Going online with GPS data

Zhanglong Cao

23 November 2016

Overview

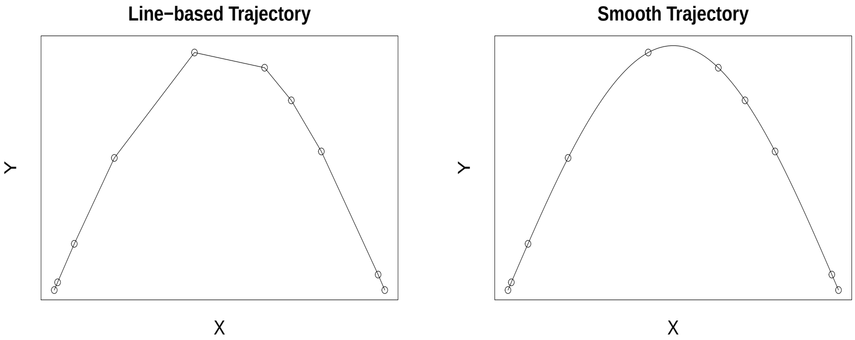
1 GPS Data and Trajectory Reconstruction

GPS units record time series data of moving objects. These data are in the form of

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

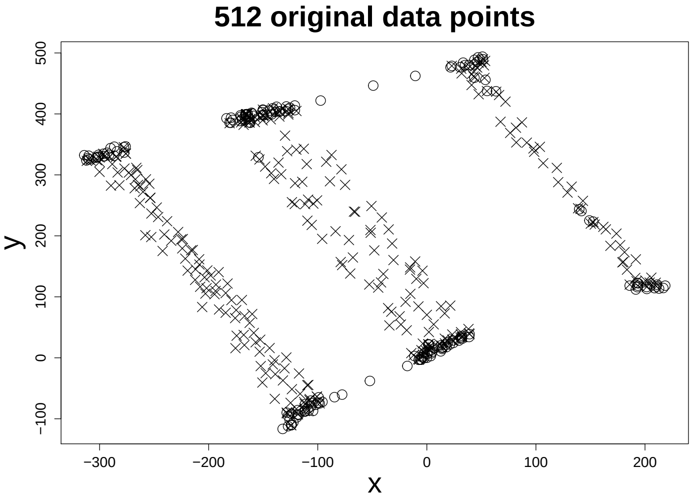
x	longitude
y	latitude
v	velocity
ω	barrier
b	boom status

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



Line−based Trajectory
Y
X
Smooth Trajectory
Y
X

Off-line: Playing with a batch of data



2.1 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

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

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

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 and second derivatives and which are cubic on each open interval* (t_i, t_{i+1}).

Solution to The New Objective Function

The tractor 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.

After solving the above equation, we have

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

Then tractor spline could be represented as

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

2.2 A New Cross Validation Method

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

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

2.3 A Spin-off Discovery

Build up a new space

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

Equipped with an appropriate inner product

$$(f,g) = f(0)g(0) + f'(0)g'(0) + \int_0^1 f''g''dx, \tag{5}$$

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

$$f(x) = d_1 + d_2 x + \sum_{j=1}^{n} c_j R_1(x_j, x) + \sum_{i=j}^{n} b_j \dot{R}_1(x_j, \cdot), \tag{6}$$

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

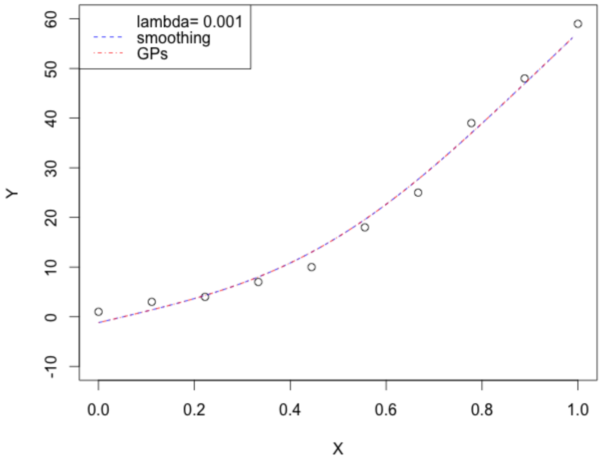
$$\mathbf{d} = (T^\top M^{-1} T)^{-1} T^\top M^{-1} \begin{bmatrix} Y \\ V \end{bmatrix},$$

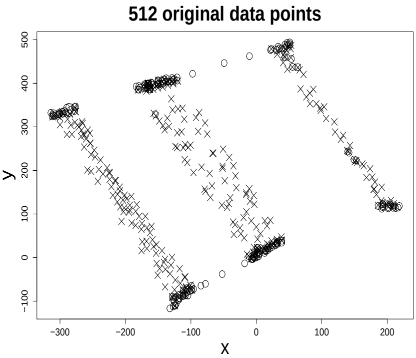
$$\begin{bmatrix} \mathbf{c} \\ \mathbf{b} \end{bmatrix} = (M^{-1} - M^{-1} T (T^\top M^{-1} T)^{-1} T^\top M^{-1}) \begin{bmatrix} Y \\ V \end{bmatrix},$$

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

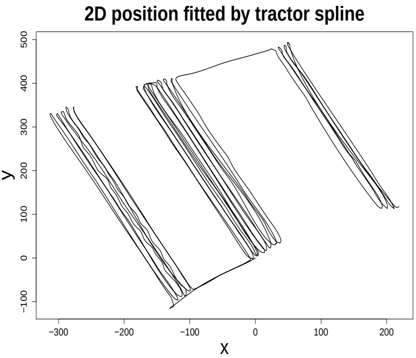
2.4 Numeric Simulation

When λ is chosen 0.001, smoothing spline and GPs return the same results.



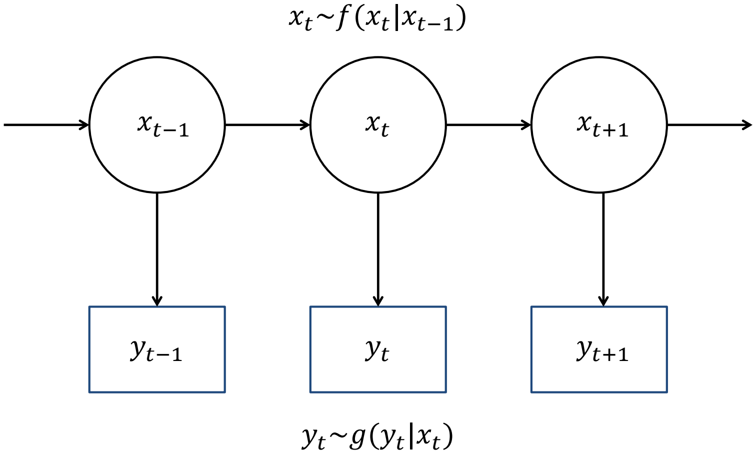


512 original data points
y
x
−100
0
100
200
300
400
500
−300
−200
−100
0
100
200



2D position fitted by tractor spline
y
x
−100
0
100
200
300
400
500
−300
−200
−100
0
100
200

On-line: Updating data instantly



3.1 Dynamic Linear Models

A state space model is

$$
\begin{aligned}
y_t &= F_t x_t + \epsilon_t, \\
x_t &= \phi_t x_{t-1} + w_t, \\
x_0 &\sim N(m_0, c_0),
\end{aligned}
$$

where $\epsilon_t \sim N(0, \sigma)$ and $w_t \sim N(0, \eta)$. x_t are hidden status and y_t are observations.

Suppose the parameter ϕ and observations $y_{1:t}$ are known, we are interested in inferring the true status x_k by maxing the probability function:

$$P(x_k|y_{1:t},\phi) = \int P(x_k|y_{1:t},\phi)P(\phi|y_{1:t})d\phi. \tag{7}$$

- If $k \leq t$, Smoothing;
- If $k = t$, Filtering;
- If $k > t$, Prediction.

The joint distribution of $x_{0:t}$ and $y_{1:t}$ is

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim N\left(0, \Sigma^{-1}\right),$$

where Σ is

$$\begin{bmatrix}
\frac{1}{L^2} + \frac{\phi^2}{\eta^2} & \frac{-\phi}{\eta^2} & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\frac{-\phi}{\eta^2} & \frac{1+\phi^2}{\eta^2} + \frac{1}{\sigma^2} & \cdots & 0 & -\frac{1}{\sigma^2} & 0 & \cdots & 0 \\
0 & \frac{-\phi}{\eta^2} & \cdots & 0 & 0 & -\frac{1}{\sigma^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{\phi^2}{\eta^2} + \frac{1}{\sigma^2} & 0 & 0 & \cdots & -\frac{1}{\sigma^2} \\
0 & -\frac{1}{\sigma^2} & \cdots & 0 & \frac{1}{\sigma^2} & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & \frac{1}{\sigma^2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -\frac{1}{\sigma^2} & 0 & 0 & \cdots & \frac{1}{\sigma^2}
\end{bmatrix}.$$

By setting $\sigma = \eta = 1$ and giving an initial value of x_0, will have $\Sigma = \begin{bmatrix} A(\phi) & -I \\ -I & I \end{bmatrix}$ and

$$\Sigma^{-1} = \begin{bmatrix} (A(\phi) - I)^{-1} & A^{-1}(\phi)(I - A^{-1}(\phi))^{-1} \\ (A(\phi) - I)^{-1} & (I - A^{-1}(\phi))^{-1} \end{bmatrix}.$$

Then

$$x_{1:t} | y_{1:t} \sim N\left(A^{-1}(\phi)y, A^{-1}(\phi)\right),$$

and the covariance of $y_{1:t}$ is $\Sigma_t = (I - A^{-1}(\phi))^{-1}$.

3.2 Parameter Estimation

The Metropolis-Hastings algorithm makes use of proposal density $Q(\phi)$ which depends on the current stat $\phi^{(t)}$. We assume that we can evaluate $P^*(\phi)$ for any ϕ. A tentative new state ϕ' is generated from the proposal density $Q(\phi'; \phi^{(t)})$. To decide whether to accept the new state, we compute the quantity

$$\alpha = \frac{P^*(\phi')}{P^*(\phi^{(t)})} \frac{Q(\phi^{(t)}; \phi')}{Q(\phi'; \phi^{(t)})}.$$

If $\alpha \geq 1$, then the new state is accepted. Otherwise, the new state is accepted with probability α.

In our case, the proposal $\phi' \sim N(\phi^{(t)}, \sigma)$, and the density Q is symmetric, so

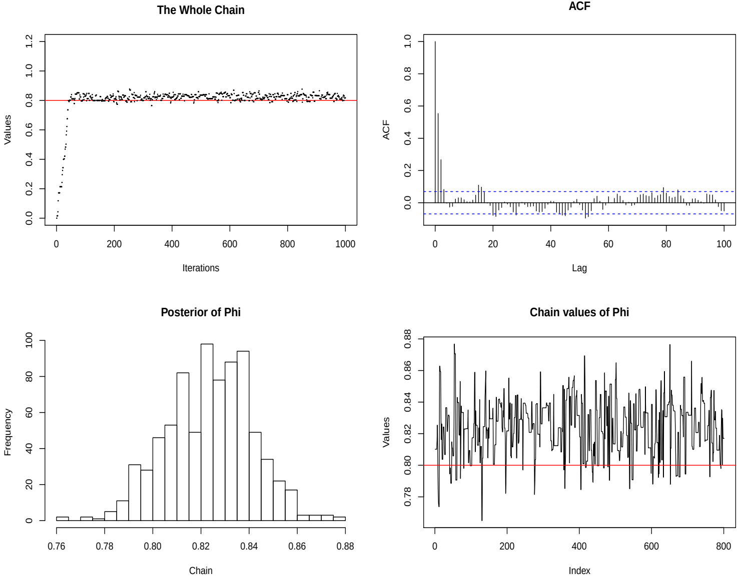
$$\alpha = \frac{P^*(\phi')}{P^*(\phi^{(t)})} = \frac{P(y_{1:t}|\phi')P(\phi')}{P(y_{1:t}|\phi^{(t)})P(\phi^{(t)})}.$$

Because $P(y_{1:t}|\phi) \sim N(0, \sigma_t)$, then the posterior of ϕ is

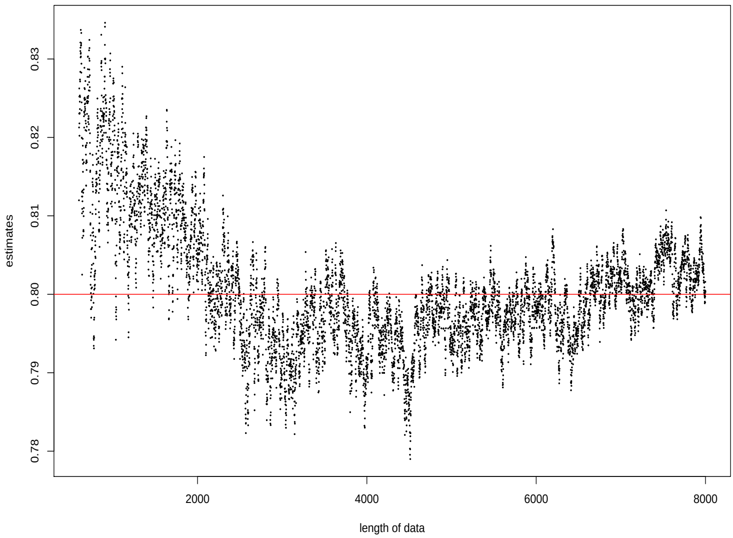
$$P(\phi|y_{1:t}) = \frac{P(y_{1:t}|\phi)P(\phi)}{P(y_{1:t})} \propto P(y_{1:t}|\phi)P(\phi)$$
$$= e^{-\frac{1}{2}y^\top (I - A^{-1}(\phi))y} \sqrt{det(I - A^{-1}(\phi))} P(\phi).$$

Additionally, $det(I - A^{-1}(\phi)) = det(A^{-1}(\phi))$, taking the Choleski decomposition of $A = LL^\top$ and ln of posterior will give

$$\ln P(\phi|y_{1:t}) = \frac{1}{2}y^\top L^{-\top} L^{-1} y - \ln \sqrt{|L^\top||L|} - \ln P(\phi)$$
$$= \frac{1}{2}u^\top u - \ln tr(L) - \ln P(\phi).$$


The Whole Chain
Values
Iterations
ACF
ACF
Lag
Posterior of Phi
Frequency
Chain
Chain values of Phi
Values
Index

Accuracy of Phi



4.1 Future Work

- Computing faster and faster.
- Estimating x_t, x_{t+1} and more.
- Deducing high-level description of vehicle motion

4.2 Acknowledgment

- 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.

Y.J. Kim and C. Gu (2004)
Smoothing spline gaussian regression: more scalable computation via efficient approximation no. 2, pp. 337-356

Blight, B.J.N. and Ott, L. (1975)
A Bayesian Approach to Model Inadequacy for Polynomial Regression.

MacKay, David JC (2003)
Information theory, inference and learning algorithms

Cappé, Olivier and Godsill, Simon J and Moulines, Eric (2007)
An overview of existing methods and recent advances in sequential Monte Carlo

The End