



**Metrische multidimensionale Skalierung  
Metrische MDS**

**P34.6**

**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**

Autor: em. Prof. Dr. Kurt Holm, Universität Linz, Österreich

## 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 Mehrfach-  
antworten` (0.8 MB)
28. `Metrische multidimensionale Skalierung (MDS)` (0,4 MB)
29. `Metrisches multidimensionales Unfolding (MDU)` (0,6 MB)
30. `Nicht-metrische multidimensionale Skalierung (MDS)` (0,5 MB)
31. `Pfadanalyse als wiederholte Regressionsanalyse` (0,7 MB)
32. `Datei-Operationen mit Almo` (1,1 MB)
33. `Wählerstromanalyse und Wahlhochrechnung` (1,6 MB)
34. `Soziometrie. Auswertung soziometrischer Daten` (0,5 MB)
35. `Konfidenzintervall und p-Wert beim Bootstrap-Verfahren` (200 KB)

## Inhaltsverzeichnis

P34.6.1 Überblick, Beispiele einer metrischen MDS .....	4
P34.6.1.1 Der Begriff "Distanz" .....	7
P34.6.2 Bedingungen für die metrische MDS .....	8
P34.6.2.1 Ähnlichkeitsmatrizen .....	8
P34.6.2.2 Korrelationsmatrix .....	9
P34.6.3 Theorie und Kalkül der metrischen MDS.....	10
P34.6.4 Vergleich mit nicht-metrischer MDS .....	11
P34.6.5 Vergleich mit SPSS-ALSCAL und PROXSCAL .....	11
P34.6.5.1 Iterativer oder algebraischer Kalkül bei metrischen Daten .....	14
P34.6.6 Eingabe in Almo .....	14
P34.6.7 Erläuterungen zu den Eingabeboxen .....	17
P34.6.8 Ausgabe aus Almo-Prog30ma .....	21
P34.6.9 Prog33m3: Paarvergleich mit anschliessender MDS .....	31
P34.6.10 Paarvergleich mit Präferenzurteil .....	33
Literatur.....	34

## P34.6.1 Überblick, Beispiele einer metrischen MDS

Ein typisches Beispiel einer MDS-Studie ist folgendes:

Die Entfernungen (in km) zwischen 5 Städten sind folgende:

	Berlin	Frankfurt	Hamburg	Köln	München
Berlin	0	548	289	576	586
Frankfurt	548	0	493	195	392
Hamburg	289	493	0	427	776
Köln	576	195	427	0	577
München	586	392	776	577	0

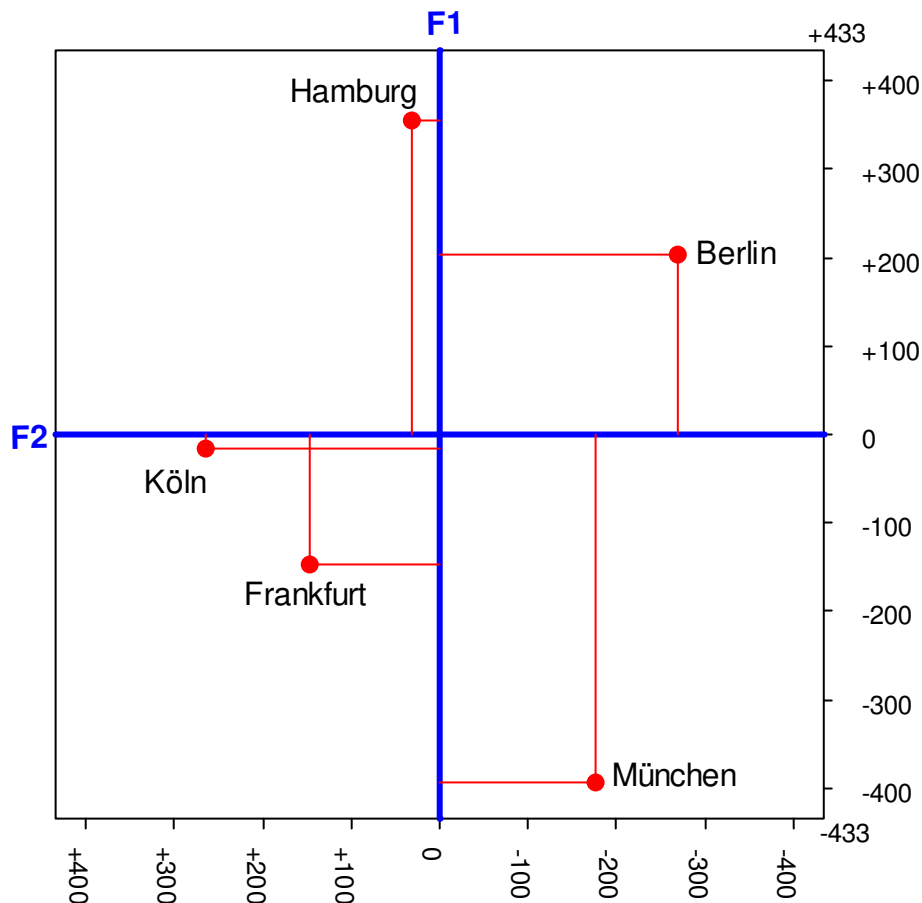
(Beispiel aus Wikipedia, Artikel "Multidimensionale Skalierung", Februar 2015)

Das Almo-Programm Prog30ma zur metrischen MDS gibt folgende "Faktorladungsmatrix" aus

		Faktor 1	Faktor 2
Berlin	V1	203.1860	-269.5053
Frankfurt	V2	-147.7360	148.0635
Hamburg	V3	353.9964	32.0311
Köln	V4	-15.7861	265.9419
München	V5	-393.6603	-176.5311

und zeichnet folgendes zweidimensionale Koordinatensystem

Faktorladungen



Das Koordinatensystem wurde um 90 Grad gegen den Uhrzeigersinn gekippt, so dass die Achse F1 nach oben zeigt. Die Grafik stimmt dann mit der Landkarte überein. Die Maßzahlen können allerdings verzerrt sein.

Wir können nun eine erste Definition der metrischen MDS geben:

*Ihre Aufgabe ist es, die Distanzen bzw. Ähnlichkeiten zwischen Objekten in einem ein- oder mehrdimensionalen Raum abzubilden.*

### Ein weiteres Beispiel: Unähnlichkeit von Automarken

Durch paarweisen Vergleich wurden die Distanzen in der Wahrnehmung verschiedener Automarken ermittelt:

	Opel	VW	Suz	Toy	Merc	BMW	Ferr	Por	Lamb	Roll
Opel	0									
VW	3.1	0								
Suzuki	5.0	4.4	0							
Toyota	3.8	3.3	3.7	0						
Mercedes	5.9	5.8	7.0	5.3	0					
BMW	5.5	5.5	7.0	4.2	2.7	0				
Ferrari	8.4	8.1	8.3	8.3	6.9	6.8	0			
Porsche	8.4	8.1	8.4	8.3	6.4	6.4	3.0	0		
Lamborghini	8.5	8.2	8.8	8.7	6.6	6.4	2.1	3.4	0	
Rolls Royce	8.5	8.6	8.9	8.2	5.8	7.0	6.6	6.8	6.3	0

(Daten aus Internetpaper "Multidimensionale Skalierung", Lehrstuhl für empirische Wirtschafts- und Sozialforschung, Fachbereich Wirtschaftswissenschaft, BUGH Wuppertal, 2001  
[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) )

Almo liefert folgendes Ergebnis:

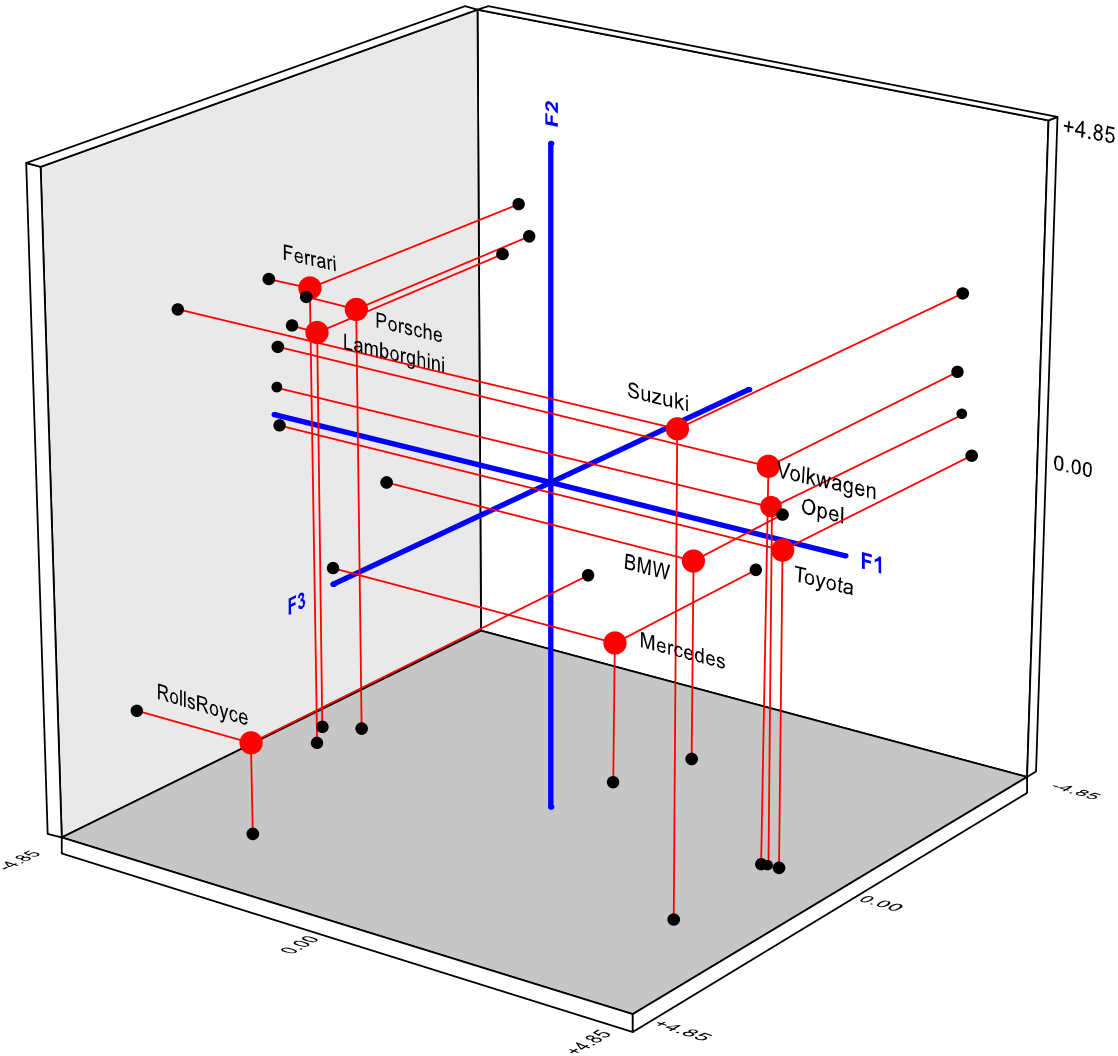
Matrix der Faktorladungen

		Faktor 1	Faktor 2	Faktor 3
Opel	V1	3.6170	0.3755	-0.0848
Volkswagen	V2	3.5083	0.9681	-0.1096
Suzuki	V3	3.5673	2.1300	2.0885
Toyota	V4	3.7712	-0.2035	-0.1224
Mercedes	V5	0.1811	-2.7555	-1.2839
BMW	V6	0.6122	-1.8120	-2.5856
Ferrari	V7	-4.1109	2.0314	0.0821
Porsche	V8	-3.9372	1.5550	-0.7375
Lamborghini	V9	-4.4048	1.2016	-0.4196
RollsRoyce	V10	-2.8042	-3.4905	3.1727

Für die 3-Faktoren-Lösung gibt Almo einen Stress von 0.11 und für die 2-Faktoren-Lösung von 0.20 aus. Der Stress-Koeffizient drückt die Güte einer Lösung aus. Je kleiner der Koeffizient umso besser die Lösung. Selbstverständlich ist dieses Gütemaß umso besser je mehr Faktoren extrahiert werden. Nach gängiger Meinung bedeutet ein Stress von 0.20, dass eine ausreichend gute Lösung gefunden wurde. Es genügt also, 2 Faktoren zu extrahieren.

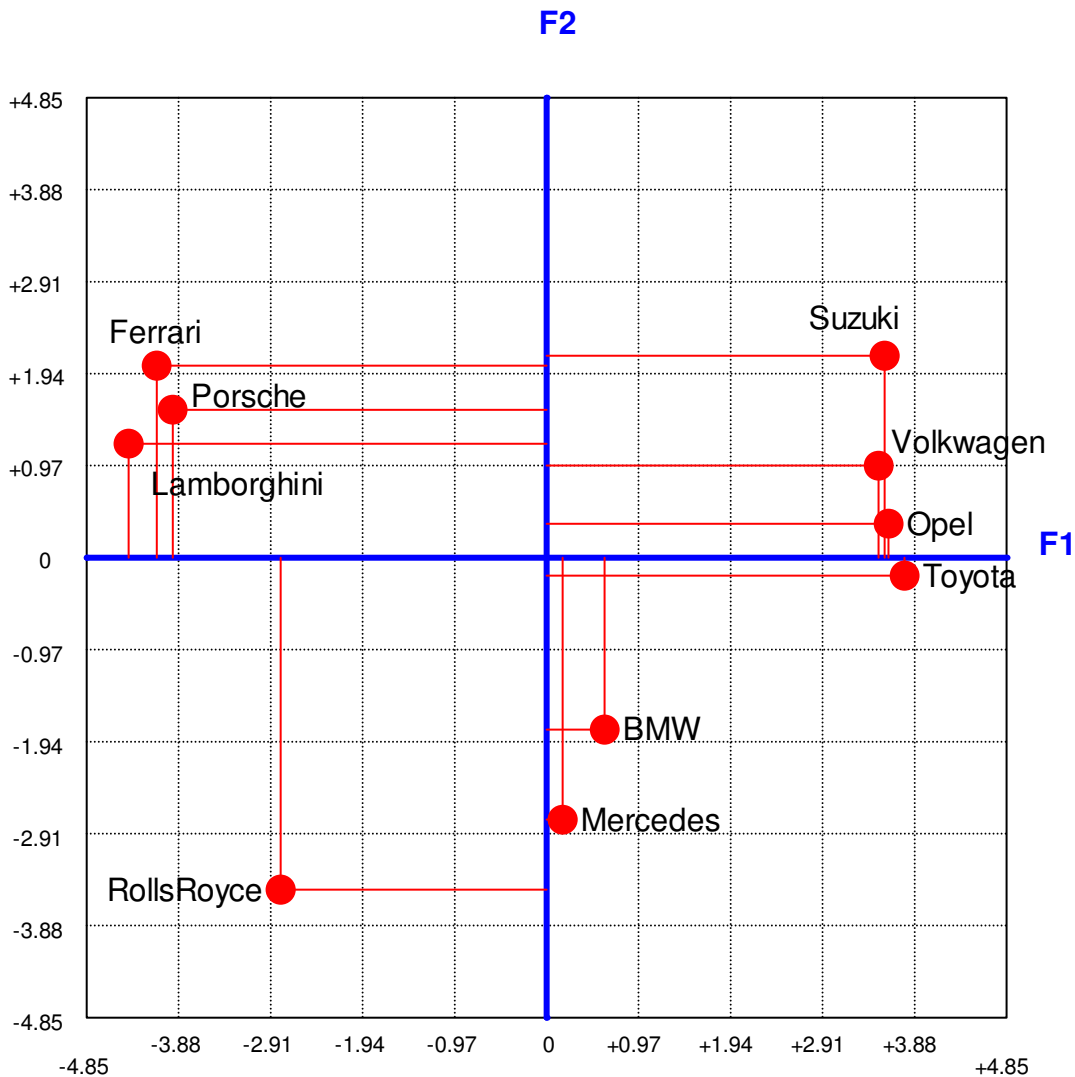
Die 3-Faktoren Lösung grafisch dargestellt:

Faktorladungen



Im Almo-Grafik-Editor geben wir nun die Grafik für die beiden ersten Faktoren aus:

## Faktorladungen



Sowohl bei der 2- als auch bei der 3-dimensionalen Darstellung wird folgendes ersichtlich:

- (1) Ferrari, Porsche, Lamborghini bilden die Gruppe der rasanten Sportwagen,
- (2) Suzuki, Volkswagen, Opel, Toyota bilden die Gruppe der "Alltagsautos"
- (3) BMW und Mercedes sind die Autos für den anspruchsvolleren Fahrer
- (4) Rolls Royce ist das edle Ausnahme-Auto

Die 2 bzw. 3 Achsen F1, F2 und F3 sind inhaltlich schwer zu interpretieren. Vielleicht ist F1 als die Achse der Sportlichkeit zu verstehen. Allerdings würde dann Rolls Royce auf der sportlichen Seite wie Ferrari usw. stehen. Das ist kaum akzeptierbar. Aber wie ist F2 zu interpretieren? Es ist dasselbe Problem wie bei der Faktorenanalyse. Die signifikanten Achsen, in unserem Beispiel F1 und F2 sind oft inhaltlich nicht interpretierbar. So dass wir bei unserer oben gegebenen Definition bleiben und mit wenigen Worten definieren :

*Die metrische MDS ist die dimensionale Darstellung einer Distanzmatrix*

### **P34.6.1.1 Der Begriff "Distanz"**

Wir werden folgende Begriffe verwenden

1. "Distanz" und synonym dazu "Unähnlichkeit"
2. "Ähnlichkeit"

Für den MDS-Kalkül müssen Ähnlichkeiten in Distanzen (bzw. Unähnlichkeiten) umgewandelt werden. Das wird noch ausgeführt. "Distanz" oder "Unähnlichkeit" ist die empirisch erhobene Entfernung zwischen den Objekten.

Bei der nicht-metrischen MDS (allgemeiner bei den iterativen MDS-Verfahren) wird hingegen streng zwischen "Unähnlichkeit" und "Distanz" unterschieden. "Distanz" ist dort die vom Modell reproduzierte Entfernung. "Unähnlichkeit" ist die empirisch erhobene Entfernung. Siehe dazu Almo-Dokument Nr 30 "Nicht-metrische multidimensionale Skalierung"

## **P34.6.2 Bedingungen für die metrische MDS**

1. Die Distanz- bzw. Ähnlichkeitsmaße müssen quantitativ sein. Sind sie ordinal, dann kann eine **nicht-metrische MDS** mit Prog34m gerechnet werden. Ob eine Maßzahl als quantitativ oder doch nur als ordinal einzuschätzen ist, entscheidet letztendlich der Forscher. In den Human- und Sozialwissenschaften ist man hier eher großzügig.
2. Die Objekte, die miteinander verglichen werden, müssen auch vergleichbar sein. Es macht kaum einen Sinn, von befragten Personen Ähnlichkeiten zwischen Automarken und Ähnlichkeiten zwischen prominenten Rennfahrern zu erfragen und aus diesen eine gemeinsame Distanz- bzw. Ähnlichkeitsmatrix zu bilden. Die Objekte einer solchen Matrix müssen vergleichbar sein. Strenger formuliert: Die einzelnen Objekte müssen Ausprägungen ein und desselben Objektklasse sein. Was *eine* Objektklasse ist entscheidet der Forscher. Ob es sinnvoll ist, den Rennfahrer X mit der Automarke Mercedes zu vergleichen, entscheidet der Forscher.

Wird allerdings eine *Korrelationsmatrix* für inkonsistenten Objekte errechnet und auf diese die metrische MDS angewendet, dann entsteht ein Ergebnis, das möglicherweise sinnvoll interpretierbar ist. Wir werden darauf weiter unten noch eingehen.

Ein Spezialfall bildet die Einführung von "Idealpunkten" beim nicht-metrischen MDS und das "multidimensionale Unfolding" (kurz: MDU), das ebenfalls in Almo enthalten ist. In den beiden Almo-Dokumenten Nr. 29 und 30 werden diese beschrieben. In beiden Verfahren werden sehr wohl verschiedene Objekte in einer Distanzmatrix zusammengebracht.

3. Die Maßeinheiten, mit denen die Objekte verglichen werden, müssen dieselben sein. In unserem Beispiel (Distanzen zwischen Städten) wäre es unzulässig, die Entfernungen für Berlin in Meilen anzugeben und für die anderen Städte in km.
4. Die metrische MDS kann auf eine Distanzmatrix, aber auch auf eine Ähnlichkeitsmatrix angewendet werden. Die Ähnlichkeitsmatrix wird in Almo programmintern in eine Distanzmatrix umgewandelt. Siehe nächsten Abschnitt.
5. Die Distanzmatrix und auch die Ähnlichkeitsmatrix muss vollständig sein. D.h. für alle Objektpaare muss ein Wert vorhanden sein. Fehlt einer oder mehrere Werte, dann muss ein Schätzwert eingesetzt werden. Das kann z.B. der Gesamtmittelwert aus dem unteren Dreieck der Matrix (ohne Diagonale) sein.
6. In der Diagonale der Distanzmatrix muss 0 stehen.

### **P34.6.2.1 Ähnlichkeitsmatrizen**

Häufig liegt nicht eine Distanzmatrix sondern eine Ähnlichkeitsmatrix vor. Für den MDS-Kalkül muss diese in eine Distanzmatrix transformiert werden. Almo macht das automatisch. Der Benutzer muss jedoch zwischen 2 Methoden der Umwandlung entscheiden.

#### *Methoden*

Almo ermittelt den maximalen Matrixwert. Das muss ein Diagonalwert sein. Dann werden alle Matrixwerte von diesem maximalen Wert subtrahiert. Die Diagonalglieder müssen dann alle =0 sein.



## Method II

Die Umformung der Ähnlichkeiten erfolgt nach folgender Formel:

$$d_{ik} = \sqrt{a_{ii} + a_{kk} - 2 * a_{ik}}$$

$a_{ik}$  = Ähnlichkeit in der Matrix-Zelle ik

$a_{ii}$  = Diagonalglied in Zeile/Spalte ii

$a_{kk}$  = Diagonalglied in Zeile/Spalte kk

$d_{ik}$  = Distanz für Matrix-Zelle ik

$2 * a_{ik}$  darf nicht größer sein als die Summe der beiden Diagonalglieder. Die Ähnlichkeitsmatrix wäre dann inkonsistent. Sie könnte nicht in eine Distanzmatrix gewandelt werden. Betrachten wir folgende Ähnlichkeitsmatrix

```
3
2 3
2 2 6
1 1 5 3
```

Das Diagonalglied in der 3. Zeile/Spalte ist 6. Das Glied in Zeile 4, Spalte 3 ist 5. Gemäß obiger Formel wäre 5 in folgendes Distanzmaß zu wandeln

$$\begin{aligned} d(4,3) &= \text{Wurzel}( a(4,4) + a(3,3) - 2*5) \\ d(4,3) &= \text{Wurzel}( 3 + 6 - 2*5) \end{aligned}$$

Es entsteht eine negative Wurzel.

In der Diagonale der Ähnlichkeitsmatrix sollte in allen Diagonalgliedern derselbe Wert stehen. Seine Größe ist beliebig. Er muss aber größer sein als der maximale Matrixwert außerhalb der Diagonale. Almo überprüft diese Bedingung und bringt bei deren Verletzung eine Meldung.

Die beiden Methoden erbringen etwas verschiedene Ergebnisse. Sie sind jedoch strukturgleich, was beim Betrachten der Punkte-Diagramme ersichtlich wird. Methode I ist eher vorzuziehen.

### P34.6.2.2 Korrelationsmatrix

Wir haben oben in Satz 2 und 3 gesagt, dass die Objekte einer MDS-Analyse vergleichbar sein müssen und in derselben Maßeinheit gemessen sein müssen. Auch wenn die Objekte in einer Korrelationsmatrix ursprünglich (vor dem Korrelieren) in verschiedenen Maßeinheiten gemessen wurden, so sind sie nun als korrelierte Objekte alle in derselben Maßeinheit gemessen, nämlich in ihrer Standardabweichung. Das ist der Vorteil der Korrelationsmatrix: Die Objekte sind nun vergleichbar, obwohl sie ursprünglich verschiedenartig waren und in verschiedenen Maßeinheiten gemessen wurden.

Siehe dazu das Programm "AutoFahr.Alm". Es wird erreicht durch Klick auf das Menü ALMO, dann Eintrag "Liste aller Almo-Programme". In diesem Programm werden dieselben Daten mit dem MDS-Verfahren analysiert, die im Masken-Programm "Prog30.m5" einer nominalen Faktorenanalyse (Korrespondenzanalyse) unterworfen werden. Dieses Programm kann aus "AutoFahr.Alm" gestartet werden. Die grafische Darstellung der Punktekonfiguration in beiden Programmen ist fast identisch.

Die Korrelationsmatrix muss in der Programm-Maske als Ähnlichkeitmatrix deklariert werden. Sie wird von Almo automatisch in eine Distanzmatrix umgedreht. Dabei sollte der Benutzer als Umwandlungsmethode eher die *Method I* einsetzen. Bei *Method II* wird gemäß obiger Umwandlungsformel

aus r= 1	der Distanzwert d= 0
aus r= 0.5	d= 1
aus r= 0	d= 1.414
aus r=-0.5	d= 1.732
aus r=-1	d= 2

Eine Distanz von d=0 entsteht bei der perfekten Korrelation r=1. Die maximale Distanz von d=2 entsteht bei der perfekten gegenläufigen Korrelation von r=-1. Die Mitte, also eine Distanz von d=1 entsteht aber nicht bei r=0 sondern bei r=0.5.

Zu beachten ist, dass eine normale Faktorisierung einer Korrelationsmatrix ein anderes Ergebnis liefert als die metrische MDS. In der Regel wird man jedoch beim Betrachten der Diagramme feststellen, dass die Punktekonfigurationen ähnlich sind.

### P34.6.3 Theorie und Kalkül der metrischen MDS

Die metrische MDS wurde wesentlich von W.S. Torgerson schon 1958 entwickelt. Die zu vergleichenden Objekte werden begriffen als Punkte in einem euklidischen Raum. Die Distanzmatrix **D** enthält also die euklidischen Distanzen zwischen diesen Punkten. Die Distanzmatrix gilt es nun zu einer Matrix **B\*** zu transformieren, deren einzelne Glieder skalare Produkte der Objektabstände sind. Nur dann ist die Matrix faktorierbar, d.h. dimensional zerlegbar. Der Ursprung des dabei entstehenden Koordinatensystems wird in den Centroid (geometrischer Mittelpunkt) aller Punkte gelegt. Zuvor wird **D** quadriert.

Nachfolgende Formel erzeugt aus der Distanzmatrix **D** die (wie wir es nennen)

quadrierte und zentrierte Distanzmatrix **B\***  
d.h. die faktorierbare MDS-Distanzmatrix **B\***.

Siehe W.S. Torgerson, 1958, S.258, Gleichung 16.

Ein einzelnes Glied  $b_{jk}^*$  wird gemäß folgender Formel gebildet:

$$b_{jk}^* = \frac{1}{2} \left( \frac{1}{n} \sum_j^n d_{jk}^2 + \frac{1}{n} \sum_k^n d_{jk}^2 - \frac{1}{n^2} \sum_j^n \sum_k^n d_{jk}^2 - d_{jk}^2 \right)$$

j = das ist der Zeilenindex  
k = das ist der Spaltenindex  
n = Zahl der Objekte

$d_{jk}^2$  = die Distanzmatrix **D** wird gliedweise quadriert. Dies ist das Glied jk aus dieser Matrix

$\frac{1}{n} \sum_j^n d_{jk}^2$  = das ist die Zeilensumme der Zeile j der gliedweise quadrierten Distanzmatrix **D** - dividiert durch die Zahl der Objekte (=der Zeilenmittelwert)

$\frac{1}{n} \sum_k^n d_{jk}^2$  = das ist die Spaltensumme der Spalte k der gliedweise quadrierten Distanzmatrix **D** - dividiert durch die Zahl der Objekte (=der Spaltenmittelwert)

$\frac{1}{n^2} \sum_j^n \sum_k^n d_{jk}^2$  = das ist die Gesamtsumme der gliedweise quadrierten Distanzmatrix **D** – dividiert durch das Quadrat der Objektzahl (=der Gesamtmittelwert)

Der weitere Kalkül verläuft dann wie bei der Faktorenanalyse. Mit einem Eigenwert-Eigenvektor-Verfahren wird die Matrix **B\*** faktoriell zerlegt.

$$\begin{aligned}
\mathbf{B}^* &= \mathbf{V} \cdot \mathbf{W} \cdot \mathbf{V}' \\
&= \mathbf{V} \cdot \text{sqrt}(\mathbf{W}) * \text{sqrt}(\mathbf{W}) \cdot \mathbf{V}' \\
&= \mathbf{L} * \mathbf{L}'
\end{aligned}$$

$\mathbf{B}^*$  = quadrierte und zentrierte Distanzmatrix  
 $\mathbf{V}$  = Eigenvektoren  
 $\mathbf{V}'$  = Transponierte von  $\mathbf{V}$   
 $\mathbf{W}$  = Diagonalmatrix der Eigenwerte  
 $\text{sqrt}(\mathbf{W})$  = Diagonalmatrix der Wurzeln aus den Eigenwerten  
 $\mathbf{L}$  = Ladungsmatrix  
 $\mathbf{L}'$  = Transponierte von  $\mathbf{L}$

Die metrische MDS ist kurz definiert:

*Die metrische MDS ist eine Faktorenanalyse der quadriert und zentrierten Distanzmatrix*

Tatsächlich wird in Almo auch das Faktorenanalyse-Programm mit einigen Einschüben verwendet. Siehe dazu die ausführliche Darstellung im Almo Dokument Nr.15 "Faktorenanalyse".

### **P34.6.4 Vergleich mit nicht-metrischer MDS**

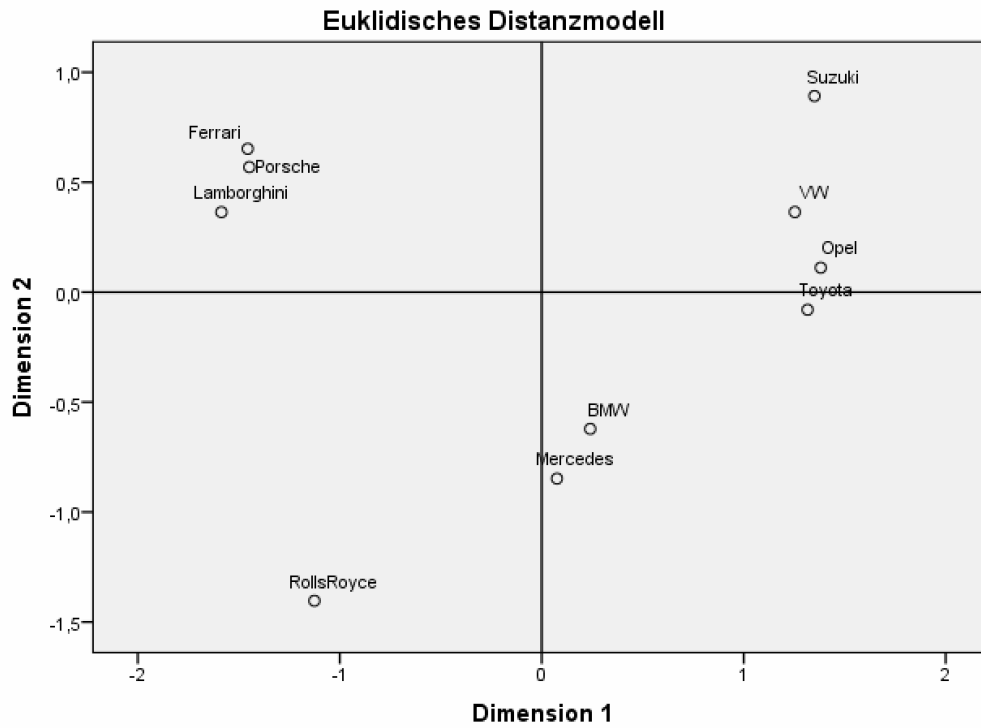
Der Vorteil der nicht-metrischen MDS (nMDS) ist es, dass sie mit ordinalen Distanzdaten (bzw. Ähnlichkeitsdaten) gerechnet werden kann. Die Ergebnisse, die die beiden Methoden erbringen, sind ähnlich. In Almo ist die nicht-metrische MDS in den Programm-Masken Prog34ma bis Prog34mc realisiert. Die Beschreibung dazu ist im Almo-Dokument Nr. 30 "Nicht-metrische multidimensionale Skalierung" enthalten. Almo verwendet den Kalkül nach Kruskal (1964). Inzwischen gibt es mehrere Varianten und Alternativen zum Kruskal-Verfahren, z.B die beiden SPSS-Programme ALSCAL und Proxscal. Diese ermöglichen es auch den speziellen iterativen Kalkül der nMDS auf metrische Distanzen anzupassen, so dass bei diesen kein Informationsverlust entsteht. Die Ergebnisse sind dann denen der metrischen MDS sehr ähnlich. Siehe dazu nachfolgend P34.6.5.1.

### **P34.6.5 Vergleich mit SPSS-ALSCAL und PROXSCAL**

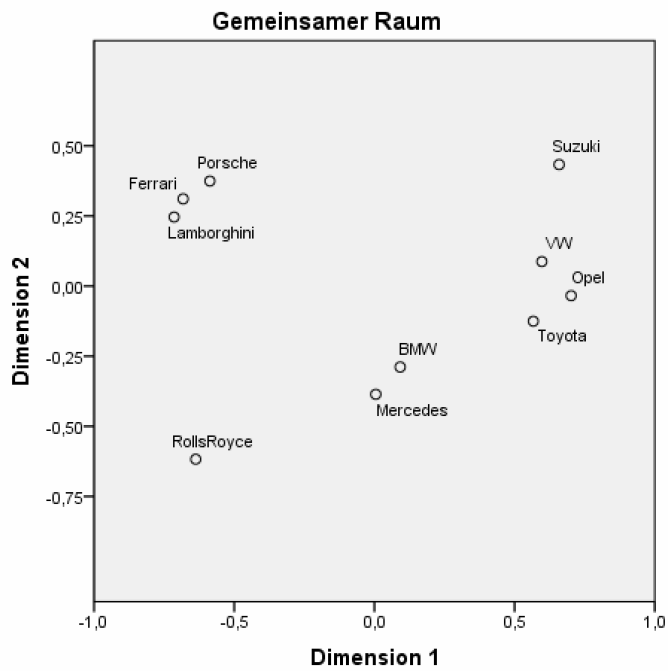
Mit ALSCAL und PROXSCAL kann eine metrische und eine ordinale MDS gerechnet werden. Beide sind iterative Verfahren. Sie verwenden nicht den klassischen ("faktorenanalytischen") MDS-Kalkül nach Torgerson. Der Kalkül von Alscal nach Takane, Young & DeLeeuw (1977) ist dokumentiert in "IBM SPSS Statistics xx Algorithms. Alscal, Proxscal und Almo erbringen sehr ähnliche, aber nicht identische Ergebnisse.

Die Ergebnisse der 3 Verfahren in Form der grafischen Darstellung als 2-dimensionale Punkte-Diagramme, sind für alle Beispiele, die wir gerechnet haben, fast identisch. Optisch lässt sich fast kein Unterschied zwischen den Diagrammen feststellen. Man vergleiche die nachfolgenden Diagramme.

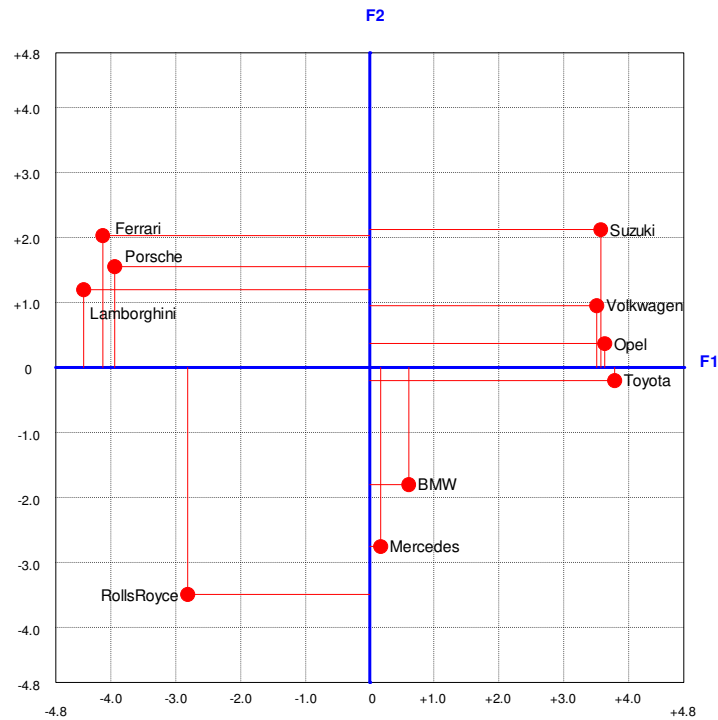
#### **Diagramm aus Alscal**



**Diagramm aus Proxscal**



**Diagramm aus Almo**



Die beiden SPSS-Verfahren können nachgerechnet werden mit folgender Distanzmatrix *Automarken\_mat.sav* oder der "umgedrehten" Ähnlichkeitsmatrix *Automarken\_Aehnl.sav*. Sie sind zu finden im Almo-Unterordner TESTDAT.

Die Punkte-Konfiguration bei Proxscal ist geringfügig anders als die bei Almo und Alscal.

Die numerischen Faktorladungen aus Almo, Alscal und Proxscal sind ungefähr proportional zueinander - jedoch nicht exakt proportional.

	ALMO		SPSS-ALSCAL		SPSS-Proxscal	
	Faktor 1	Faktor 2	Faktor 1	Faktor 2	Faktor 1	Faktor 2
Opel	3.6170	0.3755	1,3811	,1113	.702	-.034
Volkswagen	3.5083	0.9681	1,2531	,3646	.597	.087
Suzuki	3.5673	2.1300	1,3501	,8920	.659	.433
Toyota	3.7712	-0.2035	1,3163	-,0797	.567	-.125
Mercedes	0.1811	-2.7555	,0753	-,8478	.005	-.386
BMW	0.6122	-1.8120	,2400	-,6219	.092	-.289
Ferrari	-4.1109	2.0314	-1,4558	,6523	-.682	.311
Porsche	-3.9372	1.5550	-1,4482	,5688	-.587	.374
Lamborgh	-4.4048	1.2016	-1,5860	,3639	-.715	.246
RollsRoy	-2.8042	-3.4905	-1,1258	-1,4035	-.638	-.617
Stress-1	0.23		0.1		0.09	

Anmerkung: Wird bei Proxscal als Anfangs-Konfiguration *Zufallsstart(1)* angegeben, dann wird keine Lösung gefunden - bei mehreren *Zufallsstart*-Versuchen jedoch immer. Wird hingegen *Torgerson* oder *Simplex* als Anfangs-Verfahren eingesetzt, dann wird die Lösung sofort gefunden. Das weist darauf hin, dass das iterative Suchverfahren alleine, ohne gute Anfangswerte versagen kann. Das ist bei iterativen MDS-Verfahren die Regel. Wird mit Zufallswerten als Startwerten gerechnet, dann sollten immer mehrere Versuche vorgesehen werden. Bei Proxscal wird das auch explizit angeboten.

### ***P34.6.5.1 Iterativer oder algebraischer Kalkül bei metrischen Daten***

Der Stress-Koeffizient ist bei Alscal und Proxscal höher als bei Almo. Das ist häufig so. Wir wollen untersuchen, warum iterative Verfahren häufig bessere Stress-Koeffizienten liefern als algebraische. Wir beschäftigen uns dabei nur mit der metrischen nicht mit der ordinalen MDS. Für diese gibt es nur die iterativen Verfahren. Die iterativen Verfahren wurden auch ursprünglich ausdrücklich für ordinale Distanzmatrizen entwickelt und sind erst danach auf metrische ausgedehnt worden.

Bei den iterativen MDS-Verfahren wird zunächst die Zahl der Dimensionen festgelegt, z.B. 2. Wir wollen nun in einem Gedankenexperiment annehmen, dass der euklidische Raum in dem die Objekte aufgespannt werden tatsächlich 3-dimensional ist. D.h. die Distanzmatrix ist perfekt 3-dimensional.

Die iterativen Verfahren benötigen eine Anfangs-Konfiguration der Objekte im Raum. Wird bei Proxscal beispielsweise *Torgerson* oder *Simplex* als Anfangs-Konfiguration gewählt, dann wird eine erste Lösung errechnet.

In unserem Beispiel gibt Proxscal bei *Torgerson* dafür einen "normalisierten r Rohstress" von 0.241 aus. Das ist ein anderer Stress-Koeffizient als der von Almo angegebene Stress-1-Koeffizient von 0.2. Die beiden Stress-Koeffizienten sind nicht vergleichbar. Proxscal gibt die Ladungsmatrix dieser ersten Lösung nicht aus.

Ist nun der gefundene Stress-Koeffizient nicht befriedigend, dann wird dies im algebraischen metrischen MDS so interpretiert, dass 2 Dimensionen nicht genügen um die metrischen Distanzen zwischen den Objektpunkten korrekt im euklidischen Raum abzubilden. Es muss eine 3. Dimensionen hinzugenommen werden. In unserem Automarken-Beispiel verbessert sich dadurch der Stress-1 auf 0.15.

Die iterativen MDS-Verfahren hingegen versuchen nun, grafisch betrachtet, im 2-dimensionalen Raum durch Hin- und Herschieben der Objektpunkte den Stresswert zu verbessern. Dabei werden manche Distanzen kürzer, andere eventuell länger, jedoch so dass insgesamt eine Stress-Verbesserung entsteht. Zu beachten ist: Das Hin- und Herschieben erfolgt im 2-dimensionalen Sub-Raum des tatsächlich 3-dimensionalen Raums.

Das Hin- und Herschieben erfolgt nach dem von Kruskal eingeführten "Gradienten-Verfahren" (bzw. dessen Modifikationen). Dabei geht es darum bei jeder Iteration eine etwas veränderte Punkte-Konfiguration zu finden, die den Stress-Koeffizienten verbessert. Die Stress-Koeffizienten sind bei den Verfahren verschieden - mit der Folge dass sie (geringe) unterschiedliche Ergebnisse erbringen. Das Gradienten-Verfahren nach Kruskal wird im Almo-Dokument 30 "nicht-metrische MDS", Abschnitt P30.0.2 beschrieben.

Wenn nun jedoch tatsächlich, wie wir in unserem Gedankenexperiment unterstellen, eine bedeutsame 3. Dimension existiert, d.h. relevante Ladungen für den 3. Faktor gegeben sind, dann ist die Darstellung der Punkte im 2-dimensionalen Sub-Raum falsch.

Die iterativen Programme bieten dem Benutzer an, Lösungen für eine aufsteigende Zahl von Dimensionen zu rechnen. Aber es gilt: Da nur noch 3 Dimensionen grafisch gezeichnet werden können, werden in der Regel maximal nur 3 Dimensionen gerechnet.

Siehe Exkurs: Vergleich metrische MDS und iterative MDS in Almo-Dokumen Nr.30 "Nicht-metrische MDS".

## **P34.6.6 Eingabe in Almo**

Almo enthält 2 Programm-Masken zur Eingabe und Analyse einer Distanz- bzw. Ähnlichkeitsmatrix. Bei Prog30ma wird die Matrix aus einer Datei eingelesen. Bei Prog30mc schreibt der Benutzer die Matrixdaten hinter das Programm-Ende.

Man findet beide Programme durch Klick auf den Knopf "Verfahren" (unterhalb der Menüleiste), dann "MDS / MDU".

Prog30ma.Msk  
**Metrische multidimensionale Skalierung (metrische MDS)  
 Faktorisierung einer Distanzmatrix  
 mit eingelesener fertiger Distanzmatrix aus Datei**

Beispiel: Die Distanzen (in km) zwischen 5 Städten sind folgende:

	Berlin	Frankfurt	Hamburg	Köln	München
Berlin	0	548	289	576	586
Frankfur	548	0	493	195	392
Hamburg	289	493	0	427	776
Köln	576	195	427	0	577
München	586	392	776	577	0

Das Ergebnis ist eine "Faktorladungsmatrix", deren 2-dimensionale grafische Darstellung der Lage der 5 Städte auf der Landkarte entspricht.

Beispiel aus Wikipedia, Artikel "Multidimensionale Skalierung", Februar 2015

---

weitere laden durch  
 Beispiel-Programme: Doppelklick

|

Unähnlichkeit polit. Parteien	[".\Almo_BSP\PolitPartei.Alm"]
Unähnlichkeit von Automarken	[".\Almo_BSP\Automarken.Alm"]
Korrelation von Verbrechen	[".\Almo_BSP\Verbrechen.Alm"]
Ähnlichkeitsmatrix	[".\Almo_BSP\Aehnlich.Alm"]

Programm-Bedienung ---> Hilfe

Speicher fuer x Variable

Geben Sie soviele an, wie die zu faktorisierende Matrix  
 Zeilen bzw. Spalten besitzt

Vereinbare Variable= 20

Option: Weitere Vereinbarungen - nur wenn Almo dazu auffordert

Variablennamen Hilfe

Datei der Variablennamen

↔ ↵ █

Freie Namensfelder

↔ Leere alle Eingabefelder dieser Sub-Box

↔ Name1=Berlin;  
 ↔ Name2=Frankfurt;  
 ↔ Name3=Hamburg;  
 ↔ Name4=Köln;  
 ↔ Name5=München;  
 ↔ █

[...] [erzeuge zusätzliche Namensfelder](#)

Zahl der Zeilen bzw. Spalten der Matrix

↔ ↕ ↴ 5

Matrix aus Datei einlesen Hilfe

↔ ↵ "C:\Almo15\TESTDAT\StaedteEntferng.mat"

---

↕ ↴ 0 **Matrixtyp** Hilfe

0= Distanzmatrix / Unähnlichkeits-Matrix  
 1= Ähnlichkeits-Matrix. Umdrehen zu Distanzmatrix Methode I  
 2= Ähnlichkeits-Matrix. Umdrehen zu Distanzmatrix Methode II

↵ Option: Faktoren / reproduzierte Matrix

↵ Option: Rotation

↵ Option: weitere Optionen

↵ Option: "Aussehen" der auszugebenden Tabelle bzw. Matrix

↵ Grafik-Optionen

Programmende



## P34.6.7 Erläuterungen zu den Eingabeboxen

### **Eingabe-Box: Speicher für x Variable**

### **Eingabe-Box: Weitere Vereinbarungen**

Zu diesen beiden Boxen siehe Almo-Dokument 0 "Arbeiten mit Almo", Abschnitt P0.1 und P0.2

### **Eingabe-Box: Variablen Namen geben**

Sie müssen keine Variablennamen verwenden. Für die in der Eingabe-Box "Speicher für x Variable" vereinbarten Variablen erzeugt Almo automatisch die Variablennamen V1, V2, V3, etc. In diesem Falle leeren sie die Box, jeweils durch Klick auf den doppelköpfigen Pfeil (den "Raus-Rein-Knopf").

Von einem "Variablennamen" (im eigentlichen Sinne des Wortes) sprechen wir in Almo, wenn der Benutzer extra einen Namen definiert, z.B. so:

```
Name 5 =München;
```

Dies ist dann der Variablenname von V5.

Beachte:

Ein Name darf kein Blank enthalten.

Eine Namensgebung muss mit einem Semikolon abgeschlossen werden.

Siehe die ausführliche Beschreibung der Namensgebung im Almo-Dokument 0 "Arbeiten mit Almo", Abschnitt P0.3

### **Matrix aus Datei einlesen oder "Benutzer schreibt Matrix hinter Programm-Ende"**

#### **Eingabe-Box: Matrix aus Datei einlesen**

Bei der oben abgebildeten Programm-Maske Prog30ma wird die Distanz- bzw. Ähnlichkeitsmatrix aus einer Datei eingelesen.

Klicken Sie auf den 2. Knopf vor dem Eingabefeld. Almo präsentiert die Datei-Auswahl-Box. Selektieren Sie die Datei, die die Matrix enthält. Selbstverständlich muss die Matrix zuvor von Ihnen in diese Datei eingegeben worden sein.

Anmerkung: Die Matrix muß die nachfolgend dargestellte Form besitzen! Bei der metrischer MDS mit Prog30ma und auch bei verschiedenen anderen Programmen ist es empfehlenswert, die in einer Datei befindliche Matrix zu laden und zu überprüfen.

#### **Eingabe-Box in Prog30mc : Benutzer schreibt Matrix hinter Programm-Ende**

Bei der hier nicht abgebildeten Programm-Maske Prog30mc wird es dem Benutzer erlaubt, die Matrix am Programm-Ende (hinter der Programm-Maske) selbst zu schreiben

Klicken Sie auf den 2. Knopf. Es wird dann in das Eingabefeld das Wort "Eingabe" eingesetzt. Almo erwartet dann die Daten der Matrix am Programm-Ende. Schalten Sie am Programm-Ende in der Box "Schreiben der Matrixwerte" die Schreibsperre aus und schreiben Sie die Matrix-Daten hinter dem Programm.

Die Matrix-Daten müssen in folgender Form geschrieben werden. Siehe dazu auch Handbuch, Teil 2, Abschnitt 43

```
*                Stern
*                Stern
*                Stern
1.0
0.4 1.0
0.3 0.7 1.0      Dreiecksmatrix mit Diagonale
```

```

0.1 0.2 0.7 1.0
0.2 0.3 0.2 0.3 1.0
0.3 0.2 0.2 0.1 0.8 1.0
*                               Stern
*                               Stern

```

Die Eingabe besteht also aus dem unteren Dreieck mit Diagonale der Matrix und 3 Sternen vor und 2 Sternen hinter der Matrix. Die Sterne sind bloße Platzhalter für Informationen, die in anderen Almo-Verfahren, aber nicht hier bei der MDS benötigt werden.

### Matrixtyp

Im Unteren Teil der Eingabebox müssen Sie folgende Angabe machen

```

0 = Distanzmatrix bzw. Unähnlichkeitsmatrix
1 = Ähnlichkeits-Matrix. Umdrehen zu Distanzmatrix Methode I
2 = Ähnlichkeits-Matrix. Umdrehen zu Distanzmatrix Methode II

```

Die Ähnlichkeitsmatrix kann auch eine Korrelationsmatrix sein  
Die Matrizen müssen symmetrisch sein

Die Ähnlichkeitsmatrix wird von Almo programmintern  
in eine Distanzmatrix umgewandelt - nach folgenden Methoden

#### Methode I

Almo ermittelt den maximalen Matrixwert. Das muss ein Diagonalwert sein  
Dann werden alle Matrixwerte von diesem maximalen Wert subtrahiert.  
Die Diagonalglieder müssen dann alle =0 sein

#### Methode II

Die Umformung der Ähnlichkeiten erfolgt nach folgender Formel:

$$d_{ik} = \sqrt{a_{ii} + a_{kk} - 2 * a_{ik}}$$

$a_{ik}$  = Ähnlichkeit in der Matrix-Zelle ik

$a_{ii}$  = Diagonalglied in Zeile/Spalte ii

$a_{kk}$  = Diagonalglied in Zeile/Spalte kk

$d_{ik}$  = Distanz für Matrix-Zelle ik

$2 * a_{ik}$  darf nicht größer sein als die Summe der beiden Diagonalglieder. Die Ähnlichkeitsmatrix wäre dann inkonsistent.

Bei einer **Korrelationsmatrix** als Ähnlichkeitsmatrix wird bei Methode II gemäß obiger Umwandlungsformel

```

aus r= 1   der Distanzwert 0
aus r= 0.5   1
aus r= 0     1.414
aus r=-0.5   1.732
aus r=-1     2

```

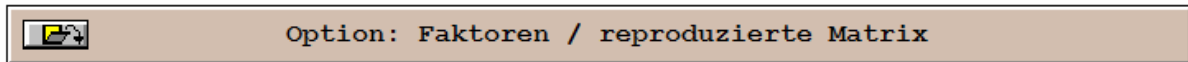
Eher vorzuziehen ist hier Methode I. Dabei wird

```

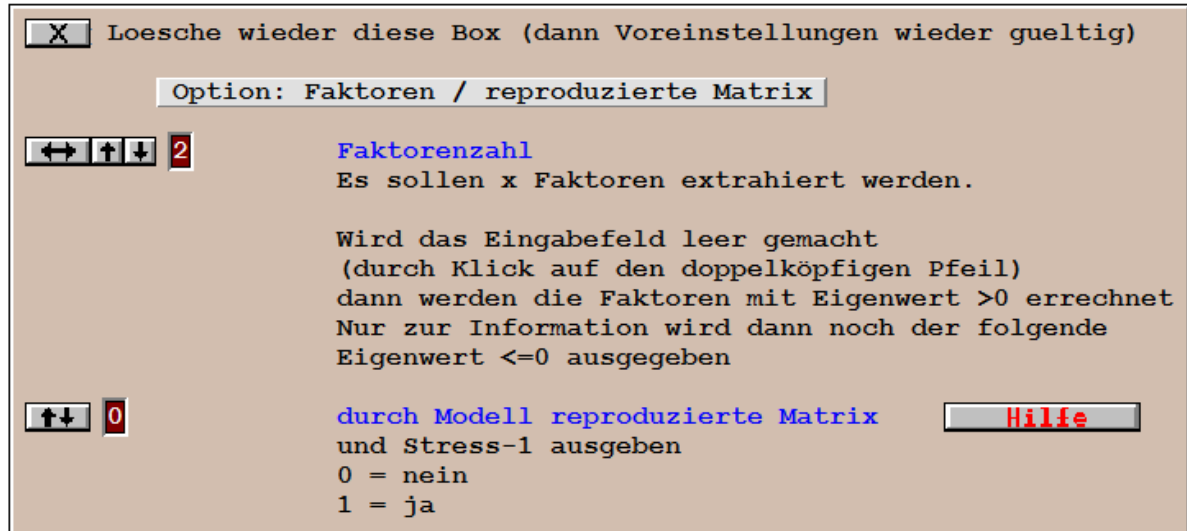
aus r= 1   der Distanzwert 0
aus r= 0.5   0.5
aus r= 0     1
aus r=-0.5   1.5
aus r=-1     2

```

## Optionsbox: Faktoren / reproduzierte Matrix



Wird die Optionsbox geöffnet, dann ist folgendes zu sehen:

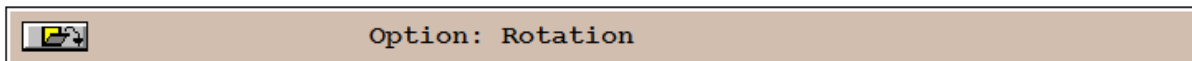


1. *Eingabefeld*: Der Benutzer kann die Zahl der Faktoren Almo überlassen. Das Eingabefeld bleibt dann leer. Klick auf den "Raus-Rein-Knopf". In diesem Falle extrahiert Almo soviele Faktoren wie Variable (Objekte) vorhanden sind. Treten negative Eigenwerte auf, dann bricht Almo das Extrahieren von Faktoren ab.

Der Benutzer kann aber auch die Zahl der Faktoren fixieren. Durch Klick auf den nach oben oder unten weisenden Pfeilknopf legt er eine Zahl fest.

2. *Eingabefeld*: Almo bildet aus der ermittelten unrotierten Faktorladung für den Faktor 1, dann für die Faktoren 1+2, dann für 1+2+3, dann für 1+2+3+4 etc. die vom Modell reproduzierte Distanzmatrix und den Stress1-Koeffizienten. Wir werden das ausführlicher erläutern, wenn wir die Ergebnis-Ausgabe kommentieren.

## Optionsbox: Rotation



Almo liefert, wie wir bei den obigen Beispielen gezeigt haben, eine 2- oder 3-dimensionale Grafik. Die Lage der Koordinatenachsen kann im Grafik-Editor

1. beliebig **gedreht** werden. Das geschieht durch Klick auf den Knopf "Kippen" in der linken Werkzeugleiste.

2. Die Punkte können auch um die Achsen **gespiegelt** werden. In der rechten Werkzeugleiste befinden sich 2 bzw. 3 Checkboxes mit der Beschriftung *F1 spiegeln*, *F2 spiegeln* usw. Dabei heisst "F1 spiegeln" dass das Vorzeichen der Ladungen auf dem Faktor F1 umgedreht wird; entsprechend auch bei F2 und F3.

Rotation bedeutet, dass die Koordinatenachse gedreht werden, wobei die Punkte aber stehen bleiben und nicht mitgedreht werden. Die Koordinatenachsen bleiben (bei der rechtwinkligen Varimax-Rotation) dabei rechtwinklig aufeinander stehen - oder der Winkel zwischen ihnen wird verändert (bei der schiefwinkligen Rotation).

Sinn der Rotation ist es eine Position zu finden, bei der die Achsen inhaltlich interpretierbar sind. Eine

Rotation ist in der Regel nur sinnvoll, wenn "Punktewolken" (Cluster) vorhanden sind, die deutlich voneinander getrennt sind und in sich gut geschlossen sind.

Die vorhandenen rechtwinkligen oder schiefwinkligen Rotationskalküle verwenden bestimmte Prinzipien, die eher bei der klassischen Faktorenanalyse als bei der MDS angebracht sind. Lesen Sie die ausführliche Darstellung im Almo-Dokument 15 "Faktorenanalyse", Abschnitt P30.1.9.

Wenn Sie aber doch rotieren wollen, dann öffnen Sie die Optionsbox. Sie sehen dann folgendes:

**Loesche wieder diese Box (dann Voreinstellungen wieder gueltig)**

**Rotation**

**Zahl der zu rotierenden Faktoren**  
keine Angabe = die extrahierten Faktoren werden rotiert  
=0 nicht rotieren

Beispiele fuer weitere mögliche Angaben:  
=2 Rotation für 2 Faktoren  
=2:4 Rotation für die ersten 2 Faktoren dann für die ersten 3 Faktoren usw. bis zu den ersten 4 Faktoren

**Rotationstyp**  
= schiefwinklig; Quartimin-Rotation mit Verbesserung durch die Gruppen-Rotation (Voreinstellung)  
= rechtwinklig; Varimax-Rotation

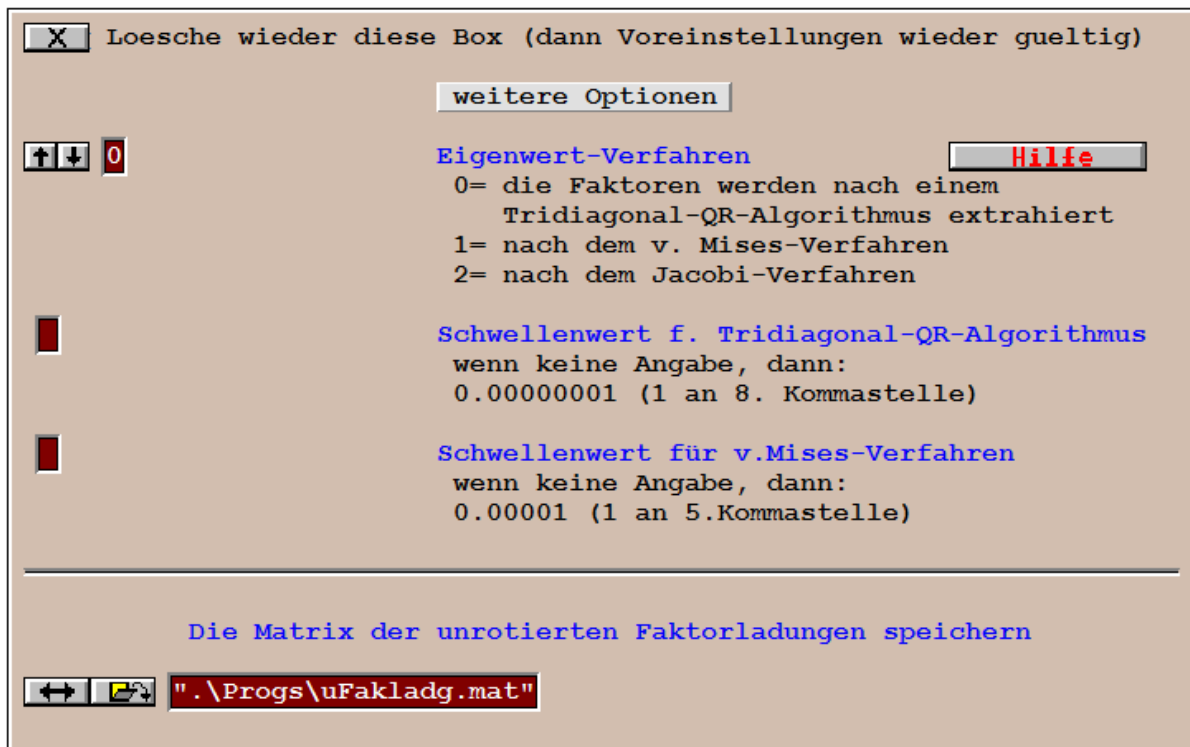
**Rotationslösung bei schiefwinkliger Rotation**  
=1 die Ergebnisse der durch die "Gruppen-Rotation" verbesserten Quartimin-Rotation werden ausgegeben (Voreinstellung)  
=2 die reine Quartimin-Lösung wird zusätzlich ausgegeben  
=3 nur Quartimin-Lösung rechnen und ausgeben

Im Almo-Dokument Nr. 15 "Faktorenanalyse" wird diese Eingabebox und der Vorgang der Rotation ausführlich dargestellt. Siehe dort die Abschnitte P30.1.4, P30.2.3 Seite 29 und P30.3.6.

### Optionsbox: Weitere Optionen

**Option: weitere Optionen**

Wenn Sie die Box öffnen, dann sehen Sie folgendes



1. Eingabefeld: Der Benutzer kann zwischen 3 Methoden der Faktorisierung der Matrix wählen.

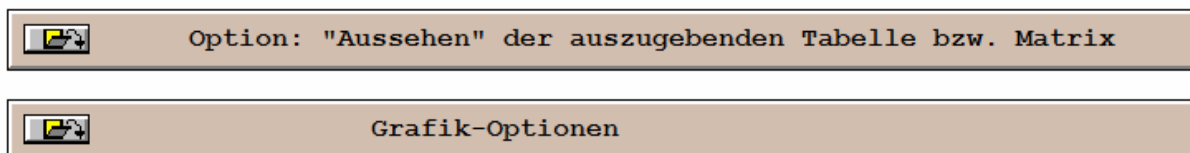
Die 3 Eigenwert-Kalküle erbringen identische Ergebnisse. In sehr seltenen Fällen kann es Abweichungen an weit hinten liegenden Kommastellen geben. Dabei ist es dann (noch seltener) möglich, dass ein Kalkül einen weiteren validen Eigenwert ausgibt.

Das schnellste Verfahren ist das Tridiagonal-QR-Verfahren. Der v. Mises-Kalkül ist der langsamste. Je nach gewähltem Eigenwert-Kalkül kann eine spaltenweise Vorzeichen-Umkehr der Eigenvektoren stattfinden. Dieser entspricht grafisch eine Spiegelung der Punkte um die Achse - und ist somit für die in Almo angebotenen Verfahren bedeutungslos.

2. und 3. Eingabefeld: Hier kann die Genauigkeit der Lösung erhöht werden. Eine Verringerung ist natürlich auch möglich. Ist aber nicht zu empfehlen.

4. Eingabefeld: Die errechnete Faktorladungsmatrix kann in eine Datei gespeichert werden.

#### Optionsbox "Aussehen" der auszugebenden Tabelle bzw. Matrix Optionsbox Grafik-Optionen



Siehe zu diesen beiden Optionsboxen das Almo-Dokument 0 "Arbeiten mit Almo", Abschnitt P0.7.3 und P0.7.4.

### P34.6.8 Ausgabe aus Almo-Prog30ma

Wir erläutern die Ausgabe am oben bereits angesprochenem Beispiel der Wahrnehmung verschiedener Automarken.

Almo gibt im 1. Teil der Ergebnisliste die eingegebene Distanzmatrix aus. Danach folgt das eigentliche Ergebnis

Zu faktorisierende MDS-Matrix  
(quadrierte und zentrierte Distanzmatrix)

		Opel	Volkwag	Suzuki	Toyota	Mercede
		V1	V2	V3	V4	V5
Opel	V1	19.92	13.94	10.04	10.99	-1.54
Volkwage	V2	13.94	17.57	11.68	11.58	-2.13
Suzuki	V3	10.04	11.68	25.16	13.98	-6.01
Toyota	V4	10.99	11.58	13.98	16.49	0.11
Mercedes	V5	-1.54	-2.13	-6.01	0.11	11.81
BMW	V6	0.39	-0.79	-6.37	4.98	7.81
Ferrari	V7	-14.11	-12.81	-10.65	-14.99	-6.69
Porsche	V8	-14.13	-12.83	-11.51	-15.01	-3.39
Lamborgh	V9	-14.55	-13.22	-14.52	-17.98	-4.26
RollsRoy	V10	-10.95	-12.98	-11.81	-10.16	4.30

		BMW	Ferrari	Porsche	Lamborg	RollsRo
		V6	V7	V8	V9	V10
Opel	V1	0.39	-14.11	-14.13	-14.55	-10.95
Volkwage	V2	-0.79	-12.81	-12.83	-13.22	-12.98
Suzuki	V3	-6.37	-10.65	-11.51	-14.52	-11.81
Toyota	V4	4.98	-14.99	-15.01	-17.98	-10.16
Mercedes	V5	7.81	-6.69	-3.39	-4.26	4.30
BMW	V6	11.11	-6.35	-3.74	-3.31	-3.73
Ferrari	V7	-6.35	22.43	17.90	20.62	4.65
Porsche	V8	-3.74	17.90	22.38	17.03	3.28
Lamborgh	V9	-3.31	20.62	17.03	23.23	6.98
RollsRoy	V10	-3.73	4.65	3.28	6.98	30.43

Wir haben oben die von Torgerson entwickelte Formel dargestellt, nach der die Distanzmatrix **D** in die *faktorisierbare MDS-Distanzmatrix B\** transformiert wird. Was wir als "zu faktorisierende MDS-Matrix" bezeichnen ist diese Matrix **B\***.

Ergebnisse aus Faktorenanalyse  
Metrische multidimensionale Skalierung (metrische MDS)

-----  
Verwendetes Eigenwert-, Eigenvektor-Verfahren: Tridiagonal-Qr-Verfahren

Schwellenwert fuer Eigenwert-Kalkuel= 1e-008

Eigenwerte der MDS-Matrix je Faktor  
112.41      36.70      23.52

Normierte Eigenvektoren je Faktor

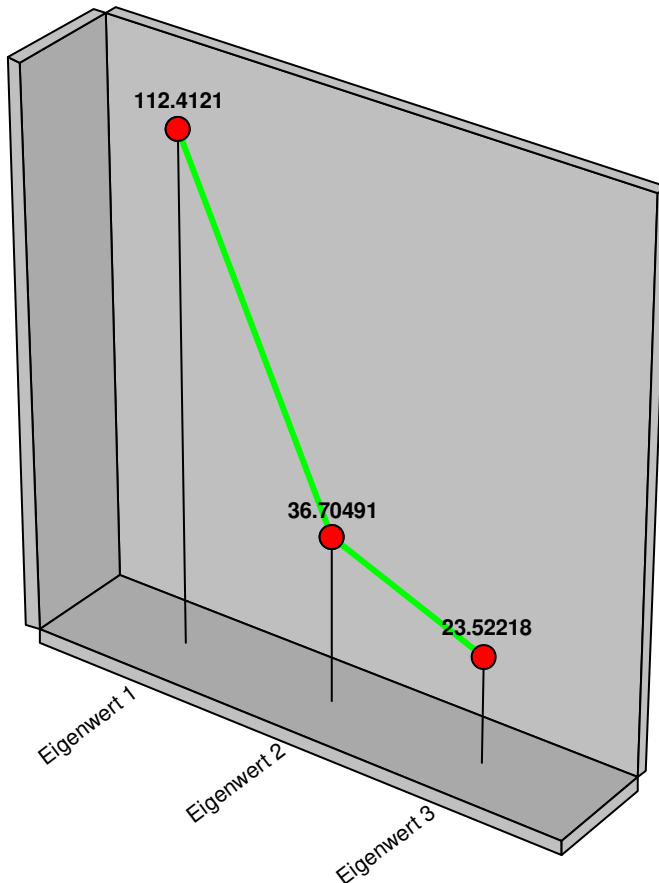
	Faktor 1	Faktor 2	Faktor 3
Opel	0.3411	0.0620	-0.0175
Volkswagen	0.3309	0.1598	-0.0226
Suzuki	0.3365	0.3516	0.4306
Toyota	0.3557	-0.0336	-0.0252
Mercedes	0.0171	-0.4548	-0.2647
BMW	0.0577	-0.2991	-0.5331
Ferrari	-0.3877	0.3353	0.0169
Porsche	-0.3713	0.2567	-0.1521
Lamborghini	-0.4155	0.1983	-0.0865
RollsRoyce	-0.2645	-0.5761	0.6542

In der Programm-Maske hatten wir die Zahl der Faktoren auf 3 beschränkt. Hätten wir das nicht getan, dann hätte Almo versucht, so viele Faktoren zu extrahieren wie Objekte vorhanden sind (also 10) . Wir hätten dann folgende Eigenwerte erhalten

1	2	3	4	5	6	7	8	9
112.41	36.70	23.52	12.43	6.25	4.80	4.01	1.82	-1.44

Nach 8 Eigenwerten bricht Almo ab, da der 9. Eigenwert negativ ist. Almo zeichnet nun folgendes "Scree-Test"-Diagramm der ersten 3 Eigenwerte

Eigenwerte



Nach dem 2. Eigenwert macht die Kurve einen Knick. Der relativ grobe "Scree-Test" legt damit nahe, dass nur 2 Eigenwerte (also Faktoren) bedeutsam sind. Man sollte jedoch nicht zu viel Vertrauen in diesen Test haben.

Bei 8 Eigenwerten hätten wir folgendes Ergebnis erhalten

Prozent Eigenwert je Faktor von gesamter Eigenwertsumme						
55.6612	18.1746	11.6471	6.1556	3.0964	2.3766	
1.9879	0.9005					

gesamte Eigenwertsumme der MDS-Matrix=	201.9577
Durch 8 Faktoren erklärte Eigenwertsumme=	201.9577
Prozentsatz der erkläerten Eigenwertsumme=	100.0000

Nach jedem weiteren extrahierten Faktor berechnet Almo aus den jeweils vorhandenen Faktorladungen die "reproduzierte" Distanzmatrix und durch Vergleich mit der eingegebenen empirischen Distanzmatrix den Koeffizienten "Stress1" (nach Kruskal).

Stress1 nach 1 extrahierten Faktor = 0.40119  
 Stress1 nach 2 extrahierten Faktoren = 0.23181  
 Stress1 nach 3 extrahierten Faktoren = 0.15962  
 Stress1 nach 4 extrahierten Faktoren = 0.09087  
 Stress1 nach 5 extrahierten Faktoren = 0.05813  
 Stress1 nach 6 extrahierten Faktoren = 0.02727  
 Stress1 nach 7 extrahierten Faktoren = 0.01596  
 Stress1 nach 8 extrahierten Faktoren = 0.00863

Wird die Optionsbox "Faktoren/reproduzierte Matrix" geöffnet und im 2. Eingabefeld eine 1 eingesetzt, dann wird zusätzlich zum Stress-Koeffizienten noch (je x Faktoren) die reproduzierte MDS-Matrix und die reproduzierte Distanzmatrix ausgegeben. Wir zeigen dies nach 3 extrahierten Faktoren:

Reproduzierte (quadrierte u. zentrierte) MDS-Matrix nach 3 Faktoren

	Spalte 1	Spalte 2	Spalte 3	Spalte 4	Spalte 5	Spalte 6	Spalte 7	Spalte 8	Spalte 9	Spalte10
1	13.2305	0.6118	1.5782	0.6096	4.7574	3.9358	7.8963	7.7082	8.0997	7.2262
2	0.6118	13.2572	0.9345	1.2142	5.1343	4.2978	7.6840	7.5493	7.9625	7.3806
3	1.5782	0.9345	21.6247	2.0868	3.2629	-0.4733	7.8503	5.9863	7.1498	15.1224
4	0.6096	1.2142	2.0868	14.2785	4.5619	3.8614	8.1828	7.9967	8.3472	6.9629
5	4.7574	5.1343	3.2629	4.5619	9.2742	4.3571	6.3238	6.9084	6.5958	-0.9991
6	3.9358	4.2978	-0.4733	3.8614	4.3571	10.3433	5.8769	7.5666	6.9373	-4.3968
7	7.8963	7.6840	7.8503	8.1828	6.3238	5.8769	21.0325	0.4465	0.8459	5.9349
8	7.7082	7.5493	5.9863	7.9967	6.9084	7.5666	0.4465	18.4632	0.8956	2.8314
9	8.0997	7.9625	7.1498	8.3472	6.5958	6.9373	0.8459	0.8956	21.0220	3.6265
10	7.2262	7.3806	15.1224	6.9629	-0.9991	-4.3968	5.9349	2.8314	3.6265	30.1130

Reproduzierte Distanzmatrix nach 3 Faktoren

	Spalte 1	Spalte 2	Spalte 3	Spalte 4	Spalte 5	Spalte 6	Spalte 7	Spalte 8	Spalte 9	Spalte10
1	0	0.6030	2.7935	0.6004	4.8007	4.4797	7.9050	7.6735	8.0711	8.1724
2	0.6030	0	2.4870	1.2009	5.1298	4.7167	7.6954	7.4949	7.9226	8.3964
3	2.7935	2.4870	0	3.2210	6.8344	6.7911	7.9367	8.0396	8.4088	8.5652
4	0.6004	1.2009	3.2210	0	4.5553	4.3167	8.1954	7.9303	8.3012	8.0559
5	4.8007	5.1298	6.8344	4.5553	0	1.6645	6.5728	5.9866	6.1185	5.4141
6	4.4797	4.7167	6.7911	4.3167	1.6645	0	6.6479	5.9539	6.2404	6.9026
7	7.9050	7.6954	7.9367	8.1954	6.5728	6.6479	0	0.9638	1.0132	6.4615
8	7.6735	7.4949	8.0396	7.9303	5.9866	5.9539	0.9638	0	0.6668	6.4831
9	8.0711	7.9226	8.4088	8.3012	6.1185	6.2404	1.0132	0.6668	0	6.1223
10	8.1724	8.3964	8.5652	8.0559	5.4141	6.9026	6.4615	6.4831	6.1223	0

Stress1 nach 3 extrahierten Faktoren = 0.15962

Die Distanzen der Distanzmatrix werden nach folgender Formel aus den Faktorladungen 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 Variablen

fz = Zahl der Faktoren

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

In obiger Formel wird die quadrierte Distanz bestimmt.

f(ik) bzw. f(jk) = Ladung von Variable 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
}
}

```

**Stress** ist ein Maß für die Relevanz der Faktoren. Von den verschiedenen Varianten dieses Koeffizienten verwendet Almo "Stress1" (nach Kruskal) . Seine Formel ist

$$Stress1 = \sqrt{\frac{sum1}{sum2}}$$

sum1=die Glieder der eingegebenen Distanzmatrix werden quadriert.  
 Dabei werden nur die Glieder der unteren Dreiecksmatrix ohne Diagonale verwendet. Diese quadrierten Glieder werden zu sum1 summiert.

sum2=aus den x extrahierten Faktorladungen wird die "reproduzierte Distanzmatrix" gebildet. Diese wird von der empirischen Distanzmatrix gliedweise subtrahiert. Es entsteht eine "Residualmatrix".  
 Deren Glieder aus der unteren Dreiecksmatrix (ohne Diagonale) werden wie bei sum1 beschrieben quadriert und zu sum2 summiert.

In der Literatur wird die Güte einer MDS-Lösung folgendermaßen beurteilt:

Stress1	Goodness of fit
-----	-----
größer 0.2	schwach, nicht genügend
0.2 bis 0.1	akzeptabel
0.1 bis 0.05	gut
kleiner 0.05	sehr gut
0	perfekt

Der Stress-Koeffizient ist kein statistischer Signifikanztest, der es dem Benutzer erlauben würde zu konstatieren, dass die gefundene Faktoren-Lösung mit einer Sicherheit von x Prozent signifikant ist. Er ist ein Maß, dass versucht eine brauchbare Lösung zu identifizieren.

Für die Berechnung von Stress wird die reproduzierte Distanzmatrix und nicht die "reproduzierte MDS-Matrix" benötigt. Wird die Optionsbox "Faktoren / reproduzierte Matrix" geöffnet, dann kann man (im 2. Eingabefeld) die Ausgabe der "reproduzierten Matrix" anfordern. Almo berechnet die reproduziert Distanzmatrix nachdem 1 Faktor extrahiert wurde, dann nachdem 2 Faktoren extrahiert wurden, dann nachdem 3 Faktoren extrahiert wurden usw. Man kann dann erkennen, dass die empirische Distanzmatrix immer genauer reproduziert wird, je mehr Faktoren extrahiert wurden. Der Stress-Koeffizient hilft dann zu entscheiden, wann genügend Faktoren extrahiert wurden. In nachfolgender Tabelle wollen wir das illustrieren.

Opel	Volkswagen	Suzuki	Toyota	Mercedes	BMW	Ferrari	Porsche	Lamborghini	RollsRo
----	-----	-----	-----	-----	----	-----	-----	-----	-----
0	3.1000	5.0000	3.8000	5.9000	5.5000	8.4000	8.4000	8.5000	8.5000
0	0.1087	0.0496	0.1543	3.4359	3.0048	7.7278	7.5541	8.0217	6.4211
nach 1 Faktor: Stress1=			<b>0.40119</b>						
-----									
0	3.1000	5.0000	3.8000	5.9000	5.5000	8.4000	8.4000	8.5000	8.5000
0	0.6025	1.7552	0.5993	4.6485	3.7167	7.9032	7.6457	8.0642	7.4951
nach 2 Faktoren: Stress1=			<b>0.23181</b>						
-----									
0	3.1000	5.0000	3.8000	5.9000	5.5000	8.4000	8.4000	8.5000	8.5000
0	0.6030	2.7935	0.6004	4.8007	4.4797	7.9050	7.6735	8.0711	8.1724
nach 3 Faktoren: Stress1=			<b>0.15962</b>						
-----									
:	:	:	:	:	:	:	:	:	:
-----									

```

0          3.1000   5.0000   3.8000   5.9000   5.5000   8.4000   8.4000   8.5000   8.5000
0          3.1073   5.0101   3.8621   5.9060   5.5430   8.4049   8.4001   8.5099   8.5014
nach 8 Faktoren: Stress1= 0.00863

```

In der jeweils 1. Zeile stehen die Werte aus der "empirischen Distanzmatrix" von Opel gegenüber den anderen Automarken. In der jeweils 2. Zeile stehen die Werte aus der "reproduzierten Distanzmatrix" nachdem 1 Faktor extrahiert wurde, dann nachdem 2 Faktoren extrahiert wurden, dann nach 3 und 8 Faktoren. In der 3. Zeile geben wir den entsprechenden Stress-Koeffizienten an. Es ist gut zu sehen, wie die Reproduktion immer besser wird. Schon nach 2 Faktoren erreichen wir einen Stress-Wert von 0.2, so dass uns nahe gelegt wird, uns mit 2 Faktoren zu begnügen.

Die Faktorladungen entstehen nun sehr einfach dadurch, dass die Werte des Eigenvektors  $k$  mit der Wurzel aus dem Eigenwert  $k$  multipliziert werden. Wir erhalten:

Matrix der Faktorladungen

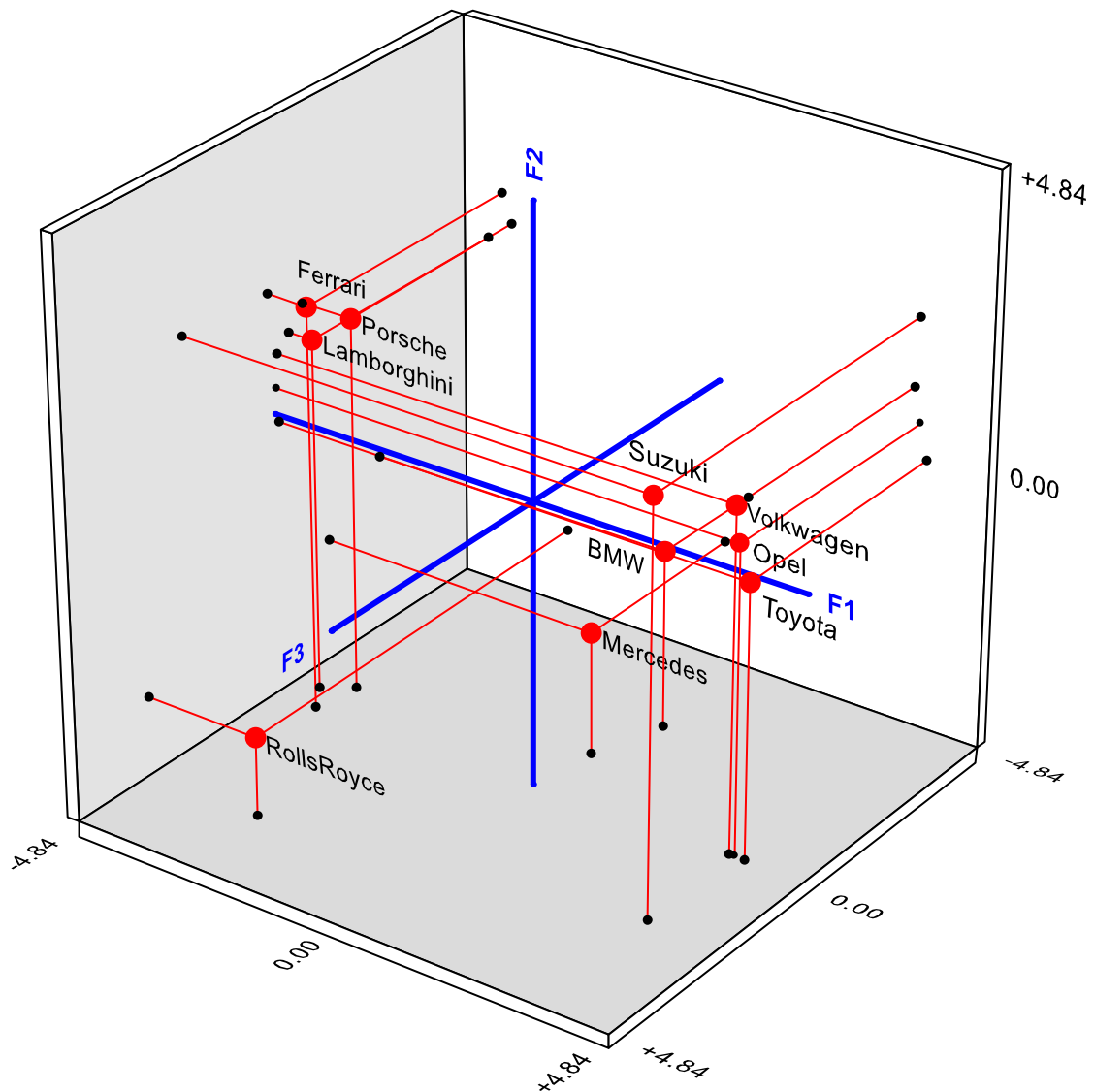
		Faktor 1	Faktor 2	Faktor 3
Opel	V1	3.62	0.38	-0.08
Volkswagen	V2	3.51	0.97	-0.11
Suzuki	V3	3.57	2.13	2.09
Toyota	V4	3.77	-0.20	-0.12
Mercedes	V5	0.18	-2.76	-1.28
BMW	V6	0.61	-1.81	-2.59
Ferrari	V7	-4.11	2.03	0.08
Porsche	V8	-3.94	1.55	-0.74
Lamborgh	V9	-4.40	1.20	-0.42
RollsRoy	V10	-2.80	-3.49	3.17

Dies ist das zentrale Ergebnis der metrischen MDS. Almo fügt noch die Mitteilung an:

Je nach gewähltem Eigenwert-Kalkuel kann eine spaltenweise Vorzeichen-Umkehr stattfinden. Dieser entspricht grafisch eine Spiegelung um die Achse - und ist somit bedeutungslos

Almo zeichnet nun folgende 3-dimensionale Grafik

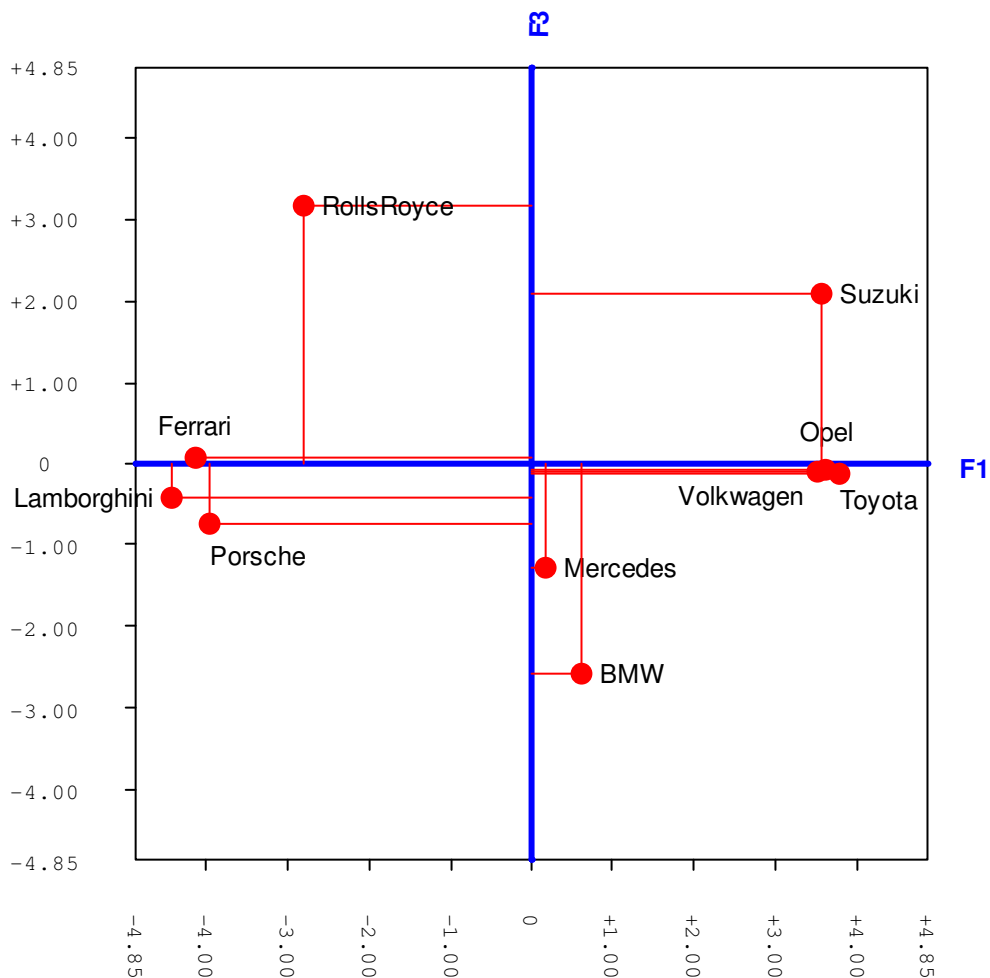
## Faktorladungen



Wir haben oben bereits versucht die 2-dimensionale Grafik zu interpretieren. Bei der 3-dimensionalen Grafik erkennt man nun, die Sonderrolle von Suzuki, BMW und besonders Rolls Royce sehr deutlich. Sie weichen im Faktor F3 erheblich von den anderen ab.

Man kann das sichtbar machen, wenn man im Grafik-Editor (nach Klick auf den großen Grafik-Knopf in der Ergebnisausgabe) in der linken Werkzeugleiste auf "Diverse Positionen" und danach auf "F1-F3 Projektion auf Bodenplatte" klickt (siehe nachfolgendes Bild des Grafik-Editors). Man erhält dann folgende Grafik:

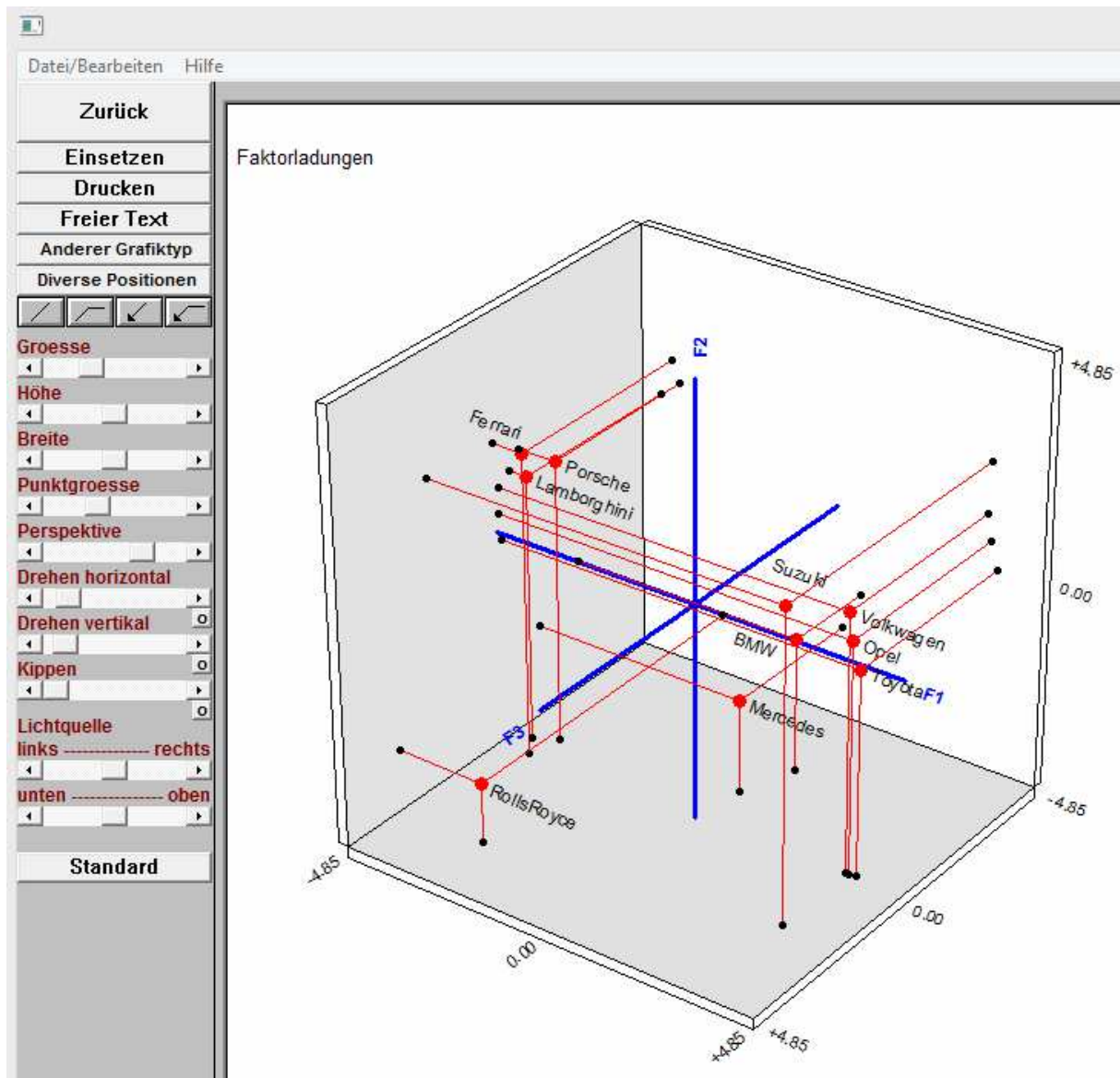
## Faktorladungen



### **Projektion auf untere Wand bzw. auf die Fläche F1-F3**

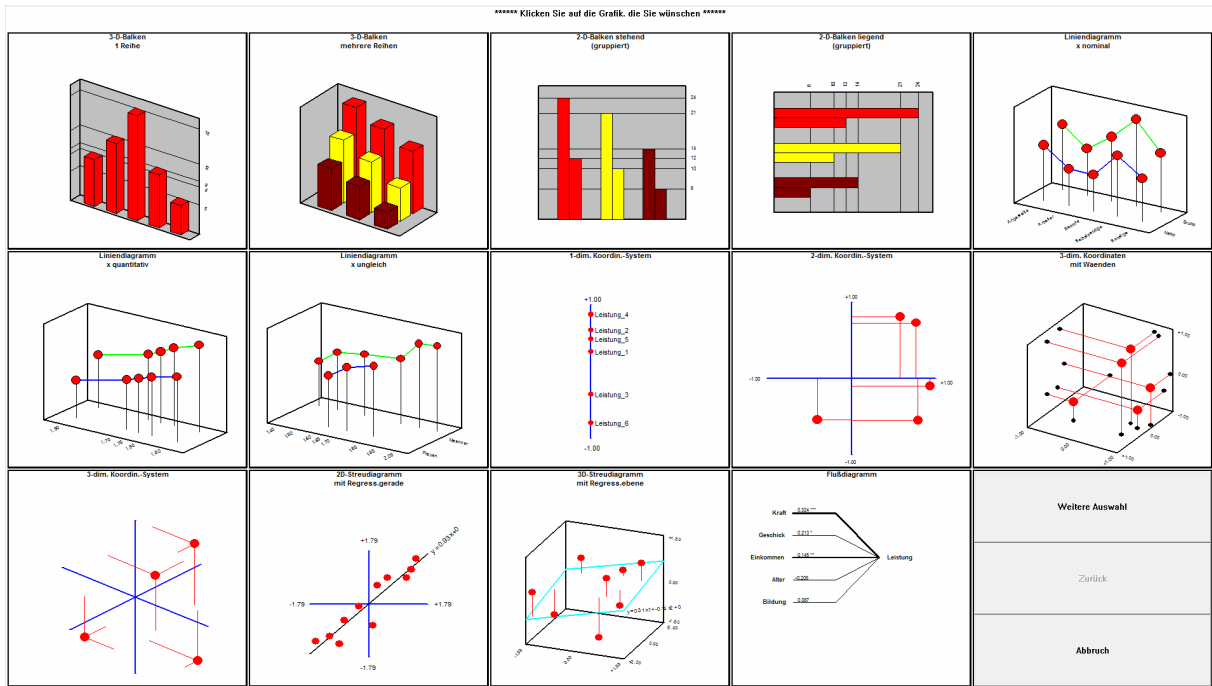
Im Faktor F3 weichen Rolls Royce und Suzuki weit nach oben ab und BMW weit nach unten. Die anderen Automarken schmiegen sich enger an die F1-Achse an. F3 ist für sie nicht relevant.

Da sowohl der "Scree-Test" als auch der Stress1-Koeffizient nahe legen, dass nur 2 Faktoren bedeutsam sind, wollen wir die 3-dimensionale Grafik auf 2 Dimensionen reduzieren. Im Grafik-Editor (von dem wir hier nur einen Ausschnitt zeigen) geschieht folgendes:



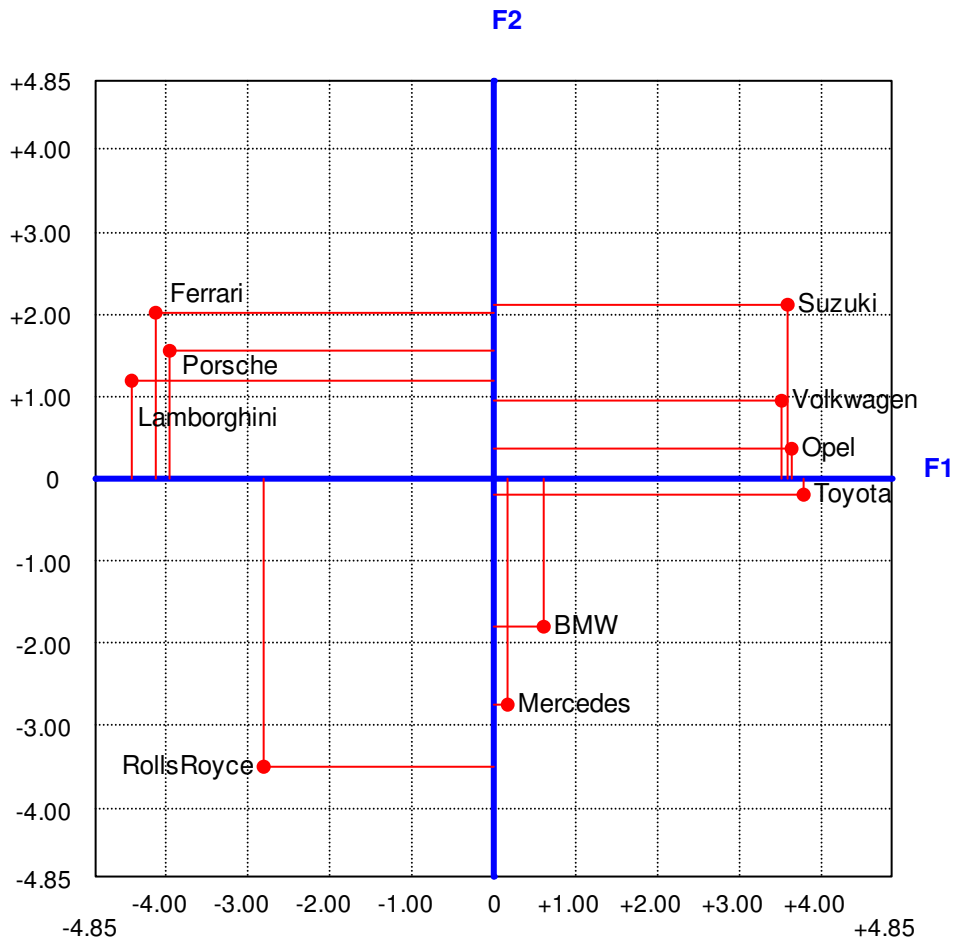
Klicken Sie in der linken Werkzeugleiste auf den Knopf "Anderer Grafiktyp". Sie sehen dann folgendes:

(Alternativ könnten Sie auch auf "Diverse Positionen" klicken und dann auf den 2-dimensionalen Grafik-Ausschnitt F1-F2)



Die Abbildung ist stark verkleinert, so dass die Beschriftung nicht lesbar ist. Klicken Sie jetzt auf die kleine Symbolgrafik "2-dim. Koordin.-System" (in der 2. Zeile und 4. Spalte). Also erzeugt dann aus den ersten beiden Faktoren eine 2-dimensionale Grafik

Faktorladungen



Diese 2-dimensionale Grafik genügt, um zu erkennen, dass die Autos typische Gruppen bilden

- (1) Ferrari, Porsche, Lamborghini bilden die Gruppe der rasanten Sportwagen,
- (2) Suzuki, Volkswagen, Opel, Toyota bilden die Gruppe der "Alltagsautos"
- (3) BMW und Mercedes sind die Autos für den anspruchsvolleren Fahrer
- (4) Rolls Royce ist das edle Ausnahme-Auto

Aus der Faktorladungsmatrix reproduziert nun Also die Distanzmatrix der Automarken

vom Modell prognostizierte Matrix der euklidischen Distanzen zwischen den Variablen (nach 8 Faktoren)

	Opel	Volkswag	Suzuki	Toyota	Mercedede
Opel	0	0.60	2.79	0.60	4.80
Volkswage	0.60	0	2.49	1.20	5.13
Suzuki	2.79	2.49	0	3.22	6.83
Toyota	0.60	1.20	3.22	0	4.56
Mercedes	4.80	5.13	6.83	4.56	0
BMW	4.48	4.72	6.79	4.32	1.66
Ferrari	7.91	7.70	7.94	8.20	6.57
Porsche	7.67	7.49	8.04	7.93	5.99
Lamborgh	8.07	7.92	8.41	8.30	6.12
RollsRoy	8.17	8.40	8.57	8.06	5.41

	BMW	Ferrari	Porsche	Lamborg	RollsRo
Opel	4.48	7.91	7.67	8.07	8.17
Volkswage	4.72	7.70	7.49	7.92	8.40
Suzuki	6.79	7.94	8.04	8.41	8.57
Toyota	4.32	8.20	7.93	8.30	8.06
Mercedes	1.66	6.57	5.99	6.12	5.41
BMW	0	6.65	5.95	6.24	6.90
Ferrari	6.65	0	0.96	1.01	6.46
Porsche	5.95	0.96	0	0.67	6.48
Lamborgh	6.24	1.01	0.67	0	6.12
RollsRoy	6.90	6.46	6.48	6.12	0

Werden die prognostizierten Distanzen mit den tatsächlichen verglichen, dann wird man einen kleinen Unterschied feststellen. Der Unterschied wird umso kleiner, je mehr Faktoren extrahiert und für die Reproduzierung der Distanzmatrix verwendet werden. Nachfolgend haben wir die Distanzen von Rolls Royce gegenüber den anderen Automarken aufgelistet. In der 1. Zeile stehen die tatsächlichen Distanzen, in der 2. die durch 8 Faktoren reproduzierten Distanzen und in der 3. Zeile dir nur durch 3 Faktoren reproduzierten Distanzen.

	Opel	Volkswag	Suzuki	Toyota	Mercedede	BMW	Ferrari	Lamborg
RollsRoy	8.5000	8.6000	8.9000	8.2000	5.8000	7.0000	6.6000	6.3000
RollsRoy	8.5014	8.6002	8.9015	8.2434	5.8153	7.0204	6.6013	6.3254
RollsRoy	8.1724	8.3964	8.5652	8.0559	5.4141	6.9026	6.4615	6.1223

### P34.6.9 Prog33m3: Paarvergleich mit anschließender MDS

Das Programm umschließt (für den Benutzer unbemerkbar) 2 Programme zu einem. Zuerst werden (z.B. aus einer Befragung von Personen) Paarvergleichs-Daten eingelesen. Diese werden zu einem Mittelwert je Vergleichspaar zusammengefasst. Aus diesen Mittelwerten wird dann eine Ähnlichkeitsmatrix geformt. Im 2. Teil wird diese Matrix dem Kalkül der metrischen MDS

unterworfen. Das Programm wird erreicht durch Klick auf den Knopf "Verfahren", dann "MDS/MDU.

Es wird folgendes (konstruierte) Beispiel gerechnet:

In einer Untersuchung sollen Befragten 6 Automarken vergleichen. Die 6 Automarken sind:

	Variablennamen
	-----
VW Golf	Golf
Opel Astra	Astra
Ford Focus	Focus
Audi A3	A3
Skoda Octavia	Octavia
BMW 1er	BMW1er

Aus den 6 Autos wurden folgende 15 Paare gebildet. Die Paare müssen in einer spezifischen Weise gebildet werden. Sie entstehen durch systematisches von oben nach unten Kombinieren

V1	Golf_Astra
V2	Golf_Focus
V3	Golf_A3
V4	Golf_Octavia
V5	Golf_BMW1er
V6	Astra_Focus
V7	Astra_A3
V8	Astra_Octavia
V9	Astra_BMW1er
V10	Focus_A3
V11	Focus_Octavia
V12	Focus_BMW1er
V13	A3_Octavia
V14	A3_BMW1er
V15	Octavia_BMW1er

Zu jedem Paar muss der Befragte nun angeben, wie ähnlich die beiden Autos sind. Die Frage könnte z.B. so lauten

Nehmen Sie einmal an,

Sie würden vor der Entscheidung stehen, jetzt ein Auto zu kaufen

Wie ähnlich sind Ihrer Meinung nach die beiden Autos ?

		nicht ähnlich		mittel		stark ähnlich						
		0	1	2	3	4	5	6	7	8	9	10
Golf - Astra	--wie sehr ähnlich											
Golf - Focus		.						.				
Golf - A3		.						.				
.		.						.				
.		.						.				
.		.						.				
Octavia - BMW1er		.						.				

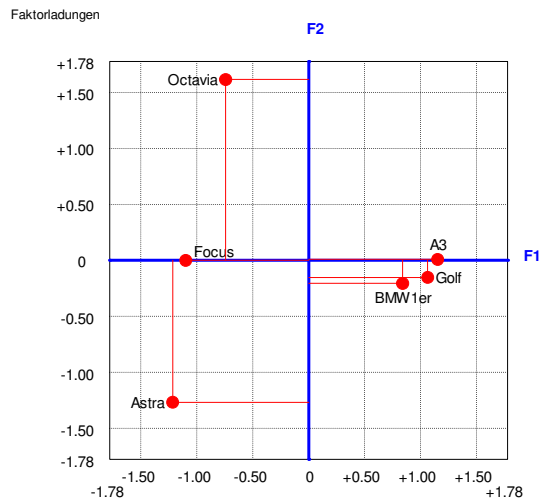
Für jeden Befragten wird ein Datensatz eingelesen, der aus diesen 15 Paarvergleichen besteht. In der Ausgabe aus Prog33m3 wird dann folgende Ähnlichkeitsmatrix ausgegeben:

	Objekt 1	Objekt 2	Objekt 3	Objekt 4	Objekt 5	Objekt 6
	Golf	Astra	Focus	A3	Octavia	BMW1er
	-----	-----	-----	-----	-----	-----
Objekt 1 Golf	-	2.7833	2.9167	5.6000	2.9167	4.6000
Objekt 2 Astra	2.7833	-	3.4500	2.6833	2.7000	3.1667
Objekt 3 Focus	2.9167	3.4500	-	3.2000	3.3000	3.2167
Objekt 4 A3	5.6000	2.6833	3.2000	-	3.2833	5.6667



Objekt 5	Octavia	2.9167	2.7000	3.3000	3.2833	-	3.1500
Objekt 6	BMW1er	4.6000	3.1667	3.2167	5.6667	3.1500	-

Schließlich wird dann folgende Grafik der Distanzen zwischen den Autos ausgegeben.



### P34.6.10 Paarvergleich mit Präferenzurteil

Es ist naheliegend mit Prog33m3 aus Präferenzurteilen eine Ähnlichkeitsmatrix zu erzeugen und diese einer MDS-Analyse zu unterwerfen. Hier ist aber Vorsicht geboten.

Die Frage in oben dargestellten Beispiel lautete "Wie ähnlich sind die beiden Autos?". Nun könnte man auch fragen:

Welches der beiden Autos ziehen Sie vor?  
 Unterstreichen Sie das Auto das Sie vorziehen  
 und geben Sie an wie stark Sie dieses Auto vorziehen.

Aus der Ähnlichkeitsfrage ist damit eine Präferenzfrage geworden.

Dabei entsteht ein Problem. Beim Paar "Golf - Astra" in unserem Beispiel könnte z.B. ein Befragter den Golf mit 8 gegenüber dem Astra präferieren. Ein anderer Befragter präferiert hingegen den Astra (vielleicht ebenfalls mit 8). Wie soll das gehandhabt werden?

Das auf Paarvergleiche angewandte MDS-Verfahren analysiert die Distanz zwischen den beiden Objekten - gleichgültig welches der beiden Objekte präferiert wird. Nun könnte man versucht sein, eine besondere Kodierung anzuwenden, wenn das 2. Objekt präferiert wird, z.B. die Antworten des Befragten dann mit -1 zu multiplizieren. Das führt, wie man es auch macht, zu einer verfälschten Distanzmatrix und damit zu falschen Ergebnissen des MDS-Verfahrens.

Präferenzfragen sind also nicht äquivalent zu Ähnlichkeitsfragen. Die MDS-Analyse untersucht Distanz- bzw. Ähnlichkeitsmatrizen. Die eigentliche Information, die die Präferenzfrage liefert, nämlich dass das Objekt i dem Objekt k vorgezogen wird, geht verloren.

Präferenzurteile sollten mit dem Verfahren des "multidimensionalen Unfolding" (kurz: MDU) analysiert werden. Siehe dazu die entsprechenden Almo-Programme Prog18m6 usw. und die ausführliche Darstellung im Almo-Dokument 29 "Metrisches multidimensionales Unfolding (MDU) "

## Literatur

Die Literatur zur MDS ist sehr umfangreich und auch sehr zerstreut über viele Bücher und Zeitschriften-Artikel. Einen Überblick über die MDS geben

**Borg, I., Groenen, P.:** Modern multidimensional scaling: theory and applications, 2005, New York, Springer, 2. Auflage

Eine kurze und sehr übersichtliche Darstellung (auch zur nicht-metrischen MDS) ist im Internet zu finden bei

**Jacoby, W. G. :** Multidimensional Scaling, an Introduction, 2012, im Internet herunter ladbar bei <http://polisci.msu.edu/jacoby/iu/mds2012/outline/2012%20IU%20MDS%20Outline,%2012-2-12.pdf>

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

**BUGH** Wuppertal: "Multidimensionale Skalierung", Lehrstuhl für empirische Wirtschafts- und Sozialforschung, Fachbereich Wirtschaftswissenschaft, 2001, im Internet herunter ladbar 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)

**Busing, Comandeur, Heiser:** Proxscal, a multidimensional scaling program for individual differences scaling with constraints, 1997, im Internet herunter ladbar bei [http://www.researchgate.net/publication/244439009\\_PROXSCAL\\_a\\_multidimensional\\_scaling\\_program\\_for\\_individual\\_differences\\_scaling\\_with\\_constraints](http://www.researchgate.net/publication/244439009_PROXSCAL_a_multidimensional_scaling_program_for_individual_differences_scaling_with_constraints)

**Kruskal J. B.:** Nonmetric multidimensional scaling, Psychometrika, 1964, 29, S. 115-131

**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.

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