第四讲 多元正态分布

2023.10.8

正态 = 线性

1. 去相关化

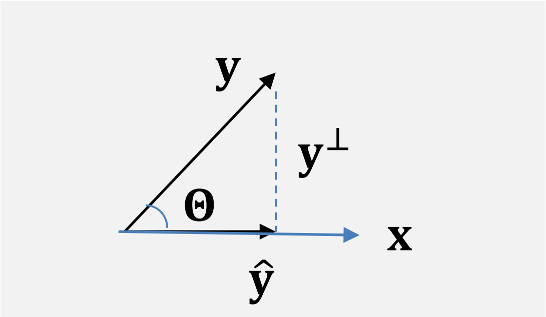
两个随机向量: $\mathbf{y}_{q\times 1}$, $\mathbf{x}_{p\times 1}$ ， $\mathrm{var}\begin{pmatrix}\mathbf{y}\\ \mathbf{x}\end{pmatrix} = \begin{pmatrix}\Sigma_{\mathbf{yy}} & \Sigma_{\mathbf{yx}}\\ \Sigma_{\mathbf{xy}} & \Sigma_{\mathbf{xx}}\end{pmatrix}$

- 去相关化: $\mathbf{y}^{\perp} = \mathbf{y} - \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x} \triangleq \mathbf{y} - \hat{\mathbf{y}}$,

$$\mathrm{var}\begin{pmatrix}\mathbf{y}^{\perp}\\ \mathbf{x}\end{pmatrix} = \begin{pmatrix}\Sigma_{\mathbf{yy\bullet x}} & 0\\ 0 & \Sigma_{\mathbf{xx}}\end{pmatrix}$$

  $\mathbf{y}^{\perp}$与$\mathbf{x}$不相关，与$\hat{\mathbf{y}} = \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$也不相关

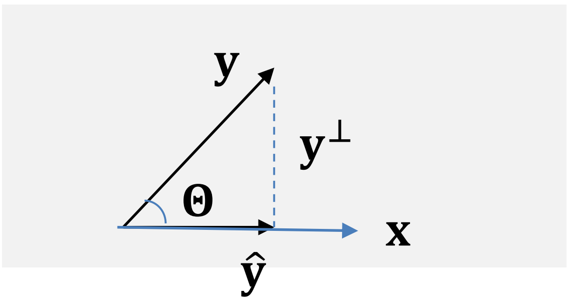
- 正交分解/直和: $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{y}^{\perp}$，$\hat{\mathbf{y}} \perp \mathbf{y}^{\perp}$ (不相关)



Recap

2. 典则系数/决定系数

正交分解 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{y}^{\perp}$两边同时求方差



- 方差分解：$\mathrm{var}(\mathbf{y}) = \mathrm{var}(\hat{\mathbf{y}}) + \mathrm{var}(\mathbf{y}^{\perp})$

$$\mathrm{var}(\hat{\mathbf{y}}) = \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}},\quad \mathrm{var}(\mathbf{y}^{\perp}) = \Sigma_{\mathbf{yy\bullet x}}$$

方差矩阵理解成“长度平方”

- $\mathrm{var}(\hat{\mathbf{y}})$相对于$\mathrm{var}(\mathbf{y})$的大小，代表$\mathbf{y},\mathbf{x}$的相关性(夹角)：

$$\Phi = [\mathrm{var}(\mathbf{y})]^{-1/2}\mathrm{var}(\hat{\mathbf{y}})[\mathrm{var}(\mathbf{y})]^{-1/2}$$

$$= \Sigma_{\mathbf{yy}}^{-1/2}(\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}})\Sigma_{\mathbf{yy}}^{-1/2}$$

形式上将Φ理解成
“$\mathrm{var}(\hat{\mathbf{y}}) \div \mathrm{var}(\mathbf{y}) = \cos^2(\boldsymbol{\Theta})$”

Recap

$$\Phi = \Sigma_{\mathbf{yy}}^{-1/2}(\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}})\Sigma_{\mathbf{yy}}^{-1/2}$$

Φ是$q \times q$矩阵，是相关系数的推广，可理解成“相关系数平方”。

对比Pearson相关系数的平方：

$$\mathrm{var}\begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} \sigma_{yy} & \sigma_{yx} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix}$$

$$\rho_{xy}^2 = \frac{\sigma_{xy}^2}{\sigma_{xx}\sigma_{yy}} = \frac{\sigma_{yx}\sigma_{xx}^{-1}\sigma_{xy}}{\sigma_{yy}}$$

- 第一典则相关系数：$\sqrt{\lambda_{max}(\Phi)}$
- 决定系数：当$q = 1$时，$\Phi = \Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}/\Sigma_{yy} \in [0,1]$ 称为决定系数，

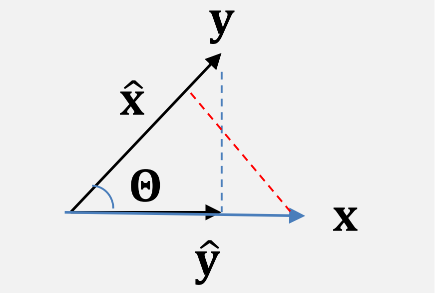
  （$\sqrt{\Phi}$为第一/唯一典则相关系数）

  特别地，若$p = q = 1$,即x, y都是一元随机变量，$\Phi = \rho_{xy}^2$

注：也可将$\mathbf{x}$向$\mathbf{y}$投影（右图），$\hat{\mathbf{x}} = \Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}\mathbf{y}$

$$\Psi = [\mathrm{var}(\mathbf{x})]^{-1/2}\mathrm{var}(\hat{\mathbf{x}})[\mathrm{var}(\mathbf{x})]^{-1/2}$$

$$= \Sigma_{\mathbf{xx}}^{-1/2}(\Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}\Sigma_{\mathbf{yx}})\Sigma_{\mathbf{xx}}^{-1/2}，\lambda_{max}(\Psi) = \lambda_{max}(\Phi)$$



3. 偏相关系数

目标：控制干扰因素$\mathbf{z}_{r\times 1}$条件下，研究随机变量 x, y的相关性。 假设

$$\mathrm{var}\begin{pmatrix} x \\ y \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{x\mathbf{z}} \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{y\mathbf{z}} \\ \Sigma_{\mathbf{z}x} & \Sigma_{\mathbf{z}y} & \Sigma_{\mathbf{zz}} \end{pmatrix},$$

高亮部分(2阶方阵)
是x, y的协方差矩阵

- 去相关化: $x^{\perp} = x - \Sigma_{x\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z}$, $y^{\perp} = y - \Sigma_{y\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z}$,

$$\mathrm{var}\begin{pmatrix} x^{\perp} \\ y^{\perp} \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \Sigma_{xx\bullet\mathbf{z}} & \Sigma_{xy\bullet\mathbf{z}} & 0 \\ \Sigma_{yx\bullet\mathbf{z}} & \Sigma_{yy\bullet\mathbf{z}} & 0 \\ 0 & 0 & \Sigma_{\mathbf{zz}} \end{pmatrix}$$

高亮部分(2阶方阵)
是消除z的影响后，
$x^{\perp}, y^{\perp}$的协方差矩阵

- x、y的偏相关系数

$$\rho_{xy\bullet\mathbf{z}} = \frac{\Sigma_{xy\bullet\mathbf{z}}}{\sqrt{\Sigma_{xx\bullet\mathbf{z}}}\sqrt{\Sigma_{yy\bullet\mathbf{z}}}}$$

一般情形（典则 + 偏相关）

将前两页方法拓展到三个随机向量：$\mathbf{x}_{p\times 1}$，$\mathbf{y}_{q\times 1}$ ，$\mathbf{z}_{r\times 1}$，

控制$\mathbf{z}$条件下研究$\mathbf{x}$， $\mathbf{y}$之间的相关性

$$\mathrm{var}\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\\ \mathbf{z}\end{pmatrix}=\begin{pmatrix}\Sigma_{\mathbf{xx}} & \Sigma_{\mathbf{xy}} & \Sigma_{\mathbf{xz}}\\ \Sigma_{\mathbf{yx}} & \Sigma_{\mathbf{yy}} & \Sigma_{\mathbf{yz}}\\ \Sigma_{\mathbf{zx}} & \Sigma_{\mathbf{zy}} & \Sigma_{\mathbf{zz}}\end{pmatrix},$$

高亮部分($p+q$阶方阵)是$(\mathbf{x},\mathbf{y})$的协方差矩阵

- 去相关化: $\mathbf{x}^{\perp}=\mathbf{x}-\Sigma_{\mathbf{xz}}\Sigma_{zz}^{-1}\mathbf{z}$，$\mathbf{y}^{\perp}=\mathbf{y}-\Sigma_{\mathbf{yz}}\Sigma_{zz}^{-1}\mathbf{z}$，

$$\mathrm{var}\begin{pmatrix}\mathbf{x}^{\perp}\\ \mathbf{y}^{\perp}\\ \mathbf{z}\end{pmatrix}=\begin{pmatrix}\Sigma_{\mathbf{xx\cdot z}} & \Sigma_{\mathbf{xy\cdot z}} & 0\\ \Sigma_{\mathbf{yx\cdot z}} & \Sigma_{\mathbf{yy\cdot z}} & 0\\ 0 & 0 & \Sigma_{\mathbf{zz}}\end{pmatrix}$$

高亮部分($p+q$阶方阵)是消除$\mathbf{z}$后，$(\mathbf{x}^{\perp},\mathbf{y}^{\perp})$的协方差矩阵

- 令 $\Phi_{\mathbf{xy\cdot z}}=\Sigma_{\mathbf{yy\cdot z}}^{-1/2}(\Sigma_{\mathbf{yx\cdot z}}\Sigma_{\mathbf{xx\cdot z}}^{-1}\Sigma_{\mathbf{xy\cdot z}})\Sigma_{\mathbf{yy\cdot z}}^{-1/2}$，

代表控制$\mathbf{z}$之后，随机向量$\mathbf{x}$和$\mathbf{y}$的相关性（对应Φ）。

当$q>1$时， $\Phi_{\mathbf{xy\cdot z}}$是$q\times q$ 矩阵.

$\Phi_{\mathbf{xy\bullet z}} = \Sigma_{\mathbf{yy\bullet z}}^{-1/2}(\Sigma_{\mathbf{yx\bullet z}}\Sigma_{\mathbf{xx\bullet z}}^{-1}\Sigma_{\mathbf{xy\bullet z}})\Sigma_{\mathbf{yy\bullet z}}^{-1/2}$，是偏相关系数的推广，可理解成“偏相关系数的平方”。

❖ 第一典则偏相关系数：$\sqrt{\lambda_{max}(\Phi_{\mathbf{xy\bullet z}})}$

度量了控制**z**条件下，随机向量**x**，**y**之间相关性大小（介于0，1之间）。

❖ $q = 1$时，$\Phi_{\mathbf{x}y\bullet\mathbf{z}} = \Sigma_{y\mathbf{x}\bullet\mathbf{z}}\Sigma_{\mathbf{xx\bullet z}}^{-1}\Sigma_{\mathbf{x}y\bullet\mathbf{z}}/\Sigma_{yy\bullet\mathbf{z}} \in [0,1]$ 称为偏决定系数，

特别地，若$p = q = 1$,即x, y都是一元随机变量，$\Phi_{xy\bullet\mathbf{z}} = \rho_{xy\bullet\mathbf{z}}^2$

如何检验？<u>大概</u>可以如下检验：

$$nr^2 > \chi^2_{|p-q|+1}(\alpha) \text{ 时否定原假设。}$$

这里r^2代表典则相关系数的平方、决定系数、典则偏相关系数的平方或偏决定系数。

注意后两者（红字）是我们自己的命名，并不一定通用，你可以调查文献中有没有这些概念以及是如何命名的。

例1. 140 个七年级学生进行了4项测试，

y_1 = arithmetic power, y_2 = arithmetic speed,

x_1 = reading power, x_2 = reading speed,

相关系数矩阵如下，我们关心的是（1）阅读(x_1, x_2)与数学(y_1, y_2)之间的相关性，（2）数学能力y_1与(x_1, x_2)的相关性,（3）y_1, y_2的偏相关系数。

$$
\Sigma = \begin{array}{c} \\ y_1 \\ y_2 \\ x_1 \\ x_2 \end{array}
\begin{array}{c}
\begin{array}{cccc} y_1 & y_2 & x_1 & x_2 \end{array} \\
\begin{pmatrix}
1.00 & 0.42 & 0.07 & 0.06 \\
0.42 & 1.00 & -0.06 & 0.24 \\
0.07 & -0.06 & 1.00 & 0.63 \\
0.06 & 0.24 & 0.63 & 1.00
\end{pmatrix}
\end{array}
= \begin{pmatrix} \Sigma_{\mathbf{yy}} & \Sigma_{\mathbf{yx}} \\ \Sigma_{\mathbf{xy}} & \Sigma_{\mathbf{xx}} \end{pmatrix},
$$

(1) (x_1, x_2)与(y_1, y_2) 的典则相关系数.

$$\Sigma_{\mathbf{yy}}^{-1/2} = \begin{pmatrix} 1 & 0.42 \\ 0.42 & 1 \end{pmatrix}^{-1/2} = \begin{pmatrix} 1.08 & -0.24 \\ -0.24 & 1.08 \end{pmatrix}$$

$$\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}} = \begin{pmatrix} 0.07 & 0.06 \\ -0.06 & 0.24 \end{pmatrix} \begin{pmatrix} 1 & 0.63 \\ 0.63 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0.07 & -0.06 \\ 0.06 & 0.24 \end{pmatrix}$$

$$\Phi = \Sigma_{\mathbf{yy}}^{-1/2}(\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}})\Sigma_{\mathbf{yy}}^{-1/2} = \begin{pmatrix} 0.01 & -0.03 \\ -0.03 & 0.15 \end{pmatrix}$$

第一典则系数：$\sqrt{\lambda_{max}(\Phi)} = 0.4$
检验：$n\lambda_{max}(\Phi) = 4.7 > \chi^2_{|p-q|+1}(0.05) = 3.84$。
数学与阅读显著相关。

$$
\Sigma = \begin{array}{c|cccc} & y_1 & y_2 & x_1 & x_2 \\ \hline y_1 & 1.00 & 0.42 & 0.07 & 0.06 \\ y_2 & 0.42 & 1.00 & -0.06 & 0.24 \\ x_1 & 0.07 & -0.06 & 1.00 & 0.63 \\ x_2 & 0.06 & 0.24 & 0.63 & 1.00 \end{array} = \begin{pmatrix} \Sigma_{\mathbf{yy}} & \Sigma_{\mathbf{yx}} \\ \Sigma_{\mathbf{xy}} & \Sigma_{\mathbf{xx}} \end{pmatrix},
$$

(2) y_1与$\mathbf{x} = (x_1, x_2)$的决定系数：

$$\phi_1 = \Sigma_{y_1y_1}^{-1/2}\left(\Sigma_{y_1\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y_1}\right)\Sigma_{y_1y_1}^{-1/2} \qquad \Sigma_{y_1y_1} = 1$$

$= (0.07\ 0.06)\begin{pmatrix} 1 & 0.63 \\ 0.63 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 0.07 \\ 0.06 \end{pmatrix} = 0.005$，典则系数$\sqrt{\phi_1} = 0.07$

同样，y_2与$\mathbf{x} = (x_1, x_2)$的决定系数：

$$\phi_2 = \Sigma_{y_2y_2}^{-1/2}\left(\Sigma_{y_2\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y_2}\right)\Sigma_{y_2y_2}^{-1/2}$$

$= (-0.06\ 0.24)\begin{pmatrix} 1 & 0.63 \\ 0.63 & 1 \end{pmatrix}^{-1}\begin{pmatrix} -0.06 \\ 0.24 \end{pmatrix} = 0.13$，典则系数$\sqrt{\phi_2} = 0.36$

(3) y_1, y_2的偏相关系数(控制阅读)

$$\Sigma_{\mathbf{yy}\bullet\mathbf{x}} = \begin{pmatrix} \Sigma_{y_1y_1\bullet\mathbf{x}} & \Sigma_{y_1y_2\bullet\mathbf{x}} \\ \Sigma_{y_2y_1\bullet\mathbf{x}} & \Sigma_{y_2y_2\bullet\mathbf{x}} \end{pmatrix} = \Sigma_{\mathbf{yy}} - \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}}$$

$$= \begin{pmatrix} 1 & 0.42 \\ 0.42 & 1 \end{pmatrix} - \begin{pmatrix} 0.07 & 0.06 \\ -0.06 & 0.24 \end{pmatrix} \begin{pmatrix} 1 & 0.63 \\ 0.63 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0.07 & -0.06 \\ 0.06 & 0.24 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0.42 \\ 0.42 & 0.87 \end{pmatrix}$$

所以 $\rho_{y_1y_2\bullet\mathbf{x}} = 0.42/\sqrt{1 \times 0.87} = 0.45$

另一种算法（参见后面偏相关系数矩阵）

$$\Omega = \Sigma^{-1} = \begin{array}{c} \\ y_1 \\ y_2 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & x_1 & x_2 \end{array} \\ \begin{pmatrix} \omega_{11} & \omega_{12} & & \\ \omega_{21} & \omega_{22} & & \\ & & & \\ & & & \end{pmatrix} \end{array} = \begin{pmatrix} 1.26 & -0.60 & & \\ -0.60 & 1.44 & & \\ & & & \\ & & & \end{pmatrix},$$

$$\rho_{y_1y_2\bullet x_1x_2} = -\omega_{12}/\sqrt{\omega_{11}\omega_{22}} = 0.45$$

(4)控制阅读速度x_2的条件下，数学 (y_1, y_2)是否与阅读能力x_1有关？ 记$w=(y_1, y_2, x_1)$

$$\Sigma = \begin{array}{c|cccc} & y_1 & y_2 & x_1 & x_2 \\ y_1 & 1.00 & 0.42 & 0.07 & 0.06 \\ y_2 & 0.42 & 1.00 & -0.06 & 0.24 \\ x_1 & 0.07 & -0.06 & 1.00 & 0.63 \\ x_2 & 0.06 & 0.24 & 0.63 & 1.00 \end{array} = \begin{pmatrix} \Sigma_{\mathbf{ww}} & \Sigma_{\mathbf{w}x_2} \\ \Sigma_{x_2\mathbf{w}} & \Sigma_{x_2x_2} \end{pmatrix},$$

$$\Sigma_{\mathbf{ww}\bullet x_2} = \Sigma_{\mathbf{ww}} - \Sigma_{\mathbf{w}x_2}\Sigma_{x_2x_2}^{-1}\Sigma_{x_2\mathbf{w}}$$

$$= \begin{pmatrix} 1 & 0.42 & 0.07 \\ 0.42 & 1 & -0.06 \\ 0.07 & -0.06 & 1 \end{pmatrix} - \begin{pmatrix} 0.06 \\ 0.24 \\ 0.63 \end{pmatrix} (0.06\ 0.24\ \ 0.63)$$

$$= \begin{array}{c} y_1 \\ y_2 \\ x_1 \end{array} \begin{pmatrix} 0.996 & 0.406 & 0.032 \\ 0.406 & 0.942 & -0.212 \\ 0.032 & -0.212 & 0.603 \end{pmatrix},$$ 基于该矩阵，计算（y_1，y_2）与x_1的决定系数

$$\Phi_{(y_1,y_2)x_1\bullet x_2} = (0.032\ \ -0.212) \begin{pmatrix} 0.996 & 0.406 \\ 0.406 & 0.942 \end{pmatrix}^{-1} \begin{pmatrix} 0.032 \\ -0.212 \end{pmatrix} /0.603 = 0.11$$

偏相关系数矩阵

分块矩阵的逆

命题1（分块矩阵的逆）. $\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}>0$ （正定）

则 $\Sigma^{-1}=\begin{pmatrix}\Sigma_{11\bullet 2}^{-1} & -\Sigma_{11\bullet 2}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\\ -\Sigma_{22\bullet 1}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22\bullet 1}^{-1}\end{pmatrix}$

其中 $\Sigma_{11\bullet 2}=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$， $\Sigma_{22\bullet 1}=\Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$，

证明1：记 $\Omega=\Sigma^{-1}=\begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix}$，由 $\Sigma\Omega=I$，

$$\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\begin{pmatrix}\Omega_{11} & \Omega_{12}\\ \Omega_{21} & \Omega_{22}\end{pmatrix}=\begin{pmatrix}I_q & 0\\ 0 & I_p\end{pmatrix}$$ （对角子矩阵阶数分别为 q, p）

$\Rightarrow$

(1) $\Sigma_{11}\Omega_{11}+\Sigma_{12}\Omega_{21}=I_q$

(2) $\Sigma_{21}\Omega_{11}+\Sigma_{22}\Omega_{21}=0\Rightarrow\Omega_{21}=-\Sigma_{22}{}^{-1}\Sigma_{21}\Omega_{11}$ 代入(1)

$$\Rightarrow\Sigma_{11}\Omega_{11}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma_{21}\Omega_{11}=I_q\Rightarrow\Omega_{11}=\left(\Sigma_{11}-\Sigma_{12}\Sigma_{22}{}^{-1}\Sigma_{21}\right)^{-1}.$$

证明2：由去相关化对角化公式

$$\begin{pmatrix} I_q & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_p \end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} I_q & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I_p \end{pmatrix}=\begin{pmatrix} \Sigma_{11\bullet 2} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$$

两边求逆: $\begin{pmatrix} I_q & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I_p \end{pmatrix}^{-1}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1}\begin{pmatrix} I_q & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_p \end{pmatrix}^{-1}=\begin{pmatrix} {\Sigma_{11\bullet 2}}^{-1} & 0 \\ 0 & {\Sigma_{22}}^{-1} \end{pmatrix}$

$$\Rightarrow \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1}=\begin{pmatrix} I_q & 0 \\ -\Sigma_{22}^{-1}\Sigma_{21} & I_p \end{pmatrix}\begin{pmatrix} \Sigma_{11\bullet 2}^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{pmatrix}\begin{pmatrix} I_q & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_p \end{pmatrix}$$

$$=\begin{pmatrix} {\Sigma_{11\bullet 2}}^{-1} & -\Sigma_{11\bullet 2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\bullet 2}^{-1} & \Sigma_{22}^{-1}+\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\bullet 2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{pmatrix}$$

由对称性知右下角等于$\Sigma_{22\bullet 1}^{-1}$.

给定$\mathbf{x}$的协方差矩阵Σ，则x_i, x_j的相关系数

$$\rho_{ij} = \sigma_{ij} / \sqrt{\sigma_{ii}\sigma_{jj}},$$

且相关系数矩阵

$$R = \left(\rho_{ij}\right) = D^{-1/2}\Sigma D^{-1/2},\ D = diag(\Sigma)。$$

如何计算偏相关系数矩阵$R_{\text{partial}} = (\rho_{ij\bullet\text{其它}})$?

偏相关系数矩阵

命题2：随机向量$\mathbf{x}_{p\times1}$的协方差矩阵为$\Sigma = \left(\sigma_{ij}\right)$，$\Omega = \Sigma^{-1} = \left(\omega_{ij}\right)$称为精度矩阵(precision matrix)，记$C = diag(\Omega)$，则 x_i, x_j的偏相关系数

$$\rho_{ij\bullet\text{其它}} = \begin{cases} -\omega_{ij} / \sqrt{\omega_{ii}\omega_{jj}} & i \neq j \\ 1 & i = j \end{cases},$$

且偏相关系数矩阵

$$R_{\text{partial}} = (\rho_{ij\bullet\text{其它}}) = 2I_p - C^{-1/2}\Omega C^{-1/2}.$$

注1：Σ元素包含了相关性信息；Σ^{-1}元素包含了偏相关性信息。

注2：基于Σ^{-1}的偏相关系数计算法则类似于基于Σ的相关系数计算法则，但相差一个符号。

注3：$\rho_{ij\bullet\text{其它}} = 0 \Leftrightarrow \Sigma^{-1}$的$(i, j)$元$= 0$，即$\omega_{ij} = 0$。

证明：证明主要基于两点事实(1)偏相关系数中出现的量$\Sigma_{ab\bullet\mathbf{z}}$在$\Sigma^{-1}$中也出现了（2）二阶方阵的逆的元素与原矩阵除了一个常数因子之外相同或差一个符号。

随机向量$\mathbf{x}=(x_1,\dots,x_p)^{\mathsf{T}}$. 下面对$i=1, j=2$计算。

记$\mathbf{w}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{z}=(x_3,\dots,x_p)^{\mathsf{T}}$, 即$\mathbf{x}=\begin{pmatrix} \mathbf{w} \\ \mathbf{z} \end{pmatrix}$

$$\Sigma=\operatorname{cov}\begin{pmatrix} \mathbf{w} \\ \mathbf{z} \end{pmatrix}=\begin{pmatrix} \Sigma_{\mathbf{ww}} & \Sigma_{\mathbf{wz}} \\ \Sigma_{\mathbf{zw}} & \Sigma_{\mathbf{zz}} \end{pmatrix}c=\begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{1\mathbf{z}} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{2\mathbf{z}} \\ \Sigma_{\mathbf{z}1} & \Sigma_{\mathbf{z}1} & \Sigma_{\mathbf{zz}} \end{pmatrix},$$

$$而\Sigma_{\mathbf{ww\bullet z}}=\Sigma_{\mathbf{ww}}-\Sigma_{\mathbf{wz}}\Sigma_{\mathbf{zz}}^{-1}\Sigma_{\mathbf{zw}}=\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}-\begin{pmatrix} \Sigma_{1\mathbf{z}} \\ \Sigma_{2\mathbf{z}} \end{pmatrix}\Sigma_{\mathbf{zz}}^{-1}(\Sigma_{\mathbf{z}1},\ \Sigma_{\mathbf{z}2})$$

$$=\begin{pmatrix} \Sigma_{11\bullet\mathbf{z}} & \Sigma_{12\bullet\mathbf{z}} \\ \Sigma_{21\bullet\mathbf{z}} & \Sigma_{22\bullet\mathbf{z}} \end{pmatrix}\overset{\text{记为}}{=}\begin{pmatrix} a & b \\ b & c \end{pmatrix},\quad 所以\rho_{12\bullet\mathbf{z}}=\frac{b}{\sqrt{ac}}$$

记$\Omega=\Sigma^{-1}$,注意到分块求逆公式（特别地，左上角与$\Sigma_{\mathbf{ww\bullet z}}$有关）：

$$\Omega=\Sigma^{-1}=\begin{pmatrix}\Sigma_{\mathbf{ww}} & \Sigma_{\mathbf{wz}}\\ \Sigma_{\mathbf{zw}} & \Sigma_{\mathbf{zz}}\end{pmatrix}^{-1}=\begin{pmatrix}\Sigma_{\mathbf{ww\bullet z}}^{-1} & *\\ * & *\end{pmatrix}=\begin{pmatrix}\begin{pmatrix}a & b\\ b & c\end{pmatrix}^{-1} & *\\ * & *\end{pmatrix}$$

$$=\begin{pmatrix}\dfrac{1}{ac-b^2}\begin{pmatrix}c & -b\\ -b & a\end{pmatrix} & *\\ * & *\end{pmatrix}$$

$$=\begin{pmatrix}\begin{pmatrix}\omega_{11} & \omega_{12}\\ \omega_{21} & \omega_{22}\end{pmatrix} & *\\ * & *\end{pmatrix}$$

关键：2×2矩阵的逆除了公共因子$\dfrac{1}{ac-b^2}$之外，其元素是原矩阵元素的置换或相反数：

$$\begin{pmatrix}a & b\\ b & c\end{pmatrix}^{-1}=\frac{1}{ac-b^2}\begin{pmatrix}c & -b\\ -b & a\end{pmatrix}$$

所以$\rho_{12\bullet\mathbf{z}}=\dfrac{b}{\sqrt{ac}}=-\dfrac{\omega_{12}}{\sqrt{\omega_{11}\omega_{22}}}$

多元正态分布

多元正态有许多很好的性质，其中最重要的（对于本课程）是

1. 多元正态假设下，偏相关系数为条件相关系数
2. 条件分布满足（启发）所有线性模型的假设

多元正态

定义：若p维随机向量$\mathbf{x}$的密度为

$$f(\mathbf{x})=\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right),\quad \mathbf{x}\in \mathrm{R}^{p}$$

其中参数$\boldsymbol{\mu}\in R^{p}$，$\Sigma$为$p\times p$正定矩阵，则$\mathbf{x}$服从多元正态分布，记为$\mathbf{x}\sim N_{p}(\boldsymbol{\mu},\Sigma)$。

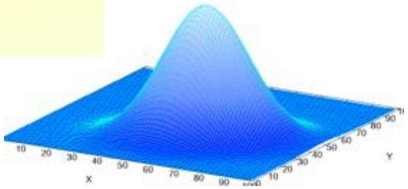
多元标准正态

$\boldsymbol{\mu}=\mathbf{0},\Sigma=I_{p}$时，$N_{p}(0,I_{p})$称作多元标准正态分布

$$\mathbf{x}\sim N_{p}(0,I_{p})\Leftrightarrow f(\mathbf{x})=\frac{1}{(2\pi)^{p/2}}e^{-\frac{\mathbf{x}^{\mathrm{T}}\mathbf{x}}{2}}=\prod_{i=1}^{p}\frac{1}{\sqrt{2\pi}}e^{-\frac{x_i^2}{2}}$$

$$\Leftrightarrow x_1,\ldots,x_p\ iid\sim N(0,1)$$

$$\Rightarrow E(x_i)=0,\mathrm{var}(x_i)=1\Rightarrow E(\mathbf{x})=\mathbf{0},\mathrm{var}(\mathbf{x})=I_{p}.$$



引理1：若 $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma)$，则对任意可逆常数矩阵$A$,常数向量$\mathbf{b}_{p\times 1}$,
$A\mathbf{x} + \mathbf{b} \sim N_p\left(A\boldsymbol{\mu} + \mathbf{b}, A\Sigma A^\top\right)$。

证明：令$\mathbf{y} = A\mathbf{x} + \mathbf{b}, \mathbf{x} = A^{-1}(\mathbf{y} - \mathbf{b})$, $Jacobi$行列式$J = |\partial \mathbf{x} / \partial \mathbf{y}| = |A^{-1}| = 1/|A|$,
$\mathbf{y}$的概率密度

$$g(\mathbf{y}) = f(A^{-1}(\mathbf{y} - \mathbf{b})) \times |J|$$

$$= \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}|A|}\exp\left(-\frac{1}{2}(A^{-1}(\mathbf{y}-\mathbf{b})-\boldsymbol{\mu})^\top \Sigma^{-1}(A^{-1}(\mathbf{y}-\mathbf{b})-\boldsymbol{\mu})\right)$$

$$= \frac{1}{(2\pi)^{p/2}|A\Sigma A^\top|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{y}-\mathbf{b}-A\boldsymbol{\mu})^\top)A\Sigma A^\top)^{-1}(\mathbf{y}-\mathbf{b}-A\boldsymbol{\mu})\right),$$

所以$\mathbf{y} \sim N_p\left(A\boldsymbol{\mu} + \mathbf{b}, A\Sigma A^\top\right)$.

注：后面定理3将说明A可以是任何行满秩的长方形矩阵。

推论1: (1) $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma) \Rightarrow \mathbf{y} = \Sigma^{-1/2}(\mathbf{x}-\boldsymbol{\mu}) \sim N_p(0, I_p)$.

(2) $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma) \Leftrightarrow \mathbf{x}$可以表示为$\mathbf{x} = B\mathbf{y} + \boldsymbol{\mu}$, 其中$\mathbf{y} \sim N_p(0, I_p)$,

方阵B使得$\Sigma = BB^{\top}$(通常$B = \Sigma^{1/2}$).

证(2): ($\Rightarrow$)取$\mathbf{y} = \Sigma^{-1/2}(\mathbf{x}-\boldsymbol{\mu})$，则$\mathbf{x} = \Sigma^{1/2}\mathbf{y} + \boldsymbol{\mu}$，$\mathbf{y} \sim N_p(0, I_p)$

($\Leftarrow$)若$\mathbf{y} \sim N_p(0, I_p)$，则由引理1，$\mathbf{x} = B\mathbf{y} + \boldsymbol{\mu} \sim N(\boldsymbol{\mu}, BB^{\top})$.

定理1:

(1) $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma) \Rightarrow E(\mathbf{x}) = \boldsymbol{\mu}, \mathrm{var}(\mathbf{x}) = \Sigma$.

(2) $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma) \Rightarrow$ 矩母函数$E(\exp(\mathbf{t}^{\top}\mathbf{x})) = \exp(\mathbf{t}^{\top}\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}^{\top}\Sigma\mathbf{t})$, 任何$\mathbf{t} \in R^p$

证明： （1）令$\mathbf{y} = \Sigma^{-1/2}(\mathbf{x}-\boldsymbol{\mu}) \sim N_p(0, I_p)$，即$\mathbf{x} = \Sigma^{1/2}\mathbf{y} + \boldsymbol{\mu}$,

因为$E(\mathbf{y}) = 0, \mathrm{var}(\mathbf{y}) = I_p$, 所以

$E(\mathbf{x}) = E(\Sigma^{1/2}\mathbf{y} + \boldsymbol{\mu}) = \boldsymbol{\mu}, \mathrm{var}(\mathbf{x}) = \mathrm{var}(\Sigma^{1/2}\mathbf{y} + \boldsymbol{\mu}) = \Sigma^{1/2}\,\mathrm{var}(\mathbf{y})\Sigma^{1/2} = \Sigma$.

(2)类似于（1）可证。

推论1用多元标准正态表示一般的多元正态，因为多元标准正态容易处理，所以这种表示方便应用（参见定理1 ），这种表示可以作为多元正态的等价定义。

定义2（奇异多元正态分布，P.L. Hsu）. 假设$\mathbf{y} \sim N_p(0, I_p)$，$B$是任一$q \times p$常数矩阵，$\boldsymbol{\mu}$是$q \times 1$常数向量，则称$\mathbf{x} = B\mathbf{y} + \boldsymbol{\mu}$的分布为奇异（singular)多元正态分布，记作$\mathbf{x} \sim N_q\left(\boldsymbol{\mu}, BB^\top\right)$。

注：当B可逆时，定义2等价于定义1.

下面讨论多元正态$\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma)$的分量的分布，为此划分

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

其中$\mathbf{x}_1$，$\boldsymbol{\mu}_1$为$q \times 1$向量，Σ_{11}为$q \times q$矩阵，

$$\boxed{\begin{array}{l}\text{引理2：若}\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma)\text{，且如上划分，则 }\mathbf{x}_1\text{与}\mathbf{x}_2\text{独立} \Leftrightarrow \Sigma_{12} = 0, \\ \text{此时 }\mathbf{x}_1 \sim N_q(\boldsymbol{\mu}_1, \Sigma_{11}), \mathbf{x}_2 \sim N_{p-q}(\boldsymbol{\mu}_2, \Sigma_{22}).\end{array}}$$

证：若$\Sigma_{12} = 0$，则容易验证密度函数可以分解成$\mathbf{x}_1 \sim N_q(\boldsymbol{\mu}_1, \Sigma_{11})$和$\mathbf{x}_2 \sim N_{p-q}(\boldsymbol{\mu}_2, \Sigma_{22})$的密度乘积。

边际与条件分布

定理2：若$\mathbf{x} \sim N_p(\boldsymbol{\mu}, \Sigma)$，且如前划分，则

(1) $\mathbf{x}_1 \sim N_q(\boldsymbol{\mu}_1, \Sigma_{11}), \mathbf{x}_2 \sim N_{p-q}(\boldsymbol{\mu}_2, \Sigma_{22})$

(2) $\mathbf{x}_1 | \mathbf{x}_2 \sim N_q\left(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2),\ \Sigma_{11\bullet 2}\right)$.

证：令去相关化变换：

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^{\perp} \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} I_q & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I_{p-q} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 - \Sigma_{12}\Sigma_{22}^{-1}\mathbf{x}_2 \\ \mathbf{x}_2 \end{pmatrix},$$

由引理1，$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim N_p\left(\begin{pmatrix} \boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11\bullet 2} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}\right),$

由引理2，$\mathbf{y}_1$，$\mathbf{y}_2$独立，且$\mathbf{x}_2 = \mathbf{y}_2 \sim N_{p-q}(\boldsymbol{\mu}_2, \Sigma_{22})$，

$$\mathbf{y}_1 \sim N_q(\boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{\mu}_2, \Sigma_{11\bullet 2}),$$

所以给定$\mathbf{x}_2$时，$\mathbf{x}_1 = \mathbf{y}_1 + \Sigma_{12}\Sigma_{22}^{-1}\mathbf{x}_2 \sim N_q(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11\bullet 2})$.

条件分布$\mathbf{x}_1|\mathbf{x}_2 \sim N_q(\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2), \Sigma_{11\bullet 2})$蕴含了如下事实

- 条件方差常数： $\mathrm{var}(\mathbf{x}_1|\mathbf{x}_2) = \Sigma_{11\bullet 2}$是常数矩阵与$\mathbf{x}_2$无关。
- 条件期望线性：$E(\mathbf{x}_1|\mathbf{x}_2) = \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$是$\mathbf{x}_2$的线性函数。
- 去相关化：$\boldsymbol{\varepsilon} = \mathbf{x}_1 - E(\mathbf{x}_1|\mathbf{x}_2) = \mathbf{x}_1 - [\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)]$,

  $\boldsymbol{\varepsilon}|\mathbf{x}_2 \sim N_q(0, \Sigma_{11\bullet 2})$ 与$\mathbf{x}_2$无关，所以$\boldsymbol{\varepsilon} \perp\!\!\!\perp \mathbf{x}_2$，$\boldsymbol{\varepsilon} \sim N_q(0, \Sigma_{11\bullet 2})$.
- 多元正态线性模型（改写去相关化）：

$$\mathbf{x}_1 = (\boldsymbol{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\boldsymbol{\mu}_2) + \Sigma_{12}\Sigma_{22}^{-1}\mathbf{x}_2 + \boldsymbol{\varepsilon} \triangleq \mathbf{a} + \mathbf{b}^{\mathrm{T}}\mathbf{x}_2 + \boldsymbol{\varepsilon},$$

  其中$\boldsymbol{\varepsilon} \perp\!\!\!\perp \mathbf{x}_2$， $\boldsymbol{\varepsilon} \sim N_q(0, \Sigma_{11\bullet 2})$。

一般的多元线性回归模型放松误差正态性：

$$\mathbf{x}_1 = \mathbf{a} + \mathbf{b}^{\mathrm{T}}\mathbf{x}_2 + \boldsymbol{\varepsilon},\ \ \boldsymbol{\varepsilon} \perp\!\!\!\perp \mathbf{x}_2,\ \ \boldsymbol{\varepsilon} \sim (0, \Sigma_{11\bullet 2})$$

定理3：若 $\mathbf{x}\sim N_p(\boldsymbol{\mu},\Sigma)$，则对任意行满秩$q\times p$常数矩阵$A$，常数向量$\mathbf{b}_{p\times 1}$，$A\mathbf{x}+\mathbf{b}\sim N_q\left(A\boldsymbol{\mu}+\mathbf{b}, A\Sigma A^{\top}\right)$。

证：若 $\mathbf{x}\sim N_p(\boldsymbol{\mu},\Sigma)$，

取$(p-q)\times p$矩阵B，使得$C=\begin{pmatrix}A\\B\end{pmatrix}$可逆，则由引理1，

$$C\mathbf{x}=\begin{pmatrix}A\mathbf{x}\\B\mathbf{x}\end{pmatrix}\sim N\left(C\boldsymbol{\mu}, C\Sigma C^{\top}\right)=N\left(\begin{pmatrix}A\boldsymbol{\mu}\\B\boldsymbol{\mu}\end{pmatrix},\begin{pmatrix}A\Sigma A^{\top} & A\Sigma B^{\top}\\B\Sigma A^{\top} & B\Sigma B^{\top}\end{pmatrix}\right),$$

由定理2，边际$A\mathbf{x}\sim N\left(A\boldsymbol{\mu}, A\Sigma A^{\top}\right)$，$A\mathbf{x}+\mathbf{b}\sim N\left(A\boldsymbol{\mu}+\mathbf{b}, A\Sigma A^{\top}\right)$。

条件独立

定理4.（1）多元正态假设下，偏相关系数等于条件相关系数。

（2）设 $\mathbf{x} \sim N_n(\boldsymbol{\mu}, \Sigma)$, $\Sigma^{-1} = (\omega_{ij})$. 若 $\omega_{ij} = 0$，则给定其它变量时，x_i与x_j条件独立。

证明：（1）假设 $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \mathbf{z} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_{\mathbf{z}} \end{pmatrix}, \ \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{1\mathbf{z}} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{2\mathbf{z}} \\ \Sigma_{\mathbf{z}1} & \Sigma_{\mathbf{z}2} & \Sigma_{\mathbf{zz}} \end{pmatrix}\right)$,

记 $\mathbf{w} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, 则 $\mathbf{w}|_{\mathbf{z}} \sim N(*, \Sigma_{\mathbf{ww}\bullet\mathbf{z}})$, 其中 $\Sigma_{\mathbf{ww}\bullet\mathbf{z}} = \mathrm{var}\left(\begin{matrix} x_1 \\ x_2 \end{matrix} \middle| \mathbf{z}\right) = \begin{pmatrix} \Sigma_{11\bullet\mathbf{z}} & \Sigma_{12\bullet\mathbf{z}} \\ \Sigma_{21\bullet\mathbf{z}} & \Sigma_{22\bullet\mathbf{z}} \end{pmatrix}$

是条件协方差矩阵。则给定 $(x_3,.., x_p) = \mathbf{z}$ 条件下，x_1和x_2的条件相关系数

$\rho_{12|\mathbf{z}} = \dfrac{\Sigma_{12\bullet\mathbf{z}}}{\sqrt{\Sigma_{11\bullet\mathbf{z}}}\sqrt{\Sigma_{22\bullet\mathbf{z}}}}$，这也是偏相关系数 $\rho_{12\bullet\mathbf{z}}$

（2）由命题2，$\omega_{12} = 0 \Leftrightarrow \rho_{12\bullet\mathbf{z}} = -\dfrac{\omega_{12}}{\sqrt{\omega_{11}\omega_{22}}} = 0$

$\overset{\text{由（1）}}{\Leftrightarrow} \rho_{12|\mathbf{z}} = 0$, 条件不相关

$\Leftrightarrow x_1$与x_2条件独立(因为正态假设下不相关 $\Leftrightarrow$ 独立)

基于下述引理3，我们给出定理4（2）的更直接的证明

引理3：若x_1和x_2的联合密度函数$f(x_1,x_2)$可分解为$f(x_1,x_2)=\varphi_1(x_1)\varphi_2(x_2)$，其中$\varphi_1$，$\varphi_2$未必是密度函数，则$x_1$和$x_2$独立。

(2)的另一个证明：

不妨设均值$\boldsymbol{\mu}=0$，设$\omega_{12}=0$，则$\mathbf{x}$的联合密度

$$f(\mathbf{x})=C\exp\left(-\frac{1}{2}\mathbf{x}^{\mathsf{T}}\Omega\mathbf{x}\right)=C\exp\left(-\frac{1}{2}\omega_{11}x_1^2-\frac{1}{2}\omega_{22}x_2^2+c_1x_1+c_2x_2+c_3\right)$$

其中C是常数，其中c_1,c_2,c_3只与$x_3,...,x_n$有关。无交叉项x_1x_2.

给定$x_3,...,x_n$(为常数)时，x_1,x_2的联合密度

$$f(x_1,x_2\,|\,x_3,...,x_n)\propto f(\mathbf{x})=C\exp\left(-\frac{1}{2}\omega_{11}x_1^2-\frac{1}{2}\omega_{22}x_2^2+c_1x_1+c_2x_2+c_3\right)$$

$=f_1(x_1)f_2(x_2)$，所以x_1,x_2条件独立。

至此，我们介绍了Pearson相关系数及其拓展

- ❑ 两个随机变量的Pearson相关系数
- ❑ 两个随机变量的偏相关系数
- ❑ 两个随机向量的典则相关系数
- ❑ 一个随机变量和一个随机向量之间的决定系数
- ❑ 控制第三个随机向量时，两个随机向量之间的偏典则相关系数/决定系数。

在定义偏相关系数和典则系数的时候，去相关化$\mathbf{y}^{\perp} = \mathbf{y} - \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$ 是基本工具。

在随机变量空间中定义内积$< x, y >= E(xy)$，去相关化实际上是正交投影或Gram-Schmidt正交化。

本次课介绍了多元正态，说明了正态情形下偏相关系数实际上是条件相关系数，偏相关系数等于0代表条件独立。另外，求解了正态分布的条件分布，该条件分布满足或启发了线性模型的所有假设。