(ENKEMNA0302) Applied Linear Algebra

Dr. Gabor FACSKO, PhD
Senior Research Fellow
facskog@gamma.ttk.pte.hu

University of Pecs, Faculty of Sciences, Institute of Mathematics and Informatics, 7624 Pecs, Ifjusag utja 6.
Wigner Research Centre for Physics, Department of Space Physics and Space Technology, 1121 Budapest, Konkoly-Thege Miklos ut 29-33.
https://facsko.ttk.pte.hu

April 17, 2025

Schedule I

- **Classes: April 17, 30, and May 7, 2025.**
- Szinguláris érték, szinguláris vektor, SVD, PCA.
- ~~Mátrixok összehasonlítása, pozitív mátrixok, nemnegatív mátrixok, irreducibilis mátrixok, SMRC, NMF.~~
- Reakcióegyenletek sztöchiometrikus rendezése.
- ~~Lineáris programozási feladatok mátrixaritmetikai megoldhatósága. (MLF?)~~
- ~~Power of matrices. Applications: linear recursions, power of incidence matrixes.~~
- ~~Gram-Schmidt ortogonalization.~~
- ~~Further applications.~~

Singular Value, Singular Vector, SVD I

- (Singular value) A generalization of the orthogonal diagonalization of symmetric matrices to arbitrary matrices by finding two orthonormal bases instead of one.
- Used for data compression and solving systems of equations.
- A generalization of orthogonal diagonalization, where singular values take the place of eigenvalues, and the spectral decomposition is replaced by the singular value decomposition (SVD).
- The matrix mapping corresponding to a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a one-to-one correspondence between the row space and the column space.
- In these two subspaces, we search for orthonormal bases in which the matrix representation of the mapping becomes diagonal.
- These are then extended to orthonormal bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.
- The same construction can be carried out in complex vector spaces.

Singular Value, Singular Vector, SVD II

- For the matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, we seek orthonormal bases $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \subset \mathbb{R}^n$ and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\} \subset \mathbb{R}^m$ such that the matrix $\mathbf{A}$ becomes diagonal in these bases. That is, there exist real values σ_i such that $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$ where $1 \leq i \leq \min(m, n)$. We are looking for mutually orthogonal vectors whose images under the transformation are also mutually orthogonal.
- Proposition: If the orthogonal vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ include at least one eigenvector of $\mathbf{A}^T\mathbf{A}$, then the vectors $\mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \in \mathbb{R}^m$ are also orthogonal. Proof: See Wettl notes.
- The mapping $A : \mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ is a one-to-one correspondence between the row space and the column space. Thus, we first seek appropriate bases only in these two subspaces. Their common dimension is equal to the rank, denoted by r.

Singular Value, Singular Vector, SVD III

- Since $\mathbf{A}^T\mathbf{A}$ is symmetric and positive semi-definite, it is orthogonally diagonalizable and its eigenvalues are non-negative. The eigenvectors span an orthonormal basis of $\mathbb{R}^n$, among which those corresponding to positive eigenvalues span the orthogonal complement of the null space, i.e. the row space, since $\mathcal{N}\left(\mathbf{A}^T\mathbf{A}\right) = \mathcal{N}\left(\mathbf{A}\right)$.
- Let this basis of the row space be denoted by $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$. None of these vectors belong to the null space, and their images under $\mathbf{A}$ are pairwise orthogonal, so the vectors $\mathbf{A}\mathbf{v}_i$ form an orthogonal basis of the column space. If $\mathbf{v}_i$ is a unit eigenvector of $\mathbf{A}^T\mathbf{A}$ with eigenvalue $\lambda_i > 0$, then $|\mathbf{A}\mathbf{v}_i| = \lambda_i$, since

$$|\mathbf{A}\mathbf{v}_i|^2 = \mathbf{v}_i^T\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \lambda_i\,|\mathbf{v}_i|^2 = \lambda_i.$$

  Let $\sigma_i = \sqrt{\lambda_i}$ and $\mathbf{u}_i = \mathbf{A}\mathbf{v}_i/\sigma_i$. Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ form an orthonormal basis of the column space.

Singular Value, Singular Vector, SVD IV

- Definition: (Singular value). For a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ of rank r, the square roots of the positive eigenvalues of the matrix $\mathbf{A}^T\mathbf{A}$ are called the singular values of $\mathbf{A}$, denoted by $\sigma_1, \sigma_2, \ldots, \sigma_r$. A unit eigenvector $\mathbf{v}$ of $\mathbf{A}^T\mathbf{A}$ corresponding to eigenvalue σ^2 is called a right singular vector of $\mathbf{A}$, and the unit vector $\mathbf{u} = \frac{1}{\sigma}\mathbf{Av}$ is called the left singular vector corresponding to $\mathbf{v}$.
- Singular values are particularly useful, for instance, when solving systems of equations where the coefficient matrix is singular.
- An important consequence of this definition is that since $|\mathbf{u}_i| = 1$ for $i = 1, 2, \ldots, r$, we have $|\mathbf{Av}_i| = \sigma_i$.
- If σ is a singular value of $\mathbf{A}$ with multiplicity t, then σ^2 is an eigenvalue of $\mathbf{A}^T\mathbf{A}$ with algebraic (and geometric) multiplicity t, so the right singular vectors corresponding to σ span a t-dimensional subspace of the row space, and the corresponding left singular vectors span a t-dimensional subspace of the column space.

Singular Value, Singular Vector, SVD V

- If $\mathbf{Av} = \sigma \mathbf{u}$, then $\mathbf{A}^T\mathbf{u} = \frac{1}{\sigma}\mathbf{A}^T\mathbf{Av} = \frac{1}{\sigma}\sigma^2\mathbf{v} = \sigma\mathbf{v}$, hence
$$\mathbf{Av} = \sigma\mathbf{u},\ \ \mathbf{A}^T\mathbf{u} = \sigma\mathbf{v}$$
form a pair of equations connecting the right and left singular vectors.
- If $\mathbf{A} \in \mathbb{C}^{n\times n}$, then the row space is replaced with its conjugate, i.e., the subspace $\mathcal{O}\left(\mathbf{A}^H\right)$, and transposition is replaced with Hermitian conjugation, so instead of $\mathbf{A}^T\mathbf{A}$, the matrix $\mathbf{A}^H\mathbf{A}$ is used.
- Wettl notes Example 10.3.
- (Singular decomposition) The singular values and vectors of a matrix define a matrix factorization called the singular value decomposition.

Singular Value, Singular Vector, SVD VI

- Form the diagonal matrix

$$\Sigma_1 = diag\,(\sigma_1, \ldots, \sigma_r) = \begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \end{pmatrix}$$

from the singular values, and define matrices $\mathbf{U}_1 = \{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ and $\mathbf{V}_1 = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ from the singular vectors. Then the equations $\mathbf{A}\mathbf{v}_i = \sigma\mathbf{u}_i$ become

$$\mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\boldsymbol{\Sigma}_1,$$

or

Singular Value, Singular Vector, SVD VII

$$\mathbf{A}\left(\mathbf{v}_1\mathbf{v}_2\ldots\mathbf{v}_r\right) = \left(\mathbf{u}_1\mathbf{u}_2\ldots\mathbf{u}_r\right)\begin{pmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \end{pmatrix}$$

Since $\mathbf{V}_1$ is semi-orthogonal and its column vectors form an ONB of the row space, $\mathbf{V}_1\mathbf{V}_1^T$ is the projection matrix onto the row space, so $\mathbf{A}\mathbf{V}_1\mathbf{V}_1^T = \mathbf{A}$, and by multiplying the equation above on the right by $\mathbf{V}_1^T$, we get $\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T$.

- Definition: (Reduced singular value decomposition and dyadic form). The decomposition of the real (complex) matrix $\mathbf{A}$ as

$$\mathbf{A} = \mathbf{U}_1\boldsymbol{\Sigma}_1\mathbf{V}_1^T \quad \left(\mathbf{A} = \mathbf{U}_1\boldsymbol{\Sigma}_1\mathbf{V}_1^H\right)$$

is called the reduced singular value decomposition, where Σ_1 is a square diagonal

Singular Value, Singular Vector, SVD VIII

matrix with positive real entries in monotonically decreasing order on the main diagonal, and $\mathbf{U}_1$ and $\mathbf{V}_1$ are semi-orthogonal (in the complex case: $\mathbf{U}_1^H\mathbf{U}_1 = \mathbf{V}_1^H\mathbf{V}_1 = \mathbf{E}_r$). Writing this product in block form with column vectors of $\mathbf{U}_1$ and row vectors of $\mathbf{V}_1^T$ yields the dyadic decomposition of $\mathbf{A}$, called the dyadic form of singular value decomposition:

$$\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T\,.$$

(In the complex case, replace transpose with Hermitian conjugate.)

⋮

Singular Value, Singular Vector, SVD IX

- Definition: (Singular value decomposition). The decomposition of the real (complex) matrix $\mathbf{A}$ as

$$\mathbf{A} = \mathbf{U\Sigma V}^T \quad \left(\mathbf{A} = \mathbf{U\Sigma V}^H\right)$$

  is called the singular value decomposition, or simply SVD, if $\mathbf{U}$ and $\mathbf{V}$ are orthogonal (unitary), and Σ is a diagonal matrix with non-negative real numbers in decreasing order along the main diagonal.
- Wettl notes: Example 10.6.
- Theorem: (Existence and uniqueness of SVD). Every real or complex matrix has a singular value decomposition. The sequence of singular values in decreasing order is unique, but the decomposition itself is not.
  Proof: See Wettl notes.

Singular Value, Singular Vector, SVD X

- ~~Geometric interpretation of SVD~~
- ~~Image of the unit sphere~~
- ~~Polar decomposition~~
- ~~Pseudoinverse~~

Principal Component Analysis I

- Principal Component Analysis (PCA) is a multivariate statistical procedure that falls under data reduction methods.
- Its essence lies in reducing the dimensions of a large dataset — whose variables are interrelated — while retaining as much of the existing variance as possible.
- This is achieved by applying an orthogonal transformation that converts the possibly correlated variables of the dataset into a set of linearly uncorrelated variables, which are called principal components.
- The number of principal components is less than or equal to the number of original variables.

Principal Component Analysis II

- The transformation is defined in such a way that the first principal component has the largest possible variance, and each succeeding component has the highest possible variance under the condition that it is orthogonal (i.e., uncorrelated) to the preceding components.
- The principal components are orthogonal because they are the eigenvectors of the covariance matrix, which is symmetric.
- Principal component analysis is sensitive to the relative scaling of the original variables.
- PCA can also be understood as fitting an n-dimensional ellipsoid to the data, where each axis of the ellipsoid represents a principal component. If one of the ellipsoid's axes is small, the variance along that axis is also small, and if we omit that axis and its corresponding principal component from the data representation, we only lose a proportionally small amount of information.

Principal Component Analysis III

- To find the axes of the ellipsoid, we first subtract the mean of each variable from the dataset to center the data around the origin. Then, we compute the covariance matrix of the data, along with its eigenvalues and the corresponding eigenvectors. Next, we orthogonalize and normalize the set of eigenvectors (ONB) to obtain unit vectors. These mutually orthogonal unit vectors can then be considered as the axes of the ellipsoid fitted to the data. The proportion of variance represented by each eigenvector can be calculated by dividing its corresponding eigenvalue by the sum of all eigenvalues.
- This procedure is sensitive to data scaling, and there is no consensus on how to optimally scale the data.

Principal Component Analysis IV

- The steps of Principal Component Analysis:
  1. Input the data and create the dataset. Let x and y be the variables.
  2. Subtract the means from each data dimension, i.e., subtract the mean $\overline{X}$ from each x value, and the mean $\overline{Y}$ from each y value.
  3. Compute the covariance matrix using the above formula. In this case, the covariance matrix will be 2×2.
  4. Calculate the eigenvectors and eigenvalues of the covariance matrix. In PCA, it is important that the eigenvectors are unit vectors (of length 1). The eigenvectors must be orthogonal to each other.
  5. Select the components and form the feature vector. Sort the eigenvectors by their eigenvalues from largest to smallest to obtain the components in order of significance. If the eigenvalues are sufficiently small, the amount of data lost is proportionally small. Then form the feature vector from the retained eigenvectors.

Principal Component Analysis V

6. Derive the new dataset. This is done by transposing the feature vector obtained in the previous step — converting it from a column vector to a row vector, with the most significant eigenvector as the first value. The resulting vector is then multiplied from the right with the transpose of the original dataset.

- Covariance: indicates the relationship between two variables.
- Covariance matrix: for a 3D vector, it is a 3×3 matrix.

The End

Thank you for your attention!