The Gaussian Copula Model

Karl B. Gregory

Texas A & M University

April 30, 2011
The Gaussian Copula Model

Theory of Gaussian Copula Model
   Latent Variable Concept
   Meaning of Copula
   Rank Likelihood for Copula Estimation

Bayesian Inference
   Gibbs Sampling
   Illustration
   Summarizing Results / Dependence Graphs

Example
Suppose we have several ordinal variables which we measure on a group of experimental units. We wish to explore the correlations between them. We cannot assign numerical values to the levels of the non-numeric ordinal variables, e.g. education level. How then to estimate their correlations?

Recap of Latent Variable Concept
Recap of Latent Variable Concept

- Suppose we have several ordinal variables which we measure on a group of experimental units.
Recap of Latent Variable Concept

- Suppose we have several ordinal variables which we measure on a group of experimental units.
- We wish to explore the correlations between them.
Recap of Latent Variable Concept

- Suppose we have several ordinal variables which we measure on a group of experimental units.
- We wish to explore the correlations between them.
- We cannot assign numerical values to the levels of the non-numeric ordinal variables, e.g. education level.
Recap of Latent Variable Concept

- Suppose we have several ordinal variables which we measure on a group of experimental units.
- We wish to explore the correlations between them.
- We cannot assign numerical values to the levels of the non-numeric ordinal variables, e.g. education level.
- How then to estimate their correlations?
Recap of Latent Variable Concept

Assume that each ordinal response is determined by the value of some unobservable Gaussian random variable.
Assume that each ordinal response is determined by the value of some unobservable Gaussian random variable.

- When an individual’s educational ambition exceeds certain thresholds, he or she pursues the next higher degree.
Recap of Latent Variable Concept

Assume that each ordinal response is determined by the value of some unobservable Gaussian random variable.

▶ When an individual’s educational ambition exceeds certain thresholds, he or she pursues the next higher degree.
▶ The number of children someone has is a manifestation of his or her (Gaussian) propensity to domesticity.
The Gaussian Copula Model

Let \( Y_1, \ldots, Y_n \) be i.i.d. random samples from a \( p \)-variate population. The latent variable model is

\[
Z_1 \ldots Z_n \sim \text{i.i.d. multivariate normal}(0, \Psi) \\
Y_{i,j} = g_j(Z_{i,j})
\]
The Gaussian Copula Model

Let $Y_1, \ldots, Y_n$ be i.i.d. random samples from a $p$-variate population. The latent variable model is

$$Z_1 \ldots Z_n \sim \text{i.i.d. multivariate normal}(0, \Psi)$$

$$Y_{i,j} = g_j(Z_{i,j})$$

The functions $g_j$ are discretizing functions of the random variables $Z_{i,j}$. They consist merely of a set of threshold values, such that when $Z_{i,j}$ exceeds the next threshold, $Y_{i,j}$ assumes the next higher value.
The Gaussian Copula Model

Let $Y_1, \ldots, Y_n$ be i.i.d. random samples from a $p$-variate population. The latent variable model is

$$Z_1 \ldots Z_n \sim \text{i.i.d. multivariate normal}(0, \Psi)$$

$$Y_{i,j} = g_j(Z_{i,j})$$

The functions $g_j$ are discretizing functions of the random variables $Z_{i,j}$. They consist merely of a set of threshold values, such that when $Z_{i,j}$ exceeds the next threshold, $Y_{i,j}$ assumes the next higher value.

$\Psi$ describes the correlations between the $p$ variates in $Y$. 
"Copula"?
“Copula”?

Cupola?
The Gaussian Copula Model

- Theory of Gaussian Copula Model
- Meaning of Copula

“Copula”? Cupola? Coppola?
“Copula”? 

Consider the marginal distribution of $Y_{i,j}$ if it is continuous:

$$F_j(y) = P(Y_{i,j} \leq y)$$

$$= P(g_j(Z_{i,j}) \leq y)$$

$$= P(Z_{i,j} \leq g_j^{-1}(y))$$

$$= \Phi(g_j^{-1}(y))$$
“Copula”?

- Consider the marginal distribution of $Y_{i,j}$ if it is continuous:

\[
F_j(y) = P(Y_{i,j} \leq y) \\
= P(g_j(Z_{i,j}) \leq y) \\
= P(Z_{i,j} \leq g_j^{-1}(y)) \\
= \Phi(g_j^{-1}(y))
\]

Thus the marginal distribution of $Y_{i,j}$ depends only on the function $g_j$. 

The “copula” model separates the parameters for the dependencies among variables from the parameters for their marginal distributions.
“Copula”? 

▶ Consider the marginal distribution of $Y_{i,j}$ if it is continuous:

$$F_j(y) = P(Y_{i,j} \leq y) = P(g_j(Z_{i,j}) \leq y) = P(Z_{i,j} \leq g_j^{-1}(y)) = \Phi(g_j^{-1}(y))$$

▶ Thus the marginal distribution of $Y_{i,j}$ depends only on the function $g_j$.

▶ The “copula” model separates the parameters for the dependencies among variables from the parameters for their marginal distributions.
Of the unknown parameters \( \Psi \) and \( g_1, \ldots, g_p \), we are only interested in \( \Psi \).
Rank Likelihood for Copula Estimation

- Of the unknown parameters $\Psi$ and $g_1, \ldots, g_p$, we are only interested in $\Psi$.
- Treat $g_1, \ldots, g_p$ as nuisance parameters.
Rank Likelihood for Copula Estimation

- Of the unknown parameters $\Psi$ and $g_1, \ldots, g_p$, we are only interested in $\Psi$.
- Treat $g_1, \ldots, g_p$ as nuisance parameters.
- We can obtain a likelihood function for $\Psi$ that does not depend on $g_1, \ldots, g_p$. 
Rank Likelihood for Copula Estimation

- Of the unknown parameters $\Psi$ and $g_1, \ldots, g_p$, we are only interested in $\Psi$.
- Treat $g_1, \ldots, g_p$ as nuisance parameters.
- We can obtain a likelihood function for $\Psi$ that does not depend on $g_1, \ldots, g_p$.
- Given our observed $Y$, we know that the matrix $Z$ must lie in the set

\[
R(Y) = \{ Z : z_{i,j} < z'_{i',j} \text{ if } y_{i,j} < y'_{i',j} \}
\]
Rank Likelihood for Copula Estimation

- Of the unknown parameters $\Psi$ and $g_1, \ldots, g_p$, we are only interested in $\Psi$.
- Treat $g_1, \ldots, g_p$ as nuisance parameters.
- We can obtain a likelihood function for $\Psi$ that does not depend on $g_1, \ldots, g_p$.
- Given our observed $Y$, we know that the matrix $Z$ must lie in the set

$$R(Y) = \{Z : z_{i,j} < z_{i',j} \text{ if } y_{i,j} < y_{i',j}\}$$

- For example, all the $z$ values for which $y = 2$ must be less than all the $z$ values for which $y = 3$. 
Visual of $\mathbf{Z} \in R(\mathbf{Y})$
$P(Z \in R(Y)|\Psi)$ is called the rank likelihood for the multivariate Gaussian copula model.
The Gaussian Copula Model

Rank Likelihood for Copula Estimation

- $P(Z \in R(Y) | \Psi)$ is called the rank likelihood for the multivariate Gaussian copula model.
- It is very difficult to compute this likelihood.
Rank Likelihood for Copula Estimation

- $P(Z \in R(Y) | \Psi)$ is called the rank likelihood for the multivariate Gaussian copula model.
- It is very difficult to compute this likelihood.
- We instead take a Bayesian approach, drawing inferences from the joint posterior $P(\Psi, Z | Z \in R(Y))$. 
Rank Likelihood for Copula Estimation

- $P(Z \in R(Y) | \Psi)$ is called the rank likelihood for the multivariate Gaussian copula model.
- It is very difficult to compute this likelihood.
- We instead take a Bayesian approach, drawing inferences from the joint posterior $P(\Psi, Z | Z \in R(Y))$.
- By choosing the right priors, we can make an MCMC approximation to $P(\Psi, Z | Z \in R(Y))$ via Gibbs sampling.
Bayesian Inference on $\Psi$

- Let $\Psi = h(\Sigma)$ where $\Sigma \sim \text{Inverse-Wishart}(\nu_0, S_0^{-1})$. 

The Gaussian Copula Model

Bayesian Inference
Bayesian Inference on $\Psi$

- Let $\Psi = h(\Sigma)$ where $\Sigma \sim \text{Inverse-Wishart}(\nu_0, S^{-1}_0)$.

- $\Psi_{i,j} = h(\Sigma) = \sigma_{i,j} / \sqrt{\sigma_i^2 \sigma_j^2}$. 
Bayesian Inference on $\Psi$

- Let $\Psi = h(\Sigma)$ where $\Sigma \sim \text{Inverse-Wishart}(\nu_0, S_0^{-1})$.
- $\Psi_{i,j} = h(\Sigma) = \sigma_{i,j} / \sqrt{\sigma_i^2 \sigma_j^2}$.
- Let $Z_1, \ldots, Z_n | \Psi \sim \text{i.i.d. multivariate normal}(0, \Psi)$. 
Bayesian Inference on $\Psi$

- Let $\Psi = h(\Sigma)$ where $\Sigma \sim \text{Inverse-Wishart}(\nu_0, S_0^{-1})$.
- $\Psi_{i,j} = h(\Sigma) = \sigma_{i,j} / \sqrt{\sigma_i^2 \sigma_j^2}$.
- Let $Z_1, \ldots, Z_n | \Psi \sim \text{i.i.d. multivariate normal}(0, \Psi)$.
- Note that our observed data $Y$ do not inform us as to the magnitude of the $Z$ elements, but only their ordering.
Bayesian Inference on $\Psi$

- Let $\Psi = h(\Sigma)$ where $\Sigma \sim \text{Inverse-Wishart}(\nu_0, S_0^{-1})$.
- $\Psi_{i,j} = h(\Sigma) = \sigma_{i,j} / \sqrt{\sigma_i^2 \sigma_j^2}$.
- Let $Z_1, \ldots, Z_n | \Psi \sim \text{i.i.d. multivariate normal}(0, \Psi)$.
- Note that our observed data $Y$ do not inform us as to the magnitude of the $Z$ elements, but only their ordering.
- Hence there is insufficient information to estimate $\Sigma$ (it is non-identifiable), but $\Psi$ is identifiable.
Bayesian Inference on $\Psi$

- Let $\Psi = h(\Sigma)$ where $\Sigma \sim \text{Inverse-Wishart}(\nu_0, S_0^{-1})$.
- $\Psi_{i,j} = h(\Sigma) = \sigma_{i,j} / \sqrt{\sigma_i^2 \sigma_j^2}$.
- Let $Z_1, \ldots, Z_n | \Psi \sim \text{i.i.d. multivariate normal}(0, \Psi)$.
- Note that our observed data $Y$ do not inform us as to the magnitude of the $Z$ elements, but only their ordering.
- Hence there is insufficient information to estimate $\Sigma$ (it is non-identifiable), but $\Psi$ is identifiable.
- It does not matter whether $Z_1, \ldots, Z_n$ are from a $N_p(0, \Sigma)$ or $N_p(0, \Psi = h(\Sigma))$. 
Setting up the Gibbs Sampler

▶ We wish to sample from the posterior distribution $P(\Psi, Z | Z \in R(Y))$. 
Setting up the Gibbs Sampler

▶ We wish to sample from the posterior distribution $P(\Psi, Z | Z \in R(Y))$.
▶ So we sample from the full conditionals:

$$
\Sigma^{(s+1)} \sim p(\Sigma | Z^{(s)}) \\
\propto p(\Sigma) \times p(Z^{(s)} | \Sigma) \\
= \text{inverse-Wishart}(\nu_0 + n, [S_0 + Z^{(s)'}Z^{(s)}]^{-1})
$$
Full Conditional of $Z$:

The full conditional distribution of $Z$ is more complex: We have to generate $n \ p \times 1$ vectors from a normal distribution such that the resulting $Z$ matrix conforms to $Z \in \mathbb{R}(Y)$.

$Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)})$
Full Conditional of Z:

The full conditional distribution of \( Z \) is more complex: We have to generate \( n \ p \times 1 \) vectors from a normal distribution such that the resulting \( Z \) matrix conforms to \( Z \in R(Y) \).

\[
Z_{i,1}^{(s+1)} \sim p(Z_{i,1}|\Sigma^{(s+1)}, Z_{i,-1}^{(s)})
\]

But we must restrict \( Z_{i,1} \) such that

\[
\max\{z_{k,1} : y_{k,1} < y_{i,1}\} < Z_{i,1} < \min\{z_{k,1} : y_{k,1} > y_{i,1}\}
\]
Full Conditional of $Z$:

The full conditional distribution of $Z$ is more complex: We have to generate $n \ p \times 1$ vectors from a normal distribution such that the resulting $Z$ matrix conforms to $Z \in \mathbb{R}(Y)$.

- $Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)})$
- But we must restrict $Z_{i,1}$ such that
  \[
  \max\{z_{k,1} : y_{k,1} < y_{i,1}\} < Z_{i,1} < \min\{z_{k,1} : y_{k,1} > y_{i,1}\}
  \]
- $Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)})I_{(a_{i1},b_{i1})}(Z_{i,1})$
Full Conditional of $Z$:

The full conditional distribution of $Z$ is more complex: We have to generate $n \ p \times 1$ vectors from a normal distribution such that the resulting $Z$ matrix conforms to $Z \in R(Y)$.

- $Z_{i,1}^{(s+1)} \sim p(Z_{i,1} \mid \Sigma^{(s+1)}, Z_{i,-1}^{(s)})$

- But we must restrict $Z_{i,1}$ such that

$$\max\{z_{k,1} : y_{k,1} < y_{i,1}\} < Z_{i,1} < \min\{z_{k,1} : y_{k,1} > y_{i,1}\}$$

- $Z_{i,1}^{(s+1)} \sim p(Z_{i,1} \mid \Sigma^{(s+1)}, Z_{i,-1}^{(s)})I(a_{i1}, b_{i1})(Z_{i,1})$

$$p(Z_{i,1} \mid \Sigma^{(s+1)}, Z_{i,-1}^{(s)}) =$$

$$N(\Sigma_{1,-1}^{(s+1)}(\Sigma_{-1,-1}^{(s+1)})^{-1}Z_{i,-1}^{(s)}, \Sigma_{1,1}^{(s+1)} - \Sigma_{1,-1}^{(s+1)}(\Sigma_{-1,-1}^{(s+1)})^{-1}\Sigma_{-1,1}^{(s+1)}) \times I(a_{i1}, b_{i1})(Z_{i,1})$$
Full Conditional of $Z$:

which leads to
Full Conditional of $Z$:

which leads to

$$Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)})I(a_{i1},b_{i1})(Z_{i,1})$$
Full Conditional of $Z$:

which leads to

$$Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,1}^{(s)})I(a_{i1}, b_{i1})(Z_{i,1})$$

$$Z_{i,2}^{(s+1)} \sim p(Z_{i,2} | \Sigma^{(s+1)}, Z_{i,1}^{(s+1)}, Z_{i,1}^{(s)})I(a_{i2}, b_{i2})(Z_{i,2})$$
Full Conditional of $Z$:

which leads to

\[
\begin{align*}
Z_{i,1}^{(s+1)} & \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)}) I(a_{i1}, b_{i1})(Z_{i,1}) \\
Z_{i,2}^{(s+1)} & \sim p(Z_{i,2} | \Sigma^{(s+1)}, Z_{i,1}^{(s+1)}, Z_{i, -(j>2)}^{(s)}) I(a_{i2}, b_{i2})(Z_{i,2}) \\
Z_{i,3}^{(s+1)} & \sim p(Z_{i,3} | \Sigma^{(s+1)}, Z_{i, -(j<3)}^{(s+1)}, Z_{i, -(j>3)}^{(s)}) I(a_{i3}, b_{i3})(Z_{i,3})
\end{align*}
\]
Full Conditional of $Z$:

which leads to

$$Z_{i,1}^{(s+1)} \sim p(Z_{i,1}|\Sigma^{(s+1)}, Z_{i,-1}^{(s)})I_{(a_{i1},b_{i1})}(Z_{i,1})$$

$$Z_{i,2}^{(s+1)} \sim p(Z_{i,2}|\Sigma^{(s+1)}, Z_{i,1}^{(s+1)}, Z_{i,-(j>2)}^{(s)})I_{(a_{i2},b_{i2})}(Z_{i,2})$$

$$Z_{i,3}^{(s+1)} \sim p(Z_{i,3}|\Sigma^{(s+1)}, Z_{i,-(j<3)}, Z_{i,-(j>3)}^{(s)})I_{(a_{i3},b_{i3})}(Z_{i,3})$$

$$Z_{i,p-2}^{(s+1)} \sim p(Z_{i,p-2}|\Sigma^{(s+1)}, Z_{i,-(j<p-2)}, Z_{i,-(j>p-2)}^{(s)})I_{(a_{i(p-2)},b_{i(p-2)})}(Z_{i,p-2})$$
Full Conditional of $Z$:

which leads to

$$Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)}) I(a_{i1},b_{i1})(Z_{i,1})$$

$$Z_{i,2}^{(s+1)} \sim p(Z_{i,2} | \Sigma^{(s+1)}, Z_{i,1}^{(s)}, Z_{i,-(j>2)}^{(s)}) I(a_{i2},b_{i2})(Z_{i,2})$$

$$Z_{i,3}^{(s+1)} \sim p(Z_{i,3} | \Sigma^{(s+1)}, Z_{i,-(j<3)}, Z_{i,-(j>3)}^{(s)}) I(a_{i3},b_{i3})(Z_{i,3})$$

$$Z_{i,p-2}^{(s+1)} \sim p(Z_{i,p-2} | \Sigma^{(s+1)}, Z_{i,-(j<p-2)}, Z_{i,-(j>p-2)}^{(s)}) I(a_{i(p-2)},b_{i(p-2)})(Z_{i,p-2})$$

$$Z_{i,p-1}^{(s+1)} \sim p(Z_{i,p-1} | \Sigma^{(s+1)}, Z_{i,-(j<p-1)}, Z_{i,-p}^{(s)}) I(a_{i(p-1)},b_{i(p-1)})(Z_{i,p-1})$$
Full Conditional of $Z$:

which leads to

\[ Z_{i,1}^{(s+1)} \sim p(Z_{i,1} | \Sigma^{(s+1)}, Z_{i,-1}^{(s)})I(a_{i1},b_{i1})(Z_{i,1}) \]

\[ Z_{i,2}^{(s+1)} \sim p(Z_{i,2} | \Sigma^{(s+1)}, Z_{i,1}^{(s+1)}, Z_{i,-(j>2)}^{(s)})I(a_{i2},b_{i2})(Z_{i,2}) \]

\[ Z_{i,3}^{(s+1)} \sim p(Z_{i,3} | \Sigma^{(s+1)}, Z_{i,-(j<3)}, Z_{i,-(j>3)}^{(s)})I(a_{i3},b_{i3})(Z_{i,3}) \]

\[ Z_{i,p-2}^{(s+1)} \sim p(Z_{i,p-2} | \Sigma^{(s+1)}, Z_{i,-(j<p-2)}, Z_{i,-(j>p-2)}^{(s)})I(a_{i(p-2)},b_{i(p-2)})(Z_{i,p-2}) \]

\[ Z_{i,p-1}^{(s+1)} \sim p(Z_{i,p-1} | \Sigma^{(s+1)}, Z_{i,-(j<p-1)}, Z_{i,-p}^{(s)})I(a_{i(p-1)},b_{i(p-1)})(Z_{i,p-1}) \]

\[ Z_{i,p}^{(s+1)} \sim p(Z_{i,p} | \Sigma^{(s+1)}, Z_{i,-p}^{(s+1)})I(a_{ip},b_{ip})(Z_{i,p}) \]
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$. 
An Illustration on “Cooked-Up” Data:

- Let $\mathbf{Y}_1, \ldots, \mathbf{Y}_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $\mathbf{Z}_1, \ldots, \mathbf{Z}_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
  1. Generate $Z_1, \ldots, Z_{30} \sim N(0, \Psi)$. 


An Illustration on “Cooked-Up” Data:

- Let $\mathbf{Y}_1, \ldots, \mathbf{Y}_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $\mathbf{Z}_1, \ldots, \mathbf{Z}_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
  1. Generate $\mathbf{Z}_1, \ldots, \mathbf{Z}_{30} \sim N(0, \Psi)$.
  2. Define $g_1, g_2, g_3$. 
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
  1. Generate $Z_1, \ldots, Z_{30} \sim N(0, \Psi)$.
  2. Define $g_1, g_2, g_3$.
  3. Pass $Z_1, \ldots, Z_{30}$ through $g_1, g_2, g_3$ to obtain $Y_1, \ldots, Y_{30}$. 
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
  1. Generate $Z_1, \ldots, Z_{30} \sim N(0, \Psi)$.
  2. Define $g_1, g_2, g_3$.
  3. Pass $Z_1, \ldots, Z_{30}$ through $g_1, g_2, g_3$ to obtain $Y_1, \ldots, Y_{30}$.
  4. Choose initial values $Z_1^0, \ldots, Z_{30}^0 \in R(\mathbf{Y}_{obs})$ and run the Gibbs sampler.

Letting $Z_1, \ldots, Z_{30}$ be $3 \times 1$ vectors of ordinal responses, they can be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$. To illustrate the Gaussian Copula method:

1. Generate $Z_1, \ldots, Z_{30} \sim N(0, \Psi)$.
2. Define $g_1, g_2, g_3$.
3. Pass $Z_1, \ldots, Z_{30}$ through $g_1, g_2, g_3$ to obtain $Y_1, \ldots, Y_{30}$.
4. Choose initial values $Z_1^0, \ldots, Z_{30}^0 \in R(\mathbf{Y}_{obs})$ and run the Gibbs sampler.
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
  1. Generate $Z_1, \ldots, Z_{30} \sim N(0, \Psi)$.
  2. Define $g_1, g_2, g_3$.
  3. Pass $Z_1, \ldots, Z_{30}$ through $g_1, g_2, g_3$ to obtain $Y_1, \ldots, Y_{30}$.
  4. Choose initial values $Z_1^0, \ldots, Z_{30}^0 \in R(Y_{obs})$ and run the Gibbs sampler.
  5. Take the mean of the correlation matrices from all the MCMC iterations.
An Illustration on “Cooked-Up” Data:

- Let $Y_1, \ldots, Y_{30}$ be $3 \times 1$ vectors of ordinal responses.
- Let them be discrete expressions of the $3 \times 1$ multivariate normal random vectors $Z_1, \ldots, Z_{30}$ via the functions $g_1, g_2, g_3$.
- To illustrate the Gaussian Copula method:
  1. Generate $Z_1, \ldots, Z_{30} \sim N(0, \Psi)$.
  2. Define $g_1, g_2, g_3$.
  3. Pass $Z_1, \ldots, Z_{30}$ through $g_1, g_2, g_3$ to obtain $Y_1, \ldots, Y_{30}$.
  4. Choose initial values $Z^0_1, \ldots, Z^0_{30} \in R(Y_{obs})$ and run the Gibbs sampler.
  5. Take the mean of the correlation matrices from all the MCMC iterations.
  6. Compare this to the sample correlation of the original set of latent variable vectors $Z_1, \ldots, Z_{30}$.
Gibbs Sampling Code

```r
S<-10000
All_Psi<-matrix(1,p,p*S)
Psi<-matrix(1,p,p)

for (s in 1:S){
  Sigma<-riwish(v0+n,S0+t(Z)%*%Z)
  for(i in 1:p){for(j in 1:p){Psi[i,j]<-Sigma[i,j]/sqrt(Sigma[i,i]*Sigma[j,j])}}

  for (i in 1:n){for(j in 1:p){
    Sz<-Sigma[j,-j]%*%solve(Sigma[-j,-j])
    sz<-sqrt(Sigma[j,j]-Sz%*%Sigma[-j,j])
    ez<-Sz%*%Z[i,-j]

    a<-max(Z[,j]<Y[i,j],na.rm=T)
    b<-min(Z[,j]>Y[i,j],na.rm=T)

    u<-runif(1, pnorm((a-ez)/sz),pnorm((b-ez)/sz))
    Z[i,j]<-ez+sz*qnorm(u)
  }}
  All_Psi[,,(p*(s-1)+1):(p*s)]<-Psi }

Mean_Psi<-All_Psi%*%(matrix(1,S,1)%x%diag(p))/S
```
Gibbs-sampled Markov chain of length 20,000 of three correlation coefficients. (n=30)
The Gaussian Copula Model

Bayesian Inference

Illustration
The Gaussian Copula Model

Bayesian Inference

Illustration
Here is the correlation matrix for the original $Z_1, \ldots, Z_{30}$ vectors and the estimate obtained for it using the Gaussian Copula method:

\[
cor(Z_{\text{orig}}) \nonumber
\]
\[
[,1] [,2] [,3] 
[1,] 1.0000000 0.8227272 -0.5764696 
[2,] 0.8227272 1.0000000 -0.8327858 
[3,] -0.5764696 -0.8327858 1.0000000 
\]

Mean_Psi: Gaussian Copula estimate of $\Psi$.

\[
[,1] [,2] [,3] 
[1,] 1.0000000 0.8062399 -0.5301972 
[2,] 0.8062399 1.0000000 -0.7937843 
[3,] -0.5301972 -0.7937843 1.0000000 
\]

cor(Z_{\text{orig}})-Mean_Psi: Discrepancy between target and estimate.

\[
[,1] [,2] [,3] 
[1,] 0.00000000 0.01648722 -0.04627241 
[2,] 0.01648722 0.00000000 -0.03900150 
[3,] -0.04627241 -0.03900150 0.00000000 
\]
Summarizing Results

▶ How do we interpret the estimated correlation matrix?

Recall that these are the estimated correlations between the latent Gaussian variables $Z_1, Z_2, Z_3$ of which we have assumed our observed ordinal values $Y_1, Y_2, Y_3$ to be discrete manifestations.

Consider regressing one latent variable on the others and examining the slope coefficients.

The estimated slope coefficients for the regression of $Z_j$ on $Z_{-j}$ comprise the vector $\Psi_j,_{-j} (\Psi_{-j},_{-j})^{-1}$.

We can compute this set of coefficients for every single MCMC iteration and create percentile intervals for each one (a la bootstrap).
Summarizing Results

- How do we interpret the estimated correlation matrix?
- Recall that these are the estimated correlations between the latent Gaussian variables $Z_1, Z_2, Z_3$ of which we have assumed our observed ordinal values $Y_1, Y_2, Y_3$ to be discrete manifestations.
Summarizing Results

- How do we interpret the estimated correlation matrix?
- Recall that these are the estimated correlations between the latent Gaussian variables $Z_1, Z_2, Z_3$ of which we have assumed our observed ordinal values $Y_1, Y_2, Y_3$ to be discrete manifestations.
- Consider regressing one latent variable on the others and examining the slope coefficients.
Summarizing Results

- How do we interpret the estimated correlation matrix?
- Recall that these are the estimated correlations between the latent Gaussian variables $Z_1, Z_2, Z_3$ of which we have assumed our observed ordinal values $Y_1, Y_2, Y_3$ to be discrete manifestations.
- Consider regressing one latent variable on the others and examining the slope coefficients.
- The estimated slope coefficients for the regression of $Z_j$ on $Z_{-j}$ comprise the vector $\Psi_{j,-j}(\Psi_{-j,-j})^{-1}$. 
Summarizing Results

- How do we interpret the estimated correlation matrix?
- Recall that these are the estimated correlations between the latent Gaussian variables $Z_1, Z_2, Z_3$ of which we have assumed our observed ordinal values $Y_1, Y_2, Y_3$ to be discrete manifestations.
- Consider regressing one latent variable on the others and examining the slope coefficients.
- The estimated slope coefficients for the regression of $Z_j$ on $Z_{-j}$ comprise the vector $\Psi_{j,-j}(\Psi_{-j,-j})^{-1}$.
- We can compute this set of coefficients for every single MCMC iteration and create percentile intervals for each one (à la bootstrap).
The Gaussian Copula Model
Bayesian Inference
Summarizing Results / Dependence Graphs

Creating Percentile Intervals for Slope Coefficients

\[
\begin{align*}
\text{Beta} & \leftarrow \text{matrix}(1, p*(p-1), S) \\
\text{for (s in 1:S) } & \\
\text{for(j in 1:p) } & \\
\text{Beta}[((j-1)*(p-1)+1):(j*(p-1)),s] & \leftarrow (\text{All_Psi}[,\text{(p*(s-1)+1):(p*s)})[j,-j] \\
\%\% \text{solve}((\text{All_Psi}[,\text{(p*(s-1)+1):(p*s})][,-j,-j]))
\end{align*}
\]

\[
\text{CIs} & \leftarrow \text{matrix}(1, p*(p-1), 2) \\
\text{rownames} & \leftarrow \text{character}(\text{length}=p*(p-1)) \\
\text{for (i in 1:(p*(p-1))) } & \\
\text{CIs[i,]} & \leftarrow (\text{as.vector(quantile(\text{Beta}[i,],c(.025,.975))})) \\
\text{for (j in 1:p) } & \\
\text{rownames}[(((p-1)*j)-(p-2)):((p-1)*j)] & \leftarrow (\text{paste("","|",c(1:p)[-j],""})) \\
\text{rownames(CIs)} & \leftarrow \text{rownames} \\
\text{colnames(CIs)} & \leftarrow (\text{c("LB","UB")})
\]

\[
\text{CIs}
\]

\begin{array}{c|cc}
\text{LB} & \text{UB} \\
\{ 1 | 2 \} & 0.6072825 & 1.6459383 \\
\{ 1 | 3 \} & -0.1774682 & 0.9726943 \\
\{ 2 | 1 \} & 0.2699097 & 0.7785032 \\
\{ 2 | 3 \} & -0.7615016 & -0.2388328 \\
\{ 3 | 1 \} & -0.1947200 & 0.9548002 \\
\{ 3 | 2 \} & -1.6038528 & -0.5530947 \\
\end{array}
Dependence Graph

The set of percentile intervals for the slope coefficients gives rise to a dependence graph.

<table>
<thead>
<tr>
<th>CIs</th>
<th>LB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>{ 1</td>
<td>2 }</td>
<td>0.6072825</td>
</tr>
<tr>
<td>{ 1</td>
<td>3 }</td>
<td>-0.1774682</td>
</tr>
<tr>
<td>{ 2</td>
<td>1 }</td>
<td>0.2699097</td>
</tr>
<tr>
<td>{ 2</td>
<td>3 }</td>
<td>-0.7615016</td>
</tr>
<tr>
<td>{ 3</td>
<td>1 }</td>
<td>-0.1947200</td>
</tr>
<tr>
<td>{ 3</td>
<td>2 }</td>
<td>-1.6038528</td>
</tr>
</tbody>
</table>

Diagram: 1 → 2, 2 → 3, 3 → 1
Example: Blood Pressure Case Study

We now use the Gaussian Copula model to explore correlations between the following ordinal-scale variables measured on a sample of 50 patients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1$</td>
<td>Exercise 1 = Low, 2 = Medium, 3 = High</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>Overweight 1 = Normal, 2 = Overweight, 3 = Obese</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>Alcohol 1 = Low, 2 = Medium, 3 = High</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>Stress 1 = Low, 2 = Medium, 3 = High</td>
</tr>
<tr>
<td>$Y_5$</td>
<td>Income 1 = Low, 2 = Medium, 3 = High</td>
</tr>
<tr>
<td>$Y_6$</td>
<td>Education 1 = Low, 2 = Medium, 3 = High</td>
</tr>
</tbody>
</table>

Data are taken from [http://www.math.yorku.ca/Who/Faculty/Ng/ssc2003/BPMain.htm](http://www.math.yorku.ca/Who/Faculty/Ng/ssc2003/BPMain.htm)
The Gaussian Copula Model

Example

Mean_Psi:

```
[1,]  1.000000000 -0.4820717 -0.06046314 -0.35346704 -0.004869353
[2,] -0.482071691  1.0000000 -0.18685984  0.51045570  0.25730921
[3,] -0.060463136 -0.1868598  1.00000000 -0.24879991 -0.26537232
[4,] -0.353467040  0.5104557 -0.24879991  1.00000000 -0.00486935
[5,]  0.257309210 -0.2494349 -0.16639406 -0.28372326  1.00000000
[6,] -0.004869353 -0.2653723 -0.11699204 -0.00565760  1.00000000
```
The Gaussian Copula Model

Example
Dependence Graph

We create a dependence graph for the variables in our dataset based on the estimated slope coefficients of the latent variables regressed upon each other.

CIs[which(CIs[,1]*CIs[,2]>0),]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>LB</th>
<th>UB</th>
</tr>
</thead>
<tbody>
<tr>
<td>{ 1</td>
<td>2 }</td>
<td>-0.850021395</td>
<td>-0.094304300</td>
</tr>
<tr>
<td>{ 2</td>
<td>1 }</td>
<td>-0.650882131</td>
<td>-0.069841757</td>
</tr>
<tr>
<td>{ 2</td>
<td>4 }</td>
<td>0.005934974</td>
<td>0.622353926</td>
</tr>
<tr>
<td>{ 2</td>
<td>6 }</td>
<td>-0.549366295</td>
<td>-0.006750942</td>
</tr>
<tr>
<td>{ 4</td>
<td>2 }</td>
<td>0.006359494</td>
<td>0.765293999</td>
</tr>
<tr>
<td>{ 6</td>
<td>2 }</td>
<td>-0.894628238</td>
<td>-0.011763248</td>
</tr>
</tbody>
</table>
Resources:

Find this presentation as well as my R code for the MCMC illustration at my website:

www.stat.tamu.edu/~kbgregory