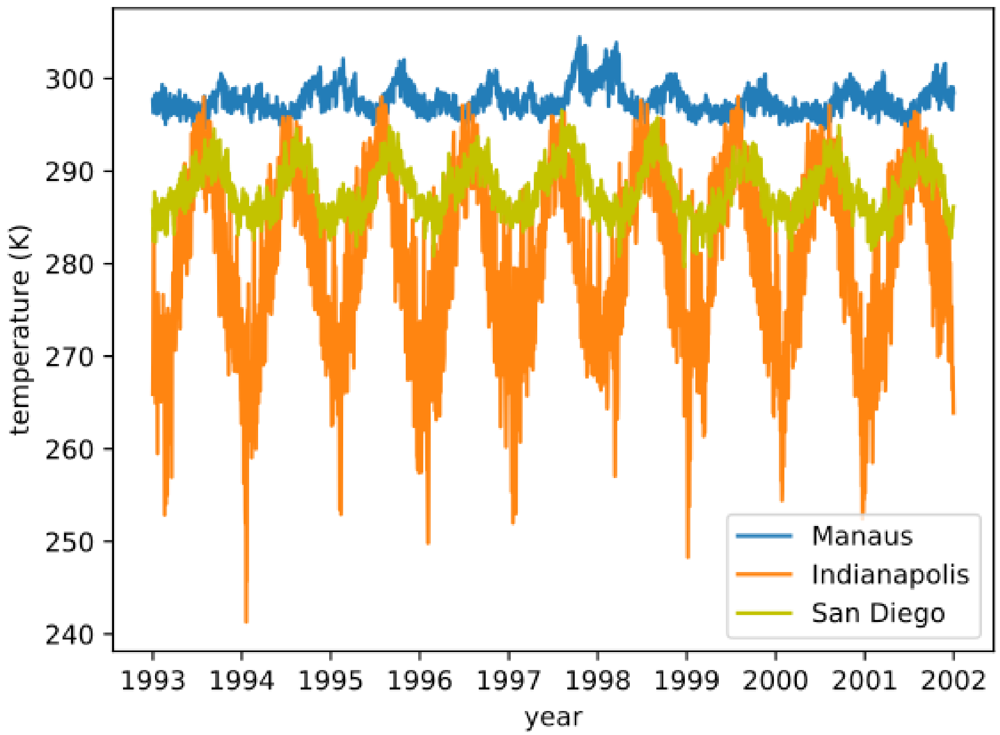
Algorithms for Estimating Trends in Global Temperature Volatility

Arash Khodadadi and Daniel McDonald

Time-series of the temperature (in Kelvin) of three cities



Modeling trend in variance

- We have 52000 of these time-series
- They have different trends and periodic components
- The periodic component is not necessarily sinusoidal.
- We don't really care about the shape of the trend and the periodic terms. We just want to get rid of them!

These observations suggest a non-parametric approach.

Modeling trend in variance

- We need to model the trend in the variance of the time-series

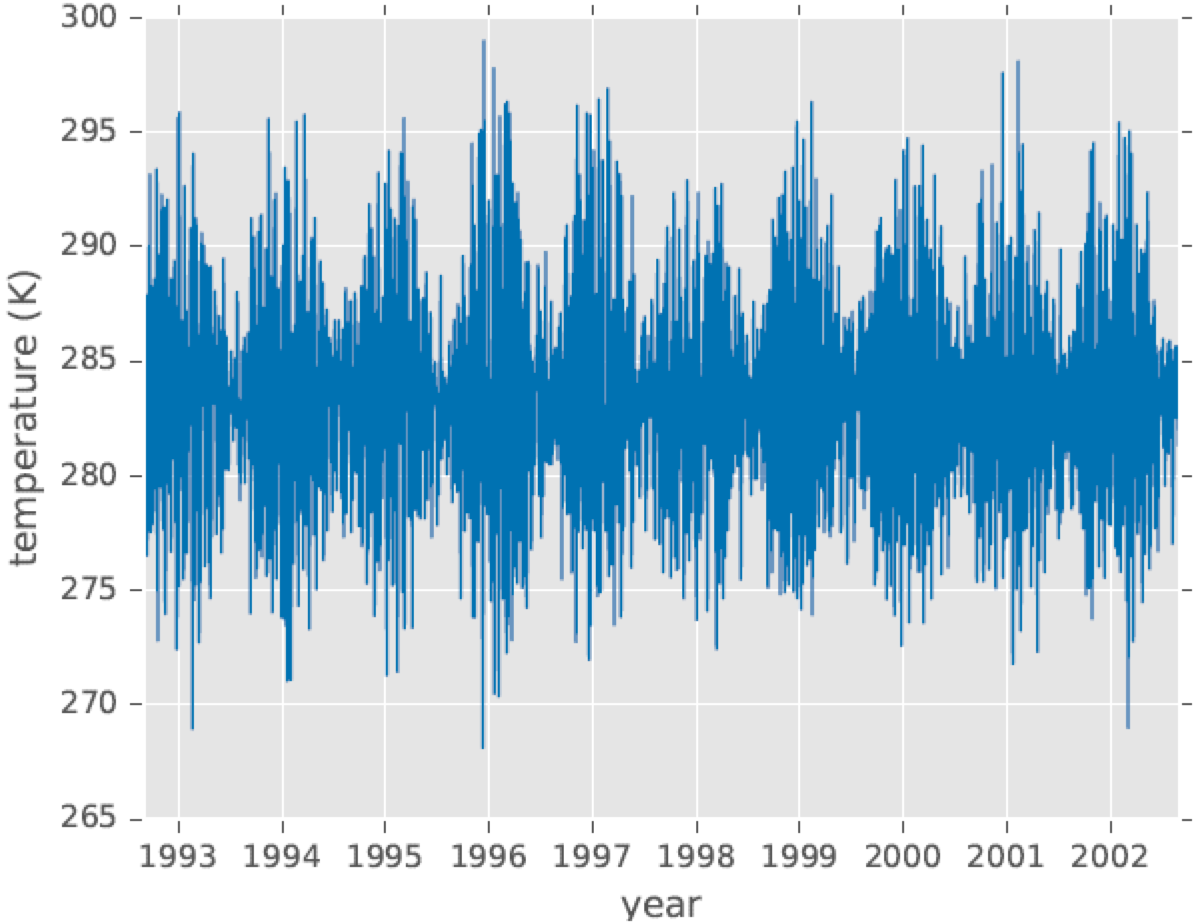
- Data is spatio-temporal

Outline

- l_1-trend filtering for removing trend and cyclic terms
- Extension to filtering the variance of a time-series
- Extension to filtering the variance of spatio-temporal data

L1 trend filtering of variance

Motivating example



L1 trend filtering of variance

$$y_t \sim N(0, exp(h_t))$$

Here, ht is a hidden (unobserved) state.

Penalized MLE estimation of variance leads to the following optimization problem:

$$\min_{\mathsf{h}} -l(\mathsf{h}, \mathsf{y}) + \lambda ||D\mathsf{h}||_1$$

$$-l(\mathsf{h}, \mathsf{y}) = \sum_t h_t + y_t^2 e^{-h_t}$$

L1 trend filtering of variance

- We solve this optimization problem using the primal-dual interior point method.

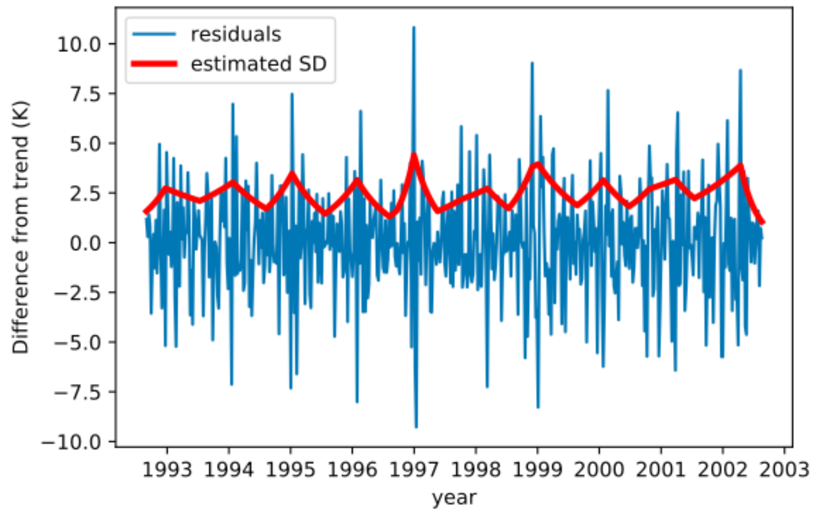
- The dual is:

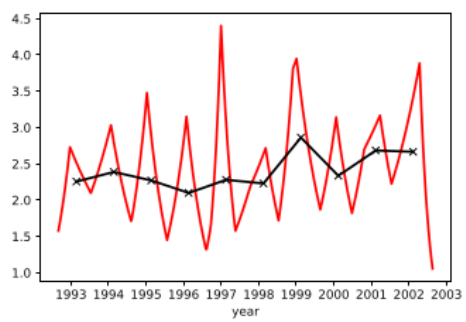
$$\min_{\nu} f^*(-D^T\nu)$$
$$s.t. \ ||\nu||_\infty \leq \lambda$$

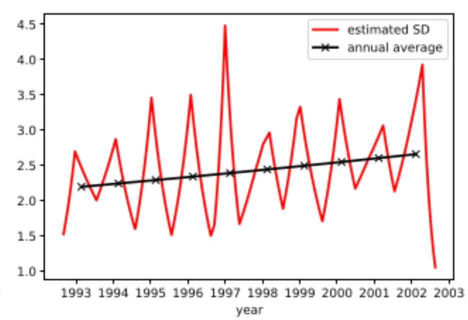
Here, $f^*(\cdot)$ is the conjugate of f and we have:

$$f^*(\mathrm{u}) = \sum_t (u_t - 1)\log\frac{y_t^2}{1-u_t} + u_t - 1$$

L1 trend filtering of variance





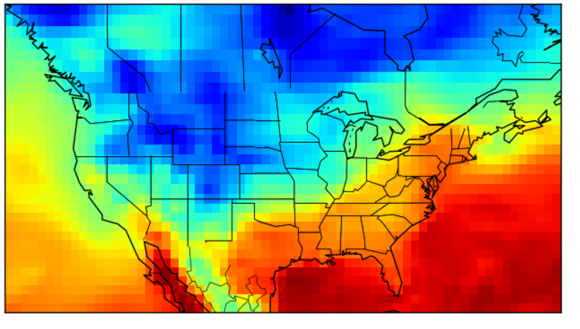


L1 trend filtering on graph

So far we have only considered temporal data.

Spatio-temporal data:

- penalize differences in the variance of neighboring points
- In particular, for each add a row to such that and another row such that .
- This results in spatially piecewise constant variance.
- Graph LASSO at each time step



L1 trend filtering on graph

Optimization:

We need to solve the same optimization problem:

$$\min_{\nu} f^*(-D^T\nu)$$
$$s.t. \ \ ||\nu||_\infty \le \lambda$$

Here, $f^*(\cdot)$ is the conjugate of f and we have:

$$f^*(\mathsf{u}) = \sum_t (u_t - 1)\log\frac{y_t^2}{1-u_t} + u_t - 1$$

If the spatial grid is $n_r \times n_c$ and the time-series are of length T, then the size of D is:

$$number\ of\ rows = 3n_r n_c T - Tn_c - 2n_r n_c$$

$$number\ of\ columns = n_r n_c T$$

L1 trend filtering on graph

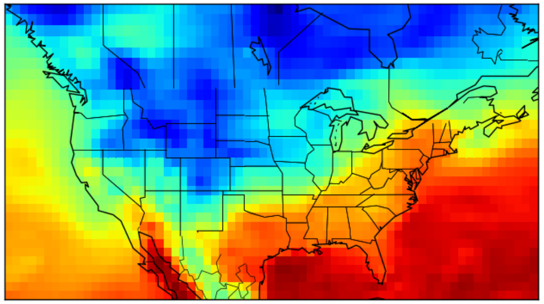
Optimization:

If the spatial grid is $n_r \times n_c$ and the time-series are of length T, then the size of D is:

$$number\ of\ rows = 3n_r n_c T - T n_c - 2n_r n_c$$

$$number\ of\ columns = n_r n_c T$$

Example: For the grid shown here $n_r = 32, n_c = 68, T = 521$ and so D is of size: **3,680,160×1,239,980 !!!**



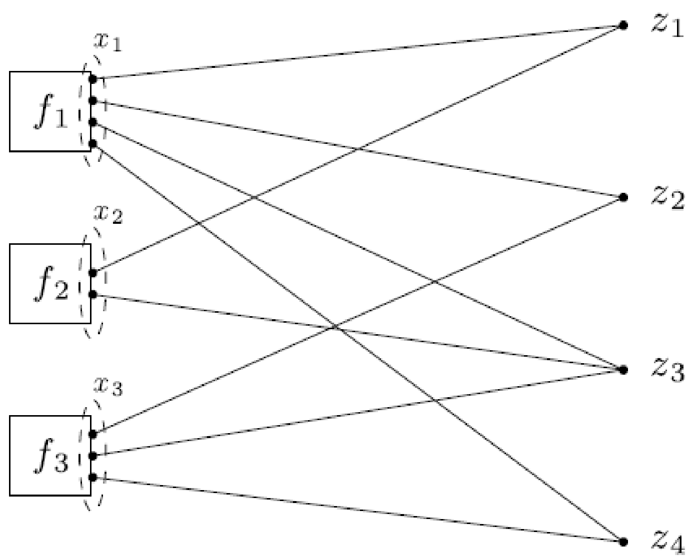
Consensus Alternative Direction Method of Multipliers (ADMM):

Consider the following optimization problem:

$$\min_{z} \sum f_i(z)$$

Now, define the *local variables* x_i such that $\sum f_i(z) = \sum f_i(x_i)$, then the above optimization is equivalent to:

$$\min_{\{x_1,\ldots,x_N\}} \sum f_i(x_i)$$
$$s.t.: x_i - z_i = 0\,, i = 1, \ldots, N$$



Consensus ADMM for L1 trend filtering on graph

$$\min_{\{x_1,\ldots,x_N\}} \sum f_i(x_i)$$
$$s.t.: x_i - z_i = 0\ , i = 1, \ldots , N$$

ADMM steps:

$$x_i^{k+1} := \underset{x_i}{\operatorname{argmin}} \left(f_i(x_i) + y_i^{kT} x_i + (\rho/2) \| x_i - \tilde{z}_i^k \|_2^2 \right)$$

$$z_g^{k+1} := (1/k_g) \sum_{\mathcal{G}(i,j)=g} (x_i^{k+1})_j,$$

$$y_i^{k+1} := y_i^k + \rho(x_i^{k+1} - \tilde{z}_i^{k+1}),$$

Consensus ADMM for -trend filtering on graph

Each x-update step has the following form:

$$\min_h \sum_{i,j,t} h_{i,j,t} + y_{i,j,t} \cdot \exp(-h_{i,j,t}) + \lambda \|Dh\|_1 + \frac{\rho}{2}\|h - \alpha\|_2^2$$

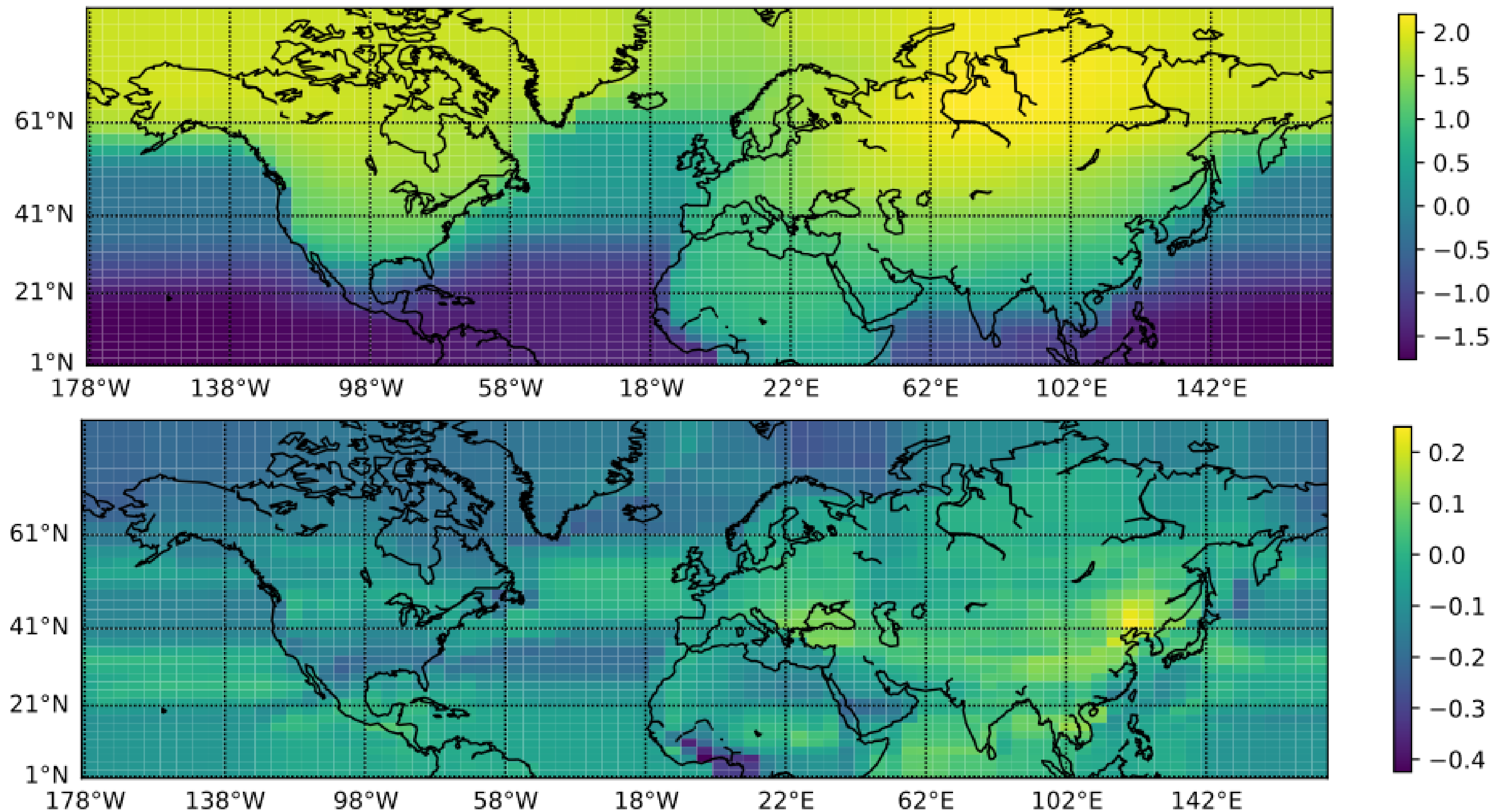
The dual is:

$$\min_\nu f^*(-D^T \nu)$$
$$s.t.\ \ ||\nu||_\infty \leq \lambda$$

Where the conjugate function is obtained as follows:

$$f^*(u) = \max_h\ u^T h - h - [diag\ (y)]^2 e^{-h} - \frac{\rho}{2}\|h - \alpha\|_2^2$$

This leads to solving equations in the form Lambert function



61°N
41°N
21°N
1°N
178°W
138°W
98°W
58°W
18°W
22°E
62°E
102°E
142°E
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4