

Rozkład SVD macierzy

Jakub Kuciński Karol Kuczmarz

18 maja 2020

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Wektor własny i wartość własna

Definicja

Niech $A \in M_{n \times n}(\mathbb{R})$. Wektor $v \in \mathbb{R}^n$ nazywamy **wektorem własnym** A , jeżeli

$$(\exists \lambda \in \mathbb{R})(Av = \lambda v)$$

Wektor własny i wartość własna

Definicja

Niech $A \in M_{n \times n}(\mathbb{R})$. Wektor $v \in \mathbb{R}^n$ nazywamy **wektorem własnym** A , jeżeli

$$(\exists \lambda \in \mathbb{R})(Av = \lambda v)$$

Definicja

Wartością własną macierzy $A \in M_{m \times n}(\mathbb{R})$ nazywamy $\lambda \in \mathbb{R}$ takie, że istnieje niezerowy wektor własny dla λ (tzn. $(\exists v \in \mathbb{R}^n \setminus \{0\})(Av = \lambda v)$).

Diagonalizacja macierzy

Twierdzenie

Jeżeli macierz $A \in M_{n \times n}(\mathbb{R})$ ma n liniowo niezależnych wektorów własnych $v_1, \dots, v_n \in \mathbb{R}^n$ i $\lambda_1, \dots, \lambda_n$ to odpowiadające im wartości własne, to macierz A możemy przedstawić jako

$$A = PDP^{-1}$$

gdzie:

- $P = (v_1, \dots, v_n)$
- $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ diagonalna.

Diagonalizacja macierzy

Uwaga

Nie wszystkie macierze można diagonalizować. Można przejść w liczby zespolone, ale dalej możemy mieć wielokrotne wartości własne. Uogólnienie – *Twierdzenie Jordana*.

Standardowy iloczyn skalarny

Niech $v, u \in \mathbb{R}^n$.

$$u \circ v = \sum_{i=1}^n u_i \cdot v_i = u^T v$$

Dodatkowe pojęcia cz. 1

Definicja

Macierz kwadratową $A \in M_{n \times n}(\mathbb{R})$ nazywamy macierzą ortogonalną, jeżeli zachodzi $AA^T = A^T A = I$ (wtedy też $A^T = A^{-1}$).

Definicja

Zbiór wektorów będący bazą przestrzeni \mathbb{R}^n taki, że wektory są parami prostopadłe i mające długość 1 nazywamy bazą ortonormalną.

Dodatkowe pojęcia cz. 2

Uwaga

Ortogonalności macierzy są równoważne warunki:

- kolumny A stanowią bazę ortonormalną \mathbb{R}^n
- wiersze A stanowią bazę ortonormalną \mathbb{R}^n

Macierze symetryczne

Do rozkładu SVD potrzebne nam będą szczególnie macierze symetryczne, czyli takie, które są niezmiennicze na operację transpozycji ($A = A^T$). Mają one kilka ciekawych własności.

Twierdzenie spektralne

Twierdzenie

Każda symetryczna macierz $A \in M_{n \times n}(\mathbb{R})$ diagonalizuje się w pewnej bazie ortonormalnej i wszystkie jej wartości własne są rzeczywiste.

Dodatnia określoność

Definicja

Macierz symetryczna $A \in M_{n \times n}(\mathbb{R})$ jest dodatnio określona, jeżeli dla każdego $v \in \mathbb{R}^n \setminus \{0\}$ zachodzi

$$v \circ Av > 0$$

Jeżeli nierówność jest słaba, to mówimy, że macierz jest dodatnio półokreślona.

Dodatnia określoność

Dodatnia określoność (półokreśloność) jest równoznaczna ze wszystkimi dodatnimi (nieujemnymi) wartościami własnymi macierzy. Z twierdzenia spektralnego mamy $A = PDP^T$. Wtedy

$$v \circ Av = v \circ PDP^T v = (P^T v) \circ D(P^T v) > 0$$

Dodatnia określoność

Dodatnia określoność (półokreśloność) jest równoznaczna ze wszystkimi dodatnimi (nieujemnymi) wartościami własnymi macierzy. Z twierdzenia spektralnego mamy $A = PDP^T$. Wtedy

$$v \circ Av = v \circ PDP^T v = (P^T v) \circ D(P^T v) > 0$$

Skoro P^T jest ortogonalna, to jest macierzą zmiany bazy, czyli powyższa równość jest tożsama z

$$v \circ Dv > 0$$

$$v \circ Dv = \sum_{i=1}^n v_i \cdot (\lambda_i v_i) = \sum_{i=1}^n \lambda_i \cdot v_i^2$$

Własności macierzy $A^T A$ i AA^T

Lemat

Dla dowolnej macierzy $A \in M_{m \times n}(\mathbb{R})$ macierze AA^T i $A^T A$ są symetryczne i dodatnio półokreślone.

Własności macierzy $A^T A$ i AA^T

Lemat

Dla dowolnej macierzy $A \in M_{m \times n}(\mathbb{R})$ macierze AA^T i $A^T A$ są symetryczne i dodatnio półokreślone.

Dowód

Macierze $A^T A$ i AA^T są odpowiednio wymiaru $n \times n$ i $m \times m$.
Najpierw pokażemy symetryczność.

$$(AA^T)^T = (A^T)^T A^T = AA^T$$

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

Własności macierzy $A^T A$ i AA^T

Dowód cd.

Pokażemy dodatnią półokreśloność $A^T A$, dowód dla macierzy AA^T jest analogiczny. Weźmy dowolny wektor $v \in \mathbb{R}^n \setminus \{0\}$.

$$v \circ A^T A v = A v \circ A v \geq 0$$

z definicji iloczynu skalarnego.

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Twierdzenie

Twierdzenie

Każdą macierz $A \in M_{m \times n}(\mathbb{R})$ można przedstawić w postaci iloczynu $A = U\Sigma V^T$, gdzie:

Twierdzenie

Twierdzenie

Każdą macierz $A \in M_{m \times n}(\mathbb{R})$ można przedstawić w postaci iloczynu $A = U\Sigma V^T$, gdzie:

- $U \in M_{m \times m}(\mathbb{R})$ jest macierzą ortogonalną
- $\Sigma \in M_{m \times n}(\mathbb{R})$ jest macierzą postaci $\begin{pmatrix} D & 0 \end{pmatrix}$ lub $\begin{pmatrix} D \\ 0 \end{pmatrix}$ dla D diagonalnej z nieujemnymi wyrazami na przekątnej uporządkowanymi nierosnąco
- $V \in M_{n \times n}(\mathbb{R})$ jest macierzą ortogonalną.

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Uproszczenie

Przeprowadzimy dowód dla założenia, że $m = n$ i wszystkie wartości własne macierzy $A^T A$ są dodatnie.

Dowód cz. I

Macierz $A^T A$ jest symetryczna, więc możemy skorzystać z twierdzenia spektralnego i zapisać ją w postaci

$$A^T A = V D V^T$$

$$V^T A^T A V = D$$

$$(A^T A)^{-1} = (V D V^T)^{-1} = V D^{-1} V^T$$

gdzie D diagonalna

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

oraz V ortogonalna.

Dowód cz. II

Ponadto bez straty ogólności możemy założyć, że wartości na przekątnej macierzy D są uporządkowane nierosnąco, czyli $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$.

Niech $U = AVD^{-\frac{1}{2}}$. Zauważmy, że

$$UD^{\frac{1}{2}}V^T = AVD^{-\frac{1}{2}}D^{\frac{1}{2}}V^T = AVIV^T = AI = A$$

Wystarczy teraz pokazać, że U jest ortogonalna. Pokażemy, że $UU^T = U^TU = I$

Dowód cz. III

$$\begin{aligned} UU^T &= AVD^{-\frac{1}{2}}(AVD^{-\frac{1}{2}})^T = AVD^{-\frac{1}{2}}D^{-\frac{1}{2}}V^T A^T = \\ &= AVD^{-1}V^T A^T = A(A^T A)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = I \\ U^T U &= (AVD^{-\frac{1}{2}})^T AVD^{-\frac{1}{2}} = \\ &= D^{-\frac{1}{2}}V^T A^T AVD^{-\frac{1}{2}} = D^{-\frac{1}{2}}DD^{-\frac{1}{2}} = I \end{aligned}$$

Zatem U jest ortogonalna.

Uwaga konstrukcyjna

Uwaga

Macierz U jest macierzą wektorów własnych macierzy AA^T .

Uwaga konstrukcyjna

Uwaga

Macierz U jest macierzą wektorów własnych macierzy AA^T .

Dowód

Mamy

$$\begin{aligned} U &= AVD^{-\frac{1}{2}} = A \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\lambda_n}} \end{pmatrix} = \\ &= \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} Av_1 & \dots & \frac{1}{\sqrt{\lambda_n}} Av_n \end{pmatrix} = \begin{pmatrix} u_1 & \dots & u_m \end{pmatrix} \end{aligned}$$

czyli $u_i = \frac{1}{\sqrt{\lambda_i}} Av_i$.

Uwaga konstrukcyjna

Pokażemy, że u_i jest wektorem własnym macierzy AA^T dla wartości własnej λ_i .

$$AA^T \frac{1}{\sqrt{\lambda_i}} Av_i = \frac{1}{\sqrt{\lambda_i}} A(A^T Av_i) = \frac{1}{\sqrt{\lambda_i}} A\lambda_i v_i = \sqrt{\lambda_i} Av_i$$
$$AA^T \frac{1}{\sqrt{\lambda_i}} Av_i = \sqrt{\lambda_i} Av_i$$

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Dwa kroki

- Bidiagonalizacja macierzy
- Rozkład SVD macierzy bidiagonalnej

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

Transformacje Householdera

Do bidiagonalizacji będziemy wykorzystywać macierze transformacji Householdera. Wyraża się je wzorem

$$A = I - 2vv^T$$

gdzie v jest wektorem długości 1. Wtedy macierz A jest macierzą ortogonalną.

Bidiagonalizacja

W pierwszym kroku mnożymy naszą macierz przez taką macierz P_1 ,
że wyzerujemy wyrazy z pierwszej kolumny.

$$P_1 \cdot \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{pmatrix} = \begin{pmatrix} * & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{pmatrix}$$

Bidiagonalizacja

Następnie z prawej strony mnożymy przez macierz Q_1 taką, że wyzerujemy wyrazy z pierwszego wiersza.

$$\begin{pmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \dots & * \end{pmatrix} \cdot Q_1 = \begin{pmatrix} * & * & 0 & \dots & 0 \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & * & * & \dots & * \end{pmatrix}$$

Bidiagonalizacja

Wyzerowanie odpowiednich elementów kolumn i wierszy w macierzy A tworzy ciąg macierzy Q_1, \dots, Q_n i P_1, \dots, P_n . Niech

$$P = P_n \dots P_1$$

$$Q = Q_1 \dots Q_n$$

wtedy $PAQ = J$, gdzie J bidiagonalna.

Metoda iteracyjna

$$S_n \dots S_1 \cdot J \cdot T_1 \dots T_n \xrightarrow{n \rightarrow \infty} \begin{pmatrix} * & & & \\ & * & & \\ & & \ddots & \\ & & & * \end{pmatrix}$$

- metoda iteracyjna – koszt pojedynczego kroku: $O(n^2)$
- metoda "bezpośrednia" – koszt całości: $O(n^3)$

Macierze obrotu Givensa

$$G(i, j, \theta) = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & -s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

gdzie $c = \cos(\theta)$ i $s = \sin(\theta)$

Macierze obrotu Givensa

Dokładniej $G(i, j, \theta)$ jest macierzą identycznościową z wyjątkiem wyrazów

$$g_{i,i} = g_{j,j} = c$$

$$g_{i,j} = -g_{j,i} = s$$

Zauważmy, że ta macierz jest ortogonalna.

Pojedyncza iteracja

Najpierw mnożymy z prawej strony przez macierz $T_2 = G(1, 2, \theta)$.

$$\begin{pmatrix} * & * & 0 & \dots & 0 \\ 0 & * & * & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * \end{pmatrix} \cdot T_2 = \begin{pmatrix} * & * & 0 & \dots & 0 \\ * & * & * & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * \end{pmatrix}$$

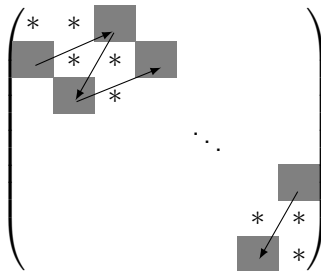
Pojedyncza iteracja

Następnie dobieramy macierz obrotu Givensa S_2 tak, by po przemnożeniu z lewej strony wyzerować element $(2, 1)$.

$$S_2 \cdot \begin{pmatrix} * & * & 0 & \dots & 0 \\ * & * & * & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * \end{pmatrix} = \begin{pmatrix} * & * & * & \dots & 0 \\ 0 & * & * & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & * \end{pmatrix}$$

Pojedyncza iteracja

Można to zobrazować następująco:



Macierze ortogonalne

Mamy macierz diagonalną, ale jak odtworzyć macierze ortogonalne?

Macierze ortogonalne

Mamy macierz diagonalną, ale jak odtworzyć macierze ortogonalne?
Należy je spamiętywać.

Macierze ortogonalne

Mamy macierz diagonalną, ale jak odtworzyć macierze ortogonalne?
Należy je spamiętywać. Mamy $J = PAQ$ i $\Sigma = SJT$, czyli
 $A = P^T J Q^T$ i $J = S^T \Sigma T^T$.

Zatem

$$A = P^T (S^T \Sigma T^T) Q^T = (P^T S^T) \Sigma (T^T Q^T)$$

Skoro iloczyn macierzy ortogonalnych jest macierzą ortogonalną, to

$$\begin{aligned} U &\leftarrow P^T S^T \\ V^T &\leftarrow T^T Q^T \end{aligned}$$

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Kompresja danych

Mamy $A = U\Sigma V^T$ (SVD). Załóżmy, że

$$\Sigma = \begin{pmatrix} \sigma_1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \sigma_r & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

Kompresja danych

Zapiszmy macierze U i V jako $U = (u_1 \ \dots \ u_m)$ i
 $V = (v_1 \ \dots \ v_n)$.

Kompresja danych

Zapiszmy macierze U i V jako $U = (u_1 \ \dots \ u_m)$ i
 $V = (v_1 \ \dots \ v_n)$.

Wtedy

$$\begin{aligned} A &= U \Sigma V^T = (u_1 \ \dots \ u_m) \Sigma \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} = \\ &= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T \end{aligned}$$

Skracanie sumy

- wartości na przekątnej macierzy Σ są uporządkowane nierosnąco
- wektory u_i oraz v_i są długości 1

$$A = U\Sigma V^T = (u_1 \quad \dots \quad u_m) \Sigma \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} =$$

$$= \sigma_1 u_1 v_1^T + \dots + \sigma_r u_r v_r^T \approx \sigma_1 u_1 v_1^T + \dots + \sigma_l u_l v_l^T$$

dla $l \leq r$.

Oszczędność pamięciowa

WAŻNE

Dane często są przechowywane w postaci bardzo dużych prostokątnych macierzy, więc przedstawienie ich w postaci zredukowanego rozkładu SVD może przynieść dużą oszczędność pamięciową.

Spis treści

- 1 Podstawowe pojęcia
- 2 Rozkład SVD
 - Twierdzenie
 - Dowód
- 3 Numeryczne wyznaczanie
- 4 Zastosowania
 - Kompresja danych
 - Inne

Pseudoodwrotność

Pseudoodwrotność macierzy $A \in M_{m \times n}(\mathbb{R})$ rozumiemy jako macierz $A^\dagger \in M_{n \times m}(\mathbb{R})$ taką, że

$$AA^\dagger = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = \bar{I}$$

Pseudoodwrotność

SVD może zostać użyte do policzenia pseudoodwrotności macierzy. Niech $A = U\Sigma V^T$. Wtedy $A^\dagger = V\Sigma^\dagger U^T$, gdzie Σ^\dagger rozumiemy jako

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

$$AA^\dagger = U\Sigma V^T V\Sigma^\dagger U^T = U\Sigma\Sigma^\dagger U^T = U\bar{I}U^T = \bar{I}$$

Rozwiązywanie równań jednorodnych

Fakt

$$(\forall x \in \mathbb{R}^n)(Ax = 0 \Leftrightarrow A^T Ax = 0)$$

Zatem wystarczy rozwiązać układ $A^T Ax = 0$. Po zdiagonalizowaniu macierzy symetrycznej (tw. spektralne) $A^T A = VDV^T$ możemy podać rozwiązanie jako przestrzeń rozpinaną przez kolumny V odpowiadające zerowym wartościom własnym.

Obraz, jądro i rząd

Fakt

Dla macierzy $A = U\Sigma V^T$:

- kolumny macierzy U odpowiadające niezerowym wartościom szczególnym rozpinają obraz przekształcenia liniowego zadanego macierzą A
- kolumny macierzy V odpowiadające zerowym wartościom własnym $A^T A$ rozpinają jądro przekształcenia liniowego zadanego macierzą A
- $\text{rank}(A) = \#$ wartości szczególnych

Źródła cz. I



G. H. Golub and C. Reinsch *Handbook Series Linear Algebra. Singular Value Decomposition and Least Squares Solutions*. Numer. Math. 14, 403–420 (1970)



G. Golub and W. Kahan, *Calculating the Singular Values and Pseudo-inverse of a Matrix*. J. Siam Numer. Anal., Ser. B, Vol. 2, No. 2. Printed in U.S.A.. 1965



Singular value decomposition. Wikipedia: The Free Encyklopedia (16.05.2020)
https://en.wikipedia.org/wiki/Singular_value_decomposition



Understanding Singular Value Decomposition and its Application in Data Science. (16.05.2020)
<https://towardsdatascience.com/understanding-singular-value-decomposition-and-its-application-in-data-science-388a54be95d>

Źródła cz. II



Singular Value Decomposition as Simply as Possible (16.05.2020)

<https://gregorygundersen.com/blog/2018/12/10/svd/>



A. Kiełbasiński and Hubert Schwetlick, *Numeryczna algebra liniowa*.

Wydawnictwo Naukowo-Techniczne



Åke Björck and Germund Dahlquist, *Numerical Methods in Scientific Computing. Volume II*.



MATH2071: LAB #9: The Singular Value Decomposition. (16.05.2020)

<http://www.math.pitt.edu/~sussmanm/2071Spring08/lab09/index.html>



Wikiwand. Singular value decomposition. (16.05.2020)

https://www.wikiwand.com/en/Singular_value_decomposition?fbclid=IwAR0_B1ohQBsQnXalAqx869TnLd2jWU2Mc2TQTwnkXkV_pMLyT3H189kIbz4#