Relationship between SVD and PCA

Relationship between SVD and PCA

- Assume A is an m by n centered matrix where rows denoted the samples and columns denote the features.
- Recall that from PCA that we can find covariance matrix by A^TA which is a symmetric matrix and can be orthogonally diagonalized.
- $\boldsymbol{A^T A = PDP^T}$ where columns of P are orthonormal eigenvectors of covariance matrix and D is a diagonal matrix of its eigenvalues.
- Also recall the new coordinates (principal components) will be calculated by finding the projection of A onto the principal directions: $\boldsymbol{Y = AP}$

Relationship between SVD and PCA

- Decompose A into its singular values/vectors: $A = U\Sigma V^T$
- $A^T A = (U\Sigma V^T)^T U\Sigma V^T$
- $= V\Sigma^T U^T U\Sigma V^T$
- $= V\Sigma^T I\Sigma V^T = V\Sigma^T \Sigma V^T$
- Compare this result with the result from pervious slide:
- $\boldsymbol{A^T A = PDP^T}$= $V\Sigma^T \Sigma V^T$ =>
- You can conclude that **P=V** and hence **columns of V are the principal directions.**
- Also you can conclude $\boldsymbol{D = \Sigma^T \Sigma}$ and hence **singular values are related to the eigenvalues of covariance matrix via $\lambda_i = \sigma_i^2$**

Relationship between SVD and PCA

- Right multiply both sides of $A = U\Sigma V^T$ by V:
- $\boldsymbol{AV = U\Sigma}$
- Compare this with previous PCA derivation $\boldsymbol{Y = AP}$
- We can conclude that $\boldsymbol{Y = U\Sigma}$ which is to say that **columns of $\boldsymbol{U\Sigma}$ are the coordinates in the new space. (a.k.a. scores).**
- Note that our derivation was based on the assumption that rows of A denote samples and columns denote features. If rows and columns of A are switched (like in your textbook), then the interpretation of U and V will also be switched.

Summary

- **Eigenvalues of the covariance matrix** are the same as the **square of the singular values of the (centered) data matrix.** The eigenvalues of the raw data matrix itself don't have this direct relationship with the singular values.
- This relationship is fundamental to understanding how PCA and SVD are connected. PCA relies on finding the eigenvectors of the covariance matrix. These eigenvectors correspond to the principal components, which are the directions of maximum variance in the data. Since the eigenvalues of the covariance matrix are related to the singular values of the data matrix, SVD can be used to efficiently perform PCA.