

作业 8：模型质量的评价

姓名（学号）

问题：汽车油耗与自重的关系

设 y 是汽车油耗（加仑/英里）， x 是汽车重量（吨），数据如下：

$$y = c(3.4, 3.8, 9.1, 2.2, 2.6, 2.9, 2.0, 2.7, 1.9, 3.4),$$

$$x = c(5.5, 5.9, 6.5, 3.3, 3.6, 4.6, 2.9, 3.6, 3.1, 4.9).$$

1. 用传统方法拟合如下回归模型，并分别做出包含模型曲线的散点图¹：

$$M1 : y \sim N(\alpha, \sigma^2)$$

$$M2 : y \sim N(\alpha + \beta x, \sigma^2)$$

$$M3 : y \sim N(\alpha + \beta_1 x + \beta_2 x^2, \sigma^2)$$

$$M4 : y \sim N(\alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3, \sigma^2)$$

$$M5 : y \sim N(\alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4, \sigma^2)$$

$$M6 : y \sim N(\alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5, \sigma^2)$$

2. 对以上模型，所有参数均采用弱信息先验，拟合贝叶斯模型；
3. 检查各个贝叶斯模型是否符合实际数据；
4. 分别用交叉验证和信息准则对六个模型进行比较。

¹提示：5 阶多项式的平滑曲线图可以在 `ggplot2` 中设置 `geom_smooth(method = lm, formula = y ~ poly(x, 5))`