ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Μεταπτυχιακό πρόγραμμα ειδίκευσης και συμπληρωματικής ειδίκευσης στη Στατιστική

25 Φεβρουαρίου 2002

ΕΞΕΤΑΣΕΙΣ -Μέρος 2^{o}

Μάθημα : "Προχωρημένες χρονολογικές σειρές"
Διδάσκοντες: Ε. Ιωαννίδης

1st Question

(3,5 points) Consider the problem of the estimation of the spectrum of a stationary process given observations $X_1,\ldots X_T$. Assume you are using as estimator a kernel-smoothed periodogramm with some bandwidth.

- ✓ Explain the problematic of choosing an appropriate bandwidth: why not just taking a very small or a very big one?

If you define a theoretically "optimal" bandwidth as the one which minimizes the

MSE= $\int\left[\frac{E\tilde{f}(\lambda)-f(\lambda)}{f(\lambda)}\right]^2 d\lambda+\int\left[\frac{\tilde{f}(\lambda)-Ef(\lambda)}{f(\lambda)}\right]^2 d\lambda$ = bias^2+variance

- ✓ Would this "optimal" bandwidth increase or decrease when the number of available observations is increased?
- ✓ If you get other observations from the same time series model would the "optimal bandwidth" change?
- ✓ If another researcher has the same number of observations from another time series model, which has a smoother spectrum than yours would his "optimal bandwidth" be different from yours?

(You do not need to remember and give convergence rates etc. A description in words showing that you have understood the problematic is sufficient).

2nd Question

(1,5 points) Assume somebody knows what leakage is and the untapered periodogramm suffers from leakage. Try to explain him in a few sentences (without technical details and proofs) that the estimator of the spectrum obtained by fitting an AR(p) density via the Yule-Walker equations suffers from the same problem. Just give him a sketch of the main ideas. Two sentences are enough.

3rd Question

(1,5 points) Assume you observe data $Z_1,\ldots,Z_T$ of which you know that they are the sum of two uncorrelated stationary time series X_t and Y_t , i.e. $Z_t = X_t + Y_t$. Show that the expectation of the periodogramm of your data $Z_1,\ldots Z_T$ is

$$EI_Z(\lambda) = EI_X(\lambda) + EI_Y(\lambda)$$

Hint: for a,b$\in$**C** you have $|a+b|^2 = |a|^2 + |b|^2 + 2\,\mathrm{Re}(a\bar{b})$

4th Question

(2,5 points) Assume you have observations $X_1,\ldots X_T$ from a gaussian series with smooth spectrum. In order to estimate the spectrum you calculate the periodogramm $I(\lambda)$ and the kernel smoothed periodogramm $\tilde{f}_b(\lambda)$ using bandwidth b. Assuming $b\to 0$ and $bT\to\infty$ and assuming that λ_T and μ_T are two sequences of frequencies in $(0,\pi)$ examine whether $Cov[I(\lambda_T), I(\mu_T)]$ and $Cov\left[\sqrt{Tb}\tilde{f}_b(\lambda_T), \sqrt{Tb}\tilde{f}_b(\mu_T)\right]$ **converge or not to 0** when $T\to\infty$, in each of the following cases:

- $\lambda_T = \lambda$ and $\mu_T = \mu$, $\mu\neq\lambda$
- λ_T - $\mu_T = 1/T$
- λ_T - $\mu_T = b$ (remember b depends on T)

Does $Cov[I(\lambda_T), I(\mu_T)]$ depend on tapering in each of the above cases?

Do $Var[I(\lambda)]$ and $Var\left[\sqrt{Tb}\tilde{f}_b(\lambda)\right]$ for some fixed λ depend on tapering?

5th Question

(3,5 points) Below you find 6 figures. Each figure is based on data $X_1,\ldots,X_{500}$ (series A) which is the same for all figures. You find the following figures (in log scale):

I. [A, non-tap]: the untapered unsmoothed and an untapered smoothed periodogram of the data
II. [A, tap]: the 30% tapered unsmoothed and a 30% tapered smoothed periodogram of the data.
III. [Diff(A), non-tap]: similar as in I., but calculated on the basis of the **4-lag** differences of the data $\Delta_4 X_t = X_t - X_{t-4}$ (instead of the data X_t themselves).
IV. [Diff(A), non-tap]: similar as in II., but calculated on the basis of the **4-lag** differences of the data $\Delta_4 X_t = X_t - X_{t-4}$ (instead of the data X_t themselves).
V. [A, tap/non-tap]: the untapered and tapered unsmoothed periodogram of the data
VI. [Diff(A), tap/non-tap]: similar as in V., but calculated on the basis of the **4-lag** differences of the data $\Delta_4 X_t = X_t - X_{t-4}$ (instead of the data X_t themselves).

Discuss your suspicions about the model which generated the data. Is it an ARMA model (if yes, what do you suspect about the roots of it's polynomials), is it a random wave, is it process with roots on the Unit circle, is it sums of the above, etc? (You may have more than one guesses about the model). In particular discuss and explain the following points on the basis of your model suspicion:

a) Compare the behavior of the tapered and non-tapered estimators (and explain their differences) at frequencies $\lambda=0$ and $\lambda=\pi/2$ (figures I and II)
b) Compare the behavior of the periodogramms $\lambda = 0, \pi/2$ with its behavior at the frequency $\lambda = 3\pi/4$. What kind of a peak do we have here?

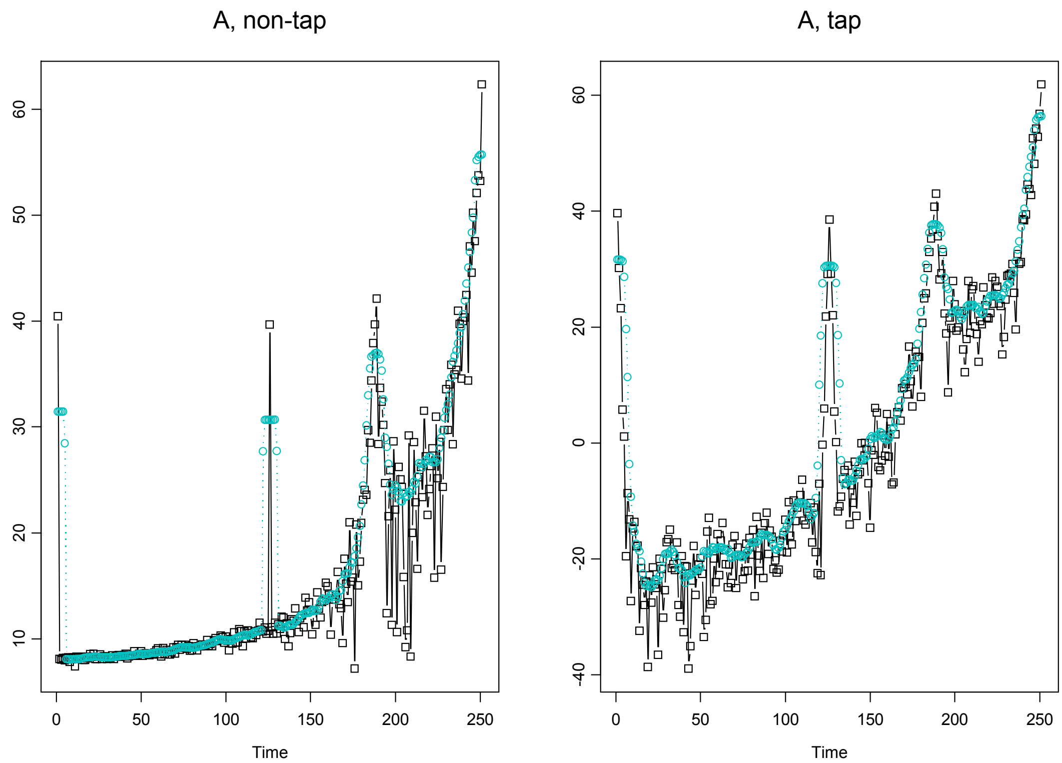
c) When looking at the filtered series $\Delta_4 X_t$ (figures III and IV) for these two frequencies ($\lambda=0$ and $\lambda=\pi/2$) the peaks not only disappeared (in both, the tapered and the non-tapered case), but –moreover-- in the tapered case there seem to be MA roots at these frequencies. Not so in the non-tapered case. Which of the two (tapered/ non-tapered) does better estimate the true spectrum of the filtered process? (see also figure VI.)

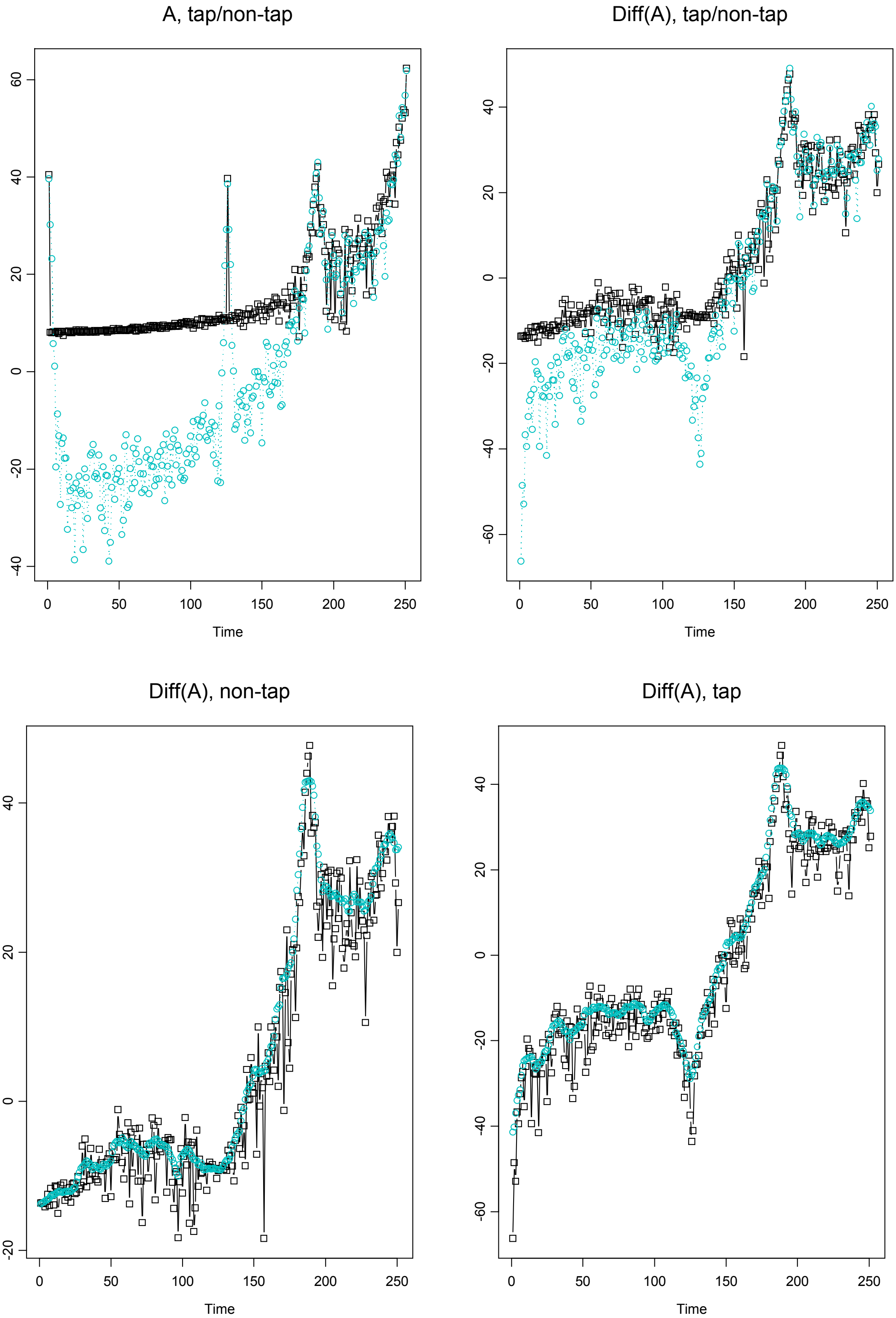
d) What kind of peak do we have at $\lambda=\pi$? Here as well the peak disappears for the filtered series. Why?

e) In figure V the tapered periodogramm is much smaller than the non-tapered one particularly for frequencies smaller than $\pi/2$. Moreover it fluctuates rapidly up and down. Do you have an explanation for these two points? Which of the two (tapered/non-tapered) does better estimate the true spectrum of the process?

Here are some theoretical points, which you may assume as known:

- ✓ If X_t is a random wave of frequency 0, π or $\pi/2$, then $\Delta_4 X_t = 0$.
- ✓ The transfer polynomial of the filter $\Delta_4 X_t$ components equals $\psi(z) = \left(z - e^{i0}\right)\left(z - e^{i\frac{\pi}{2}}\right)\left(z - e^{i\pi}\right)\left(z - e^{i-\frac{\pi}{2}}\right)$
- ✓ If X_t is a series with spectrum $f_X(\lambda)$ and $Y_t = \Delta_4 X_t$, we have $f_Y(\lambda) = f_X(\lambda) * \left|\psi\left(e^{i\lambda}\right)\right|^2$, where ψ as above.
- ✓ The statement of question 3 above




A, tap/non-tap
Diff(A), tap/non-tap
Diff(A), non-tap
Diff(A), tap
Time
Time
Time
Time