Collateralized Debt Obligation and Copula Models

Certificate in Quantitative Finance

Siyi Zhou

Topics covered in this lecture

1. a review of Collateralized Debt Obligation (CDO) and its pricing – calculation of fair spread
2. the concept of Copula and its use in modelling dependence within a multidimensional variable (e.g., default time)
3. use of copula function to generate correlated default times – sampling from a copula
4. pricing a synthetic CDO with a single-factor copula representation
5. Example of Gaussian and Student's t copulae
6. Example of one-period pricing for a synthetic CDO

By the end of this lecture you will be able to

- understand the structuring of a synthetic CDO
- understand the concept of copula as a powerful tool to model joint events
- generate joint default times by sampling with Gaussian and Student's t copulae
- price a synthetic CDO tranche using a factor model for Gaussian copula
- understand the role of default correlation in credit risk models
- understand pricing for any generic multi-name credit derivative

What is CDO?

Collateralised Debt Obligation is an asset-backed security.

It is an diversified investment in a pool of assets (e.g., bonds, loans, CDS) structured in the form of tranches: assets are 'sliced' into tranches according to their *credit risk*.

Instead of selling the assets individually, a CDO sponsor (e.g., a bank directly, asset management company, or non-financial institutions) creates an independent legal entity called Special Purpose Vehicle (SPV).

CDO Tranches

The capital structure of a typical CDO is

- Senior tranche, rated AAA
- Mezzanine tranche, rated BB to AA
- Equity tranche, unrated

The position of each tranche is determined by its attachment and detachment points.

How CDO Works

- At good times, tranche Investors will receive a regular premium (made of returns from the asset pool).
- At bad times, when the total loss reaches an attachment point of a tranche, investors in that tranche start to lose their capital, and when the total loss reaches the detachment point, the investors lose all their capital invested. No further loss can occur to them.
- Loss is applied in a reverse order of seniority: Senior tranche is protected by Mezzanine and Equity tranches, Mezzanine tranche is protected only by Equity tranche.

Motivation

First issued in the late 1980s, CDO emerged a decade later as the fastest growing sector in financial markets. Banks that are actively trading CDO were motivated by two reasons:

1. Explore arbitrage opportunities
(the majority of CDO were arbitrage-driven).
2. Offload credit risk from a bank's loan book—this achieves reduction in regulatory capital requirement and improves return on risk capital.

Cash CDO

Today, CDO is a broad term that can refer to different products. From a pricing point of view we are interested to distinguish between cash and synthetic CDO (by source of funding).

Cash CDO involves a purchase of a liquid pool of assets (e.g., corporate bonds, loans, etc.). Ownership of the assets is transferred to a Special Purpose Vehicle who acts as the issuer of CDO.

Synthetic CDO

If—rather than owning the asset pool—Special Purpose Vehicle obtains exposure by **selling** Credit Default Swaps on the reference portfolio, such CDO is referred to as synthetic.

Unfunded synthetic CDO means that investors pay only when their tranches are affected by defaults. In that case, counterparty default risk (investors' ability to pay the notional amount insured, minus recovery) must be taken into account.

CDO Markets

CDO markets is organised in two segments:

- An illiquid segment, for example Cash CDO “buy-and-hold” investors whose decisions are mainly based on ratings and yields, and
- An actively traded segment in which the underlying credit portfolio is based on the standardized CDS index, such as the iTraxx (European) or CDX (North American) index series. This market is traded by the correlation trading desks.

The net cash flows of index tranches are the same as for synthetic CDO tranches. We will study synthetic CDO pricing in this lecture.

Cash CDO and Illiquid Markets

In the illiquid segment, CDO pricing and risk measurement are difficult.

- Market-to-market pricing and hedging are generally unavailable.
- A cash CDO often has a complex waterfall structure making the cash flow from its tranches highly *path dependent*.
- Such CDO structures require active management, challenged by the difficulty of key parameter estimation (e.g., *default correlation*). Resulting market-to-model pricing is very weak.

Traditionally, investors relied on rating agencies for valuation and risk assessment. But since the credit crunch, the credibility of the agencies and models has been severely undermined.

Synthetic CDO and Liquid Index-based Markets

However, for the Synthetic CDO markets,

- The cash flows of synthetic CDO tranches are very simple and **not** path-dependent.
- Market participants can easily take long or short positions in the CDX (North American) and other index tranches as well as the underlying single-name CDS instruments.
- The basis between index tranches and single-name CDS instruments tends to stay within a reasonable range because of a strong arbitrage relationship.

Therefore the valuation and risk management in index-based markets are simpler. In this lecture, we will focus on synthetic CDO instruments, treating a CDO as (a derivative of) a portfolio of CDS contracts.

Notation

Let's get familiar with notation before deriving the pricing equation for a synthetic CDO.

- survival time for reference name i is τ_i
- loss given default for reference name i is LGD_i
- exposure at default for reference name i is EAD_i
- tranche with attachment and detachment points $[D, U]$
- settlement time (date) t_j
- tenor $\Delta = t_{j+1} - t_j$
- discount factor, taken from zero coupon bond price, $Z(t, T)$

Loss Function

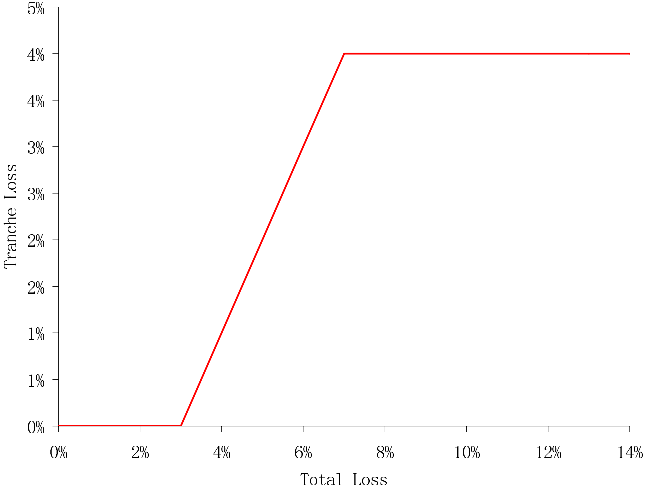
1. the loss for a reference name i by time t is $L_i(t)$
2. the total loss for entire reference pool by time t is $L(t)$
3. the loss for a tranche $[d, u]$ by time t is $L(t; d, u)$

We define the loss function using an indicator: if survival time is less then settlement time then $I\{\tau_i < t\} = 1$.

$$\begin{aligned} L_i(t) &= LGD_i \times EAD_i \times I\{\tau_i < t\} \\ L(t) &= \sum_{i=1}^{N} L_i(t) \\ L(t; u, d) &= \max\{\min\{L(t), u\} - d, 0\} \end{aligned}$$

If the total loss $L(t)$ exceeds the upper boundary (detachment point) of a tranche, then the entire tranche $(u - d)$ is wiped out.

Mezzanine Tranche Payoff



Synthetic CDO Pricing

The present value of protection leg (assuming payment in arrears) is

$$\sum_{j=1}^{M} Z(0, t_j)\left[L(t_j; d, u) - L(t_{j-1}; d, u)\right] \tag{1}$$

so the present value of premium leg is

$$s\,\Delta \sum_{j=1}^{M} Z(0, t_j)\left[(u - d) - L(t_j; d, u)\right] \tag{2}$$

where s is the fair spread (premium paid) for the tranche $[d, u]$.

Synthetic CDO Pricing (Continued)

Similarly to an interest rate swap and credit default swap pricing, present value of both legs must be equal at inception.

This usually applies under the expectation. So, the fair spread s for tranche $[d, u]$ is calculated as

$$s = \frac{\mathbb{E}\left[\sum_{j=1}^{M} Z(0,t_j)\left[L(t_j;d,u) - L(t_{j-1};d,u)\right]\right]}{\Delta\,\mathbb{E}\left[\sum_{j=1}^{M} Z(0,t_j)\left[(u-d) - L(t_j;d,u)\right]\right]} \tag{3}$$

CDO pricing means calculating the fair spread for each tranche.

Loss Distribution for CDO Portfolio

The key input for the pricing equation (3) is the expected loss.

It is natural to see the loss as a random variable that follows the joint distribution of default times for each reference name:

$$L(t) \sim F(t_1, t_2, \ldots, t_n)$$

The loss per each reference name $L_i(t)$ also depends on loss given default, as defined by a recovery rate $LGD \equiv 1 - R$.

It is possible to make reasonable assumptions about recovery rate. Therefore, the focus of the CDO pricing is to figure out the joint distribution of default times of the asset pool.

Marginal and Joint Distributions

A marginal distribution of a random variable X is defined by its **CDF**

$$F(x) = \Pr(X \leq x)$$

The joint distribution function of two random variables X_1 and X_2 is

$$F(x_1, x_2) = \Pr(X_1 \leq x_1, X_2 \leq x_2)$$

We can model the default risk of a credit portfolio if the joint default distribution function is known

$$F(t_1, t_2, \ldots, t_n) = \Pr(\tau_1 \leq t_1, \tau_2 \leq t_2, \ldots, \tau_n \leq t_n)$$

where $\tau_i \leq t_i$ means a default event.

Issues with Joint Distributions

In a bivariate case, one can see the direct link between the joint distribution and joint probability:

$$
\begin{aligned}
F(x_1, x_2) &\equiv \Pr(X_1 \leq x_1, X_2 \leq x_2) \\
&= 1 - \Pr(X_1 > x_1) - \Pr(X_2 > x_2) + \Pr(X_1 > x_1, X_2 > x_2)
\end{aligned}
$$

However, a multivariate joint distribution is a multidimensional analytical construction that

1. requires matrix representation and might not have explicit solution;
2. assumes identical marginal distributions—assuming different marginal distribution for default times leads to a highly problematic inhomogeneous joint distribution; and
3. has measures of dependence appearing in marginal distributions.

Copula Approach

The better way to isolate dependence structure among multiple variables is **a copula**.

Simply by transforming a random variable by its own CDF, we can obtain the uniform copula function as $U = F(X)$.

U is also known as *the grade* of X, and the grades distribution is always uniform $U \sim \mathrm{U}_{[0,1]}$

Copula approach separates the joint distribution into (a) marginal distributions for each variable and (b) their dependence structure. In a way, copula is 'a pure joint distribution'.

Definition of Copula

The joint distribution function of n uniform random variables $(U_1, U_2, \cdots, U_k)$ is **the Copula Function**.

$$C(u_1, u_2, \ldots, u_n; \mathbf{\Sigma}_\rho)$$

where dependence structure is defined by either a correlation matrix $\mathbf{\Sigma}_\rho$ or single parameter α that assumes equicorrelation.

For a standardised distribution $\sigma^2 = 1$, so covariance and correlation matrices are the same $\mathbf{\Sigma}_\rho \equiv \mathbf{\Sigma}$.

The copula function is expressed in terms of uniform variables u_i.

Copula as a Joint Distribution

In order to obtain the expression for the copula function in terms of u_i, we transform a random variable by its own CDF.

For the random variables $X_1, X_2, \ldots, X_n$ with a flexible choice of marginal distributions $F_1(x_1), F_2(x_2), \ldots, F_n(x_n)$

$$\begin{aligned}
C(u_1, u_2, \ldots, u_n) &\equiv \Pr(U_1 \leq u_1, U_2 \leq u_2, \ldots,\ U_n \leq u_n) \\
&= \Pr(F_1(X_1) \leq u_1, F_2(X_2) \leq u_2, \ldots,\ F_n(X_n) \leq u_n) \\
&= \Pr(X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2), \ldots,\ X_n \leq F_n^{-1}(u_n)) \\
&= \Pr(X_1 \leq x_1, X_2 \leq x_2, \ldots,\ X_n \leq x_n) \\
&\equiv F(x_1, x_2, \ldots, x_n) \quad \text{which is a joint distribution.}
\end{aligned}$$

We established the equivalence.

Sklar Theorem

Sklar proved that any joint distribution can be expressed in the form of a copula function, and if the joint distribution is continuous then the copula function is unique.

This is an inverse of the equivalence that we have just proved by moving from copula definition to a joint distribution.

In lay terms, Sklar Theorem shows that we can obtain and rely on the copula function to model dependence structure.

Copula Density Function

A useful application of the Sklar Theorem is derivation of the copula density function from a multivariate PDF $f(x_1, \ldots, x_n)$.

$$
\begin{aligned}
f(x_1, \ldots, x_n) &= \frac{\partial^n \left[C(F_1(x_1), \ldots, F_n(x_n))\right]}{\partial F_1(x_1), \ldots, \partial F_n(x_n)} \times \prod_{i=1}^{n} f_i(x_i) \\
&= c(u_1, u_2, \ldots,\ u_n) \times \prod_{i=1}^{n} f_i(x_i) \quad \Longrightarrow \\
c(u_1, u_2, \ldots,\ u_n) &= \frac{f(x_1, \ldots, x_n)}{\prod_{i=1}^{n} f_i(x_i)}
\end{aligned}
$$

Differentiating over the Copula Function C allowed to obtain the copula density function c (notice the small cap):

$$
\frac{\partial^n \left[C(F_1(x_1), \ldots, F_n(x_n))\right]}{\partial F_1(x_1), \ldots, \partial F_n(x_n)} = c\left(F_1(x_1), F_2(x_2), \ldots,\ F_n(x_n)\right)
$$

Classification of Copulae

From among the several families of copulae, the two are most often used in quantitative finance due to their tractability.

- **Elliptical** (they square a random variable)
  - Gaussian Copula
  - Student's t Copula

  In our CDO pricing discussion, we will focus on Gaussian Copula.

- **Archimedean** (Copula Functions are tractable sums of u_i)
  - Gumbel Copula
  - Clayton Copula
  - Frank Copula

  The benefit of Archimedean copulae is the modelling of n-dimensional dependence with a single parameter α, which bears the assumption of equicorrelation in a portfolio of similar assets.

Multivariate Gaussian Copula

The multivariate Gaussian Copula Function can be expressed as

$$C(u_1, u_2, \ldots, u_n) = \Phi_n \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n); \boldsymbol{\Sigma}\right)$$

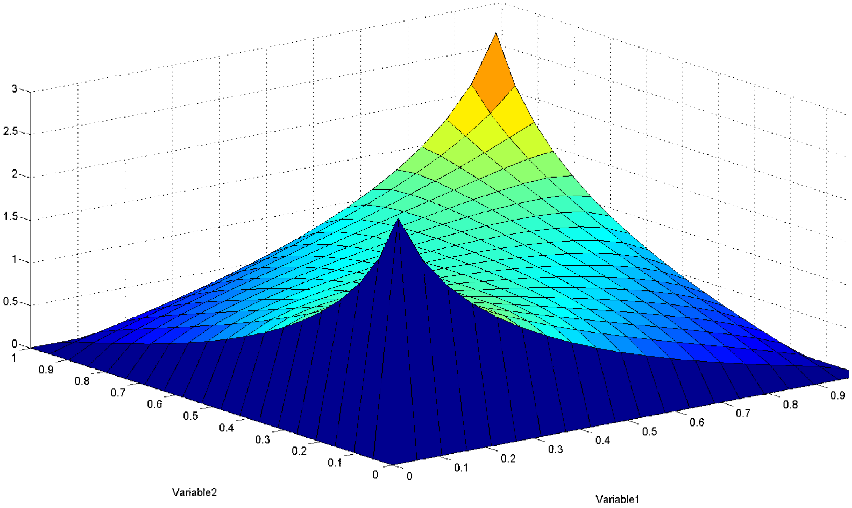
where Φ_n is CDF for the *multivariate* standard Normal distribution. This copula function have no closed-form solution. However, the n-dimensional copula density function can be expressed in matrix form as follows:

$$c\left(\Phi(x_1), \Phi(x_2), \ldots, \Phi(x_n)\right) = \frac{\frac{1}{\sqrt{2\pi^n|\Sigma|}} \exp\left(-\frac{1}{2}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right)}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right)}$$

or in terms of a vector of uniform variables U (using inverse CDF)

$$c(u_1, u_2, \ldots, u_n) = \frac{1}{\sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}\Phi^{-1}(\boldsymbol{U}')(\boldsymbol{\Sigma}^{-1} - \boldsymbol{I})\Phi^{-1}(\boldsymbol{U})\right]$$

Bivariate Gaussian Copula Density, $\rho = 0.5$



Multivariate Student's t Copula

The multivariate Student's t Copula Function can be expressed as

$$C(u_1, u_2, \ldots, u_n) = T_v\left(T_v^{-1}(u_1), T_v^{-1}(u_2), \ldots, T_v^{-1}(u_n); \mathbf{\Sigma}\right)$$

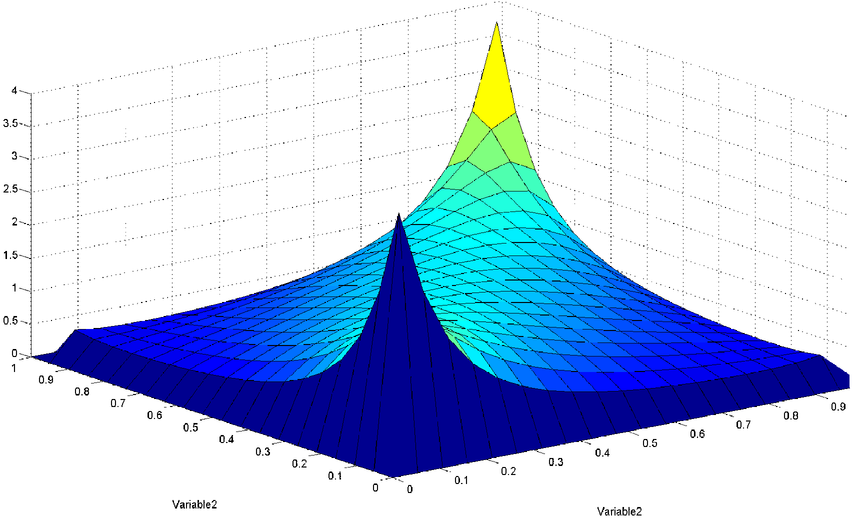
where T_v is CDF for the standard Student's t distribution, multi- or univariate as necessary, now with the degrees of freedom parameter v.

Student's t copula density function is

$$c(u_1, \ldots, u_n) = \frac{1}{\sqrt{|\Sigma|}} \frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \left(\frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}\right)^n \frac{\left(1 + \frac{T_v^{-1}(\boldsymbol{U'})\mathbf{\Sigma}^{-1} T_v^{-1}(\boldsymbol{U})}{v}\right)^{-\frac{v+n}{2}}}{\prod_{i=1}^{n}\left(1 + \frac{T_v^{-1}(u_i)^2}{v}\right)^{-\frac{v+1}{2}}}$$

where $\Gamma(v)$ is a Gamma function.

Bivariate Student's t Copula Density, $\rho = 0.5, d.f. = 3$



Joint Distribution of Default Times

Having introduced Elliptical copulae, let's return to the issue of CDO pricing and the example of *n*-asset portfolio.

Specifically to the joint distribution of default times of the asset pool.

If we transform a default time τ_i by its own CDF $u_i = F_i(\tau_i)$, then we can use standard Normal CDF to transform from uniform to Normal plane

$$x_i = \Phi^{-1}(u_i) = \Phi^{-1}(F_i(\tau_i))$$

This allows modelling dependence among default times with the Gaussian Copula.

$$C\left(F_1(\tau_i), \cdots, F_n(\tau_n)\right) = \Phi_n\left(\Phi^{-1}(F_1(\tau_1)), \ldots, \Phi^{-1}(F_n(\tau_n)); \boldsymbol{\Sigma}_\rho\right)$$

Correlated Random Variables

You can already guess that CDO pricing would rely on Monte-Carlo method and draw some standard Normal random variables Z_i.

We have to impose correlation on Z_i and obtain the Normal variables X_i correlated in the same way as default times τ_i.

The method requires a successfully estimated correlation matrix $\mathbf{\Sigma}_\rho$. Because we operate with standartised variables, let's consider how to impose correlation using decomposition of covariance matrix $\mathbf{\Sigma}$.

Selected Notation

Let $\mathbf{Z}$ be an d-dimensional vector of **independent** standard Normal variables.

$$\mathbf{Z}^T = (z_1\, z_2\, \cdots\, z_d)$$

The most computationally effective method to transform $\mathbf{Z}$ is by a linear combination of its existing values.

To impose unique weights on the values in $\mathbf{Z}$, let's define an $n \times d$ matrix, where each row represents a vector of weights.

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nd} \end{pmatrix}$$

Correlated Normal Vector

We generate an n-dimensional vector of **correlated** standard Normal variables.

$$\mathbf{X}^T = (x_1 \, x_2 \, \cdots \, x_n)$$

by pre-multiplying by the matrix $\mathbf{A}$

$$\mathbf{X} = \mathbf{A}\,\mathbf{Z} \quad (n \times d \times d \times 1)$$

where, the flexible choice of dimensions is reconciled with

$$x_i = \sum_{j=1}^{d} a_{ij} z_j$$

We have to figure out how to obtain the special matrix $\mathbf{A}$, for which we have to look at the properties of the covariance matrix of $\mathbf{X}$.

Covariance Matrix

Introducing expectation since we work with random variables, the (auto)covariance matrix of $\mathbf{X}$ is calculated as

$$\mathbf{\Sigma} = \mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right] = \mathbb{E}\begin{bmatrix} x_1^2 & x_1 x_2 & \dots & x_1 x_n \\ \vdots & \dots & \dots & \vdots \\ x_n x_1 & x_n x_2 & \cdots & x_n^2 \end{bmatrix}$$

Given that $\mathbf{X} = \mathbf{A}\,\mathbf{Z}$ the covariance matrix can be represented as

$$\mathbf{\Sigma} = \mathbb{E}\left[\mathbf{A}\mathbf{Z}\mathbf{Z}^T\mathbf{A}^T\right] = \mathbf{A}\mathbf{A}^T$$

The result shows that we can impose correlation on **any** arbitrary chosen vector $\mathbf{Z}$ while retaining the covariance structure.

Matrix Factorization

We obtain matrix $\mathbf{A}$ by Factorization or Decomposition of covariance matrix $\mathbf{\Sigma}$.

Matrix factorisation is similar to root-finding: for any numbers $x = a \times b$, a, b are *factors* of the number x. Positive numbers have a real square root, a ready candidate for a factor $x = \sqrt{x} \times \sqrt{x}$.

Similar result holds for matrices. Any symmetric positive definite or semi-definite matrix $\mathbf{\Sigma}$ can be factorized into a product of matrices in several ways.

$$\boldsymbol{Z}^{\boldsymbol{T}} \mathbf{\Sigma} \boldsymbol{Z} \geq 0$$

This technical condition for a matrix to be positive definite for any arbitrary real or complex vector $\boldsymbol{Z}$ guarantees a unique Cholesky decomposition and positive eigenvalues.

Decomposition of Covariance or Correlation Matrix

We are going to review two popular methods of matrix decomposition:

- Cholesky Decomposition

$$\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^{\boldsymbol{T}}$$

where $\boldsymbol{L}$ is the lower triangular matrix.

- Spectral Decomposition (PCA)

$$\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{T} = \mathbf{V}\boldsymbol{\Lambda}^{\frac{1}{2}}\left(\mathbf{V}\boldsymbol{\Lambda}^{\frac{1}{2}}\right)^{T}$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues and $\boldsymbol{V}$ is the vectorised matrix of eigenvectors (in columns).

Cholesky Decomposition

For any symmetric positive definite matrix, Cholesky Decomposition provides a unique solution for the lower triangular matrix.

The method is analytically tractable, and we can see how it works with an example and general algorithm.

Two-Dimensional Example

Let's introduce a two-variate covariance matrix $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

Cholesky Decomposition takes the form the product of two triangular matrices $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^{\boldsymbol{T}} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{21} \\ 0 & a_{22} \end{pmatrix}$$

with real and positive diagonal elements $a_{11} = \sigma_1^2$ and $a_{22} = \sigma_2^2$.

Two-Dimensional Example (Continued)

Completing multiplication $\boldsymbol{AA^T}$ we obtain

$$\begin{pmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{21}a_{11} & a_{21}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

Given that $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$, we have 3 equations for 3 unknowns

$$\begin{cases} a_{11}^2 = \sigma_1^2 \\ a_{21}a_{11} = \rho\sigma_1\sigma_2 \\ a_{21}^2 + a_{22}^2 = \sigma_2^2 \end{cases}$$

Solving for a_{ij} in sequence, we obtain the solution for $\boldsymbol{A}$ as

$$\begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix}$$

Two-Dimensional Example (Application)

Remember that the purpose of obtaining $\boldsymbol{A}$ was to impose correlation on an arbitrarily drawn vector of independent standard Normal variables $\boldsymbol{Z}$ by

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Z}$$

Continuing our example

$$\boldsymbol{X} = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

so that

$$\begin{aligned} x_1 &= \sigma_1 z_1 \\ x_2 &= \rho\sigma_2 z_1 + \sqrt{1-\rho^2}\sigma_2 z_2 \end{aligned}$$

x_i are converted to default times using their own CDF as $\tau_i = F^{-1}(\Phi(x_i))$.

General Algorithm for Cholesky Decomposition

For the d-dimensional symmetric matrix $\boldsymbol{\Sigma}$, we need to solve the system of equations stemming from the product $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^{\boldsymbol{T}} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$

$$\boldsymbol{\Sigma} = \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{d1} \\ & a_{22} & \cdots & a_{d2} \\ & & \ddots & \vdots \\ & & & a_{dd} \end{pmatrix} =$$

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \vdots & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \cdots & \cdots & \sigma_{dd} \end{pmatrix}$$

where, again, all diagonal elements are real and strictly positive.

General Algorithm (Continued)

Looping over the σ_{ij} by rows i and then columns j gives (starting with the first row of $\boldsymbol{A}$ by the first column of $\boldsymbol{A^T}$)

$$\begin{aligned} a_{11}^2 &= \sigma_{11} \\ a_{11}a_{21} &= \sigma_{12} \\ &\vdots \\ a_{11}a_{d1} &= \sigma_{1d} \\ a_{21}^2 + a_{22}^2 &= \sigma_{22} \\ &\vdots \\ a_{21}a_{d1} + a_{22}a_{d2} &= \sigma_{2d} \end{aligned}$$

Exactly one new entry of the matrix $\boldsymbol{A}$ appears in each equation, making it possible to solve the system in sequence.

General Algorithm (Completed)

We can express the generalised Cholesky Decomposition in the form

$$\sigma_{ij} = \sum_{k=1}^{i} a_{ik} a_{jk} \quad \forall\, j \geq i \tag{4}$$

Separate the last term as $\sum_{k=1}^{i-1} a_{ik} a_{jk} + a_{ii} a_{ji}$ and rearranging gives solution for a_{ji}

$$a_{ji} = \frac{1}{a_{ii}} \left(\sigma_{ij} - \sum_{k=1}^{i-1} a_{ik} a_{jk} \right) \quad \forall\, j \geq i$$

Solution for a_{ii} also comes from (4) by setting $j = i$ and rearranging as

$$a_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} a_{ik}^2} \quad \forall\, j = i$$

Spectral Decomposition

The properties of spectral decomposition of a symmetric matrix into its eigenvalues and eigenvectors also allow obtaining matrix $\boldsymbol{A}$

$$\mathbf{\Sigma} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}}\left(\mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}}\right)^T$$

where $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues and $\boldsymbol{V}$ is the vectorised matrix of eigenvectors (in columns). Eigenvectors are orthogonal to each other and represent the directions of zero covariance (among themselves).

We can simply notice by functional similarity that

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}}$$

Cholesky vs. Spectral: Pros and Cons

It is clear how the Cholesky Decomposition wins a computational advantage: it is a sequential solving through the organised system of equations where only one new unknown comes up at a time.

The restriction is that the Cholesky Decomposition does not work when matrix $\boldsymbol{\Sigma}$ is only positive semi-definite.

In that case, we carry out the Spectral Decomposition, which is also occasionally useful in terms of Principal Component Analysis in order to reveal the internal structure of correlation or covariance matrix. Imposing correlation by using only several principal components is also a possibility.

Simulation by Copula Procedure

To simulate random default times by means of Gaussian copula with correlation $\boldsymbol{\Sigma}$ we carry out the following Monte Carlo-like procedure:

1. Find a suitable decomposition method to obtain $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$.
2. Draw an n-dimensional vector of independent standard Normal variables $\mathbf{Z} = (z_1, \cdots, z_n)'$.
3. Compute a vector of correlated variables by $\mathbf{X} = \mathbf{AZ}$.
4. Convert the normal variables to default times using marginal CDF with individual parameters for each reference name.

$$\tau_i = F_i^{-1}\left(\Phi(x_i)\right) \tag{5}$$

The procedure represents only **one run** of Monte Carlo.
τ_i is utilised in the fair spread s calculation.

Notes on Simulation by Copula

Copula represents the joint distribution of default times

$$F(t_1, t_2, \ldots, t_n) = \Pr(\tau_1 \le t_1, \tau_2 \le t_2, \ldots, \tau_n \le t_n)$$

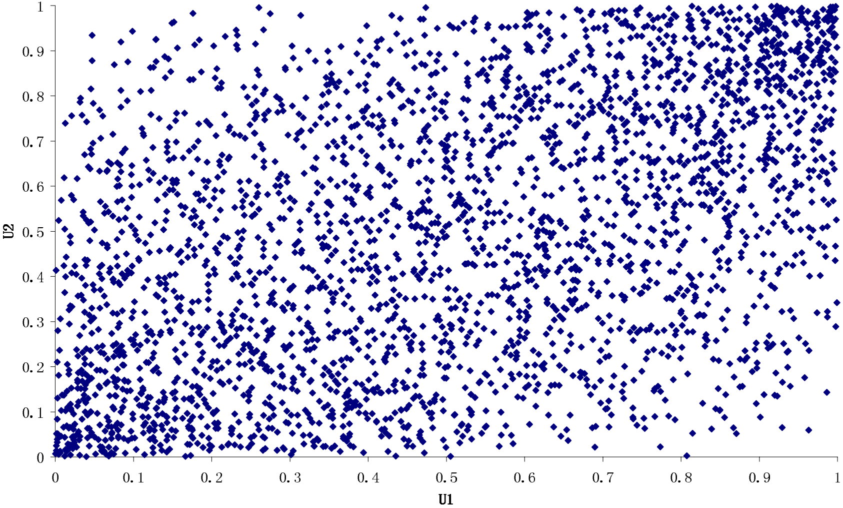
that, in turn, defines the loss distribution for credit portfolio (a CDO).

Implementation of Step 4 is more nuanced than the theory suggests: often, we have no closed-form solution for marginal distributions $F_i(\tau_i)$.

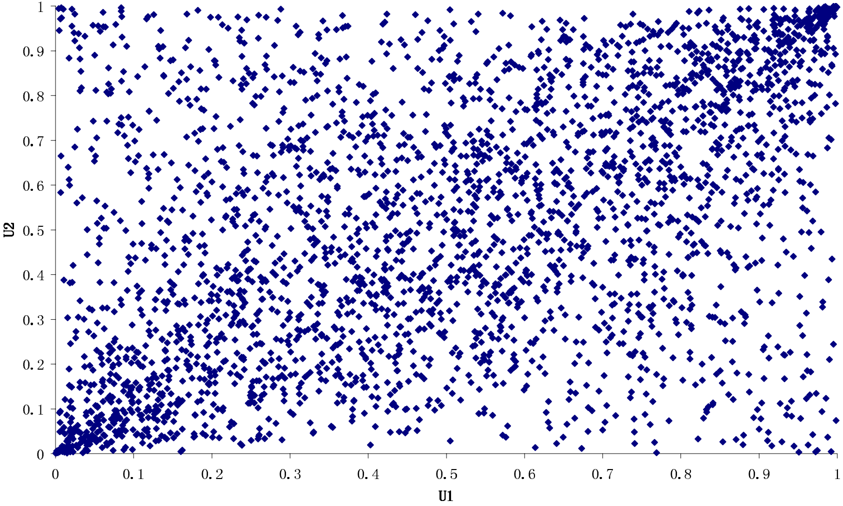
τ_i is modelled as the inter-arrival time of the Poisson Process, leading to the Exponential distribution's **CDF** $F(\tau) = 1 - e^{-\lambda\tau}$ and calibrated intensities, which are bootstrapped from the single-name CDS data.

$$\text{Equation (5) gives} \quad \tau \overset{D}{=} F^{-1}(u) = -\frac{\log(1-u)}{\lambda} \tag{6}$$

Simulated Bivariate Gaussian Copula, $\rho = 0.5$



Simulated Bivariate Student's t Copula, $\rho = 0.5, d.f. = 2$



Factor Represenation for Copula

We have considered the explicit Gaussian copula model and associated simulation approach for the vector of joint default times.

The approach definitionally relied on Copula Function. The Archimedean family have tractable copula functions, so there exists a preference in their use in research. But working with copula functions of the Elliptical family is as complicated as working with joint distributions.

<u>There exits an alternative</u>, a factor representation for the copula model with *a linear equation*. The alternative is wildly popular in practice:

- Factor representation does the same job as the copula model.
- A factorisation is typically easier to understand and operate with.
- It requires significantly less computational effort (also applies to coding mistakes and wrong parameters being passed).

Asset Value Approach

One of the first approaches to quantify the loss distribution was the Asset Value Approach, developed in 1990s in CreditMetrics, a portfolio model for evaluating credit risk. The model developed two assumptions:

- The volatility of asset value should predict the chance of default.
- The joint [default] probabilities for two firms can be derived from a knowledge of the correlation between the two firm's asset values.

Default event was modelled as a breach of threshold in **a latent variable**, a proxy to the firm's asset value. You can recognise the similarity of the approach to Structural Models of credit risk.

Latent Variable

Factor presentation proposes that for each obligor i in a credit portfolio (CDO), there exists a latent variable A_i and associated threshold d_i such

$$\begin{aligned} A_i \leq d_i &\quad\equiv\quad \text{Obligor i defaults} \\ A_i > d_i &\quad\equiv\quad \text{Obligor i does not default} \end{aligned}$$

The factorisation is formulated in linear form using a common factor Z and idiosyncratic factor (shock) ε_i (both standard Normal), linked by a factor loading w_i, a parameter that represents correlation.

$$A_i = w_i Z + \sqrt{1 - w_i^2}\varepsilon_i \qquad A_i \sim N(0,1) \tag{7}$$

The probability of default for every name can be conditioned on the common factor $F(\tau_i < t|Z)$, leading to (similar) independent default probabilities per name, instead of a joint distribution.

Default Correlation

One can recognise the factor representation as a ready Cholesky decomposition. The nuance is that instead of correlating default events, the Asset Value Approach correlates latent variables.

$$\rho_{ij} = \mathbb{C}\text{ov}[A_i, A_j] = w_i\, w_j$$

We can show that with expectations algebra

$$\begin{aligned}
\mathbb{C}\text{ov}[A_i, A_j] &= \mathbb{C}\text{ov}[w_i Z + \sqrt{1-w_i^2}\varepsilon_i,\; w_j Z + \sqrt{1-w_j^2}\varepsilon_j] \\
&= \mathbb{C}\text{ov}[w_i Z,\; w_j Z] = w_i w_j \times \mathbb{V}\text{ar}[Z] \\
&= w_i w_j \times 1
\end{aligned}$$

The conditions on the idiosyncratic factor ε_i are such that

$$\mathbb{C}\text{ov}[\varepsilon_i, \varepsilon_j] = 0, i \neq j \quad \mathbb{C}\text{ov}[Z, \varepsilon_i] = 0, \forall i$$

Simulation by Factorised Copula Procedure

To simulate the loss distribution under the Asset Value Approach, again a Monte Carlo-like procedure is easily implementable:

1. Draw a random value for the latent variable A_i for each obligor by (7)

$$A_i = w_i Z + \sqrt{1 - w_i^2}\varepsilon_i$$

   for that, one needs inputs of w_i, Z and random input for $\varepsilon_i \sim \mathrm{IN}(0, 1)$
2. For each obligor, check if it defaulted, $A_i \leq d_i$.
   In case of default, calculate loss $L_i(t) = LGD_i \times EAD_i$.
3. Aggregate individual losses into portfolio loss $L(t)$.
4. Steps 1-3 represent **one** run of Monte Carlo. Repeat the steps until there is enough data to generate a portfolio *loss distribution*.

Calibration in Factor Representation

To parameterise the factorised copula model, we need values for d_i and w_i.

We can set the threshold d_i to match some empirically estimated probability of default (i.e., bootstrapped from CDS):

$$P_i = \Pr[A_i \le d_i] = \Phi(d_i)$$

$$d_i = \Phi^{-1}(P_i) \quad \Longrightarrow \quad d_i = \Phi^{-1}(1 - \mathrm{PrSurv}_i)$$

To determine the factor loading w, we can use the advanced proxies of the true correlation among the firms' asset values. For a synthetic CDO, we can correlate quoted credit spreads. This is calibration.

A question of joint default returns us to copula, now, for latent variables:

$$P_{ij}(t) = \Pr(\tau_i \le t, \tau_j \le t) = \Pr(A_i \le d_i, A_j \le d_j) = \Phi_2(d_i, d_j; \rho_{ij})$$

CDO Pricing Sample Results

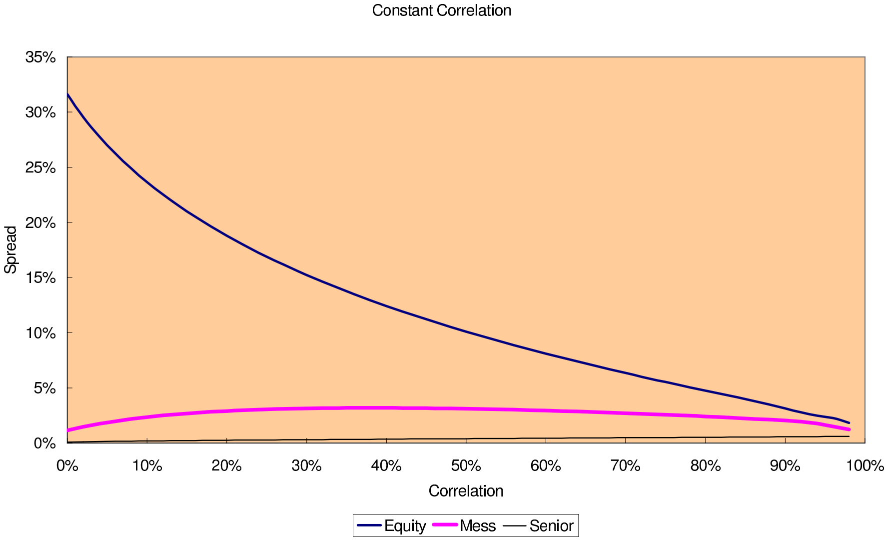
Suppose there is a homogeneous loans portfolio of 125 names with the equal notional principal per name. It is estimated that each name has a probability of default (PD) of 1% per annum and $LGD = 1 - R$ of 60%.

The maturity of this CDO is 5 years with premium payable quarterly in arrears. The interest rate is 5% per annum. Given 30% homogenous correlation among all names, the spread for each tranche is priced as

Tranche	Attachment Point	Spread
Equity	0%-3%	15.30%
Mezzanine	3%-10%	3.15%
Senior	10%-100%	0.07%

Ensuing illustration shows how Equity holders can get wiped out while receiving inadequately low premiums, given a high level default correlation.

Tranche Spread vs. Correlation



Please take away the following ideas...

- Liquid CDO markets trade in standardized CDS indicies and synthetic credit portfolios.
- Modelling default risk of a credit-exposed portfolio relies on the joint distribution of default times.
- Copula is a powerful method to isolate dependence structure, 'a pure joint distribution', while preserving the choice of marginal distributions.
- One-factor representation for Gaussian copula is intuitive and popular approach to price credit derivatives, such as a synthetic CDO or kth-to-default basket of CDS.