Rozdział 7

Układy równań liniowych

W tym rozdziale zajmiemy się rozwiązywaniem układów równań liniowych (2.2). Stosując zapis macierzowy zadanie formułujemy następująco. Dla danej macierzy (*współczynników*) $A \in \mathbf{K}^{m,n}$ oraz wektora (*wyrazów wolnych*) $\vec{b} \in \mathbf{K}^m$ należy znaleźć wszystkie wektory (*niewiadome*) $\vec{x}$ spełniające równość

$$A * \vec{x} = \vec{b}. \tag{7.1}$$

7.1 Zbiór rozwiązań

7.1.1 Twierdzenie Kroneckera-Capelliego

Mamy trzy możliwości:

(i) $\forall \vec{x} \in \mathbf{K}^n \quad A * \vec{x} \neq \vec{b} \quad \Longrightarrow$ układ jest *sprzeczny*

(ii) $\exists \vec{x} \in \mathbf{K}^n \quad A * \vec{x} = \vec{b} \quad \Longrightarrow$ układ jest *niesprzeczny*

(iii) $!\exists \vec{x} \in \mathbf{K}^n \quad A * \vec{x} = \vec{b} \quad \Longrightarrow$ układ jest *oznaczony* [1]

Twierdzenie 7.1 *(*KRONECKERA-CAPELLIEGO)
Układ $A * \vec{x} = \vec{b}$ jest niesprzeczny wtedy i tylko wtedy gdy

$$\mathrm{rz}(A) = \mathrm{rz}([A, \vec{b}]),$$

tzn. rząd macierzy A jest równy rzędowi A rozszerzonej o wektor $\vec{b}$.

[1] Symbol "$!\exists$" czytamy jako "istnieje dokładnie jeden".

Dowód. Jeśli $\mathrm{rz}([A,\vec{b}]) = \mathrm{rz}(A)$ to wektor $\vec{b}$ należy do przestrzeni rozpiętej przez wektory-kolumny macierzy A. To zaś oznacza, że $\vec{b}$ jest liniową kombinacją tych wektorów. Współczynniki tej kombinacji tworzą rozwiązanie $\vec{x}$ układu.

Z drugiej strony, jeśli istnieje $\vec{x} \in \mathbf{K}^n$ taki, że $A*\vec{x} = \vec{b}$ to $\vec{b}$ jest kombinacją liniową wektorów-kolumn macierzy A, czyli $\vec{b}$ należy do przestrzeni rozpiętej na tych wektorach. To zaś implikuje że $\mathrm{rz}([A,\vec{b}]) = \mathrm{rz}(A)$ i kończy dowód.

Możemy równoważnie stwierdzić, że układ $A*\vec{x} = \vec{b}$ jest niesprzeczny wtedy i tylko wtedy gdy $\vec{b} \in \mathcal{R}(A)$, czyli wektor wyrazów wolnych leży w obrazie macierzy współczynników.

7.1.2 Zbiór rozwiązań jako warstwa

Niech

$$\mathcal{L}(A,\vec{b}) \;=\; \{\, \vec{x} \in \mathbf{K}^n : \; A*\vec{x} = \vec{b} \,\}$$

będzie zbiorem wszystkich rozwiązań układu $A*\vec{x} = \vec{b}$.

Definicja 7.1 *Powiemy, że dwa układy, $A_1 * \vec{x} = \vec{b}_1$ oraz $A_2 * \vec{x} = \vec{b}_2$, są* równoważne *gdy mają ten sam zbiór rozwiązań, tzn. gdy*

$$\mathcal{L}(A_1,\vec{b}_1) = \mathcal{L}(A_2,\vec{b}_2).$$

Twierdzenie 7.2 *Jeśli układ $A*\vec{x} = \vec{b}$ jest niesprzeczny to zbiór rozwiązań*

$$\mathcal{L}(A,\vec{b}) \;=\; \{\, \vec{x}_0 + \vec{y} : \; \vec{y} \in \mathcal{N}(A) \,\} \;=\; W\left(\vec{x}_0, \mathcal{N}(A)\right),$$

gdzie $\vec{x}_0$ jest dowolnym rozwiązaniem szczególnym układu.

Dowód. Jeśli $\vec{x}_0$ jest rozwiązaniem szczególnym i $\vec{y} \in \mathcal{N}(A)$ to

$$A*(\vec{x}_0 + \vec{y}) = A*\vec{x}_0 + A*\vec{y} = \vec{b} + \vec{0} = \vec{b},$$

czyli $\vec{x}_0 + \vec{y}$ jest też rozwiązaniem. To zaś implikuje że $W\left(\vec{x}_0, \mathcal{N}(A)\right) \subseteq \mathcal{L}(A,\vec{b})$.

Z drugiej strony, jeśli $A*\vec{x} = \vec{b}$ to $A*(\vec{x}-\vec{x}_0) = \vec{b}-\vec{b} = \vec{0}$, czyli $(\vec{x}-\vec{x}_0) \in \mathcal{N}(A)$. A więc $\vec{x} = \vec{x}_0 + (\vec{x}-\vec{x}_0)$ jest z jednej strony rozwiązaniem układu, a z drugiej elementem warstwy $W\left(\vec{x}_0, \mathcal{N}(A)\right)$. Stąd $\mathcal{L}(A,\vec{b}) \subseteq W\left(\vec{x}_0, \mathcal{N}(A)\right)$.

7.1.3 Układy nieosobliwe

Rozpatrzmy przez chwilę układy z macierzami kwadratowymi.

Twierdzenie 7.3 *Macierz kwadratowa $A \in \mathbf{K}^{n,n}$ jest nieosobliwa wtedy i tylko wtedy gdy* $\mathrm{rz}(A) = n$.

Dowód. Wobec nierówności

$$n = \mathrm{rz}(I_n) = \mathrm{rz}(A * A^{-1}) \le \min\big(\mathrm{rz}(A), \mathrm{rz}(A^{-1})\big)$$

mamy, że jeśli A jest nieosobliwa to $\mathrm{rz}(A) = n = \mathrm{rz}(A^{-1})$. Z drugiej strony, jeśli $\mathrm{rz}(A) = n$ to kolumny A są wektorami liniowo niezależnymi. Stąd istnieje macierz $X \in \mathbf{K}^{n,n}$ taka, że $A * X = I_n$. Podobnie, istnieje $Y \in \mathbf{K}^{n,n}$ taka, że $A^T * Y = I_n$, czyli $Y^T * A = I_n$. Ponadto

$$Y^T = Y^T * I_n = (Y^T * A) * X = I_n * X = X,$$

tzn. odwrotności lewostronna i prawostronna istnieją i są sobie równe, $A^{-1} = X = Y^T$. To kończy dowód.

Wiemy, że jeśli macierz kwadratowa A jest nieosobliwa to rozwiązaniem układu $A*\vec{x} = \vec{b}$ jest $\vec{x}^* := A^{-1}*\vec{b}$ i jest to jedyne rozwiązanie, tzn. układ jest oznaczony. Ale też odwrotnie, jeśli układ $A * \vec{x} = \vec{b}$ z macierzą kwadratową A jest oznaczony dla pewnego $\vec{b}$ to macierz A jest nieosobliwa. Rzeczywiście, wtedy jądro $\mathcal{N}(A) = \{\vec{0}\}$. To znaczy, że wektory-kolumny macierzy tworzą układ liniowo niezależny i $\mathrm{rz}(A) = n$. Na podstawie twierdzenia 2.1 wnioskujemy że A jest nieosobliwa.

Wniosek 7.1 *Niech A będzie macierzą kwadratową. Układ $A * \vec{x} = \vec{b}$ jest oznaczony wtedy i tylko wtedy gdy A jest nieosobliwa.*

7.2 Efektywna metoda rozwiązania

Dla danej macierzy $A \in \mathbf{K}^{m,n}$ i wektora $\vec{b} \in \mathbf{K}^m$ chcemy zbadać, czy układ (7.1) jest niesprzeczny, a jeśli tak to znaleźć zbiór jego rozwiązań (warstwę)

$$\mathcal{L}(A, \vec{b}) = \vec{x}_0 + \mathcal{N}(A),$$

gdzie $\mathcal{N}(A) = \mathrm{span}(\vec{z}_{s+1}, \vec{z}_{s+2}, \ldots, \vec{z}_n)$ i $s = \mathrm{rz}(A)$.

7.2.1 Ogólny schemat

Najpierw wyznaczymy $s = \mathrm{rz}(A)$ i $t = \mathrm{rz}([A, \vec{b}])$, a następnie w przypadku $s = t$ skonstruujemy rozwiązanie szczególne $\vec{x}_0$ oraz bazę $\vec{z}_{s+1}, \ldots, \vec{z}_n$ jądra $\mathcal{N}(A)$.

Ogólny schemat postępowania prowadzący do postaci równania, z którego można już stosunkowo łatwo odczytać rozwiązanie jest następujący. Startując z układu wyjściowego (7.1), który oznaczymy przez (U_0), kolejno dla $k = 1, 2, \ldots, s$ konstruujemy "prostsze" i (prawie) równoważne (U_0) układy (U_k) z macierzami $A^{(k)}$ tego samego formatu co A. Macierz $A^{(s)}$ układu docelowego (U_s) okaże się trójkątną górną.

Dopuszczamy przy tym następujące operacje na układnie równań:

(i) zamiana kolejności równań (wierszy macierzy),

(ii) zamiana kolejności niewiadomych (kolumn macierzy),

(iii) dodanie do jednego z równań innego równania pomnożonego przez skalar.

Zauważmy, że operacje (i) oraz (iii) prowadzą do układów równoważnych, zaś (ii) powoduje jedynie zmianę kolejności niewiadomych. W szczególności, rząd macierzy układu nie ulega zmianie.

Schemat sprowadzania układu wyjściowego do postaci z macierzą trójkątną górną przy pomocy zdefiniowanych operacji, który teraz dokładniej opiszemy, nazywamy ELIMINACJĄ GAUSSA.

7.2.2 Eliminacja Gaussa

Załóżmy, że wykonaliśmy już k przekształceń układu wyjściowego,

$$(U_0) \to (U_1) \to \ldots \to (U_k),$$

gdzie $0 \le k \le s - 1$, otrzymując układ

$$A^{(k)} * \vec{x}^{(k)} = \vec{b}^{(k)}$$

z macierzą

$$A^{(k)} = \begin{bmatrix} R_k & T_k \\ 0 & V_k \end{bmatrix},$$

gdzie $R_k \in \mathrm{TRIU}^{k,k}$ jest kwadratową i nieosobliwą macierzą trójkątną górną. Oczywiście, założenie to jest spełnione dla $k = 0$, czyli dla układu wyjściowego, bowiem wtedy $A^{(0)} = V_0 = A$.

Krok $(k + 1)$-szy eliminacji

1. Wybieramy w V_k element różny od zera, powiedzmy $v_{p,q} \neq 0$, $k + 1 \leq p \leq m$, $k + 1 \leq q \leq n$. (Taki element istnieje, bo inaczej mielibyśmy $\mathrm{rz}(A) = \mathrm{rz}(A^{(k)}) = k < s = \mathrm{rz}(A)$.)

2. Przestawiamy wiersze $(k + 1)$-szy z p-tym oraz kolumny (niewiadome) $(k + 1)$-szą z q-tą.

3. Dokonujemy eliminacji $(k + 1)$-szej niewiadomej z równań od $(k + 1)$-szego do m-tego odejmując od równania i-tego równanie $(k + 1)$-sze pomnożone przez

$$l_{i,k+1} := a^{(k)}_{i,k+1}/a^{(k)}_{k+1,k+1}, \quad \text{dla } i = k + 1, k + 2, \ldots, m.$$

   (Tutaj $a^{(k)}_{i,j}$ oznacza element aktualnie stojący na przecięciu i-tego wiersza i j-tej kolumny macierzy układu).

Po wykonaniu $(k + 1)$-szego kroku metody dostajemy układ z macierzą $A^{(k+1)}$, która ma wyzerowane współczynniki o indeksach (i, j) dla $1 \leq j \leq k + 1$, $j < i \leq m$.

Po wykonaniu $s = \mathrm{rz}(A)$ kroków dostajemy układ (U_s) postaci

$$A^{(s)} * \vec{x}^{(s)} = \vec{b}^{(s)},$$

przy czym

$$A^{(s)} = \begin{bmatrix} R_s & T_s \\ 0 & 0 \end{bmatrix},$$

a wektor $\vec{x}^{(s)}$ różni się od $\vec{x}^{(0)}$ jedynie przestawieniem (permutacją) współrzędnych. Rzeczywiście, gdyby odpowiednia podmacierz V_s macierzy $A^{(s)}$ była niezerowa to mielibyśmy $\mathrm{rz}(A) = \mathrm{rz}(A^{(s)}) > s$.

Dodajmy jeszcze, że w przypadkach $s = 0$ i $s = \min(m, n)$ niektóre bloki macierzy $A^{(s)}$ są puste.

7.2.3 Konstrukcja rozwiązania ogólnego

Przyjmując

$$\vec{x}^{(s)} = \begin{bmatrix} \vec{x}_I^{(s)} \\ \vec{x}_{II}^{(s)} \end{bmatrix}, \qquad \vec{b}^{(s)} = \begin{bmatrix} \vec{b}_I^{(s)} \\ \vec{b}_{II}^{(s)} \end{bmatrix},$$

gdzie $\vec{x}_I^{(s)}, \vec{b}_I^{(s)} \in \mathbf{K}^s$, $\vec{x}_{II}^{(s)} \in \mathbf{K}^{n-s}$, $\vec{b}_{II}^{(s)} \in \mathbf{K}^{m-s}$, układ (U_s) możemy zapisać jako

$$\left\{ \begin{array}{rcl} R_s * \vec{x}_I^{(s)} + T_s * \vec{x}_{II}^{(s)} & = & \vec{b}_I^{(s)} \\ \vec{0} & = & \vec{b}_{II}^{(s)} \end{array} \right. .$$

Jeśli teraz $\vec{b}_{II}^{(s)} \neq \vec{0}$ to układ jest sprzeczny i nie ma rozwiązań. Jeśli zaś $\vec{b}_{II}^{(s)} = \vec{0}$ to układ (U_s) redukuje się do

$$R_s * \vec{x}_I^{(s)} + T_s * \vec{x}_{II}^{(s)} = \vec{b}_I^{(s)}.$$

Przyjmując $\vec{x}_{II}^{(s)} = \vec{0}$ otrzymujemy rozwiązanie szczególne

$$\vec{x}_0^{(s)} = \begin{bmatrix} \vec{u}_0 \\ \vec{0} \end{bmatrix},$$

gdzie $\vec{u}_0 \in \mathbf{K}^s$ jest rozwiązaniem nieosobliwego układu

$$R_s * \vec{u}_0 = \vec{b}_I^{(s)}$$

z macierzą trójkątną dolną R_s.

Bazę jądra macierzy $A^{(s)}$,

$$\mathcal{N}(A^{(s)}) = \mathcal{N}([R_s, T_s]),$$

znajdujemy rozwiązując $(n-s)$ liniowo niezależnych rozwiązań układów jednorodnych

$$R_s * \vec{x}_I^{(s)} + T_s * \vec{x}_{II}^{(s)} = \vec{0}$$

kładąc kolejno za $\vec{x}_{II}^{(s)}$ wersory $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_{n-s} \in \mathbf{K}^{n-s}$. Oznaczając

$$T_s = [\vec{t}_{s+1}, \vec{t}_{s+2}, \ldots, \vec{t}_n]$$

otrzymujemy

$$\vec{z}_j^{(s)} = \begin{bmatrix} \vec{u}_j \\ \vec{e}_{j-s} \end{bmatrix}, \qquad \text{gdzie} \quad R_s * \vec{u}_j^{(s)} = -\vec{t}_j,$$

$s+1 \le j \le n$. Ostatecznie dostajemy

$$\mathcal{L}(A^{(s)},\vec{b}^{(s)}) = \vec{x}_0 + \mathrm{span}(\vec{z}_{s+1}^{(s)},\ldots,\vec{z}_n^{(s)}).$$

Są to również wszystkie rozwiązania układu wyjściowego, z dokładnością do odpowiedniej permutacji współrzędnych.

7.3 Interpretacja macierzowa eliminacji

Zobaczymy teraz jak proces eliminacji Gaussa wygląda z punktu widzenia rachunku macierzy.

7.3.1 Analiza operacji elementarnych

Załóżmy najpierw, dla uproszczenia, że nie musimy wykonywać przestawień ani wierszy ani kolumn układu (tzn. w każdym kroku element k-ty na głównej przekątnej jest niezerowy). Wtedy eliminacja w k-tym kroku odpowiada mnożeniu macierzy $A^{(k-1)}$ z lewej strony przez macierz kwadratową

$$L_k := I_m - \vec{l}_k * \vec{e}_k^{\,T} = \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & -l_{k+1,k} & \ddots & & \\ & & & \vdots & & \ddots & \\ & & & -l_{m,k} & & & 1 \end{bmatrix} \in \mathbf{K}^{m,m},$$

gdzie $\vec{l}_k := [0,\ldots,0,l_{k+1,k},\ldots,l_{m,k}]^T$ oraz

$$l_{i,k} := \frac{a_{i,k}^{(k-1)}}{a_{k,k}^{(k-1)}}, \qquad k+1 \le i \le m. \tag{7.2}$$

Dlatego

$$\begin{array}{rcl} A^{(1)} & = & L_1 * A, \\ A^{(2)} & = & L_2 * A^{(1)} \;=\; L_2 * L_1 * A, \\ & \cdots & \\ A^{(s)} & = & L_s * A^{(s-1)} \;=\; \cdots \;=\; L_s * L_{s-1} * \cdots * L_1 * A, \end{array}$$

a stąd $A = (L_1^{-1} * \cdots * L_s^{-1}) * A^{(s)}$.

Zauważmy teraz, że

$$L_i^{-1} = (I_m - \vec{l}_i * \vec{e}_i^T)^{-1} = (I_m + \vec{l}_i * \vec{e}_i^T).$$

Rzeczywiście, wobec tego, że $\vec{e}_i^T * \vec{l}_i = 0$ mamy bowiem

$$(I_m - \vec{l}_i * \vec{e}_i^T) * (I_m + \vec{l}_i * \vec{e}_i^T) = I_m - \vec{l}_i * (\vec{e}_i^T * \vec{l}_i) * \vec{e}_i^T = I_m.$$

Stąd

$$\begin{aligned} L_1^{-1} * \cdots * L_s^{-1} &= (I_m + \vec{l}_1 * \vec{e}_1^T) * \cdots * (I_m + \vec{l}_s * \vec{e}_s^T) \\ &= I_m + \vec{l}_1 * \vec{e}_1^T + \cdots + \vec{l}_s * \vec{e}_s^T. \end{aligned}$$

Oznaczając $\hat{L} := L_1^{-1} * \cdots * L_s^{-1}$ oraz $\hat{R} := A^{(s)}$ otrzymujemy ostatecznie rozkład (faktoryzację) macierzy,

$$A = \hat{L} * \hat{R},$$

gdzie

$$\hat{L} = \begin{bmatrix} L & 0 \\ H & I \end{bmatrix} \in \mathbf{K}^{m,m}, \qquad \hat{R} = \begin{bmatrix} R & T \\ 0 & 0 \end{bmatrix} \in \mathbf{K}^{m,n},$$

$L \in \mathrm{TRIL}^{s,s}$ jest kwadratową macierzą trójkątną dolną z jedynkami na głównej przekątnej oraz $R \in \mathrm{TRIU}^{s,s}$ jest macierzą trójkątną górną.

Dodajmy jeszcze, że współczynniki $l_{i,k}$ macierzy $\hat{L}$ są dla $1 \le k \le s$, $k < i \le m$, zdefiniowane przez (7.2).

Rozpatrzmy teraz ogólny przypadek, gdy dokonujemy przestawień wierszy i/lub kolumn macierzy. Przypomnijmy, że przestawienie wierszy i-tego z j-tym jest równoważne pomnożeniu macierzy przez elementarną macierz permutacji (transpozycję) $T_{i,j}$, natomiast pomnożenie macierzy z prawej strony przez $T_{i,j}$ jest równoważne przestawieniu kolumny i-tej z j-tą. Dlatego w obecności przestawień dostajemy

$$\begin{aligned} A^{(1)} &= L_1 * T_{1,p_1} * A * T_{1,q_1}, \\ A^{(2)} &= L_2 * T_{2,p_2} * A^{(1)} * T_{2,q_2} = L_2 * T_{2,p_2} * L_1 * T_{1,p_1} * A * T_{1,q_1} * T_{2,q_2} \\ &\cdots \\ A^{(s)} &= L_s * T_{s,p_s} * \cdots * T_{2,p_2} * L_1 * T_{1,p_1} * A * T_{1,q_1} * T_{2,q_2} * \cdots * T_{s,q_s}, \end{aligned}$$

przy czym zawsze $i \le p_i$, $j \le q_j$, $1 \le i, j \le s$.

Zauważmy, że

$$T_{2,p_2} * L_1 * T_{1,p_1} = (T_{2,p_2} * L_1 * T_{2,p_2}) * T_{2,p_2} * T_{1,p_1}.$$

Dalej

$$\begin{aligned} T_{2,p_2} * L_1 * T_{2,p_2} &= T_{2,p_2} * (I_m - \vec{l}_1 * \vec{e}_1^T) * T_{2,p_2} \\ &= I_m - (T_{2,p_2} * \vec{l}_1) * (\vec{e}_1^T * T_{2,p_2}) \\ &= I_m - \vec{l}_1' * \vec{e}_1^T \\ &=: L_1^{(1)}, \end{aligned}$$

gdzie $L_1^{(1)}$ różni się od L_1 jedynie przestawieniem wyrazów 2-go i p_2-go w pierwszej kolumnie. Uogólniając to rozumowanie dostajemy

$$\begin{aligned} & L_s * T_{s,p_s} * \cdots * T_{2,p_2} * L_1 * T_{1,p_1} \\ &= L_s * T_{s,p_s} * \cdots * L_2 * L_1^{(1)} * T_{2,p_2} * T_{1,p_1} \\ &= L_s * T_{s,p_s} * \cdots * (T_{3,p_3} * L_2 * T_{3,p_3}) * (T_{3,p_3} * L_1^{(1)} * T_{3,p_3}) \\ &\qquad\qquad * T_{3,p_3} * T_{2,p_2} * T_{1,p_1} \\ &= L_s * T_{s,p_s} * \cdots * L_3 * L_2^{(2)} * L_1^{(2)} * T_{3,p_3} * T_{2,p_2} * T_{1,p_1} \\ & \ldots \\ &= (L_s^{(s-1)} * L_{s-1}^{(s-1)} * \cdots * L_3^{(s-1)} * L_2^{(s-1)} * L_1^{(s-1)}) * (T_{s,p_s} * \cdots * T_{1,p_1}). \end{aligned}$$

Definiując macierze permutacji

$$Q^{-1} = Q^T := T_{1,q_1} * \cdots * T_{s,q_s}, \qquad P := T_{s,p_s} * \cdots * T_{1,p_1},$$

otrzymujemy ostatecznie

$$A^{(s)} = L_s^{(s-1)} * \cdots * L_1^{(s-1)} * P * A * Q^T = \hat{R},$$

czyli pożądany rozkład

$$P * A * Q^T = \hat{L} * \hat{R},$$

$\hat{L} = (L_1^{(s-1)})^{-1} * \cdots * (L_s^{(s-1)})^{-1}$, $\hat{R} = A^{(s)}$.

7.3.2 Rozkład trójkątno-trójkątny macierzy

Wynikiem naszej analizy jest następujące twierdzenie.

Twierdzenie 7.4 (O ROZKŁADZIE TRÓJKĄTNO-TRÓJKĄTNYM MACIERZY) *Dla dowolnej macierzy $A \in \mathbf{K}^{m,n}$ rzędu s istnieją (na ogół niejednoznaczne) macierze permutacji $P \in \mathbf{K}^{m,m}$ i $Q \in \mathbf{K}^{n,n}$ takie, że macierz $\hat{A} = P * A * Q^T$ ma jednoznaczny rozkład trójkątno-trójkątny*

$$\hat{A} = \hat{L} * \hat{R},$$

gdzie $\hat{L} \in \mathbf{K}^{m,m}$, $\hat{R} \in \mathbf{K}^{m,n}$, $\hat{l}_{i,i} = 1 \ \forall i$,

$$\hat{L} = \begin{bmatrix} L & 0 \\ H & I \end{bmatrix}, \qquad \hat{R} = \begin{bmatrix} R & T \\ 0 & 0 \end{bmatrix},$$

$L \in \mathrm{TRIL}^{s,s}$, $R \in \mathrm{TRIU}^{s,s}$.

Dowód. Istnienie rozkładu udowodniliśmy przez konstrukcję macierzy $\hat{L}$ i $\hat{R}$ (za pomocą eliminacji Gaussa). Aby pokazać jednoznaczność, przedstawimy $\hat{A}$ w postaci blokowej,

$$\hat{A} = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix}, \qquad A_{1,1} \in \mathbf{K}^{s,s}.$$

Wtedy dla danego rozkładu $\hat{A} = \hat{L} * \hat{R}$ mamy

$$A_{1,1} = L * R, \quad A_{1,2} = L * T, \quad A_{2,1} = H * R, \quad A_{2,2} = H * T.$$

Gdyby istniał inny rozkład z macierzami $\overline{L}$ i $\overline{R}$ to mielibyśmy $\overline{L} * \overline{R} = L * R$, czyli

$$L^{-1} * \overline{L} = R * \overline{R}^{-1}.$$

Po lewej stronie ostatniej równości mamy macierz trójkątną dolną z jedynkami na przekątnej, a z prawej trójkątną górną. To wymusza $L^{-1} * \overline{L} = I_s = R * \overline{R}^{-1}$, czyli $\overline{L} = L$ i $\overline{R} = R$. Jednoznaczność pozostałych bloków w rozkładzie wynika z równości $T = L^{-1} * A_{1,2}$ i $H = A_{2,1} * R^{-1}$.

7.4 Eliminacja bez przestawień

Podamy teraz warunek konieczny i dostateczny na to, aby istniał rozkład trójkątno-trójkątny oryginalnej macierzy A, a tym samym, aby eliminacja Gaussa była wykonalna bez przestawień wierszy i/lub kolumn.

Twierdzenie 7.5 *Aby macierz* $A = (a_{i,j}) \in \mathbf{K}^{m,n}$ *rzędu* s *miała rozkład trójkątno-trójkątny bez przestawień wierszy i/lub kolumn potrzeba i wystarcza, że wszystkie macierze kątowe* $A_k = (a_{i,j})_{i,j=1}^{k} \in \mathbf{K}^{k,k}$ *dla* $k = 1, 2, \ldots, s$ *są nieosobliwe.*

Dowód. Jeśli macierz ma rozkład $A = \hat{L} * \hat{R}$ to $A_k = \hat{L}_k * \hat{R}_k$ jest nieosobliwa dla $1 \leq k \leq s$. Dowód w drugą stronę podamy przez indukcję po s. Dla $s = 1$ twierdzenie jest oczywiste, bo wtedy $a_{1,1} \neq 0$ i eliminacja pierwszej kolumny jest wykonalna. Załóżmy, że twierdzenie jest prawdziwe dla $s - 1$. Oznaczając

$$A_s = \begin{bmatrix} A_{s-1} & \vec{a} \\ \vec{c}^T & a_{s,s} \end{bmatrix},$$

mamy z założenia indukcyjnego, że istnieje rozkład $A_{s-1} = L^{(s-1)} * R^{(s-1)}$ dla pewnych nieosobliwych macierzy

$$L^{(s-1)} \in \mathrm{TRIL}^{s-1,s-1} \quad \text{oraz} \quad R^{(s-1)} \in \mathrm{TRIU}^{s-1,s-1}.$$

Oznaczając

$$L^{(s)} = \begin{bmatrix} L^{(s-1)} & \vec{0} \\ \vec{g}^T & 1 \end{bmatrix}, \qquad R^{(s)} = \begin{bmatrix} R^{(s-1)} & \vec{b} \\ \vec{0}^T & r_{s,s} \end{bmatrix},$$

i rozwiązując równanie $A^{(s)} = L^{(s)} * R^{(s)}$ otrzymujemy

$$\begin{aligned} \vec{a} &= L^{(s-1)} * \vec{b}, \\ \vec{c}^T &= \vec{g}^T * R^{(s-1)} \qquad (\text{albo} \quad (R^{(s-1)})^T * \vec{g} = \vec{c}), \\ a_{s,s} &= \vec{g}^T * \vec{b} + r_{s,s}, \end{aligned}$$

skąd kolejno wyliczamy $\vec{b}$, $\vec{g}$ i $r_{s,s}$. Rozkład trójkątno-trójkątny macierzy kątowej $A^{(s)}$ w oczywisty sposób implikuje już rozkład całej macierzy A.

Na koniec podamy ważne klasy macierzy, dla których eliminacja Gaussa jest możliwa bez przestawień wierszy i/lub kolumn. Są to macierze:

(a) *diagonalnie dominujące wierszami*

$$\mathrm{WDD}^{n,n} = \left\{ A = (a_{i,j}) \in \mathbf{K}^{n,n} : \; |a_{i,i}| > \sum_{i \neq j=1}^{n} |a_{i,j}|, \; 1 \leq i \leq n \right\}.$$

(b) *diagonalnie dominujące kolumnami*

$$\mathrm{KDD}^{n,n} = \{A \in \mathbf{K}^{n,n} : \ A^T \in \mathrm{WDD}^{n,n}\}.$$

(c) *hermitowskie dodatnio określone*

$$\mathrm{HPD}^{n,n} = \{A \in \mathbf{K}^{n,n} : \ A^H = A \text{ oraz } \forall \vec{x} \neq \vec{0}\ \vec{x}^H * A * \vec{x} > 0\}.$$

(d) *hermitowskie ujemnie określone*

$$\mathrm{HND}^{n,n} = \{A \in \mathbf{K}^{n,n} : \ (-A) \in \mathrm{HPD}^{n,n}\}.$$

Aby pokazać, że w tych przypadkach przestawienie wierszy/kulumn nie jest konieczne, wykażemy, że spełnione są założenia twierdzenia 7.5.

W przypadku (a) zauważamy, że jeśli $A \in \mathrm{WDD}^{n,n}$ to A jest nieosobliwa, $\mathcal{N}(A) = \{\vec{0}\}$. Jeśli bowiem $A * \vec{x} = \vec{0}$ i $\vec{x} \neq \vec{0}$ to dla p takiego, że $|x_p| = \|\vec{x}\|_\infty$ mamy $a_{p,p}x_p + \sum_{j \neq p} a_{p,j}x_j = 0$, a stąd

$$|a_{p,p}| \leq \sum_{j \neq p} |a_{p,j}| \left| \frac{x_j}{x_p} \right| \leq \sum_{j \neq p} |a_{p,j}|,$$

co przeczy dominacji głównej diagonali macierzy. Uzasadnienie uzupełnia obserwacja, że jeśli $A \in \mathrm{WDD}^{n,n}$ to również macierze kątowe $A_k \in \mathrm{WDD}^{k,k}$, $1 \leq k \leq n$.

Dla przypadku (b) wystarczy zauważyć, że jesłi $A \in \mathrm{KDD}^{n,n}$ to $A^T \in \mathrm{WDD}^{n,n}$, oraz wykorzystać fakt, że nieosobliwość A jest równoważna nieosobliwości A^T.

W przypadku (c) (i zupełnie podobnie w (d)) zauważamy, że każda macierz $A \in \mathrm{HPD}^{n,n}$ jest nieosobliwa. Jeśli bowiem $\vec{x} \neq \vec{0}$ i $A * \vec{x} = \vec{0}$ to $\vec{x}^H * A * \vec{x} = 0$. Ponadto, wszystkie macierze kątowe $A_k \in \mathrm{HPD}^{k,k}$, bo dla dowolnego niezerowego $\vec{y} \in \mathbf{K}^k$ mamy

$$\vec{y}^H * A_k * \vec{y} = \begin{bmatrix} \vec{y} \\ \vec{0} \end{bmatrix}^H * A * \begin{bmatrix} \vec{y} \\ \vec{0} \end{bmatrix} > 0.$$