Consider a whole list of N observations y1,y2,...,yN and compute the mean value of the posterior you would get by combining all these observations with the prior. This can be done in one of two ways.

- By forming joint likelihood from all N observations as a batch, and combine with the prior to get a joint posterior.
- Start cranking through recursively by taking y1 and combining with the prior to a get a posterior, then taking that posterior as a new prior for a new step, to be combined with observation y2, to get a new posterior, and so on. It will only take two steps or so to discover the recursion relation for the mean and variance of the posterior after having seen N > 1 observations.

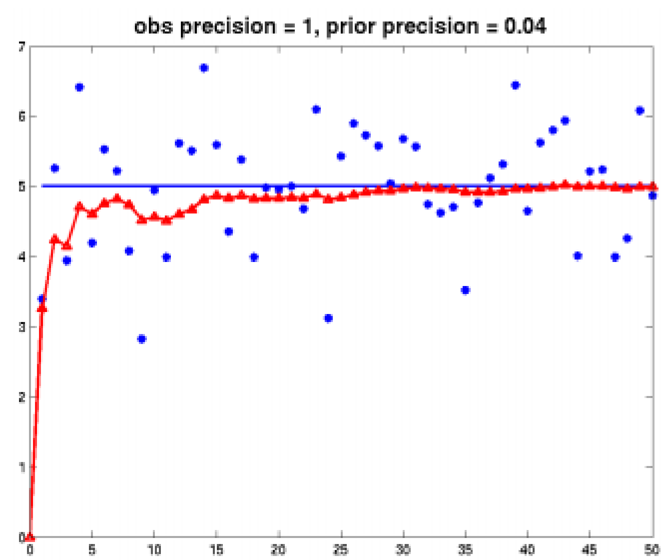
  A 1-D Normal distribution with mu standard deviation, x mean and p = $1/\sigma^2$ (precision is 1/variance) can be written as:

$$N(x|\mu, p) = c \exp\left\{\frac{1}{2}\left[p\,(x-\mu)^2\right]\right\}$$

This code implements a recursive Bayes filter as in the above example, but by really doing it a single step at a time, using the posterior computed as a result of seeing observation yt as a prior for the next time step, where you will be seeing yt+1. That is, it constantly keeps around a current mean t and current variance t (or precision t) that keep getting updated as new observation yt+1 comes in.

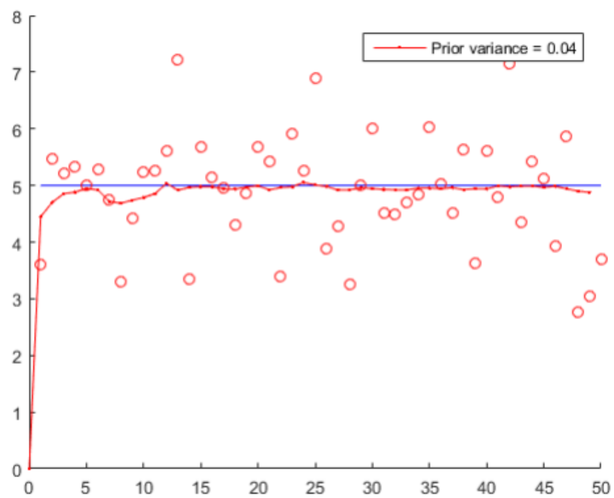
It assumes that each yt is a noisy observation of a constant state value x, and for simplicity, that these observations are independent, and have the same variance. Using the model p(x) = N(μ,σ 2) ; prior distribution yt = x+ε ; observation at time t p(ε) = N(0,s 2) ; observation noise.

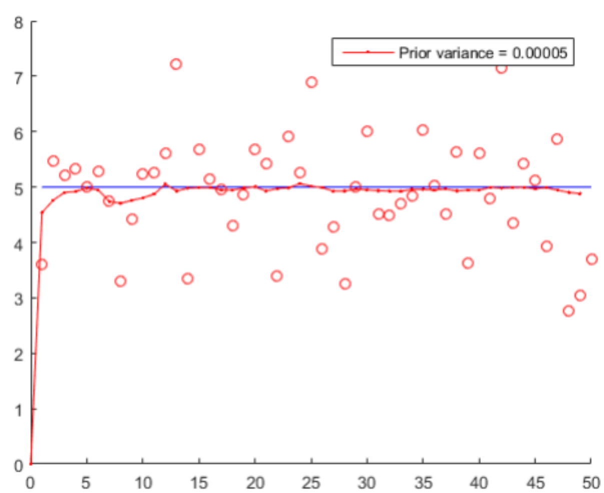
The results are plotted in a graph showing time as the x axis and the observations and estimated state values (mean of posterior distribution at time t) on the y axis, as in the figure below

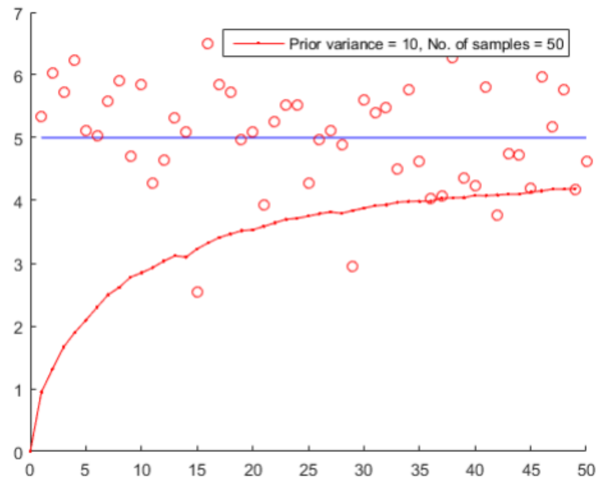


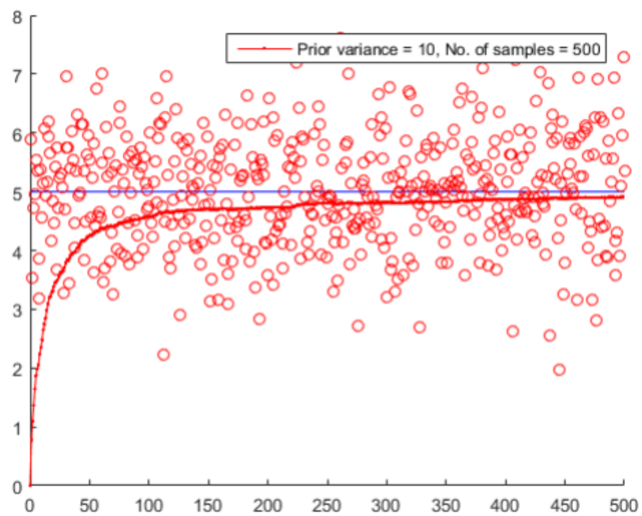
In this plot, the blue line is the true state value, blue points are the noisy observations, and red line is the estimated state value over time. This was produced for a true state value of 5, a prior distribution with mean 0 and precision 0.04 (variance 25), observation noise with mean 0 and precision 1 (variance 1), and running for 50 time steps (so there were 50 observations).

The following plots were obtained by playing around with the variance of the prior estimate. What happens if you set it low (that is, you mistakenly think you have a good prior)? (First two plots are for N = 50).






Prior variance = 10, No. of samples = 50
0
1
2
3
4
5
6
7
0 5 10 15 20 25 30 35 40 45 50


Prior variance = 10, No. of samples = 500
0
1
2
3
4
5
6
7
8
0 50 100 150 200 250 300 350 400 450 500