

Wstęp do metod numerycznych

Numeryczne zagadnienie własne

P. F. Góra

<http://th-www.if.uj.edu.pl/zfs/gora/>

2010

Zagadnienie własne

Definicja: Niech $A \in \mathbb{C}^{N \times N}$. Liczbę $\lambda \in \mathbb{C}$ nazywam *wartością własną macierzy A* , jeżeli istnieje niezerowy wektor $x \in \mathbb{C}^N$ taki, że

$$Ax = \lambda x. \quad (1)$$

Wektor x nazywam wówczas *wektorem własnym macierzy A do wartości własnej λ* .

Definicja jest sformułowana dla macierzy zespolonych, my jednak — poza przypadkiem macierzy hermitowskich — będziemy zajmować się macierzami rzeczywistymi. Należy jednak pamiętać, że także macierze rzeczywiste (niesymetryczne) mogą mieć zespolone wartości własne.

Problem poszukiwania wartości własnych macierzy nazywa się *zagadnieniem własnym*.

Równanie charakterystyczne

Równanie (1) można zapisać w postaci

$$(\mathbf{A} - \lambda \mathbb{I}) \mathbf{x} = \mathbf{0} . \quad (2)$$

Jeżeli λ jest wartością własną, równanie to musi mieć niezerowe rozwiązanie ze względu na \mathbf{x} . Jest to możliwe tylko w wypadku

$$\det (\mathbf{A} - \lambda \mathbb{I}) = 0 . \quad (3)$$

Równanie (3) nazywane jest *równaniem charakterystycznym macierzy \mathbf{A}* . Jest to równanie wielomianowe stopnia N . Widzimy, że każda wartość własna jest pierwiastkiem równania charakterystycznego, ale stwierdzenie odwrotne niekoniecznie jest prawdą.

Zbiór wszystkich rozwiązań równania (3) nazywam *widmem macierzy \mathbf{A}* .

Macierz $A \in \mathbb{C}^{N \times N}$, która ma N niezależnych liniowo wektorów własnych, nazywam *macierzą diagonalizowalną*. Łatwo sprawdzić, że jeżeli $X \in \mathbb{C}^{N \times N}$ jest macierzą, której kolumnami są kolejne, liniowo niezależne wektory własne diagonalizowalnej macierzy A , zachodzi

$$X^{-1}AX = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}. \quad (4)$$

Uwaga! Poza przypadkami, które daje się rozwiązać analitycznie i poza pewnymi przypadkami specjalnymi, które wykraczają poza zakres tego wykładu, próba poszukiwania wartości własnych macierzy poprzez rozwiązywanie jej równania charakterystycznego *na ogół jest numerycznie nieefektywna*. Trzeba znaleźć jakieś inne metody.

Transformacja podobieństwa

Zależność (4) jest szczególnym przypadkiem *transformacji podobieństwa*. Ogólnie, dwie macierze A , B nazywam macierzami podobnymi, jeżeli istnieje taka nieosobliwa macierz S , że zachodzi

$$B = S^{-1}AS. \quad (5)$$

Macierze podobne mają takie same widma, co wynika z faktu, że jeśli spełnione jest równanie (1), spełnione jest także równanie

$$S^{-1}(A - \lambda I)SS^{-1}x = 0, \quad (6)$$

a zatem wektor $S^{-1}x$ jest wektorem własnym macierzy (5) do wartości własnej λ .

Ortogonalna transformacja podobieństwa

Na ogół zamiast ogólnej transformacji podobieństwa, używa się *ortogonalnej transformacji podobieństwa*

$$\mathbf{B} = \mathbf{O}^T \mathbf{A} \mathbf{O}, \quad (7)$$

gdzie $\mathbf{O}^T \mathbf{O} = \mathbb{I}$. Co więcej, typowo wykonuje się cały ciąg transformacji $\mathbf{A}_2 = \mathbf{O}_1^T \mathbf{A}_1 \mathbf{O}_1$, $\mathbf{A}_3 = \mathbf{O}_2^T \mathbf{A}_2 \mathbf{O}_2$ itd, czyli

$$\mathbf{A}_k = \underbrace{\mathbf{O}_k^T \dots \mathbf{O}_2^T \mathbf{O}_1^T}_{\mathbf{P}_k^T} \mathbf{A}_1 \underbrace{\mathbf{O}_1 \mathbf{O}_2 \dots \mathbf{O}_k}_{\mathbf{P}_k} \quad (8)$$

\mathbf{P}_k jest złożeniem macierzy ortogonalnych, a więc samo jest macierzą ortogonalną.

Ponieważ kolejne transformacje wykonuje się w kolejności “od wewnątrz”, nie musimy pamiętać każdej z nich z osobna, a tylko złożenia, według schematu

$$\mathbf{P}_0 = \mathbb{I}, \quad \mathbf{P}_l = \mathbf{P}_{l-1} \mathbf{O}_l. \quad (9)$$

Zakumulowana macierz \mathbf{P}_k jest potrzebna do znalezienia wektorów własnych. W szczególności jeśli \mathbf{A}_1 (a zatem \mathbf{A}_2 , \mathbf{A}_3 itd) macierzą symetryczną, natomiast \mathbf{A}_k jest macierzą diagonalną (lub “numerycznie diagonalną”, to znaczy elementy pozadiagonalne są dostatecznie małe), \mathbf{P}_k jest jest macierzą, której kolumnami są wektory własne.

Uwaga: Jeżeli nie musimy znaleźć wszystkich wektorów własnych i interesują nas tylko wartości własne, **nie musimy zapamiętywać** zakumulowanych macierzy transformacji, co istotnie przyczynia się do zmniejszenia numerycznego kosztu całej procedury.

Metoda potęgowa

Niech $\mathbf{A} \in \mathbb{R}^{N \times N}$ będzie macierzą symetryczną, $\mathbf{A} = \mathbf{A}^T$. Wiadomo, że macierz taka jest diagonalizowalna, ma rzeczywiste wartości własne, a jej unormowane wektory własne tworzą bazę ortonormalną w \mathbb{R}^N . Niech bazą tą będzie $\{\mathbf{e}_i\}_{i=1}^N$, przy czym $\mathbf{A}\mathbf{e}_i = \lambda_i\mathbf{e}_i$. Przyjmijmy dodatkowo, że wartości własne macierzy \mathbf{A} są dodatnie i uporządkowane $\lambda_1 > \lambda_2 > \dots > \lambda_N > 0$.

Weźmy wektor $\mathbf{y} \in \mathbb{R}^N$. Posiada on rozkład

$$\mathbf{y} = \sum_{i=1}^N \beta_i \mathbf{e}_i. \quad (10)$$

Obliczamy

$$\mathbf{A}\mathbf{y} = \mathbf{A} \sum_{i=1}^N \beta_i \mathbf{e}_i = \sum_{i=1}^N \beta_i \mathbf{A} \mathbf{e}_i = \sum_{i=1}^N \beta_i \lambda_i \mathbf{e}_i \quad (11a)$$

$$\mathbf{A}^2 \mathbf{y} = \mathbf{A} \sum_{i=1}^N \beta_i \lambda_i \mathbf{e}_i = \sum_{i=1}^N \beta_i \lambda_i^2 \mathbf{e}_i \quad (11b)$$

.....

$$\mathbf{A}^k \mathbf{y} = \sum_{i=1}^N \beta_i \lambda_i^k \mathbf{e}_i \quad (11c)$$

Widzimy, że dla dostatecznie dużych k wyraz zawierający λ_1^k będzie dominował w sumie po prawej stronie (11c): pozostałe współczynniki będą zaniedbywalnie małe, a zatem prawa strona (11c) dla dostatecznie dużych k będzie dążyć do wektora proporcjonalnego do \mathbf{e}_1 , czyli do wektora własnego do największej wartości własnej.

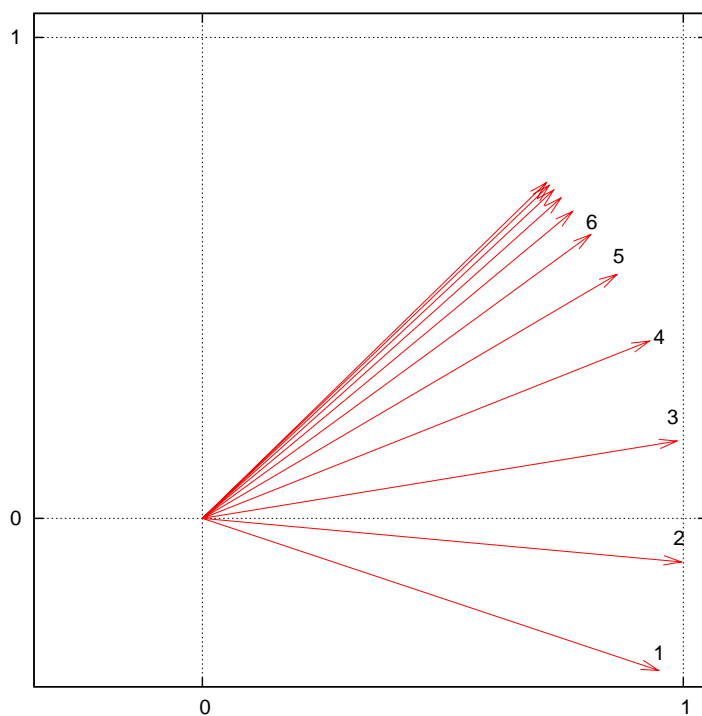
A zatem iteracja ($\|y_1\| = 1$, poza tym dowolny)

$$\begin{aligned} \mathbf{A}y_k &= \mathbf{z}_k \\ y_{k+1} &= \frac{\mathbf{z}_k}{\|\mathbf{z}_k\|} \end{aligned} \tag{12}$$

zbiega się, przy przyjętych założeniach, do unormowanego wektora własnego macierzy \mathbf{A} , odpowiadającego największej wartości własnej.

Gdy iteracja (12) zbiegnie się do punktu stałego (kolejne wektory $y_k \simeq e_1$ przestaną się zauważalnie zmieniać), wartość własną obliczamy jako $\lambda_1 = \|\mathbf{z}_k\|$.

Przykład



Kolejne wektory
generowane przez
metodę potęgową, czyli
iterację (12), dla macierzy
$$\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}.$$

Druga wartość własna

Kiedy w sumie (11c) wyraz z wektorem własnym do największej wartości własnej **nie będzie** dominować? Wtedy i tylko wtedy, gdy współczynnik odpowiadający temu wektorowi w rozwinięciu (10) będzie znikać, $\beta_1 = 0$. To sugeruje sposób na szukanie wektora własnego odpowiadającego *drugiej* co do wielkości wartości własnej: należy upewnić się, że wektor y_1 i wszystkie jego iteraty są prostopadłe do uprzednio znalezionej e_1 .

Iteracja ($\|y_1\| = 1$, $e_1^T y_1 = 0$, poza tym dowolny)

$$\begin{aligned} A y_k &= z_k \\ z_k &= z_k - e_1 (e_1^T z_k) \\ y_{k+1} &= \frac{z_k}{\|z_k\|} \end{aligned} \tag{13}$$

zbiega się, przy przyjętych założeniach, do unormowanego wektora własnego macierzy A , odpowiadającego drugiej co do wielkości wartości własnej. Drugi krok powyższego algorytmu oznacza ortogonalizację. Teoretycznie, skoro wektor startowy iteracji jest ortogonalny do e_1 , krok tn można pominąć. Tak byłoby w arytmetyce dokładnej. W arytmetyce przybliżonej zakumulowany błąd obcięcia może spowodować, że wygenerowane wektory będą miały niezerowy rzut na kierunek e_1 , który w ten sposób zacznie dominować. Dlatego reortogonalizacja, jeśli nie w każdej iteracji, to przynajmniej co kilka, jest nieodzowna, podrażając koszt numeryczny całego algorytmu.

Odwrotna metoda potęgowa

Przypuśćmy, że zamiast największej, szukamy najmniejszej (lecz większej od zera) wartości własnej. Jak łatwo sprawdzić, jeżeli $\mathbf{A}\mathbf{e}_i = \lambda_i\mathbf{e}_i$, to $\mathbf{A}^{-1}\mathbf{e}_i = \lambda_i^{-1}\mathbf{e}_i$, a zatem najmniejsza wartość własna jest największą wartością własną macierzy odwrotnej. Prowadzimy więc iterację

$$\begin{aligned}\mathbf{A}^{-1}\mathbf{y}_k &= \mathbf{z}_k \\ \mathbf{y}_{k+1} &= \frac{\mathbf{z}_k}{\|\mathbf{z}_k\|}\end{aligned}\tag{14}$$

Należy pamiętać, że zapis $\mathbf{A}^{-1}\mathbf{y}_k = \mathbf{z}_k$ należy rozumieć jako konieczność rozwiązania równania $\mathbf{A}\mathbf{z}_k = \mathbf{y}_k$. Widać, że kolejne równania rozwiązujemy korzystając z dokonanego *tylko raz* rozkładu LU (lub rozkładu Cholesky'ego, jeśli macierz jest dodatnio określona — tak, jak przy obowiązujących dotąd założeniach).

Uwagi o metodzie potęgowej

1. Jeżeli nie ograniczamy się do macierzy dodatnio określonych, dominującą wartością własną będzie wartość własna o największym module. Podobnie, dominującą wartością własną macierzy odwrotnej będzie wartość własna macierzy nieodwróconej o najmniejszym module; trzeba jednak pamiętać, że macierz odwrotna jest określona tylko wtedy, gdy żadna wartość własna nie jest zerowa.
2. Ujemne wartości własne rozpoznajemy po tym, że po ustabilizowaniu się kierunku wektora własnego, jego współrzędne zmieniają znak w kolejnych iteracjach.
3. Jeżeli przy pomocy metody potęgowej trzeba znaleźć więcej niż kilka wartości i wektorów własnych, metoda, z uwagi na konieczność reortogonalizacji, staje się bardzo kosztowna.
4. Jeżeli w widmie macierzy pojawiają się wartości własne bardzo zbliżone *co do modułu* — na przykład 2 oraz 1.99999999, ale też 1 oraz -0.99999999 — zbieżność będzie bardzo wolna, a nawet można spodziewać się stagnacji.
5. Metoda potęgowa *nie działa* dla macierzy niesymetrycznych!