Credit Spread Pricing with Correlation Stress-Test

by Dr Richard Diamond

CQF JAN 2022

1. Introduction to CQF Final Project

2. Pricing a credit product: kth-to-default Basket CDS

Preparation

- Start downloading data and planning own implementation of the project.
- It is expected you apply own understanding. It is up to you to source and generate suitable data and features.

  Equities Data tutorial (*yfinance*). If you can't get hold: **make reasonable assumptions**, even generate the data.

  For example: you can make up reasonable CDS spreads and use correlation from equities.

  Refer to the relevant CQF Lectures and do extra reading on pricing methodologies and numerical techniques.

Code Adoption

A. You can adopt code for specific tasks, not model as a whole. Amend code it for your purpose, not copy/paste.

B. Where numerical techniques intense (and not part of the central model): use ready libraries with expertise. For example, portfolio optimisation best done using quadratic optimisation routine, not stochastic gradient Solver.

C. You are welcome to implement complex numerical methods vs. use of ready solution – if able to.

Numerical Techniques

Implement as necessary, numerical techniques
from the first principles.

What to code: pricing formulae, Black-Litterman calculation, SDE simulation, matrix form regression, Engle-Granger, interpolation, numerical integration, Cholesky, t-copula formula, CDS bootstrap, features computation...

Use ready solutions for: covariance shrinkage, nearest correlation, ML numerical methods (eg, decision trees, neural nets), low latency RNs, kernel density (cdf estimation), QR-decomposition (PCA), EGARCH estimation, Johansen Procedure...

Project Report

- A full **mathematical description** of the models employed as well as numerical methods. Remember *accuracy and convergence*!
- Results presented using **a plenty of tables and figures**, which must be interpreted not just thrown at the reader.
- **Pros and cons** of a model and its implementation, together with possible improvements.
- **Demonstrate 'the specials'** of your implementation: own research, own coding of complex methods, use of the industrial-strength libraries of C++, Python.
- Instructions on how to use software if not obvious.
  The code must be thoroughly tested and well-documented.

CQF Electives

See Project Brief for the current relevant table.

Counterparty Credit Risk – choose for CR topic	CDS, survival probabilities and hazard rates reviewed. Three key numerical methods for quant finance pricing (Monte-Carlo, Binomial Trees, Finite Difference). Monte Carlo for simple LMM. Review of Module Five on Credit with a touch on the copula method. **Outcome:** covers CVA Computation clearly and reviews of credit spread pricing techniques.
Risk Budgeting – choose for PC topic	Reviews the nuance of Modern Portfolio Theory, ties in VaR and Risk Decomposition with through derivations and expectation algebra. Gives simple examples of figures you need to compute and then combine with portfolio optimisation. Risk-budgeting portfolio from Video Part 10.
Adv Risk Management – useful in general	The base you need to be a risk manager (read Coleman guide) and tools for Basel regulation: (a) weakness of risk-weighted assets, (b) extreme value theory for ES and capital requirement and (c) adjoint automatic differentiation to compute formal sensitivities. Other numericals covered are the same as Counterparty Risk. **Outcome:** this elective is best taken for your own advancement.

Final Day as advised

Don't Extend Your Luck!

1. Introduction to CQF Final Project

2. Pricing a credit product: kth-to-default Basket CDS

Introduction and Case

Case Study

In 2008 AIG Financial Products Corp had notional CDS exposure to highly rated CDSs of roughly **$450-500 billion** concentrated on banks, with about $60 billion exposed to subprime mortgages.

Viewing CDS as a leveraged purchase of bond, consider the implications.

- Outright purchase of $450 billion worth of corporate bonds (especially subordinated bank debt) would have attracted everyone's attention: C-level management, counterparties, regulators. **CDS positions did not**.

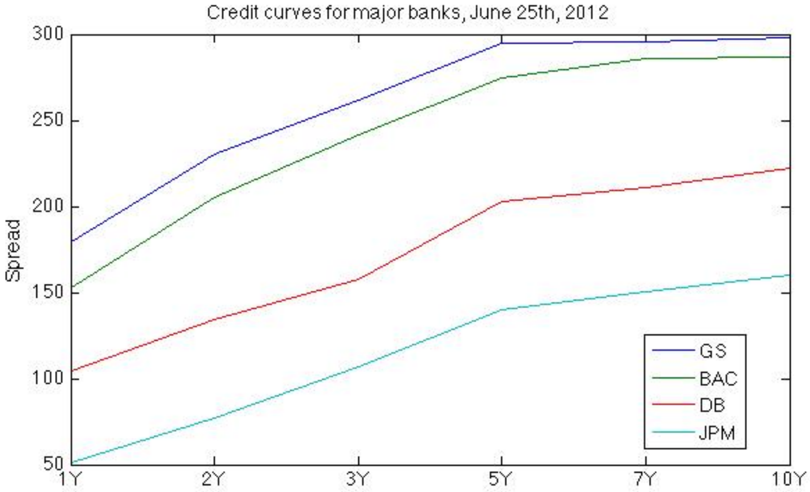
- Consider the exposure as if you were a risk manager: CDSs typically have 5Y maturities, and rates were about 5.5% in 2008.

  A five-year par bond with a rate of 5.5% has a sensitivity to credit spreads, or credit DV01, of about \$435 per **one** basis point for \$1 million notional.

  \$450 billion notional will have exposure of \$200 million per **one** basis point. A move of <u>50 bps</u> would generate roughly <u>\$10 billion</u> in losses. Mid-2007 through early 2008, spreads on five-year AAA financial issues rose from about 50 bps to about 150 bps...

Re-worked from: *A Practical Guide to Risk Management*, Coleman (2011)

Single-name CDS: credit curves



Basket CDS. Default time from intensity models

What is a **Basket CDS**?

Structured credit product, an OTC credit exotic. Product characteristics and pricing methodology give a practical insight in how to price tranched products (pools of securities), with correlation.

If the *k* number of defaults occur

- the contract terminates, the protection buyer receives

$$\mathrm{LGD} = (1 - R) \times \frac{1}{5} \times N$$

Protection seller receives $PL = s \times 10$ million per year. Paid periodically in arrears (vs. upfront fee CDS). Five reference names, 2m notional each, with maturity T = 5 years.

Potential Pricing

1st to default 'Equity'	22 bps Gaussian	t copula
2nd to default	10 bps	
5th to default	1.6 bps	

$s = \frac{DL}{PL}$ is computed separately for each k-th to default. The spread is an expectation over the joint loss distribution (with simplification).

Credit Spreads

There are spreads on standardised CDS indices, such as CDX North America and iTraxx Europe (5Y maturity) baskets of 50 top credit quality corporate names.

Why interest in synthetic credit? **Selling protection** (CDS) means the same exposure as **buying a bond** – leveraged,eg, borrowing the initial purchase price of the bond. Receive premiums until the maturity or default, and pay out the principal in case of default.

Price a fair spread for a basket of five reference names by sampling default times from both, Gaussian and Student's t copulae.

Hazard Rates Data (five snapshots)

What does default time τ_k depend on?

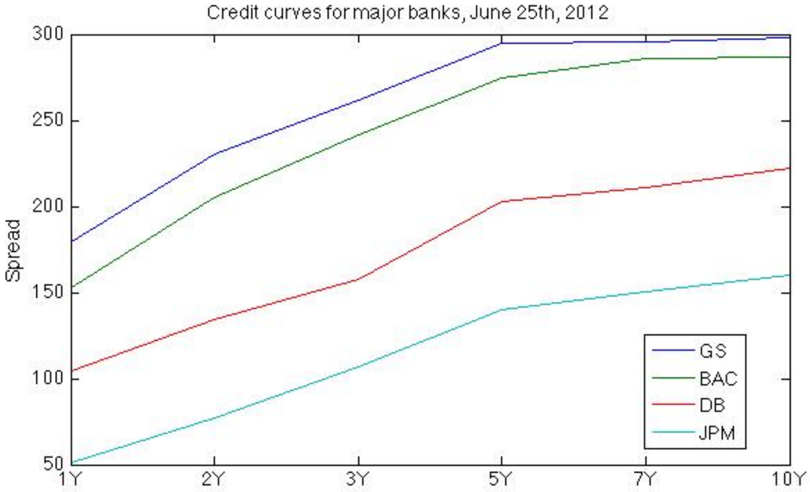
$$\boldsymbol{\tau} \sim Exp(\lambda_{1Y}, \dots, \lambda_{5Y})$$

For each reference name, we have to bootstrap hazard rates from the credit curve 'today'.

- 5 single-name par spread CDS with maturities up to 5Y.
- Matching discounting curve for 1Y,..., 5Y points.
- assume recovery rate $R = 40\%$ (there are LGD models)

CDS Lecture illustrates how to bootstrap implied survival probabilities and thus, hazard rates.

Credit Curves (term structure), bps



Survival Probabilities

The cumulative survival probability relates to hazard rate function:

$$\log P(0, t_m) = -\int_0^{t_m} \lambda_s ds = -\sum_{i=j}^{m} \lambda_j \Delta t_j \tag{1}$$

- $P(0, t_m)$ is survival probability up to time t_m
- λ_j is hazard rate between $j - 1$ to j
- Δt_j is gap between each period, likely to be 1 year

$P(0, t_m)$ is 'like' a discounting factor.

Hazard Rate for Each Tenor

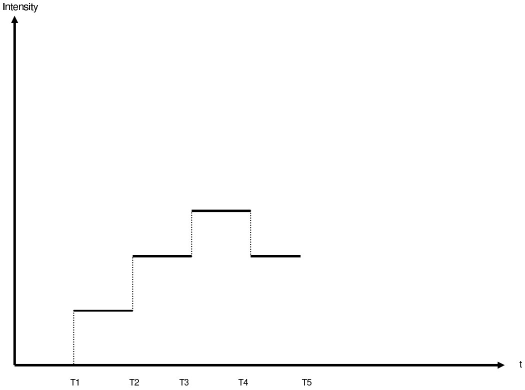
Assuming hazard rate function as piecewise constant, we bootstrap iteratively for each tenor Δt_j (year)

$$\begin{aligned}\lambda_1 &= -\frac{1}{\Delta t}\log P(0,t_1) \\ \lambda_m &= -\frac{1}{\Delta t}\log P(0,t_m) - \sum_{j=1}^{m-1}\lambda_j \end{aligned} \tag{2}$$

We can express the intensity as a ratio of survival probabilities:

$$\begin{aligned}\lambda_m &= -\frac{1}{\Delta t}\log P(0,t_m) + \frac{1}{\Delta t}\log P(0,t_{m-1}) \\ \lambda_m &= -\frac{1}{\Delta t}\log\frac{P(0,t_m)}{P(0,t_{m-1})} \end{aligned} \tag{3}$$

IHP calibrated on single-name CDS



For each name, we have a term structure $\widehat{\lambda}_{1Y}, \dots, \widehat{\lambda}_{5Y}$, a calibration to the Inhomogenous Poisson process (IHP).

Exponential inter-arrival times

- Suppose a simulation gives us correlated $(u_1, \ldots, u_5)$.
- Our task is to convert $u_i \to \tau_i$, done individually for **each** reference name by using (five) hazard rate from the curve

$$\begin{aligned}\tau_{\text{Name 1}} &\sim Exp(\lambda_{1Y}, \ldots, \lambda_{5Y}) \\ \tau_{\text{Name 2}} &\sim Exp(\lambda_{1Y}, \ldots, \lambda_{5Y})\end{aligned}$$

Exponential CDF is $u = 1 - e^{-\lambda\tau}$ so,

$$\log(1-u) = -\lambda_\tau \tau$$

There is input u but **two unknowns** λ_τ and τ.

Marginal default time (for each name)

$$\boxed{\tau = t_{m-1} + \delta t}$$

1. First, we find the year of default,
   i.e., determine that default occurs between t_{m-1} and t_m.

2. Second, we estimate the year fraction δt or use accruals.

Year of default

- Iterate adding up hazard rates λ_j

$$\tau = \inf\left\{ t > 0 \; : \; \log(1-u) \geq -\sum^{t} \lambda_m \right\}$$

where default occurs if inequality holds and

$t_{m-1} \leq \tau \leq t_m$

Comparison is done on negative scale because $\log(1-u) < 0$.

- If the inequality **holds** after adding λ_m then default occurs.

Validating example

- Using absolute values to compare on positive scale

$$|\log(1-u)| \leq \sum \lambda$$

- We construct a validation table, where small u implies a default

u	$\|log(1-u)\|$
0.90	2.3
0.50	0.69
0.25	0.2877
0.10	0.1054
0.05	0.0513

Exact default time

Exact default time $\tau = t_{m-1} + \delta t$ requires year fraction δt

$$1 - u = \exp\left(-\int_0^{t_{m-1}+\delta t} \lambda_s ds\right) = P(0, t_{m-1}) \exp\left(-\int_{t_{m-1}}^{t_{m-1}+\delta t} \lambda_s ds\right)$$

$$\log\left(\frac{1-u}{P(0, t_{m-1})}\right) = -\delta t \lambda_m$$

$$\delta t = -\frac{1}{\lambda_m} \log\left(\frac{1-u}{P(0, t_{m-1})}\right) \tag{4}$$

What matters in practice is how default event is determined and settled. Assume $\delta t = 0.5$ is called accruals.

$$\boxed{\tau = t_{m-1} + \delta t}$$

Copula. Correlation of Default Events

Marginal Distributions

Question

What is the distribution of default time is for each reference name?

Answer

It is an **Exponential Distribution** parametrised empirically by a set of five hazard rates (piecewise constant, per year).

$$
\begin{aligned}
\tau_{\text{Name 1}} &\sim Exp(\lambda_{1Y}, \ldots, \lambda_{5Y}) \\
\tau_{\text{Name 2}} &\sim Exp(\lambda_{1Y}, \ldots, \lambda_{5Y})
\end{aligned}
$$

Copula Method

A great deal of flexibility by separating:

- **Marginal distributions** for default times τ_i
- **Dependence structure** (correlation: linear, rank, calibrated based on MLE by an optimiser)

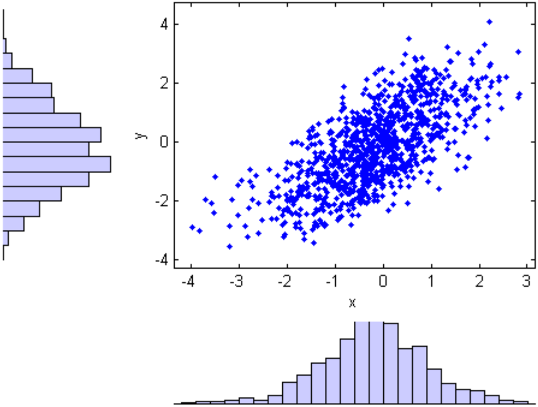
The joint distribution for k-th to default time across all reference names $\tau_k \sim F_k(t_1, t_2, \ldots, t_n)$ has **no closed form**. However,

$$F(x_1, x_2, \ldots, x_n) \equiv C(u_1, u_2, \ldots, u_n)$$

Let's review the concept of **copula**.

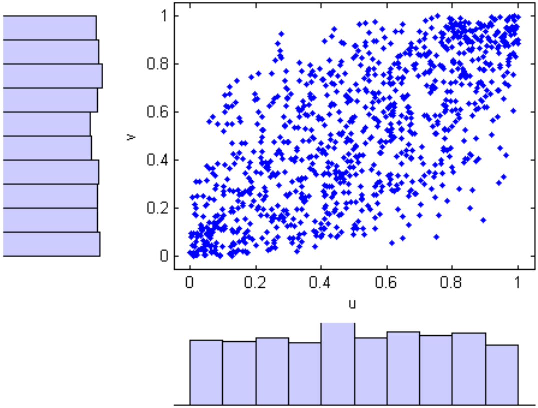
Joint Student's t Distribution

Joint distributions are cumbersome to work with and might have no analytical solution for CDF and ICDF.



Student's t Copula

Applying Student's t **CDF** to marginals means re-scaling to a uniform $[0, 1]$ projection. This should look familiar:



Sampling from copula

Copula method is a way to structure Monte Carlo: generate a correlated set $(u_1, u_2, u_3, u_4, u_5)$.

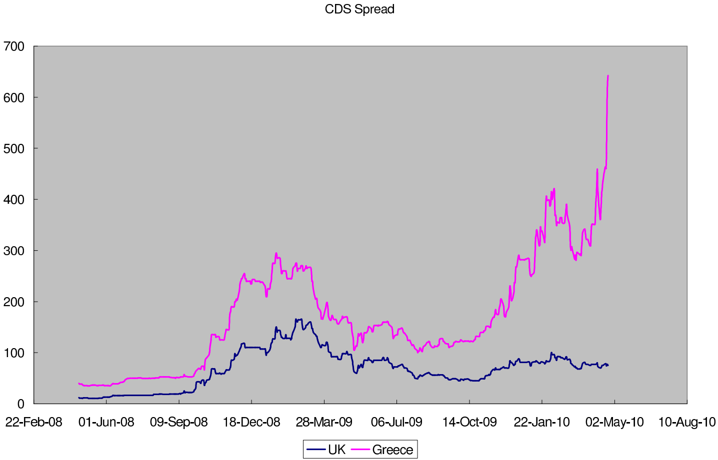
The approach is non-parametric which gives it a flexibility. The joint distribution of default time τ_k **unknown** for each case of *k-th to default*.

$$\tau_k \sim F_k(t_1, t_2, \dots, t_n)$$

- Implementation boils down to Cholesky decomposition $\widehat{\boldsymbol{\Sigma}}_\rho = \mathbf{A}\mathbf{A}'$
- We generate independent Normal RNs, and impose correlation by $\boldsymbol{X}^{Sim} = \boldsymbol{AZ}$, then convert back $\boldsymbol{U}^{Sim} = \Phi(\boldsymbol{X}^{\boldsymbol{Sim}})$

Correlation Estimation. Pseudo-samples

Historic Credit Spreads, bps



Correlating *levels* of credit spreads is spurious (unit root variables!).

Data for Default Correlation

Estimate correlation from **historical PD** data, as implied by credit spreads.

- 5Y tenor is a good reference point;
- daily or weekly changes; one-two year period.

Getting daily historical quotes for CDS and discounting a year or two back can be a challenge. Common substitutes:

- historical returns data (equity, bond yields) OR *base correlation* available from traded correlation instruments.

Distribution Fitting:

To apply correlation/transform to uniform easily we need data close to Normal histograms

- Obtain the time series for 5Y tenor point (CDS, PD, prices) and plot histograms. Non-stationary variables produce bi-modal histograms.

- Convert variables to *changes* ΔCDS, ΔPD or returns, subtracting means if necessary, and generate histograms.

  Hazard rate are already log-differences of the survival probabilities $\propto -\log \frac{P(0,t_m)}{P(0,t_{m-1})}$.

This is experimentation: demo can be provided but coding not to be shared.

Pseudo-samples: Normalising

Maximum Likelihood requires sample data transformed into uniform pseudo-samples $\mathbf{U}_t^{Hist}$. A simple recipe:

- Take data and convert into standardised changes (returns) $\mathbf{Z}$.

$$\Delta \text{CDS}, \Delta \text{PD} \rightarrow \mathbf{Z}^{\mathbf{Hist}}$$

- Under all assumptions (Normal distribution, volatility known)

$$\mathbf{U}_t^{Hist} = \Phi(\mathbf{Z}^{\mathbf{Hist}})$$

The drawback is that histograms of $\mathbf{U}_t^{Hist}$ might not be that uniform! Interferes with correlation.

Pseudo-samples: Kernel Smoothing (for pdf)

The proper recipe is $\mathbf{U}_t^{Hist} = \widehat{F}(\mathbf{PD})$ – you are not limited to PD data. Also computable $\mathbf{Z} = \Phi^{-1}(\mathbf{U})$.

This is numerically involved and involves at least two steps,

- fitting a pdf to a kernel function
- numerical integration over kernel to compute Empirical CDF $\widehat{F}()$

Matlab *ksdensity()* gives ready implementation. Use ‘ecdf’ and vary bandwidth ‘bw’ to obtain good-looking uniform histograms (each reference name). If things not uniform try more amenable data, eg, returns.

Dependence Fitting:

Plot 2D scatter plots (one reference name vs. another)

- Original data, variables such as CDS_{5Y}, prices – is this a valid move?
- Changes ΔCDS if using those.
- Pseudo-samples $\mathbf{U}$ – what does this plot represent?

The plots helps to visualise dependence, now moving on to formal estimation of 5×5 correlation matrix.

Correlation Matrix

- Once we got Normal $\mathbf{Z}$ from pseudo-samples $\mathbf{U}_t^{Hist}$ it is straightforward to calculate linear correlation matrix $\mathbf{\Sigma} = \rho(\mathbf{Z})$. Good enough for Gaussian copula.

- But t copula sampling requires **rank correlation**.

  Spearman's rho is estimated on pseudo-samples $\mathbf{\Sigma}_S = \rho(\mathbf{U})$. A separate formula $\mathbf{\Sigma}_\tau = \rho_\tau(\mathbf{X})$ is defined for Kendall's tau.

  Use $\rho = 2\sin\left(\frac{\pi}{6}\rho_S\right)$ and $\rho = \sin\left(\frac{\pi}{2}\rho_\tau\right)$ 'to linearlise' rank correlation.

The inferred linear correlation matrix is not guaranteed to be positive definite as required for Cholesky – so the nearest correlation matrix is obtained.

t copula

Fitting t copula gives an explicit exercise in **Maximum Likelihood**.

Two parameters to calibrate: correlation matrix $\widehat{\boldsymbol{\Sigma}}$ and ν, degrees of freedom. This is what ready software does and comes up with unrealistic correlations by optimisaiton.

$$\underset{\nu}{\operatorname{argmax}} \left\{ \sum_{t=1}^{T} \log c(\mathbf{u}_t^{Hist}; \nu, \widehat{\boldsymbol{\Sigma}}) \right\}$$

$\mathbf{u}'_{Hist}$ is a 1×5 row vector of observations for five reference names .

Log-likelihood of t copula

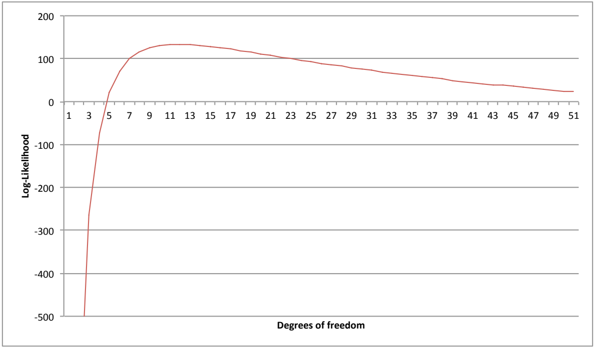


Figure: Use rank correlation matrix input, repeat calculation of the sum of log-likelihoods *for each value* of $\nu = 1 \ldots 25$ and plot.

t copula density

$$c(\mathbf{u};\nu,\widehat{\boldsymbol{\Sigma}}) = \frac{1}{\sqrt{|\Sigma|}} \frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \left(\frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}\right)^{n} \frac{\left(1+\frac{T_v^{-1}(\boldsymbol{u}')\boldsymbol{\Sigma}^{-1}T_v^{-1}(\boldsymbol{u})}{v}\right)^{-\frac{v+n}{2}}}{\prod_{i=1}^{n}\left(1+\frac{T_v^{-1}(u_i)^2}{v}\right)^{-\frac{v+1}{2}}}$$

- $\mathbf{u}'$ is a row vector of 1×5 that represents an observation of spreads (scaled) for five reference names on a given day.
- the denominator is calculated element-wise by drawing u_i from $\mathbf{u}'$.
- the nominator of the last term produces a scalar $1 \times 5 \times 5 \times 5 \times 5 \times 1$.
- in VBA use *EXP(GAMMALN())*. **ICDF** for Student's t T_v^{-1} code in Q&A, apply elementwise.

Sampling from copula. Correlated $(u_1, u_2, u_3, u_4, u_5)$

Sampling from Gaussian copula

1. Compute decomposition of correlation matrix $\widehat{\mathbf{\Sigma}} = \mathbf{A}\,\mathbf{A}'$. Use the simplest option of Cholesky decomposition if the matrix is positive definite.
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. Use Normal CDF to map to a uniform vector $\mathbf{U} = \Phi(\mathbf{X})$.

Convert each uniform variable to default time $u_i \to \tau_i$ using hazard rates structure for each name.

Inefficiency: RN uniform variables converted to Normal, correlation imposed, then we need CDF to convert back to uniform!

Differences are **rank correlation** and **chi-squared RN**.

1. Compute decomposition of correlation matrix $\widehat{\mathbf{\Sigma}} = \mathbf{A}\,\mathbf{A}'$.
2. Draw an n-dimensional vector of independent standard Normal variables $\mathbf{Z} = (z_1, \cdots, z_n)'$.
3. Draw an independent chi-squared random variable $s \sim \chi^2_\nu$. Compute n-dimensional Student's t vector $\mathbf{Y} = \mathbf{Z}/\sqrt{\frac{s}{\nu}}$.
4. Impose correlation by $\mathbf{X} = \mathbf{AY}$.
5. Map to a correlated uniform vector by $\mathbf{U} = T_v(\mathbf{X})$ using t CDF.

Why t copula?

1. If your calibrated (or reasonably assumed) d.f. $\nu = 7$.
2. The chi-squared random variable $s \sim \chi^2_\nu$ is obtained by drawing ν squared Normal random variables, separately

$$s = Z_1^2 + Z_2^2 + ... + Z_7^2$$

t copula means stronger co-movement. If simulated $u_1 = 0.1$, then $u_2 \approx 0.1$ is likely.

Think of the impact on the spread: multiple defaults together raise the spread for k-th to default, compared to Gaussian.

Where is the copula?

Question

Usually at this stage, a question is asked: "Where in these algorithms is the copula?"

Answer

What we do with imposing correlation by Cholesky result $\boldsymbol{A}$ is equivalent to **factorisation** of the copula into a set of linear equations.

Check CDO lecture for two-dimensional Cholesky solution for $\boldsymbol{A}$ (p.52)

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

Review [Reference]

We have covered how to:

- estimate appropriate correlation matrix
- sample a correlated random $U = (u_1, \dots, u_n)$ from copula
- convert each random variable into default time $u_i \to \tau_i$

In our simulation of correlated default events, their marginal distributions $\tau_i \sim Exp(\widehat{\lambda})$ are kept separately from dependence structure, a linear correlation matrix $\widehat{\boldsymbol{\Sigma}}$.

The joint distribution for k*th*-to-default time across all reference names $\tau_k \sim F_k(t_1, t_2, \dots, t_n) \equiv C(u_1, u_2, \dots, u_n)$ has been represented by a factorised copula (Cholesky linear system).

Spread Computation

kth-to-default spread

Par spread of k*th*-to-default swap is derived by equating DL $=$ PL.

$$s = \frac{< DL >}{< PL_{\not\$} >} = \frac{(1-R)\sum_{i=1}^{m} Z(0,t_i)\,(F_k(t_i) - F_k(t_{i-1}))}{\Delta t \sum_{i=1}^{m} Z(0,t_i)\,(1 - F_k(t_i))}$$

Assume, we simulated default times $(\tau_{N1}, \tau_{N2}, \tau_{N3}, \tau_{N4}, \tau_{N5})_{1..10{,}000}$

The joint distribution $\tau_k \sim F_k(t_1, t_2, \dots, t_n)$ remains **unknown**. It is likely to be a different distribution for each kth-to-default instrument.

Total Expected Loss

Structured credit pricing is simplified by introducing the **Total Expected Loss**, an expectation over the joint distribution,

$$\mathbb{E}[F_k(t)] = L_k \qquad L_i - L_{i-1} = \frac{1}{5} \times \text{Notional}$$

$$\mathbb{E}[s] = \frac{(1-R)\sum_{i=1}^{m} Z(0,t_i)\,(L_i - L_{i-1})}{\Delta t\sum_{i=1}^{m} Z(0,t_i)\,(NP - L_i)}$$

To satisfy the expectation, the fair spread is calculated using Monte-Carlo.

Upon k-th to default, the notional payment is made by protection seller for the defaulted entity.

- s is fair spread of the contract paid $\frac{1}{dt}$ times per annum until τ_k **or** maturity. $\Delta t \approx t_i - t_{i-1}$ is an accrual factor.
- summation in the spread is over m years, with $\tau = t_{m-1} + \delta t$.
- $Z(t, T)$ is a risk-free zero coupon bond price as discount factor.
- R is recovery rate, $LGD = 1 - R$.
- $NP = 1$ is notional principal. We invest $\frac{1}{5}$ of notional in each name.

Spread Computation: 1st to default, τ_1

Loss Function per time period $L_i - L_{i-1} = \frac{1}{5} \times NP$. This simplifies computation.

$$s = \frac{(1-R)Z(0,\tau_1) \times \frac{1}{5}}{Z(0,\tau_1)\,\tau_1 \times \frac{5}{5}}$$

For $\tau_1 < 5$ years, $Z(0,\tau_1)$ in the DL numerator coincides with one in PL denominator. We keep it in both expressions because we compute average PL and DL separately.

Each kth-to-default basket is priced as **a separate instrument**. Simulations can be saved and re-used but N_{sim} likely to be different. Price by both, Gaussian and t copulæ.

Spread Computation: 1st to default, table

If default time $\tau_k \geq 5$ years then $DL = 0$ but the paid premium has to be discounted. Assuming annual payment, discretisation goes

$$Z(0,1) \times 1 + Z(0,2) \times 1 + ... + Z(0,5) \times 1$$

0	$(1-R)/5$	0	0	0	$(1-R)/5$	...
$DF \times 5$	τ_1	$DF \times 5$	$DF \times 5$	$DF \times 5$	τ_1	...

Very small default times τ_k lead to large spreads and interfere with convergence. Can introduce a floor $\tau_k = \max(\hat{\tau}_k, 0.25)$.

Average DL and PL across simulations **separately**, and calculate the spread one time. Done to improve convergence.

Spread Computation: 2nd to default

2nd-to-default also protects from the loss in **single name**

$$s = \frac{(1-R)Z(0,\tau_2)\times\frac{1}{5}}{Z(0,\tau_1)(\tau_1 - 0)\times\frac{5}{5} + Z(0,\tau_2)(\tau_2-\tau_1)\times\frac{4}{5}}$$

Reference entity defaulted *before* k-th default removed from the portfolio, reducing its value by $\frac{1}{5}\times NP$ each.

The rule is consistent with the Removal of Defaulted Reference Entity provisions of ISDA Terms.

Without the removal, the denominator PL $= Z(0,\tau_2)\tau_2\times\frac{5}{5}$.

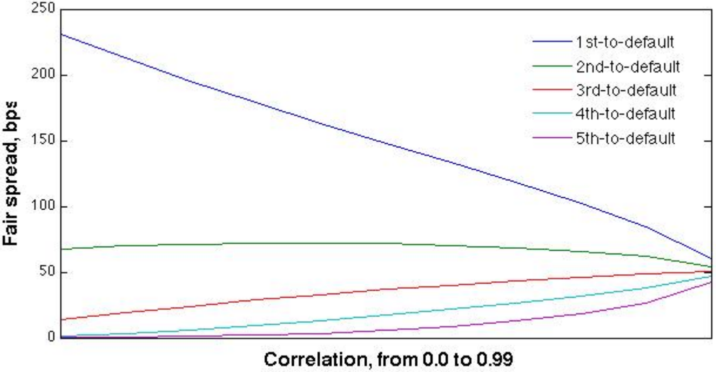
In practice, product implementation might prefer this.

Model Validation

- The fair spread for k-th to-default Basket CDS should be less than k-th -1. Why?
- **Risk and Sensitivity Analysis** of the spread is important
  1. default correlation among reference names: either stress-test by constant high/low correlation or $\pm$ percentage change in correlation from the actual estimated levels.
  2. credit quality of each individual name (change in credit spread, credit delta) as well as recovery rate.
- Correlation matrix is key input, so make sure to explain:
  1. historical sampling of default correlation matrix, and
  2. choice of the stress-testing levels of correlation, i.e., what kind of event they represent

Sensitivity to constant correlation

As default correlation increases to very high levels, spreads for different kth-to-default instruments lapse. Why?



Which default correlation levels have you obtained from linear and rank correlation measures?

Data Requirements - Reference

1. *A snapshot* of credit spreads on a given day is used in estimation of hazard rates:
   - For each reference name, the term structure of hazard rates for 1Y, 2Y,...5Y (non-cumulative) parametrises the distribution of default time τ.
2. *Historical* credit spreads data is needed for estimation of the (inferred) linear correlation matrix of PD.
   - Alternative estimation of default correlations is possible. Please see below and consult with the Q&A.
3. Discounting curve data is necessary for both, hazard rates bootstrapping and basket spread s calculation. Approximate.

Basket CDS Implementation Step-by-Step

1. For each reference name, bootstrap implied default probabilities from quoted CDS and convert them to hazard rates.
2. Estimate the appropriate inputs for 'sampling from copula', i.e., correlation matrix and degrees of freedom.
3. For each simulation, repeat the following routine:
   1. Sample a vector of correlated uniform random variables – you will need to implement sampling from both Gaussian and Student's t copula separately.
   2. Use hazard rates of each reference name to convert the corresponding uniform variable of u_i into exact default time τ_i.
   3. Based on τ_k calculate the discounted values of premium and default legs.
4. Average premium and default legs across simulations separately. Calculate the fair spread s.

END OF WORKSHOP