MACIERZ ORTOGONALNA

Jeśli macierz $\mathbf{Q}_{m\times n}$ $(m \geqslant n)$ ma kolumny wzajemnie ortogonalne, to

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{D}_n,$$

gdzie $\mathbf{D}_n$ jest macierzą diagonalną o wymiarowości n.

Jeśli kolumny $\mathbf{Q}$ są, dodatkowo, wektorami ortonormalnymi, to

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}_n$$

Macierz kwadratową $\mathbf{Q}$ nazywamy **macierzą ortogonalną**, jeśli

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I},$$

tzn. $\mathbf{Q}$ to macierz kwadratowa, której kolumny są wektorami ortonormalnymi.

Formalnie definicja powyższa określa macierz ortonormalną, ale *podana definicja przyjęła się w literaturze.*

MACIERZ ORTOGONALNA 2

Z definicji $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$ wynika bezpośrednio, że

$$\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1},$$

skąd dalej

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I},$$

tzn. w macierzy ortogonalnej kolumny są wektorami wzajemnie ortonormalnymi oraz wiersze są wektorami wzajemnie ortonormalnymi.

Lemat
Jeśli $\mathbf{Q}_{m\times n}$ – macierz o kolumnach będących wektorami ortonormalnymi, to

$$\|\mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{x}\|_2 .$$

Dowód. Z definicji normy drugiej i z własności $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}_n$ wynika

$$\forall \mathbf{x} \quad \|\mathbf{Q}\mathbf{x}\|_2 = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\mathbf{x}} = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}} = \|\mathbf{x}\|_2 .$$

Wniosek. Dla macierzy ortogonalnej obowiązuje:

$$\|\mathbf{Q}\mathbf{x}\|_2 = \|\mathbf{x}\|_2 , \quad \left\|\mathbf{Q}^{\mathrm{T}}\mathbf{x}\right\|_2 = \|\mathbf{x}\|_2 .$$

ROZKŁAD QR (ortogonalno-trójkątny)

Twierdzenie. Każdą macierz $\mathbf{A}_{m\times n}$, $(m \geqslant n)$ o liniowo niezależnych kolumnach (tj. o pełnym rzędzie równym n) można przedstawić w postaci następujących rozkładów ortogonalno- trójkątnych:

1. *Rozkład QR wąski (ekonomiczny) nieunormowany:*

$$\mathbf{A}_{m\times n} = \bar{\mathbf{Q}}_{m\times n}\bar{\mathbf{R}}_{n\times n},$$

   gdzie $\bar{\mathbf{Q}}_{m\times n}$ jest macierzą o kolumnach wzajemnie ortogonalnych (nie ortonormalnych), a $\bar{\mathbf{R}}_{n\times n}$ trójkątną górną z jedynkami na diagonali;

2. *Rozkład QR wąski (ekonomiczny):*

$$\mathbf{A}_{m\times n} = \mathbf{Q}_{m\times n}\mathbf{R}_{n\times n},$$

   gdzie $\mathbf{Q}_{m\times n}$ jest macierzą o kolumnach ortonormalnych, a $\mathbf{R}_{n\times n}$ trójkątną górną z dodatnimi elementami na diagonali;

3. *Rozkład QR:*

$$\mathbf{A}_{m\times n} = \mathbf{Q}_{m\times m}\begin{bmatrix} \mathbf{R}_{n\times n} \\ \mathbf{0} \end{bmatrix}_{m\times n} = \begin{bmatrix}\mathbf{Q}_{m\times n} \ \mathbf{Q}^2_{m\times(m-n)}\end{bmatrix}\begin{bmatrix} \mathbf{R}_{n\times n} \\ \mathbf{0} \end{bmatrix}_{m\times n},$$

   gdzie $\mathbf{Q}_{m\times m}$ jest macierzą ortogonalną.

ROZKŁAD QR (ortogonalno-trójkątny) 2

Dowód - przez konstrukcję rozkładów.

1. Niech $\mathbf{A} = [\mathbf{a}_1, ..., \mathbf{a}_n]$. Zortogonalizujemy kolumny $\mathbf{a}_i$ macierzy $\mathbf{A}$ standardowym algorytmem Grama-Schmidta, oznaczając przez $\bar{\mathbf{q}}_i$ wektory zortogonalizowane:

$$\begin{aligned}
\bar{\mathbf{q}}_1 &= \mathbf{a}_1, \quad \bar{r}_{11} \triangleq 1, \\
\bar{\mathbf{q}}_2 &= \mathbf{a}_2 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}} \mathbf{a}_2}{\bar{\mathbf{q}}_1^{\mathrm{T}} \bar{\mathbf{q}}_1} \bar{\mathbf{q}}_1 = \mathbf{a}_2 - \bar{r}_{12} \bar{\mathbf{q}}_1, \qquad \bar{r}_{22} \triangleq 1, \\
\bar{\mathbf{q}}_i &= \mathbf{a}_i - \sum_{j=1}^{i-1} \frac{\bar{\mathbf{q}}_j^{\mathrm{T}} \mathbf{a}_i}{\bar{\mathbf{q}}_j^{\mathrm{T}} \bar{\mathbf{q}}_j} \bar{\mathbf{q}}_j = \mathbf{a}_i - \sum_{j=1}^{i-1} \bar{r}_{ji} \bar{\mathbf{q}}_j, \quad \bar{r}_{ii} \triangleq 1, \qquad i = 3, ..., n.
\end{aligned}$$

Równania powyższe można przepisać w postaci:

$$\begin{aligned}
\mathbf{a}_1 &= \bar{r}_{11} \bar{\mathbf{q}}_1, \quad \bar{r}_{11} \triangleq 1, \\
\mathbf{a}_2 &= \bar{r}_{12} \bar{\mathbf{q}}_1 + \bar{r}_{22} \bar{\mathbf{q}}_2 = \sum_{j=1}^{2} \bar{r}_{j2} \bar{\mathbf{q}}_j, \qquad \bar{r}_{12} = \frac{\bar{\mathbf{q}}_1^{\mathrm{T}} \mathbf{a}_2}{\bar{\mathbf{q}}_1^{\mathrm{T}} \bar{\mathbf{q}}_1}, \quad \bar{r}_{22} \triangleq 1, \\
\mathbf{a}_i &= \sum_{j=1}^{i-1} \bar{r}_{ji} \bar{\mathbf{q}}_j + \bar{r}_{ii} \bar{\mathbf{q}}_i = \sum_{j=1}^{i} \bar{r}_{ji} \bar{\mathbf{q}}_j, \qquad \bar{r}_{ji} = \frac{\bar{\mathbf{q}}_j^{\mathrm{T}} \mathbf{a}_i}{\bar{\mathbf{q}}_j^{\mathrm{T}} \bar{\mathbf{q}}_j}, \quad \bar{r}_{ii} \triangleq 1, \quad i = 3, ..., n.
\end{aligned}$$

ROZKŁAD QR (ortogonalno-trójkątny) 3

Wykorzystamy ogólną własność iloczynu macierzy:

> *i-ta kolumna $\mathbf{a}_i$ macierzy iloczynowej $\mathbf{A} = \mathbf{BC}$ jest kombinacją liniową kolumn pierwszej macierzy iloczynu $\mathbf{B}$, ze współczynnikami będącymi elementami i-tej kolumny drugiej macierzy iloczynu $\mathbf{C}$:*

$$\mathbf{a}_i = \mathbf{B}\mathbf{c_i} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n] \begin{bmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{bmatrix} = \sum_{j=1}^{n} \mathbf{b}_j c_{ji}$$

Stosując tę własność do $\mathbf{a}_i = \sum_{j=1}^{i} \bar{\mathbf{q}}_j \bar{r}_{ji}$ można zapisać:

$$\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n] = [\bar{\mathbf{q}}_1, ..., \bar{\mathbf{q}}_n] \begin{bmatrix} 1 & \bar{r}_{12} & \ldots & \bar{r}_{1n} \\ 0 & 1 & \ldots & \bar{r}_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} = \bar{\mathbf{Q}}\bar{\mathbf{R}},$$

gdzie macierz $\bar{\mathbf{Q}}_{m\times n} = [\bar{\mathbf{q}}_1, ..., \bar{\mathbf{q}}_n]$ ma kolumny ortogonalne (nie ortonormalne), a macierz $\bar{\mathbf{R}}_{n\times n}$ jest trójkątna górna z jedynkami na diagonali, co kończy dowód 1.

ROZKŁAD QR (ortogonalno-trójkątny) 4

2. Jeśli przeprowadzimy normalizację kolumn macierzy $\bar{\mathbf{Q}}$:

$$\mathbf{Q} = \left[\frac{\bar{\mathbf{q}}_1}{\|\bar{\mathbf{q}}_1\|}, \frac{\bar{\mathbf{q}}_2}{\|\bar{\mathbf{q}}_2\|}, \ldots, \frac{\bar{\mathbf{q}}_n}{\|\bar{\mathbf{q}}_n\|} \right]$$

oraz dokonamy przekształcenia

$$\mathbf{R} = \mathbf{N}\bar{\mathbf{R}},$$

gdzie

$$\mathbf{N} = \begin{bmatrix} \|\bar{\mathbf{q}}_1\| & 0 & \ldots & 0 \\ 0 & \|\bar{\mathbf{q}}_2\| & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \|\bar{\mathbf{q}}_n\| \end{bmatrix},$$

to uzyskamy rozkład QR wąski

$$\mathbf{A}_{m\times n} = \mathbf{Q}_{m\times n}\mathbf{R}_{n\times n},$$

gdzie $\mathbf{Q}_{m\times n}$ ma kolumny ortonormalne, co kończy dowód tezy 2.

ROZKŁAD QR (ortogonalno-trójkątny) 5

3. Jeśli

– macierz $\mathbf{Q}_{m\times n}$ uzupełnimy do macierzy kwadratowej ortogonalnej $\mathbf{Q}_{m\times m}$ przez dodanie m-n kolumn – dowolnych, ale ortonormalnych wzajemnie i z kolumnami macierzy $\mathbf{Q}_{m\times n}$,

– macierz $\mathbf{R}_{n\times n}$ uzupełnimy zerami do wymiaru $m \times n$,

to uzyskamy rozkład QR (pełny):

$$\mathbf{A}_{m\times n} = \left[\mathbf{Q}_{m\times n}\ \mathbf{Q}^2_{m\times(m-n)}\right] \begin{bmatrix} \mathbf{R}_{n\times n} \\ \mathbf{0} \end{bmatrix}_{m\times n} = \mathbf{Q}_{m\times m} \begin{bmatrix} \mathbf{R}_{n\times n} \\ \mathbf{0} \end{bmatrix}_{m\times n},$$

co kończy dowód Twierdzenia.

W ogólności, *rozkład QR istnieje dla każdej macierzy* $\mathbf{A}_{m\times n},\ m \geqslant n$, *(również o rzędzie* $k < n$).

ROZKŁAD QR (ortogonalno-trójkątny) 6

Numerycznego obliczenia rozkładu QR macierzy można dokonywać:

1. Dla **macierzy o pełnym rzędzie**, metodą ortogonalizacji Grama-Schmidta:
   - 1a. rozkład QR wąski nieunormowany,
   - 1b. po normalizacji rozkład QR wąski,
   - 1c. który można dalej rozszerzyć do rozkładu QR pełnego:
     uzupełniamy macierz $\mathbf{A}_{m\times n}$ do macierzy kwadratowej

     $$\mathbf{A}_{m\times m} = [\mathbf{A}_{m\times n}\ \mathbf{A}_{m\times(m-n)}],$$

     przez dodanie jakichkolwiek $m-n$ kolumn tworzących uzupełnienie $\mathbf{A}_{m\times(m-n)}$ liniowo niezależnych z kolumnami macierzy $\mathbf{A}_{m\times n}$ (i względem siebie),
     i kontynuujemy również dla nich, kolejno, ortogonalizację Grama- Schmidta.
     Rozszerzamy $\mathbf{R}$ o zera.

   Z reguły stosujemy tzw. *zmodyfikowany algorytm Grama- Schmidta*, który ma lepsze własności numeryczne.

ROZKŁAD QR (ortogonalno-trójkątny) 7

Dla macierzy **o niepełnym (dowolnym) rzędzie** można odpowiednio rozszerzyć algorytm Grama-Schmidta;

ale zalecane jest stosowanie jednej z metod podanych poniżej:

2. metodą odbić Householdera – dostajemy od razu rozkład QR (pełny), w razie potrzeby rozkład wąski uzyskujemy przez usunięcie ostatnich $m-n$ kolumn macierzy $\mathbf{Q}_{m\times m}$ i $m-n$ ostatnich wierszy macierzy $\mathbf{R}_{m\times n}$;

3. metodą obrotów Givensa – dostajemy od razu rozkład QR (pełny), podobnie jak przy metodzie odbić Householdera. Metoda zalecana *dla macierzy rzadkich*.

ROZKŁAD QR – Przykład

Dokonamy rozkładu QR macierzy

$$\text{a)}\quad \mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2] = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}$$

Ortogonalizacja:

$$\begin{aligned} \bar{\mathbf{q}}_1 &= \mathbf{a}_1 = [1\ 2\ -2]^{\mathrm{T}}, \\ \bar{\mathbf{q}}_2 &= \mathbf{a}_2 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_2}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 = \mathbf{a}_2 - \bar{r}_{12}\bar{\mathbf{q}}_1 = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} - \frac{-9}{9}\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}. \end{aligned}$$

Stąd dostajemy rozkład wąski (nieunormowany):

$$\bar{\mathbf{Q}}_{\mathbf{A}} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ -2 & 2 \end{bmatrix}, \quad \bar{\mathbf{R}}_{\mathbf{A}} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

Normując kolumny macierzy $\bar{\mathbf{Q}}_{\mathbf{A}}$ dostajemy rozkład wąski:

$$\mathbf{N} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \quad \mathbf{Q}_{\mathbf{A}} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{2}{3} \end{bmatrix}, \quad \mathbf{R}_{\mathbf{A}} = \mathbf{N}\bar{\mathbf{R}}_{\mathbf{A}} = \begin{bmatrix} 3 & -3 \\ 0 & 3 \end{bmatrix}.$$

ROZKŁAD QR – Przykład c.d.

b)

$$\mathbf{B} = [\mathbf{a}_1\ \mathbf{a}_2\ \mathbf{a}_3] = \begin{bmatrix} 1 & 1 & 9 \\ 2 & -1 & 0 \\ -2 & 4 & 0 \end{bmatrix} \qquad \left(\bar{\mathbf{Q}}_{\mathbf{A}} \bar{\mathbf{R}}_{\mathbf{A}} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \right)$$

$\mathbf{B} = [\mathbf{A}\ \mathbf{a}_3]$, stąd wystarczy dodać ortogonalizację trzeciej kolumny:

$$\bar{\mathbf{q}}_3 = \mathbf{a}_3 - \frac{\bar{\mathbf{q}}_1^T \mathbf{a}_3}{\bar{\mathbf{q}}_1^T \bar{\mathbf{q}}_1} \bar{\mathbf{q}}_1 - \frac{\bar{\mathbf{q}}_2^T \mathbf{a}_3}{\bar{\mathbf{q}}_2^T \bar{\mathbf{q}}_2} \bar{\mathbf{q}}_2 = \mathbf{a}_3 - \bar{r}_{13} \bar{\mathbf{q}}_1 - \bar{r}_{23} \bar{\mathbf{q}}_2 = \begin{bmatrix} 9 \\ 0 \\ 0 \end{bmatrix} - \frac{9}{9} \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} - \frac{18}{9} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ -2 \end{bmatrix}$$

Dostajemy rozkład nieunormowany

$$\bar{\mathbf{Q}}_{\mathbf{B}} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 1 & -4 \\ -2 & 2 & -2 \end{bmatrix}, \quad \bar{\mathbf{R}}_{\mathbf{B}} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Normując kolumny macierzy $\bar{\mathbf{Q}}$ dostajemy:

$$\mathbf{N} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}, \quad \mathbf{Q}_{\mathbf{B}} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{bmatrix}, \quad \mathbf{R}_{\mathbf{B}} = \mathbf{N} \bar{\mathbf{R}}_{\mathbf{B}} = \begin{bmatrix} 3 & -3 & 3 \\ 0 & 3 & 6 \\ 0 & 0 & 6 \end{bmatrix}.$$

ALGORYTMY Grama-Schmidta CGS i MGS

Klasyczny algorytm Grama-Schmidta (CGS, nienormowany) ortogonalizuje kolumny macierzy kolejno, względem wszystkich już zortogonalizowanych:

$$\bar{\mathbf{q}}_1 = \mathbf{a}_1 \quad \Rightarrow \quad \bar{\mathbf{q}}_2 = \mathbf{a}_2 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_2}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 \quad \Rightarrow \quad \bar{\mathbf{q}}_3 = \mathbf{a}_3 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_3}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 - \frac{\bar{\mathbf{q}}_2^{\mathrm{T}}\mathbf{a}_3}{\bar{\mathbf{q}}_2^{\mathrm{T}}\bar{\mathbf{q}}_2}\bar{\mathbf{q}}_2 \quad \Rightarrow \quad \text{itd.} \ldots$$

Zmodyfikowany algorytm Grama Schmidta (MGS, nienormowany) ortogonalizuje wobec ostatniej zortogonalizowanej kolumny wszystkie następne kolumny:

$$\bar{\mathbf{q}}_1 = \mathbf{a}_1$$

$$\begin{array}{rcll}
\bar{\mathbf{q}}_2 & = & \mathbf{a}_2 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_2}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 & \\
\mathbf{a}_3^{(1)} & = & \mathbf{a}_3 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_3}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 \quad \Rightarrow & \bar{\mathbf{q}}_3 = \mathbf{a}_3^{(1)} - \dfrac{\bar{\mathbf{q}}_2^{\mathrm{T}}\mathbf{a}_3}{\bar{\mathbf{q}}_2^{\mathrm{T}}\bar{\mathbf{q}}_2}\bar{\mathbf{q}}_2 \quad \Rightarrow \quad \text{itd.} \\
\mathbf{a}_4^{(1)} & = & \mathbf{a}_4 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_4}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 & \mathbf{a}_4^{(2)} = \mathbf{a}_4^{(1)} - \dfrac{\bar{\mathbf{q}}_2^{\mathrm{T}}\mathbf{a}_4}{\bar{\mathbf{q}}_2^{\mathrm{T}}\bar{\mathbf{q}}_2}\bar{\mathbf{q}}_2 \\
\mathbf{a}_5^{(1)} & = & \mathbf{a}_5 - \frac{\bar{\mathbf{q}}_1^{\mathrm{T}}\mathbf{a}_5}{\bar{\mathbf{q}}_1^{\mathrm{T}}\bar{\mathbf{q}}_1}\bar{\mathbf{q}}_1 & \mathbf{a}_5^{(2)} = \mathbf{a}_5^{(1)} - \dfrac{\bar{\mathbf{q}}_2^{\mathrm{T}}\mathbf{a}_5}{\bar{\mathbf{q}}_2^{\mathrm{T}}\bar{\mathbf{q}}_2}\bar{\mathbf{q}}_2 \\
 & & \vdots & \vdots
\end{array}$$

ZMODYFIKOWANY ALGORYTM Grama-Schmidta

Zmodyfikowany algorytm Grama-Schmidta (MGS) ortogonalizuje wobec ostatniej zortogonalizowanej kolumny wszystkie następne kolumny, w pseudonotacji Matlaba:

for i = 1:n,
$\quad \bar{\mathbf{q}}_i = \mathbf{a}_i^{(i)}; \quad \bar{r}_{ii} = 1;$
$\quad d_i = \bar{\mathbf{q}}_i^{\mathrm{T}} * \bar{\mathbf{q}}_i;$
$\quad$ for j = i+1:n
$\qquad \bar{r}_{ij} = (\bar{\mathbf{q}}_i^{\mathrm{T}} * \mathbf{a}_j^{(i)})/d_i;$
$\qquad \mathbf{a}_j^{(i+1)} = \mathbf{a}_j^{(i)} - \bar{r}_{ij} * \bar{\mathbf{q}}_i;$
$\quad$ end
end

Matematycznie, algorytmy CGS i MGS są równoważne, ale MGS ma lepsze własności numeryczne.

ROZKŁAD QR zmodyfikowanym algorytmem Grama-Schmidta

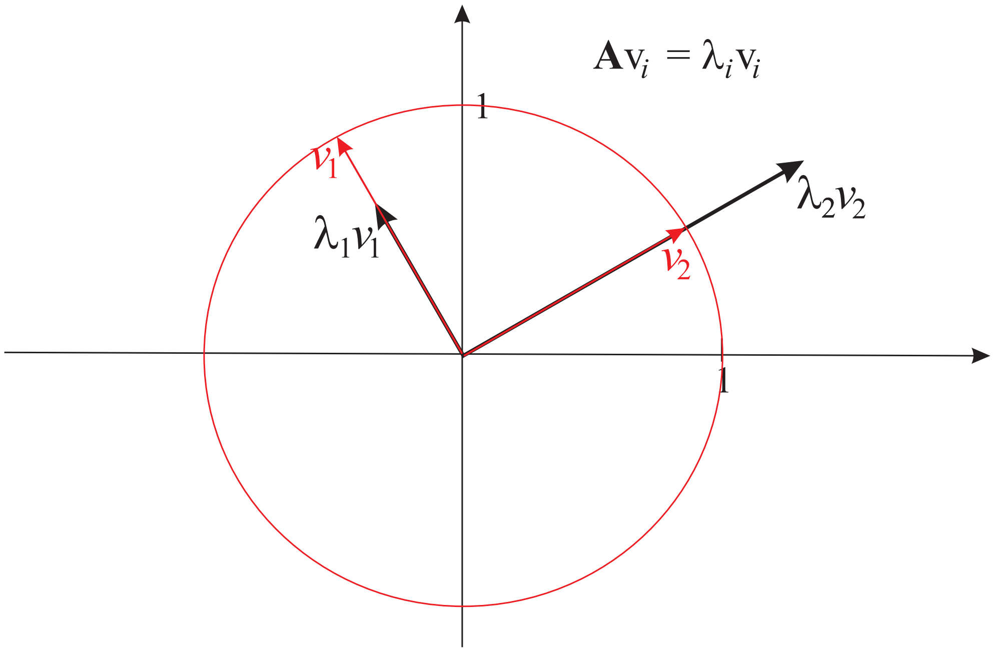
```
function [Q,R]=qrmgs(A)
%rozkład QR (wąski) zmodyfikowanym algorytmem Grama-Schmidta
%dla macierzy prostokątnych rzeczywistych i zespolonych
[m n]=size(A); Q=zeros(m,n); R=zeros(n,n); d=zeros(1,n);
  %rozkład z kolumnami Q ortogonalnymi (nie ortonormalnymi):
for i=1:n
   Q(:,i)=A(:,i);
   R(i,i)=1;
   d(i)=Q(:,i)'*Q(:,i);
   for j=i+1:n
       R(i,j)=(Q(:,i)'*A(:,j))/d(i);
       A(:,j)=A(:,j)-R(i,j)*Q(:,i);
   end
end
  %normowanie rozkładu (kolumny Q ortonormalne):
for i=1:n
    Q(:,i)=Q(:,i)/sqrt(d(i));
    R(i,i:n)=R(i,i:n)*sqrt(d(i));
end
```

WARTOŚCI WŁASNE

$$\mathbf{A}_{n\times n}: \qquad \mathbf{Av} = \lambda \mathbf{v}, \qquad \lambda - \text{wartość własna}, \ \mathbf{v} - \text{wektor własny}$$

Macierz o wymiarze n ma dokładnie n wartości własnych: $\mathbf{Av}_i = \lambda_i \mathbf{v}_i, \quad i = 1, ..., n.$

Interpretacja dla $n = 2$ i macierzy $\mathbf{A}$ symetrycznej:

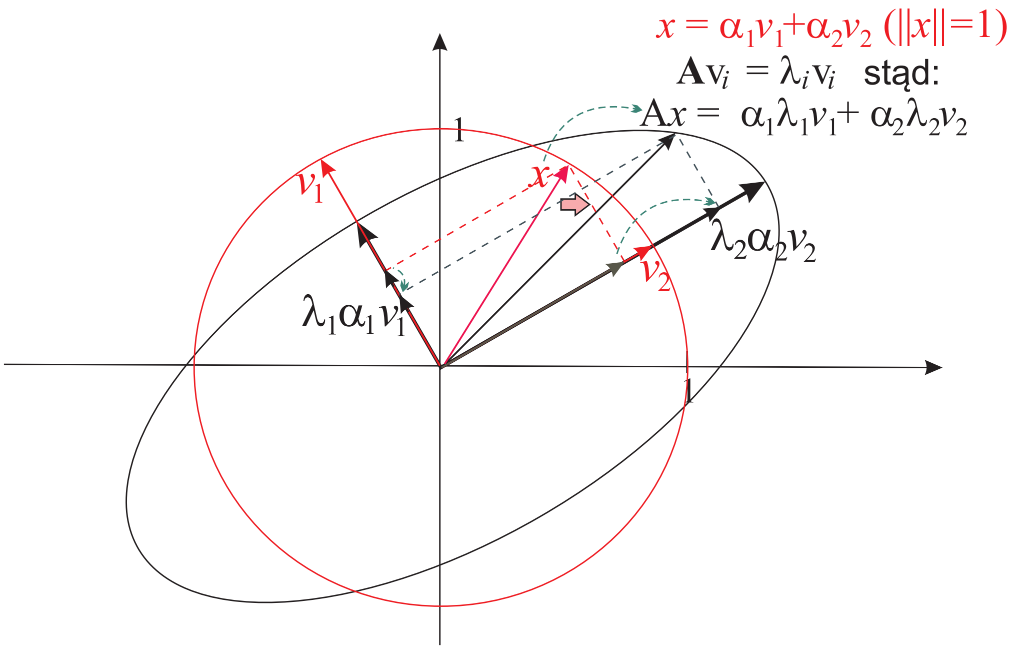


WARTOŚCI WŁASNE

$$\mathbf{A}_{n\times n}: \qquad \mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \qquad \lambda - \text{wartość własna}, \ \mathbf{v} - \text{wektor własny}$$

Macierz o wymiarze n ma dokładnie n wartości własnych: $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad i = 1, ..., n.$

Interpretacja dla $n = 2$ i macierzy $\mathbf{A}$ symetrycznej:



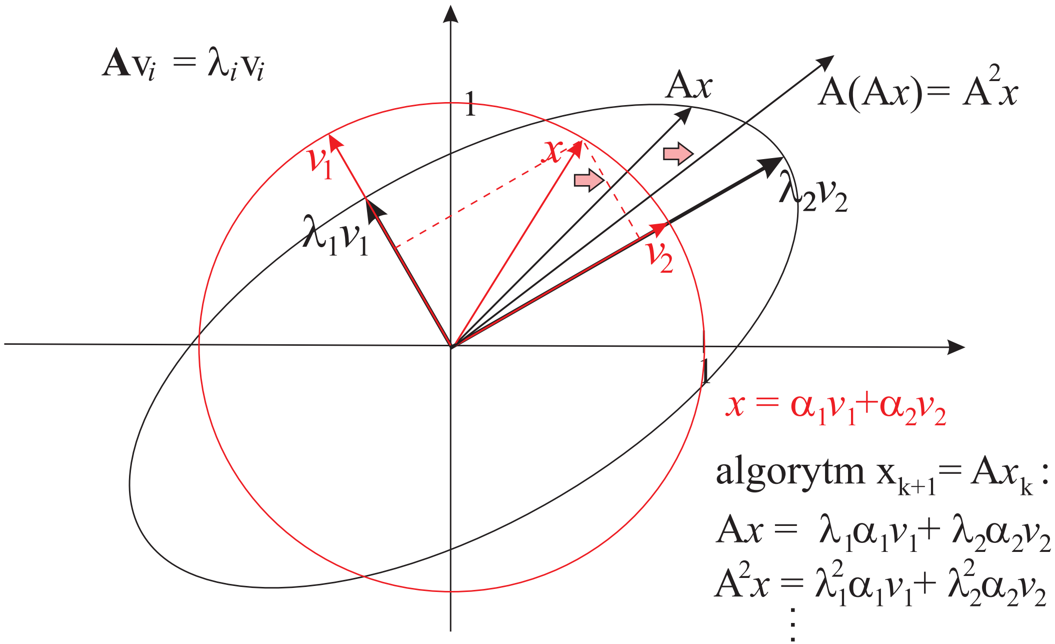
$$\|\mathbf{A}\|_2 = \max_{\|\mathbf{x}=1\|_2} \|\mathbf{A}\mathbf{x}\|_2$$
$$= \max_{\|\mathbf{x}=1\|_2} \sqrt{\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}} = \sqrt{\max_{\|\mathbf{x}=1\|_2} \mathbf{x}^T(\mathbf{A}^T\mathbf{A})\mathbf{x}} = \sqrt{\mathbf{v}_{max}^T(\lambda_{max}\mathbf{v}_{max})} = \sqrt{\lambda_{max}}$$

WARTOŚCI WŁASNE

$$\mathbf{A}_{n\times n}: \qquad \mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \qquad \lambda - \text{wartość własna}, \ \mathbf{v} - \text{wektor własny}$$

Macierz o wymiarze n ma dokładnie n wartości własnych: $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad i = 1, ..., n.$

Interpretacja dla $n = 2$ i macierzy $\mathbf{A}$ symetrycznej – zbieżność/rozbieżność:



WARTOŚCI WŁASNE 2

$$\mathbf{A}_{n\times n}: \qquad \mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \qquad \lambda - \text{wartość własna}, \ \mathbf{v} - \text{wektor własny}$$

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \iff (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{v} = 0$$

λ jest wartością własną macierzy $\mathbf{A}$ wtedy i tylko wtedy, jeśli jest pierwiastkiem *równania charakterystycznego*:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

$\mathrm{sp}(\mathbf{A})$ – zbiór wszystkich wartości własnych macierzy $\mathbf{A}$ (*spectrum*)

$$\lambda \in \mathrm{sp}(\mathbf{A}) \ \Rightarrow \ \forall \alpha \in \mathbb{R} \ (\lambda + \alpha) \in \mathrm{sp}(\mathbf{A} + \alpha\mathbf{I})$$

Macierze $\mathbf{A}$ i $\mathbf{B}$ są podobne, jeśli istnieje macierz nieosobliwa $\mathbf{S}$ taka, że

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{B}$$

Macierze podobne mają takie same wartości własne, gdyż:

$$\begin{aligned}
\det\left(\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda\mathbf{I}\right) &= \det\left(\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda\mathbf{S}^{-1}\mathbf{S}\right) \\
&= \det \mathbf{S}^{-1}(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{S} \\
&= \det \mathbf{S}^{-1} \det(\mathbf{A} - \lambda\mathbf{I}) \det \mathbf{S} \\
&= \det(\mathbf{A} - \lambda\mathbf{I})
\end{aligned}$$

WARTOŚCI WŁASNE 3

Twierdzenie. Jeśli macierz $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ jest symetryczna, to wszystkie wartości własne i wektory własne są rzeczywiste. Ponadto, istnieje zbiór n ortonormalnych wektorów własnych (istnieje zbiór wektorów własnych tworzący bazę ortonormalną przestrzeni $\mathbb{R}^n$).

Twierdzenie: Jeśli wszystkie wektory własne macierzy $\mathbf{A}_n$ są ortogonalne (ortonormalne), to można tę macierz sprowadzić do postaci diagonalnej $\mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$, przez podobieństwo.

Dowód:

$$\begin{aligned}
\mathbf{A}\mathbf{v}_i &= \lambda_i \mathbf{v}_i, \quad i = 1, ..., n, \ \text{lub w zapisie macierzowym}: \\
\mathbf{A}\mathbf{V} &= \mathbf{V}\mathrm{diag}\{\lambda_i\}, \qquad \text{gdzie } \mathbf{V} \triangleq [\mathbf{v}_1, ..., \mathbf{v}_n] \\
\mathbf{V}^T\mathbf{A}\mathbf{V} &= \mathbf{V}^T\mathbf{V}\mathrm{diag}\{\lambda_i\}
\end{aligned}$$

Ponieważ macierz $\mathbf{V}$ jest ortogonalna, to $\mathbf{V}^T\mathbf{V} = \mathbf{I}$, $\mathbf{V}^T = \mathbf{V}^{-1}$, skąd

$$\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathrm{diag}\{\lambda_i\}.$$

Wniosek: Każdą macierz symetryczną $\mathbf{A}_n$ można sprowadzić do macierzy diagonalnej $\mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$, przez podobieństwo.

WARTOŚCI WŁASNE 4

$\mathbf{A}^*$ – transpozycja macierzy i sprzężenie elementów zespolonych: $\mathbf{A}^* = \bar{\mathbf{A}}^{\mathrm{T}}$.

Macierz kwadratowa o elementach w ogólności zespolonych $\mathbf{A}$ jest:

– *hermitowska*, gdy $\mathbf{A} = \mathbf{A}^*$,

– *unitarna*, gdy $\mathbf{A}^*\mathbf{A} = \mathbf{I}$ $(\Leftrightarrow \mathbf{A}^* = \mathbf{A}^{-1})$.

Jeśli macierz o elementach w ogólności zespolonych $\mathbf{A}_n$ jest normalna, tzn.

$$\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^* \quad (\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{A}\mathbf{A}^{\mathrm{T}}),$$

to istnieje zbiór jej n ortonormalnych wektorów własnych (istnieje zbiór wektorów własnych tworzący bazę ortonormalną przestrzeni $\mathbb{R}^n$).

Twierdzenie (Schura). Każda macierz $\mathbf{A}_n$ jest podobna do macierzy trójkątnej górnej (w ogólności zespolonej) $\mathbf{R}$,

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{bmatrix} \lambda_1 & x & \cdots & x & x \\ 0 & \lambda_2 & \cdots & x & x \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \mathbf{R},$$

gdzie macierz przekształcenia $\mathbf{U}$ jest macierzą unitarną.

– **Metody wyznacznikowe**, wykorzystujące fakt, że wartości własne są zerami wielomianu charakterystycznego $\det(\lambda\mathbf{I} - \mathbf{A}) = 0$ – efektywne dla pojedynczej czy niewielu wartości własnych.

Przykłady:

1. *Metoda bisekcji wykorzystująca ciągi Sturma* – dla macierzy symetrycznych. Pojedynczą wartość własną znajdujemy metodą bisekcji, decyzje o wyborze jednego z pododcinków po każdym podziale podejmuje się łatwo korzystając z pewnych własności tzw. wielomianów Sturma. Macierz wyjściową należy najpierw sprowadzić do *postaci trójdiagonalnej* $\mathbf{T}$:

$$\mathbf{T} = \begin{bmatrix} x & x & 0 & 0 & \cdots & 0 & 0 \\ x & x & x & 0 & \cdots & 0 & 0 \\ 0 & x & x & x & \cdots & 0 & 0 \\ 0 & 0 & x & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & x & x \end{bmatrix}$$

WYBRANE METODY WYZNACZANIA WART. WŁASNYCH 2

2. *Metoda Hymana* – pierwiastki wielomianu charakterystycznego znajdujemy metodą Newtona, wykorzystując specyficzny sposób obliczania wartości wielomianu i jego pochodnej.
   Macierz wyjściową zaleca się na wstępie sprowadzić do *postaci Hessenberga* (tzn. prawie trókątnej górnej) :

$$\mathbf{H} = \begin{bmatrix} x & x & x & x & \cdots & x & x \\ x & x & x & x & \cdots & x & x \\ 0 & x & x & x & \cdots & x & x \\ 0 & 0 & x & x & \cdots & x & x \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & x & x \end{bmatrix}$$

– **Metoda Jacobiego** – diagonalizacja macierzy symetrycznej za pomocą ciągu obrotów Givensa (efektywne do wymiaru $n \cong 10$).

– **Metoda QR** – najbardziej ogólna, efektywna.

WARTOŚCI WŁASNE – METODA QR

Metoda iteracyjna, opiera się na sukcesywnym wykorzystywaniu rozkładu QR.

Metoda QR dla macierzy symetrycznych:

Zalecane (niekonieczne) wstępne sprowadzenie $\mathbf{A}$ do postaci trójdiagonalnej (postać Hessenberga macierzy symetrycznej).

Schemat pojedynczej iteracji:

$$\mathbf{A}^{(k)} = \mathbf{Q}^{(k)}\mathbf{R}^{(k)} \implies \mathbf{R}^{(k)} = \mathbf{Q}^{(k)T}\mathbf{A}^{(k)}$$
$$\mathbf{A}^{(k+1)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)} = \mathbf{Q}^{(k)T}\mathbf{A}^{(k)}\mathbf{Q}^{(k)} \ \text{– też trójdiagonalna}$$

Algorytm podstawowy:

$$\begin{aligned}
\mathbf{A}^{(1)} &= \mathbf{A} \\
\mathbf{A}^{(1)} &= \mathbf{Q}^{(1)}\mathbf{R}^{(1)} \quad \text{(faktoryzacja)} \\
\mathbf{A}^{(2)} &= \mathbf{R}^{(1)}\mathbf{Q}^{(1)} \ (= \mathbf{Q}^{(1)T}\mathbf{A}^{(1)}\mathbf{Q}^{(1)}) \\
\mathbf{A}^{(2)} &= \mathbf{Q}^{(2)}\mathbf{R}^{(2)} \quad \text{(faktoryzacja)} \\
\mathbf{A}^{(3)} &= \mathbf{R}^{(2)}\mathbf{Q}^{(2)} \ (= \mathbf{Q}^{(2)T}\mathbf{A}^{(2)}\mathbf{Q}^{(2)} = \mathbf{Q}^{(2)T}\mathbf{Q}^{(1)T}\mathbf{A}^{(1)}\mathbf{Q}^{(1)}\mathbf{Q}^{(2)}) \\
& \ \text{itd.} \\
\mathbf{A}^{(\mathbf{k})} &\longrightarrow \mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \mathrm{diag}\{\lambda_{\mathrm{i}}\}
\end{aligned}$$

WARTOŚCI WŁASNE – METODA QR 2

$$\text{Jeśli} \quad |\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0 \qquad \text{(różne wartości własne)}$$
$$\text{to} \quad a_{i+1,i}^{(k)} \longrightarrow 0 \;\; \text{liniowo z ilorazem} \left|\frac{\lambda_{i+1}}{\lambda_i}\right|,$$

czyli

$$\frac{a_{i+1,i}^{(k+1)}}{a_{i+1,i}^{(k)}} \approx \left|\frac{\lambda_{i+1}}{\lambda_i}\right|, \;\; i = 1, ..., n-1$$

Algorytm z przesunięciami (k-ta iteracja):

$$\begin{aligned}
\mathbf{A}^{(k)} - p_k\mathbf{I} &= \mathbf{Q}^{(k)}\mathbf{R}^{(k)} \qquad \Longrightarrow \qquad \mathbf{R}^{(k)} = \mathbf{Q}^{(k)T}(\mathbf{A}^{(k)} - p_k\mathbf{I}) \\
\mathbf{A}^{(k+1)} &= \mathbf{R}^{(k)}\mathbf{Q}^{(k)} + p_k\mathbf{I} \\
&= \mathbf{Q}^{(k)T}(\mathbf{A}^{(k)} - p_k\mathbf{I})\mathbf{Q}^{(k)} + p_k\mathbf{I} \\
&= \mathbf{Q}^{(k)T}\mathbf{A}^{(k)}\mathbf{Q}^{(k)}
\end{aligned}$$

Wówczas zbieżność:

$$\frac{a_{i+1,i}^{(k+1)}}{a_{i+1,i}^{(k)}} \approx \frac{|\lambda_{i+1} - p_k|}{|\lambda_i - p_k|}$$

.

WARTOŚCI WŁASNE – METODA QR 3

Wybór p_k:

bliższa $d_n^{(k)}$ wartość własna zaznaczonej podmacierzy 2×2 macierzy $\mathbf{A}^{(k)}$,

$$\text{(tródiagonalna) } \mathbf{A}^{(k)} = \left[\begin{array}{ccc|cc} d_1^{(k)} & \cdot & \cdot & \cdot & 0 \\ \cdot & \ddots & & & \\ \cdot & & d_{n-2}^{(k)} & e_{n-2}^{(k)} & 0 \\ \cline{4-5} \cdot & & e_{n-2}^{(k)} & d_{n-1}^{(k)} & e_{n-1}^{(k)} \\ 0 & \cdot & 0 & e_{n-1}^{(k)} & d_n^{(k)} \end{array}\right]$$

Iterujemy z przesunięciami aż do $e_{n-1}^{(k)} = 0$
(dla macierzy pełnej do wyzerowania ostatniego wiersza oprócz $d_n^{(k)}$).

WARTOŚCI WŁASNE – METODA QR 4

Struktura algorytmu **QR** z przesunięciami:

1. Znajdujemy wartość własną λ_n w przedstawiony sposób,
2. Opuszczamy ostatni wiersz i ostatnią kolumnę aktualnej macierzy $(\mathbf{A})^k$ (deflacja), tzn. uwzględniamy dalej podmacierz $\mathbf{A}_{n-1}^{(k)}$,

$$\mathbf{A}^{(k)} = \begin{bmatrix} d_1^{(k)} & \cdot & \cdot & \cdot & 0 \\ \cdot & \ddots & & & \\ \cdot & & d_{n-2}^{(k)} & e_{n-2}^{(k)} & 0 \\ \cdot & & e_{n-2}^{(k)} & d_{n-1}^{(k)} & 0 \\ 0 & \cdot & 0 & 0 & \lambda_n^{(k)} \end{bmatrix} \Rightarrow \mathbf{A}_{n-1}^{(k)} = \begin{bmatrix} d_1^{(k)} & \cdot & \cdot & \cdot \\ \cdot & \ddots & & \\ \cdot & & d_{n-2}^{(k)} & e_{n-2}^{(k)} \\ \cdot & \cdot & e_{n-2}^{(k)} & d_{n-1}^{(k)} \end{bmatrix}$$

3. Znajdujemy λ_{n-1}, w analogiczny sposób - tzn. przekształcamy macierz trójdiagonalną $\mathbf{A}_{n-1}^{(k)}$ aż do uzyskania $e_{n-2}^{(k)} = 0$ (przy przekształcaniu macierzy pełnej aż do wyzerowania ostatniego wiersza prócz $d_{n-1}^{(k)}$).
4. Opuszczamy ostatni wiersz i ostatnią kolumnę macierzy $\mathbf{A}_{n-1}^{(k)}$ (deflacja),

   itd. od punktu 2 do 4 dla wymiaru macierzy $n-2,\ n-3, \ldots, 2$ – do znalezienia wszystkich wartości własnych.

WARTOŚCI WŁASNE – METODA QR 5

Metoda QR dla macierzy niesymetrycznych:
identyczna struktura, tylko należy stosować arytmetykę liczb zespolonych (metoda może zawieść).

Liczby iteracji algorytmów metody QR bez przesunięć i z przesunięciami:

tolerancja	tol=0.00001		tol=0.0000001		liczba zesp. wart. własnych
alg. z przesunięciami	nie	tak	nie	tak	
macierz symetryczna 5×5	35	7	47	8	0
	154	6	failed	23	0
	91	8	124	9	0
macierz symetryczna 10×10	93	16	129	16	0
	132	24	193	76	0
	failed	15	failed	18	0
macierz niesymetr. 5×5	–	7	–	8	2
	–	55	–	78	4
	–	10	–	11	2
macierz niesymetr. 10×10	–	24	–	27	6
	–	44	–	54	4
	–	26	–	45	4

WARTOŚCI SZCZEGÓLNE

Rozważamy macierz prostokątną $\mathbf{A}_{mxn}$, $m \geqslant n$.

$(\mathbf{A}^{\mathrm{T}}\mathbf{A})_{nxn}$ jest zawsze symetryczna i dodatnio półokreślona, stąd

$$\forall \lambda \in \mathrm{sp}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) \quad \lambda_{\mathrm{i}} \geqslant 0$$

czyli istnieją pierwiastki λ_i:

$$\lambda_i = (\sigma_i)^2, \quad \sigma_i \geqslant 0, \quad i = 1, ..., n$$

σ_i to wartości szczególne (singularne, *singular values*) macierzy $\mathbf{A}$.

Wnioski:

$\mathbf{A}_{nxn}$ symetryczna i dodatnio półokreślona, to

$$\sigma_i = \lambda_i, \quad \lambda_i \in \mathrm{sp}\left(\mathbf{A}\right), \quad \mathrm{i} = 1, ..., \mathrm{n}$$

$\mathbf{A}_{nxn}$ symetryczna, to

$$\sigma_i = \left|\lambda_i\right|, \quad \lambda_i \in \mathrm{sp}\left(\mathbf{A}\right), \quad \mathrm{i} = 1, ..., \mathrm{n}$$

WARTOŚCI SZCZEGÓLNE 2

Uwaga. Wyznaczanie σ_i poprzez wyznaczanie wartości własnych macierzy $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ prowadzi do utraty dokładności, **przykład**:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ \varepsilon & 0 \\ 0 & \varepsilon \end{bmatrix}, \text{ gdzie } \varepsilon < \sqrt{\text{eps}}. \quad \text{Mamy } \mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} 1+\varepsilon^2 & 1 \\ 1 & 1+\varepsilon^2 \end{bmatrix},$$

stąd wartości szczególne macierzy $\mathbf{A}$: $\sigma_1(\mathbf{A}) = \sqrt{2+\varepsilon^2},\ \sigma_2(\mathbf{A}) = |\varepsilon|$.

Natomiast

$$fl\left(\mathbf{A}^{T}\mathbf{A}\right) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad (\text{gdyż } \varepsilon^2 < \text{eps }),$$

stąd wartości szczególne $\mathbf{A}$ liczone numerycznie jako pierwiastki wartości własnych $\mathbf{A}^{T}\mathbf{A}$ w arytmetyce zmiennopozycyjnej wynoszą $\widetilde{\sigma}_1 = \sqrt{2},\ \widetilde{\sigma}_2 = 0$.

Ale przecież $\sigma_2(\mathbf{A}) = |\varepsilon|$, co potwierdza dużą utratę dokładności,

np. dla $eps = 10^{-12}$ i $\varepsilon = 10^{-7}$ $(< \sqrt{eps} = 10^{-6})$ powinniśmy dostać $\sigma_2(\mathbf{A}) = 10^{-7}$ $(\gg eps)$, a nie 0.

Stosuje się algorytmy SVD **nie prowadzące do utraty dokładności** (np. algorytm Goluba-Reinscha).

DEKOMPOZYCJA SVD

Twierdzenie (dekompozycja SVD).

Dla dowolnej macierzy $\mathbf{A}_{mxn}$ istnieją macierze ortogonalne $\mathbf{U}_{mxm}$, $\mathbf{V}_{nxn}$ i macierz $\mathbf{\Sigma}_{m\times n}$ takie, że:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}}, \quad \text{gdzie} \;\; \mathbf{\Sigma} = \begin{bmatrix} \mathbf{D}_{n\times n} \\ \mathbf{0}_{(m-n)\times n} \end{bmatrix}_{\mathrm{m\times n}} = \left[\begin{array}{ccccc} \sigma_1 & 0 & \cdots & & 0 \\ 0 & \ddots & & & \\ \vdots & & \sigma_k & & \vdots \\ & & & \ddots & 0 \\ 0 & & \cdots & 0 & \sigma_n \\ \hline & & \mathbf{0} & & \end{array}\right],$$

gdzie:

- $\mathbf{D}_{n\times n} = \mathrm{diag}\,\{\sigma_{\mathrm{i}},\ \mathrm{i} = 1, ..., \mathrm{n}\}$, $\sigma_i \geqslant 0$, zwyczajowo $\sigma_i \geqslant \sigma_{i+1}$. (Jeśli istnieją zerowe, to $\mathbf{D}$ osobliwa, a $\mathbf{A}$ niepełnego rzędu - ilość niezerowych σ_i to rząd $k \leqslant n$ macierzy $\mathbf{A}$),
- $\mathbf{U}_{m\times m}$ macierz, której kolumnami są ortonormalne wektory własne macierzy $\mathbf{A}\mathbf{A}^{\mathrm{T}}$,
- $\mathbf{V}_{n\times n}$ macierz, której kolumnami są ortonormalne wektory własne macierzy $\mathbf{A}^{\mathrm{T}}\mathbf{A}$.

LINIOWE ZADANIE NAJMNIEJSZYCH KWADRATÓW LZNK

Dane: $\mathbf{A}_{mxn}$, $(m > n,\ \mathrm{rank}\,\mathbf{A} \leqslant \mathrm{n})$, $\mathbf{b} \in \mathbb{R}^m$.

Zadanie: znaleźć wektor $\widehat{\mathbf{x}} \in \mathbb{R}^n$ taki, że

$$\forall \mathbf{x} \in \mathbb{R}^n \qquad \|\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}\|_2 \leqslant \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \,.$$

Interpretacja geometryczna: $\mathbf{A}\widehat{\mathbf{x}}$ jest rzutem ortogonalnym wektora $\mathbf{b}$ na podprzestrzeń rozpinaną przez kolumny macierzy $\mathbf{A}$ (tzn. podprzestrzeń $\{\mathbf{y} \in \mathbb{R}^m :\ \mathbf{y} = \mathbf{A}\mathbf{x}, \mathbf{x} \in \mathbb{R}^n\}$).

Zadanie LZNK może mieć niejednoznaczne rozwiązanie, dla ujednoznacznienia wymaga się zwykle wektora rozwiązania o najmniejszej normie, tzn.

$$(\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 = \|\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}\|_2) \Rightarrow \|\widehat{\mathbf{x}}\|_2 \leqslant \|\mathbf{x}\|_2$$

lub wektora o największej liczbie zerowych składników.

LNZK jest równoważne minimalizacji następującej funkcji kwadratowej $J(x)$:

$$J(x) = (\mathbf{b} - \mathbf{A}\mathbf{x})^{\mathrm{T}}(\mathbf{b} - \mathbf{A}\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} - 2\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{b} + \mathbf{b}^{\mathrm{T}}\mathbf{b}.$$

LINIOWE ZADANIE NAJMNIEJSZYCH KWADRATÓW 2

Niech k – rząd macierzy $\mathbf{A}$, $k \leqslant n$.

Metody rozwiązywania LZNK:

a) *Macierz $\mathbf{A}$ pełnego rzędu $(k = n)$* :

Jeśli $k = n$ to $\mathbf{A}^{\mathrm{T}}\mathbf{A} > 0$, stąd funkcja $J(x) = \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} - 2\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{b} + \mathbf{b}^{\mathrm{T}}\mathbf{b}$ jest ściśle wypukła i ma jednoznaczne minimum w punkcie $J^{'}(x) = \mathbf{0}$:

$$J^{'}(x)^{\mathrm{T}} = 2\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} - 2\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{0}.$$

Wynikający stąd układ równań liniowych to *układ równań normalnych*:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b},$$

który ma jednoznaczne rozwiązanie.

Dla *słabo uwarunkowanej macierzy* $\mathbf{A}$ uwarunkowanie $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ staje się jeszcze gorsze, gdyż

$$\mathrm{cond}_2(\mathbf{A}^{\mathrm{T}}\mathbf{A}) = (\mathrm{cond}_2\mathbf{A})^2 = \sigma_1^2/\sigma_{\mathrm{n}}^2,$$

gdzie σ_1 i σ_n to największa i najmniejsza wartość szczególna macierzy $\mathbf{A}$.

Wówczas zalecane jest skorzystanie z rozkładu $\mathbf{QR}$ macierzy $\mathbf{A}$.

LINIOWE ZADANIE NAJMNIEJSZYCH KWADRATÓW 3

Wykorzystując wąski rozkład QR macierzy $\mathbf{A}$, układ równań normalnych

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b}$$

możemy zapisać w postaci

$$\mathbf{R}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Q}\mathbf{R}\mathbf{x} = \mathbf{R}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{b}.$$

$\mathbf{Q}$ ma kolumny wzajemnie ortogonalne, tzn. $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$, zaś $\mathbf{R}$ jest nieosobliwa (gdyż $k = n$), stąd dostajemy dobrze określony układ równań liniowych

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^{\mathrm{T}}\mathbf{b}.$$

Jeśli zastosujemy rozkład QR **wąski nieunormowany** (np. bezpośrednio z ortogonalizacji Grama-Schmidta), $\mathbf{A}_{m\times n} = \bar{\mathbf{Q}}_{m\times n}\bar{\mathbf{R}}_{n\times n}$, to układ równań normalnych jest postaci:

$$\bar{\mathbf{R}}^{\mathrm{T}}\bar{\mathbf{Q}}^{\mathrm{T}}\bar{\mathbf{Q}}\bar{\mathbf{R}}\mathbf{x} = \bar{\mathbf{R}}^{\mathrm{T}}\bar{\mathbf{Q}}^{\mathrm{T}}\mathbf{b},$$

skąd wynika

$$\bar{\mathbf{R}}\mathbf{x} = (\bar{\mathbf{Q}}^{\mathrm{T}}\bar{\mathbf{Q}})^{-1}\bar{\mathbf{Q}}^{\mathrm{T}}\mathbf{b},$$

gdzie $\bar{\mathbf{Q}}^{\mathrm{T}}\bar{\mathbf{Q}} = \mathrm{diag}\{d_1, d_2, \ldots, d_n\}$, $d_i = \bar{\mathbf{q}}_i^{\mathrm{T}}\bar{\mathbf{q}}_i$, $i = 1, \ldots, n$.

LINIOWE ZADANIE NAJMNIEJSZYCH KWADRATÓW 4

b) *Jeśli $k < n$ (macierz $\mathbf{A}$ niepełnego rzędu)*, to zalecany jest algorytm oparty na rozkładzie SVD macierzy $\mathbf{A}$.

Stosując SVD i własność macierzy ortogonalnej $||\mathbf{U}^{\mathrm{T}}\mathbf{y}||_2 = ||\mathbf{y}||_2$ mamy:

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 = \left\|\mathbf{b} - \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{x}\right\|_2 = \left\|\mathbf{U}^T\mathbf{b} - \mathbf{\Sigma}\left(\mathbf{V}^T\mathbf{x}\right)\right\|_2 = \left\|\widetilde{\mathbf{b}} - \mathbf{\Sigma}\widetilde{\mathbf{x}}\right\|_2,$$

gdzie $\widetilde{\mathbf{b}} = \mathbf{U}^T\mathbf{b}$, $\widetilde{\mathbf{x}} = \mathbf{V}^T\mathbf{x}$.

Oznaczmy: $\widetilde{\mathbf{b}} = [\widetilde{\mathbf{b}}_1^T\ \widetilde{\mathbf{b}}_2^T]^T$, $\widetilde{\mathbf{b}}_1 \in \mathbb{R}^n, \widetilde{\mathbf{b}}_2 \in \mathbb{R}^{m-n}$, stąd

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 = \left\|\widetilde{\mathbf{b}} - \mathbf{\Sigma}\widetilde{\mathbf{x}}\right\|_2 = \left\|\begin{array}{c} \widetilde{\mathbf{b}}_1 - \mathbf{D}\widetilde{\mathbf{x}} \\ \widetilde{\mathbf{b}}_2 \end{array}\right\|_2 = \left\|\begin{array}{c} \widetilde{b}_1 - \sigma_1\widetilde{x}_1 \\ \vdots \\ \widetilde{b}_k - \sigma_k\widetilde{x}_k \\ \widetilde{b}_{k+1} \\ \vdots \\ \widetilde{b}_m \end{array}\right\|_2,$$

gdzie $k \leqslant n$ to liczba niezerowych wartości szczególnych (rząd macierzy $\mathbf{A}$).

LINIOWE ZADANIE NAJMNIEJSZYCH KWADRATÓW 5

Norma $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$ osiąga minimum $\sqrt{\sum_{i=k+1}^{m}(\widetilde{b}_i)^2}$ dla każdego $\widetilde{\mathbf{x}}$ postaci:

$$\widetilde{\mathbf{x}} = \begin{bmatrix} \widetilde{x}_1 = \widetilde{b}_1/\sigma_1 \\ \vdots \\ \widetilde{x}_k = \widetilde{b}_k/\sigma_k \\ \widetilde{x}_{k+1} \text{ dowolne} \\ \vdots \\ \widetilde{x}_n \text{ dowolne} \end{bmatrix}$$

Rozwiązanie jednoznaczne $\widehat{\widetilde{\mathbf{x}}}$ o minimalnej normie uzyskujemy przyjmując

$$\widehat{\widetilde{\mathbf{x}}} = \begin{bmatrix} \widetilde{b}_1/\sigma_1 \\ \vdots \\ \widetilde{b}_k/\sigma_k \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\text{tzn.} \;\; \widehat{\widetilde{\mathbf{x}}} = \mathbf{\Sigma}^+\widetilde{\mathbf{b}}, \quad \mathbf{\Sigma}^+ = \left[\mathbf{D}^+_{n\times n} \;\; \mathbf{0}_{n\times(m-n)}\right], \;\; \mathbf{D}^+ = \text{diag}\{\frac{1}{\sigma_1}, ..., \frac{1}{\sigma_\text{k}}, 0, ..., 0\}$$

Stąd $\quad \mathbf{V}^T\widehat{\mathbf{x}} = \mathbf{\Sigma}^+\mathbf{U}^T\mathbf{b} \quad \Rightarrow \quad \widehat{\mathbf{x}} = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T\mathbf{b},$

$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T$ – **macierz pseudoodwrotna macierzy $\mathbf{A}$**

LINIOWE ZADANIE NAJMNIEJSZYCH KWADRATÓW – Przykład

Rozwiązać układ równań $\mathbf{A}\mathbf{x} = \mathbf{b}$, w sensie minimalizacji sumy kwadratów reszt, gdzie

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}.$$

Macierz $\mathbf{A}$ jest pełnego rzędu, stąd można rozwiązać układ równań normalnych lub skorzystać z rozkładu QR.

Układ równań normalnych:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b} \quad \Leftrightarrow \quad \begin{bmatrix} 9 & -9 \\ -9 & 18 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

skąd rozwiązanie $\mathbf{x} = [\frac{4}{3}, 1]^{\mathrm{T}}$.

Do rozwiązania metodą rozkładu QR wystarczy rozkład wąski, rozkład taki dla podanej macierzy $\mathbf{A}$ wyliczyliśmy uprzednio w przykładzie, stąd

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^{\mathrm{T}}\mathbf{b} \quad \Leftrightarrow \quad \begin{bmatrix} 3 & -3 \\ 0 & 3 \end{bmatrix} \mathbf{x} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

PRZEKSZTAŁCENIE (ODBICIE) HOUSEHOLDERA*

Przekształcenie zdefiniowane macierzą $\mathbf{P}$ postaci

$$\mathbf{P} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T, \quad \text{gdzie } \mathbf{w} \in \mathbb{R}^{\mathrm{n}}, \ \|\mathbf{w}\| = 1.$$

Własności:

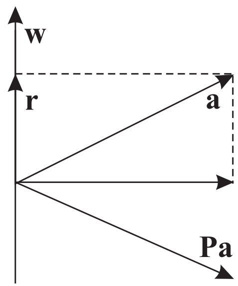
$$\begin{aligned} \mathbf{P} &= \mathbf{P}^T \quad \text{(symetria)}, \\ \mathbf{P}\mathbf{P}^T &= \mathbf{I} \quad (\text{tzn. } \mathbf{P} = \mathbf{P}^{-1}, \text{ortogonalność}), \end{aligned}$$

gdyż

$$\left(\mathbf{I} - 2\mathbf{w}\mathbf{w}^T\right)\left(\mathbf{I} - 2\mathbf{w}\mathbf{w}^T\right) = \mathbf{I} - 4\mathbf{w}\mathbf{w}^T + 4\mathbf{w}\left(\mathbf{w}^T\mathbf{w}\right)\mathbf{w}^T = \mathbf{I}$$

Interpretacja geometryczna:

$$\mathbf{y} = \mathbf{P}\mathbf{a} = \left(\mathbf{I} - 2\mathbf{w}\mathbf{w}^T\right)\mathbf{a} = \mathbf{a} - 2\left(\mathbf{w}^T\mathbf{a}\right)\mathbf{w} = \mathbf{a} - 2\mathbf{r},$$



$\mathbf{Pa}$ jest odbiciem zwierciadlanym $\mathbf{a}$ względem płaszczyzny prostopadłej do $\mathbf{w}$.
Odpowiednio dobierając $\mathbf{w}$ uzyskujemy dowolne położenie wektora $\mathbf{y} = \mathbf{Pa}$.

** materiał uzupełniający*

PRZEKSZTAŁCENIE HOUSEHOLDERA - zastosowanie

Zadanie: uzyskanie $\mathbf{Pa} \parallel e_1$, gdzie $e_1 = [1\ 0\ 0\ \dots\ 0]^T$.

Przyjmujemy $\mathbf{P} = \mathbf{I} - \frac{1}{K}\mathbf{u}\mathbf{u}^T$, $\mathbf{u} = \mathbf{a} \pm \|\mathbf{a}\|\, \mathbf{e}_1$, $K = \frac{1}{2}\|\mathbf{u}\|^2$, $(\mathbf{w} = \sqrt{2/K}\mathbf{u})$

gdzie znak dobieramy tak, aby uzyskać większą normę $\mathbf{u}$, tzn.

$$\mathbf{u} = \mathbf{a} + \mathrm{sgn}(\mathrm{a}_1)\,\|\mathbf{a}\|\,\mathbf{e}_1, \quad \text{gdzie } \mathrm{a}_1 = (\mathbf{e}_1)^{\mathrm{T}}\,\mathbf{a}.$$

Wówczas:

$$\begin{aligned}
\mathbf{Pa} &= \mathbf{a} - \frac{\mathbf{u}}{K}\left(\mathbf{a} + \mathrm{sgn}(\mathrm{a}_1)\,\|\mathbf{a}\|\,\mathbf{e}_1\right)^T \mathbf{a} \\
&= \mathbf{a} - \frac{2\mathbf{u}\left(\|\mathbf{a}\|^2 + \|\mathbf{a}\|\,|a_1|\right)}{\|\mathbf{a} + \mathrm{sgn}(\mathrm{a}_1)\,\|\mathbf{a}\|\,\mathbf{e}_1\|^2} \\
&= \mathbf{a} - \frac{2\mathbf{u}\left(\|\mathbf{a}\|^2 + \|\mathbf{a}\|\,|a_1|\right)}{\|\mathbf{a}\|^2 + \|\mathbf{a}\|^2 + 2\,\|\mathbf{a}\|\,|a_1|} \\
&= \mathbf{a} - \mathbf{u} = -\mathrm{sgn}(\mathrm{a}_1)\,\|\mathbf{a}\|\,\mathbf{e}_1
\end{aligned}$$

Nie należy liczyć $\mathbf{P}$, tylko wynik $\mathbf{Pa}$: $\mathbf{P} = (\mathbf{I} - 2\mathbf{w}\mathbf{w}^T)\mathbf{a} = \mathbf{a} - (2\mathbf{w}^T\mathbf{a})\mathbf{w}$.

Przekształcenie Householdera jest algorytmem numerycznie poprawnym.