Université catholique de Louvain – Institut de statistique, biostatistique et sciences actuarielles

Young Researchers' Day

September 24th, 2010

9^{00}	**Jonathan Jaeger**	Functional estimation in systems defined by differential equation using bayesian smoothing methods
9^{30}	**Ingrid Haff**	Pair-copula constructions of multiple dependence
10^{00}	**Louis Boulanger**	Modeling extremal dependence: Application to CDOs pricing

Coffee Break

11^{00}	**Diane Pierret**	Multivariate volatility modelling of electricity futures
11^{30}	**Boris Demeshev**	Guided local linear regression
12^{00}	**Majda Talamakrouni**	Guided censored regression

Lunch at the cafeteria

The Young Researchers' Day is held in room c.115 of the Institut de statistique, biostatistique et sciences actuarielles, Voie du Roman Pays 20, Louvain-la-Neuve.

Functional Estimation in Systems Defined by Differential Equation using Bayesian Smoothing Methods

JONATHAN JAEGER (Jonathan.Jaeger@uclouvain.be)

Differential equations are frequently used to specify models in chemical engineering, pharmacokinetics and other sciences. Current methods for parameter and functional estimations in such models use minimization techniques and numerical solver. These approaches are computationally intensive and often poorly suited for statistical inference.

Alternative estimation methods of the state function and of the dynamic model parameters were proposed by Ramsay et al. (2007). It accounts for measurements errors and elude numerical integration. That approach involves some basis function expansion of the state function and a penalty term expressed using the set of differential equations.

The methodology above may be viewed as a generalization of the penalized spline theory for which a Bayesian framework was proposed by Berry et al. (2002). We aim at providing a Bayesian framework for the more general approach described by Ramsay et al (2007). First, we present a brief introduction to dynamic models defined by systems of differential equations. We then propose a full Bayesian smoothing approach for the joint estimation of the differential equation parameters and of the state functions and extend it to a hierarchical context. We conclude the presentation by some practical examples.

References

Berry S.M., Carroll R.J. and Ruppert D. (2002): Bayesian smoothing and regression splines for measurement error problems. *Journal of the American Statistical Association 97, 160-169.*

Ramsay J.O., Hooker G., Campbell D. and Cao J. (2007): Parameter estimation for differential equations: a generalized smoothing approach. *Journal of the Royal Statistical Society, Series B 69, 741-796.*

Pair-copula constructions of multiple dependence

Ingrid Haff (Ingrid.Haff@nr.no)

Copulæ are now regularly used within fields such as finance, survival analysis and actuarial sciences. Although the list of parametric bivariate copulæ is long and varied, the choice is rather limited in higher dimensions. Therefore one has proposed a number of hierarchical, copula-based structures, among those the pair-copula construction (PCC). This hierarchical model, built exclusively from pair-copulæ, consists of various levels, incorporating an increasing number of variables in the conditioning sets. Due to its simple structure, simulation and inference are straightforward. Still, it can model a wide range of complex dependencies, as we show in two four-dimensional examples, involving financial returns and precipitation values.

Modeling Extremal Dependence: Application to CDOs pricing

Louis Boulanger (Louis.Boulanger@uclouvain.be)

Collateralized debt obligations are a special type of asset-backed securities which are characterized by a split into tranches of decreasing seniorities. Tranches correspond to different levels of risk born by the investor: senior, mezzanine, equity. They allow the originators of the underlying assets to pass credit risk (e.g. regarding mortgage loans) to another institution or individual investors via the creation of a Special Purpose Vehicle (SPV).

The evolution of the collateralized debt obligations market is closely linked to the financial crisis. First, the growing complexity and the spectacularly fast increase of issued volumes of CDOs had the following consequence: deterioration in credit quality. Particularly considering subprime mortgage credits. More defaults occurred than predicted by the models used to estimate default probabilities. Second, CDOs' pricing mainly depends on the expected losses born by each tranche so that the computation of default probabilities is the key issue for the valuation of CDO tranches. For example, lower joint default probabilities correspond to senior tranches being more attractive. On the contrary, higher joint default probabilities but with marginal probabilities remaining low will attract more equity investors as they are required to bear the first losses on the underlying portfolio.

We describe in details for the bivariate case, the model widely used in practice: the Li model, suggesting the use of a Gaussian copula to compute joint default probabilities from marginal default probabilities of two companies. Under this model, the parameter of the copula is defined as the linear correlation between the companies' values or returns. We come to the conclusion that the Gaussian copula is inappropriate because of its asymptotic independence property unable to impose a structure of dependence to the tails of a joint distribution. We prove indeed that marginal default probabilities remain small by calculating the latter from credit default swap (CDS) spreads.

Extreme-Value theory allows us to study this tail dependence. In order to compute the joint probability that BNP Paribas and Société Générale will default within five years, we build an empirical non-parametric extreme-value copula based on daily returns of those two companies. By expressing extreme-value copulas through equivalent dependence functions on one hand or through an exponent measure and a spectral representation on the other hand, we are indeed able to construct useful non-parametric estimators. We do so via the programming of these estimators on the R-project free software.

The joint default probabilities computed from the empirical extreme-value copula are higher than the ones coming out from the use of the Gaussian copula. However, as data about time-until-default variables are by definition not available, a model of best-fitting is impossible and a comparison of results can not clearly conduce to a preference in the choice of a specific copula.

Our main contribution in this area of research is that we point out that marginal probabilities being small, only the lower tail of the distribution is of interest: the Gaussian copula is thus inappropriate in this context. Another innovation is that we introduce a more complete study of the Extreme-Value Theory in the CDO context than what is currently available. We define concepts and tools appropriate to the study of tail dependence. Among them, we can list asymptotic independence, rate of convergence and the coefficient of tail dependence. They remain useful whatever the copula selected to compute joint default probabilities.

Multivariate Volatility Modelling of Electricity Futures

DIANE PIERRET (Diane.Pierret@uclouvain.be)

Following the rapid development of emerging electricity markets, trading of energy derivatives has intensified on both OTC and pool markets. The leading trading market in continental Europe is the European Energy Exchange (EEX) founded in 2002. The EEX Physical Electricity Index (Phelix) has become the reference for power prices in Europe. However, due to the intrinsic nature of electricity, futures represent a larger market than spot trading. Energy risk management uses futures to hedge against spot price risk but futures are also a substitute to spot trading for investors willing to take positions in power markets without the physical constraints linked to electricity. We are therefore interested in modelling the risks associated with a portfolio composed of different Phelix futures corresponding to different maturities. For this sake, different multivariate GARCH models are reviewed. Based on empirical evidence, the dynamic conditional correlation (DCC) class of models is selected and in particular, the modified DCC model of Aielli (2009). In a second step, different ways for introducing flexibility in the models for volatilities and conditional correlations are explored. The flexibilities concern two major axes: the possibility to allow for asset-specific dynamics and the decomposition into long term and short term components where the long term component is smoothly, slowly changing over time. Ultimately, the appropriate modelling of both the volatilities and the correlations in a multivariate context should help us in elaborating accurate joint density forecasts and risk measures for power futures portfolios.

Guided local linear regression

BORIS DEMESHEV (Boris.Demeshev@uclouvain.be)

We consider two-step estimation procedure of the unknown function $m(x)$. The first step is a parametric estimator, $\hat{m}_0(x)$. The second step is a non-parametric estimator $\hat{m}(x)$ which takes into account the result of the first step.

The goal is to build an estimator whith two properties:

- When the estimator $\hat{m}_0(x)$ is "bad" the asymptotic bias and variance are comparable with those of the classic local linear non-parametric estimator
- When the estimator $\hat{m}_0(x)$ is "good" the asymptotic bias and variance are less than those of the classic local linear non-parametric estimator

We consider three alternative estimator for the second step:

(i) $\hat{m}_{xm}(x)$ minimizes

$$\sum_i K\left(\frac{X_i - x}{h}\right)(Y_i - [\gamma_0 + \gamma_1(\hat{m}_0(X_i) - \hat{m}_0(x))])^2$$

Motivation: taylor expansion using another basis function. If this function is guessed right we win.

(ii) $\hat{m}_{mx}(x)$ minimizes

$$\sum_i K\left(\frac{\hat{m}_0(X_i) - \hat{m}_0(x)}{h}\right)(Y_i - [\gamma_0 + \gamma_1(X_i - x)])^2$$

Motivation: non-linear change of bandwidth which will change the weigth of observations. If this change is right we give higher weigth for relevant observations.

(iii) $\hat{m}_{mm}(x)$ minimizes

$$\sum_i K\left(\frac{\hat{m}_0(X_i) - \hat{m}_0(x)}{h}\right)(Y_i - [\gamma_0 + \gamma_1(\hat{m}_0(X_i) - \hat{m}_0(x))])^2$$

Motivation: the two previous motivations combined.

All these estimators have closed-form solution. We prove that in the case of "bad" $\hat{m}_0$ the estimators $\hat{m}_{mx}$ and $\hat{m}_{mm}$ may be inconsistent. We show the results of Monte-Carlo simulations.

What has been done before?

(i) Einbeck, Local fitting with power basis, 2004: study of $\hat{m}_{xm}$ estimator with non-parametric preestimator $\hat{m}_0$

(ii) Akritas, Reverse windows in nonparametric regression, 2005: study of Nadaraya-Watson analog of $\hat{m}_{mx}$ estimator with non-parametric preestimator $\hat{m}_0$

Questions to be answered:

(i) What are the asymptotic bias and variance of the three estimator when the parametric estimator is "good"?

(ii) What are the properties of these estimators in quantile regression setup?

(iii) Can these estimators deal with the "dimensionality curse"?

Guided Censored regression

MAJDA TALAMAKROUNI (Majda.Talamakrouni@uclouvain.be)

Parametrically guided nonparametric estimation is one of the most promising approaches that improves the bias of traditional nonparametric regression estimators without an increase in the variance. In the context of completely observed i.i.d. data, many techniques are available in the literature. These include Glad (1998), Fan and Ullah (1999), Gozalo and Linton (2000), Mays et al. (2001) and Martins-Filho et al. (2008).

However, a common problem in practice is the presence of censoring. We are studying the guided nonparametric estimator of the regression function when the dependent variable is subject to censoring. To deal with censoring we will simply transform the observed data in unbiased way, a technique that is largely used in the literature, see for example Delecroix et al. (2008). We will first study the case when the dependent and the censoring variable are independent and then we will move on to the more difficult and realistic case : the two variables are only conditionally independent given a covariate. We will study the properties of the new approach like the asymptotic normality, the efficiency and the robustness. The asymptotic results will also be illustrated with finite sample simulations.

References

Glad, I. K. (1998), Parametrically Guided Nonparametric Regression, *The Scandinavian Journal of Statistics. Theory and Applications, 25, 649-668.*

Fan, Y., and Ullah, A. (1999). Asymptotic Normality of a Combined Regression Estimator. *Journal of Multivariate Analysis, 71, 191-240.*

Gozalo, P., and Linton, O. (2000). Local Nonlinear Least Squares: Using Parametric Information in Nonparametric Regression. *Journal of Econometrics, 99, 63-106.*

Mays, J. E., Birch, J. B., and Starnes, B. A. (2001). Model Robust Regression: Combining Parametric, Nonparametric, and Semiparametric Methods. *Journal of Nonparametric Statistics, 13, 245-277.*

Martins-Filho, C., Mishra, S. and Ullah, A. (2008). A class of improved parametrically guided nonparametric regression estimators. *Econometric Reviews, 27, 542-573.*

Delecroix, M., Lopez, O. and Patilea, V. (2008). Nonlinear censored regression using synthetic data. *Scandinavian Journal of Statistics, 35, 248-265.*