1 generate_time_synthetic_data

We generate a multivariate spatio-temporal process $\{y_{t,i}\}_{t=0}^{T-1}$ observed at N spatial locations $\{s_i\}_{i=1}^{N}$ over T time steps. The data-generating process consists of a global mean component, latent temporal processes, a fixed spatial basis with time-varying amplitude, covariate-driven effects, and additive observation noise.

Spatial locations. Spatial locations are one-dimensional and fixed over time:

$$s_i \in \mathbb{R}, \qquad i = 1, \dots, N.$$

Spatial basis. A single spatial basis vector is constructed using a Gaussian kernel and normalized to unit ℓ_2 norm:

$$\tilde{\phi}_i = \exp(-s_i^2), \qquad \phi_i = \frac{\tilde{\phi}_i}{\left(\sum_{j=1}^{N} \tilde{\phi}_j^2\right)^{1/2}}.$$

This basis defines a fixed spatial pattern shared across all time points.

Latent temporal processes. Two independent latent AR(1) processes are introduced:

$$\eta_0 \sim \mathcal{N}(0,1), \qquad \eta_t = \rho_\eta \, \eta_{t-1} + \omega_t, \qquad \omega_t \sim \mathcal{N}(0, \sigma_\eta^2),$$

$$f_0 \sim \mathcal{N}(0,1), \qquad f_t = \rho_f \, f_{t-1} + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, \sigma_f^2),$$

for $t \geq 1$. The process η_t controls the temporal evolution of the spatial effect, while f_t represents a global temporal drift shared across all locations. *In the implementation, the innovation variances are fixed to $\sigma_\eta^2 = \sigma_f^2 = 0.1^2$.*

Covariates. At each time t and location i, a p-dimensional continuous covariate vector $x_{t,i} \in \mathbb{R}^p$ is observed.

Covariate effects. Covariates affect the target through an additive function

$$h(x_{t,i}) = \beta^\top x_{t,i} + \kappa \, \psi(x_{t,i}),$$

where β denotes fixed linear coefficients and $\psi(\cdot)$ is a predefined nonlinear transformation. The scalar $\kappa \geq 0$ controls the strength of the nonlinear contribution.

Observation noise. Observation errors are assumed to be independent Gaussian noise:

$$\varepsilon_{t,i} \sim \mathcal{N}(0, \sigma^2).$$

Observation equation. The target variable is generated according to

$$\boxed{y_{t,i} = \mu + f_t + \lambda\, \eta_t\, \phi_i + h(x_{t,i}) + \varepsilon_{t,i}}, \qquad t = 0, \dots, T-1,\ i = 1, \dots, N.$$

Parameter definitions. The parameters in the baseline data-generating process are summarized as follows:

- μ: global mean level of the target process.
- λ: amplitude (eigenvalue) of the spatial effect, controlling the overall strength of the spatial pattern.
- ρ_η: autoregressive coefficient of the latent spatial amplitude process η_t.
- ρ_f: autoregressive coefficient of the global temporal drift f_t.
- σ_η^2: innovation variance of the spatial amplitude process (fixed to 0.1^2 in the implementation).
- σ_f^2: innovation variance of the global temporal drift (fixed to 0.1^2 in the implementation).
- σ^2: observation noise variance.
- β: linear coefficients for the covariate effects.
- κ: strength of the nonlinear covariate contribution, corresponding to `non_linear_strength` in the implementation.
- p: dimension of the continuous covariate vector $x_{t,i}$.
- N: number of spatial locations.
- T: total number of time steps.

2 generate_time_NAR_synthetic_data

This data-generating process extends the baseline `generate_time_synthetic_data` by adding a nonlinear autoregressive (NAR) component with a global, time-varying gate. The spatial structure, latent temporal processes, and covariate effects are identical to the baseline model.

Spatial and latent components. Let $s_i \in \mathbb{R}$ denote the location of site i. A normalized Gaussian spatial basis is defined as

$$\tilde{\phi}_i = \exp(-s_i^2), \qquad \phi_i = \frac{\tilde{\phi}_i}{\left(\sum_{j=1}^{N} \tilde{\phi}_j^2\right)^{1/2} + 10^{-12}}.$$

Latent temporal processes follow AR(1) dynamics,

$$\eta_t = \rho_\eta \eta_{t-1} + \omega_t, \qquad f_t = \rho_f f_{t-1} + \nu_t,$$

with Gaussian innovations. *In the implementation, the innovation variances are fixed to* $\omega_t \sim \mathcal{N}(0, 0.1^2)$ *and* $\nu_t \sim \mathcal{N}(0, 0.1^2)$.

Covariates. Each site has three continuous covariates $x_{t,i} \in \mathbb{R}^3$. The covariate effect is

$$h(x_{t,i}) = 0.3\, x_{t,i}^{(1)} + 0.4\, x_{t,i}^{(2)} + 0.2\, x_{t,i}^{(3)} + \kappa\Big((x_{t,i}^{(1)})^2 + 0.5\, x_{t,i}^{(1)} x_{t,i}^{(2)} + 0.3\, \sin(x_{t,i}^{(3)}) \Big).$$

Baseline signal. The baseline spatio-temporal signal is

$$m_{t,i} = \mu + f_t + \lambda \eta_t \phi_i + h(x_{t,i}).$$

Rolling-scale gated autoregression. Let W denote the rolling window length and $w_0(t) = \max\{0, t - W\}$. Define the rolling mean and scale

$$\bar{y}_t = \frac{1}{(t - w_0(t))N} \sum_{k=w_0(t)}^{t-1} \sum_{i=1}^{N} y_{k,i}, \qquad s_t = \sqrt{\frac{1}{(t - w_0(t))N} \sum_{k=w_0(t)}^{t-1} \sum_{i=1}^{N} (y_{k,i} - \bar{y}_t)^2 + \epsilon}.$$

A global gate is constructed as

$$z_t = \frac{\frac{1}{N}\sum_{i=1}^{N} y_{t-1,i} - \bar{y}_t}{s_t}, \qquad g_t = \mathbb{I}\{z_t > \tau\}.$$

The effective autoregressive coefficient is

$$\rho_t^{\text{eff}} = \rho_y\big(\rho_{\text{low}} + (\rho_{\text{high}} - \rho_{\text{low}}) g_t\big),$$

and the nonlinear autoregressive term is

$$r_{t,i} = \rho_t^{\text{eff}}\, y_{t-1,i} + \gamma \tanh\left(\frac{y_{t-1,i}}{s_t}\right).$$

Observation equation. The final observation model is

$$y_{t,i} = m_{t,i} + r_{t,i} + \varepsilon_{t,i}, \qquad \varepsilon_{t,i} \sim \mathcal{N}(0, \sigma^2),$$

with $y_{0,i}$ initialized analogously to the baseline model.

Additional parameters for the NAR extension. The nonlinear autoregressive extension introduces the following additional parameters, which are not present in the baseline model:

- ρ_y: base linear autoregressive coefficient controlling the dependence on the previous observation $y_{t-1,i}$.
- γ: strength of the nonlinear autoregressive component based on the hyperbolic tangent function. When $\gamma = 0$, the model reduces to a gated linear autoregressive process.
- W: rolling window length used to compute the global rolling mean $\bar{y}_t$ and scale s_t.
- τ: threshold parameter for the global gate. The gate is activated when the standardized global signal z_t exceeds τ.
- ρ_{low}: scaling factor applied to ρ_y when the gate is inactive ($g_t = 0$).
- ρ_{high}: scaling factor applied to ρ_y when the gate is active ($g_t = 1$).
- ϵ: a small positive constant added to the rolling scale s_t for numerical stability (fixed to 10^{-6} in the implementation).