IV Wykład: **5 XI 10r**

Opracowano na podstawie [2, ss. 299–303], [3, ss. 19–22] i [4, ss. 204–210].

IV–1.1 Definicja (macierz Hauseholdera)
Niech $v \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$ będzie wektorem jednostkowym $v^T v = \|v\|_2 = 1$ zapisanym jako macierz jednokolumnowa. Wtedy macierzą Hauseholdera nazywamy macierz:

$$H \stackrel{\text{def}}{=} I - 2vv^T.$$

IV–1.2 Uwaga
Macierz Hauseholdera jest
1. *$H = H^T$ – symetryczna,*
2. *$H^2 = I$ – idempotentna.*

Dowód: Uzasadnienie własności **1**. jest stosunkowo proste:

$$H^T = (I - 2vv^T)^T = I - (2vv^T)^T = I - 2(v^T)^T v^T = I - 2vv^T = H.$$

Rachunek dla **2**. też jest standardowy:

$$\begin{aligned} H^2 &= H\,H = (I - 2vv^T)(I - 2vv^T) = (I - 2vv^T - 2vv^T + 4(vv^T)(vv^T)) = \\ &= (I - 4vv^T + 4v(v^T v)v^T) = (I - 4vv^T + 4v\cdot 1\cdot v^T) = I \end{aligned}$$

□

Przedstawimy teraz interpretację geometryczną działania macierzy Hauseholdera na wektor $a = [a_1, \ldots, a_n]^T$. Dla $v \in \mathbb{R}^n,\ \|v\|_2 = 1$ mamy następujące przedstawienie:

$$a = (vv^T)a + (a - vv^T a) = v(v^T a) + y. \tag{IV–1.1}$$

Ponieważ zachodzi

$$\begin{aligned} v^T y &= v^T(a - vv^T a) = v^T a - (v^T v)v^T a = v^T a - v^T(vv^T)a = \\ &= v^T a - v^T\cdot 1\cdot a = 0 \end{aligned}$$

stąd więc w (IV–1.1) przedstawiliśmy dowolny wektor a jako sumę ortogonalną wektora $(vv^T)a$ współliniowego z v oraz $y = a - (vv^T)a$ ortogonalnego do v. Dalej mamy[1]

$$Ha = a - 2vv^T a = -vv^T a + (a - vv^T a) = -vv^T a + y$$

co oznacza, że działanie macierzy H na wektor a polega na jego odbiciu względem hiperpłaszczyzny ortogonalnej do wektora v.
Dla wektora $a \in \mathbb{R}^n,\ a \neq 0$ znajdziemy jednostkowy wektor $v \in \mathbb{R}^n,\ \|v\| = 1$ o własności

$$Ha = \sigma e$$

gdzie H jest macierzą Hauseholdera wyznaczoną przez v, wektor $e \in \mathbb{R}^n$ postaci $e^T = [1, 0, \ldots, 0]^T$ oraz σ jest liczbą

$$\sigma = \begin{cases} \frac{a_1}{|a_1|}\sqrt{a^T a}, & a_1 \neq 0, \\ \sqrt{a^T a}, & a_1 = 0. \end{cases}$$

[1] W rachunkach należy zwrócić uwagę na łączność mnożenia macierzy oraz na wyrażenie $v^T a$ które jest liczbą podczas gdy vv^T jest macierzą $n \times n$.

Niech $u = a \mp \sigma e$. Wtedy zachodzi

$$u^T u = (a \mp \sigma e)^T (a \mp \sigma e) = a^T a \mp \sigma a^T e \mp \sigma e^T a + \sigma^2 e^T e = \quad \text{(IV–1.2)}$$

$$= 2(a^T a \mp |a_1| \sqrt{a^T a})$$

oraz

$$u^T a = (a \mp \sigma e)^T a = a^T a \mp |a_1| \sqrt{a^T a} = \frac{1}{2} u^T u \quad \text{(IV–1.3)}$$

gdzie w ostatniej równości skorzystaliśmy z (IV–1.2). Bez utraty[2] ogólności możemy założyć, że $\sqrt{a^T a} - |a_1| > 0$. Niech

$$v = \frac{u}{\sqrt{u^T u}}.$$

i nich H będzie macierzą Hauseholdera wyznaczoną przez ten wektor v. Wtedy z (IV–1.3) wynika, iż

$$Ha = (I - 2vv^T)a = a - 2(vv^T)a = a - \frac{2}{u^T u} u(u^T a) =$$

$$= a - \frac{2}{u^T u} u \, \frac{1}{2} u^T u = a - u = \pm \sigma e.$$

W ostatniej równości odwołaliśmy się do określenia wektora u.
Dla danej macierzy $A \in \mathbb{R}^{n \times n}$ wykazana własność jest podstawą do wyznaczenia rozkładu

$$A = QR$$

gdzie Q jest ortogonalna a R górnie trójkątna. Konstrukcja macierzy Q będzie indukcyjna względem kolumn. Niech więc

$$A = [a_1, \ldots, a_n]$$

gdzie kolumny macierz A oznaczyliśmy następująco

$$a_i = [a_{1i}, \ldots, a_{ni}]^T, \quad i = 1, \ldots, n.$$

W pierwszym kroku dla pierwszej kolumny $i = 1$ niech

$$a \stackrel{\text{def}}{=} a_1 \in \mathbb{R}^{n \times 1} \quad \text{oraz} \quad e = [1, 0, \ldots, 0]^T \in \mathbb{R}^{n \times 1}.$$

Wtedy istnieje macierz Hauseholdera $H_1 \in \mathbb{R}^{n \times n}$ o własności

$$\begin{bmatrix} \sigma_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = H_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}.$$

Jeżeli macierz A pomnożymy lewostronnie przez $Q_1 = H_1$, to otrzymamy

$$\begin{bmatrix} \sigma_1 & * & \cdots & * \\ 0 & a^{(1)}_{11} & \cdots & a^{(1)}_{1\,n-1} \\ \vdots & \vdots & & \vdots \\ 0 & a^{(1)}_{n-1\,1} & \cdots & a^{(1)}_{n-1\,n-1} \end{bmatrix} = Q_1 \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

[2] Bo inaczej $a = a_1 e$ – byłby już wymaganej postaci.

Znakiem $*$ oznaczyliśmy elementy nieistotne dla dalszych naszych rozważań. W drugim kroku za macierz $\widetilde{A}_1 \in \mathbb{R}^{(n-1)\times(n-1)}$ przyjmujemy

$$\widetilde{A}_1 = \begin{bmatrix} a^{(1)}_{11} & \cdots & a^{(1)}_{1\,n-1} \\ \vdots & \ldots & \vdots \\ a^{(1)}_{n-1\,1} & \cdots & a^{(1)}_{n-1\,n-1} \end{bmatrix}$$

i procedurę powtarzamy, aby otrzymać macierz Hauseholdera $\widetilde{H}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$ o własności

$$\begin{bmatrix} \sigma_2 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \cdots & \vdots \\ 0 & * & \cdots & * \end{bmatrix} = \widetilde{H}_2\widetilde{A}_1$$

Definiujemy nową macierz ortogonalną $Q_2 \in \mathbb{R}^{n\times n}$, $\mathbf{0}^T = [0,\ldots,0]^T \in \mathbb{R}^{(n-1)\times 1}$

$$Q_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \widetilde{H}_2 \end{bmatrix}$$

dla której zachodzi

$$Q_2Q_1A = Q_2 \begin{bmatrix} \sigma_1 & * \\ \mathbf{0} & \widetilde{A}_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 & * & * & \cdots & * \\ 0 & \sigma_2 & * & \cdots & * \\ 0 & 0 & a^{(2)}_{11} & \cdots & a^{(2)}_{1\,n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a^{(2)}_{n-2\,1} & \cdots & a^{(2)}_{n-2\,n-2} \end{bmatrix}.$$

W kolejnym kroku mamy $\widetilde{A}_2 \in \mathbb{R}^{(n-2)\times(n-2)}$ z

$$\widetilde{A}_2 = \begin{bmatrix} a^{(2)}_{11} & \cdots & a^{(2)}_{1\,n-2} \\ \vdots & \ldots & \vdots \\ a^{(2)}_{n-2\,1} & \cdots & a^{(2)}_{n-2\,n-2} \end{bmatrix}$$

i macierz Hauseholdera $\widetilde{H}_3 \in \mathbb{R}^{(n-2)\times(n-2)}$, która służy do zdefiniowania macierzy ortogonalnej Q_3 (tu $\mathbf{0}^T = [0,\ldots,0]^T \in \mathbb{R}^{(n-2)\times 1}$)

$$Q_3 = \begin{bmatrix} 1 & 0 & \mathbf{0}^T \\ 0 & 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{0} & \widetilde{H}_3 \end{bmatrix}$$

dla której zachodzi

$$Q_3Q_2Q_1A = \begin{bmatrix} \sigma_1 & * & * & * & \cdots & * \\ 0 & \sigma_2 & * & * & \cdots & * \\ 0 & 0 & \sigma_3 & * & \cdots & * \\ 0 & 0 & 0 & a^{(3)}_{11} & \cdots & a^{(3)}_{1\,n-3} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & a^{(3)}_{n-3\,1} & \cdots & a^{(3)}_{n-3\,n-3} \end{bmatrix}. \tag{IV–1.4}$$

Postępując tak dalej otrzymamy macierze ortogonalne $Q_1, \ldots, Q_{n-1}$, które po wymnożeniu z lewej strony macierzy A

$$Q_{n-1} \ldots Q_2Q_1A = R$$

dadzą macierz górnie trójkątną. Definiujemy macierz Q jako

$$Q \stackrel{\text{def}}{=} (Q_{n-1} \ldots Q_2Q_1)^T$$

Z Uwagi IV–1.2 wynikają równości

$$Q = (Q_{n-1} \ldots Q_2Q_1)^T = Q_1Q_2 \ldots Q_{n-1}, \quad Q^TQ = QQ^T = I$$

które gdy zastosujemy je do równości (IV–1.4), to otrzymamy

$$A = QQ^TA = QR.$$

Jest to szukany rozkład macierzy A.

□

Przykład z [1]: Mamy macierz

$$A = \begin{bmatrix} 1 & 7 & 5 \\ \sqrt{3} & -\sqrt{3} & \sqrt{3} \\ 0 & 2 & 2 \end{bmatrix}$$

Pierwsza macierz ortogonalna jest następująca

$$Q_1 = \begin{bmatrix} -0.5000 & -0.8660 & 0 \\ -0.8660 & 0.5000 & 0 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

po wymnożeniu z lewej strony macierzy A otrzymujemy:

$$Q_1A = \begin{bmatrix} -2.0000 & -2.0000 & -4.0000 \\ -0.0000 & -6.9282 & -3.4641 \\ 0 & 2.0000 & 2.0000 \end{bmatrix}$$

Wybieramy podmacierz

$$\widetilde{A}_1 = \begin{bmatrix} -6.9282 & -3.4641 \\ 2.0000 & 2.0000 \end{bmatrix}$$

Znajdujemy macierz Hauseholdera wyznaczoną przez pierwszą kolumnę macierzy $\widetilde{A}_1$

$$\widetilde{H}_2 = \begin{bmatrix} -0.9608 & 0.2774 \\ 0.2774 & 0.9608 \end{bmatrix}$$

Przy pomocy macierzy $\widetilde{H}_2$ budujemy macierz ortogonalną P_2:

$$Q_2 = \begin{bmatrix} 1.0000 & 0 & 0 \\ 0 & -0.9608 & 0.2774 \\ 0 & 0.2774 & 0.9608 \end{bmatrix}$$

Teraz

$$Q = Q_1 Q_2 = \begin{bmatrix} -0.5000 & 0.8321 & -0.2402 \\ -0.8660 & -0.4804 & 0.1387 \\ 0 & 0.2774 & 0.9608 \end{bmatrix}$$

oraz

$$R = \begin{bmatrix} -2.0000 & -2.0000 & -4.0000 \\ 0.0000 & 7.2111 & 3.8829 \\ -0.0000 & 0.0000 & 0.9608 \end{bmatrix}$$

I sprawdzamy $A = QR$

$$QR = \begin{bmatrix} 1.0000 & 7.0000 & 5.0000 \\ 1.7321 & -1.7321 & 1.7321 \\ 0.0000 & 2.0000 & 2.0000 \end{bmatrix}$$

oraz ortogonalność macierzy Q:

$$Q' * Q = \begin{bmatrix} 1.0000 & -0.0000 & 0.0000 \\ -0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 1.0000 \end{bmatrix}.$$

Skrypt MatLabowski przy pomocy którego wykonano obliczenia:

```
clear all;
clc;
format short;
r = sqrt(3);
A = [1 7 5; r -r r; 0 2 2];
%
a = A(:,1);
r = sign(a(1))*sqrt(a'*a);
I = eye(3);
u = a+r*I(:,1);
v = u/sqrt(u'*u);
```

```
Q1 = I-2*v*v';
A0 = Q1*A;
A1 = A0(2:3,2:3);
Q1
A0
%
A1
a = A1(:,1);
r = sign(a(1))*sqrt(a'*a);
I = eye(2);
u = a+r*I(:,1);
v = u/sqrt(u'*u);
H2 = I-2*v*v';
Q2 = [1 [0 0];[0 0]' H2];
H2
Q2
Q = Q1*Q2;
Q
R = Q'*A;
R
% sprawdzamy
Q*R
Q'*Q
```

Literatura

[1] M. Fröhner and G. Windisch, Numerical mathematics, Institut für Mathematik, BTU Cottbus, 2002,

for Students in Mathematics, and Economathematics at the University of Zielona Góra.

[2] D. Kincaid and W. Cheney, Numerical analysis, Brooks, USA, 1996,

Książka jest w bibliotece Instytutu Informatyki Uniwersytetu Wrocławskiego.

[3] R. Kress, Numerical analysis, SV, Berlin, 1998.

[4] A. Quarteroni, R. Sacco, and F. Saleri, Numerical mathematics, Springer, 2000.

Wrocław, dnia 15 listopada 2010

(–) Krzysztof Tabisz
tabisz@math.uni.wroc.pl