Post-Doctoral Position in Statistical Learning.

Online aggregation of Kalman filters.

Olivier Wintenberger
Professeur Université Pierre et Marie Curie
olivier.wintenberger@upmc.fr

Yannig Goude
Chef de projet Energy Data Lab EDF R&D
yannig.goude@edf.fr

Scientific project

Forecasting non-stationnary processes is a fundamental question for electricity management and trading on energy markets. Electriocty consumption processes, electricity prices, renewable production could have time varying behavior at a different time scale. More precisely

- In the field of load forecasting, to model a time varying portfolio of customers with arrival/departures occuring with time. There is thus a need to develop robust and adaptive statistical methods, respecting some physical or statistical cunstraints. Different methods already exist at EDF R&D for that, among them GAM or adaptive GAM perfomed well in operation (see e.g. Gaillard et al. (2015, 2016), Thouvenot et al. (2016)) but there is still some work concerning identifiability issues, estimation of the learning rate (forgetting factor) and auto-regression terms estimation.
- Modelling a small number of individual customers is an important concern for load scheduling.
- Forecasting time series for energy markets, e.g. the net imbalance volume for load balancing.

The prediction problem will be tackled by using an aggregation of multiples state-space models. State space models are very broad predictive models widely used in many application fields, see Prado & West (2010). One can think of state-space modeling for

- State variables that are lagged observations, lincluding any ARMA model,
- Functions of explicative variables such as GAM with hidden states,
- More general dynamical models with time dependent coefficients…

The Kalman algorithm is a recursive procedure with a cubic cost in the number of variables. It provides the best linear prediction given the model.

The aim of the project is to combine the predictions from multiple state-space models in an optimal way and to apply it to prediction on the electricity market.

Optimal aggregation is usually obtained thanks to Exponential Weights Algorithms; see Cesa-Bianchi, N., Lugosi, G. (2006). Recent works prove that it is necessary to include an indicator of the risk of prediction in the aggregation procedure, see Wintenberger (2016). The Kalman recursion provides such indicators, namely the conditional variances of the best linear prediction through time. These indicators will be included in the exponential weights in order to achieve an optimal aggregation, adapting the pioneer work of Leung et Barron (2006) to conditional quadratic risk.

Position description

The candidate should start as soon as October 2017. The position will be held at UPMC and will long for 1 year with a possible extension of 1 year. The salary is approx. 2200 euros net per month, depending on the experience of the candidate. The position is partly supported by EDF DataLab and the candidate will have to implement and apply the algorithm in R or Python on electricity data.

Bibliography

Cesa-Bianchi, N., Lugosi, G. (2006). Prediction, learning, and games. Cambridge university press.

Leung, G., & Barron, A. R. (2006). Information theory and mixing least-squares regressions. IEEE Transactions on Information Theory, 52(8), 3396-3410.

Gaillard, P. & Goude, Y. Antoniadis, A.; Poggi, J.-M. & Brossat, X. Forecasting Electricity Consumption by Aggregating Experts; How to Design a Good Set of Experts Modeling and Stochastic Learning for Forecasting in High Dimensions, Springer International Publishing, 2015, 217, 95-115

Gaillard, P.; Goude, Y. & Nedellec, R. Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting International Journal of Forecasting, Elsevier, 2016, 32, 1038-1050

Prado, R., West, M. (2010). Time series: modeling, computation, and inference. CRC Press.

Thouvenot, V.; Pichavant, A.; Goude, Y.; Antoniadis, A. & Poggi, J. M. Electricity Forecasting Using Multi-Stage Estimators of Nonlinear Additive Models IEEE Transactions on Power Systems, 2016, 31, 3665-3673

Wintenberger, O. (2016) Optimal learning with Bernstein Online Aggregation, Machine Learning, Online first.