

SPRAWOZDANIE - LABORATORIUM 5

Diagonalizacja macierzy metodą potęgową

Marta Dychała

14 kwietnia 2021

1 Wstęp teoretyczny

Jedną z metod iteracyjnych wyznaczania wektorów własnych macierzy \mathbf{A} jest metoda potęgowa. W metodzie tej zakłada się, że istnieje n niezależnych wektorów własnych macierzy \mathbf{A} , które stanowią bazę przestrzeni liniowej: x_1, x_2, \dots, x_n oraz że wartości własne macierzy \mathbf{A} tworzą ciąg: $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$; wtedy dla wartości własnej λ_i w m -tej iteracji spełnione jest równanie:

$$Av_0 = \sum_{i=1}^n a_i \lambda_i x_i$$
$$v_m = A^m v_0 = \sum_{i=1}^n a_i \lambda_i^m x_i, \quad (1)$$

gdzie wektor v_0 jest wektorem postaci:

$$v_0 = \sum_{i=1}^n a_i x_i.$$

Jeżeli λ_i jest dominującą wartością własną (tj. $\frac{\lambda_i}{\lambda_1} < 1, i > 1$), wówczas zachodzi równość:

$$\lim_{m \rightarrow \infty} \frac{A^m v_0}{\lambda_1^m} = a_1 x_1,$$

z której wynika, że wartość własną można obliczyć następująco:

$$\lambda_1 = \lim_{m \rightarrow \infty} \frac{y^T v_{m+1}}{y^T v_m},$$

gdzie y to dowolny wektor nieortogonalny do x_1 . W celu wyznaczenia wektora własnego x_1 , można skorzystać z przekształcenia równania (1):

$$v_m = \sum_{i=1}^n a_i \lambda_i^m x_i = \lambda_1^m \left[a_1 x_1 + \sum_{i=2}^n \left(\frac{\lambda_i}{\lambda_1} \right)^m a_i x_i \right]. \quad (2)$$

Zgodnie z przyjętymi założeniami, λ_1 jest dominującą wartością własną, w wyniku czego można uprościć równanie (2) do postaci:

$$v_m \approx \lambda_1^m a_1 x_1;$$

dlatego też unormowany wektor własny x_1 ma postać:

$$x_1 = \frac{v_m}{\|v_m\|_2}$$

Jest to jednak metoda pozwalająca wyznaczyć tylko pierwszy wektor własny. W celu wyznaczenia pozostałych wektorów własnych można skorzystać z metody redukcji macierzy, która polega na odpowiednim

zredukowaniu macierzy wejściowej po każdej iteracji. Jeżeli λ_1 jest wartością własną macierzy \mathbf{A} , a x_1 to odpowiadający jej wektor, to dla dowolnego wektora v spełniającego warunek:

$$v^T x_1 = 1$$

Macierz zredukowana

$$W_1 = A - \lambda_1 x_1 v^T \quad (3)$$

ma te same wartości co macierz \mathbf{A} oprócz λ_1 , która ma wartość 0.

Jedną z metod wyznaczania wartości własnych, która wykorzystuje redukcję macierzy jest redukcja Hotellinga. W metodzie tej za wektor v przyjmuje się lewy wektor własny przynależny do wartości własnej λ_i . Ponieważ często nie jest znana postać lewych wektorów własnych, to metoda ta jest najbardziej skuteczna w przypadku obliczania wartości własnych macierzy symetrycznej; wówczas lewe wektory mają identyczną postać, co prawe:

$$v = x_1,$$

w wyniku czego równanie (3) upraszcza się do postaci:

$$W_1 = A - \lambda_1 x_1 x_1^T,$$

co pozwala na iteracyjne wyznaczenie wszystkich wartości własnych:

$$W_0 = A$$

$$W_i = W_{i-1} - \lambda_{i-1} x_{i-1} x_{i-1}^T, \quad i = 1, 2, \dots, n-1$$

2 Zadanie do wykonania

2.1 Opis problemu

Głównym celem zajęć laboratoryjnych było wyznaczenie wartości własnych macierzy \mathbf{A} wykorzystując w tym celu metodę potęgową oraz na podstawie otrzymanych wektorów własnych - dokonanie diagonalizacji macierzy \mathbf{A} . Elementy macierzy \mathbf{A} zostały zdefiniowane zgodnie z poniższym wzorem:

$$\mathbf{A}_{ij} = \frac{1}{\sqrt{2 + |i - j|}}, \quad i, j = 0, 1, \dots, n-1, \quad (4)$$

gdzie n to wymiar macierzy \mathbf{A} . W czasie zajęć przyjęto założenie, że $n = 7$. Ponieważ macierz \mathbf{A} wyznaczona ze wzoru (4) jest macierzą symetryczną, wartości własne zostały wyznaczone iteracyjnie za pomocą poniższego algorytmu korzystającego z metody Hotellinga:

$$\begin{aligned} & W_0 = A \quad (\text{inicjalizacja macierzy iterującej}) \\ & \text{for}(k = 0; k < K_{val}; k++) \{ \\ & \quad x_k^0 = [1, 1, \dots, 1] \\ & \quad \text{for}(i = 1; i \leq IT_MAX; i++) \{ \\ & \quad \quad x_k^{i+1} = W_k x_k^i \quad (\text{inicjalizacja wektora startowego}) \\ & \quad \quad \lambda_k^i = \frac{(x_k^{i+1})^T x_k^i}{(x_k^i)^T x_k^i} \\ & \quad \quad x_k^i = \frac{x_k^{i+1}}{\|x_k^{i+1}\|_2} \\ & \quad \} \\ & \quad W_{k+1} = W_k - \lambda_k x_k^i (x_k^i)^T \quad (\text{iloczyn tensorowy}) \\ & \} \end{aligned}$$

W naszym przypadku $K_{val} = 7$. Za liczbę iteracji dla każdej z siedmiu wartości własnych λ_k przyjęto wartość parametru $IT_MAX = 12$. Ponadto za pomocą powyższego algorytmu wyznaczono macierz \mathbf{X} , której kolumnami były wektory własne x_k odpowiadające k -tym wartościom własnym:

$$X = [x_0, x_1, \dots, x_{n-1}].$$

Utworzona w ten sposób macierz \mathbf{X} została wykorzystana do utworzenia macierzy \mathbf{D} , zdefiniowanej jako iloczyn:

$$D = X^T A X. \quad (5)$$

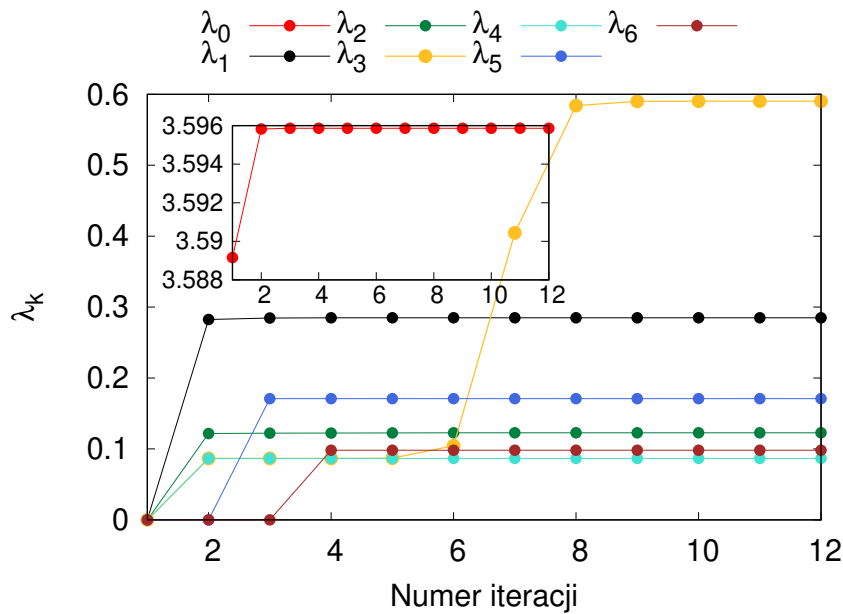
2.2 Wyniki

Wszystkie obliczenia wykonywano korzystając z typu zmiennoprzecinkowego podwójnej precyzji (double).

Wykorzystując metodę potęgową i przyjmując, że $IT_MAX = 12$, obliczono wartości własne macierzy \mathbf{A} :

$$\begin{aligned}\lambda_0 &= 3.59586 \\ \lambda_1 &= 0.284988 \\ \lambda_2 &= 0.122785 \\ \lambda_3 &= 0.59039 \\ \lambda_4 &= 0.0865954 \\ \lambda_5 &= 0.170974 \\ \lambda_6 &= 0.0981544\end{aligned}$$

Przybliżenia każdej z wartości własnych po każdej iteracji zapisano do pliku, na podstawie którego wygenerowano poniższy wykres:



Rysunek 1: Kolejne przybliżenia wartości własnych w zależności od numeru iteracji.

Otrzymane wektory własne miały postać:

$$\begin{aligned}\vec{x}_0 &= \begin{pmatrix} 0.352941 \\ 0.377935 \\ 0.392221 \\ 0.396889 \\ 0.392221 \\ 0.377935 \\ 0.352941 \end{pmatrix} & \vec{x}_1 &= \begin{pmatrix} 0.477239 \\ 0.164774 \\ -0.314852 \\ -0.540298 \\ -0.314852 \\ 0.164774 \\ 0.477239 \end{pmatrix} & \vec{x}_2 &= \begin{pmatrix} 0.361183 \\ -0.463973 \\ -0.140424 \\ 0.518765 \\ -0.140416 \\ -0.46396 \\ 0.361198 \end{pmatrix} & \vec{x}_3 &= \begin{pmatrix} 0.479171 \\ 0.446604 \\ 0.266346 \\ -2.40173e-06 \\ -0.266339 \\ -0.446612 \\ -0.479167 \end{pmatrix} \\ \vec{x}_4 &= \begin{pmatrix} 0.130602 \\ -0.338044 \\ 0.476978 \\ -0.531332 \\ 0.477006 \\ -0.337998 \\ 0.130652 \end{pmatrix} & \vec{x}_5 &= \begin{pmatrix} -0.449005 \\ 0.172676 \\ 0.518245 \\ -2.51207e-09 \\ -0.518245 \\ -0.172676 \\ 0.449005 \end{pmatrix} & \vec{x}_6 &= \begin{pmatrix} -0.262282 \\ 0.520312 \\ -0.400604 \\ -1.93979e-14 \\ 0.400604 \\ -0.520312 \\ 0.262282 \end{pmatrix}\end{aligned}$$

W ten sposób macierz \mathbf{D} otrzymana ze wzoru (5) miała postać:

$$\mathbf{D} = \begin{pmatrix} 3.59586 & -1.17462e-13 & 5.55112e-17 & -1.11022e-16 & -9.71445e-16 & -1.11022e-16 & 5.55112e-17 \\ -1.17427e-13 & 0.284988 & -6.25267e-06 & -1.97872e-11 & -3.81662e-09 & 3.46945e-17 & 6.245e-17 \\ 4.16334e-17 & -6.25267e-06 & 0.122786 & -8.59639e-06 & -0.000329107 & -1.48159e-12 & -1.21431e-16 \\ 1.94289e-16 & -1.97871e-11 & -8.59639e-06 & 0.59039 & -2.99717e-05 & -9.76996e-15 & 4.16334e-17 \\ -9.51496e-16 & -3.81662e-09 & -0.000329107 & -2.99717e-05 & 0.0865957 & -1.18642e-09 & -8.21392e-16 \\ -2.04697e-16 & 0 & -1.48159e-12 & -9.8116e-15 & -1.18642e-09 & 0.170974 & -1.83654e-09 \\ 6.245e-17 & 6.76542e-17 & -7.63278e-17 & 9.02056e-17 & -8.41341e-16 & -1.83654e-09 & 0.0981544 \end{pmatrix}$$

Otrzymana macierz \mathbf{D} różni się od spodziewanego wyniku, ponieważ pomimo faktu, że na głównej przekątnej macierzy \mathbf{D} znajdują się kolejne wartości własne, to poza nią teoretycznie powinny znajdować się same zera, ponadto wyznaczone wartości własne nie są ułożone w kolejności malejącej. Niezerowe wartości (ale bliskie zera) poza diagonalą macierzy \mathbf{D} są związane z błędami numerycznymi związanymi z zastosowaniem do obliczeń typu double, natomiast niewłaściwa kolejność wyznaczonych wartości własnych wynika z ograniczonej liczby iteracji, która okazała się być w tym przypadku niewystarczająca.

Dlatego też w poniższej tabeli znajdują się wartości IT_MAX dla których zostają wyznaczone kolejne wartości własne z dokładnością do sześciu miejsc znaczących:

Tabela 1: Wartości przybliżeń wartości własnych w zależności od liczby iteracji. Żółtym kolorem oznaczono wartości, które pozostają niezmiennie po zwiększonej wartości IT_MAX .

IT_MAX	λ_0	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
3	3.59586	0.284526	0.122311	0.0870731	0.00026213	0.590387	0.000243218
12	3.59586	0.284988	0.122785	0.59039	0.0865954	0.170974	0.0981544
54	3.59586	0.590389	0.284988	0.122787	0.170974	0.0865947	0.0981544
55	3.59586	0.59039	0.284988	0.122787	0.170974	0.0865947	0.0981544
112	3.59586	0.59039	0.284988	0.170974	0.122787	0.0865947	0.0981544
283	3.59586	0.59039	0.284988	0.170974	0.122787	0.0981544	0.0865947
400	3.59586	0.59039	0.284988	0.170974	0.122787	0.0981544	0.0865947
500	3.59586	0.59039	0.284988	0.170974	0.122787	0.0981544	0.0865947

Z powyższej tabeli można wywnioskować, że wraz ze wzrostem liczby iteracji wzrasta dokładność wyznaczenia wartości własnej. Na przykład wartość własna λ_0 zostaje wyznaczona z dokładnością do sześciu miejsc znaczących już dla $IT_MAX = 3$, natomiast λ_1 dla $IT_MAX = 55$. Po 284 iteracjach nie zauważono już większych zmian w przybliżeniach wartości własnych, co udowadniają wyznaczone przybliżenia dla wartości własnych dla $IT_MAX = 400$ oraz $IT_MAX = 500$.

3 Wnioski

Metoda potęgowa jest wyjątkowo prostą metodą do zaimplementowania na macierzy o małym wymiarze, tak jak to miało miejsce w naszym przypadku, jednakże jest ona mniej wydajna, jeżeli używa się jej na macierzach o dużych wymiarach, bowiem do jej zaimplementowania wymagana jest duża ilość pamięci.

Okazało się, że 12 iteracji potrzebne do obliczenia każdej z wartości własnej były niewystarczające, bowiem otrzymane wartości własne nie były malejące wraz ze zwiększeniem się k . W odpowiedniej kolejności ułożyły się dopiero po przyjęciu parametru $IT_MAX = 283$. Duży wpływ na otrzymane wartości własne ma postać macierzy \mathbf{W} , która zależy od liczby iteracji potrzebnej na obliczenie danej wartości własnej. Przykładowo, zanim zostanie obliczona wartość λ_6 dla $IT_MAX = 3$, macierz \mathbf{W} przekształcana jest 18-krotnie, natomiast dla $IT_MAX = 283$ - aż 1698-krotnie.

Omawiana metoda zachowuje się dosyć niestabilnie. Dowodem na poparcie tej tezy jest chociażby otrzymana macierz \mathbf{D} . Zgodnie ze wzorem (5) powinna być ona macierzą diagonalną, jednak tak nie było, bowiem poza jej główną diagonalą znajdowały się niezerowe wartości, które są skutkiem zastosowania do obliczeń typu podwójnej precyzji.