

Wyznaczanie wartości i wektorów własnych macierzy (problem własny)

Plan wykładu:

1. Pojęcia podstawowe, definicje
2. Proste metody iteracyjne
 - a) metoda potęgowa dla wartości własnej o największym module
 - b) przyśpieszanie zbieżności
 - c) redukcja macierzy metodą Hottelunga
 - d) redukcja macierzy metodą Wielandta
5. Sprowadzanie macierzy symetrycznych/hermitowskich do postaci trójdiagonalnej
 - a) metoda Hauseholdera
 - b) metoda Lanczosa
6. Sprowadzanie macierzy symetrycznych/hermitowskich do postaci macierzy Hessenberga
7. Wyznaczanie wartości i wektorów własnych macierzy trójdiagonalnych i macierzy Hessenberga
 - a) metoda bisekcji
 - b) rozkład QR metodą Hauseholdera
8. Uogólniony problem własny
9. Rozkład SVD

Pojęcia podstawowe

Często przy tworzeniu modeli matematycznych wykorzystywanych do symulacji zjawisk fizycznych czy zachowania się układu, zachodzi potrzeba rozwiązania tzw. **problemu własnego** (np. rów. Schrodingera):

$$Ax_k = \lambda_k x_k \quad A = [a_{ij}]$$

- A jest macierzą kwadratową o $n \times n$
- x_k jest **wektorem własnym** macierzy odpowiadającym **wartości własnej** λ_k

$$a_{ij}, \lambda_k, x_m^{(k)} \in C$$

Nie zawsze układ równań, którego chcemy znaleźć rozwiązanie, przyjmuje tak prostą postać. Nierzadko mamy do czynienia z tzw. **uogólnionym problemem własnym**:

$$A \cdot x = \lambda B \cdot x$$

Jeśli macierz B jest nieosobliwa to problem uogólniony można przekształcić do postaci:

$$B^{-1}Ax = \lambda x$$

Liczbę λ nazywamy wartością własną macierzy jeśli istnieje taki niezerowy wektor x dla którego zachodzi:

$$Ax = \lambda x$$

Wektor x nazywamy (prawostronnym) wektorem własnym przynależnym do wartości własnej λ . Ciąg wszystkich wartości własnych nazywamy **widmem macierzy** A i oznaczamy: **Sp(A)**. Z powyższej definicji wynika:

$$(A - \lambda I)x = 0$$

Macierz $(A - \lambda I)$ jest osobliwa, więc:

$$\det(A - \lambda I) = 0$$

Wyznacznik ten jest wielomianem stopnia n zmiennej λ :

$$\begin{aligned} w(\lambda) &= \det(A - \lambda I) \\ &= (-1)^n (\lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_0) \end{aligned}$$

Każda wartość własna λ_k jest pierwiastkiem **wielomianu charakterystycznego** macierzy A .

Def. Wartości i wektory własne macierzy transponowanej A^T nazywamy **lewostronnymi wartościami** i **lewostronnymi wektorami własnymi** macierzy A .

Wyznacznik macierzy po jej transponowaniu nie ulega zmianie. Dlatego widmo macierzy A jest równe widmu lewostronnemu.

Tw. Jeżeli λ_p jest prawostronną wartością własną macierzy, a λ_l jest jej lewostronną wartością własną oraz gdy

$$\lambda_p \neq \lambda_l$$

Wówczas wektor własny \mathbf{x}_l jest ortogonalny do lewostronnego wektora własnego \mathbf{x}_p .

$$\begin{aligned}\mathbf{x}_l^T \mathbf{x}_p &= 0 \\ \mathbf{x}_l^T A \mathbf{x}_p &= \lambda_p \mathbf{x}_l^T \mathbf{x}_p \\ \mathbf{x}_l^T A \mathbf{x}_p &= (A^T \mathbf{x}_l)^T \mathbf{x}_p = \lambda_l \mathbf{x}_l^T \mathbf{x}_p\end{aligned}$$

$$(\lambda_l - \lambda_p) \mathbf{x}_l^T \mathbf{x}_p = 0 \quad (\lambda_p \neq \lambda_l)$$

Dla macierzy symetrycznej $\mathbf{A}=\mathbf{A}^T$ wektory własne są zarazem wektorami lewostronnymi. Jeżeli więc wektory własne przynależą do różnych wartości własnych to są do siebie **ortogonalne**.

Def. Macierze A i B są podobne jeśli istnieje nieosobliwa macierz podobieństwa P, że:

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

Tw. Jeżeli macierze A i B są podobne to mają identyczne widmo wartości własnych.

Tw. Macierz $Q_{m \times n}$ ($m \geq n$) nazywamy ortogonalną jeśli:

$$Q^T Q = I_{n \times n}$$

Tw. Jeżeli macierz $Q_{n \times n}$ jest ortogonalna to:

$$Q Q^T = I_{n \times n}$$

Tw. Macierz symetryczna A jest ortogonalnie podobna do macierzy diagonalnej D:

$$Q^T A Q = D$$

Tw. Wartości własne macierzy symetrycznej są rzeczywiste.

Def. Macierz o elementach zespolonych i własności

$$A = (A^T)^* \Rightarrow a_{ii} \in R$$

nazywamy macierzą hermitowską.

Wartości własne macierzy hermitowskiej są rzeczywiste
a wektory własne są do siebie ortogonalne.

Iloczyn skalarny (wewnętrzny)

$$\langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \mathbf{b} = [\alpha_1, \alpha_2, \dots, \alpha_n] \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

Iloczyn zewnętrzny (tensorowy)

$$\mathbf{a} \mathbf{a}^T \mathbf{b} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} [\beta_1, \beta_2, \dots, \beta_n] = \begin{bmatrix} \alpha_1 \beta_1 & \alpha_1 \beta_2 & \dots & \alpha_1 \beta_n \\ \alpha_2 \beta_1 & \alpha_2 \beta_2 & \dots & \alpha_2 \beta_n \\ \dots & \dots & \dots & \dots \\ \alpha_n \beta_1 & \alpha_n \beta_2 & \dots & \alpha_n \beta_n \end{bmatrix}$$

Uwaga: iloczyn zewnętrzny to operator macierzowy, możemy taki operator wykorzystać np. w metodzie Grama-Schmidta do ortogonalizacji (operator rzutowy)

Tw. Widmo macierzy ulega przesunięciu po dodaniu do niej macierzy jednostkowej pomnożonej przez liczbę:

widmo: $Sp(A) = \{\lambda_1, \lambda_2, \dots\}$

zostaje zastąpione przez: $Sp(A + cI) = Sp(A) + c$

$$Sp(A + cI) = \{\lambda_1 + c, \lambda_2 + c, \dots\}$$

Tw. (Cayleya-Hamiltona)

Jeśli

$$w(\lambda) = \det(A - \lambda I) = 0$$

jest równaniem charakterystycznym macierzy A to

$$w(A) = 0$$

Metody wyznaczania wartości i wektorów własnych macierzy

Macierze o elementach i wartościach własnych rzeczywistych

Macierze symetryczne (w tym **hermitowskie**)

Macierze niesymetryczne

Proste metody iteracyjne

Metoda potęgowa i jej modyfikacje

Redukcja macierzy do postaci trójdzielnej

Metoda Householdera Metoda Givensa

Metoda Lanczosa (macierze rzadkie)

Redukcja do postaci diagonalnej metodą Jacobiego

Redukcja macierzy do postaci Hessenberga

Rozkład QR

Wartości i wektory własne

Metoda potęgowa wyznaczania pojedynczych wartości własnych i wektorów własnych

Założmy że istnieje n liniowo niezależnych wektorów własnych macierzy A , stanowią **bazę przestrzeni liniowej**

$$\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n\}$$

Wówczas dla dowolnego wektora \mathbf{v}_0

$$\mathbf{v}_0 = \sum_{i=1}^n a_i \mathbf{x}_i$$

Jeśli λ_i stanowią wartości własne macierzy

$$A\mathbf{v}_0 = \sum_{i=1}^n a_i \lambda_i \mathbf{x}_i$$

$$\mathbf{v}_m = A^m \mathbf{v}_0 = \sum_{i=1}^n a_i \lambda_i^m \mathbf{x}_i$$

Zakładamy, że wartości własne tworzą ciąg

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

Jeśli λ_1 jest dominującą wartością własną, oraz wektor \mathbf{v}_0 ma składową w kierunku \mathbf{x}_1 to wówczas zachodzi

$$\lim_{m \rightarrow \infty} \frac{A^m \mathbf{v}_0}{\lambda_1^m} = a_1 \mathbf{x}_1$$

Z czego można wysnuć wniosek, że wartość własną można obliczyć następująco

$$\lambda_1 = \lim_{m \rightarrow \infty} \frac{\mathbf{y}^T \mathbf{v}_{m+1}}{\mathbf{y}^T \mathbf{v}_m}$$

Dla dowolnego wektora \mathbf{y} **nieortogonalnego** do \mathbf{x}_1 . Zazwyczaj \mathbf{y} ma 1 na pozycji elementu o największym module w \mathbf{v}_{m+1} a na pozostałych 0.

Jaka jest zbieżność metody?

$$\mathbf{v}_m = \lambda_1^m \left[a_1 \mathbf{x}_1 + \sum_{i=2}^n \left(\frac{\lambda_i}{\lambda_1} \right)^m a_i \mathbf{x}_i \right]$$

Zależy od $(\lambda_i/\lambda_1)^m$ ale również od współczynników a_i czyli od wyboru \mathbf{v}_0 . Jeśli wartość własna o największym module jest zespolona to ciąg nie jest zbieżny.

Jak wyznaczyć wektor własny \mathbf{x}_1 ?

Ponieważ
$$\mathbf{v}_m \approx \lambda_1^m a_1 \mathbf{x}_1$$

więc **unormowany wektor własny** będzie miał postać

$$\mathbf{x}_1 = \frac{\mathbf{v}_m}{|\mathbf{v}_m|}$$

Jeśli wartość własna jest pierwiastkiem wielokrotnym równania charakterystycznego to metoda jest zbieżna bo składnik z λ_1 dominuje

$$\mathbf{v}_m = A^m \mathbf{v}_0 = \lambda_1^m \sum_{i=1}^k a_i \mathbf{x}_i + \sum_{i=k+1}^n \lambda_i^m a_i \mathbf{x}_i$$

Uwaga: problem pojawia się gdy $\lambda_1 = -\lambda_j$ tj. identyczne moduły generują oscylacje (wtedy wybieramy ciąg wektorów \mathbf{v}_{2k})

Przyspieszanie zbieżności - iloraz Rayleigha

Jeśli macierz jest symetryczna to jej wektory własne są ortogonalne. Załóżmy, że są również ortonormalne

$$\mathbf{x}_i^T \mathbf{x}_j = \delta_{ij}$$

Wtedy

$$\mathbf{v}_m^T A \mathbf{v}_m = \mathbf{v}_m^T \mathbf{v}_{m+1} = \sum_{i=1}^n a_i^2 \lambda_i^{2m+1}$$

Oraz

$$\mathbf{v}_m^T \mathbf{v}_m = \sum_{i=1}^n a_i^2 \lambda_i^{2m}$$

co daje lepszą zbieżność niż wariant podstawowy metody

$$\frac{\mathbf{v}_m^T \mathbf{v}_{m+1}}{\mathbf{v}_m^T \mathbf{v}_m} = \lambda_1 + O \left[\left(\frac{\lambda_i}{\lambda_1} \right)^{2m} \right]$$

Wyznaczanie pozostałych wartości i wektorów własnych w metodzie potęgowej

a) metoda redukcji wektora

Jeśli znamy wartość własną o największym module to możemy wykorzystać ten fakt przy wyznaczaniu kolejnej największej co do modułu wartości własnej tj. λ_2

$$\begin{aligned}\mathbf{w}_m &= \mathbf{v}_{m+1} - \lambda_1 \mathbf{v}_m \\ &= \sum_{i=2}^n a_i (\lambda_i - \lambda_1) \lambda_i^m \mathbf{x}_i\end{aligned}$$

Ponieważ wektory \mathbf{v}_{m+1} oraz $\lambda_1 \mathbf{v}_m$ są bliskie więc metoda może być w wielu przypadkach nieużyteczna.

b) metoda zerowania składowej

Znając λ_1 możemy zdefiniować wektor

$$\mathbf{w}_0 = (A - \lambda_1 I) \mathbf{v}_0$$

Czyli z \mathbf{v}_0 usuwamy składową w kierunku \mathbf{x}_1 . Wektor \mathbf{w}_0 nie ma składowej w kierunku \mathbf{x}_1 więc ciąg \mathbf{w}_m powinien być zbieżny do λ_2 oraz \mathbf{x}_2 . Ze względu na błędy zaokrągleń jednak usuwa się co pewną ilość iteracji składową \mathbf{x}_1 tj.

$$\mathbf{z}_m = (A - \lambda_1 I) \mathbf{w}_m$$

Aby wyznaczyć kolejne wartości własne korzystamy z wektora

$$\mathbf{w}_0 = (A - \lambda_1 I)(A - \lambda_2 I)\mathbf{v}_0$$

Taki proces staje się mało wydajny dla kolejnych wartości własnych.

c) redukcja macierzy (najskuteczniejszy sposób)

Tw. Jeśli λ_1 jest wartością własną macierzy A i \mathbf{x}_1 odpowiadającym jej wektorem własnym oraz dla dowolnego wektora \mathbf{v} o własności

$$\mathbf{v}^T \mathbf{x}_1 = 1$$

Macierz zredukowana

$$W_1 = A - \lambda_1 \mathbf{x}_1 \mathbf{v}^T$$

ma te same wartości co macierz A oprócz λ_1 która jest zerem.

przykład: redukcja Hotellinga (macierze symetryczne)

- za wektor \mathbf{v} przyjmujemy lewy wektor własny przynależny do wartości własnej λ_1 , ale na ogół nie znamy lewych wektorów
- metoda jest więc skuteczna tylko w przypadku macierzy symetrycznych, wtedy lewe wektory są identyczne z prawymi

$$\mathbf{v} = \mathbf{x}_1$$

iloczyn zewnętrzny/tensorowy
(tworzymy operator macierzowy)

$$W_1 = A - \lambda_1 \mathbf{x}_1 \mathbf{x}_1^T$$

lub rekurencyjnie

$$W_0 = A$$

$$W_i = W_{i-1} - \lambda_{i-1} \mathbf{x}_{i-1} \mathbf{x}_{i-1}^T$$

$$i = 1, 2, \dots, n - 1$$

Redukcja macierzy gęstej do prostszej postaci - macierzy trójdzielnej lub macierzy Hessenberga

Macierz pierwotną przekształcamy iteracyjnie

$$A = A_0 \rightarrow A_1 \rightarrow A_2 \rightarrow \dots \rightarrow A_m$$

$$A_i = P_i^{-1} A_{i-1} P_i, \quad i = 1, 2, \dots, m$$

tak aby końcowa macierz B była macierzą podobną do A_0

$$\begin{aligned} B &= A_m = P^{-1} A P \\ P &= P_1 P_2 \dots P_m \end{aligned}$$

- widmo spektralne A i B takie samo

a następnie w łatwy sposób wyznaczamy wartości i wektory własne macierzy B

$$\begin{aligned} B\mathbf{y} &= \lambda\mathbf{y} \\ \mathbf{x} &= P\mathbf{y} = P_1 \dots P_m \mathbf{y} \\ A\mathbf{x} &= \lambda\mathbf{x} \end{aligned}$$

Uwaga:

nakład obliczeń potrzebnych do wyznaczenia \mathbf{x}_i oraz λ_i macierzy B powinien być jak najmniejszy

Redukcja macierzy hermitowskiej do postaci trójdzielnej metodą Householdera

Pierwotna macierz hermitowska

$$A^H = A = A_0 \quad A^H = (A^T)^*$$

Transformujemy

$$A_i = P_i^{-1} A_{i-1} P_i$$

Przy pomocy **macierzy Householdera**

$$P_i = P_i^H = P_i^{-1} = I - \beta_i \mathbf{u}_i \mathbf{u}_i^H$$

Dokonujemy transformacji A_{i-1} do A_i . A_{i-1} ma postać

$$A_{i-1} = \left[\begin{array}{c|c|c} J_{i-1} & \mathbf{c} & 0 \\ \hline \mathbf{c}^H & \delta & \mathbf{a}_i^H \\ \hline 0 & \mathbf{a}_i & \tilde{A}_{i-1} \end{array} \right]$$

$$= [\alpha_{jk}]$$

$$\mathbf{a}_i = \begin{bmatrix} \alpha_{i+1,i} \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{ni} \end{bmatrix}$$

$$\left[\begin{array}{c|c} J_{i-1} & \mathbf{c} \\ \hline \mathbf{c}^H & \delta_i \end{array} \right] = \left[\begin{array}{cccccc|cc} \delta_1 & \bar{\gamma}_2 & \cdot & \cdot & \cdot & 0 & 0 \\ \gamma_2 & \ddots & \ddots & & & \cdot & \cdot \\ \cdot & \ddots & \ddots & \ddots & & \cdot & \cdot \\ \cdot & & \ddots & \ddots & \ddots & \cdot & \cdot \\ \cdot & & & \ddots & \ddots & \cdot & \cdot \\ \cdot & & & & \ddots & \bar{\gamma}_{i-1} & 0 \\ 0 & \cdot & \cdot & \cdot & \gamma_{i-1} & \delta_{i-1} & \bar{\gamma}_i \\ \hline 0 & \cdot & \cdot & \cdot & 0 & \gamma_i & \delta_i \end{array} \right]$$

Zauważmy że w kolejnym kroku należy dokonać transformacji

- a) (n-i) elementowego wektora \mathbf{a}_i
- b) macierzy kwadratowej

$$\tilde{A}_{i-1}$$

rzędu (n-i)

Trzeba utworzyć macierz (operator) Householdera rzędu (n-i)

$$\tilde{P}_i$$

taką aby transformowała wektor \mathbf{a}_i

$$\tilde{P}_i \mathbf{a}_i = k \cdot \mathbf{e}_1$$

Macierz Householdera konstruujemy następująco

$$\tilde{P}_i = I - \beta \mathbf{u} \mathbf{u}^H$$

$$\beta = \begin{cases} \frac{1}{\sigma(\sigma + |\alpha_{i+1,i}|)} & \sigma \neq 0 \\ 0 & \sigma = 0 \end{cases}$$

$$\sigma = \|\mathbf{a}_i\|_2 = \sqrt{\sum_{j=i+1}^n |\alpha_{ji}|^2}$$

$$\mathbf{u} = \begin{bmatrix} e^{i\varphi}(\sigma + |\alpha_{i+1,i}|) \\ \alpha_{i+2,i} \\ \vdots \\ \alpha_{n,i} \end{bmatrix}$$

$$k = -\sigma e^{i\varphi} \iff \alpha_{i+1,i} = e^{i\varphi} |\alpha_{i+1,i}|$$

Dla tak konstruowanego operatora Householdera

$$P_i = \left[\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & \tilde{P}_i \end{array} \right] \quad \} i - 1$$

możemy dokonać transformacji macierzy A_{i-1}

$$P_i^{-1} A_{i-1} P_i = P_i A_{i-1} P_i = \left[\begin{array}{c|c|c} J_{i-1} & \mathbf{c} & 0 \\ \hline \mathbf{c}^H & \delta_i & \mathbf{a}_i^H \tilde{P}_i \\ \hline 0 & \tilde{P}_i \mathbf{a}_i & \tilde{P}_i \tilde{A}_{i-1} \tilde{P}_i \end{array} \right] = \left[\begin{array}{ccc|c|c|c} \delta_1 & \bar{\gamma}_2 & & 0 & 0 & \\ \gamma_2 & \ddots & \ddots & & \vdots & 0 \\ & \ddots & \ddots & \bar{\gamma}_{i-1} & 0 & \\ 0 & & \gamma_{i-1} & \delta_{i-1} & \bar{\gamma}_i & \\ \hline 0 & \dots & 0 & \gamma_i & \delta_i & \tilde{\gamma}_{i+1} 0 \dots 0 \\ \hline & & & & \gamma_{i+1} & \\ & & & & 0 & \\ & & & & \vdots & \tilde{P}_i \tilde{A}_{i-1} \tilde{P}_i \\ & & 0 & & 0 & \end{array} \right]$$

$$k = -\sigma e^{i\varphi} \iff \alpha_{i+1,i} = e^{i\varphi} |\alpha_{i+1,i}| \quad \gamma_{i+1} = k$$

Pojawia się problem: mnożenie macierzy wymaga wykonania $O(n^3)$ operacji.

Jak dokonać efektywnego mnożenia macierzy (zwykle jest nieekonomiczne)?

$$\tilde{P}_i = I - \beta \mathbf{u} \mathbf{u}^H$$

$$\begin{aligned} \tilde{P}_i \tilde{A}_{i-1} \tilde{P}_i &= (I - \beta \mathbf{u} \mathbf{u}^H) \tilde{A}_{i-1} (I - \beta \mathbf{u} \mathbf{u}^H) \\ &= \tilde{A}_{i-1} - \beta \tilde{A}_{i-1} \mathbf{u} \mathbf{u}^H - \beta \mathbf{u} \mathbf{u}^H \tilde{A}_{i-1} \\ &\quad + \beta^2 \mathbf{u} \mathbf{u}^H \tilde{A}_{i-1} \mathbf{u} \mathbf{u}^H \end{aligned}$$

Definiujemy dwa wektory pomocnicze

$$\begin{aligned} \mathbf{p} &= \beta \tilde{A}_{i-1} \mathbf{u} \\ \mathbf{q} &= \mathbf{p} - \frac{\beta}{2} (\mathbf{p}^H \mathbf{u}) \mathbf{u} \end{aligned}$$

ponieważ

$$\beta \geq 0$$

oraz z faktu, że \tilde{A}_{i-1} jest hermitowska wynika

$$\mathbf{p}^H \mathbf{u} = \beta \mathbf{u}^H \tilde{A}_{i-1} \mathbf{u} = (\mathbf{p}^H \mathbf{u})^H$$

Wykorzystajmy wektory \mathbf{p} oraz \mathbf{q} podczas transformacji.

Przy użyciu \mathbf{p} i \mathbf{q} dokonujemy redukcji liczby wymaganych operacji

$$\begin{aligned}
 \tilde{P}_i \tilde{A}_{i-1} \tilde{P}_i &= \tilde{A}_{i-1} - \mathbf{p}\mathbf{u}^H - \mathbf{u}\mathbf{p}^H + \beta \mathbf{u}\mathbf{p}^H \mathbf{u}\mathbf{u}^H \\
 &= \tilde{A}_{i-1} - \mathbf{u} \left(\mathbf{p} - \frac{\beta}{2} (\mathbf{p}^H \mathbf{u}) \mathbf{u} \right)^H \\
 &\quad - \left(\mathbf{p} - \frac{\beta}{2} (\mathbf{p}^H \mathbf{u}) \mathbf{u} \right) \mathbf{u}^H \\
 &= \tilde{A}_{i-1} - \mathbf{u}\mathbf{q}^H - \mathbf{q}\mathbf{u}^H \quad \longrightarrow \quad \boxed{\tilde{P}_i \tilde{A}_{i-1} \tilde{P}_i = \tilde{A}_{i-1} - \mathbf{u}\mathbf{q}^H - \mathbf{q}\mathbf{u}^H}
 \end{aligned}$$

Co uzyskaliśmy?

Po redukcji, transformacja A_{i-1} wymaga już tylko $O[(n-i)^2]$ operacji - jest ekonomiczna.

Przeprowadzając proces przekształcania macierzy do końca dostaniemy macierz B w postaci trójdzielnej

$$B = A_{n-2} = \begin{bmatrix} \delta_1 & \tilde{\gamma}_2 & \cdots & 0 \\ \gamma_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tilde{\gamma}_n \\ 0 & \cdots & \gamma_n & \delta_n \end{bmatrix}$$

- metoda Householdera jest stabilna numerycznie
- kolejnym krokiem będzie wyznaczenie wartości i wektorów własnych B

Redukcja macierzy (rzadkiej) hermitowskiej do postaci trójdzielnej metodą Lanczosa

Naszym zadaniem jest znalezienie wartości i wektorów własnych macierzy stopnia n

$$\begin{aligned}\lambda_1, \lambda_2, \dots, \lambda_n &\in \mathbb{R} \\ \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n &\in \mathbb{C}^n\end{aligned}$$

Jeśli jednak n jest bardzo duże (np. rzędu $\sim 10^5$) a nas interesuje jedynie mały wycinek widma wartości (wektorów) własnych (np. $m=50$) to **wówczas należy zredukować macierz do postaci trójdzielnej rzędu m .**

W metodzie Lanczosa wykorzystujemy do tego celu **podprzestrzeń Kryłowa**

$$\begin{aligned}\mathbf{q} &\in \mathbb{C}^n \\ K_m(\mathbf{q}, A) &= \text{span}[\mathbf{q}, A\mathbf{q}, \dots, A^{m-1}\mathbf{q}] \\ m &\geq 1 \\ K_0(\mathbf{q}, A) &= \{\mathbf{0}\} \\ \dim K_m &= m\end{aligned}$$

- bazę w podprzestrzeni generujemy przy użyciu \mathbf{q} i A

Zakładamy że wektory rozpinające podprzestrzeń

$$K_m = \text{span}[\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m]$$

są **ortonormalne**

$$\mathbf{q}_i^H \mathbf{q}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Metoda Lanczosa

jest **metodą iteracyjną**, w każdej iteracji poszukujemy nowego wektora bazowego (wymiar podprzestrzeni zwiększa się o 1)

Sposób postępowania:

Wybieramy dowolny wektor startowy $\mathbf{q}_1 \in C^n, \mathbf{q}_1 \neq \mathbf{0}$

oraz zakładamy warunek $\gamma_1 \mathbf{q}_0 = 0$

następnie iteracyjnie wyznaczamy kolejne wektory bazy zgodnie z wzorem

$$A\mathbf{q}_i = \mathbf{y} = \gamma_i \mathbf{q}_{i-1} + \delta_i \mathbf{q}_i + \gamma_{i+1} \mathbf{q}_{i+1} \quad i \geq 1$$

gdzie współczynniki γ oraz δ znajdujemy wykorzystując **ortogonalność** wektorów \mathbf{q}

$$\delta_i = \mathbf{q}_i^H A\mathbf{q}_i$$

$$\mathbf{r}_i = A\mathbf{q}_i - \delta_i \mathbf{q}_i - \gamma_i \mathbf{q}_{i-1} \quad \longrightarrow \quad \mathbf{q}_{i+1} = \frac{\mathbf{r}_i}{\gamma_{i+1}} \Leftrightarrow \gamma_{i+1} \neq 0$$

$$\gamma_{i+1} = \|\mathbf{r}_i\|$$

Procedura zatrzyma się gdy $\gamma_{i+1} = 0$

Wówczas wymiar wygenerowanej podprzestrzeni $m = \max_i \dim K_i(\mathbf{q}, A)$

Schemat iteracyjny można zapisać macierzowo (łatwo dostrzec podobieństwo J i A)

$$Q_i = [\mathbf{q}_1, \dots, \mathbf{q}_i]$$

Q jest macierzą ortogonalną

$$Q_i^H Q_i = I$$

$$J_i = \begin{bmatrix} \delta_1 & \gamma_2 & & 0 \\ \gamma_2 & \delta_2 & \ddots & \\ & \ddots & \ddots & \gamma_i \\ 0 & & \gamma_i & \delta_i \end{bmatrix}$$

- tę macierz poddamy diagonalizacji

$$AQ_i = Q_i J_i \quad (\text{"+"} \quad \gamma_{i+1}), \quad i = 1, 2, \dots, m$$

$$Q_i^H A Q_i = J_i$$

Może się okazać (rzadko), że zostanie spełniony warunek

$$\gamma_{i+1} = 0$$

wówczas dostaniemy dokładne przybliżenie **m największych wartości własnych**

$$AQ_m = Q_m J_m$$

ale dla wybranego startowego \mathbf{q}_0 nie znajdziemy więcej wektorów własnych (wówczas schemat restartujemy używając innego wektora startowego).

- wartości własne J_m stanowią **oszacowanie** wartości własnych A

$$J_m \mathbf{z} = \lambda \mathbf{z}, \quad \mathbf{z} \neq \mathbf{0}$$

$$\mathbf{x} = Q_m \mathbf{z} \neq \mathbf{0}$$

$$A\mathbf{x} = A Q_m \mathbf{z} = Q_m J_m \mathbf{z} = \lambda Q_m \mathbf{z} = \lambda \mathbf{x}$$

Macierz Q_n jest macierzą ortogonalną, a macierz J_n jest macierzą trójdziagonalną „**podobną**” do A. Podobieństwo nie jest zupełne tzn. tylko część wartości własnych J może być zbliżona do wartości A - czyli mogą stanowić ich dobre oszacownie.

Uwagi

- 1) po redukcji A do J należy zdiagonalizować J (bisekcja/rozkład QR)
- 2) wartości własne J (a raczej ich część) mogą stanowić dobre przybliżenie wartości własnych A o **największym module - zbieżność do wartości skrajnych widma jest największa**
- 3) w celu wyznaczenia wartości własnych z innych obszarów widma należy użyć metody **iteracji odwrotnej** z jednoczesnym przesunięciem widma
- 4) Ze względu na błędy zaokrągleń wektory rozpinające podprzestrzeń Kryłowa przestają być ortogonalne (gdy rośnie wymiar podprzestrzeni). Konieczne staje się przeprowadzenie tzw. **ortogonalizacji Grama-Schmidta** nowego wektora bazy do już wyznaczonych co jest kosztowne

$$\tilde{\mathbf{q}}_{i+1} \rightarrow \mathbf{q}_{i+1} = \tilde{\mathbf{q}}_{i+1} - \sum_{j=1}^i (\mathbf{q}_j^H \tilde{\mathbf{q}}_{i+1}) \mathbf{q}_j$$

Metoda iteracji odwrotnej + przesuwanie widma

$$A\mathbf{x} = \lambda B\mathbf{x} \quad /(-\sigma B\mathbf{x})$$

$$\begin{aligned} \sigma &\in C \\ \sigma &\neq \lambda \end{aligned}$$

- przesunięcie w okolice interesującego nas punktu (może to być np. $\sigma = 0$)

$$(A - \sigma B)\mathbf{x} = (\lambda - \sigma)B\mathbf{x}, \quad (A - \sigma B)^{-1} \cdot /$$

$$\mathbf{x} = (\lambda - \sigma)(A - \sigma B)^{-1}B\mathbf{x} \quad / \cdot \frac{1}{\lambda - \sigma}$$

$$\underbrace{\frac{1}{\lambda - \sigma}}_{\mu} \mathbf{x} = \underbrace{(A - \sigma B)^{-1}B}_{C} \mathbf{x}$$

$$C\mathbf{x} = \mu\mathbf{x}$$

- do rozwiązania pozostaje nam standardowy problem, **ze zmienionym widmem wartości własnych**
- największa zbieżność dla wartości zbliżonych do sigmy

$$\mu = \frac{1}{\lambda - \sigma} \longrightarrow$$

λ_i	5.1	5.0	4.0	0.1	-0.1	-10	-20
$\sigma = 0 \rightarrow \mu_i$	0.196	0.2	0.25	10	-10	-0.1	-0.05

przykład: zastosowanie w metodzie **Lanczosa** (ale w metodzie potęgowej też zadziała)

Iteracyjnie wyznaczamy kolejne wektory $q^{(i)}$ czyli:

$$Cq^{(i)} = q^{(i+1)} \quad \rightarrow \quad (A - \sigma B)^{-1} Bq^{(i)} = q^{(i+1)}, \quad (A - \sigma B)/$$

$$Bq^{(i)} = b^{(i)} = (A - \sigma B)q^{(i+1)}$$

W każdej iteracji należy rozwiązać układ równań

$$(A - \sigma B)q^{(i+1)} = b^{(i)}$$

Układ równań rozwiążemy **szybko** jeśli:

- 1) jeśli dysponujemy rozkładem LU macierzy $(A - \sigma B)$
rozkład macierzy wykonujemy tylko raz - na początku
- 2) jeśli macierz jest rzadka i możemy zastosować metody iteracyjne
(zazwyczaj wykorzystuje się jeszcze preconditioning - przyspieszanie)

Rozwiązanie układu w każdej iteracji jest tanie dzięki czemu zastosowanie modyfikacji „**shift-invert**” pozwala zmniejszyć ilość iteracji około **10-krotnie** (a w niektórych zagadnieniach nawet 100-krotnie - gdy wartości własne są bliskie 0).

Redukcja macierzy do postaci macierzy **Hessenberga** (górnej) metodą eliminacji Gaussa

Naszym zadaniem jest przekształcenie w sposób rekurencyjny macierzy A

$$A = A_0 \rightarrow A_1 \rightarrow \dots \rightarrow A_{n-2} = B_H$$
$$A_i = P_i^{-1} A_{i-1} P_i$$

do postaci Hessenberga

$$B_H = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{bmatrix} = T + U$$

T - macierz trójkątna

U - macierz trójkątna górna

Uwagi:

1) każda macierz kwadratowa jest ortogonalnie podobna do macierzy Hessenberga

2) ponieważ macierz A jest przekształcana przez podobieństwo więc przekształcenie macierzy **hermitowskiej** do postaci Hessenberga prowadzi do uzyskania macierzy trójkątno-diagonalnej ze względu na symetrię A . Macierz trójkątno-diagonalna jest hermitowska.

b) macierzy eliminacji

$$G_j = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{j+1,j} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{nj} & & & 1 \end{bmatrix}$$

z warunkiem $l_{ij} \leq 1$

Macierz odwrotna G_j^{-1}

$$G_j^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{j+1,j} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nj} & & & 1 \end{bmatrix}$$

Jakie są skutki mnożenia macierzy G_j oraz π_{rs} z A ?

1. $\pi_{rs}^{-1} A$ - zamienia wiersze r i s w A
2. $A\pi_{rs}$ - zamienia kolumny r i s w A
2. $G_j^{-1} A$ - odejmuje wiersz j -ty przemnożony przez l_{rj} od wiersza r w A
3. AG_j - przemnaża kolumnę r -tą przez l_{rj} a następnie dodaje do kolumny j -tej w A

Założmy, że w macierzy A_{i-1} (i-1) pierwszych kolumn posiada postać Hessenberga

$$A_{i-1} = \left[\begin{array}{cccccc|c|cccc} * & \cdot & \cdot & \cdot & \cdot & * & * & \cdot & \cdot & \cdot & * \\ * & \cdot & & & & \cdot & \cdot & & & & \cdot \\ & \cdot & \cdot & & & \cdot & \cdot & & & & \cdot \\ & & \cdot & \cdot & & \cdot & \cdot & & & & \cdot \\ & & & \cdot & \cdot & \cdot & \cdot & & & & \cdot \\ 0 & & & & * & * & * & \cdot & \cdot & \cdot & * \\ \hline 0 & \cdot & \cdot & \cdot & 0 & * & * & \cdot & \cdot & \cdot & * \\ \hline & & & & & & & \cdot & & & \cdot \\ & & & & & & & \cdot & & & \cdot \\ & & & & & & & \cdot & & & \cdot \\ 0 & & & & & & * & \cdot & \cdot & \cdot & * \end{array} \right] = \left[\begin{array}{c|c|c} B_{i-1} & \mathbf{d} & \hat{A}_{i-1} \\ \hline \mathbf{c} & \delta_i & \mathbf{b} \\ \hline 0 & \mathbf{a} & \tilde{A}_{i-1} \end{array} \right] \quad \mathbf{a} = \left[\begin{array}{c} \alpha_{i+1,i} \\ \vdots \\ \alpha_{ni} \end{array} \right]$$

Dokonujemy przekształcenia

$$A_i = P_i^{-1} A_{i-1} P_i$$

macierz przekształcenia jest zdefiniowana w poniższy sposób

$$P_i = \pi_{r,i+1} G_{i+1} \quad r \geq i + 1$$

$$P_i^{-1} = G_{i+1}^{-1} \pi_{r,i+1}^{-1}$$

Macierz przekształcenia P_i nie zmienia postaci Hessenberga w $(i-1)$ pierwszych kolumnach.

Zmienia natomiast kolumnę **i-tą** tj. zeruje elementy w tej kolumnie od $i+2$ do n .

$$P_i^{-1} \cdot \begin{bmatrix} \mathbf{d} \\ \delta_i \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{d} \\ \delta_i \\ \bar{\mathbf{a}} \end{bmatrix} \quad \bar{\mathbf{a}} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Aby określić które wiersze należy zamienić wierszami, najpierw trzeba określić element o największym module w **a**

$$|\alpha_{ri}| = \max_{i+1 \leq j \leq n} |\alpha_{ji}|, \quad r \geq i+1$$

Po wyznaczeniu wartości r zamieniamy miejscami wiersze ($r \leftrightarrow i+1$) oraz kolumny ($r \leftrightarrow i+1$) w A_{i-1}

$$A' = \pi_{r,i+1}^{-1} A_{i-1} \pi_{r,i+1} = [\alpha'_{jk}]$$

Dzięki temu element o największym module w **wektorze a** znajduje się w nim na pierwszym miejscu.

Jak określić G_{j+1} ?

Warunek: trzeba wyzerować wszystkie elementy w **a** poza jego pierwszym elementem (standardowe postępowanie w metodzie Gaussa)

$$l_{j,i+1} = \begin{cases} \frac{\alpha'_{ji}}{\alpha'_{i+1,i}} & \alpha_{i+1,i} \neq 0 \\ 0 & \alpha_{i+1,i} = 0 \end{cases}$$

$$j = i+2, i+3, \dots, n$$

Przeprowadzając drugi etap przekształcenia

$$A_i = G_{i+1}^{-1} A' G_{i+1} = P_i^{-1} A_{-1} P_i$$

dostajemy macierz Hessenberga w pierwszych i kolumnach macierzy A_i .

Uwagi:

1) po $n-2$ iteracjach macierz A zostaje przekształcona do postaci **Hessenberga** (górnej)

2) metoda eliminacji Gaussa jest stabilna numerycznie i wymaga wykonania

$$M = \frac{2}{3}n^3 + O(n^2)$$

operacji mnożenia

3) redukcję macierzy do postaci Hessenberga można przeprowadzić także przy użyciu metod - **Hauseholdera (operator rzutowy)** lub **Givensa (eliminacja przez obroty)**

W metodzie **Givensa** (obroty) trzeba wykonać

$$M = \frac{10}{3}n^3 + O(n^2)$$

a w metodzie **Hausholdera**

$$M = \frac{5}{3}n^3 + O(n^2)$$

operacji mnożenia

Wyznaczanie wartości własnych macierzy

Przekształciliśmy macierz

- symetryczną do postaci **trójdiagonalnej** (Hausholder/Lanczos)
- macierz niesymetryczną do postaci **Hessenberga** (Gauss/Hauseholder/Givens)

Dla tych macierzy poszukujemy wartości i wektorów własnych.

Wyznaczanie wartości własnych macierzy trójdzielnej **metodą bisekcji**

Po przekształceniu macierzy A do postaci

$$J = \begin{bmatrix} \delta_1 & \bar{\gamma}_2 & & 0 \\ \gamma_2 & \ddots & \ddots & \\ & \ddots & \ddots & \bar{\gamma}_n \\ 0 & & \gamma_n & \delta_n \end{bmatrix}$$

ta macierz jest **nieredukowalna** jeśli spełniony jest warunek

$$\gamma_i \neq 0, \quad i = 2, \dots, n$$

w przeciwnym razie można ją zapisać w postaci

$$J = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_k \end{bmatrix}$$

i rozwiązywać problem dla mniejszych macierzy J_1, J_2, J_3, \dots
ponieważ widmo wartości J oraz ciągu J_i jest identyczne.

W **metodzie bisekcji** wykorzystujemy wielomian charakterystyczny macierzy - $W(\lambda)$.

Sposób wyznaczania wartości wielomianu charakterystycznego J_i :

1. zakładamy dowolną wartość λ

2. obliczamy wartość $W(\lambda)$ rozwijając wyznacznik względem kolejnych kolumn macierzy

$$\omega_i(\lambda) = \det(J_i - \lambda I)$$

co prowadzi do procedury rekurencyjnej

$$\begin{aligned}\omega_0(\lambda) &= 1 \\ \omega_1(\lambda) &= \delta_1 - \lambda \\ \dots & \dots \dots \\ \omega_i(\lambda) &= (\delta_i - \lambda)\omega_{i-1}(\lambda) - |\gamma_i|^2 \omega_{i-2}(\lambda) \\ i &= 2, 3, \dots, n \\ W(\lambda) &= \omega_n(\lambda)\end{aligned}$$

macierz jest hermitowska więc $\delta_i, |\gamma_i|^2 \in R$

i wszystkie wartości pośrednie też są rzeczywiste

Wyznaczanie wartości własnej w metodzie biecacji.

Wybieramy dowolną liczbę λ i obliczamy wartość wielomianu charakterystycznego rekurencyjnie zachowując informacje o ciągu

$$\omega_0(\lambda), \omega_1(\lambda), \dots, \omega_n(\lambda)$$

a następnie korzystamy z poniższych twierdzeń:

Tw. Jeżeli elementy $\gamma_2, \gamma_3, \dots, \gamma_n$ (pozadiagonalne) są niezerowe, to wartości własne macierzy J są pojedyncze.

Tw. Jeżeli elementy $\gamma_2, \gamma_3, \dots, \gamma_n$ (pozadiagonalne) są niezerowe, to ciąg wartości

$$\omega_0(\lambda), \omega_1(\lambda), \dots, \omega_n(\lambda)$$

spełnia poniższe warunki:

a) jeżeli $\omega_i(\lambda)=0$ dla pewnego $i < n$, to przyjmujemy

$$\omega_{i-1}(\lambda) \cdot \omega_{i+1}(\lambda) < 0$$

b) jeżeli $\omega_n(\lambda)=\omega(\lambda)$ jest różne od 0, to liczba zmian znaków sąsiednich liczb w ciągu **jest równa liczbie wartości własnych macierzy J mniejszych od λ .**

c) jeżeli $\omega_n(\lambda)=0$, to λ jest wartością własną macierzy J , a ponadto jest tyle wartości własnych mniejszych niż λ , ile nastąpiło zmian znaków w ciągu

zalety:

- metoda bisekcji jest bardzo dokładna
- umożliwia obliczenie wartości własnej o określonym indeksie k
liczba iteracji potrzebna do wyznaczenia λ_k wynosi

$$IT = \log_2 \frac{\beta_0 - \alpha_0}{\rho}$$

α_0, β_0 – przedział poszukiwań wartości własnej,
 ρ – dokładność wyznaczenia wartości własnej

wady:

- uzyskiwanie dużych wartości ciągu: $\omega_0(\lambda), \omega_1(\lambda), \dots, \omega_n(\lambda)$,
jeśli λ znacznie różni się od wartości własnych J

Wektory własne w metodzie bisekcji

Znając k-tą wartość własną macierzy J **wektor własny** \mathbf{x}_k wyznaczamy według wzorów:

$$x_1 = 1$$

$$x_2 = \frac{\lambda_k - \delta_1}{\gamma_2}$$

$$x_{i+1} = \frac{(\lambda_k - \delta_i)x_i - \gamma_i x_{i-1}}{\gamma_{i+1}}, \quad i = 2, 3, \dots, n-1$$

Wyznaczanie wartości i wektorów własnych przy użyciu rozkładu QR

- rozkład QR dla macierzy gęstych wymaga dużego nakładu obliczeniowego $\sim O(n^3)$.
- natomiast macierze: trójdzielna oraz Hessenberga to macierze rzadkie i rozkład QR uzyskamy wykonując $O(n)$ operacji.
- rozkład QR wykonujemy maksymalnie $n-1$ krotnie (wyjaśnienie na kolejnym slajdzie)
np. przy użyciu metody Householdera (operator rzutowy)

Prześledźmy proces iteracyjnego przekształcania macierzy ($A=Tri=Hess$)

$$A_0 = A \quad - \text{macierz pierwotna}$$

$$A_i = Q_i R_i \quad Q_i^{-1} \cdot /$$

$$Q_i^{-1} A_i = R_i \quad / \cdot Q_i$$

$$Q_i^{-1} A_i Q_i = R_i Q_i = A_{i+1} = Q_{i+1} R_{i+1}$$

$$Q_i^{-1} A_i Q_i = Q_{i+1} R_{i+1} \quad Q_{i+1}^{-1} \cdot /$$

$$Q_{i+1}^{-1} Q_i^{-1} A_i Q_i = R_{i+1} \quad / \cdot Q_{i+1}$$

$$Q_{i+1}^{-1} Q_i^{-1} A_i Q_i Q_{i+1} = R_{i+1} Q_{i+1} = A_{i+2}$$

Cofając się wstecz do $i=0$ otrzymamy

$$A_{k+1} = Q_k^{-1} Q_{k-1}^{-1} \dots Q_1^{-1} A Q_1 Q_2 \dots Q_k$$

dla skrócenia zapisu wprowadźmy oznaczenia

$$P = P_k = Q_1 Q_2 \dots Q_k$$

$$P^{-1} = P_k^{-1} = Q_k^{-1} Q_{k-1}^{-1} \dots Q_1^{-1}$$

Jeśli proces przekształcania macierzy doprowadzimy do końca wówczas otrzymamy

$$P^{-1} A P = A_{k+1} = H$$

gdzie: macierz H jest macierzą górną-trójkątną z wartościami własnymi na diagonalu

$$\lambda_i = h_{ii}$$

Wektor własny \mathbf{y}_k macierzy H odpowiadający wartości własnej λ_k wyznaczamy stosując wzory

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

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

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

Dysponując wektorami \mathbf{x}_k możemy wyznaczyć wektory własne \mathbf{y}_k wyjściowego problemu (macierz trójdzielna/Hessenberga - dla oryginalnego problemu postępujemy podobnie)

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

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

$$P^{-1}AP\mathbf{x} = H\mathbf{x} = \lambda\mathbf{x}$$

$$A(P\mathbf{x}) = \lambda P\mathbf{x}$$

$$\mathbf{y} = P\mathbf{x}$$

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

Uogólniony problem własny

- uogólniony problem własny definiujemy następująco

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

- najprościej byłoby przekształcić powyższe równanie tak, aby przeprowadzić je do zwykłego problemu własnego

$$B^{-1}A\mathbf{x} = C\mathbf{x} = \lambda\mathbf{x}$$

- problemem pozostaje jednak jak znaleźć B^{-1} ?
- w przypadku, gdy B oraz A są macierzami symetrycznymi możemy posłużyć się **rozkładem LL^T** (w ogólnym przypadku można skorzystać z rozkładu LU)

$$B = LL^T$$

$$BB^{-1} = I = LL^T(L^T)^{-1}L^{-1}$$

$$B^{-1} = (L^T)^{-1}L^{-1}$$

- wówczas wykorzystując rozkład LL^T można znaleźć macierz podobną do $B^{-1}A$

$$\begin{aligned}L^T(B^{-1}A)(L^T)^{-1} &= L^T(L^T)^{-1}L^{-1}A(L^T)^{-1} \\ &= L^{-1}A(L^T)^{-1} \\ &= G\end{aligned}$$

Dzięki temu przekształceniu, macierz G jest symetryczna jak A i posiada identyczne widmo wartości własnych (ale inne wektory własne).

$$G\mathbf{y} = \lambda\mathbf{y}$$

Jak znaleźć G ? Najpierw należy znaleźć macierz F

$$F = A(L^T)^{-1}$$

rozwiązując układ równań

$$FL^T = A \implies LF^T = A^T = A$$

a następnie wyznaczamy G

$$G = L^{-1}F$$

rozwiązując układ równań

$$LG = F$$

Rozkład LL^T wymaga wykonania $n^3/6$ mnożeń a wyznaczenie macierzy G $(2/3)n^3$. Macierz G jest symetryczna więc w celu wyznaczenia jej wartości i wektorów własnych korzystamy z metod przeznaczonych dla tej klasy macierzy.

Wektory własne macierzy A wyznaczamy przekształcając wektory macierzy G lub rozwiązując układ

$$L^T\mathbf{x} = \mathbf{y}$$

Przykład. Rozwiązanie równania Schrodingera przy użyciu bazy funkcyjnej (metoda Galerkina)

- zaczynamy od zapisania równania własnego (definiujemy problem)

$$\hat{h}\Psi = E\Psi$$

- problem ma postać różniczkową określoną przez operator energii

$$\hat{h} = \frac{(\hat{\mathbf{p}} + e\mathbf{A})^2}{2m^*} + \frac{m^*}{2}(\omega_x^2 x^2 + \omega_y^2 y^2)$$

$\hat{\mathbf{p}} = \left[-i\hbar \frac{\partial}{\partial x}, -i\hbar \frac{\partial}{\partial y} \right]$

$\mathbf{A}(x, y) = B[0, x, 0]$

$$\hat{h} = -\frac{\hbar^2}{2m^*} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + i \frac{eB\hbar}{m^*} \left(x \frac{\partial}{\partial y} \right) + \frac{e^2 B^2}{2m^*} x^2 + \frac{m^*}{2} (\omega_x^2 x^2 + \omega_y^2 y^2)$$

- problem różniczkowy musimy zamienić w problem algebraiczny

- rozwiązania możemy poszukiwać w bazie funkcji (szukamy współczynników rozwinięcia)

$$\Psi(x, y) = \sum_{m=1}^N c_m f_m(x, y)$$

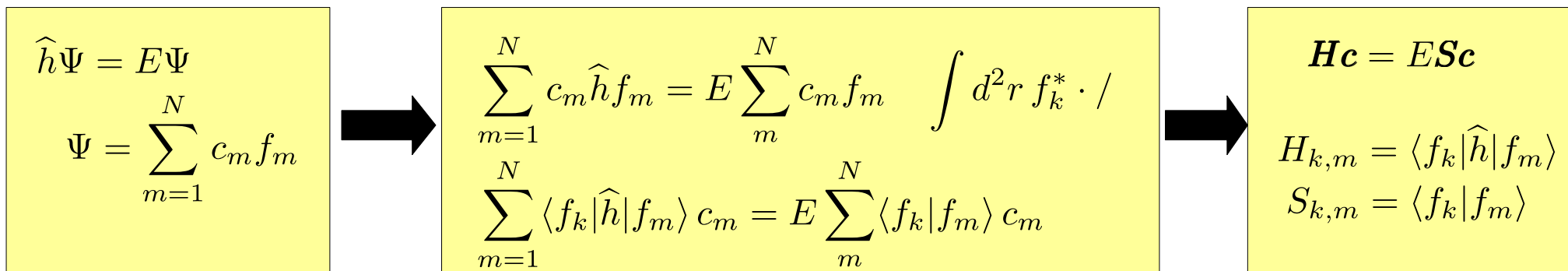
- w przykładzie użyto funkcji Gaussa

$$f_m(x, y) = \underbrace{\exp\left[-\frac{(x-x_m)^2 + (y-y_m)^2}{2\sigma^2}\right]}_{\text{funkcja Gaussa}} \underbrace{\exp\left[-i\frac{eB}{\hbar}(x+x_m)(y-y_m)\right]}_{\text{czynnik fazowy}}$$

problem różniczkowy

algebraizacja problemu

macierzowy problem uogólniony



diagonalizacja dostarcza: E_m, \mathbf{c}_m

Jeśli to możliwe całkowanie wykonujemy analitycznie przy użyciu programów do obliczeń symbolicznych typu:

Mathematica czy **Maple**

lub numerycznie - wtedy wyznaczanie elementów macierzowych może zająć więcej czasu niż diagonalizacja

$$Hc = ES c$$

$$H_{k,m} = \langle f_k | \hat{h} | f_m \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f_k^*(x, y) \hat{h}(x, y) f_m(x, y)$$

$$S_{k,m} = \langle f_k | f_m \rangle = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f_k^*(x, y) f_m(x, y)$$

Element H_{km}

```

t1 = xk**2
t2 = yk**2
t3 = xm**2
t4 = ym**2
t5 = cmplx(0.D0,4.D0)
t8 = sig**2
t14 = ym*wb
t17 = cmplx(0.D0,-4.D0)
t23 = xk*xm
t24 = 2*t23
t30 = q**2
t31 = wb**2
t32 = t31*t30
t33 = t8**2
t35 = t4*t33*t32
t38 = yk*t33*ym*t32
t41 = t33*t2*t32
t42 = ym*yk
t43 = 2*t42
t45 = t3*t33*t32
t48 = xk*xm*t33*t32
t51 = t1*t33*t32
t52 = t1+t2+t3+t4+ym*xm*t8*wb*q*t5+t8*t14*q*xk*t5+t8*yk*wb*q*xm*t1
#7-t24+yk*xk*t8*wb*q*t17+t35-2*t38+t41-t43+t45-2*t48+t51
t53 = 1/t8
t56 = exp(-t53*t52/4)
t58 = t33*sm
t63 = t30**2
t64 = t31**2
t65 = t64*t63
t66 = t33**2
t71 = sm*t5
t72 = t33*t8
t73 = ax*t72
t74 = t73*t71
t76 = t14*xk*q
t79 = t14*xm*q
t81 = bx*t72
t82 = t81*t71
t84 = sm*t17
t85 = t81*t84
t86 = yk*q
t88 = wb*xm*t86
t90 = t73*t84
t93 = t66*sm
t94 = ax*t93
t99 = wb*xk*t86
t101 = bx*t93
t120 = 2*t4*t32*t94+t99*t82+t3+t4+2*t1*t32*t101-4*t42*t32*t94+2*t2
#*t32*t94-4*t23*t32*t101-4*bx*yk*ym*t58-2*t35-2*t41
t159 = sm*t72
t165 = -2*t1*ax*t58-2*t3*ax*t58+t4*t66*t65+t2*t66*t65+4*t48+t99*t9
#0-4*t8-4*t31*t30*t72-4*bx*t159-4*ax*t159+t79*t85
t173 = -1/sm*t53*(-4*ax*xk*xm*t58+t1-2*yk*ym*t66*t65+t76*t74+t2+t7
#9*t74+t76*t82-t24+t88*t85+t88*t90+t120-2*t45-2*t51-2*xk*xm*t66*t65
#+4*t38-t43+2*t3*t32*t101-2*t2*bx*t58-2*t4*bx*t58+t3*t66*t65+t1*t66
#*t65+t165)*0.3141592653589793D1*t56/8

```

Element S_{km}

```

t1 = xk**2
t2 = yk**2
t3 = xm**2
t4 = ym**2
t5 = cmplx(0.D0,4.D0)
t8 = sig**2
t17 = cmplx(0.D0,-4.D0)
t30 = q**2
t31 = wb**2
t32 = t31*t30
t33 = t8**2
t52 = t1+t2+t3+t4+ym*xm*t8*wb*q*t5
#+t8*ym*wb*q*xk*t5+t8*yk*wb*q*xm*
#t17-2*xk*xm+yk*xk*t8*wb*q*t17+t4*t33*t32
#-2*yk*t33*ym*t32+t33*t2*t3
#2-2*ym*yk+t3*t33*t32-2*xk*xm*t33*t32+t1*t33*t32
t56 = exp(-1/t8*t52/4)

t58 = 0.3141592653589793D1*t8*t5

```

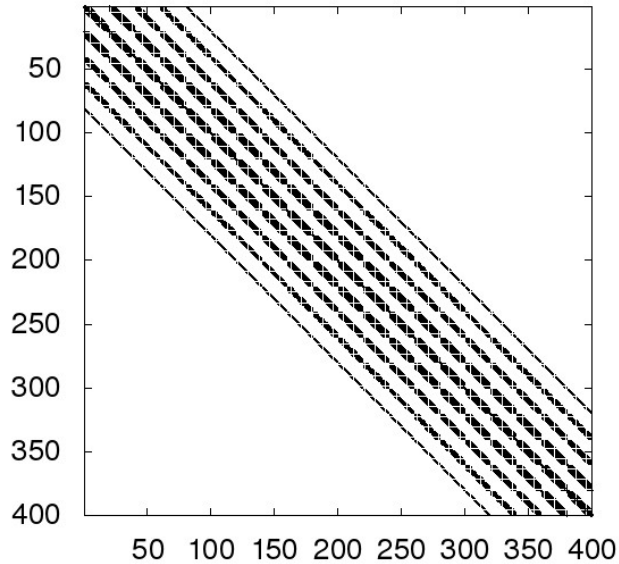
Korzystając z obu bloków możemy wyznaczyć wszystkie elementy macierzowe

- całkowanie wykonane analitycznie w programie MAPLE

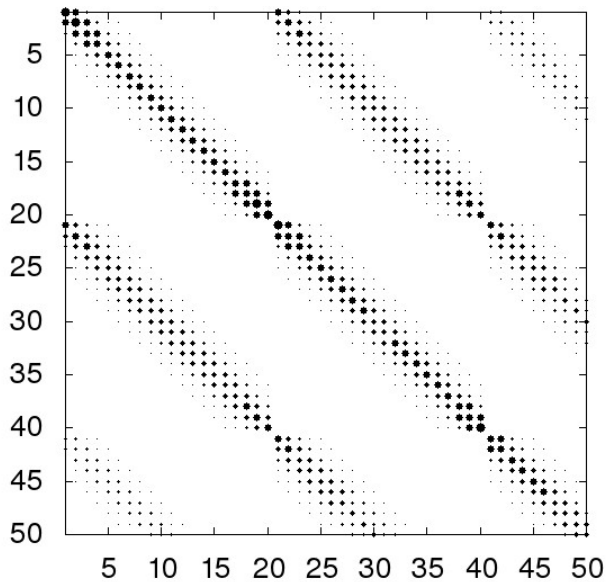
Macierze są rzadkie co można wykorzystać jeśli są duże, a my nie potrzebujemy wszystkich wartości/wektorów własnych.

Macierz Hamiltonianu

$n=20 \times 20$, $H_{km}/20000$



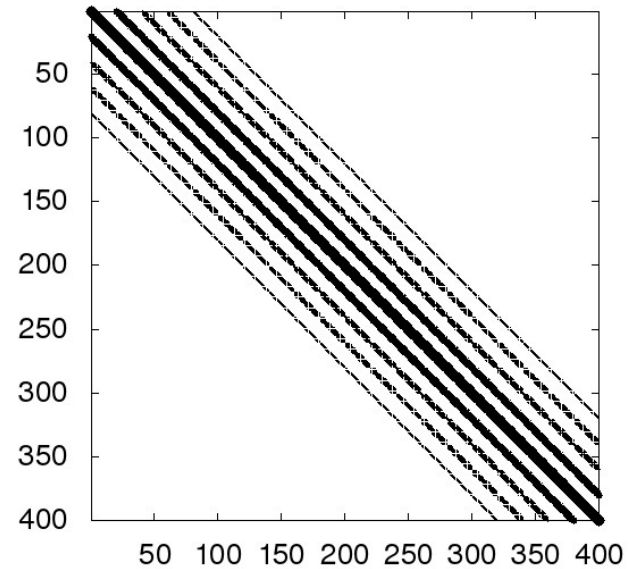
$n=20 \times 20$, $\text{sqrt}(H_{km})/20$



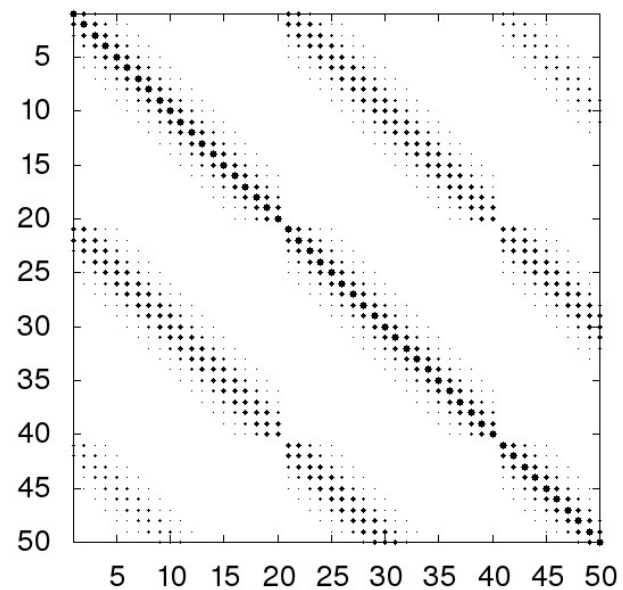
$$Hc = ESc$$

Macierz cętek przekrywania

$n=20 \times 20$, $S_{km}/20000$

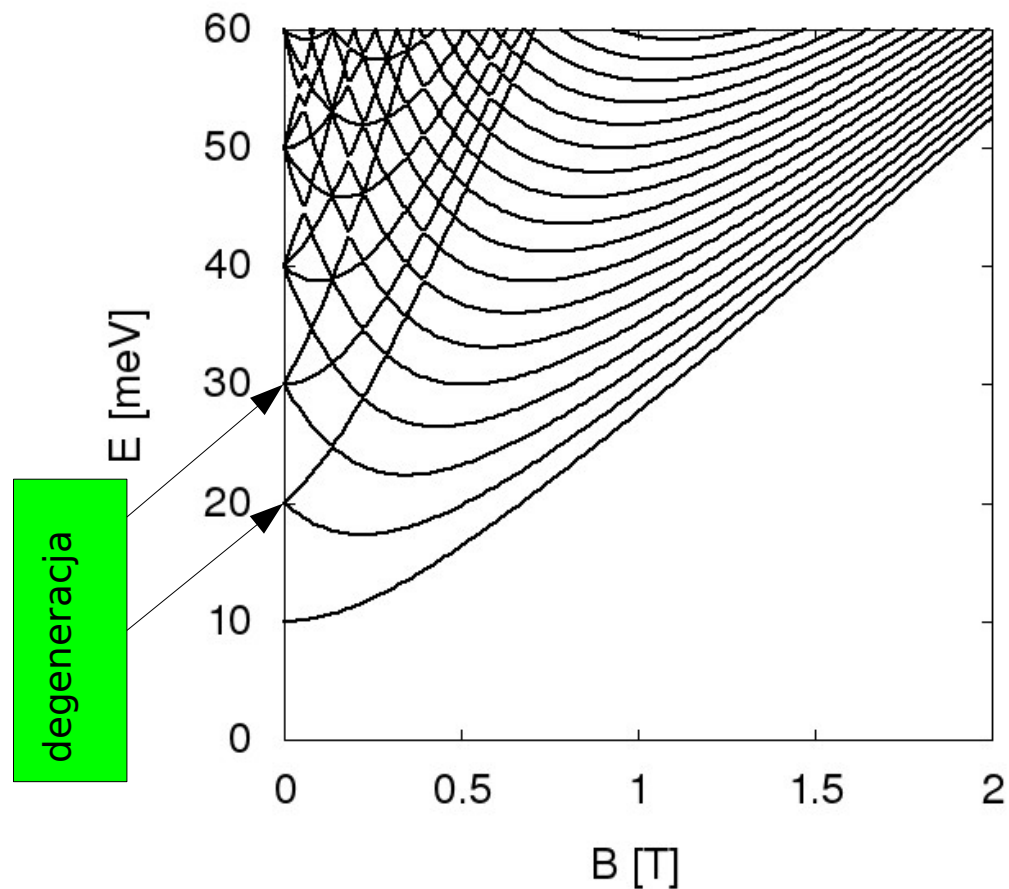
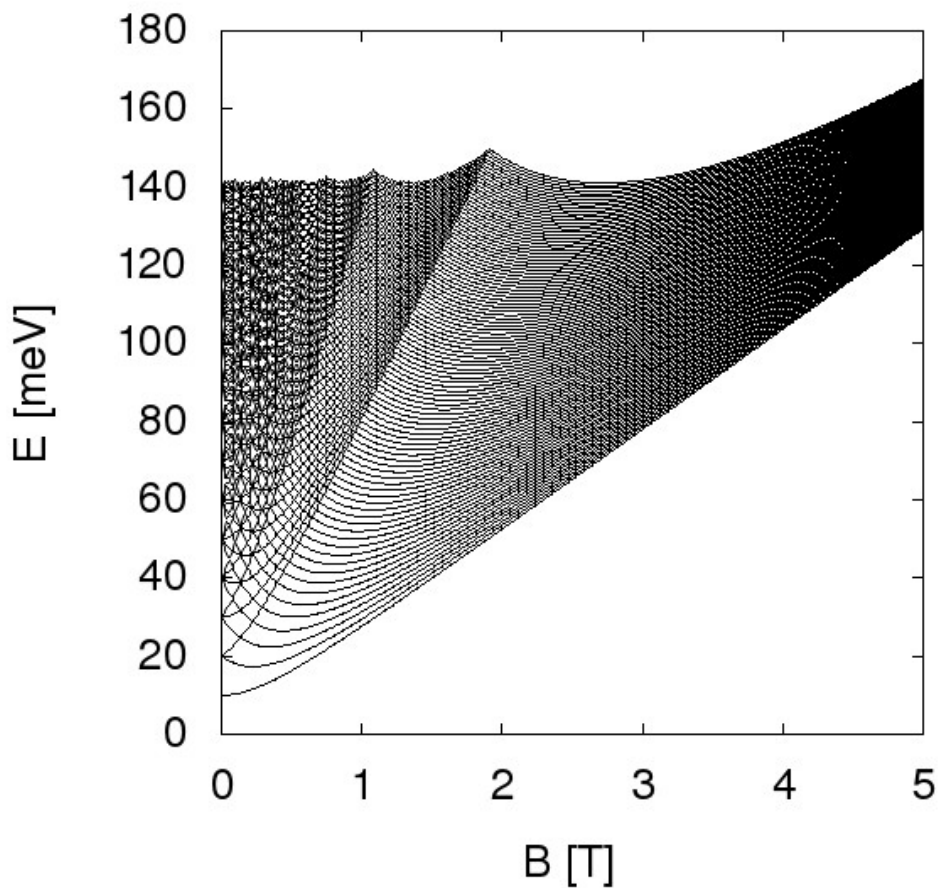


$n=20 \times 20$, $\text{sqrt}(S_{km})/200$



Widmo energii

$$\hbar\omega_x = \hbar\omega_y = 10 \text{ meV}$$



Przecinające się linie na wykresach oznaczają degenerację stanów (identyczne wartości własne macierzy H)

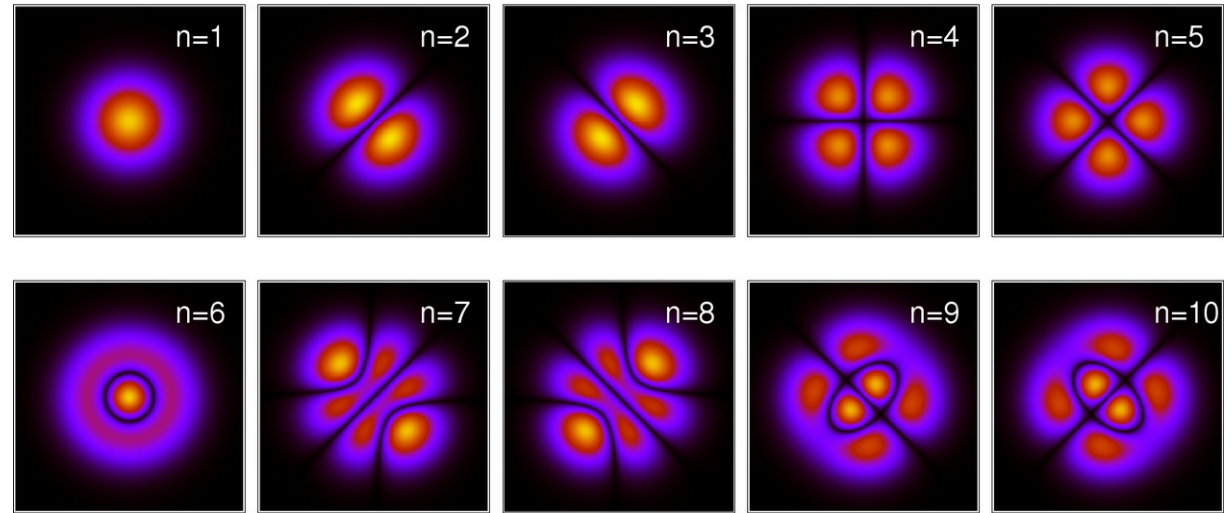
$$|\Psi(x, y)|^2, \quad B_z = 0$$

$$E[\hbar\omega_{x/y}] = 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5$$

brak symetrii

$$\hbar\omega_x = \hbar\omega_y = 10 \text{ meV}$$

stany o tej samej energii
mieszają się ze sobą

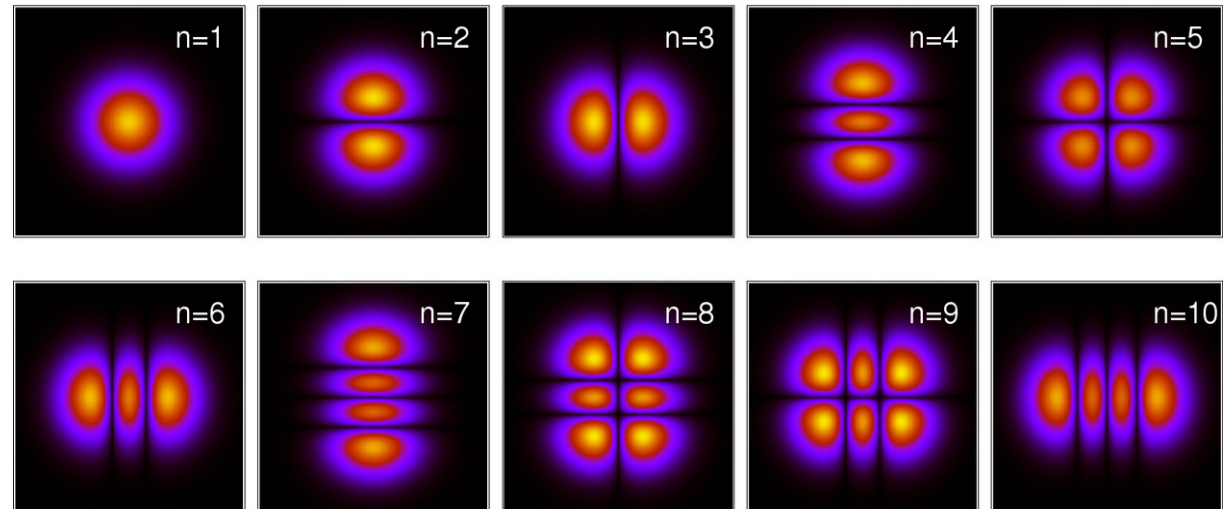


degeneracje: (2,3) (4,5,6) (7,8,9,10)

symetria rozwiązań
wymuszona

$$\hbar\omega_x = 10 \text{ meV}$$

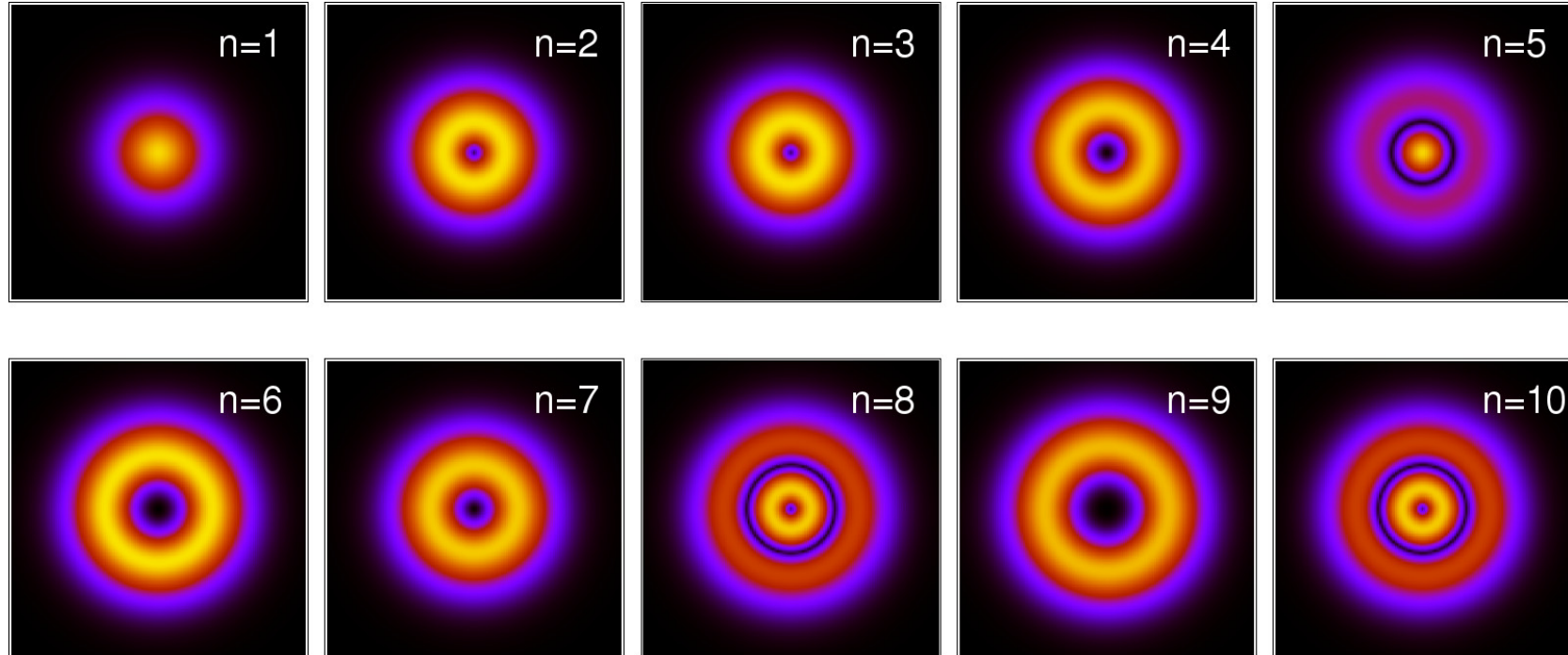
$$\hbar\omega_y = 10.01 \text{ meV}$$



degeneracja wartości własnych
zniesiona - mieszanie
stanów o tej samej
energii nie występuje

$$|\Psi(x, y)|^2, \quad B_z = 0.1 T$$

$$\hbar\omega_x = \hbar\omega_y = 10 \text{ meV}$$



Operator energii posiada symetrię obrotową
- dlatego rozwiązania muszą ją odtwarzać

stany z zerową gęstością w środku mają niezerowy moment pędu

Rozkład SVD (Singular Value Decomposition)

Macierz, w ogólnym przypadku prostokątną $R^{m \times n}$ chcemy zapisać w postaci iloczynu

$$A = U \Sigma V^T$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

$$V^T V = I$$

$$U^T U = I$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & \dots & a_{2n} \\ \vdots & & & \ddots & & \vdots \\ a_{m1} & a_{m2} & \dots & \dots & \dots & a_{mn} \end{pmatrix}$$
$$= \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1k} \\ u_{21} & u_{22} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & \dots & \dots & u_{mk} \end{pmatrix} \times \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_k & 0 & \dots & 0 \end{pmatrix}$$
$$\times \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1n} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} & \dots & v_{2n} \\ \vdots & & & \ddots & & \vdots \\ v_{k1} & v_{k2} & \dots & \dots & \dots & v_{kn} \end{pmatrix}$$

Jak wyznaczyć rozkład SVD?

$$U\Sigma V^T = A$$

$$/ \cdot V \quad U^T \cdot /$$

$$U\Sigma = AV \quad V\Sigma = U^T A$$

dla

możemy zapisać

$$j \in \{1, 2, \dots, n\}$$

$$Av_j = \sigma_j u_j \quad A^T \cdot /$$

$$A^T u_j = \sigma_j v_j$$

$$\begin{aligned} A^T Av_j &= \sigma_j A^T u_j \\ &= \sigma_j^2 v_j \end{aligned}$$

Najpierw rozwiązujemy problem własny (np. metoda Lanczosa)

$$A^T Av_j = \sigma_j^2 v_j$$

a następnie szukamy macierzy U

$$A = U\Sigma V^T \quad V \cdot /$$

$$VA = U\Sigma \quad / \cdot \Sigma^{-1} \quad \iff \sigma_j > 0$$

$$U = \Sigma^{-1}VA$$

Zastosowania rozkładu SVD:

1. Rozwiązywanie układów równań

$$Ax = b,$$

$$U\Sigma V^T x = b, \quad U^T \cdot /$$

$$(U^T U)\Sigma V^T x = U^T b, \quad \Sigma^{-1} \cdot /$$

$$V^T x = \Sigma^{-1} U^T b, \quad V \cdot /$$

$$x = V\Sigma^{-1} U^T b,$$

2. Rozwiązanie układu nadokreślonego (problem najmniejszych kwadratów)

$$Ax = b, \quad A^T \cdot /$$

$$A^T Ax = A^T b, \quad (A^T A)^{-1} \cdot /$$

$$x = A^I b,$$

A^I jest pseudoodwrotnością macierzy A .

$$A^I = (A^T A)^{-1} A^T$$

Ponieważ $m > n$, SVD ma postać

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T$$

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \quad \sigma_j > 0$$

A^I można wyrazić poprzez SVD

$$A^I = V (\Sigma^{-1} 0) U^T$$

I znaleźć rozwiązanie

$$x = V (\Sigma^{-1} 0) U^T b$$

3. Kompresja obrazu

SVD możemy zapisać jako

$$A = U\Sigma V^T$$

lub operując na wektorach

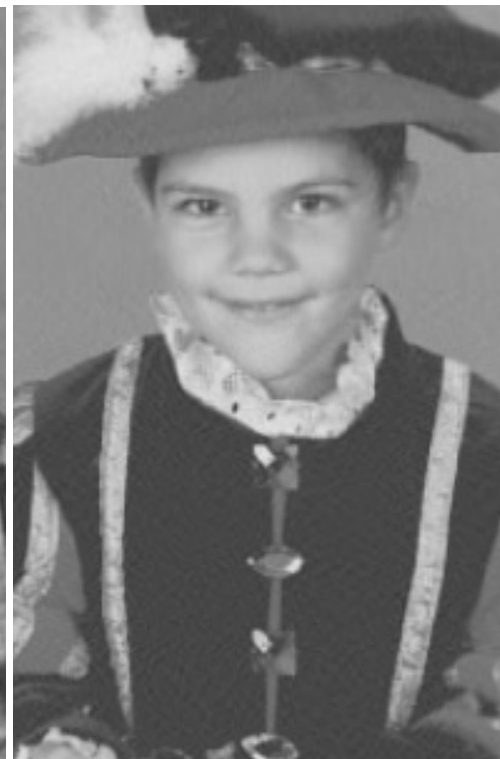
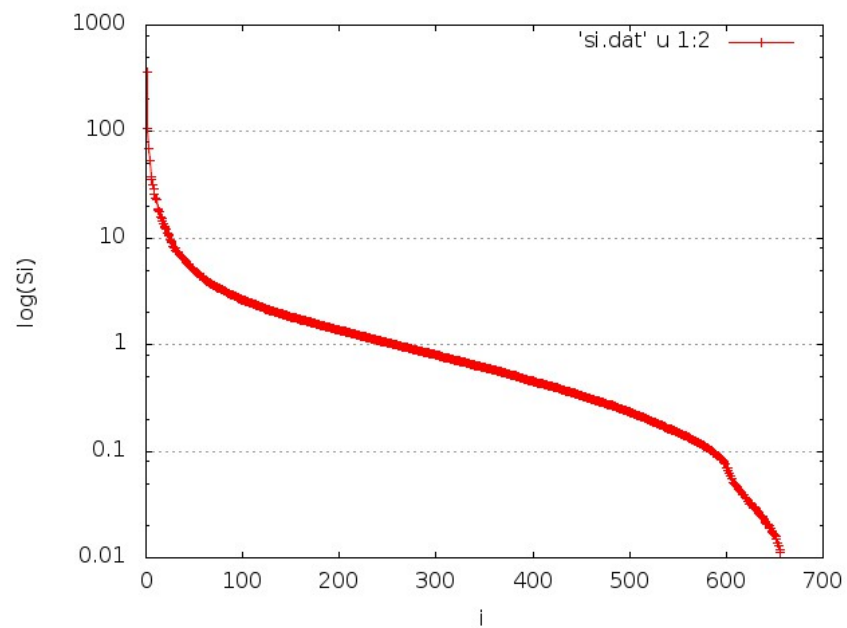
$$A = \sum_{j=1}^n \sigma_j u_j v_j^T$$

wówczas zakładając $k < n$ dostajemy przybliżenie macierzy A (k-rank)

$$A^{(k)} = \sum_{j=1}^{k < n} \sigma_j u_j v_j^T$$

Przykład na następnym slajdzie





Oryginał
887x656 pikseli

$A^{(50)}$

$A^{(100)}$

$A^{(200)}$