Incorporating Multidimensional Tail-Dependencies in the Valuation of Credit Derivatives

Noel McWilliam   Kar-Wei Loh

Working Paper, November 2008

Abstract

The need for an accurate representation of tail-risk has become increasingly acute in the wake of the credit crisis. We introduce a hyper-cubiod normal mixture copula that permits the representation of complex tail-dependence structures in a multi-dimensional setting. We outline an efficient pattern recognition calibration methodology that can identify tail-dependencies independent of the number of risk factors considered. This model is used to develop a new framework for pricing credit derivative instruments, we derive semi-analytical and analytical pricing formulae for a First-to-Default Swap and illustrate with an example valuation. Model assumptions are validated against iTraxx Series 5 equity data over an 8 year period. Identification and representation of tail-dependencies is crucial to further study of contagion dynamics, our model provides a basis for future research in this area.

Introduction

Future returns of a market variable such as an exchange rate, equity price, or interest rate are commonly held to be log-normally distributed. Intuitively we might expect this to be the case because returns over a given time horizon can be thought of as the cumulative product of smaller returns within an embedded set of smaller time horizons. Assuming that the smaller log-returns are independent and identically distributed with finite variance, then we may invoke the Central Limit Theorem to infer that the sum of the logs, and hence the overall log-return, must be normally distributed. The Gaussian distribution, ubiquitous in many areas of science, is of particular use in finance because it provides a strong theoretical foundation for many stochastic diffusion processes. It also leads to simple no-arbitrage solutions for derivative instrument pricing, such as the celebrated Black-Scholes formula. Empirical evidence shows however that the likelihood of large returns on a market variable over short to medium time intervals is typically much greater than would be implied by a normal distribution. This is manifest in a frequency plot of returns having "fatter tails" than that of a normal. Returns also show a degree of asymmetry, in contrast to a normal's symmetric bell curve shape.

Authors such as Bouchaud (1999) attribute these inconsistencies to a weakness in the common assumption of independence of returns between two non-overlapping time periods: in practice market participants modify their activities on the basis of past performance and in doing so introduce high order auto-correlations between returns at different times. This effect is likely to be enhanced in crisis situations if positive feedback mechanisms between trader behaviours strengthen these time dependencies. As such, failure to capture tail risk is never more consequential than in times of crisis and will undoubtedly lead to an underestimation of VaR and mis-pricing of tail sensitive instruments at the precise moment when investors have the greatest need for robust and accurate pricing mechanisms and risk statistics.

The task of evolving from a normal framework is problematic not least because the characteristics of dependencies between non-normal variables are much richer than can be captured via a matrix of linear correlation coefficients. For this reason there has been a significant movement away from Gaussian dependence in recent years towards copula based pricing and risk management practices. Copula are advantageous in that they can exhibit a wide range of non-linear dependence structures, and in particular are able to capture tail-dependence — a measure of the propensity of random variables to co-move in an "extreme" way.

Unfortunately many copula are unable to represent complexities within the tails. A student-t copula for example is expressed in terms of a single tail parameter that imposes tail-dependence uniformly across all dimensions. In reality different markets may show varying degrees of extreme dependence in relation to their geographical, industrial or political proximity (or any other such measure of similarity). These structures are likely to impact the pricing and risk management of tail-sensitive instruments and in particular credit derivative (CD) instruments since their value is critically dependent on the relationship between defaults, which are themselves extreme events.

It is our view that the unfolding credit crisis has as a contributing factor a fundamental mis-pricing of CD instruments that ignores the complexities of tail-dependence. Moreover in the absence of a framework in which to formulate multi-dimensional tail-dependencies one cannot quantify nor understand the phenomenon of contagion; the process by which large shocks filter through markets.

In this paper we suggest an extension of a Gaussian factor-copula asset-based structural default model (Hull and White 2001, Li 2001) in which log-return dependencies are given by a hyper-cuboid normal mixture copula. This setting permits specification of tail dependencies within and between all possible sub-spaces of the multivariate; thus conferring a significant advantage over other copula. Our work has important implications for risk management in the measurement and characterisation of tail risk and in the pricing of derivative instruments sensitive to such factors.

The Structural Factor-Copula Default Model

Single name credit derivatives are contracts on cash-flows contingent on the credit status of a single obligor. Although reliant on a simplified interpretation of credit spreads as driven only by default and recovery expectations of market participants it is common practice to price such contracts in a manner consistent with the term structure of spreads on risky bonds or market quoted CDSs underwritten on the same (or similar) obligor. In doing so these models obviate the collection and interpretation of large volumes of counterparty specific data pertaining to capital structure and avoids reliance upon unobservable parameters.

In the case of basket CDs cashflows are contingent on sequences or combinations of defaults for a collection of obligors. Here additional information is required as to the likelihood of these events, which requires calibration of a joint default distribution such that the implied marginal default characteristics are consistent with the single name counterparts. For n^{th}–to–default CDSs and CDOs one typically models each individual default as a Poisson process for which the cumulative default probability $F(\cdot)$ is given as

$$F(t) = 1 - e^{-\int_0^t h(\tau)d\tau}, \tag{1}$$

where the instantaneous hazard rate function $h(\cdot)$ is calibrated to the term structure of an appropriate credit spread.

Li (2001) assumes that the multivariate distribution of default times and the multivariate distribution of a set of associated asset returns share the same dependence structure. Since a counterparty's asset returns are not directly observable in the market it is common practice to use equities as a proxy, which are themselves assumed to be (standard) Gaussian in distribution. Hence at any given time t during the lifetime of the contract we may relate via a Copula function each marginal default probability $F_i(t), i = 1, \ldots, n$ to a set of barrier asset values

$$\beta_i = \Phi^{-1}(F_i(t)).$$

below which we assume a default of the i^{th} obligor to have occurred.

To extract the joint default/survival probabilities or any combination thereof we need to integrate the multivariate over the appropriate volume defined by the barrier values: Let $\phi_{MV}(x;0,1,\rho)$ denote the standard (zero mean, unit variance) n-dimensional multivariate normal distribution with ρ the correlation matrix and state space $x = [x_1, ..., x_n]^\top$. The probability of default by time t of all names is then

$$\int_{-\infty}^{\beta_n} \cdots \int_{-\infty}^{\beta_1} \phi_{MV}(x;0,1,\rho)\ dx_1 \dots dx_n.$$

On the basis that equity returns contain predictive information as to the likelihood of joint default, then we must surmise that failure to incorporate tail dependencies within equity return structures will undoubtedly lead to an incorrect formulation of the default risks posed. Given that defaults are themselves generally tail events then it seems pertinent to ask how one might better incorporate such dependencies into the model.

A Multivariate Normal Mixture Copula

As a means to capture tail dependence in equity returns we consider normal mixture copula sharing the general form

$$\mathbf{C}(u) = \sum_{k=1}^{m} w_k \Phi(\Phi_1^{-1}(u_1), \dots, \Phi_n^{-1}(u_n); \mu_k, \sigma_k, \rho_k)$$

$$\sum_{k=1}^{m} w_k = 1, w_k > 0$$

where $m < \infty$ is the number of component normals and $\mu_k \in \mathbb{R}^n, \sigma_k \in \mathbb{R}^n_+$ and $\rho_k \in \mathbb{R}^{n\times n}$ are the k^{th} component mean, standard deviation and correlation respectively. Here $u \in [0,1]^n$ with each $\Phi_i, i = 1, \dots, n$ a standard cumulative normal.

By characterising returns as comprising a collection of states each having its own unique profile we admit a model that incorporates the possibility of structural shifts, thus providing a means by which to accommodate crisis events into our pricing framework.

However, additional flexibility comes at a cost: the question as to the incidence and magnitude of such shifts, for example, poses a significant identification problem since the states themselves are not directly observable. Computational issues surrounding parameterisation can also be considerable with each normal component specified using a parameter set of dimension $2n + \frac{1}{2}(n^2 + n) + 1$.

In terms of our intended use for the model we must be sensitive to the pricing problems that already exist even within the Gaussian framework. In this regard, and as a means to introduce the model parameterisation, we might first suggest borrowing the notion of conditional independence, so that conditional on the state of the market $k = 1, \dots, m$ we suppose $\rho_k = I_n$.

Under this assumption we may deduce analytic solutions to the joint probability of default and survival, whereupon the latter at time t is

$$\sum_{k=1}^{m} w_k \prod_{i=1}^{n} \Phi\left(\frac{\mu_{k(X_i)} - \beta_i}{\sigma_{k(X_i)}}; 0, 1\right) \quad (2)$$

with $\mu_{k(X_i)}$ and $\sigma_{k(X_i)}$ the mean and standard deviation of the k^{th} component normal along the i^{th} dimension and $\beta_i = \Phi^{-1}_{NM_i}(F_i(t))$ is the corresponding barrier where

$$\Phi_{NM_i}(\cdot) = \sum_{k=1}^{m} w_k \Phi_i(\cdot; \mu_{k(X_i)}, \sigma_{k(X_i)})$$

is the cumulative normal of the projection of the multivariate mixture onto the i^{th} univariate margin.

It is important to note that although the component normals are uncorrelated when considered individually, correlations in the overall distribution may not be zero. Overall we note that

$$\rho_{ij} = [\sum_{k=1}^{m} w_k [\rho_{k(X_i X_j)} \sigma_{k(X_i)} \sigma_{k(X_j)} + \mu_{k(X_i)} \mu_{k(X_j)}] - \mu_{(X_i)} \mu_{(X_j)}][\sigma_{(X_i)} \sigma_{(X_j)}]^{-1} \quad (3)$$

where

$$\mu_{(X_i)} = \sum_{k=1}^{m} w_k \mu_{k(X_i)},$$
$$\sigma^2_{(X_i)} = \sum_{k=1}^{m} w_k[\sigma^2_{k(X_i)} + \mu^2_{k(X_i)}] - \mu^2_{(X_i)}$$

and $\rho_{k(X_iX_j)}$ is the k^{th} component correlation between X_i and $X_j; i \neq j = 1, \ldots, n$. This is a generalisation of the correlation mapping as given by Wang (2001).

Assuming without loss of generality [1] that the univariate mixtures each have zero mean and unit variance then the overall correlation under the assumption of conditional independence will be of the form

$$\rho_{ij} = \sum_{k=1}^{m} w_k \mu_{k(X_i)} \mu_{k(X_j)}, \tag{4}$$

which is independent of any choice for σ_k.

Thus by optimising the means and weights of the mixture it is possible to price CDs analytically while retaining at least some of the correlation structure given in the underlying data. Indeed, in this formulation one is not restricted to the inner product structure for correlation as is inherent in the standard conditional independence framework. More formally given a fixed m we consider

$$\min_{w,\mu_s} \parallel \rho_T - \rho(w, \mu_s) \parallel : \sum_{k=1}^{m} \mu_{k(X_i)} = 0,$$
$$\sum_{k=1}^{m} w_k = 1, w_k \geq 0 \tag{5}$$

where $\rho_T, \rho(w, \mu_s)$ are the "target" (i.e. market observed) and reconstructed (from relation 4) correlation matrices respectively.

One may perform this optimisation once across all names in a set of basket CDs and thus maximise the computational savings that this approach offers. Figure 1 shows the results of an optimisation performed

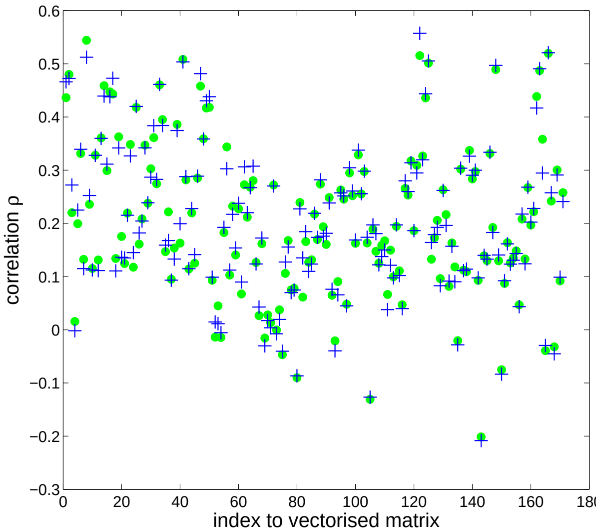

Figure 1: Optimised $m = 10$ (Analytic) and Target Correlations: Scatter Plot Compares Target (cross) with Optimal Values (dot)

on correlation data between equities corresponding to 19 of the most liquid public names in iTraxx data series 5. Correlations are calculated on a six month window over the period 29.06.2003 to 28.12.2003. The computation time was 9.19 seconds using MATLAB 7.1 on a Intel Core Duo CPU @ 2.4GHz. Values for $k = 10$ show a very good fit to the correlation data having an average absolute error of 0.0128. Given the optimised multivariate we may determine the corresponding copula which may be used as the basis for pricing CDs.

A Hypercuboid Copula

Although the unconstrained conditionally independent approach admits analytic pricing for CD instruments (such as first-to-default) while capturing correlation, it is by no means clear that it will adequately reflect the true multivariate structure since one may define many such distributions having the same correlation. This is apparent from Equation 4 which is expressed independently of the component standard deviations, and belies the fact that correlation is an

[1] Our assumption on the form of the univariate margins will not impact the marginal form of the copula, although our restriction here will generally affect the results of the optimisation.

inadequate dependence measure outside of a linear, Gaussian framework. In order to create a more detailed picture of the multivariate and hence a more accurate representation of CD prices we require additional information from the asset return distribution.

With this in mind we construct a normal mixture copula having univariate margins matched to higher order moments of empirical log-return data (see Wang 2001 for example). In this way the copula will contain information as to the magnitude and direction of large shocks within each univariate. The question then remains as to the manner in which they are related.

As we show, a "bottom-up" approach to calibration allows us to successively restrict the form of the multivariate, giving rise to a hyper-cuboid mixture whose tail dependencies should more closely reflect those embedded within the data. We begin by outlining our method for univariate calibration.

Univariate Construction

The model considered herein assumes each univariate to comprise a mixture of two normals. Let X_i denote the i^{th} risk factor then its first four non-central moments are

$$\begin{aligned} m_1 &= p\mu_0 + (1-p)\mu_1 \\ m_2 &= p\sigma_0^2 + (1-p)\sigma_1^2 + p\mu_0^2 + (1-p)\mu_1^2 \\ m_3 &= 3p\sigma_0^2\mu_0 + 3(1-p)\sigma_1^2\mu_1 + p\mu_0^3 + (1-p)\mu_1^3 \\ m_4 &= 3p\sigma_0^4 + 3(1-p)\sigma_1^4 + 6p\sigma_0^2\mu_0^2 \\ &\quad +6(1-p)\sigma_1^2\mu_1^2 + p\mu_0^4 + (1-p)\mu_1^4 \end{aligned}$$

where $p \in [0,1]$ and for clarity we denote $\mu_k = \mu_k(X_i)$ and $\sigma_k = \sigma_k(X_i), k = 0, 1$. The centralised moments are

$$\begin{aligned} \mu &= m_1 \\ \sigma &= \sqrt{(m_2 - m_1^2)} \\ s &= \frac{1}{\sigma^3}(m_3 - 3m_2m_1 + 2m_1^3) \\ k &= \frac{1}{\sigma^4}(m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4). \end{aligned}$$

In our calibration we choose the mixture parameters to match the centralised moments of empirical data, requiring the simultaneous solution of the above non-linear equations. Since we have five unknowns and four equations we will choose a value for the mixing probability p in such a way as to maximise the region of real solutions. Without detriment to the generality of our analysis for the remainder of this paper we will assume p is a fixed value uniformly defined for all univariates.

Multivariate Construction

Once the univariate margins have been calibrated it is necessary to define a dependence structure between the risk factors. Since each univariate is specified as a mixture of two normals it follows that the bivariates are mixtures of at most four, each corresponding to a unique pairing of the component states. Following this logic the resultant n-dimensional multivariate will therefore be a mixture of 2^n normals (see Figure 2 for an illustration), each "centrally" located [2] at a vertex of an n-dimensional hyper-cuboid.

If we consider each univariate as comprising a "core" and "tail" then the hyper-cuboid vertices represent the possible combinations of core and tail scenarios. Through our choice of probability weights $W = [w_1, \ldots, w_{2^n}]$ we may determine the likelihood of a joint tail movement in any subset of risk factors. However, we must be careful to ensure that our choice is concomitant with our univariate specification so that the projection of the multivariate weights onto each univariate satisfies the pre-specified mixing probabilities, namely that

$$\psi W = (1-p)\mathbf{1}_{n\times 1} \qquad (6)$$

where

$$\psi = \begin{bmatrix} \mathbf{1}_{1\times 2^{(n-1)}} \otimes [0,1] \otimes \mathbf{1}_{1\times 2^0} \\ \vdots \\ \mathbf{1}_{1\times 2^0} \otimes [0,1] \otimes \mathbf{1}_{1\times 2^{(n-1)}} \end{bmatrix}.$$

and that we satisfy the basic property

$$\sum_{i=1}^{2^n} w_i = 1 \qquad (7)$$

[2] For an interesting discussion on the topology of mixture distributions with particular reference to modality see Ray and Lindsay, 2005.

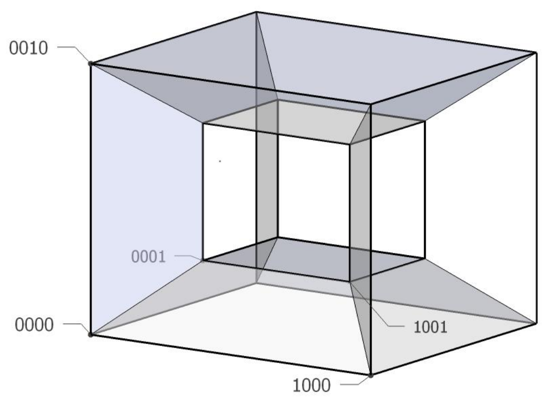

Figure 2: Hypercubiod Illustration: 3D projection of a 4D multivariate.

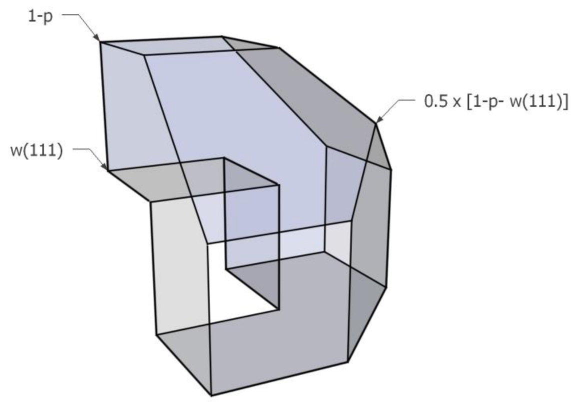

Figure 3: 3D Reachable Set for τ.

with $\otimes$ the Kronecker product and $\mathbf{1}_{i\times j} \in \mathbb{Z}^{i\times j}$ a matrix having unit entries.

Note that each column of ψ corresponds to a unique permutation of core (0) and tail (1) events, whereas the i^{th} row identifies a sub-set of the multivariate tails contributing to the tail of univariate X_i.

Since the rank of ψ is n it follows that there are many $(2^n - n - 1)$ choices of W that would satisfy our univariate constraints. We might like however to make use of bivariate data to further inform our model. In order to choose W in a manner consistent with a set of given bivariate distributions it is instructive to first determine the set of permissible bivariates.

When considered on a pair-wise basis we have $2^2 - 2 - 1 = 1$ degree of freedom in our choice of bivariate weights. Let τ_{ij} denote the probability weight associated with the joint tail of the bivariate for the ij^{th} risk-factor pairing, that is

$$\tau_{ij} = [\psi_{R:i} \times \psi_{R:j}]W$$

where here $\times$ represents the element-by-element or "group product" operator and $\psi_{R:i}$ is the i^{th} row of the matrix ψ.

Using τ_{ij} as our control variable we deduce from equations 6 and 7 the absolute bound $0 \le \tau_{ij} \le 1-p$. Clearly we have considerable flexibility in our choice for τ_{ij} when considered in isolation. However the extension to a multivariate setting imposes an additional set of relative constraints on the (symmetric) matrix $\tau \in \mathbb{R}^{n\times n}$ of unique τ_{ij} values $i, j = 1, \ldots, n$. Specifically by re-expressing the univariate constraints with respect to τ we may obtain a reachable set

$$\sum_{j=1, i\neq j}^{n} \tau_{ij} + \psi_{R:1+n-i} \times (\mathbf{2}_{2^n} - \xi)W = 1 - p.$$

for each $i = 1, \ldots, n$ and

$$\sum_{i=1}^{n-1} \sum_{j=1, i\neq j}^{n-1} \tau_{ij} + \Bigg[\mathbf{1}_{1\times 2^{n-1}} \otimes [1, 0] - \left(\sum_{i=1}^{n-1} \psi_{R:1+n-i} - \psi_1\right) \times (\xi - \mathbf{1}_{2^n})\Bigg] W = p,$$

where $\mathbf{2}_{2^n} = [2, \ldots, 2] \in \mathbb{R}^{1\times 2^n}, \xi = [\xi_1, \ldots, \xi_{2^n}] \in \mathbb{R}^{1\times 2^n}$ with

$$\begin{cases} \xi_{2^k} &= k \qquad k = 0, \ldots, n \\ \xi_{2^k+t} &= \xi_t + 1 \quad 0 < t < 2^k. \end{cases}$$

a fractal sequence in the limit as $n \to \infty$.

These equations define a region of permissible τ values in $\frac{n^2-n}{2}$ dimensional space in analogy to the region of positive definiteness for correlation coefficients in a Gaussian setting. [3] Even for $n = 3$ the region can be quite complex (See Figure 3 for an illustration).

The question of choosing W may then be thought of as optimising the above volume such that the target matrix τ_T (deduced from empirical data) lies within it or close as possible to it. Formally we have the following constrained linear optimisation problem:

$$\min_{W \in \mathbb{R}^{2^n}} \parallel \beta W - \pi_T \parallel : \psi W = \begin{bmatrix} 1-p \\ \vdots \\ 1-p \end{bmatrix}, \quad \sum_i w_i = 1, w_i \geq 0 \tag{8}$$

where $i = 1, \ldots, 2^n$, $\pi = vec\{\tau\} \triangleq [\tau_{12}, \ldots, \tau_{1n}, \tau_{23}, \ldots, \tau_{2n}, \ldots, \ldots \tau_{(n-1)n}]^\top$ with $\pi_T = vec\{\tau_T\}$ and $\beta : W \mapsto \pi$ a linear map.

While numerous techniques exist for solving the above, in our experience the problem has proven intractable via classical methods for even moderately sized systems ($n > 15$) since the dimension of the parameter space increases exponentially with n. As we have previously intimated, it is crucial that our model should be as complete as possible if we are to identify and condition upon stress patterns that may arise from (or even prelude) contagion effects. For this reason we outline in the next sub-section a method by which we may identify key multivariate tail-dependence structures in a manner independent of the dimension of the underlying system. The solutions provided by this "pattern analysis" may be used as an alternative to or means of conditioning the above optimisation problem.

Pattern Analysis

The method outlined herein provides an efficient means in which to identify a subset of columns of ψ that can explain a large degree of the variation in the matrix of bivariate tail probabilities τ_T. This is achieved by recasting the optimisation problem in 8 as one of pattern recognition. In this framework high-dimensional τ matrices are considered as high-resolution images in which we search for simple geometric shapes. In doing so we may address the optimisation problem at a suitable level of resolution independent of the original dimension, thus allowing us to consider systems of arbitrary size without the associated computational difficulties.

In this approach we partition the range of permissible τ values $[0, 1-p]$ into sub-intervals of length $\frac{1}{d}, d \in \mathbb{Z}^+$ defining for each the matrix $\tau_{T_i} = \max\{0, \min\{\tau_T, \frac{1}{d}\} - \frac{i-1}{d}\}$. For each $i = 1, \ldots, d$ we approximate the variation within τ_{T_i} as the sum of a set of "simple" matrices. In essence we consider d distinct optimisation problems for which the variation of the target values is constrained. In many practical instances this simplifies the target structure and facilitates identification of simple shapes within it.

Within each partition we look for simple non-overlapping square block structures having reflectional symmetry about the principal diagonal so that the location of the j^{th} block is completely characterised by the indices of its lower l_j and upper u_j bounds. Each block corresponds to the bivariate projection of tail dependence within a subspace of the multivariate whose indices are contiguous in $\mathbb{Z}^+$. This structure is represented by the column vector $\psi_{C:j}$ whereby $\psi_{vj} = 1$ for $l_j \leq v \leq u_j$ and $\psi_{vj} = 0$ otherwise. The associated probability weight $w_j = \frac{1}{d}$. Note that by restricting our search to non-overlapping blocks for each i we only consider subsets of ψ columns within which all elements are orthogonal.

Once the set of all such blocks have been identified we determine which columns we should "merge". To this end we calculate the ratio

$$\frac{\parallel \tau_{T_i} - \frac{1}{d}(\psi_{C:j} + \psi_{C:h})(\psi_{C:j} + \psi_{C:h})^\top \parallel}{\parallel \tau_{T_i} - \frac{1}{d}\psi_{C:j}\psi_{C:j}^\top - \frac{1}{d}\psi_{C:h}\psi_{C:h}^\top \parallel},$$

with $j = 1, \ldots, m_i; h = j+1, \ldots, m_i$ where m_i denotes the number of elements within the i^{th} set. If

[3] These same relations may also be used to express the set of permissible W with respect to a given τ_T.

the ratio is less than unity then the error associated with the merged column $\psi_j + \psi_h$ is less than that of the components when considered individually and we substitute the pair with its sum. In contrast to the original elements within the set, the sequence of indices to non-zero entries of a merged column may not be contiguous in $\mathbb{Z}^+$, thus introducing greater complexity to the kernel structures of the corresponding bivariate projections.

The reconstructed τ matrix is then composed as the union of the sets for each partition. While elements within each of the aforementioned sets are orthogonal, it is important to note that no such restriction is enforced between the sets. This allows us to capture a considerable degree of variation observed within the target.

The results of this analysis may be used as a proxy solution to the full optimisation problem or alternatively used as a means to condition the problem and reduce the dimension thereof.

Figure 4 shows the results of pattern analysis optimisation on a target matrix τ (as implied via correlation matching — see the next section for details) for the aforementioned iTraxx equity data set. Pattern analysis reduces the dimension of the optimisation (8) from $2^{19} = 524,288$ to 38. As can be seen our results provide a good representation of the target data having an average absolute error 0.0054. Pattern analysis computation time was 0.68 seconds, conditioned optimisation took 0.45 seconds.

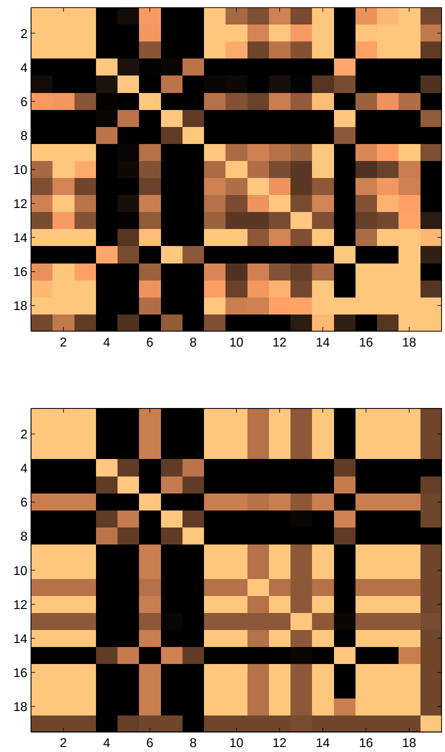

Figure 4: Pattern Analysis Example: Target τ (top) and Reconstructed τ (bottom). Data: target implied via correlations from daily log-returns for 19 liquid Itraxx Series 5 equities between 29/06/03–29/12/03.

An Example Calibration

We outline a candidate bivariate calibration technique that is some sense similar to the moment matching technique used for the univariate margins. We then provide statistical validation of both univariate and bivariate methods using real equity data.

Correlation Matching

In the analysis that follows we make use of correlation data to infer a set of bivariate tail dependencies. Although by no means necessary, for simplicity we will assume that each component normal is correlated identically to the overall distribution so that $\rho_{k(X_i X_j)} = \rho_{ij}$. Given the component univariate mixture parameters, and from (6) and (7) with $n = 2$ that $w_4 = \tau_{ij}, w_2 = w_3 = 1 - p - \tau_{ij}, w_1 = 2p - 1 - \tau_{ij}$ then may invert the mapping 3 and obtain an implied tail-dependence value which will serve as our target. Any values exceeding the permissible bounds will be rounded up or down accordingly.

Validation

Our aim is to test the basic premise of our parameterisation; namely that the risk factors under consideration are distributed in accordance with a normal mixture, and that we may infer tail dependencies via the fitted univariate mixture parameters and observed correlation values.

To this effect we evaluate the descriptive capability of our technique for univariate and bivariate time-series data. We provide s^2 goodness of fit and Kolmogorov-Smirnov (KS) statistics whereby the null hypothesis assumes the data to have originated from our parameterised model. All results are benchmarked against equivalent measures for a Geometric Brownian Motion (GBM).

Since our work is framed within the context of CD valuation we consider data for 19 of the 45 names in iTraxx Series 5. These correspond to the most liquid publicly traded equities over the period 15.05.2000 to 10.01.2008.

Method

We calibrate the univariate and bivariate margins on a rolling window of daily log-return data evaluated at increments of the same size as the window (so that the windows do not intersect and their union covers the entire period). Each empirical bivariate is compared with a mixture distribution parameterised by the *target* matrix τ_T as determined via our correlation matching procedure.

Test statistics are provided "in-sample" that is, using the same data as used for calibration. By comparing the empirical cumulative distribution against that of a normal mixture we generate KS and s^2 values that indicate goodness of fit.All tests are performed on the assumption that the underlying stochastic process is i.i.d.

Results

Figure 5 shows the results of a univariate KS test on FKI Ltd. on a moving 1 year and 6 month window respectively. These results are typical of those obtained in our analysis.

Inspection of the 1y reveals some apparent temporal progression of KS values with the highest probabilities given outside of the mid-range interval between 2002 and 2005. In virtually every instance the

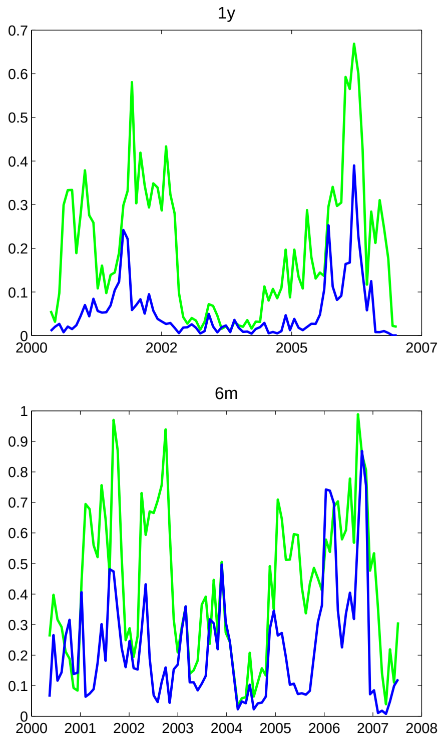

Figure 5: Univariate KS probabilties on a 1y and 6m moving window over an 8 year period, green indicates normal mixture fit, blue a normal fit. Data: FKI Ltd Daily Log-Returns.

mixture fit performs better than the normal fit; the latter — in contrast to the former — is rejected almost everywhere at the 95% confidence level. Both fits are markedly improved for the 6 month data although the normal mixture probabilities are still almost uniformly higher than for the normal.

By amalgamating the data for all 19 equities we derive the proportion of time-instances when the null hypothesis is rejected in the 1 year mixture (10.21%) and Gaussian (30.04%) model. Both values exceed the 5% level one would expect had the hypothesis been true (and the experiments independent). Unlike the normal model at 9.81%, the 6 month mixture model shows a marked improvement in this regard with only 2.87% rejected.

		GRAND MEAN	AV STDEV.
KS			
1D	Nml	0.44	0.28
	Mix	0.63	0.26
2D	Nml	0.27	0.23
	Mix	0.46	0.25
s^2 Error			
1D	Nml	0.2585	0.2583
	Mix	0.1396	0.1536
2D	Nml	0.8454	0.6323
	Mix	0.5733	0.8577

Table 1: Grand results, Normal Vs. Normal Mixture: KS probabilties and s^2 Errors for 6m moving window.

Table 1 gives the mean (across all 19 equities) of the mean (across the time-line) KS probability and s^2 errors for both univariate and bivariate data under the 6m parameterisation. We refer to the mean of the means as the "grand mean". We also provide values for the standard deviation (across the 19 equities) of these means. As can be seen the mixture distribution outperforms the normal by a substantial margin in both the KS probabilities and goodness of fit measures for the univariate and bivariate fits. While in either instance the mixture fit is rarely rejected, the univariate statistics perform better relative to the bivariates, which may reflect our choice of tail-measure parameterisation.

Credit Derivative Pricing

In the context of the popular structural model for default we make use of our hyper-cubiod mixture to extend the factor-copula approach of Li (2001). A significant advantage in this regard is the ease by which we can formulate tail-consistent prices within a semi-analytic framework. Li and Liang (2005) similarly make use of a mixture copula to price a CDO-tranche consistent CDO^2, however they do so using a simple form (components differ only in their correlation parameter) which takes no account of extreme movements or multi-dimensional tail-structure.

First To Default Swap

A First-to-Default swap (FtD) is an insurance contract against the default of a counterparty within a basket of n counterparties. The contract has provisions for premium payments as well as variable recovery rates. As with a CDS there is a fee payment and a default payment leg that need to be priced.

The FtD will have a sequence of M payable fees $F_p(t_i)$ at times $t_1, t_2, \ldots, t_M$. Given the *joint* survival probability $S(t)$ of the basket, and under the assumption that bond prices are independent of default and survival rates, then the value of this premium leg is

$$V_{pl}(t) = \sum_{m=1}^{M} Z(t, t_m) F_p(t_m) S(t_m)$$

where $Z(t, t_i)$ is the discount factor between times t and t_i.

On the event of the first default in the basket the seller makes a payment equal to the principal times a percentage of value the defaulted obligor's debt. To value the default leg we need to obtain the probability density $d_i(t)$ that obligor i will be the first to default at time t. Given that the swap matures at T:

$$V_{dl}(t) = \sum_{i=1}^{n} \int_t^T F_{d_i}(1 - \delta_i) Z(t, \tau) d_i(\tau) d\tau$$

where δ_i and F_{d_i} are the recovery rate and contingent value associated with the i^{th} obligor.

The overall spread value of the FtD Swap is the ratio of the default to premium leg values. For simplicity we ignore accrued fees compensating instances when defaults occur between payment dates.

It follows from the above that we require an expression for the joint survival function $S(t)$ and the default distribution of the n obligors. Li assumes that the joint survival distribution may be characterised by a Gaussian copula in which the correlation structure is obtained from a set of asset/equity returns. This asset based model implies that an obligor will default if its asset value X_i falls below a barrier $\beta_i(t)$. The margins of the copula are then related to the univarate survival probabilities $F_i(t), i = 1, \ldots, n$.

Typically there will not exist a closed form solution for the joint survival function and in such instances one must resort to numerical techniques, such as Monte Carlo, for their evaluation. Unfortunately the computational overhead of this approach can be prohibitive since it is often necessary to generate thousands or hundreds of thousands (if the probabilities are small) of samples for each trade in order to price with a reasonable level of accuracy. This problem is further compounded in the context of risk management, in which we must re-value under different market and credit scenarios and over different time horizons. Each valuation must therefore have minimum overhead.

For this reason the use of semi-analytic pricing has become a popular tool suitable for the purposes of risk management applications: this method assumes that the returns on each of the aforementioned assets are independent when conditioned on a given value for the market risk factor V which represents the 'state of the economy'. Each asset is correlated to V so that obligors are correlated indirectly. Given a single factor model with $V \in \mathbb{R}$ then the returns on the assets of obligor i are

$$X_i = \rho_i V + \sqrt{1-\rho_i^2}\xi_i$$

where ρ_i is the correlation of the i^{th} obligor to $V \sim N(0,1)$ which is itself independent of $\xi_i \sim N(0,1)$; a factor representing the idiosyncratic risk specific to the obligor.

Conditional on V , the survival probability for obligor i is therefore

$$P_i(t|V) = \Phi\left[\frac{\rho_i V - \beta_i(t)}{\sqrt{1-\rho_i^2}}\right]$$

with the notation $P_i(t|V) \triangleq P(\tau_i > t|V)$ where τ_i denotes this obligor's default time. The conditional survival probability for obligors $i = 1, \ldots, n$ is then given by the product of the marginal survival probabilities:

$$P(\tau_1 > t, \ldots, \tau_n > t|V) = \prod_{i=1}^{n} P_i(t|V).$$

The joint unconditional survival probability is obtained by integrating over all possible states of the economy so that

$$S(t) \triangleq P(\tau_1 > t, \ldots, \tau_n > t) = \int_{-\infty}^{\infty} P(V) \prod_{i=1}^{n} P_i(t|V) dV$$

where

$$P(V) = \frac{1}{\sqrt{2\pi}} e^{\frac{-V^2}{2}}$$

is the probability distribution for V. Putting all this together gives

$$S(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\prod_{i=1}^{n} \Phi\left[\frac{\rho_i V - \beta_i(t)}{\sqrt{1-\rho_i^2}}\right] e^{\frac{-V^2}{2}}\right] dV.$$

For the contingent default payment, a payment of $F_{d_i}(1-\delta_i)$ is made if obligor i is the first to default. Let

$$d_i(t) \triangleq P(t < \tau_1, \ldots, t < \tau_i \le t + dt, \ldots, t < \tau_n)$$

then with respect to the conditional independence framework

$$d_i(t) = \int_{-\infty}^{\infty} \frac{d}{dt}[F_i(t)] \Big|_V \prod_{j=1, j \neq i}^{n} P_j(t|V) P(V) dV$$

where

$$\frac{d}{dt}[F_i(t)]\,|_V = \frac{d}{dt}[F_i(t)] \times \frac{\phi_{X_i|V}(\Phi_{X_i}^{-1}(F_i(t)))}{\phi_{X_i}(\Phi_{X_i}^{-1}(F_i(t)))}.$$

is the differential of the default function evaluated at t and conditioned on V.

Given these expressions for the various probabilities it is now possible to value without recourse simulation and can be evaluated much more rapidly than Monte Carlo methods. We now move on to express the FtD price within our normal mixture copula framework.

Analytic FtD

Since our starting point in the construction of the copula is a normal mixture multivariate having normal mixture univariate margins it follows that

$$\beta_i(t) = \Phi_{NM_i}^{-1}(F_i(t))$$

where $\Phi^{-1}_{NM_i}$ is the inverse of the cumulative normal mixture associated with the i^{th} obligor

$$\Phi_{NM_i}(\cdot) = \sum_{k=1}^{m} w_k \Phi(\cdot; \mu_k(X_i), \sigma_k(X_i)).$$

The joint survival probability $S(t)$ is then given by Equation 2. The instantaneous probability of default of the i^{th} obligor at time t (assuming no simultaneous defaults) when no previous defaults have occurred is

$$\sum_{k=1}^{m} w_k \left[-\frac{dF_i(t)}{dt}\bigg|_k \prod_{j=1, j\neq i}^{n} \Phi\left[\frac{\mu_k(X_j)-\beta_j}{\sigma_k(X_j)}\right]\right],$$

where

$$\frac{dF_i(t)}{dt}\bigg|_k = \frac{\phi_k(\Phi^{-1}_{NM}[F_i(t)])}{\sum_{s=1}^{m} w_s\phi_s(\Phi^{-1}_{NM}[F_i(t)])} \times \frac{dF_i(t)}{dt}$$

is the default density at t conditional on k, with $\Phi_j(\cdot) \triangleq \Phi(\cdot; \mu_{j(X_i)}, \sigma_{j(X_i)})$.

Semi-Analytic FtD

The one-factor conditional independence assumption imposes an inner product structure on the associated correlation matrix. In contrast we only require that each component correlation has such a structure, whereby we assume

$$\rho_k = I_n + (\rho\rho^\top - \textbf{trace}(\rho\rho^\top))$$

for all $k = 1, \ldots, m$ where $\rho \in \mathbb{R}^n$ with $\rho_i \in [-1, 1]$. Generally speaking this will allow a greater range of correlation structures than previously available within this framework since we optimise the bivariate tails to match the empirical target correlations.

To derive the pricing formulae we condition on both k and V to obtain the following semi-analytic expression for the joint survival probability $S(t)$

$$\sum_{k=1}^{m} w_k \int_{-\infty}^{\infty} \prod_{i=1}^{n} \Phi\left[\frac{\mu_{k(X_i)} + \rho_i\sigma_{k(X_i)}V - \beta_i}{\sqrt{(1-\rho_i^2)}\sigma_{k(X_i)}}\right] e^{\frac{-V^2}{2}} dV$$

where $\mu_{k(X_i)}$ and $\sigma_{k(X_i)}$ are the mean and standard deviation of the k^{th} component normal along the i^{th} dimension, with β_i the barrier for the i^{th} obligor as before.

The first to default density at time t is

$$\sum_{k=1}^{m} w_k \int_{-\infty}^{\infty} \frac{dF(t)}{dt}\bigg|_{\{V,k\}} \prod_{j=1, j\neq i}^{n} \Phi\left[\frac{\mu_{k(X_j)} + \rho_i\sigma_{k(X_j)}V - \beta_j}{\sqrt{(1-\rho_j^2)}\sigma_{k(X_j)}}\right] e^{\frac{-V^2}{2}} dV$$

where

$$\frac{dF_i(t)}{dt}\bigg|_{\{V,k\}} = \frac{dF_i(t)}{dt}\bigg|_k \frac{\phi_{X_i|V}(\Phi^{-1}_{X_i}(F_{i|k}(t)))}{\phi_{X_i}(\Phi^{-1}_{X_i}(F_{i|k}(t)))}$$

with $F_{i|k}(t) = \Phi_{NM}(\Phi_k^{-1}[F_i(t)])$ the default distribution of the i^{th} obligor conditional on k. Figure 6 illustrates the effect of tail-dependency and skewness (of the univariate margins) on the FtD value for a simple basket of 2 obligors. Here $m = 4$, $\rho_k = 0$, a single fee of unit value is payable at T if no default occurs, the recovery rate for both obligors is zero as is the risk-free rate. Under these assumptions the value reduces to $\frac{1-S(T)}{10000\times S(T)}$. One may visualise $S(T)$ as that associated with the upper right quadrant of the bivariate mixture bisected vertically and horizontally

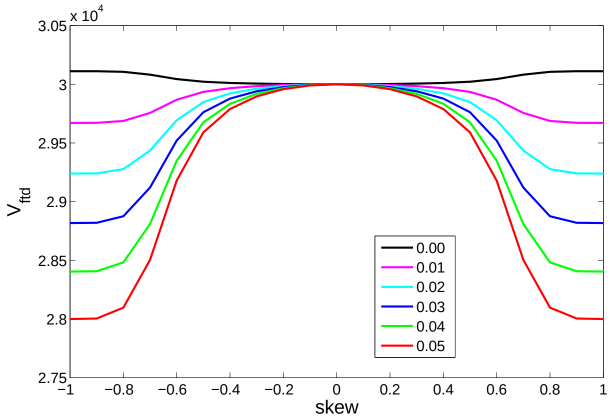

Figure 6: FtD value as a function of skewness and tail-probability $\tau \in [0, 0.05]$.

by the barriers $\beta_i; i = 1, 2$. We note that the bivariate is itself reflected about the axis anti-diagonal in accordance with the sign of the skew value s so that $S(T)(s) \equiv P(\tau_1 < T, \tau_2 < T)(-s)$.

Negative skew implies large negative movements in each univariate so that, all other things being equal, increasing τ reduces the likelihood of individual large negative movements, consequently increasing the probability of a joint default and joint survival. This latter point is significant for a FtD giving rise to an overall reduction of its value. As skew increases in magnitude this effect becomes more pronounced, hence the value decreases with respect to it.

Since we have chosen hazard rates such that the probability of survival of each individual obligor at $T = 1$ is 0.5 the thresholds will be located at the median value of the each univariate which gives rise to joint survival and default probabilities that are approximately equal; thus explaining the price symmetry about the zero skew axis. In general this will not be the case.

Conclusions

Recent market events have served to highlight the disasterous consequences of inadequate in-house risk management systems and practices. Having paid severely for under-estimating risk in the good times, it seems entirely possible that financial institutions could destabilise the global economy by over-compensating for it in the bad. As such the need for an accurate and robust model of credit risk is of critical importance.

In this paper we have put forward a generalised method for the pricing of credit derivatives in a tail-consistent framework. Our method makes use of a hyper-cubiod mixture copula that extends popular asset-based factor models. We have explored the difficulties posed regarding parameterisation in a non-normal context and devised herein an efficient means in which to identify tail dependencies within high dimensional systems. Identification of such structures is of particular concern if one is to characterise the process by which large shocks propagate within and between markets. As such in later work we will consider a dynamic extention to our approach which makes use of auto-correlation data to inform a Hidden Markov process for the time-evolution of the tails. In this way we will permit a pricing mechanism for credit instruments in a contagious world.

References

[1] J. -ph. Bouchaud, *Elements for a theory of financial risks*, Physica A, Vol. 263, (1999).

[2] J. Hull and A. White, *Valuing Credit Default Swaps II: Modelling Default Correlations*, Journal of Derivatives, Vol. 8, No. 3, (Spring 2001), pp. 12-22.

[3] D. X. Li, *On Default Correlation: A Copula Function Approach*, Journal of Fixed Income, Vol 9, (2001) pp. 43-54.

[4] D. X. Li and M. Liang *CDO^2 Pricing Using Gaussian Mixture Model with Transformation of Loss Distribution.*, Working Paper, Barclays Capital (September 2005). Source: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=890766

[5] S. Ray and B. Lindsay, *The topography of multivariate normal mixtures*, The Annals of Statistics, (2005) Vol 33, No 5. pp 2042-2065.

[6] J. Wang, *Generating daily changes in market variables using a multivariate mixture of normal distributions*, Proceedings of the 33rd winter conference on simulation, pp 283-289, Washington DC, USA, (2001). IEEE Computer Society.

Acknowledgements

The authors are deeply grateful for the guidance and support of Malcolm Carmichael, Ellen Stars and Flick Roper. We would also like to thank Dr. Simon Marshall for his assistance in coding the pattern analysis routine and also to Yimin Liu and Robert Hone for their constructive feedback on earlier drafts.