Atlantis Studies in Uncertainty Modelling, volume 3

Joint Proceedings of the 19th World Congress of the International Fuzzy Systems Association (IFSA), the 12th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), and the 11th International Summer School on Aggregation Operators (AGOP)

Laplacian Singular Values

Jiří Janeček[a] and **Irina Perfilieva**[b]

[a]Department of Mathematics, Faculty of Science, University of Ostrava,
30. dubna 22, 701 03 Ostrava, `Jiri.Janecek@osu.cz`

[b]Institute for Research and Applications of Fuzzy Modeling, University of Ostrava,
30. dubna 22, 701 03 Ostrava, `Irina.Perfilieva@osu.cz`

Abstract

In this contribution, we focus on extending the Laplacian processing used in data-driven dimensionality reduction based on weighted graphs by incorporating the concept of singular value decomposition. We indicate a novel point of view on generalized eigenvalue problem by pointing out geometric meaning of factorization matrices. We demonstrate that classical eigenvalue problem of normalized Laplacian, generalized eigenvalue problem of pure Laplacian and singular value decomposition of specific altered Laplacian form are mutually equivalent problems and discuss some of its theoretical implications.

Keywords: Singular value decomposition, Dimensionality reduction, Weighted graph, Generalized eigenvalue problem, Data processing.

Introduction

Our long-term research [3, 4] (which we follow and extend in this paper) is focused on dimensionality reduction (DR) based on analysis of Laplacian matrix of a weighted graph used in signal (and especially image) processing. Laplacian matrix is determined by a weighted adjacency matrix and derived degree matrix. These matrices describe the geometric structure of the data (in the other words, characterize the space we work with in terms of local closeness) and the importance of all data points (points closely connected with other points are represented by vertices of higher degree than less important points, e.g. outliers). The result of such DR is determined by generalized eigenvectors that correspond to smallest generalized eigenvalues — these values describe the reversed importance of eigenvector directions in which each transformed object is decomposed as the eigenvectors play the role of an alternative, data-driven space basis.

The motivation of this paper is to enrich the set of methods we use in Laplacian data processing by incorporating the singular value decomposition (SVD) to eventually discover a new perspective of the original problem (i.e. the problem of DR) and to embed our research into a broader scope. We would like to show that SVD is different but at the same time similar approach (similarity comes from special structure and square size of the Laplacian matrix) to graph-based DR and hence its procedures and results are also applicable in the case of our Laplacian formalism.

SVD is widely used in matrix factorization computations as it is able to discover "hidden features" of the input data matrix A. Assuming that each row of A represents one object (e.g. one physical body or one person) and each column represents a noisy registered value of one variable (e.g. a percentage of water content), then factorizing A into three matrices as described below can reveal for which objects one particular feature is present in a significant degree, which variables strongly correspond to this feature and how much prevalent the representation of this feature in the input data is (weakly manifesting features are considered to be a type of noise). If we look at the initial matrix as at a linear transformation, then it can be also interpreted as a composition of three transformations: rotation (or reflection) of objects, their scaling (and eventually dropping or increasing the number of coordinates) and rotation (or reflection) of variables to fit the initial ones, each of which corresponds to one of the constituent matrices of the SVD factorization of the initial matrix. This enriched interpretation creates an interesting topic as it makes a connection between our well-established Laplacian formalism and the ubiquitous algebraic and geometric operations, in the other words, it expands our point of view on DR.

Other usages and interpretations of SVD include, e.g., solving systems of linear algebraic equations (the matrix of the system need not be square — we can use SVD to construct its pseudoinverse), finding data correlations, building linear regression systems, image processing (compression, noise filtering), determining the rank and range of a matrix and, of course, DR (for each object, it finds its lower-dimensional representation in a feature space of pre-selected dimension, e.g. as in PCA). One of the main advantages of SVD is the fact that any complex matrix (symmetric, non-symmetric, square, rectangular, ...) can be decomposed by this method (on the pure theoretical level, the numerical stability is not of our interest but there are techniques used in computer computations to reduce the round-off errors caused by computing the SVD which are out of the scope of our current research). SVD therefore generalizes the utilization of spectral decomposition which is restricted to only square and at the same time diagonalizable matrices (their eigenvalues also indicate the scaling factors of the linear transformations that they represent, this time along the direction of eigenvectors that appear on both sides of the factorization which disables the distinction between projection vectors used in transformation composition which is usually considered as an advantage but at the same time this is the reason why it cannot be applied to non-square matrices and hence an opportunity for SVD arises). To sum up, the decomposition possibilities show the usefulness in applied problems.

1 Weighted Graph Induced Laplacian

Let us assume that the data (set of objects of one arbitrary nature) is given by

$$X = \{x_i \,|\, i = 1, \ldots, n\},$$

indexed by

$$I = \{1, \ldots, n\},$$

and that the closeness between any two objects x_i and x_j in X is described by $w(i,j)$ where

$$w \colon I \times I \to \mathbb{R}$$

is a non-negative, symmetric function. For the further purpose of this paper, it is assumed that the closer objects have higher closeness value (weight) than the less close ones.

If we introduce a weighted graph $G(V, E, W)$ where

$$V = I, \quad E = V \times V$$

and where weights of the edges $e_{ij} = (i,j)$ are stored in the adjacency matrix $W \in \mathbb{R}^{n \times n}$ and given by

$$w_{ij} = w(i,j),$$

then it accurately describes the data geometry. We assume that W determines a connected graph. Following the method proposed in [1], we can derive a diagonal matrix $D \in \mathbb{R}^{n \times n}$ with diagonal entries given by

$$d_{ii} = \sum_{j=1}^{n} w_{ij},$$

that describes degrees of all vertices. Laplacian matrix $L \in \mathbb{R}^{n \times n}$ of the graph G is defined as

$$L = D - W.$$

Let us note that the literature is inconsistent in the sign of Laplacians meaning that the definition $L = W - D$ is also sometimes used. Laplacian matrix L is symmetric as both of its constituent matrices, W and D, are. Moreover, it is positive semi-definite so that for any vector $\vec{y} \in \mathbb{R}^n$, it holds:

$$\vec{y}^{\top} L \vec{y} = \frac{1}{2} \sum_{i,j=1}^{n} (y_i - y_j)^2 w_{ij} \geq 0,$$

where $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$. Laplacian matrix L is in the centre of our interest as well as its generalized eigenvectors $\vec{y}$, i.e. the solutions of the generalized eigenvalue problem

$$L \vec{y} = \lambda D \vec{y}, \tag{1}$$

that are used to determine a real, lower-dimensional representation of the original data.

2 Singular Value Decomposition

A classical result (formalized e.g. in [2]) states that any real, $m \times n$ matrix A can be decomposed as:

$$A = U S V^{\top},$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices, i.e.

$$U^{\top} U = U U^{\top} = I, \quad V^{\top} V = V V^{\top} = I,$$

and S is an $m \times n$ diagonal matrix (i.e. $\forall i \neq j \colon s_{ij} = 0$) with $p = \min\{m, n\}$ diagonal entries $\sigma_1, \ldots, \sigma_p$ called the singular values of A that are usually ordered:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0.$$

(Each of these values corresponds to one of the hidden features we talked about in the Introduction — strong features are accompanied by bigger singular values while small, non-zero singular values are results of a data noise.) To emphasize that the singular value σ_i

belongs to the matrix A, we write $\sigma_i(A)$. Columns of the matrix U, $\vec{u}_j$, $j = 1,\ldots,m$, are called left singular vectors (and absolute values of its elements reveal which objects represent a corresponding hidden feature), columns of the matrix V, $\vec{v}_j$, $j = 1,\ldots,n$, are right singular vectors (absolute values of its elements tell us which variables support a particular feature).

If $m \leq n$, it holds that

$$A = \sum_{i=1}^{m} \sigma_i \vec{u}_i \vec{v}_i^\top$$

(in the case that $\mathrm{rank}(A) < m = p$, then for $i \geq \mathrm{rank}(A)$, σ_i will be zero), in the other words, the i-th row ($i = 1,\ldots,m$) of A can be expressed as

$$\sum_{r=1}^{m} \sigma_r u_{ir} \vec{v}_r^\top \,.$$

Since $A^\top A = V S^\top S V^\top$ is a diagonalizable matrix, each right singular vector $\vec{v}_j$ of A is an eigenvector of $A^\top A$ and hence

$$\mu_i(A^\top A) = \sigma_i^2(A) \,.$$

Recall that the eigenvalue problem has the form

$$A\vec{z} = \mu\vec{z} \,.$$

Each complex number μ_i for which this equation holds (i.e. exists a non-zero vector $\vec{z}$), is called the (classical) eigenvalue of A and is denoted by $\mu_i(A)$. Eigenvalues of symmetric matrices are real. Analogously it holds for left singular vectors. It shows the relation between singular values and eigenvalues — singular value of A is the square root of the eigenvalue of the iterated matrix A (i.e. of $A^\top A$ or $AA^\top$).

3 Laplacian SVD

As the Laplacian matrix L is real, its SVD is formed by two orthogonal matrices as described above but in general, these two matrices, for a genereal complex matrix, are unitary (so in our case $U^\top = U^*$ and $V^\top = V^*$). Because L is symmetric, it moreover holds that

$$L = L^\top = \overline{L^\top} = L^* \,,$$

which implies that

$$L^* L = LL^* \,,$$

and hence Laplacian matrix is normal.

Normality of L implies that U and V are unitary matrices that diagonalize L. And since L is a positive semi-definite matrix, its SVD coincides with its eigende-composition, i.e. if we rewrite L as $QTQ^\top$ (with Q orthogonal and T diagonal), then for classical eigenvectors and eigenvalues of L (solving $L\vec{z} = \mu\vec{z}$), it holds:

$$q_{\cdot,i} = z_{i,\cdot} \quad \text{and} \quad t_{ii} = \mu_i \,;$$

and we see that

$$U = V = Q, \quad S = T \quad \text{and} \quad \sigma_i(L) = \mu_i(L) \,.$$

Even though we showed this useful coincidence, the spectral decomposition itself is not of our interest as the classical eigenvalue is not used in DR.

Let us recall that the value of λ which are we interested in, is the generalized eigenvalue of the Laplacian matrix L as it (together with corresponding generalized eigenvectors) solves the equation (1).

It is also a classical eigenvalue of $D^{-1}L$ as $D^{-1}L\vec{y} = \lambda\vec{y}$. Moreover, it is also a classical eigenvalue of normalized Laplacian matrix $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ as

$$D^{-\frac{1}{2}}LD^{-\frac{1}{2}}\vec{z} = \lambda\vec{z},$$

for if $\vec{y} = D^{-\frac{1}{2}}\vec{z}$ or equivalently $\vec{z} = D^{\frac{1}{2}}\vec{y}$.

The normalized Laplacian matrix is symmetric (and real) and hence diagonalizable which means that we can express it as follows:

$$D^{-\frac{1}{2}}LD^{-\frac{1}{2}} = Q^\top TQ,$$

where $Q \in \mathbb{R}^{n \times n}$ is orthogonal and $T \in \mathbb{R}^{n \times n}$ is diagonal. Recall that $\mu_i(A)$ denotes a classical eigenvalue of any square matrix A. As

$$Q^\top TQ = Q^\top T^{\frac{1}{2}}(Q^\top T^{\frac{1}{2}})^\top \,,$$

we can see that

$$\mu_i(D^{-\frac{1}{2}}LD^{-\frac{1}{2}}) = \mu_i(Q^\top TQ) = \sigma_i^2(Q^\top T^{\frac{1}{2}}) \,,$$

which is equal to the generalized eigenvalue of L:

$$\lambda_i(L) = \sigma_i^2(Q^\top T^{\frac{1}{2}}) \,, \tag{2}$$

that we are interested in. Hence the singular value decomposition of the matrix $Q^\top T^{\frac{1}{2}}$ follows.

Recall that the diagonal entries of the matrix T are classical eigenvalues of $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ which ere equal to λ_i, so

$$t_{ii} = \lambda_i \,,$$

and hence

$$t_{ii}^{\frac{1}{2}} = \sqrt{\lambda_i} \,,$$

the diagonal entries of the matrix $T^{\frac{1}{2}}$.

We are looking for orthogonal matrices U and V and a diagonal matrix S s.t.

$$Q^\top T^{\frac{1}{2}} = USV^\top \,,$$

where

$$s_{ii} = \sigma_i \,,$$

and from (2), we get

$$s_{ii} = \sqrt{\lambda_i},$$

which proves that

$$S = T^{\frac{1}{2}}.$$

This shows, that if we diagonalize the normalized Laplacian (by computing T), the SVD will then not be necessary (S will be clear), as the generalized eigenvalues will instantly emerge.

From the above, the main result of this paper regarding processing of data of one arbitrary nature can be formulated as follows:

Proposition 1. *If data is described by a symmetric, non-negative weighted graph adjacency matrix W, D is derived degree matrix, L is derived Laplacian matrix and eigendecoposition of the normalized Laplacian matrix $D^{-\frac{1}{2}}LD^{-\frac{1}{2}}$ is equal to $Q^{\top}TQ$ producing a unique diagonal matrix T, then SVD factorization of the matrix $Q^{\top}T^{\frac{1}{2}}$ produces a unique diagonal matrix S s.t. $S = T^{\frac{1}{2}}$. Moreover, the values of S^2 are equal to generalized eigenvalues of L.*

This result gives us a new perspective of DR solution as it connects Laplacian eigenmaps (that can be found after computing the generalized eigenvalues, see Example 1) with a much more general matrix factorization technique that is applicable in a broader area of problems.

Below, we would like to illustrate the creation of a reduced data representation given by the Laplacian eigendecomposition where the space closeness is determined by a fuzzy partition.

Example 1. *Consider a set of two-dimensional points*

$$X = \big\{(1,1),(2,2),(3,4)\big\}.$$

Our purpose is to reduce the dimensionality of the given data from two to one. As a first step, let us set up a universe $[-2,6] \times [-2,7]$ and establish its fuzzy partition using the nodes $\{1,2,3\}$ along the x-axis and $\{1,2,4\}$ along the y-axis. Our basic functions A_1, A_2, A_3, B_1, B_2, B_3 are of the symmetric triangular shapes with support lengths 6. Their combinations $A_i \cdot B_j$ constitute the fuzzy partition of the universe $[-2,6] \times [-2,7]$. We can now define the closeness between any pair of points (x_i,y_i) and (x_j,y_j) in X by

$$w_{ij} = A_i(x_j) \cdot B_i(y_j).$$

Therefore, the weight matrix is

$$W = \begin{bmatrix} 1 & \frac{4}{9} & 0 \\ \frac{4}{9} & 1 & \frac{2}{9} \\ 0 & \frac{2}{9} & 1 \end{bmatrix}.$$

The corresponding diagonal matrix is equal to

$$D = \begin{bmatrix} \frac{13}{9} & 0 & 0 \\ 0 & \frac{5}{3} & 0 \\ 0 & 0 & \frac{11}{9} \end{bmatrix},$$

and the Laplacian matrix is

$$L = \begin{bmatrix} \frac{4}{9} & -\frac{4}{9} & 0 \\ -\frac{4}{9} & \frac{2}{3} & -\frac{2}{9} \\ 0 & -\frac{2}{9} & \frac{2}{9} \end{bmatrix}.$$

Its generalized eigenvalues in decreasing order are $\lambda_1 \doteq 0.674$, $\lambda_2 \doteq 0.216$ and $\lambda_3 = 0$. One-dimensional embedding of X is then determined by the generalized eigenvector $\vec{y}^2$ corresponding to λ_2 (the second smallest gen. eigenvalue), i.e. by

$$\vec{y}^2 \doteq \begin{bmatrix} -0.526 \\ -0.157 \\ 0.836 \end{bmatrix}.$$

As expected, the first two points are mapped closer together: $(1,1) \mapsto -0.526$, $(2,2) \mapsto -0.157$, while $(3,4) \mapsto 0.836$ is mapped further.

The inverse problem consists in finding approximation representation of a given data using its reduced form. We plan to discuss this problem using the technique proposed in this contribution using the SVD factorization.

4 Conclusions

We showed that the singular value decomposition of the Laplacian matrix L coincides with its eigendecomposition. We focused on lower-dimensional embedding of given (high-dimensional) points and connected it with singular values of the transformed Laplacian matrix. We showed that SVD factorization of this matrix need not be explicitly computed to find these values as they coincide with square roots of the diagonalized normalized Laplacian matrix. Showing this connection enables us to broaden the scope of our future research by considering data spaces where their geometry is described by rectangular matrices.

Our future research will be focused on the inverse problem connected with DR where we plan to use the technique of SVD decomposition.

References

[1] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation, Neural computation 15 (6) (2003) 1373–1396.

[2] G. H. Golub, C. F. Van Loan, Matrix computations, Johns Hopkins University Press, 1996.

[3] J. Janeček, I. Perfiljeva, Dimensionality reduction and its f-transform representation, in: Proceedings of the 11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019), Atlantis Press, 2019, pp. 835–842.

[4] J. Janeček, I. Perfiljeva, F-transform and dimensionality reduction: Common and different, in: 10th International Summer School on Aggregation Operators: New Trends in Aggregation Theory, Springer Nature Switzerland AG, 2019, pp. 267–278.