Comparing Two Systems Using Gaussian Copulae[1]

Samuel M. T. Ehrlichman and Shane G. Henderson

School of Operations Research and Information Engineering
Cornell University

July 13, 2008

[1]Thanks to NSF Grants DMI 0400287 and CMMI 0800688

Common Random Numbers and Copulas

- Is $EX > EY$, i.e., $E(X - Y) > 0$?

Common Random Numbers and Copulas

- Is $EX > EY$, i.e., $E(X - Y) > 0$?
- Let $X = f_U(\mathbf{U}_X), Y = g_U(\mathbf{U}_Y)$ where $\mathbf{U}_X, \mathbf{U}_Y \sim \mathcal{U}([0,1]^d)$
- Standard sampling: $\mathbf{U}_X$ is independent of $\mathbf{U}_Y$
- CRN sampling: $\mathbf{U}_X = \mathbf{U}_Y$
- $\mathrm{var}(X - Y) = \mathrm{var}X + \mathrm{var}\,Y - 2\mathrm{cov}(X, Y)$

Common Random Numbers and Copulas

- Is $EX > EY$, i.e., $E(X - Y) > 0$?
- Let $X = f_U(\mathbf{U}_X)$, $Y = g_U(\mathbf{U}_Y)$ where $\mathbf{U}_X, \mathbf{U}_Y \sim \mathcal{U}([0,1]^d)$
- Standard sampling: $\mathbf{U}_X$ is independent of $\mathbf{U}_Y$
- CRN sampling: $\mathbf{U}_X = \mathbf{U}_Y$
- $\mathrm{var}(X - Y) = \mathrm{var}X + \mathrm{var}\,Y - 2\mathrm{cov}(X, Y)$

What is essential?

Let $\mathbf{U} = (\mathbf{U}_X, \mathbf{U}_Y)$

$$\mathbf{U} = (\underbrace{U_1, U_2, \ldots, U_d}_{IID}, \underbrace{U_{d+1}, U_{d+2}, \ldots, U_{2d}}_{IID})$$

Coupling

- We want a copula that minimizes $\mathrm{var}(X - Y)$, i.e., maximizes $\mathrm{cov}(X, Y)$ subject to the marginal constraints

Coupling

- We want a copula that minimizes $\mathrm{var}(X - Y)$, i.e., maximizes $\mathrm{cov}(X, Y)$ subject to the marginal constraints
- Not a new idea...
  1. Wright and Ramsay (1979)
  2. Schmeiser and Kachitvichyanukul (1986)
  3. Devroye (1990)
  4. Glasserman and Yao (1992)
  5. Glasserman and Yao (2004)

Coupling

- We want a copula that minimizes $\mathrm{var}(X - Y)$, i.e., maximizes $\mathrm{cov}(X, Y)$ subject to the marginal constraints
- Not a new idea...
  1. Wright and Ramsay (1979)
  2. Schmeiser and Kachitvichyanukul (1986)
  3. Devroye (1990)
  4. Glasserman and Yao (1992)
  5. Glasserman and Yao (2004)

What is new ...

... is that we use a new class of copulas and have a computational method for searching over it

Gaussian Copulae

Let $\mathbf{Z} = (\mathbf{Z}_X, \mathbf{Z}_Y)$ be jointly Gaussian, standard normal marginals, covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} I_d & \mathbf{\Sigma}_{XY} \\ \mathbf{\Sigma}_{XY}^T & I_d \end{bmatrix}$$

Let $\mathbf{Z} = (\mathbf{Z}_X, \mathbf{Z}_Y)$ be jointly Gaussian, standard normal marginals, covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} I_d & \mathbf{\Sigma}_{XY} \\ \mathbf{\Sigma}_{XY}^T & I_d \end{bmatrix}$$

Set

$$\mathbf{U}_X[i] = \Phi(\mathbf{Z}_X[i]), \mathbf{U}_Y[i] = \Phi(\mathbf{Z}_Y[i])$$

for $i = 1, 2, \ldots, d$ and

Gaussian Copulae

Let $\mathbf{Z} = (\mathbf{Z}_X, \mathbf{Z}_Y)$ be jointly Gaussian, standard normal marginals, covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} I_d & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{XY}^T & I_d \end{bmatrix}$$

Set

$$\mathbf{U}_X[i] = \Phi(\mathbf{Z}_X[i]), \mathbf{U}_Y[i] = \Phi(\mathbf{Z}_Y[i])$$

for $i = 1, 2, \ldots, d$ and

$$X = f_U(\mathbf{U}_X) = f(\mathbf{Z}_X)$$
$$Y = g_U(\mathbf{U}_Y) = g(\mathbf{Z}_Y)$$

Example 1

Take $d = 2$

$$X = f(\mathbf{Z}_X) = \frac{Z_X[1] + Z_X[2]}{\sqrt{2}}$$
$$Y = g(\mathbf{Z}_Y) = Z_Y[1]$$

Example 1

Take $d = 2$

$$X = f(\mathbf{Z}_X) = \frac{Z_X[1] + Z_X[2]}{\sqrt{2}}$$

$$Y = g(\mathbf{Z}_Y) = Z_Y[1]$$

Independence copula

$\mathrm{var}(X - Y) = 2$

Example 1

Take $d = 2$

$$X = f(\mathbf{Z}_X) = \frac{Z_X[1] + Z_X[2]}{\sqrt{2}}$$
$$Y = g(\mathbf{Z}_Y) = Z_Y[1]$$

Independence copula

$\mathrm{var}(X - Y) = 2$

CRN copula

$$\mathrm{var}(X - Y) = 1 + 1 - 2/\sqrt{2}$$
$$= 2 - \sqrt{2}$$

Example 1

Take $d = 2$

$$X = f(\mathbf{Z}_X) = \frac{Z_X[1] + Z_X[2]}{\sqrt{2}}$$
$$Y = g(\mathbf{Z}_Y) = Z_Y[1]$$

Independence copula

$\mathrm{var}(X - Y) = 2$

CRN copula

$$\mathrm{var}(X - Y) = 1 + 1 - 2/\sqrt{2}$$
$$= 2 - \sqrt{2}$$

An optimal Gaussian copula:

$$\mathbf{Z}_Y[1] = \frac{Z_X[1] + Z_X[2]}{\sqrt{2}}$$
$$\mathbf{Z}_Y[2] = \frac{Z_X[1] - Z_X[2]}{\sqrt{2}}$$

so that $X = Y$ or, equivalently,

$$\boldsymbol{\Sigma}_{XY} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Example 2

Take $d = 1$

$$X = f_U(\mathbf{U}_X) = I(\mathbf{U}_X \in [0.5, 0.6])$$
$$Y = g_U(\mathbf{U}_Y) = I(\mathbf{U}_Y \in [0.7, 0.8])$$

Example 2

Take $d = 1$

$$X = f_U(\mathbf{U}_X) = I(\mathbf{U}_X \in [0.5, 0.6])$$
$$Y = g_U(\mathbf{U}_Y) = I(\mathbf{U}_Y \in [0.7, 0.8])$$

Independence copula

$$\begin{aligned}\mathrm{var}(X - Y) &= 2 \times 0.1 \times 0.9 \\ &= 0.18\end{aligned}$$

Example 2

Take $d = 1$

$$X = f_U(\mathbf{U}_X) = I(\mathbf{U}_X \in [0.5, 0.6])$$
$$Y = g_U(\mathbf{U}_Y) = I(\mathbf{U}_Y \in [0.7, 0.8])$$

Independence copula

$$\begin{aligned} \mathrm{var}(X - Y) &= 2 \times 0.1 \times 0.9 \\ &= 0.18 \end{aligned}$$

CRN copula

$$X - Y = \begin{cases} 1 & \text{w.p. } 0.1 \\ -1 & \text{w.p. } 0.1 \\ 0 & \text{w.p. } 0.8 \end{cases}$$

$\mathrm{var}(X - Y) = 0.2$

Example 2

Take $d = 1$

$$X = f_U(\mathbf{U}_X) = I(\mathbf{U}_X \in [0.5, 0.6])$$
$$Y = g_U(\mathbf{U}_Y) = I(\mathbf{U}_Y \in [0.7, 0.8])$$

Independence copula

$$\text{var}(X - Y) = 2 \times 0.1 \times 0.9$$
$$= 0.18$$

CRN copula

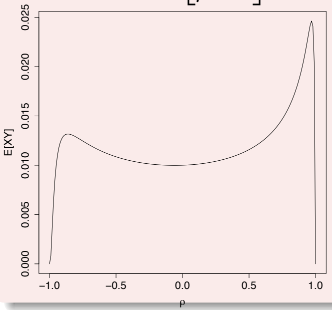
$$X - Y = \begin{cases} 1 & \text{w.p. } 0.1 \\ -1 & \text{w.p. } 0.1 \\ 0 & \text{w.p. } 0.8 \end{cases}$$

$\text{var}(X - Y) = 0.2$

Optimal Gaussian copula:

$\text{var}(X - Y) = 0.15$

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$



Optimizing the Copula

- $\min \operatorname{var}(X - Y) \Leftrightarrow \max \operatorname{cov}(X, Y) \Leftrightarrow \max Ef(\mathbf{Z}_X)g(\mathbf{Z}_Y)$

Optimizing the Copula

- min $\operatorname{var}(X - Y) \Leftrightarrow$ max $\operatorname{cov}(X, Y) \Leftrightarrow$ max $Ef(\mathbf{Z}_X)g(\mathbf{Z}_Y)$
- max $Ef\left(\left(\boldsymbol{\Sigma}^{1/2}\mathbf{N}\right)[1, \ldots, d]\right) g\left(\left(\boldsymbol{\Sigma}^{1/2}\mathbf{N}\right)[d+1, \ldots, 2d]\right)$
  subject to
  $$\boldsymbol{\Sigma} = \begin{bmatrix} I_d & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{XY}^T & I_d \end{bmatrix} \succeq 0$$

Optimizing the Copula

- min $\mathrm{var}(X - Y) \Leftrightarrow$ max $\mathrm{cov}(X, Y) \Leftrightarrow$ max $Ef(\mathbf{Z}_X)g(\mathbf{Z}_Y)$
- max $Ef\left(\left(\mathbf{\Sigma}^{1/2}\mathbf{N}\right)[1,\ldots,d]\right) g\left(\left(\mathbf{\Sigma}^{1/2}\mathbf{N}\right)[d+1,\ldots,2d]\right)$
  subject to
  $$\mathbf{\Sigma} = \begin{bmatrix} I_d & \mathbf{\Sigma}_{XY} \\ \mathbf{\Sigma}_{XY}^T & I_d \end{bmatrix} \succeq 0$$
- Nonlinear semidefinite program
- Can be tackled, but gradients of objective are tricky
- Reformulate using Cholesky factors directly by observing that $\mathbf{\Sigma}_{XY}$ is "sub-orthogonal" ...

An Alternative Formulation

Proposition

$$\mathbf{\Sigma} = \begin{bmatrix} I_d & \mathbf{\Sigma}_{XY} \\ \mathbf{\Sigma}_{XY}^T & I_d \end{bmatrix} \succeq 0 \Leftrightarrow \exists \mathbf{M}_2 : \mathbf{M}^T\mathbf{M} = I \textit{ where } \mathbf{M} := \begin{bmatrix} \mathbf{\Sigma}_{XY} \\ \mathbf{M}_2 \end{bmatrix}$$

Furthermore, $\mathbf{\Sigma}$ *is covariance matrix of*

$$\begin{bmatrix} \mathbf{Z}_X \\ \mathbf{Z}_Y \end{bmatrix} := \begin{bmatrix} \mathbf{N}[1,\ldots,d] \\ \mathbf{M}^T\mathbf{N} \end{bmatrix}$$

An Alternative Formulation

Proposition

$$\boldsymbol{\Sigma} = \begin{bmatrix} I_d & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{XY}^T & I_d \end{bmatrix} \succeq 0 \Leftrightarrow \exists \mathbf{M}_2 : \mathbf{M}^T \mathbf{M} = I \text{ where } \mathbf{M} := \begin{bmatrix} \boldsymbol{\Sigma}_{XY} \\ \mathbf{M}_2 \end{bmatrix}$$

Furthermore, $\boldsymbol{\Sigma}$ *is covariance matrix of*

$$\begin{bmatrix} \mathbf{Z}_X \\ \mathbf{Z}_Y \end{bmatrix} := \begin{bmatrix} \mathbf{N}[1, \ldots, d] \\ \mathbf{M}^T \mathbf{N} \end{bmatrix}$$

Notice that

$$\begin{bmatrix} I_d & 0_d \\ \boldsymbol{\Sigma}_{XY}^T & M_2^T \end{bmatrix} \begin{bmatrix} I_d & \boldsymbol{\Sigma}_{XY} \\ 0_d & M_2 \end{bmatrix} = \boldsymbol{\Sigma}$$

Solving the Alternative Formulation

- max $Ef(\mathbf{Z}_X)g(\mathbf{M}^T\mathbf{N})$ subject to $\mathbf{M}^T\mathbf{M} = I$
- Nonlinear optimization over a Stiefel manifold
- Gradients w.r.t. $\mathbf{M}$ easily obtained

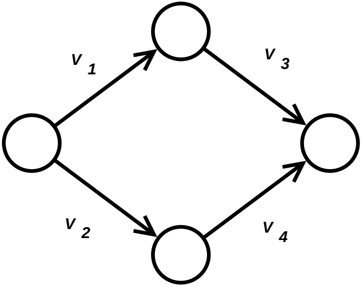
Solving the Alternative Formulation

- max $Ef(\mathbf{Z}_X)g(\mathbf{M}^T\mathbf{N})$ subject to $\mathbf{M}^T\mathbf{M} = I$
- Nonlinear optimization over a Stiefel manifold
- Gradients w.r.t. $\mathbf{M}$ easily obtained
- Use, e.g., sample-average approximation
- Sample and fix $\mathbf{N}_1, \ldots, \mathbf{N}_m$, and

$$\begin{aligned} \text{max} \quad & \frac{1}{m}\sum_{i=1}^{m} f(\mathbf{N}_i[1,\ldots,d])g(\mathbf{M}^T\mathbf{N}_i) \\ \text{subject to} \quad & \mathbf{M}^T\mathbf{M} = I \end{aligned}$$

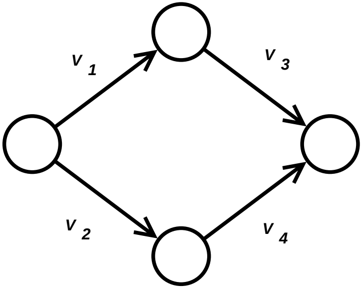
- Freely available `sgmin` in MATLAB
- Use solution to above in subsequent conditionally independent simulation

Example 3



$$X : V_1, V_2, V_3, V_4 \sim \text{ IID } \exp(1)$$
$$Y : V_1, V_2 \sim \text{ IID } \exp(1)$$
$$V_3 \sim \exp((1 + V_2)/2)$$
$$V_4 \sim \exp((1 + V_1)/2)$$

Example 3



Sampling Strategy	Variance
IID	5.3
CRN	0.57
OPT	0.28

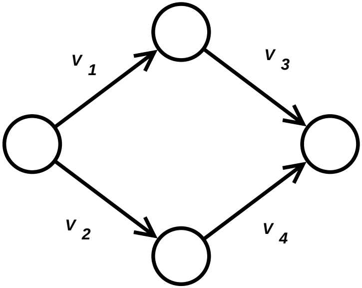
$$X : V_1, V_2, V_3, V_4 \sim \text{ IID } \exp(1)$$

$$Y : V_1, V_2 \sim \text{ IID } \exp(1)$$

$$V_3 \sim \exp((1 + V_2)/2)$$

$$V_4 \sim \exp((1 + V_1)/2)$$

Example 3



$X : V_1, V_2, V_3, V_4 \sim$ IID exp(1)

$Y : V_1, V_2 \sim$ IID exp(1)

$V_3 \sim \exp((1 + V_2)/2)$

$V_4 \sim \exp((1 + V_1)/2)$

Sampling Strategy	Variance
IID	5.3
CRN	0.57
OPT	0.28

$$\Sigma^*_{XY} = \begin{bmatrix} .958 & -.038 & .160 & .237 \\ -.037 & .960 & .239 & .141 \\ -.158 & -.238 & .957 & -.048 \\ -.239 & -.143 & -.026 & .960 \end{bmatrix}$$

An Observation for Linear Functions

- The optimal Σ_{XY} in Example 1 was orthogonal
- When is this true in general?

An Observation for Linear Functions

- The optimal Σ_{XY} in Example 1 was orthogonal
- When is this true in general?

Proposition

If f and g are linear, then an optimal Σ_{XY} is orthogonal and corresponds to a Householder transformation that "aligns" the two linear functions.

Conclusions and Future Research

- One might use more general joint distributions than CRN for comparisons
- Gaussian copula is particularly convenient
- Could use other copulas, e.g., chessboards, but computation needs to be feasible
- Simple examples demonstrate that large gains are possible

Conclusions and Future Research

- One might use more general joint distributions than CRN for comparisons
- Gaussian copula is particularly convenient
- Could use other copulas, e.g., chessboards, but computation needs to be feasible
- Simple examples demonstrate that large gains are possible
- When is optimization problem unimodal?
- Clarify connection to existing optimality results for CRN?
- More complicated (interesting?) examples?