Bonus Task Energy Market Optimisation

Group 13

10^{th} October 2021

1 Vector Auto-Regression

1.1 Decision

Consider the for-ex trading, where we do not know the future prices at a given time t. We ideally want to know the best price at the next $timet + 1$, often it is challenging to get the result before a given time $t + 1$. Since our energy market price is updated every 30 minutes, by using Vector Auto-Regression, we are able to predict the price before $t + 1$ hits. Thus, Vector Auto-Regression is computationally less expensive than ARMA. To improve this model, we could take into account external data such as Covid-19 (exogenous variable) or the interaction between Weather and Bushfire data (endogenous variable). As a result, our model might be less accurate due to the prediction errors. On the other hand, we could also try using other forecasting algorithm such as Long Short-Term Memory.

1.2 Concept

The equation for Vector Auto-Regression with 2 predictors and 195 lags are:

$$\mathrm{Y}_{1,t} = \alpha_1 + \beta_{1,1} \times Y_{1,t-1} + ... + \beta_{1,195} \times Y_{2,t-1} + \epsilon_{1,t}$$
$$\mathrm{Y}_{2,t} = \alpha_1 + \beta_{2,1} \times Y_{1,t-1} + ... + \beta_{2,195} \times Y_{2,t-1} + \epsilon_{2,t}$$

where $Y_{1,t}$ is the matrix of predicted price and $Y_{2,t}$ is the matrix of predicted Demand. Here, we are taking the predicted price for the bonus task. For the prediction results, we use the training period to obtain β and test it on 1^{th} July 2021 until 15^{th} of August 2021. To calculate the revenue, we simulate the predicted price on the mandatory task algorithm. Finally, we obtained a revenue of \$4.184 millions. This is a pretty good estimate to the mandatory task's revenue, which is \$6 millions.

1.3 Justification

Prior to fitting in the Vector Auto-regression model, we consider:

- Granger Causality Test: it is found that the optimal columns to use for Vector Auto-regression model in $market\_data.xlxs$ are Victoria Trading Price and Demand.
- Dickey-Fuller Test (Stationary Test): since all the p-values are significant, we reject $H0$ that the data is non-stationary. Thus, all the columns are stationary.
- Log-transformation: To reduce the variance of Victoria Trading Price and Demand, we consider log transformation with the formula: $log(X) + min(X) + 1$. We chose to add $min(X) + 1$ to deal with negative numbers. Since our data is already stationary, we did not use log-difference.
- Hannah-Quinn Information Criterion (Optimal Lag): we use 195 as the optimal lag.