Probabilistic Time Series Analysis: Homework 1 Solutions

Tim Kunisky

October 21, 2018

Problem 1

Consider a time series constructed as

$$x_t = 2\cos\left(2\pi \cdot \frac{t+15}{50}\right) + w_t,$$

where w_t are independent white noise variables with zero mean and variance σ^2.

(a) What is the ACF of this process?

Note that

$$\mathbb{E}x_t = 2\cos\left(2\pi \cdot \frac{t+15}{50}\right),$$

so $x_t - \mathbb{E}x_t = w_t$. Thus, the autocovariance is

$$R_x(s,t) = \mathbb{E}\left[(x_s - \mathbb{E}x_s)(x_t - \mathbb{E}x_t)\right] = \mathbb{E}[w_s w_t] = R_w(s,t) = \begin{cases} 0 & : \ s \neq t \\ \sigma^2 & : \ s = t \end{cases}$$

The autocorrelation is then

$$\rho_x(s,t) = \frac{R_x(s,t)}{\sqrt{R_x(s,s)R_x(t,t)}} = \frac{R_x(s,t)}{\sigma^2} = \begin{cases} 0 & : \ s \neq t \\ 1 & : \ s = t \end{cases}$$

(b) Simulate a series of $n = 500$ observations from this process for $\sigma^2 = 1$. Compute the sample ACF of the data you generated up to a lag of 100 and comment.

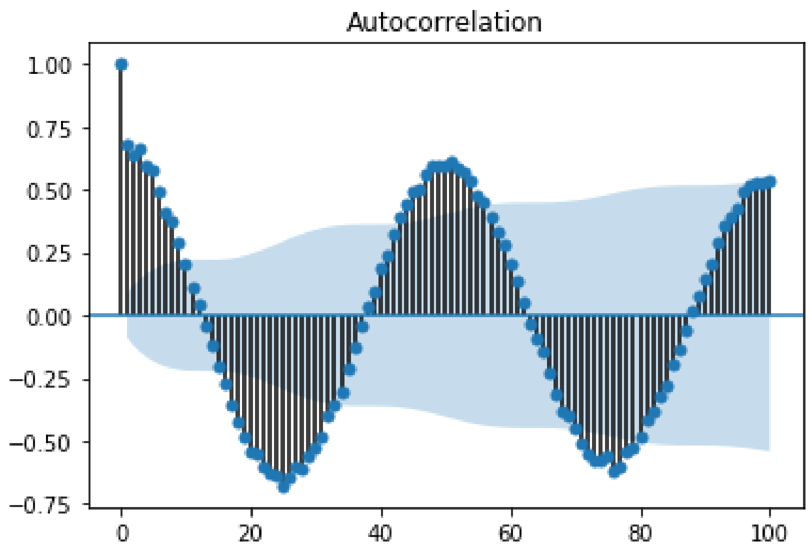
The code to generate the plot should be something like this:

```
import numpy as np
from statsmodels.graphics.tsaplots import plot_acf

n = 500
nlags = 100

t = np.arange(n)
w = np.random.normal(size=n)
x = 2.0 * np.cos(2.0 * np.pi * (t + 15.0) / 50.0) + w

plot_acf(x, lags=nlags)
```



The sample ACF plot does not resemble the theoretical ACF derived above. This is a result of the series not being stationary, whereby the sample ACF from a single sample does not estimate the theoretical ACF. That is because under stationarity, pairs $(x_t, x_{t+\ell})$ from a single sample with various t and a fixed ℓ are identically distributed, so it is reasonable to use them as samples to estimate statistics of the true joint distribution of x_t and $x_{t+\ell}$, which does not depend on t. Without stationarity, the true joint distribution of x_t and $x_{t+\ell}$ may depend on t, and pairs as above from a single sample need not be identically distributed (in our case, the most obvious indicator that stationarity fails is that the mean of x_t depends on t). Thus, it is no longer reasonable to view the pairs $(x_t, x_{t+\ell})$ for varying t as samples of the same random variable.

Problem 2

Consider the following ARMA(p, q) models:

$$\begin{aligned} x_t &= 0.8x_{t-1} - 0.2x_{t-2} + w_t - 1.1w_{t-1}, \\ y_t &= 0.7y_{t-1} - 0.1y_{t-2} + w_t - 0.2w_{t-1} \end{aligned}$$

(for the sake of clarity I will rewrite the second one with the variable "y" in these solutions).

(a) Rewrite the expressions using the backward operator.

If B is the backward operator, we have

$$\begin{aligned} (1 - 0.8B + 0.2B^2)\boldsymbol{x} &= (1 - 1.1B)\boldsymbol{w}, \\ (1 - 0.7B + 0.1B^2)\boldsymbol{y} &= (1 - 0.2B)\boldsymbol{w}. \end{aligned}$$

(b) Determine the parameters p and q for the resulting model. Identify and eliminate potential parameter redundancy.

Initially, it appears we have $p = 2$ and $q = 1$ for both models. To check for parameter redundancy, we factorize the degree 2 backward operator polynomials on the left-hand sides:

$$1 - 0.8B + 0.2B^2 = 0.2(B - (2 + i))(B - (2 - i)),$$
$$1 - 0.7B + 0.1B^2 = 0.1(B - 2)(B - 5).$$

Therefore the $\boldsymbol{x}$ model does not have parameter redundancy, so its minimal parameters are indeed $p = 2$ and $q = 1$. In the $\boldsymbol{y}$ model, on the right-hand side we have $1 - 0.2B = 0.2(5 - B)$, so we may reduce to

$$(1 - 0.5B)\boldsymbol{y} = \boldsymbol{w},$$

and the minimal parameters are $p = 1$ and $q = 0$.

(c) Are these models causal and/or invertible?

For the $\boldsymbol{x}$ model, the roots of the left-hand side polynomial are $2 \pm i$ and the only root of the right-hand side polynomial is 1.1. All of these lie outside the unit circle, so the model is both causal and invertible.

For the $\boldsymbol{y}$ model, after eliminating parameter redundancy, the right-hand side polynomial is constant and the coefficients of the left-hand side polynomial are summable (because they are finite; generally this might be an infinite power series in which case there would be something to check), so the model is invertible. The only root of the left-hand side polynomial is 2, which lies outside the unit circle, so the model is also causal.

Problem 3

Prove that if random variables U and V can be written as linear combinations of the form $U = \sum_i a_i X_i$ and $V = \sum_j b_j Y_j$, then

$$\mathsf{Cov}(U, V) = \sum_{i,j} a_i b_j \mathsf{Cov}(X_i, Y_j).$$

We compute using the definition of covariance and linearity of expectation:

$$\begin{aligned}
\mathsf{Cov}(U, V) &= \mathbb{E}\left[(U - \mathbb{E}U)(V - \mathbb{E}V)\right] \\
&= \mathbb{E}\left[\left(\sum_i a_i X_i - \sum_i a_i \mathbb{E}X_i\right)\left(\sum_j b_j Y_j - \sum_j b_j \mathbb{E}Y_j\right)\right] \\
&= \mathbb{E}\left[\left(\sum_i a_i (X_i - \mathbb{E}X_i)\right)\left(\sum_j b_j (Y_j - \mathbb{E}Y_j)\right)\right] \\
&= \mathbb{E}\left[\sum_{i,j} a_i b_j (X_i - \mathbb{E}X_i)(Y_j - \mathbb{E}Y_j)\right] \\
&= \sum_{i,j} a_i b_j \mathbb{E}\left[(X_i - \mathbb{E}X_i)(Y_j - \mathbb{E}Y_j)\right] \\
&= \sum_{i,j} a_i b_j \mathsf{Cov}(X_i, Y_j).
\end{aligned}$$

Problem 4

Consider a series of the form

$$y_t = a_2 t^2 + a_1 t + a_0 + x_t$$

where a_t are nonzero parameters and x_t is a stationary series.

(a) Repeatedly apply the difference operator ∇ to reduce $\boldsymbol{y}$ to a stationary series.

Differencing once, we have

$$\begin{aligned}(\nabla \boldsymbol{y})_t &= y_t - y_{t-1} \\ &= a_2(t^2 - (t-1)^2) + a_1(t - (t-1)) + (\nabla \boldsymbol{x})_t \\ &= a_2(2t-1) + a_1 + (\nabla \boldsymbol{x})_t.\end{aligned}$$

Since there is still a linear dependence on t, the series is still not stationary. Differencing again, we find

$$\begin{aligned}(\nabla^2 \boldsymbol{y})_t &= (\nabla \boldsymbol{y})_t - (\nabla \boldsymbol{y})_{t-1} \\ &= a_2(2t - 1 - (2(t-1) - 1)) + (\nabla^2 \boldsymbol{x})_t \\ &= 2a_2 + (\nabla^2 \boldsymbol{x})_t.\end{aligned}$$

Since $\boldsymbol{x}$ is stationary, $\nabla^2 \boldsymbol{x}$ is stationary, and therefore $\nabla^2 \boldsymbol{y}$ is stationary as well (since shifting by a constant does not affect stationarity).

(b) Rewrite the original expressions using the backward operator.

The differencing operator may be written in terms of the backward operator as $\nabla = 1 - B$. Therefore, the above result may equivalently be written as

$$(1-B)^2 \boldsymbol{y} = 2a_2 + (1-B)^2 \boldsymbol{x}.$$

(c) Comment on the generalization of this idea to mth order polynomial dependence in t.

Generally, if $p(t)$ is a degree m polynomial, then we may write $p(t) = a_m t^m + q(t)$ where $q(t)$ has degree at most $m-1$. The effect of differencing on such a polynomial is

$$\begin{aligned}p(t) - p(t-1) &= a_m(t^m - (t-1)^m) + q(t) - q(t-1) \\ &= a_m\left(t^m - \sum_{k=0}^{m} \binom{m}{k} (-1)^k t^{m-k}\right) + q(t) - q(t-1) \\ &= a_m\left(m t^{m-1} - \sum_{k=2}^{m} \binom{m}{k} (-1)^k t^{m-k}\right) + q(t) - q(t-1) \\ &= m a_m t^{m-1} + r(t)\end{aligned}$$

where $r(t)$ is some polynomial of degree at most $m-2$, and where we have used the binomial theorem. Thus, differencing a degree m polynomial reduces the degree by one and multiplies the leading coefficient by m.

Applying this inductively, we see that if $y_t = p(t) + x_t$ where $p(t)$ has degree m and leading term $a_m t^m$, then

$$(\nabla \boldsymbol{y})_t = m! \cdot a_m + (\nabla \boldsymbol{x})_t,$$

which is stationary if $\boldsymbol{x}$ is stationary.