Stat 521 notes on partial correlations, by H Joe, October 2015

This document has a collection of results for multivariate normal/Gaussian and partial correlations; they are used in a variety of graphical models.

Notation

- Random variables as denoted in upper case letters, e.g., Y_1, Y_i, Z. Random vectors are denoted as boldfaced upper case, e.g., $\boldsymbol{Z}, \boldsymbol{Y}$.
- Density functions are denoted by f, and cumulative distribution functions (cdf) are denoted by F. Subscripts indicate the random variables or the conditioning; examples are $F_{\boldsymbol{Y}}, F_{\boldsymbol{Y}_2|\boldsymbol{Y}_1}$.
- If Y, Z are random variables and $F_{Y|Z}$ is a conditional cdf, then $F_{Y|Z}(\cdot|z)$ is a conditional cdf mapping $\mathbb{R}$ to $[0,1]$ for any z that is a possible value of Z. Similar if Z is replaced by vector $\boldsymbol{Z}$.
- $\rho_{Y_1Y_2|\boldsymbol{Z}}(\boldsymbol{z})$ is the conditional correlation of the conditional distribution $F_{Y_1Y_2|\boldsymbol{Z}}(\cdot|\boldsymbol{z})$ or $[Y_1, Y_2|\boldsymbol{Z} = \boldsymbol{z}]$. Here $\boldsymbol{Z}$ could be a random variable or random vector.
- $\rho_{Y_1Y_2;\boldsymbol{Z}}$ is the partial correlation based on the formula for $(Y_1, Y_2, \boldsymbol{Z})$ being multivariate Gaussian. $\rho_{Y_1Y_2|\boldsymbol{Z}}(\cdot) \equiv \rho_{Y_1Y_2;\boldsymbol{Z}}$ for multivariate Gaussian, and a few other distributions.
- ϕ is standard normal/Gaussian density, Φ is the corresponding cdf. The normal random variable with mean μ and variance σ^2 is denoted a $\mathrm{N}(\mu, \sigma^2)$.
- ϕ_d, Φ_d are the d-variate Gaussian density and cdf respectively. The d-dimensional Gaussian random vector with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ is denoted as $\mathrm{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

1 Multivariate Gaussian density in partitioned form

1.1 Bivariate

Decomposition of bivariate normal/Gaussian as $f_{Y_1Y_2} = f_{Y_1} f_{Y_2|Y_1}$, where $f_{Y_1Y_2}$ is the bivariate Gaussian density with means μ_1, μ_2, variances σ_1, σ_2 and correlation $\rho \in (-1, 1)$.

With $\rho = \rho_{12}$, the determinant and inverse of the correlation matrix $\boldsymbol{R} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ are $\det(\boldsymbol{R}) = |\boldsymbol{R}| = (1-\rho^2)$ and $\boldsymbol{R}^{-1} = (1-\rho^2)^{-1} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$. The bivariate Gaussian density $\phi_2(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \phi_2(y_1, y_2; \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ can be written as

$$\frac{1}{(2\pi)\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp\{-\tfrac{1}{2}(1-\rho^2)^{-1}[z_1^2 + z_2^2 - 2\rho z_1 z_2]\},$$

where $z_1 = (y_1 - \mu_1)/\sigma_1$ and $z_2 = (y_2 - \mu_2)/\sigma_2$. The term in the exponent simplifies to:

$$(1-\rho^2)^{-1}\{z_1^2 + z_2^2 - 2\rho z_1 z_2\} = z_1^2 + (1-\rho^2)^{-1}\{\rho^2 z_1^2 + z_2^2 - 2\rho z_1 z_2\} = z_1^2 + (1-\rho^2)^{-1}(z_2 - \rho z_1)^2.$$

Hence

$$\begin{aligned}\phi_2(\boldsymbol{y};\boldsymbol{\mu},\sigma_1,\sigma_2,\rho) &= \sigma_1^{-1}\phi(z_1)\cdot\sigma_2^{-1}(1-\rho^2)^{-1/2}(2\pi)^{-1/2}\exp\{-\tfrac{1}{2}(1-\rho^2)^{-1}(z_2-\rho z_1)^2\} \\ &= \sigma_1^{-1}\phi(z_1)\cdot\sigma_2^{-1}(1-\rho^2)^{-1/2}\phi\big((z_2-\rho z_1)/(1-\rho^2)^{1/2}\big)\end{aligned}$$

and

$$\frac{(z_2-\rho z_1)}{(1-\rho^2)^{1/2}} = \sigma_2^{-1}(1-\rho^2)^{-1/2}\Big(y_2-\mu_2-\frac{\rho\sigma_2(y_1-\mu_1)}{\sigma_1}\Big).$$

Since in general,

$$f_{Y_1,Y_2}(y_1,y_2) = f_{Y_1}(y_1)\, f_{Y_2|Y_1}(y_2|y_1),$$

we can identify $f_{Y_2|Y_1}(y_2|y_1)$ as a $\mathrm{N}(\mu_2+\rho\sigma_2(y_1-\mu_1)/\sigma_1,\ \sigma_2^2(1-\rho^2))$ density.

1.2 Inverse of partitioned matrix

Matrix inverse for partitioned matrix (verify by multiplying the inverse by the original matrix).

For a partitioned positive definite (or invertible) symmetric matrix,

$$\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}^{-1} = \begin{pmatrix}\boldsymbol{\Sigma}_{11;2}^{-1} & -\boldsymbol{\Sigma}_{11;2}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11;2}^{-1} & \boldsymbol{\Sigma}_{22}^{-1}+\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11;2}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\end{pmatrix}, \tag{1}$$

where $\boldsymbol{\Sigma}_{11;2}=\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. By symmetry (interchange subscripts 1,2), one can also write

$$\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}^{-1} = \begin{pmatrix}\boldsymbol{\Sigma}_{11}^{-1}+\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22;1}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & -\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22;1}^{-1}\\ -\boldsymbol{\Sigma}_{22;1}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} & \boldsymbol{\Sigma}_{22;1}^{-1}\end{pmatrix}, \tag{2}$$

where $\boldsymbol{\Sigma}_{22;1}=\boldsymbol{\Sigma}_{22}-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$.

The above and the determinant result of $|\boldsymbol{\Sigma}|=|\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22;1}|$ can be derived from the following (see page 13, Muirhead (1982), *Aspects of Multivariate Statistical Theory*). Let $\boldsymbol{B}$ be a $d\times d$ (not necessarily symmetric) matrix with partitioned components with k and $d-k$ rows/columns, and suppose $\boldsymbol{B}_{22}$ is invertible. Then

$$\begin{pmatrix}\boldsymbol{I}_k & -\boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1}\\ \boldsymbol{0} & \boldsymbol{I}_{d-k}\end{pmatrix}\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_k & \boldsymbol{0}\\ -\boldsymbol{B}_{22}^{-1}\boldsymbol{B}_{21} & \boldsymbol{I}_{d-k}\end{pmatrix} = \begin{pmatrix}\boldsymbol{B}_{11}-\boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1}\boldsymbol{B}_{21} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{B}_{22}\end{pmatrix}.$$

1.3 Multivariate

Decomposition of multivariate Gaussian density: $f_{\boldsymbol{Y}_1\boldsymbol{Y}_2}=f_{\boldsymbol{Y}_1}f_{\boldsymbol{Y}_2|\boldsymbol{Y}_1}$, where the multivariate Gaussian random vector has been partitioned as $\boldsymbol{Y}_1,\boldsymbol{Y}_2$.

Partition $\boldsymbol{\Sigma}^{-1}$ as $\begin{pmatrix}\boldsymbol{\Sigma}^{11} & \boldsymbol{\Sigma}^{12}\\ \boldsymbol{\Sigma}^{21} & \boldsymbol{\Sigma}^{22}\end{pmatrix}$, and let $\boldsymbol{z}_1=\boldsymbol{y}_1-\boldsymbol{\mu}_1$, $\boldsymbol{z}_2=\boldsymbol{y}_2-\boldsymbol{\mu}_2$. With the partitioning, then using matrix inverse of a partitioned matrix in (2), the exponent of the multivariate Gaussian density is

$$\begin{aligned}&(\boldsymbol{y}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}) = \boldsymbol{z}_1^\top\boldsymbol{\Sigma}^{11}\boldsymbol{z}_1+\boldsymbol{z}_1^\top\boldsymbol{\Sigma}^{12}\boldsymbol{z}_2+\boldsymbol{z}_2^\top\boldsymbol{\Sigma}^{21}\boldsymbol{z}_1+\boldsymbol{z}_2^\top\boldsymbol{\Sigma}^{22}\boldsymbol{z}_2\\ &= \boldsymbol{z}_1^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1+\boldsymbol{z}_1^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22;1}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1-\boldsymbol{z}_1^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22;1}^{-1}\boldsymbol{z}_2-\boldsymbol{z}_2^\top\boldsymbol{\Sigma}_{22;1}^{-1}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1\\ &\qquad+\boldsymbol{z}_2^\top\boldsymbol{\Sigma}_{22;1}^{-1}\boldsymbol{z}_2\\ &= \boldsymbol{z}_1^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1+(\boldsymbol{z}_2-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1)^\top\boldsymbol{\Sigma}_{22;1}^{-1}(\boldsymbol{z}_2-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{z}_1).\end{aligned}$$

Since in general,

$$f_{\boldsymbol{Y}_1,\boldsymbol{Y}_2}(\boldsymbol{y}_1,\boldsymbol{y}_2) = f_{\boldsymbol{Y}_1}(\boldsymbol{y}_1)\, f_{\boldsymbol{Y}_2|\boldsymbol{Y}_1}(\boldsymbol{y}_2|\boldsymbol{y}_1),$$

the multiplicative constant terms of $f_{\boldsymbol{Y}_1,\boldsymbol{Y}_2}, f_{\boldsymbol{Y}_1}, f_{\boldsymbol{Y}_2|\boldsymbol{Y}_1}$ must satisfy:

$$(2\pi)^{d/2}|\boldsymbol{\Sigma}|^{1/2} = (2\pi)^{d_1/2}|\boldsymbol{\Sigma}_{11}|^{1/2} \cdot (2\pi)^{d_2/2}|\boldsymbol{\Sigma}_{22;1}|^{1/2}.$$

Hence

$$\phi_d(\boldsymbol{y};\boldsymbol{\mu},\boldsymbol{\Sigma}) = \phi_{d_1}(\boldsymbol{y}_1;\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_{22}) \cdot \phi_{d_2}\big(\boldsymbol{y}_2;\boldsymbol{\mu}_{2|1}(\boldsymbol{y}_1),\boldsymbol{\Sigma}_{22;1}\big),$$

and we can identify $f_{\boldsymbol{Y}_2|\boldsymbol{Y}_1}(\boldsymbol{y}_2|\boldsymbol{y}_1)$ as a $\mathrm{N}_{d_2}(\boldsymbol{\mu}_{2|1}(\boldsymbol{y}_1),\ \boldsymbol{\Sigma}_{22;1})$ density, where $\boldsymbol{\mu}_{2|1}(\boldsymbol{y}_1) = \boldsymbol{\mu}_2+\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{y}_1-\boldsymbol{\mu}_1)$.

2 Partial correlation and covariance

The partial correlation for multivariate Gaussian is defined as a conditional correlation.

2.1 Definition

Let $(\boldsymbol{Y}_1^T,\boldsymbol{Y}_2^T)^T \sim \mathrm{N}_d(\boldsymbol{0},\boldsymbol{\Sigma})$ with subvectors of size $d-2$ and 2, and let $\boldsymbol{\Sigma}_{11},\boldsymbol{\Sigma}_{12},\boldsymbol{\Sigma}_{21},\boldsymbol{\Sigma}_{22}$ be blocks (submatrices) of $\boldsymbol{\Sigma}$. Let $\boldsymbol{\Sigma}_{22;1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}$ be the covariance matrix of conditional distribution of $[\boldsymbol{Y}_2|\boldsymbol{Y}_1 = \boldsymbol{y}_1]$. Suppose $d \geq 3$ and $\boldsymbol{\Sigma}_{11}$ has dimension $(d-2)\times(d-2)$. From the 2×2 covariance matrix $\boldsymbol{\Sigma}_{22;1}$, obtain the partial correlation $\rho_{d-1,d;1\cdots d-2}$. With $\boldsymbol{a}_{d-1} = (\sigma_{1,d-1},\ldots,\sigma_{d-2,d-1})^\top$ and $\boldsymbol{a}_d = (\sigma_{1,d},\ldots,\sigma_{d-2,d})^\top$, then

$$\boldsymbol{\Sigma}_{22;1} = \begin{pmatrix} \sigma_{d-1,d-1} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1} & \sigma_{d-1,d} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d \\ \sigma_{d,d-1} - \boldsymbol{a}_d^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1} & \sigma_{dd} - \boldsymbol{a}_d^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d \end{pmatrix}.$$

The conditional covariance is $\sigma_{d-1,d;1\cdots d-2} = \sigma_{d-1,d} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d$. The conditional correlation is

$$\rho_{d-1,d;1\cdots d-2} = \rho_{Y_{d-1},Y_d;Y_1\cdots Y_{d-2}} = \frac{\sigma_{d-1,d} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d}{\big[\sigma_{d-1,d-1} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1}\big]^{1/2}\big[\sigma_{dd} - \boldsymbol{a}_d^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d\big]^{1/2}}.$$

If $\boldsymbol{\Sigma} = \boldsymbol{R}$ is a correlation matrix, then in the above, $\sigma_{d-1,d} = \rho_{d-1,d}$ and $\sigma_{d-1,d-1} = \sigma_{dd} = 1$. Other partial correlations can be obtained by permutating indices and working with submatrices of $\boldsymbol{\Sigma}$.

For the special case of $d-2=1$ or $d=3$ with unit variances, there is a simple resulting formula:

$$\begin{aligned} \boldsymbol{\Sigma}_{22;1} &= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} = \begin{pmatrix} 1 & \rho_{23} \\ \rho_{32} & 1 \end{pmatrix} - \begin{pmatrix} \rho_{21} \\ \rho_{31} \end{pmatrix} 1^{-1} \begin{pmatrix} \rho_{12} & \rho_{13} \end{pmatrix} = \begin{pmatrix} 1-\rho_{12}^2 & \rho_{23}-\rho_{21}\rho_{13}, \\ \rho_{32}-\rho_{31}\rho_{12} & 1-\rho_{13}^2 \end{pmatrix} \\ \rho_{23;1} &= \frac{\rho_{23}-\rho_{12}\rho_{13}}{\sqrt{(1-\rho_{12}^2)(1-\rho_{13}^2)}}. \end{aligned}$$

There is a recursive formula for partial correlations as more variables are conditioned on. This is written in the form with indices that are subsets of $\{1,\dots,d\}$.

$$\begin{aligned}
\rho_{jk;g} &= \frac{\rho_{jk}-\rho_{jg}\rho_{kg}}{\sqrt{(1-\rho_{jg}^2)(1-\rho_{kg}^2)}},\\
\rho_{jk;ga} &= \frac{\rho_{jk;a}-\rho_{jg;a}\rho_{kg;a}}{\sqrt{(1-\rho_{jg;a}^2)(1-\rho_{kg;a}^2)}},\\
\rho_{jk;gT} &= \frac{\rho_{jk;T}-\rho_{jg;T}\rho_{kg;T}}{\sqrt{(1-\rho_{jg;T}^2)(1-\rho_{kg;T}^2)}},
\end{aligned}$$

where T can consist of indices of more than one variable (or none).

2.2 Proof of recursion

The proof of the recursion in given in Anderson (1958), *An Introduction to Multivariate Statistical Analysis*, Wiley. The result is given below with the notation used in this document.

Let $\boldsymbol{Z}_g, \boldsymbol{Z}_a, \boldsymbol{Z}_h$ be subvectors of a multivariate Gaussian random vector (g for a given set of indices, a for additional indices to condition on, h for the remaining indices).

Let the covariance matrix be $\begin{pmatrix}\boldsymbol{\Sigma}_{gg} & \boldsymbol{\Sigma}_{ga} & \boldsymbol{\Sigma}_{gh}\\ \boldsymbol{\Sigma}_{ag} & \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ah}\\ \boldsymbol{\Sigma}_{hg} & \boldsymbol{\Sigma}_{ha} & \boldsymbol{\Sigma}_{hh}\end{pmatrix}$. The covariance matrix of $\boldsymbol{Z}_h$ given $\boldsymbol{Z}_g=\boldsymbol{z}_g$, $\boldsymbol{Z}_a=\boldsymbol{z}_a$ is

$$\boldsymbol{\Sigma}_{hh;ga} = \boldsymbol{\Sigma}_{hh} - \begin{pmatrix}\boldsymbol{\Sigma}_{hg} & \boldsymbol{\Sigma}_{ha}\end{pmatrix}\begin{pmatrix}\boldsymbol{\Sigma}_{gg} & \boldsymbol{\Sigma}_{ga}\\ \boldsymbol{\Sigma}_{ag} & \boldsymbol{\Sigma}_{aa}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\Sigma}_{gh}\\ \boldsymbol{\Sigma}_{ah}\end{pmatrix}$$

Also, $[\boldsymbol{Z}_a,\boldsymbol{Z}_h|\boldsymbol{Z}_g=\boldsymbol{z}_g]$ is multivariate Gaussian, with mean $\begin{pmatrix}\boldsymbol{\Sigma}_{ag}\boldsymbol{\Sigma}_{gg}^{-1}\boldsymbol{z}_g & \boldsymbol{\Sigma}_{hg}\boldsymbol{\Sigma}_{gg}^{-1}\boldsymbol{z}_g\end{pmatrix}$ and covariance matrix $\boldsymbol{\Sigma}_{;g}=\begin{pmatrix}\boldsymbol{\Sigma}_{aa;g} & \boldsymbol{\Sigma}_{ah;g}\\ \boldsymbol{\Sigma}_{ha;g} & \boldsymbol{\Sigma}_{hh;g}\end{pmatrix}$. Condition on $\boldsymbol{Z}_a=\boldsymbol{z}_a$ with this distribution. Then for $[\boldsymbol{Z}_h|\boldsymbol{Z}_g=\boldsymbol{z}_g,\boldsymbol{Z}_a=\boldsymbol{z}_a]$, the mean is $\boldsymbol{\Sigma}_{hg}\boldsymbol{\Sigma}_{gg}^{-1}\boldsymbol{z}_g+\boldsymbol{\Sigma}_{ha;g}\boldsymbol{\Sigma}_{aa;g}^{-1}(\boldsymbol{z}_a-\boldsymbol{\Sigma}_{ag}\boldsymbol{\Sigma}_{gg}^{-1}\boldsymbol{z}_g)$ and the covariance matrix $\boldsymbol{\Sigma}_{hh;g}-\boldsymbol{\Sigma}_{ha;g}\boldsymbol{\Sigma}_{aa;g}^{-1}\boldsymbol{\Sigma}_{ah;g}$. Hence we must have

$$\boldsymbol{\Sigma}_{hh;ga} = \boldsymbol{\Sigma}_{hh} - \begin{pmatrix}\boldsymbol{\Sigma}_{hg} & \boldsymbol{\Sigma}_{ha}\end{pmatrix}\begin{pmatrix}\boldsymbol{\Sigma}_{gg} & \boldsymbol{\Sigma}_{ga}\\ \boldsymbol{\Sigma}_{ag} & \boldsymbol{\Sigma}_{aa}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{\Sigma}_{gh}\\ \boldsymbol{\Sigma}_{ah}\end{pmatrix} = \boldsymbol{\Sigma}_{hh;g}-\boldsymbol{\Sigma}_{ha;g}\boldsymbol{\Sigma}_{aa;g}^{-1}\boldsymbol{\Sigma}_{ah;g}. \tag{3}$$

Consider a variable $\boldsymbol{Z}_j$ as a component of $\boldsymbol{Z}_h$, and suppose Z_a is a single variable. Then since $[(\boldsymbol{Z}_h, Z_a)|\boldsymbol{Z}_g=\boldsymbol{z}_g]$ is multivariate Gaussian,

$$\sigma_{jj;ag} := \mathrm{Var}(Z_j|Z_a=z_a,\boldsymbol{Z}_g=\boldsymbol{z}_g) = \mathrm{Var}(Z_j|\boldsymbol{Z}_g=\boldsymbol{z}_g)[1-\{\mathrm{Cor}(Z_j,Z_a|\boldsymbol{Z}_g=\boldsymbol{z}_g)\}^2] = \sigma_{jj;g}(1-\rho_{aj;g}^2). \tag{4}$$

Now assume h has cardinality of at least 2, and let j,k be two different components of h. $\sigma_{jk;ga}$ is the conditional covariance is Z_j,Z_k given $\boldsymbol{Z}_g=\boldsymbol{z}_g, Z_a=z_a$.

If Z_a is a single variable (a has cardinality 1), an off-diagonal element of (3) leads to:

$$\sigma_{jk;ga} = \sigma_{jk;a} - \sigma_{ja;g}\sigma_{ak;g}/\sigma_{aa;g}.$$

The partial correlation is

$$\rho_{jk;ga} = \sigma_{jk;ga}\big/\sqrt{\sigma_{jj;ag}\sigma_{kk;ag}}\,.$$

Substitute in (4) twice to get:

$$\rho_{jk;ga} = \frac{\sigma_{jk;g} - \sigma_{ja;g}\sigma_{ak;g}/\sigma_{aa;g}}{\sqrt{\sigma_{jj;g}(1-\rho^2_{aj;g}) \cdot \sigma_{kk;g}(1-\rho^2_{ak;g})}} = \frac{\rho_{jk;g} - \rho_{ja;g}\rho_{ak;g}}{\sqrt{(1-\rho^2_{aj;g})(1-\rho^2_{ak;g})}}.$$

3 Precision matrix of multivariate Gaussian

The precision matrix is the inverse of the covariance matrix $\boldsymbol{\Sigma}$. Some properties of $\boldsymbol{\Sigma}^{-1}$ with interpretations of partial correlations are given below.

For notation, $\rho_{Y_j,Y_k;\{Y_i:i\neq j,i\neq k\}}$ is abbreviated to $\rho_{jk;\text{rest}}$.

3.1 Partial correlation of two variables given the remaining variables

Let $d = m + q$ where $m \geq 2$ and $q \geq 2$. Suppose $\boldsymbol{\Sigma}_{11}$ is $m \times m$, $\boldsymbol{\Sigma}_{12}$ is $m \times q$, $\boldsymbol{\Sigma}_{21}$ is $q \times m$ and $\boldsymbol{\Sigma}_{22}$ is $q \times q$.

Let $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}$. Solve the equations

$$\begin{aligned}
\boldsymbol{\Sigma}_{11}\boldsymbol{A}_{11} + \boldsymbol{\Sigma}_{12}\boldsymbol{A}_{21} &= \boldsymbol{I}_m, \\
\boldsymbol{\Sigma}_{11}\boldsymbol{A}_{12} + \boldsymbol{\Sigma}_{12}\boldsymbol{A}_{22} &= \boldsymbol{0}, \\
\boldsymbol{\Sigma}_{21}\boldsymbol{A}_{11} + \boldsymbol{\Sigma}_{22}\boldsymbol{A}_{21} &= \boldsymbol{0}, \\
\boldsymbol{\Sigma}_{21}\boldsymbol{A}_{12} + \boldsymbol{\Sigma}_{22}\boldsymbol{A}_{22} &= \boldsymbol{I}_q.
\end{aligned}$$

From the second equation, $A_{12} = -\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{A}_{22}$. Plug into the fourth equation to get

$$\boldsymbol{A}_{22}^{-1} = [\boldsymbol{\Sigma}_{22}^{-1} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}] = \boldsymbol{\Sigma}_{22;1}$$

To get the partial correlation of variables $d-1$ and d given the rest, let $m = d-2$ and $q = 2$. Write $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1} = (\sigma^{jk})$. The entries of $\boldsymbol{A}_{22}^{-1}$ are

$$\begin{aligned}
\begin{pmatrix} \sigma^{d-1,d-1} & \sigma^{d-1,d} \\ \sigma^{d,d-1} & \sigma^{dd} \end{pmatrix}^{-1} &= \left[\sigma^{d-1,d-1}\sigma^{dd} - (\sigma^{d-1,d})^2\right]^{-1} \begin{pmatrix} \sigma^{dd} & -\sigma^{d-1,d} \\ -\sigma^{d,d-1} & \sigma^{d-1,d-1} \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \mathrm{Var}(Y_{d-1}|\boldsymbol{Y}_{\text{rest}}) & \mathrm{Cov}(Y_{d-1},Y_d|\boldsymbol{Y}_{\text{rest}}) \\ \mathrm{Cov}(Y_d,Y_{d-1}|\boldsymbol{Y}_{\text{rest}}) & \mathrm{Var}(Y_d|\boldsymbol{Y}_{\text{rest}}) \end{pmatrix}.
\end{aligned}$$

Convert the two covariance matrices to get a partial correlation:

$$\rho_{d-1,d;\text{rest}} = \frac{\mathrm{Cov}(Y_{d-1},Y_d|\boldsymbol{Y}_{\text{rest}})}{[\mathrm{Var}(Y_{d-1}|\boldsymbol{Y}_{\text{rest}})\ \mathrm{Var}(Y_d|\boldsymbol{Y}_{\text{rest}})]^{1/2}} = \frac{-\sigma^{d-1,d}}{[\sigma^{d-1,d-1}\sigma^{dd}]^{1/2}}.$$

By permuting indices, other partial correlations $\rho_{jk;\text{rest}}$ can be obtained from the precision matrix.

IMPORTANT: All of the partial correlations given remaining variables can be obtained by converting $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1}$ into a correlation matrix and changing the signs of the off-diagonal elements. The result also implies that σ^{jk} is opposite in sign to $\rho_{jk;\text{rest}}$ for all $j \neq k$.

3.2 Partial correlation and cofactor

Partial correlation $\rho_{jk;\text{rest}}$ in terms of cofactor. This shows that signs of off-diagonal of precision matrix are opposite in sign to the corresponding partial correlation.

Let $\boldsymbol{A} = (a_{jk})$ be the inverse of the d-dimensional correlation or covariance matrix $\boldsymbol{\Sigma}$. Let $\boldsymbol{\Sigma}^{(uv)}$ be the matrix obtained from $\boldsymbol{\Sigma}$ by removing the uth, vth rows and columns. Let $\sigma_{jk\cdot\text{rest}}$ be the partial covariance of the jth and kth variables given the remaining $d-2$ variables. Then $a_{jk} = -\sigma_{jk\cdot\text{rest}}|\boldsymbol{\Sigma}^{(jk)}||\boldsymbol{\Sigma}|^{-1}$ for $j \neq k$.

Without loss of generality, prove this for $(j,k) = (d-1,d)$.
$a_{d-1,d} = (-1)^{2d-1}|M_{d-1,d}|/|\boldsymbol{\Sigma}|$ where M_{d-1},d is the matrix with row $d-1$ and column d removed from $\boldsymbol{\Sigma}$. Then $|M_{d-1,d}| = |\boldsymbol{\Sigma}^{(d-1,d)}| \cdot \sigma_{d-1,d\cdot\text{rest}}$.

3.3 Another identity

The following is a result from pp 32–33 of Kurowicka and Cooke (2006), *Uncertainty Analysis with High Dimensional Dependence Modelling*:

$$\rho_{d-1,d;\text{rest}} = \frac{-\text{cofactor}_{d-1,d}}{[\text{cofactor}_{d-1,d-1}\text{cofactor}_{dd}]^{1/2}}.$$

The cofactor cofactor_{ij} is the determinant of the (covariance) matrix $\boldsymbol{\Sigma}$ with the ith row and jth column deleted multiplied by $(-1)^{i+j}$.

For the proof, decompose

$$\begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{a}_{d-1} & \boldsymbol{a}_d \\ \boldsymbol{a}_{d-1}^\top & \sigma_{d-1,d-1} & \sigma_{d-1,d} \\ \boldsymbol{a}_d^\top & \sigma_{d,d-1} & \sigma_{dd} \end{pmatrix} \tag{5}$$

Then

$$\text{cofactor}_{dd} = |\boldsymbol{\Sigma}_{11}|[\sigma_{d-1,d-1} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1}],$$

$$\text{cofactor}_{d-1,d-1} = |\boldsymbol{\Sigma}_{11}|[\sigma_{d,d} - \boldsymbol{a}_d^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d],$$

$$\text{cofactor}_{d-1,d} = -|\boldsymbol{\Sigma}_{11}|[\sigma_{d-1,d} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d],$$

Then the above ratio is

$$\frac{\sigma_{d-1,d} - \boldsymbol{a}_d^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1}}{[\sigma_{d-1,d-1} - \boldsymbol{a}_{d-1}^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1}]^{1/2}[\sigma_{d,d} - \boldsymbol{a}_d^\top\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d]^{1/2}},$$

which matches the previous definition of partial correlation.

One result that is used above is the following. Let $\boldsymbol{A}_{11}$ be an invertible matrix, $\boldsymbol{a}_{12}$ be a row vector and $\boldsymbol{a}_{21}$ be a column vector and a_{22} be a scalar in the partitioned matrix: $\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & a_{22} \end{pmatrix}$. Then with the determinant being invariant to elementary row operations,

$$\det\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & a_{22} \end{pmatrix} = \det\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{0}^\top & a_{22} - \boldsymbol{a}_{12}\boldsymbol{A}_{11}^{-1}\boldsymbol{a}_{21} \end{pmatrix} = \det(A_{11})(a_{22} - \boldsymbol{a}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{a}_{12}).$$

3.4 Conditional multivariate Gaussian and precision matrix

The conditional distributions of multivariate Gaussian can be expressed in terms of quantities of the precision matrix. $\boldsymbol{Y}_{-i}$ is used to denote $\boldsymbol{Y}_{\{1\cdots d\}\backslash i}$.

Let $\boldsymbol{Y} \sim \mathrm{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Let $\boldsymbol{A} = \boldsymbol{\Sigma}^{-1} = (a_{ij})$ be the precision matrix. The follow result appears in the literature in graphical models and Gaussian Markov random fields.

Then (a) the conditional expectation of Y_i given $\boldsymbol{Y}_{-i} = \boldsymbol{y}_{-i}$ is $\mu_i - a_{ii}^{-1} \sum_{j\neq i} a_{ij}(y_j - \mu_j)$, and (b) the conditional variance is a_{ii}^{-1}.

Proof for the case of $i = d$ for (a) and (b). Partition $\boldsymbol{A}$ into $\begin{pmatrix} \boldsymbol{A}_{-d,-d} & \boldsymbol{\gamma} \\ \boldsymbol{\gamma}^\top & a_{dd} \end{pmatrix}$.

Let $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{\mu}$. The quadratic form in the joint density is:

$$(\boldsymbol{y} - \boldsymbol{\mu})^\top \boldsymbol{A}(\boldsymbol{y} - \boldsymbol{\mu}) = \boldsymbol{z}_{-d}^\top \boldsymbol{A}_{-d,-d}\boldsymbol{z}_{-d} + \boldsymbol{z}_{-d}^\top \boldsymbol{\gamma} z_d + z_d \boldsymbol{\gamma}^\top \boldsymbol{z}_{-d} + a_{dd} z_d^2.$$

Complete the square for the terms involving z_d to get

$$\begin{aligned} & a_{dd}\big(z_d^2 + 2a_{dd}^{-1} z_d \boldsymbol{\gamma}^\top \boldsymbol{z}_{-d} + a_{dd}^{-2} \boldsymbol{z}_{-d}^\top \boldsymbol{\gamma}\boldsymbol{\gamma}^\top \boldsymbol{z}_{-d}\big) + \boldsymbol{z}_{-d}^\top[\boldsymbol{A}_{-d,-d} - a_{dd}^{-2}\boldsymbol{\gamma}\boldsymbol{\gamma}^\top]\boldsymbol{z}_{-d} \\ &= a_{dd}(z_d + a_{dd}^{-1}\boldsymbol{\gamma}^\top \boldsymbol{z}_{-d})^2 + \boldsymbol{z}_{-d}^\top[\boldsymbol{A}_{-d,-d} - a_{dd}^{-2}\boldsymbol{\gamma}\boldsymbol{\gamma}^\top]\boldsymbol{z}_{-d}. \end{aligned}$$

The term $\boldsymbol{z}_{-d}^\top[\boldsymbol{A}_{-d,-d} - a_{dd}^{-2}\boldsymbol{\gamma}\boldsymbol{\gamma}^\top]\boldsymbol{z}_{-d}$ is absorbed into the marginal density of $\boldsymbol{Y}_{-d}$. Hence the conditional mean of Y_d given $\boldsymbol{Y}_{-d} = \boldsymbol{y}_{-d}$ is

$$\mu_d - a_{dd}^{-1}\boldsymbol{\gamma}^\top \boldsymbol{z}_{-d} = \mu_d - a_{dd}^{-1} \sum_{j\neq d} a_{jd}(y_j - \mu_j)$$

and the conditional variance of Y_d given $\boldsymbol{Y}_{-d} = \boldsymbol{y}_{-d}$ is a_{dd}^{-1} (the precision is a_{dd}).

4 Regression

Suppose $(Y, W_1, \ldots, W_k)$ are jointly multivariate Gaussian where $k \geq 2$. Consider regression of Y on $\boldsymbol{W} = (W_1, \ldots, W_k)^\top$ as write as

$$Y = \beta_0 + \beta_1 W_1 + \cdots + \beta_k W_k + E,$$

where E is independent of $\boldsymbol{W}$. The variances and covariances are written as $\sigma_{YY}, \boldsymbol{\Sigma}_{Y,\boldsymbol{W}} \boldsymbol{\Sigma}_{\boldsymbol{W}}, \boldsymbol{\Sigma}_{\boldsymbol{W},Y}$, where $\sigma_{YY} = \mathrm{Var}(Y)$, $\boldsymbol{\Sigma}_{\boldsymbol{W}}$ is the covariance matrix of $\boldsymbol{W} = (W_1, \ldots, W_k)$ and $\boldsymbol{\Sigma}_{\boldsymbol{W},Y}$ is the vector of covariances: $\mathrm{Cov}(W_j, Y)$ The following results will be derived:

(a) the sign of the β's in terms of partial correlations, and an equation for β_j in terms of $\rho_{YW_j;\boldsymbol{W}_{-j}}$ where $\boldsymbol{W}_{-j}$ is the vector $\boldsymbol{W}$ with W_j omitted;

(b) the variance of the residual E or the conditional variance of Y given $\boldsymbol{W}$;

(c) the multiple correlation coefficient R^2 of the regression.

Proof of (a):

Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^\top$. Then $\boldsymbol{\beta} = \boldsymbol{\Sigma}_{\boldsymbol{W}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{W},Y}$, for $j = 1, \ldots, k$.

Partition $\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{W}}, \boldsymbol{\Sigma}_{\boldsymbol{W},Y}$ into a component the first variable W_1 and a component for the rest $(W_2, \ldots, W_k)$; subscripts are 1 and 2.

$$\begin{pmatrix} \beta_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{1Y} \\ \boldsymbol{\Sigma}_{2Y} \end{pmatrix}.$$

Let $\sigma_{11;2} = \sigma_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$. Then from the identity for the matrix inverse (see (1)),

$$\beta_1 = \sigma_{11;2}^{-1}\sigma_{1Y} - \sigma_{11;2}^{-1}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{2Y} = \sigma_{11;2}^{-1}[\sigma_{1Y} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{2Y}].$$

It remains to show that $\sigma_{1Y} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{2Y}$ has the same sign as $\rho_{W_1Y;W_2,\cdots,W_k}$.

For the partition of the covariance matrix that includes Y, write

$$\begin{pmatrix} \sigma_{YY} & \sigma_{Y1} & \boldsymbol{\Sigma}_{Y2} \\ \sigma_{1Y} & \sigma_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{2Y} & \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Then, from Section 2.1, the partial covariance of Y, W_1 given $W_2, \ldots, W_k$ is equal to

$$\sigma_{1Y} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{2Y}.$$

The partial correlation has the square roots of

$$\sigma_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} = \sigma_{11;2}, \quad \sigma_{YY} - \boldsymbol{\Sigma}_{Y2}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{2Y}$$

in the denominator. Hence

$$\beta_1 = \frac{\rho_{W_1Y;W_2,\cdots,W_k}(\sigma_{YY} - \boldsymbol{\Sigma}_{Y2}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{2Y})^{1/2}}{\sigma_{11;2}^{1/2}}. \tag{6}$$

Permute the indices to get a similar result for β_j and $\rho_{YW_j;\boldsymbol{W}_{-j}}$. This completes (a).

Proof of (b).

Since E is independent of $\boldsymbol{W}$, $\sigma_{YY} = \mathrm{Var}(Y) = \boldsymbol{\beta}^T\boldsymbol{\Sigma}_{\boldsymbol{W}}\boldsymbol{\beta} + \mathrm{Var}(E)$ or

$$\mathrm{Var}(E) = \sigma_{YY} - \boldsymbol{\Sigma}_{Y,\boldsymbol{W}}\boldsymbol{\Sigma}_{\boldsymbol{W}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{W},Y}.$$

If $Y, \boldsymbol{W}$ have been standardized to have mean 1, write

$$\psi^2 := 1 - \boldsymbol{R}_{Y,\boldsymbol{W}}\boldsymbol{R}_{\boldsymbol{W}}^{-1}\boldsymbol{R}_{\boldsymbol{W},Y}, \tag{7}$$

where the $\boldsymbol{R}$ vectors and matrices consist of correlations.

Proof of (c).

The multiple $R^2_{Y\cdot\boldsymbol{W}}$ is the square of maximum correlation between Y and linear combinations of $\boldsymbol{W}$, or the square of the correlation of Y and $\boldsymbol{\beta}^T\boldsymbol{W}$ with $\boldsymbol{\beta} = \boldsymbol{\Sigma}_{\boldsymbol{W}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{W},Y}$. Hence

$$R^2_{Y\cdot\boldsymbol{W}} = \frac{(\boldsymbol{\beta}^T\boldsymbol{\Sigma}_{\boldsymbol{W},Y})^2}{\sigma_{YY}\cdot\boldsymbol{\Sigma}_{Y,\boldsymbol{W}}\boldsymbol{\Sigma}_{\boldsymbol{W}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{W},Y}} = \frac{\boldsymbol{\Sigma}_{Y,\boldsymbol{W}}\boldsymbol{\Sigma}_{\boldsymbol{W}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{W},Y}}{\sigma_{YY}} = \boldsymbol{R}_{Y,\boldsymbol{W}}\boldsymbol{R}_{\boldsymbol{W}}^{-1}\boldsymbol{R}_{\boldsymbol{W},Y} = 1-\psi^2. \tag{8}$$

□

Another result following from Section 3.4 is the expression of regression coefficients in terms of the precision or inverse covariance matrix.

Let $\boldsymbol{\Sigma} = \begin{pmatrix}\sigma_{YY} & \boldsymbol{\Sigma}_{Y,\boldsymbol{W}}\\ \boldsymbol{\Sigma}_{\boldsymbol{W},Y} & \boldsymbol{\Sigma}_{\boldsymbol{W}}\end{pmatrix}$ be the covariance matrix of $\begin{pmatrix}Y\\ \boldsymbol{W}\end{pmatrix}$. Let the precision matrix be denoted as $\boldsymbol{\Sigma}^{-1} = \begin{pmatrix}a_{YY} & \boldsymbol{A}_{Y,\boldsymbol{W}}\\ \boldsymbol{A}_{\boldsymbol{W},Y} & \boldsymbol{A}_{\boldsymbol{W}}\end{pmatrix}$ and write $\boldsymbol{A}_{Y,\boldsymbol{W}} = (a_{YW_1},\ldots,a_{YW_k})$.

Then for the regression of Y on $\boldsymbol{W}$, the regression coefficient of W_j is $-a_{YW_j}/a_{YY}$.

Note that all results in this section hold for the sample counterparts with multiple regression, where population covariance matrices are replaced with sample covariance matrices and population partial correlations are replaced by sample partial correlations. The sample counterparts do not assume multivariate Gaussian.

4.1 Partial correlation in terms of regression coefficients

A result on pp 32–33 of Kurowicka and Cooke (2006) is the following.

The partial correlation has an expression that comes from regressing Y_d versus $Y_1,\ldots,Y_{d-1}$ and regressing Y_{d-1} versus $Y_1,\ldots,Y_{d-2},Y_d$. For the first regression, let $\beta_{d,d-1}$ be the regression coefficient of Y_{d-1}. For the second regression, let $\beta_{d-1,d}$ be the regression coefficient of Y_d.

Then

$$\rho_{d-1,d;\text{rest}} = \text{sign}\,(\beta_{d,d-1})\sqrt{\beta_{d,d-1}\beta_{d-1,d}}\,. \tag{9}$$

Using the partition in (5), by converting the result in (6), one gets:

$$\begin{aligned}
\beta_{d,d-1} &= \rho_{d-1,d;\text{rest}}\frac{\sigma_{dd}-\boldsymbol{a}_d^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d}{\sigma_{d-1,d-1,}-\boldsymbol{a}_{d-1}^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1}},\\
\beta_{d-1,d} &= \rho_{d-1,d;\text{rest}}\frac{\sigma_{d-1,d-1,}-\boldsymbol{a}_{d-1}^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_{d-1}}{\sigma_{dd}-\boldsymbol{a}_d^T\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{a}_d},\\
\beta_{d,d-1}\beta_{d-1,d} &= \rho^2_{d-1,d;\text{rest}}.
\end{aligned}$$

Hence (9) follows.

5 Determinant of a correlation matrix

Consider $(Y_1,\ldots,Y_d)\sim \mathrm{N}_d(\boldsymbol{0},\boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}=\boldsymbol{R}$ is a covariance matrix. The joint density can be decomposed as a product of conditional densities

$$f_{1:d} = f_1\prod_{j=2}^{d} f_{j|1:(j-1)}.$$

Since each density on the right-hand side is univariate Gaussian, say $f_{j|1:(j-1)}$ is $\mathrm{N}(\boldsymbol{\beta}_j^T(y_1,\ldots,y_{j-1}),\psi_j^2)$, then by matching the normalizing constants on the two sides

$$(2\pi)^{-d/2}|\boldsymbol{R}|^{-1/2} = (2\pi)^{-1/2}\prod_{j=2}^{d}[(2\pi)^{-1/2}\psi_j^{-1}$$

or

$$|\boldsymbol{R}| = \prod_{j=2}^{d}\psi_j^2.$$

From (8),

$$|\boldsymbol{R}| = \prod_{j=2}^{d}(1 - R^2_{Y_j\cdot(Y_1,\ldots,Y_{j-1})}). \tag{10}$$

That is, $|\boldsymbol{R}|$ is the product of terms of form $(1-R^2)$, where the R^2 come from the regressions or linear relations:

$$\begin{aligned}
Y_2 &= \beta_{21}Y_1 + \psi_2\epsilon_2 \\
Y_3 &= \beta_{31}Y_1 + \beta_{32}Y_2 + \psi_3\epsilon_3 \\
\vdots &\quad \vdots \\
Y_d &= \beta_{d1}Y_1 + \beta_{d2}Y_2 + \cdots + \beta_{d,d-1}Y_{d-1} + \psi_d\epsilon_d
\end{aligned}$$

Without proof, we state the further result that

$$1 - R^2_{Y_j\cdot(Y_1,\ldots,Y_{j-1}}(1-\rho^2_{ji_{j1}}) = \prod_{k=2}^{j-1}(1-\rho^2_{ji_{jk};i_{j1},\ldots,i_{j,k-1}}) \quad (i_{j1},\ldots,i_{j,j-1}) \text{ permutation of } (1,\ldots,j-1),$$

where $\rho_{ji_{jk};i_{j1},\ldots,i_{j,k-1}}$ is an abbreviated form for the partial correlation of Y_j and $Y_{i_{jk}}$ given $Y_{i_{j1}},\ldots,Y_{i_{j,k-1}}$. A special case is:

$$1 - R^2_{Y_j\cdot(Y_1,\ldots,Y_{j-1}} = (1-\rho^2_{j1})\prod_{k=2}^{j-1}(1-\rho^2_{jk;1:(k-1)}).$$

Subsitute the two above equations into (10) to get

$$\begin{aligned}
|\boldsymbol{R}| &= \prod_{j=2}^{d}\Big\{(1-\rho^2_{ji_{j1}}) = \prod_{k=2}^{j-1}(1-\rho^2_{ji_{jk};i_{j1},\ldots,i_{j,k-1}})\Big\} &(11)\\
&= \prod_{j=2}^{d}\Big\{(1-\rho^2_{j1})\prod_{k=2}^{j-1}(1-\rho^2_{jk;1:(k-1)})\Big\}. &(12)
\end{aligned}$$

Since the labelling of variables $1,\ldots,d$ is arbitrary, the determinant can be rewritten as a product over $(1-\rho_e^2)$ in many different forms after permuting the indices in (11). Note that there is a pattern to all cases of this formula: there are $d-1$ correlations, $d-2$ partial correlations given one variable, $d-k-1$ partial correlations given k variables for $2\le k\le d-3$, and one partial correlation given $d-2$ variables.

The link of this determinant to Bayesian network and vine diagrams will be made later.