1 Model

Setting: we consider a panel of Dirichelet observations with cross sectional units $i = 1, \ldots, N$ at time points $t = 1, \ldots, T$:

- $\boldsymbol{y}_{it} = (y_{it1}, \ldots, y_{itD})$: vector of $d = 1, \ldots, D$ time dependent proportions/shares to be modelled for each individual unit i
- The likelihood of all observations given the parameters is

$$L\left(\underline{\boldsymbol{\alpha}}; \boldsymbol{y}_{11}, \ldots, \boldsymbol{y}_{NT}\right) = \prod_{t=1}^{T} \prod_{i=1}^{N} \mathcal{B}\left(\exp\left(x_{it1}\right), \ldots, \exp\left(x_{itD}\right)\right)^{-1} \times \prod_{d=1}^{D} \left(y_{i,t,d}\right)^{\exp(x_{itd})-1} .$$

- Here, $\underline{\boldsymbol{\alpha}} = (\alpha_{11}, \ldots, \alpha_{1T}, \alpha_{21}, \ldots, \alpha_{2T}, \alpha_{N1}, \ldots, \alpha_{NT})$. For each individual unit i, shares $\alpha_{it} = (\alpha_{it1}, \ldots, \alpha_{itD})$ will be modeled componentwise as a latent stochastic processes over time on the logarithmic scale. More precisely, fixing the cross-sectional unit i, for $t = 1, \ldots, T$, and each component $d = 1, \ldots, D$ we have $\log\left(\alpha_{itd}\right) = x_{itd}$. The loglink ensures the positive scale of the Dirichelet parameters i.e. that $\alpha_{itd} > 0$.
- For each cross sectional unit i, latent state transitions of each parameter component d follows some autoregressive process over time:

$$x_{itd} = \phi_x x_{i(t-1)d} + \left(\boldsymbol{z}_{itd}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z}} + \left(\boldsymbol{z}_{itd}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z}} + \varepsilon_{itd} \ , \ \varepsilon_{itd} \sim \mathcal{N}\left(0, \sigma_{\varepsilon}^2\right) \ ,$$

  In general, the parameters $\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}, \boldsymbol{\gamma}_{\boldsymbol{Z}}, \sigma_{\varepsilon}^2$ will vary with different α_d's but for notational convenience we have omitted dependence on the index d for the moment. Regressors include $\boldsymbol{z}_{itd}^{lin}$ to model linear effects on the response as well as nonlinear covariate effects $\boldsymbol{z}_{itd}^{spl}$ via polynomial splines.

2 KZ model

2.1 Mathematical model

The latent state transitions are

$$\begin{aligned}
x_{it1} &= \phi_1 x_{i(t-1)1} + \left(\boldsymbol{z}_{it1}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_1}} + \left(\boldsymbol{z}_{it1}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_1}} + \varepsilon_1 \ , \varepsilon_1 \sim \mathcal{N}\left(0, \sigma_{\varepsilon_1}^2\right) \ , \\
x_{it2} &= \phi_2 x_{i(t-1)2} + \left(\boldsymbol{z}_{it2}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_2}} + \left(\boldsymbol{z}_{it2}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_2}} + \varepsilon_2 \ , \varepsilon_2 \sim \mathcal{N}\left(0, \sigma_{\varepsilon_2}^2\right) \ , \\
&\vdots \\
x_{itD} &= \phi_D x_{i(t-1)D} + \left(\boldsymbol{z}_{itD}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_D}} + \left(\boldsymbol{z}_{itD}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_D}} + \varepsilon_D \ , \varepsilon_D \sim \mathcal{N}\left(0, \sigma_{\varepsilon_D}^2\right) \ ,
\end{aligned}$$

where there are D different parameters $\alpha_{it1}, \ldots, \alpha_{itD}$ of the Dirichelt distribution $\forall i, t$ with link $x_{itd} = \log(\alpha_{itd})$. The parameters of the state processes are grouped as

$$\begin{aligned}
\boldsymbol{\theta}_1 &= \left(\phi_1, \boldsymbol{\beta}_{\boldsymbol{Z_1}}, \boldsymbol{\gamma}_{\boldsymbol{Z_1}}, \sigma_{\varepsilon_1}^2\right) \ , \\
\boldsymbol{\theta}_2 &= \left(\phi_2, \boldsymbol{\beta}_{\boldsymbol{Z_2}}, \boldsymbol{\gamma}_{\boldsymbol{Z_2}}, \sigma_{\varepsilon_2}^2\right) \ , \\
&\vdots \\
\boldsymbol{\theta}_D &= \left(\phi_D, \boldsymbol{\beta}_{\boldsymbol{Z_D}}, \boldsymbol{\gamma}_{\boldsymbol{Z_D}}, \sigma_{\varepsilon_D}^2\right) \ , \\
\underline{\boldsymbol{\theta}} &= \left(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D\right) \ .
\end{aligned}$$

The measurement equation is given as the dirichlet density

$$p_{\underline{\boldsymbol{\theta}}}\left(\boldsymbol{y}_{it} | x_{it1}, \ldots, x_{itD}\right) = \mathcal{B}\left(\exp\left(x_{it1}\right), \ldots, \exp\left(x_{itD}\right)\right)^{-1} \times \prod_{d=1}^{D} \left(y_{itd}\right)^{\exp(x_{itd})-1} \ .$$

This implies $\forall\ t = 1, \ldots, T$ and $\forall\ i = 1, \ldots, N$

$$\begin{aligned}
p_{\boldsymbol{\theta}_1}\left(x_{it1} | x_{i(t-1)1}, \boldsymbol{z}_{it1}^{lin}, \boldsymbol{z}_{it1}^{spl}\right) &= \mathcal{N}\left(x_{it1} \mid \mu_{it1}; \sigma_{\varepsilon_1}^2\right) \ , \\
p_{\boldsymbol{\theta}_2}\left(x_{it2} | x_{i(t-1)2}, \boldsymbol{z}_{it2}^{lin}, \boldsymbol{z}_{it2}^{spl}\right) &= \mathcal{N}\left(x_{it2} \mid \mu_{it2}; \sigma_{\varepsilon_2}^2\right) \ , \\
&\vdots \\
p_{\boldsymbol{\theta}_D}\left(x_{itD} | x_{i(t-1)D}, \boldsymbol{z}_{itD}^{lin}, \boldsymbol{z}_{itD}^{spl}\right) &= \mathcal{N}\left(x_{itD} \mid \mu_{itD}; \sigma_{\varepsilon_D}^2\right) \ , \\
p_{\underline{\boldsymbol{\theta}}}\left(\boldsymbol{y}_{it} | x_{it1}, x_{it2}, \ldots, x_{itD}\right) &= \mathbf{Dir}\left(\boldsymbol{y}_{it} | \exp\left(x_{it1}\right), \ldots, \exp\left(x_{itD}\right)\right) \ ,
\end{aligned}$$

where $\mu_{itd} = \phi_d x_{i(t-1)d} + \left(\boldsymbol{z}_{itd}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_d}} + \left(\boldsymbol{z}_{itd}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_d}}$ for $d = 1, \ldots, D$.

2.2 Gibbs-Part: univariate (for one $i = 1, \dots, N$)

W.l.o.g. we can discard the dependence on α_d and i since the derivations are similar $\forall i, d$. Thus, let $p(\boldsymbol{\theta}|x_{0:T}, \boldsymbol{y}_{1:T})$ for some $x_{0:T} \equiv x_{i(0:T)d,\cdot}$. The full probabilistic model with parameters $\boldsymbol{\theta} = (\phi, \boldsymbol{\beta_Z}, \boldsymbol{\gamma_Z}, \sigma_\varepsilon^2)$ can be factorized as

$$
\begin{aligned}
p(\boldsymbol{\theta}, x_{0:T}, \boldsymbol{y}_{1:T}) = p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}, x_{0:T})\, p(x_{0:T}, \boldsymbol{\theta}) &= \prod_{t=1}^{T} p(\boldsymbol{y}_t|x_t) \prod_{t=1}^{T} p(x_t|x_{t-1}, \boldsymbol{\theta})\, p(x_0|\boldsymbol{\theta})\, p(\boldsymbol{\theta}) \\
&= \prod_{t=1}^{T} p(\boldsymbol{y}_t|x_t) \times \frac{1}{(2\pi\sigma_\varepsilon^2)^{T/2}} \\
&\times \prod_{t=1}^{T} \exp\left(-\frac{\left(x_t - \phi x_{t-1} - \left(\boldsymbol{z}_t^{lin}\right)' \boldsymbol{\beta_Z} - \left(\boldsymbol{z}_t^{spl}\right)' \boldsymbol{\gamma_Z}\right)^2}{2\sigma_\varepsilon^2} \right) \times p(x_0|\boldsymbol{\theta}) \times p(\boldsymbol{\theta}) \ .
\end{aligned}
$$

For $\boldsymbol{\beta} = (\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}', \boldsymbol{\gamma}_{\boldsymbol{Z}}')'$ we assume a normal prior $\boldsymbol{\beta} \sim \mathcal{N}_{K_\beta + K_\gamma + 1}\left(\underline{\boldsymbol{\beta}}, \underline{\boldsymbol{\Omega}_\beta}\right)$ where the dimensions of $\boldsymbol{z}_t^{lin}$ and $\boldsymbol{z}_t^{spl}$ are given as K_β and K_γ, respectively. In our settings, it suffices to set priors according to

$$
\underline{\boldsymbol{\beta}} = (0, 0, \dots, 0) \ , \quad \underline{\boldsymbol{\Omega}_\beta} = \begin{bmatrix} \underline{\boldsymbol{\Omega}_{lin}} & \boldsymbol{0} \\ \boldsymbol{0} & \underline{\boldsymbol{\Omega}_{spl}} \end{bmatrix} \ ,
$$

$$
\underline{\boldsymbol{\Omega}_{lin}} = \begin{bmatrix} \tau_\phi^2 & 0 & \dots 0 \\ 0 & \tau_{\beta_Z}^2 & \dots 0 \\ \dots & \dots & \dots \\ 0 & 0 & \tau_{\beta_Z}^2 \end{bmatrix}_{(K_\beta+1)\times(K_\beta+1)} \ , \quad \underline{\boldsymbol{\Omega}_{spl}} = \left[splinesPrior \right]_{K_\gamma \times K_\gamma} \ .
$$

An uninformative prior at $(\phi, \boldsymbol{\beta_Z})$ would let $\tau_\phi^2, \tau_{\beta_Z}^2 \to \infty$ s.th. $\underline{\boldsymbol{\Omega}_{lin}}^{-1} \to \boldsymbol{0}$, where $\boldsymbol{0}$ is the zero matrix. For σ_ε^2 we assume an $\mathcal{IG}(a_\sigma, b_\sigma)$-prior with $a_\sigma = b_\sigma = 0.001$

$$
p\left(\sigma_\varepsilon^2\right) = \frac{b_\varepsilon^{a_\varepsilon}}{\Gamma(a_\varepsilon)} \left(\sigma_\varepsilon^2\right)^{-a_\varepsilon - 1} \exp\left(-\frac{b_\varepsilon}{\sigma_\varepsilon^2} \right) \ .
$$

Then, all conditional parameter distributions are conjugate, as will be shown now.

The full conditional for σ_ε^2 is obtained as

$$p\left(\sigma_\varepsilon^2 | x_{0:T}, \boldsymbol{y}_{1:T}, \boldsymbol{\beta}\right) = \frac{1}{\left(2\pi\sigma_\varepsilon^2\right)^{T/2}} \prod_{t=1}^{T} \exp\left(-\frac{\left(x_t - \boldsymbol{\mu}_x\right)^2}{2\sigma_\varepsilon^2}\right) \times \frac{b_\varepsilon^{a_\varepsilon}}{\Gamma\left(a_\varepsilon\right)} \left(\sigma_\varepsilon^2\right)^{-a_\varepsilon - 1} \exp\left(-\frac{b_\varepsilon}{\sigma_\varepsilon^2}\right)$$

$$\propto \left(\sigma_\varepsilon^2\right)^{-(a_\varepsilon + T/2) - 1} \times \exp\left(-\frac{1}{\sigma_\varepsilon^2}\left(b_\varepsilon + \frac{\sum_{t=1}^{T}\left(x_t - \boldsymbol{\mu}_x\right)^2}{2}\right)\right) \ ,$$

where $\boldsymbol{\mu}_x = \phi x_{t-1} + \left(\boldsymbol{z}_t^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z}} + \left(\boldsymbol{z}_t^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z}}$. With e.g. $a_\sigma = b_\sigma = 0.001$, we have

$$\sigma_\varepsilon^2 \sim \mathcal{IG}\left(a_\sigma^*, b_\sigma^*\right) \ , \ a_\sigma^* = a_\sigma + T/2 \ , \ b_\sigma^* = b_\sigma + \frac{\sum_{t=1}^{T}\left(x_t - \boldsymbol{\mu}_x\right)^2}{2} \ .$$

The full conditional for $\boldsymbol{\beta}$ is obtained easier if one changes to matrix notation. Let

$$x_{2:T} = \boldsymbol{Z}_{2:T}\boldsymbol{\beta} + \varepsilon_{2:T} \ ,$$

$$x_{2:T} = \begin{bmatrix} x_2 \\ \vdots \\ x_T \end{bmatrix} \ , \quad x_{1:T-1} = \begin{bmatrix} x_1 \\ \vdots \\ x_{T-1} \end{bmatrix} \ , \quad \varepsilon_{2:T} = \begin{bmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix} \ , \quad \mathbb{V}\left(\varepsilon_{2:T}\right) = \boldsymbol{\Omega}_\varepsilon = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \dots & 0 \\ 0 & \sigma_\varepsilon^2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ \dots & \dots & \dots & \sigma_\varepsilon^2 \end{bmatrix} \ ,$$

$$\boldsymbol{Z}_{2:T} = \left[x_{1:T-1}, \boldsymbol{Z}_{2:T}^{lin}, \boldsymbol{Z}_{2:T}^{spl}\right] \ , \quad \boldsymbol{Z}_{2:T}^{lin} = \begin{bmatrix} z_{21} & z_{22} & \dots & z_{2K_\beta} \\ z_{31} & z_{32} & \dots & z_{3K_\beta} \\ \vdots & \vdots & \ddots & \vdots \\ z_{T1} & z_{12} & \dots & z_{TK_\beta} \end{bmatrix} \ , \quad \boldsymbol{Z}_{2:T}^{spl} = \left[splines\right] \ .$$

To derive the full conditional Gibbs block for $\boldsymbol{\beta}$ we start as follows:

$$\begin{aligned} p\left(\boldsymbol{\beta}|x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma_\varepsilon^2\right) \propto \exp &\left\{-\frac{1}{2}\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}\right)' \boldsymbol{\Omega}_\varepsilon^{-1}\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}\right)\right\} \\ \times \exp &\left\{-\frac{1}{2}\left(\boldsymbol{\beta} - \underline{\boldsymbol{\beta}}\right)' \underline{\boldsymbol{\Omega}_\beta}^{-1}\left(\boldsymbol{\beta} - \underline{\boldsymbol{\beta}}\right)\right\} \end{aligned}$$

Now we have

$$\begin{aligned} \left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}\right)' \boldsymbol{\Omega}_\varepsilon^{-1}\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}\right) &= \boldsymbol{\beta}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}\boldsymbol{Z}_{2:T}\boldsymbol{\beta} - 2\boldsymbol{\beta}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}x_{2:T} \\ &\quad + x_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}x_{2:T} \\ \left(\boldsymbol{\beta} - \underline{\boldsymbol{\beta}}\right)' \underline{\boldsymbol{\Omega}_\beta}^{-1}\left(\boldsymbol{\beta} - \underline{\boldsymbol{\beta}}\right) &= \boldsymbol{\beta}\underline{\boldsymbol{\Omega}_\beta}^{-1}\boldsymbol{\beta} - 2\boldsymbol{\beta}\underline{\boldsymbol{\Omega}_\beta}^{-1}\underline{\boldsymbol{\beta}} \\ &\quad + \underline{\boldsymbol{\beta}}\underline{\boldsymbol{\Omega}_\beta}^{-1}\underline{\boldsymbol{\beta}} \ . \end{aligned}$$

But then $\boldsymbol{\beta}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}\boldsymbol{Z}_{2:T}\boldsymbol{\beta} + \boldsymbol{\beta}\underline{\boldsymbol{\Omega}_\beta}^{-1}\boldsymbol{\beta} = \boldsymbol{\beta}\left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}\boldsymbol{Z}_{2:T} + \underline{\boldsymbol{\Omega}_\beta}^{-1}\right]\boldsymbol{\beta}$ implies

$$\overline{\boldsymbol{\Omega}_\beta} = \left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}\boldsymbol{Z}_{2:T} + \underline{\boldsymbol{\Omega}_\beta}^{-1}\right]^{-1} \ ,$$

and $2\boldsymbol{\beta}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}x_{2:T} - 2\boldsymbol{\beta}\underline{\boldsymbol{\Omega}_\beta}^{-1}\underline{\boldsymbol{\beta}} = -2\boldsymbol{\beta}\left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}x_{2:T} + \underline{\boldsymbol{\Omega}_\beta}^{-1}\underline{\boldsymbol{\beta}}\right]$ implies

$$\overline{\boldsymbol{\beta}} = \overline{\boldsymbol{\Omega}_\beta} \times \left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_\varepsilon^{-1}x_{2:T} + \underline{\boldsymbol{\Omega}_\beta}^{-1}\boldsymbol{\beta}\right] \ .$$

Finally, we have

$$p\left(\boldsymbol{\beta}|x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma_\varepsilon^2\right) = \mathcal{N}_{K+1}\left(\overline{\boldsymbol{\beta}}, \overline{\boldsymbol{\Omega}_\beta}\right) \ .$$