A control theoretic approach to evaluate and inform ecological momentary interventions

Instructions for code usage

August 31, 2023

1 Introduction

In the main article, we employ concepts and measures from network control theory (NCT) to formally quantify and analyze proximal intervention effects on an individual's mental state as measured by time series of ecological momentary assessment (EMA) variables. This GitHub repository contains the complete code used for performing the analyses mentioned, and an additional tutorial notebooks. For instructional purposes, we use random synthetic time series resembling the original data. To reproduce the figures and statistical tests, the original data is available on reasonable request to the corresponding author.

2 Setting up the environment

The code requires only standard python packages, as specified in the `requirements.txt` file. We recommend that you set up a python environment specifically for this code. If you're using Windows, this is done from the Power Shell by

```
python -m venv .venv
.venv/Scripts/activate.ps1
pip install -r requirements.txt
```

3 File overview

- `tutorial.ipynb`

  Jupyter notebook that guides the user through the analyses performed in the paper, step by step, using a randomly generated synthetic EMA time series. You can easily replace the synthetic data with your own.

- `ctrl/discrete_optimal_control.py`

  Module containing all control theoretic methods.

- `ctrl/utils.py`

  Module containing additional resources for data loading, plotting, model fitting, and statistical testing.

- `figure<x>.ipynb`

  Notebook that produces figure ¡x¿ from the paper. Each notebook contains a global variable `TEST`. If `TEST==False`, the original analyses are performed and the plots are drawn as they appear in the paper, provided the original data files are placed in the directory `data`. Figure panels are saved to `figures/figure<x><panel>.png`. If `TEST==True`, analyses are carried out with randomly generated synthetic data, and figures are not saved.

- `custom_rcparams_paper.py`

  Plotting parameters.

- `model_distribution.py`

  Contains parameters of the distribution from which a model is drawn to create synthetic data.

4 Notes on the tutorial

The `tutorial.ipynb` performs all NCT analyses mentioned in the paper on a VAR(1) model fitted to a synthetic time series: controllability and impulse response measures, as well as optimal control strategies. Please note that in the article, these analyses are carried out on the VAR(1) model of each subject and centrality measures and/or distributions are reported. For the sake of simplicity, the tutorial only shows the analysis of a single model. Hence, the way results are presented differ slightly from the article.

5 Notes on the synthetic data

The method `utils.generate_random_trajectory(T, seed)` creates synthetic EMA time series in the following way:

1. A multivariate Gaussian distribution was estimated over the A and B matrices of the VAR(1) models used in the study. The estimated mean and covariance are stored in the file `model_distribution.py`. The original data set comprised 15 EMA variables and 4 EMI, thus the synthetic data will have the same size.

2. A VAR(1) model is created using $A \in \mathbb{R}^{15\times 15}$ and $B \in \mathbb{R}^{15\times 4}$ matrices drawn at random from this distribution.

3. An initial condition x_1 is drawn at random from a uniform distribution over $\{-3, -2, -1, 0, 1, 2, 3\}^{15}$.

4. A sequence of input vectors $U = u_1, \ldots, u_{T-1}$ is drawn. Every second input is an indicator vector from the set
$$\{[1, 0, 0, 0], [0, 1, 0, 0], [0, 0, 1, 0], [0, 0, 0, 1]\}$$
representing the occurrence of 1 out of 4 EMI. The others are 0. This roughly reflects EMI distributions, where uniformly drawn EMI are administered every second time an EMA is carried out.

5. A sequence $X = x_1, \ldots, x_T$ is created by iterating the model
$$Ax_t + Bu_t$$
from the initial condition $t = 1$ to $t = T$, where T is the first argument to the method. The sequence is rounded to the nearest integer vectors in $\{-3, -2, -1, 0, 1, 2, 3\}^{15}$ to mimic Likert scale EMA data.

6. The method returns X, U.

This way, the synthetic data very roughly reflect the properties of the empirical data, which are time series of Likert scales and indicator variables, and even may elicit similar results in the analyses. To obtain the same results again when drawing random synthetic data, specify the `seed` argument to `utils.generate_random_trajectory`.

6 Questions welcome

Please address any questions or comments to the corresponding author Janik Fechtelpeter.