Autoregressive systems under stationary noises

Alexandre Brouste[1] Cai Chunhao[1]
Marina Kleptsyna[1]

[1]University of Le Mans, France

March 12, Paris

Outline

Outline

Model I

Observations

$$X_t = \sum_{j=1}^{p} \theta_j X_{t-j} + \xi_t, \quad T \geq t \geq p, \quad X_j = x_j, \quad 0 \leq j < p.$$

with

- — unknown $\theta \in R^p$
- —-centered **stationary (not independent)** Gaussian sequence of noises ξ.

The goal —– to estimate θ

Model I

Observations

$$X_t = \sum_{j=1}^{p} \theta_j X_{t-j} + \xi_t, \quad T \geq t \geq p, \quad X_j = x_j, \quad 0 \leq j < p.$$

with

- — unknown $\theta \in R^p$
- —-centered **stationary (not independent**) Gaussian sequence of noises ξ.

The goal —— to estimate θ

Model I

Observations

$$X_t = \sum_{j=1}^{p} \theta_j X_{t-j} + \xi_t, \quad T \geq t \geq p, \quad X_j = x_j, \quad 0 \leq j < p.$$

with

- — unknown $\theta \in R^p$
- —-centered **stationary (not independent**) Gaussian sequence of noises ξ.

The goal —— to estimate θ

Maximum Likelihood Estimator

Question for discussion

- The structure of the Likelihood function
- The long time asymptotic properties of the MLE
- Comparison with i.i.d. noises case

Outline

Model II

Observation Model

$$\begin{cases} X_n = \sum_{j=1}^{p} \theta_j X_{n-j} + \xi_n, & n \geq p, \quad X_j = x_j,\ 0 \leq j < p \\ Y_n = \mu X_n + \tilde{\xi}_n, & n \geq 1, \quad Y_0 = y. \end{cases}$$

- the goal — to write the minimal complexity algorithm to filter X
- the question: the structure of the innovation process

Model II

Observation Model

$$\begin{cases} X_n = \sum_{j=1}^{p} \theta_j X_{n-j} + \xi_n, & n \geq p, \qquad X_j = x_j,\ 0 \leq j < p \\ Y_n = \mu X_n + \tilde{\xi}_n, & n \geq 1, \qquad Y_0 = y. \end{cases}$$

- the goal — to write the minimal complexity algorithm to filter X
- the question: the structure of the innovation process

Outline

Noises

Noises

$$E\xi_m\xi_n = c(m,n) = \rho(|\,n-m|), \quad \rho(0)=1,$$

$c(.,.)$ is positive definite function.

Assumptions about noises

assumptions, spectral density

$$\int_{-\pi}^{\pi} |\ln(f(\lambda)| \, d\lambda < \infty.$$

sufficiently, covariance

$$|\rho(t)| \leq ct^{-\alpha}, \ \alpha > 0.$$

Examples: ARMA(p,q), fGn, Mixed fGn....

Assumptions on the model

Let Θ be the companion matrix

$$\Theta = \begin{pmatrix} \theta_1 & \theta_2 & \cdots & \cdots & \theta_p \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

We suppose that the spectral radius of Θ :

$$r(\Theta) < 1.$$

Outline

Longtime Properties of MLE

Longtime Properties of MLE

- MLE $\hat{\theta}_T$ is uniformly on compacts $\mathbb{K} \subset \mathbb{R}^p \bigcap r(\Theta) < 1$ consistent.
- Uniformly asymptotically normal:

$$\sqrt{T}\left(\hat{\theta}_T - \theta\right) \xrightarrow{law} \mathcal{N}\left(0, \mathcal{I}(\theta)^{-1}\right)$$

- Uniform on $\vartheta \in \mathbb{K}$ convergence of the moments.

The Fisher information

$\mathcal{I}(\theta)$ —-the Fisher information **does not depend** on the structure of ξ.

The Fisher information

$\mathcal{I}(\theta)$ is the unique nonnegative defined solution of the Lyapounov equation

$$\mathcal{I}(\theta) = \Theta\mathcal{I}(\theta)\Theta' + bb',$$

where Θ is the companion matrix and $b = \begin{pmatrix} 1 \\ \mathbf{0}_p \end{pmatrix}$

Filtering: representation theorem

The representation Theorem holds:
Let $(\zeta_n)_{n\geq 1}$ be the $2p$-dimensional autoregressive process defined by

$$\zeta_n = A_{n-1}\zeta_{n-1} + \sigma_n b\varepsilon_n, \quad n \geq 1, \quad \zeta_0 = \begin{pmatrix} x \\ \mathbf{0}_p \end{pmatrix}$$

with

$$A_n = \begin{pmatrix} \Theta & \Theta\beta_n \\ \beta_n Id_p & Id_p \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \mathbf{0}_p \end{pmatrix}, \quad x = \begin{pmatrix} x_0 \\ \vdots \\ x_{p-1} \end{pmatrix},$$

Then the following representations hold:

$$\begin{aligned} X_n &= \sum_{m=1}^{n} K(n,m) b^* \zeta_m, \\ Z_n &= \mu b^* \zeta_n + \sigma_n \tilde{\varepsilon}_n, \quad n \geq 1. \end{aligned}$$

Outline

Stationary Gaussian sequences

Let $\xi = (\xi_n)_{n\geq 0}$ be a centered stationary Gaussian sequence.

- innovations–

$$\sigma_1\varepsilon_1 = \xi_1, \quad \sigma_n\varepsilon_n = \xi_n - \mathbf{E}(\xi_n \,|\, \xi_1, \ldots, \xi_{n-1}), \quad n \geq 2.$$

- a deterministic kernel $k = (k(n,m),\, n \geq 1,\, m \leq n)$ such that $k(n,n) = 1$ and

$$\sigma_n\varepsilon_n = \sum_{m=1}^{n} k(n,m)\xi_m.$$

- the partial correlation coefficient

$$\beta_{n-1} = -k(n,1),\ n \geq 1.$$

Levinson-Durbin algorithm

Relations between $k(.,.)$, $\rho(.)$, (β_n) and σ_n^2, $n \geq 1$:

$$\sigma_n^2 = \prod_{m=1}^{n-1} (1 - \beta_m^2) \,,$$

$$\sum_{m=1}^{n} k(n,m)\rho(m) = \beta_n \sigma_n^2 \,,$$

$$k(n+1, n+1-m) = k(n, n-m) - \beta_n k(n,m) \,.$$

Inverse Kernel

The positive definiteness of the covariance $c(.,.) \Rightarrow$ exists an inverse deterministic kernel $K = (K(n,m), n \geq 1, m \leq n)$ such that

$$\xi_n = \sum_{m=1}^{n} K(n,m) \sigma_m \varepsilon_m.$$

Actually, kernels k and K are nothing but the ingredients of the Choleski decomposition of covariance and inverse of covariance matrices.

$$\Gamma_n^{-1} = k_n D_n^{-1} k_n'; \; \Gamma_n = K_n' D_n K_n$$

Outline

Transformation of the observations

Let us introduce $(Z_n)_{n\geq 1}$ such that

$$Z_n = \sum_{m=1}^{n} k(n,m)X_m, \quad n \geq 1.$$

Then we have also

$$X_n = \sum_{m=1}^{n} K(n,m)Z_m.$$

Equivalent Markov nonergodic observation model

The initial observation model is equivalent to the model, based on the $2p$-dimensional autoregressive process :

$$\zeta_n = A_{n-1}\zeta_{n-1} + \sigma_n \, b \, \varepsilon_n, \quad n \geq 1, \quad \zeta_0 = \begin{pmatrix} x \\ \mathbf{0}_p \end{pmatrix}$$

with (Θ is the companion matrix, $(\varepsilon_n)_{n\geq 1}$ — i.i.d.)

$$A_n = \begin{pmatrix} \Theta & \Theta\beta_n \\ \beta_n Id_p & Id_p \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ \mathbf{0}_p \end{pmatrix}, \quad x = \begin{pmatrix} x_0 \\ \vdots \\ x_{p-1} \end{pmatrix},$$

Likelihood function

Likelihood function

The log-likelihood

$$\ln \mathcal{L}(Z^{(N)}, \vartheta) = -\frac{1}{2}\sum_{n=2}^{N}\left(\frac{\nu_n^\theta}{\sigma_n}\right)^2 - \frac{N}{2}\ln 2\pi - \frac{1}{2}\sum_{n=2}^{N}\ln \sigma_n^2$$

where $\nu_n^\theta = b' \left(\zeta_n - A_{n-1}\zeta_{n-1}\right)$.

MLE

MLE, $p = 1$

$$\hat{\vartheta}_N = \frac{\displaystyle\sum_{n=2}^{N} \frac{a_{n-1}^* \zeta_{n-1} Z_n}{\sigma_n^2}}{\displaystyle\sum_{n=2}^{N} \zeta_{n-1}^* \frac{a_{n-1} a_{n-1}^*}{\sigma_n^2} \zeta_{n-1}}$$

and $a_n = \begin{pmatrix} 1 \\ \beta_n \end{pmatrix}$.

Outline

Laplace Transform computations

It is sufficient to prove that

$$\lim_{N\to\infty} \mathbf{E}_\theta \exp\left(-\frac{\mu}{2N}\sum_{n=2}^{N} \zeta_{n-1}^* \mathcal{M}_{n-1} \zeta_{n-1}\right) = \exp\left(-\frac{\mu}{2}\mathcal{I}(\vartheta)\right)$$

with $\mathcal{M}_{n-1} = \frac{a_{n-1} a_{n-1}^*}{\sigma_n^2}$ and $\mathcal{I}(\theta) = (1-\vartheta^2)^{-1}$

Open questions

- Strong consistency for $p > 1$ in instable and explosive domain
- Asymptotic distribution (even for $p = 1$) in instable and explosive domain (dependence on the covariance of ξ)
- Estimation problem under partial observations