SVD und PCA

M. Gruber

KW 43, Rev.1

1 Singulärwertzerlegung (SVD)

Die Videolektion Singular Value Decomposition (the SVD) von Prof. Gilbert Strang führt in das Thema ein. Wir notieren hier die wichtigsten Punkte der 15-minütigen Einführung.

Jede $m \times n$-Matrix A hat eine Singulärwertzerlegung, d.h. man kann A in der Form

$$A = U\Sigma V^T \tag{1}$$

faktorisieren mit Matrizen U und V, deren Spalten paarweise orthogonal und auf die Länge 1 normiert sind (wir nennen solche Matrizen *orthonormal*). Der Faktor Σ ist eine Diagonalmatrix. Sie enthält die *Singulärwerte*, der Größe nach von links oben nach rechts unten absteigend angeordnet. Die Spalten von U sind die *linken Singulärvektoren*, die Spalten von V die *rechten Singulärvektoren*.

Wir werden Fälle betrachten, in denen A eine Datenmatrix ist.

Beispiel 1 (Iris) *Im Fall der Iris-Daten repräsentiert eine Zeile von A ein Iris-Exemplar. Die Spalten von A stehen für die vier Attribute sepal.length, sepal.width, petal.length, petal.width. Wenn die vier Spalten $A_{*1}, A_{*2}, A_{*3}, A_{*4}$ von A außerdem zentriert sind, d.h. wenn das arithmetische Mittel über jede Spalte null ist, ist $A^TA = n \cdot \mathbf{Cov}(A_{*i}, A_{*j})_{1\le i,j\le 4}$. Wir werden sehen, dass A^TA und auch AA^T für die Singulärwertzerlegung eine wichtige Rolle spielen.*

Wenn man annimmt, dass die Faktorisierung der Form (1) existiert, dann ist

$$A^TA = V\Sigma^TU^TU\Sigma V^T = V\Sigma^2V^T. \tag{2}$$

Das Produkt U^TU ist in der Darstellung rechts weggefallen, weil $U^TU = I$ ist (I Einheitsmatrix). Das Produkt $\Sigma\Sigma^T$ ist Σ^2, denn $\Sigma^T = \Sigma$. Was wir in (2) auf der rechten Seite sehen, ist nichts anderes als die Diagonalisierung von A^TA, also die Zerlegung bezüglich der Eigenwerte (sie stehen in Σ^2) und Eigenvektoren (sie stehen als Spalten in V). Wer eine Berechnungsmethode für Eigenwertzerlegungen symmetrischer Matrizen hat, kann also V und das Quadrat von Σ berechnen.

Ein ähnlicher Weg führt zu U, nämlich über die Matrix AA^T. Mit der Faktorisierung (1) ist nämlich

$$AA^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma^2 U^T, \tag{3}$$

nichts anderes als die Eigenwertzerlegung der symmetrischen Matrix AA^T.

Beispiel 2 (mini, aber instruktiv)

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}, \quad AA^T = \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix}.$$

$$A^T A = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = V\Sigma^2 V^T.$$

$$AA^T = \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} = U\Sigma^2 U^T.$$

$$A = U\Sigma V^T = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}. \tag{4}$$

$$A = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} \sqrt{10} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} = u_1 \sigma_1 v_1^T. \tag{5}$$

Gleichung (4) ist die Singulärwertzerlegung von A.

Gleichung (5) stellt A als Produkt seiner Hauptkomponente(n) dar.

2 Hauptkomponentenanalyse (PCA)

Die Hauptkomponentenanalyse *(Principal Component Analysis (PCA))* ist die Anwendung der Singulärwertzerlegung auf Datenmatrizen.

Die Videolektion Dimensionality Reduction: Principal Components Analysis, Part 3, von Bingni Brunton führt in das Thema ein. Als Beispiel dienen die Iris-Daten. Im Video heißt die Datenmatrix X. Wir übernehmen die Bezeichnung. Es schadet nicht, die vier Spalten von X zu zentrieren, dann ist nämlich auch die Verbindung zur Kovarianz

$$\frac{1}{150} X^T X = \begin{bmatrix} 0.681122 & -0.0390067 & 1.26519 & 0.513458 \\ -0.0390067 & 0.186751 & -0.319568 & -0.117195 \\ 1.26519 & -0.319568 & 3.09242 & 1.28774 \\ 0.513458 & -0.117195 & 1.28774 & 0.578532 \end{bmatrix}$$

hergestellt.

Die Singulärwertzerlegung der vollen Datenmatrix $X = U\Sigma V^T$ hat die Diagonalmatrix

$$\Sigma = \begin{bmatrix} 25.0899 & 0. & 0. & 0. \\ 0. & 6.00785 & 0. & 0. \\ 0. & 0. & 3.42054 & 0. \\ 0. & 0. & 0. & 1.8785 \end{bmatrix}.$$

und die Matrix

$$V = \begin{bmatrix} -0.36159 & -0.65654 & 0.580997 & 0.317255 \\ 0.0822689 & -0.729712 & -0.596418 & -0.324094 \\ -0.856572 & 0.175767 & -0.0725241 & -0.479719 \\ -0.358844 & 0.0747065 & -0.549061 & 0.751121 \end{bmatrix}.$$

der rechten Singulärbvektoren. Sie bilden die Orthonormalbasis des neuen Eigenschaften-Raumes.

Die Reduktion

Datenreduktion mit minimalem Informationsverlust erreicht man durch Einschränkung auf weniger Singulärwerte:

$$X \approx U_t \Sigma_t V_t^T$$

mit verkleinerten *(truncated)* Matrizen U_t, Σ_t, V_t.

Dimension 2

Setzen wir $t = 2$. Danit drücken wir aus, dass wir uns auf zwei Dimensionen beschränken. Aus unserer Datenmatrix X wird eine Datenmatrix X_2 mit zwei Spalten, wobei

$$X_2 = XV_2 = U_2\Sigma_2$$

mit

$$\Sigma_2 = \begin{bmatrix} 25.0899 & 0. \\ 0. & 6.00785 \end{bmatrix} \quad \text{und} \quad V_2 = \begin{bmatrix} -0.36159 & -0.65654 \\ 0.0822689 & -0.729712 \\ -0.856572 & 0.175767 \\ -0.358844 & 0.0747065 \end{bmatrix}$$

ist.

Jede Zeile von X_2 repräsentiert ein Iris-Exemplar und kann als Punkt in der Zeichenebene visualisiert werden (Abb. 1). Da wir wissen, welche Iris-Spezies der Punkt repräsentiert, können wir den Punkt entsprechend färben. Im PCA-Bild sieht man schön, dass es eine Beziehung zwischen Messdaten und Spezies gibt.

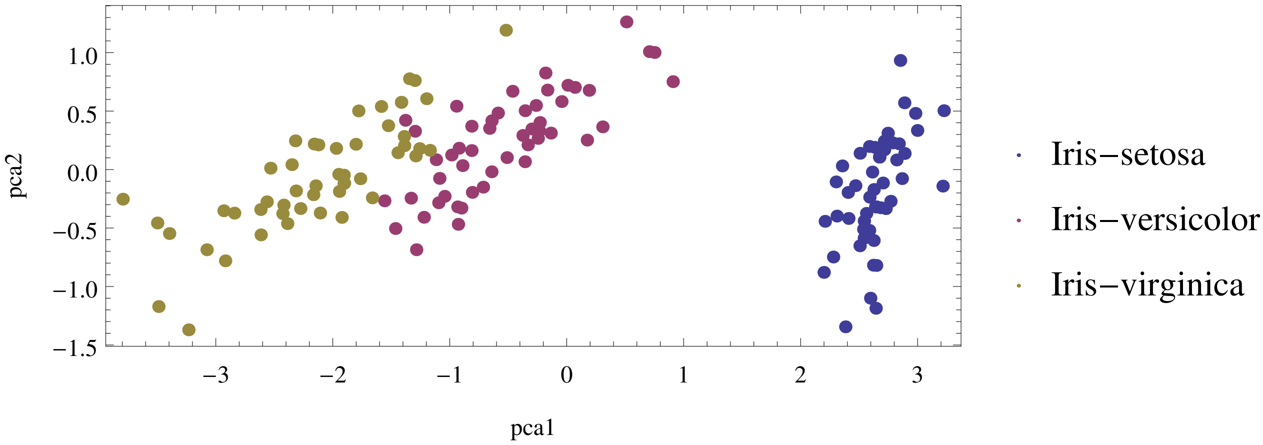

Abbildung 1: Zweidimeniionales PCA-Bild der Iris-Daten.

Dimension 1

Nun sei $t = 1$. Es ist

$$\Sigma_1 = \begin{bmatrix} 25.0899 \end{bmatrix} \quad \text{und} \quad V_1 = \begin{bmatrix} -0.36159 \\ 0.0822689 \\ -0.856572 \\ -0.358844 \end{bmatrix}.$$

Unsere Datenmatrix X_1 ist nun der Spaltenvektor XV_1. Jedes Iris-Exemplar ist nur noch durch einen skalaren Wert auf der Ordinate repräsentiert. Als Abszissenwert können wir den Fehler darstellen, den man durch die Reduktion auf eine Dimension macht. Es ist der Beitrag jedes Iris-Exemplars zur Kleinsten-Quadrate-Summe. Man sieht immer noch klar eine Beziehung zwischen Messdaten und Spezies.

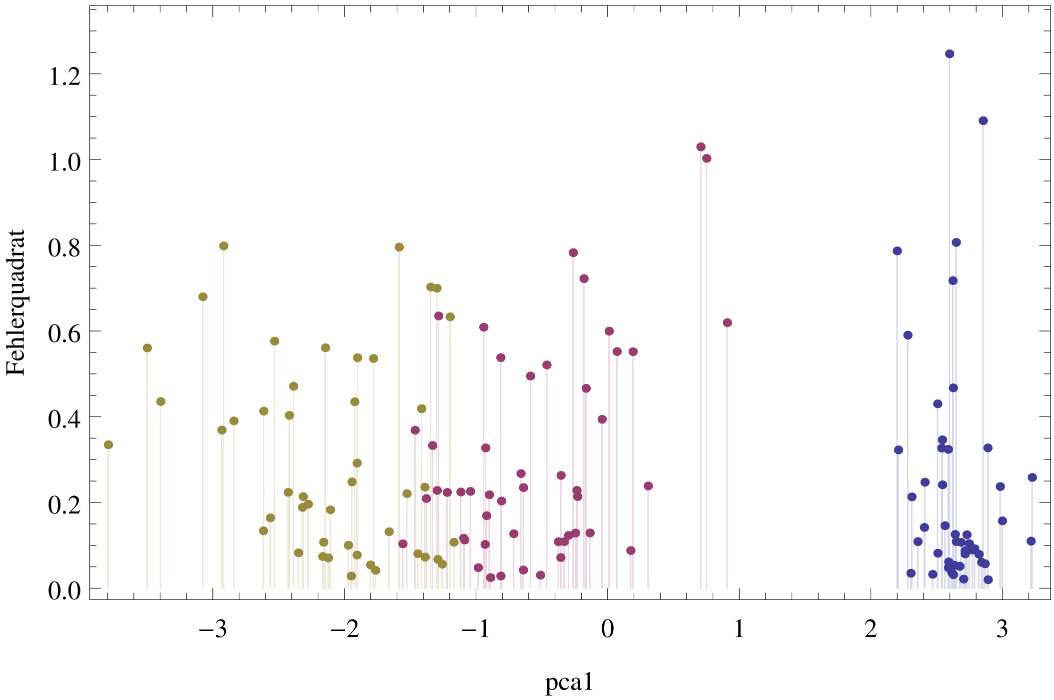

Abbildung 2: Eindimensionales PCA-Bild der Iris-Daten.