Financial Time Series

Topic 6: Fractional Integration and Long Memory Processes

Hung Chen
Department of Mathematics
National Taiwan University
5/06/2002

OUTLINE

1. Variance Estimate
2. Decomposing time series
3. Hypothesis Testing
4. Spectral Density
5. Fractional Integration and Long Memory Processes
   - ARFIMA models
   - Testing for fractional differencing
   - Estimation
6. Measures of Persistence and Trend Reversion

Variance Estimate in ARMA Models

Consider an $ARMA(p,q)$ process

$$\begin{aligned} &X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} \\ &= a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}, \end{aligned}$$

or

$$\begin{aligned} &(1 - \phi_1 B - \cdots - \phi_p B^p) X_t \\ &= (1 - \theta_1 B - \cdots - \theta_q B^q) a_t, \end{aligned}$$

i.e.

$$\phi(B) X_t = \theta(B) a_t.$$

Here $\{a_t\} \sim IID(0, \sigma^2)$.
Use either MLE or Least squares method, both approaches lead to the following result.

- Set $\hat{\beta} = (\hat{\phi}_1, \cdots, \hat{\phi}_p, \hat{\theta}_1, \cdots, \hat{\theta}_q)^T$. Then

$$\sqrt{T}(\hat{\beta} - \beta) \to N(0, \mathbf{V}(\beta)),$$

where

$$\mathbf{V}(\beta) = \sigma^2 \begin{bmatrix} E\mathbf{U}_t\mathbf{U}_t^T & E\mathbf{U}_t\mathbf{V}_t^T \\ E\mathbf{V}_t\mathbf{U}_t^T & E\mathbf{V}_t\mathbf{V}_t^T \end{bmatrix}^{-1}.$$

- autogressive processes:

$$\mathbf{U}_t = (U_t, \ldots, U_{t+1-p})^t$$

$$
\begin{aligned}
\mathbf{V}_t &= (U_t, \ldots, U_{t+1-p})^t \\
\phi(B)U_t &= a_t \\
\theta(B)V_t &= a_t.
\end{aligned}
$$

- $AR(p)$:

$$
\begin{aligned}
Var(\phi) &= \sigma^2(E\mathbf{U}_t\mathbf{U}_t^T)^{-1}, \\
EU_tU_t^T &= (EX_iX_j)_{i,j=1}^p.
\end{aligned}
$$

- $AR(1)$: $\hat{\phi}$ is $AN(\phi, T^{-1}(1-\phi^2))$.
- $AR(2)$: $(\hat{\phi}_1, \hat{\phi}_2)^T$ is

$$
AN\left(\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, T^{-1}\begin{pmatrix} 1-\phi_2^2 & -\phi_1(1+\phi_2) \\ -\phi_1(1+\phi_2) & 1-\phi_2^2 \end{pmatrix}\right)
$$

- $MA(q)$:

$$
\begin{aligned}
Var(\theta) &= \sigma^2(EV_tV_t^T)^{-1}, \\
EV_tV_t^T &= (EV_iV_j)_{i,j=1}^q.
\end{aligned}
$$

Apply the results from $AR(p)$, we have

- $MA(1)$: $\hat{\phi}$ is $AN(\theta, T^{-1}(1-\theta^2))$.
- $MA(2)$: $(\hat{\theta}_1, \hat{\theta}_2)^T$ is

$$
AN\left(\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}, T^{-1}\begin{pmatrix} 1-\theta_2^2 & -\theta_1(1+\theta_2) \\ -\theta_1(1+\theta_2) & 1-\theta_2^2 \end{pmatrix}\right)
$$

Question: How do we simulate a univariate ARIMA time series?

- Use arima.sim in Splus.
- The innovations are Gaussian by default.
- Command:
arima.sim(model, n=100, innov=NULL, n.start=100, start.innov=NULL, rand.gen=rnorm, xreg=NULL, reg.coef=NULL, ...)
- Example 1: Simulate an $ARMA(1,1)$ with standard deviation of innovations 1.

$$x < -arima.sim(100, model = list(ar = .5, ma = -.6)).$$

Example 2: Simulate an $ARIMA(0,1,1)$ with contaminated innovations.

$$rand.10wild < -function(n) \ \ ifelse(runif(n) > .90, rnorm(n), rcauchy(n))$$
$$x.wild < -arima.sim(100, model = list(ndiff = 1, ma = .6), n.start = 100, rand.gen = rand.1$$

- model: a list specifying an ARIMA model. Note that the coefficients must be provided

through the elements ar and ma (otherwise the coefficients are set to zero).
Example: $model = list(ma = c(-.5, -.25))$

- n: the length of the series to be simulated.
- innov: a univariate time series or vector of innovations to produce the series.
  If not provided, innov will be generated using rand.gen.
- n.start: the number of start-up values discarded when simulating non-stationary models.
  The start-up innovations will be generated by rand.gen if start.innov is not provided.
- start.innov: a univariate time series or vector of innovations to be used as start up values.
  Missing values are not allowed.
- rand.gen: a function which is called to generate the innovations.
  Usually, rand.gen will be a random number generator.
- xreg: a univariate or multivariate time se-

ries, or a vector, or a matrix with univariate time series per column.
These will be used as additive regression variables.

- reg.coef: a vector of regression coefficients corresponding to xreg.

Decomposition of Time Series

Suppose a time series is difference stationary.

- Unobserved component models
- Write it as
$$x_t = z_t + u_t. \tag{1}$$
trend plus noise: how and why
- What is it for?
Idea: The unobserved random walk is buried in white noise.
- Motivated Example:
What is the expected real rates of interest under the assumption of rational expectation (financial market efficiency):
  - Example 3.5, Fig. 3.8
  - z_t: unobservable expected real rate
  - z_t: a driftless random walk (under the above assumption)
  - x_t: observed real rate
  - u_t: unexpected inflation
It is a white-noise process if the market is efficient.

– x_t: Will follow the $ARIMA(0,1,1)$ process.

$$(1-B)x_t = (1-\theta B)e_t. \qquad (2)$$

Refer to Example 3.5.

- Question 1: Given only $\{x_t\}$ and its model, can z_t and u_t be identified?
- Question 2: How do we estimate these two unobserved components?
  Signal Extraction

Muth's (1960) approach:

- The trend component, z_t, is a random walk.

$$z_t = \mu + z_{t-1} + v_t.$$

- The noise component, u_t is white noise and independent of v_t.

$$u_t \sim WN(0, \sigma_u^2), \quad v_t \sim WN(0, \sigma_v^2),$$

$$E(u_t v_{t-i}) = 0 \text{ for all } i.$$

- $\triangle x_t$ is a stationary process

$$\triangle x_t = \mu + v_t + u_t - u_{t-1}. \qquad (3)$$

- ACF of $\triangle x_t$: It cuts off at lag one with coefficient

$$\rho_1 = -\frac{\sigma_u^2}{\sigma_u^2 + 2\sigma_v^2}. \tag{4}$$

  Here $\sigma_u^2 + 2\sigma_v^2$ is the variance of $\triangle x_t$.

- $-0.5 \leq \rho_1 \leq 0$

- $\kappa = \sigma_v^2/\sigma_u^2$: signal-to-noise variance ratio
  $\kappa = 0 = \sigma_v^2$: z_t is a deterministic linear trend.
  $\kappa = \infty$: x_t is a pure random walk.

- $\triangle x_t$: an $MA(1)$ process

$$\triangle x_t = \mu + e_t - \theta e_{t-1}, \tag{5}$$

  where

  - $e_t \sim WN(0, \sigma_e^2)$.
  - $\kappa = (1-\theta)^2/\theta$
  - $\theta = \left\{(\kappa + 2) - (\kappa^2 + 4\kappa)^{1/2}\right\}/2$
  - $\sigma_u^2 = \theta\sigma_e^2$

- Identifiability:
  $\hat{\sigma}_u^2$: lag one autocovariance of $\triangle x_t$
  $\hat{\sigma}_v^2$: based on the variance of $\triangle x_t$ and $\hat{\sigma}_u^2$

- MMSE estimate of z_t:
  Given $\{x_t\}_{-\infty}^{\infty}$, Pierce (1979) proposes
  $$\hat{z}_t = v_Z(B)x_t = \sum_{j=-\infty}^{\infty} v_{zj}x_{t-j}.$$
  Refer to p103 for the definition of the filter $v_Z(B)$ in general.

- Estimate of u_t:
  $$\hat{u}_t = (1 - v_Z(B))x_t$$

- Under Muth model,
  $$\begin{aligned} v_Z(B) &= \frac{\sigma_v^2}{\sigma_e^2}(1-\theta B)^{-1}(1-\theta B^{-1})^{-1} \\ &= \frac{\sigma_v^2}{\sigma_e^2}\frac{1}{1-\theta^2}\sum_{j=-\infty}^{\infty}\theta^{|j|}B^j. \end{aligned}$$

- Note that $\sigma_v^2 = (1-\theta)^2\sigma_e^2$, we have
  $$\hat{z}_t = \frac{(1-\theta)^2}{1-\theta^2}\sum_{j=-\infty}^{\infty}\theta^{|j|}x_{t-j}.$$

- How do we estimate z_t if we only have data on x_t up to $t-m$?
  Pierce (1979) proposed the following:
  For $m \geq 0$,
  $$\hat{z}_t^{(m)} = (1-\theta)B^m\sum_{j=0}^{\infty}(\theta B)^j x_t.$$

For $m < 0$,

$$\hat{z}_t^{(m)} = \frac{1-\theta}{\theta^m} B^m \sum_{j=0}^{\infty} (\theta B)^j x_t + \frac{1}{1-\theta B} \sum_{j=0}^{-m-1} \theta^j B^{-j} x_t.$$

More general form:

- $\triangle z_t = \mu + \nu(B)v_t$ and $u_t = \lambda(B)a_t$
  where v_t and a_t are independent white-noise sequences with finite variances σ_v^2 and σ_a^2 and
  where $\nu(B)$ and $\lambda(B)$ are stationary polynomials having no common roots.

- x_t will have the following form

$$\triangle x_t = \mu + \theta(B)e_t \tag{6}$$

  where $\theta(B)$ and σ_e^2 can be obtained from

$$\sigma_e^2 \frac{\theta(B)\theta(B^{-1})}{(1-B)(1-B^{-1})} \tag{7}$$
$$= \sigma_v^2 \frac{\nu(B)\nu(B^{-1})}{(1-B)(1-B^{-1})} + \sigma_a^2 \lambda(B)\lambda(B^{-1}).$$

  The parameters will not be identified in general.

- Poterba and Summers (1988) model:
  - Assume $u_t = \lambda u_{t-1} + a_t$. Then

$$\triangle x_t = \mu + v_t + (1-\lambda B)^{-1}(1-B)a_t$$

    or

$$\triangle x_t^* = (1-\lambda)\mu + (1-\lambda B)v_t + (1-B)a_t$$

where $x_t^* = (1 - \lambda B)x_t$.

- $\triangle x_t$: $ARMA(1,1)$ process

$$(1 - \lambda B) \triangle x_t = \theta_0 + (1 - \theta_1 B)e_t$$

where $e_t \sim WN(0, \sigma_e^2)$ and $\theta_0 = \mu(1 - \lambda)$.

- $\theta_1 = [2(1 + \lambda\kappa)]^{-1} \left\{2 + \kappa(1 + \lambda)^2 \right.$

$$\left. -(1 - \lambda)\left[(1 + \lambda)^2\kappa^2 + 4\kappa\right]^{1/2}\right\}$$

- $\sigma_e^2 = (\lambda\sigma_v^2 + \sigma_a^2)/\theta_1$

Example 3.5 Estimating expected real rates of interest

- Model in (2) is fitted to the real UK Treasury bill rate over the period 1952Q1 to 1995Q3.
- $\triangle x_t = (1 - 0.694B)e_t$, $\hat{\sigma}_e^2 = 7.62$
- Hence,

$$\begin{aligned} \hat{\sigma}_v^2 &= (1 - 0.694)^2 \hat{\sigma}_e^2 = 0.71 \\ \hat{\sigma}_u^2 &= 0.694 \hat{\sigma}_e^2 = 5.29. \end{aligned}$$

- The variations in the expected real rate are small compared to variations in unexpected inflation. (0.71/5.29 = 0.134)
- Exponentially weighted moving average

$$\hat{z}_t = v_Z^{(0)}(B)x_t = (1-0.694)\sum_{j=0}^{\infty}(0.694B)^j x_t.$$

- When θ is close to zero, $\hat{z}_t$ will be almost equal to the most recently observed value of x.
- Large values of θ correspond to small values of the signal-to-noise ratio.
- Unexpected inflation: $\hat{u}_t = x_t - \hat{z}_t$

- Figure 3.8:
  - The expected real rate is considerably smoother than the observed real rate. (small κ)
  - Early 50s: Expected real rate is generally negative.
  - 1956 to 1970: consistently positive
  - mid70 to mid 80: negative
  - Minimum: $1975Q1$ (peak inflation due to the OPEC price rise)
  - mid80 to present: positive
  - Fluctuations in unexpected inflation are fairly homogeneous except for the period from 1974 to 1982.

Hypothesis Testing: nested hypotheses

Consider statistical tests of $r < \ell$ independent equality restrictions on the $\ell \times 1$ parameter vector θ_0, which is being represented by the implicit side relations

$$g_j(\theta) = 0, \quad j = 1, 2, \ldots, r. \tag{8}$$

This setting is being called a nested hypothesis.

- The vector that satisfy (8) form an $(\ell - r)$-dimensional subspace Θ_0 of the parameter space Θ.
  θ_0 lies in a subspace.
- $H_0 : \theta_0 \in \Theta_0$ versus $H_a : \theta_0 \in \Theta - \Theta_0$.
- We can differentiate functions of θ at $\theta_0 \in \Theta_0$ in all directions, including those leading to a passage into the alternative parameter space $\Theta - \Theta_0$.

Likelihood Ratio test

- unconstrained maximizer: $\hat{\theta}$

$$\max_{\theta \in \Theta} L(\theta)$$

- constrained maximizer: $\tilde{\theta}$

$$\max_{\theta \in \Theta_0} L(\theta)$$

re-parametrization or applying Lagrange Multiplier method

$$\log L(\theta) - \sum_{i=1}^{r} \mu_j g_j(\theta)$$

- Form the likelihood ratio

$$\lambda = L(\tilde{\theta})/L(\hat{\theta}).$$

Under H_0, $LR = -2 \log \lambda$ is asymptotically distributed as chi-square with r degrees of freedom.

- Taylor series expansion:

$$\begin{aligned} \log L(\tilde{\theta}) \; - \; \log L(\hat{\theta}) &\approx q(\hat{\theta})^T (\tilde{\theta} - \hat{\theta}) \\ &+ \frac{1}{2} (\tilde{\theta} - \hat{\theta})^T Q(\hat{\theta}) (\tilde{\theta} - \hat{\theta}). \end{aligned}$$

- LR will serve as a test statistic for H_0.

- LR can be written as

$$\sqrt{n}(\tilde{\theta} - \hat{\theta})^T \bar{H} (\tilde{\theta} - \hat{\theta})\sqrt{n}$$

where $\bar{H}$ is the Hessian matrix.

Wald's test

- Idea: $g_j(\hat{\theta})$ should be close to $g_j(\theta_0)$ which is zero.
- Wald's test statistic:

$$\begin{aligned} W &= (g_1(\hat{\theta}), \ldots, g_r(\hat{\theta}))^T \\ &\quad (G_r(\hat{\theta})\hat{V}G_r(\hat{\theta})^T)^{-1}(g_1(\hat{\theta}), \ldots, g_r(\hat{\theta})) \end{aligned}$$

where $G_r(\theta)$: the $r \times \ell$ matrix from the derivative of $(g_1(\hat{\theta}), \ldots, g_r(\hat{\theta}))$ and
$\hat{V}$: the covariance matrix estimate of $\hat{\theta}$.

- Under H_0, W is asymptotically distributed as chi-square with r degrees of freedom.

Lagrange multiplier test

- It is also called Rao efficient score test.
- score vector:

$$q(\theta) = \partial \log L(\theta)/\partial\theta$$

- Idea: $q(\tilde{\theta})$ should be close to $q(\theta_0)$ which is zero.
- Lagrange multiplier test statistic:

$$LM = q(\tilde{\theta})^T \hat{V}(\tilde{\theta}) q(\tilde{\theta}).$$

- Under H_0, LM is asymptotically distributed as chi-square with r degrees of freedom.

Spectral Density

- Time-Domain property: The autocorrelations and the variance summarize the second order moments of a stationary process.
- Frequency-Domain properties:
- Consider $x_1, \ldots, x_T$ made at times $1, \ldots, T$ respectively.
- Express x_t as

$$T^{-1/2} \sum_{-\pi < w_j \leq \pi} a_j \exp(itw_j)$$

  where
  $w_j = 2\pi j/T$: Fourier frequencies
  a_j: random Fourier coefficients.

- spectral density: $f_x(w) = E|a_j|^2$
- Periodogram: $I(w_j)$

$$I(w_j) = T^{-1} \left| \sum_{t=1}^{T} x_t \exp(-itw_j) \right|^2 .$$

  Note that

$$\|\mathbf{x}\|^2 = \sum_j I(w_j).$$

- High values of $f(w)$: possible cyclical behavior at frequency w with the period of one cycle equalling $2\pi/w$ time units.
- The series x_t will display long memory if its spectral density, $f_x(w)$, increases without limit as the frequency w tends to zero.
- If x_t is ARFIMA, then $f_x(w)$ behaves like w^{-2d} as $w \to 0$.
  d: It parametrizes the low-frequency behavior.

Fractional Integration and Long Memory

- In the analysis of financial time series, we usually consider the order of differencing, d, is either 0 or 1.
  - $x_t \sim I(1)$: The ACF declines linearly.
  - $x_t \sim I(0)$: The ACF declines exponentially.

    Observations separated by a long time span may be assumed to be independent.
- Long persistence: Many empirically observed time series appeared to satisfy the assumption of stationarity (perhaps after some differencing transformation) but it exhibits a dependence between distant observations.
- Hurst effect (Mandlebrot and Wallis, 1969): hydrology
- Many economic time series exhibit the tendency for large values to be followed by large values of the same sign.

  The series seem to go through a succession of cycles even including long cycles whose length is comparable to the total sample size.

- Call for new models.

fractionally integrated

- Model long-term persistence.
- ARFIMA (AR Fractionally IMA)
- Consider real $d > -1$,

$$\triangle^d = (1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k \quad (9)$$
$$= 1 - dB + \frac{d(d-1)}{2!}B^2 - \frac{d(d-1)(d-2)}{3!}B^3 + \cdots$$

- How does the ARFIMA model incorporate long memory behaviour?
- Fractional white noise ($ARFIMA(0, d, 0)$ process)

$$(1-B)^d x_t = a_t.$$

- random walk versus Brownian motion
  fractional white noise versus fractional Brownian motion
- For non-integer values of d, ACF of x_t declines **hyperbolically** to zero.

The autocorrelations are given by $\rho_k = \Gamma k^{2d-1}$ where Γ is the ratio of two gamma functions.

- weakly (2nd order) stationary: $d < 0.5$
- non-stationary: $d \geq 0.5$
  $Var(x_t) = \infty$.
- Invertible: $d > -0.5$.
  The process can be written in AR form if the π weights converge, i.e. $\Sigma_{j=0}^{\infty} |\pi_j| < \infty$.

Test for Fractional Difference

Classical approach to detect the presence of long-term memory: Hurst (1951), Mandelbrot (1972)

- R/S statistic: range over standard deviation

$$R_0 = \hat{\sigma}_0^{-1} \left[\max_{1 \le i \le T} \sum_{t=1}^{i} (x_t - \bar{x}) - \min_{1 \le i \le T} \sum_{t=1}^{i} (x_t - \bar{x}) \right] \tag{10}$$

  where $\hat{\sigma}_0^2 = T^{-1} \Sigma_{t=1}^{T} (x_t - \bar{x})^2$.

- the range:
  the maximum of the partial sums of the first i deviations of x_i from the sample mean
  the minimum of the partial sums of the first i deviations of x_i from the sample mean

- Shortcoming: R/S is also sensitive to short-range dependence (short-term autocorrelation)

- Modified R/S statistic proposed in Lo (1991):

$$R_q = \hat{\sigma}_q^{-1} \left[\max_{1 \le i \le T} \sum_{t=1}^{i} (x_t - \bar{x}) - \min_{1 \le i \le T} \sum_{t=1}^{i} (x_t - \bar{x}) \right] \tag{11}$$

where

$$\hat{\sigma}_q^2 = \hat{\sigma}_0^2 \left(1 + \frac{2}{T} \sum_{j=1}^{q} w_{qj} r_j\right)$$

and $w_{qj} = 1 - j(q+1)^{-1}$ for $q < T$.
Here r_j is the sample autocorrelations.

- The asymptotic distribution of $T^{-1/2} R_q$ can be found in Lo (1991).
- This test is consistent against a class of long-range dependent alternatives that include all $ARFIMA(p, d, q)$ models with $-0.5 \leq d \leq 0.5$.
- Lo's recommendation: $q = [T^{0.25}]$
No satisfactory answer on the choice of q.

LM test of $d = 0$:

- Use the residuals from fitting an $ARIMA(p, 0, q)$ model to x_t.
- Fitted model:

$$\hat{\phi}(B) x_t = \hat{\theta}(B) \hat{a}_t$$

- LM test of $d = 0$ as the t-ratio on δ in the following regression.

$$\hat{a}_t = \sum_{i=1}^{p} \beta_i W_{t-i} + \sum_{j=1}^{q} \gamma_j Z_{t-j} + \delta K_t(m) + u_t$$

where $\hat{\theta}(B)W_t = x_t$, $\hat{\theta}(B)Z_t = \hat{a}_t$, and $K_t(m) = \Sigma_{j=1}^{m} j^{-1}\hat{a}_{t-j}$.

- Property: consistent, asymptotically normal, r obust to non-normality

- Problem: It is severely affected by autocorrelation in w_t.
  Refer to page 120 for further discussion.

Estimation of d: GPH method

- Geweke and Porter-Hudak (1983): Spectral density of x_t

$$\begin{aligned} f_x(w) &= |1-\exp(-iw)|^{-2d} f_W(w) \\ &= (4\sin^2(w/2))^{-d} f_W(w) \end{aligned}$$

  where $f_W(w)$ is the spectral density of $w_t = (1-B)^d x_t$.

- $\log(f_x(w)) = \log(f_W(w)) - d\log(4\sin^2(w/2))$

- Estimate d as (minus) the slope estimator of the regression of the periodogram $I_T(w_j)$ on a constant and $\log(4\sin^2(w/2))$ at frequencies $w_j = 2\pi j/T$, $j = 1, \ldots, K = [T^{1/2}]$.

Example 3.7 Exchange Rate and Stock Returns

- Dollar/Sterling Exchange: $I(1)$, one unit root
- FTA All Share index: $I(1)$, one unit root
- Daily returns for the S&P 500 index:
  Little evidence that the series is long memory.
  Either squared returns series or absolute returns does appear to be long memory. (Will be discussed later.)
- Goal: Check whether the returns (differences) are really stationary or whether they exhibit long memory.

Dollar/Sterling Exchange

- Use the modified R/S with $q = 9$ to the exchange rate difference.

  $$T^{-1/2}R_9 = 1.692, \quad (0.809, 1.862) : \ 95\% \ CI$$

  We cannot reject the hypothesis that exchange rate returns are short memory.

- LM test: Using the residuals from an $ARIMA(1,1,0)$ model
  t-ratios for δ were 1.03, 1.23, 1.30 and 1.21 for m set equal to 25, 50, 75 and 100 respectively.
- GPH estimate: $d = -0.07 \pm 0.08$ with $K = [T^{1/2}] = 22$

FTA All Share index

- the modified R/S with $q = 4$: $T^{-1/2}R_4 = 2.090$, significant
- LM test is not significant.
- GPH estimate: $d = 0.39 \pm 0.19$ with $K = 19$
- It should be $I(1.4$ instead of $I(1)$.

Measure of Persistence

- Capture short-run dynamics: $ARIMA$
- Suppose that x_t contains a unit root. Then
$$\triangle x_t = \mu + \sum_{j=0}^{\infty} \psi_j a_{t-j}. \qquad (12)$$
- What is the impact of a_t in period $t + k$?
For $\triangle x_{t+k}$, it is ψ_k.
For x_{t+k} is $1 + \psi_1 + \cdots + \psi_k$.
Ultimate impact on the level of x: $A(1) = \Sigma_{j=0}^{\infty} \psi_j$.
- $A(1)$: a measure of how persistent shocks to x are.
    - $A(1) = 0$: trend stationary series
    - $A(1) = 1$: random walk

mean aversion versus mean reversion

Trend Reversion

- Example 3.6 UK stock price
- Try $ARIMA(3, 1, 0)$ to the logarithms of the FTA All Share index in example 2.6.
- $A(1) = 1/0.874 = 1.144$: mean aversion