CH5115 Term Project

Name: Abhiram S
Roll no.: CH19B037

Problem 1

$S = 1000$ samples for each parameter were generated from their respective priors.

The predictor equation used is given below:

$$\hat{y}[k] = a_1 y[k-1] + b_1 u[k-1] + b_2 u[k-2]$$

Euclidean distance between $\hat{\mathbf{y}}$ and $\mathbf{y}$ (i.e. 2-norm) was chosen as the distance function for ABC rejection and tolerance on this distance was chosen as $\epsilon = 10$.

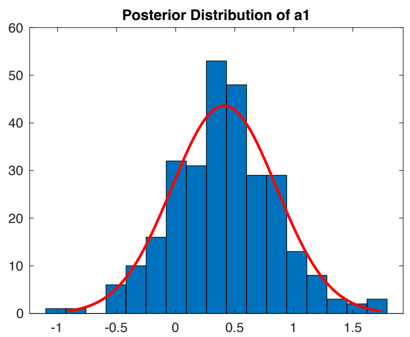
From 1000 sample points in the prior, 285 were accepted into the posterior, yielding an acceptance rate of 28.5%. This is arguably quite high, but such a tolerance was chosen so that the posteriors and residuals are distributed almost normally. This is so that the formula for Bhattacharyya distance between two normal distributions can be used without inducing much error. The results asked in parts (1), (2) and (3) are given below:

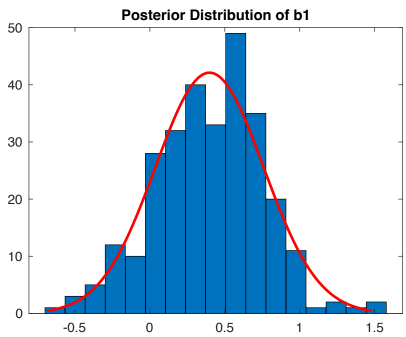
PROBLEM 1 RESULTS:

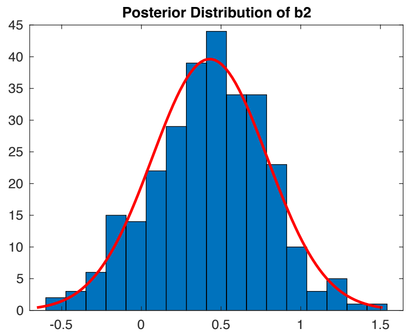
	Reduction in Uncertainty	**Point Estimate**	**Information Gain**
a1	1.7347	0.41027	8.5597e-05
b1	34.6	0.39773	0.018911
b2	56.454	0.42881	0.05033

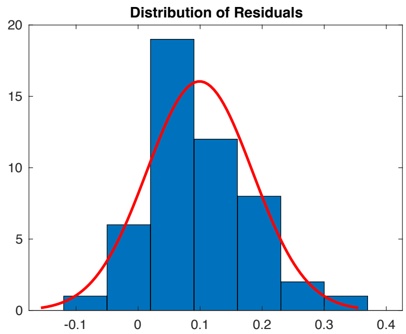
We see that a_1 has the least reduction in uncertainty and hence, has the least information gain. b_2 has the largest reduction in uncertainty and thus has the highest information gain. This is consistent and makes sense.

From the plots below, we see that the posteriors and residuals are approximately Gaussian, which further suggests that the fit is good.









Problem 2

The first order state space model is given to be:

$$x[k+1] = [-a_1]x[k] + [b_1 - a_1 b_0]u[k]$$
$$y[k] = [1]x[k] + [b_0]u[k] + e[k]$$

The corresponding predictor equations are:

$$\hat{x}[k+1] = [-a_1]\hat{x}[k] + [b_1 - a_1 b_0]u[k]$$
$$\hat{y}[k] = [1]\hat{x}[k] + [b_0]u[k]$$

The second order state model is given to be:

$$\boldsymbol{x}[k+1] = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix} \boldsymbol{x}[k] + \begin{bmatrix} b_1 - a_1 b_0 \\ b_2 - a_2 b_0 \end{bmatrix} u[k]$$
$$y[k] = [1 \quad 0]\,\boldsymbol{x}[k] + b_0 u[k] + e[k]$$

The corresponding predictor equations are:

$$\hat{\boldsymbol{x}}[k+1] = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix} \hat{\boldsymbol{x}}[k] + \begin{bmatrix} b_1 - a_1 b_0 \\ b_2 - a_2 b_0 \end{bmatrix} u[k]$$
$$\hat{y}[k] = [1 \quad 0]\,\hat{\boldsymbol{x}}[k] + [b_0]u[k]$$

$S = 5000$ samples for each parameter were generated from their respective priors. Euclidean distance between $\hat{\boldsymbol{y}}$ and $\boldsymbol{y}$ (i.e. 2-norm) was chosen as the distance function for ABC rejection for both models and tolerance on this distance was chosen as $\epsilon_1 = 45$ for the first order model and $\epsilon_2 = 60$ for the second order model. This was done because for the same ϵ, the posterior for the second order model had too few elements. Thus, the tolerance was increased for the second order model so that both posteriors have sufficient number of elements. From 5000 sample points in the prior, 120 were accepted into the posterior for the first order model and 66 were accepted into the posterior for the second order model, yielding acceptance rates of 2.4% and 1.3% respectively. This is small enough, and the results required by parts (1) and (2) are shown below:

```
PROBLEM 2 RESULTS:
Estimated First-Order State Space Model:
-0.1184 | 0.7515
 -------+-------
 1.0000 | 0.0126

Estimated Second-Order State Space Model:
-0.0970 1.0000 |  0.7384
-0.2209 0.0000 |  0.1941
 --------------+--------
 1.0000 0.0000 | -0.0120

Improvement in Fit:
    MSE (1st Order)    MSE (2nd Order)    % Reduction in MSE
    _______________    _______________    __________________

        1.1496             1.1362               1.168

Loss in Information:
          Info Gain (1st Order)    Info Gain (2nd Order)    % Loss in Info
          _____________________    _____________________    ______________

    a1           0.35383                  0.11177               68.413
    b0           0.39785                  0.24198               39.179
    b1           0.38924                  0.18521               52.419
```

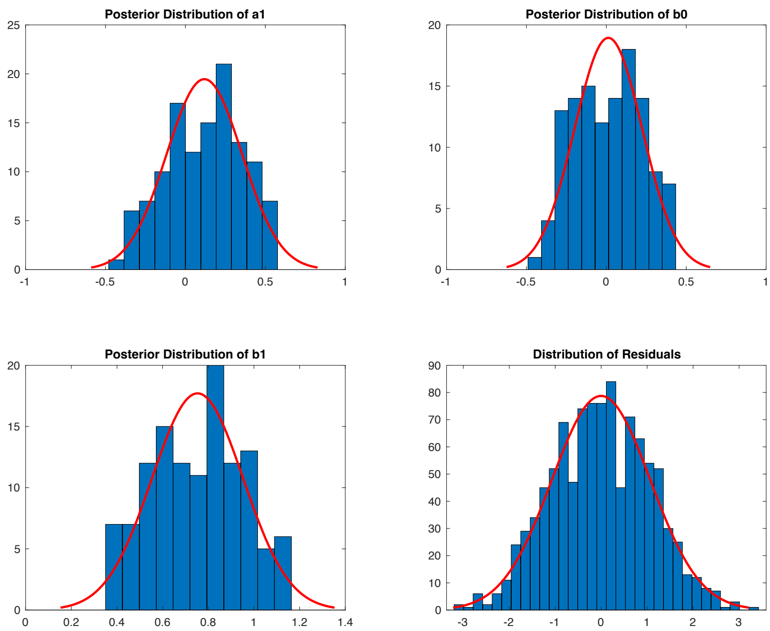
Note that the state space models are displayed in the following format, with lines separating the matrices:

$$\begin{bmatrix} \mathbf{A} & | & \mathbf{B} \\ - & - & - \\ \mathbf{C} & | & \mathbf{D} \end{bmatrix}$$

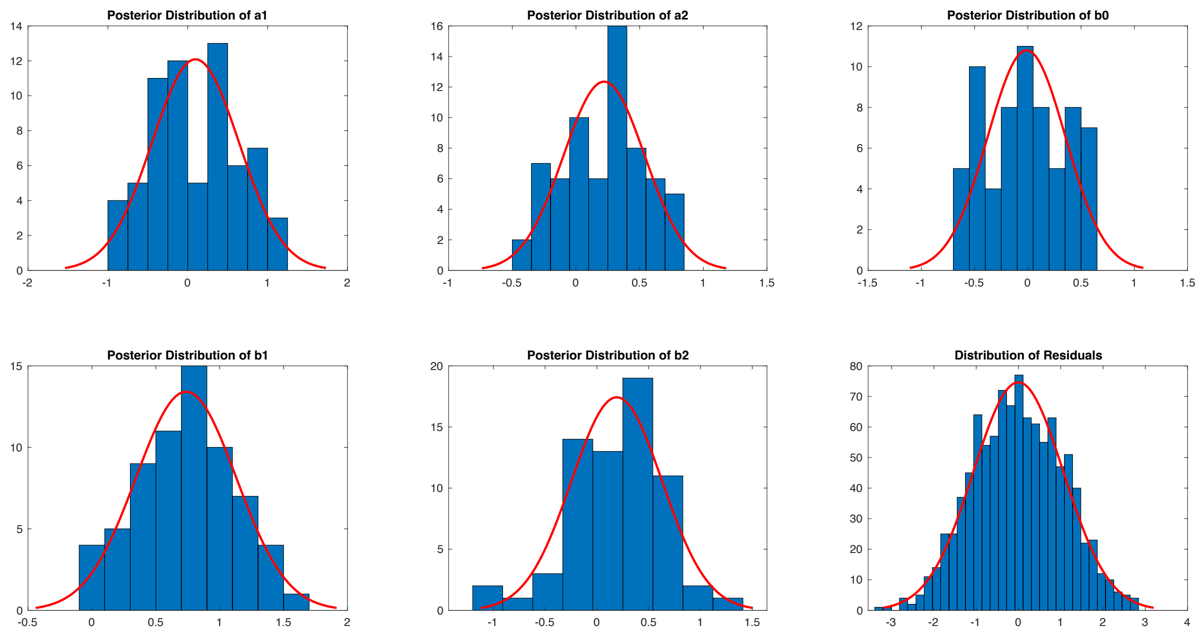
We see that there is a mere 1% reduction in mean square error while going from a first order fit to a second order fit. At the same time, we lose a lot of information in the common parameters (68%, 39% and 52% respectively for a_1, b_0 and b_1). The loss in information is not worth the marginally better fit. Thus, we can say that the first order fit is a better choice overall.

The posterior distributions and residual distributions for both models are shown below:

First Order Model



Second Order Model



We see that most of the posteriors are fairly Gaussian. The residuals are more Gaussian than the posteriors, mainly because the residuals are much larger in number.

Based on these observations, we can propose a method for order estimation, as asked in part (3). We fit models of increasing order and evaluate the reduction in error and the information loss for each order. At every step, we compare the improvement in fit with the information loss. If the fit increases drastically while the information loss is comparatively small, the tradeoff is worth it. We continue doing this. Past the optimal order, we will notice that there is very little improvement in fit while the information losses are huge. This is an indication for us to stop. The point where we encounter these diminishing returns is the best estimate of order.