Non-Linear Assignment

Name-Rahul Singh

Roll - 171111

Introduction

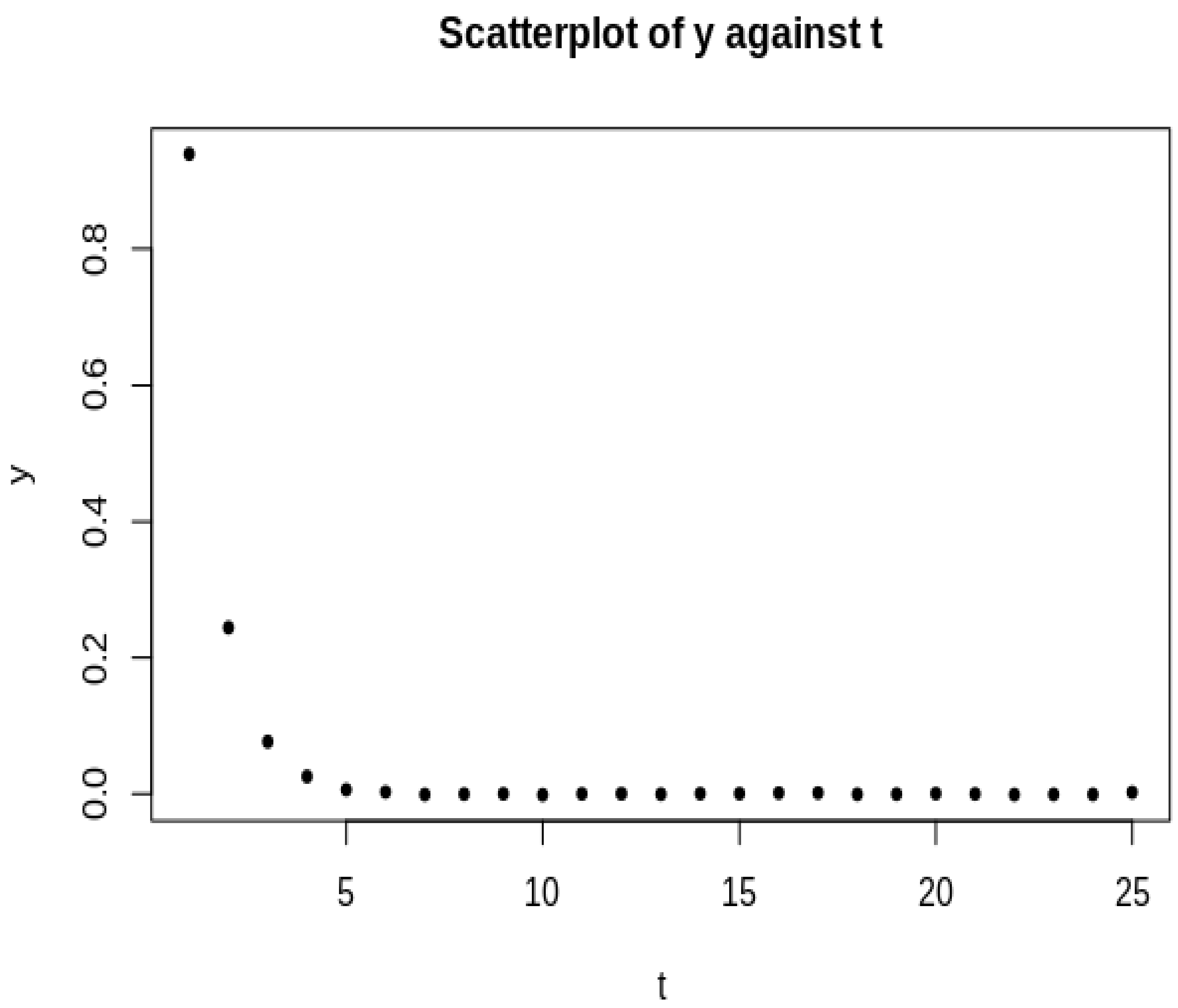
We were given 25 observations on the variable y. We are supposed to fit the model :

$y_t = \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + \varepsilon_t$

Here ε_t is a sequence of i.i.d. random variable with mean zero and finite variance.

Part-(a)

We plot the data against t and obtain the following plot:



Part-(b)

We are to use the Prony's method to check whether the rots that we obtain are real or not.

We define a matrix $\boldsymbol{Z} = \begin{bmatrix} y_1 & y_2 & \cdots & y_{k+1} \\ y_2 & & & y_{k+2} \\ \vdots & & & \vdots \\ y_{n-k} & y_{n-k+1} & \cdots & y_n \end{bmatrix}$

Here,k=2 and n=25.

After that we find the eigen vector which has the smallest eigen value of $V=Z^TZ$, that is, we select $\hat{a}$ as $V^* \hat{a} \approx 0$

Here, we obtain $\hat{a} = (0.0171,0.3601,0.9328)^T$

Then we solve the equation:

$0.0171+0.3601x+0.9328x^2=0$ ....#1

The solution of this equation are the estimates of e^{β_1} and e^{β_2} .

The solutions of the equation #1 are : 0.055 and 0.331

Thus, $\hat{\beta}_1=-2.894$ and $\hat{\beta}_2=-1.107$

To obtain the estimates of other parameters we define the matrix:

$$\boldsymbol{X} = \begin{bmatrix} e^{\beta_1} & e^{\beta_2} \\ \vdots & \vdots \\ e^{n\beta_1} & e^{n\beta_2} \end{bmatrix}$$

$\hat{\alpha} = (X^TX)^{-1}XY$

The estimates are obtained as:

$\hat{\alpha}_1=4.340$ and $\hat{\alpha}_2=2.114$

Therefore, we can obtain real estimates of the parameters using Prony's Method.

Part-(c)

After that we plot the contours.We vary both β_1 and β_2 from -3 to 0.1 at intervals of 0.1 and find the Partial Least Square estimate of α_1 and α_2 using the equation:

$\hat{\alpha}$ =$(X^TX)^{-1}XY$

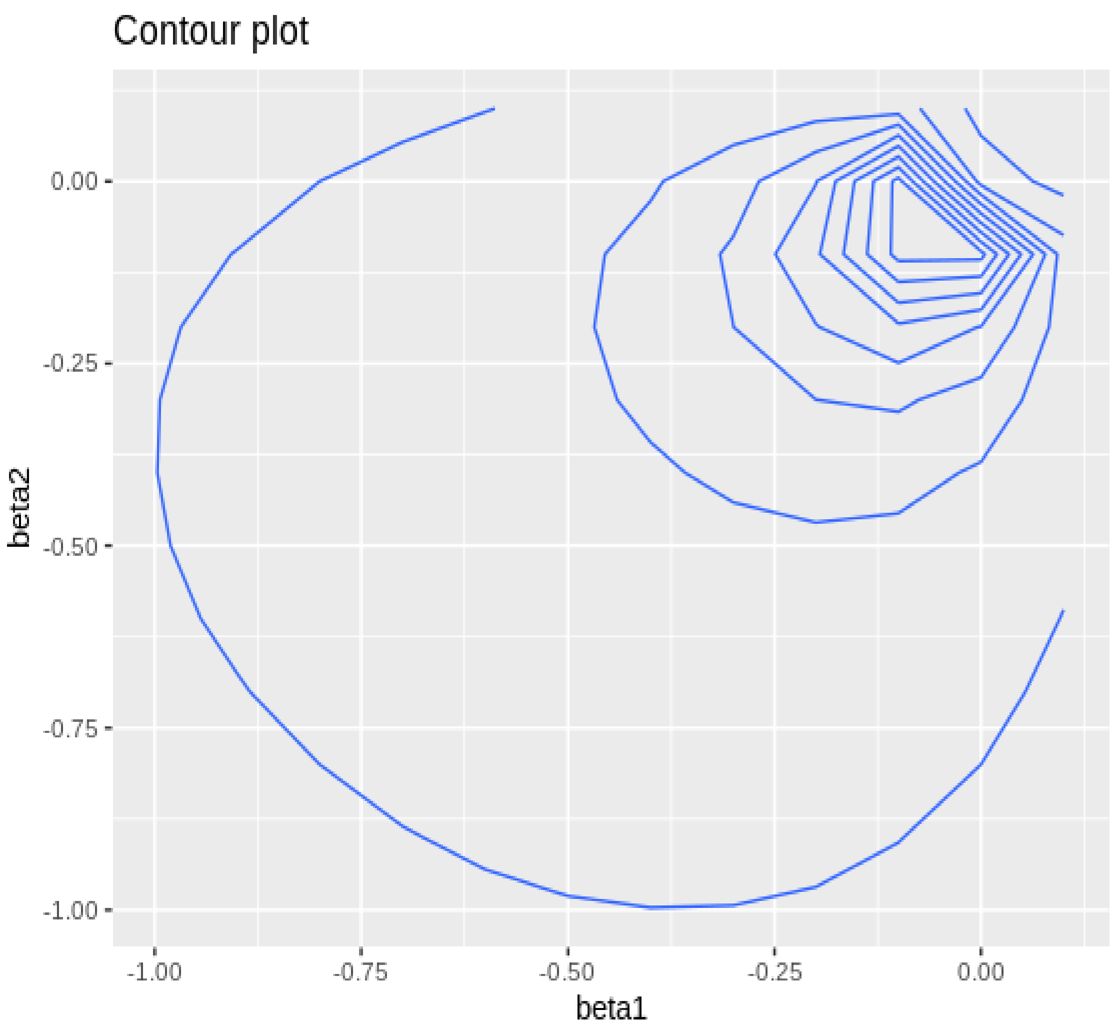
We then obtain the predicted values of y as:

$$\hat{y}_t = \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t}$$

We then find the Residual Sum od Squares(RSS)= $\sum_{i=1}^{n} (y_t - \hat{y}_t)^2$

We find RSS for all combinations of the parameters in the given range and then plot the contour plot to obtain the following:

We obtain the following contour plots:



Part-(d)

We estimate the unknown parameters using the Gauss-Newton method.
Here we define a matrix:

$\mathring{F}(\theta^a) = ((\partial f_i(\theta)/\partial \theta_j))$ where f_i denotes the functional form for y_i

where $\theta = (\alpha_1, \beta_1, \alpha_2, \beta_2)$ and j varies from 1 to 4.

We then set the approximate value of the parameters say θ^a to the estimates obtained from the Prony's method.

After that we obatin the next iterative value of the parameters as:

$\theta^{a+1} = \theta^a + \delta^a$ where:

$$\delta^a = (\mathring{F}(\theta^a)^T * \mathring{F}(\theta^a))^{-1} \mathring{F}(\theta^a)(y - f(\theta^a))$$

Thus we can iteratively obtained the values of parameters with the stopping rule as:

$$|\theta^{a+1} - \theta^a| < 0.0005$$

Thus the estimates are obtained as:

$$\hat{\alpha}_1 = 10.474 ; \hat{\beta}_1 = -4.021 ; \hat{\alpha}_2 = 2.343 ; \hat{\beta}_2 = -1.137$$

Part-(e)

We are now to estimate the values of the parameters using Osborne's Algorithm:

We need to find the eigen vector of the matrix B(g) corresponding to the smallest eigen vector.

The matrix B(g) is defined as ((b_{kj})) where:

b_{kj} =

$$(y^T U_k (G(g)^T G(g))^{-1} U_j^T y) + (y^T U_j (G(g)^T G(g))^{-1} U_k^T y) - y^T G(g)(G(g)^T G(g))^{-1}(U_k^T U_j + U_j^T U_k)(G(g)^T G(g))^{-1} G(g) y$$

We start with an estimate of $g^{(0)}=(0.01707969,-0.36000500,0.93279402)^T$ which was obtained from the Prony's Method.

From this we obtain the value of $B(g^0)$ and then calculate the eigen vector corresponding to the smallest eigen value say $g^{(1)}$. From that we find $B(g^{(1)})$.

We repeat the procedure till the stopping criteria is satisfied which is:

$|g^{(a+1)}-g^{(a)}]<0.005$

Here,

$$G^T(g)= \begin{bmatrix} g_0 & g_1 & \cdots & g_k & 0 & 0 & \cdots & 0 \\ 0 & g_0 & & & g_k & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & \cdots & g_k \end{bmatrix}$$

Here, k=2

$$U_1^T = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & & & 0 & 0 \\ \vdots & & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 \end{bmatrix}$$

$$U_2^T = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

$$U_3^T = \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

$G^T(g)=g_0U_1^T + g_1U_2^T + g_2U_3^T$

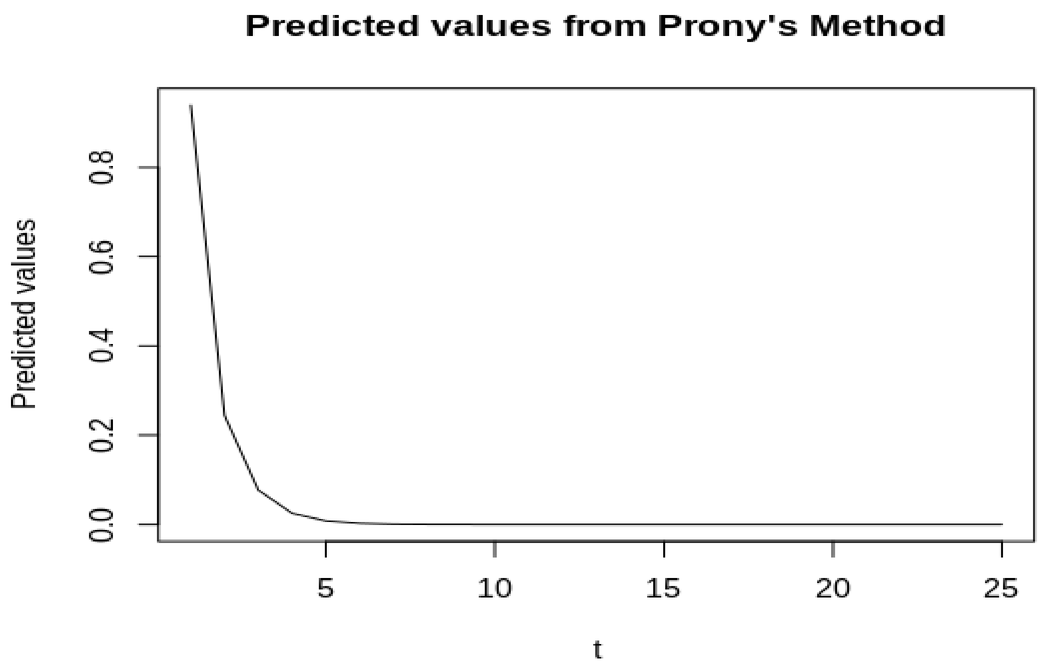
We run the procedure iteratively and obtain the estimates of the parameters as:

$$\hat{\alpha}_1=5.394\,;\hat{\beta}_1=-3.222\,;\hat{\alpha}_2=2.225\,;\hat{\beta}_2=-1.122$$

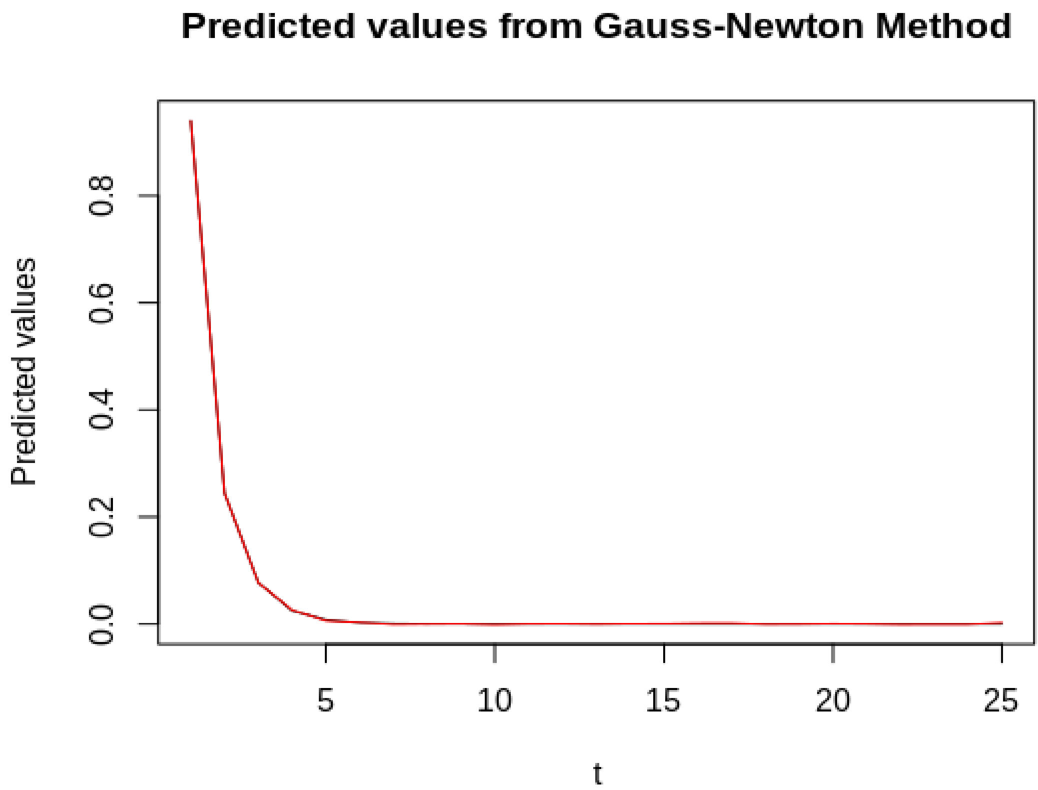
Part-(f)

We then the predicted curve that we have obtained from the three procedures:

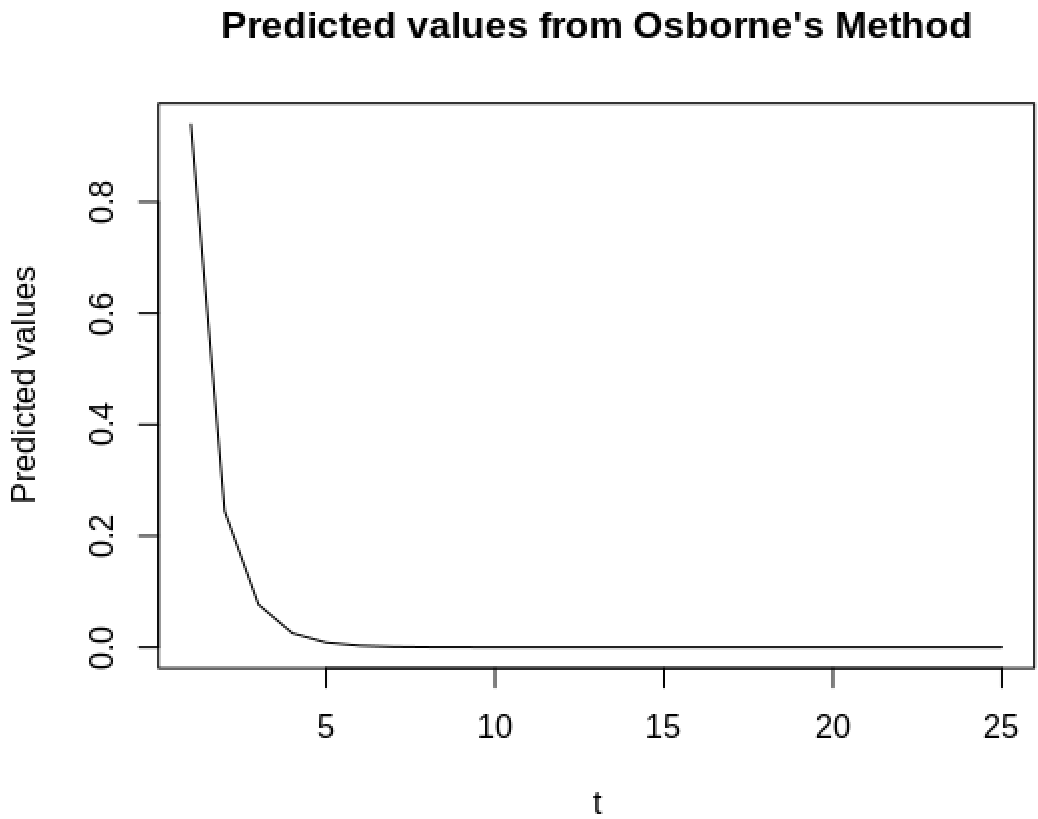
Prony's Method:



Gauss-Newton Method:



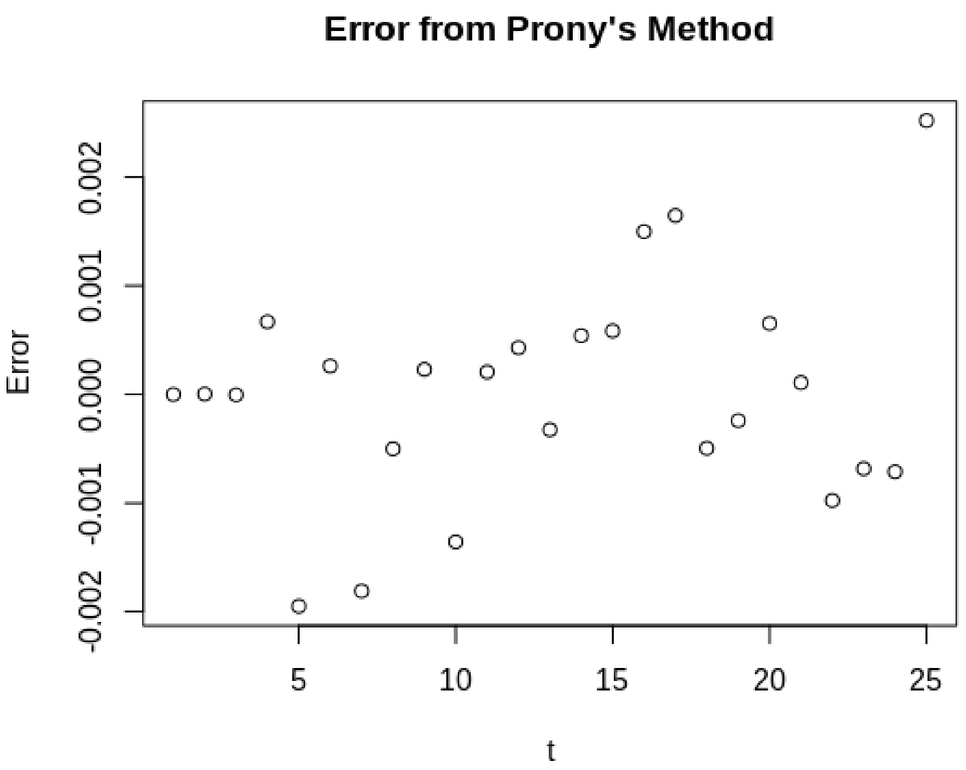
Osborne’s Method



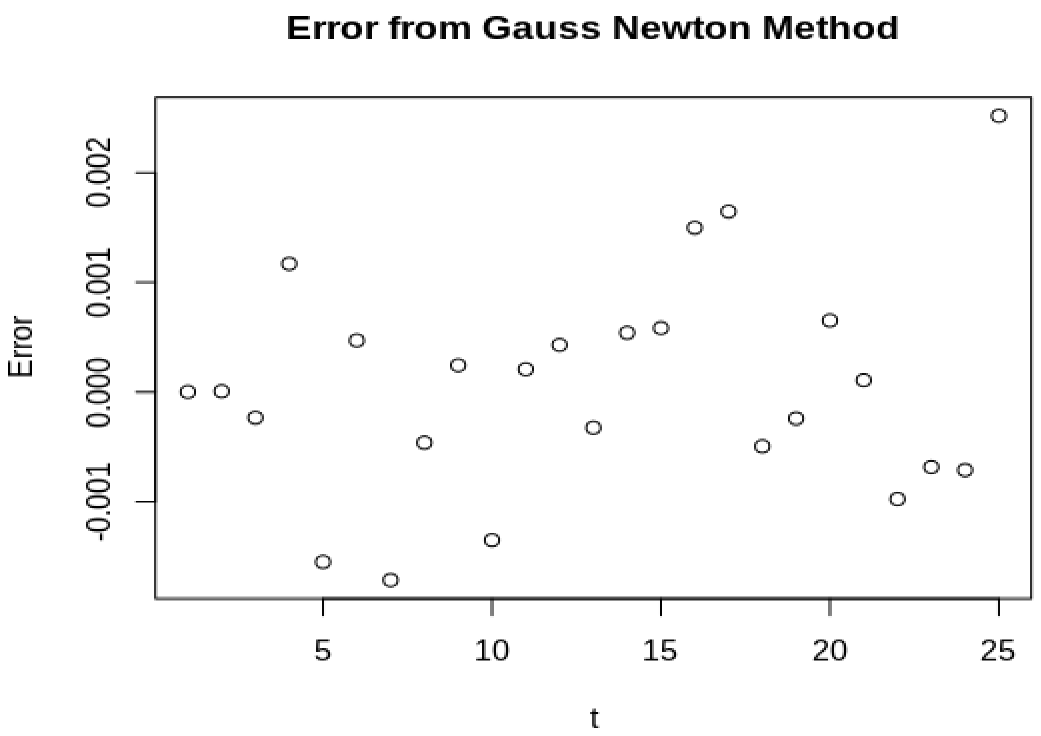
Part-(g)

We then continue to plot the errors associated in predicting the values of y in case of the three algorithms.

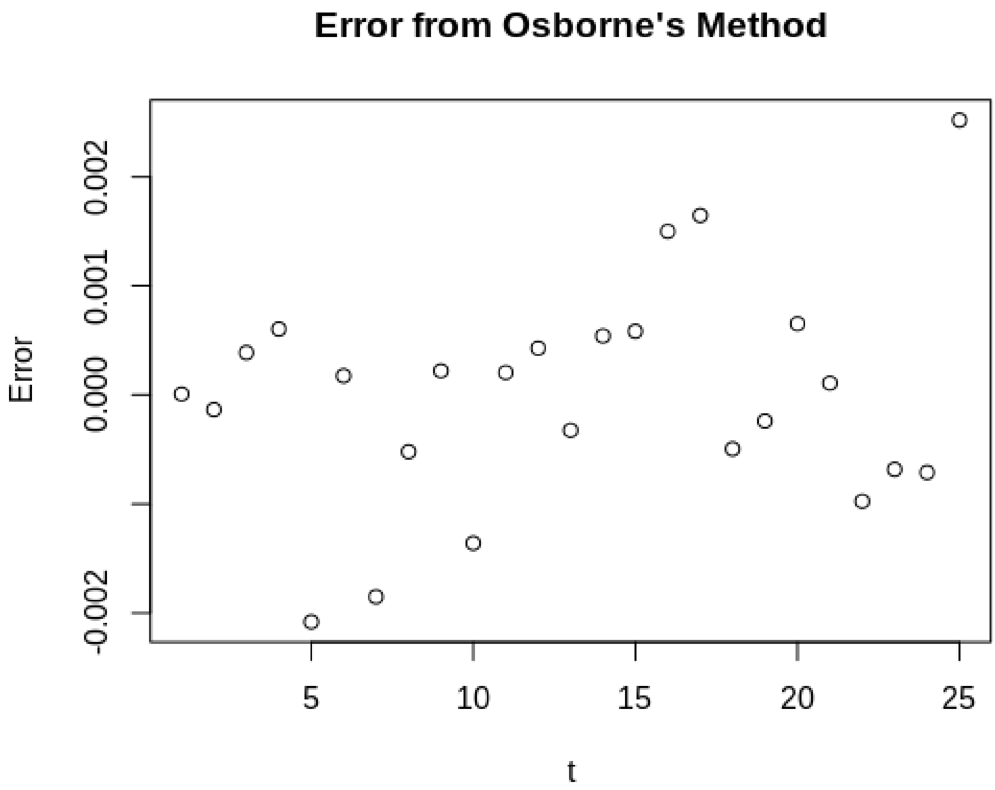
Prony’s Method



Gauss-Newton Method:



Osborne's Method:



From all the three graphs above it is clear that the error term is scattered around the point without displaying any particular pattern.

Part-(h)

We need to find the number of components using the cross-validation approach.

For this we consider models of the form:

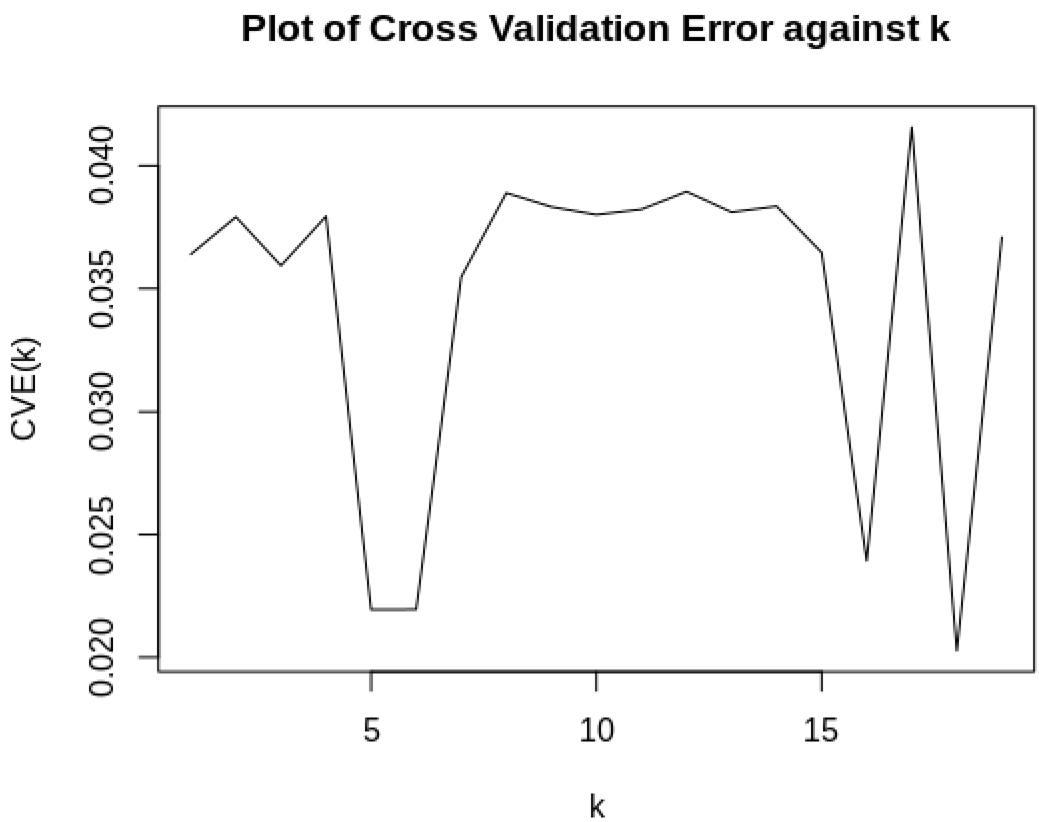
$y_t = \alpha_1 e^{\beta_1 t} + \alpha_2 e^{\beta_2 t} + \ldots + \alpha_k e^{\beta_k t} + \varepsilon_t$

Then for this model we consider the data y_{ti} where ith observation is removed and the above compartment model is condsidered for the data y_{ti} estimates of the parameters are obtained using Prony's Method. Then for a k>=2 we calculate the Cross-Validation Error for order k[CVE(k)] as:

$$CVE(k) = \left(\sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right)/n$$

where y_i is the observation that is left out and its value is predicted using the estimates of the Prony's Algorithm from the rest 24 observations.

The procedure is repeated for k=2,...,20 and the following graph is obtained:



From the graph it is clear that the CVE(k) reduces for k=5,6 and against goes up as k increases and finally for k=18 CVE(k) sees a sharp fall in the value of CVE(k).

We need to consider a trade-off as for large values of k the model becomes very complex and the parameters cannot be estimated with high precision. Therefore, for this dataset and the given compartment model the value of k can be considered to be 5 or 6 as the CVE(k) decreases and the number of parameters is not very high.

Next we need to find an confidence interval for the parameters.

$$\sqrt{(n)}(\theta-\hat{\theta})\rightarrow N_p(0,\sigma^2 n*(\mathring{F}(\theta)^T \mathring{F}(\theta))^{-1})$$

where the matrix F is same as defined above.
We evaluate the matrix F at the estimates otained from the Prony's Algorithm.
An estimate of the variance is MSE.

$$\hat{\sigma}=MSE=SSE/(n-k)$$

We obatin MSE= $1.67*10^{-6}$

From the above multivariate normal distribution we can find the marginal distributions of the parameters which will follow univariate normal distributions from which we can construct the 100(1-α)% confidence interval using the below formula:

$$(\hat{\theta}-t_{\alpha/2}\hat{\sigma}/\sqrt{(n)},\hat{\theta}+t_{\alpha/2}\hat{\sigma}/\sqrt{(n)})$$

where $t_{\alpha/2}$ denotes the upper α/2% point of a t_{n-1} distribution [Here n=25]. The estimates of the parameters is obtained from the Prony's Algorithm.
Hence the 95% Confidence Intervals of the parameters is given as:-

$$\hat{\alpha}_1 \epsilon(-1.382,10.533);\hat{\beta}_1 \epsilon(-4.913,-1.041);\hat{\alpha}_2 \epsilon(1.150,2.780);\hat{\beta}_2 \epsilon(-1.194,-1.028)$$

Thus we obtain the individual 95% CI for the parameters.