Previous Research

Zhanglong Cao

Department of Maths & Stats
University of Otago

June 5, 2018

Overview

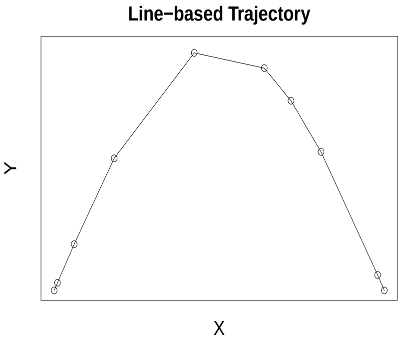
GPS points and Trajectory

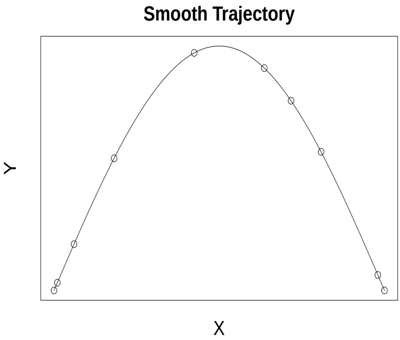
GPS units irregularly record time series data of a moving object. These data are in the form of

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

x	longitude
y	latitude
v	velocity
ω	bearing
b	boom status

Trajectory is a connection by a time series successive position recorded by GPS devices.


Line−based Trajectory
Y
X


Smooth Trajectory
Y
X

A New 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

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

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

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

and the objective function reduces to

$$\text{MSE}(\theta, \lambda, \gamma) = (\mathbf{z} - \mathbf{B}\theta)^T(\mathbf{z} - \mathbf{B}\theta) + \gamma(\mathbf{w} - \mathbf{C}\theta)^T(\mathbf{w} - \mathbf{C}\theta) + n\theta^T\mathbf{\Omega}_\lambda\theta,$$

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

Cross Validation of V-Spline

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

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

Then

$$\begin{aligned}\mathrm{CV} &= \frac{1}{n}\sum(\hat{f}^{(-i)}(t_i) - z_i)^2 \\ &= \frac{1}{n}\sum\left(\frac{\hat{f}(t_i) - z_i + \gamma\frac{T_{ii}}{1-\gamma V_{ii}}(\hat{f}'(t_i) - w_i))}{1 - S_{ii} - \gamma\frac{T_{ii}}{1-\gamma V_{ii}}U_{ii}}\right)^2\end{aligned}$$

Adjusted Penalty Term

Given a constant $\lambda = \frac{(\Delta T_1)^3}{(\Delta d_1)^2}\eta$, the penalty term becomes

$$\eta \frac{(\Delta T_1)^2}{(\Delta d_1)^2}\left((2\varepsilon_1 + \varepsilon_2)^2 + 3\varepsilon_2^2\right) = \eta \frac{(2\varepsilon_1 + \varepsilon_2)^2 + 3\varepsilon_2^2}{\bar{v}^2} \\ \sim \left(\frac{\text{discrepancy in velocity}}{\text{average velocity}}\right)^2 \tag{3}$$

which will be enormous with large measured errors in velocity v_1 or v_2 comparing to average velocity $\bar{v}$.

Hilbert Space and Reproducing Kernel for V-Spline

For any $f \in \mathbb{H}$, $f = \sum_{i=1}^{2n} \theta_i N_i(t)$. Building up a new space

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

Equipped with an appropriate inner product

$$(f, g) = f(0)g(0) + f'(0)g'(0) + \int_0^1 f'' g'' dt, \tag{4}$$

the space $\mathcal{C}^2_{p.w.}[0,1]$ is made a reproducing kernel Hilbert space.

$f \in \mathcal{C}^2_{p.w.}[0,1]$ can be written as

$$f(t) = d_1 + d_2 t + \sum_{j=1}^{n} c_j R_1(t_j, t) + \sum_{i=j}^{n} b_j \dot{R}_1(t_j, \cdot), \tag{5}$$

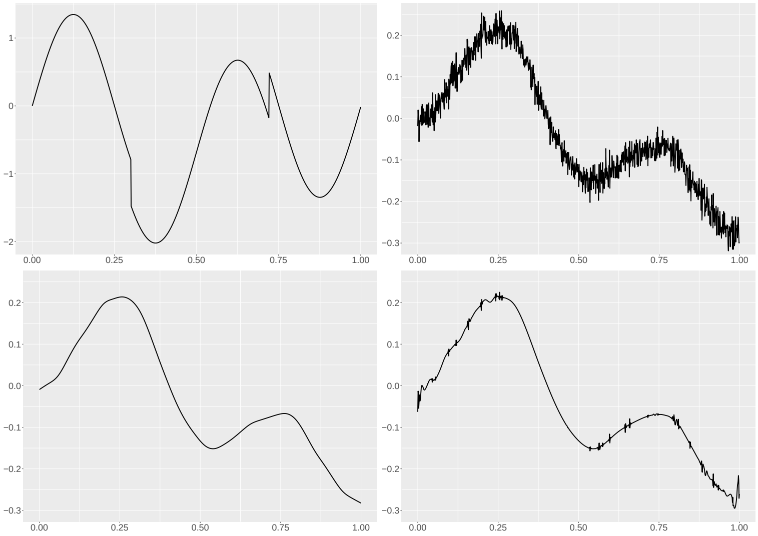
where $\mathbf{d}, \mathbf{c}$ and $\mathbf{b}$ are coefficients. $f = \phi^\top \mathbf{d} + \xi^\top \mathbf{c} + \psi^\top \mathbf{b}$, with the coefficients given by

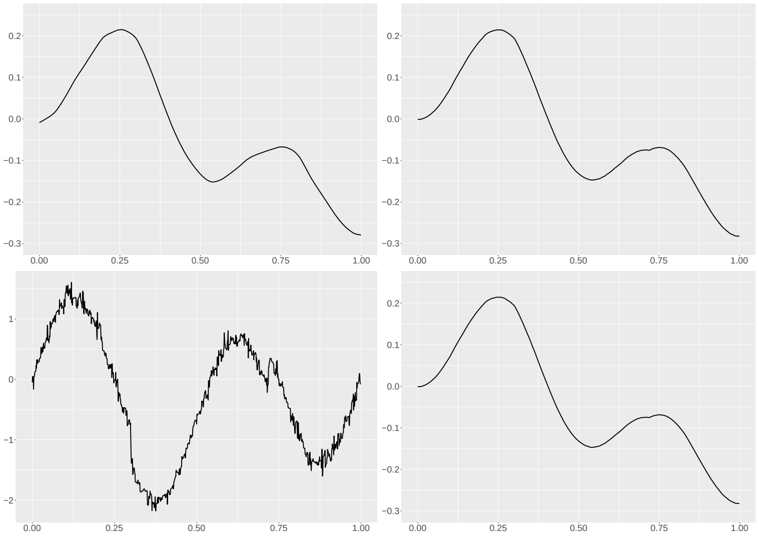
$$\mathbf{d} = (T^\top M^{-1} T)^{-1} T^\top M^{-1} \begin{bmatrix} Y \\ V \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{b} \end{bmatrix} = (M^{-1} - M^{-1} T (T^\top M^{-1} T)^{-1} T^\top M^{-1}) \begin{bmatrix} Y \\ V \end{bmatrix},$$

where $T = \begin{bmatrix} S \\ S' \end{bmatrix}$ and $M = \begin{bmatrix} Q + n\lambda I & P \\ Q' & P' + \frac{n\lambda}{\gamma} I \end{bmatrix}$.

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

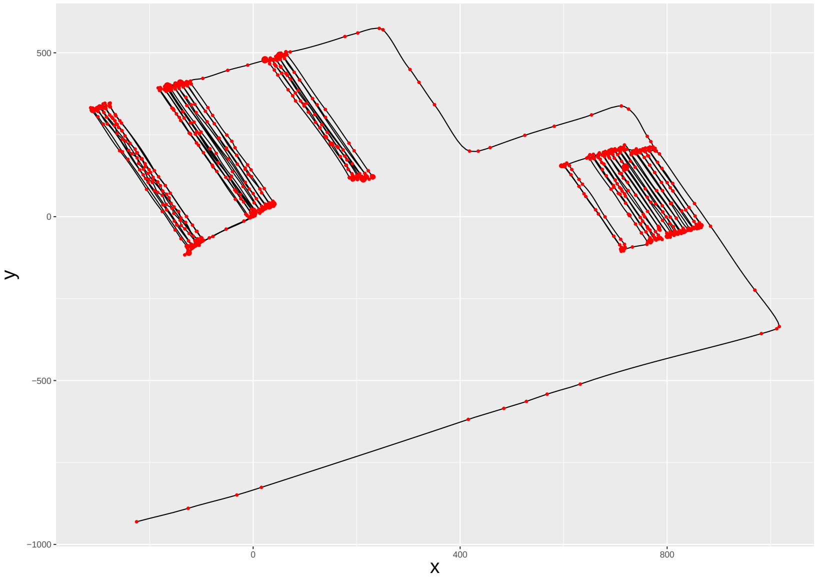


Figure: 2-dimensional reconstruction. Larger dots indicate bigger values of penalty function $\lambda(t)$.

On-line Estimation and Adaptive Sequential MCMC

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

$$\begin{cases} \mathrm{d}u_t = -\gamma u_t \mathrm{d}t + \lambda \mathrm{d}W_t', \\ \mathrm{d}x_t = u_t \mathrm{d}t + \xi \mathrm{d}W_t, \end{cases} \tag{6}$$

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

$$p(X_t \mid Y_{1:t}) = \int p(X_t \mid Y_{1:t}, \theta) p(\theta \mid Y_{1:t}) \mathrm{d}\theta \tag{7}$$

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 $\hat{p}(\theta)$

Estimation Phase

We use a Monte Carlo estimate of the target in (7):

$$p(X_t \mid Y_{1:t}) \doteq \frac{1}{N} \sum_{i=1}^{N} p(X_t \mid Y_{1:t}, \theta^{(i)}). \tag{8}$$

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.

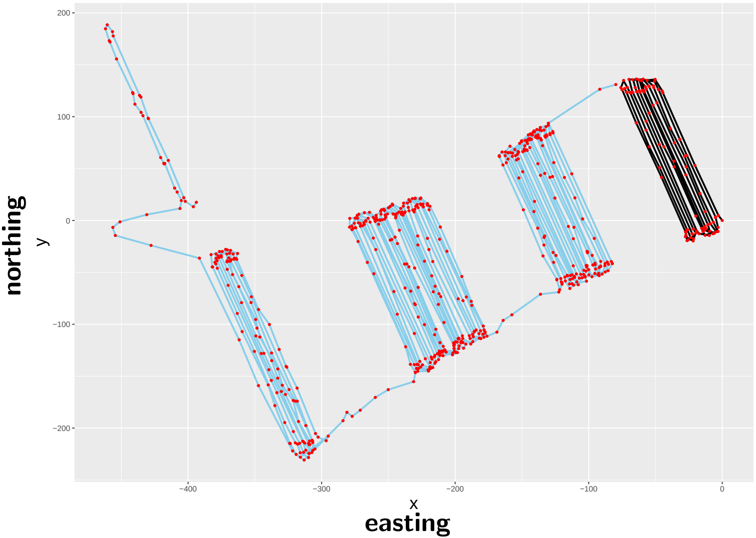


Figure: Posterior mean trajectory in two dimensions. The orange line is the reconstruction from the learning phase and the green line is the filtering from the estimation phase. Red dots are the measurements, and the two blue dots are the start and end points.

Acknowledgment

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

Ministry of Business, Innovation & Employment

TRACMAP AGRICULTURE

The End