第 23 卷 第 2 期
2008 年 6 月

天津科技大学学报
Journal of Tianjin University of Science & Technology

Vol.23 No. 2
Jun. 2008

# Mahalanobis距离在多元随机变量线性变换中的定理

张绍璞
（天津科技大学理学院，天津 300457）

**摘 要**：根据多元统计分析中 Mahalanobis 距离的计算理论,对其计算性质进行了讨论和研究,证明了多元随机变量在经过可逆线性变换后,其 Mahalanobis 距离的计算数值不变这一重要性质.并通过实际问题说明了此性质在多元统计分析,特别是在距离判别分析中的作用.

**关键词**：多元统计分析；随机向量；协方差矩阵；Mahalanobis 距离；距离判别分析
**中图分类号**：O212.4 **文献标识码**：A **文章编号**：1672-6510（2008）02-0083-04

# Theorem of Mahalanobis Distance from Multivariate Random Variables Linear Transformation

ZHANG Shao-pu
（College of Science，Tianjin University of Science & Technology，Tianjin 300457，China）

**Abstract**：According to multivariate statistical analysis of the calculation theory of Mahalanobis distance,the nature of its calculation were discussed and studied. It has proved an important nature that the numerical calculation of Mahalanobis distance of multivariate random variables is unchanged after the reversible linear transformation. And it has illustrate through the example that the nature has an effect on multivariate statistical analysis,especially on distance discriminant analysis.

**Keywords**：multivariate statistical analysis；random vector；covariance matrix；Mahalanobis distance；distance discriminantory analysis

随着社会的不断发展与科技的进步,多元统计分析在国内外也得到越来越广泛的应用.在实际中经常根据需要,利用所观测到的数据资料,对研究的对象进行分类或定性.例如在科研实验中,可以根据实验对象所测试的多项指标(如强度、寿命、耗能量、偏差程度等)来判定此实验对象的类型或级别.如果所检查测试的多项指标可以表示为多元的随机向量,则这些判定都可以用多元统计分析中的距离判别分析方法来解决.其中 Mahalanobis 距离是最重要的判别工具之一,且应用范围较广,并得到广泛的重视[1−3].由于随机变量经过线性变换后其 Mahalanobis 距离的计算非常繁杂,难度较大,为了简化繁杂的计算,更有利于实际应用,本文讨论了 Mahalanobis 距离的有关性质,给出了简化计算的方法,进一步完善了距离判别分析的理论与应用的研究.

**收稿日期**：2007−12−19；**修回日期**：2008−03−17
**作者简介**：张绍璞（1954—），男，山东人，副教授.

## 1 随机向量及协方差矩阵的性质

一般把 $n$ 个随机变量 $X_1,X_2,\cdots,X_n$ 的整体称为 $n$ 维随机向量,记为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^T$,记 $E(\boldsymbol{X})=(EX_1,EX_2,\cdots,EX_n)^T$ 为均值向量,均值向量也可简写为 $\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_n)^T$.

设 $\boldsymbol{X},\boldsymbol{Y}$ 为随机向量,$\boldsymbol{A},\boldsymbol{B}$ 为适合运算的常数矩阵,则均值向量运算具有以下性质[4]:

（1）$E(\boldsymbol{AX})=\boldsymbol{A}E(\boldsymbol{X})$;

（2）$E(\boldsymbol{AXB})=\boldsymbol{A}E(\boldsymbol{X})\boldsymbol{B}$;

（3）$E(\boldsymbol{AX}+\boldsymbol{BY})=\boldsymbol{A}E(\boldsymbol{X})+\boldsymbol{B}E(\boldsymbol{X})$.

记 $\boldsymbol{X}$ 的协方差矩阵为

$$\boldsymbol{V}=D(\boldsymbol{X})=E(\boldsymbol{X}-E\boldsymbol{X})(\boldsymbol{X}-E\boldsymbol{X})^T=$$

$$\begin{pmatrix} \mathrm{Cov}(X_1,X_1) & \mathrm{Cov}(X_1,X_2) & \cdots & \mathrm{Cov}(X_1,X_n) \\ \mathrm{Cov}(X_2,X_1) & \mathrm{Cov}(X_2,X_2) & \cdots & \mathrm{Cov}(X_2,X_n) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{Cov}(X_n,X_1) & \mathrm{Cov}(X_n,X_2) & \cdots & \mathrm{Cov}(X_n,X_n) \end{pmatrix}$$

记 $\sigma_{ij}=\mathrm{Cov}(X_i,X_j)$，则上式可简写为

$$\boldsymbol{V}=(\sigma_{ij})_{n\times n}$$

设 $\boldsymbol{A}$ 为常数矩阵，$\boldsymbol{b}$ 为常数向量，则协方差矩阵有以下性质[5]：

（1）$D(\boldsymbol{X})\geqslant 0$，即 $\boldsymbol{X}$ 的协方差矩阵是非负定矩阵；

（2）$D(\boldsymbol{X}+\boldsymbol{b})=D(\boldsymbol{X})$；

（3）$D(\boldsymbol{AX})=\boldsymbol{A}D(\boldsymbol{X})\boldsymbol{A}^T$.

## 2　Mahalanobis 距离及其作用

由于常用的欧氏距离很难处理和协调各随机变量指标（如长度、重量、时间、温度等）不同量纲之间的“大小”比较关系，另外欧氏距离也很难考虑各随机变量之间的相关性.所以在多元统计分析中，由统计学家 Mahalanobis 定义了一种距离，称为 Mahalanobis 距离，简称马氏距离.

**定义**　设 $\boldsymbol{X},\boldsymbol{Y}$ 是服从均值向量为 $\boldsymbol{\mu}$，协方差矩阵为 $n$ 阶方阵 $\boldsymbol{V}$ 的总体 $G$ 中抽取的两个随机向量样品，定义 $\boldsymbol{X},\boldsymbol{Y}$ 之间的马氏距离[6]为

$$D^2(\boldsymbol{X},\boldsymbol{Y})=(\boldsymbol{X}-\boldsymbol{Y})^T\boldsymbol{V}^{-1}(\boldsymbol{X}-\boldsymbol{Y})$$

不难看出当协方差矩阵 $\boldsymbol{V}$ 为单位矩阵 $\boldsymbol{E}$ 时，即 $\boldsymbol{V}^{-1}=\boldsymbol{E}$，则有

$$D^2(\boldsymbol{X},\boldsymbol{Y})=(\boldsymbol{X}-\boldsymbol{Y})^T\boldsymbol{V}^{-1}(\boldsymbol{X}-\boldsymbol{Y})=$$
$$(\boldsymbol{X}-\boldsymbol{Y})^T(\boldsymbol{X}-\boldsymbol{Y})=(x_1-y_1)^2+\cdots+(x_n-y_n)^2$$

即欧氏距离是马氏距离当协方差阵 $\boldsymbol{V}=\boldsymbol{E}$ 时的特例.

关于马氏距离在假设检验与判别分析中许多重要的应用，见有关文献［7—9］.

## 3　随机向量的线性变换及应用

在多元统计分析中，根据需要有时需对随机向量 $\boldsymbol{X}$ 进行随机向量线性变换 $\boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{b}$，其中 $\boldsymbol{A}$ 为可逆方阵，$\boldsymbol{b}$ 为常数向量.则随机向量 $\boldsymbol{Y}$ 是 $\boldsymbol{X}$ 的随机向量线性函数.在向量空间中，向量经过线性变换后，其欧氏距离要发生相应的改变.

下面根据实际问题来讨论关于随机向量经过线性变换后，其欧氏距离的变化和马氏距离值的计算.

**问题**　已知三维随机向量总体 $\boldsymbol{X}=\begin{pmatrix}X_1\\X_2\\X_3\end{pmatrix}$ 的协方差矩阵为 $\boldsymbol{V}_X=\begin{pmatrix}1&1&1\\1&2&1\\1&1&4\end{pmatrix}$，设随机向量函数

$$\boldsymbol{Y}=\begin{pmatrix}Y_1\\Y_2\\Y_3\end{pmatrix}=\begin{pmatrix}4X_1+X_3+1\\2X_2+X_3+2\\X_1+3X_3+3\end{pmatrix}.$$

（1）设 $\boldsymbol{X}$ 的两个样本为 $x_1=\begin{pmatrix}13\\38\\82\end{pmatrix}, x_2=\begin{pmatrix}22\\31\\74\end{pmatrix}$，求对应的样本函数值 $y_1$ 与 $y_2$，并求 $x_1$ 与 $x_2$ 的欧氏距离 $d_1$ 和 $y_1$ 与 $y_2$ 的欧氏距离 $d_2$.

因为 $\boldsymbol{Y}=\begin{pmatrix}4&0&1\\0&2&1\\1&0&3\end{pmatrix}\begin{pmatrix}X_1\\X_2\\X_3\end{pmatrix}+\begin{pmatrix}1\\2\\3\end{pmatrix}$，设 $\boldsymbol{A}=\begin{pmatrix}4&0&1\\0&2&1\\1&0&3\end{pmatrix}$，$\boldsymbol{b}=\begin{pmatrix}1\\2\\3\end{pmatrix}$，则有线性变换 $\boldsymbol{Y}=\boldsymbol{AX}+\boldsymbol{b}$，得

$$y_1=\boldsymbol{A}x_1+\boldsymbol{b}=\begin{pmatrix}4&0&1\\0&2&1\\1&0&3\end{pmatrix}\begin{pmatrix}13\\38\\82\end{pmatrix}+\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}135\\160\\262\end{pmatrix}$$

$$y_2=\boldsymbol{A}x_2+\boldsymbol{b}=\begin{pmatrix}4&0&1\\0&2&1\\1&0&3\end{pmatrix}\begin{pmatrix}22\\31\\74\end{pmatrix}+\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}163\\138\\247\end{pmatrix}$$

$$\text{因为 } x_1-x_2=\begin{pmatrix}13\\38\\82\end{pmatrix}-\begin{pmatrix}22\\31\\74\end{pmatrix}=\begin{pmatrix}-9\\7\\8\end{pmatrix}$$

$$y_1-y_2=\begin{pmatrix}135\\160\\262\end{pmatrix}-\begin{pmatrix}163\\138\\247\end{pmatrix}=\begin{pmatrix}-28\\22\\15\end{pmatrix}$$

故

$$d_1=\sqrt{(-9)^2+7^2+8^2}=\sqrt{194}$$

$$d_2=\sqrt{(-28)^2+22^2+15^2}=\sqrt{1493}$$

显然 $d_1\neq d_2$，即向量经过线性变换后，其欧氏距离要发生相应的改变.

（2）求 $\boldsymbol{Y}$ 的协方差矩阵 $\boldsymbol{V}_Y$.

根据协方差矩阵的性质，得 $\boldsymbol{Y}$ 的协方差矩阵公式为

$\boldsymbol{V}_Y = D(\boldsymbol{Y}) = D(\boldsymbol{AX}+\boldsymbol{b}) = D(\boldsymbol{AX}) = \boldsymbol{A}D(\boldsymbol{X})\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{V}_X\boldsymbol{A}^T$

得

$$\boldsymbol{V}_Y = \begin{pmatrix}4&0&1\\0&2&1\\1&0&3\end{pmatrix}\begin{pmatrix}1&1&1\\1&2&1\\1&1&4\end{pmatrix}\begin{pmatrix}4&0&1\\0&2&0\\1&1&3\end{pmatrix} = \begin{pmatrix}5&5&8\\3&5&6\\4&4&13\end{pmatrix}\begin{pmatrix}4&0&1\\0&2&0\\1&1&3\end{pmatrix} = \begin{pmatrix}28&18&29\\18&16&21\\29&21&43\end{pmatrix}$$

（3）求 $y_1$ 与 $y_2$ 的马氏距离 $D^2(y_1,y_2)$.

因为

$$\boldsymbol{V}_Y^{-1} = \begin{pmatrix}28&18&29\\18&16&21\\29&21&43\end{pmatrix}^{-1} = \begin{pmatrix}\frac{247}{1452}&\frac{-5}{44}&\frac{-43}{726}\\\frac{-5}{44}&\frac{1}{4}&\frac{-1}{22}\\\frac{-43}{726}&\frac{-1}{22}&\frac{31}{363}\end{pmatrix}$$

所以 $y_1$ 与 $y_2$ 的马氏距离

$$D^2(y_1,y_2) = (y_1-y_2)^T\ \ V_Y^{-1}(y_1-y_2) =$$

$$(-28\quad 22\quad 15)\begin{pmatrix}\frac{247}{1452}&\frac{-5}{44}&\frac{-43}{726}\\\frac{-5}{44}&\frac{1}{4}&\frac{-1}{22}\\\frac{-43}{726}&\frac{-1}{22}&\frac{31}{363}\end{pmatrix}\begin{pmatrix}-28\\22\\15\end{pmatrix} =$$

$$\begin{pmatrix}\frac{-11836}{1452}&\frac{352}{44}&\frac{1408}{726}\end{pmatrix}\begin{pmatrix}-28\\22\\15\end{pmatrix} =$$

$$\frac{331408}{1452}+\frac{352}{2}+\frac{21120}{726} = 433\frac{1}{3}.$$

## 4　在随机向量线性变换中马氏距离的性质及其应用

由上述问题的讨论可以看到，计算马氏距离 $D^2(y_1,y_2)$ 需要先求 $y_1$ 与 $y_2$，再求 $\boldsymbol{V}_Y = \boldsymbol{A}\boldsymbol{V}_X\boldsymbol{A}^T$，另外需要解逆矩阵 $\boldsymbol{V}_Y^{-1}$，最后才计算 $D^2(y_1,y_2) = (y_1-y_2)^T\boldsymbol{V}_Y^{-1}(y_1-y_2)$，计算过程较繁，经过研究，可得马氏距离一个重要性质，可简化在线性变换中求 $D^2(y_1,y_2)$ 的计算.

**定理**　设随机向量 $\boldsymbol{X}$ 是服从协方差矩阵为 $\boldsymbol{V}_X$ 的总体，随机向量 $\boldsymbol{X}$ 的线性变换函数 $\boldsymbol{Y} = \boldsymbol{AX}+\boldsymbol{b}$（$\boldsymbol{A}$ 为相应的可逆方阵，$\boldsymbol{b}$ 为常数向量），随机向量 $x_1, x_2$ 为 $\boldsymbol{X}$ 的两个样本，$y_1, y_2$ 为 $x_1, x_2$ 对应的随机向量线性变换函数值.则 $y_1$ 和 $y_2$ 的马氏距离与 $x_1$ 和 $x_2$ 的马氏距离相等.即 $D^2(y_1,y_2) = D^2(x_1,x_2)$.

**证明**：设 $\boldsymbol{V}_Y$ 是 $\boldsymbol{Y} = \boldsymbol{AX}+\boldsymbol{b}$ 的协方差矩阵，则有

$$\boldsymbol{V}_Y = D(\boldsymbol{Y}) = D(\boldsymbol{AX}+\boldsymbol{b}) = D(\boldsymbol{AX}) = \boldsymbol{A}D(\boldsymbol{X})\boldsymbol{A}^T = \boldsymbol{A}\boldsymbol{V}_X\boldsymbol{A}^T$$

$$D^2(y_1,y_2) = (y_1-y_2)^T\boldsymbol{V}_Y^{-1}(y_1-y_2) =$$

$$\left[(\boldsymbol{A}x_1+\boldsymbol{b})-(\boldsymbol{A}x_2+\boldsymbol{b})\right]^T\left(\boldsymbol{A}\boldsymbol{V}_X\boldsymbol{A}^T\right)^{-1}$$

$$\left[(\boldsymbol{A}x_1+\boldsymbol{b})-(\boldsymbol{A}x_2+\boldsymbol{b})\right] =$$

$$\left[\boldsymbol{A}(x_1-x_2)\right]^T\left[\left(\boldsymbol{A}^T\right)^{-1}\boldsymbol{V}_X^{-1}\boldsymbol{A}^{-1}\right]\left[\boldsymbol{A}(x_1-x_2)\right] =$$

$$(x_1-x_2)^T\boldsymbol{A}^T\left(\boldsymbol{A}^T\right)^{-1}\boldsymbol{V}_X^{-1}\boldsymbol{A}^{-1}\boldsymbol{A}(x_1-x_2) =$$

$$(x_1-x_2)^T\boldsymbol{V}_X^{-1}(x_1-x_2) = D^2(x_1,x_2)$$

即　$D^2(y_1,y_2) = D^2(x_1,x_2)$.

利用定理计算前述问题中的马氏距离.

由于　$x_1-x_2 = (-9\quad 7\quad 8)^T$

$$\boldsymbol{V}_X^{-1} = \begin{pmatrix}1&1&1\\1&2&1\\1&1&4\end{pmatrix}^{-1} = \begin{pmatrix}\frac{7}{3}&-1&\frac{-1}{3}\\-1&1&0\\\frac{-1}{3}&0&\frac{1}{3}\end{pmatrix}$$

所以根据定理有

$$D^2(y_1,y_2) = D^2(x_1,x_2) =$$

$$(x_1-x_2)^T\boldsymbol{V}_X^{-1}(x_1-x_2) =$$

$$(-9\quad 7\quad 8)\begin{pmatrix}\frac{7}{3}&-1&\frac{-1}{3}\\-1&1&0\\\frac{-1}{3}&0&\frac{1}{3}\end{pmatrix}\begin{pmatrix}-9\\7\\8\end{pmatrix} =$$

$$(-9\quad 7\quad 8)\begin{pmatrix}-30\frac{2}{3}\\16\\5\frac{2}{3}\end{pmatrix} = 276+112+45\frac{1}{3} = 433\frac{1}{3}$$

通过比较可知，利用定理可以简化在随机变量线性变换中求马氏距离 $D^2(y_1,y_2)$ 的有关计算.

## 5　结　语

多元统计分析是应用非常广泛的学科，其中有许

多问题涉及到马氏距离的计算.本文讨论了在随机变量经过线性变换后对其马氏距离的计算,经过对其性质的有关研究和证明,给出了可以简化马氏距离计算的有关方法,更有利于实际应用.

**参　考　文　献:**


[1] 李　昆,周晓兰. 基于 Mahalanobis 距离的运动意识分类研究[J]. 计算机工程与设计,2007(7):1601—1603.

[2] 李国宏,施鹏飞. 基于次特征值误差补偿和非对称分布的马氏距离改进算法[J]. 电子学报,2007(4):747—750.

[3] 李玉榕,项国波. 一种基于马氏距离的线性判别分析分类算法[J]. 计算机仿真,2006(8):86—88.

[4] 高惠璇. 应用多元统计分析[M]. 北京:北京大学出版社,2005:20—25.

[5] 于秀林. 多元统计分析[M]. 北京:中国统计出版社,1999:13—15.

[6] 王学民. 应用多元分析[M]. 上海:上海财经大学出版社,2004:138—143.

[7] 赵荣军,和向丽. 基于 Excel 的马氏距离计算方法[J]. 物探化探计算技术,2005(4):358—360.

[8] 张润楚. 多元统计分析[M]. 北京:科学出版社,2006:144—160.

[9] 何晓群. 多元统计分析[M]. 北京:中国人民大学出版社,2004:122—136.


(上接第 23 页)

由图可见:陈化 50 min 时观察到未长成的类似球形的不规则粒子,如图 3(a);陈化 70 min 后粒子稍大些,粒子已经显现出球形形貌,如图 3(b);陈化 110 min 后,粒子变为规则的球形,并且平均粒径随着时间的推移而增大,如图 3(c);陈化时间在 130~240 min 范围时,平均粒径基本不变. 这是因为在一定陈化时间内,由于小颗粒表面自由能大于大颗粒,其溶解度较高,所以小颗粒会溶解形成大颗粒,随着陈化时间的延长,粒径逐渐均一化,粒子也就不再进一步生长了. 因此观测到了粒径基本不变较均匀的粒子.

## 3　结　论

对于利用尿素共沉淀法制备 $Y_2(CrO_4)_3$ 均匀胶体粒子,关键在于调节金属离子和尿素的浓度配比,在适当的金属离子浓度范围内($2.0\times10^{-4}\sim4.0\times10^{-3}$ mol·$L^{-1}$)增加尿素的浓度(0.9~1.6 mol·$L^{-1}$)就可以调控粒子形态的转变.

**参　考　文　献:**


[1] 王巧巧,官月平,郎宇琪,等. 尿素水解法制备球形磁性 $Al_2O_3$ 复合材料[J]. 化学通报,2007(2):34—38.

[2] 刘　运,苗鸿雁. ZnS 纳米粒子的制备与表征[J]. 电子元件与材料,2006(7):69—71.

[3] Daniel S,Mufit A. Preparation of spherical,monosized $Y_2O_3$ precursor particle[J]. Journal of olloid and Interface Science,1988,122(1):47—59.

[4] Matijevic E. Colloid science of ceramic powders[J]. Pure and Applied Chemistry,1988,60(10):1479—1491.

[5] 郭明林,张玉亭. 铜(Ⅱ)-铬(Ⅵ)复合均匀胶体粒子的制备[J]. 物理化学学报,1998,14(10):877—880.

[6] Kratohvil S,Matijevic E. Preparation of copper compounds of different compositions and particle morphologied[J]. Journal of Materials Research,1991,6(4):766—777.

[7] 孙晓明. 低维功能纳米材料的液相合成、表征与性能研究[D]. 北京:清华大学化学学院,2005.