

Numpy - Assignment

All assignments should be submitted to the Github repository you have been assigned to. The deadline is August 16th by 11 PM. You should write a report in Latex with the solutions to the problems below. If the problem requires you to code, then the code should also be included in the report. You can add code to Latex with the `lstlisting` package (see this stack post) or the `minted` package (see this other stack post).

Problem 1 Read the `numpy.random` documentation page. Test the linear regression function using a normal distribution for the unobserved heterogeneity and some other distribution of your choice for the independent variables.

Problem 2 Extend `linreg_ols` to compute the standard error of the OLS estimates under the classical assumptions. Remember that under the classical assumptions, the standard error of the OLS estimates:

$$SE(\hat{\beta}_i) = \sqrt{s^2 \cdot [(X'X)^{-1}]_{[i,i]}}$$

where s^2 is computed from the residuals of the estimation:

$$e \equiv y - X\hat{\beta}$$

$$s^2 \equiv \frac{e'e}{n - K}$$

Remember to update the output of the function and extend its documentation.

Problem 3 Run the linear regression suggested in Equation (8) of Pace and Barry (1997).

Problem 4 Python has been around since 1990 and has a huge number of packages. Figure out if Python has a function or a package (or many packages) for estimating the parameters of a linear regression using OLS. If it does, what is the name of the function or the package? What are the outputs of the function you found?

Problem 5 When the classical assumption of homoskedasticity fails, we need a different estimators for the standard errors of the OLS estimates. White (1980) proposes a heteroskedasticity-robust estimator for the standard errors of the OLS estimates, which is now known as White's standard error. Equation 2.4.1 in Hayashi (2000) shows White's standard error:

$$\widehat{SE}(\hat{\beta}_i) \equiv \sqrt{\frac{1}{n} [S_{xx}^{-1} \hat{S} S_{xx}^{-1}]_{[i,i]}}$$

There is a slight change of notation in this part of the Hayashi (2000), and x'_i is the i th row of X (x_i is a column vector with the explanatory variables in the i th row of X). The term S_{xx} is the sample mean of $x_i x'_i$: $S_{xx} = \frac{1}{n} \sum_{i=1}^n x_i x'_i$ (Equation 2.3.6 in Hayashi (2000)). The term \hat{S} is an estimator for a matrix of fourth moments, and it is defined as $\hat{S} = \frac{1}{n} \sum_{i=1}^n e_i^2 x_i x'_i$ (Equation 2.5.1 in Hayashi (2000)), and e_i is as before (residual for the i th observation). Implement this estimator in the `linreg_ols` function. You could write a local function to estimate White's standard error in the `linreg_ols.m` file.

Problem 6 (Optional) Implement the t -test based on White's standard error. Also, compute the p -value from given the statistic.

Problem 7 Newey and West (1987) proposes another estimator to the standard errors for the OLS estimates under weaker assumptions. The estimator the authors propose is robust not only to heteroskedasticity, but also autocorrelation, and is known as the HAC standard errors. Does the package you found in Problem 4 have such a function to compute the HAC standard errors? If not, can you find some other package that does have it? If so, what is the name of the package?

Problem 8 (Optional) Implement the t -test based on the HAC standard error. Also, compute the p -value from given the statistic.

Problem 9 Extend `linreg_ols` to accept an optional input named `cov_type`, which specifies the type of standard errors estimator to report. The variable `cov_type` can take one of many values, say "Standard", "White" or "HAC". Your function should be able to check if `cov_type` is one of those types and act accordingly. If it is not, the function should raise an appropriate error alerting the user with a very clear message of what is wrong.

References

- Hayashi, F. (2000). *Econometrics*. Princeton University Press. ISBN: 9780691010182. URL: <https://books.google.com/books?id=QyIW8WUIyzcC>.
- Newey, Whitney K. and Kenneth D. West (1987). “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix”. In: *Econometrica* 55.3, p. 703. URL: <https://login.proxy.lib.duke.edu/login?url=https://search-proquest-com.proxy.lib.duke.edu/docview/214867593?accountid=10598>.
- Pace, R Kelley and Ronald Barry (1997). “Sparse spatial autoregressions”. In: *Statistics & Probability Letters* 33.3, pp. 291–297. URL: [https://doi.org/10.1016/S0167-7152\(96\)00140-X](https://doi.org/10.1016/S0167-7152(96)00140-X).
- White, Halbert (1980). “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”. In: *Econometrica* 48.4, pp. 817–838. URL: <https://www.jstor.org/stable/1912934>.