02427 Advanced Time Series Analysis E18

Computer exercise 2

Niklas Christoffer Petersen

2018-11-07

1 Part 1

Estimate the parameters in at least one of the systems generating the time series that you simulated in part 1 of computer exercise 1. Use the prediction error method (also refered to as the conditional least squares method) for the estimation.

Our model of choice it the SETAR$(2, 1, 1)$ model presented in exercise 1:

$$y_t = \begin{cases} 0.2 + y_{t-1} + \epsilon_t & \text{if } y_{t-1} < 100 \\ 10 + 0.95 y_{t-1} + \epsilon_t & \text{if } y_{t-1} \geq 100 \end{cases}$$

We load the data generated for the SETAR$(2, 1, 1)$ model in computer exercise 1, and for convenience we plot the time series as shown in Figure 1.

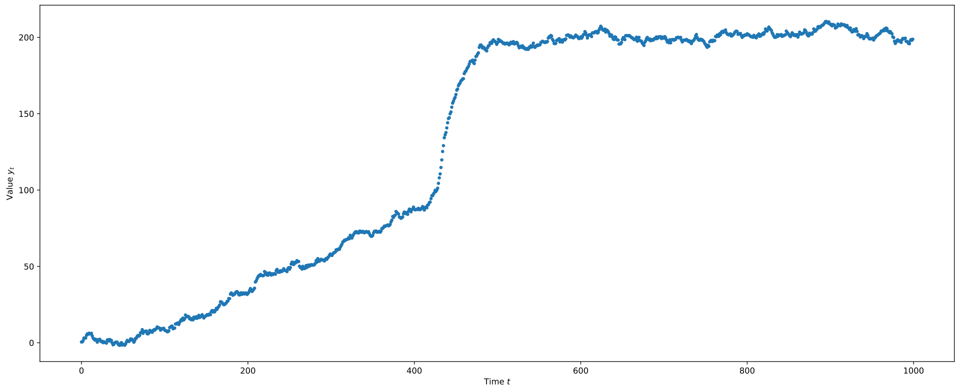


Figure 1: SETAR$(2, 1, 1)$ model from exercise 1.

We start by define our SETAR$(2, 1, 1)$ model such that the slope coefficients of each regime are paramized using θ_0 and θ_1.

$$\hat{y}_t(\Theta = [\theta_0, \theta_1, \ldots]) = \begin{cases} 0.2 + \theta_0 y_{t-1} + \epsilon_t & \text{if } y_{t-1} < 100 \\ 10 + \theta_1 y_{t-1} + \epsilon_t & \text{if } y_{t-1} \geq 100 \end{cases}$$

We define our residual function taking all the parameters of the above SETAR function, as well as the true y values, where $\Theta = [\theta_0, \theta_1]$

$$RSSSetar : \hat{y}_t(\Theta) - y_t$$

We use the `least_squares` function from the `scipy.optimize` package (Python) for our optimization. Our initial guess is that both values are 0:

$$\Theta_{init} = [0, 0]$$

The output from the optimization is shown in Figure 2.

```
Iteration     Total nfev        Cost      Cost reduction    Step norm     Optimality
   0              1         1.1629e+07                                    2.96e+04
   1              2         7.5497e+06      4.08e+06       1.00e+00       3.74e+08
   2              4         4.5226e+06      3.03e+06       4.15e-01       8.62e+08
   3              5         2.9896e+06      1.53e+06       4.15e-01       4.87e+08
   4              7         2.9320e+06      5.76e+04       4.63e-02       8.98e+07
   5             11         2.9290e+06      2.99e+03       1.81e-04       8.75e+07
   6             12         2.9274e+06      1.58e+03       4.52e-05       8.76e+07
   7             14         2.9272e+06      1.93e+02       1.13e-05       8.75e+07
   8             15         2.9264e+06      8.39e+02       2.83e-06       8.74e+07
   9             16         2.9256e+06      7.73e+02       5.65e-06       8.71e+07
  10             20         2.9256e+06      1.77e+01       1.77e-07       8.71e+07
  11             21         2.9255e+06      3.54e+01       3.53e-07       8.71e+07
  12             25         2.9255e+06      1.11e+00       1.10e-08       8.71e+07
‘xtol‘ termination condition is satisfied.
Function evaluations 25, initial cost 1.1629e+07, final cost 2.9255e+06, first-order
    optimality 8.71e+07.
```

Figure 2: Output from optimization.

We see the least squares optimization algorithm quickly finds the optimal values:

$$\theta_0 = 1.0074464779279126, \theta_1 = 0.875866248321334$$

We see, that they are close to the true values (1, .95). Better estimate could probably be achieved if lowering the termination tolerence level for the optimizer, but for know we are satisfied with the estimates.

Now also wanted (naïvly) to try estimate the regime change parameter:

$$\hat{y'}_t(\Theta = [\theta_0, \theta_1, \ldots]) = \begin{cases} 0.2 + y_{t-1} + \epsilon_t & \text{if } y_{t-1} < \theta_0 \\ 10 + .95 y_{t-1} + \epsilon_t & \text{if } y_{t-1} \geq \theta_0 \end{cases}$$

Here we are only trying to estimate a single parameter, θ_0. The output from the optimization is shown in Figure 3, and the optimizer terminates immediately, since the gradient is below the threshold.

```
Iteration     Total nfev        Cost      Cost reduction    Step norm     Optimality
   0              1         5.1584e+06                                    0.00e+00
‘gtol‘ termination condition is satisfied.
Function evaluations 1, initial cost 5.1584e+06, final cost 5.1584e+06, first-order
    optimality 0.00e+00.
```

Figure 3: Output from optimization of regime change.

2 Part 2

This part has been difficult, since the mean squared residuals very easily explode due to the exponential nature of the model and chosen parameters. So the limits for the grid has be set very close the the true values:

$\theta_0 : [.99, 1]$ (True 1)

$\theta_1 : [.90, .95]$ (True .95)

If we include all values (N = 1000) we will get the contour plot shown in Figure 4. Here we see the optimal value of the parameters will result in MSE close to zero, while just moving a small step (especially for $theta_0$) will get the MSE to increase rapidly.

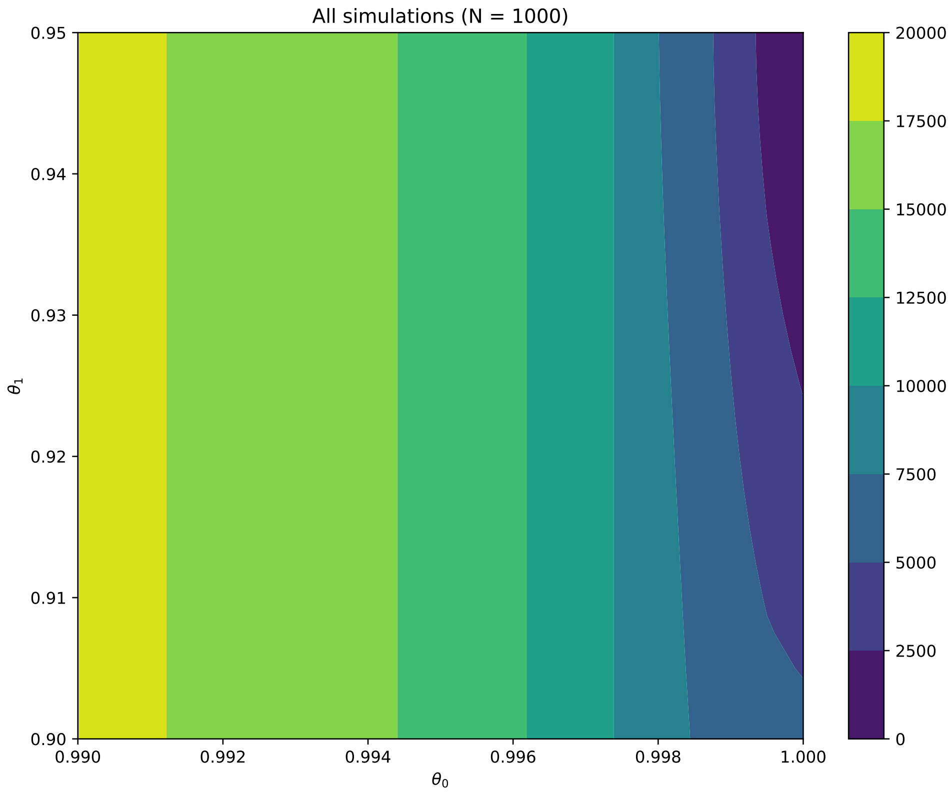


Figure 4: All simulations (N = 1000)

If we only include the first 100 values (N = 100) we will get the contour plot shown in Figure 5. Likewise if we include the last 100 values (N = 100) we will get the contour plot shown in Figure 6. Here it almost seem like $theta_1$ has no influence. I was honestly expecting more differences between the first and the last simulations, but again, this is presumably the result of a poorly chosen model for this experiment.

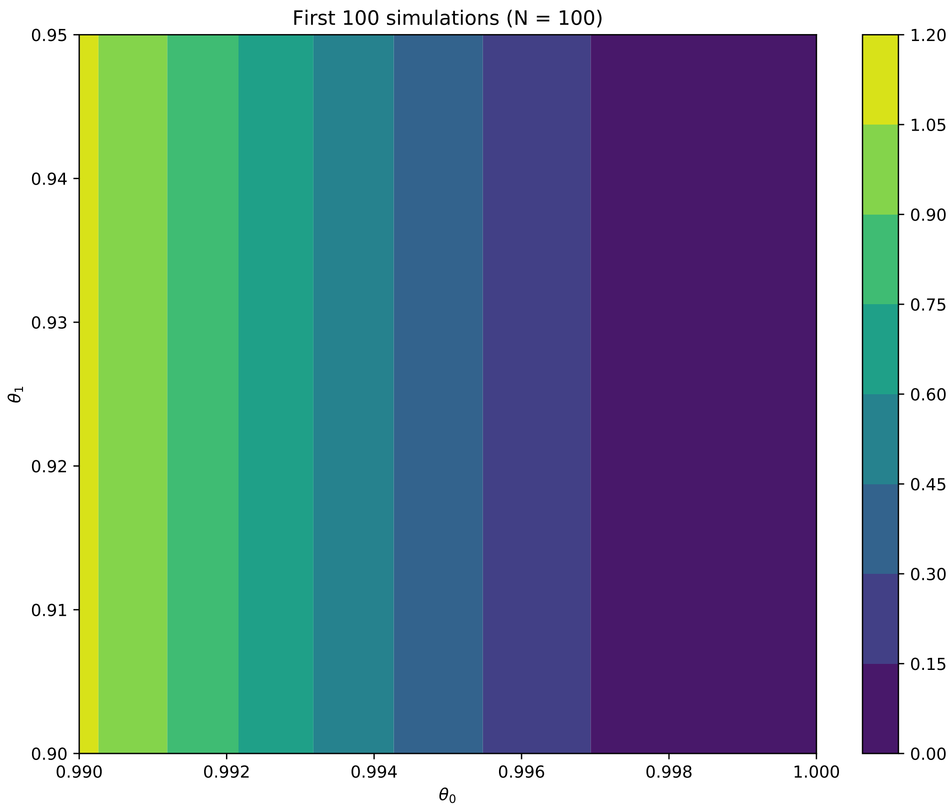


Figure 5: First 100 simulations (N = 100)

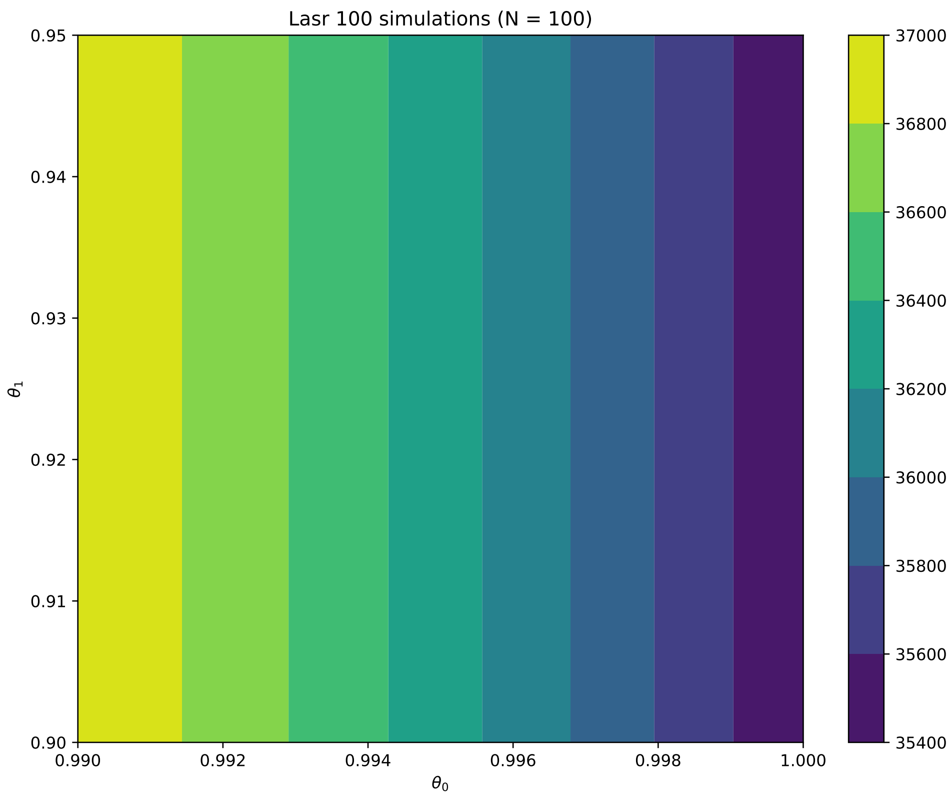


Figure 6: Last 100 simulations (N = 100)

3 Part 3

We have chosen the following model, where the observation process is a simple MA(1)-process, and the hidden/latent process is as ARMA(1, 1)-process:

$$y_t = x_t \epsilon_{t-1} + \epsilon_t$$
$$x_t = \phi x_{t-1} + \theta \zeta_{t-1} + \zeta_t$$

We write to state-space form:

$$X_t = \begin{bmatrix} \phi & \theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \zeta_{t-1} \end{bmatrix} + \begin{bmatrix} \zeta_t \\ 0 \end{bmatrix}$$
$$Y_t = \begin{bmatrix} \epsilon_{t-1} & 0 \end{bmatrix} X_t + \epsilon_t$$

For simulation we chose the following parameters and simulate 100 iterations. The simulated time series is shown in Figure 7.

$$\phi = .95$$
$$\theta = .80$$
$$\sigma_\zeta^2 = .4^2$$
$$\sigma_\epsilon^2 = 2^2$$

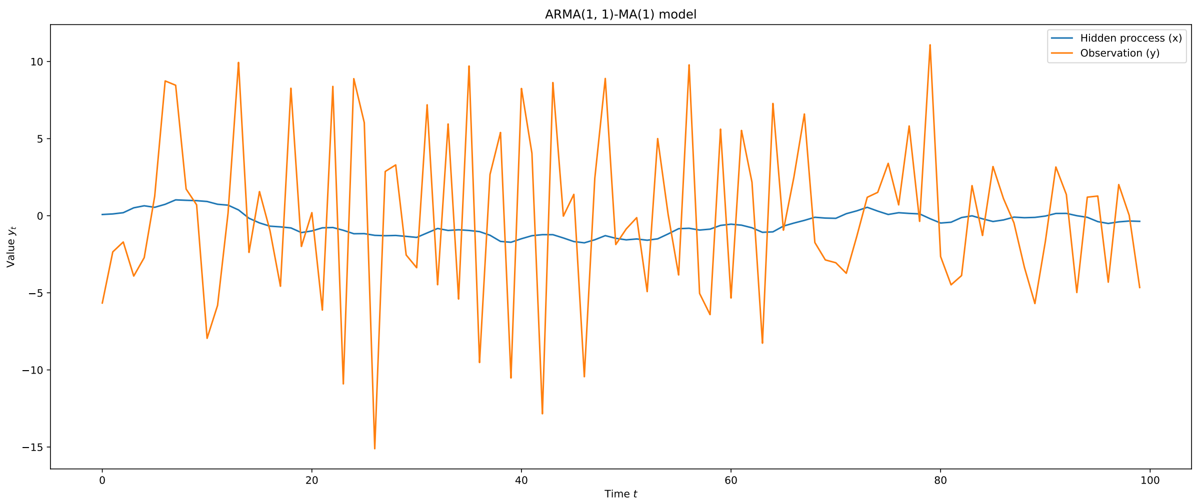


Figure 7: Simulations (N = 100)

4 Part 4

We consider the following simple state space model:

$$x_{t+1} = ax_t + v_t$$
$$y_t = x_t + e_t$$

We simulated 20 runs of each 1000 time steps as shown in Figure 8 (not that informative). We will later see, that 1000 time steps is enough for the EKF to converge.

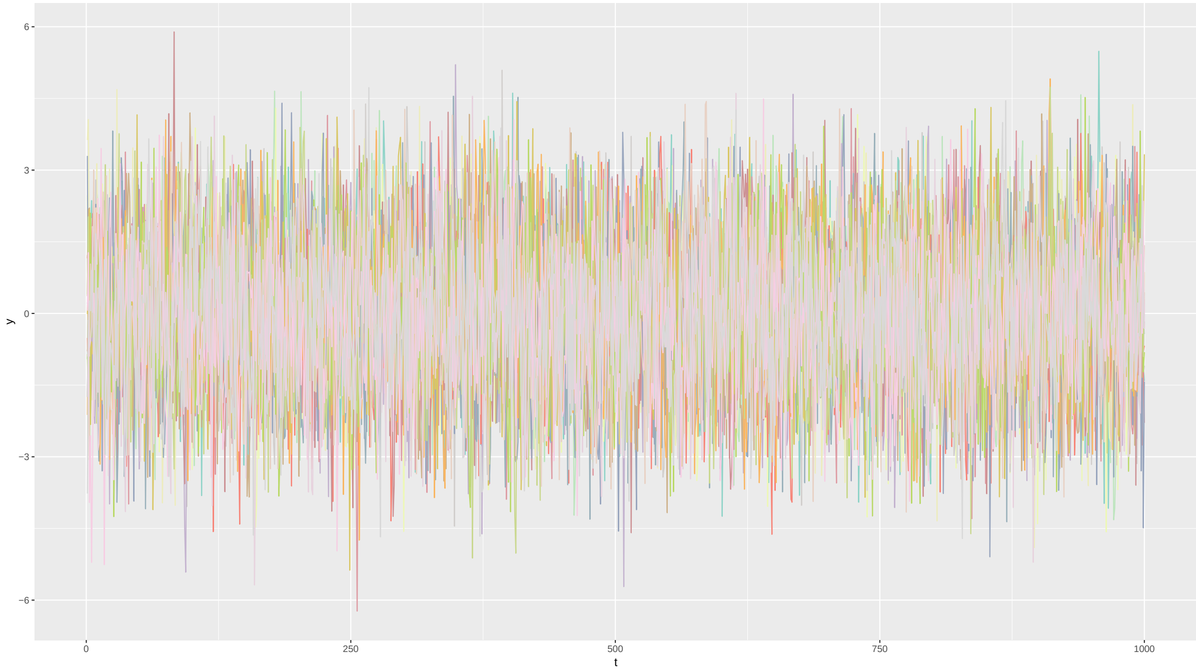


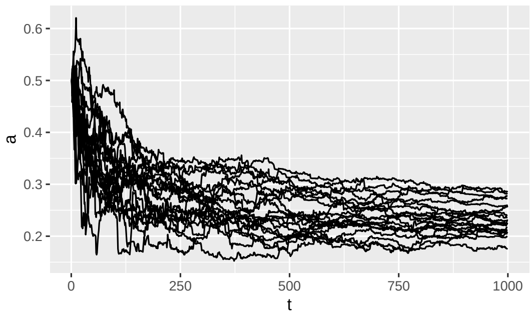
Figure 8: 20 Simulations (N = 20 x 1000)

We formulate a corresponding state space form, including the parameter a which we want to estimate:

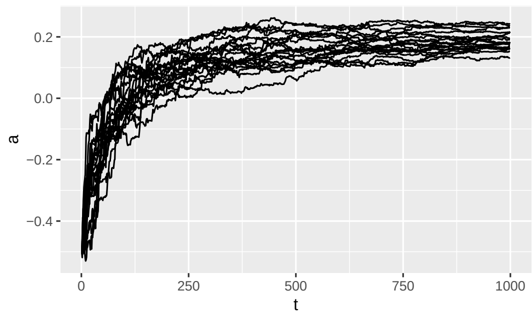
$$X_{t+1} = \begin{bmatrix} x_{t+1} \\ a_{t+1} \end{bmatrix} = \begin{bmatrix} a_t x_t \\ a_t \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \end{bmatrix}$$
$$Y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} X_t + e_t$$

Figures 9 to 12 shows the convergence for the 8 different configurations. In all of the cases we see that the EFK converge, but at different rates. However for the first configuration ($\sigma_v^2 = 10$, $\sigma_a^2 = 1$), and the third ($\sigma_v^2 = 10$, $\sigma_a^2 = 10$) it converges towards around .20 instead of the true value $a = .4$.

Table 1 summarizes the results of the last simulation step of each of the 8 different configurations, including the point and interval (95%) estimates for a. It does not seem important whether the initial guess is 0.5 or -0.5, or if $\sigma_a^2 = 1$ or $\sigma_a^2 = 10$. What seems important is whether the variance of the white noise term v is large or not. The value of σ_a^2 is only influencing the convergence in the early iterations of EKF, but not the final point or interval estimates.

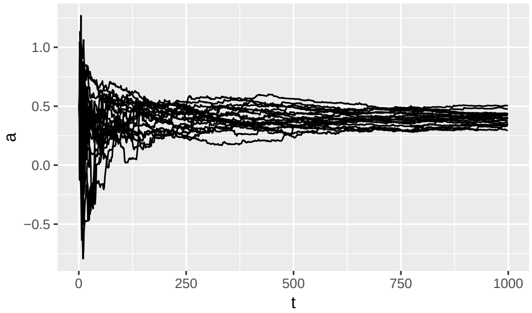


(a) $a_{init} = .5$

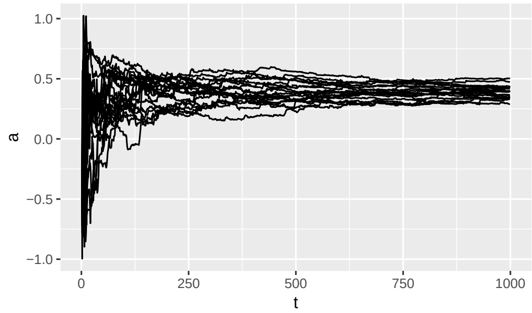


(b) $a_{init} = -.5$

Figure 9: Convergance of a for $\sigma_v^2 = 10$ and $\sigma_a^2 = 1$

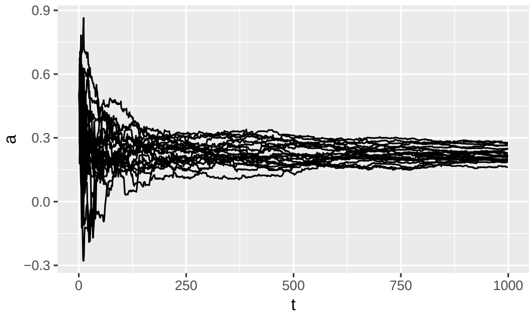


(a) $a_{init} = .5$

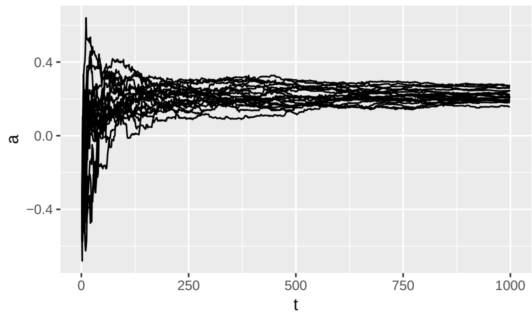


(b) $a_{init} = -.5$

Figure 10: Convergance of a for $\sigma_v^2 = 1$ and $\sigma_a^2 = 1$

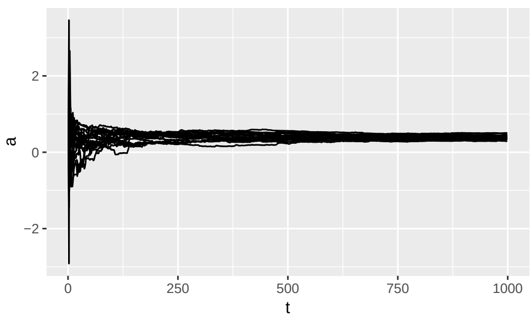


(a) $a_{init} = .5$

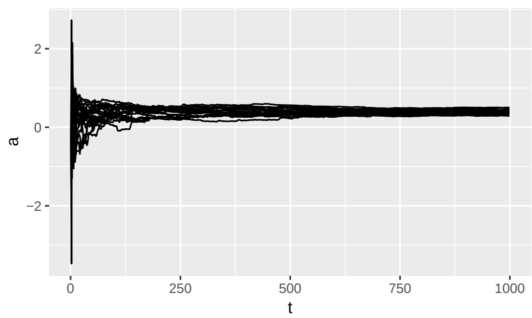


(b) $a_{init} = -.5$

Figure 11: Convergance of a for $\sigma_v^2 = 10$ and $\sigma_a^2 = 10$



(a) $a_{init} = .5$



(b) $a_{init} = -.5$

Figure 12: Convergance of a for $\sigma_v^2 = 1$ and $\sigma_a^2 = 10$

	a_{init}	σ_v^2	σ_a^2	$\hat{a}$	95% Conf.
1a	0.5	10	1	0.23	0.19–0.28
1b	-0.5	10	1	0.19	0.14–0.24
2a	0.5	1	1	0.40	0.31–0.49
2b	-0.5	1	1	0.39	0.31–0.49
3a	0.5	10	10	0.22	0.17–0.27
3b	-0.5	10	10	0.22	0.17–0.27
4a	0.5	1	10	0.39	0.31–0.49
4b	0.5	1	10	0.39	0.31–0.49

Table 1: Summary of EKF results.