The chapter Linear Methods for Regression from The Elements of Statistical Learning gives an introduction to the Linear Regression Model and one of the most famous estimation methods of Least Squares. Using Least Squares we can estimate our coefficients in such a way that the Residual Sum of Squares is minimized. The predicted values are then fitted using these estimated coefficients. The different extensions and variations of linear regression are also discussed.

There is an introduction to testing the significance of groups of coefficients simultaneously. If we have to test if a categorical variable with k levels can be excluded from the model we use the F test to check if the coefficients of the dummy variables corresponding to each level can be set to zero. F test measures the change in the Residual Sum of Squares of each additional parameter in the bigger model. We can also use the Z-score to test if a particular coefficient is zero. An absolute Z-score value of greater than two is significant at 5% level (The level which we choose in general).

However Least Square Estimation has a couple of cons. The estimates have low bias and large variance. This decreases the accuracy of our prediction which can be improved by shrinking or keeping the values of some coefficients as zero. Often we would also like a smaller subset that exhibits the strongest effects. Thus, the interpretation of the model could change as we might be willing to sacrifice some variables.

There are a different number of ways to choose this subset to restrict the linear regression model: subset selection, ridge regression and the lasso. We can use best subset regression which finds for each k ∈ {0, 1, 2,...,p} the subset of size k the best subset is the one that gives the least residual sum of squares. We can use Forward Stepwise selection to seek a good path through all the subsets. We start with the intercept and then keep adding the predictor which most improves the fit. This method is preferred as if the value of p is large it is impossible to calculate the best subset sequence, but is always possible to calculate the forward stepwise sequence. The approach also has lower variance and is more constrained in its search. The opposite to this is Backward Stepwise where we start with the full model and deletes the predictor with the least Z score. However we can only use this method when N > p. Forward Stagewise Regression is similar to Forward Stepwise Regression where none of the other variables are adjusted when a new term is added. Hence this is inefficient for higher dimensional problems as it may take more than p steps to obtain the fit.

Ridge Regression does a proportional shrinkage by shrinking the regression coefficients by imposing a penalty on their size. The ridge estimate has a Gaussian distribution, is a mode of posterior distribution and is also the posterior mean. Lasso alters each coefficient by a constant factor λ, which truncates at zero. The three methods are Bayes estimates with different priors and are maximizers of the posterior.

Least Angle Regression is intimately connected with the lasso and identifies the variable most correlated with the response. LAR moves the coefficient towards the least square value. When another variable gets closer in correlation with the residual, it joins the active set and their coefficients are moved together in a way that keeps their correlations tied and decreasing. This process continues until all variables are in the model and ends at the full least squares fit. The LAR algorithm is efficient as it requires the same order of computation as that of a single least squares fit using the p predictors.

When selection and shrinkage methods have to be applied to the multiple output case, an application of a univariate technique individually to each outcome or simultaneously to all outcomes is considered. A data reduction technique known as Canonical Correlation Analysis combines responses and the solution is computed using a generalized Singular Value Decomposition (SVD) of the sample cross-covariance matrix.

There is another LAR like algorithm known as Incremental Forward Stagewise Regression which focuses on forward stagewise regression. We consider the linear regression version of the forward stagewise boosting. A coefficient is generated by repeatedly updating the coefficient of the variable most correlated with the current residuals by a small amount Ʃ. If we let this small number tend to 0 it is identical to the lasso path. This limiting procedure is called infinitesimal forward stagewise regression or FS0. This method is more constrained than lasso and may be more useful when p >> N. The coefficient profiles are much smoother and hence have lower variance than the coefficients generated by the lasso method.

The Lasso method has been heavily studied by various people. Donoho showed that under certain assumptions on the model matrix X, if the true model is sparse, this solution identifies the correct predictors with high probability. Meinshausen showed that one can also use the lasso to select the set of non-zero predictors, apply the lasso again, but using only some selected predictors from the previous step. This is called as the relaxed lasso. Fan and Li showed that the lasso penalty function can be modified so that larger coefficients are shrunken less severely which is called the Smoothly Clipped Absolute Deviation (SCAD) penalty.

WSP is a globally renowned in professional services and its expertise in environmental consultancy and sustainability. The goals of sustainability are important to me and working for WSP would give me the opportunity to work with other professionals with the same goals and morals as me. The diverse and collaborative environment at WSP would also give me the opportunity to grow and learn new skills everyday.

My contribution for the first two years and beyond that would be to work as hard as I can to learn and develop the skills required for the role. As a graduate role I will come with a fresh outlook to situations and will also bring adaptability to various situations. I have experience working in team environments and working to achieve common goals in my past work experiences. I believe that working together as a team is the best way to efficiently achieve your goals. I also want to grow at the company to the best of my ability and thus, I will embrace all opportunities to prove myself at the company.

When I was in grade 7, I was the captain of my middle school house cricket team. This was a big deal for me as I was very passionate about cricket at the time. As I was in grade 7, most of my peers were a year above me. Alot of my peers thus, felt that they deserved to be captain as they were senior. However with time they could see that the right choice was made as my ability was clear to see and we managed to win the tournament that year. I consider this to be my greatest achievement as its the one I am most proud of.

During my time at Deloitte I was part of a team environment focused on meeting tight deadlines. I also love to play various sports like cricket, football and badminton. These two experiences have taught me the importance of teamwork and working together to achieve your goals. They have also taught me the importance of hard work, dedication and grit.

I also have volunteer experience as a FUTURES volunteer which gave me an insight into the importance of giving back to the community and the importance of research done into making the world a more advanced and better place.