

Hauptkomponentenanalyse

Jürgen Grieser

09.10.1997

1 Einleitung

Der wesentliche Bestandteil der Hauptkomponentenanalyse ist die Hauptachsentransformation. Geht man davon aus, daß man n Zeitreihen der Länge p vorgegeben hat, so stellen diese eine Punktwolke von n Punkten im p -dimensionalen Raum dar. Man kann die p skalaren Zeitreihen somit auch als eine p -dimensionale Vektorzeitreihe auffassen. Ein Vektor (und auch eine Reihe von Vektoren) kann bezüglich einer beliebigen Basis im Raum formuliert werden. Ziel der Hauptachsentransformation ist es nun, eine andere sinnvollere Basis zu verwenden, als die durch die Zeitreihen vorgegebene. Dies geschieht in zwei Schritten. Zunächst wird der Ursprung des Koordinatensystems in den Schwerpunkt der Punktwolke gesetzt. Im zweiten Schritt wird das Koordinatensystem dann so gedreht, da die erste Koordinate in Richtung der größten Varianz der Punktwolke zeigt. Damit ist die erste Hauptachse festgelegt und die Varianz in dieser Richtung ist die erste Hauptkomponente. Die nächste Drehung wird dann um diese Koordinatenachse durchgeführt, und zwar so, daß die zweite Hauptachse (die orthogonal zur ersten stehen muß) in Richtung der größten verbleibenden Varianz zeigt. Dieser Vorgang wird so oft wiederholt, bis eine neue p -dimensionale Basis geschaffen ist. Die neuen Basisvektoren werden oft empirische Orthogonalfunktionen (EOF) genannt. Nach dieser Transformation ist die Varianz der Punktwolke (und damit auch der Originalreihen) so auf neue Koordinaten (und damit auch neue Zeitreihen) verteilt, daß die Varianz dieser Reihen mit zunehmender Reihennummer abnimmt. Die neuen Zeitreihen heißen Hauptkomponenten-Zeitserien oder PC-Zeitserien. Die erste Hauptachse beschreibt die Hauptvarianz, die p -te Hauptachse die wenigste Varianz. Eine Anwendung der Hauptachsentransformation ist die Vernachlässigung von Anteilen, die wenig zur Gesamtvarianz beitragen. Dazu betrachtet man einfach nur einen Anteil der Hauptachsen. Man erhält somit die (linear) effektivste Darstellung, mit den wenigsten Datenreihen. Hat man z.B. $p = 5$ Datenreihen, und drücken die ersten zwei Hauptachsen schon 99% der Varianz aus, so reicht es, wenn man sich nicht für das verbleibende 1% Restvarianz interessiert, nur die beiden zu diesen Hauptachsen gehörenden Zeitserien zu untersuchen. Es gibt dann keine effektivere Darstellung der fünf Ausgangsreihen durch nur zwei Zeitserien.

Für Analysen ist es von besonderem Interesse zu sehen, wie schnell die Varianz in den Hauptachsen abnimmt. Konkret stellt sich dabei die Frage, wieviel unabhängige Information ist in den p Datenreihen wirklich enthalten. Eine weitere wichtige Frage ist, wie ist die Information auf die verschiedenen Zeitreihen verteilt. Diese Information steckt in der Drehmatrix, mit der vom alten System ins neue gedreht wird.

1.1 Transformationsstrategie

Gegeben sei eine Matrix $Z(t, x)$ mit $t = 1, \dots, n$ und $x = 1, \dots, p$, die aus den p Zeitreihen der Länge n besteht:

$$Z = \begin{pmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,p} \\ z_{2,1} & z_{2,2} & \cdots & z_{2,p} \\ \dots & \dots & \dots & \dots \\ z_{n,1} & z_{n,2} & \cdots & z_{n,p} \end{pmatrix} \quad (1)$$

Zunächst wird von jeder Zeitreihe der Mittelwert dieser Reihe abgezogen. Damit sind die Spaltensummen und v.a. die Gesamtsumme gleich null und man hat den Koordinatenursprung in den Schwerpunkt der Punktwolke gesetzt. Im folgenden werden also nur noch die Abweichungen vom Mittel untersucht. Die wesentliche Information darüber, wie das System zu drehen ist, steckt in der Kovarianzmatrix. Die Drehung soll so stattfinden, daß zwischen den gedrehten Koordinaten keine Kovarianz mehr besteht. Das bedeutet, daß die entstehenden Reihen linear unabhängig werden. Dies ist gleichbedeutend mit der Forderung der effektivsten Darstellung. Jede neue Koordinate enthält nur linear unabhängige neue Information. Die Forderung, daß zwischen den neuen Koordinaten keine Korrelation mehr bestehen soll, läßt sich dadurch erreichen, daß die Kovarianzmatrix diagonalisiert wird. Die zur diagonalisierten Kovarianzmatrix gehörenden Eigenvektoren sind eine Linearkombination der alten Basisvektoren, und drücken damit den Zusammenhang zwischen den alten und den neuen Zeitreihen aus. Die Eigenwerte der Kovarianzmatrix (die Diagonalen der diagonalisierten Kovarianzmatrix) geben die Varianz in dieser (neuen) Koordinatenrichtung an.

Konkret müssen zunächst also die Eigenwerte λ der Kovarianzmatrix Σ gefunden werden. Dies geschieht durch Lösen der folgenden Gleichung:

$$\det(\Sigma - \lambda I) = 0. \quad (2)$$

Dabei stellt I die Einheitsmatrix dar. Da Σ eine $p \times p$ Matrix ist, existieren p Eigenwerte. Da Σ außerdem noch reell und symmetrisch ist, sind die Eigenwerte alle reell und positiv semidefinit. Zusätzlich folgt daraus schon, daß die Matrix der Eigenvektoren orthogonal ist. Hat man die Eigenwerte nach Glg. (2) bestimmt, so kann man mit

$$(\Sigma - \lambda_i I) \cdot \vec{e}_i = 0 \quad (3)$$

den i -ten Eigenvektor \vec{e}_i bestimmen. Die x -te Komponente des i -ten Eigenvektors ist dann das Gewicht, mit dem die x -te Originalzeitreihe in die i -te Zeitreihe des gedrehten Systems eingeht.

1.2 Ein Beispiel

Als Beispiel wird hier eine Hauptkomponententransformation von vier Temperaturreihen aus dem Mittelmeergebiet durchgeführt. Um Rechenaufwand zu sparen, werden nur die Monatswerte des ersten Halbjahres 1990 verwendet. Die Werte sind in Tabelle 1 eingetragen:

Tabelle 1: Verwendete Datenreihen (Monatsmitteltemperaturen in $.1^{\circ}C$).

Station	Januar	Februar	März	April	Mai	Juni	Mittel
Athen	84	109	140	166	205	247	158.5
Gibraltar	138	157	161	165	197	223	173.5
Luqa	132	141	147	167	197	239	170.5
Palma	101	135	128	128	181	219	148.66

Der Mittelwertvektor der Punktwolke (die aus sechs Punkten im vierdimensionalen Raum besteht) ist gegeben durch $\vec{m} = (158.5; 173.5; 170.5; 148.66)$. Dann folgt für die (schon auf den Mittelwert bezogene) Matrix Z :

$$Z(t, x) = \begin{pmatrix} -74.5 & -35.5 & -38.5 & -47.66 \\ -49.5 & -16.5 & -29.5 & -13.66 \\ -18.5 & -12.5 & -23.5 & -20.66 \\ 7.5 & -8.5 & -3.5 & -20.66 \\ 46.5 & 23.5 & 26.5 & 32.33 \\ 88.5 & 49.5 & 68.5 & 70.33 \end{pmatrix} \quad (4)$$

Als nächstes ist die Kovarianzmatrix zu bestimmen. Sie ergibt sich zu:

$$\Sigma = \begin{pmatrix} 3678.701 & 1820.498 & 2406.300 & 2436.602 \\ 1820.498 & 952.698 & 1238.098 & 1318.599 \\ 2406.300 & 1238.098 & 1662.299 & 1694.200 \\ 2436.602 & 1318.599 & 1694.200 & 1861.068 \end{pmatrix} \quad (5)$$

Für die Korrelationsmatrix \mathcal{C} , die die Korrelationskoeffizienten zwischen je zwei Zeitreihen als Elemente enthält, folgt damit

$$\mathcal{C} = \begin{pmatrix} 1 & .9724 & .9731 & .9311 \\ .9724 & 1 & .9838 & .9903 \\ .9731 & .9838 & 1 & .9632 \\ .9311 & .9903 & .9632 & 1 \end{pmatrix} \quad (6)$$

Die Korrelationen zwischen den Reihen sind sehr hoch. Dies liegt darin, daß es bei allen Stationen zwischen Januar und Juni wärmer wird. Zwischen der ersten und der vierten Reihe ist die Korrelation am geringsten, zwischen der zweiten und der vierten Reihe am

höchsten. Athen und Palma hatten demnach den unterschiedlichsten, Gibraltar und Palma den ähnlichsten Temperaturverlauf.

Aus Glg. (2) folgen die vier Eigenwerte in absteigender Reihenfolge:

7937.275
180.048
37.224
.219

Das sind die Varianzen der vier zugeordneten Zeitreihen im gedrehten Raum. Wie man sieht trägt die vierte Hauptachse kaum noch Varianz bei. Die mit Abstand meiste Varianz ist in der ersten Hauptachse. Da die neuen Koordinaten orthogonal sind, kann die Varianz einfach addiert werden. Damit kann man ausdrücken welcher Anteil der Gesamtvarianz der Punktwolke in welcher Koordinate liegt. Die Varianz teilt sich demnach auf die Koordinaten wie folgt auf:

97.3330 %
2.2079 %
.4565 %
.0027 %.

Die zugeordneten (hier schon normierten) Basisvektoren sind:

$$\begin{aligned} \vec{e}_1 &= (+.6733; +.3448; +.4535; +.4713) \\ \vec{e}_2 &= (-.6539; +.2059; +.0604; +.7255) \\ \vec{e}_3 &= (-.3041; -.1625; +.8888; -.3020) \\ \vec{e}_4 &= (-.1630; +.9013; -.0271; -.4005). \end{aligned} \tag{7}$$

Man kann sich leicht davon überzeugen, daß diese Vektoren senkrecht (im Rahmen der Rechengenauigkeit) aufeinanderstehen, indem man das Skalarprodukt zwischen ihnen bildet. Ebenso einfach sieht man, daß sie normiert sind, da ihre Länge eins ist. Es ist nun möglich die Zeitreihen im gedrehten System zu berechnen. Die Zeitreihe der ersten Hauptkomponente im gedrehten System besteht aus der Summe der mit den entsprechenden Komponenten des ersten Hauptvektors multiplizierten Zeitreihen, d.h.

$$a_1(t) = .6733 z_1(t) + .3448 z_2(t) + .4535 z_3(t) + .4713 z_4(t). \tag{8}$$

1.3 Nomenklatur

Gegeben sei die Matrix Z der (p) zeitlich zentrierten (t-centered) Zeitreihen der Länge n mit den Elementen:

$$z(t, x), \quad x = 1, \dots, p \text{ und } t = 1, \dots, n. \tag{9}$$

Die $z(t, x)$ sind skalare Variablen. Ein Schnappschuß der raumzeitlichen Entwicklung zur Zeit t ist dann gegeben durch den Vektor

$$\vec{z}_x(t) = [z(t, 1); z(t, 2); \dots; z(t, p)]. \tag{10}$$

Die Zeitreihe an einem bestimmten Raumpunkt x ist gegeben durch

$$\vec{z}_t(x) = [z(1, x); z(2, x); \dots; z(n, x)]. \quad (11)$$

Damit ist die Matrix Z auch darstellbar als Vektor von Vektoren

$$Z = [\vec{z}_x(t=1), \vec{z}_x(t=2); \dots; \vec{z}_x(t=n)]^T = [\vec{z}_t(x=1), \vec{z}_t(x=2); \dots; \vec{z}_t(x=p)]. \quad (12)$$

1.4 Streumatrix

Die Matrix Z stellt n Punkte im p -dimensionalen Raum dar. Jeder Punkt ist dabei ein Zeitpunkt (man könnte das durchaus auch umgekehrt definieren). In dem p -dimensionalen Raum kann durch die Basisvektoren $\vec{e}_1; \vec{e}_2; \dots; \vec{e}_p$ eine beliebige Basis definiert werden.

$$\vec{z}_x^T(t) \cdot \vec{e}_j = \vec{e}_j^T \cdot \vec{z}_x(t) = \sum_{x=1}^p z(t, x) e_j(x) \quad (13)$$

ist dann die Projektion der durch Z beschriebenen Punktwolke in die Richtung von \vec{e}_j . Als Streuung (Scatter) in Richtung von \vec{e}_j bezeichnet man dann

$$\Psi(\vec{e}_j) \equiv \sum_{t=1}^n [\vec{z}_x^T(t) \cdot \vec{e}_j]^2 = \sum_{t=1}^n [\vec{z}_x^T(t) \cdot \vec{e}_j] [\vec{e}_j^T \cdot \vec{z}_x(t)]. \quad (14)$$

Das ist $n-1$ mal die Varianz der Projektion der Punktwolke auf die Achse mit der Richtung \vec{e}_j . Gleichung (14) kann leicht umgeformt werden zu

$$\Psi(\vec{e}_j) = \vec{e}_j^T \left[\sum_{t=1}^n \vec{z}_x(t) \vec{z}_x^T(t) \right] \vec{e}_j. \quad (15)$$

Der Term in den eckigen Klammern in Glg. (15) ist die Streu- bzw. Scattermatrix. Sie ist $n-1$ mal die Kovarianzmatrix. Für die Elemente $s_{x,x'}$ der Scattermatrix gilt:

$$s(x, x') = \sum_{t=1}^n z(t, x) z(t, x'). \quad (16)$$

In Matrixschreibweise bedeutet dies:

$$S = Z^T Z. \quad (17)$$

Die Eigenwerte der Streumatrix sind dann gegeben durch

$$S \vec{e}_j = \lambda \vec{e}_j, \text{ für alle } j. \quad (18)$$

Das zu lösende Gleichungssystem ist demnach

$$\sum_{x'=1}^p s(x, x') \vec{e}_j(x') = \lambda_j \vec{e}_j(x), \text{ für alle } x. \quad (19)$$

Mit den Matrixdarstellungen $E \equiv [\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p]$ und $L = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_p]$ und $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ folgt die Matrixnotation:

$$S E = E L. \quad (20)$$

Nun kann man die Punktwolke Z im Eigenvektorsystem E ausdrücken. Dann ist

$$Z = \sum_{j=1}^p \vec{a}_j \vec{e}_j^T. \quad (21)$$

Dabei sind die Amplitudenvektoren \vec{a}_j die PC-Zeitreihen $\vec{a}_j = [a_j(1), a_j(2), \dots, a_j(n)]^T$. Damit folgt

$$z(t, x) = \sum_{j=1}^p a_j(t) e_j(x) \quad (22)$$

Die Einführung neuer normalisierter Amplituden $\alpha_j(t) \equiv a_j(t)/\lambda_j^{1/2}$ führt zu der Singular-Value-Decomposition:

$$z(t, x) = \sum_{j=1}^p \lambda_j^{1/2} \alpha_j(t) e_j(x) \quad (23)$$

mit der Matrixschreibweise

$$Z = A' L^{1/2} E^T. \quad (24)$$

1.5 Korrelationen

Man kann zwei Arten von Korrelationen unterscheiden. Einerseits kann man nach der Korrelation zwischen der Zeitreihe an der Position x und der PC-Zeitreihe j fragen. Diese Korrelation wird hier $r_{za}(j, x)$ genannt und gibt an, wieviel zeitliche Variation des Ortes x durch die Zeitreihe des j -ten Eigenvektors angegeben ist. Andererseits kann man nach der räumlichen Korrelation zwischen dem Ursprungsfeld zur Zeit t und dem Eigenvektor j zur Zeit t fragen. Diese Korrelation wird hier als $r_{ze}(j, t)$ bezeichnet und gibt an, wie stark der j -te Eigenvektor zur Zeit t in der beobachteten räumlichen Struktur vorhanden ist. Für beide Arten der Korrelation können demnach Karten gemalt werden, von denen eine die Korrelation in Abhängigkeit von Ort und Eigenvektor ausdrückt, während die andere die Korrelation in Abhängigkeit von der Zeit und den Eigenvektoren ausdrückt. Das Produktmoment der Zeitreihe an der Position x und der PC-Zeitreihe j ist gegeben durch

$$\sum_{t=1}^n z(t, x) a_j(t) = \sum_{t=1}^n \left[\sum_{k=1}^p a_k(t) e_k(t) \right] a_j(t) = \lambda_j e_j(x). \quad (25)$$

Damit folgt für den Korrelationskoeffizienten

$$r_{za}(j, x) \equiv \frac{\sum_{t=1}^n z(t, x) a_j(t)}{\sqrt{\left[\sum_{t=1}^n z^2(t, x) \right] \left[\sum_{t=1}^n a_j^2(t) \right]}} = \frac{\lambda_j^{1/2} e_j(x)}{\sqrt{\sum_{j=1}^p \lambda_j e_j^2(x)}}. \quad (26)$$

Analog folgt wegen

$$\sum_{x=1}^p z(t, x) e_j(x) = a_j(t) = \lambda_j^{1/2} \alpha_j \quad (27)$$

für

$$r_{ze}(j, t) \equiv \frac{\sum_{x=1}^p z(t, x) e_j(x)}{\sqrt{\left[\sum_{x=1}^p z^2(t, x) \right] \left[\sum_{x=1}^p e_j^2(x) \right]}} = \frac{\lambda_j^{1/2} \alpha_j(t)}{\sqrt{\sum_{j=1}^p \lambda_j \alpha_j^2(t)}}. \quad (28)$$