■解説

一般逆行列 (2)

田辺国士

前回は一般逆行列と連立一次方程式や射影行列との代数的関連についてのべました．今回は線型写像としての一般逆行列の意味，および特異値分解との関連についてのべましょう．

§1. 線型写像としての一般逆行列

前回と同様に，n 次元縦ベクトル全体からなる実ユークリッド空間を $\boldsymbol{R}^n$ とあらわすと，$m\times n$ 行列 A は $\boldsymbol{R}^n$ から $\boldsymbol{R}^m$ への線型写像

$$A:\boldsymbol{R}^n\ni x\longrightarrow Ax\in\boldsymbol{R}^m \tag{1}$$

を定義します．以後このように定義された線型写像をもとの行列と同一視して A であらわすことにします．A の像空間と核空間をそれぞれ

$$\boldsymbol{R}(A)\equiv\{y\in\boldsymbol{R}^m\,;\,Ax=y,\ x\in\boldsymbol{R}^n\}\subset\boldsymbol{R}^m$$
$$\boldsymbol{N}(A)\equiv\{x\in\boldsymbol{R}^n\,;\,Ax=0\}\subset\boldsymbol{R}^n$$

とすると，一般に $\boldsymbol{N}(A)$ は $\boldsymbol{R}^n$ の真部分空間 ($\boldsymbol{N}(A)\subsetneq\boldsymbol{R}^n$) となります．そこで，$\boldsymbol{N}(A)$ の適当な補空間 $\boldsymbol{V}$

$$\boldsymbol{V}\oplus\boldsymbol{N}(A)=\boldsymbol{R}^n \tag{2}$$

を考えます．ただし，$\oplus$は直和をあらわします．すなわち，$\boldsymbol{R}^n$ に属する任意のベクトル x が $\boldsymbol{V}$ に属するベクトル x_1 と $\boldsymbol{N}(A)$ に属するベクトル x_2 の和に一意的にあらわされるような空間 $\boldsymbol{V}$ を考えます．(図1) $\boldsymbol{N}(A)$ の補空間 $\boldsymbol{V}$ は一意的には定まりませんが，いずれをとっても，線型写像 A を部分空間 $\boldsymbol{V}$ に制限した $\boldsymbol{V}$ から $\boldsymbol{R}(A)$ への写像

$$A_{\boldsymbol{R}(A),\boldsymbol{V}}:\boldsymbol{V}\ni x\longrightarrow Ax\in\boldsymbol{R}(A) \tag{3}$$

は全単写 (1：1かつ上への写像) になり，したがって逆写像 $A_{\boldsymbol{R}(A),\boldsymbol{V}}^{-1}$ をもちます．おおざっぱにいうと，この $A_{\boldsymbol{R}(A),\boldsymbol{V}}^{-1}$ を，定義域が $\boldsymbol{R}^m$，値域が $\boldsymbol{R}^n$ になるように，適当に拡張した線型写像が A の一般逆行列なのです．このことは反射型一般逆行列の場合に，より正確にのべることができます．

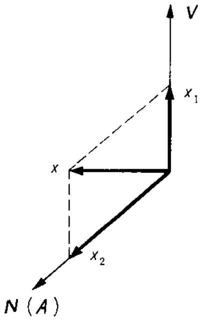

図 1

反射型一般逆行列 A_r^- に対して，$\boldsymbol{R}(A_r^-)$ と $\boldsymbol{N}(A_r^-)$ はそれぞれ $\boldsymbol{N}(A)$ と $\boldsymbol{R}(A)$ の補空間です．すなわち

$$\boldsymbol{R}(A_r^-)\oplus\boldsymbol{N}(A)=\boldsymbol{R}^n$$
$$\boldsymbol{R}(A)\oplus\boldsymbol{N}(A_r^-)=\boldsymbol{R}^m \tag{4}$$

が成り立ちます．このとき，$\boldsymbol{V}\equiv\boldsymbol{R}(A_r^-)$，$\boldsymbol{W}\equiv\boldsymbol{N}(A_r^-)$ とすると，任意の $y\in\boldsymbol{R}^m$ に対して

$$A_r^-y=A_{\boldsymbol{R}(A),\boldsymbol{V}}^{-1}P_{\boldsymbol{R}(A),\boldsymbol{W}}y \tag{5}$$

となります．ただし，$P_{\boldsymbol{R}(A),\boldsymbol{W}}$ は $\boldsymbol{W}$ に沿う $\boldsymbol{R}(A)$ の上への射影行列とします．(図2)

逆に，任意の補空間を $\boldsymbol{V}$, $\boldsymbol{W}$,

$$\left.\begin{aligned}\boldsymbol{V}\oplus\boldsymbol{N}(A)&=\boldsymbol{R}^n\\ \boldsymbol{R}(A)\oplus\boldsymbol{W}&=\boldsymbol{R}^m\end{aligned}\right. \tag{6}$$

とするとき，これらの部分空間を像空間，核空間とする一般逆行列 A^-

$$\boldsymbol{R}(A^-)=\boldsymbol{V},\ \boldsymbol{N}(A^-)=\boldsymbol{W} \tag{7}$$

が一意的に定まり，それは必ず反射型一般逆行列となります．これを $A_{\boldsymbol{V},\boldsymbol{W}}^+$ であらわすことにします．

$X=A_{\boldsymbol{V},\boldsymbol{W}}^+$ とするとき

$$XA=P_{\boldsymbol{V},\boldsymbol{N}(A)},\ AX=P_{\boldsymbol{R}(A),\boldsymbol{W}} \tag{8}$$

が成り立ちますが，逆にこれを満足する X は必ずしも一意的には定まりません．したがって，前回の式(16)のかわりに，この式によって反射型一般逆行列 $A_{\boldsymbol{V}\boldsymbol{W}}^+$ を定義することはできないことに注意してください．

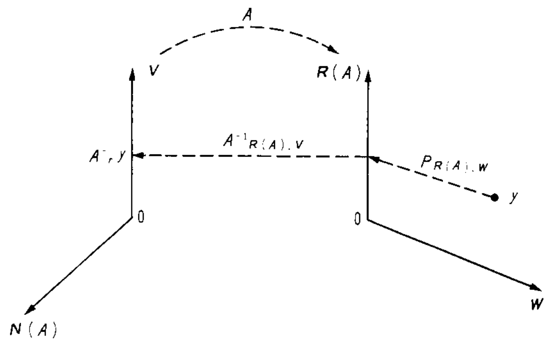

図 2

$X \in \{A_r^-\}$とするとき

$$X = A_{\boldsymbol{R}(X),\boldsymbol{N}(X)}{}^+, \quad A = X_{\boldsymbol{R}(A),\boldsymbol{N}(A)}{}^+ \tag{9}$$

が成り立ちます．また，$\boldsymbol{V}, \boldsymbol{W}$をそれぞれ$\boldsymbol{N}(A), \boldsymbol{R}(A)$の普通の内積の意味での直交補空間，すなわち

$$\boldsymbol{V} = \boldsymbol{R}(A^t), \quad \boldsymbol{W} = \boldsymbol{N}(A^t) \tag{10}$$

とすると，$A_{\boldsymbol{V},\boldsymbol{W}}{}^+$ は Moore-Penrose 逆行列A^+に一致します．

式(6)を満たす任意の補空間を$\boldsymbol{V}$，$\boldsymbol{W}$とするとき，$\boldsymbol{R}^n$と$\boldsymbol{R}^m$の内積をそれぞれ

$$\langle x_1, x_2 \rangle = x_2{}^t N x_1, \quad \langle y_1, y_2 \rangle = y_2{}^t M y_1 \tag{11}$$

ただし

$$M \equiv (P_{\boldsymbol{R}(A),\boldsymbol{W}})^t P_{\boldsymbol{R}(A),\boldsymbol{W}} + (P_{\boldsymbol{W},\boldsymbol{R}(A)})^t P_{\boldsymbol{W},\boldsymbol{R}(A)},$$
$$N \equiv (P_{\boldsymbol{V},\boldsymbol{N}(A)})^t P_{\boldsymbol{V},\boldsymbol{N}(A)} + (P_{\boldsymbol{N}(A),\boldsymbol{V}})^t P_{\boldsymbol{N}(A),\boldsymbol{V}} \tag{12}$$

と定義しなおすと，$\boldsymbol{V}$，$\boldsymbol{W}$はこの内積の意味での$\boldsymbol{N}(A)$，$\boldsymbol{R}(A)$の直交補空間になりますから，$A_{\boldsymbol{V},\boldsymbol{W}}{}^+$を$A^+$と対等のもの(あるいは$A^+$を一般化したもの)と考えることができます．実際，内積(11)から導かれる距離(ノルム)をそれぞれ$\|x\|_N \equiv \sqrt{x^t N x}$，$\|y\|_M = \sqrt{y^t M y}$とするとき，連立一次方程式

$$Ax = b \tag{13}$$

の重みつき最小二乗解，すなわち

$$\|Ax - b\|_M{}^2 \tag{14}$$

を最小にするxのなかでノルム$\|x\|_N$が最小の解は$A_{\boldsymbol{V},\boldsymbol{W}}{}^+ b$になります．

これと同様の考えから，任意の正定符号行列M, Nに対して，内積を(11)で定義することにより，最小ノルム型一般逆行列A_m^-や最小二乗型一般逆行列A_l^-に対応する一般逆行列を定義することができます．条件

$$AXA = A, \quad (NXA)^t = NXA \tag{15}$$

を満たす一般逆行列XをNノルム最小型一般逆行列とよび，$A_{m(N)}{}^-$であらわします．連立一次方程式(13)が解をもつときには，$A_{m(N)}{}^- b$は，ノルム$\|x\|_N$が最小の解となります．また，条件

$$AXA = A, \quad (MAX)^t = MAX \tag{16}$$

を満たす一般逆行列XをM最小二乗型一般逆行列とよび，$A_{l(M)}{}^-$であらわします．$A_{l(M)}{}^- b$は連立一次方程式の重みつき最小二乗解となります．

条件(15)(16)を満たす反射型一般逆行列をMoore-Penrose型一般逆行列とよび，$A_{MN}{}^+$であらわします．$X = A_{MN}{}^+$とするとき

$$A_{MN}{}^+ = A_{\boldsymbol{R}(X),\boldsymbol{N}(X)}{}^+ \tag{17}$$

となりますから，反射型一般逆行列はすべて，適当なノルムのもとでMoore-Penrose型一般逆行列であることがわかります．

§ 2. 特異値分解と一般逆行列

行列の特異値分解を用いると，一般逆行列の行列としての性質がいっそう明らかになります．

任意の$m \times n$行列$A = (a_{ij})$に対して，特異値分解とよばれるつぎのような分解が存在します．

$$A = UDV^t, \tag{18}$$

ただし，$D = (d_{ij})$は$m \times n$対角行列 $(i \neq j \Rightarrow d_{ij} = 0)$でその対角成分$d_{ii} = \sigma_i$は非負の実数で，特異値とよばれます．$U$と$V$はそれぞれ$m \times m$，$n \times n$の直交行列です．特異値は$\sigma_1 \geq \sigma_2 \geq \cdots\cdots \geq \sigma_r > 0$と順序づけられているものとします．このとき$r = \mathrm{rank}\, A$です．特異値分解(18)は行列の写像としての性質を端的に表現しています．すなわち写像A

$$A : x \longrightarrow V^t x \longrightarrow D(V^t x) \longrightarrow U(DV^t x) = Ax \tag{19}$$

の第一段と第三段はベクトルの回転・鏡映をあらわし，第二段はベクトルの成分ごとの拡大・縮小をあらわしています．

ベクトルのユークリッドノルム$\|x\|_2$から導かれる行列Aのノルムを

$$\|A\|_2 \equiv \max_x \frac{\|Ax\|_2}{\|x\|_2} \tag{20}$$

とするとき，$\|A\|_2 = \sigma_1$となります．また，行列Aのフロベニウスノルム

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}{}^2} \tag{21}$$

に対して，$\|A\|_F = \sqrt{\sum_{i=1}^r \sigma_i{}^2}$となります．いま0でない特異値$\sigma_i$をその対角成分にもつ対角行列を$\Delta$とします．

$$\Delta = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_r \end{pmatrix} \tag{22}$$

このとき，一般逆行列A^-，最小ノルム型一般逆行列A_m^-，最小二乗型一般逆行列A_l^-，反射型一般逆行列A_r^-，Moore-Penrose逆行列A^+はそれぞれつぎのようにあらわされます．

$$A^- = V \left[\begin{array}{c:c} \Delta^{-1} & G \\ \hdashline F & H \end{array}\right] U^t, \tag{23}$$

$$A_m^- = V \left[\begin{array}{c:c} \Delta^{-1} & G \\ \hdashline O & H \end{array}\right] U^t, \tag{24}$$

$$A_l^- = V \left[\begin{array}{c:c} \Delta^{-1} & O \\ \hdashline F & H \end{array}\right] U^t, \tag{25}$$

$$A_r^- = V\begin{bmatrix} \Delta^{-1} & G \\ F & F\Delta G \end{bmatrix} U^t, \tag{26}$$

$$A^+ = V\begin{bmatrix} \Delta^{-1} & O \\ O & O \end{bmatrix} U^t, \tag{27}$$

ただし，F, G, Hはそれぞれ適当な部分行列とします．直交行列Uの最初のr個の縦ベクトルが張る部分空間が$\boldsymbol{R}(A)$，直交行列Vの最後から$n-r$個の縦ベクトルが張る部分空間が$\boldsymbol{N}(A)$であることに注意すると前節でのべたことがよりいっそうはっきり理解できるでしょう．

Moore-Penrose一般逆行列の表現(27)においてはrank A，すなわち0でない特異値の数が重要な役割りを負っています．A^+の形成にあたって，非零特異値がいかに小さな値であっても，0の特異値とは峻別されます．したがって，A^+をAの要素の函数(写像)と考えるとき連続写像ではないのです．このことが，一般逆行列の計算に不安定性をもたらす最大の原因です．ある種の不適切問題から導かれる行列は観測誤差，計算誤差，離散化誤差のために小さな特異値が生じ，本来の階数よりもみかけ上高い階数になっている場合があります．このようなときには，そのMoore-Penrose逆行列はあまり役に立たないことになります．もし私たちがその行列の実質的階数kをしっているならば，それを考慮して逆行列もどきをつくると便利でしょう．行列Aの特異値分解を(28)とするとき，大きいほうのk個の特異値を対角成分にもつ$k\times k$行列を

$$\Delta_k = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \\ & & & \sigma_k \end{pmatrix} \tag{28}$$

と定義し，(27)のかわりに行列$A_k{}^+$を

$$A_k{}^+ = V\begin{bmatrix} \Delta_k{}^{-1} & O \\ O & O \end{bmatrix} V^t \tag{29}$$

と定義するならば，いまのべたような場合には役立つでしょう．この$A_k{}^+$を階数kの擬逆行列とよぶことにします．$A_k{}^+$はrank $X=k(k\leqq \text{rank } A)$となる行列$X$の中で

$$\|A-X\| \tag{30}$$

が最小となる行列A_kのMoore-Penrose逆行列です．ただし，この場合ノルムは$\|\cdot\|_2$と$\|\cdot\|_F$の両方の意味で成り立ちます．

以下にのべるMoore-Penrose逆行列の漸近的表現は，すべて特異値分解(28)を用いることにより容易に証明できます．

□ $$A^+ = \alpha\sum_{i=0}^{\infty}(I-\alpha A^tA)^iA^t = \alpha\sum_{i=0}^{\infty}A^t(I-\alpha AA^t)^i \tag{31}$$

ただし，$0<\alpha<2\|A\|_2{}^{-2}$とします．この級数は連立一次方程式(13)の最小二乗解を求める線型反復解法

$$x_{i+1} = x_i + \alpha A^t(b-Ax_i) \tag{32}$$

と密接に関連しています．ちなみに，この反復法は，方程式(13)の最小二乗解

$$A^+b + (I-A^+A)x_0 \tag{33}$$

に収束します．

□ $$A^+ = \lim_{\lambda\to+0}(\lambda I+A^tA)^{-1}A^t = \lim_{\lambda\to+0}A^t(\lambda I+AA^t)^{-1} \tag{34}$$

$R_\lambda \equiv (\lambda I+A^tA)^{-1}A^t$とおくと，$R_\lambda b$は函数

$$\|b-Ax\|_2{}^2 + \lambda\|x\|_2{}^2 \tag{35}$$

を最小とするベクトルです．この第1項は残差平方和，第2項はxのノルムの平方にλを掛けた一定の罰金(penalty)で，$\lambda\to+0$にしてゆく過程はSUMTと同じ考え方です．さきほど不必要な特異値を制御するために階数条件つき擬行列$A_k{}^+$を導入しましたが，類似のことがR_λを利用することにより可能となります．ただし，この場合は階数kのかわりに正則化パラメータλを適当に選んでやる必要があります．R_λは正則化行列とよばれ，Tikhonovによって導入されたものです．

□ $$A^+ = \sum_{i=1}^{\infty}(I+A^tA)^{-i}A^t = \sum_{i=1}^{\infty}A^t(I+AA^t)^{-i}$$

□ $$A^+ = \lim_{\mu\to\infty}\int_0^1 \exp\{-A^tA(\mu-s)\}A^t\,ds = \lim_{\mu\to\infty}\int_0^1 A^t\exp\{-AA^t(\mu-s)\}\,ds$$

参 考 文 献

[1] A. Ben-Israel and A. Charnes, SIAM J. 11, pp. 667—699, 1963.

[2] G. H. Golub and C. Reinsch, Numerische Mathematik 14, pp. 403—420, 1970.

[3] D. Showalter, Proc. Amer. Math, Soc. 18, pp. 584—586, 1967.

[4] M. Sibuya, Ann. Inst. Statist. Math. 22, pp. 543—556, 1970.

[5] K. Tanabe ; Numerische Mathematik 22, pp. 349—359, 1974.

[6] K. Tanabe ; Computation and Analysis 6—4 pp. 2—25, 1975.

[7] 田辺国士；日科技連：計算機活用セミナーテキスト pp. 41—62, 1975.

[8] A. N. Tikhonov, USSR Computational Math. and Math. Physics 5, pp. 181—188, 1965.

(たなべ・くにお　統計数理研究所)