CIVE 650 - Spatiotemporal Data Analysis (2025 Fall)

Assignment 1
Eigen traffic

Date assigned: Sept 17, 2025
Date due: Oct 1, 2025

The `Seattle_traffic_flow.mat` file has two matrices `data` and `mdata`, recording the vehicle count data collected from a sensor on a Seattle highway of year 2015. The 96×365 `data` matrix records the number of vehicles per 15 minutes; each row represents a 15-minute interval in a day and each column represents a day. The `mdata` matrix is the same as `data` but has 30% missing values. For example, the traffic speed of the highway from day 1 to day 7 can be visualized as follows:

```
load("Seattle_traffic_flow.mat")
figure, hold on
for day=1:7
    plot(0:0.25:23.75, data(:, day),...
        'DisplayName',strcat('Day',num2str(day)));
end
xlabel('Time in a day'), ylabel('#Vehicle per 15 minutes')
legend()
```

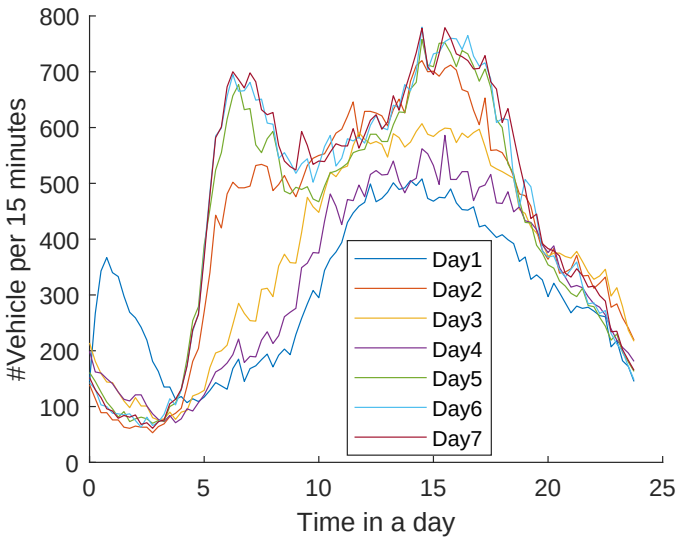

Figure 1: Traffic speed of the first seven days.

1. Perform principle component analysis by SVD to data of the first 360 days (`data(:,1:360)`). Plot the first 4 principal components (first 4 columns of U).

2. Plot all singular values in descending order and the cumulative explained variance curve (similar .

3. We denote by r the rank of truncation in SVD. Plot the data of Day-10 and Day-20 (`data(:,10), data(:,20)`) and its low rank approximation when r=10, 20, 30 and 50.

4. Fit the data of the last five days (`data(:,361:365)`) using $r = 20$. Compare the real and the fitted value.

5. Project data onto the first four PCA components. Plot the bivariate scatter plot for projection on the first 4 components (e.g., `https://plotly.com/python/pca-visualization/#visualize-all-the-principal-components`). You can label data by the day in a week. Describe the patterns in the figure.

6. `mdata` has 30% missing values. Use ALS (`https://www.mathworks.com/help/stats/pca.html#bttvehk-2`) for PCA on `mdata`. Choose $r = 10, 20, 30, 50$, respectively, and reconstruct the traffic speed of Day-10 and Day-20 from `mdata`. Compare your reconstructed data with the real values in `data`.

Note: a good reference to help with this assignment is the Eigenfaces example in Chapter 1 of *Data-driven science and engineering: Machine learning, dynamical systems, and control*, (Brunton & Kutz, 2019). Cambridge University Press.