

Multidimensionale Skalierung

Beispieldatei zur Datenanalyse

Lehrstuhl für empirische Wirtschafts- und Sozialforschung

Fachbereich Wirtschaftswissenschaft

BUGH Wuppertal

2001

Multidimensionale Skalierung - Inhalt

1. Grundlagen der Multidimensionalen Skalierung	1
1.1 Einführende Beispiele.....	1
1.1.1 Ein Beispiel zur Wahrnehmung von Farben.....	1
1.1.2 Ein Beispiel zur Wahrnehmung von Automarken	3
1.2 Vom Wahrnehmungsraum über Ähnlichkeitsurteile zur Konfiguration	6
1.3 Modellkonstellationen der MDS	9
1.3.1 Quadratische Ähnlichkeitsmatrizen (Objekt x Objekt)	9
1.3.2 Rechteckige Ähnlichkeitsmatrizen (Objekt x Subjekt)	10
1.4 Distanzmetriken der MDS	12
2. Auswahl der Objekte und Erhebung der Ähnlichkeiten	18
2.1 Auswahl der Objekte.....	18
2.2 Erhebung von Ähnlichkeitsurteilen	18
2.2.1 Ratingverfahren.....	18
2.2.2 Rangreihung	19
2.2.3 Ankerpunktmethod.....	20
2.2.4 Datenerhebung für die MDU.....	20
3. Der technische Ablauf der MDS: Von Ähnlichkeiten über Disparitäten zu Distanzen	21
3.1 Optimierung 1: Transformation der Rohdaten	23
3.2 Optimierung 2: Anpassung der Konfiguration	27
4. Durchführung der MDS mit SPSS	29
4.1 Einfache ordinale MDS.....	29
4.1.1 Dateneingabe.....	29
4.1.2 Aufruf der Prozedur und Optionen.....	30
4.1.3 Interpretation des Output	34
4.1.4 Die Gefahr der entarteten Lösung einer ordinalen MDS.....	41
4.2 Einfache Intervall-MDS.....	42
4.2.1 Aufruf der Prozedur und Optionen.....	43
4.2.2 Interpretation des Output	43
4.3 Wiederholte MDS.....	45
4.3.1 Dateneingabe.....	45
4.3.2 Aufruf der Prozedur und Optionen.....	46

4.3.3 Interpretation des Output	46
4.4 INDSCAL	50
4.4.1 Aufruf der Prozedur und Optionen.....	50
4.4.2 Interpretation des Output	51
5. Multidimensionales Entfalten (MDU): Das Idealpunktmodell zur Interpretation der Konfiguration	57
5.1 Dateneingabe.....	58
5.2 Aufruf der Prozedur und Optionen.....	58
5.3 Interpretation des Output	59
5.4 ALLBUS 1994: Parteiensympathie und Links-Rechts-Selbsteinstufung im West-Ost-Vergleich.....	60
6. Property Fitting - das Vektormodell zur Interpretation der euklidischen Konfiguration	64
7. Ausblick.....	70

1. Grundlagen der Multidimensionalen Skalierung

Die multidimensionale Skalierung (MDS) ist ein Verfahren, mit dem Objekte auf Basis ihrer Ähnlichkeit zueinander in einem (möglichst niedrig dimensionierten) Raum dargestellt werden. In diesem Sinne kann die MDS als Datenreduktionstechnik analog zur Faktorenanalyse verstanden werden. Die MDS wird daher auch Ähnlichkeitsstrukturanalyse (similarity structure analyses [SSA]) genannt. Die MDS bzw. SSA wird in diesem Zusammenhang explorativ zur Konstruktion eines (möglichst niedrig dimensionierten) Wahrnehmungsraums eingesetzt. In den meisten Anwendungen tritt dabei der ursprüngliche Zweck der MDS, nämlich die multidimensionale Skalierung vorgegebener Dimensionen in den Hintergrund. Um so wichtiger wird dann eine im Anwendungsbezug plausible Interpretation des durch die MDS erzeugten Wahrnehmungsraums. In diesem Zusammenhang kann die zusätzliche Integration unabhängig erhobener Beurteilungsdimensionen in den Wahrnehmungsraum mit Hilfe des Vektormodells eine wertvolle Interpretationshilfe darstellen.

Bevor wir mit weiteren detaillierten Erläuterungen beginnen, folgen nun zunächst zwei einfache Beispiele, die den Ablauf und das Ziel des Verfahrens veranschaulichen sollen.

1.1 Einführende Beispiele

1.1.1 Ein Beispiel zur Wahrnehmung von Farben

In einer empirischen Untersuchung haben wir 10 Personen gebeten, Urteile über die Ähnlichkeit von Farbkombinationen abzugeben. Die Menge der durch die MDS zu repräsentierenden Objekte wird aus den Farben Rot, Orange, Gelb, Grün, Blau und Violett gebildet. Den Personen wurden 15 Karten präsentiert, auf denen zwei verschiedenfarbige Kästchen abgebildet waren. Die Befragten mussten nun entscheiden, ob sich die beiden abgebildeten Farben „sehr ähnlich“ oder „nicht sehr ähnlich“ sind.

Als Ergebnis dieser Datenerhebung liegt die folgende Tabelle 1.1.1.1 vor:

Tab. 1.1.1.1: Häufigkeit der Nennung „Die Farben sind sich nicht sehr ähnlich“

	Rot	Orange	Gelb	Grün	Blau	Violett
Rot	-					
Orange	6	-				
Gelb	8	0	-			
Grün	10	8	9	-		
Blau	10	10	10	6	-	
Violett	0	7	10	9	7	-

6 Befragte waren der Ansicht, Rot und Orange seien sich nicht sehr ähnlich. Alle Befragten (10) bezeichneten Grün und Rot als nicht sehr ähnlich. Dagegen war keiner der Befragten der Meinung, Orange und Gelb seien sich nicht sehr ähnlich (usw.).

Die hier dargestellte Tabelle 1.1.1.1 wird auch als *Unähnlichkeitsmatrix* bezeichnet, da die Werte der Matrix als *Unähnlichkeiten* interpretiert werden können: Je größer die Ausprägung eines Wertes in der Matrix, desto weniger ähnlich werden die Objekte der zugehörigen Zeile und Spalte angesehen. Wie weiter unten noch erläutert wird, werden in der Regel keine Häufigkeiten in dieser Matrix erfasst, sondern (Un-)Ähnlichkeitsratings oder Rangfolgen.

Die Unähnlichkeitsmatrix stellt die Datenbasis für die nun durchzuführende Multidimensionale Skalierung (MDS) dar. Das formale Ziel der MDS ist es, die Objekte (im Beispiel: Farben) räumlich so anzuordnen, dass die Abstände (*Distanzen*) zwischen den Objekten im Raum möglichst exakt den erhobenen Unähnlichkeiten entsprechen. Dabei soll folgende Regel gelten:

Je unähnlicher zwei Farben beurteilt werden, desto größer ist ihre Distanz im Lösungsraum der MDS!

Im Idealfall hat dieser Raum, die sog. *Konfiguration*, nicht mehr als zwei Dimensionen, was die Darstellung (auf dem Papier) und die Interpretierbarkeit erheblich erleichtert.

Auf die Unähnlichkeitsmatrix in Tabelle 1.1.1.1 angewendet, erzeugt die MDS die folgende zweidimensionale Konfiguration (Abbildung 1.1.1.1):

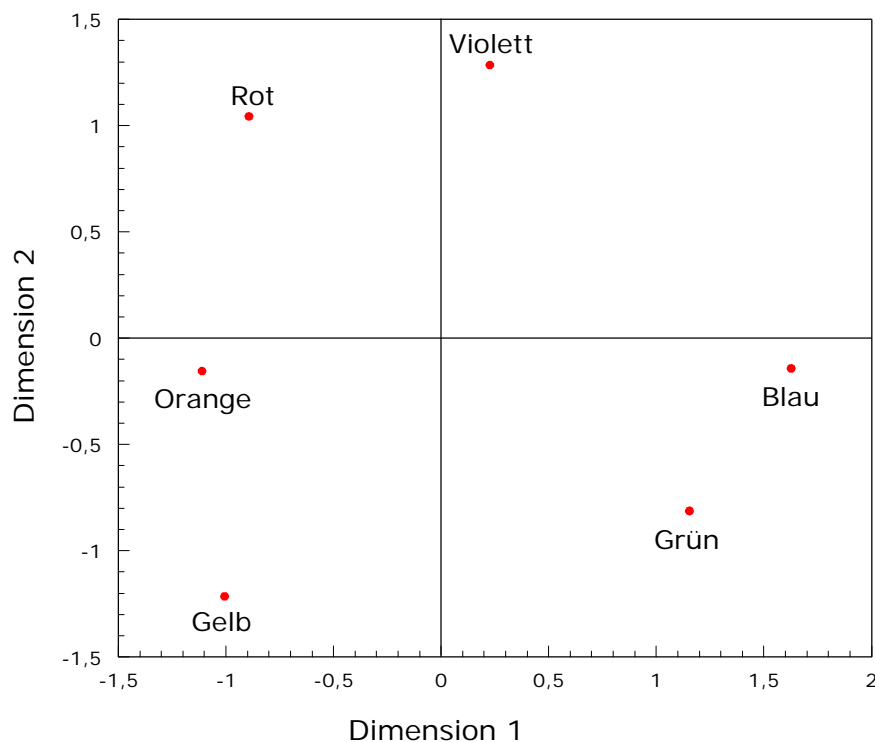


Abb. 1.1.1.1: Zweidimensionale Konfiguration von Farben
(ordinale MDS: Stress=0,04; RSQ=0,98)

Ein Vergleich der Konfiguration in Abbildung 1.1.1.1 mit Tabelle 1.1.1.1 zeigt, dass die Anpassung der Abstände im Raum an die erhobenen Unähnlichkeiten recht gut gelungen ist:

- Farben, die als nicht sehr ähnlich zueinander eingestuft wurden, liegen weit auseinander (z.B. Grün und Rot).

- Farben, die als ähnlich zueinander empfunden wurden, liegen nahe beieinander (z.B. Orange und Gelb)

Interessant an der Konfiguration in Abbildung 1.1.1.1 ist, dass die Anordnung der Farben in etwa der Anordnung im bekannten Farbkreis entspricht: Zwischen den drei Grundfarben Blau, Rot und Gelb liegen jeweils die entsprechenden Mischfarben Violett, Orange und Grün.

Die Interpretation der Konfiguration lässt Rückschlüsse auf die Kriterien zu, welche die Befragten bei der Beurteilung der Farbähnlichkeiten zugrunde gelegt haben. Es erscheint naheliegend, dass die Probanden die Ähnlichkeit der Farben auf Basis des Farbkreises beurteilt haben. Die Farben wurden also auch in der Vorstellung der Probanden nach einem bestimmten Schema räumlich angeordnet. Diese räumliche Anordnung der Objekte in der Vorstellung der Befragten wird im folgenden als *Wahrnehmungsraum* bezeichnet. Der Wahrnehmungsraum - ein theoretisches Konstrukt - ist das, was den Forscher im Zuge der MDS eigentlich interessiert. Die MDS erzeugt auf der Basis von Ähnlichkeitsurteilen eine räumliche Konfiguration von Objekten, die ein Abbild des Wahrnehmungsraumes der Befragten darstellt. Zusätzlich liefert die MDS eine Reihe von Kennziffern (z.B. Stress, RSQ), anhand derer neben der grundsätzlichen Möglichkeit auch die Güte der Messung des Wahrnehmungsraumes beurteilt werden kann.

1.1.2 Ein Beispiel zur Wahrnehmung von Automarken

In einer weiteren empirischen Untersuchung wurde die Wahrnehmung folgender Automarken untersucht: Opel, VW, Suzuki, Toyota, Mercedes, BMW, Ferrari, Porsche, Lamborghini und Rolls Royce. Dazu wurde 10 Männern die folgende Aufgabe gestellt:

„Beurteilen Sie paarweise die Ähnlichkeit der folgenden Automarken. Vergeben sie für die Ähnlichkeit eines Automarkenpaares eine Zahl im Wertebereich von 1 (= sehr ähnlich) bis 9 (= sehr unähnlich)!“

Diese Erhebungsform von Ähnlichkeiten wird als *Ratingverfahren* bezeichnet. Die Unähnlichkeitsmatrix (Tabelle 1.1.2.1) enthält die Mittelwerte der Ähnlichkeitsratings. Je höher der Mittelwert eines Automarkenpaares, desto unähnlicher wurden diese beiden Automarken im Durchschnitt von den befragten Männern eingestuft.

Tab. 1.1.2.1: Unähnlichkeitsmatrix (Mittlere Unähnlichkeit von Automarkenpaaren)

	Opel	VW	Suzuki	Toyota	Mercedes	BMW	Ferrari	Porsche	Lamborghini	Rolls Royce
Opel	-									
VW	3,1	-								
Suzuki	5,0	4,4	-							
Toyota	3,8	3,3	3,7	-						
Mercedes	5,9	5,8	7,0	5,3	-					
BMW	5,5	5,5	7,0	4,2	2,7	-				
Ferrari	8,4	8,1	8,3	8,3	6,9	6,8	-			
Porsche	8,4	8,1	8,4	8,3	6,4	6,4	3,0	-		
Lamborghini	8,5	8,2	8,8	8,7	6,6	6,4	2,1	3,4	-	
Rolls Royce	8,5	8,6	8,9	8,2	5,8	7,0	6,6	6,8	6,3	-

Die Marken Ferrari und Lamborghini wurden von den Befragten mit einer Unähnlichkeit von 2,1 am ähnlichsten eingestuft. Am wenigsten ähnlich werden Suzuki und Rolls Royce mit einem Wert von 8,9 wahrgenommen. Jeweils die drei ähnlichsten und die drei unähnlichsten Paarungen sind in Tabelle 1.1.2.1 farbig hervorgehoben.

Die Rohdaten aus Tabelle 1.1.2.1 wurden nun im Zuge einer ordinalen MDS analysiert (analog zu Beispiel 1.1.1). Das Adjektiv *ordinal* bezieht sich auf das Skalenniveau der Rohdaten. Unterstellt wird bei ordinalem Skalenniveau, dass die Meßwerte eine Rangfolge abbilden, die Abstände zwischen ihnen jedoch nicht festgelegt sind. Tabelle 1.1.2.2 enthält die Rohdaten sowie die entsprechenden Rangplätze und die Distanzen in der Konfiguration der drei unähnlichsten Paarungen:

Tab. 1.1.2.2: Rohdaten und Rangplätze von drei Automarkenpaaren

Paare	Rohdaten	Rangplatz	Distanz
Lamborghini - Toyota	8,7	33	2,97
Lamborghini - Suzuki	8,8	34	3,04
Rolls Royce - Suzuki	8,9	35	3,51

Da nur die ordinale Information der Rohdaten in der ordinalen MDS verwendet wird, ist es unerheblich, wie groß der absolute Unterschied zwischen den Unähnlichkeiten zweier Paarungen ist: Alle drei Paare unterscheiden sich um einen Rangplatz. Das bedeutet, in der Konfiguration sollte die Distanz zwischen Lamborghini und Toyota kleiner sein als die Distanz zwischen Lamborghini und Suzuki. Diese Distanz zwischen Lamborghini und Suzuki sollte wiederum geringer sein als die Distanz zwischen Rolls Royce und Suzuki. Um welchen Betrag die Distanz geringer sein sollte, wird nicht festgelegt. Wie Tabelle 1.1.2.2 zu entnehmen ist, spiegelt sich die Reihenfolge der Unähnlichkeiten in den Distanzen wider, ohne dass die Abstände zwischen den Distanzen den Abständen zwischen den Rangplätzen entsprechen. Will man überprüfen, ob das Verfahren tatsächlich nur die ordinale Information der Rohdaten verwendet, kann man anstelle

der Rohdatenmatrix die Matrix der Ränge eingeben. Das Ergebnis der MDS unterscheidet sich nicht.¹

Die ordinale MDS der Daten in Tabelle 1.1.2.1 brachte die folgende Konfiguration hervor (Abbildung 1.1.2.1):

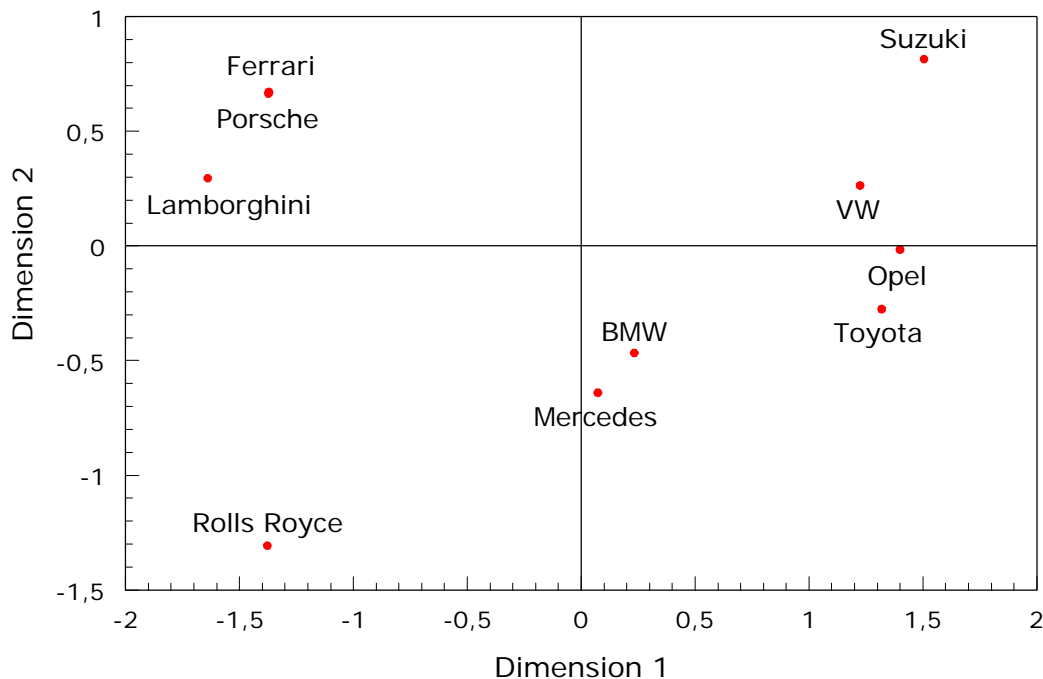


Abb. 1.1.2.1: Zweidimensionale Konfiguration von Automarken (ordinale MDS: Stress=0,06; RSQ=0,98)

Auch in diesem Beispiel zeigt ein Vergleich der Konfiguration mit den Rohdaten, dass die Distanzen in der Konfiguration die Unähnlichkeiten zwischen den Automarken recht gut widerspiegeln:

- Marken, die als sehr unähnlich eingestuft wurden, liegen weit auseinander (z.B. Rolls Royce und Suzuki).
- Marken, die als sehr ähnlich eingestuft wurden, liegen eng beieinander (z.B. BMW und Mercedes).

Es werden aber auch Abweichungen deutlich. So werden die Marken Ferrari und Porsche beinahe exakt übereinanderliegend abgebildet, obwohl Ferrari in der Rohdatenmatrix Lamborghini ähnlicher als Porsche ausgewiesen ist. Diese Abweichungen schlagen sich in der Erhöhung des Stress-Maßes und der Verringerung des RSQ nieder (Erläuterungen zu diesen Maßen weiter unten).

Wie läßt sich nun die Konfiguration der Automarken interpretieren? Es fällt auf und ist auch theoretisch plausibel, dass einige der untersuchten Automarken zu Gruppen zusammengedrückt sind, die gemeinsame Merkmale aufweisen. In Abbildung 1.1.2.2 sind diese Gruppen einge-

¹ Tatsächlich unterscheidet sich das Ergebnis der MDS auf den ersten Blick doch für die Anwendung auf die Rohdaten und die Anwendung auf die Rangdaten. Dieser Unterschied ist auf die unterschiedliche Startkonfiguration des Verfahrens zurückzuführen. Wird die Startkonfiguration vorgegeben, erzielen beide Analysen exakt die gleichen Ergebnisse.

zeichnet und mit Bezeichnungen versehen. So lassen sich Ferrari, Lamborghini und Porsche zur Gruppe der Sportwagen zusammenfassen. VW, Opel und Toyota lassen sich trotz breiter Produktpalette der Mittelklasse zurechnen. BMW und Mercedes zählen trotz der ebenfalls breiten Produktpalette eher zur gehobenen Klasse. Abgesetzt von den Gruppen haben sich als Extrempole Rolls Royce (Luxusklasse) auf der einen Seite und Suzuki (Kleinwagen) auf der anderen Seite.

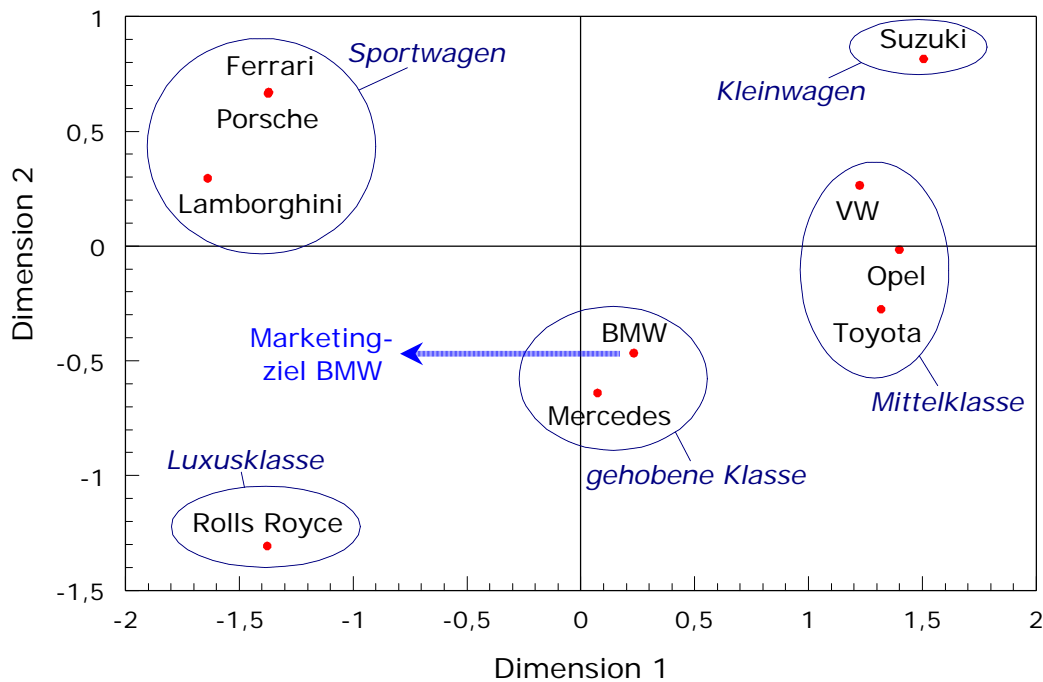


Abb. 1.1.2.2: Klassifikation der Automarken in der Konfiguration und mögliches Marketingziel aus Sicht von BMW.

Wie könnten die Ergebnisse dieser MDS vom Marketing genutzt werden? Versetzen wir uns in die Position eines Marketing-Entscheidungers bei BMW: Möglicherweise besteht das Ziel, BMW als Marke stärker von Mercedes abzusetzen - und zwar in Richtung Luxus und Sportlichkeit. Die abgebildete Konfiguration der MDS könnte als Ist-Zustand interpretiert werden.² Die Wirksamkeit der nun einzusetzenden Marketingmaßnahmen könnte durch eine Replikationsstudie überprüft werden. Sollte der Einsatz der Marketingmaßnahmen von Erfolg gekrönt sein, müsste sich die Position von BMW im Wahrnehmungsraum der Konsumenten stärker in Richtung Sportlichkeit und Luxus verschoben haben. In der Konfiguration sollte BMW in der Folge weiter links abgebildet werden (siehe Pfeil in Abbildung 1.1.2.2).

1.2 Vom Wahrnehmungsraum über Ähnlichkeitsurteile zur Konfiguration

Auf Basis der einführenden Beispiele lässt sich nun die MDS in verschiedene Komponenten zerlegen:

² Natürlich nur unter der Voraussetzung, dass die Daten der MDS repräsentativ für die Zielgruppe von BMW erhoben wurden.

1. Der Wahrnehmungsraum

Der Wahrnehmungsraum ist ein theoretisches Konstrukt und daher nicht beobachtbar. In der theoretischen Vorstellung wird der Wahrnehmungsraum betrachtet als eine räumliche Anordnung von Objekten in der Kognition der Untersuchungspersonen. Dabei wird davon ausgegangen, dass ähnliche Objekte im Wahrnehmungsraum näher beieinander liegen und unähnliche Objekte eine größere Entfernung zueinander aufweisen (wie auch immer eine solche Form von „Entfernung“ vorstellbar sein mag). Die MDS soll - einfach ausgedrückt - den Wahrnehmungsraum messen. Bereits an dieser Stelle wird deutlich, dass es keinen Sinn macht, den Wahrnehmungsraum von Objekten zu untersuchen, die den Untersuchungspersonen unbekannt sind - es gibt ihn schlichtweg nicht! Daher ist vor der eigentlichen Datenerhebung im Zuge einer MDS sicherzustellen, dass den Untersuchungspersonen alle Objekte bekannt sind.

2. Ähnlichkeitsurteile

Die Untersuchungspersonen geben Ähnlichkeitsurteile (oder Unähnlichkeitsurteile) über Paare von Objekten ab. Diese Ähnlichkeitsurteile stellen die *Rohdaten* der MDS dar. Zu beachten ist das *Skalenniveau* dieser Ähnlichkeitsurteile. In der Regel sind die Ähnlichkeitsurteile *ordinal-* oder *intervallskaliert*, seltener *ratioskaliert*.³ Auf die verschiedenen Möglichkeiten, Ähnlichkeitsurteile zu erheben, wird weiter unten eingegangen.

3. Konfiguration

In der Konfiguration sind die Objekte räumlich angeordnet. Nach Möglichkeit wird angestrebt, sich bei der räumlichen Anordnung der Objekte in der Konfiguration auf maximal zwei (in Ausnahmefällen drei) *Dimensionen* (= Skalen der MDSkalierung) zu beschränken. Die Distanzen zwischen Objekten in der Konfiguration sollen die Ähnlichkeitsurteile zu den Objektpaaren widerspiegeln. Dabei werden verschiedene *Distanzmetriken* unterschieden - die „Luftlinie“ zwischen zwei Objekten (*euklidische Metrik*) ist nur eine von verschiedenen Möglichkeiten (siehe die Ausführungen weiter unten).

Bereits in den einführenden Beispielen fällt auf, dass die absoluten Beträge der Distanzen nicht mit den absoluten Beträgen der Rohdaten übereinstimmen. Tatsächlich werden die Distanzen zwischen den Objekten in der Konfiguration auch nicht direkt den Rohdaten, sondern *Transformationen der Rohdaten*, den sogenannten *Disparitäten*, angepaßt.

Bei der Transformation der Rohdaten in Disparitäten ist das Skalenniveau der Rohdaten von Bedeutung:

- Ordinale Rohdaten → Disparitäten werden über monotone Transformation erzeugt.
- Intervallskalierte Rohdaten → Disparitäten werden über lineare Transformation erzeugt.
- Ratioskalierte Rohdaten → Disparitäten werden über proportionale⁴ Transformation erzeugt.

Tabelle 1.2.1 enthält für jede Form der Transformation ein Beispiel.

³ Zu den unterschiedlichen Skalenniveaus vgl. die Ausführungen im Skript zur Vorlesung.

⁴ Unter proportionaler Transformation wird hier die Multiplikation mit einer positiven Konstanten c verstanden ($c < 1$ entspricht einer Stauchung, $c > 1$ entspricht einer Streckung).

Tab. 1.2.1: Verschiedene Transformationen

Rohdaten	Transformationen (Disparitäten)		
	Monoton	Linear (z.B. $x^T = 2 + 3 \cdot x$)	Proportional (z.B. $x^T = 3 \cdot x$)
1	1	5	3
2	1,5	8	6
3	1,6	11	9
4	4	14	12

Konstitutiv für den Lösungsraum sind die zugrundeliegenden Dimensionen oder Achsen des Raums. Im Regelfall eines 2-dimensionalen Lösungsraums werden durch die MDS also zwei Intervallskalen erzeugt, die als Achsen den zweidimensionalen Raum aufspannen. Jedes Objekt ist durch seine Koordinaten auf diesen beiden Achsen charakterisiert, die Distanz zwischen zwei Objekten wird durch eine Distanzfunktion aus den Koordinaten der Punkte berechnet - in der Regel ist das die vertraute euklidische Metrik (vgl. Abschnitt 1.4).

Aus messtheoretischer Sicht ist hervorzuheben, dass im Zuge der ordinalen MDS eine Anhebung des Skalenniveaus stattfindet. Die in der Ähnlichkeitsmatrix enthaltene Information wird ordinal im Sinne der Indexmessung interpretiert und das Ergebnis der MDS sind die (in der Regel zwei) den Lösungsraum aufspannenden Intervallskalen. Um dies zu ermöglichen, findet bei der MDS eine Datenverdichtung statt. Bei 10 Objekten werden aus den 45 Ähnlichkeitseinstufungen für alle möglichen Paare 20 verschiedene Koordinaten für die 10 Objekte im zweidimensionalen Lösungsraum berechnet.

Als Kennziffer für die Datenverdichtung eignet sich der *Datenverdichtungskoeffizient* Q , der als Quotient der Zahl der Ähnlichkeiten⁵ und der Anzahl der Koordinaten des Lösungsraumes berechnet wird:

$$Q = \text{Zahl der Ähnlichkeiten} / \text{Zahl der Koordinaten}$$

Der Datenverdichtungskoeffizient hat in diesem Beispiel den Wert 2,5 (= 45/20) (zum Datenverdichtungskoeffizient vgl. Backhaus et al 1996, 459). Werte für $Q \geq 2$ gelten nach Backhaus et al (1996) als akzeptabel.

Es werden also (in der Regel ordinale) Informationen über Paare von Objekten erhoben, um daraus metrische (intervallskalierte) Informationen über die Objekten zu ermitteln. Nur wegen der damit einhergehenden Datenverdichtung ist eine solche Anhebung des Messniveaus überhaupt sinnvoll durchführbar. Die erreichte Güte der metrischen Repräsentation kann auf Grundlage des STRESS-Maßes beurteilt werden. Inwieweit die erhaltenen Intervallskalen als Repräsentationsmessungen angesehen werden können, soll im Ausblick (Abschnitt 7.) kurz diskutiert werden.

⁵ Gleiche Ähnlichkeiten in den Rohdaten (sog. ties) werden nur einfach gezählt. Dadurch verringert sich der Datenverdichtungskoeffizient.

1.3 Modellkonstellationen der MDS

1.3.1 Quadratische Ähnlichkeitsmatrizen (Objekt x Objekt)

In den bisherigen Ausführungen bestand die Datenbasis der MDS immer aus einer quadratischen⁶ Ähnlichkeits- oder Unähnlichkeitsmatrix. Es ist jedoch auch möglich, mehr als eine Matrix auf einmal im Zuge der MDS zu verarbeiten. Es ergeben sich so prinzipiell drei Modellkonstellationen:

1. eine Matrix und eine Konfiguration → einfache MDS (siehe 4.1 und 4.2)

Sofern nur eine Matrix von der MDS verarbeitet wird, sprechen wir im folgenden auch von einer *einfachen MDS*. Da in der Regel die Ähnlichkeitsurteile von mehr als einer Person erhoben werden, müssen diese individuellen Matrizen zunächst zu einer Matrix zusammengefasst (aggregiert) werden. Diese Matrix wird daher als *aggregierte Matrix* bezeichnet. Häufig wird zur Aggregation der Daten der Mittelwert der Ähnlichkeitsurteile für jedes Objektpaar gebildet.

2. mehr als eine Matrix und eine Konfiguration → wiederholte MDS (siehe 4.3)

Liegen mehrere Ähnlichkeitsmatrizen vor und wird auf eine Aggregation zu einer Matrix verzichtet, so wird das Verfahren als *wiederholte MDS* bezeichnet, wenn für alle individuellen Matrizen ein und dieselbe Konfiguration erzeugt wird. Es wird also unterstellt, dass alle Untersuchungspersonen den gleichen Wahrnehmungsraum besitzen. Im Zuge dieses Verfahrens kann diese Hypothese überprüft werden, denn es wird zu jeder individuellen Matrix ein Wert berechnet (STRESS), der die Güte der Anpassung an die Konfiguration ausdrückt.

3. mehr als eine Matrix und mehr als eine Konfiguration → INDSCAL (siehe 4.4)

Das *INDSCAL*-Verfahren (Skalierung individueller Differenzen) basiert ebenfalls auf mehreren Ähnlichkeitsmatrizen, erzeugt jedoch für jede Matrix eine andere Konfiguration. Allerdings unterscheiden sich die individuellen Konfigurationen nicht vollständig voneinander, sondern sie basieren auf einer einzigen Konfiguration, die lediglich durch unterschiedliche Gewichtung der Dimensionen (stauchen oder strecken) an die individuellen Matrizen angepaßt wird. Auch dieses Verfahren erzeugt für jede Matrix einen individuellen STRESS-Wert.

Den drei genannten Modellkonstellationen liegen also unterschiedliche Hypothesen und unterschiedlich starke Hypothesentests bezüglich der Ausgangsdatenlage zugrunde:

Einfache MDS: Es wird davon ausgegangen, dass die individuellen Matrizen alle auf Basis des gleichen Wahrnehmungsraumes erzeugt wurden (von Zufallsfehlern abgesehen) und daher zu einer aggregierten Matrix zusammengefaßt werden können. Diese Hypothese wird nicht weiter überprüft!

⁶ Zusätzlich sind die hier behandelten quadratischen Matrizen auch symmetrisch, das heißt die Matrix ist entlang der Diagonale gespiegelt. Es ist mit SPSS auch möglich, quadratische asymmetrische Matrizen multidimensional zu skalieren (vgl. auch die Ankerpunktmethode als Erhebungsform in Abschnitt 2.2.3). Dieser Spezialfall soll hier jedoch nicht weiter behandelt werden.

Wiederholte MDS: Es wird davon ausgegangen, dass die individuellen Matrizen alle auf Basis des gleichen Wahrnehmungsraumes erzeugt wurden. Diese Hypothese wird im Zuge des Verfahrens überprüft!

INDSCAL: Es wird davon ausgegangen, dass die individuellen Matrizen alle auf Basis ähnlicher Wahrnehmungsräume erzeugt wurden, die sich nur durch unterschiedlich gewichtete Dimensionen unterscheiden. Diese Hypothese wird im Zuge des Verfahrens überprüft!

Da sowohl bei INDSCAL als auch bei der wiederholten MDS die Unterschiede zwischen den Matrizen einem Test unterzogen werden, eignen sich diese beiden Verfahren auch, um Veränderungen von Wahrnehmungsräumen (z.B. im Zuge eines Experiments) zu überprüfen.

1.3.2 Rechteckige Ähnlichkeitsmatrizen (Objekt x Subjekt) (siehe 5.)

Neben der Erhebung von Ähnlichkeiten zwischen Objekten ist es auch möglich, die Nähe oder Distanz von Subjekten (Befragten) zu einer Reihe von Objekten zu erheben. So könnten z.B. Versuchspersonen aufgefordert werden, eine Reihe von Zeitschriften nach ihrer subjektiven Präferenz zu ordnen. Die resultierende Datenmatrix könnte so aussehen (Tabelle 1.3.2.1):

Tab. 1.3.2.1: Rangreihung von Zeitschriften nach subjektiver Präferenz (1 = am meisten präferiert) von vier Versuchspersonen (fiktive Daten).

VPN	Bild	Spiegel	Bunte	Stern	Motorwelt	Brigitte	Hörzu
1	1	7	3	5	6	4	2
2	7	1	3	4	6	2	5
3	2	6	5	4	1	7	3
4	7	1	4	2	3	5	6

Auch diese Ähnlichkeitsmatrix kann multidimensional skaliert werden. Dabei ist jedoch zu beachten, dass die Daten dieser Matrix nur zeilenweise miteinander verglichen werden können. In diesem Zusammenhang wird auch von *zeilenkonditionalen Daten* gesprochen (im Gegensatz zu matrixkonditionalen Daten in den quadratischen Matrizen von Abschnitt 1.3.1).

Die multidimensionale Skalierung solcher zeilenkonditionalen Objekt x Subjekt - Matrizen wird als *multidimensionale Entfaltung* oder *MDU* (multidimensional unfolding) bezeichnet. Die besondere Eigenart der Konfiguration einer MDU liegt darin, dass Subjekte und Objekte zusammen in ein und demselben Raum dargestellt werden. Die Konfiguration einer MDU der Daten in Tabelle 1.3.2.1 ist in Abbildung 1.3.2.1 dargestellt.

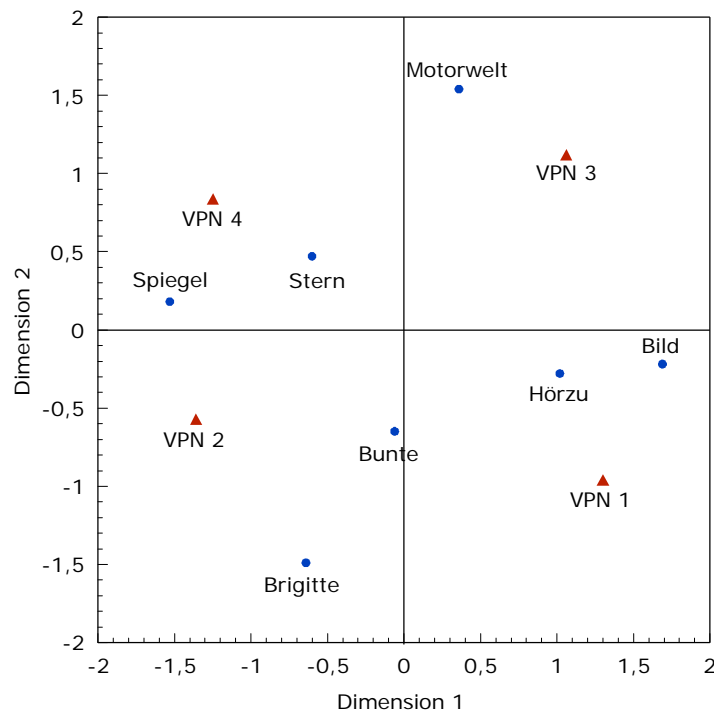


Abb. 1.3.2.1: Multidimensionale Entfaltung (MDU) von Präferenzen für Zeitschriften.

Der Interpretation der Konfiguration einer MDU liegt ein spezifisches Modell zugrunde, das sogenannte *Idealpunktmodell*. Die Subjekte (VPN 1 bis 4) werden in der Konfiguration in ihren jeweiligen Idealpunkten abgebildet. Bezogen auf das Beispiel der Zeitschriftenpräferenzen entspricht der Idealpunkt also der idealen Zeitschrift aus Sicht des Subjekts. Je weiter entfernt ein Objekt von einem Subjekt abgebildet wird, desto weiter entfernt vom Idealpunkt wurde dieses Objekt vom Subjekt wahrgenommen und um so weniger wurde es präferiert.

In der nächsten Abbildung 1.3.2.2 sind die Distanzen aller Objekte zu Subjekt VPN 4 eingezeichnet. Ein Vergleich der Reihenfolge dieser Distanzen mit den Rohdaten in Tabelle 1.3.2.1 zeigt, dass die erhobene Präferenzordnung von VPN 4 durch die Distanzen in der Konfiguration perfekt widerspiegelt wird.

Abbildung 1.3.2.2 veranschaulicht auch, warum bei dem hier angewendeten Verfahren von einer multidimensionalen Entfaltung gesprochen wird: Die Präferenzurteile von VPN 4 aus Tabelle 1.3.2.1 für jedes Objekt kann man sich als Mikadostäbchen unterschiedlicher Länge vorstellen. Kurze Mikados stellen kleine Zahlen (=Nähe) dar, lange Mikados stellen hohe Zahlen (=Distanz) dar. Diese Mikados werden in der Konfiguration im Idealpunkt von VPN 4 senkrecht aufgestellt und dann zu den Seiten hin „entfaltet“. Dabei können die Mikados frei um den Idealpunkt rotiert werden. Dieser Vorgang wird auch für die weiteren Subjekte durchgeführt - und zwar so lange, bis alle Spitzen der jeweils zum gleichen Objekt gehörenden Mikados in einem Punkt zusammentreffen.

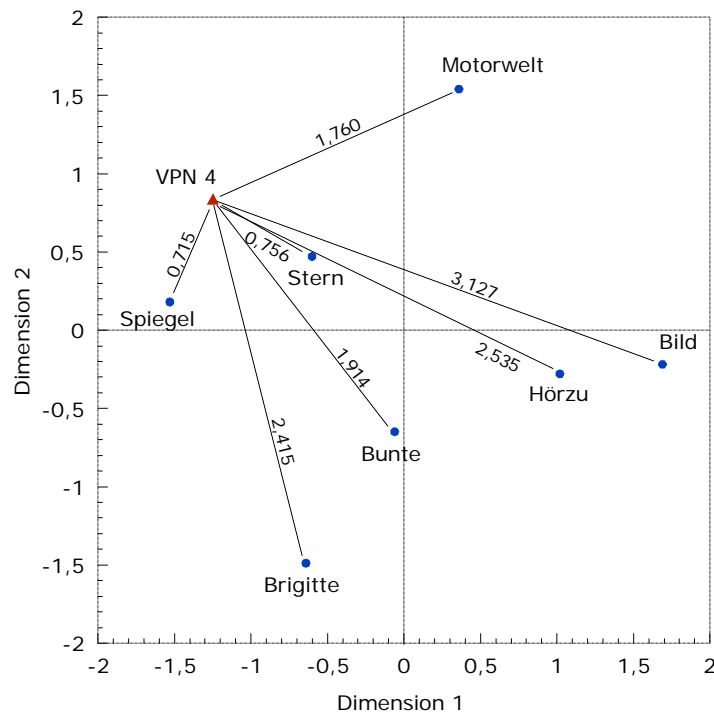


Abb. 1.3.2.2: Distanz von VPN 4 zu den Objekten in der Konfiguration.

Die MDU lässt sich genauso wie die MDS ebenfalls auf mehr als eine Rechtecksmatrix anwenden. Es gelten hier also prinzipiell die gleichen Möglichkeiten der Datenkonstellation, wie sie im vorangegangenen Abschnitt 1.3.1 aufgezeigt wurden.

1.4 Distanzmetriken der MDS

Wie bereits oben angedeutet, ist die „Luftlinie“ nicht die einzige Möglichkeit, die Distanz zwischen zwei Objekten in der Konfiguration auszudrücken. Es gibt vielmehr eine Reihe verschiedener Distanzmetriken, die jedoch alle auf dieselbe allgemeine Formel zurückzuführen sind.

Zunächst wenden wir uns jedoch der intuitiv naheliegenden Form der Distanz zwischen zwei Punkten zu: der euklidischen Distanz.

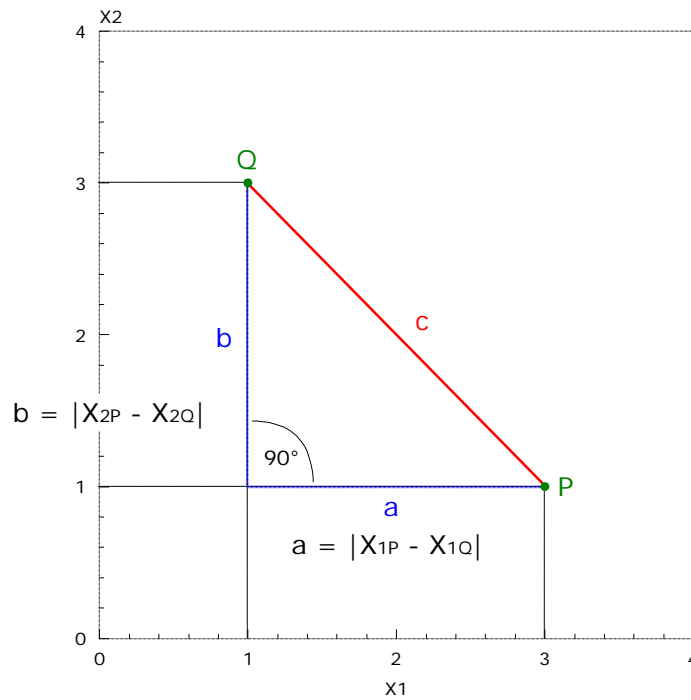


Abb. 1.4.1: Die Distanz im euklidischen Raum.

Abbildung 1.4.1 enthält zwei Punkte P und Q, die in einem zweidimensionalen Raum liegen, der von den Dimensionen X_1 und X_2 aufgespannt wird. Die euklidische Distanz („Luftlinie“) zwischen diesen beiden Punkten ist als Strecke c bezeichnet. Um die Länge der Strecke c zu bestimmen, zeichnen wir zwei weitere Strecken a und b ein, die im rechten Winkel aufeinandertreffen und mit c ein Dreieck bilden.

Der Einfachheit halber verläuft a parallel zu Dimension X_1 und b parallel zu Dimension X_2 , so dass sich die Länge der Strecken a und b leicht ablesen lässt:

$$a = |X_{1P} - X_{1Q}| \quad [= |3 - 1| = 2]$$

$$b = |X_{2P} - X_{2Q}| \quad [= |1 - 3| = 2]$$

Die Länge der Strecke c berechnet sich nun einfach nach Pythagoras:

$$c^2 = a^2 + b^2$$

$$c = \sqrt{a^2 + b^2}$$

$$c = \sqrt{(X_{1P} - X_{1Q})^2 + (X_{2P} - X_{2Q})^2} \quad [= \sqrt{(3 - 1)^2 + (1 - 3)^2} = \sqrt{8} = 2,83]$$

Wenn die Distanzen der Objekte in der Konfiguration als euklidische Distanzen berechnet werden, so gilt in diesem Raum die *euklidische Metrik*. Die allgemeine Formel der euklidischen Metrik für r -dimensionale Räume lautet:

$$\text{Euklidische Metrik: } d_{ij} = \left(\sum_{k=1}^r (x_{ik} - x_{jk})^2 \right)^{\frac{1}{2}}, \text{ mit}$$

r = Anzahl der Dimensionen
 i, j = Objekte.

Die Verwendung der euklidischen Metrik in der Konfiguration hat zwei wesentliche Vorteile:

1. Die Darstellung der Konfiguration kann intuitiv interpretiert werden, weil die Distanzen den Luftlinien zwischen den Objekten entsprechen.
2. Die Konfiguration kann um das Koordinatenkreuz gedreht oder an den Achsen gespiegelt werden, ohne dass sich die Distanzen zwischen den Objekten ändern. Die Lösung ist also *rotationsinvariant* (siehe S. 17). Das bedeutet insbesondere, dass in der euklidischen Metrik zusätzliche Geraden (Vektoren) durch den Ursprung gelegt werden können, um die Konfiguration besser interpretieren zu können (vgl. Abschnitt 6).

Vermutlich aus diesen Gründen erlaubt SPSS ausschließlich die Verwendung der euklidischen Metrik in der Konfiguration (mit Ausnahme von INDSCAL(!)).

Aus theoretischen Gründen könnte es jedoch sinnvoll sein, andere Metriken in der Konfiguration zu verwenden. Wir erinnern uns, dass die Konfiguration als Messung des Wahrnehmungsraumes interpretiert wird.

In einer (fiktiven) Untersuchung werden New Yorker Taxifahrer nach der Fahrzeit zwischen verschiedenen Hotels gefragt. Tabelle 1.4.1 enthält die durchschnittlichen Fahrzeiten in Minuten.

Tab. 1.4.1: Geschätzte Fahrzeit in Minuten zwischen Hotels in New York

	Milford Plaza	Shoreham	Belleclaire	Novotel	Pennsylvania	Lucerne	Days Inn	Gershwin
Milford Plaza	0							
Shoreham	8	0						
Belleclaire	12	14	0					
Novotel	4	4	12	0				
Pennsylvania	6	11	19	7	0			
Lucerne	13	15	1	12	19	0		
Days Inn	2	6	13	2	5	14	0	
Gershwin	11	9	23	11	4	24	10	0

Um diese Entfernungsangaben multidimensional zu skalieren und als Konfiguration einen Lageplan der Hotels zu erhalten, wäre die euklidische Metrik nicht angemessen. Der Grund liegt in der Art und Weise, wie der Wahrnehmungsraum der Taxifahrer bezogen auf die Fahrzeiten strukturiert ist.

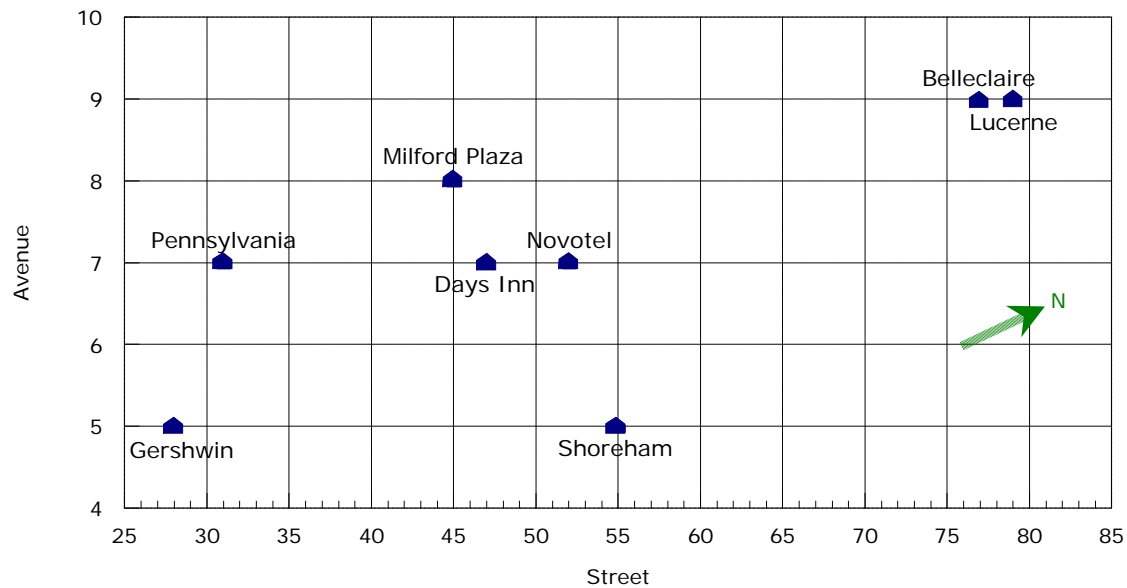


Abb. 1.4.2: Lageplan New Yorker Hotels

Die Straßen von New York (Manhattan) teilen die Stadt in Blöcke und sind durchnummeriert. Von Osten nach Westen verlaufen „Streets“, von Norden nach Süden ist das Stadtgebiet von „Avenues“ durchzogen. Abbildung 1.4.2 zeigt einen Lageplan der genannten Hotels. Die Taxifahrer können nur in rechten Winkeln von einem Hotel zum nächsten gelangen und schätzen die Fahrzeit daher nicht auf Basis der Luftlinie, sondern auf Basis der dazwischen liegenden „Streets“ und „Avenues“. Die Fahrzeit von einer Avenue zur nächsten beträgt in etwa fünfmal soviel wie die Fahrzeit zwischen zwei Straßen.

Möchte ein Taxi zum Beispiel vom Pennsylvania Hotel zum Milford Plaza, so fährt es die 7th Avenue in Richtung Norden⁷ bis zur 45th Street und biegt dann nach Westen ab. Das Hotel liegt an der Ecke 8th Avenue/45th Street. Dazu braucht es ca. 6 Minuten. Die Fahrzeit vom Pennsylvania zum Novotel ist mit etwa 7 Minuten nur geringfügig länger, obwohl die Luftlinie vom Pennsylvania zum Novotel deutlich länger ist, als zum Milford Plaza.

Die Metrik, nach der die Distanzen dieses Raumes berechnet werden, wird *City-Block-Metrik* genannt. City-Block-Distanzen für r -dimensionale Räume werden nach der folgenden Formel berechnet:

$$\text{City-Block-Metrik: } d_{ij} = \sum_{k=1}^r |x_{ik} - x_{jk}|.$$

Die City-Block-Distanz zwischen den Punkten P und Q in Abbildung 1.4.1 wird beispielsweise berechnet als $d_{pq} = a + b = 2 + 2 = 4$.

City-Block-Metrik und euklidische Metrik stellen zwei Spezialfälle der allgemeinen *Minkowski-Metrik* dar. Minkowski-Distanzen werden nach der folgenden allgemeinen Formel berechnet:

$$\text{Minkowski-Metrik: } d_{ij}^p = \left(\sum_{k=1}^r (x_{ik} - x_{jk})^p \right)^{\frac{1}{p}}.$$

⁷ Die möglichen Fahrrichtungen sind in diesem Beispiel nicht berücksichtigt.

Wird der Parameter p der Minkowski-Metrik gleich 1 gesetzt, so erhalten wir die City-Block-Metrik. Setzen wir $p = 2$, so ergibt sich die euklidische Metrik. Für $p = \infty$ erhält man die sog. Dominanzmetrik (auch Supremums-Metrik genannt)

$$\text{Dominanzmetrik: } d_{ij}^p = \max_k |x_{ik} - x_{jk}|$$

Wie aus der Urteilspsychologie bekannt, hängt die Art, wie Ähnlichkeitsurteile zustande kommen, von der Schwierigkeit der abzugebenden Urteile ab. Je schwerer die Ähnlichkeitsurteile für die Objektpaare abzugeben sind, desto höher ist der für das Urteilsverhalten ausschlaggebende Koeffizient p der allgemeinen Minkowski-Metrik. Bei schwierigen Urteilen wird das deutlich differenzierende Merkmal am stärksten gewichtet ($p \rightarrow \infty$), bei leichten Urteilen erhalten alle relevanten Merkmale ungefähr das gleiche Gewicht ($p \rightarrow 1$).

Ein weiterer Spezialfall einer Distanz-Metrik wird im oben angesprochenen INDSCAL-Verfahren angewendet. Beim INDSCAL-Verfahren wird zu jeder Ähnlichkeitsmatrix eine eigene Konfiguration erstellt. Wie bereits erläutert, gehen diese individuellen Konfigurationen auf eine gemeinsame Konfiguration zurück, deren Dimensionen individuell gewichtet (gestaucht oder gestreckt) werden. Die Formel dieser Distanz-Metrik entspricht bei SPSS weitestgehend der euklidischen Distanz⁸, mit dem Unterschied eines individuellen Gewichtungsfaktors für jede Dimension:

$$\text{INDSCAL-Metrik: } d_{mij} = \left(\sum_{k=1}^r w_{mk} (x_{ik} - x_{jk})^2 \right)^{\frac{1}{2}}, \text{ mit}$$

w_m = Gewichtungsfaktor

m = Rohdatenmatrix

i, j = Objekte

r = Dimensionen.

Zu beachten ist, dass die oben genannten positiven Eigenschaften der euklidischen Distanz (siehe S. 14) für keine andere Distanz-Metrik gelten:

- Die Interpretation der Konfiguration wird schwieriger, weil die Luftlinie (und damit der intuitive optische Eindruck eines Abstands) zwischen zwei Objekten nicht mehr der Distanz entspricht.
- Die Lage des Koordinatenkreuzes ist festgelegt. Rotationen des Koordinatenkreuzes führen zu Änderungen der Distanzen. Lediglich das Klappen der Konfiguration entlang der Dimensionsachsen ist erlaubt.

Insbesondere die zweite der beiden Eigenschaften der euklidischen Distanz wird deutlich, wenn die Einheits-„kreise“ der Distanz-Metriken betrachtet werden (siehe Abbildung 1.4.3). Nur der Einheits-„kreis“ der euklidischen Metrik ist ein echter Kreis, der bei der Rotation in sich selbst übergeht.

⁸ Prinzipiell ließe sich auch jede andere Minkowski-Metrik durch die individuelle Gewichtung im Zuge des INDSCAL-Verfahrens erweitern.

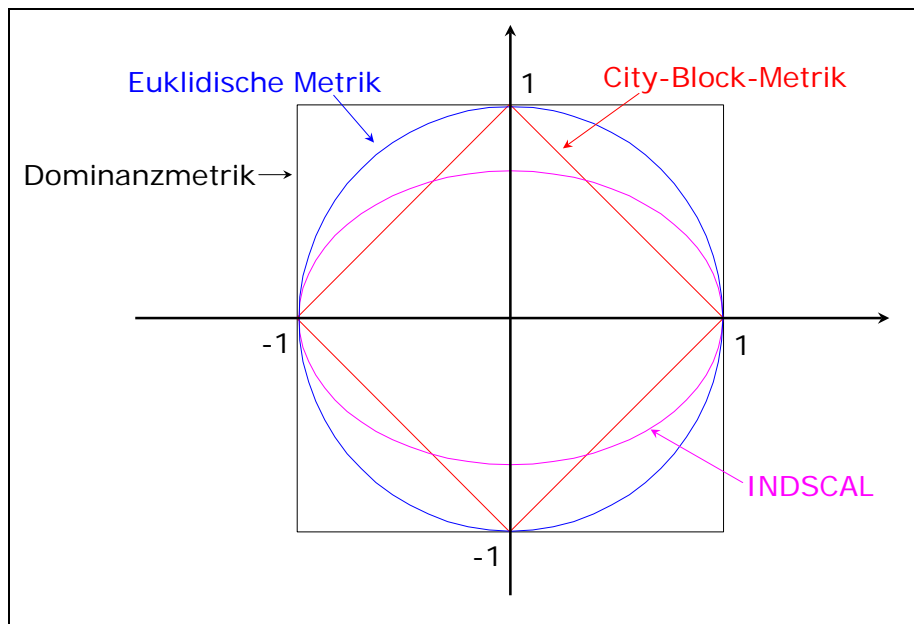


Abb. 1.4.3: Einheits-„kreise“ der euklidischen Metrik, der City-Block-Metrik und des INDSCAL-Verfahrens. Alle Punkte auf dem Einheitskreis haben eine Distanz von 1 zum Ursprung.

Bei allen anderen Metriken verändert der Einheits-„kreis“ seine Lage, wenn das Koordinatenkreuz in Abbildung 1.4.3 rotiert wird. Die Konfiguration mit euklidischer Metrik wird daher auch *rotationsinvariant* genannt.

2. Auswahl der Objekte und Erhebung der Ähnlichkeiten

2.1 Auswahl der Objekte

Bei der Auswahl der Objekte für eine MDS sind verschiedene Faktoren zu berücksichtigen:

- Die Objekte müssen den Befragten Personen bekannt sein, da sonst die Ähnlichkeiten zwischen den Objekten nicht valide beurteilt werden können.
- Die Anzahl der Ähnlichkeiten zwischen den Objekten muß die Anzahl der Koordinaten im Lösungsraum deutlich übersteigen (vgl. den Datenverdichtungskoeffizient, S. 8).
- Andererseits sollte die Zahl der Objekte nicht zu groß sein, da die Befragten sonst durch die Anzahl der abzugebenden Ähnlichkeitsurteile überfordert würden.

Da der angesprochene trade off zwischen erforderlicher Objektzahl aus Gründen der Datenanalyse und maximaler Anzahl an Objekten aus Gründen der Erhebungssituation nicht immer einfach einzuschätzen ist, sollten vor der Hauptuntersuchung auf jeden Fall Pre-Tests (evt. mit variabler Zahl von Objekten) durchgeführt werden. Möglich ist auch in einer Voruntersuchung die Bekanntheit (evt. auch Marktanteile) einer größeren Menge interessierender Objekte zu erheben, um dann in der MDS nur die bekanntesten Objekte beurteilen zu lassen.

2.2 Erhebung von Ähnlichkeitsurteilen

Die Datenbasis der MDS besteht aus Ähnlichkeitsurteilen, die in Form von Matrizen vorliegen. Diese Ähnlichkeitsurteile können auf verschiedene Art und Weise erhoben werden. Das Erhebungsverfahren hat Einfluss auf die Form der Matrix und die Konditionalität der Daten.

Die spezifische *Konditionalität* der Daten beantwortet die Frage, welche Zahlen innerhalb einer Datenmatrix miteinander vergleichbar sind und infolge dessen eine valide Information darstellen, die bei der Berechnung der MDS-Lösung berücksichtigt werden darf. Wir unterscheiden zwei Fälle:

Matrixkonditionalität: Alle Zahlen innerhalb einer Matrix sind vergleichbar. Zahlen verschiedener Matrizen sind nicht miteinander vergleichbar.

Zeilenkonditionalität: Nur Zahlen innerhalb einer Zeile sind miteinander vergleichbar, Zahlen aus verschiedenen Zeilen dürfen nicht gegenübergestellt werden.

In den nächsten Abschnitten werden die gängigen Erhebungsverfahren vorgestellt.

2.2.1 Ratingverfahren

Die Objekte werden paarweise miteinander verglichen und ihre Ähnlichkeit wird als Wert einer vorgegebenen Ratingskala ausgedrückt:

Wie ähnlich sind sich die folgenden Objektpaare auf einer Skala von 1 bis 10 (1 bedeutet sehr ähnlich, 10 bedeutet überhaupt nicht ähnlich):

	Obj 1	Obj 2	Obj 3	Obj 4
Obj 1	X			
Obj 2	2	X		
Obj 3	3	4	X	
Obj 4	1	3	5	X

Die so erhaltene Ähnlichkeitsmatrix ist **quadratisch** und **symmetrisch**, die Daten sind **matrixkonditional**.

Zwischen den Ähnlichkeitsurteilen können (und müssen ab einer bestimmten Zahl von Objektpaaren) sog. **Ties** auftreten; das bedeutet, daß zwei Objektpaare den gleichen Ähnlichkeitsgrad zugewiesen bekommen (z.B. 1-3 und 2-4).

2.2.2 Rangreihung

Es werden alle möglichen Paarungen von Objekten gebildet (und z.B. auf Karteikarten festgehalten). Diese Objektpaare werden in eine Rangreihe gebracht (in der Regel durch Sortieren).

Objektpaar	Rangplatz
Obj 1 - Obj 4	1
Obj 1 - Obj 2	2
Obj 2 - Obj 4	3
Obj 1 - Obj 3	4
Obj 2 - Obj 3	5
Obj 3 - Obj 4	6

Diese Methode ergibt folgende Ähnlichkeitsmatrix:

	Obj 1	Obj 2	Obj 3	Obj 4
Obj 1	X			
Obj 2	2	X		
Obj 3	4	5	X	
Obj 4	1	3	6	X

Die so erhaltene Ähnlichkeitsmatrix ist **quadratisch** und **symmetrisch**, die Daten sind **matrixkonditional**. Zwischen den Ähnlichkeitsurteilen können **keine Ties** auftreten. Allerdings

wird eine Sortieraufgabe mit zunehmender Zahl von Objektpaaren immer schwieriger, da die Zahl der Paare quadratisch (genauer $n \cdot (n-1)/2$) wächst.⁹

2.2.3 Ankerpunktmethode

Jedes Objekt dient gegenüber jeweils allen anderen als Vergleichsobjekt (= Ankerpunkt). Die Befragten geben zeilenweise die Reihenfolge der Ähnlichkeiten der Objekte zum Vergleichsobjekt an:

Bsp.: Geben Sie für jedes Objekt (zeilenweise) die Reihenfolge der anderen drei Objekte zu diesem Objekt an.

	Obj 1	Obj 2	Obj 3	Obj 4
Obj 1	0	2	3	1
Obj 2	1	0	3	2
Obj 3	1	2	0	3
Obj 4	1	2	3	0

Die so erhaltene Ähnlichkeitsmatrix ist **quadratisch** und **asymmetrisch**, die Daten sind **zeilenkonditional**. Das bedeutet, dass die vergebenen Rangplätze nur zeilenweise miteinander verglichen werden dürfen. Das entsprechend eingestellte MDS-Programm berücksichtigt daher auch nur die zeilenspezifische Information der Datenmatrix.

Die Ankerpunktmethode ist besonders dann geeignet, wenn einerseits viele Objekte verglichen werden sollen, aber andererseits mittels der Ratingmethode zu viele Ties erwartet werden.

2.2.4 Datenerhebung für die MDU

Für eine MDU werden keine Ähnlichkeitsdaten zwischen Objekten erhoben, sondern die Subjekte (z.B. die befragten Personen) geben Urteile zu den Objekten ab. Diese Urteile (z.B. Präferenzen) werden als Nähe bzw. Distanz der Subjekte zu den Objekten aufgefaßt. Die Subjekte können daher auch als Ankerpunkte im Sinne von 2.2.3 verstanden werden.

Als Verfahren eignen sich hier die Ratingmethode oder die Rangreihung (siehe oben). Die Datenbasis der MDU besteht aus einer **rechteckigen zeilenkonditionalen Matrix** (siehe Tabelle 1.3.2.1 auf Seite 10).

⁹ Wollte man die Rangreihe im Sinne einer Repräsentationsmessung durch Paarvergleich erstellen (vgl. Skript zur Vorlesung, Kapitel 5.1), so müssten Paare von Paaren erzeugt werden und (zumindest teilweise) beurteilt werden. Bei 10 Objekten wären dies aus 45 Paaren dementsprechend 990 ($= 45 \cdot (45-1)/2$) Paare von Paaren.

3. Der technische Ablauf der MDS: Von Ähnlichkeiten über Disparitäten zu Distanzen

In diesem Abschnitt werden die technischen Hintergründe des iterativen Algorithmus der MDS erläutert. Dieser Teil ist eher abstrakt gehalten und kann bei rein anwendungsorientiertem Interesse an der MDS weitestgehend übersprungen werden. Eine Ausnahme bilden hier allerdings die Ausführungen zu den Gütekriterien der MDS (STRESS, RSQ), die für die Ergebnisinterpretation zentral sind.

Bei der Bestimmung der Konfiguration verwendet die MDS einen iterativen Prozess. Die Grundidee dieses Prozesses ist relativ simpel:

Alle Objekte werden zunächst mehr oder weniger willkürlich im Raum angeordnet. Im nächsten Schritt werden die Distanzen zwischen den Objekten mit den Ähnlichkeiten verglichen (wobei das Skalenniveau der Ähnlichkeiten berücksichtigt wird). Wenn nun zwei Objekte im Verhältnis zu ihrer Ähnlichkeit zu weit auseinanderliegen, werden sie aufeinander zu geschoben. Sollten zwei eher unähnliche Objekte zu nahe bei einander liegen, werden sie voneinander weg bewegt. Dieser Vorgang wird so lange fortgesetzt, bis die Konfiguration der Objekte die erhobenen Ähnlichkeiten zufriedenstellend widerspiegelt.

Damit der hier geschilderte iterative Prozess einwandfrei funktioniert, müssen eine Reihe von Entscheidungen und Vorkehrungen getroffen werden:

1. Die Dimensionalität des Raumes muss vor Beginn des Iterationsprozesses festgelegt werden. Nur so ist klagestellt, in welchen Dimensionen die Objekte bewegt werden können: eindimensional entlang einer Gerade, zweidimensional in einer Ebene oder (seltener) dreidimensional im „Raum“. Lösungen mit mehr als drei Dimensionen werden in der Regel nicht angestrebt.
2. Die Startkonfiguration muss festgelegt werden. SPSS erzeugt diese Startkonfiguration per Voreinstellung automatisch.¹⁰ Die Startkonfiguration kann jedoch auch vorgegeben werden (z.B. als Ergebnis einer vorangegangenen Analyse).
3. Die Distanz-Metrik muss festgelegt werden, um die Distanzen zu berechnen. Wie schon angesprochen, stehen in SPSS nur die euklidische Metrik und das INDSCAL-Verfahren (angewendet auf die euklidische Metrik) zur Verfügung.
4. Wie bereits oben erläutert, werden die Distanzen in der Konfiguration nicht direkt mit den Rohdaten verglichen, sondern mit transformierten Rohdaten, den sogenannten Disparitäten. Die Ähnlichkeiten müssen daher in Disparitäten transformiert werden. Dazu muss das Skalenniveau der Rohdaten festgelegt werden, um zu klären, welche Transformationen erlaubt sind.
5. Es muss ein Verfahren implementiert werden, auf dessen Grundlage die Objekte in der Konfiguration verschoben werden.
6. Dazu muß ein Kriterium für die Güte der Anpassung der jeweiligen Lösung aus den transformierten Rohdaten (Disparitäten) gefunden werden. Dies geschieht durch sog. STRESS-

¹⁰ Der Algorithmus, den SPSS zur Festlegung der Startkonfiguration verwendet (Young-Householder-Torgerson-Prozedur), soll hier nicht weiter erläutert werden. Näheres dazu in SPSS (1991): *SPSS. Statistical Algorithms. 2nd Edition*. Chicago.

Maße, die in irgendeiner Form die standardisierte Abweichung der Lösungsdistanzen von den Disparitäten berechnen. Ziel des Verfahrens ist eine optimale Anpassung der MDS-Lösung an die Rohdaten, also ein möglichst geringer STRESS.

7. Es müssen Kriterien für den Abbruch des Iterationsprozess festgelegt werden. Dabei werden üblicherweise drei Informationen berücksichtigt:

- Die Anzahl der Schleifen, die der Iterationsprozeß durchlaufen hat. Hier kann ein Maximalwert festgelegt werden.
- Der Wert, den das STRESS-Maß annimmt. Hier kann ebenfalls ein Maximalwert festgelegt werden, der im Zuge der Minimierung des STRESS unterschritten werden muss.
- Die Veränderung der Konfiguration zwischen zwei aufeinanderfolgenden Iterationsschritten. Dazu wird die Veränderung des STRESS herangezogen. Auch hier wird die Iteration abgebrochen, wenn ein vorgegebener Maximalwert unterschritten wird.

Die folgende Abbildung 3.1 zeigt den Iterationsprozess der MDS im Überblick:

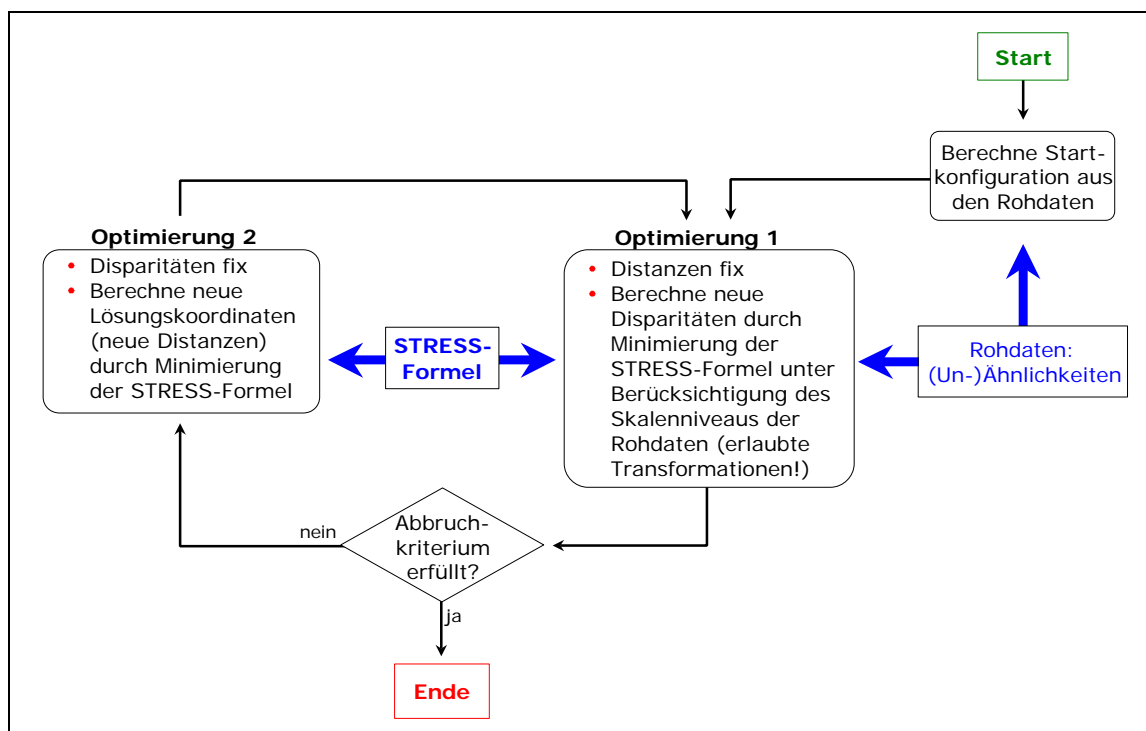


Abb. 3.1: Iterationsprozeß der MDS.

Wie in Abbildung 3.1 dargestellt, erfolgt die Anpassung der Konfiguration an die Ähnlichkeitsdaten in zwei Optimierungsschritten:¹¹

Optimierung 1: Zunächst wird geprüft, wie gut die Konfiguration an die Ausgangsdaten angepasst ist. Dabei werden die Rohdaten je nach Skalenniveau einer bestimmten Transformation unterzogen. Die daraus resultierenden Disparitäten werden mit den Distanzen verglichen. Der Unterschied zwischen Disparitäten und Distanzen wird als STRESS ausgegeben. Wenn der STRESS der Konfiguration

¹¹ Das in SPSS integrierte MDS-Programm ALSCAL verwendet dazu einen alternierenden Kleinstquadrate-Algorithmus (alternating least squares algorithm [ALS]), der von Takane, Young und de Leeuw (1977) entwickelt wurde.

klein genug ist oder sich nicht mehr wesentlich verändert hat, wird nach Optimierungsschritt 1 die Iteration abgebrochen und das Ergebnis der MDS ausgegeben.¹²

Optimierung 2: In diesem Optimierungsschritt werden die Objekte in der Konfiguration verschoben. *Diese Verschiebung erfolgt wiederum auf Grundlage des Unterschiedes zwischen Disparitäten und Distanzen.* Durch die resultierende verbesserte Anpassung der Distanzen an die Disparitäten wird der STRESS erneut verringert. Auf Optimierung 2 folgt wieder Optimierung 1, wodurch sich der STRESS in der Regel weiter verringert.

Die Einzelheiten der beiden Optimierungsschritte werden nun im Detail erläutert.

3.1 Optimierung 1: Transformation der Rohdaten

Die Transformation der Rohdaten in Disparitäten soll am Beispiel von ordinalen Rohdaten erläutert werden, das heißt, wir gehen von dem gebräuchlichen Fall einer ordinalen MDS aus. Angenommen, die Unähnlichkeit von vier Objekten A, B, C und D sei wie folgt erhoben worden (siehe Tabelle 3.1.1):¹³

Tab. 3.1.1: Rohdatenmatrix (ordinale Unähnlichkeiten)

	A	B	C	D
A	0			
B	1	0		
C	3	2	0	
D	4	6	5	0

Die Objekte A und B werden laut Tabelle 3.1.1 als am ähnlichsten zueinander eingestuft. Die Objekte B und D sind sich am unähnlichsten.

Die Optimierung beginnt mit der folgenden Startkonfiguration¹⁴ (Abbildung 3.1.1):

¹² Weiterhin wird die Iteration abgebrochen, wenn ein vorher festgelegtes Maximum an Iterationsschritten durchgeführt wurde.

¹³ Wären hier etwa die Rohdaten als Ähnlichkeitsratings erhoben worden, würden sie im Falle einer ordinalen MDS in Rangplätze umgewandelt, bevor das Optimierungsverfahren in Gang gesetzt wird.

¹⁴ Wie diese Startkonfiguration zustande gekommen ist, sei an dieser Stelle unerheblich.

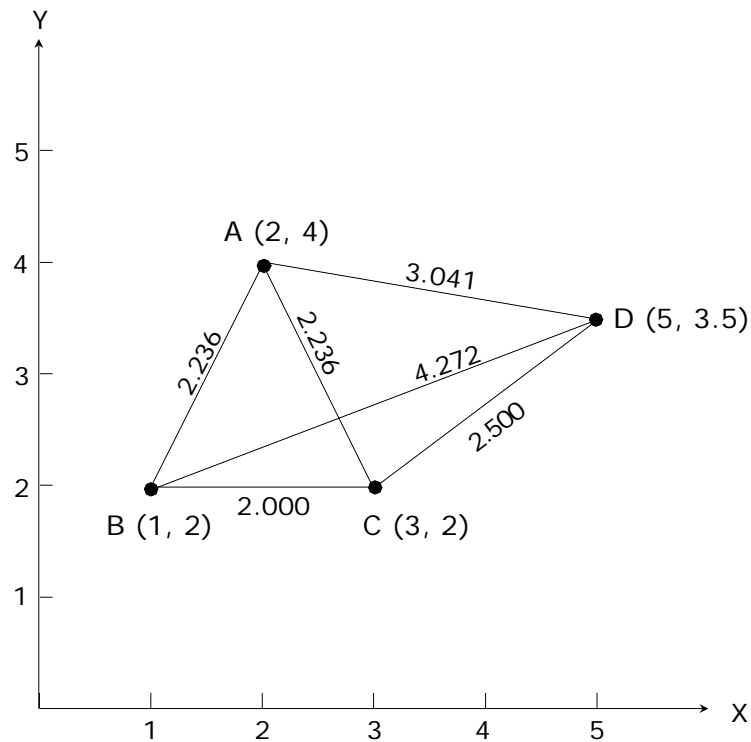


Abb. 3.1.1: Startkonfiguration einer MDS (X- und Y-Koordinaten stehen in Klammern). Die Reihenfolge der Distanzen zwischen den Objekten entspricht nicht der Reihenfolge der erhobenen Ähnlichkeiten zwischen den Objekten.

Werden die Rohdaten den Distanzen gegenüber gestellt, so ist festzustellen, dass die Reihenfolge der Distanzen nicht der Reihenfolge der Rohdaten entspricht (Tabelle 3.1.2):

Tab. 3.1.2: Distanzen in der Startkonfiguration

Objekte	Rohdaten	Distanzen ¹⁵	Reihenfolge der Distanzen
AB	1	2,236	2
BC	2	2,000	1
AC	3	2,236	2
AD	4	3,041	5
CD	5	2,500	4
BD	6	4,272	6

Weiterhin wird in Tabelle 3.1.2 deutlich, dass die Rohdaten in ihren absoluten Beträgen erheblich von den Distanzen abweichen. Da die Rohdaten ordinales Skalenniveau aufweisen, werden sie nun einer (schwach-) monotonen Transformation unterzogen – genau dies ist die Bedingung der ordinalen MDS. Als schwach monotone Transformation gilt jede Transformation der Rohdaten, welche die Reihenfolge der Daten nicht verändert. Dabei wird in Kauf genommen, dass benachbarte Rohdaten den gleichen Wert zugewiesen bekommen (daher schwach-monoton(!)). Die folgende Abbildung 3.1.2 (Shepard-Diagramm) verdeutlicht diesen Vorgang:

¹⁵ Es handelt sich um euklidische Distanzen, die nach der Formel in Abschnitt 1.4 berechnet wurden.

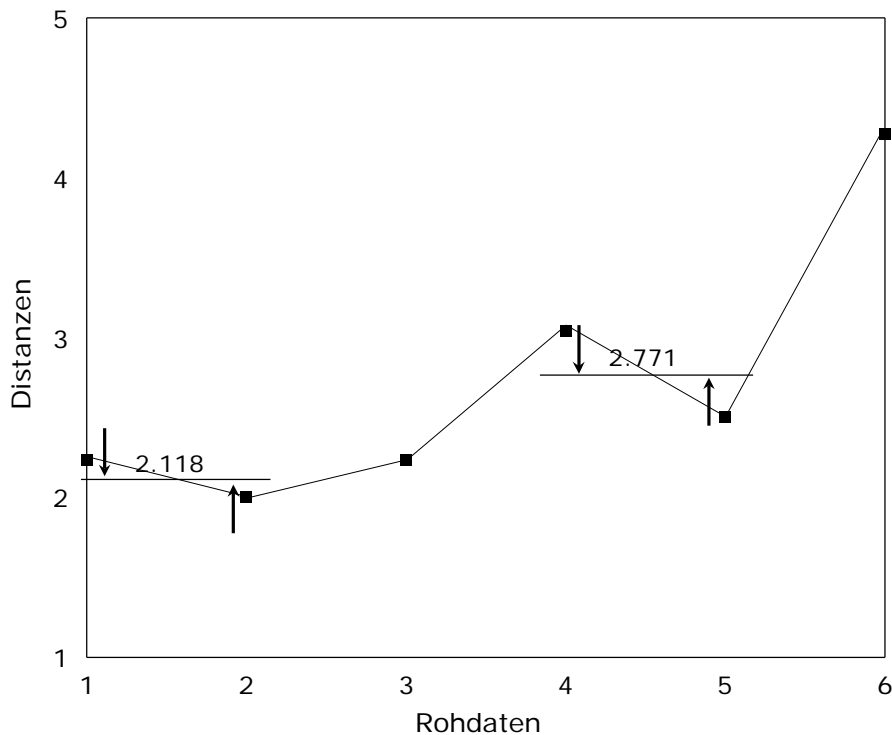


Abb. 3.1.2: Shepard-Diagramm. Die Rohdaten werden einer schwach-monotonen Transformation unterzogen und damit an die Distanzen angepaßt.

Wenn die Punkte im Shepard-Diagramm mit einer Linie verbunden werden, so ist die Abweichung von der Monotonie-Bedingung an den abwärts gerichteten Linien zwischen zwei Punkten zu erkennen. Die ordinalen Rohdaten werden nun nach folgender Regel in Disparitäten transformiert:¹⁶

1. Wenn die Monotonie-Bedingung zwischen zwei benachbarten Punkten nicht verletzt ist, so verwende die jeweiligen Distanzen als Transformation der Rohdaten!
2. Wenn die Monotonie-Bedingung zwischen zwei benachbarten Punkten verletzt ist, so verwende den Mittelwert der beiden Distanzen als Transformation der Rohdaten!

Tabelle 3.1.3 enthält die Distanzen und die über schwach-monotone Transformation gewonnenen Disparitäten. Zusätzlich wird die Berechnung des STRESS dieser Konfiguration in dieser Tabelle durchgeführt:

¹⁶ Wenn die Rohdaten intervallskaliert sind, darf keine schwach-monotone, sondern muß eine lineare Transformation durchgeführt werden. In diesem Fall werden die Disparitäten als vorhergesagte Werte einer linearen Regression der Distanzen auf die Rohdaten gewonnen. Im Shepard-Diagramm würden die Disparitäten dann als Punkte auf einer Geraden abgetragen. (Zur linearen Regression siehe die entsprechende Beispieldatei zur Datenanalyse.)

Tab. 3.1.3: Distanzen, Disparitäten und die Berechnung des STRESS

Objekt-Paare	Rohdaten (u)	Distanzen (d)	Disparitäten (d*)	Berechnung des STRESS	
				(d-d*) ²	d ²
AB	1	2,236	2,118	0,0139	4,9997
BC	2	2,000	2,118	0,0139	4,0000
AC	3	2,236	2,236	0,0000	4,9997
AD	4	3,041	2,771	0,0729	9,2477
CD	5	2,500	2,771	0,0734	6,2500
BD	6	4,272	4,272	0,0000	18,2500
			Summe	0,1742	47,7471

Benachbarte Punkte, zwischen denen die Monotonie-Bedingung verletzt ist, sind in Tabelle 3.1.3 fett hervorgehoben. Die Reihenfolge der Rohdaten wird in den Disparitäten beibehalten. Gleichzeitig sind die Disparitäten optimal an die Distanzen angepaßt. Wie die Anpassung dieser Konfiguration insgesamt zu beurteilen ist, wird über das STRESS-Maß¹⁷ ermittelt:

$$\text{STRESS} = \sqrt{\frac{\sum (d - d^*)^2}{\sum d^2}}$$

Der STRESS wird berechnet als Wurzel aus der Summe der quadrierten Abweichungen der Disparitäten d* von den Distanzen d, geteilt durch die Summe der quadrierten Distanzen. Für das Beispiel in Tabelle 3.1.3 ergibt sich ein STRESS von

$$\text{STRESS} = \sqrt{\frac{0,1742}{47,7471}} = 0,060.$$

Es existieren keine exakt festgelegten Schranken, um die Höhe des STRESS inhaltlich zu beurteilen. Als Daumenregel wird üblicherweise ein STRESS geringer als 0,1 als hervorragend bezeichnet. STRESS-Werte zwischen 0,1 und 0,2 gelten noch als akzeptabel. Letztlich gibt es keine exakten Vorgaben dafür, welchen STRESS-Wert man noch akzeptieren kann und welchen Wert man als „gut“ bezeichnen kann. Im Kontext psycho-physischer Messungen wird man dabei wesentlich strengere Maßstäbe verwenden als z.B. bei der Einstellungsmessung oder gar bei der explorativen Analyse von Wahrnehmungsräumen im Rahmen des Marketing.

„Um überhaupt eine Norm zu haben, hat man die ‚nullste aller Null-Hypothesen‘ untersucht und tausende von Zufallsdaten per MDS skaliert und dabei registriert, welche Stress-Werte sich ergeben“ (vgl. Borg/Staufenbiel 1989, S. 98f, wo auch Tabellen für den erwarteten Stress und dessen Standardabweichung in Zufallsstrukturen in Abhängigkeit von Objektzahl und Dimensionalität des Lösungsraums angegeben werden).

¹⁷ Es gibt verschiedene STRESS-Maße. Am gängigsten ist die Interpretation des STRESS-Maßes 1 von Kruskal, dessen Formel hier dargestellt ist. Das ALSCAL-Programm in SPSS verwendet zur Steuerung des Algorithmus ein verbessertes, aber auch komplexeres STRESS-Maß, das entwickelt wurde, um die Gefahr entarteter Lösungen zu verringern (vgl. 4.1.4). Die Ausgabe von SPSS enthält aber auch den Wert für das gerade geschilderte, für Beurteilungs- und Vergleichszwecke gebräuchlichste STRESS-Maß von Kruskal.

Neben dem STRESS wird ein weiteres Maß als Gütekriterium für die Anpassung der Konfiguration an die Rohdaten betrachtet. Es handelt sich hierbei um R^2 (auch abgekürzt als RSQ). R^2 ist die quadrierte Korrelation der Distanzen mit den Disparitäten und stellt ein Maß für die lineare Anpassung der Disparitäten an die Distanzen dar. R^2 wird auch als Varianz der Disparitäten interpretiert, die durch die Distanzen erklärt wird (vgl. auch Abschnitt 4.1.3). In der Praxis gelten R^2 -Werte größer als 0,9 als akzeptabel.

3.2 Optimierung 2: Anpassung der Konfiguration¹⁸

Obwohl im Beispiel aus Abschnitt 3.1 der STRESS bereits akzeptabel ist und daher die Iteration vermutlich schon zum Ende gekommen wäre, soll auf Basis des selben Beispiels erläutert werden, wie nun die Konfiguration angepasst würde.

Die Position der Objekte in der Konfiguration ist durch deren Koordinaten festgelegt. Diese Koordinaten werden nun angepasst. Eine Formel zur Anpassung der Koordinaten muss dabei zwei zentrale Fragen berücksichtigen:

1. Liegen zwei Objekte zu nahe beieinander oder zu weit auseinander?
2. In welcher Richtung liegen die beiden Objekte zueinander?

Die erste Frage lässt sich durch den Vergleich der Disparität eines Objektpaares mit der zugehörigen Distanz beantworten. Ist die Disparität kleiner als die Distanz, so liegen die Objekte noch zu weit auseinander. Ist die Disparität dagegen größer als die Distanz, so liegen die Objekte in der aktuellen Konfiguration zu eng beisammen.

Die zweite Frage wird durch die Betrachtung der Koordinaten der beiden Objekte gelöst. Im oben dargestellten Beispiel liegt der Punkt D von A aus gesehen in positiver Richtung auf der X-Achse und in negativer Richtung auf der Y-Achse. Um das Vorzeichen der Richtung von A aus zu bestimmen, wird der Koordinatenwert von A vom Koordinatenwert von D abgezogen:

$$\begin{aligned}x_D - x_A &= 5 - 2 = 3 \text{ (ist positiv),} \\y_D - y_A &= 3,5 - 4 = -0,5 \text{ (ist negativ).}\end{aligned}$$

Für die Berechnung eines neuen Koordinatenwertes x^* auf Dimension X für Objekt A in Bezug auf Objekt B wird die folgende Formel verwendet:¹⁹

$$x_A^* = x_A + a \left(1 - \frac{d_{AB}^*}{d_{AB}} \right) (x_B - x_A)$$

Im ersten Klammerausdruck wird das Verhältnis von Disparität zu Distanz berechnet und von 1 abgezogen. Wenn die Disparität mit der Distanz übereinstimmt, so wird dieser Ausdruck gleich 0 und es findet keine Anpassung der Koordinate x_A in Bezug auf Objekt B statt. Wenn die Disparität größer als die Distanz ist, wird der Ausdruck in der ersten Klammer negativ und x_A wird

¹⁸ In den folgenden Ausführungen wird der Anpassungsprozeß in Anlehnung Kruskal geschildert (Gradientenverfahren; vgl. Backhaus et al 1996, 453f). Tatsächlich verläuft die Anpassung der Konfiguration im Programmablauf von SPSS wesentlich komplizierter: Es wird ein eigener STRESS-minimierender iterativer Prozeß durchlaufen, der abbricht, sobald die neuen Koordinaten stabil bleiben (vgl. SPSS 1991, 9f).

¹⁹ Für Dimension Y gilt entsprechendes.

verringert. Dadurch wird Objekt A von Objekt B in Dimension X weggeschoben. Ob das Abrücken von Objekt A in Bezug auf Objekt B in positiver oder negativer Richtung der X-Achse erfolgen muss, wird im zweiten Klammersausdruck festgelegt.

Wie stark sich der Abstand der Objekte und das Verhältnis von Distanz und Disparität auf die Veränderung der Koordinaten auswirkt, wird vom Parameter α bestimmt. Dieser Parameter α darf nicht zu groß gewählt werden, da die Verschiebung der Objekte sonst über das Ziel hinaus schießen würde und zu einer schlechteren Konfiguration führt. Üblich sind Startwerte von $\alpha = 0,2$. Im Zuge der Iteration wird α weiter verringert.

Als Beispiel wird in der folgenden Tabelle 3.2.1 die Berechnung der neuen Koordinatenwerte für Objekt A unseres Beispiels durchgeführt:

Objekt A hat in der Startkonfiguration die folgenden Koordinaten:

$$x_A = 2$$

$$y_A = 4$$

Wir wählen $\alpha = 0,2$.

Tab. 3.2.1: Berechnung der neuen Koordinaten für Objekt A

Objekt i (x, y)	d*	d	1-(d*/d)	x_i-x_A	y_i-y_A	x^*_A	y^*_A
B (1, 2)	2,118	2,236	0,053	-1,0	-2,0	1,989	3,979
C (3, 2)	2,236	2,236	0,000	1,0	-2,0	2,000	4,000
D (5, 3.5)	2,771	3,041	0,089	3,0	-0,5	2,053	3,991
Mittelwert						2,014	3,990

Die neuen Koordinaten für A werden zunächst in Bezug auf jedes andere Objekt getrennt berechnet. Die Mittelwerte dieser jeweils drei einzelnen Werte stellen die endgültigen neuen Koordinaten von A dar.

Objekt A hat also in der angepassten Konfiguration die folgenden Koordinaten:

$$x^*_A = 2,014$$

$$y^*_A = 3,990$$

Dadurch bewegt sich A in der Konfiguration ein kleines Stück nach rechts unten. Dieser Vorgang wird analog für die Objekte B, C und D durchgeführt.

Durch die Änderung der Koordinaten der Objekte ändern sich auch die Distanzen zwischen den Objekten. Daher wird im Anschluss an die Verschiebung der Objekte in der Konfiguration direkt wieder der erste Optimierungsschritt (Neuberechnung der Disparitäten) durchgeführt.

4. Durchführung der MDS mit SPSS

In den folgenden Abschnitten wird die Durchführung einer MDS mit SPSS erläutert. Dazu verwenden wir zunächst ein (fiktives) Beispiel zur Wahrnehmung von Erfrischungsgetränken:

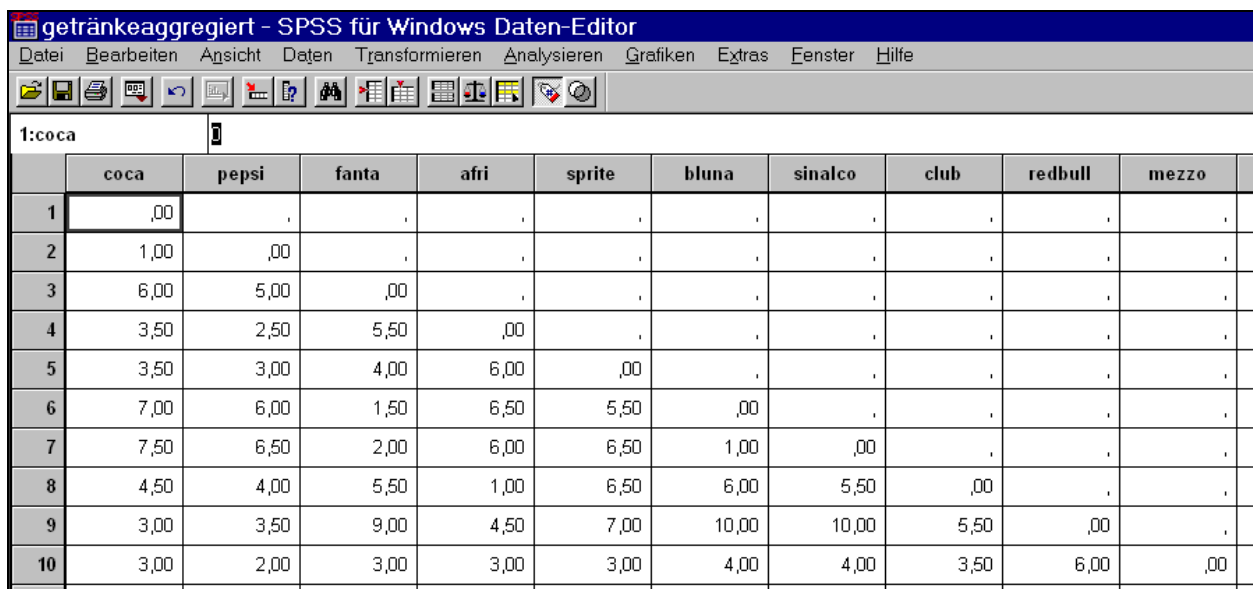
In einer Untersuchung wurden zwei Personen aufgefordert, die Ähnlichkeit von zehn Getränkemarken auf einer Ratingskala mit 10 Ausprägungen zu bewerten. Dabei bedeuteten kleine Zahlen hohe Ähnlichkeit und große Zahlen geringe Ähnlichkeit (= Distanz). Als Getränkemarken wurden ausgewählt: Coca Cola, Pepsi, Fanta, Afri Cola, Sprite, Bluna, Sinalco, Club-Cola, RedBull und Mezzo-Mix.

Bevor wir uns in unserem Beispiel den individuellen Datenmatrizen der beiden Personen zuwenden, untersuchen wir (aus didaktischen Gründen) zunächst die aggregierte Datenmatrix. Die aggregierte Ähnlichkeitsmatrix enthält die Mittelwerte der abgegebenen Ähnlichkeitsurteile für jedes Objektpaar.

4.1 Einfache ordinale MDS

4.1.1 Dateneingabe

Bevor die MDS mit SPSS durchgeführt werden kann, muss zunächst die Datenmatrix in SPSS eingegeben werden. Die Datenmatrix ist in diesem Fall quadratisch und symmetrisch. Deshalb reicht es aus, nur die untere Dreiecksmatrix einzugeben (siehe Abbildung 4.1.1.1).



The screenshot shows the SPSS Data Editor window titled 'getränkeaggregiert - SPSS für Windows Daten-Editor'. The menu bar includes 'Datei', 'Bearbeiten', 'Ansicht', 'Daten', 'Transformieren', 'Analysieren', 'Grafiken', 'Extras', 'Fenster', and 'Hilfe'. The toolbar contains various icons for file operations and data manipulation. The data grid shows a lower triangular matrix with 10 columns and 10 rows. The columns are labeled 'coca', 'pepsi', 'fanta', 'afri', 'sprite', 'bluna', 'sinalco', 'club', 'redbull', and 'mezzo'. The rows are numbered 1 to 10. The diagonal elements are all 0,00. The lower triangular elements represent similarity scores between different brands.

	coca	pepsi	fanta	afri	sprite	bluna	sinalco	club	redbull	mezzo
1	,00
2	1,00	,00
3	6,00	5,00	,00
4	3,50	2,50	5,50	,00
5	3,50	3,00	4,00	6,00	,00
6	7,00	6,00	1,50	6,50	5,50	,00
7	7,50	6,50	2,00	6,00	6,50	1,00	,00	.	.	.
8	4,50	4,00	5,50	1,00	6,50	6,00	5,50	,00	.	.
9	3,00	3,50	9,00	4,50	7,00	10,00	10,00	5,50	,00	.
10	3,00	2,00	3,00	3,00	3,00	4,00	4,00	3,50	6,00	,00

Abb. 4.1.1.1: Dateneingabe im SPSS-Daten-Editor

Zu beachten ist, dass jede Spalte und jede Zeile jeweils ein Objekt bezeichnet. Es können jedoch nur die Spalten (als Variablen) mit Bezeichnungen versehen werden. SPSS geht implizit davon aus, dass die Zeilen in der gleichen Reihenfolge wie die Spalten organisiert sind. So ent-

hält also Zeile 1 (wie Spalte 1) die Ähnlichkeitsurteile von Coca Cola zu den anderen Getränken, Zeile 2 enthält (wie Spalte 2) die Ähnlichkeitsurteile von Pepsi zu den anderen Getränken, usw. ...

4.1.2 Aufruf der Prozedur und Optionen

Sobald alle Daten eingegeben wurden, kann die Prozedur „Multidimensionale Skalierung“ ausgewählt werden. Die MDS befindet sich unter dem Menü „Analysieren“ und dem Untereintrag „Skalieren“ (siehe Abbildung 4.1.2.1).

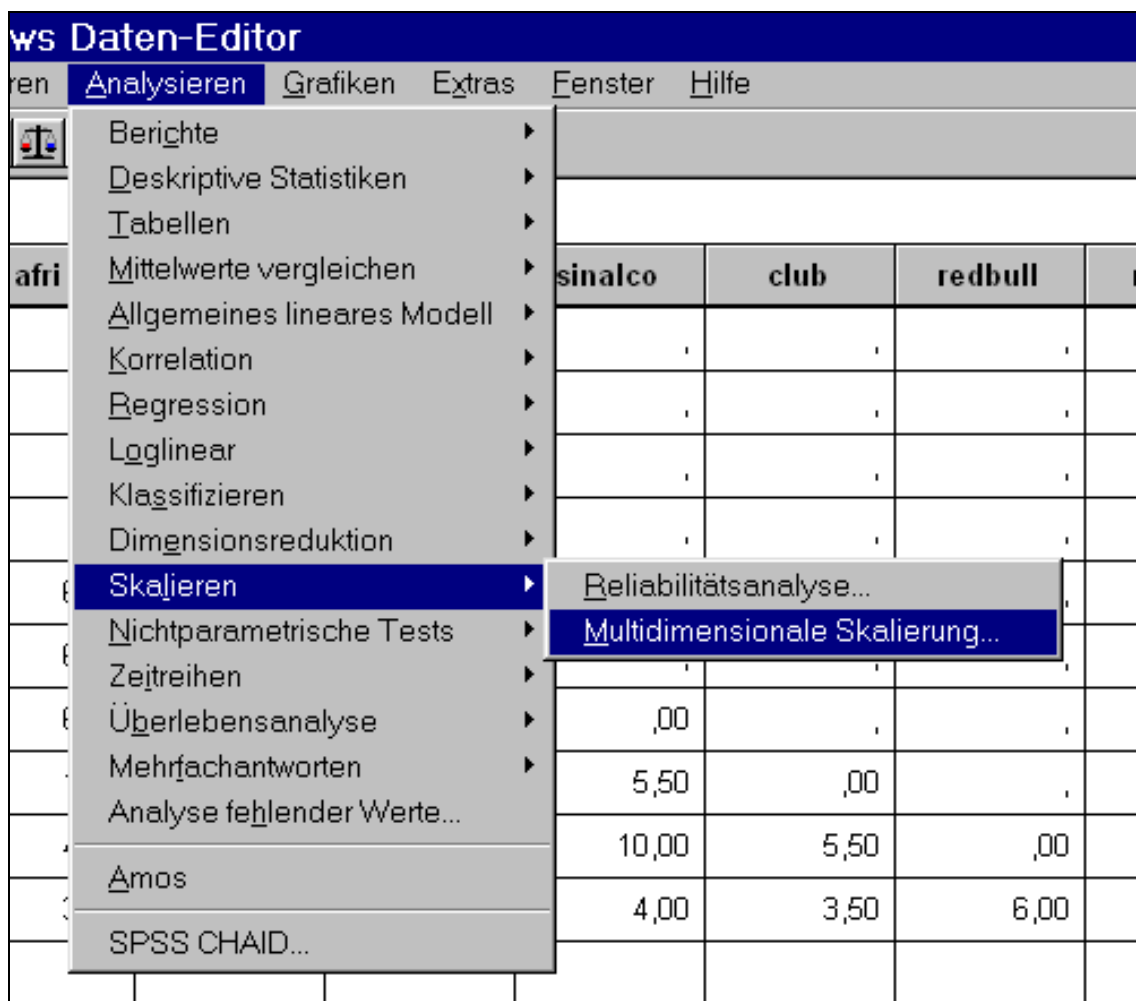


Abb. 4.1.2.1: Aufruf der MDS

Nach Auswahl der MDS erscheint das Haupteingabefeld der Prozedur (siehe Abbildung 4.1.2.2). Die wesentlichen Auswahloptionen in diesem Feld wurden von uns mit den Zahlen von 1 bis 5 gekennzeichnet.

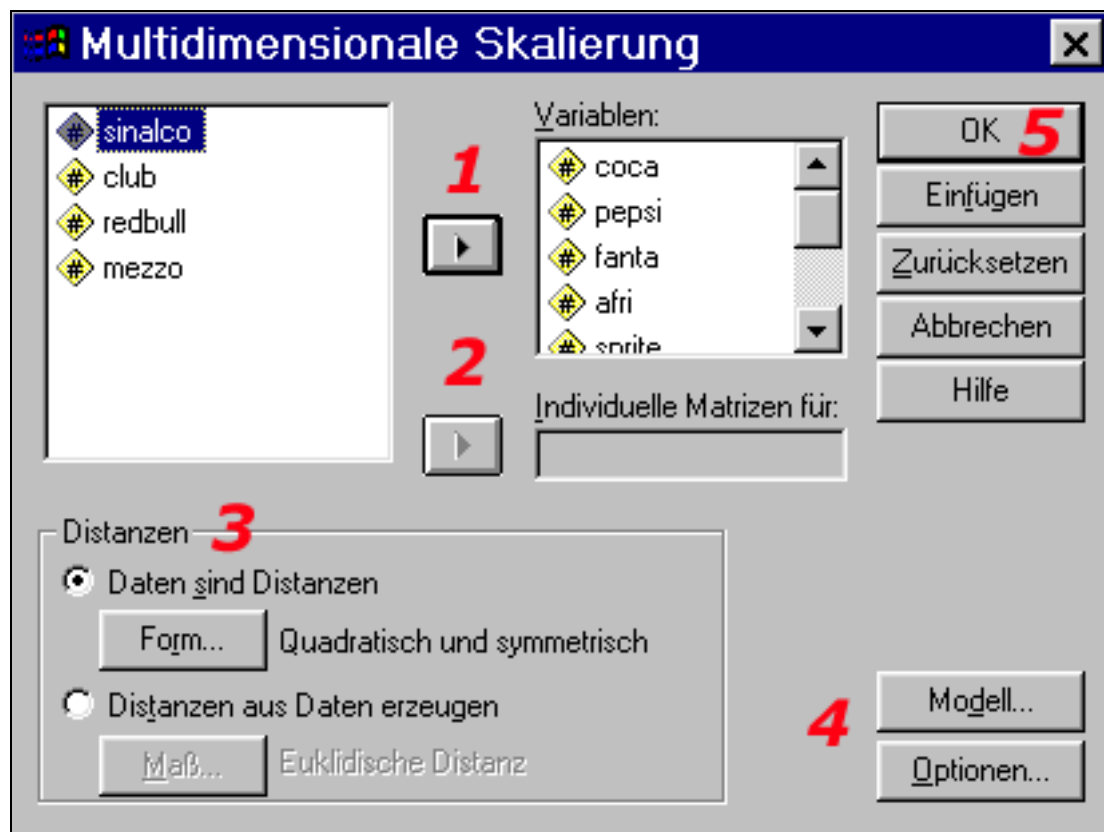


Abb. 4.1.2.2: Haupteingabefeld der SPSS-Prozedur „MDS“

Zunächst werden die zu analysierenden Objekte über die Pfeiltaste 1 vom linken Fenster in das rechte Fenster übertragen. Dabei ist darauf zu achten, dass auch wirklich alle Objekte - aber keine sonstigen Variablen - übertragen werden. Die Pfeiltaste 2 („individuelle Matrizen für ...“) hat nur eine Bedeutung, wenn die Daten noch nicht als Ähnlichkeitsmatrix vorliegen, sondern aus den Daten erst eine solche Matrix erzeugt werden soll.

Diese Möglichkeit kann im Bereich 3 ausgewählt werden. Dieser Bereich ist (etwas irreführend) mit der Überschrift „Distanzen“ versehen. Gemeint sind hier aber die Rohdaten der MDS (Ähnlichkeiten oder Unähnlichkeiten), nicht die Distanzen in der Konfiguration.²⁰ Mit der Auswahl „Daten sind Distanzen“ wird der Prozedur mitgeteilt, dass die Rohdaten wie in unserem Beispiel bereits in der Form einer Matrix vorliegen.

Im Bereich 4 kann nun das Modell genauer spezifiziert werden und es können verschiedene Optionen für die Ergebnisausgabe und das Iterationsverfahren ausgewählt werden.

Sind alle erforderlichen Einstellungen vorgenommen worden, kann die MDS durchgeführt werden. Dies geschieht in Bereich 5 entweder direkt durch die Auswahl „OK“, oder es wird erst der vollständig spezifizierte Befehl über die Taste „Einfügen“ in ein Syntaxfenster übertragen, um dann dort (gegebenenfalls weiter modifiziert) ausgeführt zu werden.

Durch die Auswahl der Taste „Form ...“ im Bereich 3 („Distanzen“) gelangen wir zu folgendem Auswahlfeld (Abbildung 4.1.2.3):

²⁰ Zu Missverständnissen kann es hier vor allem dann kommen, wenn aus den vorliegenden Daten erst eine „Distanzmatrix“ erzeugt werden soll. Unter dieser Option kann unter anderem aus einer Vielzahl unterschiedlicher Distanzmetriken (euklidische Metrik, City-Block-Metrik etc.) ausgewählt werden. Diese Distanzmetriken beziehen sich aber ausschließlich auf die Rohdaten - mit der Metrik in der Konfiguration hat diese Auswahl nichts zu tun.

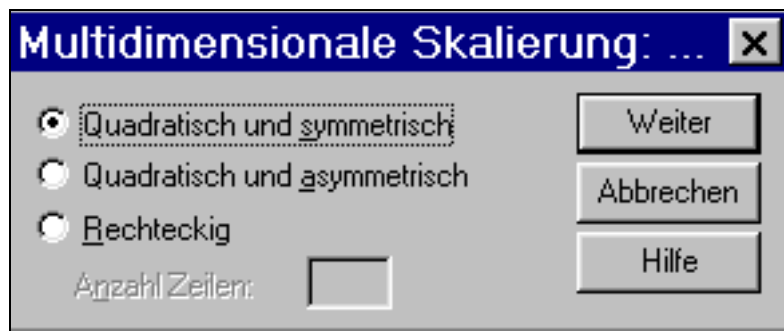


Abb. 4.1.2.3: Form der Rohdatenmatrix

An dieser Stelle wird eingestellt, welche Form die Ähnlichkeitsmatrix hat (vgl. Abschnitt 1.3). In unserem Beispiel trifft die Voreinstellung „Quadratisch und symmetrisch“ zu. Wenn an Stelle der MDS eine MDU durchgeführt werden soll (vgl. Abschnitt 1.3.2) und die Daten daher in einer Rechtecksmatrix vorliegen, wird hier „Rechteckig“ ausgewählt. Zusätzlich muss dann eingetragen werden, wie viele Zeilen (in der Regel Subjekte) die Rechtecksmatrix enthält. Die Option „Quadratisch und asymmetrisch“ wird ausgewählt, wenn die Ankerpunktmethod als Erhebungsverfahren eingesetzt wurde (vgl. Abschnitt 2.2.3) und die Datenmatrix daher zwar quadratisch, aber asymmetrisch und zeilenkonditional ist.

Durch die Auswahl der Taste „Weiter“ gelangen wir wieder zum Haupteingabefeld (Abbildung 4.1.2.2) zurück. Hier wählen wir nun im Bereich 4 die Taste „Modell ...“ und öffnen folgendes Eingabefeld (Abbildung 4.1.2.4):

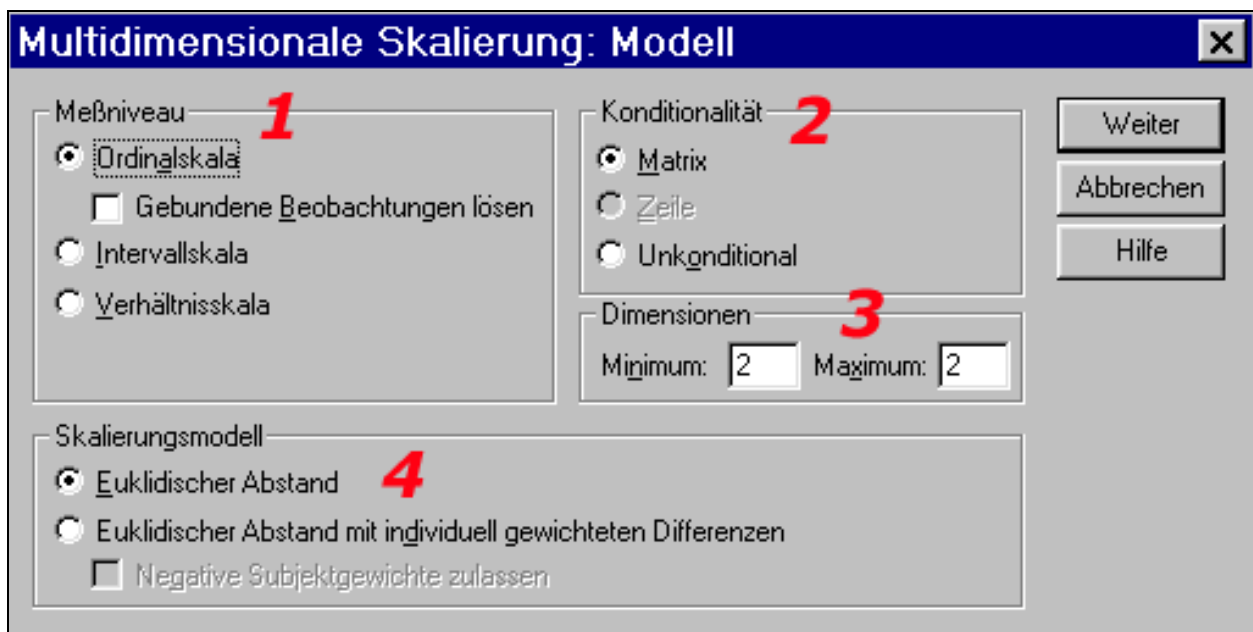


Abb. 4.1.2.4: Modelloptionen der MDS

Auch hier haben wir die Bereiche mit Zahlen versehen. Im Bereich 1 wird das „Meßniveau“ der Rohdaten festgelegt. Die Voreinstellung ist hier „Ordinalskala“ und bedeutet, dass die Ähnlichkeitsurteile der Eingabematrix als Rangordnung interpretiert werden, die absoluten Abstände zwischen benachbarten Ähnlichkeitsurteilen jedoch nicht berücksichtigt werden (vgl. die Abschnitte 1.1.2 und 1.2).

In Bereich 2 wird die „Konditionalität“ der Rohdaten angegeben (vgl. Abschnitt 2.2). Die Voreinstellung „Matrix“ ist in unserem Beispiel korrekt, da die Ähnlichkeitsratings unserer Eingabematrix innerhalb dieser Matrix vergleichbar sind. Die Einstellung „Zeile“ ist hier nicht möglich, da die Matrix bereits oben als symmetrisch definiert wurde. Wenn die Daten als Rechtecksmatrix eingegeben wurden (MDU), so ist an dieser Stelle Zeilenkonditionalität festzulegen. Gleiches gilt für eine quadratische asymmetrische Matrix, die mit der Ankerpunktmethodene erhoben wurde. Die Option „Unkonditional“ ist nur dann relevant, wenn mehr als eine Matrix eingegeben wird und die Daten auch über die einzelnen Matrizen hinaus vergleichbar sind. Dieser Fall ist eher selten, könnte aber zum Beispiel bei einer wiederholten Messung eintreten.

Im Bereich 3 wird die minimale und maximale Zahl der „Dimensionen“ der Konfiguration eingestellt. Prinzipiell kann dieser Bereich von 1 bis zur Anzahl der Objekte weniger 1 reichen. In der Praxis wird eine Konfiguration mit mehr als drei Dimensionen aufgrund der problematischen Interpretation kaum angestrebt werden. Wir belassen es bei der Voreinstellung von 2 Dimensionen (Minimum und Maximum).

Im Bereich 4 wird die Entscheidung darüber getroffen, ob alle Matrizen in ein- und derselben Konfiguration abgebildet werden sollen (einfache MDS/MDU oder wiederholte MDS/MDU), oder ob die Dimensionen für jede Matrix individuell gewichtet werden sollen (INDSCAL) (vgl. Abschnitt 1.3.1). Die Auswahl von INDSCAL („Euklidischer Abstand mit individuell gewichteten Differenzen“) stellt natürlich nur dann eine sinnvolle Option dar, wenn mehr als eine Ähnlichkeitsmatrix verarbeitet wird. In unserem Beispiel ist daher die Voreinstellung („Euklidischer Abstand“) adäquat.

Mit der Taste „Weiter“ gelangen wir wieder in das Haupteingabefeld (Abbildung 4.1.2.2) zurück und wählen dort zum Abschluss im Bereich 4 die Taste „Optionen ...“ (Abbildung 4.1.2.5):

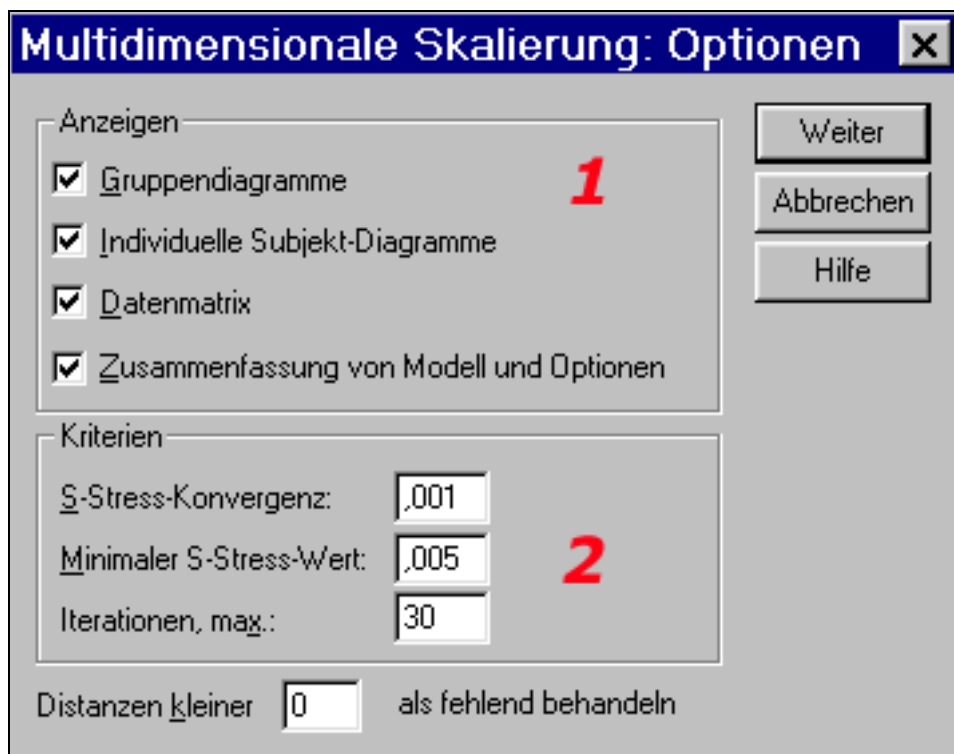


Abb. 4.1.2.5: Optionen für Output und Iterationsverfahren der MDS

Im Bereich 1 der Optionen wird eingestellt, was alles in der Ergebnisausgabe der MDS enthalten sein soll. Per Voreinstellung ist hier nichts angeklickt. Wir setzen daher ein Häkchen in jedes Feld.

Im Bereich 2 werden die „Kriterien“ für den Abbruch des Iterationsprozesses festgelegt (vgl. Abschnitt 3.). Die Voreinstellung ist hier in der Regel brauchbar:

- Die „S-Stress-Konvergenz“ von 0,001 bedeutet, dass die Iteration abgebrochen wird, wenn sich der S-STRESS von einem Iterationsschritt zum nächsten weniger als 0,001 ändert und somit die Verschiebung der Objekte in der Konfiguration nur noch minimal war.
- Der „minimale S-STRESS-Wert“ von 0,005 führt dazu, dass die Iteration abgebrochen wird, wenn der S-STRESS kleiner als 0,005 ist, was im Normalfall auf eine ausreichende Anpassung der Konfiguration schließen lässt.
- Die „maximale Anzahl von Iterationen“ wird auf 30 begrenzt, um zu verhindern, dass der Iterationsprozess in eine Endlosschleife gerät. Die Begrenzung auf 30 Iterationen kann als sehr vorsichtig bezeichnet werden - höhere Werte führen auf neueren Rechnern kaum zu nennenswerten Geschwindigkeitseinbußen.

Wenn eines der drei Kriterien erfüllt ist, wird die Iteration beendet. Insbesondere wenn die Iteration aufgrund des letzten Kriteriums abgebrochen wurde, sollte eine Erhöhung dieser Angabe in Betracht gezogen werden.

Nach der Rückkehr in das Haupteingabefeld (Taste „Weiter“) wird die MDS durch die Auswahl der Taste „OK“ ausgeführt.

4.1.3 Interpretation des Output

Aufgrund der getroffenen Einstellungen erzeugt SPSS umfangreichen Output.

Zunächst werden die eingestellten Optionen wiedergegeben, beginnend mit der Datenbasis:

Alscal Procedure Options

Data Options-

```

Number of Rows (Observations/Matrix). 10
Number of Columns (Variables) . . . 10
Number of Matrices . . . . . 1
Measurement Level . . . . . Ordinal
Data Matrix Shape . . . . . Symmetric
Type . . . . . Dissimilarity
Approach to Ties . . . . . Leave Tied
Conditionality . . . . . Matrix
Data Cutoff at . . . . . ,000000

```

Die eingegebene symmetrische Matrix hat 10 Zeilen (*rows*) und Spalten (*columns*). Es handelt sich hierbei um ordinale Unähnlichkeiten (*ordinal dissimilarity*) - je größer deren Werte, desto unähnlicher sind sich zwei Objekte. Bindungen (*ties*) werden nicht gelöst (*leave tied*). Das bedeutet, dass gleich große Unähnlichkeiten tatsächlich als gleich behandelt werden. Die Daten sind matrixkonditional, also innerhalb einer Matrix vergleichbar. Die Angabe „Data Cutoff at ,000000“ bedeutet, dass Rohdaten kleiner als Null als fehlend betrachtet werden.

Es folgen die Modelloptionen:

Model Options-

```

Model . . . . . Euclid
Maximum Dimensionality . . . . . 2
Minimum Dimensionality . . . . . 2
Negative Weights . . . . . Not Permitted

```

In der Konfiguration herrscht die euklidische Metrik (`model euclid`). Berechnet wird nur die zweidimensionale Lösung (`maximum dimensionality = minimum dimensionality = 2`). Da hier nicht das INDSCAL-Verfahren angewendet wird, ist die Angabe zur Zulässigkeit negativer Gewichte (`negative weights not permitted`) in diesem Beispiel nebensächlich.

Die nächsten Angaben beziehen sich auf die Ergebnisausgabe:

Output Options-

```

Job Option Header . . . . . Printed
Data Matrices . . . . . Printed
Configurations and Transformations . . . . . Plotted
Output Dataset . . . . . Not Created
Initial Stimulus Coordinates . . . . . Computed

```

Diese Angaben sind weitgehend selbsterklärend. Interessant ist hier aber der Hinweis auf die prinzipielle Möglichkeit, die Daten der Konfiguration in einer Datei zu speichern (`output dataset`). Diese Daten könnten dann z.B. als Startkonfiguration einer weiteren MDS verwendet werden. In unserer Analyse wurde die Startkonfiguration jedoch von SPSS berechnet (`initial stimulus coordinates computed`).

Den Abschluß der Angaben zur durchgeführten Analyse bilden die Kriterien für den Abbruch des Iterationsverfahrens:

Algorithmic Options-

```

Maximum Iterations . . . . . 30
Convergence Criterion . . . . . ,00100
Minimum S-stress . . . . . ,00500
Missing Data Estimated by . . . . . Ulbounds
Tiestore . . . . . 45

```

Zusätzlich werden noch Einstellungen zur Behandlung fehlender Werte (`missing data`) und Ties (`tiestore`) ausgegeben, auf die hier jedoch nicht weiter eingegangen werden soll.

Als nächstes werden die Rohdatenmatrizen (hier eine einzige) ausgegeben. An dieser Stelle kann noch einmal kontrolliert werden, ob diese Werte korrekt eingegeben wurden. Außerdem ist es so Dritten möglich, die Analyse zu wiederholen.

```

Raw (unscaled) Data for Subject 1
      1      2      3      4      5
1      ,000
2      1,000      ,000
3      6,000      5,000      ,000
4      3,500      2,500      5,500      ,000
5      3,500      3,000      4,000      6,000      ,000
6      7,000      6,000      1,500      6,500      5,500
7      7,500      6,500      2,000      6,000      6,500
8      4,500      4,000      5,500      1,000      6,500
9      3,000      3,500      9,000      4,500      7,000
10     3,000      2,000      3,000      3,000      3,000
      6      7      8      9      10
6      ,000
7      1,000      ,000
8      6,000      5,500      ,000
9      10,000     10,000     5,500      ,000
10     4,000      4,000      3,500      6,000      ,000

```

Die nun folgende Warnmeldung soll daran erinnern, dass die Anhebung des Skalenniveaus (ordinal → ratio) im Zuge der MDS eine ausreichende Anzahl an Rohdaten benötigt (vgl. Datenverdichtungskoeffizient in Abschnitt 2.1).

```

>Warning # 14654
>The total number of parameters being estimated (the number of stimulus
>coordinates plus the number of weights, if any) is large relative to the
>number of data values in your data matrix. The results may not be reliable
>since there may not be enough data to precisely estimate the values of the
>parameters. You should reduce the number of parameters (e.g. request
>fewer dimensions) or increase the number of observations.

>Number of parameters is 20. Number of data values is 45

```

In vorliegendem Beispiel werden 20 Koordinaten auf der Basis von 45 Unähnlichkeiten geschätzt. Der Datenverdichtungskoeffizient wäre also größer als zwei, so dass die Warnung an dieser Stelle etwas konservativ erscheint. Wenn mehr als 10 Objekte zweidimensional skaliert werden, wird diese Warnung nicht mehr ausgegeben.

Auf die Ausgabe der Warnmeldung folgen Angaben zum Verlauf des Iterationsprozesses.

Iteration history for the 2 dimensional solution (in squared distances)

Young's S-stress formula 1 is used.

Iteration	S-stress	Improvement
1	,02813	
2	,02541	,00272
3	,02459	,00082

Iterations stopped because
S-stress improvement is less than ,001000

Wir sehen, dass die Iteration drei Schleifen durchlaufen hat. Die Iteration wurde abgebrochen, weil die Objekte in der Konfiguration vom zweiten zum dritten Iterationsschritt kaum noch verschoben wurden (S-Stress²¹ improvement is less than ,001).²²

Es folgen die ersten Ergebnisse der MDS: STRESS und R².

Stress and squared correlation (RSQ) in distances

RSQ values are the proportion of variance of the scaled data (disparities) in the partition (row, matrix, or entire data) which is accounted for by their corresponding distances. Stress values are Kruskal's stress formula 1.

For matrix
Stress = ,02179 RSQ = ,99707

Der hier ausgegebene STRESS basiert auf der in Abschnitt 3.1 erläuterten Formel von Kruskal (STRESS 1) und liegt mit einem Betrag von 0,02179 deutlich unter der Grenze von 0,1. Daher kann die Anpassung als hervorragend bezeichnet werden. R² (RSQ) beträgt 0,99707. Das bedeutet, die Varianz der Disparitäten wird zu 99,7% durch die Distanzen der Konfiguration erklärt.²³ Auch dieser Wert ist höchst zufriedenstellend.

²¹ SPSS verwendet im Zuge der Iteration den S-STRESS von Young, dessen Berechnung sich vom STRESS 1 von Kruskal unterscheidet (siehe Fußnote 17).

²² Ein generelles Problem des hier geschilderten Minimierungsverfahrens ist die *Gefahr lokaler Minima*. So kann zwar mit Recht davon ausgegangen werden, dass in der Umgebung der gefundenen Lösung keine Punktconfiguration mit geringerem STRESS und damit besserer Anpassung existiert. Das bedeutet aber nicht, dass nicht möglicherweise mit einer anderen Anfangskonfiguration ein anderes, eventuell besseres, lokales Minimum und damit ein niedrigerer STRESS-Wert gefunden werden kann. Es empfiehlt sich daher, auch andere Anfangskonfigurationen auszuprobieren, um das globale Konvergenzverhalten des Algorithmus zu erkunden - insbesondere dann, wenn die inhaltliche Interpretation der Konfiguration fragwürdig erscheint. Erweisen sich dabei die gefundene Lösung und der zugehörige STRESS-Wert als stabil, kann man mit einer gewissen Sicherheit vermuten, dass die Lösung optimal ist. Man kann aber nie ganz sicher sein, tatsächlich das absolute Minimum des STRESS-Maßes gefunden zu haben.

²³ Diese Interpretation von R² wird freundlicherweise im Output gleich mitgeliefert.

Anschließend werden die Koordinaten der Objekte in der zweidimensionalen Konfiguration ausgegeben.

```

Configuration derived in 2 dimensions

      Stimulus Coordinates

      Dimension
Stimulus Stimulus   1       2
Number   Name
1       COCA      1,0732  -,4818
2       PEPSI      ,7125  -,3833
3       FANTA     -1,3131  -,2381
4       AFRI       ,6451   ,8952
5       SPRITE    -,0711  -1,4253
6       BLUNA     -1,7408  -,0561
7       SINALCO  -1,7354  ,5402
8       CLUB      ,3461   1,1895
9       REDBULL   2,2358  ,1069
10      MEZZO     -,1523  -,1472

```

Diese Koordinaten werden verwendet, um die Konfiguration der Objekte darzustellen. Sie können auch in andere Programme übernommen werden, um eine bessere graphische Darstellung zu erzielen, als dies mit SPSS möglich ist (siehe die Bemerkung zur graphischen Darstellung in SPSS weiter unten in diesem Abschnitt).

Abschließend werden die transformierten Rohdaten (Disparitäten) ausgegeben.

```

Optimally scaled data (disparities) for subject 1

      1       2       3       4       5
1      ,000
2      ,471      ,000
3      2,424      2,031      ,000
4      1,490      1,274      2,194      ,000
5      1,490      1,274      1,664      2,424      ,000
6      2,808      2,424      ,471      2,602      2,194
7      2,989      2,602      ,891      2,424      2,602
8      1,799      1,664      2,194      ,471      2,602
9      1,274      1,490      3,566      1,799      2,808
10     1,274      ,891      1,274      1,274      1,274

      6       7       8       9       10
6      ,000
7      ,471      ,000
8      2,424      2,194      ,000
9      3,987      3,987      2,194      ,000
10     1,664      1,664      1,490      2,424      ,000

```

Mit der Ausgabe der Disparitäten endet der Output im Textformat. Es folgen nun noch vier graphische Darstellungen.

Als erstes wird als Ergebnis der MDS die zweidimensionale Konfiguration ausgegeben (Abbildung 4.1.3.1):

Konfiguration des abgeleiteten Stimulus

Euklidisches Distanzmodell

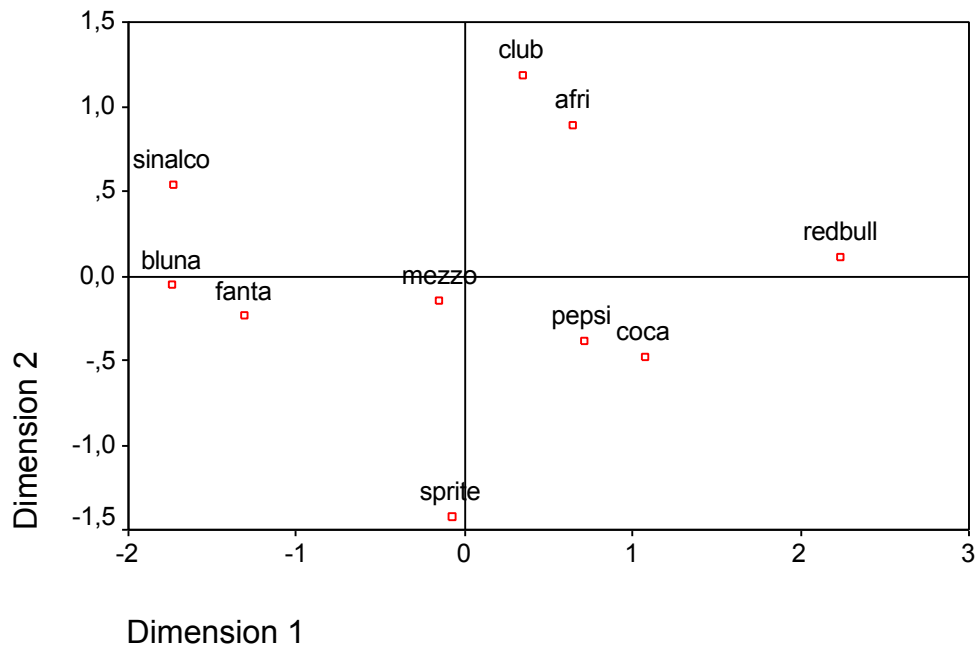


Abb. 4.1.3.1: SPSS-Graphik: Konfiguration der Objekte im zweidimensionalen Raum (Ergebnis der MDS)

Bevor die Konfiguration inhaltlich interpretiert werden soll, eine generelle Bemerkung zur Interpretation der von SPSS erzeugten Graphiken: Es ist vor jeder Interpretation der SPSS-Graphiken darauf zu achten, ob die Dimensionen der Graphik maßstabsgetreu sind: Der Abstand zwischen 0 und 1 auf Dimension 1 muss dem Abstand zwischen 0 und 1 auf Dimension 2 entsprechen, da sonst der euklidische Raum nicht korrekt interpretierbar ist. In diesem Beispiel ist das (zufällig!) der Fall - dem ist jedoch bei weitem nicht immer so. Daher bietet es sich generell an, die graphische Darstellung der Konfiguration auf Basis der ausgegebenen Koordinaten mit Hilfe eines besseren Graphikprogrammes zu erstellen.²⁴

Inhaltlich lässt sich die Konfiguration in Abbildung 4.1.3.1 wie folgt interpretieren:

Der Wahrnehmungsraum der Befragten lässt sich in sechs Gruppen von Getränken untergliedern: Orangenlimonade (Sinalco, Bluna, Fanta), Zitronenlimonade (Sprite), Cola-Fanta-Mischgetränke (Mezzo-Mix), Cola [Weltmarken] (Coca Cola, Pepsi), Cola [regionale Marken] (Club-Cola, Afri Cola) und besonders anregende Getränke (Red Bull). Diese Gruppen sind in ungefähr horizontaler Richtung von rechts nach links nach ihrem Anregungspotential angeordnet (Getränke ohne Koffein ganz links, Getränke mit mehr Koffein weiter rechts). In ungefähr vertikaler Richtung schlägt sich offenbar die Markenbekanntheit nieder. Bekannte Marken liegen weiter unten, weniger bekannte Marken liegen weiter oben.

Die drei weiteren Diagramme dienen zur Beurteilung der Güte der Anpassung der Konfiguration an die Rohdaten.

²⁴ So wurden alle Graphiken in dieser Beispieldatei (mit Ausnahme des SPSS-Outputs) mit Harvard Graphics erstellt.

Streudiagramm mit linearer Anpassung

Euklidisches Distanzmodell

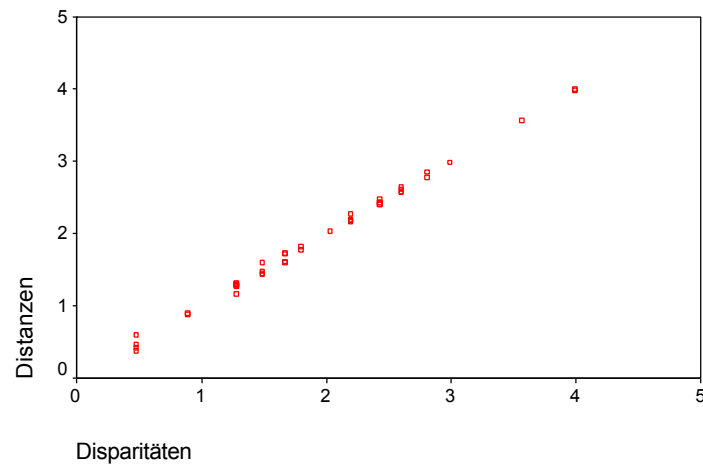


Abb. 4.1.3.2: Streudiagramm mit linearer Anpassung (Disparitäten vs. Distanzen).

Im Streudiagramm mit linearer Anpassung (Abbildung 4.1.3.2) werden die Disparitäten gegen die Distanzen geplottet. Je stärker die Punkte in dieser Graphik um eine Gerade durch den Ursprung streuen, desto größer wird R^2 .

Streudiagramm mit nicht linearer Anpassung

Euklidisches Distanzmodell

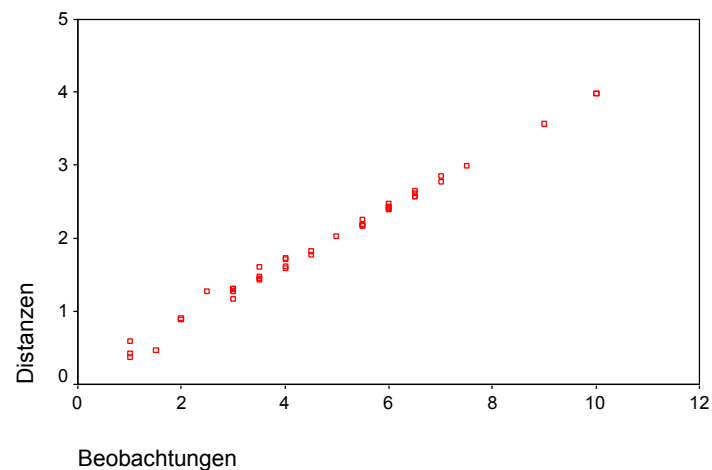


Abb. 4.1.3.3: Streudiagramm mit nicht-linearer Anpassung (Beobachtungen vs. Distanzen)

Im Streudiagramm mit nicht linearer Anpassung (Abbildung 4.1.3.3) werden die Rohdaten (Beobachtungen) gegen die Distanzen geplottet. Diese Gegenüberstellung findet im Zuge des Verfahrens eigentlich gar nicht statt, da die Rohdaten immer erst in Disparitäten transformiert werden. Zu erkennen ist in dieser Abbildung, dass die Punkte nicht monoton ansteigen, sondern die Monotoniebedingung an einigen Stellen (z.B. zwischen den beobachteten Werten 1,5 und 2) verletzt wird.

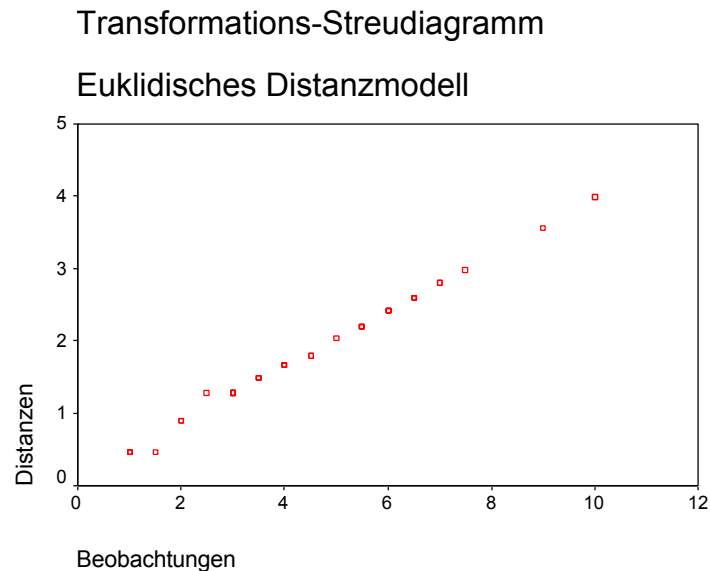


Abb. 4.1.3.4: Transformationsstreudiagramm (Beobachtungen vs. Disparitäten) [SPSS beschriftet diese Graphik falsch!]

Im Transformationsstreudiagramm (Abbildung 4.1.3.4) werden die Rohdaten gegen die Disparitäten geplottet (Achtung: SPSS beschriftet dieses Shepard-Diagramm falsch!).²⁵ Deutlich zu erkennen ist der treppenförmige Verlauf der Punkte im Shepard-Diagramm, der durch die schwach-monotone Transformation der Rohdaten in Disparitäten bei der ordinalen MDS hervorgerufen wird.

4.1.4 Die Gefahr der entarteten Lösung einer ordinalen MDS

Die monotone Transformation der Rohdaten in Disparitäten bei der ordinalen MDS induziert ein gewisses Risiko, dass die Lösung der MDS entartet. Die Gefahr der entarteten Lösung einer ordinalen MDS besteht immer dann, wenn ein Objekt (oder eine Subgruppe der Objekte) zu allen anderen Objekten die größte Unähnlichkeit aufweist. Im diesem Fall kann das STRESS-Maß beliebig verkleinert werden, je weiter das besagte Objekt von allen anderen Objekten abgerückt wird. Ein Blick auf die Formel des STRESS-Maßes (siehe S. 26) zeigt, dass die immer größer werdende Distanz dieses Objektes zu den anderen Objekten den Nenner des STRESS quadratisch anwachsen läßt, während sie im Zähler der Formel nicht von der entsprechenden Disparität abweicht und daher keinen Einfluß auf die Größe des Zählers hat.

Um das Verhalten des Algorithmus der MDS im Falle einer entarteten Lösung zu demonstrieren, haben wir das zuletzt ausgeführte Getränkebeispiel modifiziert. Es wurde ein weiteres Getränk (z.B. Mineralwasser) in die Untersuchung mit aufgenommen, welches zu allen anderen

²⁵ Wir haben den Fehler hier aus didaktischen Gründen absichtlich nicht korrigiert. Die Graphiken können jedoch mit SPSS bearbeitet werden. Dabei kann auch die Beschriftung („Disparitäten“ anstelle von „Distanzen“) korrigiert werden.

Getränken die größte Unähnlichkeit hat. Diese Unähnlichkeit wird in der erhobenen Ähnlichkeitsmatrix durch den Wert 11 repräsentiert.²⁶

Um den Effekt der entarteten Lösung deutlicher zu zeigen, haben wir die Konvergenzkriterien für den Algorithmus strenger formuliert. Als Höchstzahl an Iterationen wurde der Wert 500 vorgegeben. Als maximal erlaubte Änderung des STRESS-Maßes zwischen zwei aufeinanderfolgenden Iterationen wurde 0 eingetragen. Als maximal akzeptierter STRESS (hier S-STRESS) wurde der Wert 0,001 festgelegt.

Die Optimierung der Konfiguration wurde nach 82 Iterationen abgebrochen, weil der S-STRESS den Höchstwert von 0,001 unterschritt. Die Lösung weist einen STRESS (nach Kendall) von 0,01909 auf. R^2 beträgt 0,99954. Die Anpassung scheint also nahezu perfekt. Ein Blick auf die graphische Darstellung der Konfiguration zeigt jedoch, dass das Ergebnis dieser MDS inhaltlich wenig aussagekräftig ist (siehe Abbildung 4.1.4.1):

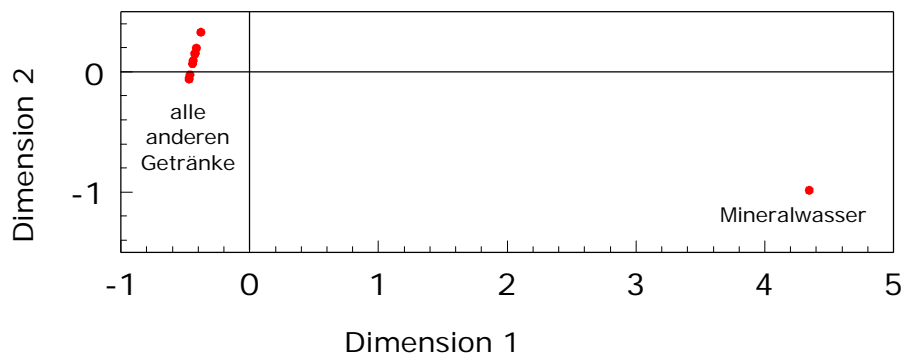


Abb. 4.1.4.1: Entartete Lösung einer ordinalen MDS.

Die einzige Information, die uns diese Konfiguration über den Wahrnehmungsraum der Befragten liefert, ist der große Unterschied, der zwischen Mineralwasser und der Gruppe aller anderen Getränke gesehen wird. Um etwas über die Konfiguration der übrigen Getränke zu erfahren, muß das Mineralwasser aus der Analyse ausgeschlossen werden (zu den Ergebnissen siehe Abschnitt 4.1.3).

In der Praxis sind entartete Lösungen häufig nicht auf den ersten Blick zu erkennen, da die Voreinstellung der Konvergenzkriterien die extreme Verzerrung der Konfiguration (wie in Abbildung 4.1.4.1) verhindert. Im Zweifelsfall ist es daher angebracht, mit den Konvergenzkriterien zu experimentieren, um eine entartete Lösung entweder auszuschließen oder aber die Ausreisser in der Konfiguration zu identifizieren und aus der weiteren Analyse auszuschließen.

4.2 Einfache Intervall-MDS

Wir analysieren nun die gleiche Datenmatrix erneut. Der einzige Unterschied zur vorangegangenen Analyse besteht im Skalenniveau der Rohdaten. Zuvor wurden die Rohdaten als ordi-

²⁶ Im Prinzip ist es beliebig, welche Zahl hier verwendet wird. Sie muss nur größer sein, als alle anderen Unähnlichkeiten in der Distanzmatrix.

nalskaliert betrachtet. In der jetzt folgenden Analyse wird unterstellt, dass die Rohdaten intervallskaliert sind.

4.2.1 Aufruf der Prozedur und Optionen

Alle Einstellungen der vorherigen Analyse können beibehalten werden, mit einer Ausnahme bei den Modelloptionen (Abbildung 4.2.1.1): Im Bereich 1 der Modelloptionen wird als Meßniveau nun „Intervallskala“ ausgewählt.

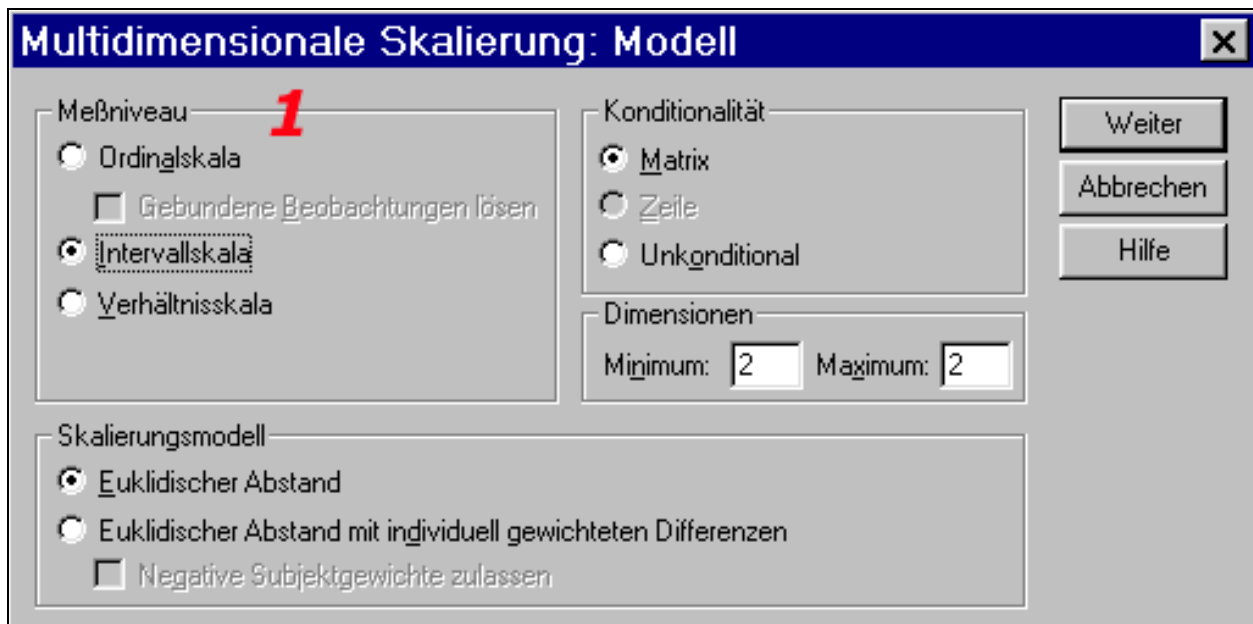


Abb. 4.2.1.1: Modelloptionen der Intervall-MDS

4.2.2 Interpretation des Output

Die Durchführung der Analyse mit den veränderten Optionen erzeugt einen zur ersten Analyse weitestgehend identischen Output. Daher sollen hier nur die wesentlichen Unterschiede kommentiert werden.

Alscal Procedure Options

Data Options-

```

Number of Rows (Observations/Matrix).    10
Number of Columns (Variables) . . . . . 10
Number of Matrices . . . . . 1
Measurement Level . . . . . Interval
Data Matrix Shape . . . . . Symmetric
Type . . . . . Dissimilarity
Approach to Ties . . . . . Leave Tied
Conditionality . . . . . Matrix
Data Cutoff at . . . . . ,000000

```

Die Ausgabe der Optionen unterscheidet sich nur in einem Punkt von der vorangegangenen Analyse: Measurement Level Intervall.

Es folgt wie oben die Ausgabe der Rohdaten. Die Warnmeldung (siehe Seite 36) unterbleibt jedoch, da die Anhebung des Skalenniveaus von Intervall- auf Rationiveau weniger problematisch ist.

Der STRESS der Lösung verschlechtert sich minimal und beträgt nun 0,02818. Gleiches gilt für R^2 mit einem Betrag von 0,99510. Diese Verschlechterung von STRESS und R^2 beim Wechsel auf ein höheres Skalenniveau der Rohdaten erfolgt fast zwangsläufig, da die erlaubten Transformationen zunehmend eingeschränkt sind und daher die Disparitäten nicht mehr so gut an die Distanzen angepasst werden können. Häufig fällt die Verschlechterung der Gütekriterien wesentlich deutlicher aus. Im hier analysierten Beispiel ist jedoch auch die Güte der Lösung der Intervall-MDS hervorragend einzustufen.

Die Konfiguration der Objekte in der Intervall-MDS (Abbildung 4.2.2.2) unterscheidet sich optisch so gut wie gar nicht von der Konfiguration der ordinalen MDS.

Konfiguration des abgeleiteten Stimulus

Euklidisches Distanzmodell

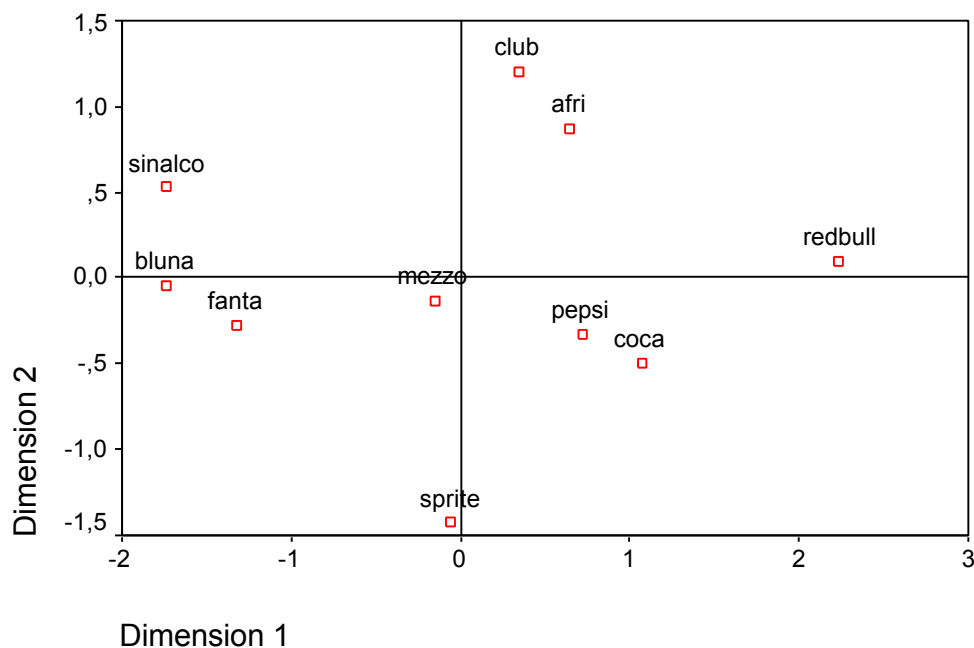


Abb. 4.2.2.2: SPSS-Graphik: Konfiguration der Objekte im zweidimensionalen Raum (Intervall-MDS)

Die Interpretation dieser Konfiguration in Abbildung 4.2.2.2 entspricht daher der Interpretation auf Seite 39.

Als einziges Diagramm zur Beurteilung der Güte der Anpassung wird bei der Intervall-MDS das Streudiagramm mit linearer Anpassung ausgegeben (Abbildung 4.2.2.3).

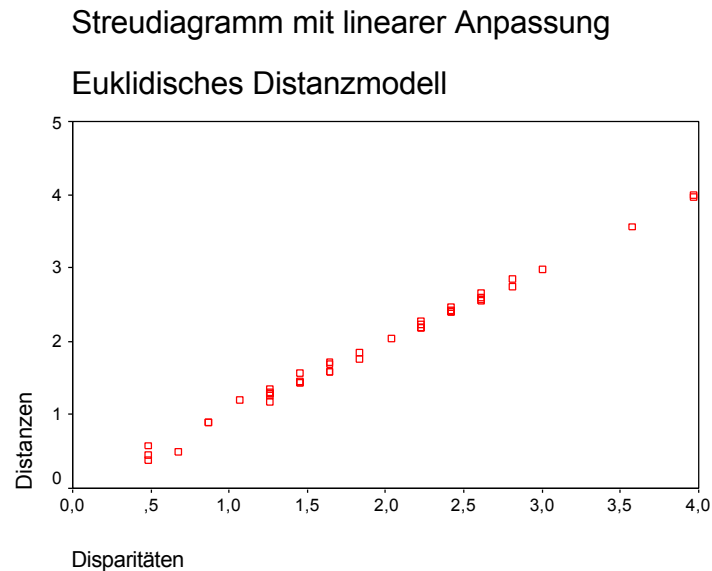


Abb. 4.2.2.3: Streudiagramm mit linearer Anpassung (Intervall-MDS).

Das Streudiagramm mit nicht-linearer Anpassung (Beobachtungen vs. Distanzen) wird nicht angezeigt, da die Disparitäten bei der Intervall-MDS lineare Transformationen der Beobachtungen sind. Daher unterscheidet sich das Streudiagramm mit linearer Anpassung (Disparitäten vs. Distanzen) kaum vom Diagramm mit nicht-linearer Anpassung (vgl. Abbildung 4.1.3.3 auf Seite 39).

4.3 Wiederholte MDS

In der Regel besteht die ursprüngliche Datenbasis einer MDS aus mehr als einer Ähnlichkeitsmatrix, da zumeist Ähnlichkeitsurteile von mehr als einer Person erhoben werden. Sofern nicht unterstellt werden kann oder soll, dass sich die Wahrnehmungsräume der Personen gleichen, ist zunächst von der Aggregation der Daten zu einer Matrix (wie unter 4.1 und 4.2) abzusehen.

Um die Äquivalenz der Wahrnehmungsräume zu überprüfen, kann eine wiederholte MDS auf Basis der individuell erhobenen Ähnlichkeitsmatrizen durchgeführt werden. Die Dateneingabe unterscheidet sich entsprechend, wie im folgenden Abschnitt dargestellt wird.

4.3.1 Dateneingabe

Im Dateneditor von SPSS werden die Ähnlichkeitsmatrizen für alle Personen untereinander eingegeben. Das zusätzliche Anlegen einer Variable mit der Versuchspersonennummer (VPN) dient der Zuordnung der Matrizen zu den Personen, ist jedoch für den Programmablauf der MDS nicht erforderlich. Abbildung 4.3.1.1 enthält die Ähnlichkeitsmatrizen der zwei Befragten des bereits in den letzten Abschnitten erläuterten Getränkemarkenbeispiels.

	vpn	labels	coca	pepsi	fanta	afri	sprite	bluna	sinalce	club	redbull	mezzo
1	1,00	coca	,00	,	,	,	,	,	,	,	,	,
2	1,00	pepsi	1,00	,00	,	,	,	,	,	,	,	,
3	1,00	fanta	6,00	5,00	,00	,	,	,	,	,	,	,
4	1,00	afri	5,00	4,00	6,00	,00	,	,	,	,	,	,
5	1,00	sprite	4,00	4,00	5,00	9,00	,00	,	,	,	,	,
6	1,00	bluna	7,00	6,00	2,00	7,00	7,00	,00	,	,	,	,
7	1,00	sinalce	8,00	7,00	3,00	6,00	8,00	1,00	,00	,	,	,
8	1,00	club	7,00	6,00	7,00	1,00	10,00	7,00	6,00	,00	,	,
9	1,00	redbull	3,00	3,00	9,00	5,00	8,00	10,00	10,00	6,00	,00	,
10	1,00	mezzo	3,00	2,00	3,00	4,00	5,00	4,00	4,00	5,00	6,00	,00
11	2,00	coca	,00	,	,	,	,	,	,	,	,	,
12	2,00	pepsi	1,00	,00	,	,	,	,	,	,	,	,
13	2,00	fanta	6,00	5,00	,00	,	,	,	,	,	,	,
14	2,00	afri	2,00	1,00	5,00	,00	,	,	,	,	,	,
15	2,00	sprite	3,00	2,00	3,00	3,00	,00	,	,	,	,	,
16	2,00	bluna	7,00	6,00	1,00	6,00	4,00	,00	,	,	,	,
17	2,00	sinalce	7,00	6,00	1,00	6,00	5,00	1,00	,00	,	,	,
18	2,00	club	2,00	2,00	4,00	1,00	3,00	5,00	5,00	,00	,	,
19	2,00	redbull	3,00	4,00	9,00	4,00	6,00	10,00	10,00	5,00	,00	,
20	2,00	mezzo	3,00	2,00	3,00	2,00	1,00	4,00	4,00	2,00	6,00	,00

Abb. 4.3.1.1: Dateneingabe bei mehr als einer Matrix.

Die individuelle Erfassung der Rohdaten in SPSS (wie hier gezeigt) ist auf jeden Fall ratsam, da sich aus dieser Datei im nachhinein leicht eine aggregierte Rohdatenmatrix erzeugen lässt (über die Menübefehle „Daten → Aggregieren“).

4.3.2 Aufruf der Prozedur und Optionen

Der Aufruf der MDS erfolgt analog zu der bereits oben geschilderten Vorgehensweise. SPSS erkennt automatisch, dass es sich um mehr als eine Ähnlichkeitsmatrix handelt und reagiert entsprechend. Die weitere Vorgehensweise wird am Beispiel einer ordinalen wiederholten MDS erläutert. Sämtliche Einstellungen der Prozedur entsprechen den Einstellungen der normalen ordinalen MDS (siehe Abschnitt 4.1.2).

4.3.3 Interpretation des Output

Der Textoutput der ordinalen wiederholten MDS entspricht bis auf vier Abweichungen der ordinalen MDS:

Alscal Procedure Options

Data Options-

```

Number of Rows (Observations/Matrix). 10
Number of Columns (Variables) . . . 10
Number of Matrices . . . . . 2
Measurement Level . . . . . Ordinal
Data Matrix Shape . . . . . Symmetric
Type . . . . . Dissimilarity
Approach to Ties . . . . . Leave Tied
Conditionality . . . . . Matrix
Data Cutoff at . . . . . ,000000

```

Die Anzahl der Matrizen (Number of Matrices) beträgt nun 2. Alle anderen Angaben zu den Optionen bleiben gleich.

Als Rohdatenmatrizen (und später auch als Disparitäten) werden nun entsprechend zwei Matrizen angezeigt.

Die wichtigste Änderung betrifft die Ausgabe von STRESS und RSQ:

Stress and squared correlation (RSQ) in distances

RSQ values are the proportion of variance of the scaled data (disparities) in the partition (row, matrix, or entire data) which is accounted for by their corresponding distances. Stress values are Kruskal's stress formula 1.

Matrix	Stress	RSQ	Matrix	Stress	RSQ
1	,106	,927	2	,177	,799

Averaged (rms) over matrices
 Stress = ,14583 RSQ = ,86331

Beide Maße (STRESS und RSQ) werden für jede einzelne Matrix ausgegeben. Im Beispiel ist zu erkennen, dass sowohl STRESS und RSQ deutlich schlechter geworden sind. Insbesondere die Anpassung der zweite Matrix an die Konfiguration ist mit einem STRESS von 0,177 und einem RSQ von 0,799 nicht mehr besonders optimal. Abschließend wird ein gemittelter STRESS über alle Matrizen ausgegeben, der zur Beurteilung der Anpassung aller Matrizen an die gemeinsame Konfiguration herangezogen werden kann. Auch dieser ist im Beispiel mit einem Betrag von 0,14583 nicht mehr besonders gut, aber noch akzeptabel (da kleiner als 0,2).

Konfiguration des abgeleiteten Stimulus

Euklidisches Distanzmodell

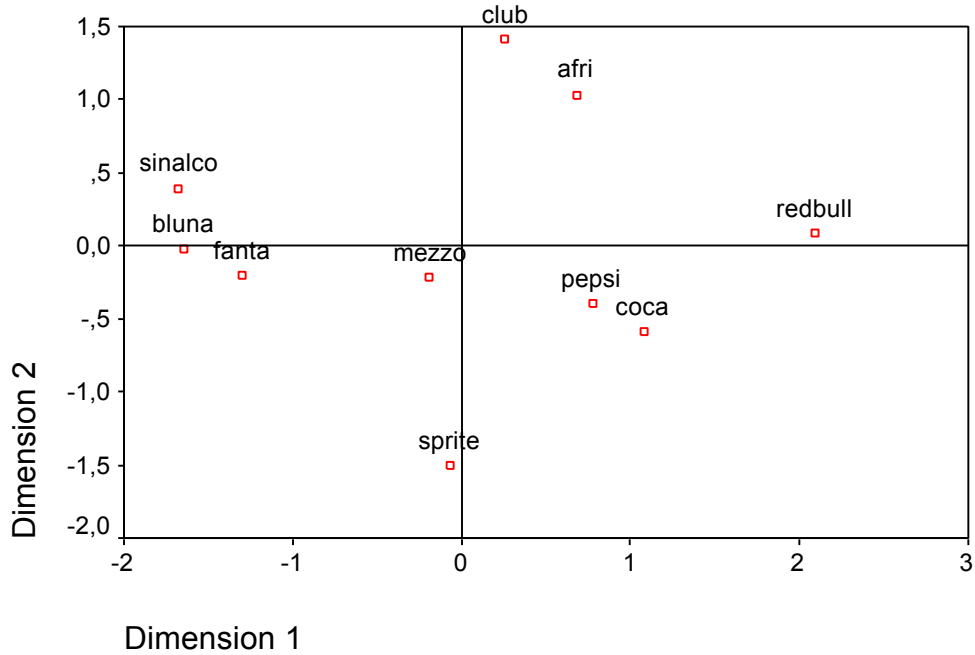


Abb. 4.3.3.1: Konfiguration der ordinalen wiederholten MDS.

In Abbildung 4.3.3.1 ist die Konfiguration der ordinalen wiederholten MDS dargestellt. Ein Vergleich mit der Lösung der ordinalen MDS, angewendet auf die aggregierte Datenmatrix, zeigt, dass sich die beiden Konfigurationen kaum voneinander unterscheiden.

Das folgende Streudiagramm mit linearer Anpassung (Abbildung 4.3.3.2) ist ebenfalls vergleichbar mit der einfachen ordinalen MDS, mit dem einzigen Unterschied, dass nun doppelt so viele Punkte ($2 \times 45 = 90$) im Diagramm dargestellt sind.

Streudiagramm mit linearer Anpassung

Euklidisches Distanzmodell

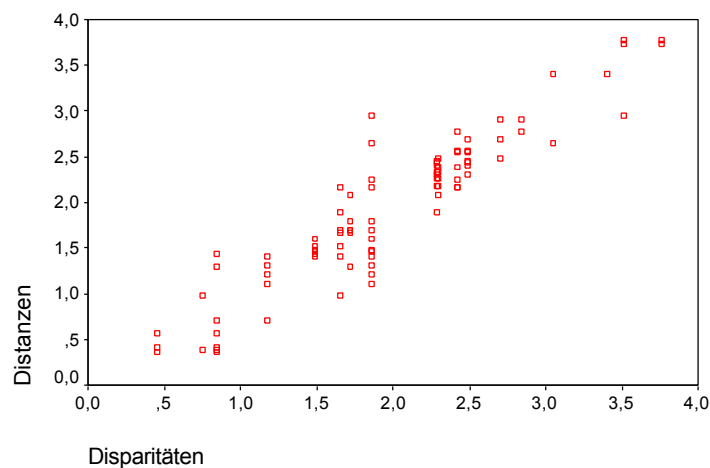


Abb. 4.3.3.2: Streudiagramm mit linearer Anpassung

Im Vergleich zum Streudiagramm mit linearer Anpassung der ordinalen MDS (Abbildung 4.1.3.2) streuen die Punkte in Abbildung 4.3.3.2 wesentlich stärker um die Gerade durch den Ursprung der Graphik, worauf der deutlich schlechtere Betrag des RSQ zurück zu führen ist.

Die Ausgabe der individuellen Streudiagramme mit nicht-linearer Anpassung wird bei mehr als einer Matrix unterdrückt. Es werden abschließend für jedes Subjekt die Transformations-Streudiagramme dargestellt (Abbildung 4.3.3.3 und 4.3.3.4).

Transformations-Streudiagramm für Subjekt 1

Euklidisches Distanzmodell

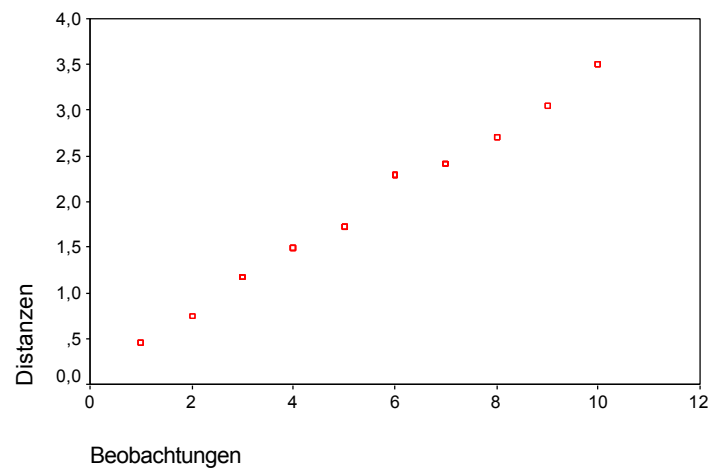


Abb. 4.3.3.3: Transformationsstreudiagramm für Subjekt 1 (Disparitäten vs. Beobachtungen)

Transformations-Streudiagramm für Subjekt 2

Euklidisches Distanzmodell

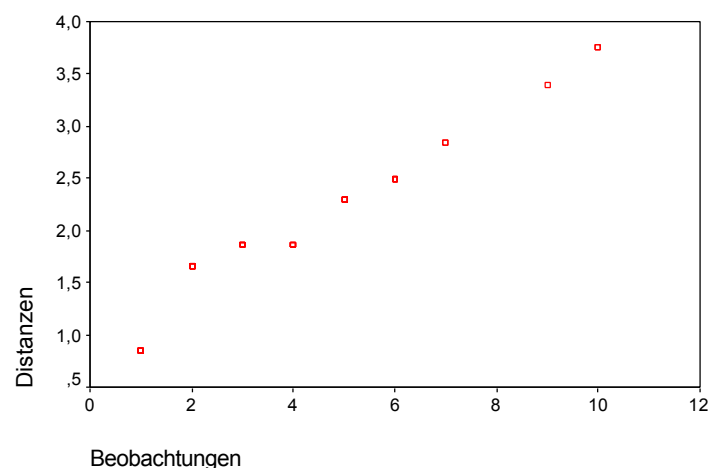


Abb. 4.3.3.4: Transformationsstreudiagramm für Subjekt 2 (Disparitäten vs. Beobachtungen)

In Abbildung 4.3.3.3 und 4.3.3.4 ist wieder zu beachten, dass die Beschriftung von SPSS falsch durchgeführt wird. Geplottet werden nicht Distanzen, sondern Disparitäten gegen Beobachtungen.

4.4 INDSCAL

Mit der wiederholten MDS (letzter Abschnitt) wird die starke Hypothese getestet, die Wahrnehmungsräume aller Subjekte seien äquivalent und daher in ein und derselben Konfiguration darstellbar. Beim INDSCAL-Verfahren wird diese Hypothese in einer schwächeren Form getestet. Als Datenbasis dienen wieder die individuellen Matrizen der Befragten. Es wird ebenfalls nur eine Konfiguration erzeugt. Allerdings wird zugelassen, dass diese „gemeinsame“ Konfiguration entlang ihrer Dimensionen individuell gestaucht oder gestreckt wird. Inhaltlich lässt sich diese individuelle Gewichtung der Dimensionen durch die Subjekte interpretieren als die individuelle Wichtigkeit der Dimensionen des Wahrnehmungsraumes für die Unterscheidbarkeit der Objekte.

Das INDSCAL-Verfahren wird nun ebenfalls am Beispiel der Beurteilung von Getränkemarken und einer ordinalen MDS erläutert.

4.4.1 Aufruf der Prozedur und Optionen

Da die Datenbasis für das INDSCAL-Verfahren der wiederholten MDS entspricht (vgl. Abschnitt 4.3.1), wenden wir uns direkt dem Aufruf der Prozedur zu. Von der einfachen ordinalen MDS unterscheidet sich der Aufruf von INDSCAL nur in einem einzigen Punkt, der in Abbildung 4.4.1.1 dargestellt ist.

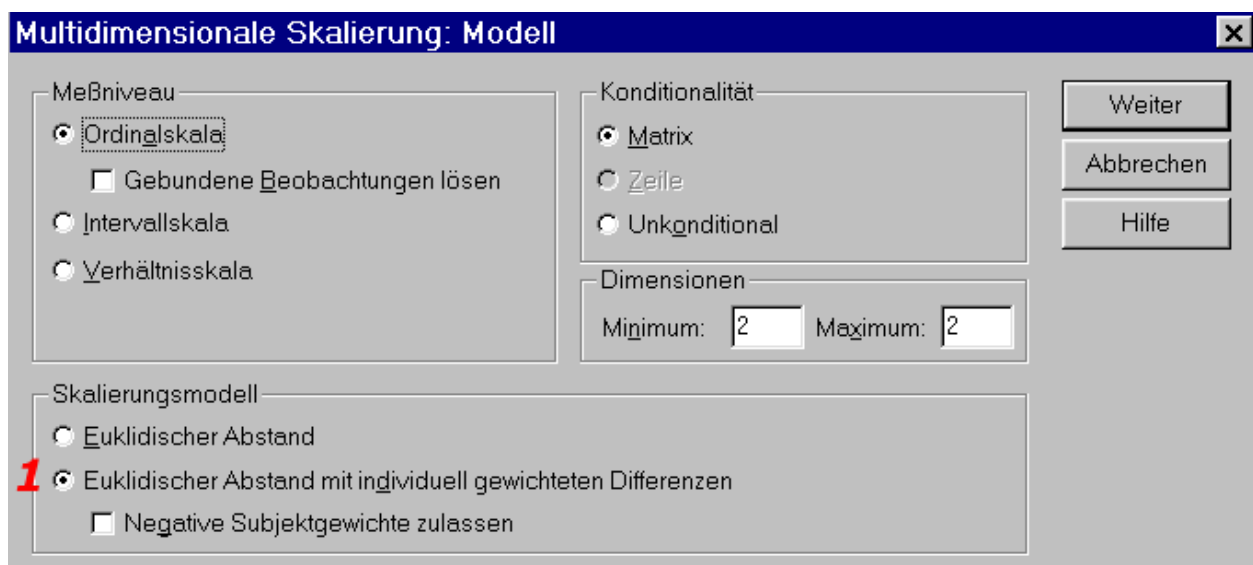


Abb. 4.4.1.1: Modell-Optionen für das INDSCAL-Verfahren.

Als Skalierungsmodell wird unter der Option „Modell“ nicht der euklidische Abstand, sondern der euklidische Abstand mit individuell gewichteten Differenzen ausgewählt. Dadurch wird das INDSCAL-Verfahren aufgerufen. Dies sollte gleichzeitig daran erinnern, dass in der Konfiguration nun nicht mehr die rotationsinvariante euklidische Metrik gilt (vgl. Seite 14 in Abschnitt 1.4). Die Option „Negative Subjektgewichte zulassen“ kann ausgewählt werden, wenn neben einer reinen Stauchung oder Streckung der Dimensionen auch die Spiegelung der Konfiguration entlang einer oder beider Dimensionen erlaubt sein soll.

4.4.2 Interpretation des Output

Die Auswahl des INDSCAL-Verfahrens wird in den von SPSS ausgegebenen Optionen durch den Eintrag `Model: Indscal` dokumentiert. Bei diesem Verfahren ist auch die Angabe `Negative Weights: Not Permitted` von Relevanz. Da negative Gewichte nicht erlaubt sind, werden die Dimensionen zwar individuell gestaucht oder gestreckt, nicht aber gespiegelt.

Alscal Procedure Options

Data Options-

```
Number of Rows (Observations/Matrix). 10
Number of Columns (Variables) . . . 10
Number of Matrices . . . . . 2
Measurement Level . . . . . Ordinal
Data Matrix Shape . . . . . Symmetric
Type . . . . . Dissimilarity
Approach to Ties . . . . . Leave Tied
Conditionality . . . . . Matrix
Data Cutoff at . . . . . ,000000
```

Model Options-

```
Model . . . . . Indscal
Maximum Dimensionality . . . . . 2
Minimum Dimensionality . . . . . 2
Negative Weights . . . . . Not Permitted
```

Der unmittelbar folgende Output entspricht den bisherigen Analysen und wird daher nicht mehr näher erläutert.

Wie bei der wiederholten MDS wird auch im Zuge des INDSCAL-Verfahrens für jede einzelne der Rohdatenmatrizen STRESS und RSQ berechnet, gefolgt von einem gemittelten STRESS und RSQ für die Anpassung der Konfiguration über alle Matrizen.

Stress and squared correlation (RSQ) in distances

RSQ values are the proportion of variance of the scaled data (disparities) in the partition (row, matrix, or entire data) which is accounted for by their corresponding distances. Stress values are Kruskal's stress formula 1.

Matrix	Stress	RSQ	Matrix	Stress	RSQ
1	,040	,988	2	,042	,993

Averaged (rms) over matrices
Stress = ,04123 RSQ = ,99051

Vergleichen wir die STRESS- und RSQ-Werte dieser Analyse mit der vorangegangenen wiederholten MDS, so fällt die Anpassung nun erheblich günstiger aus. Der STRESS ist sowohl für die individuellen Matrizen als auch im Durchschnitt kleiner als 0,05 und damit beinahe genauso gut, wie der STRESS der einfachen ordinalen MDS.

Die individuelle Gewichtung der Dimensionen hat also gegenüber der wiederholten MDS zu einer deutlich besseren Anpassung der Konfiguration an die Rohdaten geführt. Damit wird die Hypothese gestützt, dass die befragten Subjekte einen ähnlichen Wahrnehmungsraum der Getränkemarken besitzen, aber die den Raum konstituierenden Dimensionen unterschiedlich gewichten.

Im Anschluß an die (hier nicht dargestellten) Koordinaten der Objekte in der Konfiguration werden die individuellen Gewichte der Dimensionen ausgegeben:

Subject weights measure the importance of each dimension to each subject.
Squared weights sum to RSQ.

A subject with weights proportional to the average weights has a weirdness of zero, the minimum value.

A subject with one large weight and many low weights has a weirdness near one.

A subject with exactly one positive weight has a weirdness of one, the maximum value for nonnegative weights.

Subject Weights			
Subject Number	Weirdness	Dimension	
		1	2
1	,4196	,7932	,5993
2	,7848	,9943	,0629
Overall importance of each dimension:		,8089	,1816

Die beiden Subjekte 1 und 2 gewichten die Dimensionen des Lösungsraumes offenbar unterschiedlich. Der Faktor, mit dem die Dimensionen individuell angepasst werden, ergibt sich aus der Quadratwurzel der Gewichte:

- Um die Konfiguration optimal an die Rohdatenmatrix von Subjekt 1 anzupassen, wird Dimension 1 mit dem Faktor 0,891 (= Wurzel(0,7932)) multipliziert. Dimension 2 wird mit dem Faktor 0,774 (= Wurzel(0,5993)) gewichtet.
- Die Konfiguration ist optimal an die Rohdaten von Subjekt 2 angepaßt, wenn Dimension 1 mit 0,997 und Dimension 2 mit 0,251 multipliziert wird.

Setzt man die Quadratwurzel der Gewichte der Dimension zueinander ins Verhältnis, so ergibt sich, dass Subjekt 1 die erste Dimension ungefähr 1,15-mal ($0,891/0,774$) so stark gewichtet wie die zweite Dimension. Dies ist ein nahezu ausgewogenes Verhältnis. Subjekt 2 differenziert dagegen die Getränkemarken wesentlich einseitiger auf Basis von Dimension 1. Die erste Dimension wird hier 3,97-mal ($0,997/0,251$) so stark gewichtet wie Dimension 2. Wichtig ist, dass nur dieses Verhältnis der Gewichte zueinander herangezogen werden darf, um die Gewichtung durch die Subjekte zu interpretieren. Nicht erlaubt ist dagegen, die Gewichte der Subjekte jeweils einer Dimension gegenüber zu stellen.

Diese Beschränkung hat folgenden Grund: Die Gewichte geben nicht nur Auskunft darüber, wie stark ein Subjekt die Dimensionen gewichtet, sondern gleichzeitig, wie gut die Konfiguration nach der Gewichtung an die Rohdatenmatrix des Subjekts angepaßt ist. Auf diese Interpretation der Gewichte wird bereits von SPSS im Textoutput hingewiesen. So sumieren sich die Quadrate der Dimensionsgewichte für jedes Subjekt zu RSQ. Subjekt 1 hat ein RSQ von $0,988 = 0,629 + 0,359 = 0,7932^2 + 0,5993^2$. Subjekt 2 hat ein RSQ von $0,993 = 0,9890 + 0,004 = 0,9943^2 + 0,0629^2$ (vgl. auch Abbildung 4.4.2.3).

Angenommen, zwei Subjekte hätten die folgenden Gewichte für zwei Dimensionen :

	Dim 1	Dim 2
Subjekt A	0,707	0,707
Subjekt B	0,500	0,500

Es ist leicht erkennbar, dass beide Subjekte A und B die Dimensionen der Konfiguration im gleichen Verhältnis zueinander gewichten (1 : 1). Die Konfiguration ist also für Subjekt A und B die gleiche. Allerdings gelingt die Anpassung der Konfiguration für Subjekt A perfekt ($RSQ = 1,0 = 0,707^2 + 0,707^2$), für Subjekt B dagegen sehr schlecht ($RSQ = 0,5 = 0,5^2 + 0,5^2$). Eine Interpretation im Sinne der Aussage „*Subjekt 1 gewichtet Dimension 1 mit 0,707 stärker als Subjekt 2 mit 0,500*“ wäre dagegen offensichtlich falsch!

Die Ausgabe der *Weirdness* für jedes Subjekt dient dazu, Subjekte mit einer ungewöhnlichen Gewichtung der Dimensionen zu identifizieren. Je näher der *Weirdness*-Index an 1 liegt, desto stärker weicht der Vektor der Gewichtung der Dimensionen durch dieses Subjekt vom durchschnittlichen Vektor der Gewichte ab. Auf diese Weise lassen sich Ausreisser identifizieren, die eine *Weirdness* nahe 1 aufweisen. Solche Fälle könnten dann in weiteren Analysen getrennt untersucht werden.

Konfiguration des abgeleiteten Stimulus

Euklid. Distanzmodell für individuelle Unterschiede (gewichtet)

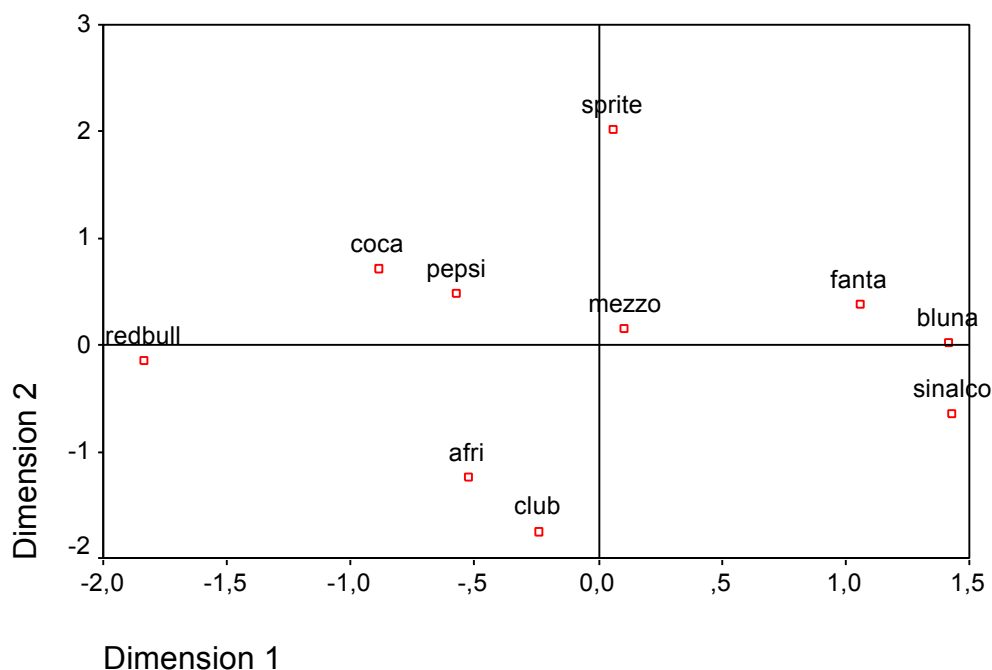


Abb. 4.4.2.1: Gemeinsame Konfiguration des INDSCAL-Verfahrens (ordinale MDS)

Die gemeinsame Konfiguration als Lösung des INDSCAL-Verfahrens (Abbildung 4.4.2.1) ähnelt sehr stark den Konfigurationen der voran gegangenen Beispiele. Der Raum ist lediglich entlang Dimension 1 und Dimension 2 gespiegelt. Entlang Dimension 1 nimmt der Grad der anregenden Wirkung der Getränke von links nach rechts ab. Die zweite Dimension läßt sich von oben nach unten vielleicht im Sinne abnehmender Konsumhäufigkeit interpretieren. Es sei an dieser Stelle nochmals daran erinnert, dass die hier dargestellte gemeinsame Konfiguration nicht in dieser Form auf die Subjekte übertragen werden kann, sondern dass für jedes Subjekt auf Grundlage der Gewichte eine eigene Konfiguration existiert. Dieser Umstand unterscheidet das INDSCAL-Verfahren von allen anderen Anwendungen der MDS.

Die abgeleiteten Subjektgewichte sind graphisch in der nächsten Abbildung 4.4.2.2 dargestellt.

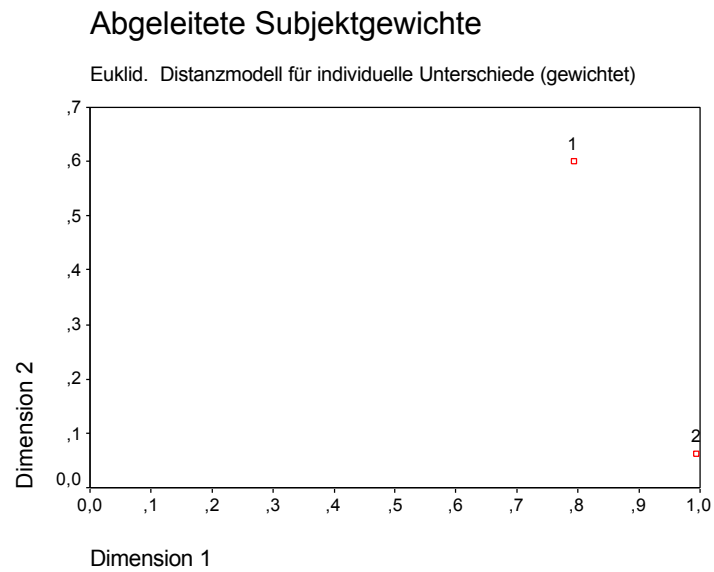


Abb. 4.4.2.2: Abgeleitete Subjektgewichte (original SPSS-Graphik)

Da die von SPSS erzeugte Graphik nur eingeschränkt bearbeitet werden kann und optische Nachteile aufweist, haben wir die Graphik der abgeleiteten Subjektgewichte mit einem anderen Programm erneut erzeugt (vgl. Abbildung 4.4.2.3)

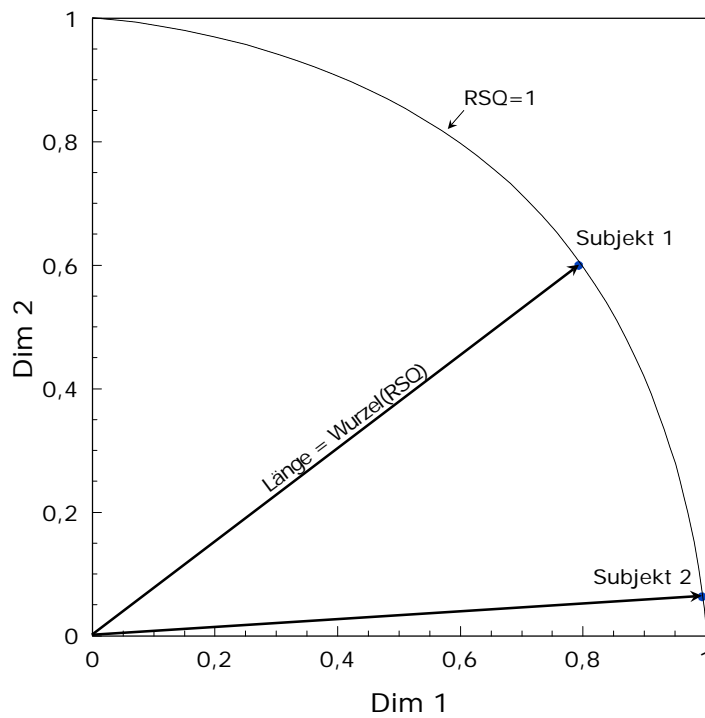


Abb. 4.4.2.3: Abgeleitete Subjektgewichte (ergänzte graphische Darstellung)

In Abbildung 4.4.2.3 wurden die Punkte für Subjekt 1 und 2 mit einem Vektor durch den Ursprung verbunden. Die Winkel dieser Vektoren (Verhältnisse der Gewichte; s.o.) lassen sich miteinander vergleichen. Die Länge dieser Vektoren entspricht der Wurzel aus RSQ. Je länger ein Vektor ist, desto höher ist das RSQ der entsprechenden Subjektmatrix. Die maximale Länge der Vektoren ($RSQ = 1$) ist durch den Halbkreis in der Graphik angedeutet. Subjekte, die weiter innen im Halbkreis liegen, weisen eine schlechtere Anpassung der Konfiguration an die Rohdaten auf. In diesem Beispiel ist allerdings die Anpassung der Konfiguration beider Subjekte hervorragend.

Inhaltlich wird deutlich, dass Subjekt 2 die Differenzierung der Getränkemarken vor allem auf Basis des Grades der anregenden Wirkung vornimmt (Dim 1), während Subjekt 1 neben der anregenden Wirkung auch die Konsumhäufigkeit zur Beurteilung der Getränkemarken heranzieht (vgl. die Interpretation der Konfiguration weiter oben).

Die folgenden Streudiagramme mit linearer Anpassung und monotoner Transformation (Abbildung 4.4.2.4, 4.4.2.5, 4.4.2.6) wurden bereits in den letzten Beispielen erläutert. Vor allem im Vergleich zur Lösung der wiederholten MDS des vorangegangenen Abschnitts fällt die deutlich geringere Streuung Punkte im Streudiagramm mit linearer Anpassung um die Gerade durch den Ursprung auf.

Streudiagramm mit linearer Anpassung

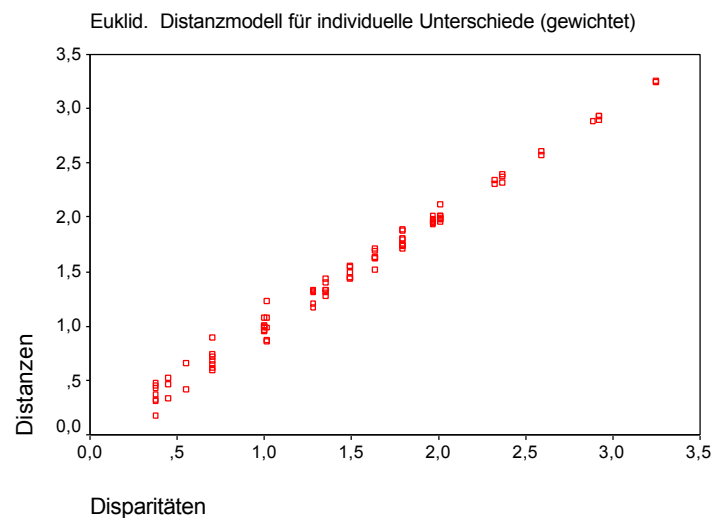


Abb. 4.4.2.4: Streudiagramm mit linearer Anpassung (INDSCAL)

Transformations-Streudiagramm für Subjekt 1

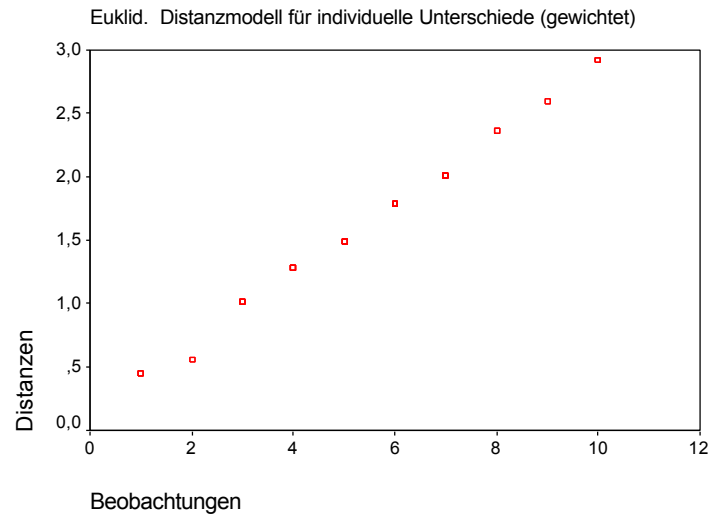


Abb. 4.4.2.5: Transformations-Streudiagramm für Subjekt 1 (INDSCAL)

Transformations-Streudiagramm für Subjekt 2

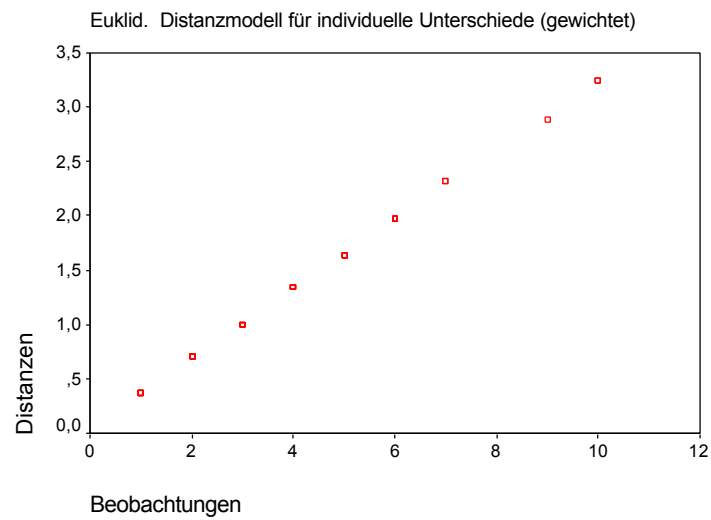


Abb. 4.4.2.6: Transformations-Streudiagramm für Subjekt 2 (INDSCAL)

5. Multidimensionales Entfalten²⁷ (MDU): Das Idealpunktmodell zur Interpretation der Konfiguration

Wie bereits in Abschnitt 1.3.2 ausgeführt, lassen sich nicht nur quadratische Ähnlichkeitsmatrizen, sondern auch rechteckige Ähnlichkeitsmatrizen multidimensional skalieren. Das in diesem Zusammenhang verwendete Verfahren wird als multidimensionales Entfalten (MDU) bezeichnet. Die Interpretation der Konfiguration einer MDU folgt dem Idealpunktmodell.

Als (hypothetisches) Beispiel zur Datenanalyse betrachten wir den Parteienraum:

Sechs Personen wurde die folgende Aufgabe gestellt: "Ordnen Sie die folgenden sechs Parteien in der Reihenfolge der Sympathie, die Sie für die Partei empfinden: CDU, CSU, SPD, FDP, Grüne, PDS". Auf diese Weise wurde die folgende Datenmatrix erhoben:

Tab. 5.1: Rangordnung von Parteien nach Sympathie

	Rangplatz					
	CDU	CSU	SPD	FDP	Grüne	PDS
Person 1	2	1	4	3	5	6
Person 2	1	2	4	3	5	6
Person 3	2	4	3	1	5	6
Person 4	5	6	1	2	3	4
Person 5	5	6	2	4	1	3
Person 6	5	6	3	4	2	1

Person 1 ist die CSU am stärksten sympathisch, für die PDS empfindet Person 1 am wenigsten Sympathie. Person 6 ist genau der gegenteiligen Auffassung und empfindet am meisten Sympathie für die PDS ...

Um auf Grundlage dieser Datenbasis (rechteckige zeilenkonditionale Matrix) eine angemessene Konfiguration des gemeinsamen Wahrnehmungsraumes der sechs befragten Personen zu erzeugen, eignet sich hier das Idealpunktmodell. Im Idealpunktmodell wird davon ausgegangen, dass jede Person einen individuellen Idealpunkt im Wahrnehmungsraum hat. Die Objekte der Konfiguration liegen um so näher an diesem Punkt, je positiver sie von der Person beurteilt werden. Weniger "ideale" Objekte werden entsprechend weiter weg vom Idealpunkt der Person abgebildet.

²⁷ engl.: multidimensional unfolding.

5.1 Dateneingabe

Die Daten für das Idealpunktmodell werden entsprechend der Tabelle 5.1 in den Dateneditor von SPSS übernommen (auf eine ausführliche Darstellung soll an dieser Stelle verzichtet werden).

5.2 Aufruf der Prozedur und Optionen

Die MDU wird genauso wie die MDS über das Menü (Analysieren → Skalieren → Multidimensionale Skalierung) aufgerufen. Die nun zunächst folgenden Schritte sind in Abbildung 5.2.1 dargestellt.

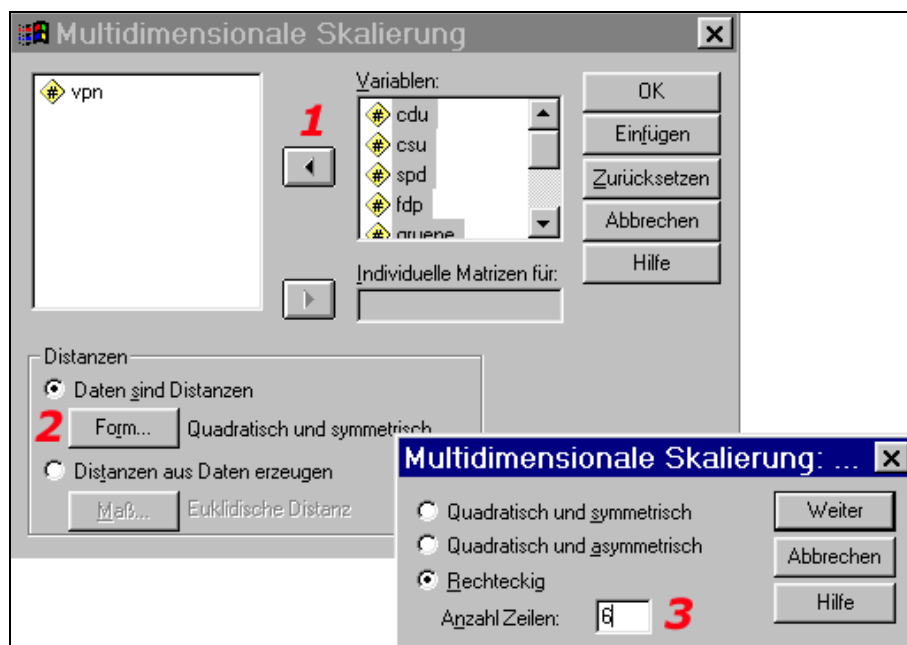


Abb. 5.2.1: Anforderung einer MDU

Zunächst werden (wie bei der MDS) die Objekte nach links in das Variablenfenster übertragen (1). Als nächstes muß angegeben werden, welche Form die Datenmatrix hat (2). Die Voreinstellung "quadratisch und symmetrisch" ist hier nicht korrekt. Stattdessen muß angegeben werden, dass es sich um eine rechteckige Datenmatrix mit 6 Zeilen (pro Person eine!) handelt (3).

Zusätzlich muß die Datenmatrix unter dem Unterpunkt "Modell" als "zeilenkonditional" ausgewiesen werden (vgl. Abbildung 4.1.2.4 auf Seite 32). An dieser Stelle wird auch die Dimensionalität der Konfiguration festgelegt. In diesem Beispiel kann die Hypothese aufgestellt werden, dass der Parteienraum eindimensional (links - rechts) ist. Deshalb tragen wir hier als minimale und maximale Anzahl an Dimensionen eine 1 ein.²⁸

Alle übrigen Einstellungen können analog zu den bisherigen Beispielen vorgenommen werden.

²⁸ Als Beispiel für eine zweidimensionale MDU siehe Abschnitte 1.3.2 und 5.4.

5.3 Interpretation des Output

Da wir Zeilenkonditionalität vorgegeben haben, erhalten wir nun für jede Zeile (=Person) einen eigenen STRESS- und RSQ-Wert.

```

Matrix 1
(Row Stimuli Only)
Stimulus  Stress  RSQ  Stimulus  Stress  RSQ
1          ,009  1,000  2          ,000  1,000
3          ,043  ,998  4          ,092  ,992
5          ,074  ,995  6          ,000  1,000
Averaged (rms) over stimuli
Stress =  ,051      RSQ =  ,997

```

Zusätzlich werden gemittelter STRESS- und RSQ-Wert zur Beurteilung der Konfiguration insgesamt ausgegeben. Sowohl die individuellen, als auch die gemittelten Werte weisen auf eine hinreichend gelungene Anpassung der Konfiguration hin.

```

Stimulus Coordinates
Stimulus  Dimension
Number    Name      1
Column
1         CDU     -1,0553
2         CSU     -1,4070
3         SPD      ,4073
4         FDP     -,2339
5         GRUENE  1,0543
6         PDS     1,3373
Row
1         -1,2103
2         -1,1080
3         -,5532
4         ,5423
5         1,0263
6         1,2002

```

Als Spaltenkoordinaten (Column) erhalten wir wie gehabt die Koordinaten der Objekte im Wahrnehmungsraum. Zusätzlich erhalten wir als Zeilenkoordinaten (Row) die Lage der Idealpunkte der Personen im Wahrnehmungsraum. Abbildung 5.3.1 enthält die Konfiguration in graphischer Darstellung.

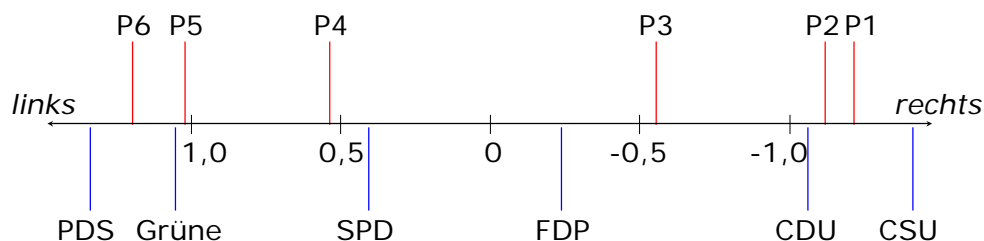


Abb. 5.3.1: Konfiguration der eindimensionalen ordinalen MDU zur Parteiensympathie.

Die graphische Darstellung der Objekte und Personen in ein und derselben Konfiguration läßt sich als Idealpunktmodell interpretieren: Der Idealpunkt von Person 3 (P3) liegt im Parteien-

raum zwischen FDP und CDU (näher an der FDP), der Idealpunkt von Person 4 (P4) liegt nahe bei der SPD, usw.

Der Vorgang der Entfaltung wird am Beispiel der Person 3 in der folgenden Abbildung 5.3.2 dargestellt:

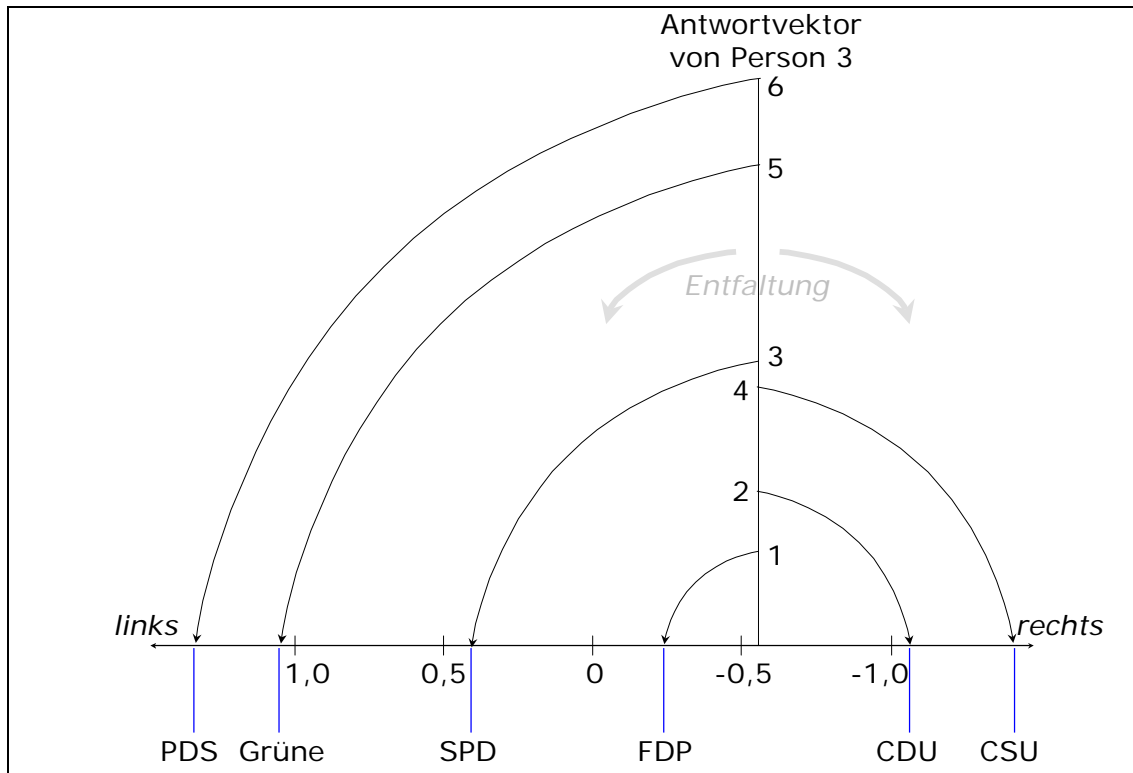


Abb. 5.3.2: "Entfaltung" des Antwortvektors von Person 3 auf die Dimensionsachse der Konfiguration.

Der Antwortvektor von Person 3 steht senkrecht zur Konfiguration im Idealpunkt und wird in die Konfiguration "entfaltet". Die Projektion des Antwortvektors auf die Konfiguration ist für diese Person nur an einer Stelle leicht inkonsistent. Die SPD wurde von Person 3 an dritter Stelle genannt, liegt aber vom Idealpunkt aus gesehen an vierter Stelle hinter der CSU. Entsprechend ist der STRESS für Person 3 auch nicht Null, sondern 0,043 (siehe Output oben).

5.4 ALLBUS 1994: Parteiensympathie und Links-Rechts-Selbsteinstufung im West-Ost-Vergleich

Zum Abschluß dieses Kapitels soll die Aussagekraft einer MDU an einem empirischen Beispiel auf Basis des für die deutsche Wohnbevölkerung repräsentativen ALLBUS 1994 demonstriert werden.

Im Zuge der Erhebung des ALLBUS 1994 wurde die Frage nach der Sympathie für die Parteien des Bundestages gestellt ("Wie sympathisch ist Ihnen die CDU [bzw. CSU, SPD, FDP, Bündnis90/Grüne, PDS, Republikaner]?"). Die Personen konnten auf einer Skala von -5 (sehr unsympathisch) bis +5 (sehr sympathisch) antworten (Sympathiethermometer). Ebenfalls erho-

ben wurde die Links-Rechts-Selbsteinstufung der Personen auf einer Skala von 1 (links) bis 10 (rechts).

In Vorbereitung zur eigentlichen MDU wurden die Sympathiethermometer der Parteien für jede Ausprägung der Links-Rechts-Selbsteinstufung aggregiert - und zwar für Ost- und Westdeutschland getrennt. Zuvor wurden die Sympathiethermometer recodiert, so dass niedrige Ausprägungen Sympathie und hohe Ausprägungen Antipathie anzeigen. Der so erzeugte Datensatz enthält 20 Zeilen mit Mittelwerten für jedes Sympathiethermometer. Für jede der 10 Ausprägungen der Links-Rechts-Selbsteinstufung existiert je eine Zeile für Ost- und Westdeutschland. Je kleiner ein Mittelwert in einer Zeile ist, desto sympathischer wird diese Partei von den Befragten dieser Zeile im Durchschnitt empfunden (siehe Tabelle 5.4.1).

Tab. 5.4.1: Erhebungsgebiet, Links-Rechts-Selbsteinstufung* und Mittelwerte der Parteiensympathie**

Gebiet	Links-Rechts	CDU	SPD	CDU	FDP	B90/Grüne	Reps	PDS
West	1	6,16	2,72	6,74	5,72	4,95	8,91	8,00
West	2	6,85	3,11	7,66	6,32	3,81	9,39	7,51
West	3	6,50	3,47	7,38	6,28	4,09	9,38	7,99
West	4	6,17	3,47	6,92	5,83	4,38	9,35	8,17
West	5	5,02	4,22	5,61	5,48	5,47	8,97	8,59
West	6	4,43	4,63	5,04	5,39	5,94	8,55	8,74
West	7	3,66	4,75	4,44	4,99	6,11	8,26	8,70
West	8	3,41	5,04	3,83	5,11	6,49	7,69	8,59
West	9	3,53	4,85	3,88	6,56	7,32	7,79	8,88
West	10	4,20	6,40	3,87	5,23	7,43	6,00	8,70
Ost	1	8,67	5,91	9,00	7,79	5,03	10,00	2,76
Ost	2	7,33	4,07	7,66	5,82	4,33	9,43	3,41
Ost	3	6,71	4,01	7,72	5,94	4,21	9,76	4,78
Ost	4	6,25	4,52	7,16	5,81	4,79	9,49	5,67
Ost	5	5,43	4,44	6,27	5,20	5,11	9,03	6,51
Ost	6	4,75	4,66	5,68	4,96	5,26	8,76	7,45
Ost	7	4,36	5,38	5,38	5,36	5,44	7,72	7,87
Ost	8	4,24	4,59	5,34	5,41	5,52	6,24	7,21
Ost	9	5,86	5,14	7,00	6,29	4,57	5,71	6,71
Ost	10	6,56	6,44	5,78	6,89	5,33	1,44	7,78
* von 1 (links) bis 10 (rechts)								
** je kleiner der Mittelwert, desto sympathischer die Partei								

Diese zeilenkonditionale Datenmatrix wurde nun im Zuge einer ordinalen zweidimensionalen MDU analysiert. Die Stress- und RSQ-Werte sind in der folgenden Tabelle 5.4.2 aufgeführt.

Tab. 5.4.2: Stress und RSQ der zweidimensionalen Lösung				
	West		Ost	
Links-Rechts	Stress	RSQ	Stress	RSQ
1	0,000	1,000	0,000	1,000
2	0,001	1,000	0,071	0,995
3	0,000	1,000	0,024	0,999
4	0,000	1,000	0,017	1,000
5	0,054	0,997	0,000	1,000
6	0,055	0,997	0,046	0,998
7	0,061	0,996	0,053	0,997
8	0,087	0,992	0,190	0,965
9	0,079	0,994	0,204	0,969
10	0,04	0,998	0,059	0,997

Insgesamt erscheinen die Werte für Stress und RSQ der Zeilenpunkte akzeptabel. Dies drückt sich auch im über alle Zeilen gemittelten Stress in Höhe von 0,077 aus. Lediglich die Zeilen 8 und 9 für Ostdeutschland haben relativ schlechte Stresswerte.

Die Konfiguration dieser MDU ist in vielfältiger Hinsicht interessant (siehe Abbildung 5.4.1). Aufeinanderfolgende Ausprägungen der Links-Rechts-Skala sind mit Linien verbunden, um den Verlauf der Skala von links nach rechts verfolgen zu können.

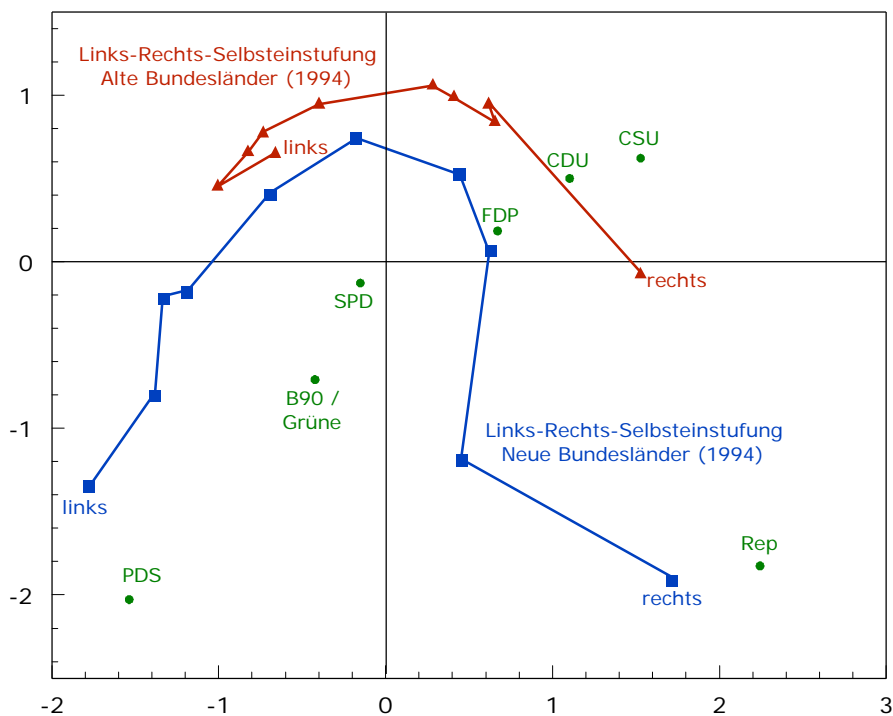


Abb. 5.4.1: Zweidimensionale MDU für Parteisympathie und Links-Rechts-Selbsteinstufung in Ost- und Westdeutschland (Daten: ALLBUS 1994).

Betrachten wir zunächst die Lage der Parteien in der Konfiguration (grüne Punkte):

- Die horizontale Dimension läßt sich im Sinne der üblichen Vorstellung vom Parteienspektrum interpretieren (siehe auch das fiktive Beispiel in den ersten Abschnitten dieses Kapitels). Die vertikale Dimension gliedert den Parteienraum (in etwa) nach konventionellen und weniger konventionellen Parteien. Die "Protestparteien" PDS und REP liegen deutlich abgesetzt von den etablierten Parteien Bündnis90/Grünen, SPD, FDP, CDU und CSU im unteren Bereich der Skala.
- Auch eine alternative Interpretation wäre denkbar: PDS, Bündnis90/Grüne, SPD, FDP, CDU und CSU liegen alle ungefähr auf einer Linie. Auch diese Linie könnte als konventionelles Parteienspektrum interpretiert werden, von dem sich die rechtsextremen Republikaner deutlich absetzen.

Betrachten wir nun die Idealpunkte der Ausprägungen der Links-Rechts-Selbsteinstufung:

Die Linie durch die westdeutschen Idealpunkte liegt oberhalb der ostdeutschen Linie und ist wesentlich weniger gekrümmt. Folgt man der oben zuerst genannten Interpretation der Lage der Parteien, so kann festgehalten werden, dass 1994 in Westdeutschland die Protestparteien links wie rechts als wenig sympathisch empfunden wurden, während in Ostdeutschland in den linken und rechten Randgruppen eine wesentlich stärkere Sympathie für diese Parteien empfunden wurde. Folgte man der zweiten Interpretation und zählte die PDS zu den etablierten Parteien des Parteienspektrums, so verliefen die Linien der West- und Ostdeutschen relativ ähnlich, mit dem Unterschied, dass die Ostdeutschen stärker mit dem "linken" Parteienspektrum sympathisierten.

6. Property Fitting - das Vektormodell zur Interpretation der euklidischen Konfiguration

Die Interpretation der Konfiguration einer MDS erfolgt in der Regel auf Basis von Hypothesen über die Struktur des Wahrnehmungsraumes der Befragten. Sofern keine zusätzlichen Informationen erhoben wurden, müssen diese Hypothesen aus der Theorie abgeleitet werden. In der Praxis des Marketing wird eine solche Theorie zumeist weniger fundiert sein, sondern eher ad hoc auf der Basis von Intuition, eigenen Erfahrungen und Markenkenntnissen gebildet werden. Ein Beispiel für eine solche ad-hoc-Interpretation einer Konfiguration stellen die Ausführungen zur MDS von Automobilmarken unter 1.1.2 dar.

Sofern die Möglichkeit besteht, zusätzliche Informationen bezüglich der Beurteilung der Objekte durch die Befragten zu erheben, bietet es sich an, diese Zusatzinformationen bei der Interpretation der Konfiguration zu nutzen. Die in diesem Zusammenhang angewendete Technik wird als *Property Fitting* oder *Vektormodell* bezeichnet.

Das Property Fitting vollzieht sich in folgenden Schritten:

1. Erhebung von Ähnlichkeitsdaten für Paare von Objekten (wie bei jeder MDS!)
2. Erhebung von zusätzlicher Eigenschaftsbeurteilungen der Objekte durch die Befragten (z.B. Qualität, Präferenz, Preiswürdigkeit etc.)
3. Durchführung der MDS (in der Regel einer einfachen MDS auf Basis der aggregierten Ähnlichkeitsmatrix). In der Konfiguration muss die euklidische Metrik gelten!
4. Regression der Eigenschaftsvektoren auf die Dimensionen der MDS
5. Einzeichnen der Eigenschaftsvektoren in die Konfiguration auf Basis der Regressionskoeffizienten
6. Interpretation der Lage der Objekte in der Konfiguration durch das Fällen des Lots auf die Eigenschaftsvektoren.

Die Schritte werden nun im einzelnen anhand eines Fallbeispiels erläutert:

In einer empirischen Untersuchung zur Wahrnehmung und Beurteilung von Jeansmarken wurden 8 Befragte aufgefordert, die Ähnlichkeit von je zwei Jeansmarken auf einer Skala von 1 (sehr ähnlich) bis 10 (sehr unähnlich) zu beurteilen. Dazu wurden die folgenden Jeansmarken jeweils paarweise abgefragt: Levi's, Diesel, Rocky (H&M), Boss, Replay, Versace, Jinglers (C&A), Joop und Calvin Klein.

Zusätzlich wurden die Befragten gebeten zu beurteilen, wie stark die folgenden Eigenschaften auf die Jeansmarken zutreffen (verwendet wurde eine Skala von 1 (trifft überhaupt nicht zu) bis 10 (trifft voll und ganz zu)): „hohe Qualität“, „preisgünstig“, „klassisch“ und „modern“.

Tabelle 6.1 enthält die Mittelwerte der Eigenschaftsbeurteilungen der Jeansmarken durch die 8 Befragten:

Tab. 6.1: Mittelwerte der Eigenschaftsbeurteilungen der Jeansmarken durch 8 Befragte

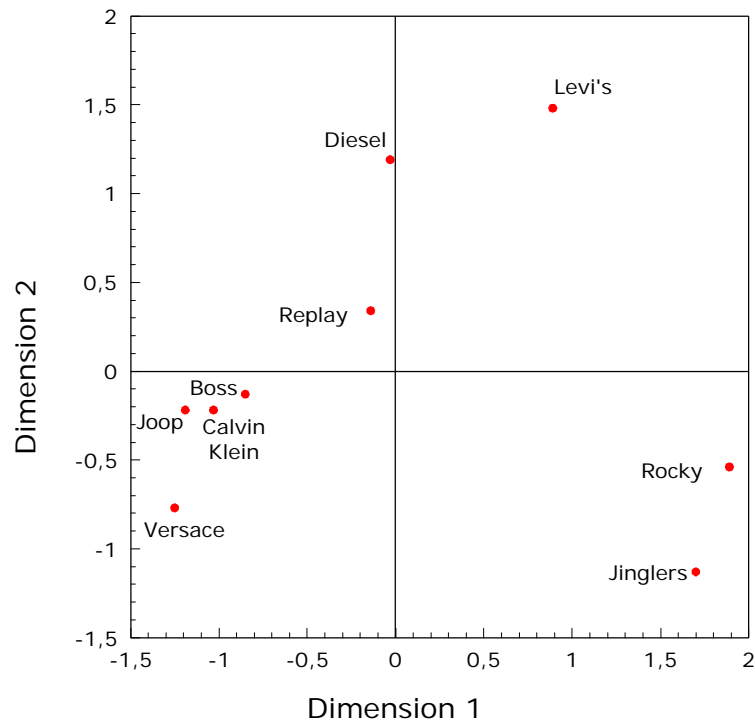
	Hohe Qualität	preisgünstig	klassisch	Modern
Levi's	9,00	3,50	9,13	8,25
Diesel	8,13	3,50	7,50	7,00
Rocky	4,13	8,75	5,00	6,63
Boss	8,25	2,00	6,13	4,38
Replay	7,50	3,00	5,38	6,50
Versace	7,63	1,13	5,13	5,75
Jinglers	4,00	8,38	5,13	3,50
Joop	7,75	4,63	5,63	5,63
Calv. Klein	7,88	3,13	5,75	7,13

Auf Basis der aggregierten Ähnlichkeitsmatrix wurde eine ordinale MDS durchgeführt (vgl. zum Ablauf analog Abschnitt 4.1). Der STRESS der zweidimensionalen Lösung war mit einem Wert von 0,0417 sehr gut. Die Konfiguration der MDS wird durch die folgenden Koordinaten aufgespannt (Tabelle 6.2):

Tab. 6.2: Koordinaten der Objekte in der zweidimensionalen Lösung

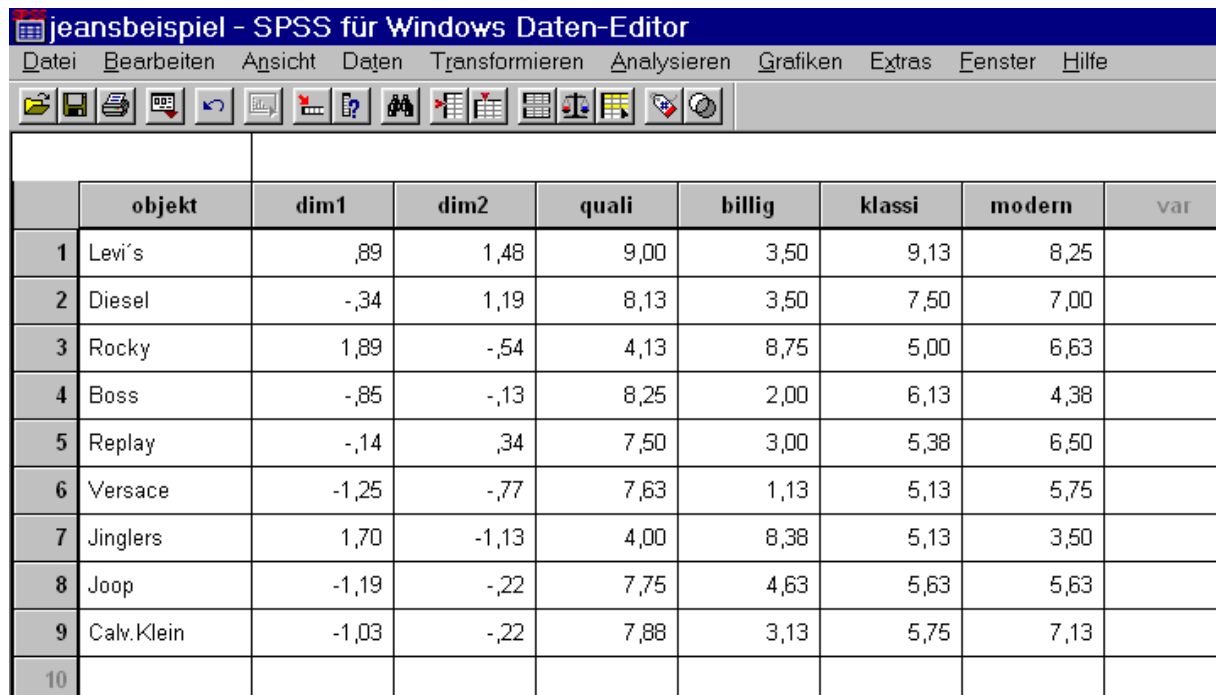
	Dimension 1	Dimension 2
Levi's	0,89	1,48
Diesel	-0,03	1,19
Rocky (H&M)	1,89	-0,54
Boss	-0,85	-0,13
Replay	-0,14	0,34
Versace	-1,25	-0,77
Jinglers (C&A)	1,70	-1,13
Joop	-1,19	-0,22
Calvin Klein	-1,03	-0,22

Die Konfiguration der Jeansmarken ist in Abbildung 6.1 dargestellt.



6.1: Konfiguration der Jeansmarken (zweidimensionale ordinale MDS)

Auf Basis der Konfiguration lassen sich die Jeansmarken grob in vier Kategorien einteilen: Klassische Marken (Diesel, Levi's), Handelsmarken (Rocky H&M, Jinglers C&A), Designerjeans (Boss, Joop, Calvin Klein, Versace) sowie die Marke Replay, die sich zwischen den Klassikern und den Designerjeans befindet. In diese Konfiguration sollen nun Vektoren der zusätzlich erhobenen Eigenschaftsbeurteilungen eingezeichnet werden. Dazu wird zunächst ein neues SPSS-Datenfenster erzeugt. In dieses Datenfenster (siehe Abbildung 6.2) werden als Variablen die Koordinaten der Objekte (dim1 und dim2) sowie die Mittelwerte der Eigenschaftsbeurteilungen (quali, billig, klassi, modern) eingetragen.



	objekt	dim1	dim2	quali	billig	klassi	modern	var
1	Levi's	,89	1,48	9,00	3,50	9,13	8,25	
2	Diesel	-,34	1,19	8,13	3,50	7,50	7,00	
3	Rocky	1,89	-,54	4,13	8,75	5,00	6,63	
4	Boss	-,85	-,13	8,25	2,00	6,13	4,38	
5	Replay	-,14	,34	7,50	3,00	5,38	6,50	
6	Versace	-1,25	-,77	7,63	1,13	5,13	5,75	
7	Jinglers	1,70	-1,13	4,00	8,38	5,13	3,50	
8	Joop	-1,19	-,22	7,75	4,63	5,63	5,63	
9	Calv.Klein	-1,03	-,22	7,88	3,13	5,75	7,13	
10								

Abb. 6.2: SPSS-Datenfenster zur Berechnung des Vektor-Modells

Nun wird für jede der Eigenschaften eine lineare Regressionsanalyse durchgeführt.²⁹ Die beiden Dimensionen sind dabei die unabhängigen Variablen, anhand derer die Varianz der Eigenschaftsbeurteilungen erklärt werden soll.

Die folgenden Tabellen enthalten exemplarisch den Output von SPSS für die lineare Regression der Eigenschaft „hohe Qualität“ auf die beiden Dimensionen.

Tab. 6.3: Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,960	,921	,895	,5837

Die Modellzusammenfassung in Tabelle 6.3 gibt Aufschluß darüber, wie gut die Konfiguration der Objekte (repräsentiert durch die beiden Dimensionen) geeignet ist, die Varianz der entsprechenden Eigenschaftsbeurteilung zu erklären. Als Maß für die erklärte Varianz wird R^2 berechnet und aufgrund der geringen Anzahl an Freiheitsgraden korrigiert. Das korrigierte R^2 stellt die Grundlage für die Entscheidung dar, ob die Eigenschaft als Vektor in die Konfiguration eingezeichnet werden kann oder nicht. In diesem Fall erklären die beiden Dimensionen 89,5% der Eigenschaft „hohe Qualität“. Dieser Wert ist zufriedenstellend. Daher wird diese Eigenschaftsbeurteilung als Vektor in die Konfiguration gelegt (siehe Abbildung 6.3).

Tabelle 6.4 enthält die errechneten Regressionskoeffizienten.

²⁹ Zur linearen Regressionsanalyse siehe ausführlich unsere entsprechende Beispieldatei.

Tab. 6.4 Koeffizienten

		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
Modell		B	Standardfehler	Beta		
1	(Konstante)	7,105	,195		36,504	,000
	DIM1	-,999	,168	-,683	-5,947	,001
	DIM2	1,318	,239	,634	5,514	,001

Für das Vektormodell sind in erster Linie die nicht standardisierten B-Koeffizienten der beiden Dimensionen von Interesse. Sie werden später als Koordinaten für den zu zeichnenden Vektor verwendet. Der Vektor verläuft im Diagramm als Gerade durch den Ursprung und den Punkt, der durch diese beiden Koordinaten festgelegt ist - und zwar als Pfeil in Richtung des Punktes.

Die standardisierten Beta-Koeffizienten geben in der Höhe ihres Betrages Aufschluss über die Erklärungskraft der einzelnen Dimensionen. Im dargestellten Fall tragen beide Dimensionen in etwa zu gleichen Teilen zur Erklärung der Varianz der abhängigen Variablen bei.

Die Konstante der Regressionsgleichung kann als Betrag interpretiert werden, den der Eigenschaftsvektor im Ursprung des Koordinatensystems aufweist.

Die folgende Tabelle 6.5 zeigt die wichtigsten Ergebnisse der Regressionsanalysen für alle vier Variablen im Überblick.

Tab. 6.5: Ergebnisse der Regressionanalysen auf die vier Eigenschaftsbeurteilungen

Eigenschaft	Korrigiertes R-Quadrat	B (DIM1)	B (DIM2)
Hohe Qualität	0,895	-0,999	1,318
Preisgünstig	0,727	1,757	-0,961
Klassisch	0,750	0,119	1,431
modern	0,387	0,013	1,240

Tabelle 6.5 ist zu entnehmen, dass die Varianz der Eigenschaften „hohe Qualität“ und „klassisch“ gut (> 75%) durch die Konfiguration erklärt wird. Die Varianz der Eigenschaften „preisgünstig“ und „modern“ wird dagegen kaum durch die Konfiguration der Objekte erklärt. Daher wird auf die Darstellung der letztgenannten Eigenschaften als Vektoren in der Konfiguration verzichtet.

Abbildung 6.3 enthält die Konfiguration mit den beiden eingezeichneten Vektoren „hohe Qualität“ und „klassisch“.

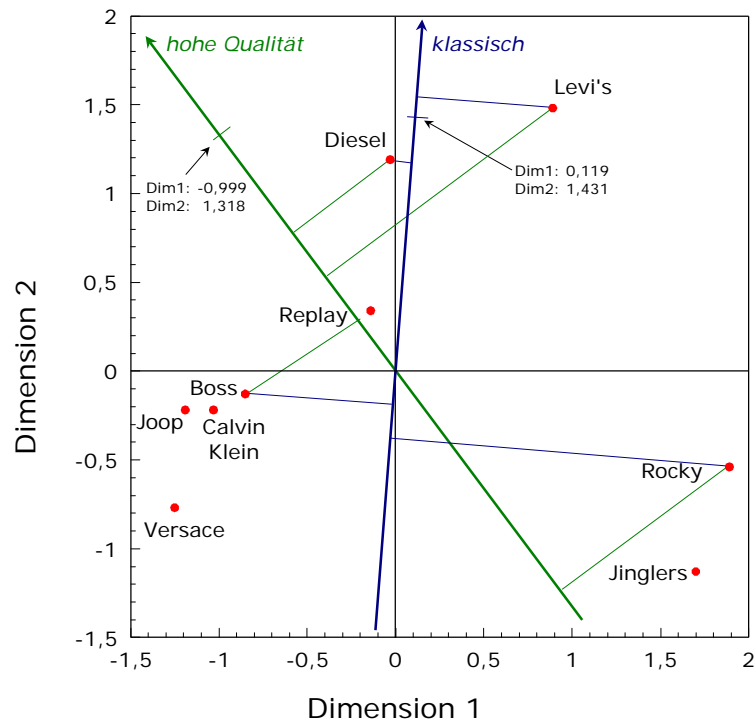


Abb. 6.3: Konfiguration der Objekte und Vektoren des Property Fitting.

Der Vektor *klassisch* verläuft in der Konfiguration relativ steil nach rechts oben.³⁰ Je weiter eine Jeansmarke in der Richtung dieses Vektors in der Konfiguration abgebildet wird, desto „klassischer“ wurde sie eingestuft. Dabei kommt es nicht auf die Nähe der Objekte zu den Vektoren an. Entscheidend ist die Lage des Lots von den Objekten auf den Vektor. So liegt das Lot von Levi's auf dem Vektor „klassisch“ weiter aussen als das Lot von Diesel auf diesem Vektor. Levi's wird also klassischer eingestuft als Diesel (obwohl Diesel „näher“ am Vektor *klassisch* liegt). Die Qualität von Diesel wird dagegen höher als die Qualität von Levi's eingeschätzt. Die Qualität von Boss wird ähnlich wie die Qualität von Replay eingeschätzt. Die beiden Handelsmarken - Rocky von H&M sowie Jingers von C&A - werden qualitativ am schlechtesten beurteilt.

³⁰ Zur Orientierung sind die beiden Punkte, die gemeinsam mit dem Ursprung die Richtung der Vektoren vorgeben, gekennzeichnet.

7. Ausblick

Wie gezeigt, eignet sich die MDS hervorragend zur explorativen Analyse subjektiver Wahrnehmungsräume. Dabei leisten das Vektor- und das Idealpunktmodell wertvolle Hilfen bei der Interpretation der gefundenen Lösungsräume. Weiterhin bietet das INDSCAL-Verfahren die Möglichkeit der Skalierung individueller Unterschiede in der Wahrnehmung. Gerade im Hinblick auf die Marktforschung ist die MDS daher vielfach einsetzbar zur

- Untersuchung eines Markenraumes
- Analyse der Positionierung einer bestimmten Marke
- Entwicklung einer Zielvorstellung für die Neupositionierung der Marke im Kontext konkurrierender Marken
- Überprüfung des Neupositionierungsziels nach Einsatz geeigneter Marketing-Instrumente

Darüber hinaus kann die MDS aber auch als Skalierungsverfahren im eigentlichen Sinne verstanden werden. Die Stärke der MDS in diesem Zusammenhang liegt darin, aus der ordinalen Ausgangsinformation über die Ähnlichkeit von Paaren von Objekten einen möglichst niedrig dimensionierten Wahrnehmungsraum zu konstruieren (vgl. die Diskussion der Datenverdichtung in Abschnitt 1.2). Die den Wahrnehmungsraum aufspannenden Dimensionen können als Intervallskalen im Sinne der *Repräsentationsmessung* verstanden werden, da über das STRESS-Maß ein Konsistenztest für das Meßmodell vorliegt (vgl. Skript zur Vorlesung, Kap. 5.1). Allerdings sollte dazu eine theoretisch fundierte Interpretation der erhobenen Skalen vorliegen. Zusätzlich ist genauer zu klären, welche Anforderungen im konkreten Fall an den Wert des STRESS-Maßes gestellt werden sollen. Die üblichen Kriterien für einen befriedigenden oder ausreichenden Fit (üblicherweise bei 2 Dimensionen) reichen in diesem Fall sicher nicht aus.