Correlation Sensitivity and State Dependence

Dr. SiYi Zhou

CQF, Fitch Learning

2016 Jan

The main topics covered in this lecture

- The meaning of correlation, in particular linear correlation.
- Understand when to use and not to use linear correlation.
- Analyzing correlation sensitivity and state dependence in finite difference , examples using simple equity exotic and structured products.
- Correlation sensitivity and state dependence in credit risk modelling.
- An uncertain correlation model for Mezzanine tranche.

By the end of this lecture you will be able to

- Understand the concept of correlation in quantitative finance.
- Use and interpret linear correlation coefficient with care.
- Implement advanced correlation modeling in equity and credit derivatives.

Correlation Overview

Correlation plays a very important role in quantitative finance, and has a long history backdated to Markowitz portfolio theory in 1950s. It has since been used extensively in financial derivative market. In general it exists in almost all markets once there is a multi-dimensional problem. For example,

- An equity derivative with stochastic volatility.
- A 3 factor Libor Market Model with 20 underlying forward Libor rates.
- A CDX index with 125 reference names.

These situations all have something in common. They all require the input of parameter(s) measuring the relationship between two or more variables.

The Quant Tool Box

- So the quant searches around in his toolbox for some mathematical device that can be used to model the relationship between two or more variables. Unfortunately the number of such tools is so limited and in most cases it end up with relying on correlation.
- It's not that correlation is particularly brilliant, as we'll see, but it is easy to understand. Unfortunately, it is also easy to misunderstand.

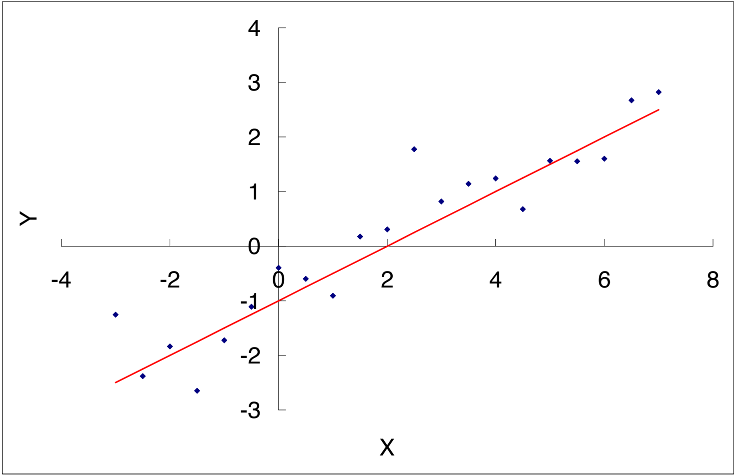
What's Correlation?

- In plain English, correlation measures to what degree two or more variables are related together.
- So far there is no mention to what kind of relationship exits between these variable.

Linear relationship

- A linear relationship occurs when variables are directly *proportional* to one another.
- It can be represented on a graph as a straight line.
- Or the value would "average out" to be a straight line.
- Mathematically, if the relationship between X and Y is linear, then the ratio of change in Y and change in X is constant(gradient,slope) $\forall(X, Y)$.

Linear Relationship



Linear Correlation

The most common way to quantify linear relationship is by *linear correlation coefficient* .

Definition

The linear correlation of two random variables (X, Y) is

$$\rho_{xy} = \frac{\mathrm{Cov}(x, y)}{\sqrt{\mathrm{var}(x)\mathrm{var}(y)}}$$

Intuitively, ρ_{xy} determines to what degree that the two variables are "proportional(linearly related)" to one another. In other words, how much the relationship can be approximated by a straight line.

Interpreting Linear Correlation Coefficient

The linear correlation coefficient is bounded between $[-1, 1]$.

- $\rho_{xy} = 1$.
  Perfect positive correlation, straight linear with positive slope.
- $\rho_{xy} = -1$.
  Perfect negative correlation, straight linear with negative slope.
- $\rho_{xy} = 0$.
  Uncorrelated, no linear relationship is detected.
- How to interpret, for example, $\rho_{xy} = 50\%$?
  In linear regression analysis, how good the regression line fits the data is measured by the *coefficient of determination* $R^2 = 25\%$, i.e., 25% of variance of Y is explained by X.. Mathematically it equals the square of correlation coefficient.

R2 is the correlation squared. so, if correlation is 50%, r2 is 25%. So, in a regression analysis People often interpret R^2 as measure of 'goodness of fit.' The higher the R2 - the better the model. or power of the exponential variable to explain the independent variable. The better part of X to explain Y. Higher correlation is better in our analysis, 25% of the variance of Y can be explained by X (independent variable). 25% of variance of Y explained by the linear model X.

Misuse of Linear Correlation

- In finance, linear correlation has been widely used for many years.
- In some markets such as equity and fixed-income its use is well developed and statistically sound.
- However, even if you totally understand the meaning of it, you still could misuse it for many different reasons.

Case 1: Non Linear Relationship

- You could make a mistake in the first place, that the relationship between X and Y isn't linear at all.
- Using linear correlation coefficient can be very misleading. For example, you calculated that $\rho_{xy} \approx 0$, in fact the true relationship should be quadratic.
- But how to figure out what kind of relationship exists between two or more variables? It is not an easy task!

How would you know if X & Y have a linear relationship or not?

Non-Linear Relationship

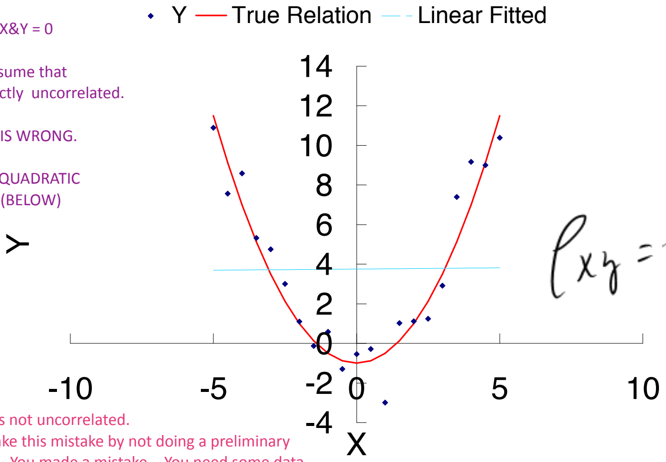
Graphed here, X & Y have some relationship.. It's certainly non-linear. It's quadratic relationship.
Having the Data. You can always calculate the Linear correlation coefficient. In example below, Correlation rho of X&Y = 0

Rho between X&Y = 0

that would assume that
X&Y are perfectly uncorrelated.

CLEARLY THIS IS WRONG.

THEY HAVE A QUADRATIC
Relationship (BELOW)



$\rho_{xy} = 0$

So, Clearly it's not uncorrelated.
You could make this mistake by not doing a preliminary data analysis. You made a mistake - You need some data analysis first. Theory, or plotting, etc.

Case 2: Non-normal Distribution

Second case: You can misuse the correlation is if you calculate the correlation for variables which are NOT elliptic distribution;

The linear correlation coefficient is only suitable for elliptical distribution, such as Normal and Student's t distribution. However it is inappropriate for other distributions.

ELLIPTIC distribution: meaning it's either a NORMAL distribution or a Student T's distribution, or some other distribution that's similar. I.E. Not symmetrical or some other weird distribution.

A simple example is Log-normal Random Walk.

$$S(T) = \mathcal{LN}\left[\log S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma^2 T\right]$$

Example - stock price = S(T) is a lognormal random walk

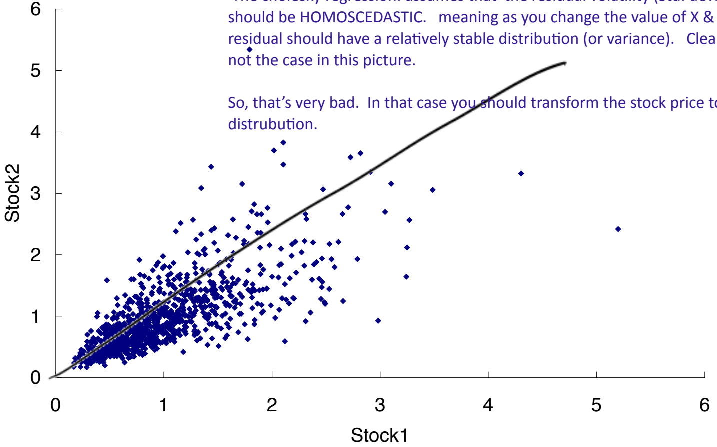
So, in theory, we assume stock price follows lognormal random walk. After taking LOG. It then becomes NORMAL. but regionally it's a LOGNORMAL distribution. And it's highly skewed distribution.

Bi-variate Log-normal

Random simulation (stock 1 & stock 2). Then, having the data you can always calculate the linear correlation coefficient. There's a connection between the correlation and the regression. In the bivariate case, if you calculate the correlation, it implies you're trying to fit a straight line on the picture. A regression line. This is not desired feature. If you look below. WHY?

The cholesky regression. assumes that the residual volatility (Std. deviation), should be HOMOSCEDASTIC. meaning as you change the value of X & Y. the residual should have a relatively stable distribution (or variance). Clearly. this is not the case in this picture.

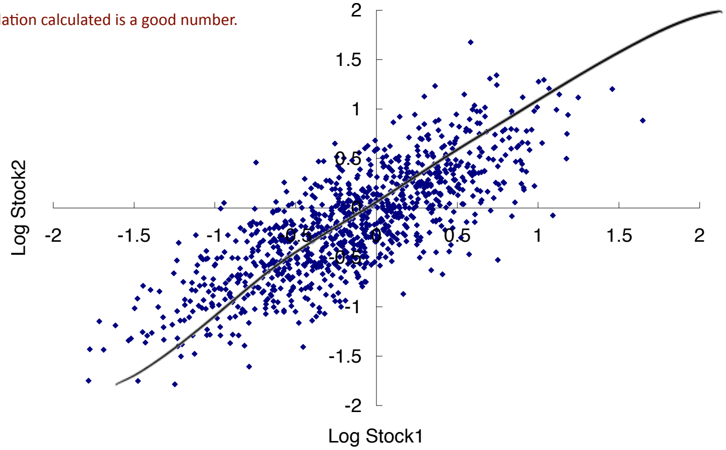
So, that's very bad. In that case you should transform the stock price to normal distrubution.



It's is more appropriate work with log stock price.

To convert the LOGNORMAL stock price into NORMAL. simply by taking LOG of the Stock prices. Which gives you example 2 which is more ideal to work with. In this case, this regression (below) is HOMOSCEDADISTIK.

And the correlation calculated is a good number.



Linear Correlation of Log-normal

Mathematically, what happens is .

(mathematics) Sufficiently operationalizable or useful to allow a mathematical calculation to proceed toward a solution.
(computer science) Of a decision problem, algorithmically solvable fast enough to be practically relevant, typically in polynomial time.

As the moment function of Log-normal is tractable, one can use it to calculate the linear correlation coefficient . Suppose

X1 & X2 are Normal distribution

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix} \right),$$

and

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \exp \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$

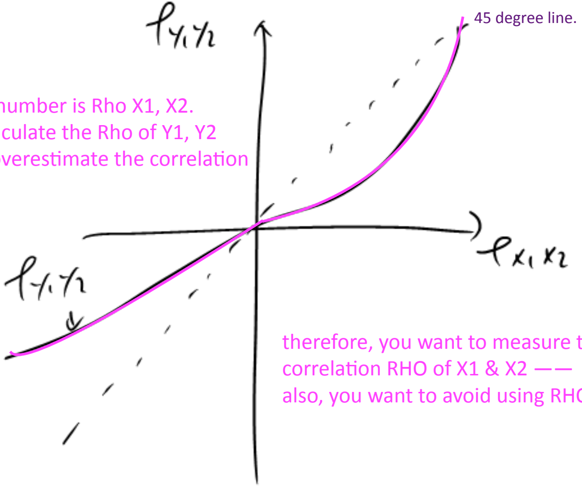
So

$$\rho_{y_1 y_2} = \frac{e^{\sigma_1\sigma_2\rho} - 1}{\sqrt{\left(e^{\sigma_1^2} - 1\right)\left(e^{\sigma_2^2} - 1\right)}}. \neq \rho_{x_1 x_2}$$

if $\sigma_1 = \sigma_2 = 1$

Sigma 1 = Sigma 2 = 1

$$\rho_{Y_1 Y_2} = \frac{e^{\rho} - 1}{e - 1}$$



the correct number is Rho X1, X2.
so, if you calculate the Rho of Y1, Y2
you would overestimate the correlation

therefore, you want to measure the
correlation RHO of X1 & X2 —— and
also, you want to avoid using RHO Y1/Y2

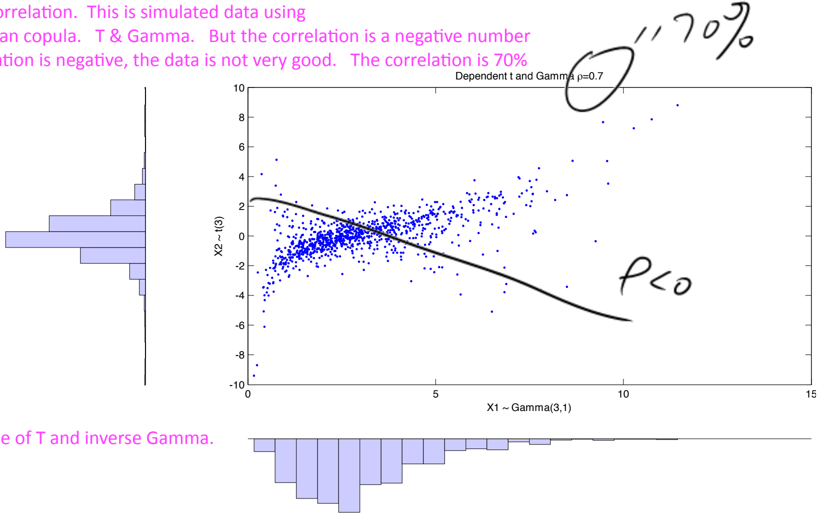
Another Example

- The previous example in fact wasn't that bad since the linear correlation coefficient of Log-normal was not too far away from that of normal, and it is analytically tractable. However there are many other situations in which the linear correlation coefficient could easily distort the true relationship.
- For example, dependence between Student's t and Gamma distribution. If one fits a linear relationship then it can result in a misunderstanding.

Dependent Gamma and Students' t

If fits a straight line, which direction is correct?

70% correlation. This is simulated data using
Gaussian copula. T & Gamma. But the correlation is a negative number
correlation is negative, the data is not very good. The correlation is 70%



Inverse of T and inverse Gamma.

Rank Correlation

Ways to get around problems, instead of LINEAR CORRELATION, we can use RANK CORRELATION. The good thing about RANK rho is that it is invariant!

In this situation rank correlation outperforms linear correlation, nonetheless it is invariant subject to non-linear monotonic transformations.

Roughly speaking, rank correlation measures the degree to which large or small values of one random variable associate with large or small values of another. However, unlike the linear correlation coefficient , they measure the association only in terms of ranks. As a consequence, the rank correlation is preserved under any monotonic transformation. *Kendall's tau* or *Spearman's rho* are more useful in describing the dependence between random variables, because they are invariant to the choice of marginal distributions.

BASICALLY. It is the correlation based on RANKS. We don't need to worry about distance between two variables, but we care about RANK.

However in general even if marginal distributions and correlation are given, dependence(relationship) is not uniquely determined. Which implies knowing correlation can not fully depict the picture of dependence.

Spearman's rho

Gaussian Copula is still linear, so, you can't use the RANK correlation in that distribution.

Spearman's rho is closely linked to the concept of linear correlation. It is defined as the linear correlation of associated CDFs.

Definition

The Spearman's rho of two random variables is

$$\rho_S = \frac{\mathrm{Cov}(F_x(x), F_y(y))}{\sqrt{F_x(x)F_y(y)}} = 12\mathbb{E}\left(F_x(x)F_y(y)\right) - 3$$

Where F_x and F_y are marginal distribution of x and y respectively.

Spearman's rho is invariant with respect to monotonic transformations.

Kendall's tau

Kendall's tau is based on the ideal of *concordance* and *dis-concordance.* Suppose pairs (x, y) and (x', y') are drawn from a joint distribution $F_2(x, y)$, they are concordant if $x > x'$ and $y > y'$; they are dis-concordant if $x > x'$ and $y < y'$ or vice versa.

Definition

Kendall's tau is the difference between probability of concordance and dis-concordance

$$\begin{aligned} \rho_K &= \Pr\left[(x - x')(y - y') > 0\right] - \Pr\left[(x - x')(y - y') < 0\right] \\ &= 4\mathbb{E}\left(F_2(x, y)\right) - 1 \end{aligned}$$

Like Spearman's rho, it is invariant with respect to monotonic transformations.

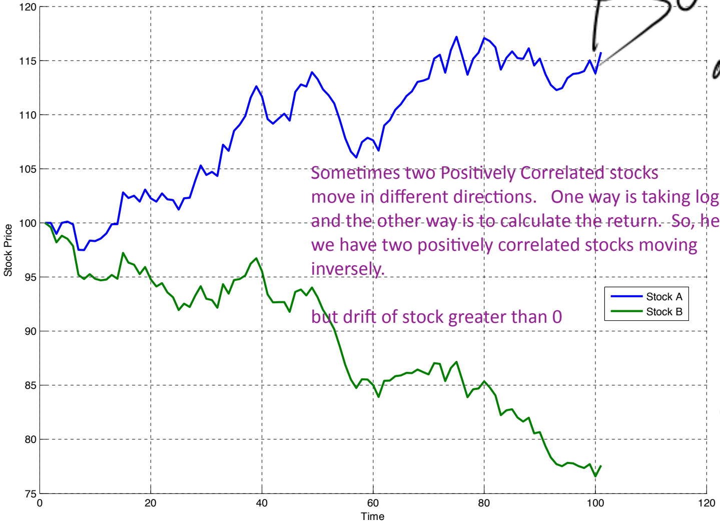
Case 3: Intuition and Timescales

- When we think of two assets that are highly correlated then we are tempted to think of them both moving along side by side almost.
- Surely if one is growing rapidly then so must the other?
- Similarly, if two assets are highly negatively correlated then they go in opposite direction?

What is the correlation for this picture?

LOW CORRELATION

$\rho > 0$



$drift_1 > 0$

Sometimes two Positively Correlated stocks move in different directions. One way is taking log and the other way is to calculate the return. So, here we have two positively correlated stocks moving inversely.

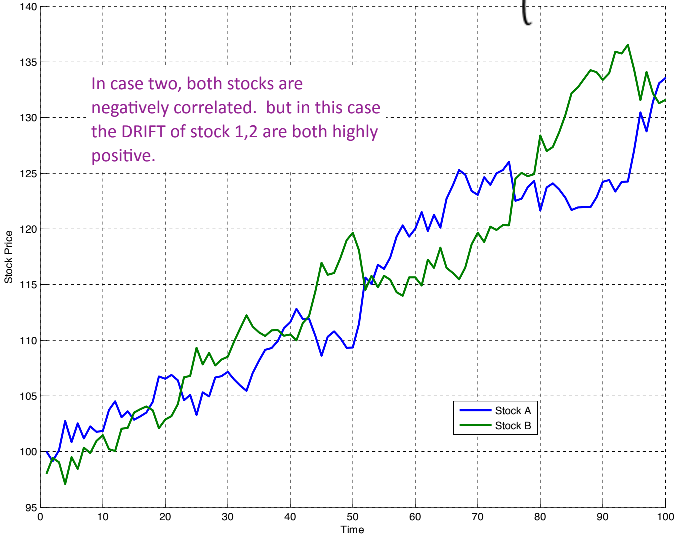
but drift of stock greater than 0

$drift_2 < 0$

And This Picture?

$\rho < 0 \quad drift_{1,2} > 0$

HIGH CORRELATION



In case two, both stocks are negatively correlated. but in this case the DRIFT of stock 1,2 are both highly positive.

What Went Wrong?

If you're a hedge fund, select the stock by clever drift, vs. correlations. CAPM

Our intuition tells us that the correlation is very high in the first picture and low in the second picture, but it isn't true.

Correlation is about what happens at the smallest, technically infinitesimal if think of Itô calculus, time scale. it is not about the "big picture" direction.

For example, if we are interested in how assets behave over some finite time horizon in the Markowitz portfolio theory framework, then we still use correlation even though we typically don't care about short timescales(at least in theory). Really we ought to be modelling drift better, and any longer-term interaction between two assets might be represented by a clever drift term(in the stochastic differential equation sense).

If Hedged

However, if we are hedging an option that depends on two or more underlying assets then, conversely, we don't care about direction (because we are hedging), only about dynamics over the hedging timescale.

The use of correlation may then be easier to justify.

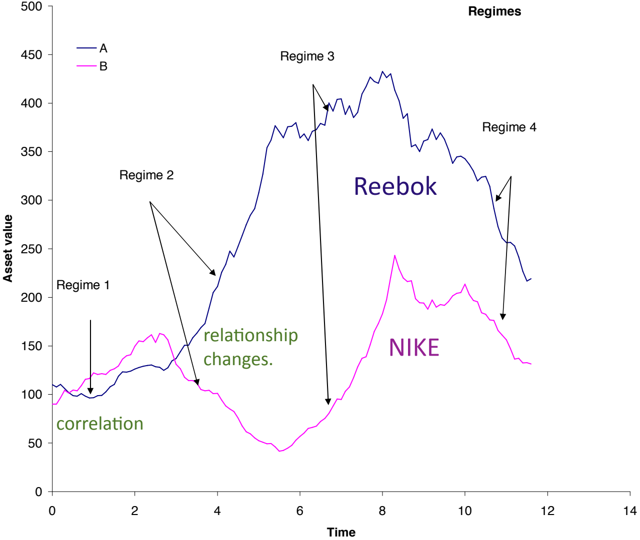
Case4: Stability

In a real life situation, the relationship between two or more variables is hardly stable. It is difficult to model interesting and realistic dynamic using a simple concept like correlation.

In the figure below is plotted the share prices against time of two makers of running shoes.

Correlation is a static constant number, relationship is dynamic. not static.

Share prices of two shoe makers



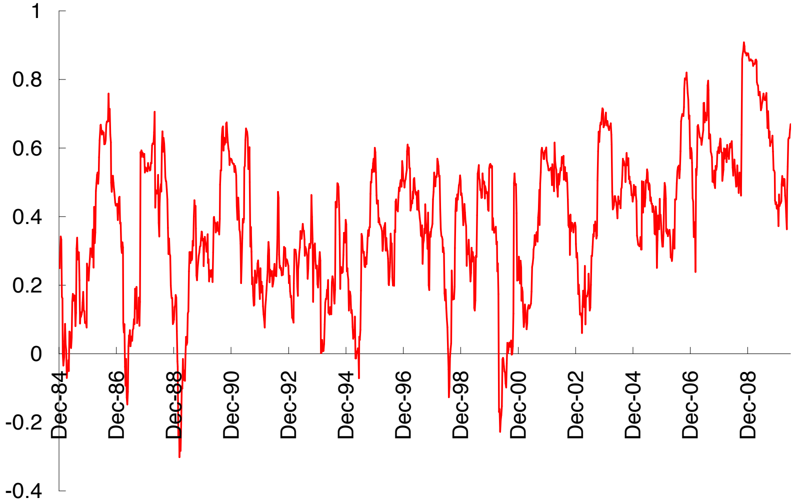
Competition

1. In Regime 1 both shares are rising, the companies are seen as great investments, who doesn't want to go jogging and get healthy? See how both shares rise together.
2. In Regime 2 company A has just employed a hotshot basketball player to advertise their wares, stock A grows even faster but the greater competition causes share B to fall.
3. News Flash! Top basketball player in sex and drugs scandal! Parents stop buying brand A and switch to brand B. The competition is on the other foot so to speak. Stock A now not doing so well and stock B recovers, they are taking away income from A.
4. Another News Flash! Top advocate of jogging drops dead. . . while jogging! Sales of both brands fall together.

Why different correlation?

- There are many ways to calculate correlation, One can always get a number to represent correlation, but often these numbers varies.
- It may caused by structure change of the true relationship.
- Or different data and data window that have been used in the calculation.
- It also depends on what correlation estimation model has been used.

Rolling 6-month Correlation: SP500 vs Nikkie250



Case 5: Causality

Correlation means, it's equal. Bilateral. Bi-directional relationship where one asset does not dominate the other. NIKE doesn't outperform REEBOK. If this is the case, then that relationship is referred to as CAUSALITY RELATIONSHIP.

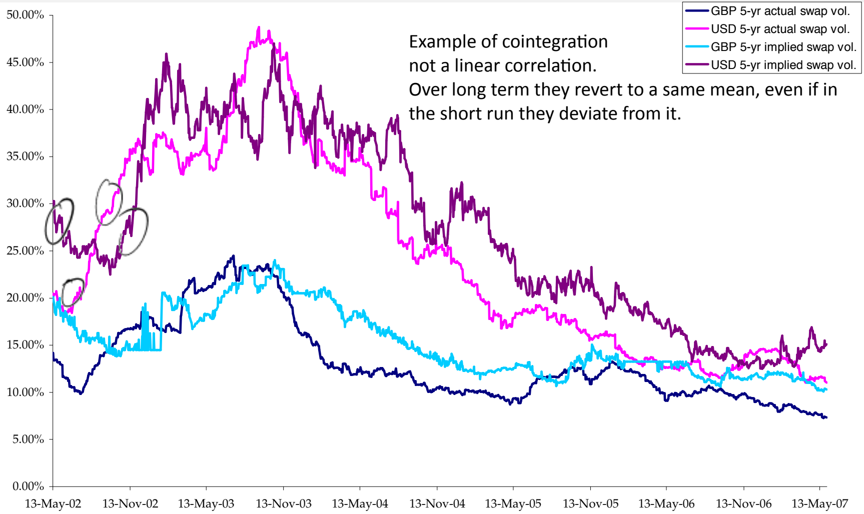
Previous graph shows the historic correlation between these US and Japanese stock market which moves randomly between -0.2 to 0.9. However, when the correlation was high, do we really believe in this stronger relation?

In practice, there will also be some delay between trading in one stock and trading in the other.

If company A is much bigger and better known than company B then company's A's stocks will trade first and there may be a short delay until people "join the dots" and think that stock B might be affected as well.

This is causality, and not something that correlation models.

Implied vol vs actual vol



Case 6: Co-integration

- Two series can't wonder off in opposite directions for very long without coming back to a mean distance eventually.
- But it doesn't mean that on a daily basis the two prices have to move in synchrony at all.
- Technically speaking, a set of processes is defined as co-integrated if a linear combination of them is stationary.
- Many time series are non-stationary, such random walk, but they do move together overtime.
- This implies that these series are bound by some relationship in the long run.
- Modeling this long term or equilibrium relationship by correlation is inappropriate.

Summary

What should we do?

When modelling correlation, you should ask the following questions

- Is the relationship linear?
- Are the distributions normal?
- What is the time horizon? If long term, should take trend into consideration?
- How stable is the relationship?
- Is there any causality or co-integration?

Good News and Bad News

- As you can see that the modelling relationship even between two entities can be so fascinating and challenging, and it can be modelled using all sorts of interesting mathematics. One thing is for sure and that is such relationship are certainly not captured by a correlation of 0.6.
- If you like modelling then it is great news, you have a blank canvas on which to express your ideals.
- But if you have to work with correlation on a day-to-day basis it is definitely bad news.

In the rest of this lecture we will be addressing how to cope with this problem by means of uncertain correlation.

An Exotic Option

Contract: Two-asset, worst-of, knockout option

- There are two underlyings.
- Payoff is $\min(S_1, S_2)$.
- Knocks out if ever $S_1 \geq X_1$ or $S_2 \geq X_2$.

finite difference method

Classical pricing of this is via partial differential equation :

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 V}{\partial S_1^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 V}{\partial S_1\, \partial S_2} + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 V}{\partial S_2^2}$$

$$+ rS_1 \frac{\partial V}{\partial S_1} + rS_2 \frac{\partial V}{\partial S_2} - rV = 0$$

$\downarrow$

ρ?

(assuming lognormal, correlated random walks), with

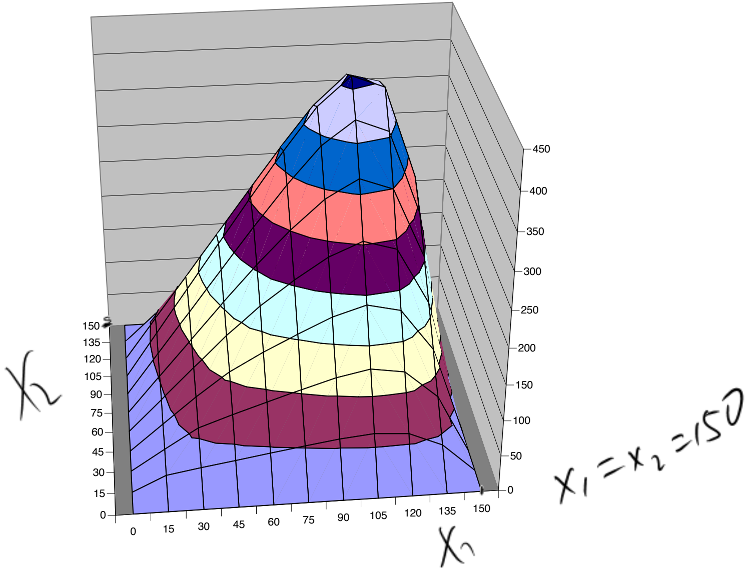
$$V(S_1, S_2, T) = \min(S_1, S_2),$$

and

$$V(X_1, S_2, t) = V(S_1, X_2, t) = 0.$$

$V(S_1 = x_1, S_2, t) = V(S_1, S_2 = X_2, t) = 0.$

Value of a two-asset, worst-of, knockout option



Sensitivity Analysis

Traditionally as part of the risk management of such a contract you would look at the sensitivity of the value to the parameters.

For example, in this plot the correlation was 0.8, one might see how the price varies as ρ varies from 0.3 to 1.0, say. This is like calculating $\frac{\partial V}{\partial \rho}$.

But this will only be totally meaningful if the value of the contract is monotonic in this parameter.

Cross Gamma

By looking at the governing PDE we see that correlation multiplies a term of the form

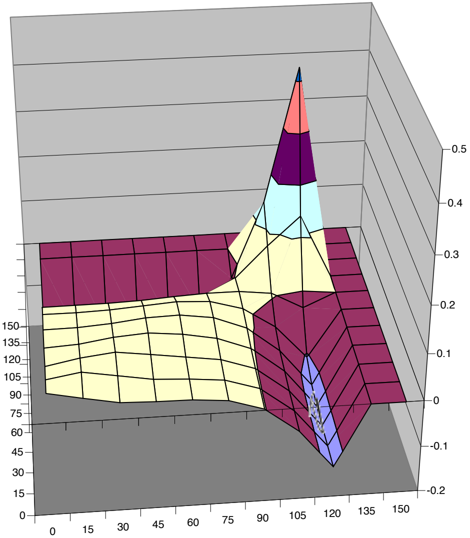
$$\frac{\partial^2 V}{\partial S_1 \, \partial S_2}.$$

If this term is single signed then the above analysis would be useful.

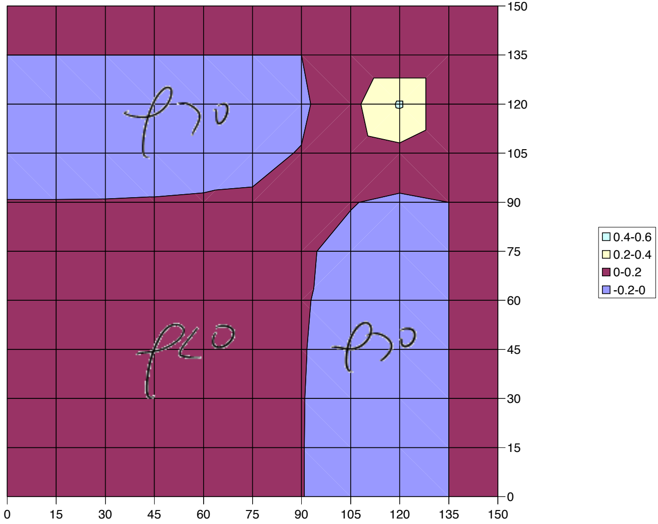
If it is not single signed then the analysis can be very misleading and therefore dangerous!

So let's calculate this term!

Cross gamma of a two-asset, worst-of, knockout option



Contour plot of cross gamma



An Equity Structured Product

Let's now turn to the valuation of an equity structured product which mimics the payoff of a CDO structure.

We will use a factor copula model to address a credit product of CDO shortly in later section.

Assumptions

- There are three underlying assets, $i = 1, 2, 3$.
- Each S_i follows a lognormal random walk. The random walks need not be lognormal since we are solving numerically, this choice is more for tradition than anything.
- Volatilities are σ_i, constants.
- Drifts will all be the interest rate r. This choice can be interpreted as either a) risk neutral, perhaps because the underlyings are traded, b) parsimony because we don't want to be sidetracked by extra parameters or c) irrelevant perhaps because the time horizon is too short or because other parameters and the model itself are far from perfect. This is a detail which we won't be concerned with.
- Correlations are ρ_{ij}. These parameters will be the focus of our attention.

Valuation

Being ambiguous to what the payoff function is for the time being, we know that the value of this product $V(S_1, S_2, S_3, t)$ satisfies the partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i\,\partial S_j} + \sum_{i=1}^{3} r S_i \frac{\partial V}{\partial S_i} - rV = 0.$$

Tranched Product

To mimic the behavior of a CDO we assume there 3 tranches in this structured product, with unit of notional principle for each tranche.

We also assume that S_i would be knocked-out (default) if it falls below a threshold A_i at the maturity. We could also consider the case of default being if *ever* the variable falls below that level *before* maturity.

This helps us to define the final and the boundary conditions for each tranche.

Final and boundary conditions

- The Senior Tranche will thus have final condition

$$V(S_1, S_2, S_3, T) = \begin{cases} 0, & \text{if all } S_i \text{ are below their threshold;} \\ 1, & \text{else.} \end{cases}$$

Senior tranche pays 0 if all of the lower tranches default. Otherwise payoff is 1

- The Mezzanine Tranche will have final condition

$$V(S_1, S_2, S_3, T) = \begin{cases} 0, & \text{if any two of } S_i \text{ are below their threshold;} \\ 1, & \text{else.} \end{cases}$$

Mezzanine tranche pays 0 if all of the lower tranches default. otherwise 1

- And the Equity Tranche will have final condition

$$V(S_1, S_2, S_3, T) = \begin{cases} 0, & \text{if any one of } S_i \text{ are below their threshold;} \\ 1, & \text{else.} \end{cases}$$

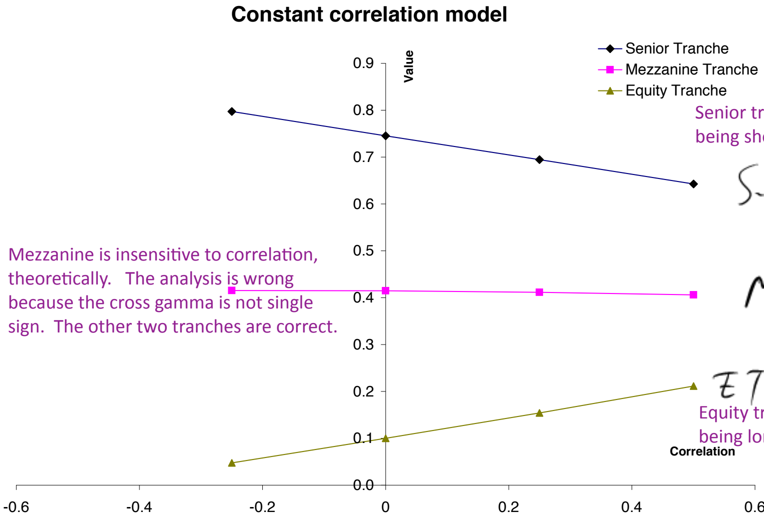
These problems are easily solved by finite differences.

Results for constant correlation

Let's see how changes in correlation affect the value of the three tranches. To keep things simple we will have all volatilities the same, $\sigma_i = 0.2$, the interest rate is $r = 0.05$, maturity $T = 3$, all thresholds the same $A_i = 1$.

For the figure below we assume that all three correlation, ρ_{12}, ρ_{23} and ρ_{31} are the same, and then find the values of each tranche as a function of that correlation.

The results are qualitatively similar to those from classical copula models.



Senior tranche is like being short Correlation

S.T.

MT

ET

Mezzanine is insensitive to correlation, theoretically. The analysis is wrong because the cross gamma is not single sign. The other two tranches are correct.

Equity tranche is like being long Correlation

Correlation Risk

Traditionally one would say that the Mezzanine Tranche exhibits little sensitivity to the chosen correlation. As correlation varies from -0.25 to +0.5 the value of this tranche hardly changes at all, falling from 0.415 to 0.406, a range of 0.009.

However, we know that insensitivity to a *constant* parameter can be the result of non-single-signed greeks and can therefore hide serious risks. This is precisely the problem here. In the Mezzanine Tranche the cross gamma terms may not be single signed.

Non-single-signed cross gamma

The cross gamma terms $\frac{\partial^2 V}{\partial S_i \, \partial S_j}$ for $i \neq j$ can change sign throughout the S_1-S_2-S_3-t domain. When this is the case increasing correlations uniformly will increase V in part of the domain while decreasing it elsewhere.

There will then be regions in which the value hardly changes at all, giving the impression of insensitivity to the parameter. The same happens when the correlation is everywhere decreased.

state-dependent correlation

To evaluate the real dependence on a parameter such as correlation one first increases the parameter where the cross gamma is positive and decreases it where it is negative. This will be the 'best of both worlds' giving the best-case value. Then one does the opposite, decreasing the parameter value where the cross gamma is positive and increasing it where it is negative. This gives the worst-case value.

This models a **state-dependent correlation**, but one where the state is chosen during the (numerical) solution to give the best or worst possible values.

Complications

In the problem we are addressing here there is an added complication because there are three correlations, ρ_{12}, ρ_{23} and ρ_{31}, each of which multiply a different cross gamma in the governing partial differential equation. We have two possible ways of dealing with this.

Not-quite-the-best/worst case

First assume that all correlations are the same, ρ, and then choose the value of ρ depending on the sign of the combination of cross gammas which it multiplies.

If we are interested in finding the sensitivity of value to a range of correlations $\rho^- \leq \rho \leq \rho^+$ then we choose ρ to be either ρ^- or ρ^+ depending on the sign of

$$\boxed{\sum_{i=1}^{3} \sum_{j=1, j\neq i}^{3} \sigma_i \sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \, \partial S_j}.}$$

Genuine best/worst case

A better 'best' and a worse 'worst' than the above can be found by treating each correlation independently.

So we would choose each ρ_{ij} to be either ρ^- or ρ^+ depending on the sign of each $\frac{\partial^2 V}{\partial S_i\, \partial S_j}$. (It is not quite this simple because we also have a constraint on the resulting correlation 'matrix' that it be positive definite everywhere.) In the numerical results that follow it is these extreme cases we have considered.

$\rho < \rho^+$

ρ^-

rho+ upper boundary

rho- lower boundary

Best case Scenario

when cross gamma is positive, we use high correlation

$$\text{best} \begin{cases} \frac{\partial^2 V}{\partial S_1 \partial S_2} > 0 & \rho^+ \\ \frac{\partial^2 V}{\partial S_1 \partial S_2} < 0 & \rho^- \end{cases}$$

when cross gamma is low, we use low correlation

when cross gamma is low, we use Low correlation

$$\text{worst} \begin{cases} \frac{\partial^2 V}{\partial S_1 \partial S_2} > 0 & \rho^- \\ \frac{\partial^2 V}{\partial S_1 \partial S_2} < 0 & \rho^+ \end{cases}$$

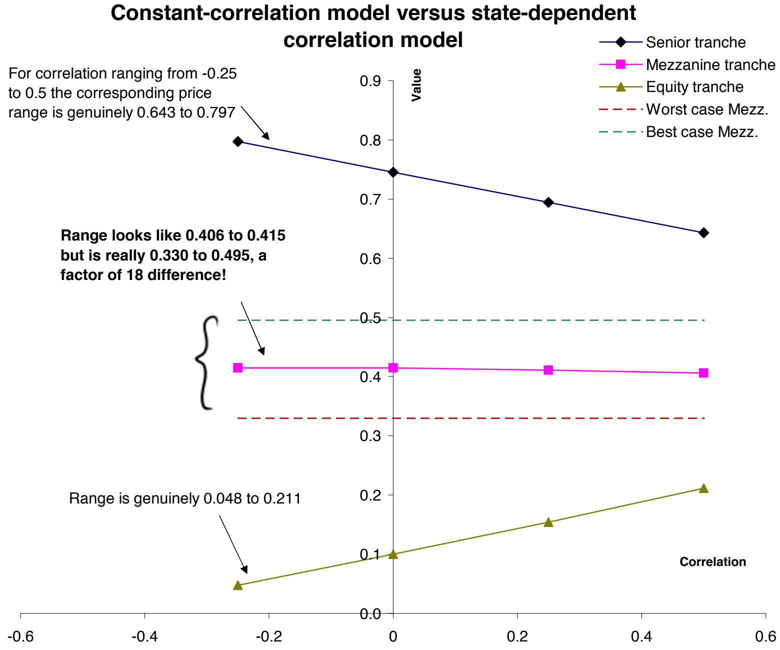
when cross gamma is low, we use high correlation

Results for best and worst cases

Once one has a finite-difference method that will solve the linear case of constant correlation then it is easy to modify it to the non-linear best and worst cases. (Note that one has to be careful with implicit finite-difference methods because the nonlinearity can lead to the wrong solution.)

This is why we have chosen to value using finite-differences rather than Monte Carlo since the latter is unsuited to such non-linear problems. But it is also the reason why we are restricted to having few underlyings!

Solving the same problem as before for best and worst cases we get results shown here.


Constant-correlation model versus state-dependent correlation model
Senior tranche
Mezzanine tranche
Equity tranche
Worst case Mezz.
Best case Mezz.
For correlation ranging from -0.25 to 0.5 the corresponding price range is genuinely 0.643 to 0.797
Range looks like 0.406 to 0.415 but is really 0.330 to 0.495, a factor of 18 difference!
Range is genuinely 0.048 to 0.211
Value
Correlation
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
-0.6
-0.4
-0.2
0
0.2
0.4
0.6

Senior and Equity Tranche Result

- For the Senior Tranche the naive, constant correlation analysis gives a range of values from 0.643 to 0.797. The more sophisticated best/worst cases are exactly the same. This is because all cross gammas are singled signed. They are actually all negative everywhere as seen by the value decreasing as the correlation rises.
- For the Equity Tranche the naive, constant correlation analysis gives a range of values from 0.048 to 0.211. The more sophisticated best/worst cases are the same. This is again because all cross gammas are singled signed. This time they are all positive everywhere as seen by the value increasing as the correlation rises.

Therefore classical sensitivity analysis will give correct results for both senior and equity tranche.

Mezzanine Tranche Result

The interesting case, and the one where the danger lies, is the Mezzanine Tranche.

The naive, constant correlation analysis gives a range of values from **0.406** to 0.**415**.

The more sophisticated best/worst cases are totally different however. The range is then **0.330** to **0.495**. This is because cross gammas are *not* singled signed, and this is why the Mezzanine Tranche appears very flat in the traditional sensitivity analysis.

But the true range of values is **18** *times as much, and this is why the Mezzanine Tranche is dangerous.*

Homogenous portfolio

A portfolio is homogenous if each name in the portfolio shares the same default probability, recovery rate, correlation and notional principal.

The loss distribution for homogenous portfolio has closed form solution.

Default probability

NORMAL COPULA = 1 factor COPULA CDO

Z is a common asset

Recall the one factor normal copula model, the asset value of a firm i is driven by

Asset i

Each asset has it's asset value normalized

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

Risk factor Z

idiosyncratic risk factor epsilon i

We can drop subscript off since the reference portfolio is homogenous. The default probability for any i is

F(t) = Cum Dist Function

$$F(t) = \Pr[\tau < t] = \Pr[A < d(t)],$$

Prob default

asset value is smaller than default threshold d(t)

where $d(t)$ is default threshold.

Because of the assumption of normality across all i

$$\boxed{F(t) = \Phi\left(d(t)\right),}$$

Because A is normal.
the probability of default is PHI (d(t))

from where the default threshold can be derived

$$d(t) = \Phi^{-1}(F(t)).$$

Conditional default probability

Conditional on common factor Z, the default probability will be

$$
\begin{aligned}
F(t|Z) &= \Pr\left[\overbrace{wZ + \sqrt{1-w^2}\,\varepsilon}^{A} < d(t) \mid Z\right] \\
&= \Pr\left[\varepsilon < \frac{d(t) - w\,Z}{\sqrt{1-w^2}} \mid Z\right] \\
&= \Phi\left(\frac{d(t) - w\,Z}{\sqrt{1-w^2}}\right).
\end{aligned}
$$

w^2 is correlation

Define $\rho = w^2$, then

$$F(t|Z) = \Phi\left(\frac{d(t) - \sqrt{\rho}\,Z}{\sqrt{1-\rho}}\right). \tag{2}$$

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

$$A_j = w_j Z + \cancel{\sqrt{1-w_j^2}\,\varepsilon_j}$$

$$\rho_{ij} = \mathrm{Cov}(A_i, A_j) = w_i w_j \,\mathrm{Var}(Z)$$

$$= w_i w_j$$

$$w_i = w \quad \forall i$$

$$\rho_{ij} = w^2 = \rho$$

$$F(t \mid Z) = \Phi\left(\frac{d - \sqrt{\rho}\, Z}{\sqrt{1-\rho}} \right)$$

Conditional independence

Conditioning on Z, the default of an obligor will be independent of any other obligor. As a result, the number of default conditional on Z follows binomial distribution.

Let N the the size of the reference portfolio, and K be the number of default occurred before time t, so

$$K \sim \mathbf{Binomial}\left(N, F(t|Z)\right),$$

hence

$$\mathbf{Pr}[K = k|Z] = \binom{N}{k} F(t|Z)^k \left(1 - F(t|Z)\right)^{N-k}. \tag{3}$$

$$A_i = \boxed{wZ} + \sqrt{1-w^2}\,\varepsilon_i$$

$$A_j = \boxed{wZ} + \sqrt{1-w^2}\,\varepsilon_j$$

Conditional on Z

A_i & A_j are independant

because ε_i & ε_j are independent.

Conditional PD $F(t|Z)$ is same for all i

$$K \mid Z \sim \text{Binomial}\left(N, F(t|Z)\right)$$

Unconditional Distribution

The unconditional distribution function of K can be derived by iterated expectation,

$E(X) = E(E(X|Y))$

$$
\begin{aligned}
\mathbf{Pr}[K=k] &= \mathbb{E}\left(\mathbf{Pr}[K=k|Z]\right) \qquad (4)\\
&= \int_{-\infty}^{\infty} \binom{N}{k} F(t|Z)^k \left(1-F(t|Z)\right)^{N-k} \, d\,\Phi(Z).
\end{aligned}
$$

Note, this integral must be solved numerically.

$Z \sim N(0,1)$

Φ CDF

Loss distribution

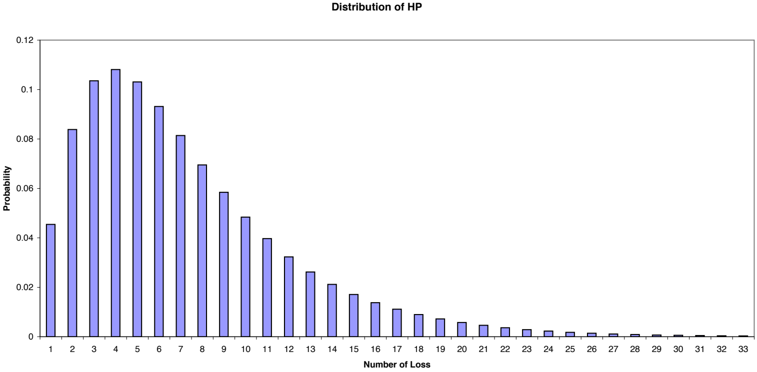
Equation (**??**) is the distribution function for the number of default, we need to translate it to portfolio loss distribution.

Denote $L(t)$ the percentage of loss for the reference portfolio by time t, and θ is the recovery rate. So when there is K defaults in the portfolio, the loss will be (assume unit NP):

$$L(t) = \frac{K(1-\theta)}{N}.$$

Thus the loss distribution function is

$$\mathbf{Pr}\left[L(t) = \frac{k(1-\theta)}{N}\right] = \mathbf{Pr}[K = k].$$


Distribution of HP
Probability
0.12
0.1
0.08
0.06
0.04
0.02
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
Number of Loss

Large Homogenous Portfolio

If the number of reference names N of the underlying portfolio becomes reasonably large (as it effectively is for a typical CDO), the distribution function for the portfolio loss can be further simplified. This allows for computation of the portfolio loss distribution without resorting to either Monte Carlo simulation nor numerical schemes to solve integral (**??**).

Fraction of default

Define the fraction of default as

$$Y = \frac{K}{N}.$$

According to CLT, conditional on Z, Y is approximately normal with

$$\begin{aligned} \mathbb{E}[Y|Z] &= \frac{\mathbb{E}(K|Z)}{N} = F(t|Z) \\ \mathbf{Var}(Y|Z) &= \frac{\mathbf{Var}(K|Z)}{N^2} = \frac{F(t|Z)(1-F(t|Z))}{N}. \end{aligned}$$

So

$$\lim_{N\to\infty} Y = F(t|Z).$$

Loss distribution for LHP

Define the cumulative distribution function of $Y(z)$ is G, then

$$
\begin{aligned}
G(y) &= \mathbf{Pr}[Y(z) \le y] = \mathbf{Pr}[F(t|Z) \le y] \\
&= \mathbf{Pr}\left[\Phi\left(\frac{d(t) - \sqrt{\rho}\,Z}{\sqrt{1-\rho}}\right) \le y\right] \\
&= \mathbf{Pr}\left[Z \le \frac{\sqrt{1-\rho}\Phi^{-1}(y) - d(t)}{\sqrt{\rho}}\right] \\
&= \Phi\left(\frac{\sqrt{1-\rho}\Phi^{-1}(y) - d(t)}{\sqrt{\rho}}\right).
\end{aligned}
$$

Loss distribution for LHP

To derive ultimate loss distribution use the fact that percentage of loss equals to fraction of loss multiply by loss given default, i.e.,

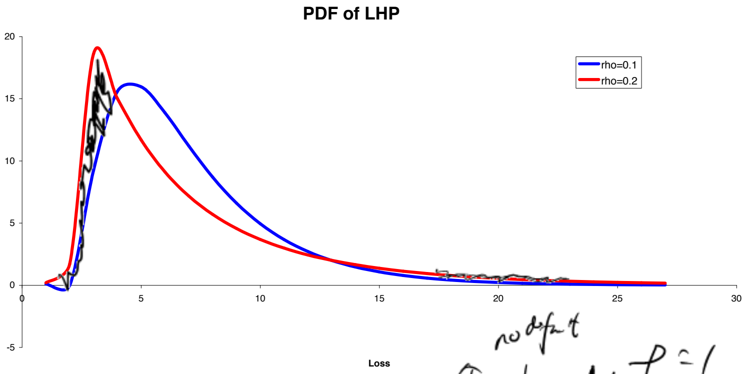
$$\Rightarrow \quad L(t) = Y(1-\theta),$$

So the distribution function of $L(t)$ will be

$$\mathbf{Pr}[L(t) \leq l] = G\left(\frac{l}{1-\theta}\right). \qquad \text{CDF}$$

The PDF of the loss distribution of LHP will be PDF

$$f(y) = \sqrt{\frac{1-\rho}{\rho}} \exp\left(-\frac{1}{2\rho}\left(\sqrt{1-\rho}\,\Phi^{-1}(y) - d(t)\right)^2 + \frac{1}{2}\left(\Phi^{-1}(y)\right)^2\right).$$



no default

default

$P = 1$

Expected base tranche Loss

By using analytic distribution for LHP, one can use to calculate the expected base tranche loss(tranche with no subordination), for example the equity tranche loss function $L(t; 0, l)$.

$$
\begin{aligned}
\mathbb{E}\left[L(t;0,l)\right] &= \mathbb{E}\left[L(t)\mathbf{I}\{L(t) < l\} + l\,\mathbf{I}\{L(t) \geq l\}\right] \\
&= \mathbb{E}\left[(1-\theta)F(t|Z)\mathbf{I}\{L(t) < l\}\right] + l\,\Phi(-a) \\
&= (1-\theta)\mathbb{E}\left[F(t|Z)\mathbf{I}\{Z > -a\}\right] + l\,\Phi(-a) \\
&= (1-\theta)\Phi_2\left(d(t), a, -\sqrt{\rho}\right) + l\,\Phi(-a),
\end{aligned}
$$

where

$$
a = \frac{\sqrt{1-\rho}\,\Phi^{-1}(\frac{l}{1-\theta}) - d(t)}{\sqrt{\rho}},
$$

and $\Phi_2(x, y; r)$ is bivariate normal distribution with correlation parameter r.

Expected Loss for general tranche

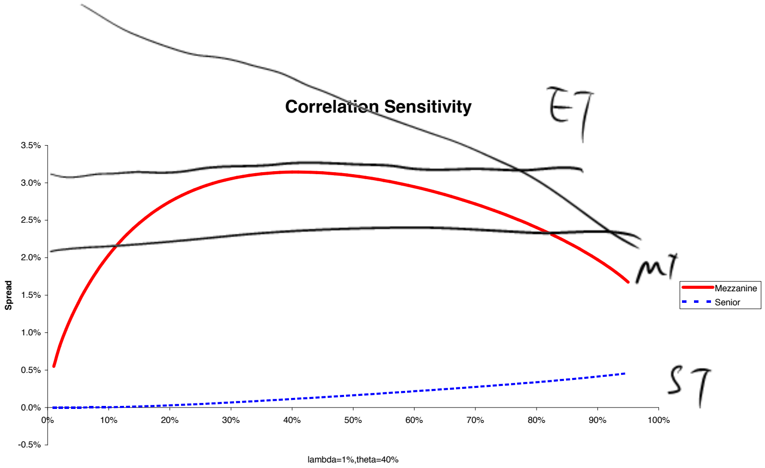
One can use the expected base tranche loss to calculate general tranche loss like this

$$\mathbb{E}\left[L(t; l_1, l_2)\right] = \mathbb{E}\left[L(t; 0, l_2)\right] - \mathbb{E}\left[L(t; 0, l_1)\right].$$

Therefore the closed for solution for tranche spread can be obtained by using this formula

$$s = \frac{\sum_{j=1}^{M} P(0, t_j)\left[\mathbb{E}[L(t_j; d, u)] - \mathbb{E}[L(t_{j-1}; d, u)]\right]}{\Delta \sum_{j=1}^{M} P(0, t_j)\left[1 - \mathbb{E}[L(t_j; d, u)]\right]}. \quad (5)$$

Note: $L(t_j; u, d)$ is redefined to be percentage of tranche loss.


Correlation Sensitivity
ET
MT
ST
Spread
3.5%
3.0%
2.5%
2.0%
1.5%
1.0%
0.5%
0.0%
-0.5%
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
lambda=1%,theta=40%
Mezzanine
Senior

Static Model

The one factor normal copula is

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

However this model is static, in the sense that the output of the model is only based on what has been given at the beginning, and nothing is supposed to change since then. This is an undesirable feature, as we always want to account for rich dynamics in our model.

Implied Correlation

Normal copula model has been the market standard model, in which all variables that affect tranche spread are typically know apart from default correlation.

Therefore, given the spread of a tranche one can use the model to backout the correlation parameter that best coincides with the spread.

This correlation is called **Implied Correlation** or **Compound Correlation**.

Drawbacks of Implied Correlation

The concept of implied correlation is very appealing in terms of its similarities with implied volatility in equity option. However it suffers a few drawbacks.

1. Theoretically all tranches should have identical implied correlation, but in practice we observe correlation skew.
2. For mezzanine tranche its spread is not unique.
3. And it is difficult to calibrate, i.e., for certain spread implied correlation might not exist.

Base Correlation

In this respect the market uses implied base correlations and the Gaussian copula model in much the same way as it uses implied volatilities in equity model.

For example to price exotic option volatility implied from vanilla option is used.

However the base correlation also has problems. It is unable to guarantee positive monotonicity of expected accumulated loss, as a result the principle of no-arbitrage is failed.

Uncertain Correlation

Here we propose an uncertain correlation model, compare with other models it benefits from the following

1. It stays in the framework of factor copula model, requires little change,
2. hence very easy to implement both in static and dynamic model.
3. It generates a wider range of tranche spread.
4. It can be made to comply with different economic and default scenarios.

Dynamic Model

The factor copula model is so popular so that no one can afford to abandon it. Luckily this can be modified to accommodate rich dynamics.

$$A_i(t) = w_i Z(t) + \sqrt{1 - w_i^2}\, \varepsilon_i(t). \tag{7}$$

State Dependent Correlation

The the assumption that default correlation is a constant is inappropriate. It can depend on many thing.

One scenario is that correlation may depend on the common factor $Z(t)$.

Intuitively the common factor represents the general condition of the financial markets or the wider economy. If there is a financial crisis or economic recession, we would expect to more firms defaulting. Regulators often worry about systemic risk. In this case, default is contagious. The default of a large organization may spread out causing further defaults. So a lower value of $Z(t)$ implies higher default correlation.

Or it could be the other way around if two firms are competitors and one of them defaults.

Assumptions of uncertain correlation model

To accommodate the above possibilities without leaving the copula framework, we shall consider the model of uncertain correlation. The model has the following characteristics:

- The correlation between all underlyings is the same, same as standard copula model.
- That correlation lies between ρ_l and ρ_u.
- That correlation can vary as a function of Z.
- The dependence of correlation on Z is a step function, flipping from ρ_l to ρ_u (or *vice versa*) as Z goes above a threshold Z_k.
- The threshold Z_k is chosen to give each tranche its lowest value.

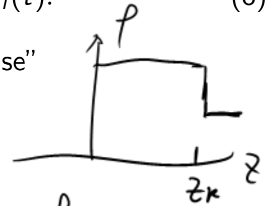
Step correlation function

Based on above assumption and dynamic equation (**??**) can be rewritten as

$$A_i(t) = \sqrt{\rho(z)}Z(t) + \sqrt{1-\rho(z)}\,\varepsilon_i(t). \tag{8}$$

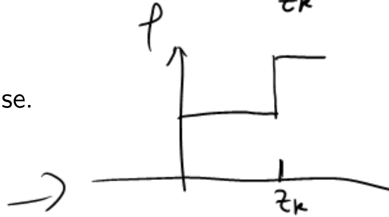
where $\rho(z)$ is a step function such that for "worst case"

$$\begin{cases} \rho = \rho_l, & z > z_k; \\ \rho = \rho_h, & \text{otherwise.} \end{cases}$$



and "best case"

$$\begin{cases} \rho = \rho_h, & z > z_k; \\ \rho = \rho_l, & \text{otherwise.} \end{cases}$$



Constant correlation model

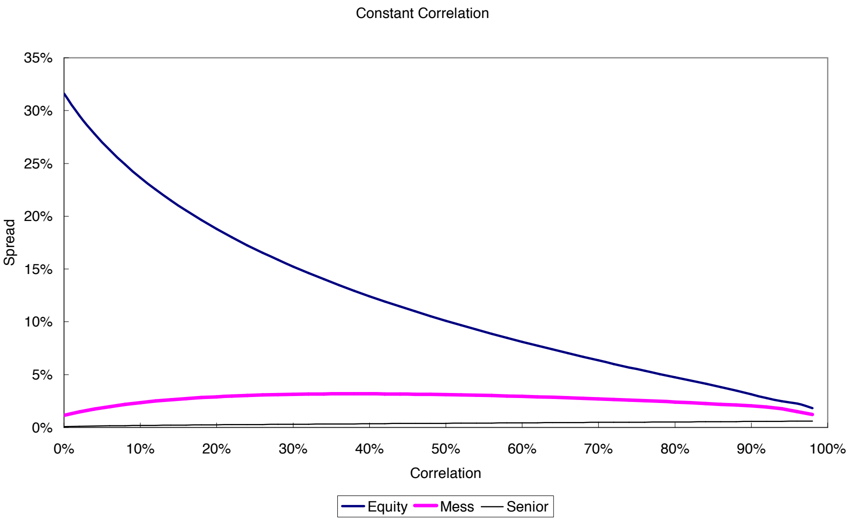
We test best-worst case scenarios on Mezzanine tranche using both constant and uncertain correlation model.

In the constant correlation model we calculate tranche spread with different correlation values varying between 0 and 1.

We assume there are 125 names, each with constant intensity of 1%, recovery rate of 40%.

The CDO has typical 3 tranches:$0-3\%, 3\%-10\%, 10\%-100\%$, 5 years time to maturity, and quarterly premium payments.

Interest is constant 5%.


Constant Correlation
Spread
35%
30%
25%
20%
15%
10%
5%
0%
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Correlation
Equity
Mess
Senior

Implement uncertain correlation model

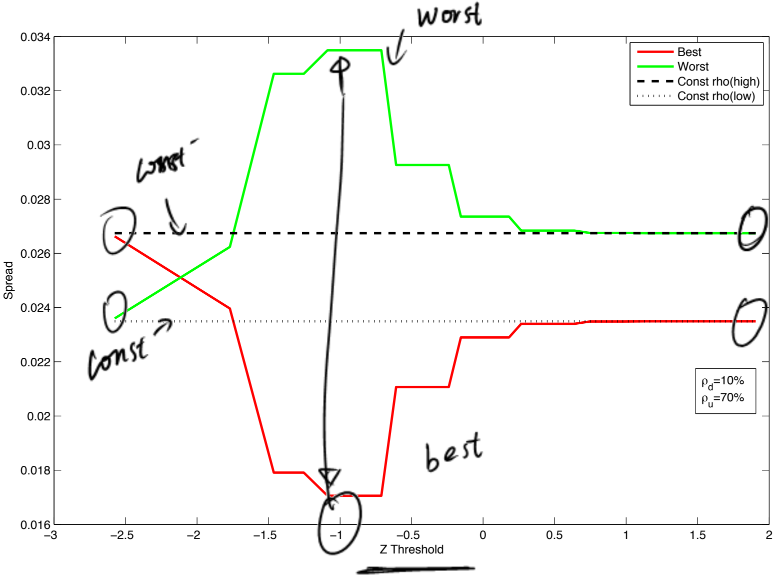
In uncertain correlation model we assume $\rho_l = 10\%$ and $\rho_l = 70\%$. We calculate tranche spread with different correlation values depending on the value of Z. We repeat the calculation with different values of threshold which triggers jump in correlation. The calculation consists of the following steps:

1. Calculate conditional default probability $F(t, z; \rho(z))$.
2. Calculate portfolio loss distribution for all time steps.
3. Calculate expected tranche loss for all time steps.
4. Calculate tranche spreads.

Results of uncertain correlation model

The next figure plots Mezzanine tranche spread against different triggers. Note the dotted lines give tranche spreads when correlations are constants, the solid lines are the worst case and best case. There are several observations:

1. Green and red lines start from corresponding constant correlation. This is because when $z_k = -\infty$, correlation always jumps up to ρ_h in best case(red line case). Whereas it always jumps down to ρ_l in worst case(red line).
2. Green and red lines ultimately settle down to the opposite of constant correlation. For example, when z_k is very large, the green line settles down to lower correlation, whereas the red line settles up to higher correlation.
3. The distance formed by best and worst case in the uncertain correlation model is significantly wider than that formed by the constant correlation model.


worst
Worst
const
best
Spread
Z Threshold
Best
Worst
Const rho(high)
Const rho(low)
ρd=10%
ρu=70%

Please take away the following important ideas

- Correlation is not a panacea for every dependency problem.
- Like volatility, correlation sensitivity to a derivative contract could appear low, but it might be its cross gamma isn't single signed.
- Uncertain correlation with best-worst case analysis can help us to get more sensible results.