, für das ein wie oben beschriebenes Funktional $\mathcal{H}$ existiert. Wegen der Definition von $\mathcal{J}(y)$ erhalten wir somit

$$\mathcal{H}(y) \geq a > \mathcal{J}(y) \geq \mathcal{H}(y),$$

also einen Widerspruch.
Daher sei $(x,a) \in \mathcal{X} \times \mathbb{R}$ mit $\mathcal{F}(x) > a$ beliebig, d.h. $(x,a) \notin$ epi $\mathcal{F}$. Einerseits ist mit dom $\mathcal{F} \neq \emptyset$ auch epi $\mathcal{F} \neq \emptyset$. Andererseits folgt aus der Unterhalbstetigkeit und Konvexität von $\mathcal{F}$ sowie den Lemmata (A.2.8) und (A.3.2), dass epi $\mathcal{F}$ abgeschlossen und konvex ist. Da $(\mathcal{X} \times \mathbb{R}, \langle\cdot,\cdot\rangle_{\mathcal{X}\times\mathbb{R}})$ nach Bemerkung (A.3.6) obendrein ein Hilbertraum ist, ist damit der Satz von Mazur für Hilberträume, also der nachfolgende Satz (A.3.7), anwendbar. Demnach existieren ein $(y,b) \in \mathcal{X} \times \mathbb{R}$ und ein $C \in \mathbb{R}$, sodass

$$\langle(y,b),(x,a)\rangle_{\mathcal{X}\times\mathbb{R}} > C \quad \text{und} \quad \langle(y,b),(z,d)\rangle_{\mathcal{X}\times\mathbb{R}} \leq C \ \text{ für alle } (z,d) \in \text{epi } \mathcal{F}$$

ist. Dies impliziert

$$\langle y,x\rangle + ba = \langle (y,b),(x,a)\rangle_{\mathcal{X}\times\mathbb{R}} > C \geq \langle (y,b),(z,d)\rangle_{\mathcal{X}\times\mathbb{R}} = \langle y,z\rangle + bd \tag{A.14}$$

für alle $(z,d) \in \operatorname{epi} \mathcal{F}$. Wenn wir nun ein $(z,d) \in \operatorname{epi} \mathcal{F} \neq \emptyset$ wählen, so ist auch $(z,e) \in \operatorname{epi} \mathcal{F}$ für alle $e \geq d$. Daher gilt die Ungleichung (A.14) bei festgehaltenem z für jedes $e \geq d$ anstelle von d, sodass $b \leq 0$ sein muss. Jetzt definieren wir noch das Funktional

$$\mathcal{I} : \mathcal{X} \times \mathbb{R} \to \mathbb{R}, \quad \mathcal{I}(z,d) := \langle (y,b),(z,d)\rangle_{\mathcal{X}\times\mathbb{R}} = \langle y,z\rangle + bd,$$

das offensichtlich stetig und linear ist, und unterscheiden zwei Fälle bzgl. b:

<u>*1. Fall:*</u> $b < 0$
Dann sei

$$\mathcal{H} : \mathcal{X} \to \mathbb{R}, \quad \mathcal{H}(z) := \frac{1}{b} \cdot \big(\mathcal{I}(x,a) - \langle y,z\rangle\big).$$

Nach Konstruktion ist $\mathcal{H}$ affin, stetig und es erfüllt $\mathcal{H}(x) = a$. Zudem ist mit $z \in \operatorname{dom} \mathcal{F}$ offenbar $(z, \mathcal{F}(z)) \in \operatorname{epi} \mathcal{F}$, sodass uns die Ungleichung (A.14) für $d = \mathcal{F}(z)$ wegen $b < 0$ gerade

$$\mathcal{H}(z) = \frac{1}{b} \cdot \big(\underbrace{\mathcal{I}(x,a)}_{=\ \langle y,x\rangle + ba} - \langle y,z\rangle\big) \leq \frac{1}{b} \cdot \Big[\big(\langle y,z\rangle + b\mathcal{F}(z)\big) - \langle y,z\rangle\Big] = \mathcal{F}(z)$$

liefert. Hingegen gilt $\mathcal{H}(z) \leq \mathcal{F}(z)$ für $z \notin \operatorname{dom} \mathcal{F}$ ohnehin schon, d.h. insgesamt erfüllt $\mathcal{H}$ alle gewünschten Eigenschaften.

<u>*2. Fall:*</u> $b = 0$
Die Ungleichung (A.14) reduziert sich somit darauf, dass

$$\langle y,x\rangle > C \geq \langle y,z\rangle \tag{A.15}$$

für alle $(z,d) \in \operatorname{epi} \mathcal{F}$ bzw. einfach $z \in \operatorname{dom} \mathcal{F}$ ist. Weil $z = x$ diese strikte Ungleichung nicht erfüllt, muss $x \notin \operatorname{dom} \mathcal{F}$, also $\mathcal{F}(x) = \infty$, gelten. Außerdem ergibt sich folgende

Zwischenbehauptung: Es existiert ein affines, stetiges Funktional $\mathcal{G}$ mit $\mathcal{G} \leq \mathcal{F}$.

Dazu: Sei $\hat{x} \in \operatorname{dom} \mathcal{F} \neq \emptyset$, wozu wir zunächst ein $\hat{a} \in \mathbb{R}$ mit $\hat{a} < \mathcal{F}(\hat{x})$ finden können. Anschließend gehen wir wie oben vor und der Satz von Mazur über Hilberträumen liefert wieder ein $(\hat{y}, \hat{b}) \in \mathcal{X} \times \mathbb{R}$ und ein $\hat{C} \in \mathbb{R}$ wie oben. Mit denselben Argumenten folgt dann $\hat{b} \leq 0$. Jedoch ist hier $\hat{b} = 0$ ausgeschlossen; denn ansonsten ergibt sich die zu (A.15) analoge Ungleichung

$$\langle \hat{y}, \hat{x}\rangle > \hat{C} \geq \langle \hat{y}, z\rangle$$

für alle $z \in \operatorname{dom} \mathcal{F}$. Dies impliziert dann aber mit dem gleichen Argument (direkt nach Ungleichung (A.15)), dass $\hat{x} \notin \operatorname{dom} \mathcal{F}$ ist, sodass wir einen Widerspruch erhalten.
Demnach muss $\hat{b} < 0$ sein, sodass uns der 1. Fall dieses $\mathcal{G}$ beschert. Damit ist die Zwischenbehauptung gezeigt.

Neben diesem $\mathcal{G}$ seien $\alpha > 0$ beliebig und $\beta \in \mathbb{R}$ mit $\langle y, x\rangle > \beta > C$, sodass dies zusammen mit Ungleichung (A.15)

$$\beta > \langle y, z\rangle \tag{A.16}$$

für alle $z \in \operatorname{dom} \mathcal{F}$ liefert. So definieren wir das Funktional

$$\mathcal{H}_\alpha : \mathcal{X} \to \mathbb{R}, \quad \mathcal{H}_\alpha(z) := \mathcal{G}(z) + \alpha(\langle y, z\rangle - \beta),$$

das nach Konstruktion affin und stetig ist. Dann ist zum Einen nach Ungleichung (A.16) und wegen $\mathcal{G} \leq \mathcal{F}$ ebenfalls

$$\mathcal{H}_\alpha(z) = \mathcal{G}(z) + \alpha(\underbrace{\langle y, z\rangle - \beta}_{<\,0}) < \mathcal{G}(z) \leq \mathcal{F}(z)$$

für alle $z \in \operatorname{dom} \mathcal{F}$, wobei $\mathcal{H}_\alpha(z) \leq \mathcal{F}(z)$ wiederum auch für alle $z \notin \operatorname{dom} \mathcal{F}$ gilt. Zum Anderen finden wir wegen der ursprünglichen Anforderung an β ein hinreichend großes $\alpha_0 > 0$, sodass obendrein

$$\mathcal{H}_{\alpha_0}(x) = \mathcal{G}(x) + \alpha_0(\underbrace{\langle y, x\rangle - \beta}_{>\,0}) \geq a$$

ist, d.h. $\mathcal{H} := \mathcal{H}_{\alpha_0}$ genügt allen gewünschten Anforderungen. □

Über Hilberträumen können wir den Satz von Mazur wie folgt formulieren.

(A.3.7) Satz (Mazur)
Sei $(\mathcal{X}, \langle\cdot,\cdot\rangle)$ sogar ein Hilbertraum und $M \subseteq \mathcal{X}$ nicht-leer, abgeschlossen und konvex. Dann existiert zu jedem $x_0 \notin M$ ein $y \in \mathcal{X}$ und ein $C \in \mathbb{R}$, sodass

$$\langle y, x_0\rangle > C \qquad \textit{und} \qquad \langle y, x\rangle \leq C \quad \textit{für alle } x \in M$$

gilt. ◇

Beweis
Zunächst definieren wir zu einem $z \in M \neq \emptyset$ die Menge $N := M - z$, die nach den Annahmen bzgl. M ebenfalls nicht-leer, abgeschlossen und konvex ist. Darüber hinaus ist

$$0 = z - z \in N$$

und $w := x_0 - z \notin N$; denn sonst wäre $x_0 = w + z \in N + z = M$, das aber hier ausgeschlossen wird. Somit ist die klassische Variante des Satzes von Mazurs, siehe dazu [10, Satz (5.3.1)], anwendbar, wonach ein lineares und stetiges Funktional $\mathcal{F}$ in seinem Dualraum $\mathcal{X}'$ existiert, sodass

$$\mathcal{F}(w) > 1 \qquad \text{und} \qquad \mathcal{F}(v) \leq 1 \quad \text{für alle } v \in N \tag{A.17}$$

erfüllt ist. Nach dem Riesz'schen Darstellungssatz finden wir nun ein $y \in \mathcal{X}$ mit $\mathcal{F}(x) = \langle y, x \rangle$ für alle $x \in \mathcal{X}$. In (A.17) übersetzt sich dies dann zu

$$\langle y, w \rangle > 1 \qquad \text{und} \qquad \langle y, v \rangle \leq 1 \quad \text{für alle } v \in N.$$

Aufgrund der Beziehung $N = M - z$ erhalten wir daraus schließlich

$$\langle y, x_0 \rangle - \langle y, z \rangle > 1 \qquad \text{und} \qquad \langle y, x \rangle - \langle y, z \rangle \leq 1 \quad \text{für alle } x \in M,$$

sodass für $C := 1 + \langle y, z \rangle$ die Behauptung folgt. □

Beweis von Satz (A.3.4)
(i) „ $\leq$ “ Wegen der Definition der Fenchel-Konjugierten ist zunächst

$$\mathcal{F}^*(y) = \sup_{z \in \mathcal{X}} \{\langle z, y \rangle - \mathcal{F}(z)\} \geq \langle x, y \rangle - \mathcal{F}(x)$$

für alle $x, y \in \mathcal{X}$. Außerdem ist $\mathcal{F}^* > -\infty$ aufgrund der Voraussetzung, dass dom $\mathcal{F}$ nicht-leer ist, in Kombination mit Lemma (A.3.3). Zu gegebenen $x, y \in \mathcal{X}$ folgt dann mit einer einfachen Fallunterscheidung, ob $\mathcal{F}$ an der Stelle x bzw. $\mathcal{F}^*$ an der Stelle y einen endlichen Wert oder eben ∞ annimmt, dass auch

$$\mathcal{F}(x) \geq \langle x, y \rangle - \mathcal{F}^*(y)$$

für alle $x, y \in \mathcal{X}$ ist. Dies impliziert gerade, dass

$$\mathcal{F}(x) \geq \sup_{y \in \mathcal{X}} \{\langle x, y \rangle - \mathcal{F}^*(y)\} = \mathcal{F}^{**}(x)$$

für alle $x \in \mathcal{X}$ gilt, und somit erhalten wir $\mathcal{F} \geq \mathcal{F}^{**}$.

„ $\geq$ “ Zuerst ergibt sich aus Lemma (A.3.5) die Identität

$$\mathcal{F} = \sup \{\mathcal{G} \mid \mathcal{G} : \mathcal{X} \to \mathbb{R} \text{ affin und stetig mit } \mathcal{G} \leq \mathcal{F}\}. \tag{A.18}$$

Diese ist der Schlüssel für unsere anschließende Argumentation. Also sei mit $\mathcal{G} : \mathcal{X} \to \mathbb{R}$ ein beliebiges affines und stetiges Funktional gegeben, das $\mathcal{G} \leq \mathcal{F}$ erfüllt. Aufgrund der Definition affiner Funktionale, siehe Lemma (A.3.5), finden wir mithilfe des Riesz'schen Darstellungssatzes ein $y \in \mathcal{X}$, sodass $\mathcal{G}(x) = \langle y, x \rangle + \mathcal{G}(0)$ für alle $x \in \mathcal{X}$ ist, wobei ab jetzt $C := -\mathcal{G}(0)$ sei. Nach Annahme ist dann

$$\langle y, x \rangle - C = \mathcal{G}(x) \leq \mathcal{F}(x)$$

für alle $x \in \mathcal{X}$. Mit einer erneuten Fallunterscheidung bzgl. der Werte, die $\mathcal{F}$ an der Stelle eines beliebigen $x \in \mathcal{X}$ annimmt, folgt daraus

$$C \geq \langle y, x \rangle - \mathcal{F}(x)$$

für alle $x \in \mathcal{X}$ und daher

$$C \geq \sup_{x \in \mathcal{X}} \{\langle y, x \rangle - \mathcal{F}(x)\} = \mathcal{F}^*(y).$$

Dies liefert aber wiederum, dass

$$\mathcal{G}(x) = \langle y, x \rangle - C \leq \langle y, x \rangle - \mathcal{F}^*(y) \leq \mathcal{F}^{**}(x)$$

für alle $x \in \mathcal{X}$ gilt. Insgesamt ist damit $\mathcal{G} \leq \mathcal{F}^{**}$ und, da $\mathcal{G}$ beliebig gewählt war, erhalten wir mit der Identität (A.18)

$$\mathcal{F}^{**} \geq \sup \{\mathcal{G} \mid \mathcal{G} : \mathcal{X} \to \mathbb{R} \text{ affin und stetig mit } \mathcal{G} \leq \mathcal{F}\} = \mathcal{F}.$$

(ii) Nach Lemma (A.3.3) ist schon bekannt, dass $\mathcal{F}^* > -\infty$ gilt sowie $\mathcal{F}^*$ konvex und unterhalbstetig ist.
Schließlich bleibt zu zeigen, dass dom $\mathcal{F}^*$ nicht-leer ist. Dazu nehmen wir das Gegenteil an, d.h. es sei $\mathcal{F}^* \equiv \infty$. Dann muss

$$-\infty \equiv \mathcal{F}^{**} \underset{\text{(i)}}{=} \mathcal{F}$$

sein, das aber der Anforderung $\mathcal{F} > -\infty$ widerspricht. $\square$

Literaturverzeichnis

[1] C. Acerbi, D. Tasche. ON THE COHERENCE OF EXPECTED SHORTFALL. J. Banking Finance 26, S. 1487-1503, 2002
URL: https://www.researchgate.net/profile/Dirk_Tasche/publication/222405274_On_the_Coherence_of_Expected_Shortfall/links/59ea35404585151983c7eae4/On-the-Coherence-of-Expected-Shortfall.pdf?origin=publication_detail

[2] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath. COHERENT MEASURES OF RISK. Math. Finance 9, S. 203-227, 1999
URL: https://people.math.ethz.ch/∼delbaen/ftp/preprints/CoherentMF.pdf

[3] Martin Brokate. KONVEXE ANALYSIS. Skript zur gleichnamigen Vorlesung. Zentrum Mathematik, TU München, 2009
URL: http://www-m6.ma.tum.de/∼brokate/con_ss09.pdf

[4] F. Delbaen. Entwurf: COHERENT RISK MEASURES. Vorlesungsmitschrift, Pisa, 2000
URL: http://www.math.ethz.ch/∼delbaen/ftp/preprints/PISA007.pdf

[5] F. Delbaen. COHERENT RISK MEASURES ON GENERAL PROBABILITY SPACES. **Aus:** K. Sandmann, P.J. Schönbucher. ADVANCES IN FINANCE AND STOCHASTICS. ESSAYS IN HONOUR OF DIETER SONDERMANN. Springer 2002: S. 1-37, Berlin, Heidelberg, New York
URL: https://people.math.ethz.ch/∼delbaen/ftp/preprints/RiskMeasuresGeneralSpaces.pdf

[6] P. Embrechts. EXTREME VALUE THEORY: POTENTIAL AND LIMITATIONS AS AN INTEGRATED RISK MANAGEMENT TOOL. Working paper, ETH Zürich, 2000
URL: https://people.math.ethz.ch/∼embrecht/ftp/evtpot.pdf

[7] Martin Frank. NUMERISCHE ANALYSIS I. Skript zur gleichnamigen Vorlesung. Center for Computational Engineering Science Mathematics Division, RWTH Aachen, 2016

[8] Udo Kamps und Erhard Cramer. STOCHASTIK II. Skript zur gleichnamigen Vorlesung. Institut für Statistik und Wirtchaftsmathematik, RWTH Aachen, 2016

[9] Stanislaus Maier-Paape, Qiji Jim Zhu. A General Framework for Portfolio Theory - Part I: Theory and Various Models. Risks, 6(2):53, 2018
URL: http://www.mdpi.com/2227-9091/6/2/53/pdf

[10] Stanislaus Maier-Paape. Funktionalanalysis. Skript zur gleichnamigen Vorlesung. Institut für Mathematik, RWTH Aachen, 2018
URL: http://www.instmath.rwth-aachen.de/∼maier/lecture_notes/Funkionalanalysis_Maier-Paape.pdf

[11] R. Tyrell Rockafellar. Convex Analysis. Princeton University Press 1970, Princeton (New Jersey)

[12] R. Tyrell Rockafellar, Stan Uryasev und Michael Zabarankin. Generalized deviations in risk analysis. Finance Stochast. 10, S. 51 - 74, 2006
URL: http://www.ise.ufl.edu/uryasev/files/2011/11/devia_fanance_an_stochastics.pdf

[13] R. Tyrell Rockafellar, Stan Uryasev und Michael Zabarankin. Master funds in portfolio analysis with general deviation measures. J. Banking Finance 30, S. 743-778, 2006
URL: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID615823_code223822.pdf?abstractid=615823&mirid=1&type=2