Zad. 10. Wyznaczanie rozkładu SVD macierzy prostokątnej

e-mail: andrzej.kedziorski@fizyka.umk.pl
tel.: 56611-3274
pokój: 485B
http://www.fizyka.umk.pl/~tecumseh/EDU/MNII/

Zadanie 10

Napisz program wyznaczający rozkład SVD macierzy prostokątnej $\mathbf{A}$. Na podstawie rozkładu SVD wyznaczyć macierz pseudoodwrotną do macierzy $\mathbf{A}$.

Rozkład wg wartości szczególnych/osobliwych (SVD)

SVD (*Singular Value Decomposition*) to metoda pozwalająca na wyodrębnienie liniowo niezależnych kolumn z macierzy prostokątnej $m \times n$.

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

- $\boldsymbol{U}$ i $\boldsymbol{V}$ są macierzami ortogonalnymi o wymiarach $m \times m$ i $n \times n$
- $\boldsymbol{\Sigma}$ jest macierzą diagonalną $m \times n$ z nieujemnymi elementami rzeczywistymi σ_i[15] uporządkowanymi malejąco, tzn.

$$\sigma_1 \geqslant \sigma_2 \ldots \geqslant \sigma_{\min(m,n)} \geqslant 0$$

σ_i to wartości szczególne $\boldsymbol{A}$, a $\min(m,n)$ kolumn macierzy $\boldsymbol{U}$ i $\boldsymbol{V}$ są lewymi i prawymi wektorami szczególnymi $\boldsymbol{A}$. Mamy

$$\boldsymbol{A}\boldsymbol{v}_j = \sigma_j \boldsymbol{u}_j, \quad \boldsymbol{A}^T\boldsymbol{u}_j = \sigma_j \boldsymbol{v}_j, \quad \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{v}_j = \sigma_j^2 \boldsymbol{v}_j, \quad \boldsymbol{A}\boldsymbol{A}^T\boldsymbol{u}_j = \sigma_j^2 \boldsymbol{u}_j$$

[15] Tutaj σ_i nie mają nic wspólnego z odchyleniem standardowym danych doświadczalnych!

Prostokątna macierz diagonalna - przykład 4×3

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{pmatrix} \quad \mathbf{\Sigma}^{Tr} = \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \end{pmatrix}$$

$$\mathbf{\Sigma}^{Tr}\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$$

$$\mathbf{\Sigma}\mathbf{\Sigma}^{Tr} = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Rozkład SVD - przekształcenia (1)

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n}\,\mathbf{V}^{Tr}_{n\times n} \quad \Big| \cdot \mathbf{V}_{n\times n} \qquad (\mathbf{V}^{Tr}\mathbf{V} = \mathbf{1})$$

$$\mathbf{A}_{m\times n}\,\mathbf{V}_{n\times n} = \mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n}$$

$$\sum_{k=1}^{n} A_{ik} V_{kj} = \underbrace{\sum_{\ell=1}^{\min(m,n)} U_{i\ell}\sigma_{\ell}\delta_{\ell j}}_{U_{ij}\sigma_j} \qquad i = 1,\ldots,m,\ j = 1,\ldots,n$$

$$\Downarrow$$

$$\mathbf{A}\mathbf{v}_j = \sigma_j \mathbf{u}_j, \quad j = 1,\ldots,\min(m,n)$$

$\mathbf{v}_j$ oraz $\mathbf{u}_j$ to kolumny macierzy $\mathbf{V}$ i $\mathbf{U}$

Rozkład SVD - przekształcenia (2)

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n}\,\mathbf{V}^{Tr}_{n\times n}$$

$$\mathbf{A}^{Tr}_{n\times m} = (\mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n}\,\mathbf{V}^{Tr}_{n\times n})^{Tr} = (\mathbf{V}^{Tr}_{n\times n})^{Tr}(\mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n})^{Tr}$$

$$= \mathbf{V}_{n\times n}\,\mathbf{\Sigma}^{Tr}_{n\times m}\mathbf{U}^{Tr}_{m\times m}$$

$$\mathbf{A}^{Tr}_{n\times m} = \mathbf{V}_{n\times n}\,\mathbf{\Sigma}^{Tr}_{n\times m}\mathbf{U}^{Tr}_{m\times m} \qquad |\cdot\mathbf{U}_{m\times m} \qquad (\mathbf{U}^{Tr}\mathbf{U} = \mathbf{1})$$

$$\mathbf{A}^{Tr}_{n\times m}\,\mathbf{U}_{m\times m} = \mathbf{V}_{n\times n}\,\mathbf{\Sigma}^{Tr}_{n\times m}$$

$$\sum_{k=1}^{m}(\mathbf{A}^{Tr})_{ik}U_{kj} = \underbrace{\sum_{\ell=1}^{\min(m,n)} V_{i\ell}\sigma_\ell\delta_{\ell j}}_{V_{ij}\sigma_j} \qquad i = 1,\ldots,n,\ j = 1,\ldots,m$$

$$\Downarrow$$

$$\mathbf{A}^{Tr}\mathbf{u}_j = \sigma_j\mathbf{v}_j, \quad j = 1,\ldots,\min(m,n)$$

$\mathbf{v}_j$ oraz $\mathbf{u}_j$ to kolumny macierzy $\mathbf{V}$ i $\mathbf{U}$

Rozkład SVD - przekształcenia (3)

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\,\boldsymbol{\Sigma}_{m\times n}\,\mathbf{V}^{Tr}_{n\times n} \qquad \mathbf{A}^{Tr}_{n\times m} = \mathbf{V}_{n\times n}\,\boldsymbol{\Sigma}^{Tr}_{n\times m}\mathbf{U}^{Tr}_{m\times m}$$

$$\mathbf{A}^{Tr}\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^{Tr}\underbrace{\mathbf{U}^{Tr}\,\mathbf{U}}_{\mathbf{1}}\boldsymbol{\Sigma}\mathbf{V}^{Tr} = \mathbf{V}\boldsymbol{\Sigma}^{Tr}\boldsymbol{\Sigma}\mathbf{V}^{Tr}$$

$$\mathbf{A}^{Tr}\mathbf{A} = \mathbf{V}\boldsymbol{\Sigma}^{Tr}\boldsymbol{\Sigma}\mathbf{V}^{Tr} \quad \big| \cdot \mathbf{V}$$

$$(\mathbf{A}^{Tr}\mathbf{A})_{n\times n}\mathbf{V}_{n\times n} = \mathbf{V}_{n\times n}(\boldsymbol{\Sigma}^{Tr}\boldsymbol{\Sigma})_{n\times n}$$

$$\sum_{k=1}^{n}(\mathbf{A}^{Tr}\mathbf{A})_{ik}V_{kj} = \underbrace{\sum_{\ell=1}^{\min(m,n)} V_{i\ell}\sigma_\ell^2\delta_{\ell j}}_{V_{ij}\sigma_j^2}$$

$$\Downarrow$$

$$\mathbf{A}^{Tr}\mathbf{A}\,\mathbf{v}_j = \sigma_j^2\,\mathbf{v}_j, \qquad j = 1, \ldots, \min(m, n)$$

Kolumna $\mathbf{v}_j$ jest wektorem własnym symetrycznej macierzy $\mathbf{A}^{Tr}\mathbf{A}$ przy wartości własnej σ_j^2

Rozkład SVD - przekształcenia (4)

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n}\,\mathbf{V}^{Tr}_{n\times n} \qquad \mathbf{A}^{Tr}_{n\times m} = \mathbf{V}_{n\times n}\,\mathbf{\Sigma}^{Tr}_{n\times m}\mathbf{U}^{Tr}_{m\times m}$$

$$\mathbf{A}\mathbf{A}^{Tr} = \mathbf{U}\mathbf{\Sigma}\underbrace{\mathbf{V}^{Tr}\mathbf{V}}_{\mathbf{1}}\mathbf{\Sigma}^{Tr}\mathbf{U}^{Tr} = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^{Tr}\mathbf{U}^{Tr}$$

$$\mathbf{A}\mathbf{A}^{Tr} = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^{Tr}\mathbf{U}^{Tr} \quad | \cdot \mathbf{U}$$

$$(\mathbf{A}\mathbf{A}^{Tr})_{m\times m}\mathbf{U}_{m\times m} = \mathbf{U}_{m\times m}(\mathbf{\Sigma}\mathbf{\Sigma}^{Tr})_{m\times m}$$

$$\sum_{k=1}^{m}(\mathbf{A}\mathbf{A}^{Tr})_{ik}U_{kj} = \underbrace{\sum_{\ell=1}^{\min(m,n)} U_{i\ell}\sigma_\ell^2\delta_{\ell j}}_{U_{ij}\sigma_j^2}$$

$$\Downarrow$$

$$\mathbf{A}\mathbf{A}^{Tr}\,\mathbf{u}_j = \sigma_j^2\,\mathbf{u}_j, \qquad j = 1, \ldots, \min(m,n)$$

Kolumna $\mathbf{u}_j$ jest wektorem własnym symetrycznej macierzy $\mathbf{A}\mathbf{A}^{Tr}$ przy wartości własnej σ_j^2

Zastosowania SVD

- Norma euklidesowa
$$\|\boldsymbol{A}\|_2 = \max_{x \neq 0} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sigma_{\max}$$
- Rząd macierzy. W teorii rząd macierzy jest określony przez liczbę niezerowych wartości szczególnych. W praktyce przez liczbę wartości szczególnych, które są większe od zadanego progu (*rząd numeryczny* macierzy).

Zastosowania SVD

- Wskaźnik uwarunkowania macierzy (prostoktnych i kwadratowych)

$$\mathrm{cond}(\boldsymbol{A}) = \sigma_{\mathrm{max}}/\sigma_{\mathrm{min}}$$

  Dla macierzy kwadratowych ten wskaźnik jest miarą osobliwości, a dla macierzy prostokątnych – liczby liniowo zależnych kolumn.

- Rozwiązywanie równań liniowych dla zagadnień najmniejszych kwadratów. Rozwiązanie minimalizujące normę euklidesową dla równania $\boldsymbol{A}\boldsymbol{x} \approx \boldsymbol{b}$ wynosi

$$\boldsymbol{x} = \sum_{\sigma_i \neq 0} \frac{\boldsymbol{u}_i^T \boldsymbol{b}}{\sigma_i} \boldsymbol{v}_i$$

  SVD jest szczególnie przydatna przy rozwiązywaniu zagadnień źle uwarunkowanych lub o macierzach o obniżonym rzędzie, gdyż w sumie można opuścić wyrazy ze zbyt małymi wartościami szczególnymi (rozwiązanie jest mniej czułe na zaburzenie danych).

Zastosowania SVD

- Pseudoodwrotność macierzy. Pseudoodwrotnością skalarnej wartości σ jest $1/\sigma$, jeśli $\sigma \neq 0$ lub – w przeciwnym przypadku – zero. To pozwala zdefiniować pseudoodwrotność macierzy diagonalnej Σ^+, która na diagonalnej ma pseudoodwrotności wartości szczególnych. Wówczas

$$\boldsymbol{A}^+ = \boldsymbol{V}\boldsymbol{\Sigma}^+\boldsymbol{U}^T$$

  $\boldsymbol{A}^+$ istnieje zawsze niezależnie od tego, czy macierz jest kwadratowa, czy pełnego rzędu. Jeśli macierz jest kwadratowa i nieosobliwa, to jest ona równoważna macierzy odwrotnej $\boldsymbol{A}^{-1}$.

  Rozwiązanie w sensie minimalizowania normy euklidesowej zagadnienia $\boldsymbol{A}\boldsymbol{x} \approx \boldsymbol{b}$ jest równe $\boldsymbol{A}^+\boldsymbol{b}$.

Pseudoodwrotność $\mathbf{\Sigma}$ - przykład 4×3

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \\ 0 & 0 & 0 \end{pmatrix} \qquad \Rightarrow \qquad \mathbf{\Sigma}^+ = \begin{pmatrix} \frac{1}{\sigma_1} & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_3} & 0 \end{pmatrix}$$

gdzie założono, że $\sigma_i \neq 0, \ i = 1, 2, 3$

$$\mathbf{\Sigma}\mathbf{\Sigma}^+ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathbf{\Sigma}^+\mathbf{\Sigma} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Pseudodwrotność $\mathbf{A}_{m\times n}$, gdzie $m \geqslant n$ oraz $\sigma_{\min} \neq 0$

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\,\mathbf{\Sigma}_{m\times n}\,\mathbf{V}^{Tr}_{n\times n} \qquad \mathbf{A}^{+}_{n\times m} = \mathbf{V}_{n\times n}\,\mathbf{\Sigma}^{+}_{n\times m}\,\mathbf{U}^{Tr}_{m\times m}$$

$$\mathbf{A}^{+}\mathbf{A} = \mathbf{V}\,\mathbf{\Sigma}^{+}\,\underbrace{\mathbf{U}^{Tr}\,\mathbf{U}}_{\mathbf{1}}\,\mathbf{\Sigma}\,\mathbf{V}^{Tr} = \mathbf{V}\,\underbrace{\mathbf{\Sigma}^{+}\,\mathbf{\Sigma}}_{\mathbf{1}}\,\mathbf{V}^{Tr} = \mathbf{V}\mathbf{V}^{Tr} = \mathbf{1}_{n\times n}$$

$$\mathbf{A}\mathbf{A}^{+} = \mathbf{U}\,\mathbf{\Sigma}\,\underbrace{\mathbf{V}^{Tr}\,\mathbf{V}}_{\mathbf{1}}\,\mathbf{\Sigma}^{+}\,\mathbf{U}^{Tr} = \mathbf{U}\,\underbrace{\mathbf{\Sigma}\,\mathbf{\Sigma}^{+}}_{\tilde{\mathbf{1}}}\,\mathbf{U}^{Tr} = \mathbf{U}\tilde{\mathbf{1}}\mathbf{U}^{Tr} = \tilde{\mathbf{1}}_{m\times m}$$

$$\tilde{\mathbf{1}} = \begin{pmatrix} \mathbf{1}_{n\times n} & \mathbf{0}_{n\times(m-n)} \\ \mathbf{0}_{(m-n)\times n} & \mathbf{0}_{(m-n)\times(m-n)} \end{pmatrix}_{m\times m}$$

Zastosowania SVD

- Bazy ortonormalne. Kolumny $\boldsymbol{V}$ odpowiadające zerowym wartościom σ tworzą bazę przestrzeni zerowej $\boldsymbol{A}$ (jądro przekształcenia). Pozostałe wektory $\boldsymbol{V}$ tworzą bazę ortogonalnego uzupełnienia przestrzeni zerowej. Kolumny $\boldsymbol{U}$ odpowiadające niezerowym wartościom szczególnym tworzą bazę przestrzeni będącej obrazem przekształcenia określonego przez macierz $\boldsymbol{A}$. Pozostałe kolumny stanowią bazę ortogonalnego uzupełnienia tej przestrzeni.

Zastosowania SVD

- Przybliżanie macierzy przez macierz o mniejszym rzędzie:

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^T = \sigma_1 \boldsymbol{E}_1 + \sigma_2 \boldsymbol{E}_2 + \ldots + \sigma_n \boldsymbol{E}_n$$

  Każda macierz $\boldsymbol{E}_i = \underline{u}_i \boldsymbol{v}_i^T$ jest rzędu 1.

  Zwarte przybliżenie macierzy $\boldsymbol{A}$ – suma z pominięciem wyrazów odpowiadających małym wartościom σ_i.

  Jeśli $\boldsymbol{A}$ jest przybliżona przez sumę zbudowaną przy pomocy k największych wartości szczególnych, to otrzymujemy przybliżenie $\boldsymbol{A}$ najlepsze w sensie normy Frobeniusa (norma euklidesowa macierzy traktowanej jako wektor w przestrzeni $\mathcal{R}^{mn}$).

  Takie przybliżenie jest przydatne przy przetwarzaniu obrazów, kompresji danych, kryptografii, itp.

„Naturalna” metoda rozkładu SVD macierzy $\mathbf{A}_{m\times n}$

„Naturalna” metoda: rozwiązać zagadnienie własne dla mniejszej macierzy symetrycznej

1. $m \geqslant n: \mathbf{A}^{Tr}\mathbf{A}\,\mathbf{v}_j = \sigma_j^2\,\mathbf{v}_j, \quad j = 1,\dots,n$
2. $m < n: \ \mathbf{A}\mathbf{A}^{Tr}\,\mathbf{u}_k = \sigma_k^2\,\mathbf{u}_k, \quad k = 1,\dots,m$

Problemy

- Najmniejsze wartości własne, a w konsekwencji σ_i, są narażone na **duże błędy numeryczne**
- Metoda QR ma problem ze zbieżnością, gdy $\sigma_i = \sigma_{i+1}$
- W metodzie QR wykorzystującej ortogonalizację Grama-Schmidta problem, gdy $\sigma_i \approx 0 \Rightarrow$ rozwiązać zagadnienie własne dla $\mathbf{A}^{Tr}\mathbf{A} - s\mathbf{1}$ (albo $\mathbf{A}\mathbf{A}^{Tr} - s\mathbf{1}$), gdzie np. $s = 1$; lepiej zastosować inną metodę rozkładu QR np. używając transformacji Householdera
- Ortogonalizacja Grama-Schmidta (GS) ma swoje dodatkowe problemy... zob. zmodyfikowana metoda GS (MGS)

Metoda rozkładu SVD macierzy $\mathbf{A}_{m\times n}$, zał. $m \geqslant n$

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\mathbf{\Sigma}_{m\times n}\mathbf{V}^{Tr}_{n\times n}$$

1. Dokonać rozkładu $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, gdzie $\mathbf{P}$ i $\mathbf{Q}$ to macierze ortogonalne, a $\mathbf{J}$ jest macierzą dwudiagonalną
2. Dokonać rozkładu SVD macierzy $\mathbf{J}_{m\times n} = \mathbf{X}_{m\times m}\mathbf{\Sigma}_{m\times n}\mathbf{Y}^{Tr}_{n\times n}$, gdzie $\mathbf{X}$ i $\mathbf{Y}$ to macierze ortogonalne, a macierz $\mathbf{J}$ ma takie same wartości szczególne co macierz $\mathbf{A}$
3. Obliczyć macierze ortogonalne $\mathbf{U}$ i $\mathbf{V}$

$$\begin{aligned}\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{Q}^{Tr} &= \mathbf{P}(\mathbf{X}\mathbf{\Sigma}\mathbf{Y}^{Tr})\mathbf{Q}^{Tr}\\ &= (\mathbf{P}\mathbf{X})\mathbf{\Sigma}(\mathbf{Q}\mathbf{Y})^{Tr},\end{aligned}$$

więc

$$\mathbf{U} = \mathbf{P}\mathbf{X} \qquad \mathbf{V} = \mathbf{Q}\mathbf{Y}$$

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

Macierz dwudiagonalna

$$
J = \left(\begin{array}{ccccccc}
\alpha_1 & \beta_1 & 0 & \cdot & \cdot & \cdot & 0 \\
 & \alpha_2 & \beta_2 & 0 & \cdot & \cdot & \cdot \\
 & \mathbf{0} & \cdot & \cdot & \cdot & & \\
 & & & \cdot & \cdot & \cdot & \\
 & & & & \cdot & \cdot & \\
 & & & & & \cdot & \beta_{n-1} \\
 & & & & & & \alpha_n \\
 & & & \mathbf{0} & & &
\end{array}\right) \begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \left.\vphantom{\mathbf{0}}\right\} (m-n)\times n \end{array} .
$$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

Procedura korzystająca z transformacji Hauseholdera

- Tworzymy kolejne macierze $\mathbf{A}^{(1)}$, $\mathbf{A}^{(3/2)}$, ... $\mathbf{A}^{(k)}$, $\mathbf{A}^{(k+1/2)}$, ... $\mathbf{A}^{(n)}$, $\mathbf{A}^{(n+1/2)}$

$$\begin{aligned}
\mathbf{A}^{(1)} &= \mathbf{A} \\
\mathbf{A}^{(k+1/2)} &= \mathbf{P}^{(k)}\mathbf{A}^{(k)}, \quad k = 1, 2, \ldots, n \\
\mathbf{A}^{(k+1)} &= \mathbf{A}^{(k+1/2)}\mathbf{Q}^{(k)}, \quad k = 1, 2, \ldots, n-1, \ \text{właściwie do } n-2.
\end{aligned}$$

  Zauważ, że $\mathbf{J} = \mathbf{A}^{(n+1/2)}$, ale $\mathbf{J}$ można skonstruować osobno...

- Symetryczne i ortogonalne macierze Hauseholdera $\mathbf{P}^{(k)}$ i $\mathbf{Q}^{(k)}$

$$\begin{aligned}
\mathbf{P}^{(k)} &= \mathbf{1} - 2\mathbf{x}^{(k)}(\mathbf{x}^{(k)})^{Tr}, \quad (\mathbf{x}^{(k)})^{Tr}\mathbf{x}^{(k)} = 1 \\
\mathbf{Q}^{(k)} &= \mathbf{1} - 2\mathbf{y}^{(k)}(\mathbf{y}^{(k)})^{Tr}, \quad (\mathbf{y}^{(k)})^{Tr}\mathbf{y}^{(k)} = 1
\end{aligned}$$

  wyznaczamy z warunków

$$\begin{aligned}
\mathbf{P}^{(k)} &: \quad A^{(k+1/2)}_{ik} = 0, \quad i = k+1, \ldots, m \\
\mathbf{Q}^{(k)} &: \quad A^{(k+1)}_{kj} = 0, \quad j = k+2, \ldots, n
\end{aligned}$$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

Macierz prawie dwudiagonalna

$$
A^{(k+1)} = \begin{pmatrix}
\alpha_1 & \beta_1 & 0 & \cdot & \cdot & & & & & \\
0 & \alpha_2 & \beta_2 & 0 & \cdot & & & & & \\
\cdot & 0 & \cdot & \cdot & \cdot & & & \text{\Huge 0} & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & & & \\
 & & & & \alpha_k & \beta_k & & & & \\
 & & & & & x & x & \cdot & \cdot & \cdot \\
 & & \text{\Huge 0} & & & x & x & \cdot & \cdot & \cdot \\
 & & & & & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & & & & & \cdot & \cdot & \cdot & \cdot & \cdot
\end{pmatrix}
$$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

Wyznaczenie wektora jednostkowego $\mathbf{x}^{(k)}$ i elementu diagonalnego α_k

$$s_k = \sqrt{\sum_{i=k}^{m} |A^{(k)}_{ik}|^2}$$

$$\alpha_k = -s_k \operatorname{sign}(A^{(k)}_{kk})$$

$$x^{(k)}_i = 0, \quad i < k$$

$$x^{(k)}_k = \sqrt{\frac{1}{2}\left(1 + \frac{|A^{(k)}_{kk}|}{s_k}\right)}$$

$$x^{(k)}_i = c_k\, A^{(k)}_{i,k}, \quad i = k+1, \ldots, m$$

gdzie

$$c_k = \frac{\operatorname{sign}(A^{(k)}_{kk})}{2 s_k\, x^{(k)}_k}$$

Jeżeli $s_k = 0$, wtedy $\alpha_k = 0$ i $\mathbf{x}^{(k)} = \mathbf{0}$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

Wyznaczenie wektora jednostkowego $\mathbf{y}^{(k)}$ i elementu pozadiagonalnego β_k

$$t_k = \sqrt{\sum_{j=k+1}^{n} |A_{kj}^{(k+1/2)}|^2}$$

$$\beta_k = -t_k \,\mathrm{sign}(A_{k,k+1}^{(k+1/2)})$$

$$y_j^{(k)} = 0, \quad j \leqslant k$$

$$y_{k+1}^{(k)} = \sqrt{\frac{1}{2}\left(1 + \frac{|A_{k,k+1}^{(k+1/2)}|}{t_k}\right)}$$

$$y_j^{(k)} = d_k\, A_{k,j}^{(k+1/2)}, \quad j = k+2, \ldots, n$$

gdzie

$$d_k = \frac{\mathrm{sign}(A_{k,k+1}^{(k+1/2)})}{2 t_k\, y_{k+1}^{(k)}}$$

Jeżeli $t_k = 0$, wtedy $\beta_k = 0$ i $\mathbf{y}^{(k)} = \mathbf{0}$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

$$
\begin{aligned}
\mathbf{A}^{(1)} &= \mathbf{A}\\
\mathbf{A}^{(k+1/2)} &= \mathbf{P}^{(k)}\mathbf{A}^{(k)}, \quad k = 1, 2, \dots, n\\
\mathbf{A}^{(k+1)} &= \mathbf{A}^{(k+1/2)}\mathbf{Q}^{(k)}, \quad k = 1, 2, \dots, n-1.\\
\mathbf{A}^{(n+1/2)} &= \mathbf{J}
\end{aligned}
$$

$$\Downarrow$$

$$
\begin{aligned}
\mathbf{A}^{(3/2)} &= \mathbf{P}^{(1)}\mathbf{A}^{(1)}\\
\mathbf{A}^{(2)} &= \mathbf{P}^{(1)}\mathbf{A}^{(1)}\mathbf{Q}^{(1)}\\
\mathbf{A}^{(5/2)} &= \mathbf{P}^{(2)}\mathbf{P}^{(1)}\mathbf{A}^{(1)}\mathbf{Q}^{(1)}\\
\mathbf{A}^{(3)} &= \mathbf{P}^{(2)}\mathbf{P}^{(1)}\mathbf{A}^{(1)}\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\\
&\vdots\\
\mathbf{A}^{(n+1/2)} &= \mathbf{P}^{(n)}\dots\mathbf{P}^{(2)}\mathbf{P}^{(1)}\mathbf{A}^{(1)}\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\dots\mathbf{Q}^{(n-1)}
\end{aligned}
$$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

$$
\begin{aligned}
\mathbf{A}^{(1)} &= \mathbf{A}\\
\mathbf{A}^{(k+1/2)} &= \mathbf{P}^{(k)}\mathbf{A}^{(k)}, \quad k = 1, 2, \dots, n\\
\mathbf{A}^{(k+1)} &= \mathbf{A}^{(k+1/2)}\mathbf{Q}^{(k)}, \quad k = 1, 2, \dots, n-1.\\
\mathbf{A}^{(n+1/2)} &= \mathbf{J}
\end{aligned}
$$

$$\Downarrow$$

$$
\begin{aligned}
&\left(\mathbf{P}^{(n)}\right)^{Tr} \mid \quad \mathbf{J} = \mathbf{P}^{(n)}\dots\mathbf{P}^{(2)}\mathbf{P}^{(1)}\mathbf{A}\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\dots\mathbf{Q}^{(n-1)} \quad \mid \left(\mathbf{Q}^{(n-1)}\right)^{Tr}\\
&\left(\mathbf{P}^{(n-1)}\right)^{Tr} \mid \quad \mathbf{P}^{(n)}\mathbf{J}\mathbf{Q}^{(n-1)} = \mathbf{P}^{(n-1)}\dots\mathbf{P}^{(2)}\mathbf{P}^{(1)}\mathbf{A}\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\dots\mathbf{Q}^{(n-2)} \quad \mid \left(\mathbf{Q}^{(n-2)}\right)^{Tr}\\
&\mathbf{P}^{(n-1)}\mathbf{P}^{(n)}\mathbf{J}\mathbf{Q}^{(n-1)}\mathbf{Q}^{(n-2)} = \mathbf{P}^{(n-2)}\dots\mathbf{P}^{(2)}\mathbf{P}^{(1)}\mathbf{A}\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}\dots\mathbf{Q}^{(n-3)}\\
&\qquad\vdots\\
&\mathbf{P}^{(2)}\dots\mathbf{P}^{(n-1)}\mathbf{P}^{(n)}\mathbf{J}\mathbf{Q}^{(n-1)}\mathbf{Q}^{(n-2)}\dots\mathbf{Q}^{(2)} = \mathbf{P}^{(1)}\mathbf{A}\mathbf{Q}^{(1)}\\
&\underbrace{\mathbf{P}^{(1)}\mathbf{P}^{(2)}\dots\mathbf{P}^{(n-1)}\mathbf{P}^{(n)}}_{\mathbf{P}}\,\mathbf{J}\,\underbrace{\mathbf{Q}^{(n-1)}\mathbf{Q}^{(n-2)}\dots\mathbf{Q}^{(2)}\mathbf{Q}^{(1)}}_{\mathbf{Q}^{Tr}} = \mathbf{A}
\end{aligned}
$$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Rozkład $\mathbf{A}_{m\times n} = \mathbf{P}_{m\times m}\mathbf{J}_{m\times n}\mathbf{Q}^{Tr}_{n\times n}$, zał. $m \geqslant n$

$$\mathbf{A}^{(1)} = \mathbf{A}$$
$$\mathbf{A}^{(k+1/2)} = \mathbf{P}^{(k)}\mathbf{A}^{(k)}, \quad k = 1, 2, \ldots, n$$
$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k+1/2)}\mathbf{Q}^{(k)}, \quad k = 1, 2, \ldots, n-1.$$
$$\mathbf{A}^{(n+1/2)} = \mathbf{J}$$

$$\Downarrow$$

Macierze ortogonalne

$$\mathbf{P} = \mathbf{P}^{(1)}\,\mathbf{P}^{(2)}\ldots\mathbf{P}^{(n)}$$

$$\mathbf{Q}^{Tr} = \mathbf{Q}^{(n-1)}\ldots\mathbf{Q}^{(2)}\,\mathbf{Q}^{(1)}$$

Przy okazji wyznaczania wektorów $\mathbf{x}^{(k)}$ i $\mathbf{y}^{(k)}$ obliczyliśmy elementy diagonalne α_k i pozadiagonalne β_k macierz dwudiagonalnej $\mathbf{J}$

G. Golub, W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, J. SIAM Numer. Anal. B 2, 205-224 (1965).

Dygresja: Równość wartości szczególnych macierzy $\mathbf{J}$ i $\mathbf{A}$

Mamy następujące rozkłady macierzy

$$\mathbf{A} = \mathbf{U\Sigma V}^{Tr}, \qquad \mathbf{A} = \mathbf{PJQ}^{Tr}, \qquad \mathbf{J} = \mathbf{X\Sigma' Y}^{Tr}.$$

Pokażemy, że $\mathbf{\Sigma} = \mathbf{\Sigma}'$:

$$\begin{aligned} \mathbf{P}^{Tr} \times \mid \quad \mathbf{PJQ}^{Tr} &= \mathbf{U\Sigma V}^{Tr} \quad \mid \times \mathbf{Q} \\ \mathbf{J} &= \mathbf{P}^{Tr}\mathbf{U\Sigma V}^{Tr}\mathbf{Q} \end{aligned}$$

Ponieważ $\mathbf{P}^{Tr}\mathbf{U}$ i $\mathbf{V}^{Tr}\mathbf{Q}$ są macierzami ortogonalnymi, to uzyskaliśmy rozkład SVD macierzy $\mathbf{J}$, więc z jednoznaczności[1] SVD mamy w szczególności $\mathbf{\Sigma}' = \mathbf{\Sigma}$.

[1] SVD nie jest w pełni jednoznaczny, ale dla danej macierzy mamy na pewno jeden zbiór wartości szczególnych.

Rozkład SVD macierzy dwudiagonalnej $\mathbf{J}_{m\times n} = \mathbf{X}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\mathbf{Y}^{Tr}_{n\times n}$

1. Znaleźć wartości własne λ_i symetrycznej macierzy trójdiagonalnej $\mathbf{K}_{n\times n} = \mathbf{J}^{Tr}\mathbf{J}$, gdzie $\lambda_{n-i+1} = \sigma_i^2$ oraz $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_n$
2. Znaleźć wektory własne $\mathbf{y}_i$ macierzy $\mathbf{K}$, tzn. takie, że $\mathbf{K}\mathbf{y}_i = \sigma_i^2\mathbf{y}_i$; $\mathbf{y}_i$ jest i-tą kolumną ortogonalnej macierzy $\mathbf{Y}$
   (zobacz zad. 9)
   Uwaga: Każdy wektor własny $\mathbf{y}_i$ musi być unormowany w sensie normy Euklidesowej, tzn.
   $$\|\mathbf{y}\|_2 = \sqrt{\sum_{j=1}^{n} |y_j|^2} = 1$$
   Normowanie wektora: $\mathbf{y}_i \rightarrow \frac{\mathbf{y}_i}{\|\mathbf{y}_i\|_2}$
3. Znaleźć kolumny $\mathbf{x}_j$ macierzy $\mathbf{X}$
   $$\mathbf{x}_j = \frac{1}{\sigma_j}\mathbf{J}\,\mathbf{y}_j, \quad \sigma_j > 0$$

Na razie problemy w punktach 2. i 3., gdy $\sigma_j \sim 0$...

Wartości własne symetrycznej i trójdiagonalnej macierzy **K**

Macierz symetryczna trójdiagonalna

$$\mathbf{K}_{n\times n} = \begin{pmatrix} c_1 & b_2 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ b_2 & c_2 & b_3 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & b_3 & c_3 & b_4 & 0 & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \dots & 0 & b_k & c_k & b_{k+1} & 0 & \dots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & b_{n-2} & c_{n-2} & b_{n-1} & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & b_{n-1} & c_{n-1} & b_n \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & b_n & c_n \end{pmatrix}$$

W. Barth, R. S. Martin and J. H. Wilkinson, *Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection*, Numerische Mathematik 9, 386-393 (1967) .

Wartości własne symetrycznej i trójdiagonalnej macierzy $\mathbf{K}$

Oznaczenia

- $\mathbf{c} = [c_1, c_2, \ldots, c_n]^{Tr}$ - wektor reprezentujący główną diagonalę macierzy $\mathbf{K}_{n\times n}$
- $\mathbf{b} = [0, b_2, \ldots, b_n]^{Tr}$ - wektor reprezentujący elementy pozadiagonalne macierzy $\mathbf{K}_{n\times n}$

W. Barth, R. S. Martin and J. H. Wilkinson, *Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection*, Numerische Mathematematik 9, 386-393 (1967) .

Wartości własne symetrycznej i trójdiagonalnej macierzy **K**

Twierdzenie:
Wartości własne λ_i macierzy **K** znajdują się w przedziale $[\lambda_{min}, \lambda_{max}]$, gdzie $(b_1 = b_{n+1} = 0)$

$$\lambda_{min} = \min_{i=1,\dots,n} \left[c_i - (|b_i| + |b_{i+1}|)\right]$$
$$\lambda_{max} = \max_{i=1,\dots,n} \left[c_i + (|b_i| + |b_{i+1}|)\right],$$

W. Barth, R. S. Martin and J. H. Wilkinson, *Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection*, Numerische Mathematik 9, 386-393 (1967) .

Wartości własne symetrycznej i trójdiagonalnej macierzy **K**

Twierdzenie (∼Sturma):
Mamy ciąg (∼Sturma)

$$q_1(\lambda) = c_1 - \lambda$$
$$q_i(\lambda) = (c_i - \lambda) - b_i^2 / q_{i-1}(\lambda), \quad i = 2, \ldots, n$$

Liczba $a(\lambda)$, będąca liczbą ujemnych wyrazów ciągu $\{q_i(\lambda)\}_{i=1}^{n}$ jest równa liczbie wartości własnych macierzy **K** mniejszych od λ

Uwaga techniczna: gdy $q_{i-1} = 0$, wtedy $q_i(\lambda) = (c_i - \lambda) - |b_i|/\epsilon_{mach}$, gdzie ϵ_{mach} jest najmniejszą wartością ϵ, dla której $1 + \epsilon > 1$.

W. Barth, R. S. Martin and J. H. Wilkinson, *Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection*, Numerische Mathematik 9, 386-393 (1967) .

Wartości własne symetrycznej i trójdiagonalnej macierzy **K**

Bisekcja poszukująca wartości własne $\lambda_k \in [\lambda_{min}, \lambda_{max}]$, $k = n, n-1, \ldots, 1$ macierzy **K**. Niech $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$.
Na początku $x_l = \lambda_{min}$, $x_u = \lambda_{max}$

1. Jeżeli $x_u - x_l < \epsilon$, to kończymy
2. Wyznaczamy środek przedziału (bisekcja) $x = (x_l + x_u)/2$
3. Jeżeli $a(x) < k$ to $x_l = x$, w przeciwnym wypadku $x_u = x$

Uwagi:

- Można **bardzo precyzyjnie** wyznaczyć wszystkie wartości własne [2]
- Na podstawie bieżących obliczeń, można przy okazji wyciągać informacje dla pozostałych wartości własnych, tzn. zacieśniać ich przedziały przeszukiwania (zob. na następnej stronie)

W. Barth, R. S. Martin and J. H. Wilkinson, *Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection*, Numerische Mathematik 9, 386-393 (1967) .

[2] Z ograniczoną precyzją dla $\lambda_k \sim 0$

Wartości własne symetrycznej i trójdiagonalnej macierzy **K**

Modyfikacja bisekcji

- Wektor $\mathbf{w} = [w_1, \ldots, w_n]^{Tr}$ przechowuje bieżące dolne ograniczenia na wartości własne
- Wektor $\mathbf{x} = [x_1, \ldots, x_n]^{Tr}$ przechowuje bieżące górne ograniczenia na wartości własne, docelowo będzie zawierał wyznaczone wartości własne
- Na początku $w_i = \lambda_{min}$, $x_i = \lambda_{max}$.
- Bisekcja dla $k = n, \ldots, 1$, gdzie $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$.
  1. $x_l = w_k$, $x_u = x_k$, jeżeli $x_u - x_l < \epsilon$, to kończymy
  2. Wyznaczamy środek przedziału (bisekcja) $x_k = (x_l + x_u)/2$
  3. Jeżeli $a(x_k) < k$ to $x_l = x_k$, w przeciwnym wypadku $x_u = x_k$. Przy okazji jeżeli $a(x_k) < k$, to x_k ogranicza wartości $\lambda_1, \ldots, \lambda_a$ z góry, jednocześnie ogranicza z dołu wartości $\lambda_{a+1}, \lambda_{a+2} \ldots, \lambda_k$.
     - Oznacza to, że dla $i = 1, \ldots, a$ jeżeli $x_i > x_k$, to $x_i = x_k$.
     - Także to oznacza, że dla $i = a + 1, \ldots, k$ jeżeli $w_i < x_k$, to $w_i = x_k$.

W. Barth, R. S. Martin and J. H. Wilkinson, *Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection*, Numerische Mathematik 9, 386-393 (1967) .

Do zrobienia

1. Wczytać prostokątną macierz $\mathbf{A}_{m \times n}$ o elementach rzeczywistych
2. Wyznaczyć macierze $\mathbf{U}_{m \times m}$, $\mathbf{\Sigma}_{m \times n}$ $\mathbf{V}_{n \times n}$
3. Sprawdzić czy $\mathbf{A} = \mathbf{U\Sigma V}^{Tr}$
4. Wyznaczyć macierz pseudoodwrotną $\mathbf{A}^{+} = \mathbf{V\Sigma}^{+}\mathbf{U}^{Tr}$, gdzie

$$\mathbf{A}^{+}_{n \times m},\ \mathbf{V}_{n \times n},\ \mathbf{\Sigma}^{+}_{n \times m},\ \mathbf{U}^{Tr}_{m \times m}$$

5. Sprawdzić, czy $\mathbf{A}^{+}\mathbf{A} = \mathbf{1}_{n \times n}$ (czy $\mathbf{AA}^{+} = \mathbf{1}_{m \times m}$?) (spr. $\mathbf{\Sigma}^{+}\mathbf{\Sigma}$ oraz $\mathbf{\Sigma\Sigma}^{+}$)
6. Czy do zastosowań rozkładu SVD potrzebujemy wszystkich kolumn większej macierzy ortogonalnej? Jeżeli nie, to których?

Przykłady

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

$$\mathbf{A}_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 6 & 5 \end{pmatrix} \qquad \mathbf{A}_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{pmatrix}$$

$$\mathbf{A}_4 = \begin{pmatrix} 21 & 0 & 770 & 0 & 50666 \\ 0 & 770 & 0 & 50666 & 0 \\ 770 & 0 & 50666 & 0 & 3956810 \\ 0 & 50666 & 0 & 3956810 & 0 \\ 50666 & 0 & 3956810 & 0 & 335462666 \end{pmatrix}$$

Dodatek: Rozkład SVD macierzy prostokątnych z wykorzystaniem metody QR

Rozkład SVD macierzy prostokątnej

$$\mathbf{A}_{N\times M} = \mathbf{U}_{N\times N}\mathbf{\Sigma}_{N\times M}\mathbf{V}^{Tr}_{M\times M}$$

Wyznaczamy mniejszą z macierzy ortogonalnych **V** albo **U**, tzn.

$1^\circ \; N \geqslant M$

$$\left(\mathbf{A}^{Tr}\mathbf{A}\right)\mathbf{v}_k = \sigma_k^2\mathbf{v}_k, \;\; k = 1, ..., M$$

$$\Downarrow$$

$$\mathbf{V} = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; ... \;\; \mathbf{v}_M]$$

$2^\circ \; N < M$

$$\left(\mathbf{A}\mathbf{A}^{Tr}\right)\mathbf{u}_j = \sigma_j^2\mathbf{u}_j, \;\; j = 1, ..., N$$

$$\Downarrow$$

$$\mathbf{U} = [\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; ... \;\; \mathbf{u}_N]$$

Kolejność kolumn w **V** (albo **U**) jest wyznaczona relacją

$$\sigma_1 > \sigma_2 > ... > \sigma_r, \quad r = \min(N, M)$$

Jeżeli powyższe równanie własne macierzy symetrycznych $\mathbf{A}^{Tr}\mathbf{A}$ (albo $\mathbf{A}\mathbf{A}^{Tr}$) rozwiązujemy metodą QR z ortogonalizacją Grama-Schmidta, to dla macierzy symetrycznych, kolumny macierzy ortogonalnej $\mathbf{Q} = \mathbf{Q}_1\mathbf{Q}_2...\mathbf{Q}_{k_{max}}$, są wektorami własnymi rozpatrywanej macierzy symetrycznej, tzn. w naszym przypadku $\mathbf{Q} = \mathbf{V}$ (albo $\mathbf{Q} = \mathbf{U}$) o ile kolumny macierzy **Q** zostały przesortowane tak, że $\sigma_1 > \sigma_2 > \sigma_3 > ... > \sigma_r$. Macierze ortogonalne $\mathbf{Q}_k$ pochodzą z k-tej iteracji metody QR znajdowania wartości własnych macierzy kwadratowych.

Dodatek: Rozkład SVD macierzy prostokątnych z wykorzystaniem metody QR

Rozkład SVD macierzy prostokątnej

$$\mathbf{A}_{N\times M} = \mathbf{U}_{N\times N}\mathbf{\Sigma}_{N\times M}\mathbf{V}^{Tr}_{M\times M}$$

Wyznaczamy mniejszą z macierzy ortogonalnych $\mathbf{V}$ albo $\mathbf{U}$, tzn.

1° $N \geqslant M$

$$\left(\mathbf{A}^{Tr}\mathbf{A}\right)\mathbf{v}_k = \sigma_k^2\mathbf{v}_k,\ \ k = 1, ..., M$$

$$\Downarrow$$

$$\mathbf{V} = [\mathbf{v}_1\ \mathbf{v}_2\ ...\ \mathbf{v}_M]$$

2° $N < M$

$$\left(\mathbf{A}\mathbf{A}^{Tr}\right)\mathbf{u}_j = \sigma_j^2\mathbf{u}_j,\ \ j = 1, ..., N$$

$$\Downarrow$$

$$\mathbf{U} = [\mathbf{u}_1\ \mathbf{u}_2\ ...\ \mathbf{u}_N]$$

Kolejność kolumn w $\mathbf{V}$ (albo $\mathbf{U}$) jest wyznaczona relacją

$$\sigma_1 > \sigma_2 > ... > \sigma_r, \quad r = \min(N, M)$$

Jeżeli powyższe równanie własne macierzy symetrycznych $\mathbf{A}^{Tr}\mathbf{A}$ (albo $\mathbf{A}\mathbf{A}^{Tr}$) rozwiązujemy metodą QR z ortogonalizacją Grama-Schmidta, to możemy napotkać na następujące trudności

1. $\sigma_k = \sigma_{k+1}$ $\Rightarrow$ mogą wystąpić problemy ze zbieżnością metody QR
2. $\sigma_r \approx 0 \Rightarrow$ ortogonalizacja Grama-Schmidta będzie narażona na duże błędy numeryczne lub zupełnie zawiedzie
   Wskazówka: rozwiązać równanie własne dla $\mathbf{A}^{Tr}\mathbf{A} + s\mathbf{1}$ (albo $\mathbf{A}\mathbf{A}^{Tr} + s\mathbf{1}$), gdzie np. $s \sim 1$

Dodatek: Wyznaczenie drugiej (większej) macierzy ortogonalnej

Rozkład SVD macierzy prostokątnej

$$\mathbf{A}_{N\times M} = \mathbf{U}_{N\times N}\mathbf{\Sigma}_{N\times M}\mathbf{V}^{Tr}_{M\times M}$$

$1^\circ \quad N \geqslant M$

Obliczamy M początkowych kolumn macierzy $\mathbf{U}$ wg wzoru

$$\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i, \quad \sigma_i > 0$$

$$\mathbf{A}^{Tr}_{M\times N} = \mathbf{V}_{M\times M}\mathbf{\Sigma}^{Tr}_{M\times N}\mathbf{U}^{Tr}_{N\times N}$$

$2^\circ \quad N < M$

Obliczamy N początkowych kolumn macierzy $\mathbf{V}$ wg wzoru

$$\mathbf{v}_i = \frac{1}{\sigma_i}\mathbf{A}^{Tr}\mathbf{u}_i, \quad \sigma_i > 0$$

Co się dzieje, gdy $\sigma_r = 0$ (lub $\sigma_r \approx 0$) dla $r \leqslant \min(N, M)$?

$$\sigma_r\mathbf{u}_r = \mathbf{A}\mathbf{v}_r, \quad \sigma_r = 0$$

$$\Updownarrow$$

$$\mathbf{A}\mathbf{v}_r = \mathbf{0}$$

$$\sigma_r\mathbf{v}_r = \mathbf{A}^{Tr}\mathbf{u}_r, \quad \sigma_r = 0$$

$$\Updownarrow$$

$$\mathbf{A}^{Tr}\mathbf{u}_r = \mathbf{0}$$

Kolumna $\mathbf{u}_r$ (lub $\mathbf{v}_r$) wyznaczamy z warunków dla $i = 1, 2, ..., r-1$

$$\mathbf{u}_i^{Tr}\mathbf{u}_r = \delta_{ir}$$

$$\mathbf{v}_i^{Tr}\mathbf{v}_r = \delta_{ir}$$

Dodatek: Wyznaczenie drugiej (większej) macierzy ortogonalnej

Rozkład SVD macierzy prostokątnej

$$\mathbf{A}_{N\times M} = \mathbf{U}_{N\times N}\mathbf{\Sigma}_{N\times M}\mathbf{V}^{Tr}_{M\times M}$$

$1^\circ \ N \geqslant M$

Obliczamy M początkowych kolumn macierzy $\mathbf{U}$ wg wzoru

$$\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i, \quad \sigma_i > 0$$

$$\mathbf{A}^{Tr}_{M\times N} = \mathbf{V}_{M\times M}\mathbf{\Sigma}^{Tr}_{M\times N}\mathbf{U}^{Tr}_{N\times N}$$

$2^\circ \ N < M$

Obliczamy N początkowych kolumn macierzy $\mathbf{V}$ wg wzoru

$$\mathbf{v}_i = \frac{1}{\sigma_i}\mathbf{A}^{Tr}\mathbf{u}_i, \quad \sigma_i > 0$$

Kolumny $\mathbf{u}_i$ (lub $\mathbf{v}_i$), gdzie $i > \min(N, M)$, wyznaczamy z warunków dla $j = 1, 2, ..., i-1$

$$\mathbf{u}_j^{Tr}\mathbf{u}_i = \delta_{ji}, \ i = M+1, ..., N$$

$$\mathbf{v}_j^{Tr}\mathbf{v}_i = \delta_{ji}, \ i = N+1, ..., M$$

Dodatek: „Uzupełnienie” macierzy ortogonalnej

- $\mathbf{Q}$ jest macierzą kwadratową $n \times n$
- Znamy $m < n$ początkowych kolumn $\mathbf{q}_i$, gdzie $i = 1, \ldots, m$, macierzy $\mathbf{Q}$ spełniających warunki

$$\mathbf{q}_i^{Tr} \mathbf{q}_j = \delta_{ij}, \quad i, j = 1, 2, .., m$$

- Chcemy, żeby cała macierz była ortogonalna tzn., że wszystkie kolumny są wzajemnie otrogonalne i unormowane do 1
  1. Wypełniamy (nieznane) kolumny $\mathbf{q}_k$, gdzie $k > m$ liczbami (pseudo)losowymi, aby uniknąć liniowych zależności (co zrobić jeżeli okaże się, że dana kolumna $\mathbf{q}_k$, gdzie $k > m$, jest jednak liniowo zależna?)
  2. Kolejno dla kolumn $\mathbf{q}_k$, gdzie $k = m + 1, \ldots, n$, dokonujemy ortogonalizacji Grama-Schmidta, tzn.

$$\mathbf{q}'_k = \mathbf{q}_k - \sum_{i=1}^{k-1} (\mathbf{q}_k^{Tr} \mathbf{q}_i) \mathbf{q}_i \qquad \text{ortogonalizacja}$$

$$\mathbf{q}''_k = \frac{\mathbf{q}'_k}{\|\mathbf{q}'_k\|_2} \qquad \text{normalizacja}$$

gdzie

$$\|\mathbf{q}'_k\|_2 = \sqrt{\mathbf{q}'^{Tr}_k \mathbf{q}'_k} = \sqrt{\sum_{i=1}^{n} (q'_{ik})^2}$$