

Vektoren und Matrizen

Übungen 2

Allgemeine Betrachtungen zur Vektor- und Matrixrechnung: Es könnte hilfreich sein, sich den Sinn der Anwendung der Vektor- und Matrixrechnung in der multivariaten Analyse klar zu machen.

Gegeben sei eine $(m \times n)$ -Matrix $X = (x_{ij})$ von zentrierten Messungen von n Variablen an m Personen. Aus verschiedenen Gründen ist es von Interesse, die Kovarianzen oder Korrelationen zwischen den Variablen durch latente Variable zu "erklären", und der übliche Ansatz dazu ist

$$x_{ij} = a_{j1}L_{i1} + a_{j2}L_{i2} + \cdots + a_{jn}L_{in}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n \quad (1)$$

wobei die L_{i1}, \dots, L_{in} Werte von unbekanntem, "latentem" Variablen repräsentieren und die a_{jk} als Gewichte interpretiert werden können, mit denen diese Variablen in die Messungen x_{ij} eingehen. L_{ik} ist etwa ein Maß für die Ausprägung der k -ten latenten Variablen bei der i -ten Person in einer Stichprobe von Personen, und a_{jk} repräsentiert das Ausmaß, in dem der j -te Test die k -te latente Variable erfasst.

Es gibt $m \cdot n$ "Datenpunkte" x_{ij} , aber $m \cdot n$ unbekannte Parameter L_{ik} plus $n \cdot n$ unbekannte Parameter a_{jk} . Für $m = 100$ und $n = 20$ ergeben sich dann 2400 freie Parameter, die irgendwie aus den Daten bestimmt werden müssen. Die a_{jk} und L_{ik} sind Parameter, weil sie als Konstanten in einem Modell erscheinen, dass durch die rechte Seite von (1) spezifiziert wird: es wird angenommen, dass die latenten Variablen zwar durch die a_{jk} gewichtet, abgesehen davon aber additiv in die Bestimmung der X_{ij} eingehen. Man spricht von einem *linearen Modell*. Die Annahme eines solchen Modells ist nicht denkbare, aber es ist das einfachste Modell, dass man annehmen kann. Die Parameter heißen *frei*, weil ihre Werte nicht durch irgendwelche Nebenbedingungen festgelegt sind und aus den Daten geschätzt werden müssen.

Die Anzahl der zu schätzenden Parameter ist größer als die Anzahl der x_{ij} -Werte, also kann keine der bekannten Parameterschätzverfahren wie etwa die Methode der Kleinsten Quadrate (minimiere $Q = \sum_{i,j} (x_{ij} - \hat{x}_{ij})^2$, \hat{x}_{ij} durch die rechte Seite von (1) gegeben) zum Erfolg führen. Es müssen also Randbedingungen für die Parameter eingeführt werden.

Ein erster Schritt ist der Übergang von der Betrachtung einzelner Messwerte zu Gruppen von Messwerten: die Messwerte einer Person werden zu einem n -dimensionalen Vektor $\tilde{\mathbf{x}}_i$ zusammengefasst, dessen Komponenten die zentrierten Messwerte der i -ten Person sind: $x_{i1}, x_{i2}, \dots, x_{in}$. Ebenso können die a_{jk} und L_{ik} zu Vektoren \mathbf{a}_j und $\tilde{\mathbf{L}}_i$ zusammengefasst werden, und die Variablen können durch die Vektoren $\mathbf{x}_j = (x_{1j}, \dots, x_{mj})'$ repräsentiert werden. Durch die Möglichkeit der Verknüpfung von Vektoren als Linearkombinationen und Skalarprodukte werden Beziehungen zwischen den Daten und den Parametern darstellbar, die auf der Ebene einzelner Datenpunkte bzw. einzelner Parameterwerte kaum oder nur sehr schwer umständlich deutlich werden. So kann man fordern, dass die Vektoren $\tilde{\mathbf{X}}_i$ und den $\tilde{\mathbf{L}}_i$ durch eine Rotation ineinander überführbar sein sollen, weil dann die Konfiguration der Punkte, die Personen abbilden, im Raum der latenten Variablen der Konfiguration im Datenraum entspricht, d.h. die Relationen zwischen den Personen können im Raum der latenten Variablen betrachtet werden. Diese Rotation kann auf der Datenebene nur sehr umständlich durch Koordinatentransformationen bewerkstelligt werden, in die der Winkel, um den rotiert werden soll, explizit eingeht. Im Rahmen der Vektor- und Matrixrechnung wird nur eine zunächst unbekannte Matrix S postuliert derart, dass $\tilde{\mathbf{x}}_i = S\tilde{\mathbf{L}}_i$ für alle i , und da S eine Rotation bewirken soll, folgt unmittelbar, dass S eine orthonormale Matrix ist, – die Vektoren $\tilde{\mathbf{x}}_i$ sind damit als Linearkombinationen von Vektoren gekennzeichnet, die orthogonal und normiert sind, und darüber hinaus folgt, dass $T\tilde{\mathbf{x}}_i = \tilde{\mathbf{L}}_i$ sein muß mit $T = S'$. Damit sind die Spaltenvektoren von S und T noch nicht explizit gegeben, aber einige wenige weitere Betrachtungen auf der Ebene von Vektoren und Matrizen zeigen, dass T die Matrix der Eigenvektoren der Kovarianz- oder Korrelationsmatrix sein muß, d.h. dass für die Spalten von T die Beziehung $R\mathbf{t}_k = \lambda_k \mathbf{t}_k$, $k = 1, \dots, n$ gelten muß. Diese Vektoren sind numerisch bestimmbar; man kann sagen, dass T mit den Daten mitgeliefert wird. All diese Schlußfolgerungen resultieren aus Betrachtungen über *Strukturen* von Daten und Parametermengen, die ohne den Vektor- und Matrixbegriff kaum zugänglich gewesen wären. Man hat also den Übergang

Daten \longrightarrow Vektoren \longrightarrow Matrizen

Dieser Übergang bedeutet den Übergang von einzelnen Zahlen zu allgemeinen Strukturen, die zwischen den Zahlen existieren. Die Regeln zur Verknüpfung von Vektoren und Matrizen implizieren wiederum syntaktischen Operationen (Linearkombinationen, Skalarprodukte, Matrixmultiplikationen), die die genannten Einsichten in die Strukturen dieser Merkmale liefert, zu denen man ohne den Vektor- und Matrizenkalkül nur schwer gelangen würde. Dieser Kalkül ist keine *l'art pour l'art*-Konstruktion abgehobener, von der Realität entfremdeter Mathematiker, sondern ist als eine Art Werkzeugkasten für das Lösen komplexer Probleme über Jahrzehnte hinweg anhand konkreter Probleme entwickelt worden.

Aufgaben:

1. **Aufgabe:** Es sei A eine $(m \cdot n)$ -Matrix und I_n sei die $(n \cdot n)$ -Einheitsmatrix, d.h. die Zeilen- und Spaltenvektoren bestehen aus den Einheitsvektoren \mathbf{e}_j , $j = 1, \dots, n$. Vergewissern Sie sich, dass $AI_n = A$, und $I_m A = A$, I_m die $(m \times m)$ -Einheitsmatrix.

2. **Aufgabe:** Es sei

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}, \quad Y = \frac{1}{x_{11}x_{22} - x_{12}x_{21}} \begin{pmatrix} x_{22} & -x_{12} \\ -x_{21} & x_{11} \end{pmatrix}.$$

Zeigen Sie, dass das Produkt XY gleich der (2×2) -Einheitsmatrix I_2 ist; für die Matrix Y gilt $Y = X^{-1}$, X^{-1} ist die zu X *inverse Matrix* (oder einfach *die Inverse*).

Welche Voraussetzung muß erfüllt sein, damit die Inverse existiert?

3. **Aufgabe:** Wie zeigt man, dass *alle* 2-dimensionalen Vektoren als Linearkombinationen von zwei beliebigen, aber linear unabhängigen Vektoren \mathbf{x}_1 und \mathbf{x}_2 dargestellt werden können?
4. **Aufgabe:** X sei wie in der vorangegangenen Aufgabe definiert. Gilt insbesondere einerseits

$$Y = XA, \quad Y = [\mathbf{y}_1, \mathbf{y}_2], \quad X = [\mathbf{x}_1, \mathbf{x}_2], \quad A = [\mathbf{a}_1, \mathbf{a}_2]$$

und andererseits

$$X = YB, \quad B = [\mathbf{b}_1, \mathbf{b}_2],$$

und $\mathbf{y}_1, \mathbf{y}_2$ seien linear unabhängig. Welche Relation existiert zwischen den Matrizen A und B ?

Hinweis: Den Begriff der inversen Matrix nicht vergessen!

5. **Aufgabe:** Gegeben seien die 5-dimensionalen Einheitsvektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (\mathbf{e}_1 hat eine 1 als erste Komponente, \mathbf{e}_2 hat eine 1 als zweite Komponente, \mathbf{e}_3 hat die 1 als dritte Komponente); \mathbf{x} sei eine Linearkombination dieser Vektoren. Hat \mathbf{x} Komponenten, die gleich Null sind, und falls ja, welche sind es? (Vergl Satz 1.4, p. 17 im Skriptum). Es sei $T = (t_{ij})$ eine (5×5) -Matrix, $t_{ij} \neq 0$ für alle i, j . Es sei $\mathbf{y} = T\mathbf{x}$. Wieviele Komponenten hat \mathbf{y} , die gleich Null sind?
6. **Aufgabe:** Es seien zehn 100-dimensionale Vektoren \mathbf{x}_j , $j = 1, \dots, 10$ gegeben; jeder Vektor \mathbf{x}_j sei als Linearkombination von 3 orthonormalen 100-dimensionalen Vektoren $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3$ darstellbar:

$$\mathbf{x}_j = a_{1j}\mathbf{L}_1 + a_{2j}\mathbf{L}_2 + a_{3j}\mathbf{L}_3.$$

Zeigen Sie, dass die Koeffizienten a_{jk} einer Korrelation zwischen \mathbf{x}_j und \mathbf{L}_k entsprechen (vergl. Seite 22, Skriptum \rightarrow orthonormale Basisentwicklung). Wieviele Dimensionen hat der Teilraum, in dem die \mathbf{x}_j liegen?

7. **Aufgabe:** Gegeben sei die Matrix $L = [\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3]$, wobei die \mathbf{L}_k paarweise orthogonal seien. Welche Form hat die Matrix $L'L$, und wie lautet die Matrixgleichung, mit der Sie eine Normalisierung der \mathbf{L}_k erreichen?

Hinweis: Wiederholen Sie, wie eine Normalisierung von Vektoren durchgeführt wird (s. Seite 11, Skriptum), und das Resultat der Multiplikation einer Matrix mit einer Diagonalmatrix (s. Seite 26, Skriptum).

8. **Aufgabe:** In Gleichung (1) wurden die zentrierten Werte x_{ij} durch das Skalarprodukt $\mathbf{a}'_j \mathbf{L}_i$ "erklärt". Welche Bedeutung hat der Spezialfall, dass \mathbf{a}_j und \mathbf{L}_i orthogonal sind, und wann nimmt x_{ij} einen maximalen Wert an?

Hinweis: $\mathbf{x}'\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos \theta$ und $|\cos \theta| \leq 1$.