

ÜBER DIE GEWINNUNG EINER
PROPORTIONSMATRIX
AUS EINER
RANGFREQUENZMATRIX

von

Jürgen KRIZ

Forschungsbericht No. 10

Dezember 1967

Ich danke Herrn Dipl.Ing. Horst WEGSCHEIDER
für zahlreiche Hinweise und die sorgfältige
Durchsicht des Manuskriptes unter dem
Gesichtspunkt eines Mathematikers.

Einleitung

1931 wurde von THURSTONE ein Verfahren veröffentlicht, das gestattet, eine Proportionsmatrix direkt aus einer Rangfrequenzmatrix zu berechnen¹ (exakt gesagt: zu schätzen). Dies Verfahren wird u.A. auch von GUILFORD 1954 und SIXTL 1967 beschrieben und ist heute ein durchaus gängiges Verfahren in der Skalierung.

In dem Verfahren von THURSTONE werden Wahrscheinlichkeiten als unabhängig behandelt (und daher multipliziert), von denen gezeigt werden kann, daß sie abhängig sind.

Dagegen soll ein Verfahren vorgeschlagen werden, das die Abhängigkeit dieser Wahrscheinlichkeiten berücksichtigt. Für jedes $P_{j>i}$ der Proportionsmatrix wird dabei eine obere und eine untere Grenze exakt bestimmt, und dann der Wert für $P_{j>i}$ in diesem Intervall geschätzt. Das vorgeschlagene Verfahren ist weniger aufwendig als das von THURSTONE. Ferner gestattet es eine direkte Angabe des maximalen Fehlers für jedes geschätzte $P_{j>i}$.

¹Die Begriffe werden auf den folgenden Seiten erläutert.

Die Methode der Rangordnungen.

Bei diesem Verfahren haben N Vpn (oder eine Vp N mal) die Aufgabe, m Objekte hinsichtlich ihres Ausprägungsgrades auf einer Dimension (z.B. Schönheit) in eine Rangreihe zu bringen. Wesentlich dabei ist, daß genau so viele Ränge verteilt werden müssen, wie Objekte vorhanden sind, d.h. die Bildung von "ties" ist nicht erlaubt.

Aus jeder Rangreihe lassen sich $\binom{m}{2}$ Ordinalurteile über die Objekte gewinnen. Z.B. folgt aus: $A > C > B > D$:
 $A > C$, $A > B$, $A > D$, $C > B$, usw. Die N Vpn liefern also insgesamt $N \binom{m}{2}$ Ordinalurteile. Diese kann man in einer $m \times m$ Matrix zusammenfassen, wobei in der i -ten Zeile und j -ten Spalte die Häufigkeit ($f_{j > i}$) steht, mit der $j > i$ geurteilt wurde (maximal N). Eine solche Matrix heißt kombinierte Dominanzmatrix. In dieser Arbeit wird als Proportionsmatrix eine kombinierte Dominanzmatrix bezeichnet, in deren Zellen nicht absolute, sondern relative Häufigkeiten stehen.

Wesentlich weniger aufwendig als $N \binom{m}{2}$ Ordinalurteile zu gewinnen und zu einer kombinierten Dominanzmatrix zusammenzustellen ist es, auszuzählen, wie oft dem Objekt j der Rang i zugeteilt wurde. Wenn dies für alle Objekte und alle Ränge getan wird, lassen sich die Ergebnisse wieder in einer $m \times m$ Matrix zusammenfassen. In der i -ten Zeile und j -ten Spalte dieser Matrix steht die Häufigkeit, mit der dem Objekt j der Rang i zugeteilt wurde. Eine solche Matrix heißt Rangfrequenzmatrix.

Um Skalenwerte für die Objekte nach THURSTONE's "Law of comparative Judgment" zu berechnen, ist es notwendig, eine Angabe über $P_{j > i}$ zu haben. Diese ist in der Proportionsmatrix direkt enthalten, in der Rangfrequenzmatrix dagegen nicht. Daher entwickelte THURSTONE 1931 ein

Verfahren, das eine Angabe über $P_{j>i}$ aus einer Rangfrequenzmatrix erlaubt. Betont sei an dieser Stelle, daß es sich um ein Schätzverfahren handelt, da $P_{j>i}$ nicht mehr eindeutig aus der Rangfrequenzmatrix reproduzierbar ist (wenn man von ganz speziellen Matrizen - z.B. der Diagonalmatrix - absieht).

Darstellung und Kritik von THURSTONE's Methode

Das Verfahren von THURSTONE wird hier nur soweit, wie es für die Kritik notwendig ist, dargestellt - ansonsten sei auf die Originalarbeit von THURSTONE verwiesen. THURSTONE argumentiert folgendermaßen:

"Let there be n specimens in the series to be arranged in rank order by N subjects. Let A und B be two of these specimens and let a_1 = frequency with which specimen A is placed in rank 1 by the N subjects, P_{a1} = proportion of the N subjects who place specimen A in rank 1, P_{b1} = proportion who place specimen B in rank 1, and similarly for the other specimens and the other rank orders. Since these values of P may be regarded as probabilities, we have $P_{b2} + P_{b3} + P_{b4} + \dots + P_{bn} = P_{b>1}$ = probability that any subject at random (or any one judgment of a single subject) will place B in a rank higher than rank 1. Hence $P_{a1} \cdot P_{b>1}$ = probability that any subject, chosen at random, will place A in rank 1, and B in a higher rank. Similarly, $P_{a2} \cdot P_{b>2}$ = probability that a subject will place A in rank 2 and B in a higher rank.

In general, this product may be written $P_{ak} \cdot P_{b>k}$ = probability that A will be perceived in rank k and that B will be perceived in a higher rank than k ."

Der Satz, der für unsere Kritik am wesentlichsten ist, wurde im Text unterstrichen: P_{ak} und $P_{b>k}$ miteinander zu multiplizieren ist nur sinnvoll, wenn beide Wahrscheinlichkeiten stochastisch unabhängig sind. Das ist hier aber nicht der Fall. Um dies leichter zu erläutern, betrachten wir folgendes - scheinbar analoge - Beispiel:

Man habe zwei Würfel A und B. Die Wahrscheinlichkeit P_{a1} mit A eine "1" zu würfeln, ist $1/6$. Die Wahrscheinlichkeit $P_{b>1}$ mit B eine Zahl größer als "1" zu würfeln, ist $5/6$. Somit ist die Wahrscheinlichkeit mit A eine "1" und mit B eine Zahl größer als "1" zu würfeln $= P_{a1} \cdot P_{b>1} = 1/6 \cdot 5/6 = 5/36$.

Im Fall der Rangfrequenzmatrix ist P_{a1} die Wahrscheinlichkeit, mit der A der Rangplatz "1" zugeordnet und $P_{b>1}$ die Wahrscheinlichkeit, mit der B ein Rangplatz größer als "1" zugeordnet wird. Der wesentliche Unterschied zu dem Würfelbeispiel besteht aber darin, daß man weiß, daß jede Vpn, die A den Rangplatz "1" zugeteilt hat, automatisch B einen höheren Rangplatz geben muß (denn "ties" sind ausgeschlossen). Es gilt nun:

$$(1) \quad P(\{r(A)=1\} \cap \{r(B) > 1\}) = P(r(A)=1) \cdot P(r(B) > 1 / r(A)=1)$$

da (2) $P(r(B) > 1 / r(A)=1) = 1$

gilt (3) $P(\{r(A)=1\} \cap \{r(B) > 1\}) = P(r(A)=1)$

daraus folgt:

Die Wahrscheinlichkeit, daß A der Rangplatz "1" und B ein höherer zugeordnet wird, ist P_{a1} und nicht $P_{a1} \cdot P_{b>1}$. D.h. sofern $P_{b>1} \neq 1$ ist, unterschätzt THURSTONE die Wahrscheinlichkeit.

Für Rangplatz "2" läßt sich nicht mehr so argumentieren. Sofern $P_{b1} \neq 0$ ist, müssen nicht alle Vpn, die A den Rangplatz "2" gaben, unbedingt B einen höheren geben. Sofern aber $P_{a2} > P_{b1}$ ist, müssen mindestens $(P_{a2} - P_{b1}) \cdot N$ Vpn A den Rangplatz "2" und B einen höheren gegeben haben (denn nur $P_{b1} \cdot N$ Vpn konnten B einen niedrigeren gegeben haben).

Es zeigt sich also, daß die Wahrscheinlichkeiten P_{ak} und $P_{b>k}$ nicht unabhängig voneinander sind und daher das Produkt dieser beiden Wahrscheinlichkeiten nicht die Wahrscheinlichkeit des gemeinsamen Auftretens (Produktereignisses) liefert.

Das Verfahren des Unsicherheitsintervalles

Das Verfahren geht von der Tatsache aus, daß sich aus einer Rangfrequenzmatrix zwar nicht mehr eindeutig die Rangordnungen aller V_{pn} - und somit $P_{a>b}$ - ermitteln läßt (von Sonderfällen abgesehen), wohl aber ein Teil der Aussagen eindeutig rekonstruierbar ist. Ein Beispiel soll dies veranschaulichen:

Nehmen wir an, 10 V_{pn} ranken 3 Bilder A, B und C hinsichtlich ihrer Schönheit, wobei 3 der höchste Rang sein soll (d.h. dem schönsten der drei Bilder zugeordnet werden soll). Die Ergebnisse könnten etwa so aussehen:

V_p	A	B	C
1	1	3	2
2	1	3	2
3	1	2	3
4	2	3	1
5	1	3	2
6	2	1	3
7	3	2	1
8	1	3	2
9	2	3	1
10	3	2	1

Die Rangfrequenzmatrix dieser Daten sieht so aus:

Rang	A	B	C
1	5	1	4
2	3	3	4
3	2	6	2

Aus dieser Rangfrequenzmatrix lassen sich nun folgende Aussagen gewinnen:

- 1) 5 Vpn haben A den Rangplatz "1" gegeben. Diese müssen B einen höheren Rang gegeben haben. 3 Vpn gaben A den Rangplatz "2"; da nur 1 Vp B den Rangplatz "1" gegeben hat, müssen von diesen 3 Vpn mindestens 2 Vpn B einen höheren Rangplatz gegeben haben. Es haben also mindestens 7 Vpn B \succ A geurteilt.

Ähnlich ergeben sich die weiteren Aussagen:

- 2) mindestens 2 Vpn haben A \succ B geurteilt
- 3) " 4 " " C \succ A "
- 4) " 3 " " A \succ C "
- 5) " 2 " " C \succ B "
- 6) " 7 " " B \succ C "

Diese Aussagen lassen sich folgendermaßen kombinieren:

$$\begin{array}{l} \text{I} \quad 2 \leq f_{a \succ b} \leq 3 \quad ; \quad 7 \leq f_{b \succ a} \leq 8 \\ \text{II} \quad 3 \leq f_{a \succ c} \leq 6 \quad ; \quad 4 \leq f_{c \succ a} \leq 7 \\ \text{III} \quad 7 \leq f_{b \succ c} \leq 8 \quad ; \quad 2 \leq f_{c \succ b} \leq 3 \end{array}$$

Formalisieren wir unsere bisherigen Überlegungen, so ergibt sich (wie sich leicht zeigen läßt):

$$(4) \quad \text{Min}(f_{j \succ i}) = \text{Max}_{0 \leq k \leq m} \left\{ f_{\text{cum}}(k, i) - f_{\text{cum}}(k-1, j) \right\}$$

$$\begin{aligned} (5) \quad \text{Max}(f_{j \succ i}) &= N - \text{Min}(f_{i \succ j}) \\ &= N - \text{Max}_{0 \leq k \leq m} \left\{ f_{\text{cum}}(k, j) - f_{\text{cum}}(k-1, i) \right\} \end{aligned}$$

wobei gelten soll: $f_{\text{cum}}(0,j) = f_{\text{cum}}(0,k) = 0$

f_{ki} = Wert der Rangfrequenzmatrix in Zeile k
und Spalte i

$f_{\text{cum}}(k,i)$ = $\sum_{z=1}^k f_{zi}$

$\text{Min}(f_{j>i})$ = Kleinstes Element der Menge, die aus
den Elementen $(f_{j>i})$ aller kombinierten
Dominanzmatrizen besteht, die aus der
Rangfrequenzmatrix gebildet werden
können.

Es läßt sich also für jedes $f_{j>i}$ eine obere und eine
untere Grenze bestimmen, zwischen denen der "wahre" Wert
liegt. Die untere Grenze stellt die minimale, die obere
die maximale Anzahl Vpn dar, die $j>i$ geurteilt haben
können. Die Länge des Unsicherheitsintervalles
($N - \text{Min}(f_{i>j}) - \text{Min}(f_{j>i})$) entspricht der Anzahl
der Vpn, über die keine Aussage über die Beziehung
zwischen i und j mehr reproduziert werden kann.

Die Summe $D = \text{Min}(f_{i>j}) + \text{Min}(f_{j>i})$ entspricht somit der
Anzahl der Vpn, deren Urteil hinsichtlich i und j
rekonstruierbar ist. Es läßt sich leicht zeigen, daß bei N
Vpn und m zu beurteilenden Reizen diese Summe D mindestens
den Wert $2N/m$ haben muß, (und somit die Intervallbildung
sinnvoll ist).

Beweis: (angedeutet)

Wir greifen aus einer Rangfrequenzmatrix mit m Objekten
($m > 2$) 2 Spalten A und B heraus. Sind die Werte in dieser
Submatrix über die Ränge gleichverteilt, ergibt sich:

Rang	A	B
1	N/m	N/m
2	N/m	N/m
.	.	.
.	.	.
k	N/m	N/m
.	.	.
.	.	.
m	N/m	N/m

Es gilt:

$$N/m \leq f_{a > b} \leq N - N/m ; \quad N/m \leq f_{b > a} \leq N - N/m ; \quad D = 2N/m$$

Gezeigt wird, daß D nur größer werden kann, wenn die Werte von einer Gleichverteilung über die Ränge abweichen: Stände in Spalte B und Zeile k der Wert $N/m + \Delta N$, so muß, damit die Spaltensumme N erhalten bleibt, der Wert in mindestens einer, höchstens allen weiteren Zellen der Spalte B kleiner werden. Es gilt somit: ($\Delta N > 0$):

$$(6) \quad f_{\text{cum}}(k-1, b) = (k-1)(N/m) - v \Delta N \quad 0 \leq v \leq 1$$

nach (4) ergibt sich:

$$\begin{aligned} \text{Min}(f_{a > b}) &= f_{\text{cum}}(k, b) - f_{\text{cum}}(k-1, a) \\ &= (k-1)(N/m) - v \Delta N + N/m + \Delta N - (k-1)(N/m) \\ &= N/m + (1-v) \Delta N \end{aligned}$$

$$\begin{aligned} \text{Min}(f_{b > a}) &= f_{\text{cum}}(k, a) - f_{\text{cum}}(k-1, b) \\ &= (k-1)(N/m) + N - (k-1)(N/m) - v \Delta N \\ &= N/m + v \Delta N \end{aligned}$$

$$\underline{D = 2N/m + \Delta N}$$

D ist somit (unabhängig von v) um ΔN größer geworden. Stände in Spalte B und Zeile k der Wert $N/m - \Delta N$, gilt analog; ($\Delta N > 0$):

$$(6a) \quad f_{\text{cum}}(k-1, b) = (k-1)(N/m) + v \Delta N \quad ; \quad 0 \leq v \leq 1$$

nach (4) ergibt sich:

$$\begin{aligned} \text{Min}(f_{b > a}) &= f_{\text{cum}}(k+1, a) - f_{\text{cum}}(k, b) \\ &= (k+1)(N/m) - k(N/m) - \Delta N + v \Delta N \\ &= N/m + (1-v) \Delta N \end{aligned}$$

$$\text{wobei } f_{\text{cum}}(m+1, a) = f_{\text{cum}}(m, a) = N$$

$$\begin{aligned} \text{Min}(f_{a > b}) &= f_{\text{cum}}(k-1, b) - f_{\text{cum}}(k-2, a) \\ &= (k-1)(N/m) + v \Delta N - (k-2)(N/m) \\ &= N/m + v \Delta N \end{aligned}$$

$$\text{wobei } f_{\text{cum}}(-1, a) = f_{\text{cum}}(0, a) = 0$$

$$\underline{D = 2N/m + \Delta N}$$

D wird also in jedem Falle größer. q.e.d.

Zur Bestimmung von $P_{a > b}$

Wir stellten fest, daß für $f_{a > b}$ eine obere und untere Grenze berechnet werden und die Unsicherheitsintervalllänge maximal $N - 2N/m$ betragen kann. Diese Feststellungen sind auch dann noch korrekt, wenn man die absoluten Häufigkeiten in relative verwandelt (indem durch N dividiert wird) und damit die dazugehörigen Wahrscheinlichkeiten schätzt. Man erhält somit für $P_{a > b}$ eine obere und eine

untere Grenze, die Länge des Intervalles kann maximal $1-2/m$ betragen.

Dies erlaubt uns, eine Aussage über den maximalen Fehler bei der Schätzung von $P_{a>b}$ zu machen, denn es gilt:

$$(7) \quad E_{\max} = \frac{P_{\max} - P_{\min}}{2} + \left| \frac{P_{\max} + P_{\min}}{2} - \hat{P} \right|$$

E_{\max} = maximaler Fehler

P_{\max} = obere Grenze für $P_{a>b}$

P_{\min} = untere Grenze für $P_{a>b}$

\hat{P} = geschätzter $P_{a>b}$ -Wert

wobei es offenbar gleichgültig ist, nach welchem Verfahren \hat{P} geschätzt wurde, d.h. auf diese Weise ließe sich auch der maximale Fehler für jeden Wert, der nach THURSTONE's Verfahren erhalten wird, bestimmen.

Nach welchem Verfahren nun \hat{P} innerhalb des beschriebenen Unsicherheitsintervalles bestimmt wird, soll in dieser Arbeit nicht festgelegt werden. Es erscheint dem Autor besser, drei Verfahren, die sich besonders anbieten, aufzuführen:

1) Methode des minimalen maximalen Fehlers.

\hat{P} wird als arithmetischer Mittelwert von P_{\max} und P_{\min} bestimmt, also $\hat{P} = (P_{\max} + P_{\min})/2$. Diese Methode hat neben ihrem geringen Aufwand den Vorteil, daß der maximale Fehler minimiert wird. Diejenigen V_{pn} , über die keine Aussage gemacht werden konnte, werden durch diese Methode so behandelt, als ob für $P_{a>b}$ und $P_{b>a}$ die gleiche

Wahrscheinlichkeit bestände, was um so zweifelhafter ist, je mehr P von .50 abweicht. Die Streuung der \hat{P} -Werte um .50 wird somit geringer sein, als die der "wahren" P-Werte.

2) Methode der repräsentativen Erfassung

Betrachtet man diejenigen Vpn, deren Aussagen hinsichtlich a und b rekonstruiert werden konnten, als repräsentativ für alle Vpn, ergibt sich \hat{P} aus

$$P_{a > b} = \frac{\text{Min}(f_{a > b})}{\text{Min}(f_{a > b}) + \text{Min}(f_{b > a})}$$

Sind $\text{Min}(f_{a > b})$ und $\text{Min}(f_{b > a})$ klein, so bewirkt schon ein geringer Unterschied zwischen beiden bei dieser Methode eine große Abweichung vom Mittelwert ($P = .50$). Die Streuung der \hat{P} -Werte wird somit bei dieser Methode größer sein, als die der "wahren" P-Werte.

3) Methode des z-Mittelwertes

Im Gegensatz zum 1. Verfahren wird hier das arithmetische Mittel der zu P_{\max} und P_{\min} gehörigen z-Werte gebildet. \hat{P} ist dann der zum gemittelten z-Wert dazugehörige P-Wert. (Für $P_{\min} = 0$ bzw. $P_{\max} = 1$ wird z willkürlich -3 bzw. +3 gesetzt) Die Streuung der \hat{P} -Werte ist größer als im 1. und kleiner als im 2. Verfahren.

Zusammenfassung

Es konnte gezeigt werden, daß das Verfahren von THURSTONE zur Gewinnung einer Proportionsmatrix aus einer Rangfrequenzmatrix abhängige Wahrscheinlichkeiten als unabhängig behandelt und daher nicht ganz einwandfrei ist.

Es wurde ein Verfahren vorgeschlagen, das die Abhängigkeit der Wahrscheinlichkeiten berücksichtigt. Das Verfahren gestattet, für je zwei Objekte j und i der Rangfrequenzmatrix exakt ein Intervall anzugeben, in dem der "wahre" $P_{j > i}$ -Wert liegen muß.

Für die Schätzung von $P_{j > i}$ in diesem Intervall wurden drei Methoden kurz angeführt. Der Autor ist jedoch der Meinung, daß es noch einer befriedigenderen Lösung für dieses Problem bedarf.

Literatur

GUILFORD, J.P.

"Psychometric Methods"
McGraw-Hill Book Company,
New York, Toronto, London, 1954 ²

SIXTL, F.

"Meßmethoden der Psychologie"
Verlag Julius Beltz, Weinheim 1967

THURSTONE, L.L.

"Rank order as a Psychometric Method", J.Exp.Psychol. 14, No.3,
June 1931, 187-201