Universität
Augsburg
University

Asymptotische Struktureigenschaften des iterativen proportionalen Anpassungsverfahrens

Dissertation zur Erlangung des
akademischen Grades Dr. rer. nat.

eingereicht an der Mathematisch-
Naturwissenschaftlich-Technischen
Fakultät der Universität Augsburg

von
Christoph Gietl

Augsburg, Dezember 2014

Erstgutachter: Prof. Dr. Friedrich Pukelsheim
Zweitgutachter: Prof. Dr. Lothar Heinrich
Mündliche Prüfung: Freitag, 30. Januar 2015

Abstract

This thesis analyses the asymptotic behaviour of two variants of the iterative proportional fitting procedure (IPF procedure), the two-dimensional IPF procedure and the multidimensional IPF procedure.

The two-dimensional IPF procedure takes as input a non-negative matrix and positive row and column marginals. It then generates a sequence of matrices, called the IPF sequence, by alternately scaling rows and columns. Three approaches to an analysis of convergence are discussed. The first approach is based on information divergence, the second on an inverted geometric mean and the third on an L_1-error function. By using the information divergence approach, necessary and sufficient conditions for convergence of the IPF sequence are shown. For the case of convergence, continuous dependence of the limit of the IPF sequence on the input matrix is proven. Under certain restrictions, continuous dependence of the limit on the marginals is also shown. For the case of divergence, convergence of the IPF subsequence along the row scaling steps and convergence of the IPF subsequence along the column scaling steps are proven. The two limit points of the IPF sequence are characterised using information divergence. Under certain restrictions, continuous dependence of the limit points on the input matrix and the marginals is shown. For both cases, the connectedness structure of the input matrix and the limit points of the IPF sequence is analysed. Original results of this thesis regarding the two-dimensional IPF procedure include convergence of the IPF subsequences as well as the characterisation of the limit points and the analysis of their connectedness structure in the case of divergence. The continuity statements for both cases are original as well.

The multidimensional IPF procedure takes as input a non-negative table and several positive marginal tables. It then generates a sequence of tables, called the IPF sequence, by cyclical scaling. By using the information divergence approach, necessary and sufficient conditions for convergence of the IPF sequence are shown. For the case of convergence, continuous dependence of the limit of the IPF sequence on the input table is proven. For the case of divergence, numerical simulations suggest convergence of each IPF subsequence along which the same marginal table is fitted. This conjecture is proven for two special cases. For the general case, non-existence of a universal variational characterisation of the limits of these IPF subsequences is shown. Original results of this thesis regarding the multidimensional IPF procedure include convergence of the IPF subsequences in one of the special cases and the non-existence of a universal variational characterisation of the limits of these IPF subsequences in the general case of divergence. The continuity statement for the case of convergence is original as well.

Zusammenfassung

Diese Dissertation analysiert das asymptotische Verhalten von zwei Varianten des iterativen proportionalen Anpassungsverfahrens (IPF-Verfahrens, iterative proportional fitting procedure), dem zweidimensionalen IPF-Verfahren und dem mehrdimensionalen IPF-Verfahren.

Das zweidimensionale IPF-Verfahren verlangt als Eingabe eine nichtnegative Ausgangsmatrix sowie positive Zeilen- und Spaltenmarginalien. Davon ausgehend wird durch abwechselnde Skalierung der Zeilen und der Spalten eine Matrizenfolge errechnet, die sogenannte IPF-Folge. Drei Ansätze zur Konvergenzanalyse werden diskutiert. Der erste Ansatz basiert auf der Informationsdivergenz, der zweite auf einem invertierten geometrischen Mittel und der dritte auf einem L_1-Fehlerfunktional. Mithilfe des informationstheoretischen Ansatzes werden notwendige und hinreichende Bedingungen für die Konvergenz der IPF-Folge gezeigt. Für den Konvergenzfall wird die stetige Abhängigkeit des Grenzwerts der IPF-Folge von der Ausgangsmatrix bewiesen. Unter gewissen Einschränkungen wird auch die stetige Abhängigkeit des Grenzwerts von den Marginalien gezeigt. Für den Divergenzfall werden die Konvergenz der IPF-Teilfolge entlang der Zeilenskalierungsschritte und die Konvergenz der IPF-Teilfolge entlang der Spaltenskalierungsschritte bewiesen. Die beiden Häufungspunkte der IPF-Folge werden informationstheoretisch charakterisiert. Unter gewissen Einschränkungen wird die stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix und den Marginalien gezeigt. Für beide Fälle wird die Zusammenhangsstruktur der Ausgangsmatrix und der Häufungspunkte der IPF-Folge untersucht. Neue Resultate dieser Dissertation zum zweidimensionalen IPF-Verfahren sind die Konvergenz der IPF-Teilfolgen sowie die informationstheoretische Charakterisierung der Häufungspunkte und die Analyse ihrer Zusammenhangsstruktur im Divergenzfall. Ebenfalls neu sind die Stetigkeitsaussagen in beiden Fällen.

Das mehrdimensionale IPF-Verfahren verlangt als Eingabe eine nichtnegative Ausgangstafel und mehrere positive Marginaltafeln. Davon ausgehend wird durch zyklische Skalierung eine Tafelfolge errechnet, die sogenannte IPF-Folge. Mithilfe des informationstheoretischen Ansatzes werden notwendige und hinreichende Bedingungen für die Konvergenz der IPF-Folge gezeigt. Für den Konvergenzfall wird die stetige Abhängigkeit des Grenzwerts der IPF-Folge von der Ausgangstafel bewiesen. Für den Divergenzfall legen numerische Simulationen die Konvergenz jeder IPF-Teilfolge nahe, entlang welcher an dieselbe Marginaltafel angepasst wird. Diese Hypothese wird in zwei Spezialfällen bewiesen. Für den allgemeinen Fall wird die Nichtexistenz einer universellen variationellen Charakterisierung der Grenzwerte dieser IPF-Teilfolgen gezeigt. Neue Resultate dieser Dissertation zum mehrdimensionalen IPF-Verfahren sind die Konvergenz der IPF-Teilfolgen in einem der beiden Spezialfälle sowie die Nichtexistenz einer universellen variationellen Charakterisierung der Grenzwerte dieser IPF-Teilfolgen im allgemeinen Divergenzfall. Ebenfalls neu ist die Stetigkeitsaussage im Konvergenzfall.

Danksagung

Hiermit möchte ich mich ganz herzlich bei allen bedanken, die mich bei der Anfertigung dieser Dissertation unterstützt haben.

Mein Dank gilt dabei insbesondere Friedrich Pukelsheim, der mich auf das Thema des iterativen proportionalen Anpassungsverfahrens aufmerksam gemacht hat und der dieses Dissertationsprojekt betreut hat. Die motivierende Atmosphäre in unseren Besprechungen, die zahlreichen Anregungen und Ratschläge und die stets konstruktive Kritik waren für mich ein großer Ansporn. Des Weiteren möchte ich mich bei Fabian Reffel bedanken für die sehr gute Zusammenarbeit der letzten Jahre. Unsere gemeinsame Forschung zum zweidimensionalen iterativen proportionalen Anpassungsverfahren wurde zu einem Grundpfeiler dieser Dissertation. Neben den langen Diskussionen vor der Bürotafel werden mir vor allem unsere gemeinsamen Konferenzreisen in Erinnerung bleiben. Mein Dank gilt auch den anderen Kollegen am Institut für Mathematik und am Institut für Informatik der Universität Augsburg, insbesondere Christian Bräu, Andreas Käufl, Jan Natolski, Kai-Friederike Oelbermann und Patrick Roocks, die stets ein offenes Ohr für meine fachlichen Fragen hatten und die durch ihr konstruktives Feedback zur Entstehung dieser Dissertation beigetragen haben.

Außerdem möchte ich mich bei den Forschern am Ústav teorie informace a automatizace in Prag bedanken, deren Gastfreundschaft ich während eines zweimonatigen Forschungsaufenthalts sowie weiterer Kurzbesuche in Prag kennen und schätzen lernen durfte. Mein besonderer Dank gilt dabei meinem Gastgeber, František Matúš. Unsere Gespräche waren mir eine große Hilfe dabei, dieses Dissertationsprojekt im Kontext der aktuellen Forschung zu verorten.

Darüber hinaus bedanke ich mich bei der Studienstiftung des deutschen Volkes, die dieses Dissertationsprojekt, den Forschungsaufenthalt in Prag und einige Konferenzreisen zum großen Teil finanziert hat. Auch die ideelle Förderung der Studienstiftung war eine große Bereicherung.

Ganz besonders bedanken möchte ich mich bei meiner Familie und bei meinen Freunden für die Unterstützung, die ich während meines Promotionsstudiums durch sie erfahren habe. Zu guter Letzt möchte ich mich bei Sissi bedanken, die stets ein offenes Ohr für meine Sorgen und Probleme hatte und die seit über drei Jahren mein Leben bereichert.

Inhaltsverzeichnis

Einleitung

Das iterative proportionale Anpassungsverfahren (IPF-Verfahren, iterative proportional fitting procedure) ist ein Algorithmus zur Anpassung von Matrizen oder höherdimensionalen Tafeln an vorgegebene Marginalien. Er wurde von Kruithof (1937, S. E20 ff.) eingeführt und von Deming und Stephan (1940) bekannt gemacht. Ausgehend von einer vorgegebenen Ausgangstafel errechnet das IPF-Verfahren durch zyklische Skalierung eine Tafelfolge, die sogenannte IPF-Folge. Die zentrale Anwendung des IPF-Verfahrens ist die Berechnung der Maximum-Likelihood-Schätzer in loglinearen Kontingenztafelmodellen in der Statistik (vgl. Agresti 1990, Abschnitt 6.5). Weitere Anwendungen sind das Ranking von Webseiten in der Informatik (vgl. Knight 2008, Abschnitt 5) und die Schätzung von Passagierflüssen in Verkehrsnetzen im Operations Research (vgl. McCord u. a. 2010). Erneute Aktualität erfährt das IPF-Verfahren durch die Einführung von biproportionalen Sitzzuteilungsmethoden bei Parlamentswahlen (vgl. Balinski und Pukelsheim 2006). Die mathematische Analyse des IPF-Verfahrens konzentriert sich bisher auf die Untersuchung der Konvergenz der IPF-Folge. Diese Dissertation erweitert den Fokus dieser Analyse und untersucht die asymptotischen Struktureigenschaften des IPF-Verhaltens. Dies schließt neben der Analyse der Konvergenz der IPF-Folge auch die Untersuchung des asymptotischen Verhaltens der IPF-Folge im Divergenzfall ein. Außerdem werden im Konvergenzfall der Grenzwert und im Divergenzfall die Häufungspunkte der IPF-Folge analysiert.

Dabei werden zwei Varianten des IPF-Verfahrens unterschieden: Das zweidimensionale IPF-Verfahren passt eine vorgegebene Ausgangsmatrix an vorgegebene Zeilen- und Spaltenmarginalien an. Das mehrdimensionale IPF-Verfahren passt eine vorgegebene Ausgangstafel an vorgegebene eindimensionale und auch höherdimensionale Marginaltafeln an. Die vorliegende Dissertation ist dementsprechend in zwei Teile gegliedert. Teil I besteht aus den Kapiteln 1 bis 5 und befasst sich mit dem zweidimensionalen IPF-Verfahren. Teil II besteht aus den Kapiteln 6 bis 11 und befasst sich mit dem mehrdimensionalen IPF-Verfahren. Am Ende eines jeden Kapitels steht ein Abschnitt mit dem Titel „Kommentare und Referenzen“. Dort werden alternative Bezeichnungen, Notationen, Definitionen und Beweisansätze diskutiert und die Quellen der verwendeten Begriffe, Beweise und Resultate erörtert. Bereits an dieser Stelle sei angemerkt, dass die Unterscheidung zwischen dem zweidimensionalen IPF-Verfahren und dem mehrdimensionalen IPF-Verfahren nicht disjunkt ist. Stattdessen ist das zweidimensionale IPF-Verfahren als niedrigdimensionaler Spezialfall in der Formulierung des mehrdimensionalen IPF-Verfahrens enthalten. Die Ausgangsmatrix übernimmt dabei die Rolle einer zweidimensionalen Ausgangstafel, die Zeilenmarginalien die Rolle einer eindimensionalen Marginaltafel und die Spaltenmarginalien die Rolle einer weiteren eindimensionalen Marginaltafel. Wie in der Mathematik üblich ist der niedrigdimensionale Spezialfall anschaulicher und leichter zugänglich als der Allgemeinfall. Außerdem können für den Spezialfall zahlreiche Aussagen gezeigt werden, die nicht auf den Allgemeinfall verallgemeinert werden können. Aus diesem Grund bietet es sich an, das zweidimensionale IPF-Verfahren vorab in Teil I zu diskutieren und alle interessanten Fragestellungen zunächst für diesen Spezialfall zu unersuchen. Das mehrdimensionale IPF-Verfahren wird dann in Teil II als Verallgemeinerung des zweidimensionalen IPF-Verfahrens eingeführt und es wird analysiert, welche der für das zweidimensionale IPF-Verfahren gezeigten Aussagen verallgemeinert werden können.

In Kapitel 1 werden das zweidimensionale IPF-Verfahren sowie die zugehörigen biproportionalen Anpassungsprobleme vorgestellt und ihre grundlegenden Eigenschaften diskutiert. Diese Diskussion erfolgt anhand der vom IPF-Verfahren errechneten Matrizenfolge sowie der ebenfalls vom IPF-Verfahren errechneten Multiplikatorfolge und verzichtet auf aufwendigere Ansätze. In Kapitel 2 werden die drei in der einschlägigen Literatur vorherrschenden Ansätze zur Konvergenzanalyse des zweidimensionalen IPF-Verfahrens erörtert. Der erste Ansatz basiert auf der Informationsdivergenz (I-Divergenz). Die I-Divergenz und die darauf aufbauenden I-Projektionen werden vorgestellt. Mit der Stetigkeit von I_1-Projektionen wird das erste neue Resultat dieser Dissertation bewiesen. Anschließend werden notwendige und hinreichende Bedingungen für die Konvergenz des IPF-Verfahrens gezeigt und mithilfe der sogenannten Flussungleichungen konkretisiert. Der zweite Ansatz basiert auf einem invertierten geometrischen Mittel. Es wird gezeigt, dass dieser Ansatz äquivalent ist zum ersten Ansatz. Der dritte Ansatz basiert auf einem L_1-Fehlerfunktional und dient der Messung des Fortschritts des IPF-Verfahrens bei der Anpassung an die Zeilen- und Spaltenmarginalien. In Kapitel 3 wird das asymptotische Verhalten des IPF-Verfahrens im Divergenzfall untersucht. Als zweites neues Resultat dieser Dissertation werden die Konvergenz der IPF-Teilfolge entlang der Zeilenskalierungsschritte und die Konvergenz der IPF-Teilfolge entlang der Spaltenskalierungsschritte bewiesen. In Kapitel 4 werden die Häufungspunkte der IPF-Folge analysiert. Als neues Resultat dieser Dissertation wird die gemeinsame Zerlegung der Häufungspunkte und der Ausgangsmatrix gezeigt. Ebenfalls als neues Resultat werden die Häufungspunkte informationstheoretisch charakterisiert. Anschließend wird die Abhängigkeit der Limesmatrix beziehungsweise der Häufungspunkte von der Ausgangsmatrix und den Marginalien untersucht. Für den Konvergenzfall wird die stetige Abhängigkeit der Limesmatrix von der Ausgangsmatrix gezeigt. Für den Divergenzfall folgt daraus unter gewissen Regularitätsbedingungen die stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix. Unter Beschränkung auf sogenannte biproportionale Direktanpassungsprobleme wird die stetige Abhängigkeit der Limesmatrix von der Ausgangsmatrix und den Marginalien bewiesen. Ebenfalls unter gewissen Regularitätsbedingungen folgt daraus für den Divergenzfall die stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix und den Marginalien. Diese Stetigkeitsaussagen stellen ebenfalls neue Resultate dar. Kapitel 5 fasst die Ergebnisse zum zweidimensionalen IPF-Verfahren zusammen und benennt offene Probleme.

In Kapitel 6 wird das mehrdimensionalen IPF-Verfahren vorgestellt. Anschließend wird in Kapitel 7 der informationstheoretische Ansatz zur Konvergenzanalyse des mehrdimensionalen IPF-Verfahrens angewendet. Als neues Resultat dieser Dissertation wird gezeigt, dass der L_1-Ansatz zur Konvergenzanalyse nicht auf das mehrdimensionale IPF-Verfahren mit mehr als zwei Marginaltafeln übertragen werden kann. Für den Konvergenzfall wird in Kapitel 8 gezeigt, dass die Limestafel stetig von der Ausgangstafel abhängt. Dies stellt ebenfalls ein neues Resultat dieser Dissertation dar. In Kapitel 9 wird das asymptotische Verhalten des mehrdimensionalen IPF-Verfahrens im Divergenzfall anhand von zwei Klassen von Beispielen erörtert. Neben den bereits bekannten Beispielen von Asci und Piccioni (2003) mit zweidimensionalen Marginaltafeln wird eine neue Klasse von Beispielen mit eindimensionalen Marginaltafeln diskutiert. Für beide Beispielklassen wird die Konvergenz der IPF-Teilfolgen bewiesen, entlang welcher an dieselbe Marginaltafel angepasst wird. In Kapitel 10 wird die Hypothese aufgestellt, dass diese Konvergenzaussage allgemein für das mehrdimensionale IPF-Verfahren gilt. Die Konvergenz dieser IPF-Teilfolgen impliziert die Aussage, dass ihre Grenzwerte einen Zyklus bezüglich des mehrdimensionalen IPF-

Verfahrens bilden. Es wird gezeigt, dass eine variationelle Charakterisierung dieser Grenzwerte nicht möglich ist. Dies stellt das zentrale Resultat dieses Teils der Dissertation dar. Anschließend werden weitere Ansätze zum Beweis der Hypothese diskutiert. Kapitel 11 fasst die Ergebnisse zum mehrdimensionalen IPF-Verfahren zusammen und benennt offene Probleme.

Mit Ausnahme von Abschnitt 4.4 sind alle in dieser Dissertation auftretenden Vektoren, Matrizen und Tafeln nichtnegativ. Der Träger eines Vektors $x = (x_1, \dots, x_k) \in [0; \infty)^k$ wird mit $\operatorname{supp}(x) := \{i \in \{1, \dots, k\} \mid x_i > 0\}$ bezeichnet. Der Träger einer Matrix S oder einer Tafel S wird in analoger Weise mit $\operatorname{supp}(S)$ bezeichnet. Das Symbol „$\ll$“ steht für maßtheoretische Dominanz. Das heißt, für zwei Vektoren $x, y \in [0; \infty)^k$ ist die Schreibweise $x \ll y$ gleichbedeutend mit $\operatorname{supp}(x) \subseteq \operatorname{supp}(y)$. Das Symbol „$\equiv$“ steht für maßtheoretische Äquivalenz. Das heißt, für zwei Vektoren $x, y \in [0; \infty)^k$ ist die Schreibweise $x \equiv y$ gleichbedeutend mit $\operatorname{supp}(x) = \operatorname{supp}(y)$. Für Matrizen und Tafeln sind die Symbole „$\ll$“ und „$\equiv$“ in analoger Weise definiert. Die Unterscheidung zwischen Nulleinträgen und positiven Einträgen ist für die Analyse des IPF-Verfahrens von zentraler Bedeutung, da Nulleinträge starr sind und durch Skalierung nicht verändert werden können. Positive Einträge hingegen können durch Skalierung jeden positiven Wert annehmen und im Grenzwert auch gegen 0 konvergieren. Steht ein „+“ anstelle eines Index eines Vektors, einer Matrix oder einer Tafel, so ist damit die Summe über alle infrage kommenden Indizes gemeint. Für einen Vektor $r \in (0; \infty)^k$ gilt beispielsweise $r_+ = \sum_{i=1}^k r_i$. Genauso bezeichnet eine Menge anstelle eines Index die Summe über alle in dieser Menge enthaltenen Indizes, $r_I = \sum_{i \in I} r_i$. Folgen von Vektoren, Matrizen oder Tafeln werden entweder in Funktionenschreibweise, $(A(t))$, oder mit hochgestelltem Index, (A^n), dargestellt. Konvexitätsparameter $\alpha \in [0; 1]$ werden ebenfalls hochgestellt, $x^\alpha = (1 - \alpha)x^0 + \alpha x^1$. Hochgestellte und abgeklammerte Mengen bezeichnen den Teilvektor bestehend aus den Einträgen, deren Indizes in der jeweiligen Menge enthalten sind, $r^{(I_1)} = (r_i)_{i \in I_1}$. Potenzen von Vektoren, Matrizen oder Tafeln treten in dieser Dissertation nicht auf. Das Produktzeichen „$\prod$“ steht je nach Kontext entweder für das Produkt reeller Zahlen oder für das kartesische Produkt von Mengen. Mit Ausnahme von Kapitel 10 wird nicht zwischen Zeilen- und Spaltenvektoren unterschieden. Das Symbol „T“ bezeichnet je nach Kontext entweder einen Zeitparameter oder die Transposition eines Vektors oder einer Matrix. Werden die Funktionen exp und ln auf einen Vektor angewendet, so ist diese Abbildung komponentenweise zu verstehen. Dies wird durch die Schreibweise mit geschweiften Klammern verdeutlicht, $\exp\{(x_1, \dots, x_n)^T\} = (\exp(x_1), \dots, \exp(x_n))^T$. Die Schreibweise $\measuredangle(x, y, z)$ bezeichnet den gerichteten Winkel aus dem Intervall $[0; 2\pi)$, den die Halbgeraden $[yx$ und $[yz$ einschließen.

Teil I.

Asymptotische Struktureigenschaften des zweidimensionalen iterativen proportionalen Anpassungsverfahrens (IPF-Verfahrens)

1. Zweidimensionales IPF-Verfahren

In diesem Kapitel wird das zweidimensionale iterative proportionale Anpassungsverfahren (IPF-Verfahren, iterative proportional fitting procedure) präsentiert. Dazu wird im ersten Abschnitt das IPF-Verfahren beschrieben. Der zweite Abschnitt stellt die Begriffe der biproportionalen Limesanpassung und der biproportionalen Direktanpassung vor. Im dritten Abschnitt wird der Zusammenhangsbegriff für Matrizen definiert. Darauf aufbauend werden notwendige und hinreichende Bedingungen dafür gezeigt, dass eine Limesanpassung eine Direktanpassung ist. Der vierte Abschnitt beschreibt, wie biproportionale Anpassungsprobleme mit dünn besetzter Ausgangsmatrix in kleinere biproportionale Anpassungsprobleme zerlegt werden können. Der fünfte Abschnitt zeigt eine Möglichkeit zur Normierung biproportionaler Anpassungsprobleme.

1.1. Definition des Verfahrens

Das *zweidimensionale IPF-Verfahren* benötigt als Eingabe eine nichtnegative Matrix $A \in [0; \infty)^{k \times \ell}$ mit positiven Zeilensummen, $a_{i+} > 0$, und positiven Spaltensummen, $a_{+j} > 0$, sowie zwei positive Vektoren $r \in (0; \infty)^k$ und $c \in (0; \infty)^\ell$. Die Matrix A wird im Folgenden die *Ausgangsmatrix* genannt, während $r = (r_1, \ldots, r_k)$ als *Zeilenmarginalien* und $c = (c_1, \ldots, c_\ell)$ als *Spaltenmarginalien* bezeichnet werden.

Zur Initialisierung des IPF-Verfahrens wird $A(0) := A$ gewählt. Davon ausgehend wird durch abwechselnde Ausführung der folgenden beiden Schritte die *IPF-Folge* $(A(t))$ berechnet:

Zeilenanpassung: Ungerade Schritte $t + 1$ passen die Zeilensummen an die Zeilenmarginalien an. Zu diesem Zweck werden alle Einträge derselben Zeile mit demselben Zeilenmultiplikator multipliziert gemäß der Formel

$$a_{ij}(t + 1) := \frac{r_i}{a_{i+}(t)} \cdot a_{ij}(t) \text{ für alle Einträge } (i, j). \tag{1.1.1}$$

Spaltenanpassung: Gerade Schritte $t+2$ passen die Spaltensummen an die Spaltenmarginalien an. Zu diesem Zweck werden alle Einträge derselben Spalte mit demselben Spaltenmultiplikator multipliziert gemäß der Formel

$$a_{ij}(t + 2) := \frac{c_j}{a_{+j}(t + 1)} \cdot a_{ij}(t + 1) \text{ für alle Einträge } (i, j). \tag{1.1.2}$$

Per Induktion kann man zeigen, dass für alle Schritte $t \geq 0$ die Ungleichung $a_{ij}(t) > 0$ genau dann gilt, wenn $a_{ij} > 0$ gilt. Folglich bleiben alle Zeilensummen $a_{i+}(t)$ und alle Spaltensummen $a_{+j}(t)$ stets positiv. Das IPF-Verfahren ist somit wohldefiniert. Außerdem sind alle Matrizen der IPF-Folge maßtheoretisch äquivalent zur Ausgangsmatrix,

$$A(t) \equiv A \text{ für alle Schritte } t \geq 0. \tag{1.1.3}$$

Das *IPF-Verfahren konvergiert*, wenn die IPF-Folge $(A(t))$ konvergiert.

1.2. Biproportionale Anpassungsprobleme

Im Konvergenzfall hat die Limesmatrix $B^* := \lim_{t\to\infty} A(t)$ drei entscheidende Eigenschaften: Erstens erfüllt sie die Spaltenmarginalien. Das heißt, $b^*_{+j} = c_j$ gilt für alle Spalten j. Zweitens erfüllt sie die Zeilenmarginalien. Das heißt, $b^*_{i+} = r_i$ gilt für alle Zeilen i. Drittens ist sie biproportional zur Ausgangsmatrix A. Das heißt, es existieren Multiplikatorfolgen $\rho_i(t) > 0$ und $\sigma_j(t) > 0$, sodass $b_{ij} = \lim_{t\to\infty} \rho_i(t) a_{ij} \sigma_j(t)$ für alle Einträge (i, j) gilt. Diese Multiplikatorfolgen können berechnet werden als die Produkte der in den Gleichungen (1.1.1) und (1.1.2) auftretenden Multiplikatoren. Alle Matrizen B, die die drei genannten Eigenschaften erfüllen, werden im Folgenden als *biproportionale Limesanpassung von A an c und r* bezeichnet. Gemäß der dritten Eigenschaft gilt $B \ll A$. Pukelsheim (2014, Theorem 1) zeigt, dass die biproportionale Limesanpassung von A an c und r eindeutig ist.

Satz 1.2.1 (Eindeutigkeit biproportionaler Limesanpassungen, vgl. Pukelsheim 2014, Theorem 1)**.** *Sei $A \in [0; \infty)^{k\times\ell}$ eine Ausgangsmatrix, seien $r \in (0; \infty)^k$ Zeilenmarginalien und seien $c \in (0; \infty)^\ell$ Spaltenmarginalien im Sinne von Abschnitt 1.1. Weiter seien B und $\tilde{B}$ biproportionale Limesanpassungen von A an c und r. Dann gilt $B = \tilde{B}$.*

Beweis. Wir nehmen an, es gilt $B \neq \tilde{B}$. Dann gilt $B - \tilde{B} \neq 0$. Die Zeilensummen von $B - \tilde{B}$ und die Spaltensummen von $B - \tilde{B}$ sind jeweils gleich 0. Dies erlaubt uns die Konstruktion eines Zyklus von Einträgen

$$(i_1, j_1), (i_1, j_2), (i_2, j_2), (i_2, j_3), \ldots, (i_{q-1}, j_{q-1}), (i_{q-1}, j_q), (i_q, j_q), (i_q, j_1), \tag{1.2.1}$$

entlang welchem die Matrix $B - \tilde{B}$ ein alternierendes Vorzeichen hat. Dazu wählen wir als erstes einen Eintrag (i_1, j_1), für den $b_{i_1 j_1} > \tilde{b}_{i_1 j_1}$ gilt. Da die Zeilensummen von $B - \tilde{B}$ gleich 0 sind, existiert in der Zeile i_1 ein Eintrag (i_1, j_2), für den $b_{i_1 j_2} < \tilde{b}_{i_1 j_2}$ gilt. Da auch die Spaltensummen von $B - \tilde{B}$ gleich 0 sind, existiert in der Spalte j_2 ein Eintrag (i_2, j_2), für den $b_{i_2 j_2} > \tilde{b}_{i_2 j_2}$ gilt. Diese Suche setzen wir fort, bis wir auf eine Spalte j_Q treffen, die bereits zuvor aufgetreten ist. Das heißt, es existiert ein $P < Q$, sodass $j_P = j_Q$ gilt. Dann verwerfen wir die Einträge (i_1, j_1) bis (i_{P-1}, j_P), nummerieren die verbleibenden Einträge neu und erhalten so den Zyklus (1.2.1). Da aus $a_{ij} = 0$ die Gleichung $b_{ij} = \tilde{b}_{ij} = 0$ folgt, sind alle Einträge der Matrix A entlang dieses Zyklus positiv.

Seien nun $(\rho_i(t))$ und $(\sigma_j(t))$ die zur Limesanpassung B gehörenden Multiplikatorfolgen und seien $(\tilde{\rho}_i(t))$ und $(\tilde{\sigma}_j(t))$ die zur Limesanpassung $\tilde{B}$ gehörenden Multiplikatorfolgen. Dann gilt für alle Schritte $t \geq 0$ die Gleichungskette

$$\prod_{p=1}^{q} \frac{a_{i_p j_p}}{a_{i_p j_{p+1}}} = \prod_{p=1}^{q} \frac{\dfrac{a_{i_p j_p}}{\rho_{i_p}(t)\sigma_{j_p}(t)}}{\dfrac{a_{i_p j_{p+1}}}{\rho_{i_p}(t)\sigma_{j_{p+1}}(t)}} = \prod_{p=1}^{q} \frac{\dfrac{a_{i_p j_p}}{\tilde{\rho}_{i_p}(t)\tilde{\sigma}_{j_p}(t)}}{\dfrac{a_{i_p j_{p+1}}}{\tilde{\rho}_{i_p}(t)\tilde{\sigma}_{j_{p+1}}(t)}}, \tag{1.2.2}$$

wobei wir $j_{q+1} := j_1$ wählen. Mit dem Grenzübergang $t \to \infty$ folgt

$$\prod_{p=1}^{q} \frac{a_{i_p j_p}}{a_{i_p j_{p+1}}} = \prod_{p=1}^{q} \frac{b_{i_p j_p}}{b_{i_p j_{p+1}}} = \prod_{p=1}^{q} \frac{\tilde{b}_{i_p j_p}}{\tilde{b}_{i_p j_{p+1}}}. \tag{1.2.3}$$

Nach Konstruktion des Zyklus (1.2.1) gilt aber $b_{i_p j_p} > \tilde{b}_{i_p j_p}$ und $b_{i_p j_{p+1}} < \tilde{b}_{i_p j_{p+1}}$ für alle $p = 1, \ldots, q$. Dies steht im Widerspruch zur Gleichung (1.2.3). Folglich ist die Annahme $B \neq \tilde{B}$ falsch und es gilt $B = \tilde{B}$. □

Wenn die zur biproportionalen Limesanpassung B von A an c und r gehörenden Multiplikatorfolgen $(\rho_i(t))$ und $(\sigma_j(t))$ durch konstante Multiplikatoren $\rho_i > 0$ und $\sigma_j > 0$ ersetzt werden können, dann heißt B die *biproportionale Direktanpassung von A an c und r*.

Die im vorangegangen Abschnitt definierte Eingabe (A, c, r) des IPF-Verfahrens wird im Folgenden als *biproportionales Anpassungsproblem* bezeichnet. Wenn die biproportionale Limesanpassung von A an c und r existiert, dann heißt (A, c, r) ein *biproportionales Limesanpassungsproblem*. Wenn die biproportionale Direktanpassung von A an c und r existiert, dann heißt (A, c, r) ein *biproportionales Direktanpassungsproblem*.

1.3. Direktheit biproportionaler Limesanpassungsprobleme

Bei Vorliegen eines biproportionalen Limesanpassungsproblems (A, c, r) mit zugehörender biproportionaler Limesanpassung B liegt die Frage nahe, ob B eine Direktanpassung von A ist. Für die Beantwortung dieser Frage ist es hilfreich, sich entlang der positiven Einträge von B durch die Zeilen und Spalten bewegen zu können, ähnlich wie bei der Konstruktion des Zyklus (1.2.1). Dieser Bewegung sind allerdings Grenzen gesetzt, falls B unzusammenhängend ist im Sinne der folgenden Definition.

Definition 1.3.1 (Zusammenhang von Matrizen, vgl. Pukelsheim 2014, Abschnitt 2)**.** Sei $S \in [0; \infty)^{k \times \ell}$ eine beliebige nichtnegative Matrix. Dann heißt S *zusammenhängend*, wenn S nicht unzusammenhängend ist. Die Matrix S heißt *unzusammenhängend*, wenn eine Zeilenpermutation und eine Spaltenpermutation existieren, sodass S Blockdiagonalform annimmt,

$$S = \begin{array}{c c} & \begin{array}{cc} J & J' \end{array} \\ \begin{array}{c} I \\ I' \end{array} & \begin{pmatrix} S^{(I,J)} & 0 \\ 0 & S^{(I',J')} \end{pmatrix} \end{array}, \tag{1.3.1}$$

wobei mindestens eine der beiden Teilmengen $I \subseteq \{1, \ldots, k\}$ oder $J \subseteq \{1, \ldots, \ell\}$ eine nichtleere und echte Teilmenge ist. Der Strich bezeichnet dabei das Komplement der jeweiligen Teilmenge.

Falls die Limesanpassung B jedoch zusammenhängend ist, so können wir uns entlang der positiven Einträge von B durch alle Zeilen und durch alle Spalten bewegen. Dies ermöglicht den Beweis des folgenden Satzes.

Satz 1.3.2 (Direktheit biproportionaler Limesanpassungen bei zusammenhängender Ausgangsmatrix, vgl. Pukelsheim 2014, Theorem 2)**.** *Sei (A, c, r) ein biproportionales Limesanpassungsproblem, dessen Ausgangsmatrix A zusammenhängend ist. Weiter sei B die biproportionale Limesanpassung von A an c und r. Dann sind die folgenden drei Aussagen äquivalent:*

(i) Die Limesanpassung B ist die Direktanpassung von A an c und r.

(ii) Die Limesanpassung B ist maßtheoretisch äquivalent zu A, $B \equiv A$.

(iii) Die Limesanpassung B ist zusammenhängend.

Beweis. (i) $\Rightarrow$ (ii) : Sei B die biproportionale Direktanpassung von A an c und r. Dann existieren Zeilenmultiplikatoren $\rho_i > 0$ und Spaltenmultiplikatoren $\sigma_j > 0$, sodass $b_{ij} = \rho_i a_{ij} \sigma_j$ für alle Einträge (i, j) gilt. Folglich gilt $b_{ij} = 0$ genau dann, wenn $a_{ij} = 0$ gilt, und B ist maßtheoretisch äquivalent zu A.

(ii) $\Rightarrow$ (iii) : Sei B maßtheoretisch äquivalent zu A. Dann gilt $b_{ij} = 0$ genau dann, wenn $a_{ij} = 0$ gilt. Also folgt aus dem Zusammenhang von A der Zusammenhang von B.

(iii) $\Rightarrow$ (i) : Sei B zusammenhängend. Dann können wir durch einen Scan-Algorithmus Zeilenmultiplikatoren $\rho_i > 0$ und Spaltenmultiplikatoren $\sigma_j > 0$ ermitteln, sodass $b_{ij} = \rho_i a_{ij} \sigma_j$ für alle Einträge (i, j) gilt. Dazu normieren wir als erstes die Multiplikatorfolgen $(\rho_i(t))$ und $(\sigma_j(t))$, indem wir $\tilde{\rho}_i(t) := \rho_i(t)/\rho_1(t)$ und $\tilde{\sigma}_j(t) := \sigma_j(t) \cdot \rho_1(t)$ wählen. Folglich gilt $\tilde{\rho}_1(t) = 1$ für alle Schritte $t \geq 0$. Dementsprechend wählen wir $\rho_1 := 1$. Als nächstes betrachten wir alle Spalten j, für die $b_{1j} > 0$ gilt, und setzen

$$\sigma_j := \frac{b_{1j}}{\rho_1 a_{1j}} = \frac{a_{1j} \lim_{t\to\infty} \tilde{\rho}_1(t)\tilde{\sigma}_j(t)}{a_{1j} \lim_{t\to\infty} \tilde{\rho}_1(t)} = \lim_{t\to\infty} \tilde{\sigma}_j(t). \tag{1.3.2}$$

Folglich gilt $b_{1j} = \rho_1 a_{1j} \sigma_j$ für alle bereits gescannten Spalten j. Als nächstes betrachten wir alle noch nicht gescannten Zeilen i, für die $b_{ij} > 0$ für eine bereits gescannte Spalte j gilt, und setzen

$$\rho_i := \frac{b_{ij}}{a_{ij}\sigma_j} = \frac{a_{ij} \lim_{t\to\infty} \tilde{\rho}_i(t)\tilde{\sigma}_j(t)}{a_{ij} \lim_{t\to\infty} \tilde{\sigma}_j(t)} = \lim_{t\to\infty} \tilde{\rho}_i(t). \tag{1.3.3}$$

Folglich gilt $b_{ij} = \rho_i a_{ij} \sigma_j$ für alle bereits gescannten Zeilen i und alle bereits gescannten Spalten j. Anschließend wenden wir uns wieder den noch nicht gescannten Spalten zu, danach den noch nicht gescannten Zeilen. Der Scan-Algorithmus terminiert nach maximal $k + \ell$ Schritten. Sei nun I die Menge der gescannten Zeilen und J die Menge der gescannten Spalten. Aufgrund der Terminierung des Scan-Algorithmus gilt bis auf Zeilen- und Spaltenpermutation

$$B = \begin{pmatrix} B^{(I,J)} & 0 \\ 0 & B^{(I',J')} \end{pmatrix}. \tag{1.3.4}$$

Da B nach Voraussetzung zusammenhängend ist, müssen $I = \{1, \ldots, k\}$ und $J = \{1, \ldots, \ell\}$ gelten. Folglich gilt $b_{ij} = \rho_i a_{ij} \sigma_j$ für alle Einträge (i, j). Damit ist B die biproportionale Direktanpassung von A an c und r. □

Der soeben gezeigte Satz 1.3.2 nennt zwei jeweils notwendige und hinreichende Kriterien dafür, dass die biproportionale Limesanpassung B der Ausgangsmatrix A eine Direktanpassung ist. Diese Kriterien gelten jedoch nur unter der Bedingung, dass A zusammenhängend ist.

Nach Definition 1.3.1 kann aber jede unzusammenhängende Matrix S in zwei Blöcke $S^{(I,J)}$ und $S^{(I',J')}$ zerlegt werden. Ist einer dieser Blöcke unzusammenhängend, so kann er weiter zerlegt werden. Auf diese Weise erhalten wir schrittweise bis auf Zeilen- und Spaltenpermutation eine

Zerlegung

$$S = \begin{pmatrix} S^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & S^{(I_M,J_M)} \end{pmatrix}, \tag{1.3.5}$$

deren Blöcke $S^{(I_1,J_1)}, S^{(I_2,J_2)}, \ldots, S^{(I_M,J_M)}$ jeweils zusammenhängend sind. Diese Blöcke werden daher im Folgenden als die *Zusammenhangskomponenten der Matrix S* bezeichnet.

Die Zerlegung der Ausgangsmatrix A in ihre Zusammenhangskomponenten ermöglicht den Beweis der folgenden Verallgemeinerung von Satz 1.3.2.

Satz 1.3.3 (Direktheit biproportionaler Limesanpassungen)**.** *Sei (A, c, r) ein biproportionales Limesanpassungsproblem. Weiter sei B die biproportionale Limesanpassung von A an c und r. Dann sind die folgenden beiden Aussagen äquivalent:*

(i) Die Limesanpassung B ist die Direktanpassung von A an c und r.

(ii) Die Limesanpassung B ist maßtheoretisch äquivalent zu A, $B \equiv A$.

Beweis. Sei (A, c, r) ein biproportionales Limesanpassungsproblem und sei B die zugehörige biproportionale Limesanpassung. Weiter seien $A^{(I_1,J_1)}$, $A^{(I_2,J_2)}$, $\ldots$, $A^{(I_M,J_M)}$ die Zusammenhangskomponenten von A. Da $B \ll A$ gilt, können wir B in die Blöcke $B^{(I_1,J_1)}$, $B^{(I_2,J_2)}$, $\ldots$, $B^{(I_M,J_M)}$ zerlegen und erhalten so die gemeinsame Zerlegung

$$A = \begin{pmatrix} A^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A^{(I_M,J_M)} \end{pmatrix} \text{ und } B = \begin{pmatrix} B^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B^{(I_M,J_M)} \end{pmatrix}. \tag{1.3.6}$$

Die Marginalien c und r können wir entsprechend der Zerlegung von A in die Komponenten $c^{(J_1)}$, $c^{(J_2)}, \ldots, c^{(J_M)}$ und $r^{(I_1)}, r^{(I_2)}, \ldots, r^{(I_M)}$ zerlegen. Folglich ist $B^{(I_m,J_m)}$ für alle $m = 1, \ldots, M$ die biproportionale Limesanpassung von $A^{(I_m,J_m)}$ an $c^{(J_m)}$ und $r^{(I_m)}$. Des Weiteren ist B genau dann eine Direktanpassung von A, wenn $B^{(I_m,J_m)}$ für alle $m = 1, \ldots, M$ eine Direktanpassung von $A^{(I_m,J_m)}$ ist. Dies ist nach Satz 1.3.2 genau dann der Fall, wenn $B^{(I_m,J_m)} \equiv A^{(I_m,J_m)}$ für alle $m = 1, \ldots, M$ gilt. Dies ist wiederum genau dann der Fall, wenn $B \equiv A$ gilt. □

1.4. Zerlegung biproportionaler Anpassungsprobleme

Bei vielen in der Praxis auftretenden biproportionalen Anpassungsproblemen (A, c, r) ist die Ausgangsmatrix A dünn besetzt. Das heißt, A enthält nur wenige positive Einträge. In diesen Fällen bietet sich eine Überprüfung des Zusammenhangs von A an. Ist A unzusammenhängend, so kann A in ihre Zusammenhangskomponenten $A^{(I_1,J_1)}$, $A^{(I_2,J_2)}$, $\ldots$, $A^{(I_M,J_M)}$ zerlegt werden. Wie im Beweis von Satz 1.3.3 können die Marginalien c und r in die entsprechenden Komponenten

$c^{(J_1)}, c^{(J_2)}, \ldots, c^{(J_M)}$ und $r^{(I_1)}, r^{(I_2)}, \ldots, r^{(I_M)}$ zerlegt werden. Auf diese Weise erhalten wir die biproportionalen Anpassungsprobleme $(A^{(I_m,J_m)}, c^{(J_m)}, r^{(I_m)})$, $m = 1, \ldots, M$. Die separate Ausführung des IPF-Verfahrens auf diesen Anpassungsproblemen liefert dasselbe Resultat wie die Ausführung des IPF-Verfahrens auf dem ursprünglichen Anpassungsproblem (A, c, r). Außerdem ist die Existenz der biproportionalen Limesanpassung von A an c und r äquivalent zur Existenz der biproportionalen Limesanpassungen von $A^{(I_m,J_m)}$ an $c^{(J_m)}$ und $r^{(I_m)}$ für alle $m = 1, \ldots, M$. Ebenso ist die Direktheit der biproportionalen Limesanpassung von A äquivalent zur Direktheit der biproportionalen Limesanpassungen von $A^{(I_m,J_m)}$ für alle $m = 1, \ldots, M$.

Bei biproportionalen Limesanpassungsproblemen (A, c, r) mit unzusammenhängender Ausgangsmatrix A können wir, wie im Beweis von Satz 1.3.3, neben den Marginalien c und r auch die Limesanpassung B entsprechend der Ausgangsmatrix A zerlegen. Auf diese Weise erhalten die gemeinsame Zerlegung

$$A = \begin{pmatrix} A^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A^{(I_M,J_M)} \end{pmatrix} \text{ und } B = \begin{pmatrix} B^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B^{(I_M,J_M)} \end{pmatrix}. \tag{1.4.1}$$

Dabei gilt es jedoch zu beachten, dass die einzelnen Blöcke $B^{(I_1,J_1)}, B^{(I_2,J_2)}, \ldots, B^{(I_M,J_M)}$ im Allgemeinen nicht zusammenhängend sind. Möchten wir stattdessen eine gemeinsame Zerlegung der Ausgangsmatrix A und der biproportionalen Limesanpassung B bestimmen, bei der B in seine Zusammenhangskomponenten zerfällt, so benötigen wir das folgende Lemma.

Lemma 1.4.1 (Gemeinsame Zerlegung, vgl. Gietl 2009, Lemma 3.6)**.** *Sei (A, c, r) ein biproportionales Limesanpassungsproblem mit einer zusammenhängenden Ausgangsmatrix A und einer unzusammenhängenden biproportionalen Limesanpassung B. Dann existieren eine nichtleere und echte Zeilenteilmenge $I \subsetneq \{1, \ldots, k\}$, $I \neq \emptyset$, sowie eine nichtleere und echte Spaltenteilmenge $J \subsetneq \{1, \ldots, \ell\}$, $J \neq \emptyset$, sodass die Matrizen A und B die Form*

$$A = \begin{pmatrix} A^{(I,J)} & 0 \\ A^{(I',J)} & A^{(I',J')} \end{pmatrix} \text{ und } B = \begin{pmatrix} B^{(I,J)} & 0 \\ 0 & B^{(I',J')} \end{pmatrix} \tag{1.4.2}$$

annehmen.

Beweis. Der Beweis erfolgt in vier Schritten. In Schritt I wählen wir die Teilmengen I und J. Dass die so gewählten Teilmengen I und J sowie ihre Komplemente I' und J' nichtleer sind, zeigen wir in Schritt II. In Schritt III beweisen wir die Aussagen $B^{(I',J)} = 0$ und $B^{(I,J')} = 0$. Die Aussage $A^{(I,J')} = 0$ zeigen wir schließlich in Schritt IV.

I. Seien $(\rho_i(t))$ und $(\sigma_j(t))$ die zur Limesanpassung B gehörenden Multiplikatorfolgen. Wir normieren diese Multiplikatorfolgen, indem wir $\rho_{\min}(t) := \min_i \rho_i(t) > 0$ und damit $\tilde{\rho}_i(t) := \rho_i(t)/\rho_{min}(t) \geq 1$ sowie $\tilde{\sigma}_j(t) := \sigma_j(t) \cdot \rho_{\min}(t) > 0$ wählen. Folglich gilt

$$b_{ij} = \lim_{t\to\infty} \tilde{\rho}_i(t) \cdot a_{ij} \cdot \tilde{\sigma}_j(t) \text{ für alle Einträge } (i, j). \tag{1.4.3}$$

Von den normierten Multiplikatorfolgen ausgehend wählen wir $J := \{j \mid \lim_{t\to\infty} \tilde{\sigma}_j(t) = 0\}$ und $I := \{i \mid \exists j \in J : b_{ij} > 0\}$.

II. Für die biproportionale Limesanpassung B gilt $B \ll A$. Da B jedoch unzusammenhängend und A zusammenhängend ist, existiert ein Eintrag (i_0, j_0) mit $a_{i_0 j_0} > 0$ und $b_{i_0 j_0} = 0$. Folglich gilt

$$\lim_{t\to\infty} \tilde{\rho}_{i_0}(t) \cdot \tilde{\sigma}_{j_0}(t) = 0. \tag{1.4.4}$$

Mit $\tilde{\rho}_{i_0}(t) \geq 1$ für alle Schritte t folgt

$$\limsup_{t\to\infty} \tilde{\sigma}_{j_0}(t) \leq \limsup_{t\to\infty} \tilde{\rho}_{i_0}(t) \cdot \tilde{\sigma}_{j_0}(t) = 0 \tag{1.4.5}$$

und somit $J \neq \emptyset$. Nach Wahl von I gilt außerdem $B^{(I',J)} = 0$ und somit $b_{IJ} = b_{+J} = r_J > 0$. Daraus folgt $I \neq \emptyset$.

Sei nun $i_1 \in I$ beliebig. Dann existiert eine Spalte $j_1 \in J$ mit $b_{i_1 j_1} > 0$ und damit auch $a_{i_1 j_1} > 0$. Folglich gilt

$$\lim_{t\to\infty} \tilde{\rho}_{i_1}(t) \cdot \tilde{\sigma}_{j_1}(t) = \frac{b_{i_1 j_1}}{a_{i_1 j_1}} \in (0; \infty). \tag{1.4.6}$$

Weil $\lim_{t\to\infty} \tilde{\sigma}_{j_1}(t) = 0$ nach Wahl von J gilt, folgt $\lim_{t\to\infty} \tilde{\rho}_{i_1}(t) = \infty$. Also gilt

$$\lim_{t\to\infty} \tilde{\rho}_i(t) = \infty \text{ für alle Zeilen } i \in I. \tag{1.4.7}$$

Da $\tilde{\rho}_{\min}(t) = 1$ für alle Schritte t gilt, ist die Menge I' nichtleer. Wegen $B^{(I',J)} = 0$ gilt außerdem $b_{I'J'} = b_{I'+} = r_{I'} > 0$ und somit $J' \neq \emptyset$.

III. Wie bereits im vorangegangenen Schritt erwähnt, gilt $B^{(I',J)} = 0$.

Wir nehmen an, es existiert ein Eintrag $(i_2, j_2) \in I \times J'$ mit $b_{i_2 j_2} > 0$. Dann gilt $a_{i_2 j_2} > 0$ und damit

$$\lim_{t\to\infty} \tilde{\rho}_{i_2}(t) \cdot \tilde{\sigma}_{j_2}(t) = \frac{b_{i_2 j_2}}{a_{i_2 j_2}} \in (0; \infty). \tag{1.4.8}$$

Da $i_2 \in I$ und somit $\lim_{t\to\infty} \tilde{\rho}_{i_2}(t) = \infty$ nach Gleichung (1.4.7) gilt, folgt die Gleichung $\lim_{t\to\infty} \tilde{\sigma}_{j_2}(t) = 0$ und schließlich $j_2 \in J$ im Widerspruch zur Annahme. Also gilt $B^{(I,J')} = 0$.

IV. Wir nehmen an, es existiert ein Eintrag $(i_3, j_3) \in I \times J'$ mit $a_{i_3 j_3} > 0$. Dann gilt

$$\lim_{t\to\infty} \tilde{\rho}_{i_3}(t) \cdot \tilde{\sigma}_{j_3}(t) = 0 \tag{1.4.9}$$

und somit

$$\limsup_{t\to\infty} \tilde{\sigma}_{j_3}(t) \leq \limsup_{t\to\infty} \tilde{\rho}_{i_3}(t) \cdot \tilde{\sigma}_{j_3}(t) = 0. \tag{1.4.10}$$

Folglich gilt $j_3 \in J$ im Widerspruch zur Annahme. Also gilt $A^{(I,J')} = 0$. □

Das soeben gezeigte Lemma 1.4.1 und die in Gleichung (1.4.1) beschriebene Zerlegung ermöglichen den Beweis des folgenden Satzes.

Satz 1.4.2 (Gemeinsame Zerlegung, vgl. Gietl 2009, Satz 3.7). *Sei (A, c, r) ein biproportionales Limesanpassungsproblem. Weiter sei B die biproportionale Limesanpassung von A an c und r. Dann existieren disjunkte Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1, \ldots, k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \ldots, \ell\}$, sodass die Matrizen A und B die Form*

$$A = \begin{pmatrix} A^{(I_1,J_1)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A^{(I_M,J_1)} & \cdots & \cdots & A^{(I_M,J_M)} \end{pmatrix} \text{ und } B = \begin{pmatrix} B^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B^{(I_M,J_M)} \end{pmatrix} \tag{1.4.11}$$

annehmen. Dabei sind für alle $m = 1, \ldots, M$ die Diagonalblöcke $A^{(I_m,J_m)}$ und $B^{(I_m,J_m)}$ jeweils zusammenhängend.

Beweis. Der Beweis erfolgt per vollständiger Induktion über die Größe $g := \max\{k, \ell\}$ des biproportionalen Limesanpassungsproblems (A, c, r). In trivialer Weise gilt die Behauptung für alle biproportionalen Limesanpassungsprobleme der Größe 1. Die Behauptung gelte nun für alle biproportionalen Limesanpassungsprobleme der Größen 1 bis γ. Sei (A, c, r) daher ein biproportionales Limesanpassungsproblem der Größe $\gamma + 1 = \max\{k, \ell\}$ und sei B die zugehörige biproportionale Limesanpassung. Wir zeigen im Folgenden, dass die Matrizen A und B in der Form (1.4.11) mit zusammenhängenden Diagonalblöcken zerlegt werden können. Dabei unterscheiden wir drei Fälle. Die Behauptung folgt dann induktiv.

I. Sei B zusammenhängend. Dann ist auch A zusammenhängend. Folglich erfüllen die uneigentlichen Zerlegungen $I_1 := \{1, \ldots, k\}$ und $J_1 := \{1, \ldots, \ell\}$ die Anforderungen.

II. Sei B unzusammenhängend und A unzusammenhängend. Dann existieren nichtleere und echte Teilmengen $I \subsetneq \{1, \ldots, k\}$, $I \neq \emptyset$, und $J \subsetneq \{1, \ldots, \ell\}$, $J \neq \emptyset$, sodass die Matrizen A und B die Form

$$A = \begin{pmatrix} A^{(I,J)} & 0 \\ 0 & A^{(I',J')} \end{pmatrix} \text{ und } B = \begin{pmatrix} B^{(I,J)} & 0 \\ 0 & B^{(I',J')} \end{pmatrix} \tag{1.4.12}$$

annehmen. Dabei ist die Matrix $B^{(I,J)}$ die biproportionale Limesanpassung des Anpassungsproblems $(A^{(I,J)}, c^{(J)}, r^{(I)})$. Folglich zerfallen die Matrizen $A^{(I,J)}$ und $B^{(I,J)}$ nach Induktionsvoraussetzung entsprechend der Form (1.4.11), wobei deren Diagonalblöcke zusammenhängend sind. Die Matrizen $A^{(I',J')}$ und $B^{(I',J')}$ zerfallen in analoger Weise. Einsetzen der beiden Zerfallsstrukturen in die Form (1.4.12) führt zum gewünschten Ergebnis.

III. Sei schließlich B unzusammenhängend und A zusammenhängend. Dann existieren nach Lemma 1.4.1 nichtleere und echte Teilmengen $I \subsetneq \{1, \ldots, k\}$, $I \neq \emptyset$, und $J \subsetneq \{1, \ldots, \ell\}$, $J \neq \emptyset$, sodass die Matrizen A und B die Form (1.4.2) annehmen. Dabei ist die Matrix $B^{(I,J)}$ die biproportionale Limesanpassung des Anpassungsproblems $(A^{(I,J)}, c^{(J)}, r^{(I)})$. Folglich zerfallen die Matrizen $A^{(I,J)}$ und $B^{(I,J)}$ nach Induktionsvoraussetzung entsprechend der Form (1.4.11), wobei deren Diagonalblöcke zusammenhängend sind. Die Matrizen $A^{(I',J')}$ und $B^{(I',J')}$ zerfallen in analoger Weise. Einsetzen der beiden Zerfallsstrukturen in die Form (1.4.2) führt zum gewünschten Ergebnis. □

1.5. Normierung biproportionaler Anpassungsprobleme

Bei der Analyse des IPF-Verfahrens ist es häufig von Vorteil, die Ausgangsmatrix A und die Marginalien c und r zu normieren,

$$\tilde{A} := \frac{1}{a_{++}} \cdot A, \qquad \tilde{c} := \frac{1}{c_{+}} \cdot c, \qquad \tilde{r} := \frac{1}{r_{+}} \cdot r. \tag{1.5.1}$$

Die ausgehend von dem *normierten biproportionalen Anpassungsproblem* $(\tilde{A}, \tilde{c}, \tilde{r})$ errechnete IPF-Folge $(\tilde{A}(t))$ entspricht dann einer gliedweisen Normierung der ausgehend von dem ursprünglichen biproportionalen Anpassungsproblem (A, c, r) errechneten IPF-Folge $(A(t))$.

Satz 1.5.1 (Normierung biproportionaler Anpassungsprobleme)**.** *Sei (A, c, r) ein biproportionales Anpassungsproblem und sei $(\tilde{A}, \tilde{c}, \tilde{r})$ das zugehörige normierte biproportionale Anpassungsproblem. Weiter sei $(A(t))$ die ausgehend von (A, c, r) errechnete IPF-Folge und sei $(\tilde{A}(t))$ die ausgehend von $(\tilde{A}, \tilde{c}, \tilde{r})$ errechnete IPF-Folge. Dann gelten für alle geraden Schritte $t \geq 0$ die Gleichungen*

$$A(t+1) = r_{+} \cdot \tilde{A}(t+1) \qquad \textit{und} \qquad A(t+2) = c_{+} \cdot \tilde{A}(t+2). \tag{1.5.2}$$

Beweis. Der Beweis erfolgt per Induktion über die geraden Schritte $t \geq 0$. Nach Voraussetzung gilt $A = a_{++} \cdot \tilde{A}$ sowie $c = c_{+} \cdot \tilde{c}$ und $r = r_{+} \cdot \tilde{r}$. Gemäß Definition des IPF-Verfahrens folgt daraus für alle Einträge (i, j) die Aussage

$$\begin{aligned} a_{ij}(1) &= \frac{r_i}{a_{i+}} \cdot a_{ij} = \frac{r_{+} \cdot \tilde{r}_i}{a_{++} \cdot \tilde{a}_{i+}} \cdot \left(a_{++} \cdot \tilde{a}_{ij}\right) \\ &= r_{+} \cdot \frac{\tilde{r}_i}{\tilde{a}_{i+}} \cdot \tilde{a}_{ij} = r_{+} \cdot \tilde{a}_{ij}(1). \end{aligned} \tag{1.5.3}$$

Daraus folgt wiederum für alle Einträge (i, j) die Aussage

$$\begin{aligned} a_{ij}(2) &= \frac{c_j}{a_{+j}(1)} \cdot a_{ij}(1) = \frac{c_{+} \cdot \tilde{c}_j}{r_{+} \cdot \tilde{a}_{+j}(1)} \cdot \left(r_{+} \cdot \tilde{a}_{ij}(1)\right) \\ &= c_{+} \cdot \frac{\tilde{c}_j}{\tilde{a}_{+j}(1)} \cdot \tilde{a}_{ij}(1) = c_{+} \cdot \tilde{a}_{ij}(2). \end{aligned} \tag{1.5.4}$$

Somit sind die Gleichungen (1.5.2) für den Schritt $t = 0$ gezeigt.

Gelten die Gleichungen (1.5.2) nun für einen beliebigen geraden Schritt t. Mit der Definition des IPF-Verfahrens folgt daraus für alle Einträge (i, j) die Aussage

$$\begin{aligned} a_{ij}(t+3) &= \frac{r_i}{a_{i+}(t+2)} \cdot a_{ij}(t+2) = \frac{r_{+} \cdot \tilde{r}_i}{c_{+} \cdot \tilde{a}_{i+}(t+2)} \cdot \left(c_{+} \cdot \tilde{a}_{ij}(t+2)\right) \\ &= r_{+} \cdot \frac{\tilde{r}_i}{\tilde{a}_{i+}(t+2)} \cdot \tilde{a}_{ij}(t+2) = r_{+} \cdot \tilde{a}_{ij}(t+3). \end{aligned} \tag{1.5.5}$$

Daraus folgt wiederum für alle Einträge (i, j) die Aussage

$$\begin{aligned} a_{ij}(t+4) &= \frac{c_j}{a_{+j}(t+3)} \cdot a_{ij}(t+3) = \frac{c_{+} \cdot \tilde{c}_j}{r_{+} \cdot \tilde{a}_{+j}(t+3)} \cdot \left(r_{+} \cdot \tilde{a}_{ij}(t+3)\right) \\ &= c_{+} \cdot \frac{\tilde{c}_j}{\tilde{a}_{+j}(t+3)} \cdot \tilde{a}_{ij}(t+3) = c_{+} \cdot \tilde{a}_{ij}(t+4). \end{aligned} \tag{1.5.6}$$

Somit sind die Gleichungen (1.5.2) für den Schritt $t + 2$ gezeigt. Die Behauptung folgt induktiv. □

1.6. Kommentare und Referenzen

Die in Abschnitt 1.1 gegebene Definition des zweidimensionalen IPF-Verfahrens sowie die Bezeichnung „IPF-Verfahren“ orientieren sich an Pukelsheim (2014, Abschnitt 3). Anstelle von Zeilen- und von Spaltenmultiplikatoren verwendet Pukelsheim jedoch Zeilen- und Spaltendivisoren. Zugunsten der besseren Vergleichbarkeit mit den Werken anderer Autoren und einer übersichtlicheren Notation wird hier den Multiplikatoren der Vorzug gegeben. Das zweidimensionale IPF-Verfahren wurde von Kruithof (1937, S. E20 ff.) eingeführt und von Deming und Stephan (1940) bekannt gemacht. Während Kruithof das IPF-Verfahren als „methode der dubbele factoren“ bezeichnet, sprechen Deming und Stephan von der „method of iterative proportions“. Darüber hinaus sind laut Pukelsheim (2014, Abschnitt 1.1) die Bezeichnungen „matrix raking“, „matrix scaling“ und „RAS method“ gebräuchlich.

Die in Abschnitt 1.2 gegebenen Definitionen von biproportionalen Limesanpassungen und von biproportionalen Direktanpassungen sowie der Beweis des Eindeutigkeitssatzes 1.2.1 wurden von Pukelsheim (2014, Abschnitt 2) übernommen. Die Existenz biproportionaler Limesanpassungen wurde von zahlreichen Autoren untersucht. Ein Überblick findet sich bei Rothblum und Schneider (1989, Abschnitte 1 und 3). In diesem Kapitel wurde auf derartige Existenzaussagen verzichtet. Stattdessen wird die Existenz biproportionaler Limesanpassungen im Anschluss an die Konvergenzanalyse des zweidimensionalen IPF-Verfahrens diskutiert (siehe Satz 2.3.3 (i), S. 34).

Definition 1.3.1 und der Beweis des Satzes 1.3.2 wurden ebenfalls von Pukelsheim (2014, Abschnitt 2) übernommen.

Lemma 1.4.1 und Satz 1.4.2 wurden bereits in Gietl (2009, Abschnitt 3.4) gezeigt. Die sowohl in Abschnitt 1.4 als auch in Gietl (2009, Abschnitt 3.4) angegebenen Beweise folgen der Idee des Beweises von Theorem 2 bei Bacharach (1970). Durch die Normierung der Multiplikatoren und die Verwendung des Limes superior werden jedoch die bei Bacharach auftretenden Probleme mit nicht existierenden Häufungspunkten umgangen. Eine mit Lemma 1.4.1 vergleichbare Aussage findet sich außerdem bei Balinski und Demange (1989, Lemma 4).

Satz 1.5.1 wurde in ähnlicher Form bereits in Gietl (2009, Abschnitt 3.1) gezeigt. Die Idee, die Analyse des IPF-Verfahrens auf normierte Ausgangsmatrizen sowie normierte Marginalien zu beschränken, geht mindestens auf Csiszár (1975, Korollar 3.3) zurück. Csiszár verzichtet jedoch auf eine Diskussion dieser Einschränkung.

2. Konvergenzanalyse des zweidimensionalen IPF-Verfahrens

In diesem Kapitel werden drei Ansätze zur Konvergenzanalyse des zweidimensionalen IPF-Verfahrens vorgestellt. Der erste und für diese Dissertation zentrale Ansatz basiert auf der Informationsdivergenz (I-Divergenz). Die I-Divergenz wird im ersten Abschnitt diskutiert. Darauf aufbauend werden im zweiten Abschnitt I_1-Projektionen und I_2-Projektionen untersucht. Der dritte Abschnitt präsentiert den informationstheoretischen Ansatz zur Konvergenzanalyse. Die dort gezeigten notwendigen und hinreichenden Bedingungen für die Konvergenz des IPF-Verfahrens werden anschließend im vierten Abschnitt mithilfe der Flussungleichungen konkretisiert. Der zweite Ansatz basiert auf einem invertierten geometrischen Mittel und wird im fünften Abschnitt vorgestellt. Es wird gezeigt, dass dieser Ansatz äquivalent ist zum I-Divergenz-Ansatz. Der dritte Ansatz basiert auf einem L_1-Fehlerfunktional und wird im sechsten Abschnitt diskutiert. Er dient der Messung des Fortschritts des IPF-Verfahrens bei der Anpassung an die Zeilen- und Spaltenmarginalien.

2.1. Informationsdivergenz (I-Divergenz)

Die Informationsdivergenz spielt eine entscheidende Rolle in der Analyse des IPF-Verfahrens. Ihre formale Definition erfordert ein wenig Vorarbeit. Die Funktion $\varphi : (0;\infty)^2 \to (-\infty;\infty)$ sei definiert gemäß

$$\varphi(x,y) := x \ln \frac{x}{y}. \tag{2.1.1}$$

Aus der Konvexität der Abbildung $z \mapsto -\ln z$ folgt für alle $(x,y), (\tilde{x},\tilde{y}) \in (0;\infty)^2$ und für alle $\alpha \in [0;1]$ die Ungleichung

$$\begin{aligned}
&\varphi((1-\alpha)x + \alpha\tilde{x}, (1-\alpha)y + \alpha\tilde{y}) \\
&\qquad = ((1-\alpha)x + \alpha\tilde{x}) \ln \frac{(1-\alpha)x + \alpha\tilde{x}}{(1-\alpha)y + \alpha\tilde{y}} \\
&\qquad = ((1-\alpha)x + \alpha\tilde{x}) \left(-\ln\left(\frac{(1-\alpha)x}{(1-\alpha)x + \alpha\tilde{x}}\frac{y}{x} + \frac{\alpha\tilde{x}}{(1-\alpha)x + \alpha\tilde{x}}\frac{\tilde{y}}{\tilde{x}}\right)\right) \\
&\qquad \le (1-\alpha)x\left(-\ln\frac{y}{x}\right) + \alpha\tilde{x}\left(-\ln\frac{\tilde{y}}{\tilde{x}}\right) \\
&\qquad = (1-\alpha)\varphi(x,y) + \alpha\varphi(\tilde{x},\tilde{y}).
\end{aligned} \tag{2.1.2}$$

Folglich ist die Funktion φ konvex. Weiter sei $\overline{\varphi} : [0;\infty)^2 \to (-\infty;\infty]$ der Abschluss von φ im Sinne von Rockafellar (1972, Abschnitt 7). Dann ist $\overline{\varphi}$ konvex und halbstetig von unten und es gilt

$$\overline{\varphi}(x,y) := \liminf_{(\tilde{x},\tilde{y})\to(x,y)} \varphi(\tilde{x},\tilde{y}) = \begin{cases} x\ln\frac{x}{y} & \text{falls } x>0 \text{ und } y>0,\\ \infty & \text{falls } x>0 \text{ und } y=0,\\ 0 & \text{falls } x=0 \text{ und } y>0,\\ 0 & \text{falls } x=0 \text{ und } y=0. \end{cases} \tag{2.1.3}$$

Für festes $x > 0$ ist die Funktion $\overline{\varphi}(x,\cdot)$ strikt konvex auf $[0;\infty)$, denn $y \mapsto -\ln y$ ist strikt konvex. Für festes $y > 0$ ist die Funktion $\overline{\varphi}(\cdot,y)$ strikt konvex auf $[0;\infty)$, denn $x \mapsto x\ln x$ ist strikt konvex und $x \mapsto x\ln y$ ist linear. Des Weiteren ist die Funktion $\overline{\varphi}(x,y)$ für festes $x \geq 0$ stetig in $y \in [0;\infty)$. Im Fall $x = 0$ ist die Funktion $\overline{\varphi}(x,y)$ konstant und im Fall $x > 0$ folgt die Behauptung aus der Stetigkeit von $y \mapsto -\ln y$.

Definition 2.1.1 (I-Divergenz, vgl. Kullback und Leibler 1951, Abschnitt 2)**.** Seien $x \in [0;\infty)^k$ und $y \in [0;\infty)^k$ beliebig. Dann ist die *Informationsdivergenz (I-Divergenz) von x relativ zu y* definiert als

$$\mathrm{I}(x|y) := \sum_i \overline{\varphi}(x_i,y_i) = \begin{cases} \sum_{i:\,x_i>0} x_i \ln\frac{x_i}{y_i} & \text{falls } x \ll y,\\ \infty & \text{sonst.} \end{cases} \tag{2.1.4}$$

Für Matrizen ist die I-Divergenz in analoger Weise definiert. Die Summation erfolgt in diesem Fall über alle Einträge beziehungsweise über alle positiven Einträge.

Streng genommen müsste man zwischen den einzelnen I-Divergenzen $\mathrm{I}_k : [0;\infty)^k \times [0;\infty)^k \to (-\infty;\infty]$, $k \in \mathbb{N}$, sowie $\mathrm{I}_{k\times\ell} : [0;\infty)^{k\times\ell} \times [0;\infty)^{k\times\ell} \to (-\infty;\infty]$, $k,\ell \in \mathbb{N}$, unterscheiden, da es sich bei ihnen um Funktionen mit verschiedenen Definitionsbereichen handelt. Im Folgenden wird jedoch auf diese Unterscheidung verzichtet. Stattdessen werden alle I-Divergenzen mit I bezeichnet. Welche Funkion I_k oder $\mathrm{I}_{k\times\ell}$ gemeint ist, kann an den Argumenten der I-Divergenz abgelesen werden.

Aus Gleichung (2.1.4) geht hervor, dass die I-Divergenz $\mathrm{I}(x|y)$ genau dann endlich ist, wenn $x \ll y$ gilt. Die oben diskutierten Konvexitäts- und Stetigkeitseigenschaften der Funktion $\overline{\varphi}$ übertragen sich wie folgt auf die I-Divergenz.

Satz 2.1.2 (Konvexität und Stetigkeit der I-Divergenz, vgl. Gietl und Reffel 2013a, Theorem 2.1)**.** *Für die oben definierte I-Divergenz* $\mathrm{I} : [0;\infty)^k \times [0;\infty)^k \to (-\infty;\infty]$ *gilt:*

(i) Die Abbildung $(x,y) \mapsto \mathrm{I}(x|y)$ *ist konvex auf* $[0;\infty)^k \times [0;\infty)^k$.

(ii) Für festes $x \in [0;\infty)^k$ *ist die Abbildung* $y \mapsto \mathrm{I}(x|y)$ *strikt konvex auf* $\{y \in [0;\infty)^k \mid x \equiv y\}$.

(iii) Für festes $y \in [0;\infty)^k$ *ist die Abbildung* $x \mapsto \mathrm{I}(x|y)$ *strikt konvex auf* $\{x \in [0;\infty)^k \mid x \ll y\}$.

(iv) Die Abbildung $(x,y) \mapsto \mathrm{I}(x|y)$ *ist halbstetig von unten in allen* $(x,y) \in [0;\infty)^k \times [0;\infty)^k$.

(v) Für festes $x \in [0;\infty)^k$ *ist die Abbildung* $y \mapsto \mathrm{I}(x|y)$ *stetig in allen* $y \in [0;\infty)^k$.

Für die I-Divergenz zwischen Matrizen gelten die Aussagen (i) bis (v) in analoger Weise.

Beweis. Die Aussagen (i) bis (v) folgen aus der obigen Diskussion der Funktion $\overline{\varphi}$. Die Einschränkung $x \ll y$ in den Aussagen (ii) und (iii) schließt alle Fälle aus, in denen die I-Divergenz $\mathrm{I}(x|y)$ unendlich ist. Die zusätzliche Einschränkung $x \gg y$ in Aussage (ii) schließt alle Fälle aus, in denen sich zwei Vektoren y und $\tilde{y}$ nur außerhalb des Trägers von x unterscheiden. □

Der folgende Satz impliziert mehrere hilfreiche Ungleichungen zur I-Divergenz. Er besagt, dass die I-Divergenz nicht zunimmt, wenn Einträge aggregiert werden.

Satz 2.1.3 (Aggregationsungleichung für Vektoren, vgl. Gietl und Reffel 2013a, Lemma 2.2)**.** *Für alle Vektoren $x, y \in [0;\infty)^k$ gilt die Ungleichung*

$$\mathrm{I}(x|y) \geq \overline{\varphi}(x_+, y_+) = \mathrm{I}(x_+|y_+). \tag{2.1.5}$$

Im Fall $x \ll y$ gilt Gleichheit genau dann, wenn $x_i/y_i = x_+/y_+$ für alle $i \in \operatorname{supp}(y)$ gilt. Im Fall $x \not\ll y$ gilt Gleichheit genau dann, wenn $y_+ = 0$ gilt.

Beweis. Gilt $x_+ = 0$, so sind beide Seiten gleich 0. Gilt $x_+ > 0$ und $x \not\ll y$, so ist die linke Seite gleich ∞. Gilt $x_+ > 0$ und $x \ll y$, so folgt $y_+ > 0$. Mit der Konvexität und der Monotonie der Abbildung $z \mapsto -\ln z$ gilt dann

$$\begin{aligned} \mathrm{I}(x|y) &= \sum_{i:\, x_i > 0} x_i \ln \frac{x_i}{y_i} = x_+ \sum_{i:\, x_i>0} \frac{x_i}{x_+} \left(-\ln \frac{y_i}{x_i}\right) \\ &\geq x_+ \left(-\ln\left(\sum_{i:\, x_i>0} \frac{x_i}{x_+}\frac{y_i}{x_i}\right)\right) \geq x_+ \ln \frac{x_+}{y_+} = \overline{\varphi}(x_+, y_+). \end{aligned} \tag{2.1.6}$$

Die Gleichheitsaussage für den Fall $x \ll y$ folgt aus der strikten Konvexität der Abbildung $z \mapsto -\ln z$. □

Korollar 2.1.4 (Nichtnegativität der I-Divergenz zwischen Vektoren)**.** *Für alle Vektoren $x, y \in [0;\infty)^k$ mit übereinstimmenden Summen $x_+ = y_+$ ist die I-Divergenz $\mathrm{I}(x|y)$ nach unten durch 0 beschränkt,*

$$\mathrm{I}(x|y) \geq 0. \tag{2.1.7}$$

Gleichheit gilt genau dann, wenn $x = y$ gilt.

Aus der in Satz 2.1.3 gezeigten Aggregationsungleichung für Vektoren können die folgenden Aggregationsungleichungen für Matrizen abgeleitet werden.

Satz 2.1.5 (Aggregationsungleichungen für Matrizen, vgl. Gietl und Reffel 2013a, Theorem 2.3)**.** *Für alle Matrizen $S, T \in [0;\infty)^{k\times\ell}$ gelten die folgenden drei Ungleichungen:*

(i) Es gilt die Summenungleichung

$$\mathrm{I}(S|T) \geq \mathrm{I}(s_{++}|t_{++}). \tag{2.1.8}$$

Im Fall $S \ll T$ gilt Gleichheit genau dann, wenn $s_{ij}/t_{ij} = s_{++}/t_{++}$ für alle $(i,j) \in \operatorname{supp}(T)$ gilt. Im Fall $S \not\ll T$ gilt Gleichheit genau dann, wenn $t_{++} = 0$ gilt.

(ii) Es gilt die Zeilensummenungleichung

$$\mathrm{I}(S|T) \geq \mathrm{I}((s_{1+},\ldots,s_{k+})|(t_{1+},\ldots,t_{k+})). \tag{2.1.9}$$

Im Fall $S \ll T$ gilt Gleichheit genau dann, wenn $s_{ij}/t_{ij} = s_{i+}/t_{i+}$ für alle $(i,j) \in \operatorname{supp}(T)$ gilt. Im Fall $S \not\ll T$ gilt Gleichheit genau dann, wenn $(s_{1+},\ldots,s_{k+}) \not\ll (t_{1+},\ldots,t_{k+})$ gilt.

(iii) Es gilt die Spaltensummenungleichung

$$\mathrm{I}(S|T) \geq \mathrm{I}((s_{+1},\ldots,s_{+\ell})|(t_{+1},\ldots,t_{+\ell})). \tag{2.1.10}$$

Im Fall $S \ll T$ gilt Gleichheit genau dann, wenn $s_{ij}/t_{ij} = s_{+j}/t_{+j}$ für alle $(i,j) \in \operatorname{supp}(T)$ gilt. Im Fall $S \not\ll T$ gilt Gleichheit genau dann, wenn $(s_{+1},\ldots,s_{+\ell}) \not\ll (t_{+1},\ldots,t_{+\ell})$ gilt.

Beweis. (i) Die Summenungleichung und die dazugehörige Gleichheitsaussage folgen unmittelbar aus Satz 2.1.3.

(ii) Nach Definition 2.1.1 ist die I-Divergenz $\mathrm{I}(S|T)$ gleich der Summe aller I-Divergenzen $\mathrm{I}((s_{i1},\ldots,s_{i\ell})|(t_{i1},\ldots,t_{i\ell}))$ der einzelnen Zeilen i. Die einzelnen Summanden können gemäß Satz 2.1.3 jeweils nach unten durch $\mathrm{I}(s_{i+}|t_{i+})$ abgeschätzt werden. Insgesamt erhalten wir also

$$\begin{aligned} \mathrm{I}(S|T) &= \sum_i \sum_j \overline{\varphi}(s_{ij},t_{ij}) = \sum_i \mathrm{I}((s_{i1},\ldots,s_{i\ell})|(t_{i1},\ldots,t_{i\ell})) \\ &\geq \sum_i \mathrm{I}(s_{i+}|t_{i+}) = \mathrm{I}((s_{1+},\ldots,s_{k+})|(t_{1+},\ldots,t_{k+})). \end{aligned} \tag{2.1.11}$$

Gleichheit gilt genau dann, wenn in allen Abschätzungen Gleichheit gilt.

(iii) Die Spaltensummenungleichung und die dazugehörige Gleichheitsaussage folgen in analoger Weise. □

Korollar 2.1.6 (Nichtnegativität der I-Divergenz zwischen Matrizen)**.** *Für alle Matrizen $S, T \in [0;\infty)^{k\times\ell}$ mit übereinstimmenden Summen $s_{++} = t_{++}$ ist die I-Divergenz $\mathrm{I}(S|T)$ nach unten durch 0 beschränkt,*

$$\mathrm{I}(S|T) \geq 0. \tag{2.1.12}$$

Gleichheit gilt genau dann, wenn $S = T$ gilt.

Für Wahrscheinlichkeitsvektoren, das heißt für nichtnegative Vektoren p mit $p_+ = 1$, und für Wahrscheinlichkeitsmatrizen, das heißt für nichtnegative Matrizen P mit $p_{++} = 1$, kann eine weitere untere Schranke für die I-Divergenz gezeigt werden.

Satz 2.1.7 (Pinsker-Ungleichung, vgl. Cover und Thomas 2006, Lemma 11.6.1)**.** *Für alle Wahrscheinlichkeitsvektoren $p, q \in [0;1]^k$ gilt die Ungleichung*

$$\mathrm{I}(p|q) \geq \frac{1}{2}\,\|p-q\|_1^2, \tag{2.1.13}$$

wobei $\|p-q\|_1 := \sum_i |p_i - q_i|$ den L_1-Abstand von p und q bezeichnet. Ungleichung (2.1.13) *gilt in analoger Weise für die I-Divergenz und den L_1-Abstand zwischen Matrizen.*

Beweis. Gilt $p = q$, so sind beide Seiten gleich 0. Gilt $p \neq q$ und $p \not\ll q$, so ist die linke Seite gleich ∞. Gilt $p \neq q$ und $p \ll q$, so wählen wir wählen wir $I := \{i \,|\, p_i < q_i\}$. Dann gilt $\emptyset \subsetneq I \subsetneq \{1, \ldots, k\}$ und $0 \leq p_I < q_I < 1$. Den Vektor p zerlegen wir dementsprechend in die Komponenten $p^{(I)}$ und $p^{(I')}$ und den Vektor q in die Komponenten $q^{(I)}$ und $q^{(I')}$. Folglich gilt nach Definition 2.1.1 und Satz 2.1.3 die Ungleichung

$$\mathrm{I}(p|q) = \mathrm{I}(p^{(I)}|q^{(I)}) + \mathrm{I}(p^{(I')}|q^{(I')}) \geq \overline{\varphi}(p_I, q_I) + \overline{\varphi}(p_{I'}, q_{I'}). \tag{2.1.14}$$

Außerdem gilt

$$\|p - q\|_1 = \sum_{i \in I} |p_i - q_i| + \sum_{i \in I'} |p_i - q_i| = p_I - q_I - p_{I'} + q_{I'} = 2(p_I - q_I). \tag{2.1.15}$$

Daraus folgt

$$\mathrm{I}(p|q) - \frac{1}{2}\, \|p - q\|_1^2 \geq \overline{\varphi}(p_I, q_I) + \overline{\varphi}(p_{I'}, q_{I'}) - 2(p_I - q_I)^2. \tag{2.1.16}$$

Der verbleibende Beweis unterscheidet zwei Fälle.

I. Im Fall $0 = p_I < q_I < 1$ gilt

$$\mathrm{I}(p|q) - \frac{1}{2}\, \|p - q\|_1^2 \geq \overline{\varphi}(p_I, q_I) + \overline{\varphi}(p_{I'}, q_{I'}) - 2(p_I - q_I)^2 = -\ln(1 - q_I) - 2q_I^2. \tag{2.1.17}$$

Für $q_I = 0$ nimmt die rechte Seite den Wert 0 an. Da für uns jedoch nur die Werte $q_I \in (0; 1)$ interessant sind, betrachten wir die Ableitung der rechten Seite. Für diese Ableitung gilt

$$\frac{\partial}{\partial q_I} \left(-\ln(1 - q_I) - 2q_I^2 \right) = \frac{1}{1 - q_I} - 4q_I = \frac{1 - 4q_I(1 - q_I)}{1 - q_I} \geq 0. \tag{2.1.18}$$

Folglich ist die rechte Seite stets nichtnegativ und es gilt

$$\mathrm{I}(p|q) \geq \frac{1}{2}\, \|p - q\|_1^2\,. \tag{2.1.19}$$

II. Im Fall $0 < p_I < q_I < 1$ gilt

$$\begin{aligned} \mathrm{I}(p|q) - \frac{1}{2}\, \|p - q\|_1^2 &\geq \overline{\varphi}(p_I, q_I) + \overline{\varphi}(p_{I'}, q_{I'}) - 2(p_I - q_I)^2 \\ &= p_I \ln \frac{p_I}{q_I} + (1 - p_I) \ln \frac{1 - p_I}{1 - q_I} - 2(p_I - q_I)^2. \end{aligned} \tag{2.1.20}$$

Für $q_I = p_I$ nimmt die rechte Seite den Wert 0 an. Da für uns jedoch nur die Werte $q_I \in (p_I; 1)$ interessant sind, betrachten wir die Ableitung der rechten Seite nach q_I. Für diese Ableitung gilt

$$\begin{aligned} &\frac{\partial}{\partial q_I} \left(p_I \ln \frac{p_I}{q_I} + (1 - p_I) \ln \frac{1 - p_I}{1 - q_I} - 2(p_I - q_I)^2 \right) \\ &\qquad = -\frac{p_I}{q_I} + \frac{1 - p_I}{1 - q_I} + 4(p_I - q_I) \geq 0. \end{aligned} \tag{2.1.21}$$

Folglich ist die rechte Seite stets nichtnegativ und es gilt

$$\mathrm{I}(p|q) \geq \frac{1}{2}\, \|p - q\|_1^2\,. \tag{2.1.22}$$

Damit ist die Ungleichung (2.1.13) gezeigt. □

Die soeben gezeigte Pinsker-Ungleichung erlaubt den Schluss von der Konvergenzaussage $\mathrm{I}(p^n|q) \to 0$ für $n \to \infty$ auf die Konvergenzaussage $\|p^n - q\|_1 \to 0$ und somit auf die Konvergenz der Vektorenfolge (p^n) gegen den Vektor q. Diese Eigenschaft spielt eine entscheidende Rolle bei der informationstheoretischen Analyse des IPF-Verfahrens.

2.2. I-Projektionen

Im Mittelpunkt der informationstheoretischen Analyse des IPF-Verfahrens steht die Minimierung der I-Divergenz. Diese Minimierung kann auf zwei Arten erfolgen. Bei den I_1-Projektionen wird die I-Divergenz im ersten Argument minimiert.

Definition 2.2.1 (I_1-Projektion, vgl. Csiszár 1975, S. 147)**.** Sei $\mathcal{E}$ eine beliebige Teilmenge von $[0;\infty)^k$ und sei $y \in [0;\infty)^k$ ein beliebiger Vektor. Dann heißt $x \in \mathcal{E}$ eine *I_1-Projektion von y auf $\mathcal{E}$*, wenn gilt

$$\mathrm{I}(x|y) = \inf_{\tilde{x}\in\mathcal{E}} \mathrm{I}(\tilde{x}|y) < \infty. \tag{2.2.1}$$

Für Matrizen ist die I_1-Projektion in analoger Weise definiert.

Im Folgenden werden einige Eigenschaften von I_1-Projektionen von Vektoren diskutiert. Diese Eigenschaften gelten in analoger Weise für I_1-Projektionen von Matrizen. Für die Existenz einer I_1-Projektion auf eine kompakte Menge gilt das folgende notwendige und hinreichende Kriterium.

Satz 2.2.2 (Existenz von I_1-Projektionen auf kompakte Mengen)**.** *Sei $\mathcal{E}$ eine kompakte Teilmenge von $[0;\infty)^k$ und sei $y \in [0;\infty)^k$ ein beliebiger Vektor. Dann existiert eine I_1-Projektion von y auf $\mathcal{E}$ genau dann, wenn ein Vektor $\tilde{x} \in \mathcal{E}$ mit $\tilde{x} \ll y$ existiert.*

Beweis. Nach Voraussetzung ist die Menge $\mathcal{E}$ kompakt. Da die Abbildung $\tilde{x} \mapsto \mathrm{I}(\tilde{x}|y)$ halbstetig von unten ist (siehe Satz 2.1.2 (iv)), wird das Infimum in Gleichung (2.2.1) angenommen (vgl. Bauschke und Combettes 2011, Theorem 1.28). Dieses Infimum ist genau dann endlich, wenn ein Vektor $\tilde{x} \in \mathcal{E}$ existiert, sodass $\mathrm{I}(\tilde{x}|y) < \infty$ gilt. Dies ist genau dann der Fall, wenn $\tilde{x} \ll y$ gilt. □

Für konvexe Mengen ist die Eindeutigkeit der I_1-Projektion sichergestellt, wie der folgende Satz zeigt.

Satz 2.2.3 (Eindeutigkeit von I_1-Projektionen auf konvexe Mengen, vgl. Csiszár 1975, S. 147)**.** *Sei $\mathcal{E}$ eine konvexe Teilmenge von $[0;\infty)^k$ und sei $y \in [0;\infty)^k$ ein beliebiger Vektor. Dann ist die I_1-Projektion von y auf $\mathcal{E}$ eindeutig.*

Beweis. Nach Voraussetzung ist die Menge $\mathcal{E}$ konvex. Da die Abbildung $\tilde{x} \mapsto \mathrm{I}(\tilde{x}|y)$ strikt konvex in allen $\tilde{x} \in [0;\infty)^k$ mit $\tilde{x} \ll y$ ist (siehe Satz 2.1.2 (iii)), ist die I_1-Projektion von y auf $\mathcal{E}$ eindeutig. □

Wie die beiden vorangegangenen Sätze zeigen, existiert für kompakte und konvexe Mengen stets eine eindeutige I_1-Projektion, sofern diese Mengen einen Vektor enthalten, der den Ausgangsvektor dominiert. In diesem Fall kann außerdem gezeigt werden, dass die I_1-Projektion stetig ist.

Satz 2.2.4 (Stetigkeit von I_1-Projektionen auf kompakte und konvexe Mengen, vgl. Gietl und Reffel 2013c, Theorem 2)**.** *Sei $\mathcal{E}$ eine kompakte und konvexe Teilmenge von $[0;\infty)^k$. Weiter sei (y^n) eine konvergente Folge von Vektoren in $[0;\infty)^k$ mit dem Grenzwert $y \in [0;\infty)^k$. Außerdem seien x^n die I_1-Projektionen von y^n auf $\mathcal{E}$ für alle $n \in \mathbb{N}$ und x die I_1-Projektion von y auf $\mathcal{E}$. Dann konvergiert die Folge (x^n) gegen x,*

$$\lim_{n\to\infty} x^n = x. \tag{2.2.2}$$

Beweis. Sei $x^* := \lim_{m\to\infty} x^{n_m} \in \mathcal{E}$ ein beliebiger Häufungspunkt der Folge (x^n). Dann gilt mit der unteren Halbstetigkeit der I-Divergenz (siehe Satz 2.1.2 (iv))

$$\mathrm{I}(x^*|y) = \mathrm{I}(\lim_{m\to\infty} x^{n_m} | \lim_{m\to\infty} y^{n_m}) \leq \liminf_{m\to\infty} \mathrm{I}(x^{n_m}|y^{n_m}). \tag{2.2.3}$$

Da der Vektor x^{n_m} nach Voraussetzung die I_1-Projektion von y^{n_m} auf $\mathcal{E}$ ist, gilt außerdem $\mathrm{I}(x^{n_m}|y^{n_m}) \leq \mathrm{I}(x|y^{n_m})$ für alle $m \in \mathbb{N}$ und somit

$$\liminf_{m\to\infty} \mathrm{I}(x^{n_m}|y^{n_m}) \leq \liminf_{m\to\infty} \mathrm{I}(x|y^{n_m}). \tag{2.2.4}$$

Mit der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v)) folgt

$$\liminf_{m\to\infty} \mathrm{I}(x|y^{n_m}) = \mathrm{I}(x|\lim_{m\to\infty} y^{n_m}) = \mathrm{I}(x|y). \tag{2.2.5}$$

Damit gilt $\mathrm{I}(x^*|y) \leq \mathrm{I}(x|y)$. Da der Vektor x nach Voraussetzung die I_1-Projektion von y auf $\mathcal{E}$ ist, folgt $\mathrm{I}(x^*|y) = \mathrm{I}(x|y)$ und $x^* = x$. Also konvergiert die Folge (x^n) gegen x. □

Während bei den I_1-Projektionen die I-Divergenz im ersten Argument minimiert wird, wird bei den I_2-Projektionen die I-Divergenz im zweiten Argument minimiert.

Definition 2.2.5 (I_2-Projektion)**.** Sei $x \in [0;\infty)^k$ ein beliebiger Vektor und sei $\mathcal{F}$ eine beliebige Teilmenge von $[0;\infty)^k$. Dann heißt $y \in \mathcal{F}$ eine *I_2-Projektion von x auf $\mathcal{F}$*, wenn gilt

$$\mathrm{I}(x|y) = \inf_{\tilde{y}\in\mathcal{F}} \mathrm{I}(x|\tilde{y}) < \infty. \tag{2.2.6}$$

Für Matrizen ist die I_2-Projektion in analoger Weise definiert.

Wie der folgende Satz zeigt, sind die Bedingungen für die Existenz einer I_2-Projektion symmetrisch zu den Bedingungen für die Existenz einer I_1-Projektion. Die Bedingungen gelten in analoger Weise für I_2-Projektionen von Matrizen.

Satz 2.2.6 (Existenz von I_2-Projektionen auf kompakte Mengen)**.** *Sei $x \in [0;\infty)^k$ ein beliebiger Vektor und sei $\mathcal{F}$ eine beliebige Teilmenge von $[0;\infty)^k$. Dann existiert eine I_2-Projektion von x auf $\mathcal{F}$ genau dann, wenn ein Vektor $\tilde{y} \in \mathcal{F}$ mit $x \ll \tilde{y}$ existiert.*

Beweis. Nach Voraussetzung ist die Menge $\mathcal{F}$ kompakt. Da die Abbildung $\tilde{y} \mapsto \mathrm{I}(x|\tilde{y})$ halbstetig von unten ist (siehe Satz 2.1.2 (iv)), wird das Infimum in Gleichung (2.2.6) angenommen (vgl. Bauschke und Combettes 2011, Theorem 1.28). Dieses Infimum ist genau dann endlich, wenn ein Vektor $\tilde{y} \in \mathcal{F}$ existiert, sodass $\mathrm{I}(x|\tilde{y}) < \infty$ gilt. Dies ist genau dann der Fall, wenn $x \ll \tilde{y}$ gilt. □

Im Gegensatz zu den Existenzbedingungen sind die Bedingungen für die Eindeutigkeit von I_2-Projektionen nicht symmetrisch zu den Bedingungen für die Eindeutigkeit von I_1-Projektionen. Wie Matúš (1997, Abschnitt 4) zeigt, treten stattdessen bei den I_2-Projektionen sogenannte log-konvexe Mengen an die Stelle konvexer Mengen. Die in diesem Abschnitt gezeigte Stetigkeit von I_1-Projektionen kann ebenfalls nicht auf I_2-Projektionen übertragen werden, da die I-Divergenz im Allgemeinen nur im zweiten Argument stetig ist.

2.3. Informationstheoretische Konvergenzanalyse

Im Folgenden wird die I-Divergenz zur Analyse der Konvergenz des IPF-Verfahrens angewendet. Durch Skalierung passt jeder Schritt des IPF-Verfahrens entweder die Zeilen an, um die Zeilenmarginalien zu erfüllen, oder die Spalten, um die Spaltenmarginalien zu erfüllen. Mithilfe der I-Divergenz wird untersucht, wie sich die angepassten Matrizen zu allen anderen Matrizen verhalten, die die Nulleinträge der Ausgangsmatrix übernehmen und die jeweiligen Marginalien erfüllen.

Sei daher $\mathcal{C}$ die Menge aller Matrizen, die die Spaltenmarginalien c erfüllen, und sei $\mathcal{R}$ die Menge aller Matrizen, die die Zeilenmarginalien r erfüllen,

$$\mathcal{C} := \left\{ C \in [0;\infty)^{k\times\ell} \,\middle|\, \forall j : c_{+j} = c_j \right\}, \tag{2.3.1}$$

$$\mathcal{R} := \left\{ R \in [0;\infty)^{k\times\ell} \,\middle|\, \forall i : r_{i+} = r_i \right\}. \tag{2.3.2}$$

Außerdem sei $\mathcal{E}$ die Schnittmenge von $\mathcal{C}$ und $\mathcal{R}$,

$$\mathcal{E} := \mathcal{C} \cap \mathcal{R}. \tag{2.3.3}$$

Des Weiteren bezeichne $\mathcal{C}^{\ll A}$ die Menge aller Matrizen in $\mathcal{C}$, die von der Ausgangsmatrix A dominiert werden, und $\mathcal{R}^{\ll A}$ die Menge aller Matrizen in $\mathcal{R}$, die von der Ausgangsmatrix A dominiert werden,

$$\mathcal{C}^{\ll A} := \{ C \in \mathcal{C} \,|\, C \ll A \}, \tag{2.3.4}$$

$$\mathcal{R}^{\ll A} := \{ R \in \mathcal{R} \,|\, R \ll A \}. \tag{2.3.5}$$

Schließlich sei $\mathcal{E}^{\ll A}$ die Schnittmenge von $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$,

$$\mathcal{E}^{\ll A} := \mathcal{C}^{\ll A} \cap \mathcal{R}^{\ll A}. \tag{2.3.6}$$

Die so definierten Mengen $\mathcal{C}$, $\mathcal{R}$ und $\mathcal{E}$ sowie $\mathcal{C}^{\ll A}$, $\mathcal{R}^{\ll A}$ und $\mathcal{E}^{\ll A}$ sind kompakt und konvex. Außerdem ist die ungeradzahlige IPF-Teilfolge $(A(t+1))_{t=0,2,4,\ldots}$ in $\mathcal{R}^{\ll A} \subseteq \mathcal{R}$ und die geradzahlige IPF-Teilfolge $(A(t+2))_{t=0,2,4,\ldots}$ in $\mathcal{C}^{\ll A} \subseteq \mathcal{C}$ enthalten. Konvergiert das IPF-Verfahren, so ist die Limesmatrix $B^* := \lim_{t\to\infty} A(t)$ in $\mathcal{E}^{\ll A} \subseteq \mathcal{E}$ enthalten. Die Bedingung $\mathcal{E}^{\ll A} \neq \emptyset$ ist also notwendig für die Konvergenz des IPF-Verfahrens.

Für die weiteren Untersuchungen zur Konvergenz des IPF-Verfahrens legen die in den Korollaren 2.1.4 und 2.1.6 gezeigten Nichtnegativitätseigenschaften sowie der Satz 1.5.1 (siehe S. 20) eine Beschränkung auf normierte biproportionale Anpassungsprobleme (A, c, r) nahe. Diese Anpassungsprobleme bestehen aus einer Wahrscheinlichkeitsmatrix $A \in [0;1]^{k\times\ell}$ mit $a_{++} = 1$ und zwei Wahrscheinlichkeitsvektoren $c \in (0;1)^\ell$ und $r \in (0;1)^k$ mit $c_+ = r_+ = 1$.

Lemma 2.3.1 (Zeilen- und Spaltenanpassungen, vgl. Cramer 2000, Theoreme 3.4 (i) und 3.8 (i))**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem, $(A(t))$ die zugehörige IPF-Folge und $\mathcal{C}^{\ll A}$ sowie $\mathcal{R}^{\ll A}$ die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen. Dann gelten für alle geraden Schritte $t \geq 0$ die folgenden vier Aussagen:*

(i) Für alle Matrizen $R \in \mathcal{R}^{\ll A}$ gilt

$$\mathrm{I}(R|A(t)) = \mathrm{I}(R|A(t+1)) + \mathrm{I}(A(t+1)|A(t)) \text{ und} \tag{2.3.7}$$

$$\mathrm{I}(A(t)|R) \geq \mathrm{I}(A(t)|A(t+1)). \tag{2.3.8}$$

Gleichheit gilt in Ungleichung (2.3.8) *genau dann, wenn $R = A(t+1)$ gilt.*

(ii) Die Zeilenanpassung $A(t+1)$ ist die I_1-Projektion und die I_2-Projektion von $A(t)$ auf $\mathcal{R}^{\ll A}$.

(iii) Für alle Matrizen $C \in \mathcal{C}^{\ll A}$ gilt

$$\mathrm{I}(C|A(t+1)) = \mathrm{I}(C|A(t+2)) + \mathrm{I}(A(t+2)|A(t+1)) \text{ und} \tag{2.3.9}$$

$$\mathrm{I}(A(t+1)|C) \geq \mathrm{I}(A(t+1)|A(t+2)). \tag{2.3.10}$$

Gleichheit gilt in Ungleichung (2.3.10) *genau dann, wenn $C = A(t+2)$ gilt.*

(iv) Die Spaltenanpassung $A(t+2)$ ist die I_1-Projektion und die I_2-Projektion von $A(t+1)$ auf $\mathcal{C}^{\ll A}$.

Beweis. Im Folgenden werden nur die Aussagen (i) und (ii) gezeigt. Der Beweis der Aussagen (iii) und (iv) erfolgt in analoger Weise.

(i) Gemäß Definition der Menge $\mathcal{R}^{\ll A}$ wird die Matrix R von der Ausgangsmatrix A dominiert. Die Matrizen $A(t)$ und $A(t+1)$ sind, wie in Abschnitt 1.1 (siehe S. 12) gezeigt, maßtheoretisch äquivalent zu A. Folglich gilt $R \ll A(t) \equiv A(t+1)$. Daraus folgt nach Definition des IPF-Verfahrens

$$\begin{aligned}
\mathrm{I}(R|A(t)) &= \sum_{i,j:\, r_{ij}>0} r_{ij} \ln \frac{r_{ij}}{a_{ij}(t)} \\
&= \sum_{i,j:\, r_{ij}>0} r_{ij} \ln \frac{r_{ij}}{a_{ij}(t+1)} + \sum_{i,j:\, r_{ij}>0} r_{ij} \ln \frac{r_i}{a_{i+}(t)} \\
&= \mathrm{I}(R|A(t+1)) + \sum_i r_i \ln \frac{r_i}{a_{i+}(t)} \\
&= \mathrm{I}(R|A(t+1)) + \sum_{i,j:\, a_{ij}(t+1)>0} a_{ij}(t+1) \ln \frac{a_{ij}(t+1)}{a_{ij}(t)} \\
&= \mathrm{I}(R|A(t+1)) + \mathrm{I}(A(t+1)|A(t)).
\end{aligned} \tag{2.3.11}$$

Gemäß der Zeilensummenungleichung (siehe Satz 2.1.5 (ii)) und der Definition des IPF-Verfahrens gilt

$$\begin{aligned}\mathrm{I}(A(t)|R) \geq \mathrm{I}((a_{1+}(t),\ldots,a_{k+}(t))|r) &= \sum_i a_{i+}(t)\ln\frac{a_{i+}(t)}{r_i}\\ &= \sum_{i,j\colon a_{ij}(t)>0} a_{ij}(t)\ln\frac{a_{ij}(t)}{a_{ij}(t+1)} = \mathrm{I}(A(t)|A(t+1)).\end{aligned} \tag{2.3.12}$$

(ii) Nach Aussage (i) gilt für alle $R \in \mathcal{R}^{\ll A}$ die Gleichung (2.3.7). Daraus folgt mit der Nichtnegativität der I-Divergenz (siehe Korollar 2.1.6) für alle $R \in \mathcal{R}^{\ll A}$ die Ungleichung $\mathrm{I}(R|A(t)) \geq \mathrm{I}(A(t+1)|A(t))$. Folglich ist $A(t+1)$ die I_1-Projektion von $A(t)$ auf $\mathcal{R}^{\ll A}$.

Nach Aussage (i) gilt außerdem für alle $R \in \mathcal{R}^{\ll A}$ die Ungleichung (2.3.8). Folglich ist $A(t+1)$ die I_2-Projektion von $A(t)$ auf $\mathcal{R}^{\ll A}$. □

Mithilfe der im vorangegangenen Lemma 2.3.1 gezeigten Gleichungen (2.3.7) und (2.3.9) wird im Folgenden bewiesen, dass die Bedingung $\mathcal{E}^{\ll A} \neq \emptyset$ nicht nur notwendig sondern auch hinreichend ist für die Konvergenz des IPF-Verfahrens.

Satz 2.3.2 (Konvergenz des zweidimensionalen IPF-Verfahrens, vgl. Brown, Chase und Pittenger 1993, Theorem 3.1)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem, $(A(t))$ die zugehörige IPF-Folge und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Wahrscheinlichkeitsmatrizen. Dann gilt:*

(i) Die IPF-Folge $(A(t))$ konvergiert genau dann, wenn die Menge $\mathcal{E}^{\ll A}$ nichtleer ist.

Im Konvergenzfall gelten für die Limesmatrix $B^* := \lim_{t\to\infty} A(t)$ die folgenden vier Aussagen:

(ii) Für alle $P \in \mathcal{E}^{\ll A}$ gilt $P \ll B^* \ll A$.

(iii) Für alle $P \in \mathcal{E}^{\ll A}$ gilt die Gleichung

$$\mathrm{I}(P|A) = \mathrm{I}(P|B^*) + \mathrm{I}(B^*|A). \tag{2.3.13}$$

(iv) Die Limesmatrix B^* ist die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$.

(v) Es gilt $B^* \equiv A$ genau dann, wenn eine Matrix $P \in \mathcal{E}^{\ll A}$ existiert, für die $P \equiv A$ gilt.

Beweis. (i) Sei zunächst die IPF-Folge $(A(t))$ konvergent und sei $B^* := \lim_{t\to\infty} A(t)$ ihre Limesmatrix. Dann ist B^* in $\mathcal{E}^{\ll A}$ enthalten.

Sei nun die Menge $\mathcal{E}^{\ll A} = \mathcal{C}^{\ll A} \cap \mathcal{R}^{\ll A}$ nichtleer. Aus den in Lemma 2.3.1 gezeigten Gleichungen (2.3.7) und (2.3.9) folgt

$$\forall P \in \mathcal{E}^{\ll A}\ \forall t \geq 0: \quad \mathrm{I}(P|A(t)) = \mathrm{I}(P|A(t+1)) + \mathrm{I}(A(t+1)|A(t)). \tag{2.3.14}$$

Wiederholte Anwendung dieser Aussage liefert

$$\forall P \in \mathcal{E}^{\ll A}\ \forall t \geq 0: \quad \mathrm{I}(P|A) = \mathrm{I}(P|A(t+1)) + \sum_{\tau=0}^{t} \mathrm{I}(A(\tau+1)|A(\tau)). \tag{2.3.15}$$

Sei $P \in \mathcal{E}^{\ll A}$ beliebig. Nach Definition der Menge $\mathcal{E}^{\ll A}$ wird P von der Ausgangsmatrix A dominiert und es gilt $\mathrm{I}(P|A) < \infty$ und somit

$$\mathrm{I}(A(t+1)|A(t)) \to 0 \text{ für } t \to \infty. \tag{2.3.16}$$

Mit der Pinsker-Ungleichung (siehe Satz 2.1.7) folgt

$$\|A(t+1) - A(t)\|_1 \to 0 \text{ für } t \to \infty. \tag{2.3.17}$$

Da die IPF-Folge $(A(t))$ in der kompakten Menge $[0;1]^{k\times\ell}$ enthalten ist, besitzt sie mindestens einen Häufungspunkt. Sei B^* ein solcher Häufungspunkt. Dann existiert eine Schrittfolge (t_n), sodass gilt

$$A(t_n) \to B^* \text{ für } n \to \infty. \tag{2.3.18}$$

Mit der Konvergenzaussage (2.3.17) folgt daraus

$$A(t_n + 1) \to B^* \text{ für } n \to \infty. \tag{2.3.19}$$

Mindestens eine der beiden Schrittfolgen (t_n) und $(t_n + 1)$ enthält unendlich viele gerade Schritte. Folglich gilt $B^* \in \mathcal{C}^{\ll A}$. Analog gilt $B^* \in \mathcal{R}^{\ll A}$ und somit $B^* \in \mathcal{E}^{\ll A}$.

Nach Gleichung (2.3.14) gilt damit

$$\forall t \geq 0: \quad \mathrm{I}(B^*|A(t)) = \mathrm{I}(B^*|A(t+1)) + \mathrm{I}(A(t+1)|A(t)). \tag{2.3.20}$$

Somit ist die Folge $(\mathrm{I}(B^*|A(t)))$ monoton fallend. Mit der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v)) folgt

$$\lim_{t\to\infty} \mathrm{I}(B^*|A(t)) = \lim_{n\to\infty} \mathrm{I}(B^*|A(t_n)) = \mathrm{I}(B^*|B^*) = 0. \tag{2.3.21}$$

Daraus folgt mit der Pinsker-Ungleichung (siehe Satz 2.1.7) die Konvergenz der IPF-Folge $(A(t))$ gegen B^*.

(ii) Sei die IPF-Folge $(A(t))$ konvergent und sei $B^* := \lim_{t\to\infty} A(t) \in \mathcal{E}^{\ll A}$ ihre Limesmatrix. Weiter sei $P \in \mathcal{E}^{\ll A}$ beliebig. Dann gilt gemäß Definition der Menge $\mathcal{E}^{\ll A}$ die Ungleichung $\mathrm{I}(P|A(0)) = \mathrm{I}(P|A) < \infty$. Nach Gleichung (2.3.14) ist die Folge $(\mathrm{I}(P|A(t)))$ außerdem monoton fallend. Mit der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v)) folgt daraus

$$\mathrm{I}(P|B^*) = \lim_{t\to\infty} \mathrm{I}(P|A(t)) \leq \mathrm{I}(P|A(0)) < \infty. \tag{2.3.22}$$

Dies impliziert die Aussage $P \ll B^* \ll A$.

(iii) Die Aussage $P \ll B^* \ll A$ erlaubt die Rechnung

$$
\begin{aligned}
&\mathrm{I}(P|B^*) + \mathrm{I}(B^*|A) \\
&\qquad = \sum_{i,j:\, p_{ij}>0} p_{ij} \ln \frac{p_{ij}}{b^*_{ij}} + \sum_{i,j:\, b^*_{ij}>0} b^*_{ij} \ln \frac{b^*_{ij}}{a_{ij}} \\
&\qquad = \sum_{i,j:\, p_{ij}>0} p_{ij} \ln \frac{p_{ij}}{a_{ij}} + \sum_{i,j:\, p_{ij}>0} p_{ij} \ln \frac{a_{ij}}{b^*_{ij}} + \sum_{i,j:\, b^*_{ij}>0} b^*_{ij} \ln \frac{b^*_{ij}}{a_{ij}} \\
&\qquad = \mathrm{I}(P|A) + \sum_{i,j:\, b^*_{ij}>0} \left(b^*_{ij} - p_{ij}\right) \ln \frac{b^*_{ij}}{a_{ij}}. \qquad (2.3.23)
\end{aligned}
$$

Da B^* die Limesmatrix der IPF-Folge $(A(t))$ ist, gilt

$$
\begin{aligned}
\sum_{i,j:\, b^*_{ij}>0} \left(b^*_{ij} - p_{ij}\right) \ln \frac{b^*_{ij}}{a_{ij}} &= \lim_{t\to\infty} \sum_{i,j:\, b^*_{ij}>0} \left(b^*_{ij} - p_{ij}\right) \ln \frac{a_{ij}(t)}{a_{ij}} \\
&= \lim_{t\to\infty} \sum_{\tau=0}^{t-1} \sum_{i,j:\, b^*_{ij}>0} \left(b^*_{ij} - p_{ij}\right) \ln \frac{a_{ij}(\tau+1)}{a_{ij}(\tau)}. \qquad (2.3.24)
\end{aligned}
$$

Nach Definition des IPF-Verfahrens ist der Term $\ln(a_{ij}(\tau+1)/a_{ij}(\tau))$ für alle geraden Schritte $\tau \geq 0$ innerhalb der Zeilen i und für alle ungeraden Schritte $\tau \geq 1$ innerhalb der Spalten j konstant. Da die Matrizen B^* und P jedoch beide in $\mathcal{E}^{\ll A}$ enthalten sind, gilt $b^*_{i+} = p_{i+}$ für alle Zeilen i sowie $b^*_{+j} = p_{+j}$ für alle Spalten j und damit

$$
\forall \tau \geq 0: \quad \sum_{i,j:\, b^*_{ij}>0} \left(b^*_{ij} - p_{ij}\right) \ln \frac{a_{ij}(\tau+1)}{a_{ij}(\tau)} = 0. \qquad (2.3.25)
$$

Daraus folgt Gleichung (2.3.13).

(iv) Nach Aussage (iii) gilt für alle $P \in \mathcal{E}^{\ll A}$ die Gleichung (2.3.13). Daraus folgt mit der Nichtnegativität der I-Divergenz (siehe Korollar 2.1.6) für alle $P \in \mathcal{E}^{\ll A}$ die Ungleichung $\mathrm{I}(P|A) \geq \mathrm{I}(B^*|A)$. Folglich ist B^* die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$.

(v) Sei die IPF-Folge $(A(t))$ konvergent und sei $B^* := \lim_{t\to\infty} A(t) \in \mathcal{E}^{\ll A}$ ihr Grenzwert. Gilt $B^* \equiv A$, so erfüllt B^* die Voraussetzungen an P. Existiert eine Matrix $P \in \mathcal{E}^{\ll A}$ mit $P \equiv A$, so folgt mit Aussage (ii) die Aussage $P \equiv B^* \equiv A$. □

Die Limesmatrix der IPF-Folge $(A(t))$ ist die biproportionale Limesanpassung der Ausgangsmatrix A an die Marginalien c und r (siehe Abschnitt 1.2, S. 13). Somit folgt aus der Konvergenz des IPF-Verfahrens die Existenz der biproportionalen Limesanpassung. Umgekehrt ist die biproportionale Limesanpassung stets in der Menge $\mathcal{E}^{\ll A}$ enthalten. Also impliziert die Existenz der biproportionalen Limesanpassung auch die Konvergenz des IPF-Verfahrens. Die Existenz biproportionaler Limes- und Direktanpassungen kann damit wie folgt charakterisiert werden.

Satz 2.3.3 (Existenz biproportionaler Limes- und Direktanpassungen)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Wahrscheinlichkeitsmatrizen. Dann gelten die folgenden beiden Aussagen:*

(i) Die biproportionale Limesanpassung B von A an c und r existiert genau dann, wenn die Menge $\mathcal{E}^{\ll A}$ nichtleer ist.

(ii) Die biproportionale Direktanpassung B von A an c und r existiert genau dann, wenn eine Matrix $P \in \mathcal{E}^{\ll A}$ existiert, für die $P \equiv A$ gilt.

Beweis. (i) Existiert die biproportionale Limesanpassung B von A an c und r, so ist diese in der Menge $\mathcal{E}^{\ll A}$ enthalten. Ist die Menge $\mathcal{E}^{\ll A}$ nichtleer, so konvergiert nach Satz 2.3.2 (i) die IPF-Folge $(A(t))$ und ihre Limesmatrix $B := \lim_{t\to\infty} A(t)$ ist die biproportionale Limesanpassung von A an c und r.

(ii) Existiert die biproportionale Direktanpassung B von A an c und r, so ist diese in der Menge $\mathcal{E}^{\ll A}$ enthalten und es gilt $B \equiv A$. Existiert eine Matrix $P \in \mathcal{E}^{\ll A}$, für die $P \equiv A$ gilt, so konvergiert nach Satz 2.3.2 (i) die IPF-Folge $(A(t))$ und ihr Grenzwert $B := \lim_{t\to\infty} A(t)$ ist die biproportionale Limesanpassung von A an c und r. Außerdem gilt mit Satz 2.3.2 (ii) die Aussage $B \equiv A$. Gemäß Satz 1.3.3 (siehe S. 16) ist B damit die biproportionale Direktanpassung von A an c und r. □

Der folgende Satz zeigt, dass biproportionale Limesanpassungen an c und r und I_1-Projektionen auf $\mathcal{E}^{\ll A}$ äquivalent sind.

Satz 2.3.4 (Äquivalenz von biproportionalen Limesanpassungen und I_1-Projektionen)**.** *Sei (A, c, r) ein normiertes biproportionales Limesanpassungsproblem und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Wahrscheinlichkeitsmatrizen. Dann gelten die folgenden beiden Aussagen:*

(i) Die biproportionale Limesanpassung von A an c und r existiert genau dann, wenn die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$ existiert.

(ii) Im Existenzfall ist die biproportionale Limesanpassung von A an c und r gleich der I_1-Projektion von A auf $\mathcal{E}^{\ll A}$.

Beweis. (i) Sei B die biproportionale Limesanpassung von A an c und r. Dann gilt $B \in \mathcal{E}^{\ll A}$. Daraus folgt nach Satz 2.2.2 die Existenz der I_1-Projektion von A auf $\mathcal{E}^{\ll A}$.

Sei nun B die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$. Dann gilt $B \in \mathcal{E}^{\ll A}$. Daraus folgt nach Satz 2.3.3 (i) die Existenz der biproportionalen Limesanpassung von A an c und r.

(ii) Sei B die biproportionale Limesanpassung von A an c und r. Dann konvergiert das IPF-Verfahren nach Satz 2.3.2 (i) gegen B. Folglich ist die Limesanpassung B nach Satz 2.3.2 (iv) die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$. □

2.4. Flussungleichungen

In Satz 2.3.2 (i) wurde gezeigt, dass das IPF-Verfahren genau dann konvergiert, wenn die Menge $\mathcal{E}^{\ll A}$ nichtleer ist. Das heißt, wenn eine Matrix P existiert, die von der Ausgangsmatrix A dominiert wird und die Marginalien c und r erfüllt. Dieses Existenzkriterium kann mithilfe der sog. Flussungleichungen überprüft werden. Dabei bezeichnet $J_A(I) := \{j \in \{1, \dots, \ell\} \mid \exists i \in I : a_{ij} > 0\}$ die *Menge der in A mit den Zeilen I verbundenen Spalten*. Eine Normierung des biproportionalen Anpassungsproblems (A, c, r) ist nicht erforderlich. Die Frage, ob die Menge $\mathcal{E}^{\ll A}$ leer ist, kann wie folgt entschieden werden.

Satz 2.4.1 (Flussungleichungen, vgl. Hershkowitz, Hoffman und Schneider 1997, Theorem 3.2 und Bemerkung 3.5)**.** *Sei (A, c, r) ein biproportionales Anpassungsproblem und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Matrizen. Dann gelten die folgenden beiden Aussagen:*

(i) Die Menge $\mathcal{E}^{\ll A}$ ist genau dann nichtleer, wenn die Gleichung $c_+ = r_+$ gilt und für alle Zeilenteilmengen $I \subseteq \{1, \dots, k\}$ die Flussungleichungen *$r_I \leq c_{J_A(I)}$ gelten.*

(ii) Falls die Ausgangsmatrix A zusammenhängend ist, existiert eine Matrix $S \in \mathcal{E}^{\ll A}$, für die $S \equiv A$ gilt, genau dann, wenn die Gleichung $c_+ = r_+$ gilt und für alle nichtleeren und echten Zeilenteilmengen $I \subsetneq \{1, \dots, k\}$, $I \neq \emptyset$, die strikten Flussungleichungen *$r_I < c_{J_A(I)}$ gelten.*

Beweis. Siehe Hershkowitz, Hoffman und Schneider (1997, Theorem 3.2 und Bemerkung 3.5). □

Die Flussungleichungen sind somit ein notwendiges und hinreichendes Kriterium für die Konvergenz des zweidimensionalen IPF-Verfahrens. Die strikten Flussungleichungen sind hingegen ein notwendiges und hinreichendes Kriterium für die Direktheit der errechneten biproportionalen Limesanpassung.

2.5. Konvergenzanalyse mithilfe des geometrischen Mittels

Neben dem in Abschnitt 2.3 diskutierten informationstheoretischen Ansatz zur Konvergenzanalyse existieren noch weitere Ansätze. Der im Folgenden präsentierte Ansatz basiert auf einem invertierten geometrischen Mittel und geht auf Pretzel (1980) zurück. Bei näherer Betrachtung erweist er sich allerdings als äquivalent zum informationstheoretischen Ansatz. Mithilfe der Ungleichung vom arithmetischen und geometrischen Mittel zeigt Pretzel das folgende Lemma.

Lemma 2.5.1 (Geometrisches Mittel, vgl. Pretzel 1980, Lemma 1)**.** *Sei $a \in (0; \infty)^k$ und sei die Funktion $f : (0; \infty)^k \to (0; \infty)$ definiert gemäß*

$$f(x) := \prod_i \left(\frac{a_i}{x_i}\right)^{a_i}. \tag{2.5.1}$$

Dann gelten die folgenden beiden Aussagen:

(i) Für alle $x \in (0; \infty)^k$ mit $x_+ = a_+$ gilt $f(x) \geq 1$.

(ii) Für alle Folgen (x^n) in $(0; \infty)^k$ mit $x_+^n = a_+$ gilt $\lim_{n\to\infty} f(x^n) = 1$ genau dann, wenn $\lim_{n\to\infty} x_i^n = a_i$ für alle Einträge i gilt.

Beweis. Siehe Pretzel (1980, Lemma 1). □

Lemma 2.5.1 ermöglicht Pretzel einen kurzen und eleganten Beweis des folgenden Satzes.

Satz 2.5.2 (Randsummen der Häufungspunkte, vgl. Pretzel 1980, Proposition 2)**.** *Sei (A, c, r) ein biproportionales Anpassungsproblem und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Matrizen. Weiter sei B ein beliebiger Häufungspunkt der zugehörigen IPF-Folge $(A(t))$. Dann sind die Randsummen von B gleich den Marginalien c und r,*

$$\forall i: \quad b_{i+} = r_i, \qquad\qquad \forall j: \quad b_{+j} = c_j. \tag{2.5.2}$$

Beweis. Sei $S \in \mathcal{E}^{\ll A}$ beliebig. Davon ausgehend wählen wir

$$\forall n \geq 0: \quad f_n := \prod_{i,j:\, s_{ij}>0} \left(\frac{a_{ij}(2n+1)}{a_{ij}}\right)^{s_{ij}} \quad \text{und} \tag{2.5.3}$$

$$\forall n \geq 0: \quad g_n := \prod_{i,j:\, s_{ij}>0} \left(\frac{a_{ij}(2n+2)}{a_{ij}}\right)^{s_{ij}}. \tag{2.5.4}$$

Nach Lemma 2.5.1 (i) gelten dann die Ungleichungen

$$\forall n \geq 0: \quad \frac{g_n}{f_n} = \prod_{i,j:\, s_{ij}>0} \left(\frac{a_{ij}(2n+2)}{a_{ij}(2n+1)}\right)^{s_{ij}} = \prod_j \left(\frac{c_j}{a_{+j}(2n+1)}\right)^{c_j} \geq 1 \text{ und} \tag{2.5.5}$$

$$\forall n \geq 0: \quad \frac{f_{n+1}}{g_n} = \prod_{i,j:\, s_{ij}>0} \left(\frac{a_{ij}(2n+3)}{a_{ij}(2n+2)}\right)^{s_{ij}} = \prod_i \left(\frac{r_i}{a_{i+}(2n+2)}\right)^{r_i} \geq 1. \tag{2.5.6}$$

Damit sind die Folgen (f_n) und (g_n) monoton steigend,

$$f_0 \leq g_0 \leq f_1 \leq g_1 \leq f_2 \leq g_2 \leq \cdots. \tag{2.5.7}$$

Wir wählen nun

$$L := \frac{\max_i r_i}{\min_{i,j:\, a_{ij}>0} a_{ij}}. \tag{2.5.8}$$

Somit gilt

$$\forall n \geq 0: \quad f_n = \prod_{i,j:\, s_{ij}>0} \left(\frac{a_{ij}(2n+1)}{a_{ij}}\right)^{s_{ij}} \leq \prod_{i,j:\, s_{ij}>0} L^{s_{ij}} \leq L^{r_+} < \infty. \tag{2.5.9}$$

Also sind die Folgen (f_n) und (g_n) monoton steigend und nach oben beschränkt,

$$f_0 \leq g_0 \leq f_1 \leq g_1 \leq f_2 \leq g_2 \leq \cdots \leq L^{r_+} < \infty. \tag{2.5.10}$$

Infolgedessen gilt

$$\lim_{n\to\infty} \prod_j \left(\frac{c_j}{a_{+j}(2n+1)}\right)^{c_j} = \lim_{n\to\infty} \frac{g_n}{f_n} = 1 \text{ und} \tag{2.5.11}$$

$$\lim_{n\to\infty} \prod_i \left(\frac{r_i}{a_{i+}(2n+2)}\right)^{r_i} = \lim_{n\to\infty} \frac{f_{n+1}}{g_n} = 1. \tag{2.5.12}$$

Nach Lemma 2.5.1 (ii) folgt daraus

$$\forall j: \quad \lim_{n\to\infty} a_{+j}(2n+1) = c_j \text{ und} \tag{2.5.13}$$

$$\forall i: \quad \lim_{n\to\infty} a_{i+}(2n+2) = r_i. \tag{2.5.14}$$

Da alle Matrizen $A(2n+1)$ die Zeilenmarginalien r und alle Matrizen $A(2n+2)$ die Spaltenmarginalien c erfüllen, folgt die Behauptung. □

Gemäß dem soeben gezeigten Satz 2.5.2 ist jeder Häufungspunkt $B^* = \lim_{n\to\infty} A(t_n)$ der IPF-Folge $(A(t))$ eine biproportionale Limesanpassung von A an c und r. Mit der Eindeutigkeit biproportionaler Limesanpassungen (siehe Satz 1.2.1, S. 13) folgt daraus die Konvergenz der IPF-Folge $(A(t))$. Da Pretzel aber anscheinend nur die Eindeutigkeit biproportionaler Direktanpassungen kannte (vgl. Pretzel 1980, Proposition 1), benötigt dieser noch weitere Aussagen, um die Konvergenz der IPF-Folge zu zeigen (vgl. Pretzel 1980, Proposition 3 und 4).

Hinter dem kurzen und eleganten Konvergenzbeweis von Pretzel verbirgt sich der in Abschnitt 2.3 diskutierte informationstheoretische Ansatz. Die in Lemma 2.5.1 behandelte Funktion f ist die Exponentialfunktion der I-Divergenz relativ zu a,

$$f(x) = \exp\left(\sum_i a_i \ln \frac{a_i}{x_i}\right) = \exp\left(\mathrm{I}(a|x)\right). \tag{2.5.15}$$

Die untere Schranke $f(x) \geq 1$ ist somit äquivalent zu der in Korollar 2.1.4 gezeigten unteren Schranke $\mathrm{I}(a|x) \geq 0$. Die in Lemma 2.5.1 (ii) gezeigte Äquivalenz folgt aus der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v)) sowie der Pinsker-Ungleichung (siehe Satz 2.1.7). Für die im Beweis von Satz 2.5.2 verwendeten Folgen (f_n) und (g_n) gelten die Gleichungen

$$\forall n \geq 0: \quad f_n = \exp\left(\mathrm{I}(S|A) - \mathrm{I}(S|A(2n+1))\right) \text{ und} \tag{2.5.16}$$

$$\forall n \geq 0: \quad g_n = \exp\left(\mathrm{I}(S|A) - \mathrm{I}(S|A(2n+2))\right). \tag{2.5.17}$$

Dementsprechend gelten

$$\begin{aligned} \forall n \geq 0: \quad \frac{g_n}{f_n} &= \exp\left(\mathrm{I}(S|A(2n+1)) - \mathrm{I}(S|A(2n+2))\right) \\ &= \exp\left(\mathrm{I}(c|(a_{+1}(2n+1), \ldots, a_{+\ell}(2n+1)))\right) \geq 1 \text{ und} \end{aligned} \tag{2.5.18}$$

$$\begin{aligned} \forall n \geq 0: \quad \frac{f_{n+1}}{g_n} &= \exp\left(\mathrm{I}(S|A(2n+2)) - \mathrm{I}(S|A(2n+3))\right) \\ &= \exp\left(\mathrm{I}(r|(a_{1+}(2n+2), \ldots, a_{k+}(2n+2)))\right) \geq 1. \end{aligned} \tag{2.5.19}$$

Im Fall normierter biproportionaler Anpassungsprobleme (A, c, r) kann als obere Schranke für die Folgen (f_n) und (g_n) anstelle von $L^{r_+} = L$ auch $\exp\left(\mathrm{I}(S|A)\right)$ verwendet werden,

$$f_1 \leq g_1 \leq f_2 \leq g_2 \leq f_3 \leq g_3 \leq \cdots \leq \exp\left(\mathrm{I}(S|A)\right) < \infty. \tag{2.5.20}$$

Unabhängig von der Wahl der oberen Schranke impliziert die resultierende Konvergenz der Folgen (g_n/f_n) und (f_{n+1}/g_n) gegen 1 die Konvergenz der Folgen $(\mathrm{I}(c|(a_{+1}(2n+1), \ldots, a_{+\ell}(2n+1))))$

und $(\mathrm{I}(r|(a_{1+}(2n+2),\dots,a_{k+}(2n+2))))$ gegen 0. Mit der Pinsker-Ungleichung (siehe Satz 2.1.7) folgen daraus die Gleichungen (2.5.13) und (2.5.14) und damit die Behauptung von Satz 2.5.2. Durch Nachrechnen erhalten wir außerdem

$$\forall n \geq 0: \ \mathrm{I}(A(2n+2)|A(2n+1)) = \mathrm{I}(c|(a_{+1}(2n+1),\dots,a_{+\ell}(2n+1))) \text{ und} \tag{2.5.21}$$

$$\forall n \geq 0: \ \mathrm{I}(A(2n+3)|A(2n+2)) = \mathrm{I}(r|(a_{1+}(2n+2),\dots,a_{k+}(2n+2))). \tag{2.5.22}$$

Damit sind auch die im informationstheoretischen Ansatz verwendeten Gleichungen (2.3.14) in abgeschwächter Form in den Ungleichungen (2.5.18) und (2.5.19) enthalten.

2.6. L_1-Ansatz zur Konvergenzanalyse

Zur Messung des Fortschritts des zweidimensionalen IPF-Verfahrens bei der Anpassung an die Zeilen- und Spaltenmarginalien schlägt Pukelsheim (2014) den L_1-Ansatz zur Konvergenzanalyse vor. Der L_1-Ansatz verzichtet auf die Normierung des biproportionalen Anpassungsproblems (A,c,r). Er basiert auf dem für alle Matrizen $S \in [0;\infty)^{k\times\ell}$ definierten *L_1-Fehler*

$$f(S) := \sum_i \left|s_{i+} - r_i\right| + \sum_j \left|s_{+j} - c_j\right|. \tag{2.6.1}$$

Für alle Matrizen $S \in [0;\infty)^{k\times\ell}$ gilt $f(S) \geq 0$. Gleichheit gilt genau dann, wenn die Matrix S die Zeilenmarginalien r und die Spaltenmarginalien c erfüllt. Die Anpassungsgüte der IPF-Folge $(A(t))$ wird im L_1-Ansatz mithilfe der *L_1-Fehlerfolge* $(f(A(t)))$ untersucht. Wie das folgende Lemma zeigt, ist die L_1-Fehlerfolge $(f(A(t)))$ monoton fallend und beschränkt.

Lemma 2.6.1 (Monotonie und untere Schranken der L_1-Fehlerfolge, vgl. Pukelsheim 2014, Lemma 2)**.** *Sei (A,c,r) ein biproportionales Anpassungsproblem und $(A(t))$ die zugehörige IPF-Folge. Dann gelten die folgenden beiden Aussagen:*

(i) Die L_1-Fehlerfolge $(f(A(t)))$ fällt monoton,

$$\forall t \geq 1: \quad f(A(t)) \geq f(A(t+1)). \tag{2.6.2}$$

(ii) Die L_1-Fehlerfolge $(f(A(t)))$ ist nach unten beschränkt,

$$\forall t \geq 1: \quad f(A(t)) \geq \max_{I\subseteq\{1,\dots,k\}} \left(r_I - c_{J_A(I)} + c_{J_A(I)'} - r_{I'}\right). \tag{2.6.3}$$

Beweis. (i) Sei $t \geq 2$ ein beliebiger gerader Schritt. Dann gilt für den L_1-Fehler $f(A(t))$ die Aussage

$$f(A(t)) = \sum_i \left|a_{i+}(t) - r_i\right| = \sum_i \Big|\sum_j \big(a_{ij}(t) - a_{ij}(t+1)\big)\Big|. \tag{2.6.4}$$

Da die Matrix $A(t+1)$ nach Definition des IPF-Verfahrens durch Zeilenskalierung aus $A(t)$ hervorgeht, hat die Differenz $a_{ij}(t)-a_{ij}(t+1)$ ein innerhalb einer jeden Zeile i gleichbleibendes

Vorzeichen. Daraus folgt

$$\begin{aligned}\sum_i \Big| \sum_j \big(a_{ij}(t) - a_{ij}(t+1)\big) \Big| &= \sum_i \sum_j \big|a_{ij}(t) - a_{ij}(t+1)\big| \\ &\geq \sum_j \Big| \sum_i \big(a_{ij}(t) - a_{ij}(t+1)\big) \Big| \\ &= \sum_j \big|c_j - a_{+j}(t+1)\big| = f(A(t+1)). \end{aligned} \tag{2.6.5}$$

Somit gilt für alle geraden Schritte $t \geq 2$ die Ungleichung $f(A(t)) \geq f(A(t+1))$. Für alle ungeraden Schritte $t \geq 1$ folgt die Ungleichung in analoger Weise.

(ii) Sei $t \geq 2$ ein beliebiger gerader Schritt und sei $I \subseteq \{1, \ldots, k\}$ eine beliebige Zeilenteilmenge. Dann gilt

$$\sum_{i \in I} (a_{i+}(t) - r_i) + \sum_{i \in I'} (a_{i+}(t) - r_i) = a_{++}(t) - r_+ = c_+ - r_+ \tag{2.6.6}$$

und damit

$$\sum_{i \in I'} (a_{i+}(t) - r_i) = c_+ - r_+ + \sum_{i \in I} (r_i - a_{i+}(t)). \tag{2.6.7}$$

Daraus folgt

$$\begin{aligned} f(A(t)) = \sum_i \big|a_{i+}(t) - r_i\big| &\geq \sum_{i \in I} (r_i - a_{i+}(t)) + \sum_{i \in I'} (a_{i+}(t) - r_i) \\ &= c_+ - r_+ + 2 \sum_{i \in I} (r_i - a_{i+}(t)). \end{aligned} \tag{2.6.8}$$

Für die Summe gilt

$$\begin{aligned} \sum_{i \in I} (r_i - a_{i+}(t)) &= r_I - a_{I+}(t) = r_I - a_{I J_A(I)}(t) \\ &= r_I - a_{+J_A(I)}(t) + a_{I' J_A(I)}(t) \geq r_I - c_{J_A(I)}. \end{aligned} \tag{2.6.9}$$

Also gilt insgesamt

$$f(A(t)) \geq c_+ - r_+ + 2\left(r_I - c_{J_A(I)}\right) = r_I - c_{J_A(I)} - r_{I'} + c_{J_A(I)'}. \tag{2.6.10}$$

Da diese Abschätzung für alle Zeilenteilmengen $I \subseteq \{1, \ldots, k\}$ gilt und der L_1-Fehler $f(A(t+1))$ für jeden ungeraden Schritt $t+1$ gemäß Aussage (i) nach unten durch $f(A(t+2))$ abgeschätzt werden kann, ist somit Aussage (ii) gezeigt. □

Aus dem soeben bewiesenen Lemma 2.6.1 folgt die Konvergenz der L_1-Fehlerfolge $(f(A(t)))$. Der folgende Satz fasst die aus der informationstheoretischen Analyse bekannten Konvergenzkriterien für das zweidimensionale IPF-Verfahren zusammen und zeigt, dass der Grenzwert der L_1-Fehlerfolge genau dann gleich 0 ist, wenn die IPF-Folge konvergiert.

Satz 2.6.2 (Konvergenz des zweidimensionalen IPF-Verfahrens, vgl. Pukelsheim 2014, Theorem 3)**.** *Sei (A, c, r) ein biproportionales Anpassungsproblem, $(A(t))$ die zugehörige IPF-Folge und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Matrizen. Dann sind die folgenden fünf Aussagen äquivalent:*

(i) Die IPF-Folge $(A(t))$ konvergiert.

(ii) Die biproportionale Limesanpassung B von A an c und r existiert.

(iii) Die Menge $\mathcal{E}^{\ll A}$ ist nichtleer.

(iv) Es gilt $r_+ = c_+$ und für alle Zeilenteilmengen $I \subseteq \{1, \ldots, k\}$ gelten die Flussungleichungen $r_I \leq c_{J_A(I)}$.

(v) Die L_1-Fehlerfolge $(f(A(t)))$ konvergiert gegen 0.

Beweis. (i) ⇔ (iii) : Diese Äquivalenz wurde in Satz 2.3.2 (i) gezeigt.

(ii) ⇔ (iii) : Diese Äquivalenz wurde in Satz 2.3.3 (i) gezeigt.

(iii) ⇔ (iv) : Diese Äquivalenz wurde in Satz 2.4.1 (i) gezeigt.

(i) ⇒ (v) : Sei die IPF-Folge $(A(t))$ konvergent. Dann ist die Limesmatrix $B := \lim_{t\to\infty} A(t)$ die biproportionale Limesanpassung von A an c und r und es gilt

$$\lim_{t\to\infty} f\Big(A(t)\Big) = f\Big(\lim_{t\to\infty} A(t)\Big) = f\Big(B\Big) = 0. \tag{2.6.11}$$

(v) ⇒ (i) : Die L_1-Fehlerfolge $(f(A(t)))$ konvergiere gegen 0. Da die Matrizen $A(t)$ für alle Schritte $t \geq 1$ in der kompakten Menge $[0; \max(c_+, r_+)]^{k\times\ell}$ enthalten sind, besitzt die IPF-Folge $(A(t))$ mindestens einen Häufungspunkt. Sei $B^* = \lim_{n\to\infty} A(t_n)$ ein solcher Häufungspunkt. Dann gilt mit der Stetigkeit des L_1-Fehlerfunktionals und der Monotonie der L_1-Fehlerfolge (siehe Lemma 2.6.1 (i)) die Gleichung

$$f\Big(B^*\Big) = f\Big(\lim_{n\to\infty} A(t_n)\Big) = \lim_{n\to\infty} f\Big(A(t_n)\Big) = \lim_{t\to\infty} f\Big(A(t)\Big) = 0. \tag{2.6.12}$$

Der Häufungspunkt B^* erfüllt somit die Zeilenmarginalien r und die Spaltenmarginalien c und ist folglich eine biproportionale Limesanpassung von A an c und r. Da die biproportionale Limesanpassung nach Satz 1.2.1 (siehe S. 13) eindeutig ist, hat die IPF-Folge $(A(t))$ genau einen Häufungspunkt und ist damit konvergent. □

Der L_1-Ansatz zur Konvergenzanalyse ermöglicht es also, die Konvergenz der IPF-Folge an der L_1-Fehlerfolge abzulesen. Konvergiert die L_1-Fehlerfolge gegen 0, so konvergiert auch die IPF-Folge. Konvergiert die L_1-Fehlerfolge gegen einen positiven Wert, so divergiert die IPF-Folge. Um die Konvergenz des IPF-Verfahrens zu entscheiden ohne die IPF-Folge zu errechnen, sind jedoch weiterhin die aus der informationstheoretischen Analyse bekannten Konvergenzkriterien erforderlich.

2.7. Kommentare und Referenzen

Die Diskussion der I-Divergenz in Abschnitt 2.1 findet sich in verkürzter Darstellung bei Gietl und Reffel (2013a, Abschnitt 2). Die I-Divergenz geht auf Kullback und Leibler (1951) zurück und wurde nicht zuletzt durch Kullback (1959) populär. Sie ist unter anderem auch unter den Bezeichnungen „Kullback-Leibler information", „information for discrimination", „information gain" und „relative entropy" bekannt (vgl. Csiszár 1975, S. 146). Die einschlägige Literatur definiert die I-Divergenz entweder nur für Wahrscheinlichkeitsmaße (vgl. Csiszár 1975) oder fügt einen Korrekturterm für endliche Maße hinzu (vgl. Csiszár und Tusnády 1984). Selbstverständlich könnten die hier untersuchten nichtnegativen Vektoren und Matrizen als endliche Maße auf endlichen Räumen betrachtet werden. Aus Gründen der Einfachheit verzichtet Definition 2.1.1 jedoch auf den Korrekturterm. Infolgedessen gilt es zu beachten, dass die Nichtnegativität der I-Divergenz nur für Vektoren und Matrizen mit übereinstimmenden Summen gilt (siehe Korollare 2.1.4 und 2.1.6). Für letztere Vektoren und Matrizen ist die I-Divergenz genau dann gleich 0, wenn beide Vektoren beziehungsweise beide Matrizen übereinstimmen. Somit kann die I-Divergenz als „Abstand" dienen. Sie ist jedoch keine Metrik und kann auch durch keine sinnvolle Funktion in eine solche transformiert werden (vgl. Csiszár 1964, Satz 2). Eine ausführliche Diskussion der I-Divergenz zwischen Wahrscheinlichkeitsmaßen auf beliebigen Messräumen sowie ihrer statistischen Interpretationen findet sich bei Gietl (2011, Kapitel 2). Der Beweis der Pinsker-Ungleichung (siehe Satz 2.1.7) wurde von Cover und Thomas (2006, Lemma 11.6.1) übernommen. Die Bezeichnung „Pinsker-Ungleichung" bezieht sich auf Pinsker (1963, Kapitel I, Abschnitt 2.3), der eine schwächere Ungleichung zeigen konnte. Die Ungleichung in ihrer heutigen Form wurde unabhängig voneinander von Kullback (1967), Csiszár (1967, Theorem 4.1) und Kemperman (1969, Theorem 6.1) bewiesen. Laut Tsybakov (2009, S. 132) ist daher auch die Bezeichnung „Kullback-Csiszár-Kemperman inequality" üblich.

Die in Abschnitt 2.2 untersuchten I_1-Projektionen wurden von Csiszár (1975, S. 147) eingeführt. Csiszár bezeichnet sie als „I-projections". Die hier verwendeten Bezeichnungen „I_1-Projektion" und „I_2-Projektion" richten sich nach Cramer (2000, Definition 2.1). Die in Satz 2.2.4 untersuchte Stetigkeit von I_1-Projektionen wurde erstmals von Gietl und Reffel (2013c, Theorem 2) gezeigt. Dieses Resultat wurde außerdem in Gietl und Reffel (2013b, Theorem 2.4) veröffentlicht. Sowohl in Gietl und Reffel (2013b) als auch in Gietl und Reffel (2013c) werden jedoch anstelle der I_1-Projektionen sogenannte f-Projektionen untersucht. Dabei handelt es sich um eine Familie von Projektionen, die die I_1-Projektionen enthält und die auf den sogenannten f-Divergenzen basiert. Bei den f-Divergenzen handelt es sich wiederum um eine Familie von Divergenzen, die die I-Divergenz enthält.

Die in Abschnitt 2.3 vorgestellte informationstheoretische Konvergenzanalyse basiert auf der Arbeit mehrerer Autoren. Als erstes haben Ireland und Kullback (1968) die I-Divergenz zur Analyse des IPF-Verfahrens verwendet. Wie Csiszár (1975, S. 155) anmerkt, ist ihr Konvergenzbeweis jedoch unvollständig. Der erste vollständige Konvergenzbeweis findet sich ebenfalls bei Csiszár (1975, Theorem 3.2). In Lemma 2.3.1 könnte man die Menge $\mathcal{C}^{\ll A}$ durch $\mathcal{C}$ und die Menge $\mathcal{R}^{\ll A}$ durch $\mathcal{R}$ ersetzen. Allerdings könnten dann in den Gleichungen (2.3.7) und (2.3.9) jeweils beide Seiten unendlich sein. Die Sätze 2.3.2 und 2.3.3 könnten ebenfalls angepasst werden. Die in Lemma 2.3.1 gezeigten Gleichungen (2.3.7) und (2.3.9) gehen auf Csiszár (1975, Theorem 2.2) zurück. Sie sind ein informationstheoretisches Analogon zum Satz des Pythagoras und werden daher auch als „Pythagorean identit[ies]" (Brown, Chase und Pittenger 1993, S. 6) oder als „three points propert[ies]“ (Csiszár und Tusnády 1984, S. 215) bezeichnet. Die ebenfalls in Lemma 2.3.1 gezeigten Ungleichungen (2.3.8) und (2.3.10) wurden von Cramer (2000, Theorem 3.8 (i)) übernommen. Die in Satz 2.3.2 zusammengefassten Resultate wurden erstmals von Csiszár (1975, Theorem 3.2) gezeigt. Für seinen Beweis entwickelt Csiszár vorab eine Informationsgeometrie (Abschnitt 2). Dabei beweist er eine mit den Gleichungen (2.3.7) und (2.3.9) vergleichbare Pythagoras-Ungleichung für beliebige I_1-Projektionen (Theorem 2.2) und zeigt eine hinreichende Bedingung dafür, dass die Konkatenation der I_1-Projektion auf eine beliebige Menge $\mathcal{E}$ und der I_1-Projektion auf eine Teilmenge $\mathcal{E}_1 \subseteq \mathcal{E}$ der alleinigen I_1-Projektion auf die Teilmenge $\mathcal{E}_1$ entspricht (Theorem 2.3). Der Beweis selbst nutzt dann die Existenz der I_1-Projektion B von A auf die Menge $\mathcal{E}^{\ll A}$, um die Konvergenzaussage (2.3.17) zu zeigen und daraus die Zugehörigkeit aller Häufungspunkte B^* der IPF-Folge $(A(t))$ zur Menge $\mathcal{E}^{\ll A}$ zu folgern. Die Konkatenationseigenschaft und die Pythagoras-Ungleichung ermöglichen anschließend den Beweis, dass jeder Häufungspunkt B^* gleich der I_1-Projektion B ist, und somit den Beweis der Konvergenz der IPF-Folge. Der in Abschnitt 2.3 angegebene Beweis von Satz 2.3.2 folgt allerdings der Argumentation von Brown, Chase und Pittenger (1993, Theorem 3.1). Brown, Chase und Pittenger verzichten zunächst auf eine Diskussion der Existenz der I_1-Projektion. Stattdessen folgern sie aus der Existenz einer Matrix $P \in \mathcal{E}^{\ll A}$ die Konvergenz der IPF-Folge. Für deren Grenzwert B^* zeigen sie dann die Pythagoras-Gleichung (2.3.13). Aus dieser Gleichung folgt, dass B^* die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$ ist. Satz 2.3.3 wurde unter anderem bereits von Pukelsheim (2014, Theorem 2 und Theorem 3) gezeigt. Der in Abschnitt 2.3 angegebene Beweis basiert jedoch im Gegensatz zu den Beweisen von Pukelsheim auf den Ergebnissen der informationstheoretischen Konvergenzanalyse.

Die in Abschnitt 2.4 eingeführte Notation orientiert sich an Pukelsheim (2014). Ein Beweis von Satz 2.4.1 findet sich bei Hershkowitz, Hoffman und Schneider (1997, Theorem 3.2 und Bemerkung 3.5). Dabei geht Aussage (i) auf Rothblum und Schneider (1989, Theorem 3) zurück. Aussage (ii) kann ebenfalls bei Rothblum und Schneider (1989, Theorem 2) gefunden werden, geht aber auf Brualdi (1968, Theorem 2.1) und auf Menon und Schneider (1969, Korollar 4.2) zurück. Das zentrale Hilfsmittel in den genannten Aufsätzen von Brualdi und von Hershkowitz, Hoffman und Schneider ist das aus der Netzwerktheorie bekannte Zirkulationstheorem von Hoffman. Aus dieser Tatsache resultiert die Bezeichnung „Flussungleichungen“.

Die in Abschnitt 2.5 vorgestellte Konvergenzanalyse mithilfe des geometrischen Mittels geht, wie bereits erwähnt, auf Pretzel (1980) zurück. Dementsprechend orientiert sich die in Lemma 2.5.1 und Satz 2.5.2 neu hinzugekommene Notation an Pretzel (1980). Einige kleine Anpassungen sind der Tatsache geschuldet, dass Pretzel in seiner Definition des IPF-Verfahrens nicht mit einer Zeilenanpassung sondern mit einer Spaltenanpassung beginnt.

Der in Abschnitt 2.6 vorgestellte L_1-Ansatz wurde, wie bereits erwähnt, von Pukelsheim (2014) vorgeschlagen. Dementsprechend orientiert sich die durch das L_1-Fehlerfunktional neu hinzugekommene Notation an Pukelsheim (2014). Der Beweis von Lemma 2.6.1 und der Schritt „(v) $\Rightarrow$ (i)“ im Beweis von Satz 2.6.2 wurden ebenfalls von Pukelsheim (2014) übernommen. In früheren Fassungen von Pukelsheim (2014) schlug Pukelsheim einen Beweisansatz für Satz 2.6.2 vor, der ausschließlich auf dem L_1-Fehlerfunktional basiert und keinerlei Gebrauch von der I-Divergenz oder ähnlichen Funktionalen macht. Dieser Ansatz erwies sich jedoch als nicht zielführend, da die Konvergenz der L_1-Fehlerfolge gegen die in Lemma 2.6.1 (ii) gezeigte untere Schranke nicht ohne Rückgriff auf die I-Divergenz gezeigt werden konnte. In der finalen Fassung beweist Pukelsheim Satz 2.6.2 stattdessen mithilfe der von Gietl und Reffel (2013a) für alle biproportionalen Anpassungsprobleme (A, c, r) gezeigten Konvergenz der geradzahligen IPF-

Teilfolge $(A(t))_{t=0,2,4,\ldots}$ und der ungeradzahligen IPF-Teilfolge $(A(t+1))_{t=0,2,4,\ldots}$. Der Beweis dieser Konvergenzaussagen basiert allerdings ebenfalls auf der I-Divergenz und wird im nächsten Kapitel vorgestellt.

3. Divergenzanalyse des zweidimensionalen IPF-Verfahrens

Nachdem im vorangegangenen Kapitel notwendige und hinreichende Kriterien für die Konvergenz des IPF-Verfahrens gezeigt wurden, wird in diesem Kapitel das asymptotische Verhalten des IPF-Verfahrens im Divergenzfall untersucht. Der erste Abschnitt stellt das IPF-Verfahren als eine abwechselnde Minimierung der I-Divergenz dar. Im zweiten Abschnitt werden eine Drei-Punkte-Gleichung und eine Vier-Punkte-Ungleichung gezeigt. Im dritten Abschnitt wird schließlich bewiesen, dass die IPF-Teilfolge entlang der geraden Schritte und die IPF-Teilfolge entlang der ungeraden Schritte jeweils konvergieren und dass die IPF-Folge somit maximal zwei Häufungspunkte hat.

3.1. Abwechselnde Minimierung der I-Divergenz

In Lemma 2.3.1 (siehe S. 30) wurde gezeigt, dass jede Zeilenanpassung zugleich der I_1-Projektion und der I_2-Projektion auf $\mathcal{R}^{\ll A}$ entspricht und dass jede Spaltenanpassung zugleich der I_1-Projektion und der I_2-Projektion auf $\mathcal{C}^{\ll A}$ entspricht. Die Konvergenzanalyse in Abschnitt 2.3 (siehe S. 29) macht jedoch nur von der Charakterisierung als I_1-Projektion Gebrauch. Bei der Analyse des asymptotischen Verhaltens des IPF-Verfahrens im Divergenzfall nutzen wir zusätzlich die Charakterisierung als I_2-Projektion und betrachten das IPF-Verfahren als ein Verfahren abwechselnder I_1-Projektionen und I_2-Projektionen. Das heißt, wir betrachten jede Spaltenanpassung $A(t+2)$ als die I_1-Projektion von $A(t+1)$ auf $\mathcal{C}^{\ll A}$ und jede Zeilenanpassung $A(t+1)$ als die I_2-Projektion von $A(t)$ auf $\mathcal{R}^{\ll A}$.

Zur Vereinfachung der Darstellung erlauben wir neben Matrizen auch die beiden Mengen $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ als Argumente der I-Divergenz. Für beliebige Wahrscheinlichkeitsmatrizen P mit $P \ll A$ definieren wir

$$\mathrm{I}(\mathcal{C}^{\ll A}|P) := \inf_{C\in\mathcal{C}^{\ll A}} \mathrm{I}(C|P), \tag{3.1.1}$$

$$\mathrm{I}(P|\mathcal{R}^{\ll A}) := \inf_{R\in\mathcal{R}^{\ll A}} \mathrm{I}(P|R) \text{ und} \tag{3.1.2}$$

$$\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) := \inf_{(C,R)\in\mathcal{C}^{\ll A}\times\mathcal{R}^{\ll A}} \mathrm{I}(C|R). \tag{3.1.3}$$

Aufgrund der Kompaktheit der Mengen $\mathcal{C}^{\ll A}$, $\mathcal{R}^{\ll A}$ sowie $\mathcal{C}^{\ll A}\times\mathcal{R}^{\ll A}$ und der unteren Halbstetigkeit der Abbildungen $\mathrm{I}(\cdot|P)$, $\mathrm{I}(P|\cdot)$ sowie $\mathrm{I}(\cdot|\cdot)$ (siehe Satz 2.1.2 (iv), S. 23) werden die drei Infima angenommen (vgl. Bauschke und Combettes 2011, Theorem 1.28). Die I-Divergenz $\mathrm{I}(\mathcal{C}^{\ll A}|P)$ ist genau dann endlich, wenn alle Spaltensummen von P positiv sind. In diesem Fall gilt $\mathrm{I}(C|P) = \mathrm{I}(\mathcal{C}^{\ll A}|P)$ genau dann, wenn C die Spaltenanpassung von P ist. Die I-Divergenz $\mathrm{I}(P|\mathcal{R}^{\ll A})$ ist für alle Wahrscheinlichkeitsmatrizen P mit $P \ll A$ endlich. Falls die Zeilensummen von P positiv sind, gilt $\mathrm{I}(P|R) = \mathrm{I}(P|\mathcal{R}^{\ll A})$ genau dann, wenn R die Zeilenanpassung von P ist. Die I-Divergenz $\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ ist stets endlich. Im Allgemeinen existieren jedoch mehrere Paare $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$, für die $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ gilt.

Satz 3.1.1. *Sei* $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ *ein Matrizenpaar, das die Gleichung* $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ *erfüllt. Dann gelten die folgenden vier Aussagen:*

(i) Die Matrix C ist die Spaltenanpassung von R an die Spaltenmarginalien c.

(ii) Die Matrix R ist die Zeilenanpassung von C an die Zeilenmarginalien r.

(iii) Die Matrizen C und R sind maßtheoretisch äquivalent.

(iv) Es gilt $c_{i+} > 0$ für alle Zeilen i und $r_{+j} > 0$ für alle Spalten j.

Beweis. Nach Voraussetzung gilt $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) < \infty$ und somit $\mathrm{I}(C|R) = \mathrm{I}(C|\mathcal{R}^{\ll A}) = \mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) < \infty$. Die Aussagen (i) bis (iv) folgen aus der obigen Diskussion der I-Divergenzen $\mathrm{I}(C|\mathcal{R}^{\ll A})$ und $\mathrm{I}(\mathcal{C}^{\ll A}|R)$.

(i) Aus der Endlichkeit der I-Divergenz $\mathrm{I}(\mathcal{C}^{\ll A}|R)$ folgt die Positivität der Spaltensummen r_{+j} von R. Da außerdem $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|R)$ gilt, ist C die Spaltenanpassung von R an die Spaltenmarginalien c.

(ii) Nachdem C die Spaltenanpassung von R an die Spaltenmarginalien c ist, sind die Zeilensummen c_{i+} von C positiv. Da außerdem $\mathrm{I}(C|R) = \mathrm{I}(C|\mathcal{R}^{\ll A})$ gilt, ist R ist die Zeilenanpassung von C an die Zeilenmarginalien r.

(iii) Nachdem C die Spaltenanpassung von R an die Spaltenmarginalien c ist, sind C und R maßtheoretisch äquivalent.

(iv) Die Positivität der Randsummen wurde bereits in den Beweisen der Aussagen (i) und (ii) gezeigt. □

3.2. Mehr-Punkte-Eigenschaften

Für die Divergenzanalyse werden neben der Charakterisierung von Zeilenanpassungen und Spaltenanpassungen als I_1-Projektionen und I_2-Projektionen die folgenden Mehr-Punkte-Eigenschaften benötigt.

Satz 3.2.1 (Mehr-Punkte-Eigenschaften, vgl. Gietl und Reffel 2013a, Theorem 4.1)**.** *Sei* (A, c, r) *ein normiertes biproportionales Anpassungsproblem,* $(A(t))$ *die zugehörige IPF-Folge und* $\mathcal{C}^{\ll A}$ *sowie* $\mathcal{R}^{\ll A}$ *die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen. Außerdem seien* $C \in \mathcal{C}^{\ll A}$ *und* $R \in \mathcal{R}^{\ll A}$ *so gewählt, dass* $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ *gilt. Dann gelten für alle geraden Schritte* $t \geq 0$ *die folgenden beiden Mehr-Punkte-Eigenschaften:*

(a) die Drei-Punkte-Gleichung

$$\mathrm{I}(C|A(t+2)) + \mathrm{I}(A(t+2)|A(t+1)) = \mathrm{I}(C|A(t+1)) \textit{ und} \tag{3.2.1}$$

(b) die Vier-Punkte-Ungleichung

$$\mathrm{I}(C|A(t+3)) \leq \mathrm{I}(C|A(t+2)) + \mathrm{I}(C|R). \tag{3.2.2}$$

Beweis. Nach Definition des IPF-Verfahrens sowie der Menge $\mathcal{C}^{\ll A}$ gilt

$$C \ll A \equiv A(t+1) \equiv A(t+2) \equiv A(t+3). \tag{3.2.3}$$

(a) Die Drei-Punkte-Gleichung ist ein Spezialfall der in Lemma 2.3.1 (iii) (siehe S. 30) gezeigten Gleichung (2.3.9).

(b) Nach Definition des IPF-Verfahrens gilt

$$a_{ij}(t+3) = \frac{r_i}{a_{i+}(t+2)} \cdot a_{ij}(t+2) \text{ für alle Einträge } (i,j). \tag{3.2.4}$$

Daraus folgt mit der Positivität von c_{i+}, r_i und $a_{i+}(t+2)$ für alle Zeilen i und mit der Zeilensummenungleichung (siehe Satz 2.1.5 (ii), S. 24) sowie der Nichtnegativität der I-Divergenz (siehe Korollar 2.1.4, S. 24) die Ungleichung

$$\begin{aligned}
\mathrm{I}(C|A(t+3)) &= \sum_{i,j:\, c_{ij}>0} c_{ij} \ln \frac{c_{ij}}{\frac{r_i}{a_{i+}(t+2)} \cdot a_{ij}(t+2)} \\
&= \sum_{i,j:\, c_{ij}>0} c_{ij} \ln \frac{c_{ij}}{a_{ij}(t+2)} + \sum_i c_{i+} \ln \frac{a_{i+}(t+2)}{r_i} \\
&= \mathrm{I}(C|A(t+2)) + \sum_i c_{i+} \ln \frac{c_{i+}}{r_i} - \sum_i c_{i+} \ln \frac{c_{i+}}{a_{i+}(t+2)} \\
&\leq \mathrm{I}(C|A(t+2)) + \mathrm{I}(C|R).
\end{aligned} \tag{3.2.5}$$

Damit ist die Vier-Punkte-Ungleichung gezeigt. □

Wenn wir die I-Divergenz als einen „Abstand“ betrachten, können wir die beiden Mehr-Punkte-Eigenschaften wie folgt interpretieren: Die Matrix $A(t+1) \in \mathcal{R}^{\ll A}$ wird an die Spaltenmarginalien angepasst. Dadurch fällt ihr „Abstand“ zur Matrix $C \in \mathcal{C}^{\ll A}$ um den Wert $\mathrm{I}(A(t+2)|A(t+1))$. Anschließend wird die resultierende Matrix $A(t+2) \in \mathcal{C}^{\ll A}$ wieder an die Zeilenmarginalien angepasst. Infolgedessen steigt der „Abstand“ zur Matrix C im Allgemeinen wieder, jedoch maximal um den „Mindestabstand“ $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ der beiden Mengen $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$.

3.3. Häufungspunkte der IPF-Folge

Satz 3.3.1 (Häufungspunkte der IPF-Folge, vgl. Gietl und Reffel 2013a, Theorem 5.3)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem, $(A(t))$ die zugehörige IPF-Folge und $\mathcal{C}^{\ll A}$ sowie $\mathcal{R}^{\ll A}$ die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen. Dann gelten die folgenden vier Aussagen:*

(i) Die geradzahlige IPF-Teilfolge $(A(t))_{t=0,2,4,\ldots}$ konvergiert.

(ii) Die ungeradzahlige IPF-Teilfolge $(A(t+1))_{t=0,2,4,\ldots}$ konvergiert.

(iii) *Die IPF-Folge* $(A(t))$ *hat maximal zwei Häufungspunkte.*

(iv) *Die Grenzwerte* $C^* := \lim_{t=0,2,4,\ldots} A(t)$ *und* $R^* := \lim_{t=0,2,4,\ldots} A(t+1)$ *erfüllen die Gleichung*

$$\mathrm{I}(C^*|R^*) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}). \tag{3.3.1}$$

Beweis. Der Beweis erfolgt in drei Schritten.

I. Seien $C \in \mathcal{C}^{\ll A}$ und $R \in \mathcal{R}^{\ll A}$ so gewählt, dass $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ gilt. Nach Definition des IPF-Verfahrens sowie der Mengen $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ gilt

$$\forall t \geq 0: \quad C \ll A \equiv A(t) \text{ und } R \ll A \equiv A(t). \tag{3.3.2}$$

Aus der Vier-Punkte-Ungleichung und der Drei-Punkte-Gleichung (siehe Satz 3.2.1) folgt die Ungleichung

$$\begin{aligned}\forall t = 2,4,6\ldots: \quad \mathrm{I}(C|A(t)) &\geq \mathrm{I}(C|A(t+1)) - \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})\\ &= \mathrm{I}(C|A(t+2)) + \mathrm{I}(A(t+2)|A(t+1)) - \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}).\end{aligned} \tag{3.3.3}$$

Wiederholte Anwendung dieser Ungleichung liefert

$$\begin{aligned}&\forall t = 2,4,6,\ldots:\\ &\mathrm{I}(C|A(2)) \geq \mathrm{I}(C|A(t)) + \sum_{\tau=4,6,8,\ldots,t} \Big(\mathrm{I}(A(\tau)|A(\tau-1)) - \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})\Big).\end{aligned} \tag{3.3.4}$$

Für die unendliche Reihe folgt daraus

$$\mathrm{I}(C|A(2)) \geq \sum_{\tau=4,6,8,\ldots} \Big(\mathrm{I}(A(\tau)|A(\tau-1)) - \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})\Big). \tag{3.3.5}$$

Da $\mathrm{I}(C|A(2))$ jedoch endlich ist und für alle geraden Schritte $\tau \geq 2$ die Ungleichung $\mathrm{I}(A(\tau)|A(\tau-1)) \geq \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ gilt, folgt

$$\lim_{t=2,4,6,\ldots} \mathrm{I}(A(t)|A(t-1)) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}). \tag{3.3.6}$$

II. Wir wählen eine geradzahlige Schrittfolge (t_n), sodass die IPF-Teilfolgen $(A(t_n-1))$ und $(A(t_n))$ konvergieren. Außerdem setzen wir $R^* := \lim_{n\to\infty} A(t_n-1) \in \mathcal{R}^{\ll A}$ und $C^* := \lim_{n\to\infty} A(t_n) \in \mathcal{C}^{\ll A}$. Daraus folgt mit der unteren Halbstetigkeit der I-Divergenz (siehe Satz 2.1.2 (iv), S. 23) und mit der Gleichung (3.3.6) die Ungleichung

$$\begin{aligned}\mathrm{I}(C^*|R^*) &= \mathrm{I}(\lim_{n\to\infty} A(t_n)|\lim_{n\to\infty} A(t_n-1))\\ &\leq \liminf_{n\to\infty} \mathrm{I}(A(t_n)|A(t_n-1)) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}).\end{aligned} \tag{3.3.7}$$

Somit gilt $\mathrm{I}(C^*|R^*) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$.

III. Diese Gleichheit erlaubt die Anwendung von Ungleichung (3.3.3) auf die in Schritt II gewählten Matrizen $C^* \in \mathcal{C}^{\ll A}$ und $R^* \in \mathcal{R}^{\ll A}$. Damit ist die Folge $(\mathrm{I}(C^*|A(t)))_{t=2,4,6,\dots}$ monoton fallend. Mit der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v), S. 23) folgt

$$\lim_{t=2,4,6,\dots} \mathrm{I}(C^*|A(t)) = \lim_{n\to\infty} \mathrm{I}(C^*|A(t_n)) = \mathrm{I}(C^*|\lim_{n\to\infty} A(t_n)) = \mathrm{I}(C^*|C^*) = 0. \tag{3.3.8}$$

Daraus folgt mit der Pinsker-Ungleichung (siehe Satz 2.1.7, S. 25) die Aussage

$$\lim_{t=2,4,6,\dots} A(t) = C^*. \tag{3.3.9}$$

Somit ist Aussage (i) gezeigt. Mit der Stetigkeit der Zeilenanpassung folgt daraus Aussage (ii). Aussage (iii) folgt unmittelbar aus den Aussagen (i) und (ii). Aussage (iv) wurde bereits in Schritt II gezeigt. □

3.4. Kommentare und Referenzen

Im Gegensatz zum Konvergenzfall wurde dem Fall der Divergenz des IPF-Verfahrens in der einschlägigen Literatur bisher nur wenig Beachtung zuteil. Die in diesem Kapitel präsentierte Divergenzanalyse wurde erstmals in Gietl und Reffel (2013a) vorgestellt. Die sowohl hier als auch dort angegebenen Beweise basieren auf den Ergebnissen von Csiszár und Tusnády (1984) zur alternierenden Minimierung der I-Divergenz. Während Gietl und Reffel beliebige biproportionale Anpassungsprobleme erlauben, beschränkt sich die Darstellung in diesem Kapitel auf die Untersuchung normierter biproportionaler Anpassungsprobleme. Diese Beschränkung ist, wie in Satz 1.5.1 (siehe S. 20) gezeigt wurde, keine echte Einschränkung. Sie vereinfacht jedoch die Argumentation, da die I-Divergenz für Wahrscheinlichkeitsmatrizen nach unten durch 0 beschränkt ist (siehe Korollar 2.1.6, S. 25).

Die zu Beginn von Abschnitt 3.1 getroffene Entscheidung, Spaltenanpassungen als I_1-Projektionen und Zeilenanpassungen als I_2-Projektion zu betrachten, ist willkürlich. Genauso könnte man Zeilenanpassungen als I_1-Projektionen und Spaltenanpassungen als I_2-Projektion betrachten. Die restliche Argumentation müsste dann allerdings angepasst werden. Die Aussage von Satz 3.1.1 wurde von Fabian P. Reffel in einem unveröffentlichten Manuskript für den Spezialfall der beiden Häufungspunkte C^* und R^* der IPF-Folge bewiesen. Die Beweisidee konnte jedoch unverändert für den allgemeineren Fall der hier betrachteten Matrizenpaare übernommen werden.

Auf die Voraussetzung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ in Satz 3.2.1 wurde in dem entsprechenden Theorem 4.1 in Gietl und Reffel (2013a) verzichtet. Stattdessen werden die beiden Mehr-Punkte-Eigenschaften dort für beliebige Matrizen $C \in \mathcal{C}$ und $R \in \mathcal{R}$ gezeigt. Die Forderung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ ermöglicht jedoch die in Abschnitt 3.2 beschriebene anschauliche Interpretation der Mehr-Punkte-Eigenschaften.

Die im Beweis von Satz 3.3.1 gezeigte Aussage (3.3.6) wird auch als „convergence lemma“ bezeichnet (vgl. Gietl und Reffel 2013a, Lemma 5.2). Streng genommen zeigen Gietl und Reffel (2013a) jedoch nicht die Aussage (3.3.6) sondern die Gleichung $\lim_{t=2,4,6,\dots} \mathrm{I}(A(t)|A(t+1)) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$. Dafür wird die Fünf-Punkte-Eigenschaft benötigt. Diese Eigenschaft folgt aber aus der Summation der Drei-Punkte-Gleichung und der Vier-Punkte-Ungleichung sowie einer zusätzlichen Abschätzung. Die ebenfalls im Beweis von Satz 3.3.1 gezeigte Aussage über die Monotonie der Folge $(\mathrm{I}(C^*|A(t)))_{t=2,4,6,\dots}$ wird auch als „monotonicity lemma“ bezeichnet (vgl. Gietl und Reffel 2013a, Lemma 5.1).

4. Strukturanalyse der Häufungspunkte des zweidimensionalen IPF-Verfahrens

Auf die Untersuchung der IPF-Folge in den vorangegangenen Kapiteln folgt in diesem Kapitel die Analyse ihrer Häufungspunkte. Dazu werden im ersten Abschnitt die Zusammenhangsstruktur und die Randsummen der Häufungspunkte untersucht. Im zweiten Abschnitt werden die Häufungspunkte mittels der I-Divergenz charakterisiert. Der dritte Abschnitt zeigt für den Konvergenzfall die stetige Abhängigkeit der Limesmatrix von der Ausgangsmatrix mithilfe der Stetigkeit der I_1-Projektion. Für den Divergenzfall folgt daraus unter gewissen Regularitätsbedingungen die stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix. Im vierten Abschnitt wird mithilfe eines optimierungstheoretischen Ansatzes und unter Beschränkung auf Direktanpassungsprobleme die stetige Abhängigkeit der Limesmatrix von der Ausgangsmatrix und den Marginalien bewiesen. Ebenfalls unter gewissen Regularitätsbedingungen folgt daraus für den Divergenzfall die stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix und den Marginalien.

4.1. Zusammenhangsstruktur und Randsummen der Häufungspunkte

In Satz 3.3.1 (iv) (siehe S. 46) wurde gezeigt, dass die beiden Häufungspunkte C^* und R^* der IPF-Folge die Gleichung $\mathrm{I}(C^*|R^*) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ erfüllen. Somit sind C^* und R^* maßtheoretisch äquivalent (siehe Satz 3.1.1 (iii), S. 45). Darauf aufbauend zeigt der folgende Satz, dass C^* und R^* dieselbe blockdiagonale Zusammenhangsstruktur besitzen und dass von dieser auf eine Blockdreiecksform der Ausgangsmatrix A geschlossen werden kann.

Satz 4.1.1 (Zusammenhangsstruktur der Häufungspunkte)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem und seien $C^* := \lim_{t=0,2,4,\dots} A(t)$ und $R^* := \lim_{t=0,2,4,\dots} A(t+1)$ die Häufungspunkte der zugehörigen IPF-Folge $(A(t))$. Dann existieren disjunkte Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1, \dots, k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \dots, \ell\}$, sodass die Matrizen A, C^* und R^* die Form*

$$A = \begin{pmatrix} A^{(I_1,J_1)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A^{(I_M,J_1)} & \dots & \dots & A^{(I_M,J_M)} \end{pmatrix} \text{ sowie} \tag{4.1.1}$$

$$C^* = \begin{pmatrix} C^{*(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & C^{*(I_M,J_M)} \end{pmatrix},\ R^* = \begin{pmatrix} R^{*(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R^{*(I_M,J_M)} \end{pmatrix} \tag{4.1.2}$$

annehmen. Dabei sind für alle $m = 1, \dots, M$ *die Diagonalblöcke* $A^{(I_m, J_m)}$, $C^{*(I_m, J_m)}$ *und* $R^{*(I_m, J_m)}$ *jeweils zusammenhängend.*

Beweis. Der Häufungspunkt C^* ist die biproportionale Limesanpassung der Ausgangsmatrix A an die Zeilenmarginalien $(c^*_{1+}, \dots, c^*_{k+})$ und an die Spaltenmarginalien c. Folglich existieren nach Satz 1.4.2 (siehe S. 19) disjunkte Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1, \dots, k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \dots, \ell\}$, sodass die Matrizen A und C^* die in den Gleichungen (4.1.1) und (4.1.2) dargestellte Form annehmen. Da die Häufungspunkte C^* und R^* nach Satz 3.1.1 (iii) (siehe S. 45) maßtheoretisch äquivalent sind, hat auch R^* die in Gleichung (4.1.2) dargestellte Form. □

Die soeben gezeigte Zusammenhangsstruktur der Häufungspunkte ermöglicht den Beweis des folgenden Satzes über die Randsummen der Häufungspunkte.

Satz 4.1.2 (Randsummen der Häufungspunkte, vgl. Reffel 2014, Lemma 9.3)**.** *Sei* (A, c, r) *ein normiertes biproportionales Anpassungsproblem und seien* $C^* := \lim_{t=0,2,4,\dots} A(t)$ *und* $R^* := \lim_{t=0,2,4,\dots} A(t+1)$ *die Häufungspunkte der zugehörigen IPF-Folge* $(A(t))$*. Außerdem seien* $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1, \dots, k\}$ *und* $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \dots, \ell\}$ *Zerlegungen im Sinne von Satz 4.1.1. Dann gelten für alle Zusammenhangskomponenten* $m = 1, \dots, M$ *die beiden Aussagen*

$$\forall i \in I_m: \quad c^*_{i+} = r_i \cdot \frac{c_{J_m}}{r_{I_m}} \text{ und} \tag{4.1.3}$$

$$\forall j \in J_m: \quad r^*_{+j} = c_j \cdot \frac{r_{I_m}}{c_{J_m}}. \tag{4.1.4}$$

Beweis. Da C^* die Spaltenanpassung von R^* an c ist und da R^* die Zeilenanpassung von C^* an r ist (siehe Satz 3.1.1, S. 45), gilt

$$c^*_{ij} = \frac{c_j}{r^*_{+j}} \cdot r^*_{ij} = \frac{c_j}{r^*_{+j}} \cdot \frac{r_i}{c^*_{i+}} \cdot c^*_{ij} \text{ für alle Einträge } (i, j). \tag{4.1.5}$$

Folglich gilt

$$\frac{c^*_{i+}}{r_i} = \frac{c_j}{r^*_{+j}} \text{ für alle Einträge } (i, j) \text{ mit } c^*_{ij} > 0. \tag{4.1.6}$$

Aufgrund der Zusammenhangsstruktur der Matrix C^* ist der Quotient c_j / r^*_{+j} somit über alle Spalten $j \in J_m$ derselben Zusammenhangskomponente m konstant. Also gilt

$$\forall m = 1, \dots, M \ \forall i \in I_m: \quad \frac{c^*_{i+}}{r_i} = \frac{c_{J_m}}{r^*_{+J_m}}. \tag{4.1.7}$$

Nach Satz 4.1.1 gilt $r^*_{+J_m} = r_{I_m}$ und somit Gleichung (4.1.3). Die Gleichung (4.1.4) folgt in analoger Weise. □

4.2. Informationstheoretische Charakterisierung der Häufungspunkte

Wenn das IPF-Verfahren konvergiert, dann minimiert die Limesmatrix der IPF-Folge die I-Divergenz $\mathrm{I}(P|A)$ unter allen Matrizen $P \in \mathcal{E}^{\ll A} = \mathcal{C}^{\ll A} \cap \mathcal{R}^{\ll A}$ (siehe Satz 2.3.2 (iii), S. 31). Im Divergenzfall dagegen ist die Schnittmenge $\mathcal{C}^{\ll A} \cap \mathcal{R}^{\ll A}$ leer. Nichtsdestotrotz sollen im Folgenden die Häufungspunkte C^* und R^* der IPF-Folge mittels der I-Divergenz charakterisiert werden. Nach Satz 3.3.1 (iv) (siehe S. 46) gilt $\mathrm{I}(C^*|R^*) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$. Das heißt, das Häufungspunktepaar (C^*, R^*) minimiert die I-Divergenz $\mathrm{I}(C|R)$ unter allen Matrizenpaaren $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$. Dies ist jedoch keine eindeutige Charakterisierung der Häufungspunkte, da die I-Divergenz $\mathrm{I}(C|R)$ auch für andere Matrizenpaare $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ den Wert $\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ annehmen kann. Der folgende Satz zeigt zunächst, wie der Wert $\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ bestimmt werden kann.

Satz 4.2.1 (Minimale I-Divergenz)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem, seien $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen und seien $C^* := \lim_{t=0,2,4,\dots} A(t)$ und $R^* := \lim_{t=0,2,4,\dots} A(t+1)$ die Häufungspunkte der zugehörigen IPF-Folge $(A(t))$. Außerdem seien $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1, \dots, k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \dots, \ell\}$ Zerlegungen im Sinne von Satz 4.1.1. Dann gilt*

$$\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) = \mathrm{I}(C^*|R^*) = \sum_j c_j \ln \frac{c_j}{r^*_{+j}} = \sum_{m=1}^{M} c_{J_m} \ln \frac{c_{J_m}}{r_{I_m}}. \tag{4.2.1}$$

Beweis. Nach Satz 3.1.1 (i) (siehe S. 45) ist C^* die Spaltenanpassung von R^* an c. Folglich gilt

$$\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) = \mathrm{I}(C^*|R^*) = \sum_{i,j:\, c^*_{ij}>0} c^*_{ij} \ln \frac{c_j}{r^*_{+j}} = \sum_j c_j \ln \frac{c_j}{r^*_{+j}} = \sum_{m=1}^{M} \sum_{j \in J_m} c_j \ln \frac{c_j}{r^*_{+j}}. \tag{4.2.2}$$

Gemäß Satz 4.1.2 folgt außerdem

$$\sum_{m=1}^{M} \sum_{j \in J_m} c_j \ln \frac{c_j}{r^*_{+j}} = \sum_{m=1}^{M} \sum_{j \in J_m} c_j \ln \frac{c_{J_m}}{r_{I_m}} = \sum_{m=1}^{M} c_{J_m} \ln \frac{c_{J_m}}{r_{I_m}}. \tag{4.2.3}$$

Damit ist die Gleichungskette (4.2.1) gezeigt. □

Als nächstes werden die Matrizenpaare $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ charakterisiert, die die Gleichung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ erfüllen. Aufgrund von Satz 3.1.1 (i) (siehe S. 45) können wir uns dabei auf die Matrizen $R \in \mathcal{R}^{\ll A}$ konzentrieren. Gilt $\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ für eine Matrix $R \in \mathcal{R}^{\ll A}$, so folgt $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ für eine Matrix $C \in \mathcal{C}^{\ll A}$ genau dann, wenn C die Spaltenanpassung von R an die Spaltenmarginalien c ist. Das folgende Lemma ermöglicht es, die entscheidende Gleichung $\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ zu überprüfen, ohne die beiden Minima $\mathrm{I}(\mathcal{C}^{\ll A}|R)$ und $\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ zu bestimmen.

Lemma 4.2.2 (Spaltensummen)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem und seien $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen. Außerdem sei $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \dots, \ell\}$ eine Zerlegung im Sinne von Satz 4.1.1. Dann gilt für alle $R \in \mathcal{R}^{\ll A}$ die Äquivalenzaussage*

$$\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) \quad \Leftrightarrow \quad \forall m = 1, \dots, M \;\forall j \in J_m : \quad r_{+j} = c_j \cdot \frac{r_{I_m}}{c_{J_m}}. \tag{4.2.4}$$

Beweis. Der Beweis unterscheidet zwei Fälle.

I. Falls $r_{+j} = 0$ für eine der Spalten j gilt, gelten $\mathrm{I}(\mathcal{C}^{\ll A}|R) = \infty > \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ und
$$\exists m \in \{1, \dots, M\}\ \exists j \in J_m : \quad r_{+j} = 0 < c_j \cdot \frac{r_{I_m}}{c_{J_m}}. \tag{4.2.5}$$
Damit ist die Äquivalenzaussage (4.2.4) für diesen Fall gezeigt.

II. Falls $r_{+j} > 0$ für alle Spalten j gilt, gilt $\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(C|R)$, wobei C die Spaltenanpassung von R an c ist. Daraus folgt die Gleichung
$$\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(C|R) = \sum_{i,j:\, c_{ij}>0} c_{ij} \ln \frac{c_j}{r_{+j}} = \sum_j c_j \ln \frac{c_j}{r_{+j}} = \mathrm{I}(c|(r_{+1}, \dots, r_{+\ell})). \tag{4.2.6}$$
Des Weiteren gilt nach Satz 4.2.1 gilt die Gleichung
$$\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) = \sum_j c_j \ln \frac{c_j}{r^*_{+j}} = \mathrm{I}(c|(r^*_{+1}, \dots, r^*_{+\ell})). \tag{4.2.7}$$
Aus diesen beiden Gleichungen folgt die Ungleichung
$$\mathrm{I}(c|(r_{+1}, \dots, r_{+\ell})) = \mathrm{I}(\mathcal{C}^{\ll A}|R) \geq \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) = \mathrm{I}(c|(r^*_{+1}, \dots, r^*_{+\ell})). \tag{4.2.8}$$
Darüber hinaus folgt die Äquivalenzaussage
$$\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) \quad\Leftrightarrow\quad \mathrm{I}(c|(r_{+1}, \dots, r_{+\ell})) = \mathrm{I}(c|(r^*_{+1}, \dots, r^*_{+\ell})). \tag{4.2.9}$$
Wir betrachten nun die konvexe Menge
$$\Gamma := \left\{(\tilde{r}_{+1}, \dots, \tilde{r}_{+\ell}) \,\middle|\, \widetilde{R} \in \mathcal{R}^{\ll A}\right\}. \tag{4.2.10}$$
Für alle $\gamma \in \Gamma$ mit $\gamma_j = 0$ für einen Eintrag j gilt
$$\mathrm{I}(c|\gamma) = \infty > \mathrm{I}(c|(r^*_{+1}, \dots, r^*_{+\ell})). \tag{4.2.11}$$
Für alle anderen $\gamma \in \Gamma$ folgt aus Ungleichung (4.2.8) die Ungleichung
$$\mathrm{I}(c|\gamma) \geq \mathrm{I}(c|(r^*_{+1}, \dots, r^*_{+\ell})). \tag{4.2.12}$$
Aufgrund der strikten Konvexität der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (ii), S. 23) ist $(r^*_{+1}, \dots, r^*_{+\ell}) \in \Gamma$ somit das eindeutige Minimum der Funktion $\mathrm{I}(c|\gamma)$ über alle $\gamma \in \Gamma$. Daraus folgt die Äquivalenzaussage
$$\mathrm{I}(c|(r_{+1}, \dots, r_{+\ell})) = \mathrm{I}(c|(r^*_{+1}, \dots, r^*_{+\ell})) \quad\Leftrightarrow\quad (r_{+1}, \dots, r_{+\ell}) = (r^*_{+1}, \dots, r^*_{+\ell}). \tag{4.2.13}$$
Nach Satz 4.1.2 gilt
$$\begin{aligned}&(r_{+1}, \dots, r_{+\ell}) = (r^*_{+1}, \dots, r^*_{+\ell})\\ &\qquad\Leftrightarrow\quad \forall m = 1, \dots, M\ \forall j \in J_m : \quad r_{+j} = c_j \cdot \frac{r_{I_m}}{c_{J_m}}.\end{aligned} \tag{4.2.14}$$
Die Äquivalenzaussagen (4.2.9), (4.2.13) und (4.2.14) implizieren die Äquivalenzaussage (4.2.4). Damit ist die Äquivalenzaussage (4.2.4) auch für diesen Fall gezeigt. □

Das soeben gezeigte Lemma 4.2.2 ermöglicht gemeinsam mit Satz 3.1.1 (siehe S. 45) die folgende Charakterisierung der Matrizenpaare $(C,R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$, die die Gleichung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ erfüllen.

Satz 4.2.3 (Minimierende Matrizenpaare)**.** *Sei (A,c,r) ein normiertes biproportionales Anpassungsproblem und seien $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen. Außerdem seien $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1,\ldots,k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1,\ldots,\ell\}$ Zerlegungen im Sinne von Satz 4.1.1. Des Weiteren sei $(C,R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ beliebig. Dann sind die folgenden drei Aussagen äquivalent:*

(i) Es gilt $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$.

(ii) Die Matrix C ist die Spaltenanpassung von R an die Spaltenmarginalien c und für die Spalten der Matrix R gilt

$$\forall m = 1,\ldots,M \ \forall j \in J_m : \quad r_{+j} = c_j \cdot \frac{r_{I_m}}{c_{J_m}}. \tag{4.2.15}$$

(iii) Die Matrix R ist die Zeilenanpassung von C an die Zeilenmarginalien r und für die Zeilen der Matrix C gilt

$$\forall m = 1,\ldots,M \ \forall i \in I_m : \quad c_{i+} = r_i \cdot \frac{c_{J_m}}{r_{I_m}}. \tag{4.2.16}$$

Darüber hinaus impliziert jede dieser drei Aussagen die folgende Aussage:

(iv) Bis auf geeignete Zeilen- und Spaltenvertauschungen gilt

$$C = \begin{pmatrix} C^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & C^{(I_M,J_M)} \end{pmatrix}, R = \begin{pmatrix} R^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R^{(I_M,J_M)} \end{pmatrix}. \tag{4.2.17}$$

Beweis. Seien $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1,\ldots,k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1,\ldots,\ell\}$ Zerlegungen im Sinne von Satz 4.1.1 und sei $(C,R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ beliebig. Dann hat die Ausgangsmatrix A nach Satz 4.1.1 Blockdreiecksform. Diese Blockdreiecksform überträgt sich nach Definition der Mengen $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ auf die Matrizen C und R, sodass gilt

$$C = \begin{pmatrix} C^{(I_1,J_1)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ C^{(I_M,J_1)} & \cdots & \cdots & C^{(I_M,J_M)} \end{pmatrix}, R = \begin{pmatrix} R^{(I_1,J_1)} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ R^{(I_M,J_1)} & \cdots & \cdots & R^{(I_M,J_M)} \end{pmatrix}. \tag{4.2.18}$$

(i) $\Rightarrow$ (ii) : Gelte Aussage (i). Somit gilt $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$. Folglich ist nach Satz 3.1.1 (i) (siehe S. 45) die Matrix C die Spaltenanpassung von R an die Spaltenmarginalien c. Außerdem gilt die Gleichung $\mathrm{I}(\mathcal{C}^{\ll A}|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$. Daraus folgt mit Lemma 4.2.2 die Gleichung (4.2.15). Damit gilt Aussage (ii).

(ii) ⇒ (iii) : Gelte Aussage (ii). Somit ist C ist die Spaltenanpassung von R an die Spaltenmarginalien c und es gilt die Gleichung (4.2.15). Aus Gleichung (4.2.15) folgt

$$\forall m = 1, \ldots, M: \quad r_{+J_m} = c_{J_m} \cdot \frac{r_{I_m}}{c_{J_m}} = r_{I_m}. \tag{4.2.19}$$

Nach Definition der Menge $\mathcal{R}^{\ll A}$ gilt außerdem

$$\forall m = 1, \ldots, M: \quad r_{I_m+} = r_{I_m}. \tag{4.2.20}$$

Daraus folgt mit Gleichung (4.2.18) die Aussage

$$R = \begin{pmatrix} R^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R^{(I_M,J_M)} \end{pmatrix}. \tag{4.2.21}$$

Da C die Spaltenanpassung von R an die Spaltenmarginalien c ist, folgt außerdem

$$C = \begin{pmatrix} C^{(I_1,J_1)} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & C^{(I_M,J_M)} \end{pmatrix}. \tag{4.2.22}$$

Des Weiteren folgt mit Gleichung (4.2.15) die Aussage

$$\forall m = 1, \ldots, M \ \forall i \in I_m \ \forall j \in J_m: \quad c_{ij} = r_{ij} \cdot \frac{c_j}{r_{+j}} = r_{ij} \cdot \frac{c_{J_m}}{r_{I_m}}. \tag{4.2.23}$$

Folglich gilt

$$\begin{aligned} \forall m = 1, \ldots, M \ \forall i \in I_m: \quad c_{i+} &= \sum_{j \in J_m} c_{ij} = \sum_{j \in J_m} r_{ij} \cdot \frac{c_{J_m}}{r_{I_m}} = \\ &= r_{iJ_m} \cdot \frac{c_{J_m}}{r_{I_m}} = r_{i+} \cdot \frac{c_{J_m}}{r_{I_m}} = r_i \cdot \frac{c_{J_m}}{r_{I_m}}. \end{aligned} \tag{4.2.24}$$

Somit ist die Gleichung (4.2.16) gezeigt. Außerdem folgt aus den Gleichungen (4.2.23) und (4.2.24) die Aussage

$$\forall m = 1, \ldots, M \ \forall i \in I_m \ \forall j \in J_m: \quad r_{ij} = c_{ij} \cdot \frac{r_{I_m}}{c_{J_m}} = c_{ij} \cdot \frac{r_i}{c_{i+}}. \tag{4.2.25}$$

Also ist R die Zeilenanpassung von C an die Zeilenmarginalien r. Damit gilt Aussage (iii).

(iii) ⇒ (i) : Gelte Aussage (iii). Somit ist R die Zeilenanpassung von C an die Zeilenmarginalien r und es gilt die Gleichung (4.2.16). Daraus folgt

$$\begin{aligned} \mathrm{I}(C|R) &= \sum_{i,j:\, c_{ij}>0} c_{ij} \ln \frac{c_{i+}}{r_i} = \sum_{m=1}^{M} \sum_{i \in I_m} c_{i+} \ln \frac{c_{i+}}{r_i} \\ &= \sum_{m=1}^{M} \sum_{i \in I_m} \left(r_i \cdot \frac{c_{J_m}}{r_{I_m}} \right) \ln \frac{c_{J_m}}{r_{I_m}} = \sum_{m=1}^{M} c_{J_m} \ln \frac{c_{J_m}}{r_{I_m}}. \end{aligned} \tag{4.2.26}$$

Mit Satz 4.2.1 folgt weiter

$$\mathrm{I}(C|R) = \sum_{m=1}^{M} c_{J_m} \ln \frac{c_{J_m}}{r_{I_m}} = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}). \tag{4.2.27}$$

Damit gilt Aussage (i).

(iii) ⇒ (iv) : Diese Folgerung wurde bereits im Beweisschritt „(ii) ⇒ (iii)“ gezeigt. □

Im Gegensatz zu den Blöcken $C^{*(I_m,J_m)}$ und $R^{*(I_m,J_m)}$ in Satz 4.1.1 sind die Blöcke $C^{(I_m,J_m)}$ und $R^{(I_m,J_m)}$ in Satz 4.2.3 (iv) nicht notwendig zusammenhängend. Sie können stattdessen weiter zerfallen. Dies zeigt das folgende Beispiel.

Beispiel 4.2.4 (Weiterer Zerfall). Gegeben seien die Zeilenmarginalien $r = (0.3, 0.3, 0.4)$ und die Spaltenmarginalien $c = (0.4, 0.4, 0.2)$ sowie die Ausgangsmatrix

$$A = \begin{pmatrix} 0.1 & 0.1 & 0 \\ 0.1 & 0.1 & 0 \\ 0 & 0 & 0.6 \end{pmatrix}. \tag{4.2.28}$$

Die daraus resultierende IPF-Folge $(A(t))$ hat die beiden Häufungspunkte

$$C^* = \begin{pmatrix} 0.2 & 0.2 & 0 \\ 0.2 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{pmatrix} \text{ und } R^* = \begin{pmatrix} 0.15 & 0.15 & 0 \\ 0.15 & 0.15 & 0 \\ 0 & 0 & 0.4 \end{pmatrix}. \tag{4.2.29}$$

Daraus folgt

$$\mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}) = \mathrm{I}(C^*|R^*) = 4 \cdot 0.2 \ln \frac{0.2}{0.15} + 0.2 \ln \frac{0.2}{0.4} = 0.8 \ln \frac{4}{3} + 0.2 \ln \frac{1}{2}. \tag{4.2.30}$$

Wir wählen nun

$$C := \begin{pmatrix} 0.4 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \end{pmatrix} \in \mathcal{C}^{\ll A} \text{ und } R := \begin{pmatrix} 0.3 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.4 \end{pmatrix} \in \mathcal{R}^{\ll A}. \tag{4.2.31}$$

Für das so gewählte Matrizenpaar (C, R) gilt

$$\mathrm{I}(C|R) = 2 \cdot 0.4 \ln \frac{0.4}{0.3} + 0.2 \ln \frac{0.2}{0.4} = 0.8 \ln \frac{4}{3} + 0.2 \ln \frac{1}{2} = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A}). \tag{4.2.32}$$

Das Matrizenpaar (C, R) erfüllt die Gleichung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ und übernimmt folglich die Nullblöcke von (C^*, R^*). Der linke obere Block zerfällt allerdings weiter.

Während der vorangegangene Satz 4.2.3 die gemeinsamen Eigenschaften aller die Gleichung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ erfüllenden Matrizenpaare $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ beschreibt, zeigt der folgende Satz die charakteristischen Eigenschaften des Häufungspunktepaares (C^*, R^*).

Satz 4.2.5 (Charakterisierung)**.** *Sei (A, c, r) ein normiertes biproportionales Anpassungsproblem, seien $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ die zugehörigen Mengen von Wahrscheinlichkeitsmatrizen und seien $C^* := \lim_{t=0,2,4,\ldots} A(t)$ und $R^* := \lim_{t=0,2,4,\ldots} A(t+1)$ die Häufungspunkte der zugehörigen IPF-Folge $(A(t))$. Außerdem sei $(C, R) \in \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A}$ ein Matrizenpaar, das die Gleichung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ erfüllt. Dann gilt*

$$\mathrm{I}(C|A) \geq \mathrm{I}(C^*|A). \tag{4.2.33}$$

Gleichheit gilt genau dann, wenn $C = C^*$ gilt.

Beweis. Der Häufungspunkt $C^* \in \mathcal{C}^{\ll A}$ ist die biproportionale Limesanpassung der Ausgangsmatrix A an die Zeilenmarginalien $(c^*_{1+}, \ldots, c^*_{k+})$ und die Spaltenmarginalien c. Nach Satz 2.3.4 (ii) (siehe S. 34) ist C^* somit die I_1-Projektion von A auf die Menge

$$\widetilde{\mathcal{E}}^{\ll A} := \left\{ P \in \mathcal{C}^{\ll A} \,\middle|\, \forall i : p_{i+} = c^*_{i+} \right\}. \tag{4.2.34}$$

Nach Satz 4.2.3 ist die Matrix $C \in \mathcal{C}^{\ll A}$ ebenfalls in der Menge $\widetilde{\mathcal{E}}^{\ll A}$ enthalten. Folglich gilt Ungleichung (4.2.33). Wegen der Eindeutigkeit der I_1-Projektion (siehe Satz 2.2.3, S. 27) gilt Gleichheit genau dann, wenn $C = C^*$ gilt. □

Mit den Sätzen 4.2.3 und 4.2.5. ist die angestrebte informationstheoretische Charakterisierung der Häufungspunkte erreicht: Zwei Matrizen $C \in \mathcal{C}^{\ll A}$ und $R \in \mathcal{R}^{\ll A}$ sind genau dann die Häufungspunkte der IPF-Folge, wenn sie erstens die Gleichung $\mathrm{I}(C|R) = \mathrm{I}(\mathcal{C}^{\ll A}|\mathcal{R}^{\ll A})$ erfüllen und zweitens unter allen diese Gleichung erfüllenden Matrizenpaaren die minimale I-Divergenz $\mathrm{I}(C|A)$ aufweisen. Das folgende Beispiel unterstreicht die Bedeutung von Satz 4.2.5.

Beispiel 4.2.6 (Charakterisierung). Wir betrachten erneut die Ausgangsmatrix A sowie die Marginalien c und r aus Beispiel 4.2.4. Neben den dort betrachteten Matrizenpaaren (C^*, R^*) und (C, R) existieren noch weitere Matrizenpaare, die die I-Divergenz minimieren. Die I-Divergenz $\mathrm{I} : \mathcal{C}^{\ll A} \times \mathcal{R}^{\ll A} \to [0; \infty]$ wird von allen Paaren (C_α, R_α), $\alpha \in [0; 1]$, minimiert, für die gilt

$$C_\alpha := \begin{pmatrix} 0.4(1-\alpha) & 0.4\alpha & 0 \\ 0.4\alpha & 0.4(1-\alpha) & 0 \\ 0 & 0 & 0.2 \end{pmatrix} \text{ und } R_\alpha := \begin{pmatrix} 0.3(1-\alpha) & 0.3\alpha & 0 \\ 0.3\alpha & 0.3(1-\alpha) & 0 \\ 0 & 0 & 0.4 \end{pmatrix}. \tag{4.2.35}$$

Allerdings gilt für alle $\alpha \in [0; 1]$ die Gleichung

$$\mathrm{I}(C_\alpha|A) = 0.8 \ln 4 + 0.2 \ln \frac{1}{3} + 0.8\left((1-\alpha)\ln(1-\alpha) + \alpha \ln \alpha\right). \tag{4.2.36}$$

Diese I-Divergenz wird eindeutig minimiert von $C_{0.5} = C^*$.

4.3. Stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix

Bei den Einträgen der Ausgangsmatrix handelt es sich in vielen Anwendungen des IPF-Verfahrens um beobachtete Werte, die gewisse Fehler beinhalten können. Dementsprechend stellt sich die Frage, ob die biproportionale Limesanpassung stetig von der Ausgangsmatrix abhängt. Der folgende Satz nutzt die Äquivalenz von biproportionalen Limesanpassungen und I_1-Projektionen (siehe Satz 2.3.4, S. 34) sowie die Stetigkeit von I_1-Projektionen (siehe Satz 2.2.4, S. 28), um diese Abhängigkeit zu zeigen.

Satz 4.3.1 (Stetige Abhängigkeit der biproportionalen Limesanpassung von der Ausgangsmatrix)**.** *Sei (A, c, r) ein biproportionales Anpassungsproblem und sei B die zugehörige biproportionale Limesanpassung gemäß Abschnitt 1.2 (siehe S. 13). Weiter sei (A^n) eine Folge von Matrizen in $[0; \infty)^{k\times\ell}$, die gegen A konvergiert. Dann gelten die folgenden beiden Aussagen:*

(i) Für hinreichend große $n \in \mathbb{N}$ hat das biproportionale Anpassungsproblem (A^n, c, r) eine biproportionale Limesanpassung B^n.

(ii) Die Folge (B^n) der biproportionalen Limesanpassungen konvergiert gegen B,

$$\lim_{n\to\infty} B^n = B. \tag{4.3.1}$$

Beweis. (i) Wir betrachten die zu den biproportionalen Anpassungsproblemen (A, c, r) und (A^n, c, r), $n \in \mathbb{N}$, gehörenden Matrizenmengen

$$\mathcal{E}^{\ll A} = \left\{S \in [0; \infty)^{k\times\ell} \,\middle|\, S \ll A \text{ und } \forall i : s_{i+} = r_i \text{ und } \forall j : s_{+j} = c_j\right\} \text{ und} \tag{4.3.2}$$

$$\mathcal{E}^{\ll A^n} = \left\{S \in [0; \infty)^{k\times\ell} \,\middle|\, S \ll A^n \text{ und } \forall i : s_{i+} = r_i \text{ und } \forall j : s_{+j} = c_j\right\}. \tag{4.3.3}$$

Nach Voraussetzung ist B die biproportionale Limesanpassung von A an c und r. Dementsprechend ist B in der Menge $\mathcal{E}^{\ll A}$ enthalten. Für hinreichend große $n \in \mathbb{N}$ gilt weiter $A \ll A^n$ und damit $\mathcal{E}^{\ll A} \subseteq \mathcal{E}^{\ll A^n}$. Folglich ist B ein Element der Menge $\mathcal{E}^{\ll A^n}$. Nach Satz 2.6.2 (siehe S. 40) existiert somit eine biproportionale Limesanpassung B^n von A^n an c und r.

(ii) Ohne Beschränkung der Allgemeinheit seien die Anpassungsprobleme (A, c, r) und (A^n, c, r), $n \in \mathbb{N}$, normiert. Dann ist die Limesanpassung B nach Satz 2.3.4 (ii) (siehe S. 34) die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$. Da die I-Divergenz $\mathrm{I}(S|A)$ für alle nicht von A dominierten Matrizen S unendlich ist, ist B außerdem die I_1-Projektion von A auf die Menge

$$\mathcal{E} = \left\{S \in [0; \infty)^{k\times\ell} \,\middle|\, \forall i : s_{i+} = r_i \text{ und } \forall j : s_{+j} = c_j\right\}. \tag{4.3.4}$$

In analoger Weise sind die Limesanpassungen B^n die I_1-Projektionen von A^n auf $\mathcal{E}$. Aus der Stetigkeit der I_1-Projektion auf $\mathcal{E}$ (siehe Satz 2.2.4, S. 28) folgt die Konvergenz der Matrizenfolge (B^n) gegen die Matrix B. □

Im Divergenzfall stellt sich die Frage nach der stetigen Abhängigkeit der Häufungspunkte der IPF-Folge von der Ausgangsmatrix. Wie der folgende Satz zeigt, kann diese Frage unter gewissen Regularitätsbedingungen mit Ja beantwortet werden.

Satz 4.3.2 (Stetige Abhängigkeit der Häufungspunkte der IPF-Folge von der Ausgangsmatrix). *Sei (A, c, r) ein biproportionales Anpassungsproblem und seien $C^* := \lim_{t=0,2,4,\dots} A(t)$ und $R^* := \lim_{t=0,2,4,\dots} A(t+1)$ die Häufungspunkte der zugehörigen IPF-Folge $(A(t))$. Weiter sei (A^n) eine Folge von Matrizen in $[0;\infty)^{k\times\ell}$ mit positiven Zeilensummen, $a^n_{i+} > 0$, und positiven Spaltensummen, $a^n_{+j} > 0$, die gegen A konvergiert. Außerdem bezeichnen $C^{n,*} := \lim_{t=0,2,4,\dots} A^n(t)$ und $R^{n,*} := \lim_{t=0,2,4,\dots} A^n(t+1)$ die Häufungspunkte der zu dem biproportionalen Anpassungsproblem (A^n, c, r) gehörenden IPF-Folgen $(A^n(t))$. Dann gilt: Wenn die Matrizen $C^{n,*}$ und $R^{n,*}$ dieselbe Zusammenhangsstruktur haben wie die Matrizen C^* und R^*, dann konvergiert die Häufungspunktefolge $(C^{n,*})$ gegen C^* und die Häufungspunktefolge $(R^{n,*})$ gegen R^*.*

Beweis. Da die Häufungspunkte $C^{n,*}$ und $R^{n,*}$ nach Voraussetzung dieselbe Zusammenhangsstruktur haben wie die Häufungspunkte C^* und R^*, existieren gemeinsame Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1, \dots, k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1, \dots, \ell\}$ im Sinne von Satz 4.1.1. Damit gelten nach Satz 4.1.2 für alle $n \in \mathbb{N}$ und für alle Zusammenhangskomponenten $m = 1, \dots, M$ die beiden Aussagen

$$\forall i \in I_m : \quad c^{n,*}_{i+} = c^*_{i+} = r_i \cdot \frac{c_{J_m}}{r_{I_m}} \text{ und} \tag{4.3.5}$$

$$\forall j \in J_m : \quad r^{n,*}_{+j} = r^*_{+j} = c_j \cdot \frac{r_{I_m}}{c_{J_m}}. \tag{4.3.6}$$

Folglich ist der Häufungspunkt C^* die biproportionale Limesanpassung der Ausgangsmatrix A an die Zeilenmarginalien $(c^*_{1+}, \dots, c^*_{k+})$ und die Spaltenmarginalien c und die Häufungspunkte $C^{n,*}$ sind die biproportionalen Limesanpassungen der Ausgangsmatrizen A^n an dieselben Marginalien. Also konvergiert nach Satz 4.3.1 die Häufungspunktefolge $(C^{n,*})$ gegen C^*. Die Konvergenz der Häufungspunktefolge $(R^{n,*})$ gegen R^* folgt analog. □

4.4. Stetige Abhängigkeit der Häufungspunkte von der Ausgangsmatrix und den Marginalien

Im vorangegangenen Abschnitt wurde die stetige Abhängigkeit der biproportionalen Limesanpassung B von der Ausgangsmatrix A untersucht. Genau wie die Matrix A können auch die Marginalien c und r Fehler enthalten. Der im vorangegangenen Abschnitt verwendete Ansatz lässt allerdings keine Veränderung der Marginalien c und r zu, da die Menge $\mathcal{E}$ von diesen Marginalien abhängt und die für den Beweis der Stetigkeit der I_1-Projektion benötigte Ungleichung $\mathrm{I}(x^{n_m}|y^{n_m}) \leq \mathrm{I}(x|y^{n_m})$ (beziehungsweise $\mathrm{I}(B^{n_m}|A^{n_m}) \leq \mathrm{I}(B^{n_m}|A)$ in Matrizenschreibweise) für veränderliches $\mathcal{E}$ nicht garantiert werden kann. Um die Abhängigkeit der biproportionalen Limesanpassung B von dem gesamten Anpassungsproblem (A, c, r) zu untersuchen, wird daher in diesem Abschnitt ein neuer Ansatz vorgestellt. Die Äquivalenz von biproportionalen Limesanpassungen und I_1-Projektionen (siehe Satz 2.3.4, S. 34) erlaubt die Interpretation jedes Limesanpassungsproblems (A, c, r) als ein Minimierungsproblem, dessen Zielfunktion und dessen Nebenbedingungen von den Parametern A, c und r abhängen. Diese Interpretation motiviert die Betrachtung sogenannter *verallgemeinerter konvexer Optimierungsprobleme*. Diese Probleme werden von Rockafellar (1972, Abschnitt 29) untersucht. Dabei verwendet Rockafellar das Konzept der Bifunktion. Dieses Konzept hat ursprünglich nichts mit biproportionalen Anpassungsproblemen zu tun. Bei Beschränkung auf Direktanpassungsprobleme ermöglicht es jedoch den Beweis der stetigen Abhängigkeit der Limesmatrix der IPF-Folge von der Ausgangsmatrix und den Marginalien.

Definition 4.4.1 (Bifunktion, vgl. Rockafellar 1972, Abschnitt 29)**.** Eine *Bifunktion von* $\mathbb{R}^p$ *nach* $\mathbb{R}^q$ ist eine Abbildung F, die jedem Vektor $u \in \mathbb{R}^p$ eine Funktion $Fu : \mathbb{R}^q \to [-\infty; +\infty]$ zuweist. Diese Funktion Fu bildet wiederum jeden Vektor $x \in \mathbb{R}^q$ auf einen Wert $(Fu)(x) \in [-\infty; +\infty]$ ab. Die Bifunktion F heißt *konvex*, wenn die Funktion $(u, x) \mapsto (Fu)(x)$ konvex ist. Der *eigentliche Definitionsbereich der Bifunktion* F, $\operatorname{dom} F$, ist die Menge aller Vektoren $u \in \mathbb{R}^p$, für die die Funktion Fu nicht konstant unendlich ist,

$$\operatorname{dom} F := \left\{u \in \mathbb{R}^p \,\middle|\, \exists x \in \mathbb{R}^q : (Fu)(x) < \infty\right\}. \tag{4.4.1}$$

Die Funktion $(\inf F) : \mathbb{R}^p \to [-\infty; +\infty]$, $u \mapsto \inf_{x\in\mathbb{R}^q}(Fu)(x)$, die jeden Parameter $u \in \mathbb{R}^p$ auf den zugehörigen Optimalwert $\inf_{x\in\mathbb{R}^q}(Fu)(x) \in [-\infty; +\infty]$ abbildet, wird als *Perturbationsfunktion* bezeichnet. Des Weiteren beschreibt der Operator int das *Innere einer Menge*. Für jede beliebige Teilmenge $\mathcal{M}$ eines euklidischen Vektorraums gilt

$$\operatorname{int} \mathcal{M} := \{m \in \mathcal{M} \mid \exists \epsilon > 0 : m + B(0, \epsilon) \subseteq \mathcal{M}\}. \tag{4.4.2}$$

Dabei bezeichnet $B(0, \epsilon) := \{x \mid \|x\| \leq \epsilon\}$ die ϵ-Kugel um den Nullpunkt.

Die entscheidende Rolle bei der Analyse der Abhängigkeit der biproportionalen Limesanpassung vom gesamten Anpassungsproblem wird der Perturbationsfunktion $\inf F$ zuteil. Deren Stetigkeit wird von Rockafellar (1972) untersucht.

Satz 4.4.2 (Stetigkeit der Perturbationsfunktion, vgl. Rockafellar 1972, Korollar 29.1.5)**.** *Sei F eine konvexe Bifunktion von $\mathbb{R}^p$ nach $\mathbb{R}^q$ und gelte* $0 \in \operatorname{int}(\operatorname{dom} F) \subseteq \mathbb{R}^p$. *Dann existiert eine offene, konvexe Umgebung $\mathcal{U}$ von* $0 \in \mathbb{R}^p$*, auf der die Perturbationsfunktion* $(\inf F) : \mathbb{R}^p \to [-\infty; +\infty]$ *stetig ist.*

Beweis. Siehe Rockafellar (1972, Korollar 29.1.5). □

Um Satz 4.4.2 anwenden zu können, beschränken wir uns auf Folgen $((A^n, c^n, r^n))$ von Anpassungsproblemen, die gegen ein Direktanpassungsproblem (A, c, r) konvergieren. Des Weiteren nehmen wir an, dass $A \equiv A^n$ und $c^n_+ = r^n_+$ für alle $n \in \mathbb{N}$ gilt. Somit verlieren wir einen Freiheitsgrad und variieren nur c^n und $(r^n_1, \ldots, r^n_{k-1})$, während wir $r^n_k := c^n_+ - (r^n_1 + \cdots + r^n_{k-1})$ setzen. Das Hauptresultat dieses Abschnitts lautet dann wie folgt.

Satz 4.4.3 (Stetige Abhängigkeit der biproportionalen Direktanpassung von der Ausgangsmatrix und den Marginalien)**.** *Sei (A, c, r) ein biproportionales Direktanpassungsproblem und sei B die zugehörige biproportionale Direktanpassung gemäß Abschnitt 1.2 (siehe S. 13). Weiter sei $((A^n, c^n, r^n))$ eine Folge von biproportionalen Anpassungsproblemen, die gegen (A, c, r) konvergiert. Außerdem gelte $A^n \equiv A$ und $c^n_+ = r^n_+$ für alle $n \in \mathbb{N}$. Dann gelten die folgenden beiden Aussagen:*

(i) Für hinreichend große $n \in \mathbb{N}$ hat das biproportionale Anpassungsproblem (A^n, c^n, r^n) eine biproportionale Direktanpassung B^n.

(ii) *Die Folge* (B^n) *der biproportionalen Direktanpassungen konvergiert gegen* B,

$$\lim_{n\to\infty} B^n = B. \tag{4.4.3}$$

Beweis. Der Beweis erfolgt in sechs Schritten. Schritt I zeigt Aussage (i). In Schritt II definieren wir eine Bifunktion F. Schritt III zeigt, dass diese Bifunktion konvex ist, und Schritt IV zeigt, dass $0 \in \operatorname{int}(\operatorname{dom} F)$ gilt. Diese beiden Eigenschaften ermöglichen die Anwendung von Satz 4.4.2 in Schritt V. Schritt VI zeigt schließlich Aussage (ii). Ohne Beschränkung der Allgemeinheit nehmen wir für diesen Beweis an, dass die Matrizen A und A^n, $n \in \mathbb{N}$, zusammenhängend sind. Die Kurzschreibweise $\mathbb{R}^{\operatorname{supp}(A)}$ bezeichnet im Folgenden den Untervektorraum der von A dominierten Matrizen,

$$\mathbb{R}^{\operatorname{supp}(A)} := \left\{ S \in \mathbb{R}^{k\times\ell} \,\middle|\, \operatorname{supp}(S) \subseteq \operatorname{supp}(A) \right\} = \left\{ S \in \mathbb{R}^{k\times\ell} \,\middle|\, S \ll A \right\}. \tag{4.4.4}$$

I. Da (A, c, r) ein biproportionales Direktanpassungsproblem ist, gelten nach Satz 2.4.1 (ii) (siehe S. 35) die Gleichung $c_+ = r_+$ sowie die strikten Flussungleichungen

$$\forall I \subsetneq \{1, \ldots, k\},\ I \neq \emptyset : \quad r_I < c_{J_A(I)}. \tag{4.4.5}$$

Nachdem außerdem die Folge (c^n) gegen c konvergiert und die Folge (r^n) gegen r konvergiert, folgen daraus für hinreichend große $n \in \mathbb{N}$ die Ungleichungen

$$\forall I \subsetneq \{1, \ldots, k\},\ I \neq \emptyset : \quad r^n_I < c^n_{J_A(I)} = c^n_{J_{A^n}(I)}. \tag{4.4.6}$$

Die Gleichheit gilt, da nach Voraussetzung $A^n \equiv A$ für alle $n \in \mathbb{N}$ gilt. Weil nach Voraussetzung außerdem $c^n_+ = r^n_+$ für alle $n \in \mathbb{N}$ gilt, folgt mit Satz 2.4.1 (ii) (siehe S. 35) und Satz 2.3.3 (ii) (siehe S. 34) für hinreichend große $n \in \mathbb{N}$ die Existenz der biproportionalen Direktanpassung B^n von A^n an c^n und r^n. Damit ist Aussage (i) gezeigt.

II. Sei $u = (\Delta A, \Delta c, \Delta r) \in \mathbb{R}^{\operatorname{supp}(A)+\ell+(k-1)}$. Weiter sei die Bifunktion F von $\mathbb{R}^{\operatorname{supp}(A)+\ell+(k-1)}$ nach $\mathbb{R}^{k\times\ell}$ definiert gemäß

$$(Fu)(x) := \begin{cases} \mathrm{I}(x|A+\Delta A) + \delta(x|\mathcal{E}_{\Delta c,\Delta r}) & \text{falls } A + \Delta A \geq 0 \text{ und } x \geq 0, \\ \infty & \text{sonst,} \end{cases} \tag{4.4.7}$$

wobei

$$\delta(x|\mathcal{E}_{\Delta c,\Delta r}) := \begin{cases} 0 & \text{falls } x \in \mathcal{E}_{\Delta c,\Delta r}, \\ \infty & \text{sonst,} \end{cases} \tag{4.4.8}$$

und

$$\begin{aligned} \mathcal{E}_{\Delta c,\Delta r} := \Big\{ x \in \mathbb{R}^{k\times\ell}_{\geq 0} \,\Big|\, & \forall i \in \{1, \ldots, k-1\} : \ x_{i+} = r_i + \Delta r_i \\ & \text{und } \forall j \in \{1, \ldots, \ell\} : \ x_{+j} = c_j + \Delta c_j \Big\}. \end{aligned} \tag{4.4.9}$$

III. Nach Satz 2.1.2 (i) (siehe S. 23) ist die I-Divergenz auf ihrem gesamten Definitionsbereich konvex. Folglich ist $\mathrm{I}(x|A + \Delta A)$ als Funktion von $u = (\Delta A, \Delta c, \Delta r)$ und x konvex.

Seien $u^0, u^1 \in \mathbb{R}^{\mathrm{supp}(A)+\ell+(k-1)}$ und $x^0, x^1 \in \mathbb{R}^{k\times\ell}$ sowie $\alpha \in (0; 1)$ beliebig. Außerdem sei $u^\alpha := (1-\alpha)u^0 + \alpha u^1$ und $x^\alpha := (1-\alpha)x^0 + \alpha x^1$. Falls $x^0 \notin \mathcal{E}_{\Delta c^0, \Delta r^0}$ oder $x^1 \notin \mathcal{E}_{\Delta c^1, \Delta r^1}$ gilt, folgt

$$\delta(x^\alpha|\mathcal{E}_{\Delta c^\alpha, \Delta r^\alpha}) \leq \infty = (1-\alpha)\delta(x^0|\mathcal{E}_{\Delta c^0, \Delta r^0}) + \alpha\delta(x^1|\mathcal{E}_{\Delta c^1, \Delta r^1}). \tag{4.4.10}$$

Falls $x^0 \in \mathcal{E}_{\Delta c^0, \Delta r^0}$ und $x^1 \in \mathcal{E}_{\Delta c^1, \Delta r^1}$ gilt, folgt

$$\begin{aligned} \forall i \in \{1, \ldots, k-1\}: \quad x^\alpha_{i+} &= (1-\alpha)x^0_{i+} + \alpha x^1_{i+} \\ &= (1-\alpha)(r_i + \Delta r^0_i) + \alpha(r_i + \Delta r^0_i) \\ &= r_i + \Delta r^\alpha_i. \end{aligned} \tag{4.4.11}$$

In analoger Weise folgt

$$\forall j \in \{1, \ldots, \ell\}: \quad x^\alpha_{+j} = c_j + \Delta c^\alpha_j. \tag{4.4.12}$$

Also gilt $x^\alpha \in \mathcal{E}_{\Delta c^\alpha, \Delta r^\alpha}$ und damit auch

$$\delta(x^\alpha|\mathcal{E}_{\Delta c^\alpha, \Delta r^\alpha}) = 0 = (1-\alpha)\delta(x^0|\mathcal{E}_{\Delta c^0, \Delta r^0}) + \alpha\delta(x^1|\mathcal{E}_{\Delta c^1, \Delta r^1}). \tag{4.4.13}$$

Folglich ist $\delta(x|\mathcal{E}_{\Delta c, \Delta r})$ als Funktion von $u = (\Delta A, \Delta c, \Delta r)$ und x konvex.

Somit ist die Funktion $(u, x) \mapsto (Fu)(x)$ als Summe konvexer Funktionen konvex und F ist eine konvexe Bifunktion.

IV. Sei m die minimale Differenz der strikten Flussungleichungen (4.4.5),

$$m := \min\left\{c_{J_A(I)} - r_I \,\middle|\, I \subsetneq \{1, \ldots, k\}, I \neq \emptyset\right\} > 0. \tag{4.4.14}$$

Weiter sei $u = (\Delta A, \Delta c, \Delta r)$ ein beliebiges Element der Menge

$$\begin{aligned} \widetilde{\mathcal{U}} := \left(-\min_{(i,j)\in\mathrm{supp}(A)} \frac{a_{ij}}{2}; \infty\right)^{\mathrm{supp}(A)} &\times \left[-\frac{m}{4k\ell}; \frac{m}{4k\ell}\right]^\ell \\ &\times \left[-\frac{m}{4(k-1)k}; \frac{m}{4(k-1)k}\right]^{k-1}. \end{aligned} \tag{4.4.15}$$

Außerdem sei

$$\Delta r_k := \Delta c_+ - \sum_{i=1}^{k-1} \Delta r_i. \tag{4.4.16}$$

Daraus folgt $A + \Delta A \geq 0$ und $\mathrm{supp}\,(A + \Delta A) = \mathrm{supp}\,(A)$ sowie

$$|\Delta r_k| \leq |\Delta c_+| + \sum_{i=1}^{k-1} |\Delta r_i| \leq \ell \cdot \frac{m}{4k\ell} + (k-1)\cdot\frac{m}{4(k-1)k} = \frac{m}{2k}. \tag{4.4.17}$$

Somit gilt $\Delta r_i \leq m/(2k)$ für alle Zeilen $i \in \{1,\dots,k\}$ und $\Delta c_j \leq m/(2\ell)$ für alle Spalten $j \in \{1,\dots,\ell\}$. Daraus folgt

$$\begin{aligned}\forall I \subsetneq \{1,\dots,k\},\ I \neq \emptyset: \quad \Delta r_I - \Delta c_{J_A(I)} &\leq (k-1)\cdot\frac{m}{2k} + \ell\cdot\frac{m}{2\ell}\\ &= m - \frac{m}{2k} < c_{J_A(I)} - r_I. \end{aligned} \tag{4.4.18}$$

Dies impliziert

$$\begin{aligned}\forall I \subsetneq \{1,\dots,k\},\ I \neq \emptyset: \quad r_I + \Delta r_I &< c_{J_A(I)} + \Delta c_{J_A(I)}\\ &= c_{J_{A+\Delta A}(I)} + \Delta c_{J_{A+\Delta A}(I)}. \end{aligned} \tag{4.4.19}$$

Da nach Wahl von Δr_k außerdem $c_+ + \Delta c_+ = r_+ + \Delta r_+$ gilt, folgt mit Satz 2.4.1 (ii) (siehe S. 35) die Existenz einer Matrix $x \in \mathcal{E}_{\Delta c,\Delta r} \subset \mathbb{R}^{k\times\ell}$ mit $x \equiv A + \Delta A$. Für diese Matrix x gilt

$$(Fu)(x) = \mathrm{I}(x|A+\Delta A) + \delta(x|\mathcal{E}_{\Delta c,\Delta r}) < \infty. \tag{4.4.20}$$

Damit ist der Vektor $u = (\Delta A, \Delta c, \Delta r)$ im eigentlichen Definitionsbereich der Bifunktion F enthalten. Es gilt also $u \in \operatorname{dom} F$. Da $u \in \widetilde{\mathcal{U}}$ beliebig gewählt wurde, folgt $\widetilde{\mathcal{U}} \subseteq \operatorname{dom} F$ und damit insbesondere $0 \in \operatorname{int}(\operatorname{dom} F)$.

V. In den vorangegangenen beiden Schritten wurde gezeigt, dass die Bifunktion F konvex ist und dass $0 \in \operatorname{int}(\operatorname{dom} F)$ gilt. Nach Satz 4.4.2 existiert folglich eine offene, konvexe Umgebung $\mathcal{U}$ von $0 \in \mathbb{R}^{\operatorname{supp}(A)+\ell+(k-1)}$, auf der die Perturbationsfunktion

$$(\Delta A, \Delta c, \Delta r) = u \mapsto \inf_{x\in\mathbb{R}^{k\times\ell}} (Fu)(x) = \inf_{x\in\mathcal{E}_{\Delta c,\Delta r}} \mathrm{I}(x|A+\Delta A) \tag{4.4.21}$$

stetig ist.

VI. Für alle $n \in \mathbb{N}$ wählen wir

$$\Delta a_{ij}^n := a_{ij}^n - a_{ij} \text{ für alle Einträge } (i,j) \in \operatorname{supp}(A), \tag{4.4.22}$$

$$\Delta c_j^n := c_j^n - c_j \text{ für alle Spalten } j \in \{1,\dots,\ell\} \text{ und} \tag{4.4.23}$$

$$\Delta r_i^n := r_i^n - r_i \text{ für alle Zeilen } i \in \{1,\dots,k-1\}. \tag{4.4.24}$$

Nach Satz 2.3.4 (ii) (siehe S. 34) ist die Direktanpassung B die I_1-Projektion von A auf $\mathcal{E}_{0,0}$ und die Direktanpassungen B^n sind die I_1-Projektionen von A^n auf $\mathcal{E}_{\Delta c^n,\Delta r^n}$. Somit gilt

$$\mathrm{I}(B|A) = \inf_{x\in\mathcal{E}_{0,0}} \mathrm{I}(x|A) \text{ und } \mathrm{I}(B^n|A^n) = \inf_{x\in\mathcal{E}_{\Delta c^n,\Delta r^n}} \mathrm{I}(x|A+\Delta A^n). \tag{4.4.25}$$

Mit der im vorangegangenen Schritt gezeigten Stetigkeit der Perturbationsfunktion folgt die Gleichung

$$\begin{aligned}\lim_{n\to\infty} \mathrm{I}(B^n|A^n) &= \lim_{n\to\infty} \inf_{x\in\mathcal{E}_{\Delta c^n,\Delta r^n}} \mathrm{I}(x|A+\Delta A^n)\\ &= \inf_{x\in\mathcal{E}_{0,0}} \mathrm{I}(x|A) = \mathrm{I}(B|A). \end{aligned} \tag{4.4.26}$$

Nach Voraussetzung konvergiert $b^n_{++} = c^n_+$ für $n \to \infty$ gegen c_+. Also ist die Folge (B^n) in einer kompakten Teilmenge von $[0;\infty)^{k\times\ell}$ enthalten. Sei nun $B^* := \lim_{m\to\infty} B^{n_m} \in \mathcal{E}_{0,0}$ ein beliebiger Häufungspunkt dieser Folge. Dann gilt mit der unteren Halbstetigkeit der I-Divergenz (siehe Satz 2.1.2 (iv), S. 23) und Gleichung (4.4.26) die Ungleichung

$$\mathrm{I}(B^*|A) = \mathrm{I}(\lim_{m\to\infty} B^{n_m}|\lim_{m\to\infty} A^{n_m}) \leq \liminf_{m\to\infty} \mathrm{I}(B^{n_m}|A^{n_m}) = \mathrm{I}(B|A). \tag{4.4.27}$$

Da B die I_1-Projektion von A auf $\mathcal{E}_{0,0}$ ist, folgt $\mathrm{I}(B^*|A) = \mathrm{I}(B|A)$ und $B^* = B$. Folglich gilt

$$\lim_{n\to\infty} B^n = B. \tag{4.4.28}$$

Damit ist Aussage (ii) gezeigt. □

Wie im vorangegangenen Abschnitt stellt sich auch hier die Frage nach der Übertragbarkeit der soeben für die Limesmatrix im Konvergenzfall gezeigten Ergebnisse auf die Häufungspunkte im Divergenzfall. Der folgende Satz zeigt, dass dies unter gewissen Regularitätsbedingungen möglich ist.

Satz 4.4.4 (Stetige Abhängigkeit der Häufungspunkte der IPF-Folge von der Ausgangsmatrix und den Marginalien)**.** *Sei (A, c, r) ein biproportionales Anpassungsproblem und seien $C^* := \lim_{t=0,2,4,\dots} A(t)$ und $R^* := \lim_{t=0,2,4,\dots} A(t+1)$ die Häufungspunkte der zugehörigen IPF-Folge $(A(t))$. Weiter sei $((A^n, c^n, r^n))$ eine Folge von biproportionalen Anpassungsproblemen, die gegen (A, c, r) konvergiert und für die $A^n \equiv A$ gilt. Außerdem bezeichnen $C^{n,*} := \lim_{t=0,2,4,\dots} A^n(t)$ und $R^{n,*} := \lim_{t=0,2,4,\dots} A^n(t+1)$ die Häufungspunkte der zu dem Anpassungsproblem (A^n, c^n, r^n) gehörenden IPF-Folgen $(A^n(t))$. Dann gilt: Wenn die Matrizen $C^{n,*}$ und $R^{n,*}$ dieselbe Zusammenhangsstruktur haben wie die Matrizen C^* und R^*, dann konvergiert die Häufungspunktefolge $(C^{n,*})$ gegen C^* und die Häufungspunktefolge $(R^{n,*})$ gegen R^*.*

Beweis. Da die Matrizen $C^{n,*}$ und $R^{n,*}$ dieselbe Zusammenhangsstruktur haben wie die Matrizen C^* und R^*, existieren gemeinsame Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1,\dots,k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1,\dots,\ell\}$ im Sinne von Satz 4.1.1. Die resultierenden Blöcke $C^{*(I_m,J_m)}$, $m = 1,\dots,M$ und $C^{n,*(I_m,J_m)}$, $m = 1,\dots,M$, sind gemäß Satz 4.1.1 zusammenhängend. Nach Satz 1.3.2 (siehe S. 14) sind sie somit Direktanpassungen der entsprechenden Blöcke von (A, c, r) beziehungsweise von (A^n, c^n, r^n) bei skalierten Marginalien. Folglich konvergieren nach Satz 4.4.3 (ii) für alle $m = 1,\dots,M$ die Zusammenhangskomponenten $C^{n,*(I_m,J_m)}$ gegen die Zusammenhangskomponente $C^{*(I_m,J_m)}$. Dementsprechend konvergiert die Häufungspunktefolge $(C^{n,*})$ gegen C^*. Die Konvergenz der Häufungspunktefolge $(R^{n,*})$ gegen R^* folgt analog. □

4.5. Kommentare und Referenzen

Satz 4.1.1 ist bisher unveröffentlicht, basiert jedoch auf Satz 1.4.2 (siehe S. 19). Man beachte daher auch die Kommentare und Referenzen in Abschnitt 1.6 (siehe S. 21). Satz 4.1.1 trifft keine Aussage darüber, wie die Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1,\dots,k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1,\dots,\ell\}$ bestimmt werden können. Liegt einer der beiden Häufungspunkte C^* oder R^* vor, so können wir dessen Zusammenhangskomponenten ermitteln und erhalten somit Zerlegungen $\tilde{I}_1 \uplus \tilde{I}_2 \uplus$

$\cdots \uplus \tilde{I}_M = \{1,\ldots,k\}$ und $\tilde{J}_1 \uplus \tilde{J}_2 \uplus \cdots \uplus \tilde{J}_M = \{1,\ldots,\ell\}$. Durch geschicktes Permutieren der Mengen $\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_M$ und der Mengen $\tilde{J}_1, \tilde{J}_2, \ldots, \tilde{J}_M$ erhalten wir daraus die Zerlegungen $I_1 \uplus I_2 \uplus \cdots \uplus I_M = \{1,\ldots,k\}$ und $J_1 \uplus J_2 \uplus \cdots \uplus J_M = \{1,\ldots,\ell\}$ gemäß Satz 4.1.1. Zur Bestimmung der Häufungspunkte C^* oder R^* ist es jedoch im Allgemeinen erforderlich, das IPF-Verfahren anzuwenden und die Limesmatrizen $\lim_{t=0,2,4,\ldots} A(t)$ und $\lim_{t=0,2,4,\ldots} A(t+1)$ zu ermitteln. Dabei ist es wichtig und zugleich schwierig, zwischen echten Nulleinträgen und „numerischen Nulleinträgen", also sehr kleinen Werten, in diesen Limesmatrizen zu unterscheiden. Als Alternative schlägt Aas (2014) einen Algorithmus vor, der es erlaubt, ausgehend von einem Anpassungsproblem (A, c, r) die gemeinsame Zerlegung der Ausgangsmatrix und der Häufungspunkte zu bestimmen, ohne das IPF-Verfahren anzuwenden. Satz 4.1.2 wurde bereits von Reffel (2014, Lemma 9.3) gezeigt. Eine vergleichbare Aussage wurde in einer früheren Fassung von Pukelsheim (2014) gezeigt.

Die in Abschnitt 4.2 präsentierten Ergebnisse sind bisher unveröffentlicht.

Die Ergebnisse der Abschnitte 4.3 und 4.4 wurden im Wesentlichen bereits in Gietl und Reffel (2013b) vorgestellt. Abschnitt 4.3 wurde außerdem vorab in Gietl und Reffel (2013c) veröffentlicht. Man beachte daher auch die Kommentare zu Gietl und Reffel (2013b) und Gietl und Reffel (2013c) in Abschnitt 2.7 (siehe S. 41). Die Sätze 4.3.2 und 4.4.4 sind im Gegensatz zu den anderen Ergebnissen bisher unveröffentlicht, basieren aber ebenfalls auf gemeinsamer Arbeit mit Fabian P. Reffel. Für den Spezialfall biproportionaler Limesanpassungsprobleme (A, c, r) mit quadratischer Ausgangsmatrix $A \in [0;\infty)^{k\times k}$ sowie Einheitsmarginalien $c = r = (1,\ldots,1) \in (0;\infty)^k$ wurde die in Satz 4.3.1 bewiesene stetige Abhängigkeit der biproportionalen Limesanpassung B von der Ausgangsmatrix A bereits von Sinkhorn (1972, Korollar 3) gezeigt. Sinkhorns Beweis basiert jedoch auf Resultaten zu doppelt stochastischen Matrizen und deren Permanenten. Aus diesem Grund erscheint eine Verallgemeinerung dieses Beweises auf beliebige biproportionale Limesanpassungsprobleme nicht möglich. Für Folgen von biproportionalen Direktanpassungsproblemen $((A^n, c^n, r^n))$ mit positiven Ausgangsmatrizen $A^n \in (0;\infty)^{k\times\ell}$ und $c^n_+ = r^n_+$, die gegen ein biproportionales Direktanpassungsproblem (A, c, r) mit positiver Ausgangsmatrix $A \in (0;\infty)^{k\times\ell}$ konvergieren, zeigen Balinski und Demange (1989, Korollar zu Theorem 3) die Konvergenz der Limesanpassungen (B^n) gegen die Limesanpassung B. Dieser Fall ist in Satz 4.4.3 enthalten. Außerdem zeigen Balinski und Demange (1989, Theorem 3), dass bei Verzicht auf die Einschränkung $A \in (0;\infty)^{k\times\ell}$ zugunsten von $A \in [0;\infty)^{k\times\ell}$ die Limesanpassungen (B^n) weiterhin gegen eine Matrix $\widetilde{B}$ konvergieren, die von der Ausgangsmatrix A dominiert wird und die Marginalien c und r erfüllt. Balinski und Demange lassen jedoch die Frage unbeantwortet, ob es sich bei diesem Grenzwert $\widetilde{B}$ um die Limesanpassung von A an c und r handelt.

5. Zusammenfassung der Ergebnisse zum zweidimensionalen IPF-Verfahren

In Teil I dieser Dissertation konnte ein umfassender Überblick über das zweidimensionale IPF-Verfahren und seine asymptotischen Struktureigenschaften gegeben werden. Zu diesem Zweck wurden in Kapitel 1 das zweidimensionale IPF-Verfahren und die biproportionalen Anpassungsprobleme eingeführt und einige grundlegende Eigenschaften diskutiert. Diese Diskussion erfolgte nur anhand der Matrizenfolge und der Multiplikatorfolge und ohne Rückgriff auf aufwendigere Ansätze, wie zum Beispiel den informationstheoretischen Ansatz. In Kapitel 2 wurden die drei in der einschlägigen Literatur vorherrschenden Ansätze zur Konvergenzanalyse des zweidimensionalen IPF-Verfahrens diskutiert. Als Hinführung zum informationstheoretischen Ansatz wurden I-Divergenzen und, darauf aufbauend, I-Projektionen von Vektoren und von Matrizen eingeführt und ihre wichtigsten Eigenschaften diskutiert. Dort wurde mit der Stetigkeit von I_1-Projektionen das erste neue Resultat dieser Dissertation bewiesen. Anschließend wurde in kompakter Darstellung die Grundidee des informationstheoretischen Ansatzes zur Konvergenzanalyse herausgearbeitet und die wichtigsten Implikationen für biproportionale Anpassungsprobleme gezeigt. Als zweites neues Resultat dieser Dissertation wurde die Äquivalenz des auf dem geometrischen Mittel basierenden Ansatzes zum informationstheoretischen Ansatz gezeigt. Der informationstheoretische Ansatz kristallisierte sich damit als der zentrale Ansatz für die Untersuchung des IPF-Verfahrens heraus. In Kapitel 3 wurde mit der Konvergenz der IPF-Teilfolgen im Divergenzfall ein weiteres neues Resultat dieser Dissertation bewiesen. Darauf aufbauend konnten in Kapitel 4 weitere neue Resultate gezeigt werden, nämlich die gemeinsame Zerlegung der Häufungspunkte und der Ausgangsmatrix des IPF-Verfahrens und die informationstheoretische Charakterisierung der Häufungspunkte. Als letztes neues Resultat zum zweidimensionalen IPF-Verfahren wurde die stetige Abhängigkeit der Limesmatrix beziehungsweise der Häufungspunkte von der Ausgangsmatrix und den Marginalien gezeigt.

Als offenes Problem bezüglich des zweidimensionalen IPF-Verfahrens verbleibt die Frage nach der Verallgemeinerbarkeit einiger Stetigkeitsaussagen. So stellt sich die Frage, unter welchen Bedingungen in Satz 4.3.2 (siehe S. 59) auf die Forderung nach der Direktheit des biproportionalen Anpassungsproblems (A, c, r) verzichtet werden kann. Des Weiteren stellt sich die Frage, ob auf die in Satz 4.4.3 (siehe S. 60) und in Satz 4.4.4 (siehe S. 64) geforderten Regularitätsbedingungen verzichtet werden kann.

In Teil II dieser Dissertation werden die asymptotischen Struktureigenschaften des mehrdimensionalen IPF-Verfahrens untersucht. Während beim zweidimensionalen IPF-Verfahren Matrizen, also zweidimensionale Tafeln, durch zyklische Skalierung an zwei eindimensionale Marginaltafeln angepasst werden, passt das mehrdimensionale IPF-Verfahren Tafeln der Dimension $d \geq 2$ an $\eta \geq 2$ eindimensionale oder höherdimensionale Marginaltafeln an. Ziel der Analyse ist es, möglichst viele der in Teil I gezeigten Aussagen über das zweidimensionale IPF-Verfahren auf das mehrdimensionale IPF-Verfahren zu übertragen.

Teil II.

Asymptotische Struktureigenschaften des mehrdimensionalen iterativen proportionalen Anpassungsverfahrens (IPF-Verfahrens)

6. Mehrdimensionales IPF-Verfahren

In diesem Kapitel wird das mehrdimensionale iterative proportionale Anpassungsverfahren (IPF-Verfahren, iterative proportional fitting procedure) vorgestellt. Dazu wird im ersten Abschnitt das Verfahren beschrieben. Im zweiten Abschnitt wird eine Möglichkeit zur Normierung der Eingabe und der Ausgabe des Verfahrens gezeigt.

6.1. Definition des Verfahrens und mehrdimensionaler Anpassungsprobleme

Das in Teil I untersuchte zweidimensionale IPF-Verfahren kann zu einem mehrdimensionalen Verfahren verallgemeinert werden. An die Stelle der zweidimensionalen Matrizen $S \in [0;\infty)^{k\times\ell}$ treten dabei Tafeln $S \in [0;\infty)^{k_1\times\cdots\times k_d}$ der Dimension $d \geq 2$ mit der *Indexmenge* $\Omega := \prod_{\delta=1}^{d}\{1,\ldots,k_\delta\}$. Diese Tafeln verlangen nach einer platzsparenden und von der Wahl der Dimension d unabhängigen Notation. Die Menge $[0;\infty)^{k_1\times\cdots\times k_d}$ wird daher im Folgenden mit $[0;\infty)^{\Omega}$ bezeichnet. An die Stelle der Indexpaare (i,j) bei Matrizen treten bei Tafeln $S \in [0;\infty)^{\Omega}$ die *Indexvektoren* $\omega \in \Omega$. Folglich gilt $S = (s_\omega)_{\omega\in\Omega}$. Des Weiteren muss der Begriff des Randes formalisiert werden. Während die Randsummen einer Matrix nur aus den Zeilensummen und den Spaltensummen bestehen, verfügt jede Tafel der Dimension $d > 2$ über eine Vielzahl von Rändern. Jede nichtleere und echte Teilmenge $D \subsetneq \{1,\ldots,d\}$, $D \neq \emptyset$, ist ein *Rand der Tafel* $S \in [0;\infty)^{\Omega}$. Die Menge $\Omega_D := \prod_{\delta\in D}\{1,\ldots,k_\delta\}$ wird als die *Indexmenge des Randes* D bezeichnet. Zu jedem Rand D gehört außerdem der *Komplementärrand* $D' := \{1,\ldots,d\} \setminus D$. Dementsprechend kann jeder Indexvektor $\omega \in \Omega$ in die beiden Komponenten $\omega^{(D)} \in \Omega_D$ und $\omega^{(D')} \in \Omega_{D'}$ zerlegt werden. Umgekehrt können zwei beliebige Indexvektoren $\omega^* \in \Omega_D$ und $\omega^{**} \in \Omega_{D'}$ stets zu einem Indexvektor $\tilde\omega \in \Omega$ mit $\tilde\omega^{(D)} = \omega^*$ und $\tilde\omega^{(D')} = \omega^{**}$ zusammengefasst werden. Dieser Indexvektor $\tilde\omega$ wird im Folgenden unter Vernachlässigung der Reihenfolge mit (ω^*,ω^{**}) bezeichnet. Die zum Rand D der Tafel S gehörenden Randsummen werden in der *Randtafel* $S^{(D)} \in [0;\infty)^{\Omega_D}$ zusammengefasst. Die Einträge $s^{(D)}_{\omega^*}$ dieser Randtafel $S^{(D)}$ sind definiert gemäß

$$\forall \omega^* \in \Omega_D: \quad s^{(D)}_{\omega^*} := \sum_{\omega^{**}\in\Omega_{D'}} s_{(\omega^*,\omega^{**})}. \tag{6.1.1}$$

An die Stelle der eindimensionalen Zeilenmarginalien $r \in (0;\infty)^k$ und der eindimensionalen Spaltenmarginalien $c \in (0;\infty)^\ell$ treten Tafeln $M^{(D)} \in (0;\infty)^{\Omega_D}$ der Dimension $|D| \in \{1,\ldots,d-1\}$.

Das *mehrdimensionale IPF-Verfahren* benötigt als Eingabe eine nichtnegative Tafel $A \in [0;\infty)^{\Omega}$, wobei $\Omega = \prod_{\delta=1}^{d}\{1,\ldots,k_\delta\}$ und $d \geq 2$ gelten, sowie $\eta \geq 2$ Ränder $D_i \subsetneq \{1,\ldots,d\}$, $D_i \neq \emptyset$, $i = 1,\ldots,\eta$. Für die zugehörigen Randtafeln $A^{(i)} := A^{(D_i)}$ muss $A^{(i)} \in (0;\infty)^{\Omega_i}$ gelten. Dabei bezeichnet $\Omega_i := \Omega_{D_i} = \prod_{\delta\in D_i}\{1,\ldots,k_\delta\}$ die Indexmenge des Randes D_i. Außerdem benötigt das mehrdimensionale IPF-Verfahren als Eingabe zu jedem Rand D_i eine positive Tafel $M^{(i)} \in$

$(0;\infty)^{\Omega_i}$. Die Tafel A wird im Folgenden die *Ausgangstafel* genannt, während die Tafeln $M^{(i)}$, $i = 1, \ldots, \eta$, als *Marginaltafeln* bezeichnet werden.

Zur Initialisierung des IPF-Verfahrens wird $A(0) := A$ gewählt. Davon ausgehend wird durch iterative Ausführung der folgenden η Schritte die *IPF-Folge* $(A(t))$ berechnet:

Der Schritt $s\eta + 1$ passt die Randtafel $A^{(1)}(s\eta)$ an die Marginaltafel $M^{(1)}$ an. Zu diesem Zweck werden alle Einträge $a_\omega(s\eta)$, die dieselbe Indexkomponente $\omega^{(D_1)}$ besitzen, mit demselben Multiplikator multipliziert gemäß der Formel

$$\forall \omega \in \Omega : \quad a_\omega(s\eta + 1) := \frac{m^{(1)}_{\omega^{(D_1)}}}{a^{(1)}_{\omega^{(D_1)}}(s\eta)} \cdot a_\omega(s\eta). \tag{6.1.2}$$

Der Schritt $s\eta + 2$ passt die Randtafel $A^{(2)}(s\eta + 1)$ an die Marginaltafel $M^{(2)}$ an. Zu diesem Zweck werden alle Einträge $a_\omega(s\eta + 1)$, die dieselbe Indexkomponente $\omega^{(D_2)}$ besitzen, mit demselben Multiplikator multipliziert gemäß der Formel

$$\forall \omega \in \Omega : \quad a_\omega(s\eta + 2) := \frac{m^{(2)}_{\omega^{(D_2)}}}{a^{(2)}_{\omega^{(D_2)}}(s\eta + 1)} \cdot a_\omega(s\eta + 1). \tag{6.1.3}$$

Die Schritte $s\eta + 3$ bis $s\eta + \eta - 1$ erfolgen in analoger Weise.

Der Schritt $s\eta + \eta$ passt die Randtafel $A^{(\eta)}(s\eta + \eta - 1)$ an die Marginaltafel $M^{(\eta)}$ an. Zu diesem Zweck werden alle Einträge $a_\omega(s\eta + \eta - 1)$, die dieselbe Indexkomponente $\omega^{(D_\eta)}$ besitzen, mit demselben Multiplikator multipliziert gemäß der Formel

$$\forall \omega \in \Omega : \quad a_\omega(s\eta + \eta) := \frac{m^{(\eta)}_{\omega^{(D_\eta)}}}{a^{(\eta)}_{\omega^{(D_\eta)}}(s\eta + \eta - 1)} \cdot a_\omega(s\eta + \eta - 1). \tag{6.1.4}$$

Per Induktion kann man zeigen, dass für alle Schritte $t \geq 0$ die Ungleichung $a_\omega(t) > 0$ genau dann gilt, wenn $a_\omega > 0$ gilt. Folglich bleiben alle Randsummen bezüglich der Ränder $D_1, \ldots, D_\eta$ stets positiv,

$$\forall t \geq 0 \quad \forall i = 1, \ldots, \eta \quad \forall \omega^* \in \Omega_i : \quad a^{(i)}_{\omega^*}(t) > 0. \tag{6.1.5}$$

Das IPF-Verfahren ist somit wohldefiniert. Außerdem sind alle Tafeln der IPF-Folge maßtheoretisch äquivalent zur Ausgangstafel,

$$\forall t \geq 0 : \quad A(t) \equiv A. \tag{6.1.6}$$

Die in diesem Abschnitt definierte Eingabe $(A, M^{(1)}, \ldots, M^{(\eta)})$ des IPF-Verfahrens wird im Folgenden als *mehrdimensionales Anpassungsproblem* bezeichnet. Das *IPF-Verfahren konvergiert*, wenn die IPF-Folge $(A(t))$ konvergiert.

Beispiel 6.1.1 (Zweidimensionales IPF-Verfahren als Spezialfall des mehrdimensionalen IPF-Verfahrens). Sei (A, c, r) ein biproportionales Anpassungsproblem und $(A(t))$ die zugehörige IPF-Folge im Sinne von Kapitel 1 (siehe S. 12). Dann ist die Ausgangsmatrix $A \in [0;\infty)^{k\times\ell}$ eine Ausgangstafel $A \in [0;\infty)^{\Omega}$ der Dimension $d = 2$ mit der Indexmenge $\Omega = \{1, \dots, k\} \times \{1, \dots, \ell\}$. Der Zeilenrand von A ist gleich dem Rand $D_1 := \{1\}$ und die Zeilensummen von A sind gleich der Randtafel $A^{(1)} \in (0;\infty)^{\Omega_1}$ mit der Indexmenge $\Omega_1 = \{1, \dots, k\}$. Der Spaltenrand A ist gleich dem Rand $D_2 := \{2\}$ und die Spaltensummen von A sind gleich der Randtafel $A^{(2)} \in (0;\infty)^{\Omega_2}$ mit der Indexmenge $\Omega_2 = \{1, \dots, \ell\}$. Weiter sei $(A, M^{(1)}, M^{(2)})$ das mehrdimensionale Anpassungsproblem mit den Rändern D_1 und D_2 sowie den Marginaltafeln $M^{(1)} := r$ und $M^{(2)} := c$. Dann ist die vom mehrdimensionalen IPF-Verfahren ausgehend von $(A, M^{(1)}, M^{(2)})$ errechnete Tafelfolge $(\widetilde{A}(t))$ gleich der vom zweidimensionalen IPF-Verfahren errechneten Matrizenfolge $(A(t))$. Somit ist das zweidimensionale IPF-Verfahren ein niedrigdimensionaler Spezialfall des mehrdimensionalen IPF-Verfahrens. Für diesen Spezialfall gelten $d = 2$ und $\eta = 2$ sowie $D_1 = \{1\}$ und $D_2 = \{2\}$.

6.2. Normierung mehrdimensionaler Anpassungsprobleme

Wie beim zweidimensionalen IPF-Verfahren ist es auch bei der Analyse des mehrdimensionalen IPF-Verfahrens häufig von Vorteil, die Ausgangstafel A und die Marginaltafeln $M^{(i)}$, $i = 1, \dots, \eta$, zu normieren,

$$\widetilde{A} := \frac{1}{\sum_{\omega\in\Omega} a_\omega} \cdot A \qquad \text{und} \qquad \forall i = 1, \dots, \eta : \quad \widetilde{M}^{(i)} := \frac{1}{\sum_{\omega^*\in\Omega_i} m^{(i)}_{\omega^*}} \cdot M^{(i)}. \tag{6.2.1}$$

Die ausgehend von dem *normierten mehrdimensionalen Anpassungsproblem* $(\widetilde{A}, \widetilde{M}^{(1)}, \dots, \widetilde{M}^{(\eta)})$ errechnete IPF-Folge $(\widetilde{A}(t))$ entspricht dann einer gliedweisen Normierung der ausgehend von dem ursprünglichen mehrdimensionalen Anpassungsproblem $(A, M^{(1)}, \dots, M^{(\eta)})$ errechneten IPF-Folge $(A(t))$.

Satz 6.2.1 (Normierung mehrdimensionaler Anpassungsprobleme)**.** *Sei $(A, M^{(1)}, \dots, M^{(\eta)})$ ein mehrdimensionales Anpassungsproblem und sei $(\widetilde{A}, \widetilde{M}^{(1)}, \dots, \widetilde{M}^{(\eta)})$ das zugehörige normierte mehrdimensionale Anpassungsproblem. Weiter sei $(A(t))$ die ausgehend von $(A, M^{(1)}, \dots, M^{(\eta)})$ errechnete IPF-Folge und $(\widetilde{A}(t))$ die ausgehend von $(\widetilde{A}, \widetilde{M}^{(1)}, \dots, \widetilde{M}^{(\eta)})$ errechnete IPF-Folge. Dann gilt*

$$\forall s \in \mathbb{N}_0 \quad \forall i = 1, \dots, \eta : \quad A(s\eta + i) = \left(\sum_{\omega^*\in\Omega_i} m^{(i)}_{\omega^*} \right) \cdot \widetilde{A}(s\eta + i). \tag{6.2.2}$$

Beweis. Der Beweis erfolgt per Induktion über $s \in \mathbb{N}_0$ sowie $i \in \{1, \dots, \eta\}$ und ist in drei Schritte gegliedert. In Schritt I zeigen wir, dass Gleichung (6.2.2) für $s = 0$ und $i = 1$ gilt. In Schritt II nehmen wir an, dass Gleichung (6.2.2) für ein beliebiges $s \in \mathbb{N}_0$ und ein beliebiges $i \in \{1, \dots, \eta - 1\}$ gilt. Daraus folgern wir die Gültigkeit von Gleichung (6.2.2) für s und $i + 1$. In Schritt III nehmen wir schließlich an, dass Gleichung (6.2.2) für ein beliebiges $s \in \mathbb{N}_0$ und $i = \eta$ gilt. Daraus folgern wir die Gültigkeit von Gleichung (6.2.2) für $s + 1$ und $i = 1$. Damit ist Gleichung (6.2.2) für alle $s \in \mathbb{N}_0$ und für alle $i \in \{1, \dots, \eta\}$ gezeigt.

I. Nach Voraussetzung gelten

$$A = \left(\sum_{\omega \in \Omega} a_\omega \right) \cdot \widetilde{A} \qquad \text{und} \qquad \forall i = 1, \ldots, \eta : \quad M^{(i)} = \left(\sum_{\omega^* \in \Omega_i} m_{\omega^*}^{(i)} \right) \cdot \widetilde{M}^{(i)}. \tag{6.2.3}$$

Daraus folgt gemäß Definition des IPF-Verfahrens für alle $\omega \in \Omega$ die Aussage

$$\begin{aligned} a_\omega(1) &= \frac{m^{(1)}_{\omega^{(D_1)}}}{a^{(1)}_{\omega^{(D_1)}}} \cdot a_\omega = \frac{\left(\sum_{\omega^* \in \Omega_1} m^{(1)}_{\omega^*}\right) \cdot \tilde{m}^{(1)}_{\omega^{(D_1)}}}{\left(\sum_{\omega \in \Omega} a_\omega\right) \cdot \tilde{a}^{(1)}_{\omega^{(D_1)}}} \cdot \left(\sum_{\omega \in \Omega} a_\omega \right) \cdot \tilde{a}_\omega \\ &= \left(\sum_{\omega^* \in \Omega_1} m^{(1)}_{\omega^*} \right) \cdot \frac{\tilde{m}^{(1)}_{\omega^{(D_1)}}}{\tilde{a}^{(1)}_{\omega^{(D_1)}}} \cdot \tilde{a}_\omega = \left(\sum_{\omega^* \in \Omega_1} m^{(1)}_{\omega^*} \right) \cdot \tilde{a}_\omega(1). \end{aligned} \tag{6.2.4}$$

Damit ist Gleichung (6.2.2) für $s = 0$ und $i = 1$ gezeigt.

II. Gleichung (6.2.2) gelte für ein beliebiges $s \in \mathbb{N}_0$ und ein beliebiges $i \in \{1, \ldots, \eta - 1\}$. Daraus folgt gemäß Definition des IPF-Verfahrens für alle $\omega \in \Omega$ die Aussage

$$\begin{aligned} a_\omega(s\eta + i + 1) &= \frac{m^{(i+1)}_{\omega^{(D_{i+1})}}}{a^{(i+1)}_{\omega^{(D_{i+1})}}(s\eta + i)} \cdot a_\omega(s\eta + i) \\ &= \frac{\left(\sum_{\omega^* \in \Omega_{i+1}} m^{(i+1)}_{\omega^*}\right) \cdot \tilde{m}^{(i+1)}_{\omega^{(D_{i+1})}}}{\left(\sum_{\omega^* \in \Omega_i} m^{(i)}_{\omega^*}\right) \cdot \tilde{a}^{(i+1)}_{\omega^{(D_{i+1})}}(s\eta + i)} \cdot \left(\sum_{\omega^* \in \Omega_i} m^{(i)}_{\omega^*} \right) \cdot \tilde{a}_\omega(s\eta + i) \\ &= \left(\sum_{\omega^* \in \Omega_{i+1}} m^{(i+1)}_{\omega^*} \right) \cdot \frac{\tilde{m}^{(i+1)}_{\omega^{(D_{i+1})}}}{\tilde{a}^{(i+1)}_{\omega^{(D_{i+1})}}(s\eta + i)} \cdot \tilde{a}_\omega(s\eta + i) \\ &= \left(\sum_{\omega^* \in \Omega_{i+1}} m^{(i+1)}_{\omega^*} \right) \cdot \tilde{a}_\omega(s\eta + i + 1). \end{aligned} \tag{6.2.5}$$

Somit gilt Gleichung (6.2.2) für s und $i + 1$.

III. Gleichung (6.2.2) gelte für ein beliebiges $s \in \mathbb{N}_0$ und $i = \eta$. Daraus folgt gemäß Definition des IPF-Verfahrens für alle $\omega \in \Omega$ die Aussage

$$
\begin{aligned}
&a_\omega((s+1)\eta+1)\\
&\quad= \frac{m^{(1)}_{\omega^{(D_1)}}}{a^{(1)}_{\omega^{(D_1)}}((s+1)\eta)} \cdot a_\omega((s+1)\eta)\\
&\quad= \frac{\left(\sum_{\omega^*\in\Omega_1} m^{(1)}_{\omega^*}\right)\cdot \tilde{m}^{(1)}_{\omega^{(D_1)}}}{\left(\sum_{\omega^*\in\Omega_\eta} m^{(\eta)}_{\omega^*}\right)\cdot \tilde{a}^{(1)}_{\omega^{(D_1)}}((s+1)\eta)} \cdot \left(\sum_{\omega^*\in\Omega_\eta} m^{(\eta)}_{\omega^*}\right)\cdot \tilde{a}_\omega((s+1)\eta)\\
&\quad= \left(\sum_{\omega^*\in\Omega_1} m^{(1)}_{\omega^*}\right)\cdot \frac{\tilde{m}^{(1)}_{\omega^{(D_1)}}}{\tilde{a}^{(1)}_{\omega^{(D_1)}}((s+1)\eta)} \cdot \tilde{a}_\omega((s+1)\eta)\\
&\quad= \left(\sum_{\omega^*\in\Omega_1} m^{(1)}_{\omega^*}\right)\cdot \tilde{a}_\omega((s+1)\eta+1). \qquad (6.2.6)
\end{aligned}
$$

Somit gilt Gleichung (6.2.2) für $s+1$ und $i = 1$. □

6.3. Kommentare und Referenzen

Die in Abschnitt 6.1 eingeführte Notation ermöglicht die formale Diskussion des mehrdimensionalen IPF-Verfahrens ohne Festlegung der Dimension d der Ausgangstafel. Dies wird erreicht durch die Verwendung von flexiblen Indexvektoren $\omega \in \Omega = \prod_{\delta=1}^{d}\{1,\ldots,k_\delta\}$ in Anlehnung an Gokhale und Kullback (1978a,b) und Kullback und Keegel (1984) anstelle von unflexiblen Indextupeln, wie zum Beispiel (i, j, k, ℓ). Der Begriff „Tafel“ wurde von Bishop, Fienberg und Holland (1975), Agresti (1990) und Christensen (1990) übernommen. Er ermöglicht eine klare Abgrenzung von den in Teil I diskutierten „Matrizen“. Andere Autoren verwenden stattdessen den Begriff „multidimensional matrices“ (Bapat 1982). Wie in Beispiel 6.1.1 gezeigt wurde, ist das in Teil I untersuchte zweidimensionale IPF-Verfahren ein niedrigdimensionaler Spezialfall des in Abschnitt 6.1 definierten mehrdimensionalen IPF-Verfahrens, für den $d = 2$ und $\eta = 2$ gilt. Dementsprechend sind alle „positiven“ Aussagen, die in den folgenden Kapiteln über das mehrdimensionale IPF-Verfahren gezeigt werden, Verallgemeinerungen der entsprechenden Aussagen in Teil I. „Negative“ Aussagen oder Gegenbeispiele, nach denen bestimmte Eigenschaften des zweidimensionalen IPF-Verfahrens nicht auf das mehrdimensionale IPF-Verfahren übertragen werden können, bedürfen dagegen immer der Einschränkung $d > 2$ oder $\eta > 2$.

Die in Abschnitt 6.2 diskutierte Normierung mehrdimensionaler Anpassungsprobleme ist eine Verallgemeinerung der in Abschnitt 1.5 (siehe S. 20) eingeführten Normierung biproportionaler Anpassungsprobleme. Man beachte daher die Kommentare und Referenzen in Abschnitt 1.6 (siehe S. 21).

Im Gegensatz zu Kapitel 1 (siehe S. 12) wurde in diesem Kapitel keine Unterscheidung zwischen Limesanpassungen und Direktanpassungen eingeführt. Während die Definitionen der Limesanpassung und der Direktanpassung leicht auf mehrdimensionale Anpassungsprobleme übertragen werden könnten, gestaltet sich die Charakterisierung von mehrdimensionalen Limesanpassungen und Direktanpassungen schwierig. Dies liegt daran, dass unklar ist, wie der in Definition 1.3.1 (siehe S. 14) eingeführte Zusammenhangsbegriff für Matrizen auf mehrdimensionale Tafeln verallgemeinert werden soll. Dementsprechend unklar bleibt auch, wie eine Verallgemeinerung der in Abschnitt 1.4 (siehe S. 16) beschriebenen Zerlegung biproportionaler Anpassungsprobleme aussehen könnte.

7. Konvergenzanalyse des mehrdimensionalen IPF-Verfahrens

In diesem Kapitel werden zwei Ansätze zur Konvergenzanalyse des mehrdimensionalen IPF-Verfahrens diskutiert. Der erste Ansatz basiert erneut auf der Informationsdivergenz (I-Divergenz). Dementsprechend werden in den ersten beiden Abschnitten die I-Divergenz und die darauf aufbauenden I-Projektionen auf beliebige Tafeln verallgemeinert. Im dritten Abschnitt wird die informationstheoretische Konvergenzanalyse des zweidimensionalen IPF-Verfahrens auf das mehrdimensionale IPF-Verfahren verallgemeinert. Der vierte Abschnitt zeigt, dass der L_1-Ansatz zur Konvergenzanalyse des zweidimensionalen IPF-Verfahrens nicht auf das mehrdimensionale IPF-Verfahren mit $\eta > 2$ Marginaltafeln übertragen werden kann.

7.1. Informationsdivergenz (I-Divergenz)

Die Definition der Informationsdivergenz zwischen Vektoren und zwischen Matrizen (siehe Definition 2.1.1, S. 23) kann ohne Probleme auf beliebige Tafeln der Dimension $d \geq 2$ verallgemeinert werden.

Definition 7.1.1 (I-Divergenz, vgl. Kullback und Leibler 1951, Abschnitt 2)**.** Sei Ω eine Indexmenge im Sinne von Abschnitt 6.1 (siehe S. 68) und seien $S \in [0;\infty)^\Omega$ und $T \in [0;\infty)^\Omega$ beliebig. Dann ist die *Informationsdivergenz (I-Divergenz) von S relativ zu T* definiert als

$$\mathrm{I}(S|T) := \sum_{\omega\in\Omega} \overline{\varphi}(s_\omega, t_\omega) = \begin{cases} \sum_{\omega\in\Omega:\, s_\omega>0} s_\omega \ln \frac{s_\omega}{t_\omega} & \text{falls } S \ll T, \\ \infty & \text{sonst.} \end{cases} \tag{7.1.1}$$

Die in Satz 2.1.2 (siehe S. 23) gezeigten Konvexitäts- und Stetigkeitseigenschaften gelten in analoger Weise für die I-Divergenz zwischen höherdimensionalen Tafeln; ebenso die in Satz 2.1.7 (siehe S. 25) gezeigte Pinsker-Ungleichung. Die Aggregationsungleichungen für Matrizen (siehe Satz 2.1.5, S. 24) übertragen sich wie folgt auf höherdimensionale Tafeln.

Satz 7.1.2 (Aggregationsungleichungen für Tafeln)**.** *Für alle Tafeln $S, T \in [0;\infty)^\Omega$ mit $\Omega = \prod_{\delta=1}^{d}\{1,\ldots,k_\delta\}$, $d \geq 2$, gelten die folgenden Ungleichungen:*

(i) Es gilt die Summenungleichung

$$\mathrm{I}\Big(S\Big|T\Big) \geq \mathrm{I}\Big(\sum_{\omega\in\Omega} s_\omega\Big|\sum_{\omega\in\Omega} t_\omega\Big). \tag{7.1.2}$$

Im Fall $S \ll T$ gilt Gleichheit genau dann, wenn $s_\omega/t_\omega = (\sum_{\tilde\omega\in\Omega} s_{\tilde\omega})/(\sum_{\tilde\omega\in\Omega} t_{\tilde\omega})$ für alle $\omega \in \operatorname{supp}(T)$ gilt. Im Fall $S \not\ll T$ gilt Gleichheit genau dann, wenn $\sum_{\omega\in\Omega} t_\omega = 0$ gilt.

(ii) Für alle Ränder $D \subsetneq \{1, \ldots, d\}$, $D \neq \emptyset$, gilt die Randsummenungleichung

$$\mathrm{I}\left(S \middle| T\right) \geq \mathrm{I}\left(S^{(D)} \middle| T^{(D)}\right). \tag{7.1.3}$$

Im Fall $S \ll T$ gilt Gleichheit genau dann, wenn gilt

$$\forall (\omega^*, \omega^{**}) \in \operatorname{supp}(T): \quad \frac{s_{(\omega^*,\omega^{**})}}{t_{(\omega^*,\omega^{**})}} = \frac{\sum_{\tilde{\omega}^{**} \in \Omega_{D'}} s_{(\omega^*,\tilde{\omega}^{**})}}{\sum_{\tilde{\omega}^{**} \in \Omega_{D'}} t_{(\omega^*,\tilde{\omega}^{**})}}. \tag{7.1.4}$$

Im Fall $S \not\ll T$ gilt Gleichheit genau dann, wenn $S^{(D)} \not\ll T^{(D)}$ gilt.

Beweis. (i) Die Summenungleichung (7.1.2) und die dazugehörige Gleichheitsaussage folgen unmittelbar aus Satz 2.1.3 (siehe S. 24).

(ii) Nach Definition 7.1.1 ist die I-Divergenz $\mathrm{I}(S|T)$ gleich der Summe sämtlicher I-Divergenzen $\mathrm{I}((s_{(\omega^*,\omega^{**})})_{\omega^{**} \in \Omega_{D'}} | (t_{(\omega^*,\omega^{**})})_{\omega^{**} \in \Omega_{D'}})$, $\omega^* \in \Omega_D$. Gemäß Aussage (i) können die Summanden nach unten jeweils durch $\mathrm{I}(\sum_{\omega^{**} \in \Omega_{D'}} s_{(\omega^*,\omega^{**})} | \sum_{\omega^{**} \in \Omega_{D'}} t_{(\omega^*,\omega^{**})})$ abgeschätzt werden. Insgesamt erhalten wir also

$$\begin{aligned}
\mathrm{I}\left(S \middle| T\right) &= \sum_{\omega^* \in \Omega_D} \sum_{\omega^{**} \in \Omega_{D'}} \overline{\varphi}\left(s_{(\omega^*,\omega^{**})}, t_{(\omega^*,\omega^{**})}\right) \\
&= \sum_{\omega^* \in \Omega_D} \mathrm{I}\left(\left(s_{(\omega^*,\omega^{**})}\right)_{\omega^{**} \in \Omega_{D'}} \middle| \left(t_{(\omega^*,\omega^{**})}\right)_{\omega^{**} \in \Omega_{D'}}\right) \\
&\geq \sum_{\omega^* \in \Omega_D} \mathrm{I}\left(\sum_{\omega^{**} \in \Omega_{D'}} s_{(\omega^*,\omega^{**})} \middle| \sum_{\omega^{**} \in \Omega_{D'}} t_{(\omega^*,\omega^{**})}\right) = \mathrm{I}\left(S^{(D)} \middle| T^{(D)}\right).
\end{aligned} \tag{7.1.5}$$

Gleichheit gilt genau dann, wenn in allen Abschätzungen Gleichheit gilt. □

Korollar 7.1.3 (Nichtnegativität der I-Divergenz zwischen mehrdimensionalen Tafeln)**.** *Für alle Tafeln $S, T \in [0;\infty)^\Omega$ mit übereinstimmenden Summen $\sum_{\omega \in \Omega} s_\omega = \sum_{\omega \in \Omega} t_\omega$ ist die I-Divergenz* $\mathrm{I}(S|T)$ *nach unten durch* 0 *beschränkt,*

$$\mathrm{I}(S|T) \geq 0. \tag{7.1.6}$$

Gleichheit gilt genau dann, wenn $S = T$ gilt.

7.2. I-Projektionen

Wie die Definition der I-Divergenz zwischen Vektoren und zwischen Matrizen können auch die Definitionen der I_1-Projektionen und der I_2-Projektionen von Vektoren und von Matrizen (siehe Abschnitt 2.2, S. 27) ohne Probleme auf höherdimensionale Tafeln übertragen werden.

Definition 7.2.1 (I_1-Projektion, vgl. Csiszár 1975, S. 147)**.** Sei Ω eine Indexmenge im Sinne von Abschnitt 6.1 (siehe S. 68). Weiter sei $\mathcal{E}$ eine beliebige Teilmenge von $[0;\infty)^\Omega$ und $T \in [0;\infty)^\Omega$ eine beliebige Tafel. Dann heißt $S \in \mathcal{E}$ eine *I_1-Projektion von T auf $\mathcal{E}$*, wenn gilt

$$\mathrm{I}(S|T) = \inf_{\widetilde{S} \in \mathcal{E}} \mathrm{I}(\widetilde{S}|T) < \infty. \tag{7.2.1}$$

Definition 7.2.2 (I_2-Projektion)**.** Sei Ω eine Indexmenge im Sinne von Abschnitt 6.1 (siehe S. 68). Weiter sei $S \in [0;\infty)^\Omega$ eine beliebige Tafel und $\mathcal{F}$ eine beliebige Teilmenge von $[0;\infty)^\Omega$. Dann heißt $T \in \mathcal{F}$ eine *I_2-Projektion von S auf $\mathcal{F}$*, wenn gilt

$$\mathrm{I}(S|T) = \inf_{\widetilde{T}\in\mathcal{F}} \mathrm{I}(S|\widetilde{T}) < \infty. \tag{7.2.2}$$

Die in Abschnitt 2.2 (siehe S. 27) diskutierten Eigenschaften von I-Projektionen von Vektoren gelten in analoger Weise für I-Projektionen von höherdimensionalen Tafeln.

7.3. Informationstheoretische Konvergenzanalyse

Wie bereits in Abschnitt 2.3 (siehe S. 29) wird im Folgenden die I-Divergenz zur Analyse der Konvergenz des IPF-Verfahrens angewendet. Durch Skalierung passt jeder Schritt des mehrdimensionalen IPF-Verfahrens die Einträge der aktuellen Tafel an, um die durch eine Marginaltafel vorgegebenen Randsummen zu erfüllen. Mithilfe der I-Divergenz wird untersucht, wie sich die angepassten Tafeln zu allen anderen Kandidatentafeln verhalten, die die Nulleinträge der Ausgangstafel übernehmen und die durch die jeweilige Marginaltafel vorgegebenen Randsummen erfüllen.

Für alle $i = 1, \ldots, \eta$ sei daher $\mathcal{E}_i$ die Menge aller Tafeln, die die durch die Marginaltafel $M^{(i)}$ vorgegebenen Randsummen erfüllen,

$$\forall i = 1, \ldots, \eta : \quad \mathcal{E}_i := \left\{ S \in [0;\infty)^\Omega \,\middle|\, S^{(i)} = M^{(i)} \right\}. \tag{7.3.1}$$

Außerdem sei $\mathcal{E}$ die Schnittmenge von $\mathcal{E}_1, \ldots, \mathcal{E}_\eta$,

$$\mathcal{E} := \mathcal{E}_1 \cap \cdots \cap \mathcal{E}_\eta. \tag{7.3.2}$$

Des Weiteren bezeichne $\mathcal{E}_i^{\ll A}$ für alle $i = 1, \ldots, \eta$ die Menge aller Tafeln in $\mathcal{E}_i$, die von der Ausgangstafel A dominiert werden,

$$\forall i = 1, \ldots, \eta : \quad \mathcal{E}_i^{\ll A} := \{ S \in \mathcal{E}_i \mid S \ll A \}. \tag{7.3.3}$$

Schließlich sei $\mathcal{E}^{\ll A}$ die Schnittmenge von $\mathcal{E}_1^{\ll A}, \ldots, \mathcal{E}_\eta^{\ll A}$,

$$\mathcal{E}^{\ll A} := \mathcal{E}_1^{\ll A} \cap \cdots \cap \mathcal{E}_\eta^{\ll A}. \tag{7.3.4}$$

Die so definierten Mengen $\mathcal{E}_1, \ldots, \mathcal{E}_\eta$ und $\mathcal{E}$ sowie $\mathcal{E}_1^{\ll A}, \ldots, \mathcal{E}_\eta^{\ll A}$ und $\mathcal{E}^{\ll A}$ sind kompakt und konvex. Außerdem ist für alle $i = 1, \ldots, \eta$ die IPF-Teilfolge $(A(s\eta + i))_{s\in\mathbb{N}_0}$ in $\mathcal{E}_i^{\ll A} \subseteq \mathcal{E}_i$ enthalten. Konvergiert das IPF-Verfahren, so ist die Limestafel $B^* := \lim_{t\to\infty} A(t)$ in $\mathcal{E}^{\ll A} \subseteq \mathcal{E}$ enthalten. Die Bedingung $\mathcal{E}^{\ll A} \neq \emptyset$ ist also notwendig für die Konvergenz des IPF-Verfahrens.

Für die weiteren Untersuchungen zur Konvergenz des IPF-Verfahrens legen die in Korollar 2.1.4 (siehe S. 24) und Korollar 7.1.3 gezeigten Nichtnegativitätseigenschaften sowie der Satz 6.2.1 (siehe S. 70) eine Beschränkung auf normierte mehrdimensionale Anpassungsprobleme $(A, M^{(1)}, \ldots, M^{(\eta)})$ nahe. Diese Anpassungsprobleme bestehen aus einer Wahrscheinlichkeitstafel $A \in [0;1]^\Omega$ mit $\sum_{\omega\in\Omega} a_\omega = 1$ und $\eta \geq 2$ Wahrscheinlichkeitstafeln $M^{(i)} \in (0;1)^{\Omega_i}$ mit $\sum_{\omega^*\in\Omega_i} m_{\omega^*}^{(i)} = 1$.

Lemma 7.3.1 (Randanpassungen, vgl. Cramer 2000, Theoreme 3.4 (i) und 3.8 (i))**.** *Bezeichne* $(A, M^{(1)}, \ldots, M^{(\eta)})$ *ein normiertes mehrdimensionales Anpassungsproblem,* $(A(t))$ *die zugehörige IPF-Folge und* $\mathcal{E}_1^{\ll A}, \ldots, \mathcal{E}_\eta^{\ll A}$ *die zugehörigen Mengen von Wahrscheinlichkeitstafeln. Dann gelten für alle* $s \in \mathbb{N}_0$ *und für alle* $i = 1, \ldots, \eta$ *die folgenden beiden Aussagen:*

(i) Für alle Tafeln $P \in \mathcal{E}_i^{\ll A}$ *gilt*

$$\mathrm{I}(P|A(s\eta+i-1)) = \mathrm{I}(P|A(s\eta+i)) + \mathrm{I}(A(s\eta+i)|A(s\eta+i-1)) \text{ und} \tag{7.3.5}$$

$$\mathrm{I}(A(s\eta+i-1)|P) \geq \mathrm{I}(A(s\eta+i-1)|A(s\eta+i)). \tag{7.3.6}$$

Gleichheit gilt in Ungleichung (7.3.6) *genau dann, wenn* $P = A(s\eta+i)$ *gilt.*

(ii) Die Randanpassung $A(s\eta+i)$ *ist die* I_1*-Projektion und die* I_2*-Projektion von* $A(s\eta+i-1)$ *auf* $\mathcal{E}_i^{\ll A}$.

Beweis. (i) Gemäß Definition der Menge $\mathcal{E}_i^{\ll A}$ wird die Tafel P von der Ausgangstafel A dominiert. Die Tafeln $A(s\eta+i-1)$ und $A(s\eta+i)$ sind, wie in Abschnitt 6.1 (siehe S. 68) gezeigt, maßtheoretisch äquivalent zu A. Folglich gilt $P \ll A(s\eta+i-1) \equiv A(s\eta+i)$. Daraus folgt nach Definition des IPF-Verfahrens

$$\begin{aligned}
\mathrm{I}(P|A(s\eta+i-1)) &= \sum_{\substack{\omega\in\Omega:\\ p_\omega>0}} p_\omega \ln \frac{p_\omega}{a_\omega(s\eta+i-1)} \\
&= \sum_{\substack{\omega\in\Omega:\\ p_\omega>0}} p_\omega \ln \frac{p_\omega}{a_\omega(s\eta+i)} + \sum_{\substack{\omega\in\Omega:\\ p_\omega>0}} p_\omega \ln \frac{m^{(i)}_{\omega^{(D_i)}}}{a^{(i)}_{\omega^{(D_i)}}(s\eta+i-1)} \\
&= \mathrm{I}(P|A(s\eta+i)) + \sum_{\omega^*\in\Omega_i} m^{(i)}_{\omega^*} \ln \frac{m^{(i)}_{\omega^*}}{a^{(i)}_{\omega^*}(s\eta+i-1)} \\
&= \mathrm{I}(P|A(s\eta+i)) + \sum_{\substack{\omega\in\Omega:\\ a_\omega(s\eta+i)>0}} a_\omega(s\eta+i) \ln \frac{a_\omega(s\eta+i)}{a_\omega(s\eta+i-1)} \\
&= \mathrm{I}(P|A(s\eta+i)) + \mathrm{I}(A(s\eta+i)|A(s\eta+i-1)).
\end{aligned} \tag{7.3.7}$$

Gemäß der Randsummenungleichung (siehe Satz 7.1.2 (ii)) und der Definition des IPF-Verfahrens gilt

$$\begin{aligned}
\mathrm{I}(A(s\eta+i-1)|P) &\geq \mathrm{I}(A^{(i)}(s\eta+i-1)|M^{(i)}) \\
&= \sum_{\omega^*\in\Omega_i} a^{(i)}_{\omega^*}(s\eta+i-1) \ln \frac{a^{(i)}_{\omega^*}(s\eta+i-1)}{m^{(i)}_{\omega^*}} \\
&= \sum_{\substack{\omega\in\Omega:\\ a_\omega(s\eta+i-1)>0}} a_\omega(s\eta+i-1) \ln \frac{a_\omega(s\eta+i-1)}{a_\omega(s\eta+i)} \\
&= \mathrm{I}(A(s\eta+i-1)|A(s\eta+i)).
\end{aligned} \tag{7.3.8}$$

(ii) Nach Aussage (i) gilt für alle $P \in \mathcal{E}_i^{\ll A}$ die Gleichung (7.3.5). Daraus folgt mit der Nichtnegativität der I-Divergenz (siehe Korollar 7.1.3) für alle Tafeln $P \in \mathcal{E}_i^{\ll A}$ die Ungleichung $\mathrm{I}(P|A(s\eta+i-1)) \geq \mathrm{I}(A(s\eta+i)|A(s\eta+i-1))$. Folglich ist $A(s\eta+i)$ die I_1-Projektion von $A(s\eta+i-1)$ auf $\mathcal{E}_i^{\ll A}$.

Nach Aussage (i) gilt außerdem für alle $P \in \mathcal{E}_i^{\ll A}$ die Ungleichung (7.3.6). Folglich ist $A(s\eta+i)$ die I_2-Projektion von $A(s\eta+i-1)$ auf $\mathcal{E}_i^{\ll A}$. □

Mithilfe der im vorangegangenen Lemma 7.3.1 gezeigten Gleichung (7.3.5) wird im Folgenden bewiesen, dass die Bedingung $\mathcal{E}^{\ll A} \neq \emptyset$ nicht nur notwendig sondern auch hinreichend ist für die Konvergenz des IPF-Verfahrens.

Satz 7.3.2 (Konvergenz des mehrdimensionalen IPF-Verfahrens, vgl. Brown, Chase und Pittenger 1993, Theorem 3.1)**.** *Sei $(A, M^{(1)}, \ldots, M^{(\eta)})$ ein normiertes mehrdimensionales Anpassungsproblem, $(A(t))$ die zugehörige IPF-Folge und $\mathcal{E}^{\ll A}$ die zugehörige Menge von Wahrscheinlichkeitstafeln. Dann gilt:*

(i) Die IPF-Folge $(A(t))$ konvergiert genau dann, wenn die Menge $\mathcal{E}^{\ll A}$ nichtleer ist.

Im Konvergenzfall gelten für die Limestafel $B^* := \lim_{t\to\infty} A(t)$ die folgenden vier Aussagen:

(ii) Für alle $P \in \mathcal{E}^{\ll A}$ gilt $P \ll B^* \ll A$.

(iii) Für alle $P \in \mathcal{E}^{\ll A}$ gilt die Gleichung

$$\mathrm{I}(P|A) = \mathrm{I}(P|B^*) + \mathrm{I}(B^*|A). \tag{7.3.9}$$

(iv) Die Limestafel B^* ist die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$.

(v) Es gilt $B^* \equiv A$ genau dann, wenn eine Matrix $P \in \mathcal{E}^{\ll A}$ existiert, für die $P \equiv A$ gilt.

Beweis. (i) Sei zunächst die IPF-Folge $(A(t))$ konvergent und sei $B^* := \lim_{t\to\infty} A(t)$ ihre Limestafel. Dann ist B^* in $\mathcal{E}^{\ll A}$ enthalten.

Sei nun die Menge $\mathcal{E}^{\ll A} = \mathcal{E}_1^{\ll A} \cap \cdots \cap \mathcal{E}_\eta^{\ll A}$ nichtleer. Aus der in Lemma 7.3.1 gezeigten Gleichung (7.3.5) folgt

$$\forall P \in \mathcal{E}^{\ll A}\ \forall t \geq 0: \ \mathrm{I}(P|A(t)) = \mathrm{I}(P|A(t+1)) + \mathrm{I}(A(t+1)|A(t)). \tag{7.3.10}$$

Wiederholte Anwendung dieser Aussage liefert

$$\forall P \in \mathcal{E}^{\ll A}\ \forall t \geq 0: \ \mathrm{I}(P|A) = \mathrm{I}(P|A(t+1)) + \sum_{\tau=0}^{t} \mathrm{I}(A(\tau+1)|A(\tau)). \tag{7.3.11}$$

Sei $P \in \mathcal{E}^{\ll A}$ beliebig. Nach Definition der Menge $\mathcal{E}^{\ll A}$ wird P von der Ausgangstafel A dominiert und es gilt $\mathrm{I}(P|A) < \infty$ und somit

$$\mathrm{I}(A(t+1)|A(t)) \to 0 \text{ für } t \to \infty. \tag{7.3.12}$$

Mit der Pinsker-Ungleichung (siehe Satz 2.1.7, S. 25) folgt

$$\|A(t+1) - A(t)\|_1 \to 0 \text{ für } t \to \infty. \tag{7.3.13}$$

Da die IPF-Folge $(A(t))$ in der kompakten Menge $[0;1]^\Omega$ enthalten ist, besitzt sie mindestens einen Häufungspunkt. Sei B^* ein solcher Häufungspunkt. Dann existiert eine Schrittfolge (t_m), sodass gilt

$$A(t_m) \to B^* \text{ für } m \to \infty. \tag{7.3.14}$$

Mit der Konvergenzaussage (7.3.13) folgen daraus die Aussagen

$$\begin{aligned} A(t_m+1) &\to B^* \text{ für } m \to \infty, \\ A(t_m+2) &\to B^* \text{ für } m \to \infty, \\ &\vdots \\ A(t_m+\eta-1) &\to B^* \text{ für } m \to \infty. \end{aligned} \tag{7.3.15}$$

Mindestens eine der Schrittfolgen $(t_m), \dots, (t_m+\eta-1)$ enthält unendlich viele Vielfache von η. Folglich gilt $B^* \in \mathcal{E}_\eta^{\ll A}$. Analog gelten $B^* \in \mathcal{E}_1^{\ll A}, \dots, B^* \in \mathcal{E}_{\eta-1}^{\ll A}$ und somit $B^* \in \mathcal{E}^{\ll A}$.

Nach Gleichung (7.3.10) gilt damit

$$\forall t \geq 0: \ \mathrm{I}(B^*|A(t)) = \mathrm{I}(B^*|A(t+1)) + \mathrm{I}(A(t+1)|A(t)). \tag{7.3.16}$$

Somit ist die Folge $(\mathrm{I}(B^*|A(t)))$ monoton fallend. Mit der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v), S. 23) folgt

$$\lim_{t\to\infty} \mathrm{I}(B^*|A(t)) = \lim_{m\to\infty} \mathrm{I}(B^*|A(t_m)) = \mathrm{I}(B^*|B^*) = 0. \tag{7.3.17}$$

Daraus folgt mit der Pinsker-Ungleichung (siehe Satz 2.1.7, S. 25) die Konvergenz der IPF-Folge $(A(t))$ gegen B^*.

(ii) Sei die IPF-Folge $(A(t))$ konvergent und sei $B^* := \lim_{t\to\infty} A(t) \in \mathcal{E}^{\ll A}$ ihre Limestafel. Weiter sei $P \in \mathcal{E}^{\ll A}$ beliebig. Dann gilt gemäß Definition der Menge $\mathcal{E}^{\ll A}$ die Ungleichung $\mathrm{I}(P|A(0)) = \mathrm{I}(P|A) < \infty$. Nach Gleichung (7.3.10) ist die Folge $(\mathrm{I}(P|A(t)))$ außerdem monoton fallend. Mit der Stetigkeit der I-Divergenz im zweiten Argument (siehe Satz 2.1.2 (v), S. 23) folgt daraus

$$\mathrm{I}(P|B^*) = \lim_{t\to\infty} \mathrm{I}(P|A(t)) \leq \mathrm{I}(P|A(0)) < \infty. \tag{7.3.18}$$

Dies impliziert die Aussage $P \ll B^* \ll A$.

(iii) Die Aussage $P \ll B^* \ll A$ erlaubt die Rechnung

$$\begin{aligned}\mathrm{I}(P|B^*) + \mathrm{I}(B^*|A) &= \sum_{\substack{\omega\in\Omega:\\ p_\omega>0}} p_\omega \ln \frac{p_\omega}{b^*_\omega} + \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} b^*_\omega \ln \frac{b^*_\omega}{a_\omega} \\ &= \sum_{\substack{\omega\in\Omega:\\ p_\omega>0}} p_\omega \ln \frac{p_\omega}{a_\omega} + \sum_{\substack{\omega\in\Omega:\\ p_\omega>0}} p_\omega \ln \frac{a_\omega}{b^*_\omega} + \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} b^*_\omega \ln \frac{b^*_\omega}{a_\omega} \\ &= \mathrm{I}(P|A) + \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} \left(b^*_\omega - p_\omega\right) \ln \frac{b^*_\omega}{a_\omega}. \end{aligned} \tag{7.3.19}$$

Da B^* die Limestafel der IPF-Folge $(A(t))$ ist, gilt

$$\begin{aligned}\sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} \left(b^*_\omega - p_\omega\right) \ln \frac{b^*_\omega}{a_\omega} &= \lim_{s\to\infty} \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} \left(b^*_\omega - p_\omega\right) \ln \frac{a_\omega(s\eta)}{a_\omega} \\ &= \lim_{s\to\infty} \sum_{\sigma=0}^{s-1} \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} \left(b^*_\omega - p_\omega\right) \ln \frac{a_\omega(\sigma\eta+\eta)}{a_\omega(\sigma\eta)} \\ &= \lim_{s\to\infty} \sum_{\sigma=0}^{s-1} \sum_{i=0}^{\eta-1} \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} \left(b^*_\omega - p_\omega\right) \ln \frac{a_\omega(\sigma\eta+i+1)}{a_\omega(\sigma\eta+i)}. \end{aligned} \tag{7.3.20}$$

Für alle $\sigma \in \mathbb{N}_0$ und für alle $i = 0, \ldots, \eta - 1$ ist nach Definition des IPF-Verfahrens der Term $\ln(a_\omega(\sigma\eta + i + 1)/a_\omega(\sigma\eta + i))$ über alle $\omega \in \Omega$ mit derselben Komponente $\omega^{(D_{i+1})}$ konstant. Da die Tafeln B^* und P jedoch beide in $\mathcal{E}^{\ll A} \subseteq \mathcal{E}^{\ll A}_{i+1}$ enthalten sind, gilt

$$\forall \omega^* \in \Omega_{i+1} = \Omega_{D_{i+1}} : \quad \sum_{\omega^{**}\in\Omega_{D'_{i+1}}} b^*_{(\omega^*,\omega^{**})} = \sum_{\omega^{**}\in\Omega_{D'_{i+1}}} p_{(\omega^*,\omega^{**})}. \tag{7.3.21}$$

Damit gilt

$$\forall s \in \mathbb{N}_0 \quad \forall i = 0, \ldots, \eta - 1 \qquad \sum_{\substack{\omega\in\Omega:\\ b^*_\omega>0}} \left(b^*_\omega - p_\omega\right) \ln \frac{a_\omega(\sigma\eta+i+1)}{a_\omega(\sigma\eta+i)} = 0. \tag{7.3.22}$$

Daraus folgt Gleichung (7.3.9).

(iv) Nach Aussage (iii) gilt für alle $P \in \mathcal{E}^{\ll A}$ die Gleichung (7.3.9). Daraus folgt mit der Nichtnegativität der I-Divergenz (siehe Korollar 7.1.3) für alle $P \in \mathcal{E}^{\ll A}$ die Ungleichung $\mathrm{I}(P|A) \geq \mathrm{I}(B^*|A)$. Folglich ist B^* die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$.

(v) Sei die IPF-Folge $(A(t))$ konvergent und sei $B^* := \lim_{t\to\infty} A(t) \in \mathcal{E}^{\ll A}$ ihr Grenzwert. Gilt $B^* \equiv A$, so erfüllt B^* die Voraussetzungen an P. Existiert eine Tafel $P \in \mathcal{E}^{\ll A}$ mit $P \equiv A$, so folgt mit Aussage (ii) die Aussage $P \equiv B^* \equiv A$. □

7.4. L_1-Ansatz zur Konvergenzanalyse

Zur Messung des Fortschritts des mehrdimensionalen IPF-Verfahrens bei der Anpassung an die Marginaltafeln bietet sich eine Verallgemeinerung des in Abschnitt 2.6 (siehe S. 38) beschriebenen L_1-Ansatzes an. Wieder wird auf die Normierung des Anpassungsproblems $(A, M^{(1)}, \ldots, M^{(\eta)})$ verzichtet. Der zugehörige *L_1-Fehler* wird für alle Tafeln $S \in [0;\infty)^\Omega$ definiert gemäß

$$f(S) := \sum_{i=1}^{\eta} \left\| S^{(i)} - M^{(i)} \right\|_1 . \tag{7.4.1}$$

Dabei bezeichnet $\|S^{(i)} - M^{(i)}\|_1$ den L_1-Abstand der Randtafel $S^{(i)}$ und der Marginaltafel $M^{(i)}$. Für alle Tafeln $S \in [0;\infty)^\Omega$ gilt $f(S) \geq 0$. Gleichheit gilt genau dann, wenn die Tafel S alle durch die Marginaltafeln $M^{(i)}$, $i = 1, \ldots, \eta$, vorgegebenen Randsummen erfüllt. Die Anpassungsgüte der IPF-Folge $(A(t))$ wird mithilfe der *L_1-Fehlerfolge* $(f(A(t)))$ untersucht. Wie das folgende Lemma zeigt, ist die L_1-Fehlerfolge $(f(A(t)))$ im Fall $\eta = 2$ monoton fallend und beschränkt.

Satz 7.4.1 (Monotonie der L_1-Fehlerfolge bei Anpassungsproblemen mit $\eta = 2$ Marginaltafeln)**.** *Sei $(A, M^{(1)}, M^{(2)})$ ein mehrdimensionales Anpassungsproblem mit zwei Marginaltafeln und $(A(t))$ die zugehörige IPF-Folge. Dann fällt die L_1-Fehlerfolge $(f(A(t)))$ monoton,*

$$\forall t \geq 1: \quad f(A(t)) \geq f(A(t+1)). \tag{7.4.2}$$

Beweis. Seien D_1 und D_2 die zu $M^{(1)}$ und $M^{(2)}$ gehörenden Ränder. Dann wählen wir

$$D_{00} := \{1, \ldots, d\} \setminus D_1 \setminus D_2, \qquad D_{11} := D_1 \setminus D_2, \tag{7.4.3}$$

$$D_{12} := D_1 \cap D_2, \qquad D_{22} := D_2 \setminus D_1. \tag{7.4.4}$$

Folglich gilt $D_{00} \uplus D_{11} \uplus D_{12} \uplus D_{22} = \{1, \ldots, d\}$. Weiter bezeichnen wir die zu D_{00}, D_{11}, D_{12} und D_{22} gehörenden Indexmengen mit Ω_{00}, Ω_{11}, Ω_{12} und Ω_{22}. Außerdem sei $t \geq 2$ ein beliebiger gerader Schritt. Dann gilt für den L_1-Fehler $f(A(t))$ die Aussage

$$\begin{aligned} f(A(t)) &= \left\| A^{(1)}(t) - M^{(1)} \right\|_1 = \sum_{\omega^{12} \in \Omega_{12}} \sum_{\omega^{11} \in \Omega_{11}} \left| a^{(1)}_{(\omega^{11},\omega^{12})}(t) - m^{(1)}_{(\omega^{11},\omega^{12})} \right| \\ &= \sum_{\omega^{12} \in \Omega_{12}} \sum_{\omega^{11} \in \Omega_{11}} \left| \sum_{\omega^{22} \in \Omega_{22}} \sum_{\omega^{00} \in \Omega_{00}} a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t) - a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t+1) \right|. \end{aligned} \tag{7.4.5}$$

Da die Tafel $A(t+1)$ nach Definition des IPF-Verfahrens durch Skalierung entlang des Randes $D_1 = D_{11} \uplus D_{12}$ aus $A(t)$ hervorgeht, hängt das Vorzeichen der Differenz $a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t) - a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t+1)$ nur von ω^{11} und ω^{12} ab. Daraus folgt

$$\sum_{\omega^{12} \in \Omega_{12}} \sum_{\omega^{11} \in \Omega_{11}} \left| \sum_{\omega^{22} \in \Omega_{22}} \sum_{\omega^{00} \in \Omega_{00}} a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t) - a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t+1) \right|$$

$$
\begin{aligned}
&= \sum_{\omega^{12}\in\Omega_{12}} \sum_{\omega^{11}\in\Omega_{11}} \sum_{\omega^{22}\in\Omega_{22}} \sum_{\omega^{00}\in\Omega_{00}} \left| a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t) - a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t+1) \right| \\
&\geq \sum_{\omega^{12}\in\Omega_{12}} \sum_{\omega^{22}\in\Omega_{22}} \Big| \sum_{\omega^{11}\in\Omega_{11}} \sum_{\omega^{00}\in\Omega_{00}} a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t) - a_{(\omega^{00},\omega^{11},\omega^{12},\omega^{22})}(t+1) \Big| \\
&= \sum_{\omega^{12}\in\Omega_{12}} \sum_{\omega^{22}\in\Omega_{22}} \left| m^{(2)}_{(\omega^{12},\omega^{22})} - a^{(2)}_{(\omega^{12},\omega^{22})}(t+1) \right| \\
&= \left\| M^{(2)} - A^{(2)}(t+1) \right\|_1 = f(A(t+1)) . \qquad (7.4.6)
\end{aligned}
$$

Somit gilt für alle geraden Schritte $t \geq 2$ die Ungleichung $f(A(t)) \geq f(A(t+1))$. Für alle ungeraden Schritte $t \geq 1$ folgt die Ungleichung in analoger Weise. □

Das folgende Beispiel zeigt, dass die L_1-Fehlerfolge für mehrdimensionale Anpassungsprobleme $(A, M^{(1)}, \ldots, M^{(\eta)})$ mit $\eta = 3$ Marginaltafeln im Allgemeinen nicht monoton fallend ist. Der L_1-Ansatz kann daher nicht auf das mehrdimensionale IPF-Verfahren mit einer beliebigen Anzahl $\eta \geq 2$ von Marginaltafeln verallgemeinert werden.

Beispiel 7.4.2 (Nichtmonotones Verhalten der L_1-Fehlerfolge bei Anpassungsproblemen mit $\eta = 3$ Marginaltafeln). Sei $\Omega := \{1, 2\}^3$ und sei die dreidimensionale Tafel $A \in [0; \infty)^{\Omega}$ definiert gemäß

$$
a_\omega := \begin{cases} 1 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 1, \\ 3 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 2, \\ 0 & \text{sonst.} \end{cases} \qquad (7.4.7)
$$

Weiter seien die Ränder $D_1 := \{1\}$, $D_2 := \{2\}$ und $D_3 := \{3\}$ gegeben. Die zugehörigen eindimensionalen Marginaltafeln seien definiert gemäß $M^{(1)} := (9, 1)$, $M^{(2)} := (5, 5)$ und $M^{(3)} := (6, 4)$. Der zu dem Anpassungsproblem $(A, M^{(1)}, M^{(2)}, M^{(3)})$ gehörende L_1-Fehler berechnet sich dann für alle Tafeln $S \in [0; \infty)^{\Omega}$ gemäß der Formel

$$
\begin{aligned}
f(S) &= \left\| S^{(1)} - M^{(1)} \right\|_1 + \left\| S^{(2)} - M^{(2)} \right\|_1 + \left\| S^{(3)} - M^{(3)} \right\|_1 \\
&= \left| s_{111} + s_{112} + s_{121} + s_{122} - 9 \right| + \left| s_{211} + s_{212} + s_{221} + s_{222} - 1 \right| \\
&\quad + \left| s_{111} + s_{112} + s_{211} + s_{212} - 5 \right| + \left| s_{121} + s_{122} + s_{221} + s_{222} - 5 \right| \\
&\quad + \left| s_{111} + s_{121} + s_{211} + s_{221} - 6 \right| + \left| s_{112} + s_{122} + s_{212} + s_{222} - 4 \right| . \qquad (7.4.8)
\end{aligned}
$$

Aufgrund der Diagonalstruktur der Ausgangstafel A sind die IPF-Teilfolgen $(A(3s+1))$, $(A(3s+2))$ und $(A(3s+3))$ konstant. Für alle $s \in \mathbb{N}_0$ gilt

$$
a_\omega(3s+1) = \begin{cases} 9 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 1, \\ 1 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 2, \\ 0 & \text{sonst;} \end{cases} \qquad (7.4.9)
$$

$$
a_\omega(3s+2) = \begin{cases} 5 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 1, \\ 5 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 2, \\ 0 & \text{sonst;} \end{cases} \qquad (7.4.10)
$$

$$a_\omega(3s+3) = \begin{cases} 6 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 1, \\ 4 & \text{falls } \omega_1 = \omega_2 = \omega_3 = 2, \\ 0 & \text{sonst.} \end{cases} \tag{7.4.11}$$

Die Teilfolgen $(f(A(3s+1)))$, $(f(A(3s+2)))$ und $(f(A(3s+3)))$ der L_1-Fehlerfolge sind dementsprechend ebenfalls konstant. Für alle $s \in \mathbb{N}_0$ gilt

$$f(A(3s+1)) = |9-9| + |1-1| + |9-5| + |1-5| + |9-6| + |1-4| = 14, \tag{7.4.12}$$

$$f(A(3s+2)) = |5-9| + |5-1| + |5-5| + |5-5| + |5-6| + |5-4| = 10 \text{ und} \tag{7.4.13}$$

$$f(A(3s+3)) = |6-9| + |4-1| + |6-5| + |4-5| + |6-6| + |4-4| = 8. \tag{7.4.14}$$

Die L_1-Fehlerfolge $(f(A(t)))$ springt also zwischen den Werten 14, 10 und 8 und ist somit nicht monoton fallend.

7.5. Kommentare und Referenzen

Die in den Abschnitten 7.1 und 7.2 vorgestellten I-Divergenzen und I-Projektionen von Tafeln sind Verallgemeinerungen der in Abschnitt 2.1 (siehe S. 22) und in Abschnitt 2.2 (siehe S. 27) diskutierten I-Divergenzen und I-Projektionen von Matrizen.

Die in Abschnitt 7.3 beschriebene informationstheoretische Konvergenzanalyse des mehrdimensionalen IPF-Verfahrens basiert, wie bereits die informationstheoretische Konvergenzanalyse des zweidimensionalen IPF-Verfahrens, auf der Arbeit von Csiszár (1975, Theorem 3.2) und Brown, Chase und Pittenger (1993, Theorem 3.1). In Lemma 7.3.1 könnte man die Mengen $\mathcal{E}_i^{\ll A}$, $i = 1, \ldots, \eta$, durch $\mathcal{E}_i$ ersetzen. Allerdings könnten dann in Gleichung (7.3.5) beide Seiten unendlich sein. Satz 7.3.2 könnte ebenfalls angepasst werden. Die in Satz 7.3.2 zusammengefassten Resultate wurden erstmals von Csiszár (1975, Theorem 3.2) gezeigt. Für die Unterschiede zwischen Csiszárs Beweis und dem Beweis in Abschnitt 7.3, der der Argumentation von Brown, Chase und Pittenger (1993, Theorem 3.1) folgt, sei auf die Kommentare und Referenzen in Abschnitt 2.7 (siehe S. 41) verwiesen. Das Theorem 3.1 von Brown, Chase und Pittenger ist jedoch allgemeiner als Satz 7.3.2 (i), da Brown, Chase und Pittenger nicht nur zyklische Skalierungsfolgen $\sigma = (1, 2, \ldots, \eta-1, \eta, 1, 2, \ldots, \eta-1, \eta, \ldots)$ betrachten. Stattdessen erlaubt die von ihnen untersuchte Variante des IPF-Verfahrens beliebige Skalierungsfolgen $\sigma \subseteq \{1, \ldots, \eta\}$. Für den Fall, dass jedes Element der Menge $\{1, \ldots, \eta\}$ unendlich oft in der Skalierungsfolge σ auftritt und dass $\mathcal{E}^{\ll A} \neq \emptyset$ gilt, zeigen Brown, Chase und Pittenger die Konvergenz der resultierenden IPF-Folge. Da der Fokus dieser Dissertation auf dem Divergenzfall $\mathcal{E}^{\ll A} = \emptyset$ liegt und Simulationen bei nichtzyklischer Skalierungsfolge im Divergenzfall keinerlei Regelmäßigkeit der IPF-Folge erkennen lassen, wird diese Verallgemeinerung des mehrdimensionalen IPF-Verfahrens hier nicht näher behandelt. Im Gegensatz zu den Konvergenzbedingungen des zweidimensionalen IPF-Verfahrens (siehe Satz 2.3.2 (i), S. 31) kann für die Konvergenzbedingungen des mehrdimensionalen IPF-Verfahrens (siehe Satz 7.3.2 (i)) kein einfacheres Kriterium angegeben werden. Dies liegt daran, dass kein mehrdimensionales Analogon der Flussungleichungen (siehe Abschnitt 2.4, S. 35) bekannt ist. Die in Abschnitt 2.5 (siehe S. 35) beschriebene Konvergenzanalyse des zweidimensionalen IPF-Verfahrens mithilfe des geometrischen Mittels kann dagegen auf das mehrdimensionale IPF-Verfahren übertragen werden. Aufgrund der Äquivalenz dieses Ansatzes zum informationstheoretischen Ansatz wurde hier jedoch auf eine ausführliche Beschreibung dieser Verallgemeinerung verzichtet.

Die in Abschnitt 7.4 präsentierten Überlegungen zur Verallgemeinerbarkeit des in Abschnitt 2.6 (siehe S. 38) diskutierten L_1-Ansatzes zur Konvergenzanalyse des zweidimensionalen IPF-Verfahrens sind bisher unveröffentlicht.

8. Strukturanalyse der Grenzwerte des mehrdimensionalen IPF-Verfahrens

Auf die Konvergenzanalyse des mehrdimensionalen IPF-Verfahrens im vorangegangenen Kapitel folgt in diesem Kapitel die Strukturanalyse der Limestafel. Dazu wird im ersten Abschnitt die stetige Abhängigkeit der Limestafel von der Ausgangstafel gezeigt. Der zweite Abschnitt diskutiert die Verallgemeinerbarkeit der weiteren die Limesmatrix betreffenden Aussagen von Kapitel 4 (siehe S. 50).

8.1. Stetige Abhängigkeit der Limestafel von der Ausgangstafel

Wie beim zweidimensionalen IPF-Verfahren stellt sich auch beim mehrdimensionalen IPF-Verfahren die Frage nach der stetigen Abhängigkeit der Limestafel von der Ausgangstafel. Der folgende Satz nutzt die Stetigkeit von I_1-Projektionen (siehe Satz 2.2.4, S. 28), um diese Abhängigkeit zu zeigen.

Satz 8.1.1 (Stetige Abhängigkeit der Limestafel von der Ausgangstafel)**.** *Sei* $(A, M^{(1)}, \dots, M^{(\eta)})$ *ein mehrdimensionales Anpassungsproblem, für das das IPF-Verfahren konvergiert, und sei* $B := \lim_{t\to\infty} A(t)$ *die Limestafel der zugehörigen IPF-Folge* $(A(t))$*. Weiter sei* (A^n) *eine Folge von Tafeln in* $[0;\infty)^\Omega$*, die gegen A konvergiert. Dann gelten die folgenden beiden Aussagen:*

(i) Für hinreichend große $n \in \mathbb{N}$ *konvergiert die zum Anpassungsproblem* $(A^n, M^{(1)}, \dots, M^{(\eta)})$ *gehörende IPF-Folge* $(A^n(t))$ *gegen eine Tafel* B^n*.*

(ii) Die Folge (B^n) *der Limestafeln konvergiert gegen die Limestafel B,*

$$\lim_{n\to\infty} B^n = B. \tag{8.1.1}$$

Beweis. Die Anpassungsprobleme $(A, M^{(1)}, \dots, M^{(\eta)})$ und $(A^n, M^{(1)}, \dots, M^{(\eta)})$, $n \in \mathbb{N}$, seien ohne Beschränkung der Allgemeinheit normiert. Außerdem seien $\mathcal{E}$ sowie $\mathcal{E}^{\ll A}$ und $\mathcal{E}^{\ll A^n}$, $n \in \mathbb{N}$, die zugehörigen Mengen von Wahrscheinlichkeitstafeln.

(i) Nach Voraussetzung ist B die Limestafel der IPF-Folge $(A(t))$. Dementsprechend ist B in der Menge $\mathcal{E}^{\ll A}$ enthalten. Für hinreichend große $n \in \mathbb{N}$ gilt weiter $A \ll A^n$ und damit $\mathcal{E}^{\ll A} \subseteq \mathcal{E}^{\ll A^n}$. Folglich gilt $B \in \mathcal{E}^{\ll A^n}$. Nach Satz 7.3.2 (i) (siehe S. 78) konvergiert somit die IPF-Folge $(A^n(t))$. Die Limestafel der IPF-Folge $(A^n(t))$ bezeichnen wir im Folgenden mit B^n.

(ii) Nach Satz 7.3.2 (iv) (siehe S. 78) ist die Limestafel B die I_1-Projektion von A auf $\mathcal{E}^{\ll A}$. Da die I-Divergenz $\mathrm{I}(S|A)$ für alle nicht von A dominierten Tafeln S unendlich ist, ist B außerdem die I_1-Projektion von A auf die Menge $\mathcal{E}$. In analoger Weise sind die Limestafeln B^n die I_1-Projektionen von A^n auf $\mathcal{E}$. Aus der Stetigkeit der I_1-Projektion auf $\mathcal{E}$ (siehe Satz 2.2.4, S. 28) folgt die Konvergenz der Tafelfolge (B^n) gegen die Tafel B. □

8.2. Weitere Struktureigenschaften der Limestafel

Im vorangegangenen Abschnitt wurde die stetige Abhängigkeit der Limestafel B des mehrdimensionalen IPF-Verfahrens von der Ausgangstafel A untersucht. Genau wie A können auch die Marginaltafeln $M^{(i)}, i = 1, \ldots, \eta$, Fehler enthalten. Für die Untersuchung der stetigen Abhängigkeit der Limestafel des mehrdimensionalen IPF-Verfahrens von den Marginaltafeln ist jedoch kein zielführender Ansatz bekannt. Der in Abschnitt 4.4 (siehe S. 59) verwendete Ansatz zur Untersuchung der stetigen Abhängigkeit der Limesmatrix von der Ausgangsmatrix und den Marginalien basiert auf den Flussungleichungen. Da für die Flussungleichungen kein mehrdimensionales Analogon bekannt ist, kann dieser Ansatz nicht verallgemeinert werden.

Die in Abschnitt 4.1 (siehe S. 50) beschriebene Analyse der Zusammenhangsstruktur der Limesmatrix basiert auf dem Zusammenhangsbegriff für Matrizen. Eine Verallgemeinerung auf das mehrdimensionale IPF-Verfahren würde einen Zusammenhangsbegriff für mehrdimensionale Tafeln voraussetzen. Wie bereits in Abschnitt 6.3 (siehe S. 72) erwähnt wurde, ist jedoch unklar, wie die Definition dieses Zusammenhangsbegriffs aussehen sollte.

8.3. Kommentare und Referenzen

Satz 8.1.1 ist eine Verallgemeinerung von Satz 4.3.1 (siehe S. 58). Beide Sätze basieren auf der in Satz 2.2.4 (siehe S. 28) gezeigten Stetigkeit von I_1-Projektionen. Man beachte daher die Kommentare und Referenzen in Abschnitt 2.7 (siehe S. 41) und in Abschnitt 4.5 (siehe S. 64).

9. Divergenzanalyse des mehrdimensionalen IPF-Verfahrens

Nachdem in den letzten beiden Kapiteln die Konvergenz des mehrdimensionalen IPF-Verfahrens analysiert wurde, wird in diesem Kapitel das asymptotische Verhalten des mehrdimensionalen IPF-Verfahrens im Divergenzfall untersucht. Diese Untersuchung erfolgt anhand von zwei anschaulichen $2 \times 2 \times 2$-Beispielen. Für mehrdimensionale Anpassungsprobleme mit einer positiven $2 \times 2 \times 2$-Ausgangstafel und drei identischen, doppelt symmetrischen 2×2-Marginaltafeln zeigt der erste Abschnitt, dass die zugehörige IPF-Folge maximal drei Häufungspunkte hat. Der zweite Abschnitt zeigt dieselbe Aussage für mehrdimensionale Anpassungsprobleme mit einer $2{\times}2{\times}2$-Ausgangstafel mit maximal drei positiven Einträgen und drei eindimensionalen Marginaltafeln.

9.1. Mehrdimensionale Anpassungsprobleme mit positiver $2 \times 2 \times 2$-Ausgangstafel und identischen, doppelt symmetrischen 2×2-Marginaltafeln

In diesem Kapitel werden, wie bereits in Beispiel 7.4.2 (siehe S. 82), mehrdimensionale Anpassungsprobleme mit $2 \times 2 \times 2$-Ausgangstafeln untersucht. Nach der in Abschnitt 6.1 (siehe S. 68) eingeführten Notation gilt somit $\Omega = \{1, 2\} \times \{1, 2\} \times \{1, 2\}$. Statt $S \in [0; \infty)^{\Omega}$ wird jedoch der Einfachheit halber ab jetzt für alle Tafeln die Schreibweise $S \in [0; \infty)^{2\times2\times2}$ verwendet. Die acht Einträge einer jeden Tafel $S \in [0; \infty)^{2\times2\times2}$ werden mit s_{111} bis s_{222} bezeichnet und angeordnet gemäß der Vorschrift

$$S = \left(\begin{array}{cc|cc} s_{111} & s_{121} & s_{112} & s_{122} \\ s_{211} & s_{221} & s_{212} & s_{222} \end{array}\right). \tag{9.1.1}$$

Die Einträge s_{ijk} mit gemeinsamem Index i werden als *Zeile* bezeichnet; die Einträge s_{ijk} mit gemeinsamem Index j als *Spalte* und die Einträge s_{ijk} mit gemeinsamem Index k als *Lage*.

Speziell in diesem Abschnitt werden Anpassungsprobleme $(A, M^{(1)}, M^{(2)}, M^{(3)})$ untersucht, die aus einer beliebigen positiven Ausgangstafel $A \in (0; \infty)^{2\times2\times2}$ und drei identischen, doppelt symmetrischen zweidimensionalen Marginaltafeln $M^{(1)}$, $M^{(2)}$ und $M^{(3)}$ bestehen. Die Marginaltafeln $M^{(1)}$, $M^{(2)}$ und $M^{(3)}$ beziehen sich dabei auf die Ränder $D_1 := \{1, 2\}$, $D_2 := \{2, 3\}$ und $D_3 := \{1, 3\}$. Sie sind definiert gemäß

$$M^{(1)} := M^{(2)} := M^{(3)} := M := \begin{pmatrix} \epsilon & \frac{1}{2} - \epsilon \\ \frac{1}{2} - \epsilon & \epsilon \end{pmatrix} \in (0; \infty)^{2\times2}, \tag{9.1.2}$$

wobei $\epsilon \in (0; 1/2)$ gilt. Diese Klasse von mehrdimensionalen Anpassungsproblemen geht auf Asci und Piccioni (2003) zurück.

Die zugehörige Menge $\mathcal{E}^{\ll A} = \mathcal{E}_1^{\ll A} \cap \mathcal{E}_2^{\ll A} \cap \mathcal{E}_3^{\ll A}$ ist gleich der Menge aller Tafeln $S \in [0;\infty)^{2\times2\times2}$, für die gilt

$$S = \left(\begin{matrix} s_{111} & \epsilon - s_{111} \\ \epsilon - s_{111} & s_{111} - 2\epsilon + \frac{1}{2} \end{matrix}\middle| \begin{matrix} \epsilon - s_{111} & s_{111} - 2\epsilon + \frac{1}{2} \\ s_{111} - 2\epsilon + \frac{1}{2} & 3\epsilon - s_{111} - \frac{1}{2} \end{matrix}\right). \tag{9.1.3}$$

Der Eintrag s_{111} kann dabei frei gewählt werden. Jedoch muss die Bedingung $S \in [0;\infty)^{2\times2\times2}$ berücksichtigt werden. Deshalb gilt

$$s_{111} \geq 0, \qquad \epsilon - s_{111} \geq 0, \qquad s_{111} - 2\epsilon + \tfrac{1}{2} \geq 0, \qquad 3\epsilon - s_{111} - \tfrac{1}{2} \geq 0. \tag{9.1.4}$$

Äquivalent gilt

$$0 \leq s_{111}, \qquad s_{111} \leq \epsilon, \qquad 2\epsilon - \tfrac{1}{2} \leq s_{111}, \qquad s_{111} \leq 3\epsilon - \tfrac{1}{2}. \tag{9.1.5}$$

Diese vier Ungleichungen können zu einer Ungleichungskette zusammengefasst werden,

$$\max\{0, 2\epsilon - 1/2\} \leq s_{111} \leq \min\{\epsilon, 3\epsilon - 1/2\}. \tag{9.1.6}$$

Für $\epsilon \in (0; 1/6)$ ist die Menge $\mathcal{E}^{\ll A}$ folglich leer. Für $\epsilon = 1/6$ enthält die Menge $\mathcal{E}^{\ll A}$ als einziges Element die Tafel

$$S = \left(\begin{matrix} 0 & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} \end{matrix}\middle| \begin{matrix} \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & 0 \end{matrix}\right). \tag{9.1.7}$$

Für $\epsilon \in (1/6; 1/2)$ enthält die Menge $\mathcal{E}^{\ll A}$ überabzählbar viele Tafeln.

Bei Anwendung auf das mehrdimensionale Anpassungsproblem $(A, M^{(1)}, M^{(2)}, M^{(3)})$ berechnet das in Abschnitt 6.1 (siehe S. 68) definierte IPF-Verfahren ausgehend von $A(0) := A$ die IPF-Folge $(A(t))$ durch die iterative Ausführung der Schritte

$$a_{ijk}(3s+1) := \frac{m_{ij}}{a_{ij+}(3s)} \cdot a_{ijk}(3s) \text{ für alle Einträge } (i,j,k), \tag{9.1.8}$$

$$a_{ijk}(3s+2) := \frac{m_{jk}}{a_{+jk}(3s+1)} \cdot a_{ijk}(3s+1) \text{ für alle Einträge } (i,j,k) \text{ und} \tag{9.1.9}$$

$$a_{ijk}(3s+3) := \frac{m_{ik}}{a_{i+k}(3s+2)} \cdot a_{ijk}(3s+2) \text{ für alle Einträge } (i,j,k). \tag{9.1.10}$$

Für die Analyse des IPF-Verfahrens empfiehlt sich folglich die Betrachtung der Operatoren

$$T_1(S) := S^*, \text{ wobei } s^*_{ijk} := \frac{m_{ij}}{s_{ij+}} \cdot s_{ijk} \text{ für alle Einträge } (i,j,k), \tag{9.1.11}$$

$$T_2(S) := S^*, \text{ wobei } s^*_{ijk} := \frac{m_{jk}}{s_{+jk}} \cdot s_{ijk} \text{ für alle Einträge } (i,j,k), \text{ und} \tag{9.1.12}$$

$$T_3(S) := S^*, \text{ wobei } s^*_{ijk} := \frac{m_{ik}}{s_{i+k}} \cdot s_{ijk} \text{ für alle Einträge } (i,j,k). \tag{9.1.13}$$

Die so definierten Operatoren T_1, T_2 und T_3 sind auf der gesamten Menge $[0;\infty)^{2\times2\times2}$ wohldefiniert mit Ausnahme der Ränder, für die mindestens eine der jeweiligen Randsummen s_{ij+}, s_{+jk} oder s_{i+k} den Wert 0 annimmt. Auf diesen Definitionsmengen sind die Operatoren T_1, T_2 und T_3 stetig.

Für den Fall $\epsilon \in [1/6; 1/2)$ zeigen Asci und Piccioni (2003) die Konvergenz des IPF-Verfahrens.

Satz 9.1.1 (Konvergenz der IPF-Folge, vgl. Asci und Piccioni 2003, Proposition 1)**.** *Bezeichne* $(A, M^{(1)}, M^{(2)}, M^{(3)})$ *das oben beschriebene mehrdimensionale Anpassungsproblem und sei* $\epsilon \in [1/6; 1/2)$. *Dann konvergiert die zugehörige IPF-Folge* $(A(t))$.

Beweis. Ohne Beschränkung der Allgemeinheit sei die Ausgangstafel A normiert. Gemäß der obigen Diskussion ist die zu dem Anpassungsproblem $(A, M^{(1)}, M^{(2)}, M^{(3)})$ gehörende Menge $\mathcal{E}^{\ll A} = \mathcal{E}_1^{\ll A} \cap \mathcal{E}_2^{\ll A} \cap \mathcal{E}_3^{\ll A}$ nichtleer. Daraus folgt nach Satz 7.3.2 (i) (siehe S. 78) die Konvergenz der IPF-Folge $(A(t))$. □

Für den Fall $\epsilon \in (0; 1/6]$ zeigen Asci und Piccioni (2003) die Konvergenz der IPF-Teilfolgen $(A(3s))$, $(A(3s + 1))$ und $(A(3s + 2))$.

Satz 9.1.2 (Konvergenz der IPF-Teilfolgen, vgl. Asci und Piccioni 2003, Theorem 2)**.** *Bezeichne* $(A, M^{(1)}, M^{(2)}, M^{(3)})$ *das oben beschriebene mehrdimensionale Anpassungsproblem und sei* $\epsilon \in (0; 1/6]$. *Weiter sei* $(A(t))$ *die zugehörige IPF-Folge. Dann sind die IPF-Teilfolgen* $(A(3s))$, $(A(3s + 1))$ *und* $(A(3s + 2))$ *konvergent. Für die Häufungspunkte* $B(0) := \lim_{s\to\infty} A(3s)$, $B(1) := \lim_{s\to\infty} A(3s + 1)$ *und* $B(2) := \lim_{s\to\infty} A(3s + 2)$ *gilt*

$$B(0) = \left(\begin{array}{cc|cc} 0 & \epsilon & \frac{-\epsilon+1-\sqrt{-3\epsilon^2+2\epsilon}}{2} & \frac{-\epsilon+\sqrt{-3\epsilon^2+2\epsilon}}{2} \\ \frac{-\epsilon+\sqrt{-3\epsilon^2+2\epsilon}}{2} & \frac{-\epsilon+1-\sqrt{-3\epsilon^2+2\epsilon}}{2} & \epsilon & 0 \end{array}\right), \tag{9.1.14}$$

$$B(1) = \left(\begin{array}{cc|cc} 0 & \frac{-\epsilon+\sqrt{-3\epsilon^2+2\epsilon}}{2} & \epsilon & \frac{-\epsilon+1-\sqrt{-3\epsilon^2+2\epsilon}}{2} \\ \frac{-\epsilon+1-\sqrt{-3\epsilon^2+2\epsilon}}{2} & \epsilon & \frac{-\epsilon+\sqrt{-3\epsilon^2+2\epsilon}}{2} & 0 \end{array}\right) \quad \textit{und} \tag{9.1.15}$$

$$B(2) = \left(\begin{array}{cc|cc} 0 & \frac{-\epsilon+1-\sqrt{-3\epsilon^2+2\epsilon}}{2} & \frac{-\epsilon+\sqrt{-3\epsilon^2+2\epsilon}}{2} & \epsilon \\ \epsilon & \frac{-\epsilon+\sqrt{-3\epsilon^2+2\epsilon}}{2} & \frac{-\epsilon+1-\sqrt{-3\epsilon^2+2\epsilon}}{2} & 0 \end{array}\right). \tag{9.1.16}$$

Beweis. Im Fall $\epsilon = 1/6$ konvergiert die IPF-Folge $(A(t))$ gemäß Satz 9.1.1. Da die Menge $\mathcal{E}^{\ll A}$ in diesem Fall nur ein einziges Element S enthält, gilt $S = \lim_{t\to\infty} A(t)$. Außerdem gilt $S = B(0) = B(1) = B(2)$.

Im Fall $\epsilon \in (0; 1/6)$ nehmen wir ohne Beschränkung der Allgemeinheit an, dass die Ausgangstafel A normiert ist. Der Beweis erfolgt dann in sechs Schritten. In Schritt I definieren wir eine Funktion $\Lambda : (0; \infty)^{2\times2\times2} \to (0; \infty)$ und zeigen, dass die resultierende Folge $(\Lambda(A(t)))_{t\in\mathbb{N}}$ monoton steigend ist. Daraus folgern wir in Schritt II die Aussagen $\lim_{t\to\infty} a_{111}(t) = 0$ und $\lim_{t\to\infty} a_{222}(t) = 0$. In Schritt III definieren wir eine Funktion $f : [0; 1/2 - \epsilon] \to [0; 1/2 - \epsilon]$ und beweisen, dass für jeden Häufungspunkt b der Folge $(a_{211}(6s))$ auch der Punkt $f^{(6)}(b)$ ein Häufungspunkt derselben Folge ist, wobei $f^{(6)}$ die sechsfache Hintereinanderausführung der Funktion f bezeichnet. Schritt IV zeigt, dass die Folge $(f^{(6h)}(b))$ gegen den eindeutigen Fixpunkt x^* der Funktion f konvergiert. Das Verhalten der Funktion f in der Nähe von x^* wird in Schritt V untersucht. Daraus folgern wir in Schritt VI die Konvergenz der Folge $(a_{211}(6s))$ und somit schließlich die Konvergenz der verbleibenden Einträge der Tafeln der IPF-Teilfolgen.

I. Die Funktion $\Lambda : (0; \infty)^{2\times2\times2} \to (0; \infty)$ sei definiert gemäß

$$\Lambda(S) := s_{111}^{3\epsilon-\frac{1}{2}} \cdot (s_{112}s_{121}s_{211})^{\frac{1}{2}-2\epsilon} \cdot (s_{122}s_{212}s_{221})^{\epsilon}. \tag{9.1.17}$$

Für alle Wahrscheinlichkeitstafeln $P \in (0;1)^{2\times 2\times 2}$ gilt folglich

$$\begin{aligned}
\frac{\Lambda(T_1(P))}{\Lambda(P)} &= \left(\frac{m_{11}}{p_{11+}}\right)^{3\epsilon-\frac{1}{2}} \cdot \left(\frac{m_{11}}{p_{11+}}\frac{m_{12}}{p_{12+}}\frac{m_{21}}{p_{21+}}\right)^{\frac{1}{2}-2\epsilon} \cdot \left(\frac{m_{12}}{p_{12+}}\frac{m_{21}}{p_{21+}}\frac{m_{22}}{p_{22+}}\right)^{\epsilon} \\
&= \left(\frac{m_{11}}{p_{11+}}\right)^{\epsilon} \cdot \left(\frac{m_{12}}{p_{12+}}\right)^{\frac{1}{2}-\epsilon} \cdot \left(\frac{m_{21}}{p_{21+}}\right)^{\frac{1}{2}-\epsilon} \cdot \left(\frac{m_{22}}{p_{22+}}\right)^{\epsilon} \\
&= \left(\frac{m_{11}}{p_{11+}}\right)^{m_{11}} \cdot \left(\frac{m_{12}}{p_{12+}}\right)^{m_{12}} \cdot \left(\frac{m_{21}}{p_{21+}}\right)^{m_{21}} \cdot \left(\frac{m_{22}}{p_{22+}}\right)^{m_{22}} \\
&= \exp(\mathrm{I}(M|(p_{ij+}))) \geq 1. \qquad (9.1.18)
\end{aligned}$$

Analog gilt

$$\frac{\Lambda(T_2(P))}{\Lambda(P)} = \exp(\mathrm{I}(M|(p_{+jk}))) \geq 1 \text{ und} \qquad (9.1.19)$$

$$\frac{\Lambda(T_3(P))}{\Lambda(P)} = \exp(\mathrm{I}(M|(p_{i+k}))) \geq 1. \qquad (9.1.20)$$

Somit ist die Folge $(\Lambda(A(t)))_{t\in\mathbb{N}}$ monoton steigend und es gilt

$$0 < \Lambda(A(1)) \leq \Lambda(A(2)) \leq \Lambda(A(3)) \leq \cdots. \qquad (9.1.21)$$

Außerdem gelten die Äquivalenzaussagen

$$\Lambda(T_1(P)) = \Lambda(P) \quad\Leftrightarrow\quad \mathrm{I}(M|(p_{ij+})) = 0 \quad\Leftrightarrow\quad M = (p_{ij+}), \qquad (9.1.22)$$

$$\Lambda(T_2(P)) = \Lambda(P) \quad\Leftrightarrow\quad \mathrm{I}(M|(p_{+jk})) = 0 \quad\Leftrightarrow\quad M = (p_{+jk}) \text{ und} \qquad (9.1.23)$$

$$\Lambda(T_3(P)) = \Lambda(P) \quad\Leftrightarrow\quad \mathrm{I}(M|(p_{i+k})) = 0 \quad\Leftrightarrow\quad M = (p_{i+k}). \qquad (9.1.24)$$

Daraus folgt

$$\Lambda(T_1(P)) = \Lambda(P) \quad\Leftrightarrow\quad T_1(P) = P, \qquad (9.1.25)$$

$$\Lambda(T_2(P)) = \Lambda(P) \quad\Leftrightarrow\quad T_2(P) = P \text{ und} \qquad (9.1.26)$$

$$\Lambda(T_3(P)) = \Lambda(P) \quad\Leftrightarrow\quad T_3(P) = P. \qquad (9.1.27)$$

II. Wir nehmen an, es existieren eine Konstante $\gamma > 0$ und eine IPF-Teilfolge $(A(t_\ell))$, sodass gilt

$$\forall \ell \in \mathbb{N}: \quad a_{111}(t_\ell) \geq \gamma. \qquad (9.1.28)$$

Des Weiteren nehmen wir an, dass ein Indexvektor $(i,j,k) \in \{1,2\}^3 \setminus \{(1,1,1),(2,2,2)\}$ existiert, für den eine Teilfolge von $(a_{ijk}(t_\ell))$ gegen 0 geht. Dann geht auch die entsprechende Teilfolge von $(\Lambda(A(t_\ell)))$ gegen 0. Dies steht jedoch im Widerspruch zur Ungleichungskette (9.1.21). Somit ist die zweite Annahme falsch und es existiert ein $\delta \in (0;\gamma]$, sodass für alle Indexvektoren $(i,j,k) \in \{1,2\}^3 \setminus \{(1,1,1),(2,2,2)\}$ gilt

$$\forall \ell \in \mathbb{N}: \quad a_{ijk}(t_\ell) \geq \delta. \qquad (9.1.29)$$

Folglich ist die IPF-Teilfolge $(A(t_\ell))$ in der kompakten Menge $[\delta; 1]^7 \times [0; 1]$ enthalten. Sei $(A(\tau_\ell))$ eine konvergente Teilfolge von $(A(t_\ell))$ und sei $B^* := \lim_{\ell\to\infty} A(\tau_\ell)$ ihre Limestafel. Falls die Folge (τ_ℓ) unendlich viele Vielfache von 3 enthält, ersetzen wir $(A(\tau_\ell))$ durch eine gemeinsame Teilfolge von $(A(\tau_\ell))$ und $(A(3s))$ und wählen $F := T_3 \circ T_2 \circ T_1$. Mit der Stetigkeit der Funktionen T_1, T_2, T_3 und Λ auf $[\delta; 1]^7 \times [0; 1]$ und der Monotonie der Folge $(\Lambda(A(t)))_{t\in\mathbb{N}}$ folgt dann

$$\begin{aligned}\Lambda(F(B^*)) &= \Lambda(F(\lim_{\ell\to\infty} A(\tau_\ell))) = \lim_{\ell\to\infty} \Lambda(F(A(\tau_\ell))) = \lim_{\ell\to\infty} \Lambda(A(\tau_\ell + 3)) \\ &= \lim_{\ell\to\infty} \Lambda(A(\tau_\ell)) = \Lambda(\lim_{\ell\to\infty} A(\tau_\ell)) = \Lambda(B^*). \end{aligned} \tag{9.1.30}$$

Also gilt

$$\Lambda(T_3 \circ T_2 \circ T_1(B^*)) = \Lambda(T_2 \circ T_1(B^*)) = \Lambda(T_1(B^*)) = \Lambda(B^*). \tag{9.1.31}$$

Daraus folgt mit Schritt I die Aussage $T_3 \circ T_2 \circ T_1(B^*) = T_2 \circ T_1(B^*) = T_1(B^*) = B^*$. Somit erfüllt die Tafel $B^* \in [\delta; 1]^7 \times [0; 1]$ die durch die Marginaltafeln $M^{(1)}$, $M^{(2)}$ und $M^{(3)}$ vorgegebenen Randsummen und ist folglich in der Menge $\mathcal{E}$ enthalten. Falls die Folge (τ_ℓ) dagegen nur endlich viele Vielfache von 3 enthält, enthält (τ_ℓ) unendlich viele Elemente der Folge $(3s + 1)$ oder unendlich viele Elemente der Folge $(3s + 2)$. In beiden Fällen erfolgt der Beweis der Aussage $B^* \in \mathcal{E}$ analog zum oben diskutierten Fall. Die Aussage $B^* \in \mathcal{E} = \mathcal{E}^{\ll A}$ steht jedoch im Widerspruch zur Voraussetzung, da die Menge $\mathcal{E} = \mathcal{E}^{\ll A}$ nach Voraussetzung leer ist. Also ist auch die erste Annahme falsch und es gilt

$$\lim_{t\to\infty} a_{111}(t) = 0. \tag{9.1.32}$$

Durch Wiederholung der Schritte I und II bei Vertauschung der Indizes 1 und 2 erhalten wir außerdem die Aussage

$$\lim_{t\to\infty} a_{222}(t) = 0. \tag{9.1.33}$$

III. Aus den soeben gezeigten Grenzwerten der Folgen $(a_{111}(t))$ und $(a_{222}(t))$ folgt aufgrund der durch die Marginaltafeln $M^{(1)}$, $M^{(2)}$ und $M^{(3)}$ vorgegebenen Randsummen die Grenzwertaussage

$$\begin{aligned}\lim_{s\to\infty} a_{212}(6s) &= \lim_{s\to\infty} a_{112}(6s + 1) = \lim_{s\to\infty} a_{122}(6s + 2)) = \lim_{s\to\infty} a_{121}(6s + 3)) \\ &= \lim_{s\to\infty} a_{221}(6s + 4)) = \lim_{s\to\infty} a_{211}(6s + 5)) = \epsilon. \end{aligned} \tag{9.1.34}$$

Als nächstes untersuchen wir die Folge $(a_{211}(6s))$. Dazu definieren wir die Funktion $g : [0; 1/2 - \epsilon] \times (0; \infty) \to [0; 1/2 - \epsilon]$ gemäß

$$g(x, t) := \left(\tfrac{1}{2} - \epsilon\right) \frac{t}{x + t}. \tag{9.1.35}$$

Für alle $s \in \mathbb{N}_0$ gilt folglich

$$a_{212}(6s + 1) = g(a_{211}(6s), a_{212}(6s)), \tag{9.1.36}$$

$$a_{112}(6s + 2) = g(a_{212}(6s + 1), a_{112}(6s + 1)), \tag{9.1.37}$$

$$a_{122}(6s+3) = g(a_{112}(6s+2), a_{122}(6s+2)), \tag{9.1.38}$$

$$a_{121}(6s+4) = g(a_{122}(6s+3), a_{121}(6s+3)), \tag{9.1.39}$$

$$a_{221}(6s+5) = g(a_{121}(6s+4), a_{221}(6s+4)) \text{ und} \tag{9.1.40}$$

$$a_{211}(6s+6) = g(a_{221}(6s+5), a_{211}(6s+5)). \tag{9.1.41}$$

Da die Folge $(a_{211}(6s))$ in der kompakten Menge $[0; 1/2 - \epsilon]$ enthalten ist, existiert eine konvergente Teilfolge $(a_{211}(6s_\ell))$ mit dem Grenzwert

$$b := \lim_{\ell \to \infty} a_{211}(6s_\ell) \in \left[0; \tfrac{1}{2} - \epsilon\right]. \tag{9.1.42}$$

Mit der Gleichung (9.1.36), der Stetigkeit der Funktion g sowie der Grenzwertaussage (9.1.34) folgt daraus

$$\lim_{\ell \to \infty} a_{212}(6s_\ell + 1) = \lim_{\ell \to \infty} g(a_{211}(6s_\ell), a_{212}(6s_\ell)) = g(b, \epsilon). \tag{9.1.43}$$

Wir definieren daher die Funktion $f : [0; 1/2 - \epsilon] \to [0; 1/2 - \epsilon]$ gemäß

$$f(x) := g(x, \epsilon) = \frac{\epsilon(1 - 2\epsilon)}{2(x + \epsilon)}. \tag{9.1.44}$$

Die h-fache Anwendung der Funktion f bezeichnen wir mit $f^{(h)}$. Auf diese Weise erhalten wir die Aussagen

$$\lim_{\ell \to \infty} a_{212}(6s_\ell + 1) = f(b), \qquad \lim_{\ell \to \infty} a_{112}(6s_\ell + 2) = f^{(2)}(b), \tag{9.1.45}$$

$$\lim_{\ell \to \infty} a_{122}(6s_\ell + 3) = f^{(3)}(b), \qquad \lim_{\ell \to \infty} a_{121}(6s_\ell + 4) = f^{(4)}(b), \tag{9.1.46}$$

$$\lim_{\ell \to \infty} a_{221}(6s_\ell + 5) = f^{(5)}(b) \text{ und} \qquad \lim_{\ell \to \infty} a_{211}(6s_\ell + 6) = f^{(6)}(b). \tag{9.1.47}$$

Folglich ist $f^{(6)}(b)$ ebenfalls ein Häufungspunkt der Folge $(a_{211}(6s))$. Durch Wiederholung dieser Argumentation erhalten wir schließlich die Aussage, dass alle Werte $f^{(6h)}(b)$, $h \in \mathbb{N}_0$, Häufungspunkte der Folge $(a_{211}(6s))$ sind.

IV. Als nächstes analysieren wir die Folge $(f^{(6h)}(b))$. Die Funktion $f : [0; 1/2 - \epsilon] \to [0; 1/2 - \epsilon]$ ist stetig und streng monoton fallend. Außerdem gilt $f(0) = 1/2 - \epsilon > 0$ sowie $f(1/2 - \epsilon) = \epsilon(1 - 2\epsilon) < 1/2 - \epsilon$. Somit hat die Funktion f genau einen Fixpunkt x^*. Für diesen Fixpunkt gilt

$$x^* = \frac{-\epsilon + \sqrt{-3\epsilon^2 + 2\epsilon}}{2} \in \left(0; \tfrac{1}{2} - \epsilon\right). \tag{9.1.48}$$

Außerdem gilt

$$\forall x \in \left[0; \tfrac{1}{2} - \epsilon\right]: \quad f^{(2)}(x) = \left(\tfrac{1}{2} - \epsilon\right) \frac{x + \epsilon}{x + \frac{1}{2}} \text{ und} \tag{9.1.49}$$

$$f^{(2)\prime}(x) = \frac{\left(\frac{1}{2} - \epsilon\right)^2}{\left(x + \frac{1}{2}\right)^2} \leq (1 - 2\epsilon)^2 < 1. \tag{9.1.50}$$

Folglich ist die Funktion $f^{(2)} : [0; 1/2-\epsilon] \to [0; 1/2-\epsilon]$ eine Kontraktion mit x^* als einzigem Fixpunkt. Die Folge $(f^{(6h)}(b))$ konvergiert somit nach dem Banachschen Fixpunktsatz gegen x^*. Da alle Folgenglieder $f^{(6h)}(b)$ Häufungspunkte der Folge $(a_{211}(6s))$ sind, ist x^* ebenfalls ein Häufungspunkt der Folge $(a_{211}(6s))$.

V. Wir untersuchen nun die Funktionen f und $g(\cdot, t)$ in der Nähe des Punktes x^*. Für die Ableitung von f im Punkt x^* gilt

$$\left|f'(x^*)\right| = \frac{\epsilon(1-2\epsilon)}{2(x^*+\epsilon)^2} = \frac{1-2\epsilon}{1-\epsilon+\sqrt{2\epsilon-3\epsilon^2}} < 1. \tag{9.1.51}$$

Da f stetig differenzierbar und streng monoton fallend ist, existiert ein $\Delta > 0$, sodass gilt

$$\forall \delta \in (0; \Delta] : \quad x^* - \delta < f(x^* + \delta) < f(x^* - \delta) < x^* + \delta. \tag{9.1.52}$$

Da außerdem $\lim_{t \to \epsilon} g(x, t) = g(x, \epsilon) = f(x)$ für alle $x \in [0; 1/2 - \epsilon]$ gilt, folgt

$$\begin{aligned} &\forall \delta \in (0; \Delta] \quad \exists \gamma > 0 \quad \forall t \in (\epsilon - \gamma; \epsilon + \gamma) : \\ &\qquad x^* - \delta < g(x^* + \delta, t) < g(x^* - \delta, t) < x^* + \delta. \end{aligned} \tag{9.1.53}$$

Für alle $t \in (0; \infty)$ ist die Funktion $g(\cdot, t)$ streng monoton fallend. Somit gilt

$$\begin{aligned} &\forall \delta \in (0; \Delta] \quad \exists \gamma > 0 \quad \forall x \in (x^* - \delta; x^* + \delta) \quad \forall t \in (\epsilon - \gamma; \epsilon + \gamma) : \\ &\qquad g(x, t) \in (x^* - \delta; x^* + \delta). \end{aligned} \tag{9.1.54}$$

VI. Die soeben gezeigte Aussage über das Verhalten der Funktion g in der Nähe des Punktes (x^*, ϵ) ermöglicht uns den Beweis der Konvergenz der Folge $(a_{211}(6s))$. Sei dazu $(a_{211}(6s^*_\ell))$ eine Teilfolge, die gegen x^* konvergiert. Weiter sei $\Theta > 0$ beliebig. Dann wählen wir $\delta' := \min\{\Delta, \Theta\}$. Daraus folgt

$$\begin{aligned} &\exists \gamma > 0 \quad \forall x \in (x^* - \delta'; x^* + \delta') \quad \forall t \in (\epsilon - \gamma; \epsilon + \gamma) : \\ &\qquad g(x, t) \in (x^* - \delta'; x^* + \delta'). \end{aligned} \tag{9.1.55}$$

Anschließend wählen wir ein $s' \in \mathbb{N}_0$, sodass gilt

$$\begin{aligned} \forall s \geq s' : \quad & a_{212}(6s), a_{112}(6s+1), a_{122}(6s+2)), \\ & a_{121}(6s+3)), a_{221}(6s+4)), a_{211}(6s+5)) \in (\epsilon - \gamma; \epsilon + \gamma). \end{aligned} \tag{9.1.56}$$

Außerdem wählen wir ein $\ell' \in \mathbb{N}_0$, sodass $s^*_{\ell'} \geq s'$ und $a_{211}(6s^*_{\ell'}) \in (x^* - \delta'; x^* + \delta')$ gelten. Des Weiteren wählen wir $s'' := s_{\ell'} \geq s'$. Dann gilt

$$a_{212}(6s''+1) = g(a_{211}(6s''), a_{212}(6s'')) \in (x^* - \delta'; x^* + \delta'), \tag{9.1.57}$$

$$a_{112}(6s''+2) = g(a_{212}(6s''+1), a_{112}(6s''+1)) \in (x^* - \delta'; x^* + \delta'), \tag{9.1.58}$$

$$a_{122}(6s''+3) = g(a_{112}(6s''+2), a_{122}(6s''+2)) \in (x^* - \delta'; x^* + \delta'), \tag{9.1.59}$$

$$a_{121}(6s''+4) = g(a_{122}(6s''+3), a_{121}(6s''+3)) \in (x^* - \delta'; x^* + \delta'), \tag{9.1.60}$$

$$a_{221}(6s''+5) = g(a_{121}(6s''+4), a_{221}(6s''+4)) \in (x^* - \delta'; x^* + \delta') \text{ und} \tag{9.1.61}$$

$$a_{211}(6s''+6) = g(a_{221}(6s''+5), a_{211}(6s''+5)) \in (x^* - \delta'; x^* + \delta'). \tag{9.1.62}$$

Daraus folgt induktiv die Aussage

$$\forall s \geq s'' : \quad a_{211}(6s) \in (x^* - \delta'; x^* + \delta') \subseteq (x^* - \Theta; x^* + \Theta). \tag{9.1.63}$$

Zusammenfassend gilt somit

$$\forall \Theta > 0 \quad \exists s'' \in \mathbb{N}_0 \quad \forall s \geq s'' : \quad a_{211}(6s) \in (x^* - \Theta; x^* + \Theta). \tag{9.1.64}$$

Daraus folgt

$$\lim_{s\to\infty} a_{211}(6s) = x^* = \frac{-\epsilon + \sqrt{-3\epsilon^2 + 2\epsilon}}{2}. \tag{9.1.65}$$

Mit den Gleichungen (9.1.36) bis (9.1.40) und der Gleichungskette (9.1.34) folgt daraus

$$\lim_{s\to\infty} a_{212}(6s+1) = \lim_{s\to\infty} g(a_{211}(6s), a_{212}(6s)) = g(x^*, \epsilon) = x^*, \tag{9.1.66}$$

$$\lim_{s\to\infty} a_{112}(6s+2) = \lim_{s\to\infty} g(a_{212}(6s+1), a_{112}(6s+1)) = g(x^*, \epsilon) = x^*, \tag{9.1.67}$$

$$\lim_{s\to\infty} a_{122}(6s+3) = \lim_{s\to\infty} g(a_{112}(6s+2), a_{122}(6s+2)) = g(x^*, \epsilon) = x^*, \tag{9.1.68}$$

$$\lim_{s\to\infty} a_{121}(6s+4) = \lim_{s\to\infty} g(a_{122}(6s+3), a_{121}(6s+3)) = g(x^*, \epsilon) = x^* \text{ und} \tag{9.1.69}$$

$$\lim_{s\to\infty} a_{221}(6s+5) = \lim_{s\to\infty} g(a_{121}(6s+4), a_{221}(6s+4)) = g(x^*, \epsilon) = x^*. \tag{9.1.70}$$

Aufgrund der durch die Marginaltafeln $M^{(1)}$, $M^{(2)}$ und $M^{(3)}$ vorgegebenen Randsummen folgen außerdem die Aussagen

$$\lim_{s\to\infty} a_{221}(6s) = \tfrac{1}{2} - \epsilon - x^*, \qquad \lim_{s\to\infty} a_{211}(6s+1) = \tfrac{1}{2} - \epsilon - x^*, \tag{9.1.71}$$

$$\lim_{s\to\infty} a_{112}(6s+2) = \tfrac{1}{2} - \epsilon - x^*, \qquad \lim_{s\to\infty} a_{112}(6s+3) = \tfrac{1}{2} - \epsilon - x^*, \tag{9.1.72}$$

$$\lim_{s\to\infty} a_{221}(6s+4) = \tfrac{1}{2} - \epsilon - x^* \text{ und} \qquad \lim_{s\to\infty} a_{121}(6s+5) = \tfrac{1}{2} - \epsilon - x^*. \tag{9.1.73}$$

Durch Wiederholung der Schritte III bis VI bei Vertauschung der Indizes 1 und 2 erhalten wir schließlich die Konvergenzaussagen für die verbleibenden Einträge der Tafeln der IPF-Teilfolgen $(A(6s))$, $(A(6s+1))$, $(A(6s+2))$, $(A(6s+3))$, $(A(6s+4))$ und $(A(6s+5))$. □

9.2. Mehrdimensionale Anpassungsprobleme mit dünn besetzter $2 \times 2 \times 2$-Ausgangstafel und eindimensionalen Marginaltafeln

In diesem Abschnitt werden mehrdimensionale Anpassungsprobleme untersucht, die aus einer $2\times2\times2$-Ausgangstafel und eindimensionalen Marginaltafeln bestehen. Diese Anpassungsprobleme sind nur interessant, wenn drei Marginaltafeln gegeben sind. Sind dagegen beispielsweise nur die zu den Rändern $D_1 := \{1\}$, $D_2 := \{2\}$ gehörenden Marginaltafeln $M^{(1)}$ und $M^{(2)}$ gegeben, so bleiben alle Verhältnisse $a_{ij1}(t)/a_{ij2}(t)$, $i, j \in \{1; 2\}$, über alle Schritte t konstant. Das mehrdimensionale Anpassungsproblem kann dann zu einem biproportionalen Anpassungsproblem mit der 2×2-Ausgangsmatrix (a_{ij+}) reduziert werden.

Des Weiteren sind die oben genannten mehrdimensionalen Anpassungsprobleme nur dann interessant, wenn die Ausgangstafel Nulleinträge enthält. Ansonsten ist die Konvergenz des IPF-Verfahrens durch Satz 7.3.2 (i) (siehe S. 78) sichergestellt, da die Menge $\mathcal{E}^{\ll A} = \mathcal{E}$ die Tafel $S \in [0;1]^{2\times2\times2}$ enthält, die definiert ist gemäß

$$s_{ijk} := m_i^{(1)} \cdot m_j^{(2)} \cdot m_k^{(3)} \text{ für alle Einträge } (i,j,k). \tag{9.2.1}$$

Alle $2\times2\times2$-Ausgangstafeln mit genau zwei positiven Einträgen können bis auf Symmetrie zum Fall

$$A = \left(\begin{array}{cc|cc} a_{111} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{222} \end{array}\right) \tag{9.2.2}$$

zusammengefasst werden. Diese Ausgangstafeln resultieren gemäß Definition des mehrdimensionalen IPF-Verfahrens in konstanten IPF-Teilfolgen $(A(3s))$, $(A(3s+1))$ und $(A(3s+2))$ und sind daher ebenfalls wenig interessant.

Bei $2\times2\times2$-Ausgangstafeln mit genau drei positiven Einträgen müssen zwei Fälle unterschieden werden. Die Fälle, in denen zwei positive Einträge in zwei Rändern übereinstimmen, können bis auf Symmetrie zum Fall

$$A = \left(\begin{array}{cc|cc} a_{111} & 0 & a_{112} & 0 \\ 0 & a_{221} & 0 & 0 \end{array}\right) \tag{9.2.3}$$

zusammengefasst werden. In diesem Fall sind die Zeilenskalierungen überflüssig, da Zeilenskalierung und anschließende Spaltenskalierung dasselbe Ergebnis liefern wie direkte Spaltenskalierung. Das nur aus Spaltenanpassungen und Lagenanpassungen bestehende IPF-Verfahren kann auf das zweidimensionale IPF-Verfahren zurückgeführt werden. Die Fälle, in denen alle Paare von positiven Einträgen in höchstens einem Rand übereinstimmen, können bis auf Symmetrie zum Fall

$$A = \left(\begin{array}{cc|cc} a_{111} & 0 & 0 & a_{122} \\ 0 & a_{221} & 0 & 0 \end{array}\right) \tag{9.2.4}$$

zusammengefasst werden.

Für den Rest dieses Abschnitts betrachten wir daher ausschließlich mehrdimensionale Anpassungsprobleme $(A, M^{(1)}, M^{(2)}, M^{(3)})$ bestehend aus einer Ausgangstafel $A \in [0;\infty)^{2\times2\times2}$ gemäß der obigen Gleichung (9.2.4) sowie eindimensionalen Marginaltafeln $M^{(1)}, M^{(2)}, M^{(3)} \in (0;\infty)^2$. Um die Konvergenz der zugehörigen IPF-Teilfolgen zu zeigen, erweist sich die Betrachtung der reellen Möbiustransformationen als hilfreich. Diese Möbiustransformationen sind auf $\mathbb{R} \cup \{\infty\}$ definiert, der Einpunktkompaktifizierung der reellen Zahlen. Diese Einpunktkompaktifizierung ist vergleichbar mit der Erweiterung der komplexen Zahlen $\mathbb{C}$ zur Riemannschen Zahlenkugel $\mathbb{C} \cup \{\infty\}$. Im Gegensatz zur erweiterten reellen Achse $\mathbb{R} \cup \{-\infty; +\infty\}$ werden die reellen Zahlen $\mathbb{R}$ bei der Einpunktkompaktifizierung nur um einen einzigen Punkt im Unendlichen erweitert.

Definition 9.2.1 (Reelle Möbiustransformationen, vgl. Beardon 1991, Abschnitt 1.2)**.** Gegeben seien die reellen Zahlen $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ mit $\alpha\delta - \beta\gamma \neq 0$. Im Fall $\gamma = 0$ ist die zugehörige *reelle Möbiustransformation* $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ definiert gemäß

$$f(x) := \begin{cases} \frac{\alpha x+\beta}{\gamma x+\delta} = \frac{\alpha}{\delta}x + \frac{\beta}{\delta} \in \mathbb{R} & \text{falls } x \in \mathbb{R}, \\ \infty & \text{falls } x = \infty. \end{cases} \tag{9.2.5}$$

Im Fall $\gamma \neq 0$ ist die zugehörige *reelle Möbiustransformation* $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ definiert gemäß

$$f(x) := \begin{cases} \frac{\alpha x+\beta}{\gamma x+\delta} \in \mathbb{R} \setminus \left\{\frac{\alpha}{\gamma}\right\} & \text{falls } x \in \mathbb{R} \setminus \left\{-\frac{\delta}{\gamma}\right\}, \\ \infty & \text{falls } x = -\frac{\delta}{\gamma}, \\ \frac{\alpha}{\gamma} & \text{falls } x = \infty. \end{cases} \tag{9.2.6}$$

Die Werte α, β, γ und δ heißen die *Parameter der reellen Möbiustransformation* f.

Im Fall $\gamma = 0$ stellt die Bedingung $\alpha\delta - \beta\gamma \neq 0$ sicher, dass der Nenner $\gamma x + \delta = \delta$ ungleich 0 ist. Außerdem stellt diese Bedingung in beiden Fällen sicher, dass die Funktion $x \mapsto (\alpha x + \beta)/(\gamma x + \delta)$ nicht konstant ist.

Satz 9.2.2 (Bijektivität, Stetigkeit und Differenzierbarkeit reeller Möbiustransformationen)**.** *Seien* $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ *reelle Zahlen mit* $\alpha\delta - \beta\gamma \neq 0$. *Dann ist die zugehörige reelle Möbiustransformation* $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ *bijektiv und stetig. Im Fall* $\gamma = 0$ *ist* f *außerdem differenzierbar auf* $\mathbb{R}$ *mit der Ableitung* $f'(x) = \alpha/\delta \neq 0$. *Im Fall* $\gamma \neq 0$ *ist* f *differenzierbar auf* $\mathbb{R} \setminus \{-\delta/\gamma\}$ *mit der Ableitung* $f'(x) = (\alpha\delta - \beta\gamma)/(\gamma x + \delta)^2 \neq 0$.

Beweis. Im Fall $\gamma = 0$ gilt $f(x) = (\alpha/\delta)x + (\beta/\delta)$ für alle $x \in \mathbb{R}$, wobei $\alpha/\delta \neq 0$ gilt. Folglich bildet f die reellen Zahlen $\mathbb{R}$ bijektiv und stetig auf sich selbst ab. Für alle reellen Folgen (x_s) mit $\lim_{s\to\infty} x_s = \infty$ gilt außerdem

$$\lim_{s\to\infty} f(x_s) = \lim_{s\to\infty} \frac{\alpha}{\delta} x_s + \frac{\beta}{\delta} = \infty = f(\infty). \tag{9.2.7}$$

Somit ist f auch im Punkt $x = \infty$ stetig.

Im Fall $\gamma \neq 0$ gilt für alle $x \in \mathbb{R} \setminus \{-\delta/\gamma\}$ und für alle $y \in \mathbb{R} \setminus \{\alpha/\gamma\}$ die Äquivalenzaussage

$$\begin{aligned} \frac{\alpha x + \beta}{\gamma x + \delta} = y \quad &\Leftrightarrow \quad \alpha x + \beta = \gamma x y + \delta y \\ &\Leftrightarrow \quad (\alpha - \gamma y)x = \delta y - \beta \\ &\Leftrightarrow \quad x = \frac{\delta y - \beta}{-\gamma y + \alpha}. \end{aligned} \tag{9.2.8}$$

Somit ist $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ bijektiv. Als Komposition stetiger Funktionen ist f außerdem stetig auf $\mathbb{R} \setminus \{-\delta/\gamma\}$. Für alle Folgen (x_s) in $\mathbb{R} \setminus \{-\delta/\gamma\}$ mit $\lim_{s\to\infty} x_s = -\delta/\gamma$ gilt außerdem

$$\lim_{s\to\infty} \alpha x_s + \beta = -\frac{\alpha\delta}{\gamma} + \beta = -\frac{\alpha\delta - \beta\gamma}{\gamma} \neq 0, \tag{9.2.9}$$

$$\lim_{s\to\infty} \gamma x_s + \delta = -\frac{\gamma\delta}{\gamma} + \delta = 0 \text{ und somit} \tag{9.2.10}$$

$$\lim_{s\to\infty} f(x_s) = \lim_{s\to\infty} \frac{\alpha x_s + \beta}{\gamma x_s + \delta} = \infty = f\left(\frac{-\delta}{\gamma}\right). \tag{9.2.11}$$

Für alle Folgen (x_s) in $\mathbb{R} \setminus \{-\delta/\gamma\}$ mit $\lim_{s\to\infty} x_s = \infty$ gilt

$$\lim_{s\to\infty} f(x_s) = \lim_{s\to\infty} \frac{\alpha x_s + \beta}{\gamma x_s + \delta} = \frac{\alpha}{\gamma} = f(\infty). \tag{9.2.12}$$

Somit ist f auch in den Punkten $x = -\delta/\gamma$ und $x = \infty$ stetig. □

Der folgende Satz zeigt, dass jede Komposition von reellen Möbiustransformationen wieder eine reelle Möbiustransformation ist.

Satz 9.2.3 (Komposition reeller Möbiustransformationen, vgl. Needham 1997, Abschnitt 3.V.3)**.** *Seien $\alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R}$, $i = 1, 2$, reelle Zahlen mit $\alpha_i\delta_i - \beta_i\gamma_i \neq 0$ und seien $f_i : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ die zugehörigen reellen Möbiustransformationen. Außerdem sei $f := f_2 \circ f_1$ die Komposition von f_1 und f_2. Dann ist f die reelle Möbiustransformation mit den Parametern $\alpha := \alpha_1\alpha_2 + \beta_2\gamma_1$, $\beta := \alpha_2\beta_1 + \beta_2\delta_1$, $\gamma := \alpha_1\gamma_2 + \gamma_1\delta_2$ und $\delta := \beta_1\gamma_2 + \delta_1\delta_2$.*

Beweis. Seien $\alpha_i, \beta_i, \gamma_i, \delta_i \in \mathbb{R}$, $i = 1, 2$, reelle Zahlen mit $\alpha_i\delta_i - \beta_i\gamma_i \neq 0$ und seien $f_i : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ die zugehörigen reellen Möbiustransformationen. Außerdem sei $f := f_2 \circ f_1$ die Komposition von f_1 und f_2. Des Weiteren sei $g : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ die reelle Möbiustransformation mit den Parametern $\alpha := \alpha_1\alpha_2 + \beta_2\gamma_1$, $\beta := \alpha_2\beta_1 + \beta_2\delta_1$, $\gamma := \alpha_1\gamma_2 + \gamma_1\delta_2$ und $\delta := \beta_1\gamma_2 + \delta_1\delta_2$. Die reelle Möbiustransformation g ist wohldefiniert, da

$$\begin{aligned}\alpha\delta - \beta\gamma &= (\alpha_1\alpha_2 + \beta_2\gamma_1)(\beta_1\gamma_2 + \delta_1\delta_2) - (\alpha_2\beta_1 + \beta_2\delta_1)(\alpha_1\gamma_2 + \gamma_1\delta_2) \\ &= \alpha_1\alpha_2\delta_1\delta_2 + \beta_1\beta_2\gamma_1\gamma_2 - \alpha_2\beta_1\gamma_1\delta_2 - \alpha_1\beta_2\gamma_2\delta_1 \\ &= (\alpha_1\delta_1 - \beta_1\gamma_1)(\alpha_2\delta_2 - \beta_2\gamma_2) \neq 0 \text{ gilt.} \end{aligned} \tag{9.2.13}$$

Im Folgenden zeigen wir, dass die Funktionen f und g in fast allen Punkten $x \in \mathbb{R} \cup \{\infty\}$ übereinstimmen. Dazu wählen wir zunächst $M := \mathbb{R}$. Falls $\gamma_1 \neq 0$ gilt, entfernen wir den Punkt $-\delta_1/\gamma_1 \in \mathbb{R}$ aus der Menge M. Falls $\gamma_2 \neq 0$ gilt, entfernen wir den Punkt $f_1^{-1}(-\delta_2/\gamma_2) \in \mathbb{R} \cup \{\infty\}$ aus der Menge M. Dann gilt für alle $x \in M$ die Aussage

$$\begin{aligned} f(x) = f_2(f_1(x)) &= \frac{\alpha_2 \frac{\alpha_1 x + \beta_1}{\gamma_1 x + \delta_1} + \beta_2}{\gamma_2 \frac{\alpha_1 x + \beta_1}{\gamma_1 x + \delta_1} + \delta_2} = \frac{\alpha_1\alpha_2 x + \alpha_2\beta_1 + \beta_2\gamma_1 x + \beta_2\delta_1}{\alpha_1\gamma_2 x + \beta_1\gamma_2 + \gamma_1\delta_2 x + \delta_1\delta_2} \\ &= \frac{(\alpha_1\alpha_2 + \beta_2\gamma_1)x + (\alpha_2\beta_1 + \beta_2\delta_1)}{(\alpha_1\gamma_2 + \gamma_1\delta_2)x + (\beta_1\gamma_2 + \delta_1\delta_2)} = \frac{\alpha x + \beta}{\gamma x + \delta} = g(x). \end{aligned} \tag{9.2.14}$$

Also stimmen die Funktionen f und g in fast allen Punkten $x \in \mathbb{R} \cup \{\infty\}$ überein. Da beide Funktionen stetig sind, folgt $f = g$. □

Für die Untersuchung des IPF-Verfahrens im Fall mehrdimensionaler Anpassungsprobleme mit einer Ausgangstafel gemäß Gleichung (9.2.4) und drei eindimensionalen Marginaltafeln erweisen sich insbesondere diejenigen reellen Möbiustransformationen als hilfreich, deren Parameter neben $\alpha\delta - \beta\gamma \neq 0$ auch die Bedingungen $\gamma \neq 0$ und $(\alpha - \delta)^2 + 4\beta\gamma > 0$ erfüllen. Der folgende Satz zeigt, dass jede Möbiustransformation dieser Art genau zwei Fixpunkte besitzt.

Satz 9.2.4 (Reelle Möbiustransformationen mit zwei Fixpunkten, vgl. Beardon 1991, Abschnitt 1.2)**.** *Seien $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ reelle Zahlen mit $\alpha\delta - \beta\gamma \neq 0$, $\gamma \neq 0$ und $(\alpha - \delta)^2 + 4\beta\gamma > 0$. Außerdem sei $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ die zugehörige reelle Möbiustransformation. Dann gilt:*

(i) f hat genau zwei Fixpunkte $x_1, x_2 \in \mathbb{R}$, $x_1 \neq x_2$, mit

$$x_{1/2} = \frac{(\alpha - \delta) \pm \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma} \qquad \textit{und} \qquad f'(x_1) \cdot f'(x_2) = 1. \tag{9.2.15}$$

Außerdem gilt genau eine der folgenden drei Aussagen:

(ii) $f'(x_1) = f'(x_2) = -1$ *und* $f \circ f = \mathrm{id}$.

(iii) $|f'(x_1)| < 1$ *und* $\forall x \in \mathbb{R} \cup \{\infty\} \setminus \{x_2\} : \lim_{s\to\infty} f^{(s)}(x) = x_1$.

(iv) $|f'(x_2)| < 1$ *und* $\forall x \in \mathbb{R} \cup \{\infty\} \setminus \{x_1\} : \lim_{s\to\infty} f^{(s)}(x) = x_2$.

Beweis. Da $\gamma \neq 0$ gilt, ist $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ definiert gemäß Gleichung (9.2.6). Folglich gilt

$$\begin{aligned} f(x) = x \quad &\Leftrightarrow \quad x \in \mathbb{R} \setminus \left\{-\frac{\delta}{\gamma}\right\} \text{ und } \frac{\alpha x + \beta}{\gamma x + \delta} = x \\ &\Leftrightarrow \quad x \in \mathbb{R} \setminus \left\{-\frac{\delta}{\gamma}\right\} \text{ und } \alpha x + \beta = \gamma x^2 + \delta x \\ &\Leftrightarrow \quad x \in \mathbb{R} \setminus \left\{-\frac{\delta}{\gamma}\right\} \text{ und } 0 = \gamma x^2 + (\delta - \alpha)x - \beta \\ &\Leftrightarrow \quad x \in \mathbb{R} \setminus \left\{-\frac{\delta}{\gamma}\right\} \text{ und } x = \frac{(\alpha - \delta) \pm \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma}. \end{aligned} \tag{9.2.16}$$

Außerdem gilt

$$\begin{aligned} \frac{(\alpha - \delta) \pm \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma} = -\frac{\delta}{\gamma} \quad &\Leftrightarrow \quad (\alpha - \delta) \pm \sqrt{(\alpha - \delta)^2 + 4\beta\gamma} = -2\delta \\ &\Leftrightarrow \quad \alpha + \delta = \mp\sqrt{(\alpha - \delta)^2 + 4\beta\gamma} \\ &\Leftrightarrow \quad (\alpha + \delta)^2 = (\alpha - \delta)^2 + 4\beta\gamma \\ &\Leftrightarrow \quad 4\alpha\delta = 4\beta\gamma \\ &\Leftrightarrow \quad \alpha\delta - \beta\gamma = 0. \end{aligned} \tag{9.2.17}$$

Da nach Voraussetzung $\alpha\delta - \beta\gamma \neq 0$ gilt, folgt

$$f(x) = x \quad \Leftrightarrow \quad x = \frac{(\alpha - \delta) \pm \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma}. \tag{9.2.18}$$

Nach Satz 9.2.2 gilt außerdem

$$\begin{aligned} f'(x_1) \cdot f'(x_2) &= \frac{\alpha\delta - \beta\gamma}{\left(\gamma \frac{(\alpha-\delta)-\sqrt{(\alpha-\delta)^2+4\beta\gamma}}{2\gamma} + \delta\right)^2} \cdot \frac{\alpha\delta - \beta\gamma}{\left(\gamma \frac{(\alpha-\delta)+\sqrt{(\alpha-\delta)^2+4\beta\gamma}}{2\gamma} + \delta\right)^2} \\ &= \frac{4\alpha\delta - 4\beta\gamma}{\left((\alpha + \delta) - \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}\right)^2} \cdot \frac{4\alpha\delta - 4\beta\gamma}{\left((\alpha + \delta) + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}\right)^2} \\ &= \frac{(4\alpha\delta - 4\beta\gamma)^2}{((\alpha + \delta)^2 - (\alpha - \delta)^2 - 4\beta\gamma)^2} = \frac{(4\alpha\delta - 4\beta\gamma)^2}{(4\alpha\delta - 4\beta\gamma)^2} = 1. \end{aligned} \tag{9.2.19}$$

Damit ist Aussage (i) gezeigt.

Wir wählen nun $g : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ gemäß

$$g(x) := \begin{cases} \frac{x-x_1}{x-x_2} \in \mathbb{R} \setminus \{1\} & \text{falls } x \in \mathbb{R} \setminus \{x_2\}, \\ \infty & \text{falls } x = x_2, \\ 1 & \text{falls } x = \infty. \end{cases} \tag{9.2.20}$$

Somit ist g eine reelle Möbiustransformation. Für die Umkehrabbildung $g^{-1} : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ gilt

$$g^{-1}(x) := \begin{cases} \frac{x_2 x - x_1}{x-1} \in \mathbb{R} \setminus \{x_2\} & \text{falls } x \in \mathbb{R} \setminus \{1\}, \\ \infty & \text{falls } x = 1, \\ x_2 & \text{falls } x = \infty. \end{cases} \tag{9.2.21}$$

Somit ist die Umkehrabbildung g^{-1} ebenfalls eine reelle Möbiustransformation. Nach Satz 9.2.3 ist auch die Komposition $h := g \circ f \circ g^{-1}$ eine reelle Möbiustransformation. Für h gilt

$$h(0) = g(f(g^{-1}(0))) = g(f(x_1)) = g(x_1) = 0 \text{ und} \tag{9.2.22}$$

$$h(\infty) = g(f(g^{-1}(\infty))) = g(f(x_2)) = g(x_2) = \infty. \tag{9.2.23}$$

Folglich existiert ein $\tilde{\alpha} \in \mathbb{R} \setminus \{0\}$, sodass gilt

$$h(x) := \begin{cases} \tilde{\alpha} x \in \mathbb{R} & \text{falls } x \in \mathbb{R}, \\ \infty & \text{falls } x = \infty. \end{cases} \tag{9.2.24}$$

Außerdem gilt nach der Kettenregel und der Umkehrregel die Gleichungskette

$$\tilde{\alpha} = h'(0) = g'(f(g^{-1}(0))) \cdot f'(g^{-1}(0)) \cdot (g^{-1})'(0) = g'(x_1) \cdot f'(x_1) \cdot (g^{-1})'(0) = f'(x_1). \tag{9.2.25}$$

Für den Rest des Beweises unterscheiden wir vier Fälle.

I. Im Fall $\alpha = 1$ gilt $f'(x_1) = f'(x_2) = 1$ und $h = \mathrm{id}$. Daraus folgt $f = g^{-1} \circ h \circ g = \mathrm{id}$. Dies ist jedoch nicht möglich, da nach Voraussetzung $\gamma \neq 0$ gilt. Folglich kann dieser Fall nicht eintreten.

II. Im Fall $\alpha = -1$ gilt $f'(x_1) = f'(x_2) = -1$ und $h^2 = \mathrm{id}$. Daraus folgt $f^2 = g^{-1} \circ h^2 \circ g = \mathrm{id}$.

III. Im Fall $|\alpha| < 1$ gilt $|f'(x_1)| < 1$ und

$$\begin{aligned} \forall x \in \mathbb{R} \cup \{\infty\} \setminus \{x_2\} : \quad \lim_{s \to \infty} f^{(s)}(x) &= \lim_{s \to \infty} g^{-1}\left(h^{(s)}(g(x))\right) \\ &= g^{-1}\left(\lim_{s \to \infty} \tilde{\alpha}^s \cdot g(x)\right) = g^{-1}(0) = x_1. \end{aligned} \tag{9.2.26}$$

IV. Im Fall $|\alpha| > 1$ gilt $|f'(x_2)| < 1$ und

$$\begin{aligned} \forall x \in \mathbb{R} \cup \{\infty\} \setminus \{x_1\} : \quad \lim_{s \to \infty} f^{(s)}(x) &= \lim_{s \to \infty} g^{-1}\left(h^{(s)}(g(x))\right) \\ &= g^{-1}\left(\lim_{s \to \infty} \tilde{\alpha}^s \cdot g(x)\right) = g^{-1}(\infty) = x_2. \end{aligned} \tag{9.2.27}$$

Damit ist gezeigt, dass genau eine der drei Aussagen (ii), (iii) oder (iv) gilt. □

Der soeben gezeigte Satz 9.2.4 ermöglicht den Beweis des folgenden Satzes.

Satz 9.2.5 (Konvergenz der IPF-Teilfolgen)**.** *Sei $(A, M^{(1)}, M^{(2)}, M^{(3)})$ ein mehrdimensionales Anpassungsproblem bestehend aus einer Ausgangstafel*

$$A = \left(\begin{matrix} a_{111} & 0 \\ 0 & a_{221} \end{matrix}\,\middle|\,\begin{matrix} 0 & a_{122} \\ 0 & 0 \end{matrix}\right) \in [0; \infty)^{2\times 2\times 2} \tag{9.2.28}$$

und eindimensionalen Marginaltafeln $M^{(1)}, M^{(2)}, M^{(3)} \in (0; \infty)^2$. Außerdem sei $(A(t))$ die zugehörige IPF-Folge. Dann sind die IPF-Teilfolgen $(A(3s))$, $(A(3s + 1))$ und $(A(3s + 2))$ konvergent.

Beweis. Der Beweis erfolgt in vier Schritten. In Schritt I zeigen wir, dass der Eintrag $a_{111}(1)$ durch eine Funktion vom Typ $x \mapsto (\alpha x + \beta)/(\gamma x + \delta)$ auf den Eintrag $a_{111}(4)$ abgebildet wird. Deren Parameter α, β, γ und δ hängen nur von den Marginaltafeln $M^{(1)}$, $M^{(2)}$ und $M^{(3)}$ ab. Schritt II zeigt, dass es sich bei dieser Abbildung um eine reelle Möbiustransformation handelt, die die Voraussetzungen von Satz 9.2.4 erfüllt. Anschließend zeigen wir in Schritt III, dass nur einer der beiden Fixpunkte dieser Möbiustransformation als Wert für die Einträge $a_{111}(3s + 1)$, $s \in \mathbb{N}_0$, infrage kommt. In Schritt IV zeigen wir schließlich, dass die Folge $(a_{111}(3s + 1))$ gegen diesen Fixpunkt konvergiert.

I. Das Anpassungsproblem $(A, M^{(1)}, M^{(2)}, M^{(3)})$ sei ohne Beschränkung der Allgemeinheit normiert. Folglich existieren Parameter $\rho, \sigma, \tau \in (0; 1)$, sodass gilt

$$M^{(1)} = \begin{pmatrix} \rho \\ 1 - \rho \end{pmatrix}, \qquad M^{(2)} = \begin{pmatrix} 1 - \sigma \\ \sigma \end{pmatrix}, \qquad M^{(3)} = \begin{pmatrix} \tau \\ 1 - \tau \end{pmatrix}. \tag{9.2.29}$$

Nach Definition des IPF-Verfahrens existiert außerdem ein $x \in (0; \rho)$, sodass gilt

$$A(1) = \left(\begin{matrix} x & 0 \\ 0 & -\rho + 1 \end{matrix}\,\middle|\,\begin{matrix} 0 & -x + \rho \\ 0 & 0 \end{matrix}\right). \tag{9.2.30}$$

Für die nächsten Tafeln der IPF-Folge gilt dann gemäß Definition des IPF-Verfahrens

$$A(2) = \left(\begin{matrix} -\sigma + 1 & 0 \\ 0 & \frac{-\rho\sigma+\sigma}{-x+1} \end{matrix}\,\middle|\,\begin{matrix} 0 & \frac{-\sigma x+\rho\sigma}{-x+1} \\ 0 & 0 \end{matrix}\right), \tag{9.2.31}$$

$$A(3) = \left(\begin{matrix} \frac{(\sigma\tau-\tau)x+(-\sigma\tau+\tau)}{(\sigma-1)x+(-\rho\sigma+1)} & 0 \\ 0 & \frac{-\rho\sigma\tau+\sigma\tau}{(\sigma-1)x+(-\rho\sigma+1)} \end{matrix}\,\middle|\,\begin{matrix} 0 & -\tau + 1 \\ 0 & 0 \end{matrix}\right) \text{ und} \tag{9.2.32}$$

$$A(4) = \left(\begin{matrix} \frac{(\rho\sigma\tau-\rho\tau)x+(-\rho\sigma\tau+\rho\tau)}{(\sigma-1)x+(\rho\sigma\tau-\rho\sigma-\sigma\tau+1)} & 0 \\ 0 & -\rho + 1 \end{matrix}\,\middle|\,\begin{matrix} 0 & \frac{(-\rho\sigma\tau+\rho\sigma+\rho\tau-\rho)x+(\rho^2\sigma\tau-\rho^2\sigma-\rho\tau+\rho)}{(\sigma-1)x+(\rho\sigma\tau-\rho\sigma-\sigma\tau+1)} \\ 0 & 0 \end{matrix}\right). \tag{9.2.33}$$

Der Eintrag $a_{111}(1) = x$ wird also durch die drei darauffolgenden Schritte des IPF-Verfahrens auf $a_{111}(4) = ((\rho\sigma\tau - \rho\tau)x + (-\rho\sigma\tau + \rho\tau))/((\sigma - 1)x + (\rho\sigma\tau - \rho\sigma - \sigma\tau + 1))$ abgebildet.

II. Um zu zeigen, dass es sich bei dieser Abbildung um eine reelle Möbiustransformation handelt, die die Voraussetzungen von Satz 9.2.4 erfüllt, wählen wir

$$\alpha := \rho\sigma\tau - \rho\tau \qquad \text{und} \qquad \beta := -\rho\sigma\tau + \rho\tau \text{ sowie} \tag{9.2.34}$$
$$\gamma := \sigma - 1 \qquad \text{und} \qquad \delta := \rho\sigma\tau - \rho\sigma - \sigma\tau + 1. \tag{9.2.35}$$

Dann gilt

$$\alpha\delta - \beta\gamma = \rho(1-\rho)\sigma(1-\sigma)\tau(1-\tau) > 0, \tag{9.2.36}$$
$$\gamma = \sigma - 1 < 0 \text{ und} \tag{9.2.37}$$
$$\begin{aligned}(\alpha-\delta)^2 + 4\beta\gamma = {} & \rho^2\sigma^2 + \rho^2\tau^2 + \sigma^2\tau^2 - 2\rho^2\sigma\tau - 2\rho\sigma^2\tau - 2\rho\sigma\tau^2 \\ & + 8\rho\sigma\tau - 2\rho\sigma - 2\rho\tau - 2\sigma\tau + 1.\end{aligned} \tag{9.2.38}$$

Um zu beweisen, dass $(\alpha - \delta)^2 + 4\beta\delta$ für jede Wahl von $\rho, \sigma, \tau \in (0; 1)$ positiv ist, betrachten wir diesen Term als Funktion von $\rho, \sigma, \tau \in [0; 1]$. Wir wählen also

$$\begin{aligned}g(\rho,\sigma,\tau) := {} & \rho^2\sigma^2 + \rho^2\tau^2 + \sigma^2\tau^2 - 2\rho^2\sigma\tau - 2\rho\sigma^2\tau - 2\rho\sigma\tau^2 \\ & + 8\rho\sigma\tau - 2\rho\sigma - 2\rho\tau - 2\sigma\tau + 1.\end{aligned} \tag{9.2.39}$$

Auf dem Rand des Definitionsbereichs $[0; 1]^3$ von g gilt folglich

$$g(0,\sigma,\tau) = (\sigma\tau - 1)^2 \geq 0 \quad \text{und} \quad g(1,\sigma,\tau) = (\sigma-1)^2(\tau-1)^2 \geq 0, \tag{9.2.40}$$
$$g(\rho,0,\tau) = (\rho\tau - 1)^2 \geq 0 \quad \text{und} \quad g(\rho,1,\tau) = (\rho-1)^2(\tau-1)^2 \geq 0 \text{ sowie} \tag{9.2.41}$$
$$g(\rho,\sigma,0) = (\rho\sigma - 1)^2 \geq 0 \quad \text{und} \quad g(\rho,\sigma,1) = (\rho-1)^2(\sigma-1)^2 \geq 0. \tag{9.2.42}$$

Somit ist g auf dem Rand von $[0; 1]^3$ nichtnegativ. Als nächstes suchen wir alle lokalen Extrema von g im Inneren von $[0; 1]^3$. Dazu betrachten wir die Ableitung $h(\rho, \sigma, \tau) := (\partial/\partial\rho)g(\rho, \sigma, \tau)$. Für diese Ableitung gilt

$$h(\rho,\sigma,\tau) = \frac{\partial}{\partial\rho} g(\rho,\sigma,\tau) = 2\rho\sigma^2 + 2\rho\tau^2 - 4\rho\sigma\tau - 2\sigma^2\tau - 2\sigma\tau^2 + 8\sigma\tau - 2\sigma - 2\tau. \tag{9.2.43}$$

Um herauszufinden, welche Werte h im Inneren von $[0; 1]^3$ annimmt, betrachten wir zuerst den Rand von $[0; 1]^3$. Dort gilt

$$h(0,\sigma,\tau) = -2\sigma^2\tau - 2\sigma\tau^2 + 8\sigma\tau - 2\sigma - 2\tau, \tag{9.2.44}$$
$$h(1,\sigma,\tau) = -2(\sigma+\tau)(\sigma-1)(\tau-1) \leq 0, \tag{9.2.45}$$
$$h(\rho,0,\tau) = 2(\rho\tau - 1)\tau \leq 0, \tag{9.2.46}$$
$$h(\rho,1,\tau) = 2(\rho-1)(\tau-1)^2 \leq 0, \tag{9.2.47}$$
$$h(\rho,\sigma,0) = 2(\rho\sigma - 1)\sigma \leq 0 \text{ und} \tag{9.2.48}$$
$$h(\rho,\sigma,1) = 2(\rho-1)(\sigma-1)^2 \leq 0. \tag{9.2.49}$$

Die Einschränkung $m(\sigma, \tau) := h(0, \sigma, \tau)$ der Funktion h auf den Rand $\rho = 0$ bedarf näherer Betrachtung. Auf dem Rand ihres Definitionsbereichs $[0; 1]^2$ gilt wiederum

$$m(0, \tau) = -2\tau \leq 0 \qquad \text{und} \qquad m(1, \tau) = -2(\tau - 1)^2 \leq 0 \text{ sowie} \tag{9.2.50}$$

$$m(\sigma, 0) = -2\sigma \leq 0 \qquad \text{und} \qquad m(\sigma, 1) = -2(\sigma - 1)^2 \leq 0. \tag{9.2.51}$$

Somit ist m auf dem Rand von $[0; 1]^2$ nichtnegativ. Als nächstes suchen wir alle lokalen Extrema der Funktion m im Inneren von $[0; 1]^2$. Eine notwendige Voraussetzung hierfür ist

$$0 = \frac{\partial}{\partial \sigma} m(\sigma, \tau) - \frac{\partial}{\partial \tau} m(\sigma, \tau) = 2(\sigma + \tau - 4)(\sigma - \tau). \tag{9.2.52}$$

Diese Voraussetzung ist in $[0; 1]^2$ genau dann erfüllt, wenn $\sigma = \tau$ gilt. In diesem Fall gilt

$$m(\sigma, \tau) = m(\sigma, \sigma) = -4\sigma(\sigma - 1)^2. \tag{9.2.53}$$

Folglich sind alle lokalen Extrema der Funktion m im Inneren von $[0; 1]^2$ negativ. Da die Funktion m auch auf dem Rand von $[0; 1]^2$ überall nichtpositiv ist, gilt auf ihrem gesamten Definitionsbereich

$$h(0, \sigma, \tau) = m(\sigma, \tau) \leq 0. \tag{9.2.54}$$

Somit ist die Funktion h überall auf dem Rand von $[0; 1]^3$ nichtpositiv.

Als nächstes suchen wir alle lokalen Extrema der Funktion h im Inneren von $[0; 1]^3$. Eine notwendige Voraussetzung hierfür ist

$$0 = \frac{\partial}{\partial \rho} h(\rho, \sigma, \tau) = 2\sigma^2 + 2\tau^2 - 4\sigma\tau = 2(\sigma - \tau)^2. \tag{9.2.55}$$

Diese Voraussetzung ist genau dann erfüllt, wenn $\sigma = \tau$ gilt. In diesem Fall gilt

$$h(\rho, \sigma, \tau) = h(\rho, \sigma, \sigma) = -4\sigma^3 + 8\sigma^2 - 4\sigma = -4\sigma(\sigma - 1)^2. \tag{9.2.56}$$

Somit sind alle lokalen Extrema von h im Inneren von $[0; 1]^3$ negativ. Folglich ist die Funktion $h = \partial g / \partial \rho$ im gesamten Inneren von $[0; 1]^3$ negativ.

Also verfügt die Funktion g über keinerlei lokale Extrema im Inneren von $[0; 1]^3$. Damit ist $(\alpha - \delta)^2 + 4\beta\delta = g(\rho, \sigma, \tau)$ für alle $\rho, \sigma, \tau \in (0; 1)$ positiv.

III. Sei $f : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ die reelle Möbiustransformation mit den in den Gleichungen (9.2.34) und (9.2.35) definierten Parametern α, β, γ und δ. Dann gilt

$$\forall s \in \mathbb{N}_0 : \quad a_{111}(3s + 1) = f^{(s)}(x). \tag{9.2.57}$$

Nach Satz 9.2.4 hat f genau zwei Fixpunkte $x_1, x_2 \in \mathbb{R}$. Für diese Fixpunkte gilt

$$x_1 = \frac{(\alpha - \delta) + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma} \quad \text{und} \quad x_2 = \frac{(\alpha - \delta) - \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma}. \tag{9.2.58}$$

Da nach Schritt II die Ungleichungen $2\gamma = 2(\sigma - 1) < 0$ und $(\alpha - \delta)^2 + 4\beta\gamma > 0$ gelten, folgt $x_1 < x_2$. Als nächstes betrachten wir die Terme

$$\alpha - \delta = \rho\sigma - \rho\tau + \sigma\tau - 1 < \rho - \rho\tau + \sigma - 1 = -(\rho - 1)(\sigma - 1) < 0 \text{ und} \tag{9.2.59}$$

$$4\beta\gamma = -4\rho(\sigma - 1)^2\tau < 0. \tag{9.2.60}$$

Folglich gilt $x_1 > 0$ und damit $0 < x_1 < x_2$. Um zu entscheiden, welche Fixpunkte im Intervall $(0; \rho)$ liegen, untersuchen wir die Differenzen

$$x_1 - \rho = \frac{(\alpha - \delta - 2\gamma\rho) + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma} \text{ und} \tag{9.2.61}$$

$$x_2 - \rho = \frac{(\alpha - \delta - 2\gamma\rho) - \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}}{2\gamma}. \tag{9.2.62}$$

Dazu betrachten wir die Differenz

$$\begin{aligned}(\alpha - \delta - 2\gamma\rho)^2 - ((\alpha - \delta)^2 + 4\beta\gamma) &= 4\gamma(\rho(\gamma\rho - \alpha + \delta) - \beta)\\ &= 4\rho(\rho - 1)(\sigma - 1)(\tau - 1) < 0.\end{aligned} \tag{9.2.63}$$

Folglich gilt $x_1 - \rho < 0$ und $x_2 - \rho > 0$ und somit schließlich $0 < x_1 < \rho < x_2$.

IV. Nach Satz 9.2.4 gilt genau eine der drei dort genannten Aussagen (ii), (iii) oder (iv). Um herauszufinden, welche der drei Aussagen gilt, betrachten wir die beiden Ableitungen $f'(x_1)$ und $f'(x_2)$. Nach Satz 9.2.2 gilt

$$f'(x_1) = \frac{\alpha\delta - \beta\gamma}{(\gamma x_1 + \delta)^2} = \frac{4(\alpha\delta - \beta\gamma)}{\left((\alpha + \delta) + \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}\right)^2} \text{ und} \tag{9.2.64}$$

$$f'(x_2) = \frac{\alpha\delta - \beta\gamma}{(\gamma x_2 + \delta)^2} = \frac{4(\alpha\delta - \beta\gamma)}{\left((\alpha + \delta) - \sqrt{(\alpha - \delta)^2 + 4\beta\gamma}\right)^2}. \tag{9.2.65}$$

Da nach Schritt II die Ungleichungen $\alpha\delta - \beta\gamma > 0$ und $(\alpha - \delta)^2 + 4\beta\gamma > 0$ gelten, gilt $f'(x_1) = f'(x_2)$ genau dann, wenn $\alpha + \delta = 0$ gilt. Für die Summe $\alpha + \delta$ gilt jedoch

$$\begin{aligned}\alpha + \delta &= 2\rho\sigma\tau - \rho\sigma - \rho\tau - \sigma\tau + 1\\ &= \rho(1 - \sigma)(1 - \tau) + (1 - \rho)\sigma(1 - \tau)\\ &\qquad + (1 - \rho)(1 - \sigma)\tau + (1 - \rho)(1 - \sigma)(1 - \tau)\\ &> 0.\end{aligned} \tag{9.2.66}$$

Somit gilt $f'(x_1) \neq f'(x_2)$. Folglich kann Aussage (ii) von Satz 9.2.4 nicht gelten. Da nach Definition des IPF-Verfahrens für alle $s \in \mathbb{N}_0$ die Aussage $f^{(s)}(x) = a_{111}(3s + 1) \in (0; \rho)$ gilt und nach Schritt III die Ungleichung $\rho < x_2$ gilt, kann Aussage (iv) von Satz 9.2.4 ebenfalls nicht gelten. Somit muss Aussage (iii) gelten. Also gilt

$$\lim_{s\to\infty} a_{111}(3s + 1) = \lim_{s\to\infty} f^{(s)}(x) = x_1. \tag{9.2.67}$$

Aus der Konvergenz der Folge $(a_{111}(3s+1))$ folgt die Konvergenz der Folge $(A(3s+1))$ und damit auch die Konvergenz der Folgen $(A(3s+2))$ und $(A(3s))$. □

9.3. Kommentare und Referenzen

Die in Abschnitt 9.1 vorgestellte Analyse des IPF-Verfahrens bei positiver $2\times 2\times 2$-Ausgangstafel und identischen, doppelt symmetrischen 2×2-Marginaltafeln geht auf Asci und Piccioni (2003) zurück. Die Notation wurde jedoch an die vorangegangenen Kapitel angepasst. Für den Beweis von Satz 9.1.1 verweisen Asci und Piccioni auf Lauritzen (1996, Theorem 4.13). Der hier angegebene Beweis basiert stattdessen auf der in Abschnitt 7.3 vorgestellten informationstheoretischen Konvergenzanalyse. Der Beweis von Satz 9.1.2 wurde von Asci und Piccioni übernommen. Dabei hängt insbesondere der Beweisschritt II stark von der Struktur der Marginaltafeln ab. Eine Verallgemeinerung ist daher nicht ohne Weiteres möglich.

Die in Abschnitt 9.2 präsentierten Ergebnisse zum IPF-Verfahren bei dünn besetzter $2\times 2\times 2$-Ausgangstafel und eindimensionalen Marginaltafeln sind bisher unveröffentlicht. Die in Definition 9.2.1 eingeführten Möbiustransformationen werden üblicherweise für komplexe Zahlen definiert und untersucht (vgl. Beardon 1991, Abschnitt 1.2; Needham 1997, Abschnitt 3.I.1). Dabei werden sowohl die reellen Parameter α, β, γ und δ als auch die reelle Variable x durch komplexe Zahlen ersetzt. Die hier getroffene Einschränkung auf reelle Zahlen vereinfacht jedoch die Aussagen und Beweise der Sätze 9.2.2, 9.2.3 und 9.2.4. Eine Verallgemeinerung von Satz 9.2.5 auf Ausgangstafeln mit vier oder mehr Einträgen erscheint schwierig, da in diesem Fall rationale Funktionen höheren Grades an die Stelle der Möbiustransformationen treten. Deren Iteration ist deutlich schwerer zu handhaben.

10. Unmöglichkeit der variationellen Charakterisierung der Limeszyklen des mehrdimensionalen IPF-Verfahrens

In diesem Kapitel wird die Möglichkeit einer variationellen Charakterisierung der Häufungspunkte des mehrdimensionalen IPF-Verfahrens diskutiert, die vergleichbar ist mit der in Abschnitt 4.2 eingeführten informationstheoretischen Charakterisierung der Häufungspunkte des zweidimensionalen IPF-Verfahrens. Dazu wird im ersten Abschnitt erörtert, wie eine derartige Charakterisierung aussehen kann. Im zweiten Abschnitt werden die geometrischen Eigenschaften der I-Divergenz auf dem Wahrscheinlichkeitssimplex diskutiert. Darauf aufbauend wird im dritten Abschnitt die Nichtexistenz der gewünschten Charakterisierung bewiesen. Im vierten Abschnitt werden weitere Ansätze zur Divergenzanalyse des mehrdimensionalen IPF-Verfahrens diskutiert.

10.1. Variationelle Charakterisierung der Limeszyklen

In Kapitel 3 (siehe S. 44) wurde für alle biproportionalen Anpassungsprobleme (A, c, r) die Konvergenz der zugehörigen IPF-Teilfolgen $(A(2s + 1))$ und $(A(2s + 2))$ bewiesen. Des Weiteren wurde in Kapitel 9 (siehe S. 87) für einige spezielle dreidimensionale Anpassungsprobleme $(A, M^{(1)}, M^{(2)}, M^{(3)})$ die Konvergenz der zugehörigen IPF-Teilfolgen $(A(3s + 1))$, $(A(3s + 2))$ und $(A(3s + 3))$ gezeigt. Für allgemeine mehrdimensionale Anpassungsprobleme $(A, M^{(1)}, \dots, M^{(\eta)})$ legen umfangreiche Simulationen ebenfalls die Konvergenz der zugehörigen IPF-Teilfolgen $(A(s\eta + i))$, $i = 1, \dots, \eta$, nahe.

Hypothese 10.1.1 (Konvergenz der IPF-Teilfolgen)**.** *Sei $(A, M^{(1)}, \dots, M^{(\eta)})$ ein mehrdimensionales Anpassungsproblem und sei $(A(t))$ die zugehörige IPF-Folge. Dann konvergieren die IPF-Teilfolgen $(A(s\eta + i))$, $i = 1, \dots, \eta$.*

Wenn die IPF-Teilfolgen im Sinne von Hypothese 10.1.1 konvergieren, dann bilden ihre Grenzwerte einen Zyklus bezüglich der I_1-Projektionen auf die zu $(A, M^{(1)}, \dots, M^{(\eta)})$ gehörenden Mengen $\mathcal{E}_i^{\ll A}$, $i = 1, \dots, \eta$. Dies zeigt der folgende Satz.

Satz 10.1.2 (Zykluseigenschaft der Häufungspunkte)**.** *Sei $(A, M^{(1)}, \dots, M^{(\eta)})$ ein normiertes mehrdimensionales Anpassungsproblem, $(A(t))$ die zugehörige IPF-Folge und $\mathcal{E}_i^{\ll A}$, $i = 1, \dots, \eta$, die zugehörigen Mengen von Wahrscheinlichkeitstafeln. Außerdem seien die IPF-Teilfolgen $(A(s\eta + i))$, $i = 1, \dots, \eta$, konvergent und $B^*(i) := \lim_{s\to\infty} A(s\eta + i) \in \mathcal{E}_i^{\ll A}$, $i = 1, \dots, \eta$, ihre Grenzwerte. Dann ist $B^*(1)$ die I_1-Projektion von $B^*(\eta)$ auf $\mathcal{E}_1^{\ll A}$. Des Weiteren ist $B^*(i)$ für alle $i = 2, \dots, \eta$ die I_1-Projektion von $B^*(i-1)$ auf $\mathcal{E}_i^{\ll A}$.*

Beweis. Für alle $s \in \mathbb{N}_0$ ist die Tafel $A(s\eta + 1)$ nach Lemma 7.3.1 (ii) (siehe S. 77) gleich der I_1-Projektion von $A(s\eta)$ auf $\mathcal{E}_1^{\ll A}$. Aufgrund der Stetigkeit der I_1-Projektion auf $\mathcal{E}_1^{\ll A}$ (siehe

Satz 2.2.4, S. 28) ist folglich der Grenzwert $B^*(1) = \lim_{s\to\infty} A(s\eta + 1)$ gleich der I_1-Projektion von $B^*(\eta) = \lim_{s\to\infty} A(s\eta)$ auf $\mathcal{E}_1^{\ll A}$. In analoger Weise können wir für alle $i = 2, \ldots, \eta$ zeigen, dass der Grenzwert $B^*(i)$ gleich der I_1-Projektion von $B^*(i-1)$ auf $\mathcal{E}_i^{\ll A}$ ist. □

Aufgrund der soeben gezeigten Zykluseigenschaft liegt es nahe, die zu den Mengen $\mathcal{E}_i^{\ll A}$, $i = 1, \ldots, \eta$, gehörende Menge von Zyklen zu untersuchen. Seien daher $\mathcal{E}_i$, $i = 1, \ldots, \eta$, beliebige nichtleere, konvexe, abgeschlossene Mengen von Wahrscheinlichkeitstafeln. Dann ist die *zu den Mengen* $\mathcal{E}_i$*,* $i = 1, \ldots, \eta$*, gehörende Menge* $\mathrm{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta)$ *von Zyklen* definiert gemäß

$$\begin{aligned}
&\mathrm{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta) \\
&:= \left\{ (p^{(1)}, \ldots, p^{(\eta)}) \in \mathcal{E}_1 \times \cdots \times \mathcal{E}_\eta \;\middle|\; \begin{array}{l} p^{(1)} \text{ ist } \mathrm{I}_1\text{-Projektion von } p^{(\eta)} \text{ auf } \mathcal{E}_1, \\ p^{(2)} \text{ ist } \mathrm{I}_1\text{-Projektion von } p^{(1)} \text{ auf } \mathcal{E}_2, \\ p^{(3)} \text{ ist } \mathrm{I}_1\text{-Projektion von } p^{(2)} \text{ auf } \mathcal{E}_3, \\ \qquad\vdots \\ p^{(\eta)} \text{ ist } \mathrm{I}_1\text{-Projektion von } p^{(\eta-1)} \text{ auf } \mathcal{E}_\eta \end{array} \right\}.
\end{aligned} \tag{10.1.1}$$

Für alle Zyklen $(p^{(1)}, \ldots, p^{(\eta)}) \in \mathrm{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta) \subseteq \mathcal{E}_1 \times \cdots \times \mathcal{E}_\eta$ gilt aufgrund der in Gleichung (10.1.1) geforderten I_1-Projektionseigenschaften die Aussage

$$p^{(1)} \ll p^{(\eta)} \ll p^{(\eta-1)} \ll \cdots \ll p^{(3)} \ll p^{(2)} \ll p^{(1)}. \tag{10.1.2}$$

Daraus folgt

$$p^{(1)} \equiv p^{(2)} \equiv p^{(3)} \equiv \cdots \equiv p^{(\eta-1)} \equiv p^{(\eta)}. \tag{10.1.3}$$

Also ist die Existenz eines Tupels $(p^{(1)}, \ldots, p^{(\eta)}) \in \mathcal{E}_1 \times \cdots \times \mathcal{E}_\eta$ mit $p^{(1)} \equiv \cdots \equiv p^{(\eta)}$ eine notwendige Voraussetzung für die Existenz eines Zyklus zu den Mengen $\mathcal{E}_i$, $i = 1, \ldots, \eta$.

Für das zweidimensionale IPF-Verfahren und die zugehörigen Mengen $\mathcal{C}^{\ll A}$ und $\mathcal{R}^{\ll A}$ wurde in Satz 4.2.3 (siehe S. 54) gezeigt, dass gilt

$$\mathrm{cyc}\left(\mathcal{C}^{\ll A}, \mathcal{R}^{\ll A}\right) = \operatorname*{arg\,min}_{C \in \mathcal{C}^{\ll A}, R \in \mathcal{R}^{\ll A}} \mathrm{I}(C|R). \tag{10.1.4}$$

Folglich bietet es sich an, für das mehrdimensionale IPF-Verfahren eine vergleichbare *variationelle Charakterisierung der Zyklen* zu suchen, das heißt eine Funktion $\Phi : \overline{\mathcal{H}}^\eta \to \mathbb{R}$, sodass für alle geordneten Familien $(\mathcal{E}_1, \ldots, \mathcal{E}_\eta)$ von abgeschlossenen, konvexen Teilmengen von $\overline{\mathcal{H}}$ mit mindestens einem Tupel $(p^{(1)}, \ldots, p^{(\eta)}) \in \mathcal{E}_1 \times \cdots \times \mathcal{E}_\eta$, das $p^{(1)} \equiv \cdots \equiv p^{(\eta)}$ erfüllt, gilt

$$\mathrm{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta) = \operatorname*{arg\,min}_{p^{(1)} \in \mathcal{E}_1, \ldots, p^{(\eta)} \in \mathcal{E}_\eta} \Phi\left(p^{(1)}, \ldots, p^{(\eta)}\right). \tag{10.1.5}$$

Dabei bezeichnet $\overline{\mathcal{H}}$ die Menge aller Wahrscheinlichkeitstafeln.

10.2. Informationsdivergenz auf dem Wahrscheinlichkeitssimplex

Für die nachfolgenden Überlegungen spielt die kartesische Struktur der Wahrscheinlichkeitstafeln keine Rolle. Daher empfiehlt es sich, die Wahrscheinlichkeitstafeln als Wahrscheinlichkeitsvektoren im $\mathbb{R}^k$ zu betrachten, wobei $k \geq 3$ gilt. Dann ist die Menge der Wahrscheinlichkeitstafeln gleich dem *abgeschlossenen Wahrscheinlichkeitssimplex* $\overline{\mathcal{H}} := \{p \in [0;\infty)^k \mid p_1 + \cdots + p_k = 1\}$. Die Menge der positiven Wahrscheinlichkeitstafeln ist gleich dem *offenen Wahrscheinlichkeitssimplex* $\mathcal{H} := \{p \in (0;\infty)^k \mid p_1 + \cdots + p_k = 1\}$. Für die Untersuchung der I_1-Projektionen ist die Funktion $\mathrm{I}(\cdot|q) : \overline{\mathcal{H}} \to [0;\infty]$ bei gegebenem $q \in \overline{\mathcal{H}}$ von besonderem Interesse. Für den Fall $q \in \mathcal{H}$ zeigt das folgende Lemma einige Eigenschaften der Funktion $\mathrm{I}(\cdot|q) : \overline{\mathcal{H}} \to [0;\infty)$.

Lemma 10.2.1 (Strikte Konvexität, Stetigkeit und Differenzierbarkeit der I-Divergenz im ersten Argument)**.** *Sei $q \in \mathcal{H}$. Dann gelten für die Funktion $\mathrm{I}(\cdot|q) : \overline{\mathcal{H}} \to [0;\infty)$ die folgenden vier Aussagen:*

(i) Die Funktion $\mathrm{I}(\cdot|q)$ ist strikt konvex.

(ii) Die Funktion $\mathrm{I}(\cdot|q)$ ist stetig.

(iii) Die Funktion $\mathrm{I}(\cdot|q)$ ist auf $\mathcal{H}$ beliebig oft differenzierbar.

(iv) Die Funktion $\mathrm{I}(\cdot|q)$ nimmt ihr eindeutiges Minimum im Punkt q an.

Beweis. (i) Da nach Voraussetzung $q \in \mathcal{H}$ und damit $\mathrm{supp}(q) = \{1,\ldots,k\}$ gilt, folgt aus Satz 2.1.2 (iii) (siehe S. 23) die strikte Konvexität der Funktion $\mathrm{I}(\cdot|q)$.

(ii) Gemäß Definition 2.1.1 (siehe S. 23) gilt

$$\forall p \in \overline{\mathcal{H}} : \quad \mathrm{I}(p|q) = \sum_i \overline{\varphi}(p_i, q_i). \tag{10.2.1}$$

Da nach Voraussetzung $q \in \mathcal{H}$ und damit $q_i > 0$ für alle $i = 1,\ldots,k$ gilt, folgt für alle Summanden $\overline{\varphi}(p_i, q_i)$ die Gleichung

$$\overline{\varphi}(p_i, q_i) = \begin{cases} p_i \ln \frac{p_i}{q_i} & \text{falls } p_i > 0, \\ 0 & \text{falls } p_i = 0. \end{cases} \tag{10.2.2}$$

Damit ist $\overline{\varphi}(p_i, q_i)$ stetig in p_i. Folglich ist $\mathrm{I}(p|q)$ stetig in p.

(iii) Gemäß Definition 2.1.1 (siehe S. 23) gilt

$$\forall p \in \mathcal{H} : \quad \mathrm{I}(p|q) = \sum_i p_i \ln \frac{p_i}{q_i} = \sum_i p_i \ln p_i - \sum_i p_i \ln q_i. \tag{10.2.3}$$

Folglich ist $\mathrm{I}(\cdot|q)$ auf $\mathcal{H}$ beliebig oft differenzierbar.

(iv) Gemäß Korollar 2.1.4 (siehe S. 24) ist $\mathrm{I}(p|q)$ nach unten durch 0 beschränkt. Das Minimum wird genau dann angenommen, wenn $p = q$ gilt. □

Ebenfalls von Interesse ist im folgenden Abschnitt die Einschränkung der Funktion $\mathrm{I}(\cdot|q)$ auf eine Strecke durch $q \in \mathcal{H}$.

Lemma 10.2.2 (I-Divergenz auf Strecken)**.** *Sei $q \in \mathcal{H}$ und sei h ein Richtungsvektor in $\mathcal{H}$, das heißt, gelte $h \in \mathbb{R}^k \setminus \{0\}$ mit $h_1 + \cdots + h_k = 0$. Dann gilt die Aussage:*

(i) Die Gerade $q + \mathbb{R}h$ schneidet das abgeschlossene Wahrscheinlichkeitssimplex $\overline{\mathcal{H}}$ in einer Strecke,

$$\exists \alpha_{\min} < 0 \quad \exists \alpha_{\max} > 0 : \quad (q + \mathbb{R}h) \cap \overline{\mathcal{H}} = q + [\alpha_{\min}; \alpha_{\max}]\, h. \tag{10.2.4}$$

Für die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ gelten die folgenden sechs Aussagen:

(ii) Die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ ist strikt konvex auf $[\alpha_{\min}; \alpha_{\max}]$.

(iii) Die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ ist stetig auf $[\alpha_{\min}; \alpha_{\max}]$.

(iv) Die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ ist beliebig oft differenzierbar auf $(\alpha_{\min}; \alpha_{\max})$.

(v) Die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ nimmt ihr eindeutiges Minimum im Punkt $\alpha = 0$ an.

(vi) Die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ ist streng monoton fallend auf $[\alpha_{\min}; 0]$.

(vii) Die Funktion $\alpha \mapsto \mathrm{I}(q + \alpha h|q)$ ist streng monoton steigend auf $[0; \alpha_{\max}]$.

Beweis. Aussage (i) folgt aus der Konvexität und der Kompaktheit der Menge $\overline{\mathcal{H}}$ sowie der Wahl der Vektoren q und h. Die Aussagen (ii) bis (v) folgen aus Lemma 10.2.1. Die Aussagen (vi) und (vii) folgen aus den Aussagen (ii) bis (v). □

Ein interessanter Spezialfall der Funktion $\mathrm{I}(\cdot|q) : \overline{\mathcal{H}} \to [0; \infty)$ ist die I-Divergenz $\mathrm{I}(\cdot|\mu)$ relativ zur Gleichverteilung $\mu := (\frac{1}{k}, \ldots, \frac{1}{k})^T \in \mathcal{H}$.

Lemma 10.2.3 (I-Divergenz relativ zur Gleichverteilung)**.** *Sei $\mu \in \mathcal{H}$ die Gleichverteilung. Dann gelten die folgenden beiden Aussagen:*

(i) Die Funktion $\mathrm{I}(\cdot|\mu)$ nimmt auf dem abgeschlossenen Wahrscheinlichkeitssimplex $\overline{\mathcal{H}}$ nur Werte von 0 bis $\ln k$ an,

$$\forall p \in \overline{\mathcal{H}} : \quad \mathrm{I}(p|\mu) \in [0; \ln k]\,. \tag{10.2.5}$$

Das Maximum $\ln k$ wird nur in den Eckpunkten $e^{(1)} := (1, 0, \ldots, 0)^T$ bis $e^{(k)} := (0, \ldots, 0, 1)^T$ von $\overline{\mathcal{H}}$ angenommen.

(ii) Die Funktion $\mathrm{I}(\cdot|\mu)$ nimmt auf dem Rand $\partial\mathcal{H}$ des Wahrscheinlichkeitssimplex nur Werte von $\ln \frac{k}{k-1}$ bis $\ln k$ an,

$$\forall p \in \partial\mathcal{H} : \quad \mathrm{I}(p|\mu) \in \left[\ln \frac{k}{k-1}; \ln k\right]. \tag{10.2.6}$$

Das Minimum $\ln \frac{k}{k-1}$ wird nur in den Punkten $(0, \frac{1}{k-1}, \ldots, \frac{1}{k-1})^T$ bis $(\frac{1}{k-1}, \ldots, \frac{1}{k-1}, 0)^T$ angenommen.

Beweis. (i) Für die Eckpunkte $e^{(1)}, \dots, e^{(k)}$ gilt

$$\mathrm{I}(e^{(1)}|\mu) = \dots = \mathrm{I}(e^{(k)}|\mu) = \ln k. \tag{10.2.7}$$

Für alle anderen Punkte $p \in \overline{\mathcal{H}} \setminus \{e^{(1)}, \dots, e^{(k)}\}$ folgt aufgrund der strikten Konvexität der Funktion $\mathrm{I}(\cdot|q)$ (siehe Lemma 10.2.1 (i)) die Ungleichung

$$\begin{aligned}\mathrm{I}(p|\mu) &= \mathrm{I}(p_1 e^{(1)} + \dots + p_k e^{(k)}|\mu) \\ &< p_1\,\mathrm{I}(e^{(1)}|\mu) + \dots + p_k\,\mathrm{I}(e^{(k)}|\mu) = (p_1 + \dots + p_k)\ln k = \ln k.\end{aligned} \tag{10.2.8}$$

(ii) Sei $p \in \partial\mathcal{H}$. Dann gilt ohne Beschränkung der Allgemeinheit $p_1 = 0$. Daraus folgt mit der Aggregationsungleichung für Vektoren (siehe Satz 2.1.3, S. 24) die Ungleichung

$$\begin{aligned}\mathrm{I}(p|\mu) &= \mathrm{I}\left(p_1\middle|\frac{1}{k}\right) + \mathrm{I}\left((p_2, \dots, p_k)\middle|\left(\frac{1}{k}, \dots, \frac{1}{k}\right)\right) = \mathrm{I}\left((p_2, \dots, p_k)\middle|\left(\frac{1}{k}, \dots, \frac{1}{k}\right)\right) \\ &\geq \mathrm{I}\left(p_2 + \dots + p_k\middle|\frac{1}{k} + \dots + \frac{1}{k}\right) = \mathrm{I}\left(1\middle|\frac{k-1}{k}\right) = \ln\frac{k}{k-1}.\end{aligned} \tag{10.2.9}$$

Gleichheit gilt genau dann, wenn $p_i/(1/k) = 1/((k-1)/k)$ für alle $i = 2, \dots, k$ gilt. Dies ist genau dann der Fall, wenn $p_i = \frac{1}{k-1}$ für alle $i = 2, \dots, k$ gilt. □

Das soeben gezeigte Lemma 10.2.3 ist für die Diskussion der Funktion $\mathrm{I}(\cdot|\mu)$ von zentraler Bedeutung. Da $\mathrm{I}(\cdot|\mu)$ auf dem Rand $\partial\mathcal{H}$ nach unten durch $\ln\frac{k}{k-1}$ beschränkt ist, nimmt die Funktion $\alpha \mapsto \mathrm{I}(\mu + \alpha h|\mu)$ bei gegebenem Richtungsvektor h im Sinne von Lemma 10.2.2 jeden Wert im Intervall $[0; \ln\frac{k}{k-1}]$ an. Die Interpretation der I-Divergenz $\mathrm{I}(p|\mu)$ als „Abstand" des Vektors p zur Gleichverteilung μ erlaubt die Definition von I_1-Kugeln und I_1-Sphären um μ.

Definition 10.2.4 (I_1-Kugel und I_1-Sphäre, vgl. Csiszár 1975, S. 146 f.)**.** Sei $\mu \in \mathcal{H}$ die Gleichverteilung und sei $\rho \in [0; \infty)$. Dann heißt die Menge $\mathcal{B}(\mu, \rho) := \{p \in \overline{\mathcal{H}} \mid \mathrm{I}(p|\mu) \leq \rho\}$ die *I_1-Kugel um μ mit Radius ρ*. Die Menge $\mathcal{S}(\mu, \rho) := \{p \in \overline{\mathcal{H}} \mid \mathrm{I}(p|\mu) = \rho\}$ heißt die *I_1-Sphäre um μ mit Radius ρ*.

Aus den zuvor diskutierten Eigenschaften der Funktion $\mathrm{I}(\cdot|\mu)$ folgen einige wichtige Eigenschaften der I_1-Kugeln $\mathcal{B}(\mu, \rho)$.

Lemma 10.2.5 (Abgeschlossenheit, Konvexität und Lage von I_1-Kugeln)**.** *Sei $\mu \in \mathcal{H}$ die Gleichverteilung. Dann gelten die folgenden beiden Aussagen:*

(i) Für alle Radien $\rho \in [0; \infty)$ ist die I_1-Kugel $\mathcal{B}(\mu, \rho)$ abgeschlossen und konvex.

(ii) Für alle Radien $\rho \in [0; \ln\frac{k}{k-1})$ ist die I_1-Kugel $\mathcal{B}(\mu, \rho)$ eine echte Teilmenge des offenen Wahrscheinlichkeitssimplex,

$$\forall \rho \in \left[0; \ln\frac{k}{k-1}\right): \quad \mathcal{B}(\mu, \rho) \subsetneq \mathcal{H}. \tag{10.2.10}$$

Beweis. (i) Aus der Stetigkeit und der Konvexität der Funktion $\mathrm{I}(\cdot|\mu)$ (siehe Lemma 10.2.1) folgt die Abgeschlossenheit und die Konvexität der I_1-Kugel $\mathcal{B}(\mu, \rho)$.

(ii) Sei $\rho \in [0; \ln \frac{k}{k-1})$ beliebig. Nach Definition der I_1-Kugel gilt $\mathcal{B}(\mu, \rho) \subseteq \overline{\mathcal{H}}$. Da gemäß Lemma 10.2.3 (ii) für alle $p \in \partial\mathcal{H}$ die Ungleichung $\mathrm{I}(p|\mu) \geq \ln \frac{k}{k-1}$ und damit $p \notin \mathcal{B}(\mu, \rho)$ gilt, folgt $\mathcal{B}(\mu, \rho) \subseteq \mathcal{H}$. Gemäß Lemma 10.2.3 (i) gilt außerdem $\mathrm{I}(e^{(1)}|\mu) = \ln k > \ln \frac{k}{k-1}$. Aufgrund der Stetigkeit der Funktion $\mathrm{I}(\cdot|\mu)$ (siehe Lemma 10.2.1 (ii)) existiert folglich ein $p \in \mathcal{H}$, für das $\mathrm{I}(p|\mu) > \ln \frac{k}{k-1}$ und damit $p \notin \mathcal{B}(\mu, \rho)$ gilt. Somit gilt schließlich $\mathcal{B}(\mu, \rho) \subsetneq \mathcal{H}$. □

Von besonderem Interesse ist im folgenden Abschnitt die Schnittmenge einer beliebigen I_1-Sphäre $\mathcal{S}(\mu, \rho)$ mit Radius $\rho \in (0; \ln \frac{k}{k-1})$ mit einem beliebigen zweidimensionalen affinen Unterraum $\mathcal{A}$ der affinen Hülle $\mathrm{aff}(\mathcal{H}) = \{x \in \mathbb{R}^k \mid x_1 + \cdots + x_k = 1\}$ des Wahrscheinlichkeitssimplex $\mathcal{H}$, der den Punkt μ enthält. Sei daher $\{c^{(1)}, c^{(2)}\}$ eine Orthonormalbasis des zugehörigen linearen Unterraums $\mathcal{A} - \mu$. Dann kann $\{c^{(1)}, c^{(2)}\}$ zu einer Orthonormalbasis $\{c^{(1)}, \ldots, c^{(k)}\}$ des $\mathbb{R}^k$ erweitert werden, für die gilt

$$c^{(k)} = \left(\sqrt{\frac{1}{k}}, \ldots, \sqrt{\frac{1}{k}} \right)^T. \tag{10.2.11}$$

Darüber hinaus sei

$$C := \begin{pmatrix} - & {c^{(1)}}^T & - \\ & \vdots & \\ - & {c^{(k)}}^T & - \end{pmatrix} \in \mathbb{R}^{k \times k} \quad \text{und} \quad C_{3:k} := \begin{pmatrix} - & {c^{(3)}}^T & - \\ & \vdots & \\ - & {c^{(k)}}^T & - \end{pmatrix} \in \mathbb{R}^{(k-2) \times k}. \tag{10.2.12}$$

Dann gilt

$$\mathcal{A} = \mu + \mathbb{R}c^{(1)} + \mathbb{R}c^{(2)} = \{x \in \mathbb{R}^k \mid C_{3:k}\, x = C_{3:k}\, \mu\}. \tag{10.2.13}$$

Das folgende Lemma zeigt, dass die Schnittmenge $\mathcal{A} \cap \mathcal{S}(\mu, \rho)$ als geschlossene und stetig differenzierbare Kurve parametrisiert werden kann.

Lemma 10.2.6 (Schnitt von I_1-Sphären und Ebenen)**.** *Sei $\mathcal{A}$ ein zweidimensionaler affiner Unterraum von* $\mathrm{aff}(\mathcal{H})$*, der den Punkt μ enthält, und sei $\{c^{(1)}, \ldots, c^{(k)}\}$ die zugehörige Orthonormalbasis des $\mathbb{R}^k$ sowie C und $C_{3:k}$ die zugehörigen Matrizen. Weiter sei $\mathcal{S}(\mu, \rho)$ eine beliebige I_1-Sphäre um μ mit Radius $\rho \in (0; \ln \frac{k}{k-1})$. Dann gilt:*

(i) Es existiert eine geschlossene, stetig differenzierbare und nach Bogenlänge parametrisierte Kurve $z : [0; T] \to (0; \infty)^k$ mit

$$\mathrm{Bild}\, z = \mathcal{A} \cap \mathcal{S}(\mu, \rho). \tag{10.2.14}$$

(ii) Für die Ableitung $z'(t)$ der Kurve $z : [0; T] \to (0; \infty)^k$ gilt

$$\forall t \in [0; T]: \quad z'(t) = \left\| C^T R C \ln\{z(t)\} \right\|_2^{-1} \cdot C^T R C \ln\{z(t)\}. \tag{10.2.15}$$

Dabei gilt

$$R := \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} \in \mathbb{R}^{k \times k}. \tag{10.2.16}$$

Beweis. (i) Nach Voraussetzung gilt

$$\mathcal{A} = \left\{ \mu + r \left(\cos(\omega) c^{(1)} + \sin(\omega) c^{(2)} \right) \,\middle|\, r \in [0;\infty),\ \omega \in [0;2\pi) \right\}. \tag{10.2.17}$$

Anstelle der Schnittmenge $\mathcal{A} \cap \mathcal{S}(\mu, \rho)$ betrachten wir zunächst die Schnittmenge $\mathcal{A} \cap \mathcal{B}(\mu, \rho_{\max})$ mit $\rho_{\max} := \ln \frac{k}{k-1} > \rho$. Außerdem sei

$$K := \left\{ (r,\omega) \in [0;\infty) \times [0;2\pi) \,\middle|\, \mu + r \left(\cos(\omega) c^{(1)} + \sin(\omega) c^{(2)} \right) \in \mathcal{B}(\mu, \rho_{\max}) \right\}. \tag{10.2.18}$$

Gemäß der Lemmata 10.2.2 und 10.2.3 (ii) existiert zu jedem Winkel $\omega \in [0;2\pi)$ ein Radius $R(\omega) > 0$, sodass gilt

$$\mathrm{I}\left(\mu + R(\omega) \left(\cos(\omega) c^{(1)} + \sin(\omega) c^{(2)} \right) \middle| \mu \right) = \ln \frac{k}{k-1} = \rho_{\max}. \tag{10.2.19}$$

Aufgrund von Lemma 10.2.2 gilt außerdem

$$K = \bigcup_{\omega \in [0;2\pi)} [0; R(\omega)] \times \{\omega\}. \tag{10.2.20}$$

Sei $F : K \to [0;\rho_{\max}] \times [0;2\pi)$ definiert gemäß

$$F(r,\omega) := \begin{pmatrix} \mathrm{I}\left(\mu + r \left(\cos(\omega) c^{(1)} + \sin(\omega) c^{(2)} \right) \middle| \mu \right) \\ \omega \end{pmatrix}. \tag{10.2.21}$$

Weiter sei $\omega \in [0;2\pi)$ beliebig. Dann ist die Funktion $F_1(\cdot,\omega)$ nach Lemma 10.2.2 stetig und streng monoton steigend. Außerdem gilt $F_1(0,\omega) = 0$ und $F_1(R(\omega),\omega) = \rho_{\max}$. Folglich ist die Funktion F bijektiv. Sei daher $G : [0;\rho_{\max}] \times [0;2\pi) \to K$ die Umkehrfunktion von F.

Für die Ableitung der Funktion F gilt

$$\mathrm{D}_F(r,\omega) = \begin{pmatrix} \frac{\partial}{\partial r} F_1(r,\omega) & \frac{\partial}{\partial \omega} F_1(r,\omega) \\ 0 & 1 \end{pmatrix}. \tag{10.2.22}$$

Somit ist die Funktion F in allen Punkten $(r,\omega) \in K$ mit $\mu + r(\cos(\omega)c^{(1)} + \sin(\omega)c^{(2)}) \in \operatorname{int} \mathcal{B}(\mu, \rho_{\max}) \subsetneq \mathcal{H}$ stetig differenzierbar. Falls $r > 0$ gilt, gilt außerdem

$$\frac{\partial}{\partial r} F_1(r,\omega) > 0. \tag{10.2.23}$$

In diesem Fall ist die Matrix $\mathrm{D}_F(r,\omega)$ invertierbar. Nach dem Satz von der Umkehrabbildung ist die Funktion G somit in allen Punkten $(\tilde{\rho},\omega) \in (0;\rho_{\max}) \times [0;2\pi)$ stetig differenzierbar.

Als nächstes betrachten wir die Kurve $\tilde{z} : [0;2\pi] \to (0;\infty)^k$, wobei

$$\tilde{z}(\omega) := \mu + G_1(\rho,\omega) \left(\cos(\omega) c^{(1)} + \sin(\omega) c^{(2)} \right). \tag{10.2.24}$$

Wegen der Periodizität der Winkelfunktionen ist die Kurve $\tilde{z}$ geschlossen. Aufgrund der obigen Diskussion der Funktionen F und G ist $\tilde{z}$ außerdem stetig differenzierbar und es gilt

$$\operatorname{Bild} \tilde{z} = \mathcal{A} \cap \mathcal{S}(\mu, \rho) \subsetneq (0;\infty)^k. \tag{10.2.25}$$

Die Inklusion $\mathcal{A} \cap \mathcal{S}(\mu, \rho) \subsetneq (0; \infty)^k$ folgt mit Lemma 10.2.5 (ii) aus der Voraussetzung $\rho \in (0; \ln \frac{k}{k-1})$. Da $\measuredangle(\tilde{z}(0), \mu, \tilde{z}(\omega)) = \omega$ für alle $\omega \in [0; 2\pi)$ gilt, folgt $\tilde{z}'(\omega) \neq 0$ für alle $\omega \in [0; 2\pi]$. Folglich kann $\tilde{z} : [0; 2\pi] \to (0; \infty)^k$ nach Bogenlänge reparametrisiert werden und wir erhalten eine geschlossene, stetig differenzierbare und nach Bogenlänge parametrisierte Kurve $z : [0; T] \to (0; \infty)^k$ mit

$$\operatorname{Bild} z = \mathcal{A} \cap \mathcal{S}(\mu, \rho). \tag{10.2.26}$$

(ii) Die Abbildung $\ell : (0; \infty)^k \to \mathbb{R}$ sei definiert gemäß

$$\ell(x) := \sum_i x_i \ln x_i = \mathrm{I}(x|\mu) + \left(\sum_i x_i\right) \ln \frac{1}{k}. \tag{10.2.27}$$

Für die Ableitung von ℓ gilt dann

$$\forall x \in (0; \infty)^k : \quad \mathrm{D}_\ell(x) = (1 + \ln x_1, \ldots, 1 + \ln x_k). \tag{10.2.28}$$

Außerdem gilt

$$\forall t \in [0; T] : \quad \ell(z(t)) = \ell(z(0)). \tag{10.2.29}$$

Daraus folgt mit der Kettenregel

$$\begin{aligned} \forall t \in [0; T] : \quad 0 &= \mathrm{D}_{\ell \circ z}(t) = \mathrm{D}_\ell(z(t)) \cdot \mathrm{D}_z(t) \\ &= (1 + \ln z_1(t), \ldots, 1 + \ln z_k(t)) \begin{pmatrix} z_1'(t) \\ \vdots \\ z_k'(t) \end{pmatrix}. \end{aligned} \tag{10.2.30}$$

Somit gilt

$$\forall t \in [0; T] : \quad z'(t) \perp \begin{pmatrix} 1 + \ln z_1(t) \\ \vdots \\ 1 + \ln z_k(t) \end{pmatrix}. \tag{10.2.31}$$

Da $z(t) \in \mathcal{A} = \mu + \mathbb{R}c^{(1)} + \mathbb{R}c^{(2)}$ für alle $t \in [0; T]$ gilt, folgt außerdem

$$\forall t \in [0; T] : \quad z'(t) \in \operatorname{span}\{c^{(1)}, c^{(2)}\} \text{ und damit} \tag{10.2.32}$$

$$\forall t \in [0; T] \quad \forall i = 3, \ldots, k : \quad z'(t) \perp c^{(i)}. \tag{10.2.33}$$

Sei $\widetilde{R} := \operatorname{diag}(1, 1, 0, \ldots, 0) \in \mathbb{R}^{k \times k}$ und sei $R^* := \operatorname{diag}(0, 0, 1, \ldots, 1) \in \mathbb{R}^{k \times k}$. Dann gilt

$$\forall x \in \mathbb{R}^k : \quad x = C^T \widetilde{R} C x + C^T R^* C x, \text{ wobei} \tag{10.2.34}$$

$$\forall x \in \mathbb{R}^k : \quad C^T \widetilde{R} C x \in \operatorname{span}\{c^{(1)}, c^{(2)}\} \text{ und} \tag{10.2.35}$$

$$\forall x \in \mathbb{R}^k : \quad C^T R^* C x \in \operatorname{span}\{c^{(3)}, \ldots, c^{(k)}\}. \tag{10.2.36}$$

Aus den Gleichungen (10.2.33) und (10.2.36) folgt die Aussage

$$\forall t \in [0;T]: \quad z'(t) \perp C^T R^* C \begin{pmatrix} 1+\ln z_1(t) \\ \vdots \\ 1+\ln z_k(t) \end{pmatrix}. \tag{10.2.37}$$

Daraus folgt mit den Gleichungen (10.2.31) und (10.2.34) schließlich

$$\forall t \in [0;T]: \quad z'(t) \perp C^T \widetilde{R} C \begin{pmatrix} 1+\ln z_1(t) \\ \vdots \\ 1+\ln z_k(t) \end{pmatrix} = C^T \widetilde{R} C \ln\{z(t)\}. \tag{10.2.38}$$

Als nächstes zeigen wir, dass der Vektor $C^T \widetilde{R} C \ln\{z(t)\}$ stets ungleich 0 ist. Dazu nehmen wir an, dass ein $t \in [0;T]$ existiert, sodass gilt

$$C^T \widetilde{R} C \ln\{z(t)\} = 0. \tag{10.2.39}$$

Dann folgt daraus mit den Gleichungen (10.2.34) und (10.2.36) die Aussage

$$\ln\{z(t)\} = C^T R^* C \ln\{z(t)\} \in \operatorname{span}\{c^{(3)}, \dots, c^{(k)}\}. \tag{10.2.40}$$

Daraus folgt wiederum die Existenz eines Vektors $\tilde{\lambda} = (\tilde{\lambda}_3, \dots, \tilde{\lambda}_k)^T \in \mathbb{R}^{k-2}$, für den gilt

$$\ln\{z(t)\} = \tilde{\lambda}_3 c^{(3)} + \dots + \tilde{\lambda}_k c^{(k)} = C_{3:k}^T \tilde{\lambda}. \tag{10.2.41}$$

Folglich gilt

$$z(t) = \exp\left\{C_{3:k}^T \tilde{\lambda}\right\}. \tag{10.2.42}$$

Zugleich gilt aber

$$z(t) \in \mathcal{A} = \{x \in \mathbb{R}^k \mid C_{3:k} x = C_{3:k} \mu\}. \tag{10.2.43}$$

Wir definieren die Funktion $f : \mathbb{R}^{k-2} \to \mathbb{R}^{k-2}$ gemäß

$$f(\lambda) := C_{3:k} \exp\left\{C_{3:k}^T \lambda\right\}. \tag{10.2.44}$$

Dann folgt aus den Gleichungen (10.2.42) und (10.2.43) die Aussage

$$\begin{aligned} f\left(\tilde{\lambda}\right) &= C_{3:k} \exp\left\{C_{3:k}^T \tilde{\lambda}\right\} \\ &= C_{3:k} z(t) = C_{3:k} \mu = C_{3:k} \exp\left\{C_{3:k}^T \hat{\lambda}\right\} = f\left(\hat{\lambda}\right), \text{ wobei} \end{aligned} \tag{10.2.45}$$

$$\hat{\lambda} := (0, \dots, 0, -\sqrt{k} \ln k)^T \in \mathbb{R}^{k-2} \text{ gilt.} \tag{10.2.46}$$

Außerdem ist die Funktion f überall stetig differenzierbar und es gilt

$$\forall \lambda \in \mathbb{R}^{k-2}: \quad \mathrm{D}_f(\lambda) = C_{3:k} \operatorname{diag}\left(\exp\left\{C_{3:k}^T \lambda\right\}\right) C_{3:k}^T. \tag{10.2.47}$$

Da die Jacobi-Matrix $\mathrm{D}_f(\lambda)$ für alle $\lambda \in \mathbb{R}^{k-2}$ positiv definit ist, ist die Funktion $f : \mathbb{R}^{k-2} \to \mathbb{R}^{k-2}$ injektiv. Daraus folgt mit Gleichung (10.2.45) die Aussage $\tilde{\lambda} = \hat{\lambda}$ und somit schließlich

$$z(t) = \exp\left\{C_{3:k}^T \tilde{\lambda}\right\} = \exp\left\{C_{3:k}^T \hat{\lambda}\right\} = \mu. \tag{10.2.48}$$

Dies steht im Widerspruch zu
$$\mathrm{I}(z(t)|\mu) = \rho > 0. \tag{10.2.49}$$
Also gilt
$$\forall t \in [0;T]: \quad C^T \widetilde{R} C \ln\{z(t)\} \neq 0. \tag{10.2.50}$$

Zusammenfassend gilt also
$$\forall t \in [0;T]: \quad z'(t) \perp C^T \widetilde{R} C \ln\{z(t)\} \in \operatorname{span}\{c^{(1)}, c^{(2)}\} \setminus \{0\}. \tag{10.2.51}$$

Da nach Gleichung (10.2.32) außerdem $z'(t) \in \operatorname{span}\{c^{(1)}, c^{(2)}\}$ für alle $t \in [0;T]$ gilt, folgt die Aussage
$$\forall t \in [0;T]: \quad z'(t) \in \mathbb{R} \cdot C^T R C \ln\{z(t)\}. \tag{10.2.52}$$
Dabei ist die Matrix $R \in \mathbb{R}^{k\times k}$ gemäß Gleichung (10.2.16) definiert. Da die Kurve $z : [0;T] \to (0;\infty)^k$ nach Bogenlänge parametrisiert ist, gilt $\|z'(t)\|_2 = 1$ für alle $t \in [0;T]$. Das Vorzeichen von $z'(t)$ folgt aus der Umlaufrichtung von z um den Punkt μ. Folglich gilt Gleichung (10.2.15). □

Ebenfalls von Interesse ist die Minimierung der I-Divergenz auf den Tangenten der Kurve $z : [0;T] \to (0;\infty)^k$.

Lemma 10.2.7 (I-Divergenz auf Tangenten von I_1-Sphären)**.** *Sei $z : [0;T] \to (0;\infty)^k$ die Kurve aus Lemma 10.2.6 und sei $t \in [0;T]$ beliebig. Dann gelten die folgenden beiden Aussagen:*

(i) Die Tangente $z(t) + \mathbb{R}z'(t)$ schneidet das abgeschlossene Wahrscheinlichkeitssimplex $\overline{\mathcal{H}}$ in einer Strecke,
$$\exists \alpha_{\min} < 0 \quad \exists \alpha_{\max} > 0: \quad (z(t) + \mathbb{R}z'(t)) \cap \overline{\mathcal{H}} = z(t) + [\alpha_{\min};\alpha_{\max}]\, z'(t). \tag{10.2.53}$$

(ii) Auf der Strecke $z(t) + [\alpha_{\min};\alpha_{\max}]z'(t)$ nimmt die Funktion $\mathrm{I}(\cdot|\mu)$ ihr Minimum im Punkt $z(t)$ an,
$$\forall \alpha \in [\alpha_{\min};\alpha_{\max}] \setminus \{0\}: \quad \mathrm{I}(z(t) + \alpha z'(t)|\mu) > \mathrm{I}(z(t)|\mu). \tag{10.2.54}$$

Beweis. (i) Die Gerade $z(t) + \mathbb{R}z'(t)$ erfüllt die Voraussetzungen von Lemma 10.2.2. Somit folgt Aussage (i) aus Lemma 10.2.2 (i).

(ii) Die Abbildung $\ell : (0;\infty)^k \to \mathbb{R}$ sei definiert gemäß
$$\ell(x) := \sum_i x_i \ln x_i = \mathrm{I}(x|\mu) + \left(\sum_i x_i\right) \ln \frac{1}{k}. \tag{10.2.55}$$

Für den Gradienten von ℓ gilt dann
$$\forall x \in (0;\infty)^k: \quad \nabla \ell(x) = \begin{pmatrix} 1 + \ln x_1 \\ \vdots \\ 1 + \ln x_k \end{pmatrix}. \tag{10.2.56}$$

Aufgrund der strikten Konvexität von ℓ folgt daraus

$$\forall y \in (0;\infty)^k \setminus \{z(t)\}: \quad \ell(y) > \ell(z(t)) + \langle \nabla\ell(z(t)), y - z(t)\rangle. \tag{10.2.57}$$

Daraus folgt wiederum

$$\begin{aligned}\forall \alpha \in (\alpha_{\min};\alpha_{\max}) \setminus \{0\}: \quad \ell(z(t)+\alpha z'(t)) &> \ell(z(t)) + \alpha\langle \nabla\ell(z(t)), z'(t)\rangle \\ &= \ell(z(t)).\end{aligned} \tag{10.2.58}$$

Die letzte Gleichheitsaussage in der Ungleichungskette (10.2.58) folgt aus der Tatsache, dass die beiden Vektoren $z'(t)$ und $\nabla\ell(z(t)) = (1+\ln z_1(t), \ldots, 1+\ln z_k(t))^T$, wie in Gleichung (10.2.31) gezeigt, senkrecht aufeinander stehen. Aus der Ungleichung (10.2.58) folgt schließlich die zu zeigende Ungleichung (10.2.54). □

10.3. Unmöglichkeitssatz

In der Definition der variationellen Charakterisierungen $\Phi : \overline{\mathcal{H}}^\eta \to \mathbb{R}$ der Zyklen (siehe Abschnitt 10.1) tritt die I-Divergenz nur indirekt auf in Gestalt der Menge $\operatorname{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta)$. Die Definition dieser Menge basiert wiederum auf den I_1-Projektionen. Nichtsdestotrotz besteht ein enger Zusammenhang zwischen der Funktion Φ und der I-Divergenz. So müssen alle Funktionen $\Phi(p^{(1)}, \ldots, p^{(\eta-1)}, \cdot)$ mit $p^{(1)} \equiv \cdots \equiv p^{(\eta-1)}$ über jeder abgeschlossenen, konvexen Menge $\mathcal{E}_\eta \subseteq \overline{\mathcal{H}}$ mit mindestens einem Element $p^{(\eta)}$, das $p^{(\eta)} \equiv p^{(1)}$ erfüllt, ihr Minimum in der I_1-Projektion von $p^{(\eta-1)}$ auf $\mathcal{E}_\eta$ annehmen, um die Menge $\operatorname{cyc}(\{p^{(1)}\}, \ldots, \{p^{(\eta-1)}\}, \mathcal{E}_\eta)$ korrekt zu charakterisieren. Für den Fall $p^{(\eta-1)} = \mu$ zeigt das folgende Lemma, dass der Wert der Funktion $\Phi(p^{(1)}, \ldots, p^{(\eta-1)}, \cdot)$ im Punkt $p^{(\eta)}$ im Wesentlichen von der I-Divergenz $\mathrm{I}(p^{(\eta)}|\mu)$ abhängt.

Lemma 10.3.1. *Sei $\overline{\mathcal{H}}$ das abgeschlossene Wahrscheinlichkeitssimplex im $\mathbb{R}^k$, wobei $k \geq 3$ gilt, und sei $\mu \in \mathcal{H}$ die Gleichverteilung. Weiter sei $\varphi : \overline{\mathcal{H}} \to \mathbb{R}$ eine Funktion, die über jeder abgeschlossenen, konvexen Menge $\mathcal{E} \subseteq \overline{\mathcal{H}}$, die $\mathcal{E} \cap \mathcal{H} \neq \emptyset$ erfüllt, ihr Infimum im Punkt $\mathrm{P}_\mathcal{E}\,\mu$ annimmt. Dabei bezeichne $\mathrm{P}_\mathcal{E}\,\mu$ die I_1-Projektion von μ auf $\mathcal{E}$. Dann gelten die folgenden drei Aussagen:*

(i) Die Funktion φ ist auf der I_1-Kugel $\mathcal{B}(\mu, \ln\frac{k}{k-1})$ entropisch wachsend,

$$\forall x, y \in \mathcal{B}\left(\mu, \ln\frac{k}{k-1}\right): \quad \mathrm{I}(x|\mu) < \mathrm{I}(y|\mu) \Rightarrow \varphi(x) \leq \varphi(y). \tag{10.3.1}$$

(ii) Falls $\mathrm{P}_\mathcal{E}\,\mu$ für alle abgeschlossenen, konvexen Mengen $\mathcal{E} \subseteq \overline{\mathcal{H}}$, die $\mathcal{E} \cap \mathcal{H} \neq \emptyset$ erfüllen, der einzige Punkt ist, in dem die Funktion φ ihr Minimum annimmt, dann ist φ auf $\mathcal{B}(\mu, \ln\frac{k}{k-1})$ streng entropisch wachsend,

$$\forall x, y \in \mathcal{B}\left(\mu, \ln\frac{k}{k-1}\right): \quad \mathrm{I}(x|\mu) < \mathrm{I}(y|\mu) \Rightarrow \varphi(x) < \varphi(y). \tag{10.3.2}$$

(iii) Die Funktion φ ist auf fast allen I_1-Sphären $\mathcal{S}(\mu, \rho)$, $\rho \in [0; \ln\frac{k}{k-1}]$, konstant.

Beweis. (i) Seien $x, y \in \mathcal{B}(\mu, \ln \frac{k}{k-1})$ und gelte $\mathrm{I}(x|\mu) < \mathrm{I}(y|\mu)$. Nach Voraussetzung nimmt die Funktion $\varphi : \overline{\mathcal{H}} \to \mathbb{R}$ ihr Infimum im Punkt $\mathrm{P}_{\overline{\mathcal{H}}}\, \mu = \mu$ an. Im Fall $x = \mu$ gilt folglich $\varphi(x) = \varphi(\mu) \leq \varphi(y)$. Für den Fall $x \neq \mu$ erfolgt der Beweis der Aussage $\varphi(x) \leq \varphi(y)$ in fünf Schritten.

I. Sei $\mathcal{A}$ ein zweidimensionaler affiner Unterraum von $\mathrm{aff}(\mathcal{H})$, der die Punkte μ, x und y enthält, und sei $C \in \mathbb{R}^{k \times k}$ die zugehörige Matrix im Sinne von Abschnitt 10.2. Weiter sei $\rho := \mathrm{I}(x|\mu) \in (0; \ln \frac{k}{k-1})$. Dann existiert nach Lemma 10.2.6 eine geschlossene und nach Bogenlänge parametrisierte Kurve $z : [0; T] \to (0; \infty)^k$ mit $z(0) = z(T) = x$ und

$$\mathrm{Bild}\, z = \mathcal{A} \cap \mathcal{S}(\mu, \rho). \tag{10.3.3}$$

Die Kurve $z : [0; T] \to (0; \infty)^k$ ist außerdem stetig differenzierbar mit

$$\forall t \in [0; T] : \quad z'(t) = \left\| C^T R C \ln \{z(t)\} \right\|_2^{-1} \cdot C^T R C \ln \{z(t)\}\,. \tag{10.3.4}$$

Dabei ist die Matrix $R \in \mathbb{R}^{k \times k}$ gemäß Gleichung (10.2.16) definiert.

Folglich löst die Kurve $z : [0; T] \to (0; \infty)^k$ das autonome Anfangswertproblem

$$\begin{aligned} &z'(t) = f(z(t)) \text{ und } z(0) = x, \text{ wobei} \\ &f(z) := \left\| CRC^T \ln \{z\} \right\|_2^{-1} \cdot CRC^T \ln \{z\} \text{ gilt.} \end{aligned} \tag{10.3.5}$$

Die so definierte Funktion f ist eine Verkettung der Funktionen $g(z) := CRC^T \ln \{z\}$ und $h(z) := \|z\|_2^{-1} \cdot z$. Wie im Beweis von Lemma 10.2.6 (ii) gezeigt, gilt $g(z) = CRC^T \ln \{z\} \neq 0$ für alle $z \in \mathcal{H} \cap \mathcal{A} \setminus \{\mu\}$. Folglich ist $f(z) = h(g(z))$ für alle $z \in \mathcal{H} \cap \mathcal{A} \setminus \{\mu\}$ wohldefiniert. Da die Funktionen g und h auf ihrem gesamten Definitionsbereich stetig differenzierbar sind, ist nach der Kettenregel auch die Funktion $f : \mathcal{H} \cap \mathcal{A} \setminus \{\mu\} \to \mathcal{A} - \mu$ stetig differenzierbar. Folglich ist f lokal Lipschitz-stetig. Nach dem globalen Existenz- und Eindeutigkeitssatz (vgl. Aulbach 2004, Satz 2.4.1) ist die Kurve $z : [0; T] \to (0; \infty)^k$ damit die eindeutige Lösung des Anfangswertproblems (10.3.5).

II. Als eindeutige Lösung des Anfangswertproblems (10.3.5) kann die Kurve $z : [0; T] \to (0; \infty)^k$ mithilfe des expliziten Euler-Verfahrens beliebig genau durch einen Polygonzug angenähert werden (vgl. Strehmel, Weiner und Podhaisky 2012, Kapitel 2). Zuvor „verlängern“ wir jedoch die Kurve $z : [0; T] \to (0; \infty)^k$, indem wir jedem Zeitpunkt $t \in (T; 2T]$ den Punkt $z(t) := z(t - T) \in (0; \infty)^k$ zuordnen. Auf diese Weise erhalten wir eine geschlossene, reguläre Kurve $z : [0; 2T] \to (0; \infty)^k$ mit $z(0) = z(T) = z(2T) = x$, die den Punkt μ mit konstanter Geschwindigkeit genau zweimal umläuft.

Nun wenden wir das explizite Euler-Verfahren an. Dazu zerlegen wir das Intervall $[0; 2T]$ in die Teilintervalle $[t_n; t_{n+1}]$, $n = 0, \ldots, N-1$, wobei gilt

$$0 = t_0 < t_1 < t_2 < \cdots < t_N = 2T. \tag{10.3.6}$$

Die Werte $h_n := t_{n+1} - t_n$, $n = 0, \ldots, N-1$, bezeichnen wir im Folgenden als die *Schrittweiten*, den Wert $h_{\max} := \max_{n=0,\ldots,N-1} h_n$ als die *maximale Schrittweite*. Für jeden Zeitpunkt t_n, $n = 0, \ldots, N$, bestimmen wir eine Näherungslösung $u_n \approx z(t_n)$ gemäß der Vorschrift

$$u_0 := x \text{ und} \tag{10.3.7}$$

$$u_{n+1} := u_n + h_n f(u_n) \text{ für alle } n = 0, \ldots, N-1. \tag{10.3.8}$$

Indem wir die resultierenden Punkte $u_0, \ldots, u_N$ durch Strecken verbinden, erhalten wir einen Polygonzug $z^* : [0; 2T] \to (0; \infty)^k$ mit den Eckpunkten $z^*(t_n) = u_n \approx z(t_n)$, $n = 0, \ldots, N$.

Dieser Polygonzug $z^* : [0; 2T] \to (0; \infty)^k$ konvergiert in dem folgenden Sinn gegen die Kurve $z : [0; 2T] \to (0; \infty)^k$: Geht die maximale Schrittweite $h_{\max}$ gegen 0, so geht auch der maximale Fehler $\max_{n=0,\ldots,N} \|z^*(t_n) - z(t_n)\|_2$ gegen 0. (Vgl. Strehmel, Weiner und Podhaisky 2012, Folgerung 2.2.2)

III. Nach der im vorangegangen Schritt gewählten Konstruktion gilt $u_0 = x = z(0)$ und $u_1 = u_0 + h_0 f(u_0) = z(0) + h_0 z'(0)$. Somit gilt gemäß Lemma 10.2.7 die Ungleichung $\mathrm{I}(u_0|\mu) < \mathrm{I}(u_1|\mu)$. Diese Argumentation kann analog fortgesetzt werden. Auf diese Weise erhalten wir die Ungleichungskette

$$\mathrm{I}(x|\mu) = \mathrm{I}(u_0|\mu) < \mathrm{I}(u_1|\mu) < \cdots < \mathrm{I}(u_{N-1}|\mu) < \mathrm{I}(u_N|\mu). \tag{10.3.9}$$

Der Abstand $\|u_N - x\|_2$ ist jedoch nach oben durch den maximalen Fehler beschränkt,

$$\|u_N - x\|_2 = \|z^*(t_N) - z(t_N)\|_2 \leq \max_{n=0,\ldots,N} \|z^*(t_n) - z(t_n)\|_2 . \tag{10.3.10}$$

Aufgrund der Stetigkeit der Funktion $\mathrm{I}(\cdot|\mu)$ (siehe Lemma 10.2.1 (ii)) können wir folglich die maximale Schrittweite $h_{\max}$ und damit auch die Schrittzahl N so wählen, dass gilt

$$\mathrm{I}(x|\mu) = \mathrm{I}(u_0|\mu) < \mathrm{I}(u_1|\mu) < \cdots < \mathrm{I}(u_{N-1}|\mu) < \mathrm{I}(u_N|\mu) < \mathrm{I}(y|\mu). \tag{10.3.11}$$

IV. Da die Kurve $z : [0; 2T] \to (0; \infty)^k$ den Punkt μ genau zweimal vollständig umläuft, umläuft der so konstruierte Polygonzug $z^* : [0; 2T] \to (0; \infty)^k$ den Punkt μ mindestens einmal vollständig. Also existiert ein minimales $\tau \in [0; 2T]$, sodass $\measuredangle(x, \mu, z^*(\tau)) = \measuredangle(x, \mu, y)$ gilt. Des Weiteren existiert ein $n \in \{1, \ldots, N\}$, sodass $t_{n-1} < \tau \leq t_n$ gilt. Falls $\tau < t_n$ gilt, ersetzen wir den Wert t_n durch τ und den Punkt u_n durch $z^*(\tau)$. Auf diese Weise erhalten wir einen Polygonzug $z^* : [0; \tau] \to (0; \infty)^k$, für dessen Eckpunkte $u_0, \ldots, u_n$ die Gleichungen $u_0 = x$ und $\measuredangle(x, \mu, u_n) = \measuredangle(x, \mu, y)$ gelten. Außerdem gilt die Ungleichungskette

$$\mathrm{I}(x|\mu) = \mathrm{I}(u_0|\mu) < \mathrm{I}(u_1|\mu) < \cdots < \mathrm{I}(u_{n-1}|\mu) < \mathrm{I}(u_n|\mu) < \mathrm{I}(y|\mu). \tag{10.3.12}$$

V. Nach Voraussetzung nimmt die Funktion $\varphi : \overline{\mathcal{H}} \to \mathbb{R}$ über jeder abgeschlossenen, konvexen Menge $\mathcal{E} \subseteq \overline{\mathcal{H}}$, die $\mathcal{E} \cap \mathcal{H} \neq \emptyset$ erfüllt, ihr Infimum im Punkt $\mathrm{P}_{\mathcal{E}}\,\mu$ an. Die I_1-Projektion $\mathrm{P}_{\mathcal{E}_0}\,\mu$ von μ auf die Strecke $\mathcal{E}_0 := [u_0; u_1] = \{z(0) + \alpha z'(0) \,|\, \alpha \in [0; h_0]\}$ ist nach Lemma 10.2.7 gleich dem Punkt $z(0) = u_0$. Da die Strecke $\mathcal{E}_0$ außerdem eine abgeschlossene und konvexe Teilmenge von $\overline{\mathcal{H}}$ ist und $\mathcal{E}_0 \cap \mathcal{H} \neq \emptyset$ gilt, nimmt die Funktion $\varphi : \overline{\mathcal{H}} \to \mathbb{R}$ ihr Infimum über der Strecke $\mathcal{E}_0$ im Punkt $\mathrm{P}_{\mathcal{E}_0}\,\mu = u_0$

an. Somit gilt $\varphi(x) = \varphi(u_0) \leq \varphi(u_1)$. Durch analoge Argumentation erhalten wir die Ungleichungskette

$$\varphi(x) = \varphi(u_0) \leq \varphi(u_1) \leq \cdots \leq \varphi(u_n). \tag{10.3.13}$$

Als letztes betrachten wir die Strecke $\mathcal{E}_n := [u_n; y] \subsetneq [\mu; y]$. Da die I-Divergenz $\mathrm{I}(\cdot|\mu)$ auf $\mathcal{E}_n$ gemäß Lemma 10.2.2 (vii) streng monoton steigend ist, gilt $\mathrm{P}_{\mathcal{E}_n}\,\mu = u_n$. Folglich gilt $\varphi(u_n) \leq \varphi(y)$ und somit schließlich

$$\varphi(x) = \varphi(u_0) \leq \varphi(u_1) \leq \cdots \leq \varphi(u_n) \leq \varphi(y). \tag{10.3.14}$$

Damit ist Aussage (i) gezeigt.

(ii) Falls $\mathrm{P}_{\mathcal{E}}\,\mu$ für alle abgeschlossenen, konvexen Mengen $\mathcal{E} \subseteq \overline{\mathcal{H}}$, die $\mathcal{E} \cap \mathcal{H} \neq \emptyset$ erfüllen, der einzige Punkt ist, in dem die Funktion φ ihr Minimum annimmt, dann folgt

$$\varphi(x) = \varphi(u_0) < \varphi(u_1) < \cdots < \varphi(u_n) < \varphi(y). \tag{10.3.15}$$

(iii) Nach Voraussetzung nimmt die Funktion $\varphi : \overline{\mathcal{H}} \to \mathbb{R}$ ihr Infimum im Punkt $\mathrm{P}_{\overline{\mathcal{H}}}\,\mu = \mu$ an. Folglich ist φ nach unten durch $\varphi(\mu) \in \mathbb{R}$ beschränkt. Um zu zeigen, dass φ auf fast allen I_1-Sphären $\mathcal{S}(\mu,\rho)$, $\rho \in [0; \ln\frac{k}{k-1}]$, konstant ist, definieren wir die beiden Funktionen $a : [0; \ln\frac{k}{k-1}] \to \mathbb{R}$ und $b : [0; \ln\frac{k}{k-1}] \to \mathbb{R} \cup \{+\infty\}$ gemäß

$$a(\rho) := \inf_{z \in \mathcal{S}(\mu,\rho)} \varphi(z) \qquad \text{und} \qquad b(\rho) := \sup_{z \in \mathcal{S}(\mu,\rho)} \varphi(z). \tag{10.3.16}$$

Nach Aussage (i) gilt dann

$$\forall \rho, \sigma \in \left[0; \ln\frac{k}{k-1}\right] : \quad \rho < \sigma \Rightarrow a(\rho) \leq b(\rho) \leq a(\sigma) \leq b(\sigma). \tag{10.3.17}$$

Also sind die Funktionen a und b monoton steigend. Folglich besitzt a nach dem Satz von Froda in jedem Punkt $\rho \in (0; \ln\frac{k}{k-1})$ einen linksseitigen Grenzwert $a(\rho^-)$ und einen rechtsseitigen Grenzwert $a(\rho^+)$, für die $a(\rho^-) \leq a(\rho) \leq a(\rho^+)$ gilt. Ebenso besitzt b in jedem Punkt $\rho \in (0; \ln\frac{k}{k-1})$ einen linksseitigen Grenzwert $b(\rho^-)$ und einen rechtsseitigen Grenzwert $b(\rho^+)$, für die $b(\rho^-) \leq b(\rho) \leq b(\rho^+)$ gilt. Aus der Ungleichungskette (10.3.17) folgt außerdem $a(\rho^-) = b(\rho^-)$ sowie $a(\rho^+) = b(\rho^+)$ für alle $\rho \in (0; \ln\frac{k}{k-1})$ und damit die Aussage

$$\forall \rho \in \left(0; \ln\frac{k}{k-1}\right) : \quad a(\rho^-) = b(\rho^-) \leq a(\rho) \leq b(\rho) \leq a(\rho^+) = b(\rho^+). \tag{10.3.18}$$

Der Satz von Froda besagt außerdem, dass die Menge aller Punkte $\rho \in (0; \ln\frac{k}{k-1})$ mit $a(\rho^-) < a(\rho^+)$ höchstens abzählbar ist. Folglich gilt die Gleichung $a(\rho) = b(\rho)$ für fast alle $\rho \in [0; \ln\frac{k}{k-1}]$. Also ist die Funktion φ auf fast allen I_1-Sphären $\mathcal{S}(\mu,\rho)$, $\rho \in [0; \ln\frac{k}{k-1}]$, konstant. □

Das soeben gezeigte Lemma 10.3.1 ermöglicht den Beweis der Aussage, dass die Zyklen von I_1-Projektionen auf mehr als zwei konvexe Mengen nicht in der gewünschten Weise charakterisiert werden können.

Satz 10.3.2 (Unmöglichkeit der variationellen Charakterisierung der Limeszyklen des mehrdimensionalen IPF-Verfahrens)**.** *Sei $\overline{\mathcal{H}}$ das abgeschlossene Wahrscheinlichkeitssimplex im $\mathbb{R}^k$, wobei $k \geq 3$ gilt. Weiter sei $\eta \geq 3$. Dann existiert keine Funktion $\Phi : \overline{\mathcal{H}}^\eta \to \mathbb{R}$, sodass für alle geordneten Familien $(\mathcal{E}_1, \ldots, \mathcal{E}_\eta)$ von abgeschlossenen, konvexen Teilmengen von $\overline{\mathcal{H}}$ mit mindestens einem Tupel $(p^{(1)}, \ldots, p^{(\eta)}) \in \mathcal{E}_1 \times \cdots \times \mathcal{E}_\eta$, das $p^{(1)} \equiv \cdots \equiv p^{(\eta)}$ erfüllt, gilt*

$$\operatorname{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta) = \operatorname*{arg\,min}_{p^{(1)} \in \mathcal{E}_1, \ldots, p^{(\eta)} \in \mathcal{E}_\eta} \Phi\left(p^{(1)}, \ldots, p^{(\eta)}\right). \tag{10.3.19}$$

Beweis. Wir nehmen an, dass eine Funktion $\Phi : \overline{\mathcal{H}}^\eta \to \mathbb{R}$ mit der gewünschten Eigenschaft existiert. Diese Annahme führen wir in zwei Schritten zum Widerspruch. Dazu zeigen wir in Schritt I, dass die Funktionen $\Phi(z, \mu, \ldots, \mu, \cdot)$, $z \in \mathcal{H}$, auf fast allen I_1-Sphären $\mathcal{S}(\mu, \rho)$, $\rho \in [0; \ln \frac{k}{k-1}]$, konstant sind. Anschließend wählen wir in Schritt II die Punkte $x^{(\alpha_i)}, x^{(\beta_i)} \in \overline{\mathcal{H}}$, $i = 1, 2$, entlang einer Geraden durch den Punkt μ und folgern schließlich die Ungleichung $\Phi(x^{(\alpha_1)}, \mu, \ldots, \mu, x^{(\beta_1)}) < \Phi(x^{(\alpha_1)}, \mu, \ldots, \mu, x^{(\beta_1)})$.

I. Sei $z \in \mathcal{H}$ beliebig. Dann wählen wir $\mathcal{E}_1 := \{z\}$ und $\mathcal{E}_2 := \cdots := \mathcal{E}_{\eta-1} := \{\mu\}$. Folglich gilt für jede abgeschlossene, konvexe Menge $\mathcal{E}_\eta \subseteq \overline{\mathcal{H}}$, die $\mathcal{E}_\eta \cap \mathcal{H} \neq \emptyset$ erfüllt, die Aussage

$$\begin{aligned} \operatorname*{arg\,min}_{p^{(1)} \in \mathcal{E}_1, \ldots, p^{(\eta)} \in \mathcal{E}_\eta} \Phi\left(p^{(1)}, \ldots, p^{(\eta)}\right) &= \operatorname{cyc}(\mathcal{E}_1, \ldots, \mathcal{E}_\eta) \\ &= \left\{(z, \mu, \ldots, \mu, \mathrm{P}_{\mathcal{E}_\eta} \mu)\right\}. \end{aligned} \tag{10.3.20}$$

Also nimmt die Funktion $\varphi := \Phi(z, \mu, \ldots, \mu, \cdot)$ auf allen abgeschlossenen, konvexen Mengen $\mathcal{E} \subseteq \overline{\mathcal{H}}$ ihr Infimum im Punkt $\mathrm{P}_\mathcal{E}\, \mu$ an. Nach Lemma 10.3.1 (iii) ist die Funktion $\varphi = \Phi(z, \mu, \ldots, \mu, \cdot)$ somit auf fast allen I_1-Sphären $\mathcal{S}(\mu, \rho)$, $\rho \in [0; \ln \frac{k}{k-1}]$, konstant.

II. Als nächstes betrachten wir die Strecke $\{x^{(\alpha)} \mid \alpha \in [\alpha_{\min}; \alpha_{\max}]\}$. Dabei wählen wir

$$\forall \alpha \in [\alpha_{\min}; \alpha_{\max}] : \quad x^{(\alpha)} := \mu + \alpha(n-1, -1, \ldots, -1)^T. \tag{10.3.21}$$

Außerdem wählen wir $\alpha_{\min} < 0$ und $\alpha_{\max} > 0$ so, dass gilt

$$\mathrm{I}(x^{(\alpha_{\min})}|\mu) = \mathrm{I}(x^{(\alpha_{\max})}|\mu) = \ln \frac{k}{k-1}. \tag{10.3.22}$$

Gemäß Lemma 10.2.2 ist die Funktion $\alpha \mapsto \mathrm{I}(x^{(\alpha)}|\mu)$ auf dem Intervall $[\alpha_{\min}; \alpha_{\max}]$ stetig und außerdem streng monoton fallend auf dem Teilintervall $[\alpha_{\min}; 0]$ und streng monoton steigend auf dem Teilintervall $[0; \alpha_{\max}]$. Als nächstes wählen wir $\alpha_1 \in (\alpha_{\min}; 0)$ und $\alpha_2 \in (0; \alpha_{\max})$ so, dass gilt

$$\mathrm{I}(x^{(\alpha_1)}|\mu) = \mathrm{I}(x^{(\alpha_2)}|\mu). \tag{10.3.23}$$

Nach Schritt I sind die Funktionen $\Phi(x^{(\alpha_1)}, \mu, \ldots, \mu, \cdot)$ und $\Phi(x^{(\alpha_2)}, \mu, \ldots, \mu, \cdot)$ auf fast allen I_1-Sphären $\mathcal{S}(\mu, \rho)$, $\rho \in [0; \ln \frac{k}{k-1}]$, konstant. Wir wählen daher ein $\tilde{\rho} \in (\mathrm{I}(x^{(\alpha_1)}|\mu); \ln \frac{k}{k-1})$, sodass diese beiden Funktionen auf $\mathcal{S}(\mu, \tilde{\rho})$ konstant sind. Anschließend wählen wir $\beta_1 \in (\alpha_{\min}; \alpha_1)$ und $\beta_2 \in (\alpha_2; \alpha_{\max})$, sodass gilt

$$\mathrm{I}(x^{(\beta_1)}|\mu) = \mathrm{I}(x^{(\beta_2)}|\mu) = \tilde{\rho}. \tag{10.3.24}$$

Folglich gilt

$$\alpha_{\min} < \beta_1 < \alpha_1 < 0 < \alpha_2 < \beta_2 < \alpha_{\max}. \tag{10.3.25}$$

Außerdem gelten nach Lemma 10.2.2 die beiden Gleichungen

$$\operatorname{cyc}\left(\left[x^{(\alpha_1)}; x^{(\alpha_2)}\right], \{\mu\}, \ldots, \{\mu\}, \left\{x^{(\beta_1)}\right\}\right) = \left\{\left(x^{(\alpha_1)}, \mu, \ldots, \mu, x^{(\beta_1)}\right)\right\} \text{ und} \tag{10.3.26}$$

$$\operatorname{cyc}\left(\left[x^{(\alpha_1)}; x^{(\alpha_2)}\right], \{\mu\}, \ldots, \{\mu\}, \left\{x^{(\beta_2)}\right\}\right) = \left\{\left(x^{(\alpha_2)}, \mu, \ldots, \mu, x^{(\beta_2)}\right)\right\}. \tag{10.3.27}$$

Daraus folgt mit der Annahme und der Konstanz der Funktionen $\Phi(x^{(\alpha_1)}, \mu, \ldots, \mu, \cdot)$ und $\Phi(x^{(\alpha_2)}, \mu, \ldots, \mu, \cdot)$ auf $\mathcal{S}(\mu, \tilde{\rho})$ der Widerspruch

$$\begin{aligned}
\Phi\left(x^{(\alpha_1)}, \mu, \ldots, \mu, x^{(\beta_1)}\right) &< \Phi\left(x^{(\alpha_2)}, \mu, \ldots, \mu, x^{(\beta_1)}\right) \\
&= \Phi\left(x^{(\alpha_2)}, \mu, \ldots, \mu, x^{(\beta_2)}\right) \\
&< \Phi\left(x^{(\alpha_1)}, \mu, \ldots, \mu, x^{(\beta_2)}\right) \\
&= \Phi\left(x^{(\alpha_1)}, \mu, \ldots, \mu, x^{(\beta_1)}\right).
\end{aligned} \tag{10.3.28}$$

Also existiert keine Funktion $\Phi : \overline{\mathcal{H}}^{\eta} \to \mathbb{R}$ mit der gewünschten Eigenschaft. □

10.4. Weitere Ansätze zur Divergenzanalyse

Die soeben gezeigte Unmöglichkeit der variationellen Charakterisierung der Limeszyklen des mehrdimensionalen IPF-Verfahrens schließt einen einfachen Beweis für Hypothese 10.1.1 aus, welcher auf der schrittweisen Minimierung eines universellen, für alle Anpassungsprobleme $(A, M^{(1)}, \ldots, M^{(\eta)})$ gültigen Funktionals basiert. Auch andere naheliegende Ansätze scheiden aus. So erfüllen die I_1-Projektionen im Allgemeinen keine Kontraktionseigenschaft. Dies zeigt das folgende Beispiel.

Beispiel 10.4.1 (Nichtkontrahierende I_1-Projektionen). Gegeben seien die Zeilenmarginalien $r = (0.9, 0.1)$ sowie die Wahrscheinlichkeitsmatrizen

$$P := \begin{pmatrix} 0.09 & 0.01 \\ 0.45 & 0.45 \end{pmatrix} \qquad \text{und} \qquad \widetilde{P} := \begin{pmatrix} 0.01 & 0.09 \\ 0.45 & 0.45 \end{pmatrix}. \tag{10.4.1}$$

Des Weiteren sei $\mathcal{R}$ die zugehörige Menge von Wahrscheinlichkeitsmatrizen $R \in [0;1]^{2\times 2}$, die die Zeilenmarginalien r erfüllen. Dann ist die I_1-Projektion auf $\mathcal{R}$ äquivalent zur Zeilenanpassung an r (siehe Lemma 2.3.1 (ii), S. 30). Dementsprechend gelten für die I_1-Projektion $P(1)$ von P auf $\mathcal{R}$ und für die I_1-Projektion $\widetilde{P}(1)$ von $\widetilde{P}$ auf $\mathcal{R}$ die Gleichungen

$$P(1) = \begin{pmatrix} 0.81 & 0.09 \\ 0.05 & 0.05 \end{pmatrix} \qquad \text{und} \qquad \widetilde{P}(1) := \begin{pmatrix} 0.09 & 0.81 \\ 0.05 & 0.05 \end{pmatrix}. \tag{10.4.2}$$

Folglich gilt für alle L_p-Normen $\|\cdot\|_p$ mit $p \in [1; \infty]$ die Ungleichung

$$\left\|P - \widetilde{P}\right\|_p < \left\|P(1) - \widetilde{P}(1)\right\|_p. \tag{10.4.3}$$

Der Abstand der Matrizen P und $\widetilde{P}$ ist also unabhängig von der Wahl der L_p-Norm kleiner als der Abstand ihrer I_1-Projektionen $P(1)$ und $\widetilde{P}(1)$. Also sind I_1-Projektionen im Allgemeinen nichtkontrahierend. Beispiele für nichtkontrahierende I_1-Projektionen mehrdimensionaler Wahrscheinlichkeitstafeln können in analoger Weise konstruiert werden.

Auch die beiden in Kapitel 9 (siehe S. 87) vorgestellten Ansätze können nicht ohne Weiteres auf beliebige Anpassungsprobleme $(A, M^{(1)}, \ldots, M^{(\eta)})$ verallgemeinert werden. Die Probleme, die diese Verallgemeinerungen verhindern, wurden in Abschnitt 9.3 (siehe S. 104) diskutiert. Folglich bleibt Hypothese 10.1.1 bis auf Weiteres unbewiesen.

10.5. Kommentare und Referenzen

Hypothese 10.1.1 wurde für mehrdimensionale Anpassungsprobleme $(A, M^{(1)}, \ldots, M^{(\eta)})$ bestehend aus einer Gleichverteilung A und bestimmten Marginaltafeln $M^{(i)}$, $i = 1, \ldots, \eta$, bereits von Vomlel (1999, Hypothese 3.1) aufgestellt. Umfangreiche Simulationen legen jedoch nahe, dass diese Hypothese für beliebige mehrdimensionale Anpassungsprobleme gilt. Satz 10.1.2 ist bisher unveröffentlicht, basiert jedoch auf Satz 2.2.4, (siehe S. 28). Man beachte daher auch die Kommentare und Referenzen in Abschnitt 2.7 (siehe S. 41). Die in Abschnitt 10.1 eingeführte Bezeichnung „variationelle Charakterisierung" orientiert sich an Baillon, Combettes und Cominetti (2012). Streng genommen müsste man von einer „universellen variationellen Charakterisierung" sprechen, da die zugehörige Menge der Zyklen für alle Familien von Teilmengen durch dieselbe Funktion charakterisiert werden soll.

Die Einschränkung $k \geq 3$ zu Beginn von Abschnitt 10.2 ist keine wirkliche Einschränkung, da bereits die kleinstmöglichen dreidimensionalen Anpassungsprobleme über eine $2 \times 2 \times 2$-Ausgangstafel und somit über acht Einträge in der Ausgangstafel verfügen. Selbst wenn wir uns bei der Vektorisierung der Ausgangstafel auf ihre positiven Einträge beschränken, ist die Forderung $k \geq 3$ sinnvoll, da Ausgangstafeln mit nur zwei positiven Einträgen nur von geringem Interesse sind, wie in Abschnitt 9.2 (siehe S. 94) gezeigt wurde. Die in den Lemmata 10.2.1, 10.2.2 und 10.2.3 gezeigten Eigenschaften der I-Divergenz bei festem zweiten Argument folgen aus den in Abschnitt 2.1 (siehe S. 22) diskutierten grundlegenden Eigenschaften der I-Divergenz. Der in Definition 10.2.4 eingeführte Begriff der I_1-Sphäre geht auf Csiszár (1975, S. 146 f.) zurück. Im Gegensatz zu der hier gewählten Sprechweise definiert Csiszár jedoch die „I-sphere" als das Innere der I_1-Kugel. Die in Lemma 10.2.5 (i) gezeigte Abgeschlossenheit der I_1-Kugeln gilt sowohl bezüglich der Standardtopologie des $\mathbb{R}^k$ als auch bezüglich der Standardtopologie des affinen Unterraums $\mathrm{aff}(\mathcal{H})$.

Die in Abschnitt 10.3 vorgestellten Ergebnisse sind bisher unveröffentlicht, basieren jedoch auf der Argumentation von Baillon, Combettes und Cominetti (2012). Baillon, Combettes und Cominetti zeigen die Unmöglichkeit der variationellen Charakterisierung der Zyklen orthogonaler Projektionen auf mehr als zwei nichtleere, abgeschlossene, konvexe Mengen in einem beliebigen reellen Hilbertraum. Das hier gezeigte Lemma 10.3.1 ist ein Analogon zu Theorem 2.1 von Baillon, Combettes und Cominetti. Bei diesem Analogon treten die I_1-Projektionen an die Stelle von orthogonalen Projektionen. Der Beweis von Aussage (i) ist hier jedoch deutlich aufwendiger als bei Baillon, Combettes und Cominetti, da die Approximation der hier betrachteten Schnittmengen von I_1-Sphären und Ebenen durch Polygonzüge entschieden mühsamer ist als die Approximation der dort betrachteten Kreise. Die in den Aussagen (i) und (ii) verwendeten Begriffe „entropisch wachsend“ und „streng entropisch wachsend“ beziehen sich auf die „relative entropy“, eine alternative Bezeichnung für die I-Divergenz (siehe Abschnitt 2.7, S. 41). Die im Beweis von Aussage (iii) verwendete Bezeichnung „Satz von Froda“ orientiert sich an Baillon, Combettes und Cominetti. In vielen Standardwerken zur Analysis wird dieser Satz ohne Angabe eines Namens aufgeführt. Der hier gezeigte Satz 10.3.2 ist ein Analogon zu Theorem 2.3 von Baillon, Combettes und Cominetti. Der zugehörige Beweis erfolgt ebenfalls analog zum Beweis von Baillon, Combettes und Cominetti. Die Geometrie des Wahrscheinlichkeitssimplex und der I_1-Sphären erfordert jedoch einige Anpassungen.

11. Zusammenfassung der Ergebnisse zum mehrdimensionalen IPF-Verfahren

Das in Kapitel 5 ausgerufene Ziel, möglichst viele der in Teil I bewiesenen Aussagen über das zweidimensionale IPF-Verfahren auf das mehrdimensionale IPF-Verfahren zu verallgemeinern, wurde teilweise erreicht. Für manche Aussagen wurde stattdessen jedoch gezeigt, dass ihre Verallgemeinerung auf das mehrdimensionale IPF-Verfahren unmöglich oder zumindest schwierig ist. Die in Kapitel 6 eingeführte flexible Notation erlaubt die formale Diskussion des mehrdimensionalen IPF-Verfahrens ohne Festlegung der Dimensionen der Ausgangstafel und der Marginaltafeln. Danach wurde in Kapitel 7 der informationstheoretische Ansatz zur Konvergenzanalyse auf das mehrdimensionale IPF-Verfahren angewendet. Als neues Resultat dieser Dissertation wurde gezeigt, dass der L_1-Ansatz zur Konvergenzanalyse nicht auf das mehrdimensionale IPF-Verfahren verallgemeinert werden kann. Für den Konvergenzfall wurde in Kapitel 8 bewiesen, dass die Limestafel stetig von der Ausgangstafel abhängt. Dieses Resultat ist ebenfalls ein neues Resultat dieser Dissertation. In Kapitel 9 wurde das asymptotische Verhalten des mehrdimensionalen IPF-Verfahrens im Divergenzfall anhand von zwei Klassen von Beispielen mit dreidimensionalen Ausgangstafeln untersucht. Neben den bereits bekannten Beispielen von Asci und Piccioni (2003) mit zweidimensionalen Marginaltafeln wurde eine neue Klasse von Beispielen mit eindimensionalen Marginaltafeln diskutiert und für diese die Konvergenz der IPF-Teilfolgen gezeigt. In Kapitel 10 wurde die Hypothese aufgestellt, dass jede IPF-Teilfolge, entlang welcher an dieselbe Marginaltafel angepasst wird, konvergiert. Anschließend wurde gezeigt, dass eine variationelle Charakterisierung der Grenzwerte dieser IPF-Teilfolgen nicht möglich ist. Dies stellt das zentrale Resultat dieses Teils der Dissertation dar.

Der Beweis der Hypothese über die Konvergenz der IPF-Teilfolgen im Divergenzfall bleibt daher ein offenes Problem. Weitere offene Probleme sind im Konvergenzfall die Analyse der Abhängigkeit der Limestafel von den Marginaltafeln und die Analyse der Zusammenhangsstruktur der Limestafel. Dabei stellt sich allerdings zuerst die Frage nach der Definition des Zusammenhangsbegriffs für Tafeln der Dimension $d > 2$. Im Divergenzfall bleibt die Frage nach einer andersartigen nützlichen Charakterisierung der Häufungspunkte ein offenes Problem. Auch die Frage nach der stetigen Abhängigkeit der Häufungspunkte von der Ausgangstafel und von den Marginaltafeln sowie die Frage nach der Zusammenhangsstruktur der Häufungspunkte bleiben offen.

Anhang

Literaturverzeichnis

Aas, Erik (2014). “Limit points of the iterative scaling procedure”. *Annals of Operations Research* 215, 15–23. DOI: 10.1007/s10479-013-1416-2 (zitiert auf S. 65).

Agresti, Alan (1990). *Categorical Data Analysis*. New York: Wiley (zitiert auf den S. 8, 72).

Asci, Claudio und Mauro Piccioni (2003). “A note on the IPF algorithm when the marginal problem is unsolvable”. *Kybernetika* 39, 731–737. URL: http://www.kybernetika.cz/content/2003/6/731 (zitiert auf den S. 9, 87–89, 104, 123).

Aulbach, Bernd (2004). *Gewöhnliche Differenzialgleichungen. 2. Auflage*. Heidelberg: Spektrum (zitiert auf S. 116).

Bacharach, Michael O. L. (1970). “The solution of the biproportional constrained matrix problem”. In: *Biproportional Matrices & Input-Output Change*. Cambridge: Cambridge University Press. Kap. 4, S. 43–59 (zitiert auf S. 21).

Baillon, Jean-Bernard, Patrick L. Combettes und Roberto Cominetti (2012). “There is no variational characterization of the cycles in the method of periodic projections”. *Journal of Functional Analysis* 262, 400–408. DOI: 10.1016/j.jfa.2011.09.002 (zitiert auf den S. 121, 122).

Balinski, Michel L. und Gabrielle Demange (1989). “An axiomatic approach to proportionality between matrices”. *Mathematics of Operations Research* 14, 700–719. DOI: 10.1287/moor.14.4.700 (zitiert auf den S. 21, 65).

Balinski, Michel L. und Friedrich Pukelsheim (2006). „Matrices and politics“. In: *Festschrift for Tarmo Pukkila on his 60th Birthday*. Hrsg. von E. P. Liski u. a. Department of Mathematics, Statistics und Philosophy, University of Tampere, S. 233–242 (zitiert auf S. 8).

Bapat, Ravindra (1982). “D_1AD_2 theorems for multidimensional matrices”. *Linear Algebra and its Applications* 48, 437–442. DOI: 10.1016/0024-3795(82)90125-2 (zitiert auf S. 72).

Bauschke, Heinz H. und Patrick L. Combettes (2011). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer (zitiert auf den S. 27, 28, 44).

Beardon, Alan F. (1991). *Iteration of Rational Functions. Complex Analytic Dynamical Systems*. New York: Springer (zitiert auf den S. 95, 97, 104).

Bishop, Yvonne M. M., Stephen E. Fienberg und Paul W. Holland (1975). *Discrete Multivariate Analysis. Theory and Practice*. Cambridge Mass.: MIT Press (zitiert auf S. 72).

Brown, Jack B., Phillip J. Chase und Arthur O. Pittenger (1993). “Order independence and factor convergence in iterative scaling”. *Linear Algebra and its Applications* 190, 1–38. DOI: 10.1016/0024-3795(93)90218-D (zitiert auf den S. 31, 41, 42, 78, 83).

Brualdi, Richard A. (1968). “Convex sets of non-negative matrices”. *Canadian Journal of Mathematics* 20, 144–157. DOI: 10.4153/CJM-1968-016-9 (zitiert auf S. 42).

Christensen, Ronald (1990). *Log-Linear Models*. New York: Springer (zitiert auf S. 72).

Cover, Thomas M. und Joy A. Thomas (2006). *Elements of Information Theory. Second Edition*. Hoboken N. J.: Wiley (zitiert auf den S. 25, 41).

Cramer, Erhard (2000). “Probability measures with given marginals and conditionals: I-projections and conditional iterative proportional fitting”. *Journal for Stochastic Methods and Models* 18, 311–329 (zitiert auf den S. 30, 41, 42, 77).

Csiszár, Imre (1964). „Über topologische und metrische Eigenschaften der relativen Information der Ordnung α“. In: *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*. (Liblice, 5.–13. Juni 1962). Hrsg. von Jaroslav Kozesnik. Prag: Publishing House of the Czechoslovak Academy of Sciences, S. 63–73 (zitiert auf S. 41).

– (1967). “Information-type measures of difference of probability distributions and indirect observations”. *Studia Scientiarium Mathematicarum Hungarica* 2, 299–318 (zitiert auf S. 41).

– (1975). “*I*-divergence geometry of probability distributions and minimization problems”. *Annals of Probability* 3, 146–158. DOI: 10.1214/aop/1176996454 (zitiert auf den S. 21, 27, 41, 42, 75, 83, 109, 121).

Csiszár, Imre und Gábor Tusnády (1984). “Information geometry and alternating minimization procedures”. *Statistics & Decisions* Supplement Issue 1, 205–237 (zitiert auf den S. 41, 42, 48).

Deming, W. Edwards und Frederick F. Stephan (1940). “On a least squares adjustment of a sampled frequency table when the expected marginal totals are known”. *Annals of Mathematical Statistics* 11, 427–444. DOI: 10.1214/aoms/1177731829 (zitiert auf den S. 8, 21).

Gietl, Christoph (2009). *Struktur der Häufungspunkte beim iterativen proportionalen Anpassungsverfahren*. Bachelorarbeit. Institut für Mathematik, Universität Augsburg (zitiert auf den S. 17, 19, 21).

– (2011). *Kullback/Leibler-Divergenz und ihre Anwendungen in der Statistik*. Masterarbeit. Institut für Mathematik, Universität Augsburg (zitiert auf S. 41).

Gietl, Christoph und Fabian P. Reffel (2013a). “Accumulation points of the iterative proportional fitting procedure”. *Metrika* 76, 783–798. DOI: 10.1007/s00184-012-0415-7 (zitiert auf den S. 23, 24, 41, 42, 45, 46, 48, 49).

– (2013b). *Continuity of f-Projections and Applications to the Iterative Proportional Fitting Procedure*. Preprint. Institut für Mathematik, Universität Augsburg. URL: http://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/index/index/docId/2365 (zitiert auf den S. 41, 65).

– (2013c). “Continuity of f-projections on discrete spaces”. In: *Geometric Science of Information. First International Conference, GSI 2013, Paris, France, August 28–30, 2013. Proceedings*. Hrsg. von Frank Nielsen und Frédéric Barbaresco. Berlin: Springer, S. 519–524. DOI: 10.1007/978-3-642-40020-9_57 (zitiert auf den S. 28, 41, 65).

Gokhale, Dattaprabhakar V. und Solomon Kullback (1978a). *The Information in Contingency Tables*. New York: Dekker (zitiert auf S. 72).

– (1978b). “The minimum discrimination information approach in analyzing categorical data”. *Communications in Statistics. Theory and Methods* 7, 987–1005. DOI: 10.1080/03610927808827687 (zitiert auf S. 72).

Hershkowitz, Daniel, Alan J. Hoffman und Hans Schneider (1997). “On the existence of sequences and matrices with prescribed partial sums of elements”. *Linear Algebra and its Applications* 265, 71–92. DOI: 10.1016/S0024-3795(96)00535-6 (zitiert auf den S. 35, 42).

Ireland, C. T. und Solomon Kullback (1968). “Contingency tables with given marginals”. *Biometrika* 55, 179–188. DOI: 10.1093/biomet/55.1.179 (zitiert auf S. 41).

Kemperman, Johannes H. B. (1969). “On the optimum rate of transmitting information”. *Annals of Mathematical Statistics* 40, 2156–2177. DOI: 10.1214/aoms/1177697293 (zitiert auf S. 41).

Knight, Philip A. (2008). “The Sinkhorn-Knopp algorithm: convergence and applications”. *SIAM Journal on Matrix Analysis and Applications* 30, 261–275. DOI: `10.1137/060659624` (zitiert auf S. 8).

Kruithof, J. (1937). „Telefoonverkeersrekening”. *De Ingenieur* 52, E15–E25 (zitiert auf den S. 8, 21).

Kullback, Solomon (1959). *Information Theory and Statistics*. New York: Wiley. Nachdrucke New York: Dover Publications, 1968 und Gloucester Mass.: Peter Smith, 1978 (zitiert auf S. 41).

– (1967). “A lower bound for discrimination information in terms of variation”. *IEEE Transactions on Information Theory* 13, 126–127. DOI: `10.1109/TIT.1967.1053968` (zitiert auf S. 41).

Kullback, Solomon und John C. Keegel (1984). “Categorical data problems using information theoretic approach”. In: *Nonparametric Methods*. Hrsg. von Paruchuri R. Krishnaiah und Pranab K. Sen. Handbook of Statistics 4. Amsterdam: North-Holland. Kap. 35, S. 831–871. DOI: `10.1016/S0169-7161(84)04037-2` (zitiert auf S. 72).

Kullback, Solomon und Richard A. Leibler (1951). “On information and sufficiency”. *Annals of Mathematical Statistics* 22, 79–86. DOI: `10.1214/aoms/1177729694` (zitiert auf den S. 23, 41, 74).

Lauritzen, Steffen L. (1996). *Graphical Models*. Oxford: Clarendon Press (zitiert auf S. 104).

Matúš, František (1997). “On *I*-divergence geometry of affine and log-affine sets”. unveröffentlichtes Manuskript (zitiert auf S. 29).

McCord, Mark u. a. (2010). “Iterative proportional fitting procedure to determine bus route passenger origin-destination flows”. *Transportation Research Record* 2145, 59–65. DOI: `10.3141/2145-07` (zitiert auf S. 8).

Menon, M. V. und Hans Schneider (1969). “The spectrum of a nonlinear operator associated with a matrix”. *Linear Algebra and its Applications* 2, 321–334. DOI: `10.1016/0024-3795(69)90034-2` (zitiert auf S. 42).

Needham, Tristan (1997). *Visual Complex Analysis*. Oxford: Clarendon Press (zitiert auf den S. 97, 104).

Pinsker, Mark S. (1963). *Information und Informationsstabilität zufälliger Größen und Prozesse*. Aus dem Russischen übers. von Kurt Nawrotzki. Berlin: Deutscher Verlag der Wissenschaften (zitiert auf S. 41).

Pretzel, Oliver (1980). “Convergence of the iterative scaling procedure for non-negative matrices”. *Journal of the London Mathematical Society* 2, 379–384. DOI: `10.1112/jlms/s2-21.2.379` (zitiert auf den S. 35–37, 42).

Pukelsheim, Friedrich (2014). “Biproportional scaling of matrices and the iterative proportional fitting procedure”. *Annals of Operations Research* 215, 269–283. DOI: `10.1007/s10479-013-1468-3` (zitiert auf den S. 13, 14, 21, 38, 40, 42, 65).

Reffel, Fabian P. (2014). *Konvergenzverhalten des iterativen proportionalen Anpassungsverfahrens im Fall kontinuierlicher Maße und im Fall diskreter Maße*. Berlin: Logos (zitiert auf den S. 51, 65).

Rockafellar, Ralph T. (1972). *Convex Analysis*. Princeton N. J.: Princeton University Press (zitiert auf den S. 23, 59, 60).

Rothblum, Uriel G. und Hans Schneider (1989). “Scalings of matrices which have prespecified row sums and column sums via optimization”. *Linear Algebra and its Applications* 114/115, 737–764. DOI: `10.1016/0024-3795(89)90491-6` (zitiert auf den S. 21, 42).

Sinkhorn, Richard (1972). “Continuous dependence of A in the D_1AD_2 theorems”. *Proceedings of the American Mathematical Society* 32, 395–398. DOI: 10.1090/S0002-9939-1972-0297792-4 (zitiert auf S. 65).

Strehmel, Karl, Rüdiger Weiner und Helmut Podhaisky (2012). *Numerik gewöhnlicher Differentialgleichungen. Nichtsteife, steife und differential-algebraische Gleichungen. 2., überarbeitete und erweiterte Auflage*. Wiesbaden: Vieweg+Teubner (zitiert auf den S. 116, 117).

Tsybakov, Alexandre B. (2009). *Introduction to Nonparametric Estimation*. New York: Springer (zitiert auf S. 41).

Vomlel, Jiří (1999). *Methods of Probabilistic Knowledge Integration*. Dissertation. Faculty of Electrical Engineering, Czech Technical University, Prag. URL: http://staff.utia.cas.cz/vomlel/prace.ps (zitiert auf S. 121).

Stichwortverzeichnis