Credit Risk Models

6. Modeling dependence and copulas

Andrew Lesniewski

Baruch College
New York

Fall 2016

Outline

Portfolio credit risk

- We now turn to the issues of modeling portfolios of credit risky assets.
- The central problem of portfolio credit risk modeling is understanding the impact of the dependence between credit events of the individual names on the portfolio as a whole.
- Exposure to portfolio credit risk is common in banking and insurance businesses.
- Various commonly traded structures such as *collateralized loan obligations* (CLOs) are exposed to portfolio credit risk.
- There is a class of credit derivatives, called *synthetic CDOs*, which are options on the portfolio losses in baskets of CDSs.

Portfolio loss function

- Let us consider portfolio consisting of risky N names. We fix a time horizon T and let l_i denote the *loss given default* (LGD), defined as par less the recovery value, for name i. As usual, let τ_i denote the default time for name i.
- Let $L(T)$ denote the total portfolio loss over the time horizon T. Thus

$$L(T) = \sum_{i=1}^{N} 1_{\tau_i \leq T}\, l_i .$$

- The expected loss is thus given by

$$\begin{aligned} \mathrm{E}[L(T)] &= \sum_{i=1}^{N} \mathrm{E}[1_{\tau_i \leq T}]\, l_i \\ &= \sum_{i=1}^{N} Q_i(T)\, l_i , \end{aligned}$$

where $Q_i(T)$ is the default probability of name i.

Portfolio loss function

- The expected loss can often be estimated from CDS markets (which allows us to infer risk-neutral default probabilities) or from rating (using historical probabilities).
- The primary driver of loss distributions is the dependence among the defaults of the individual names.
- Intuitively, higher dependence among the credit events lead to more extreme events with multiple defaults. At the same time, higher dependence increases the likelihood of events with few defaults.
- As an illustration, consider $N = 10$, with $l_i = 1$ and $Q_i(T) = 0.1$, for all i.

Portfolio loss function

- If all names are completely independent:

$$\mathrm{P}(L(T)=k)=\binom{N}{k}(0.1)^k(0.9)^{N-k},\ \text{where } n=0,\ldots,10.$$

In particular,

$$\begin{aligned}\mathrm{P}(L(T)=0)&=(0.9)^{10}=0.349,\\ \mathrm{P}(L(T)=N)&=(0.1)^{10}=10^{-10}.\end{aligned}$$

- If all names are perfectly dependent (if one firm defaults, they all default):

$$\begin{aligned}\mathrm{P}(L(T)=0)&=0.9,\\ \mathrm{P}(L(T)=k)&=0.0,\ \text{for all } k=1,\ldots,N-1,\\ \mathrm{P}(L(T)=N)&=0.1.\end{aligned}$$

- Consequently, more dependence implies higher probabilities of tail events.

Pearson correlation coefficient

- Consider N random variables $X_1, \ldots, X_N$ with joint distribution $H(X_1, \ldots, X_N)$. Their Pearson correlation matrix ρ is defined by

$$\rho_{ij} = \frac{\operatorname{Cov}(X_i, X_j)}{\sqrt{\operatorname{Var}(X_i)}\sqrt{\operatorname{Var}(X_j)}}\,.$$

- For example, if $X_i = 1_{\tau_i < T}$, then

$$\begin{aligned} \mathrm{E}[X_i^2] &= \mathrm{E}[X_i] \\ &= Q_i\,(T)\,, \end{aligned}$$

and so

$$\begin{aligned} \operatorname{Cov}(X_i, X_j) &= \mathrm{P}(\tau_i < T, \tau_j < T) - Q_i(T)Q_j(T), \\ \operatorname{Var}(X_i) &= Q_i(T)S_i(T). \end{aligned}$$

Correlation coefficient

- The Pearson correlation coefficient is a flawed measure of dependence:
    (i) It is not natural outside of Gaussian distributions.
    (ii) It only measures linear relationships.
    (iii) Perfectly dependent random variables do not necessarily have a correlation of 1. For example, if $X \sim N(0, 1)$, then X and $Y \triangleq X^2$ are perfectly dependent, but their correlation coefficient is 0. The correlation coefficient between X and $Z \triangleq X^3$ is $3/\sqrt{15} \approx 0.775$.
    (iv) Correlation is not invariant under monotone transformations of the random variables.
- Some other measures include concordance measures such as Kendall's tau and Spearman's rho.

Kendall's tau

- Let $\{(x_1, y_1), \ldots, (x_n, y_n)\}$ be a sample of observations from a pair of random variables X and Y with joint distribution $H(X, Y)$. Of the $\binom{n}{2}$ distinct pairs (x_i, y_i), (x_j, y_j), let c be the number of *concordant* pairs (if $x_i < x_j$ and $y_i < y_j$, or $x_i > x_j$ and $y_i > y_j$), and let d be the number of *discordant* pairs (if $x_i < x_j$ and $y_i > y_j$, or $x_i > x_j$ and $y_i < y_j$).
- Then *Kendall's* tau is defined as

$$\widehat{\tau}_{X,Y} = \frac{c-d}{c+d}\,.$$

- The probabilistic version of this sample definition is

$$\tau_K(X, Y) = \mathrm{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathrm{P}((X_1 - X_2)(Y_1 - Y_2) < 0),$$

where (X_1, Y_1) and (X_2, Y_2) are independent pairs of random variables drawn from H.

Spearman's rho

- We consider three pairs (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) of independent pairs of random variables drawn from the joint distribution H. Their Spearman' rho is defined by

$$\rho_S(X, Y) = 3\big(\mathsf{P}((X_1 - X_2)(Y_1 - Y_3) > 0) - \mathsf{P}((X_1 - X_2)(Y_1 - Y_3) < 0)\big).$$

- Spearman's rho is thus the difference of the probability of concordance and the probability of discordance for the vectors of random variables (X_1, Y_1) and (X_2, Y_3). These two pairs have the same margins, but the former has joint distribution H while the components of the latter are independent.

Definition of a copula

- Copulas are devices for constructing joint (multivariate) probability distributions out of marginal (univariate) probability distributions.
- To motivate the definition, we consider uniformly distributed random variables $U_i \in [0, 1]$, $i = 1, \ldots, N$.
- Their joint distribution is a function $C : [0, 1]^N \to [0, 1]$ such that

$$C(u_1, \ldots, u_N) = \mathrm{P}(U_1 \leq u_1, \ldots, U_N \leq u_N).$$

- Note that C satisfies the following conditions:
  (i)

$$\begin{aligned} C(u_1, \ldots, 0, \ldots, u_N) &= 0, \\ C(1, \ldots, 1, u_i, 1, \ldots, 1) &= u_i. \end{aligned}$$

  (ii) *Volume monotonicity condition*, which is elementary but somewhat cumbersome to write down. We shall formulate it explicitly in the $N = 2$ case only, see [2] or [3] for the general case.

Definition of a copula

- If $N = 2$, condition (ii) reads as follows: for $a_1 \leq b_1$ and $a_2 \leq b_2$,

$$C(b_1, b_2) - C(a_1, b_2) - C(b_1, a_2) + C(a_1, a_2) \geq 0.$$

- This is a consequence of the fact that $\mathrm{P}(a_1 \leq U_1 \leq b_1, a_2 \leq U_2 \leq b_2) \geq 0$, i.e.

$$\begin{aligned} 0 &\leq \int_{a_1}^{b_1} \int_{a_2}^{b_2} dC(u_1, u_2) \\ &= C(b_1, b_2) - C(a_1, b_2) - C(b_1, a_2) + C(a_1, a_2), \end{aligned}$$

  where

$$dC(u_1, u_2) \triangleq \frac{\partial^2 C(u_1, u_2)}{\partial u_2 \partial u_2} du_1 du_2.$$

- *Definition*: Any function $C : [0, 1]^N \to [0, 1]$ satisfying conditions (i) and (ii) is called a *copula*.
- A copula with $N = 2$ is called a *bivariate copula*.

Simple copulas

- The independence copula is defined by

$$C(u_1, \ldots, u_N) = \prod_{i=1}^{N} u_i .$$

- The dependence copula is defined by

$$C(u_1, \ldots, u_N) = \min_i u_i .$$

- The anti-dependence copula can be defined for $N = 2$:

$$C(u_1, u_2) = \max(u_1 + u_2 - 1, 0).$$

Joint probability distributions and copulas

- Consider N continuous random variables $X_1, \ldots, X_N$ with monotone increasing marginal probability distributions $F_1(x), \ldots, F_N(x)$, i.e. $F_i(x) = \mathsf{P}(X_i \leq x)$, for $i = 1, \ldots, N$. Let $H(X_1, \ldots, X_N)$ denote their joint probability distribution, i.e.

$$H(x_1, \ldots, x_N) = \mathsf{P}(X_1 \leq x_1, \ldots, X_N \leq x_N).$$

- For each i, consider the random variable $U_i = F_i(X_i)$. Clearly, $U_i \in [0, 1]$. Since

$$\begin{aligned}\mathsf{P}(U_i \leq u) &= \mathsf{P}(X_i \leq F_i^{-1}(u)) \\ &= F_i(F_i^{-1}(u)) \\ &= u,\end{aligned}$$

U_i is uniformly distributed on $[0, 1]$.

Sklar's theorem

- The function
$$C(u_1, \ldots, u_N) = H(F_1^{-1}(u_1), \ldots, F_N^{-1}(u_N))$$
defines a copula.
- Conversely, given a copula C, the function
$$H(x_1, \ldots, x_N) = C(F_1(x_1), \ldots, F_N(x_N))$$
defines a joint probability function with marginals $F_1, \ldots, F_N$.
- The last two statements form the content of *Sklar's theorem*.
- Actually, Sklar's theorem is a bit more general than what we formulated above.

Gaussian copula

- Consider an N-dimensional Gaussian distribution Φ_ρ with correlation matrix ρ.
- A random vector Z distributed according to Φ_ρ has zero mean and each of its components has variance one.
- The Gaussian copula is defined by

$$C_\rho(u_1, \ldots, u_N) = \Phi_\rho(N^{-1}(u_1), \ldots, N^{-1}(u_N)),$$

where $N(x)$ is the standard uniform Gaussian distribution function.

Gaussian copula

- Explicitly, in the bivariate case,

$$C_\rho(u_1, u_2) = \frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \int_{-\infty}^{N^{-1}(u_1)} \int_{-\infty}^{N^{-1}(u_2)} \exp\Big\{-\frac{s^2 - 2\rho_{12} su + u^2}{2(1-\rho_{12}^2)}\Big\} ds du.$$

- Note that the correlation matrix ρ is the driver of dependence for the Gaussian copula.
- This is one of the reason's of the popularity of the Gaussian copula: its dependence parameters are the correlation coefficients between the random variables (which are not necessarily Gaussian).

Student t copula

- Let Z be a random vector distributed according to Φ_ρ, and let Y be a scalar random variable distributed according to the χ^2-distribution with ν degrees of freedom, $S \sim \chi^2_\nu$.
- The Student t copula $C_{\nu,\rho}$ with ν degrees of freedom is defined as the copula corresponding to the N-dimensional random variable $X = \sqrt{\nu/S}Z$,

$$C_{\nu,\rho}(u_1, \ldots, u_n) = t_{\nu,\rho}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_N)),$$

where $t_{\nu,\rho}$ is the distribution of $\sqrt{\nu/S}Z$, and t_ν is the distribution of $\sqrt{\nu/S}Z_1$.

Student t copula

- Explicitly, in the bivariate case,

$$C_{\nu,\rho}(u_1,u_2)=\frac{1}{2\pi\sqrt{1-\rho_{12}^2}}\int_{-\infty}^{t_\nu^{-1}(u_1)}\int_{-\infty}^{t_\nu^{-1}(u_2)}\Big(1+\frac{s^2-2\rho_{12}su+u^2}{\nu(1-\rho_{12}^2)}\Big)^{-\frac{\nu+2}{2}}dsdu.$$

- The Student t copula has fatter tails than the Gaussian copula (the smaller ν, the fatter the tails).
- It is used to model situations where tail events are believed to play a role.

Archimedean copulas

- These are simple copulas that have interesting properties but (so far) little use in finance.
- A function $\phi : [0, 1] \to [0, \infty)$ is called a *copula generator* if it satisfies the properties:
  (i) It is monotone decreasing,
  (ii) $\phi(0) = \infty$,
  (iii) $\phi(1) = 0$.
- Given a copula generator, and Archimedean copula is defined by

$$C(u_1, \ldots, u_N) = \phi^{-1}(\phi(u_1) + \ldots + \phi(u_N)).$$

Archimedean copulas

- Below are some of the more popular Archimedean copulas.
- *Independence copula*:
$$\phi(u) = -\log u.$$
- *Clayton copula*: for $\theta > 0$,
$$\begin{aligned} \phi(u) &= u^{-\theta} - 1, \\ C_{Clayton}(u_1, \ldots, u_N) &= (u_1^{-\theta} + \ldots u_N^{-\theta} - N + 1)^{-1/\theta}. \end{aligned}$$

Archimedean copulas

- *Gumbel copula*: for $\theta > 1$,

$$\phi(u) = (-\log u)^{\theta},$$
$$C_{Gumbel}(u_1, \ldots, u_N) = \exp\Big\{ -\big((-\log u_1)^{\theta} + \ldots + (-\log u_N)^{\theta}\big)^{1/\theta} \Big\}.$$

- *Frank copula*: for $\theta \in \mathbb{R}$,

$$\phi(u) = -\log \frac{e^{-\theta u} - 1}{e^{-\theta} - 1},$$
$$C_{Frank}(u_1, \ldots, u_N) = -\frac{1}{\theta} \log \Big(1 + \frac{(e^{-\theta u_1} - 1) \ldots (e^{-\theta_N u} - 1)}{(e^{-\theta} - 1)^{N-1}} \Big).$$

Copulas and dependence measures

- Let (X, Y) be a vector of continuous random variables with copula C. Then
  - Kendall's tau is given by

$$\tau_K(X, Y) = 4 \iint_{[0,1]^2} C(u, v) dC(u, v) - 1.$$

(i) Spearman's rho is given by

$$\begin{aligned}\rho_S(X, Y) &= 12 \iint_{[0,1]^2} uv dC(u, v) - 3 \\ &= 12 \iint_{[0,1]^2} C(u, v) du dv - 3.\end{aligned}$$

(ii) Note that these results do not depend on the marginals of X and Y.

Copulas and dependence measures

- For Clayton's copula,

$$\tau_K = \frac{\theta}{\theta + 2}\,.$$

- For Gumbel's copula,

$$\tau_K = 1 - \frac{1}{\theta}\,.$$

- For Gaussian and Student t copulas, the following relation holds:

$$\rho = 2\sin(\pi\rho_S/6)\,.$$

Tail dependence

- The concept of tail dependence relates to the amount of dependence in the upper right quadrant tail or lower left quadrant tail of a bivariate distribution. This concept is relevant for the study of dependence between tail ($\equiv$ extreme) events.
- It turns out that tail dependence between two continuous random variables X and Y is a copula property and hence the amount of tail dependence is invariant under strictly increasing transformations of X and Y.
- Consider a bivariate copula is such that the limit

$$\lim_{u \to 1} (1 - 2u + C(u,u))/(1-u) \triangleq \lambda_U$$

exists. Then C has *upper tail dependence* if $0 < \lambda_U \leq 1$, and no upper tail dependence if $\lambda_U = 0$

- Similarly, if

$$\lim_{u \to 0} C(u,u)/u \triangleq \lambda_L.$$

exists, then C has *lower tail dependence* if $0 < \lambda_L \leq 1$, and no lower tail dependence if $\lambda_L = 0$.

Tail dependence

- The definition is justified by the following calculation

$$\begin{aligned}\lambda_U &= \lim_{u\to 1} \mathrm{P}(U_1 > u | U_2 > u)\\ &= \lim_{u\to 1} \mathrm{P}(U_2 > u | U_1 > u).\end{aligned}$$

  However,

$$\lim_{u\to 1} \mathrm{P}(U_1 > u | U_2 > u) = \lim_{x\to\infty} \mathrm{P}(X > x | Y > x).$$

  This shows that λ_U measures the co-dependence of X and Y at the right tail: $\lambda_U > 0$ if and only if the conditional probabilities $\mathrm{P}(X > x | Y > x)$ and $\mathrm{P}(Y > x | X > x)$ are positive for large values of x.

- A similar argument can be made for λ_L.

Tail dependence

- For the bivariate Gaussian copula, $|\rho_{12}| < 1$,

$$\lambda_U = 0,$$
$$\lambda_L = 0.$$

- For the bivariate Student t copula, $|\rho_{12}| < 1$,

$$\lambda_U = 2t_{\nu+1}(-\sqrt{\nu+1}\sqrt{(1-\rho_{12})/(1+\rho_{12})}),$$
$$\lambda_L = \lambda_U.$$

- For the bivariate Clayton copula, $\theta > 0$,

$$\lambda_U = 0,$$
$$\lambda_L = 2^{1/\theta}$$

- For the bivariate Gumbel copula, $\theta > 1$,

$$\lambda_U = 2 - 2^{1/\theta},$$
$$\lambda_L = 0.$$

Marshall-Olkin copula

- Consider a two-component system where the components are subject to shocks, which are fatal to one or both components. We assume that these shocks arrive at times T_1, T_2, and T_{12} as independent Poisson processes, with intensities of λ_1, λ_2, and λ_{12}, respectively.
- In credit applications, the components are credit names, and λ_1, λ_2, and λ_{12} are default intensities. The shared shock λ_{12} can be interpreted as global shock, that affects multiple names simultaneously.
- Let $\tau_1 = \min(T_1, T_{12})$ and $\tau_2 = \min(T_2, T_{12})$ denote the lifetimes of the two components.
- The joint survival probability is given by

$$\begin{aligned} \mathrm{P}(\tau_1 > t_1, \tau_2 > t_2) &= \mathrm{P}(T_1 > t_1)\mathrm{P}(T_2 > t_2)\mathrm{P}(T_{12} > \max(t_1, t_2)) \\ &= \exp(-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)). \end{aligned}$$

- The marginal survival probabilities are

$$\begin{aligned} \mathrm{P}(\tau_1 > t_1) &= \exp(-(\lambda_1 + \lambda_{12})t_1), \\ \mathrm{P}(\tau_2 > t_1) &= \exp(-(\lambda_2 + \lambda_{12})t_2). \end{aligned}$$

Marshall-Olkin copula

- We now find the copula C_{MO} for this distribution. To this end, we express $\mathrm{P}(\tau_1 > t_1, \tau_2 > t_2)$ in terms of $\mathrm{P}(\tau_1 > t_1)$ and $\mathrm{P}(\tau_2 > t_2)$
- Since

$$\max(t_1, t_2) = t_1 + t_2 - \min(t_1, t_2),$$

we have

$$\mathrm{P}(\tau_1 > t_1, \tau_2 > t_2) = \mathrm{P}(\tau_1 > t_1)\mathrm{P}(\tau_2 > t_2)\min(\exp(\lambda_{12} t_1), \exp(\lambda_{12} t_2)).$$

- Set

$$\begin{aligned} u_1 &= \mathrm{P}(\tau_1 > t_1), \\ u_2 &= \mathrm{P}(\tau_2 > t_2), \end{aligned}$$

and

$$\begin{aligned} \alpha_1 &= \lambda_{12}/(\lambda_1 + \lambda_{12}), \\ \alpha_2 &= \lambda_{12}/(\lambda_2 + \lambda_{12}). \end{aligned}$$

Marshall-Olkin copula

- Then

$$
\begin{aligned}
\exp(\lambda_{12} t_1) &= u_1^{-\alpha_1}, \\
\exp(\lambda_{12} t_2) &= u_2^{-\alpha_2}.
\end{aligned}
$$

- This leads to the following copula:

$$
\begin{aligned}
C_{MO}(u_1, u_2) &= u_1 u_2 \min(u_1^{-\alpha_1}, u_2^{-\alpha_2}) \\
&= \min(u_1^{1-\alpha_1} u_2, u_1 u_2^{1-\alpha_2}).
\end{aligned}
$$

- Notice that the Marshall-Olkin copula is an example of a *survival copula*, which expresses the joint survival probability in terms of marginal survival probabilities. This is unlike the copulas discussed above, which express the joint event probabilities in terms of marginal event probabilities.

Marshall-Olkin copula

- The Marshall-Olkin copula can be extended to more than two names in a straightforward manner.
- However, its complexity increases with the number of components. There are different kinds of shared shocks: some will affect all names, some will affect only subsets (like those in a particular industry).
- In practice, the specification of these shocks and their intensities is impractical, as the combinatorics get complicated, and the Marshall-Olkin copula is not used often in practice.
- Elements of the MO copula approach, specifically the idea of shared shocks, are used in combination with other models.

Algorithm for simulating Gaussian copula

- Consider the N-dimensional Gaussian distribution $N(0, \rho)$ with mean 0 and correlation matrix ρ.
- Find the Cholesky decomposition of ρ, i.e. $\rho = LL^{\mathrm{T}}$, where L is an $N \times N$-dimensional lower triangular matrix. Then repeat the following steps a desired number of times:
  - (i) *Step 1*. Simulate N independent samples $Z_1, \ldots, Z_N$ from $N(0, 1)$.
  - (ii) *Step 2*. Calculate $X_1 = LZ_1, \ldots, X_N = LZ_N$.
  - (iii) *Step 3*. Convert this sample to a correlated N-dimensional uniform vector $U_1 = N(X_1), \ldots, U_N = N(X_N)$.
- Then $(U_1, \ldots, U_N) \sim C_\rho$.

Algorithm for simulating Student t copula

- The algorithm is similar to the one above (Gaussian copula). Sampling from a χ^2-distribution can be done by standard acceptance-rejection techniques or by numerically inverting the cumulative distribution function
- Find the Cholesky decomposition L of ρ, and repeat the following steps:
  - (i) *Step 1*. Simulate N independent samples $Z_1, \ldots, Z_N$ from from $N(0,1)$.
  - (ii) *Step 2*. Simulate a random variate S from χ^2_ν independent of $Z_1, \ldots, Z_N$.
  - (iii) *Step 3*. Calculate $Y = LZ$.
  - (iv) *Step 4*. Set $X = \sqrt{\nu/S}Y$.
  - (v) *Step 5*. Set $U_i = t(X_i)$, for $i = 1, \ldots, N$.
- Then $(U_1, \ldots, U_N) \sim C_{\nu\rho}$.

References

Embrechts, P., Lindskog, F., and McNeil, A.: Modelling dependence with copulas and applications to risk management, in *Handbook of Heavy Tailed Distributions in Finance*, Elsevier (2003).

Joe, H.:Dependence Modeling with Copulas, Chapman & Hall (2014).

Nelsen, R. B.: *An Introduction to Copulas*, Springer (1999).