Gaussian Process Conditional Copulas

José Miguel Hernández–Lobato[1],

Februrary 11, 2014

joint work with

David Lopez-Paz[1,2], James Robert Lloyd[1], Daniel Hernández–Lobato[3] and Zoubin Ghahramani[1]

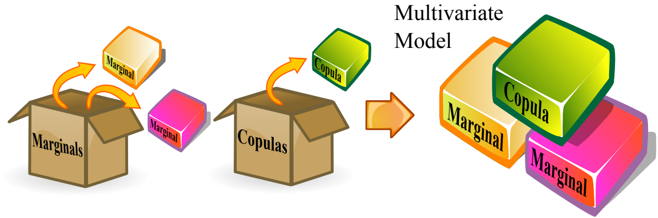
[1] Cambridge University.
[2] Max Planck Institute for Intelligent Systems
[3] Universidad Autónoma de Madrid.

Separating Marginals and Dependencies Using Copulas

Sklar's Theorem: any joint distribution can be written in terms of its copula and its univariate marginals [Sklar, 1959].

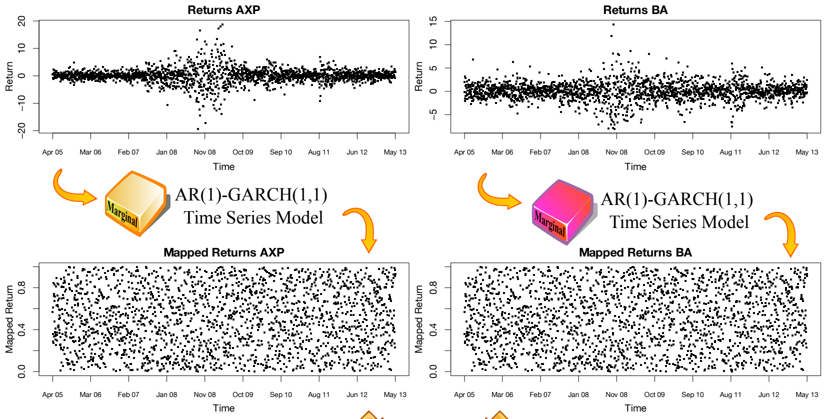
$$p(x_1, x_2) = \underbrace{c(u_1, u_2)}_{\text{copula}} \underbrace{p_1(x_1)\, p_2(x_2)}_{\text{marginals}},$$

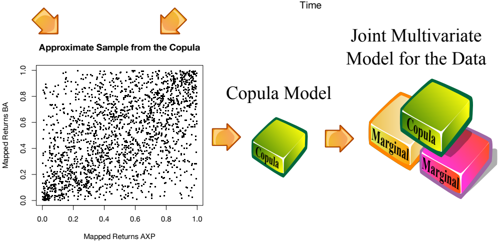
where $u_i := P_i(x_i)$ and P_i denotes the marginal cdf for x_i.



We can use different copulas and different univariate marginals as building blocks for new multivariate models.

Construction and Estimation of Multivariate Models



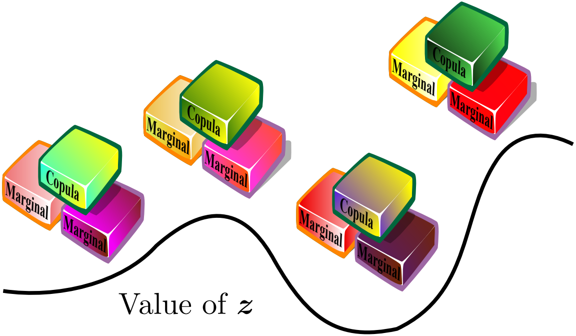


Copulas easily extend univariate models to the multivariate regime.

Copulas simplify the estimation of multivariate models [Joe, 2005].

Conditional Copulas

The distribution of x_1 and x_2 may depend on a covariate vector $\mathbf{z}$.



The copula framework can be extended to conditional distributions [Patton, 2006]. In this case,

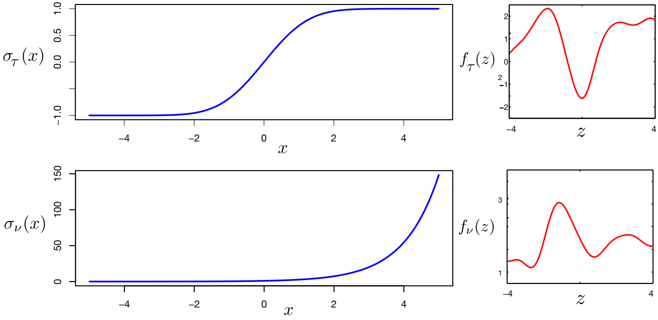
$$p(x_1, x_2|\mathbf{z}) = \underbrace{c(u_1, u_2|\mathbf{z})}_{\substack{\text{conditional}\\ \text{copula}}} \underbrace{p_1(x_1|\mathbf{z})\, p_2(x_2|\mathbf{z})}_{\substack{\text{conditional}\\ \text{marginals}}},$$

where $u_i := P_i(x_i|\mathbf{z})$ and P_i denotes the marginal cdf for x_i given $\mathbf{z}$.

Semiparametric Conditional Copulas

Main Assumption: $c(u_1, u_2|\mathbf{z})$ is given by a **parametric** copula model $c_{\text{par}}[u_1, u_2|\theta_1(\mathbf{z}), \ldots, \theta_k(\mathbf{z})]$ specified by k scalar parameters $\theta_1, \ldots, \theta_k$ that are functions of $\mathbf{z}$. We select $\theta_i(\mathbf{z}) = \sigma_i[f_i(\mathbf{z})]$, where f_i is a real function and σ_i is a link function that maps $\mathbb{R}$ to a set Θ_i of valid configurations for θ_i.

Example: conditional Student's t copula. $c_{\text{student}}[u_1, u_2|\tau(\mathbf{z}), \nu(\mathbf{z})]$, $\tau(\mathbf{z}) = \sigma_\tau[f_\tau(\mathbf{z})]$, $\nu(\mathbf{z}) = \sigma_\nu[f_\nu(\mathbf{z})]$, $\sigma_\tau(x) = 2\Phi(x) - 1$, $\sigma_\nu(x) = \exp(x)$.



Our objective is to estimate $f_1, \ldots, f_k$ from data.

Expectation Propagation

EP approximates the unnormalized joint $p(\mathbf{f}_1, \ldots, \mathbf{f}_k, \mathcal{D}_{u_1,u_2} | \mathcal{D}_\mathbf{z})$ with an unnormalized Gaussian $\mathcal{Q}(\mathbf{f}_1, \ldots, \mathbf{f}_k) = k \prod_{j=1}^{k} \mathcal{N}(\mathbf{f}_j | \mathbf{m}_j, \mathbf{V}_j)$.

Copy Likelihood (Exact Factor) $g_i(f_{1i}, \ldots, f_{ki})$ — Gaussian Prior

$$p(\mathbf{f}_1, \ldots, \mathbf{f}_k, \mathcal{D}_{U,V} | \mathcal{D}_\mathbf{Z}) = \left[\prod_{i=1}^{n} \overbrace{c_{\text{par}} \left[u_{1i}, u_{2i} | \sigma_1(f_{1i}), \ldots, \sigma_k(f_{ki}) \right]}^{g_i(f_{1i}, \ldots, f_{ki})} \right] \left[\prod_{j=1}^{k} \mathcal{N}(\mathbf{f}_j | \mathbf{m}_\mathbf{Z}^{(j)}, \mathbf{V}_{\mathbf{Z},\mathbf{Z}}^{(j)}) \right]$$

EP Projection ↓   Exact ↓

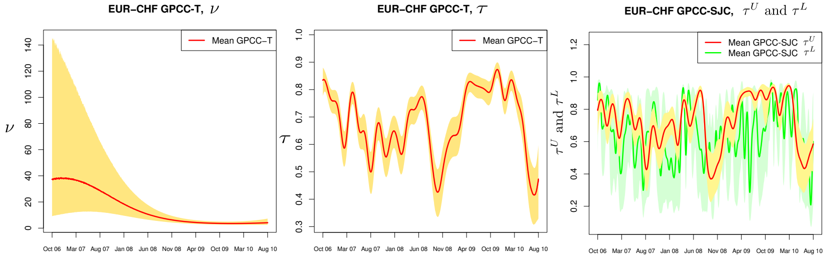
$$\mathcal{Q}(\mathbf{f}_1, \ldots, \mathbf{f}_k) = \underbrace{\left[\prod_{i=1}^{n} \tilde{k}_i \left[\prod_{j=1}^{k} \mathcal{N}(f_{ji} |, \tilde{m}_{ji}, \tilde{v}_{ji}) \right] \right]}_{\tilde{g}_i(f_{1i}, \ldots, f_{ki})} \left[\prod_{j=1}^{k} \mathcal{N}(\mathbf{f}_j | \mathbf{m}_\mathbf{Z}^{(j)}, \mathbf{V}_{\mathbf{Z},\mathbf{Z}}^{(j)}) \right]$$

Approximate Gaussian Factor

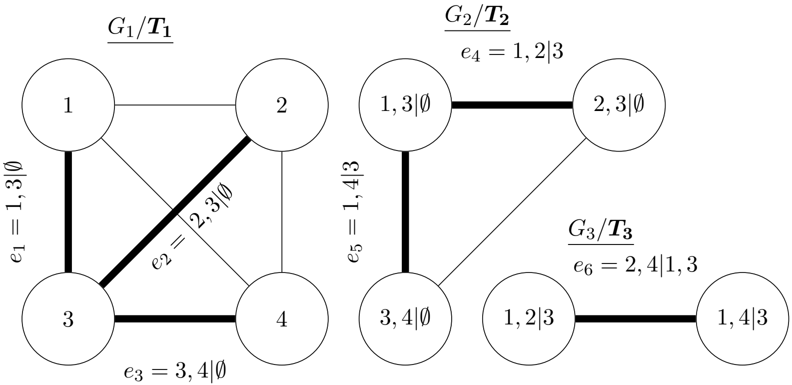
EP adjusts $\tilde{g}_i$ by minimizing $\text{KL}[g_i \mathcal{Q}^{\setminus i} || \tilde{g}_i \mathcal{Q}^{\setminus i}]$, where $\mathcal{Q}^{\setminus i}$ is the ratio of $\mathcal{Q}$ and $\tilde{g}_i$. EP with sparse GPs [Naish-Guzman et al. 2008].

Results on Foreign Exchange Time Series

Method	AUD	CAD	JPY	NOK	SEK	EUR	GBP	NZD
GPCC-G	0.1260	0.0562	**0.1221**	0.4106	0.4132	0.8842	0.2487	0.1045
GPCC-T	**0.1319**	**0.0589**	0.1201	**0.4161**	**0.4192**	**0.8995**	**0.2514**	**0.1079**
GPCC-SJC	0.1168	0.0469	0.1064	0.3941	0.3905	0.8287	0.2404	0.0921
HMM	0.1164	0.0478	0.1009	0.4069	0.3955	0.8700	0.2374	0.0926
TVC	0.1181	0.0524	0.1038	0.3930	0.3878	0.7855	0.2301	0.0974
DSJCC	0.0798	0.0259	0.0891	0.3994	0.3937	0.8335	0.2320	0.0560
CONST-G	0.0925	0.0398	0.0771	0.3413	0.3426	0.6803	0.2085	0.0745
CONST-T	0.1078	0.0463	0.0898	0.3765	0.3760	0.7732	0.2231	0.0875
CONST-SJC	0.1000	0.0425	0.0852	0.3536	0.3544	0.7113	0.2165	0.0796



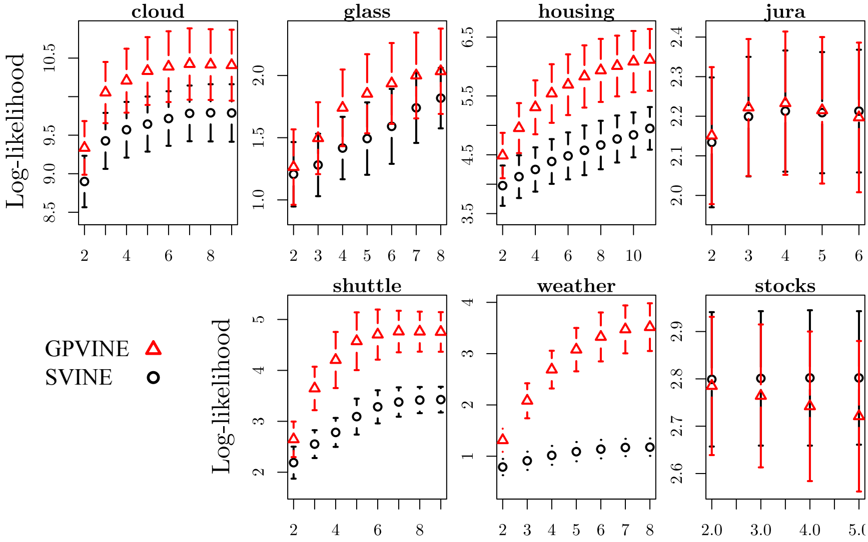
Extension to Higher Dimensions: Vine Factorizations



$$c_{1234} = \underbrace{\underbrace{c_{13|\emptyset}}_{e_1} \underbrace{c_{23|\emptyset}}_{e_2} \underbrace{c_{34|\emptyset}}_{e_3}}_{T_1} \underbrace{\underbrace{c_{12|3}}_{e_4} \underbrace{c_{14|3}}_{e_5}}_{T_2} \underbrace{\underbrace{c_{24|13}}_{e_6}}_{T_3}$$

[Lopez-Paz et al. 2013]

Results for Fully Conditional Vines



To Conclude...

Summary:

- We have proposed a new framework for conditional dependencies based on parametric copula models and GPs.
- Can be easily extended to higher dimensions using vines.

Open questions:

- Advantages/disadvantages of copulas over factor models?

Future directions:

- Can we go beyond the semiparametric approach?
- What about dependencies between latent variables?

Potential additional applications:

- Meteorology: wind and rain forecast.
- Ecosystem Informatics: associations among soil nutrients.
- Economics: optimal bundle pricing [Letham et al. 2014].

Useful References I

- J. M. Hernndez-Lobato, J. R. Lloyd and D. Hernndez-Lobato. Gaussian Process Conditional Copulas with Applications to Financial Time Series. In Advances in Neural Information Processing Systems 26 (NIPS), 2013.
- A. Sklar. Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statis. Univ. Paris, 8(1):229–231, 1959.
- H. Joe. Asymptotic efciency of the two-stage estimation method for copula-based models. J. Multivar. Anal., 94(2):401–419, 2005.
- A. J. Patton. Modelling asymmetric exchange rate dependence. International economic review, 47(2):527–556, 2006.
- E. Snelson and Z. Ghahramani. Sparse Gaussian processes using pseudo-inputs. In NIPS, 1257–1264, 2006.

Useful References II

- A. Naish-Guzman and S. B. Holden. The generalized FITC approximation. In NIPS, 1057–1064, 2008.
- E. Jondeau and M. Rockinger. The copula-garch model of conditional dependencies: An international stock market application. Journal of international money and nance, 25(5):827–853, 2006.
- J. M. Hernandez-Lobato, D. Hernández-Lobato, and A. Suárez. GARCH processes with non-parametric innovations for market risk estimation. In ICANN, 718–727, 2007.
- D. Lopez-Paz, J. M. Hernández-Lobato and Z. Ghahramani. Gaussian Process Vine Copulas for Multivariate Dependence, In ICML, 2013.
- B. Letham, W. Sun, A. Sheopuri. Latent Variable Copula Inference for Bundle Pricing from Retail Transaction Data. JMLR W&CP 32 (1) : 217–225, 2014.

Thank you for your attention!