Computergestützte Mathematik zur Linearen Algebra

Singulärwertzerlegung

Christiane Helzel

4. Februar 2016

Symmetrische Matrizen

Erinnerung: Für $A \in \mathbb{R}^{n,n}$ symmetrisch, dann gilt

- alle Eigenwerte von A sind reell
- es gibt eine Orthonormalbasis $\{q_1, \ldots, q_n\}$ von Eigenvektoren von A:

$$Q^T A Q = D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

  wobei $Q = [q_1 \mid \ldots \mid q_n]$ orthogonal ($Q^T Q = QQ^T = I$)
- Es gilt: $\mathrm{Rang}(A) = p$ genau dann, wenn genau p Eigenwerte λ_j von Null verschieden sind.

Eigenschaften von $A^T A$

Lemma. Es sei $A \in \mathbb{R}^{m,n}$, $m \geq n$.

1. $A^T A$ ist symmetrisch und positiv semidefinit.
2. $A^T A$ ist genau dann positiv definit, wenn $\operatorname{kern}(A) = \{0\}$.
3. In jedem Fall gilt

$$\begin{aligned} \operatorname{kern}(A^T A) &= \operatorname{kern}(A) \\ \operatorname{bild}(A^T A) &= \operatorname{bild}(A^T) = \operatorname{kern}(A)^\perp \end{aligned}$$

Beweis

- $A^T A$ symmetrisch ist offensichtlich. Ferner ist

$$x^T A^T Ax = (Ax)^T (Ax) = ||Ax||^2 \geq 0 \quad \text{für alle } x,$$

d.h. $A^T A$ positiv semidefinit und $\mathrm{kern}(A^T A) \subset \mathrm{kern}(A)$. Wegen $\mathrm{kern}(A) \subset \mathrm{kern}(A^T A)$ folgt $\mathrm{kern}(A^T A) = \mathrm{kern}(A)$.

- $\mathrm{bild}(A^T A) \subset \mathrm{bild}(A^T)$ ist klar, Gleichheit folgt aus

$$\begin{aligned} \dim \mathrm{bild}(A^T A) &= n - \dim \mathrm{kern}(A^T A) = n - \dim \mathrm{kern}(A) \\ &= \mathrm{rang}(A) = \dim \mathrm{bild}(A) = \dim \mathrm{bild}(A^T). \end{aligned}$$

- Seien $z \in \mathrm{bild}(A^T)$ und $x \in \mathrm{kern}(A)$ beliebig, dann gibt es $y \in \mathbb{R}^m$ mit $z = A^T y$ und $x^T z = x^T A^T y = (Ax)^T y = 0$.

Singulärwertzerlegung

Es sei $A \in \mathbb{R}^{m,n}$, $m \geq n$ mit $\mathrm{Rang}(A) = p$. $\lambda_1, \ldots, \lambda_n$ seien die absteigend sortierten Eigenwerte von $A^T A$

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p > \lambda_{p+1} = \ldots = \lambda_n = 0$$

und $v_1, \ldots, v_n \in \mathbb{R}^n$ sei eine Orthonormalbasis von Eigenvektoren (von $A^T A$). Definiere

$$u_i = \frac{1}{\sqrt{\lambda_i}} A v_i, \qquad i = 1, \ldots, p,$$

dann gilt

$$u_i^T u_j = \frac{1}{\sqrt{\lambda_i}} \frac{1}{\sqrt{\lambda_j}} (A v_i)^T A v_j = \frac{1}{\sqrt{\lambda_i \lambda_j}} v_i^T (A^T A) v_j = \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} v_i^T v_j = \delta_{ij}$$

Singulärwertzerlegung II

Damit ist $\{u_1, \ldots, u_p\}$ eine Orthonormalbasis von Bild(A), denn

$$\dim \operatorname{bild}(A) = \operatorname{rang}(A) = p$$

Ergänze diese durch weitere $m - p$ Vektoren $u_{p+1}, \ldots, u_m$ zu einer Orthonormalbasis von $\mathbb{R}^m$.

$$\begin{aligned} A^T u_i &= \frac{1}{\sqrt{\lambda_i}} A^T A v_i = \sqrt{\lambda_i} v_i, && i = 1, \ldots, p \\ A^T u_i &= 0 && i = p+1, \ldots, m \end{aligned}$$

Singulärwertzerlegung III

Definition und Satz. Jede Matrix $A \in \mathbb{K}^{m,n}$ mit $\mathrm{rang}(A) = p$ besitzt eine Singulärwertzerlegung, d.h. ein System

$$\{\sigma_i, u_j, v_k \mid i = 1, \dots, p,\ j = 1, \dots, m,\ k = 1, \dots, n\}$$

mit $\sigma_1 \geq \sigma_2 \geq \dots, \geq \sigma_p > 0$ und Orthonormalbasen $\{u_j\}_{j=1}^m$ und $\{v_k\}_{k=1}^n$ des $\mathbb{K}^m$ bzw. $\mathbb{K}^n$, wobei

$$\begin{aligned} Av_i &= \sigma_i u_i, & A^T u_i &= \sigma_i v_i, & i &= 1, \dots, p, \\ Av_k &= 0, & A^T u_j &= 0, & j, k &> p \end{aligned}$$

σ_i heißen Singulärwerte von A, v_i rechte und u_i linke Singulärvektoren.

$$A = U \Sigma V^* \ \text{ mit } u_i = U(:, i) \ \text{ und } v_i = V(:, i)$$

Singulärwertzerlegung IV

In Matrizenschreibweise:

$$U = [u_1, \ldots, u_m] \in \mathbb{R}^{m,m}, \qquad V = [v_1, \ldots, v_n] \in \mathbb{R}^{n,n}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & \sigma_p & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix}$$

Dann gilt $U^T U = I$, $V^T V = I$ und

$$A = U\Sigma V^T, \qquad A^T = V\Sigma^T U^T, \qquad \Sigma = U^T A V,$$

$$A = \sum_{i=1}^{p} \sigma_i u_i v_i^T, \qquad A^T = \sum_{i=1}^{p} \sigma_i v_i u_i^T$$

Geometrische Interpretation

Durchlaufen die Vektoren

$$x = \alpha_i v_i + \alpha_j v_j + \alpha_k v_k, \qquad ||x||^2 = \alpha_i^2 + \alpha_j^2 + \alpha_k^2 = 1$$

die Einheitskugel des Unterraums $\text{span}\{v_i, v_j, v_k\}$, dann durchlaufen ihre Bilder

$$Ax = \sigma_i \alpha_i u_i + \sigma_j \alpha_j u_j + \sigma_k \alpha_k u_k, = \beta_i u_i + \beta_j u_j + \beta_k u_k,$$

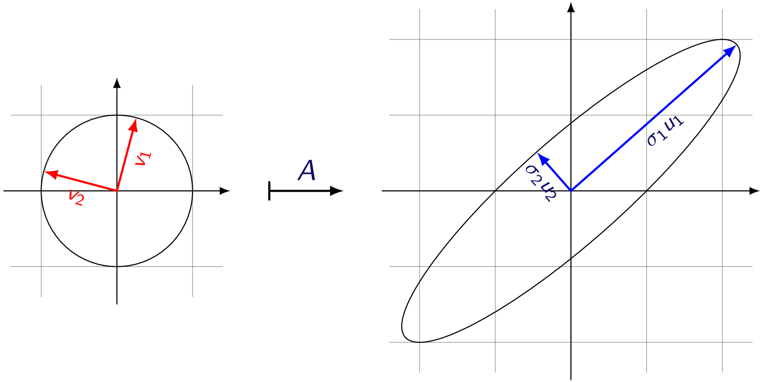
ein Ellipsoid in dem durch u_i, u_j und u_k aufgespannten Teilraum des $\mathbb{R}^m$, denn

$$\frac{1}{\sigma_i^2}\beta_i^2 + \frac{1}{\sigma_j^2}\beta_j^2 + \frac{1}{\sigma_k^2}\beta_k^2 = \alpha_i^2 + \alpha_j^2 + \alpha_k^2 = 1$$

(Ellipsoid mit Scheitelpunkten $(\pm\sigma_i, 0, 0)$, $(0, \pm\sigma_j, 0)$, $(0, 0, \pm\sigma_k)$ in den zu (u_i, u_j, u_k) gehörenden Koordinaten.)

Beispiel

$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$, Matlab: `[U,S,V] = svd(A)`



Eigenwert- / Singulärwertzerlegung

Für $A \in \mathbb{R}^{n,n}$ diagonalisierbar gibt es $X \in \mathbb{C}^{n,n}$ nicht singulär mit

$$A = X\Lambda X^{-1}, \qquad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n,n}$$

Für $A \in \mathbb{R}^{m,n}$ beliebig gibt es orthogonale Matrizen $U \in \mathbb{R}^{m,m}$, $V \in \mathbb{R}^{n,n}$ mit

$$A = U\Sigma V^T, \qquad \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m,n}$$

mit $\sigma_i \geq 0$.

Anwendung

Satz. Sei $A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$ und $A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$. Dann gilt

$$||A - A_k||_2 \leq ||A - B||_2$$

für alle Matrizen B mit $\operatorname{rang} B = k$ und $||A - A_k|| = \sigma_{k+1}$.
Es gilt auch $A_k = U\Sigma_k V^T$ mit $\Sigma_k = \operatorname{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, k)$

Anwendung: Datenkompression