

文章编号: 0529-6579 (1999) 03-0012-05

四元数矩阵微分及其在精确分布上的应用^{*}

滕成业, 李绍明

(中山大学数学系, 广东 广州 510275)

摘 要: 讨论了四元数矩阵的外微分形式, 得出四元数矩阵变换下 Jacobi 行列式的一些结果, 利用这些结果, 简化了四元数 Wishart 分布的推导, 并且在求得四元数 Stiefel 流形体积的基础上, 进一步导出其特征根分布

关键词: 四元数; 外微分式; 四元数正态分布; 四元数 Wishart 分布

中图分类号: O 212 2 4 **文献标识码:** A

Hamilton 于 1843 年最先发现了四元数. 从代数学观点看, 实数域的代数扩域(体)只有 3 种: 实数域、复数域及四元数体. 由于四元数乘法的不可交换性, 使得对它的研究要比对实数、复数的研究困难得多. 但是, 近 20 年来, 特别在中国就代数学领域, 对四元数的研究, 尤其是四元数矩阵论的研究, 已经成为一个研究热点.

1975 年, Anderson^[1] 首先研究了四元数正态模型, 涉及到四元数正态分布及协方差的极大似然估计, 但他所运用的工具并不好; 另外, 他所选择的同构实矩阵有错, 所以, 无法给出四元数正态分布的分析表达式. 1997 年滕成业等^[2] 通过建立四元数与实矩阵的同构, 直接用实分析的方法研究四元数正态分布, 并且通过特征函数的反演, 得到了 Wishart 分布的密度函数.

本文首先讨论了四元数矩阵的外微分形式, 并且求出四元数上 Stiefel 流形的体积, 这些结果在四元数统计分析中, 对导出各种精确分布起到重要的作用. 利用这些结果, 可直接推导四元数 Wishart 分布的密度函数, 从而简化了文[2] 的推导过程, 并进一步推导出 Wishart 矩阵的特征根分布.

1 四元数矩阵变换下矩阵的外微分式之间的关系

记 $Q^{n \times n} = \{X \mid X_r + X_i \cdot i + X_j \cdot j + X_k \cdot k, \text{ 其中 } X_r, X_i, X_j, X_k \in \mathbb{R}^{m \times n}\}$ 为四元数矩阵集合. $\forall X \in Q^{n \times n}$, X 的范数记为 $\|X\| = [\text{tr}(X^* \cdot X)]^{1/2}$, 此处 X^* 指矩阵 X 的共轭转置. 这样使 $Q^{n \times n}$ 成为度量空间, 能建立相应的拓扑结构, 且 $Q^{n \times n}$ 与 $\mathbb{R}^{4m \times n}$ 是同胚的. 由此可把 $Q^{n \times n}$ 作为 $4mn$ 维流形, 在 $Q^{n \times n}$ 的每一点处定义切空间、微分空间(余切空间)、相应的

^{*} 基金项目: 中山大学高等学术基金(98M4)资助项目

收稿日期: 1998-12-02 作者简介: 滕成业, 男, 1940 年生, 教授.

外微分式, 并且可定义流形上的外微分及积分^[3].

定义1 设 $X \in Q^{m \times n}$, $X = X_r + X_i \circ i + X_j \circ j + X_k \circ k$, 其中 $X_r, X_i, X_j, X_k \in R^{m \times n}$,

则 $dX \stackrel{\text{def}}{=} dX_r + dX_i \circ i + dX_j \circ j + dX_k \circ k$.

定义2 上述 X 的 $4mn$ -外微分式定义为

$$(dX) \stackrel{\text{def}}{=} (dX_r) \wedge (dX_i) \wedge (dX_j) \wedge (dX_k).$$

其中, $(dX_r), (dX_i), (dX_j), (dX_k)$ 分别为实矩阵 X_r, X_i, X_j, X_k 的 mn -外微分式, 符号“ $\wedge$ ”指外积.

定理1 若 $X, Y \in Q^{m \times n}$ 为矩阵变元, $B \in Q^{n \times n}$, $C \in Q^{m \times n}$ 为矩阵常量, 且 $X = BYC$, 则 X 的 $4mn$ -外微分式为: $(dX) = [\det_r B]^n \circ [\det_r C]^m (dY)$

$$\text{其中, 实矩阵 } {}_rB = \begin{bmatrix} B_r & B_i & B_j & B_k \\ -B_i & B_r & -B_k & B_j \\ -B_j & B_k & B_r & -B_i \\ -B_k & -B_j & B_i & B_r \end{bmatrix}, \quad {}_rC = \begin{bmatrix} C_r & C_i & C_j & C_k \\ -C_i & C_r & -C_k & C_j \\ -C_j & C_k & C_r & -C_i \\ -C_k & -C_j & C_i & C_r \end{bmatrix}$$

证明 利用文[2]定理1及文[4]定理2.1.4, 经观察第1个子竖列变化可得结论.

定义3 若 $X \in Q^{n \times n}$, 且为自共轭矩阵, 即: $X^* = X$.

$$X = X_r + X_i \circ i + X_j \circ j + X_k \circ k$$

易知, X_r 为实对称矩阵, X_i, X_j, X_k 为反对称实矩阵. 此时 X 的 $(2m^2 - m)$ -外微分式定义为

$$(dX) = \left(\bigwedge_{p \leq q}^m dX_{pq} \right) \wedge \left(\bigwedge_{p < q}^m dX_{ipq} \right) \wedge \left(\bigwedge_{p < q}^m dX_{jpq} \right) \wedge \left(\bigwedge_{p < q}^m dX_{kpq} \right).$$

定理2 若有自共轭矩阵变元 $X, Y \in Q^{m \times m}$, $B \in Q^{m \times m}$ 为非奇异的矩阵常量, 且 $X = B \circ Y \circ B^*$. 则 X 的 $(2m^2 - m)$ -外微分式为

$$(dX) = [\det_r \beta]^{(2m-1)/2} (dY)$$

证明 类似定理1, 经详细计算可得结论. 证明过程略.

定义4 若 $X \in Q^{m \times m}$ 为 Hermit 反对称四元数矩阵, $X = X_r + X_i \circ i + X_j \circ j + X_k \circ k$.

易知, X_r 为反对称实矩阵, X_i, X_j, X_k 为实对称矩阵. 由此定义 X 的 $(2m^2 + m)$ -外微分式定义为

$$(dX) \stackrel{\text{def}}{=} (dX_r) \wedge d(X_i) \wedge d(X_j) \wedge d(X_k)$$

定理3 若 $X, Y \in Q^{m \times m}$, 且均为 Hermit 反对称矩阵变元, 且 $X = B \circ Y \circ B^*$, $B \in Q^{m \times n}$ 为非奇异的矩阵常量, 则 X 的 $(2m^2 + m)$ -外微分式定义为

$$(dX) = [\det_r \beta]^{(2m+1)/2} (dY)$$

定理4 若 $X, Y \in Q^{n \times m}$ 为可逆的自共轭变元, $Y = X^{-1}$, 则 Y 的 $(2m^2 - m)$ -外微分式为

$$(dY) = [\det_r X]^{-(2m-1)/2} (dX)$$

证明 由 $X \circ Y = I$ 知, $dX \circ Y + X \circ dY = 0$, 即

$$dY = -X^{-1} \circ dX \circ X^{-1}$$

由定理2可得结论.

定理5 设 $Z_{n \times m} = H_1 \circ T$, $n \geq m$, T 为 m 阶主对角线元素为正的上三角矩阵, H_1 为

$n \times m$ 列正交矩阵: $H^* \circ H_1 = I$. 设 H_2 为 $n \times (n-m)$ 列正交矩阵, 且满足

$$H = [H_1 \vdots H_2] = [h_1 \cdots h_m \vdots h_{m+1} \cdots h_n], H^* \circ H = I$$

则 $Z_{n \times m}$ 的 $4mn$ -微分式为

$$(dZ) = \prod_{i=1}^m t_{ii}^{4(n-i)+3} (dT) \wedge (H_1^* dH_1)$$

其中, $(H_1^* dH_1) \stackrel{\text{def}}{=} \bigwedge_{i=1}^m \bigwedge_{j=i}^m (h_j^* dh_i)$

证明 由 $Z = H_1 \circ T$, 有 $dZ = dH_1 \circ T + H_1 \circ dT$, 故

$$H^* dZ = \begin{bmatrix} H_1^* \\ H_2^* \end{bmatrix} \circ dZ = \begin{bmatrix} H_1^* dH_1 \circ T + dT \\ H_2^* dH_1 \circ T \end{bmatrix}$$

计算此式两边的外微分式可得

$$(H^* dZ) = (dZ), (H_1^* dH_1 \circ T + dT) = \prod_{i=1}^m t_{ii}^{4(n-i)+3} \bigwedge_{i=1}^n \bigwedge_{j=i}^m (h_j^* dh_i) \wedge (dT)$$

$$(H_2^* dH_1) \circ T = \prod_{i=1}^m t_{ii}^{4(n-m)} \bigwedge_{j=m+1}^n \bigwedge_{i=1}^m (h_j^* dh_i)$$

由此可得定理结论.

定理 6 设 $V_{n,m} = \{H_1 \mid H_1 \in Q^{n \times m}, \text{ 且 } H_1^* \circ H_1 = I, \text{ 其中, } n \geq m\}$, 则有

$$\int_{V_{n,m}} (H_1^* dH_1) = \frac{2^m \pi^{2mn-m^2+m}}{\prod_{i=1}^m \Gamma[2(n-i+1)]} \stackrel{\text{def}}{=} \text{Vol}(V_{n,m})$$

其中, $(H_1^* dH_1)$ 为定理 5 所定义.

证明 设 $Z \in Q^{n \times m}$ 为四元数随机矩阵, 且它的所有元素的实分量相互独立, 均服从 $N(0, 1)$ 分布则有

$$\int_{-\infty < z_{pq} < +\infty} \cdots \int \text{etr}\left(-\frac{1}{2} Z^* \circ Z\right) (dZ) = (2\pi)^{2mn}$$

令 $A = Z^* Z$, $Z = H_1 T$, $H_1 \in V_{n,m}$, T 为对角线为正的上三角四元数矩阵.

把 $\text{tr}(Z^* Z) = \sum_{i,j} t_{ij}^* t_{ij}$ 及定理 5 微分式代入上式并且在积分中积出 (dT) 部分, 比较剩余部分即可得出结论.

定理 7 设 $Z \in Q^{n \times m}$ 为列满秩的四元数矩阵, $n \geq m$. $Z = H_1 T$, H_1 为列正交阵, T 为对角线元素为正的上三角四元数矩阵, $A = Z^* Z = T^* T$, 则 Z 的 $4mn$ -外微分式为

$$(dZ) = 2^{-m} [\det A]^{2(n-m)+1} (dA) \wedge (H_1^* dH_1)$$

证明 由文[2]引理 5.1 知 $(dT) = 2^{-m} \prod_{i=1}^m t_{ii}^{-4(m-i)-1} (dA)$ 代入定理 5 右端, 利用 $\det A = \prod_{i=1}^m t_{ii}^2$, 即可得结论.

定理 8 设 $A \in Q^{n \times m}$ 为非奇异自共轭矩阵变元, $A = H \Lambda H^*$, H 为 Hermit 正交矩阵, Λ 为对角矩阵, 且对角线元素 $\lambda_1 > \lambda_2 > \cdots > \lambda_m > 0$, 则 A 的 $(2m^2 - m)$ -外微分式

$$(dA) = \prod_{i < j}^m (\lambda_i - \lambda_j)^4 \bigwedge_{i < j}^m (h_j^* dh_i) \bigwedge_{i=1}^m d\lambda_i$$

证明 由 $A = H\Delta H^*$,

有 $dA = dH \circ \Delta \circ H^* + Hd\Delta \circ H^* + H\Delta dH^*$

所以 $H^* dAH = H^* \circ dH \circ \Delta - \Delta \circ H^* dH + d\Delta$

经计算 $H^* dH\Delta - \Delta H^* dH$ 的 $2m(m-1)$ -外微分式为

$$\prod_{i < j}^m (\lambda_i - \lambda_j)^4 \bigwedge_{i < j}^m (h_j^* dh_i)$$

进一步考察 $d\Delta$ 的 m -外微分式 $(d\Delta) = \bigwedge_{i=1}^m d\lambda_i$, 由此可得定理结论.

2 四元数 Wishart 矩阵及其特征根分布的密度函数

定义 5 设 $Z_1, \dots, Z_n$ 为 m 维独立同分布的四元数随机列向量,

$$Z_i \sim N(0, \Sigma), 1 \leq i \leq n, \text{ 令 } A = \sum_{i=1}^n Z_i \circ Z_i^*$$

则称 A 为四元数 Wishart 矩阵, 记 $A \sim W_m(n, \Sigma)$.

定理 9 上述定义的 Wishart 矩阵 A 的密度函数为

$$pdf(A) = \frac{2^{2mn} (\det \Sigma)^{-2n}}{\pi^{m(m-1)} \prod_{i=1}^m \Gamma(2(n-i+1))} \cdot [\det A]^{2(n-m)+1} \cdot \exp\{-2\text{Retr}(\Sigma^{-1}A)\}$$

证明 记 $Z = [Z_1 \cdots Z_n]$, 由文[2] 定理 3.1 可得 Z 的密度函数为

$$pdf(Z) = \frac{2^{2mn}}{\pi^{m(m-1)}} (\det \Sigma)^{-2n} \cdot \exp\{-2\text{Retr}(\Sigma^{-1}Z \circ Z^*)\}$$

由定理 7 可得

$$pdf(Z)(dZ) = \frac{2^{2mn-n}}{\pi^{2mn}} [\det \Sigma]^{-2n} \cdot [\det A]^{2(n-m)+1} \cdot \exp\{-2\text{Retr}(\Sigma^{-1}A)\} (dA) \wedge (H_1^* dH)$$

上式对 $(H_1^* dH_1)$ 在 $V_{n,m}$ 上积分, 利用定理 6, 即可得出定理结论.

定理 10 若 $A \sim W_m(n, \Sigma)$, 且 $\Sigma = I$, 则 A 的特征根 $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$ 的分布密度函数为

$$\frac{2^{2mn} \pi^{2m}}{\prod_{i=1}^m \Gamma(2(n-i+1)) \cdot \prod_{i=1}^m \Gamma(2(m-i+1))} \prod_{i=1}^m \lambda_i^{2(n-m)+1} \cdot \prod_{i < j}^m (\lambda_i - \lambda_j)^4 \cdot \exp\left(-2 \sum_{i=1}^m \lambda_i\right)$$

证明 由定理 8, 把 (dA) 的表达式代入 $pdf(A)(dA)$ 的表达式中, 利用定理 6 的结论, 对 $H \in V_{m,m}$ 积分, 可得定理结论.

定理 11 若 $A \sim W_m(n, \Sigma)$, 且 $\Sigma \neq \lambda I$, 则 A 的特征根 $\lambda_1 > \cdots > \lambda_m > 0$ 的分布密度函数为

$$\frac{2^{2mn} \pi^{2m}}{\prod_{i=1}^m \Gamma(2(n-i+1)) \cdot \prod_{i=1}^m \Gamma(2(m-i+1))} \cdot \prod_{i=1}^m \lambda_i^{2(n-m)+1} \cdot \prod_{i < j}^m (\lambda_i - \lambda_j)^4 \cdot \sum_{k=0}^{\infty} \frac{(-2)^k}{k!} \cdot \sum_{\kappa} \frac{C_{\kappa}(\Sigma^{-1}) \cdot C_{\kappa}(\Lambda)}{C_{\kappa}(I)}$$

其中, $C_{\kappa}(\circ)$ 为 κ 的不超过 m 的划分 $\mathbb{I}_{\kappa}$ 所对应的带状多项式. $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$.

证明 $pdf(A)(dA) \propto [\det A]^{2(n-m)+1} \exp\{-2\text{Retr}(\Sigma^{-1}A)\} (dA)$

由定理 8 进一步可得

$$\propto \prod_{i=1}^m \lambda_i^{2(n-m)+1} \exp\{-2\text{Re tr}(\Sigma^{-1} \mathbf{H} \wedge \mathbf{H}^*)\} \circ \prod_{i < j}^m (\lambda_i - \lambda_j)^4 \bigwedge_{i < j}^m (\mathbf{h}_j^* d\mathbf{h}_i) \bigwedge_{i=1}^m (d\lambda_i)$$

记 $(d\mathbf{H}) \stackrel{\text{def}}{=} \frac{\bigwedge_{i < j}^m (\mathbf{h}_i^* d\mathbf{h}_j)}{\text{Vol}(\mathbf{V}_{m,m})}$, 由定理 6, 上式右边对 $\mathbf{H} \in \mathbf{V}_{m,m}$ 积分, 得

$$\begin{aligned} & \propto \prod_{i=1}^m \lambda_i^{2(n-m)+1} \circ \prod_{i < j}^m (\lambda_i - \lambda_j)^4 \int_{\mathbf{H} \in \mathbf{V}_{m,m}} \exp\{-2\text{Re tr}(\Sigma^{-1} \mathbf{H} \wedge \mathbf{H}^*)\} \circ (d\mathbf{H}) \bigwedge_{i=1}^m (d\lambda_i) \\ & \propto \prod_{i=1}^m \lambda_i^{2(n-m)+1} \circ \prod_{i < j}^m (\lambda_i - \lambda_j)^4 \sum_{k=0}^{\infty} \frac{(-2)^k}{k!} \circ \int_{\mathbf{H} \in \mathbf{V}_{m,m}} [\text{Re tr}(\Sigma^{-1} \mathbf{H} \wedge \mathbf{H}^*)]^k \circ (d\mathbf{H}) \bigwedge_{i=1}^m (d\lambda_i) \end{aligned}$$

利用文[5] 引理 5, 可得 $\int_{\mathbf{H} \in \mathbf{V}_{m,m}} [\text{Re tr}(\Sigma^{-1} \mathbf{H} \wedge \mathbf{H}^*)]^k \circ (d\mathbf{H}) = \sum_{\kappa} \frac{C_{\kappa}(\Sigma^{-1}) C_{\kappa}(\Lambda)}{C_{\kappa}(I)}$ 代入上式, 即可得定理结论.

参考文献:

- [1] ANDERSON S. Invariant normal models[J]. Ann Statist, 1975(3): 132~154.
- [2] TENG C Y, FAN K T. Statistical analysis based on normal distribution of quaternion[A]. Contemporary multivariate analysis and its application[D]. Hong Kong, 1997. 19~27.
- [3] 陈省身, 陈维桓. 微分几何讲义[M]. 北京: 北京大学出版社, 1983. 1~100.
- [4] MUIRHEAD R J. Aspect of multivariate statistical theory[M]. New York: John Wiley & Sons, 1982. 50~107.
- [5] TENG C Y, FAN H B, DENG W C. The generalized noncentral Wishart distribution[J]. J Math Research and Exposition, 1989(9): 479~488.

Exterior Differential Forms on Quaternion Matrices and Its Application

*
TENG Cheng-ye, LI Shao-ming

Abstract: It is frequent that quaternion matrix transformations are used in deriving many important quaternion distributins. In these cases, several exterior differential forms on quaternion matrices given in this paper are crucial in the derivation of many density functions. We also give the volume of the Stiefel manifold on the quaternion. By using these results, we directly derive the density function of Wishart matrix, which dramatically simplifies the existing derivation of Wishart distribution. Finally, we acquire the density functin of latent roots of quaternion Wishart matrix.

Keywords: quaternion; quaternion normal distribution; quaternion Wishart distribution; exterior differential form

* Department of Mathematics, Zhongshan University, Guangzhou 510275, China