Lund University
Centre for Mathematical Sciences
Mathematical Statistics

Stationary stochastic processes
for scientists and engineers
October 21, 2013

Computer exercise 1 in Stationary stochastic processes

The purpose of this exercise is to study the estimation of the expected value, covariance function and spectral density for some process realizations, both simulated and from real measurements.

1 Preparations

- Download and read the folder `Matlabhints`.
- Carefully read through the entire computer exercise.
- Study Chapters 2 & 4, and also 9.1 - 9.2.3 in the textbook thoroughly.
- Answer the exercises in the question dictionary below; (use additional sheets, if needed); you are expected to be able to explain your answers during the exercise.

1.1 Question dictionary

1. Consider a Gaussian process with frequencies $f_k = \{5, 10\}$ and variances $\sigma_k^2 = \{2, 2\}$, formed as

$$X(t) = \sum_{k=1}^{2} A_k \cos(2\pi f_k t + \phi_k)$$

where $\phi_k \sim U(0, 2\pi)$ and $A_k \sim \text{Rayleigh}(\sigma_k^2)$ are independent stochastic variables, i.e., the probability function is

$$f_R(x; \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad x \geq 0,$$

where $\mathsf{E}[R] = \sigma\sqrt{\frac{\pi}{2}}$ and $\mathsf{V}[R] = \frac{4-\pi}{2}\sigma^2$.

What is the covariance function of the process? What is its spectral density? *(hint: Solve Exercise 4.2 and read Chapter 2.3.4 in the textbook.)*

Your answer:

2. What is a consistent estimate? What is a linear ergodic process? Completely ergodic?

   Your answer:

3. Find an expression for the variance of $\widehat{m}$, $\mathsf{V}[\widehat{m}]$, where $\widehat{m}$ is an estimate of the expected value, m, of a stationary process, based on the observations $x(1), x(2), \dots, x(n)$.

   Your answer:

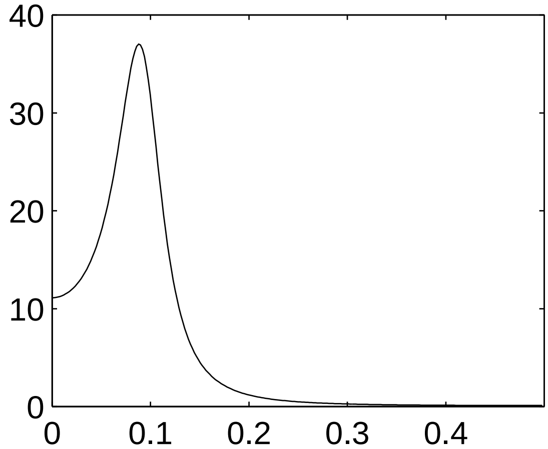
4. Suggest an estimate of the covariance function, $r_x(\tau)$.

   Your answer:

5. Write down two expressions for how to calculate the periodogram of a time series.

   Your answer:

6. Consider a process with the power spectral density in the following figure.



What would the periodogram of a typical realisation of this process look like? Can you remark the variance of the estimate for different frequencies? How will the estimate be affected by a limited number of data samples?

Your answer:

2 To start the MATLAB exercise

- Download the m-files from the folder `ssproutines` and the data from `sspdata` and make them accessible from your MATLAB window.
- It is often convenient to gather MATLAB commands in a program script; the MATLAB editor can be used for this purpose. This is started with the command `edit` in the MATLAB window.

3 Estimation of the expected value, covariance function and spectral density

3.1 Estimation of expected value

Load the file `data1.mat` using the command `load data1`. The file contains a realization of 100 measurements of white noise with the unknown expected value m (save your MATLAB code).

1. Plot the sequence

```
>> plot(data1.x)
```

Q. Does it seems like this process has a zero mean?

2. Estimate the mean with

```
>> mean(data1.x)
```

Q. What value did you get? Is this a reliable estimate? (Hint: Is the mean estimator implemented by MATLAB, unbiased and consistent for white noise?)

3. Assume that we know that the measurements are independent.

Q. Derive an expression for $\mathsf{V}[\widehat{m}]$ ***exploiting independence.***

Q. Write an expression for a 95% confidence interval of m***.***

Use the expressions you derived to estimate $\mathsf{V}[\widehat{m}]$ and give a 95 % confidence interval of m (hint: use the MATLAB function `std`).

Q. What values did you get?

3.2 Estimation of the covariance function

Load the file covProc which contains a realisation of an unknown process. Plot y_t against y_{t-k} for different values of k, e.g. start with the commands.

```
>> k=1
>> plot(covProc(1:end-k),covProc(k+1:end),'.')
```

and change to k=2 and k=3 and examine the differences between the plots.

Q. What do these plots "scatter plots" represent?

Finally, estimate the covariance function with the WAFO routine dat2cov, which takes as input a matrix with times in first column and data in second column; see the help text.

```
>>  t=(1:length(covProc))';
>>  x=[t covProc];
>>  r=dat2cov(x,20,2)
>>  r.R(1:4)         % r.R(1) is an estimate of the variance
>>  v=std(covProc)^2 % v is another estimate
```

Q. What values did you get? How do these compare with the plots above? And why is there a small difference between the two estimates of the variance?

3.3 Spectrum estimate of a sum of harmonics

By means of spekgui

Amongst the exercise files in ssproutines there is a function spekgui that you can use to estimate the covariance functions and spectral densities for some processes. You can find the help text for spekgui in the Matlabhints .

Start the function with the command

```
>> spekgui
```

The function `simsum2` simulates a Gaussian process with a discrete spectral density with just two frequencies, $f_k = \{5, 10\}$, and the variances $\sigma_k^2 = \{2, 2\}$. A new realization is simulated each time you call the function. The current realization is saved in the MATLAB-variable `data`. Import this to `spekgui` and analyze it using the periodogram.

The function `simsum` simulates a Gaussian sum of random cosines with arbitrary frequencies.

Q. Draw a rough sketch of the spectral density estimate obtained using the periodogram. Do the peaks have equal heights?

To explain this effect, study the variance of the periodogram estimates by simulating new realizations using `simsum` and import into `spekgui`. Investigate how the spectral estimates change.

Q. Based on the investigations above, what can you say about the variance of the periodogram estimate? (Recall preparation question 6.)

By means of `dat2spec`

The WAFO routine `dat2spec` is included in `ssproutines`. It works on a data matrix with time in first column and data in second; see the help text. In anticipation of Computer exercise 3 you can run the following command:

```
>>  S = dat2spec(x,50,[],1,[],[],[],'f')
```

The argument (50) in the call is a lag-parameter that stabilizes the spectrum estimate; see further Computer exercise 3 and Chapter 9 in the text book. `dat2spec` gives confidence limits for the true spectrum.

4 Student in a symphony orchestra

4.1 Keynotes and overtones

The sound from most acoustic instruments consist of a fundamental frequency, often termed a keynote, and some overtones. The phases of the overtones typically depend on the instrument and are partly correlated with the swinging of the keynote. This, together with the relation between the power of the overtones, produces the perceived sound of the instrument.

If the keynote has frequency f_0, what are the frequencies of the overtones? This will depend on the type of instrument, but for string instruments, the overtones can be well represented[1] as

$$f_k = k f_0,$$

with $k = 1, 2, \ldots$. Load the files `cello.mat` and `trombone.mat`. These files contain the signals of a tone played by a cellist and a trombonist at The Academic Orchestra in Lund[2]. You can listen to the tones by using the command `soundsc(cello.x)`.

Import the data (`cello` or `trombone`) into `spekgui` and estimate the spectral densities using some appropriate method (e.g., using Welch's method with 2-3 overlapping windowed sequences; this is given in `spekgui` as a parameter). Examine the result using both a linear and logarithmic scale.

1. ***What are the frequencies of the cello and trombone keynotes?***

2. ***Do the overtones appear at integer multiples of the keynotes?***

3. ***How many overtones can you see for the cello and the trombone sounds?*** (hint: use the logarithmic scale).

[1] It is worth noting that the stiffness of the string will actually produce some frequency offsets such that the overtones will not be exact multiples of the fundamental frequency. A more precise model of the overtones taking the string stiffness into account can be found as

$$f_k = k f_0 \sqrt{1 + Bk^2}$$

where B is a positive stiffness parameter.

[2] Founded 1745, the orchestra still today plays at academic seremonies and give regular symphony concerts.

4. The sounds were recorded with a really bad tape-recorder, and contain a lot of noise.

   Q. Can you see a strong noise peak at a particular frequency? (Hint: The tape recorder was not battery charged.)

5. The sounds from two different instruments that are played simultaneously can be assumed to be independent.

   Q. Use this to find out how many cellists you need to drown out the sound from one trombonist. [3] (Hint: As the process has mean zero, the power is equal to the variance of the process. Also, the variances will sum as independent variables.)

   It is worth recalling that the string section is much bigger than the brass section in a symphony orchestra.

4.2 Aliasing

Start with studying the spectrum of the cello using `spekgui`. Then, create a down-sampled realisation by extracting every second sample from the original signal (save your MATLAB code)

```
>> n=2;
>> cello2.x=cello.x(1:n:end);
>> cello2.dt=cello.dt*n;
```

1. Import `cello2` to `spekgui` and study the spectrum.

   Q. Has the spectrum changed? How has the spectrum range changed? At what frequencies do aliased peaks (if any) appear?

[3] Both the musicians nearly played at maximum volume from their instruments at the recording time.

2. Examine the `trombone` process in the same way. Try some different values of `n` for different down-sampling.

   Q. How much slower must this signal be sampled to give an aliasing in the spectrum?

   A correct down-sampling, without aliasing, is obtained if the signal is low-pass filtered before the down-sampling. This can be done using the MATLAB-function `decimate` in the signal processing toolbox.

```
>>  cello2.x=decimate(cello.x,2);
>>  % Works only with toolbox 'signal'
>>  cello2.dt=cello.dt*2;
```

   Q. Are there still any aliased peaks in the spectrum?