Executive Summary for Task 4

October 2, 2025

For Task 4, we have chosen The Real Estate Select Sector SPDR Fund (XLRE) for analysis. The 30 largest holdings of XLRE were identified using data provided by State Street Investment Management1. The data was downloaded in CSV format and is already sorted in descending order. We then record the tickers of the top 30 holdings. Historical data for these holdings, covering approximately six months from March 23, 2025, to September 23, 2025, was obtained from Yahoo Finance for further analysis. To analyze the data, we first calculated and standardized the price return data. Then, we applied Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) to the standardized covariance matrix to understand data further.

The first key finding is, from PCA, we found 74% of total variance will be explained by first 3 PC. While the top ten principal components explain 90% of the total variance, with immaterial contribution for subsequent components. This indicates that the majority of the ETF's systematic risk is driven by a relatively small number of factors. Additionally, both PCA and SVD produce identical eigenvalues, and their eigenvectors are equivalent, differing only in sign, aligning with the theoretical concepts discussed in Module 3, Lecture 4.

0.1 Methodology

We first set up the essential libraries: OS for importing data, yfinance to pull XLRE data, pandas and NumPy for data manipulation, and Seaborn and Matplotlib for plotting. Our goal is to fetch ETF prices, compute returns, and analyze their risk structure. Next we're pulling 6 months of adjusted closing prices for XLRE's 30 largest holdings using Yahoo Finance from period : March 23, 2025, to September 23, 2025 . Forming the basis for calculating daily returns and analyzing the ETF's risk and return. We've converted the raw price series into daily returns, which represent the percentage change in price. This transformation is important because daily returns provide a standardized format to compare the relative performance of different holdings for XLRE. Using raw prices, which vary across holdings due to differing share prices, would affect the results based on the magnitude of the share price. The second transformation is to apply a common statistical technique which transform the variable into a standard scale. It converts a variable so that it has a mean of 0 and a standard deviation of 1. It is important since PCA is a variance maximizing procedure, we want the variable having a same scale to uncover the maximum components and for fair comparison.

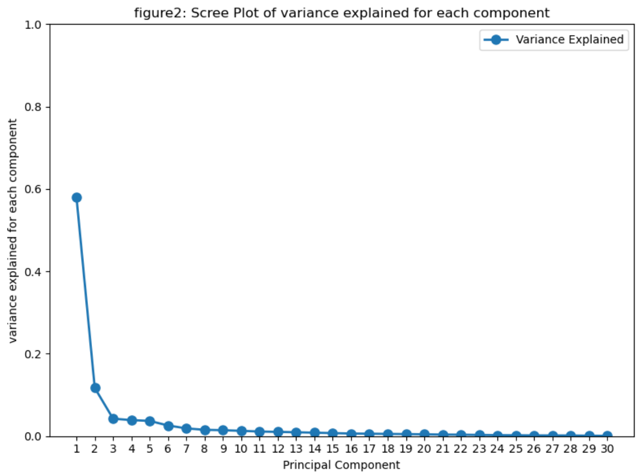

Figure 1:

0.2 PCA Results

	Eigenvalues	Explained	Cumulative Explained
1	17.426312	58.09%	58.09%
2	3.514055	11.71%	69.80%
3	1.275606	4.25%	74.05%
4	1.168617	3.90%	77.95%
5	1.109116	3.70%	81.65%
6	0.779281	2.60%	84.24%
7	0.565794	1.89%	86.13%
8	0.462744	1.54%	87.67%
9	0.434802	1.45%	89.12%
10	0.386216	1.29%	90.41%

74% of total variance will be explained by first 3 PC, if we want to explain 80% of the variance we could select the top 5 PCs. If we want more than 90% of explained variance , we can include top 10 PCs, after that the increment are extremely immaterial.

0.3 SVD results

We perform SVD on standardized data, decomposing it into three components: U, a 127 x 127 orthonormal matrix; S, a 127 x 30 diagonal matrix with positive real numbers on the diagonal; and V a 30 x 30 orthonormal matrix.

The Singular values are :

[4.17448325 1.8745815 1.12942724 1.08102544 1.05314552 0.88276912 0.75219259 0.68025267 0.65939491 0.6214627 0.58067326 0.55895637 0.53315542 0.50283076 0.47109492 0.43491749 0.42084228 0.40438441 0.38847708 0.36812626 0.34160575 0.32937165 0.29648478 0.2689663 0.25305384 0.22984405 0.22607186 0.21473141 0.1801843 0.14362366]

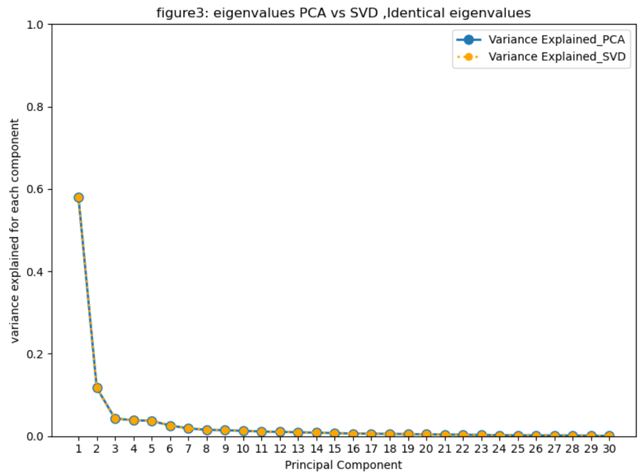

Figure 2:

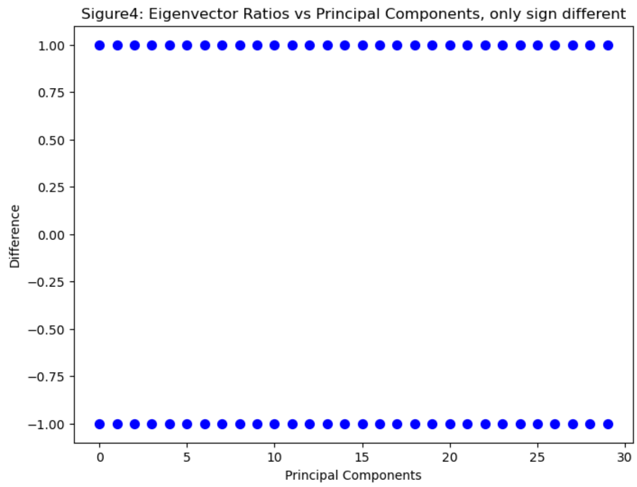

Figure 3:

To show the connection between SVD and principal component analysis (PCA), we follow what discussed in module 3 lesson 4, The squared singular values in matrix are eigen values of covariance matrix. We plot the results in figure 3, which show they are indeed the same. As discussed, the transpose of the V matrix in SVD represents the eigenvectors. As expected, when compared with the eigenvectors from PCA, they differ only by sign. The results are plotted in Figure 4, the plot the ratio for each components of PCs , the magnitude of ratio is always 1 .