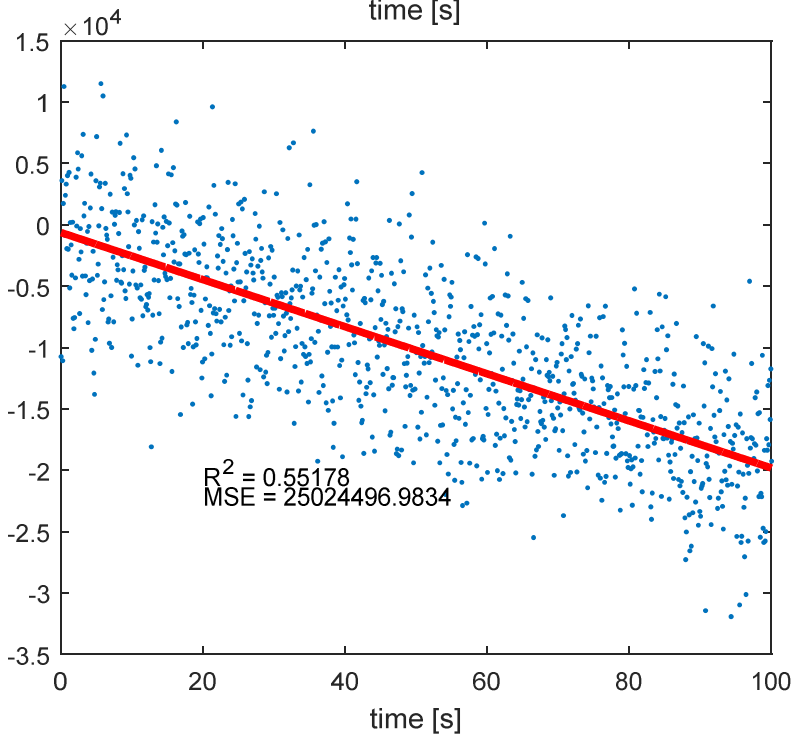
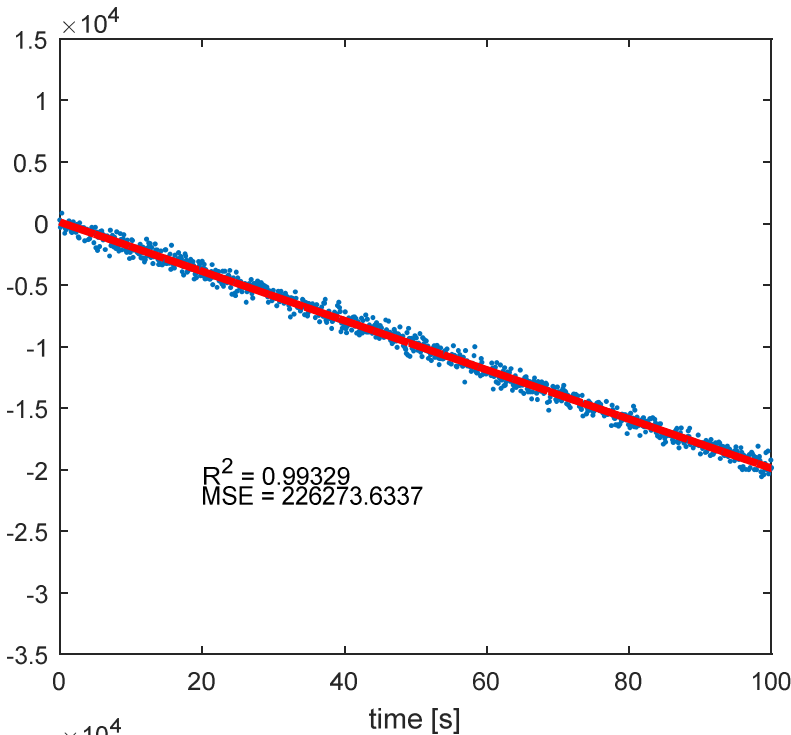


1. Escriba una función que reciba dos vectores (x y) y estime la línea que mejor se ajusta a la data. Los outputs de la función son: la pendiente de la recta (b), el intercepto de la recta (a), MSE y R^2 . Escriba una versión de su función usando loops y otra usando vectores. Presente una prueba de escritorio cuando el input es $x = [1 \ 2 \ 3 \ 4]$ y $y = [8 \ 3 \ 1 \ -5]$.

2. Escriba un programa que lea el archivo “largedat.txt” y calcule la recta que mejor se ajuste a la data, MSE y R^2 . La frecuencia de muestreo del archivo es 10 Hz. Repita para “largedat2.txt”.



LINEAR REGRESSION AND GOODNESS OF FIT

Least Squares

$$y = \hat{a} + \hat{b}x, \text{ where}$$

$$y\text{-intercept}, \hat{a} = \bar{y} - \hat{b}\bar{x}$$

$$\text{and slope}, \hat{b} = S_{xy}/S_{xx}$$

$$S_{xy} = \sum_{i=1}^n x_i y_i - (1/n) \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)$$

$$S_{xx} = \sum_{i=1}^n x_i^2 - (1/n) \left(\sum_{i=1}^n x_i \right)^2$$

$$\bar{y} = (1/n) \left(\sum_{i=1}^n y_i \right)$$

$$\bar{x} = (1/n) \left(\sum_{i=1}^n x_i \right)$$

where

n = sample size

S_{xx} = sum of squares of x

S_{yy} = sum of squares of y

S_{xy} = sum of x - y products

Standard Error of Estimate (S_e^2):

$$S_e^2 = \frac{S_{xx} S_{yy} - S_{xy}^2}{S_{xx} (n - 2)} = MSE, \text{ where}$$

$$S_{yy} = \sum_{i=1}^n y_i^2 - (1/n) \left(\sum_{i=1}^n y_i \right)^2$$

Confidence Interval for Intercept ($\hat{a}$):

$$\hat{a} \pm t_{\alpha/2, n-2} \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right) MSE}$$

Confidence Interval for Slope ($\hat{b}$):

$$\hat{b} \pm t_{\alpha/2, n-2} \sqrt{\frac{MSE}{S_{xx}}}$$

Sample Correlation Coefficient (R) and Coefficient of Determination (R^2):

$$R = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

$$R^2 = \frac{S_{xy}^2}{S_{xx} S_{yy}}$$