Statistical Data Analysis

Dr. Jana de Wiljes

1. Februar 2022

Universität Potsdam

Singular value decomposition

Singular value decomposition

Before: for quadratic matrices we had Eigenvalues and Eigenvectors that can be used to diagonalise a matrix

$x^5 + 3x^2 + 2x + 2$

$(x - x_0)(x - x_1) \cdot (\ldots)$

$p(x)$

Now:

- similar concept for non quadratic matrices
- the corresponding scalars are called Singular values which opposed to the Eigenvalues are **always** real
- Although similarity exist singular value decomposition is not an generalization of Eigenvalues/Eigenvector approach
- the rank of a matrix can be determined in a numerical stable way

Definition

- $\mathrm{GL}(n, \mathbb{R})$ general linear group of degree n is the set of $n \times n$ invertible matrices
- $\mathrm{O}(n) = \left\{ Q \in \mathrm{GL}(n, \mathbb{R}) \mid Q^\mathsf{T} Q = Q Q^\mathsf{T} = I \right\}.$

$$\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

Singular value decomposition

Theorem

$m < n$ $\quad$ $(m<n)$

Let $A \in \mathbb{R}^{m\times n}$ *be a matrix. Then* $\sigma_1, \ldots, \sigma_p \in \mathbb{R}$ *with* $\sigma_1 \geq \cdots \geq \sigma_p \geq 0$ *as well as* $U \in O(m)$ *and* $V \in O(n)$ *exist, so that*

$A = S\, D\, S^{-1}$

$$U^t A V = \Sigma := \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \\ \vdots & & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix},$$

$m = n$

$n < m$

$m < n$

wobei $p = \min(m, n)$*. The values* σ_i *are called* **singular values** *of* A*. A representation of the form* $A = U\Sigma V^t$ *is called* **singular value decomposition** *(***SVD***).*

Example

- For a quadratic matrix:

$$A_1 = \begin{pmatrix} 4 & 12 \\ 12 & 11 \end{pmatrix} = \begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix} \cdot \begin{pmatrix} 20 & 0 \\ 0 & 5 \end{pmatrix} \cdot \begin{pmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{pmatrix}$$

- SDV of orthogonal matrizes:

$$A_2 = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

-

$$A_3 = \begin{pmatrix} 0.36 & 1.60 & 0.48 \\ 0.48 & -1.20 & 0.64 \end{pmatrix} = \begin{pmatrix} 0.8 & 0.6 \\ -0.6 & 0.8 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0.6 & 0 & 0.8 \\ -0.8 & 0 & 0.6 \end{pmatrix}$$

Remark

- Note that

$$\operatorname{rang}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0, \quad \operatorname{rang}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1.$$

- yet the two eigenvalues in both cases are 0 and 0. The singular values on the other hand are 0, 0 and 0, 1 respectively, i.e., in this case the eigenvalues do not tell you anything about the rank of the matrix but the number of singular values of the matrix correspond to its rank

- Consider for $\varepsilon > 0$:

$$A = \begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix}.$$

Since $\chi_A(t) = t^2 - \varepsilon$ the corresponding eigenvalues are $\pm\sqrt{\varepsilon}$. The singular values are $\sigma_1 = 1, \sigma_2 = \varepsilon$ and for ε converging towards 0, the rank of matrix is converging towards 1

Proof

Proof to show $U^T A V = \Sigma$ ex. $\boxed{m < n}$

$\boxed{A = U \Sigma V^T}$

We will construct a SVD of A: • consider $\boxed{B = A^t \cdot A}$ (note that B is a real-valued symmetric with real eigenvalues λ_i

• $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_n$ and let $\{v_1, .., v_n\}$ be a basis of orthnomal Eigenvectors corresponding to λ_i (for example using Gram Schmidt)

$\boxed{A^t = A^\perp}$

$i < n$

$\langle a_1, \dots a_i \rangle = \langle v_1, \dots v_i \rangle$

$\langle a_1, \dots, a_n \rangle = \langle v_1 \dots v_n \rangle$

$(x - \lambda_1)^0$ $\langle v_n \rangle = \langle a_n \rangle$

• λ_i are non-negative $v_i^t \cdot B \cdot v_i = \lambda_i \underbrace{v_i^t \cdot v_i} = \lambda_i$

$x = \boxed{Av_i} = 0$

$x^t \cdot x$ $(A \cdot v_i)^t$

$\rightarrow \boxed{v_i^t B \cdot v_i} = \boxed{v_i^t A^t A v_i} = \boxed{\langle Av_i, Av_i \rangle} \ge 0$ (B is semi positive definite)

↑ using rules of scalar product $\boxed{\langle x, x \rangle \ge 0}$

• rank(B) = rank(A) = r ⇒ the first $\lambda_1, \dots, \lambda_r$ are positive (i.e. $\neq 0$)

Construct the following vectors for $\boxed{i = 1, \dots, r}$ $r < m$

$$\boxed{u_i := \frac{1}{\sqrt{\lambda_i}} A v_i}$$

and augment those by $m - r$ orthonormal vectors $\boxed{u_{r+1}, \dots, u_m}$ that are also $\boxed{\text{orthonormal}}$ with respect to $u_1, \dots, u_r$ to basis of the $\mathbb{R}^m$

to show u_i u_j are orthonormal for $i, j \in \{1, \dots, r\}$

$$\boxed{u_i^T u_j} = \frac{1}{\sqrt{\lambda_i \lambda_j}} v_i^T \underbrace{\boxed{A^T A}}_{\substack{B \\ \lambda_j \cdot v_j}} v_j = \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} \boxed{v_i^T \cdot v_j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

$\sqrt{\lambda_i^2}$ $\quad$ $\boxed{1 \quad i = j}$ $\quad$ $0 \quad i = j$

Set Matrices U and V to be: $U = \begin{pmatrix} | & & | \\ u_1 & \dots & u_m \\ | & & | \end{pmatrix}$ and $V = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$

$$\boxed{U \Sigma V^T} = \sum_{i=1}^{r} \sqrt{\lambda_i}\, u_i v_i^t = \sum_{i=1}^{r} A v_i \cdot v_i^t = A \cdot \underbrace{\sum_{i=1}^{n} v_i v_i^t}_{I} = A \cdot I = A$$

$+0 \dots +0$ $\quad$ $n-r$ $\quad$ v_1 $\quad$ v_1^t

where $\sigma_i := \sqrt{\lambda_i}$ for $i = 1, \dots, r$ and $\sigma_i = 0$ $i = r+1, \dots, p$

$\boxed{A^T A v_j = 0}$

$j > r$

$n \times 1 \cdot 1 \times n$

$(n \times n)$

$$\begin{pmatrix} \sigma_1 & 0 & \dots & \sigma_r & & 0 \\ 0 & & & & 0 & 0 \dots \end{pmatrix} \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix}$$

$\begin{pmatrix} 1 & 0 \\ 0 & \end{pmatrix}$

$\begin{pmatrix} 0 & 1 \end{pmatrix}$

B

$\square$

Proof

Singular value decomposition

Theorem

Let $A = U\Sigma V$ be the singular value decomposition of $A \in \mathbb{R}^{m\times n}$ with singular values $\sigma_1 \geq \cdots \geq \sigma_p$ für $p = \min(m,n)$. Let $u_1, \ldots, u_m$ and $v_1, \ldots, v_n$ denote the columns of U and V respectively. Then the following holds:

- $\boxed{Av_i = \sigma_i u_i}$ *and* $\boxed{A^t u_i = \sigma_i v_i}$ *für* $i = 1, 2, \ldots, p$.
- *For* $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0$ *follows that* $\operatorname{rang} A = r$. *Furthermore,*

$$\boxed{\mathbf{Ker}(A) = \langle v_{r+1}, \ldots, v_n\rangle} \text{ und } \boxed{\mathbf{Im}(A) = \langle u_1, \ldots, u_r\rangle}.$$

- *the squares $\sigma_1^2, \ldots, \sigma_p^2$ of the singular values are the eigenvalues of $\boxed{A^t A}$ and of $\boxed{AA^t}$ to the corresponding eigen vectors $\boxed{v_1, \ldots, v_p}$ and $\boxed{u_1, \ldots, u_p}$ respectively.*

B

Golub-Reinsch algorithm

$m \times n$ $\quad n < n$

$\|Ax - b\|$

$\begin{pmatrix} R \\ 0 \end{pmatrix} x = Qb$

Input: A$\in \mathbb{R}^{n\times m}$, $m \leq n$, ϵ

- $\begin{bmatrix} B \\ 0 \end{bmatrix} = (U_1, \ldots, U_n)^\top A(V_1 \ldots V_n - 2)$ where U_i and V_j are householder transformations
- Set q=0
- **while** $(q < n)$
  1. set $B(i, i+1) = 0$ if for any $i = 1, \ldots, n-1$ $B(i, i+1) \leq \epsilon(|B(i,i)| + |B(i+1, i+1)|)$
  2. Determine the smallest p and the largest q so that B can be blocked as

$$B = \begin{bmatrix} B_{1,1} & 0 & 0 \\ 0 & B_{2,2} & 0 \\ 0 & 0 & B_{3,3} \end{bmatrix} \tag{1}$$

$p \times p$ $\quad q \times q$ $\quad n-p-q \times n-p-q$

where $B_{3,3}$ is diagonal and $B_{2,2}$ has no zero superdiagonal entry.

  3. If $q = n$, set $\Sigma =$ the diagonal portion of B **STOP**
  4. If for $i = p+1, \ldots, n-q-1$ $B_{i,i} = 0$, then
     - Apply Givens rotations so that $B_{i,i+1} = 0$ and $B_{2,2}$ is still upper bidiagonal.
  5. else Golub Kahan SVD step: This step is essentially applying the QR method to the symmetric tridiagonal matrix $T = BB$

$O(n^3)$

Σ

Householder transformation

$A \in \mathbb{R}^{\boxed{n \times m}}$ $m<n$

σ

x x x

$$H_1$$

$$U_1^* A = \begin{bmatrix} \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix} \longrightarrow U_1^* A V_1 = \begin{bmatrix} x & \mathbf{x} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}$$

4

5

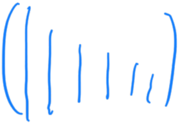
$$\longrightarrow U_2^* U_1^* A V_1 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{0} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{0} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{0} & \mathbf{x} & \mathbf{x} \end{bmatrix} \longrightarrow U_2^* U_1^* A V_1 V_2 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & x & \mathbf{x} & \mathbf{0} \\ 0 & 0 & \mathbf{x} & \mathbf{x} \\ 0 & 0 & \mathbf{x} & \mathbf{x} \\ 0 & 0 & \mathbf{x} & \mathbf{x} \end{bmatrix}$$

$$\longrightarrow U_3^* U_2^* U_1^* A V_1 V_2 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & x & x & 0 \\ 0 & 0 & \mathbf{x} & \mathbf{x} \\ 0 & 0 & \mathbf{0} & \mathbf{x} \\ 0 & 0 & \mathbf{0} & \mathbf{x} \end{bmatrix} \longrightarrow U_4^* U_3^* U_2^* U_1^* A V_1 V_2 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & x & x & 0 \\ 0 & 0 & x & x \\ 0 & 0 & 0 & \mathbf{x} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix} = B.$$

Householder transformation

$A \in \mathbb{R}^{m\times n}$ H

$$A = \begin{pmatrix} 2 & 4 & -4 \\ 1 & 1 & 2 \\ 2 & -3 & 0 \end{pmatrix}$$



$$a_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \leadsto \|a_1\| = \sqrt{4+1+4} = 3$$

$$v^{(1)} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + \mathrm{sign}(2) \cdot 3 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix}$$

$$v^{(1)} = a_1 + \mathrm{sgn}(a_{11})\, \|a_1\| \cdot e_1$$

$$v^{(1)T} v^{(1)} = 25 + 1 + 4 = 30$$

$$H = \underbrace{E}_{Id} - \frac{2\, v^{(1)} \cdot v^{(1)T}}{v^{(1)T} \cdot v^{(1)}}$$

$$v^{(1)} \cdot v^{(1)T} = \begin{pmatrix} 5 \\ 1 \\ 2 \end{pmatrix} (5\ 1\ 2) = \begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 2 & 4 \end{pmatrix}$$

$$H^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{2}{30} \begin{pmatrix} 25 & 5 & 10 \\ 5 & 1 & 2 \\ 10 & 2 & 4 \end{pmatrix} = \begin{pmatrix} -2/3 & -1/3 & -2/3 \\ -1/3 & 14/15 & -2/15 \\ -2/3 & -2/15 & 11/15 \end{pmatrix}$$

$$A^{(1)} = H^{(1)} A = \begin{pmatrix} -3 & -1 & 2 \\ 0 & 0 & 16/5 \\ 0 & -5 & 12/5 \end{pmatrix}$$

$$v^{(2)} = a_1' + \mathrm{sgn}(0) \cdot 5 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ -5 \end{pmatrix}$$

$$H' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{2}{50} \begin{pmatrix} 25 & -25 \\ -25 & 25 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Householder transformation

$$H^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}} \end{pmatrix}$$

$$H'_2$$

$$R = H^{(2)} \cdot A^{(1)}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -3 & -1 & 2 \\ 0 & 0 & 16/5 \\ 0 & -5 & 12/5 \end{pmatrix}$$

$$= \begin{pmatrix} -3 & -1 & 2 \\ 0 & -5 & 12/2 \\ 0 & 0 & 16/5 \end{pmatrix}$$

$$R = \underbrace{H^{(2)} \cdot H^{(1)}}_{Q^T} A$$

$$Q^T Q = Id$$

$$A = QR$$

$$Q = H^{(1)T} H^{(2)T}$$

Householder transformation

Householder transformation

$$
U_1^* A = \begin{bmatrix} \mathbf{x} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ \mathbf{0} & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix} \quad \longrightarrow \quad U_1^* A V_1 = \begin{bmatrix} x & \mathbf{x} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \end{bmatrix}
$$

$$
\longrightarrow \quad U_2^* U_1^* A V_1 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & \mathbf{x} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{0} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{0} & \mathbf{x} & \mathbf{x} \\ 0 & \mathbf{0} & \mathbf{x} & \mathbf{x} \end{bmatrix} \quad \longrightarrow \quad U_2^* U_1^* A V_1 V_2 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & x & \mathbf{x} & \mathbf{0} \\ 0 & 0 & \mathbf{x} & \mathbf{x} \\ 0 & 0 & \mathbf{x} & \mathbf{x} \\ 0 & 0 & \mathbf{x} & \mathbf{x} \end{bmatrix}
$$

$$
\longrightarrow \quad U_3^* U_2^* U_1^* A V_1 V_2 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & x & x & 0 \\ 0 & 0 & \mathbf{x} & \mathbf{x} \\ 0 & 0 & \mathbf{0} & \mathbf{x} \\ 0 & 0 & \mathbf{0} & \mathbf{x} \end{bmatrix} \quad \longrightarrow \quad U_4^* U_3^* U_2^* U_1^* A V_1 V_2 = \begin{bmatrix} x & x & 0 & 0 \\ 0 & x & x & 0 \\ 0 & 0 & x & x \\ 0 & 0 & 0 & \mathbf{x} \\ 0 & 0 & 0 & \mathbf{0} \end{bmatrix} = B.
$$

Singular value decomposition

Remark: For symmetric matrices A the singular values are the absolut values of the eigenvalues of A. In case all eigenvalues are non-negativ, $\boxed{A = S^t \Lambda S}$ is the SVD.

Pseudoinverse

Definition

Let $A \in \mathbb{R}^{m \times n}$. A matrix $A^+ \in \mathbb{R}^{n \times m}$ is called the **pseudoinverse** of A, if $\forall b \in \mathbb{R}^m$ the vector $\boxed{x = A^+ b}$ is the solution of the minimalisation problem

A^{-1}

$$\text{Find } x \text{, so that } \boxed{\|b - Ax\|_2} \text{ is minimal}$$

i.e., $\underbrace{\|b - \boxed{AA^+}_{Id}\, b\|}_{0} = \min_{x \in \mathbb{R}^n} \|b - Ax\|.$

$Ax = b$

$x = A^{-1} b$

$\tilde{x} = A^+ b$

Motivation

Note: for a quadratic invertible matrix A the pseudoinverse is: $A^+ = A^{-1}$

Application: in case the system $Ax = b$ does not have a solution, it is possible to obtain the best approximation $\tilde{x} = A^+ b$ via the pseudoinverse A^+ i.e., the one that minimizes the error $\|Ax - b\|$ (note that is the solution of the least squares problem.

Note that A^+ can be consider as a lineare mapping. Then the following holds

- $$AA^+ \colon \mathbb{R}^m \to \mathbf{Im}(A)$$
  is the orthogonal projection to image of A and
- $$A^+A \colon \mathbb{R}^n \to (\mathbf{Ker}A)^\perp$$
  ist the orthogonal projection to the orthogonal complement von the kernel of A.

SVD and Pseudoinverse

Theorem

Let $A \in \mathbb{R}^{m \times n}$ and let $A = U\Sigma V^t$ be the corresponding singular value decomposition with singular value $\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0$. Then we define

$$\Sigma^+ = \begin{pmatrix} \frac{1}{\sigma_1} & & & 0 \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ 0 & & & 0 \end{pmatrix}$$

and the matrix $A^+ = V\Sigma^+U^t \in \mathbb{R}^{n \times m}$ is the pseudo inverse of A.

$\boxed{\|Ax - b\|}$