$$\theta^*(t) = \arg\min_{\theta(t)} \int_0^T g(x(t), \theta(t))\ dt \quad \approx \quad \arg\min_{\theta(t)} \sum_{t=0}^{T} g(x_t, \theta(t))$$

error of the model on $[0, T]$

$\theta(t)$ is constant on clusters

$$[\Theta^*, \Gamma^*(t)] = \arg\min_{\Theta, \Gamma} \int_0^T \sum_{k=1}^{K} \gamma_k(t) \cdot g(x(t), \theta_k)\ dt \quad \approx \quad \arg\min_{\Theta, \Gamma} \sum_{t=0}^{T} \sum_{k=1}^{K} \gamma_{k,t} \cdot g(x_t, \theta_k)$$

$\theta_1, \theta_2, \ldots, \theta_k$ - model parameters on clusters

$\gamma_1(t), \gamma_2(t), \ldots, \gamma_k(t)$ - cluster indicator functions

$\theta_1 \quad \theta_2 \quad \theta_3 \quad \gamma_1(t) \quad \gamma_2(t) \quad \gamma_3(t)$

continuous real-valued functions $\gamma_k(t) \in [0, 1]$

smooth $\gamma_k(t) \quad \Rightarrow \quad$ regularisation

$$\gamma_k(t) \in \{0, 1\}, \quad \gamma_k(t) = \begin{cases} 1 & \text{if } k\text{-th cluster is active in } t \\ 0 & \text{if } k\text{-th cluster is inactive in } t \end{cases}$$

$$0 \leq \gamma_k(t) \leq 1, \quad \forall t : \sum_{k=1}^{K} \gamma_k(t) = 1$$

$$\text{s.t. } \ 0 \leq \gamma_k(t) \leq 1, \ \forall t : \sum_{k=1}^{K} \gamma_k(t) = 1$$

$$\int_0^T \sum_{k=1}^{K} \gamma_k(t) \cdot g(x(t), \theta_k)\ dt + \varepsilon^2 \sum_{k=1}^{K} \int_0^T (\partial_t \gamma_k(t))^2 dt \quad \approx \quad \sum_{t=0}^{T} \sum_{k=1}^{K} \gamma_{k,t} \cdot g(x_t, \theta_k) + \varepsilon^2 \sum_{k=1}^{K} \sum_{t=0}^{T-1} (\gamma_{k,t+1} - \gamma_{k,t})^2$$

$$[\Theta^*, \Gamma^*] = \arg\min_{\Theta, \Gamma} \sum_{t=0}^{T} \sum_{k=1}^{K} \gamma_{k,t} \cdot g(x_t, \theta_k) + \varepsilon^2 \sum_{k=1}^{K} \sum_{t=0}^{T-1} (\gamma_{k,t+1} - \gamma_{k,t})^2$$

$$\text{s.t. } \ 0 \leq \gamma_{k,t} \leq 1, \ \forall t : \sum_{k=1}^{K} \gamma_{k,t} = 1$$

$L(\Theta, \Gamma)$

set feasible initial approximation Γ_0

while $\|L(\Theta_{it},\Gamma_{it}) - L(\Theta_{it-1},\Gamma_{it-1})\| > \varepsilon_L$
solve $\Theta_{it} = \arg\min_{\Theta} L(\Theta,\Gamma_{it})$ *(with fixed* Γ_{it}*)*
solve $\Gamma_{it} = \arg\min_{\Gamma} L(\Theta_{it},\Gamma)$ *(with fixed* Θ_{it}*)*
$it = it + 1$
endwhile

$$\begin{array}{rcl} \theta_1 & = & [0,0,0] \\ \theta_2 & = & [1,0,0] \\ \theta_3 & = & [0,1,0] \end{array}$$

$$\sum_{t=0}^{T} \|x_t - \theta(t)\|^2 \ \rightarrow \ \min$$

$y^t \in \{0,1\}$: binary data

$P_X^t = (P(x_1^t),...,P(x_n^t),1)^T$ probability vector

Log-likelihood at time t:

$$\mathcal{L}(t) = y^t \log\left(\Lambda_X(t)^T P_X^t\right) + (1-y^t)\log\left(1-\Lambda_X(t)^T P_X^t\right)$$

Want to solve the estimation problem

$$\max_{\Lambda_X(t):\ t\in[0,T]} \int_0^T \mathcal{L}(t)dt$$

such that:

$$0 < \Lambda_X(t)^T P_X^t < 1, \quad \forall t$$

$$Y^t = a + bY^{t-1} + cX^{t-1} + \text{noise}$$

$$c \cong 0 \quad c \neq 0 \quad \Lambda \cong 0 \quad \Lambda = P[y^t|x^t] \quad \varepsilon = P[y^t \text{ and NOT } x^t] \quad y^t \quad X \quad x^t \quad Y$$

$$X \quad \text{not } X \quad Y \quad \{y^t\}, t=0,\dots,T \quad \{x^t\}, t=0,\dots,T$$

$$\begin{array}{rcl} P[y^t] & = & P[y^t \text{ and } x^t] + P[y^t \text{ and NOT } x^t] \\ & = & \Lambda P[x^t] + \varepsilon \end{array}$$

X has a causality impact on Y if
for any t, event y^t is happening if (and only if) event x^t happened.

fit a stochastic model for Y and X, e.g. a linear VAR-X:

If model with $c \cong 0$ is more statistically-significant than the one with $c \neq 0$

$$X \text{ has no } \textbf{Granger-causality} \text{ impact on } Y$$

Law of the Total Probability:

If $\Lambda \cong 0$ there is no causality impact from X on Y.

$$\begin{gathered} P[y^t] = \Lambda_1(t)P[x_1^t] + \Lambda_2(t)P[x_2^t] + \cdots + \Lambda_n(t)P[x_n^t] \\ +P[y^t \text{ and not}(x_1^t \text{ or } \ldots \text{ or } x_n^t)] \end{gathered}$$

where $\Lambda_i(t) = P[y^t|x_i^t]$ is possibly time-dependent due to unresolved scales.

If $\Lambda_i \cong 0$ there is no causality impact from X_i on Y.