Lecture 34

Long Memory ARMA and Fractional Differencing

Arnab Hazra

Paleoclimatic Glacial Varves (recap)

- Melting glaciers deposit yearly layers of sand and silt during the spring melting seasons.
- This can be reconstructed yearly over a period ranging from the time deglaciation began in New England (about 12,600 years ago) to the time it ended (about 6,000 years ago).
- Such sedimentary deposits, called varves, can be used as proxies for paleoclimatic parameters, such as temperature, because, in a warm year, more sand and silt are deposited from the receding glacier.
- Because the variation in thicknesses increases in proportion to the amount deposited, a logarithmic transformation could remove the nonstationarity observable in the variance as a function of time.

Varve versus log-varve data

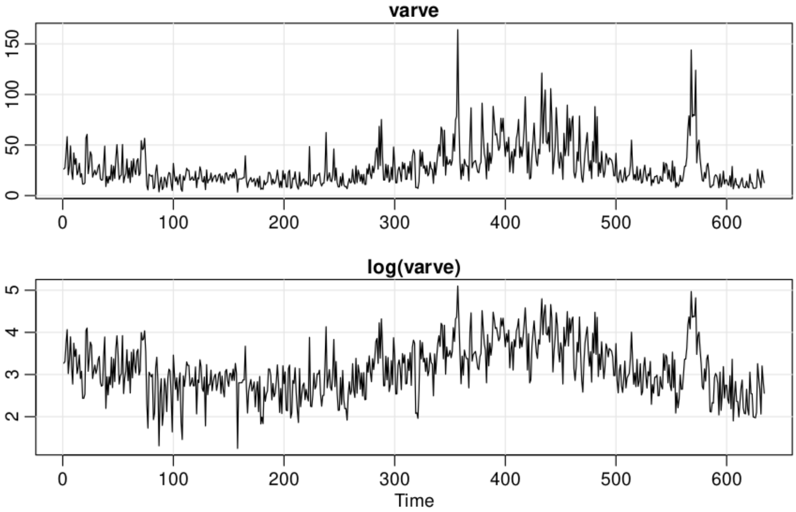
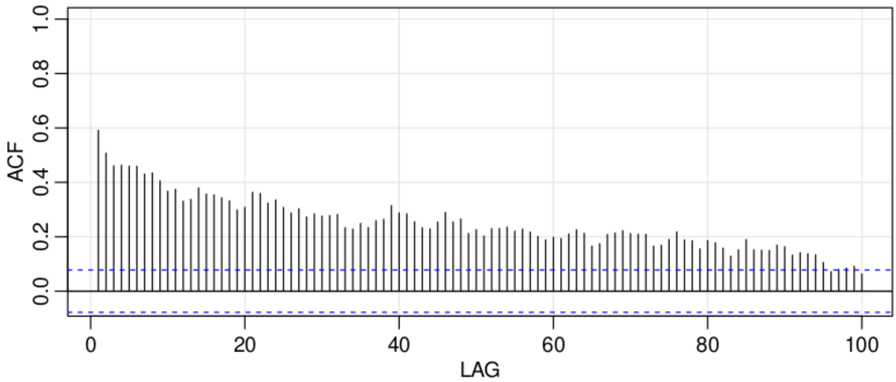


Fig. 2.7. *Glacial varve thicknesses (top) from Massachusetts for* $n = 634$ *years compared with* log *transformed thicknesses (bottom).*

ACF of log verve data



ACF and PACF of $\nabla X_t = \nabla \log(\text{Varve}_t)$

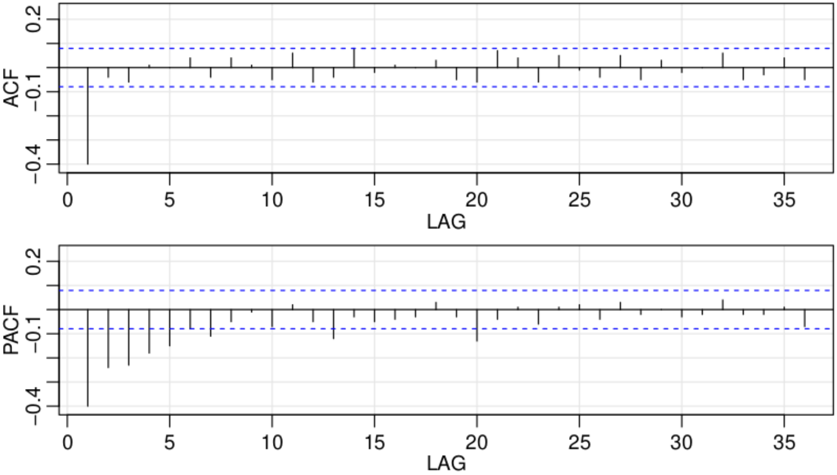


Fig. 3.9. *ACF and PACF of transformed glacial varves.*

An MA(1) process seems reasonable for ∇X_t

ARIMA(0,1,1) versus ARIMA(1,1,1)

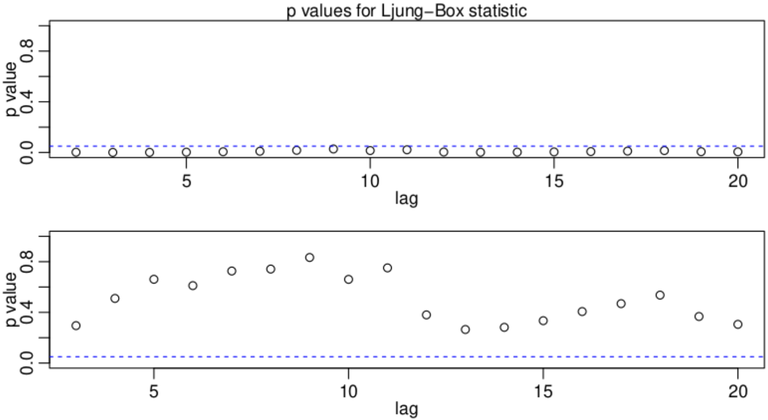


Fig. 3.17. *Q-statistic p-values for the ARIMA(0, 1, 1) fit (top) and the ARIMA(1, 1, 1) fit (bottom) to the logged varve data.*

Definition of ARMA models (recap)

- A autoregressive moving average model of order (p, q), abbreviated $ARMA(p, q)$, is of the form

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}.$$

- Here X_t is stationary, $W_t \sim WN(0, \sigma_W^2)$, and $\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ are constants with $\phi_p, \theta_q \neq 0$.
- The parameters p and q are called the autoregressive and the moving average orders, respectively.
- If X_t has a nonzero mean μ, we set $\alpha = (1 - \phi_1 - \ldots - \phi_p)\mu$ and write the model as

$$X_t = \alpha + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \ldots + \theta_q W_{t-q}.$$

Coefficients of $\psi(z)$ and $\pi(z)$ (recap)

- An ARMA(p, q) model is defined by $\phi(B)X_t = \theta(B)W_t$.
- An ARMA(p, q) model is said to be causal, if $X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j} = \psi(B)W_t$.
- An ARMA(p, q) model is said to be invertible, if $\pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} = W_t$.
- The coefficients ψ_j's can be determined by solving $\psi(z) = \theta(z)/\phi(z)$, where $|z| \leq 1$.
- The coefficients π_j's can be determined by solving $\pi(z) = \phi(z)/\theta(z)$, where $|z| \leq 1$.

Short-memory process

- The conventional ARMA(p, q) process is often referred to as a short-memory process because the coefficients in the representation

$$X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$$

obtained by solving $\phi(z)\psi(z) = \theta(z)$, are dominated by exponential decay.
- This result implies the ACF of the short memory process satisfies $\rho(h) \to 0$ exponentially fast as $h \to \infty$.
- When the sample ACF of a time series decays slowly, we need to difference the series until it seems stationary.
- For the log-varve data, $\nabla X_t = \phi \nabla X_{t-1} + W_t + \theta W_{t-1}$.

Fractional differencing

- The use of the first difference $\nabla X_t = (1 - B)X_t$ can be a too severe modification due to overdifferencing of the original process.
- The easiest way to generate a series is using the difference operator $(1 - B)^d$ for fractional values of d, say, $0 < d < 0.5$; here $(1 - B)^d X_t = W_t$.
- The fractionally differenced series, for $|d| < 0.5$, is often called fractional noise (except when d is zero) and d becomes a parameter along with σ_W^2.
- Such processes occur in hydrology and in environmental series, such as the varve data ($\hat{d} = 0.384$).
- Such series tend to exhibit sample autocorrelations that are not necessarily large, but persist for a long time.

Properties of fractional differencing

- To investigate its properties, we can use the binomial expansion ($d > -1$) to write

$$W_t = (1-B)^d X_t = \sum_{j=0}^{\infty} \pi_j B^j X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$$

where

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}.$$

- Similarly ($d < 1$), we can write

$$X_t = (1-B)^{-d} W_t = \sum_{j=0}^{\infty} \psi_j B^j W_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$$

where

$$\psi_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}.$$

Properties of fractional differencing (contd.)

- Using the previous representation and after some nontrivial manipulations, it can be shown that the ACF of X_t is

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)} \sim h^{2d-1}$$

  with $\Gamma(x+1) = x\Gamma(x)$ being the gamma function.

- From this we see that for $0 < d < 0.5$, $\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty$ and hence the term long memory process.

- The terms $\pi_j(d)$ can be derived using $\pi_0(d) = 1$ with

$$\pi_{j+1}(d) = (j-d)\pi_j(d)/(j+1)$$

  for $j = 0, 1, \ldots$

π-weights for log varve data

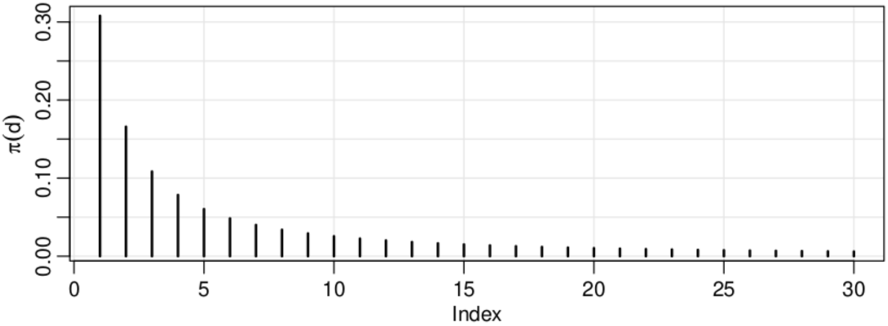


Fig. 5.2. *Coefficients* $\pi_j(.384)$, $j = 1, 2, \ldots, 30$ *in the representation* (5.7).

The fractional differencing is done using the `fracdiff` package.

Unit root test

- Consider a causal AR(1) process $X_t = \phi X_{t-1} + W_t$.
- A unit root test provides a way to test whether X_t is a random walk (the null case) as opposed to a causal process (the alternative).
- That is, it provides a procedure for testing $H_0 : \phi = 1$ versus $H_1 : |\phi| < 1$.
- The test statistic used here is $T(\hat{\phi} - 1)$ and it is called Dickey-Fuller (DF) statistic, where $\hat{\phi}$ is lag-1 sample ACF.
- The asymptotic distribution of the test statistic requires Brownian motion (and hence skipped).

Thank you!