1 Some Results In Time Series Modeling And Analysis

1.1 Extended Sample $A.C.F$'s and Model Identification

Tsay & Tiao - JASA '79 1984-84-96

Consider the possibly non-stationary $ARMA\,(p{,}q)$ model:

$$\Phi\,(B)\,Z_t = \theta\,(B)\,a_t$$

$$\begin{aligned}\Phi\,(B) &= 1-\psi_1 B ... -\psi_p B^p \\ &= U\,(B)\,\psi\,(B) \\ U\,(B) &= 1-u_1 B ... - u_d B^d\end{aligned}$$

$$\begin{aligned}\psi\,(B) &= 1-\psi_1 B ... -\psi_{p-d} B^{p-d} \\ \theta\,(B) &= 1-\theta_1 B ... -\theta_q B^q\end{aligned}$$

B is the backshift operator $\{a_t\}$ $i.i.d.\ Gaussian\ \left(0{,}\sigma_a^2\right)$.

Zeros of $u\,(B)$ - **on** unit circle

Zeros of $\psi\,(B)$ - **outside** unit circle

Zeros of $\theta\,(B)$ - **outside** unit circle

$\Phi\,(B)$ and $\theta\,(B)$ have no common factors.

For Z_t non-stationary, Z_t begins at t_0.

If $U\,(B) = (1-B)^d$, the model is equivalent to an $ARIMA\,(p-d,\,d,\,q)$

$O.L.S.$ estimates of ψ are **consistent** if:

$$d = 0, \qquad q = 0.$$

Mann & Wald '43.

Objective:

A) Define **consistent**, **iterative** $O.L.S.$ estimates in the general case. -

B) "Identify" the model.

Consider:

$$Z_t = \sum_{l=1}^{p} \Phi_{l(p)}{}^{(0)} Z_{t-l} + e_{p,t}^{(0)}$$

$$t = p+1,\ p+2, \ldots n$$

Let $\hat{\Phi}_{l(p)}{}^{(0)}$ be the $O.L.S.$ estimate and $\hat{e}_{p,t}^{(0)}$ is the $O.L.S.$ residual; i.e.

$$\hat{e}_{p,t}^{(0)} = Z_t - \sum_{l=1}^{p} \hat{\Phi}_{l(p)}{}^{(0)} Z_{t-l}$$

contains information about the $M.A.$ component.

Define:

$$Z_t = \sum_{l=1}^{p} \Phi_{l(p)}{}^{(1)} Z_{t-l} + \beta_{1(p)}{}^{(1)} \hat{e}_{p,t-1}^{(0)} + e_{p,t}^{(1)}$$

$$t = p+2, \ldots\ n$$

$$\hat{e}_{p,t}^{(1)} = Z_t - \sum_{l=1}^{p} \hat{\Phi}_{l(p)}{}^{(1)} Z_{t-l} - \hat{\beta}_{1(p)}{}^{(1)} \hat{e}_{p,t-1}^{(0)}$$

This can be extended to $\hat{e}_{p,t}^{(q)}$ if there is a qth. order

$MA(\bullet)$.

$$\hat{\Phi}_{l(p)}{}^{(q)} \overset{P}{\longrightarrow} \Phi_l, \qquad l = 1,\ldots\ p.$$

Determining the order of the equations uses **consistency** and the fact that a pure $MA(q)$ process has a sample $a.c.f.$ that "truncates" at lag q.

1.2 The Extended SACF (ESACF)

If $p = 0$,

$$Z_t \sim MA(q)$$

$$r_j(0) \doteq 0, \qquad \forall\ j > q, \qquad p = 0$$

If $p = 1$,

$$Z_t \sim ARMA\ (1,q)$$

$W_t = Z_t - \Phi_1 Z_{t-1}$ is (asymptotically) $MA(q)$, so that

$$r_j(1)\ \doteq\ 0, \qquad \forall\ j > q, \qquad p = 1$$

etc.

$$r_j(1) = r_j\left(W_{1,t}{}^{(j)}\right)$$

and

$$r_j(k) = r_j\left(W_{k,t}{}^{(j)}\right)$$

$$W_{k,t}{}^{(j)} = Z_t - \sum \Phi_{l(k)}{}^{(j)} Z_{t-l}$$

If p, q are unknown, the model tends to overfit both elements.

$$r_j(k) \neq 0 \text{ for } 0 \leq j - q \leq k - p$$
$$r_j(k) = 0 \text{ for } j - q > k - p \leq 0$$

1.3 A Comment About Sample Auto-Covariance

Let

$$\widetilde{S}_\tau \equiv \frac{1}{N-|\tau|} \sum_{1}^{N-|\tau|} \left(X_t - \overline{X}\right) \ \left(X_{t+|\tau|} - \overline{X}\right)$$

and

$$\hat{S}_\tau \equiv \frac{1}{N} \sum_{1}^{N-|\tau|} \left(X_t - \overline{X}\right) \ \left(X_{t+|\tau|} - \overline{X}\right)$$

$$= \left(1 - \frac{|\tau|}{N}\right) \hat{S}_\tau$$

Let $\tilde{S}'_\tau$, $\hat{S}'_\tau$ be the same functions but with μ, $E\left\{X_t\right\} = \mu$, replacing $\overline{X}$.

$\hat{S}'_\tau$ is <u>called</u> the "biased estimator".

If $\{X_t\}$ is white noise,

$$S_\tau = cov\left(X_t, X_{t+|\tau|}\right) = 0, \qquad \forall\, \tau \neq 0$$

therefore

$$E\left\{\widetilde{S}_\tau\right\} = \frac{-S_0}{N}, \qquad E\left\{\widehat{S}_\tau\right\} = -\left(1 - \frac{|\tau|}{N}\right)\frac{S_0}{N}$$

$$\sum_{\tau = -(N-1)}^{N-1} \widehat{S}_\tau = 0,$$

so contain a linear constraint not satisfied by S_τ.

For $\tau = 0$,

$$\hat{S}_0 = \tilde{S}_0 = \frac{1}{N} \sum_1^N \left(X_t - \overline{X} \right)^2$$

$$E \left\{ \hat{S}_0 \right\} = S_0 - var \left(\overline{X} \right);$$

i.e. $\hat{S}_0$ underestimates the variance.

1.4 Bootstrapping in Regression

Freedman & Peters, JASA, 1989, 97-106.
c.f. Lele - JRSS, B. 1991, 53, 253-
Brad Efron, The Jacknife, Bootstrap, and Other Resampling Plans. SIAM, 1982

The principle involved is to treat a sample probability distribution function as if it were the theoretical probability distribution, and sample independently from it in order to provide a sequence of observations that can be **averaged**.

Key is the **assumption** of **independence** of the joint distribution, and that the requested sampling procedure should maintain that under independence.

But we still rely on $F_n(x)$, the sampling probability distribution function, not being a "low probability outcome."

Now, if we are dealing with a time series, we **cannot** use the bootstrap **directly**, except in trivial cases.

With time series we can:

1) Use a **model** to **reduce** the problem implicitly to resampling of "uncorrelated" **residuals**.

2) Resample **functions** of the data that are uncorrelated, or for which the correlation procedure can be allowed for.

- **Lele** procedure.

1.5 The Freedman - Peters Solution

$$\underset{q}{Y_t} = Y_{t-1}\underset{q*q}{B} + X_t\underset{p*q}{C} + \varepsilon_t$$

Let

$$\widehat{\varepsilon} = Y_t - Y_{t-1}\widehat{B} - X_t\widehat{C}$$

We assume $\{\varepsilon_t\}$ is $i.i.d.$

$\widehat{\varepsilon}_t$ estimates $\{\varepsilon_t\}$ with probability distribution function

$$\mu = \frac{1}{n} \quad \forall t,$$

i.e. $\{\widehat{\varepsilon}_t\}$ are **equally likely**!

.

The data used are:

$$\{Y_t\}, \{X_t\}, Y_0$$

We sample ε^* from μ and calculate:

$$Y_t^* = Y_{t-1}^*\hat{B} + X_t\hat{C} + \varepsilon_t^*$$

$$Y_0^* = Y_0, \qquad \forall t.$$

We re-estimate $\widehat{B}^*$, $\widehat{C}^*$ using $\{Y_t^*\}$, $\{X_t\}$, Y_0.

A major finding is that conventional (analytic) asymptotics can be off by a factor of 3! - in context of $G.L.S.$

e.q.

$$Y = X_B + \varepsilon, \qquad cov\left(\varepsilon\right) = \sum$$

$$\hat{B}_{glc} = \left(X^{'} \sum^{-1} X\right)^{-1} X^{'} \sum^{-1} Y$$

See the articles mentioned.

(Saved under a:,v: TimeSeries)