

Bruno Ebner

Zur Asymptotik eines mit quadratischen Abhängig- keiten operierenden Tests auf multivariate Normal- verteilung



Scientific
Publishing

Bruno Ebner

**Zur Asymptotik eines mit quadratischen Abhängigkeiten
operierenden Tests auf multivariate Normalverteilung**

Zur Asymptotik eines mit quadratischen Abhängigkeiten operierenden Tests auf multivariante Normalverteilung

von
Bruno Ebner

Dissertation, Karlsruher Institut für Technologie
Fakultät für Mathematik
Tag der mündlichen Prüfung: 10.02.2010
Referenten: Prof. Dr. N. Henze, Prof. Dr. L. Baringhaus

Impressum

Karlsruher Institut für Technologie (KIT)
KIT Scientific Publishing
Straße am Forum 2
D-76131 Karlsruhe
www.ksp.kit.edu

KIT – Universität des Landes Baden-Württemberg und nationales
Forschungszentrum in der Helmholtz-Gemeinschaft



Diese Veröffentlichung ist im Internet unter folgender Creative Commons-Lizenz
publiziert: <http://creativecommons.org/licenses/by-nc-nd/3.0/de/>

KIT Scientific Publishing 2010
Print on Demand

ISBN 978-3-86644-549-9



Zur Asymptotik eines mit quadratischen Abhängigkeiten operierenden Tests auf multivariate Normalverteilung

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

an der Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Dipl.-Math. oec. Bruno Ebner
aus Aachen

Tag der mündlichen Prüfung:
Referent:
Korreferent:

10. Februar 2010
Prof. Dr. N. Henze
Prof. Dr. L. Baringhaus

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit am Institut für Stochastik des Karlsruher Instituts für Technologie (KIT). Mein herzlicher Dank richtet sich an all diejenigen, die mich bei der Entstehung tatkräftig unterstützt haben:

- an Herrn Prof. Dr. N. Henze für seine Unterstützung während der Entstehung der vorliegenden Arbeit. Ohne die vielen wertvollen Hinweise und hilfreichen Diskussionen wäre diese Arbeit in dieser Form nicht möglich gewesen,
- an Herrn Prof. Dr. L. Baringhaus für die freundliche Übernahme des Korreferats,
- an Herrn Dr. B. Klar und Frau Dr. W. Lao für viele hilfreiche und konstruktive Ratschläge, die diese Arbeit während ihrer Entstehung begleitet haben,
- an Herrn Dr. M. Folkers für viele hilfreiche Hinweise zur Literatur und für die hervorragende mathematische Bibliothek, welche keine Wünsche offen lässt,
- an Herrn Dipl.-Math. oec. D. Engel für die R -Unterstützung,
- an Herrn Dr. M. Ruggenthaler und Herrn A. Rohde für geduldiges Korrekturlesen,
- an meine Eltern, meinen Bruder, meine Freundin und meinen Freundeskreis dafür, dass Sie mich immer unterstützt und viel Verständnis gezeigt haben,
- an das gesamte Institut für Stochastik, insbesondere an Frau Michaela Regelin und Frau Tatjana Dominic, für das angenehme Arbeitsklima während der Jahre am Institut.

Karlsruhe, den 10. August 2010

Inhaltsverzeichnis

1	Einleitung	1
2	Die Teststatistik von Cox und Small	5
2.1	Affine Invarianz der Teststatistik T_n	8
2.2	Analyse der Teststatistik T_n	10
3	Konvergenz ohne geschätzte Parameter unter sphärischer Symmetrie	15
4	Konvergenz mit standardisierten Daten unter sphärischer Symmetrie	25
4.1	Konvergenzaussagen für $C_{n,k}$	25
4.2	Konvergenzaussagen für D_n	33
4.3	Konvergenzaussagen für W_n	36
4.4	Die Grenzverteilung der Teststatistik T_n	40
5	Beispiele	41
5.1	Die Grenzverteilung unter H_0	41
5.2	Multivariate Pearson Typ II-Verteilung	42
5.3	Multivariate Pearson Typ VII-Verteilung	43
6	Konvergenzaussagen unter benachbarten Alternativen	47
7	Konvergenzaussagen unter festen Alternativen	55
7.1	Stochastische Konvergenz der Teststatistik T_n unter festen Alternativen	56
7.2	Konvergenzaussagen für $C_{n,k}$	60
7.3	Konvergenzaussagen für D_n	67
7.4	Konvergenzaussagen für W_n	70
8	Simulationen	73
8.1	Implementierung des Funktionals $\max_{b \in \mathcal{S}^{d-1}}(\cdot)$	73
8.1.1	Gleichverteilte Datenpunkte auf der Einheitssphäre \mathcal{S}^{d-1}	74
8.2	Empirische Quantile unter der Nullhypothese	76
8.3	Klassische Verfahren zum Testen auf multivariate Normalverteilung	77
8.4	Simulationen unter festen Alternativen	80
8.4.1	Multivariate Pearson Typ II und Typ VII-Verteilung	80
8.4.2	Multivariate Burr, Pareto und Logistische Verteilungen	81
8.5	Vergleich der Güte der betrachteten Anpassungstests	83

9	Abschließende Bemerkungen	89
A	Technische Berechnungen	91
A.1	Gemischte Momente von Komponenten von $U \sim \mathcal{U}(\mathcal{S}^{d-1})$	92
A.2	Kovarianzen	93
A.3	Beweis von Theorem 3.1.7	100
A.4	Die Entropiebedingung in \mathcal{S}^{d-1}	103
B	Schwache Konvergenz im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$	105
B.2	Verteilungskonvergenz in $C(\mathcal{S}^{d-1}, \mathbb{R})$	106
B.3	Straffheit in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$	107
B.4	Der zentrale Grenzwertsatz in $C(\mathcal{S}^{d-1}, \mathbb{R})$	108
C	Kugeloberflächenfunktionen	109
C.2	Legendre-Polynome	110
C.3	Das Funk-Hecke Theorem	112
C.4	Eine Methode zur Berechnung von $E((b'U)^r(c'U)^s)$	112
D	Aussagen der asymptotischen Statistik	115
D.1	Benachbarte Alternativen	115
	Literaturverzeichnis	119

Tabellenverzeichnis

8.1	Größenordnung von \mathcal{U}_n für verschiedene n und verschiedene Dimensionen d	75
8.2	Empirische 95%-Quantile unter der Nullhypothese für die Teststatistik von Cox und Small	76
8.3	Empirische 99%-Quantile unter der Nullhypothese für die Teststatistik von Cox und Small	76
8.4	Empirische 95%-Quantile unter der Nullhypothese für die Schiefestatistik von Mardia für die Stichprobenumfänge $n = 25, 50, 100$	77
8.5	Empirische 95%-Quantile unter der Nullhypothese für die Schiefestatistik von Malkovich und Afifi für die Stichprobenumfänge $n = 25, 50, 100$	78
8.6	Empirische 95%-Quantile unter der Nullhypothese für die BHEP-Teststatistik für $n = 50$ in Abhängigkeit des Parameters $\beta > 0$	79
8.7	Empirische 95%-Quantile unter der Nullhypothese für die Teststatistik von Székely und Rizzo für die Stichprobenumfänge $n = 25, 50, 100$	79
8.8	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 2, 3, 5, 10$)	80
8.9	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für Pearson Typ VII-Verteilungen mit Parameter a ($n = 25; d = 2, 3, 5, 10$)	81
8.10	Erzeugungsvorschriften für multivariat Burr-, Pareto- und Logistisch verteilte Zufallsvektoren	82
8.11	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für multivariate Burr-, Pareto- und Logistische Verteilungen ($n = 25; d = 2, 3, 5, 10$)	83
8.12	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für multivariate Burr, Pareto und Logistische Verteilungen ($n = 50; d = 2, 3, 5, 10$)	83
8.13	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 2$)	84
8.14	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 3$)	85
8.15	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 5$)	85
8.16	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ für Pearson Typ VII-Verteilungen mit Parameter a ($n = 25; d = 2$)	87
8.17	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ für Pearson Typ VII-Verteilungen mit Parameter a bei einem Stichprobenumfang von ($n = 25; d = 3$)	87
8.18	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ für Pearson Typ VII-Verteilungen mit Parameter a ($n = 25; d = 5$)	87
8.19	Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für multivariate Burr, Pareto und Logistische Verteilungen ($n = 25; d = 2, 3, 5$)	88
A.1	Explizite Werte bestimmter Oberflächenintegrale	92

Abbildungsverzeichnis

8.1	Punktwolken von einer multivariaten Burr-, Pareto- und Logistische-Verteilung mit gleichverteilten Komponenten (links) und logistisch-verteilten Komponenten (rechts)	86
8.2	Punktwolken von einer multivariaten Burr-, Pareto- und Logistische-Verteilung mit Burr-verteilten Komponenten (links) und Pareto-verteilten Komponenten (rechts)	86

Kapitel 1

Einleitung

In der induktiven Statistik werden Beobachtungen x_1, \dots, x_n als Realisierungen von Zufallsvariablen X_1, \dots, X_n aufgefasst. Um Schlussfolgerungen aus den beobachteten Daten zu ziehen, die über deren reine Beschreibung hinausgehen, sind Annahmen über den zugrunde liegenden Zufallsmechanismus nötig. Demnach ist ein stochastisches Modell zu spezifizieren, welches die Verteilung von X_1, \dots, X_n beschreibt. Dabei ist ein gewisser Trade-Off unvermeidbar, denn einerseits lässt eine sehr genaue Spezifikation des Modells weitreichendere Schlüsse zu, andererseits läuft man dadurch aber Gefahr, falsche Aussagen zu treffen, falls die Daten durch das Modell doch nicht richtig erfasst werden. In diesem Fall wären möglichst schwache Annahmen zu bevorzugen. Die Statistik schlägt für dieses Dilemma einen Kompromiss vor: man wählt zunächst einen Modellrahmen und überprüft anschließend mit statistischen Methoden, ob dieser Rahmen mit den vorhandenen Daten verträglich ist. Falls dies bejaht werden kann, so können auf das Rahmenmodell aufbauende statistische Verfahren angewandt werden und über die reine Beschreibung der Daten hinausgehende Schlüsse gezogen werden.

Die Überprüfung des Modellrahmens erfolgt mit sogenannten Anpassungstests. Dabei treffen wir im Folgenden die grundlegende Annahme, dass X_1, \dots, X_n unabhängige identisch verteilte Zufallsvariable sind. Diese Voraussetzung kann in der Praxis durch kontrollierte, wiederholte Experimente zumindest näherungsweise gewährleistet werden. Wir bezeichnen mit Θ den Parameterraum einer Verteilungsfamilie $\{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$, die durch einen endlichdimensionalen Vektor ϑ parametrisiert ist, und schreiben \mathbb{P}^{X_1} für die unbekannte Wahrscheinlichkeitsverteilung von X_1 . Dann testet ein Anpassungstest die Hypothese

$$H_0 : \mathbb{P}^{X_1} \in \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$$

gegen die allgemeine Alternative

$$H_1 : \mathbb{P}^{X_1} \notin \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}.$$

Ein solches Verfahren beruht auf einer geeigneten Teststatistik, also einer Funktion der Daten, welche zu einem so genannten Testentscheid führt. Liegt dabei der Wert dieser Funktion in einem noch zu definierenden Ablehnbereich, so wird die Hypothese verworfen. Um einen Ablehnbereich angeben zu können, muss die Teststatistik unter der Hypothese H_0 mit wahrscheinlichkeitstheoretischen Mitteln untersucht werden. In der Regel wird die Teststatistik

auf einem Merkmal aufbauen, das die Verteilungsfamilie charakterisiert. So kann ein Vergleichskriterium zur Überprüfung, ob die vorliegenden Daten aus $\{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ stammen können oder nicht, geschaffen werden.

Viele der klassischen Anpassungstests für univariate Daten basieren auf der Tatsache, dass die Verteilungsfunktion $F_\vartheta(\cdot)$ durch den Parametervektor ϑ eindeutig bestimmt ist. Es liegt also nahe, ϑ durch die Realisierung eines konsistenten Schätzers $\hat{\vartheta}_n$ zu ersetzen und diese als Modellparameter in die Verteilungsfunktion einzusetzen. Ein geeignetes Abstandsmaß auf Funktionenräumen zwischen der Funktion $F_{\hat{\vartheta}_n}(\cdot)$ und der empirischen Verteilungsfunktion $F_n(t) := \frac{1}{n} \sum_{j=1}^n \mathbf{1}\{X_j \leq t\}$ führt dann zur Gruppe der klassischen Anpassungstests, welche als prominente Vertreter den Kolmogorov-Smirnov-Test und den Cramér-von Mises Test beinhaltet. Da diese Verfahren nur schwer auf multivariate Daten verallgemeinert werden können, beruhen moderne Verfahren auf anderen Merkmalen.

Ein wichtiger Modellrahmen in der Statistik ist die Lokations-Skalenfamilie der multivariaten Normalverteilung, also die Familie $\mathcal{N}_d := \{\mathcal{N}_d(\mu, \Sigma) : \mu \in \mathbb{R}^d, \Sigma \in \mathbb{R}^{d \times d} \text{ positiv definit}\}$ mit dem Lokations- und Skalenparameter $\vartheta = (\mu, \Sigma)$. Wir beschränken uns in dieser Arbeit auf Anpassungstests auf multivariate Normalverteilung. Ein besonderes Merkmal der Normalverteilung ist, dass sie durch den Lokations- und Skalenparameter $\vartheta = (\mu, \Sigma)$ eindeutig bestimmt ist. Da mit der Kovarianzmatrix lediglich der lineare Zusammenhang zwischen den einzelnen Komponenten eines d -variater normalverteilter Zufallsvektors erfasst wird, liegen demnach keine quadratischen oder höheren Abhängigkeiten vor. Dieses Merkmal nutzt der in dieser Arbeit betrachtete Test von Cox und Small aus. Dabei wird durch eine empirische Standardisierung der Daten (Zufallsvektoren) X_1, \dots, X_n mit Hilfe des Stichprobenmittels \bar{X}_n und der empirischen Kovarianzmatrix S_n durch $Y_{n,j} := S_n^{-\frac{1}{2}}(X_j - \bar{X}_n)$ für $j = 1, \dots, n$ ein Vergleich ohne einen Schätzer $\hat{\vartheta}_n$ durchgeführt, da in diesem Fall unter H_0 ohne Beschränkung der Allgemeinheit eine Standardnormalverteilung angenommen werden kann.

Da die Klasse \mathcal{N}_d der nichtdegenerierten d -dimensionalen Normalverteilungen abgeschlossen bezüglich nichtsingulärer affiner Transformationen ist, ist die affine Invarianz eine natürliche Forderung an eine Teststatistik für das vorgelegte Anpassungstestproblem. Dabei nennt man eine Teststatistik T_n affin invariant, falls für jede nicht singuläre $d \times d$ -Matrix A und jeden Vektor $\beta \in \mathbb{R}^d$

$$T_n(AX_1 + \beta, \dots, AX_n + \beta) = T_n(X_1, \dots, X_n)$$

gilt. Ein weiteres Kriterium, das ein „guter“ Test auf multivariate Normalverteilung erfüllen sollte, ist die globale Konsistenz, also die Konsistenz gegenüber jeder alternativen Verteilung. Dabei versteht man unter der Konsistenz eines Tests gegenüber einer Alternative, dass die Wahrscheinlichkeit für das Verwerfen der Hypothese, mit wachsendem Stichprobenumfang gegen 1 konvergiert. Für weitere Interpretationen zu den Themen affine Invarianz und Konsistenz eines Anpassungstests auf multivariate Normalverteilung verweisen wir auf den Übersichtsartikel von Henze [16].

Der in dieser Arbeit untersuchte Test auf multivariate Normalverteilung von Cox und Small ist affin invariant, aber wie sich in den Kapiteln 4 und 7 zeigt nicht global konsistent.

Die vorliegende Dissertation ist wie folgt gegliedert: Kapitel 2 führt die Teststatistik T_n von Cox und Small ein und zeigt durch eine kurze Analyse von T_n Möglichkeiten für die Herleitung von wahrscheinlichkeitstheoretischen Konvergenzaussagen auf. Danach gliedert sich die Arbeit in mehrere Kapitel zum Konvergenzverhalten der Teststatistik unter bestimmten Verteilungsannahmen. Im Detail geht es in Kapitel 3 um das theoretische Grundgerüst der Arbeit. Im vereinfachten Fall nichtstandardisierter sphärisch-symmetrisch verteilter Daten wird unter einer gewissen Momentenbedingung an die Radienverteilung die schwache Konvergenz des stochastischen Prozesses $W_n(\cdot)$ aus (2.13) in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ nachgewiesen und der probabilistische Zusammenhang zu T_n hergestellt. Kapitel 4 erweitert die Ergebnisse von Kapitel 3 auf den Fall standardisierter Daten und gibt die Grenzverteilung der Teststatistik unter einer sphärisch-symmetrischen Verteilungsannahme an. In Kapitel 5 werden explizite Beispiele zu den gemachten Aussagen angegeben, unter anderem die H_0 -Verteilung der Teststatistik T_n . Gegenstand des Kapitels 6 ist die Analyse des Grenzverhaltens der Teststatistik T_n unter benachbarten Alternativen, und es wird die schwache Konvergenz des stochastischen Prozesses $W_n(\cdot)$ gegen einen um einen bestimmten Shiftparameter verschobenen Grenzprozess gezeigt. Das Kapitel 7 schließt die theoretische Arbeit mit einem Resultat über die stochastische Konvergenz der Teststatistik gegen einen von der jeweiligen zugrunde liegenden festen Alternative abhängenden Parameter ab. Kapitel 8 bestätigt die theoretischen Ergebnisse in einer Simulationsstudie und stellt ein Verfahren zur Implementierung des Maximumfunktionals vor. Den Abschluss der Arbeit bilden ein umfangreicher Anhang mit technischen Berechnungen, ergänzenden Ergebnissen und Literaturangaben.

Um Unklarheiten zu vermeiden, soll auf einige grundlegende Vereinbarungen und Notationen hingewiesen werden. In dieser Arbeit sei d die Dimension des zugrunde liegenden Vektorraums \mathbb{R}^d . Es bezeichne $\|\cdot\|$ die Euklidnorm und $\mathcal{S}^{d-1} := \{x \in \mathbb{R}^d : \|x\| = 1\}$ die Einheitssphäre im \mathbb{R}^d . Des Weiteren steht $|\mathcal{S}^{d-1}| = 2\pi^{\frac{d}{2}} / \Gamma(\frac{d}{2})$ für die Oberfläche von \mathcal{S}^{d-1} ; dabei ist mit $\Gamma(x)$, $x > 0$, die Gammafunktion gemeint. Die Betafunktion wird mit $B(x, y)$, $x, y > 0$, bezeichnet. Wir schreiben $\lambda^d(\cdot)$ für das d -dimensionale Lebesgue-Maß und $\mathcal{S}^{d-1}(\cdot)$ für das sphärische Lebesgue-Maß auf \mathcal{S}^{d-1} . Vektoren in \mathbb{R}^d werden als Spaltenvektoren notiert, und für $x \in \mathbb{R}^d$ bezeichne x' den zugehörigen transponierten Vektor. Die Komponenten von $x \in \mathbb{R}^d$ werden, sofern nicht anders angegeben, mit Indizes, also $x' = (x_1, \dots, x_d)$, geschrieben. Weiter notieren wir für die $d \times d$ -Einheitsmatrix I_d und beim Nullvektor $0 \in \mathbb{R}^d$ verzichten wir, falls keine Verwechslungsgefahr besteht, auf den Dimensionsindex. Alle im Folgenden betrachteten Zufallsvariablen sind auf einem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ definiert. Der \mathbb{R}^d sei stets mit der Borel- σ -Algebra \mathcal{B}^d versehen. Des Weiteren wird mit $\mathcal{N}_d(\mu, \Sigma)$ die d -variate Normalverteilung mit Erwartungswertvektor $\mu \in \mathbb{R}^d$ und Kovarianzmatrix Σ , mit \mathcal{N}_d die Klasse der nichtdegenerierten d -dimensionalen Normalverteilungen und mit $\mathcal{U}(\mathcal{S}^{d-1})$ die Gleichverteilung auf der Einheitssphäre bezeichnet. Für die univariate Normalverteilung wird auf den Dimensionsindex verzichtet. Analog schreiben wir χ_t^2 für die Chi-Quadrat Verteilung mit t Freiheitsgraden, $t \in \mathbb{N}$. Für die Gleichheit in Verteilung zweier Zufallselemente X, Y wird $X \stackrel{d}{=} Y$ geschrieben; wie üblich schreiben wir $\xrightarrow{\mathcal{D}}$, $\xrightarrow{\mathbb{P}}$ und $\xrightarrow{\text{f.s.}}$ für Verteilungs-, stochastische und fast sichere Konvergenz. Ein stochastisch beschränkter Term wird mit $O_{\mathbb{P}}(1)$ bezeichnet, analog steht $o_{\mathbb{P}}(1)$ für einen stochastisch gegen 0 strebenden Term. Für den Raum der reellwertigen orthogonalen $d \times d$ -Matrizen wird $\mathcal{O}(d)$ geschrieben. Der Raum der reellwertigen $d \times d$ -Matrizen $\mathbb{R}^{d \times d}$ sei mit der Schur-Norm

$\|\cdot\|_S$ versehen. Unter einem standardisierten d -dimensionalen Zufallsvektor X verstehen wir, dass er zentriert ist und die Einheitsmatrix als Kovarianzmatrix hat, dass also $E(X) = 0$ und $E(XX') = I_d$ gilt.

Kapitel 2

Die Teststatistik von Cox und Small

Es sei X ein d -dimensionaler Zufallsvektor mit Verteilung \mathbb{P}^X , welche als absolut stetig bezüglich des d -dimensionalen Borel-Lebesgue Maßes vorausgesetzt sei. Ein klassisches Problem der multivariaten Statistik besteht darin, anhand von n unabhängigen und identisch verteilten Realisierungen X_1, \dots, X_n von X die Hypothese

$$H_0 : \mathbb{P}^X \in \mathcal{N}_d,$$

dass X eine nichtausgeartete d -dimensionale Normalverteilung besitzt, gegen die allgemeine Alternative

$$H_1 : \mathbb{P}^X \notin \mathcal{N}_d$$

zu testen.

In ihrem Artikel von 1978 stellen Cox und Small in [7] eine Statistik für dieses Testproblem vor und motivieren diese wie folgt: Bekanntlich ist eine d -variate Normalverteilung $\mathcal{N}_d(\mu, \Sigma)$ durch den Erwartungswertvektor $\mu \in \mathbb{R}^d$ und die Kovarianzmatrix $\Sigma \in \mathbb{R}^{d \times d}$ eindeutig bestimmt. Da die Kovarianzmatrix lediglich den linearen Zusammenhang zwischen den einzelnen Komponenten eines d -variater normalverteilter Zufallsvektors $Z \sim \mathcal{N}_d(\mu, \Sigma)$ widerspiegelt, schlagen Cox und Small vor, das Auftreten von nichtlinearen Abhängigkeiten in der zugrunde liegenden Stichprobe zu untersuchen. Dabei betrachten sie dasjenige Paar linearer Kombinationen der ursprünglichen Daten, welches die größtmögliche Krümmung in der Regression aufweist.

Zur Konstruktion einer geeigneten Teststatistik, die die oben genannten Überlegungen abdeckt, bezeichne für beliebige $a, b \in \mathcal{S}^{d-1}$ die Größe $\gamma(a, b)$ den Kleinst-Quadrat-Regressionskoeffizienten von $a'X$ und $(b'X)^2$, der sich als Lösung des Optimierungsproblems

$$\min_{\delta, \varepsilon, \gamma} E(a'X - \delta - \varepsilon b'X - \gamma(b'X)^2)^2 \quad (2.1)$$

ergibt. Ohne Beschränkung der Allgemeinheit wird der Zufallsvektor X als standardisiert vorausgesetzt. Dann gilt $E(a'X) = 0$ und $E(a'X)^2 = 1$ für alle $a \in \mathcal{S}^{d-1}$. Um das Optimierungsproblem in (2.1) zu lösen, definieren wir für $a, b \in \mathcal{S}^{d-1}$ und $(\delta, \varepsilon, \gamma) \in \mathbb{R}^3$ durch

$$\tau_{a,b}(\delta, \varepsilon, \gamma) := 1 - 2\varepsilon E(a'Xb'X) - 2\gamma E(a'X(b'X)^2) + \delta^2 + 2\delta\varepsilon + \varepsilon^2 + 2\varepsilon\gamma E(b'X)^3 + \gamma^2 E(b'X)^4$$

eine Abbildung $\tau_{a,b} : \mathbb{R}^3 \rightarrow \mathbb{R}$. Das Polynom $\tau_{a,b}$ muss auf lokale Extremstellen untersucht werden. Es gilt für die partiellen Ableitungen

$$\begin{aligned}\frac{\partial \tau_{a,b}}{\partial \delta}(\delta, \varepsilon, \gamma) &= 2\delta + 2\gamma \\ \frac{\partial \tau_{a,b}}{\partial \varepsilon}(\delta, \varepsilon, \gamma) &= -2E(a'Xb'X) + 2\varepsilon + 2\gamma E(b'X)^3 \\ \frac{\partial \tau_{a,b}}{\partial \gamma}(\delta, \varepsilon, \gamma) &= -2E(a'X(b'X)^2) + 2\delta + 2\varepsilon E(b'X)^3 + 2\gamma E(b'X)^4.\end{aligned}$$

Die Bestimmung der stationären Punkte von $\tau_{a,b}$ ergibt für $\gamma = \gamma(a, b)$ den Wert

$$\gamma(a, b) = \frac{E(a'X(b'X)^2) - E(a'Xb'X)E(b'X)^3}{E(b'X)^4 - 1 - (E(b'X)^3)^2}.$$

Die Hessematrix des Polynoms $\tau_{a,b}$ lautet für beliebiges $(\delta, \varepsilon, \gamma) \in \mathbb{R}^3$

$$H(\tau_{a,b}, (\delta, \varepsilon, \gamma)) = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 2E(b'X)^3 \\ 2 & 2E(b'X)^3 & 2E(b'X)^4 \end{pmatrix}.$$

Sie ist genau dann positiv definit, wenn

$$\det(H(\tau_{a,b}, (\delta, \varepsilon, \gamma))) = 8 \left(E(b'X)^4 - 1 - (E(b'X)^3)^2 \right) > 0$$

gilt. Dies ist wegen der Cauchy-Schwarz-Ungleichung und

$$\begin{aligned}E(b'X)^4 - 1 - (E(b'X)^3)^2 &= \text{Var}(b'X)^2 - (\text{Cov}(b'X, (b'X)^2))^2 \\ &> \text{Var}(b'X)^2 - \text{Var}(b'X)\text{Var}(b'X)^2 = 0\end{aligned}$$

stets erfüllt. Dabei verwenden wir hier die scharfe Cauchy-Schwarz-Ungleichung, da $(b'X)$ und $(b'X)^2$ offensichtlich nicht linear abhängig sind. Definiert man

$$\eta(a, b) := \gamma(a, b) \left(E(b'X)^4 - 1 - (E(b'X)^3)^2 \right)^{\frac{1}{2}}$$

so kann die Größe $\eta^2(a, b)$ als Anteil des quadratischen Faktors der Kleinste-Quadrate-Regression von $a'X$ auf $b'X$ und $(b'X)^2$ an der Varianz von $a'X$ interpretiert werden. Um den maximalen Anteil bestimmen zu können, muss $\eta^2(a, b)$ unter der Nebenbedingung $\|a\| = 1$ bezüglich a maximiert werden. Der Nenner von $\eta^2(a, b)$ hängt nur von b ab und wird im Folgenden nicht aufgeführt. Zur Vereinfachung sei

$$\begin{aligned}\xi(a, b) &:= E(a'X(b'X)^2) - E(a'Xb'X)E(b'X)^3 \\ &= \sum_{j_1, j_2, j_3=1}^d a_{j_1} b_{j_2} b_{j_3} E(X_{j_1} X_{j_2} X_{j_3}) \\ &\quad - \sum_{j_1, j_2=1}^d a_{j_1} b_{j_2} E(X_{j_1} X_{j_2}) \sum_{j_1, j_2, j_3=1}^d b_{j_1} b_{j_2} b_{j_3} E(X_{j_1} X_{j_2} X_{j_3})\end{aligned}$$

gesetzt. Zur Bestimmung der stationären Punkte der Lagrangefunktion

$$L(a, b, \lambda) = \xi(a, b) - \frac{1}{2}\lambda(\|a\|^2 - 1), \quad a, b \in \mathcal{S}^{d-1}, \lambda \in \mathbb{R},$$

muss für $j = 1, \dots, d$

$$\frac{\partial L}{\partial a_j}(a, b, \lambda) = \sum_{j_1, j_2=1}^d b_{j_1} b_{j_2} E(X_j X_{j_1} X_{j_2}) - b_j \sum_{j_1, j_2, j_3=1}^d b_{j_1} b_{j_2} b_{j_3} E(X_{j_1} X_{j_2} X_{j_3}) - \lambda a_j = 0$$

gelten. Multiplikation der j -ten Gleichung mit b_j und Summation über alle $j = 1, \dots, d$ liefert die erste Bedingung

$$\lambda \sum_{j=1}^d a_j b_j = 0,$$

und Multiplikation der j -ten Gleichung mit a_j und Summation über alle $j = 1, \dots, d$ ergibt die zweite Bedingung

$$\xi(a, b) - \lambda = 0.$$

Da offensichtlich für beliebige $a, b \in \mathcal{S}^{d-1}$ die Beziehung $\xi(a, b) \neq 0$ erfüllt ist, kann nach der zweiten Bedingung λ nicht gleich 0 sein. Daraus und aus der ersten Bedingung folgt, dass $a'b = 0$ ist und demnach a orthogonal zu b sein muss. Weiter gilt für $j = 1, \dots, d$

$$a_j = (\xi(a, b))^{-1} \left[\sum_{j_1, j_2=1}^d b_{j_1} b_{j_2} E(X_j X_{j_1} X_{j_2}) - b_j \sum_{j_1, j_2, j_3=1}^d b_{j_1} b_{j_2} b_{j_3} E(X_{j_1} X_{j_2} X_{j_3}) \right]. \quad (2.2)$$

Für $\max_{a \in \mathcal{S}^{d-1}} \xi^2(a, b)$ gilt mit (2.2) und $a'b = 0$

$$\begin{aligned} \max_{a \in \mathcal{S}^{d-1}} \xi^2(a, b) &= \sum_{j, j_1, j_2, j_3, j_4=1}^d b_{j_1} b_{j_2} b_{j_3} b_{j_4} E(X_j X_{j_1} X_{j_2}) E(X_j X_{j_3} X_{j_4}) \\ &\quad - \left(\sum_{j_1, j_2, j_3=1}^d b_{j_1} b_{j_2} b_{j_3} E(X_{j_1} X_{j_2} X_{j_3}) \right)^2 \\ &= \|E(X(b'X)^2)\|^2 - (E(b'X)^3)^2. \end{aligned}$$

Damit ergibt sich für $\eta^2(b) := \max_{a \in \mathcal{S}^{d-1}} \eta^2(a, b)$ die Beziehung

$$\eta^2(b) = \frac{\|E(X(b'X)^2)\|^2 - (E(b'X)^3)^2}{E(b'X)^4 - 1 - (E(b'X)^3)^2}.$$

Das Funktional, auf dem der Test von Cox and Small aufbaut, ist

$$\mathcal{T}(\mathbb{P}^X) = \max_{b \in \mathcal{S}^{d-1}} \eta^2(b). \quad (2.3)$$

Eine Vereinfachung dieser Darstellung des Funktionals ist außer unter der Annahme spezieller Verteilungen für den Zufallsvektor X nicht mehr möglich.

Im Folgenden verwenden wir wie in Cox und Small [7] oder in Henze [16] das oben angegebene Funktional, um eine Teststatistik zu konstruieren. Auf Grund der Annahme der Standardisierung des Zufallsvektors X werden für die Teststatistik die standardisierten Daten

$$Y_{n,j} := S_n^{-\frac{1}{2}}(X_j - \bar{X}_n), \quad j = 1, \dots, n, \quad (2.4)$$

verwendet. Dabei bezeichnen $\bar{X}_n := \frac{1}{n} \sum_{j=1}^n X_j$ das empirische Mittel und

$$S_n := \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)'$$

die empirische Kovarianzmatrix von X_1, \dots, X_n , sowie $S_n^{-\frac{1}{2}}$ die symmetrische positiv definite Quadratwurzel von S_n^{-1} . Zur Sicherung der \mathbb{P} -fast sicheren Regularität der empirischen Kovarianzmatrix S_n (siehe [11]) sei im Folgenden der Stichprobenumfang n echt größer als die Dimension d , und unter der Verteilung \mathbb{P}^X habe jede $(d-1)$ -dimensionale Hyperebene des \mathbb{R}^d das Maß 0.

Die auf dem Funktional \mathcal{T} aus (2.3) aufbauende Teststatistik T_n ist

$$T_n := \max_{b \in \mathcal{S}^{d-1}} \eta_n^2(b), \quad (2.5)$$

mit

$$\eta_n^2(b) := \frac{\left\| n^{-1} \sum_{j=1}^n Y_{n,j} (b' Y_{n,j})^2 \right\|^2 - \left(n^{-1} \sum_{j=1}^n (b' Y_{n,j})^3 \right)^2}{n^{-1} \sum_{j=1}^n (b' Y_{n,j})^4 - 1 - \left(n^{-1} \sum_{j=1}^n (b' Y_{n,j})^3 \right)^2}. \quad (2.6)$$

Unter der Hypothese ist das Funktional \mathcal{T} gleich 0, so dass eine Ablehnung der Hypothese für große Werte von T_n erfolgt.

Im folgenden Unterabschnitt zeigen wir die affine Invarianz von T_n .

2.1 Affine Invarianz der Teststatistik T_n

In diesem Abschnitt zeigen wir die affine Invarianz der in (2.5) vorgestellten Teststatistik T_n . Da T_n standardisierte Daten verwendet, betrachten wir zunächst den Zusammenhang zwischen den affin linear transformierten standardisierten Daten und den ursprünglichen standardisierten Daten $Y_{n,j}$ für $j = 1, \dots, n$. Im Folgenden seien A eine invertierbare $d \times d$ -Matrix und $\beta \in \mathbb{R}^d$. Setzen wir die affin transformierte Stichprobe $AX_1 + \beta, \dots, AX_n + \beta$ in das empirische Mittel und in die empirische Kovarianzmatrix ein, erhalten wir die transformierten Analoga, also

$$\tilde{\bar{X}}_n := \frac{1}{n} \sum_{j=1}^n \{AX_j + \beta\} = A\bar{X}_n + \beta$$

und

$$\tilde{S}_n := \frac{1}{n} \sum_{j=1}^n (AX_j + \beta - A\bar{X}_n - \beta)(AX_j + \beta - A\bar{X}_n - \beta)' = AS_n A'.$$

Definieren wir weiter

$$\tilde{Y}_{n,j} := \tilde{S}_n^{-\frac{1}{2}} (AX_j + \beta - \tilde{X}_n), \quad j = 1, \dots, n, \quad (2.7)$$

so zeigt das folgende Lemma den Zusammenhang zwischen den transformierten, standardisierten Daten $\tilde{Y}_{n,j}$ und den ursprünglichen standardisierten Daten $Y_{n,j}$.

Lemma 2.1.1

Es seien A eine invertierbare $d \times d$ -Matrix und $\beta \in \mathbb{R}^d$. Für die wie in (2.7) transformierten Daten gilt: Es existiert eine orthogonale Matrix $\tilde{A} \in \mathcal{O}(d)$ mit der Eigenschaft

$$\tilde{Y}_{n,j} = \tilde{A} Y_{n,j}, \quad j = 1, \dots, n.$$

Die Matrix \tilde{A} ist gegeben durch

$$\tilde{A} := (AS_n A')^{-\frac{1}{2}} AS_n^{\frac{1}{2}}.$$

Beweis:

Es gilt für $j = 1, \dots, n$

$$\begin{aligned} \tilde{Y}_{n,j} &= \tilde{S}_n^{-\frac{1}{2}} (AX_j + \beta - \tilde{X}_n) \\ &= (AS_n A')^{-\frac{1}{2}} (AX_j - A\bar{X}_n) \\ &= (AS_n A')^{-\frac{1}{2}} AS_n^{\frac{1}{2}} S_n^{-\frac{1}{2}} (X_j - \bar{X}_n) \\ &= (AS_n A')^{-\frac{1}{2}} AS_n^{\frac{1}{2}} Y_{n,j}. \end{aligned}$$

Weiter ist die Matrix $\tilde{A} := (AS_n A')^{-\frac{1}{2}} AS_n^{\frac{1}{2}}$ orthogonal, da $\tilde{A}' \tilde{A} = I_d$ gilt. \square

Wie das vorherige Lemma zeigt, bewirkt eine affine Transformation der Daten X_1, \dots, X_n lediglich eine Drehung der standardisierten Daten $Y_{n,1}, \dots, Y_{n,n}$ um den Ursprung. Wegen $\tilde{A} \in \mathcal{O}(d)$ liefert der eben gezeigte Zusammenhang mit der Festsetzung $\tilde{A}' b =: \tilde{b} \in \mathcal{S}^{d-1}$

$$\begin{aligned} \left\| \frac{1}{n} \sum_{j=1}^n \tilde{Y}_{n,j} (b' \tilde{Y}_{n,j})^2 \right\|^2 &= \left\| \frac{1}{n} \sum_{j=1}^n \tilde{A} Y_{n,j} (b' \tilde{A} Y_{n,j})^2 \right\|^2 \\ &= \left(\tilde{A} \frac{1}{n} \sum_{j=1}^n Y_{n,j} (\tilde{b}' Y_{n,j})^2 \right)' \left(\tilde{A} \frac{1}{n} \sum_{j=1}^n Y_{n,j} (\tilde{b}' Y_{n,j})^2 \right) \\ &= \left\| \frac{1}{n} \sum_{j=1}^n Y_{n,j} (\tilde{b}' Y_{n,j})^2 \right\|^2. \end{aligned}$$

Weiter gilt für $r \in \mathbb{N}$ nach einfacher Rechnung

$$(b' \tilde{Y}_{n,j})^r = (\tilde{b}' Y_{n,j})^r.$$

Setzt man dies alles zusammen, ergibt sich nach affiner Transformationen der Daten für die Größe $\eta_n^2(b)$ die Beziehung

$$\begin{aligned}\eta_n^2(b) &= \frac{\left\| n^{-1} \sum_{j=1}^n \tilde{Y}_{n,j} (b' \tilde{Y}_{n,j})^2 \right\|^2 - \left(n^{-1} \sum_{j=1}^n (b' \tilde{Y}_{n,j})^3 \right)^2}{n^{-1} \sum_{j=1}^n (b' \tilde{Y}_{n,j})^4 - 1 - \left(n^{-1} \sum_{j=1}^n (b' \tilde{Y}_{n,j})^3 \right)^2} \\ &= \frac{\left\| n^{-1} \sum_{j=1}^n Y_{n,j} (\tilde{b}' Y_{n,j})^2 \right\|^2 - \left(n^{-1} \sum_{j=1}^n (\tilde{b}' Y_{n,j})^3 \right)^2}{n^{-1} \sum_{j=1}^n (\tilde{b}' Y_{n,j})^4 - 1 - \left(n^{-1} \sum_{j=1}^n (\tilde{b}' Y_{n,j})^3 \right)^2} \\ &= \eta_n^2(\tilde{b}).\end{aligned}$$

Es folgt die affine Invarianz der Teststatistik T_n aus der Invarianz der Maximalwertbildung von $\eta_n(\cdot)$ über der d -dimensionalen Einheitskugel \mathcal{S}^{d-1} gegenüber Drehungen, also

$$\max_{b \in \mathcal{S}^{d-1}} \eta_n^2(b) = \max_{b \in \mathcal{S}^{d-1}} \eta_n^2(Db),$$

falls $D \in \mathcal{O}(d)$. Dabei ist zu beachten, dass $\eta_n^2 \in C(\mathcal{S}^{d-1}, \mathbb{R})$ gilt und \mathcal{S}^{d-1} ein kompakter Träger ist, so dass das Maximum also insbesondere endlich ist.

2.2 Analyse der Teststatistik T_n

Bevor wir zu konkreten Aussagen zum Konvergenzverhalten und zur Konsistenz eines mit T_n aus (2.5) operierenden Tests auf multivariate Normalverteilung kommen, wollen wir die Teststatistik analysieren und die Methodik vorstellen, die in den folgenden Kapiteln Verwendung findet. Formal besteht die Teststatistik T_n aus zwei Komponenten,

1. der für jedes $b \in \mathcal{S}^{d-1}$ definierten Größe $\eta_n^2(b)$ und
2. dem Funktional $C(\mathcal{S}^{d-1}, \mathbb{R}) \rightarrow \mathbb{R}$, $g \mapsto \max_{b \in \mathcal{S}^{d-1}} g(b)$.

Wir bezeichnen mit $C(\mathcal{S}^{d-1}, \mathbb{R})$ den Banachraum der stetigen, reellwertigen Funktionen mit Definitionsbereich \mathcal{S}^{d-1} . Für topologische Eigenschaften des Raumes $C(\mathcal{S}^{d-1}, \mathbb{R})$ siehe Anhang B. Da es sich bei dem Maximumfunktional um ein stetiges Funktional auf dem Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$ handelt, ist der Abbildungssatz (siehe [21], S.188) anwendbar. Um Aussagen bezüglich der schwachen Konvergenz der Teststatistik T_n treffen zu können, betrachten wir die Größe $\eta_n^2(\cdot)$ als Zufallselement im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$. Unter gewissen noch zu präzisierenden strukturellen Annahmen an die zugrunde liegende Verteilung gilt folgender Sachverhalt.

Proposition 2.2.1

Es existiert ein zentriertes Zufallselement η^2 in $C(\mathcal{S}^{d-1}, \mathbb{R})$ mit Kovarianzkern $\sigma_\eta(b, c)$, $b, c \in \mathcal{S}^{d-1}$ für das

$$n\eta_n^2(\cdot) \xrightarrow{\mathcal{D}} \eta^2(\cdot)$$

für $n \rightarrow \infty$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$ gilt.

Aus dieser Proposition ergibt sich mit Hilfe des Abbildungssatzes und der Stetigkeit des Maximumfunktionals die

Folgerung 2.2.2

Für die Teststatistik T_n aus (2.5) gilt

$$nT_n = n \max_{b \in \mathcal{S}^{d-1}} \eta_n^2(b) \xrightarrow{\mathcal{D}} \max_{b \in \mathcal{S}^{d-1}} \eta^2(b).$$

Es bleibt das Zufallselement η_n^2 in Anteile aufzuteilen, über die man in $C(\mathcal{S}^{d-1}, \mathbb{R})$ Aussagen machen kann.

1. Zunächst betrachten wir den Zähler der Größe η_n^2 aus (2.6) und setzen für $b \in \mathcal{S}^{d-1}$

$$V_n(b) := \left\| \frac{1}{n} \sum_{j=1}^n Y_{n,j} (b' Y_{n,j})^2 \right\|^2 - \left(\frac{1}{n} \sum_{j=1}^n (b' Y_{n,j})^3 \right)^2. \quad (2.8)$$

Durch elementare Umformungen (siehe Anhang A) erhalten wir die Darstellung

$$V_n(b) = \frac{1}{n^2} \sum_{j_1, j_2=1}^n h_b(Y_{n,j_1}, Y_{n,j_2}), \quad b \in \mathcal{S}^{d-1}, \quad (2.9)$$

mit dem vom Vektor $b \in \mathcal{S}^{d-1}$ abhängigen symmetrischen Kern $h_b : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, welcher durch

$$h_b(x, y) := (b'x)^2 (b'y)^2 (x'y - b'xb'y), \quad x, y \in \mathbb{R}^d, \quad b \in \mathcal{S}^{d-1}, \quad (2.10)$$

definiert ist. Es handelt sich demnach bei $V_n(b)$ um eine V -Statistik mit Kern h_b .

Im Folgenden bezeichne $\varphi_k(x)$ die Projektion des Vektors $x \in \mathbb{R}^d$ auf dessen k -te Komponente. Des Weiteren definieren wir für $k \in \{1, \dots, d\}$ die Abbildung $f_k : \mathbb{R}^d \times \mathcal{S}^{d-1} \rightarrow \mathbb{R}$ durch

$$(x, b) \mapsto f_k(x, b) := (b'x)^2 \varphi_k(x) \quad (2.11)$$

und setzen

$$C_{n,k}(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n f_k(Y_{n,j}, b) \quad \text{und} \quad D_n(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n (b'Y_{n,j})^3.$$

Mit diesen Bezeichnungen gilt dann für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned}
nV_n(b) &= \frac{1}{n} \sum_{j_1, j_2=1}^n h_b(Y_{n, j_1}, Y_{n, j_2}) \\
&= \frac{1}{n} \sum_{j_1, j_2=1}^n (b'Y_{n, j_1})^2 (b'Y_{n, j_2})^2 (Y_{n, j_1}' Y_{n, j_2} - b'Y_{n, j_1} b'Y_{n, j_2}) \\
&= \frac{1}{n} \sum_{j_1, j_2=1}^n (b'Y_{n, j_1})^2 (b'Y_{n, j_2})^2 \left(\sum_{k=1}^d \varphi_k(Y_{n, j_1}) \varphi_k(Y_{n, j_2}) - b'Y_{n, j_1} b'Y_{n, j_2} \right) \\
&= \frac{1}{n} \sum_{j_1, j_2=1}^n \left(\sum_{k=1}^d f_k(Y_{n, j_1}, b) f_k(Y_{n, j_2}, b) - (b'Y_{n, j_1})^3 (b'Y_{n, j_2})^3 \right) \\
&= \sum_{k=1}^d \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n f_k(Y_{n, j}, b) \right)^2 - \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n (b'Y_{n, j})^3 \right)^2
\end{aligned}$$

und somit

$$nV_n(b) = \sum_{k=1}^d (C_{n, k}(b))^2 - (D_n(b))^2, \quad b \in \mathcal{S}^{d-1}. \quad (2.12)$$

Es sei weiter durch

$$W_n(b) := \begin{pmatrix} C_{n, 1}(b) \\ C_{n, 2}(b) \\ \vdots \\ C_{n, d}(b) \\ D_n(b) \end{pmatrix}, \quad b \in \mathcal{S}^{d-1}, \quad (2.13)$$

eine Familie $(W_n(b))_{n>d}$ von $(d+1)$ -dimensionalen Zufallsvektoren definiert. Der $(d+1)$ -dimensionale Zufallsvektor $W_n(b)$ steht in funktionalem Zusammenhang mit dem in (2.12) dargestellten Zufallselement $nV_n(b)$. Zur Erläuterung betrachten wir die symmetrische $(d+1) \times (d+1)$ -Matrix

$$B := \begin{pmatrix} I_d & 0 \\ 0 & -1 \end{pmatrix},$$

sowie die stetige Abbildung

$$\Psi_B : \mathbb{R}^{d+1} \rightarrow \mathbb{R}, \quad x \mapsto \Psi_B(x) := x' B x. \quad (2.14)$$

Mit Hilfe der Abbildung Ψ_B gilt für $b \in \mathcal{S}^{d-1}$ die algebraische Gleichheit

$$nV_n(b) = \Psi_B(W_n(b)).$$

Um die schwache Konvergenz von $nV_n(\cdot)$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$ anzugeben, reicht es unter Berücksichtigung des Abbildungssatzes aus, die schwache Konvergenz des multivariaten Multiparameter-Prozesses $W_n(\cdot)$ in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ zu untersuchen.

2. Es bleibt der Nenner der Größe η_n^2 aus (2.6) zu untersuchen. Dabei setzen wir

$$\mathcal{A}_n(b) := n^{-1} \sum_{j=1}^n (b'Y_{n,j})^4 - 1 - \left(n^{-1} \sum_{j=1}^n (b'Y_{n,j})^3 \right)^2.$$

Strukturell werden wir in diesem Fall versuchen, mit dem Gesetz großer Zahlen auf die stochastische Konvergenz des Zufallselements $\mathcal{A}_n(\cdot)$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$ gegen eine Konstante zu schließen.

Die Analyse der Teststatistik schließen wir mit einer Bemerkung zur Implementation der Teststatistik T_n in einer Programmiersprache ab: Das Maximumfunktional aus 2. auf der Oberfläche der Einheitssphäre lässt sich nicht effizient berechnen. Hier müssen alternative Möglichkeiten zur Approximation gefunden werden, siehe Abschnitt 8.1.

Kapitel 3

Konvergenz ohne geschätzte Parameter unter sphärischer Symmetrie

In diesem Kapitel werden Konvergenzaussagen für die Teststatistik T_n aus (2.5) unter der Einschränkung nicht geschätzter Parameter gemacht. Dabei betrachten wir bei Berechnung der Größe η_n^2 aus (2.6) anstelle der empirisch standardisierten Daten $Y_{n,j}$ vereinfachend die Daten $X_j, j \in \{1, \dots, n\}$. Im Folgenden seien X_1, \dots, X_n unabhängig identisch sphärisch-symmetrisch verteilte d -dimensionale Zufallsvektoren mit der Eigenschaft $E\|X_1\|^6 < \infty$. Auf Grund der affinen Invarianz der Statistik T_n von Cox und Small können wir X_1 ohne Beschränkung der Allgemeinheit als standardisiert voraussetzen, d.h. es gelte $EX_1 = 0$ und $\text{Cov}(X_1, X_1) = EX_1X_1' = I_d$. Des Weiteren gelte $n > d$. Unter diesen Voraussetzungen gilt

$$E\|X_1\|^2 = d.$$

Abkürzend schreiben wir

$$m_l = E\|X_1\|^l, \quad l \geq 1,$$

für die l -ten Momente der Radienverteilung von X_1 . Im Folgenden soll die schwache Konvergenz der Zufallselemente $nV_n(\cdot)$ aus (2.12) im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$ gegen ein Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$ gezeigt werden. Hierzu betrachten wir für $k \in \{1, \dots, d\}$ die stochastischen Prozesse $C_{n,k}(\cdot)$ und $D_n(\cdot)$ als Zufallselemente auf dem Raum $C(\mathcal{S}^{d-1}, \mathbb{R})$.

Lemma 3.1.1

Es existiert ein Gaußprozess $C_k(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitssphäre mit stetigen Pfaden, Erwartungswertvektor 0 und dem Kovarianzkern

$$\varrho_k(b, c) = \frac{m_6}{d(d+2)(d+4)} [2(b'c)^2 + 8b_k c_k (b'c) + 2(c_k^2 + b_k^2) + 1] \quad (3.1)$$

($b, c \in \mathcal{S}^{d-1}$). Betrachtet man $C_k(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$, so gilt

$$C_{n,k}(\cdot) \xrightarrow{\mathcal{D}} C_k(\cdot).$$

Beweis:

Um die Kovarianzstruktur des Zufallselements C_k zu bestimmen, reicht es, die Kovarianzstruktur von $C_{1,k}$ zu berechnen. Im Folgenden wird auf den Index des Zufallsvektors X_1 verzichtet. Mit $X \stackrel{d}{=} RU$, wobei $U \sim \mathcal{U}(\mathcal{S}^{d-1})$ und R eine positive reellwertige von U stochastisch unabhängige Zufallsvariable ist, gilt für $b, c \in \mathcal{S}^{d-1}$

$$\begin{aligned} \text{Cov}(f_k(X, b), f_k(X, c)) &= E[f_k(X, b)f_k(X, c)] \\ &= m_6 E[(b'U)^2(c'U)^2\varphi_k^2(U)]. \end{aligned}$$

Ferner gilt für $b, c, u \in \mathbb{R}^d$

$$\begin{aligned} (b'u)^2(c'u)^2 &= \sum_{(i,j,l,m) \in \{1, \dots, d\}^4} b_i b_j c_l c_m u_i u_j u_l u_m \\ &= \sum_{j=1}^d b_j^2 c_j^2 u_j^4 + \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i^2 c_j^2 u_i^2 u_j^2 + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i b_j c_i c_j u_i^2 u_j^2 + \mathcal{W}_u. \end{aligned}$$

Dabei stellt \mathcal{W}_u eine Summe dar, bei der in jedem Summanden mindestens eine ungerade Potenz einer Komponente von u vorkommt. Aus Symmetriegründen gilt somit

$$\int_{\mathcal{S}^{d-1}} \mathcal{W}_u u_k^2 d\mathcal{S}^{d-1}(u) = 0.$$

Mit der Abkürzung $\zeta_{d,2}$ aus Anhang A.1 und den Werten aus Tabelle A.1 gilt

$$\begin{aligned} E[(b'U)^2(c'U)^2\varphi_k^2(U)] &= \frac{1}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} (b'u)^2(c'u)^2 u_k^2 d\mathcal{S}^{d-1}(u) \\ &= \frac{1}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} \left[\sum_{i=1}^d b_i^2 c_i^2 u_i^4 + \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i^2 c_j^2 u_i^2 u_j^2 \right. \\ &\quad \left. + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i b_j c_i c_j u_i^2 u_j^2 + \mathcal{W}_u \right] u_k^2 d\mathcal{S}^{d-1}(u) \\ &= \frac{1}{|\mathcal{S}^{d-1}|} \left[\sum_{i=1}^d b_i^2 c_i^2 \int_{\mathcal{S}^{d-1}} u_i^4 u_k^2 d\mathcal{S}^{d-1}(u) \right. \\ &\quad + \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i^2 c_j^2 \int_{\mathcal{S}^{d-1}} u_i^2 u_j^2 u_k^2 d\mathcal{S}^{d-1}(u) \\ &\quad \left. + 2 \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i b_j c_i c_j \int_{\mathcal{S}^{d-1}} u_i^2 u_j^2 u_k^2 d\mathcal{S}^{d-1}(u) \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{|\mathcal{S}^{d-1}|} \left[3\zeta_{d,2} \sum_{i=1}^d b_i^2 c_i^2 + 12\zeta_{d,2} b_k^2 c_k^2 \right. \\
&\quad + \zeta_{d,2} \sum_{\substack{i,j=1 \\ i \neq j}}^d b_i^2 c_j^2 + 2\zeta_{d,2}(b_k^2 + c_k^2) - 4\zeta_{d,2} b_k^2 c_k^2 \\
&\quad \left. + 2\zeta_{d,2} \sum_{\substack{i,j=1 \\ i \neq j \neq k}}^d b_i b_j c_i c_j + 12\zeta_{d,2} b_k c_k \sum_{\substack{j=1 \\ j \neq k}}^d b_j c_j \right] \\
&= \frac{1}{d(d+2)(d+4)} [2(b'c)^2 + 8b_k c_k (b'c) + 2(b_k^2 + c_k^2) + 1].
\end{aligned}$$

Daraus folgt die behauptete Kovarianzstruktur. Um zu zeigen, dass es einen Gaußprozess $C_k(\cdot)$ mit Kovarianzkern $\varrho_k(b, c)$ gibt, so dass $C_{n,k}(\cdot) \xrightarrow{\mathcal{D}} C_k(\cdot)$ konvergiert, reicht es nach [2], S. 172 aus, die Gültigkeit der Beziehung

$$\int_0^1 (\log N(\xi))^{\frac{1}{2}} d\xi < \infty \quad (3.2)$$

zu zeigen. Dabei gibt $N(\xi)$ für gegebenes $\xi > 0$ die kleinste positive natürliche Zahl m an, so dass \mathcal{S}^{d-1} durch genau m messbare Teilmengen überdeckt werden kann, wobei diese Teilmengen höchstens den Durchmesser ξ haben dürfen. Bedingung (3.2) ist für den metrischen Raum $(\mathcal{S}^{d-1}, \|\cdot\|)$ erfüllt, siehe Anhang A.4. Dazu gilt für $b, c \in \mathcal{S}^{d-1}$ mit der Cauchy-Schwarzschen Ungleichung und der Dreiecksungleichung

$$\begin{aligned}
|f_k(X, b) - f_k(X, c)| &= |\varphi_k(X) ((b'X)^2 - (c'X)^2)| \\
&= |\varphi_k(X)| |(b'X)^2 - (c'X)^2| \\
&\leq \|X\| |(b'X) - (c'X)| |(b'X) + (c'X)| \\
&\leq 2\|X\|^3 \|b - c\|.
\end{aligned} \quad (3.3)$$

Da $E\|X\|^6 < \infty$ vorausgesetzt wurde, folgt die Behauptung. \square

Lemma 3.1.2

Es existiert ein zentrierter Gaußprozess $D(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitssphäre mit stetigen Pfaden und Kovarianzkern

$$\varrho(b, c) = \frac{m_6}{d(d+2)(d+4)} (6(b'c)^3 + 9(b'c)), \quad b, c \in \mathcal{S}^{d-1}. \quad (3.4)$$

Betrachtet man $D(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$, so gilt

$$D_n(\cdot) \xrightarrow{\mathcal{D}} D(\cdot).$$

Beweis:

Seien $b, c \in \mathcal{S}^{d-1}$. Mit $X \stackrel{d}{=} RU$, wobei $U \sim \mathcal{U}(\mathcal{S}^{d-1})$ und R eine positive, reellwertige und

von U stochastisch unabhängige Zufallsvariable ist, gilt

$$\begin{aligned} \text{Cov}((b'X)^3, (c'X)^3) &= E[(b'X)^3(c'X)^3] \\ &= m_6 E[(b'U)^3(c'U)^3]. \end{aligned}$$

Der Zufallsvektor U ist gleichverteilt auf der Einheitssphäre und somit verteilungsinvariant unter Drehungen. Also gibt es eine $d \times d$ -Matrix $D \in \mathcal{O}(d)$ mit $D'D = I_d$ und $Dc = e_1$ und $D'b =: \bar{b} \in \mathcal{S}^{d-1}$. Dabei bezeichnet e_1 den ersten Einheitsvektor des \mathbb{R}^d . Mit $\|\bar{b}\| = 1$, Symmetrieargumenten und der Linearität des Erwartungswertes folgt

$$\begin{aligned} E[(b'U)^3(c'U)^3] &= E[(\bar{b}'U)^3 U_1^3] \\ &= E\left[\sum_{i,j,l=1}^d \bar{b}_i \bar{b}_j \bar{b}_l U_i U_j U_l U_1^3\right] \\ &= E\left[\bar{b}_1^3 U_1^6 + 3\bar{b}_1 U_1^4 \sum_{j,l=2}^d \bar{b}_j \bar{b}_l U_j U_l\right] \\ &= \bar{b}_1^3 E(U_1^6) + 3\bar{b}_1 E(U_1^4 U_2^2) \sum_{j=2}^d \bar{b}_j^2 \\ &= \bar{b}_1^3 E(U_1^6) + 3\bar{b}_1 E(U_1^4 U_2^2) (1 - \bar{b}_1^2) \\ &= \bar{b}_1^3 (E(U_1^6) - 3E(U_1^4 U_2^2)) + 3\bar{b}_1 E(U_1^4 U_2^2). \end{aligned}$$

Es gilt nach Tabelle A.1

$$E(U_1^6) = \frac{15}{d(d+2)(d+4)}, \quad (3.5)$$

$$E(U_1^4 U_2^2) = \frac{3}{d(d+2)(d+4)}. \quad (3.6)$$

Mit $\bar{b}_1 = b'c$, (3.5) und (3.6) folgt die behauptete Kovarianzstruktur. Für die Existenz des Grenzprozesses verwenden wir wieder das Korollar 7.17 aus [2], S. 172. Die Entropiebedingung ist wie im letzten Beweis erfüllt, und es gilt für $b, c \in \mathcal{S}^{d-1}$ mit der Cauchy-Schwarzschen Ungleichung und der Dreiecksungleichung die Abschätzung

$$\begin{aligned} |(b'X)^3 - (c'X)^3| &= |(b'X)((b'X)^2 - (c'X)^2) + (c'X)^2((b'X) - (c'X))| \\ &\leq |(b'X)| |((b'X)^2 - (c'X)^2)| + (c'X)^2 |(b'X) - (c'X)| \\ &\leq 3\|X\|^3 \|b - c\|. \end{aligned}$$

Da $E\|X\|^6 < \infty$ vorausgesetzt wurde, folgt die Behauptung. \square

Der Kovarianzkern des Grenzprozesses $D(\cdot)$ hängt lediglich von dem Winkel zwischen zwei Elementen von \mathcal{S}^{d-1} ab, so dass man mit Hilfe der Theorie der Kugeloberflächenfunktionen leicht eine alternative Darstellung des Prozesses angeben kann.

Bemerkung 3.1.3

1. Der zentrierte Gaußsche Grenzprozess $D(b)$, $b \in \mathcal{S}^{d-1}$, kann in der Form

$$D(b) = \sqrt{\alpha_{1,d}} \sum_{k=1}^{N(d,3)} \varphi_{3,k}(b) N_{3,k} + \sqrt{\alpha_{2,d}} \sum_{k=1}^{N(d,1)} \varphi_{1,k}(b) N_{1,k}, \quad b \in \mathcal{S}^{d-1}, \quad (3.7)$$

dargestellt werden, wobei

$$\begin{aligned} \alpha_{1,d} &= \frac{72m_6}{d^2(d+2)^2(d+4)^2}, \\ \alpha_{2,d} &= \frac{3(d+5)m_6}{d^2(d+2)^2(d+4)}. \end{aligned}$$

Dabei sind alle in (3.7) auftretenden Zufallsvariablen $N_{3,k}$ ($k \in \{1, \dots, N(d,3)\}$) und $N_{1,k}$ ($k \in \{1, \dots, N(d,1)\}$) unabhängig und je standardnormalverteilt. Die Größe $N(d, j)$ gibt die Dimension von $\mathcal{V}_j(d)$, dem Raum der d -dimensionalen harmonischen Kugelflächenfunktionen vom Grad j an (für weitere Details zu Kugelflächenfunktionen siehe auch Anhang C). Die Funktionen $\varphi_{j,k} \in \mathcal{V}_j(d)$ mit $j \in \{1, 3\}$ und $k \in \{1, \dots, N(d, j)\}$ sind linear unabhängige harmonische Kugelflächenfunktionen vom Grad j , welche orthonormal im Sinne der Gleichverteilung $\mathcal{U}(\mathcal{S}^{d-1})$ sind.

2. Für die Grenzprozesse C_k mit $k \in \{1, \dots, d\}$ ist es schwer, eine alternative stochastische Darstellung anzugeben, da der Kovarianzkern $\varrho_k(b, c)$, $b, c \in \mathcal{S}^{d-1}$ nicht nur von dem Winkel zwischen den Vektoren b und c abhängt.

Im Folgenden definieren wir für $k, l \in \{1, \dots, d\}$ die Abbildungen $\Phi_{k,l}^{(1)} : \mathcal{S}^{d-1} \times \mathcal{S}^{d-1} \rightarrow \mathbb{R}$ und $\Phi_k^{(2)} : \mathcal{S}^{d-1} \times \mathcal{S}^{d-1} \rightarrow \mathbb{R}$ durch

$$\Phi_{k,l}^{(1)}(b, c) := \text{Cov}(C_{1,k}(b), C_{1,l}(c)), \quad b, c \in \mathcal{S}^{d-1}, \quad (3.8)$$

$$\Phi_k^{(2)}(b, c) := \text{Cov}(D_1(b), C_{1,k}(c)), \quad b, c \in \mathcal{S}^{d-1}. \quad (3.9)$$

Weiter sei $\Phi^{(3)} : \mathcal{S}^{d-1} \times \mathcal{S}^{d-1} \rightarrow \mathbb{R}$ durch

$$\Phi^{(3)}(b, c) := \text{Cov}(D_1(b), D_1(c)), \quad b, c \in \mathcal{S}^{d-1}$$

definiert. Man beachte, dass für $k \in \{1, \dots, d\}$ die Abbildungen $\Phi_{k,k}^{(1)}(b, c)$ mit dem Kovarianzkern $\varrho_k(b, c)$ aus (3.1) und die Abbildung $\Phi^{(3)}(b, c)$ mit dem Kovarianzkern $\varrho(b, c)$ aus (3.4) identisch sind. Unter den gemachten Voraussetzungen gilt

$$\Phi_{k,l}^{(1)}(b, c) = \frac{2m_6}{d(d+2)(d+4)} (2(b'c)(b_k c_l + c_k b_l) + (b_k b_l + c_k c_l)), \quad (3.10)$$

$$\Phi_k^{(2)}(b, c) = \frac{3m_6}{d(d+2)(d+4)} (2b_k (b'c)^2 + 2c_k (b'c) + b_k). \quad (3.11)$$

Da die Beweise dieser expliziten Darstellungen etwas technisch sind, wurden sie zugunsten einer besseren Lesbarkeit in den Anhang verlagert, siehe hierzu Anhang A.2.

Theorem 3.1.4

Unter den Voraussetzungen von oben existiert ein zentrierter $(d+1)$ -dimensionaler Gaußprozess $W(b)$, $b \in \mathcal{S}^{d-1}$, auf \mathcal{S}^{d-1} mit stetigen Pfaden und dem Kovarianzmatrixkern

$$\sigma(b, c) := \begin{pmatrix} \Phi_{1,1}^{(1)}(b, c) & \Phi_{1,2}^{(1)}(b, c) & \cdots & \Phi_{1,d}^{(1)}(b, c) & \Phi_1^{(2)}(b, c) \\ \Phi_{2,1}^{(1)}(b, c) & \Phi_{2,2}^{(1)}(b, c) & \cdots & \Phi_{2,d}^{(1)}(b, c) & \Phi_2^{(2)}(b, c) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Phi_{d,1}^{(1)}(b, c) & \Phi_{d,2}^{(1)}(b, c) & \cdots & \Phi_{d,d}^{(1)}(b, c) & \Phi_d^{(2)}(b, c) \\ \Phi_1^{(2)}(c, b) & \Phi_2^{(2)}(c, b) & \cdots & \Phi_d^{(2)}(c, b) & \Phi^{(3)}(b, c) \end{pmatrix}, \quad b, c \in \mathcal{S}^{d-1}. \quad (3.12)$$

Betrachtet man den in (2.13) definierten Prozess $W_n(\cdot)$ als zufälliges Element auf dem Raum $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$, so gilt

$$W_n(\cdot) \xrightarrow{\mathcal{D}} W(\cdot).$$

Beweis:

In Theorem 3.1.1 und in Theorem 3.1.2 wurde die schwache Konvergenz der stochastischen Prozesse $C_{n,k}$ und D_n im Banachraum der stetigen Funktionen $C(\mathcal{S}^{d-1}, \mathbb{R})$ gezeigt. Da Verteilungskonvergenz auf separablen, metrischen Räumen Straffheit impliziert und nach Theorem B.3.1 komponentenweise Straffheit die Straffheit im kartesischen Produkt der Räume sichert, ist die Familie (W_n) der Zufallselemente in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ straff. Zum Nachweis der schwachen Konvergenz in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ müssen die endlichdimensionalen Verteilungen der Familie (W_n) schwach gegen eine bestimmte Grenzverteilung konvergieren. Die oben angegebene Kovarianzstruktur und eine direkte Anwendung des multivariaten zentralen Grenzwertsatzes liefern das Ergebnis des folgenden Lemmas.

Lemma 3.1.5

Es gilt für $r \in \mathbb{N}$ und $\beta_1, \dots, \beta_r \in \mathcal{S}^{d-1}$

$$\begin{pmatrix} W_n(\beta_1) \\ W_n(\beta_2) \\ \vdots \\ W_n(\beta_r) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_{r(d+1)} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \Sigma_{\beta_1, \dots, \beta_r} \right).$$

Dabei ist $\Sigma_{\beta_1, \dots, \beta_r}$ die durch

$$\Sigma_{\beta_1, \dots, \beta_r} := \begin{pmatrix} \sigma(\beta_1, \beta_1) & \sigma(\beta_1, \beta_2) & \cdots & \sigma(\beta_1, \beta_r) \\ \sigma(\beta_2, \beta_1) & \sigma(\beta_2, \beta_2) & \cdots & \sigma(\beta_2, \beta_r) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(\beta_r, \beta_1) & \sigma(\beta_r, \beta_2) & \cdots & \sigma(\beta_r, \beta_r) \end{pmatrix}$$

definierte $r(d+1) \times r(d+1)$ -Matrix, und für $i, j \in \{1, \dots, r\}$ ist $\sigma(\beta_i, \beta_j)$ wie in (3.12) definiert.

Somit konvergiert nach Theorem B.2.2 die Familie (W_n) in $\times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ gegen einen Grenzprozess W , dessen Verteilung durch die Grenzverteilungen der endlichdimensionalen Verteilungen eindeutig bestimmt ist. \square

Wie Abschnitt 2.2 schon angedeutet, gilt mit Hilfe der Abbildung Ψ_B aus (2.14) für $b \in \mathcal{S}^{d-1}$ die algebraische Gleichheit

$$nV_n(b) = \Psi_B(W_n(b)).$$

Auf Grund der Stetigkeit von Ψ_B und der schwachen Konvergenz des stochastischen Prozesses $W_n(\cdot)$ in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ ergibt sich nach Anwendung des Abbildungssatzes die

Folgerung 3.1.6

Für die Folge $(nV_n(\cdot))_{n>d}$ $C(\mathcal{S}^{d-1}, \mathbb{R})$ -wertiger Zufallselemente gilt

$$nV_n(\cdot) \xrightarrow{\mathcal{D}} \Psi_B(W(\cdot)).$$

Wie das folgende Theorem zeigt, lässt sich für festes $b \in \mathcal{S}^{d-1}$ die Verteilung von $\Psi_B(W(b))$ explizit angeben. Der Beweis ist in Anhang A.3 zu finden.

Theorem 3.1.7

Mit den Voraussetzungen und Bezeichnungen von oben gilt für jedes $b \in \mathcal{S}^{d-1}$

$$\Psi_B(W(b)) \stackrel{d}{=} \frac{3m_6}{d(d+2)(d+4)} \sum_{j=1}^{d-1} N_j^2,$$

wobei N_1, \dots, N_{d-1} unabhängige, identisch standardnormalverteilte Zufallsvariablen sind.

Das folgende Lemma zeigt den Zusammenhang zwischen der Größe η_n^2 und dem zentrierten $(d+1)$ -dimensionalen stochastischen Prozess $W_n(\cdot)$. Dabei ist zu beachten, dass wir in diesem Fall vereinfachend in η_n^2 anstatt der skalierten Residuen $Y_{n,j}$ die Daten X_j verwenden.

Lemma 3.1.8

Unter den Voraussetzungen von Theorem 3.1.4 gilt mit $\gamma := \frac{3m_4}{d(d+2)}$ für $n \rightarrow \infty$

$$\sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{\gamma - 1} \Psi_B(W_n(b)) - n\eta_n^2(b) \right| \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Zunächst verwenden wir, dass für $b \in \mathcal{S}^{d-1}$ und $l \in \mathbb{N}$, falls $m_l < \infty$, mit dem Funk-Hecke

Theorem (siehe C.3.1)

$$\begin{aligned}
E(b'X_1)^l &= m_l E(b'U_1)^l \\
&= \frac{m_l}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} (b'u)^l d\mathcal{S}^{d-1} \\
&= \frac{m_l}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} (b'u)^{l-1} P_1(d; b'u) d\mathcal{S}^{d-1} \\
&= \frac{m_l}{|\mathcal{S}^{d-1}|} \lambda P_1(d; \|b\|^2) \\
&= \frac{m_l}{|\mathcal{S}^{d-1}|} \lambda
\end{aligned}$$

gilt, wobei

$$\lambda := |\mathcal{S}^{d-2}| \int_{-1}^1 P_1(d; t) t^{l-1} (1-t^2)^{\frac{d-3}{2}} dt = |\mathcal{S}^{d-2}| \int_{-1}^1 t^l (1-t^2)^{\frac{d-3}{2}} dt$$

eine Konstante und $P_m(d; \cdot)$ das d -dimensionale Legendre-Polynom vom Grad m ist (siehe hierzu C.2). Speziell gilt

$$\begin{aligned}
E(b'X_1)^3 &= 0, \\
E(b'X_1)^4 &= \gamma,
\end{aligned}$$

und insbesondere ist unter den gemachten Voraussetzungen $E(b'X_1)^4 < \infty$. Dann gilt nach dem starken Gesetz großer Zahlen

$$\begin{aligned}
\frac{1}{n} \sum_{j=1}^n (b'X_j)^3 &\xrightarrow{\text{f.s.}} 0, \\
\frac{1}{n} \sum_{j=1}^n (b'X_j)^4 &\xrightarrow{\text{f.s.}} \gamma.
\end{aligned} \tag{3.13}$$

Wir betrachten die Zerlegung

$$\frac{1}{\gamma-1} \Psi_B(W_n(b)) - n\eta_n^2(b) = Z_n^{(1)}(b) + Z_n^{(2)}(b),$$

wobei die stochastischen Prozesse durch

$$\begin{aligned}
Z_n^{(1)}(b) &= \frac{1}{\gamma-1} \left(\Psi_B(W_n(b)) - n \left(\left\| \frac{1}{n} \sum_{j=1}^n X_j (b'X_j)^2 \right\|^2 - \left(\frac{1}{n} \sum_{j=1}^n (b'X_j)^3 \right)^2 \right) \right), \\
Z_n^{(2)}(b) &= n \left(\left\| \frac{1}{n} \sum_{j=1}^n X_j (b'X_j)^2 \right\|^2 - \left(\frac{1}{n} \sum_{j=1}^n (b'X_j)^3 \right)^2 \right) \times \\
&\quad \times \left(\frac{1}{\gamma-1} - \frac{1}{\frac{1}{n} \sum_{j=1}^n (b'X_j)^4 - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b'X_j)^3 \right)^2} \right)
\end{aligned}$$

gegeben sind. Der stochastische Prozess $Z_n^{(1)}(b)$ ist nach den Überlegungen der Gleichung (2.12) fast überall 0. Weiter gilt mit Lemma 3.1.6

$$\sup_{b \in \mathcal{S}^{d-1}} \left| n \left(\left\| \frac{1}{n} \sum_{j=1}^n X_j (b' X_j)^2 \right\|^2 - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 \right)^2 \right) \right| = O_{\mathbb{P}}(1).$$

Sei

$$\mathcal{A}_n(b) := \frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 \right)^2.$$

Damit folgt

$$\begin{aligned} \left| \frac{1}{\gamma - 1} - \frac{1}{\mathcal{A}_n(b)} \right| &= \left| \frac{\mathcal{A}_n(b) - (\gamma - 1)}{(\gamma - 1)\mathcal{A}_n(b)} \right| \\ &= \frac{1}{\sqrt{n}} \frac{1}{|\gamma - 1|} \frac{1}{|\mathcal{A}_n(b)|} \left| \sqrt{n} (\mathcal{A}_n(b) - (\gamma - 1)) \right| \\ &\leq \frac{1}{\sqrt{n}} \frac{1}{|\gamma - 1|} \frac{1}{\inf_{b \in \mathcal{S}^{d-1}} |\mathcal{A}_n(b)|} \sup_{b \in \mathcal{S}^{d-1}} \left| \sqrt{n} (\mathcal{A}_n(b) - (\gamma - 1)) \right|. \end{aligned}$$

Mit der inversen Dreiecksungleichung ergibt sich

$$\begin{aligned} |\mathcal{A}_n(b)| &= |\mathcal{A}_n(b) - (\gamma - 1) + \gamma - 1| \\ &\geq |\gamma - 1| - |\mathcal{A}_n(b) - (\gamma - 1)| \end{aligned}$$

also

$$\inf_{b \in \mathcal{S}^{d-1}} |\mathcal{A}_n(b)| \geq |\gamma - 1| - \frac{1}{\sqrt{n}} \sup_{b \in \mathcal{S}^{d-1}} \left| \sqrt{n} (\mathcal{A}_n(b) - (\gamma - 1)) \right|.$$

Da $\sup_{b \in \mathcal{S}^{d-1}} \left| \sqrt{n} (\mathcal{A}_n(b) - (\gamma - 1)) \right|$ straff ist, folgt $\sup_{b \in \mathcal{S}^{d-1}} |Z_n^{(2)}(b)| = o_{\mathbb{P}}(1)$ und somit die Behauptung. \square

Kapitel 4

Konvergenz mit standardisierten Daten unter sphärischer Symmetrie

In diesem Kapitel sollen Konvergenzaussagen für die Teststatistik T_n unter Berücksichtigung der standardisierten Daten $Y_{n,j}$ gemacht werden. Als Grundannahme seien die d -variaten Zufallsvektoren X_1, \dots, X_n unabhängig identisch sphärisch-symmetrisch verteilt, und es gelte $EX_1 = 0$, $\text{Cov}(X_1, X_1) = E(X_1 X_1') = I_d$, sowie $E\|X_1\|^6 < \infty$. Die skalierten Residuen $Y_{n,j}$ seien wie in (2.4) definiert. Zunächst betrachten wir wie im vorhergehenden Kapitel den Zähler der gebrochenrationalen Größe $\eta_n^2(b)$. Hierzu bezeichnen analog für $k \in \{1, \dots, d\}$ und $b \in \mathcal{S}^{d-1}$

$$C_{n,k}(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n f_k(Y_{n,j}, b), \quad (4.1)$$

$$D_n(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n (b' Y_{n,j})^3, \quad (4.2)$$

$$W_n(b) := \begin{pmatrix} C_{n,1}(b) \\ \vdots \\ C_{n,d}(b) \\ D_n(b) \end{pmatrix} \quad (4.3)$$

die stochastischen Prozesse auf \mathcal{S}^{d-1} , welche auf den empirisch standardisierten Daten aufbauen.

4.1 Konvergenzaussagen für $C_{n,k}$

Das Problem bei dem oben definierten stochastischen Prozess $C_{n,k}$ liegt darin, dass in den standardisierten Daten $Y_{n,j}$, sowohl der Schätzer S_n für die Kovarianzmatrix als auch der Schätzer \bar{X}_n für den Erwartungswertvektor auftreten. Auf Grund dessen liegt keine Summe von unabhängigen Zufallsvariablen vor. Ziel dieses Abschnitts ist es, die Schätzer herauszurechnen, was mit Hilfe einer mehrdimensionalen Taylorentwicklung ersten Grades versucht wird. Im Folgenden sei kurz

$$\Delta_{n,j} := Y_{n,j} - X_j = \left(S_n^{-\frac{1}{2}} - I_d \right) X_j - S_n^{-\frac{1}{2}} \bar{X}_n \quad (4.4)$$

gesetzt. Weiter definieren wir den Hilfsprozess

$$\tilde{C}_{n,k}(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n [f_k(X_j, b) + 2\varphi_k(X_j)(b'X_j)(b'\Delta_{n,j}) + (b'X_j)^2\varphi_k(\Delta_{n,j})]. \quad (4.5)$$

Zur Trennung der relevanten Anteile aus dem Restterm $\Delta_{n,j}$ zerlegen wir den Hilfsprozess in drei Summen, welche separat betrachtet werden. Es gilt

$$\tilde{C}_{n,k}(b) = \tilde{C}_{n,k}^{(1)}(b) + 2\tilde{C}_{n,k}^{(2)}(b) + \tilde{C}_{n,k}^{(3)}(b),$$

wobei

$$\begin{aligned} \tilde{C}_{n,k}^{(1)}(b) &:= \frac{1}{\sqrt{n}} \sum_{j=1}^n f_k(X_j, b), \\ \tilde{C}_{n,k}^{(2)}(b) &:= \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi_k(X_j)(b'X_j)(b'\Delta_{n,j}), \\ \tilde{C}_{n,k}^{(3)}(b) &:= \frac{1}{\sqrt{n}} \sum_{j=1}^n (b'X_j)^2\varphi_k(\Delta_{n,j}). \end{aligned}$$

Da der Restterm $\Delta_{n,j}$ lediglich in den beiden letzten Hilfsprozessen auftritt, werden nur diese betrachtet. Man beachte, dass $\tilde{C}_{n,k}^{(1)}$ dem stochastischen Prozess $C_{n,k}$ aus dem letzten Kapitel entspricht. Nach kurzer Rechnung und mit $\varphi_k(X_j) = e_k'X_j$ erhält man die Identitäten

$$\begin{aligned} \tilde{C}_{n,k}^{(2)}(b) &= b' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi_k(X_j)(b'X_j)X_j - \left(b'S_n^{-\frac{1}{2}}\bar{X}_n \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi_k(X_j)(b'X_j), \\ \tilde{C}_{n,k}^{(3)}(b) &= e_k' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n (b'X_j)^2X_j - \varphi_k \left(S_n^{-\frac{1}{2}}\bar{X}_n \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n (b'X_j)^2. \end{aligned}$$

Hiermit kommen wir zum nächsten Hilfsprozess

$$C_{n,k}^*(b) = \frac{1}{\sqrt{n}} \sum_{j=1}^n [f_k(X_j, b) - 2b'X_jb_k - e_k'X_j]; \quad (4.6)$$

hierbei und im Folgenden sind die unter den Voraussetzungen explizit berechenbaren Erwartungswertvektoren

$$\begin{aligned} E(\varphi_k(X_1)(b'X_1)X_1) &= 0, \\ E((b'X_1)^2X_1) &= 0, \end{aligned}$$

sowie die Erwartungswerte

$$\begin{aligned} E(\varphi_k(X_1)(b'X_1)) &= b_k, \\ E(b'X_1)^2 &= 1 \end{aligned}$$

zu beachten.

Theorem 4.1.1

Es existiert ein zentrierter Gaußprozess $C_k^*(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitskugel mit stetigen Pfaden und dem Kovarianzkern

$$\begin{aligned} \rho_k^*(b, c) &= \frac{m_6}{d(d+2)(d+4)} [2(b'c)^2 + 8b_k c_k (b'c) + 2(c_k^2 + b_k^2) + 1] \\ &\quad - \frac{2m_4}{d(d+2)} [4b_k c_k (b'c) + c_k^2 + b_k^2] + 4b_k c_k (b'c) - 1 \end{aligned} \quad (4.7)$$

($b, c \in \mathcal{S}^{d-1}$). Betrachtet man $C_k^*(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$, so gilt

$$C_{n,k}(\cdot) \xrightarrow{\mathcal{D}} C_k^*(\cdot).$$

Beweis:

Der Beweis des Theorems unterteilt sich in drei Teile. Mit der Norm $\|\cdot\|_C$ aus (B.2) in $C(\mathcal{S}^{d-1}, \mathbb{R})$ ist das Folgende zu zeigen:

1. $C_{n,k}^* \xrightarrow{\mathcal{D}} C_k^*$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$
2. $\|C_{n,k} - \tilde{C}_{n,k}\|_C \xrightarrow{\mathbb{P}} 0$ und
3. $\|\tilde{C}_{n,k} - C_{n,k}^*\|_C \xrightarrow{\mathbb{P}} 0$

für $n \rightarrow \infty$. Dann folgt die Behauptung direkt aus Theorem 4.1 aus [6].

Der Nachweis der Behauptungen 1.-3. ist Gegenstand der folgenden Lemmata.

Lemma 4.1.2

Für das in (4.6) definierte Zufallselement $C_{n,k}^*$ gilt die schwache Konvergenz $C_{n,k}^* \xrightarrow{\mathcal{D}} C_k^*$ im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$.

Beweis:

Es gilt $C_{n,k}^*(b) = \frac{1}{\sqrt{n}} \sum_{j=1}^n C_{j,k}^{**}(b)$ mit

$$C_{j,k}^{**}(b) := f_k(X_j, b) - 2b'X_j b_k - \varphi_k(X_j), \quad j = 1, \dots, n. \quad (4.8)$$

Wir betrachten die unabhängigen identisch verteilten Zufallsvariablen $C_{1,k}^{**}, \dots, C_{n,k}^{**}$ als Zufallselemente im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$. Die Zufallselemente sind zentriert, und es gilt $E\|C_{1,k}^{**}\|^2 < \infty$. Um den Kovarianzkern von C_k^* zu bestimmen, reicht es aus, den Kovarianzkern von $C_{1,k}^{**}$ zu berechnen. Es gilt für $b, c \in \mathcal{S}^{d-1}$

$$\begin{aligned} E(C_{1,k}^{**}(b)C_{1,k}^{**}(c)) &= E[f_k(X_1, b) - 2b'X_1 b_k - \varphi_k(X_1)] [f_k(X_1, c) - 2c'X_1 c_k - \varphi_k(X_1)] \\ &= E(f_k(X_1, b)f_k(X_1, c)) - 2c_k E(f_k(X_1, b)c'X_1) \\ &\quad - E(f_k(X_1, b)\varphi_k(X_1)) - 2b_k E(f_k(X_1, c)b'X_1) \\ &\quad + 4b_k c_k E((b'X_1)(c'X_1)) + 2b_k E((b'X_1)\varphi_k(X_1)) \\ &\quad - E(f_k(X_1, c)\varphi_k(X_1)) + 2c_k E((c'X_1)\varphi_k(X_1)) + E(\varphi_k(X_1))^2. \end{aligned}$$

Unter Benutzung von $\varrho_k(b, c)$ aus (3.1) gilt

$$E(f_k(X_1, b)f_k(X_1, c)) = \varrho_k(b, c), \quad (4.9)$$

$$E(f_k(X_1, b)c'X_1) = \frac{m_4}{d(d+2)}(2b_k(b'c) + c_k), \quad (4.10)$$

$$E(f_k(X_1, b)\varphi_k(X_1)) = 1 + 2b_k^2, \quad (4.11)$$

$$E((b'X_1)(c'X_1)) = b'c, \quad (4.12)$$

$$E((b'X_1)\varphi_k(X_1)) = b_k, \quad (4.13)$$

und mit einfachen Symmetrieargumenten und kurzem Rechnen folgt die behauptete Kovarianzstruktur. Die Entropiebedingung (3.2) ist in diesem Fall wie im Beweis zu Theorem 3.1.1 erfüllt. Ferner gilt mit der Cauchy-Schwarzschen Ungleichung, der Dreiecksungleichung und (3.3) für $b, c \in \mathcal{S}^{d-1}$

$$\begin{aligned} |C_{1,k}^{**}(b) - C_{1,k}^{**}(c)| &= |f_k(X_1, b) - 2b'X_1b_k - f_k(X_1, c) + 2c'X_1c_k| \quad (4.14) \\ &\leq |f_k(X_1, b) - f_k(X_1, c)| + 2|c'X_1c_k - b'X_1b_k| \\ &\leq 2\|X_1\|^3\|b - c\| + |c'X_1|c_k - b_k| + |b_k||c'X_1 - b'X_1| \\ &\leq 2(\|X_1\|^3 + \|X_1\|)\|b - c\|. \end{aligned}$$

Mit $E\|X_1\|^6 < \infty$ folgt nach [2], S. 172, Korollar 7.17, die behauptete schwache Konvergenz gegen den Grenzprozess C_k^* in $C(\mathcal{S}^{d-1}, \mathbb{R})$. \square

Lemma 4.1.3

Für die in (4.1) und (4.5) definierten Zufallselemente gilt

$$\|C_{n,k} - \tilde{C}_{n,k}\|_C \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Mit $\Delta_{n,j}$ aus (4.4) ergibt sich

$$Y_{n,j} = X_j + \Delta_{n,j}.$$

Weiter gilt mit der Linearität der Projektion φ_k

$$\begin{aligned} f_k(Y_{n,j}, b) &= (b'Y_{n,j})^2\varphi_k(Y_{n,j}) \\ &= (b'\Delta_{n,j} + b'X_j)^2(\varphi_k(\Delta_{n,j}) + \varphi_k(X_j)) \\ &= f_k(X_j, b) + (b'X_j)^2\varphi_k(\Delta_{n,j}) + 2(b'\Delta_{n,j})(b'X_j)\varphi_k(X_j) \\ &\quad + \xi_{n,k,j}^{(1)}(b) + \xi_{n,k,j}^{(2)}(b), \end{aligned}$$

mit $|\xi_{n,k,j}^{(1)}(b)| \leq \|\Delta_{n,j}\|^3$ und $|\xi_{n,k,j}^{(2)}(b)| \leq 3\|X_j\|\|\Delta_{n,j}\|^2$. Wie in [17], S. 9, oder in [15], S. 14, schon gezeigt wurde, ist

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^2 = o_{\mathbb{P}}(1). \quad (4.15)$$

Mit $\Delta_{n,j} = \left(S_n^{-\frac{1}{2}} - I_d\right)X_j - S_n^{-\frac{1}{2}}\bar{X}_n$ und der Dreiecksungleichung folgt

$$\|\Delta_{n,j}\|^2 \leq \|(S_n^{-\frac{1}{2}} - I_d)X_j\|^2 + 2\|(S_n^{-\frac{1}{2}} - I_d)X_j\|\|S_n^{-\frac{1}{2}}\bar{X}_n\| + \|S_n^{-\frac{1}{2}}\bar{X}_n\|^2$$

und damit

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=1}^n \|X_j\| \|\Delta_{n,j}\|^2 &\leq \|S_n^{-\frac{1}{2}} - I_d\|_S^2 \frac{1}{\sqrt{n}} \sum_{j=1}^n \|X_j\|^3 \\
&\quad + \|S_n^{-\frac{1}{2}} - I_d\|_S \|S_n^{-\frac{1}{2}} \bar{X}_n\| \frac{2}{\sqrt{n}} \sum_{j=1}^n \|X_j\|^2 \\
&\quad + \|S_n^{-\frac{1}{2}} \bar{X}_n\|^2 \frac{1}{\sqrt{n}} \sum_{j=1}^n \|X_j\| \\
&\leq \frac{1}{\sqrt{n}} \|\sqrt{n} (S_n^{-\frac{1}{2}} - I_d)\|_S^2 \frac{1}{n} \sum_{j=1}^n \|X_j\|^3 \\
&\quad + 2 \|\sqrt{n} (S_n^{-\frac{1}{2}} - I_d)\|_S \|S_n^{-\frac{1}{2}} \bar{X}_n\| \frac{1}{n} \sum_{j=1}^n \|X_j\|^2 \\
&\quad + \frac{1}{\sqrt{n}} \|S_n^{-1}\|_S \|\sqrt{n} \bar{X}_n\|^2 \frac{1}{n} \sum_{j=1}^n \|X_j\| \\
&= o_{\mathbb{P}}(1).
\end{aligned} \tag{4.16}$$

Demnach folgt

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^3 &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^2 \|(S_n^{-\frac{1}{2}} - I_d) X_j - S_n^{-\frac{1}{2}} \bar{X}_n\| \\
&\leq \|S_n^{-\frac{1}{2}} - I_d\|_S \frac{1}{\sqrt{n}} \sum_{j=1}^n \|X_j\| \|\Delta_{n,j}\|^2 \\
&\quad + \|S_n^{-\frac{1}{2}} \bar{X}_n\| \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^2 \\
&= o_{\mathbb{P}}(1).
\end{aligned} \tag{4.17}$$

Zusammengesetzt ergibt dies für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned}
|C_{n,k}(b) - \tilde{C}_{n,k}(b)| &= \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\xi_{n,k,j}^{(1)}(b) + \xi_{n,k,j}^{(2)}(b) \right) \right| \\
&\leq \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^3 + \frac{3}{\sqrt{n}} \sum_{j=1}^n \|X_j\| \|\Delta_{n,j}\|^2.
\end{aligned}$$

Also ist

$$\begin{aligned}
\|C_{n,k} - \tilde{C}_{n,k}\|_C &= \sup_{b \in \mathcal{S}^{d-1}} |C_{n,k}(b) - \tilde{C}_{n,k}(b)| \\
&\leq \sup_{b \in \mathcal{S}^{d-1}} \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^3 + \frac{3}{\sqrt{n}} \sum_{j=1}^n \|X_j\| \|\Delta_{n,j}\|^2 \right\} \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^3 + \frac{3}{\sqrt{n}} \sum_{j=1}^n \|X_j\| \|\Delta_{n,j}\|^2 \xrightarrow{\mathbb{P}} 0,
\end{aligned}$$

und damit folgt die Behauptung. \square

Lemma 4.1.4

Für die in (4.5) und (4.6) definierten Zufallselemente gilt

$$\|\tilde{C}_{n,k} - C_{n,k}^*\|_C \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Nach Einsetzen von $\Delta_{n,j} = \left(S_n^{-\frac{1}{2}} - I_d\right) X_j - S_n^{-\frac{1}{2}} \bar{X}_n$ gilt für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned} \tilde{C}_{n,k}(b) - C_{n,k}^*(b) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n [2\varphi_k(X_j)(b'X_j)(b'\Delta_{n,j}) + (b'X_j)^2\varphi_k(\Delta_{n,j}) + 2b'X_jb_k \\ &\quad + e_k'X_j] \\ &= 2b' \left(S_n^{-\frac{1}{2}} - I_d\right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi_k(X_j)(b'X_j)X_j \\ &\quad - 2b'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi_k(X_j)(b'X_j) \\ &\quad + e_k' \left(S_n^{-\frac{1}{2}} - I_d\right) \frac{1}{\sqrt{n}} \sum_{j=1}^n (b'X_j)^2X_j - e_k'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n (b'X_j)^2 \\ &\quad + 2b_kb' \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j + e_k' \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j \\ &= \mathcal{A}_{n,k}(b) - \mathcal{B}_{n,k}(b) - \sqrt{n}2b'S_n^{-\frac{1}{2}}\bar{X}_nb_k + \mathcal{C}_{n,k}(b) - \mathcal{D}_{n,k}(b) \\ &\quad - \sqrt{n}e_k'S_n^{-\frac{1}{2}}\bar{X}_n + 2b_kb'\sqrt{n}\bar{X}_n + e_k'\sqrt{n}\bar{X}_n \\ &= \mathcal{A}_{n,k}(b) - \mathcal{B}_{n,k}(b) + \mathcal{C}_{n,k}(b) - \mathcal{D}_{n,k}(b) + \mathcal{E}_n(b) + \mathcal{F}_n(b) \end{aligned}$$

mit

$$\mathcal{A}_{n,k}(b) := 2b' \left(S_n^{-\frac{1}{2}} - I_d\right) \frac{1}{\sqrt{n}} \sum_{j=1}^n [\varphi_k(X_j)(b'X_j)X_j],$$

$$\mathcal{B}_{n,k}(b) := 2b'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n [\varphi_k(X_j)(b'X_j) - b_k],$$

$$\mathcal{C}_{n,k}(b) := e_k' \left(S_n^{-\frac{1}{2}} - I_d\right) \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b'X_j)^2X_j],$$

$$\mathcal{D}_{n,k}(b) := e_k'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b'X_j)^2 - 1]$$

und

$$\mathcal{E}_n(b) := 2b' \left[-\sqrt{n}S_n^{-\frac{1}{2}}\bar{X}_n + \sqrt{n}\bar{X}_n\right] b_k,$$

$$\mathcal{F}_n(b) := e_k' \left[-\sqrt{n}S_n^{-\frac{1}{2}}\bar{X}_n + \sqrt{n}\bar{X}_n\right].$$

Da die Dreiecksungleichung

$$\|\tilde{C}_{n,k} - C_{n,k}^*\|_C \leq \|\mathcal{A}_{n,k}\|_C + \|\mathcal{B}_{n,k}\|_C + \|\mathcal{C}_{n,k}\|_C + \|\mathcal{D}_{n,k}\|_C + \|\mathcal{E}_n\|_C + \|\mathcal{F}_n\|_C$$

liefert, genügt es die stochastische Konvergenz der sechs Summanden auf der rechten Seite gegen 0 zu zeigen. Dies wird sukzessive in den folgenden Unterabschnitten (a)-(f) gezeigt.

(a) Mit den unabhängigen identisch verteilten und zentrierten Zufallsvektoren

$$\tilde{\mathcal{A}}_k(X_j, b) := \varphi_k(X_j)(b'X_j)X_j,$$

$j \in \{1, \dots, n\}$, gilt punktweise für $b \in \mathcal{S}^{d-1}$

$$|\mathcal{A}_{n,k}(b)| \leq 2 \left\| S_n^{-\frac{1}{2}} - I_d \right\|_S \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{A}}_k(X_j, b) \right\|,$$

und hiermit

$$\begin{aligned} \|\mathcal{A}_{n,k}\|_C^2 &= \sup_{b \in \mathcal{S}^{d-1}} |\mathcal{A}_{n,k}(b)|^2 \\ &\leq 4 \left\| S_n^{-\frac{1}{2}} - I_d \right\|_S^2 \sup_{b \in \mathcal{S}^{d-1}} \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{A}}_k(X_j, b) \right\|^2. \end{aligned}$$

Da $\left\| S_n^{-\frac{1}{2}} - I_d \right\|_S^2$ fast sicher gegen 0 konvergiert, reicht es die Straffheit von

$$\sup_{b \in \mathcal{S}^{d-1}} \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{A}}_k(X_j, b) \right\|^2$$

zu zeigen. Um dies einzusehen betrachten wir den zentrierten stochastischen Prozess

$$\tilde{\mathcal{A}}_k(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{A}}_k(X_j, b), \quad b \in \mathcal{S}^{d-1}$$

in $C(\mathcal{S}^{d-1}, \mathbb{R})$. Auf Grund der Voraussetzungen existiert der Kovarianzmatrixkern $E\left(\tilde{\mathcal{A}}_k(X_1, b)\tilde{\mathcal{A}}_k(X_1, c)'\right)$. Weiter gilt mit der Cauchy-Schwarzschen Ungleichung für $b, c \in \mathcal{S}^{d-1}$

$$\begin{aligned} \left\| \tilde{\mathcal{A}}_k(X_1, b) - \tilde{\mathcal{A}}_k(X_1, c) \right\| &= \left\| \varphi_k(X_1)(b'X_1)X_1 - \varphi_k(X_1)(c'X_1)X_1 \right\| \\ &\leq |\varphi_k(X_1)| \|X_1\|^2 \|b - c\|. \end{aligned}$$

Aus Theorem B.4.1 folgt die schwache Konvergenz von $\tilde{\mathcal{A}}_k$ gegen einen zentrierten Gaußprozess \mathcal{A}_k in $C(\mathcal{S}^{d-1}, \mathbb{R}^d)$ und nach dem Abbildungssatz auch die schwache Konvergenz

$$\sup_{b \in \mathcal{S}^{d-1}} \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{A}}_k(X_j, b) \right\|^2 \xrightarrow{\mathcal{D}} \sup_{b \in \mathcal{S}^{d-1}} \|\mathcal{A}_k(b)\|^2.$$

Da Verteilungskonvergenz Straffheit impliziert, folgt die Behauptung.

- (b) Eine ähnliche Argumentation führt mit der Straffheit von $S_n^{-\frac{1}{2}}$ und der fast sicheren Konvergenz von \bar{X}_n gegen 0 zu

$$\begin{aligned} \|\mathcal{B}_{n,k}\|_C &= \sup_{b \in \mathcal{S}^{d-1}} |\mathcal{B}_{n,k}(b)| \\ &\leq 2 \sup_{b \in \mathcal{S}^{d-1}} \left\{ b'(S_n^{-\frac{1}{2}} \bar{X}_n) \right\} \sup_{b \in \mathcal{S}^{d-1}} \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{B}}_k(X_j, b) \right\} \xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Dabei bezeichnen

$$\tilde{\mathcal{B}}_k(X_j, b) := \varphi_k(X_j)(b'X_j) - b_k, \quad j = 1, \dots, n,$$

unabhängige, identisch verteilte und zentrierte Zufallsvariablen. Wie bei (a) bleibt zu zeigen, dass $\sup_{b \in \mathcal{S}^{d-1}} \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{B}}_{n,k}(X_j, b) \right\}$ straff ist. Wir betrachten den zentrierten, stochastischen Prozess

$$\tilde{B}_k(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n \tilde{\mathcal{B}}_k(X_j, b), \quad b \in \mathcal{S}^{d-1}$$

in $C(\mathcal{S}^{d-1}, \mathbb{R})$. Weiter gilt mit der Dreiecksungleichung und der Cauchy-Schwarzschen Ungleichung

$$\begin{aligned} \left| \tilde{\mathcal{B}}_k(X_1, b) - \tilde{\mathcal{B}}_k(X_1, c) \right| &= |\varphi_k(X_1)(b'X_1) - b_k - \varphi_k(X_1)(c'X_1) + c_k| \\ &\leq |\varphi_k(X_1)| |b'X_1 - c'X_1| + |c_k - b_k| \\ &\leq (|\varphi_k(X_1)| \|X_1\| + 1) \|b - c\|. \end{aligned}$$

Die Behauptung folgt aus Theorem B.4.1.

- (c) Analog zu (a).
 (d) Analog zu (b).
 (e) Es ist für $b \in \mathcal{S}^{d-1}$

$$|\mathcal{E}_n(b)| \leq 2 \|I_d - S_n^{-\frac{1}{2}}\|_S \|\sqrt{n} \bar{X}_n\| |b_k|.$$

Demnach gilt

$$\begin{aligned} \|\mathcal{E}_n\|_C &= \sup_{b \in \mathcal{S}^{d-1}} |\mathcal{E}_n(b)| \\ &\leq 2 \|I_d - S_n^{-\frac{1}{2}}\|_S \|\sqrt{n} \bar{X}_n\| \sup_{b \in \mathcal{S}^{d-1}} |b_k| \xrightarrow{\mathbb{P}} 0, \end{aligned}$$

denn $\sqrt{n} \bar{X}_n$ ist straff, $\sup_{b \in \mathcal{S}^{d-1}} |b_k| < 1$ und $S_n^{-\frac{1}{2}}$ geht fast sicher gegen I_d .

- (f) Analog zu (e). □

4.2 Konvergenzaussagen für D_n

Im Folgenden soll das Grenzverhalten des stochastischen Prozesses D_n aus (4.2) analysiert werden. Die Methodik des vorangehenden Abschnittes führt auch in diesem Fall zum Ziel. Wir werden zunächst eine Taylorapproximation durchführen, um dann im nächsten Schritt die relevanten Anteile herauszufiltern.

Theorem 4.2.1

Es existiert ein zentrierter Gaußprozess $D^*(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitssphäre mit stetigen Pfaden und dem Kovarianzkern

$$\rho^*(b, c) = \frac{m_6}{d(d+2)(d+4)}(6(b'c)^3 + 9(b'c)) - \frac{18m_4}{d(d+2)}(b'c) + 9(b'c) \quad (4.18)$$

($b, c \in \mathcal{S}^{d-1}$). Betrachtet man $D^*(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$, so gilt

$$D_n(\cdot) \xrightarrow{\mathcal{D}} D^*(\cdot).$$

Beweis:

Wir definieren den Hilfsprozess

$$\tilde{D}_n(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b'X_j)^3 + 3(b'X_j)^2(b'\Delta_{n,j})]. \quad (4.19)$$

Unter Benutzung der Identität $\Delta_{n,j} = (S_n^{-\frac{1}{2}} - I_d) X_j - S_n^{-\frac{1}{2}} \bar{X}_n$ kommen wir zum zweiten Hilfsprozess

$$D_n^*(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b'X_j)^3 - 3b'X_j]. \quad (4.20)$$

Zum Beweis des Theorems verwenden wir wieder drei Schritte und zeigen der Reihe nach

1. $D_n^* \xrightarrow{\mathcal{D}} D^*$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$
2. $\|D_n - \tilde{D}_n\|_C \xrightarrow{\mathbb{P}} 0$ und
3. $\|\tilde{D}_n - D_n^*\|_C \xrightarrow{\mathbb{P}} 0$.

Zur Übersichtlichkeit werden die Einzelergebnisse in drei Lemmata aufgeteilt.

Lemma 4.2.2

Für das in (4.20) definierte Zufallselement D_n^* gilt die schwache Konvergenz $D_n^* \xrightarrow{\mathcal{D}} D^*$ im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$.

Beweis:

Es gilt $D_n^*(b) = \frac{1}{\sqrt{n}} \sum_{j=1}^n D_j^{**}(b)$ mit

$$D_j^{**}(b) := (b'X_j)^3 - 3b'X_j, \quad (j = 1, \dots, n).$$

Mit der gleichen Argumentation wie im Beweis zu Lemma 4.1.2 reicht es, die Kovarianzstruktur von D_1^{**} zu bestimmen. Mit $\varrho(b, c)$ aus (3.4) und

$$E((b'X_1)(c'X_1)^3) = \frac{3m_4}{d(d+2)}(b'c)$$

(siehe Bemerkung C.4.1) gilt für $b, c \in \mathcal{S}^{d-1}$

$$\begin{aligned} E(D_1^{**}(b)D_1^{**}(c)) &= E((b'X_1)^3(c'X_1)^3) - 3E((b'X_1)(c'X_1)^3) \\ &\quad - 3E((b'X_1)^3(c'X_1)) + 9E((b'X_1)(c'X_1)) \\ &= \varrho(b, c) - \frac{18m_4}{d(d+2)}(b'c) + 9(b'c). \end{aligned}$$

Weiter gilt

$$\begin{aligned} |D_1^{**}(b) - D_1^{**}(c)| &= |(b'X_j)^3 - 3b'X_j - (c'X_j)^3 + 3c'X_j| \\ &\leq 3(\|X_1\|^3 + \|X_1\|)\|b - c\| \end{aligned} \quad (4.21)$$

und wegen $E\|X_1\|^6 < \infty$ folgt nach [2], S. 172, Korollar 7.17 die Behauptung. \square

Lemma 4.2.3

Für die in (4.2) und (4.19) definierten Zufallselemente gilt

$$\|D_n - \tilde{D}_n\|_C \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Es gilt mit $Y_{n,j} = X_j + \Delta_{n,j}$ für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned} (b'Y_{n,j})^3 &= (b'\Delta_{n,j} + b'X_j)^3 \\ &= (b'X_j)^3 + 3(b'X_j)^2(b'\Delta_{n,j}) + 3(b'\Delta_{n,j})^2(b'X_j) + (b'\Delta_{n,j})^3 \\ &= (b'X_j)^3 + 3(b'X_j)^2(b'\Delta_{n,j}) + \eta_{n,j}^{(1)}(b) + \eta_{n,j}^{(2)}(b) \end{aligned}$$

mit $|\eta_{n,j}^{(1)}(b)| \leq 3\|\Delta_{n,j}\|^2\|X_j\|$ und $|\eta_{n,j}^{(2)}(b)| \leq \|\Delta_{n,j}\|^3$. Wie im Beweis zu Lemma 4.1.3 gilt für $b \in \mathcal{S}^{d-1}$ nach der Dreiecksungleichung

$$\begin{aligned} |D_n(b) - \tilde{D}_n(b)| &= \left| \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\eta_{n,j}^{(1)}(b) + \eta_{n,j}^{(2)}(b) \right) \right| \\ &\leq \frac{3}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^2\|X_j\| + \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^3. \end{aligned}$$

Es folgt

$$\begin{aligned} \|D_n - \tilde{D}_n\|_C &= \sup_{b \in \mathcal{S}^{d-1}} |D_n(b) - \tilde{D}_n(b)| \\ &\leq \frac{3}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^2 \|X_j\| + \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^3. \end{aligned}$$

Nach (4.17) und (4.16) konvergiert der letzte Term stochastisch gegen 0. \square

Lemma 4.2.4

Für die in (4.19) und (4.20) definierten Zufallselemente gilt

$$\|\tilde{D}_n - D_n^*\|_C \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Die Identität $\Delta_{n,j} = (S_n^{-\frac{1}{2}} - I_d) X_j - S_n^{-\frac{1}{2}} \bar{X}_n$ liefert für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned} \tilde{D}_n(b) - D_n^*(b) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n [3(b' X_j)^2 (b' \Delta_{n,j}) + 3b' X_j] \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^n [3(b' X_j)^2 (b' (S_n^{-\frac{1}{2}} - I_d) X_j - b' S_n^{-\frac{1}{2}} \bar{X}_n) + 3b' X_j] \\ &= 3b' (S_n^{-\frac{1}{2}} - I_d) \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b' X_j)^2 X_j] - 3b' S_n^{-\frac{1}{2}} \bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n (b' X_j)^2 \\ &\quad + 3b' \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j \\ &= 3b' (S_n^{-\frac{1}{2}} - I_d) \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b' X_j)^2 X_j] \\ &\quad - 3b' S_n^{-\frac{1}{2}} \bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b' X_j)^2 - 1] - 3\sqrt{n} b' S_n^{-\frac{1}{2}} \bar{X}_n + 3b' \sqrt{n} \bar{X}_n \\ &= 3(\mathcal{A}_n(b) - \mathcal{B}_n(b) - \mathcal{C}_n(b)) \end{aligned}$$

mit

$$\begin{aligned} \mathcal{A}_n(b) &:= b' (S_n^{-\frac{1}{2}} - I_d) \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b' X_j)^2 X_j], \\ \mathcal{B}_n(b) &:= b' S_n^{-\frac{1}{2}} \bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b' X_j)^2 - 1], \\ \mathcal{C}_n(b) &:= b' (\sqrt{n} (S_n^{-\frac{1}{2}} - I_d) \bar{X}_n). \end{aligned}$$

Nach der Dreiecksungleichung gilt

$$\|\tilde{D}_n - D_n^*\|_C \leq 3 (\|\mathcal{A}_n\|_C + \|\mathcal{B}_n\|_C + \|\mathcal{C}_n\|_C),$$

also genügt es die stochastische Konvergenz der drei Summanden auf der rechten Seite gegen 0 zu zeigen. Um dies einzusehen, verwende man die gleiche Argumentation wie im letzten Teil des Beweises zu Lemma 4.1.4. \square

Bemerkung 4.2.5

Beim Grenzprozess D^* in $C(\mathcal{S}^{d-1}, \mathbb{R})$ handelt es sich um denselben Grenzprozess wie in Theorem 2.1 aus [4], S. 53. In jener Arbeit werden in Proposition 2.3 und 2.5 zwei alternative stochastische Darstellungen für den Grenzprozess D^* angegeben.

4.3 Konvergenzaussagen für W_n

In diesem Abschnitt soll das Konvergenzverhalten des multivariaten stochastischen Prozesses W_n untersucht werden. Hierfür werden die vorangehenden Ergebnisse verwendet und erweitert. Analog zu den Hilfsprozessen $\tilde{C}_{n,k}$, $C_{n,k}^*$, \tilde{D}_n und D_n^* definieren wir für $b \in \mathcal{S}^{d-1}$ die Hilfsprozesse

$$\tilde{W}_n(b) := \begin{pmatrix} \tilde{C}_{n,1}(b) \\ \vdots \\ \tilde{C}_{n,d}(b) \\ \tilde{D}_n(b) \end{pmatrix} \quad \text{und} \quad W_n^*(b) := \begin{pmatrix} C_{n,1}^*(b) \\ \vdots \\ C_{n,d}^*(b) \\ D_n^*(b) \end{pmatrix}. \quad (4.22)$$

Im folgenden Theorem werden abkürzend für $k, l \in \{1, \dots, d\}$ mit $k \neq l$ und $b, c \in \mathcal{S}^{d-1}$ die Abbildungen

$$\begin{aligned} \Phi_{k,l}^{(1,*)}(b, c) &:= \text{Cov}(C_{1,k}^{**}(b), C_{1,l}^{**}(c)), \\ \Phi_k^{(2,*)}(b, c) &:= \text{Cov}(D_1^{**}(b), C_{1,k}^{**}(c)), \end{aligned}$$

verwendet. Unter den gemachten Voraussetzungen gilt

$$\begin{aligned} \Phi_{k,l}^{(1,*)}(b, c) &= \Phi_{k,l}^{(1)}(b, c) - \frac{4m_4}{d(d+2)} (2b_k c_l (b'c) + c_k c_l + b_k b_l) \\ &\quad + 4b_k c_l (b'c) + 2c_k c_l + 2b_k b_l \end{aligned} \quad (4.23)$$

$$\Phi_k^{(2,*)}(b, c) = \Phi_k^{(2)}(b, c) - \frac{6m_4}{d(d+2)} (2c_k (b'c) + b_k) + 6c_k (b'c) + 3b_k. \quad (4.24)$$

Die Beweise zu den eben gemachten Aussagen finden sich im Anhang A.2. Die schwache Konvergenz gegen den Gaußschen multivariaten Grenzprozess W^* zeigt das folgende Theorem.

Theorem 4.3.1

Es existiert ein zentrierter $(d + 1)$ -variater Gaußprozess $W^*(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitssphäre mit stetigen Pfaden und einer durch den Matrixkern

$$\sigma^*(b, c) = \begin{pmatrix} \rho_1^*(b, c) & \Phi_{1,2}^{(1,*)}(b, c) & \cdots & \Phi_{1,d}^{(1,*)}(b, c) & \Phi_1^{(2,*)}(b, c) \\ \Phi_{2,1}^{(1,*)}(b, c) & \rho_2^*(b, c) & \cdots & \Phi_{2,d}^{(1,*)}(b, c) & \Phi_2^{(2,*)}(b, c) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Phi_{d,1}^{(1,*)}(b, c) & \Phi_{d,2}^{(1,*)}(b, c) & \cdots & \rho_d^*(b, c) & \Phi_d^{(2,*)}(b, c) \\ \Phi_1^{(2,*)}(c, b) & \Phi_2^{(2,*)}(c, b) & \cdots & \Phi_d^{(2,*)}(c, b) & \rho^*(b, c) \end{pmatrix}$$

($b, c \in \mathcal{S}^{d-1}$) gegebenen Kovarianzstruktur. Betrachtet man $W^*(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$, so gilt

$$W_n(\cdot) \xrightarrow{\mathcal{D}} W^*(\cdot).$$

Beweis:

Wie schon in den Beweisen der vorangehenden Theoreme können wir den Beweis in drei Teile zerlegen. Es ist

1. $W_n^* \xrightarrow{\mathcal{D}} W^*$ in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$
2. $\|W_n - \tilde{W}_n\| \xrightarrow{\mathbb{P}} 0$ und
3. $\|\tilde{W}_n - W_n^*\| \xrightarrow{\mathbb{P}} 0$

für $n \rightarrow \infty$ zu zeigen.

1. Die Kovarianzstruktur ergibt sich durch direktes Berechnen von $\sigma^*(b, c) = E(W_1^*(b)W_1^*(c)')$ (siehe hierzu Anhang A.2). Die Entropiebedingung für [2], S. 172, Korollar 7.17, ist auch hier erfüllt, und es gilt für $b, c \in \mathcal{S}^{d-1}$ mit (4.14) und (4.21)

$$\begin{aligned} \|W_1^*(b) - W_1^*(c)\| &= \sqrt{\sum_{k=1}^d |C_{1,k}^{**}(b) - C_{1,k}^{**}(c)|^2 + |D^{**}(b) - D^{**}(c)|^2} \\ &\leq \sqrt{4d + 9(\|X_1\|^3 + \|X_1\|)} \|b - c\|. \end{aligned}$$

Da nach Voraussetzung $E\|X_1\|^6 < \infty$ ist, folgt die Behauptung.

2. Es gilt für $b \in \mathcal{S}^{d-1}$

$$\|W_n(b) - \tilde{W}_n(b)\|^2 = \sum_{k=1}^d |C_{n,k}(b) - \tilde{C}_{n,k}(b)|^2 + |D_n(b) - \tilde{D}_n(b)|^2.$$

Mit Lemma 4.1.3 und 4.2.3 folgt

$$\begin{aligned}
\|W_n - \tilde{W}_n\|^2 &= \sup_{b \in \mathcal{S}^{d-1}} \|W_n(b) - \tilde{W}_n(b)\|^2 \\
&= \sup_{b \in \mathcal{S}^{d-1}} \left\{ \sum_{k=1}^d |C_{n,k}(b) - \tilde{C}_{n,k}(b)|^2 + |D_n(b) - \tilde{D}_n(b)|^2 \right\} \\
&\leq \sum_{k=1}^d \sup_{b \in \mathcal{S}^{d-1}} |C_{n,k}(b) - \tilde{C}_{n,k}(b)|^2 + \sup_{b \in \mathcal{S}^{d-1}} |D_n(b) - \tilde{D}_n(b)|^2 \\
&= \sum_{k=1}^d \|C_{n,k} - \tilde{C}_{n,k}\|_C^2 + \|D_n - \tilde{D}_n\|_C^2 \xrightarrow{\mathbb{P}} 0
\end{aligned}$$

für $n \rightarrow \infty$.

3. Analog zu dem eben Gezeigten folgt die stochastische Konvergenz gegen 0 aus der Dreiecksungleichung und der stochastischen Konvergenz der stochastischen Hilfsprozesse in Lemma 4.1.4 und Lemma 4.2.4 gegen 0.

□

Die Verbindung zwischen dem stochastischen Prozess W_n und der Größe η_n^2 aus (2.6) wird durch das folgende Theorem aufgezeigt. Dabei sei Ψ_B die in (2.14) definierte Abbildung.

Theorem 4.3.2

Für $\gamma = \frac{3m_4}{d(d+2)}$ gilt

$$\sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{\gamma - 1} \Psi_B(W_n(b)) - n\eta_n^2(b) \right| \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Wir können zum Beweis des Theorems dieselbe Argumentation wie im Beweis zu Lemma 3.1.8 verwenden. Im Folgenden sei $b \in \mathcal{S}^{d-1}$. Es bleibt noch zu überprüfen, ob

$$\begin{aligned}
\sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b'Y_{n,j})^4 - \gamma \right| &\xrightarrow{\mathbb{P}} 0 \text{ und} \\
\sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b'Y_{n,j})^3 \right| &\xrightarrow{\mathbb{P}} 0
\end{aligned} \tag{4.25}$$

für $n \rightarrow \infty$ gilt. Dies wird in den folgenden Zeilen gezeigt. Es ist für $j = 1, \dots, n$

$$\begin{aligned}
(b'Y_{n,j})^4 &= (b'(\Delta_{n,j} + X_j))^4 \\
&= (b'X_j)^4 + 4(b'X_j)^3(b'\Delta_{n,j}) + 6(b'X_j)^2(b'\Delta_{n,j})^2 + 4(b'X_j)(b'\Delta_{n,j})^3 + (b'\Delta_{n,j})^4 \\
&= (b'X_j)^4 + \eta_{n,j}(b)
\end{aligned}$$

mit

$$|\eta_{n,j}(b)| \leq \|\Delta_{n,j}\|^4 + 4\|X_j\|\|\Delta_{n,j}\|^3 + 6\|X_j\|^2\|\Delta_{n,j}\|^2 + 4\|X_j\|^3\|\Delta_{n,j}\|. \quad (4.26)$$

Ferner gilt mit (4.15)

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n \|\Delta_{n,j}\|^4 &\leq \frac{1}{n} \sum_{j_1, j_2=1}^n \|\Delta_{n,j_1}\|^2 \|\Delta_{n,j_2}\|^2 \\ &= \frac{1}{\sqrt{n}} \sum_{j_1=1}^n \|\Delta_{n,j_1}\|^2 \frac{1}{\sqrt{n}} \sum_{j_2=1}^n \|\Delta_{n,j_2}\|^2 \\ &= \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\|^2 \right)^2 \xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Für den zweiten Summanden in (4.26) ergibt sich

$$\frac{1}{n} \sum_{j=1}^n \|X_j\|\|\Delta_{n,j}\|^3 \leq \frac{1}{\sqrt{n}} \sum_{j_1=1}^n \|X_{j_1}\|\|\Delta_{n,j_1}\|^2 \frac{1}{\sqrt{n}} \sum_{j_2=1}^n \|\Delta_{n,j_2}\|, \quad (4.27)$$

wobei nach (4.16) $\frac{1}{\sqrt{n}} \sum_{j_1=1}^n \|X_{j_1}\|\|\Delta_{n,j_1}\|^2 = o_{\mathbb{P}}(1)$ ist. Weiter ist mit (4.4)

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{j=1}^n \|\Delta_{n,j}\| &\leq \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[\left\| \left(S_n^{-\frac{1}{2}} - I_d \right) X_j \right\| + \left\| S_n^{-\frac{1}{2}} \bar{X}_n \right\| \right] \\ &\leq \left\| \sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right) \right\| \left\| \frac{1}{n} \sum_{j=1}^n X_j \right\| + \left\| S_n^{-\frac{1}{2}} \right\| \left\| \sqrt{n} \bar{X}_n \right\|. \end{aligned}$$

Auf Grund der Straffheit von $\sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right)$ und $\sqrt{n} \bar{X}_n$, der fast sicheren Konvergenz $S_n^{-\frac{1}{2}} \xrightarrow{\text{f.s.}} I_d$ und von $\frac{1}{n} \sum_{j=1}^n \|X_j\| \xrightarrow{\text{f.s.}} E\|X_1\| < \infty$ gilt in (4.27)

$$\frac{1}{n} \sum_{j=1}^n \|X_j\|\|\Delta_{n,j}\|^3 = o_{\mathbb{P}}(1).$$

Den dritten Summanden in (4.26) schreiben wir mit Hilfe der Darstellung von $\Delta_{n,j}$ aus (4.4) wie folgt: Es gilt

$$\begin{aligned} \frac{1}{n} \sum_{j=1}^n \|X_j\|^2 \|\Delta_{n,j}\|^2 &\leq \frac{2}{n} \sum_{j=1}^n \left[\left\| S_n^{-\frac{1}{2}} - I_d \right\|_S^2 \|X_j\|^4 + \left\| S_n^{-\frac{1}{2}} \bar{X}_n \right\|^2 \|X_j\|^2 \right] \\ &= 2 \left\| S_n^{-\frac{1}{2}} - I_d \right\|_S^2 \frac{1}{n} \sum_{j=1}^n \|X_j\|^4 + 2 \left\| S_n^{-\frac{1}{2}} \bar{X}_n \right\|^2 \frac{1}{n} \sum_{j=1}^n \|X_j\|^2 \\ &\xrightarrow{\mathbb{P}} 0 \end{aligned}$$

für $n \rightarrow \infty$. Die Asymptotik des vierten Summanden erhalten wir durch die Abschätzung

$$\frac{1}{n} \sum_{j=1}^n \|X_j\|^3 \|\Delta_{n,j}\| \leq \|S_n^{-\frac{1}{2}} - I_d\|_S \frac{1}{n} \sum_{j=1}^n \|X_j\|^4 + \|S_n^{-\frac{1}{2}} \bar{X}_n\| \frac{1}{n} \sum_{j=1}^n \|X_j\|^3 \xrightarrow{\mathbb{P}} 0.$$

Insgesamt führt (3.13) zu

$$\begin{aligned} \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b' Y_{n,j})^4 - \gamma \right| &= \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n [(b' X_j)^4 + \eta_{n,j}(b)] - \gamma \right| \\ &\leq \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - \gamma \right| + \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n \eta_{n,j}(b) \right| \\ &\xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Unter Benutzung der Ergebnisse aus dem Beweis von Lemma 4.2.3 kann analog wie oben die stochastische Konvergenz in (4.25) gezeigt werden. \square

4.4 Die Grenzverteilung der Teststatistik T_n

In diesem Abschnitt wird für die Teststatistik T_n aus (2.5) die Grenzverteilung unter Annahme einer sphärisch-symmetrischen Grundverteilung angegeben.

Theorem 4.4.1

Seien X_1, \dots, X_n sphärisch-symmetrisch verteilte Zufallsvektoren mit $E\|X\|^6 < \infty$. Dann gilt für die in (2.5) angegebene Teststatistik T_n von Cox und Small mit γ wie in Theorem 4.3.2

$$nT_n \xrightarrow{\mathcal{D}} \frac{1}{\gamma - 1} \max_{b \in \mathcal{S}^{d-1}} \Psi_B(W^*(b))$$

für $n \rightarrow \infty$. Dabei sind Ψ_B die Abbildung aus (2.14) und $W^*(\cdot)$ der $(d+1)$ -dimensionale zentrierte Gaußsche Grenzprozess aus Theorem 4.3.1.

Beweis:

Unter den Vorarbeiten trivial. \square

Kapitel 5

Beispiele

Anhand bestimmter Verteilungsannahmen sollen im Folgenden die in den Kapiteln 3 und 4 hergeleiteten Ergebnisse verdeutlicht werden. Wir beschränken uns bei den Beispielen auf Verteilungen aus der Klasse der sphärisch-symmetrischen Verteilungen (für Details siehe [13]).

5.1 Die Grenzverteilung unter H_0

In diesem Abschnitt wird die Grenzverteilung unter der Hypothese H_0 untersucht. Auf Grund der affinen Invarianz der Teststatistik reicht es, die Grenzverteilung unter der Hypothese der Standardnormalverteilung zu charakterisieren. Demzufolge seien X_1, \dots, X_n unabhängige identisch $\mathcal{N}_d(0, I_d)$ -verteilte Zufallsvektoren. Unter diesen Voraussetzungen gilt nach [13], S. 77, für $r > 0$

$$E\|X_1\|^{2r} = 2^r \frac{\Gamma\left(\frac{d+2r}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}.$$

Mit dieser Formel folgt

$$\begin{aligned} m_4 &= d(d+2), \\ m_6 &= d(d+2)(d+4). \end{aligned} \tag{5.1}$$

Setzen wir diese Momente in die explizit angegebenen Formeln in (4.7), (4.18), (4.23) und (4.24) ein, so folgt für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$

$$\rho_{k, \mathcal{N}_d}^*(b, c) = 2(b'c)^2 + 4b_k c_k (b'c), \tag{5.2}$$

$$\rho_{\mathcal{N}_d}^*(b, c) = 6(b'c)^3, \tag{5.3}$$

$$\Phi_{k, l, \mathcal{N}_d}^{(1,*)}(b, c) = 4c_k b_l (b'c),$$

$$\Phi_{k, \mathcal{N}_d}^{(2,*)}(b, c) = 6b_k (b'c)^2.$$

Damit können wir den Kovarianzmatrixkern des Grenzprozesses W^* angeben. Es gilt

$$\sigma_{\mathcal{N}_d}^*(b, c) = 2(b'c)^2 \begin{pmatrix} I_d & 3b \\ 3c' & 3(b'c) \end{pmatrix} + 4(b'c) \begin{pmatrix} cb' & 0 \\ 0 & 0 \end{pmatrix}, \quad b, c \in \mathcal{S}^{d-1}. \tag{5.4}$$

5.2 Multivariate Pearson Typ II-Verteilung

Ein Zufallsvektor X ist Pearson Typ II verteilt ($X \sim MP_{II_d}(a, \mu, \Delta)$), falls X die Dichte

$$g(x) = \frac{\Gamma\left(\frac{d}{2} + a + 1\right)}{\Gamma(a + 1)\pi^{\frac{d}{2}}|\Delta|^{\frac{1}{2}}}(1 - (x - \mu)' \Delta^{-1}(x - \mu))^a,$$

besitzt. Dabei gilt $(x - \mu)' \Delta^{-1}(x - \mu) \leq 1$, $a > -1$, sowie $\mu \in \mathbb{R}^d$ und Δ ist eine symmetrische, positiv semidefinite $d \times d$ Matrix (siehe [13], Abschnitt 3.4 und [20], Abschnitt 6.2). Es gilt $E(X) = \mu$, und die Kovarianzmatrix von X lautet

$$E((X - \mu)(X - \mu)') = \frac{1}{d + 2a + 2}\Delta = \Sigma. \quad (5.5)$$

Die Zufallsvariable $(X - \mu)' \Delta^{-1}(X - \mu)$ ist nach [13], S. 89, Beta-verteilt mit den Parametern $\frac{d}{2}$ und $a + 1$. Für die Momente der Beta-verteilten Zufallsvariablen $(X - \mu)' \Delta^{-1}(X - \mu)$ gilt mit $r > 0$

$$E((X - \mu)' \Delta^{-1}(X - \mu))^r = \frac{B\left(\frac{d}{2} + r, a + 1\right)}{B\left(\frac{d}{2}, a + 1\right)} = \prod_{j=0}^{r-1} \frac{d + 2j}{d + 2a + 2 + 2j}. \quad (5.6)$$

Unter Beachtung von (5.5) wenden wir die Formel aus (5.6) für gerade Momente an und erhalten

$$\begin{aligned} m_{2r} &= E((X - \mu)' \Sigma^{-1}(X - \mu))^r \\ &= (d + 2a + 2)^r E((X - \mu)' \Delta^{-1}(X - \mu))^r. \end{aligned}$$

Für $r \in \{1, 2\}$ ergibt sich speziell

$$\begin{aligned} m_4 &= d(d + 2) \frac{d + 2a + 2}{d + 2a + 4}, \\ m_6 &= d(d + 2)(d + 4) \frac{(d + 2a + 2)^2}{(d + 2a + 4)(d + 2a + 6)}. \end{aligned}$$

Zur übersichtlicheren Schreibweise führen wir die von der Dimension d und dem Parameter a abhängenden Konstanten

$$\begin{aligned} \kappa_{d,a}^{(1)} &:= \frac{(d + 2a + 2)^2}{(d + 2a + 4)(d + 2a + 6)} \quad \text{und} \\ \kappa_{d,a}^{(2)} &:= \frac{d + 2a + 2}{d + 2a + 4} \end{aligned}$$

ein. Daraus ergibt sich für die in (4.7), (4.18), (4.23) und (4.24) angegebenen Größen

$$\begin{aligned}
\rho_{k,MPH_d}^*(b, c) &= \kappa_{d,a}^{(1)} [2(b'c)^2 + 8b_k c_k (b'c) + 2(c_k^2 + b_k^2) + 1] \\
&\quad - 2\kappa_{d,a}^{(2)} [4b_k c_k (b'c) + c_k^2 + b_k^2] + 4b_k c_k (b'c) - 1, \\
\rho_{MPH_d}^*(b, c) &= \kappa_{d,a}^{(1)} (6(b'c)^3 + 9(b'c)) - 18\kappa_{d,a}^{(2)} (b'c) + 9(b'c), \\
\Phi_{k,l,MPH_d}^{(1,*)}(b, c) &= 2\kappa_{d,a}^{(1)} (2(b'c)(b_k c_l + c_k b_l) + (b_k b_l + c_k c_l)) \\
&\quad - 4\kappa_{d,a}^{(2)} (2b_k c_l (b'c) + c_k c_l + b_k b_l) \\
&\quad + 4b_k c_l (b'c) + 2c_k c_l + 2b_k b_l, \\
\Phi_{k,MPH_d}^{(2,*)}(b, c) &= 3\kappa_{d,a}^{(1)} (2b_k (b'c)^2 + 2c_k (b'c) + b_k) \\
&\quad - 3\kappa_{d,a}^{(2)} (4c_k (b'c) + 2b_k) + 6c_k (b'c) + 3b_k.
\end{aligned}$$

Der Kovarianzmatrixkern des Grenzprozesses W^* ist also unter MPH_d für $b, c \in \mathcal{S}^{d-1}$ gegeben durch

$$\begin{aligned}
\sigma_{MPH_d}^*(b, c) &= \kappa_{d,a}^{(1)} \left[(2(b'c)^2 + 1) \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2(bb' + cc') & 3b \\ 3c' & 0 \end{pmatrix} \right. \\
&\quad \left. + 6(b'c)^2 \begin{pmatrix} 0_d & b \\ c' & (b'c) \end{pmatrix} + (b'c) \begin{pmatrix} 4(bc' + cb') & 6c \\ 6b' & 9 \end{pmatrix} \right] \\
&\quad - 2\kappa_{d,a}^{(2)} \left[(b'c) \begin{pmatrix} 4bc' & 6c \\ 6b' & 9 \end{pmatrix} + \begin{pmatrix} 2(bb' + cc') & 3b \\ 3c' & 0 \end{pmatrix} - A(b, c) \right] \\
&\quad + (b'c) \begin{pmatrix} 4bc' & 6c \\ 6b' & 9 \end{pmatrix} + \begin{pmatrix} 2(bb' + cc') & 3b \\ 3c' & 0 \end{pmatrix} - \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} - 2A(b, c),
\end{aligned}$$

wobei 0_d die $d \times d$ -Nullmatrix ist und $A(b, c) := \text{diag}(b_1^2 + c_1^2, \dots, b_d^2 + c_d^2, 0)$ gesetzt ist.

Bemerkung 5.2.1

Die multivariate Pearson Typ II-Verteilung geht beim Grenzübergang $a \rightarrow \infty$ in die multivariate Normalverteilung über. In diesem Fall gilt

$$\begin{aligned}
\kappa_{d,\infty}^{(1)} &:= \lim_{a \rightarrow \infty} \frac{(d + 2a + 2)^2}{(d + 2a + 4)(d + 2a + 6)} = 1 \quad \text{und} \\
\kappa_{d,\infty}^{(2)} &:= \lim_{a \rightarrow \infty} \frac{d + 2a + 2}{d + 2a + 4} = 1.
\end{aligned}$$

Einsetzen und Kürzen liefert das erwartete Ergebnis

$$\sigma_{MPH_d}^*(b, c) = \sigma_{N_d}^*(b, c).$$

5.3 Multivariate Pearson Typ VII-Verteilung

Ein Zufallsvektor X ist Pearson Typ VII verteilt ($X \sim MPVII_d(a, N, \mu, \Delta)$), falls X die Dichte

$$g(x) = \frac{\Gamma(N)}{\Gamma\left(N - \frac{d}{2}\right) (\pi a)^{\frac{d}{2}} |\Delta|^{\frac{1}{2}}} \left(1 + \frac{1}{a} (x - \mu)' \Delta^{-1} (x - \mu)\right)^{-N}$$

besitzt. Dabei gilt $N > \frac{d}{2}$, $a > 0$, $\mu \in \mathbb{R}^d$, und Δ ist eine symmetrische positiv semidefinite $d \times d$ -Matrix (siehe [13], Abschnitt 3.3 und [20], Abschnitt 6.2). Der Erwartungswertvektor und die Kovarianzmatrix sind gegeben durch $E(X) = \mu$ und

$$E((X - \mu)(X - \mu)') = \frac{a}{2N - d - 2} \Delta = \Sigma \quad (5.7)$$

für $N > \frac{d}{2} + 1$. Die Zufallsvariable $(X - \mu)' \Delta^{-1} (X - \mu)$ besitzt nach [13], S. 82, Formel (3.25), die Dichte

$$g_R(x) = \frac{1}{B\left(\frac{d}{2}, N - \frac{d}{2}\right)} a^{-\frac{d}{2}} x^{\frac{d}{2}-1} \left(1 + \frac{x}{a}\right)^{-N}, \quad x > 0. \quad (5.8)$$

Seien X_1, \dots, X_n unabhängig identisch $MPVII_d(a, N, \mu, \Delta)$ -verteilt. Damit gilt für $r > 0$ mit den üblichen Rechenregeln für die Beta- und Gammafunktion

$$\begin{aligned} E((X - \mu)' \Delta^{-1} (X - \mu))^r &= \int_0^\infty \frac{1}{B\left(\frac{d}{2}, N - \frac{d}{2}\right)} a^{-\frac{d}{2}} x^{\frac{d}{2}-1+r} \left(1 + \frac{x}{a}\right)^{-N} dx \\ &= \frac{1}{B\left(\frac{d}{2}, N - \frac{d}{2}\right)} \int_0^\infty a^{r-1} \left(\frac{x}{a}\right)^{\frac{d}{2}-1+r} \left(1 + \frac{x}{a}\right)^{-N} dx \\ &= \frac{a^r}{B\left(\frac{d}{2}, N - \frac{d}{2}\right)} \int_0^\infty v^{\frac{d}{2}-1+r} (1+v)^{-N} dv \\ &= a^r \frac{B\left(\frac{d}{2} + r, N - \frac{d}{2} - r\right)}{B\left(\frac{d}{2}, N - \frac{d}{2}\right)} \\ &= a^r \frac{\Gamma\left(\frac{d}{2} + r\right) \Gamma\left(N - \frac{d}{2} - r\right)}{\Gamma\left(\frac{d}{2}\right) \Gamma\left(N - \frac{d}{2}\right)} \\ &= a^r \frac{\prod_{j=0}^{r-1} \left(\frac{d}{2} + j\right)}{\prod_{j=1}^r \left(N - \frac{d}{2} - j\right)} \\ &= a^r \prod_{j=0}^{r-1} \frac{d + 2j}{2N - d - 2j - 2}. \end{aligned}$$

Das r -te Moment von $(X - \mu)' \Sigma^{-1} (X - \mu)$ ist nach dem eben Berechneten und (5.7) gegeben durch

$$\begin{aligned} m_{2r} &= E((X - \mu)' \Sigma^{-1} (X - \mu))^r \\ &= \left(\frac{2N - d - 2}{a}\right)^r E((X - \mu)' \Delta^{-1} (X - \mu))^r \\ &= (2N - d - 2)^r \prod_{j=0}^{r-1} \frac{d + 2j}{2N - d - 2j - 2}. \end{aligned}$$

Für $r \in \{2, 3\}$ ergibt sich also speziell

$$m_4 = d(d+2) \frac{2N-d-2}{2N-d-4},$$

$$m_6 = d(d+2)(d+4) \frac{(2N-d-2)^2}{(2N-d-4)(2N-d-6)}.$$

Auch hier führen wir abkürzend die beiden von der Dimension d und dem Parameter N abhängigen Konstanten

$$\kappa_{d,N}^{(1)} := \frac{(2N-d-2)^2}{(2N-d-4)(2N-d-6)} \quad \text{und}$$

$$\kappa_{d,N}^{(2)} := \frac{2N-d-2}{2N-d-4}$$

ein. Es ergeben sich die zur Charakterisierung des Kovarianzmatrixkerns benötigten Größen

$$\begin{aligned} \rho_{k,MPVII_d}^* (b, c) &= \kappa_{d,N}^{(1)} [2(b'c)^2 + 8b_k c_k (b'c) + 2(c_k^2 + b_k^2) + 1] \\ &\quad - 2\kappa_{d,N}^{(2)} [4b_k c_k (b'c) + c_k^2 + b_k^2] + 4b_k c_k (b'c) - 1, \\ \rho_{MPVII_d}^* (b, c) &= \kappa_{d,N}^{(1)} (6(b'c)^3 + 9(b'c)) - 18\kappa_{d,N}^{(2)} (b'c) + 9(b'c), \\ \Phi_{k,l,MPVII_d}^{(1,*)} (b, c) &= 2\kappa_{d,N}^{(1)} (2(b'c)(b_k c_l + c_k b_l) + (b_k b_l + c_k c_l)) \\ &\quad - 4\kappa_{d,N}^{(2)} (2b_k c_l (b'c) + c_k c_l + b_k b_l) \\ &\quad + 4b_k c_l (b'c) + 2c_k c_l + 2b_k b_l, \\ \Phi_{k,MPVII_d}^{(2,*)} (b, c) &= 3\kappa_{d,N}^{(1)} (2b_k (b'c)^2 + 2c_k (b'c) + b_k) \\ &\quad - 3\kappa_{d,N}^{(2)} (4c_k (b'c) + 2b_k) + 6c_k (b'c) + 3b_k. \end{aligned}$$

Die eben angegebenen Größen unterscheiden sich nur in den Konstanten von den Kovarianzausdrücken aus dem Beispiel für die Pearson Typ II-Verteilung, so dass der Kovarianzkern $\sigma_{MPVII_d}^* (b, c)$ dieselbe Gestalt hat wie der Kovarianzkern $\sigma_{MPII_d}^* (b, c)$, wobei wir $\kappa_{d,a}^{(j)}$ durch $\kappa_{d,N}^{(j)}$ für $j = 1, 2$ ersetzen.

Bemerkung 5.3.1

1. Die multivariate Pearson Typ VII-Verteilung beinhaltet für $N = \frac{d+1}{2}$ die multivariate Cauchy Verteilung. Dabei existiert das erste Moment $E(X)$ für $X \sim MPVII_d(a, \frac{d+1}{2}, \mu, \Delta)$ nicht.
2. Die multivariate Pearson Typ VII-Verteilung geht beim Grenzübergang $a \rightarrow \infty$ in die multivariate Normalverteilung über. In diesem Fall gilt

$$\kappa_{d,\infty}^{(1)} := \lim_{N \rightarrow \infty} \frac{(2N-d-2)^2}{(2N-d-4)(2N-d-6)} = 1 \quad \text{und}$$

$$\kappa_{d,\infty}^{(2)} := \lim_{N \rightarrow \infty} \frac{2N-d-2}{2N-d-4} = 1.$$

Auch hier liefert Einsetzen und Kürzen das erwartete Ergebnis

$$\sigma_{MPVII_d}^* (b, c) = \sigma_{N_d}^* (b, c).$$

Kapitel 6

Konvergenzaussagen unter benachbarten Alternativen

In diesem Kapitel untersuchen wir das Konvergenzverhalten der Teststatistik T_n unter benachbarten Alternativen. Zur Präzisierung betrachten wir ein Dreiecksschema X_{n1}, \dots, X_{nn} , $n \geq d + 1$, von zeilenweise unabhängig identisch verteilten d -dimensionaler Zufallsvektoren, mit zugrunde liegender Lebesgue-Dichte

$$f_n(x) = \phi(x) \left(1 + \frac{h(x)}{\sqrt{n}} \right), \quad x \in \mathbb{R}^d. \quad (6.1)$$

Dabei bezeichnet ϕ die Dichte der d -dimensionalen Standardnormalverteilung, und h ist eine beschränkte messbare Funktion mit der Eigenschaft $\int_{\mathbb{R}^d} h(x)\phi(x) dx = 0$. Wir nehmen weiter an, dass n groß genug ist, um die Nichtnegativität von f_n zu gewährleisten. Da Verwechslungen ausgeschlossen werden können, verwenden wir wie in den letzten Abschnitten die Standardnotationen für den empirischen Mittelwert, die empirische Kovarianzmatrix und die empirisch standardisierten Daten, schreiben also

$$\begin{aligned} \bar{X}_n &= \frac{1}{n} \sum_{j=1}^n X_{nj}, \\ S_n &= \frac{1}{n} \sum_{j=1}^n (X_{nj} - \bar{X}_n)(X_{nj} - \bar{X}_n)', \\ Y_{n,j} &= S_n^{-\frac{1}{2}}(X_{nj} - \bar{X}_n), \quad j = 1, \dots, n. \end{aligned}$$

Bevor Aussagen über die Grenzverteilung der Teststatistik unter benachbarten Alternativen gemacht werden können, formulieren und beweisen wir ein wichtiges Hilfsresultat. Hierzu betrachten wir die Wahrscheinlichkeitsmaße

$$\mathbb{P}^{(n)} := \bigotimes_{j=1}^n (\phi \lambda^d) \quad \text{und} \quad \mathbb{A}^{(n)} := \bigotimes_{j=1}^n (f_n \lambda^d)$$

auf dem Messraum $(\mathfrak{X}_n, \mathfrak{B}_n) := \bigotimes_{j=1}^n (\mathbb{R}^d, \mathcal{B}^d)$ und den Dichtequotienten $L_n := \frac{d\mathbb{A}^{(n)}}{d\mathbb{P}^{(n)}}$.

Lemma 6.1.1

Unter der Folge von Wahrscheinlichkeitsmaßen $(\mathbb{P}^{(n)})_{n \in \mathbb{N}}$ gilt für $n \rightarrow \infty$

$$\log L_n \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{\sigma^2}{2}, \sigma^2\right),$$

mit $\sigma^2 := \int_{\mathbb{R}^d} h^2(x) \phi(x) dx < \infty$.

Beweis:

Unter Berücksichtigung der Taylorreihenentwicklung des Logarithmus um den Entwicklungspunkt $x_0 = 1$ gilt

$$\begin{aligned} \log L_n(X_{n1}, \dots, X_{nm}) &= \sum_{j=1}^n \log\left(1 + \frac{h(X_{nj})}{\sqrt{n}}\right) \\ &= \sum_{j=1}^n \left\{ \frac{h(X_{nj})}{\sqrt{n}} - \frac{h^2(X_{nj})}{2n} + R_{n,j} \right\}, \end{aligned}$$

wobei

$$R_{n,j} = \frac{1}{3!} \frac{2}{(1 + \eta_{n,j})^3} \frac{h^3(X_{n,j})}{n^{\frac{3}{2}}}, \quad (6.2)$$

mit $|\eta_{n,j}| \leq \frac{|h(X_{n,j})|}{\sqrt{n}}$. Wegen der Beschränktheit der Funktion h gilt $\sum_{j=1}^n R_{n,j} = o_{\mathbb{P}^{(n)}}(1)$.

Weiter ist für $n \rightarrow \infty$

$$E\left(\frac{1}{n} \sum_{j=1}^n h^2(X_{n,j})\right) \rightarrow \sigma^2 \text{ und } V\left(\frac{1}{n} \sum_{j=1}^n h^2(X_{n,j})\right) \rightarrow 0.$$

Also folgt nach dem zentralen Grenzwertsatz von Lindeberg-Feller die Behauptung. \square

Für die Beweise der beiden folgenden Theoreme verwenden wir einige Aussagen aus der Theorie benachbarter Verteilungsfolgen (siehe [29], Kapitel 6.3 und Anhang D.1). Hierbei ist zu beachten, dass wegen Lemma 6.1.1 nach Le Cams erstem Lemma (siehe Theorem D.1.3) die Verteilungsfolge $\mathbb{A}^{(n)}$ wechselseitig benachbart zu $\mathbb{P}^{(n)}$ ist.

Um den Schreibaufwand übersichtlich zu halten, schreiben wir

$$\mathcal{G}_k : \mathbb{R}^d \times \mathcal{S}^{d-1} \rightarrow \mathbb{R}, (x, a) \mapsto \mathcal{G}_k(x, a) := f_k(x, a) - 2a'xa_k - \varphi_k(x),$$

mit f_k und φ_k wie in (2.11). Dabei gilt für den Hilfsprozess $C_{n,k}^*$ aus (4.6)

$$C_{n,k}^*(b) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathcal{G}_k(X_{n,j}, b), \quad b \in \mathcal{S}^{d-1}.$$

Theorem 6.1.2

Unter den oben gemachten Voraussetzungen gilt unter dem Dreiecksschema X_{n1}, \dots, X_{nn}

$$C_{n,k} \xrightarrow{\mathcal{D}} C_k^* + \mathcal{C}_k$$

in $C(\mathcal{S}^{d-1}, \mathbb{R})$, wobei C_k^* ein zentrierter Gaußprozess in $C(\mathcal{S}^{d-1}, \mathbb{R})$ mit Kovarianzkern $\rho_{k, \mathcal{N}_d}^*$ aus (5.2) ist. Der Shiftparameter \mathcal{C}_k ist gegeben durch

$$\mathcal{C}_k(b) = \int_{\mathbb{R}^d} \mathcal{G}_k(u, b) h(u) \phi(u) du, \quad b \in \mathcal{S}^{d-1}. \quad (6.3)$$

Beweis:

Der Beweis ist in zwei Schritte gegliedert. Zunächst ist Folgendes zu zeigen:

- (a) Unter $\mathbb{P}^{(n)}$ gilt $\lim_{n \rightarrow \infty} \text{Cov}(C_{n,k}^*(b), \log L_n) = \mathcal{C}_k(b)$,
- (b) Für $l \in \mathbb{N}$ und Vektoren $\hat{b}_1, \dots, \hat{b}_l \in \mathcal{S}^{d-1}$ konvergiert unter $\mathbb{P}^{(n)}$ die Folge der gemeinsamen Verteilungen von $C_{n,k}^*(\hat{b}_1), \dots, C_{n,k}^*(\hat{b}_l)$ und $\log L_n$ gegen die $(l+1)$ -dimensionale Normalverteilung

$$\mathcal{N}_{l+1} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{\sigma^2}{2} \end{pmatrix}, \begin{pmatrix} \Sigma_{\hat{b}_1, \dots, \hat{b}_l} & c_k \\ c_k' & \sigma^2 \end{pmatrix} \right),$$

mit $\Sigma_{\hat{b}_1, \dots, \hat{b}_l} := \left(\rho_{k, \mathcal{N}_d}^*(\hat{b}_{j_1}, \hat{b}_{j_2}) \right)_{1 \leq j_1, j_2 \leq l}$ und $c_k := \left(\mathcal{C}_k(\hat{b}_1), \dots, \mathcal{C}_k(\hat{b}_l) \right)'$, wobei $\rho_{k, \mathcal{N}_d}^*$ der Kovarianzkern aus (5.2) ist.

Unter der Annahme, dass (a) und (b) gelten, folgt dann mit Lemma 6.1.1 und Le Cams drittem Lemma (siehe D.1.5), dass die endlichdimensionalen Verteilungen von $C_{n,k}^*$ unter $\mathbb{A}^{(n)}$ gegen die endlichdimensionalen Verteilungen des verschobenen Gaußschen Grenzprozesses $C_k^* + \mathcal{C}_k$ konvergieren. Da $C_{n,k}^*$ unter $\mathbb{P}^{(n)}$ straff ist und $\mathbb{P}^{(n)}$ und $\mathbb{A}^{(n)}$ benachbart sind, ist $C_{n,k}^*$ auch unter $\mathbb{A}^{(n)}$ straff, und es gilt unter $\mathbb{A}^{(n)}$

$$C_{n,k}^* \xrightarrow{\mathcal{D}} C_k^* + \mathcal{C}_k.$$

Aus Lemma 4.1.3 und Lemma 4.1.4 folgt

$$\|C_{n,k} - C_{n,k}^*\|_C \xrightarrow{\mathbb{P}^{(n)}} 0.$$

Wegen der Benachbarkeit von $\mathbb{P}^{(n)}$ und $\mathbb{A}^{(n)}$ gilt auch

$$\|C_{n,k} - C_{n,k}^*\|_C \xrightarrow{\mathbb{A}^{(n)}} 0$$

und demnach die Behauptung. Es bleiben die Aussagen (a) und (b) zu zeigen.

(a) Es gilt wegen $E(C_{n,k}^*(b)) = 0$ unter $\mathbb{P}^{(n)}$ für $b \in \mathcal{S}^{d-1}$ mit (6.2)

$$\begin{aligned} \text{Cov}(C_{n,k}^*(b), \log L_n) &= E(C_{n,k}^*(b) \log L_n) \\ &= E \left(C_{n,k}^*(b) \left[\sum_{j=1}^n \left\{ \frac{h(X_{nj})}{\sqrt{n}} - \frac{h^2(X_{nj})}{2n} + R_{n,j} \right\} \right] \right), \end{aligned}$$

mit $R_{n,j}$ wie in (6.2). Dabei existiert wegen $|h(X_{nj})|^3 \leq \|h\|^3$ und $|\eta_{n,j}| \leq \frac{\|h\|}{\sqrt{n}}$ eine Konstante $C > 0$ mit

$$\begin{aligned} \sum_{j=1}^n |R_{n,j}| &= \frac{2}{3!} \frac{1}{n^{\frac{3}{2}}} \sum_{j=1}^n |h(X_{nj})|^3 \frac{1}{(1 + \eta_{n,j})^3} \\ &\leq \frac{C}{\sqrt{n}}. \end{aligned}$$

Also ist

$$\left| E \left(C_{n,k}^*(b) \sum_{j=1}^n R_{n,j} \right) \right| \leq \frac{C}{\sqrt{n}} E |C_{n,k}^*(b)| \leq \frac{C}{\sqrt{n}} \sqrt{E (C_{n,k}^*(b))^2} \longrightarrow 0$$

für $n \rightarrow \infty$. Wegen

$$\frac{1}{2n} E \left(C_{n,k}^*(b) \sum_{j=1}^n h^2(X_{nj}) \right) \longrightarrow 0$$

für $n \rightarrow \infty$ und

$$\begin{aligned} \frac{1}{\sqrt{n}} E \left(C_{n,k}^*(b) \sum_{j=1}^n h(X_{nj}) \right) &= \frac{1}{n} \sum_{j_1, j_2=1}^n E (\mathcal{G}_k(X_{nj_1}, b) h(X_{nj_2})) \\ &= E (\mathcal{G}_k(X_{n1}, b) h(X_{n1})) \\ &= \int_{\mathbb{R}^d} \mathcal{G}_k(u, b) h(u) \phi(u) du \\ &= \mathcal{C}_k(b) \end{aligned}$$

folgt die behauptete Grenzwertaussage.

(b) Die Grenzverteilung des $(l+1)$ -dimensionalen Zufallsvektors

$$\left(C_{n,k}^*(\hat{b}_1), \dots, C_{n,k}^*(\hat{b}_l), \log L_n \right)'$$

ergibt sich direkt aus dem multivariaten zentralen Grenzwertsatz, wobei die Parameter der Normalverteilung durch die Aussagen von Theorem 4.1.1 und Lemma 6.1.1 eindeutig bestimmt sind. \square

In analoger Weise können wir unter derselben Folge von Alternativen $\mathbb{A}^{(n)}$ eine Konvergenzaussage für den stochastischen Prozess D_n aus (4.2) formulieren. Hierzu schreiben wir

$$\mathcal{J} : \mathbb{R}^d \times \mathcal{S}^{d-1} \rightarrow \mathbb{R}, (x, a) \mapsto \mathcal{J}(x, a) := (a'x)^3 - 3a'x.$$

Damit lässt sich der Hilfsprozess D_n^* aus (4.20) durch

$$D_n^*(b) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathcal{J}(X_j, b), \quad b \in \mathcal{S}^{d-1},$$

darstellen.

Theorem 6.1.3

Unter den oben gemachten Voraussetzungen gilt unter dem Dreiecksschema X_{n1}, \dots, X_{nn}

$$D_n \xrightarrow{\mathcal{D}} D^* + \mathcal{D}$$

in $C(\mathcal{S}^{d-1}, \mathbb{R})$, wobei D^* ein zentrierter Gaußprozess in $C(\mathcal{S}^{d-1}, \mathbb{R})$ mit Kovarianzkern $\rho_{\mathcal{N}_d}^*$ aus (5.3) ist. Der Shiftparameter \mathcal{D} ist gegeben durch

$$\mathcal{D}(b) = \int_{\mathbb{R}^d} \mathcal{J}(u, b) h(u) \phi(u) \, du, \quad b \in \mathcal{S}^{d-1}. \quad (6.4)$$

Beweis:

Analog zum Beweis von Theorem 6.1.2. □

Bemerkung 6.1.4

Definieren wir den stochastischen Prozess

$$Z_n(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathcal{J}(X_{nj}, b),$$

so erhalten wir denselben stochastischen Prozess wie in [4], S. 53. Für diesen Prozess gilt wegen $\|Z_n - D_n\|_C \xrightarrow{\mathbb{P}} 0$ unter $\mathbb{P}^{(n)}$ (siehe hierzu die Lemmata 4.2.3 und 4.2.4) die gleiche Konvergenzaussage wie für D_n in Theorem 6.1.3. Daraus folgt das Konvergenzverhalten des Prozesses Z_n unter benachbarten Alternativen. Setzen wir

$$b_{1,n}(u) = \frac{\left(\frac{1}{n} \sum_{j=1}^n (u' X_j - u' \bar{X}_n)^3 \right)^2}{(u' S_n u)^3},$$

so ist die Teststatistik von Malkovich und Afifi durch

$$b_{1,n}^M = \max_{u \in \mathcal{S}^{d-1}} b_{1,n}(u)$$

gegeben, siehe hierzu auch Abschnitt 8.3. Unter den Voraussetzungen von Theorem 6.1.3 und unter Berücksichtigung des Theorems 2.6 aus [4] und der Benachbarkeit der Alternativen gilt

$$nb_{1,n}^M \xrightarrow{\mathcal{D}} \max_{u \in \mathcal{S}^{d-1}} (D^*(u) + \mathcal{D}(u))^2,$$

mit $\mathcal{D}(u)$ wie in (6.4).

Mit den eben gezeigten Resultaten lassen sich weiterführende Erkenntnisse zum asymptotischen Verhalten des multivariaten Multiparameterprozesses W_n aus (4.3) unter benachbarten Alternativen gewinnen. Dabei bezeichne

$$\mathcal{W}(b) = \begin{pmatrix} \mathcal{C}_1(b) \\ \vdots \\ \mathcal{C}_d(b) \\ \mathcal{D}(b) \end{pmatrix}, \quad b \in \mathcal{S}^{d-1}, \quad (6.5)$$

mit $\mathcal{C}_1(b), \dots, \mathcal{C}_d(b)$ wie in (6.3) und $\mathcal{D}(b)$ aus (6.4). Wie das folgende Theorem zeigt, konvergiert der stochastische Prozess W_n unter den gemachten Voraussetzungen gegen einen um $\mathcal{W}(b)$ verschobenen Grenzprozess W^* .

Theorem 6.1.5

Unter dem Dreiecksschema X_{n1}, \dots, X_{nn} ist unter den gemachten Voraussetzungen

$$W_n \xrightarrow{\mathcal{D}} W^* + \mathcal{W}$$

in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$, wobei W^* ein zentrierter Gaußprozess in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ mit Kovarianzkern $\sigma_{\mathcal{N}_d}^*(b, c)$ aus (5.4) ist. Der Shiftparametervektor \mathcal{W} ist gegeben durch (6.5).

Beweis:

Aus den Beweisen der Theoreme 6.1.2 und 6.1.3 gilt unter $\mathbb{P}^{(n)}$ für $k = 1, \dots, d$

1. $\lim_{n \rightarrow \infty} \text{Cov}(C_{n,k}^*(b), \log L_n) = \mathcal{C}_k(b)$,
2. $\lim_{n \rightarrow \infty} \text{Cov}(D_n^*(b), \log L_n) = \mathcal{D}(b)$

für $b \in \mathcal{S}^{d-1}$. Weiter gilt für $l \in \mathbb{N}$ und $\hat{b}_1, \dots, \hat{b}_l \in \mathcal{S}^{d-1}$, dass die gemeinsame Verteilung von $W_n(\hat{b}_1), \dots, W_n(\hat{b}_l)$ und $\log L_n$ gegen die $l(d+1)+1$ -dimensionale Normalverteilung

$$\mathcal{N}_{l(d+1)+1} \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{\sigma^2}{2} \end{pmatrix}, \begin{pmatrix} \sigma_{\hat{b}_1, \dots, \hat{b}_l} & w \\ w' & \sigma^2 \end{pmatrix} \right),$$

mit der $l(d+1) \times l(d+1)$ Matrix $\sigma_{\hat{b}_1, \dots, \hat{b}_l} := \left(\sigma_{\mathcal{N}_d}^*(\hat{b}_{j_1}, \hat{b}_{j_2}) \right)_{1 \leq j_1, j_2 \leq l}$ und dem $l(d+1)$ -dimensionalen Vektor $w := \left(\mathcal{W}(\hat{b}_1)', \dots, \mathcal{W}(\hat{b}_l)' \right)'$ konvergieren, wobei $\sigma_{\mathcal{N}_d}^*$ der Kovarianzmatrixkern aus (5.4) ist. Es folgt die Behauptung unter derselben Argumentation wie im Beweis zu Theorem 6.1.2. \square

Aus den Aussagen der eben vorgestellten Theoreme lässt sich das Verhalten der Teststatistik T_n unter benachbarten Alternativen charakterisieren.

Theorem 6.1.6

Es sei das Dreiecksschema X_{n1}, \dots, X_{nn} , $n \geq d + 1$, mit zugrunde liegender Lebesgue-Dichte f_n aus (6.1) gegeben. Dann gilt für die in (2.5) angegebene Teststatistik von Cox und Small

$$nT_n \xrightarrow{\mathcal{D}} \frac{1}{2} \max_{b \in \mathcal{S}^{d-1}} \Psi_B(W^*(b) + \mathcal{W}(b))$$

für $n \rightarrow \infty$. Dabei sind Ψ_B die Abbildung aus (2.14), $W^*(\cdot)$ der $(d + 1)$ -dimensionale zentrierte Gaußsche Grenzprozess aus Theorem 4.3.1 und $\mathcal{W}(\cdot)$ der Shiftparametervektor aus (6.5).

Beweis:

Es gilt nach dem Beweis zu Theorem 4.4.1 und mit (5.1)

$$\sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b'Y_{n,j})^4 - 3 \right| \xrightarrow{\mathbb{P}^{(n)}} 0 \quad \text{und} \quad \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b'Y_{n,j})^3 \right| \xrightarrow{\mathbb{P}^{(n)}} 0$$

für alle $b \in \mathcal{S}^{d-1}$. Wegen der Benachbarkeit von $\mathbb{A}^{(n)}$ und $\mathbb{P}^{(n)}$ gelten die stochastischen Konvergenzen auch unter $\mathbb{A}^{(n)}$. Die Behauptung folgt aus dem Lemma von Slutsky und zweifacher Anwendung des Abbildungssatzes. \square

Kapitel 7

Konvergenzaussagen unter festen Alternativen

In diesem Kapitel werden die Konvergenzaussagen der vorangehenden Kapitel unter Aufgabe der Annahme der sphärischen Symmetrie an die zugrunde liegende Verteilung verallgemeinert. Hierzu seien im Folgenden X_1, \dots, X_n unabhängig identisch verteilte d -dimensionale Zufallsvektoren mit $E\|X_1\|^6 < \infty$. Weiter belege die Verteilung von X_1 jede $(d-1)$ -dimensionale Hyperebene des \mathbb{R}^d mit Masse 0 und es gelte $n > d$, damit die empirische Kovarianzmatrix S_n fast sicher nicht singulär ist. Im Hinblick auf die affine Invarianz der Teststatistik können die Zufallsvektoren o.B.d.A. als standardisiert angenommen werden, d.h. es gilt $E(X_1) = 0$ und $E(X_1 X_1') = I_d$. Unter den getroffenen Annahmen existieren alle auftretenden gemischten Momente. Das Problem im Vergleich zu den Ausführungen aus Kapitel 4 sind die unter einer solch allgemeinen Verteilungsannahme nicht mehr explizit angebbaren Ausdrücke für die auftretenden Erwartungswerte und der Verlust der Zentrierung der betrachteten stochastischen Prozesse.

Der Übersichtlichkeit halber werden im Folgenden die in den theoretischen Ausführungen verwendeten Abbildungen definiert. Dabei bezeichnen für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned}\mu_1^\star(b) &:= E(\varphi_k(X_1)(b'X_1)X_1), \\ \mu_2^\star(b) &:= E((b'X_1)^2 X_1),\end{aligned}$$

d -dimensionale Vektoren und

$$\begin{aligned}\mu_3^\diamond(b) &:= E(b'X_1)^3, \\ \mu_4^\diamond(b) &:= E(b'X_1)^4,\end{aligned}$$

reelle Zahlen. Wir weisen darauf hin, dass unter diesen Verteilungsannahmen die gemischten Momente

$$E(\varphi_k(X_1)(b'X_1)) = b_k \quad \text{und} \quad E(b'X_1)^2 = 1$$

explizit angebar sind, und dass zwischen $\mu_3^\diamond(b)$ und $\mu_2^\star(b)$ die Beziehung

$$\mu_3^\diamond(b) = E(b'X_1)^3 = E((b'X_1)^2(b'X_1)) = b'E((b'X_1)^2 X_1) = b'\mu_2^\star(b)$$

besteht. Weiter gilt für $b \in \mathcal{S}^{d-1}$ die Identität

$$E(f_k(X_1, b)) = \varphi_k(\mu_2^\star(b)) = e'_k \mu_2^\star(b) = \varphi_k \circ \mu_2^\star(b). \quad (7.1)$$

7.1 Stochastische Konvergenz der Teststatistik T_n unter festen Alternativen

Damit die stochastische Konvergenz der Teststatistik T_n gezeigt werden kann, muss das Grenzverhalten von T_n unter festen Alternativen untersucht werden. Hierbei geht auch wie in den vorhergehenden Kapiteln die schwache Konvergenz des in (4.3) definierten, multivariaten Multiparameterprozesses W_n entscheidend in die Betrachtungen ein. Unter den gemachten Annahmen muss W_n ebenfalls zentriert werden. Dabei bezeichne

$$\mu(b) := E \begin{pmatrix} f_1(X_1, b) \\ \vdots \\ f_d(X_1, b) \\ (b'X_1)^3 \end{pmatrix}, \quad b \in \mathcal{S}^{d-1}, \quad (7.2)$$

wobei nach den Voraussetzungen $\mu(b)$ existiert und $\|\mu(b)\| < \infty$ für alle $b \in \mathcal{S}^{d-1}$ gilt (für eine explizite Darstellung von μ siehe auch (7.16)). Der Beweis des folgenden Theorems benötigt die Straffheit von $\sqrt{n}(W_n(\cdot) - \mu(\cdot))$ in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$. Auf Grund einer besseren Lesbarkeit wird dies in Theorem 7.4.1 im Anschluss gezeigt.

Die im Folgenden gemachten Aussagen beruhen auf der stochastischen Konvergenz der Teststatistik unter einer festen Alternative gegen eine von dieser Alternative abhängenden Konstante. Wie das folgende Theorem zeigt, ist diese durch

$$\mathcal{L} := \max_{b \in \mathcal{S}^{d-1}} \frac{\Psi_B(\mu(b))}{\mu_4^\star(b) - 1 - (\mu_3^\star(b))^2}$$

gegeben.

Theorem 7.1.1

Für die Teststatistik T_n aus (2.5) gilt

$$T_n \xrightarrow{\mathbb{P}} \mathcal{L}$$

für $n \rightarrow \infty$.

Beweis:

Es gilt

$$\begin{aligned} nT_n &= \max_{b \in \mathcal{S}^{d-1}} \frac{\Psi_B(W_n(b))}{\mathcal{V}_n(b)} \\ &= \max_{b \in \mathcal{S}^{d-1}} \left\{ \frac{1}{\mathcal{V}_n(b)} \left(\Psi_B(W_n(b) - \sqrt{n}\mu(b)) + 2\sqrt{n}\mu(b)'B(W_n(b) - \sqrt{n}\mu(b)) \right. \right. \\ &\quad \left. \left. + n\mu(b)'B\mu(b) \right) \right\} \\ &= \max_{b \in \mathcal{S}^{d-1}} \left\{ \frac{1}{\mathcal{V}_n(b)} (\mathcal{W}_n(b) + \mathcal{Z}_n(b)) \right\} \end{aligned}$$

mit

$$\begin{aligned}\mathscr{W}_n(b) &:= \Psi_B(W_n(b) - \sqrt{n}\mu(b)), \\ \mathscr{Y}_n(b) &:= 2\sqrt{n}\mu(b)'B(W_n(b) - \sqrt{n}\mu(b)) + n\mu(b)'B\mu(b),\end{aligned}$$

und

$$\mathscr{V}_n(b) := \frac{1}{n} \sum_{j=1}^n (b'Y_{n,j})^4 - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b'Y_{n,j})^3 \right)^2$$

für $b \in \mathcal{S}^{d-1}$. Nach Theorem 7.4.1 gilt

$$\sqrt{n} \left(\frac{1}{\sqrt{n}} W_n(\cdot) - \mu(\cdot) \right) \xrightarrow{\mathcal{D}} W^\star(b)$$

mit einem zentrierten Gaußprozess W^\star in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$. Insbesondere gilt nach dem Abbildungssatz

$$\mathscr{W}_n(\cdot) \xrightarrow{\mathcal{D}} \Psi_B(W^\star(\cdot))$$

und wegen der Straffheit der Folge \mathscr{W}_n folgt

$$\frac{1}{n} \mathscr{W}_n(\cdot) \xrightarrow{\mathbb{P}} 0 \tag{7.3}$$

in $C(\mathcal{S}^{d-1}, \mathbb{R})$. In ähnlicher Weise schließen wir auf

$$\frac{1}{n} \mathscr{Y}_n(\cdot) \xrightarrow{\mathbb{P}} \mu(\cdot)'B\mu(\cdot) = \Psi_B(\mu(\cdot)). \tag{7.4}$$

Für $\mathscr{V}_n(b)$ ist mit Hilfe der Identitäten für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned}(b'Y_{n,j})^4 &= (b'X_j)^4 + \eta_{n,j}(b), \\ (b'Y_{n,j})^3 &= (b'X_j)^3 + \tilde{\eta}_{n,j}(b),\end{aligned}$$

wobei

$$\begin{aligned}|\eta_{n,j}(b)| &\leq \|\Delta_{n,j}\|^4 + 4\|X_j\| \|\Delta_{n,j}\|^3 + 6\|X_j\|^2 \|\Delta_{n,j}\|^2 + 4\|X_j\|^3 \|\Delta_{n,j}\|, \\ |\tilde{\eta}_{n,j}(b)| &\leq \|\Delta_{n,j}\|^3 + 3\|X_j\| \|\Delta_{n,j}\|^2 + 3\|X_j\|^2 \|\Delta_{n,j}\|,\end{aligned}$$

mit $\Delta_{n,j}$ wie in (4.4), gilt. Aus

$$\begin{aligned}\mathscr{V}_n(b) &= \frac{1}{n} \sum_{j=1}^n [(b'X_j)^4 + \eta_{n,j}(b)] - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b'X_j)^3 + \tilde{\eta}_{n,j}(b) \right)^2 \\ &= \frac{1}{n} \sum_{j=1}^n (b'X_j)^4 + \frac{1}{n} \sum_{j=1}^n \eta_{n,j}(b) - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b'X_j)^3 \right)^2 \\ &\quad - \frac{2}{n} \sum_{j=1}^n (b'X_j)^3 \frac{1}{n} \sum_{l=1}^n \tilde{\eta}_{n,l}(b) - \left(\frac{1}{n} \sum_{j=1}^n \tilde{\eta}_{n,j}(b) \right)^2\end{aligned}$$

und den Resultaten aus Lemma 4.1.3 und Theorem 4.3.2 folgt

$$\sup_{b \in \mathcal{S}^{d-1}} \left| \mathcal{V}_n(b) - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 \right)^2 \right) \right| \xrightarrow{\mathbb{P}} 0. \quad (7.5)$$

Wir betrachten nun den durch

$$\mathcal{X}_n(b) := \frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - 1 - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 \right)^2, \quad b \in \mathcal{S}^{d-1},$$

definierten stochastischen Prozess auf der Einheitssphäre. Weiter bezeichnen wir mit $\tilde{\mu}$ die stetige Abbildung

$$\tilde{\mu} : \mathcal{S}^{d-1} \rightarrow \mathbb{R}, \quad b \mapsto \tilde{\mu}(b) := \mu_4^\diamond(b) - 1 - (\mu_3^\diamond(b))^2.$$

Wegen (7.5) gilt

$$\|\mathcal{X}_n - \mathcal{V}_n\|_C \xrightarrow{\mathbb{P}} 0,$$

und nach dem Gesetz großer Zahlen und dem Abbildungssatz folgt

$$\begin{aligned} \|\mathcal{X}_n - \tilde{\mu}\|_C &= \sup_{b \in \mathcal{S}^{d-1}} |\mathcal{X}_n(b) - \tilde{\mu}(b)| \\ &= \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - \mu_4^\diamond(b) + (\mu_3^\diamond(b))^2 - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 \right)^2 \right| \\ &\leq \sup_{b \in \mathcal{S}^{d-1}} \left| \frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - \mu_4^\diamond(b) \right| + \sup_{b \in \mathcal{S}^{d-1}} \left| (\mu_3^\diamond(b))^2 - \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 \right)^2 \right| \\ &\xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Dabei folgt die stochastische Konvergenz aus der schwachen Konvergenz der Prozesse

$$\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^4 - \mu_4^\diamond(b) \right), \quad b \in \mathcal{S}^{d-1},$$

und

$$\sqrt{n} \left(\frac{1}{n} \sum_{j=1}^n (b' X_j)^3 - \mu_3^\diamond(b) \right), \quad b \in \mathcal{S}^{d-1},$$

gegen entsprechende Grenzprozesse in $C(\mathcal{S}^{d-1}, \mathbb{R})$. Damit gilt

$$\mathcal{V}_n(\cdot) \xrightarrow{\mathbb{P}} \tilde{\mu}(\cdot) \quad (7.6)$$

in $C(\mathcal{S}^{d-1}, \mathbb{R})$. Zusammengesetzt ergibt dies nach Anwendung des Abbildungssatzes

$$T_n = \max_{b \in \mathcal{S}^{d-1}} \left\{ \frac{1}{\mathcal{Y}'_n(b)} \left(\frac{1}{n} \mathcal{W}_n(b) + \frac{1}{n} \mathcal{Z}_n(b) \right) \right\} \xrightarrow{\mathbb{P}} \mathcal{L}$$

und das war zu zeigen. \square

Die richtige Interpretation der Konstanten \mathcal{L} ist in diesem Fall wichtig für Konsistenzaussagen des mit T_n als Prüfgröße operierenden Tests. Für den Zähler von \mathcal{L} können einfache Aussagen getroffen werden.

Bemerkung 7.1.2

1. Für $b \in \mathcal{S}^{d-1}$ gilt

$$\Psi_B(\mu(b)) = \sum_{k=1}^d [E((b'X_1)^2 e'_k X_1)]^2 - [E(b'X_1)^3]^2 = \|E((b'X_1)^2 X_1)\|^2 - [E(b'X_1)^3]^2,$$

und wegen der Cauchy-Schwarzschen Ungleichung gilt

$$\|E((b'X_1)^2 X_1)\|^2 = \|b\|^2 \|E((b'X_1)^2 X_1)\|^2 \geq |b' E((b'X_1)^2 X_1)|^2 = [E(b'X_1)^3]^2.$$

Somit gilt

$$\Psi_B(\mu(b)) \geq 0, \quad b \in \mathcal{S}^{d-1}$$

und folglich auch

$$\max_{b \in \mathcal{S}^{d-1}} \Psi_B(\mu(b)) \geq 0.$$

Auf Grund der Abschätzung durch die Cauchy-Schwarzsche Ungleichung können wir die folgende Aussagen treffen. Die Größe $\Psi_B(\mu(b))$ ist genau dann gleich 0, falls

- (a) $E((b'X_1)^2 X_1) = 0$ oder
- (b) $E((b'X_1)^2 X_1)$ und $b \in \mathcal{S}^{d-1}$ linear abhängig sind.

2. Unter der Hypothese gilt $\mathcal{L} = 0$.

3. Falls eine sphärisch-symmetrische Alternative mit existierendem sechsten Moment der Radialverteilung zugrunde liegt (siehe Theorem 4.4.1) gilt $\mathcal{L} = 0$.

Es bleibt zu klären, unter welchen Alternativen \mathcal{L} den Wert 0 annimmt.

Mit Hilfe des Theorems 7.1.1 lassen sich erste Konsistenzaussagen treffen. Für feste Alternativen bei denen ein $\tilde{b} \in \mathcal{S}^{d-1}$ existiert, so dass

$$\frac{\Psi_B(\mu(\tilde{b}))}{\mu_4^\star(\tilde{b}) - 1 - (\mu_3^\star(\tilde{b}))^2} > 0$$

gilt, ist der mit T_n als Prüfgröße operierende Test konsistent, da für jedes feste $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}(nT_n > \varepsilon) = 1$$

gilt. Nach Theorem 4.4.1 ist die Teststatistik T_n offensichtlich nicht konsistent gegen sphärisch-symmetrische Alternativen mit existierendem sechsten Moment der Radienverteilung, da sie dasselbe Grenzverhalten aufweist wie unter der Hypothese H_0 .

Die Verteilungskonvergenz der Teststatistik T_n unter festen Alternativen konnte nicht gezeigt werden. Allerdings liegt auf Grund der Berechnungen zu Theorem 7.1.1 die folgende Vermutung nahe.

Vermutung 7.1.3

Unter den gemachten Voraussetzungen von oben gilt

$$\sqrt{n}(T_n - \mathcal{L}) \xrightarrow{\mathcal{D}} 2 \max_{b \in \mathcal{S}^{d-1}} \mu(b)' BW^\star(b),$$

wobei W^\star der zentrierte Gaußprozess in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ aus Theorem 7.4.1 ist.

Die in diesem Kapitel gemachten Aussagen beruhen auf der schwachen Konvergenz des geeignet zentrierten und normierten stochastischen Prozesses $W_n(\cdot)$ im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$ gegen ein zentriertes Gaußsches Zufallselement W^\star mit bestimmter Kovarianzstruktur. Diese Aussage wird in den folgenden Unterabschnitten bewiesen. Wie in den Kapiteln 3, 4 und 6 nähern wir uns dieser Aufgabe über die schwache Konvergenz der Komponentenprozesse.

7.2 Konvergenzaussagen für $C_{n,k}$

Dieser Abschnitt soll das Konvergenzverhalten des stochastischen Prozesses $C_{n,k}$ unter den oben gemachten Voraussetzungen klären. Das Vorgehen ist dem in Kapitel 4.1 ähnlich. Wir betrachten den Hilfsprozess aus (4.5)

$$\tilde{C}_{n,k}(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n [f_k(X_j, b) + 2\varphi_k(X_j)(b'X_j)(b'\Delta_{n,j}) + (b'X_j)^2\varphi_k(\Delta_{n,j})]$$

und den den Voraussetzungen angepassten Hilfsprozess

$$C_{n,k}^\star(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ f_k(X_j, b) - 2(b'X_j)b_k - b'(X_jX_j' - I_d)\mu_1^\star(b) - \varphi_k(X_j) - \frac{1}{2}e_k'(X_jX_j' - I_d)\mu_2^\star(b) \right\}.$$

Man beachte, dass der Hilfsprozess $C_{n,k}^\star(b)$ mit dem in (4.6) definierten Hilfsprozess $C_{n,k}^*(b)$ im folgenden Zusammenhang steht: Es gilt

$$C_{n,k}^\star(b) = C_{n,k}^*(b) - \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ b'(X_jX_j' - I_d)\mu_1^\star(b) + \frac{1}{2}e_k'(X_jX_j' - I_d)\mu_2^\star(b) \right\}. \quad (7.7)$$

Um im vorliegenden Fall den Kovarianzkern des im folgenden Theorem auftretenden Grenzprozesses $C_k^{\star\star}(\cdot)$ von $C_{n,k}^{\star\star}(\cdot)$ aufschreiben zu können, führen wir die folgenden Notationen ein. Zunächst setzen wir für die reellwertigen gemischten Momente

$$\begin{aligned}\Xi_{j_1, j_2, j_3}^{(k)}(b, c) &:= E((b'X)^{j_1}(c'X)^{j_2}(e'_k X)^{j_3}), \\ \Upsilon_{4,1}(b, c) &:= E(b'X_1 X_1' \mu_1^{\star\star}(b) c'X_1 X_1' \mu_1^{\star\star}(c)), \\ \Upsilon_{4,2}^{(k)}(b, c) &:= E(e'_k X_1 X_1' \mu_2^{\star\star}(b) c'X_1 X_1' \mu_1^{\star\star}(c)), \\ \Upsilon_{4,3}^{(k)}(b, c) &:= E(e'_k X_1 X_1' \mu_2^{\star\star}(b) e'_k X_1 X_1' \mu_2^{\star\star}(c)),\end{aligned}$$

und für die vektorwertigen Momente

$$\begin{aligned}\Upsilon_{3,1}(b) &:= E(b'X_1 X_1' \mu_1^{\star\star}(b) X_1), \\ \Upsilon_{3,2}^{(k)}(b) &:= E(e'_k X_1 X_1' \mu_2^{\star\star}(b) X_1), \\ \Upsilon_{5,1}(b, c) &:= E(b'X_1 X_1' \mu_1^{\star\star}(b) (c'X_1)^2 X_1), \\ \Upsilon_{5,2}^{(k)}(b, c) &:= E(e'_k X_1 X_1' \mu_2^{\star\star}(b) (c'X_1)^2 X_1).\end{aligned}$$

Theorem 7.2.1

Es existiert ein zentrierter Gaußprozess $C_k^{\star\star}(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitskugel mit stetigen Pfaden und dem Kovarianzkern

$$\begin{aligned}\rho_k^{\star\star}(b, c) &= \Xi_{2,2,2}^{(k)}(b, c) - 2b_k \Xi_{1,2,1}^{(k)}(b, c) - \Xi_{0,2,2}^{(k)}(b, c) - 2c_k \Xi_{2,1,1}^{(k)}(b, c) - \Xi_{2,0,2}^{(k)}(b, c) \\ &\quad - \frac{5}{4} \Xi_{0,2,1}^{(k)}(b, c) \Xi_{2,0,1}^{(k)}(b, c) - \frac{1}{2} e'_k \left(2\Upsilon_{5,1}(b, c) + 2\Upsilon_{5,1}(c, b) + \Upsilon_{5,2}^{(k)}(b, c) \right. \\ &\quad \left. + \Upsilon_{5,2}^{(k)}(c, b) \right) + \left(c_k c + \frac{1}{2} e_k \right)' \left(2\Upsilon_{3,1}(b) + \Upsilon_{3,2}^{(k)}(b) \right) \\ &\quad + \left(b_k b + \frac{1}{2} e_k \right)' \left(2\Upsilon_{3,1}(c) + \Upsilon_{3,2}^{(k)}(c) \right) + \frac{1}{4} \left(4\Upsilon_{4,1}(b, c) + \Upsilon_{4,3}^{(k)}(b, c) \right. \\ &\quad \left. + 2\Upsilon_{4,2}^{(k)}(c, b) + 2\Upsilon_{4,2}^{(k)}(b, c) \right) + 4b_k c_k (b'c) + 2(b_k^2 + c_k^2) + 1\end{aligned}$$

($b, c \in \mathcal{S}^{d-1}$). Betrachtet man $C_k^{\star\star}(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$, so gilt

$$\sqrt{n} \left(\frac{1}{\sqrt{n}} C_{n,k}^{\star\star}(\cdot) - \varphi_k \circ \mu_2^{\star\star}(\cdot) \right) \xrightarrow{\mathcal{D}} C_k^{\star\star}(\cdot).$$

Beweis:

Der Beweis des Theorems unterteilt sich wie im Beweis zu Theorem 4.1.1 in drei Teile. Es ist zu zeigen:

1. $\sqrt{n} \left(\frac{1}{\sqrt{n}} C_{n,k}^{\star\star} - \varphi_k \circ \mu_2^{\star\star} \right) \xrightarrow{\mathcal{D}} C_k^{\star\star}$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$,
2. $\|C_{n,k} - \tilde{C}_{n,k}\|_C \xrightarrow{\mathbb{P}} 0$ und
3. $\|\tilde{C}_{n,k} - C_{n,k}^{\star\star}\|_C \xrightarrow{\mathbb{P}} 0$

für $n \rightarrow \infty$. Dann folgt die Behauptung direkt aus Theorem 4.1 aus [6]. Man beachte, dass die zweite Aussage schon in Lemma 4.1.3 gezeigt wurde. Dabei beachte man, dass die Annahme der sphärischen Symmetrie nicht verwendet wurde.

Lemma 7.2.2

Für das in (7.7) definierte Zufallselement $C_{n,k}^{\star}$ gilt die schwache Konvergenz

$$\sqrt{n} \left(\frac{1}{\sqrt{n}} C_{n,k}^{\star} - \varphi_k \circ \mu_2^{\star} \right) \xrightarrow{\mathcal{D}} C_k^{\star}$$

im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$.

Beweis:

Es gilt $C_{n,k}^{\star}(b) = \frac{1}{\sqrt{n}} \sum_{j=1}^n Z_{j,k}(b)$ mit

$$\begin{aligned} Z_{j,k}(b) &:= f_k(X_j, b) - 2(b'X_j)b_k - b'(X_jX_j' - I_d)\mu_1^{\star}(b) - \varphi_k(X_j) \\ &\quad - \frac{1}{2}e_k'(X_jX_j' - I_d)\mu_2^{\star}(b), \end{aligned}$$

$j = 1, \dots, n$. Da die Zentrierung im vorliegenden Fall nicht gegeben ist, müssen die Zufallselemente $Z_{1,k}, \dots, Z_{n,k}$ noch geeignet zentriert werden. Unter Verwendung der in (7.1) gegebenen Identität und den getroffenen Voraussetzungen $E(X_1) = 0$ und $E(X_1X_1') = I_d$ gilt für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned} E(Z_{1,k}(b)) &= E \left(f_k(X_1, b) - 2(b'X_1)b_k - b'(X_1X_1' - I_d)\mu_1^{\star}(b) \right. \\ &\quad \left. - \varphi_k(X_1) - \frac{1}{2}e_k'(X_1X_1' - I_d)\mu_2^{\star}(b) \right) \\ &= E(f_k(X_1, b)) \\ &= \varphi_k(\mu_2^{\star}(b)). \end{aligned}$$

Definiere für $j = 1, \dots, n$ und $b \in \mathcal{S}^{d-1}$

$$\mathcal{Z}_{j,k}(b) := Z_{j,k}(b) - \varphi_k(\mu_2^{\star}(b)). \quad (7.8)$$

Wir betrachten die unabhängigen identisch verteilten Zufallsvariablen $\mathcal{Z}_{1,k}, \dots, \mathcal{Z}_{n,k}$ als Zufallselemente im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$. Die Zufallselemente sind zentriert, und es gilt $E\|\mathcal{Z}_{1,k}\|_C^2 < \infty$. Um den Kovarianzkern des Zufallselements C_k^{\star} zu bestimmen, reicht es den Kovarianzkern von $\mathcal{Z}_{1,k}$ zu berechnen. Dabei gilt

$$\text{Cov}(\mathcal{Z}_{1,k}(b), \mathcal{Z}_{1,k}(c)) = E(Z_{1,k}(b)Z_{1,k}(c)) - \varphi_k(\mu_2^{\star}(b))\varphi_k(\mu_2^{\star}(c)).$$

Weiter gilt für $b, c \in \mathcal{S}^{d-1}$

$$\begin{aligned}
E(Z_{1,k}(b)Z_{1,k}(c)) &= E \left\{ \left[f_k(X_1, b) - 2(b'X_1)b_k - b'(X_1X_1' - I_d)\mu_1^{\star}(b) - \varphi_k(X_1) \right. \right. \\
&\quad \left. \left. - \frac{1}{2}e'_k(X_1X_1' - I_d)\mu_2^{\star}(b) \right] \cdot \left[f_k(X_1, c) - 2(c'X_1)c_k \right. \right. \\
&\quad \left. \left. - c'(X_1X_1' - I_d)\mu_1^{\star}(c) - \varphi_k(X_1) - \frac{1}{2}e'_k(X_1X_1' - I_d)\mu_2^{\star}(c) \right] \right\} \\
&= E(f_k(X_1, b)f_k(X_1, c)) - 2b_kE((b'X_1)f_k(X_1, c)) \\
&\quad - E(b'(X_1X_1' - I_d)\mu_1^{\star}(b)f_k(X_1, c)) - E(\varphi_k(X_1)f_k(X_1, c)) \\
&\quad - \frac{1}{2}E(e'_k(X_1X_1' - I_d)\mu_2^{\star}(b)f_k(X_1, c)) - 2c_kE(f_k(X_1, b)(c'X_1)) \\
&\quad + 4b_kc_kE((b'X_1)(c'X_1)) + 2c_kE(b'(X_1X_1' - I_d)\mu_1^{\star}(b)(c'X_1)) \\
&\quad + 2c_kE(\varphi_k(X_1)(c'X_1)) + c_kE(e'_k(X_1X_1' - I_d)\mu_2^{\star}(b)(c'X_1)) \\
&\quad - E(f_k(X_1, b)c'(X_1X_1' - I_d)\mu_1^{\star}(c)) \\
&\quad + 2b_kE((b'X_1)c'(X_1X_1' - I_d)\mu_1^{\star}(c)) \\
&\quad + E(b'(X_1X_1' - I_d)\mu_1^{\star}(b)c'(X_1X_1' - I_d)\mu_1^{\star}(c)) \\
&\quad + E(\varphi_k(X_1)c'(X_1X_1' - I_d)\mu_1^{\star}(c)) - E(f_k(X_1, b)\varphi_k(X_1)) \\
&\quad + \frac{1}{2}E(e'_k(X_1X_1' - I_d)\mu_2^{\star}(b)c'(X_1X_1' - I_d)\mu_1^{\star}(c)) \\
&\quad + 2b_kE((b'X_1)\varphi_k(X_1)) + E(b'(X_1X_1' - I_d)\mu_1^{\star}(b)\varphi_k(X_1)) \\
&\quad + E(\varphi_k^2(X_1)) + \frac{1}{2}E(e'_k(X_1X_1' - I_d)\mu_2^{\star}(b)\varphi_k(X_1)) \\
&\quad - \frac{1}{2}E(f_k(X_1, b)e'_k(X_1X_1' - I_d)\mu_2^{\star}(c)) \\
&\quad + b_kE((b'X_1)e'_k(X_1X_1' - I_d)\mu_2^{\star}(c)) \\
&\quad + \frac{1}{2}E(\varphi_k(X_1)e'_k(X_1X_1' - I_d)\mu_2^{\star}(c)) \\
&\quad + \frac{1}{2}E(b'(X_1X_1' - I_d)\mu_1^{\star}(b)e'_k(X_1X_1' - I_d)\mu_2^{\star}(c)) \\
&\quad + \frac{1}{4}E(e'_k(X_1X_1' - I_d)\mu_2^{\star}(b)e'_k(X_1X_1' - I_d)\mu_2^{\star}(c)) \\
&= \Xi_{2,2,2}^{(k)}(b, c) - 2b_k\Xi_{1,2,1}^{(k)}(b, c) - \Xi_{0,2,2}^{(k)}(b, c) - 2c_k\Xi_{2,1,1}^{(k)}(b, c) - \Xi_{2,0,2}^{(k)}(b, c) \\
&\quad + 4b_kc_k(b'c) + 2(b_k^2 + c_k^2) + 1 - e'_k\Upsilon_{5,1}(b, c) + b'\mu_1^{\star}(b)\Xi_{0,2,1}^{(k)}(b, c) \\
&\quad - \frac{1}{2}e'_k\Upsilon_{5,2}^{(k)}(b, c) + \frac{1}{2}e'_k\mu_2^{\star}(b)\Xi_{0,2,1}^{(k)}(b, c) + 2c_kc'\Upsilon_{3,1}(b) + c_kc'\Upsilon_{3,2}^{(k)}(b) \\
&\quad - e'_k\Upsilon_{5,1}(c, b) + c'\mu_1^{\star}(c)\Xi_{2,0,1}^{(k)}(b, c) + \Upsilon_{4,1}(b, c) - b'\mu_1^{\star}(b)c'\mu_1^{\star}(c) \\
&\quad + e'_k\Upsilon_{3,1}(c) + \frac{1}{2}\Upsilon_{4,2}^{(k)}(b, c) - \frac{1}{2}e'_k\mu_2^{\star}(b)c'\mu_1^{\star}(c) + e'_k\Upsilon_{3,1}(b) + \frac{1}{2}e'_k\Upsilon_{3,2}^{(k)}(b) \\
&\quad - \frac{1}{2}e'_k\Upsilon_{5,2}^{(k)}(c, b) + \frac{1}{2}e'_k\mu_2^{\star}(c)\Xi_{2,0,1}^{(k)}(b, c) + 2b_kb'\Upsilon_{3,1}(c) + b_kb'\Upsilon_{3,2}^{(k)}(c) \\
&\quad + \frac{1}{2}\Upsilon_{4,2}^{(k)}(c, b) - \frac{1}{2}b'\mu_1^{\star}(b)e'_k\mu_2^{\star}(c) + \frac{1}{2}e'_k\Upsilon_{3,2}^{(k)}(c) + \frac{1}{4}\Upsilon_{4,3}^{(k)}(b, c) \\
&\quad - \frac{1}{4}e'_k\mu_2^{\star}(b)e'_k\mu_2^{\star}(c).
\end{aligned}$$

Unter Berücksichtigung der Identitäten $b'\mu_1^\star(b) = \Xi_{2,0,1}^{(k)}(b, c) = e'_k\mu_2^\star(b)$ und $c'\mu_1^\star(c) = \Xi_{0,2,1}^{(k)}(b, c) = e'_k\mu_2^\star(c)$ ergibt kurzes Zusammenfassen

$$\begin{aligned}
E(Z_{1,k}(b)Z_{1,k}(c)) &= \Xi_{2,2,2}^{(k)}(b, c) - 2b_k\Xi_{1,2,1}^{(k)}(b, c) - \Xi_{0,2,2}^{(k)}(b, c) - 2c_k\Xi_{2,1,1}^{(k)}(b, c) \\
&\quad - \Xi_{2,0,2}^{(k)}(b, c) + 2\Xi_{0,2,1}^{(k)}(b, c)\Xi_{2,0,1}^{(k)}(b, c) - e'_k(\Upsilon_{5,1}(b, c) + \Upsilon_{5,1}(c, b)) \\
&\quad - \frac{1}{2}e'_k(\Upsilon_{5,2}^{(k)}(b, c) + \Upsilon_{5,2}^{(k)}(c, b)) + (2c_kc + e_k)'\Upsilon_{3,1}(b) \\
&\quad + \left(c_kc + \frac{1}{2}e_k\right)'\Upsilon_{3,2}^{(k)}(b) + (2b_kb + e_k)'\Upsilon_{3,1}(c) \\
&\quad + \left(b_kb + \frac{1}{2}e_k\right)'\Upsilon_{3,2}^{(k)}(c) + \frac{1}{4}(4\Upsilon_{4,1}(b, c) + \Upsilon_{4,3}^{(k)}(b, c) + 2\Upsilon_{4,2}^{(k)}(c, b) \\
&\quad + 2\Upsilon_{4,2}^{(k)}(b, c)) - \frac{1}{4}(2b'\mu_1^\star(b)e'_k\mu_2^\star(c) + e'_k\mu_2^\star(b)e'_k\mu_2^\star(c) \\
&\quad + 4b'\mu_1^\star(b)c'\mu_1^\star(c) + 2e'_k\mu_2^\star(b)c'\mu_1^\star(c)) + 4b_kc_k(b'c) + 2(b_k^2 + c_k^2) + 1 \\
&= \Xi_{2,2,2}^{(k)}(b, c) - 2b_k\Xi_{1,2,1}^{(k)}(b, c) - \Xi_{0,2,2}^{(k)}(b, c) - 2c_k\Xi_{2,1,1}^{(k)}(b, c) - \Xi_{2,0,2}^{(k)}(b, c) \\
&\quad - \frac{1}{4}\Xi_{0,2,1}^{(k)}(b, c)\Xi_{2,0,1}^{(k)}(b, c) - \frac{1}{2}e'_k(2\Upsilon_{5,1}(b, c) + 2\Upsilon_{5,1}(c, b) + \Upsilon_{5,2}^{(k)}(b, c) \\
&\quad + \Upsilon_{5,2}^{(k)}(c, b)) + \left(c_kc + \frac{1}{2}e_k\right)'\left(2\Upsilon_{3,1}(b) + \Upsilon_{3,2}^{(k)}(b)\right) \\
&\quad + \left(b_kb + \frac{1}{2}e_k\right)'\left(2\Upsilon_{3,1}(c) + \Upsilon_{3,2}^{(k)}(c)\right) + \frac{1}{4}\left(4\Upsilon_{4,1}(b, c) + \Upsilon_{4,3}^{(k)}(b, c) \right. \\
&\quad \left. + 2\Upsilon_{4,2}^{(k)}(c, b) + 2\Upsilon_{4,2}^{(k)}(b, c)\right) + 4b_kc_k(b'c) + 2(b_k^2 + c_k^2) + 1.
\end{aligned}$$

Insgesamt folgt

$$\begin{aligned}
\text{Cov}(\mathcal{Z}_{1,k}(b), \mathcal{Z}_{1,k}(c)) &= \Xi_{2,2,2}^{(k)}(b, c) - 2b_k\Xi_{1,2,1}^{(k)}(b, c) - \Xi_{0,2,2}^{(k)}(b, c) - 2c_k\Xi_{2,1,1}^{(k)}(b, c) \\
&\quad - \Xi_{2,0,2}^{(k)}(b, c) - \frac{5}{4}\Xi_{0,2,1}^{(k)}(b, c)\Xi_{2,0,1}^{(k)}(b, c) - \frac{1}{2}e'_k(2\Upsilon_{5,1}(b, c) \\
&\quad + 2\Upsilon_{5,1}(c, b) + \Upsilon_{5,2}^{(k)}(b, c) + \Upsilon_{5,2}^{(k)}(c, b)) + \left(c_kc + \frac{1}{2}e_k\right)'\left(2\Upsilon_{3,1}(b) \right. \\
&\quad \left. + \Upsilon_{3,2}^{(k)}(b)\right) + \left(b_kb + \frac{1}{2}e_k\right)'\left(2\Upsilon_{3,1}(c) + \Upsilon_{3,2}^{(k)}(c)\right) + \frac{1}{4}\left(4\Upsilon_{4,1}(b, c) \right. \\
&\quad \left. + \Upsilon_{4,3}^{(k)}(b, c) + 2\Upsilon_{4,2}^{(k)}(c, b) + 2\Upsilon_{4,2}^{(k)}(b, c)\right) + 4b_kc_k(b'c) \\
&\quad + 2(b_k^2 + c_k^2) + 1.
\end{aligned}$$

Also gilt nach Theorem B.4.1 die Behauptung. \square

Lemma 7.2.3

Für die in (4.5) und (7.7) definierten Zufallselemente gilt

$$\|\tilde{C}_{n,k} - C_{n,k}^\star\|_C \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Unter Benutzung der Identität $\Delta_{n,j} = \left(S_n^{-\frac{1}{2}} - I_d\right) X_j - S_n^{-\frac{1}{2}} \bar{X}_n$ gilt für $b \in \mathcal{S}^{d-1}$

$$\begin{aligned} \tilde{C}_{n,k}(b) - C_{n,k}^\star(b) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[2\varphi_k(X_j)(b'X_j)(b'\Delta_{n,j}) + (b'X_j)^2\varphi_k(\Delta_{n,j}) + 2(b'X_j)b_k \right. \\ &\quad \left. + b'(X_jX_j' - I_d)\mu_1^\star(b) + \varphi_k(X_j) + \frac{1}{2}e_k'(X_jX_j' - I_d)\mu_2^\star(b) \right] \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left[2\varphi_k(X_j)(b'X_j) \left(b' \left(\left(S_n^{-\frac{1}{2}} - I_d \right) X_j - S_n^{-\frac{1}{2}} \bar{X}_n \right) \right) \right. \\ &\quad \left. + (b'X_j)^2\varphi_k \left(\left(S_n^{-\frac{1}{2}} - I_d \right) X_j - S_n^{-\frac{1}{2}} \bar{X}_n \right) \right. \\ &\quad \left. + 2(b'X_j)b_k + b'(X_jX_j' - I_d)\mu_1^\star(b) + \varphi_k(X_j) \right. \\ &\quad \left. + \frac{1}{2}e_k'(X_jX_j' - I_d)\mu_2^\star(b) \right] \\ &= 2b' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ \varphi_k(X_j)(b'X_j)X_j - \mu_1^\star(b) \right\} \\ &\quad - 2b'S_n^{-\frac{1}{2}}\bar{X}_n\sqrt{n}b_k + 2b' \left(S_n^{-\frac{1}{2}} - I_d \right) \sqrt{n}\mu_1^\star(b) \\ &\quad - 2b'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ \varphi_k(X_j)(b'X_j) - b_k \right\} \\ &\quad + e_k' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b'X_j)^2X_j - \mu_2^\star(b) \right\} \\ &\quad + 2b_kb' \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j + e_k' \left(S_n^{-\frac{1}{2}} - I_d \right) \sqrt{n}\mu_2^\star(b) - \sqrt{n}e_k'S_n^{-\frac{1}{2}}\bar{X}_n \\ &\quad - e_k'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b'X_j)^2 - 1 \right\} + \frac{1}{\sqrt{n}} \sum_{j=1}^n \varphi_k(X_j) \\ &\quad + b' \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_jX_j' - I_d)\mu_1^\star(b) + \frac{1}{2}e_k' \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_jX_j' - I_d)\mu_2^\star(b) \\ &= \mathcal{A}_{n,k}(b) - \mathcal{B}_{n,k}(b) + \mathcal{E}_{n,k}(b) - \mathcal{F}_{n,k}(b) + \mathcal{M}_n(b) \\ &\quad - \mathcal{N}_n(b) + \mathcal{O}_{n,k}(b) - \mathcal{P}_{n,k}(b) \end{aligned}$$

mit

$$\begin{aligned}\mathcal{A}_{n,k}(b) &:= 2b' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \{ \varphi_k(X_j) (b' X_j) X_j - \mu_1^*(b) \}, \\ \mathcal{B}_{n,k}(b) &:= 2b' S_n^{-\frac{1}{2}} \bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \{ \varphi_k(X_j) (b' X_j) - b_k \}, \\ \mathcal{E}_{n,k}(b) &:= e'_k \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \{ (b' X_j)^2 X_j - \mu_2^*(b) \}, \\ \mathcal{F}_{n,k}(b) &:= e'_k S_n^{-\frac{1}{2}} \bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \{ (b' X_j)^2 - 1 \},\end{aligned}$$

und

$$\begin{aligned}\mathcal{M}_n(b) &:= 2b' \left(\sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right) + \frac{1}{2\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d) \right) \mu_1^*(b), \\ \mathcal{N}_n(b) &:= 2b_k b' (S_n^{-\frac{1}{2}} - I_d) \sqrt{n} \bar{X}_n, \\ \mathcal{O}_{n,k}(b) &:= e'_k \left(\sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right) + \frac{1}{2\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d) \right) \mu_2^*(b), \\ \mathcal{P}_{n,k}(b) &:= e'_k (S_n^{-\frac{1}{2}} - I_d) \sqrt{n} \bar{X}_n.\end{aligned}$$

Mit der Dreiecksungleichung und dem eben Gezeigten gilt

$$\begin{aligned}\|\tilde{C}_{n,k} - C_{n,k}^*\|_C &\leq \|\mathcal{A}_{n,k}\|_C + \|\mathcal{B}_{n,k}\|_C + \|\mathcal{E}_{n,k}\|_C + \|\mathcal{F}_{n,k}\|_C + \|\mathcal{M}_n\|_C \\ &\quad + \|\mathcal{N}_n\|_C + \|\mathcal{O}_{n,k}\|_C + \|\mathcal{P}_{n,k}\|_C.\end{aligned}$$

Dieselben Methoden und Schlussfolgerungen wie im Beweis zu Lemma 4.1.4 führen zu der stochastischen Konvergenz gegen 0 der Summanden $\|\mathcal{A}_{n,k}\|_C$, $\|\mathcal{B}_{n,k}\|_C$, $\|\mathcal{E}_{n,k}\|_C$, $\|\mathcal{F}_{n,k}\|_C$, $\|\mathcal{N}_n\|_C$ und $\|\mathcal{P}_{n,k}\|_C$ auf der rechten Seite der obigen Ungleichung. Es bleibt $\|\mathcal{M}_n\|_C \xrightarrow{\mathbb{P}} 0$ und $\|\mathcal{O}_{n,k}\|_C \xrightarrow{\mathbb{P}} 0$ zu zeigen. Vereinfachend schreiben wir

$$\tilde{\mathcal{M}}_n := \sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right) + \frac{1}{2\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d).$$

Hiermit gilt

$$|\mathcal{M}_n(b)| = |2b' \tilde{\mathcal{M}}_n \mu_1^*(b)| \leq 2\|b\| \|\tilde{\mathcal{M}}_n\|_S \|\mu_1^*(b)\| = 2\|\tilde{\mathcal{M}}_n\|_S \|\mu_1^*(b)\|.$$

Da nach Voraussetzung $\mu_1^*(b)$ existiert und eine stetige Funktion ist, gilt $\sup_{b \in S^{d-1}} \|\mu_1^*(b)\| < \infty$ und damit reicht es, die stochastische Konvergenz von $\|\tilde{\mathcal{M}}_n\|_S$ zu untersuchen. Mit der Identität (vergleiche [17], S. 9)

$$\sqrt{n}(S_n - I_d) = \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d) - \sqrt{n} \bar{X}_n \bar{X}_n'$$

und $\sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right) = -\sqrt{n} (S_n - I_d) S_n^{-1} (S_n^{-\frac{1}{2}} + I_d)^{-1}$ gilt

$$\tilde{\mathcal{M}}_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d) \left[\frac{1}{2} I_d - S_n^{-1} (S_n^{-\frac{1}{2}} + I_d)^{-1} \right] + \sqrt{n} \bar{X}_n \bar{X}_n' S_n^{-1} (S_n^{-\frac{1}{2}} + I_d)^{-1},$$

also mit der Dreiecksungleichung

$$\begin{aligned} \|\tilde{\mathcal{M}}_n\|_S &\leq \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d) \right\|_S \left\| \frac{1}{2} I_d - S_n^{-1} (S_n^{-\frac{1}{2}} + I_d)^{-1} \right\|_S \\ &\quad + \|\sqrt{n} \bar{X}_n \bar{X}_n'\|_S \left\| S_n^{-1} (S_n^{-\frac{1}{2}} + I_d)^{-1} \right\|_S. \end{aligned}$$

Da sowohl $\frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j X_j' - I_d)$ als auch $\sqrt{n} \bar{X}_n$ straff ist und \bar{X}_n fast sicher gegen 0 bzw. S_n^{-1} und $S_n^{-\frac{1}{2}}$ fast sicher gegen I_d konvergieren, gilt $\|\tilde{\mathcal{M}}_n\|_S \xrightarrow{\mathbb{P}} 0$ und damit $\|\mathcal{M}_n\|_C \xrightarrow{\mathbb{P}} 0$. Weiter ist

$$|\mathcal{O}_{n,k}(b)| = |e_k' \tilde{\mathcal{M}}_n \mu_2^{\star}(b)| \leq \|\tilde{\mathcal{M}}_n\|_S \|\mu_2^{\star}(b)\|, \quad b \in \mathcal{S}^{d-1}.$$

Nach Voraussetzung existiert $\mu_2^{\star}(b)$, ist stetig in b und demnach gilt unter derselben Schlussweise wie eben $\|\mathcal{O}_{n,k}\|_C \xrightarrow{\mathbb{P}} 0$ und insgesamt die Behauptung. \square

7.3 Konvergenzaussagen für D_n

Dieser Abschnitt soll das Konvergenzverhalten des stochastischen Prozesses D_n unter den oben gemachten Voraussetzungen aufzeigen. Wir betrachten hierzu den Hilfsprozess

$$\tilde{D}_n(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n [(b' X_j)^3 + 3(b' X_j)^2 (b' \Delta_{n,j})]$$

aus (4.19) und den angepassten Hilfsprozess

$$D_n^{\star}(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b' X_j)^3 - 3(b' X_j) - \frac{3}{2} b' (X_j X_j' - I_d) \mu_2^{\star}(b) \right\}. \quad (7.9)$$

Um im folgenden Theorem den Kovarianzkern des Grenzprozesses $D^{\star}(\cdot)$ in übersichtlicher Form aufschreiben zu können, führen wir analog zum letzten Abschnitt abkürzende Bezeichnungen für die auftretenden gemischten Momente ein. Wir setzen für die reellwertigen gemischten Momente

$$\Xi_{j_1, j_2}(b, c) := E \left((b' X_1)^{j_1} (c' X_1)^{j_2} \right), \quad (7.10)$$

$$\Upsilon_{4,2}(b, c) := E \left(b' X_1 X_1' \mu_2^{\star}(b) c' X_1 X_1' \mu_2^{\star}(c) \right) \quad (7.11)$$

und für die vektorwertigen gemischten Momente

$$\Upsilon_{3,3}(b) := E \left(b' X_1 X_1' \mu_2^{\star}(b) X_1 \right), \quad (7.12)$$

$$\Upsilon_{5,4}(b, c) := E \left(b' X_1 X_1' \mu_2^{\star}(b) (c' X_1)^2 X_1 \right). \quad (7.13)$$

Theorem 7.3.1

Es existiert ein zentrierter Gaußprozess $D^\star(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitssphäre mit stetigen Pfaden und dem Kovarianzkern

$$\begin{aligned} \rho^\star(b, c) &= \Xi_{3,3}(b, c) - 3(\Xi_{1,3}(b, c) + \Xi_{3,1}(b, c)) - \frac{3}{2}(c' \Upsilon_{5,4}(b, c) + b' \Upsilon_{5,4}(c, b)) \\ &\quad + \frac{9}{2}(c' \Upsilon_{3,3}(b) + b' \Upsilon_{3,3}(c)) + \frac{9}{4} \Upsilon_{4,2}(b, c) - \frac{1}{4} \mu_3^\star(b) \mu_3^\star(c) + 9(b'c). \end{aligned}$$

($b, c \in \mathcal{S}^{d-1}$). Betrachtet man $D^\star(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$, so gilt

$$\sqrt{n} \left(\frac{1}{\sqrt{n}} D_n(\cdot) - \mu_3^\star(\cdot) \right) \xrightarrow{\mathcal{D}} D^\star(\cdot).$$

Beweis:

Der Beweis des Theorems unterteilt sich wie im Beweis zu Theorem 4.1.1 und Theorem 7.2.1 in drei Teile. Es ist zu zeigen:

1. $\sqrt{n} \left(\frac{1}{\sqrt{n}} D_n^\star - \mu_3^\star \right) \xrightarrow{\mathcal{D}} D^\star$ in $C(\mathcal{S}^{d-1}, \mathbb{R})$,
2. $\|D_n - \tilde{D}_n\|_C \xrightarrow{\mathbb{P}} 0$ und
3. $\|\tilde{D}_n - D_n^\star\|_C \xrightarrow{\mathbb{P}} 0$

für $n \rightarrow \infty$. Dann folgt die Behauptung direkt aus Theorem 4.1 aus [6]. Man beachte, dass die zweite Aussage schon in Lemma 4.2.3 gezeigt wurde und dort die Annahme der sphärischen Symmetrie nicht verwendet wurde.

Lemma 7.3.2

Für das in (7.9) definierte Zufallselement D_n^\star gilt die schwache Konvergenz

$$\sqrt{n} \left(\frac{1}{\sqrt{n}} D_n^\star - \mu_3^\star \right) \xrightarrow{\mathcal{D}} D^\star \quad (7.14)$$

im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$.

Beweis:

Analog zum Beweis des Lemmas 7.2.2 ist $D_n^\star(b) := \frac{1}{\sqrt{n}} \sum_{j=1}^n Z_j(b)$ mit

$$Z_j(b) := (b' X_j)^3 - 3(b' X_j) - \frac{3}{2} b' (X_j X_j' - I_d) \mu_2^\star(b), \quad j = 1, \dots, n.$$

Mit $EZ_1(b) = \mu_3^\star(b)$ für $b \in \mathcal{S}^{d-1}$ gilt

$$\begin{aligned}
E[Z_1(b)Z_1(c)] &= E \left[(b'X_1)^3(c'X_1)^3 - 3(b'X_1)(c'X_1)^3 - \frac{3}{2}b'(X_1X_1' - I_d)\mu_2^\star(b)(c'X_1)^3 \right. \\
&\quad - 3(b'X_1)^3(c'X_1) + 9(b'X_1)(c'X_1) + \frac{9}{2}b'(X_1X_1' - I_d)\mu_2^\star(b)(c'X_1) \\
&\quad - \frac{3}{2}c'(X_1X_1' - I_d)\mu_2^\star(c)(b'X_1)^3 + \frac{9}{2}c'(X_1X_1' - I_d)\mu_2^\star(c)(b'X_1) \\
&\quad \left. + \frac{9}{4}b'(X_1X_1' - I_d)\mu_2^\star(b)c'(X_1X_1' - I_d)\mu_2^\star(c) \right] \\
&= \Xi_{3,3}(b, c) - 3(\Xi_{1,3}(b, c) + \Xi_{3,1}(b, c)) - \frac{3}{2}(c'\Upsilon_{5,4}(b, c) + b'\Upsilon_{5,4}(c, b)) \\
&\quad + \frac{9}{2}(c'\Upsilon_{3,3}(b) + b'\Upsilon_{3,3}(c)) + \frac{9}{4}\Upsilon_{4,2}(b, c) + \frac{3}{4}\mu_3^\star(b)\mu_3^\star(c) + 9(b'c).
\end{aligned}$$

Definiere für $j = 1, \dots, n$ und $b \in \mathcal{S}^{d-1}$

$$\mathcal{Z}_j(b) := Z_j(b) - \mu_3^\star(b). \quad (7.15)$$

Wir betrachten die unabhängigen identisch verteilten Zufallsvariablen $\mathcal{Z}_1, \dots, \mathcal{Z}_n$ als Zufallselemente im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$. Insgesamt folgt

$$\text{Cov}(\mathcal{Z}_1(b), \mathcal{Z}_1(c)) = E[Z_1(b)Z_1(c)] - \mu_3^\star(b)\mu_3^\star(c).$$

Also gilt nach Theorem B.4.1 die Behauptung. \square

Lemma 7.3.3

Für die in (4.19) und (7.9) definierten Zufallselemente gilt

$$\|\tilde{D}_n - D_n^\star\|_C \xrightarrow{\mathbb{P}} 0.$$

Beweis:

Wie im Beweis zu Lemma 7.2.3 verwenden wir die Identität $\Delta_{n,j} = \left(S_n^{-\frac{1}{2}} - I_d\right) X_j -$

$S_n^{-\frac{1}{2}}\bar{X}_n$, und damit gilt

$$\begin{aligned}
\tilde{D}_n - D_n^\star &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ 3(b'X_j)^2(b'\Delta_{n,j}) + 3(b'X_j) + \frac{3}{2}b'(X_jX_j' - I_d)\mu_2^\star(b) \right\} \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ 3(b'X_j)^2 \left(b' \left((S_n^{-\frac{1}{2}} - I_d) X_j - S_n^{-\frac{1}{2}}\bar{X}_n \right) \right) \right. \\
&\quad \left. + 3(b'X_j) + \frac{3}{2}b'(X_jX_j' - I_d)\mu_2^\star(b) \right\} \\
&= 3b' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b'X_j)^2 X_j - \mu_2^\star(b) \right\} + \sqrt{n}3b' \left(S_n^{-\frac{1}{2}} - I_d \right) \mu_2^\star(b) \\
&\quad - 3b'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b'X_j)^2 - 1 \right\} - \sqrt{n}3b'S_n^{-\frac{1}{2}}\bar{X}_n \\
&\quad + \frac{3}{\sqrt{n}} \sum_{j=1}^n (b'X_j) + 3b' \frac{1}{2\sqrt{n}} \sum_{j=1}^n (X_jX_j' - I_d)\mu_2^\star(b) \\
&= \mathcal{A}_n(b) - \mathcal{B}_n(b) + \mathcal{E}_n(b) - \mathcal{F}_n(b)
\end{aligned}$$

mit

$$\begin{aligned}
\mathcal{A}_n(b) &:= 3b' \left(S_n^{-\frac{1}{2}} - I_d \right) \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b'X_j)^2 X_j - \mu_2^\star(b) \right\}, \\
\mathcal{B}_n(b) &:= 3b'S_n^{-\frac{1}{2}}\bar{X}_n \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\{ (b'X_j)^2 - 1 \right\}, \\
\mathcal{E}_n(b) &:= 3b' \left(\sqrt{n} \left(S_n^{-\frac{1}{2}} - I_d \right) + \frac{1}{2\sqrt{n}} \sum_{j=1}^n (X_jX_j' - I_d) \right) \mu_2^\star(b), \\
\mathcal{F}_n(b) &:= 3b' \left(S_n^{-\frac{1}{2}} - I_d \right) \sqrt{n}\bar{X}_n.
\end{aligned}$$

Weiter gilt

$$\|\tilde{D}_n - D_n^\star\|_C \leq \|\mathcal{A}_n\|_C + \|\mathcal{B}_n\|_C + \|\mathcal{E}_n\|_C + \|\mathcal{F}_n\|_C.$$

Die gleichen Argumente wie in Lemma 4.1.4 und Lemma 7.2.3 führen zur stochastischen Konvergenz gegen 0 der vier Summanden auf der rechten Seite der Ungleichung, und damit ist die Behauptung gezeigt. \square

7.4 Konvergenzaussagen für W_n

In diesem Abschnitt soll die asymptotische Verteilung des multivariaten Multiparameterprozesses W_n hergeleitet werden. Dabei werden die eben gezeigten Ergebnisse verwendet. Um die Kovarianzmatrixstruktur vollständig beschreiben zu können, müssen die paarweisen Abhängigkeiten zwischen den Prozessen $C_{n,k}$ für $k = 1, \dots, d$ und D_n berechnet werden. Es

bezeichne für $b, c \in \mathcal{S}^{d-1}$ und $k, l = 1, \dots, d$ mit der Schreibweise aus (7.8) und (7.15)

$$\begin{aligned}\Phi_{k,l}^{\star}(b, c) &:= \text{Cov}(\mathcal{L}_{1,k}(b), \mathcal{L}_{1,l}(c)), \\ \Phi_k^{\star}(b, c) &:= \text{Cov}(\mathcal{L}_{1,k}(b), \mathcal{L}_1(c)).\end{aligned}$$

Um diese Kovarianzen darstellen zu können, müssen weitere Bezeichnungen für gemischte Momente eingeführt werden. Es seien für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$ die reellwertigen gemischten Momente

$$\begin{aligned}\Xi_{j_1, j_2, j_3, j_4}^{(k,l)}(b, c) &:= E((b'X_1)^{j_1}(c'X_1)^{j_2}(e'_k X_1)^{j_3}(e'_l X_1)^{j_4}), \\ \Upsilon_{4,3}(b, c) &:= E(b'X_1 X_1' \mu_1^{\star}(b) c'X_1 X_1' \mu_2^{\star}(c)), \\ \Upsilon_{4,4}^{(k)}(b, c) &:= E(e'_k X_1 X_1' \mu_2^{\star}(b) c'X_1 X_1' \mu_2^{\star}(c)), \\ \Upsilon_{4,3}^{(k,l)}(b, c) &:= E(e'_k X_1 X_1' \mu_2^{\star}(b) e'_l X_1 X_1' \mu_2^{\star}(c)).\end{aligned}$$

Dann gilt

$$\begin{aligned}\Phi_{k,l}^{\star}(b, c) &:= \Xi_{2,2,1,1}^{(k,l)}(b, c) - 2b_k \Xi_{1,2,1}^{(l)}(b, c) - e'_l \Upsilon_{5,1}(b, c) + b' \mu_1^{\star}(b) \varphi_l \circ \mu_2^{\star}(c) \\ &\quad - \Xi_{0,2,1,1}^{(k,l)}(b, c) - \frac{1}{2} e'_l \Upsilon_{5,2}^{(k)}(b, c) - \frac{1}{4} e'_k \mu_2^{\star}(b) e'_l \mu_2^{\star}(c) \\ &\quad - 2c_l \Xi_{2,1,1}^{(k)}(b, c) + 4b_k c_l (b'c) + 2c_l c' \Upsilon_{3,1}(b) + 2c_l c_k + c_l c' \Upsilon_{3,2}^{(k)}(b) \\ &\quad - e'_k \Upsilon_{5,1}(c, b) + c' \mu_1^{\star}(c) \varphi_k \circ \mu_2^{\star}(b) + 2b_k b' \Upsilon_{3,1}(c) + \Upsilon_{4,1}(b) - b' \mu_1^{\star}(b) c' \mu_1^{\star}(c) \\ &\quad + e'_k \Upsilon_{3,1}(c) + \frac{1}{2} \Upsilon_{4,2}^{(k)}(b, c) - \frac{1}{2} e'_k \mu_2^{\star}(b) c' \mu_1^{\star}(c) - \Xi_{2,0,1,1}^{(k,l)}(b, c) + 2b_k b_l \\ &\quad + e'_l \Upsilon_{3,1}(b) + \frac{1}{2} e'_l \Upsilon_{3,2}^{(k)}(b) - \frac{1}{2} e'_k \Upsilon_{5,2}^{(l)}(c, b) \\ &\quad + b_k b' \Upsilon_{3,2}^{(l)}(c) + \frac{1}{2} \Upsilon_{4,2}^{(l)}(c, b) - \frac{1}{2} b' \mu_1^{\star}(b) e'_l \mu_2^{\star}(c) \\ &\quad + \frac{1}{2} e'_k \Upsilon_{3,2}^{(l)}(c) + \frac{1}{4} \Upsilon_{4,3}^{(k,l)}(b, c),\end{aligned}$$

$$\begin{aligned}\Phi_k^{\star}(b, c) &:= \Xi_{2,3,1}^{(k)}(b, c) - 2b_k \Xi_{1,3}^{(k)}(b, c) - \Xi_{0,3,1}^{(k)}(b, c) - c' \Upsilon_{5,1}(b, c) + b' \mu_1^{\star}(b) \mu_3^{\star}(c) \\ &\quad - \frac{1}{2} c' \Upsilon_{5,2}^{(k)}(b, c) - \frac{1}{2} \varphi_k \circ \mu_2^{\star}(b) \mu_3^{\star}(c) - 3\Xi_{2,1,1}^{(k)}(b, c) + 6b_k (b'c) + 3c_k \\ &\quad + 3c' \Upsilon_{3,1}(b) + \frac{3}{2} c' \Upsilon_{3,2}^{(k)}(b) - \frac{3}{2} e'_k \Upsilon_{5,4}(c, b) + \frac{3}{2} \varphi_k \circ \mu_2^{\star}(b) c' \mu_2^{\star}(c) \\ &\quad + 3b_k b' \Upsilon_{3,3}(c) + \frac{3}{2} \Upsilon_{4,3}(b, c) - \frac{3}{2} b' \mu_1^{\star}(b) c' \mu_2^{\star}(c) \\ &\quad + \frac{3}{2} e'_k \Upsilon_{3,3}(c) + \frac{3}{4} \Upsilon_{4,4}^{(k)}(b, c) - \frac{3}{4} \varphi_k \circ \mu_2^{\star}(b) c' \mu_2^{\star}(c).\end{aligned}$$

Die Rechnungen hierzu sind zu Gunsten einer besseren Lesbarkeit in Anhang A.2 zu finden. Wie in (7.2) schon angedeutet, bezeichnen wir den Erwartungswert des Prozesses W_n mit

$$\mu(b) = E(W_1(b)) = \begin{pmatrix} \varphi_1 \circ \mu_2^{\star}(b) \\ \vdots \\ \varphi_d \circ \mu_2^{\star}(b) \\ \mu_3^{\star}(b) \end{pmatrix} = \begin{pmatrix} \mu_2^{\star}(b) \\ \mu_3^{\star}(b) \end{pmatrix}, \quad b \in \mathcal{S}^{d-1}. \quad (7.16)$$

Theorem 7.4.1

Es existiert ein zentrierter Gaußprozess $W^\star(b)$, $b \in \mathcal{S}^{d-1}$, auf der Einheitskugel mit stetigen Pfaden und dem Kovarianzmatrixkern

$$\sigma^\star(b, c) = \begin{pmatrix} \rho_1^\star(b, c) & \Phi_{1,2}^\star(b, c) & \cdots & \Phi_{1,d}^\star(b, c) & \Phi_1^\star(b, c) \\ \Phi_{2,1}^\star(b, c) & \rho_2^\star(b, c) & \cdots & \Phi_{2,d}^\star(b, c) & \Phi_2^\star(b, c) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Phi_{d,1}^\star(b, c) & \Phi_{d,2}^\star(b, c) & \cdots & \rho_d^\star(b, c) & \Phi_d^\star(b, c) \\ \Phi_1^\star(c, b) & \Phi_2^\star(c, b) & \cdots & \Phi_d^\star(c, b) & \rho^\star(b, c) \end{pmatrix}$$

für $b, c \in \mathcal{S}^{d-1}$. Betrachtet man $W^\star(\cdot)$ als Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$, so gilt

$$\sqrt{n} \left(\frac{1}{\sqrt{n}} W_n(\cdot) - \mu(\cdot) \right) \xrightarrow{\mathcal{D}} W^\star(\cdot).$$

Beweis:

Der Beweis verläuft analog zum Beweis von Theorem 4.3.1. Der Kovarianzmatrixkern des Prozesses W^\star ergibt sich aus den einzelnen Kovarianzen der Komponentenprozesse und deren Kombinationen. \square

Kapitel 8

Simulationen

In diesem Kapitel sollen die theoretischen Ergebnisse der letzten Kapitel durch eine Simulationsstudie anhand von verschiedenen Beispielverteilungen ergänzt werden. Alle verwendeten Teststatistiken wurden in der Statistiksoftware R implementiert, wobei einige in der Programmbibliothek MASS zur Verfügung gestellte Routinen verwendet wurden.

Wie schon in der Einführung kurz angedeutet, ist die Implementierung der Teststatistik T_n aus (2.5) durch die erforderliche Berechnung des Maximums auf der d -dimensionalen Einheitssphäre nicht exakt umsetzbar. Um dennoch Aussagen treffen zu können, werden wir eine Methode zur Approximation des Maximums vorstellen. Im Anschluss bestimmen wir mit der vorgestellten Methode empirische Quantile der Teststatistik T_n mit Hilfe einer Monte-Carlo Simulation.

8.1 Implementierung des Funktionals $\max_{b \in \mathcal{S}^{d-1}}(\cdot)$

Da das Funktional $\max_{b \in \mathcal{S}^{d-1}}(\cdot)$ im vorliegenden Fall nicht analytisch berechenbar ist, muss eine möglichst gute Approximation gefunden werden. Da die Einheitssphäre kompakt ist und wir uns auf dem Raum der stetigen Funktionen befinden, wird das Maximum angenommen. Es liegt also nahe, endlich viele Punkte auf der Einheitssphäre zu betrachten, die Funktion an diesen Datenpunkten auszuwerten und anschließend das erhaltene Maximum als Annäherung an das theoretische Maximum der Funktion zu verwenden.

Konkret seien $m_d \in \mathbb{N}$ und $u_1, \dots, u_{m_d} \in \mathcal{S}^{d-1}$ paarweise verschiedene Punkte auf der Einheitssphäre. Weiterhin sei $\Psi \in C(\mathcal{S}^{d-1}, \mathbb{R})$ eine stetige Funktion. Wie oben erläutert, wählen wir für das betrachtete Funktional die Approximation

$$\max_{b \in \mathcal{S}^{d-1}} \Psi(b) \approx \max\{\Psi(u_1), \dots, \Psi(u_{m_d})\}.$$

Es bleibt zu klären, wie man $m_d \in \mathbb{N}$ und $u_1, \dots, u_{m_d} \in \mathcal{S}^{d-1}$ am Besten wählt, um eine möglichst gute Annäherung an den wahren Wert des Funktionals zu erhalten. Es ist klar, dass wenn man nur „genügend“ viele Punkte gleichmäßig auf der Oberfläche verteilt, die Approximation des Maximums beliebig genau wird. Im Hinblick auf die durchzuführende

Monte-Carlo Simulation mit mehreren tausend Durchläufen muss allerdings ein möglichst effizientes Verfahren zur Bestimmung der Punkte u_1, \dots, u_{m_d} verwendet werden.

8.1.1 Gleichverteilte Datenpunkte auf der Einheitssphäre \mathcal{S}^{d-1}

Die Erzeugung von Realisierungen gleichverteilter Zufallsvektoren auf der Einheitssphäre ist effizient durchführbar, da für einen normalverteilten Zufallsvektor $N \sim \mathcal{N}_d(0, I_d)$

$$\frac{N}{\|N\|} \sim \mathcal{U}(\mathcal{S}^{d-1})$$

gilt. Allgemeiner stimmt diese Aussage auch für radial-symmetrische Verteilungen (siehe hierzu [9], S. 225f.). Wir beschränken uns in dieser Simulationsstudie auf die Normalverteilung, da sie in \mathbb{R} einfach erzeugt werden kann.

Im Folgenden wollen wir über die Güte der mit dem vorgestellten Verfahren erhaltenen Approximation des Maximums eine Aussage machen. Wie beim sogenannten Fluch der hohen Dimensionen bekannt ist, werden bei einer hohen Dimension d vergleichsweise viele Punkte benötigt, um eine vertretbare Genauigkeit der Abdeckung der Einheitssphäre zu erreichen. Da die Abschätzung zur Güte nicht trivial ist, möchten wir auf ein probabilistisches Resultat von Janson in [19] kurz eingehen, das über die $(d-1)$ -dimensionale sphärische Kappe mit maximalem Radius, welche vollständig in \mathcal{S}^{d-1} eingebettet werden kann und keinen von n unabhängigen und gleichverteilten Punkten auf der Einheitssphäre \mathcal{S}^{d-1} beinhaltet, eine asymptotische Verteilungsaussage macht. In diesem Sinne seien $U_1, \dots, U_n \sim \mathcal{U}(\mathcal{S}^{d-1})$ unabhängige, identisch verteilte Zufallsvektoren. Der klassische Fall für $d=2$ stellt die Frage nach der Verteilung des längsten Kreissegments, welches keinen der Punkte U_1, \dots, U_n enthält. Bezeichnen wir die Länge dieses Kreissegments mit L_n , so gilt nach Lévy, siehe [22],

$$nL_n - \log(n) \xrightarrow{\mathcal{D}} Z,$$

wobei die Zufallsgröße Z einer Gumbel-Verteilung folgt, es gilt also

$$P(Z \leq x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}.$$

Devroye zeigt in [8] für diesen Fall die Schranken

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{nL_n - \log(n)}{\log(\log(n))} &= 0, \\ \limsup_{n \rightarrow \infty} \frac{nL_n - \log(n)}{\log(\log(n))} &= 2 \end{aligned}$$

fast sicher. Für $d \geq 3$ gilt ein ähnliches Resultat, welches wir im Folgenden kurz vorstellen. Zunächst führen wir einige Bezeichnungen ein. Es sei $K_r(c)$ eine $(d-1)$ -dimensionale sphärische Kappe um den Vektor $c \in \mathcal{S}^{d-1}$ vom Radius $r > 0$ in \mathcal{S}^{d-1} , also

$$K_r(c) := \{b \in \mathcal{S}^{d-1} : d_{\mathcal{S}}(c, b) \leq r\},$$

wobei die Metrik $d_{\mathcal{S}}$ in (B.1) gegeben ist. Mit $\omega^{d-1}(\cdot)$ bezeichnen wir das $(d-1)$ -dimensionale Hausdorff-Maß. Es gilt $\omega^{d-1}(\mathcal{S}^{d-1}) = |\mathcal{S}^{d-1}| = 2\pi^{\frac{d}{2}} / \Gamma(\frac{d}{2})$. Weiter sei

$$l_n := \sup \{r > 0 : \exists x \in \mathcal{S}^{d-1} \text{ mit } K_r(x) \subset \mathcal{S}^{d-1} \setminus \{U_j\}_{j \in \{1, \dots, n\}}\}.$$

der maximale Radius einer sphärischen Kappe, die keinen der Punkte U_1, \dots, U_n enthält. Da wir uns mit der Oberfläche der maximalen sphärischen Kappe beschäftigen und nicht mit deren Lage, halten wir im folgenden ein $x_0 \in \mathcal{S}^{d-1}$ fest, und schreiben

$$\mathcal{U}_n := \omega^{d-1}(K_{l_n}(x_0)).$$

Theorem 8.1.1

Mit der obigen Notation gilt

$$n\mathcal{U}_n - \log(n) - (d-2)\log(\log(n)) - \log \beta_d \xrightarrow{\mathcal{D}} Z,$$

wobei die Zufallsgröße Z einer Gumbel-Verteilung genügt und die Konstante β_d durch

$$\beta_d := \frac{1}{(d-1)!} \left(\frac{\sqrt{\pi} \Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2})} \right)^{d-2}$$

gegeben ist. Weiter sind

$$\liminf_{n \rightarrow \infty} \frac{n\mathcal{U}_n - \log(n)}{\log(\log(n))} = d-2,$$

$$\limsup_{n \rightarrow \infty} \frac{n\mathcal{U}_n - \log(n)}{\log(\log(n))} = d$$

fast sicher.

Beweis:

Siehe [19], S. 275 f. □

Mit Hilfe von Theorem 8.1.1 lässt sich das Verhalten von \mathcal{U}_n abschätzen, so dass \mathcal{U}_n ungefähr mit der Größenordnung $(\log(n) + (d-2)\log(\log(n)))/n$ fällt. Dies bedeutet, dass das größte $d-1$ -dimensionale Volumen sich durch folgende Tabelle abschätzen lässt. Diese

n	$d=2$	$d=3$	$d=5$	$d=10$
1000	0.01077	0.01271	0.01657	0.02623
7000	0.00189	0.00220	0.00282	0.00438

Tabelle 8.1: Größenordnung von \mathcal{U}_n für verschiedene n und verschiedene Dimensionen d

Überlegungen haben in den folgenden Simulationen zu der entsprechenden Wahl bei der Anzahl der verwendeten Punkte zur Diskretisierung des Maximums geführt.

Bemerkung 8.1.2

In einer weiteren Simulationsstudie wurde ein alternatives Verfahren zur Approximation des Maximums verwendet: Es wird ein gleichmäßiges Netz an Datenpunkten auf der

Einheitssphäre erzeugt, für Details siehe [12], S.166 f., womit das Maximum durch die so gewonnene Diskretisierung angenähert wird. Bei einem direkten Vergleich der Studien hat sich gezeigt, dass die Ergebnisse beider Methoden vergleichbar sind und ähnliche Werte produzieren. In der Tendenz sind die kritischen Werte der Teststatistik T_n im folgenden Abschnitt mit der Zufallsmethode größer, so dass wir diese im Folgenden verwenden.

8.2 Empirische Quantile unter der Nullhypothese

Wie wir in Kapitel 4 gezeigt haben, konvergiert die Teststatistik unter der Nullhypothese in Verteilung gegen ein Funktional eines Gaußprozesses auf der Oberfläche der Einheitssphäre mit bestimmter Kovarianzstruktur. Da uns keine Möglichkeit bekannt ist, für die sich ergebende Grenzverteilung Quantile zu bestimmen, welche bei einer Testdurchführung als Ablehngrenzen verwendet werden könnten, werden empirische Quantile mit Hilfe einer Monte-Carlo-Simulation berechnet. Dabei wurden für die Dimensionen $d = 2$ und $d = 3$ jeweils 100000 Durchläufe berechnet bei einer Abdeckung der Oberfläche der Einheitskugel von 1000 Datenpunkten. Für die Dimensionen $d = 5$ und $d = 10$ wurde die Anzahl der Durchläufe auf Grund der hohen Rechenzeiten auf 25000 beschränkt. Um bei hohen Dimensionen eine einigermaßen gute Approximation gewährleisten zu können, wurden 7000 Datenpunkte auf \mathcal{S}^{d-1} zur Bestimmung des Maximums berechnet. Die nachfolgenden Tabellen 8.2 - 8.3 geben in Abhängigkeit des Stichprobenumfangs für verschiedene Ablehnniveaus die entsprechenden simulierten Quantile an.

n	$d = 2$	$d = 3$	$d = 5$	$d = 10$
25	0.3150	0.4988	0.7536	0.9540
50	0.1561	0.2577	0.3842	0.6176
100	0.0697	0.1314	0.2221	0.3452
200	0.0382	0.0659	0.1140	0.1835

Tabelle 8.2: Empirische 95%-Quantile unter der Nullhypothese für die Teststatistik von Cox und Small

n	$d = 2$	$d = 3$	$d = 5$	$d = 10$
25	0.4171	0.5883	0.7939	0.9833
50	0.2158	0.3174	0.4747	0.6540
100	0.1031	0.1661	0.2588	0.3741
200	0.0838	0.0545	0.1328	0.2017

Tabelle 8.3: Empirische 99%-Quantile unter der Nullhypothese für die Teststatistik von Cox und Small

Dabei fällt auf, dass sich mit wachsendem Stichprobenumfang die Teststatistik mehr um 0 konzentriert, was auf Grund der stochastischen Konvergenz der Teststatistik unter der Hypothese gegen 0 nicht verwunderlich ist.

8.3 Klassische Verfahren zum Testen auf multivariate Normalverteilung

Ein Vergleich verschiedener Testverfahren für das gleiche Testproblem benötigt weitere Teststatistiken. Wir verwenden in den folgenden Simulationen die in der Literatur etablierten Tests. Auf Grund des breiten Spektrums an verschiedenen Teststatistiken für das vorgelegte Anpassungstestproblem konzentrieren wir uns auf affin invariante Teststatistiken. Die in den folgenden Tabellen 8.4 - 8.7 angegebenen empirischen Quantile entstammen einer Monte-Carlo Simulation mit 100000 Durchläufen.

1. *Mardias-Schiefestatistik*: Mardia führte in [24, 25] eine multivariate Verallgemeinerung der univariaten Schiefe ein. In diesem Sinne bezeichne

$$b_{n,d}^{(1)} = \frac{1}{n^2} \sum_{j,k=1}^n (Y_{n,j}' Y_{n,k})^3$$

mit $Y_{n,j}$ wie in (2.7) für $j = 1, \dots, n$ das von Mardia vorgestellte nichtnegative invariante Schiefe-Maß. Unter der Hypothese H_0 gilt

$$nb_{n,d}^{(1)} \xrightarrow{\mathcal{D}} 6\mathcal{L},$$

wobei \mathcal{L} eine $\chi_{\frac{1}{6}d(d+1)(d+2)}^2$ -verteilte Zufallsvariable ist, siehe [24].

n	$d = 2$	$d = 3$	$d = 5$
25	1.753	3.457	9.358
50	0.781	1.997	5.440
100	0.544	1.061	2.869

Tabelle 8.4: Empirische 95%-Quantile unter der Nullhypothese für die Schiefestatistik von Mardia für die Stichprobenumfänge $n = 25, 50, 100$

2. *Malkovich und Afifi-Schiefestatistik*: Malkovich und Afifi führten als Maß für die multivariate Schiefe eines Zufallsvektors X mit Erwartungswertvektor μ und nichtsingulärer Kovarianzmatrix Σ in [23] die Größe

$$\beta_M^{(1)} = \max_{u \in \mathcal{S}^{d-1}} \frac{[E(u'(X - \mu))^3]^2}{(u'\Sigma u)^3}$$

ein. Die auf dieser Größe aufbauende Teststatistik lautet

$$b_{n,M}^{(1)} = \max_{u \in \mathcal{S}^{d-1}} b_n^{(1)}(u),$$

mit

$$b_n^{(1)}(u) = \frac{\left[\frac{1}{n} \sum_{j=1}^n (u'(X_j - \bar{X}_n))^3 \right]^2}{(u'S_n u)^3}, \quad u \in \mathcal{S}^{d-1}.$$

Konvergenzaussagen für $b_{n,M}^{(1)}$ sind in [4] zu finden. In dem Artikel wurde gezeigt, dass unter sphärisch-symmetrischer Verteilungsannahme und existierendem sechsten Moment der Radienverteilung des Zufallsvektors X

$$nb_{n,M}^{(1)} \xrightarrow{\mathcal{D}} \max_{u \in \mathcal{S}^{d-1}} \mathcal{L}^2(u),$$

wobei $\mathcal{L}(\cdot)$ ein zentrierter Gaußprozess auf \mathcal{S}^{d-1} mit stetigen Pfaden und Kovarianzkern

$$\rho^*(u, v) = \frac{6m_6}{d(d+2)(d+4)} (u'v)^3 + 9 \left(1 - \frac{2m_4}{d(d+2)} + \frac{m_6}{d(d+2)(d+4)} \right) (u'v)$$

für $u, v \in \mathcal{S}^{d-1}$. Dabei handelt es sich um denselben Grenzprozess wie in Theorem 4.2.1. Wir möchten auf die Aussagen unter benachbarten Alternativen für die Schiefestatistik von Malkovich und Afifi aus Bemerkung 6.1.4 der vorliegenden Arbeit hinweisen. Da in diesem Fall keine direkte Berechnung von kritischen Werten zum Testentscheid des Anpassungstests möglich ist, geben wir in der folgenden Tabelle empirische kritische Werte zu einem Ablehnniveau von $\alpha = 0.05$ an.

n	$d = 2$	$d = 3$	$d = 5$
25	1.348	2.093	3.590
50	1.018	1.194	1.858
100	0.414	0.613	1.000

Tabelle 8.5: Empirische 95%-Quantile unter der Nullhypothese für die Schiefestatistik von Malkovich und Afifi für die Stichprobenumfänge $n = 25, 50, 100$

3. *BHEP-Test*: Die sogenannten BHEP (Baringhaus-Henze-Epps-Pulley)-Tests sind eine Klasse invarianter und konsistenter Anpassungstests auf multivariate Normalverteilung, welche auf der empirischen charakteristischen Funktion basieren. Die allgemeinste Version wurde von Henze und Zirkler in [18] vorgestellt. Die BHEP-Statistik lautet in Abhängigkeit des Parameters $\beta > 0$

$$T_{n,\beta} = \int_{\mathbb{R}^d} \left| \Psi_n(t) - \exp\left(-\frac{\|t\|^2}{2}\right) \right|^2 \varphi_\beta(t) dt,$$

wobei

$$\Psi_n(t) = \frac{1}{n} \sum_{k=1}^n \exp(it'Y_{n,k})$$

die empirische charakteristische Funktion ($i^2 = -1$) und $Y_{n,k}$ für $k = 1, \dots, n$ die skalierten Residuen aus (2.7) sind. Weiter ist

$$\varphi_\beta(t) = (2\pi\beta^2)^{-\frac{d}{2}} \exp\left(-\frac{\|t\|^2}{2\beta^2}\right)$$

die Dichte der multivariaten Normalverteilung $\mathcal{N}_d(0, \beta^2 I_d)$. Anstatt der oben angegebenen Integraldarstellung lässt sich $T_{n,\beta}$ auch als

$$T_{n,\beta} = \frac{1}{n^2} \sum_{j,k=1}^n \exp\left(-\frac{\beta^2}{2} \|Y_{n,j} - Y_{n,k}\|^2\right) - 2(1 + \beta^2)^{-\frac{d}{2}} \frac{1}{n} \sum_{j=1}^n \exp\left(-\frac{\beta^2 \|Y_{n,j}\|^2}{2(1 + \beta^2)}\right) + (1 + 2\beta^2)^{-\frac{d}{2}}$$

schreiben, wobei diese Darstellung für Simulationen deutlich leichter zu implementieren ist. Die Hypothese wird für große Werte von $T_{n,\beta}$ verworfen. In der folgenden Tabelle sind für verschiedene Parameter $\beta > 0$ die empirischen 95%-Quantile der Teststatistik für einen Stichprobenumfang von $n = 50$ angegeben (siehe hierzu auch [18], S. 3606).

β	$d = 2$	$d = 3$	$d = 5$
0.5	0.078	0.119	0.211
1	0.548	0.678	0.849
3	1.184	1.096	1.021

Tabelle 8.6: Empirische 95%-Quantile unter der Nullhypothese für die BHEP-Teststatistik für $n = 50$ in Abhängigkeit des Parameters $\beta > 0$

Die Klasse der BHEP-Statistiken und ihre asymptotischen Eigenschaften wurde ausführlich untersucht in [18, 15, 17], für eine kurze Zusammenfassung siehe auch [16].

4. *Test von Székely und Rizzo:* Székely und Rizzo stellen in [28] einen Test auf multivariate Normalverteilung unter Verwendung des euklidischen Abstands vor. Sie verworfen die Hypothese H_0 für große Werte der Teststatistik

$$T_n^{(SR)} = n \left(\frac{2}{n} \sum_{j=1}^n E_{Z_1} \|Y_{n,j} - Z_1\| - E \|Z_1 - Z_2\| - \frac{1}{n^2} \sum_{j,k=1}^n \|Y_{n,j} - Y_{n,k}\| \right),$$

wobei $Y_{n,j}$ wie in (2.7) für $j = 1, \dots, n$ und Z_1, Z_2 unabhängige $\mathcal{N}_d(0, I_d)$ -verteilte Zufallsvektoren sind. Weiter sind Z_1, Z_2 unabhängig von den Zufallsvektoren X_1, \dots, X_n . Konvergenzaussagen für die Teststatistik $T_n^{(SR)}$ liegen noch nicht vor.

n	$d = 2$	$d = 3$	$d = 5$
25	0.943	1.046	1.240
50	0.959	1.065	1.264
100	0.967	1.076	1.274

Tabelle 8.7: Empirische 95%-Quantile unter der Nullhypothese für die Teststatistik von Székely und Rizzo für die Stichprobenumfänge $n = 25, 50, 100$

Ein guter Überblick über die verwendeten und auch weitere Teststatistiken für den Anpassungstest auf multivariate Normalverteilung wird in [16] gegeben. Dabei ist zu beachten, dass sowohl die Schiefestatik von Mardia als auch die Schiefestatik von Malkovich und Afifi inkonsistent gegenüber bestimmten sphärisch-symmetrischen Alternativen mit existierendem sechsten Moment der Radienverteilung sind (siehe [5, 4]). Da die untersuchte Teststatistik von Cox und Small ähnliche Schwächen aufweist, wurden diese Teststatistiken bewusst zum Vergleich angeführt.

Für aktuellere Arbeiten zum Thema Anpassungstests auf multivariate Normalverteilung wollen wir auf [1, 27] hinweisen.

8.4 Simulationen unter festen Alternativen

In diesem Abschnitt soll die Güte der vorgeschlagenen Teststatistik anhand von verschiedenen Alternativen untersucht werden. Hinweise zu Eigenschaften und zur Erzeugung der verwendeten Verteilungen sind in [20] zu finden. Es wurden für die Dimensionen $d = 2, 3$ jeweils 10000 und für $d = 5, 10$ jeweils 5000 Simulationsdurchläufe ermittelt.

8.4.1 Multivariate Pearson Typ II und Typ VII-Verteilung

Die schon in Kapitel 5 verwendeten multivariaten Verteilungsfamilien Pearson Typ II und Pearson Typ VII werden auch in den Simulationen als Beispielverteilungen für sphärisch-symmetrische Alternativen betrachtet. Hierbei ist wegen der Inkonsistenz der Teststatistik T_n unter gewissen sphärisch-symmetrischen Verteilungen eine geringe Güte der Teststatistik von Cox und Small zu erwarten.

a	$d = 2$	$d = 3$	$d = 5$	$d = 10$
-0.5	0.0309	0.0225	0.0292	0.0326
-0.25	0.0266	0.0219	0.0224	0.0310
0	0.0218	0.0207	0.0262	0.0302
0.25	0.0205	0.0184	0.0286	0.0284
0.5	0.0210	0.0225	0.0338	0.0316
0.75	0.0218	0.0221	0.0318	0.0296
1	0.0214	0.0236	0.0266	0.0274
2	0.0256	0.0240	0.0316	0.0292
3	0.0295	0.0263	0.0360	0.0324
4	0.0286	0.0321	0.0372	0.0324
5	0.0298	0.0325	0.0286	0.0372
10	0.0379	0.0411	0.0378	0.0328
15	0.0433	0.0392	0.0372	0.0360
20	0.0413	0.0397	0.0396	0.0266

Tabelle 8.8: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für Pearson Typ II-Verteilungen mit Parameter a ($n = 25$; $d = 2, 3, 5, 10$)

Die multivariate Pearson Typ II-Verteilung konvergiert bei wachsendem Parameter a gegen die multivariate Normalverteilung, siehe Bemerkung 5.2.1. Dies erklärt die in Tabelle 8.8 angedeutete Konvergenz der Güte der Teststatistik T_n gegen das 5%-Niveau. Da die Teststatistik allerdings unter bestimmten sphärisch-symmetrischen Alternativen dasselbe asymptotische Verhalten aufweist, bestätigen die konstante Nichteinhaltung des Niveaus in Tabelle 8.8 die schon theoretisch angedeuteten Probleme der Teststatistik solche Alternativen zu erkennen.

a	$d = 2$	$d = 3$	$d = 5$	$d = 10$
1	0.8570	0.9506	0.9720	0.8390
2	0.5731	0.7169	0.7696	0.3256
3	0.3843	0.4921	0.5132	0.1330
4	0.2954	0.3622	0.3556	0.0826
5	0.2156	0.2701	0.2684	0.0662
10	0.1102	0.1284	0.1100	0.0420
15	0.0871	0.0901	0.0850	0.0398
20	0.0741	0.0756	0.0670	0.0366

Tabelle 8.9: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für Pearson Typ VII-Verteilungen mit Parameter a ($n = 25$; $d = 2, 3, 5, 10$)

Bei den Simulationen der Pearson Typ VII-Verteilungen standen die Parameter a und N durch $N = \frac{a+d}{2}$ in Zusammenhang. Diese Form der Pearson Typ VII-Verteilung wird auch als generalisierte multivariate t -Verteilung mit a Freiheitsgraden bezeichnet. Die Teststatistik von Cox und Small erkennt bei kleinen Dimensionen die Alternative noch sehr gut, für steigendes a und für höhere Dimensionen nimmt allerdings die Güte deutlich ab. Für die deutliche Niveauunterschreitung bei $d = 10$ ist in diesem Fall wahrscheinlich der recht geringe Stichprobenumfang von $n = 25$ verantwortlich. Wie schon bei der Klasse der Pearson Typ II-Verteilungen konvergiert auch die Pearson Typ VII-Verteilung für $a \rightarrow \infty$ gegen die multivariate Normalverteilung, siehe dazu Bemerkung 5.3.1, was auch die bei wachsendem Parameter a abfallende Güte der Teststatistik T_n hin zum 5%-Niveau erklärt.

8.4.2 Multivariate Burr, Pareto und Logistische Verteilungen

Als Beispiele für Verteilungen, welche nicht zur Klasse der elliptisch-symmetrischen oder sphärisch-symmetrischen Verteilungen gehören, verwenden wir die multivariate Burr, Pareto und Logistische Verteilungen. Wie in [20], S. 160 f. nachzulesen, ist die Standardform der allen drei multivariaten Verteilungen zugrunde liegenden Wahrscheinlichkeitsdichte in Abhängigkeit des Parameters $a > 0$ gegeben durch

$$f_a(x_1, \dots, x_d) = \frac{\Gamma(a+d)}{\Gamma(a)a^d} \prod_{j_1=1}^d x_{j_1}^{-\frac{1}{a}-1} \left[\sum_{j_2=1}^d x_{j_2}^{-\frac{1}{a}} - d + 1 \right]^{-a-d} \mathbf{1}_{\{(x_1, \dots, x_d) \in [0,1]^d\}}. \quad (8.1)$$

Die zur Dichte f_a gehörende Verteilungsfunktion ist für $(x_1, \dots, x_d) \in [0, 1]^d$ gegeben durch

$$F_a(x_1, \dots, x_d) = \left[\sum_{j=1}^d x_j^{-\frac{1}{a}} - d + 1 \right]^{-a}.$$

Die entsprechenden multivariaten Burr-, Pareto- und Logistische Verteilungen lassen sich durch einfache Koordinatentransformationen aus der Wahrscheinlichkeitsdichte f_a aus (8.1) direkt ableiten. Dabei beschreibt der Parameter a in (8.1) die Abhängigkeit in den Komponenten der erzeugten Zufallsvektoren. So gilt zum Beispiel im bivariaten Fall für $(x_1, x_2) \in [0, 1]^2$

$$\lim_{a \rightarrow \infty} F_a(x_1, x_2) = x_1 x_2, \quad (8.2)$$

$$\lim_{a \rightarrow 0} F_a(x_1, x_2) = \min(x_1, x_2), \quad (8.3)$$

was in (8.2) der Unabhängigkeit der Komponenten und in (8.3) der totalen Abhängigkeit der Komponenten entspricht. Die Erzeugung der einzelnen Komponenten eines multivariaten Burr, Pareto und Logistisch verteilten Zufallsvektors ist in der folgenden Tabelle nachzulesen (siehe auch [20], S. 170). Dabei seien Y_1, \dots, Y_d unabhängig identisch $Exp(1)$ -verteilte Zufallsvariablen und X eine weitere unabhängige $\Gamma(\alpha, 1)$ -verteilte Zufallsvariable.

Marginale Verteilung	Erzeugung der Komponente X_j
Gleichv.	$\left(1 + \frac{Y_j}{X}\right)^{-a}$
Logistische V.	$-\ln\left(\frac{Y_j}{X}\right)$
Burr V.	$\left[1 - \left(1 + \frac{Y_j}{X}\right)^{-a}\right]^{-\frac{1}{a}} - 1$
Pareto V.	$\left[1 - \left(1 + \frac{Y_j}{X}\right)^{-a}\right]^{-\frac{1}{a}}$
Normal V.	$\Phi^{-1}\left[\left(1 + \frac{Y_j}{X}\right)^{-a}\right]$

Tabelle 8.10: Erzeugungsvorschriften für multivariat Burr-, Pareto- und Logistisch verteilte Zufallsvektoren

Für die verwendete Simulation wurden der Parameter a auf 1 gesetzt. In den Abbildungen 8.1 und 8.2 werden die entsprechenden erzeugten Verteilungen im Fall von $n = 1000$ Realisierungen anhand einer Punktwolke visualisiert. Dabei werden insbesondere die stark besetzten Enden der Burr- und Pareto-Verteilungen deutlich.

Die in Tabelle 8.11 dargestellten Ergebnisse weisen eine sehr hohe Güte für die marginalen Burr- und Pareto-Verteilungen auf. Am schlechtesten wird die Gleichverteilung erkannt. Die Logistische Verteilung wird, obwohl sie am ehesten einer elliptisch-symmetrischen Verteilung gleicht (siehe Abbildung 8.1) noch besser erkannt. Wie schon bei den Pearson Typ II- und Pearson Typ VII-Verteilungen zu erkennen war, fällt bei der marginalen Gleichverteilung und der marginalen Logistischen Verteilung bei der Dimension $d = 10$ die Ablehnhäufigkeit unter das Testniveau von $\alpha = 0.05$. Dies lässt sich auch hier auf den für die Dimension

Marginale Verteilung	$d = 2$	$d = 3$	$d = 5$	$d = 10$
Gleichv.	0.1174	0.1311	0.124	0.047
Logistische V.	0.2264	0.223	0.1534	0.04
Burr V.	0.9973	0.9999	1	0.9386
Pareto V.	0.9948	1	1	0.939

Tabelle 8.11: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für multivariate Burr-, Pareto- und Logistische Verteilungen ($n = 25$; $d = 2, 3, 5, 10$)

Marginale Verteilung	$d = 2$	$d = 3$	$d = 5$	$d = 10$
Gleichv.	0.1994	0.211	0.3116	0.0752
Logistische V.	0.4664	0.5646	0.7514	0.391
Burr V.	1	1	1	1
Pareto V.	1	1	1	1

Tabelle 8.12: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der Teststatistik T_n für multivariate Burr, Pareto und Logistische Verteilungen ($n = 50$; $d = 2, 3, 5, 10$)

$d = 10$ verhältnismäßig geringen Stichprobenumfang von $n = 25$ zurückführen. Diese Schlussfolgerung legen die Ergebnisse aus Tabelle 8.12 nahe, da bei dem dort verwendeten Stichprobenumfang von $n = 50$ deutlich bessere Ergebnisse erzielt wurden.

8.5 Vergleich der Güte der betrachteten Anpassungstests

In diesem Abschnitt wollen wir die Ergebnisse der Simulationsstudie vergleichen. In der Monte-Carlo Simulation wurden ein Stichprobenumfang von $n = 25$ und verschiedene Dimensionen $d = 2, 3, 5$ betrachtet. Bei den Dimensionen $d = 2$ und $d = 3$ wurden 10000 Simulationsdurchläufe durchgeführt, bei der Dimension $d = 5$ wurde auf Grund hoher Rechenzeiten die Zahl der Simulationsdurchläufe auf 5000 beschränkt. Das Maximumfunktional in der Schiefestatistik von Malkovich und Afifi wurde mit denselben Methoden wie bei der Teststatistik von Cox und Small implementiert. Auch hier wurden für die Dimensionen $d = 2, 3$ jeweils 1000 Punkte auf der Oberfläche der Einheitssphäre generiert und für die Dimension $d = 5$ wurden 7000 Punkte betrachtet. Bei der BHEP-Statistik wurde der Glättungsparameter $\beta = 1$ gesetzt.

Bei zugrunde liegender Pearson Typ II-Verteilung schneiden die Schiefestatistik von Mardia $b_{n,d}^{(1)}$ und die Schiefestatistik von Malkovich und Afifi $b_{n,M}^{(1)}$ erwartungsgemäß extrem schlecht ab. Die Teststatistik T_n von Cox und Small hält wie in Tabelle 8.8 zwar nicht das Niveau ein, lehnt aber tendenziell etwas öfter die Hypothese ab. Die BHEP-Statistik und die Teststatistik von Székely und Rizzo erzielen vor allem für kleine Werte des Parameters a bessere Ergebnisse. Für wachsendes a aber fallen auch sie unter das Niveau, was aber mit der Ähnlichkeit der Pearson Typ II-Verteilung mit der multivariaten Normalverteilung zusammenhängt (siehe Bemerkung 5.2.1). Die Güte aller betrachteten Teststatistiken konvergiert mit wachsendem

a	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
-0.5	0.0008	0.0015	0.3665	0.5099	0.0309
-0.25	0.0007	0.0017	0.1766	0.2689	0.0266
0	0.0005	0.0010	0.0945	0.1466	0.0218
0.25	0.0007	0.0013	0.0616	0.0994	0.0205
0.5	0.0009	0.0012	0.0432	0.0728	0.0210
0.75	0.0013	0.0025	0.0359	0.0611	0.0218
1	0.0019	0.0035	0.0307	0.0491	0.0214
2	0.0042	0.0054	0.0243	0.0363	0.0256
3	0.0081	0.0097	0.0256	0.0373	0.0295
4	0.0108	0.0117	0.0241	0.0344	0.0286
5	0.0132	0.0152	0.0269	0.0375	0.0298
10	0.0240	0.0251	0.0307	0.0414	0.0379
15	0.0307	0.0310	0.0327	0.0416	0.0433
20	0.0356	0.0360	0.0336	0.0455	0.0413

Tabelle 8.13: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 2$)

Parameter a wie erwartet gegen das 5%-Niveau.

Unter sphärisch-symmetrischer Alternative mit Pearson Typ VII-verteilten Zufallsvektoren erkennen alle betrachteten Teststatistiken die Alternativen sehr gut, wobei bei wachsendem Parameter a , die Güte aller Teststatistiken abfällt. Dies ist auch hier durch die Ähnlichkeit der Verteilung mit der multivariaten Normalverteilung zu erklären.

Bei den hier betrachteten nicht elliptisch- bzw. sphärisch-symmetrischen Alternativen diversifizieren die Ergebnisse etwas mehr. Falls als Marginalverteilungen der Komponenten der Zufallsvektoren die Gleichverteilung angenommen wird, schneiden die beiden Schiefestatistiken von Mardia und Malkovich und Afifi sehr schlecht ab und können hier wie bei den Pearson Typ II-Verteilungen das Niveau eindeutig nicht halten. Die Teststatistik von Cox und Small setzt sich hier deutlich von den Schiefestatistiken ab (siehe Tabelle 8.11), kann aber nicht mit der BHEP-Statistik oder der Teststatistik von Székely und Rizzo mithalten. Liegen hingegen Logistische Marginalverteilungen vor, so kehrt sich das beobachtete Ergebnis um, da die beiden Schiefestatistiken am Besten abschneiden, gefolgt von der Statistik von Cox und Small. In diesem Fall erzielen $T_{n,\beta}$ und $T_n^{(SR)}$ die geringste Ablehnhäufigkeit. Bei den beiden extremen Fällen der Burr- und Pareto-Marginalverteilungen lehnen die betrachteten Statistiken fast immer ab.

a	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
-0.5	0.0001	0.0002	0.2190	0.2151	0.0225
-0.25	0.0001	0.0005	0.1155	0.1168	0.0219
0	0.0002	0.0004	0.0707	0.0725	0.0207
0.25	0.0008	0.0016	0.0508	0.0515	0.0184
0.5	0.0008	0.0018	0.0374	0.0413	0.0225
0.75	0.0005	0.0019	0.0338	0.0364	0.0221
1	0.0006	0.0018	0.0284	0.0340	0.0236
2	0.0028	0.0045	0.0223	0.0260	0.0240
3	0.0046	0.0057	0.0223	0.0269	0.0263
4	0.0065	0.0094	0.0217	0.0267	0.0321
5	0.0100	0.0115	0.0216	0.0276	0.0325
10	0.0192	0.0243	0.0245	0.0325	0.0411
15	0.0257	0.0286	0.0327	0.0398	0.0392
20	0.0297	0.0313	0.0311	0.0407	0.0397

Tabelle 8.14: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 3$)

a	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
-0.5	0	0.0001	0.0895	0.0237	0.0292
-0.25	0	0.0004	0.0622	0.0183	0.0224
0	0.0001	0.0006	0.0448	0.0144	0.0262
0.25	0.0001	0.0004	0.0346	0.0127	0.0286
0.5	0.0003	0.0004	0.0304	0.0105	0.0338
0.75	0.0003	0.0007	0.0258	0.0117	0.0318
1	0.0004	0.0004	0.0239	0.0099	0.0266
2	0.0015	0.0007	0.0199	0.0131	0.0316
3	0.0026	0.0035	0.0187	0.0153	0.0360
4	0.0054	0.0039	0.0170	0.0164	0.0372
5	0.0063	0.0043	0.0182	0.0173	0.0286
10	0.0149	0.0107	0.0202	0.0255	0.0378
15	0.0198	0.0155	0.0215	0.0301	0.0372
20	0.0270	0.0180	0.0252	0.0365	0.0396

Tabelle 8.15: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für Pearson Typ II-Verteilungen mit Parameter a ($n = 25; d = 5$)

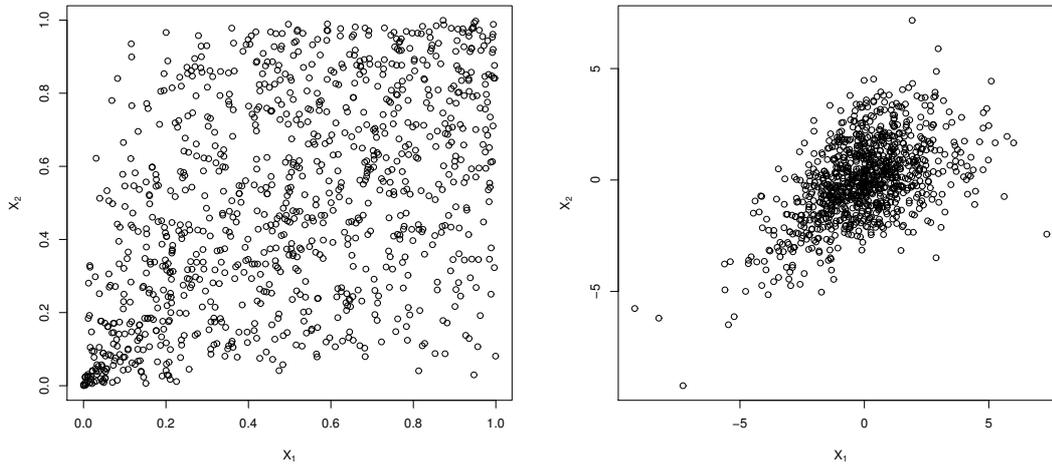


Abbildung 8.1: Punktwolken von einer multivariaten Burr-, Pareto- und Logistische-Verteilung mit gleichverteilten Komponenten (links) und logistisch-verteilten Komponenten (rechts)

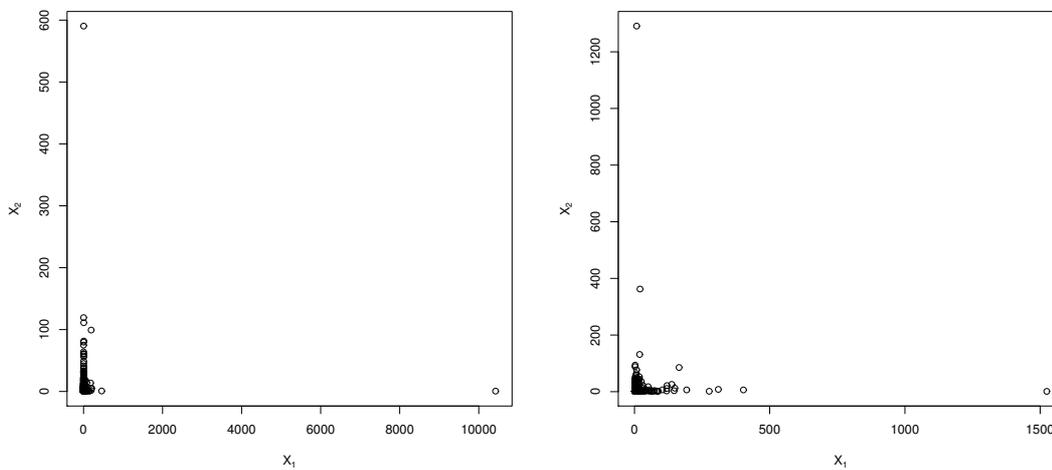


Abbildung 8.2: Punktwolken von einer multivariaten Burr-, Pareto- und Logistische-Verteilung mit Burr-verteilten Komponenten (links) und Pareto-verteilten Komponenten (rechts)

a	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
1	0.9664	0.9607	0.9875	0.9892	0.8570
2	0.7859	0.7664	0.8089	0.8132	0.5731
3	0.5924	0.5738	0.5708	0.5805	0.3843
4	0.4563	0.4377	0.4092	0.4146	0.2954
5	0.3643	0.3509	0.2890	0.2970	0.2156
10	0.1817	0.1750	0.1226	0.1332	0.1102
15	0.1252	0.1227	0.0830	0.0900	0.0871
20	0.1042	0.1058	0.0722	0.0806	0.0741

Tabelle 8.16: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ für Pearson Typ VII-Verteilungen mit Parameter a ($n = 25; d = 2$)

a	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
1	0.9930	0.9865	0.9970	0.9980	0.9506
2	0.8947	0.8613	0.8914	0.9126	0.7169
3	0.7383	0.6876	0.6712	0.7130	0.4921
4	0.5885	0.5462	0.4933	0.5443	0.3622
5	0.4752	0.4295	0.3534	0.4092	0.2701
10	0.2342	0.2127	0.1340	0.1735	0.1284
15	0.1554	0.1446	0.0841	0.1124	0.0901
20	0.1275	0.1180	0.0727	0.0966	0.0756

Tabelle 8.17: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ für Pearson Typ VII-Verteilungen mit Parameter a bei einem Stichprobenumfang von ($n = 25; d = 3$)

a	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
1	0.9998	0.9972	0.9992	0.9996	0.9720
2	0.9682	0.9302	0.9388	0.9690	0.7696
3	0.8878	0.8136	0.7682	0.8614	0.5132
4	0.7670	0.6734	0.5670	0.7222	0.3556
5	0.6484	0.5530	0.4022	0.5658	0.2684
10	0.3158	0.2672	0.1348	0.2470	0.1100
15	0.2098	0.1794	0.0762	0.1468	0.0850
20	0.1560	0.1336	0.0600	0.1136	0.0670

Tabelle 8.18: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ für Pearson Typ VII-Verteilungen mit Parameter a ($n = 25; d = 5$)

Dimension	Marginale Verteilung	$b_{n,d}^{(1)}$	$b_{n,M}^{(1)}$	$T_{n,\beta}$	$T_n^{(SR)}$	T_n
$d = 2$	Gleichv.	0.0134	0.0156	0.1723	0.2629	0.1174
	Logistische V.	0.3084	0.2590	0.1741	0.1788	0.2264
	Burr V.	0.9998	0.9996	1	1	0.9973
	Pareto V.	0.9999	0.9996	1	1	0.9948
$d = 3$	Gleichv.	0.0200	0.0257	0.2118	0.2560	0.1311
	Logistische V.	0.3805	0.2841	0.1804	0.2280	0.2230
	Burr V.	1	0.9999	1	1	0.9999
	Pareto V.	1	1	1	1	1
$d = 5$	Gleichv.	0.0306	0.0314	0.2454	0.2218	0.1240
	Logistische V.	0.4342	0.2918	0.1498	0.2760	0.1534
	Burr V.	1	1	1	1	1
	Pareto V.	1	1	1	1	1

Tabelle 8.19: Ablehnhäufigkeit zum Niveau $\alpha = 0.05$ der klassischen Testverfahren für multivariate Burr, Pareto und Logistische Verteilungen ($n = 25; d = 2, 3, 5$)

Kapitel 9

Abschließende Bemerkungen

Die vorliegende Arbeit lässt die Bestimmung der asymptotischen Grenzverteilung der Teststatistik T_n unter festen Alternativen offen. Das für diese Fragestellung zu lösende Problem ist die richtige Zentrierung von T_n , was sich wegen des Maximumfunktionals als sehr diffizil darstellt. In diesem Sinne konnte Vermutung 7.1.3 nicht bewiesen werden.

Die Teststatistik T_n von Cox und Small lässt einige Verallgemeinerungen zu. Da T_n das Vorhandensein von quadratischen Abhängigkeiten untersucht, könnte man analog eine Prüfgröße zur Untersuchung von höheren, wie z.B. kubischen Abhängigkeiten konstruieren. Allerdings ist schon im kubischen Fall mit größeren Komplikationen zu rechnen, da das für eine solche Statistik grundlegende Funktional als Lösung des Optimierungsproblems noch unhandlicher sein wird, als es bei der Untersuchung der quadratischen Abhängigkeiten ohnehin schon ist. Weiter sind die für die Praxis auftretenden Probleme der Inkonsistenz der Teststatistik T_n gegenüber bestimmten sphärisch-symmetrischen Alternativen und die schwierige numerische Berechnung auch im kubischen Fall zu erwarten.

In natürlicher Weise bietet sich die Teststatistik T_n auch als Prüfgröße eines Anpassungstests auf die affin invariante Klasse der elliptisch-symmetrischen Verteilungen an. Das Hypothesentestproblem ist in diesem Fall gegeben durch die Hypothese

$$H_0 : \mathbb{P}^X \text{ gehört zur Klasse der elliptisch-symmetrischen Verteilungen}$$

gegen die Alternative

$$H_1 : \mathbb{P}^X \text{ gehört nicht zur Klasse der elliptisch-symmetrischen Verteilungen.}$$

Theorem 4.4.1 liefert zwar für die Teststatistik T_n eine Aussage bezüglich deren asymptotischer Verteilung bei Vorliegen einer festen sphärisch-symmetrischen Verteilung unter einer Momentenbedingung, für die Konstruktion eines auf T_n basierenden Tests, der asymptotisch ein vorgegebenes Testniveau über ganz H_0 (gegebenenfalls unter Einschränkung auf Verteilungen, die einer Momentenrestriktion unterliegen) einhält, ist jedoch ein geeignetes Bootstrapverfahren erforderlich.

Anhang A

Technische Berechnungen

Im folgenden Lemma wird die in (2.9) angegebene Darstellung der Teststatistik als V -Statistik gezeigt.

Lemma A.0.1

Für $V_n(b)$ aus (2.8) gilt die Darstellung

$$V_n(b) = \frac{1}{n^2} \sum_{j_1, j_2=1}^n h_b(Y_{n, j_1}, Y_{n, j_2}), \quad b \in \mathcal{S}^{d-1},$$

wobei

$$h_b(x, y) := (b'x)^2(b'y)^2(x'y - b'xb'y), \quad x, y \in \mathbb{R}^d, \quad b \in \mathcal{S}^{d-1}.$$

Beweis:

Es gilt

$$\begin{aligned} V_n(b) &= \left\| \frac{1}{n} \sum_{j=1}^n Y_{n, j} (b'Y_{n, j})^2 \right\|^2 - \left(\frac{1}{n} \sum_{j=1}^n (b'Y_{n, j})^3 \right)^2 \\ &= \frac{1}{n^2} \left[\left(\sum_{j=1}^n Y_{n, j} (b'Y_{n, j})^2 \right)' \left(\sum_{l=1}^n Y_{n, l} (b'Y_{n, l})^2 \right) - \sum_{j_1, j_2=1}^n (b'Y_{n, j_1})^3 (b'Y_{n, j_2})^3 \right] \\ &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n [(b'Y_{n, j_1})^2 (b'Y_{n, j_2})^2 Y_{n, j_1}' Y_{n, j_2} - (b'Y_{n, j_1})^3 (b'Y_{n, j_2})^3] \\ &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n h_b(Y_{n, j_1}, Y_{n, j_2}). \end{aligned}$$

□

A.1 Gemischte Momente von Komponenten von $U \sim \mathcal{U}(\mathcal{S}^{d-1})$

In diesem Abschnitt werden explizite Ausdrücke für speziell benötigte Oberflächenintegrale angegeben. Diese Oberflächenintegrale bilden das Fundament zur Berechnung gewisser Erwartungswerte im Zusammenhang mit sphärisch-symmetrisch verteilten Zufallsvektoren. Zur Berechnung verwenden wir z.B. die in [3] auf S. 41 f. angegebenen Verfahren. Danach gilt für $d \geq 2$ und $\gamma_1, \dots, \gamma_d \in \mathbb{N}_0$

$$\int_{\mathcal{S}^{d-1}} (\xi_1^2)^{\gamma_1} \dots (\xi_d^2)^{\gamma_d} d\mathcal{S}^{d-1}(\xi) = \left(d + 2 \sum_{j=1}^d \gamma_j \right) \frac{\prod_{j=1}^d \Gamma\left(\gamma_j + \frac{1}{2}\right)}{\Gamma\left(1 + \frac{d}{2} + \sum_{j=1}^d \gamma_j\right)}. \quad (\text{A.1})$$

Nach einigen Rechnungen ergeben sich mit den Abkürzungen

$$\zeta_{d,j} := |\mathcal{S}^{d-1}| \prod_{k=0}^j (d+2k)^{-1}, \quad j \in \mathbb{N}, \quad (\text{A.2})$$

die Werte in der folgenden Tabelle. Alternativ kann man in Spezialfällen die Methoden aus

Integral	Wert
$\int_{\mathcal{S}^{d-1}} u_1^2 u_2^2 d\mathcal{S}^{d-1}(u)$	$\zeta_{d,1}$
$\int_{\mathcal{S}^{d-1}} u_1^4 d\mathcal{S}^{d-1}(u)$	$3\zeta_{d,1}$
$\int_{\mathcal{S}^{d-1}} u_1^2 u_2^2 u_3^2 d\mathcal{S}^{d-1}(u)$	$\zeta_{d,2}$
$\int_{\mathcal{S}^{d-1}} u_1^4 u_2^2 d\mathcal{S}^{d-1}(u)$	$3\zeta_{d,2}$
$\int_{\mathcal{S}^{d-1}} u_1^6 d\mathcal{S}^{d-1}(u)$	$15\zeta_{d,2}$

Tabelle A.1: Explizite Werte bestimmter Oberflächenintegrale

C.4 bzw. die in [13], S. 72, in Theorem 3.3 angegebene Formel

$$E\left(\prod_{j=1}^d U_j^{m_j}\right) = \begin{cases} \frac{1}{\left(\frac{d}{2}\right)!} \prod_{j=1}^d \frac{(2l_j)!}{4^{l_j} (l_j)!}, & \text{falls } m_j = 2l_j \text{ gerade, } j = 1, \dots, d, m = 2l; \\ 0, & \text{falls ein } m_j \text{ ungerade;} \end{cases} \quad (\text{A.3})$$

für die gemischten Momente der Komponenten von $U = (U_1, \dots, U_d) \sim \mathcal{U}(\mathcal{S}^{d-1})$ verwenden. Dabei seien kurz $m := \sum_{j=1}^d m_j$, sowie für $t \in \mathbb{R}$ $t^{[l]} := t(t+1) \cdots (t+l-1)$ gesetzt.

A.2 Kovarianzen

In den folgenden Lemmata werden explizite Ausdrücke für die in (3.8) und (3.9) aufgeführten Kovarianzen angegeben. Es gelten jeweils die gleichen Voraussetzungen an die Zufallsvektoren X_1, \dots, X_n wie in den jeweiligen Kapiteln.

Lemma A.2.1

Für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$ gilt

$$\Phi_{k,l}^{(1)}(b, c) = \frac{2m_6}{d(d+2)(d+4)} (2(b'c)(b_k c_l + c_k b_l) + (b_k b_l + c_k c_l)). \quad (\text{A.4})$$

Beweis:

Es gilt für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$

$$\Phi_{k,l}^{(1)}(b, c) = E(f_k(X, b) f_l(X, c)) = m_6 E((b'U)^2 (c'U)^2 U_k U_l).$$

Weiter gilt mit Symmetrieüberlegungen

$$\begin{aligned} E((b'U)^2 (c'U)^2 U_k U_l) &= E\left(\sum_{j_1, j_2, j_3, j_4=1}^d b_{j_1} b_{j_2} c_{j_3} c_{j_4} U_{j_1} U_{j_2} U_{j_3} U_{j_4} U_k U_l\right) \\ &= E\left(2(b_k^2 c_k c_l + b_k b_l c_k^2) U_k^4 U_l^2 + 2(b_l^2 c_l c_k + b_l b_k c_l^2) U_l^4 U_k^2 \right. \\ &\quad \left. + U_k^2 U_l^2 \sum_{\substack{j=1 \\ j \neq k \neq l \neq j}}^d (2(b_k b_l c_j^2 + c_k c_l b_j^2) \right. \\ &\quad \left. + 4(b_j c_j b_k c_l + b_j c_j c_k b_l)) U_j^2\right) \\ &= 2\left[(b_k^2 c_k c_l + b_k b_l c_k^2) E(U_k^4 U_l^2) + (b_l^2 c_l c_k + b_l b_k c_l^2) E(U_l^4 U_k^2) \right. \\ &\quad \left. + \sum_{\substack{j=1 \\ j \neq k \neq l \neq j}}^d (b_k b_l c_j^2 + c_k c_l b_j^2) E(U_k^2 U_l^2 U_j^2) \right. \\ &\quad \left. + 2 \sum_{\substack{j=1 \\ j \neq k \neq l \neq j}}^d b_j c_j (b_k c_l + c_k b_l) E(U_k^2 U_l^2 U_j^2) \right]. \end{aligned}$$

Nach Tabelle A.1 gilt für $j_1, j_2, j_3 \in \{1, \dots, d\}$ mit $j_1 \neq j_2 \neq j_3 \neq j_1$

$$E(U_{j_1}^4 U_{j_2}^2) = \frac{1}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} u_1^4 u_2^2 d\mathcal{S}^{d-1}(u) = \frac{3}{d(d+2)(d+4)}, \quad (\text{A.5})$$

$$E(U_{j_1}^2 U_{j_2}^2 U_{j_3}^2) = \frac{1}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} u_1^2 u_2^2 u_3^2 d\mathcal{S}^{d-1}(u) = \frac{1}{d(d+2)(d+4)}. \quad (\text{A.6})$$

Mit der üblichen Schreibweise des Skalarprodukts ergibt sich

$$\begin{aligned} \Phi_{k,l}^{(1)}(b, c) &= \frac{6m_6}{d(d+2)(d+4)} [(b_k^2 c_k c_l + b_k b_l c_k^2) + (b_l^2 c_l c_k + b_l b_k c_l^2)] \\ &\quad + \frac{2m_6}{d(d+2)(d+4)} (b_k b_l (1 - c_k^2 - c_l^2) + c_k c_l (1 - b_k^2 - b_l^2)) \\ &\quad + 2(b_k c_l + c_k b_l)(b'c - b_k c_k - b_l c_l) \\ &= \frac{2m_6}{d(d+2)(d+4)} (2(b'c)(b_k c_l + c_k b_l) + (b_k b_l + c_k c_l)) \end{aligned}$$

und daraus folgt die Behauptung. □

Lemma A.2.2

Für $b, c \in \mathcal{S}^{d-1}$ und $k \in \{1, \dots, d\}$ gilt

$$\Phi_k^{(2)}(b, c) = \frac{3m_6}{d(d+2)(d+4)} (2b_k (b'c)^2 + 2c_k (b'c) + b_k). \quad (\text{A.7})$$

Beweis:

Seien $b, c \in \mathcal{S}^{d-1}$ und $k \in \{1, \dots, d\}$ beliebig aber fest. Dann gilt

$$\Phi_k^{(2)}(b, c) = E((b'X)^3 f_k(X, c)) = m_6 E((b'U)^3 (c'U)^2 U_k).$$

Es existiert eine orthogonale $d \times d$ -Matrix $D \in \mathcal{O}(d)$ mit folgender Eigenschaft: Für festes $l \in \{1, \dots, d\}$ mit $l \neq k$ gilt

$$\begin{pmatrix} b'U \\ c'U \\ e_k'U \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} e_l'U \\ \xi'U \\ \tau'U \end{pmatrix}.$$

Dabei bezeichnen $e_k, e_l \in \mathcal{S}^{d-1}$ den k -ten bzw. l -ten Einheitsvektor in \mathbb{R}^d , und es gilt

$$\begin{aligned} e_l &=: Db, \\ \xi &=: Dc, \\ \tau &=: De_k. \end{aligned}$$

Damit folgt mit der Drehinvarianz der Gleichverteilung auf der Einheitssphäre

$$\begin{aligned}
E((b'U)^3(c'U)^2U_k) &= E((b'D'DU)^3(c'D'DU)^2(e'_kD'DU)) \\
&= E(((Db)'U)^3((Dc)'U)^2((De_k)'U)) \\
&= E((\xi'U)^2(\tau'U)U_l^3) \\
&= E\left(\sum_{j_1, j_2, j_3=1}^d \tau_{j_1} \xi_{j_2} \xi_{j_3} U_{j_1} U_{j_2} U_{j_3} U_l^3\right) \\
&= E\left(\tau_l \xi_l^2 U_l^6 + U_l^4 \sum_{\substack{j=1 \\ j \neq l}}^d (\tau_l \xi_j^2 + 2\xi_l \tau_j \xi_j) U_j^2\right) \\
&= \tau_l \xi_l^2 E(U_l^6) + \sum_{\substack{j=1 \\ j \neq l}}^d (\tau_l \xi_j^2 + 2\xi_l \tau_j \xi_j) E(U_l^4 U_j^2).
\end{aligned}$$

Mit dieser Darstellung gilt

$$\begin{aligned}
\Phi_k^{(2)}(b, c) &= \frac{3m_6}{d(d+2)(d+4)} \left[5\tau_l \xi_l^2 + \sum_{\substack{j=1 \\ j \neq l}}^d (\tau_l \xi_j^2 + 2\xi_l \tau_j \xi_j) \right] \\
&= \frac{3m_6}{d(d+2)(d+4)} [5\tau_l \xi_l^2 + \tau_l(1 - \xi_l^2) + 2\xi_l((\tau' \xi) - \tau_l \xi_l)] \\
&= \frac{3m_6}{d(d+2)(d+4)} [2\tau_l \xi_l^2 + \tau_l + 2\xi_l(\tau' \xi)].
\end{aligned}$$

Nach Rücktransformation folgt unter Beachtung von

$$\begin{aligned}
\tau_l &= e'_l \tau = e'_l D' D \tau = b'_k e_k = b_k, \\
\xi_l &= e'_l \xi = e'_l D' D \xi = b'_k c = b_k c, \\
\tau' \xi &= \tau' D' D \xi = e'_k c = c_k,
\end{aligned}$$

die Behauptung. □

Lemma A.2.3

Es gilt für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$

$$\Phi_{k,l}^{(1,*)}(b, c) = \Phi_{k,l}^{(1)}(b, c) - \frac{4m_4}{d(d+2)} (2b_k c_l (b'_k c) + c_k c_l + b_k b_l) + 4b_k c_l (b'_k c) + 2c_k c_l + 2b_k b_l. \tag{A.8}$$

Dabei bezeichnet $\Phi_{k,l}^{(1)}(b, c)$ die in (A.4) angegebene Abbildung.

Beweis:

Für $b, c \in \mathcal{S}^{d-1}$ gilt

$$\begin{aligned}
\Phi_{k,l}^{(1,*)}(b, c) &= \text{Cov}(C_{1,k}^{**}(b), C_{1,l}^{**}(c)) \\
&= E[(f_k(X_1, b) - 2b'X_1b_k - \varphi_k(X_1))(f_l(X_1, c) - 2c'X_1c_l - \varphi_l(X_1))] \\
&= E(f_k(X_1, b)f_l(X_1, c)) - 2b_kE(b'X_1f_l(X_1, c)) - E(\varphi_k(X_1)f_l(X_1, c)) \\
&\quad - 2c_lE(f_k(X_1, b)c'X_1) + 4b_kc_lE(b'X_1c'X_1) + 2c_lE(\varphi_k(X_1)c'X_1) \\
&\quad - E(f_k(X_1, b)\varphi_l(X_1)) + 2b_kE(b'X_1\varphi_l(X_1)) + E(\varphi_k(X_1)\varphi_l(X_1)).
\end{aligned}$$

Die oben auftretenden Erwartungswerte lassen sich wie folgt berechnen:

$$\begin{aligned}
E(f_k(X_1, b)f_l(X_1, c)) &= \Phi_{k,l}^{(1)}(b, c), \\
E(b'X_1f_l(X_1, c)) &= \frac{m_4}{d(d+2)}(2c_l(b'c) + b_l), \\
E(\varphi_k(X_1)f_l(X_1, c)) &= \frac{2m_4}{d(d+2)}c_kc_l, \\
E(b'X_1c'X_1) &= b'c, \\
E(\varphi_k(X_1)c'X_1) &= c_k, \\
E(\varphi_k(X_1)\varphi_l(X_1)) &= 0.
\end{aligned}$$

Die restlichen Erwartungswerte ergeben sich aus Symmetrieüberlegungen aus den oben angegebenen. Zusammengesetzt ergibt dies die behauptete Kovarianzstruktur. \square

Lemma A.2.4

Es gilt für $b, c \in \mathcal{S}^{d-1}$ und $k \in \{1, \dots, d\}$

$$\Phi_k^{(2,*)}(b, c) = \Phi_k^{(2)}(b, c) - \frac{3m_4}{d(d+2)}(4c_k(b'c) + 2b_k) + 6c_k(b'c) + 3b_k.$$

Dabei bezeichnet $\Phi_k^{(2)}(b, c)$ die in (A.7) angegebene Abbildung.

Beweis:

Mit $b, c \in \mathcal{S}^{d-1}$ ist

$$\begin{aligned}
\Phi_k^{(2,*)}(b, c) &= \text{Cov}(D_1^{**}(b), C_{1,k}^{**}(c)) \\
&= E[((b'X_1)^3 - 3b'X_1)(f_k(X_1, c) - 2c_k(c'X_1) - \varphi_k(X_1))] \\
&= E((b'X_1)^3 f_k(X_1, c)) - 2c_kE((b'X_1)^3 c'X_1) - E((b'X_1)^3 \varphi_k(X_1)) \\
&\quad - 3E(b'X_1 f_k(X_1, c)) + 6c_kE((b'X_1)(c'X_1)) + 3E(b'X_1 \varphi_k(X_1)).
\end{aligned}$$

Ferner gelten die Identitäten

$$\begin{aligned} E((b'X_1)^3 f_k(X_1, c)) &= \Phi_k^{(2)}(b, c), \\ E((b'X_1)^3 c'X_1) &= \frac{3m_4}{d(d+2)} b'c, \\ E((b'X_1)^3 \varphi_k(X_1)) &= \frac{3m_4}{d(d+2)} b_k. \end{aligned}$$

Die übrigen Erwartungswerte sind in (4.9)-(4.13) zu finden. Durch kurzes Nachrechnen lässt sich die Behauptung bestätigen. \square

Lemma A.2.5

Es gilt für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$

$$\begin{aligned} \Phi_{k,l}^{\star\star}(b, c) &= \Xi_{2,2,1,1}^{(k,l)}(b, c) - 2b_k \Xi_{1,2,1}^{(l)}(b, c) - e'_l \Upsilon_{5,1}(b, c) + b' \mu_1^{\star\star}(b) \varphi_l \circ \mu_2^{\star\star}(c) \\ &\quad - \Xi_{0,2,1,1}^{(k,l)}(b, c) - \frac{1}{2} e'_l \Upsilon_{5,2}^{(k)}(b, c) - \frac{1}{4} e'_k \mu_2^{\star\star}(b) e'_l \mu_2^{\star\star}(c) \\ &\quad - 2c_l \Xi_{2,1,1}^{(k)}(b, c) + 4b_k c_l (b'c) + 2c_l c' \Upsilon_{3,1}(b) + 2c_l c_k + c_l c' \Upsilon_{3,2}^{(k)}(b) \\ &\quad - e'_k \Upsilon_{5,1}(c, b) + c' \mu_1^{\star\star}(c) \varphi_k \circ \mu_2^{\star\star}(b) + 2b_k b' \Upsilon_{3,1}(c) + \Upsilon_{4,1}(b) - b' \mu_1^{\star\star}(b) c' \mu_1^{\star\star}(c) \\ &\quad + e'_k \Upsilon_{3,1}(c) + \frac{1}{2} \Upsilon_{4,2}^{(k)}(b, c) - \frac{1}{2} e'_k \mu_2^{\star\star}(b) c' \mu_1^{\star\star}(c) - \Xi_{2,0,1,1}^{(k,l)}(b, c) + 2b_k b_l \\ &\quad + e'_l \Upsilon_{3,1}(b) + \frac{1}{2} e'_l \Upsilon_{3,2}^{(k)}(b) - \frac{1}{2} e'_k \Upsilon_{5,2}^{(l)}(c, b) \\ &\quad + b_k b' \Upsilon_{3,2}^{(l)}(c) + \frac{1}{2} \Upsilon_{4,2}^{(l)}(c, b) - \frac{1}{2} b' \mu_1^{\star\star}(b) e'_l \mu_2^{\star\star}(c) \\ &\quad + \frac{1}{2} e'_k \Upsilon_{3,2}^{(l)}(c) + \frac{1}{4} \Upsilon_{4,3}^{(k,l)}(b, c). \end{aligned}$$

Beweis:

Um den Kovarianzkern angeben zu können, benötigen wir die abkürzenden Schreibweisen aus Kapitel 7. Damit folgt für $b, c \in \mathcal{S}^{d-1}$ und $k, l \in \{1, \dots, d\}$ mit $k \neq l$

$$\begin{aligned} \text{Cov}(\mathcal{Z}_{1,k}(b), \mathcal{Z}_{1,l}(c)) &= E \left[(Z_{1,k}(b) - \varphi_k \circ \mu_2^{\star\star}(b)) (Z_{1,l}(c) - \varphi_l \circ \mu_2^{\star\star}(c)) \right] \\ &= E (Z_{1,k}(b) Z_{1,l}(c)) - \varphi_k \circ \mu_2^{\star\star}(b) \varphi_l \circ \mu_2^{\star\star}(c). \end{aligned}$$

Weiter ist

$$\begin{aligned}
E(Z_{1,k}(b)Z_{1,l}(c)) &= E \left[(f_k(X_1, b) - 2(b'X_1)b_k - b'(X_1X_1' - I_d)\mu_1^*(b) - \varphi_k(X_1) \right. \\
&\quad \left. - \frac{1}{2}e'_k(X_1X_1' - I_d)\mu_2^*(b) \right) \cdot (f_l(X_1, c) - 2(c'X_1)c_l \\
&\quad \left. - c'(X_1X_1' - I_d)\mu_1^*(c) - \varphi_l(X_1) - \frac{1}{2}e'_l(X_1X_1' - I_d)\mu_2^*(c) \right) \Big] \\
&= E(f_k(X_1, b)f_l(X_1, c)) - 2b_kE((b'X_1)f_l(X_1, c)) \\
&\quad - E(b'(X_1X_1' - I_d)\mu_1^*(b)f_l(X_1, c)) - E(\varphi_k(X_1)f_l(X_1, c)) \\
&\quad - \frac{1}{2}E(e'_k(X_1X_1' - I_d)\mu_2^*(b)f_l(X_1, c)) - 2c_lE(f_k(X_1, b)(c'X_1)) \\
&\quad + 4b_kc_lE((b'X_1)(c'X_1)) + 2c_lE(b'(X_1X_1' - I_d)\mu_1^*(b)(c'X_1)) \\
&\quad + 2c_lE(\varphi_k(X_1)(c'X_1)) + c_lE(e'_k(X_1X_1' - I_d)\mu_2^*(b)(c'X_1)) \\
&\quad - E(f_k(X_1, b)c'(X_1X_1' - I_d)\mu_1^*(c)) \\
&\quad + 2b_kE((b'X_1)c'(X_1X_1' - I_d)\mu_1^*(c)) \\
&\quad + E(b'(X_1X_1' - I_d)\mu_1^*(b)c'(X_1X_1' - I_d)\mu_1^*(c)) \\
&\quad + E(\varphi_k(X_1)c'(X_1X_1' - I_d)\mu_1^*(c)) \\
&\quad + \frac{1}{2}E(e'_k(X_1X_1' - I_d)\mu_2^*(b)c'(X_1X_1' - I_d)\mu_1^*(c)) \\
&\quad - E(f_k(X_1, b)\varphi_l(X_1)) + 2b_kE((b'X_1)\varphi_l(X_1)) \\
&\quad + E(b'(X_1X_1' - I_d)\mu_1^*(b)\varphi_l(X_1)) + E(\varphi_k(X_1)\varphi_l(X_1)) \\
&\quad + \frac{1}{2}E(e'_k(X_1X_1' - I_d)\mu_2^*(b)\varphi_l(X_1)) \\
&\quad - \frac{1}{2}E(f_k(X_1, b)e'_l(X_1X_1' - I_d)\mu_2^*(c)) \\
&\quad + b_kE((b'X_1)e'_l(X_1X_1' - I_d)\mu_2^*(c)) \\
&\quad + \frac{1}{2}E(b'(X_1X_1' - I_d)\mu_1^*(b)e'_l(X_1X_1' - I_d)\mu_2^*(c)) \\
&\quad + \frac{1}{2}E(\varphi_k(X_1)e'_l(X_1X_1' - I_d)\mu_2^*(c)) \\
&\quad + \frac{1}{4}E(e'_k(X_1X_1' - I_d)\mu_2^*(b)e'_l(X_1X_1' - I_d)\mu_2^*(c)) .
\end{aligned}$$

Mit den Bezeichnungen aus Abschnitt 7.4 gilt

$$\begin{aligned}
E(Z_{1,k}(b)Z_{1,l}(c)) &= \Xi_{2,2,1,1}^{(k,l)}(b,c) - 2b_k\Xi_{1,2,1}^{(l)}(b,c) - e'_l\Upsilon_{5,1}(b,c) + b'\mu_1^\star(b)\varphi_l \circ \mu_2^\star(c) \\
&\quad - \Xi_{0,2,1,1}^{(k,l)}(b,c) - \frac{1}{2}e'_l\Upsilon_{5,2}^{(k)}(b,c) + \frac{1}{2}e'_k\mu_2^\star(b)\varphi_l \circ \mu_2^\star(c) \\
&\quad - 2c_l\Xi_{2,1,1}^{(k)}(b,c) + 4b_kc_l(b'c) + 2c_l c' \Upsilon_{3,1}(b) + 2c_l c_k + c_l c' \Upsilon_{3,2}^{(k)}(b) \\
&\quad - e'_k \Upsilon_{5,1}(c,b) + c' \mu_1^\star(c) \varphi_k \circ \mu_2^\star(b) + 2b_k b' \Upsilon_{3,1}(c) + \Upsilon_{4,1}(b) \\
&\quad - b' \mu_1^\star(b) c' \mu_1^\star(c) + e'_k \Upsilon_{3,1}(c) + \frac{1}{2} \Upsilon_{4,2}^{(k)}(b,c) - \frac{1}{2} e'_k \mu_2^\star(b) c' \mu_1^\star(c) \\
&\quad - \Xi_{2,0,1,1}^{(k,l)}(b,c) + 2b_k b_l + e'_l \Upsilon_{3,1}(b) + \frac{1}{2} e'_l \Upsilon_{3,2}^{(k)}(b) - \frac{1}{2} e'_k \Upsilon_{5,2}^{(l)}(c,b) \\
&\quad + \frac{1}{2} e'_l \mu_2^\star(c) \varphi_k \circ \mu_2^\star(b) + b_k b' \Upsilon_{3,2}^{(l)}(c) + \frac{1}{2} \Upsilon_{4,2}^{(l)}(c,b) - \frac{1}{2} b' \mu_1^\star(b) e'_l \mu_2^\star(c) \\
&\quad + \frac{1}{2} e'_k \Upsilon_{3,2}^{(l)}(c) + \frac{1}{4} \Upsilon_{4,3}^{(k,l)}(b,c) - \frac{1}{4} e'_k \mu_2^\star(b) e'_l \mu_2^\star(c)
\end{aligned}$$

und damit folgt nach kurzem Zusammenfassen die Behauptung. \square

Lemma A.2.6

Es gilt für $b, c \in \mathcal{S}^{d-1}$ und $k \in \{1, \dots, d\}$

$$\begin{aligned}
\Phi_k^\star(b,c) &= \Xi_{2,3,1}^{(k)}(b,c) - 2b_k\Xi_{1,3}(b,c) - \Xi_{0,3,1}^{(k)}(b,c) - c' \Upsilon_{5,1}(b,c) + b' \mu_1^\star(b) \mu_3^\star(c) \\
&\quad - \frac{1}{2} c' \Upsilon_{5,2}^{(k)}(b,c) - \frac{1}{2} \varphi_k \circ \mu_2^\star(b) \mu_3^\star(c) - 3\Xi_{2,1,1}^{(k)}(b,c) + 6b_k(b'c) + 3c_k \\
&\quad + 3c' \Upsilon_{3,1}(b) + \frac{3}{2} c' \Upsilon_{3,2}^{(k)}(b) - \frac{3}{2} e'_k \Upsilon_{5,4}(c,b) + \frac{3}{2} \varphi_k \circ \mu_2^\star(b) c' \mu_2^\star(c) \\
&\quad + 3b_k b' \Upsilon_{3,3}(c) + \frac{3}{2} \Upsilon_{4,3}(b,c) - \frac{3}{2} b' \mu_1^\star(b) c' \mu_2^\star(c) \\
&\quad + \frac{3}{2} e'_k \Upsilon_{3,3}(c) + \frac{3}{4} \Upsilon_{4,4}^{(k)}(b,c) - \frac{3}{4} \varphi_k \circ \mu_2^\star(b) c' \mu_2^\star(c).
\end{aligned}$$

Beweis:

Um den Kovarianzkern angeben zu können, benötigen wir die abkürzenden Schreibweisen aus Kapitel 7. Damit folgt für $b, c \in \mathcal{S}^{d-1}$ und $k \in \{1, \dots, d\}$

$$\begin{aligned}
\text{Cov}(\mathcal{Z}_{1,k}(b), \mathcal{Z}_1(c)) &= E[(Z_{1,k}(b) - \varphi_k \circ \mu_2^\star(b))(Z_1(c) - \mu_3^\star(c))] \\
&= E[Z_{1,k}(b)Z_1(c)] - \varphi_k \circ \mu_2^\star(b) \mu_3^\star(c),
\end{aligned}$$

mit

$$\begin{aligned}
Z_{1,k}(b) &= f_k(X_1, b) - 2(b'X_1)b_k - b'(X_1X_1' - I_d)\mu_1^\star(b) - \varphi_k(X_1) \\
&\quad - \frac{1}{2}e'_k(X_1X_1' - I_d)\mu_2^\star(b), \\
Z_1(c) &= (c'X_1)^3 - 3(c'X_1) - \frac{3}{2}c'(X_1X_1' - I_d)\mu_2^\star(c).
\end{aligned}$$

Weiter ist

$$\begin{aligned}
E [Z_{1,k}(b)Z_1(c)] &= E \left[\left(f_k(X_1, b) - 2(b'X_1)b_k - b'(X_1X_1' - I_d)\mu_1^\star(b) - \varphi_k(X_1) \right. \right. \\
&\quad \left. \left. - \frac{1}{2}e'_k(X_1X_1' - I_d)\mu_2^\star(b) \right) \cdot \left((c'X_1)^3 - 3(c'X_1) \right. \right. \\
&\quad \left. \left. - \frac{3}{2}c'(X_1X_1' - I_d)\mu_2^\star(c) \right) \right] \\
&= E \left(f_k(X_1, b)(c'X_1)^3 \right) - 2b_k E \left((b'X_1)(c'X_1)^3 \right) - E \left(\varphi_k(X_1)(c'X_1)^3 \right) \\
&\quad - E \left(b'(X_1X_1' - I_d)\mu_1^\star(b)(c'X_1)^3 \right) + 3E \left(\varphi_k(X_1)(c'X_1) \right) \\
&\quad - \frac{1}{2}E \left(e'_k(X_1X_1' - I_d)\mu_2^\star(b)(c'X_1)^3 \right) + 6b_k E \left((b'X_1)(c'X_1) \right) \\
&\quad - 3E \left(f_k(X_1, b)(c'X_1) \right) + 3b_k E \left((b'X_1)c'(X_1X_1' - I_d)\mu_2^\star(c) \right) \\
&\quad + 3E \left(b'(X_1X_1' - I_d)\mu_1^\star(b)(c'X_1) \right) + \frac{3}{2}E \left(e'_k(X_1X_1' - I_d)\mu_2^\star(b)(c'X_1) \right) \\
&\quad - \frac{3}{2}E \left(f_k(X_1, b)c'(X_1X_1' - I_d)\mu_2^\star(c) \right) \\
&\quad + \frac{3}{2}E \left(b'(X_1X_1' - I_d)\mu_1^\star(b)c'(X_1X_1' - I_d)\mu_2^\star(c) \right) \\
&\quad + \frac{3}{2}E \left(\varphi_k(X_1)c'(X_1X_1' - I_d)\mu_2^\star(c) \right) \\
&\quad + \frac{3}{4}E \left(e'_k(X_1X_1' - I_d)\mu_2^\star(b)c'(X_1X_1' - I_d)\mu_2^\star(c) \right) \\
&= \Xi_{2,3,1}^{(k)}(b, c) - 2b_k \Xi_{1,3}(b, c) - \Xi_{0,3,1}^{(k)}(b, c) - c'\Upsilon_{5,1}(b, c) + b'\mu_1^\star(b)\mu_3^\star(c) \\
&\quad - \frac{1}{2}c'\Upsilon_{5,2}^{(k)}(b, c) + \frac{1}{2}\varphi_k \circ \mu_2^\star(b)\mu_3^\star(c) - 3\Xi_{2,1,1}^{(k)}(b, c) + 6b_k(b'c) + 3c_k \\
&\quad + 3c'\Upsilon_{3,1}(b) + \frac{3}{2}c'\Upsilon_{3,2}^{(k)}(b) - \frac{3}{2}e'_k\Upsilon_{5,4}(c, b) + \frac{3}{2}\varphi_k \circ \mu_2^\star(b)c'\mu_2^\star(c) \\
&\quad + 3b_k b'\Upsilon_{3,3}(c) + \frac{3}{2}\Upsilon_{4,3}(b, c) - \frac{3}{2}b'\mu_1^\star(b)c'\mu_2^\star(c) \\
&\quad + \frac{3}{2}e'_k\Upsilon_{3,3}(c) + \frac{3}{4}\Upsilon_{4,4}^{(k)}(b, c) - \frac{3}{4}\varphi_k \circ \mu_2^\star(b)c'\mu_2^\star(c)
\end{aligned}$$

und daraus folgt die Behauptung. \square

A.3 Beweis von Theorem 3.1.7

Die Abbildung $\Psi_B(W(b)) = W(b)'BW(b)$ ist für festes $b \in \mathcal{S}^{d-1}$ eine quadratische Form, wobei $W(b)$ ein $(d+1)$ -dimensionaler normalverteilter Zufallsvektor mit Erwartungswertvektor 0 und Kovarianzmatrix $\sigma(b, b)$ ist (siehe (3.12)). Um die Verteilung von $\Psi_B(W(b))$ für festes $b \in \mathcal{S}^{d-1}$ zu bestimmen, verwenden wir das Korollar 2.1 aus [10]:

Korollar A.3.1 (Korollar 2.1 aus [10])

Seien $Y \sim \mathcal{N}_d(0, \Sigma)$, A eine reelle symmetrische Matrix mit Rang d , $r = \text{Rang}(\Sigma A \Sigma)$ und $\lambda_1, \dots, \lambda_r$ die von Null verschiedenen Eigenwerte der Matrix $A \Sigma$. Dann gilt

$$Y'AY \stackrel{d}{=} \sum_{j=1}^r \lambda_j N_j^2,$$

wobei N_1, \dots, N_r unabhängige standard-normalverteilte Zufallsvariablen sind.

Damit die Voraussetzungen des obigen Korollars erfüllt sind, müssen die Eigenwerte der $(d+1) \times (d+1)$ -Matrix $B \Sigma_b$ berechnet werden.

Lemma A.3.2

Für die Eigenwerte der $(d+1) \times (d+1)$ -Matrix $B \Sigma_b$ gilt:

1. Es gibt genau $d-1$ von Null verschiedene reelle Eigenwerte.
2. Die von Null verschiedenen Eigenwerte $\lambda_{j,d}$, $j \in \{1, \dots, d-1\}$, sind alle gleich und hängen nur von d ab. Ferner gilt

$$\lambda_{j,d} = \frac{3}{d(d+2)(d+4)} m_6.$$

Beweis:

Um das charakteristische Polynom von $B \Sigma_b$ zu berechnen, betrachten wir zunächst für $c_1, \dots, c_d \in \mathbb{R}$ das Polynom

$$p_d(x) := \det \begin{pmatrix} 1-x & 0 & \cdots & 0 & 5c_1 \\ 0 & 1-x & \cdots & 0 & 5c_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-x & 5c_d \\ c_1(\frac{4}{5}x-1) & c_2(\frac{4}{5}x-1) & \cdots & c_d(\frac{4}{5}x-1) & -5-x \end{pmatrix} \quad (\text{A.9})$$

Für das Polynom $p_d(x)$ gilt folgendes Lemma.

Lemma A.3.3

Für $c = (c_1, \dots, c_d) \in \mathbb{R}^d$ lautet das Polynom

$$p_d(x) = (1-x)^{d-1} \left(x^2 - 5 \left(1 - \sum_{j=1}^d c_j^2 \right) + 4x \left(1 - \sum_{j=1}^d c_j^2 \right) \right). \quad (\text{A.10})$$

Insbesondere gilt für $c \in \mathcal{S}^{d-1}$

$$p_d(x) = x^2(1-x)^{d-1}.$$

Beweis:

Die zweite Aussage folgt wegen $\|c\| = 1$ durch direktes Einsetzen in (A.9). Die erste folgt durch vollständige Induktion über d . Dabei ergibt sich der Induktionsanfang $d = 2$ aus

$$\begin{aligned}
 p_2(x) &= \det \begin{pmatrix} 1-x & 0 & 5c_1 \\ 0 & 1-x & 5c_2 \\ c_1(\frac{4}{5}x-1) & c_2(\frac{4}{5}x-1) & -5-x \end{pmatrix} \\
 &= (1-x) \det \begin{pmatrix} 1-x & 5c_2 \\ c_2(\frac{4}{5}x-1) & -5-x \end{pmatrix} + 5c_1 \det \begin{pmatrix} 0 & 1-x \\ c_1(\frac{4}{5}x-1) & -5-x \end{pmatrix} \\
 &= (1-x)(x^2 + 5(c_1^2 + c_2^2 - 1) - 4x(c_1^2 + c_2^2 - 1)).
 \end{aligned}$$

Da $c_1, \dots, c_{d+1} \in \mathbb{R}$ beliebig sind und es auf die Indizierung nicht ankommt, erhält man durch Entwicklung nach der ersten Zeile und Verwendung der Induktionsvoraussetzung (A.10)

$$\begin{aligned}
 p_{d+1}(x) &= \det \begin{pmatrix} 1-x & 0 & \cdots & 0 & 5c_{d+1} \\ 0 & 1-x & \cdots & 0 & 5c_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-x & 5c_d \\ c_{d+1}(\frac{4}{5}x-1) & c_1(\frac{4}{5}x-1) & \cdots & c_d(\frac{4}{5}x-1) & -5-x \end{pmatrix} \\
 &= (1-x)p_d(x) + (-1)^{d+2}5c_{d+1} \cdots \\
 &\quad \cdot \det \begin{pmatrix} 0 & 1-x & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-x & 0 \\ 0 & 0 & \cdots & 0 & 1-x \\ c_{d+1}(\frac{4}{5}x-1) & c_1(\frac{4}{5}x-1) & \cdots & c_{d-1}(\frac{4}{5}x-1) & c_d(\frac{4}{5}x-1) \end{pmatrix} \\
 &= (1-x)p_d(x) + (-1)^{d+2}5c_{d+1}(-1)^{d+1}c_{d+1} \left(\frac{4}{5}x-1\right) \cdots \\
 &\quad \cdot \det \begin{pmatrix} 1-x & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1-x & 0 \\ 0 & 0 & \cdots & 1-x \end{pmatrix} \\
 &= (1-x)p_d(x) - 4c_{d+1}^2x(1-x)^d + 5c_{d+1}^2x(1-x)^d,
 \end{aligned}$$

und nach Zusammenfassen folgt die Behauptung. \square

Mit Hilfe des Polynoms in (A.9) lassen sich die Eigenwerte von $B\Sigma_b$ leicht bestimmen. Es gilt für das charakteristische Polynom $g(x)$ von $B\Sigma_b$

$$g(x) = \det(B\Sigma_b - xI_{d+1}) = p_d(\gamma_d^{-1}x),$$

wobei

$$\gamma_d := \frac{3}{d(d+2)(d+4)}m_6.$$

Also sind die Eigenwerte $\lambda_1, \dots, \lambda_{d-1}$ von $B\Sigma_b$ wie behauptet. \square

Die Verteilung von $\Psi_B(W(b))$ für festes $b \in \mathcal{S}^{d-1}$ folgt nun direkt aus Korollar A.3.1 mit den eben berechneten Eigenwerten.

A.4 Die Entropiebedingung in \mathcal{S}^{d-1}

In diesem Abschnitt wollen wir aufzeigen, warum die öfters benutzte Entropiebedingung in (3.2) erfüllt ist. Es ist also

$$\int_0^1 (\log N(\xi))^{\frac{1}{2}} d\xi < \infty$$

zu zeigen. Dabei ist für $\xi > 0$

$$N(\xi) := \inf\{m \in \mathbb{N} : \cup_{j=1}^m A_j = \mathcal{S}^{d-1}, A_j \subset \mathcal{S}^{d-1}, \text{diam}(A_j) \leq \xi, j = 1, \dots, m\}.$$

Die Größe $N(\xi)$ gibt also in Abhängigkeit von $\xi > 0$ die minimale Anzahl an messbaren Mengen A_j , $j = 1, \dots, m$, mit einem Durchmesser von höchstens ξ an, welche benötigt werden, um die Einheitskugel \mathcal{S}^{d-1} zu überdecken. Es handelt sich um ein uneigentliches Integral, da für kleiner werdendes ξ die Funktion $N(\xi)$ gegen ∞ strebt. Um die Existenz des Integrals zu zeigen, schätzen wir die Funktion $N(\cdot)$ geeignet nach oben ab. Hierfür betrachten wir zunächst den abgeschlossenen d -dimensionalen Würfel

$$W_d := [-1, 1]^d.$$

Offensichtlich ist die Einheitskugel \mathcal{S}^{d-1} vollständig in W_d enthalten. Wir bezeichnen mit $\partial W_d := \{x \in W_d : \|x\|_\infty = 1\}$ den Rand des Würfels W_d , wobei $\|\cdot\|_\infty$ die Maximumnorm in \mathbb{R}^d ist. Im Folgenden versuchen wir die Überdeckung der Einheitskugel mit einer Überdeckung der Oberfläche des Würfels W_d nach oben abzuschätzen. Mit vollständiger Induktion ist leicht einzusehen, dass für $d \geq 3$

$$2 \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} = |\mathcal{S}^{d-1}| \leq |\partial W_d| = d(d-1)2^{d-1}$$

gilt. Wir definieren analog zur Abbildung $N(\xi)$ eine Abbildung $M(\xi)$ durch

$$M(\xi) := \inf\{m \in \mathbb{N} : \cup_{j=1}^m A_j = \partial W_d, A_j \subset \partial W_d, \text{diam}(A_j) \leq \xi, j = 1, \dots, m\}.$$

Mit Hilfe der metrischen Projektion sieht man leicht ein, dass für $\xi > 0$

$$N(\xi) \leq M(\xi)$$

gilt. Ein d -dimensionaler Würfel W_d hat bekanntlich $d(d-1)2^{d-3}$ Flächen. Jede Fläche des Würfels W_d lässt sich für $a > 0$ durch $\lceil \frac{2}{a} \rceil^2$ Quadrate der Kantenlänge a überdecken, so dass für W_d wegen der Kantenlänge 2 und $\xi = \sqrt{2}a$ gilt

$$M(\xi) \leq d(d-1)2^{d-3} \left\lceil \frac{2}{a} \right\rceil^2 = d(d-1)2^d \left\lceil \frac{1}{\xi} \right\rceil^2.$$

Dabei bezeichnen wir mit $\lceil \cdot \rceil$ die obere Gaußklammer. Damit folgt

$$\begin{aligned}
 \int_0^1 (\log N(\xi))^{\frac{1}{2}} d\xi &\leq \int_0^1 (\log M(\xi))^{\frac{1}{2}} d\xi \\
 &\leq \int_0^1 \left(\log \left(d(d-1)2^d \frac{1}{\xi^2} \right) \right)^{\frac{1}{2}} d\xi \\
 &= \int_0^1 (-2 \log(\xi) + \alpha_d)^{\frac{1}{2}} d\xi \\
 &= 2e^{\frac{\alpha_d}{2}} \int_{\alpha_d}^{\infty} \sqrt{t} e^{-\frac{t}{2}} dt \\
 &\leq 2e^{\frac{\alpha_d}{2}} \int_0^{\infty} \sqrt{t} e^{-\frac{t}{2}} dt \\
 &= 2e^{\frac{\alpha_d}{2}} \sqrt{2\pi} < \infty
 \end{aligned}$$

mit $\alpha_d = \log(d(d-1)2^d) > 0$.

Anhang B

Schwache Konvergenz im Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$

Im Folgenden bezeichne $\|\cdot\|$ die Euklidnorm auf \mathbb{R}^d . Durch

$$d_{\mathcal{S}}(b, c) := \|b - c\|, \quad b, c \in \mathcal{S}^{d-1}, \quad (\text{B.1})$$

wird die Einheitssphäre \mathcal{S}^{d-1} zu einem metrischen Raum $(\mathcal{S}^{d-1}, d_{\mathcal{S}})$, für den in der Folge kurz \mathcal{S}^{d-1} geschrieben wird. Des Weiteren sei $C(\mathcal{S}^{d-1}, \mathbb{R})$ der Vektorraum der stetigen Funktionen $f : \mathcal{S}^{d-1} \rightarrow \mathbb{R}$. Mit der Norm

$$\|f\|_C := \sup_{b \in \mathcal{S}^{d-1}} |f(b)|, \quad f \in C(\mathcal{S}^{d-1}, \mathbb{R}), \quad (\text{B.2})$$

definieren wir die Metrik

$$d_C(f, g) := \|f - g\|_C, \quad f, g \in C(\mathcal{S}^{d-1}, \mathbb{R}).$$

Der normierte Raum $(C(\mathcal{S}^{d-1}, \mathbb{R}), \|\cdot\|_C)$ ist vollständig und demnach ein Banachraum. Der Banachraum $(C(\mathcal{S}^{d-1}, \mathbb{R}), \|\cdot\|_C)$ wird abkürzend mit $C(\mathcal{S}^{d-1}, \mathbb{R})$ bezeichnet. In der vorliegenden Arbeit sei $C(\mathcal{S}^{d-1}, \mathbb{R})$ mit der Borel- σ -Algebra \mathcal{B} versehen. Zunächst müssen wir die Begriffe kompakt, relativ kompakt und straff definieren.

Definition B.1.1

Eine Teilmenge C' von $C(\mathcal{S}^{d-1}, \mathbb{R})$ heißt kompakt, wenn jede unendliche Folge $(f_n)_{n \in \mathbb{N}}$ von Elementen in C' einen Häufungspunkt in C' hat.

Definition B.1.2

Eine Teilmenge C' von $C(\mathcal{S}^{d-1}, \mathbb{R})$ heißt relativ kompakt, wenn der Abschluss von C' kompakt ist.

Definition B.1.3

Sei A eine beliebige nichtleere Indexmenge. Eine Menge $(\mathbb{P}_t : t \in A)$ von Wahrscheinlichkeitsmaßen auf $C(\mathcal{S}^{d-1}, \mathbb{R})$ heißt *straff*, falls zu jedem $\varepsilon \in (0, 1)$ eine kompakte Menge $K \subset C(\mathcal{S}^{d-1}, \mathbb{R})$ existiert, so dass

$$\inf_{t \in A} \mathbb{P}_t(K) \geq 1 - \varepsilon$$

gilt.

Im Folgenden wird die Straffheit einer Folge $(X_n)_{n \in \mathbb{N}}$ von $C(\mathcal{S}^{d-1}, \mathbb{R})$ -wertigen Zufallselementen verwendet. In diesem Fall ist gemeint, dass die Zufallselemente X_n die Verteilung \mathbb{P}_n für $n \in \mathbb{N}$ besitzen und die Menge $(\mathbb{P}_n : n \in \mathbb{N})$ von Wahrscheinlichkeitsmaßen auf $C(\mathcal{S}^{d-1}, \mathbb{R})$ straff ist.

B.2 Verteilungskonvergenz in $C(\mathcal{S}^{d-1}, \mathbb{R})$

In diesem Abschnitt sollen Hilfsmittel zum Nachweis der Verteilungskonvergenz von stochastischen Prozessen auf $C(\mathcal{S}^{d-1}, \mathbb{R})$ zusammengestellt werden. Wir sagen, eine Familie von $C(\mathcal{S}^{d-1}, \mathbb{R})$ -wertigen Zufallselementen $(X_n)_{n \in \mathbb{N}}$ konvergiere schwach, bzw. konvergiere in Verteilung, falls es ein $C(\mathcal{S}^{d-1}, \mathbb{R})$ -wertiges Zufallselement X gibt, so dass für jede stetige und beschränkte Funktion $f : C(\mathcal{S}^{d-1}, \mathbb{R}) \rightarrow \mathbb{R}$

$$\lim_{n \rightarrow \infty} E(f(X_n)) = E(f(X))$$

erfüllt ist. In diesem Fall schreiben wir $X_n \xrightarrow{\mathcal{D}} X$. Weiter gilt für die endlichdimensionalen Verteilungen von Zufallselementen in $C(\mathcal{S}^{d-1}, \mathbb{R})$ der folgende Satz.

Theorem B.2.1

Seien X_1 und X_2 Zufallselemente in $C(\mathcal{S}^{d-1}, \mathbb{R})$. X_1 und X_2 haben genau dann die gleiche Verteilung, wenn für $k \in \mathbb{N}$ und für alle $b_1, \dots, b_k \in \mathcal{S}^{d-1}$ für die k -dimensionalen Zufallsvektoren gilt

$$\begin{pmatrix} X_1(b_1) \\ \vdots \\ X_1(b_k) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} X_2(b_1) \\ \vdots \\ X_2(b_k) \end{pmatrix}.$$

Beweis:

Für einen ausführlichen Beweis siehe z.B. [21], S. 195.

Theorem B.2.2

Seien $(X_n)_{n \in \mathbb{N}}$ eine Familie von Zufallselementen und X ein weiteres Zufallselement in $C(\mathcal{S}^{d-1}, \mathbb{R})$. Falls

1. die endlichdimensionalen Verteilungen von X_n gegen die von X konvergieren, und

2. die Familie $(X_n)_{n \in \mathbb{N}}$ straff ist,

folgt

$$X_n \xrightarrow{\mathcal{D}} X.$$

Beweis:

Für $n \in \mathbb{N}$ sei \mathbb{P}_n die Verteilung von X_n . Sei $(X_{n_k})_{k \in \mathbb{N}}$ eine beliebige Teilfolge von $(X_n)_{n \in \mathbb{N}}$. Dann existieren eine weitere Teilfolge $(X_{n_{k_j}})_{j \in \mathbb{N}}$ und ein Wahrscheinlichkeitsmaß \mathbb{P} , so dass für jede beschränkte stetige Funktion $f : C(\mathcal{S}^{d-1}, \mathbb{R}) \rightarrow \mathbb{R}$ gilt

$$\int f d\mathbb{P}_{n_{k_j}} \longrightarrow \int f d\mathbb{P}$$

für $j \rightarrow \infty$. Es folgt aus der Konvergenz endlichdimensionaler Verteilungen, dass unabhängig von der speziellen Wahl der Teilfolge $(X_{n_k})_{k \in \mathbb{N}}$ die endlichdimensionalen Grenzverteilungen immer dieselben sind. Weiter gilt nach Theorem B.2.1, dass alle Grenzverteilungen gleich sein müssen. \square

B.3 Straffheit in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1})$

Zur Überprüfung der Straffheit in $C(\mathcal{S}^{d-1}, \mathbb{R}^{d+1}) = \times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R})$ hilft folgendes Theorem.

Theorem B.3.1

Eine Familie $(\mathbb{P}_n : n \in \mathbb{N})$ von Wahrscheinlichkeitsmaßen auf dem Produktraum $\times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R})$ ist genau dann straff, wenn die Familien der marginalen Wahrscheinlichkeitsmaße straff in den jeweiligen Banachräumen $C(\mathcal{S}^{d-1}, \mathbb{R})$ sind.

Beweis:

Sei die Familie $(\mathbb{P}_n : n \in \mathbb{N})$ von Wahrscheinlichkeitsmaßen auf dem Produktraum $\times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R})$ straff. Dann existiert zu jedem $\varepsilon > 0$ eine kompakte Teilmenge $K_\varepsilon \subset \times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R})$ für die

$$\mathbb{P}_n(K_\varepsilon) > 1 - \varepsilon, \quad n \in \mathbb{N}$$

gilt. Definiere für jedes $j \in \{1, \dots, d+1\}$ durch $\pi_j : \times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R}) \rightarrow C(\mathcal{S}^{d-1}, \mathbb{R})$,

$$f = (f_1, \dots, f_{d+1}) \mapsto \pi_j(f) = f_j,$$

die Projektionsabbildungen in den j -ten Banachraum $C(\mathcal{S}^{d-1}, \mathbb{R})$ des kartesischen Produktes. Man beachte, dass die Projektionsabbildungen stetig sind. Definiere weiter für $j \in \{1, \dots, d+1\}$

$$K'_{j,\varepsilon} := \pi_j(K_\varepsilon), \tag{B.3}$$

wobei die Mengen $K'_{j,\varepsilon}$ kompakt in $C(\mathcal{S}^{d-1}, \mathbb{R})$ sind. Man beachte, dass $K_\varepsilon \subset \pi_j^{-1}(K'_{j,\varepsilon})$ gilt. Für $j \in \{1, \dots, d+1\}$ und $n \in \mathbb{N}$ gilt wegen (B.3) und der Straffheit der Familie $(\mathbb{P}_n : n \in \mathbb{N})$

$$\mathbb{P}_n(\pi_j^{-1}(K'_{j,\varepsilon})) > 1 - \varepsilon.$$

Somit ist die Familie $(\mathbb{P}_n \pi_j^{-1} : n \in \mathbb{N})$ von Wahrscheinlichkeitsmaßen straff in $C(\mathcal{S}^{d-1}, \mathbb{R})$.

Sei umgekehrt für jedes $j \in \{1, \dots, d+1\}$ die Familie $(\mathbb{P}_n \pi_j^{-1} : n \in \mathbb{N})$ der marginalen Wahrscheinlichkeitsmaße straff auf $C(\mathcal{S}^{d-1}, \mathbb{R})$. Dann existiert zu jedem $\varepsilon_j > 0$ ein kompaktes $K_{j,\varepsilon_j} \subset C(\mathcal{S}^{d-1}, \mathbb{R})$ mit

$$\mathbb{P}_n \pi_j^{-1}(K_{j,\varepsilon_j}) > 1 - \varepsilon_j, \quad n \in \mathbb{N}.$$

Weiter sei $K_\varepsilon := K_{1,\varepsilon_1} \times \dots \times K_{d+1,\varepsilon_{d+1}}$. Die Menge K_ε ist kompakt in $\times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R})$. Mit $\pi_j, j \in \{1, \dots, d+1\}$, wie oben und den Eigenschaften von Wahrscheinlichkeitsmaßen gilt

$$\begin{aligned} \mathbb{P}_n(K_\varepsilon) &= \mathbb{P}_n(\pi_1^{-1} K_{1,\varepsilon_1} \cap \dots \cap \pi_{d+1}^{-1} K_{d+1,\varepsilon_{d+1}}) \\ &\geq 1 - \sum_{j=1}^{d+1} [1 - \mathbb{P}_n \pi_j^{-1}(K_{j,\varepsilon_j})] \\ &\geq 1 - \delta_\varepsilon, \end{aligned} \tag{B.4}$$

wobei $\delta_\varepsilon = \sum_{j=1}^{d+1} \varepsilon_j > 0$ ist. Da man zu jedem $\delta_\varepsilon > 0$ ein kompaktes K_ε findet, so dass (B.4) gilt, ist die Familie $(\mathbb{P}_n : n \in \mathbb{N})$ straff auf $\times_1^{d+1} C(\mathcal{S}^{d-1}, \mathbb{R})$. \square

B.4 Der zentrale Grenzwertsatz in $C(\mathcal{S}^{d-1}, \mathbb{R})$

Die folgende Aussage ist in [2] auf S. 172 zu finden. Zunächst werden einige benötigte Bezeichnungen eingeführt. Wir schreiben $N(\xi), \xi > 0$, für die kleinste positive natürliche Zahl m , so dass \mathcal{S}^{d-1} durch genau m Teilmengen überdeckt werden kann, wobei diese Teilmengen höchstens den Durchmesser ξ haben, siehe hierzu auch Abschnitt A.4.

Theorem B.4.1

Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge zentrierter unabhängig identisch verteilter $C(\mathcal{S}^{d-1}, \mathbb{R})$ -wertiger Zufallselemente mit $E(X_n(b))^2 < \infty$ für ein $b \in \mathcal{S}^{d-1}$. Weiter existiere eine nichtnegative Zufallsvariable M mit $E(M^2) < \infty$, und es gelte die Entropiebedingung

$$\int_0^1 (\log N(\xi))^{\frac{1}{2}} d\xi < \infty.$$

Ferner gelte

$$|X_1(b) - X_1(c)| \leq M \|b - c\|.$$

Unter diesen Voraussetzungen definiert der Kovarianzkern von X_1 einen zentrierten Gaußschen Grenzprozess X in $C(\mathcal{S}^{d-1}, \mathbb{R})$ mit der Eigenschaft

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n X_j \xrightarrow{\mathcal{D}} X.$$

Anhang C

Kugeloberflächenfunktionen

In diesem Anhang werden grundlegende Definitionen und Sätze über Kugeloberflächenfunktionen zusammengefasst. Die folgenden Definitionen kann man z.B. in [14] nachlesen. Im Folgenden seien $d, n \in \mathbb{N}$. Wir betrachten Polynome P vom Grad n in d reellen Variablen.

Definition C.1.1

Ein Polynom heißt homogen, falls alle Monome, aus denen das Polynom besteht, den gleichen Grad haben.

Ein homogenes Polynom P lässt sich mit der Multi-Index Schreibweise durch

$$P : \mathbb{R}^d \rightarrow \mathbb{R}, \quad (x_1, \dots, x_d) \mapsto P(x_1, \dots, x_d) = \sum_{\|j_{(d)}\|_1 = n} a_{j_1, \dots, j_d} x_1^{j_1} x_2^{j_2} \cdots x_d^{j_d}$$

darstellen, wobei $j_{(d)} := (j_1, \dots, j_d)$ und $\|j_{(d)}\|_1 := j_1 + \cdots + j_d$ bezeichnet. Die Folge der reellen Koeffizienten a_{j_1, \dots, j_d} ist dabei eindeutig bestimmt. Weiter bezeichne mit Δ den Laplace-Operator

$$\Delta := \left(\frac{\partial}{\partial x_1} \right)^2 + \cdots + \left(\frac{\partial}{\partial x_d} \right)^2.$$

Definition C.1.2

Ein homogenes Polynom P mit der Eigenschaft $\Delta P = 0$ heißt harmonisch.

Definition C.1.3

Eine d -dimensionale harmonische Kugeloberflächenfunktion vom Grad n ist die Restriktion eines harmonischen Polynoms vom Grad n in d Variablen auf \mathcal{S}^{d-1} .

Es bezeichne $\mathcal{V}_m(d)$ den Raum der d -dimensionalen harmonischen Kugeloberflächenfunktionen vom Grad m , $m \in \mathbb{N}_0$, und $N(d, m) := \dim(\mathcal{V}_m(d))$. Schließlich sei $I_{m,d} := \{1, \dots, N(d, m)\}$ gesetzt. Die Dimension des Raumes $\mathcal{V}_m(d)$ lässt sich durch folgende Formel angeben.

Lemma C.1.4

Es gilt

$$N(d, m) = \binom{d+m-1}{m} - \binom{d+m-3}{m-2} = \frac{2m+d-2}{m+d-2} \binom{d+m-2}{d-2}.$$

Dabei wird der zweite Binomialkoeffizient gleich 0 gesetzt, falls $m-2 < 0$ ist. Für $d = 2$ und $m = 0$ wird der Bruch auf der rechten Seite der Gleichung als 1 definiert.

Beweis:

Siehe [14], Seite 65.

Der Raum $\mathcal{V}_m(d)$ ist ein Teilraum von $C(\mathcal{S}^{d-1}, \mathbb{R})$, dem Raum der stetigen Funktionen $\mathcal{S}^{d-1} \rightarrow \mathbb{R}$. Definiere ein Skalarprodukt auf $C(\mathcal{S}^{d-1}, \mathbb{R})$ für $f, g \in C(\mathcal{S}^{d-1}, \mathbb{R})$ durch

$$\langle f, g \rangle := \int_{\mathcal{S}^{d-1}} f(\xi)g(\xi) d\mathcal{S}^{d-1}(\xi).$$

Mit der durch das Skalarprodukt induzierten Norm

$$\|f\| := (\langle f, f \rangle)^{\frac{1}{2}},$$

ist $C(\mathcal{S}^{d-1}, \mathbb{R})$ ein Prä-Hilbertraum.

C.2 Legendre-Polynome

In diesem Abschnitt werden mehrdimensionale Analoga zu den klassischen Legendre-Polynomen eingeführt. Details hierzu sind in [26] und [14] zu finden.

Theorem C.2.1

Sei $m \geq 0$ beliebig aber fest. Dann existiert genau ein Polynom $P_m(d; \cdot)$ vom Grad m mit der folgenden Eigenschaft: Ist $\varphi_1, \dots, \varphi_{N(d,m)}$ eine beliebige Orthonormalbasis von $\mathcal{V}_m(d)$, dann gilt

$$\sum_{j=1}^{N(d,m)} \varphi_j(u)\varphi_j(v) = \frac{N(d,m)}{|\mathcal{S}^{d-1}|} P_m(d; u'v).$$

Für festes $v \in \mathcal{S}^{d-1}$ bildet die Funktion $P_m(d; v' \cdot) : \mathcal{S}^{d-1} \rightarrow \mathbb{R}$, $u \mapsto P_m(d; v'u)$, eine d -dimensionale harmonische Kugeloberflächenfunktion vom Grad m .

Beweis:

siehe [14], Seite 80.

Das in Theorem C.2.1 eingeführte Polynom $P_m(d; \cdot)$ heißt d -dimensionales Legendre-Polynom vom Grad m . In [26] auf Seite 16 wird eine explizite Darstellung der mehrdimensionalen Legendre-Polynome angegeben. Es gilt

$$P_m(d; t) = m! \Gamma \left(\frac{d-1}{2} \right) \sum_{l=0}^{\lfloor \frac{m}{2} \rfloor} \left(-\frac{1}{4} \right)^l \frac{(1-t^2)^l t^{m-2l}}{l!(m-2l)! \Gamma \left(l + \frac{d-1}{2} \right)}, \quad (\text{C.1})$$

wobei $\lfloor \cdot \rfloor$ die untere Gaußklammer bezeichnet. Mit (C.1) lassen sich die ersten sieben mehrdimensionalen Legendrepolynome leicht direkt angeben.

$$\begin{aligned} P_0(d; t) &= 1, \\ P_1(d; t) &= t, \\ P_2(d; t) &= \frac{d}{d-1} t^2 - \frac{1}{d-1}, \\ P_3(d; t) &= \frac{d+2}{d-1} t^3 - \frac{3}{d-1} t, \\ P_4(d; t) &= \frac{(d+2)(d+4)}{(d+1)(d-1)} t^4 - \frac{6(d+2)}{(d+1)(d-1)} t^2 + \frac{3}{(d+1)(d-1)}, \\ P_5(d; t) &= \frac{(d+4)(d+6)}{(d+1)(d-1)} t^5 - \frac{10(d+4)}{(d+1)(d-1)} t^3 + \frac{15}{(d+1)(d-1)} t, \\ P_6(d; t) &= \frac{(d+4)(d+6)(d+8)}{(d-1)(d+1)(d+3)} t^6 - \frac{15(d+4)(d+6)}{(d-1)(d+1)(d+3)} t^4 \\ &\quad + \frac{45(d+4)}{(d-1)(d+1)(d+3)} t^2 - \frac{15}{(d-1)(d+1)(d+3)}. \end{aligned}$$

Im Umkehrschluss lässt sich jede k -Potenz einer reellen Zahl t in der Form

$$t^k = \sum_{j=1}^k c_{j,d} P_j(d; t),$$

schreiben, wobei die reellen Koeffizienten $c_{j,d}$ nur von j und d abhängen. Speziell gilt

$$\begin{aligned} t^0 &= P_0(d; t), \\ t^1 &= P_1(d; t), \\ t^2 &= \frac{d-1}{d} P_2(d; t) + \frac{1}{d} P_0(d; t), \\ t^3 &= \frac{d-1}{d+2} P_3(d; t) + \frac{3}{d+2} P_1(d; t), \\ t^4 &= \frac{d^2-1}{(d+2)(d+4)} P_4(d; t) + \frac{6(d-1)}{d(d+4)} P_2(d; t) + \frac{3}{d(d+2)}, \\ t^5 &= \frac{(d+1)(d-1)}{(d+4)(d+6)} P_5(d; t) + \frac{10(d-1)}{(d+2)(d+6)} P_3(d; t) + \frac{15}{(d+2)(d+4)} P_1(d; t), \\ t^6 &= \frac{(d-1)(d+1)(d+3)}{(d+4)(d+6)(d+8)} P_6(d; t) + \frac{15(d-1)(d+1)}{(d+2)(d+4)(d+8)} P_4(d; t) \\ &\quad + \frac{45(d-1)}{d(d+4)(d+6)} P_2(d; t) + \frac{15}{d(d+2)(d+4)}. \end{aligned}$$

C.3 Das Funk-Hecke Theorem

Das folgende Theorem ermöglicht die Auswertung bestimmter sphärischer Oberflächenintegrale.

Theorem C.3.1 (*FUNK-HECKE Theorem*)

Seien $a \in \mathcal{S}^{d-1}$, $\varphi_m \in \mathcal{V}_m(d)$ und $g \in C([-1, 1])$. Dann gilt

$$\int_{\mathcal{S}^{d-1}} g(a'\eta)\varphi_m(\eta)d\mathcal{S}^{d-1}(\eta) = \lambda_m\varphi_m(a),$$

wobei

$$\lambda_m := |\mathcal{S}^{d-2}| \int_{-1}^1 P_m(d;t)g(t)(1-t^2)^{\frac{d-3}{2}} dt.$$

Hierbei bezeichnet $P_m(d;t)$, $t \in \mathbb{R}$, das d -dimensionale Legendre-Polynom vom Grad m .

Beweis:

siehe z.B. [26], Seite 30.

C.4 Eine Methode zur Berechnung von $E((b'U)^r(c'U)^s)$

In diesem Abschnitt wird für $r, s \in \mathbb{N}$, beliebige $b, c \in \mathcal{S}^{d-1}$ und $U \sim \mathcal{U}(\mathcal{S}^{d-1})$ eine allgemeine Methode zur Berechnung des Erwartungswertes $E((b'U)^r(c'U)^s)$ vorgestellt. Es bezeichne für $m \in \mathbb{N}$ mit $P_m(d;t)$ das d -dimensionale Legendre Polynom vom Grad m . Bekanntlich gilt für beliebige $s \in \mathbb{N}$

$$t^s = \sum_{j=0}^s c_{j,d} P_j(d;t), \quad (t \in \mathbb{R}), \quad (\text{C.2})$$

wobei die reellen Koeffizienten $c_{j,d}$ eindeutig bestimmt sind und nur von j und d abhängen. Im Folgenden beachte man, dass für festes $a \in \mathcal{S}^{d-1}$ die Abbildung $\mathcal{S}^{d-1} \rightarrow \mathbb{R}$, $x \mapsto P_m(d;a'x)$ eine harmonische Kugeloberflächenfunktion vom Grad m ist. Mit (C.2) und dem Funk-Hecke Theorem gilt

$$\begin{aligned} E((b'U)^r(c'U)^s) &= \frac{1}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} (b'u)^r (c'u)^s d\mathcal{S}^{d-1}(u) \\ &= \frac{1}{|\mathcal{S}^{d-1}|} \int_{\mathcal{S}^{d-1}} (b'u)^r \sum_{j=0}^s c_{j,d} P_j(d;c'u) d\mathcal{S}^{d-1}(u) \\ &= \frac{1}{|\mathcal{S}^{d-1}|} \sum_{j=0}^s \int_{\mathcal{S}^{d-1}} (b'u)^r P_j(d;c'u) d\mathcal{S}^{d-1}(u) \\ &= \frac{1}{|\mathcal{S}^{d-1}|} \sum_{j=0}^s c_{j,d} \lambda_{j,d,r} P_j(d;c'b), \end{aligned}$$

mit

$$\lambda_{j,d,r} := |\mathcal{S}^{d-2}| \int_{-1}^1 P_j(d;t) t^r (1-t^2)^{\frac{d-3}{2}} dt.$$

Bemerkung C.4.1

Wie die obige Herleitung zeigt, sind die gemischten Momente selbst bei verschiedenen Exponenten $r, s \in \mathbb{N}$ symmetrisch in $b, c \in \mathcal{S}^{d-1}$, da diese nur als Skalarprodukt im Argument der Legendre-Polynome auftreten und die reellwertigen Konstanten $c_{j,d}$ und $\lambda_{j,d,r}$, $j = 1, \dots, s$, nur von $d, r, s \in \mathbb{N}$ abhängen.

Beispiel C.4.2

Sei speziell $r = s = 4$. Um den gesuchten Erwartungswert explizit angeben zu können, müssen die Koeffizienten $\lambda_{j,d,4}$ und $c_{j,d}$ für $j \in \{0, \dots, 4\}$ angegeben werden. Man beachte, dass auf Grund von Symmetrieüberlegungen für ungerade j die Koeffizienten $\lambda_{j,d,4}$ verschwinden. Es gilt

$$\begin{aligned} \lambda_{0,d,4} &= \frac{3}{d(d+2)} |\mathcal{S}^{d-1}|, \\ \lambda_{2,d,4} &= \frac{12}{d(d+2)(d+4)} |\mathcal{S}^{d-1}|, \\ \lambda_{4,d,4} &= \frac{24}{d(d+2)(d+4)(d+6)} |\mathcal{S}^{d-1}|. \end{aligned}$$

Weiter gilt für $j \in \{0, 2, 4\}$

$$\begin{aligned} c_{0,d} &= \frac{3}{d(d+2)}, \\ c_{2,d} &= \frac{6(d-1)}{d(d+4)}, \\ c_{4,d} &= \frac{(d-1)(d+1)}{(d+2)(d+4)}. \end{aligned}$$

Eingesetzt ergibt dies

$$\begin{aligned} E(((b'U)^4(c'U)^4)) &= \left(\frac{3}{d(d+2)} \right)^2 P_0(d; c'b) + \frac{72(d-1)}{d^2(d+2)(d+4)^2} P_2(d; c'b) \\ &\quad + \frac{24(d-1)(d+1)}{d(d+2)^2(d+4)^2(d+6)} P_4(d; c'b) \\ &= \frac{24}{d(d+2)(d+4)(d+6)} (b'c)^4 + \frac{72}{d(d+2)(d+4)(d+6)} (b'c)^2 \\ &\quad + \frac{9}{d(d+2)(d+4)(d+6)}. \end{aligned}$$

Beispiel C.4.3

Sei speziell $r = 6, s = 4$. In diesem Fall lauten die Koeffizienten mit geradem Index j

$$\begin{aligned}\lambda_{0,d,6} &= \frac{15}{d(d+2)(d+4)} |\mathcal{S}^{d-1}|, \\ \lambda_{2,d,6} &= \frac{90}{d(d+2)(d+4)(d+6)} |\mathcal{S}^{d-1}|, \\ \lambda_{4,d,6} &= \frac{360}{d(d+2)(d+4)(d+6)(d+8)} |\mathcal{S}^{d-1}|.\end{aligned}$$

Auch in diesem Fall sind alle Koeffizienten $\lambda_{j,d,6}$ mit ungeradem j gleich 0. Da die Koeffizienten $c_{j,d}$ nicht von r abhängen, sind sie mit den Koeffizienten aus Beispiel C.4.2 identisch. Einsetzen ergibt hier

$$E((b'U)^6(c'U)^4) = \frac{1}{d(d+2)(d+4)(d+6)(d+8)} (360(b'c)^4 + 540(b'c)^2 + 45).$$

Beispiel C.4.4

Sei speziell $r = 6, s = 6$. In diesem Fall benötigen wir nur noch den Koeffizienten mit Index 6, da die restlichen auftretenden Koeffizienten mit denen aus Beispiel C.4.3 übereinstimmen. Es gilt

$$\lambda_{6,d,6} = \frac{720}{d(d+2)(d+4)(d+6)(d+8)(d+10)} |\mathcal{S}^{d-1}|.$$

Auch in diesem Fall sind alle Koeffizienten $\lambda_{j,d,6}$ mit ungeradem j aus Symmetriegründen gleich 0. Die Koeffizienten $c_{j,d}$ mit $j \in \{0, 2, 4, 6\}$ lauten

$$\begin{aligned}c_{0,d} &= \frac{15}{d(d+2)(d+4)}, \\ c_{2,d} &= \frac{45(d-1)}{d(d+4)(d+6)}, \\ c_{4,d} &= \frac{15(d-1)(d+1)}{(d+2)(d+4)(d+8)}, \\ c_{6,d} &= \frac{(d-1)(d+1)(d+3)}{(d+4)(d+6)(d+8)}.\end{aligned}$$

Zusammengesetzt ergibt dies

$$E(((b'U)^6(c'U)^6)) = \frac{1}{d(d+2)(d+4)(d+6)(d+8)(d+10)} (720(b'c)^6 + 5400(b'c)^4 + 4050(b'c)^2 + 225)$$

Bemerkung C.4.5

Zur Berechnung der Koeffizienten $\lambda_{j,d,r}$ wurde das Computeralgebra System Maple 11 verwendet.

Anhang D

Aussagen der asymptotischen Statistik

D.1 Benachbarte Alternativen

Die hier aufgeführten Definitionen und Sätze zur Benachbarkeit von Verteilungsfolgen können in [29] Kapitel 6.3 nachgelesen werden.

Sei $(\mathfrak{X}, \mathfrak{B})$ ein Messraum und $(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)}) := (\times_{j=1}^n \mathfrak{X}_j, \times_{j=1}^n \mathfrak{B}_j)$ das n -fache kartesische Produkt von Messräumen $(\mathfrak{X}_j, \mathfrak{B}_j)$, $j = 1, \dots, n$. Weiter sei $\mathfrak{M}(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)})$ der Raum aller Wahrscheinlichkeitsmaße auf $(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)})$.

Der Totalvariationsabstand zweier Wahrscheinlichkeitsmaße $u, v \in \mathfrak{M}(\mathfrak{X}, \mathfrak{B})$ ist definiert durch

$$\|u - v\|_{TV} := \sup_{B \in \mathfrak{B}} |u(B) - v(B)|.$$

Der Hellinger-Abstand zweier Wahrscheinlichkeitsmaße u und v , welche von einem Maß ν dominiert werden, also $u \ll \nu$ und $v \ll \nu$, ist gegeben durch

$$H(u, v) := \left[\frac{1}{2} \int \left(\sqrt{\frac{du}{d\nu}} - \sqrt{\frac{dv}{d\nu}} \right)^2 d\nu \right]^{\frac{1}{2}}.$$

Definition D.1.1

Für $n \in \mathbb{N}$ seien $u_n, v_n \in \mathfrak{M}(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)})$ beliebige Wahrscheinlichkeitsmaße. Dann heißt:

- a) (v_n) zu (u_n) benachbart, kurz: $(v_n) \triangleleft (u_n)$, genau dann, wenn gilt:

$$\forall (B_n) \in (\mathfrak{B}_{(n)}) : u_n(B_n) \rightarrow 0 \Rightarrow v_n(B_n) \rightarrow 0;$$

speziell heißen (v_n) und (u_n) wechselseitig benachbart, kurz: $(v_n) \triangleleft \triangleright (u_n)$, genau dann, wenn gilt:

$$(v_n) \triangleleft (u_n) \quad \text{und} \quad (u_n) \triangleleft (v_n);$$

- b) (v_n) zu (u_n) nicht trennend, wenn ein $0 < c < 1$ existiert mit

$$\forall (B_n) \in (\mathfrak{B}_{(n)}) : u_n(B_n) \rightarrow 0 \Rightarrow \limsup_{n \rightarrow \infty} v_n(B_n) \leq c;$$

c) (v_n) zu (u_n) vollständig trennend, kurz $(v_n)\Delta(u_n)$, genau dann, wenn gilt:

$$\exists(B_n) \in (\mathfrak{B}_{(n)}) : u_n(B_n) \rightarrow 0 \text{ und } v_n(B_n) \rightarrow 1.$$

Theorem D.1.2

Für $n \in \mathbb{N}$ seien $u_n, v_n \in \mathfrak{M}(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)})$. Dann gilt:

- a) $\lim_{n \rightarrow \infty} \|u_n - v_n\|_{TV} = 0 \Rightarrow (v_n) \triangleleft (u_n)$.
- b) $\limsup_{n \rightarrow \infty} \|u_n - v_n\|_{TV} < 1 \Leftrightarrow (u_n) \text{ und } (v_n) \text{ nicht trennend.}$
- c) $\lim_{n \rightarrow \infty} \|u_n - v_n\|_{TV} = 1 \Rightarrow (u_n)\Delta(v_n)$.

Diese Aussagen bleiben richtig, wenn man die Totalvariation $\|u_n - v_n\|_{TV}$ durch den Hellinger-Abstand $H(u_n, v_n)$ ersetzt.

Theorem D.1.3 (1. Le Cam Lemma)

Für $n \in \mathbb{N}$ seien $u_n, v_n \in \mathfrak{M}(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)})$ und $L_n = \frac{dv_n}{du_n}$. Dann ist

$$\log L_n \xrightarrow{\mathcal{D}_{u_n}} \mathcal{N}\left(-\frac{\kappa^2}{2}, \kappa^2\right)$$

mit $\kappa^2 \in [0, \infty)$ hinreichend für die wechselseitige Benachbarkeit $(v_n) \triangleleft (u_n)$ und die Gültigkeit von

$$\log L_n \xrightarrow{\mathcal{D}_{v_n}} \mathcal{N}\left(\frac{\kappa^2}{2}, \kappa^2\right).$$

Im Folgenden sei $l \geq 1$.

Theorem D.1.4 (Le Cam)

Für $n, l \in \mathbb{N}$ seien $u_n, v_n \in \mathfrak{M}(\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)})$ beliebige Wahrscheinlichkeitsmaße, $L_n := \frac{dv_n}{du_n}$ ein Dichtequotient von v_n bzgl. u_n und $T_n : (\mathfrak{X}_{(n)}, \mathfrak{B}_{(n)}) \rightarrow (\mathbb{R}^l, \mathcal{B}^l)$ eine Statistik. Weiter seien $\tilde{F}, \hat{F} \in \mathfrak{M}(\mathbb{R}^{l+1}, \mathcal{B}^{l+1})$, $(a, c, \kappa^2) \in \mathbb{R}^l \times \mathbb{R}^l \times [0, \infty)$ und $\mathcal{S} \in \mathbb{R}^{l \times l}$ eine positiv semidefinite Matrix. Dann gilt:

a)

$$\begin{aligned} & \begin{pmatrix} T_n \\ L_n \end{pmatrix} \xrightarrow{\mathcal{D}_{u_n}} \tilde{F} \quad \text{und} \quad \int_{\mathbb{R}^{l+1}} y \, d\tilde{F}(t, y) = 1 \\ \Rightarrow & \begin{pmatrix} T_n \\ L_n \end{pmatrix} \xrightarrow{\mathcal{D}_{v_n}} \tilde{G}, \quad \tilde{G}(C) := \int_C y \, d\tilde{F}(t, y), \quad C \in \mathcal{B}^{l+1}; \end{aligned}$$

b)

$$\begin{aligned} & \begin{pmatrix} T_n \\ \log L_n \end{pmatrix} \xrightarrow{\mathcal{D}_{un}} \hat{F} \quad \text{und} \quad \int_{\mathbb{R}^{l+1}} e^z d\hat{F}(t, z) = 1 \\ \Rightarrow & \begin{pmatrix} T_n \\ \log L_n \end{pmatrix} \xrightarrow{\mathcal{D}_{vn}} \hat{G}, \quad \hat{G}(C) := \int_C e^z d\hat{F}(t, z), \quad C \in \mathcal{B}^{l+1}; \end{aligned}$$

c)

$$\begin{aligned} & \begin{pmatrix} T_n \\ \log L_n \end{pmatrix} \xrightarrow{\mathcal{D}_{un}} \mathcal{N}_{l+1} \left(\begin{pmatrix} a \\ -\frac{\kappa^2}{2} \end{pmatrix}, \begin{pmatrix} \mathcal{S} & c \\ c' & \kappa^2 \end{pmatrix} \right) \\ \Rightarrow & \begin{pmatrix} T_n \\ \log L_n \end{pmatrix} \xrightarrow{\mathcal{D}_{vn}} \mathcal{N}_{l+1} \left(\begin{pmatrix} a + c \\ \frac{\kappa^2}{2} \end{pmatrix}, \begin{pmatrix} \mathcal{S} & c \\ c' & \kappa^2 \end{pmatrix} \right). \end{aligned} \quad (\text{D.1})$$

Korollar D.1.5 (3. Le Cam Lemma)

Existieren Vektoren $a, c \in \mathbb{R}^l$ sowie eine Zahl $\kappa^2 \in [0, \infty)$ und eine positiv semidefinite Matrix $\mathcal{S} \in \mathbb{R}^{l \times l}$ derart, dass für die gemeinsame Verteilung von T_n und $\log L_n$ gemäß (D.1) asymptotisch normal ist, so gilt

$$T_n \xrightarrow{\mathcal{D}_{vn}} \mathcal{N}_l(a + c, \mathcal{S}).$$

Literaturverzeichnis

- [1] J. A. V. Alva and E. G. Estrada. A generalization of shapiro-wilk's test for multivariate normality. *Commun. Stat., Theory Methods*, 38(11):1870–1883, 2009.
- [2] A. Araujo and E. Giné. *The Central Limit Theorem for Real and Banach Valued Random Variables*. Wiley, 1980.
- [3] J. A. Baker. Integration over spheres and the divergence theorem for balls. *The American Mathematical Monthly*, 104(1):33–47, 1997.
- [4] L. Baringhaus and N. Henze. Limit distributions for measures of multivariate skewness and kurtosis based on projections. *Journal of Multivariate Analysis*, 38(1):51–69, 1991.
- [5] L. Baringhaus and N. Henze. Limit distributions for mardia's measure of multivariate skewness. *Annals of Statistics*, 20(4):1889–1902, 1992.
- [6] P. Billingsley. *Convergence of Probability Measures*. Wiley, 1st edition, 1968.
- [7] D.R. Cox and N.J.H. Small. Testing multivariate normality. *Biometrika*, 65:263–272, 1978.
- [8] L. Devroye. Laws of the iterated logarithm for order statistics of uniform spacings. *The Annals of Probability*, 9:860–867, 1981.
- [9] L. Devroye. *Non-Uniform Random Variate Generation*. Springer, 1986.
- [10] J. J. Dik and M. C. M. de Gunst. The distribution of general quadratic forms in normal variables. *Statistica Neerlandica*, 39:14–26, 1985.
- [11] M. Eaton and M. Perlman. The non-singularity of generalized sample covariance matrices. *Annals of Statistics*, 1(4):710–717, 1973.
- [12] K.T. Fang and Y. Wang. *Number theoretic methods in statistics*. Chapman & Hall, 1994.
- [13] S. Fang, K. Kotz and K. Ng. *Symmetric Multivariate and Related Distributions*. Chapman and Hall, 1990.
- [14] H. Groemer. *Geometric Applications of Fourier Series and Spherical Harmonics*. Cambridge University Press, 1996.

- [15] N. Gürtler. *Asymptotische Untersuchungen zur Klasse der BHEP-Tests auf multivariate Normalverteilung mit festem und variablen Glättungsparameter*. PhD thesis, Universität Karlsruhe (TH).
- [16] N. Henze. Invariant tests for multivariate normality: a critical review. *Statistical Papers*, 43:467–506, 2002.
- [17] N. Henze and T. Wagner. A class of invariant consistent tests for multivariate normality. *Journal of Multivariate Analysis*, 61(1):1–23, 1997.
- [18] N. Henze and B. Zirkler. A class of invariant and consistent tests for multivariate normality. *Communications in Statistics - Theory and Methods*, 19(10):3595–3617, 1990.
- [19] S. Janson. Maximal spacings in several dimensions. *The Annals of Probability*, 15(1):274–280, 1987.
- [20] M. E. Johnson. *Multivariate Statistical Simulation*. Wiley, 1987.
- [21] D. Khoshnevisan. *Multiparameter Processes*. Springer, 2002.
- [22] P. Lévy. Sur la division d’un segment par des points choisis au hasard. *Comptes Rendus de l’Académie des sciences*, 208:147–149, 1939.
- [23] J.F. Malkovich and A.A. Afifi. On tests of multivariate normality. *Journal of the American Statistical Association*, 68(341):176–179, 1973.
- [24] K. V. Mardia. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3):519–530, 1970.
- [25] K. V. Mardia. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya Series B*, 36(2):115–128, 1974.
- [26] C. Müller. *Analysis of Spherical Symmetries in Euclidean Spaces*. Springer, 1998.
- [27] J. Pudielko. On a new affine invariant and consistent test for multivariate normality. *Probab. Math. Stat.*, 25(1):43–54, 2005.
- [28] G. J. Székely and M. L. Rizzo. A new test for multivariate normality. *Journal of Multivariate Analysis*, 93:58–80, 2005.
- [29] H. Witting and U. Müller-Funk. *Mathematische Statistik II*. Teubner, 1995.

ISBN 978-3-86644-549-9

