MFE 409 Lecture 6 Extra Material
Correlations

Valentin Haddad

Spring 2025

UCLA Anderson

Lecture Objectives

- Quick reminder of the math of correlation
- The copula model of correlation
- Three models of credit risk incorporating correlation

Outline

1. Basic Concepts
2. Copulas
3. Applications to Credit Risk
   - Vasicek Model
   - Credit Risk Plus
   - CreditMetrics

Covariance and Correlation

- Covariance

$$\text{cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

- Correlation

$$\rho = \text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$$

  - $\rho \in [-1, 1]$
  - $\rho = 0 \rightarrow X$ and Y are uncorrelated

Correlation and Dependence

- X and Y are independent if

$$\forall x, \quad f(Y|X=x) = f(Y)$$

  - f: pdf

Correlation and Dependence

- X and Y are independent if

$$\forall x, \quad f(Y|X=x) = f(Y)$$

  - f: pdf

- Example 1: X and Y are uncorrelated but not independent

- Example 2: X and Y are independent but not uncorrelated

Estimating Correlation

- Assume we are at date n and we have m past observations
- Covariance:

$$\mathsf{cov}_n = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}y_{n-i}$$

- Variance:

$$\mathsf{var}_{y,n} = \frac{1}{m}\sum_{i=1}^{m} y_{n-i}^2$$

- Correlation:

$$\rho_n = \frac{\mathsf{cov}_n}{\sqrt{\mathsf{var}_{x,n}\mathsf{var}_{y,n}}}$$

Other Estimation Techniques

- Always focus on covariances and variances, then compute correlation
- EWMA

$$\text{cov}_n = \lambda \text{cov}_{n-1} + (1-\lambda) x_{n-1} y_{n-1}$$

- GARCH

$$\text{cov}_n = \omega + \alpha x_{n-1} y_{n-1} + \beta \text{cov}_{n-1}$$

Consistency Condition for Covariances

- With N variables, variance-covariance matrix Σ, dimension $N \times N$

- Σ must be positive semi-definite

$$\forall w, \quad w'\Sigma w \geq 0$$

Consistency Condition for Covariances

- With N variables, variance-covariance matrix Σ, dimension $N \times N$

- Σ must be positive semi-definite

$$\forall w, \quad w'\Sigma w \geq 0$$

- Important to check that estimation method satisfies this condition
  - Use same weighting for variances and covariances
  - EWMA ok, GARCH more tricky

Consistency Condition for Symmetric Distributions

- Assume n variables such that $\forall i \neq j,\ \text{corr}(X_i, X_j) = \rho$
- What are the possible values of ρ?

Multivariate Normal Distributions

- X_1 and X_2 multivariate normal with means μ_i, variance σ_i^2 and correlation ρ
- Conditional on X_1, X_2 is normal with:

Multivariate Normal Distributions

- X_1 and X_2 multivariate normal with means μ_i, variance σ_i^2 and correlation ρ
- Conditional on X_1, X_2 is normal with:

$$\begin{aligned} \text{mean} \quad & \mu_2 + \rho\sigma_2 \frac{X_1 - \mu_1}{\sigma_1} \\ \text{std. deviation} \quad & \sigma_2\sqrt{1-\rho^2} \end{aligned}$$

- Provides a way to simulate correlated normal distributions

Multivariate Normal Distributions

- X_1 and X_2 multivariate normal with means μ_i, variance σ_i^2 and correlation ρ
- Conditional on X_1, X_2 is normal with:

$$\text{mean} \quad \mu_2 + \rho\sigma_2 \frac{X_1 - \mu_1}{\sigma_1}$$

$$\text{std. deviation} \quad \sigma_2\sqrt{1-\rho^2}$$

- Provides a way to simulate correlated normal distributions
- 1-factor model with n variables, where F and Z_i are independent standard normal:

$$X_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

Outline

1. Basic Concepts

2. Copulas

3. Applications to Credit Risk
   - Vasicek Model
   - Credit Risk Plus
   - CreditMetrics

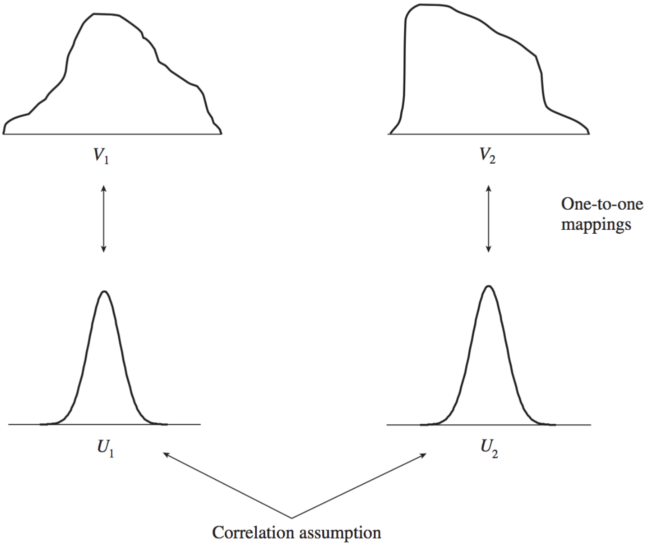
Copula: Underlying Idea

- Two variables with complicated marginal distributions, want to model their correlation

- *Step 1*: Transform the variables into a known distribution, e.g. normal distribution

- *Step 2*: Assume the transformed variables are distributed from the multivariate version of the known distribution

Copula



Gaussian Copula

- Two variables X_1 and X_2 with marginal cdf G_1 and G_2
- *Step 1*: Want transformed variables X_1 and X_2 to be standard normal:

$$\begin{aligned} G_1(X_1) &= \mathcal{N}(Y_1) \\ Y_1 &= \mathcal{N}^{-1}[G_1(X_1)] \\ Y_2 &= \mathcal{N}^{-1}[G_2(X_2)] \end{aligned}$$

- *Step 2*: Assume that Y_1 and Y_2 have correlation ρ: joint cdf $\mathcal{N}(.,.;\rho)$

Gaussian Copula

- Two variables X_1 and X_2 with marginal cdf G_1 and G_2
- *Step 1*: Want transformed variables X_1 and X_2 to be standard normal:

$$\begin{aligned} G_1(X_1) &= \mathcal{N}(Y_1) \\ Y_1 &= \mathcal{N}^{-1}[G_1(X_1)] \\ Y_2 &= \mathcal{N}^{-1}[G_2(X_2)] \end{aligned}$$

- *Step 2*: Assume that Y_1 and Y_2 have correlation ρ: joint cdf $\mathcal{N}(.,.;\rho)$
- Joint cdf:

$$\begin{aligned} G(x_1, x_2) &= \mathcal{N}(y_1, y_2; \rho) \\ &= \mathcal{N}(\mathcal{N}^{-1}[G_1(X_1)], \mathcal{N}^{-1}[G_2(X_2)]; \rho) \end{aligned}$$

Other Copula Models

- Student copula: Replace normal distribution by bivariate Student's t
  - Builds in more tail dependence

- Multivariate copula: multivariate distribution for N variables

- Factor copula model: assume that the transformed variables follow:

$$Y_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

Outline

1. Basic Concepts
2. Copulas
3. Applications to Credit Risk
   - Vasicek Model
   - Credit Risk Plus
   - CreditMetrics

Vasicek Model of Loan Portfolio

- Developed at KMV in 1987
- Application of the one-factor Gaussian copula model
- Used to compute Basel II capital requirements

Vasicek Model of Loan Portfolio

- Developed at KMV in 1987
- Application of the one-factor Gaussian copula model
- Used to compute Basel II capital requirements
- Assume a portfolio of many loans, and define T_i as the time when company i defaults.
- Given a horizon T, we define the probability of default at time T:

$$\text{PD} = \mathbb{P}(T_i < T)$$

  - If defaults are uncorrelated, PD will be the fraction of loans which actually default ex post

Vasicek Model: Gaussian Factor Copula

- Map T_i to a variable U_i which is standard normal

- Assume that:

$$U_i = \sqrt{\rho}F + \sqrt{1-\rho}Z_i$$

  - ρ is the *copula correlation*

Worst-Case Default Rate

- In this model the fraction of loans which default is random and depends of the realization of the aggregate factor
  - Law of large numbers for the residuals Z_i

Worst-Case Default Rate

- In this model the fraction of loans which default is random and depends of the realization of the aggregate factor
  - Law of large numbers for the residuals Z_i

- We are interested in the **worst-case default rate**: probability of individual default ($=$ fraction of loans which defaults) that will not be exceeded in with probability c
  - This is the VaR at confidence level c if each loan pays 1 and there is no recovery.

Worst-Case Default Rate: Derivation

- We are interested in:

$$\mathbb{P}(T_i < T | \text{bad state}) = \mathbb{P}(U_i < U | \text{bad state})$$

Worst-Case Default Rate: Derivation

- We are interested in:

$$\mathbb{P}(T_i < T | \text{bad state}) = \mathbb{P}(U_i < U | \text{bad state})$$

- Transforming the threshold T into U using unconditional probability:

$$\mathbb{P}(T_i < T) = \mathbb{P}(U_i < U)$$

$$\text{PD} = \mathcal{N}(U)$$

Worst-Case Default Rate: Derivation

- We are interested in:

$$\mathbb{P}(T_i < T|\text{bad state}) = \mathbb{P}(U_i < U|\text{bad state})$$

- Transforming the threshold T into U using unconditional probability:

$$\mathbb{P}(T_i < T) = \mathbb{P}(U_i < U)$$

$$\text{PD} = \mathcal{N}(U)$$

- Worst-case scenarios correspond to low values of F:

$$\mathbb{P}[F > \underline{F}] = c$$

$$\underline{F} = \mathcal{N}^{-1}(1-c)$$

$$= -\mathcal{N}^{-1}(c)$$

Worst-Case Default Rate: Derivation

- Using the definition of U_i

Worst-Case Default Rate: Derivation

- Using the definition of U_i

$$
\begin{aligned}
U_i &< U \\
\Leftrightarrow \sqrt{\rho}F + \sqrt{1-\rho}Z_i &< U \\
\Leftrightarrow Z_i &< \frac{U - \sqrt{\rho}F}{\sqrt{1-\rho}}
\end{aligned}
$$

Worst-Case Default Rate: Derivation

- Using the definition of U_i

$$
\begin{aligned}
U_i &< U \\
\Leftrightarrow \sqrt{\rho} F + \sqrt{1-\rho} Z_i &< U \\
\Leftrightarrow Z_i &< \frac{U - \sqrt{\rho} F}{\sqrt{1-\rho}}
\end{aligned}
$$

- Worst-case default rate:

$$
\begin{aligned}
\mathsf{WCDR} &= \mathbb{P}(T_i < T | \underline{F}) \\
&= \mathbb{P}(U_i < U | \underline{F}) \\
&= \mathcal{N}\left(\frac{U - \sqrt{\rho} \underline{F}}{\sqrt{1-\rho}} \right) = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(\mathsf{PD}) + \sqrt{\rho} \mathcal{N}^{-1}(c)}{\sqrt{1-\rho}} \right)
\end{aligned}
$$

Estimating PD and ρ

- The previous formula is valid for any level c: reveals the cdf F of default rate:

$$\mathsf{DR} = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(\mathsf{PD}) + \sqrt{\rho}\mathcal{N}^{-1}(F(\mathsf{DR}))}{\sqrt{1-\rho}}\right)$$

$$F(\mathsf{DR}) = \mathcal{N}\left(\frac{\sqrt{1-\rho}\mathcal{N}^{-1}(\mathsf{DR}) - \mathcal{N}^{-1}(\mathsf{PD})}{\sqrt{\rho}}\right)$$

Estimating PD and ρ

- The previous formula is valid for any level c: reveals the cdf F of default rate:

$$\text{DR} = \mathcal{N}\left(\frac{\mathcal{N}^{-1}(\text{PD}) + \sqrt{\rho}\mathcal{N}^{-1}(F(\text{DR}))}{\sqrt{1-\rho}}\right)$$

$$F(\text{DR}) = \mathcal{N}\left(\frac{\sqrt{1-\rho}\mathcal{N}^{-1}(\text{DR}) - \mathcal{N}^{-1}(\text{PD})}{\sqrt{\rho}}\right)$$

- Differentiate to obtain the pdf f:

$$f(\text{DR}) = \sqrt{\frac{1-\rho}{\rho}} \exp\left[\frac{1}{2}\left[\left(\mathcal{N}^{-1}(\text{DR})\right)^2 - \left(\frac{\sqrt{1-\rho}\mathcal{N}^{-1}(\text{DR}) - \mathcal{N}^{-1}(\text{PD})}{\sqrt{\rho}}\right)^2\right]\right]$$

- Maximum likelihood using historical data on default frequencies

Credit Risk Plus

- Developped by Credit Suisse Financial Products in 1997, to deal with the fact that defaults are rare
- Assume probability of default of each loan is q, and there are n loans, with independent defaults:

$$\begin{aligned}\mathbb{P}(m \text{ defaults}) &= \frac{n!}{m!(n-m)!} q^m (1-q)^{n-m} \\ &\approx \frac{e^{-qn}(qn)^m}{m!}\end{aligned}$$

Credit Risk Plus

- Developped by Credit Suisse Financial Products in 1997, to deal with the fact that defaults are rare
- Assume probability of default of each loan is q, and there are n loans, with independent defaults:

$$\begin{aligned}\mathbb{P}(m \text{ defaults}) &= \frac{n!}{m!(n-m)!} q^m (1-q)^{n-m} \\ &\approx \frac{e^{-qn}(qn)^m}{m!}\end{aligned}$$

- Uncertainty about aggregate conditions: assume qn follows a gamma distribution with mean μ and standard deviation σ:

$$\mathbb{P}(m \text{ defaults}) = p^m (1-p)^\alpha \frac{\Gamma(m+\alpha)}{\Gamma(m+1)\gamma(\alpha)}$$

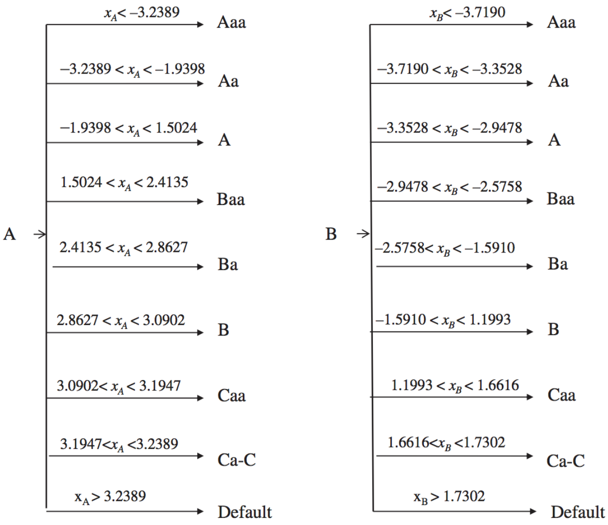
  - $\alpha = \mu^2/\sigma^2$, $p = \sigma^2/(\mu^2+\sigma^2)$

Credit Risk Plus In Practice

1. Sample overall default rate
2. Sample individual default rates using a model mapping overall to individual for each category
3. Sample individual defaults
4. Sample loss given default
5. Aggregate to the portfolio and obtain VaR

CreditMetrics

- Developed by JP Morgan in 1997



CreditMetrics

- Gaussian Copula on rating transitions

- Correlation across ratings: using correlation of equity returns

- Allows to capture credit risk, but also the market risk component of credit portfolio

Concluding Remarks

- Correlations at the heart of estimating risk for a portfolio: bad scenarios are cases where *everything* goes wrong
- Normal distribution makes it easy to deal with correlations, but not adapted to credit events
- Copula is a way to recover a bit of the simplicity of normal distributions

Concluding Remarks

- Correlations at the heart of estimating risk for a portfolio: bad scenarios are cases where *everything* goes wrong
- Normal distribution makes it easy to deal with correlations, but not adapted to credit events
- Copula is a way to recover a bit of the simplicity of normal distributions ...
- But it is far from perfect!
  - Struggles to match both low average individual probabilities and high defaults in a crash
  - Not designed to capture dynamics
- Another route: hazard models: dynamics of hazard rates
- Still no great model out there