Network Experimentation Project

Mine Su Erturk, Eray Turkel

September 21, 2020

1   Introduction

There is a decision maker conducting experiments on a network environment. An 'experiment' in the context of our model is treating a node in the network. Each node responds to the treatment differently, depending on their (observable) characteristics.

The decision maker's goal is learning the optimal treatment allocation over the network by sequentially conducting local experiments on different nodes. Because of the network setting, treating one node creates spillovers on neighboring nodes.

Finally, the distinguishing feature of our model is that every experiment creates negative externalities on previous experiments. We model this as a cost for contaminating the treatment regime decided by a previously run experiment.

1.1   Model

There are N units (individuals, firms, locations) connected on an undirected, unweighted graph. Denote each unit on the graph by $i \in \{1, \ldots, N\}$.

The units have observable characteristics, $\{x_1, \ldots, x_N\}$ with $x_i \in \mathbb{R}^k$. The experimenter's objective depends on the responses of the nodes to the treatment regime, denoted $\{Y_1, \ldots, Y_N\}$ with $Y \in \mathbb{R}$. The decision maker can run experiments sequentially, at time periods $t \in \{1, \ldots, T\}$. Let w_{it} denote the treatment assignment for node i during the period of experimentation t, with $w_{it} \in \{0, 1\}$, and denote the vector of treatment assignments at time t by $w_t = [w_{1t}, \ldots, w_{Nt}] \in \{0, 1\}^N$. We assume that the experimentation is local, in the sense that every period, only one node can be treated. Thus for all time periods t, $\sum_{i \leqslant N} w_{it} = 1$.

Each node's outcome depends on their characteristics, the treatment they receive and the exposure to the treatment through their neighbors. Y is assumed to be a linear function of the unit characteristics and the treatment regime.

Define the exposure for node i at period t as e_{it}, which is (for simplicity) assumed to be:

$$e_{it} = \frac{\sum_{j \leqslant N} \mathbf{1}(j \in nhbd(i)) w_{jt}}{\sum_{j \leqslant N} \mathbf{1}(j \in nhbd(i))}$$

We assume a linear response function taking the following form:

1

$$Y_{it}(w_{it}, e_{it}) = \beta^T x_i + \Gamma^T x_i e_{it} + \epsilon_i,$$

where $\beta, \Gamma \in \mathbb{R}^k$, and for all nodes i, $\epsilon_i \sim N(0, \sigma^2)$.

The decision maker has a prior over the value of the vector Γ and updates their beliefs through sequential experimentation. The DM's goal is to choose a treatment assignment policy over the network based on the value of Γ. Denote the final treatment assignment for node i by w_i, dropping the time subscript. We denote a policy by $\pi(\Gamma)$, where $\pi(\Gamma) = [w_1, \ldots, w_N] \in \{0,1\}^N$. We assume that the number of nodes that can be treated during the final assignment can be at most D, i.e., $\sum_{i \leqslant N} w_i \leqslant D$.

At every period t, the DM's experiment creates a cost, because the experiment interacts with a previous experimenter's treatment allocation. We assume that the cost is proportional to the number of units exposed to the treatment, and once a unit is exposed, future exposures of that unit do not increase the cost any further. Define the total cost of experimentation at period t by:

$$c_t = c \sum_{i \leqslant N} \mathbf{1}\left(e_{it} > 0 \text{ and } e_{is} = 0, \forall s < t\right).$$

Hence, the total cost incurred through experimentation by the decision maker can be written as:

$$C = \sum_{t \leqslant T} c_t.$$

We assume that the DM can incur a cost of at most $\bar{C}$ during experimentation. Therefore, we require that any feasible policy π satisfies $C \leqslant \bar{C}$. Then, as a function of the cost budget $\bar{C}$, we define the value of a policy as follows:

$$V(\pi, \bar{C}) = \sum_{i \leqslant N} \Gamma^T x_i e_i(\pi),$$

where $e_i(\pi)$ denotes the exposure function at the last period, induced by the treatment allocation under policy π, and $\bar{C}$ is the cost budget.

The DM does not observe the vector Γ but has an imperfect estimate of it learned through sequential experimentation, which we will call $\hat{\Gamma}$. We then formulate the DM's optimization problem:

$$Max_{\hat{\pi}} : E\left(\sum_{i \leqslant N} \hat{\Gamma}^T x_i e_i(\hat{\pi})\right) \quad \text{s.t. } 1^T \hat{\pi} \leqslant D. \tag{1}$$

1.2 Last Period Deployment Problem

Let $X_{K \times N}$ denote the matrix of observed covariates for the nodes. Let $\hat{\Gamma}_{1 \times K}$ be the vector of estimated coefficients.

Define the vector of exposure values at time t as $\bar{e}_t = [\mathbf{1}(e_{1t} > 0), \ldots, \mathbf{1}(e_{Nt} > 0)]_{1 \times N}$ and let c_t denote the cost vector at time t. We can recursively write:

$$c_0 = 0_{1 \times N}$$

$$c_1 = \bar{e}_1$$

$$c_2 = (c_1 + e_2) - (Diag(c_1)\bar{e}_2)$$

$$\vdots$$

$$c_n = (c_{n-1} + \bar{e}_n) - (Diag(c_{n-1})\bar{e}_n),$$

where $Diag(c_n)$ denotes the diagonal matrix with the entries of c_n on its diagonal entries. This recursive formulation captures the notion that the cost of second and future exposures to the experiment are zero.

We can also define the exposure vector e in terms of the adjacency matrix of the network, $A_{N \times N}$ (with self-edges), and the treatment allocation vector $\pi \in \{0,1\}^N$.

Representing e as a vector, we have:

$$Diag(A\mathbf{1}_N)^{-1}A\pi = e,$$

where $\mathbf{1}_N$ is an $(N \times 1)$ vector of 1's.

Therefore, we can rewrite the maximization problem (1) as:

$$Max_{\hat{\pi}} : E\left(\left(\hat{\Gamma}^T X Diag(A\mathbf{1}_N)^{-1}A\right)\hat{\pi}\right)$$

subject to:

$$1^T\hat{\pi} \leqslant D$$

$$\forall i, \hat{\pi}_i \in \{0,1\}$$

1.3   Regret

Let $z(\hat{\Gamma})$ denote the optimal objective function value of the optimization problem (1) given that the DM has estimated the coefficient vector as $\hat{\Gamma}$. Note that an oracle with full knowledge of the coefficient vector Γ would be able to achieve $z(\Gamma)$. Thus, we can define the regret as follows:

$$R = z(\Gamma) - z(\hat{\Gamma})$$

.

Theorem 1 *The regret is upper bounded by*

$$R \leqslant (\Gamma - \hat{\Gamma})K \tag{2}$$

where K is a constant that depends on the adjacency matrix A, and the allocation budget D.