Varying smoothing penalties for distributed lag models.

Alastair Rushworth (alastair.rushworth@strath.ac.uk), University of Strathclyde, UK

Why distributed lag models?

Some events (inputs) don't always have immediate consequences on an outcome (outputs). Instead, their effects are spread over the following time intervals.

Simple distributed lag model

Given a set of inputs $(x_1, \ldots, x_n)$ and outputs $(y_1, \ldots, y_n)$, a simple distributed lag model is

$$\mathbb{E}(y_i) = \sum_{j=0}^{p} x_{i-j}\beta_j, \quad i = p+1, p+2, \ldots, n$$

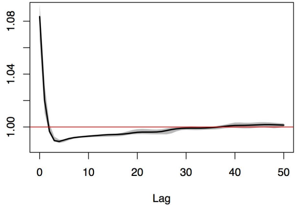
Typically, β_j are restricted to ensure model is realistic

Example: temperature and health

Effects of extreme heat observed over following time periods

Elderly and already ill are particularly vulnerable

Temperature has non-linear influence



Why structured penalties?

Distributed lag curves should reflect reality

- Lag curves should be smooth
- As lag increases, influence decays to 0
- Greater wiggliness at short versus long lags

Combinations of these incorporated into smoothers proposed in literature.

Choosing *p* is problematic

- Maximum lag *p* is usually fixed in advance, however results are sensitive to this choice.
- If lag curve really does decay gradually to zero, 'hard' choice of p doesn't have a clear interpretation

Key challenge: can smoothness and maximum lag p be handled automatically, without user input?

Proposed modelling strategy:

- Choose large p (too many lags)
- Use automatic, variable smoothing to ensure curve is expressed correctly
- Variable smoothing does the work of ensuring curve decays gradually to 0

Distributed lags with adaptive penalties

Basis for lag curve

Use a rich B-spline basis to smooth over lags

$$y_i \sim \mathrm{N}\left(\sum_{j=0}^{p} x_{i-j}\beta_j,\ \sigma^2\right), \quad i = p+1, p+2, \ldots, n$$

$$\beta_j = \sum_{k=1}^{K} B_k(j) b_k$$

Generalised smoothing for basis coefficients

Generalised random walk over basis coefficients allows varying smoothness at different lags.

$$\pi(\mathbf{b}|\lambda) \propto \exp\left(-\frac{1}{2}\sum_{k=1}^{K-1}\lambda_i (b_{k+1} - b_k)^2\right)$$

Further prior to smooth weights

- $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_{K-1})$ are parameters associated with each pair of neighbouring spline coefficients
- Sensible to assume these do not change rapidly: implement a further first-order random walk smoothing prior

$$\boldsymbol{\lambda}|\zeta^2 \sim \mathrm{N}(\mathbf{0}, \zeta^2\mathbf{K}^{-1})$$

Varying smoothing penalties for distributed lag models

Alastair Rushworth (alastair.rushworth@strath.ac.uk), University of Strathclyde, UK

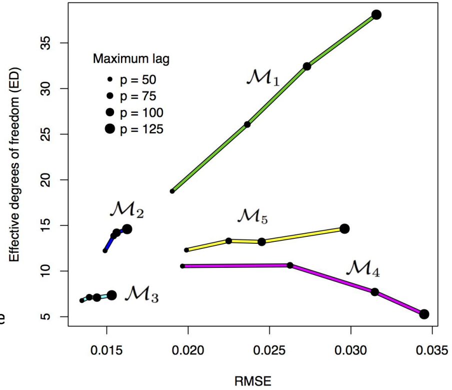
Simulation study - recovering the lag curve with the ‘wrong’ *p*

Compare 5 models with different values of *p*, *the maximum number of lags*

- *p = 50 (truth), 75, 100, 125*
- Different lag curve scenarios
- Measure RMSE (lag curve recovery)
- Measure effective degrees of freedom (models complexity)

Results

- Adaptivity ensures smaller RMSE and effective parameters
- Smoothing adapted automatically across lags - user doesn’t have to choose
- Provides confidence that choosing larger *p* avoids overfitting



Application to temperature and mortality in Greater London

Mortality data (2005 - 2014)

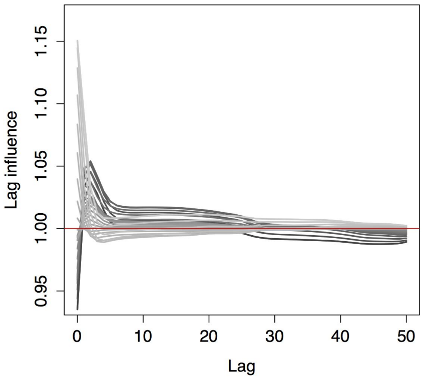
- Total daily deaths in Greater London
- Deaths broken in each of 20 five-year age categories

Covariates

- daily air temperature data
- relative humidity
- air quality (NO2 & PM10)
- national weekly influenza deaths

Key feature: both lag curves and effect of temperature are non-linear functions.

Right: Each curve is a lag function for a different temperature. Lighter greys are warmer temperatures, and darker greys are cooler.



Conclusions

New strategy for fitting lag curves using adaptive smoothing proposed

- Simpler models (by effective degrees of freedom)
- Automatic, lag-dependent smoothing
- Reduces the problem of selecting the maximum lag
- Can be fitted in STAN

Temperature and mortality

- ***Future work: how do lagged effects vary across age categories?***
- ***More important: temperature does not act independently of other meteorological variables. This should appear within the lag function as an ‘experienced temperature’.***

References

Gasparrini, A., Armstrong, B., & Kenward, M. G. (2010). Distributed lag non−linear models. Statistics in medicine, 29(21), 2224-2234.

Muggeo, V. M. (2010). Analyzing temperature effects on mortality within the R environment: the constrained segmented distributed lag parameterization. *Journal of Statistical Software*, *32*(12), 1-17.

Obermeier, V., Scheipl, F., Heumann, C., Wassermann, J., & Küchenhoff, H. (2015). Flexible distributed lags for modelling earthquake data. Journal of the Royal Statistical Society: Series C (Applied Statistics), 64(2), 395-412.

Rushworth, A. M., Bowman, A. W., Brewer, M. J., & Langan, S. J. (2013). Distributed lag models for hydrological data. *Biometrics*, *69*(2), 537-544.

Welty, L. J., Peng, R. D., Zeger, S. L., & Dominici, F. (2009). Bayesian distributed lag models: estimating effects of particulate matter air pollution on daily mortality. Biometrics, 65(1), 282-291.