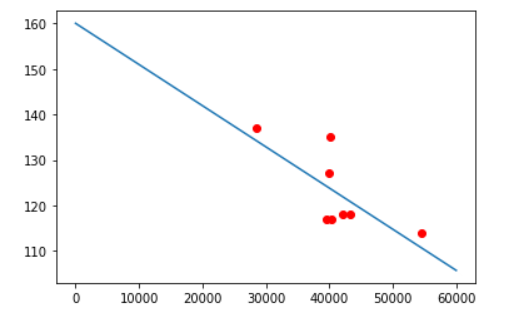
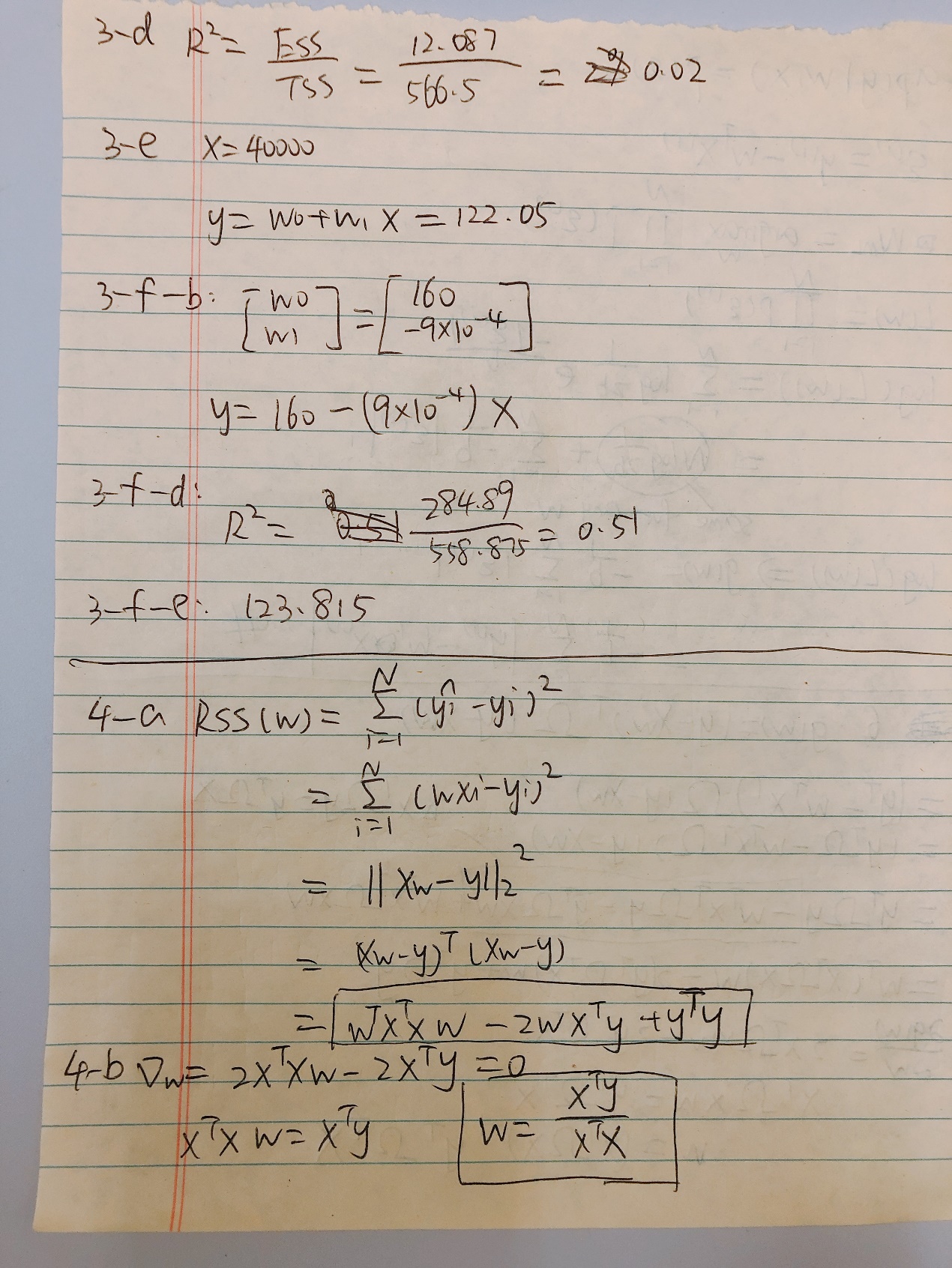


3-c before removing the noise

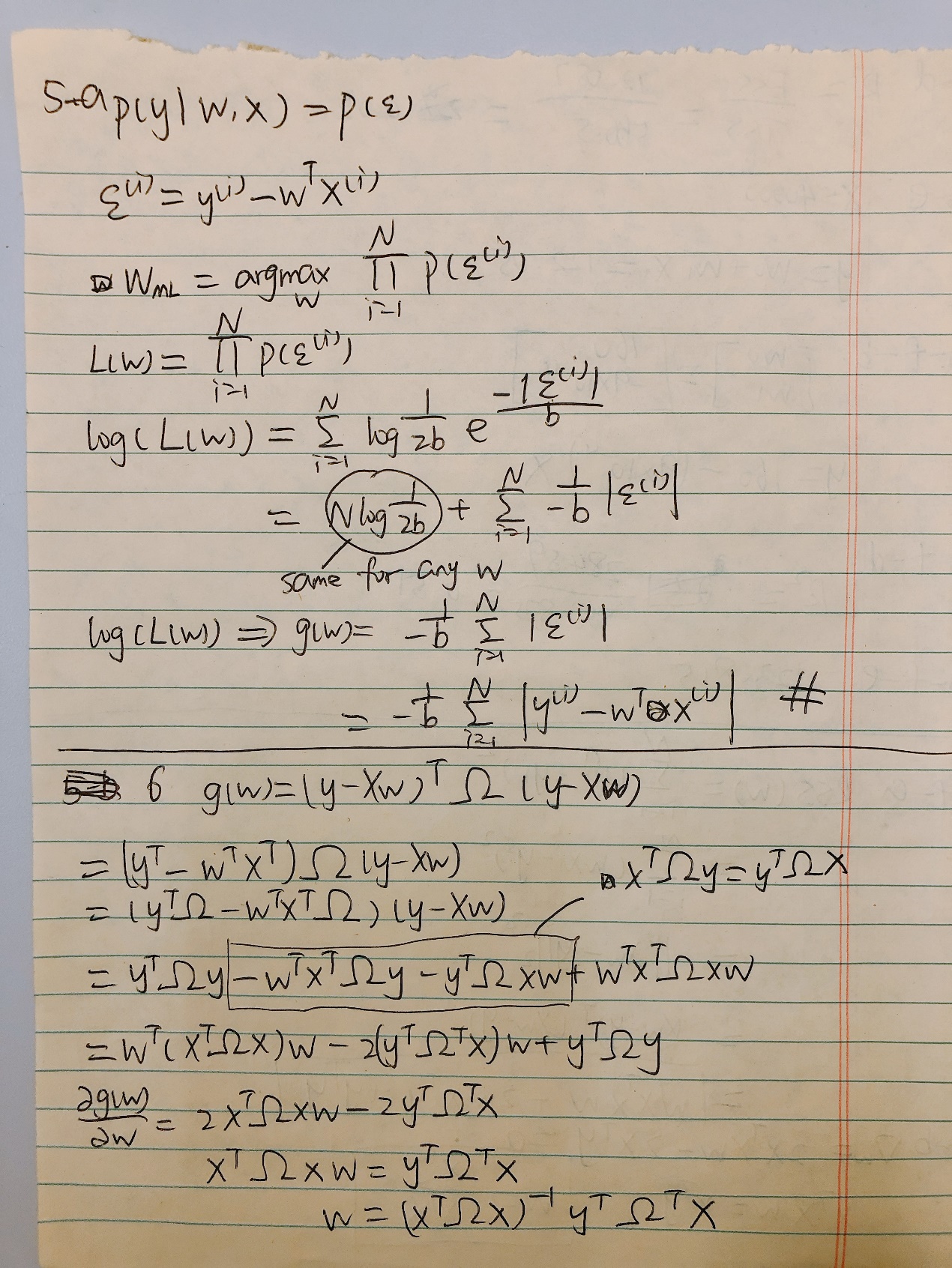


3-f-c after removing the noise

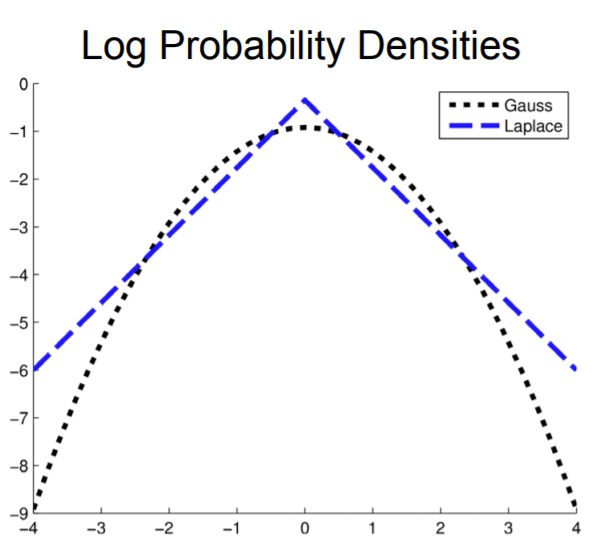


4-b

X is N\*1, y is N\*1, therefore X.T.dot(X) = a value, X.T.dot(y) is also a value.



5-b



The probability of noise data appearing in Laplace Distribution is higher than Gaussian Distribution, therefore the Laplace model would be more robust than the Gaussian.