Multivariate Methods Assignment: Canadian Weather

Group 6: Mona Bazmawe, Monika Helak, Sebastian Tanco, Zane Kliesmete

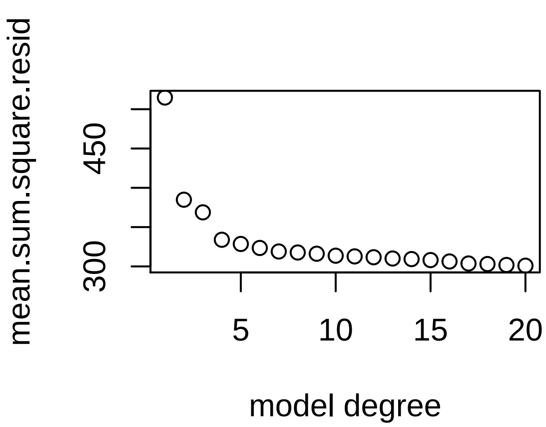
May 23rd, 2022

1. Introduction

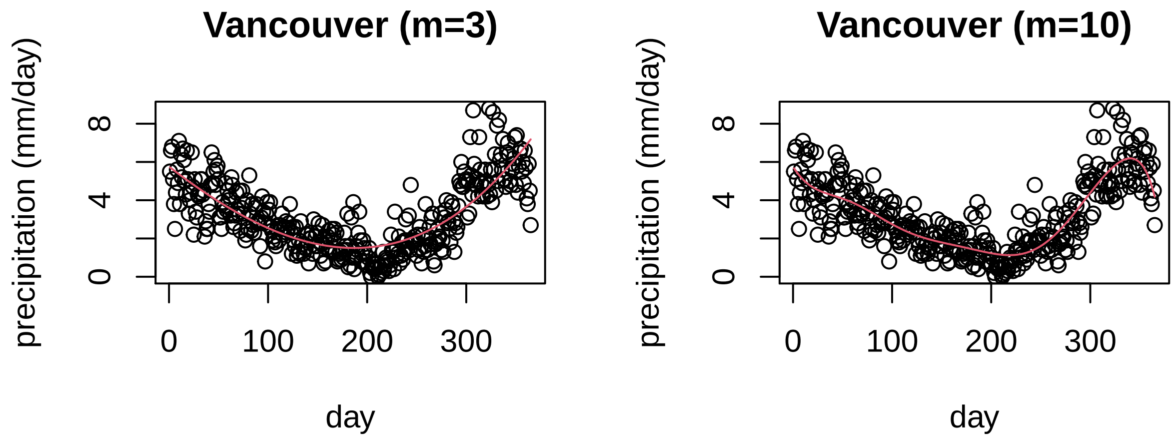
Climate can be difficult to forecast. One of the possible approaches to studying climate includes looking at weather data and finding similarities in weather patterns across different cities. We can accelerate policy decisions relating to climate change by implementing similar solutions in cities with similar weather patterns. In the following analysis we utilize a two-dimensional representation to help us identify cities with similar weather patterns and visually quantify and contrast their differences. The data set used consists of 35 cities in Canada with one year (365 days) of precipitation data.

2. Functional Data Analysis

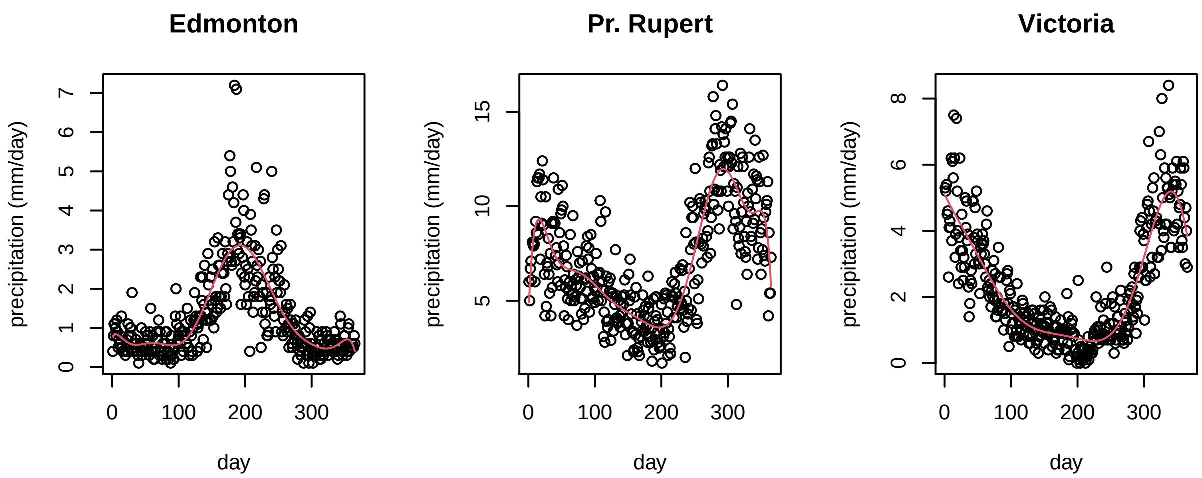
In order to perform a polynomial fitting we first needed to determine the optimal degree d for the polynomial function. The below plot shows the sum of squared residuals vs. the model degree *m*. We can tell that with a higher degree polynomial a better fit is achieved.



From the above plot we chose two values as the degrees of the polynomial and compared the fit of the model to the data. Looking at *m=3* and *m=10*, we can clearly see that the *m=10* polynomial fits closer to the data than *m=3* when setting the location equal to Vancouver.

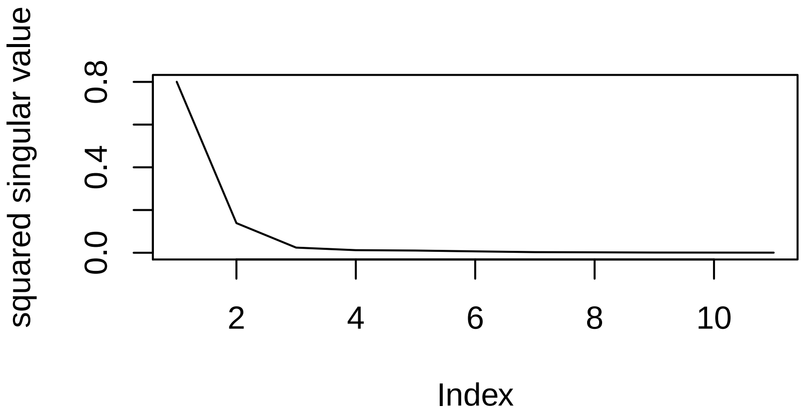


After looking at a few more locations namely Edmonton, Prince Rupert and Victoria, we concluded that *m=10* appears to be a good choice of degree for the polynomial fitting to the data.



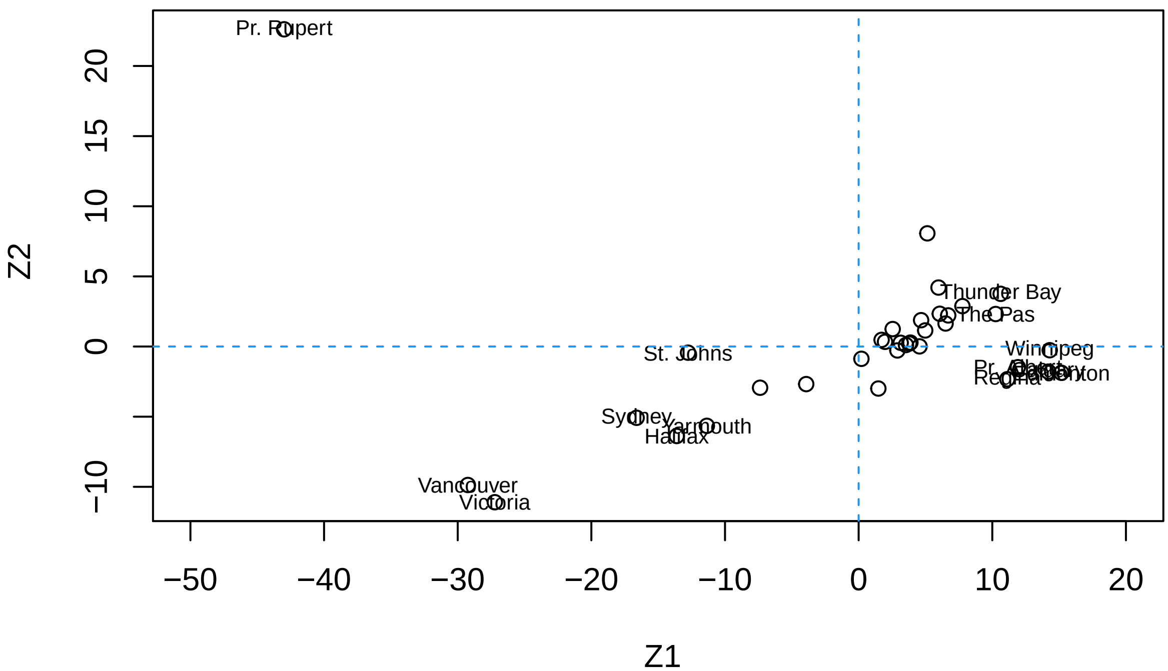
3. Multidimensional Scaling

Next, MSD is performed. First, we apply the column centering on the parameters matrix and then we obtain the SVD of this matrix. The plot below shows that in the first few dimensions we capture most of the information in the theta matrix.



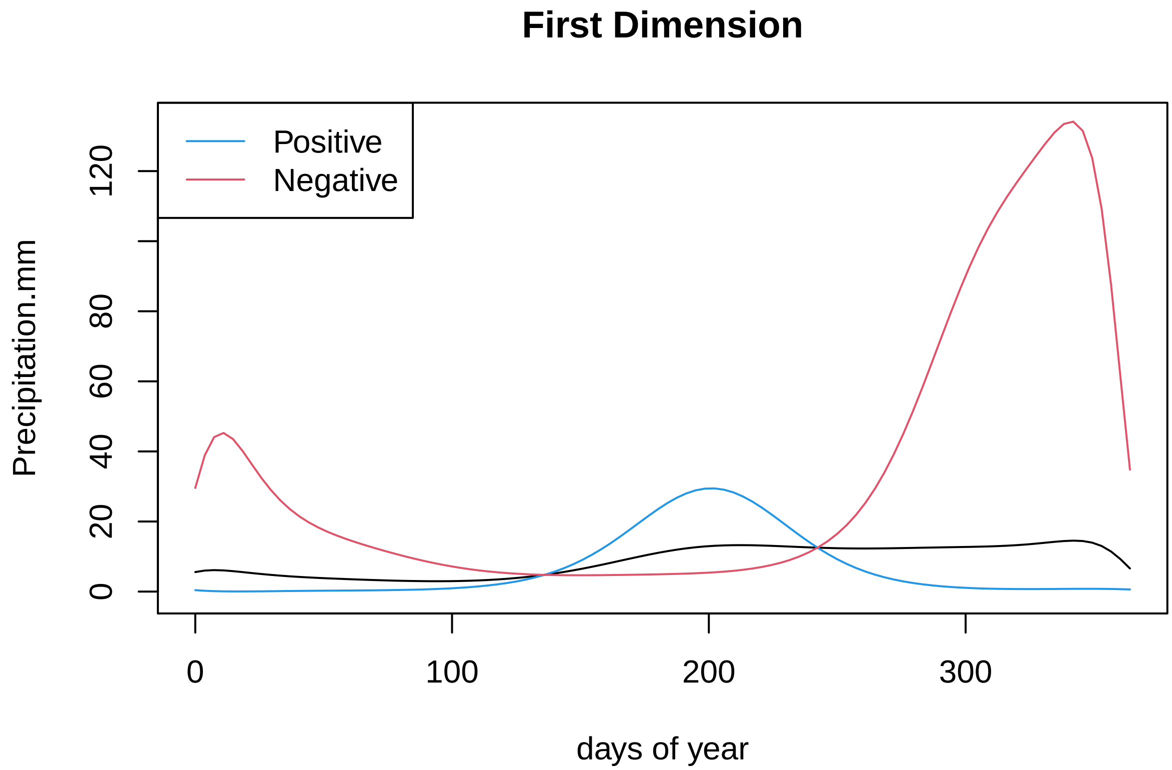
4. Functional Biplot

Here we construct the scores matrix Z_k with $k=2$ and then plot these scores. Each city has a score in each dimension which is captured in the 2x2 scatter plot. From this plot we can visualize which cities share similar precipitation characteristics in the first two dimensions which capture the most variability in the data.



It can be seen that the origin ((0, 0)) corresponds to the average precipitation/day function. The plot also shows that there are several cities with positive scores in the first dimension (e.g. Edmonton, Winnipeg and Scheffervll). The city of Prince Rupert stands out having a large positive score in the second dimension but a large negative score in the first dimension.

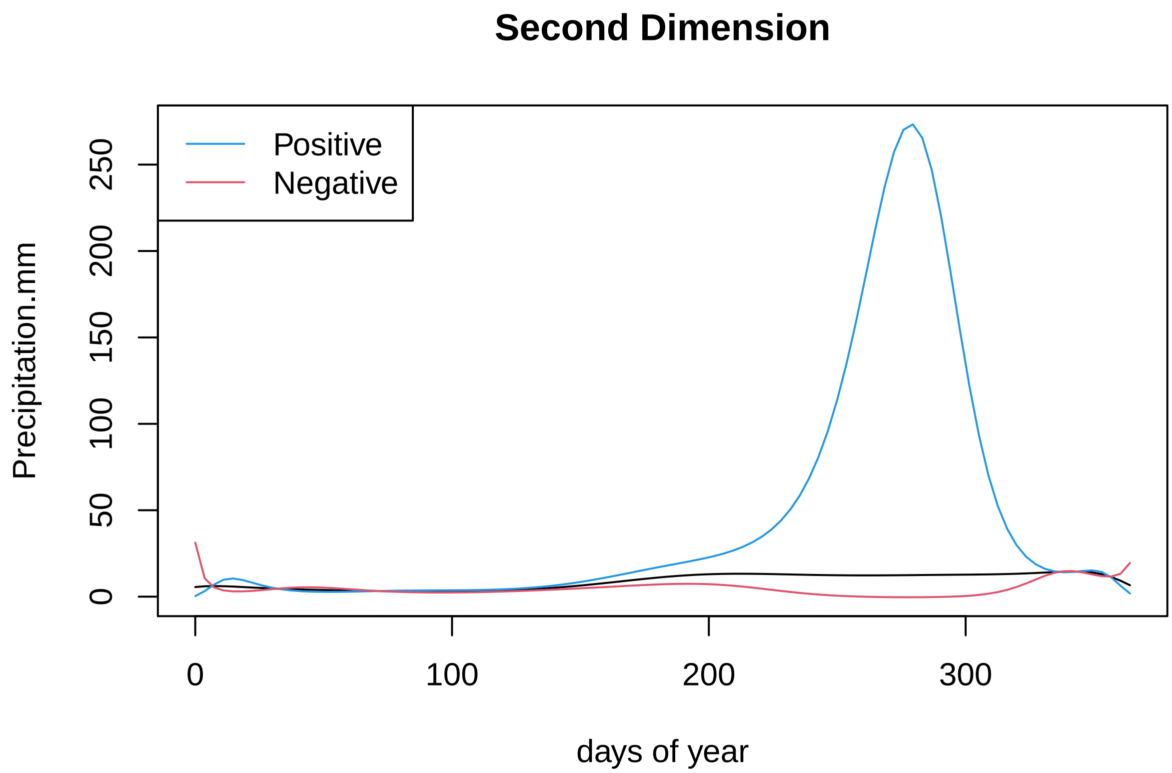
To better understand what large Z_1 or Z_2 scores mean, we will back-transform the SVD to the original function space.



We can conclude that cities that have a large negative score in the first dimension tend to have high precipitation at the end and beginning of the year, in other words, during the winter months. On the other hand, cities with a large positive score have their high precipitation in the middle of the year during the Canadian summer.

Referring back to the functional biplot, we can interpret that particularly Prince Rupert, Vancouver and Sydney have high precipitation during the winter, while the cities Edmonton, Winnipeg, Prince Albert and Regina have high precipitation in the summer.

We can repeat the procedure for the second dimension.



The graph allows us to conclude that cities with large positive scores in the second dimension have high amounts of precipitation in the second half of the year, particularly around autumn time. Cities with large negative scores get slightly less precipitation than average and do not experience high precipitation in the autumn. From the score plot we can see that Prince Rupert has the largest positive score in the second dimension. This isn't surprising as Prince Rupert also had a high score in the first dimension, which corresponded to a very high winter precipitation. High scores in both the first and second dimension point towards the fact that precipitation in Prince Rupert is unusually high at various points during the year as compared to other cities in Canada.