Kalman Filter

Carl Richardson

28th February 2021

Contents

1 Introduction

The Kalman Filter (KF) is a method used to estimate the Gaussian probability density function (pdf) of an unobserved state in the setting of a linear system with Gaussian noise. The approach incorporates knowledge of the system in the form of a State Space (SS) model. As the state of a dynamic system changes with time, the estimated pdf of the state changes too. During the interval between the previous and next observation, the KF updates it's estimate of the pdf in two stages. First, it uses the knowledge of the system dynamics and previous observations to predict the mean and covariance of the pdf; then once the observation arrives, the predictions are corrected such that the expected error between the true state and it's posterior estimate is minimised. This report investigates the estimation performance of the KF when applied to a second order Auto Regressive (AR) process, shown by equation 1. The observed output of the system is corrupted by some measurement noise and the previous observations are known. The goal of this investigation is to estimate the unknown parameters of the model used to generate the data. To formulate the problem as a SS model, the parameters of the AR process were treated as the state variable. The parameters of the process were assumed to only vary slightly due to process noise; hence the decision to model the state transitions as a random walk. The output equation of the SS model represents the AR process. This formulation of the SS model is shown by equation 2. The performance of the KF, using this SS model, was investigated when the AR process had constant and time varying parameters to illustrate the difference in performance.

$$x(n) = \sum_{k=1}^{2} \theta_k(n)x(n-k) + v(n) \text{ where } v(n) \sim \mathcal{N}(0,\, R) \tag{1}$$

$$\begin{gathered} \theta_{\mathbf{n}} = \theta_{\mathbf{n-1}} + \mathbf{w_n} \text{ where } \mathbf{w}(n) \sim \mathcal{N}(\mathbf{0},\, \mathbf{Q}) \\ x(n) = \mathbf{x_n^T}\theta_{\mathbf{n}} + v(n) \text{ where } v(n) \sim \mathcal{N}(0,\, R) \end{gathered} \tag{2}$$

2 AR Process with time invariant parameters

In this section, the true parameters of the AR process were set as 0.3 and 0.5 respectively; however, when conducting these experiments, they were treated as unknown. Unless otherwise stated, the default values used in these experiments were as follows: measurement variance of the AR process was 0.1 and this was assumed to be known, the process covariance was $0.01I$ and the initial value of theta was $(-1, 4)$. Simulations were conducted over 400 iterations. Figure 1 shows the default generated data.

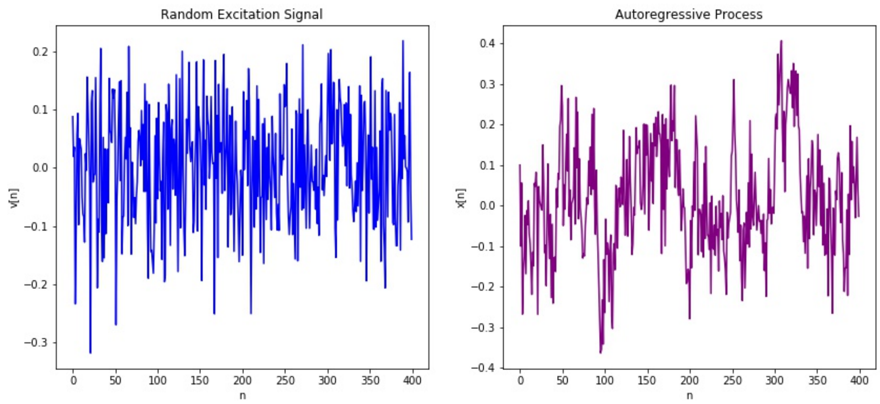

Figure 1: Generated AR process when R = 0.1

The KF was implemented using the equations shown in the lab document. To initialise the algorithm, random estimates of the parameters had to be made. Figure 2 (left) illustrates the evolution of the MSE between the true and estimated parameters given different initial conditions. The slowest convergence time took 141 iterations when a tolerance of 0.1 was set. The fastest convergence time was 67 iterations, this occurred when the distance between the true and randomly initialised estimates was shortest. Clearly, if any information about the parameters is available to make an informed choice of initial conditions, the speed of convergence could be greatly reduced. Figure 2 (right) also depicts the evolution of the MSE when the variance of the measurement noise was changed. In each case, the measurement noise was assumed to be known. The results show that within 50 iterations, the KF was able to transition from the initial estimate to approximate the parameters well in all cases. After 50 iterations, the measurement noise changed over time significantly more as the variance increased. When it was 0.6, the maximum MSE was 0.44 which translates to a true error of 0.66; whereas when it was 0.2, the maximum MSE was 0.089 which translates to a true error of 0.30. The acceptable tolerance will depend on the application.

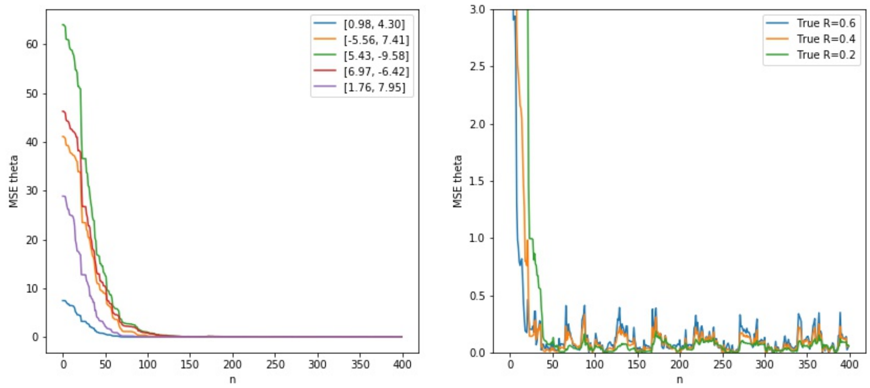

Figure 2: Evolution of the MSE given different initial conditions (left). Evolution of the MSE when the variance of the measurement noise was varied (right)

As mentioned above, the KF is able to estimate the Gaussian pdf of the parameters at each instant in time. The form of the pdf is known to be Gaussian, so it must estimate its mean and covariance. When computing the MSE, the mean of the pdf was used as the estimate of the parameters. Figure 3 shows how the estimate of the mean and variance of each parameter changes over time. As expected, the parameter values move towards their true values, but their associated variances fluctuate between 0.3 and 0.5. This was also to be expected as, by definition, the covariance of the pdf is the covariance of the error between the true and estimated parameters.

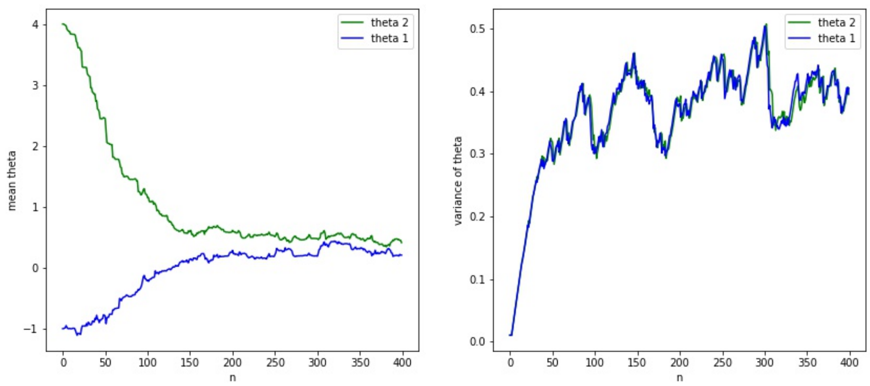

Figure 3: Evolution of the parameter estimates and their associated variance using default values

In plenty of applications, the measurement variance will not be known. To simulate this, the value of the measurement variance used by the AR process was kept as default whilst the estimated value used by the KF was modified as a hyper parameter. Figure 4 (left) shows how the MSE of the parameters evolved over time for different estimates of the measurement variance. As the deviation from the true measurement variance increased, so did the time taken for the estimates to converge. However, in each case, when the tolerance was set to 0.1, convergence occurred. Another hyper parameter to be tuned was the process covariance matrix. In each case, the covariate terms were set to zero, as the parameters were independent. Figure 4 (right) shows the evolution of the MSE for different values of the variance terms of this matrix. Different distributions of the same variance produced very similar results. When the total variance was 0.6, the convergence time took between 28 and 30 iterations. This suggests the convergence time was independent of the distribution of variance. For the case where the total variance was 0.1, this took 52 iterations to converge. Interestingly, this suggests that increasing the process noise variance leads to faster convergence.

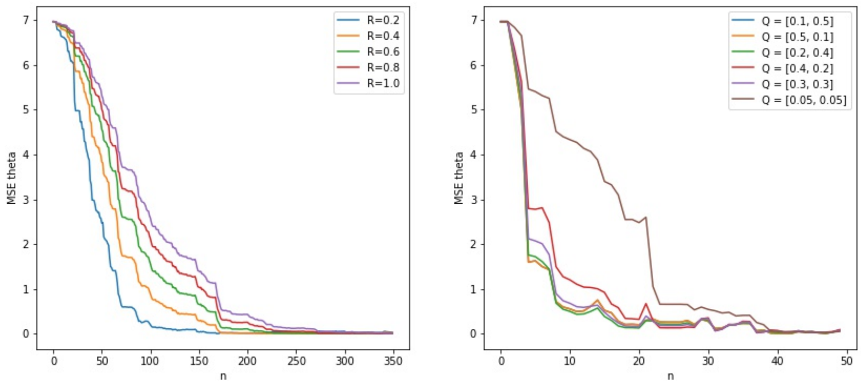

Figure 4: Tuning the hyper parameters R (left) and Q (right)

3 AR Process with time varying parameters

In this section, the true parameters of the AR process changed over time, as shown by figure 5. Unless otherwise stated, the default values used in these experiments were as follows: measurement variance of the AR process was 0.1 and this was assumed to be known, the process covariance was $0.01I$ and the initial value of theta was $(-1, 4)$. Simulations were conducted over 400 iterations. The default generated data is included in the appendix.

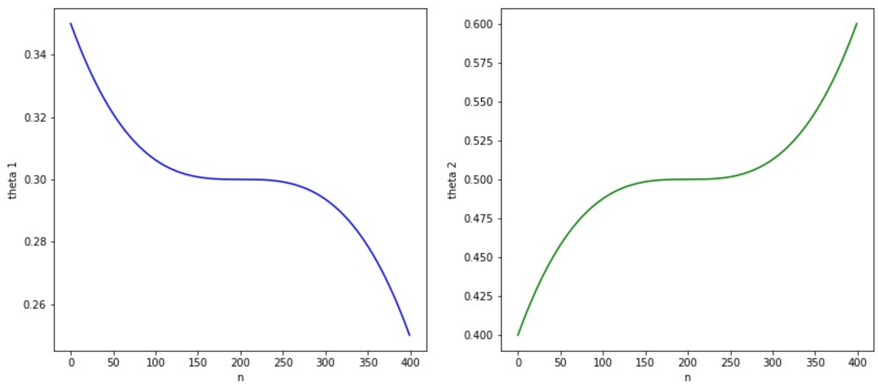

Figure 5: Time varying parameters

The effect the time varying (TV) parameters had on the speed of convergence was investigated by computing the time evolution of the MSE, for the same set of initial conditions as in the time invariant (TIV) experiments. Again, when a threshold of 0.1 was used, the MSE converged under all initial conditions, as shown by figure 6. Like before, the fastest convergence time occurred when the distance between the initialised parameters and the true parameters was shortest; however, in this case, this just meant at time zero as the parameters were TV. In this case, the fastest convergence time was 99 iterations and the slowest was 251 iterations. Compared to when the parameters were TIV, this was an increase of 1.48 and 1.78 respectively. This highlights that the KF found it more challenging to estimate the distribution of parameters, but was still able to do so.

Figure 7 shows how the estimate of the Gaussian pdf parameters changed over time. On the left, the figure shows how the mean of the TV and TIV cases varied. As can be seen, the trajectories were nearly identical. This may be caused by the mean of the TV case being the same as the TIV case. The right part of the figure shows how the covariance terms changed. Again, the trajectory of the variance terms were very similar to the TIV case. This suggests, the KF was equally confident in its estimate of the TV parameters as it was the TIV parameters. Furthermore, since the parameters were correlated, the cross covariance term was included. A small negative covariance was expected as a small move towards the mean in theta 1, occurred at the same time as a move of twice the distance occurred in theta 2. Since the deviation of theta 1 and theta 2, from their respective means, always had opposite polarity, this was expected.

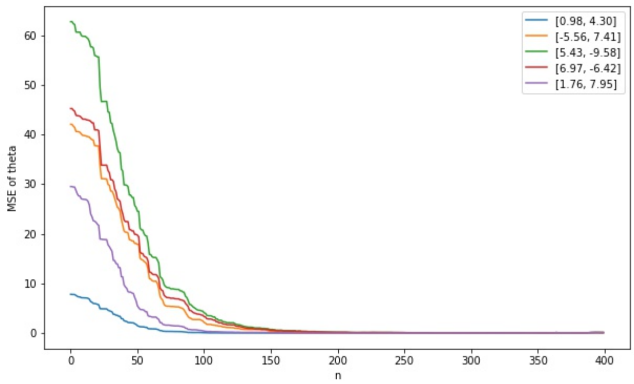

Figure 6: Evolution of the MSE given different initial conditions

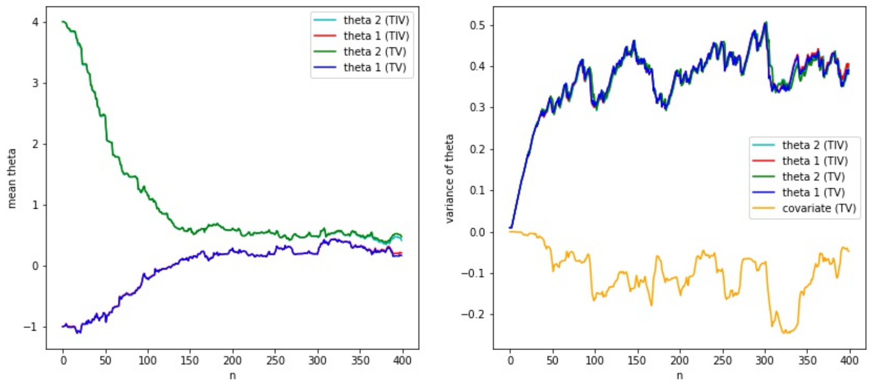

Figure 7: Evolution of the time invariant (TIV) and time varying (TV) parameter estimates and their associated variance using default values

Finally, the effect of tuning the hyper parameters of the SS model was considered for the TV case. On the left of figure 8, the estimate of the true measurement variance was treated as unknown. As the estimate moved further from the true value, the convergence time for the MSE increased. The only case when convergence occurred within 400 iterations was when $R = 0.2$, this increased from 122 iterations to 355 iterations. In the TIV case, the convergence time of the 5 experiments had a range of 147 iterations, whereas in the TV case, the same 5 experiments had a range of 559. This suggests that tuning the estimate of the measurement variance was particularly important in the TV case. On the right of figure 8, the process covariance was tuned. When the total variance was 0.6, the convergence time ranged between 22 and 30 iterations; when it was 0.1, the convergence time was 52 iterations. In both cases, this was very similar to the results when the parameters were TIV. In fact, when the variance was 0.1, it was identical. Like before, these results indicate that increasing the process variance leads to faster convergence.

4 Conclusion

The KF was capable of estimating the unobserved parameters of the second order AR process with a maximum variance of 0.5. Initialising the algorithm with parameter estimations closer to the true values improved the convergence time, as did increasing the process variance. It would be worth investigating the effect of increasing the process variance further to see if its effect hits a maximum before increasing convergence time. In addition, investigating how covariance terms, which reflect the relationship between the parameters, effects convergence time would be interesting. Finally, poor estimates of the measurement variance noticeably increased the convergence time, in particular, in the case of TV parameters.

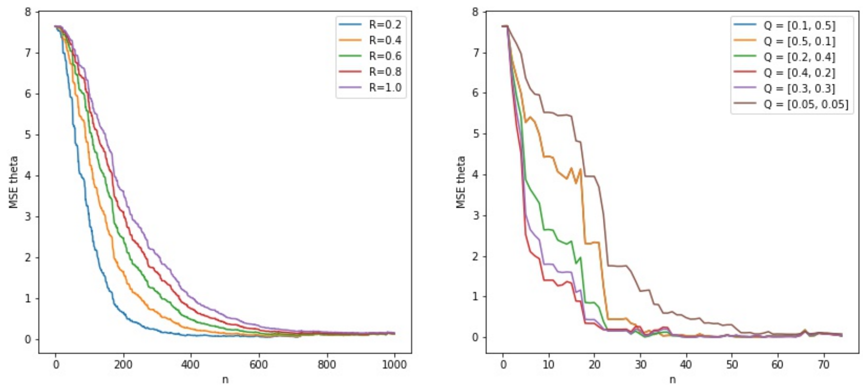

Figure 8: Tuning the hyper parameters R (left) and Q (right)

5 Appendix

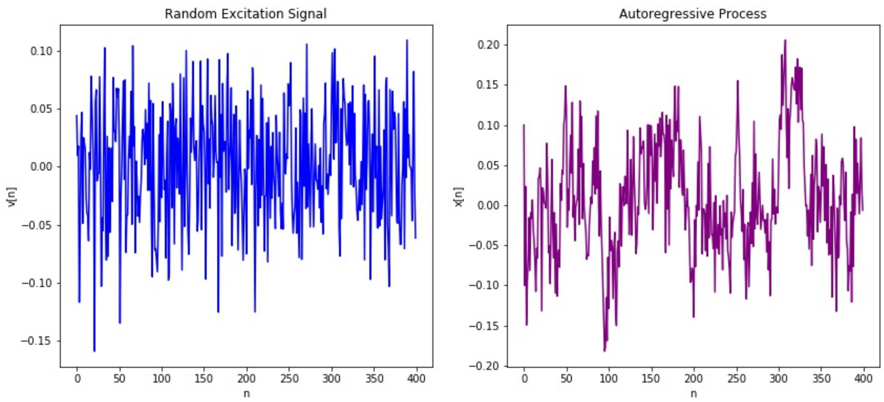

Figure 9: Time varying AR process