On l_1 Mean and Variance Filtering

Bo Wahlberg, Cristian Rojas, Mariette Annergren

Main Results:

- An l_1, Total Variation (TV), regularized Maximum Likelihood (ML) method to segment a time series with respect to changes in the mean or in the variance.
- We show that, in this setting, variance estimation of $\{y_t\}$ is equivalent to mean estimation of $\{y_t^2\}$.

Why?

- Estimating means, trends and variances in time series data are of fundamental importance in a variety of areas. Typically done to pre-process data before estimation of, for example, parametric models.
- For non-stationary data it is important to detect changes in the mean and the variance in order to segment the data into stationary subsets.

How?

Traditionally: Moving window sample mean and variance estimation. Hypothesis test based change detection.

Here: The l_1 sparseness approach. Penalize the difference between consecutive variables.

Mean Estimation

Data: $\{y_1, \dots y_N\}$

Model: $y_t \sim \mathcal{N}(m_t, 1)$, **where** $m_{t+1} = m_t$ **often**

Method: $$\min_{m_t} \left[\frac{1}{2} \sum_{t=1}^{N} (y_t - m_t)^2 + \lambda \sum_{t=1}^{N} |m_t - m_{t-1}| \right]$$

ML+TV: Related to fused lasso, l_1 trend filtering and total variation denoising

Variance Estimation

Data: $\{y_1, \dots y_N\}$

Model: $y_t \sim \mathcal{N}(0, \sigma_t^2)$, **where** $\sigma_{t+1} = \sigma_t$ **often**

ML: $$\min_{\sigma_t > 0} \frac{1}{2} \left[\sum_{t=1}^{N} \ln(\sigma_t^2) + \sum_{t=1}^{N} \frac{y_t^2}{\sigma_t^2} \right]$$

Concave + Convex! Standard trick: $\eta_t = -1/(2\sigma_t^2)$

Method: $$\min_{\eta_t < 0} \left[\frac{1}{2} \left(\sum_{t=1}^{N} -\ln(-\eta_t) - \sum_{t=1}^{N} 2\eta_t y_t^2 \right) + \lambda \sum_{t=1}^{N} |\eta_t - \eta_{t-1}| \right]$$

ML+TV: Convex optimization problem related to graphical lasso.

Equivalence

The variance estimation problem for $\{y_t\}$ has same subgradient (first order) optimality conditions as the mean estimation problem for $\{y_t^2\}$.

Proof idea:

$$\frac{d}{d\eta_t} \left[\sum_{t=1}^{N} -\ln(-\eta_t) - \sum_{t=1}^{N} 2\eta_t y_t^2 \right] = \frac{-1}{\eta_t} - 2y_t^2 = 2(\sigma_t^2 - y_t^2)$$

$$= \frac{d}{d\sigma_t^2} \left[\sum_{t=1}^{N} [y_t^2 - \sigma_t^2]^2 \right]$$

$\eta_t - \eta_{t-1} = \dfrac{\sigma_t^2 - \sigma_{t-1}^2}{2\sigma_t^2 \sigma_{t-1}^2}$ $\Rightarrow$ *Ordering is preserved* $\Rightarrow$

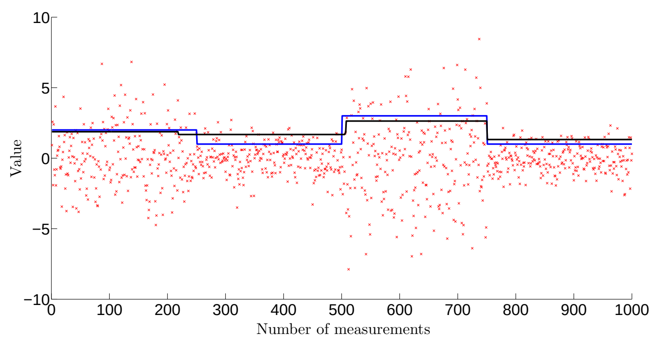
$$\frac{d}{d\eta_t} \left[\sum_{t=2}^{N} |\eta_t - \eta_{t-1}| \right] = \frac{d}{d\eta_t} \left[|\eta_t - \eta_{t-1}| + |\eta_{t+1} - \eta_t| \right]$$

{depends only on the signs of $(\eta_t - \eta_{t-1})$ and $(\eta_{t+1} - \eta_t)$, which equal the signs of $(\sigma_t^2 - \sigma_{t-1}^2)$ and $(\sigma_{t+1}^2 - \sigma_t^2)$}

$$= \frac{d}{d\sigma_t^2} \left[|\sigma_t^2 - \sigma_{t-1}^2| + |\sigma_{t+1}^2 - \sigma_t^2| \right] = \frac{d}{d\sigma_t^2} \left[\sum_{t=2}^{N} |\sigma_t^2 - \sigma_{t-1}^2| \right]$$

Ongoing and Future Work

- The vector valued covariance matrix case: $(n+1)n/2$ variables per n-dimensional sample.
- Alternating Direction Method of Multipliers (ADMM) convex optimization algorithm with linear complexity.
- Statistical analysis and applications.



Estimated variance (black line), true variance (blue line) and measurements (red crosses).

Automatic Control Lab and ACCESS, KTH Royal Institute of Technology, Stockholm, Sweden.

The research leading to these results has received funding from Swedish Research Council, the ACCESS Linnaeus Center at KTH and the European Research Council under the European Community's Seventh Framework program (FP7 2007-2013) / ERC Grant Agrement N. 267381