VI SINGULÄRWERTZERLEGUNG UND PSEUDOINVERSE

S. BARTELS, 18.12.2013

VI.A. **Singulärwertzerlegung.** Die symmetrische und positiv semi-definite Matrix $A^\top A \in \mathbb{R}^{n\times n}$ für $A \in \mathbb{R}^{m\times n}$ spielt eine wichtige Rolle im Ausgleichsproblem. Sie ist diagonalisierbar und es existiert eine Orthonormalbasis bestehend aus Eigenvektoren $v_1, v_2, \dots, v_n$ mit zugehörigen Eigenwerten

$$\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_p > \lambda_{p+1} = \dots = \lambda_n = 0$$

mit $0 \le p \le n$, wobei Eigenwerte gemäß ihrer Vielfachheit gegebenenfalls mehrfach aufgezählt werden. Für $i = 1, 2, \dots, p$ definieren wir $u_i = \lambda_i^{-1/2} A v_i$. Für $1 \le i, j \le p$ gilt dann

$$\begin{aligned} u_i^\top u_j &= \lambda_i^{-1/2}\lambda_j^{-1/2}(Av_i)^\top(Av_j) = (\lambda_i\lambda_j)^{-1/2} v_i^\top (A^\top A v_j) \\ &= (\lambda_i\lambda_j)^{-1/2} v_i^\top(\lambda_j v_j) = \lambda_j(\lambda_i\lambda_j)^{-1/2} v_i^\top v_j = \delta_{ij}. \end{aligned}$$

Die Vektoren $(u_1, u_2, \dots, u_p)$ bilden also ein Orthonormalsystem im $\mathbb{R}^m$, genauer eine Orthonormalbasis des Unterraums $\operatorname{Im} A$. Wir ergänzen es durch Vektoren $(u_{p+1}, u_{p+2}, \dots, u_m)$ zu einer Orthonormalbasis $(u_1, u_2, \dots, u_m)$ des $\mathbb{R}^m$. Es gilt

$$A^\top u_i = \lambda_i^{-1/2} A^\top A v_i = \lambda_i^{1/2} v_i, \quad i = 1, 2, \dots, p.$$

Aus $\ker A^\top A = \ker A$ folgt $\operatorname{Im} A = \operatorname{span}\{u_1, u_2, \dots, u_p\}$ und somit die Inklusion $\{u_{p+1}, \dots, u_m\} \subset (\operatorname{Im} A)^\perp = \ker A^\top$ beziehungsweise

$$A^\top u_i = 0, \quad i = p+1, \dots, m.$$

Unter Verwendung von $\sigma_i = \lambda_i^{1/2}$, $i = 1, 2, \dots, p$, erhalten wir folgenden Satz.

Satz VI.1. *Sei $A \in \mathbb{R}^{m\times n}$. Dann existieren Zahlen $\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_p > 0$ und Orthonormalbasen $(u_i)_{i=1,\dots,m}$ des $\mathbb{R}^m$ und $(v_j)_{j=1,\dots,n}$ des $\mathbb{R}^n$ mit den Eigenschaften*

$$\begin{aligned} &Av_i = \sigma_i u_i, \quad && A^\top u_i = \sigma_i v_i, \quad && i = 1, 2, \dots, p, \\ &Av_j = 0, && A^\top u_k = 0, && j = p+1, \dots, n,\ k = p+1, \dots, m. \end{aligned}$$

Die Zahlen σ_i^2, $i = 1, 2, \dots, p$, sind genau die von Null verschiedenen Eigenwerte von $A^\top A$. Für

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

gilt $U \in O(m)$ und $V \in O(n)$ und mit

$$\Sigma = \begin{bmatrix} \sigma_1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_p & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \in \mathbb{R}^{m \times n}$$

folgt

$$A = U\Sigma V^\top = \sum_{i=1}^p \sigma_i u_i v_i^\top, \quad A^\top = V\Sigma^\top U^\top = \sum_{i=1}^p \sigma_i v_i u_i^\top.$$

Die Faktorisierung heißt Singulärwertzerlegung *und im Englischen* singular value decomposition (SVD).

Beweis. Die Aussagen folgen unmittelbar aus der Konstruktion. □

VI.B. **Pseudoinverse.** Mit Hilfe der Singulärwertzerlegung lässt sich der Begriff der inversen Matrix auf nicht-reguläre und nicht-quadratische Matrizen verallgemeinern.

Definition VI.2. *Ist $A = U\Sigma V^\top$ die Singulärwertzerlegung von A und $\Sigma^+ \in \mathbb{R}^{n \times m}$ definiert durch*

$$\Sigma^+ = \begin{bmatrix} \sigma_1^{-1} & & & 0 \\ & \ddots & & \vdots \\ & & \sigma_p^{-1} & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times m},$$

so heißt $A^+ = V\Sigma^+ U^\top = \sum_{i=1}^p \sigma_i^{-1} v_i u_i^\top \in \mathbb{R}^{n \times m}$ Pseudoinverse *oder* Moore–Penrose-Inverse *von A.*

Bemerkungen VI.3. *(i) Nach Wahl der Vektoren u_i, v_j gilt $\ker A^+ = \ker A^\top$ und $\operatorname{Im} A^+ = \operatorname{Im} A^\top$.*
(ii) Die Matrix A^+ ist die eindeutig bestimmte Lösung in $\mathbb{R}^{n \times m}$ der algebraischen Gleichungen

$$AXA = A, \quad XAX = X, \quad (AX)^\top = AX, \quad (XA)^\top = XA.$$

Beispielsweise gilt wegen $U^\top U = I_m$ und $V^\top V = I_n$, dass

$$A^+ A A^+ = (V\Sigma^+ U^\top)(U\Sigma V^\top)(V\Sigma^+ U^\top) = V\Sigma^+\Sigma\Sigma^+ U^\top = V\Sigma^+ U^\top = A^+.$$

Mit der Pseudoinversen lässt sich das Ausgleichsproblem lösen.

Satz VI.4. *Der Vektor $A^+ b$ ist unter allen Lösungen des Ausgleichsproblems diejenige mit minimaler Norm.*

Beweis. Siehe Vorlesung. □

Bemerkung VI.5. *Gilt* $\operatorname{rank} A = n \le m$*, so folgt aus* $A^+ b = (A^\top A)^{-1} A^\top b$ *für alle* $b \in \mathbb{R}^m$*, dass* $A^+ = (A^\top A)^{-1} A^\top$ *und insbesondere* $A^+ = A^{-1}$ *falls* $n = m$*.*