Prognosen bei partieller Information*

Von G. Menges und E. Kofler

Es wird gezeigt, wie man zur Erzielung von Wirtschaftsprognosen alle verfügbaren Informationen, auch wenn sie unvollständig sind, ausnutzen kann. Evtl. verbliebene Lücken werden mit entscheidungstheoretischen Hilfsmitteln geschlossen.

1. Definitionen und allgemeine Aufgabenstellung

Alle wirtschaftlichen Prognosen sind im Grunde stochastisch, d. h. ziehen den Wahrscheinlichkeitsgedanken heran, aber die stochastische Interpretation der Prognosen umgreift eine breite Skala; von dem einen Extrem, bei dem die Wahrscheinlichkeit explizit gar nicht in Erscheinung tritt („sachlogische" Prognosen; vgl. *Blind* 1964 und *Menges* 1966), bis zum Extremfall der vollständigen, exakten, numerischen Kenntnis der Prognosewahrscheinlichkeiten, wie er in der Ökonometrie meist gegeben ist oder zumindest angenommen wird. Den beiden Extremen wurde Aufmerksamkeit gewidmet, besonders dem Fall der numerisch exakt spezifizierten ökonometrischen Prognosen (*Menges* 1967).

Zwischen den beiden Extremen liegen aber viele Fälle, bei denen die Wahrscheinlichkeiten nicht vollständig oder nicht exakt oder nicht numerisch (kardinal) bekannt sind. Wir bezeichnen solche Fälle von unvollständigen, unscharfen oder nicht-numerischen (nicht-kardinalen, aber z. B. ordinalen) Wahrscheinlichkeiten zusammenfassend als Fälle *partieller Information.* Vier typische Beispiele sind die folgenden:

a) Unvollständigkeit: Gegeben sind m mögliche zukünftige Zustände $z_1, z_2, \ldots, z_m$ mit den entsprechenden Wahrscheinlichkeiten $p_1, p_2, \ldots, p_m$; k ($k < m$) Wahrscheinlichkeiten $p_1, p_2, \ldots, p_k$ sind exakt bekannt, die restlichen Wahrscheinlichkeiten $p_{k+1}, \ldots, p_m$ sind nur in

* Einführungsreferat auf der Gründungssitzung des Ausschusses für Ökonometrie des Vereins für Socialpolitik am 24. März 1977 in Bad Homburg v. d. H. Die vorliegende Fassung ist etwas ausführlicher als der Vortragstext. Außerdem wird hier auf einige Einwände eingegangen, die in der Diskussion aufgetreten waren.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Form $\sum_{j=k+1}^{m} p_j = 1 - \sum_{j=1}^{k} p_j$ bekannt (Nullinformation bzgl. dieser restlichen Wahrscheinlichkeiten).

b) Unschärfe (i. e. S.): Einige oder alle Wahrscheinlichkeiten sind nur in Form von Intervallen

$$\alpha_j \leqq p_j \leqq \beta_j \; (j \in J, J = \{1, \ldots, m\}) \text{ bekannt.}$$

(Immer gilt für alle p_j: $p_j \in [0,1]$ und $\sum_{j=1}^{m} p_j = 1$.)

c) Ordinale Messung: Von einigen oder allen Wahrscheinlichkeiten ist nur eine schwache Ordnung bekannt

$$p_{j1} \leqq p_{j2} \leqq \cdots$$

Außerdem schließt der Fall der partiellen Information Mischformen von a, b, und c sowie einige andere Typen, auf die wir hier nicht näher eingehen, ein. Überhaupt werden wir uns hauptsächlich auf den Fall c) der ordinalen Messung konzentrieren, da dieser am plausibelsten und am einfachsten zu behandeln ist. Allgemein gesprochen ist der Fall der partiellen Information dadurch charakterisiert, daß irgendeine Unbestimmtheit in den Konstituenten eines Prognoseproblems vorliegt.

Wir sind der Auffassung, daß jedes Prognoseproblem der Praxis streng genommen ein solches unter partieller Information ist.

Die Grenzfälle der partiellen Information sind:

a) Nullinformation. Hier ist kein (stochastisch auswertbares) A-priori-Wissen vorhanden, und es ist auch keine statistische Inferenz (Schätzung) möglich oder durchführbar.

b) Vollständige Information. Hier ist das (stochastisch auswertbare) A-priori-Wissen so stark, daß alle Konstituenten des Prognoseproblems exakt bekannt sind; insbesondere ist die Wahrscheinlichkeitsverteilung über den möglichen zukünftigen Zuständen exakt bekannt (nicht etwa der Zustand selbst).

In der statistischen Entscheidungstheorie wird der Grenzfall der Nullinformation als Fall totaler Ungewißheit bezeichnet, der Grenzfall der vollständigen Information als Risikofall, manchmal auch als Bernoullifall.

Als Ziel der Behandlung des Prognoseproblems, d. h. als Lösung, betrachten wir in allen Fällen die Ermittlung einer Punktprognose. Bei dieser Auffassung führt die Behandlung aller drei Informationsfälle

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

zu entscheidungstheoretischen Begriffen und Lösungen, und zwar im einzelnen zu den folgenden:

Nullinformation: Hier wird das Prognoseproblem als Zweipersonen-Nullsummenspiel des Prognostikers gegen die unbekannte zukünftige Welt aufgefaßt; die Lösung erfolgt nach dem Maximin-Kriterium (bzw. Minimax-Kriterium, wenn die Konsequenzen in Verlusteinheiten gemessen sind).

Vollständige Information: Hier erfolgt die Lösung des Prognoseproblems mit Hilfe des Bernoullikriteriums.

Partielle Information: Das Vorgehen teilt sich hier in zwei Schritte:

a) Reduktion der Ungewißheit durch Heranziehung zusätzlicher Informationen. Für diese Aufgabe wurde von uns (*Kofler/Menges* 1976) die Theorie der linearen partiellen Information (*LPI*-Theorie) entwickelt.
b) Überwindung der verbleibenden Ungewißheit durch das $\mathrm{MaxE}_{\min}$-Kriterium. Es kann als *LPI*-Fall des Bernoulliprinzips betrachtet werden.

Wegen der exakten mathematischen Darstellung der Theorie der linearen partiellen Information verweisen wir auf *Kofler/Menges* (1976); hier versuchen wir, eine intuitive und beispielhafte Darstellung zu geben.

2. Theorie der linearen partiellen Information

Wir fassen die den möglichen zukünftigen Zuständen $z_i, \ldots, z_m$ entsprechenden Wahrscheinlichkeiten $p_1, \ldots, p_m$ zum Verteilungsvektor ϱ zusammen. Alle möglichen Verteilungen ϱ bilden im $\mathbf{R}^m$ ein $(m-1)$-dimensionales Verteilungssimplex $S^{(m)}$. Für $m = 3$ kann $S^{(3)}$ als baryzentrisches Dreieck dargestellt werden (Fig. 1 a). Jeder Punkt des Dreiecks hat drei nicht-negative (baryzentrische) Koordinaten, deren Summe stets gleich 1 ist.

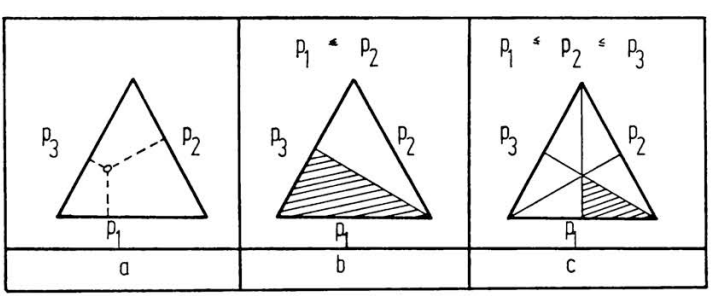

Fig. 1: *LPI* bei $m = 3$, schwache Ordnung

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Der Fall vollständiger Information bedeutet, daß man von einem bestimmten Punkt des Dreiecks weiß, daß er die wahre Verteilung repräsentiert.

Der Fall der Nullinformation bedeutet, daß alle Punkte des Dreiecks mögliche Verteilungen repräsentieren. Der Fall der partiellen Information bedeutet, daß man von einem echten Teilgebiet des Dreiecks (allgemein von $S^{(m)}$) weiß, daß es einen Punkt enthält, der die wahre Verteilung repräsentiert. Ist z. B. bekannt, daß der Zustand z_2 eher eintreten wird als der Zustand z_1, während die Wahrscheinlichkeit für den Zustand z_3 unbekannt ist (außer der Trivialität $p_3 = 1 - (p_1 + p_2)$), d. h. man kennt die schwache Partialordnung

$$p_1 \leqq p_2 \ ,$$

dann bedeutet diese Kenntnis, daß man weiß, daß der „wahre Punkt" (die wahre Verteilung) in dem in Fig. 1 b schraffierten Teilgebiet des Dreiecks liegt. Ist hingegen die schwache Ordnung vollständig bekannt, d. h. weiß man, daß

$$p_1 \leqq p_2 \leqq p_3$$

gilt, dann bedeutet dieses Wissen, daß der „wahre Punkt" (die wahre Verteilung) in dem in Fig. 1 c schraffierten Teilgebiet des Dreiecks liegt. (In 1 c ist die Ungewißheit stärker als in 1 b reduziert.)

Der oben angedeutete Fall, daß die drei Wahrscheinlichkeiten in Intervallform gegeben sind:

$$\alpha_j \leqq p_j \leqq \beta_j \ (j = 1, 2, 3)$$

ist graphisch in Fig. 2 dargestellt.

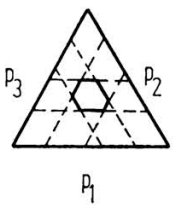

Fig. 2: *LPI* bei $m = 3$, Intervallangaben

Der verbleibende Ungewißheitsrest ist in diesem Beispiel das schraffierte Sechseck innerhalb des baryzentrischen Dreiecks, d. h. man weiß, daß der „wahre Punkt", d. h. die wahre Prognoseverteilung, in diesem Sechseck liegt.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Im allgemeinen definieren wir eine partielle Information als echtes Teilgebiet des Verteilungssimplexes $S^{(m)}$, wobei wir uns auf *lineare* partielle Informationen beschränken (tatsächlich kann man zeigen, daß dies im strengen Sinn keine Beschränkung der Allgemeinheit ist).

Definition: Eine lineare partielle Information *LPI* grenzt im Verteilungssimplex $S^{(m)}$ ein konvexes Polyeder P ein, welches ein echtes Teilgebiet von $S^{(m)}$ ist.

$$LPI : P \subset S^{(m)} \; .$$

Degeneriert P zu einem einzelnen Punkt, liegt der Grenzfall vollständiger Information vor; hingegen würde $P = S^{(m)}$ den Grenzfall der Nullinformation bedeuten. Beide Grenzfälle sind jedoch in unserer Betrachtung ausgeschlossen.

Wichtig für das folgende ist noch die Matrix der Eckpunkte des konvexen Polyeders, die wir mit $M\,(LPI)$ bezeichnen, die Eckpunkte mit $\varrho^{(e)}$. Wir betrachten wieder zwei Beispiele:

1. Beispiel: $m = 2$, Intervallangaben

$$LPI\colon\; 0{,}3 \leqq p_1 \leqq 0{,}4;$$

$$M\,(LPI) = \begin{bmatrix} 0{,}3 & 0{,}4 \\ 0{,}7 & 0{,}6 \end{bmatrix} .$$

2. Beispiel: $m = 3$, schwache Ordnung

$$LPI\colon\; p_1 \leqq p_2 \leqq p_3$$

$$M\,(LPI) = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{2} & 1 \end{bmatrix} .$$

Für den allgemeinen Fall vollständiger schwacher Ordnung

$$LPI\colon\; p_1 \leqq p_2 \leqq \ldots \leqq p_m$$

beweist man:

$$M\,(LPI) = \begin{bmatrix} \frac{1}{n} & 0 & \cdots & 0 \\ \frac{1}{n} & \frac{1}{n-1} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n-1} & \cdots & 1 \end{bmatrix} .$$

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Bestimmung der Eckpunktematrix ist im Fall vollständiger schwacher Ordnung also sehr einfach, in anderen Fällen, z. B. auch bei Intervallangaben, kann die Bestimmung von $M(LPI)$ sehr rechenaufwendig werden.

3. Das $\mathrm{MaxE}_{\min}$-Prinzip

Wir betrachten, wie oben angekündigt, das Prognoseproblem nunmehr als Entscheidungsproblem mit den Konstituenten:

(1) $A = \{a_i \mid i = 1, \ldots, n\}$ Strategiemenge des Prognostikers. (Oft bedeutet bei Prognoseproblemen a_i die Strategie, z_j mit $i = j$ als wahr zu betrachten).

(2) $Z = \{z_j \mid j = 1, \ldots, m\}$ Zustandsmenge (Menge der möglichen zukünftigen Zustände).

(3) $[u_{ij}]$; $i = 1, \ldots, n$; $j = 1, \ldots, m$; Nutzenmatrix; $u_{ij} \equiv u(a_i, z_j)$ gibt den Nutzen für den Prognostiker an, wenn er die Prognosestrategie a_i wählt, während z_j der wahre zukünftige Zustand ist.

(4) LPI: $\{\varrho \mid \varrho \in P; P \subset S^{(m)}\}$.

Mit LPI bezeichnen wir der Einfachheit halber sowohl die lineare partielle Information selbst als auch das durch diese Information definierte konvexe Polyeder im Verteilungssimplex $S^{(m)}$.

(5) $R^{(e)} = \{\varrho^{(e)}\}$ Menge der Eckpunkteverteilungen; diese wird zur Defifinition des LPI-Nullsummenspiels benötigt.

Das Prognoseproblem ist nunmehr wie folgt gegeben:

$$[A; Z; LPI; [u_{ij}]], \quad i = 1, \ldots, n, \quad j = 1, \ldots, m .$$

Wenn die wahre Prognoseverteilung $\bar{\varrho} = (p_i, \ldots, p_m)$ bekannt wäre (Fall der vollständigen Information), so könnte das Prognoseproblem mit Hilfe des Bernoulliprinzips wie folgt gelöst werden: Man wähle die Prognosestrategie $\bar{a} \in A$, für welche gilt:

$$\sum_{j=1}^{m} u(\bar{a}, z_j)\, p_j = \max_{a_i \in A} \sum_{j=1}^{m} u(a_i, z_j)\, p_j . \tag{1}$$

Wenn nach Ausschöpfung aller A-priori- und A-posteriori-Kenntnisse das Verteilungssimplex $S^{(m)}$ auf den Bereich der LPI, d. h. auf das diesem entsprechende konvexe Polyeder $P(\varrho)$, reduziert ist, dann kann (in einer natürlichen Erweiterung des Bernoulliprinzips) das Prognoseproblem nach dem $\mathrm{MaxE}_{\min}$-Prinzip gelöst werden: Man wähle die Prognosestrategie $a_i{*} \in A$, welche die im LPI-Bereich minimale Nutzenerwartung maximiert:

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

$$\min_{\varrho \in LPI} \sum_{j=1}^{m} u_i *_j p_j = \max_{a_i \in A} \min_{\varrho \in LPI} \sum_{j=1}^{m} u_{ij} p_j = v(a_i *) \ . \tag{2}$$

Die axiomatische Begründung des $\text{MaxE}_{\min}$-Prinzips findet sich in *Kofler/Menges* (1976) § 26.

Die Lösung von (2) auf Grund des $\text{MaxE}_{\min}$-Prinzips bei gegebener *LPI* (ϱ) ist mit der Lösung des sog. *LPI*-Nullsummenspiels äquivalent. Das Nullsummenspiel $[\{a_i\}; \{\varrho^{(e)}\}, B(a_i, \varrho^{(e)})]$ hat die Spielmatrix

$$B(a_i, \varrho^{(e)}) = [u_{ij}] \cdot M(LPI) \ . \tag{3}$$

Der Beweis findet sich in *Kofler/Menges* (1976) § 27.

Es ist zu bemerken, daß im allgemeinen die Lösung von (3) im Bereich der gemischten Strategien zu finden ist.

Die rechte Seite von (2), nämlich $v(a_i *)$, bedeutet den sog. Entscheidungswert der Entscheidungssituation. Der Entscheidungswert gibt die maximal gewährleistete Nutzenerwartung an.

4. Informationen über Informationen

Partielle Informationen können, wie alle Informationen, exakt sein oder fragwürdig; auch kann es oft sinnvoll sein, irgendwelche Informationen nach ihrer Zuverlässigkeit oder Glaubwürdigkeit zu bewerten. Wenn z. B. drei Forschungsinstitute unterschiedliche Prognosen (ob in Form vollständiger oder partieller Information, spielt dabei keine Rolle) aufgestellt haben, so wird der Wirtschaftspolitiker, der diese Prognosen zum Gegenstand politischer Entscheidungen macht, eine Bewertung vornehmen müssen; z. B. indem er die ihm zuverlässigst scheinende auswählt und die restlichen vernachlässigt, irgendeinen Mittelwert bestimmt oder die drei Prognosen einzeln bewertet, so daß sie schwach geordnet werden können.

Sobald eine Information auf irgendeiner Stufe des Informationssystems partiell im oben definierten Sinn ist, kann das Problem als „*LPI* höherer Ordnung" aufgefaßt und gelöst werden.

Z. B. gehe es um drei mögliche zukünftige Zustände z_1 (z. B. Fallen des Dollarkurses), z_2 (z. B. Gleichbleiben des Dollarkurses) und z_3 (z. B. Steigen des Dollarkurses) mit den entsprechenden Wahrscheinlichkeiten p_1, p_2 und p_3. Zwei Institute seien mit der Frage befaßt, und zwar prognostiziere

Institut 1:

$$LPI_1^1 : p_1 \leqq p_2 \leqq p_3$$

Institut 2:

$$LPI_2^1 : p_1 \leqq p_3 \leqq p_2 \ .$$

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Bewertung durch den Wirtschaftspolitiker laute:

$$LPI^2 : \pi\,(LPI_1^1) \leqq \pi\,(LPI_2^1)\ ,$$

d. h. der Wirtschaftspolitiker gibt der Prognose des ersten Instituts die höhere Wahrscheinlichkeit (oder Plausibilität oder Glaubwürdigkeit). Möglicherweise ist er aber seiner Sache nicht so sicher und wünscht nur auszusagen, daß er es zwar für gerechtfertigt hält, Institut 1 eher zu vertrauen als Institut 2, aber nicht in der unbedingten Form von LPI^2, sondern nur in der Form, daß er aussagt:

$$LPI_1^2 : \Phi\,(\pi\,(LPI_1^1) \leqq \pi\,(LPI_2^1)) = \omega \qquad (\omega \text{ unscharf})\ ,$$

$$LPI_2^2 : \Phi\,(\pi\,(LPI_1^1) > \pi\,(LPI_2^1)) = 1 - \omega\ .$$

Dann bildet das Paar $(\omega, 1 - \omega)$ eine LPI^3, d. h. eine LPI dritter Ordnung, wobei ω die (große) Wahrscheinlichkeit (oder Glaubwürdigkeit etc.) dafür ist, daß LPI_1^2 richtig ist, während $1 - \omega$ die (kleine) Wahrscheinlichkeit ist, daß LPI_2^2 richtig ist:

$$LPI^3:\quad 1 - \omega \leqq \omega \geqq \frac{1}{2}\ .$$

Ein wichtiger Satz der LPI-Theorie, den wir an anderer Stelle exakt zu beweisen beabsichtigen[1], besagt: sind LPI^1, LPI^2 und LPI^3 stochastisch unabhängig voneinander und hierarchisch geordnet, dann kann das komplexe Entscheidungsproblem der Art (LPI^1, LPI^2, LPI^3) in ein Entscheidungsproblem der einstufigen Art $LPI^{(1)}$ so überführt werden, daß die $\mathrm{MaxE_{min}}$-Lösungen von $LPI^{(1)}$ und (LPI^1, LPI^2, LPI^3) äquivalent sind. Dieser Satz gilt für jeden endlichen Komplexizitätsgrad.

Die LPI-Theorie, vornehmlich in ihrer komplexen Ausgestaltung, kann auf verschiedene andere Weisen zur Lösung von Prognoseproblemen verwendet werden. Wir wollen auf zwei besonders wichtig erscheinende Anwendungen näher eingehen, die LPI-Inferenz für Prognosen und die Vorherbestimmung der vorherbestimmten Variablen.

5. LPI-Inferenz

Sätze der Wahrscheinlichkeitsrechnung können grundsätzlich auf LPI erweitert werden (*Kofler/Menges* 1977 b). Mit Hilfe des entsprechend erweiterten Satzes über totale Wahrscheinlichkeiten gewinnt man ein für Prognoseprobleme wichtiges Instrument der Inferenz.

---

[1] Statistische Hefte, 20. Jg., 1979, vermutlich Heft 2 oder 3.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Satz für die LPI-Inferenz:

Sei P das der *LPI* entsprechende konvexe Polyeder, $\varrho(P)$ die Menge aller möglichen Verteilungen im *LPI*-Bereich: $\varrho(P) = \{\varrho \mid \varrho \in P\}$. P wird in r disjunkte konvexe Polyeder zerlegt:

$$P = P_1 + P_2 + \ldots + P_r \,. \tag{4}$$

Fig. 3 veranschaulicht diesen Sachverhalt für $m = 3$ und $r = 2$ im baryzentrischen Dreieck, wobei wieder, wie in Fig. 1, als *LPI:* $p_1 \leqq p_2 \leqq p_3$ angenommen wurde.

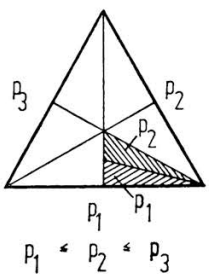

Fig. 3: *LPI* bei m = 3, schwache Ordnung, Zerlegung in 2 disjunkte Polyeder

In Fig. 3 ist der schraffierte Teil des Verteilungssimplexes in die zwei disjunkten konvexen Polyeder P_1 und P_2 zerlegt.

Sei $\varrho(P_i)$, $i = 1, \ldots, r$, die Menge aller möglichen Verteilungen ϱ im P_i-Bereich: $\varrho(P_i) = \{\varrho \mid \varrho \in P_i\}$. Sei weiterhin $W(P_i)$ die Wahrscheinlichkeit, daß die Verteilung ϱ dem P_i-Bereich angehört: $W(P_i) = W(\varrho \in P_i)$.

Dann gilt

$$\varrho(P) = \sum_{i=1}^{r} W(P_i) \cdot \varrho(P_i) \,. \tag{5}$$

Jedes P_i kann als eine Urne vorgestellt werden. Aber im Gegensatz zu der klassischen Interpretation des Satzes über totale Wahrscheinlichkeiten, wo jeder Urne eine feste Verteilung zugeordnet ist und wo die Wahrscheinlichkeiten für das Ziehen bestimmter Urnen gegeben sind, haben wir im obigen Modell Urnen mit innerhalb des *LPI*-Bereichs variablen Verteilungen und entsprechend *LPI*-variablen Wahrscheinlichkeiten für das Ziehen bestimmter Urnen.

Der Beweis ist eine natürliche Erweiterung des Satzes über totale Wahrscheinlichkeiten, der selbst ein Spezialfall von (5) ist. In diesem Spezialfall schrumpft jede Urne zu einer festen Verteilung zusammen; auch die Wahrscheinlichkeiten für das Ziehen bestimmter Urnen werden fest.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Algorithmisch betrachtet ist es einfacher, wenn man

1. jeden Zustand separat betrachtet,
2. die *LPI* als Intervalle auffaßt und
3. die Zerlegung auf zwei Klassen ($r = 2$) beschränkt sein läßt.

$\varrho\,(P_i)$ entspricht einer *LPI* und ist bei bekannter Zerlegung ebenfalls bekannt.

Hingegen ist $W\,(P_i)$ nur auf Grund einer zusätzlichen Information, z. B. einer Stichprobenentnahme σ, als *LPI* interpretierbar, evtl. zusammen mit einer Konfidenzzahl. Für $\varrho\,(P)$ liegt unter diesen Voraussetzungen gleichfalls eine *LPI* vor, wie in *Kofler/Menges* (1976) (Satz 21.1 auf S. 120 ff.) bewiesen ist.

Danach besteht eine Analogie zu der Bayes'schen Inferenz. In der Tat liefert der Satz 5 ein Inferenzinstrument ähnlich dem Bayes'schen. Wir bezeichnen sie als *LPI*-Inferenz, da sie sich im *LPI*-Bereich abspielt. Die Bayes'sche Inferenz kann aber auch unmittelbar in *LPI*-Form ausgedrückt werden, z. B. wenn die A-priori-Wahrscheinlichkeiten und/oder die Likelihoods, wie oft in der Praxis, als Intervalle gegeben sind. Sie sind als solche *LPI*'s und die Transformation in A-posteriori-Wahrscheinlichkeiten führt wieder zu *LPI*'s, d. h. zu (numerisch direkt spezifizierbaren) Intervallen, wenn die A-priori-Wahrscheinlichkeiten und die Likelihoods in Intervallform gegeben waren. Analoges gilt, wenn die Unvollständigkeit der Information darin besteht, daß die A-priori-Wahrscheinlichkeiten und/oder Likelihoods numerisch unbekannt sind, wenn sie aber jeweils schwach geordnet werden können; dann ergibt sich für die A-posteriori-Wahrscheinlichkeiten ebenfalls eine schwache Ordnung.

Der Ausdruck $W\,(P_i)$ besitzt in der Tat eine gewisse Analogie zu der Likelihood im Sinne *R. A. Fishers*. Wir bezeichnen daher $W\,(P_i)$ als *LPI*-Likelihood. Sie gibt nach erfolgter Beobachtung σ die Wahrscheinlichkeit an, daß die Verteilung ϱ dem P_i-Bereich angehört.

Es wäre zwar naheliegend, entsprechend im Ausdruck $\varrho\,(P_i)$ eine A-priori-Wahrscheinlichkeit zu erblicken. Aber dies wäre nicht gerechtfertigt, da das A-priori-Wissen natürlich bereits in P enthalten ist.

Aber es zeigt sich eine andere interessante Beziehung bezüglich A-priori- und A-posteriori-Verteilungen. Vor der Beobachtung σ haben wir eine *LPI*, deren zugehörige Verteilung wir $\varrho\,(P)_v$ nennen, wobei der Suffix v auf „vorher, a priori, nämlich vor σ", hinweist. Auf Grund der Beobachtung σ wird eine Zerlegung der Form (4) realisiert, und es wird jetzt festgestellt, mit welcher Wahrscheinlichkeit die Verteilung ϱ einem der P_i angehört. Mit anderen Worten: Es werden die $W\,(P_i)$ bestimmt

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

bzw. neu bestimmt. Das resultirende neue konvexe Polyeder hat als Menge aller möglichen Verteilungen $\varrho(P)_N$, wobei der Suffix N auf „nachher, a posteriori, d. h. nach der Beobachtung σ" hinweist.

$\varrho(P)_N$ ist zugleich auch eine Prognose, wenn die Beobachtung σ entsprechend zukunftsbezogen ist, d. h. z. B. die Vorherbestimmung einer exogenen Variablen darstellt oder auch einen Konjunkturindikator, wie der Auftragseingangsindex. Jedes neue konvexe Polyeder ist im jeweils vorhergehenden enthalten (in der Tat konnten wir einen entsprechenden Schachtelungssatz beweisen[2]). Man kann sich nun ein beliebig langes Kettenverfahren denken, bei dem jede Information über die Zukunft zu einem neuen, kleineren und enger auf die Zukunft bezogenen konvexen Polyeder führt. Daraus folgt, daß die Einschachtelung schließlich — bei unendlicher Fortsetzung — zum Grenzfall der festen Verteilung (Bernoullifall) führt, der hier also als der Grenzfall der festen Prognoseverteilung aufgefaßt werden kann. Die in unserem Buch bewiesenen Sätze über Kettenverfahren (z. B. S. 125 ff.) lassen sich allesamt analog auf die *LPI*-Inferenz anwenden.

So ist $\varrho(P)_N$ einer ersten Stufe zugleich das $\varrho(P)_v$ der zweiten Stufe, allgemein

$$\varrho(P)_N^{(\nu)} = \varrho(P)_v^{(\nu+1)}\ ;\ \nu = 1, 2, \ldots \tag{6}$$

Dieses Verfahren kann solange fortgeführt werden, wie zusätzliche Informationen hinzutreten. Praktisch wird man zwar nie den Grenzfall erreichen, aber man kann sich ihm beliebig nähern.

In der Diskussion zu unserem Vortrag wurden im Zusammenhang mit der Zerlegung von P in disjunkte Teile und ebenso im Hinblick auf den sequentiellen Charakter der Einschachtelung einige z. T. sehr interessante Einwände vorgebracht. Insbesondere war die Frage nach der Optimalität der Zerlegung sehr anregend, und wir haben uns inzwischen noch einige Gedanken zu dieser Frage gemacht. Wir möchten zu den Einwänden wie folgt Stellung nehmen:

Wir nennen eine Zerlegung eine „normale Partition", wenn sie unendlich ist und der größte Partitionsteil gegen Null strebt. Für den Fall normaler Partition beweisen wir, daß bei Vernachlässigung der Informationsbeschaffungskosten unser *LPI*-Schachtelungsprozeß zum Bernoullifall, d. h. dem Grenzfall der festen Verteilung, führt. Wir bezeichnen diesen Fall ohne Informationsbeschaffungskosten, d. h., daß die Schachtelung zum Bernoullifall führt, als „ideale Konsequenz".

---

[2] *Kofler* und *Menges* (1976), S. 125 ff.. S. 157 ff.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Wie steht es aber, wenn wir die Informationsbeschaffungskosten mit in die Betrachtung einbeziehen und außerdem, wenn nur endliche Prozesse berücksichtigt werden? Zum Zwecke der Kostenminimierung muß die Anzahl N_0 der Versuche (Beobachtungen) möglichst niedrig gehalten werden, und zwar in Relation zur Zahl n der Zustände, d. h. wir müssen auch die Zahl der Zustände klein halten. Wir hoffen beweisen zu können, daß dies erreichbar ist, indem wir dort und nur dort weiter teilen, wo die relative Häufigkeit der Beobachtungen am größten ist. Die Gefahr einer Fehlentscheidung wird — wenn auch nicht vermieden — so doch reduziert, da wir die jeweiligen Komplemente mit den jeweiligen Komplementärwahrscheinlichkeiten berücksichtigen.

Im Falle des sequentiellen Vorgehens empfiehlt sich die Aufstellung eines Entscheidungsbaumes derart, daß an jedem Verzweigungspunkt nur zwei Äste („stop" oder „weiterteilen") abgehen (duale Partition) und an den Endpunkten des ganzen Baumes die Auszahlungen unter Berücksichtigung der Informationsbeschaffungskosten und der jeweiligen Partitionsstrategie festgestellt werden. Die Auszahlungen sind die Netto-Informationswerte. Jede Partitionsstrategie wird durch einen Streckenzug repräsentiert (in Fig. 4 ist ein solcher dick eingezeichnet).

Der optimale Streckenzug ergibt sich auf Grund des bekannten Rollback-Verfahrens. Die Netto-Informationswerte können dabei auch für alle Eckpunkte (also nicht für die Endpunkte) gesondert berechnet werden.

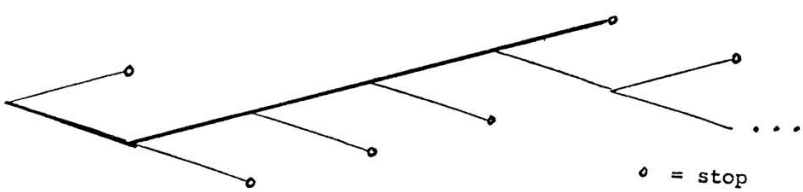

Fig. 4: Entscheidungsbaum und Partitionsstrategie

Man kann zeigen:

1. Jede Partition läßt sich auf eine duale Partition zurückführen.

2. Die duale Partition ist ökonomisch (im Sinne des Netto-Informationswertes) genau dann, wenn die Wahrscheinlichkeiten für die beiden Äste gleich groß sind[3].

[3] Vorausgesetzt, daß Gleichverteilung vorliegt. Liegt sie nicht vor, gilt die obige Aussage in Näherung.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Die Optimalität der Partition ist damit (ähnlich wie bei der Bayesschen Inferenz) vom Stichprobenergebnis abhängig.

Eine andere interessante Konsequenz der *LPI*-Inferenz ist „die transformierte *LPI* ohne Konfidenzzahl". Wir können nämlich den *LPI*-Bereich mit einer Konfidenzzahl γ und den Komplementärbereich, d. h. den ganzen Verteilungssimplex minus den durch die jeweilige *LPI* gegebenen konvexen Polyeder, mit der Komplementärwahrscheinlichkeit $1 - \gamma$ versehen und auf diese Zerlegung des ganzen Verteilungssimplexes mit Satz 5 das *LPI*-Inferenzinstrument anwenden.

In der Diskussion wurde auf die Gefahr von Fehlentscheidungen aufmerksam gemacht, da mit der durch die Konfidenzzahl γ gegebenen Wahrscheinlichkeit $1 - \gamma$ die wahre Verteilung nicht im *LPI*-Bereich liegt. In der Tat empfehlen wir, neben der Aussage über das Konfidenzintervall auch die Aussage über das komplementäre Intervall (mit der Komplementärwahrscheinlichkeit $1 - \gamma$) zu berücksichtigen. Durch die oben beschriebene Transformation verschwindet dann die Konfidenzzahl, sie wird gleichsam von der transformierten *LPI* aufgesogen.

Mit der Prognose $\varrho\,(P)_N$, die im einfachsten Prognosefall als eine Prognose der Zustandsverteilung aufgefaßt werden kann, sind automatisch andere Prognosen verbunden. Eine wichtige betrifft den Entscheidungswert. Wie oben gezeigt wurde, gelangen wir zum Entscheidungswert der Prognose auf Grund des $\mathrm{MaxE}_{\min}$-Prinzips.

Man kann natürlich für jede Stufe der *LPI*-Kette den Entscheidungswert bestimmen, der in unserer Interpretation die Bruttoprognosebewertung ist, und an der Differenz aufeinanderfolgender Prognosebewertungen den semantischen Informationswert neu hinzutretender Informationen (Stichproben- oder sonstiger Informationen) für die Prognose bemessen.

Bei der Bestimmung von $\varrho\,(P)_v$ kann man eine Art von Prinzip der „revealed preferences" anwenden. Wenn nämlich z. B. durch Befragung festgestellt wurde, daß a_i* die optimale Strategie ist, die auf Grund des Bernoulliprinzips gewählt wurde, dann ergibt sich (aus der Befragung)

$$\forall i \; \Sigma\, u_{ij}\, p_j \leqq \Sigma\, u_{i*j}\, p_j \tag{7a}$$

und (7a) bestimmt auch eine *LPI* für die Zustandsverteilung, z. B. in der Situation

$$[\{a_1, a_2\} \;;\; \{z_1, z_2\} \;;\; [\begin{smallmatrix} 2 & 3 \\ 1 & 4 \end{smallmatrix}]] \tag{7b}$$

führt die Information, daß a_1 als optimal gewählt wurde, zur *LPI:* $p_1 > p_2$.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

6. Ein numerisches Beispiel: Vorherbestimmung

Als Beispiel betrachten wir einen Fall der Vorherbestimmung, für den es bisher keine alternative Methode gibt. Der Einfachheit halber nehmen wir zwei Zustände an: „z_1 = schlecht" mit Wahrscheinlichkeit p_1, „z_2 = gut" mit Wahrscheinlichkeit p_2; a_1: pessimistische Vorherbestimmung, a_2: optimistische Vorherbestimmung.

Beim Zusammentreffen des in Wahrheit schlechten zukünftigen Zustandes z_1, z. B. Fortdauer der Rezession, mit der pessimistischen Vorherbestimmung a_1 entstehe kein Verlust, ebensowenig beim Zusammentreffen des in Wahrheit guten zukünftigen Zustandes z_2 (Konjunkturaufschwung) mit der optimistischen Vorherbestimmung a_2. In den beiden Fällen außerhalb der Diagonalen entstehe indessen ein Verlust; im Falle (a_2, z_1) sei er dreimal so groß wie im Falle (a_1, z_2).

Die Nutzenmatrix $[u_{ij}]$ sei somit gegeben durch

	z_1	z_2
a_1	0	-1
a_2	-3	0

Die in Intervallform gegebene Menge der A-priori-Verteilungen sei

$$\{\varrho\,(P) = \varrho\,(p_1, p_2)\,|\,\frac{3}{8} \leqq p_1 \leqq \frac{5}{8}\,, p_2 = 1 - p_1\}\ .$$

Die Zerlegung von P wählen wir symmetrisch um die Intervallmitte, in der Hoffnung, damit die *LPI* optimal aufnahmebereit für die Beobachtung zu machen:

$$P = P_1 + P_2 = \left[\frac{3}{8}, \frac{4}{8}\right] + \left[\frac{4}{8}, \frac{5}{8}\right]\ .$$

Die Beobachtung σ stelle einen Konjunkturindikator dar, er falle „gut" aus. Daraus folgt:

$$\sigma:\quad W\,(P_1) \leqslant W\,(P_2) \Rightarrow t \leqslant 1 - t \Rightarrow 0 \leqslant t \leqslant \frac{1}{2}\ .$$

Die Menge der A-posteriori-Verteilungen ist

$$\varrho\,(P)_N = \{\varrho\,|\,\varrho\varepsilon\,[\gamma, \delta]\}\ ,$$

wobei γ und δ die beiden Intervallgrenzen sind. Wir finden sie numerisch, indem wir bilden

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

$$\gamma = \min_{t \varepsilon [0, \frac{1}{2}]} \left[\frac{3}{8} t + (1 - t) \frac{4}{8} \right] = \frac{7}{16}$$

$$\gamma = \max_{t \varepsilon [0, \frac{1}{2}]} \left[\frac{4}{8} t + (1 - t) \frac{5}{8} \right] = \frac{10}{16} ,$$

woraus für $\varrho (P)_N$ folgt

$$\varrho (P)_N = \left[\frac{7}{16}, \frac{10}{16} \right] .$$

Die beiden Spielmatrizen B_v und B_N ergeben sich durch Nachmultiplikation der Nutzenmatrix mit der Extremalpunktematrix

$$B_v = \begin{bmatrix} 0 & -1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{5}{8} & \frac{3}{8} \end{bmatrix} = \begin{bmatrix} -\frac{5}{8} & -\frac{3}{8} \\ -\frac{9}{8} & -\frac{15}{8} \end{bmatrix} \Rightarrow \left[a_1 ; -\frac{5}{8} \right]$$

$$B_N = \begin{bmatrix} 0 & -1 \\ -3 & 0 \end{bmatrix} \begin{bmatrix} \frac{7}{16} & \frac{10}{16} \\ \frac{9}{16} & \frac{6}{16} \end{bmatrix} = \begin{bmatrix} -\frac{9}{16} & -\frac{6}{16} \\ -\frac{21}{16} & -\frac{30}{16} \end{bmatrix} \Rightarrow \left[a_1 ; -\frac{9}{16} \right]$$

Beide Spiele weisen einen Sattelpunkt bei (a_1, z_1) auf und identifizieren die pessimistische Vorherbestimmung als optimale im Sinne des $\text{MaxE}_{\min}$-Prinzips (trotz des gut ausgefallenen Konjunkturindikators).

Beim Vergleich des A-priori-Spiels mit dem A-posteriori-Spiel zeigt sich, daß die Prognosebewertung von $-\frac{10}{16}$ auf $-\frac{9}{16}$ gestiegen ist. Der semantische Wert der Information σ für die Vorherbestimmung beläuft sich daher auf $\frac{1}{16}$.

Wir hoffen, mit diesem Beispiel sowohl das grundsätzliche Vorgehen der *LPI*-Methode als auch ihre Anwendung bei der Vorherbestimmung verdeutlicht zu haben.

7. Ausblick

Die Anwendung der *LPI*-Methodik ist natürlich nicht auf die Vorherbestimmung der exogenen Variablen beschränkt. Die zuvor dargestellte *LPI*-Inferenz richtet sich ja auch in erster Linie auf die Bestimmung der Prognoseverteilung, also der Verteilung der zukünftigen Realisationen der endogenen Variablen. Da wir in der Ökonometrie keine sog. direkten Prognosen machen können, hat jedes Prognoseproblem in der Ökonometrie natürlich eine Inferenzphase. Die *LPI*-Methode erlaubt aber nun, durch entsprechende Zerlegungen der Form (4), durch wie-

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

derholte Anwendungen der *LPI*-Inferenz nach Satz (5) und durch die Bemessung der Zerlegungen an zukunftsrelevanten Variablen die Inferenzphase ganz und gar auf die Prognose hin zu orientieren. Sie erweist sich damit, aber auch in anderer Hinsicht, als der weichen Modellbildung verwandt, jedenfalls dort, wo diese direkt auf die sog. Prädiktorspezifikation gerichtet ist (*Wold* 1975).

Doch ist die Behandlung des Prognoseproblems bei unvollständiger Information mit Hilfe der *LPI*-Methodik nicht auf die Inferenz und die Vorherbestimmung der exogenen Variablen beschränkt. Weitere Anwendungsmöglichkeiten liegen in der Behandlung fehlerbehafteter Variablen, und zwar stets dann, wenn die Verteilung der Meßfehler nicht genau bekannt ist, wenn man aber etwas über die Wahrscheinlichkeiten weiß, mit denen die Fehler auftreten können.

Sodann möchten wir darauf hinweisen, daß auch die Verbindung von Prognose und wirtschaftspolitischer Maßnahme ein fruchtbares Anwendungsfeld der *LPI*-Methodik darstellen. Die möglichen zukünftigen endogenen Variablenwerte spielen dann direkt die Rolle von Zuständen und die wirtschaftspolitischen Maßnahmen die Rolle von Strategien.

Die *LPI*-Methodik vermittelt auch eine Hilfe bei einer der wohl unangenehmsten Aufgaben im Rahmen entscheidungstheoretischer Problemlösungen (vgl. *Menges* 1974): Bei der Behandlung von Unbestimmtheiten in der Nutzenmatrix selbst. Sofern diese Unbestimmtheiten in stochastischer, linearer Form ausgedrückt werden können, führt das Modell der Zweipersonen-Nullsummenspiele in seiner *LPI*-Form zu einer Lösung auch dieses Problems, wie wir in unserem Beitrag zur Demaria-Festschrift gezeigt haben (*Menges/Kofler* 1978).

Der Grundgedanke ist hier wie bei allen Anwendungen der *LPI*-Methodik derselbe: Jedes A-priori-Wissen und jede empirische Kenntnis so weit wie nur irgend möglich auszunutzen und die verbleibende Lücke durch Entscheidungen zu schließen. Der Vorteil der *LPI*-Methodik liegt hauptsächlich darin, für jede Art von Information empfänglich zu sein und jede Information gemäß ihrer Schärfe bzw. Unschärfe zu verwenden.

Zusammenfassung

Wir widmen uns den praktisch sehr häufigen Fällen, in denen Prognosen auf Grund partieller Information aufzustellen sind. Partielle Information bedeutet, daß die Prognosewahrscheinlichkeiten a) unvollständig, b) unscharf oder c) nicht in zahlenmäßiger (kardinaler) Form gegeben sind. Die von uns vorgeschlagene Methodik der *LPI* (Lineare Partielle Information) erlaubt, jede Art von Information gemäß ihres Partialitätsgrades auszunutzen. Die je verbleibende Lücke wird durch Entscheidung geschlossen; hierfür steht das $\mathrm{MaxE}_{\mathrm{min}}$-Prinzip zur Verfügung. Die Methodik wird anhand einiger Anwendungsfälle, darunter einem numerischen, exemplifiziert.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/

Summary

We devote ourselves to the practically very frequent cases in which forecastings have to be made on the basis of partial information. Partial information means that the forecasting probabilities are a) incomplete, b) imprecise (fuzzy) or c) not given in quantitative (cardinal) forms. The methodology of *LPI* (Linear Partial Information) allows for the utilization of any kind of information according to its degree of indeterminateness. The ever remaining gaps are filled by decisions; here the $\text{Max}E_{\text{min}}$-principle is applicable. The methodology is illustrated by some examples including a numerical one.

Literatur

Blind, A. (1964), Das derzeitige Verhältnis zwischen Statistik und Nationalökonomie. In: Das Verhältnis der Wirtschaftswissenschaft zur Rechtswissenschaft, Soziologie und Statistik; Schriften des Vereins für Socialpolitik, N. F. Bd. 33, Berlin, S. 337 - 360.

Kofler, E. und G. *Menges* (1976), Entscheidungen bei unvollständiger Information. Springer-Verlag Berlin-Heidelberg- New York.

— (1977 a), Stochastic linearisation of indeterminateness. Towards a theory of games and decisions under partial information. In: R. Henn und O. Moeschlin (ed.), Mathematical Economics and Game Theory. Essays in Honor of Oskar Morgenstern. Springer-Verlag, Berlin-Heidelberg-New York, pp. 53 - 63.

— (1977 b), Die Strukturierung von Unbestimmtheiten und eine Verallgemeinerung des Axiomensystems von Kolmogoroff für unbestimmte Wahrscheinlichkeiten. Statistische Hefte, 18. Jg., Nr. 4, S. 298 - 302.

Menges, G. (1966), Statistik und Wirtschaftsprognose. In: A. Blind (Hrg.), Umrisse einer Wirtschaftsstatistik. Festgabe für Paul Flaskämper zur 80. Wiederkehr seines Geburtstages. Hamburg, S. 50 - 71.

— (1967), Ökonometrische Prognosen. Westdeutscher Verlag. Köln und Opladen.

— (1974), Wie gut sind Prognosen? Mitteilungen aus der Arbeitsmarkt- und Berufsforschung. 7. Jg., Nr. 3, S. 242 - 250.

Menges, G. and E. *Kofler* (1978), Linear stochastic structuring of indeterminate decision matrices. In: T. Bagiotti and G. Franco (ed.), Pioneering Economics. International Essays in Honour of Giovanni Demaria, Padova, pp. 691 - 699.

Wold, H. (1975), Soft modelling by latent variables: The non-linear iterative partial least squares (NIPALIS) approach. In: Perspectives in Probability and Statistics. J. Gami (ed.), New York, pp. 117 - 142.

DOI https://doi.org/10.3790/schm.100.1.1 | Generated on 2025-07-26 01:38:22
OPEN ACCESS | Licensed under CC BY 4.0 | https://creativecommons.org/about/cclicenses/