ECON 706: PROBLEM SET 4

DANIEL PFEFFER

Problem 1. Estimating a Trend-Stationary Model

(i) Download quarterly U.S. real GDP from the FRED database maintained by the Federal Reserve Bank of St. Louis for 1985:Q1 to 2019:Q4.

Solution. Query real GDP data for 1985:Q1 to 2019:Q4 from the FRED.

```python
import numpy as np
import pandas as pd
import pandas_datareader.data as web
import datetime

# 1985:Q1 to 2019:Q4
start = datetime.datetime(1985, 1, 1)
end = datetime.datetime(2019, 12, 31)

# Quarterly U.S. real GDP from the FRED
ts = web.DataReader('GDPC1', 'fred', start, end)
```

□

(ii) What do we mean by quarterly GDP? Goods and services produced within the quarter, or the goods and services that would be produced in the calendar year if the output during the quarter prevails over 12 months?

Solution. According to the FRED, the GDPC1 series is measured in "Billions of Chained 2009 Dollars, Seasonally Adjusted Annual Rate". So quarterly GDP refers to the goods and services that would be produced in the calendar year if the output during the quarter prevails over 12 months. □

(iii) Now consider the following model:

$$y_t = \gamma_1 + \gamma_2 t + \tilde{y}_t, \qquad \tilde{y}_t = \rho \tilde{y}_{t-1} + \eta_t, \qquad \eta_t \sim \text{iid } N(0, \sigma_\eta^2), \tag{1}$$

where $y_t := \log \text{GDP}_t$. Rewrite the model as

$$y_t = \beta_1 + \beta_2 t + \beta_3 y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{iid } N(0, \sigma_\varepsilon^2). \tag{2}$$

What is the relationship between $(\gamma_1, \gamma_2, \rho, \sigma_\eta^2)$ and $(\beta_1, \beta_2, \beta_3, \sigma_\varepsilon^2)$?

Solution. Using lag operators, we can write

$$(1 - \rho L)y_t = \eta_t$$

and express (1) as

$$y_t = \gamma_1 + \gamma_2 t + (1 - \rho L)^{-1} \eta_t,$$

1

provided $(1-\rho L)^{-1}$ exists. Then applying the quasi-difference operator $\Delta^\rho := (1-\rho L)$ to both sides gives

$$\begin{aligned}
\Delta^\rho y_t &= (1-\rho L)(\gamma_1 + \gamma_2 t + (1-\rho L)^{-1}\eta_t) \\
&= \gamma_1 + \gamma_2 t - \rho\gamma_1 - \rho\gamma_2(t-1) + \eta_t \\
&= (1-\rho)\gamma_1 + \rho\gamma_2 + (1-\rho)\gamma_2 t + \eta_t
\end{aligned}$$

or

$$y_t = (1-\rho)\gamma_1 + \rho\gamma_2 + (1-\rho)\gamma_2 t + \rho y_{t-1} + \eta_t.$$

Hence $(\beta_1, \beta_2, \beta_3, \sigma_\varepsilon^2) = ((1-\rho)\gamma_1 + \rho\gamma_2, (1-\rho)\gamma_2, \rho, \sigma_\eta^2)$. $\qquad\square$

(iv) Estimate (2) by OLS. What is your estimate of the average growth rate of output (be precise about the units of measurement)? What is the estimated autocorrelation of the fluctuations around the deterministic trend?

Solution. The following code estimates (2) by OLS.

```
import statsmodels.api as sm

# Add constant
ts = sm.add_constant(ts)

# Add trend variable
ts['trend'] = ts.const.cumsum()

# Log transformation
ts['y'] = np.log(ts['GDPC1'])

# Add lag term
ts['y_lag1'] = ts['y'].shift()

# Drop first observation
ts = ts[1:]

# Fit model by OLS
mod = sm.OLS(ts['y'], ts[['const', 'trend', 'y_lag1']])
res = mod.fit()
```

Since γ_2 in (1) can be interpreted as the quarterly growth rate, an estimate of average quarterly growth rate can be obtained through the relation in (i), which implies $\gamma_2 = \beta_2/(1-\beta_3)$. Using the the OLS regression results in Table 1, the estimated average quarterly growth rate of approximately 0.42% or 1.5% per year, a poor estimate, since the actual average annual growth rate is roughly 2-3%.

To obtain the estimated autocorrelation of the fluctuations around the deterministic trend, since $\beta_3 = \rho$, we obtain an estimate of approximately 0.988, which indicates that the estimated process is highly persistent. This evidence suggests that this trend-stationary model is not a good choice and that the real GDP process perhaps contains a stochastic tend component.

$\qquad\square$

TABLE 1. OLS regression results of Equation (2).

	$\log \text{GDP}_t$
const	0.117
	(0.105)
trend	0.000
	(0.000)
$\log \text{GDP}_{t-1}$	0.988***
	(0.012)

* $p < .1$, ** $p < .05$,
*** $p < .01$

(v) Use your OLS estimates of β and iterate (2) forward starting from y_1 and setting the error terms to zero. Plot the resulting trajectory and setting the error terms to zero. Plot the resulting trajectory and compare it to plots of $\tilde{y}_t = \hat{\beta}_1 + \hat{\beta}_2 t$ and $\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 t + \hat{\beta}_3 y_{t-1}$, where now you use the observed value y_{t-1}. Discuss.

Solution. The following code generates Figure 1.

```
# Generate trajectory
beta1 = res.params.values[0]
beta2 = res.params.values[1]
beta3 = res.params.values[2]
y = [ts['y'][0]]
for t in range(1, len(ts)):
    y.append(beta1 + beta2*ts['trend'][t] + beta3*y[t-1])

# Generate y_tilde
y_tilde =  beta1 + beta2*ts['trend']

# Plot results
import matplotlib.pyplot as plt

fig, [ax1, ax2] = plt.subplots(1, 2, figsize=(12,4))

ax1.plot(mod.data.orig_endog, label='$y_t$')
ax1.plot(res.fittedvalues, label='$\hat y_t$')
ax1.plot(ts.index, y, label='Trajectory')
ax1.legend()

ax2.plot(ts.index, y)
ax2.plot(res.fittedvalues)
ax2.plot(y_tilde, color='r', label='$\tilde y_t$')
ax2.legend()

plt.show()
```

The left panel in Figure 1 shows the fitted values of the model tracks real GDP relatively closely. Moreover, the resulting trajectory implied by OLS estimates of (2) tracks the upward trend in GDP. The right panel in Figure 1 shows that if the autocorrelation is omitted, GDP is essentially zero.
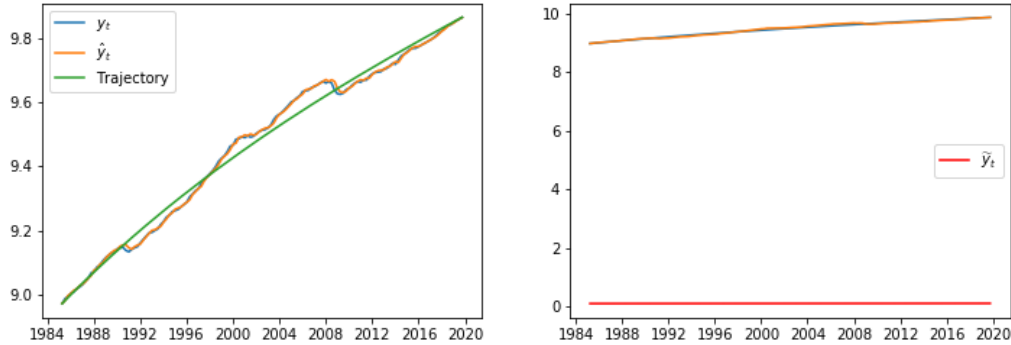

FIGURE 1

□

Problem 2. Frequentist Asymptotics – Part I

Consider the AR(1) model

$$y_t = \phi y_{t-1} + \eta_t, \qquad \eta_t \sim \text{iid N}(0, \sigma_\eta^2),$$

where $|\phi| < 1$.

(i) What is the difference between a conditional and unconditional likelihood function?

Solution. Let $y = (y_1, \ldots, y_T)'$, and $\theta = (\phi, \sigma_\eta^2)'$. The conditional likelihood function $p(y \mid y_0, \theta)$ assumes y_0 is known and conditions on this initial observation. The unconditional likelihood function $p(y, y_0 \mid \theta) = p(y \mid y_0, \theta)p(y_0 \mid \theta)$ is the joint density of the entire sample and therefore contains more "information".

The first-order Markovian structure and iid Gaussian innovations of the AR(1) yields a conditional likelihood function of the factorized form

$$p(y \mid y_0, \theta) = \prod_{t=1}^{T} p(y^t \mid y^{t-1}, y_0, \theta) = \frac{1}{(2\pi\sigma_\eta^2)^{T/2}} \exp\left(-\frac{1}{2\sigma_\eta^2} \sum_{t=1}^{T}(y_t - \phi y_{t-1})^2\right),$$

where $y^{t_k} = (y_1, \ldots, y_{t_k})'$.

To obtain the unconditional likelihood function, notice that if the process $\{y_t\}$ was initialized in the infinite past, then the assumption that $|\phi| < 1$ implies that $y_0 \sim N(0, \sigma_\eta^2/(1 - \phi))$, and hence

$$p(y, y_0 \mid \theta) = \frac{(1 - \phi^2)^{1/2}}{(2\pi\sigma_\eta^2)^{(T+1)/2}} \exp\left(-\frac{1}{2\sigma_\eta^2} \sum_{t=1}^{T}(y_t - \phi y_{t-1})^2 + \frac{(1 - \phi^2)}{2\sigma_\eta^2}y_0^2\right)$$

□

(ii) Derive the conditional MLE of ϕ, denoted $\hat{\phi}$.

Solution. The conditional log likelihood is

$$\log p(y \mid y_0, \theta) = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma_\eta^2 - \frac{1}{2\sigma_\eta^2} \sum_{t=1}^{T} (y_t - \phi y_{t-1})^2.$$

To find the MLE of ϕ we only solve

$$\min_{\phi} \sum_{t=1}^{T} (y_t - \phi y_{t-1})^2,$$

since maximization of $p(y \mid y_0, \theta)$ with respect to ϕ only depends on this expression. The first order condition necessary condition for is

$$\sum_{t=1}^{T} (y_t - \phi y_{t-1}) y_{t-1} = 0 \iff \sum_{t=1}^{T} y_t y_{t-1} - \phi \sum_{t=1}^{T} y_{t-1}^2 = 0$$

$$\iff \sum_{t=1}^{T} y_t y_{t-1} = \phi \sum_{t=1}^{T} y_{t-1}^2.$$

Provided the variance of y_{t-1} is nonzero, the MLE of ϕ is

$$\hat{\phi} = \frac{\sum_{t=1}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2} = \phi + \frac{T^{-1} \sum_{t=1}^{T} y_{t-1} \eta_t}{T^{-1} \sum_{t=1}^{T} y_{t-1}^2}.$$

$\square$

(iii) Show that $\hat{\phi}$ is a consistent estimator of ϕ.

Solution. If η_t is strictly stationary and ergodic and $|\phi| < 1$, then y_t, y_t^2, $y_t y_{t-1}$, and $y_{t-1} \eta_t$ are all measurable transformations of η_t. Then by the law of large numbers for strictly stationary and ergodic processes (White, 1984), the limit of the denominator is

$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \to \mathrm{E}\, y_{t-1}^2 = \frac{\sigma_\eta^2}{1 - \phi^2}.$$

Now we derive the limit of the numerator. By independence, $y_{t-1} \eta_t$ is a martingale difference sequence, i.e., $\mathrm{E}\, y_{t-1} \eta_t = 0$ with second moment

$$\mathrm{E}\, y_{t-1}^2 \eta_t^2 = \mathrm{E}\, y_{t-1}^2 \, \mathrm{E}\, \eta_t^2 = \frac{\sigma_\eta^2}{1 - \phi^2} \sigma_\eta^2.$$

So again by the law of large numbers for strictly stationary and ergodic processes

$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1} \eta_t \to \mathrm{E}\, y_{t-1} \eta_t = 0.$$

Combining these results implies $\hat{\phi} - \phi \to 0$. $\square$

(iv) Derive the limiting distribution of $\hat{\phi}$.

Solution. By the central limit theorem for martingale difference sequences (White, 1984),

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1} \eta_t \Rightarrow N\left(0, \frac{\sigma_\eta^4}{1 - \phi^2}\right)$$

and therefore

$$\sqrt{T}(\hat{\phi} - \phi) \Rightarrow N(0, V),$$

where $V = \sigma_\eta^4/(1-\phi^2) \cdot (\sigma_\eta^2/(1-\phi^2))^{-2} = 1 - \phi^2$. □

(v) Suppose data is generated from an AR(2) model

$$y_t = \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \varepsilon_t, \qquad \varepsilon_t \sim \text{iid N}(0, \sigma_\varepsilon^2), \tag{3}$$

but the econometrician estimates the above AR(1) model. What are the probability limits of $\hat{\phi}$ and $\hat{\sigma}_\eta^2$ in terms of $(\gamma_1, \gamma_2, \sigma_\varepsilon^2)$?

Solution. Assume that the coefficients γ_1 and γ_2 satisfy the stationarity conditions. By the law of large numbers for strictly stationary and ergodic processes, the denominator converges to

$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \to \gamma(0) = \frac{1 - \gamma_2}{(1 + \gamma_2)((1 - \gamma_2)^2 - \gamma_1^2)} \sigma_\varepsilon^2,$$

and the numerator converges to

$$\frac{1}{T} \sum_{t=1}^{T} y_t y_{t-1} \to \gamma(1) = \frac{\gamma_1}{1 - \gamma_2} \gamma(0).$$

So

$$\hat{\phi} \to \frac{\gamma_1}{1 - \gamma_2}$$

To derive $\hat{\sigma}_\eta^2$, first differentiate the conditional log likelihood with respect to σ_η^2, which yields the first-order necessary condition

$$-\frac{T}{2\sigma_\eta^2} - \frac{1}{2\sigma_\eta^4} \sum_{t=1}^{T} (y_t - \phi y_{t-1})^2 = 0,$$

which gives

$$\hat{\sigma}_\eta^2 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{\phi} y_{t-1})^2$$

$$= \frac{1}{T} \sum_{t=1}^{T} (y_t^2 - 2\hat{\phi} y_t y_{t-1} + \hat{\phi}^2 y_{t-1}^2)$$

$$= \frac{1}{T} \sum_{t=1}^{T} y_t^2 - 2\hat{\phi} \frac{1}{T} \sum_{t=1}^{T} y_t y_{t-1} + \hat{\phi}^2 \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2$$

$$= \frac{1}{T} \sum_{t=1}^{T} y_t^2 - 2 \left(\frac{T^{-1} \sum_{t=1}^{T} y_t y_{t-1}}{T^{-1} \sum_{t=1}^{T} y_{t-1}^2} \right) \frac{1}{T} \sum_{t=1}^{T} y_t y_{t-1} + \left(\frac{T^{-1} \sum_{t=1}^{T} y_t y_{t-1}}{T^{-1} \sum_{t=1}^{T} y_{t-1}^2} \right)^2 \frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2$$

$$= \frac{1}{T} \sum_{t=1}^{T} y_t^2 - \frac{(T^{-1} \sum_{t=1}^{T} y_t y_{t-1})^2}{T^{-1} \sum_{t=1}^{T} y_{t-1}^2}.$$

By the law of large numbers,

$$\frac{1}{T} \sum_{t=1}^{T} y_t^2 \to \text{E} \, y_t^2 = \gamma(0)$$

and

$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \to \gamma(0),$$

and, by the continuous mapping theorem and law of large numbers,

$$\left(\frac{1}{T}\sum_{t=1}^{T} y_t y_{t-1}\right)^2 \to (\mathrm{E}\, y_t y_{t-1})^2 = \gamma(0)^2 \left(\frac{\gamma_1}{1-\gamma_2}\right)^2.$$

Hence

$$\hat{\sigma}_\eta^2 \to \gamma(0) - \left(\frac{\gamma_1}{1-\gamma_2}\right)^2 \gamma(0) = \frac{\sigma_\varepsilon^2}{1-\gamma_2^2},$$

since

$$\gamma(0) - \left(\frac{\gamma_1}{1-\gamma_2}\right)^2 \gamma(0) = \gamma(0)\left(1 - \frac{\gamma_1^2}{(1-\gamma_2)^2}\right)$$

$$= \frac{(1-\gamma_2)\sigma_\varepsilon^2}{(1+\gamma_2)((1-\gamma_2)^2 - \gamma_1^2)} \frac{(1-\gamma_2)^2 - \gamma_1^2}{(1-\gamma_2)^2} = \frac{\sigma_\varepsilon^2}{1-\gamma_2^2}.$$

$\square$

Problem 3. Bayesian Analysis – Part II

Consider the AR(1) model

$$y_t = \phi y_{t-1} + \eta_t, \qquad \eta_t \sim \text{iid } \mathrm{N}(0, \sigma_\eta^2).$$

(i) Consider the following prior for ϕ: $\phi \sim N(0, \tau^2)$. Show that the posterior distribution of ϕ is normal and derive its mean and its variance.

Solution. Combining the likelihood function with prior distribution gives the posterior distribution:

$$p(\phi \mid y) \propto p(y \mid \phi)p(\phi)$$

$$\propto \exp\left\{-\frac{1}{2}\left(\sum_{t=1}^{T}(y_t - \phi y_{t-1})^2 + \frac{\phi^2}{\tau^2}\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\sum_{t=1}^{T} y_t^2 - 2\phi\sum_{t=1}^{T} y_t y_{t-1} + \phi^2\left(\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2}\right)\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\frac{\phi^2 - 2\phi\sum_{t=1}^{T} y_t y_{t-1}(\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2})^{-1}}{(\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2})^{-1}}\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\frac{(\phi - \sum_{t=1}^{T} y_t y_{t-1}(\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2})^{-1})^2}{(\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2})^{-1}}\right\},$$

which is of the form of a normal distribution:

$$\phi \mid y \sim N(\bar{\phi}, \overline{V}),$$

where

$$\bar{\phi} = \frac{\sum_{t=1}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2}}$$

$$\overline{V} = \left(\sum_{t=1}^{T} y_{t-1}^2 + \tau^{-2}\right)^{-1}.$$

$\square$

(ii) What happens to the posterior distribution if you let τ go to zero or infinity?

Solution. Letting the prior standard deviation go to zero $\tau \to 0$,

$$\bar{\phi} = \frac{\tau^2 \sum_{t=1}^T y_t y_{t-1}}{\tau^2 \sum_{t=1}^T y_{t-1}^2 + 1} \to 0,$$

and

$$\overline{V} = \frac{\tau^2}{\sum_{t=1}^T y_{t-1}^2 + 1} \to 0,$$

meaning that the posterior mean and variance converge to their prior mean and variance.

Letting $\tau \to \infty$,

$$\bar{\phi} \to \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2} = \hat{\phi},$$

and

$$\overline{V} \to \left(\sum_{t=1}^T y_{t-1}^2 \right)^{-1}.$$

□

(iii) Suppose the goal is to forecast y_{T+2} based on information up until time T, given by the sample $y = (y_1, \ldots, y_T)'$. Show that under the loss function

$$L(y_{T+2}, \hat{y}_{T+2|T}) = (y_{T+2} - \hat{y}_{T+2|T})^2, \tag{4}$$

where y_{T+2} is the actual value and $\hat{y}_{T+2|T}$ is the predicted value, the optimal (minimizing posterior expected loss) forecast is given by

$$\hat{y}_{T+2|T}^* = \mathrm{E}(y_{T+2} \mid y) = \mathrm{E}_y\, y_{T+2} \tag{5}$$

Solution. To verify this claim, consider basing $\hat{y}_{T+2|T}^*$ on any function g not equal to $\mathrm{E}_y\, y_{T+2}$. Under g, adding and subtracting the conditional expectation, the expected posterior loss is

$\mathrm{E}_y (y_{T+2} - g(y))^2$

$= \mathrm{E}_y (y_{T+2} - \mathrm{E}_y\, y_{T+2} + \mathrm{E}_y\, y_{T+2} - g(y))^2$

$= \mathrm{E}_y (y_{T+2} - \mathrm{E}_y\, y_{T+2})^2 + 2\, \mathrm{E}_y ((y_{T+2} - \mathrm{E}_y\, y_{T+2})\, \mathrm{E}_y (y_{T+2} - g(y)) + \mathrm{E}_y\, \mathrm{E}_y (y_{T+2} - g(y))^2$

If we condition on $y_{T+2|T}$ in the cross product term and take expectations, then

$$\mathrm{E}_y (y_{T+2} - \mathrm{E}_y (y_{T+2} \mid y_{T+2|T}))\, \mathrm{E}_y ((y_{T+2} \mid y_{T+2|T}) - g(y)) \mid y_{T+2|T})$$
$$= \mathrm{E}_y ((y_{T+2} \mid y_{T+2|T}) - g(y))\, \mathrm{E}_y (y_{T+2} - \mathrm{E}_y (y_{T+2} \mid y_{T+2|T}))$$
$$= \mathrm{E}_y ((y_{t+2} \mid y_{T+2|T}) - g(y)) \cdot 0 = 0,$$

where the last equality follows from the law of iterated projections. Our expected quadratic loss is now

$$\mathrm{E}_y (y_{T+2} - g(y))^2 = \mathrm{E}_y (y_{T+2} - \mathrm{E}_y\, y_{T+2})^2 + \mathrm{E}_y (y_{T+2} - g(y))^2,$$

which attains a minimum when g is such that $g(y) = \mathrm{E}_y\, y_{T+2}$. □

(iv) Using the results from (i), calculate the optimal two-step ahead predictor for the estimated AR(1) model.

Solution. Since η_τ for $\tau > T$ is independent of $\phi \mid y$, we have

$$
\begin{aligned}
\hat{y}_{T+2|T}^* &= \mathrm{E}_y(\phi^2 y_T + \phi \eta_T + \eta_{T+1}) \\
&= y_T \, \mathrm{E}_y \, \phi^2 + \mathrm{E}_y \, \phi \, \mathrm{E}_y \, \eta_T + \mathrm{E}_y \, \eta_{T+1} \\
&= (\overline{V} + \phi^2) y_T.
\end{aligned}
$$

$\square$