Power Forecasting and Control discussion notes

Waleed Alhaddad

- Check the how reasonable are the obtained parameters through the quadratic variation

$$\sum_{i,j}(X_{i,j} - X_{i,j})^2 \approx 2\theta_t \alpha \sum_{i,j} X_{i,j}(1 - X_{i,j}) \tag{1}$$

- Apply the model to the french data
  - We are waiting for the details of the evolution of the maximum installed capacity to be able to normalize.
  - Exploit the pyramid structure of the forecast updates, as the french recompute the forecast every hour.
  - We are also waiting to know if the forecasting technology changed during those years and when. This will enable us to pick upon the particularities of the different forecasting technologies. Also, to check if the newest forecasting technology performs better in practice.
- Understand, apply and extend the "R" package DiffusionRgqd to our case of multiple paths and see how it performs in comparison.
- Plot the likelihood as two "profile likelihood" plots as common in the statistics literature. This means fix one of the parameters, optimize in the other and plot. For instance, it can be expressed in terms varying σ as

$$R(\sigma) = \frac{\max_\mu \mathcal{L}(\mu, \sigma)}{\mathcal{L}(\hat{\mu}, \hat{\sigma})} \tag{2}$$

  And similarly, we plot $R(\mu)$.
- The issue with the power curtailing or human intervention is still open. The ideas so far:

- set a threshold to θ_t to prevent the process from following such sudden changes. However, this involves choosing some value of θ_t and a threshold. Also, the data is not refined enough to identify such changes from natural wind speed changes (as currently its recorded at intervals of one hour)
- Introduce a jumping process, adds complexity to the model but might enhance it overall. This will enable us to capture the sudden changes.
- Regime switching, similar to the first point, but in this case we can assume that θ_t follows some Markov Chain instead of hard setting thresholds.

- Double check that the Feller condition is satisfied to ensure the positivity of the process.

- Try to see if we can make the system ergodic, especially when we have higher frequency data. This is done when $\Delta n \to \infty$ where Δ is the interval between each sample of n total samples.

  Then we can leverage existing theoretical results regarding the asymptotic behavior of our model.

- Possibility of adding regularizers on θ_t that way we can manipulate it and differentiate it easily for further theoretical work.

- Assign an a distribution on the initial point and optimized it separately. By that I mean, optimize it without involving the likelihood and its optimization.

- Once we have higher frequency data, we can use Ahmad's MLE estimators to find the parameters of the model in a simpler and faster way. Regarding this point, I have his papers from 2012 and 2013 which are in the shared Dropbox folder.

- Another motivation of the project is that power production companies are not allowed to dump more power into the grid than authorized. Otherwise, they have to pay hefty fines as they endanger the grid by using it above its capacity.

- It may also be possible to obtain the aggregated power production of France for both controllable and non-controllable energy sources. This will benefit us in two ways, we can identify human intervention more easily and it can be used as data and a test case for Renzo's optimal

control ongoing project. Currently this data set is sparse and has some negative values we couldn't understand. (see Dropbox shared folder under "Realized_ Production")