Singular value decomposition (SVD) and spectral decomposition

1. Definition of SVD

Let X be $n \times p$ matrix. The SVD of X is given as $X = U\Lambda^{1/2} V^T$, where U is a $n \times n$ orthonormal matrix, $\Lambda^{1/2}$ is a $n \times p$ diagonal matrix and V is a $p \times p$ orthonormal matrix.

2. The column vectors of U are the eigenvectors of XX^T. The column vectors of V are the eigenvectors of X^TX. The nonzero diagonal elements of $\Lambda^{1/2}$ are the square root of the eigenvalues of either XX^T or X^TX (In fact, we can show that the eigenvalues of XX^T and X^TX are the same).

3. Spectral decomposition from the SVD

Let $\Sigma = X^TX$. Then, the SVD implies that $\Sigma = V\Lambda V^T$, where $\Lambda = \Lambda^{1/2}\Lambda^{1/2}$. For a symmetric matrix Σ, Σ can be decomposed as the multiplication of $\Sigma = X^TX$

4. Geometrical meaning of SVD.

In this section, for any $n \times p$ matrix M, M_k is the $n \times k$ matrix whose jth column is equivalent to the jth column of the matrix M for $1 \le j \le k$.

Suppose a $n \times p$ matrix X is given and the SVD of X is $X = U\Lambda^{1/2} V^T$. Consider $L(A) = |X - AV^T|_2^2$ for $n \times p$ matrix A, where $|A|_2^2$ is the sum of the squares of all elements of a given matrix A. Then, the SVD imples that $\mathrm{argmin}_A L(A) = U\Lambda^{1/2}$.

For $1 \le k \le p$, let $L_k(A) = |X - AV_k^T|_2^2$, and let $\widehat{A}_k = \mathrm{argmin}_A |X - AV_k^T|_2^2$. Then, we have $\hat{A} = (U\Lambda^{1/2})_k$. That is, the column vectors of $\widehat{A}_k$ are $\lambda_{(j)}^{1/2} U^j$, where U^j is the j th column vector of U. Note that $\widehat{A}_k$ is the matrix of PC loadings when only k many PCs are used to explain X.

That is, the SVD is pursuing the projection of each row vector of X (each instance) to the span space of the column vectors of V_k. The corresponding coefficients are the corresponding row vector of $\hat{A}$.

Now, the question is what happens if we project the rows of X onto a k -dimensional linear sub-space other than the span space of the column vectors of

V_k. Suppose G is the $n \times k$ matrix whose column vectors span a k-dimensional linear subspace other than the span space of the column vectors of V_k. Let $L_G(A) = \|X - AG^T\|_2^2$. Then, it can be proven that $\mathrm{argmin}_A L_{V_k}(A) \leqq \mathrm{argmin}_A L_G(A)$ for any matrix G. That is, the span space of the eigenvectors is the most efficient space to explain X.

An application of the SVD is to compress given data into the corresponding eigenvectors and PC loadings.