Applications of Distance Correlation to Time Series : 距離の相関係数の時系列解析への応用

Muneya Matsui : 松井 宗也

KakenSympo@Kanazawa : 科研費シンポ@金沢

1. Dec. / 2018

Flow chart of my talk

Informative example: Amazon daily stock returns

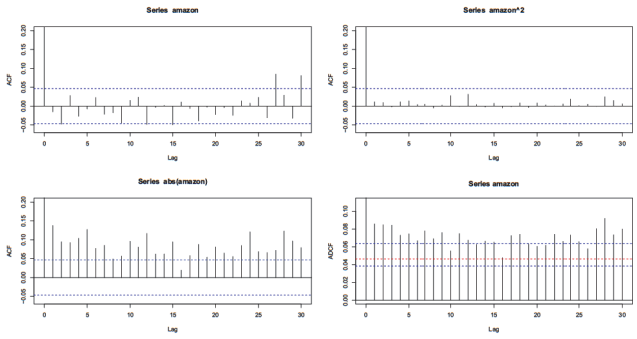

Figure : ACF and ADCF of (X_t) from 05/16/1997 to 06/16/2004. Up left: ACF of (X_t); Up right: ACF of (X_t^2); Low left: ACF of $(|X_t|)$; Low right: ADCF of (X_t), the 5%, 50%, 95% confidence bounds of ADCF from randomly permuting the data (mimic iid case).

Definition of DCVF and DCF and literature:

Given vectors $\boldsymbol{X} \in \mathbb{R}^p$ and $\boldsymbol{Y} \in \mathbb{R}^q$, the *distance covariance* (DCVF) between $\boldsymbol{X}$ and $\boldsymbol{Y}$ w.r.t. a suitable measure μ on $\mathbb{R}^{p+q}$ is given by

$$T(X,Y;\mu) = \int_{\mathbb{R}^{p+q}} \left|\varphi_{X,Y}(s,t) - \varphi_X(s)\,\varphi_Y(t)\right|^2 \mu(ds,dt)\,, \quad (1.1)$$

where ch.f. of any r.v. $\boldsymbol{Z} \in \mathbb{R}^d$ is denoted by $\varphi_{\boldsymbol{Z}}(\boldsymbol{t}) = \mathbb{E}[\mathrm{e}^{\,i\langle t,Z\rangle}]$.

Testing independence.

Székely et al. (2007) : $\mu(ds,dt) = |s|^{-2}|t|^{-2}dsdt$
Székely and Rizzo (2014,2009) : $\mu(ds,dt) = c_{p,q}|s|^{-\alpha-p}|t|^{-\alpha-q}dsdt$
with $\alpha \in (0,2)$. with this choice distance correlation (DCF) $T(X,Y;\mu)/(T(X,X;\mu)T(Y,Y;\mu))^{1/2}$ is invariant relative to scale and orthogonal transformations.

Time series: Zhou (2012), Fokianos and Pitsillou (2016), and Hong (1999), DMMPs : Davis, Matsui, Mikosch and Wan (2018).
Stochastic Processes: Matsui, Mikosch and Samorodnitsky (2017).

Definition of DCVF and DCF and literature:

Classical measure of correlation:
Kendall rank correlation coefficient
Spearman's rank correlation coefficient
can not detect all relations
(could be zero even when $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not independent)

Test of independence: empirical dist. func.
KolmogorovSmirnov type test e.g. Bulm et al. (1961)
Cramérvon Mises type test (Hoeffding's test of independence)
low quality of the empirical dist. func. when dimensions are high
Gretton and Györfi (2010)

Equivalent to Hilbert-Schmidt Independence Criterion (HSIC)
related to data analysis using reproducing kernel Hilbert space (RKHS)

Condition for existence:

Let $(\boldsymbol{X}',\boldsymbol{Y}')$ be an independent copy of $(\boldsymbol{X},\boldsymbol{Y})$, and let $\boldsymbol{Y}''$ and $\boldsymbol{Y}'''$ be independent copies of $\boldsymbol{Y}$ which are also independent of $(\boldsymbol{X},\boldsymbol{Y}),(\boldsymbol{X}',\boldsymbol{Y}')$. We have

$$\begin{aligned} &T(X,Y;\mu) \\ &= \int_{\mathbb{R}^{p+q}} \mathbb{E}\Big[\mathrm{e}^{\,i\langle s,X-X'\rangle + i\langle t,Y-Y'\rangle} + \mathrm{e}^{\,i\langle s,X-X'\rangle}\,\mathrm{e}^{\,i\langle t,Y''-Y'''\rangle} \\ &\qquad - \mathrm{e}^{\,i\langle s,X-X'\rangle + i\langle t,Y-Y''\rangle} - \mathrm{e}^{\,-i\langle s,X-X'\rangle - i\langle t,Y-Y''\rangle}\Big]\mu(ds,dt)\,. \end{aligned}$$

$T(X,Y;\mu)$ is not always well-defined...
What is the condition for μ ?
Lemmas 2.1 2.3 2.6 in DMMPs

Examples : explicit cases

1. μ has density w given by $w(s,t) = c_{p,q}\,|s|^{-\alpha-p}\,|t|^{-\alpha-q}$,

$$T = \mathbb{E}[|X-X'|^\alpha\,|Y-Y'|^\alpha] + \mathbb{E}[|X-X'|^\alpha]\,\mathbb{E}[|Y-Y'|^\alpha] - 2\,\mathbb{E}[|X-X'|^\alpha\,|Y-Y''|^\alpha]\,.$$

2. independent symmetric Z_1 and Z_2 with multivariate β-stable for some $\beta \in (0,2]$. joint ch.f $\varphi_{Z_1,Z_2}(x,y) = \mathrm{e}^{-(|x|^\beta+|y|^\beta)}$,

$$T = \mathbb{E}[\mathrm{e}^{-(|X-X'|^\beta+|Y-Y'|^\beta)}] + \mathbb{E}[\mathrm{e}^{-|X-X'|^\beta}]\,\mathbb{E}[\mathrm{e}^{-|Y-Y'|^\beta}] - 2\,\mathbb{E}[\mathrm{e}^{-(|X-X'|^\beta+|Y-Y''|^\beta)}]\,.$$

3. sub-Gaussian $\alpha/2$-stable random vectors with ch.f. $-\log\varphi_{Z_1,Z_2}(x,y) = |(x,y)'\Sigma(x,y)|^{\alpha/2}$, where Σ is the covariance matrix of $(x,y) \in \mathbb{R}^{p+q}$ for any $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$. Then

$$T = \mathbb{E}\big[|(X-X',Y-Y')'\Sigma\,(X-X',Y-Y')|^{\alpha/2} + [|(X-X',Y''-Y''')'\Sigma\,(X-X',Y''-Y''')|^{\alpha/2} - 2[|(X-X',Y-Y'')'\Sigma\,(X-X',Y-Y'')|^{\alpha/2}\big]\,.$$

Empirical DCV for time series

empirical DCV for a stationary time series $((X_t,Y_t))$ with generic element (X,Y) where $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$:

$$T_n(X,Y;\mu) = \int_{\mathbb{R}^{p+q}} |\varphi^n_{X,Y}(s,t) - \varphi^n_X(s)\,\varphi^n_Y(t)|^2\,\mu(ds,dt)\,,$$

where the empirical ch.f. is given by
$\varphi^n_{X,Y}(s,t) = \frac{1}{n}\sum_{j=1}^n \mathrm{e}^{i\langle s,X_j\rangle + i\langle t,Y_j\rangle}$, $n \geq 1$, and
$\varphi^n_X(s) = \varphi^n_{X,Y}(s,0)$ and $\varphi^n_Y(s) = \varphi^n_{X,Y}(0,t)$.

Consistency Theorem 3.1 in DMMPs
Consider a stationary ergodic time series $((X_j,Y_j))_{j=1,2,\ldots}$ with values in $\mathbb{R}^{p+q}$ and assume one of the existence conditions are satisfied. Then

$$T_n(X,Y;\mu) \overset{\text{a.s.}}{\to} T(X,Y;\mu) \quad \text{as} \quad n \to \infty.$$

Testing serial dependence

cross-distance covariance function (CDCVF) of $((X_t,Y_t))$:

$$T^{X,Y}_\mu(h) = T(X_0,Y_h;\mu)\,, \qquad h \in \mathbb{Z}\,,$$

auto-distance covariance function (ADCVF) of (X_t) :

$$T^X_\mu(h) = T^{X,X}_\mu(h)\,, \qquad h \in \mathbb{Z}\,.$$

The empirical versions $T^X_{n,\mu}$ and $T^{X,Y}_{n,\mu}$ are defined correspondingly. e.g. replace $\varphi^n_{X,Y}(s,t)$ in the definition of $T_n(X,Y;\mu)$ by

$$\varphi^n_{X_0,Y_h}(s,t) = \frac{1}{n}\sum_{j=1}^{n-h} \mathrm{e}^{\,i\langle s,X_j\rangle + i\langle t,Y_{j+h}\rangle}, \quad n \geq h+1\,,$$

Corresponding correlation func. (CDCF) and (ADCF) respectively ;

$$R^{X,Y}_\mu(h) = \frac{T^{X,Y}_\mu(h)}{\sqrt{T^X_\mu(0)\,T^Y_\mu(0)}} \qquad \textit{and} \qquad R^X_\mu(h) = \frac{T^X_\mu(h)}{T^X_\mu(0)}.$$

Consistency and weak convergence are also proven.

ADCVF and ADCF of fitted residuals from $\mathrm{AR}(p)$ process

$$\mathrm{AR}(p): \qquad X_t = \sum_{k=1}^p \phi_k X_{t-k} + Z_t\,, \quad t = 0,\pm1,\ldots,$$

where (Z_t) is an iid with $\mathbb{E}[|Z|^\kappa] < \infty$, $\kappa > 0$ and if $\kappa \geq 1$ mean 0. write AR(p): $Z_t = X_t - \phi^T \mathbf{X}_{t-1}$, where $\phi = (\phi_1,\ldots,\phi_p)^T$, $p \geq 1$ and $\mathbf{X}_t = (X_t,\ldots,X_{t-p+1})^T$.

$$\sqrt{n}\,(\widehat{\phi}-\phi) \overset{d}{\to} \mathbf{Q} \sim N(0,\sigma^2\Gamma_p^{-1}), \tag{3.2}$$

where $\Gamma_{n,p} \overset{\text{a.s.}}{\to} \Gamma_p = \big(\gamma_X(j-k)\big)_{1\leq j,k\leq p}$ The residuals

$$\widehat{Z}_t = X_t - \widehat{\phi}^T\,\mathbf{X}_{t-1} = (\phi-\widehat{\phi})^T\,\mathbf{X}_{t-1} + Z_t\,, \quad t = p+1,\ldots,n \tag{3.3}$$

$$T^{\widehat{Z}}_{n,\mu}(h) = \int_{\mathbb{R}} |C^{\widehat{Z}}_n(s,t)|^2\,\mu(ds,dt),$$

where $C^{\widehat{Z}}_n(s,t) = \varphi^n_{\widehat{Z},\widehat{Z}_{+h}}(s,t) - \varphi^n_{\widehat{Z}}(s)\varphi^n_{\widehat{Z}_{+h}}(t)$.

Weak convergence of T_n

Assume that $((X_j,Y_j))$ is a strictly stationary with values in $\mathbb{R}^{p+q}$
(α_h) : α-mixing with rate satisfies $\sum_h \alpha_h^{1/r} < \infty$ for some $r > 1$.
$u = 2r/(r-1)$, $X = (X^{(1)},\ldots,X^{(p)})$, $Y = (Y^{(1)},\ldots,Y^{(q)})$.
Assume that X_0 and Y_0 are independent and for some $\alpha \in (u/2,u]$, $\epsilon \in [0,1/2)$ and $\alpha' \leq \min(2,\alpha)$, the following hold:

$$\mathbb{E}[|X|^\alpha + |Y|^\alpha] < \infty, \quad \mathbb{E}\Big[\prod_{l=1}^p |X^{(l)}|^\alpha\Big] < \infty, \quad \mathbb{E}\Big[\prod_{l=1}^q |Y^{(l)}|^\alpha\Big] < \infty,$$

$$\int_{\mathbb{R}^{p+q}} (1\wedge|s|^{\alpha'(1+\epsilon)/u})(1\wedge|t|^{\alpha'(1+\epsilon)/u})\,\mu(ds,dt) < \infty.$$

Then

$$n\,T_n(X,Y;\mu) \overset{d}{\to} \|G\|^2_\mu = \int_{\mathbb{R}^{p+q}} |G(s,t)|^2\,\mu(ds,dt)\,,$$

where G is a complex-valued mean-zero Gaussian process.

Weak convergence of $T^{\widehat{Z}}_{n,\mu}(h)$

Consider a causal $\mathrm{AR}(p)$ process with iid noise (Z_t). Assume

$$\int_{\mathbb{R}^2} (1\wedge|s|^2)(1\wedge|t|^2)\mu(ds,dt) + (s^2+t^2)1(|s|\wedge|t|>1)\mu(ds,dt) < \infty.$$

If $\sigma^2 = \mathrm{Var}(Z) < \infty$ (omit infinite variance case), then

$$n\,T^{\widehat{Z}}_{n,\mu}(h) \overset{d}{\to} \|G_h+\xi_h\|^2_\mu \quad \text{and} \quad n\,R^{\widehat{Z}}_{n,\mu}(h) \overset{d}{\to} \frac{\|G_h+\xi_h\|^2_\mu}{T^Z_\mu(0)}\,,$$

where (G_h,ξ_h) are jointly Gaussian limit random fields on $\mathbb{R}^2$.
The idea of PF:

$$T^{\widehat{Z}}_{n,\mu}(h) = \int_{\mathbb{R}} |C^{\widehat{Z}}_n(s,t)|^2\,\mu(ds,dt),$$

$\sqrt{n}\,(C^Z_n, C^{\widehat{Z}}_n - C^Z_n) \overset{d}{\to} (G_h,\xi_h)$, $\sqrt{n}\,C^{\widehat{Z}}_n \overset{d}{\to} G_h+\xi_h$.
The original Székely's measures do not fit...

AR simulations 1

Indep. rep. of a $\mathrm{AR}(10)$ time series $(X_t)_{t=1,\ldots,1000}$ with $Z_t \sim N(0,1)$. $\mu := \mu_1 \times \mu_2$, where $\mu_i \sim N(0, 0.5)$. Approximate the limit distribution $\|G_h + \xi_h\|_\mu^2 \,/\, T_\mu^Z(0)$ of $n\, R_{n,\mu}^{\widehat{Z}}(h)$.

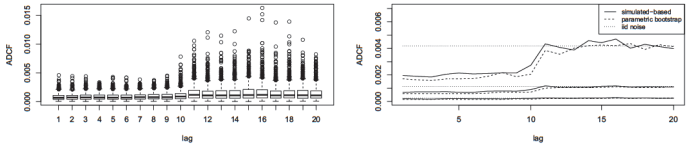

Figure : Left: Box-plots from 1000 independent replications. Right: 5%, 50%, 95% empirical quantiles of $n\, R_{n,\mu}^{\widehat{Z}}(h)$ based on simulated residuals, on resampled residuals and on iid noise.

The asymptotic variance of the ADCF of the residuals is smaller than that of iid noise at initial lags, and gradually increases. Similar to that of ACF Ch.9.4 of B&D. Parametric bootstrap provides a good approximation.

AR simulations 2

AR(10) with $t_{1.5}$. Left: box-plots of $n\, R_{n,\mu}^{\widehat{Z}}(h)$. Right: quantiles of $n\, R_{n,\mu}^{\widehat{Z}}(h)$ and $n\, R_{n,\mu}^{Z}(h)$, both $\xrightarrow{d} \|G_h\|_\mu^2 \,/\, T_\mu^Z(0)$. The quantiles of $\|G_h\|_\mu^2 \,/\, T_\mu^Z(0)$ can be approximated naively by bootstrapping the fitted residuals $(\widehat{Z}_t)$. $\mu = \mu_1 \times \mu_2$, with each $\mu_i \sim N(0, 0.5)$.

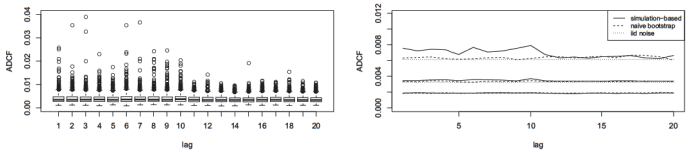

Figure : Distribution of $n\, R_{n,\mu}^{\widehat{Z}}(h)$ for residuals of AR process with $t_{1.5}$ innovations. Left: lag-wise box-plots. Right panel: empirical 5%, 50%, 95% quantiles from simulated residuals, from resampled residuals and from iid noise.

The agreement is reasonably good.

AR simulations 3

Example illustrating the limitation of using the measure by Székely et al. with $Z_t \sim$the symmetric $G(.2, .5)$.

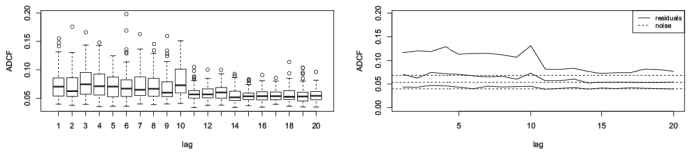

Figure : Distribution of $n\, R_{n,\mu}^{\widehat{Z}}(h)$, $n = 1000$ for residuals of AR process with a symmetric Gamma(0.2,0.5) noise. Left: box-plots from 500 indep. rep. Right: empirical 5%, 50%, 95% quantiles from simulated residuals and from iid noise.

first 10 lags are spread out compared to those at lags greater than 10. This illustrates the problem with using the Székely's measure as a weight function applied to the residuals.

Wind Speed Data

Wind speeds at Kilkenny's synoptic meteorological station in Ireland.

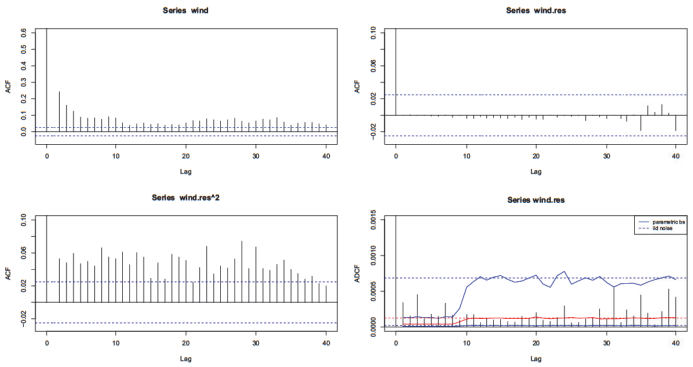

Figure : ACF and ADCF, and AR(9) fitted residuals, ACF of the series, the residuals and the residual squares. The 5%, 50%, 95% confidence bounds of ADCF for fitted residuals from 1000 parametric bootstraps.

Wind Speed Data

some doubt on the validity of an AR(9) model with iid noise for this data. $\Rightarrow$ consider a GARCH(1,1) model applied to the residuals

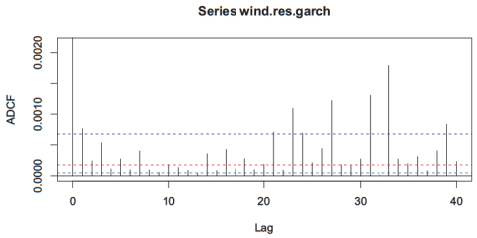

Figure : ADCF of the residuals of Kilkenny wind speed time series from AR(9)-GARCH fitting and the 5%, 50%, 95% confidence bounds of ADCF for iid noise from 1000 random permutations.

a periodic AR model, which allows for periodicity in both the AR parameters and white noise variance might be a more desirable model.

References

Davis, R.A., Matsui, M., Mikosch, T. and Wan, P. (2018) Applications of distance correlation to time series. *Bernoulli* **24**, 3087–3116.

Dehling, H., Matsui, M., Mikosch, T., Samorodnitsky, G. and Tafakori, L. (2018) Distance covariance for discretized stochastic processes. *arXiv:1806.09369*.

Fokianos, K. and Pitsillou, M. (2016) Consistent testing for pairwise dependence in time series. *Technometrics* **59**, 262–270.

Gretton, A. and Györfi, L. (2010) Consistent nonparametric tests of independence. *J. Mach. Learn. Res.* **11**, 1391–1423.

Hong, Y. (1999) Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach. *J. Amer. Statist. Assoc.* **94:448**, 1201–1220.

Matsui, M., Mikosch, T. and Samorodnitsky, G. (2017) Distance covariance for stochastic processes. *Probab. Math. Statist.* **37**, 355-372.

Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007) Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35**, 2769–2794.

Székely, G.J. and Rizzo, M.L. (2009) Brownian distance covariance. *Ann. Appl. Stat.* **3**, 1236–1265.

Zhou, Z. (2012) Measuring non linear dependence in time-series, a distance correlation approach. *J. Time Series Anal.* **33**, 438–457.