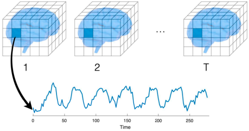
Flexible Large-Scale Time Series Regression

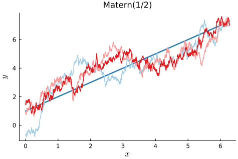
Neuroimaging data

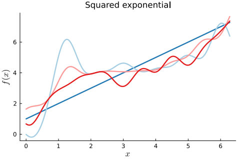
- 100,000 time series, one for each brain voxel.
- **complicated noise structure**. Breathing, heartbeats.



Automatic flexible time series regression for each voxel

$$y(t) = \boldsymbol{x}(t)^\top \boldsymbol{\beta} + \varepsilon(t)$$
$$\varepsilon(t) \sim \mathrm{GP}(0, k(t, t'))$$





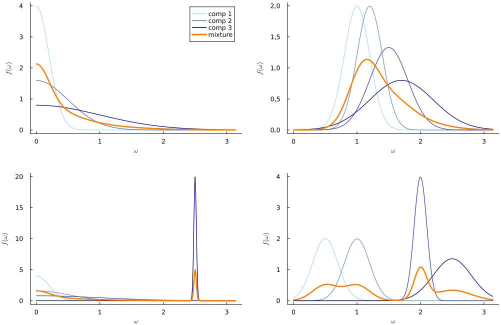
Spectral mixture kernels

- Kernel of stationary process $\iff$ spectral density

$$k(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} f(\omega)\, d\omega$$

- Spectral mixture kernel: $f(\omega)$ is mixture of Gaussians

$$f_{\boldsymbol{\theta}}(\omega) = \sum_{j=1}^{J} \nu_j \cdot \mathcal{N}\left(\omega \mid \mu_j, \sigma_j^2\right), \quad \boldsymbol{\theta} = (\nu_j, \mu_j, \sigma_j^2)_{j=1}^{J}$$



Bayesian inference

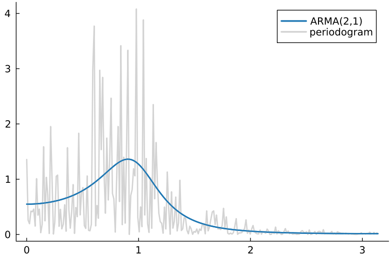
- **Time domain likelihood** for discretely sampled data

$$\boldsymbol{y}|\boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{X} \sim \mathcal{N}\left(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}(\boldsymbol{\theta})\right)$$

  $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is the noise covariance obtained from the kernel $k_{\boldsymbol{\theta}}(\tau)$.

- How to infer the hyperparameters $\boldsymbol{\theta}$?
- **Whittle likelihood** from **periodogram data**

$$\hat{f}_k \overset{\text{indep}}{\sim} \text{Exp}(f(\omega_k))$$



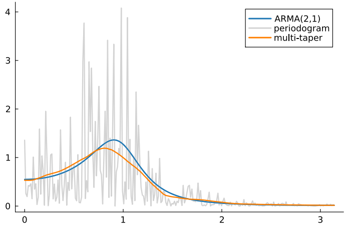
Bayesian inference

- **Robust Whittle likelihood** from **multi-taper estimate**

$$\tilde{f}_k \overset{\text{indep}}{\sim} \text{Gamma}(\alpha_k, f(\omega_k)/\alpha_k)$$

$$\alpha_k = \exp(\gamma_0 + \gamma_1 \cdot n_k)$$

and n_k is number of tapers at ω_k.



- **Regularization priors** to avoid overfitting:
  - **Horseshoe**-like shrinkage prior for **mixture weights** ν_j
  - Cliff-distribution to **penalize small component variances**