Algebraic methods for correlated assets

Jason Morton
jason@math.stanford.edu

Stanford University
Joint work-in-progress with I. Filiz, X. Guo and B. Sturmfels

April 11, 2008
Practitioners' Financial Mathematics Seminar
University of Minnesota

1 Introduction

2 Statistical Dependence

3 Graphical models

4 Graphical models for correlated default

My Perspective on Financial Mathematics

Mathematics:

- Mathematician studying geometry (algebraic, polyhedral, computational) of statistical dependence

Financial:

- Former M&A banker at Credit Suisse (First Boston)
- Manage a nonprofit endowment
- Consultant to fund-of-hedge-funds

1 Introduction

2 Statistical Dependence

3 Graphical models

4 Graphical models for correlated default

Statistical Dependence Among Assets

- How do the variables (esp. price, return) associated with assets depend on each other? We must make simplifying assumptions here.
- Gross structure: are X and Y
  - Correlated $\rho(X, Y) \neq 0$?
  - Independent $X \perp\!\!\!\perp Y$, $P(x, y) = P(x)P(y)$?
  - Independent given Z, $X \perp\!\!\!\perp Y|Z$, $P(x, y|z) = P(x|z)P(y|z)$?
- Fine structure: given that X and Y are not independent or correlated, how do they depend on each other? What is the dimension, i.e. how many parameters should we use to measure it? How can we understand the model intuitively?

Choosing Dependence Structures

- The space of possible dependence structures is very large!
- We need to make a simplifying assumption that is
  - **Tractable**: we can analyze it mathematically
  - **Interpretable**: humans can grasp what it means
  - **Estimable**: we can fit a model
  - **Flexible**: flexible enough to fit reality, but no more flexible than that
- Working definition of simple: has geometry
- In the vast majority of cases, the chosen structure is some variation on the multivariate normal distribution (Markowitz asset allocation, Black-Sholes option pricing, normal copulas, etc. - they are the workhorse of finance)
- But a lot of money has been lost assuming multivariate normality. Multi-sigma events happen all the time!

"We were seeing things that were 25-standard deviation moves, several days in a row," David Viniar, Goldman CFO, on losses in their Global Alpha fund, August 2007

Choosing Dependence Structures

- The space of possible dependence structures is very large!
- We need to make a simplifying assumption that is
  - **Tractable**: we can analyze it mathematically
  - **Interpretable**: humans can grasp what it means
  - **Estimable**: we can fit a model
  - **Flexible**: flexible enough to fit reality, but no more flexible than that
- Working definition of simple: has geometry
- In the vast majority of cases, the chosen structure is some variation on the multivariate normal distribution (Markowitz asset allocation, Black-Sholes option pricing, normal copulas, etc. - they are the workhorse of finance)
- But a lot of money has been lost assuming multivariate normality. Multi-sigma events happen all the time!

"We were seeing things that were 25-standard deviation moves, several days in a row," David Viniar, Goldman CFO, on losses in their Global Alpha fund, August 2007

Choosing Dependence Structures

- The space of possible dependence structures is very large!
- We need to make a simplifying assumption that is
  - **Tractable**: we can analyze it mathematically
  - **Interpretable**: humans can grasp what it means
  - **Estimable**: we can fit a model
  - **Flexible**: flexible enough to fit reality, but no more flexible than that
- Working definition of simple: has geometry
- In the vast majority of cases, the chosen structure is some variation on the multivariate normal distribution (Markowitz asset allocation, Black-Sholes option pricing, normal copulas, etc. - they are the workhorse of finance)
- But a lot of money has been lost assuming multivariate normality. Multi-sigma events happen all the time!

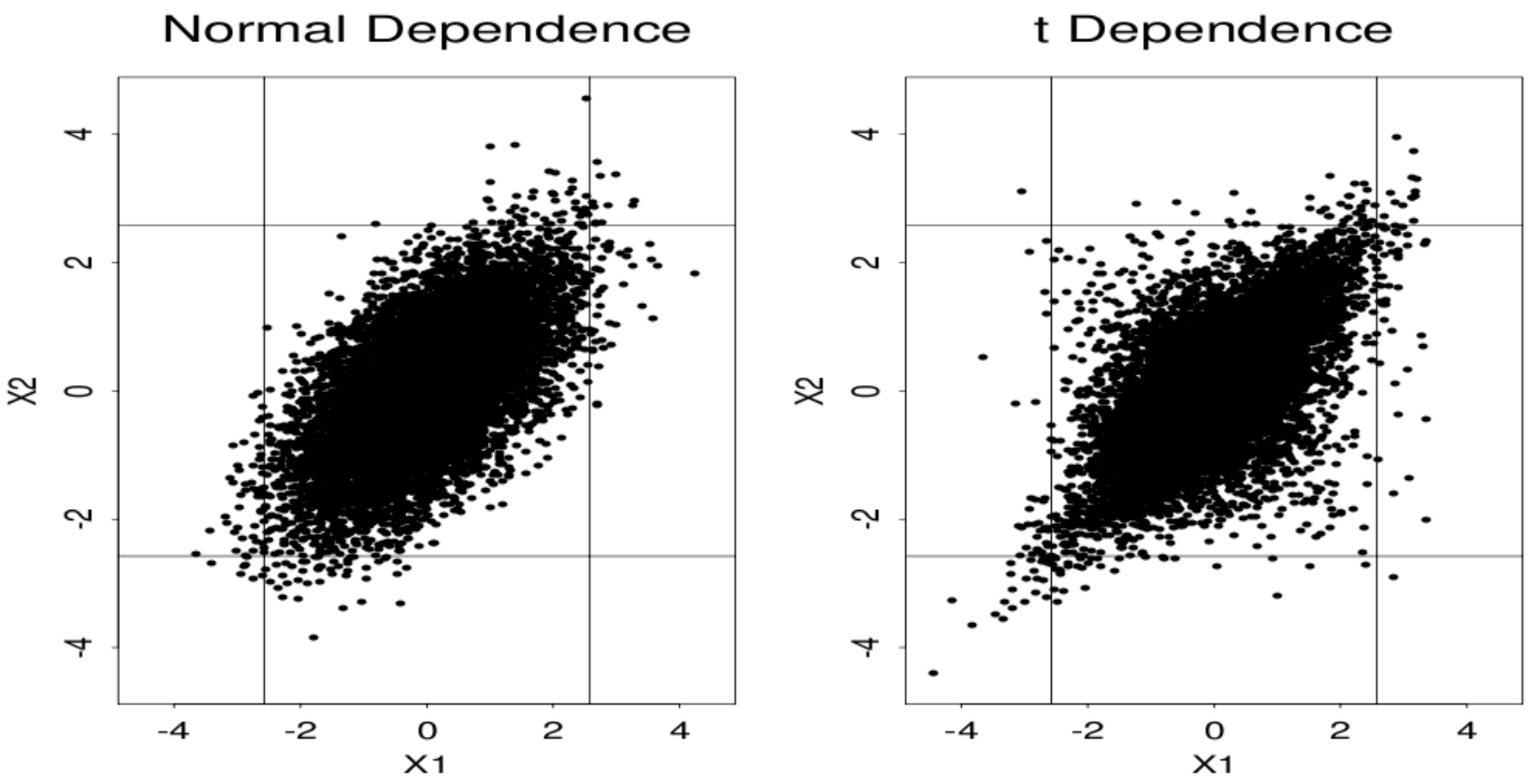
"We were seeing things that were 25-standard deviation moves, several days in a row," David Viniar, Goldman CFO, on losses in their Global Alpha fund, August 2007

Choosing Dependence Structures

- The space of possible dependence structures is very large!
- We need to make a simplifying assumption that is
  - **Tractable**: we can analyze it mathematically
  - **Interpretable**: humans can grasp what it means
  - **Estimable**: we can fit a model
  - **Flexible**: flexible enough to fit reality, but no more flexible than that
- Working definition of simple: has geometry
- In the vast majority of cases, the chosen structure is some variation on the multivariate normal distribution (Markowitz asset allocation, Black-Sholes option pricing, normal copulas, etc. - they are the workhorse of finance)
- But a lot of money has been lost assuming multivariate normality. Multi-sigma events happen all the time!

"We were seeing things that were 25-standard deviation moves, several days in a row," David Viniar, Goldman CFO, on losses in their Global Alpha fund, August 2007

Normal vs. t distribution



[Frey, McNeil, and Nyfeler 2001]

- Above have identical mean, covariance, and correlation (0.7)
- One promising direction is to study elliptical distributions and alternate copulas

Choosing Dependence Structures

- We want to talk about some other simplifying choices of dependence structure that come from geometry, algebra, and topology.
- On a good day, these are also Tractable, Interpretable, Estimable, and Flexible*
- Intuition:
  - while the space of possible dependence structures is very large,
  - the dimension of the space of possible dependencies is not
- Model this intuition and accomplish dimension reduction with geometry: graphs, linear spaces, varieties, manifolds, topological spaces
- Beyond fat tails - explore dependence space
  - e.g. multimodality is appropriate in some cases

Conditinal independence and multivariate normal geometry

These make extensive use of linear spaces.

- $X \sim N(0, \Sigma)$ iff there is an embedding of vectors v^i, one for each X_i in $\mathbb{R}^n$, so that $\sigma_{ij} = \langle v_i | v_j \rangle$.
- don't really need n dimensions, just $r << n$
- the n variables are conditionally independent given the r factors
- $X_i \perp\!\!\!\perp X_j | X_K$ if the projections of v_i and v_j to $K^\perp$ are orthogonal; algebraically the almost principal minor $\det \Sigma_{iK,jK}$ vanishes
- More complicated conditional independence models are also possible. These are also called covariance selection models.
- Most easily interpreted among these are the graphical models

1 Introduction

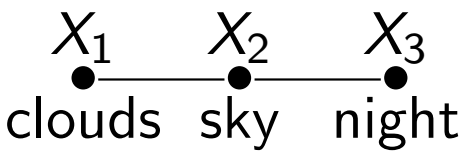
2 Statistical Dependence

3 Graphical models

4 Graphical models for correlated default

Graphical models in General

Graphical models describe the *dependence structure among random variables* with a *graph* $G = (V, E)$, together with parameters associated to each clique in the graph.



- Lack of an edge denotes conditional independence; graph separation and conditional independence coincide

Undirected graphical models

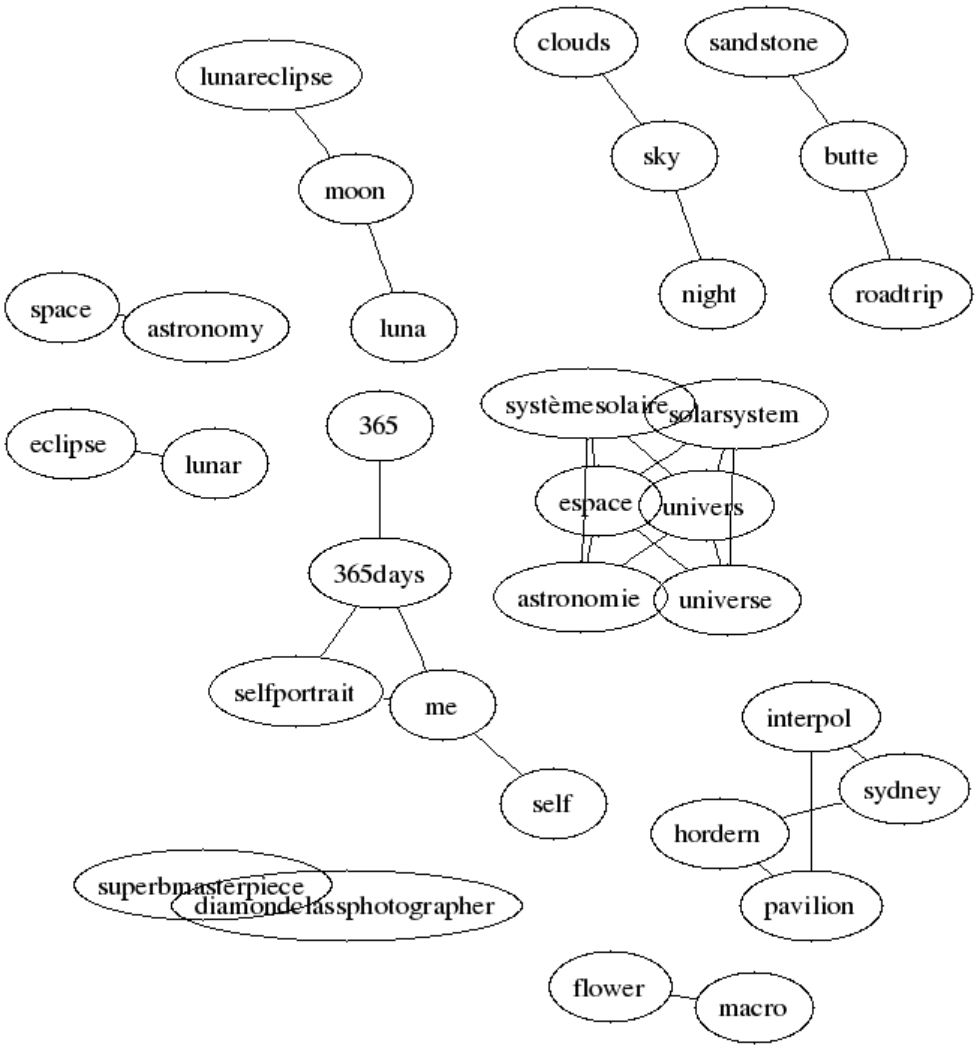
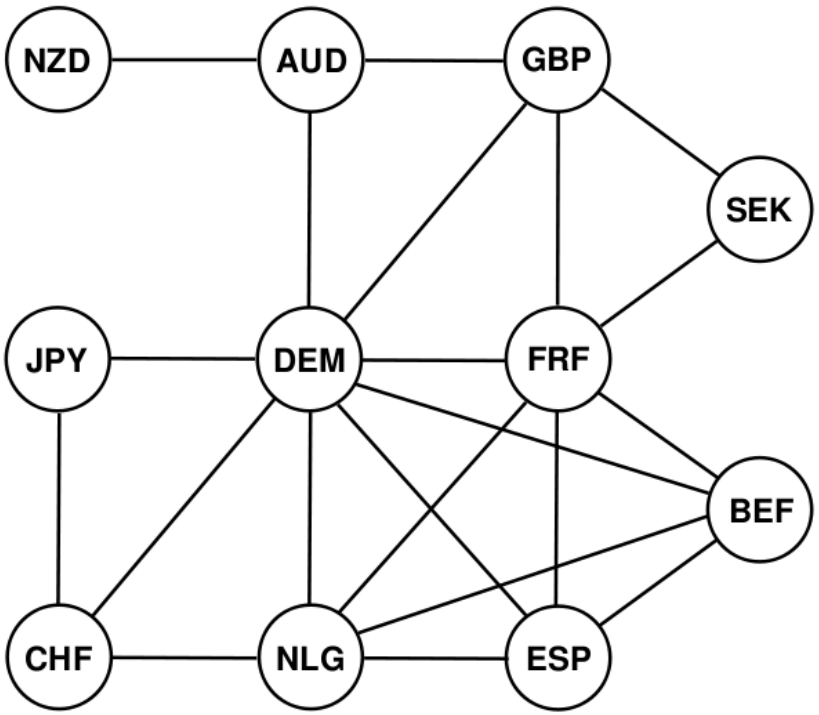


Figure: Part of a Binary model for photo tags on flickr with 2,786 tags and 500 observations

Gaussian Graphical models

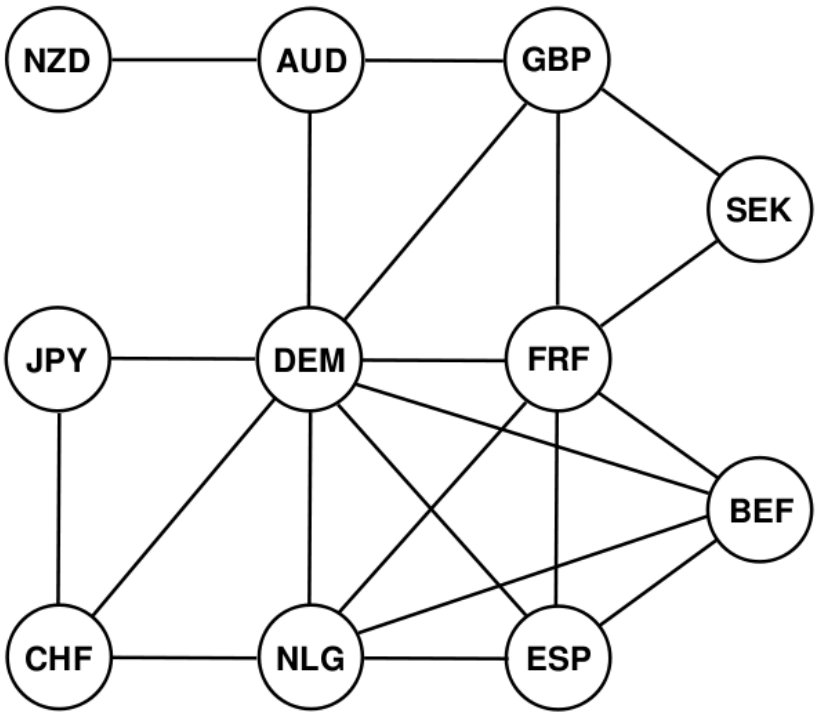
- Undirected graphical models (covariance selection models) for p-multivariate Gaussian random variables.



[Carvalho and West 2006]

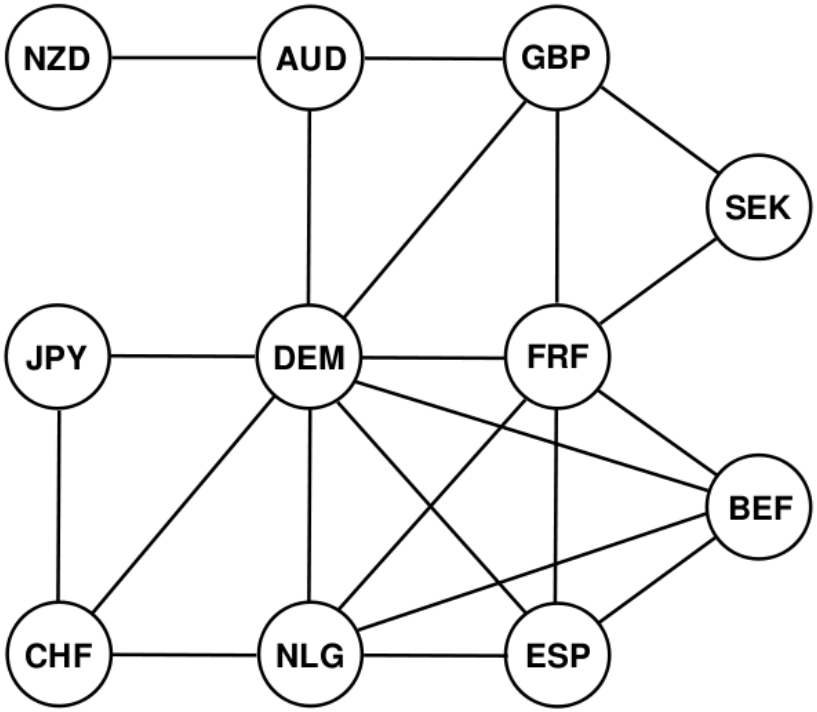
Gaussian Graphical models

- Undirected graphical models (covariance selection models) for p-multivariate Gaussian random variables.



[Carvalho and West 2006]

Gaussian Graphical models



Combinatorial structure can be described with conditional independence statements:

$$X_1 \perp\!\!\!\perp X_3 | X_2 \text{ i.e. } P(x_1, x_3 | x_2) = P(x_1 | x_2) P(x_3 | x_2)$$

these CI models are interesting mathematical systems (semigraphoids). CI statements and factorization duality.

1 Introduction

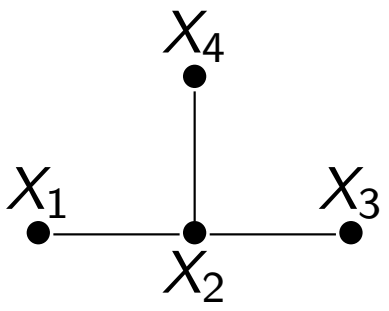
2 Statistical Dependence

3 Graphical models

4 Graphical models for correlated default

Graphical model for default - one period

- Graph (V, E)



- Firm represented by a random variable at each node in the graph
- $X_i = 1$ if firm is in default, $X_i = 0$ otherwise
- Nodes have a parameter η_u that adjusts credit quality
- Edges (u, v) have a parameter $\eta_{u,v}$ that represents the strength of the relationship
  - same as logistic regression parameters
- Graph(ical) models, Ising models, Markov random fields, closely related to logistic regression models

Graphical model for default - one period

- One node per firm with associated binary random variable X_i, $(i \in V)$ with $X_i = 1$ for default and $X_i = 0$ for survival.
- Parameters $\eta_i \in \mathbb{R}$ for each node and $\eta_{uv} \in \mathbb{R}$ for each edge

Definition

The joint probability distribution of the random variable $X := (X_1, \ldots, X_M)$ *is given by*

$$P(X = w) = \frac{1}{Z} \cdot \exp\left(\sum_{i \in V} \eta_i w_i + \sum_{(u,v) \in E} \eta_{uv} w_u w_v \right)$$

Here $w = (w_1, \ldots, w_M)$ runs over $\{0,1\}^M$, Z is the normalization constant or *partition function* defined by

$$Z = \sum_{w \in \{0,1\}^M} \exp\Big(\sum_{i \in V} \eta_i w_i + \sum_{(u,v) \in E} \eta_{uv} w_u w_v \Big).$$

Graphical model for default - one period

- Sufficient statistics are the marginals for single firm default rates (e.g. from spreads, credit ratings, or historical rates) and pairwise correlations / co-occurance
- Toric models - have existence and uniqueness of the MLE fixing single (default rates) and pairwise (default correlations) marginals
- Commutative algebra and Gröbner bases theory tells us how to sample from outcomes with these marginals to do MCMC (Diaconis and Sturmfels, Develin and Sullivant)
  - Is the observed distribution exceptional (e.g. by agreeing with the model) among those with consistent marginals?
  - Price contingent claims

Graphical model for default - one period

- Sufficient statistics are the marginals for single firm default rates (e.g. from spreads, credit ratings, or historical rates) and pairwise correlations / co-occurance
- Toric models - have existence and uniqueness of the MLE fixing single (default rates) and pairwise (default correlations) marginals
- Commutative algebra and Gröbner bases theory tells us how to sample from outcomes with these marginals to do MCMC (Diaconis and Sturmfels, Develin and Sullivant)
  - Is the observed distribution exceptional (e.g. by agreeing with the model) among those with consistent marginals?
  - Price contingent claims

Graphical model for default - one period

- Sufficient statistics are the marginals for single firm default rates (e.g. from spreads, credit ratings, or historical rates) and pairwise correlations / co-occurance
- Toric models - have existence and uniqueness of the MLE fixing single (default rates) and pairwise (default correlations) marginals
- Commutative algebra and Gröbner bases theory tells us how to sample from outcomes with these marginals to do MCMC (Diaconis and Sturmfels, Develin and Sullivant)
  - Is the observed distribution exceptional (e.g. by agreeing with the model) among those with consistent marginals?
  - Price contingent claims

Graphical model for default - one period

- Sufficient statistics are the marginals for single firm default rates (e.g. from spreads, credit ratings, or historical rates) and pairwise correlations / co-occurance
- Toric models - have existence and uniqueness of the MLE fixing single (default rates) and pairwise (default correlations) marginals
- Commutative algebra and Gröbner bases theory tells us how to sample from outcomes with these marginals to do MCMC (Diaconis and Sturmfels, Develin and Sullivant)
  - Is the observed distribution exceptional (e.g. by agreeing with the model) among those with consistent marginals?
  - Price contingent claims

Existence and uniqueness of the MLE

Data for assessment of default correlation is usually assumed to identify the following two sets of sufficient statistics:

- The *marginal default probability* $P(X_i = 1)$ is assumed to be known for each firm $i \in V$.
- The *pairwise linear correlation*

$$\rho_{uv} = \frac{P(X_u = X_v = 1) \; - \; P(X_u{=}1) \cdot P(X_v{=}1)}{\sqrt{P(X_u{=}1) \cdot (1 - P(X_u{=}1)) \cdot P(X_v{=}1) \cdot (1 - P(X_v{=}1))}}$$

  for all pairs of firms u and v that share an edge in E.

Existence and uniqueness of the MLE

Knowledge of the marginal default probabilities and the pairwise linear correlations is equivalent to knowledge of the following set of $N + |E|$ sufficient statistics:

- Single node marginals $P_i := P(X_i = 1)$ for all $i \in V$
- Double node marginals $P_{uv} := P(X_u = X_v = 1)$ for all $(u, v) \in E$.

The $N+|E|$ sufficient statistics uniquely specify the $N+|E|$ model parameters η_i and η_{uv}.

Theorem

The single and double node marginals P_i and P_{uv} arising from the graphical model, for some choice of parameters η_i and η_{uv}, are characterized by a finite set of linear inequalities. If these linear inequalities are satisfied then the model parameters can be recovered uniquely from the single and double node marginals.

Existence and uniqueness of the MLE

Knowledge of the marginal default probabilities and the pairwise linear correlations is equivalent to knowledge of the following set of $N + |E|$ sufficient statistics:

- Single node marginals $P_i := P(X_i = 1)$ for all $i \in V$
- Double node marginals $P_{uv} := P(X_u = X_v = 1)$ for all $(u, v) \in E$.

The $N+|E|$ sufficient statistics uniquely specify the $N+|E|$ model parameters η_i and η_{uv}.

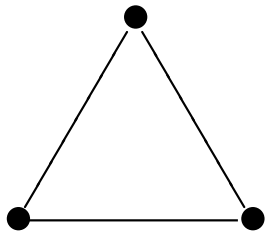
Theorem

The single and double node marginals P_i and P_{uv} arising from the graphical model, for some choice of parameters η_i and η_{uv}, are characterized by a finite set of linear inequalities. If these linear inequalities are satisfied then the model parameters can be recovered uniquely from the single and double node marginals.

Moment map of toric geometry

As a Toric Model

Let G be the triangle



$+$

- with $V = \{1, 2, 3\}$ and
- $E = \big\{\{1, 2\}, \{1, 3\}, \{2, 3\}\big\}$.
- The marginals from our model are characterized by the following 16 linear equalities:

$$P_i \geq P_{ij} \geq 0 \text{ for all } i, j,$$
$$P_1 + P_2 \leq P_{12} + 1\,,\ P_1 + P_3 \leq P_{13} + 1\,,\ P_2 + P_3 \leq P_{23} + 1\,,$$
$$P_1 + P_{23} \leq P_{12} + P_{13}\,,\ P_2 + P_{13} \leq P_{12} + P_{23}$$
$$P_3 + P_{12} \leq P_{13} + P_{23}\,, P_1 + P_2 + P_3 \leq P_{12} + P_{13} + P_{23} + 1.$$

As a Toric Model

The solution set of these 16 linear inequalities in the six-dimensional *cyclic polytope* $C(6,8)$ which is the convex hull of the columns of the matrix

$$\begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_{12} \\ P_{13} \\ P_{23} \end{array} \begin{array}{c} \begin{array}{cccccccc} p_{000} & p_{001} & p_{010} & p_{011} & p_{100} & p_{101} & p_{110} & p_{111} \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \tag{1}$$

If these inequalities are satisfied then there exist unique parameters $\eta_1, \eta_2, \eta_3,\ \eta_{12}, \eta_{13},\ \eta_{23}$ such that the distribution (1) has the given marginals $P_1, P_2, P_3, P_{12}, P_{13}, P_{23}$.

Estimating the graph

Now we know that given the graph, we can at least theoretically find parameters to fit the data.

But given some data, we also need to learn the graph, as well as the parameters. This is hard:

- the richest model is the $(p-1)$-simplex (or all edges), so we need some sparsity constraints or regularization strategy
- the number of graphs or simplicial complexes grows quickly
- evaluating the partition function (and so the likelihood) is hard

Estimating the graph

Now we know that given the graph, we can at least theoretically find parameters to fit the data.

But given some data, we also need to learn the graph, as well as the parameters. This is hard:

- the richest model is the $(p-1)$-simplex (or all edges), so we need some sparsity constraints or regularization strategy
- the number of graphs or simplicial complexes grows quickly
- evaluating the partition function (and so the likelihood) is hard

Estimating the graph

Now we know that given the graph, we can at least theoretically find parameters to fit the data.

But given some data, we also need to learn the graph, as well as the parameters. This is hard:

- the richest model is the $(p-1)$-simplex (or all edges), so we need some sparsity constraints or regularization strategy
- the number of graphs or simplicial complexes grows quickly
- evaluating the partition function (and so the likelihood) is hard

Structure Learning in Graphical Models

There are many approaches to this structure learning or model selection problem, including

- searching in model space, evaluating the likelihood (e.g. greedily attaching explanatory edges)
- constraint-based approaches that test conditional independence (algebraic statistics provides new invariants that may be helpful here)
- Bayesian approaches (0-centered prior on the simplex parameters), and
- add edge-penalizing term to the maximum likelihood problem.

Structure Learning in Graphical Models

There are many approaches to this structure learning or model selection problem, including

- searching in model space, evaluating the likelihood (e.g. greedily attaching explanatory edges)
- constraint-based approaches that test conditional independence (algebraic statistics provides new invariants that may be helpful here)
- Bayesian approaches (0-centered prior on the simplex parameters), and
- add edge-penalizing term to the maximum likelihood problem.

Structure Learning in Graphical Models

There are many approaches to this structure learning or model selection problem, including

- searching in model space, evaluating the likelihood (e.g. greedily attaching explanatory edges)
- constraint-based approaches that test conditional independence (algebraic statistics provides new invariants that may be helpful here)
- Bayesian approaches (0-centered prior on the simplex parameters), and
- add edge-penalizing term to the maximum likelihood problem.

Structure Learning in Graphical Models

There are many approaches to this structure learning or model selection problem, including

- searching in model space, evaluating the likelihood (e.g. greedily attaching explanatory edges)
- constraint-based approaches that test conditional independence (algebraic statistics provides new invariants that may be helpful here)
- Bayesian approaches (0-centered prior on the simplex parameters), and
- add edge-penalizing term to the maximum likelihood problem.

Structure Learning in Graphical Models

There are many approaches to this structure learning or model selection problem, including

- searching in model space, evaluating the likelihood (e.g. greedily attaching explanatory edges)
- constraint-based approaches that test conditional independence (algebraic statistics provides new invariants that may be helpful here)
- Bayesian approaches (0-centered prior on the simplex parameters), and
- add edge-penalizing term to the maximum likelihood problem.

Structure Learning in Graphical Models

Most model selection methodologies, in attempting to decide which simplices to include, introduce a sparsity parameter t.

- This parameter might be the weight on the penalty term, a threshold α, or in the Bayesian case a parameter in the prior.
- It is sometimes trained on a held-out data set or its correct asymptotic order is found analytically.

Particularly interesting are ℓ_1-regularized methods

$$\min_{\eta} -\ell(\eta) + t\|\eta\|_1$$

where $-\ell(\eta)$ is the negative log likelihood of the data and η the parameter vector. These methods can have good sample complexity.

To keep things simple and get some analytical results, we'll assume a sector model, where each firm connects only to one sector or market and the sectors or markets are linked.

Structure Learning in Graphical Models

Most model selection methodologies, in attempting to decide which simplices to include, introduce a sparsity parameter t.

- This parameter might be the weight on the penalty term, a threshold α, or in the Bayesian case a parameter in the prior.
- It is sometimes trained on a held-out data set or its correct asymptotic order is found analytically.

Particularly interesting are ℓ_1-regularized methods

$$\min_{\eta} -\ell(\eta) + t\|\eta\|_1$$

where $-\ell(\eta)$ is the negative log likelihood of the data and η the parameter vector. These methods can have good sample complexity.

To keep things simple and get some analytical results, we'll assume a sector model, where each firm connects only to one sector or market and the sectors or markets are linked.

Structure Learning in Graphical Models

Most model selection methodologies, in attempting to decide which simplices to include, introduce a sparsity parameter t.

- This parameter might be the weight on the penalty term, a threshold α, or in the Bayesian case a parameter in the prior.
- It is sometimes trained on a held-out data set or its correct asymptotic order is found analytically.

Particularly interesting are ℓ_1-regularized methods

$$\min_{\eta} -\ell(\eta) + t\|\eta\|_1$$

where $-\ell(\eta)$ is the negative log likelihood of the data and η the parameter vector. These methods can have good sample complexity.

To keep things simple and get some analytical results, we'll assume a sector model, where each firm connects only to one sector or market and the sectors or markets are linked.

Graphical model - sector version

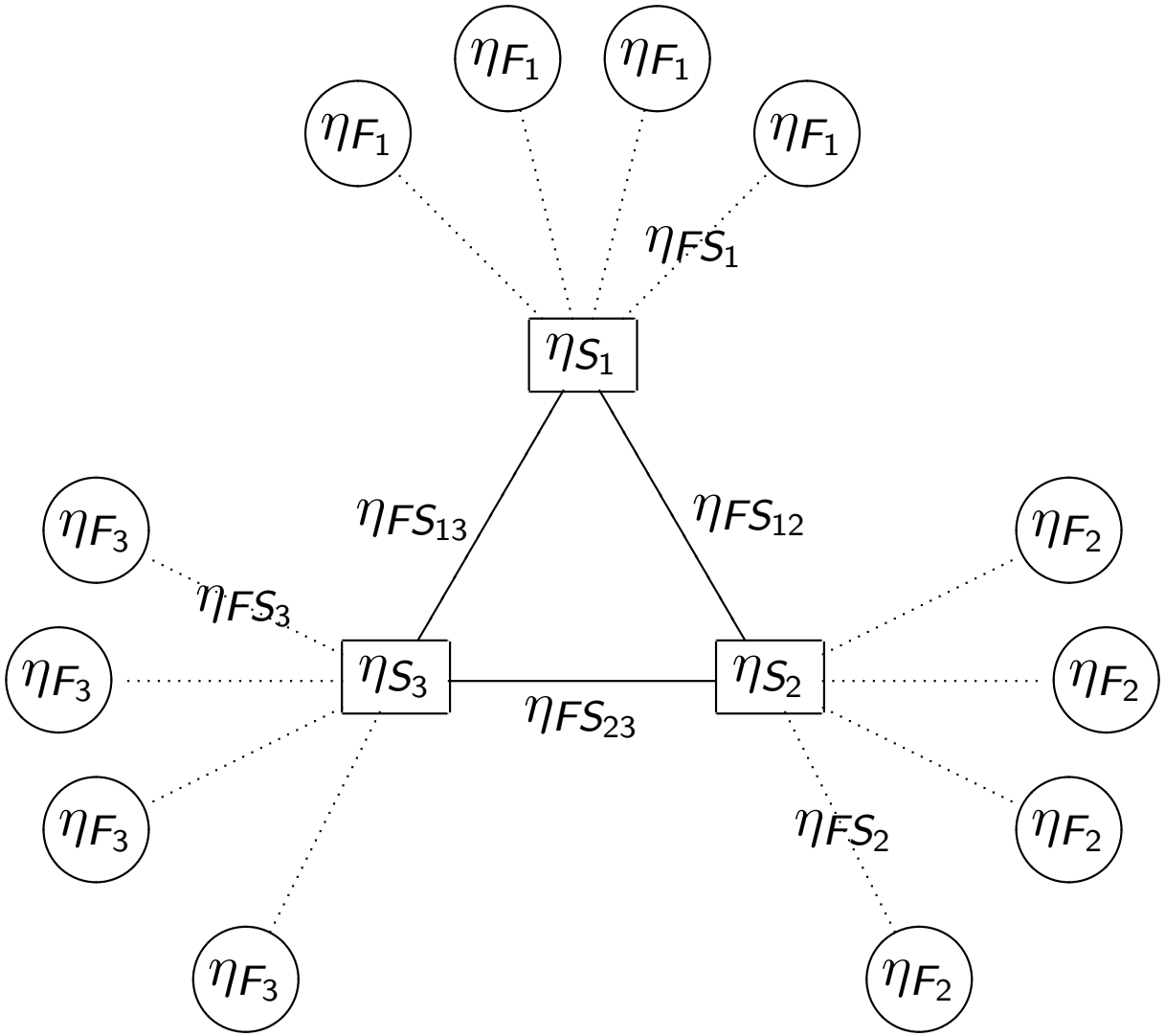
We specify a sector-firm structure for the graph.

- Set $M = N + S$, where nodes $1, \ldots, N$ represent individual firms and $N+1, \ldots, N+S$ represent individual industry sectors
- The joint probability distribution for $(X_1, \cdots, X_N)$ is given by marginalizing over the hidden sector variables:

$$P(x_1, \cdots, x_N) := \sum_{s \in \{0,1\}^S} Q(\quad X_1 = x_1, \cdots, X_N = x_N, \quad X_{N+1} = s_1, \cdots, X_{N+S} = s_{N+S}).$$

Where Q is the distribution from the fully observed graph model.

Graphical model - sector version



Graphical model - sector version

Simplifying assumptions:

- Each firm belongs to one sector $j = 1, 2, \cdots, S$
- Each firm node i has a single edge, which connects to its respective sector node.
- For any particular sector node j, all firm nodes that connect to it have the same node weight η_{F_j} and same edge weight η_{FS_j}.
- Sector nodes are allowed to have different node weights η_{S_j}, and they can connect to each other with different edge weights $\eta_{N+u,N+v}$.
- In the spirit of copula and latent variable models that assume uniform pairwise correlation

Graphical model - sector version

Our distribution now reduces to

$$P(x_1, \ldots, x_N) = \frac{1}{Z_S} \sum_{s \in \{0,1\}^S} \exp\left(\sum_{j=1}^{S} s_j \eta_{S_j} + s_j n_j \eta_{FS_j} + n_j \eta_{F_j} \right)$$
$$\exp\left(\sum_{(u,v):u,v \in \{1,\cdots,S\}} s_u s_v \eta_{N+u,N+v} \right)$$

where $n_j := \sum_{i:\eta_{i,N+j} \neq 0} x_i$ is the number of defaulting firms in sector j.

Graphical model - sector version

Proposition

For this model, the distribution of the number of defaults, $P\left(\sum_{i=1}^{N} X_i = n\right) =$, *is given by*

$$\frac{1}{Z_S} \sum_{(n_1,\ldots,n_S):\sum_j n_j = n} \sum_{s\in\{0,1\}^S} \exp\left(\sum_{u,v\in[S]} s_u s_v \eta_{N+u,N+v}\right)$$
$$\prod_j \binom{N_j}{n_j} \exp\left(s_j \eta_{S_j} + s_j n_j \eta_{FS_j} + n_j \eta_{F_j}\right),$$

where the normalizing constant Z_S *is*

$$\sum_{s\in\{0,1\}^S} \exp\left(\sum_{j=1}^{S} s_j \eta_{S_j} + \sum_{u,v\in[S]} s_u s_v \eta_{uv}\right) \prod_{j=1}^{S} \left(1 + e^{\eta_{F_j} + s_j \eta_{FS_j}}\right)^{N_j}.$$

Graphical model - sector version

This probability distribution on the number of defaults can be written as a mixture of binomial random variables.
When $S = 1$,

$$P\left(X_1 = x_1, \cdots, X_N = x_N\right) = \frac{1}{Z_1}\left(e^{\eta_F \sum_i x_i} + e^{\eta_S + (\eta_{FS} + \eta_F)\sum_i x_i}\right)$$

$$P\left(\sum_{i=1}^{N} X_i = n\right) = \frac{1}{Z_1}\binom{N}{n}\left[e^{n\eta_F} + e^{\eta_S + n\eta_F + n\eta_{FS}}\right]$$

$$Z_1 = \left(1 + e^{\eta_F}\right)^N + e^{\eta_S}\left(1 + e^{\eta_F + \eta_{FS}}\right)^N$$

Graphical model - sector version

Proposition

When $S = 1$,

$$\sum_{i=1}^{N} X_i \stackrel{d}{=} YB_1 + (1-Y)B_2,$$

where

$$\begin{aligned} Y &\sim Bernoulli\left(\frac{e^{\eta_S}\left(1+e^{\eta_F}e^{\eta_{FS}}\right)^N}{Z_1}\right), \\ B_1 &\sim Binomial\left(\frac{e^{\eta_F}e^{\eta_{FS}}}{1+e^{\eta_F}e^{\eta_{FS}}}, N\right), \\ B_2 &\sim Binomial\left(\frac{e^{\eta_F}}{1+e^{\eta_F}}, N\right), \end{aligned}$$

and Y, B_1, B_2 *are independent.*

Graphical model - sector version

This means that

- Tails of the loss distribution can be adjusted by shifting the binomial random variable with higher mean.
- The loss distribution may be bimodal. Bimodality can be explained by a contagion effect among firms. Having a high number of defaults may make it more likely for neighbouring firms to default.

Graphical model - multiperiod version - one sector

Each period:

- Some firms newly default based on the model (conditioned on state of linked firms/sectors)
- Some firms emerge from default / are restructured/ leave the model with probability p_R
- Compute a Markov chain with states No. of defaults, transition probabilities from model
- CDO prices from the Markov chain

Multiperiod Markov Chain

- Model is then a two-state discrete time Markov chain, as Number of Defaults, Number Remaining $= (D_{t_{k+1}}, N_{t_{k+1}})$ only depends on (D_{t_k}, N_{t_k}). Let CD_t is the number of firms among D_{t-1} that have not been removed.
- An entry of the transition matrix $P_{(D_{t_k}, N_{t_k}) \to (D_{t_{k+1}}, N_{t_{k+1}})}$ of the Markov chain is given by:

$$
\begin{aligned}
= \quad & \binom{CD_{t_k}}{N_{t+k} - N_{t_{k+1}}} p_R^{N_{t_k} - N_{t_{k+1}}} (1 - p_R)^{D_{t_k} - N + N_{t_{k+1}}} \\
& P(D_{t_{k+1}} - D_{t_k}; N_{t_k} - CD_{t_k}, \eta_S + CD_{t_k}\eta_{FS}, \eta_{FS}, \eta_F)
\end{aligned}
$$

with $D_0 = 0, N_0 = N$ and
$N \geq D_{t_{k+1}} \geq D_{t_k} \geq 0 \text{ and } N \geq N_{t_k} \geq N_{t_{k+1}} \geq 0.$

Basic CDO Pricing

A Collateralized Debt Obligation (CDO) is a portfolio of defaultable instruments (loans, credits, bonds or default swaps), whose credit risk is sold to investors who agree to bear the losses in the portfolio, in return for a periodic payment.

- A CDO is usually sold in *tranches*, which are specified by their *attachment points* K_L and *detachment points* K_U as a percentage of total notional of the portfolio.
- The holder of a tranche is responsible for covering all losses in excess of K_L percent of the notional, up to K_U percent.
- In return, the premiums he receives are adjusted according to the remaining notional he is responsible for.
- In the case of popularly traded tranches on the North American Investment Grade Credit Default Swap Index (CDX.NA.IG), the tranches are named *equity, mezzanine, senior, senior, super-senior* with attachment and detachment points of $0-3, 3-7, 7-10, 10-15, 15-30$ respectively.

Basic CDO Pricing

- Given an underlying portfolio, and fixed attachment and detachment points for all tranches, the pricing problem is the determination of the spreads (periodic payment percentages) s_l for all tranches.
- Assume the default-free interest rate is independent of the credit risk of securities in the portfolio.

Basic CDO Pricing

- Denote total notional of the portfolio M, periodic payment dates $0 = t_0, t_1, \ldots, t_K$, period $t_{k+1} - t_k$ by γ.
- Let C_{t_k} be the total percentage of loss in the portfolio by time t_k, the attachment (detachment) point for tranche l by K_{L_l} (K_{U_l}), and the discount factor from t_0 to t_k by $\beta(t_0, t_k)$

Then the percentage loss $C_{l,t}$ suffered by the holders of tranche l up to time t is given by:

$$C_{l,t} := \min\{C_t, K_{U_l}\} - \min\{C_t, K_{L_l}\}$$

The value at time t_0 of payments received by the holder of tranche l is then:

$$\sum_{k=1}^{K} \beta(t_0, t_k) s_l \gamma M \mathbb{E}[K_{U_l} - K_{L_l} - C_{l,t_k}]$$

Basic CDO Pricing

The value at time t_0 of payments made by the holder of tranche l is given by:

$$\sum_{k=1}^{K} \beta(t_0, t_k) M \mathbb{E}[C_{l,t_k} - C_{l,t_{k-1}}]$$

Now set the premium s_l such that the value of payments received is equal to the value of payments made:

$$s_l = \frac{\sum_{k=1}^{K} \beta(t_0, t_k) \left(\mathbb{E}[C_{l,t_k}] - \mathbb{E}[C_{l,t_{k-1}}] \right)}{\sum_{k=1}^{K} \beta(t_0, t_k) \gamma \left(K_{U_l} - K_{L_l} - \mathbb{E}[C_{l,t_k}] \right)}$$

We only need the distribution for the percentage loss C_{t_k} for $k = 1, \ldots, K$ to price.

Basic CDO Pricing

Denoting the k-step transition matrix for the Markov chain with P^k and the number of losses at the k-th step by L^k:

$$P\left(C_{t_k} = \frac{m}{N}\right) = P\left(L^k = m\right) = \sum_{n=0}^{N} P^k_{(0,N)\to(m,n)}$$

Thus the spreads in our model are given by

Proposition

Given the inputs of the yield curve β, attachment (K_{L_l}) and detachment (K_{U_l}) points, and the implied Markov transition matrix P, the spread s_l of tranche l is given by

$$\frac{\sum_{k=1}^{K} \beta(t_0, t_k) \left(\sum_{m=0}^{N} \phi(m, N, l) \sum_{n=0}^{N} P^k_{(0,N)\to(m,n)} - P^{k-1}_{(0,N)\to(m,n)}\right)}{\sum_{k=1}^{K} \beta(t_0, t_k)\gamma \left(K_{U_l} - K_{L_l} - \sum_{m=0}^{N} \phi(m, N, l) \sum_{n=0}^{N} P^k_{(0,N)\to(m,n)}\right)}$$

where $\phi(m, N, l) = \left(\min\{\frac{m}{N}, K_{U_l}\} - \min\{\frac{m}{N}, K_{L_l}\}\right)$.

Comparison with normal copula

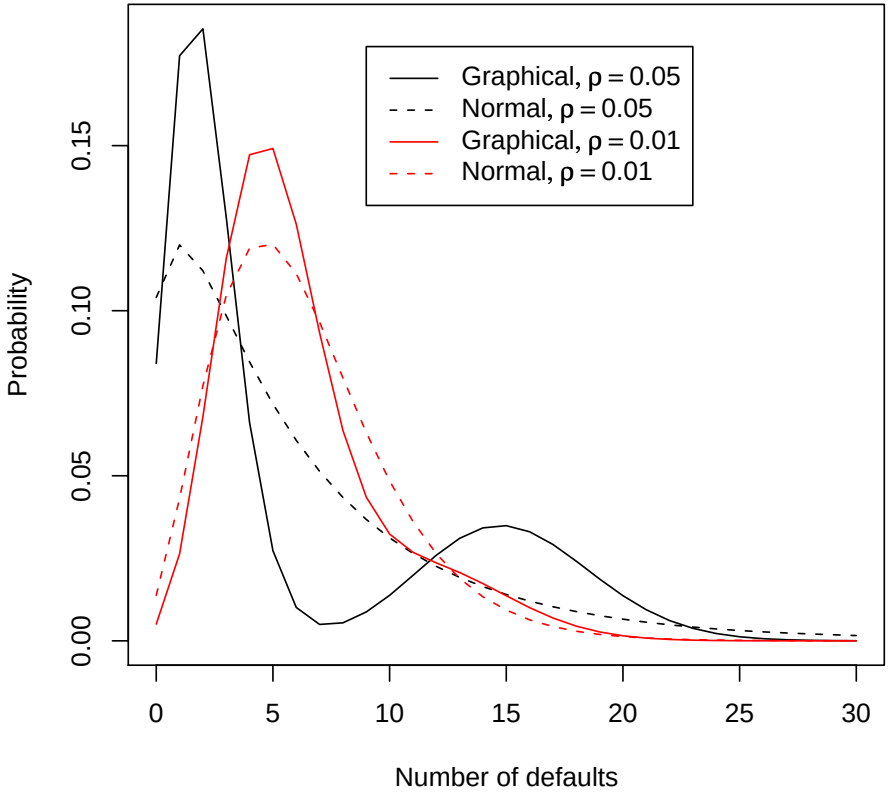


Figure: Comparison of one-factor normal copula and single sector graphical models

Comparison with normal copula

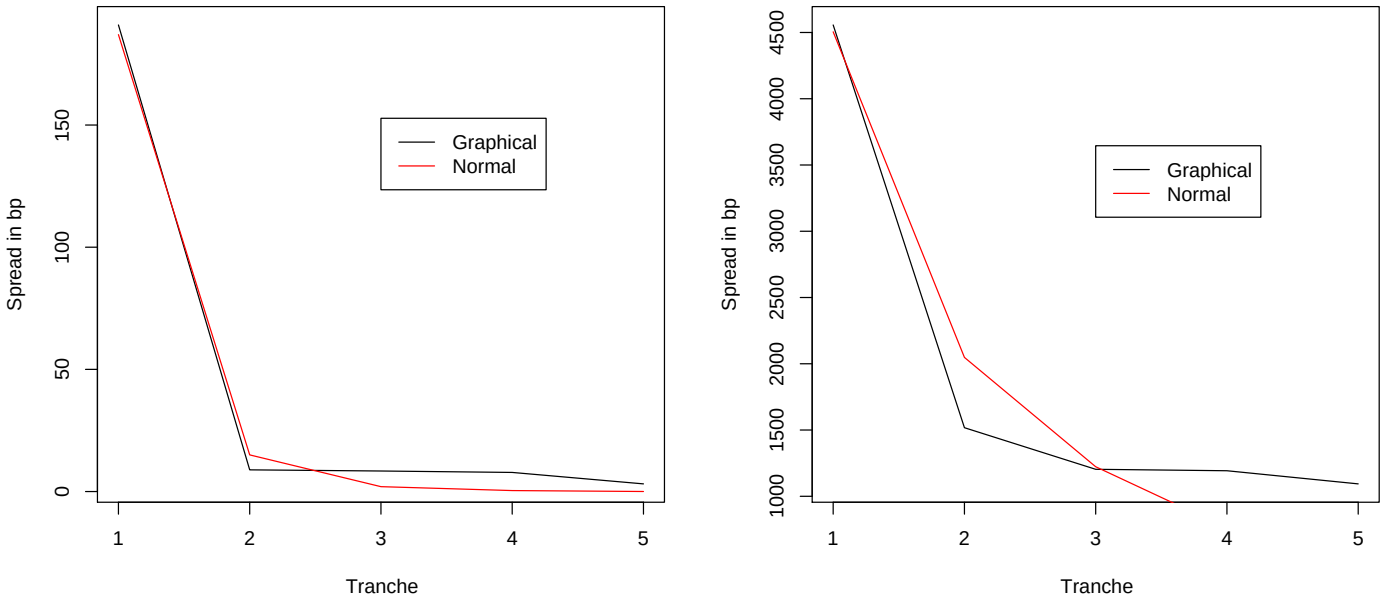


Figure: Tranche spreads for graphical and normal copula models

End

End

Feel free to contact me:

jason@math.stanford.edu