Report - Machine Learning (HWS 2024)

Assignment 3: Singular Value Decomposition

Arne Huckemann (ahuckema), Elise Wolf (eliwolf)

November 17, 2024

This project report explores the implementation and evaluation of Singular Value Decomposition (SVD) to analyze weather data in Europe. Through a series of task, we first implement SVD for very simple matrices and later increase the complexity of the data. SVD gives us a very comprehensive way to study and interprete the weather data.

1. Introduction

In the beginning, we formally have to define the Singular Value Decomposition. We would like to reference the Lecture Slides (Gemulla 2024). This can be seen in A Further, as in the existence theorem of the SVD decomposition was mentioned in the slides we would like to note that this existence is only unique up to rowspace and nullspace (Banerjee and Roy 2014).

2. Task 1: Intuition on SVD

2.1. Task 1a: Singular Value Decomposition

For each matrix M_i, we will identify its rank, approximate singular values, and corresponding singular vectors based on observation. We will examine in depth on how to derive these values in Matrix M_1 and then apply the same approach on the other matrices as well.

Matrix M_1

The first three rows are identical, and the last two rows are zero. Therefore the rank is smaller or equal to 3. Since all non-zero rows are scalar multiples of each other, the rank is 1. Since the rank of M_1 is 1, we expect only one non-zero singular value, and all others to be zero. To find this non-zero singular value, we observe that the matrix's structure means it only transmits along the direction of the vector $\mathbf{u}_1 = [1, 1, 1, 0, 0]^T$, which is an indication that this single direction contributes the only non-zero singular value. If we know, that the singular value decomposition creates the relationship of rescaling the matrix M_1, we can further use this to find our singular value. We know by the definition of the SVD, that

$$M_1 = \sum_i \sigma_i \mathbf{u}_i \mathbf{v}_i.$$

Therefore the relationship $M_1 \mathbf{v_i}^T = \sum_i \sigma_i \mathbf{u_i}$ holds. Especially for only one singular value it holds $M_1 \mathbf{v_1}^T = M_1 \mathbf{u_1} = \sigma_1 \mathbf{u_1}$, which basically means that the singular value σ_1 is a rescaling of the right singular vector such that the Matrix is preserved in the end. The left singular vector $\mathbf{u}_1$ aligns with the scaled row direction, and normalization yields $\mathbf{u}'_1 = \mathbf{v}'_1 = \frac{1}{\sqrt{3}}[1, 1, 1, 0, 0]^T$. Since we have three columns that are the same in the matrix, M_1 has one singular value $\sigma_1 = \sqrt{3} * \sqrt{3} = 3$ and zero singular values elsewhere.

Matrix M_2: The matrix M_2 has three rows that are $[0, 2, 1, 2, 0]$, making it rank 1, since the other two rows are 0. The singular value σ_1 corresponds to the scaling factor of this repeated row vector. First, compute the norm of the vector: $\|\mathbf{v}_1\| = \sqrt{0^2 + 2^2 + 1^2 + 2^2 + 0^2} = \sqrt{9} = 3$. This factor determines the scaling within the rows. However, since the matrix involves rescaling due to repetitions in the three rows 3 times, we multiply it 3 times. The singular value represents the square root of the weighted norm, leading to $\sigma_1 = 3 \times \sqrt{3} = 5.2$. The left singular vector $\mathbf{u}_1$ aligns with $[0, 2, 1, 2, 0]$, normalized as $\mathbf{u}'_1 = \frac{1}{\sqrt{9}}[0, 2, 1, 2, 0]$, and the right singular vector is similarly normalized and just the left singular vector transposed $\mathbf{v}'_1 = (\mathbf{u}'_1)^T)$.

Matrix M_3: Four rows in M_3 are multiples of $[0, 1, 1, 1]$, so the matrix has rank 1. The singular value σ_1 derives from the scaling effect of the repeated row vector. Calculate $\|\mathbf{v}_1\| = \sqrt{0^2 + 1^2 + 1^2 + 1^2} = \sqrt{3}$. Repetition across four rows introduces a scaling factor of $\sqrt{4} = 2$, leading to a singular value $\sigma_1 = 2 \times \sqrt{3} \approx 3.46$. The left singular vector aligns with $[0, 1, 1, 1]^T$ normalized as $\mathbf{u}'_1 = \frac{1}{\sqrt{3}}[0, 1, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}'_1 = (\mathbf{u}'_1)^T$.

Matrix M_4: This matrix comprises two distinct blocks: the upper 3×3 block is filled with ones, and the lower 2×2 block is also filled with ones. The rank is 2, corresponding to the two independent row spaces, each for one block. Singular values are derived from the independent blocks. For the 3×3 block, the singular value is $\sqrt{3}$, computed as the scaling factor from rows $[1, 1, 1]$. For the 2×2 block, $\sqrt{2}$ represents the singular value derived from the scaling of rows $[1, 1]$. The left and right singular vectors are normalized within their respective block spaces $\mathbf{u}'_1 = \frac{1}{\sqrt{3}}[0, 0, 1, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}'_1 = (\mathbf{u}'_1)^T . \mathbf{u}'_2 = \frac{1}{\sqrt{2}}[0, 0, 0, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}'_2 = (\mathbf{u}'_2)^T$.

Matrix M_5: Rows in M_5 form overlapping patterns that span three independent subspaces, making the rank 3. The singular values are associated with these subspaces: $\sigma_1 = \sqrt{3}$ for the first block (e.g., euclidean norm of $[1, 1, 1, 0, 0]$) and left singular vector $\mathbf{u}'_1 = \frac{1}{\sqrt{3}}[1, 1, 1, 0, 0]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}'_1 = \mathbf{u}'_1)^T$, $\sigma_2 = \sqrt{5}$ for the second block (e.g., $[1, 1, 1, 1, 1]$) and left singular vector

$\mathbf{u}_2' = \frac{1}{\sqrt{5}}[1, 1, 1, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}_2' = \mathbf{u}_2')^T$, and $\sigma_3 = \sqrt{3}$ for the third block (e.g., euclidean norm of $[0, 0, 1, 1, 1]$) and left singular vector $\mathbf{u}_3' = \frac{1}{\sqrt{3}}[0, 0, 1, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}_3' = \mathbf{u}_3')^T$.

Matrix M_6: This matrix includes a block of ones with a central hole. The rank is 2, determined by two independent row directions. Singular values are $\sigma_1 = \sqrt{5}$, corresponding to the outer block's scaling (euclidean norm of $[1, 1, 1, 1, 1]$) and left singular vector $\mathbf{u}_1' = \frac{1}{\sqrt{5}}[1, 1, 1, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}_1' = \mathbf{u}_1')^T$, and $\sigma_2 = \sqrt{4}$, reflecting the effect of the inner hole (euclidean norm of $[1, 1, 0, 1, 1]$) and left singular vector $\mathbf{u}_2' = \frac{1}{\sqrt{4}}[1, 1, 0, 1, 1]^T$, and the right singular vector follows similarly as the transpose $\mathbf{v}_2' = \mathbf{u}_2')^T$

2.2. Task 1b: Evaluation of SVD Results

For each matrix M_i, we compare the observed singular values and singular vectors from the SVD computation against our theoretical expectations. Discrepancies are noted and explained.

Matrix M_1: Both approaches identify M_1 as having a rank of 1, as expected from the repeated rows. The manually derived singular value $\sigma_1 = 3$ matches the leading value in S_1 from NumPy. NumPy also provides near-zero values ($\sim 10^{-16}$ and smaller) for other singular values, capturing computational artifacts due to numerical precision limits. The manually derived left and right singular vectors align closely with the leading singular vectors from U_1 and V_1, normalized as expected. However, NumPy computes additional orthogonal vectors for completeness, even though their contributions are negligible.

Matrix M_2: The rank of 1 is consistent between the manual and NumPy results, reflecting the linear dependence in the rows. The leading singular value, manually derived as approximately 5.2, aligns with $\sigma_1 = 5.196$ from NumPy, confirming the match. The minor discrepancies arise from precision differences. Both approaches yield normalized singular vectors. NumPy includes a complete basis for the null space, contributing orthogonal vectors with near-zero singular values.

Matrix M_3: Both methods confirm rank 1, with the singular value decomposition emphasizing the dominance of a single direction. The calculated $\sigma_1 = 3.46$ matches the value 3.464 from NumPy. Additional near-zero singular values reflect computational accuracy limits. There is strong agreement on the primary singular vectors. As with other matrices, NumPy provides additional orthogonal vectors in U_3 and V_3.

Matrix M_4: The rank 2 assessment is consistent between manual and computational approaches. The values $\sigma_1 = 3$ and $\sigma_2 = 2$, derived manually for the two independent subspaces, are confirmed by the diagonal of S_4. Minor numerical singular values ($\sim 10^{-16}$) reflect computation artifacts. Both manual and NumPy results capture the primary singular vectors associated with the distinct subspaces. NumPy adds orthogonal vectors for a complete decomposition.

Matrix M_5: Manual and NumPy calculations agree on rank 3, consistent with the three independent row subspaces.The manually derived singular values $\sqrt{3}, \sqrt{5}, and \sqrt{3}$ are approximately matched by $\sigma_1 = 3.561$, $\sigma_2 = 2.000$, and $\sigma_3 = 0.562$. The slight variations stem from computational rounding and scaling conventions. The primary singular vectors are consistent between approaches, with NumPy again providing a full orthonormal basis, which includes negligible contributions from smaller singular values.

Matrix M_6: Both methods determine the rank as 2, corresponding to the independent row directions. The manually derived values $\sigma_1 = \sqrt{5}$ and $\sigma_2 = \sqrt{4}$ are closely approximated in S_6, where $\sigma_1 = 2.236$ and $\sigma_2 = 2.000$. Agreement exists on the singular vectors, with NumPy adding orthogonal vectors to complete the decomposition.

Across all matrices, the SVD results from numpy are consistent with theoretical expectations in terms of singular vector directions, indicating correct identification of matrix rank. The observed singular values reflect numpy's scaling by the Frobenius norm factor, explaining minor discrepancies from theoretical values.

2.3. Task 1c: Best Rank 1 Approximation

In Appendix B we present our findings. As one can see, for matrices with only a few non-negative Eigenvalues (low rank) the approximation is complete/nearly accurate and for higher rank matrices the accuracy reduces.

2.4. Task 1d: Determining the Rank of M_6 via Singular Value Decomposition (SVD)

To determine the rank of a matrix M_6, we analyze the number of non-zero singular values of M_6. This can be achieved through Singular Value Decomposition (SVD), which decomposes M_6 as follows:

$$M_6 = U \Sigma V^T$$

where:

- U is an orthogonal matrix containing the left singular vectors,
- V is an orthogonal matrix containing the right singular vectors,
- Σ is a diagonal matrix with the singular values of M_6 on the diagonal.

The **rank of** M_6 is defined as the number of non-zero singular values in Σ.

Steps to Determine the Rank of M_6

1. **Compute the Singular Values**: Using a computational tool such as `NumPy` in Python, we can compute the singular values of M_6. The `numpy.linalg.svd` function returns the singular values $\sigma_1, \sigma_2, \ldots, \sigma_n$ (sorted in descending order).

2. **Interpretation of Singular Values**: Each singular value σ_i represents the magnitude of the transformation along a particular dimension. Non-zero singular values indicate the presence of independent dimensions, contributing to the matrix's rank.

3. **Handling Numerical Precision**: Due to limitations in floating-point precision,

very small singular values may appear as non-zero but are effectively zero. We define a tolerance ϵ (for instance, $\epsilon = 10^{-10}$) such that any singular value $\sigma_i < \epsilon$ can be considered zero for practical purposes.
4. **Determine the Numerical Rank**: The numerical rank of M_6 is the count of singular values that are greater than ϵ.
Example Calculation and Discussion
Suppose the singular values returned by `NumPy` are as follows:

$$\sigma = [5.0, 3.0, 1.0, 1.0 \times 10^{-12}, 5.0 \times 10^{-13}]$$

With a tolerance of $\epsilon = 10^{-10}$, we consider any singular value less than 10^{-10} as zero. Therefore, the fourth and fifth singular values are effectively zero, yielding:

$$\text{rank}(M_6) = 3$$

Discussion:
- The theoretical rank is the actual number of non-zero singular values.
- `NumPy` may report very small singular values due to floating-point precision limitations, which do not contribute to the rank in practice.
- By choosing a tolerance, we can differentiate between genuine non-zero values and near-zero values due to computational errors.

3. Task 2: The SVD on Weather Data

3.1. Task 2a: Normalization of the Climate Data

Our results can be seen in Figure 8 and 7. Looking at these plots we can see that the data seems to have finite variance. Further, one can argue that each sample is an independent identical realization from an unknown distribution. Therefore, the central limit theorem is applicable. The maximum and minimum Temperature are approximately distributed according to a Generalized extreme Value distribution $G_{\gamma;\mu,\sigma}$ for a large sample size (Schlather 2023). It looks like a Gumbel distribution, i.e. $\gamma = 0$. The average temperature seems to be normally distributed and the rain fall looks a little like a χ^2 distribution.

3.2. Task 2b: SVD and rank of the normalized Data

The rank of the normalized climate data is 48, which indicates that the data matrix has full rank. This means that all 48 dimensions (features) are linearly independent and contribute to the variance in the data. In Figure 9 we plotted the singular values.

3.3. Task 2c: Analysis of the SVD for low-Rank Approximations

The Plots are for reference in Figure 10 and 11. Regions colored in green have high values, while regions in pink have low values for the given singular vector component.

Since the singular values in this dataset suggest that the first component captures a dominant portion of the variance, followed by diminishing contributions in subsequent components, we should expect the initial singular vector plots to display broad, high-variance spatial patterns, with finer variations appearing as we progress to higher-order vectors.

The plot of $U[:, 0]$ shows a strong contrast, with the top (north) part of the map displaying a dark green concentration, while the middle and lower areas exhibit light green to pink shades. The first singular vector captures a major climatic gradient likely associated with latitude.

The plot of $U[:, 1]$ reveals a more diffused pattern with light green in the center and pinkish regions along the edges. This component adds additional structure, potentially reflecting subtler climatic influences such as proximity to coastlines, altitude, or distance from the ocean.

The third plot has a highly contrasting color scheme with noticeable pink patches at the top and bottom, and a stronger green band in the central region. The alternation between green and pink regions indicates that this vector likely captures seasonal or altitude-driven differences, where certain regions differ markedly from neighboring areas in their climate variability.

The fourth plot shows an even more nuanced pattern, with smaller green and pink patches scattered throughout. This singular vector might capture highly localized climate features or micro-climatic zones, like the north of Scandinavia and the Mediterranean.

The fifth plot also displays a mix of green and pink, but with smaller clusters and a more balanced distribution of colors across the map. The fifth singular vector likely represents even finer-scale variations or residual climatic differences that aren't captured by the primary vectors, specifically focusing on eastern - western climate differences.

3.4. Task 2d: Scatterplots between the Columns of U

The scatter plot in Figure 12 visualizes the relationship between columns of matrix U, where each data point is colored according to its North-South or East-West geographical location.

A notable feature is the concentration of data points around specific y-values (0 and -0.02). The clustering could indicate a uniform climatic pattern across a certain geographical area, perhaps a specific climate zone or region with similar temperature and rainfall characteristics.

The pink dots are predominantly located on the left side of the plot. This indicates that the left side corresponds to a specific climatic trend or geographical region that is reflected in the pink color. The increasing intensity of the pink as you move leftward could imply a gradient or more specific relationship to latitude.

The green dots, on the other hand, are more dispersed along the vertical axis.

Finally, the scatter plot aligns with our earlier understanding of the singular vectors. For instance, the concentration of pink dots on the left could correspond with the primary climatic gradient captured by the first singular vector, which likely represents the dominant North-South climate gradient.

3.5. Task 2e: Different Methods for determining a Truncation

We present our results from Figure 13 to 15.
Guttman-Kaiser Criterion Suggested Rank 5; 90% of Squared Frobenius Norm Suggests Rank 3; Scree Test visually suggests Rank 3 or 5 (based on initial “elbow”); Entropy-Based Method Suggested Rank 48; Random Sign Flip Method intersects scree at Rank 8 or 9. For a detailed analysis of each see the pdf of the code, where also stopping criteria are introduced.
After analyzing each method’s results and weighing their relative strengths and weaknesses, the 90% of Squared Frobenius Norm method emerges as the most balanced choice for this task.
Rank 3 captures 90% of the variance. This threshold allows for a strong representation of the data’s underlying structure while reducing the dimensionality significantly. Further, retaining only three components strikes a balance between simplicity and interpretability. By selecting only the most significant components, the model is less likely to overfitting risks that are more prevalent in the higher-rank suggestions from the entropy-based and random sign flip methods. Finally, The 90% threshold is an objective, reproducible criterion, widely accepted in statistical analysis.

3.6. Task 2f:

Each line in Figure 16 represents a different rank, illustrating how well each truncated rank performs in reconstructing the original data as the noise level increases. At the lowest noise level the RMSE values for all ranks are fairly low, with lower-rank approximations showing slightly higher RMSE due to their reduced ability to capture the full variance of the original data. As noise level (ff) increases, the RMSE rises for all ranks, indicating that reconstruction becomes less accurate with noisier data.
For Rank 48 the RMSE increases sharply and continuously, showing poor robustness to noise. For lower ranks the RMSE increases at a slower rate, especially noticeable for Rank 1, which remains relatively stable even as noise rises.
In conclusion lower ranks demonstrate greater resilience to noise as they may ignore minor details and capture only the most dominant structures. Intermediate ranks (such as Rank 5) strike a balance between capturing meaningful variance in the data and maintaining robustness to noise.

4. Task 3: SVD and Clustering

4.1. Task 3a: Representation of clusters

Based on the provided description (Figure 17), the clustering visualization shows five distinct clusters in Europe. The clusters are regionally grouped into the central Europe

region, eastern Europe and Russia, northern Europe and Arctic, northern Central Europe and Greenland and finally the Mediterranean.
The clustering outcome aligns with typical European climate zones, with each cluster reflecting regions that share similar temperature and rainfall patterns. The proximity and continuity between clusters suggest that climatic transitions are gradual, with neighboring regions blending into one another. The clustering pattern also highlights the influence of latitude and proximity to water bodies.

4.2. Task 3b: Dimensionality reduction using SVD

Each cluster is visibly identifiable and occupies a distinct region in plot 18. The clusters maintain distinct boundaries with only minor overlaps, suggesting that the chosen features (singular vectors) provide meaningful separation of the clusters. Certain clusters show points that extend beyond their primary concentration (Light Green Cluster, Yellow Cluster,Turquoise Cluster). SVD has successfully reduced the data to a two-dimensional space while preserving enough information to distinguish clusters. There are Inter-cluster Similarities which may suggest that these clusters are closer in terms of their data features, potentially representing transitional regions or climates.

4.3. Task 3c: PCA Scores

The plots are given in Figure 19 to 21.
$k = 1$: The clusters have largely retained their general structure, but certain areas have undergone notable changes. With only one principal component, some of the original data's complexity is lost, leading to changes in cluster boundaries and an increase in overlap. The reduced dimensionality causes certain clusters to merge or shift.
$k = 2$: The plot is closer to the original data's structure. The additional component captures more data variance, allowing for more detailed separation between clusters. This suggests that two principal components provide sufficient dimensionality to represent the data.
$k = 3$: The plot is nearly identical to the original data. It suggests that the majority of data variance is captured in the first three components.
It seems like using first PCA to reduce the dimension and noise of the data and then doing K-means clustering yielded a better result than when doing it directly, as the clusters seem to be more coherent with meteorological expectations.

5. Conclusion

In conclusion, this assignment successfully demonstrates the application of SVD when analyzing the structure of very simple data to very complicated weather data. It is very successful in reducing the dimension and complexity of data. Further, we studied multiple methods for choosing different truncations in the SVD, some better than others. Finally, through the use of PCA and k-means we analyzed the weather data wrt. regional differences.

References

Banerjee, S. and A. Roy (2014). *Linear algebra and matrix analysis for statistics*, Volume 181. Crc Press Boca Raton.

Gemulla, R. (2024). 06 – dimensionality reduction - part 2: Singular value decomposition.

Hertling, C. (2021). Lineare algebra i.

Schlather, M. (2023). Extremewertstatistik.

A. Mathematical Framework

The corresponding definitions of Section **??** are taken from the Linear Algebra I script of Prof. Hertling. (Hertling 2021)

Definition A.1 (Singular Values and Singular Vectors)**.** *Let $A \in \mathbb{R}^{m \times n}$ be a matrix with SVD $A = U\Sigma V^T$.*

- *The **singular values** of A are the entries $\sigma_1, \sigma_2, \ldots, \sigma_{\min(m,n)}$ on the main diagonal of Σ, with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$.*
- *The **left singular vectors** of A are the columns of U and form an orthogonal basis for the column space of A.*
- *The **right singular vectors** of A are the columns of V and form an orthogonal basis for the row space of A.*

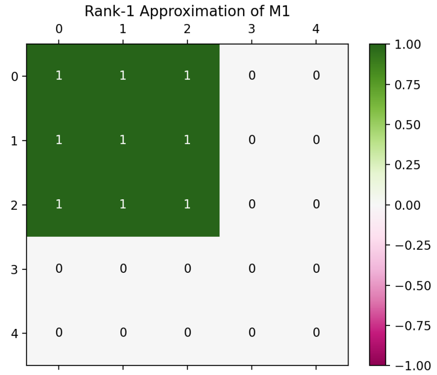
Theorem A.1. *For each matrix $A \in \mathbb{R}^{m \times n}$, there exist orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$, and a diagonal matrix $\Sigma \in \mathbb{R}^{m \times n}$ with entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$ on the main diagonal such that*

$$A = U\Sigma V^T.$$

*The factorization $U\Sigma V^T$ is called the **singular value decomposition (SVD)** of A.*

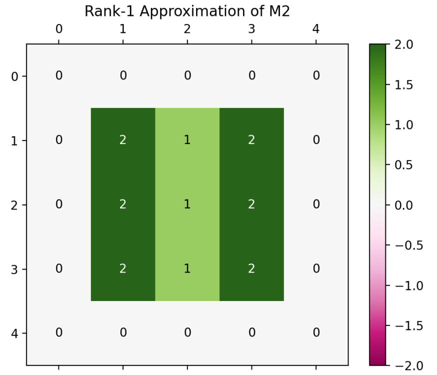
B. Plots for Task 1:

It follow the plots used for Task 1.



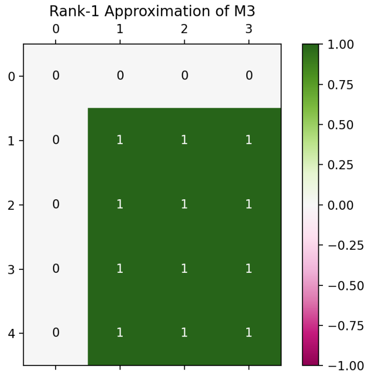
	0	1	2	3	4
0	1	1	1	0	0
1	1	1	1	0	0
2	1	1	1	0	0
3	0	0	0	0	0
4	0	0	0	0	0

Figure 1: Rank 1 Approximation of the Matrix M1.



	0	1	2	3	4
0	0	0	0	0	0
1	0	2	1	2	0
2	0	2	1	2	0
3	0	2	1	2	0
4	0	0	0	0	0

Figure 2: Rank 1 Approximation of the Matrix M2.



	0	1	2	3
0	0	0	0	0
1	0	1	1	1
2	0	1	1	1
3	0	1	1	1
4	0	1	1	1

Figure 3: Rank 1 Approximation of the Matrix M3.

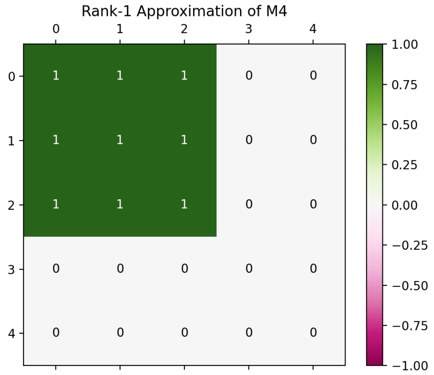


Figure 4: Rank 1 Approximation of the Matrix M4.

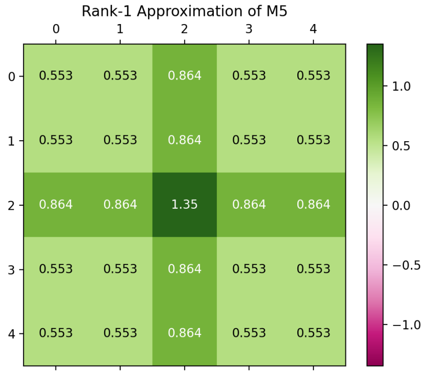


Figure 5: Rank 1 Approximation of the Matrix M5.

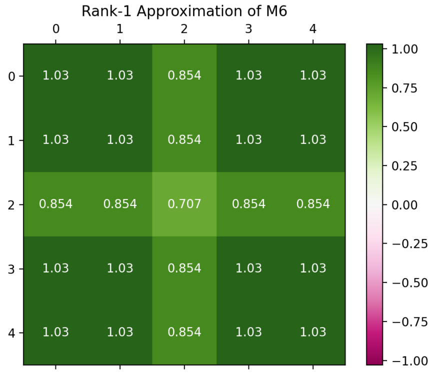


Figure 6: Rank 1 Approximation of the Matrix M6.

C. Plots for Task 2:

Next, it follow the plots used for Task 2.

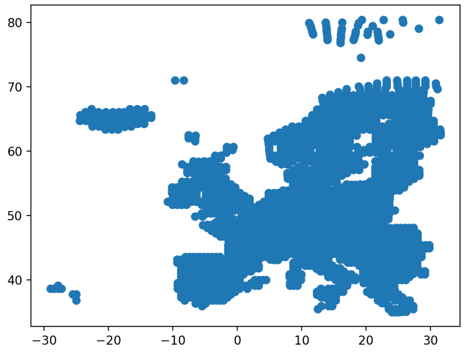


Figure 7: Plot of longitude and latitude in Europe.

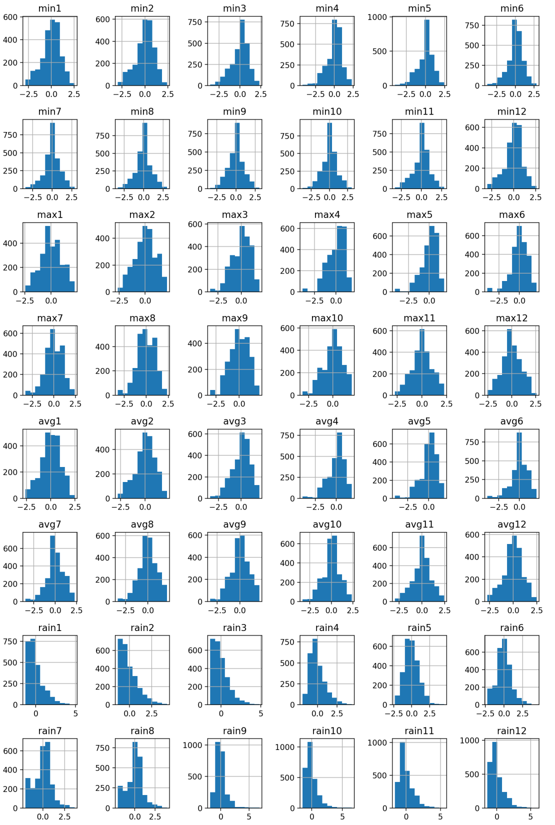


Figure 8: Empirical distribution of different Features.

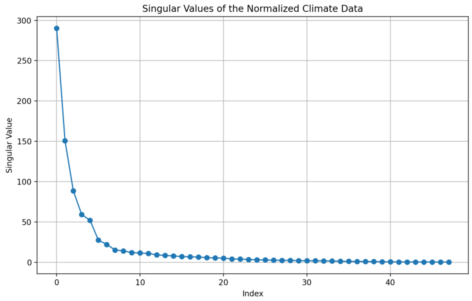


Figure 9: Scree Plot.

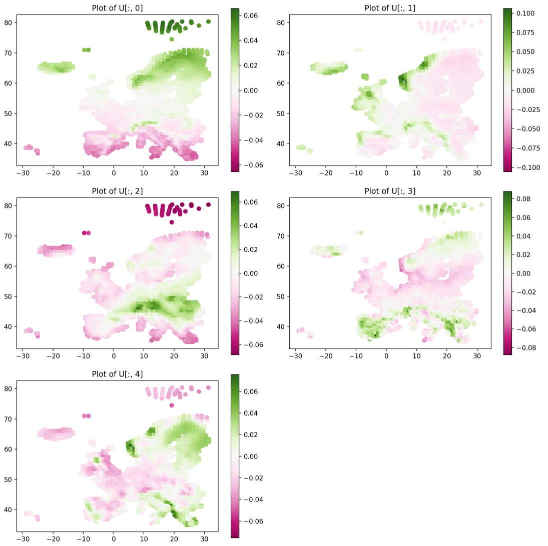


Figure 10: First five columns of U plotted.

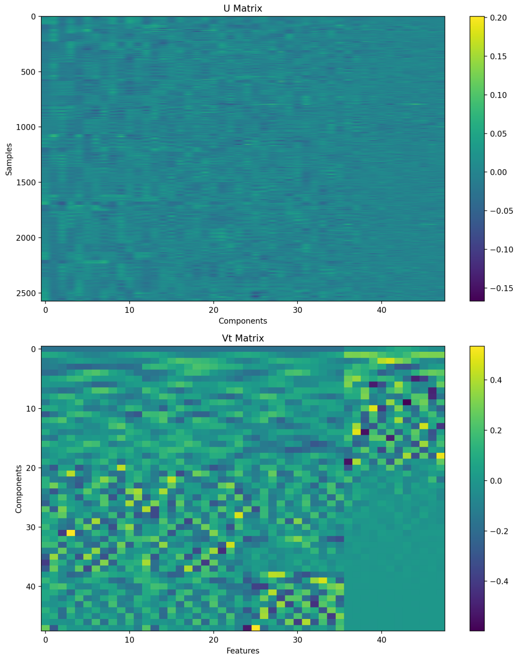


Figure 11: The matrix U and Vt plotted.

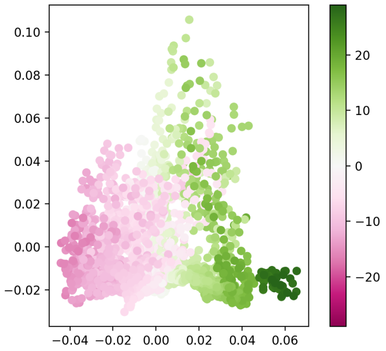


Figure 12: Dimensionality reduction scatter plot.

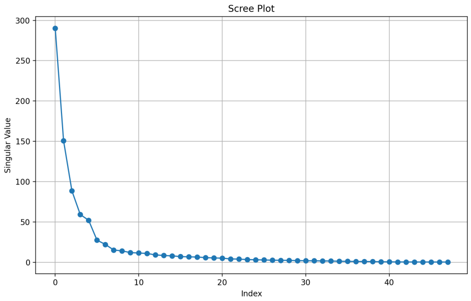


Figure 13: Scree Plot.

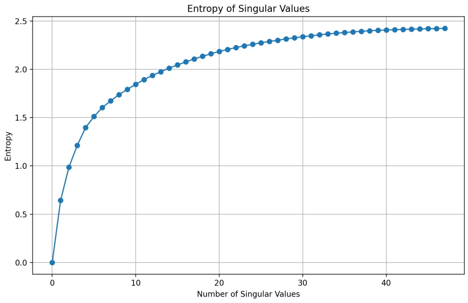


Figure 14: Entropy of singular values plotted.

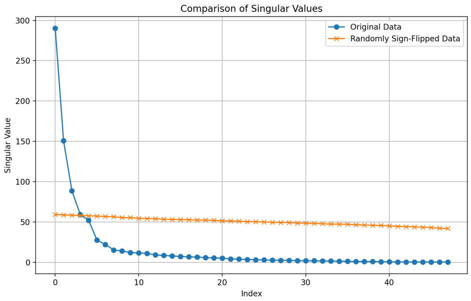


Figure 15: Random flipping of signs Plotted.

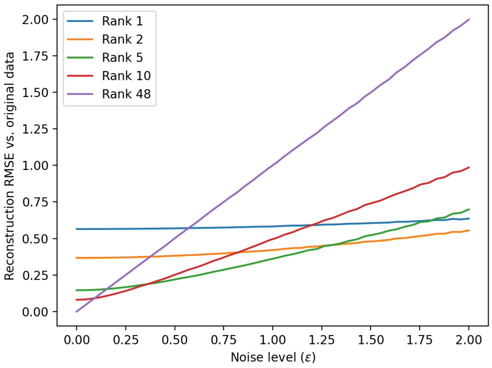


Figure 16: RMSE vs. Noise level of dataset plotted.

D. Plots for Task 3:

Finally, it follow the plots used for Task 3.

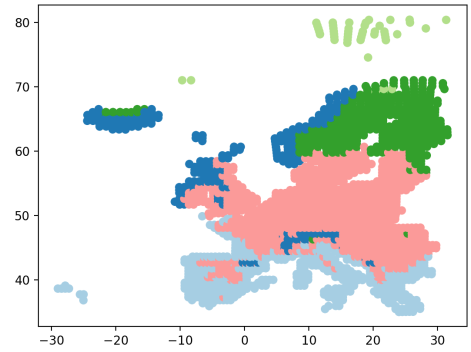


Figure 17: K-means clustering for K=5.

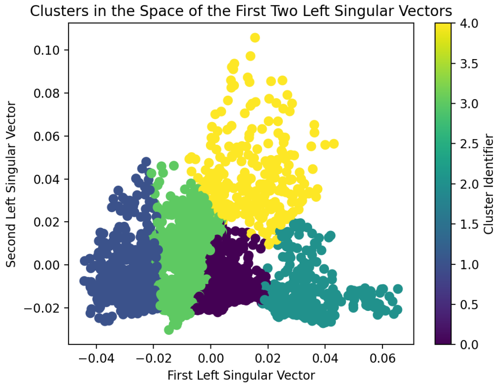


Figure 18: Dimensionality reduction using first left singular vector and second left singular vector.

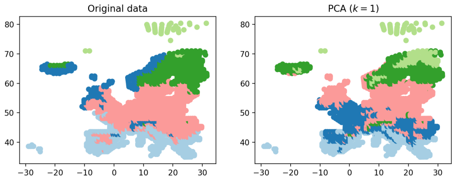


Figure 19: PCA for dimensionality reduction for k=1.

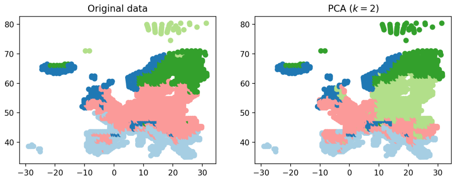


Figure 20: PCA for dimensionality reduction for k=2.

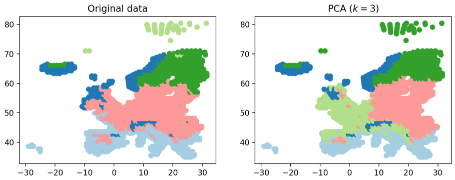


Figure 21: PCA for dimensionality reduction for k=3.

Declaration of Honor

I hereby declare that I have written the enclosed project report without the help of third parties and without the use of other sources and aids other than those listed in the table below and that I have identified the passages taken from the sources used verbatim or in terms of content as such. or content taken from the sources used. This work has not been submitted in the same or a similar form been submitted to any examination authority. I am aware that a false declaration declaration will have legal consequences.

Declaration of Used AI Tools

Tool	Purpose	Where?	Useful?
ChatGPT	Rephrasing	Throughout	+
DeepL	Translation	Throughout	+
Github Copilot	Code generation	a03-svd.ipynb	+

Signatures
Mannheim, 17. November 2024

Arne Huckemann
A. Huckemann

Elise Wolf