OBLICZANIE exp OD MACIERZY

MACIEJ BORODZIK

Streszczenie. Te notatki zawierają w miarę ścisłe i, mam nadzieję, również w miarę przejrzyste wyjaśnienie, dlaczego działa algorytm wyliczania e do macierzy. W pierwszym kroku wyjaśniamy, dlaczego działa dla macierzy o różnych wartościach własnych, drugi pokazuje, że takie macierze są gęste. Trzy następne kroki pokazują, że współczynniki są ciągłe jako funkcje o wartości własnych. W ostatnim kroku rozprawiamy się z problemem, że wartości własne macierzy A nie zależą w sposób ciągły od macierzy.

To jest pierwsza wersja notatek, może zawierać pewne literówki, bądź niedopowiedzenia. Będę wdzięczny za konstruktywne uwagi.

Wprowadzenie

Niech A będzie macierzą o wymiarach $n\times n$ i wartościach własnych $\lambda_1, \ldots, \lambda_k$, przy czym krotność wartości własnej λ_i jest równa m_i. Mamy więc $\sum m_i = n$.

Rozważmy wielomian

$$P(x) = a_1 x^{n-1} + \cdots + a_n,$$

gdzie a_i są tak dobrane, aby zachodziło:

$$\begin{array}{ccccc} P(\lambda_1) = e^{\lambda_1} & \ldots & \frac{\partial}{\partial\lambda} P|_{\lambda=\lambda_1} = e^{\lambda_1} & \ldots & \frac{\partial^{m_1-1}}{\partial\lambda^{m_1-1}} P|_{\lambda=\lambda_1} = e^{\lambda_1} \\ \ldots & & \ldots & & \ldots \\ P(\lambda_k) = e^{\lambda_k} & \ldots & \frac{\partial}{\partial\lambda} P|_{\lambda=\lambda_k} = e^{\lambda_k} & \ldots & \frac{\partial^{m_k-1}}{\partial\lambda^{m_k-1}} P|_{\lambda=\lambda_k} = e^{\lambda_1} \end{array} \tag{1}$$

Twierdzimy, że wtedy

$$e^A = a_1 A^{n-1} + a_2 A^{n-2} + \cdots + a_{n-1} A + a_n I. \tag{2}$$

Celem tych krótkich notatek jest przedstawienie skomplikowanego dowodu tego faktu.

Krok 1. Macierze diagonalne i diagonalizowalne

Dla macierzy diagonalnej o wartościach własnych $\lambda_1, \ldots, \lambda_n$ mamy

$$e^A = \begin{pmatrix} e^{\lambda_1} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2} & \ldots & 0 \\ \ldots & & & \\ 0 & 0 & \ldots & e^{\lambda_n} \end{pmatrix}$$

Z drugiej strony,

$$a_1A^{n-1}+\cdots+a_nI=\begin{pmatrix} P(\lambda_1) & 0 & \dots & 0\\ 0 & P(\lambda_2) & \dots & 0\\ \dots & & & \\ 0 & 0 & \dots & P(\lambda_n)\end{pmatrix}$$

Oznacza to, że jeśli tylko spełnione jest wyjściowe równanie — które w tym przypadku ma postać — $e^{\lambda_i}=P(\lambda_i)$ (pod warunkiem, że wartości własne są różne) — to obie strony równania (2) się zgadzają.

Dla macierzy diagonalizowalnych sytuacja jest podobna: obserwujemy, że jeśli dla macierzy A zachodzi (2), to dla dowolnego sprzężenia $B=CAC^{-1}$, też zachodzi (2). Wynika to z faktu, że $e^{CAC^{-1}}=Ce^AC^{-1}$.

Oznacza to, że teza zgadza się na podzbiorze macierzy diagonalizowalnych.

Krok 2. Gęstość macierzy diagonalizowalnych

W tym kroku pokażemy, że macierze diagonalizowalne są gęste.

Lemat 1. *Jeśli macierz A ma parami różne wartości własne, to jest diagonalizowalna.*

Dowód lematu, na poziomie szkoły ponadgimnazjalnej, pomijamy.

Twierdzenie 2. *Podzbiór macierzy takich, że ich wielomian charakterystyczny ma pierwiastek podwójny jest brzegowy (tj. ma puste wnętrze) i domknięty.*

Twierdzenie 2 będzie wynikało z dwóch następnych lematów.

Lemat 3. *Niech $Q(x_1,\dots,x_n)$ będzie niestałym wielomianem od n–zmiennych. Wtedy zbiór $\{Q=0\}$ jest domknięty i brzegowy.*

Lemat 4. *Istnieje taki wielomian $H(a_{11},\dots,a_{nn})$, że macierz $A=\{a_{ij}\}_{i,j=1}^n$ ma podwójną wartość własną wtedy i tylko wtedy, gdy $H(a_{11},\dots,a_{nn})=0$.*

Dowód Lematu 3. Domkniętość zbioru $\{Q=0\}$ wynika z ciągłości funkcji wielomianowej. Przypuśćmy, że istnieje zbiór otwarty U taki, że $Q|_U\equiv 0$. Wybierzmy $x_0\in U$. Niech $v\in\mathbb{R}^n$. Wtedy funckja $Q_v(t)=Q(x_0+tv)$ jest wielomianem jednej zmiennej, który zeruje się w pewnym otoczeniu 0. Stąd $Q_v\equiv 0$, czyli Q jest równy zero na każdej prostej, która przechodzi przez x_0. Stąd $Q\equiv 0$. □

Dowód Lematu 4. Po pierwsze współczynniki wielomianu charakterystycznego $p_1,\dots,p_n$ są w oczywisty sposób wielomianami od współczynników macierzy. Niech więc

$$C(t)=t^n+p_1t^{n-1}+\cdots+p_n$$

będzie wielomianem charakterystycznym macierzy. Wtedy C ma pierwiastek podwójny wtedy i tylko wtedy, gdy $NWD(C(t),C'(t))\neq 1$. Dla dwóch wielomianów

$$a(t)=a_rt^r+a_{r-1}t^{r-1}+\cdots+a_0$$
$$b(t)=b_st^s+b_{s-1}t^{s-1}+\cdots+b_0$$

określamy przekształcenie $M : W_{s-1} \times W_{r-1} \to W_{s+r-1}$, gdzie W_k jest przestrzenią liniową wielomianów od t stopnia co najwyżej k. Przekształcenie M zadane jest wzorem

$$M(\phi, \psi) = a \cdot \phi + b \cdot \psi.$$

M jest przekształceniem liniowym między przestrzeniami liniowymi tego samego wymiaru (mianowicie $r + s$). Ponadto, przy wyborze standardowej bazy w przestrzeniach wielomianów, macierz przekształcenia M zapisuje się wzorem

$$\begin{pmatrix} a_0 & 0 & 0 & 0 & \dots & 0 & 0 & b_0 & 0 & \dots & 0 \\ a_1 & a_0 & 0 & 0 & \dots & 0 & 0 & b_1 & b_0 & \dots & 0 \\ a_2 & a_1 & a_0 & 0 & \dots & 0 & 0 & b_2 & b_1 & \dots & 0 \\ a_3 & a_2 & a_1 & a_0 & \dots & 0 & 0 & b_3 & b_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a_r & a_{r-1} & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & a_r & 0 & 0 & \dots & b_s \end{pmatrix}$$

Wyznacznik tej macierzy nazywa się *rezultantą* wielomianów a, b i oznacza się go często $r(a, b)$.

Obserwujemy, że $r(a, b) = 0$ wtedy i tylko wtedy, gdy M nie jest "na". To zaś oznacza, że wielomiany a i b mają nietrywialny wspólny dzielnik.

Rezultantę wielomianów C i C' nazywa się często *dyskryminantą* wielomianu C i oznacza $d(C)$ (z dokładnością do stałej). Widzimy, że $d(C)$ jest wielomianem od współczynników wielomianu C. □

Tak więc twierdzenie jest udowodnione.

Krok 3. Ciągłość współczynników wielomianu P. Bez degeneracji.

Aby pokazać, że (2) zachodzi dla wszystkich macierzy A bez wyjątku, chcemy zaobserwować, że obie strony równości zmieniają się w sposób ciągły, jeśli zmieniamy macierz A. Jako, że wiemy, że zbiór, na którym (2) zachodzi, jest gęsty (na podstawie kroków 1 i 2), wystarczy to do dowodu głównej równości.

W tym kroku założymy, że ciąg $\lambda^{(k)} = (\lambda_1^{(m)}, \dots, \lambda_k^{(m)}$ zbiega do $\lambda^{(0)} = (\lambda_1^0, \dots, \lambda_k^0)$, ale $m_1, \dots, m_k$ pozostają stałe, w szczególności $\lambda_i^0 \neq \lambda_j^0$ dla $i \neq j$.

Twierdzenie 5. *Jeśli* $\mathbf{a}^{(k)} = (a_1^{(k)}, \dots, a_n^{(k)})$ *są rozwiązaniami równania* (1) *dla* $\lambda^{(k)}$, *to* $\mathbf{a}^{(k)} \to \mathbf{a}^{(0)}$.

Dowód. Rozpiszmy porządnie układ równań (1)

$$
\begin{aligned}
e^{\lambda_1} &= a_1\lambda_1^{n-1} + a_2\lambda_1^{n-2} + \dots + a_n \\
e^{\lambda_1} &= (n-1)a_1\lambda_1^{n-2} + (n-2)a_2\lambda_1^{n-3} + \dots + a_{n-1} \\
\dots \\
e^{\lambda_1} &= \frac{(n-1)!}{(n-m_1)!}a_1\lambda_1^{n-m_1-1} + \frac{(n-2)!}{(n-m_1-1)!}a_2\lambda_1^{n-m_1-2} + \dots + \\
&+ (m_1-1)!a_{n-m_1+1} \\
\dots \\
e^{\lambda_k} &= \frac{(n-1)!}{(n-m_k)!}a_1\lambda_k^{n-m_k-1} + \frac{(n-2)!}{(n-m_k-1)!}a_2\lambda_1^{n-m_k-2} + \dots + \\
&+ (m_k-1)!a_{n-m_k+1}.
\end{aligned}
\tag{3}
$$

Na mocy twierdzenia o funkcji uwikłanej, równania (1) definiują funkcje $a_1, \dots, a_n$ jako funkcje różniczkowalne od $\lambda_1, \dots, \lambda_k$ wtedy i tylko wtedy, gdy wyznacznik macierzy $\mathcal{M}$

$$
\begin{pmatrix}
\lambda_1^{n-1} & \lambda_1^{n-2} & \lambda_1^{n-3} & \dots & \lambda_1 & 1 \\
(n-1)\lambda_1^{n-2} & (n-2)\lambda_1^{n-3} & (n-3)\lambda_1^{n-4} & \dots & 1 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \\
\frac{(n-1)!}{(n-m_1)!}\lambda_1^{n-m_1} & \frac{(n-2)!}{(n-m_1-1)!}\lambda_1^{n-m_1-1} & \frac{(n-3)!}{(n-m_1-2)!}\lambda_1^{n-m_1-2} & \dots & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \\
\lambda_k^{n-1} & \lambda_k^{n-2} & \lambda_k^{n-3} & \dots & \lambda_k & 1 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \\
\frac{(n-1)!}{(n-m_k)!}\lambda_k^{n-m_k} & \frac{(n-2)!}{(n-m_k-1)!}\lambda_k^{n-m_k-1} & \frac{(n-3)!}{(n-m_k-2)!}\lambda_k^{n-m_k-2} & \dots & 0 & 0
\end{pmatrix}
$$

jest różny od zera. Dla $m_1 = \dots = m_k = 1$, macierz $\mathcal{M}$ jest macierzą Vandermonde'a i jej wyznacznik jest równy $\prod_{i<j}(\lambda_i - \lambda_j)$. W szczególności jest on różny od zera, jeśli wartości własne λ_i są parami różne. W ogólności wyznacznik $\det \mathcal{M}$ można policzyć przez indukcję po m_i.

Lemat 6. *Wyznacznik macierzy $\mathcal{M}$ jest proporcjonalny do iloczynu $\prod_{i<j}(\lambda_i - \lambda_j)^{m_i m_j}$ z wymierną stałą proporcjonalności.*

Dowód Lematu 6. Przypuśćmy, że $m_1, \dots, m_k$ są poszeregowane w kolejności malejącej, oraz dla $i > l$ zachodzi $m_i = 1$. Pokażemy, że jeśli teza lematu zachodzi dla $(m_1, \dots, m_l, 1, \dots, 1)$, to zachodzi również dla $(m_1, \dots, m_j + 1, \dots, m_l, 1, \dots, 1)$, gdzie ilość jedynek jest zmniejszona. Odpowiada to sytuacji, gdy $\lambda_k \to \lambda_j$. Dla ustalenia uwagi przyjmiemy, że $j = 1$.

Niech v_i będzie wektorem poziomym $(\lambda_i^{n-1}, \lambda_i^{n-2}, \dots, \lambda_i, 1)$. Ponadto, niech $v_{ij} = \frac{d^j}{d\lambda_i^j} v_i$. Macierz $\mathcal{M}$ możemy zapisać jako

$$\begin{pmatrix} v_{10} \\ v_{11} \\ \dots \\ v_{1,m_1-1} \\ \dots \\ v_{k0}. \end{pmatrix} \tag{4}$$

(Przypomnijmy, że przyjęliśmy, iż $m_k = 1$.)

Piszemy teraz $\lambda_k = \lambda_1 + h$. Wtedy

$$v_{k0} = v_{10} + hv_{11} + \frac{h^2}{2!} v_{12} + \dots + \frac{h^{m_1}}{m_1!} v_{1m_1} + O(h^{m_1+1}). \tag{5}$$

Wstawiając (5) jako ostatni wiersz i korzystając z faktu, że wyznacznik nie ulega zmianie przy dodaniu do wiersza kombinacji liniowej innych wierszy, uzyskujemy

$$\det \mathcal{M} = \frac{h^{m_1}}{m_1!} \det \begin{pmatrix} v_{10} \\ v_{11} \\ \dots \\ v_{1,m_1-1} \\ \dots \\ v_{1,m_1} \end{pmatrix} + O(h^{m_1+1}). \tag{6}$$

Skoro teraz $\mathcal{M}$ jest proporcjonalny do

$$\prod_{i<j<k} (\lambda_i - \lambda_j)^{m_i m_j} \cdot \prod_{i=2}^{k-1} (\lambda_i - \lambda_k)^{m_i} \cdot (\lambda_1 - \lambda_k)^{m_1},$$

wyznacznik macierzy

$$\begin{pmatrix} v_{10} \\ v_{11} \\ \dots \\ v_{1,m_1-1} \\ \dots \\ v_{1,m_1} \end{pmatrix}$$

jest proporcjonalny do

$$\prod_{i<j<k} (\lambda_i - \lambda_j)^{m_i m_j} \cdot \prod_{i=2}^{k-1} (\lambda_i - \lambda_1 - h)^{m_i},$$

z dokładnością do wyrazów $O(h)$. Przejście z h do zera kończy dowód kroku indukcyjnego. $\square$

Lemat 6 kończy dowód twierdzenia. $\square$

Uwaga 7. Z dowodu wynika, że dla dowolnego $R > 0$, na zbiorze $|\lambda_i - \lambda_j| < \frac{1}{R}$, $|\lambda_i| < R$, funkcje $a_1, \dots, a_n$ są lipschitzowskie. Istotnie, współczynniki równania (3) są ograniczone z góry, zaś wyznacznik układu jest oddzielony od zera na mocy Lematu 6.

Krok 4. Degeneracja równań

W tej części udowodnimy następujące

Twierdzenie 8. *Niech $\lambda_1^{(i)}, \dots, \lambda_k^{(i)}$ będzie ciągiem zadanym przez*

$$\begin{aligned} \lambda_1^{(i)} &= \lambda_1 \\ &\dots \\ \lambda_{k-1}^{(i)} &= \lambda_{k-1} \\ \lambda_k^{(i)} &= \lambda_1 + h, \end{aligned}$$

gdzie $h = h_i$ zbiega do zera. Wtedy $\mathbf{a}^{(i)}$ (czyli rozwiązanie układu (1) dla $\lambda^{(i)}$) dąży do $\mathbf{a}$, gdzie $\mathbf{a}$ jest rozwiązaniem układu (1) z parametrami $\lambda_1, \dots, \lambda_{k-1}$ i krotnościami $m_1 + m_k, m_2, \dots, m_{k-1}$. Ponadto, jeśli dla $1 \leq i \neq j \leq k-1$ zachodzi $|\lambda_i - \lambda_j| > 1/R$ oraz $|\lambda_i| < R$, to $||\mathbf{a}^{(i)} - \mathbf{a}|| = O(h_i)$, czyli zbieżność jest jednostajna po zbiorach ograniczonych, jeśli λ_i są rozdzielone.

Dowód. Dla uproszczenia przyjmijmy najpierw, że $m_k = 1$. Niech tak, jak poprzednio $v_i = (\lambda_i^{n-1}, \dots, \lambda_i, 1)$ oraz $v_{ij} = \frac{d^j}{d\lambda_i^j} v_i$. Mamy

$$v_{k0} = v_{10} + h v_{11} + \frac{h^2}{2!} v_{12} + \dots + \frac{h^{m_1}}{m_1!} v_{1m_1} + h^{m_1+1} \cdot \text{coś}, \tag{7}$$

gdzie wyraz „coś” szacuje się przez supremum $(m_1 + 1)-$szej pochodnej v_1 w pewnym otoczeniu λ_1.

Z drugiej strony możemy napisać, że

$$e_k^{\lambda} = e^{\lambda_1 + h} = e^{\lambda_1} + h e^{\lambda_1} + \frac{h^2}{2!} e^{\lambda_1} + \dots + \frac{h^{m_1}}{m_1!} e^{\lambda_1} + h^{m_1+1} \cdot \text{coś}. \tag{8}$$

Ponownie wyraz „coś” możemy doszacować przez supremum $(m_1 + 1)-$szej pochodnej funkcji e^{λ_1}. Wstawiając (7) i (8) do ostatniego z równań układu (3) otrzymujemy ($\mathbf{a} = (a_1, \dots, a_n)$ a $v \cdot \mathbf{a}$ oznacza iloczyn skalarny, co znakomicie upraszcza zapis).

$$\begin{aligned} \left(e^{\lambda_1} - v_{10} \cdot \mathbf{a}\right) + h\left(e^{\lambda_1} - v_{11} \cdot \mathbf{a}\right) + \frac{h^2}{2!}\left(e^{\lambda_1} - v_{12} \cdot \mathbf{a}\right) + \dots + \\ + \frac{h^{m_1-1}}{(m_1-1)!}\left(e^{\lambda_1} - v_{1,m_1-1} \cdot \mathbf{a}\right) + \frac{h^{m_1}}{m_1!}\left(e^{\lambda_1} - v_{1,m_1} \cdot \mathbf{a}\right) = \\ = O(h^{m_1+1}) \cdot \mathbf{a} + O(h^{m_1+1}). \end{aligned} \tag{9}$$

Teraz pierwsze wyrażenie w nawiasie znika na mocy pierwszego równania z serii (3). Dalej, drugi nawias jest zerowy dzięki drugiemu równaniu z układu (3). Postępując w ten sposób sprowadzamy równanie (9) do postaci

$$\frac{h^{m_1}}{m_1!}\left(e^{\lambda_1} - v_{1,m_1}\cdot\mathbf{a}\right) = O(h^{m_1+1})\cdot\mathbf{a} + O(h^{m_1+1}).$$

Dzieląc przez $\frac{h^{m_1}}{m_1!}$ otrzymujemy równanie, które znowu przypomina ostatnie równanie z (3):

$$e^{\lambda_1} + O(h) = (v_{1,m_1} + O(h))\cdot\mathbf{a}.$$

Ponieważ wyznacznik macierzy

$$\begin{pmatrix} v_{10} \\ v_{11} \\ \dots \\ v_{1,m_1-1} \\ v_{20} \\ \dots \\ v_{k-1,m_{k-1}} \\ v_{1,m_1} \end{pmatrix}$$

jest różny od zera na mocy Lematu 6, będzie on nadal niezerowy jeśli do ostatniego wiersza dodamy człon $O(h)$, gdy h jest dostatecznie małe. W związku z tym dla dostatecznie małego h rozwiązania równania

$$\begin{pmatrix} e^{\lambda_1} & = v_{10}\cdot\mathbf{a} \\ e^{\lambda_1} & = v_{11}\cdot\mathbf{a} \\ \dots & \\ e^{\lambda_1} & = v_{1,m_1-1}\cdot\mathbf{a} \\ e^{\lambda_1} & = v_{20}\cdot\mathbf{a} \\ \dots e^{\lambda_1} & = v_{k-1,m_{k-1}-1}\cdot\mathbf{a} \\ e^{\lambda_1} + O(h) & = (v_{1,m_1} + O(h))\cdot\mathbf{a} \end{pmatrix}$$

zależą w sposób ciągły od parametru h. Ponadto, o ile $|\lambda_i - \lambda_j| > 1/R$ oraz $|\lambda_i| < R$, zaś $|h| < \frac{1}{2R}$, różnica pomiędzy rozwiązaniem dla 0 a rozwiązaniem dla h szacuje się przez stałą niezależną od λ pomnożoną przez h.

Jeśli teraz $m_k > 1$, postępujemy podobnie. Musimy tylko rozwijać nie tylko v_{k0}, lecz także v_{ki} dla $i \leq m_k - 1$ w szereg Taylora. Całe rozumowanie jest bardzo podobne, tylko żmudniejsze. $\square$

Krok 5. Ciągłość w zależności od wartości własnych

Twierdzenie 9. *Niech $\lambda^{(i)} \to \lambda$. Wtedy $\mathbf{a}^{(i)} \to \mathbf{a}$. W tym przypadku oznaczamy $\lambda = (\lambda_1, \dots, \lambda_n)$, gdzie wartości własne są od razu wypisane z krotnościami.*

Będziemy starali się zastosować Krok 3 i Krok 4. Do tego celu, niech dla ustalonego rozbicia $I = I_1 \cup \dots \cup I_s$ zbioru $\{1, \dots, n\}$ określimy zbiór

$$K_I = \{\lambda \in \mathbb{C}^n \colon \forall_{t\in\{1,\dots,s\}}, \forall_{i,j\in I_t}\lambda_i = \lambda_j\}.$$

Dla $I = \{1\} \cup \{2\} \cup \ldots \cup \{n\}$ zbiór K_I pokrywa się z całą przestrzenią. W zbiorze rozbić wprowadzamy częściowy porządek $I \leq J$, gdy każdy element rozbicia I zawiera się w pewnym elemencie rozbicia J. W oczywisty sposób, jeśli $I \leq J$, to $K_J \subset K_I$. Istotnie, K_J jest wyznaczane przez więcej warunków.

Oznaczmy jeszcze

$$K_I^0 = K_I \setminus \bigcup_{\substack{J \geq I \\ J \neq I}} K_J.$$

Inaczej $K_I^0 = \{\lambda \in K_I \colon \lambda_i \neq \lambda_j \text{ gdy } i \in I_{t_1},\ j \in I_{t_2} \text{ oraz } t_1 \neq t_2\}$. Zbiory K_I^0 pokrywają całą przestrzeń $\mathbb{C}^n$ oraz jest ich skończenie wiele. Przechodząc do podciągu możemy założyć, że wszystkie $\lambda^{(i)}$ siedzą w tym samym zbiorze K_I^0. Wtedy granica należy do pewnego K_J^0, dla $J > I$.

Jeśli $J = I$, nie ma degeneracji i Krok 3 pokazuje, że $\mathbf{a}(\lambda^{(i)})$ zbiegają do $\mathbf{a}(\lambda)$. Przypuśćmy zatem, że $J > I$. Oznacza to, że istnieją takie pary indeksów $(a_1, b_1), \ldots, (a_l, b_l)$, że $\lambda^{(i)}_{a_s} - \lambda^{(i)}_{b_s}$ zbiega do zera, gdy $i \to \infty$.

Niech R będzie tak dobrane, że $|\lambda_i - \lambda_j| > 1/R$ dla i, j nie należących do tego samego podzbioru podziału J oraz $|\lambda_i| > R$. Możemy takie R dobrać, bo dobieramy je tylko do punktu granicy.

Ustalmy $\varepsilon > 0$. Z faktu, że $\lambda^{(i)} \to \lambda$ wnioskujemy, że dla dostatecznie dużych i wielkość

$$\tilde{\lambda}^{(i)} = (\tilde{\lambda}^{(i)}_1, \ldots, \tilde{\lambda}^{(i)}_n),$$

gdzie

$$\tilde{\lambda}^{(i)}_r = \begin{cases} \lambda^{(i)}_t & \text{gdy } \exists s \colon (t, r) = (a_s, b_s) \\ \lambda^{(i)}_r & \text{w przeciwnym przypadku.} \end{cases}$$

jest bliższa niż ε zarówno $\lambda^{(i)}$ oraz λ. Ponadto z definicji $\tilde{\lambda}^{(i)} \in K_J^0$. Zwiększając być może i możemy założyć, że $|\tilde{\lambda}^{(i)}_k - \tilde{\lambda}^{(i)}_l| > 1/R$ oraz $|\tilde{\lambda}^{(i)}_k| < R$.

Oznacza to, w myśl Uwagi 7,

$$||\mathbf{a}(\lambda) - \mathbf{a}(\tilde{\lambda}^{(i)})|| < \delta$$

dla dowolnie z góry zadanego δ (dobieramy i do δ).

Ponadto, z tego samego powodu (był on umieszczony na końcu Kroku 4), dla dostatecznie dużych i:

$$||\mathbf{a}(\lambda^{(i)}) - \mathbf{a}(\tilde{\lambda}^{(i)})|| < \delta.$$

To oznacza, że

$$||\mathbf{a}(\lambda^{(i)}) - \mathbf{a}(\lambda)|| < 2\delta.$$

A więc prawie mamy ciągłość. Pozostaje jeszcze uzasadnić, że wybierając podciąg zawarty stale w K_I^0 nic nie ograniczamy. Robi się to przez podobne szacowanie.

Krok 6. a_i jako funkcje od macierzy

Wartości własne macierzy A nie są niestety ciągłymi funkcjami od A. Wynika to z tego, że musimy dobrać arbitralnie kolejność wartości własnych. Można to zaobserwować na następującym, nietrywialnym, przykładzie macierzy 2x2 o współczynnikach rzeczywistych.

Niech $v = (1, 0)$ i $w = (0, 1)$. Niech też $\lambda = 1$, $\mu = 2$. Gdy t zmienia się od 0 do 1 wektor v_t robi ćwierć obrotu w lewo i przechodzi na wektor w. Wektor w_t robi trzy czwarte obrotu w lewo i przechodzi na wektor v. Wielkości λ i μ zmieniają się zgodnie ze wzorem $\lambda_t = 1 + t$, $\mu_t = 2 - t^2$ tak, że dla wszystkich $t \in [0, 1]$ $\lambda_t \neq \mu_t$.

Rozważmy przekształcenie $\phi_t : \mathbb{R}^2 \to \mathbb{R}^2$, które przeprowadza wektor v_t na $\lambda_t v_t$ zaś wektor w_t na $\mu_t w_t$. Niech A_t będzie macierzą przekształcenia ϕ_t w bazie standardowej. Oczywiście $A_1 = A_0$, bo ϕ_1 i ϕ_0 zgadzają się na obrazie wektorów bazowych $(1, 0)$ i $(0, 1)$. Ponadto wartościami własnymi macierzy A_t są (λ_t, μ_t). Jeśli chcemy, żeby wartości własne były ciągłe, musimy dla wszystkich t trzymać tę samą kolejność wartości własnych. W tym miejscu korzystamy z tego, że $\lambda_t \neq \mu_t$, a więc „brakuje pretekstu” żeby zamienić wartości własne miejscami. Ale $(\lambda_1, \mu_1) \neq (\lambda_0, \mu_0)$. Po przejściu jednego obrotu wartości własne zmieniły miejsce.

Niemniej jednak przynajmniej dla λ_i parami różnych $\mathbf{a}$ jest funkcją symetryczną od λ_i. Skoro $\mathbf{a}$ jest ciągła i symetryczna na zbiorze gęstym, jest symetryczna w ogóle.

Lemat 10. *Niech $x_1, \dots, x_n$ będą jakimiś współrzędnymi w $\mathbb{C}^n$. Niech p_i będą wielomianami określonymi przez*

$$x^n + p_1 x^{n-1} + \cdots + p_n = (x + x_1)(x + x_2) \cdots (x + x_n).$$

Wtedy dla każdej funkcji symetrycznej $f(x_1, \dots, x_n)$ istnieje dokładnie jedna funkcja $g(p_1, \dots, p_n)$ taka że $f(x_1, \dots, x_n) = g(p(x))$. Ponadto, jeśli f jest ciągła, to g również.

Dowód. Rozważmy relację równoważności $(x_1, \dots, x_n) \sim (y_1, \dots, y_n)$ wtedy, gdy wektory x i y różnią się tylko o permutację indeksów. Odwzorowanie $P\colon (x_1, \dots, x_n) \to (p_1, \dots, p_n)$, $P\colon \mathbb{C}^n \to \mathbb{C}^n$ skleja dwa punkty wtedy i tylko wtedy, gdy są one równoważne. A więc jest to odwzorowanie ilorazowe w sensie teorii mnogości. Inaczej mówiąc, każda funkcja f, która skleja punkty symetryczne zapisuje się jako złożenie g z P.

To samo rozumowanie przechodzi, jeśli zauważymy, że P jest ciągłe jako przekształcenie ilorazowe w sensie topologii. □

W związku z tym współczynniki $\mathbf{a} = (a_1, \dots, a_n)$ zapisują się jednoznacznie jako funkcje od współczynników wielomianu charakterystycznego danej macierzy. Ponadto, w myśl powyższego lematu, są one ciągłe.

Skoro współczynniki wielomianu charakterystycznego macierzy są ciągłymi funkcjami od jej współczynników, $\mathbf{a}$ są ciągłymi funkcjami od macierzy.

Maciej Borodzik