• Bivariate skew-normal copula and conditional distributions

Derivations and notation are written carefully to match the notation in Azzalini's book, and his `sn` R package, in order to code the bivariate skew-normal copula (density, cdf, conditional cdfs, conditional quantiles).

The 3-parameter skew-normal copula has 3 correlation parameters $\gamma_1, \gamma_2, \rho$ that satisfy the positive definite matrix constraint when these are put into a $3 \times 3$ correlation matrix.

Let $\mathbf{R} = \begin{pmatrix} 1 & \gamma_1 & \gamma_2 \\ \gamma_1 & 1 & \rho \\ \gamma_2 & \rho & 1 \end{pmatrix}$. Let $(Z_0, Z_1, Z_2)^T$ be trivariate normal with zero mean vector and covariance matrix $\mathbf{R}$. The bivariate skew-normal distribution comes from $(Y_1, Y_2)^T := [(Z_1, Z_2)^T | Z_0 > 0]$, with marginal distributions from $Y_j \stackrel{d}{=} [Z_j | Z_0 > 0]$ for $j = 1, 2$. To match Azzalini's notation, let $\alpha_j = \gamma_j/(1 - \gamma_j^2)^{1/2}$.

To get the conditional distributions of the bivariate skew-normal, it is useful to consider more generally: $(Y_1, Y_2)^T := [(Z_1, Z_2)^T | Z_0 > -\tau]$.

Univariate marginal density (general notation followed by substitution for normal densities):

$$\begin{aligned}
\Pr(Y_j > y) &= [\overline{F}_{Z_0}(-\tau)]^{-1} \int_{-\tau}^{\infty} \overline{F}_{Z_j|Z_0}(y|z)\, f_{Z_0}(z)\, dz \\
f_{Y_j}(y) &= [\overline{F}_{Z_0}(-\tau)]^{-1} \int_{-\tau}^{\infty} f_{Z_j|Z_0}(y|z)\, f_{Z_0}(z)\, dz \\
&= [\overline{F}_{Z_0}(-\tau)]^{-1} \int_{-\tau}^{\infty} f_{Z_0|Z_j}(z|y)\, f_{Z_j}(y)\, dz \\
&= [\overline{F}_{Z_0}(-\tau)]^{-1} f_{Z_j}(y) \overline{F}_{Z_0|Z_j}(-\tau|y) \\
&= [\Phi(\tau)]^{-1} \phi(y)\, \overline{\Phi}\big((-\tau - \gamma_j y)/(1 - \gamma_j^2)^{1/2}\big) = [\Phi(\tau)]^{-1} \phi(y)\, \Phi\big((\tau + \gamma_j y)/(1 - \gamma_j^2)^{1/2}\big) \\
&= [\Phi(\tau)]^{-1} \phi(y)\, \Phi(\alpha_{0j} + \alpha_j y),
\end{aligned}$$

where $\alpha_{0j} = \tau/(1 - \gamma_j^2)^{1/2}$.

Azzalini's notation $\mathrm{SN}(\xi, \omega, \alpha, \tau)$ where $\gamma = \alpha/\sqrt{1 + \alpha^2}$, $\alpha = \gamma/\sqrt{1 - \gamma^2}$ and $1 - \gamma^2 = (1 + \alpha^2)^{-1}$. This is obtained by adding a location $\xi$ and scale parameter $\omega$: $V \sim \mathrm{SN}(\xi, \omega, \alpha, \tau)$ with $V \stackrel{d}{=} [\xi + \omega Z_1 | Z_0 > -\tau]$ where $(Z_0, Z_1)^T$ is bivariate normal with zero means, unit variances and correlation $\gamma \in (-1, 1)$. The density is

$$\begin{aligned}
f_{SN}(v) &= [\Phi(\tau)]^{-1} \phi\Big(\frac{v - \xi}{\omega}\Big)\, \Phi\Big(\frac{\tau + \gamma\big(\frac{v-\xi}{\omega}\big)}{(1 - \gamma^2)^{1/2}}\Big) \\
&= [\Phi(\tau)]^{-1} \omega^{-1} \phi\Big(\frac{v - \xi}{\omega}\Big)\, \Phi\Big(\alpha_0 + \alpha\Big(\frac{v - \xi}{\omega}\Big)\Big),
\end{aligned} \tag{1}$$

where $\alpha_0 = \tau/(1 - \gamma^2)^{1/2}$.

Bivariate marginal density:

$$\Pr(Y_1 > y_1, Y_2 > y_2) = [\overline{F}_{Z_0}(-\tau)]^{-1} \int_{-\tau}^{\infty} \overline{F}_{Z_1 Z_2|Z_0}(y_1, y_2|z)\, f_{Z_0}(z)\, dz$$

$$
\begin{aligned}
f_{Y_1,Y_2}(y_1,y_2) &= [\overline{F}_{Z_0}(-\tau)]^{-1}\int_{-\tau}^{\infty} f_{Z_1Z_2|Z_0}(y_1,y_2|z)\, f_{Z_0}(z)\,dz \\
&= [\overline{F}_{Z_0}(-\tau)]^{-1}\int_{0}^{\infty} f_{Z_0|Z_1Z_2}(z|y_1,y_2)\, f_{Z_1Z_2}(y_1,y_2)\,dz \\
&= [\overline{F}_{Z_0}(-\tau)]^{-1} f_{Z_1Z_2}(y_1,y_2)\overline{F}_{Z_0|Z_1Z_2}(-\tau|y_1,y_2) \\
&= [\Phi(\tau)]^{-1}\phi_2(y_1,y_2;\rho)\,\Phi\big((\tau+\beta_1y_1+\beta_2y_2)/\sigma\big),
\end{aligned}
$$

where $\beta_1 = (\gamma_1 - \rho\gamma_2)(1-\rho^2)^{-1}$, $\beta_2 = (\gamma_2 - \rho\gamma_1)(1-\rho^2)^{-1}$, and

$$
\sigma^2 = 1 - (\gamma_1^2+\gamma_2^2-2\rho\gamma_1\gamma_2)/(1-\rho^2) = \frac{1-\rho^2-\gamma_1^2-\gamma_2^2+2\rho\gamma_1\gamma_2}{(1-\rho^2)} = \frac{\det(\mathbf{R})}{(1-\rho^2)},
$$

and $\phi_2(\cdot;\rho)$ is the bivariate normal density with zero means, unit variances and correlation $\rho$. Special case with $\tau = 0$ leads to

$$
f_{Y_1,Y_2}(y_1,y_2) = 2\phi_2(y_1,y_2;\rho)\,\Phi\big((\beta_1y_1+\beta_2y_2)/\sigma\big).
$$

For the `sn` R package and Azzalini's book, the parametrization is $\mathbf{\Omega} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ and $\boldsymbol{\alpha} = \mathbf{\Omega}^{-1}\boldsymbol{\gamma}/(1-\boldsymbol{\gamma}^T\mathbf{\Omega}^{-1}\boldsymbol{\gamma})^{1/2}$ where $\boldsymbol{\gamma} = (\gamma_1,\gamma_2)^T$. The inverse mapping of $\boldsymbol{\alpha}$ to $\boldsymbol{\gamma}$ is $\boldsymbol{\gamma} = (1+\boldsymbol{\alpha}^T\mathbf{\Omega}\boldsymbol{\alpha})^{-1/2}\mathbf{\Omega}\boldsymbol{\alpha}$ making use of $1-\boldsymbol{\gamma}^T\mathbf{\Omega}^{-1}\boldsymbol{\gamma} = (1+\boldsymbol{\alpha}^T\mathbf{\Omega}\boldsymbol{\alpha})^{-1}$.

Conditional density $[Y_2|Y_1=y_1]$ for general $\tau$. This will be shown to have distribution $\mathrm{SN}(\xi_{2\cdot1},\omega_{2\cdot1},\alpha_{tdo},\tau_{tdo})$ where all parameters are expressed in terms of $\gamma_1,\gamma_2,\rho,\tau$. Then similarly, the conditional density $[Y_2|Y_1=y_1]$ has distribution $\mathrm{SN}(\xi_{1\cdot2},\omega_{1\cdot2},\alpha_{1\cdot2},\tau_{1\cdot2})$

$$
\begin{aligned}
f_{Y_2|Y_1}(y_2|y_1) &= f_{Y_1,Y_2}(y_1,y_2)/f_{Y_1}(y_1) \\
&= \frac{\phi_2(y_1,y_2;\rho)\,\Phi\big((\tau+\beta_1y_1+\beta_2y_2)/\sigma\big),}{\phi(y_1)\,\Phi(\alpha_{01}+\alpha_1y_1)} \\
&= \frac{\phi\big(\frac{y_2-\rho y_1}{(1-\rho^2)^{1/2}}\big)\,\Phi\big(\frac{(\tau+\beta_1y_1+\beta_2y_2)}{\sigma}\big),}{(1-\rho^2)^{1/2}\,\Phi(\alpha_{01}+\alpha_1y_1)} \\
&= [\Phi(\tau_{2\cdot1})]^{-1}\omega_{2\cdot1}^{-1}\phi\Big(\frac{y_2-\xi_{2\cdot1}}{\omega_{2\cdot1}}\Big)\,\Phi\Big(\tau_{2\cdot1}(1+\alpha_{2\cdot1}^2)^{1/2}+\alpha_{2\cdot1}\big(\frac{y_2-\xi_{2\cdot1}}{\omega_{2\cdot1}}\big)\Big),
\end{aligned}
$$

Matching parameters with (1) leads to:

$$
\begin{aligned}
\omega_{2\cdot1}^2 &= 1-\rho^2 \\
\tau_{2\cdot1} &= \alpha_{01}+\alpha_1y_1 = (\tau+\gamma_1y_1)/(1-\gamma_1^2)^{1/2}, \quad \alpha_{01} = \tau/(1-\gamma_1^2)^{1/2}, \quad \alpha_1 = \gamma_1/(1-\gamma_1^2)^{1/2} \\
\xi_{2\cdot1} &= \rho y_1 \\
\alpha_{2\cdot1} &= \beta_2\omega_{2\cdot1}/\sigma = \frac{(\gamma_2-\rho\gamma_1)}{\sigma(1-\rho^2)^{1/2}} \\
\gamma_{2\cdot1} &= \alpha_{2\cdot1}/(1+\alpha_{2\cdot1}^2)^{1/2} \\
1-\gamma_{2\cdot1}^2 &= (1+\alpha_{2\cdot1}^2)^{-1}.
\end{aligned}
$$

Then one must also have

$$\frac{\tau + \beta_1 y_1}{\sigma} = \tau_{2\cdot 1}/(1-\gamma_{2\cdot 1}^2)^{1/2} - \alpha_{2\cdot 1}\xi_{2\cdot 1}/\omega_{2\cdot 1} = \tau_{2\cdot 1}(1+\alpha_{2\cdot 1}^2)^{1/2} - \beta_2\rho y_1/\sigma,$$

or

$$\tau + \beta_1 y_1 = \tau_{2\cdot 1}(1+\alpha_{2\cdot 1}^2)^{1/2}\sigma - \beta_2\rho y_1. \tag{2}$$

Next notice that $(1+\alpha_{2\cdot 1}^2)\sigma^2$ simplifies:

$$\begin{aligned}
(1+\alpha_{2\cdot 1}^2)\sigma^2 &= \sigma^2 + (\gamma_2 - \rho\gamma_1)^2/(1-\rho^2) \\
&= \frac{(1-\rho^2-\gamma_1^2-\gamma_2^2+2\rho\gamma_1\gamma_2) + (\gamma_2^2 - 2\rho\gamma_1\gamma_2 + \rho^2\gamma_1^2)}{(1-\rho^2)} \\
&= \frac{(1-\rho^2-\gamma_1^2+\rho^2\gamma_1^2)}{(1-\rho^2)} = 1-\gamma_1^2
\end{aligned}$$

Hence

$$\begin{aligned}
\tau_{2\cdot 1}(1+\alpha_{2\cdot 1}^2)^{1/2}\sigma - \beta_2\rho y_1 &= \tau_{2\cdot 1}(1-\gamma_1^2)^{1/2} - \frac{(\gamma_2-\rho\gamma_1)\rho y_1}{1-\rho^2} \\
&= (\tau+\gamma_1 y_1) - \frac{(\gamma_2-\rho\gamma_1)\rho y_1}{1-\rho^2} \\
&= \frac{(1-\rho^2)(\tau+\gamma_1 y_1) - (\gamma_2-\rho\gamma_1)\rho y_1}{1-\rho^2} \\
&= \frac{(1-\rho^2)\tau + \gamma_1 y_1 - \rho^2\gamma_1 y_1 - (\gamma_2-\rho\gamma_1)\rho y_1}{1-\rho^2} \\
&= \frac{(1-\rho^2)\tau + (\gamma_1 - \rho\gamma_2)y_1}{1-\rho^2} \\
&= \tau + \beta_1 y_1
\end{aligned}$$

For the copula of bivariate skew-normal with $\tau = 0$, let the bivariate cdf be

$$F_{12}(y_1, y_2; \rho, \gamma_1, \gamma_2),$$

with marginal cdfs $F_1(y_1; \alpha_1)$ and $F_2(y_2; \alpha_2)$ $(\xi = 0, \omega = 1, \tau = 0)$ and density

$$f_{12}(y_1, y_2; \rho, \gamma_1, \gamma_2).$$

Then $\tau_{2\cdot 1} = \alpha_1 y = \gamma_1 y/(1-\gamma_1^2)^{1/2}$, and $\omega_{2\cdot 1}, \xi_{2\cdot 1}, \alpha_{2\cdot 1}, \gamma_{2\cdot 1}$ are as defined above. The conditional cdfs are:

$$F_{2|1}(y_2|y_1; \xi_{2\cdot 1}, \omega_{2\cdot 1}, \alpha_{2\cdot 1}, \tau_{2\cdot 1}), \quad F_{1|2}(y_1|y_2; \xi_{1\cdot 2}, \omega_{1\cdot 2}, \alpha_{1\cdot 2}, \tau_{1\cdot 2}),$$

and they are extended skew-normal. The copula cdf is

$$C(u_1, u_2) = F_{12}\big(F_1^{-1}(u_1; \alpha_1), F_2^{-1}(u_2; \alpha_2); \rho, \gamma_1, \gamma_2\big),$$

the copula density is

$$c(u_1,u_2)=\frac{f_{12}\big(F_1^{-1}(u_1;\alpha_1),F_2^{-1}(u_2;\alpha_2);\rho,\gamma_1,\gamma_2\big)}{f_1\big(F_1^{-1}(u_1;\alpha_1);\alpha_1\big)f_2\big(F_2^{-1}(u_2;\alpha_2);\alpha_2\big)},$$

and conditional cdfs are

$$C_{2|1}(u_2|u_1)=F_{2|1}\big(F_2^{-1}(u_2;\alpha_2)|F_1^{-1}(u_1;\alpha_1);\ \xi_{2\cdot 1},\omega_{2\cdot 1},\alpha_{2\cdot 1},\tau_{2\cdot 1}\big),$$

$$C_{1|2}(u_1|u_2)=F_{1|2}\big(F_1^{-1}(u_1;\alpha_1)|F_2^{-1}(u_2;\alpha_2);\ \xi_{1\cdot 2},\omega_{1\cdot 2},\alpha_{1\cdot 2},\tau_{1\cdot 2}\big),$$