Po co nam wartości własne?

Tadeusz KOŹNIEWSKI[*]

Rzeczywistą **macierzą** $m \times n$ nazywamy prostokątną tablicę

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

o m rzędach poziomych, zwanych wierszami, oraz n rzędach pionowych, zwanych kolumnami, w której a_{ij} są liczbami rzeczywistymi. W skrócie piszemy: $A = [a_{ij}]$. Zbiór wszystkich rzeczywistych macierzy $m \times n$ oznaczamy $\mathbf{M}_{m\times n}$. Iloczynem macierzy $A = [a_{ij}] \in \mathbf{M}_{m\times n}$ i macierzy $B = [b_{ij}] \in \mathbf{M}_{n\times k}$ nazywamy macierz $C = [c_{ij}] \in \mathbf{M}_{m\times k}$, której wyrazy dane są wzorami $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$, dla $i = 1, \ldots, m$, $j = 1, \ldots, k$. Taką macierz C oznaczamy $A \cdot B$.

Na przykład:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 & 3 & 1 \\ 7 & 1 & 0 & 4 \\ 0 & 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 1 & 0 & 4 \\ 9 & 8 & 7 & 7 \end{bmatrix}.$$

Dla macierzy

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & 1 \end{bmatrix} \in \mathbf{M}_{n\times n}$$

(jedynki na przekątnej, poza nią zera) zachodzi $M \cdot I = M = I \cdot M$ dla każdej macierzy $M \in \mathbf{M}_{n\times n}$. Macierz I nazywa się macierzą jednostkową. Dla każdych macierzy $A, B, C \in \mathbf{M}_{n\times n}$ zachodzi równość $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. Wspólną wartość obu stron powyższej równości oznaczamy $A \cdot B \cdot C$ (i podobnie dla większej liczby macierzy).

W zastosowaniach algebry liniowej ważną rolę odgrywa sytuacja, gdy daną macierz $A \in \mathbf{M}_{n\times n}$ wielokrotnie mnoży się przez nią samą, to znaczy oblicza się jej potęgę $A^m = A \cdot A \cdot \ldots \cdot A$. Wielokrotne mnożenie macierzy, jakie się tu wykonuje, jest często czasochłonne, a w zastosowaniach praktycznych kosztowne. Czytelnik może, na przykład, spróbować obliczyć A^{100} dla $A = \left[\begin{smallmatrix} -1 & 2 \\ -6 & 6 \end{smallmatrix}\right]$, obliczając kolejno $A^2, A^3, \ldots$ Czy można to zrobić prościej? Tak! Służą do tego **wartości własne**.

Zacznijmy od następującej obserwacji. Jeśli macierz $B \in \mathbf{M}_{n\times n}$ jest postaci

$$B = \begin{bmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & r_n \end{bmatrix}$$

(taką macierz nazywamy **diagonalną**), to

$$B^2 = \begin{bmatrix} r_1^2 & 0 & \dots & 0 \\ 0 & r_2^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & r_n^2 \end{bmatrix}$$

i ogólnie

$$B^m = \begin{bmatrix} r_1^m & 0 & \dots & 0 \\ 0 & r_2^m & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & r_n^m \end{bmatrix}$$

dla dowolnej liczby naturalnej m. Na przykład $\left[\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right]^m = \left[\begin{smallmatrix} 2^m & 0 \\ 0 & 3^m \end{smallmatrix}\right]$. Zatem potęgowanie macierzy diagonalnych jest łatwe. Potęgowanie dowolnej macierzy $A \in \mathbf{M}_{n\times n}$ bardzo by się więc uprościło, gdyby można je było sprowadzić do potęgowania macierzy diagonalnej.

Mówimy, że liczba r jest wartością własną macierzy $A \in \mathbf{M}_{n\times n}$, jeśli istnieją liczby rzeczywiste $x_1, \ldots, x_n$, nie wszystkie równe zeru, spełniające równość

$$A \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} rx_1 \\ \vdots \\ rx_n \end{bmatrix}.$$

Ciąg $(x_1, \ldots, x_n)$ nazywamy wówczas **wektorem własnym** macierzy A o wartości własnej r.
Na przykład dla $A = \left[\begin{smallmatrix} -1 & 2 \\ -6 & 6 \end{smallmatrix}\right]$ mamy $\left[\begin{smallmatrix} -1 & 2 \\ -6 & 6 \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 4 \\ 6 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} -1 & 2 \\ -6 & 6 \end{smallmatrix}\right] \cdot \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 3 \\ 6 \end{smallmatrix}\right]$, więc wektor $(2, 3)$ jest wektorem własnym o wartości własnej 2, a wektor $(1, 2)$ jest wektorem własnym o wartości własnej 3.

Można udowodnić, że jeśli macierz $A \in \mathbf{M}_{n\times n}$ ma n różnych wartości własnych $r_1, r_2, \ldots, r_n$ oraz $v_1, v_2, \ldots, v_n$ są wektorami własnymi o tych wartościach własnych, to macierz $C \in \mathbf{M}_{n\times n}$, której kolumnami są kolejno $v_1, v_2, \ldots, v_n$, ma niezerowy wyznacznik. Na przykład dla rozpatrywanej powyżej macierzy $A = \left[\begin{smallmatrix} -1 & 2 \\ -6 & 6 \end{smallmatrix}\right]$ i jej wektorów własnych $(2, 3)$, $(1, 2)$ otrzymujemy $C = \left[\begin{smallmatrix} 2 & 1 \\ 3 & 2 \end{smallmatrix}\right]$. I faktycznie wyznacznik $\det C = 4 - 3 = 1$ jest różny od zera.

Zauważmy, że jeśli $v_1, v_2, \ldots, v_n$ są wektorami własnymi macierzy $A \in \mathbf{M}_{n\times n}$ o wartościach własnych $r_1, r_2, \ldots, r_n$ i $C = [c_{ij}] \in \mathbf{M}_{n\times n}$ jest macierzą o kolumnach $v_1, v_2, \ldots, v_n$, to z powyższej definicji wektora własnego dostajemy, że

$$A \cdot \begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{nj} \end{bmatrix} = \begin{bmatrix} r_j c_{1j} \\ r_j c_{2j} \\ \vdots \\ r_j c_{nj} \end{bmatrix}$$

dla $j = 1, \ldots, n$. Zestawiając razem tych n równości, dostajemy $A \cdot C = C \cdot B$, gdzie

$$B = \begin{bmatrix} r_1 & 0 & \dots & 0 \\ 0 & r_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & & \dots & r_n \end{bmatrix}.$$

Skorzystamy teraz z faktu, że dla każdej macierzy $C \in \mathbf{M}_{n\times n}$ o niezerowym wyznaczniku istnieje taka macierz $D \in \mathbf{M}_{n\times n}$, że $C \cdot D = I$ (macierz jednostkowa). Wówczas zachodzi też $D \cdot C = I$. Macierz D nazywamy macierzą **odwrotną** do C

[*]Instytut Matematyki, Uniwersytet Warszawski

i oznaczamy C^{-1}. Na przykład dla macierzy $C = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$ mamy $C^{-1} = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$, bo faktycznie $\begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$. Mnożąc obie strony równości $A \cdot C = C \cdot B$ z prawej przez C^{-1} i uwzględniając, że $A \cdot C \cdot C^{-1} = A \cdot I = A$, dostajemy $A = C \cdot B \cdot C^{-1}$. Wówczas dla każdego naturalnego m mamy

$$\begin{aligned} A^m &= A \cdot A \cdot \ldots \cdot A = \\ &= C \cdot B \cdot C^{-1} \cdot C \cdot B \cdot C^{-1} \cdot \ldots \cdot C \cdot B \cdot C^{-1} = \\ &= C \cdot B \cdot I \cdot B \cdot I \cdot \ldots \cdot I \cdot B \cdot C^{-1} = \\ &= C \cdot B \cdot B \cdot \ldots \cdot B \cdot C^{-1} = C \cdot B^m \cdot C^{-1}. \end{aligned}$$

Na przykład dla $A = \begin{bmatrix} -1 & 2 \\ -6 & 6 \end{bmatrix}$ mamy $C = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$ oraz $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$. Stąd $A^{100} = C \cdot B^{100} \cdot C^{-1} =$
$= \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2^{100} & 0 \\ 0 & 3^{100} \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 2^{101} & 3^{100} \\ 3 \cdot 2^{100} & 2 \cdot 3^{100} \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix} =$
$= \begin{bmatrix} 2^{102} - 3^{101} & -2^{101} + 2 \cdot 3^{100} \\ 3 \cdot 2^{101} - 2 \cdot 3^{101} & -3 \cdot 2^{100} + 4 \cdot 3^{100} \end{bmatrix}$.

Konkluzja. Aby obliczyć A^m dla danej macierzy $A \in \mathbf{M}_{n \times n}$, wystarczy znaleźć macierz diagonalną $B \in \mathbf{M}_{n \times n}$ oraz macierz $C \in \mathbf{M}_{n \times n}$ takie, że $A = C \cdot B \cdot C^{-1}$. Wówczas obliczenie $A^m = C \cdot B^m \cdot C^{-1}$ jest proste, bo potęgowanie macierzy diagonalnej B jest proste. W przypadku, gdy macierz $A \in \mathbf{M}_{n \times n}$ ma n różnych wartości własnych $r_1, r_2, \ldots, r_n$, za macierz C można wziąć macierz, której kolumnami są kolejno wektory własne o wartościach własnych $r_1, r_2, \ldots, r_n$. A co jeśli $A \in \mathbf{M}_{n \times n}$ nie ma n różnych wartości własnych? O tym w następnym artykule o potęgowaniu macierzy i twierdzeniu Jordana.

Zauważmy na koniec, że mnożenie macierzy motywowane jest następującą podstawową interpretacją geometryczną. Oznaczmy przez $\mathbb{R}^n$ zbiór wszystkich n-wyrazowych ciągów $(x_1, \ldots, x_n)$ liczb rzeczywistych. Przestrzenie $\mathbb{R}^2$ i $\mathbb{R}^3$ to dobrze znane ze szkoły płaszczyzna i przestrzeń 3-wymiarowa. Funkcję $f : \mathbb{R}^n \to \mathbb{R}^m$ nazywamy przekształceniem liniowym, jeśli można ją zapisać wzorem $f(x_1, \ldots, x_n) =$ $= (a_{11}x_1 + \ldots + a_{1n}x_n,\ a_{21}x_1 + \ldots + a_{2n}x_n, \ldots$ $\ldots, a_{m1}x_1 + \ldots + a_{mn}x_n)$ dla pewnej macierzy $A = [a_{ij}] \in \mathbf{M}_{m \times n}$. Na przykład funkcja $f : \mathbb{R}^2 \to \mathbb{R}^2$ zadana wzorem $f(x_1, x_2) = (-x_1 + 2x_2, -6x_1 + 6x_2)$ jest przekształceniem liniowym o macierzy $A = \begin{bmatrix} -1 & 2 \\ -6 & 6 \end{bmatrix} \in \mathbf{M}_{2 \times 2}$. Czytelnik łatwo sprawdzi, że złożenie $g \circ f : \mathbb{R}^k \to \mathbb{R}^m$ przekształcenia liniowego $f : \mathbb{R}^k \to \mathbb{R}^n$ o macierzy $B \in \mathbf{M}_{n \times k}$ z przekształceniem liniowym $g : \mathbb{R}^n \to \mathbb{R}^m$ o macierzy $A \in \mathbf{M}_{m \times n}$ jest przekształceniem liniowym o macierzy $A \cdot B$. Przekształcenia liniowe odgrywają kluczową rolę w wielu zastosowaniach matematyki.

Macierze oczami informatyka

Jakub RADOSZEWSKI

Jak informatyk potęguje macierze?

Poza wyjątkowymi sytuacjami, w których potrzebna jest zwarta postać wzoru na n-tą potęgę macierzy A (wymiaru d), do obliczenia A^n informatycy stosują zazwyczaj o wiele mniej skomplikowaną metodę niż wyznaczanie wartości własnych – szybkie potęgowanie binarne. Jest to algorytm rekurencyjny, wynikający wprost z poniższych wzorów:

$$A^{2k} = (A^k)^2, \qquad A^{2k+1} = (A^k)^2 \cdot A.$$

Innymi słowy, aby wyznaczyć A^{2k} (A^{2k+1}) dla $k \in \mathbb{Z}_+$, obliczamy rekurencyjnie A^k, po czym wartość A^{2k} (A^{2k+1}) otrzymujemy już za pomocą jednego mnożenia bądź dwóch. Złożoność czasowa tej metody, zakładając najprostszy możliwy algorytm mnożenia macierzy, to $O(d^3 \log n)$.

A o tym, po co informatycy potęgują macierze, opowiem pokrótce w kolejnych paragrafach.

Ciągi zdefiniowane rekurencyjnie

Załóżmy, że dany jest ciąg $(a_i)_{i=1}^{\infty}$, zdefiniowany za pomocą następującej liniowej zależności rekurencyjnej:

$$a_1 = 1, \quad a_2 = 1, \quad a_3 = 2, \quad a_n = 2a_{n-1} + 3a_{n-2} + a_{n-3} + 4.$$

Wyznaczenie n-tego wyrazu tego ciągu (powiedzmy, modulo pewna nieduża liczba M) w złożoności czasowej $O(n)$ to dla kogoś obeznanego z programowaniem dynamicznym bułka z masłem. Ale jak obliczyć bilionowy czy trylionowy wyraz tego ciągu? Przyjrzyjmy się wektorowi, który zawiera pewne trzy kolejne wyrazy ciągu a_i oraz stałą 4 (tę z zależności rekurencyjnej):

$$v_n = \begin{bmatrix} a_n \\ a_{n-1} \\ a_{n-2} \\ 4 \end{bmatrix}.$$