

Parameters identification of a population of systems based on EM algorithm

Levy Batista^{1,3}, Thierry Bastogne^{1,2,3}, El-Hadi Djermoune¹

¹CRAN CNRS UMR 7039 BP 70239, F-54506 Vandoeuvre-lès-Nancy Cedex, France, ²INRIA BIGS, BP 70239, F-54506 Vandoeuvre-lès-Nancy Cedex, France, ³CYBERnano, Telecom Nancy, 193 av. Paul Muller, 54602 Villers-lès-Nancy

Context

- Most of biological responses measured during experiments exhibit **non-linear behaviors** and **mixed effects models** are often used to better characterize the responses.
- Parameter and confidence intervals estimation involve numerical integration, linearization or stochastic approximation algorithms [1].
- Instead of modeling the response we propose a method in which each biological process is regarded as a dynamical system with input-output variables and cofactors. We show that **this approach allows to use classical linear methods**.

The hierarchical ARX Model

First stage

The first stage is described as an Auto-Regressive with eXogenous input (**ARX**) model structure of order n_a , n_b and n_d as described in [2] for a given subject $i \in [1, \dots, I]$:

$$y_i \sim \mathcal{N}(\phi_i z_i, \mathbf{I}\sigma_e^2) \quad (1)$$

where

- $\phi_i = [-y_i q^{-1}, \dots, -y_i q^{-2}, u_i q^{-n_d}, \dots, u_i q^{-n_b-n_d}]$ and q^{-1} is the delay operator
- $z_i = [a_{1,i}, a_{2,i}, \dots, b_{nb,i}]^T \in \mathbb{R}^{(n_a+n_b)}$

with $y_i \in \mathbb{R}^K$ the output signal, n_d the input-output delay and $a_{i,j}$ and $b_{i,l}$ are the unknown parameters of the individual i .

Second stage

$$z_i \sim \mathcal{N}(C_i \theta_g, \Omega) \quad (2)$$

where

- $\theta_g = [a_{1,1}, \dots, b_{nb,1}, \dots, a_{1,n_g}, \dots, b_{nb,n_g}]^T \in \mathbb{R}^{(n_g(n_a+n_b))}$
- C_i is a covariables matrix $\in \mathbb{R}^{(n_a+n_b) \times n_g(n_a+n_b)}$
- $\text{diag}(\Omega) = [\sigma_{a,1}^2, \dots, \sigma_{a,n_a}^2, \sigma_{b,1}^2, \dots, \sigma_{b,n_b}^2]^T$

The parameters to be estimated are :

$$\Theta = (\theta_g, \Omega, \sigma_e^2) \quad (3)$$

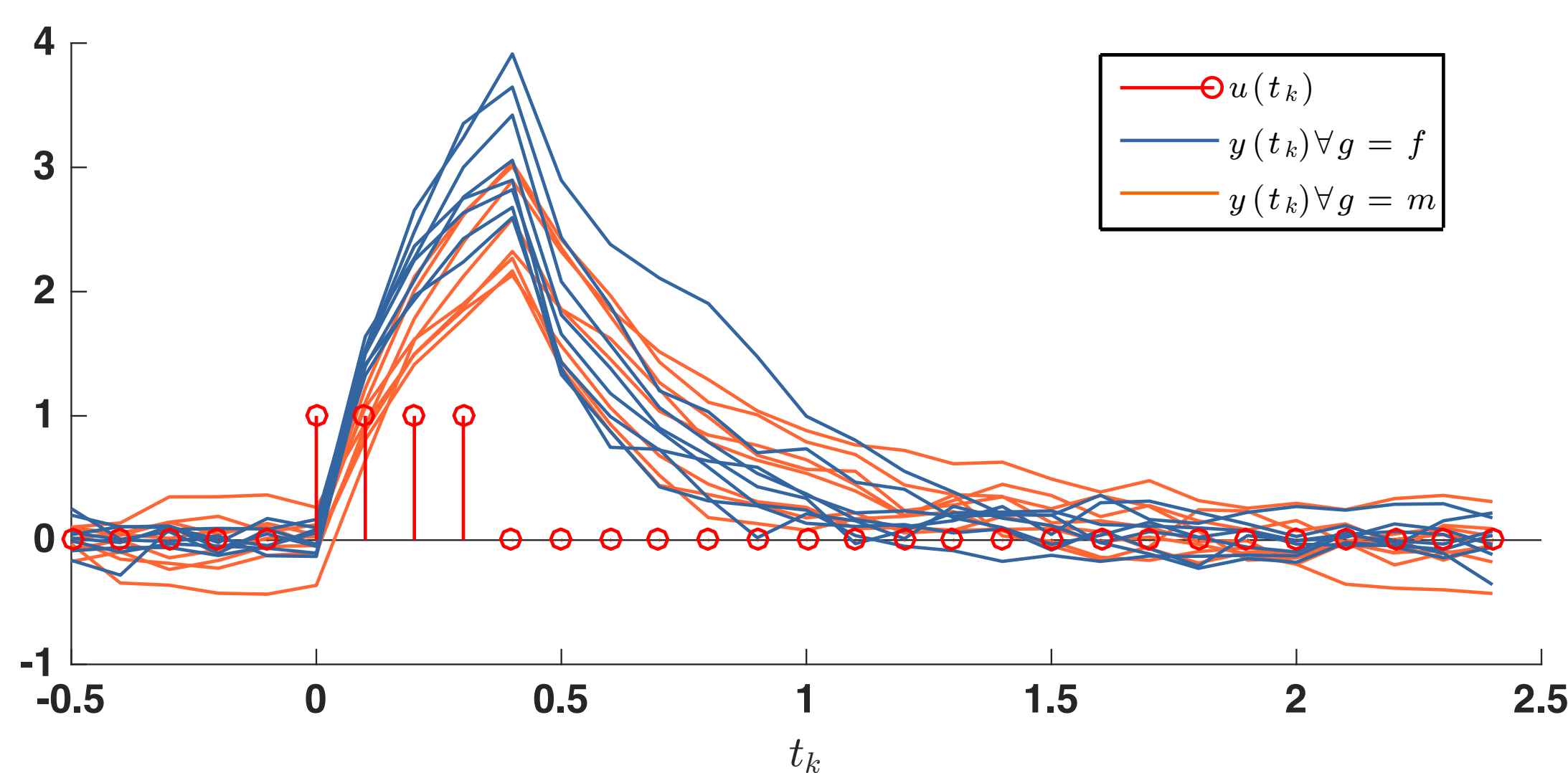


FIGURE 1: A simulation of the model with 8 individuals per group

Parameter estimation

EM Algorithm : Since the model is linear, Expectation-Maximisation algorithm can be used in order to maximise the likelihood of the complete data (y_i, z_i) . [3].

- Expectation : calculate the sufficient statistics :

$$\begin{aligned} \hat{t}_{1,i}^{(s)} &\triangleq \mathbb{E}(z_i | y_i, \Theta^{(s-1)}), & \hat{t}_{3,i}^{(s)} &\triangleq \mathbb{E}(e_i | y_i, \Theta^{(s-1)}) \\ \hat{t}_{2,i}^{(s)} &\triangleq \text{var}(z_i | y_i, \Theta^{(s-1)}), & \hat{t}_{4,i}^{(s)} &\triangleq \text{var}(e_i | y_i, \Theta^{(s-1)}) \end{aligned}$$

from the distribution of the complete data vector

$$y_i^{(c)} = \begin{pmatrix} y_i \\ z_i \end{pmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \phi_i C_i \theta_g^{(s)} \\ C_i \theta_g^{(s)} \end{bmatrix}, \begin{bmatrix} V^{(s)} & \phi_i \Omega^{(s)} \\ \Omega^{(s)} \phi_i^T & \Omega^{(s)} \end{bmatrix} \right) \quad (4)$$

where $\Theta^{(s)}$ is the parameters estimated at iteration s , and $V^{(s)} = \phi_i \Omega^{(s)} \phi_i^T + \mathbf{I}_N \sigma_e^{2(s)}$

- Maximisation : since all individuals are independent from each other the likelihood is written :

$$p(y, z; \Theta) = \prod_i p(y_i | z_i; \Theta) p(z_i; \Theta), \quad (5)$$

which after maximisation leads to :

$$\hat{\theta}_g^{(s)} = \left(\sum_i C_i^T C_i \right)^{-1} \sum_i C_i^T \hat{t}_{1,i}^{(s)} \quad (6)$$

$$\hat{\Omega}^{(s)} = \frac{1}{I} \sum_i (\hat{t}_{1,i}^{(s)} - C_i \theta_g^{(s)}) \cdot (\hat{t}_{1,i}^{(s)} - C_i \theta_g^{(s)}) + \hat{t}_{2,i}^{(s)} \quad (7)$$

$$\hat{\sigma}_e^{2(s)} = \frac{1}{NI} \sum_i \hat{t}_{3,i}^{(s)T} \hat{t}_{3,i}^{(s)} + \text{Tr}(\hat{t}_{4,i}^{(s)}) \quad (8)$$

This iterative schemes is repeated until convergence, then based on Louis' principle, [4] the Fisher information matrix is computed to provide an upper bound of the estimated variance matrix

Simulation evaluation

The model used for simulation is :

$$y_i = -a_{1,g} y_i q^{-1} - a_{2,g} y_i q^{-2} + b_{1,g} u_i q^{-1} + e(k) \quad (9)$$

where the index g corresponds to the group of the individual, here g can be m or f . Values used for simulation are :

$$\begin{bmatrix} a_{1,m} \\ a_{2,m} \\ a_{3,m} \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.05 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} a_{1,f} \\ a_{2,f} \\ a_{3,f} \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.05 \\ 1.5 \end{bmatrix}$$

and :

$$\begin{bmatrix} \sigma_{a_1}^2 \\ \sigma_{a_2}^2 \\ \sigma_{b_1}^2 \end{bmatrix} = \begin{bmatrix} 0.001 \\ 0.001 \\ 0.01 \end{bmatrix} \text{ and } \sigma_e^2 = 0.01$$

An example of simulation with $I_m = 8$ and $I_f = 8$ is presented in figure 1

Results

$I_m = 4, I_f = 4$			$I_m = 10, I_f = 10$		
Parameters	Estimate	SE (%)	Parameters	Estimate	SE (%)
$a_{1,m}$	0.6772	3	$a_{1,m}$	0.6603	1
$a_{2,m}$	0.0521	38	$a_{2,m}$	0.0865	10
$b_{1,m}$	1.0077	2	$b_{1,m}$	1.0105	1
$a_{1,f}$	0.5198	4	$a_{1,f}$	0.4910	1
$a_{2,f}$	0.0816	18	$a_{2,f}$	0.1107	6
$b_{1,f}$	1.4909	2	$b_{1,f}$	1.5082	1
$\sigma_{a_1}^2$	0.0020	35	$\sigma_{a_1}^2$	0.0016	14
$\sigma_{a_2}^2$	0.0008	57	$\sigma_{a_2}^2$	0.0013	14
$\sigma_{b_1}^2$	0.0035	36	$\sigma_{b_1}^2$	0.0065	11
σ_e^2	0.0103	0	σ_e^2	0.0098	0

$I_m = 100, I_f = 100$			$I_m = 100, I_f = 10$		
Parameters	Estimate	SE (%)	Parameters	Estimate	SE (%)
$a_{1,m}$	0.6868	0	$a_{1,m}$	0.6875	0
$a_{2,m}$	0.0613	1	$a_{2,m}$	0.0606	1
$b_{1,m}$	1.0010	0	$b_{1,m}$	1.0004	0
$a_{1,f}$	0.5030	0	$a_{1,f}$	0.5331	1
$a_{2,f}$	0.0969	1	$a_{2,f}$	0.0697	8
$b_{1,f}$	1.4995	0	$b_{1,f}$	1.5041	1
$\sigma_{a_1}^2$	0.0013	2	$\sigma_{a_1}^2$	0.0015	3
$\sigma_{a_2}^2$	0.0012	2	$\sigma_{a_2}^2$	0.0008	5
$\sigma_{b_1}^2$	0.0088	1	$\sigma_{b_1}^2$	0.0082	2
σ_e^2	0.0101	0	σ_e^2	0.0100	0

SE is the standard error in percent calculated from the standard deviation estimated from the Fisher information matrix.

Discussion

Mixed-effects are introduced in the ARX structure, EM algorithm is used to estimate the parameters of the system

Advantages

- This approach allows to fit non-linear responses using a linear model : we model the hidden system instead of the response itself,
- Any kind of input can be used
- Systems of any order can be computed

Drawbacks

- An input is needed
- Constant sampling rate is required
- Parameters have no physical meanings

References

- M. Lavielle, *Mixed Effects Models for the Population Approach. Models, Tasks, Methods & Tools*. Chapman & Hall/CRC Biostatistics Series, 2014.
- L. Ljung, *System Identification, Theory for the User*. PreWiley-Hall, 1987.
- A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the royal statistical society, Serie B*, vol. 39, no. 1, pp. 1-38, 1977.
- T. A. Louis, "Finding the observed information matrix when using the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 44, no. 2, pp. pp. 226-233, 1982.