Some further properties and applications of local Gaussian approximation

D. Tjøstheim, with G. Berentsen, B. Støve, V. Lacal, H. Otneim

UNIVERSITETET I BERGEN
Matematisk institutt

February 2016

References

- Tjøstheim and Hufthammer (2013): Local Gaussian correlation: A new measure of dependence, J. of Econometrics.
- Berentsen, Støve, Tjøstheim, Nordbø (2014): Recognizing and visualizing copulas, Insurance: Mathematics and Economics.
- Støve, Tjøstheim and Hufthammer (2014) Using local Gaussian correlation in a nonlinear re-examination of financial contagion, J. Empirical Finance.
- Støve, Tjøstheim (2014) Measuring asymmetries in financial returns: an empirical investigation using local Gaussian correlation, Oxford: Nonlinear Econometrics.
- Berentsen, Kleppe, Tjøstheim (2014) locGauss package. Journal of Statistical Software.
- Lacal and Tjøstheim (2015) Local Gaussian approximation and tests of serial independence (+ bivariate time series).
- Otneim and Tjøstheim (2015) Multivariate local Gaussian density estimation (+conditional density).

Global measures of dependence

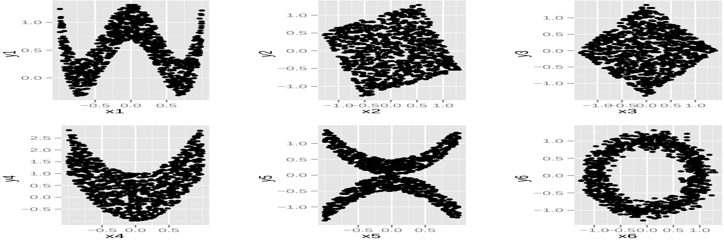
Pearson's correlation The most widely used (linear) measure of dependence for bivariate data is the **Pearson's correlation**

$$\rho = \mathrm{corr}(X_1, X_2) = \frac{E(X_1 - \mu_1)(X_2 - \mu_2)}{\sigma_1 \sigma_2} \quad (1)$$

- Completely characterize the dependence when the variables are Gaussian

Global measures of dependence

Pearson's correlation However, if the variables are not Gaussian the Pearson's correlation may be zero even if the variables are highly dependent:



Source: wikipedia.org "Correlation and dependence".

Nonparametric regression based local measure. Bjerve and Doksum (1993)?

Return data

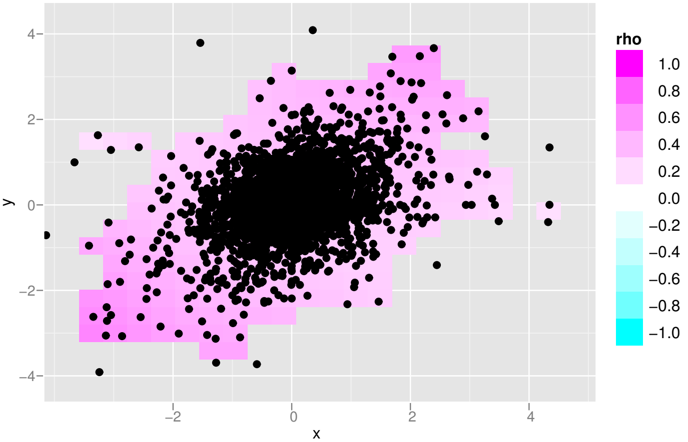


Figure: x = returns from FTSE100, y = returns from S&P500, in % (from May 1990 to December 1999). Global correlation: 0.36.

Return data

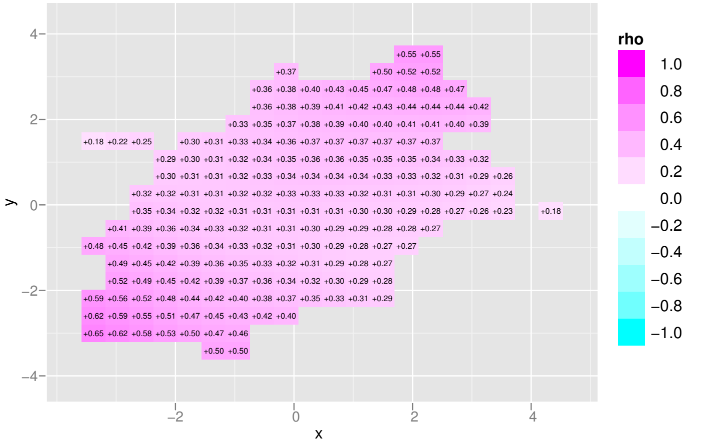


Figure: x = returns from FTSE100, y = returns from S&P500, in % (from May 1990 to December 1999). Global correlation: 0.36.

Conditional correlation

- $\{X_{1t}, X_{2t}\}$ log returns
- Conditional correlation (semi-correlation) restricted to the set A

$$\rho_A = \text{corr}(X_{1t}, X_{2t} | (X_{1t}, X_{2t}) \in A)$$

- Exceedance correlation

$$A = (-\infty, -c] \times [-\infty, -c]$$

Reference: Forbes and Rigobon (2002) Hong et al (2007), Joe (New copula book 2014)

Conditional correlation

- Bias for (X_{1t}, X_{2t}) Gaussian
- Example: $\rho = 0.40$

$$A = \{(X_1, X_2) | X_1 > q_{75}\}$$

  implies $\rho_A = 0.21$
- $\rho_A \to 0$ for increasing quantiles q
- Very misleading for local portfolio analysis

Alternative: Holland and Wang (1987), Jones (1996)

Outline of talk

- Study a local version of correlation
- Local dependence measure should depend on the values of X_1 and X_2. Local in value, not in time!
- Local Gaussian approximation
- Population value of local correlation
- Estimation of local autocorrelation and crosscorrelation
- Recognizing copulas
- Estimating densities and conditional densities in higher dimensions

Local Gaussian approximation

Idea: Given a density function $f(x_1, x_2)$, approximate f locally by a bivariate Gaussian distribution, i.e., at the point (x_1, x_2), or in a neighbourhood, fit a bivariate Gaussian density $\phi(\theta(x_1, x_2), v_1, v_2)$ given by

$$\frac{1}{2\pi\sigma_1(x_1,x_2)\sigma_2(x_1,x_2)\sqrt{1-\rho^2(x_1,x_2)}}\times$$

$$\exp\left\{-\frac{1}{1-\rho^2(x_1,x_2)}\left\{\frac{[v_1-\mu_1(x_1,x_2)]^2}{\sigma_1^2(x_1,x_2)}\right.\right.$$

$$\left.\left.-2\rho(x_1,x_2)\frac{[v_1-\mu_1(x_1,x_2)]}{\sigma_1(x_1,x_2)}\frac{[v_2-\mu_2(x_1,x_2)]}{\sigma_2(x_1,x_2)}+\frac{[v_2-\mu_2(x_1,x_2)]^2}{\sigma_2^2(x_1,x_2)}\right\}\right\}$$

Local correlation is given by $\rho(x_1, x_2)$. We also get a local mean and a local variance

- **But representation is highly non-unique**

Existence and uniqueness

- Local least squares: Minimize the penalty function

$$\int K_h(v - x)[\phi(v, \theta(x)) - f(v)]^2 dv$$

with $\theta(x) = [\mu_1(x); \mu_2(x), \sigma_1^2(x), \sigma_2^2(x), \rho(x)]$.

- Local Kullback-Leibler: Minimize

$$\int K_h(v - x)[\phi(v, \theta(x)) - \log \phi(v, \theta(x)) f(v)] dv$$

leading to

$$\int K_h(v - x) \frac{\partial}{\partial \theta_i} \log \phi(v, \theta(x))[f(v) - \phi(v, \theta(x))] dv = 0$$

and then let bandwidth $h \to 0$.

- Neighbourhood-free definition using smoothness conditions (Berentsen, Cao, Francisco-Fernandez, Tjøstheim 2016)

Existence and uniqueness

- If (X_1, X_2) is Gaussian, then local quantities equal global ones.
- Step functions of Gaussians

$$X = \sum_{i=1}^{k} (a_i + A_i Z) 1(Z \in R_i)$$

  where $Z \sim \mathcal{N}(0, I_2)$
- Explicitly calculated for exchangeable copula along diagonal

Properties

- f Gaussian implies $\rho(x_1, x_2) \equiv \rho$.
- f Gaussian copula with arbitrary marginals implies $\rho(x_1, x_2) \equiv \rho$.
- Range $-1 \leq \widehat{\rho}_h(x_1, x_2) \leq 1$, $-1 \leq \rho_h(x_1, x_2) \leq 1$, $-1 \leq \rho(x_1, x_2) \leq 1$
- $X_2 = g(X_1)$ deterministic implies $\rho(x_1, x_2) = 1$ if $g'(x_1) > 0$, $\rho(x_1, x_2) = -1$ if $g'(x_1) < 0$
- Independence $\Rightarrow \rho(x_1 x_2) \equiv 0$.

  If $\mu_1(x_1, x_2) = \mu_1(x_1)$, $\sigma_1(x_1, x_2) = \sigma_1(x_1)$ and similarly for μ_2 and σ_2, then $\rho(x_1, x_2) \equiv 0$ implies independence.

Time series case

- Stationary time series $\{X_t\}$
- Pairwise lag k: Take $X_1 = X_t$ and $X_2 = X_{t-k}$
- Local Gaussian autocorrelation: $\rho_k(x_1, x_2) = \rho_{X_t, X_{t-k}}(x_1, x_2)$ depends on k but not on t.
- Gaussian time series: $\rho_k(x_1, x_2) \equiv \rho_k$
- Multivariate time series. Local crosscorrelation.

Canonical local autocorrelation

- Standard normal marginals: $Z_t = \Phi^{-1}(F(X_t))$
- Local parameter vector: $\theta = [0,0,1,1,\rho_{c,k})$
- Pseudo standard normals: $Z_t^{(n)} = \Phi^{-1}(F_n(X_t))$.
- Local canonical correlation $\rho_{c,k}$ can be estimated by local likelihood from the pseudo standard normal variables.
- Does not depend on marginals

Estimation by local likelihood

General problem: Estimate a density function $f(x)$ by a known parametric family $g(x, \theta)$, where $\theta = \theta(x)$ is estimated locally, so that

$$\widehat{f}(x) = g(x, \widehat{\theta}(x))$$

Emphasis has been on estimation of $f(x)$, not on $\theta(x)$.

References: Hjort and Jones (1996), Loader(1996), Gourieroux and Jasiak (2012).

Local likelihood

$$L_n(\theta) = \frac{1}{n}\sum_i K_h(X_i - x)\log\phi(X_i,\theta) - \int K_h(v-x)\phi(v,\theta)dv$$

$$\begin{aligned}
0 = \dot{L}_n(\theta) &= \frac{1}{n}\sum_i K_h(X_i - x)\frac{\dot{\phi}(X_i,\theta)}{\phi(X_i,\theta)} \\
&- \int K_h(v-x)\frac{\dot{\phi}(v,\theta)}{\phi(v,\theta)}\phi(v,\theta)dv \\
&\to \int K_h(v-x)\frac{\dot{\phi}(v,\theta)}{\phi(v,\theta)}[f(v)-\phi(v,\theta)]dv \\
&\sim \frac{\dot{\phi}(x,\theta)}{\phi(x,\theta)}[f(x)-\phi(x,\theta)+O(h^2)] = 0
\end{aligned}$$

which implies that $\phi(x,\theta(x)) - f(x) = O(h^2)$

Asymptotic distribution

Under regularity conditions:

$$(nh_1h_2)^{1/2}(\hat{\theta}_{n,h}(x) - \theta_h(x)) \xrightarrow{d} \mathcal{N}(0, J_h^{-1}M_h(J_h)^{-1})$$

where for $u(x,\theta) = \nabla \log \phi(x,\theta)$

$$J_h = \int K_h(v-x)u(v,\theta(x)u^T(v,\theta(x))\phi(v,\theta(x))dv$$
$$-\int K_h(v-x)\nabla u(v,\theta(x)[f(v) - \phi(v,\theta(x))]dv$$

$$M_h = h_1h_2\int k_h^2(v-x)u(v,\theta(x))u^T(v,\theta(x))f(v)dv$$
$$-h_1h_2\int K_h(v-x)u(v,\theta(x))f(v)dv\int K_h(v-x)u^T(v,\theta(x))f(v)dv$$

- $h \to 0$. Singularity. Higher order Taylor expansion. Large variance. Requires much smoothing. Difficult to compute

Regularity conditions

- Regularity conditions for consistency and asymptotic normality as $n \to \infty$ and $h \to 0$:

- (i) Geometric ergodicity of $\{X_t\}$
- (ii) $\log n / n h_1^5 h_2^5 \to 0$
- (iii) K Lipschitz
- (iv) $\phi(\cdot, \theta(x))$ nondegenerate and

$$E\left[K_h(X_t - x)\frac{\partial}{\partial \theta_j}\phi(X_t, \theta(x))\right]^\gamma < \infty$$

for some $\gamma > 2$.

Choice of bandwidth

- Use local likelihood on $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots X_n$ to find $\hat{\theta}_{-i}(x)$
- Global likelihood cross validation: Maximize

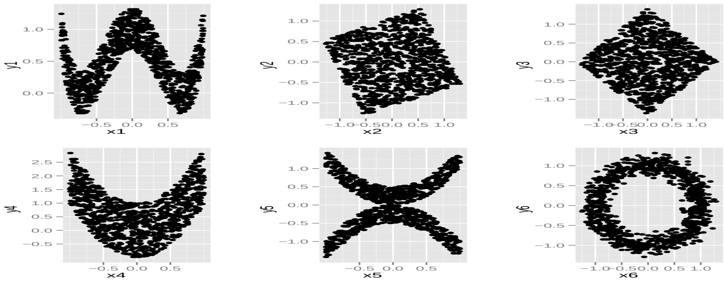
$$\mathrm{Cv}(h) = \sum_{i=1}^{n} \log[\phi(X_i, \hat{\theta}_{-i}(X_i)]w(X_i)$$

Applications

- Description of local dependence
- Tests of independence, including serial dependence
- Tail dependence
- Identification of copula model
- Tests of normality
- Local principal components
- Test of contagion
- Asymmetry in local correlation
- Local spectral estimation
- Multivariate densities in higher dimensions
- Conditional densities

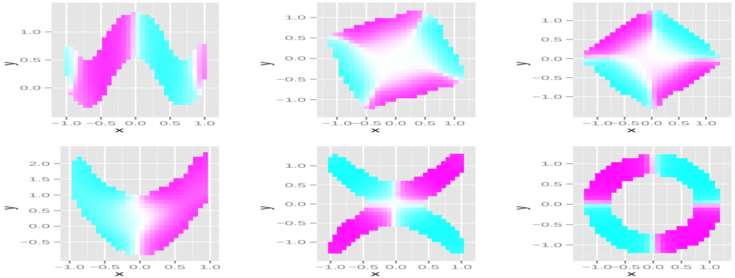
Visualizing nonlinear dependence and tests of independence

Example: wikipedia.org



Visualizing nonlinear dependence and tests of independence

Example: wikipedia.org



Tests of serial independence in time series

- Independence $\Rightarrow \rho_k(x_1, x_2) \equiv 0$. Independence $\Leftrightarrow \rho_{c,k}(x_1, x_2) \equiv 0$
- Test functional: $T_k = \int g(\rho_k(x))dF(x)$ with estimate $\hat{T}_k = \frac{1}{n}\sum g(\rho_k(X_t, X_{t-k}))$
- Consistency: Under regularity conditions $\hat{T}_k \overset{a.s.}{\to} T_k$.
- Asymptotic normality: $\sqrt{n}[C_k(A_h)]^{-1/2}(\hat{T}_k - T_k) \overset{d}{\to} \mathcal{N}(0,1)$, where $C_k = O(1)$ in the canonical case and $C_k \sim (h_1 h_2)^{-2}$ in the non-canonical case.
- Bootstrap: Essential in practice. Its validity shown in Lacal and Tjostheim (2015).
- Local crosscorrelation for bivariate series and the block bootstrap in Lacal and Tjøstheim (2016).

Simulation experiments

- GARCH: $X_t = Z_t\sqrt{h_t}, \;\; h_t = 1 + c(X_{t-1}^2 + h_{t-1})$
- ARCH: $X_t = Z_t\sqrt{h_t}, \;\; h_t = 1 + cX_{t-1}^2$
- AR: $X_t = cX_{t-1} + Z_t$
- EAR:$X_t = ce^{-X_{t-1}^2} + Z_t$
- In all cases Z_t is Gaussian, $g(\rho) = \rho^2$ and the canonical correlation has been used. We have compared the local correlation (solid line), the Brownian distance correlation (dashed line) and the Pearson correlation (dotted line).

Power

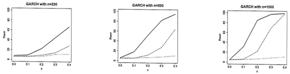


Figure 1

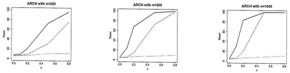


Figure 2

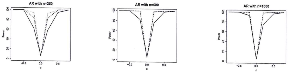


Figure 3

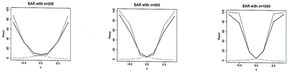
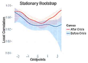
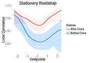


Figure 4

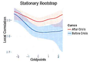
Local correlation indices



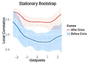
(a) CAC-DAX



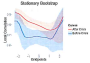
(b) CAC-FTSE



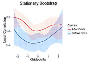
(c) DAX-FTSE



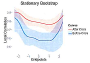
(d) DAX-OBX



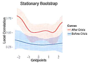
(e) FTSE-OBX



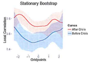
(f) FTSE-S&P



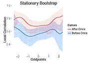
(g) OBX-CAC



(h) OBX-S&P

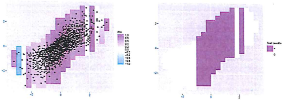


(i) S&P-CAC



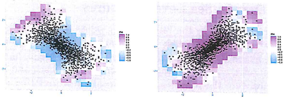
(j) S&P-DAX

Real data: Darwin Sea Level Pressure



(a) (b)

~~Figure 7~~: Local Gaussian correlation (7a) and dependence map (7b) for the Darwin sea level pressures.



(a) (b)

~~Figure 8~~: Local Gaussian correlation with $k = 6$ (8a) and $k = 12$ (8b) for the Exchange Rate.

Multivariate density estimation

- LGDE: Local Gaussian Density Estimate

$$f(x) \sim \psi(x, \mu(x), \Sigma(x))$$

- : First: Transformation of marginals

$$Z_j = (\Phi^{-1}(\hat{F}_1(X_{j1})), \ldots, \Phi^{-1}(\hat{F}_p(X_{jp})))^T$$

- Second: New joint

$$f_Z(z) = f(F_1^{-1}(\Phi(z_1)), \ldots, F_p^{-1}(\Phi(z_p))) \prod_{i=1}^{p} \{d/dzF_i^{-1}(z_i)\phi(z_i)$$

- Original density

$$f(x) = f_Z(\Phi^{-1}(F_1(x_1), \ldots, \Phi^{-1}(F_p(x_p)))) \prod_{i=1}^{p} \frac{f_i(x_i)}{\phi(\Phi^{-1}(F(x_i)))}$$

Multivariate density estimation

- For transformed data

$$\psi(z,\theta) = \psi(z,R) = (2\pi)^{-p/2}|R|^{-1/2}\exp\Big\{-\frac{1}{2}z^T R^{-1} z\Big\}$$

  where R is the correlation matrix $R = \{\rho_{ij}\}$.
- Simplification

$$\rho_{ij}(z_1,\ldots,z_p) = \rho_{ij}(z_i,z_j)$$

  which is estimated using local likelihood on (Z_i, Z_j) only and with population quantity ρ_{ij} being only dependent on $f_{Z_i,Z_j}(z_i,z_j)$.

Multivariate density estimation

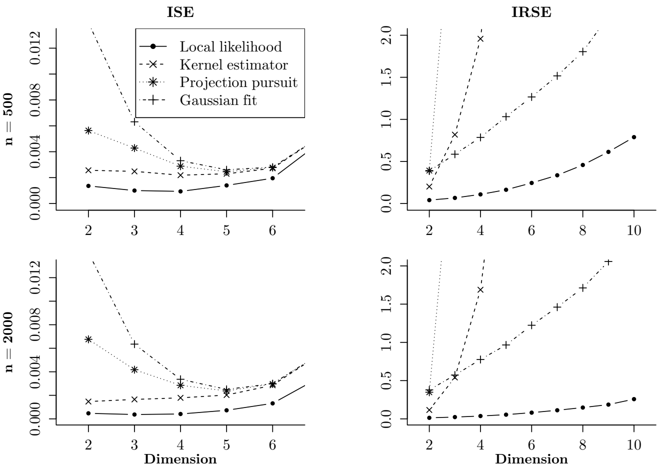
- Convergence rate: $\hat{\rho}_{ij}$ has convergence rate $(nh^2)^{1/2}$ towards ρ_{ij}
- Density estimate

$$\hat{f}_0(x) = \hat{f}_{Z,0}(\Phi^{-1}(\hat{F}_1(x_1)), \ldots, \Phi^{-1}(\hat{F}_p(x_p))) \prod_{i=1}^{p} \frac{\hat{f}_i(x_i)}{\phi(\Phi^{-1}(\hat{F}_i(x_i)))}$$

which has convergence rate $(nh^2)^{1/2}$, but which is generally (due to the simplification) estimating an *approximation* of f, not f itself.

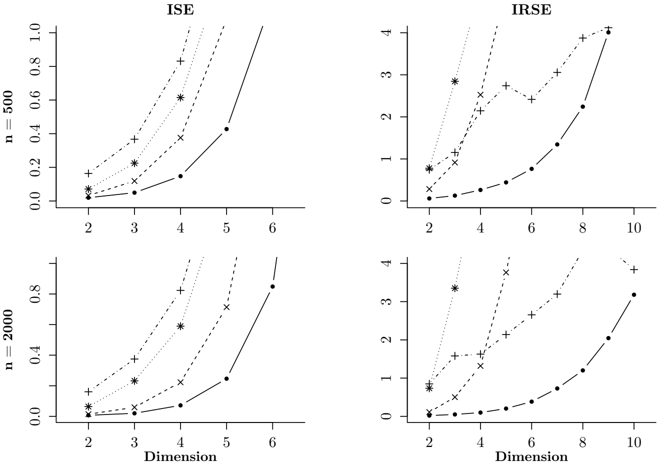
Multivariate density estimation

Example: Density estimation, normal copula, chi square marginals



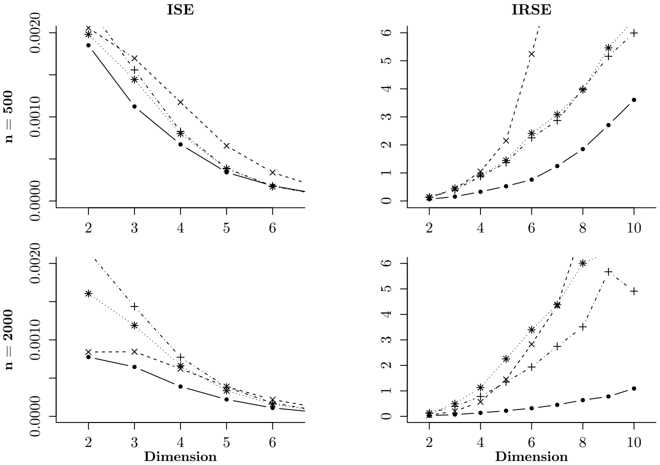
Multivariate density estimation

Example: Density estimation, t-copula, log normal marginals



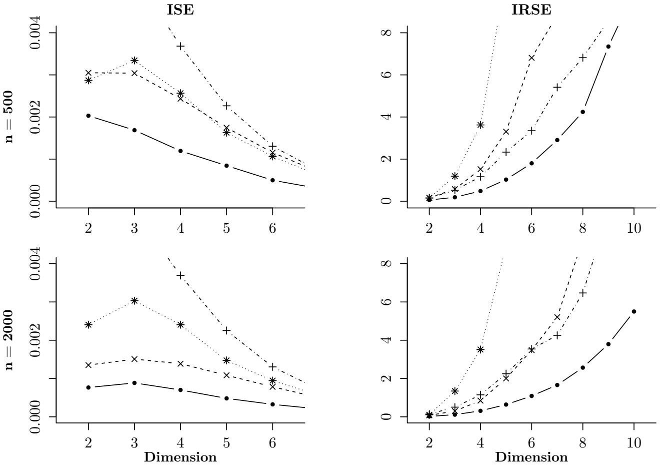
Multivariate density estimation

Example: Density estimation, Clayton copula, t marginals



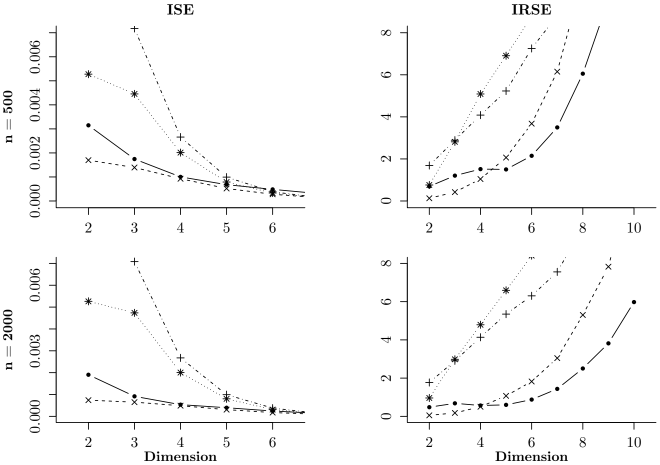
Multivariate density estimation

Example: Density estimation, Frank Copula, t-marginals



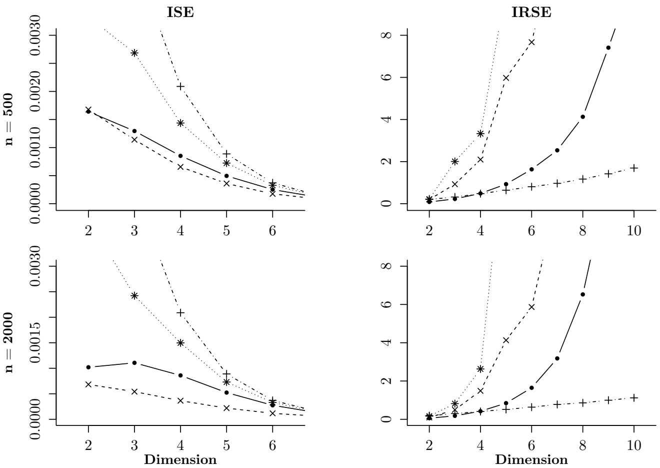
Multivariate density estimation

Example: Density estimation, mixed normal



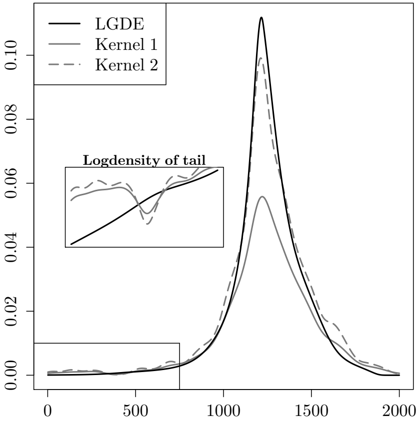
Multivariate density estimation

Example: Density estimation t4



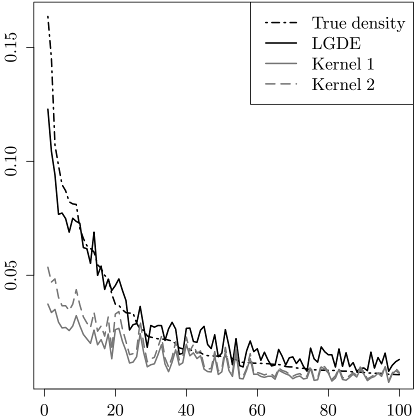
Multivariate density estimation

Example: Density estimation, central axis, 5-dim real data



Multivariate density estimation

Example: Density estimation, real data, pair copula model

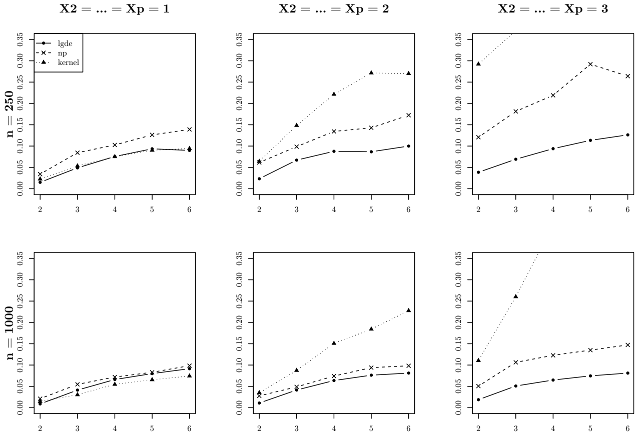


Conditional density

- Conditional density $f(x_1|x_2)$
- Kernel estimator $\hat{f}(x_1|x_2) = \frac{\hat{f}(x_1,x_2)}{\hat{f}(x_2)}$
- Conditional density local Gaussian: $\frac{\psi(u_1,u_2,\rho(x_1,x_2))}{\psi(u_2,\rho(x_2))}$ where $u = x$, and where again the conditional density is Gaussian with conditional mean $\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$ and covariance $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Using this seems to work very well in practice.

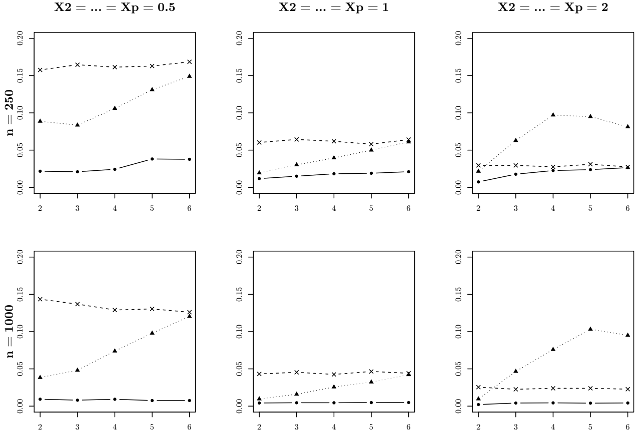
Conditional density

Example: Integrated square error cond. dens.
$X_1|X_2,\ldots,X_p = 1,2,3$, exp. marginals, Joe Copula



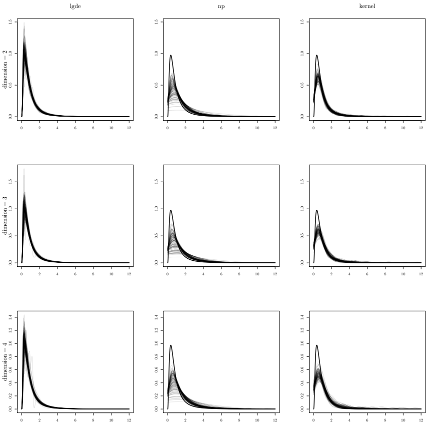
Conditional density

Example: Integrated square error cond. dens $X_1|X_2, ..., X_p = 0.5{,}1{,}2$ $n = 250, n = 1000$, two first components log-normal with a t(1) copula, the rest independent from first two and multivariate t(5)



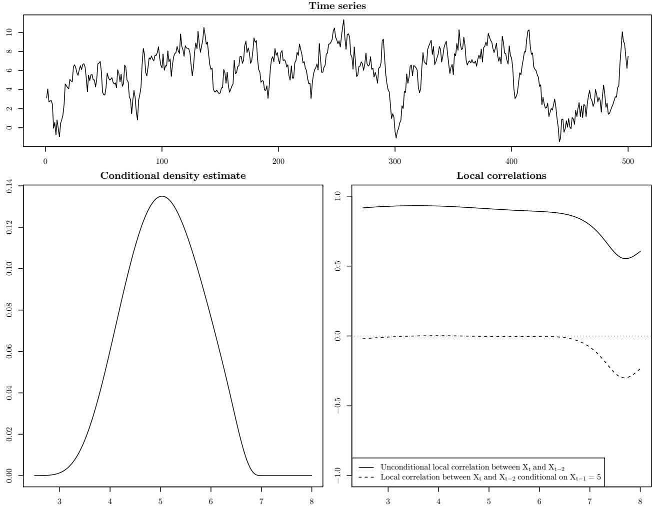
Conditional density

Example: Densities cond.dens. $X_1|X_2, ..., X_p = 0.5, n = 500$, two first components log-normal with a t(1) copula, the rest independent from first two and multivariate t(5)



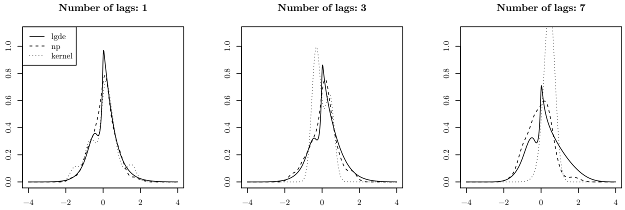
Conditional density

Example: Time series $X_t = 0.8X_{t-1} + 0.5\sqrt{|X_{t-1}|} + e_t$. Local unconditional and conditional correlation $(X_t, X_{t-2})|X_{t-1}$,



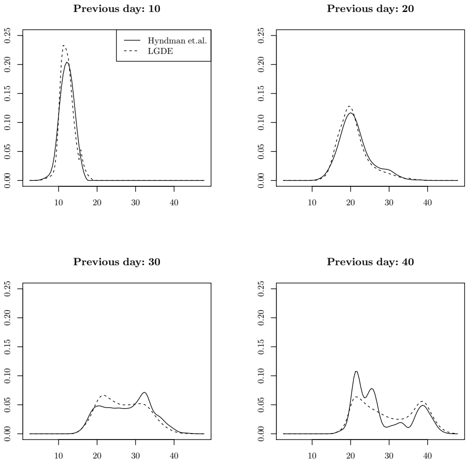
Conditional density

Example: Cond. US log-returns given 1, 3, 7 preceding days equal to -1



Conditional density

Example: Melbourne temperature data, with estimated conditional density of the maximum daily air temperature, given a preceding recording of 10, 20, 30 and 40 degrees Celsius, respectively



Conditional density

VaR: Value at Risk

	Level		
Method	0.005	0.01	0.05
LGDE	0.014	0.017	0.072
np	0.084	0.097	0.161
Kernel	0.117	0.134	0.187
Gaussian	0.045	0.064	0.125

Table: Proportion of observations exceeding the estimated VaR

Summary

- Local Gaussian correlation gives a much more detailed picture than ordinary correlation
- Population value can be defined
- It can be estimated by local likelihood
- Applications: Dependence testing, copula characterization, contagion, asymmetries in financial markets, density estimation