Analysis of Time Series, L13

Rolf Larsson

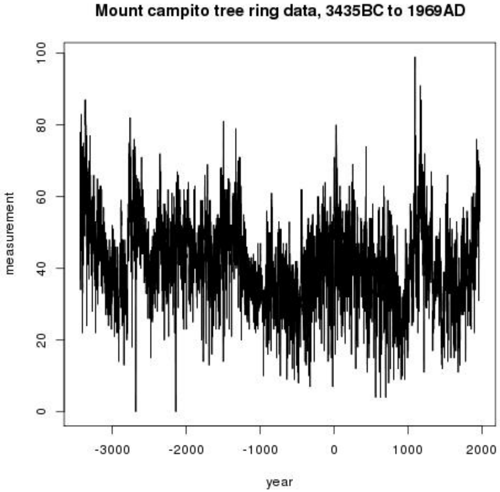
Uppsala University

29 april 2025

Today

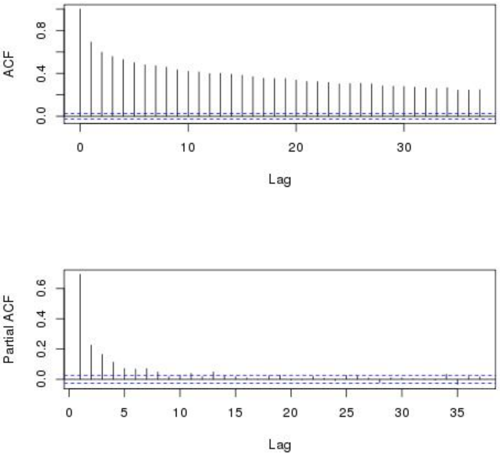
- 5.1: Long memory ARMA and Fractional Differencing
- 5.4: Threshold models
- Menti

Fractional differencing

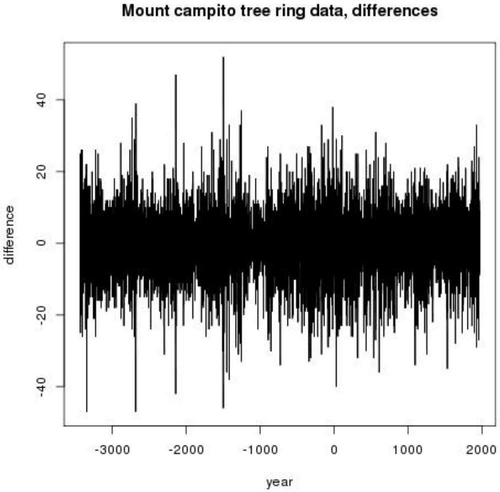


Fractional differencing

Mount Campito, ACF (slowly decreasing) and PACF (not cutting off)

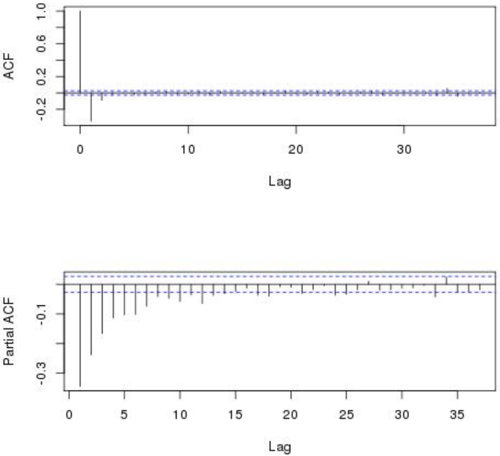


Fractional differencing



Fractional differencing

Mount Campito differences, ACF (cuts off?) and PACF



Fractional differencing

- Consider the model $(1-B)^d x_t = w_t$:
  - Random walk for $d = 1$
  - White noise for $d = 0$
  - What about $0 < d < 1$?
- WLOG: Restrict to $|d| < 1/2$.
- Binomial expansion

$$(1-B)^d = 1 - dB + \binom{d}{2} B^2 - \binom{d}{3} B^3 + \ldots = \sum_{j=0}^{\infty} \pi_j B^j,$$

where (why?)

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}, \quad \Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx.$$

Fractional differencing

- MA representation

$$x_t = (1-B)^{-d} w_t = \sum_{j=0}^{\infty} \psi_j(d) w_{t-j},$$

where

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

- Inserting in $\gamma(h) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j(d)\psi_{j+h}(d)$, it may be proved that

$$\gamma(h) = \sigma_w^2 \frac{\Gamma(h+d)\Gamma(1-2d)}{\Gamma(h+1-d)\Gamma(1-d)\Gamma(d)}.$$

Fractional differencing

- Hence (why?),

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h+1-d)\Gamma(d)}.$$

- Notation: $f(z) \sim g(z)$ means that $\frac{f(z)}{g(z)} \to 1$ as $z \to \infty$.
- Stirling's formula: $\Gamma(z) \sim \sqrt{2\pi} z^{z-1/2} e^{-z}$ as $z \to \infty$ implies (why?)

$$\rho(h) \sim Ch^{2d-1}, \quad \text{as } h \to \infty.$$

- Hence, for $d \geq 0$ (why?),

$$\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty.$$

Estimation

- Recall: $(1-B)^d x_t = w_t$. Assume that w_t is normal white noise.
- The MLE of d is found by maximum likelihood, approximately obtained by minimizing

$$Q(d) = \sum_t w_t^2.$$

- Use the R package fracdiff!

Prediction

- AR representation

$$w_t = (1-B)^d x_t = \sum_{j=0}^{\infty} \pi_j(d) x_{t-j},$$

where

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

- Truncated forecast

$$\tilde{x}_{n+m}^n = -\sum_{j=1}^{m-1} \pi_j(\hat{d}) \tilde{x}_{n+m-j}^n - \sum_{j=m}^{\infty} \pi_j(\hat{d}) x_{n+m-j}.$$

Prediction

- MA representation

$$x_t = (1-B)^{-d} w_t = \sum_{j=0}^{\infty} \psi_j(d) w_{t-j},$$

where

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

- MSPE

$$P_{n+m}^n = \hat{\sigma}_w^2 \sum_{j=0}^{m-1} \psi_j^2(\hat{d}).$$

Spectral estimation

- Spectral density based on the AR representation

$$f(\omega) = \frac{\sigma_w^2}{|\pi(e^{-2\pi i\omega})|^2} = \frac{\sigma_w^2}{|\sum_{k=0}^{\infty} \pi_k e^{-2\pi i k\omega}|^2}.$$

- It follows that (why?)

$$f(\omega) = \sigma_w^2 \{4\sin^2(\pi\omega)\}^{-d},$$

i.e. for ω small, $f(\omega) \approx \sigma_w^2 \{4(\pi\omega)^2\}^{-d} = C\omega^{-2d}$.

- Parametric spectral estimate

$$\hat{f}(\omega) = \hat{\sigma}_w^2 \{4\sin^2(\pi\omega)\}^{-\hat{d}},$$

i.e. for ω small, $\hat{f}(\omega) \approx C\omega^{-2\hat{d}}$.

Spectral estimation

- $\hat{f}(\omega) \approx C\omega^{-2\hat{d}}$ for ω small.
- Hence,
$$\log\{\hat{f}(\omega)\} \approx \log C - 2\hat{d}\log\omega.$$
- This suggests an alternative estimation method:
Estimate d via a linear regression of $\log\{\hat{f}(\omega)\}$ on $\log\omega$ for ω "small".

Fitting ARMA(2,2) to differenced Mount Campito data:

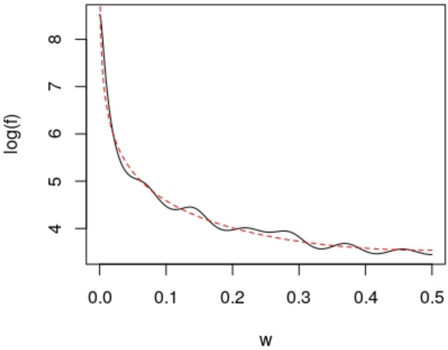
```
> a1=arima(dy,order=c(2,0,2),include.mean=FALSE);a1

Call:
arima(x = dy, order = c(2, 0, 2), include.mean = FALSE)

Coefficients:
         ar1      ar2      ma1     ma2
      1.1512  -0.2216  -1.7007  0.7059
s.e.  0.0367   0.0260   0.0322  0.0313

sigma^2 estimated as 63.93:  log likelihood = -18902.95,
aic = 37815.9
```

Fitting $(1-B)^d x_t = w_t$ to demeaned Mount Campito data (in `x`):

```
> xm=x-mean(x)
> library(fracdiff)
> m=fracdiff(xm)
> m$d
[1] 0.4472056
> m$sigma
[1] 7.991763
```

Tree rings

Estimated log spectral densities:
Parametric AR estimate (AR(14)) in black
Parametric fractional estimate dashed in red



R code for the plot:

```
> s=spec.ar(x,plot=FALSE)
> f=m$sigma^2*(4*sin(pi*s$freq)^2)^(-m$d)
> plot(s$freq,log(s$spec),type='l',xlab='w',ylab='log(f)')
> lines(s$freq,log(f),lty=2,col='red')
```

General ARFIMA

The general ARFIMA(p, d, q) model is defined as

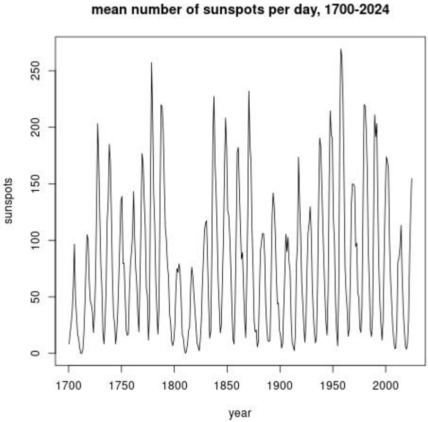
$$\phi(B)\nabla^d(x_t - \mu) = \theta(B)w_t$$

where

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p, \quad \theta(B) = 1 + \theta_1 B + ... + \theta_q B^q.$$

Threshold models

Can the behavior of a series be different above and below a certain threshold?



Threshold models

Threshold AutoRegressive (TAR) model.

- General: If the regions R_1,...,R_r are mutually exclusive and exhaustive, we may define

$$x_t = \alpha^{(j)} + \phi_1^{(j)} x_{t-1} + ... + \phi_p^{(j)} x_{t-p} + w_t^{(j)},$$

  where $(x_{t-1}, ..., x_{t-p}) \in R_j$, $j = 1, 2, ..., r$.

- Special case with $r = 2$, $p = 2$:

$$x_t = \alpha^{(j)} + \phi_1^{(j)} x_{t-1} + \phi_2^{(j)} x_{t-2} + w_t^{(j)},$$

  where $R_1 = \{x_{t-1} < c\}$, $R_2 = \{x_{t-1} \geq c\}$, $j = 1, 2$.

Threshold models

- Special case with $r = 2$, $p = 2$:

$$x_t = \alpha^{(j)} + \phi_1^{(j)} x_{t-1} + \phi_2^{(j)} x_{t-2} + w_t^{(j)},$$

where $R_1 = \{x_{t-1} < c\}$, $R_2 = \{x_{t-1} \geq c\}$, $j = 1, 2$.

- Let $\delta = 1$ if $x_{t-1} \geq c$ and 0 otherwise.
- Equivalent to regression model:

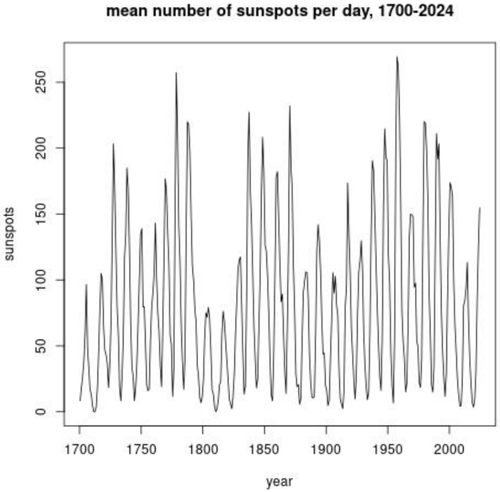
$$\begin{aligned} x_t =& \alpha^{(1)}(1-\delta) + \phi_1^{(1)}(1-\delta)x_{t-1} + \phi_2^{(1)}(1-\delta)x_{t-2} \\ &+ \alpha^{(2)}\delta + \phi_1^{(2)}\delta x_{t-1} + \phi_2^{(2)}\delta x_{t-2} + w_t, \end{aligned}$$

where

$$w_t = (1-\delta)w_t^{(1)} + \delta w_t^{(2)} \sim N(0, \sigma^2)$$

with $\sigma^2 = (1-\delta)^2\sigma_1^2 + \delta^2\sigma_2^2, \quad \sigma_j^2 = \mathrm{var}(w_t^{(j)}), \, j = 1, 2.$

Threshold models



Fit the TAR model

$$x_t = \alpha^{(j)} + \phi_1^{(j)} x_{t-1} + \phi_2^{(j)} x_{t-2} + w_t^{(j)},$$

where $R_1 = \{x_{t-1} < 75\}$, $R_2 = \{x_{t-1} \geq 75\}$, $j = 1, 2$.

In R:

```
> length(x)
[1] 325
> x0=x[seq(3,325)]
> x1=x[seq(2,324)]
> x2=x[seq(1,323)]
> d=(sign(x1-74.99)+1)/2
> x11=(1-d)*x1;x12=d*x1;x21=(1-d)*x2;x22=d*x2;d2=1-d;
> m=lm(x0~0+d2+x11+x21+d+x12+x22);summary(m)
```

```
Call:
lm(formula = x0 ~ 0 + d2 + x11 + x21 + d + x12 + x22)

Residuals:
Min       1Q  Median      3Q     Max
-69.277 -15.507  -3.303  13.873  86.909

Coefficients:
    Estimate Std. Error t value Pr(>|t|)
d2  23.25044    3.46327   6.713 8.78e-11 ***
x11  1.79571    0.11271  15.932  < 2e-16 ***
x21 -1.02830    0.07749 -13.271  < 2e-16 ***
d   34.28875    6.42086   5.340 1.77e-07 ***
x12  1.23227    0.05477  22.499  < 2e-16 ***
x22 -0.59786    0.04488 -13.322  < 2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 24.73 on 317 degrees of freedom
Multiple R-squared:  0.9405,Adjusted R-squared:  0.9394
F-statistic: 835.6 on 6 and 317 DF,  p-value: < 2.2e-16
```

Threshold models

Smooth Transition AutoRegressive (STAR) model:

- $$\begin{aligned} x_t &= \alpha^{(1)} + \phi_1^{(1)} x_{t-1} + \ldots + \phi_p^{(1)} x_{t-p} \\ &\quad + (\alpha^{(2)} + \phi_1^{(2)} x_{t-1} + \ldots + \phi_p^{(2)} x_{t-p}) f(x_{t-1}) + w_t, \end{aligned}$$

  where

  $$f(x) = \frac{1}{1 + e^{(c-x)/\eta}}.$$

- Approaches a TAR model as $\eta \searrow 0$.
- Explanation: for η small, $f(x) \approx 0$ if $x < c$, $f(x) \approx 1$ if $x > c$.

```
> m=nls(x0~a1+f11*x1+f21*x2+(a2+f12*x1+f22*x2)*1/(1+exp((75-x1)/eta)),
+       start=list(a1=23.37,f11=1.79,f21=-1.026,a2=11.06,f12=-0.557,f22=0.4
> summary(m)

Formula: x0 ~ a1 + f11 * x1 + f21 * x2 + (a2 + f12 * x1 + f22 * x2) *
1/(1 + exp((75 - x1)/eta))

Parameters:
    Estimate Std. Error t value Pr(>|t|)
a1  23.37309    3.46832   6.739 7.56e-11 ***
f11  1.78672    0.11257  15.872  < 2e-16 ***
f21 -1.02582    0.07761 -13.218  < 2e-16 ***
a2  11.29601    7.37206   1.532    0.126
f12 -0.55610    0.12523  -4.441 1.24e-05 ***
f22  0.42731    0.08979   4.759 2.96e-06 ***
eta  0.10543    0.90941   0.116    0.908
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 24.78 on 316 degrees of freedom
Number of iterations to convergence: 5
Achieved convergence tolerance: 3.163e-06
```

News of today

- Long memory models:
  - Fractional difference
  - Estimating the d parameter
  - Prediction via AR representation
  - MSPE via MA representation
  - Spectral estimation via AR representation
- Threshold models:
  - TAR
  - STAR