Stochastic Tools Homework 2

Instructor: William K. George, 2013-01-29

References: Chapters 2, 8 and Appendix B of WKG's 'Lectures in Turbulence for the 21st Century'

Access the data set described at the end of this document and carry out the analysis described below. This data was acquired by Murat Tutkun (and the Wallturb team) using a hot-wire anemometer in a boundary layer of the LML wind tunnel.

Perform the following statistical analyses:

Mean and variance

1. Estimate the mean and variance using all blocks of data.
2. Now subdivide the data into four parts and compute the mean and variance of each part separately. These probably will not be exactly the same, but you will be asked to discuss this below.

Correlation Analysis

3. Estimate the two-time correlation directly from the data by averaging the products of two velocities at different times.
4. Estimate the statistical error for several time lags and show them on your plot.
5. Now estimate the integral time scale by integrating under the autocorrelation function.
6. Now fit the oscullating parabola to your autocorrelation and estimate from it the Taylor microscale.
7. Does this look like it satisfies the criteria for a stationary random process? Why or why not?

Variability of Mean and Variance

8. You have all the information to compute the variability for each of your estimators (it will be different for the entire record and the shorter versions), and use it to place error bounds on your results. The means and variances are all likely to be different.

9. Use your variability estimates to decide whether these numbers are really different (meaning the mean and variances changed during the run), or whether you can account for the difference by just the statistical errors arising from the finite length of your record. (Hint: In the variability estimate you have to decide whether to use sqrt (1/N) for the number of independent samples, or sqrt [(2 Int/T)] where T is the block length and Int is the integral scale. Explain the difference, and then decide which to use, explaining your choice.

10. Now recompute the statistics above, but this time using the 'optimal' amount of data; i.e., skip all the data you don't need. Compare your answers to those you obtained above and explain any differences. Which estimators should you place the most confidence in?

Directions for accessing and reading data (from Matthieu)

```
The data for the homework can be downloaded from this address :

http://lml.univ-lille1.fr/lml/perso/marquillie/downloads/velocity_data.zip

(to unzip the data under linux the command is : unzip velocity_data.zip)

Information about the data given by Murat :

Boundary layer hot wire data at y+ = 445 in ASCII format.
The data is stored in an array of 131072 x 100.
The columns are separated by a tab deliminator.

Freestream velocity = 10 m/s
Number of samples = 131072 = 217
Sampling frequency = 30000 Hz
Record length = 4.369 sec
Number of blocks = 100


The fortran program below (from Matthieu) reads all the data and
computes the mean value.


program test
  implicit none
  integer,parameter :: n=13107200
  real(8) :: x(n)

  open(10,file='velocity_data.out')
  read(10,*)x
```

```
  print*,sum(x)/n

end program test
```