数理解析研究所講究録
941 巻 1996 年 156-169

20.

一般逆行列の計算アルゴリズムとその証明

齊藤敏明 (工学院大学・工)
牧野潔夫 (工学院大学・工)

20.1 はじめに

一般逆行列の計算アルゴリズムはかなり知られていて、その方法は各種の論文・単行本に書かれている。特に野田・泉田・越智 [14] は各種のアルゴリズム、その計算量および実際の処理系への implement と内容が豊富である。しかしその証明まで示された論文・単行本はわずかで、Krishunamurty の本 [6] にも証明されていない方法 ([5],[14],[6] 等) が各種ある。ここではいくつかの一般逆行列の計算アルゴリズムの証明をできるだけ初等的に示すことを目的とする。

予備知識はできるだけ少なくし、東京大学出版会から発行されている教科書"線形代数入門"[1] とその演習書"線形代数演習"[2] に書かれているもの以外は仮定しない。

数学の本質の半分は証明にあるといってもよく、また証明の中から新しいアルゴリズムが見つかることもしばしばある。したがって一般逆行列の各種計算法の証明をまとめておくことも少しは意義のあることのように思われる。

20.2 計算アルゴリズムとその証明

はじめに以下で用いるものを列記する。(m,n) 型行列 A に対して、${}^tA$ は A の転置行列、$\mathrm{Tr}(A)$ は A の Trace、$\mathrm{rank}(A)$ は A の rank をあらわす。また、(m,n) 型行列

$$D = \begin{pmatrix} d_1 & & & O \\ & \ddots & & \\ & & d_r & \\ O & & & O \end{pmatrix}$$

を $\mathrm{diag}(d_1, \ldots, d_r, 0, \ldots, 0; m, n)$ とあらわす。U, Vは直交行列、E(または E_n) は (n 次) 単位行列、Oは零行列をあらわす。以下実行列のみを扱うが、転置行列を随伴行列、直交行列をユニタリ行列におきかえれば複素行列でも以下の証明は成立する。

定義 1 (Moore-Penrose 逆行列,[2],p.68, 問 31) A を任意の (m,n) 型行列としたとき、次の 4 条件をみたす (n,m) 型行列 A^+が唯一存在する。

$$AA^+A = A \tag{1}$$

$$A^+AA^+ = A^+ \tag{2}$$

$$^t(AA^+) = AA^+ \tag{3}$$

$$^t(A^+A) = A^+A \tag{4}$$

この A^+を Moore-Penrose 逆行列という。

20.2.1 階数分解による方法

定理 1 (階数分解,[2],p.68, 問 31) (m,n) 型行列 A が $\mathrm{rank}(A) = r$のとき、rank が rの (m,r) 型行列 B および rank が rの (r,n) 型行列 Cを用いて $A = BC$とあらわせる。このとき、A の Moore-Penrose 逆行列 A^+は $A^+ = {}^tC({}^tBA{}^tC)^{-1}{}^tB = {}^tC(C{}^tC)^{-1}({}^tBB)^{-1}{}^tB$ とあらわせる。

■定理 1 を改良した方法

まず A に左から基本変形行列 Pをかけて上三角行列にする。これにさらに左から P^{-1}をかけると、$A = P^{-1}(PA) = (P_1, P_2)(PA)$ (P^{-1}の r列目までを P_1, $r+1$ 列目以降を P_2とする) となる。ここで、(PA) の $r+1$ 行目以降はすべて 0 なので、上式においてこの部分との積に対応する P_2は O行列におきかえることができる。よって、Bを (m,r) 型行列 P_1、Cを (PA) の r行までの (r,n) 型行列とすれば、$A = BC$と書ける。

そのあと定理 1 の A^+の式に代入する。

20.2.2 特異値分解による方法

定理 2 (特異値分解,[2],p.187, 問 11) 任意の (m,n) 型行列 $A(\mathrm{rank}(A) = r)$ に対し、${}^tAA$ の 0 でない固有値の正の平方根を$\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}$ とすると、(m,n) 型行列

$$D = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}, 0, \ldots, 0; m, n) \tag{5}$$

と適当な m 次直交行列 U および n 次直交行列 V が存在して、$A = UD^tV$ とかける。これを A の特異値分解という。

定理 3 A の特異値分解を $A = UD^tV$ とすると、A の Moore-Penrose 逆行列 A^+ は次のようにあらわせる。

$$\begin{aligned} A^+ &= VD^{+t}U \\ D^+ &= \mathrm{diag}(\sqrt{\lambda_1}^{-1}, \ldots, \sqrt{\lambda_r}^{-1}, 0, \ldots, 0; n, m) \end{aligned} \tag{6}$$

■証明

Moore-Penrose 逆行列の 4 条件 (1)~(4) を示すのはやさしい。（証明終）

20.2.3 Decell-Leverrier(Penrose) の方法

定理 4 (Newton の公式,[3],p.292) n 次正方行列 A について、その特性多項式 $\varphi_A(t)$ は次のように表せる。

$$\varphi_A(t) = |A - tE_n| = (-1)^n[t^n - p_1t^{n-1} - p_2t^{n-2} - \cdots - p_n]$$

このとき、$s_k = \sum_{l=1}^{n} \lambda_l^k$ とすると次の Newton の公式が成立する。

$$kp_k = s_k - p_1s_{k-1} - \cdots - p_{k-1}s_1 \ (k = 1, 2, \cdots, n) \tag{7}$$

■証明

数学的帰納法により示す。

1) $n = 1$ のとき、$\varphi_A(t) = a - t,\ p_1 = a,\ 1{\cdot}p_1 = a,\ s_1 = a^1$ であるから成立する。

2) n まで正しいとして、$n+1$ のときを示す。λ_{n+1} を λ とおく。$n+1$ のときの s_k, p_k を s'_k, p'_k とおくと、$s'_k = s_k + \lambda^k,\ p'_k = p_k - \lambda p_{k-1}\ (2 \leq k \leq n)$ となる。

また $p'_{n+1} = \lambda p_n,\ p'_1 = p_1 + \lambda$ となる。なぜなら、$s'_k = s_k + \lambda^k$ はすぐわかる。また

$$\begin{aligned} p'_k &= (-1)^k \sum \lambda_{i_1} \cdots \lambda_{i_k} \\ &= (-1)^k \sum_{\lambda_i \neq \lambda} \lambda_{i_1} \cdots \lambda_{i_k} - (-1)^{k-1} \lambda \sum_{\lambda_i \neq \lambda} \lambda_{i_1} \cdots \lambda_{i_{k-1}} \\ &= p_k - \lambda p_{k-1} \end{aligned}$$

である。

(a) $k \leq n$ のとき

$s'_k - p'_1 s'_{k-1} - \cdots - p'_{k-1} s'_1$

$$
\begin{aligned}
&= (s_k+\lambda^k)-(p_1+\lambda)(s_{k-1}+\lambda^{k-1})-(p_2-\lambda p_1)(s_{k-2}+\lambda^{k-2})-\cdots \\
&\quad -(p_l-\lambda p_{l-1})(s_{k-l}+\lambda^{k-l})-(p_{l+1}-\lambda p_l)(s_{k-l-1}+\lambda^{k-l-1}) \\
&\quad -\cdots-(p_{k-1}-\lambda p_{k-2})(s_1+\lambda) \\
&= s_k+\lambda^k-(p_1 s_{k-1}+p_2 s_{k-2}+\cdots+p_{k-1}s_1) \\
&\quad -\lambda^k-\lambda^{k-1}p_1+\lambda^{k-1}p_1-\lambda^{k-2}p_2+\lambda^{k-2}p_2-\cdots-\lambda p_{k-1}+\lambda^2 p_{k-2} \\
&\quad -\lambda(s_{k-1}-s_{k-2}p_1-\cdots-s_{k-l}p_{l-1}-\cdots-p_{k-2}s_1) \\
&= s_k-p_1 s_{k-1}-p_2 s_{k-2}-\cdots-p_{k-1}s_1-\lambda(s_{k-1}-\cdots-p_{k-2}s_1)-\lambda p_{k-1} \\
&= kp_k-\lambda(k-1)p_{k-1}-\lambda p_{k-1}=kp_k-\lambda k p_{k-1}=kp'_k
\end{aligned}
$$

となり $s'_k-p'_1 s'_{k-1}-\cdots-p'_{k-1}s'_1=kp'_k$が成立する。

(b) $k=n+1$ のとき

$\varphi_A(t)=(-1)^n[t^{n+1}-p'_1 t^n-\cdots-p'_n t-p'_{n+1}]$ より、

$0=\varphi_A(\lambda_i)=(-1)^n[\lambda_i^{n+1}-p'_i\lambda_i^n-\cdots-p'_n\lambda-p'_{n+1}]$ $(i=1,\ldots,n+1)$

ここで、$i=1,\ldots,n+1$ で和をとると $0=(-1)^n[s'_{n+1}-p'_1 s'_n-\cdots-(n+1)p'_{n+1}]$ となる。ゆえに $(n+1)p'_{n+1}=s'_{n+1}-p'_1 s'_n-\cdots-p'_n s'_1$ となる。

以上により与式が成立する。(証明終)

定理 5 (Decell-Leverrier(Penrose) の方法,[5],p.63,5.3.1,[9])(m,n) 型行列 A について $M={}^tAA$ とおき以下のプロセスを行う。

$$
\begin{array}{llllllllll}
M_1 &=& M, & q_1 &=& \mathrm{Tr}(M_1)/1, & B_1 &=& M_1-q_1E_n \\
M_2 &=& MB_1, & q_2 &=& \mathrm{Tr}(M_2)/2, & B_2 &=& M_2-q_2E_n \\
\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots&\cdots \\
M_r &=& MB_{r-1}, & q_r &=& \mathrm{Tr}(M_r)/r, & B_r &=& M_r-q_rE_n
\end{array}
\tag{8}
$$

このとき $M_1\neq 0, M_2\neq 0,\ldots,M_{r-1}\neq 0$ かつ $MB_r=O$ ならば、次の3点がわかる。

$$
\mathrm{Tr}(MB_r)\neq 0,\ \mathrm{rank}(M)=\mathrm{rank}(A)=r,\ A^+=\frac{B_{r-1}{}^tA}{q_r} \tag{9}
$$

■証明

$M_l, B_l, MB_l(l=1,\ldots,r)$ が対称であること、および $MB_l=B_lM$であることを数学的帰納法により示す。

1) $l=1$ のとき

$M_1={}^tAA$ より対称である。$B_1=A_1-q_1E_n={}^tXX-q_1E_n$だから対称である。さらに $MB_1={}^tAA({}^tAA-q_1E_n)={}^tAA{}^tAA-q_1{}^tAA$ より対称である。また $B_1M=(M_1-q_1E_n)M=(M-q_1E_n)M=M(M-q_1E_n)=MB_1$ である。

2) $l = k-1$ まで正しいとすると、

$M_k = MB_{k-1}$であるから、${}^tM_k = {}^tB_{k-1}{}^tM$となる。帰納法の仮定によりこれは

$B_{k-1}M = MB_{k-1} = M_k$となり、$M_k$は対称である。$B_k = M_k - q_kE_n$だから対称である。また

$MB_k = MM_k - q_kM = MB_{k-1}M - q_kM$であり、$M = M_1$は対称である。さらに

${}^t(MB_k) = {}^t(MB_{k-1}M) = {}^tM{}^tB_{k-1}{}^tM = MB_{k-1}M = MB_k$であるから、$MB_k$も対称となる。

$\mathrm{rank}(A) = r$とすると、$\mathrm{rank}(M) = \mathrm{rank}({}^tAA) = r$となる。ゆえに、行列 Mの特性多項式$\varphi_M(x)$ は

$p_{r+1} = \cdots = p_n = 0$ より

$$\varphi_M(x) = (-1)^n(x^n - p_1x^{n-1} - p_2x^{n-2} - \cdots - p_rx^{n-r})$$

となる。ここで、前にあげた Newton の公式より、$q_i = p_i (i = 1, \cdots, r, \cdots, n)$ に注意する。

Mの最小多項式を $f_M(x)$ として、次の定理を用いる。

「Mが対角化可能であることと Mの最小多項式 $f_M(x)$ が重根をもたないことは同値である ([1],p.198, 系 3.4)。」

これにより、Mは対称行列であるから対角化可能であり、$f_M(x)$ は重根をもたない。つまり、

$\varphi_M(x) = (-1)^nx^{n-r}(x^r - p_1x^{r-1} - \cdots - p_r)$ であり、$f_A(x)|\varphi_A(x)$、$f_A(x)$ が重根をもたないので、

$$f_M(x)|(x^r - p_1x^{r-1} - \cdots - p_r)x$$

となる。一方、$\mathrm{rank}(M) = r$, ${}^tM = M$より、

$$M \sim \mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0; n, n)\ (\text{ただし、}\lambda_i \neq 0)$$

となる。$\varphi_M(x) = x^{n-r}\prod_i(x - \lambda_i)$ だから $f_M(x) = x\prod(x - \lambda_i)$ (λ_iのうち同じものは省く)

よって、$p_r = \pm\prod_i(-\lambda_i)$ より、$p_r \neq 0$ が示される。以上で、$p_r = q_r = \mathrm{Tr}(MB_{r-1}) \neq 0$ が示された。さらに、$f_M(M) = 0$ であるから

$0 = (M^r - p_1M^{r-1} - \cdots - p_rE_n)M = M(M^r - p_1M^{r-1} - \cdots - p_rE_n) = MB_r$ が示される。

また、$Y = \frac{1}{q_r}B_{r-1}{}^tA$ とおくと、M, B_{r-1}, MB_{r-1}は対称行列であるから

$MB_{r-1} = {}^t(MB_{r-1}) = B_{r-1}M$が示され、以下の式がわかる。

$$\begin{aligned} AYA &= \tfrac{1}{q_r}AB_{r-1}{}^tAA = \tfrac{1}{q_r}AB_{r-1}M = \tfrac{1}{q_r}AMB_{r-1} \\ &= \tfrac{1}{q_r}AM_r = \tfrac{1}{q_r}A(B_r + q_rE_n) = \tfrac{1}{q_r}AB_r + A \end{aligned}$$

ここで、$MB_r = O$ つまり ${}^tAAB_r = O$ より $AB_r = O$ が示される。ゆえに、$AYA = A$ が成り立つ。同様にして、

$$YAY = \tfrac{1}{q_r^2}B_{r-1}B_r{}^tA + Y = Y$$

が成り立つ。また次式も成り立つ。

$$\begin{aligned} {}^t(AY) &= \tfrac{1}{q_r}{}^t(AB_{r-1}{}^tA) = \tfrac{1}{q_r}AB_{r-1}{}^tA = AY \\ {}^t(YA) &= \tfrac{1}{q_r}{}^t(B_{r-1}{}^tAA) = \tfrac{1}{q_r}{}^t(B_{r-1}M) = \tfrac{1}{q_r}(MB_{r-1}) \\ &= \tfrac{1}{q_r}B_{r-1}M = YA \end{aligned}$$

以上より、この Yが Moore-Penrose 逆行列の条件 (1)〜(4) をみたす。(証明終)

20.2.4 Grevilleの方法

定理 6 (Greville の方法,[5],[12]) A を (m,n) 型行列とする。A を列ベクトル表示して

$A=[a_1,a_2,\cdots,a_n]$ とおく。ここで、$A_{n-1}=[a_1,a_2,\cdots,a_{n-1}]$ とおくと、A^+は次式の形に書くことができる。

$$A^+=\begin{pmatrix} A_{n-1}^+-A_{n-1}^+a_nb_n^+ \\ b_n^+ \end{pmatrix} \tag{10}$$

ここで、b_n^+は m 次元列ベクトル b_nの Moore-Penrose 逆行列であり、次式で与えられる。

$$b_n=\begin{cases} 0 & (a_n=0\text{ のとき}) \\ (E_m-A_{n-1}A_{n-1}^+)a_n & (a_n\neq A_{n-1}A_{n-1}^+a_n\text{のとき}) \\ \dfrac{[1+{}^ta_n(A_{n-1}{}^tA_{n-1})^+a_n](A_{n-1}{}^tA_{n-1})^+a_n}{{}^ta_n(A_{n-1}{}^tA_{n-1})^+(A_{n-1}{}^tA_{n-1})^+a_n} & (a_n=A_{n-1}A_{n-1}^+a_n\neq 0\text{ のとき}) \end{cases} \tag{11}$$

$$\text{また}\quad b_n^+=\begin{cases} 0 & (b_n=0\text{ のとき}) \\ \dfrac{{}^tb_n}{{}^tb_nb_n} & (b_n\neq 0\text{ のとき}) \end{cases} \tag{12}$$

■証明

E(または E_m) は (m 次) 単位行列をあらわす。$a_n=0$ のときはすぐにわかる。$a_n\neq 0$ のときを示す。Moore-Penrose 逆行列の条件を帰納的に示す。(10) の右辺を Yとおく。

1. $a_n\neq A_{n-1}A_{n-1}^+a_n$のとき
   (a) $AYA=A$ の証明
      i. $n=1$ のとき
      $A_1=a_1$だから、(12) より $AA^+A=a_1\dfrac{{}^ta_1}{{}^ta_1a_1}a_1=a_1$となり、成立する。
      ii. $n-1$ まで正しいと仮定すると (10) より、

$$\begin{aligned} AY &= \begin{pmatrix} A_{n-1}^+-A_{n-1}^+a_nb_n^+ \\ b_n^+ \end{pmatrix}(A_{n-1},a_n) \\ &= A_{n-1}A_{n-1}^+-A_{n-1}A_{n-1}^+a_nb_n^++a_nb_n^+ \\ &= A_{n-1}A_{n-1}^++(E_m-A_{n-1}A_{n-1}^+)a_nb_n^+ \end{aligned}$$

ここで $b_n=(E_m-A_{n-1}A_{n-1}^+)a_n$ より $AA^+=A_{n-1}A_{n-1}^++b_nb_n^+$

これより

$$AYA = (A_{n-1}A_{n-1}^++b_nb_n^+)(A_{n-1},a_n)$$

$$= \quad (A_{n-1}A_{n-1}^{+}A_{n-1} + b_n b_n^{+} A_{n-1}\ ,\ A_{n-1}A_{n-1}^{+}a_n + b_n b_n^{+} a_n) \qquad (13)$$

以下、各成分を計算していく。まず (13) の第 1 成分について

$$\begin{aligned} b_n b_n^{+} A_{n-1} &= (E_m - A_{n-1}A_{n-1}^{+})\frac{a_n{}^t b_n}{{}^t b_n b_n}A_{n-1} \\ &= \frac{1}{{}^t b_n b_n}(E_m - A_{n-1}A_{n-1}^{+})a_n{}^t a_n{}^t(E_m - A_{n-1}A_{n-1}^{+})A_{n-1} \\ &= \frac{1}{{}^t b_n b_n}(E_m - A_{n-1}A_{n-1}^{+})a_n{}^t a_n(A_{n-1} - A_{n-1}A_{n-1}^{+}A_{n-1}) \\ &= O \qquad (14) \end{aligned}$$

また、第 2 成分について

$$\begin{aligned} A_{n-1}A_{n-1}^{+}a_n + b_n b_n^{+} a_n - a_n &= --(E_m - A_{n-1}A_{n-1}^{+})a_n + b_n b_n^{+} a_n \\ &= --b_n + b_n b_n^{+} a_n = -b_n(E_m - b_n^{+} a_n) \qquad (15) \\ &= --b_n \frac{{}^t b_n b_n - {}^t b_n a_n}{{}^t b_n b_n} \end{aligned}$$

ここで、(15) の分子について

$$\begin{aligned} {}^t b_n b_n - {}^t b_n a_n &= {}^t a_n(E_m - A_{n-1}A_{n-1}^{+})(E_m - A_{n-1}A_{n-1}^{+})a_n \\ &\quad -{}^t a_n(E_m - A_{n-1}A_{n-1}^{+})a_n \end{aligned} \qquad (16)$$

(16) の第 1 項において

$$\begin{aligned} (E_m - A_{n-1}A_{n-1}^{+})(E_m - A_{n-1}A_{n-1}^{+}) &= E_m - 2A_{n-1}A_{n-1}^{+} + A_{n-1}A_{n-1}^{+}A_{n-1}A_{n-1}^{+} \\ &= E_m - 2A_{n-1}A_{n-1}^{+} + A_{n-1}A_{n-1}^{+} \qquad (17) \\ &= E_m - A_{n-1}A_{n-1}^{+} \end{aligned}$$

よって、(16),(17) より ${}^t b_n b_n - {}^t b_n a_n = 0$ (18)

したがって、(15),(18) より $A_{n-1}A_{n-1}^{+}a_n + b_n b_n^{+} a_n - a_n = 0$ (19)

よって、(13),(14),(19) より $AYA = A$ が成立する。

(b) ${}^t(YA) = YA$ の証明

i. $n = 1$ のとき

容易に確かめられる。

ii. $n-1$ まで正しいと仮定すると (10) より

$$YA = \begin{pmatrix} A_{n-1}^{+} - A_{n-1}^{+}a_n b_n^{+} \\ b_n^{+} \end{pmatrix}(A_{n-1}, a_n)$$

$$= \begin{pmatrix} A_{n-1}^{+}A_{n-1} - A_{n-1}^{+}a_n b_n^{+} A_{n-1} & A_{n-1}^{+}a_n - A_{n-1}^{+}a_n b_n^{+} a_n \\ b_n^{+}A_{n-1} & b_n^{+}a_n \end{pmatrix} \quad (20)$$

まず、"(1,1) 成分"について

$$\begin{aligned} A_{n-1}^{+}a_n b_n^{+}A_{n-1} &= \frac{1}{{}^t b_n b_n} A_{n-1}^{+}a_n {}^t a_n (E_m - A_{n-1}A_{n-1}^{+})A_{n-1} \\ &= \frac{1}{{}^t b_n b_n} A_{n-1}^{+}a_n {}^t a_n (A_{n-1} - A_{n-1}A_{n-1}^{+}A_{n-1}) = O \end{aligned}$$

これより、

$${}^t(A_{n-1}^{+}A_{n-1}) - {}^t(A_{n-1}^{+}a_n b_n^{+}A_{n-1}) = (A_{n-1}^{+}A_{n-1}) \quad (21)$$

また"(1,2) 成分"について

$$\begin{aligned} A_{n-1}^{+}a_n - A_{n-1}^{+}a_n b_n^{+}a_n &= A_{n-1}^{+}a_n(E_m - b_n^{+}a_n) = A_{n-1}^{+}a_n(E_m - \frac{{}^t b_n}{{}^t b_n b_n}a_n) \\ &= \frac{1}{{}^t b_n b_n} A_{n-1}^{+}a_n({}^t b_n b_n - {}^t b_n a_n) \end{aligned}$$

$$\text{ここで、(18) より } {}^t(A_{n-1}^{+}a_n - A_{n-1}^{+}a_n b_n^{+}a_n) = 0 \quad (22)$$

また (2,1) 成分について

$${}^t(b_n^{+}A_{n-1}) = {}^t(\frac{{}^t b_n}{{}^t b_n b_n}A_{n-1}) = {}^t\left\{\frac{{}^t a_n}{{}^t b_n b_n}(E_m - A_{n-1}A_{n-1}^{+})A_{n-1}\right\} = 0 \quad (23)$$

$$\text{"(2,2) 成分"について } {}^t(b_n^{+}a_n) = \frac{{}^t a_n(E_m - A_{n-1}A_{n-1}^{+})a_n}{{}^t b_n b_n} = \frac{{}^t b_n a_n}{{}^t b_n b_n} = b_n^{+}a_n \quad (24)$$

したがって、(20),(21),(22), (23),(24) より、 ${}^t(YA) = YA$

また、$YAY = Y$, ${}^t(AY) = AY$の場合も同様にして示すことができる。

2. $a_n = A_{n-1}A_{n-1}^{+}a_n \neq 0$ のとき

ここでは、$AYA = A$ のみ示す。

(a) $n = 1$ のとき

容易に確かめられる。

(b) $n-1$ まで正しいと仮定すると

$$\begin{aligned} AY &= A_{n-1}A_{n-1}^{+} + (E_m - A_{n-1}A_{n-1}^{+})a_n b_n^{+} \\ &= A_{n-1}A_{n-1}^{+} + (E_m - A_{n-1}A_{n-1}^{+})A_{n-1}A_{n-1}^{+}a_n b_n^{+} = A_{n-1}A_{n-1}^{+} \end{aligned}$$

これより次式が成立する。

$$\begin{aligned} AYA &= (A_{n-1}A_{n-1}^{+})(A_{n-1}, a_n) &= (A_{n-1}A_{n-1}^{+}A_{n-1}, A_{n-1}A_{n-1}^{+}a_n) \\ &= (A_{n-1}, a_n) &= A \end{aligned}$$

また、$YAY=Y$, ${}^t(AY)=AY$, ${}^t(YA)=YA$ の場合も同様にして示すことができる。(証明終)

20.2.5 Hermiteの方法

定理 7 (Hermite の方法,[4],p.72, 定理 3.21,[15]) A を (m,n) 型行列とする。

1. $M=(A^tA)^2$とおく。M^-を Mの一般逆行列(つまり (1) のみをみたす行列)とすると A^+は次の形であらわせる。

$$A^+={}^tAM^-A^tA \tag{25}$$

2. $N=({}^tAA)^2$とおくと、A^+は次の形であらわせる。

$$A^+={}^tAAN^{-t}A \tag{26}$$

■証明

1. の場合について述べる。([5],p.67, 問 29) により、M^-は (m,n) 型行列 Zを用いて $M^-=M^+MM^++Z-M^+MZMM^+$ とあらわすことができる。A の特異値分解を $A=UD^tV$ (Uは m 次直交行列、Vは n 次直交行列、$D=\mathrm{diag}(\sqrt{\lambda_1},\ldots,\sqrt{\lambda_r},0,\ldots,0;m,n)$) とすれば $M=U\mathrm{diag}(\lambda_1^2,\ldots,\lambda_r^2,0,\ldots,0;m,m)^tU$ となり $M^+=U\mathrm{diag}(\lambda_1^{-2},\ldots,\lambda_r^{-2},0,\ldots,0;m,m)^tU$ となるから、

$$\begin{aligned}M^- \;=\;& U\mathrm{diag}(\lambda_1^{-2},\ldots,\lambda_r^{-2},0,\ldots,0;m,m)^tU+Z\\ & -U\mathrm{diag}(1,\ldots,1,0,\ldots,0;m,m)^tUZU\mathrm{diag}(1,\ldots,1,0,\ldots,0;m,m)^tU\end{aligned}$$

とかける。また $A^tA=UD^tD^tU$ であるから、

$$\begin{aligned}&{}^tAM^-A^tA\\ &= V^tD^tUU\mathrm{diag}(\lambda_1^{-2},\ldots,\lambda_r^{-2},0,\ldots,0;m,m)^tUUD^tD^tU+V^tD^tUZUD^tD^tU\\ &\quad -V^tD^tUU\mathrm{diag}(1,\ldots,1,0,\ldots,0;m,m)^tUZU\cdot\\ &\quad \cdot\mathrm{diag}(1,\ldots,1,0,\ldots,0;m,m)D^tD^tU\\ &= V\mathrm{diag}(\sqrt{\lambda_1}^{-1},\ldots,\sqrt{\lambda_r}^{-1},0,\ldots,0;n,m)^tU\end{aligned}$$

となり、定理 3により A^+と一致する。2. の場合も同様である。(証明終)

[注意]

M^-は${}^tM$を階数標準形に変更する行列を用いて簡単に求められる。すなわち、

$P^tM^tR=\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}$ (Pを m 次正則行列, Rを n 次正則行列, E_rを r次単位行列とする),

$r=\mathrm{rank}(M)$ とすると $M^-={}^tP^t\begin{pmatrix}E_r & O\\ O & O\end{pmatrix}R$ となるのは容易にわかる。

20.2.6 Ben-Israel,Wersan,Nobleの方法

定理 8 (Ben-Israel,Wersan,Nobleの方法,[7],p.65,Method 2,[8]) A を (m,n) 型行列とする。任意の (m,m) 型正則行列 M, 任意の n 次直交行列 P に対して A^+ は次の形であらわせる。

$$A^+ = P(M^tAAP)^+M^tA \tag{27}$$

■証明

e_i を n 次元縦ベクトル空間の i-th の基本ベクトルとする。[[2],p69, 問 36] により、${}^tAAx = {}^tAe_i$の解は $Ax = e_i$ $(i = 1,\ldots,n)$ の最小近似解であり、この中でノルム最小のものが $x = A^+e_i$となる。${}^tAAx = {}^tAe_i$ は $M^tAAPP^{-1}x = M^tAe_i$ とかける。$y = P^{-1}x$ とおくと、yは $M^tAAPy = M^tAe_i$の解であり、この逆も成立する。またPはn次直交行列だから、$\|y\| = \|x\|$ である。ゆえに、${}^tAAx = {}^tAe_i$ のノルム最小解は A^+e_iで、$M^tAAPy = M^tAe_i$のノルム最小解は $(M^tAAP)^+M^tAe_i$ であるから $P^{-1}A^+e_i = (M^tAAP)^+M^tAe_i$ となる。ゆえに
$A^+e_i = P(M^tAAP)^+M^tAe_i$ $(i = 1,\ldots,n)$ よって、$A^+ = P(M^tAAP)^+M^tA$ となる。(証明終)

20.2.7 Albertの方法

定理 9 (Albertの方法,[7],p.19,(3,4)theorem) 任意の (m,n) 型行列 A について、その Moore-Penrose 逆行列は次のようにあらわすことができる。

$$A^+ = \lim_{x\to 0} {}^tA(A^tA + x^2E_m)^{-1} \tag{28}$$

■証明

A を (m,n) 型行列とする。A の特異値分解を $A = UD^tV$ (Uを m 次直交行列、Vを n 次直交行列、$D = \mathrm{diag}(\sqrt{\lambda_1},\ldots,\sqrt{\lambda_r},0,\ldots,0;m,n)$ とする) とおく。まず (28) の左辺の極限の中を計算すると

$$\begin{aligned} {}^tA(A^tA + x^2E_m)^{-1} &= {}^t(UD^tV)\left\{(UD^tV^t(UD^tV) + x^2E_m\right\}^{-1} \\ &= V^tD^tU\left\{U(D^tD + x^2E_m)^tU\right\}^{-1} \end{aligned}$$

これより、

$$
\begin{aligned}
&{}^tA(A{}^tA + x^2E_m)^{-1} \\
&= V{}^tD(D{}^tD + x^2E_m)^{-1}{}^tU \\
&= V\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0; n, m) \cdot \\
&\quad \cdot\mathrm{diag}(\frac{1}{\lambda_1 + x^2}, \frac{1}{\lambda_2 + x^2}, \ldots, \frac{1}{\lambda_r + x^2}, \frac{1}{x^2}, \ldots, \frac{1}{x^2}; m, m){}^tU \\
&= V\mathrm{diag}(\frac{\sqrt{\lambda_1}}{\lambda_1 + x^2}, \frac{\sqrt{\lambda_2}}{\lambda_2 + x^2}, \ldots, \frac{\sqrt{\lambda_r}}{\lambda_r + x^2}, 0, \ldots, 0; n, m){}^tU \\
&\longrightarrow V\mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}, 0, \ldots, 0; n, m){}^tU \quad (x \to 0) \\
&= VD{}^tU
\end{aligned}
$$

定理3より、これは A^+ である。(証明終)

20.2.8 Ben-Israelの収束計算法

定理 10 (Ben-Israelの収束計算法,[5],p.64,5.3.2) A を (m,n) 型行列とし、$\mathrm{rank}(A) = r$とする。${}^tAA$ の最大固有値をλ_rとし、$2 < \alpha < \frac{2}{\lambda_r}$とおく。このとき $X_0 = \alpha{}^tA,\ X_{k+1} = X_k(2E_m - AX_k)$ とおくと、A^+は次の形で書ける。

$$A^+ = \lim_{k\to\infty} X_k \tag{29}$$

■証明

E_mを m 次単位行列とする。X_kが収束することを示す。A の特異値分解を $A = UD{}^tV$ (Uは m 次直交行列、Vは n 次直交行列、$D = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}, 0, \ldots, 0; m, n)$ $(0 < \sqrt{\lambda_1} \leq \sqrt{\lambda_2} \leq \cdots \leq \sqrt{\lambda_r})$) とすると、
$X_k = U\mathrm{diag}(d_{k1}, \ldots, d_{kr}, 0, \ldots, 0; m, n){}^tV$ $(k = 1, 2, \ldots)$ と書けること、および
$d_{k+1l} = 2d_{kl} - \sqrt{\lambda_l}d_{kl}^2$となることが X_kの漸化式より容易にわかる。$\lim_{k\to\infty} d_{k\ l}$ が存在すればよい。以下 d_{kl}の index lを省いて d_kとあらわす。つまり、$d_{k+1} = 2d_k - \sqrt{\lambda}d_k^2,\ d_0 = \alpha\sqrt{\lambda}$ となる。このとき、

1. $0 < \sqrt{\lambda}d_k < 1$ $(k = 1, 2, \ldots)$
2. $d_k < d_{k+1}$ $(k = 1, 2, \ldots)$

が成立することが帰納法で次のようにしてわかる。

1. $0 < \sqrt{\lambda}d_k < 1$ $(k = 1, 2, \ldots)$
   (a) $k = 1$ のとき

$$
\begin{aligned}
d_1 &= 2d_0 - \sqrt{\lambda}d_0^2 &&= d_0(2-\sqrt{\lambda}d_0) &&= \tfrac{1}{\sqrt{\lambda}}\sqrt{\lambda}d_0(2-\sqrt{\lambda}d_0) \\
&\leq \tfrac{1}{\sqrt{\lambda}}\tfrac{(\sqrt{\lambda}d_0+2-\sqrt{\lambda}d_0)^2}{4} &&= \tfrac{1}{\sqrt{\lambda}}
\end{aligned}
$$

また、$d_0 > 0,\ 2-\sqrt{\lambda}d_0 = 2-\alpha\lambda \geq 2-\frac{2\lambda}{\lambda_r} > 0$ である。よって $0 < \sqrt{\lambda}d_1 < 1$ となる。

(b) kまで成立したとき

$d_{k+1} = d_k(2-\sqrt{\lambda}d_k) = \frac{1}{\sqrt{\lambda}}\sqrt{\lambda}d_k(2-\sqrt{\lambda}d_k) \leq \frac{1}{\sqrt{\lambda}}\frac{\{\sqrt{\lambda}d_k+(2-\sqrt{\lambda}d_k)\}^2}{4} = \frac{1}{\sqrt{\lambda}}$ であり、

$d_k > 0,\ 2-\sqrt{\lambda}d_k > 2-1 > 0$ となる。よって $0 < \sqrt{\lambda}d_{k+1} < 1$ となる。

2. $d_{k+1} - d_k = d_k - \sqrt{\lambda}d_k^2 = d_k(1-\sqrt{\lambda}d_k) > 0\ (k = 1, 2, \ldots)$

1,2 より $\lim_{k\to\infty} d_k = d$ が存在する。$d_0 = \lambda \neq 0$ であるから、$d = 2d - \sqrt{\lambda}d^2$ より $d = \frac{1}{\sqrt{\lambda}}$ となる。
よって、定理 3 より $X_k \longrightarrow U\mathrm{diag}(\sqrt{\lambda_1}^{-1}, \ldots, \sqrt{\lambda_r}^{-1}, 0, \ldots, 0; m, n)^t V = A^+$ となる。（証明終）

[注意]

$d_k = 2d_{k-1} - \sqrt{\lambda}d_{k-1}^2$ より

$1-\sqrt{\lambda}d_k = (1-\sqrt{\lambda}d_{k-1})^2 = (1-\sqrt{\lambda}d_{k-2})^{2^2} = \cdots = (1-\sqrt{\lambda}d_0)^{2^{k-1}}$

よって $d_k = \frac{1}{\sqrt{\lambda}} - \frac{1}{\sqrt{\lambda}}(1-\alpha\sqrt{\lambda})^{2^k}$ ゆえに $|d_k - d| = \frac{1}{\sqrt{\lambda}}|1-\alpha\sqrt{\lambda}|^{2^k}$ となる。よって
$\max_{1\leq l\leq r}|1-\alpha\sqrt{\lambda_l}|$ が最小になるようにαを定めると収束が速い。$\sqrt{\lambda_1} \leq \sqrt{\lambda_2} \leq \cdots \sqrt{\lambda_r}$より、
$\alpha = \frac{2}{\sqrt{\lambda_1}+\sqrt{\lambda_r}}$とすると、この条件がみたされる。このとき

$|1-\alpha\sqrt{\lambda_l}| = \left|1 - \frac{2\sqrt{\lambda_l}}{\sqrt{\lambda_1}+\sqrt{\lambda_r}}\right| = \left|\frac{(\sqrt{\lambda_r}-\sqrt{\lambda_l})+(\sqrt{\lambda_1}-\sqrt{\lambda_l})}{\sqrt{\lambda_1}+\sqrt{\lambda_r}}\right|$ となる。

20.2.9 直接$x = A^+b$を求める方法

既に述べたが、$x = A^+b$ は$^tAAx = {}^tb$ の解で $\|x\|$ が最小である ([2],p69, 問 36)。これを用いて次の定理が示される。

定理 11 (直接 $x = A^+b$ を求める方法,[4],p.72, 定理 21, 注) $x = A^+b$ は

$$
\begin{pmatrix} E_m & {}^tAA \\ {}^tAA & O \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} O \\ {}^tAb \end{pmatrix} \tag{30}
$$

の解である。

■証明

$x = A^+b$ は$^tAAx = {}^tAb$ の条件のもとで $\|x\|$ を最小にするベクトルであるから、
$f(x, y) = \|x\|^2 - 2(y, {}^tAAx - {}^tAb)$ とおくと、
$\frac{\partial f}{\partial x_i} = 2x_i - 2({}^tAAy)_i = 0,\ \frac{\partial f}{\partial y_i} = ({}^tAAx - {}^tAb)_i = 0$ の解である (Lagrange の未定係数法)。ただし、$(y)_i$はベクトル yの第 i 成分である。よって $x = {}^tAAy,\ {}^tAAx = {}^tAb$ となる。これより (30) が示される。（証明終）

20.3 最後に

野田 [14] の数学的な補完として、代表的な一般逆行列の計算アルゴリズムの初等的な証明を示した。文献 [10],[11],[13], [16],[17] については現時点で入手できていないので、参考にできなかった。

参考文献

[1]斉藤正彦, 基礎数学 1 線形代数入門, 東京大学出版会 (1966)

[2]斉藤正彦, 基礎数学 4 線形代数演習, 東京大学出版会 (1985)

[3]古屋茂・小国力, 線形代数の計算法 (上・下), 産業図書 (1971)

[4]柳井晴夫・竹内啓, 射影行列・一般逆行列・特異値分解, 東京大学出版会 (1983)

[5]半谷裕彦・川口健一, 計算力学と CAE シリーズ 5 形態解析 一般逆行列とその応用, 培風館 (1991)

[6]Gregory,R.T.and Krishnamurthy,E.V.:Methods and Applications of Error-Free Computation,Springer-Verlag(1984)

[7]Albert,A.:Regression and the Moore-Penrose Pseudoinverse, Academic Press(1972)

[8]Ben-Israel,A.and Wersan,S.J.:An Elimination Method for Computing the Generalized Inverse of an Arbitrary Complex Matrix,J.ACM,Vol.10, pp.532-537(1963)

[9]Decell,H.P.,Jr.:An Application of the Cayley-Hamilton Theorem to Generalized Matrix Inversion ,SIAM Review,Vol.7,No.4, pp.526-528(1965)

[10]Frawley,W.J.:Computer Generation of Symbolic Generalized Inverses and Applications to Physics and Data Analysis,in Applications of Computer Algebra(Pavelle,R.ed.),pp.415-426,Kluwer Academic Pub.(1985)

[11]Glassey,C.R.:An Orthogonalization Method of Computing the Generalized Inverse of Matrix,ORC-66-10,Operations Research Center,Univ. of California,Berkeley(1966)

[12]Greville,T.N.E.:Some Applications of the Pseudoinverse of a Matrix,SIAM Review,Vol.2,No.1,pp.15-22(1960)

[13]Noble,B.:Methods for Computing the Moore-Penrose Generalized Inverse and Related Matters,in Generalized Inverse and Applications (Nashed,M.Z.ed.),pp.245-301,Academic Press(1973)

[14]Noda,M.,Izumida,M.,Ochi,M.(野田松太郎・泉田正則・越智正明): "一般逆行列の数式処理システムによる直接解法とその評価", 情報処理学会論文誌, Vol.30,No.11,pp.1376-1384(1989-11)

[15]Rao,T.M.,Subramanian,K.and Krishnamurthy,E.V.:Residue Arithmetic Algorithms for Exact

Computations of g-inverses of Matrices,SIAM J.Numer. Anal.,Vol.13,No.2,pp.155-171(1976)
[16]Shinozaki,N.,Sibuya,M. and Tanabe,K.:Numerical Algorithm for the Moore-Penrose Inverse of a Matrix:Direct Methods,Ann.Inst.Stat.Math., Vol.24,No.1,pp.193-203(1972)
[17]Shinozaki,N.,Sibuya,M. and Tanabe,K.:Numerical Algorithm for the Moore-Penrose Inverse of a Matrix:Iterative Methods,Ann.Inst.Stat. Math., Vol.24,No.3,pp.621-629(1972)