ExerciseSimulatedData

Before you begin !!!

- Copy **cp dcm.dat origdcm.dat**
In this way, you can always
- Copy **cp origdcm.dat dcm.dat**
back to its original form after this exercise. If you lose the original version of **dcm.dat**, just download it again from the home-page after this exercise.

Introduction

Line 2 in control file **dcm.dat** is parameter **RealData**.

Main point of this exercise is that
- If **RealData= 1**, DCM analyses **real data**.
- If **RealData $\neq$ 1**, DCM creates and analyses **simulated data**.

Simulation213.dat

Here, is one example of what happens, if **RealData=0$\neq$1**. Try to repeat this **Simulation213.dat** procedure in your own directory **/DCMtest**.
- Copy **cp dcm.dat Simulation213.dat**
- Edit **emacs Simulation213.dat** to the form

```
 1  = Tag          = Simulation213
 2  = RealData     = 0
 3  = file1        = TestData.dat
 4  = dummy        = -99.999
 5  = K1           = 2
 6  = K2           = 1
 7  = K3           = 3
 8  = nL           = 60
 9  = nS           = 30
10  = c            = 0.20
11  = TestStat     = 1
12  = PMIN         = 1.0
13  = PMAX         = 2.0
14  = Rounds       = 2
```

```
15  = NonLinear    = 1
16  = SimT         = 1
17  = SimN         = 500
18  = SimSN        = 100
19  = SimDT        = 4.0
20  = SimMany      = 0
21  = SimRounds    = 30
22  = SimDF        = 0.05
23  = SimDA        = 0.5
24  = PrintScreen  = 1.0
```

- Copy **cp Simulation213.dat dcm.dat**
- Run **python dcm.py**
- This produces **following files**

Simulation213SimulatedData.dat = Simulated data
Simulation213Model.dat = Model for simulated data
Simulation213AllBeta.dat = Free parameters of the above model
Simulation213Residuals.dat = Residuals of the above model
Simulation213Params.dat = Parameters of the above model

- This also produces **following figures**

Simulation213gsim.eps = Model used to create simulated data
Simulation213z.eps = Periodograms for simulated data
Simulation213gdet.eps = Model detected for simulated data
Simulation213fA.eps = Frequency and amplitude correlations in model detected for simulated data

- **Note: When you re-run python dcm.py, all the above-mentioned files and figures are always different in every simulation.**

- Two signal model $g(t, K_1, K_2, K_3)$ having orders K_1 = **K1=2**, K_2 = **K2=1** and K_3 = **K3=3** is used to create n = **SimN = 500** observations uniformly distributed between 0 and ΔT = **SimDT=4**. Simulated signal periods P_1 = **P1** and P_2 = **P2** are drawn from a uniform random distribution between $P_{\min}$ = **PMIN= 1** and $P_{\max}$ = **PMAX=2**. Signal

to noise ratio of simulated data is SN = **SimSN=100** (**Paper I**: Eq. 22). Simulated data are stored to file **Simulation213SimulatedData.dat**.

- **Simulated model**, and simulated data created with this model, are shown in **Simulation213gsim.eps** (**Figure 1**).
- **Detected model** is shown in **Simulation213gdet.eps** (**Figure 2**).
- The following information is printed to the screen:

Note! This information changes when python dcm.py is re-run.

```
('Simulated P = ', array([ 1.26657201,  1.48480063]))
('Long: ', 'Detected P=', array([ 1.26315789]), array([ 1.48148148]))
('Long: ', 'ZMIN=', array([ 1.12656647]))
('Short: ', 'Detected P=', array([ 1.26849894]), array([ 1.48148148]))
('Short: ', 'ZMIN=', array([ 1.02704989]))
('Final ZMIN =', 0.99294999345653645)
('Bootsrap round ', 1, ' out of ', 2)
('Bootsrap round ', 2, ' out of ', 2)
                                      n               500
                                     T1 4.3730114000e-03
                                     DT 3.9751782720e+00
                                     my -3.3039650387e-01
                                     sy 4.2614894054e-01
                                  sigma 1.1447361455e-02
                                     K1                2
                                     K2                1
                                     K3                3
                                      p               10
                                   PMIN 1.0000000000e+00
                                   PMAX 2.0000000000e+00
                                     nL               60
                                     nS               30
       CHI2 4.9297484475e+02   gives ZMIN 9.9294999346e-01
          R 1.0316146062e-01   gives ZMIN 1.4363945184e-02
  ..................................................
                  F1 7.8983852600e-01 +/- 1.0244350000e-04
                  P1 1.2660815687e+00 +/- 1.6420631906e-04
                  A1 5.5459082712e-01 +/- 8.7598018835e-05     SN      48.45
              T1MIN1 1.0819170003e-01 +/- 6.1948071179e-04
              T1MIN2              ... +/-              ...
              T1MAX1 7.4123248437e-01 +/- 5.3737755227e-04
              T1MAX2              ... +/-              ...
  ..................................................
                  F2 6.7321712000e-01 +/- 7.1364500000e-05
                  P2 1.4854048869e+00 +/- 1.5745450234e-04
                  A2 6.5823870983e-01 +/- 7.7924890192e-05     SN      57.50
              T2MIN1 2.3312536398e-01 +/- 7.1851878186e-04
              T2MIN2              ... +/-              ...
              T2MAX1 9.7582780744e-01 +/- 6.3979153069e-04
              T2MAX2              ... +/-              ...
```

```
.....................................................
                    i              BETA[i]
                    1         F1  7.89839e-01 1.02443e-04
                    2         F2  6.73217e-01 7.13645e-05
                    3        B11 -2.40961e-01 1.15641e-04
                    4        C11 -1.37226e-01 1.15073e-04
                    5        B21 -1.87362e-01 1.93139e-04
                    6        C21 -2.70584e-01 1.80924e-04
                    7         M0 -4.51501e-01 2.86890e-04
                    8         M1  4.21836e-01 6.13636e-04
                    9         M2  4.61402e-01 9.02195e-05
                   10         M3 -4.58386e-01 1.05082e-04
```

- Simulated periods are $P_1 = 1.26657201$ and $P_2 = 1.48480063$
- Detected periods are $P_1 = 1.26315789$ and $P_2 = 1.48148148$
- **Main conclusion** is that DCM can detect the two signal model used to create these simulated data. You can test this result by analysing other arbitrary simulated data samples. Just run **python dcm.py** several times, and check what happens. There are of course some limitations for detecting the correct periodicities, and those limitations will be discussed later.

ExerciseSimulatedData: Problem

Copy this file **SimulationXXX.dat** from homepage

```
 1  = Tag          = SimulationXXX
 2  = RealData     = YYY
 3  = file1        = TestData.dat
 4  = dummy        = -99.999
 5  = K1           = YYY
 6  = K2           = YYY
 7  = K3           = YYY
 8  = nL           = 60
 9  = nS           = 30
10  = c            = 0.20
11  = TestStat     = 1
12  = PMIN         = YYY
13  = PMAX         = YYY
14  = Rounds       = 2
15  = NonLinear    = 1
16  = SimT         = 1
17  = SimN         = YYY
18  = SimSN        = YYY
19  = SimDT        = YYY
20  = SimMany      = 0
21  = SimRounds    = 30
22  = SimDF        = 0.05
```

```
23  = SimDA        = 0.5
24  = PrintScreen  = 1.0
```

- File **SimulationXXX.dat** contains **nine** places marked **YYY**. Edit these nine places in **SimulationXXX.dat** to simulate data created from a two signal model $g(t)$ having orders $K_1 = 2$, $K_2 = 2$ and $K_3 = 1$. These $n = 100$ simulated observations should be uniformly distibuted between 0 and $\Delta T = 10$. Draw simulated signal periods P_1 and P_2 from a uniform random distribution between $P_{\min} = 1.5$ and $P_{\max} = 4.0$. Fix signal to noise ratio of simulated data to SN = 30.
- Copy **cp SimulationXXX.dat dcm.dat**
- Run **python dcm.py**
- Check your figures with commands
evince SimulationXXXgsim.eps &
evince SimulationXXXgdet.eps &
- Run **python dcm.py** several times to check how these figures change, and how good the period analysis results are.
- Send your **SimulationXXX.dat** to assistant.

Bonus point question: Increase number of simulated observations to $n = 300$ and signal to noise ratio to SN = 300. Do you notice any qualitative changes in the results? Use less than one hundred words in our answer.

Warning: When you do the above bonus point n and SN changes, you have to change contents of **SimulationXXX.dat**. Check that you send the **SimulationXXX.dat** version $n = 100$ and SN = 30 to the assistant.

Figure 1: **Simulation213gsim.eps**

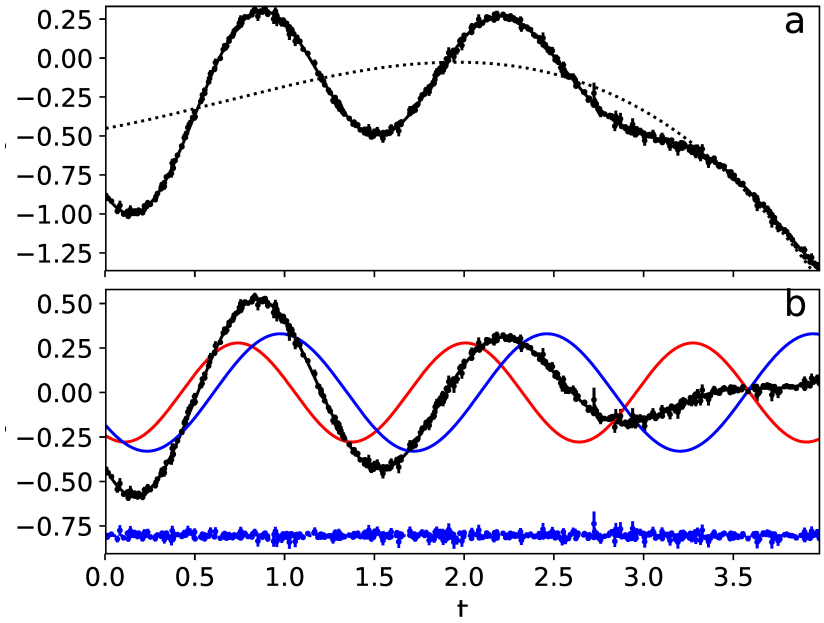


Figure 2: **Simulation213gdet.eps**

