1 Model specification

1.1 Parameter space

Consider a linear dynamic panel model with n individuals, each with T length of observations, where each agent belongs to one of the discrete types of unobserved heterogeneity indexed by for some given $\mathcal{X} = \{1, 2, \cdots, M\}$ where $M \in \mathbb{N}$. The probability of belonging to jth component is determined by $\alpha_j \in \Theta_\alpha$ for some $\Theta_\alpha \subset [0, 1]$ such that $\sum_{j=1}^{M} \alpha_j = 1$. The observation on period t from ith agent whose latent variable is $x_i \in \{1, ..., M\}$, y_{it}, is generated by

$$y_{it} = \overline{\mathbf{y}}'_{it}\boldsymbol{\rho}_{x_i} + \mathbf{x}'_{it}\boldsymbol{\beta}_{x_i} + \mathbf{z}'_{it}\gamma + \epsilon_{it} \tag{1}$$

where ϵ_{it} is an iid error term generated by $N(\mu_{x_i}, \sigma_{x_i})$ such that $\mu_{x_i} \in \Theta_\mu$ for some $\Theta_\mu \subset \mathbb{R}$ and $\sigma_{x_i} \in \Theta_\sigma$ for some $\Theta_\sigma \subset \mathbb{R}_+$. We assume that coefficients are not random so that $\boldsymbol{\beta}_{x_i} \in \Theta_{\boldsymbol{\beta}}, \boldsymbol{\rho}_{x_i} \in \Theta_{\boldsymbol{\rho}}, \gamma \in \Theta_\gamma$ for some $\Theta_{\boldsymbol{\beta}} \subset \mathbb{R}^q, \Theta_{\boldsymbol{\rho}} \subset \mathbb{R}^s, \Theta_\gamma \subset \mathbb{R}^p$. $\overline{\mathbf{y}}'_{it}$ represents s collection of previous observations; $\overline{\mathbf{y}}'_{it} := (y_{i(t-1)}, y_{i(t-2)}, \cdots, y_{i(t-s)})$.

Let $\boldsymbol{\vartheta}_j$ denote the collection of model specifications from jth component in a vector form

$$\boldsymbol{\vartheta}_j = (\alpha_j, \boldsymbol{\rho}'_j, \boldsymbol{\beta}'_j, \mu_j, \sigma_j)'$$

with $\alpha_j \in \Theta_\alpha, \boldsymbol{\rho}_j \in \Theta_{\boldsymbol{\rho}}, \boldsymbol{\beta}_j \in \Theta_{\boldsymbol{\beta}}, \mu_j \in \Theta_\mu, \sigma_j \in \Theta_\sigma$. Then the full model specification of a M-component mixture dynamic panel model can be represented by

$$\boldsymbol{\theta} = (\gamma, \boldsymbol{\vartheta}'_1, \cdots, \boldsymbol{\vartheta}'_M)'$$

for some $\gamma \in \Theta_\gamma$. We denote the space of $\boldsymbol{\theta}$ as Θ.

1.2 Likelihood function

Collect the T observations from an ith individual as $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})$, and let $\mathcal{Y}_n = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n)$ denote the collection of observations from all individuals. Then, the likelihood function of the model given $\boldsymbol{\theta} \in \Theta$ is

$$L_n(\mathcal{Y}_n, \vartheta) = \prod_{i=1}^{n} L(\mathbf{y}_i, \theta, \alpha)$$

with

$$L(\mathbf{y}_i, \boldsymbol{\theta}) = \sum_{j=1}^{M} \alpha_j f(\mathbf{y}_i; \boldsymbol{\vartheta}_j)$$

$$f(\mathbf{y}_i; \boldsymbol{\vartheta}_j) = \prod_{t=1}^{T} \frac{1}{\sigma_j} \phi \left(\frac{y_{it} - \mu_j - \overline{\mathbf{y}}'_{k-1}\boldsymbol{\rho} - \mathbf{x}'_{it}\boldsymbol{\beta}_{x_t} - \mathbf{z}'_{it}\gamma}{\sigma_j} \right)$$

where ϕ is the density function of a standard normal random variable.

2 Quick Example

In mixPanel, a model specification is represented as a list. A model specification can be randomly created by calling `GenerateMDPTheta`:

```
set.seed(1234)
# generates AR(2)-MDP(3) model
theta <- GenerateMDPTheta(M = 3, s = 2)
```

Given a model specification, a random sample can be generated by calling `GenerateMDPSample`. Users can specify the number of individuals (`N`) and the length of individual time series (`T`) as well.

```
# generates N = 60, T = 10 random sample with specification above
sample <- GenerateMDPSample(theta = theta, N = 60, T = 10)
```

Calling `GenerateMDPSample` returns a list with the following items: `y`, `y.sample`, `y.lagged`, and `MDP.model`.

- y is a $(T + s) \times N$ matrix whose column represents observations from each individual. This includes the lagged variables generated for autoregression in initial s members. ith row represents observations in ith period.

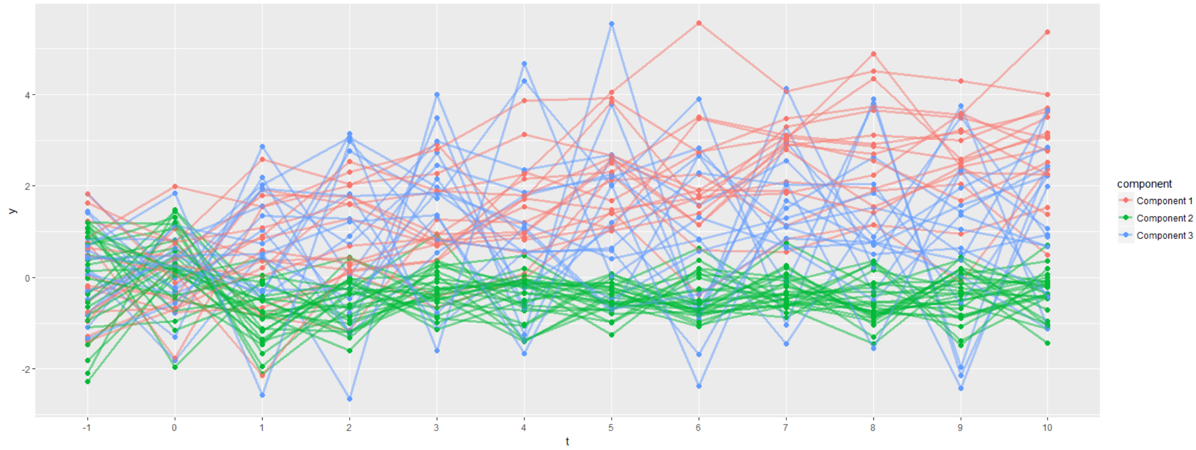
- `y.sample` is a $T \times N$ matrix whose column represents observations from each individual. Unlike `y`, this excludes the lagged variables generated for autoregression in initial s members.

- `y.lagged` is a $T \times N(1+s)$ blocked matrix partitioned by $(1+s)$ columns. Each jth block represents observations in jth individual, whose k subcolumn is a $k-1$ lagged variable.

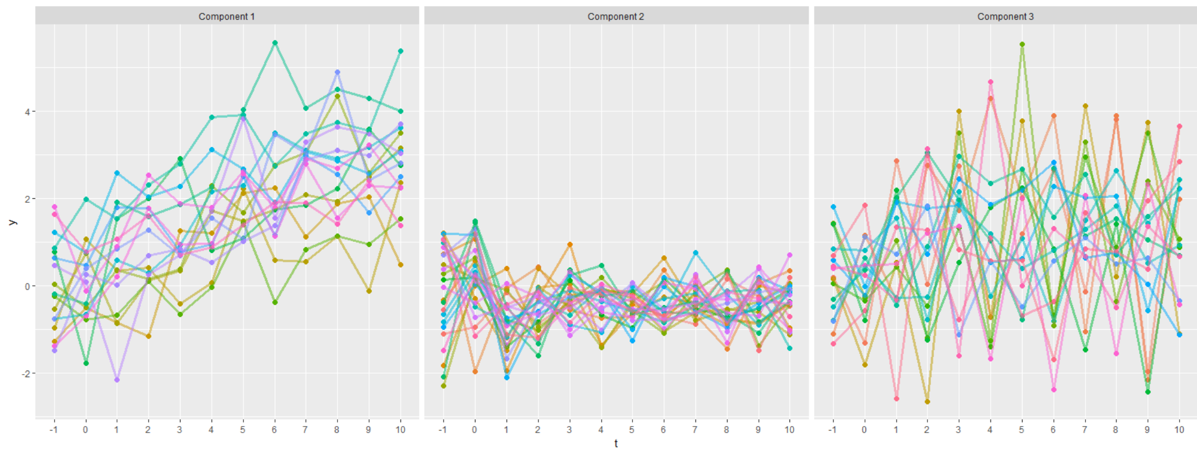
- `MDP.model` is an object of `MDP.model` class that represents a mixture dyanmic panel model.

Instances with `MDP.model` class can be used to generate plots by calling `PlotMDPModel`. The following code will generate two plots in Figure 1:

```
mdp.model <- sample$MDP.model # extract MDP.model from the sample
PlotMDPModel(mdp.model = mdp.model)
# create plots for each component
PlotMDPModel(mdp.model = mdp.model, separate = TRUE)
```



(a) With `separate = FALSE`



(b) With `separate = TRUE`

Figure 1: Plots generated by `PlotMDPModel`.