Least Squares Problems : Identification of An Ecological System (4h)

Objective

Show how least squares formulation can be used to identify the parameter of a predator-prey ecological dynamic system. Emphasize the benefit of using the recursive Kalman filter solution when real time identification is needed under dynamically varying context.

Problem Statement

The dynamic of a tow-species (predator/prey) system is given by the following ordinary differential equations :

$$\frac{\dot{P}}{P} = r(1-\frac{P}{K}) - sQ \qquad (1)$$

$$\frac{\dot{Q}}{Q} = -\alpha + \beta P \qquad (2)$$

where :

- P is the size of the prey population
- Q is the size of the predator population
- r is the growing rate of prey in the absence of predator.
- K is the carrying capacity of the environment
- s is the death rate of the prey per predator unit.
- α is the death rate of the predator in the absence of prey
- β is the grow rate of the predator per unit prey

It is assumed that one disposes of approximate measurement of P and Q (estimation based on partial counting process). These measurements are acquired periodically with a sampling period denoted by $\tau > 0$. The measurement instants are denoted by t_k. More precisely :

$$t_k = k\tau \quad ; \quad k \in \{1, \dots, N\}$$

The aim of these working sessions is to use the measurements data given by :

$$Y|_1^N = \Big\{P(t_k), Q(t_k)\Big\}_{k=1}^N \qquad (3)$$

in order to derive a faithful estimation of the model parameters, namely : r, K, s, α and β. This model can be used to regulate the size of the species populations (this control task is not considered in the present working sessions).

1. Using the notation :

$$v := \begin{pmatrix} r & \frac{r}{K} & s & \alpha & \beta \end{pmatrix}^T \in \mathbb{R}^5 \qquad (4)$$

show that the equations (1)-(2) can be written in the following form :

$$\frac{\dot{P}}{P} = [y_1(P,Q)]^T \cdot v \tag{5}$$

$$\frac{\dot{Q}}{Q} = [y_2(P,Q)]^T \cdot v \tag{6}$$

give the explicit expressions of the vector valued functions $y_1(\cdot,\cdot)$ and $y_2(\cdot,\cdot)$

2. Assuming that the measurement period is sufficiently small, show that the equations (5)-(6) can be approximately written in the following form for all measurement instant $k \in \{1,\dots,N-1\}$:

$$[L(P(t_k),Q(t_k),\tau)] \cdot v \approx q(P(t_k),Q(t_k),P(t_{k+1}),Q(t_{k+1})) \tag{7}$$

give the explicit expressions of $L \in \mathbb{R}^{2\times 5}$ and $q \in \mathbb{R}^2$ as functions of their arguments.

3. By writing (7) for all $k \in \{1,\dots,N-1\}$, show that the vector of parameters v satisfies a least squares set of equations of the form :

$$[A(Y|_1^N)] \cdot v = B(Y|_1^N) \tag{8}$$

Give the explicit expressions of $A(\cdot) \in \mathbb{R}^{(2(N-1))\times 5}$ and $B(\cdot) \in \mathbb{R}^{2(N-1)}$.

4. Write a Matlab function that uses the following syntax :

$$[A,B] = ComputeMatrices(Y,\tau)$$

where Y is a matrix containing N lines and 2 columns (the measurement data with the measurements of P in the first column and those of Q in the second)n τ is the sampling period while A and B are the least squares matrices defined in (8).

5. Execute your function using the following script :

```
>>Ytest=reshape(cos((1:1:10)'*0.05),5,2);
>>[Atest,Btest]=ComputeMatrices(Ytest,0.1);
```

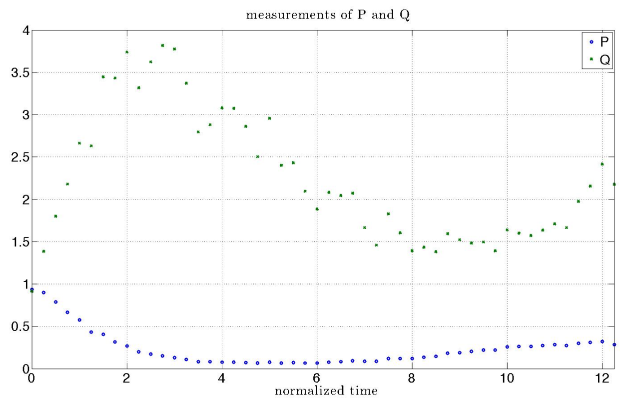
and check that the results are as follows

```
>> [Atest,Btest]

ans =

    0.0999   -0.0998   -0.0954         0         0   -0.0037
         0         0         0   -0.0955    0.0954   -0.0160
    0.0995   -0.0990   -0.0935         0         0   -0.0062
         0         0         0   -0.0939    0.0935   -0.0183
    0.0989   -0.0978   -0.0911         0         0   -0.0087
         0         0         0   -0.0921    0.0911   -0.0206
    0.0980   -0.0961   -0.0882         0         0   -0.0112
         0         0         0   -0.0900    0.0882   -0.0229
```

Note that these matrices have nothing to do with our problem. They just enable to check your function.

6. **Short Measurement campaign**

A first (rather short) measurement campaign has been done and the related measurements are saved in the mat file *data1.mat*. The resulting measurements are shown on the figure beside for P and Q.



   (a) load the mat file *data1.mat* to your workspace and compute the resulting least squares matrices A and B using the data Y and τ it contains.

   (b) Does the measurement theoretically enable the vector v to be recovered ?

   (c) Solve the least squares problem and use the resulting solution v^* to estimate the value of the model parameters r, K, s, α and β.

   (d) What is the norm of the residual associated to the solution ?

7. Check your solution by simulating the system model (1)-(2) using the parameter values you have found and starting with the same initial conditions as those contained in the measurement data *data1.mat*. Plot on the same figure the simulated evolutions of P and Q and the experimental measurements.

8. **Long Measurement Campaign**

In order to get deeper insight on the model behaviour, a longer campaign of measurements has been undertaken the result of which are stored in the file *data2.mat*.

Answer the same questions 6.(a) to 6.(d) and 7 using the *data2.mat* instead of *data1.mat*. Does the result seems satisfactory ?

9. **Recursive least squares**

In this last question, we shall use the long campaign measurements contained in the file *data2.mat* in order to estimate the parameter using a recursive least squares estimation procedure (see course) with the following blocs of successive pairs (A_j, b_j) :

$$A_j = A(Y|_j^{j+N_p-1}) \quad ; \quad b_j := A(Y|_j^{j+N_p-1}) \quad ; \quad j \in \{1, \dots, N - N_p + 1\} \tag{9}$$

where N is the number of measurements in *data2.mat* while $N_p = 50$ is the dimension of measurements taking into account at each step of the recursive least squares.

Remember that with this procedure, no estimation can begin before N_p measurements are available. Therefore, we consider during the first $N_p - 1$ sampling instants that the value of the parameters are equal to the values we've estimated during the short measurement campaign. The forthcoming values are obtained from the recursive least squares procedure (see course).

Compute and plot the estimated curves for all the 5 parameters during the long measurement campaign for a forgetting factor $\lambda = 0.5$.
[hint : use subplot to gather all the plots in a reduced space]

10. Based on these curves, can you explain the discrepancy between the measured behaviour and the simulated one ?