Solutions to Selected Problems in Time Series Analysis

JaR

June 4, 2020

This document contains solutions to selected problems in

> Brockwell, P. J., Davis, R. A., and Fienberg, S. E. (1991). *Time series: theory and methods*. Springer Science & Business Media.

Readers are encouraged to provide suggestions to improve the solutions and to report any mistake or typo that may be found.

HW9 Problem 1 (4.12) It suffices to minimize $|1-1.317\mathrm{e}^{-\mathrm{i}\lambda}+0.634\mathrm{e}^{-2\mathrm{i}\lambda}|^2$. See Figure 1. Many mistakely took 0 as the answer. The corresponding period should be $2\pi/\lambda$ (which is not important actually).

HW9 Problem 4 (5.3) Hanxiang Shen gave a simple answer. Let

$$\gamma(h) = \begin{cases} 1, & h = 0, \\ 1/2, & h \neq 0. \end{cases}$$

Then it is straightforward to verify that γ is an even and positive definite function.

HW9 Problem 5 (5.5) Using the equations provided in Example 5.2.1 (p. 173), we have

$$\begin{aligned} v_0 &= (1+\theta^2)\sigma^2, \\ v_n &= (1+\theta^2 - v_{n-1}^{-1}\theta^2\sigma^2)\sigma^2, \end{aligned}$$

and one can easily show that $v_n \geq \sigma^2$. Also, we have

$$v_n - v_{n-1} = \sigma^4\theta^2 \frac{v_{n-1} - v_{n-2}}{v_{n-1}v_{n-2}} \leq 0,$$

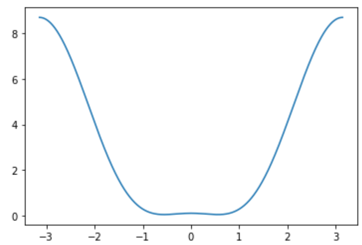
which follows that $\lim v_n$ exists and is σ^2, and that $\theta_{n1} \to \theta$.

For problem (a), according to Example 5.2.1, we have

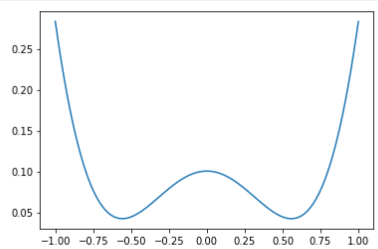
$$\hat{X}_n = \sigma^2\theta(X_{n-1} - \hat{X}_{n-1})/v_{n-2}.$$

In [1]:
```python
import numpy as np
import matplotlib.pyplot as plt

from numpy import sin, cos, pi
from scipy.optimize import minimize
```

In [2]:
```python
def f(x):
    return ( (1 - 1.317*cos(x) + 0.634*cos(2*x))**2
             + (1.317*sin(x) - 0.634*sin(2*x))**2 )

x = np.linspace(-pi, pi, 1000)
y = f(x)
plt.plot(x, y)
plt.show()
```



In [3]:
```python
x = np.linspace(-1, 1, 1000)
y = f(x)
plt.plot(x, y)
plt.show()
```



In [4]:
```python
minimize(f, 0.5)
```

Out[4]:
```
      fun: 0.042337226938543
 hess_inv: array([[0.70782066]])
      jac: array([4.22820449e-07])
  message: 'Optimization terminated successfully.'
     nfev: 15
      nit: 4
     njev: 5
   status: 0
  success: True
        x: array([0.55751665])
```

Figure 1: HW9 Problem 1 (4.12) in Python

After easy calculation,

$$\begin{aligned}\|X_n - \hat{X}_n - Z_n\| &= \left\| \left(\theta - \frac{\sigma^2\theta}{v_{n-2}}\right) Z_{n-1} - \frac{\sigma^2\theta}{v_{n-2}}(X_{n-1} - \hat{X}_{n-1} - Z_{n-1}) \right\| \\ &\le \theta \left(1 - \frac{\sigma^2}{v_{n-2}}\right) \|Z_{n-1}\| + \frac{\sigma^2\theta}{v_{n-2}} \left\| X_{n-1} - \hat{X}_{n-1} - Z_{n-1} \right\|.\end{aligned}$$

For any $\varepsilon > 0$, there exists N such that for any $n > N$,

$$\theta \left(1 - \frac{\sigma^2}{v_{n-2}}\right) \|Z_{n-1}\| < \varepsilon.$$

Finally,

$$\|X_{n+h} - \hat{X}_{n+h} - Z_{n+h}\| \le \varepsilon \frac{1}{1-\theta} + \theta^h \left\| X_n - \hat{X}_n - Z_n \right\| \to 0 \quad (h \to \infty).$$

HW8 Problem 1 (2.12) By the uniqueness of the projection. (See p. 53.)

HW8 Problem 3 (2.18) (a) Please refer to any textbook on functional analysis. (b) Stationarity guarantees that $\{X_t\}$ is a sequence in some Hilbert space (L^2), then apply (a).

HW8 Problem 4 (4.3) Note that when $h \neq 0$,

$$\frac{\sin ah}{h} = \frac{1}{2} \int_{-a}^{a} \mathrm{e}^{\mathrm{i}hv}\, \mathrm{d}v.$$

HW5 Problem 1 (7.5) Use formula (7.2.5) to compute the asymptotic covariance matrix of $\hat{\rho}(1)$, $\hat{\rho}(2)$ for an AR(1) process. What is the behaviour of the asymptotic correlation of $\hat{\rho}(1)$ and $\hat{\rho}(2)$ as $\phi \to \pm 1$?
Solution. First we have $\rho(h) = \phi^{|h|}$. When using (7.2.5) to calculate the asymptotic covariance, note that when $k = 1$ and $i = 2$ or $j = 2$, $\rho(-1) = \phi$ rather than ϕ^{-1}. (Some students made mistakes here.)

For the 2nd question, note that covariance and correlation are different concepts. (Some students made mistakes here.)

For the answer,

$$w_{1,1} = 1 - \phi^2, \quad , w_{2,2} = 1 + 2\phi^2 - 3\phi^4, \quad w_{1,2} = w_{2,1} = 2\phi - 2\phi^3.$$

And the asymptotic correlation is $2\phi/\sqrt{1 + 3\phi^2}$.

HW5 Problem 3 Suppose that $\{X_t\}$ is the AR(1) process,

$$X_t - \mu = \phi(X_{t-1} - \mu) + Z_t, \quad \{Z_t\} \sim \mathrm{IID}(0, \sigma^2),$$

where $|\phi| < 1$. Find constants $a_n > 0$ and b_n such that $\exp(\bar{X}_n)$ is $\mathrm{AN}(a_n, b_n)$.

Solution. By Theorem 7.1.2, $\bar{X}_n$ is asymptotically normal, then using Proposition 6.4.1 to obtain a_n and b_n.

This might seem strange at first sight, since when X is Gaussian, then $\exp(X)$ is nonnegative, which cannot be Gaussian. But the result above is true when it comes to asymptotic normality as the covariance converges to 0.

A Common Mistake in HW4 If $X_n \xrightarrow{\mathrm{d}} X$ and $Y_n \xrightarrow{\mathrm{d}} Y$, it is not necessary that $(X_n, Y_n) \xrightarrow{\mathrm{d}} (X, Y)$. See Proposition 6.3.1 (Carmer-Wold Device) and Remark 1 on page 205 of the textbook for further details.

For example, HW4 Problem 4 (6.24) (b), noting that $X_t = (Z_t Z_{t+1}, \ldots, Z_t Z_{t+h})'$ is h-dependent, one can employ Carmer-Wold Device and Theorem 6.4.2 (CLT for strictly stationary h-dependent sequences) to complete the proof.

HW4 Problem 1 (6.12) Suppose that X_n is $\mathrm{AN}(\mu, \sigma_n^2)$ where $\sigma_n^2 \to 0$. Show that $X_n \xrightarrow{\mathbb{P}} \mu$.

Proof. I shall present slightly different proofs below.

Proof 1, by Zhenhang Bao.

首先不妨假设$\frac{X_n-\mu}{\sigma_n}$依分布收敛到Y，且Y服从标准正态分布。下面用反证法来证明X_n依概率收敛到μ。

若不然，则存在一个ϵ_0和μ_0严格大于0，以及一列X_{n_k}（下面不妨仍记为X_n）使得

$$\mathbb{P}\left(\left|X_n - \mu\right| > \epsilon_0\right) > \mu_0$$

注意到$\frac{X_n-\mu}{\sigma_n}$依分布收敛到Y，即有对任意的$t > 0$，成立

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{X_n - \mu}{\sigma_n}\right| < t\right) = \mathbb{P}\left(|Y| < t\right)$$

下面取t足够大，使得$\mathbb{P}(|Y| < t) > 1 - \frac{\mu_0}{2}$，而此时上面等式的左边，注意到$\sigma_n \to 0$，结合我们的假设可知，存在一个$N$，任意的$n > N$，有

$$\mathbb{P}\left(\left|\frac{X_n - \mu}{\sigma_n}\right| < t\right) \leq 1 - \mu_0$$

而这与依分布收敛矛盾。于是假设不成立，即X_n依概率收敛到μ。

Proof 2, by Yinsheng Chai.

Proof. Since $Z_n := \sigma_n^{-1}(X_n - \mu) \xrightarrow{d} Z$ where $Z \sim N(0,1)$, we may claim that $Z_n = O_{\mathbb{P}}(1)$ as $n \to \infty$. In fact, because a finite set of random variables (of course only one) is bounded in probability, $\forall \varepsilon > 0, \exists\, \delta_0(\varepsilon) > 0$, such that $\mathbb{P}(|Z| > \delta_0(\varepsilon)) < \varepsilon$. Besides,

$$\lim_{n\to\infty} \mathbb{P}(|Z_n| > \delta_0(\varepsilon)) = \mathbb{P}(|Z| > \delta_0(\varepsilon)),$$

hence $\exists\, N > 0$, such that $\forall n > N, \mathbb{P}(|Z_n| > \delta_0(\varepsilon)) < \varepsilon$. For $n \le N, \{X_1, \cdots, X_N\}$ is finite and thus $\exists\, \delta_1(\varepsilon) > 0$ such that $\mathbb{P}(|Z| > \delta_1(\varepsilon)) < \varepsilon$. Let $\delta(\varepsilon) = \max\{\delta_0(\varepsilon), \delta_1(\varepsilon)\}$ and we have

$$\mathbb{P}(|Z_n| > \delta(\varepsilon)) < \varepsilon, \quad \forall n,$$

which means $Z_n = O_{\mathbb{P}}(1)$, i.e.

$$\forall\, \varepsilon > 0,\, \exists\, \delta(\varepsilon) > 0,\, \mathbb{P}(|\sigma_n^{-1}(X_n - \mu)| > \delta(\varepsilon)) < \varepsilon,$$

which is equivalent to $\mathbb{P}(|X_n - \mu| > a_n(\varepsilon)) < \varepsilon$, where $a_n = \sigma_n \delta(\varepsilon) \downarrow 0$. Therefore, $X_n - \mu = o_{\mathbb{P}}(1)$, i.e. $X_n \xrightarrow{\mathbb{P}} \mu$. □

Proof 3, similar with Proof 2.

Since $Z_n = \sigma_n^{-1}(X_n - \mu) \xrightarrow{\mathrm{d}} Z$ where $Z \sim \mathrm{N}(0,1)$, we can know that $Z_n = O_p(1)$.[1] In fact, by Skorokhod's representation theorem, there exist random variables $\{Y_n\}$ and Y on a common probability space such that $Y \stackrel{\mathrm{d}}{=} Z$, $Y_n \stackrel{\mathrm{d}}{=} Z_n$ for any n, and such that $Y_n \to Y$ a.s.. Therefore $Y_n = Y + o_p(1) = O_p(1) + o_p(1) = O_p(1)$ which implies $Z_n = O_p(1)$. Hence for any $\varepsilon > 0$ and $M > 0$, there exist $\delta > 0$ and $N = N(\varepsilon, M, \delta) > 0$ such that for any $n > N$, we have $\sigma_n < M/\delta$ and that

$$\mathbb{P}(|X_n - \mu| > M) \le \mathbb{P}(|X_n - \mu| > \delta\sigma_n) < \varepsilon,$$

which follows the result. □

HW4 Problem 2 (6.14) If $X_n = (X_{n1}, \ldots, X_{nm})' \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Sigma)$ and $\Sigma_n \xrightarrow{\mathbb{P}} \Sigma$ where Σ is non-singular, show that $X_n' \Sigma_n^{-1} X_n \xrightarrow{\mathrm{d}} \chi^2(m)$.

A Common Mistake Most of the students take Σ_n as the covariance matrix of X_n, even though the exerice only suggests that $\Sigma_n \xrightarrow{\mathbb{P}} \Sigma$ which implies that Σ_n should be an arbitrary random matrix which is never guranteed to be symmetric or invertible. From another perspective, if Σ_n is the covariance matrix, then the condition should be written as $\Sigma_n \to \Sigma$ since they are not random.

In my opinion, the meaning of this exercise is that, when we have a consistent estimate Σ_n of the covariance Σ, we can know that $X_n' \Sigma_n^{-1} X_n \xrightarrow{\mathrm{d}} \chi^2(m)$ which might be used to construct confidence intervals.

The proof below is a little cumbersome and might be beyond the syllabus, so I think it's okay that you only work out the problem when Σ_n is the covariance matrix, and skip the proof for the general case.

[1]For your information, this is called, in measure theory, Prokhorov's theorem.

Proof. If you assume Σ_n is the covariance matrix, then see the proof by Zhenhang Bao.

先不妨假设题目中涉及的所有的矩阵均非奇异，这是因为$\Sigma_n \to \Sigma$，而Σ非异。于是由协方差矩阵的性质可知所有的矩阵都是正定阵，从而存在正定阵$\{B_n\}$以及B，满足

$$\Sigma_n = B_n^2 \qquad \Sigma = B^2 \qquad B_n \to B$$

（以上矩阵的收敛均可理解为依概率收敛） 而X_n依分布收敛到Y，Y服从均值为0的高斯分布。由依分布收敛的性质可知，$B_n^{-1}X_n^T$依分布收敛到$B^{-1}Y^T$。从而，$X_n^T\Sigma_n^{-1}X_n$依分布收敛到$Y^T\Sigma^{-1}Y$。最后再注意到$Y^T\Sigma^{-1}Y$服从$\chi^2(m)$的分布，即得证。

If Σ_n is not assumed to be the covariance matrix, the proof below is given by Yinsheng Chai.

Proof. Suppose $\boldsymbol{X}_0 \sim N(0, \Sigma)$, $\chi \sim \chi^2(m)$ and $\phi_X(t)$ is the characteristic function of X. Then

$$|\phi_{\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_n}(t) - \phi_\chi(t)| \le I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= |\phi_{\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_n}(t) - \phi_{\boldsymbol{X}_0^\top \Sigma_n^{-1} \boldsymbol{X}_0}(t)|, \\ I_2 &= |\phi_{\boldsymbol{X}_0^\top \Sigma_n^{-1} \boldsymbol{X}_0}(t) - \phi_{\boldsymbol{X}_0^\top \Sigma^{-1} \boldsymbol{X}_0}(t)|, \\ I_3 &= |\phi_{\boldsymbol{X}_0^\top \Sigma^{-1} \boldsymbol{X}_0}(t) - \phi_\chi(t)|. \end{aligned}$$

We claim that $I_1, I_2, I_3 \to 0 (n \to \infty)$. In fact,

$$I_1 \le |\phi_{\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_n}(t) - \phi_{\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_0}(t)| + |\phi_{\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_0}(t) - \phi_{\boldsymbol{X}_0^\top \Sigma_n^{-1} \boldsymbol{X}_0}(t)| \to 0$$

from the Cramer-Wold device, $I_2 \to 0$ because

$$|\Sigma_n - \Sigma| = |\Sigma| \cdot |\Sigma_n| \cdot |\Sigma^{-1} - \Sigma_n^{-1}|$$

and $\Sigma_n^{-1}\boldsymbol{X}_0 \xrightarrow{d} \Sigma^{-1}\boldsymbol{X}_0$, $\boldsymbol{X}_0^\top \Sigma_n^{-1} \boldsymbol{X}_0 \xrightarrow{d} \boldsymbol{X}_0^\top \Sigma^{-1} \boldsymbol{X}_0$. $I_3 \to 0$ because

$$\boldsymbol{X}_0^\top \Sigma^{-1} \boldsymbol{X}_0 = (\Sigma^{-\frac{1}{2}} \boldsymbol{X}_0)^\top (\Sigma^{-\frac{1}{2}} \boldsymbol{X}_0)$$

and $\Sigma^{-\frac{1}{2}} \boldsymbol{X}_0 \sim N(\boldsymbol{0}, \boldsymbol{I}_m)$, $\boldsymbol{X}_0^\top \Sigma^{-1} \boldsymbol{X}_0 \sim \chi^2(m)$. Thus $|\phi_{\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_n}(t) - \phi_\chi(t)| \to 0$ as $n \to \infty$, i.e. $\boldsymbol{X}_n^\top \Sigma_n^{-1} \boldsymbol{X}_n \xrightarrow{d} \chi^2(m)$. □

Professor Yuwei Zhao provided the following proof.

Since $\det(\Sigma)$ is a continuous function of Σ. If Σ is non-singular, there exists a large number $N > 0$ such that $|\det(\Sigma_n)| > 0$, which implies that the inverse of Σ_n exists. Since Σ is a non-singular covariance matrix, there exists a matrix T such that $T'T = \Sigma$ with $|\det(T)| > 0$ and $T^{-1} = T$. Notice that $X_n'\Sigma_n^{-1}X_n \in \mathbb{R}$ and we have the trace of $X_n'\Sigma_n^{-1}X_n$ i.e. $\operatorname{tr}(X_n'\Sigma_n^{-1}X_n)$, is exactly $X_n'\Sigma_n^{-1}X_n$. We also have

$$\operatorname{tr}(X_n'\Sigma_n^{-1}X_n) = \operatorname{tr}(X_n'T'T\Sigma_n^{-1}T'TX_n) = \operatorname{tr}\Big(\big(T\Sigma_n^{-1}T'\big)\big\{TX_nX_n'T'\big\}\Big).$$

The first product $T\Sigma_n^{-1}T'$ converge in probability to the identity matrix, and $TX_nX_n'T'$ converge in distribution to the m-dim vector of independent $\chi^2(1)$ distributed random variables.

Originally the proof I wrote is as follows. In general, when Σ_n is singular, we may temporarily take Σ_n^{-1} as the MoorePenrose inverse. Since Σ is non-singular, given some matrix norm $\|\cdot\|$, there exists a constant $r > 0$ such that for any Σ_n satisfying $\|\Sigma_n - \Sigma\| \leq r$, Σ_n is non-singular. (The proof can be found, for example, in Section 5.2 of this note.) Write

$$X_n'\Sigma_n^{-1}X_n = X_n'\Sigma_n^{-1}X_n 1_{\{\|\Sigma_n - \Sigma\| \leq r\}} + X_n'\Sigma_n^{-1}X_n 1_{\{\|\Sigma_n - \Sigma\| > r\}}.$$

Noting that $\Sigma_n \xrightarrow{\mathbb{P}} \Sigma$ which implies the second term in RHS is $o_p(1)$, by Slutsky's theorem, it suffices to show

$$X_n'\Sigma_n^{-1}X_n 1_{\{\|\Sigma_n - \Sigma\| \leq r\}} \xrightarrow{\mathrm{d}} X'\Sigma^{-1}X,$$

where $X \sim \mathrm{N}(0, \Sigma)$. In order to apply continuous mapping theorem, we need to show

$$(\Sigma_n^{-1} 1_{\{\|\Sigma_n - \Sigma\| \leq r\}}, X_n) \xrightarrow{\mathrm{d}} (\Sigma^{-1}, X).$$

In fact, by portmanteau lemma and the argument here, the result holds. (You may circumvent some difficulty by revising this proof and imitating Professor Zhao's argument.) □

HW3 Problem 3 (6.6) Let $\{X_t\}$ be a stationary process with mean zero and an absolutely summable autocovariance function $\gamma(\cdot)$ such that $\sum_{h=-\infty}^{\infty} \gamma(h) = 0$. Show that $n \operatorname{Var}(\bar{X}_n) \to 0$ and hence that $n^{1/2}\bar{X}_n \xrightarrow{\mathbb{P}} 0$.

Proof. The autocovariance function is absolutely summable, so for any $\varepsilon > 0$, there exists $K > 0$ such that for any $H > K$, we have

$$\sum_{|h|>H} |\gamma(h)| < \varepsilon.$$

And since $\sum_h \gamma(h) = 0$, there exists $M > 0$ such that for any $n > M$,

$$\left|\sum_{|h|<n} \gamma(h)\right| < \varepsilon.$$

Moreover, given H, there exists $N > 0$ such that for any $n > N$,

$$\frac{1}{n}\left|\sum_{|h|\le H} |h|\gamma(h)\right| < \varepsilon.$$

Finally, noting that $\{X_t\}$ is stationary with zero mean, we have for $n > \max\{H+1, N, M\}$,

$$\begin{aligned}
|n\,\mathrm{Var}(\bar{X}_n)| &= \left|\sum_{h=-(n-1)}^{n-1}\left(1-\frac{|h|}{n}\right)\gamma(h)\right| \\
&\le \left|\sum_{|h|<n}\gamma(h)\right| + \frac{1}{n}\left|\sum_{|h|\le H}|h|\gamma(h)\right| + \sum_{|h|=H+1}^{n-1}\frac{|h|}{n}|\gamma(h)| \\
&< \varepsilon + \varepsilon + \sum_{|h|>H}|\gamma(h)| < 3\varepsilon.
\end{aligned}$$

It follows that $n^{1/2}\bar{X}_n \to 0$ in L^2 which implies convergence in probability. □

See another proof using dominated convergence theorem in Theorem 7.1.1.