Cross-market High-dimensional Nonstationary Coupling Learning

Longbing Cao

The Data Science Lab
University of Technology Sydney, Australia

www.datasciences.org
Nonstationarity

• Stationarity
  
  same distribution

\[(Z_t, \ldots, Z_{t+m}) (Z_{t+k}, \ldots, Z_{t+m+k})\]

• Weak/non-stationarity

key difference

\[\Delta(q) = \sup_{h \in H} \left| \mathcal{L}(h, Z_1^T) - \sum_{t=1}^{T} q_t \mathcal{L}(h, Z_1^{t-1}) \right|\]

dependence between events decaying with \(k\).

\[\beta(k) = \sup_{n} \mathbb{E}_{B \in \sigma_{n+k}} \left[ \sup_{A \in \sigma_{n+k}} \left| P[A \mid B] - P[A] \right| \right] \to 0\]

NIPS16 tutorial: Theory and algorithms for forecasting non-stationary time series
Nonstationary couplings

• The change of one nonstationary variable is coupled with the change of another one

• Couplings: Association, correlation, dependence, causality, latent relations, etc.
Nonstationary heterogeneous couplings

Z Wu, etc. (2007) On the trend, detrending, and variability of nonlinear and nonstationary time series

D. Rahardja (2005) X-Charts versus X / MR Chart Combinations: IID Cases and Non-IID Cases
Main challenges

Financial variables are heterogeneous, coupled w.r.t. various structures

Multi-source, heterogeneous couplings and structures

Asymmetric, non-normal and nonlinear

Cross market dependence is often embedded with nonlinear, non-normal and asymmetric dependences

High dimensional, redundant, noisy

Dynamic and transitional

Generally more than 20 financial variables leading to over ten thousand features for a group of time series in a time window

Financial variables and their influence across markets are dynamic/volatile, transitional
Example 1: Copula-based Dependence Structure Modeling

J Xu, et al. Copula-Based High Dimensional Cross-Market Dependence Modeling
W Wei et al. Modeling Asymmetry and Tail Dependence among Multiple Variables by Using Partial Regular Vine
W Wei et al. Model the Complex Dependence Structures of Financial Variables by Using Canonical Vine
An example

Positive and negative correlations

- Stock markets and exchange rate markets are dependent
- Dependences may be diverse
- Cross-market couplings have to be considered in multivariate modeling
An example

Asymmetric and tail dependence

Gumbel Copula ($\theta = 3.35$) between S&P500 and STOXX50E

Frank Copula ($\theta = 28.5$) between FTSE100 and GBP

Daily returns of S&P 500 and EUR/USD (01/01/2008 – 31/12/2010)

W Wei et al. Modeling Asymmetry and Tail Dependence among Multiple Variables by Using Partial Regular Vine
Challenges

- Multiple heterogeneous financial variables
- Stylized fact: fat tail and asymmetric correlations
- Dependence strength and structure
Copula-based dependence modeling

- Modeling joint distribution between a group of random variables

\[
F_1(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))
\]

\[
C(u_1, u_2, \ldots, u_n) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n))
\]

\[
f(x_1, x_2, \ldots, x_n) = c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)) \prod_{i=1}^{n} f_i(x_i)
\]
Partial vine for structural dependence

- Weighted partial vine copula (WPVC)

- Partial D vine tree structure for asymmetric dependence in tail risk
- Bivariate copula with different types of tail dependencies
- Truncation with conditional independence (vs. correlations) for high-dimensional
Partial correlation

• Partial D vine structure

\[ \rho_{1,2;3,\ldots,n} = \frac{\rho_{1,2;3,\ldots,n-1} - \rho_{1,n;3,\ldots,n-1} \cdot \rho_{2,n;3,\ldots,n-1}}{\sqrt{1 - \rho_{1,n;3,\ldots,n-1}^2} \cdot \sqrt{1 - \rho_{2,n;3,\ldots,n-1}^2}} \]

• Lower and upper tail dependence coefficients

\[
\lambda_L = \lim_{u \to 0} \Pr\{U_1 \leq u, \ldots, U_n \leq u \mid U_n \leq u\} \\
= \lim_{u \to 0} \frac{C(u,\ldots,u)}{u} \\
\lambda_U = \lim_{u \to 0} \Pr\{U_1 > 1-u, \ldots, U_n > 1-u \mid U_n > 1-u\} \\
= \lim_{u \to 0} \frac{C(1-u,\ldots,1-u)}{u}
\]

• 0
• \( \in (0, 1) \)
Partial D vine tree construction

Weighted partial D vine tree construction

• 6 variables, 20 partial correlations

• The smallest PC as the edge of T5: \{A,F;B,C,D,E\}, conditioned set \{A,F\} and conditioning set \{B,C,D,E\}

• Nodes in T5:
  - Constraint sets \{A,B,C,D,E\} and \{F,B,C,D,E\}
  - Find the smallest PC for each constraint set and treat them as nodes, e.g. \{A,E;B,C,D\} and \{B,F;C,D,E\}

• T4, T3

• Best D vine:

\[
D = \prod_{i,j}(1 - W_i \rho_{i,j}^2 \delta(i,j)) \\
\text{argmax}(-\ln(D))
\]

\[
W(h) = \begin{cases} 
0.5 \times \frac{e^{-m_0(k-h)}}{\sum_{i=1}^{k-1} e^{-m_0(k-i)}}, & h \in [1, k]; \\
0.5 \times \frac{e^{-m_0(h-k)}}{\sum_{i=k+1}^{N-1} e^{-m_0(i-k)}}, & h \in (k, N-1]. 
\end{cases}
\]
Parameter estimation

• Maximum Log-Likelihood estimation to estimate the parameters of copula constructed w.r.t. vine structures

\[
L(\xi; x) = \sum_{j=1}^{n} \left\{ \sum_{i=1}^{p} \ln f_i(x_{i,j}; \phi_i) + \ln c(F_1(x_1, n), F_1(x_2, n), \ldots, F_p(x_p, n); \theta) \right\}
\]

\[
\xi = (\phi_1, \ldots, \phi_p, \theta)
\]

Estimate parameters for marginal distributions:

\[
L_m(\phi : x) = \sum_{i=1}^{p} \sum_{j=1}^{n} \ln f_i(x_{i,j}; \phi_i)
\]

\[\hat{\phi} = \arg\max_{\phi} L_m(\phi : x)\]

Estimate parameter in copula:

\[
L_c(\theta; u, \phi) = \sum_{i=1}^{p} \ln(c(F_1(x_1, n), \ldots, F_p(x_p, n); \theta))
\]

\[\hat{\theta} = \arg\max_{\theta} L_c(\theta; u, \phi)\]
Case study

- Backtesting VaR – Log-likelihood ratio
- VaR forecasting of portfolio return

<table>
<thead>
<tr>
<th>1 − α</th>
<th>POF</th>
<th>( LR_{IC}^2 )</th>
<th>( LR_{IC}^2 )</th>
<th>( LR_{CC}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>99%</td>
<td>1.08%</td>
<td>0.0324</td>
<td>(0.857)</td>
<td>(0.514)</td>
</tr>
<tr>
<td>95%</td>
<td>26</td>
<td>0.382 (0.536)</td>
<td>0.188 (0.868)</td>
<td>0.570</td>
</tr>
<tr>
<td>90%</td>
<td>52</td>
<td>1.090 (0.294)</td>
<td>1.582 (0.254)</td>
<td>2.683</td>
</tr>
<tr>
<td>99%</td>
<td>52</td>
<td>2.186 (0.139)</td>
<td>2.221 (0.136)</td>
<td>4.408</td>
</tr>
<tr>
<td>90%</td>
<td>57</td>
<td>1.303 (0.547)</td>
<td>3.533 (0.176)</td>
<td>4.896</td>
</tr>
<tr>
<td>99%</td>
<td>10</td>
<td>3.376 (0.066)</td>
<td>1.843 (0.175)</td>
<td>5.318</td>
</tr>
<tr>
<td>95%</td>
<td>57</td>
<td>2.17% (0.451)</td>
<td>1.133 (0.584)</td>
<td>2.578</td>
</tr>
<tr>
<td>90%</td>
<td>57</td>
<td>2.19% (0.584)</td>
<td>3.533 (0.618)</td>
<td>4.896</td>
</tr>
<tr>
<td>99%</td>
<td>11</td>
<td>4.770 (0.029)</td>
<td>1.439 (0.230)</td>
<td>6.209</td>
</tr>
<tr>
<td>90%</td>
<td>26</td>
<td>1.99% (0.547)</td>
<td>4.896 (0.033)</td>
<td>0.086</td>
</tr>
<tr>
<td>99%</td>
<td>29</td>
<td>0.662 (0.628)</td>
<td>0.042 (0.838)</td>
<td>0.704</td>
</tr>
<tr>
<td>90%</td>
<td>59</td>
<td>2.17% (0.320)</td>
<td>1.469 (0.035)</td>
<td>6.251</td>
</tr>
<tr>
<td>99%</td>
<td>59</td>
<td>2.39% (0.000)</td>
<td>5.429 (0.002)</td>
<td>471.533</td>
</tr>
<tr>
<td>90%</td>
<td>59</td>
<td>189.570 (0.000)</td>
<td>15.333 (0.004)</td>
<td>195.903</td>
</tr>
</tbody>
</table>

- 25 indicators: 8 exch rates, 13 indices, 3 commodity, 1 commodity index
- In-sample: 5 years; out-of-sample: 10 years
- ARMA(1; 1)-GARCH(1; 1): stocks
- AR(1)-GARCH(1; 1): exchange rate
Example 2: Deep modeling of financial couplings

W Cao et al. Deep Modeling Complex Couplings within Financial Markets
Market couplings

Figure 1: A demonstration of complex couplings between financial markets

\[ \theta_i = \{ \otimes_{j=1}^I (m_{ij}) \} \]

\[ \eta = \{ \oplus_{i=1}^I (\theta_i) \} \]

\[ \theta_{i,t} = \{ m_{i,j,[t-n,t-1]} \}_{j=1}^I \]

\[ \eta_{i,t} = \{ \theta_{i,[t-n,t-1]} \}_{i=1}^I \]
Two-layer market coupling learning

Figure 2: Modeling framework of CTDBN. Here, the demonstration shows two heterogeneous financial markets, stock and currency. The first-layer are CGRBMs to model the intra-market couplings while CCRBM are built on the first layer to model inter-market couplings.
Intra/inter-market couplings

Conditional Gaussian Restricted Boltzmann Machines: CGRBM intra-market couplings

\[ P(v, h | u) = \frac{\exp(-E(v, h, u))}{Z} \]

\[ E(v, h, u) = -\frac{v^T W h}{\sigma} - u^T \mathbf{A} v - u^T B h + \frac{(v - \mathbf{a})^T (v - \mathbf{a})}{2\sigma^2} - b^T h \]

\[ P(h_f = 1 | v, u) = s(b_f + u^T B_{i,f} + v^T W_{i,f}/\sigma) \]

\[ P(v_d | v, u) = \mathcal{N}(a_d + u^T A_{i,d} + \sigma W_{d,h} h, \sigma^2) \]

Coupled Conditional Restricted Boltzmann Machines: CCRBM inter-market couplings

\[ E(\{\theta_{i,t}\}, \eta_t; \{\theta_{i,<t}\}) = -b^T \eta_t - \sum_{i=1}^{I} a_{i,T}^T \theta_{i,t} - \sum_{i=1}^{I} \theta_{i,t}^T W_i \eta_t - \sum_{i=1}^{I} \theta_{i,<t}^T A_i \theta_{i,t} - \sum_{i=1}^{I} \theta_{i,<t}^T B_i \eta_t \]

\[ \theta_{i,<t} = [\theta_{i,t-1}, \theta_{i,t-2}, \ldots, \theta_{i,t-n}] \]

\[ P(\theta_{i,f,t} = 1 | \eta_h_t; \{\theta_{i,<t}\}) = s(a_{i,f} + \theta_{i,<t}^T (A_i)_{i,f} + (W_i)_{i,f} \eta_t) \]

\[ P(\eta_h_t = 1 | \{\theta_{i,t}\}; \{\theta_{i,<t}\}) = s(b_h + \sum_{i=1}^{I} \theta_{i,t}^T (W_i)_{i,h} + \sum_{i=1}^{I} \theta_{i,<t}^T (B_i)_{i,h}) \]

Figure 3: (a) A CGRBM to model intra-market coupling at time \( t \); (b) A CCRBM to model inter-market coupling at time \( t \).
Return prediction

Table 2: Performance of comparative methods in US, China and India markets

<table>
<thead>
<tr>
<th>Model</th>
<th>Stock Accuracy</th>
<th>Currency Accuracy</th>
<th>Stock ARR</th>
<th>Currency ARR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>US</td>
<td>China</td>
<td>India</td>
<td>US</td>
</tr>
<tr>
<td>ARIMA</td>
<td>0.5357</td>
<td>0.5071</td>
<td>0.5029</td>
<td>0.5471</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.5643</td>
<td>0.55</td>
<td>0.5196</td>
<td>0.6</td>
</tr>
<tr>
<td>ANN</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5752</td>
<td>0.6235</td>
</tr>
<tr>
<td>CHMM</td>
<td>0.6533</td>
<td>0.6214</td>
<td>0.5852</td>
<td>0.6471</td>
</tr>
<tr>
<td>CGRBM</td>
<td>0.6357</td>
<td>0.6235</td>
<td>0.5898</td>
<td>0.6565</td>
</tr>
<tr>
<td>CTDBN</td>
<td><strong>0.6729</strong></td>
<td><strong>0.6324</strong></td>
<td><strong>0.6258</strong></td>
<td><strong>0.6734</strong></td>
</tr>
</tbody>
</table>
Example 3:
Coupled Hidden Markov Model-based Dependence Modeling

W. Cao, et al. Financial Crisis Forecasting via Coupled Market State Analysis
L. Cao, et al. Coupled Behavior Analysis with Applications
Challenges

• Couplings within/between financial markets
• Couplings within/between financial variables
• Heterogeneity between markets, between financial variables
Modeling within/between-market couplings

- Temporal/periodical transition
- State transition within a market
- Coupling between markets
- CHMM for both transitions/influence within and between sequences
Modeling within/between-market couplings

- CHMM-LR:
  1. couplings between equity market and commodity market \( C(E,C) \);
  2. couplings between equity market and interest market \( C(E,I) \);
  3. couplings between commodity market and interest market \( C(C,I) \).

- Select financial variables
- Modeling their within/between sequence couplings
- Forecasting
Case study: financial crisis

![chart showing financial indicators](image)

Fig. 7: Technical Performance of Various Approaches

![chart showing market behavior during 2008 Financial Crisis](image)

Fig. 8: Market Counterpart Behavior during 2008 Financial Crisis ($w = 2$)

<table>
<thead>
<tr>
<th>TABLE II: Selected Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise Coupling</td>
</tr>
<tr>
<td>$C(E, C)$</td>
</tr>
<tr>
<td>$C(C, I)$</td>
</tr>
<tr>
<td>$C(E, I)$</td>
</tr>
</tbody>
</table>
More examples:
Deep Multivariate Coupling Learning
Deep financial representation

\[ v_v = W_v([h_n; x_s]) \]

\[ h^t_w = LSTM(h^{t-1}_w, x^t_w) \]

\[ N^{t+1} = \sigma(\sum_r D_r^{-\frac{1}{2}} A_r D_r^{-\frac{1}{2}} H^t_r W^t + W_h H^t) \]

---News Headline---
Japan Railway says change of president and chairman

Figure 1: A brief description of our proposed LSTM-RGCN model. Each node in the graph represents one stock. A node can be attached with none or several overnight news text. The dashed lines indicate the two relations that connects stocks (A). The news is first encoded with the node feature encoder (B). Then the node embedding is fed into our proposed LSTM-RGCN model to make use of the correlation graph structure (C). Note that LSTM-RGCN can have multiple layers. Finally, the node vectors are used to predict the overnight stock price movement (D).

W Li et al. Modeling the Stock Relation with Graph Network for Overnight Stock Movement Prediction, IJCAI20
Neural-dependence coupled networks

Copula variational autoencoder for mixed data

\[ F(\theta, \phi; x^{(n)}) = \mathbb{E}_{q_{\phi}(z^{(n)}|x^{(n)})} \left[ \log p_{\theta}(x^{(n)}|z^{(n)}) \right] - \text{KL} \left[ q_{\phi}(z^{(n)}|x^{(n)}) \parallel p_{\theta}(z^{(n)}) \right], \]

\[ p_{\theta}(x|z) = p_{\theta}(x^c, x^s|z) = c_{\Psi}(\cdot) \prod_{i=1}^{d_c} \mathcal{N}(\mu_i, \sigma_i^2) \prod_{i=1}^{d_s} \prod_{j=1}^{J} \beta_i^{(x^s_j=j)}, \]

S. Suh and S. Choi. Gaussian Copula Variational Autoencoders for Mixed Data
Concluding remarks
Coupling complexities in finance

• Macro/micro-variables across markets
• Structured/unstructured variables
• Within/between variable couplings
• Couplings: dependence, correlation, hidden relation, etc.
• Nonstationary, heterogeneous, multiscale, stylistic, ...
Heterogenous and hierarchical couplings

L Cao. Coupling Learning of Complex Interactions
www.datasciences.org

C Zhu et al. Unsupervised Heterogeneous Coupling Learning for Categorical Representation
Some recent work on AI in FinTech

- IJCAI2020 Special track on AI in FinTech
  [https://www.ijcai.org/Proceedings/2020/](https://www.ijcai.org/Proceedings/2020/)

- IEEE Intelligent Systems special issues on AI and FinTech

- DSAA2020 journal track on Data Science and AI in FinTech
  [dsaa2020.dsaa.co](dsaa2020.dsaa.co)

- L Cao. AI in FinTech: A Research Agenda

- L Cao. AI in Finance: A Review