Copula and CDO Implementation

Siyi Zhou

Lecture 5.6

CQF June 2008

In this lecture

1. Briefly review the definition and major property of CDO
2. Introduce synthetic CDO and its general pricing framework
3. Revise basics of copula and how it can be used to model dependence among multi-dimension variable.
4. One factor normal copula model to price synthetic CDO

By the end of this lecture you will be able to

- understand what CDO is and how it works
- see how copula model is used to price credit derivatives whose underlying is a portfolio.
- Implement one factor normal copula to price synthetic CDO in spreadsheet.

What is CDO?

CDO is a type of asset backed security, it is an investment on a pool of assets. Instead of investing directly on the asset pool like many other investments, the issuer of CDO , SPV(Special Purpose Vehicle) repack the asset pool and slices it into tranches according to it's credit risk. The tranches of a typical CDO are

- Senior tranche
- Mezzanine tranche
- Equity tranche

How CDO works

- at good times, tranche Investors will receive premium (returns for asset pool) in turn periodically.

- at bad times, loss will be applied in reverse order of seniority

- The senior tranche is protected by the subordinated tranches

Synthetic CDO

Synthetic CDO does not hold cash asset, instead, it holds a portfolio of CDS, so it is called synthetic.

It also can be unfunded, which means investors only pays when default affecting their tranches. In this case, the counterparty default risk must be taken into account by risk managers.

The pricing of synthetic CDO is simpler than cash CDO since we don't need to worry about cash modelling, which conveniently helps us focus on modelling the default risk of the asset pool.

Notation

Let's get familiar with some notation first.

- survival time τ_i
- loss given default LGD_i
- exposure at default EAD_i
- tranche $[D, U]$
- settlement date t_j
- tenor $\Delta = t_{j+1} - t_j$
- discount factor $P(t, T)$

Loss function

1. the loss for obligor i by time t is $L_i(t)$,

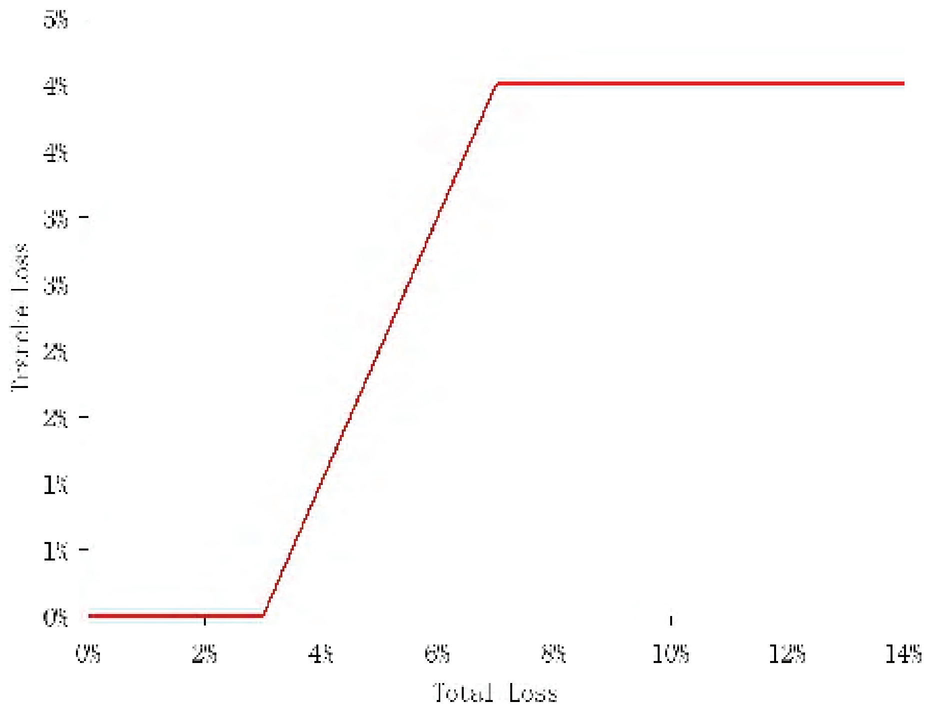
2. the total loss by time t is $L(t)$,

3. the loss for tranche $[d, u]$ by time t is $L(t; u, d)$.

So

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

Mezzanine Tranche Payoff Diagram



Synthetic CDO Pricing

The present value of protection leg is

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

Assuming the fair tranche spread is s, so the present value of premium leg is

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

Synthetic CDO Pricing

Same principle as pricing swaps, present value of both legs must be equal.

Now take expectation of present value of both legs and equate them the fair spread s for tranche (d, u) is determined.

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

Portfolio loss distribution

From equation 3, one can see that the key input to price a CDO tranche is the loss distribution of the reference porfolio. To derive the loss distribution, need know two important things:

1. Distribution of joint default

2. Specification of default for each obligor, i.e. EAD and LGD

Suppose that it is not too difficult to make assumptions on loss give default and exposure at default for each obligor, so the main task to is to figure out what the joint default distribution of the asset pool is.

Marginal and joint distribution

By definition marginal distribution of a random variable X is

$$F(x) = \text{Pr}(X \leq x),$$

The joint distribution function of two random variables X and Y is

$$F(x, y) = \text{Pr}(X \leq x, Y \leq y).$$

We can model default risk of a credit portfolio if we know joint default distribution function

$$F(t_1, t_2, \cdots, t_n) = \text{Pr}(\tau_1 \leq t_1, \tau_2 \leq t_2, \cdots, \tau_n \leq t_n).$$

Problem with joint distribution

Directly work on joint distribution is inconvenient, because

- Marginal distribution are different, conventional joint distribution only accepts homogenous marginal distribution.

- Extension to higher dimension maybe difficult.

- Measures of dependence may appear in marginal distribution.

Copula Approach

Instead of directly working on joint distribution which must be horrendous, the better way is to use copula function. Unlike joint distribution function, copula function can separate marginal distribution and their association completely, as a result, by copula, one can conveniently mix marginal distributions together with certain dependence structure to become a joint distribution.

Let's see how one can do that.

Definition of Copula

In words, a copula is a function that mixes univariate marginal distributions to become their full multivariate distribution.

Definition [Copula function]: For k uniform random variables $(U_1, U_2, \cdots, U_k)$, the joint distribution function

$$C(u_1, u_2, \cdots, u_k; \rho)$$

is called a copula function.

Copula links to joint distribution

Since the distribution function of a random variable is uniformed distributed, so copula function can be used to link marginal distribution with a joint distribution.

Suppose there are random variables $X_1, X_2, \cdots, X_k$, with distributions $F_1, F_2, \cdots, F_k$, then

$$C(F_1(x_1), F_2(x_2), \cdots, F_k(x_k)) = F(x_1, x_2, \cdots, x_n)$$

.

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

Multivariate normal Copula

Definition [Normal copula] Let Φ_k be the k-variate normal distribution function and Φ be univariate normal distribution function, the k-variate normal copula function is

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

Multivariate t Copula

Definition [Student's t Copula]: Let T_v be the k-variate Student's t distribution function with v degrees of freedom, then the multivariate Student's t Copula is

$$C(u_1, u_2, \cdots, u_k) = T_v\left(T_v^{-1}(u_1), T_v^{-1}(u_2), \cdots, T_v^{-1}(u_k); \Sigma\right)$$

Normal copula to simulate joint default time

In our example of n assets portfolio, we assume, for each obligor i, the marginal distribution is $F_i(\tau_i)$, and somehow the correlation is successfully estimated, so the normal copula for the joint survival time is

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

in which we define

$$x_i = \Phi^{-1}(F_1(\tau_i)).$$

We need to find out what τ_i is, $\forall i = 1, 2, \cdots, n$ through normal copula function. The first step is to generate correlated random variables x_is.

How to generate correlated multivariate normal distribution?

Short answer is, instead of generating correlated multivariate normal variables directly, one can generate independent normal variables and then convert them into correlated ones according to predetermined correlation matrix.

Let's see how to do it mathematically.

Set up Notation

Let's denote

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

be an independent d-dimensional normal vector.

The most effective method to create correlated normal vector is just by linearly combining independent normal vector. So introduce a $n \times d$ matrix

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

where each row represents a vector of weights allocated to elements in $\mathbf{Z}$.

Correlated Normal

Define a new vector

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

such that

$$\mathbf{X} = \mathbf{A}\,\mathbf{Z},$$

where

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

Vector $\mathbf{X}$ is then a linear combination of independent normal vector $\mathbf{Z}$. Note the dimension is changed from d to n.

Covariance Matrix

Define the covariance matrix of $\mathbf{X}$ be $\mathbf{\Sigma}$, then by definition it is,

$$\mathbf{\Sigma} = E\left[\begin{pmatrix} x_1^2 & x_1x_2 & \cdots & x_1x_n \\ \vdots & \cdots & \cdots & \vdots \\ x_nx_1 & x_nx_2 & \cdots & x_n^2 \end{pmatrix}\right] = E\left[\mathbf{X}\mathbf{X}^T\right]$$

Plug $\mathbf{X} = \mathbf{A}\,\mathbf{Z}$ into above equation come up with

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

So given Covariance matrix, if one can find $\mathbf{A}$, then by multiply independent vector $\mathbf{Z}$ by $\mathbf{A}$ end up with correlated vector $\mathbf{X}$.

Matrix Factorization

The act of Finding matrix $\mathbf{A}$ is called Matrix Factorization or Decomposition.

We all know that non-negative numbers have real square root, whereas negative number doesn't.

Similar result holds for matrices. Any symmetric at least semi-positive definite matrix, like $\mathbf{\Sigma}$ can be factorized. But the solution is not unique.

Methods to decompose covariance matrix

We are going to introduce two popular methods of decomposing correlation matrix,

- Cholesky Factorization

- Spectral Decomposition

Cholesky Factorization

The basic ideal of Cholesky Factorization is very easy, it claims that any symmetric positive definite matrix can be factorized in the form of triangular matrices.

2 Dimension Example

The best way to see this via looking example. Let's suppose that Σ is a two-dimensional matrix:

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

Cholesky Factorization takes the form:

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

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

It will end up with 3 equations for 3 unknowns like this:

$$\begin{cases} A_{11}^2 = \sigma_{11} \\ A_{21}A_{11} = \sigma_{12} \\ A_{21}^2 + A_{22}^2 = \sigma_{22} \end{cases}$$

One can solve for A_{ij} sequentially, the answer is.

$$\begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{22}\rho & \sigma_{22}\sqrt{1-\rho^2} \end{pmatrix}$$

d Dimension Algorithm

For the case of a d-dimension covariance matrix $\mathbf{\Sigma}$, we need to solve

$$\begin{pmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \vdots & \vdots & \ddots & \\ A_{d1} & A_{d2} & \cdots & A_{dd} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{d1} \\ & A_{22} & \cdots & A_{d2} \\ & & \ddots & \vdots \\ & & & A_{dd} \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ & \sigma_{22} & \cdots & \sigma_{2d} \\ & & \ddots & \vdots \\ & & & \sigma_{dd} \end{pmatrix}$$

Traversing the σ_{ij} by looping over i and then j produces,

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

Exactly one new entry of the A matrix appears in each equation, making it possible to solve for the individual entries sequentially.

More compactly, from the basic identity,

$$\sigma_{ij} = \sum_{k=1}^{i} A_{ik}A_{jk} \quad j \geq i,$$

We get have basic identity

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

and

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

This formulae make a simply recursion to find Cholesky factor.

Spectral Decomposition

spectral decomposition mainly relies on the fact that eigenvector of a symmetric matrix is orthogonal to each other.

With basic linear algebra, the spectral decomposition of of a symmetric matrix takes form

$$\mathbf{\Sigma} = \mathbf{V\Lambda V}^T$$

$\mathbf{V}$ is a matrix which collects eigenvectors in its column, and $\mathbf{\Lambda}$ is a diagonal matrix with its diagonal elements are eigenvalues of $\mathbf{\Sigma}$.

If $\mathbf{\Sigma}$ is positive semi-definite, it can be expressed as

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

Thus

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

Pros and cons

Cholesky Factorization has particular structure providing a computational advantage. Spectral Decomposition doesn't have it and hence isn't faster than Cholesky.

In addition to Cholesky's inability to deal with semi-definite matrix, Spectral Decomposition do however have a statistical interpretation that is occasionally useful, that is related to Principal Component Analysis.

Simulation procedure by using normal copula

The procedure for generating random default times from normal copula with correlation maxtrix Σ proceeds as follows

1. Find a suitable (e.g. Cholesky) decomposition A from Σ, such that $\Sigma = AA'$.

2. Draw a N-dimensional independent standard normal vector $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_n)'$.

3. Let $X = AZ$ to obtain correlated normal vector.

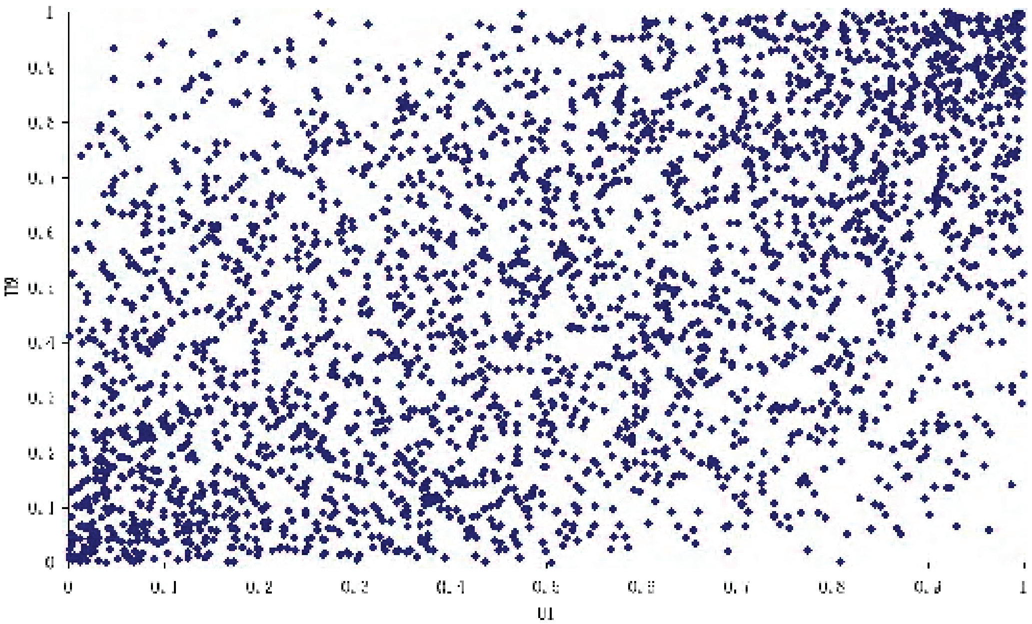
4. Calculate default time use the following equation

$$\tau_i = F^{-1}(\Phi(x_i)) \tag{4}$$

5. repeat 2 to 4 many times

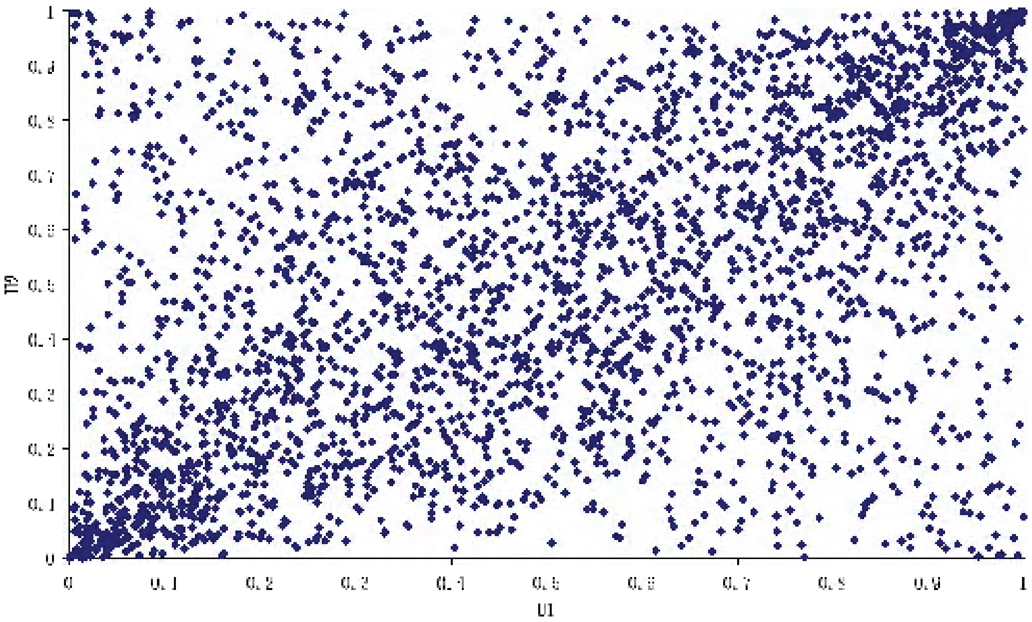
Normal copula plot

$\rho = 0.5$



Student t copula plot

$\rho = 0.5$



Factor copula model

For the time being we have seen copula models which are explicitly based on copula function. Actually there exits another type of copula model which has factor representation, these models have its advantages and are widely used by financial industry.

Factor model does the same job like explicit copula model to obtain joint distribution by mixing marginal distributions with measure of dependence. However, it is more easy to understand and less computation effort.

Asset Value Approach- one factor normal copula

One way to obtain loss distribution is asset value approach which was developed by CreditMetrics in 1990s.

The heart of this method is the assumption that, for each obligor, there exists an latent variable which determines the occurrence of default event.

Latent variable

For each obligor i, there exists a latent variable A_i and its associated threshold d_i such that

$$\begin{aligned} \text{Obligor i default} &\iff A_i \leq d_i \\ \text{Obligor i not default} &\iff A_i > d_i \end{aligned} \tag{5}$$

Where

$$A_i = w_i Z + \sqrt{1 - w_i^2} \varepsilon_i \tag{6}$$

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0, i \neq j; \quad \text{cov}(Z, \varepsilon_i) = 0, \forall i$$

Where Z and ε_i are standard normal variables. So by construction A_i is also standard normal. w_i is called factor loading or sensitivity which ultimately links to correlation.

Default correlation

Instead of directly imposing the structure on default correlations themselves, the asset value approach represents it by imposing structure on latent variables.

$$\rho_{ij} = \mathsf{cov}(A_i, A_j) = w_i\, w_j \tag{7}$$

Above can be easily shown by following:

$$\begin{aligned}\mathsf{cov}(A_i, A_j) &= \mathsf{cov}(w_i Z + \sqrt{1 - w_i^2}\varepsilon_i,\, w_j Z + \sqrt{1 - w_j^2}\varepsilon_j) \\ &= \mathsf{cov}(w_i Z,\, w_j Z) = w_i w_j \mathsf{var}(Z) \\ &= w_i w_j\end{aligned}$$

Implementation:Monte Carlo simulation

One can easily implement asset value approach to obtain loss distribution by Monte Carlo simulation

1. Using (6), randomly draw latent variable for each obligor

2. for each obligor check if it defaulted according to (5). If yes, determine individual loss

3. aggregate individual losses into portfolio loss

4. repeat steps 1 to 3 many times to arrive at portfolio loss distribution

Parametrization

Before implementation,we should know how to parameterize this model, i.e. choosing the d's and w's. We can set threshold d such that it results in the default probability that we have estimated. To do that we use the following

$$P_i = \text{Pr[default]} = \text{Pr}[A_i \leq d_i] = \Phi(d_i) \tag{8}$$

Then

$$d_i = \Phi^{-1}(P_i) \tag{9}$$

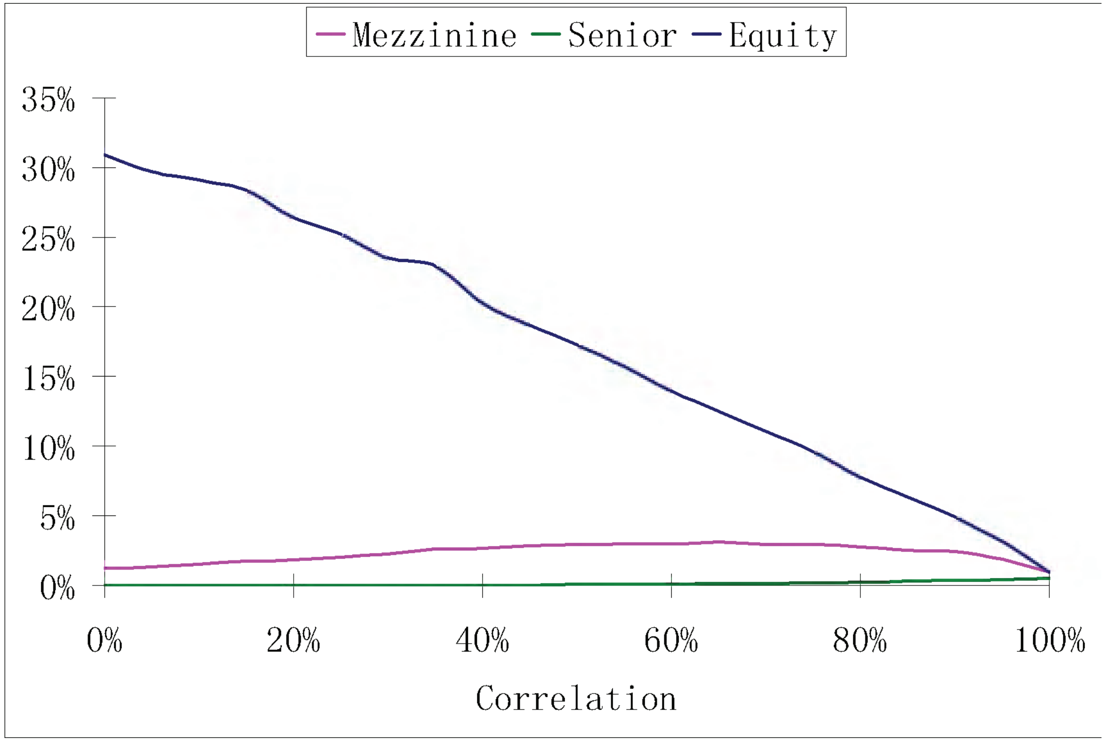
To determine the factor loadings, we could go to the root of this method and estimate correlations of asset values.

Sample results

Suppose there is a homogenous loan portfolio consists of 100 names with 100 notional principal each. We further assume each name has default rate 1% a year and LGD is 60%. The CDO based on the reference loan portfolio is 5 years with premium payable quarterly in arrear, and interest rate is 5%.

Tranche	Attachment Point	PD	EL	Spread
Equity	0%-3%	95.18%	70.76%	23.99%
Mezzanine	3%-10%	36.58%	11.38%	2.29%
Senior	10%-100%	1.42%	0.03%	0.0064%

Tranche spread against Correlation



Summary

1. Joint default distribution is a key input for modeling default risk of a portfolio.

2. Copula is a powerful method to model joint default.

3. One factor normal copula model is an intuitively appealing and easy to implement to price credit derivative such as nth-to-default CDS and synthetic CDO.