Smoothing through State-Space Models for Stream Networks

William J. Coar
DEPARTMENT OF STATISTICS
COLORADO STATE UNIVERSITY

April 21, 2006

Joint work with F. Jay Breidt, Colorado State University.
Outline

1. A smoothing problem → traditional spline smoother
2. Our goal
3. Stream networks
   (a) Local linear trend
   (b) State-space representation
4. Kalman recursions
5. Connection to a discrete spline smoother
6. Numerical example
One path on a stream network
How much to smooth?
Maybe a better smooth
Smooth $\mu(t)$ that minimizes a penalized least squares criterion function

$$\sum_{t=1}^{n} (y(t) - \mu(t))^2 + \lambda \sum_{t=1}^{n} (\nabla^2 \mu(t))^2$$

where $\nabla^2 \mu(t)$ is twice differenced $\mu(t)$.

- Smoothness determined by $\lambda$
- Choice of $\lambda$?
  1. Cross-validation
  2. Function of variance components in Local Linear Trend
- Local Linear Trend (LLT) is a state-space model
- Spline obtained as the Kalman smooth using this state-space representation
A new problem → two paths merge
Our Goals

Adapt time series methods to smooth the network

- Define a Local Linear Trend model
- Determine its state-space representation
- Implement Kalman recursions
- Construct a “spline” smoother on the network
Diagram of a stream network

- Reach characterized by (Strahler) order
- Except for first order reaches, each reach \( k \) has two parents \( u_1 \) and \( u_2 \)
- Some reaches have grandparents
- Process on reach \( k \) depends on the state at each parent
- Natural time-like ordering → downstream flow
- Merging with each step downstream
- For simplicity, assume equally spaced discrete locations
State-space model and Local Linear Trend

State-space model on network:

\[ Y(k) = G_k X(k) + W(k) \]

\[ X(k) = F_{k,u_1} X(u_1) + F_{k,u_2} X(u_2) + V(k) \]  

\[ (X_t = F_t X_{t-1} + V_{t-1}) \]

Local Linear Trend model:

\[ Y(k) = X(k) + W(k) \]

\[ X(k) = \frac{1}{2} (X(u_1) + X(u_2)) + B(k) + V(k) \]  

\[ (X_t = X_{t-1} + B_{t-1} + V_{t-1}) \]

\[ B(k) = \frac{1}{2} (B(u_1) + B(u_2)) + U(k) \]  

\[ (B_t = B_{t-1} + U_{t-1}) \]

with state-space components

\[ X(k) = \begin{bmatrix} X(k) \\ B(k) \end{bmatrix} \]

\[ F_{k,u_i} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1/2 \end{bmatrix} \]

\[ V(k) = \begin{bmatrix} V(k) + U(k) \\ U(k) \end{bmatrix} \]

**SPECIAL CASE:** \( V(k) = 0, \sigma_v^2 = 0 \)
Smoothing via Kalman recursions

**Downstream predict, filter, predict, filter...**

Given upstream information, predict via

\[
X^p(k) = F_{k,u_1} X^f(u_1) + F_{k,u_2} X^f(u_2) \\
\Omega^p_k = F_{k,u_1} \Omega^f_{u_1} F^T_{k,u_1} + F_{k,u_2} \Omega^f_{u_2} F^T_{k,u_2} + Q_t,
\]

Filter once observation is obtained

\[
X^f(k) = X^p(k) + \Omega^p_k G^T_k \Delta_k^{-1} (Y(k) - G_k X^p(k)) \\
\Omega^f_k = \Omega^p_k - \Omega^p_k G^T_k \Delta_k^{-1} G_k \Omega^p_k.
\]

where \( \Delta_k = G_k \Omega^p_k G^T_k + R_k \).

**Upstream smooth**

\[
\begin{bmatrix}
X^s(u_1) \\
X^s(u_2)
\end{bmatrix} = \begin{bmatrix}
X^f(u_1) \\
X^f(u_2)
\end{bmatrix} + \begin{bmatrix}
\Theta(u_1, k) \\
\Theta(u_2, k)
\end{bmatrix} (X^s(k) - X^p(k))
\]

where \( \Theta(u_i, k) = \Omega^f_{u_i} F^T_{k,u_i} (\Omega^p_k)^{-1} \).

**RESULT: Smoothed estimates** \( E(X|Y) \).
Conditional mean

Is conditional mode for Gaussian

Posterior mode: most probable $X$ given $Y$, the mode of $p(X|Y)$

Maximize $\log p(X|Y)$ with respect to $X$

• Equivalent to maximizing $\log p(Y, X)$ with respect to $X$

• Maximize

$$
-\frac{1}{2\sigma_w^2} \sum_{k=1}^{n} (Y(k) - X(k))^2 - \frac{1}{2\sigma_u^2} \sum_{k=1}^{n} (\nabla^2 X(k))^2.
$$

where $\nabla^2 X(k) = U(k)$
Conditional mode is Penalized Least Squares

- or equivalently,

\[
\sum_{k=1}^{n} (Y(k) - X(k))^2 + \frac{\sigma_w^2}{\sigma_u^2} \sum_{k=1}^{n} (\nabla^2 X(k))^2.
\]

(as was used for traditional spline)

- This defines a \textit{spline smoother} on a stream network through LLT
- Obtain estimate of smoothness parameter \( \lambda = \frac{\sigma_w^2}{\sigma_u^2} \) by MLE \( \hat{\lambda} \)
- Obtain \( E[X|Y] \) via Kalman smoother
Series of first order reaches keep merging - random inputs with every step.

For first order reaches,

- \( X(k) = m_0 + B(k) \), \( B(k) = b_0 + U(k) \)

Unknown initial conditions

- Moment estimators for \( m_0 \) and \( b_0 \)
- Naive estimators for initial prediction error variance
Example - The data

![Graph showing data points labeled Y(k) against reaches downstream.]

- The data
- Reaches downstream
- Y(k)
- Points are plotted on a graph with X-axis labeled 'Reaches downstream' and Y-axis labeled 'Y(k)'.
Example 1: $\hat{\lambda} = 1.18$ - estimated initial conditions
Results

Estimation of initial conditions

- Moment type estimators
- ML estimators?
- Try 0 with diffuse prior
- Sensitivity to initial prediction error variances

Impact of initial conditions

- With larger initial prediction error variance, more weight on observed $Y(k)$
Further work on State-Space Models

State-space model for stream network:

\[ Y(k) = G_k X(k) + W(k) \]
\[ X(k) = F_{k,u_1} X(u_1) + F_{k,u_2} X(u_2) + V(k) \]

General form is very flexible

- Can be multivariate
- A time component can be added, but process driven by flow
- State matrices are location dependent

Describe a large class of dependencies

- Class of ARMA(p,q) models can be defined
- More general structural models (LLT)
Work in progress

- Adapted state-space to a stream network
- Defined ARMA(p,q) and other structural models on a stream network
- Developed of Kalman recursions for this state-space representation
- Likelihood in terms of innovations
- An EM algorithm for missing values
- Starting to look at real data
The work reported here was developed under the STAR Research Assistance Agreement CR-829095 awarded by the U.S. Environmental Protection Agency (EPA) to Colorado State University. This presentation has not been formally reviewed by EPA. The views expressed here are solely those of the presenter and STARMAP, the Program he represents. EPA does not endorse any products or commercial services mentioned in this presentation.