第二章 多元正态分布

首都师范大学
数学科学学院
统计学系
Department of Statistics

2.1 随机向量

- §2.1.1 一元分布
- §2.1.2 多元分布
- §2.1.3 数字特征
- §2.1.4 欧氏距离和马氏距离
- §2.1.5 随机向量的变换
- *§2.1.6 特征函数

§2.1.2 多元分布

- 一、多元概率分布
- 二、两个常用的离散型多元分布
- 三、多元概率密度函数
- 四、边缘分布
- 五、条件分布
- 六、独立性

一、多元概率分布

- 若向量X的分量都是随机变量，则X为随机向量。
- 随机向量 $X=\left(X_1,X_2,\cdots,X_p\right)'$ 的分布函数：

$$F\left(x_1,x_2,\cdots,x_p\right)=P\left(X_1\le x_1,X_2\le x_2,\cdots,X_p\le x_p\right)$$

三、多元概率密度函数

- 多元连续型随机向量的情形：

$$F(x_1,\cdots,x_p)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_p} f(t_1,\cdots,t_p)\,\mathrm{d}t_1\cdots\mathrm{d}t_p$$

$$f(x_1,\cdots,x_p)=\frac{\partial^p}{\partial x_1\cdots\partial x_p}F(x_1,\cdots,x_p)$$

- 多元密度$f(x_1,\cdots,x_p)$的性质：

(1) $f(x_1,\cdots,x_p)\geq 0$, 对一切实数$x_1,\cdots,x_p$;

(2) $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1,\cdots,x_p)\,\mathrm{d}x_1\cdots\mathrm{d}x_p=1$。

四、边缘分布

- 设X是p维随机向量，由它的$q(<p)$ 个分量组成的向量$X_{(1)}$的分布称为X的关于$X_{(1)}$的边缘分布。
- 设 $X_{(1)}=\left(x_1,\cdots,x_q\right)'$，则对连续型的分布，有

$$f_{(1)}(x_1,\cdots,x_q)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,\cdots,x_p)\,\mathrm{d}x_{q+1}\cdots\mathrm{d}x_p$$

五、条件分布

- 设$X=\left(X_1,\cdots,X_p\right)'$是$p$维连续型的随机向量，在给定 $X_{(2)}=x_{(2)}=\left(x_{q+1},\cdots,x_p\right)'\left(f_{(2)}\left(x_{(2)}\right)>0\right)$的条件下，$X_{(1)}=\left(X_1,\cdots,X_q\right)'$ 的条件密度定义为

$$f\left(x_1,\cdots,x_q \mid x_{q+1},\cdots,x_p\right)=\frac{f\left(x_1,\cdots,x_p\right)}{f_{(2)}\left(x_{q+1},\cdots,x_p\right)}$$

或表达为

$$f\left(\boldsymbol{x}_{(1)} \mid \boldsymbol{x}_{(2)}\right)=\frac{f\left(\boldsymbol{x}\right)}{f_{(2)}\left(\boldsymbol{x}_{(2)}\right)}$$

六、独立性

- 两个连续型随机向量的独立

$$f(x,y)=f_x(x)\cdot f_y(y)$$

- n个连续型随机向量的独立

$$f(x_1,\cdots,x_n)=f_1(x_1)\cdots f_n(x_n)$$

- 随机向量之间的取值互不影响，则认为它们是相互独立的。

§2.1.3 数字特征

- 一、数学期望（均值）
- 二、协方差矩阵
- 三、相关矩阵
- *四、广义方差

一、数学期望（均值）

- 随机向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_p)'$ 的数学期望

$$E(\boldsymbol{x})=\left[E(x_1),E(x_2),\cdots,E(x_p)\right]'$$

记为 $\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_p)'$ 。

- 随机矩阵 $\boldsymbol{X}=(x_{ij})$ 的数学期望

$$E(\boldsymbol{X})=\left(E(x_{ij})\right)=\begin{pmatrix} E(x_{11}) & E(x_{12}) & \cdots & E(x_{1q}) \\ E(x_{21}) & E(x_{22}) & \cdots & E(x_{2q}) \\ \vdots & \vdots & & \vdots \\ E(x_{p1}) & E(x_{p2}) & \cdots & E(x_{pq}) \end{pmatrix}$$

随机矩阵$\boldsymbol{X}$数学期望的性质

- (1) 设a为常数，则

$$E(a\boldsymbol{X})=aE(\boldsymbol{X})$$

- (2) 设$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$为常数矩阵，则

$$E(\boldsymbol{AXB}+\boldsymbol{C})=\boldsymbol{A}E(\boldsymbol{X})\boldsymbol{B}+\boldsymbol{C}$$

- 特别地，对于随机向量$\boldsymbol{x}$，有

$$E(\boldsymbol{Ax})=\boldsymbol{A}E(\boldsymbol{x})$$

- (3) 设$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$为$n$个同阶的随机矩阵，则

$$E(\boldsymbol{X}_1+\boldsymbol{X}_2+\cdots+\boldsymbol{X}_n)=E(\boldsymbol{X}_1)+E(\boldsymbol{X}_2)+\cdots+E(\boldsymbol{X}_n)$$

二、协方差矩阵

- 协方差定义为

$$\mathrm{Cov}(x,y)=E\left[x-E(x)\right]\left[y-E(y)\right]$$

- 若 Cov(x,y)=0，则称 x 和 y 不相关。
- 两个独立随机变量一定不相关，但两个不相关的随机变量未必独立。
- 当 x=y 时，协方差即为方差：

$$\mathrm{Cov}(x,x)=V(x)$$

- $\boldsymbol{x}=\left(x_1,x_2,\cdots,x_p\right)'$ 和 $\boldsymbol{y}=\left(y_1,y_2,\cdots,y_q\right)'$ 的协方差矩阵（简称协差阵）定义为

$$\operatorname{Cov}(\boldsymbol{x},\boldsymbol{y})=\begin{pmatrix}\operatorname{Cov}(x_1,y_1) & \operatorname{Cov}(x_1,y_2) & \cdots & \operatorname{Cov}(x_1,y_q)\\ \operatorname{Cov}(x_2,y_1) & \operatorname{Cov}(x_2,y_2) & \cdots & \operatorname{Cov}(x_2,y_q)\\ \vdots & \vdots & & \vdots\\ \operatorname{Cov}(x_p,y_1) & \operatorname{Cov}(x_p,y_2) & \cdots & \operatorname{Cov}(x_p,y_q)\end{pmatrix}$$

$$=E\begin{pmatrix}x_1-E(x_1)\\ \vdots\\ x_p-E(x_p)\end{pmatrix}\left(y_1-E(y_1),\cdots,y_q-E(y_q)\right)$$

$$=E\left[\boldsymbol{x}-E(\boldsymbol{x})\right]\left[\boldsymbol{y}-E(\boldsymbol{y})\right]'$$

- $\mathrm{Cov}(\boldsymbol{x},\boldsymbol{y})=\left[\mathrm{Cov}(\boldsymbol{y},\boldsymbol{x})\right]'$
- 若Cov($\boldsymbol{x}$,$\boldsymbol{y}$)=**0**，则称$\boldsymbol{x}$和$\boldsymbol{y}$不相关。
- 两个独立随机向量必然不相关，反之则不然。
- $\boldsymbol{x}$=$\boldsymbol{y}$时的协差阵Cov($\boldsymbol{x}$,$\boldsymbol{x}$)称为$\boldsymbol{x}$的协差阵，记

$$V(\boldsymbol{x})=E\left[\boldsymbol{x}-E(\boldsymbol{x})\right]\left[\boldsymbol{x}-E(\boldsymbol{x})\right]'$$

$$=\begin{pmatrix} V(x_1) & \mathrm{Cov}(x_1,x_2) & \cdots & \mathrm{Cov}(x_1,x_p) \\ \mathrm{Cov}(x_2,x_1) & V(x_2) & \cdots & \mathrm{Cov}(x_2,x_p) \\ \vdots & \vdots & & \vdots \\ \mathrm{Cov}(x_p,x_1) & \mathrm{Cov}(x_p,x_2) & \cdots & V(x_p) \end{pmatrix}$$

V($\boldsymbol{x}$)亦记作 $\boldsymbol{\Sigma}=(\sigma_{ij})$，其中 $\sigma_{ij}=\mathrm{Cov}(x_i,x_j)$。

- $\boldsymbol{\Sigma}$包含了$\boldsymbol{x}$各分量的方差，也包含了每两个分量之间的协方差，是对称矩阵。

- 例1. 随机向量一分为二后，协差阵分为四块：

$$V\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}=\begin{pmatrix}V(\boldsymbol{x}) & \mathrm{Cov}(\boldsymbol{x},\boldsymbol{y})\\ \mathrm{Cov}(\boldsymbol{y},\boldsymbol{x}) & V(\boldsymbol{y})\end{pmatrix}$$

其中，对角块为子向量的协差阵，非对角块为两个子向量之间的协差阵。

协差阵的性质

- （1）非负定性，即 $\boldsymbol{\Sigma} \geq 0$。

➢ 推论　若$|\boldsymbol{\Sigma}| \neq 0$，则 $\boldsymbol{\Sigma} > 0$。

- （2）设$\boldsymbol{A}$为常矩阵，$\boldsymbol{b}$为常向量，则

$$V(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}) = \boldsymbol{A}V(\boldsymbol{x})\boldsymbol{A}'$$

➢ 当p=1时，则有：

$$V(ax+b) = a^2V(x)$$

- 例2.　$|\boldsymbol{\Sigma}| = 0 \Leftrightarrow \boldsymbol{x}$的分量之间存在线性关系（以概率1）。

- 实际中，如某一指标是其他一些指标的汇总值，通常可删去“多余”指标来确保$|\boldsymbol{\Sigma}|\neq 0$。假定 $\boldsymbol{\Sigma}>0$并不失一般性，以保证 $\boldsymbol{\Sigma}^{-1}$存在。

- 例3. 设随机向量 $x=(x_1,x_2,x_3)'$ 的数学期望和协方差矩阵分别为

$$\boldsymbol{\mu}=\begin{pmatrix}5\\-2\\7\end{pmatrix} \quad \text{和} \quad \boldsymbol{\Sigma}=\begin{pmatrix}4 & 1 & 2\\1 & 9 & -3\\2 & -3 & 25\end{pmatrix}$$

令 $y_1=2x_1-x_2+4x_3$，$y_2=x_2-x_3$，$y_3=x_1+3x_2-2x_3$，试求 $y=(y_1,y_2,y_3)'$ 的数学期望和协方差矩阵。

- （3）设$\boldsymbol{A}$和$\boldsymbol{B}$为常数矩阵，则

$$\mathrm{Cov}\left(\boldsymbol{A}\boldsymbol{x},\boldsymbol{B}\boldsymbol{y}\right)=\boldsymbol{A}\,\mathrm{Cov}\left(\boldsymbol{x},\boldsymbol{y}\right)\boldsymbol{B}'$$

- （4）设 $\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_n$和$\boldsymbol{B}_1,\boldsymbol{B}_2,\cdots,\boldsymbol{B}_m$为常数矩阵，则

$$\mathrm{Cov}\left(\sum_{i=1}^{n}\boldsymbol{A}_i\boldsymbol{x}_i,\sum_{j=1}^{m}\boldsymbol{B}_j\boldsymbol{y}_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{m}\boldsymbol{A}_i\,\mathrm{Cov}\left(\boldsymbol{x}_i,\boldsymbol{y}_j\right)\boldsymbol{B}_j'$$

➢ 推论 $$\operatorname{Cov}\left(\sum_{i=1}^{n} x_i, \sum_{j=1}^{m} y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \operatorname{Cov}\left(x_i, y_j\right)$$

➢ 证明 （先证推论）

$$\operatorname{Cov}\left(\sum_{i=1}^{n} x_i, \sum_{j=1}^{m} y_j\right)$$

$$= E\left[\sum_{i=1}^{n} x_i - E\left(\sum_{i=1}^{n} x_i\right)\right]\left[\sum_{j=1}^{m} y_j - E\left(\sum_{j=1}^{m} y_j\right)\right]'$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} E\left[x_i - E\left(x_i\right)\right]\left[y_j - E\left(y_j\right)\right]'$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} \operatorname{Cov}\left(x_i, y_j\right)$$

$$\operatorname{Cov}\left(\sum_{i=1}^{n} \boldsymbol{A}_i \boldsymbol{x}_i, \sum_{j=1}^{m} \boldsymbol{B}_j \boldsymbol{y}_j\right)=\sum_{i=1}^{n} \sum_{j=1}^{m} \operatorname{Cov}\left(\boldsymbol{A}_i \boldsymbol{x}_i, \boldsymbol{B}_j \boldsymbol{y}_j\right)$$

$$=\sum_{i=1}^{n} \sum_{j=1}^{m} \boldsymbol{A}_i \operatorname{Cov}\left(\boldsymbol{x}_i, \boldsymbol{y}_j\right) \boldsymbol{B}_j'$$

- （5）设$k_1,k_2,\cdots,k_n$是n个常数，$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$是$n$个相互独立的$p$维随机向量，则

$$V\left(\sum_{i=1}^{n} k_i \boldsymbol{x}_i\right)=\sum_{i=1}^{n} k_i^2 V\left(\boldsymbol{x}_i\right)$$

三、相关矩阵

- 随机变量x和y的相关系数定义：

$$\rho=\rho(x,y)=\frac{\mathrm{Cov}(x,y)}{\sqrt{V(x)V(y)}}$$

- $\boldsymbol{x}=(x_1,x_2,\cdots,x_p)'$和$\boldsymbol{y}=(y_1,y_2,\cdots,y_q)'$的相关阵定义：

$$\rho(\boldsymbol{x},\boldsymbol{y})=\begin{pmatrix} \mathrm{Cov}(x_1,y_1) & \mathrm{Cov}(x_1,y_2) & \cdots & \mathrm{Cov}(x_1,y_q) \\ \mathrm{Cov}(x_2,y_1) & \mathrm{Cov}(x_2,y_2) & \cdots & \mathrm{Cov}(x_2,y_q) \\ \vdots & \vdots & & \vdots \\ \mathrm{Cov}(x_p,y_1) & \mathrm{Cov}(x_p,y_2) & \cdots & \mathrm{Cov}(x_p,y_q) \end{pmatrix}$$

- 若 $\rho(\boldsymbol{x},\boldsymbol{y})=0$，则$\boldsymbol{x}$和$\boldsymbol{y}$不相关。
- $\boldsymbol{x}=\boldsymbol{y}$时，$\rho(\boldsymbol{x},\boldsymbol{x})$称为$\boldsymbol{x}$的相关阵，记作 $\boldsymbol{R}=(\rho_{ij})$，这里 $\rho_{ij}=\rho(x_i,x_j)$，$\rho_{ii}=1$。即

$$\boldsymbol{R}=\begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{pmatrix}$$

- $\boldsymbol{R}=(\rho_{ij})$和 $\boldsymbol{\Sigma}=(\sigma_{ij})$之间有关系：

$$\boldsymbol{R}=\boldsymbol{D}^{-1}\boldsymbol{\Sigma}\boldsymbol{D}^{-1}$$

其中 $\boldsymbol{D} = \mathrm{diag}\left(\sqrt{\sigma_{11}}, \sqrt{\sigma_{22}}, \cdots, \sqrt{\sigma_{pp}}\right)$

$\boldsymbol{R}$和 $\boldsymbol{\Sigma}$的相应元素之间的关系式：

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$$

标准化变换

- 处理数据时，常因变量单位不同而需要变换，最常用的标准化变换是

$$x_i^* = \frac{x_i - \mu_i}{\sqrt{\sigma_{ii}}}, \quad i = 1, 2, \cdots, p$$

- 记 $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_p^*)'$，于是

$$E\left(\boldsymbol{x}^*\right) = \mathbf{0}, \quad V\left(\boldsymbol{x}^*\right) = \boldsymbol{R}$$

即标准化后的协差阵正好是原始向量的相关阵。相关阵$\boldsymbol{R}$也是一个非负定阵。

§2.1.4 欧氏距离和马氏距离

- 一、欧氏距离
- 二、马氏距离

一、欧氏距离

- $\boldsymbol{x}=\left(x_1,x_2,\cdots,x_p\right)'$和$\boldsymbol{y}=\left(y_1,y_2,\cdots,y_p\right)'$

  之间欧氏距离:

$$d\left(\boldsymbol{x},\boldsymbol{y}\right)=\sqrt{\left(x_1-y_1\right)^2+\left(x_2-y_2\right)^2+\cdots+\left(x_p-y_p\right)^2}$$

- 平方欧氏距离:

$$d^2\left(\boldsymbol{x},\boldsymbol{y}\right)=\left(x_1-y_1\right)^2+\left(x_2-y_2\right)^2+\cdots+\left(x_p-y_p\right)^2$$

$$=\left(\boldsymbol{x}-\boldsymbol{y}\right)'\left(\boldsymbol{x}-\boldsymbol{y}\right)$$

- $\boldsymbol{x}=\left(x_1,x_2,\cdots,x_p\right)'$ 到总体 π 的平方欧氏距离定义:

$$d^2\left(\boldsymbol{x},\pi\right)=\left(\boldsymbol{x}-\boldsymbol{\mu}\right)'\left(\boldsymbol{x}-\boldsymbol{\mu}\right)$$

$$=\left(x_1-\mu_1\right)^2+\left(x_2-\mu_2\right)^2+\cdots+\left(x_p-\mu_p\right)^2$$

平均大小 $E\left(x_1-\mu_1\right)^2$ $E\left(x_2-\mu_2\right)^2$ … $E\left(x_p-\mu_p\right)^2$

等于 $V\left(x_1\right)$ $V\left(x_2\right)$ … $V\left(x_p\right)$

不适合直接使用欧氏距离的例子

- 下面是各国家和地区男子径赛记录的数据（1984年）：

国家和地区	100米（秒）	200米（秒）	400米（秒）	800米（分）	1500米（分）	5000米（分）	10000米（分）	马拉松（分）
阿根廷	10.39	20.81	46.84	1.81	3.7	14.04	29.36	137.72
澳大利亚	10.31	20.06	44.84	1.74	3.57	13.28	27.66	128.3
奥地利	10.44	20.81	46.82	1.79	3.6	13.26	27.72	135.9
比利时	10.34	20.68	45.04	1.73	3.6	13.22	27.45	129.95
百慕大	10.28	20.58	45.91	1.8	3.75	14.68	30.55	146.62
巴西	10.22	20.43	45.21	1.73	3.66	13.62	28.62	133.13
缅甸	10.64	21.52	48.3	1.8	3.85	14.45	30.28	139.95
加拿大	10.17	20.22	45.68	1.76	3.63	13.55	28.09	130.15
智利	10.34	20.8	46.2	1.79	3.71	13.61	29.3	134.03
中国	10.51	21.04	47.3	1.81	3.73	13.9	29.13	133.53
哥伦比亚	10.43	21.05	46.1	1.82	3.74	13.49	27.88	131.35
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

一、欧氏距离

- 如果单位不全相同，欧氏距离一般没意义。
- 即使单位相同，但如果各分量的变异性差异很大，则变异性大的分量在欧氏距离的平方和中起决定性作用，变异性小的分量却几乎不起什么作用。
- 应用中，常须作标准化变换，再计算欧氏距离。
- 令 $$x_i^* = \frac{x_i - \mu_i}{\sqrt{\sigma_{ii}}}, i = 1, \cdots, p, \quad \boldsymbol{x}^* = \left(x_1^*, \cdots, x_p^*\right)'$$

则 $$d^2\left(\boldsymbol{x}^*, \pi\right) = \boldsymbol{x}^{*\prime}\boldsymbol{x}^* = x_1^{*2} + \cdots + x_p^{*2}$$

- 由于 $E\left(x_i^{*2}\right)=V\left(x_i^{*}\right)=1, i=1,2,\cdots,p$，故平方和 $\left(x_1-\mu_1\right)^2+\cdots+\left(x_p-\mu_p\right)^2$ 中各项平均取值均为1，各分量起的平均作用一样。

- 欧氏距离经变量的标准化之后能够消除各变量的单位或方差差异的影响，但不能消除变量之间相关性的影响，有时用欧氏距离不合适。为此引入由印度著名统计学家马哈拉诺比斯（Mahalanobis，1936年）提出的“马氏距离”的概念。

二、马氏距离

- $\boldsymbol{x}=\left(x_1,x_2,\cdots,x_p\right)'$和$\boldsymbol{y}=\left(y_1,y_2,\cdots,y_p\right)'$之间的平方马氏距离定义:

$$d^2(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}-\boldsymbol{y})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{y})$$

- $\boldsymbol{x}=\left(x_1,x_2,\cdots,x_p\right)'$ 到总体 π 的平方马氏距离定义:

$$d^2(\boldsymbol{x},\pi)=(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$$

- 特点（1） 马氏距离不受变量单位影响。

比例单位变换

- 如$\boldsymbol{x}$的分量是长度、重量、速度、费用和用时等，则变量的单位变换可表达为

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} c_1 x_1 \\ c_2 x_2 \\ \vdots \\ c_p x_p \end{pmatrix} = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_p \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x}$$

其中 $\boldsymbol{C} = \mathrm{diag}\left(c_1, c_2, \cdots, c_p\right),\ c_i > 0,\ i = 1, 2, \cdots, p$

带有常数项的单位变换

- 例4. 摄氏温度与华氏温度的换算公式：

  $F=(C\times 9/5)+32$， $C=(F-32)\times 5/9$

  其中F：华氏温度，C：摄氏温度。

$$
\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} c_1 x_1 + b_1 \\ c_2 x_2 + b_2 \\ \vdots \\ c_p x_p + b_p \end{pmatrix} = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_p \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}
$$

$$
\boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{b}
$$

- 证明 x_1,x_2经单位变换后为y_1,y_2，即有

$$y_1 = Cx_1 + b, \quad y_2 = Cx_2 + b$$

$$\Sigma_y = C\Sigma_x C'$$

$$\begin{aligned}
&(y_1 - y_2)' \Sigma_y^{-1} (y_1 - y_2) \\
&= \left[(Cx_1 + b) - (Cx_2 + b)\right]' (C\Sigma_x C')^{-1} \left[(Cx_1 + b) - (Cx_2 + b)\right] \\
&= (x_1 - x_2)' C'C'^{-1} \Sigma_x^{-1} C^{-1} C (x_1 - x_2) \\
&= (x_1 - x_2)' \Sigma_x^{-1} (x_1 - x_2)
\end{aligned}$$

- **特点（2）**马氏距离是 $\boldsymbol{x}$和$\boldsymbol{y}$经“标准化”之后的欧氏距离，即

$$d^2(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x}^*-\boldsymbol{y}^*)'(\boldsymbol{x}^*-\boldsymbol{y}^*)$$

其中 $\boldsymbol{x}^*=\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{x}-\boldsymbol{\mu}),\boldsymbol{y}^*=\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{y}-\boldsymbol{\mu})$ ，

它们的均值皆为**0**，协差阵皆为单位阵$\boldsymbol{I}$。

- **特点（3）** 若 $\boldsymbol{\Sigma} = \mathrm{diag}\left(\sigma_{11}, \sigma_{22}, \cdots, \sigma_{pp}\right)$，则

$$d^2(\boldsymbol{x}, \boldsymbol{y}) = \frac{\left(x_1 - y_1\right)^2}{\sigma_{11}} + \frac{\left(x_2 - y_2\right)^2}{\sigma_{22}} + \cdots + \frac{\left(x_p - y_p\right)^2}{\sigma_{pp}}$$

即当各分量不相关时马氏距离即为各分量经标准化后的欧氏距离。