回归分析 (01714601)
课程主页：http://staff.ustc.edu.cn/~ynyang/lm2020

第四讲 偏相关系数

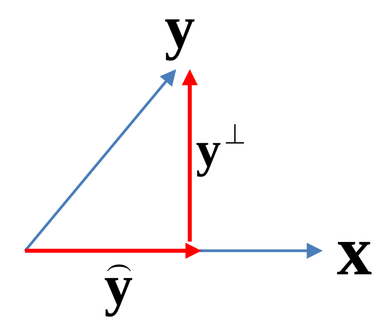
2020.2.28

随机向量的“不相关化”（续）

随机向量$\mathbf{y}$, $\mathbf{x}$的协方差矩阵为$\mathrm{cov}(\mathbf{y},\mathbf{x})=\Sigma_{\mathbf{yx}}$,

正交分解：$\mathbf{y}=\mathbf{y}^{\perp}+\widehat{\mathbf{y}}$,

$\widehat{\mathbf{y}}=\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$, $\mathbf{y}^{\perp}=\mathbf{y}-\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$



- $\mathbf{y}^{\perp}$: $\mathbf{y}$关于$\mathbf{x}$的不相关化，$\mathrm{cov}(\mathbf{y}^{\perp},\mathbf{x})=0$.
- $\widehat{\mathbf{y}}$: $\mathbf{y}$中能被$\mathbf{x}$解释的部分（$\mathbf{y}$在$\mathbf{x}$上的投影).

方差分解：$\underset{\text{总方差}}{\mathrm{var}(\mathbf{y})} = \underset{\text{与}\mathbf{x}\text{有关的部分}}{\mathrm{var}(\hat{\mathbf{y}})} + \underset{\text{与}\mathbf{x}\text{无关的部分}}{\mathrm{var}(\mathbf{y}^{\perp})}$

$$\Sigma_{\mathbf{yy}} = \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}} + \Sigma_{\mathbf{yy}\bullet\mathbf{x}}$$

我们用矩阵

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

代表$\mathrm{var}(\hat{\mathbf{y}}) = \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}}$在总方差$\mathrm{var}(\mathbf{y}) = \Sigma_{\mathbf{yy}}$中的占比,

B是矩阵，我们用B的最大特征根的平方根$\lambda_{\max}^{1/2}(B)$

度量$\mathbf{x},\mathbf{y}$的相关程度,称为第一典则相关系数.

注：$C = \Sigma_{\mathbf{xx}}^{-1/2}\left(\Sigma_{\mathbf{xy}}\Sigma_{\mathbf{yy}}^{-1}\Sigma_{\mathbf{yx}}\right)\Sigma_{\mathbf{xx}}^{-1/2}$代表了$\mathbf{y}$对$\mathbf{x}$的解释比率，注意$C, B$有共同的非0特征根,所以典则相关系数关于$\mathbf{x},\mathbf{y}$是对称的。

y为随机变量$(q=1)$时，可以证明不相关化具有最优性。

命题5. y为随机变量$(q=1)$, $\mathbf{x}$为$p\times 1$随机向量, $\operatorname{cov}\begin{pmatrix}\mathbf{x}\\ y\end{pmatrix}=\begin{pmatrix}\Sigma_{\mathbf{xx}} & \Sigma_{\mathbf{x}y}\\ \Sigma_{y\mathbf{x}} & \Sigma_{yy}\end{pmatrix}$.

对任何$\mathbf{a}\in R^p$, $\mathbf{a}^{\top}\mathbf{x}$与$y$的相关系数$\rho(\mathbf{a}^{\top}\mathbf{x},y)$在$\mathbf{a}=\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}$时达到最大。

证明:因为$\operatorname{cov}(\mathbf{a}^{\top}\mathbf{x},y)=\mathbf{a}^{\top}\Sigma_{\mathbf{x}y}$, $\operatorname{var}(\mathbf{a}^{\top}\mathbf{x})=\mathbf{a}^{\top}\Sigma_{\mathbf{xx}}\mathbf{a}$, $\operatorname{var}(y)=\Sigma_{yy}$,所以

$$\rho(\mathbf{a}^{\top}\mathbf{x},y)=\frac{\mathbf{a}^{\top}\Sigma_{\mathbf{x}y}}{\sqrt{\mathbf{a}^{\top}\Sigma_{\mathbf{xx}}\mathbf{a}}\sqrt{\Sigma_{yy}}}$$

令向量$\mathbf{b}=\Sigma_{\mathbf{xx}}^{1/2}\mathbf{a}\Rightarrow\mathbf{a}=\Sigma_{\mathbf{xx}}^{-1/2}\mathbf{b}$

$$=\frac{\mathbf{b}^{\top}\Sigma_{\mathbf{xx}}^{-1/2}\Sigma_{\mathbf{x}y}}{\sqrt{\mathbf{b}^{\top}\mathbf{b}}\sqrt{\Sigma_{yy}}}=\frac{\mathbf{b}^{\top}\mathbf{c}}{\sqrt{\mathbf{b}^{\top}\mathbf{b}}\sqrt{\Sigma_{yy}}}$$

记向量$\mathbf{c}=\Sigma_{\mathbf{xx}}^{-1/2}\Sigma_{\mathbf{x}y}$

由Cauchy-Schwartz不等式，

$$\left|\mathbf{b}^{\mathsf{T}}\mathbf{c}\right| \le \sqrt{\mathbf{b}^{\mathsf{T}}\mathbf{b}}\sqrt{\mathbf{c}^{\mathsf{T}}\mathbf{c}} = \sqrt{\mathbf{b}^{\mathsf{T}}\mathbf{b}}\sqrt{\Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}} \qquad \mathbf{c}^{\mathsf{T}}\mathbf{c} = \Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y},$$

所以 $\left|\rho(\mathbf{a}^{\mathsf{T}}\mathbf{x}, y)\right| = \dfrac{\mathbf{b}^{\mathsf{T}}\mathbf{c}}{\sqrt{\mathbf{b}^{\mathsf{T}}\mathbf{b}}\sqrt{\Sigma_{yy}}} \le \sqrt{\dfrac{\Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}}{\Sigma_{yy}}}$

等号成立但且仅当$\hat{\mathbf{b}} \propto \mathbf{c} = \Sigma_{\mathbf{xx}}^{-1/2}\Sigma_{\mathbf{x}y}$, $\hat{\mathbf{a}} \propto \Sigma_{\mathbf{xx}}^{-1/2}\hat{\mathbf{b}} = \Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}$,

此时，$\hat{\mathbf{a}}^{\mathsf{T}}\mathbf{x} \propto \Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$与$y$的相关系数最大。

最大相关系数 $\max_{\mathbf{a}\in R^p} \left|\rho(\mathbf{a}^\top \mathbf{x}, y)\right| = \sqrt{\dfrac{\Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}}{\Sigma_{yy}}}$ = 称为(第一)典则相关系数.

多元分析中处理$\mathbf{y},\mathbf{x}$都是向量的情形 (参见Page3, $B = \dfrac{\Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}}{\Sigma_{yy}}$)。

方差分解/勾股定理：$\Sigma_{yy} = \Sigma_{yy\bullet\mathbf{x}} + \Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}$

$R^2 = \dfrac{\Sigma_{y\mathbf{x}}\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{x}y}}{\Sigma_{yy}}$ 在回归分析中称为多重相关系数平方，它度量了随机变量y与随机向量$\mathbf{x}$之间的相关性。

副产品: 矩阵的分块对角化和分块矩阵的逆矩阵

$$\mathbf{y}^{\perp} = \mathbf{y} - \Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1}\mathbf{x}$$

从$\begin{pmatrix}\mathbf{y}\\ \mathbf{x}\end{pmatrix}$到$\begin{pmatrix}\mathbf{y}^{\perp}\\ \mathbf{x}\end{pmatrix}$的线性变换为:
$$\begin{pmatrix}\mathbf{y}^{\perp}\\ \mathbf{x}\end{pmatrix} = \begin{pmatrix} I_p & -\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1} \\ 0 & I_q \end{pmatrix}\begin{pmatrix}\mathbf{y}\\ \mathbf{x}\end{pmatrix}$$

两边求方差:

$$\operatorname{cov}\begin{pmatrix}\mathbf{y}^{\perp}\\ \mathbf{x}\end{pmatrix} = \begin{pmatrix} I_p & -\Sigma_{\mathbf{yx}}\Sigma_{\mathbf{xx}}^{-1} \\ 0 & I_q \end{pmatrix}\begin{pmatrix} \Sigma_{\mathbf{yy}} & \Sigma_{\mathbf{yx}} \\ \Sigma_{\mathbf{xy}} & \Sigma_{\mathbf{xx}} \end{pmatrix}\begin{pmatrix} I_p & 0 \\ -\Sigma_{\mathbf{xx}}^{-1}\Sigma_{\mathbf{xy}} & I_q \end{pmatrix} = \begin{pmatrix} \Sigma_{\mathbf{yy}\bullet\mathbf{x}} & 0 \\ 0 & \Sigma_{\mathbf{xx}} \end{pmatrix}$$

Σ的分块对角化.

所以协方差矩阵的对角化，对应于随机向量的不相关化

命题6（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 $\Sigma_{11\bullet 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$， $\Sigma_{22\bullet 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$，

下面我们只用到：

Σ 的左上角为 Σ_{11}，

Σ^{-1} 的左上角为 $\Sigma_{11\bullet 2}^{-1} = \left(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)^{-1}$

略

证明：对角化：

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

两边求逆:

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

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

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

由对称性知右下角的$\Sigma_{22}^{-1}+\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\bullet 2}^{-1}\Sigma_{12}\Sigma_{22}^{-1}$应该等于$\Sigma_{22\bullet 1}^{-1}$,同样左下角

$-\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11\bullet 2}^{-1}=-\Sigma_{22\bullet 1}^{-1}\Sigma_{21}\Sigma_{11}^{-1}.$

偏相关系数的定义

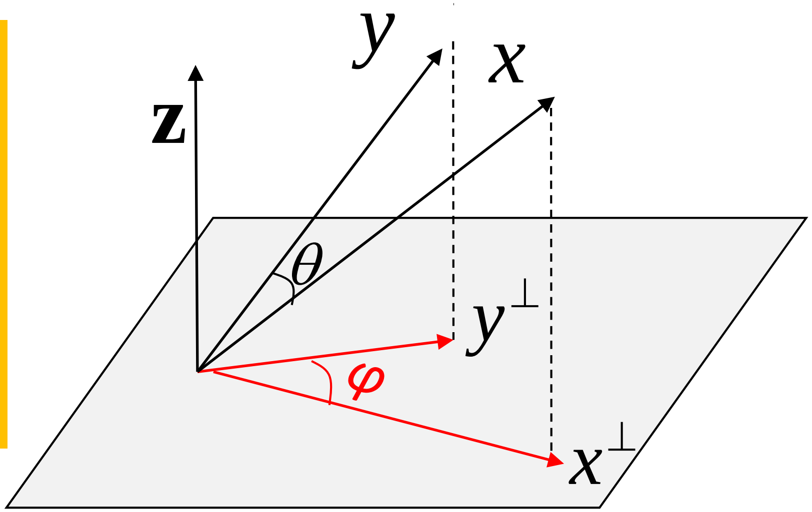
设y, x为1维随机变量，$\mathbf{z}$为随机向量或变量，令

$$y^{\perp} = y - \Sigma_{y\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z}, \quad x^{\perp} = x - \Sigma_{x\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z}$$

定义：偏相关系数

$$\rho_{yx\bullet\mathbf{z}} = \rho_{y^{\perp}x^{\perp}} = \frac{\operatorname{cov}(y^{\perp}, x^{\perp})}{\sqrt{\operatorname{var}(x^{\perp})\operatorname{var}(y^{\perp})}}$$

$$= \cos(\varphi)$$



记号：$\Sigma_{ab} = \mathrm{cov}(\mathrm{a}, \mathrm{b})$,

$$\Sigma_{ab\bullet c} = \Sigma_{ab} - \Sigma_{ac}\Sigma_{cc}^{-1}\Sigma_{cb}$$

容易验证 $\mathrm{cov}(x^{\perp}, y^{\perp}) = \Sigma_{xy\bullet\mathbf{z}}$, $\mathrm{var}(x^{\perp}) = \Sigma_{xx\bullet\mathbf{z}}, \mathrm{var}(y^{\perp}) = \Sigma_{yy\bullet\mathbf{z}}$,

我们得到由原始协方差矩阵计算偏相关系数的公式：

$$\boxed{\text{命题7.}\quad \rho_{xy\bullet\mathbf{z}} = \frac{\Sigma_{xy\bullet\mathbf{z}}}{\sqrt{\Sigma_{xx\bullet\mathbf{z}}}\sqrt{\Sigma_{yy\bullet\mathbf{z}}}}}$$

注：样本偏相关系数的计算公式除了$\bullet z$外与相关系数格式相同,

记号"$\bullet z$"表示控制了z，消除了z的影响。

特例：当x, y, z都是一维随机变量时,相关系数矩阵

$$R = \begin{pmatrix} 1 & \rho_{xy} & \rho_{xz} \\ \rho_{xy} & 1 & \rho_{yz} \\ \rho_{xz} & \rho_{yz} & 1 \end{pmatrix}$$

,则偏相关系数$\rho_{xy\bullet z}$具有如下形式

$$\rho_{xy\bullet z} = \frac{\rho_{xy} - \rho_{xz}\rho_{yz}}{\sqrt{1-\rho_{xz}^2}\sqrt{1-\rho_{yz}^2}}$$

其中ρ_{ab}代表随机变量a, b的Pearson相关系数.

例1（续）．阅读能力(y)、身高(x)和年龄(z)的相关系数矩阵如下，求偏相关系数 $\rho_{xy\bullet z}$

$$
\begin{array}{cc}
 & \begin{array}{ccc} \text{x身高} & \text{y阅读} & \text{z年龄} \end{array} \\
\begin{array}{c} x \\ y \\ z \end{array} & \begin{pmatrix} 1 & 0.50 & 0.8 \\ 0.50 & 1 & 0.7 \\ 0.8 & 0.7 & 1 \end{pmatrix}
\end{array}
$$

$$\rho_{xy\bullet z} = \frac{\rho_{xy} - \rho_{xz}\rho_{yz}}{\sqrt{1-\rho_{xz}^2}\sqrt{1-\rho_{yz}^2}} = \frac{0.50 - 0.8\times 0.7}{\sqrt{1-0.8^2}\sqrt{1-0.7^2}} = -0.14$$

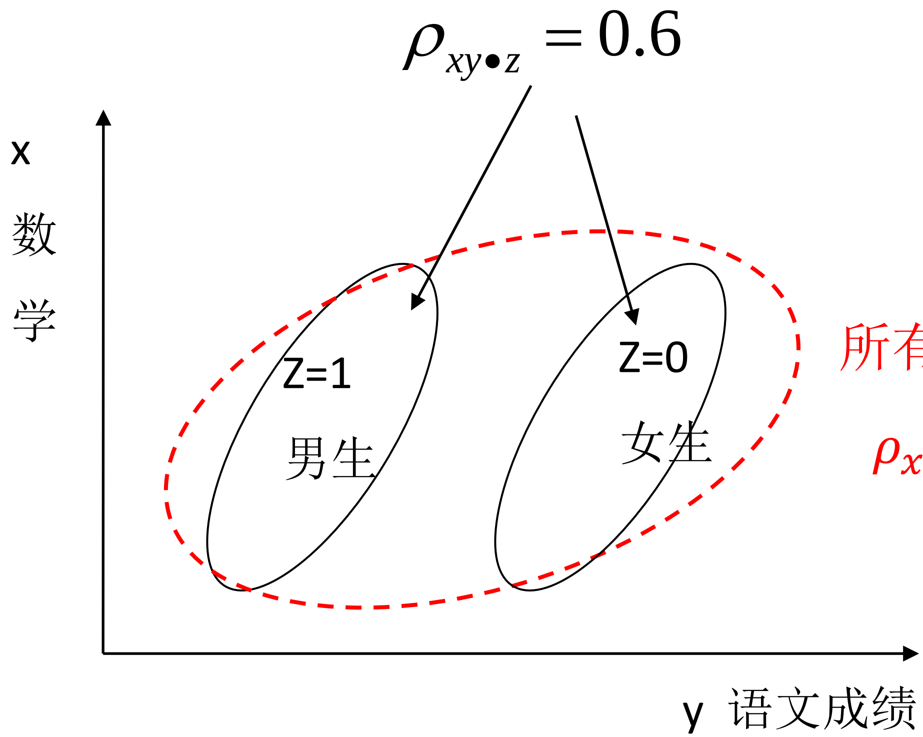
即，控制年龄之后，阅读能力和身高相关系数为-0.14。

例2. 高中男、女生的数学成绩 x与语文成绩 y 的相关系数都为0. 6，假设男女生数学成绩没有差异，但女生的语文成绩要好于男生。如果男女生混合在一起，得到的相关系数是大于、等于还是小于0. 6？

以$z = 1、0$分别代表男、女。已知$\rho_{xy \bullet z} = 0.6$

$\rho_{xz} = 0$ (数学与性别不相关)

故$0.6 = \rho_{xy \bullet z} = \dfrac{\rho_{xy}}{\sqrt{1 - \rho_{yz}^2}} \geq \rho_{xy}$



所有数据的分布轮廓更分散

$\rho_{xy} < 0.6$

偏相关系数的检验

$$H_0: \rho_{xy\bullet\mathbf{z}} = 0$$

记样本偏相关系数为$r_{xy\bullet z}$，设变量总个数为p，样本量为n。

假设数据iid服从联合正态，检验统计量在原假设下

$$T = \sqrt{n-p}\,\frac{r_{xy\bullet\mathbf{z}}}{\sqrt{1-r_{xy\bullet\mathbf{z}}^2}} \sim t_{n-p}$$

数据iid但不假设正态，检验统计量z在原假设下近似地

$$z = \sqrt{n-p} \times r_{xy\bullet\mathbf{z}} \sim N(0,1)，\quad n \to \infty$$

注：p=2时，即为相关系数的检验

偏相关系数的计算

给定一个协方差矩阵或相关系数矩阵，如何高效计算偏相关系数？

命题8：随机向量$\mathbf{x}_{p\times1}$的协方差矩阵为$\Sigma=(\sigma_{ij})$，$\Omega=\Sigma^{-1}=(\omega_{ij})$，

- x_i, x_j的相关系数$\rho_{ij}=\dfrac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$.
- x_i, x_j的偏相关系数 $\rho_{ij\bullet\text{其它}}=-\dfrac{\omega_{ij}}{\sqrt{\omega_{ii}\omega_{jj}}}$，$i\neq j$; $\rho_{ii\bullet\text{其它}}=1$.

这说明随机向量$\mathbf{x}$的任意两个分量的偏相关系数的求解法则类似于相关系数的求法但相差一个符号.

证明：下面对$i=1, j=2$证明.

随机向量$\mathbf{x}=(x_1,\ldots,x_p)^{\mathsf{T}}$，记$\mathbf{w}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$，$\mathbf{z}=(x_3,\ldots,x_p)^{\mathsf{T}}$，

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

$\mathbf{w}$关于$\mathbf{z}$的“不相关化”：$\mathbf{w}^{\perp}=\begin{pmatrix}x_1^{\perp}\\ x_2^{\perp}\end{pmatrix}=\begin{pmatrix}x_1-\Sigma_{1\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z}\\ x_2-\Sigma_{2\mathbf{z}}\Sigma_{\mathbf{zz}}^{-1}\mathbf{z}\end{pmatrix}$

$\mathbf{w}^{\perp}$的2×2方差-协方差矩阵：

$$\operatorname{cov}(\mathbf{w}^{\perp})=\Sigma_{\mathbf{ww}\bullet\mathbf{z}}=\begin{pmatrix}\Sigma_{11\bullet\mathbf{z}} & \Sigma_{12\bullet\mathbf{z}}\\ \Sigma_{21\bullet\mathbf{z}} & \Sigma_{22\bullet\mathbf{z}}\end{pmatrix}\hat{=}\begin{pmatrix}a & b\\ b & c\end{pmatrix},\quad \rho_{12\bullet\mathbf{z}}=\frac{b}{\sqrt{ac}}$$

记$\Omega=\Sigma^{-1}$,注意到分块求逆公式：

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

即Ω左上角的2×2矩阵

$$\begin{pmatrix}\omega_{11} & \omega_{12}\\ \omega_{21} & \omega_{22}\end{pmatrix}=\begin{pmatrix}a & b\\ b & c\end{pmatrix}^{-1}=\frac{1}{ac-b^2}\begin{pmatrix}c & -b\\ -b & a\end{pmatrix}$$

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

高斯图模型

定义：若n维随机向量$\mathbf{x}$的密度为

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

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

命题9(边际与条件分布)：若$\mathbf{x}\sim N_p(\boldsymbol{\mu},\Sigma)$，分块($\mathbf{x}_1$长度为$k$, $\mathbf{x}_2$长度为$p-k$)

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

则 $\mathbf{x}_1\sim N_k(\boldsymbol{\mu}_1,\Sigma_{11}),\mathbf{x}_2\sim N_{n-k}(\boldsymbol{\mu}_2,\Sigma_{22})$, 且

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

key：条件方差 $\mathrm{var}(\mathbf{x}_1\,|\,\mathbf{x}_2)=\Sigma_{11\bullet 2}=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

命题10. 多元正态假设下，偏相关系数等于条件相关系数。

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

$$\text{记}\mathbf{w}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\text{则}\ \mathbf{w}\,|_{\mathbf{z}}\sim N\left(*,\Sigma_{\mathbf{ww}\bullet\mathbf{z}}\right)$$

其中

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

即条件方差 $\operatorname{var}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\Bigg|\,\mathbf{z} = \Sigma_{\mathbf{ww}\bullet\mathbf{z}} = \begin{pmatrix} \Sigma_{11\bullet\mathbf{z}} & \Sigma_{12\bullet\mathbf{z}} \\ \Sigma_{21\bullet\mathbf{z}} & \Sigma_{22\bullet\mathbf{z}} \end{pmatrix}$

则给定$(x_3,..,x_p)=\mathbf{z}$条件下，x_1和x_2的条件相关系数$\rho_{12|\mathbf{z}}$

$$\rho_{12|\mathbf{z}} = \frac{\operatorname{cov}(x_1, x_2 \mid \mathbf{z})}{\sqrt{\operatorname{var}(x_1 \mid \mathbf{z})\operatorname{var}(x_2 \mid \mathbf{z})}} = \frac{\Sigma_{12\bullet\mathbf{z}}}{\sqrt{\Sigma_{11\bullet\mathbf{z}}}\sqrt{\Sigma_{22\bullet\mathbf{z}}}}$$

$$= \rho_{12\bullet\mathbf{z}}$$

命题11：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：由偏相关系数与协方差矩阵逆之间的关系知，

$$\omega_{ij} = 0 \Leftrightarrow \rho_{ij\bullet\text{其它}} = -\frac{\omega_{ij}}{\sqrt{\omega_{ii}\omega_{jj}}} 0 \overset{\text{命题10}}{\Leftrightarrow} \rho_{ij|\text{其它}} = 0, \text{条件不相关}$$

$\Leftrightarrow x_i$与x_j条件独立(正态假设下不相关 $\Leftrightarrow$ 独立)

证明2：不妨设均值$\boldsymbol{\mu}=0$, 设$\omega_{12}=0$, 则$\mathbf{x}$的联合密度

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

其中C是常数，其中c_1,c_2,c_3只与$x_3,\ldots,x_n$有关。

给定$x_3,\ldots,x_n$(为常数)时，x_1,x_2的联合密度

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

$$=\tilde{C}\exp\left(-\frac{1}{2}\omega_{11}x_1^2+c_1x_1\right)\exp\left(-\frac{1}{2}\omega_{22}x_2^2+c_2x_2\right),$$

其中c_1,c_2,c_3(它们只与$x_3,\ldots,x_n$有关)在$x_3,\ldots,x_n$给定的条件下是常数，所以x_1,x_2条件独立。

高斯图模型：

以节点代表多元正态随机向量的各个分量，若$\omega_{ij}=0$，即i, j分量条件独立，则它们之间无连线，否则连线。