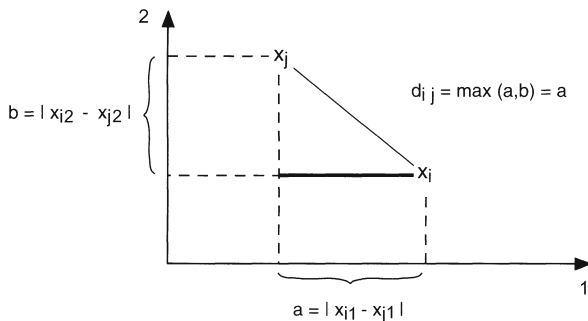


Distanz« geht auf die Situation eines Autofahrers zurück, der in einer Stadt (mit rechtwinklig verlaufenden Straßen) die Distanz zwischen Start und Ziel kalkuliert. Da die »Luftliniendistanz« nicht befahrbar ist (dies wäre die euklidische Distanz), setzt sich die Fahrstrecke aus zwei rechtwinkligen Straßenabschnitten zusammen – der City-Block-Distanz.

Für die Ermittlung einer Distanz nach der Supremummetrik setzen wir  $r \rightarrow \infty$ . Die allgemeine Distanzgleichung vereinfacht sich dann zu

$$d_{ij} = \max(|x_{ik} - x_{jk}|).$$

Diese Distanz entspricht – wie die folgende Abbildung für  $n=2$  verdeutlicht – der maximalen Koordinatendifferenz:



Da  $a > b$  ist, resultiert für die Supremumdistanz  $d_{ij} = a$ . Die Distanz entspricht der »dominierenden« Koordinatendifferenz (Dominanzmetrik).

**Bedeutung verschiedener Metriken.** Das NMDS-Verfahren bestimmt nicht nur die optimale Dimensionszahl, sondern auch diejenige Metrik, die den Ähnlichkeitsurteilen der Untersuchungsteilnehmer vermutlich zugrunde lag. Diese Metriken werden gelegentlich zur Beschreibung psychologisch unterscheidbarer Urteilsprozesse herangezogen.

Ähnlichkeitspaarvergleiche eignen sich vorzugsweise für die Skalierung komplexer, durch viele Merkmale charakterisierbarer Objekte. Die Instruktion, nach der die Untersuchungsteilnehmer die Paarvergleiche durchführen, sagt nichts darüber aus, nach welchen Kriterien die Ähnlichkeiten einzustufen sind. Dies bleibt den Untersuchungsteilnehmern selbst überlassen. Sie können beispielsweise die zu vergleichenden Objekte sorgfältig hinsichtlich einzelner Merkmale analysieren, um dann

Merkmal für Merkmal die Gesamtähnlichkeit aufzubauen. Dieses Vorgehen käme einer durch die City-Block-Metrik charakterisierten Urteilsweise sehr nahe.

Es sind auch Ähnlichkeitsurteile denkbar, die nur ein – gewissermaßen ins Auge springendes – Merkmal beachten, das die zu vergleichenden Objekte am stärksten differenziert. Diese Urteilsweise ließe sich durch die Supremummetrik beschreiben. In entsprechender Weise sind Zwischenwerte zu interpretieren: Im Bereich  $r > 2$  überwiegen »spezifisch-akzentuierende« und im Bereich  $r < 2$  »analytisch-kumulierende« Urteilsweisen (vgl. Bortz, 1975b).

Wie Wender (1969) zeigte, hängt die Art, wie Ähnlichkeitsurteile zustande kommen, auch von der Schwierigkeit der Paarvergleichsaufgabe ab: Je schwerer die Paarvergleichsurteile zu erstellen sind, desto höher ist der für das Urteilsverhalten typische Metrikkoeffizient. Bei schweren Paarvergleichen werden die deutlich differenzierenden Merkmale stärker gewichtet als die weniger differenzierenden Merkmale, und bei leichten Paarvergleichsurteilen erhalten alle relevanten Merkmale ein ähnliches Gewicht. Weitere Hinweise zur psychologischen Interpretation des Metrikparameters geben Cross (1965); Micko und Fischer (1970) sowie Shepard (1964). Methodenkritische Überlegungen zur Interpretation verschiedener Metriken liegen von Beals et al. (1968), Bortz (1974, 1975a), Wender (1969) und Wolfrum (1976a,b) vor.

### Die Analyse individueller Differenzen (INDSCAL)

Die Charakterisierung des Urteilsverhaltens durch einen Metrikparameter ist hilfreich für die Fragestellung, ob die beachteten Urteilsdimensionen gleich oder verschieden stark gewichtet wurden. Die Frage, wie stark ein Urteiler eine bestimmte Urteilsdimension gewichtet, wird damit jedoch nicht befriedigend beantwortet. Hierfür ist ein Verfahren einschlägig, das unter der Bezeichnung INDSCAL (Individual Scaling von Carroll & Chang, 1970) bekannt wurde.

Ausgangsmaterial sind die durch eine Urteilergruppe im Paarvergleich (oder in einem vergleichbaren Verfahren) bestimmten Ähnlichkeiten zwischen den zu skalierenden Objekten. Das Verfahren ermittelt neben der für alle Urteiler gültigen Reizkonfiguration (»Group Stimulus Space«) für jeden Urteiler einen individuellen Satz

von Gewichten, der angibt, wie stark die einzelnen Urteilsdimensionen gewichtet wurden.

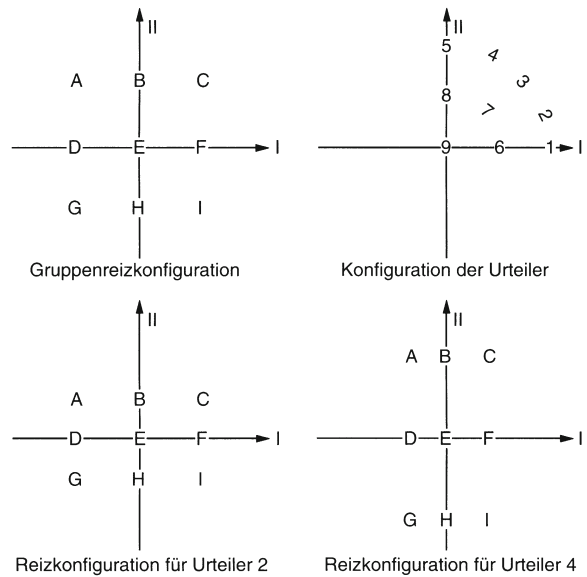
Die Besonderheit liegt also darin, dass das Verfahren es gestattet, für jeden Urteiler die relative Bedeutung oder Gewichtung der Dimensionen der Reizkonfiguration zu ermitteln. Diese individuellen Dimensionsgewichte geben an, wie stark ein Urteiler die einzelnen Dimensionen »streckt« oder »staucht«. Urteilsdimensionen, die ein Urteiler nicht beachtet, erhalten ein Gewicht von Null.

■ Abb. 4.5 veranschaulicht den INDSCAL-Ansatz an einem hypothetischen Beispiel. Die Gruppenreizkonfiguration zeigt die Position von neun Reizen auf zwei Dimensionen. Der Konfiguration der Urteiler ist zu entnehmen, wie jeder Urteiler die beiden Urteilsdimensionen gewichtet hat.

Während für den ersten und den sechsten Urteiler nur die erste und für den fünften und den achten Urteiler nur die zweite Dimension relevant ist, haben der zweite, dritte, vierte und siebente Urteiler beide Dimensionen – allerdings in unterschiedlichem Ausmaß – berücksichtigt. Der neunte Urteiler hat die beiden Urteilsdimensionen der Gruppenreizkonfiguration überhaupt nicht berücksichtigt. Für ihn waren anscheinend bei den Ähnlichkeitsschätzungen Merkmale ausschlaggebend, die für die anderen Urteiler keine Rolle spielten (**urteilerspezifische Merkmalsdimensionen**). Möglich ist allerdings auch, dass dieser Urteiler fehlerhafte bzw. zufällige Urteile abgab.

Die beiden unteren Grafiken in ■ Abb. 4.5 zeigen die Reizkonfiguration aus der Sicht des zweiten und des vierten Urteilers. Der zweite Urteiler streckt (bzw. gewichtet) die erste Dimension und der vierte Urteiler die zweite Dimension stärker als der Durchschnitt aller Urteiler.

Ausführlichere Hinweise zum mathematischen Aufbau dieses Verfahrens sind in der Originalarbeit von Carroll und Chang (1970), bei Carroll (1972), Borg (1981) oder verkürzt bei Ahrens (1974, S. 148 ff.) und bei Kühn (1976, S. 105 ff.) zu finden. Über weiterführende, an das INDSCAL-Modell angelehnte Verfahren informieren Carroll und Wish (1974). EDV-Hinweise enthält ► Anhang D. Als Anwendungsbeispiele für das INDSCAL-Verfahren seien die Arbeiten von Bortz (1975b) über Differenzierungsmöglichkeiten emotionaler und rationaler Urteile, von Wish und Carroll (1974) über individuelle Differenzen in der Wahrnehmung



■ **Abb. 4.5.** Hypothetisches Beispiel einer INDSCAL-Analyse. (Nach Carroll, 1972)

und im Urteilsverhalten sowie von Wish et al. (1972) über unterschiedliche Wahrnehmungen der Ähnlichkeit von Nationen erwähnt.

#### 4.2.4 Ratingskalen

Während bei Rangordnungen und Paarvergleichen von den Untersuchungsteilnehmern ordinale Urteile abzugeben sind, können mittels Ratingskalen (engl. rating = Einschätzung) auf unkomplizierte Weise Urteile erzeugt werden, die als intervallskaliert interpretiert werden können (► S. 181 f.). Ratingskalen zählen zu den in den Sozialwissenschaften am häufigsten verwendeten, aber auch umstrittensten Erhebungsinstrumenten (zur Geschichte der Ratingskala vgl. McReynolds & Ludwig, 1987). Die Industrie verwendet Ratings zur Bewertung von Arbeitsplätzen oder zur Personalauslese (vgl. z. B. Landy & Farr, 1980; Marcus & Schuler, 2001), Lehrer bewerten und benoten die Leistungen ihrer Schüler, Ärzte und Psychologen stufen das Verhalten psychisch Erkrankter ein; die Liste der Beispiele ließe sich mühelos verlängern. Im Folgenden befassen wir uns mit

- verschiedenen Varianten von Ratingskalen,
- messtheoretischen Problemen bei Ratingskalen,