Technical Note:

Identifying Markov models for hypothetical financial inclusion processes

Fredy Vides[1, a)]

Department of Applied Mathematics, Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras

(Dated: December 10, 2024)

In this document, some mathematical and computational techniques for the identification of Markov models for socioeconomic processes, are presented. For this technical note only a hypothetical financial inclusion process will be considered.

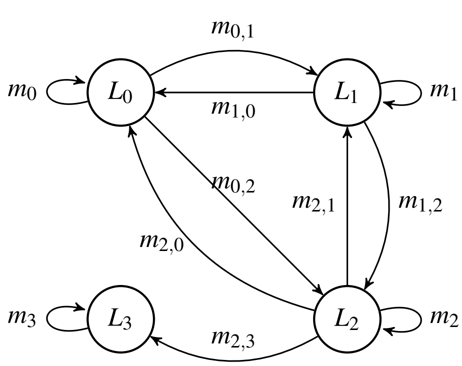
I. CONCEPTUALIZATION

Given a hypothetical financial inclusion process (**FIP**), we approach the predictive analytics model corresponding to the financial inclusion data as a structure preserving modeling method, so that some suitable hypotheses corresponding to levels of participation of the population in the financial system, can be implemented in the computational model.

At this point of the project, we will consider the following four levels of participation:

- L_0: At this level there is no participation in the financial system.
- L_1: At this level a user has access to services from a rural box, while still being at risk of moving to level L_0.
- L_2: At this level a user has access to financial services from cooperatives, financial societies or banks, while still being at risk of moving to level L_0 or L_1.
- L_3: At this level a user has full and permanent access to any financial service from any financial institution, and no significant risk of moving to lower levels of participation.

For this stage of the project, we will consider an initial configuration for potential movements between participation levels to be determined by the graph described in (I.1).



(I.1)

a) Electronic mail: fredy.vides@unah.edu.hn

II. A MARKOV MODEL FOR HYPOTHETICAL HONDURAN FINANCIAL INCLUSION PROCESSES

Let us write $p(t)$ to denote the array of four non-negative numbers

$$p(t) := \begin{bmatrix} p_0(t) & p_1(t) & p_2(t) & p_3(t) \end{bmatrix}^\top$$

corresponding to departmental population participation proportions corresponding to the four levels considered in §I, that have been measure/predicted for t months after a national financial inclusion strategy (NFIS) was initiated. For this study, we will consider switched discrete-time Markov models of the form:

$$(I - M_{\tau(t)}(q))p(t) = e_p(t),\ t \geq 0. \tag{II.1}$$

where q denotes the lag operator, $M_{\tau(t)} \in \mathbb{R}^{4\times 4}[z]$ for each discrete time step t, and $\tau : \mathbf{Z}^+ \to \mathbf{Z}^+$ denotes some switching law to be identified along with the corresponding $M_{\tau(t)}$, in terms of the residual signal $e_p(t)$, for n validity periods $[t_0, t_1) \cap \mathbf{Z}, \ldots, [t_{n-1}, t_n) \cap \mathbf{Z}$ where n is some prescribed positive integer.

Based on the structural constraints considered in the graph presented in (I.1), the structure of the matrix coefficients of each matrix polynomial $M_{\tau(t)}$ will be determined by the expression

$$M_\tau(z) := \begin{bmatrix} m_0(\tau) & m_{1,0}(\tau) & m_{2,0}(\tau) & 0 \\ m_{0,1}(\tau) & m_1(\tau) & m_{2,1}(\tau) & 0 \\ m_{0,2}(\tau) & m_{1,2}(\tau) & m_2(\tau) & 0 \\ 0 & 0 & m_{2,3}(\tau) & m_3(\tau) \end{bmatrix} z, \tag{II.2}$$

for an arbitrary variable z.

In addition the coefficients of each coefficient matrix in (II.2) should satisfy some constraints related to the transition probabilities between the participation levels of the population of each Honduran department in the financial system.

III. MODEL IDENTIFICATION SKETCH FOR FINANCIAL INCLUSION SYNTHETIC DATA

In this section we will assume that $p(0) = \begin{bmatrix} 0.6 & 0.09 & 0.3 & 0.01 \end{bmatrix}^\top$, and that $M_{\tau(t)}$ remains constant for a certain validity period $[0, T) \cap \mathbf{Z}$ for some integer $T > 0$. Let us consider a sample $\Sigma_4 := \{p(t)\}_{t=0}^{3} \subset \Sigma$ from the time series $\Sigma := \{p(t) : t \geq 0\}$, determined by the recurrence relation (II.1) and the previously considered initial state $p(0)$.

In order to compute a model of the form $f(q(t)) = \hat{M}_0 q(t)$ to identify the discrete-time system (II.1), one can take advantage of the structure of the dynamics matrix in (II.1) induced by the graph in (I.1), to restate the identification problem as a structured matrix equation corresponding to a quadratic optimization problem[2] (§7.8) of the form of (II.1), that can be described by the expression:

$$\begin{bmatrix} P_0^\top P_0 & C^\top \\ C & 0 \end{bmatrix} \begin{bmatrix} y_p \\ \lambda \end{bmatrix} = \begin{bmatrix} P_0^\top P_1 \\ d \end{bmatrix} \tag{III.1}$$

Where the columns of P_0 have been computed in terms of the elements in a sample $\{p(0), p(1), p(2)\}$ and the elements of a basis $\mathbb{B} = \{B_1, \ldots, B_m\} \subset \mathbb{R}^{4\times 4}$ determined by the structure of the dynamics matrices to be identified, and where $P_1 = \mathrm{col}(p(1), p(2), p(3))$. With $Cy_p - d$ representing the relations corresponding to the constraints $\sum_{i=1}^4 \hat{m}_{ij} = 1$, $1 \leq j \leq 4$, for $\hat{M}_0 = [\hat{m}_{ij}] = \sum_{s=1}^m y_s(p) B_s$ determined by the components of $y_p = [y_s(p)]$ of a solution to (III.1).

By using the jupyter notebook `SPMarkovModelID.ipynb` written by F. Vides, one can compute a predictive model $L(q(t)) = \hat{M}_0 q(t)$ based on a subsample $\{p(0), \ldots, p(3)\}$ of Σ applying the ideas previously considered.

The graphical representation corresponding to the predictions obtained with the model $L(q(t)) = \hat{M}_0 q(t)$ are shown in Figure 1. The predicted values $\hat{\Sigma}_{61} = \{\hat{p}(0), \ldots, \hat{p}(60)\}$ are computed using $\hat{M}_0$ and $q(0) = \hat{p}(0) = p(0)$ according to equation (II.1).

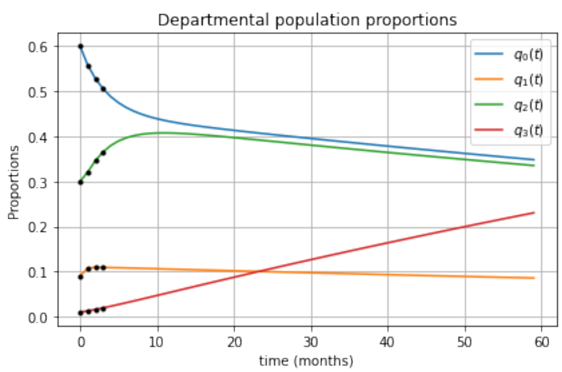


Figure 1: Graphical representation of model's prevision 60 months after NFIS was initiated, for some Honduran department under consideration. The black dots denote model's training data.

IV. CONCLUDING REMARKS

- Both the graph in (I.1) and the corresponding predictive model determined by (II.1) evolve over time in order to guarantee an accurate description/prediction of the hypothetical process behavior driven by financial inclusion data observations.
- The model determined by (II.1) describes/predicts the evolution of each departmental population participation proportions corresponding to each one of the four levels considered in §I, ranging from no participation to full and permanent participation.

ACKNOWLEDGMENT

The author wishes to thank Bryan Barnett for interesting discussions, that have been very helpful for the preparation of this report.

REFERENCES

[1] S. Boyd, L. Vandenberghe (2018). Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares. Cambridge University Press.
[2] A. Quarteroni, F. Saleri, P. Gervasio (2014). Scientific Computing with MATLAB and Octave, 4th ed. Springer-Verlag Berlin Heidelberg.
[3] F. Vides (2021). Métodos Numéricos y Modelación Computacional. Libro electrónico de lecturas. `https://cadds-lab.github.io/MNMC.pdf`