Spline-based approach to infer farm vehicle trajectories

Jerome Cao

University of Otago

November 18, 2015

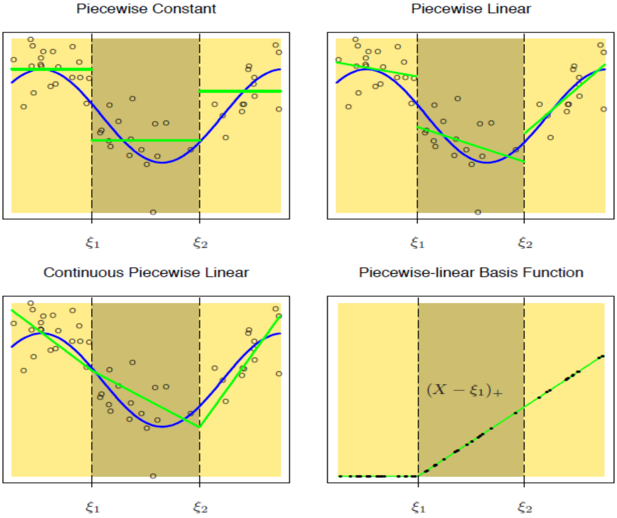
Overview

Piecewise Polynomials and Splines

Linear regression, linear discriminant analysis, logistic regression and separating hyperplanes all rely on a linear model. However, it is extremely unlikely that the true function $f(X)$ is actually linear in X.
Denote by $h_m(X) : \mathbb{R}^p \to \mathbb{R}$ the mth transformation of X, $m = 1, ..., M$. We then model

$$f(X) = \sum_{m=1}^{M} \beta_m h_m(X).$$

a linear basis expansion in X.



$[T.Hastie, R.Tibshirani, J.Friedman, andJ.Franklin(2001)]$

Smoothing Spline Problems

Consider the following problem: among all functions $f(x)$ with two continuous derivatives, find one that minimizes the penalized residual sum of squares

$$RSS(f,\lambda) = \sum_{i=1}^{n}(f(x_i) - y_i)^2 + \lambda \int_0^1 (f''(x))^2 dx \qquad (1)$$

where λ is a fixed smoothing parameter, (x_i, y_i), $i = 1, \cdots, n$ are observed data and $0 < x_1 < \cdots < x_n < 1$.

$\lambda = 0$: f can be any function that interpolates the data.

$\lambda = \infty$: the simple least squares line fit.

The above function could be solved by a natural cubic spline

$$f(x) = \sum_{i=1}^{n} N_i(x)\theta_i \tag{2}$$

where the $N_i(x)$ are an N-dimensional set of basis functions for representing this family of natural splines.

Then equation (1) becomes

$$RSS(\theta, \lambda) = (\mathbf{Y} - \mathbf{N}\theta)^T(\mathbf{Y} - \mathbf{N}\theta) + \lambda\theta^T\mathbf{\Omega}\theta$$

where $\{\mathbf{N}\}_{ij} = N_j(x_i)$ and $\{\mathbf{\Omega}\}_{jk} = \int N_j''(t)N_k''(t)dt$.
Then the solution is

$$\hat{\theta} = (\mathbf{N}^T\mathbf{N} + \lambda\mathbf{\Omega})^{-1}\mathbf{N}^T\mathbf{Y}.$$

So the fitted smoothing spline is given by

$$\hat{f}(x) = \sum_{j=1}^{n} N_j(x)\hat{\theta}_j.$$

Using Cross Validation to Find Parameter λ

The next step is finding best parameter λ. Considering data (t_i, x_i), $i = 1, \cdots, n$. The leave-one-out cross validation error is

$$CV = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{f}^{(-i)}(t_i))^2$$

where $\hat{f}^{(-i)}(t)$ is the smoothing spline without the datapoint (ti, xi). So the best parameter λ is the one that could minimize CV-errors.

A more complicated situation: 2D, velocity, piecewise λ

Tractors working on the orchard use booms to spread pesticide. When boom is down, indicating that tractors goes straight and slower, and up means the boom is not working and tractors goes faster or having a turn.
In this case, we divide smoothing parameter λ, which controls the smoothness of a spline, into two parameters λ_d and λ_u.
So the minimized CV returns three parameters λ_d, λ_u and η.

"Tractor" Spline and The Objective Function

If we have some knots, such that $a < t_1 < \cdots < t_n < b$, and $z_i = (x_i, y_i)$, $w_i = (u_i, v_i)$, for $i = 1, 2, \cdots, n$, and a positive piecewise parameter $\lambda(t)$, which will control penalty functions, then the equation

$$J[f] = \sum_{i=1}^{n} |f(t_i) - z_i|^2 + \eta \sum_{i=1}^{n} |f'(t_i) - w_i|^2 + \sum_{i=1}^{n} \lambda_i \int_{\tau_i}^{\tau_{i+1}} |f''(t)|^2 dt$$

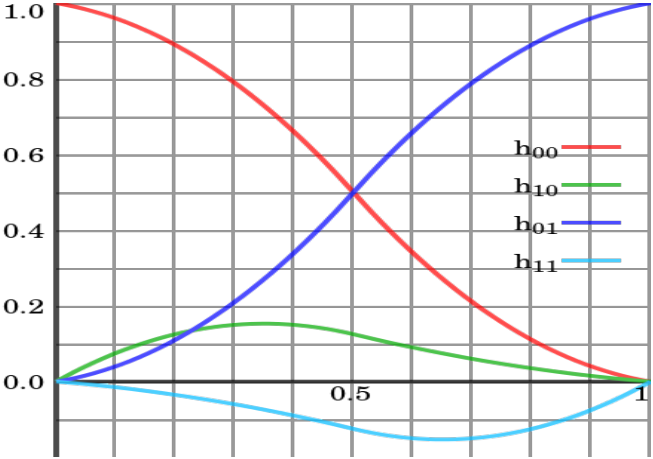
is minimised by a tractor spline, which is linear outside the knots.

Hermite Spline

On the unit interval $(0, 1)$, given a starting point $\boldsymbol{p}_0$ at $t = 0$ and an ending point $\boldsymbol{p}_1$ at $t = 1$ with starting tangent $\boldsymbol{m}_0$ at $t = 0$ and ending tangent $\boldsymbol{m}_1$ at $t = 1$, the polynomial can be defined by

$$\boldsymbol{p}(t) = (2t^3 - 3t^2 + 1)\boldsymbol{p}_0 + (t^3 - 2t^2 + t)\boldsymbol{m}_0 + (-2t^3 + 3t^2)\boldsymbol{p}_1 + (t^3 - t^2)\boldsymbol{m}_1$$

where $t \in [0, 1]$. The interpolant in each subinterval is a linear combination of these four functions.



[Cubic Hermite spline from Wikipedia]

"Tractor" Spline Basis

On an arbitrary interval $[t_i, t_{i+1}]$, we have

$$h_{00}^{(i)}(t) = \begin{cases} 2(\frac{t-t_i}{t_{i+1}-t_i})^3 - 3(\frac{t-t_i}{t_{i+1}-t_i})^2 + 1, & t_i \leq t \leq t_{i+1} \\ 0 & otherwise \end{cases}$$

$$h_{10}^{(i)}(t) = \begin{cases} \frac{(t-t_i)^3}{(t_{i+1}-t_i)^2} - 2\frac{(t-t_i)^2}{t_{i+1}-t_i} + (t - t_i), & t_i \leq t \leq t_{i+1} \\ 0 & otherwise \end{cases}$$

$$h_{01}^{(i)}(t) = \begin{cases} -2(\frac{t-t_i}{t_{i+1}-t_i})^3 + 3(\frac{t-t_i}{t_{i+1}-t_i})^2, & t_i \leq t \leq t_{i+1} \\ 0 & otherwise \end{cases}$$

$$h_{11}^{(i)}(t) = \begin{cases} \frac{(t-t_i)^3}{(t_{i+1}-t_i)^2} - \frac{(t-t_i)^2}{t_{i+1}-t_i}, & t_i \leq t \leq t_{i+1} \\ 0 & otherwise \end{cases}$$

We define functions $N_1 = h_{00}^{(1)}$, $N_2 = h_{10}^{(1)}$, $N_{2n-1} = h_{01}^{(n)}$, $N_{2n} = h_{11}^{(n)}$. For all $k = 1, 2, \ldots, n-2$ we define N_{2k+1} by

$$N_{2k+1}(t) = \begin{cases} h_{01}^{(k)} + h_{00}^{(k+1)} & t \neq t_{k+1} \\ 1 & t = t_{k+1}. \end{cases}$$

and $N_{2k+2} = h_{11}^{(k)} + h_{10}^{(k+1)}$.
We then have

$$N_i(t_j) = \begin{cases} 1 & \text{if } i = 2j-1 \\ 0 & \text{otherwise.} \end{cases}$$

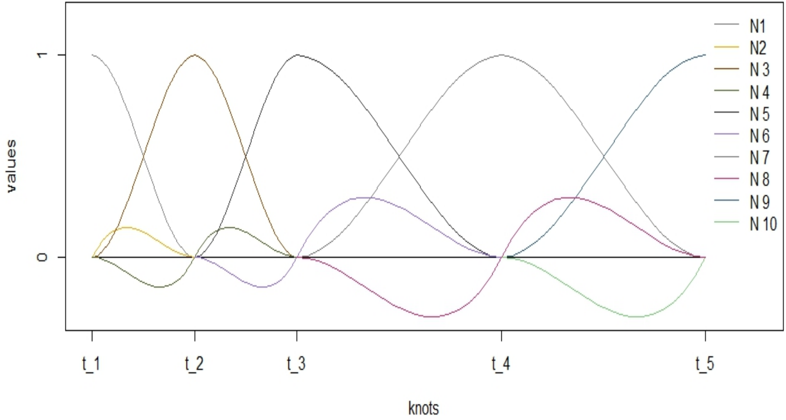
and

$$N_i'(t_j) = \begin{cases} 1 & \text{if } i = 2j \\ 0 & \text{otherwise.} \end{cases}$$

Theorem

The functions $N_1, \ldots, N_{2n}$ *provide a basis for the set of functions on* $[t_1, t_n]$ *which are continuous, have continuous first derivatives and which are cubic on each open interval* (t_i, t_{i+1}).

Assume that vector $t = (0, 1, 2, 4, 6)$, then we could draw a graph of its basis functions. These basis functions construct the tractor spline on interval $(0, 6)$.



The tractor spline $f(t)$ is a linear combination of basis functions

$$f(t) = \sum_{j=1}^{2n} \theta_j N_j(t)$$

and the $RSS(f, \lambda, \eta)$ could reduces to

$$RSS(\theta, \lambda, \eta) = (\mathbf{z} - \mathbf{B}\theta)^T(\mathbf{z} - \mathbf{B}\theta) + \eta(\mathbf{w} - \mathbf{C}\theta)^T(\mathbf{w} - \mathbf{C}\theta) + \theta^T \mathbf{\Omega}_\lambda \theta,$$

where $\mathbf{z} = \{z(x_i, y_i)\}$ are the knots and $\mathbf{w} = \{(u_i, v_i)\}$ are the tangent at knots.

After solving the above equation, we have

$$\hat{\theta} = (\mathbf{B}^T\mathbf{B} + \eta\mathbf{C}^T\mathbf{C} + \mathbf{\Omega}_\lambda)^{-1}(\mathbf{B}^T\mathbf{z} + \mathbf{C}^T\mathbf{w}). \tag{3}$$

Then tractor spline could be represented as

$$\hat{f}(t) = \sum_{j=1}^{2n} \hat{\theta}_j N_j(t) \tag{4}$$

Cross Validation

Because $\hat{f}$ and $\hat{f}'$ could be written in the form of

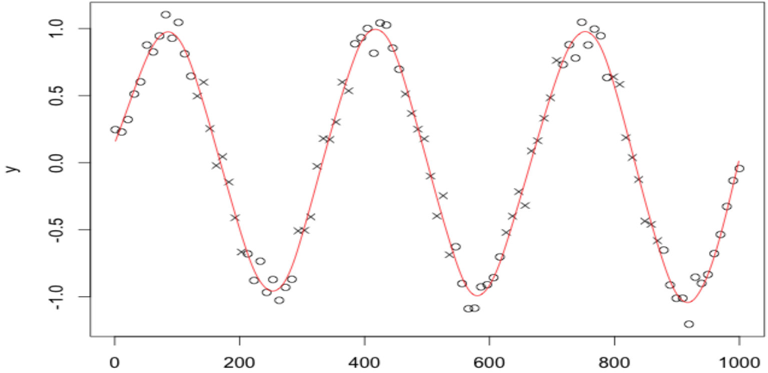
$$\hat{f} = B\hat{\theta} = Sz + \eta Tw$$
$$\hat{f}' = C\hat{\theta} = Uz + \eta Vw$$

Then

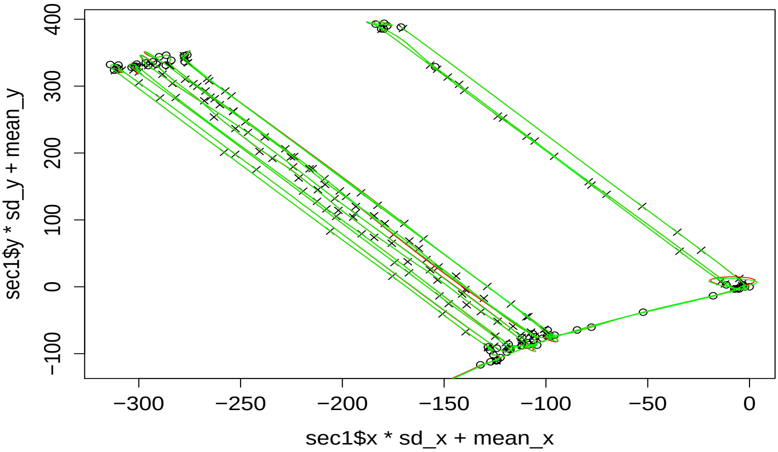
$$CV = \frac{1}{N}\sum(\hat{f}^{(-i)}(t_i) - z_i)^2$$
$$= \frac{1}{N}\sum\left(\frac{\hat{f}(t_i) - z_i + \eta\frac{T_{ii}}{1-\eta V_{ii}}(\hat{f}'(t_i) - w_i))}{1 - S_{ii} - \eta\frac{T_{ii}}{1-\eta V_{ii}}U_{ii}}\right)^2$$

1D simulation

Considering $f(t) = sin(\frac{2\pi t}{300}) + \varepsilon$, where $0 \leq t \leq 1000$ and $\varepsilon \sim N(0, 0.2)$. Using this method, we could get $\lambda_d = 1.52 \times 10^5, \lambda_u = 0.65 \times 10^5$, $\eta = 28.83$ and CV score is 1.97×10^{-2}.



Real Data


sec1$y * sd_y + mean_y
400
300
200
100
0
−100
−300
−250
−200
−150
−100
−50
0
sec1$x * sd_x + mean_x

Future Work

- Working on Gaussian Process Regression in parallel to smoothing splines
- Moving from batch inference to online inference
- Deducing high-level description of farm vehicle motion

Acknowledgement

- Ministry of Business, Innovation & Employment
  - Grant UOOX 1208
- TracMap
- Dr Matthew Parry
- Professor David Bryant

Ministry of Business, Innovation & Employment

TRACMAP AGRICULTURE

References

T. Hastie, R. Tibshirani, J. Friedman, and J. Franklin (2001)
The Elements of Statistical Learning: data mining, inference and prediction.

P. J. Green. (1994)
Nonparametric regression and generalized linear models : a roughness penalty approach

Y.J. Kim and C. Gu (2004)
Smoothing spline gaussian regression: more scalable computation via efficient approximation
Journal of the Royal Statistical Society: Series B (Statistical Methodology) vol. 66, no. 2, pp. 337-356

The End