COMP6212 Assignment 4 : Shakib-Bin Hamid 25250094 sh3g12

1 Kalman Filter

Lets suppose there a data generative process is a linear of its past values, i.e. the current process process parameters are a linear combination of previous process parameters. In this respect we can calculate the process parameters by minimising the least squared error between the prediction and the true values, $\boldsymbol{\omega} = (\boldsymbol{Y}^t\boldsymbol{Y})^{-1}\boldsymbol{Y}^t\boldsymbol{f}$, where $\boldsymbol{Y}$ is the input data and $\boldsymbol{f}$ is the target. But this method requires calculating a new $\boldsymbol{Y}^{-1}$ (a $O(n^3)$ operation) when we receive any new data. If the process is generating a time series data (getting data one at a time), then online updating the parameters can be incredibly costly.

However, it can be shown that we can calcuate new parameters without calculating a new inverse, $\boldsymbol{\omega}_n = \boldsymbol{\omega}_{n-1} + L \times (f_n - \boldsymbol{\omega}_{n-1}^t y_n)$. It means that we can calculate new parameters by updating the old parameters with the scaled prediction error on new data using the old parameters. Of course we need to assume some initial values for our parameters. But, we still do not account for any noise in our process.

In a Kalman Filter, we may assume a process $\boldsymbol{\omega}_n = \boldsymbol{A}\boldsymbol{\omega}_{n-1} + \boldsymbol{\epsilon}(n)$, where the process noise $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{Q})$. Here we can assume that our parameters are a linear transformation of previous parameters corrupted by Gaussian noise. Similarly, we can assume that our observation $f_n = \boldsymbol{\omega}_n \boldsymbol{y}_n + \xi(n)$, where observation noise $\xi \sim \mathcal{N}(0, R)$.

[1] claims that we can predict the monthly S&P 500 index using a Kalman filter. In order to verify his claim, I have built Algorithm 1 -

Algorithm 1 S&P 500 Index Prediction using a Kalman Filter

```
procedure KALMAN($\boldsymbol{s}, o, \alpha, R$)
                ▷ $\boldsymbol{s}$ = index, $o$ = order, $\alpha$ = is a constant, R = observation noise variance
    $N = size(\boldsymbol{s}, 1)$;                                  ▷ Number of months
    $\boldsymbol{Y} = zeros(N - o + 1, o)$;                        ▷ Input data in $o$ order windows
    $\boldsymbol{W} = zeros(N - o + 1, o)$;                        ▷ Weights at every point
    $\boldsymbol{e} = zeros(N - o + 1, 1)$;                        ▷ Prediction error at every point
    $\boldsymbol{K} = zeros(N - o + 1, o)$;                        ▷ Kalman gain at every point

    $\boldsymbol{A} = \boldsymbol{I}_o$                            ▷ $\boldsymbol{I}_o$ is the $o$ order Identity
    $\boldsymbol{Q} = \alpha\boldsymbol{I}_o$                      ▷ Process noise covariance

    $\boldsymbol{P} = ones(o)$                                     ▷ Inital value of $\boldsymbol{P}$
    $\boldsymbol{W}(1, :) = o^{-1}ones(1, o)$

    for $i \leftarrow 1, N - o$ do
        $\boldsymbol{Y}(n, :) = \boldsymbol{s}(i : i + o - 1)$     ▷ Input is last $o$ months' index

        $\hat{\boldsymbol{w}} = \boldsymbol{A}\boldsymbol{W}(n - 1, :)^t$     ▷ Prediction
        $\hat{\boldsymbol{P}} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^t + \boldsymbol{Q}$

        $\boldsymbol{e}(n) = \boldsymbol{s}(n + o - 1)$ - $\boldsymbol{Y}(n, :)\hat{\boldsymbol{w}}$     ▷ Prediction Error
        $\boldsymbol{K}(n, :) = (\hat{\boldsymbol{P}}\boldsymbol{Y}(n, :)^t)/(R + \boldsymbol{Y}(n, :)\hat{\boldsymbol{P}}\boldsymbol{Y}(n, :)^t)$     ▷ Kalman Gain

        $\boldsymbol{W}(n, :) = \boldsymbol{W}(n - 1, :) + \boldsymbol{K}(n, :)\boldsymbol{e}(n)$     ▷ Correction
        $\boldsymbol{P} = (\boldsymbol{I}_o - \boldsymbol{K}(n, :)^t\boldsymbol{Y}(n, :))\boldsymbol{P}$
```

Here the input at each point in time is past $o \in 2, ..., N-1$ index values. Observation noise R is taken as the residual variance from an autoregression model using `ar` function of Matlab's System Identification Toolbox. Process noise covariance needed to be tuned via α. From Figure 3a (where I have plotted the cumulative sum of prediction error as a function of α for a 3rd order filter), we can say that 10^{-3} is a reasonable value for α.

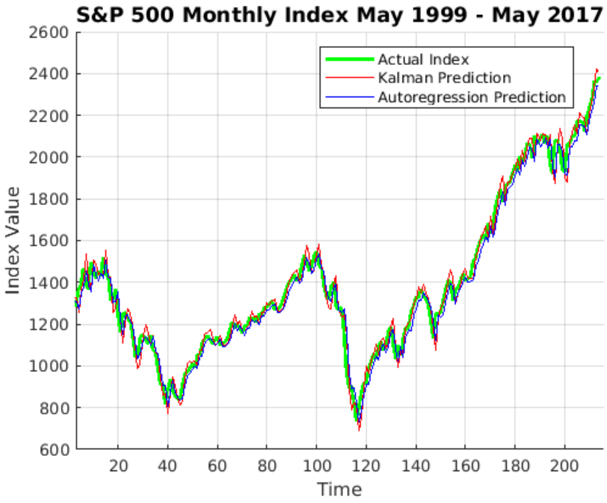
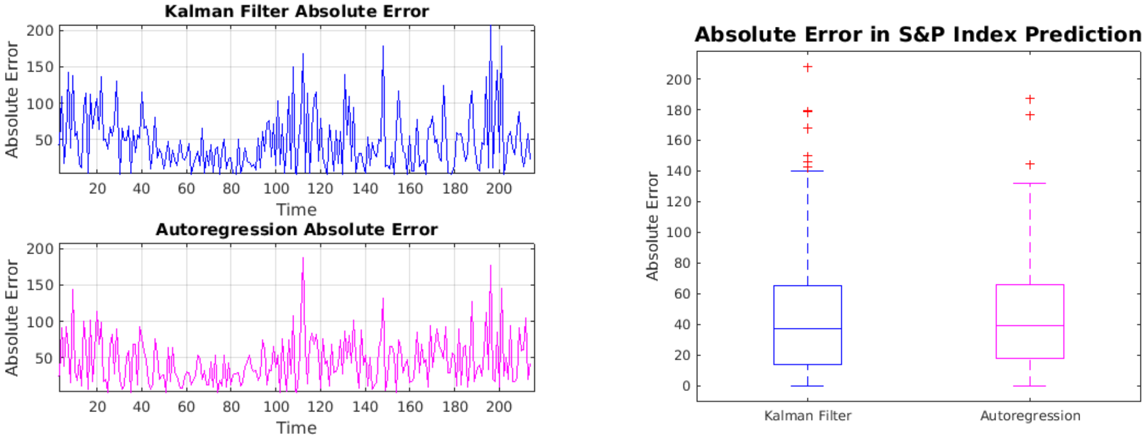


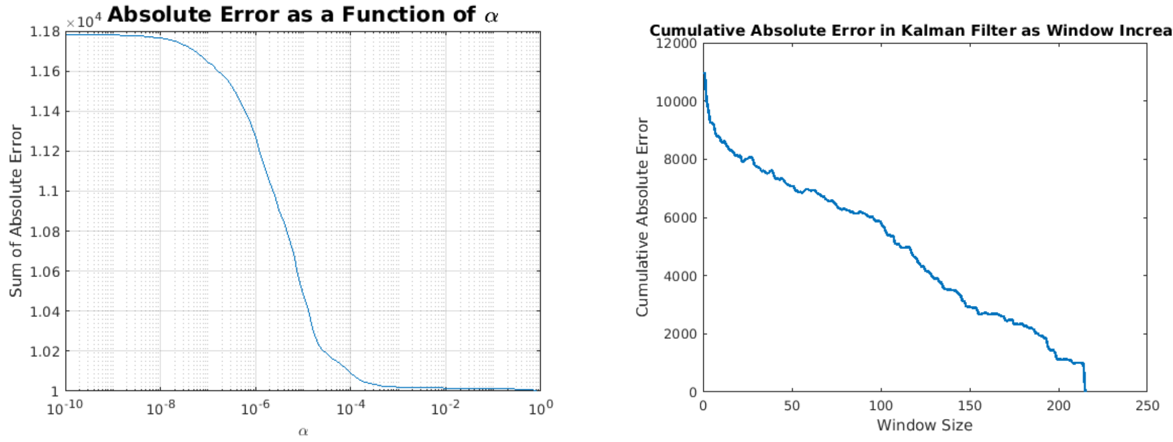
Figure 1: Kalman & Autoregression Prediction



(a) Index Prediction Absolute Error

(b) Index Prediction Error Comparison

Figure 2: Prediction Error



(a) Cumulative Absolute Error as a Function of α

(b) Index Prediction Error Comparison

Figure 3: Cumulative Absolute Error as a function of o and α

References

[1] N. Mahler, "Modeling the S & P 500 index using the Kalman filter and the LagLasso," in *Machine Learning for Signal Processing, 2009. MLSP 2009. IEEE International Workshop on, Sept 2009*, pp. 16.