Assignment 4
MVE155 Statistical inference

Yankun Xu (940630-0237)
yankun@student.chalmers.se

---

1. Simple linear regression

This assignment is related to the simple linear regression. Based on the data, we should firstly estimate the A and B according to the fitted linear regression model $y = b_0 + b_1 x$. In this formula, $b_1 = \frac{r s_y}{s_x}(r = \frac{s_{xy}}{s_x s_y})$ and $b_0 = \bar{y} - b_1 x$. Due to the second law of thermodynamics $\ln(\text{pressure}) = A + \frac{B}{T}$, we need pre-process the data in order to make y=ln(pressure) and x = 1/T. In this case, $b_0 = A$ and $b_1 = B$. Then we will get A = 18.1847 and B = -21260.

The standard errors of A and B are calculated by the formula: $s_A = \frac{s\sqrt{\Sigma x_i^2}}{s_x\sqrt{n(n-1)}}$ and $s_B = \frac{s}{s_x\sqrt{n(n-1)}}$, where $s^2 = \frac{n-1}{n-2} s_y^2 (1 - r^2)$. Then we will get $s_A = 0.1419$ and $s_B = 133.4540$. The 95% confidence intervals for A and B can be calculated by formula $A(\text{or } B) \pm t_{n-2}\left(\frac{\alpha}{2}\right) \cdot s_A(\text{or } s_B)$. In this case, n = 32 and $\alpha = 0.05$. We will finally get 95%CI for A is (17.985 18.474), and for B is (-21532 -20987).

The final step is to check the residuals. We can calculate the residuals by $e_i = y_i - \hat{y}_i$, where $\hat{y}_i = A + Bx_i$. Figure1 indicates that residuals are randomly distributed versus temperature, we could conclude that there is no relationship between residuals and temperature. Figure2 which is qqplot showing quantiles of residuals versus standard normal distribution, indicates that these residuals are close to normal distribution with zero mean, this is an ideal condition. Based on these two plot, we can believe that our estimated model is good enough.

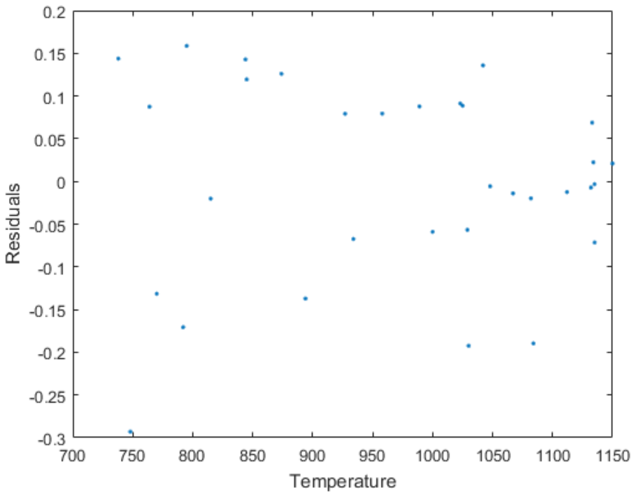

Figure 1. Residuals versus temperature

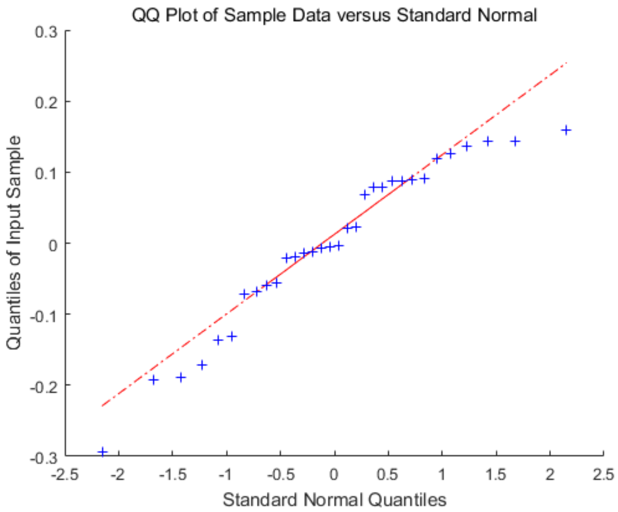

Figure 2. Examine distribution of residuals