1. Perform a Deming regression using the provided data set (HW2\_data1.txt), and compare the results with a regular regression analysis.

Note: yy is the observed outcome, and xx is the observed predictor, and x\_star is the “unobserved” true x value. Your analysis should be based on the observed data (xx and yy), and x\_star could be used to evaluate the effectiveness of Deming regression in taking errors-in-variable into consideration.

1. For the following matrices and vectors, ***manually*** conduct the indicated operations (you can use software to verify your results if you want)
2. Find the determinants of and.

Note the relationships across columns in C or D. Particularly, any individual column in C cannot be obtained as a linear combination of other columns in C. However, the third column of D can be obtained as the average of first two columns (or in other words, ½\*first column + ½\*second column). This illustrates the situation when a matrix has a zero determinant.

1. Compute
2. Find the inverse of
3. In a diabetes study, 1123 subjects were recruited, and a number of clinical traits and information were collected, including (see data in the attached file “HW2\_data2.txt”, which is the same data set used for HW1):

Sex: male/female Age: age of the study subject

bmi: body mass index fbg: fasting blood glucose

fins: fasting insulin hba1c: hemoglobin A1c

tg: total glyceride tcho: total cholesterol

hdl: high density lipoprotein ldl: low density lipoprotein

Particularly, the investigators are interested in the effects of fbg (X1) and tg (X2) on hba1c (Y). In addition, hba1c > 6.5 is considered to be diabetic. So another question is whether the same model can be used to characterize the relationship between predictors and hba1c for all individuals, or whether two different models are needed: one for people with diabetes, and the other for those without diabetes. Now, you are assigned to analyze the data, and are asked to complete the following tasks.

Consider a multiple linear regression model between hba1c and fbg + tg.

1. Given that and (with the order of intercept, fbg, tg, in the model). Manually obtain estimates for the intercept, and coefficients for fbg and tg.
2. Manually find the variances for and , and their correlation coefficient. Check your results using software of your choice.
3. For the two ways in entering the predictors into the model (fbg first, then tg, vs tg first, then fbg), summarize in a table for the sequential and partial sums of squares, respectively, for fbg and tg.