Linear Regression Models

We consider the following simple linear model for the data

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \qquad \text{with } \epsilon_i \overset{\textit{iid}}{\sim} N(0, \sigma^2),\ i = 1, ..., n$$

or equivalently

$$y_i \mid x_i \sim N\Big(\beta_0 + \beta_1 x_i, \sigma^2\Big)$$

using vector notation, we have

$$\mathbf{y} = \mathbf{x}\beta + \epsilon\ , \qquad \mathbf{y} \mid \mathbf{x}\beta \sim N(\mathbf{x}\beta, \sigma^2)$$

Note: The parameters are β_0, β_0 and the variance σ^2 assumed to be constant accross all observations. They are assumed to be fixed and unknown. The x_i are assumed to be fixed and the y_i are assumed to be random, because of the random error terms e_i.

Classical approach to Bayesian Linear Regression Models

As usual in Bayesian statistics, the posterior distribution is proportional to the likelihood times the prior distributions of the parameters, so that we have $\pi(\beta, \sigma^2 \mid \mathbf{y}) \propto L(\beta, \sigma^2)\ \pi(\beta, \sigma^2)$.

Many choices for priors are available. For example, if the following noninformative reference priors are choosen $\pi(\beta \mid \sigma^2) \propto 1$ and $\pi(\sigma^2) \propto \frac{1}{\sigma^2}$, then, the posterior, up to proportionality, becomes

$$
\begin{aligned}
\pi(\beta, \sigma^2 \mid \mathbf{y}) &\propto \prod_{i=1}^{n} L(\beta, \sigma^2)\ \pi(\beta, \sigma^2) \\
&\propto (\sigma^2)^{-n/2}\ \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \big(y_i - \beta_0 - \beta_1 x_i\big)^2 \right] \left(\frac{1}{\sigma^2} \right) \\
&\propto (\sigma^2)^{-(n+2)/2}\ \exp\left[-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{x}\beta)^T (\mathbf{y} - \mathbf{x}\beta)) \right]
\end{aligned}
$$

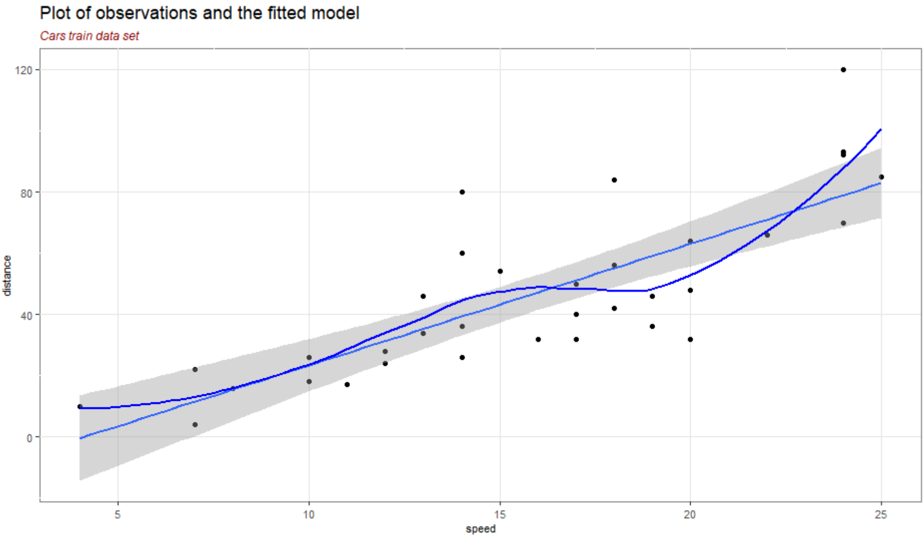
Inference for the Bayesian Linear Regression Models

As usual in Bayesian statistics, the posterior distribution is proportional to the likelihood times the prior distributions of the parameters, so that we have $\pi(\beta, \sigma^2 \mid \mathbf{y}) \propto L(\beta, \sigma^2)\ \pi(\beta, \sigma^2)$.

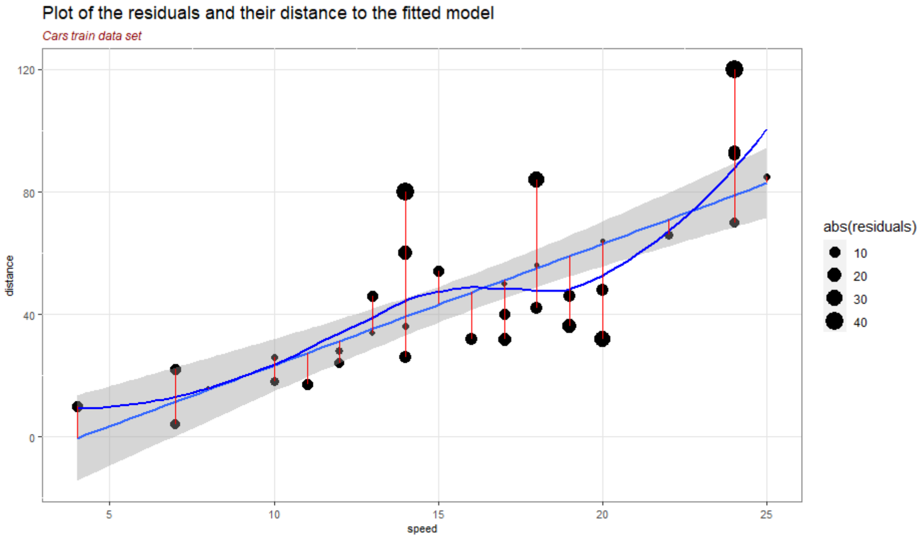
Many choices for priors are available. For example, if the following noninformative reference priors are choosen $\pi(\beta \mid \sigma^2) \propto 1$ and $\pi(\sigma^2) \propto \frac{1}{\sigma^2}$, then, the posterior, up to proportionality, becomes

$$\begin{aligned}
\pi(\beta, \sigma^2 \mid \mathbf{y}) &\propto \prod_{i=1}^{n} L(\beta, \sigma^2)\ \pi(\beta, \sigma^2) \\
&\propto (\sigma^2)^{-n/2}\ \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(y_i - \beta_0 - \beta_1 x_i\right)^2\right]\left(\frac{1}{\sigma^2}\right) \\
&\propto (\sigma^2)^{-(n+2)/2}\ \exp\left[-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{x}\beta)^T(\mathbf{y} - \mathbf{x}\beta))\right]
\end{aligned}$$

Frequentist LM, observations



Frequentist LM, residuals



References

The R Project for Statistical Computing:
https://www.r-project.org/

Python:
https://www.python.org/

course notes