Szymon Toruńczyk

1 Wartości własne oraz wektory własne macierzy

Niech A będzie kwadratową macierzą $n \times n$. Wówczas A wyznacza przekształcenie liniowe przestrzeni $\mathbb{R}^n$ w siebie. Niech $v \in \mathbb{R}^n$ będzie pewnym niezerowym wektorem oraz niech $L = \{t \cdot v : t \in \mathbb{R}\}$ będzie prostą wyznaczoną przez ten wektor.

Definicja. Jeżeli przekształcenie A przekształca prostą L w siebie, to mówimy, że v jest *wektorem własnym* przekształcenia A. Oznacza to, że

$$A \cdot v = \lambda \cdot v$$

dla pewnej liczby rzeczywistej λ, zwanej *wartością własną* związaną z wektorem własnym v.

Przykład 1. Przekształcenie A o macierzy $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ przekształca każdy wektor v na $0 = 0 \cdot v$, a zatem każdy wektor v jest wektorym własnym macierzy A o wartości własnej 0.

Przykład 2. Przekształcenie A o macierzy $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ przekształca każdą z osi bazowych w siebie, zatem odpowiadają one wektorom własnym. Tak więc, wektory $(1,0,0), (0,1,0)$ oraz $(0,0,1)$ są wektorami własnymi macierzy A. Odpowiadające im wartości własne to $1, 1, 5$. Ale także wektor $(1,1,0)$ jest wektorem własnym o wartości własnej 1.

Przykład 3. Obrót o kąt α jest przekształceniem o macierzy

$$R_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Jeżeli α nie jest całkowitą wielokrotnością liczby π, to obrót o kąt α nie zachowuje żadnej prostej w $\mathbb{R}^2$ (bo obraca ją), więc R_α nie ma (rzeczywistych) wartości własnych.

Załóżmy, że wektor v jest wektorem własnym macierzy A o wartości własnej λ, tzn.

$$A \cdot v - \lambda \cdot v = 0. \tag{1}$$

Niech I będzie macierzą identycznościową, tzn. taką, która ma jedynki na przekątnej i zera poza nią. Wektor $\lambda \cdot v$ możemy również przedstawić jako iloczyn macierzy $\lambda \cdot I$ przez wektor v. Macierz $\lambda \cdot I$ to jest macierz mająca λ na przekątnej i zera poza nią.

A zatem, równanie (1) można równoważnie zapisać tak:

$$(A - \lambda \cdot I) \cdot v = 0. \tag{2}$$

Niech A_λ oznacza macierz $A - \lambda \cdot I$. Powstaje ona z macierzy A przez odjęcie wartości λ na przekątnej.

Równanie 2 mówi, że $A_\lambda \cdot v = 0$, tzn. v leży w jądrze A_λ. Ponieważ założyliśmy, że $v \neq 0$, to oznacza to, że A_λ nie jest różnowartościowe, więc nie jest bijekcją. Innymi słowy, stąd wynika, że $\det A_\lambda = 0$, tzn. macierz A_λ jest zdegenerowana. Funkcję $w_A(\lambda) = \det A_\lambda$ nazywa się *wielomianem charakterystycznym* macierzy A (faktycznie jest to wielomian). Ponieważ $\det A_\lambda = 0$ wtedy i tylko wtedy, gdy istnieje niezerowy wektor v taki, że $A_\lambda \cdot v = 0$, otrzymujemy następujący wniosek.

Wniosek. λ jest wartością własną macierzy A wtedy i tylko wtedy, gdy $\det A_\lambda = 0$, czyli gdy λ jest pierwiastkiem wielomianu charakterystycznego w_A.

Przykład 4. Dla macierzy A z drugiego przykładu, macierz A_λ wynosi

$$\begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{pmatrix}.$$

Wtedy wyznacznik macierzy A_λ to $(1-\lambda)(1-\lambda)(5-\lambda)$. A zatem,

$$w_A(\lambda) = (1-\lambda)(1-\lambda)(5-\lambda).$$

Pierwiastkami tego wielomianu są liczby 1 oraz 5. A zatem, wartości własne macierzy A to 1 oraz 5, zgodnie z obserwacją w drugim przykładzie.

Przykład 5. Niech R będzie macierzą obrotu o ustalony kąt α, tak jak w przykładzie 3. Wówczas

$$\begin{aligned} w_R(\lambda) &= \det \begin{pmatrix} \cos\alpha - \lambda & \sin\alpha \\ -\sin\alpha & \cos\alpha - \lambda \end{pmatrix} \\ &= \cos^2(\alpha) - 2\cos\alpha \cdot \lambda + \lambda^2 + \sin^2\alpha \\ &= 1 - 2\cos\alpha \cdot \lambda + \lambda^2. \end{aligned}$$

Łatwo sprawdzić, że ten wielomian kwadratowy ma rzeczywiste pierwiastki wtedy i tylko wtedy, gdy $\cos\alpha = \pm 1$, czyli gdy α jest wielokrotnością π. Jest to zgodne z obserwacją w przykładzie 3.

Wielomian charakterystyczny w_A jest wielomianem stopnia n (gdzie n to rozmiar macierzy A). Zatem może on mieć najwyżej n rzeczywistych pierwiastków. Tak więc, macierz A ma najwyżej n wartości własnych. Jak widać na poprzednim przykładzie, tych wartości własnych może być mniej. Jednak dla niektórych macierzy, wszystkie wartości własne są rzeczywiste, tak jak widzieliśmy w przykładzie 3.

Twierdzenie 1. *Jeżeli macierz A jest symetryczna (tzn. $A = A^T$) to wszystkie pierwiastki wielomianu charakterystycznego w_A są rzeczywiste, tzn. A ma n wartości własnych (licząc z krotnościami).*

Przykład 6. Niech

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wtedy $w_A(\lambda) = (2-\lambda)^2 - 1$. Pierwiastki w_A to $\lambda = 1$ oraz $\lambda = 3$, zatem A ma wartości własne 1 i 3. Teraz możemy poszukać wektorów własnych związanych z tymi wartościami własnymi.

Dla $\lambda = 1$, istnieje wektor $v \neq 0$ taki, że $A_\lambda \cdot v = 0$ i on jest właśnie wektorem własnym o wartości własnej 1. A zatem, v jest taki, że

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot v = 0.$$

Przykładowy wektor który spełnia tę zależność to wektor $(1, -1)$. Jest on więc wektorem własnym o wartości własnej 1.

Podobnie, dla wartości własnej 3, wektor własny spełnia zależność

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \cdot v = 0.$$

Takim wektorem jest na przykład wektor $(1, 1)$.

Dla $k = 1, 2, \ldots, n$, niech A_k oznacza macierz A ograniczoną do pierwszych k kolumn oraz wierszy.

Twierdzenie 2. *Niech A będzie macierzą symetryczną. Następujące warunki są równoważne.*

1. $v^T \cdot A \cdot v > 0$ *dla* $v \neq 0$

2. $\det A_k > 0$ *dla* $k = 1, 2, \dots, n$ *(*kryterium Sylvestera*)*

3. *A ma dodatnie wartości własne*

4. w_A *ma dodatnie pierwiastki.*

Definicja. Gdy symetryczna macierz A spełnia któryś z powyższych warunków, to mówimy, że A jest *dodatnio określona.* Jeżeli A jest taka, że $-A$ jest dodatnie określona, to mówimy, że A jest *ujemnie określona.*

Twierdzenie 3. *Niech A będzie macierzą symetryczną. Następujące warunki są równoważne.*

1. *A jest ujemnie określona*

2. $v^T \cdot A \cdot v < 0$ *dla* $v \neq 0$

3. $(-1)^k \det A_k > 0$ *dla* $k = 1, 2, \dots, n$

4. *A ma ujemne wartości własne*

5. w_A *ma ujemne pierwiastki.*

Uwaga. W punkcie 3 powyższego kryterium pojawia się znak $(-1)^k$. Wynika to z tego, że jeśli B jest macierzą $k \times k$, to $\det(-B) = (-1)^k \det B$.

Definicja. Jeżeli symetryczna macierz A ma nieujemne (odpowiednio, niedodatnie) wartości własne, to mówimy, że jest *dodatnio półokreślona* (odpowiednie, *ujemnie półokreślona*).

Uwaga. Kryterium Sylvestera *nie zachodzi* dla macierzy półokreślonych tzn. nie jest prawdą, że jeśli $\det A_k \geq 0$ dla $k = 1, 2, \dots, n$ to A ma nieujemne wartości własne to A jest dodatnio półokreślona.

Definicja. Macierz A nazywa się *nieokreśloną* jeżeli ma zarówno dodatnie, jak i ujemne wartości własne. A zatem, A jest nieokreślona, jeśli nie jest ani dodatnio ani ujemnie półokreślona.

Przykład 7.

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wyznaczniki minorów głównych to 0 oraz 2, ale ta macierz nie jest dodatnio półokreślona, ponieważ jej wartości własne to $1 + \sqrt{2}$ oraz $1 - \sqrt{2}$ i są one różnych znaków. A zatem macierz A jest nieokreślona.

Uwaga. Wartości własne macierzy grają kluczową rolę w mechanice kwantowej, teorii spektralnej grafów, teorii łańcuchów Markowa, algorytmie PageRank (Google), algorytmach kwantowych, algorytmach randomizacyjnych (poprzez związki z teorią spektralną grafów), sztucznej inteligencji (np. rozpoznawanie twarzy) i wielu innych. Tłumaczą one takie zjawiska, jak rozchodzenie się fal oraz np. czyste tony (dzwięki).

2 Analiza puktów krytycznych funkcji różniczkowalnej

Niech $f\colon \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathcal{C}^2$ na $\mathbb{R}^n$. Niech p będzie punktem krytycznym funkcji $f\colon \mathbb{R}^n \to \mathbb{R}$, tzn $\mathrm{grad} f(p) = 0$, lub równoważnie, $Df(p) = 0$.

Chcemy przeanalizować zachowanie funkcji f w okolicy punktu p.

W tym celu badamy *Hessian* funkcji f w punkcie p, czyli macierz, która na pozycji (i, j) ma wartość $\frac{\partial^2 f}{\partial x_i \partial x_j}$. Macierz ta oznaczać będziemy symbolem $Hf(p)$.

Twierdzenie 4. *Jeżeli f jest klasy $\mathcal{C}^2$ w otoczeniu punktu p, to macierz $Hf(p)$ jest symetryczna.*

Analiza wartości własnych oraz wektorów własnych macierzy $Hf(p)$ mówi nam o zachowaniu funkcji f w okolicy punktu p. Intuicyjnie, jeżeli wektor v jest wektorem własnym macierzy $Hf(p)$ o wartości własnej λ, to znaczy to, że przyrost f w kierunku v jest rzędu $\frac{1}{2}\lambda \cdot t^2$ (wyrazów rzędu t, czyli liniowych, nie ma, ponieważ p jest punktem spłaszczenia f.)

W szczególności, jeżeli we wszystkie strony przyrost jest dodatni, tzn. $Hf(p)$ ma dodatnie wartości własne, to f ma minimum lokalne w punkcie p, a jeśli $Hf(p)$ ma ujemne wartości własne, to f ma maksimum lokalne w punkcie p. Jeżeli zaś $Hf(p)$ ma zarówno dodatnie jak i ujemne wartości własne, to w jednym z kierunków, f rośnie, a w drugim maleje, więc f nie ma ektremum lokalnego w p.

A zatem, mamy następujące twierdzenie.

Twierdzenie 5 (Kryterium drugiego rzędu istnienia ekstremów)**.** *Załóżmy, że $f\colon \mathbb{R}^n \to \mathbb{R}$ jest klasy $\mathcal{C}^2$ w otoczeniu punktu p oraz, że $\mathrm{grad} f(p) = 0$. Wówczas:*

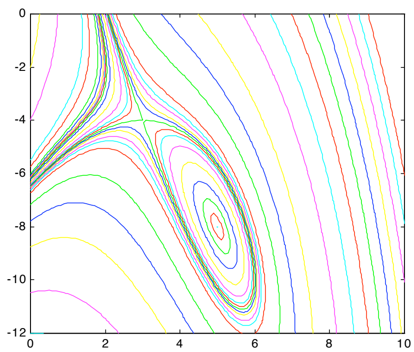
- *Jeżeli $Hf(p)$ jest dodatnio określona, to f ma minimum lokalne w p*
- *Jeżeli $Hf(p)$ jest ujemnie określona, to f ma maksimum lokalne w p*

- *Jeżeli $Hf(p)$ jest nieokreślona, to f nie ma ekstremum lokalnego w p.*

Przykład 8. Rozpatrzmy funkcję $f\colon \mathbb{R}^2 \to \mathbb{R}$ zadaną wzorem

$$f(x,y) = \frac{1}{3}x^3 - 3x^2 + \frac{y^2}{4} + xy + 13x - y + 2.$$

Ma ona dwa punkty krytyczne, $(3,-4)$ oraz $(5,-8)$, z wartościami f w tych punktach, odpowiednio, 19 oraz $17\frac{2}{3}$. Wykres konturowy funkcji f wygląda tak:



Widać z wykresu, że $(5,-8)$ jest ekstremum lokalnym a $(3,-4)$ jest punktem siodłowym. Potwierdzamy to analizą Hessianu f który w punkcie (x,y) wynosi $Hf(x,y) = \begin{pmatrix} 2x-6 & 1 \\ 1 & \frac{1}{2} \end{pmatrix}$.

W punkcie $(5,-8)$, Hessian wynosi wynosi $\begin{pmatrix} 4 & 1 \\ 1 & \frac{1}{2} \end{pmatrix}$. Macierz ta jest dodatnio określona (wyznaczniki minorów głównych wynoszą 4 oraz 1, więc są dodatnie), więc $(5,-8)$ jest minimum lokalnym funkcji f.

Możemy również obliczyć, że wartości własne tej macierzy to $1/4(9+\sqrt{65})$ oraz $1/4(9-\sqrt{65})$, a odpowiadające im wektory własne to $(-1/2+1/4(9+\sqrt{65}),1)$ oraz $(-1/2+1/4(9-\sqrt{65},1)$. Można stąd wywnioskować, że funkcja f najszybciej rośnie (względem punktu $(5,-8)$) w kierunku pierwszego z tych wektorów (gdyż pierwsza wartość własna jest większa).

W punkcie $(3,-4)$, Hessian wynosi $B = \begin{pmatrix} 0 & 1 \\ 1 & \frac{1}{2} \end{pmatrix}$.

Ta macierz nie jest ani dodatnio, ani ujemnie określona, więc pozostaje nam stwierdzić, czy jest ona nieokreślona.

Wartości własne macierzy B to $\frac{1}{4}\left(1+\sqrt{17}\right)$ oraz $\frac{1}{4}\left(1-\sqrt{17}\right)$ i są one różnych znaków, więc $(5,-8)$ nie jest ekstremum lokalnym funkcji f.

Ponadto, wektory własne macierzy B to $(-1/2 + 1/4(1 + \sqrt{17}), 1)$ oraz $(-1/2 + 1/4(1 - \sqrt{17}), 1)$. A zatem, f najszybciej rośnie/maleje w tych kierunkach.