



**Metrisches multidimensionales Unfolding  
Metrisches MDU**

**P34.7**

**Kurt Holm**

[www.almo-statistik.de](http://www.almo-statistik.de)  
[holm@almo-statistik.de](mailto:holm@almo-statistik.de)  
[kurt.holm@jku.at](mailto:kurt.holm@jku.at)

**2015**

## Weitere Almo-Dokumente

Die folgenden Dokumente können alle kostenlos von der Handbuchseite in

[www.almo-statistik.de](http://www.almo-statistik.de)

heruntergeladen werden

0. Arbeiten\_mit\_Almo.PDF (1 MB)
- 1a. Eindimensionale Tabellierung.PDF (1.8 MB)
- 1b. Zwei- und drei-dimensionale Tabellierung.PDF (1.1 MB)
2. Beliebig-dimensionale Tabellierung.PDF (1.7 MB)
3. Nicht-parametrische Verfahren.PDF (0.9 MB)
4. Kanonische Analysen.PDF (1.8 MB)  
Diskriminanzanalyse.PDF (1.8 MB)  
enthält: Kanonische Korrelation, Diskriminanzanalyse, bivariate Korrespondenzanalyse, optimale Skalierung
5. Korrelation.PDF (1.4 MB)
6. Allgemeine multiple Korrespondenzanalyse.PDF (1.5 MB)
7. Allgemeines ordinale Rasch-Modell.PDF (0.6 MB)
- 7a. Wie man mit Almo ein Rasch-Modell rechnet.PDF (0.2 MB)
8. Tests auf Mittelwertsdifferenz, t-Test.PDF (1,6 MB)
9. Logitanalyse.pdf (1,2MB) enthält Logit- und Probitanalyse
10. Koeffizienten der Logitanalyse.PDF (0,06 MB)
11. Daten-Fusion.PDF (1,1 MB)
12. Daten-Imputation.PDF (1,3 MB)
13. ALM Allgemeines Lineares Modell.PDF (2.3 MB)
- 13a. ALM Allgemeines Lineares Modell II.PDF (2.7 MB)
14. Ereignisanalyse: Sterbetafel-Methode, Kaplan-Meier-Schätzer, Cox-Regression.PDF (1,5 MB)
15. Faktorenanalyse.PDF (1,6 MB)
16. Konfirmatorische Faktorenanalyse.PDF (0,3 MB)
17. Clusteranalyse.PDF (3 MB)
18. Pisa 2012 Almo-Daten und Analyse-Programme.PDF (17 KB)
19. Guttman- und Mokken-Skalierung.PFD (0.8 MB)
20. Latent Structure Analysis.PDF (1 MB)
21. Statistische Algorithmen in C (80 KB)
22. Conjoint-Analyse (PDF 0,8 MB)
23. Ausreisser entdecken (PDF 170 KB)
24. Statistische Datenanalyse Teil I, Data Mining I
25. Statistische Datenanalyse Teil II, Data Mining II
26. Statistische Datenanalyse Teil III, Arbeiten mit Almo-Datenanalyse-System
27. Mehrfachantworten, Tabellierung von Fragen mit Mehrfachantworten (0.8 MB)
28. Metrische multidimensionale Skalierung (MDS) (0,5 MB)
29. Metrisches multidimensionales Unfolding (MDU) (0,6 MB)
30. Nicht-metrische multidimensionale Skalierung (MDS) (0,5 MB)
31. Pfadanalyse.PDF (0,7 MB)
32. Datei-Operationen mit Almo (1,1 MB)

## Inhaltsverzeichnis

<i>P34.7.1 Überblick, Beispiele einer metrischen MDU</i> .....	4
P34.7.1.1 Die Distanzmatrix $D$ .....	4
P34.7.1.2 Die Ladungsmatrix $F$ .....	5
P34.7.1.3 Die Grafik der Ladungsmatrix $F$ .....	6
P34.7.1.4 Die gesamte reproduzierte Distanzmatrix $gD^*$ .....	7
<i>P34.7.2 Kalkül der MDU</i> .....	8
P34.7.2.1 Die quadriert und doppelt zentrierte Distanzmatrix $X$ .....	8
P34.7.2.2 Die Singulärwert-Zerlegung von $X$ .....	9
P34.7.2.3 Die "vorläufige" MDU-Lösung.....	9
P34.7.2.4 Die endgültige MDU-Lösung durch den Schönemann-Kalkül.....	10
P34.7.2.5 Die "Güte" der MDU-Lösung - gemessen durch den Stress-Koeffizienten .	12
<i>P34.7.3 Eingabe in Almo-Programm Prog18m6</i> .....	13
P34.7.3.1 Erläuterungen zu den Eingabe-Boxen.....	16
P34.7.3.2 Daten für die MDU .....	18
P34.7.3.3 Problem: Spaltenzahl größer Zeilenzahl .....	25
<i>P34.7.4 Eingabe in Almo-Programm Prog18m7</i> .....	28
<i>P34.7.5 Ausgabe aus Prog18m6 und Prog18m7</i> .....	28
<i>P34.7.6 MDU-Analyse einer perfekten, fehlerfreien Distanzmatrix</i> .....	35
<i>P34.7.7 Vergleich Almo mit Alscal und Prefscal</i> .....	35
<i>P34.7.8 Anhang</i> .....	40
P34.7.6.1 Notation: Matrix-Bezeichnungen.....	40
P34.7.6.1 Singulärwert-Zerlegung einer rechteckigen nicht-symmetrischen Matrix ..	40
<i>Literatur</i> .....	41

### P34.7.1 Überblick, Beispiele einer metrischen MDU

Betrachten wir ein Beispiel. Eine Stichprobe von Personen wird darüber befragt, wie sympathisch sie die 5 Parteien ParteiA bis ParteiE finden. Die Personen werden nach Geschlecht und 4 Bildungsstufen in 8 Gruppen zusammengefasst. Für jede Gruppe wird der Mittelwert ihrer Sympathie gegenüber den 5 Parteien ermittelt.

Die MDU verlangt, dass Distanzen und nicht Präferenzen bzw. Sympathien als Daten in ihren Kalkül eingegeben werden. Die Sympathien gegenüber den Parteien wurden mit 1 bis 9 gemessen. Sie werden nun sehr einfach in Distanzen umgewandelt, indem sie "umgedreht", d.h. von 9 subtrahiert werden. Dann wird noch der kleinste Wert addiert. Es kann auch irgend eine andere Zahl verwendet werden, sie muss nur größer sein als der höchste Sympathie-Wert. Negative Distanzen dürfen jedoch nicht entstehen. Also addiert in diesem Falle automatisch einen Wert hinzu.

#### P34.7.1.1 Die Distanzmatrix *D*

So entsteht folgende Distanzmatrix **D**

Partei					Personen-Gruppe
A	B	C	D	E	PG
V1	V2	V3	V4	V5	V6
---	---	---	---	---	-----
4.0	3.0	3.0	4.0	6.0	1
3.0	9.0	9.0	2.0	4.0	2
3.0	5.0	6.0	4.0	2.0	3
2.0	4.0	4.0	3.0	3.0	4
4.0	2.0	3.0	5.0	5.0	5
4.0	1.0	2.0	5.0	5.0	6
3.0	9.0	9.0	3.0	2.0	7
4.0	6.0	7.0	4.0	2.0	8

Die Distanzmatrix ist Peter Schönemann (1970) entnommen, der wesentlich die metrische MDU mitentwickelt hat und dessen Kalkül in Almo verwendet wird.

Die MDU ist im Bereich der mathematisch-statistischen Sozialwissenschaften entwickelt worden. Sie ist jedoch ein allgemeines Verfahren, das überall da angewendet werden kann, wo zwischen zwei verschiedenen Objektgruppen Distanzinformationen vorliegen.

Die MDU benötigt als Eingabe nur die Distanzmatrix zwischen den Zeilen-Variablen (den 8 Personengruppen) und den Spalten-Variablen (den 5 Parteien). Wie diese entstand, ist für den Kalkül irrelevant. Die 6. Spalte (die Personengruppe, kurz: PG) haben wir zur Illustration hinzugefügt. In Almo ist es möglich,

1. aus Daten eine Distanz- bzw. Präferenzmatrix zu formen und für diese eine MDU zu rechnen - wobei die Daten selbst schon eine Distanz- bzw. Präferenzmatrix sein können.

Das geschieht mit der Programm-Maske Prog18m6

2. in einem Rechengang,

- (a) Personen nach bestimmten Merkmalen zu gruppieren,
- (b) je Personengruppe den Mittelwert ihrer Präferenzen bzw. Distanzen zu ermitteln, auf diese Weise die Distanzmatrix zu bilden und
- (c) dann eine MDU zu rechnen.

Das geschieht mit der Programm-Maske Prog18m7

Zur Terminologie:

*Spaltenobjekte* oder *Spaltenvariable* = das sind die Spalten der Distanzmatrix, im Beispiel: die 5 Parteien, gegenüber denen Präferenz-Urteile abgegeben werden.

*Zeilenobjekte* oder *Zeilenvariable* = das sind die Zeilen der Distanzmatrix, im Beispiel: die 8 Personengruppen, die eine Präferenz gegenüber den 5 Parteien äußern.

### ***P34.7.1.2 Die Ladungsmatrix F***

Für obige Distanzmatrix liefert ALMO folgende Ladungsmatrix **F** als Ergebnis, das hier zunächst gekürzt wiedergegeben wird.

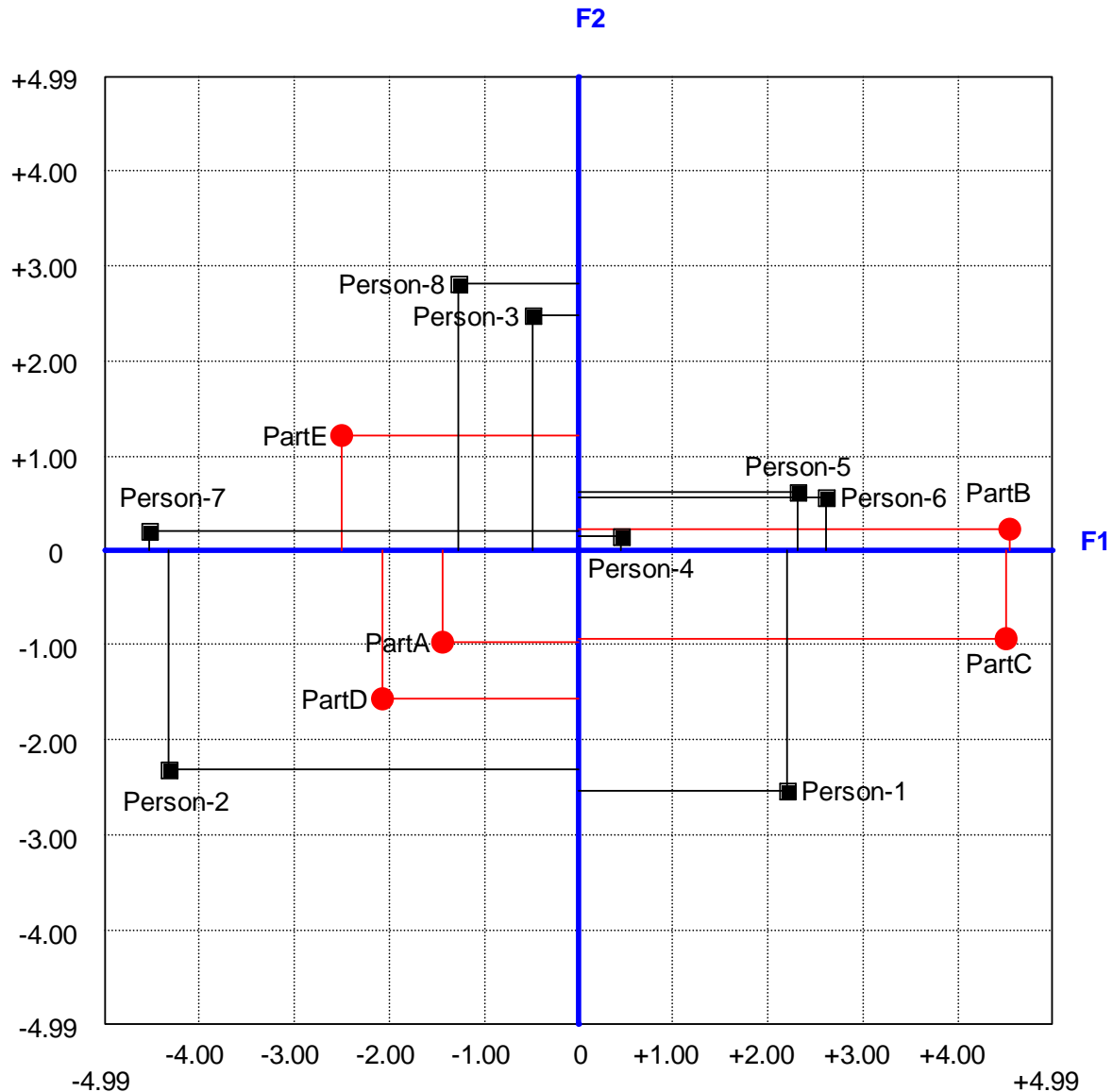
gemeinsame Matrix **F**  
aus Spaltenobjekten (Zeile 1 bis 5)  
und Zeilenobjekten (Zeile 6 bis 13)

	Faktor 1	Faktor 2
PartA	-1.4382	-0.9737
PartB	4.5376	0.2255
PartC	4.5017	-0.9314
PartD	-2.0639	-1.5698
PartE	-2.5055	1.2202
Person-1	2.1970	-2.5389
Person-2	-4.3112	-2.3061
Person-3	-0.4993	2.4875
Person-4	0.4525	0.1589
Person-5	2.3104	0.6234
Person-6	2.6130	0.5715
Person-7	-4.5262	0.2137
Person-8	-1.2678	2.8193

Ähnlich der Faktorenanalyse erhält man eine *Ladungsmatrix*, die dann noch grafisch als 2-dimensionales Koordinatensystem dargestellt wird

### P34.7.1.3 Die Grafik der Ladungsmatrix F

Ladungen der Spaltenvariablen (rote Punkte)  
und der Zeilenvariablen (schwarze Punkte)



Die Distanz-Urteile der 8 Personengruppen gegenüber den 5 Parteien lassen sich also in 2 Dimensionen arithmetisch und grafisch darstellen. Das ist Sinn und Zweck der MDU. Wir können z.B. erkennen, dass die Personengruppe 1 (etwa junge, hochgebildete Personen) eine große Distanz zur Partei E haben. Am dichtesten liegen sie bei Partei C. Informationen dieser Art sind natürlich auch aus der Distanzmatrix ablesbar. Durch die grafische Darstellung werden sie jedoch veranschaulicht. Außerdem leistet die MDU (vergleichbar der Faktorenanalyse) eine dimensionale Datenreduktion. In der Regel, aber nicht immer, sind die gefundenen Dimensionen inhaltlich benennbar. In unserem Beispiel könnte vielleicht die eine Dimension als die "Links-Rechts-Dimension" und die zweite als geographische Nord-Süd-

Dimension bezeichnet werden.

Zur Terminologie: Wir verwenden die Begriffe *Dimension* und *Faktor* synonym.

#### **P34.7.1.4 Die gesamte reproduzierte Distanzmatrix $gD^*$**

Die Leistung der MDU geht aber noch weiter. Aus der Grafik werden auch die Distanzen zwischen den Parteien ersichtlich und auch zwischen den Personengruppen. So erkennt man z.B. dass Partei C und E am weitesten voneinander entfernt sind. Bei den Personen sieht man, dass Person-1 und Person-8 am weitesten voneinander entfernt sind. Das sind Informationen, die aus der Distanzmatrix nicht unmittelbar ersichtlich werden.

Wir können nun aus der oben angegebenen Ladungsmatrix sehr einfach die Distanzen reproduzieren. Also liefert folgende *gesamte reproduzierte Distanzmatrix  $gD^*$*

	<b>I (=DB)</b>					<b>II (=D*')</b>							
	ParA	ParB	ParC	ParD	ParE	PG-1	PG-2	PG-3	PG-4	PG-5	PG-6	PG-7	PG-8
ParA	0	6.095	5.940	0.864	2.440	3.958	3.167	3.586	2.204	4.075	4.336	3.308	3.797
ParB	6.095	0	1.157	6.841	7.113	3.622	9.204	5.522	4.086	2.262	1.955	9.064	6.358
ParC	5.940	1.157	0	6.597	7.330	2.810	8.919	6.058	4.193	2.687	2.414	9.100	6.881
ParD	0.864	6.841	6.597	0	2.825	4.370	2.365	4.349	3.053	4.893	5.144	3.040	4.461
ParE	2.440	7.113	7.330	2.825	0	6.020	3.962	2.373	3.143	4.853	5.159	2.258	2.022
PG-1	3.958	3.622	2.810	4.370	6.020	0	6.512	5.704	3.213	3.164	3.138	7.265	6.381
PG-2	3.167	9.204	8.919	2.365	3.962	6.512	0	6.124	5.364	7.241	7.498	2.529	5.961
PG-3	3.586	5.522	6.058	4.349	2.373	5.704	6.124	0	2.516	3.372	3.655	4.624	0.837
PG-4	2.204	4.086	4.193	3.053	3.143	3.213	5.364	2.516	0	1.915	2.199	4.979	3.168
PG-5	4.075	2.262	2.687	4.893	4.853	3.164	7.241	3.372	1.915	0	0.307	6.849	4.198
PG-6	4.336	1.955	2.414	5.144	5.159	3.138	7.498	3.655	2.199	0.307	0	7.148	4.485
PG-7	3.308	9.064	9.100	3.040	2.258	7.265	2.529	4.624	4.979	6.849	7.148	0	4.172
PG-8	3.797	6.358	6.881	4.461	2.022	6.381	5.961	0.837	3.168	4.198	4.485	4.172	0

**III (=D\*)**

**IV (=DA)**

In der Submatrix **III** ist die Distanzmatrix enthalten - so wie sie vom Modell reproduziert wird. Beim Vergleich mit der tatsächlichen Distanzmatrix **D** erkennt man, dass zwar Unterschiede bestehen, die aber nicht sehr groß sind. Der Koeffizient "Stress-1" (nach Kruskal) ist ein Gütemaß dafür, wie gut die reproduzierte und die tatsächliche Distanzmatrix übereinstimmen. Er ist in unserem Beispiel mit 0.065 sehr gut. Je kleiner dieser Koeffizient, umso besser. 0.0 entsteht wenn empirische und reproduzierte Distanzmatrix übereinstimmen.

Überraschend ist, dass die MDU in Submatrix **I** auch die Differenzen zwischen den Parteien und in Submatrix **IV** die Differenzen zwischen den Befragten-Gruppen reproduziert. Die Submatrix **II** ist die Transponierte zu **III**.

Wir bezeichnen im folgenden **I** mit **DB**, **II** mit **D\*'**, **III** mit **D\*** und **IV** mit **DA**

Die wesentliche Ursache dafür, dass die Reproduktion der empirischen Distanzmatrix **D** nicht exakt gelingt, liegt darin, dass im Kalkül der MDU gewissermaßen Faktoren verloren gehen. **D** wird durch die MDU in 2 Faktoren zerlegt. Tatsächlich ist **D** jedoch in 4 Faktoren zerlegbar. Wir werden das noch zeigen. Der MDU-Kalkül leistet dies jedoch nicht. Das gilt eigentlich für alle empirischen Distanzmatrizen. Man kann jedoch Distanzmatrizen simulieren, die durch die MDU exakt reproduziert werden. Wir werden das in Abschnitt P34.7.6 tun.

Schönemann (1970) gibt die maximal extrahierbare Faktorenzahl mit folgender Formel an:

$$max = (\sqrt{8 * z + 1} - 3) / 2$$

z = die grössere Zahl der Zeilen- oder Spaltenobjekte

Für die Beispieldaten ist  $max \leq 2.53$ . Es können also maximal nur 2 Faktoren extrahiert werden.

### P34.7.2 Kalkül der MDU

Für die metrische MDU gibt es mehrere verschiedene Kalküle, die zwar sehr ähnliche aber nicht identische Ergebnisse erbringen. Dass es verschiedene Kalküle gibt, hat seine Ursache darin, dass es nicht "den einzig richtigen" gibt. Wir werden noch deutlich machen, warum das so ist. Jeder Kalkül besteht aus einem Näherungsverfahren, das je nach Datenlage manchmal etwas bessere, manchmal etwas schlechtere Ergebnisse als seine Konkurrenten liefert. Es gibt jedoch mit dem Stress-Koeffizienten ein Kriterium, das die Güte der Lösung beurteilt.

Wir werden den in Almo verwendeten Kalkül nach Peter Schönemann (1970) kurz darstellen. Sein Charakteristikum ist, dass er eine algebraische Lösung erbringt, was aber nicht heißt, dass diese zur einzig richtigen Lösung führt oder auch nur die beste ist.

#### P34.7.2.1 Die quadriert und doppelt zentrierte Distanzmatrix *X*

Die Distanzmatrix **D**, wie sie oben abgebildet ist, wird zuerst gliedweise quadriert und danach doppelt zentriert. So entsteht die nachfolgend abgebildete Matrix **X**. Das geschieht in folgender Weise, die wir hier nicht begründen (siehe dazu Schönemann (1970 oder 1969, S.12, Punkt [2])). Wir zeigen dies in Form einer kurzen Anweisung in der Programmiersprache C

```
for (i=1; i<=Zeilen; i++)
  for (j=1; j<=Spalten; j++)
    X[i][j]= - (D[i][j]*D[i][j]-ZeilMW[i]-SpaltMW[j]+ GesamtMW) / 2.0;
```

```
X      = quadriert und doppelt zentrierte Distanzmatrix
D      = Distanzmatrix
ZeilMW = Zeilen-Mittelwerte aus quadrierter Distanzmatrix
SpaltMW = Spalten-Mittelwert aus quadrierter Distanzmatrix
GesamtMW= Gesamt-Mittelwert aus quadrierter Distanzmatrix
```

Die Distanzmatrix wird gliedweise quadriert. In jeder Zeile wird der Zeilen-Mittelwert und in jeder Spalte wird der Spalten-Mittelwert von jedem einzelnen Element subtrahiert. Zu allen Matrix-Elementen wird dann noch der Gesamt-Mittelwert addiert. Das Ganze mit 2 dividiert und die Vorzeichen noch umgedreht. Dadurch entstehen in **X** neue Distanzmaße, die Schönemann "quasi-skalare Produkte" nennt, die nun relativ zu den Zentroiden der beiden Zeilen- und Spaltenobjekten ausgedrückt sind. Für unsere Beispieldaten entsteht folgende Matrix:

quadrierte und doppelt zentrierte Distanzmatrix **X**

-4.4125	8.9625	10.9625	-2.8500	-12.6625
9.5875	-16.5375	-14.5375	13.6500	7.8375
-0.5125	1.3625	-2.1375	-2.4500	3.7375
-1.6125	2.2625	4.2625	-2.5500	-2.3625
-5.1125	10.7625	10.2625	-8.0500	-7.8625



-5.9125	11.4625	11.9625	-8.8500	-8.6625
8.8875	-17.2375	-15.2375	10.4500	13.1375
-0.9125	-1.0375	-5.5375	0.6500	6.8375

Wir bezeichnen im folgenden diesen Matrixtyp kurz mit

***qdz-Distanzmatrix***

wobei **qdz** natürlich *quadriert* und *doppelt zentriert* bedeutet.

**P34.7.2.2 Die Singulärwert-Zerlegung von X**

Die so gewonnene Matrix wird nach dem Verfahren der "Singulärwert-Zerlegung" zerlegt in

die Matrix der Links-Eigenvektoren **U** und  
 die Matrix der Rechts-Eigenvektoren **V** und  
 der Diagonalmatrix der Wurzeln der Eigenwerte **W**

Es gilt also:

$$(1) \mathbf{X} = \mathbf{U} * \mathbf{W} * \mathbf{V}'$$

Siehe dazu Anhang: Singulärwert-Zerlegung einer rechteckigen nicht-symmetrischen Matrix  
 Siehe auch P34.7.6.1 Notation: Matrix-Bezeichnungen.

**P34.7.2.3 Die "vorläufige" MDU-Lösung**

Es wird weiter aufgespalten in

$$(2) \mathbf{G} = \mathbf{U} * \text{Wurzel}(\mathbf{W})$$

$$(3) \mathbf{H}' = \text{Wurzel}(\mathbf{W}) * \mathbf{V}'$$

(der Einfachapostroph bei **V'** und **H'** bedeutet: transponiert)

**G** = das ist nun die (vorläufige) Ladungsmatrix der Zeilenvariablen (= der Zeilenobjekte)  
**H** = das ist die (vorläufige) Ladungsmatrix der Spaltenvariablen (=der Spaltenobjekte)

Bei der "Schönemann-Variante" (wie wir das nennen) wird in gleicher Weise wie oben in Formel (2) und (3) aufgespalten.

Beim originalen Schönemann-Kalkül wird **U\*W\*V'** anders aber völlig äquivalent aufgespalten in

$$(4) \mathbf{G} = \mathbf{U}$$

$$(5) \mathbf{H}' = \mathbf{W} * \mathbf{V}'$$

Der originale Schönemann-Kalkül und seine Variante sind absolut äquivalent. Wird die Ergebnismatrix, d.h. die endgültige gemeinsame Matrix von Zeilen- und Spaltenobjekten normiert und orthogonalisiert, dann sind sie voll identisch.

Für die Beispieldaten werden 5 Eigenwerte ( $\geq 0$ ) für **W** gewonnen. Der 5. Eigenwert ist =0. Die Matrizen der Links- und Rechts-Eigenvektoren besitzen also 4 Spalten (Faktoren), demzufolge auch **G** und **H**.

Nun liegt es nahe die beiden aus der Zerlegung von  $U*W*V'$  gewonnenen Ladungsmatrizen  $G$  und  $H$  in einer gemeinsamen Matrix zusammenzufassen. Wir bezeichnen diese Matrix als "vorläufige" Ladungsmatrix  $F0$ . Sie ist noch nicht die *endgültige* gemeinsame Ladungsmatrix.

vorläufige gemeinsame Matrix  $F0$

aus Matrix  $H$  (der Spaltenobjekte) =  $V*$ Wurzel( $W$ ) in Zeile 1 bis 5  
und Matrix  $G$  (der Zeilenobjekte) =  $U*$ Wurzel( $W$ ) in Zeile 6 bis 13

Faktor 1	Faktor 2	Faktor 3	Faktor 4	
-2.1191	-0.8678	0.7514	-0.8698	
4.0557	0.9724	-0.9744	-0.4522	
3.9715	-0.7766	0.9252	0.5610	<b>Matrix H</b>
-2.7849	-1.7718	-0.9570	0.4126	
-3.1233	2.4438	0.2547	0.3485	
2.5215	-1.9157	-0.7174	0.2215	
-3.8289	-1.5896	-0.3579	-0.4174	
-0.1241	1.4811	-0.1175	-0.6877	
0.8115	-0.0847	0.7062	0.5713	<b>Matrix G</b>
2.6228	0.1736	0.2590	-0.4616	
2.9181	0.1312	0.5132	0.0005	
-4.0461	0.0832	0.8122	0.2584	
-0.8748	1.7209	-1.0978	0.5150	

Das bedeutet - grafisch betrachtet - dass die Koordinatensysteme für  $G$  und  $H$  aufeinander gelegt werden. Das ist jedoch nicht zulässig, da die beiden Koordinatensysteme keinen gemeinsamen Ursprung besitzen, gegeneinander "verdrehen" und verschieden skaliert sind. Werden sie aufeinander gelegt, dann müssen sie zuvor *transformiert* werden und gegeneinander verschoben werden, so dass ihr Ursprung und ihre Achsen aufeinander liegen.

#### ***P34.7.2.4 Die endgültige MDU-Lösung durch den Schönemann-Kalkül***

*Transformiert* heisst: Es gilt eine Transformationsmatrix  $T$  und einen "Verschiebe"-Vektor  $a$  zum "Verdrehen", Skalieren und Verschieben zum gemeinsamen Ursprung zu finden. Peter Schönemann hat eine aufwändige algebraische Lösung entwickelt, durch die  $T$  und  $a$  gefunden wird. Das MDU-Programm in Almo ist danach programmiert.

Die endgültigen Ladungsmatrizen entstehen dann entweder aus dem Schönemann-Kalkül oder seiner Variante gemäß

$$(6) \quad \mathbf{A} = \mathbf{G} * \mathbf{T}$$

$$(7) \quad \mathbf{B} = \mathbf{H} * \text{inv}(\mathbf{T})' - \mathbf{a} * \mathbf{J}$$

$T$  = das ist die gesuchte Transformationsmatrix

$\text{inv}(T)$  = das ist die Inverse von  $T$

$a$  = das ist ein Vektor bestehend aus so vielen Werten, wie Faktoren extrahiert wurden

$J$  = das ist eine Matrix in der gleichen Größe wie  $B$ , die nur mit 1 gefüllt ist

$a*J$  = das ist dann eine Matrix in der gleichen Größe wie  $B$ , die in jeder Spalte  $i$  denselben Wert (aus  $a$ ) besitzt und somit zu einer Verschiebung von  $H * \text{inv}(T)'$  führt so dass der Ursprung von  $B$  auf den von  $A$  zu liegen kommt.

Werden **A** und **B** in einer gemeinsamen Matrix hintereinander gestellt, dann entsteht die endgültige gemeinsame Matrix **F**.

Der unumgängliche Nachteil des Schönemann-Verfahrens und auch anderer konkurrierender MDU-Verfahren ist es, dass bei dem "Rechenweg" zu **T** ein Rangverlust eintritt. D.H. die beiden Ladungsmatrizen **A** und **B** besitzen weniger Faktoren als die "vorläufigen" Matrizen **G** und **H**. In unserem Beispiel besitzen **G** und **H** 4 Faktoren, **A** und **G** hingegen nur noch 2. Das gilt für alle empirischen Distanzmatrizen.

**A** und **B** werden zur endgültigen gemeinsamen Ladungsmatrix **F** hintereinander gestellt. Aus **F** wird dann die Gesamt-Distanzmatrix **gD\*** reproduziert. In unserem Beispiel ist dies eine 13\*13-Matrix. Beide Ergebnis-Matrizen haben wir bereits oben abgebildet. Wie oben gezeigt wurde, ist die Submatrix **III** dann die reproduzierte Distanzmatrix **D\*** zwischen den Zeilenobjekten einerseits und den Spaltenobjekten andererseits.

Die Distanzen in **gD\*** werden nach folgender Formel aus **F** ermittelt:

$$(8) \quad d_{ij}^2 = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^{fz} (f_{ik} - f_{jk})^2$$

m = Zahl der Spalten- plus Zeilenobjekte

fz = Zahl der Faktoren

d(ij) = Distanz für Objekt i gegenüber Objekt j

In obiger Formel wird die quadrierte Distanz bestimmt.

f(ik) bzw. f(jk) = Ladung von Spalten- und Zeilenobjekt i bzw. j auf Faktor k

Das kleine C-Programm für diese Formel lautet:

```
for (i=1; i<=m; i++)
  for (j=1; j<=m; j++)
    d[i][j] = .0; //zuerst auf .0 setzen

for (i=2; i<=m; i++) { //nur fuer unteres Dreieck
  for (j=1; j<i; j++) {
    for (k=1; k<=fz; k++)
      d[i][j] += (f[i][k]-f[j][k]) * (f[i][k]-f[j][k]);
    d[i][j] = sqrt(d[i][j]);
    d[j][i] = d[i][j]; //spiegeln
  }
}
```

### Die Transformationsmatrix **T**

Das Charakteristikum des Schönemann-Verfahrens besteht darin, dass die Transformationsmatrix **T** und der Vektor **a** so berechnet werden, dass die quadrierten Differenzen zwischen den Gliedern der Distanzmatrix **D** und der reproduzierten Matrix **D\*** sich zu einem möglichst kleinen Wert summieren. Diese Transformationsmatrix darf nicht mit der Rotationsmatrix aus der Faktorenanalyse verwechselt werden. Bei der Rotation werden - unter Beibehaltung der Punkte-Konfiguration die Koordinatenachsen orthogonal oder schiefwinklig gedreht. Die Transformationsmatrix **T** der MDU verändert die Punkte-Konfiguration. Die Distanzen zwischen den Punkten werden verändert und dies nicht um einen konstanten Faktor, der für alle paarweisen Distanzen gleich ist. Die Distanzen, berechnet gemäß Formel (8), zwischen den Objekten in Matrix **G** sind andere als die für Matrix **A** (= **G\*T**). Dies ist auch

verständlich, da **G** aus der quadrierten und doppelt zentrierten Distanzmatrix **X** und nicht aus der empirischen Distanzmatrix **D** entstanden ist.

#### **P34.7.2.5 Die "Güte" der MDU-Lösung - gemessen durch den Stress-Koeffizienten**

Die Güte der MDU-Lösung wird durch den Vergleich der originalen Distanzmatrix **D** mit der reproduzierten **D\*** festgestellt.

Je besser es der MDU gelingt die originale empirische Distanzmatrix zu reproduzieren, d.h. je ähnlicher die beiden Distanzmatrixen **D** und **D\*** sind umso besser ist die MDU-Lösung. Der Stress-1-Koeffizient nach Kruskal ist eine Formalisierung dieses Vergleichs. Er wird nach folgender Formel berechnet:

$$\text{Stress-1} = \sqrt{\text{sum2} / \text{sum1}}$$

sum1= die Glieder der originalen Distanzmatrix werden quadriert. Diese quadrierten Glieder werden zu sum1 summiert.

sum2 = aus der MDU-Lösung wird die reproduzierte Distanzmatrix **D\*** gebildet. Diese wird von der originalen Distanzmatrix **D** gliedweise subtrahiert. Es entsteht die "Residualmatrix". Dann werden alle Glieder wie bei sum1 beschrieben quadriert und zu sum2 summiert.

Der Stress-1-Koeffizient drückt also aus, wie groß die vom Modell nicht erklärte residuale Distanzsumme ist - gemessen an der gesamten Distanzsumme. Also gibt dann noch zusätzlich den Koeffizienten Stress-2 nach Kruskal aus. Dieser besteht aus n Stress-Koeffizienten je Zeile und einem aus diesen gebildeten durchschnittlichen pauschalen Stress-Wert. Stress-2 wird in folgender Weise ermittelt:

Die Formel für den ( quadrierten ) Stress-2-Koeffizienten ( nach Kruskal ) ist

$$(9) \quad S^2 = \frac{1}{n} \sum_i \frac{\sum_j (d_{ij}^* - d_{ij})^2}{\sum_j (d_{ij} - M_i)^2}$$

Notation und Formeln:

S = Stress-2 (beachte: in Formel 9 wird der *quadrierte* Stresswert bestimmt)

n = Zahl der Zeilen der Matrix

m = Zahl der Spalten

i = Index für die n Zeilen der Matrix

die Summen mit Index i laufen von i=1 bis i=n

j = Index für die m Spalten der Matrix

die Summen mit Index j laufen von j=1 bis j=m

**D** = die empirische Distanzmatrix

d = ein Element aus Distanzmatrix D

d(ij) = das Element aus D in Zeile i und Spalte j

M(i) = der Zeilenmittelwert aus den m Elementen d(ij) in Zeile i

**D\*** = die vom Modell reproduzierte Distanzmatrix

$d^*$  = ein Element aus der reproduzierten Distanzmatrix  $D^*$   
 $d^*(ij)$  = das Element aus  $D^*$  in Zeile  $i$  und Spalte  $j$

Formel 9 wird in mehreren aufeinander folgenden Schritten berechnet. Dies sind

$$(9a) \quad r(ij) = [d(ij) - d^*(ij)] * [d(ij) - d^*(ij)]$$

$r(ij)$  ist also die quadrierte Differenz von  $d$  und  $d^*$

$$(9b) \quad a(ij) = [d(ij) - M(i)] * [d(ij) - M(i)]$$

$a(ij)$  ist also der quadrierte "Abweichungswert" der  $d(ij)$  von ihrem Zeilenmittelwert

$$(9c) \quad \text{sum1}(i) = \text{Summe der } m \text{ } r(ij) \text{ in Zeile } i$$

$$(9d) \quad \text{sum2}(i) = \text{Summe der } m \text{ } a(ij) \text{ in Zeile } i$$

$$(9e) \quad \text{stress}(i) = \text{sum1}(i) / \text{sum2}(i)$$

$\text{stress}(i)$  ist also ein (quadrierter) Stress-Koeffizient für die Zeile  $i$

$$(9f) \quad \text{str} = \text{Summe der } n \text{ } \text{stress}(i)$$

$$(9g) \quad \text{Stress-2} = \text{Wurzel}(\text{str} / n)$$

Almo gibt neben dem pauschalen Stress-2-Koeffizienten aus 9g noch je Zeile  $i$  den Stress-Koeffizienten

$$(9h) \quad \text{Wurzel}(\text{stress}(i))$$

aus

In der Literatur werden mehrere verschiedene Stress-Koeffizienten verwendet. Siehe dazu Busing (2010). So werden neben Stress-1 noch verwendet: Stress-2, S-Stress-1, S-Stress-2, penalisierter Stress. Diese Koeffizienten haben teilweise nicht nur die Aufgabe, die Güte der MDU-Lösung in einer Zahl auszudrücken, sondern auch noch bei einem Iterations-Verfahren als Abbruch-Kriterium zu dienen. Ein Beispiel dafür ist das Programm *Prefscal* in SPSS.

Bei unserem Beispiel ergibt sich als Stress-1 ein Wert von 0.0657 und folgende Stress-2-Werte:

Stress-2 nach Kruskal

bezogen auf die Distanzen zwischen Zeilenobjekten und Spaltenobjekten

Stress je Zeile

1	=	0.3061
2	=	0.0682
3	=	0.2966
4	=	0.1978
5	=	0.1739
6	=	0.3069
7	=	0.0603
8	=	0.1615

Stress-2 nach 2 extrahierten Faktoren = 0.2178

### P34.7.3 Eingabe in Almo-Programm Prog18m6

In Almo werden 2 Programm-Masken für das metrische Unfolding angeboten.

**Prog18m6** rechnet mit einer aus einer Datei eingelesenen Distanz oder Präferenzmatrix eine MDU.

**Prog18m7** bildet zuerst aus Distanz- oder Präferenz-Urteilen einzelner Individuen eine Matrix und rechnet dann eine MDU.

Prog18m6.Msk  
 Metrisches multidimensionales Unfolding  
 (metrisches MDU)  
 mit eingegebener fertiger Distanz- oder Präferenz-Matrix

Präferenz- oder Distanz-Urteile von Subjekten gegenüber Objekten werden faktorisiert und grafisch in einem (mehrdimensionalen) Koordinatensystem als Punkte dargestellt

Beispiel:  
 8 Personen (oder Personengruppen) werden darüber befragt, wie stark sie die 5 Parteien Partei A bis Partei E präferieren  
 Ihre Präferenz-Urteile werden zu Distanz-Urteilen "umgedreht"  
 Es entsteht folgende Distanzmatrix

		Partei				
		A	B	C	D	E
V2	V3	V4	V5	V6		
		---	---	---	---	---
		4.0	3.0	3.0	4.0	6.0
		3.0	9.0	9.0	2.0	4.0
		3.0	5.0	6.0	4.0	2.0
		2.0	4.0	4.0	3.0	3.0
		4.0	2.0	3.0	5.0	5.0
		4.0	1.0	2.0	5.0	5.0
		3.0	9.0	9.0	3.0	2.0
		4.0	6.0	7.0	4.0	2.0

Matrix entnommen aus Schönemann (1970)

Form der Distanzmatrix ---> Hilfe

Die Personen bilden die Zeilen der Matrix und die Parteien die Spalten. Die Personen sind die "Zeilenobjekte" und die Parteien die "Spaltenobjekte"

Siehe auch  
 MDU mit gruppierten Daten ---> [".\Almo\_Msk\Prog18m7.Alm"]  
 MDU von Auto-Präferenzen ---> [".\Almo\_Bsp\Nakanishi.Alm"]

Siehe Almo-Dokument "Metrisches multidimensionales Unfolding"

Speicher fuer x Variable Hilfe

Vereinbare Variable= 100

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

**Variablenamen**

Datei der Variablenamen

zeige = Namensdatei in Output zeigen  
leer = nicht zeigen

---

**Freie Namensfelder**

Leere alle Eingabefelder dieser Sub-Box

Name 1=PartA  
 Name 2=PartB  
 Name 3=PartC  
 Name 4=PartD  
 Name 5=PartE

erzeuge zusätzliche Namensfelder

---

**Variablenamen in Datei speichern**

Eingabefeld leer = nicht speichern

**Datei aus der gelesen wird**

bei Datei-Problemen

"C:\Almo15\TESTDAT\FuenfParteien.fre"

frei Format der Daten

V1:5 der Datensatz enthält diese Variablen  
Bei Format DIREKT schreiben Sie: alle\_v

Wenn Dateiformat FIX oder Nicht-Standard-FREI

**Spaltenvariable (Spaltenobjekte)**


PartA,PartB,PartC,PartD,PartE

---


**Distanz oder Präferenz**


0 = die Spaltenvariablen sind Distanzen  
 1 = die Spaltenvariablen sind Präferenzen


**Einstellungen**



**Faktorenzahl**
  
 Es sollen x Faktoren extrahiert werden.


Wird das Eingabefeld leer gemacht  
 (durch Klick auf den doppelköpfigen Pfeil)  
 dann wird die maximal mögliche Faktorenzahl  
 extrahiert



**Kalkül**
  
 1 = nach Schönemann  
 0 = Schönemann-Variante



**Ergebnis-Matrix normieren**  
 1 = spaltenweise auf Mittelwert 0 normieren  
 0 = nicht normieren



**Ergebnis-Matrix orthogonalisieren**
  
 1 = Ergebnis-Matrix orthogonalisieren  
 0 = nicht



Option: weitere Einstellungen


Option: Ein- und Ausschliessen von Untersuchungseinheiten


Option: Umkodierungen und Kein-Wert-Angaben


Option: Untersuchungseinheiten gewichten


Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix


Grafik-Optionen

### ***P34.7.3.1 Erläuterungen zu den Eingabe-Boxen***

Im Beispiel wurde eine Distanzmatrix mit 8 Zeilen und 5 Spalten eingelesen. Die Matrix ist Schönemann(1970 oder 1969) entnommen. Wir wollen unterstellen, dass 8 Personen (oder 8 Personengruppen) darüber befragt wurden, wie positiv sie 5 Parteien beurteilen. Danach wurden ihre Urteile zu negativen Urteilen, also Distanzen "umgedreht".

#### **Eingabe-Box: Vereinbare x Variable.**

Siehe dazu Almo-Dokument 0, Arbeiten mit Almo, Abschnitt P0.1  
Eigentlich würde es ausreichen 8 Variable zu vereinbaren.

#### **Eingabe-Box: Option "Weitere Vereinbarungen".**

Siehe dazu Almo-Dokument 0, Arbeiten mit Almo, Abschnitt P0.2  
Die Optionsbox braucht für unser Beispiel nicht geöffnet zu werden



### Eingabe-Box: Variablennamen

Siehe dazu Almo-Dokument 0, Arbeiten mit Almo, Abschnitt P0.3

Den Spaltenobjekten, d.h. den 5 Parteien können Namen gegeben werden. Das ist nicht obligatorisch, jedoch empfehlenswert, da dann die Ergebnisse besser zu lesen sind. Für die Zeilenvariablen ist es nicht möglich, Namen zu vergeben.

### Eingabe-Box: Datei aus der gelesen wird

Siehe dazu Almo-Dokument 0, Arbeiten mit Almo, Abschnitt P0

The image shows two dialog boxes from the Almo software. The top dialog box is titled "Datei aus der gelesen wird" and contains the following elements: a file path input field with the text "C:\Almo15\TESTDAT\FuenfParteien.fre", a "Format der Daten" dropdown menu currently set to "frei", and a "V1:5" input field for variables. There are "Hilfe" buttons and a "bei Datei-Problemen" label. The bottom dialog box is titled "Wenn Dateiformat FIX oder Nicht-Standard-FREI" and also has a "Hilfe" button.

#### 1. Eingabefeld:

Geben Sie in der Box „Datei aus der gelesen wird“ zuerst den Namen der Datei an, in der sich Ihre Daten befinden. Das machen Sie dadurch, dass Sie diesen Namen in das Eingabefeld schreiben oder einfach dadurch, dass Sie auf den Knopf mit dem Öffne-Symbol klicken.

Es erscheint dann eine Dialogbox, in der Sie gefragt werden, ob Sie die Datei anschauen wollen oder ob Sie in das Eingabefeld einen (neuen) Namen einsetzen wollen. Klicken Sie auf "Namen einsetzen". Es erscheint dann die gewohnte Datei-Auswahlbox, in der Sie die Datei suchen, in der sich Ihre Daten befinden.

#### 2. Eingabefeld: *Format der Daten*

Klicken Sie auf den Knopf vor dem Eingabefeld und wählen Sie jenes Formate aus, in dem Ihre Daten geschrieben sind.

Almo unterscheidet 3 Formate, in denen die Daten vorliegen können. Siehe die ausführliche Darstellung im Almo-Handbuch, Teil 2, Abschnitt 9.2 und 14.

#### 1. Format FREI

Dieses Format wird gelegentlich auch "ASCII-Format" oder "free-field"-Format genannt.

Beispiel: ..... 1 2 23.5 Beamter 23, 0, 4 .....

Die Werte eines Datensatzes sind durch ein oder mehrere Blanks (Leertasten) und/oder ein Komma und/oder das Tabulator-Zeichen voneinander getrennt. Der Punkt dient als Dezimalpunkt. Wir werden im folgenden noch zwischen „Standard-FREI“ und „Nicht-Standard-FREI“ unterscheiden. Letzteres Format bezeichnen wir auch mit „erweitert FREI“. In diesem ist es z.B. möglich das Komma als Dezimalzeichen einzusetzen. Zu diesem Zweck muss die nachfolgende Optionsbox "Wenn Dateiformat fix oder nicht-Standard-frei" geöffnet werden und dort entsprechend vermerkt werden. Dann kann aber das Komma als Trennzeichen zwischen den Werten nicht verwendet werden. Es bleibt nur noch das Blank bzw. beliebig viele Blanks als Trennzeichen.

## 2. Format FIX

Die Werte stehen ohne Trennzeichen direkt hintereinander.

Beispiel: .....1223.5Beamter2304.....

Die „Breite“ der einzelnen Variablenwerte muß noch mitgeteilt werden. Dazu muss die nachfolgende Optionsbox "Wenn Dateiformat Fix oder Nicht-Standard-Frei" geöffnet werden.

## 3. Format DIREKT

Dies ist ein Almo-spezifisches binäres Format. Es ist sehr schnell, erlaubt den DIREKT-Zugriff auf ausgewählte Datensätze (sogar Variable) und ermöglicht das platzsparende Speichern der Daten auf Speichermedien mit 1 bis 8 Byte (je nach Größe des Variablenwertes). Daten in diesem Format müssen zuvor mit Almo im Format FREI oder FIX geschrieben worden sein. Sie werden dann durch Prog00mr in das Format DIREKT übertragen.

### 3. Eingabefeld: Der Datensatz enthält diese Variable:

Geben Sie z.B. V1:20 an, wenn Ihr Datensatz 20 Variable umfasst. Befinden sich Ihre Daten im Format DIREKT, dann schreiben Sie "alle\_V" oder klicken Sie auf den Knopf vor dem Eingabefeld und treffen Sie Ihre Auswahl.

### P34.7.3.2 Daten für die MDU

Almo bildet aus einer eingelesenen **Datenmatrix** eine **Distanzmatrix**, die dann dem MDU-Kalkül unterworfen wird.

Daten, die von Almo eingelesen und verarbeitet werden, müssen die Form der Datenmatrix besitzen.

#### Beispiel einer Datenmatrix

1	2	4.0	3.0	3.0	4.0	6.0	-3.234	7.2
1	3	3.0	9.0	9.0	2.0	4.0	-4.02	5.2
1	4	3.0	5.0	6.0	4.0	2.0	10.7348	6.2
1	1	2.0	4.0	4.0	3.0	3.0	2	5.2
1	2	4.0	2.0	3.0	5.0	5.0	2.0	7.1
1	3	4.0	1.0	2.0	5.0	5.0	-1.00	5.1
1	4	3.0	9.0	9.0	3.0	2.0	4	6.3
2	1	4.0	6.0	7.0	4.0	2.0	7.55	4.3

Eine Zeile enthält die Daten einer Untersuchungseinheit (z.B. einer Person). Die Variablenwerte sind durch beliebig viele Blanks voneinander getrennt. Möglich wäre auch das Komma. Als Dezimalzeichen wird der Punkt verwendet. Wir bezeichnen eine Zeile auch als "Datensatz" oder "Datenvektor" einer Untersuchungseinheit. Eine Spalte enthält die Werte einer Variablen. Die Spalten-Nummer entspricht der Variablen-Nummer. In der 3. Spalte steht also Variable V3.

Im Almo-MDU-Programm Prog18m6 werden aus einer derartigen Datenmatrix die Variablen herausgegriffen, die dann die Distanz- bzw. Präferenzmatrix bilden. In obiger Datenmatrix entsteht aus den 5 Variablen V3 bis V7 die Distanzmatrix für das Beispiel 3 aus Schönemann, das wir in der Programm-Maske Prog18m6 verwenden.

Die Datenmatrix kann auch nur aus den Variablen bestehen, welche die Distanz- bzw. Präferenzwerte der Untersuchungseinheiten enthalten.

Beispiel:

```
4.0 3.0 3.0 4.0 6.0
3.0 9.0 9.0 2.0 4.0
3.0 5.0 6.0 4.0 2.0
2.0 4.0 4.0 3.0 3.0
4.0 2.0 3.0 5.0 5.0
4.0 1.0 2.0 5.0 5.0
3.0 9.0 9.0 3.0 2.0
4.0 6.0 7.0 4.0 2.0
```

In diesem Falle sind Datenmatrix und Distanzmatrix identisch

Almo liest alle Zeilen (=Datensätze) ein. In obiger Datenmatrix sind dies 8. So entsteht also eine 8\*5 Distanz- bzw. Präferenzmatrix.

Würde die Datenmatrix 1000 Zeilen umfassen, dann würde eine 1000\*5 Distanz-bzw. Präferenzmatrix entstehen. Ob eine solche riesige Matrix noch rechenbar ist, kann nur ausprobiert werden. Sicherlich sind aber deren Ergebnisse nur schwer zu interpretieren. Die Grafik dürfte nur aus einem "Punktegewimmel" bestehen.

In Prog18m7 bietet Almo an, die Untersuchungseinheiten nach bestimmten Merkmalen zu gruppieren und dann mit den Gruppen als Zeilenobjekten eine MDU zu rechnen. Beispielsweise könnten in einer Untersuchung zur Präferenz von 1000 Personen hinsichtlich von 5 Automarken, diese Personen nach den "Gruppierungsvariablen" Geschlecht, Bildungsstufen (4 Kategorien) und Einkommensgruppe (5 Kategorien) in  $2*4*5=40$  Zeilenobjekte zusammengefasst werden. Eine Gruppierung ist allerdings nur dann sinnvoll, wenn die Mitglieder einer Gruppe in ihren Distanz- bzw. Präferenz-Urteilen gut übereinstimmen und sich von den Mitgliedern anderer Gruppen deutlich unterscheiden; kurz formuliert: wenn sie ein *Cluster* bilden. Siehe dazu Almo-Dokument 17 "Clusteranalyse".

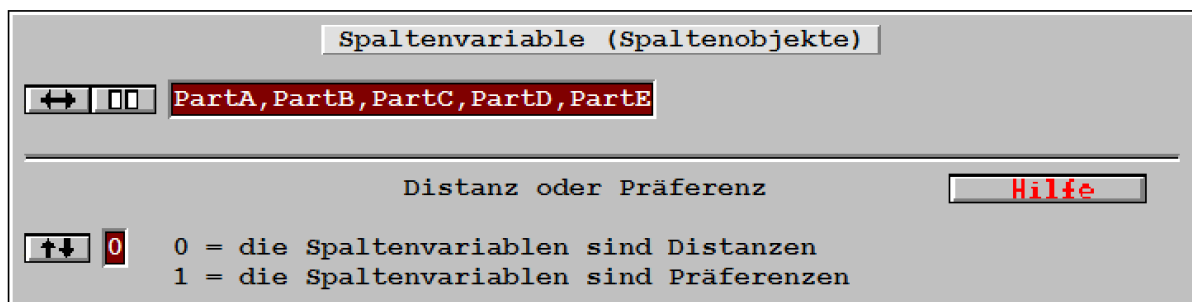
#### Eingabebox: Wenn Dateiformat ....



Siehe dazu Almo-Dokument 0 "Arbeiten mit Almo", Abschnitt P0.4.1

#### Eingabebox: Spaltenvariable (Spaltenobjekte)

Hier werden die Variable angegeben, die die Spalten der Matrix bilden



Normalerweise sind dies die Objekte, über die Distanz- bzw. Präferenz-Urteile gesprochen worden sind. Es ist jedoch auch zulässig, dass die Untersuchungseinheiten (die Personen, die diese Urteile fällen) als Spaltenvariable eingelesen werden. Dieser Fall tritt auf, wenn die

Zahl der Untersuchungseinheiten kleiner ist als die Zahl der zu beurteilenden Objekte. Wichtig für die MDU ist, dass die Zahl der Zeilen nicht kleiner ist als die Zahl der Spalten. Wenn dieser Fall auftritt, dann muss die Distanzmatrix transponiert werden. Dieses Problem wird noch ausführlich in Abschnitt P34.7.3.2 behandelt.

Im 2. Eingabefeld muss dann noch angegeben werden, ob die eingelesene Matrix Distanzen oder Präferenzen enthält

Der MDU-Kalkül erwartet eine Distanzmatrix. Je größer der Zahlenwert in einer Zelle ij der Distanzmatrix umso größer ist die Distanz zwischen Subjekt i und Objekt j.

Bestehen die vorliegenden Daten aus Präferenz-Urteilen von Subjekten gegenüber Objekten, dann ist das kein Problem. Je größer der Zahlenwert in einer Zelle ij der Präferenzmatrix umso kleiner ist die Distanz zwischen Subjekt i und Objekt j.

Es muss also einfach die Kodierung umgedreht werden, um aus einer Präferenzmatrix eine Distanzmatrix zu machen Das leistet Almo automatisch, wenn in der Eingabebox "Spaltenvariable" im 2. Eingabefeld 1 eingesetzt wird. Almo ermittelt zuerst den maximalen und minimalen Präferenzwert. Die einfache Formel ist

$$\text{Distanzwert} = \text{max.Praef.wert} - \text{Praef.wert} + \text{min.Praef.wert}$$

Beispiel:

Die Präferenzwerte reichen von 1 bis 9 Dann wird aus dem Präferenzwert 9 der Distanzwert 1 und aus dem kleinsten Präferenzwert 1 der Distanzwert 9. Aus 7 wird die die Distanz 3.

Die Präferenzwerte reichen von -5 über 0 nach +5. Dann wird z.B. aus -5 der Distanzwert +5 und aus +5 der Distanzwert -5

Bei dieser Art der Umkodierung kann also ein *negativer Distanzwert* auftreten. Negative Wert in der Distanzmatrix sind unzulässig. Almo überprüft ob negative Werte auftreten und addiert (programmintern) dann zu alle Matrixwerten einen Zahlenwert, so dass der kleinste auftretende Wert zu 0 wird.

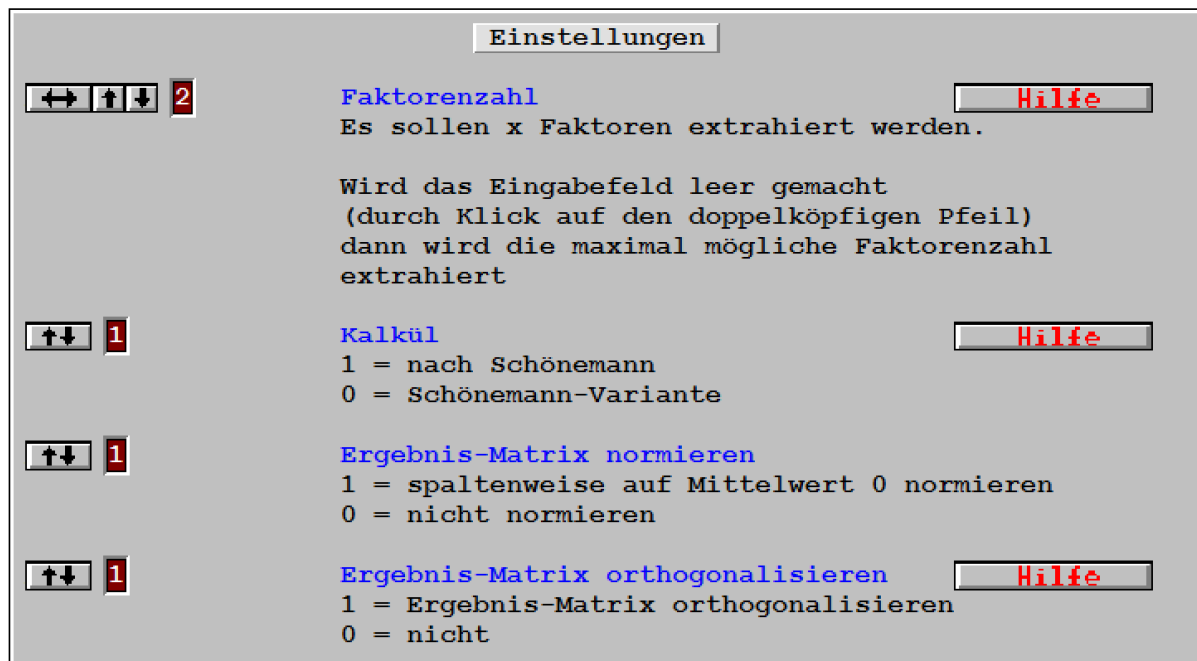
Will der Benutzer eine andere Art der Umkodierung erreichen, z.B. möchte er erreichen, dass in unserem Zahlenbeispiel ein kleinster Distanzwert von 1 entsteht, dann kann er dies selbst in folgender Weise tun:

- a. Die einzulesenden Daten werden in der Box "Spaltenvariable" im 2. Eingabefeld mit "0" als "Distanzmatrix" deklariert.
- b. Die Box "Umkodierungen und Kein-Wert-Angaben" wird geöffnet. Die mittlere Sub-Box wird geöffnet. Dann wird in ein Eingabefeld z.B. folgende Umkodierungs-Anweisung geschrieben

$$V1:5 (* -1; +5; +1)$$

Die Spaltenvariablen V1 bis V5 werden zuerst mit -1 multipliziert, danach wird 6 hinzu addiert. Wenn die Präferenzwerte in V1 bis V5 von -5 bis +5 kodiert sind, dann werden sie dadurch auf 11 bis 1 umgedreht. Die Wertespanne ist sowohl bei den originalen Daten, wie auch bei den umkodierten =10. Im 1. Fall beginnt sie mit -5 im zweiten mit 1.

## Eingabebox: Einstellungen



### 1. Eingabefeld: Faktorenzahl

Das Eingabefeld kann leer gelassen werden. Dann berechnet Almo so viele Faktoren wie gemäß nachfolgender Formel möglich sind. Der Benutzer kann aber auch eine Faktorenzahl festlegen. In der Regel wird man nicht mehr als 2 oder 3 (bei großen Matrizen) angeben. Die maximale Faktorenzahl ist

$$\text{maxFaktorzahl} \leq ( \text{Wurzel}(8 \cdot g + 1) - 3 ) / 2$$

$g$  ist die grössere Zahl der Zeilen- oder Spalten-Objekte.

Bei einer  $8 \times 5$  Distanzmatrix ist  $g=8$ . Die maximale Faktorenzahl ist dann

$$\text{maxFaktorzahl} \leq ( \text{Wurzel}(8 \cdot 8 + 1) - 3 ) / 2 \leq 2.53..$$

also =2. Folgende Vorgehensweise wird empfohlen: Zuerst sollte das Eingabefeld leer gemacht werden, also Almo die maximale Faktorenzahl rechnen lassen. Dabei ist es möglich, dass Almo die Faktorenzahl selbständig weiter reduziert. Das hat seine Ursache in einem Rangverlust bei der Errechnung der Transformationsmatrix **T**. Siehe oben Abschnitt P34.7.2.

In sehr seltenen Fällen, kann es geschehen, dass eine offenkundig unsinnige Lösung entsteht. Der Benutzer sollte dann selbst eine kleinere Faktorenzahl fixieren. Als Beispiel, bei dem ein solches Problem aufgetreten ist, siehe das Beispielprogramm "Nakanishi.Alm". Es wird gefunden durch Klick auf das Menü "Almo/Liste aller Almo-Programme"

### 2. Eingabefeld: Kalkül

Im Almo-MDU-Programm wird der Kalkül von Schönemann (1970 oder 1969) verwendet. Wir haben oben in Abschnitt P34.7.2 dargestellt, dass es für diesen Kalkül 2 Varianten gibt.

Scheinbar unterschiedliche Ergebnisse der beiden Kalküle verschwinden, wenn die Ergebnisse normiert werden.

### 3. Eingabefeld: Ergebnis-Matrix normalisieren

Das bedeutet sehr einfach, dass für jede Spalte der gemeinsamen Ladungsmatrix von Zeilen- und Spalten-Objekten der Mittelwert errechnet wird und von jedem Spaltenwert dann sein Mittelwert abgezogen wird. Die normierte gemeinsame Ladungsmatrix besitzt dann Spalten-Mittelwerte von 0.

### 4. Eingabefeld: Ergebnis-Matrix orthogonalisieren

Die gemeinsame (aus den Matrizen der Zeilen- und Spalten-Objekte bestehende) Ladungsmatrix, die aus dem Schönemann-Kalkül und seiner Variante entsteht, ist nicht notwendigerweise orthogonal. Grafisch betrachtet heißt dies, dass die Achsen des Koordinatensystems nicht rechtwinklig aufeinander stehen. Sie stehen allerdings nur sehr geringfügig schiefwinklig aufeinander. Durch die Orthogonalisierung wird das korrigiert - ohne, dass die Punktekonfiguration verändert wird. Insbesondere die Distanzen zwischen den Punkten der Zeilen- und Spalten-Objekte bleiben gleich - so dass auch der Stress-1-Koeffizient nicht verändert wird.

Die Orthogonalisierung verläuft in folgenden Schritten.

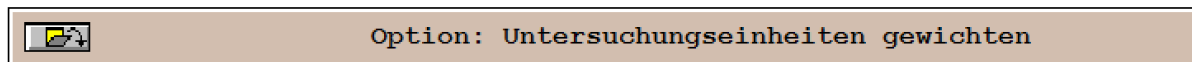
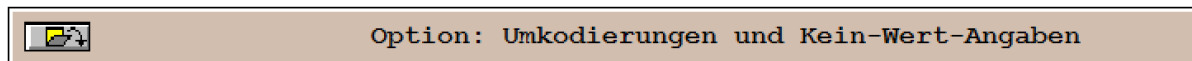
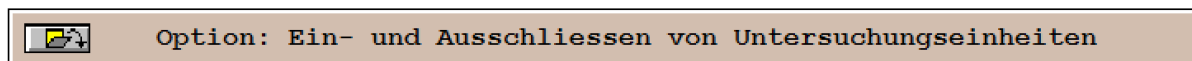
- 1.) Bilde  $\mathbf{R} = \mathbf{F}' * \mathbf{F}$
- 2.) Zerlege  $\mathbf{R}$  in  $\mathbf{S} * \mathbf{W} * \mathbf{S}'$  wobei  $\mathbf{S}$  die Matrix der Eigenvektoren ist und  $\mathbf{W}$  die Diagonalmatrix der Eigenwerte
- 3.) die neue orthogonalisierte Ladungsmatrix ist dann  $\mathbf{F} * \mathbf{S}$

Im originalen Schönemann-Kalkül wird nicht orthogonalisiert. Schönemann begnügt sich mit der Normierung der Ergebnis-Matrix. Wenn der Almo-Benutzer bei unseren Beispieldaten die Orthogonalisierung verhindert, dann erhält er aus Almo exakt, die Ergebnisse, die Schönemann (1970) angibt.

### **Optionsbox: Ein- und Ausschliessen von Untersuchungseinheiten**

### **Optionsbox: Umkodierungen und Kein-Wert-Angaben**

### **Optionsbox: Untersuchungseinheiten gewichten**



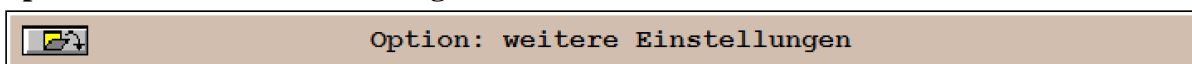
Zu diesen 3 Optionsboxen siehe Almo-Dokument Nr 0 "Arbeiten mit Almo"

Abschnitt P0.7.1

Abschnitt P0.5

Abschnitt P0.7.2

### **Optionsbox: Weitere Einstellungen**



Diese Box beinhaltet einige sehr bedeutsame Erweiterungsmöglichkeiten der Almo-MDU.

Wird die Optionsbox geöffnet, dann sieht man folgendes

XLoesche wieder diese Box (dann Voreinstellungen wieder gueltig)

weitere Einstellungen

**Zwischenergebnisse**

0 = nicht ausgeben (Voreinstellung)  
1 = Zwischenergebnisse aus Schönemann-Kalkül ausgeben  
2 = sonstige Zwischenergebnisse ausgeben  
3 = alle Zwischenergebnisse

**Lösung für vorläufige Zerlegung der Distanzmatrix**

1 = zusätzlich ausgeben   
0 = nicht (Voreinstellung)

**gesamte Distanzmatrix über  
Zeilen- u. Spalten-Objekte ausgeben**

1 = ausgeben  
0 = nicht (Voreinstellung)

---

Matrix in Datei speichern

**welcher Matrix soll gespeichert werden ?**

0 = wird nicht gespeichert (Voreinstellung)  
1 = die eingegebene Distanzmatrix speichern  
2 = die eingegebene Präferenzmatrix speichern  
3 = die Distanzmatrix transponiert speichern  
4 = die gemeinsame Ladungsmatrix speichern

".\Progs\Distanz.mat"

---

**Automatische Transponierung der Matrix  
wenn Zahl der Zeilenobjekte kleiner als  
Zahl der Spaltenobjekte ist**

0 = erlaube automatische Transponierung (Voreinstellung)  
1 = verhindere automatische Transponierung

---

**Basisstatistiken ausgeben**

1= ja  
0= nein

### 1. Eingabefeld: Zwischenergebnisse

Wird **1** eingesetzt, dann werden die Ergebnisse aus dem Schönemann-Kalkül (bzw. dessen Variante) Schritt für Schritt ausgegeben. Unsere Beispiel-Matrix ist das Schoenemann-Beispiel 3. Der Benutzer kann dann mit den bei Schoenemann (1970) etwas spärlichen Zwischenergebnissen vergleichen und er kann weiterhin die bei Schönemann im Abschnitt 4 "Computational notes" aufgeführten Rechenschritte nachvollziehen.

Wird **2** eingesetzt, dann werden noch weitere zusätzliche Zwischenergebnisse ausgegeben.

Wird **3** eingesetzt, dann werden alle Zwischenergebnisse ausgegeben. Es kann dann allerdings sehr viel an Ergebnistext von Almo produziert werden, so dass der Benutzer diese Ausgabe besser erst dann anfordert, wenn er in die Almo-MDU genügend eingearbeitet ist.

2. Eingabefeld: Lösung für vorläufige Zerlegung der Distanzmatrix (vorläufige MDU-Lösung)  
Bei der "vorläufigen MDU-Lösung" geschieht folgendes: Die quadrierte und doppelt zentrierte Datenmatrix (wir bezeichnen sie mit  $\mathbf{X}$ ) wird der Singulärwert-Zerlegung unterzogen, d.h. in die Matrix der Links-Eigenvektoren  $\mathbf{U}$  und der Rechts-Eigenvektoren  $\mathbf{V}$  und der Diagonalmatrix der Wurzeln der Eigenwerte  $\mathbf{W}$  zerlegt. Also:

$$\mathbf{X} = \mathbf{U} * \mathbf{W} * \mathbf{V}'$$

(der Apostroph bei  $\mathbf{V}$  und  $\mathbf{H}$  bedeutet: transponiert)

Dann wird weiter aufgespalten in

$$\begin{aligned}\mathbf{G} &= \mathbf{U} * \text{Wurzel}(\mathbf{W}) \\ \mathbf{H}' &= \text{Wurzel}(\mathbf{W}) * \mathbf{V}'\end{aligned}$$

$\mathbf{G}$  = das ist nun die (vorläufige) Ladungsmatrix der Zeilenvariablen (= der Zeilenobjekte)

$\mathbf{H}$  = das ist die (vorläufige) Ladungsmatrix der Spaltenvariablen (=der Spaltenobjekte)

Werden die beiden Matrizen  $\mathbf{G}$  und  $\mathbf{H}$  zu einer gemeinsamen Matrix hintereinander gestellt so entsteht die Matrix  $\mathbf{F0}$ . Das ist die "vorläufige MDU-Lösung".

Das Problem ist nun folgendes: Werden die beiden Ladungsmatrizen  $\mathbf{G}$  und  $\mathbf{H}$  in einer gemeinsamen Matrix zusammengefasst, dann bedeutet dies - grafisch betrachtet - dass die beiden Koordinatensysteme für  $\mathbf{G}$  und  $\mathbf{H}$  aufeinander gelegt werden. Das ist jedoch nicht zulässig, da die Koordinatensysteme keinen gemeinsamen Ursprung besitzen und verschieden skaliert sind. Werden sie aufeinander gelegt, dann müssen sie zuvor "transformiert" werden und gegeneinander verschoben werden, so dass ihr Ursprung aufeinander liegt. Das geschieht beim Schönemann-Verfahren und seiner Variante in einem aufwendigen Kleinste-Quadrate-Kakül, der zur endgültigen MDU-Lösung  $\mathbf{F}$  führt.

Obwohl die vorläufige MDU-Lösung  $\mathbf{F0}$  nicht korrekt ist, so liefert sie doch einige bedeutsame Ergebnisse.

1. In der Regel besitzt sie bei empirischen Daten den vollen Rang. Die Multiplikation  $\mathbf{F0} * \mathbf{F0}'$  reproduziert die quadrierte und doppelt zentrierte Distanzmatrix  $\mathbf{X}$  exakt.

2. Wird aus  $\mathbf{F0}$  die Distanzmatrix gemäß Formel (8) reproduziert, dann wird die Distanzmatrix  $\mathbf{DH}$  der Abstände zwischen den Spaltenobjekten und die Distanzmatrix  $\mathbf{DG}$  der Abstände zwischen den Zeilenobjekten reproduziert.

3. Eingabefeld: Gesamte Distanzmatrix ausgeben

Die gesamte Distanzmatrix besteht, wie oben in Abschnitt P34.7.1 gezeigt, aus  $\mathbf{D}$ ,  $\mathbf{DA}$  und  $\mathbf{DB}$ . Wird in das Eingabefeld 1 eingesetzt, dann werden diese Matrizen als Gesamtmatrix ausgegeben. Die Matrix kann je nach Zahl der Zeilen- und Spaltenobjekte sehr umfangreich werden.

4. und 5. Eingabefeld: Matrix in Datei speichern

Folgende Möglichkeiten werden alternativ angeboten:

- 0 = wird nicht gespeichert. Dies ist auch die Voreinstellung.
- 1 = die eingegebene Distanzmatrix wird in eine Datei gespeichert
- 2 = die eingegebene Präferenzmatrix wird gespeichert



- 3 = die Distanzmatrix wird transponiert und als solche gespeichert, danach wird das Prog. beendet. Es wird also kein Ergebnis ausgegeben
- 4 = die gemeinsame Ladungsmatrix speichern

Zur Eingabe von 3:

Ist die eingegebene Matrix eine Distanzmatrix, dann wird dies in der transponierten Form gespeichert. Also beendet dann die Berechnung. Es werden keine Ergebnisse ausgegeben.

Wurde eine Präferenzmatrix eingegeben, dann wird diese zuerst in eine Distanzmatrix "umgedreht". Diese wird transponiert gespeichert. Dann wird das Programm ohne Ausgabe von Ergebnissen beendet. Beachte: Es wird also keine transponierte Präferenzmatrix sondern eine transponierte Distanzmatrix gespeichert.

### **6. Eingabefeld: Automatische Transponierung**

Wenn die Zahl der Spalten größer ist als die der Zeilen, dann wird die Distanzmatrix von Almo automatisch transponiert. Für den Benutzer gibt es 2 Möglichkeit

0 = der Benutzer erlaubt die Transponierung. Das ist auch die Voreinstellung.

1 = der Benutzer verhindert die Transponierung. Almo rechnet dann mit der "falschen" Distanzmatrix und liefert für diese auch die Ergebnisse.

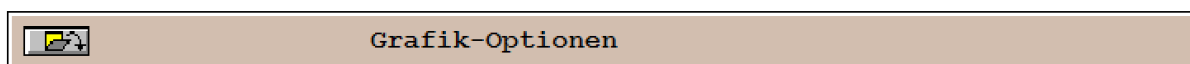
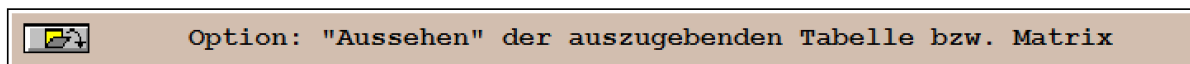
In nachfolgendem Abschnitt "P34.7.3.2 Problem: Spaltenzahl größer Zeilenzahl" wird ausführlich auf dieses Problem eingegangen.

### *7. Eingabefeld: Basisstatistiken ausgeben*

Bei Eingabe von 1 werden die Mittelwerte und Standardabweichungen der Spaltenvariablen ausgegeben.

### **Optionsbox: Aussehen der auszugebenden Tabelle**

#### **Optionsbox: Grafik-Optionen**



Zu diesen 2 Optionsboxen siehe Almo-Dokument Nr 0 "Arbeiten mit Almo" Abschnitt P0.7.3 und Abschnitt P0.7.4

### **P34.7.3.3 Problem: Spaltenzahl größer Zeilenzahl**

Wenn Die Distanzmatrix mehr Spalten als Zeilen besitzt, dann muss die Matrix transponiert werden. Die MDU-Ergebnisse sind für eine Matrix und ihre Transponierte verschieden. Würde die Matrix in ihrer "falschen", nicht-transponierten Form mit dem MDU-Kalkül analysiert werden, dann entstünden falsche Ergebnisse. Die "vorläufigen" Ergebnisse **G** und **H** aus der Singulärwert-Zerlegung von **X** aus der richtigen, d.h. transponierten und der falschen, nicht-transponierten Form sind allerdings identisch. Das bedeutet, dass die "falschen" Ergebnisse erst durch den eigentlichen Schönemann-Kalkül, d.h. den "Weg" zur Transformationsmatrix **T** und **a\*J** und der Multiplikation **A=G\*T** und **B = H\*inv(T)'-a\*J** entstehen.

Almo erkennt, dass die Spaltenzahl grösser ist als die Zeilenzahl. Es transponiert dann automatisch die Matrix. Der Benutzer kann das verhindern, wenn er im 6. Eingabefeld der Optionsbox "Weitere Einstellungen" eine 1 einsetzt. Dies hat aber nur dann einen Sinn, wenn

er die weiter unten beschriebene "alternative Vorgehensweise" einschlagen will.

Der in Almo programmierte Kalkül nach Schönemann erfordert es, dass die Zeilenzahl grösser/gleich der Spaltenzahl ist. Dabei spielt es keine Rolle, welche der beiden, Subjekte oder Objekte, die Zeilen oder die Spalten der Datenmatrix bilden. Ist jedoch die Zeilenzahl kleiner, dann muss die Matrix transponiert werden.

Wir wollen an einem Beispiel zeigen, dass die Ergebnisse, die Almo ausgibt, bei der richtigen und der falschen Form verschieden sind. Wir verwenden dazu die Beispiel-Matrix 3 von Schönemann, die in Prog18m6 eingesetzt ist.

	8*5 Distanzmatrix in der richtigen Form					5*8 Distanzmatrix in der falsche Form							
	4	3	3	4	6	4	3	3	2	4	4	3	4
	3	9	9	2	4	3	9	5	4	2	1	9	6
	3	5	6	4	2	3	9	6	4	3	2	9	7
	2	4	4	3	3	4	2	4	3	5	5	3	4
	4	2	3	5	5	6	4	2	3	5	5	2	2
	4	1	2	5	5								
	3	9	9	3	2								
	4	6	7	4	2								
		Faktor1	Faktor2			Faktor1	Faktor2						
Spaltobj1	-1.4382	-0.9737			3.2866	1.3404							
Spaltobj2	4.5376	0.2255			-5.1170	-0.6226							
Spaltobj3	4.5017	-0.9314			-5.1581	1.4344							
Spaltobj4	-2.0639	-1.5698			4.1289	2.3830							
Spaltobj5	-2.5055	1.2202			4.9793	-2.5927							
Zeilobj-1	2.1970	-2.5389			-2.0463	1.2473							
Zeilobj-2	-4.3112	-2.3061			2.5312	1.3214							
Zeilobj-3	-0.4993	2.4875			-0.2054	-1.4959							
Zeilobj-4	0.4525	0.1589			-0.8490	-0.2157							
Zeilobj-5	2.3104	0.6234			-2.1618	-0.5357							
Zeilobj-6	2.6130	0.5715			-2.3741	-0.5160							
Zeilobj-7	-4.5262	0.2137			2.6541	-0.0896							
Zeilobj-8	-1.2678	2.8193			0.3316	-1.6584							
Stress-1:		0.0657				0.3573							

Die Ladungsmatrizen sind erheblich verschieden. Der Stress-Koeffizient der falschen Form ist erheblich schlechter als der der richtigen Form. Interessant ist der Vergleich mit anderen MDU-Programmen. Das SPSS-Programm PREFSCAL liefert bei der falschen Form mit 0.3473 ebenfalls einen sehr schlechten Stress-Wert. Bei ALSCAL ist das Ergebnis für falsche und richtige Form identisch. Es ist zu vermuten, dass Alscal die falsche Form intern in die richtige transponiert und dann erst rechnet.

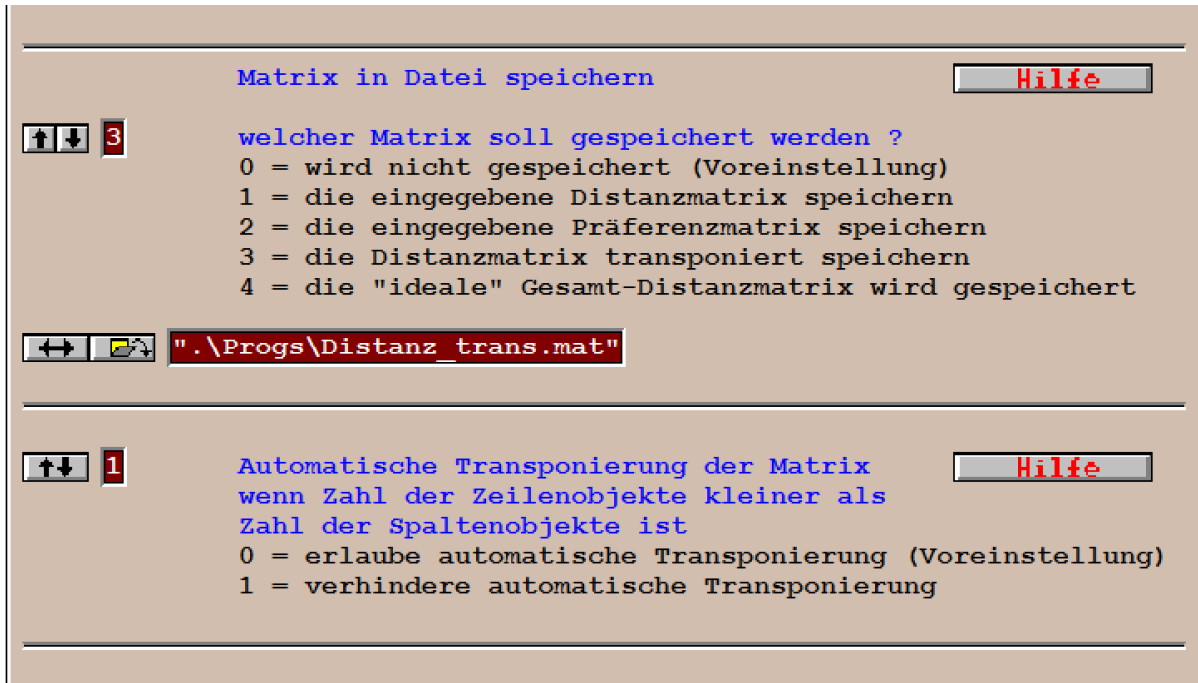
Die automatische Transponierung hat folgenden Nachteil: Die Spaltenobjekte sind nun zu Zeilenobjekten geworden und umgekehrt. Die Variablen-Nummern und, falls der Benutzer Variablenamen vergeben hat, auch diese, können dann von Almo in den Ergebnis-Tabellen nicht mehr eingesetzt werden. Almo behilft sich dadurch, dass es als Namen schlicht und einfach die Worte *Zeilobj-1*, *Zeilobj-2*,... usw. und *Spaltobj-1*, *Spaltobj-2*,...usw. in der Beschriftung der Tabellen verwendet. Wenn der Benutzer jedoch Wert auf eine gut lesbare Ergebnis-Ausgabe legt, dann ist folgende Alternative empfehlenswert.

### Alternative Vorgehensweise

1. Der Benutzer verhindert die automatische Transponierung der Matrix, indem er in der Optionsbox "weitere Einstellungen" in das 6. Eingabefeld

"automatische Transponierung ..." eine 1 einsetzt.

2. Im 5. Eingabefeld "Matrix in Datei speichern" wird eine 3 (=transponierte Distanzmatrix speichern) eingesetzt und im 5. Eingabefeld wird ein Dateiname angegeben. Der Name ist beliebig. Empfehlenswert ist es in den Ordner "Progs" zu speichern. Die Optionsbox "weitere Einstellungen" hat dann folgendes Aussehen



Almo reagiert in folgender Weise:

- a. Ist die Matrix eine Präferenzmatrix, dann wird diese zuerst in eine Distanzmatrix "umgedreht". Dieser Schritt entfällt selbstverständlich wenn die eingegebene Matrix bereits eine Distanzmatrix ist.
  - b. Die Distanzmatrix wird in die vom Benutzer im 5. Eingabefeld angegebene Datei in transponierter Form gespeichert
  - c. dann wird das Programm - ohne weiter zu rechnen - beendet.
3. Der Benutzer verfügt nun über die transponierte Distanzmatrix. Er rechnet nun mit Prog18m6 ein neues Programm.
    - a. In der Box "Datei aus der gelesen wird " muss der Name der transponierten Matrix eingetragen werden
    - b. in der Box "Spaltenvariable (Spaltenobjekte)" wird geschrieben:  $V1:x$  , wobei für x die Zahl der Spalten der transponierten Variablen einzusetzen ist. Der Typ der transponierten Matrix ist nun 0 (Distanzmatrix)
    - c. Wenn Variablennamen verwendet werden sollen, dann muss in der Box "Freie Namensfelder" den Spaltenvariablen, die zuvor die Zeilenvariablen waren, Namen gegeben werden.
  4. Das so veränderte Programm wird gerechnet

### P34.7.4 Eingabe in Almo-Programm Prog18m7

Das Programm erzeugt zuerst aus den Mittelwerten gruppierter Daten eine Distanz- bzw. Präferenz-Matrix. Danach wird ein metrisches MDU gerechnet

Personen geben Präferenz- oder Distanz-Urteile gegenüber Objekten ab. Die Personen werden nach Gruppierungs-Variablen (z.B. Beruf und Geschlecht) zusammengefasst.

Beispiel: Folgende Datenmatrix liegt vor

Gruppierungs-Variable		Spalten-Objekte				
Geschlecht	Beruf	Part.A	Part.B	Part.C	Part.D	Part.E
V1	V2	V2	V3	V4	V5	V6
----	----	---	---	---	---	---
1	2	5	6	6	5	3
2	1	5	7	6	4	4
1	4	7	5	5	6	6
1	2	6	0	0	7	5
2	2	5	8	7	4	4
1	3	6	4	3	5	7
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

Die Person in der 1. Zeile ist weiblich (Code: 1) und von Beruf Angestellte (Code: 2). Ihre Präferenz gegenüber der Partei A beträgt 5 Einheiten, gegenüber der Partei B 6 Einheiten usw. Personen derselben Merkmalsgruppe werden zusammengefasst. Ihre mittlere Präferenz gegenüber den 5 Parteien wird errechnet. So entsteht eine Präferenzmatrix für gruppierte Daten, die dann dem MDU-Verfahren unterworfen wird.

Die Zahl der Gruppierungsvariablen ist nicht beschränkt. Es kann auch nur eine Variable zur Gruppierung eingesetzt werden

Prog18m7 besitzt gegenüber dem ausführlich erläuterten Prog18m6 nur folgende zusätzliche Eingabebox

**Gruppierungs-Variable (Zeilenobjekte)**

**Geschlecht, Beruf**

Sie können eine oder mehrere Variable eingeben  
 oder auch das Eingabefeld leer lassen
 **Hilfe**

Hier werden die oben beschriebenen beiden Gruppierungsvariablen eingetragen. Wenn der Benutzer diesen beiden keine Variablennamen gegeben hat, dann müssen die Variablennummern eingeschrieben werden. Das wären im Beispiel:

V1, 2

### P34.7.5 Ausgabe aus Prog18m6 und Prog18m7

Die Ausgabe der beiden Programme ist identisch. Bei Prog18m7 werden nur vorneweg

folgende zusätzliche Informationen ausgegeben:

Für Analyse ausgewählte Variable

V1 Geschlecht: weibl männl  
 V2 Beruf: Arbeiter Angestellter Beamter Selständig  
 V3 PartA  
 V4 PartB  
 V5 PartC  
 V6 PartD  
 V7 PartE

V1 wird auch bezeichnet mit A  
 die Ausprägungen (bzw.Dummies) mit  
 A1 =weibl  
 A2 =männl

V2 wird auch bezeichnet mit B  
 die Ausprägungen (bzw.Dummies) mit  
 B1 =Arbeiter  
 B2 =Angestellter  
 B3 =Beamter  
 B4 =Selständig

Die Zeilenobjekte entstehen aus der Kombination  
 der Ausprägungen (bzw. Dummies) der Gruppierungsvariablen V1 V2  
 Sie werden in den Ergebnis-Tabellen der MDU in folgender Weise bezeichnet:

- 1 A1B1
- 2 A1B2
- 3 A1B3
- 4 A1B4
- 5 A2B1
- 6 A2B2
- 7 A2B3
- 8 A2B4

\*\*\*\*\* Erläuterung: In den Ergebnis-Tabellen der Almo-MDU werden die Zeilenobjekte mit diesen Bezeichnungen benannt.

Mittelwerte der Präferenzen  
 (arithmetisches Mittel)

Geschl Beruf		PartA	PartB	PartC	PartD	PartE
		V3	V4	V5	V6	V7
weibl	Arbeit	5.3000	6.3000	6.2800	5.2600	3.2200
	Angest	6.2333	0.2500	0.2000	7.2000	5.2500
	Beamte	6.3500	4.1833	3.2500	5.2167	7.2000
	Selstä	7.2333	5.2833	5.2833	6.2167	6.2333
männl	Arbeit	5.2333	7.2167	6.2667	4.3167	4.2500
	Angest	5.2333	8.1833	7.2000	4.2333	4.2500
	Beamte	6.2000	0.2667	0.2333	6.1667	7.3000
	Selstä	5.2333	3.2500	2.2500	5.2333	7.2833
Gesamtmittel		5.8894	4.3255	3.8191	5.4851	5.6745

Danach werden die eigentlichen MDU-Ergebnisse ausgegeben. Wir zeigen diese jedoch für **Prog18m6**, dessen Eingabe wir oben ausführlich erläutert haben.

## Ergebnisse aus Prog18m6

Zuerst werden die vom Benutzer gewählten Einstellungen und "weiteren Einstellungen" wiederholt. Dann kommen die eigentlichen Ergebnisse. Im folgenden zeigen wir die Ergebnisse der Standardausgabe ohne Zwischenergebnisse.

Distanzmatrix

-----

Distanzen zwischen Zeilenobjekten und Spaltenobjekten

	PartA	PartB	PartC	PartD	PartE
ZeilObj 1	4.0000	3.0000	3.0000	4.0000	6.0000
ZeilObj 2	3.0000	9.0000	9.0000	2.0000	4.0000
ZeilObj 3	3.0000	5.0000	6.0000	4.0000	2.0000
ZeilObj 4	2.0000	4.0000	4.0000	3.0000	3.0000
ZeilObj 5	4.0000	2.0000	3.0000	5.0000	5.0000
ZeilObj 6	4.0000	1.0000	2.0000	5.0000	5.0000
ZeilObj 7	3.0000	9.0000	9.0000	3.0000	2.0000
ZeilObj 8	4.0000	6.0000	7.0000	4.0000	2.0000

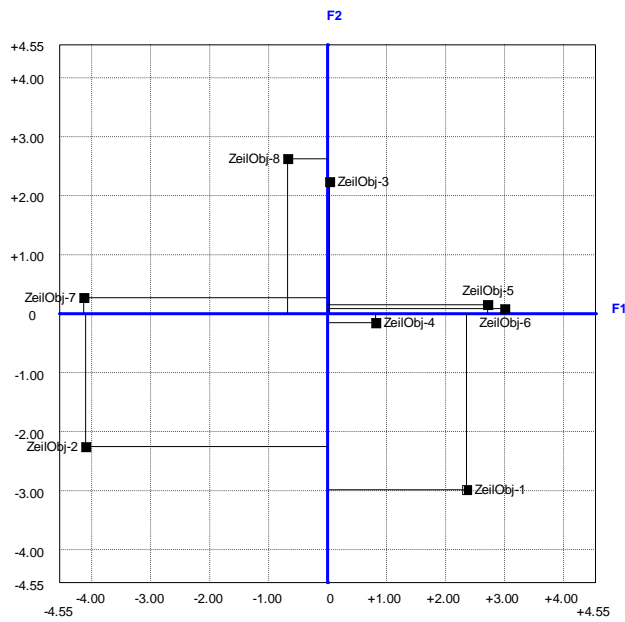
\*\*\*\*\* Erläuterung: Da den Spaltenobjekten in der Programm-Maske Namen gegeben wurden, können diese nun von Almo in der Matrixausgabe verwendet werden. Den Zeilenobjekten kann kein Name gegeben werden.

-----  
Lösung aus Zerlegung der quadrierten und doppelt zentrierten Distanzmatrix  
nach Kalkül von Schönemann  
-----

Matrix A der Zeilenvariablen

	Faktor 1	Faktor 2
ZeilObj-1	2.3569	-2.9798
ZeilObj-2	-4.1150	-2.2544
ZeilObj-3	0.0493	2.2365
ZeilObj-4	0.8219	-0.1575
ZeilObj-5	2.7097	0.1648
ZeilObj-6	3.0074	0.0902
ZeilObj-7	-4.1384	0.2745
ZeilObj-8	-0.6918	2.6256

Ladungen der Zeilenvariablen



\*\*\*\*\* Erläuterung: Die Ladungsmatrix der Zeilenobjekte und das entsprechende 2-dimensionale Koordinatensystem werden ausgegeben.

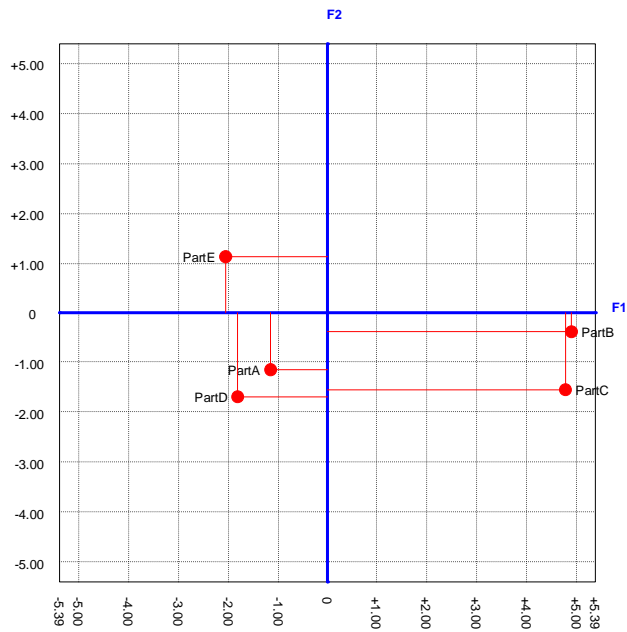
Danach erfolgt die entsprechende Ausgabe für die Spaltenobjekte

Matrix B der Spaltenvariablen

-----

	Faktor 1	Faktor 2
PartA	-1.1493	-1.1435
PartB	4.9003	-0.4007
PartC	4.7767	-1.5516
PartD	-1.8183	-1.6905
PartE	-2.0472	1.1250

Ladungen der Spaltenvariablen



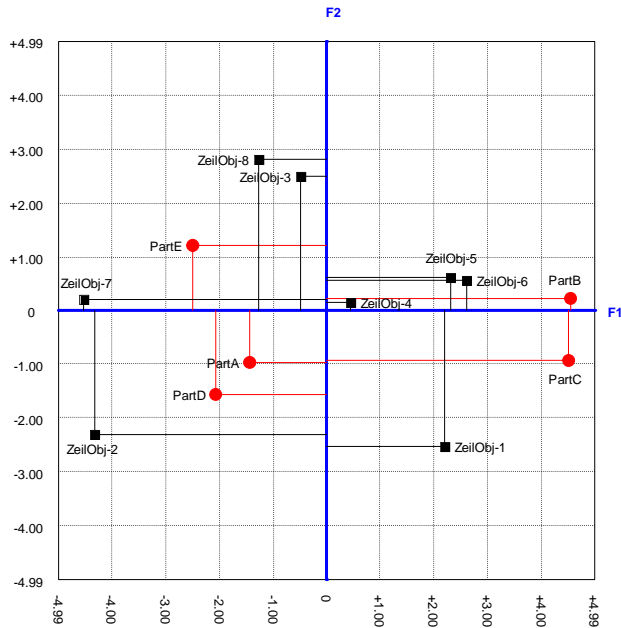
endgültige gemeinsame normierte und orthogonalisierte Matrix

-----  
 aus Spaltenobjekten B (Zeile 1 bis 5)  
 und Zeilenobjekten A (Zeile 6 bis 13)

	Faktor 1	Faktor 2
PartA	-1.4382	-0.9737
PartB	4.5376	0.2255
PartC	4.5017	-0.9314
PartD	-2.0639	-1.5698
PartE	-2.5055	1.2202
ZeilObj-1	2.1970	-2.5389
ZeilObj-2	-4.3112	-2.3061
ZeilObj-3	-0.4993	2.4875
ZeilObj-4	0.4525	0.1589
ZeilObj-5	2.3104	0.6234
ZeilObj-6	2.6130	0.5715
ZeilObj-7	-4.5262	0.2137
ZeilObj-8	-1.2678	2.8193



Ladungen der Spaltenvariablen (rote Punkte)  
und der Zeilenvariablen (schwarze Punkte)



\*\*\*\*\*Erläuterung: Die Ladungsmatrix der Spalten- und Zeilenobjekte werden zusammengefasst und die entsprechende Grafik gezeichnet.

Matrix der vom Modell reproduzierten Distanzen  
Distanzen zwischen Zeilenobjekten und Spaltenobjekten

	PartA	PartB	PartC	PartD	PartE
ZeilObj 1	3.9579	3.6222	2.8099	4.3697	6.0204
ZeilObj 2	3.1669	9.2038	8.9195	2.3648	3.9618
ZeilObj 3	3.5863	5.5215	6.0580	4.3486	2.3729
ZeilObj 4	2.2040	4.0856	4.1934	3.0530	3.1427
ZeilObj 5	4.0747	2.2624	2.6868	4.8934	4.8528
ZeilObj 6	4.3359	1.9555	2.4137	5.1438	5.1594
ZeilObj 7	3.3084	9.0638	9.1002	3.0404	2.2575
ZeilObj 8	3.7968	6.3584	<b>6.8814</b>	4.4607	2.0221

\*\*\*\*\* Erläuterung: Die Güte der MDU-Lösung wird durch den Vergleich der empirischen, d.h. der eingegebenen und der durch das Modell reproduzierten Distanzmatrix sichtbar. Der Stress-1-Koeffizient drückt diesen Vergleich in einer Zahl aus. Je kleiner diese ist, umso besser ist die Reproduktion gelungen. Der gefundene Wert von 0.0657 ist ein sehr guter Wert.

Der Stress-2-Koeffizient entsteht als Durchschnitt aus den Stress-Koeffizienten je Zeile. Er ist mit Stress-1 nicht direkt vergleichbar.

Stress-1 nach Kruskal  
bezogen auf die Distanzen zwischen Zeilenobjekten und Spaltenobjekten

Summe der quadrierten Differenzen zwischen  
reproduzierter und empirischer Distanzmatrix= 3.7842  
Stress-1 nach 2 extrahierten Faktoren= 0.0657

Stress-2 nach Kruskal  
bezogen auf die Distanzen zwischen Zeilenobjekten und Spaltenobjekten  
Stress je Zeile

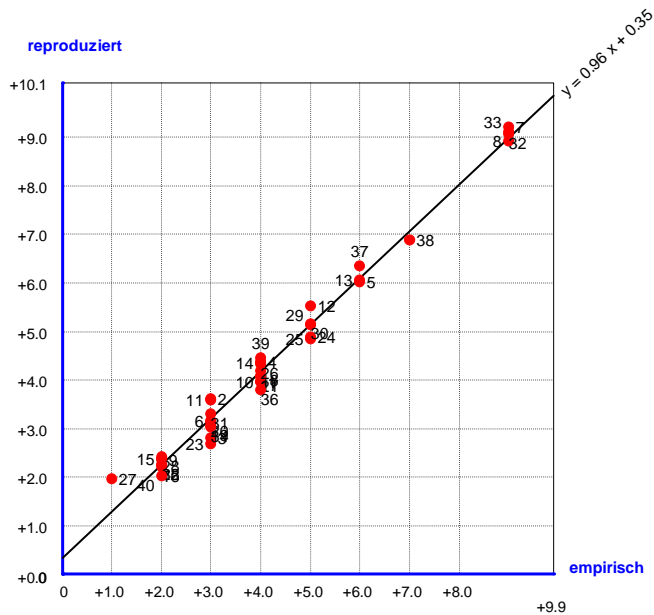
- 1 = 0.3061
- 2 = 0.0682
- 3 = 0.2966

4 = 0.1978  
 5 = 0.1739  
 6 = 0.3069  
 7 = 0.0603  
 8 = 0.1615

Stress-2 nach 2 extrahierten Faktoren= 0.2178

Almo zeichnet dann noch folgendes Streudiagramm

Streudiagramm



In diesem Diagramm werden die Distanzen aus der empirischen Distanzmatrix  $D$  und aus der durch das MDU-Modell reproduzierten Distanzmatrix  $D^*$  verglichen. Der Punkt 38 im Streudiagramm entspricht den beiden Werten in der 38. Zelle der beiden Matrizen. In  $D$  steht dort der Wert 7.0 und in  $D^*$  6.8814. Die Zellen der Matrix werden von links nach rechts, eine Zeile nach der anderen, durchnummeriert.

Der Punkt 38 besitzt die Koordinaten 7.0 auf der x-Achse und 6.8814 auf der y-Achse. Die x-Achse wird in der Grafik mit "empirisch" und die y-Achse mit "reproduziert" bezeichnet.

Der Idealzustand ist erreicht, wenn alle Punkte auf der Regressionsgeraden liegen und diese die Gleichung  $y = 1.0 * x + 0$  besitzt. Je enger die Punkte um die Regressionsgerade herum liegen umso besser.

### Ausgabe mit Zwischenergebnissen Modus 1

Wird die Optionsbox "weitere Einstellungen" geöffnet und *Zwischenergebnisse=1* gesetzt, dann werden die Zwischenschritte des Schönemann-Kalküls ausgegeben.

### Ausgabe mit "vorläufiger" MDU-Lösung

Wird die Optionsbox "weitere Einstellungen" geöffnet und *Lösung für vorläufige Zerlegung der Distanzmatrix = 1* gesetzt, dann werden die Singulär-Zerlegung der quadriert und doppelt zentrierten Distanzmatrix  $X$  und die ersten Schritte des Kalküls ausgegeben. Siehe die Gleichungen 1 bis 5 im Abschnitt P34.7.2 "Kalkül der MDU". Im nachfolgenden Exkurs P34.7.6 wird die Ergebnis-Ausgabe erläutert.

## P34.7.6 MDU-Analyse einer perfekten, fehlerfreien Distanzmatrix

Die nachfolgende Simulation kann mit dem Beispielprogramm "perfDist.Alm" nachgerechnet werden. Das Programm findet man im Menü "Almo/Liste aller Almo-Programme". Es kann auch aus der Programm-Maske Prog18m6 heraus gestartet werden.

Folgende 8\*5-Distanzmatrix wird analysiert:

3.9579	3.6222	2.8099	4.3697	6.0204
3.1669	9.2038	8.9195	2.3649	3.9618
3.5863	5.5215	6.0580	4.3486	2.3729
2.2040	4.0856	4.1934	3.0530	3.1427
4.0747	2.2624	2.6868	4.8934	4.8528
4.3359	1.9555	2.4137	5.1438	5.1594
3.3084	9.0638	9.1002	3.0404	2.2575
3.7968	6.3584	6.8814	4.4607	2.0221

Dies ist keine empirisch gewonnene Distanzmatrix. Sie wurde aus folgender 2-faktorieller Ladungsmatrix erzeugt

Faktor 1	Faktor 2
-----	-----
-1.438240	-0.973718
4.537593	0.225515
4.501671	-0.931425
-2.063912	-1.569846
-2.505509	1.220239
2.197010	-2.538927
-4.311220	-2.306137
-0.499342	2.487494
0.452515	0.158901
2.310438	0.623418
2.612958	0.571550
-4.526202	0.213677
-1.267759	2.819261

Mit Hilfe von Prog30mf wurde aus dieser Ladungsmatrix obige Distanzmatrix erzeugt. Die so erzeugte Distanzmatrix ist dadurch fehlerfrei und perfekt. Anders formuliert: Diese Distanzmatrix muss, wenn sie nach dem Singulärwert-Verfahren zerlegt wird, genau 2 Faktoren besitzen.

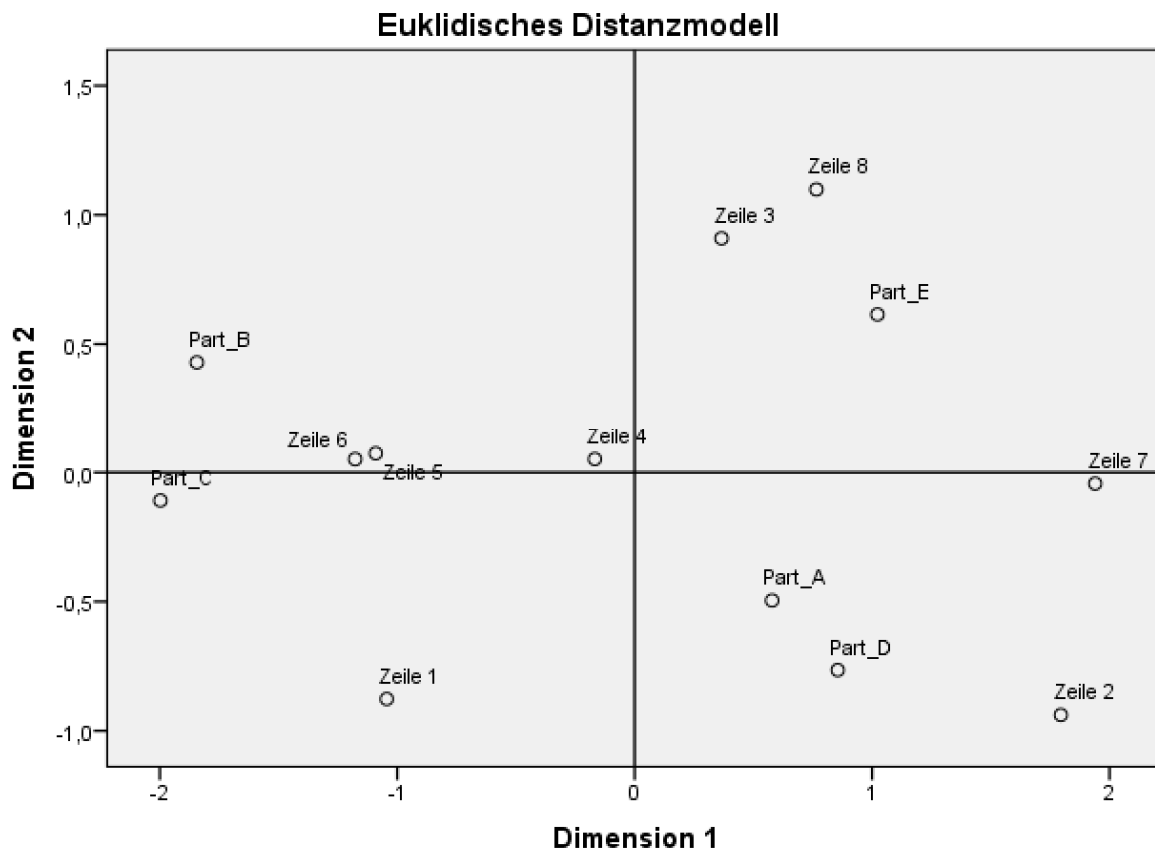
Das Almo-MDU-Programm, das nach dem Schönemann-Kalkül programmiert ist, reproduziert diese Distanzmatrix exakt. Es entsteht somit ein Stress-1-Wert von .0. Das Streudiagramm liefert eine ideale Regressionsgleichung von  $y=1.0 \cdot x+0$ . Der Benutzer kann sich davon überzeugen, wenn er das erwähnte Beispielprogramm "perfDist.Alm" rechnet.

Eine Anmerkung: Die zur Simulation verwendete Ladungsmatrix ist die beim Schönemann-Beispiel 3 (in Prog18m6) entstehende Ergebnis-Matrix. Wir haben diese jedoch nur mit 6 Kommastellen in das Prog30mf (das die Distanzmatrix erzeugt) eingegeben - so dass die Distanzmatrix nicht absolut perfekt 2-faktoriell ist.

## P34.7.7 Vergleich Almo mit Alscal und Prefscal

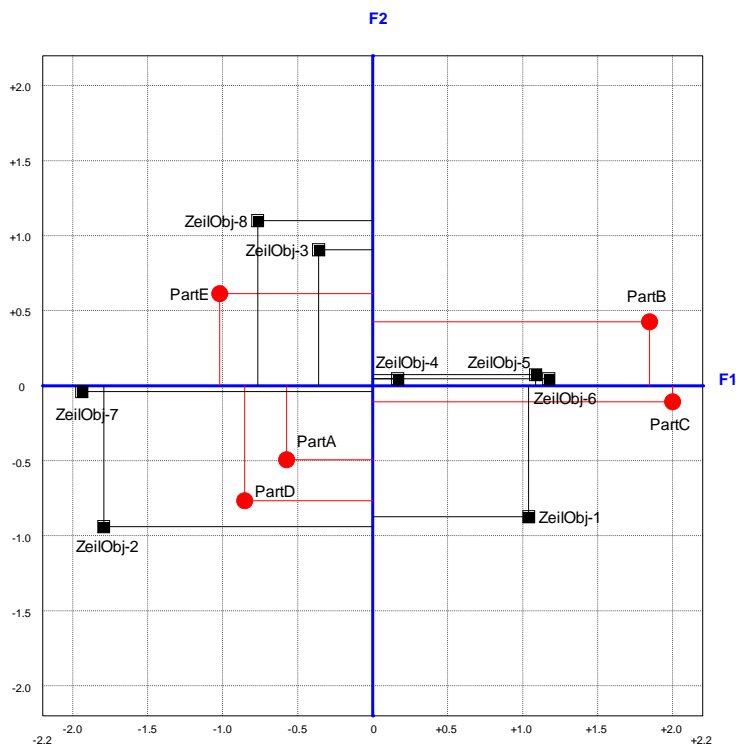
Die wesentlichen Ergebnisse aus Almo sollen mit den beiden SPSS-MDU-Programmen "Alscal" und "Prefscal" verglichen werden. Wir rechnen mit diesen 3 MDU-Programmen unsere Beispieldaten. Die "sav"-Datei für SPSS ist unter ".\Testdat\FuenfParteien.sav" zu finden. Wir wollen zuerst die grafischen Darstellungen der Ladungsmatrizen vergleichen

## Grafik aus Alscal

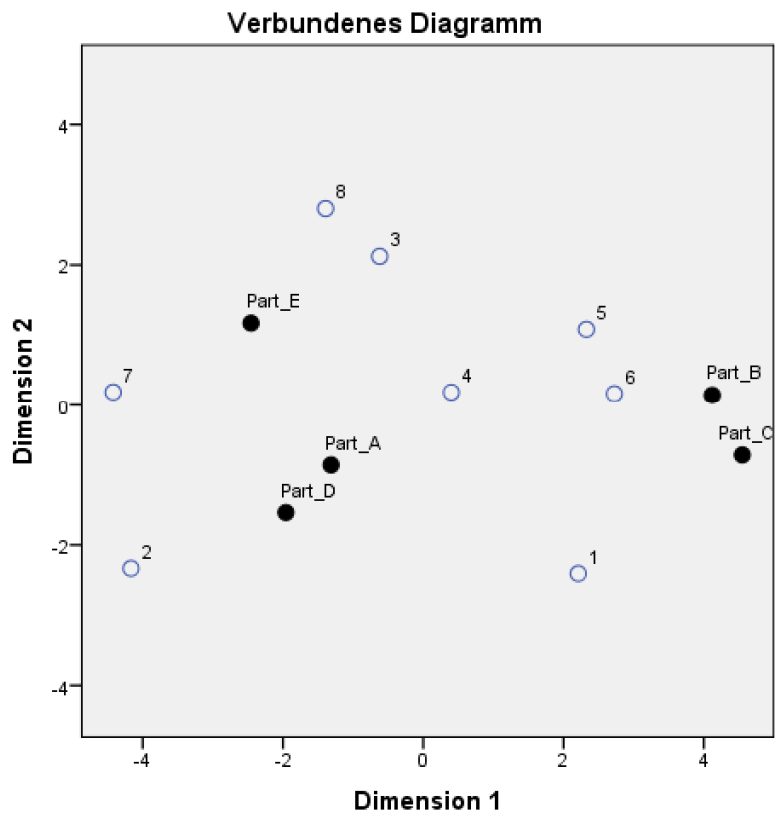


## Grafik aus Alscal (im Almo-Grafik-Editor um Achse F1 gespiegelt)

Die Alscal-Grafik wird um die Dimension 1 gespiegelt (Vorzeichen-Umkehr) damit sie mit der Almo- und Prefscal-Grafik optisch besser verglichen werden kann. Dazu wurde der Almo-Grafik-Editor verwendet.

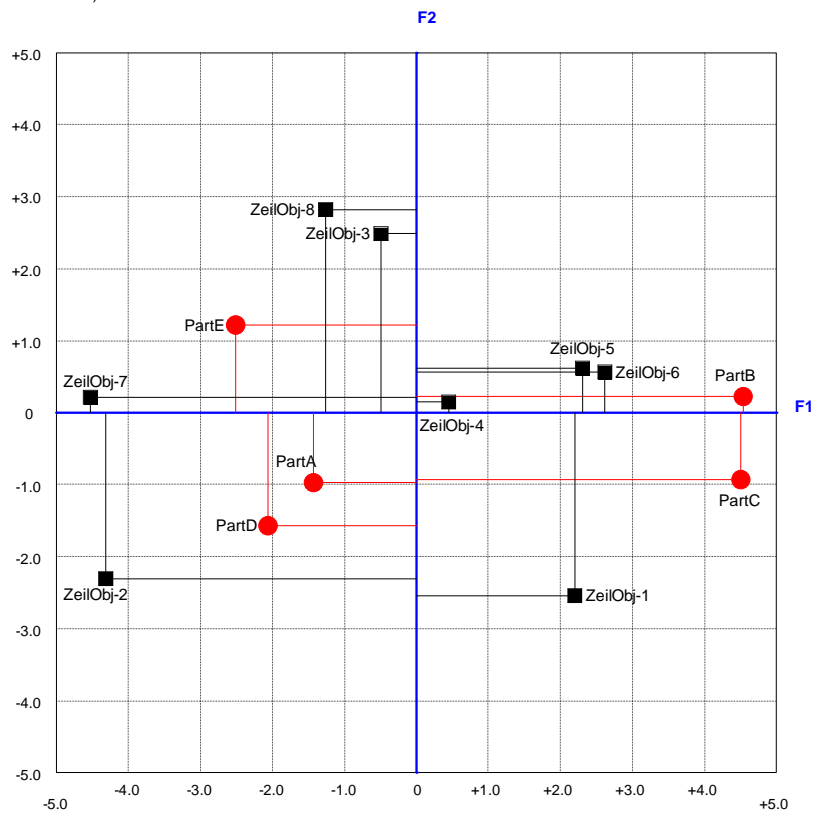


## Grafik aus Prefscal



## Grafik aus Almo

Ladungen der Spaltenvariablen (rote Punkte)  
und der Zeilenvariablen (schwarze Punkte)



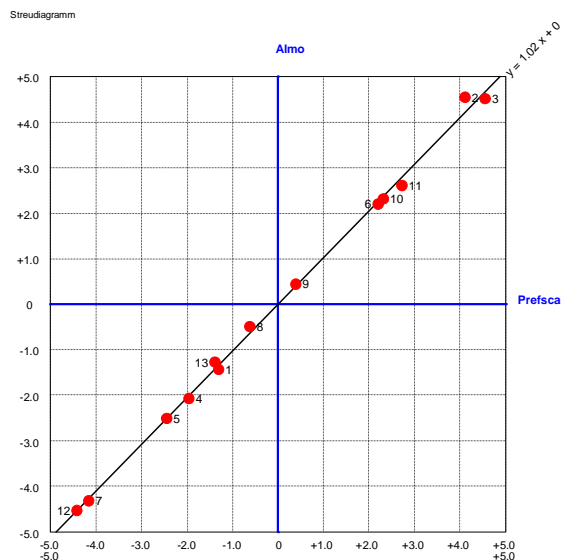
Die 5 Parteien PartA bis PartE werden in Alscal und Prefscal mit F1 bis F5 bezeichnet. Die Punktekonfigurationen sind außerordentlich ähnlich. Auffällig sind die kleineren Masseneinheiten für die beiden Achsen bei Alscal.

Die Ladungsmatrizen für die 3 Verfahren sind:

	Almo		Prefscal		Alscal	
	Faktor 1	Faktor 2	Faktor 1	Faktor 2	Faktor 1	Faktor 2
PartA	-1.4382	-0.9737	-1,315	-0,860	,5825	-,5000
PartB	4.5376	0.2255	4,120	0,143	-1,8719	,3099
PartC	4.5017	-0.9314	4,548	-0,718	-1,9869	-,3947
PartD	-2.0639	-1.5698	-1,958	-1,537	,8757	-,7806
PartE	-2.5055	1.2202	-2,455	1,172	1,0255	,6654
ZeilObj-1	2.1970	-2.5389	2,210	-2,407	-,9862	-,8751
ZeilObj-2	-4.3112	-2.3061	-4,167	-2,334	1,7879	-,8298
ZeilObj-3	-0.4993	2.4875	-0,621	2,123	,3808	,9580
ZeilObj-4	0.4525	0.1589	0,400	0,183	-,1446	,0191
ZeilObj-5	2.3104	0.6234	2,326	1,083	-1,0455	,1343
ZeilObj-6	2.6130	0.5715	2,725	0,164	-1,2757	,1060
ZeilObj-7	-4.5262	0.2137	-4,421	0,184	1,8520	,0494
ZeilObj-8	-1.2678	2.8193	-1,392	2,804	,8066	1,1385

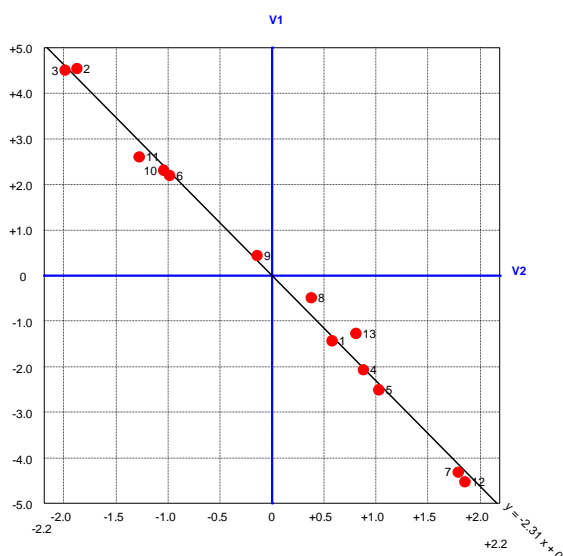
Stress-1 (Kruskal) nach 2 extrahierten Faktoren  
 Almo = 0.0657  
 Prefscal = 0.0438  
 Alscal = 0.0700

Wir haben den 1. Faktor von Almo und Prefscal in einem Streudiagramm gezeichnet, und eine Regressionsgerade mit einer Steigung von 1.02 gefunden. Die Ladungen von Almo auf dem 1. Faktor sind also durchschnittlich 1.02 mal grösser als die von Prefscal. Wie in nachfolgender Grafik zu sehen ist, liegen die Punkte auf oder sehr dicht bei der Regressionsgeraden.



Beim Vergleich von Almo und Alscal fällt auf, dass der 1. Faktor eine Vorzeichen-Umkehrung aufweist (was bedeutungslos ist) und dass die Ladungen von Almo durchschnittlich 2.31 mal grösser sind als die von Alscal - wobei, wie beim Vergleich der Grafiken zu sehen war, die Punktekonfigurationen der beiden fast gleich sind. Mit den Ladungen von Alscal kann deswegen auch nicht korrekt die reproduzierte Distanzmatrix  $D^*$  nachgerechnet werden.

Streudiagramm



Die Ergebnisse aus den 3 Programmen sind sehr ähnlich. Das ist nicht immer so. Der Benutzer rechne das Almo-Beispielprogramm "AllbusPart.Alm". Es rechnet eine MDU-Analyse mit der  $20 \times 7$  Matrix "WuppPartPraef.fre" der Allbus-Umfrage zur Parteiensympathie 1994 in West- und Ostdeutschland. Im Ordner "Testdat" ist die SPSS-Datei "WuppPartPraef.sav" enthalten. Zum Vergleich kann damit eine MDU-Analyse mit Alscal und Prefscal gerechnet werden. Die Ergebnisse aus den 3 Programmen sind zwar ähnlich aber nicht so überzeugend ähnlich wie bei unseren obigen Beispieldaten. So weicht Prefscal recht deutlich von Alscal und Almo ab, die zueinander ähnlicher sind. Unser Eindruck ist, dass die Ähnlichkeiten umso geringer sind, umso größer die Distanzmatrizen sind. Alscal und Prefscal sind iterative Verfahren. Sie versuchen eigens konstruierte aber verschiedene Stress-Koeffizienten zu minimieren. Siehe dazu Busing (2010). Dabei ist es kaum möglich zu entscheiden, welches Kriterium das bessere ist. So ist es verständlich, dass sie auch verschiedene Ergebnisse erbringen.

Eine Möglichkeit wäre es, die "Leistungsfähigkeit" konkurrierender Verfahren dadurch zu messen, dass man den Stress-1-Koeffizienten (nach Kruskal) vergleicht. Die iterativen, "Stress-minimierenden" Verfahren erzielen häufig gegenüber dem algebraischen Schönemann-Verfahren einen etwas besseren Stress-1-Wert. Sie erreichen dies auch dadurch, dass sie (grafisch gesehen) die Punkte der Zeilenobjekte hin- und herschieben, d.h. die Distanzen zwischen den Zeilenobjekten im Verlauf der Rechen-Iterationen gemäß ihrem Minimierungskriterium ständig verändern. Entsprechend so auch bei den Spaltenobjekten. Da die wahren Distanzen zwischen den Zeilenobjekten und den Spaltenobjekten nicht bekannt sind, kann nicht entschieden werden, wie legitim diese Vorgehensweise ist. Siehe dazu im Almo-Dokument Nr.28 "Metrische MDS", Abschnitt P34.6.5.1 und im Almo-Dokumen Nr.30 "Nicht-metrische MDS", Exkurs: Vergleich metrische MDS und iterative MDS.

Einen wichtigen Vorteil haben Alscal und Prefscal gegenüber dem Schönemann-Kalkül: Sie können auch ordinale Distanzen analysieren.

## P34.7.8 Anhang

### P34.7.6.1 Notation: Matrix-Bezeichnungen

**A** = Ladungsmatrix der Zeilenobjekten  
**a\*J** = "Verschiebe"-Matrix (von **H** zu **B**). Siehe Gleichung 7  
**B** = Ladungsmatrix der Spaltenobjekte  
**D** = empirische Matrix der euklidischen Distanzen zwischen Zeilenobjekten einerseits und Spaltenobjekten andererseits  
**D0** = Matrix der Distanzen zwischen Zeilenobjekten einerseits und Spaltenobjekten andererseits aus "vorläufiger Lösung"  
**D\*** = vom MDU-Modell reproduzierte Matrix der Distanzen zwischen Zeilen- und Spalten-Objekten  
**DA** = Matrix der Distanzen zwischen den Zeilenobjekten  
**DB** = Matrix der Distanzen zwischen den Spaltenobjekten  
**DG** = Distanzen zwischen Zeilenobjekten aus "vorläufiger Lösung"  
**DH** = Distanzen zwischen Spaltenobjekten aus "vorläufiger Lösung"  
**F** = gemeinsame Ladungsmatrix der Spalten- und Zeilenobjekten  
**F0** = gemeinsame Ladungsmatrix der Spalten- und Zeilenobjekten aus "vorläufiger Lösung"  
**G** = Ladungsmatrix der Zeilenobjekten aus "vorläufiger Lösung"  
**H** = Ladungsmatrix der Spaltenobjekten aus "vorläufiger Lösung"  
**gD\*** = vom MDU-Modell reproduzierte Gesamt-Distanzmatrix zwischen allen Objekten  
**T** = Transformationsmatrix (von **G** zu **A**). Siehe Gleichung 6  
**U** = Links-Eigenvektoren von **X**  
**V** = Rechtss-Eigenvektoren von **X**  
**W** = Wurzel der Eigenwerte von **X**  
**X** = quadriert und doppelt zentrierte Distanzmatrix

### P34.7.6.1 Singulärwert-Zerlegung einer rechteckigen nicht-symmetrischen Matrix

**X** ist eine n\*m- Rechteck-Matrix. Sie wird zerlegt in

$$(1) \mathbf{X} = \mathbf{L} \cdot \mathbf{W} \cdot \mathbf{R}'$$

**L** = Links-Eigenvektoren  
**R'** = Rechts-Eigenvektoren (transponiert)  
**W** = die Diagonalmatrix der Wurzeln der Eigenwerte  
**n** = Zahl der Zeilen von **X**  
**m** = Zahl der Spalten von **X**

**X'\*X** hat Dimension m\*m.

**X\*X'** hat Dimension n\*n

beide sind symmetrische Matrizen.

**X\*X'** muss im Verlauf des Kalkuels nicht ermittelt werden,

jedoch die m\*m Matrix **X'\*X**

Es ist für den Kalkül nicht notwendig aber vorteilhaft, wenn die Spaltenzahl m von **X** kleiner ist als die Zeilenzahl n. Die Rechenzeit ist dann kürzer.

Die Matrix **X'\*X** wird mit einem gängigen Eigenwert-Eigenvektor-Verfahren (in Almo mit dem Jacobi-Verfahren) zerlegt. Dabei entstehen **R** und **W** aus obiger Gleichung 1.



$p$  = Rang von  $\mathbf{X}'\mathbf{X}$  (Zahl der Eigenwerte  $>0$ )  
besitzt  $\mathbf{X}'\mathbf{X}$  den vollen Rang, dann ist  $p=m$   
 $\mathbf{W}$  = die Diagonalmatrix der Wurzeln der Eigenwerte von  $\mathbf{X}'\mathbf{X}$  ( $p \times p$ )  
diese sind identisch mit denen von  $\mathbf{X}\mathbf{X}'$ .  
 $\mathbf{R}$  = die Matrix der Rechts-Eigenvektoren von  $\mathbf{X}'\mathbf{X}$  ( $m \times p$ )

Die Matrix der Links-Eigenvektoren  $\mathbf{L}$  entsteht aus

$$(2) \mathbf{L} = \mathbf{X}\mathbf{R}^*(1/\mathbf{W})$$

$(1/\mathbf{W})$  = das ist die Diagonalmatrix  $\mathbf{W}$  mit Kehrwerten in der Diagonale  
 $\mathbf{L}$  ist eine  $n \times p$ -Matrix

Wird die Matrix  $\mathbf{X}$  transponiert, dann sind die Eigenvektoren  $\mathbf{L}$  und  $\mathbf{R}$   
gegenüber der nicht-transponierten Matrix zwar vertauscht aber wertegleich.  
Auch die Eigenwerte  $\mathbf{W}$  sind identisch.

## Literatur

Die Literatur zur MDU ist sehr umfangreich und auch sehr zerstreut über viele Bücher und Zeitschriften-Artikel. Sie wird sehr häufig zusammen mit der MDS dargestellt. Einen sehr guten Überblick über die MDS ( leider kaum über die MDU) gibt

**Borg, I., Groenen, P.:** Modern multidimensional scaling: theory and applications, 2005, New York, Springer.

und speziell über die (iterative, stress-minimierende) MDU

**Busing, Frank:** Advances in multidimensional Unfolding, 2010, im Internet herunterladbar bei <https://openaccess.leidenuniv.nl/bitstream/handle/1887/15279/binnenwerk.pdf?sequence=16>

---

**Almo-Handbuch:** Sozialwissenschaftliche Skalierungsverfahren, Abschnitt P34 "Nichtmetrische multidimensionale Skalierung"

**Borg, I.:** Multidimensionale Skalierung, in Wolf/Best(Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse, 2010, VS Verlag Wiesbaden

**BUGH** Wuppertal, Lehrstuhl für empirische Wirtschafts- und Sozialforschung: "Multidimensionale Skalierung", 2001, im Internet herunterladbar bei [http://www2.informatik.uni-osnabrueck.de/marc/lectures/zra\\_ss03/prgdat/mds.pdf](http://www2.informatik.uni-osnabrueck.de/marc/lectures/zra_ss03/prgdat/mds.pdf)

**Elff, Martin:** "munfold". R-Programm zu MDU nach dem Kalkül von Schoenemann zu finden im Internet unter <http://www.elff.eu/software/#munfold>

**Leeuw, Jan de :** Multidimensional Unfolding, 2004, im Internet herunterladbar bei [http://gifi.stat.ucla.edu/janspubs/2004/reports/deleeuw\\_R\\_04i.pdf](http://gifi.stat.ucla.edu/janspubs/2004/reports/deleeuw_R_04i.pdf)

**Schönemann, Peter H.:** On metric multidimensional unfaltung, Ohio State University, Research Bulletin RB-69-93, im Internet herunterladbar <http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.1969.tb00773.x/pdf>

**Schönemann, Peter H.:** On Metric Multidimensional Unfolding, 1970, *Psychometrika*, 35, 349–366

**Takane, Y., Young, F. W., DeLeeuw, J.** (1977). Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika*, 42, 7-67.

**Torgerson, Warren S.:** Theory and Method of Scaling, New York, Wiley, 1958

**Young, F. W., Y. Takane, and R. Lewyckyj.** 1978. ALSCAL: A nonmetric multidimensional scaling program with several different options, *Behavioral Research Methods and Instrumentation*, 10, 451–453.