TEXAS A&M UNIVERSITY

Department of Statistics

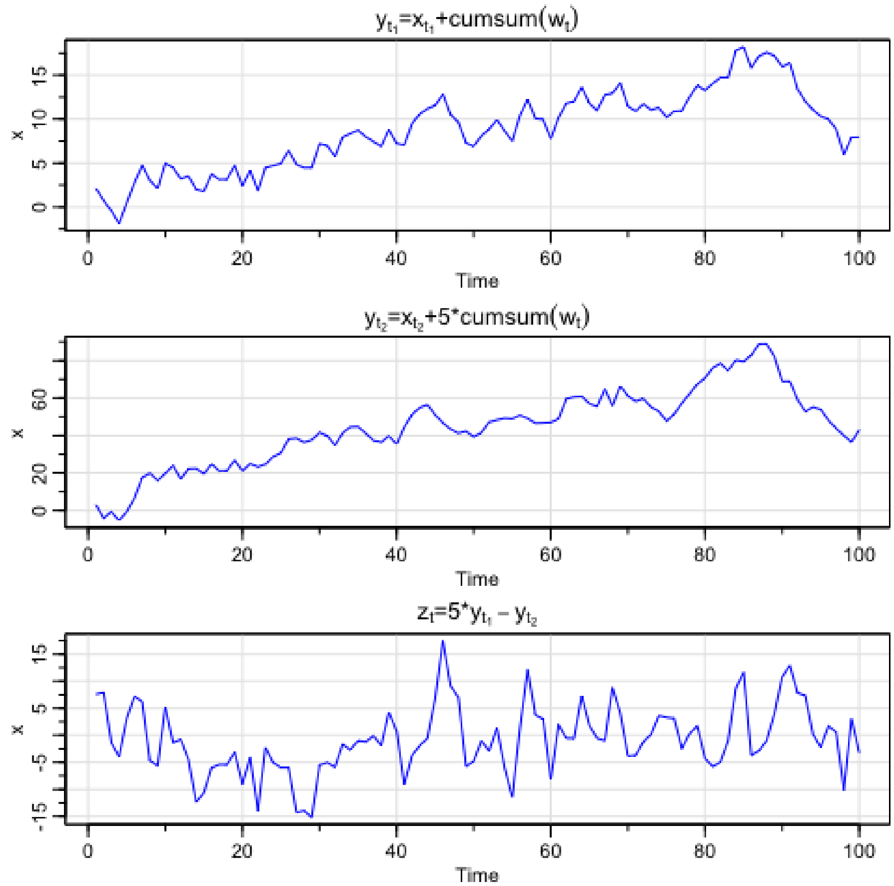
STAT626 - Methods in Time Series Analysis
Homework #5

Salih Kilicli

July 3, 2019

Problem 1.1 (a): **Simulate and plot $n = 100$ values of the three time series $\{y_{t_1}\}, \{y_{t_2}\}$ and $z_t = 5y_{t_1} - y_{t_2}$. Do they appear to be stationary?**

Solution: Clearly, the plots of $\{y_{t_1}\}, \{y_{t_2}\}$ have a changing variance in time, therefore they are not stationary. Whereas, z_t oscillates around 0 and the variance is more stable. It is most likely a stationary process.



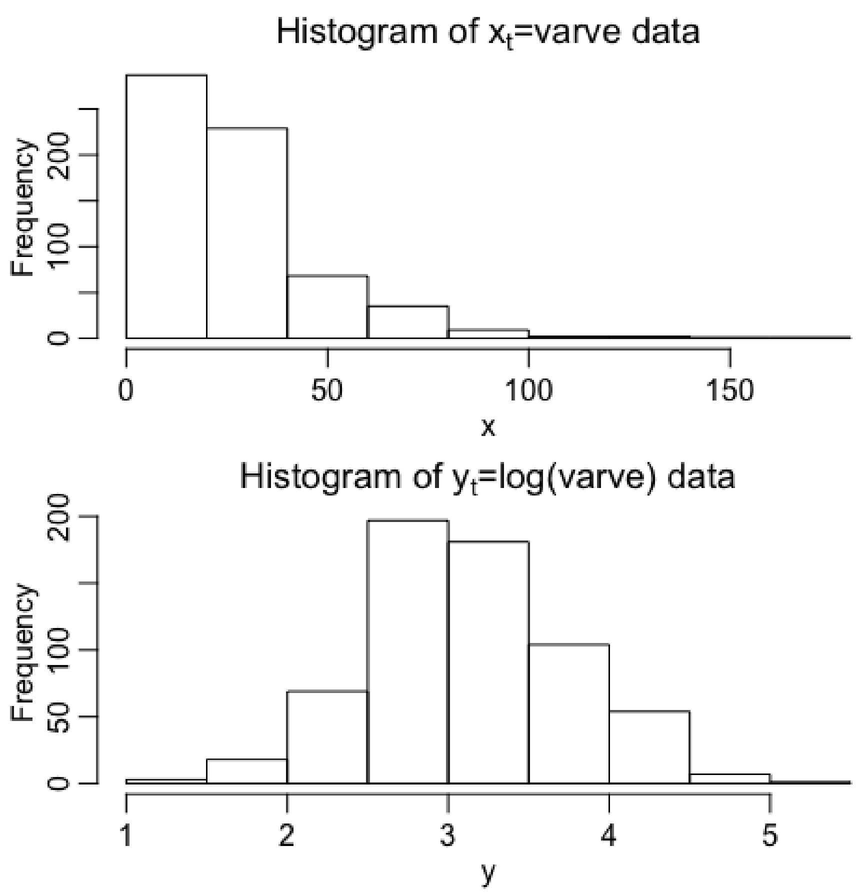
Problem 3.6: **The glacial varve record plotted in Figure 3.9 exhibits some nonstationarity that can be improved by transforming to logarithms and some additional nonstationarity that can be corrected by differencing the logarithms.**

(a) **Argue that the glacial varves series, say x_t, exhibits heteroscedasticty by computing the sample variance over the first half and second half of the data. Argue that transformation $y_t = log(x_t)$ stabilizes the variance over the series. Plot the histograms of x_t and y_t to see whether the approximation to normality is improved by transforming the data.**

(b) **Plot the series y_t. Do any time intervals, of the order 100 years, exist where one can observe behavior comparable to that observed in the global temperature records in Figure 1.2?**

(c) **Examine sample ACF of y_t and comment.**

(d) **Compute the difference $u_t = y_t - y_{t-1}$, examine its plot and sample ACF, and argue that differencing the logged varve data produces a reasonably stationary series. Can you think of a practical interpretation for u_t?**



Solution: (a) Sample variances of the first and second half of the data, respectively,

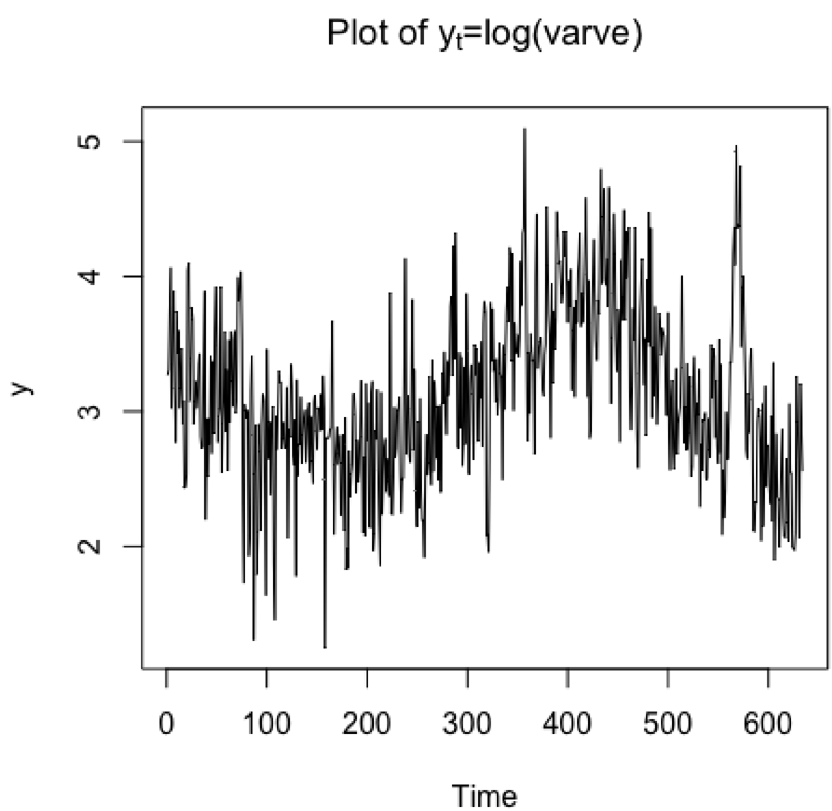
$$Svar(x_t^1) = 133.457, \quad Svar(x_t^2) = 594.490.$$

There is an excessive increase in the sample variance of data from first to second half which implies there exists a heteroscedastiscty in the varve data. Moreover, sample variances of the first and second half of the logged data are;

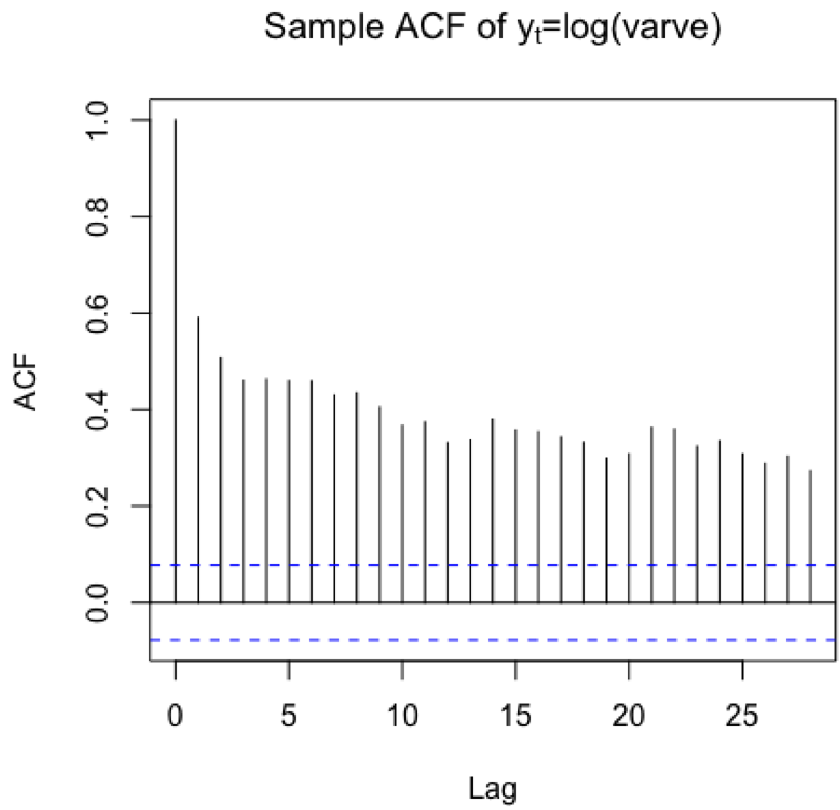
$$Svar(log(x_t^1)) = 0.27, \quad Svar(log(x_t^2)) = 0.45$$

The increase in the sample variances between two halves of the data decreased substantially when the data is logged. Therefore, log transformation stabilizes the sample variance over the series. The plots of histograms of x_t and $y_t = log(x_t)$ are given above. The second plot resembles the shape of a normal distribution and shows that the log transformation also improves the normality of the data.

(b) The increasing trend between the years $200 - 400$ is comparable to the increasing trend of the last 40 years of Global temperature data. Moreover, in general, both of the plots have nonlinear trend with sharp upwards and downwards.

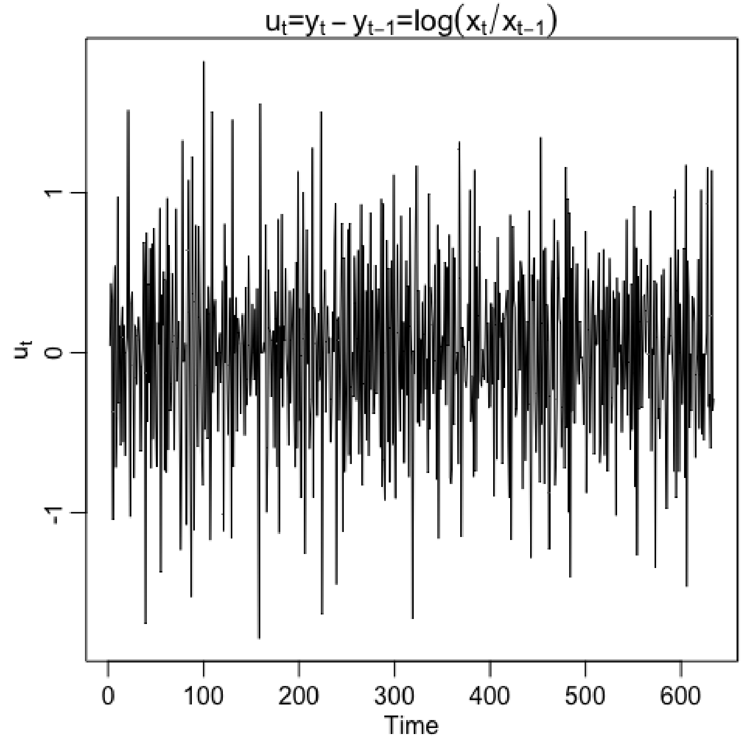


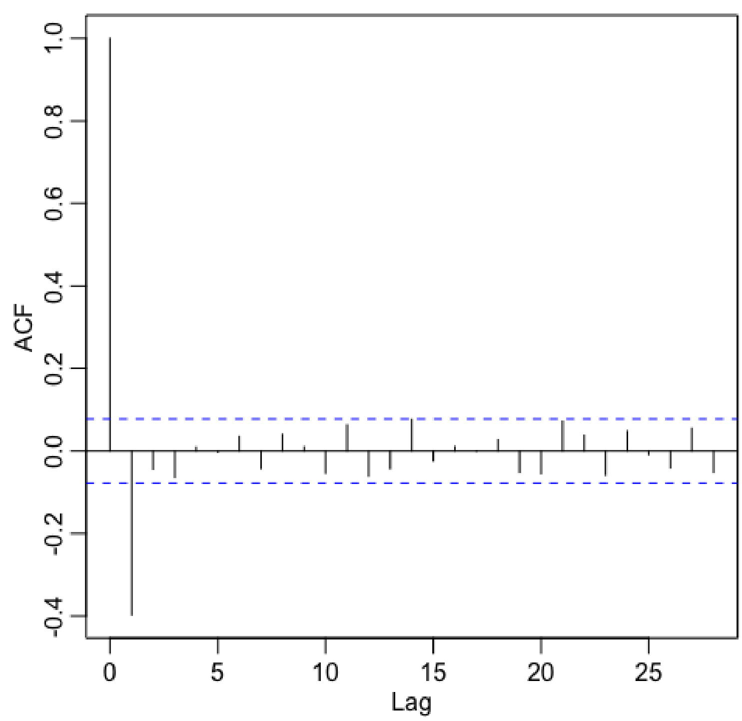
(c) Sample ACF of y_t is given below. Observing the sample ACF, we can see that as the lag increases ACF decreases which means the correlation between terms with long distance time periods are less then closer ones. Moreover, since it is dependent on time, it implies time series is not stationary.



(d) The plot of itself and plot of sample ACF of difference series $u_t = y_t - y_{t-1}$ are given below. The plot of u_t appears to be very close to plot of a stationary series since variance is finite and the series randomly oscillate around 0 without time dependency.

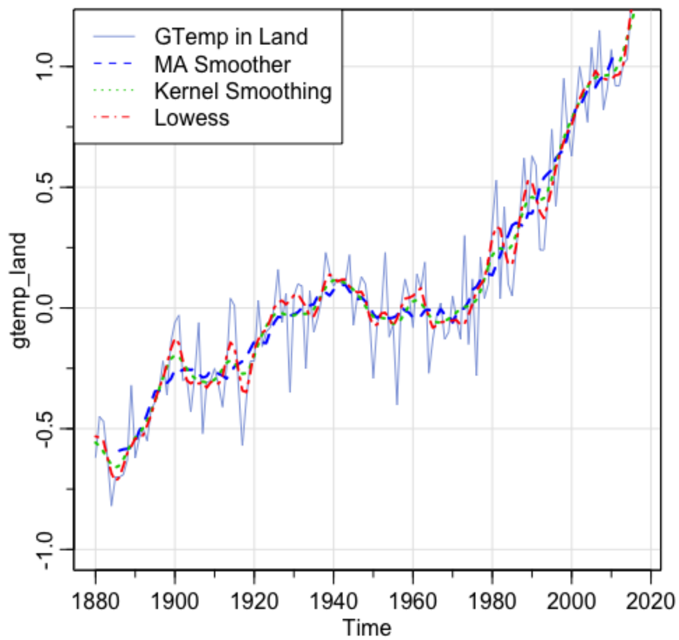
Moreover, comparing the sample ACF of u_t to sample ACF of y_t, we can see that distribution of ACF is more random and doesn't look like it depends on time. Both of the plots imply that differencing the logged data produces a reasonably stationary series. u_t s pretty similar to Eq (3.23) that gives a stationary process differencing two non-stationary series.



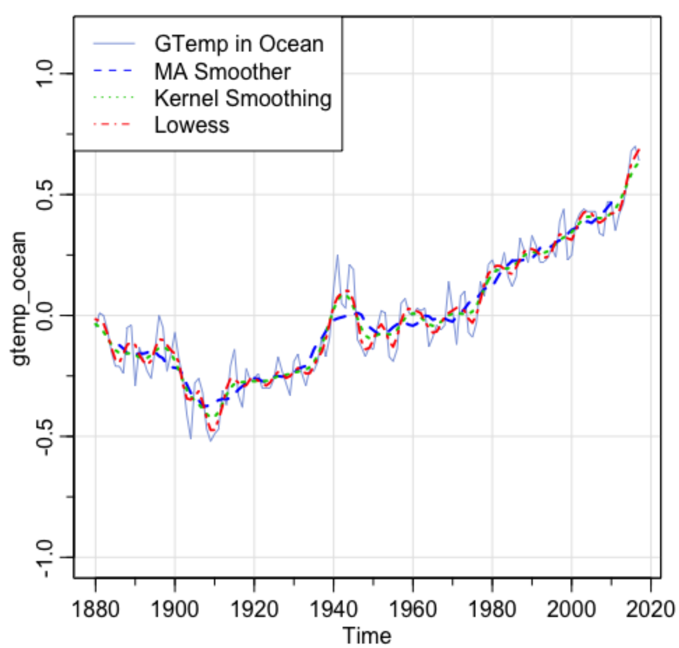


Problem 3.7: **Use the three different smoothing techniques described in Example 3.16, Example 3.17, and Example 3.18, to estimate the trend in the global temperature series displayed in Figure 1.2. Comment.**

Solution: The plot of three different smoothing techniques with global temperature in land surface data is given below. Both of the plots given below show most of the information was captured by all 3 method, whereas lowess look better on capturing non-linear patterns with sharp upward and downwards. Here kernel smoothing bandwidth is $b = 6$.

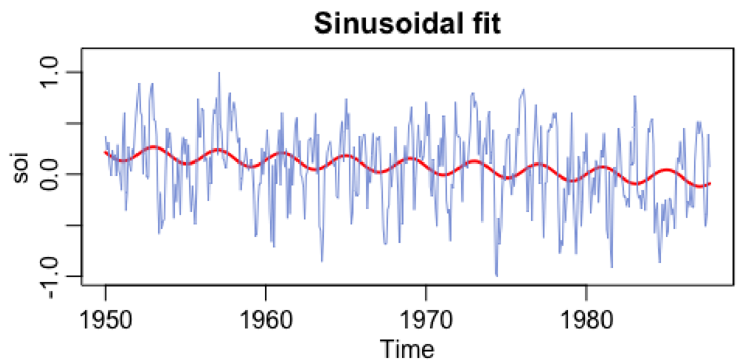


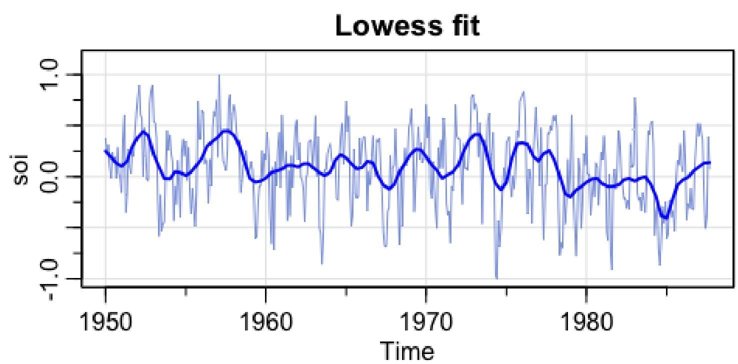
Similarly, the plot of three different smoothing techniques with global temperature in ocean surface data is given below.



Problem 3.8: **In Section 3.3, we saw that the El Nino/La Nina cycle was approximately 4 years. To investigate whether there is a strong 4-year cycle, compare a sinusoidal (one cycle for every four years) fit to the SOI to a lowess fit (as in Example 3.18). In the sinusoidal fit, include a term for the trend. Discuss the results.**

Solution: The plot of sinusoidal fit on the SOI data is given below. The upward and downward cycles in the sinusoidal fit actually parallel with the data, which supports the strong 4-year cycle.





The plot of lowess fit on the SOI data is given above.

HOMEWORK 5 - STAT 626

① Salih KILICLI

Problem I: $\{w_{t1}\}, \{w_{t2}\}, \{w_t\}$ independent $WN(0,1)$ series, and define:

$$y_{t1} = x_{t1} + \sum_{j=1}^{t} w_j, \quad y_{t2} = x_{t2} + 5\sum_{j=1}^{t} w_j$$

where $x_{t1} = 0.5\, x_{t-1,1} + w_{t1}$, $x_{t2} = 0.9\, x_{t-1,2} + w_{t2}$.

(b) Compute the autocovariance function of $\{y_{t1}\}$. Is it stationary?

$$\gamma_{y_{t1}}(t+h,t) = \gamma(y_{t1}(t+h), y_{t1}(t)) = \operatorname{cov}\Big(x_{t1}(t+h) + \sum_{j=1}^{t+h} w_j,\ x_{t1}(t) + \sum_{j=1}^{t} w_j\Big)$$

$$= \operatorname{cov}(x_{t1}(t+h), x_{t1}(t)) + \min\{t+h,t\} \operatorname{cov}(w_j, w_j)$$

NOTE: $\gamma_{AR(1)}(h) = \dfrac{\sigma_w^2 \phi^h}{1-\phi^2}$

$$= \frac{\overset{1}{\cancel{\sigma_w^2}} \cdot (0.5)^h}{1-(0.5)^2} + \overset{1}{\cancel{\sigma_w^2}} \cdot \min\{t+h,t\} = \frac{(0.5)^h}{0.75} + \min\{t+h,t\}$$

Since $\gamma_{y_{t1}}(t+h,t)$ is dependent of time, $\{y_{t1}\}$ is not stationary.

(c) Compute the autocovariance function of $\{z_t\}$ where $z_t = 5y_{t1} - y_{t2}$. Is it stationary?

$$\gamma_z(t+h,t) = \gamma(5y_{t1}(t+h) - y_{t2}(t+h),\ 5y_{t1}(t) - y_{t2}(t))$$

$$= \operatorname{cov}(5x_{t1}(t+h) - x_{t2}(t+h),\ 5x_{t1}(t) - x_{t2}(t))$$

$$= 25\operatorname{cov}(x_{t1}(t+h), x_{t1}(t)) + \operatorname{cov}(x_{t2}(t+h), x_{t2}(t))$$

$$= 25\,\gamma_{x_{t1}}(h) + \gamma_{x_{t2}}(h) = 25\cdot\left(\frac{(0.5)^h}{0.75}\right) + \left(\frac{(0.9)^h}{0.19}\right)$$

Since $\gamma_z(h)$ is independent of time and a function of lag (h) only, z_t is stationary.

(d) Compare and explain your findings in (a)-(c) regarding (non)stationarity of the three time series involved.

$$\left.\begin{aligned} \gamma_{y_{t1}}(t+h,t) &= \frac{(0.5)^h}{0.75} + \min\{t+h,t\} \\ \gamma_{y_{t2}}(t+h,t) &= \frac{(0.9)^h}{0.19} + (\min\{t+h,t\})\,25 \end{aligned}\right\} \text{both non-stationary}$$

However, for $z_t = 5y_{t1} - y_{t2}$, $\gamma_z(t+h,t) = \dfrac{25\,(0.5)^h}{0.75} + \dfrac{(0.9)^h}{0.19} = \gamma_z(h)$ stationary.

(e) Compute $\gamma_{y_{t1},y_{t2}}$ (CCF), are the two time series $\{y_{t1}\},\{y_{t2}\}$ jointly stationary?

$$\left.\begin{aligned} M_1 = E[y_{t1}] = E[x_{t1} + \textstyle\sum w_j] = 0 \quad &\text{since } E[x_{t1}] = 0.5\,E[x_{t1}] \Rightarrow E[x_{t1}] = 0 \\ M_2 = E[y_{t2}] = E[x_{t2} + 5\textstyle\sum w_j] = 0 \quad &\text{since } E[x_{t2}] = 0.9\,E[x_{t2}] \Rightarrow E[x_{t2}] = 0 \end{aligned}\right\} \text{and } x_{t1}, x_{t2} \text{ stationary}$$

$$\gamma_{y_{t1},y_{t2}} = \operatorname{cov}(y_{t1}(t+h), y_{t2}(t)) = E[(y_{t1}(t+h) - \overset{0}{\cancel{M_1}})(y_{t2}(t) - \overset{0}{\cancel{M_2}})] = E[y_{t1}(t+h)\, y_{t2}(t)]$$

$$= E\Big[\Big(x_{t1}(t+h) + \sum_{j=1}^{t+h} w_j\Big)\Big(x_{t2}(t) + 5\sum_{i=1}^{t} w_i\Big)\Big] = E[x_{t1}(t+h)\, x_{t2}(t)] + 5\sum_{j=1}^{t+h}\sum_{i=1}^{t} E[w_j w_i]$$

$5 \cdot \sigma_w^2 \min\{t+h,t\}$

$$= E[(0.5\, x_{t1}(t+h) \cancel{+ w_{t1}})(0.9\, x_{t2}(t-1) + w_{t2})] + 5\min\{t+h,t\}$$

0 assuming x_{t1}, x_{t2} indep.

(e) (Rest) $\gamma_{y_{t1}, y_{t2}}(h) = 5\min\{t+h, t\}$ dependent of time.

$$\rho(y_{t1}(t+h), y_{t2}(t)) = \frac{\gamma(y_{t1}(t+h), y_{t2}(t))}{\sqrt{var(y_{t1})\, var(y_{t2})}} = \frac{5\min\{t+h, t\}}{\sqrt{\left(\frac{1}{1-(0.5)^2}+t\right)\left(\frac{1}{1-(0.9)^2}+25t\right)}}$$ is dependent of time t.

Therefore, $\{y_{t1}\}$ and $\{y_{t2}\}$ are not jointly stationary. (It is also clear from the fact that y_{t1} and y_{t2} are not stationary.)

Problem II. | 3.5 Show that $x_t - x_{t-1} = \delta + w_t + y_t - y_{t-1}$ is stationary where w_t is white noise and y_t is stationary, and independent of each other.

i) $E[x_t - x_{t-1}] = E[\delta + w_t + y_t - y_{t-1}] = \delta + \cancel{\underbrace{E[y_t]}_{\mu_y} - \underbrace{E[y_{t-1}]}_{\mu_y}}^{0} = \delta$ } since y_t is stationary.

ii) $\gamma_{(x_t - x_{t-1})}(t+h, t) = cov(\delta + w_{t+h} + y_{t+h} - y_{t+h-1}, \delta + w_t + y_t - y_{t-1})$

$$= \underbrace{cov(w_{t+h}, w_t)}_{\begin{cases}\sigma_w^2, & \text{if } h=0\\ 0, & \text{otherwise}\end{cases}} + \underbrace{cov(y_{t+h}, y_t)}_{+\gamma_y(h)} - \underbrace{cov(y_{t+h}, y_{t-1})}_{-\gamma_y(h+1)} - \underbrace{cov(y_{t+h-1}, y_t)}_{-\gamma_y(h-1)} + \underbrace{cov(y_{t+h-1}, y_{t-1})}_{+\gamma_y(h)}$$

$$= \begin{cases} \sigma_w^2 + 2\gamma_y(0) - 2\gamma_y(1), & h=0 \\ \underbrace{2\gamma_y(h) - \gamma_y(h+1) - \gamma_y(h-1)}_{\gamma_z(h) \text{ where } z_t = y_t - y_{t-1}}, & h \neq 0 \end{cases}$$ (since $\gamma_y(1) = \gamma_y(-1)$)

Since both (i) and (ii) independent of time and (ii) is a function of lag (h), $(x_t - x_{t-1})$ is stationary.

Solutions of 3.6 - 3.7 - 3.8 will be in a seperate (Latex scan).

4.2 $\{w_t\} \sim WN(0, \sigma_w^2)$ ande $|\phi|<1$ constant. Consider the process $x_0 = w_0$, and $x_t = \phi x_{t-1} + w_t$ $t = 1, 2, \ldots$

(a) Show that $x_t = \sum_{j=0}^{t} \phi^j w_{t-j}$ for $\forall t = 0, 1, \ldots$ $\quad t=0 \Rightarrow \boxed{x_0 = w_0}$

$t=1 \Rightarrow x_1 = \phi x_0 + w_1 = \phi^0 w_1 + \phi w_0$

$t=2 \Rightarrow x_2 = \phi x_1 + w_2 = \phi^0 w_2 + \phi w_1 + \phi^2 w_0$

$t=3 \Rightarrow x_3 = \phi x_2 + w_3 = \phi^0 w_3 + \phi w_2 + \phi^2 w_1 + \phi^3 w_0$

$\vdots$

$$x_t = \phi^0 w_t + \phi^1 w_{t-1} + \cdots + \phi^{t-1} w_1 + \phi^t w_0 = \boxed{\sum_{j=0}^{t} \phi^j w_{t-j}}$$

(c) Show that, for $t = 0, 1, \ldots$ $var(x_t) = \frac{\sigma_w^2}{1-\phi^2}(1 - \phi^{2(t+1)})$

$$var(x_t) = var\left(\sum_{j=0}^{t} \phi^j w_{t-j}\right) = \sum_{j=0}^{t} \phi^{2j}\, var(w_{t-j}) = \sigma_w^2 \underbrace{(1 + \phi^2 + (\phi^2)^2 + \cdots + (\phi^2)^t)}_{= \frac{1-(\phi^2)^{t+1}}{1-\phi^2}}$$

$$= \frac{\sigma_w^2}{1-\phi^2}\left[1 - \phi^{2(t+1)}\right]$$

4.2 b) Find the $E[x_t]$.

$$E[x_t] = E\left[\sum_{j=0}^{t} \phi^j w_{t-j}\right] = \sum_{j=0}^{t} \phi^j \, \cancel{E[w_{t-j}]}^{\,0} = \sum_{j=0}^{t} \phi^j \cdot 0 = 0$$

c) is solved in the previous page.

d) Show that, for $h \geq 0$ $\quad cov(x_{t+h}, x_t) = \phi^h var(x_t)$.

$$cov(x_{t+h}, x_t) = cov(\phi x_{t+h-1} + w_{t+h}, x_t) = cov(\phi^2 x_{t+h-2} + \phi w_{t+h-1} + w_{t+h}, x_t)$$

$$\underset{\text{h times back}}{= \cdots} = cov\left(\phi^h x_t + \sum_{j=0}^{h-1} \phi^j w_{(t+h)-j}, x_t\right)$$

$$= \phi^h cov(x_t, x_t) + \sum_{j=0}^{h-1} \phi^j \, \cancel{cov(w_{(t+h)-j}, x_t)}^{\,0} = \boxed{\phi^h var(x_t)} \quad \text{for } h \geq 0.$$

e) Is x_t stationary?

$$\gamma_x(t+h, t) = cov(x_{t+h}, x_t) = \phi^h var(x_t) = \phi^h \frac{\sigma_w^2}{1-\phi^2}(1-\phi^{2(t+1)}) \rightarrow \text{time dependent}$$

Since $\gamma_x(t+h, t)$ is time dependent, x_t is not stationary.

f) Argue that, as $t \to \infty$, the process becomes stationary, so in a sense, x_t is "asymptotically stationary."

$$\lim_{t\to\infty} \gamma_x(t+h, t) = \lim_{t\to\infty} \frac{\phi^h \sigma_w^2}{1-\phi^2}(1-\phi^{2(t+1)}) = \frac{\phi^h \sigma_w^2}{1-\phi^2}\Big(1 - \underbrace{\lim_{t\to\infty} \phi^{2(t+1)}}_{=0 \text{ since } |\phi|<1,\ \phi^{2(t+1)} \to 0 \text{ as } t\to\infty}\Big)$$

$$= \frac{\phi^h \sigma_w^2}{1-\phi^2}$$

Therefore, $\gamma_x(h)$ is now just a function of lag (h) and since $E[x_t]=0$ also independent of time x_t is "asymptotically stationary."

g) Comment on how you use these results to simulate n observations of a stationary Gaussian AR(1) model from simulated iid $N(0,1)$ values.

Since as $t \to \infty$ the process becomes stationary, initial effects are strong in the stationarity. Therefore, it makes more sense to generate more observations than needed and then discarding the beginning in order to reduce the effect of initial terms.

h) Now suppose $x_0 = w_0/\sqrt{1-\phi^2}$. Is this process stationary?

$$x_1 = \phi x_0 + w_1 = \frac{\phi w_0}{\sqrt{1-\phi^2}} + w_1 \qquad \boxed{x_t = \phi x_{t-1} + w_t}$$

$\vdots$

$$x_t = \frac{\phi^t w_0}{\sqrt{1-\phi^2}} + \phi^{t-1} w_1 + \cdots + \phi^0 w_t = \phi^t w_0 + \sqrt{1-\phi^2}\left(\sum_{j=0}^{t-1} \phi^j w_{t-j}\right)$$

$$= \frac{1-\phi^{2[(t-1)+1]}}{1-\phi^2}$$

$$var(x_t) = \frac{\phi^{2t}\sigma_w^2 + (1-\phi^2)\sum_{j=0}^{t-1} \phi^{2j}\sigma_w^2}{(1-\phi^2)} = \frac{\sigma_w^2\left(\phi^{2t} + \cancel{(1-\phi^2)}\dfrac{(1-\phi^{2(t-1+1)})}{\cancel{(1-\phi^2)}}\right)}{(1-\phi^2)}$$

$$= \frac{\sigma_w^2}{(1-\phi^2)} \Rightarrow \gamma_x(t+h, t) = \frac{\phi^h \sigma_w^2}{(1-\phi^2)}$$ independent of time, function of lag h.

Therefore, x_t is stationary $\left|\text{ if } x_0 = \frac{w_0}{\sqrt{1-\phi^2}}\right.$

Problem III: Compute the mean and autocovariance functions of

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j} - \frac{\phi}{(1-\phi^2)} w_{t+1}, \quad |\phi|<1$$

and decide if it is stationary. Is the process causal?

I. method:

(i) $E[x_t] = \sum_{j=0}^{\infty} \phi^j E[w_{t-j}] - \frac{\phi}{(1-\phi^2)} E[w_{t+1}] = 0$

(ii) $\gamma_x(t+h,t) = \text{cov}(x_{t+h}, x_t) = E\left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j} - \frac{\phi}{(1-\phi^2)} w_{t+h+1}\right)\left(\sum_{i=0}^{\infty} \phi^i w_{t-i} - \frac{\phi}{(1-\phi^2)} w_{t+1}\right)\right]$

$$= E\left[\left(\frac{-\phi}{(1-\phi^2)} w_{t+h+1} + w_{t+h} + \phi^1 w_{t+h-1} + \phi^2 w_{t+h-2} + \cdots\right)\left(\frac{-\phi}{(1-\phi^2)} w_{t+1} + w_t + \phi^1 w_{t-1} + \phi^2 w_{t-2} + \cdots\right)\right]$$

$$= \begin{cases} \frac{\phi^2}{(1-\phi^2)^2}\sigma_w^2 + \underbrace{\underbrace{\sum_{i=0}^{\infty} \phi^{2i}}_{\frac{1}{(1-\phi^2)}}}_{\text{multiplying by } (1-\phi^2) \text{ and dividing yields} \Rightarrow \frac{\phi^2\sigma_w^2}{(1-\phi^2)^2} + \frac{(1-\phi^2)\sigma_w^2}{(1-\phi^2)^2}} \sigma_w^2, & h=0 \\ 0, & h\neq 0 \end{cases} = \begin{cases} \frac{\sigma_w^2}{(1-\phi^2)^2}, & h=0 \\ 0, & h\neq 0 \end{cases}$$

since

$$h=\pm 1 \quad \gamma_x(\pm 1) = \frac{-\phi\sigma_w^2}{(1-\phi^2)} + \underbrace{(\phi^1+\phi^3+\cdots+\phi^{2i+1})}_{\sum_{i=0}^{\infty}\phi^{2i+1} = \phi\sum_{i=0}^{\infty}\phi^{2i} = \frac{\phi}{(1-\phi^2)}}\sigma_w^2 = \frac{-\phi\sigma_w^2}{(1-\phi^2)} + \frac{\phi\sigma_w^2}{(1-\phi^2)} = 0$$

$$h=\pm 2 \quad \gamma_x(\pm 2) = \frac{-\phi^2\sigma_w^2}{(1-\phi^2)} + \underbrace{(\phi^2+\phi^4+\cdots+\phi^{2i+2})}_{\sum_{i=0}^{\infty}\phi^{2i+2} = \phi^2\sum_{i=0}^{\infty}\phi^{2i} = \frac{\phi^2}{(1-\phi^2)}}\sigma_w^2 = \frac{-\phi^2\sigma_w^2}{(1-\phi^2)} + \frac{\phi^2\sigma_w^2}{(1-\phi^2)} = 0$$

Therefore, $\gamma_x(h)=0$ for $h\neq 0$ and since $E[x_t]$ is constant and independent of time, $\gamma_x(h)$ is independent of time and (fixed - function of lag) x_t is a <u>stationary</u> process. Moreover, it is <u>not</u> causal since it depends on <u>future</u> white noise term w_{t+1}.

Problem IV: (Bonus) Let x_t be the standard AR(p) model with coefficients $\phi_1, \dots, \phi_p$. Show that (5)

$$\nabla x_t = \gamma x_{t-1} + \left(\sum_{j=1}^{p-1} \psi_j \nabla x_{t-j} \right) + w_t$$

where $\gamma = \left(\sum_{j=1}^{p} \phi_j \right) - 1$ and $\psi_j = - \sum_{i=j+1}^{p} \phi_i$ where $j = 1, 2, \dots, p-1$.

I believe there are 2 typos in the question. Blue parts, the parts I fixed.

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \dots + \phi_{p-1} x_{t-(p-1)} + \phi_p x_{t-p} + w_t$$

$$= (\phi_1 + \phi_2 + \dots + \phi_p) x_{t-1} - (\phi_2 + \phi_3 + \dots + \phi_p) x_{t-1}$$

$$+ (\phi_2 + \phi_3 + \dots + \phi_p) x_{t-2} - (\phi_3 + \phi_4 + \dots + \phi_p) x_{t-2}$$

$$+ (\phi_3 + \phi_4 + \dots + \phi_p) x_{t-3} - (\phi_4 + \phi_5 + \dots + \phi_p) x_{t-3}$$

$$\vdots$$

$$+ (\phi_{p-1} + \phi_p) x_{t-(p-1)} - \phi_p x_{t-(p-1)} + \phi_p x_{t-p} + w_t$$

Now, substracting x_{t-1} from both sides and collecting the terms with same coefficients yields;

$$\nabla x_t = x_t - x_{t-1} = \left[\left(\sum_{i=1}^{p} \phi_i \right) - 1 \right] x_{t-1} - \sum_{i=2}^{p} \phi_i \underbrace{(x_{t-1} - x_{t-2})}_{\nabla x_{t-1}} - \sum_{i=3}^{p} \phi_i \underbrace{(x_{t-2} - x_{t-3})}_{\nabla x_{t-2}}$$

$$- \dots - \sum_{i=p-1}^{p} \phi_i \underbrace{(x_{t-(p-2)} - x_{t-(p-1)})}_{\nabla x_{t-(p-2)}} - \sum_{i=p} \phi_i \underbrace{(x_{t-(p-1)} - x_{t-p})}_{\nabla x_{t-(p-1)}}$$

$$= \left[\left(\sum_{i=1}^{p} \phi_i \right) - 1 \right] x_{t-1} \left(\sum_{j=1}^{p-1} \left(- \sum_{i=j+1}^{p} \phi_i \right) \nabla x_{t-j} \right) + w_t$$

$$\boxed{\nabla x_t = \gamma x_{t-1} + \left(\sum_{j=1}^{p-1} \psi_j \nabla x_{t-j} \right) + w_t}$$

■