Zagadnienia własne

Janusz Szwabiński

17 grudnia 2002

Zagadnienia własne

Zagadnienie własne

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

λ - wartość własna
x - wektor własny

Pojęcia podstawowe (1)

Twierdzenie 1 *Liczba λ jest wartością własną macierzy $\mathbf{A}$ wtedy i tylko wtedy, jeśli jest pierwiastkiem wielomianu charakterystycznego*

$$w_n(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$$

macierzy $\mathbf{A}$.

Definicja Wektory i wartości własne macierzy transponowanej $\mathbf{A}^{\mathrm{T}}$ nazywamy odpowiednio lewostronnymi wektorami i lewostronnymi wartościami własnymi macierzy $\mathbf{A}$.

Pojęcia podstawowe (2)

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

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 $\lambda_1, \ldots, \lambda_n$ tworzą widmo macierzy $\mathbf{A}_{n\times n}$, to widmem macierzy $p(\mathbf{A})$ są liczby $p(\lambda_1), \ldots, p(\lambda_n)$. 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 $\mathbf{A}$ i $\mathbf{B}$ są podobne, to*

$$\mathrm{Sp}(\mathbf{A}) = \mathrm{Sp}(\mathbf{B}).$$

Definicja Macierz $\mathbf{A}$ nazywamy unitarną, jeśli

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

Definicja Macierz $\mathbf{B}$ nazywamy ortogonalną, jeśli

$$\mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}.$$

Twierdzenie 6 *Wartości własne macierzy symetrycznej $\mathbf{A}$ 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* $\mathbf{A}$ *istnieją takie macierze* $\mathbf{P}$, *że*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{J}_1 & & & 0 \\ & \mathbf{J}_2 & & \\ & & \ddots & \\ & & & \mathbf{J}_r \end{pmatrix},$$

gdzie macierze $\mathbf{J}_1, \ldots \mathbf{J}_r$ *są klatkami Jordana. Powyższą postać nazywamy postacią kanoniczną Jordana macierzy* $\mathbf{A}$.

Pojęcia podstawowe (7)

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

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

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

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

gdzie $\mathrm{Sp}(\mathbf{A}) + c$ *oznacza zbiór liczb* $\lambda_1 + c, \ldots, \lambda_n + c$.

Pojęcia podstawowe (8)

Twierdzenie 11 *(Gershgorina)*

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

$$\mathbf{C}_i = \{z : |z - a_{ii}| \le \sum_{j=1, j\neq i}^{n} |a_{ij}|\},$$

to

$$\mathrm{Sp}(\mathbf{A}) \subset \bigcup_{i=1}^{n} \mathbf{C}_i.$$

Pojęcia podstawowe (9)

Twierdzenie *(ciąg dalszy)*

(b) Jeżeli k kół $\mathbf{C}_i$ *tworzy zbiór rozłączny z pozostałymi kołami, to w zbiorze tym leży dokładnie k wartości własnych macierzy* $\mathbf{A}$.

Twierdzenie 12 *Jeżeli elementy macierzy* $\mathbf{A}$ *są niemniejsze niż elementy macierzy* $\mathbf{B}$, *to* $\rho(\mathbf{A}) \ge \rho(\mathbf{B})$.

Zaburzenia wartości własnych (1)

$$\mathbf{A}x = \lambda x \quad \rightarrow \quad \bar{\mathbf{A}}x = \bar{\lambda}x$$
$$\bar{\lambda} = \lambda + \delta\lambda$$
$$\bar{\mathbf{A}} = \mathbf{A} + \delta\mathbf{A}$$

Metoda stabilna numerycznie:

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

Zaburzenia wartości własnych (2)

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

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

Zaburzenia wartości własnych (3)

Definicja Odległością zbiorów spektralnych $\mathrm{Sp}(\mathbf{A})$ i $\mathrm{Sp}(\mathbf{B})$ 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 $\mathrm{Sp}(\mathbf{A})$, a $\gamma_i^{(p)}$ - kolejnymi wartościami własnymi ze zbioru $\mathrm{Sp}(\mathbf{B})$, które ustawiono w innym porządku, oznaczonym indeksem p. Odległość tą będziemy oznaczać przez

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

Zaburzenia wartości własnych (4)

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

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

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

$$\mathrm{dist}(\gamma, \mathrm{Sp}(\mathbf{A})) \le \alpha,$$

gdzie

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

Zaburzenia wartości własnych (6)

Macierz $\mathbf{A}$ jest symetryczna:

$\Rightarrow$ $\mathbf{P}$ jest ortogonalna

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

$\Rightarrow$ zagadnienie własne jest zawsze zadaniem dobrze uwarunkowanym

---

Zaburzenia wartości własnych (7)

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

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

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

Zaburzenia wektorów własnych (1)

Niech

$$\begin{aligned} \mathbf{A}x_i &= \lambda_i x_i \\ (\mathbf{A}+\delta\mathbf{A})y_i &= \gamma_i y_i \end{aligned}$$

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

$$\begin{aligned} &\bar{\lambda} \quad - \quad \text{obliczona wartość własna} \\ &\bar{x} \quad - \quad \text{obliczony wektor własny} \end{aligned}$$

oraz

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

(r obliczamy dokładnie!)

Zaburzenia wektorów własnych (4)

Twierdzenie 17 *(Wilkinsona) Jeżeli macierz* $\mathbf{A}$ *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* $\mathbf{A}$*. Jeżeli wartości własne macierzy* $\mathbf{A}$ *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* $\mathbf{A}$*, 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

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

⇒ $\mathbf{B}$ jest symetryczna

⇒ $\mathbf{C}$ jest skośnosymetryczna

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

Zagadnienie własne macierzy hermitowskiej (3)

Podobnie, dla wektora własnego:

$$x = y + iz$$

⇒ zagadnienie własne ma postać:

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

czyli

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

⇒ 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 trójdiagonalnej,
2. rozwiązanie zagadnienia własnego dla macierzy trójdiagonalnej

Przekształcenie zachowuje wartości własne ...

⇒ wartości własne są takie same dla macierzy trójdiagonalnej i macierzy wyjściowej

⇒ wektory własne obu macierzy związane za pomocą ortogonalnej macierzy przekształcenia

Wartości własne macierzy trójdiagonalnej (1)

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

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

$$
\begin{aligned}
w_0(\lambda) &= 1, \\
w_1(\lambda) &= b_1 - \lambda, \\
w_i(\lambda) &= (b_i - \lambda) w_{i-1}(\lambda) - a_i^2 w_{i-2}(\lambda)
\end{aligned}
$$

Wartości własne macierzy trójdiagonalnej (2)

Twierdzenie 18 *Jeżeli elementy $a_2, \ldots, a_n$ 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 trójdiagonalnej (3)

Twierdzenie 20 *Jeżeli elementy $a_2, \ldots, a_n$ są niezerowe, to ciąg wartości $w_0(\lambda), \ldots, w_n(\lambda)$ 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 $w_0(\lambda), \ldots, w_n(\lambda)$ jest równa liczbie wartości własnych macierzy $\mathbf{T}$ mniejszych od λ,*
3. *jeżeli $w_n(\lambda) = 0$, to λ jest wartością własną macierzy $\mathbf{T}$, a ponadto jest tyle wartości własnych mniejszych od λ, ile nastąpiło zmian znaków w ciągu $w_0(\lambda), \ldots, w_{n-1}(\lambda)$.*

Wartości własne macierzy trójdiagonalnej (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 trójdiagonalnej (5)

Własności metody:

- dokładność wyznaczonej wartości własnej nie zależy w istotny sposób od wartości liczb $w_0(\lambda), \ldots, w_n(\lambda)$
- 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 trójdiagonalnej

Jeśli $a_i \neq 0$ dla każdego i:

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

W przeciwnym razie

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

Trójdiagonalizacja 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 $z^{(1)} = (a_{21}, \ldots, a_{n1})^{\mathrm{T}}$ do postaci

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

Trójdiagonalizacja macierzy symetrycznej (2)

- obliczamy macierz

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

Trójdiagonalizacja 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 $z^{(2)} = (a_{21}^{(2)}, \ldots, a_{(n-1)1}^{(2)})^{\mathrm{T}}$ do postaci

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

Trójdiagonalizacja macierzy symetrycznej (4)

- obliczamy macierz

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

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

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

Trójdiagonalizacja macierzy symetrycznej (5)

Konkretna postać macierzy ortogonalnych:

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

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

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

Trójdiagonalizacja 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

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

Trójdiagonalizacja macierzy symetrycznej (7)

Przykład (ciąg dalszy)
Stąd

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

co prowadzi do

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

Trójdiagonalizacja 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 trójdiagonalna
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 $\mathbf{A}$,

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

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 $x_0, x_2, \ldots, x_{2j}, \ldots$ jest zbieżny do wektora x

$\Rightarrow$ ciąg wartości $m_1, m_3, \ldots$ jest zbieżny do m

Jeśli
$$\|x_{2j+2} - x_{2j}\| \to 0,$$
to
$$\mathbf{A}x = mx$$
W przeciwnym razie
$$\|x_{2j+2} - x_{2j}\| \to 2,$$
i
$$\mathbf{A}x = -mx$$

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 $\lambda_1, \ldots, \lambda_n$ macierzy $\mathbf{A}$ 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, $\ldots$ jest zbieżny (zakładamy przy tym, że działania są wykonywane dokładnie).*

Algorytm QR dla macierzy Hessenberga (1)

Definicja Górną macierzą Hessenberga nazywamy macierz
$$\mathbf{H} = \mathbf{T} + \mathbf{U},$$
gdzie $\mathbf{T}$ jest macierzą trójdiagonalną, a $\mathbf{U}$ - macierzą trójkątną górną.

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

$$\mathbf{A} = \mathbf{QR}.$$

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 $\mathbf{Q}^{(i)}$ oraz $\mathbf{H}^{(i)}$, $i = 1, 2, \dots$ są macierzami Hessenberga

$\Rightarrow$ jeśli wartości własne macierzy $\mathbf{H}$ są rzeczywiste i różne (także co do modułu), wówczas elementy pod diagonalą macierzy $\mathbf{H}^{(i)}$ zanikają i macierze te upodobniają się do macierzy trójkątnej

$\Rightarrow$ wszystkie macierze $\mathbf{H}^{(i)}$ są podobne, ponieważ

$$\mathbf{H}^{(i+1)} = \left(\mathbf{Q}^{(i)}\right)^{\mathrm{T}} \mathbf{H}^{(i)}\mathbf{Q}^{(i)},$$

$\Rightarrow$ elementy na diagonali są zbieżne do wartości własnych macierzy $\mathbf{H}$

Algorytm QR dla macierzy Hessenberga (4)

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

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

Wówczas w granicy otrzymamy np. macierz

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

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

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

gdzie $\mathbf{H}$ jest poszukiwaną macierzą Hessenberga

Załóżmy, 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 $a^{(i)}_{i+1,i},\ \ldots,\ a^{(i)}_{n,i}$. 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 kolumny o numerach k oraz $i+1$.

Sprowadzanie macierzy do postaci Hessenberga (3)

2. Obliczamy mnożniki

$$m^{(i)}_j=\frac{a^{(i)}_{j,i}}{a^{(i)}_{i+1,i}},\quad j=i+2,\ i+3,\ \ldots,\ n.$$

3. Od j–tego wiersza odejmujemy $(i+1)$–szy wiersz pomnożony przez $m^{(i)}_j$ oraz do $(i+1)$–szej kolumny dodajemy j-tą kolumnę pomnożoną przez $m^{(i)}_j$, $j=i+2,\ldots,n$. Uzyskana macierz $\mathbf{A}^{(i+1)}$ jest podobna do $\mathbf{A}^{(i)}$, a elementy $a^{(i+1)}_{i+2,i}$, $a^{(i+1)}_{i+3,i}$, $\ldots$, $a^{(i+1)}_{n,i}$ 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 $\mathbf{A}_{m\times n}$, $m\leq n$ posiadającą $k\leq n$ liniowo niezależnych kolumn można przedstawić w postaci iloczynu*

$$\mathbf{A}=\mathbf{QR},$$

gdzie $\mathbf{Q}_{m\times k}$ jest macierzą ortogonalną, a $\mathbf{R}_{k\times n}$ jest macierzą trapezoidalną górną.

Definicja Macierz $\mathbf{R}$ 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:

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

Zakładamy, że kolumna a_1 niezerowa:

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

gdzie

$$\mathbf{Q}^{(1)} = \left[\frac{a_1}{\|a_1\|_2}\right], \quad \mathbf{R}^{(1)} = [\|a_1\|_2]$$

Rozkład macierzy na iloczyn **QR** (3)

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

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

mamy

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

gdzie

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

Rozkład macierzy na iloczyn **QR** (4)

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

$$\begin{aligned} r_{i+1} &= \left(\mathbf{Q}^{(i)}\right)^{\mathrm{T}} a_{i+1} \\ p_{i+1} &= a_{i+1} - \mathbf{Q}^{(i)} r_{i+1} \\ s_{i+1} &= \|p_{i+1}\|_2, \end{aligned}$$

oraz, jeżeli $s_{i+1} = 0$, to

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

w przeciwnym razie

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

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

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 $\mathbf{A}$

- wyznaczamy najpierw macierz $\mathbf{P}$ taką, że

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

- wyznaczamy ciąg macierzy $\mathbf{H}^{(i)}$

- równocześnie z macierzą

$$\mathbf{H}^{(i+1)} = \left(\mathbf{Q}^{(i)}\right)^{\mathrm{T}} \mathbf{H}^{(i)} \mathbf{Q}^{(i)}$$

  wyznaczamy macierz pomocniczą

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

Wektory własne w metodzie **QR** (2)

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

$$\begin{aligned}
x_j^{(i)} &= 0, \quad j = n, n-2, \ldots, i+1, \\
x_i^{(i)} &= 1, \\
x_j^{(i)} &= -\frac{\left(\sum_{k=j+1}^{i} h_{jk}^{(N)} x_k^{(i)}\right)}{h_{jj}^{(N)} - h_{ii}^{(N)}}, \quad j = i-1, i-2, \ldots, 1.
\end{aligned}$$

Wektory własne macierzy **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 - trójdiagonalizacja 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)`
