Metoda potęgowa & Rachunek różniczkowy macierzowy

Maciej Paszyński

Katedra Informatyki
Akademia Górniczo-Hutnicza w Krakowie
home.agh.edu.pl/paszynsk

Jeśli używasz fragmentów tego wykładu, zacytuj źródło

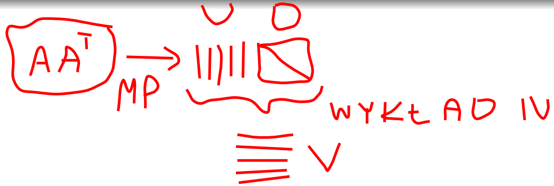
Pytanie na dzisiaj

$A = UDV$

Jaka jest relacja między **wartosciami osobliwymi** których używa metoda Singular Value Decomposition dla macierzy kwadratowej A a **wartosciami własnymi** macierzy A

$\boxed{AA^T}$ $\quad A^TA$

Program



Proszę napisać program który używając metody potęgowej znajduje SVD zadanej macierzy kwadratowej

Dla $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ szukam z i λ by $Az = \lambda z$

NIE ZEROWY

Zaczynamy od wybranego losowo $z^{(1)}$, na przykład $z^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Obliczamy $w^{(1)} = Az^{(1)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Wybieramy $\lambda^{(1)} =$ największa współrzędna z wektora $w^{(1)}$, czyli $\lambda^{(1)} = max_j |w_j^{(1)}|$. U nas $\lambda^{(1)} = 1$. Błąd iteracji NA MODUŁ

$$e^{(1)} = \|w^{(1)} - \lambda^{(1)} z^{(1)}\| = \| \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \| = \| \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \| = 1 > \epsilon$$

$\sqrt{0^2 + (-1)^2 + 0^2}$

WARUNEK STOPU

JAK O TO RESTART

Następnie $z^{(2)} = w^{(1)}/\lambda^{(1)}$. U nas $z^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} /1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ i iterujemy

$$z^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$[\,1 \quad 0 \quad 1\,]$$

$$w^{(2)} = Az^{(2)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

$\lambda^{(2)} = max_j|w_j^{(2)}| = 2.$

Błąd iteracji

$$e^{(2)} = \|w^{(2)} - \lambda^{(2)}z^{(2)}\| = \| \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \| = \| \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} \|_{\infty} = 2 > \epsilon \qquad \sqrt{(-2)^2} = 2$$

Następnie $z^{(3)} = w^{(2)}/\lambda^{(2)} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix} /2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ i iterujemy dalej

$$z^{(3)} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$[1 \quad -1 \quad 1]$

$$w^{(3)} = Az^{(3)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \\ 3 \end{bmatrix}$$

$\lambda^{(3)} = max_j |w_j^{(3)}| = 4.$

Błąd iteracji

$$e^{(3)} = \|w^{(3)} - \lambda^{(3)} z^{(3)}\| = \| \begin{bmatrix} 3 \\ -4 \\ 3 \end{bmatrix} - 4 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \| = \| \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} \|_\infty = 1 > \epsilon$$

1 = w normie $\|\cdot\|_\infty$

$\sqrt{(-1)^2+(-1)^2} = \sqrt{2}$

Następnie $z^{(4)} = w^{(3)}/\lambda^{(3)} = \begin{bmatrix} 3 \\ -4 \\ 3 \end{bmatrix} /4 = \begin{bmatrix} 0.75 \\ -1 \\ 0.75 \end{bmatrix}$ i iterujemy dalej

$$z^{(4)} = \begin{bmatrix} 0.75 \\ -1 \\ 0.75 \end{bmatrix}$$

$[0.75 \quad -1 \quad 0.75]$

$$w^{(4)} = Az^{(4)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0.75 \\ -1 \\ 0.75 \end{bmatrix} = \begin{bmatrix} 2.5 \\ -3.5 \\ 2.5 \end{bmatrix}$$

$\lambda^{(4)} = max_j |w_j^{(4)}| = 3.5.$

Błąd iteracji $e^{(4)} = \|w^{(4)} - \lambda^{(4)} z^{(4)}\| = \| \begin{bmatrix} 2.5 \\ -3.5 \\ 2.5 \end{bmatrix} - 3.5 \begin{bmatrix} 0.75 \\ -1 \\ 0.75 \end{bmatrix} \| =$

$\| \begin{bmatrix} -0.5 \\ 0.5 \\ -0.5 \end{bmatrix} \|_\infty = 0.5 > \epsilon$

$\|\cdot\|_2 = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2} < 1$

Następnie $z^{(5)} = w^{(4)}/\lambda^{(4)} = \begin{bmatrix} 2.5 \\ 3.5 \\ 2.5 \end{bmatrix} /3.5 = \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix}$ i iterujemy

dalej

Metoda potęgowa (power method) iteracja 5

$$z^{(5)} = \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix}$$

$$w^{(5)} = Az^{(5)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix} = \begin{bmatrix} 2.428 \\ -3.428 \\ 2.428 \end{bmatrix}$$

$\lambda^{(5)} = max_j|w_j^{(5)}| = 3.428.$

$A z \cong \lambda z$

Błąd iteracji $e^{(5)} = \|w^{(5)} - \lambda^{(5)} z^{(5)}\| =$

$$\left\| \begin{bmatrix} 2.428 \\ -3.428 \\ 2.428 \end{bmatrix} - 3.428 \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 0.02 \\ 0.0005 \\ 0.02 \end{bmatrix} \right\|_{\infty} = 0.02 < 0.1$$

$\|\cdot\|_2 = 0.028$

$\wedge$

0.1

Znaleźliśmy największą wartość własną $\lambda = 3.428$ oraz wektor własny $z = \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix}$

$Az \approx \lambda z$

Metoda potęgowa (power method) sprawdzenie

Znaleźliśmy największą wartość własną $\lambda = 3.428$

oraz odpowiadający mu wektor własny $z = \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix}$

$Az \approx \lambda z$

$$Az = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix} = \begin{bmatrix} 2.428 \\ -3.428 \\ 2.428 \end{bmatrix}$$

$$\lambda z = 3.428 \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix} = \begin{bmatrix} 2.448 \\ -3.428 \\ 2.448 \end{bmatrix}$$

dla dokładności $\epsilon = 0.1$

Szukamy drugiej wartosci i wektora własnego

MODYFIKACJA A USUWAMY λ_1 z

Obliczamy

$$E = A - \lambda(z/||z||z^T/||z||)$$

gdzie $Az = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0.714 \\ -1.0 \\ 0.714 \end{bmatrix} = \begin{bmatrix} 2.428 \\ -3.428 \\ 2.428 \end{bmatrix}$ $= W^{|S|}$

$$(Az)Az^T = \begin{bmatrix} 2.428 \\ -3.428 \\ 2.428 \end{bmatrix} \begin{bmatrix} 2.428 & -3.428 & 2.428 \end{bmatrix} =$$

$$\begin{bmatrix} 2.428^2 & -3.428 * 2.428 & 2.428^2 \\ -3.428 * 2.428 & (-3.428)^2 & -3.428 * 2.428 \\ 2.428^2 & -3.428 * 2.428 & 2.428^2 \end{bmatrix} =$$

$$\begin{bmatrix} 5.895 & -8.323 & 5.895 \\ -8.323 & 11.751 & -8.323 \\ 5.895 & -8.323 & 5.895 \end{bmatrix}$$

Szukamy drugiej wartosci i wektora własnego

Obliczamy

$$E = A - (Az)(Az)^T$$

gdzie $E = A - (Az)(Az)^T =$

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} - \begin{bmatrix} 5.895 & -8.323 & 5.895 \\ -8.323 & 11.751 & -8.323 \\ 5.895 & -8.323 & 5.895 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 - 5.895 & -1 - (-8.323) & 0 - 5.895 \\ -1 - (-8.323) & 2 - 11.751 & -1 - (-8.323) \\ 0 - 5.895 & -1 - (-8.323) & 2 - 5.895 \end{bmatrix} =$$

$$\begin{bmatrix} -3.895 & 7.323 & -5.895 \\ 7.323 & -9.751 & 7.323) \\ -5.895 & 7.323 & -3.895 \end{bmatrix} = E$$

Stosujemy teraz metodę potęgową względem E żeby znaleźć drugi wektor i wartosć własną macierzy A

Pochodne skalarów względem wektorów i wektorów względem skalarów

$x \in \mathcal{R}^n$ wektor, y skalar, wówczas $\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}$

Przykład $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $y = x_1 x_2 x_3$, policzyć

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \frac{\partial y}{\partial x_3} \end{bmatrix} = \begin{bmatrix} \frac{\partial (x_1 x_2 x_3)}{\partial x_1} \\ \frac{\partial (x_1 x_2 x_3)}{\partial x_2} \\ \frac{\partial (x_1 x_2 x_3)}{\partial x_3} \end{bmatrix} = \begin{bmatrix} x_2 x_3 \\ x_1 x_3 \\ x_1 x_2 \end{bmatrix}$$

$y \in \mathcal{R}^n$ wektor, x skalar, wówczas $\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \cdots & \frac{\partial y_n}{\partial x} \end{bmatrix}$

Przykład $y = \begin{bmatrix} x & x^2 & x^3 \end{bmatrix}$, policzyć

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} & \frac{\partial y_2}{\partial x} & \frac{\partial y_3}{\partial x} \end{bmatrix} = \begin{bmatrix} \frac{\partial (x)}{\partial x} & \frac{\partial (x^2)}{\partial x} & \frac{\partial (x^3)}{\partial x} \end{bmatrix} = \begin{bmatrix} 1 & 2x & 3x^2 \end{bmatrix}$$

Pochodne wektorów względem wektorów

$x \in \mathcal{R}^n, y \in \mathcal{R}^m$ to wektory

$$\frac{\partial y}{\partial x} = \left.\overbrace{\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_n}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix}}^{\text{FUNKCJE}}\right\}\text{ZMIENNE}$$

Przykład $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1^2 - x_2 \\ x_3^2 + 3x_2 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, policzyć

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} \\ \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} = \begin{array}{c} \begin{bmatrix} \frac{\partial(x_1^2 - x_2)}{\partial x_1} & \frac{\partial(x_3^2 + 3x_2)}{\partial x_1} \\ \frac{\partial(x_1^2 - x_2)}{\partial x_2} & \frac{\partial(x_3^2 + 3x_2)}{\partial x_2} \\ \frac{\partial(x_1^2 - x_2)}{\partial x_3} & \frac{\partial(x_3^2 + 3x_2)}{\partial x_3} \end{bmatrix} \end{array} \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} = \begin{bmatrix} 2x_1 & 0 \\ -1 & 3 \\ 0 & 2x_3 \end{bmatrix}$$

Własności 1/2

1)$y = x^T A x$

wówczas

$\frac{\partial y}{\partial x} = Ax + A^T x$

Jeśli A jest symetryczna, czyli $A = A^T$, wówczas

$\frac{\partial y}{\partial x} = Ax + A^T x = 2Ax$

Ponadto

$\frac{\partial}{\partial x}\frac{\partial y}{\partial x} = 2A^T$

Jeśli A jest symetryczna, czyli $A = A^T$, wówczas

$\frac{\partial}{\partial x}\frac{\partial y}{\partial x} = 2A$

Własności 2/2

2) $y = Ax$ wówczas $\frac{\partial y}{\partial x} = A^T$
Jeśli A jest symetryczna, czyli $A = A^T$, wówczas $\frac{\partial y}{\partial x} = A$

3) $y = x^T A$ wówczas $\frac{\partial y}{\partial x} = A$

4) $y = x^T x$ wówczas $\frac{\partial y}{\partial x} = 2x$

5)$x \in \mathcal{R}^n$ wektor, $y \in \mathcal{R}^r$ wektor, $z \in \mathcal{R}^m$ wektor, oraz $z = y(x)$
wówczas $\frac{\partial z}{\partial x} = \frac{\partial y}{\partial x} \frac{\partial x}{\partial y}$