Nov 21 Update Report

2022-11-19

Outline

- Recap
- Show results
- Discussion
- Adding in outlier term

Recap 1

- Setup:
  - Let y_t be the daily case count at day t
  - Then $y_t \sim Pois(r_t * w_t)$, where $w_t = \sum_a y_{t-a} w_a$
- Objective function:
  - $argmin_r \frac{1}{n}(\sum_{i=1} w_i r_i - y_i log(w_i r_i)) + \lambda||Dr||_1$
- Scaled Augmented Lagrangian:
  - Let $Dr = z$, adding penalty for being not equal
  - $L(r, u, z) = \frac{1}{n}(\sum_{i=1} w_i r_i - y_i log(w_i r_i)) + \lambda||z||_1 + \frac{\rho}{2}||Dr - z + u||_2^2 + \frac{\rho}{2}||u||_2^2$
- Update step for r
  - $r \leftarrow argmin_r \frac{1}{n}(\sum_{i=1} w_i r_i - y_i log(w_i r_i)) + \frac{\rho}{2}||Dr - z + u||_2^2$

Recap 2

- Linearize the update step of r
- If penalizing Dr: $r \leftarrow argmin_r \frac{1}{n}(\sum_{i=1} -w_i r_i + y_i log(w_i r_i)) + \rho r^T(D^T D r^o - D^T z + D^T u) + \frac{\mu}{2}||r - r^o||_2^2$
- If penalizing $Dlog(r)$: $r \leftarrow argmin_r \frac{1}{n}(\sum_{i=1} -w_i r_i + y_i log(w_i r_i)) + \rho r^T(D^T D r^o - D^T z + D^T u)(r^o)^{-1} + \frac{\mu}{2}||r - r^o||_2^2$

Pseudocode for summary

Initialize r^o, u^o, z^o Until converge

- Find r that solves $\frac{d}{dr}\frac{1}{n}(\sum_{i=1} -w_i r_i + y_i log(w_i r_i)) + \rho r^T(D^T D r^o - D^T z + D^T u) + \frac{\mu}{2}||r - r^o||_2^2$
- $z = sign(z^o) * (|z^o| - (D * r - u))$
- $u = u^o + Dr - z$

Finalizing r update

- KKT stationarity condition:

$$\frac{d}{dr}\frac{1}{n}(\sum_{i=1} -w_i r_i + y_i log(w_i r_i)) + \rho r^T(D^T D r^o - D^T z + D^T u) + \frac{\mu}{2}||r - r^o||_2^2$$
$$\implies -\frac{y_i}{n}\frac{1}{r_i} + \mu r_i + \rho(D^T D r^o - D^T z + D^T u)_i - \mu r_i^o + \frac{w_i}{n} = 0$$

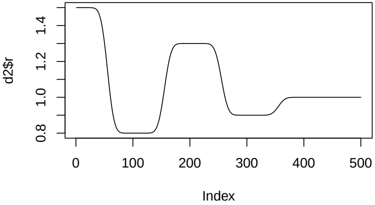
- Multiply both side by r_i, because r_i non-zero, and solve using quadratic equation
- Similarly, if penalize $log(r)$, then

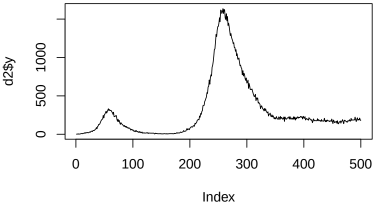
$$\implies -\frac{y_i}{n}\frac{1}{r_i} + \mu r_i + \rho(D^T D r^o - D^T z + D^T u)_i (r_i^o)^{-1} - \mu r_i^o + \frac{w_i}{n} = 0$$

- Question: Quadratic equation has two solutions
- "If this is satisfied uniquely (i.e., above problem has a unique minimizer), then the corresponding point must be the primal solution" - Geoff Gordon & Ryan Tibshirani's lecture slide

Synthetic dataset

Reference synthetic dataset



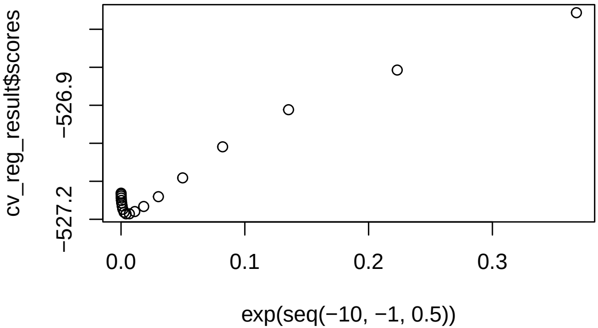


Trend Filter CV

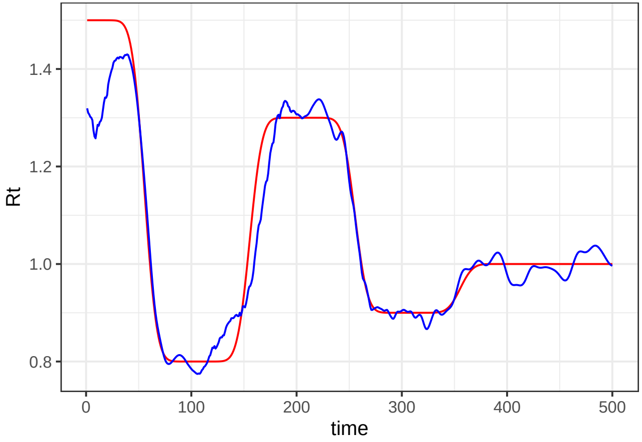
- Optimal ρ chosen by cv (same as last time).

```
## [1] 0.004086771
```

- Notice the scores are very similar

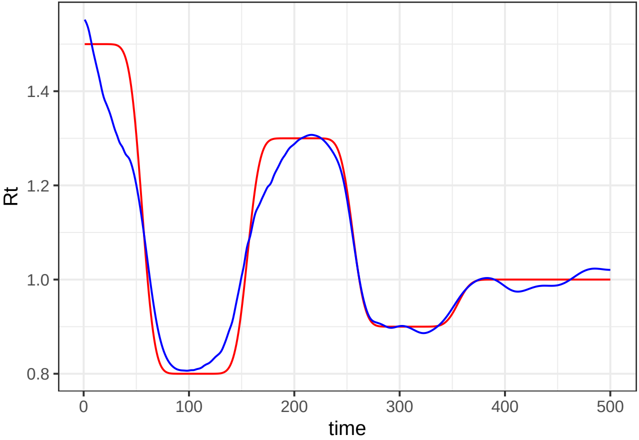


Trend Filter Fit

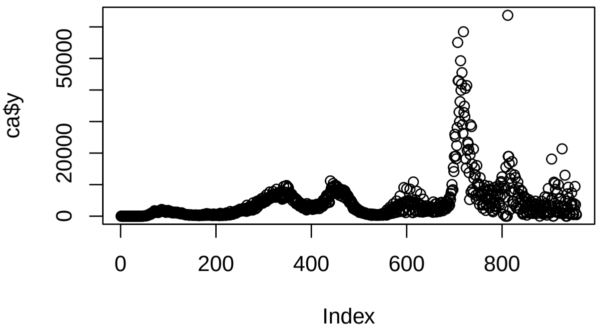


Trend Filter Fit, pick ρ myself

- Since scores are similar, pick a ρ that looks better

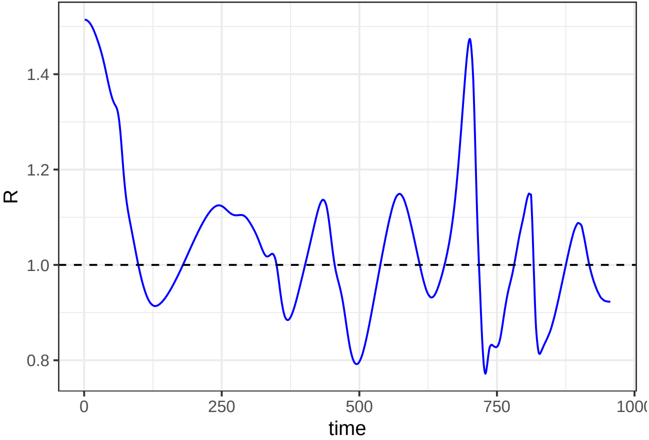


Canadian case

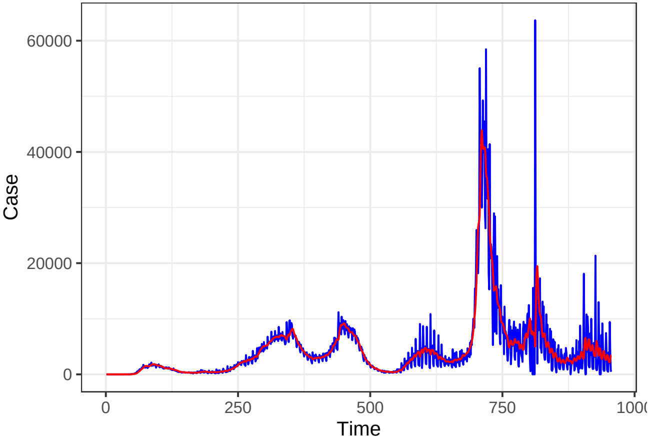


CA Trend Filter fit

- $\rho = 0.1$

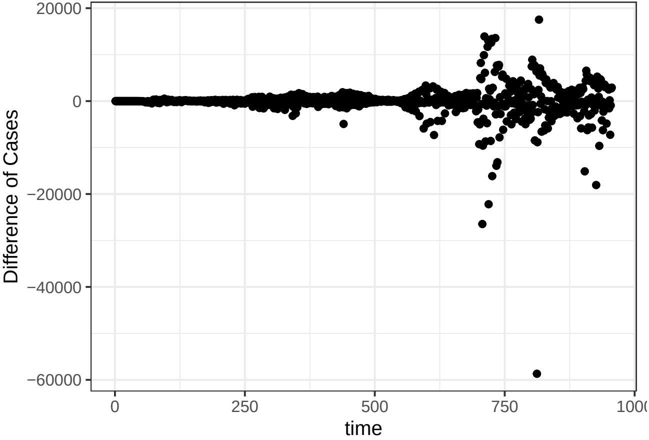


One day ahead prediction



Diagnostics plot

- Is residual plots appropriate for non-parametric regression?



Check for additions

- Model outlier (Done)
- Maximizing log likelihood of Poisson or Normal of true case count given predicted.
  - Is it reasonable to use a negative binomial distribution?
- Difference matrix D here is assumed to be of lag 1 and order 1
  - Higher order D makes sense? Or should we make it so that the degree can be chosen via CV.

Justifying outlier term

Is modeling an outlier term necessary?

- Pascal et al. "Nonsmooth convex optimization ... against low quality data"
- $y_t \sim Pois(r_t * w_t + o_t)$
- Then add L1 penalty $\sum |o_t|_1$
- Modeling outliers gives better residuals. Differences between predicted and true case count complemented by the outlier term.

Adding outlier term

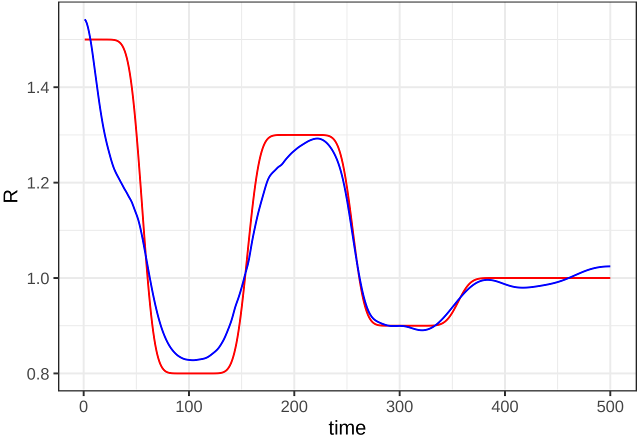
- $y_t \sim Pois(r_t * w_t + o_t)$
- For simplicity, use L2 penalty on o_t.

$$L(r, o, u, z) =$$
$$\frac{1}{n}(\sum_{i=1} w_i r_i + o_i - y_i log(w_i r_i + o_i)) + \gamma||o||_2^2 + \lambda||z||_1 + \frac{\rho}{2}||Dr - z + u||_2^2 + \frac{\rho}{2}||u||_2^2$$

- r update step is changed slightly
- o step: Find o that solves $\frac{1}{n}(\sum_{i=1} w_i r_i + o_i - y_i log(w_i r_i + o_i)) + \gamma||o||_2^2$

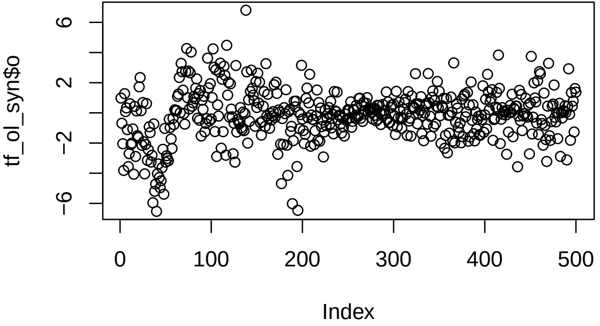
Synthetic dataset

- Here, $\rho = 5e-2$, and $\gamma = 1e-4$ are chosen randomly, CV for this with outlier term is under construction.

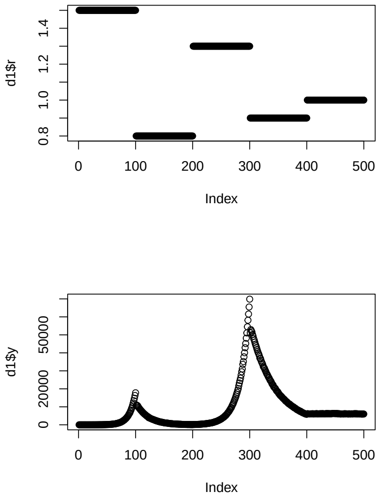


Outliers

- Modeling outliers not very significant

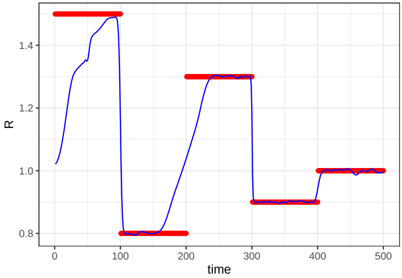


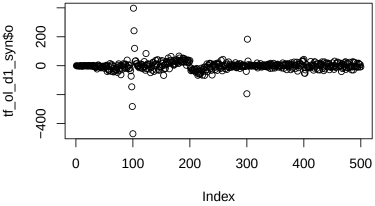
Synthetic dataset, sudden changes



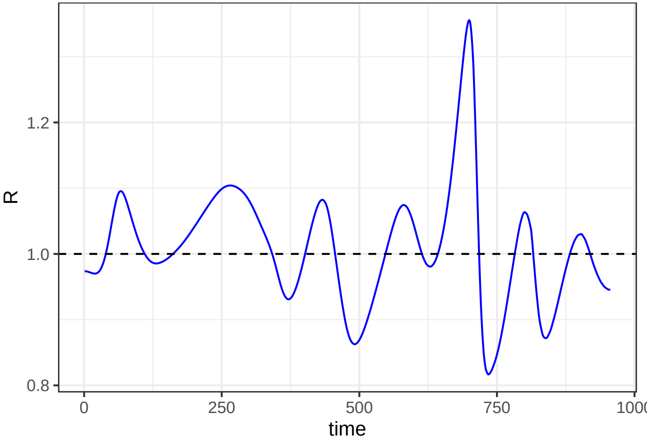
Fit and outlier

- Red is truth, blue is predicted

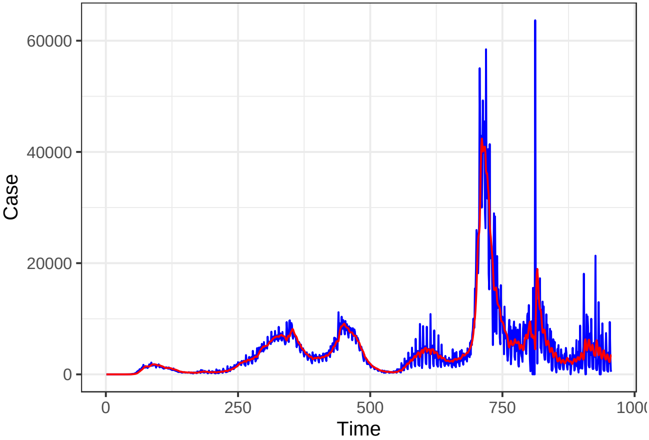




Result

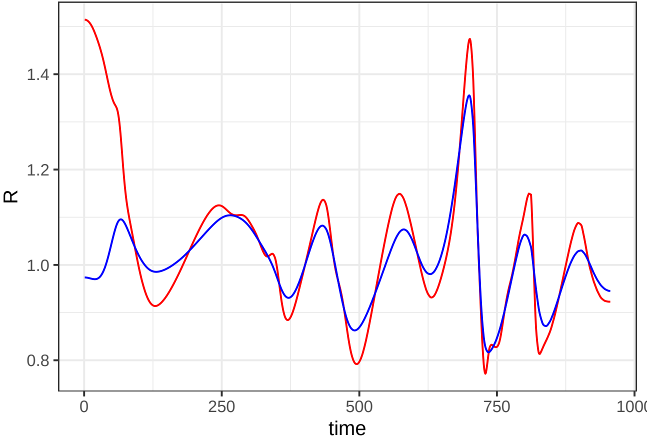


One day



Compare

- Then compare two methods. Blue, with outlier term modeled



Next step

- OOP of trend filter class