URL hittp://www.stat.duke.edu/~clyde/wavelets.htm

Clyde and George: ISDS DP 98-21 and DP 99-06

Duke University

Institute of Statistics and Decision Sciences

Merlise Clyde

ESTIMATION FOR WAVELETS

FLEXIBLE EMPIRICAL BAYES
OUTLINE

1. Bayesian Hierarchical Model for Wavelets
2. Threshold and Multiple Shrinkage Estimators
3. Empirical Bayes Prior Distributions
4. Models for Outliers
5. Performance
6. Open problems
Wavelet Regression Model:  
\[ D_f \cdot g \theta \equiv W_f \cdot (d_f \cdot g \theta) \]

\( I = \hat{f} \cdot \hat{w} = \hat{w} \cdot I \) where wavelet coefficients of \( f \) are \( \hat{w} \) of \( I \), orthonormal matrix \( \hat{w} \cdot \hat{w} \cdot I \).

Apply Discrete Wavelet Transform (DWT):

Signal observed with error \( e \):
\[ e + \hat{f} = \hat{w} \]

Unknown Signal:
\[ \hat{g} \]

Observations:
\[ \hat{y} \]

Non-parametric Regression Model:
MODEL
GLINT SERIES

Glint Radar Signal
Discrete Wavelet Transformation
Sparse Representation: $\mathcal{g}$

$0 = \mathcal{g}^T \mathcal{f}$ when $0 = \mathcal{g}^T \mathcal{f}$

Bernoulli Prior

$\mathcal{g} \sim \mathcal{B}(\mathcal{f})$

Normal Errors

$\mathcal{g} \sim \mathcal{N}(0, \mathcal{f})$

Hierarchical Model in Wavelet Domain
\[ f_{\mathcal{M}}(\mathbf{A}) = f_{\mathcal{M}}(\mathbf{A})^{\nu} - 1 \]

The posterior probability of a model is

\[ f_{\mathcal{M}} \left( \frac{c + 1}{c} \frac{\frac{m}{2} \frac{a}{l}}{1} \right) \exp\left( \frac{c + 1}{c} \frac{m}{2} \frac{a}{l} - 1 \right) = f_{\mathcal{O}} \]

where \( f_{\mathcal{O}} \) is the posterior odds that the coefficient is non-zero:

\[ \frac{f_{\mathcal{O}} + 1}{f_{\mathcal{O}}} = (\mathbf{A}|I = \gamma f_{\mathcal{O}})^\nu \]

Because of the conditional independence structure in the prior and orthogonality of the DWT, the \( \gamma \) is a posterior independent Bernoulli random variables.

\( \gamma \) is a posterior Distribution of \( \gamma \)
Transform estimates via IDWT to estimate ρ:

\[
\begin{align*}
0.5 > \frac{c}{d} | I = \gamma \lambda \nu & \quad \iff \quad 0 = \gamma \lambda \\
0.5 \leq \frac{c}{d} | I = \gamma \lambda \nu & \quad \iff \quad 1 = \gamma \lambda 
\end{align*}
\]

where

\[
\frac{c}{d} \gamma \lambda \nu + I = \gamma \lambda 
\]

under squared error loss with model selection.

Posterior Mean Given Highest Probability model (optimal):

\[
\frac{c}{d} | I = \gamma \lambda \nu = \gamma \lambda 
\]

Bayes Threshold and Shrinkage Estimators

Threshold Estimator:

\[
\frac{c}{d} \gamma \lambda \nu + I = \gamma \lambda 
\]

Posterior Mean (minimizes posterior expected squared error loss):
Closed Formed Iterative Solutions

4. Repeat 2 and 3 until convergence

(\text{New estimates of } (C',m', \omega)^2)

Maximization of complete data likelihood

3. M-Step:

\text{Expectionation of missing data given estimates of } (C',m', \omega)^2

2. E-Step:

(\forall \gamma (A, c) = \log \sum \sum_{m} (D, M, \gamma) + (m, \gamma, C', \omega)_{\gamma}^{m, \gamma, C', \omega})

EM algorithm

Bayesian estimation of \gamma and \omega, and \omega, from the marginal likelihood of the data
\( \frac{\gamma_j^l \gamma_j^l}{(1+\gamma_j^l)\gamma_j^l} = (1+\gamma_j^l)\gamma_j^l \)

\[
\left( I - \frac{\gamma_j^l \gamma_j^l}{(1+\gamma_j^l)\gamma_j^l} \right) \max = (1+\gamma_j^l)\gamma_j^l
\]

\[
\frac{\gamma_j^l \gamma_j^l}{(1+\gamma_j^l)\gamma_j^l - 1} = (1+\gamma_j^l)\gamma_j^l
\]

**M-Step:**

\[ \gamma_j^l \text{ is the posterior mean of } \gamma_j^l \]

\[
(\lambda | I = \gamma_j^l \gamma_j^l) \psi = \frac{\gamma_j^l \psi + I}{\gamma_j^l \psi} = (1+\gamma_j^l)\gamma_j^l = (\gamma_j^l \gamma_j^l) \mathcal{E}
\]

**E-Step:**
EB Estimates of $c_j^{1/2}$

- **doppler**
- **blocks**
- **heavisine**
- **bumps**
EB Estimates of \( w \)
Normal Error MSE Comparisons

Boxes for different conditions:
- **Doppler:**
  - EB
  - EMS
  - AMS
  - ETS
  - ATSS

- **Blocks:**
  - EB
  - EMS
  - AMS
  - ETS
  - ATSS

- **Heavisine:**
  - EB
  - EMS
  - AMS
  - ETS
  - ATSS

- **Bumps:**
  - EB
  - EMS
  - AMS
  - ETS
  - ATSS
Standard Normal Model \( \mathcal{N} \) and \( \mathcal{G} \) have heavy tailed distributions and \( \mathcal{G} \) having gamma distributions

\[
\begin{align*}
\text{Bernoulli}(\theta) & \sim \mathcal{N} \\
\left( \frac{\theta}{\theta^*}, 0 \right) N & \sim \mathcal{D} \\
\left( \frac{\theta}{\theta^*}, 0 \right) N & \sim \mathcal{G} \\
\mathcal{G} + \mathcal{G} & = \mathcal{D}
\end{align*}
\]

Hierarchical Model for Outliers
<table>
<thead>
<tr>
<th>But Dependent</th>
<th>Uncorrelated</th>
<th>Independent</th>
<th>Normal-Normal</th>
<th>Bivariate t</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t^n$ errors</td>
<td>$t^n$ priors</td>
<td>$t^n$ errors</td>
<td>$t^n$ errors</td>
<td>$t^n$ errors</td>
</tr>
<tr>
<td>Cauchy prior</td>
<td>Normal priors</td>
<td>Normal errors</td>
<td>Normal errors</td>
<td>Normal errors</td>
</tr>
<tr>
<td>$\mathcal{N}(1/2, 2/\nu)$</td>
<td>$\mathcal{N}(1/2, 2/\nu)$</td>
<td>$\mathcal{N}(1/2, 2/\nu)$</td>
<td>$\mathcal{N}(1/2, 2/\nu)$</td>
<td>$\mathcal{N}(1/2, 2/\nu)$</td>
</tr>
<tr>
<td>$\mathbb{Y}_* \equiv \mathbb{Y}$</td>
<td>$\mathbb{Y}_* \equiv \mathbb{Y}$</td>
<td>$\mathbb{Y}_* \equiv \mathbb{Y}$</td>
<td>$\mathbb{Y}_* \equiv \mathbb{Y}$</td>
<td>$\mathbb{Y}_* \equiv \mathbb{Y}$</td>
</tr>
<tr>
<td>Distibutions</td>
<td>$t^n$ and $\mathbb{E}_* t^n$</td>
<td>$t^n$ and $\mathbb{E}_* t^n$</td>
<td>$t^n$ and $\mathbb{E}_* t^n$</td>
<td>$t^n$ and $\mathbb{E}_* t^n$</td>
</tr>
</tbody>
</table>

Label
Normal Error MSE Comparisons
Student t Errors in Wavelet Domain - MSE Comparisons
Student's t Errors in Data Domain - MSE Comparisons
Persistence

Clustering of Coefficients

Conditional Independent Priors – Tree Structures

2. Wavelet Packets and Basis Selection

Empirical Wavelet Coefficients are Dependent

1. Non-Decimated Wavelet Transform

Open Problems