

Metody numeryczne

Zagadnienia własne

Janusz Szwabiński

`szwabin@ift.uni.wroc.pl`

Zagadnienia własne

1. Pojęcia podstawowe
2. Zaburzenia wartości i wektorów własnych
3. Wartości własne macierzy hermitowskiej
4. Metody ogólne dla macierzy o elementach rzeczywistych
 - Metoda Kryłowa
 - Metoda potęgowa
 - Algorytm **QR** dla macierzy Hessenberga
 - Srowadzenie macierzy do postaci Hessenberga
 - Rozkład macierzy na iloczyn **QR**
 - Wektory własne w metodzie **QR**
5. GNU Scientific Library (GSL)

Zagadnienie własne

$$\mathbf{A}x = \lambda x$$

λ - wartość własna

x - wektor własny

Pojęcia podstawowe (1)

Twierdzenie 1 *Liczba λ jest wartością własną macierzy A wtedy i tylko wtedy, jeśli jest pierwiastkiem wielomianu charakterystycznego*

$$w_n(\lambda) = \det(A - \lambda I)$$

macierzy A .

Definicja Wektory i wartości własne macierzy transponowanej A^T nazywamy odpowiednio lewostronnymi wektorami i lewostronnymi wartościami własnymi macierzy A .

Pojęcia podstawowe (2)

Definicja Pierwiastki $\lambda_1, \dots, \lambda_n$ wielomianu charakterystycznego (z uwzględnieniem ewentualnych wielokrotności tych pierwiastków) nazywamy widmem macierzy $\mathbf{A}_{n \times n}$. Zbiór liczb $\lambda_1, \dots, \lambda_n$ oznaczamy przez $\text{Sp}(\mathbf{A})$.

Twierdzenie 2 *Niech λ_L i λ_P będą odpowiednio lewostronną wartością własną i wartością własną macierzy \mathbf{A} , a x_L i x_P odpowiadającymi tym wartościom lewostronnym wektorem własnym i wektorem własnym. Jeżeli $\lambda_l \neq \lambda_P$, to wektory x_L i x_P są ortogonalne, czyli*

$$\langle x_l, x_P \rangle = 0.$$

Pojęcia podstawowe (3)

Twierdzenie 3 *Jeżeli $p(t)$ jest wielomianem zmiennej t , a liczby $\lambda_1, \dots, \lambda_n$ tworzą widmo macierzy $\mathbf{A}_{n \times n}$, to widmem macierzy $p(\mathbf{A})$ są liczby $p(\lambda_1), \dots, p(\lambda_n)$. Wektory własne macierzy \mathbf{A} są wektorami własnymi macierzy $p(\mathbf{A})$.*

Twierdzenie 4 (Cayleya–Hamiltona) *Jeżeli $w_n(\lambda)$ jest wielomianem charakterystycznym macierzy \mathbf{A} , to $w_n(\mathbf{A})$ jest macierzą zerową.*

Definicja Mówimy, że macierze \mathbf{A} i \mathbf{B} są podobne, jeśli istnieje taka nieosobliwa macierz \mathbf{P} , zwana macierzą podobieństwa, że

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}.$$

Pojęcia podstawowe (4)

Twierdzenie 5 *Jeżeli macierze A i B są podobne, to*

$$\text{Sp}(A) = \text{Sp}(B).$$

Definicja Macierz A nazywamy unitarną, jeśli

$$A^\dagger A = I.$$

Definicja Macierz B nazywamy ortogonalną, jeśli

$$B^T B = I.$$

Twierdzenie 6 *Wartości własne macierzy symetrycznej A i jej wektory własne są rzeczywiste.*

Twierdzenie 7 *Każda macierz jest unitarnie podobna do macierzy trójkątnej.*

Pojęcia podstawowe (5)

Definicja Macierz kwadratową $k \times k$ postaci

$$\mathbf{J} = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$

nazywamy klatką Jordana rzędu k . Liczba λ może być zespolona.

Pojęcia podstawowe (6)

Twierdzenie 8 *Dla każdej macierzy A istnieją takie macierze P , że*

$$P^{-1}AP = \begin{pmatrix} J_1 & & 0 \\ & J_2 & \\ & & \ddots \\ & & & J_r \end{pmatrix},$$

gdzie macierze J_1, \dots, J_r są klatkami Jordana. Powyższą postać nazywamy postacią kanoniczną Jordana macierzy A .

Pojęcia podstawowe (7)

Twierdzenie 9 (Schura) *Suma kwadratów modułów wartości własnych jest ograniczona od góry przez kwadrat normy euklidesowej, tzn.*

$$\sum_{i=1}^n |\lambda_i|^2 \leq \|\mathbf{A}\|_E^2.$$

Twierdzenie 10 *Dodanie do macierzy \mathbf{A} macierzy jednostkowej pomnożonej przez skalar przesuwają widmo macierzy, gdyż dla dowolnej liczby c mamy*

$$\text{Sp}(\mathbf{A} + c\mathbf{I}) = \text{Sp}(\mathbf{A}) + c,$$

gdzie $\text{Sp}(\mathbf{A}) + c$ oznacza zbiór liczb $\lambda_1 + c, \dots, \lambda_n + c$.

Pojęcia podstawowe (8)

Twierdzenie 11 (*Gershgorina*)

(a) *Jeśli C_i oznaczają koła domknięte na płaszczyźnie zespolonej o środkach w punktach a_{ii} i promieniach równych sumie modułów elementów z danego wiersza spoza diagonal, tzn.*

$$C_i = \{z : |z - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}|\},$$

to

$$\text{Sp}(\mathbf{A}) \subset \bigcup_{i=1}^n C_i.$$

Pojęcia podstawowe (9)

Twierdzenie *(ciąg dalszy)*

(b) Jeżeli k kół C_i tworzy zbiór rozłączny z pozostałymi kołami, to w zbiorze tym leży dokładnie k wartości własnych macierzy A .

Twierdzenie 12 *Jeżeli elementy macierzy A są niemniejsze niż elementy macierzy B , to $\rho(A) \geq \rho(B)$.*

Zaburzenia wartości własnych (1)

$$\mathbf{A}x = \lambda x \rightarrow \bar{\mathbf{A}}x = \bar{\lambda}x$$

$$\bar{\lambda} = \lambda + \delta\lambda$$

$$\bar{\mathbf{A}} = \mathbf{A} + \delta\mathbf{A}$$

Metoda stabilna numerycznie:

- zastępcze zaburzenie $\delta\mathbf{A}$ dąży do zera przy zwiększeniu dokładności obliczeń,
- $\delta\mathbf{A} \rightarrow 0$ pociąga za sobą $\delta\lambda \rightarrow 0$

Zaburzenia wartości własnych (2)

Definicja Odległością liczby z od zbioru $\text{Sp}(\mathbf{A}) = \{\lambda_1, \dots, \lambda_n\}$ nazywamy liczbę

$$\text{dist}(z, \text{Sp}(\mathbf{A})) = \min_{i=1, \dots, n} |z - \lambda_i|.$$

Zaburzenia wartości własnych (3)

Definicja Odległością zbiorów spektralnych $\text{Sp}(\mathbf{A})$ i $\text{Sp}(\mathbf{B})$ dwóch macierzy $\mathbf{A}_{n \times n}$ i $\mathbf{B}_{n \times n}$ nazywamy liczbę

$$\min_p \max_{i=1, \dots, n} |\lambda_i - \gamma_i^{(p)}|,$$

gdzie λ_i są kolejnymi elementami uporządkowanego zbioru $\text{Sp}(\mathbf{A})$, a $\gamma_i^{(p)}$ - kolejnymi wartościami własnymi ze zbioru $\text{Sp}(\mathbf{B})$, które ustawiono w innym porządku, oznaczonym indeksem p . Odległość tą będziemy oznaczać przez

$$\|\text{Sp}(\mathbf{A}) - \text{Sp}(\mathbf{B})\|_\infty.$$

Zaburzenia wartości własnych (4)

Twierdzenie 13 *Jeżeli zaburzenie $\delta\mathbf{A}$ elementów macierzy \mathbf{A} dąży do zera, to $\text{Sp}(\mathbf{A} + \delta\mathbf{A})$ dąży do $\text{Sp}(\mathbf{A})$ w tym sensie, że*

$$\|\text{Sp}(\mathbf{A}) - \text{Sp}(\mathbf{A} + \delta\mathbf{A})\|_{\infty} \rightarrow 0.$$

Zaburzenia wartości własnych (5)

Twierdzenie 14 (*Bauera–Fikego*) *Jeżeli macierz \mathbf{A} jest podobna do macierzy diagonalnej \mathbf{D} , tzn. istnieje nieosobliwa macierz podobieństwa taka, że*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D},$$

to dla dowolnej wartości własnej γ macierzy $\mathbf{A} + \delta\mathbf{A}$ obowiązuje oszacowanie

$$\text{dist}(\gamma, \text{Sp}(\mathbf{A})) \leq \alpha,$$

gdzie

$$\alpha = \|\mathbf{P}^{-1}\|_2 \|\mathbf{P}\|_2 \|\delta\mathbf{A}\|_2.$$

Zaburzenia wartości własnych (6)

Macierz A jest symetryczna:

$\Rightarrow P$ jest ortogonalna

$\Rightarrow \|P^{-1}\|_2 \|P\|_2 = 1$

\Rightarrow zagadnienie własne jest zawsze zadaniem dobrze uwarunkowanym

Zaburzenia wartości własnych (7)

Twierdzenie 15 *Jeżeli $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$ oraz $\gamma \in \text{Sp}(\mathbf{A} + \delta\mathbf{A})$, to dla dowolnego zaburzenia $\delta\mathbf{A}$ jest spełnione oszacowanie*

$$\text{dist}(\gamma, \text{Sp}(\mathbf{A})) \leq k \max(\alpha, \sqrt[k]{\alpha}),$$

gdzie k to maksymalny rząd klatek Jordana w postaci kanonicznej macierzy, natomiast α zdefiniowane jest jak wyżej.

Zaburzenia wektorów własnych (1)

Niech

$$\mathbf{A}x_i = \lambda_i x_i$$

$$(\mathbf{A} + \delta\mathbf{A})y_i = \gamma_i y_i$$

Definicja Niech \mathbb{X}_i oznacza podprzestrzeń własną macierzy \mathbf{A} odpowiadającą wartości własnej λ_i . Załóżmy, że $\|y_i\|_2 = 1$. Miarą zaburzenia wektora własnego y_i (będącego wektorem własnym macierzy $\mathbf{A} + \delta\mathbf{A}$) względem wektorów własnych macierzy \mathbf{A} odpowiadających wartości własnej λ_i nazywamy liczbę

$$\rho(y_i, x_i) = \min_{x \in \mathbb{X}_i} \|y_i - x\|_2.$$

Zaburzenia wektorów własnych (2)

Twierdzenie 16 *Jeżeli macierz \mathbf{A} jest symetryczna oraz wartość własna λ_i jest odległa od różnych od niej wartości własnych macierzy \mathbf{A} o $d > 0$, to dla zaburzeń $\delta\mathbf{A}$ takich, że*

$$\|\delta\mathbf{A}\|_2 \leq d,$$

wektor własny y_i macierzy $\mathbf{A} + \delta\mathbf{A}$ odpowiadający wartości własnej γ_i jest obliczony z zaburzeniem

$$\rho(y_i, x_i) \leq \frac{\|\delta\mathbf{A}\|_2}{d - \|\delta\mathbf{A}\|_2}.$$

Zaburzenia wektorów własnych (3)

Niech

$\bar{\lambda}$ - obliczona wartość własna

\bar{x} - obliczony wektor własny

oraz

$$\|\bar{x}\|_2 = 1, \quad r = A\bar{x} - \bar{\lambda}\bar{x}$$

(r obliczamy dokładnie!)

Zaburzenia wektorów własnych (4)

Twierdzenie 17 (Wilkinsona) *Jeżeli macierz A jest symetryczna oraz $\|r\|_2 = \rho$, to w przedziale $\langle \bar{\lambda} - \rho, \bar{\lambda} + \rho \rangle$ znajduje się co najmniej jedna wartość własna macierzy A . Jeżeli wartości własne macierzy A znajdujące się poza tym przedziałem są odległe od liczby $\bar{\lambda}$ co najmniej o $d > \rho$, to*

$$\bar{\rho}(\bar{x}, x) \leq \frac{\rho}{d},$$

gdzie $\bar{\rho}(\bar{x}, x)$ jest błędem wektora własnego \bar{x} , obliczonym zgodnie z definicją, tzn. $\bar{\rho}(\bar{x}, x) = \min_{x \in \mathbb{X}} \|\bar{x} - x\|_2$, a \mathbb{X} jest podprzestrzenią rozpiętą na wektorach własnych macierzy A , odpowiadających wartościom własnym z przedziału $\langle \bar{\lambda} - \rho, \bar{\lambda} + \rho \rangle$.

Zagadnienie własne macierzy hermitowskiej (1)

$$\mathbf{A}^\dagger = \mathbf{A}$$

Własności:

1. wszystkie wartości własne są rzeczywiste,
2. wektory własne można sprowadzić do postaci ortogonalnej,
3. macierz hermitowską można sprowadzić do macierzy diagonalnej o tych samych wartościach własnych za pomocą przekształcenia przez podobieństwo z macierzą unitarną

Zagadnienie własne macierzy hermitowskiej (2)

Niech

$$\mathbf{A} = \mathbf{B} + i\mathbf{C}, \quad \mathbf{B} \text{ i } \mathbf{C} - \text{macierze rzeczywiste}$$

\Rightarrow \mathbf{B} jest symetryczna

\Rightarrow \mathbf{C} jest skośnosymetryczna

$$\begin{aligned} B_{ij} &= B_{ji}, \\ C_{ij} &= -C_{ji}. \end{aligned}$$

Zagadnienie własne macierzy hermitowskiej (3)

Podobnie, dla wektora własnego:

$$x = y + iz$$

\Rightarrow zagadnienie własne ma postać:

$$(\mathbf{B} + i\mathbf{C})(y + iz) = \lambda(y + iz)$$

czyli

$$\begin{pmatrix} \mathbf{B} & -\mathbf{C} \\ \mathbf{C} & \mathbf{B} \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \lambda \begin{pmatrix} y \\ z \end{pmatrix}$$

\Rightarrow zagadnienie własne macierzy hermitowskiej równoważne zagadnieniu własnemu symetrycznej macierzy rzeczywistej

Symetryczne macierze rzeczywiste

Sposób postępowania:

1. przekształcenie macierzy symetrycznej przez podobieństwo z macierzą ortogonalną do macierzy trójdagonalnej,
2. rozwiązanie zagadnienia własnego dla macierzy trójdagonalnej

Przekształcenie zachowuje wartości własne ...

- ⇒ wartości własne są takie same dla macierzy trójdagonalnej i macierzy wyjściowej
- ⇒ wektory własne obu macierzy związane za pomocą ortogonalnej macierzy przekształcenia

Wartości własne macierzy trójdzielnej (1)

$$T = \begin{pmatrix} b_1 & a_2 & & & 0 \\ a_2 & b_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & b_{n-1} & a_n \\ 0 & & & a_n & b_n \end{pmatrix}$$

Wielomian charakterystyczny $w_n(\lambda)$ wyznaczamy ze wzoru:

$$w_0(\lambda) = 1,$$

$$w_1(\lambda) = b_1 - \lambda,$$

$$w_i(\lambda) = (b_i - \lambda)w_{i-1}(\lambda) - a_i^2 w_{i-2}(\lambda)$$

Wartości własne macierzy trójdzielnej (2)

Twierdzenie 18 *Jeżeli elementy a_2, \dots, a_n są niezerowe, to wartości własne macierzy \mathbf{T} są pojedyncze.*

Twierdzenie 19 *Wszystkie pierwiastki równania $w_n(\lambda) = 0$ leżą w przedziale $\langle -\|\mathbf{T}\|_\infty, \|\mathbf{T}\|_\infty \rangle$. Przy tym*

$$\|\mathbf{T}\|_\infty = \max \left\{ \sum_{j=1}^n |T_{ij}| \right\}.$$

Wartości własne macierzy trójdzielnej (3)

Twierdzenie 20 *Jeżeli elementy a_2, \dots, a_n są niezerowe, to ciąg wartości $w_0(\lambda), \dots, w_n(\lambda)$ spełnia warunki:*

1. *jeżeli $w_i(\lambda) = 0$ dla pewnego $i < n$, to*

$$w_{i-1}(\lambda)w_{i+1}(\lambda) < 0,$$

2. *jeżeli $w_n(\lambda) \neq 0$, to liczba zmian znaków sąsiednich liczb $w_0(\lambda), \dots, w_n(\lambda)$ jest równa liczbie wartości własnych macierzy T mniejszych od λ ,*
3. *jeżeli $w_n(\lambda) = 0$, to λ jest wartością własną macierzy T , a ponadto jest tyle wartości własnych mniejszych od λ , ile nastąpiło zmian znaków w ciągu $w_0(\lambda), \dots, w_{n-1}(\lambda)$.*

Wartości własne macierzy trójdzielnej (4)

Metoda bisekcji:

1. obliczamy normę macierzy,
2. połowimy przedział $\langle -\|\mathbf{T}\|_\infty, \|\mathbf{T}\|_\infty \rangle$ i wyznaczamy znaki wielomianów $w_i(0)$
3. dzielimy przedziały $\langle -\|\mathbf{T}\|_\infty, 0 \rangle$ i $\langle 0, \|\mathbf{T}\|_\infty \rangle$ na pół i badamy znaki wielomianów w każdym punkcie podziału;
4. postępowanie kontynuujemy do momentu, aż każda wartość własna będzie umiejscowiona w oddzielnym przedziale o długości

$$\frac{\|\mathbf{T}\|_\infty}{2^{l-1}}$$

gdzie l to liczba kroków

Wartości własne macierzy trójdzielnej (5)

Własności metody:

- dokładność wyznaczonej wartości własnej nie zależy w istotny sposób od wartości liczb $w_0(\lambda), \dots, w_n(\lambda)$
- jeśli wykonaliśmy m podziałów, to odległość środka ostatniego przedziału od λ_k można szacować jako

$$|t - \lambda_k| < (2,06\epsilon + 2^{-m})\|\mathbf{T}\|_\infty$$

Wektory własne macierzy trójdzielnej

Jeśli $a_i \neq 0$ dla każdego i :

$$\begin{aligned}x_1 &= 1, \\x_2 &= \frac{\lambda - b_1}{a_2}, \\x_{i+1} &= \frac{(\lambda - b_i)x_i - a_i x_{i-1}}{a_{i+1}}, \quad i = 2, 3, \dots, n-1.\end{aligned}$$

W przeciwnym razie

- dzielimy T na trójdzielne bloki z $a_i \neq 0$,
- obliczamy wartości własne bloków - są one również wartościami własnymi macierzy T
- wektory własne bloków po uzupełnieniu zerami są wektorami własnymi macierzy T

Trójdagonalizacja macierzy symetrycznej (1)

Metoda Householdera:

- niech

$$\mathbf{P}^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}^{(1)} \end{pmatrix},$$

gdzie $\mathbf{H}^{(1)}$ sprowadza wektor $z^{(1)} = (a_{21}, \dots, a_{n1})^T$ do postaci

$$\alpha^{(1)} \|z^{(1)}\|_2 (1, 0, \dots, 0)_{1 \times (n-1)}^T$$

Trójdagonalizacja macierzy symetrycznej (2)

- obliczamy macierz

$$\mathbf{M}^{(2)} = \mathbf{P}^{(1)} \mathbf{A} \mathbf{P}^{(1)} = \begin{pmatrix} b_1 & a_1 & 0 & & \dots & 0 \\ a_1 & & & & & \\ 0 & & \mathbf{A}_{(n-1) \times (n-1)}^{(2)} & & & \\ \vdots & & & & & \\ 0 & & & & & \end{pmatrix}$$

Trójdagonalizacja macierzy symetrycznej (3)

- określamy macierz

$$\mathbf{P}^{(2)} = \begin{pmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & \mathbf{H}^{(2)} & \end{pmatrix},$$

gdzie $\mathbf{H}^{(2)}$ sprowadza wektor $z^{(2)} = (a_{21}^{(2)}, \dots, a_{(n-1)1}^{(2)})^T$ do postaci

$$\alpha^{(2)} \|z^{(2)}\|_2 (1, 0, \dots, 0)_{1 \times (n-2)}^T$$

Trójdagonalizacja macierzy symetrycznej (4)

- obliczamy macierz

$$\mathbf{M}^{(3)} = \mathbf{P}^{(2)}\mathbf{M}^{(2)}\mathbf{P}^{(2)} = \begin{pmatrix} b_1 & a_1 & 0 & \dots & 0 \\ a_1 & b_2 & a_2 & \dots & 0 \\ 0 & a_2 & & & \\ \vdots & \vdots & & \mathbf{A}_{(n-2) \times (n-2)}^{(3)} & \\ 0 & 0 & & & \end{pmatrix}$$

Po $(n - 2)$ takich przekształceniach uzyskamy poszukiwaną macierz trójdagonalną

$$\mathbf{M}^{(n-2)} = \mathbf{T}$$

Trójdagonalizacja macierzy symetrycznej (5)

Konkretna postać macierzy ortogonalnych:

$$\mathbf{P}^{(k)} = \mathbf{I} - \frac{1}{\eta^{(k)}} w^{(k)} (w^{(k)})^T,$$

przy czym l -ta składowa wektora $w^{(k)}$ jest równa

$$w_l^{(k)} = \begin{cases} 0, & l \leq k, \\ A_{k,k+1}^{(k-1)} + \alpha^{(k)}, & l = k + 1, \\ A_{kl}^{(k-1)}, & l \geq k + 2, \end{cases}$$

oraz

$$\alpha^{(k)} = \pm \sqrt{\sum_{l=k+1}^n \left[A_{kl}^{(k-1)} \right]^2}, \quad \eta^{(k)} = \alpha^{(k)} \left[\alpha^{(k)} + A_{k,k+1}^{(k-1)} \right]$$

Trójdagonalizacja macierzy symetrycznej (6)

Przykład

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 3 & 1 & 2 \\ 4 & 2 & 1 \end{pmatrix}$$

Obliczamy

$$\alpha^{(1)} = \sqrt{\sum_{l=2}^3 A_{1l}^2} = 5, \quad \eta^{(1)} = \alpha^{(1)}(\alpha^{(1)} + A_{12}) = 40,$$

oraz

$$w_1^{(1)} = 0, \quad w_2^{(1)} = A_{21} + \alpha^{(1)} = 8, \quad w_3^{(1)} = A_{31} = 4$$

Trójdagonalizacja macierzy symetrycznej (7)

Przykład (ciąg dalszy)

Stąd

$$\mathbf{P}^{(1)} = \mathbf{I} - \frac{1}{\eta^{(1)}} w^{(1)} (w^{(1)})^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{pmatrix},$$

co prowadzi do

$$\mathbf{P}^{(1)} \mathbf{A} \mathbf{P}^{(1)} = \begin{pmatrix} 1 & -5 & 0 \\ -5 & \frac{73}{25} & \frac{14}{25} \\ 0 & \frac{14}{25} & -\frac{23}{25} \end{pmatrix}$$

Trójdagonalizacja macierzy symetrycznej (8)

$$\|\mathbf{T} - \tilde{\mathbf{T}}\|_E \leq 2k_1 n^2 \epsilon (1 + k_1 \epsilon n)^{2n} \|\mathbf{A}\|_E,$$

$\tilde{\mathbf{T}}$ - dokładna macierz trójdagonalna

k_1 - stała zależna od sposobu zaokrąglania

Liczba działań

- $M = \frac{2}{3}n^3 + O(n^2)$ mnożeń
- $D = \frac{2}{3}n^3 + O(n^2)$ dodawań
- $P = n - 2$ pierwiastkowań

Metoda Kryłowa (1)

Przykład

$$\mathbf{A} = \begin{pmatrix} 1,00 & 0,01 \\ 0,01 & 1,00 \end{pmatrix}$$

Macierz jest symetryczna - zagadnienie własne jest na pewno dobrze uwarunkowane:

$$\lambda_1 = 1,01, \quad \lambda_2 = 0,99$$

Jeśli

$$\delta\mathbf{A} = \begin{pmatrix} \epsilon_1 & \epsilon_2 \\ \epsilon_3 & \epsilon_4 \end{pmatrix},$$

a $\lambda_1(\epsilon)$ jest większą z wartości własnych macierzy $\mathbf{A} + \delta\mathbf{A}$, to

Metoda Kryłowa (2)

Przykład (ciąg dalszy)

$$\lambda_1(\epsilon) = 1,01 - \frac{1}{2}\epsilon_1 + \epsilon_2 + \epsilon_3 - \frac{1}{2}\epsilon_4 + O(\max\{|\epsilon_1|, |\epsilon_2|, |\epsilon_3|, |\epsilon_4|\}).$$

Jeśli jednak wyznaczymy wielomian charakterystyczny macierzy A ,

$$w_2(\lambda) = \lambda^2 - 2\lambda + 0,9999,$$

to większy z pierwiastków równania

$$\lambda^2 + (-2 + \delta_1)\lambda + (0,9999 + \delta_2) = 0$$

spełnia warunek

$$\lambda_1(\delta) = 1,01 - 50,5\delta_1 - 50\delta_2 + O(\max\{|\delta_1|, |\delta_2|\})$$

Metoda potęgowa (1)

Algorytm:

1. Dla $i = 0$, wybieramy dowolny wektor x_0 taki, że $\|x_0\|_\infty = 1$. Ustalamy maksymalną liczbę iteracji $IMAX$.
2. Obliczamy

$$v_{i+1} = \mathbf{A}x_i, \quad m_{i+1} = \|v_{i+1}\|_\infty.$$

Jeżeli $m_{i+1} = 0$, przerywamy obliczenia, jeśli nie, to

$$x_{i+1} = \frac{v_{i+1}}{m_{i+1}}.$$

3. Jeśli $i + 1 \geq IMAX$, przerywamy obliczenia. W przeciwnym razie zwiększamy i o jeden i wracamy do punktu 2.

Metoda potęgowa (2)

Ciąg wektorów $x_0, x_2, \dots, x_{2j}, \dots$ jest zbieżny do wektora x

\Rightarrow ciąg wartości m_1, m_3, \dots jest zbieżny do m

Jeśli

$$\|x_{2j+2} - x_{2j}\| \rightarrow 0,$$

to

$$Ax = mx$$

W przeciwnym razie

$$\|x_{2j+2} - x_{2j}\| \rightarrow 2,$$

i

$$Ax = -mx$$

Metoda potęgowa (3)

Niech

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$$

przy czym

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Metoda potęgowa (4)

Twierdzenie 21 *Jeśli wśród wartości własnych $\lambda_1, \dots, \lambda_n$ macierzy A nie ma różnych wartości własnych o równych modułach (tzn. nie można wskazać takich $\lambda_i \neq \lambda_j$, że $|\lambda_i| = |\lambda_j|$), to przy dowolnym wyborze wektora startowego x_0 ciąg wektorów x_0, x_2, \dots jest zbieżny (zakładamy przy tym, że działania są wykonywane dokładnie).*

Algorytm QR dla macierzy Hessenberga (1)

Definicja Górną macierzą Hessenberga nazywamy macierz

$$H = T + U,$$

gdzie T jest macierzą trójdziagonalną, a U - macierzą trójkątną górną.

Twierdzenie 22 *Każdą macierz kwadratową A można przedstawić w postaci iloczynu macierzy ortogonalnej Q i macierzy trójkątnej górnej R ,*

$$A = QR.$$

Algorytm QR dla macierzy Hessenberga (2)

Określamy ciąg macierzy

$$\mathbf{H} = \mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \dots$$

w następujący sposób:

$$\begin{aligned}\mathbf{H}^{(i)} &= \mathbf{Q}^{(i)} \mathbf{R}^{(i)}, \\ \mathbf{H}^{(i+1)} &= \mathbf{R}^{(i)} \mathbf{Q}^{(i)}\end{aligned}$$

dla $i = 1, 2, \dots$

Algorytm QR dla macierzy Hessenberga (3)

- \Rightarrow wszystkie macierze $Q^{(i)}$ oraz $H^{(i)}$, $i = 1, 2, \dots$ są macierzami Hessenberga
- \Rightarrow jeśli wartości własne macierzy H są rzeczywiste i różne (także co do modułu), wówczas elementy pod diagonalą macierzy $H^{(i)}$ zanikają i macierze te upodobniają się do macierzy trójkątnej
- \Rightarrow wszystkie macierze $H^{(i)}$ są podobne, ponieważ

$$H^{(i+1)} = (Q^{(i)})^T H^{(i)} Q^{(i)},$$

- \Rightarrow elementy na diagonalu są zbieżne do wartości własnych macierzy H

Algorytm QR dla macierzy Hessenberga (4)

Jeżeli macierz H ma pojedyncze wartości własne $\lambda_1, \dots, \lambda_n$ takie, że są między nimi pary sprzężone, to nie wszystkie elementy pod diagonalą macierzy $H^{(i)}$ dążą do zera, lecz

$$h_{j,j-1}^{(i)} h_{j+1,j}^{(i)} \xrightarrow{i \rightarrow \infty} 0.$$

Wówczas w granicy otrzymamy np. macierz

$$H^{(\infty)} = \begin{pmatrix} * & * & \cdot & \cdot & \cdot \\ * & * & \cdot & \cdot & \cdot \\ 0 & 0 & * & \cdot & \cdot \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{pmatrix}$$

Algorytm QR dla macierzy Hessenberga (5)

Uwaga:

- W przypadku ogólnym $\mathbf{H}^{(i)}$ nie musi dążyć do postaci trójkątnej lub prawie trójkątnej.
- Jeżeli zbieżność jest wolna, można zastosować przesunięcie wartości własnych w następujący sposób:

$$\begin{aligned}\mathbf{H}^{(i)} - k_i \mathbf{I} &= \mathbf{Q}^{(i)} \mathbf{R}^{(i)}, \\ \mathbf{H}^{(i+1)} &= \mathbf{R}^{(i)} \mathbf{Q}^{(i)} + k_i \mathbf{I}.\end{aligned}$$

Dzięki odpowiedniemu dobraniu przesunięcia k_i można uzyskać dużo szybszą zbieżność elementów macierzy $\mathbf{H}^{(i)}$.

Sprowadzanie macierzy do postaci Hessenberga (1)

- metoda Givensa,
- metoda Householdera,
- metoda eliminacji Gaussa.

Sprowadzanie macierzy do postaci Hessenberga (2)

Na macierzy $\mathbf{A}_{n \times n}$ dokonujemy $(n - 2)$ przekształceń, uzyskując kolejno macierze

$$\mathbf{A} = \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(n-2)} = \mathbf{H},$$

gdzie \mathbf{H} jest poszukiwaną macierzą Hessenberga

Założmy, że mamy już wyznaczoną macierz $\mathbf{A}^{(i)}$, której $(i - 1)$ początkowych kolumn jest równe $(i - 1)$ kolumnom macierzy \mathbf{H} :

1. Wybieramy element o największym module spośród $a_{i+1,i}^{(i)}$, \dots , $a_{n,i}^{(i)}$. Jeśli są same elementy zerowe, to przyjmujemy $\mathbf{A}^{(i)} = \mathbf{A}^{(i+1)}$ i przystępujemy do wyznaczenia kolejnej macierzy. Gdy element o maksymalnym module znajduje się w k -tym wierszu ($k \geq i$), przestawiamy wiersze i i kolumny o numerach k oraz $i + 1$.

Sprowadzanie macierzy do postaci Hessenberga (3)

2. Obliczamy mnożniki

$$m_j^{(i)} = \frac{a_{j,i}^{(i)}}{a_{i+1,i}^{(i)}}, \quad j = i + 2, i + 3, \dots, n.$$

3. Od j -tego wiersza odejmujemy $(i + 1)$ -szy wiersz pomnożony przez $m_j^{(i)}$ oraz do $(i + 1)$ -szej kolumny dodajemy j -tą kolumnę pomnożoną przez $m_j^{(i)}$, $j = i + 2, \dots, n$. Uzyskana macierz $\mathbf{A}^{(i+1)}$ jest podobna do $\mathbf{A}^{(i)}$, a elementy $a_{i+2,i}^{(i+1)}, a_{i+3,i}^{(i+1)}, \dots, a_{n,i}^{(i+1)}$ są równe zeru, co powoduje, że i -ta kolumna macierzy $\mathbf{A}^{(i+1)}$ jest równa i -tej kolumnie macierzy \mathbf{H} .

Rozkład macierzy na iloczyn QR (1)

Twierdzenie 23 *Każdą macierz $A_{m \times n}$, $m \leq n$ posiadającą $k \leq n$ liniowo niezależnych kolumn można przedstawić w postaci iloczynu*

$$A = QR,$$

gdzie $Q_{m \times k}$ jest macierzą ortogonalną, a $R_{k \times n}$ jest macierzą trapezoidalną górną.

Definicja Macierz R nazywamy trapezoidalną górną, jeżeli wszystkie elementy tej macierzy znajdujące się pod diagonalą są równe zeru,

$$r_{ij} = 0, \quad i > j.$$

Rozkład macierzy na iloczyn QR (2)

Ortogonalizacja Grama–Schmidta:

$A^{(1)}, A^{(2)}, \dots, A^{(n)}$ - macierze złożone z 1, 2, ..., n
początkowych kolumn macierzy A

Zakładamy, że kolumna a_1 niezerowa:

$$A^{(1)} = Q^{(1)} R^{(1)},$$

gdzie

$$Q^{(1)} = \begin{bmatrix} \frac{a_1}{\|a_1\|_2} \end{bmatrix}, \quad R^{(1)} = [\|a_1\|_2]$$

Rozkład macierzy na iloczyn QR (3)

Jeżeli wyznaczyliśmy już rozkład $\mathbf{A}^{(i)} = \mathbf{Q}^{(i)}\mathbf{R}^{(i)}$, to dla macierzy

$$\mathbf{A}^{(i+1)} = [\mathbf{A}^{(i)} a_{i+1}]$$

mamy

$$\mathbf{A}^{(i+1)} = \mathbf{Q}^{(i+1)}\mathbf{R}^{(i+1)},$$

gdzie

$$\begin{aligned}\mathbf{Q}^{(i+1)} &= [\mathbf{Q}^{(i)} q_{i+1}] \\ \mathbf{R}^{(i+1)} &= \begin{bmatrix} \mathbf{R}^{(i)} & r_{i+1} \\ 0 & s_{i+1} \end{bmatrix}\end{aligned}$$

Rozkład macierzy na iloczyn QR (4)

Wektory q_{i+1} i r_{i+1} , oraz liczba s_{i+1} dane są wzorami:

$$r_{i+1} = \left(Q^{(i)} \right)^T a_{i+1}$$

$$p_{i+1} = a_{i+1} - Q^{(i)} r_{i+1}$$

$$s_{i+1} = \|p_{i+1}\|_2,$$

oraz, jeżeli $s_{i+1} = 0$, to

$$Q^{(i+1)} = Q^{(i)},$$

w przeciwnym razie

$$q_{i+1} = \frac{p_{i+1}}{s_{i+1}}$$

$(i = 1, 2, \dots, n - 1).$

Rozkład macierzy na iloczyn QR (5)

Nakład obliczeń:

- $M = n^3 + n$ mnożenia
- $D = n^3 - \frac{1}{2}n^2 - \frac{1}{2}n$ dodawania
- $P = n$ pierwiastkowania

Wektory własne w metodzie QR (1)

Dla danej macierzy A

- wyznaczamy najpierw macierz P taką, że

$$P^{-1}AP = H$$

- wyznaczamy ciąg macierzy $H^{(i)}$
- równocześnie z macierzą

$$H^{(i+1)} = (Q^{(i)})^T H^{(i)} Q^{(i)}$$

wyznaczamy macierz pomocniczą

$$Z^{(i+1)} = Z^{(i)} Q^{(i)}, \quad Z^{(1)} = P$$

Wektory własne w metodzie QR (2)

Gdy macierz $\mathbf{H}^{(N)}$ jest dostatecznie bliska macierzy trójkątnej górnej, obliczamy jej wektory własne ze wzoru:

$$x_j^{(i)} = 0, \quad j = n, n-2, \dots, i+1,$$

$$x_i^{(i)} = 1,$$

$$x_j^{(i)} = -\frac{\left(\sum_{k=j+1}^i h_{jk}^{(N)} x_k^{(i)}\right)}{h_{jj}^{(N)} - h_{ii}^{(N)}}, \quad j = i-1, i-2, \dots, 1.$$

Wektory własne macierzy \mathbf{A} będą dane wzorem

$$y^{(k)} = Z^{(N)} x^{(k)}$$

GNU Scientific Library

Zasoby w sieci:

<http://www.gnu.org/software/gsl>

Użycie:

- `#include <gsl/gsl_math.h>`
- `gcc -o test test.c -L/usr/lib -lgsl -lgslcblas`

GSL - Wektory

Alokacja:

- `gsl_vector * gsl_vector_alloc(size_t n)`
- `gsl_vector * gsl_vector_calloc(size_t n)`
- `void gsl_vector_free(gsl_vector * v)`

Dostęp do elementów:

- `double gsl_vector_get(const gsl_vector * v, size_t i)`
- `void gsl_vector_set(gsl_vector * v, size_t i, double x)`

GSL - Macierze

Alokacja:

- `gsl_matrix * gsl_matrix_alloc(size_t n1, size_t n2)`
- `gsl_matrix * gsl_matrix_calloc(size_t n1, size_t n2)`
- `void gsl_matrix_free(gsl_matrix * m)`

Dostęp do elementów macierzy

- `double gsl_matrix_get(const gsl_matrix * m,
size_t i, size_t j)`
- `void gsl_matrix_set(gsl_matrix * m,
size_t i, size_t j, double x)`

GSL - trójdagonalizacja i rozkład QR

- `int gsl_linalg_symmtd_decomp(gsl_matrix * A,
gsl_vector * tau)`
- `int gsl_linalg_symmtd_unpack_T(const gsl_matrix * A,
gsl_vector * diag, gsl_vector * subdiag)`
- `int gsl_linalg_QR_decomp(gsl_matrix * A, gsl_vector * tau)`