CQF Final Project Topics: Credit Analysis, Portfolio Construction

January 2016 Cohort – Workshop II

1. Introduction to CQF Final Project
2. Pricing a synthetic credit portfolio: a kth-to-default Basket CDS
3. Portfolio Construction using the Black-Litterman Model

Implement ONE topic plus CVA Component.

Topic and Design Choice

- Start collecting data and planning how to complete the project.
- It is up to you to source and clean the suitable input data.

  We will review the data necessary for each topic, further description and links are in the Project Brief. Any suitable alternative data can be used, including pre-simulated data.

- Model validation (test cases, sensibility checks) is part of the task assigned.

  We will cover suggestions in these slides as well as Project Q&A.

  - In particular, set your option strikes and maturities, missing forward rates, CDS spreads.

Topic and Design Choice - Numerical Techniques

- Implementation of numerical techniques from the first principles is the purpose. Pricing of a derivative/credit instrument/robust allocation is the result.
  - In particular, make a choice of methods of curve fitting, numerical integration, and random numbers generation.
- Refer to the relevant CQF Lectures and do extra reading on pricing methodologies, quant finance models and numerical techniques.

Project Report

- A full **mathematical description** of the models employed as well as any numerical methods. Include mathematical measures of *accuracy and convergence*.
- Results presented using **a plenty of tables and figures**, together with sensible interpretation.
- **Pros and cons** of a particular model and its implementation, together with possible improvements.
- **Demonstrate 'the specials'** of your implementation: own research, re-coding of numerical methods, using the industrial-strength libraries of C++, Python or VBA + NAG.
- Instructions on how to use software if that is not obvious. The code must be thoroughly tested and well-documented.

Project Topics: Pricing a Credit Portfolio (spread)

The numerical techniques and pricing are applicable to fair value estimation for structured debt/CDO/CLO.

Why interest in synthetic credit? Selling protection means the same exposure as holding the bond.

> **k*th*-to-default Basket CDS**. Price a fair spread for a basket of five reference names by sampling default times from Gaussian and Student's t copulae.
>
> The spread is calculated as an expectation over the joint loss distribution.

Fair Spread for A Basket CDS

What is a Basket CDS?

The standardised CDS indices CDX North America and iTraxx Europe (5Y maturity) are issued tranches and liquid. Less need in structuring a customised Basket CDS.

However, the product characteristics and pricing methodology give a practical insight into fair valuation of credit portfolio products.

kth-to-default swap is an OTC credit exotic. If the k-th default occurs

- protection payment is made for that entity, and
- the contract terminates,
- the loss is always $L = \frac{1}{5} N$.

k-th

1st-to-default BDS

Five reference names, 2m notional each, with maturity T = 5 years.

Assume premium (fair spread) at 75bp **estimated** as $s = \frac{DL}{PL}$

- The protection seller receives $PL = 75bp \times 10m = 75,000$ per year. Paid periodically in arrears (vs. upfront fee CDS).
- If one of reference names defaults at time $\tau_{k=1} < 5$, the protection buyer receives

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

Hazard Rates Data (five snapshots)

par spread

What does default time τ_k depend on?

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

$SPL = DL \Rightarrow S = \frac{DL}{PL}$

$\tau > 5Y$

$DL = 0$

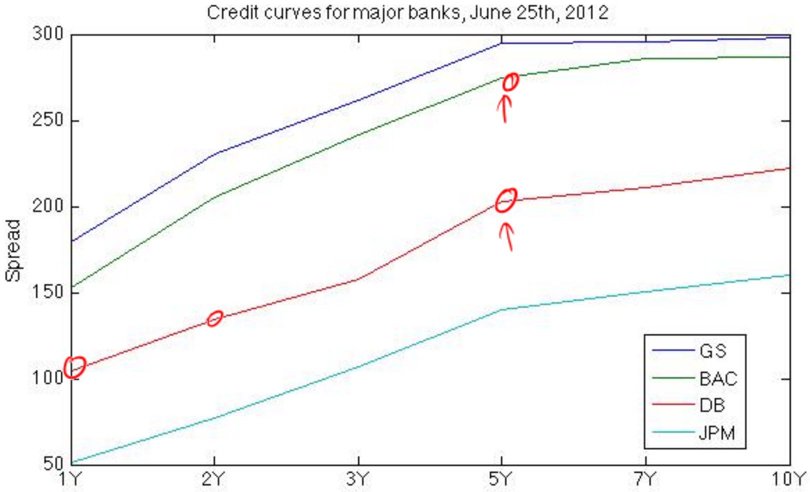
For each reference name *separately* we have to bootstrap hazard rates (default intensities) from the credit curve 'today'.

- 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.

CDS $\rightarrow$ Pr Surv $P(0, t_m)$ $\rightarrow$ Hazard Rates

Credit Curves (term structure), bps



Survival Probabilities

Once bootstrapped, the cumulative survival probability for each reference name is related to hazard rate function as follows:

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

exp exp

$$Pr\ Surv = e^{-\int_0^{t_m} \lambda_s ds}$$

- $P(0, t_m)$ 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

Exact default time is obtained by integration over the hazard rate function – but we need a discretized scheme.

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:

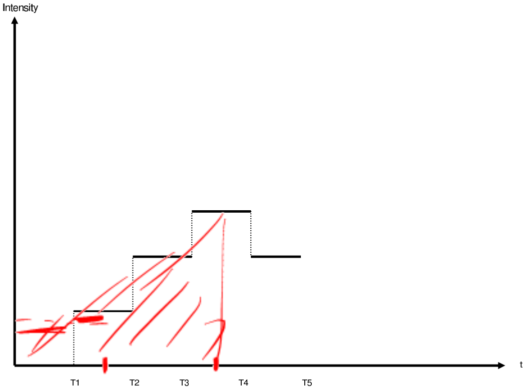
$\Delta t = 1Y$

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

$P(0, t_m)$

1Y	0.99
2Y	0.975
...	...

IHP calibrated on single-name CDS



For each name, we have a term structure $\widehat{\lambda}_{1Y}, \ldots, \widehat{\lambda}_{5Y}$.

Exponential inter-arrival times

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

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

Exponential CDF is $u \stackrel{D}{=} F(\tau)$ and $u = 1 - e^{-\lambda\tau}$ so, the inverse $\tau \stackrel{D}{=} F^{-1}(u)$ gives, CDF

ICDF

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

$\lambda(\tau)\tau$

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.

   $\delta t = 0.5$ 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

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

0.20

0.10

0.02

14-24

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

u	$\lvert log(1-u) \rvert$
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

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

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

Pr(0, τ)

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

cum

non-cum. data

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

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

τ = 2.33

Marginal Distributions

Question

What we have covered thus far describes the properties of the marginals. What are they?

Answer

The marginal distribution of default time is *different* for each reference name. It is **Exponential Distribution** parametrised empirically by a set of five piecewise constant hazard rates.

$\underset{\substack{\uparrow}}{\tau_i} \sim Exp(\widehat{\lambda}_{i,1Y}, \lambda_{i,2Y}, \ldots \lambda_{i,5Y})$

i - reference name

Copula Method

The copula method allows a great deal of flexibility:

- The method separates **dependence structure** from **marginal distributions** for default times τ_i, which we parametrised empirically, each by a term structure of $\widehat{\lambda}_{i,1Y}, \dots, \widehat{\lambda}_{i,5Y}$.

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

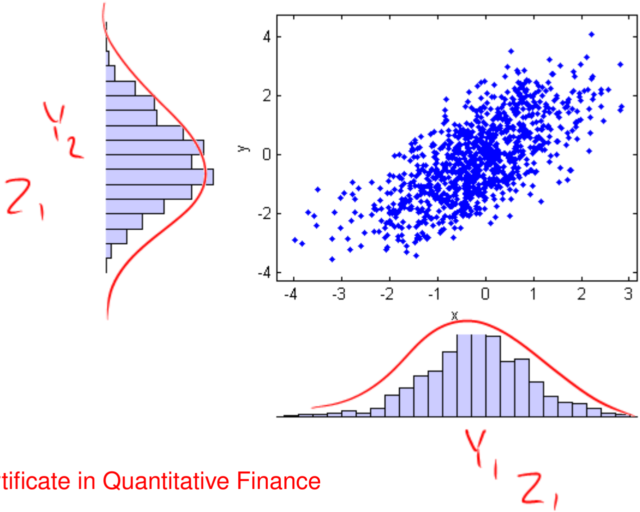
$$F(x_1, x_2, \dots, x_n) \equiv C(u_1, u_2, \dots, u_n)$$

Joint CDF

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:

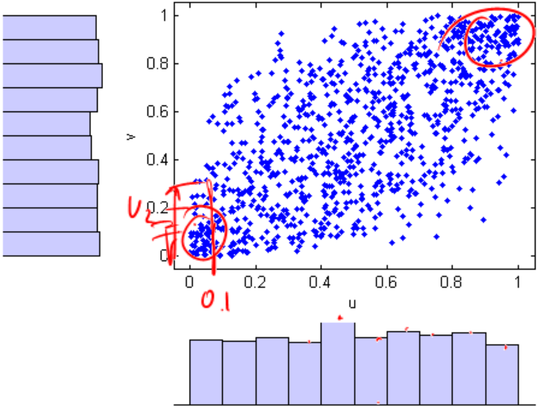
$U_1 = F(y_1)$

ν = d.f.

low d.f.

"fat tails"

U_2



U_1

Sampling from copula

Copula example above comes from the draws of correlated random variables (joint t distribution). This is parametric.

But the joint distribution of default times $\tau_k \sim F_k(...)$ is **unknown**.

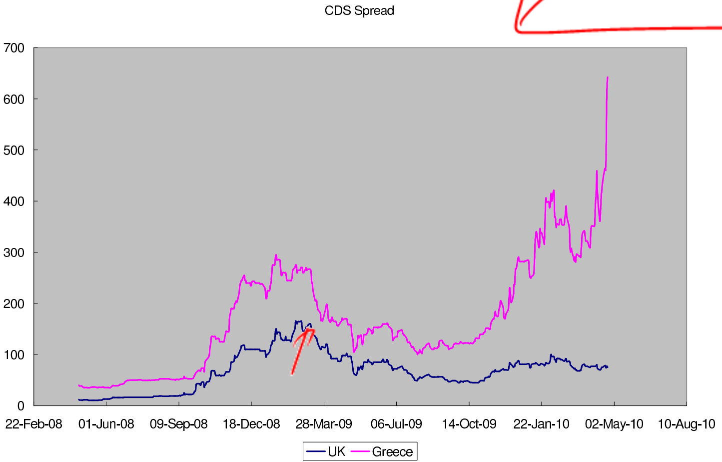
- We **sample from copula** correlated random values $U = (u_1, \dots, u_n)$. 'Copula method' refers to imposing correlation.
- Implementation boils down to Cholesky decomposition $\widehat{\mathbf{\Sigma}}_\rho = \mathbf{AA}'$ and correlating by $\mathbf{X} = \mathbf{AZ}$.

Numerical approach keeps flexibility. $F_k(t_1, t_2, \dots, t_n)$ is unknown.

Historic Credit Spreads, bps

JPM

5Y



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

Data for Default Correlation

Correlation matrix is estimated from **historical PD** data, as implied by credit spreads.

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

ΔCDS vs ΔPD, Δs

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

- historical returns data, i.e., bonds or equity (weekly log rtns)
- *base correlation* available from traded correlation instruments, often the same among all reference names.

DB Research

Distribution Fitting Experiment:

- Obtain the time series for 5Y tenor point (CDS, PD, hazard rates) and plot individual histograms. Are the variables bi-modal?
- Convert variables to *changes* (ΔCDS, ΔPD) and study the new histograms.

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

Pseudo-samples

Canonical Maximum Likelihood assumes that sample data is transformed into uniform pseudo-samples $\mathbf{U}_t^{Hist}$. How do we get that?

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

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

- Under all assumptions (Normal distribution, volatility known)

$$\mathbf{U} = \Phi(\mathbf{Z})$$

- The proper steps are $\mathbf{U} = \widehat{F}(\mathbf{PD})$ and $\mathbf{Z} = \Phi^{-1}(\mathbf{U})$.

  But obtaining $\widehat{F}()$ means first, fitting a pdf from empirical data and second numerical integration for each point.

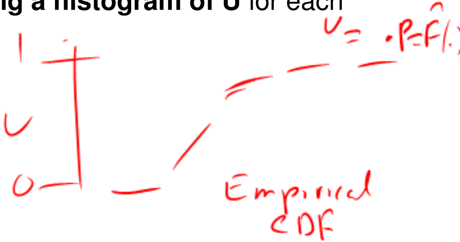
Pseudo-samples (cont.)

"Gaussian kernel"
pdf as function

Matlab's *ksdensity()* allows direct fitting $\mathbf{U} = \widehat{F}(\mathbf{PD})$ via Empirical CDF. Further details on **kernel density estimation** are in the Project Q&A.

Uniformity Check:

- **Test your smoothing by plotting a histogram of U** for each reference name (column).



Dependence Fitting Experiment:

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

- pseudo-samples **U** ← uniformity check , 2D
- changes data (e.g., ΔCDS) 2D
- original variables data (e.g., CDS_{5Y})

What do you observe?

Correlation Matrix

- Once we got Normal **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.

  ΔCDS

  linear correlation

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

  U

  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.

  To convert rank into linear correlation the special formulae are used $\rho = 2\sin\left(\frac{\pi}{6}\rho_S\right)$ and $\rho = \sin\left(\frac{\pi}{2}\rho_\tau\right)$.

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

Things are not as simple with fitting t copula.

There are two parameters to obtain by likelihood maximisation on pseudo data $\mathbf{U}$:

- inferred correlation matrix $\widehat{\mathbf{\Sigma}}$ and degrees of freedom ν.

t copula: MLE in detail

Keeping correlation matrix constant, calculate the total log-likelihood(s) *for the range* of values $\nu = 1 \dots 25$ and plot as a function.

d.f.

Log-likelihood has familiar form of 'contributions to density' from each observation

by row

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

from U

$\rho(U)$

where $\mathbf{U}'$ is a 1×5 row vector of observations for five reference names at a given time t (day/week).

$$\sum_{t=1}^{N_{obs}} \log c_t$$

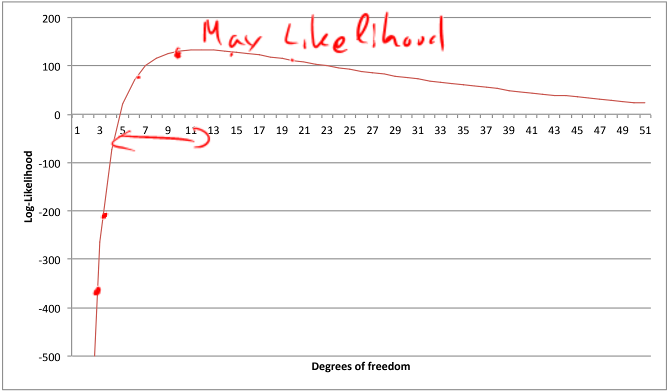


Figure: Log-Likelihood function for t copula vs d.f. parameter. Scale depends on the original transformation into **U** (the log-likelihood can be negative).

Student's t copula density

Inverse CDF

$U_{1\times 5}$

$m=5$

$1\times5 \times 5 \times 5 \times 5 \times 1$

$$c(\mathbf{U};\nu,\widehat{\boldsymbol{\Sigma}}) = \frac{1}{\sqrt{|\Sigma|}}\frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)}\left(\frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}\right)^{n}\frac{\left(1+\frac{T_v^{-1}(\boldsymbol{U}')\boldsymbol{\Sigma}^{-1}T_v^{-1}(\boldsymbol{U})}{v}\right)^{-\frac{v+n}{2}}}{\prod_{i=1}^{n=5}\left(1+\frac{T_v^{-1}(u_i)^2}{v}\right)^{-\frac{v+1}{2}}}$$

$U=(u_1, u_2, u_3, u_4, u_5)$

Taking derivative of $\frac{\partial \log c}{\partial \nu}$ is possible but solving the ensuing equation is not feasible. d.f. estimation is done numerically.

This copula density is not guaranteed to be within $[0, 1]$.

t copula density explained

To calculate the Student's t copula density function we need

- Gamma function $\Gamma(v)$, available in VBA via *GammaLn()*.
- the inverse of Student's t **CDF** T_v^{-1}. There is no ready Excel function – see solution via Beta function in Peter Jaekel's textbook (p.14). VBA code uses *BetaDist()*.
- $\mathbf{U}'$ is a row vector of 1×5 that represents an observation of spreads (scaled) for five reference names on a given day.
- the nominator of the last term produces a scalar by dimensions $1 \times 5 \times 5 \times 5 \times 5 \times 1$.
- the denominator is calculated element-wise by drawing u_i from $\mathbf{U}'$ for $i = 1..n$, $n = 5$.

Sampling from Gaussian copula

Nocred Correlation

1. Compute decomposition of correlation matrix $\widehat{\boldsymbol{\Sigma}} = \mathbf{A}\,\mathbf{A}'$.
   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 $\mathbf{U} = \Phi(\mathbf{X})$.

$U = (u_1, \ldots, u_5)'$ CDF Normal

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

$= \tau_1, \ldots, \tau_5$

Exp($\hat{\lambda}$) Exp($\hat{\lambda}$)

Sampling from Student's t copula

Differences to the Gaussian copula underlined. rank correlation

1. Compute decomposition of correlation matrix $\widehat{\mathbf{\Sigma}} = \mathbf{A}\,\mathbf{A}'$.
2. Draw an n-dimensional vector of independent standard Normal variables $\mathbf{Z} = (z_1, \cdots, z_n)'$. $\nu = 3$ $S = Z_1^2 + Z_2^2 + Z_3^2$
3. <u>Draw an independent chi-squared random variable $s \sim \chi^2_\nu$.</u>
   <u>Compute n-dimensional Student's t vector $\mathbf{Y} = \mathbf{Z}/\sqrt{\frac{s}{\nu}}$.</u>
4. Impose correlation by $\mathbf{X} = \mathbf{AY}$.
5. Map to a correlated uniform vector by $\underline{\mathbf{U} = T_\nu(\mathbf{X})}$ using t CDF. cdf t

t copula means stronger co-movement. If one variable $u_1 = 0.1$, the other is more likely to be $u_2 = 0.1$ than say, 0.6. This must produce simulations $\mathbf{U}$ with multiple defaults.

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)

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

Overview

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 $u_i \rightarrow \tau_i$

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).

kth-to-default spread

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

$$s = \frac{< DL >}{< PL_{\$} >} = \frac{(1-R)\sum_{i=1}^{m} Z(0,t_i)\ \overbrace{(F_k(t_i) - F_k(t_{i-1}))}}{\Delta t \sum_{i=1}^{m} Z(0,t_i)\ (1 - F_k(t_i))}$$

assuming the instrument exists over $m = 1, 2, 3, 4, 5$ years.

From a copula, we simulated default times

$(\tau_{N1}, \tau_{N2}, \tau_{N3}, \tau_{N4}, \tau_{N5})_{1..10,000}$

But, the joint distribution for k*th*-to-default time across all reference names $\tau_k \sim F_k(t_1, t_2, \dots, t_n)$ remains **unknown**.

k-th

1-st ...

2-nd ...

Total Expected Loss

- As seen in CDO Lecture, the Loss Function is defined as an expectation over the joint distribution $L_k = \mathbb{E}[F_k(t)]$.

$$\mathbb{E}[s] = \frac{(1 - R) \sum_{i=1}^{m} Z(0, t_i) \overbrace{(L_i - L_{i-1})}^{1/5}}{\Delta t \sum_{i=1}^{m} Z(0, t_i) (NP - L_i)}$$

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

  It is possible to use simulated default times and build Loss Distribution, separately for 1*st*-, 2*nd*-to-default, etc.

Spread Calculation Explained

Upon k-th default at time τ_k, 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 periods which corresponds to $\tau = t_{m-1} + \delta t$ notation.
- $Z(t, T)$ is a risk-free zero coupon bond price as discount factor.
- R is recovery rate, $LGD = 1 - R$.
- NP is notional principal (scaled as $NP = 1$).
  We invest $\frac{1}{5}$ of notional in each reference name.

Spread Computation: 1st and 2nd-to-default

BDS protects loss per time period $L_i - L_{i-1} = \frac{1}{5} \times NP$. The formulae are for continuous time (no payment frequency, $\tau_k < 1$).

40%

$U_1, U_2, \ldots, U_5$
Gaussian + copula

- 1st-to-default

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

$\tau_1, \tau_2, \ldots, \tau_5 > 5$

$\tau_1, \tau_2, \tau_3, \ldots < 5$

$s = 0$

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

s = as above

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

$s \times N * Z(0,\tau_i)\, \tau_i$

Each <u>kth</u>-to-default basket is priced as **a separate instrument**, with its own set of simulations.

Spread Computation: 1st-to-default annual payment

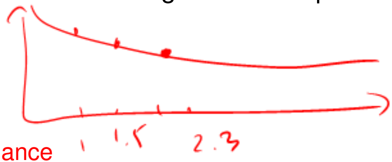
To discretise calculation according to payment frequency, for example, per annum.

- Assuming the 1st default occurs into fourth year, $\tau_k > 3$

$$s = \frac{(1-R)Z(0,\tau_1) \times \frac{1}{5}}{[Z(0,t_1)\Delta t_1 + Z(0,t_2)\Delta t_2 + Z(0,t_3)\Delta t_3 + Z(0,\tau_1)\delta t] \times \frac{5}{5}}$$

where $\tau = t_{m-1} + \delta t$, $\Delta t = 1$ and accrual factor $\delta t = 0.5$.

With the exact default time, say $\delta t = 0.31$, accrual is continuous and the interpolation (fitting) over discounting curve is required. Interpolate over log DF.



$s = \frac{\langle DL \rangle}{\langle PL \rangle}$

1. **Average** *DL* **and** *PL* **across simulations separately, and calculate a spread after.** Done to improve convergence.

2. If default time $\tau_k \geq 5$ years then $DL = 0$ and $\therefore s = 0$.
   Very small default times τ_k **lead to large spreads** and interfere with convergence. Can introduce a floor $\tau_k = \max(\hat{\tau}_k, 0.25)$.

3. Reference entities that have defaulted *before* k*th* default are 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.

Model Validation

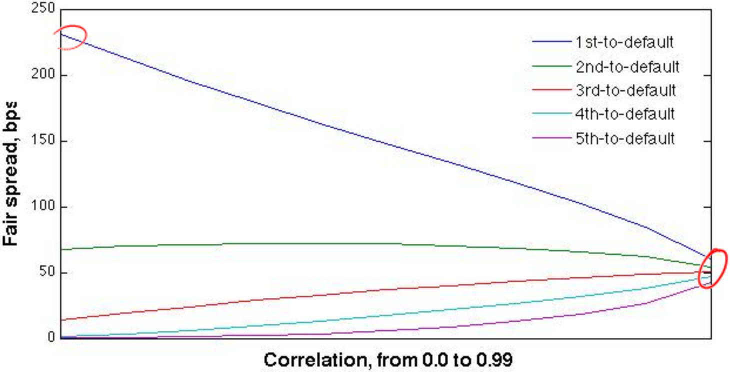
- The fair spread for kth-to-default Basket CDS should be less than kth-1. Why?

- **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.

1. Introduction to CQF Final Project
2. Pricing a synthetic credit portfolio: a kth-to-default Basket CDS
3. Portfolio Construction using the Black-Litterman Model

Project Topics

Portfolio Construction. The Black-Litterman Model allows to introduce the views into portfolio optimisation improving robustness and sensibility of allocations.

w^*

a. Using market benchmarks of your choice construct the prior, define input views, and estimate *expected* posterior returns.

b. Obtain allocations by variance minimisation and at least one more **alternative index of satisfaction** (SR, VaR/CVaR, Tracking Error). Experiment with different levels of risk aversion.

c. [Optional] Improve robustness of inputs with time series analysis techniques (GARCH, correlation robustness).

d. [Optional] Explore a multi-period allocation. That naturally leads to the back-testing of an investment into a factor strategy.

Portfolio Construction using the Black-Litterman Model

Black-Litterman Derivation

The pathbreaking approach by Black and Litterman (1990) allows to obtain more satisfactory results by blending the market-driven allocation with views.

Confidence in the view portfolios and uncertainty on the reference market model are set explicitly.

The mean-variance optimization takes the expected returns and covariances as inputs. One has to employ additional robustness techniques to avoid ‘garbage in, garbage out’.

Non-robust Allocations

To say that sample estimates are very **inefficient** is an understatement. Estimates vary from sample to sample. Ensuing mean-variance optimisation is far from robust.

*Plugging historical mean and variance into a mean-variance optimizer and implementing its portfolio advice is a **terrible** guide to investing. Practically anything does better.* $1/N$ *does better.*

John Cochrane

$w = \frac{1}{N_{assets}}$

$w \leftarrow \Sigma$

RMT

Estimates Shrinkage

If you have a more trustworthy estimate apply the shrinkage,

$$\mu = (1 - s)\,\hat{\mu} + s\,\pi$$

The technique is useful **a.** when you more conservative return expectation than a recent excellent quarter/year or vice versa,
b. the recent data comes from a particularly volatile period.

1. The sample estimate weighted with the smaller $(1 - s)$, typically this is recent data.

2. For market indices and factors, the average daily return is likely to negligible; for a FTSE100 index $\hat{\mu} \approx 0$.

Variance Smoothing

Volatility estimates are plagued by heteroskedasticity – the arbitrary sample of, say 60 days, might have a few particularly low returns causing the volatility estimate to spike.

Adding correlation to this problem leads to very involved methods for covariance matrix (Dynamic Conditional Correlation, PCA GARCH).

But we can go for a simpler recipe to apply univariate GARCH to each asset separately and re-construct covariance matrix

$$\mathbf{\Sigma}_{ij} = \rho_{ij}\,\sigma_i^{GARCH}\,\sigma_j^{GARCH}$$

1. Consider weekly/monthly vs. daily returns.
2. Correlation coefficient $\rho \in [-1, 1]$ has no timescale but varies with sample size. Larger samples give smoother rolling estimate.
3. Original GARCH model designed to forecast the next day's volatility σ_{t+1}, so consideration should be given to the model choice if estimating from weekly/monthly returns.

Uncertainty of estimates

Prior | Views $V|\mu$ = Posterior

BL assumes the distribution of excess returns $f_\mu(\mu)$ is Normal

$$\mu \sim N(\underset{\text{Eq. Return}}{\pi}, \underset{\text{Error}}{\tau\boldsymbol{\Sigma}}) \tag{5}$$

$\tau = \frac{1}{252}$

Equilibrium returns π are our 'best estimate' for expected returns.

$\tau = \frac{1}{T}$

$\tau\boldsymbol{\Sigma}$ is effectively **the estimation error** on μ.

Uncertainty of estimates, $\tau\boldsymbol{\Sigma}$ is BL input $O(10^{-4})$ to $O(10^{-5})$. We choose $\tau = 0.4$, higher than suggested $0.01 - 0.05$ range.

We start with 'an oversimplified example' from
The Black Litterman Approach guide by Attilio Meucci (2010):

Six market indicies of Italy, Spain, Switzerland, Canada, US and Germany with similar annualised historic volatilities

$$\boldsymbol{\sigma} = (21\%,\ 24\%,\ 24\%,\ 25\%,\ 29\%,\ 31\%)$$

with and correlations in the range $0.4 \leq \rho \leq 0.8$.

This is more informative than the covariance matrix, recovered as

$$\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\sigma})\,\textbf{Corr}\,\text{diag}(\boldsymbol{\sigma})$$

BL Inputs: expected returns

We take **market allocations $\tilde{\mathbf{w}}$** from a benchmark index

$$\tilde{w}' = (4\%, 4\%, 5\%, 8\%, 71\%, 8\%)$$

and calculate the equilibrium returns with risk-aversion $\lambda = 1.12$

$$\boldsymbol{\pi}' = (6\%, 7\%, 9\%, 8\%, 17\%, 10\%)$$

$\boldsymbol{\pi}$ is **the mean** of the reference distribution (**the prior**).

$$\boldsymbol{\mu} \sim N(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma})$$

How did we calculate $\tilde{\boldsymbol{w}} \to \pi$? Begin with the optimisation

$$\underset{\boldsymbol{w}}{\operatorname{argmax}} \{\boldsymbol{w}'\pi - \lambda \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}\}$$

Solve **reverse optimisation** problem to obtain equilibrium returns π from market index allocations $\tilde{\boldsymbol{w}}$

$$\boxed{\pi = 2\lambda\boldsymbol{\Sigma}\tilde{\boldsymbol{w}}}$$

Direct optimisation gives solution for allocations $\boldsymbol{w}^*$,

$$\boxed{\boldsymbol{w}^* = \frac{1}{2\lambda}\boldsymbol{\Sigma}^{-1}\pi}$$

Choice of the prior

How do we choose a prior? Say, we invest in Emerging Markets.

- **Do not** naively optimise the mean-variance over estimates calculated from a sample.
- Form equilibrium allocations π using the weights of the relevant MSCI capped index $\tilde{\boldsymbol{w}}$.
- Calculating weights by market cap is possible but gives a bias. If not benchmarked, one can start with allocations $\tilde{\boldsymbol{w}} = 1/N$.

Devise a suitable prior by referring to weights in specialised indicies (ready multi-asset indicies are rare).

- MSCI, S&P Dow Jones, FTSE for equities. Barclays Capital Aggregate for bonds. Markit iTraxx, SovX and CDX North America for credit.

Equilibrium returns π anchor the expectations and determine a vast majority of optimal portfolio allocations. This is an empirical finding.

Overweight and underweight the market

Findings and Analysis section of the project report must give a comparison as follows (see Table 6 in Guide by Thomas Idzorek):

- equilibrium allocations for the posterior vs. the prior $(\boldsymbol{\mu}_{BL} - \boldsymbol{\pi})$
- devised optimal allocations and market index weights $(\mathbf{w}^* - \tilde{\mathbf{w}})$

You should only hold something different than market weights if you are identifiably different than the market average investor.

John Cochrane

Views Distribution

Definition: a view is a statement on the market (returns) that can potentially clash with the reference market distribution.

BL operates with views expressed on expected returns $\boldsymbol{P\mu}$. The distribution of views is also Normal

$$\boldsymbol{P\mu} \sim N(\boldsymbol{v}, \boldsymbol{\Omega}) \tag{6}$$

Error

where $\boldsymbol{P}$ is a 'pick matrix' with indicator elements 0 or 1, and

$$\boldsymbol{\Omega} = \operatorname{diag}\left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}'\right) \tag{7}$$

is the uncertainty on the views.

Views Example

Within our example of IT, ES, CH, CA, US and DE indices, let's express two views

- the market in Spain will rise 12% (*absolute view*)
- the spread US-Germany will drop 10% (*relative view*)
  Germany outperforms the US

$$
\begin{array}{c}
\boldsymbol{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \\
\begin{array}{cccccc} & & & & \text{US} & \text{DE} \end{array}
\end{array}
$$

$$
\boldsymbol{v} = \begin{pmatrix} 12\% \\ -10\% \end{pmatrix} \quad \textit{annualised}
$$

The common notation is K views on N assets so, the pick matrix $\boldsymbol{P}$ has $K \times N$ dimensions.

Views based on volatility

It is a sensible approach to have a quantifiable base to the views in market volatility of the k*th* view, $k = 1, \dots, K$.

$\eta = 2$

$$\begin{aligned} \text{CI} &= \text{Expectation} \quad \pm \quad \eta \times \text{Standard Deviation} \\ v_k &= (\boldsymbol{P}\pi)_k + \eta_k \sqrt{(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}')_{k,k}} \end{aligned}$$

where η_k is selected from $\{-2, -1, 1, 2\}$ meaning 'very bearish', 'bearish', 'bullish', or 'very bullish' view respectively.

We incorporated variance (the second moment) into the views.

Updating Multivariate Density

Using multivariate Normal density $\boldsymbol{\mu} \sim N(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma})$ (matrix form *pdf*)

$$f_{\boldsymbol{\mu}}(\boldsymbol{\mu}) = \frac{|(\tau\boldsymbol{\Sigma})|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\pi})'(\tau\boldsymbol{\Sigma})^{-1}(\boldsymbol{\mu} - \boldsymbol{\pi})\right)$$

The views can be written $\boldsymbol{v} \overset{D}{=} \boldsymbol{P}\boldsymbol{\mu} + \epsilon$ where $\epsilon \sim N(0, \boldsymbol{\Omega})$.

$$\boldsymbol{V}|\boldsymbol{\mu} \sim N(\boldsymbol{P}\boldsymbol{\mu}, \boldsymbol{\Omega})$$

$$f_{\boldsymbol{V}|\boldsymbol{\mu}}(\boldsymbol{v}) = \frac{|\boldsymbol{\Omega}|^{-\frac{1}{2}}}{(2\pi)^{\frac{K}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{v} - \boldsymbol{P}\boldsymbol{\mu})'\boldsymbol{\Omega}^{-1}(\boldsymbol{v} - \boldsymbol{P}\boldsymbol{\mu})\right)$$

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$ Normal pdf

We are interested in **the posterior** distribution of returns updated with the views $\mu|\boldsymbol{V}$

Bayes Rule

$$\boxed{f_{\mu|\boldsymbol{V}}(\boldsymbol{\mu}) = \frac{f_{\mu,\boldsymbol{V}}(\boldsymbol{\mu},\boldsymbol{v})}{f_{\boldsymbol{V}}(\boldsymbol{v})}} = \frac{f_{\boldsymbol{V}|\mu}(\boldsymbol{v})\, f_{\mu}(\boldsymbol{\mu})}{\int f_{\boldsymbol{V}|\mu}(\boldsymbol{v})\, f_{\mu}(\boldsymbol{\mu}) d\boldsymbol{\mu}} \tag{8}$$

$$f_{V|\mu} = \frac{f_{\mu,V}(\mu,V)}{f_\mu(\mu)}$$

in lieu of unconditional pdf

Bayes Rule for updating the prior $f_\mu(\boldsymbol{\mu})$ with the views information is at the core of Black-Litterman derivation.

BL Solution

By substituting the respective Normal *pdfs* into the Bayes formula (8) and much tedious working, the solution for **the posterior**

$$f_{\boldsymbol{\mu}|\boldsymbol{v}}(\boldsymbol{\mu}) \sim N(\boldsymbol{\mu}_{BL}, \boldsymbol{\Sigma}^{\mu}_{BL})$$

Posterior

$$\boxed{\boldsymbol{\mu}_{BL} = \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}'\Omega^{-1}\boldsymbol{P}\right)^{-1} \left((\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \boldsymbol{P}'\Omega^{-1}\boldsymbol{v}\right)}$$

$$\boldsymbol{\Sigma}^{\mu}_{BL} = \left((\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}'\Omega^{-1}\boldsymbol{P}\right)^{-1} \quad \text{with } \boldsymbol{\Omega} = \operatorname{diag}\left(\boldsymbol{P}(\tau\boldsymbol{\Sigma})\boldsymbol{P}'\right)$$

where $\boldsymbol{\Sigma}^{\mu}_{BL} \sim O(\tau\boldsymbol{\Sigma})$ is the error on $\boldsymbol{\mu}_{BL}$ (not the covariance used in optimisation).

BL Solution for computation

Attillio Meucci (2010) suggested more **computationally stable** formulation (less matrix inversions), suitable for factors

for computation

$$\boxed{\boldsymbol{\mu}_{BL} = \boldsymbol{\pi} + \tau\boldsymbol{\Sigma}\boldsymbol{P}'\left(\tau\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}' + \boldsymbol{\Omega}\right)^{-1}(\boldsymbol{v} - \boldsymbol{P}\pi)} \tag{9}$$

$$\boldsymbol{\Sigma}_{BL} = (1+\tau)\boldsymbol{\Sigma} - \tau^2\boldsymbol{\Sigma}\boldsymbol{P}'\left(\tau\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{P}' + \boldsymbol{\Omega}\right)^{-1}\boldsymbol{P}\boldsymbol{\Sigma} \tag{10}$$

This formulation is more intuitive:

$(1+\tau)\boldsymbol{\Sigma}$ is the most distorted covariance of the posterior when views are not informative. Total covariance is now $\boldsymbol{\Sigma}_{BL} = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}^{\mu}_{BL}$.

Allocations

Obtain the allocations $\boldsymbol{w}^*$ by optimisation that uses the posterior (updated vector of expected returns $\boldsymbol{\mu_{BL}}$)

$$\underset{\boldsymbol{w}}{\operatorname{argmax}} \{\boldsymbol{w}'\boldsymbol{\mu_{BL}} - \lambda \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}\}$$

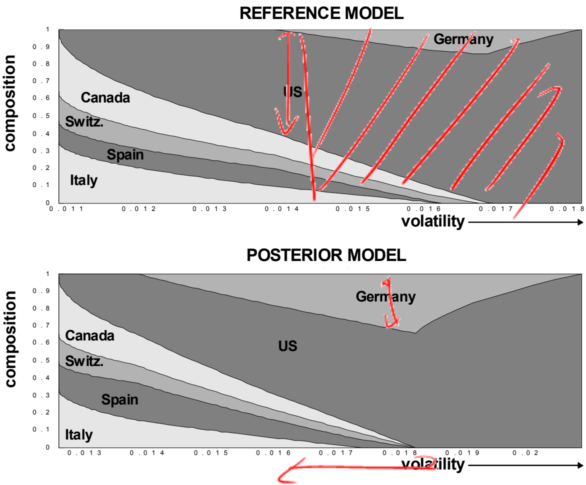
Min Var.

The explicit solution to optimisation problem without constraints is

original.

$$\boldsymbol{w}^* = \frac{1}{2\lambda}\boldsymbol{\Sigma}_{BL}^{-1}\boldsymbol{\mu_{BL}} \quad \text{or} \quad \boldsymbol{w}^* = \frac{1}{2\lambda}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{\mu_{BL}}$$

Optimise using a numerical routine because variance minimisation is not the only way to obtain optimal allocations.



From: *The Black Litterman Approach* guide by Attilio Meucci (2010).

Efficient Frontier

The above way of looking at the Efficient Frontier, via changing allocations, allows us to evaluate their robustness.

Expressing views reduces ‘the corner solution’, a case of investing all in the most earning/most risky asset (the US index).

Impact of the views. Compare allocations from the prior vs. posterior.

- allocation to Spain increases (up to certain volatility level);
- allocation to Germany increases (since US is expected to grow 10% less than Germany)

Risk aversion λ is a constant but **changing it affects allocations**. Study how allocations change for

- $\lambda = 0.01/2$ for to a near-Kelly investor (many concentrated bets)
- $\lambda = 2.24/2$ as suggested by BL (average investor, market)
- $\lambda = 6/2$ for a risk-averse investor (trustee).

Why $/2$? The unmodified values are for optimisation over $\frac{1}{2}\boldsymbol{\lambda} \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}$.

In these slides, $\boldsymbol{w}^* = \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\lambda}}\, \hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\, \mu_{\boldsymbol{BL}}$ so use the modified values.

Index of Satisfaction

We obtain optimal allocations by solving variance minimisation

$$\text{Portfolio Mean} - \lambda \times \text{Portfolio Variance}$$ − Port Kurtosis.

But that is only one method to define **a trade-off between two moments:** expected return $\mathbb{E}[X_w]$ and its standard deviation $\mathrm{SD}[X_w]$.

We can construct other trade-offs, the most common is

$$SR_w = \frac{\mathbb{E}[X_w]}{\mathrm{SD}[X_w]}$$

BL used the Sharpe Ratio 0.5 for the broad market index. Hedge funds aim for strategies to have the SR > 1.

Kinds of optimisation

Diagnose and choose at least two more kinds of optimisation *other* than an unconstrained variance minimisation.

- Sharpe Ratio maximisation
- VaR or Expected Shortfall (CVaR) minimisation
- Tracking error (TR) minimisation
- ... with appropriate constraints

... the pursuit of optimal allocation strategies has focused on fixing the excessive sensitivity to the input parameters.

Attilio Meucci

Alternative Optimisations

1. Sharpe Ratio maximisation (see next slide)

$$\underset{\boldsymbol{w}}{\operatorname{argmax}} \frac{\boldsymbol{w}'\boldsymbol{\mu_{BL}}}{\sqrt{\lambda\, \boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w}}} \qquad \underset{\boldsymbol{w}}{\operatorname{argmax}} \frac{\boldsymbol{w}'\boldsymbol{\mu_{BL}}}{\sqrt{\lambda\, \boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w}} \times \text{Factor}}$$

Tangency Portfolio allocations explicitly operate with excess returns and a budget constraint,

$$\underset{\boldsymbol{w}}{\operatorname{argmax}} \frac{\boldsymbol{w}'\boldsymbol{\mu_{BL}} - r_f}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}} = \frac{\mu_\Pi - r_f}{\sigma_\Pi}$$

$$\text{s.t.} \quad \boldsymbol{w}'\boldsymbol{1} = 1$$

$$\boldsymbol{w}^* = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\boldsymbol{1})}{\boldsymbol{1}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\boldsymbol{1})}$$

We leave as an exercise to find out the place of risk-aversion λ.

Alternative Optimisations (Cont.)

2. VaR as variance minimisation or minimisation with a constraint (risk budgeting)

$$\underset{\boldsymbol{w}}{\operatorname{argmax}} \left\{ \boldsymbol{w}'\mu_{\boldsymbol{BL}} - \sqrt{\lambda\, \boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w}} \times \text{Factor} \right\}$$

$$\underset{\boldsymbol{w}}{\operatorname{argmax}} \left\{ \boldsymbol{w}'\mu_{\boldsymbol{BL}} - \lambda\, \boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w} \right\} \quad \text{s.t. } \sqrt{\lambda \boldsymbol{w}'\boldsymbol{\Sigma}^{-1}\boldsymbol{w}} = \text{constant}/\text{Factor}$$

for VaR the factor is $\Phi^{-1}(1-c)$, for Expected Shortfall (CVaR) $\frac{1}{1-c}\int_0^{1-c} \Phi^{-1}(\gamma)d\gamma$.

Present results of both, analytical solution and numerical optimisation. Here, you can drop λ for simplicity.

Advanced Topics

These topics explain forefront, proper measures of diversification:

- **Effective Number of Bets**
- Factorisation (PCA, Minimum-Torsion Transform)

These topics generalise BL approach to conditioning by entropy pooling with exponential and kernel smoothing:

- **Entropy Pooling** with Flexible Views

  Replaces mean-variance optimisation. Particularly important is correlation stress-testing

- **Flexible Probabilities** (particularly, crisp conditioning)

Markowitz MPT Model Risk

Unconstrained mean-variance optimisation is of little value.

There is also **estimation risk**: the use of naive estimators of expected risk and return (i.e., historical sample mean) in optimization generates the Efficient Frontier that is far from robust.

Known Market

Let's generate a sample of $N = 10$ assets with ranked volatilities

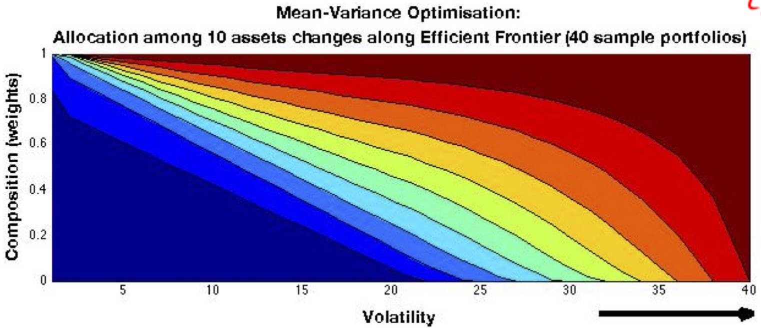
$$\boldsymbol{\sigma} = (5\%, 8.9\%, 12.8\%, 16.7\%, 20.6\%, 24.4\%, 28.3\%, 32.2\%, 36.1\%, 40\%)$$

Set correlation $\rho = 0.3$ for all assets, and introduce obtain expected returns as a proportion to each asset's volatility.

$$\boldsymbol{\mu} = (0.9\%, 1.6\%, 2.4\%, 3.3\%, 4.2\%, 5.2\%, 6.2\%, 7.25\%, 8.4\%, 9.55\%)$$

This is our **knowledge set** of the market. No estimation error.

$w^* = \frac{1}{2\lambda} \Sigma^{-1} \mu$



This is how allocations behave as acceptable portfolio risk increases: without constraints we end up investing in the most risky asset alone.

Sampling from mulativariate Normal

1. Using the **knowledge set** of our market (expected returns, covariance), we generate the multivariate Normal random values for time series with $T = 256$ observations

   This is simulated history of returns for correlated assets.

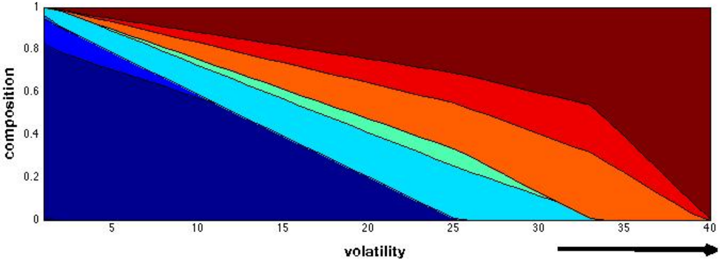
2. Then, estimate **sample covariance and mean** and optimise using these (instead of true parameters)

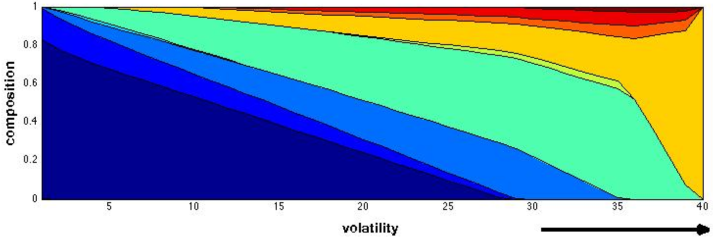
$$\mathbf{\Sigma} \to \hat{\mathbf{\Sigma}} \qquad \mu \to \hat{\mu}$$

Estimation error looks very small, less than half of a per cent.

$$\tau = \frac{1}{T} = \frac{1}{256} = 0.0039 \approx 0.4\%$$

Both frontiers come form **the same knowledge** about the market.





Before we conclude, the words of wisdom from Fischer Black:

"It is better to estimate a model than to test it. Best of all, though, is to explore a model."

r.diamond@cqf.com

Submission date is Monday,

25 July 2016.

Don’t Extend Your Luck!