Inference and Characterization of Planar Trajectories

Zhanglong Cao

Department of Mathematics & Statistics
University of Otago

4th July 2018

Main Contributions in The Past Three Years

- Non-parametric V-spline for trajectory reconstruction
- Correspondence between V-spline and Bayes estimate
- A practical adaptive MCMC sampler for combined state and parameter estimation

GPS points and Trajectory

GPS units irregularly record time series data of a moving object. These data are in the form of

$$P = \{x_t, y_t, v_t, \omega_t, b_t, \cdots | t \in \mathbb{R}\}. \tag{1}$$

x	longitude
y	latitude
v	velocity
ω	bearing
b	boom status

A **trajectory** is a curve relating (not necessarily connecting) a sequence of successive positions.

A New Objective Function

Given n observation times such that $a < t_1 < \cdots < t_n < b$, with associated position data y_i, velocity data v_i, for $i = 1, 2, \cdots, n$, and a positive piecewise parameter $\lambda(t)$, which will control penalty functions, then the equation

$$J[f] = \frac{1}{n}\sum_{i=1}^{n} |f(t_i) - y_i|^2 + \frac{\gamma}{n}\sum_{i=1}^{n} |f'(t_i) - v_i|^2 + \sum_{i=1}^{n}\int_{t_i}^{t_{i+1}} \lambda(t)|f''(t)|^2 dt \quad (2)$$

is minimized by a V-spline, which is linear outside the knots.

Solution to The New Objective Function

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

$$f(t) = \sum_{j=1}^{2n} N_j(t)\theta_j \tag{3}$$

and the objective function becomes

$$nJ[f] = (\mathbf{y} - B\theta)^T(\mathbf{y} - B\theta) + \gamma(\mathbf{v} - C\theta)^T(\mathbf{v} - C\theta) + n\theta^T\Omega_\lambda\theta. \tag{4}$$

Therefore, the solution is

$$\hat{\theta} = \left(B^\top B + \gamma C^\top C + n\Omega_\lambda\right)^{-1}\left(B^\top \mathbf{y} + \gamma C^\top \mathbf{v}\right). \tag{5}$$

Cross Validation of V-Splines

The formula of cross validation score is

$$\mathrm{CV} = \frac{1}{n}\sum_i (\hat{f}^{(-i)}(t_i) - y_i)^2 \tag{6}$$

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

$$\begin{aligned} \hat{f} &= B\hat{\theta} = S\mathbf{y} + \gamma T\mathbf{v} \\ \hat{f}' &= C\hat{\theta} = U\mathbf{y} + \gamma V\mathbf{v} \end{aligned} \tag{7}$$

Then for V-spline,

$$\mathrm{CV} = \frac{1}{n}\sum_i \left(\frac{\hat{f}(t_i) - y_i + \gamma \frac{T_{ii}}{1-\gamma V_{ii}}(\hat{f}'(t_i) - v_i))}{1 - S_{ii} - \gamma \frac{T_{ii}}{1-\gamma V_{ii}} U_{ii}} \right)^2 \tag{8}$$

Reproducing Kernel Hilbert Space for V-Splines

In the space

$$\mathcal{C}^2_{p.w.}[0,1] = \{f : f, f' \text{ continuous, } f'' \text{ piecewise continuous on } [0,1]\}.$$

given an appropriate inner product

$$(f,g) = f(0)g(0) + f'(0)g'(0) + \int_0^1 f''g''dt, \tag{9}$$

the space $\mathcal{C}^2_{p.w.}[0,1]$ is made a reproducing kernel Hilbert space. The reproducing kernel is $R(s,t) = R_0(s,t) + R_1(s,t)$, where

$$R_0(s,t) = 1 + st \tag{10}$$

$$R_1(s,t) = \int_0^1 (s-u)_+(t-u)_+du \tag{11}$$

are both non-negative definite themselves.

Bayes Estimate

In a generic process, we know that $p(\mathbf{y}, \mathbf{v} \mid f) = N(f, \sigma^2)$. However, we are more interested in f given measurements, such as

$$p(f \mid \mathbf{y}, \mathbf{v}) \propto p(\mathbf{y}, \mathbf{v} \mid f) p(f) \tag{12}$$

where $f \sim GP(0, \Sigma)$ is a Gaussian process prior. Particularly, in $\mathcal{C}^2_{p.w.}[0,1]$

$$\Sigma \propto \rho \begin{bmatrix} R_0(s,t) & \dot{R}_0(s,t) \\ R_0'(s,t) & \dot{R}_0'(s,t) \end{bmatrix} + \begin{bmatrix} R_1(s,t) & \dot{R}_1(s,t) \\ R_1'(s,t) & \dot{R}_1'(s,t) \end{bmatrix} + \begin{bmatrix} n\lambda I & 0 \\ 0 & \frac{n\lambda}{\gamma} I \end{bmatrix} \tag{13}$$

Posterior Mean

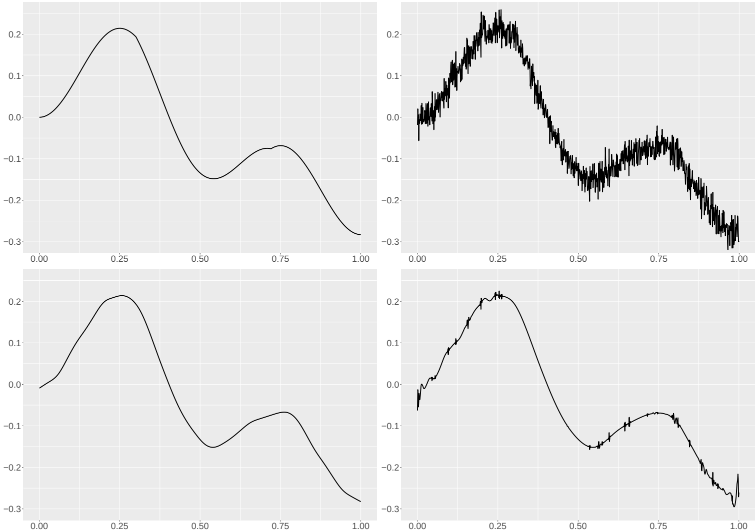
The posterior mean $E(f \mid \mathbf{y}, \mathbf{v})$ is $\hat{f} = \phi^\top \mathbf{d} + \xi^\top \mathbf{c} + \psi^\top \mathbf{b}$, with the coefficients given by

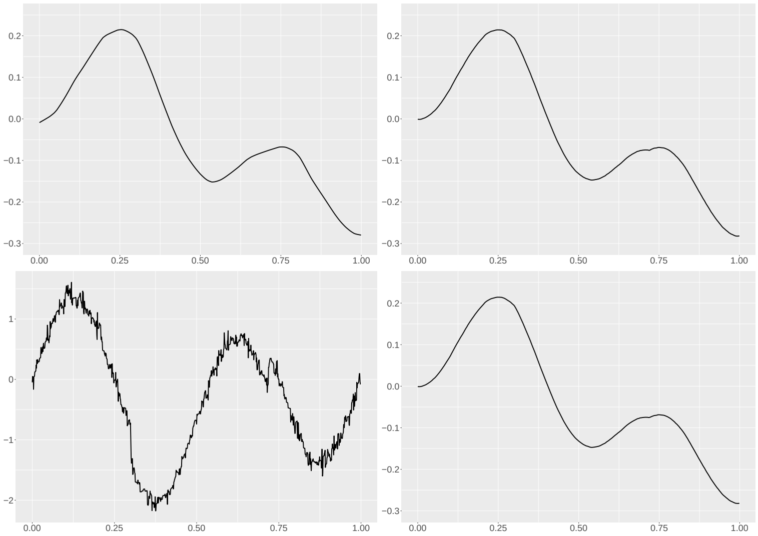
$$\mathbf{d} = \left(T^\top M^{-1} T\right)^{-1} T^\top M^{-1} \begin{bmatrix} \mathbf{y} \\ \mathbf{v} \end{bmatrix}, \tag{14}$$

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{b} \end{bmatrix} = \left(M^{-1} - M^{-1} T \left(T^\top M^{-1} T\right)^{-1} T^\top M^{-1}\right) \begin{bmatrix} \mathbf{y} \\ \mathbf{v} \end{bmatrix}, \tag{15}$$

where $M = \begin{bmatrix} Q + n\lambda I & P \\ Q' & P' + \frac{n\lambda}{\gamma} I \end{bmatrix}$ (when $\rho \to \infty$).
The posterior mean is the solution to the objective function and is correspondence to (7) if $\lambda(t)$ is constant and the sequence is converted to $[0, 1]$.

Simulation and Application




0.2
0.1
0.0
−0.1
−0.2
−0.3
0.00
0.25
0.50
0.75
1.00
1
0
−1
−2

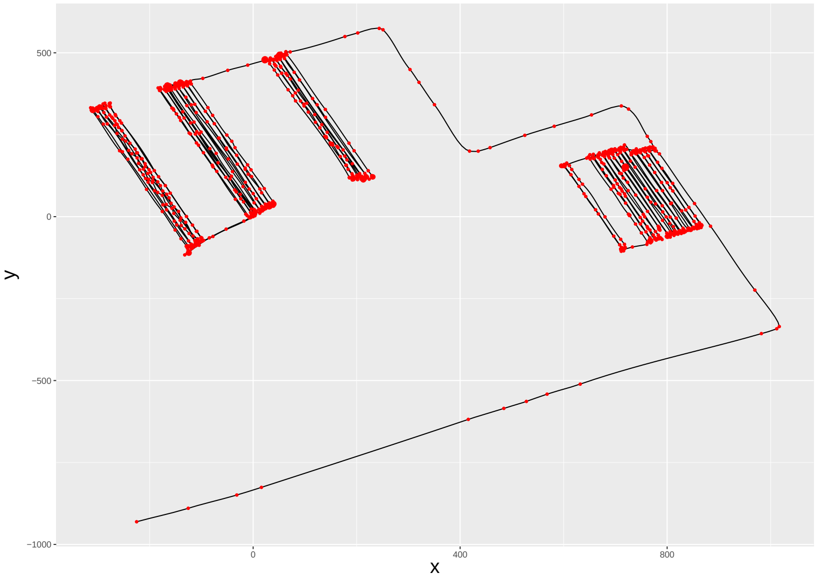


Figure: 2-dimensional reconstruction. Larger dots indicate bigger values of penalty function $\lambda(t)$.

On-line Estimation and Adaptive Sequential MCMC

The forward map for the states is based on an Ornstein-Uhlenbeck process,

$$\begin{cases} \mathrm{d}u_t = -\gamma u_t \mathrm{d}t + \lambda \mathrm{d}W'_t, \\ \mathrm{d}x_t = u_t \mathrm{d}t + \xi \mathrm{d}W_t, \end{cases} \tag{16}$$

so that γ^{-1} is roughly the time scale over which the velocity remains informative in the absence of subsequent observations. In our application, $\gamma^{-1} \approx 60$s.
The set of parameters to be estimated is $\theta = \{\gamma, \xi^2, \lambda^2, \sigma^2, \tau^2\}$. The filtering for states is

$$p(X_t \mid Y_{1:t}) = \int p(X_t \mid Y_{1:t}, \theta) p(\theta \mid Y_{1:t}) \mathrm{d}\theta \tag{17}$$

and $p(\theta \mid Y_{1:t}) \propto p(Y_{1:t} \mid \theta) p(\theta)$.

Learning Phase

In the learning phase, a cheap Gaussian surrogate $\hat{p}(\theta)$ is obtained that will be used for a *delayed-acceptance Metropolis-Hastings* sampler in the estimation phase.
Specifically, a self-tuning random walk Metropolis-Hastings algorithm, in which the parameters are updated one at a time, is used to obtain the mean and covariance structure of $\hat{p}(\theta)$

Figure: Visualization of the parameters correlation matrix, which is found in learning phase.

Estimation Phase

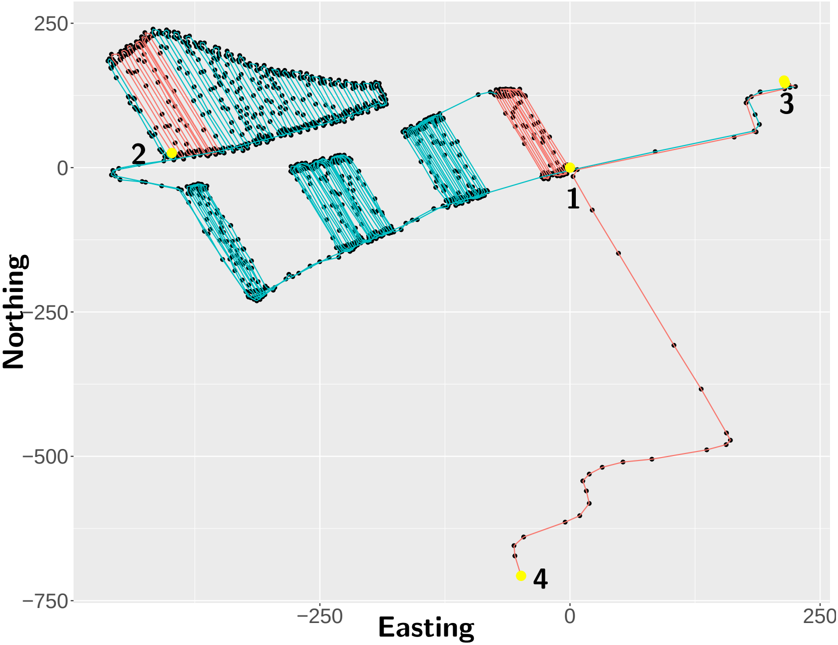
We use a Monte Carlo estimate of the target in (17):

$$p(X_t \mid Y_{1:t}) \doteq \frac{1}{N} \sum_{i=1}^{N} p(X_t \mid Y_{1:t}, \theta^{(i)}). \tag{18}$$

To maintain sampling speed, we introduce a *Sliding Window MCMC* algorithm that retains only the last L observations. This is also advantageous in our application as the parameters are often slowly varying in time. To maintain sampling efficiency, if the acceptance ratio α drops below a predefined threshold, the mean of $\hat{p}(\theta)$ is updated (the covariance is kept the same).

Application

A tractor moving on an orchard is mounted with a GPS-enabled unit, which records and transmits data to a remote server. The data is an irregularly spaced time series of longitude, latitude, speed and bearing. In online mode, the sliding window algorithm is able to infer the tractor's position within seconds with an uncertainty of ≈ 0.5m. A window L of 100 is chosen to maintain a balance between sampling speed and acceptable estimation errors.



250
0
−250
−500
−750
−250
0
250
Northing
Easting
1
2
3
4

Further videos at:
V-spline reconstruction
Adaptive MCMC

Amendments

- Chapter one: separate description for GPS data.
- Chapter two: larger labels in figures.
- Chapter three: the conjecture has been resolved.
- Chapter four: fixed typographical errors.
- Chapter five: explicit explanation, better figures.
- Others: references are properly cited; typographical errors are fixed.

Chapter one: separate description for GPS data

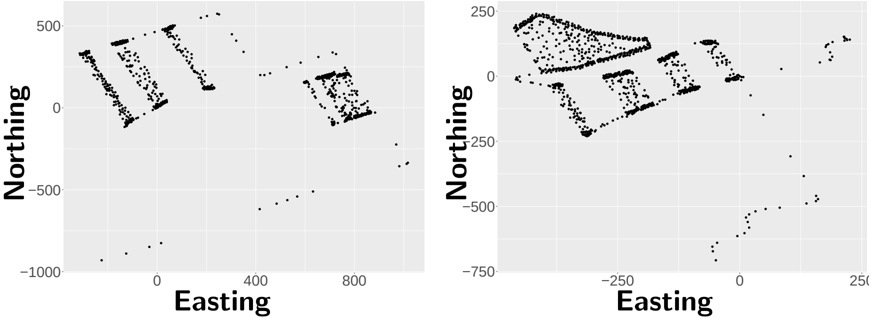


Figure: Examples of GPS data. Observed positions y_t are shown. In trajectory reconstruction, the y_t are combined with velocity information v_t and operating characteristics b_t to infer actual positions x_s, for times of interest s.

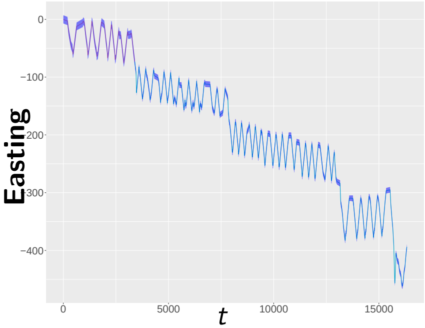
Chapter three: a new inner product

$$\langle f, g \rangle = f(0)g(0) + f'(0)g'(0) + \sum_{i=1}^{n-1} w_i^{-1} \int_{t_i}^{t_{i+1}} f''(t)g''(t)dt, \qquad (19)$$

on $\mathcal{C}^{(2)}_{\text{p.w.}}[0,1]$, where $\sum_i w_i = 1$ and $w_i > 0$. In fact, $w_i = \frac{\lambda_i}{\sum_i \lambda_i}$.

Chapter five: explicit explanation and better figures

- Remove repeated equations.
- Introduction to a new concept: delayed-acceptance.
- A homogeneous linear state-space model explains the joint distribution of $X, Y \sim N(0, \Sigma)$.
- New figures for uncertainty



Acknowledgment

- Ministry of Business, Innovation & Employment
  – Grant UOOX 1208
- TracMap
- Dr Matthew Parry
- Professor David Bryant
- All examiners and convenor
- Greg and Chris

Ministry of Business, Innovation & Employment

TRACMAP AGRICULTURE