CENTRUM OBLICZENIOWE
POLSKIEJ AKADEMII NAUK

MIECZYSŁAW WARMUS

UOGÓLNIONE ODWROTNOŚCI MACIERZY

WARSZAWA 1972
PAŃSTWOWE WYDAWNICTWO NAUKOWE

CENTRUM OBLICZENIOWE
POLSKIEJ AKADEMII NAUK

MIECZYSŁAW WARMUS

UOGÓLNIONE ODWROTNOŚCI MACIERZY

WARSZAWA 1972
PAŃSTWOWE WYDAWNICTWO NAUKOWE

W ostatnich latach w niektórych działach matematyki, zwłaszcza w algebrze liniowej i statystyce matematycznej i ich zastosowaniach, coraz częściej używa się pojęcia uogólnionych odwrotności macierzy, a więc odwrotności istniejących dla wszystkich macierzy prostokątnych. W literaturze matematycznej brak było jednolitej teorii tych odwrotności. Praca niniejsza jest próbą wypełnienia tej luki. Podano w niej zasadnicze znane fakty. Uzupełniono je wieloma nowymi twierdzeniami. Niektóre dowody twierdzeń znanych uproszczono.

Oprócz rozważań teoretycznych praca zawiera efektywny algorytm do obliczania uogólnionych odwrotności macierzy, przykłady zastosowań, jak również konkretne przykłady numeryczne. Wykład jest zrozumiały dla wszystkich czytelników znających podstawy rachunku macierzowego.

REDAKTOR WYDAWNICZY
CENTRUM OBLICZENIOWEGO PAN
Jan Lipski

Printed in Poland

Państwowe Wydawnictwo Naukowe
Oddział w Łodzi 1972

Wydanie I. Nakład 910+90 egz. Ark. wyd. 3.25. Ark. druk. 4.25.
Papier pism. kl. III, 71 g, 70×100. Podpisano do druku 5 XII 1972 r.
Druk ukończono w grudniu 1972 r. Zam 306. D-10. Cena zł 12.—

Zakład Graficzny Wydawnictw Naukowych
Łódź, ul. Gdańska 162

WSTĘP

W pracy niniejszej rozważamy macierze prostokątne o wyrazach zespolonych. Symbol A^* oznacza sprzężoną transpozycję macierzy A. Można jednak również traktować wszystkie macierze jako macierze o wyrazach rzeczywistych. Wtedy symbol A^* oznacza zwykłą transpozycję macierzy A.

Niech A i T będą dowolnymi macierzami prostokątnymi o tej samej liczbie wierszy. Równanie macierzowe

$$A\Xi = T \tag{1}$$

w przypadku, gdy A jest macierzą kwadratową nieosobliwą, ma zawsze rozwiązanie $\Xi = A^{-1}T$, gdzie A^{-1} oznacza odwrotność macierzy A. Rozwiązanie to jest jedyne. W szczególności T może być wektorem kolumnowym.

C. R. Rao [1] w naturalny sposób uogólnił pojęcie odwrotności $\mathcal{X}$ dla dowolnej macierzy prostokątnej A żądając, aby dla każdej macierzy T, dla której równanie (1) nie jest sprzeczne, macierz $\Xi = \mathcal{X}T$ była rozwiązaniem (być może nie jedynym). Okazuje się, że tak zdefiniowana odwrotność $\mathcal{X}$ istnieje dla każdej macierzy prostokątnej A. Będziemy ją nazywać odwrotnością uogólnioną albo krócej g-odwrotnością.

Narzucając na g-odwrotność dodatkowe warunki, jak to zrobili E. H. Moore [2] i R. Penrose [3], określa się odwrotność dowolnej macierzy prostokątnej jednoznacznie. Będziemy ją nazywać odwrotnością Moore`a-Penrosego albo krócej MP-odwrotnością. Zatem MP-odwrotność jest szczególnym przypadkiem g-odwrotności.

G. Peters i J. H. Wilkinson [6] pokazali, że MP-odwrotność dowolnej macierzy prostokątnej A można określić w naturalny sposób, jako taką macierz, która między macierzami $\mathcal{X}$ minimizującymi normę euklidesową $\|b - A\mathcal{X}b\|$, gdzie b jest wektorem kolumnowym, ma najmniejszą normę.

W pracy niniejszej podano podstawowe wiadomości o g-odwrotnościach. Niektóre z nich są znane, inne stanowią uzupełnienie dokonane przez autora. Po raz pierwszy podano efektywny wzór na wszystkie możliwe g-odwrotności danej macierzy, algorytm do ich obliczania, następnie efektywne wzory i algorytmy dla g-odwrotności spełniających poszczególne warunki dodatkowe. Niektóre z twierdzeń znanych otrzymały nowe prostsze dowody oparte o nowo wyprowadzone własności. Udowodniono, że klasa macierzy $\mathcal{X}$ minimizujących normę $\| b - A\mathcal{X}b \|$ pokrywa się z klasą g-odwrotności L-symetrycznych.

Jak dotychczas brak było wykończonej teorii g-odwrotności. Praca niniejsza jest próbą wypełnienia tej luki.

1. OZNACZENIA

Macierze będziemy oznaczać dużymi literami łacińskimi lub greckimi. Małe litery będą oznaczać liczby. Jeśli będzie potrzebna znajomość wymiarów macierzy, będą one podawane jako para wskaźników w nawiasach kwadratowych przy symbolu macierzy. Jeśli, na przykład, macierz M ma wymiary $p \times q$, to będziemy ją oznaczać również symbolem $M_{[p,q]}$. Odwrotnie, macierz wprowadzoną symbolem $M_{[p,q]}$ będzie można - bez każdorazowego objaśniania - oznaczać po prostu symbolem M.

Wymiar macierzy kwadratowych będzie można podawać tylko jednym wskaźnikiem, określającym stopień macierzy. Na przykład symbol $N_{[k]}$ będzie oznaczać macierz kwadratową k-go stopnia. W szczególności symbol $I_{[k]}$ będzie zawsze oznaczać macierz jednostkową k-go stopnia.

Macierze zerowe będziemy oznaczać symbolem 0 z ewentualnym podaniem wymiarów, na przykład $0_{[p,q]}$.

Wyrazy macierzy będziemy oznaczać tymi samymi literami, co samą macierz, ale małymi, zawsze z podaniem wskaźników wiersza i kolumny. A więc, na przykład, symbol m_{kl} będzie oznaczać wyraz macierzy M.

Symbolem $\varrho(M)$ będziemy oznaczać rząd macierzy M.

W celu uproszczenia wzorów będziemy dopuszczać symbole $M_{[0,q]}$, $M_{[p,0]}$ czy $M_{[0,0]}$, czyli symbole "macierzy" o liczbie wierszy lub kolumn równej zeru, przyjmując, że

$$A_{[m,0]}B_{[0,n]} = 0_{[m,n]}$$

$$A_{[m,0]} = B_{[m,0]} = 0_{[m,0]}$$

$$A_{[0,n]} = B_{[0,n]} = 0_{[0,n]}$$

$$A_{[0,0]} = B_{[0]} = 0_{[0]} = I_{[0]}$$

Zgodnie z powyższą umową macierz $B_{[0]}$ ma odwrotność, jest bowiem sama dla siebie odwrotnością

$$B_{[0]}B_{[0]} = I_{[0]}$$

Rząd macierzy o liczbie wierszy lub kolumn równej zeru jest zawsze 0.

Macierz sprzężoną do macierzy M będziemy oznaczać symbolem $\bar{M}$ a macierz transponowaną symbolem M'. Wobec tego jest zawsze

$$M^* = (\bar{M})' = \overline{(M')}$$

a dla macierzy rzeczywistych

$$M^* = M', \quad \bar{M} = M.$$

Znaczek ■ będzie oznaczać zakończenie dowodu.

2. NORMA MACIERZY

W niniejszej pracy <u>normą macierzy</u> $A_{[m,n]}$ będziemy nazywać normę euklidesową tej macierzy, tzn. liczbę rzeczywistą nieujemną $\|A\|$ taką, że

$$\|A\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 = \mathrm{Tr}(AA^*)$$

Z powyższej definicji wynika, że

$$\|A + B\|^2 = \mathrm{Tr}(A + B)(A^* + B^*) =$$

$$= \|A\|^2 + \|B\|^2 + \mathrm{Tr}(AB^*) + \mathrm{Tr}(BA^*) =$$

$$= \|A\|^2 + \|B\|^2 + 2\,\mathrm{Tr}\,\mathrm{Re}(AB^*)$$

Ponieważ z definicji kwadrat normy macierzy jest sumą kwadratów wyrazów tej macierzy, więc

$$\|A\| = 0 \Leftrightarrow A = 0$$

Wynika stąd m. in., że

$$AA^* = 0 \Leftrightarrow A = 0$$

Z powyższych własności normy macierzy będziemy korzystać w dalszym ciągu pracy.

3. TWIERDZENIA O DIAGONALIZACJI

W pracy niniejszej w sposób istotny będą wykorzystywane niektóre znane twierdzenia o diagonalizacji macierzy. Ze względu na to, że ich dowody mają charakter konstrukcyjny i stwarzają ideę przewodnią dla odpowiednich algorytmów, podamy potrzebne twierdzenia wraz z pełnymi dowodami.

Twierdzenie 1

Dla każdej macierzy $A_{[m,n]}$ istnieją takie nieosobliwe macierze $B_{[m]}$ i $C_{[n]}$, że

$$BAC = J \tag{2}$$

gdzie

$$J = \left[\begin{array}{c|c} I_{[r]} & 0_{[r,n-r]} \\ \hline 0_{[m-r,r]} & 0_{[m-r,n-r]} \end{array}\right], \quad r = \varrho(A) \tag{3}$$

Dowód.

Niech

$$p = \min(m,n) \tag{4}$$

Wobec tego rząd r macierzy A spełnia nierówność

$$0 \leqslant r \leqslant p \tag{5}$$

Gdy macierz A jest zerowa, tzn. $r = 0$, można przyjąć, że $B = I_{[m]}$, $C = I_{[n]}$, $J = 0_{[m,n]}$ i wtedy zachodzi (2). Wobec tego wystarczy ograniczyć się do przypadku, gdy

$$1 \leqslant r \leqslant p \tag{6}$$

Niech

$$A^{(k)} = \left[\begin{array}{c|cccc} I_{[k]} & & & 0_{[k,n-k]} & \\ \hline & a_{k+1,k+1}^{(k)} & \cdots & & a_{k+1,n}^{(k)} \\ 0_{[m-k,k]} & \vdots & & \vdots & \vdots \\ & a_{m,k+1}^{(k)} & \cdots & & a_{m,n}^{(k)} \end{array}\right]$$

(7)
$$k = 0,1,\ldots,r$$

$$A^{(0)} = A$$

Załóżmy najpierw, że istnieje wyraz $a_{u,v}^{(k)} \neq 0$, gdzie $k<u\leq m$, $k<v\leq n$. Wprowadzamy macierz kwadratową $B_1^{(k)}$ m-go stopnia według wzoru

$$B_1^{(k)} = I_{[m]}, \quad \text{gdy} \quad u = k+1$$

(8)

$$B_1^{(k)} = \begin{array}{cl} \begin{array}{ccccc} & k+1 & & u & \end{array} & \\ \begin{bmatrix} \begin{smallmatrix}1&&\\&\ddots&\\&&1\end{smallmatrix} & & & & \\ & 0 & & 1 & \\ & & \begin{smallmatrix}1&&\\&\ddots&\\&&1\end{smallmatrix} & & \\ & 1 & & 0 & \\ & & & & \begin{smallmatrix}1&&\\&\ddots&\\&&1\end{smallmatrix} \end{bmatrix} & \begin{array}{l} \\ k+1 \\ \\ u \\ \\ \end{array} \end{array}$$

$$\text{gdy} \quad u \neq k+1$$

oraz macierz kwadratową $C_1^{(k)}$ n-go stopnia według wzoru

$$C_1^{(k)} = \begin{array}{c} \begin{array}{ccccc} & & k+1 & & \end{array} \\ \left[\begin{array}{ccccc} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \frac{1}{a_{u,v}^{(k)}} & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 1 \end{array}\right] \end{array} \begin{array}{c} \\ \\ \\ k+1, \\ \\ \\ \end{array} \quad \text{gdy} \quad v = k+1$$

(9)

$$C_1^{(k)} = \begin{array}{c} \begin{array}{ccccccc} & & k+1 & & & v & \end{array} \\ \left[\begin{array}{ccccccccc} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & \frac{1}{a_{u,v}^{(k)}} & & & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{array}\right] \end{array} \begin{array}{c} \\ \\ \\ k+1 \\ \\ \\ \\ v \\ \\ \\ \end{array} , \quad \text{gdy} \quad v \neq k+1$$

Możemy teraz utworzyć macierz postaci

$$(10) \qquad B_1^{(k)} A^{(k)} C_1^{(k)} = \left[\begin{array}{c|cccc} I_{[k]} & & 0_{[k,n-k]} & & \\ \hline & 1 & \beta_{k+1,k+2}^{(k)} & \dots & \beta_{k+1,n}^{(k)} \\ & \beta_{k+2,k+1}^{(k)} & \beta_{k+2,k+2}^{(k)} & \dots & \beta_{k+2,n}^{(k)} \\ 0_{[m-k,k]} & \vdots & \vdots & & \vdots \\ & \beta_{m,k+1}^{(k)} & \beta_{m,k+2}^{(k)} & \dots & \beta_{m,n}^{(k)} \end{array}\right]$$

Wprowadzamy teraz macierz kwadratową $B_2^{(k)}$ m-go stopnia

$$B_2^{(k)} = I_{[m]}, \qquad \text{gdy} \quad m = k + 1$$

$$(11) \qquad B_2^{(k)} = \left[\begin{array}{cccccc} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -\beta_{k+2,k+1}^{(k)} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\beta_{m,k+1}^{(k)} & & & 1 \end{array}\right] \quad \text{gdy} \quad m \neq k + 1$$

oraz macierz kwadratową $C_2^{(k)}$ n-go stopnia według wzoru

$$C_2^{(k)} = I_{[n]}, \qquad \text{gdy} \quad n = k+1$$

(12)
$$C_2^{(k)} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & -\beta_{k+1,k+2}^{(k)} & \cdots & -\beta_{k+1,n}^{(k)} \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}, \quad \text{gdy} \quad n \neq k+1$$

Mamy wtedy możliwość określenia macierzy

(13)
$$A^{(k+1)} = B_2^{(k)} B_1^{(k)} A^{(k)} C_1^{(k)} C_2^{(k)}$$

ponieważ pomnożenie lewostronne macierzy (10) przez macierz (11) powoduje wyzerowanie wyrazów pod główną przekątną w $(k+1)$-ej kolumnie, a pomnożenie prawostronne macierzy $B_2^{(k)} B_1^{(k)} A^{(k)} C_1^{(k)}$ przez macierz (12) powoduje wyzerowanie wyrazów nad główną przekątną w $(k+1)$-ym wierszu, przy niezmienionych pierwszych $k+1$ kolumnach.

Zauważmy, że wszystkie macierze $B_1^{(k)}$, $C_1^{(k)}$, $B_2^{(k)}$, $C_2^{(k)}$ są nieosobliwe, wobec czego na mocy (13)

(14)
$$\varrho\left(A^{(k+1)}\right) = \varrho\left(A^{(k)}\right)$$

Gdy nie istnieje w macierzy (7) wyraz $a_{u,v}^{(k)} \neq 0$, gdzie $k < u \leq m$, $k < v \leq n$, macierz (7) ma postać

(15)
$$A^{(k)} = \left[\begin{array}{c|c} I_{[k]} & 0_{[k,n-k]} \\ \hline 0_{[m-k,k]} & 0_{[m-k,n-k]} \end{array}\right]$$

i wobec tego

(16)
$$\varrho\left(A^{(k)}\right) = k$$

Niech $A^{(s)}$ będzie pierwszą z kolei macierzą w ciągu $A^{(0)}, \ldots, A^{(r)}$ mającą postać (15). Wobec (14) jest

$$\varrho\left(A^{(s)}\right) = \varrho(A) = r$$

i na mocy (16) $s = r$. Wtedy na mocy (3)

$$(17) \qquad A^{(r)} = J$$

Kładąc

$$(18) \qquad B = B_2^{(r-1)} B_1^{(r-1)} \dots B_2^{(0)} B_1^{(0)}$$

otrzymujemy macierz B kwadratową nieosobliwą m-go stopnia, a kładąc

$$(19) \qquad C = C_1^{(0)} C_2^{(0)} \dots C_1^{(r-1)} C_2^{(r-1)}$$

macierz kwadratową nieosobliwą n-go stopnia, takie że

$$A^{(r)} = BA^{(0)}C$$

czyli na mocy (17) otrzymujemy (2). ■

Dla danej macierzy A wzór (2) na ogół nie określa macierzy B i C jednoznacznie. Z punktu widzenia dalszych rozważań jest to jednak nieistotne. Będziemy zakładać, że B i C są ustalonymi macierzami spełniającymi warunek (2).

W dalszym ciągu będą potrzebne różne wzory związane z macierzami B i C. Wobec tego zajmiemy się teraz wyprowadzeniem tych wzorów. Niech

$$(20) \qquad B = \left[\begin{array}{c} F_{[r,m]} \\ \hdashline G_{[m-r,m]} \end{array}\right], \quad B^{-1} = \left[\begin{array}{c:c} \Gamma_{[m,r]} & \Delta_{[m,m-r]} \end{array}\right],$$

$$C = \left[\begin{array}{c:c} D_{[n,r]} & E_{[n,n-r]} \end{array}\right], \quad C^{-1} = \left[\begin{array}{c} \Phi_{[r,n]} \\ \hdashline \Psi_{[n-r,n]} \end{array}\right]$$

gdzie - jak poprzednio - r oznacza rząd macierzy A.

Ze wzorów $B^{-1}B = I_{[m]}$ oraz $CC^{-1} = I_{[n]}$ wynika, że

$$(21) \qquad \Gamma F + \Delta G = I_{[m]}, \qquad D\Phi + E\Psi = I_{[n]}.$$

Ze wzoru $BB^{-1} = I_{[m]}$ wynika, że

(22) $F\Gamma = I_{[r]}$, $F\Delta = 0_{[r,m-r]}$, $G\Gamma = 0_{[m-r,r]}$, $G\Delta = I_{[m-r]}$

a ze wzoru $C^{-1}C = I_{[n]}$

(23) $\Phi D = I_{[r]}$, $\Phi E = 0_{[r,n-r]}$, $\Psi D = 0_{[n-r,r]}$, $\Psi E = I_{[n-r]}$

Wzór (2) jest równoważny wzorowi

(24) $$BA = JC^{-1}$$

Ale

$$BA = \left[\begin{array}{c} FA \\ \hdashline GA \end{array}\right], \qquad JC^{-1} = \left[\begin{array}{c} \Phi \\ \hdashline 0_{[m-r,n]} \end{array}\right]$$

Otrzymujemy stąd wzory

(25) $$FA = \Phi, \qquad GA = 0_{[m-r,n]}$$

Wzór (2) jest także równoważny wzorowi

(26) $$AC = B^{-1}J$$

Ponieważ

$$AC = \left[\begin{array}{c:c} AD & AE \end{array}\right], \qquad B^{-1}J = \left[\begin{array}{c:c} \Gamma & 0_{[m,n-r]} \end{array}\right]$$

zatem

(27) $$AD = \Gamma, \qquad AE = 0_{[m,n-r]}$$

Wreszcie wzór (2) jest równoważny wzorowi

(28) $$A = B^{-1}JC^{-1}$$

skąd - biorąc pod uwagę, że $(B^{-1}J)C^{-1} = \Gamma\Phi$ - otrzymujemy

(29) $$\Gamma\Phi = A$$

a ze wzorów (29), (27) i (25)

$$(30) \qquad ADFA = A$$

Z faktu, że macierze B i C są nieosobliwe, wynika, że dla każdej macierzy A obie macierze Γ i Φ są rzędu r. Wobec tego dla każdej macierzy A macierze $\Gamma^*\Gamma$ oraz $\Phi\Phi^*$ są obie macierzami kwadratowymi nieosobliwymi stopnia i rzędu r i mają macierze odwrotne $(\Gamma^*\Gamma)^{-1}$ i $(\Phi\Phi^*)^{-1}$. Również na mocy (29) dla każdej macierzy A macierz

$$\Gamma^* A \Phi^* = (\Gamma^*\Gamma)(\Phi\Phi^*)$$

jest macierzą kwadratową nieosobliwą stopnia i rzędu r i ma macierz odwrotną

$$(\Gamma^* A \Phi^*)^{-1} = (\Phi\Phi^*)^{-1}(\Gamma^*\Gamma)^{-1}$$

Skorzystamy z tego w dalszych rozważaniach.

Obecnie z łatwością udowodnimy drugie twierdzenie o diagonalizacji.

Twierdzenie 2

Dla każdej macierzy $A_{[m,n]}$ rzędu r istnieją takie macierze $F_{[r,m]}$ i $D_{[n,r]}$ rzędu r, że

$$(31) \qquad FAD = I_{[r]}$$

Dowód

Wzór (31) wynika bezpośrednio z pierwszego wzoru (25) i pierwszego wzoru (23), ponieważ z faktu, że macierze B i C są nieosobliwe, wynika, że macierze F i D są obie rzędu r. ■

4. OKREŚLENIE g-ODWROTNOŚCI I ICH PODSTAWOWE WŁASNOŚCI

Definicja

Odwrotnością uogólnioną albo g-odwrotnością macierzy $A_{[m,n]}$ nazywamy każdą taką macierz $\mathcal{X}$, że dla każdej macierzy Y, dla której równanie