1. Let the matrices A, B, and C have dimensions permitting matrix multiplication of A and B, and B and C. Let U=AB and V = BC. Demonstrate with a mathematical derivation using the definition of matrix multiplication as a dot product, that UC = AV (show that any arbitrary element of UC equals the matching element of AV).

Solution:

Let be an matrix, be an matrix, and be a matrix.

Matrix is defined as . So, the dimensions of are

Matrix is defined as . So, the dimensions of are .

The product of matrices and , , is a matrix of dimensions .

The product of matrices and , , is a matrix of dimensions .

Thus, matrices and have the same dimensions. We now have to show that an arbitrary element of , equals the matching element of .

The element is given by the dot product of the -th row of and the -th column of . As can be seen from the dimensions of matrix , there are elements in each of its rows. And there are elements in each column of . Using the summation operator, the dot product of the -th row of and the -th column of can be written as

But since , is the dot product of the -th row of and -th column of Each row of has elements, and each column of also has elements. Therefore, . Substituting this for expression for in , we get

We could rearrange the two summation operators in the above expression for to get

In the above expression, is the dot product of the -th row of and the -th column of (each of them has elements), giving us the element . Replacing by in the expression for , we get

because is the dot product of the -th row of A and -th column of (each has elements) giving us the -th row and -th column of , i.e., .

Hence, we have proved that .

1. Suppose you are fitting data to a quadratic polynomial from *x* data ranging from 0 to 10 using OLS. Based on the estimates of , you would like to predict the value of η at =20, i.e., to extrapolate the curve. How would you calculate the uncertainty in your extrapolation? Use Matrix notation.

Solution:

Our multiple regression model is **,**

where is a (11 ) column vector of observations of the dependent variable

is a (3 column vector of regression coefficients:  **=**

is a (11 matrix of a column of ones (for the intercept) and two independent variables and .

[1 0 0

1 1 1

1 2 4

1 3 9

1 4 16

= 1 5 25

1 6 36

1 7 49

1 8 64

1 9 81

1 10 100]

is a (11 column vector of error terms whose mean is zero and satisfies the classical assumptions of no autocorrelation and homoscedasticity.

The OLS estimate of is:

To get the means, variances, and covariances of the OLS estimators, we do the following:

After these preliminary results, we get to the main problem.

Our prediction of at =20 is [1 20 400]

Variance of [1 20 400][1 20 400]

[0.5804 -0.2203 0.0175

Using Matlab,  **=** -0.2203 0.1256 -0.0117

0.0175 -0.0117 0.0012]

Finally, [1 20 400] = 56.0117

Therefore, variance of = 56.0117

1. Consider the parameter estimate vector

with estimated covariance matrix . Consider an arbitrary row vector .

1. Write the general formula for the in the matrix form.
2. Using matrices and the result of a, compute the variance of by expressing the scalar in matrix form.
3. Using matrices, compute the variance of by expressing the scalar in matrix form.
4. Derive an expression for the *approximate* variance of . Why can’t you find a numerical answer here, but you could for parts b and c above.

Solution:

1. = Var
2. For the function , and . Therefore, Var( = (1 -1) = 3
3. For the function , and . Therefore, Var( = (1 1) = 7

Let’s do a first-order Taylor series expansion of around the means of and . Let their means be and , respectively.

Since ,

= (since the expansion is around the means of and )

Therefore, we have

Applying the variance operator to both sides of the above equation, we have

2

This is our approximation for the variance of the product of two random variables.

If we apply the same method to any linear combination of two random variables, the first-order Taylor series will give us the exact value of the function, and it is not an approximation like it is in this case. Therefore, we can find an exact numerical answer in b and c but not in d.

1. In Matlab, simulate Gaussian data with the following characteristics:

Perform 100 replicate experiments. In each case, use WLS to estimate the parameters (feel free to use your exact knowledge of the data noise). Do this 2 different ways: 1) using the correct 3-parameter model, 2) using the incorrect 2-parameter model where you drop the β3 term.

a) Produce one graph showing sample data, and the 2 fits

b) Calculate the sample mean of the parameter estimates from both models and compare to the true values

c) Calculate the sample standard deviation of the parameter estimates in each case and compare.

Solution:

a)



b) Correct 3-parameter model:

True beta: [100.0, 100.0, 100.0]

Mean estimates: [103.8818, 98.4490, 94.9741]

Incorrect 2-parameter model with dropped β3 term:

True beta: [100.0, 100.0]

Mean estimates: [165.4079, 158.0896]

In the correct 3-parameter model, the sample mean of beta estimates is very close to the true betas. The Weighted Least Squares (WLS) in the correctly specified model gives unbiased estimates of the parameters. Not only this, in the presence of heteroscedasticity, WLS gives the Best Linear Unbiased Estimates (BLUE).

However, in the incorrect 2-paramter model with dropped β3 term, the sample mean of beta estimates is very different from the true beta. OLS and WLS give biased estimates when a relevant explanatory variable is dropped from the regression.

c) Correct 3-parameter model:

True beta: [100.0, 100.0, 100.0]

Mean estimates: [103.8818, 98.4490, 94.9741]

**Std deviations: [43.1076, 45.4926, 64.7647]**

Incorrect 2-parmater model with dropped β3 term:

True beta: [100.0, 100.0]

Mean estimates: [165.4079, 158.0896]

**Std deviations: [4.2443, 27.2735]**

We see that the standard deviation of parameter estimates is much less in the incorrectly specified 2-parmter model than the standard deviations in the correct 3-parmeter model. The standard errors of the remaining variables in a model with dropped variables can go down if the dropped variable is highly correlated with the retained variable. In our case, the dropped variable, , has a correlation coefficient of 0.9753 with the retained variable, in the sample. Given this very high correlation, the effect of the dropped variable on dependent variable is absorbed by the retained variable, leading to a smaller variance of the disturbance term. On the other hand, if there was no correlation or very little correlation between the dropped and retained variables, we would expect the standard error to go up in the incorrectly specified regression, as in this case the variability of the dropped variable adds to the variability of the error term in the regression.

In summary, even though the standard errors are smaller in the incorrect 2-parmater model, the estimates we get are biased (coming from a biased estimator).

5. Simulate the following 2-exponential isotope decay problem

where the decay constants (α’s) are known, but the amount of each isotope (β’s) are unknown. In Matlab, perform 1000 replications of Poisson data (use the function *poissrnd*) and for each, solve for **b**OLS and **bWLS.** using the OLS and WLS matrix formulas. For the purpose of WLS, assume you have perfect knowledge of the true data variance.

a) Calculate sample mean for OLS and WLS and compare to underlying true means.

b) Calculate the sample standard deviations of the parameter estimates for OLS and WLS. Discuss whether one method is better than the other method and why.

c) Calculate the predicted standard errors of the estimates in each case and compare to the sample standard deviations

d) Calculate the predicted correlation coefficient between β1 and β2 and compare to the sample correlation coefficient