Elicitation of Experts' Knowledge for Functional Linear Regression

Grollemund, P.-M., Abraham, C., Pudlo, P. and Baragatti, M.
paul-marie.grollemund@umontpellier.fr

UNIVERSITÉ DE MONTPELLIER

Montpellier SupAgro

Abstract

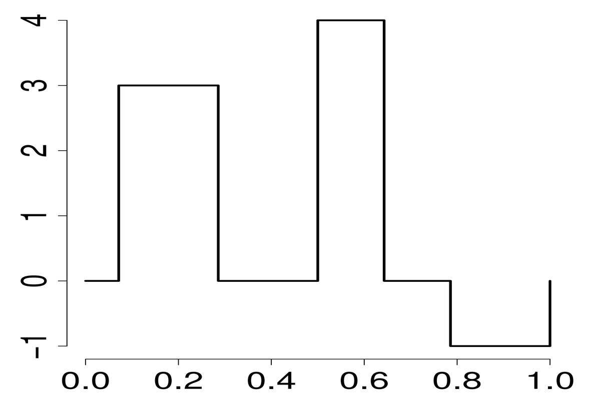
We present an approach to elicit experts' knowledge about the Bliss model [1] which is a parcimonious Bayesian Functional Linear Regression model. We derive an informative prior from elicited information and we define weights to tune prior information contribution on the estimators.

R Package

Available at
`github.com/pmgrollemund/bliss/`

Sparse Step function

$$\beta_\theta(t) = \sum_{k=1}^{K} b_k \mathbf{1}_{[a_k,b_k]}(t) \qquad (1)$$



Bliss

Bayesian Functional **L**inear Regression with **S**parse **S**tep functions aggregates "weak learners" in a Bayesian way.

Data $\mathcal{D}$**:** $y_i, x_i(\cdot)$ where $x_i(\cdot) \in L^2([0,1])$

Model: see [2, 5]

$$y_i | x_i(\cdot), \theta \sim \mathcal{N}\left(\mu + \int_0^1 x_i(t)\beta_\theta(t)\, dt, \sigma^2\right) \qquad (2)$$

where $\beta_\theta(\cdot)$ is given in (1).

Noninformative Prior $\pi_0(\cdot)$ **and MCMC:** [1]

Bayesian Estimator with L^2**-loss:**

$$\hat{\beta}(t) = \int \pi_0(\theta|\mathcal{D})\beta_\theta(t)\, \mathrm{d}\theta \qquad (3)$$

where $\pi_0(\cdot|\mathcal{D})$ is the posterior and β_θ as in (1).

Elicitation

Aim:
Collect experts' knowledge [4] about the coefficient function

Various experts provide: $\mathcal{D}^e = (y_i^e, x_i^e(\cdot), c_i^e)$ for $i = 1, \ldots, n_e$

- pseudo data: y_i^e and $x_i^e(\cdot)$
- certainty: c_i^e for each pseudo observation

Informative prior

Informative prior is defined as a fractional posterior [3] for which:

- the *initial* prior is the vaguely noninformative prior $\pi_0(\cdot)$ and
- the pseudo data model is (2)

$$\pi_E(\theta) = \pi(\theta|\mathcal{D}^1, \ldots, \mathcal{D}^E; w) \propto \pi_0(\theta) \prod_{e=1}^{E} p(\mathcal{D}^e|\theta; w) \quad \text{where} \quad p(\mathcal{D}^e|\theta; w) = \prod_{i=1}^{n_e} p(y_i^e | x_i^e(\cdot), \theta)^{w_i^e}$$

Interpretation : a sequential learning approach

- from initial prior, learn from pseudo data to derive an informative prior
- from informative prior, learn from observed data to derive a posterior

Tuning Weights

Naive approach: $w_i^e = c_i^e$

- Overconfidence Bias, see [4]

Deriving weights from important properties:

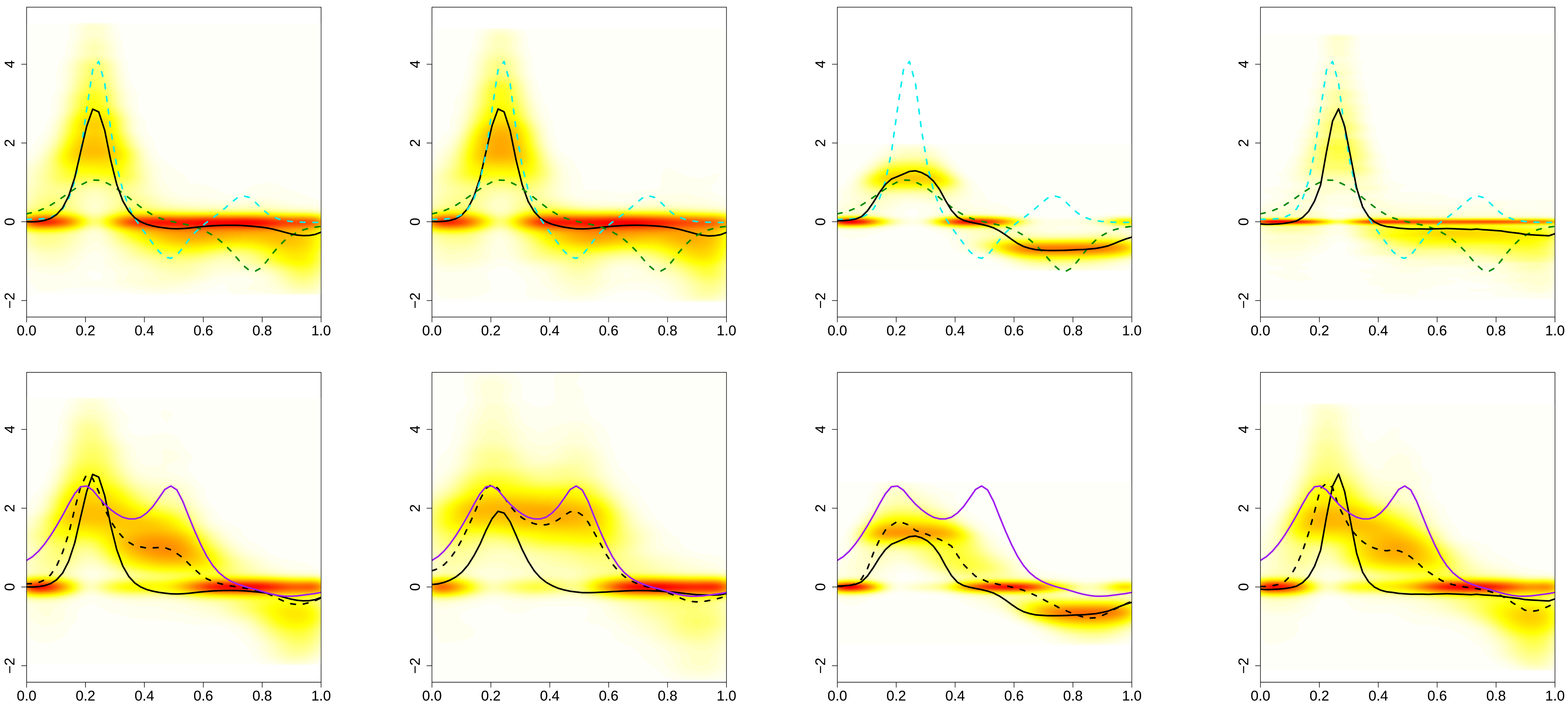
- Experts' interactions $\quad (r_{e,f} \in [0,1]$ is the dependence between expert e and expert $f)$
- pseudo data weight $\leq$ observed data weight

$$w_i^e = c_i^e \times \text{(interaction)} \times \text{(lower ps. data weight)} = c_i^e \times \frac{1}{1 + \sum_{f \neq e} r_{e,f}} \times \frac{n}{n_e \times E}$$

Illustrations

Prior

Posterior



two independent experts

low certainty ($c_i^e \approx 0$)

many dependent experts (naive approach)

many dependent experts (tuned weights)

The marginal posterior distributions of $\beta_\theta(t)$ (for each t) are represented using heat maps. Red (resp. white) colour is used to represent high (resp. low) posterior densities.
Prior expectation for each expert; Prior expectation; Posterior expectation; Posterior expectation without prior information

References

[1] Grollemund, P.-M., Abraham, C., Baragatti, M. and Pudlo, P. (2018) Bayesian Functional Linear Regression with Sparse Step functions, *Bayesian Analysis*.
[2] Morris (2015) Functional regression, *Annual Review of Statistics and Its Application*.
[3] O'Hagan (1995) Fractional Bayes factors for model comparison, *JRSSB*.
[4] O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., . . . and Rakow, T. (2006) Uncertain judgements: eliciting experts' probabilities, *John Wiley & Sons*.
[5] Reiss, P., Goldsmith, J., Shang, H.L. and Ogden, T. R. (2015) Methods for scalar-on-function regression, *International Statistical Review*.