View Reviews

Paper ID
3656

Paper Title
Modeling trend in temperature volatility using generalized LASSO

Reviewer #1

Questions

1. Please provide an "overall score" for this submission.
4: An okay submission, but not good enough; a reject. I vote for rejecting this submission, although I would not be upset if it were accepted.

2. Please provide a "confidence score" for your assessment of this submission.
5: You are absolutely certain about your assessment. You are very familiar with the related work.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.
Overall Comments:
The paper extends ideas from trend filtering, which in the literature is developed in the least-squares setting, i.e. Eq1 (c.f. Tibshirani 2014). These ideas are themselves a development of work similar to Tibshirani's 2005 paper "Sparsity and smoothness via the fused lasso". While it is true that much work does focus on trends in the mean structure, there is a very large literature on modelling non-stationary variances, for instance in a regularised smoothing setting similar to (Monti et al. 2014, NeuroImage, Gibberd et al. 2017, JCGS, Hallac et al. ACM), or alternatively modelling the localised auto-covariance via wavelet based models (Nason et al. 2000, RSSB, Gibberd et al. 2016 IEEE-SSP), The authors cite Hallac et al. 2017 which is similar to the Monti/Gibberd papers, and uses a likelihood specifically designed to capture time-varying second-order structure (i.e. covariance/precision) in a multivariate signal. In respect of this work, the proposed method seems to be a simplification, i.e. it studies changes in a univariate signal although sampled at points in space and time. While it may be true that in the context of climate science extracting trends in the variance is a relatively unexplored problem, this is definitely not the case more generally. At a minimum the paper should mention and discuss the relation to this previous work.

Strengths:

- The paper is relatively well written, it provides a real-world and well developed application of statistical/machine learning methods.

Weaknesses:

- References to previous literature which model time-varying, and/or spatially-varying volatility are quite severely lacking (see overall comments).

- The algorithm development is not particularly novel, rather the application of methods laid out in Boyd et al. 2011. Similar algorithms have been developed for fusing in hyper-spectral imaging c.f. Iordache et al.
- The authors point out that temporal and spatial correlations between the observations are ignored in previous work.. "Third and most importantly, temporal and spatial correlations between the observations are ignored". However, the proposed model does not incorporate any mechanism for spatial auto-correlation. The likelihood term still assumes observations are independently drawn, even though parameters may be "fused".
- There is limited comparison to alternative methods, the simulated data example is rather confusing, and of limited scope. I.e. how does performance vary as a function of dimensions? Why is the sampled "spatial" dimensionality so low, i.e. 5x7? How does this contrast with the application in say Figure 4.
- What data is used to construct Figure 4? What is the dimensionality of the spatial observations?

Quality: 2/5
Clarity: 3/5
Orginality: 2/5
Significane: 2/5

References:

Nason et al. "Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum." JRSSB, 1–28, 2000

Tibshirani et al. "Sparsity and smoothness via the fused lasso", RSSB, 2005

Iordache et al. "Total variation spatial regularization for sparse hyperspectral unmixing." IEEE Transactions on Geoscience and Remote Sensing, 50(11 PART1):4484–4502, 2012

Monti et al. "Estimating time-varying brain connectivity networks from functional MRI time series." NeuroImage, 103, 427–443, 2014

Gibberd et al. "Regularised Estimation of 2D-Locally Stationary Wavelet Processes". IEEE Workshop on Statistical Signal Processing (SSP), 2016

Gibberd et al. "Regularized Estimation of Piecewise Constant Gaussian Graphical Models : The Group-Fused Graphical Lasso." JCGS, 2017

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?
2: Somewhat confident

Reviewer #2

Questions

1. Please provide an "overall score" for this submission.
6: Marginally above the acceptance threshold. I tend to vote for accepting this submission, but rejecting it would not be that bad.

2. Please provide a "confidence score" for your assessment of this submission.

4: You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.

This paper presented a method to study the temperature variance over time and space. Starting from a probabilistic model following an earlier in ref[12] that models temperature mean, this paper models the temperature variance as an optimization problem that consists of an L1-norm regularization term to impose temporal smoothness. After this, the optimization objective was future developed with two additional L1-norm regularization terms to take in account spatial continuity. To solve the resulted optimization objective in Eq. (3), two solutions were derived using the ADMM method in Section 3. Comprehensive experiments have been conducted to validate the proposed method in the application of temperature trend estimation.

+ Strengths:
1. Overall this paper is clearly written and well organized.
2. The application problem studied in this paper should be interesting in climate research. From the application perspective, this paper studies temperature variance, but not temperature mean, which may be more valuable to study the impacts of extreme weathers.
3. The application problem is reasonably formalized as an optimization problem that consists of three regularization terms using L1-norms. To solve the optimization problem, two solutions are derived using the ADMM framework.
4. Experimental evaluations are comprehensive. The results have demonstrated the correctness of the hypothesis and the usefulness in the application to study temperature volatility.

- Weaknesses:
1. From machine learning perspective, the main contribution of this paper is to formalize the application problem and to derive the mathematical solutions. For the former, as mentioned above, the problem has been reasonably formulated. For the latter, the optimization objective in Eq. (3) proposed in this paper indeed is just an L1-norm regularized minimization problem. In the past couple decades, such problems have been extensively studied and many nice solution frameworks had been proposed, such as reweighted method, proximal method, etc. My question is that, among many possible ways to derive the solution algorithms, why do the authors choose to derive the solution algorithm using the ADMM framework? Is this better than other solution frameworks, such as reweighted method introduced in the below paper, in terms of convergence rate/optimal solution? If yes, some empirical studies may be necessary.
Nie et. al, Efficient and Robust Feature Selection via Joint L2,1-Norms Minimization, NIPS 2010.
In recent researches, for example as in the below paper, it has been recognized that the convergence of the solution algorithms derived by the ADMM method should be carefully studied before one can claim it really converges. As can be seen in the middle panel of Figure 2, the objective value of the derived solution algorithm does not monotonically decrease during iterations, though it finally converges.
Hong et. al, On the linear convergence of the alternating direction method of multipliers, Mathematical Programming, 2016.
Thus, some discussion on the algorithm selection may be useful to make the paper more convincing.

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?

2: Somewhat confident

Reviewer #3

Questions

1. Please provide an "overall score" for this submission.
4: An okay submission, but not good enough; a reject. I vote for rejecting this submission, although I would not be upset if it were accepted.

2. Please provide a "confidence score" for your assessment of this submission.
4: You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.
The paper presents a method to estimate trends in variance or volatility of correlated spatiotemporal data. The problem is addressed by modeling variance with smoothness constraints over space and time and two different ADMM algorithms were developed. Experiments are conducted over simulated data and spatiotemporal climate data with GARCH(1,1) as baselines.

- The idea of extending trend filtering to the spatiotemporal domain and finding efficient algorithms to solve a larger constrained optimization problem is interesting.

- In Section 1, the paper attributes climate change to change in variance. However, in fig 5, the proposed method does not show much change in variance (-0.2 - 0.1). It would help if the authors add discussions about this.

- Only GARCH(1,1) is compared in the experiments. It would be helpful to judge the merit of the proposed method if more relevant spatio-temporal methods/ models can be compared with.

- Fig 3 is hard to read as the lines are densely populated. It is very hard to figure out the upward trend and if it is even statistically significant.

- The authors argue to use variance as the proxy for volatility, however, it would be nice to see a brief review of what other metrics can be used or has been used (e.g., extreme values)

- I would like to see some discussion about applications/extensions to finance or other domains. They are not very straightforward.

- Notations: S_k in alg 1 is not defined. In alg 2 soft thresholding operation is unclear.

Typos / errors:
- line 12: relavant -> relevant

- line 230: The top row of Figure 3 -> figure 3 has only one row
- line 81: likelihood should be $-(h_t + y_t^2 \exp{-h_t})$.
- line 105/110: N_s should be $T.(2n_r.n_c - n_r - n_c)$ and accordingly number of rows of D_{ST} will change too
- Figure 3: There is something wrong about y-axis, how is it possible to go below 0 K?

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?

2: Somewhat confident