Laplacian Matrix for Dimensionality Reduction and Clustering

— Exercises with Solutions —

Laurenz Wiskott
Institut für Neuroinformatik
Ruhr-Universität Bochum, Germany, EU

28 September 2019

Contents

These exercises complement my corresponding lecture notes, and there is a version with and one without solutions. The table of contents of the lecture notes is reproduced here to give an orientation when the exercises can be reasonably solved. For best learning effect I recommend to first seriously try to solve the exercises yourself before looking into the solutions.
More teaching material is available at https://www.ini.rub.de/PEOPLE/wiskott/Teaching/Material/.

1 Introduction

2 Intuition

2.1 Heat diffusion analogy of Laplacian eigenmaps

2.2 Heat diffusion analogy of spectral clustering

2.3 Heat diffusion equation for connected heat reservoirs

2.4 Laplacian matrix

2.4.1 Exercise: Laplacian matrix is positive semi-definite

1. Create a small undirected graph with four vertices without loops and edge weights one and calculate its Laplacian.

   Solution: A graph with four vertices and edges $e_1 = (1,2), e_2 = (2,3), e_3 = (2,4)$ with weight one has the Laplacian

$$\boldsymbol{L} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \end{array} \tag{1}$$

2. For a directed graph $G = (V, E)$ without loops with vertices $V = \{v_1, ..., v_n\}$ and edges $E = \{e_1, ..., e_m\}$ (where edges are ordered pairs $e_i = (v_j, v_k)$ indicating a connection to node v_j from node v_k) the $n \times m$ incidence matrix $B = (b_{ij})$ is defined as:

$$b_{ij} := \begin{cases} +1 & \text{if } e_j = (v_i, v_x) \\ 0 & \text{if } v_i \notin e_j \\ -1 & \text{if } e_j = (v_x, v_i) \end{cases} \tag{2}$$

   with v_x being an arbitrary vertex. (Sometimes the transpose is defined as the incidence matrix.)

   Treat the undirected graph above as a directed graph, where the edges always go from the lower to the higher index vertex, and calculate its incidence matrix.

   Solution:

$$\boldsymbol{B} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{ccc} e_1 & e_2 & e_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \end{array} \tag{3}$$

   Extra question: Loops are edges that connect a vertex with itself. Why is it important for an incidence matrix that the graph has no loops?

3. Show that $\boldsymbol{L} = \boldsymbol{B}\boldsymbol{B}^T$. Argue also why this is generally the case, not only in this concrete example.

Solution: First we consider the concrete example:

$$\boldsymbol{L} \stackrel{?}{=} \boldsymbol{B}\boldsymbol{B}^T \tag{4}$$

$$\stackrel{(3)}{=} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \tag{5}$$

$$= \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \tag{6}$$

$$\stackrel{(1)}{=} \boldsymbol{L} \tag{7}$$

Generally, the Laplacian matrix is the degree matrix $\boldsymbol{D}$ (which is a diagonal matrix with the number of connections per vertex) minus the adjacency matrix $\boldsymbol{A}$ (which simply indicates with a $+1$ if two vertices are connected, assuming the connecting weights are just $+1$).

The rows of the incidence matrix $\boldsymbol{B}$ contain a $+1$ or -1 for each edge that the corresponding vertex has. The inner product of each row vector with itself therefore indicates the degree of the corresponding vertex, which yields the entries of $\boldsymbol{D}$ on the diagonal.

The inner product between two different row vectors is either -1, if the two corresponding vertices are connected by an edge or zero if they are not. The reason is that each column vector represents an edge with exactly one $+1$ and one -1. No two column vectors have the two ± 1s at the same position, because then they would be the same edge. So two row vectors can have at most one component occupied in common, indicating the edge between these two vertices. Since one component is -1 the other $+1$, the product is -1, yielding the -1s in the off-diagonal entries of the Laplacian.

Extra question: How does the result change if one inverts an edge, making it point from a higher to a lower index node?

4. Show that $\boldsymbol{L} = \boldsymbol{M}\boldsymbol{M}^T$ also holds for a Laplacian matrix with weights not equal to $+1$, thus for a graph with general weighting of the edges, with an appropriately chosen matrix $\boldsymbol{M}$.

   Solution: If the weights of the graph are indicated by W_{ij}, then one can simply chose $\pm\sqrt{W_{ij}}$ in $\boldsymbol{M}$ where there is a ± 1 in $\boldsymbol{B}$ to indicate an edge. Since the inner products between row vectors of the incidence matrix multiply only the entries comming from the same edge, matrix $\boldsymbol{L}$ ends up having $-\sqrt{W_{ij}}\sqrt{W_{ij}} = -W_{ij}$ in the off-diagonal entries and the sum over all W_{ij} for fixed i (or j) in the diagnoal entries.

5. Prove that $\boldsymbol{L}$ is positive semi-definite.

   Solution: This is obvious now since

$$\boldsymbol{x}^T\boldsymbol{L}\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{M}\underbrace{\boldsymbol{M}^T\boldsymbol{x}}_{=:\boldsymbol{y}} = \boldsymbol{y}^T\boldsymbol{y} = \|\boldsymbol{y}\|^2 \geq 0 \tag{8}$$

2.5 Solution of the heat diffusion equation

2.5.1 Exercise: Eigenvectors and -values of the Laplacian matrix

Consider the Laplacian matrix

$$\boldsymbol{L} = \begin{pmatrix} a & -a & 0 \\ -a & a+b & -b \\ 0 & -b & b \end{pmatrix} \tag{1}$$

with $0 < a, b$.

1. Solve the ordinary eigenvalue problem $\boldsymbol{L}\boldsymbol{u}_i = \gamma_i \boldsymbol{u}_i$ for $i = 1, 2, 3$.

   Hint 1: A somewhat tedious calculation yields

$$
\begin{aligned}
0 \stackrel{!}{=} |\boldsymbol{L} - \gamma \boldsymbol{I}| &= (a-\gamma)(a+b-\gamma)(b-\gamma) - (a-\gamma)b^2 - (b-\gamma)a^2 && (2)\\
\Longleftarrow \qquad \gamma_2 &= (a+b) - \sqrt{a^2 - ab + b^2} && (3)\\
\vee \quad \gamma_3 &= (a+b) + \sqrt{a^2 - ab + b^2} && (4)
\end{aligned}
$$

   You may take this for granted to also find γ_1 and $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and $\boldsymbol{u}_3$.

   Hint 2: One can easily make hypotheses about $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ by looking at $\boldsymbol{L}$. Do not try to calculate them.

   Solution: The roots of the characteristic polynomial are

$$
\begin{aligned}
0 &\stackrel{!}{=} |\boldsymbol{L} - \gamma \boldsymbol{I}| && (5)\\
&= \begin{vmatrix} a-\gamma & -a & 0 \\ -a & a+b-\gamma & -b \\ 0 & -b & b-\gamma \end{vmatrix} && (6)\\
&= (a-\gamma)(a+b-\gamma)(b-\gamma) - (a-\gamma)b^2 - (b-\gamma)a^2 && (7)\\
&= +aab - aa\gamma + abb - ab\gamma - a\gamma b + a\gamma\gamma - \gamma ab + \gamma a\gamma - \gamma bb + \gamma b\gamma + \gamma\gamma b - \gamma\gamma\gamma \\
&\quad - ab^2 + \gamma b^2 - ba^2 + \gamma a^2 && (8)\\
&= -\gamma^3 + 2\gamma^2 a + 2\gamma^2 b - 3\gamma ab && (9)\\
&= -\gamma(\gamma^2 - 2\gamma(a+b) + 3ab) && (10)\\
\Longleftarrow \qquad \gamma_1 &= 0 && (11)\\
\vee \qquad 0 &= \gamma^2 - 2\gamma(a+b) + 3ab && (12)\\
&= \gamma^2 - 2\gamma(a+b) + (a+b)^2 - (a+b)^2 + 3ab && (13)\\
&= (\gamma - (a+b))^2 - (a^2 - ab + b^2) && (14)\\
\Longleftrightarrow \quad \gamma - (a+b) &= \pm\sqrt{a^2 - ab + b^2} && (15)\\
\Longleftrightarrow \qquad \gamma_2 &= (a+b) - \sqrt{a^2 - ab + b^2} && (16)\\
\vee \quad \gamma_3 &= (a+b) + \sqrt{a^2 - ab + b^2} && (17)
\end{aligned}
$$

   For $\gamma_1 = 0$ we can chose $\boldsymbol{u}_1 = (1, 1, 1)^T$, since

$$
\boldsymbol{0} \stackrel{?}{=} (\boldsymbol{L} - \gamma_1)\boldsymbol{u}_1 = \begin{pmatrix} a & -a & 0 \\ -a & a+b & -b \\ 0 & -b & b \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (18)
$$

   For $\gamma = \gamma_{2,3} = (a+b) \pm \sqrt{a^2 - ab + b^2}$ we can see from the first and second component that we can chose $\boldsymbol{u}_{2,3} = (a/(a-\gamma), 1, b/(b-\gamma))^T$, since

$$\begin{aligned}
\mathbf{0} &\stackrel{?}{=} (\boldsymbol{L}-\gamma)\boldsymbol{u}_{2,3} && (19)\\
&= \begin{pmatrix} a-\gamma & -a & 0 \\ -a & a+b-\gamma & -b \\ 0 & -b & b-\gamma \end{pmatrix} \begin{pmatrix} \frac{a}{a-\gamma} \\ 1 \\ \frac{b}{b-\gamma} \end{pmatrix} && (20)\\
&= \begin{pmatrix} 0 \\ -\frac{a^2}{a-\gamma} + (a+b-\gamma) - \frac{b^2}{b-\gamma} \\ 0 \end{pmatrix} && (21)
\end{aligned}$$

For the second component we verify

$$\begin{aligned}
0 &\stackrel{?}{=} -\frac{a^2}{a-\gamma} + (a+b-\gamma) - \frac{b^2}{b-\gamma} && (22)\\
\iff \quad 0 &= -a^2(b-\gamma) + (a+b-\gamma)(a-\gamma)(b-\gamma) - b^2(a-\gamma) && (23)\\
&\stackrel{(7)}{=} 0 && (24)
\end{aligned}$$

We also verify that the eigenvectors $\boldsymbol{u}_i$ are orthogonal to each other.

For $\boldsymbol{u}_1$ versus $\boldsymbol{u}_{2,3}$ we find

$$\begin{aligned}
0 &\stackrel{?}{=} \boldsymbol{u}_1^T \boldsymbol{u}_{2,3} && (25)\\
&= (1,1,1) \begin{pmatrix} \frac{a}{a-\gamma} \\ 1 \\ \frac{b}{b-\gamma} \end{pmatrix} && (26)\\
&= \frac{a}{a-\gamma} + 1 + \frac{b}{b-\gamma} && (27)\\
\iff \quad 0 &= a(b-\gamma) + (a-\gamma)(b-\gamma) + b(a-\gamma) && (28)\\
&= ab - a\gamma + ab - a\gamma - \gamma b + \gamma^2 + ba - b\gamma && (29)\\
&= \gamma^2 - 2\gamma(a+b) + 3ab && (30)\\
&\stackrel{(12)}{=} 0 && (31)
\end{aligned}$$

For $\boldsymbol{u}_2$ versus $\boldsymbol{u}_3$ we find

$$
\begin{aligned}
0 &\stackrel{?}{=} \boldsymbol{u}_2^T \boldsymbol{u}_3 && (32)\\
&= \left(\frac{a}{a-\gamma_2}, 1, \frac{b}{b-\gamma_2}\right) \begin{pmatrix} \frac{a}{a-\gamma_3} \\ 1 \\ \frac{b}{b-\gamma_3} \end{pmatrix} && (33)\\
&= \frac{a^2}{(a-\gamma_2)(a-\gamma_3)} + 1 + \frac{b^2}{(b-\gamma_2)(b-\gamma_3)} && (34)\\
&= \frac{a^2}{a^2 - a(\gamma_2+\gamma_3) + \gamma_2\gamma_3} + 1 + \frac{b^2}{b^2 - b(\gamma_2+\gamma_3) + \gamma_2\gamma_3} && (35)\\
&\stackrel{(16,17)}{=} \frac{a^2}{a^2 - a\,2(a+b) + [(a+b)^2 - (a^2 - ab + b^2)]} + 1 \\
&\quad + \frac{b^2}{b^2 - b\,2(a+b) + [(a+b)^2 - (a^2 - ab + b^2)]} && (36)\\
&= \frac{a^2}{a^2 - [2a^2 + 2ab] + [a^2 + 2ab + b^2 - a^2 + ab - b^2]} + 1 \\
&\quad + \frac{b^2}{b^2 - [2b^2 + 2ab] + [a^2 + 2ab + b^2 - a^2 + ab - b^2]} && (37)\\
&= \frac{a^2}{ab - a^2} + 1 + \frac{b^2}{ab - b^2} && (38)\\
&= -\frac{a}{a-b} + 1 + \frac{b}{a-b} && (39)\\
&= -\frac{a-b}{a-b} + 1 && (40)\\
&= 0 && (41)
\end{aligned}
$$

Extra question: What can you say about the sign and relative value of the components of $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ depending on a and b?

2.5.2 Exercise: Laplacian matrix for disconnected graphs

1. Show that a Laplacian matrix of a graph with N disconnected subgraphs, i.e. subgraphs that have no edges between them, has at least N eigenvectors with eigenvalue zero.

   Solution: Consider first a Laplacian matrix of a connected graph, i.e. $N = 1$. The Laplacian matrix is defined as

$$
\begin{aligned}
&& \boldsymbol{L} &:= \boldsymbol{D} - \boldsymbol{W} && (1)\\
&\iff & L_{ij} &= D_{ii}\delta_{ij} - W_{ij} && (2)
\end{aligned}
$$

   with

$$
\begin{aligned}
D_{ii} &:= \sum_j W_{ij} = \sum_j W_{ji} \quad \text{(since } \boldsymbol{W} \text{ is symmetric)} && (3)\\
W_{ii} &= 0 && (4)
\end{aligned}
$$

With this we have

$$D_{ii} \overset{(3)}{=} \sum_j W_{ij} \quad \forall i \tag{5}$$

$$\iff \quad L_{ii} \overset{(2,4)}{=} -\sum_{j \neq i} L_{ij} \tag{6}$$

$$\iff \quad \sum_j L_{ij} = 0 \tag{7}$$

$$\iff \quad \boldsymbol{L}\mathbf{1} = 0 \cdot \mathbf{1} \tag{8}$$

with $\mathbf{1} = (1, 1, 1, ..., 1)^T$. Thus, $\mathbf{1}$ is always an eigenvector of $\boldsymbol{L}$ with eigenvalue 0.

If the graph has more than one disconnected subgraphs, the indices of the vertices can be reordered such that $\boldsymbol{L}$ has block structure with N blocks. The consideration above holds for each block individually. Thus, there is one eigenvector of the form $(1, 1, 1, ..., 1, 0, 0, 0, ..., 0)^T$, $N - 2$ of the form $(0, 0, 0, ..., 0, 1, 1, 1, ..., 1, 0, 0, 0, ..., 0)^T$, and one of the form $(0, 0, 0, ..., 0, 1, 1, 1, ..., 1)^T$, all with eigenvalue zero.

These eigenvectors are already orthogonal to each other but must still be normalized to yield an orthonormal set of eigenvectors.

2. Argue why there are no more than N eigenvectors with eigenvalue zero.

   Hint: You may use the relation

$$\frac{1}{2} \sum_{ij} (u_i - u_j)^2 W_{ij} = \boldsymbol{u}^T \boldsymbol{L} \boldsymbol{u} \tag{9}$$

   Solution: Any vector that has constant values within the diconnected subgraphs can be written as a linear combination of the eigenvectors given above. Thus, any additional eigenvector $\boldsymbol{u}_i$ must introduce variation within at least one subgraph. Because of (9) this also means that $0 < \boldsymbol{u}_i^T \boldsymbol{L} \boldsymbol{u}_i$. This in turn implies that the corresponding eigenvalue γ_i is greater zero because if it were zero, $\boldsymbol{L}\boldsymbol{u}_i = \mathbf{0}$ and $\boldsymbol{u}_i^T \boldsymbol{L} \boldsymbol{u}_i$ could not be greater than zero. Thus, there are no more eigenvectors with eigenvalue zero.

3. Do the results above also hold for the generalized eigenvalue equation

$$\boldsymbol{L}\boldsymbol{w} = \lambda \boldsymbol{D}\boldsymbol{w} \tag{10}$$

   Solution: If the eigenvalue is zero then the generalized eigenvalue equation of the pair $\boldsymbol{L}$ and $\boldsymbol{D}$ becomes the ordinary eigenvalue equation of $\boldsymbol{L}$. Thus, eigenvectors with eigenvalue zero are common to both equations. For the same reason there are no eigenvectors with eigenvalue zero for one but not the other equation. All other eigenvectors therefore have non-zero eigenvalues. Thus, the considerations above also hold for the generalized eigenvalue equation.

3 Formalism

3.1 Simple graphs

3.2 Matrix representation

3.3 Optimization problem

3.4 Associated eigenvalue problem

3.4.1 Exercise: Objective function of the Laplacian matrix

Given the Laplacian matrix

$$\begin{aligned} \boldsymbol{L} &:= \boldsymbol{D} - \boldsymbol{W} && (1) \\ \iff L_{ij} &= D_{ii}\delta_{ij} - W_{ij} && (2) \end{aligned}$$

with symmetric $W_{ij} = W_{ji}$ and

$$D_{ii} := \sum_j W_{ij} = \sum_j W_{ji} \tag{3}$$

Show that

$$\frac{1}{2}\sum_{i,j}(y_i - y_j)^2 W_{ij} = \boldsymbol{y}^T \boldsymbol{L} \boldsymbol{y} \tag{4}$$

Solution:

$$\begin{aligned}
& \frac{1}{2}\sum_{i,j}(y_i - y_j)^2 W_{ij} && (5) \\
= & \frac{1}{2}\sum_{i,j}(y_i^2 - 2y_i y_j + y_j^2) W_{ij} && (6) \\
= & \frac{1}{2}\sum_{i,j} y_i^2 W_{ij} - \sum_{i,j} y_i y_j W_{ij} + \frac{1}{2}\sum_{i,j} y_j^2 W_{ij} && (7) \\
= & \frac{1}{2}\sum_{i,j} y_i^2 W_{ij} - \sum_{i,j} y_i y_j W_{ij} + \frac{1}{2}\sum_{i,j} y_j^2 W_{ji} \quad \text{(since } W_{ij} \text{ is symmetric)} && (8) \\
= & \frac{1}{2}\sum_{i,j} y_i^2 W_{ij} - \sum_{i,j} y_i y_j W_{ij} + \frac{1}{2}\sum_{i,j} y_i^2 W_{ij} \quad \text{(just swapping the indices)} && (9) \\
= & \sum_{i,j} y_i^2 W_{ij} - \sum_{i,j} y_i y_j W_{ij} && (10) \\
= & \sum_i y_i^2 \sum_j W_{ij} - \sum_{i,j} y_i y_j W_{ij} && (11) \\
\overset{(3)}{=} & \sum_i y_i^2 D_{ii} - \sum_{i,j} y_i y_j W_{ij} && (12) \\
= & \sum_{i,j} y_i (D_{ii}\delta_{ij} - W_{ij}) y_j && (13) \\
\overset{(2)}{=} & \sum_{i,j} y_i L_{ij} y_j && (14) \\
= & \boldsymbol{y}^T \boldsymbol{L} \boldsymbol{y} && (15)
\end{aligned}$$

and thus minimizing $(1/2)\sum_{i,j}(y_i - y_j)^2 W_{ij}$ is equivalent to minimizing $\boldsymbol{y}^T\boldsymbol{L}\boldsymbol{y}$.

This also implies that $\boldsymbol{L}$ is positive semi-definite if $0 \le W_{ij}\,\forall i,j$.

3.4.2 Exercise: Generalized eigenvalue problem

Consider the generalized eigenvalue problem

$$\boldsymbol{A}\boldsymbol{u}_i \quad = \quad \lambda_i \boldsymbol{B}\boldsymbol{u}_i \tag{1}$$

with some real $N \times N$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. The λ_i are the right-eigenvalues and the $\boldsymbol{u}_i$ are the (non-zero) right-eigenvectors. To find corresponding left-eigenvalues μ_i and left-eigenvectors $\boldsymbol{v}_i$, one has to solve the equation

$$\boldsymbol{v}_i^T\boldsymbol{A} \quad = \quad \mu_i \boldsymbol{v}_i^T\boldsymbol{B} \tag{2}$$

1. Show that left- and right-eigenvalues are identical.

   Solution: For the right-eigenvalues we find the condition that

$$\begin{aligned}
& \boldsymbol{A}\boldsymbol{u}_i \stackrel{(1)}{=} \lambda_i \boldsymbol{B}\boldsymbol{u}_i && (3)\\
\iff \quad & (\boldsymbol{A} - \lambda_i \boldsymbol{B})\boldsymbol{u}_i = \boldsymbol{0} && (4)\\
\implies \quad & |\boldsymbol{A} - \lambda_i \boldsymbol{B}| = 0 && (5)
\end{aligned}$$

   For the left-eigenvalues we find analogously

$$\begin{aligned}
& \boldsymbol{v}_i^T\boldsymbol{A} \stackrel{(2)}{=} \mu_i \boldsymbol{v}_i^T\boldsymbol{B} && (6)\\
\iff \quad & \boldsymbol{v}_i^T(\boldsymbol{A} - \mu_i \boldsymbol{B}) = \boldsymbol{0}^T && (7)\\
\implies \quad & |\boldsymbol{A} - \mu_i \boldsymbol{B}| = 0 && (8)
\end{aligned}$$

   Since the conditions for λ_i and μ_i are identical, the (sorted) left- and right-eigenvalues are identical and are simply called eigenvalues λ_i.

2. Show that $\boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{u}_i = 0$ as well as $\boldsymbol{v}_j^T\boldsymbol{B}\boldsymbol{u}_i = 0$ for $\lambda_i \neq \lambda_j$.

   Hint: Consider (1) and (2) simultaneously with different eigenvalues.

   Solution: We consider the right- and left-eigenvalue equation for i and j, respectively,

$$\begin{aligned}
& \boldsymbol{A}\boldsymbol{u}_i \stackrel{(1)}{=} \lambda_i \boldsymbol{B}\boldsymbol{u}_i \qquad |\boldsymbol{v}_j^T\cdot && (9)\\
\wedge \quad & \boldsymbol{v}_j^T\boldsymbol{A} \stackrel{(2)}{=} \lambda_j \boldsymbol{v}_j^T\boldsymbol{B} \qquad |\cdot\boldsymbol{u}_i && (10)\\
\implies \quad & \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{u}_i = \lambda_i \boldsymbol{v}_j^T\boldsymbol{B}\boldsymbol{u}_i && (11)\\
\wedge \quad & \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{u}_i = \lambda_j \boldsymbol{v}_j^T\boldsymbol{B}\boldsymbol{u}_i && (12)\\
\iff \quad & \boldsymbol{v}_j^T\boldsymbol{A}\boldsymbol{u}_i = 0 && (13)\\
\wedge \quad & \boldsymbol{v}_j^T\boldsymbol{B}\boldsymbol{u}_i = 0 \qquad (\text{since } \lambda_i \neq \lambda_j) && (14)
\end{aligned}$$

   Notice that this is even true if either $\lambda_i = 0$ or $\lambda_j = 0$, but not both, of course.

3. Show that for symmetric $\boldsymbol{A}$ and $\boldsymbol{B}$ the right-eigenvectors are also left-eigenvectors.

   Solution: This is easy to see if we transpose the right-eigenvalue equation.

$$
\begin{aligned}
& \boldsymbol{A}\boldsymbol{u}_i \overset{(1)}{=} \lambda_i \boldsymbol{B}\boldsymbol{u}_i \qquad |.^T && (15)\\
\iff \quad & \boldsymbol{u}_i^T \boldsymbol{A}^T = \lambda_i \boldsymbol{u}_i^T \boldsymbol{B}^T && (16)\\
\iff \quad & \boldsymbol{u}_i^T \boldsymbol{A} = \lambda_i \boldsymbol{u}_i^T \boldsymbol{B} \qquad (\text{since } \boldsymbol{A} = \boldsymbol{A}^T \text{ and } \boldsymbol{B} = \boldsymbol{B}^T) && (17)\\
\iff \quad & (2) && (18)
\end{aligned}
$$

   Thus, right-eigenvectors are also left-eigenvectors and we can simply speak of eigenvectors.

4. Show that for symmetric $\boldsymbol{A}$ and $\boldsymbol{B}$ we have $\boldsymbol{u}_j^T \boldsymbol{A}\boldsymbol{u}_i = 0$ as well as $\boldsymbol{u}_j^T \boldsymbol{B}\boldsymbol{u}_i = 0$ for $\lambda_i \neq \lambda_j$.

   Solution: This follows trivially from the last two points of this exercise. This is not generally true for non-symmetric matrices.

5. For symmetric $\boldsymbol{A}$ and $\boldsymbol{B}$ it is convenient to normalize the eigenvectors such that $\boldsymbol{u}_i^T \boldsymbol{B}\boldsymbol{u}_i = 1$. Assume the eigenvectors form a basis, i.e. they are complete, and you want to represent an arbitrary vector $\boldsymbol{y}$ wrt this basis, i.e.

$$
\boldsymbol{y} = \sum_i \alpha_i \boldsymbol{u}_i \tag{19}
$$

   Which constraint on the α_i follows from the constraint $\boldsymbol{y}^T \boldsymbol{B}\boldsymbol{y} = 1$?

   Solution:

$$
\begin{aligned}
1 &\overset{!}{=} \boldsymbol{y}^T \boldsymbol{B}\boldsymbol{y} && (20)\\
&= \sum_{i,j} \alpha_i \alpha_j \underbrace{\boldsymbol{u}_i^T \boldsymbol{B}\boldsymbol{u}_j}_{=\delta_{ij}} && (21)\\
&= \sum_i \alpha_i^2 && (22)
\end{aligned}
$$

6. Assume $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric and you want to minimize (or maiximze) $\boldsymbol{y}^T \boldsymbol{A}\boldsymbol{y}$ under the constraint $\boldsymbol{y}^T \boldsymbol{B}\boldsymbol{y} = 1$. What is the solution?

   Hint: Use ansatz (19) and assume $\boldsymbol{u}_i^T \boldsymbol{B}\boldsymbol{u}_i = 1$.

   Solution: We want to minimize (or maximize)

$$
\begin{aligned}
\boldsymbol{y}^T \boldsymbol{A}\boldsymbol{y} &\overset{(19)}{=} \sum_{i,j} \alpha_i \alpha_j \boldsymbol{u}_i^T \boldsymbol{A}\boldsymbol{u}_j && (23)\\
&\overset{(1)}{=} \sum_{i,j} \alpha_i \alpha_j \lambda_j \underbrace{\boldsymbol{u}_i^T \boldsymbol{B}\boldsymbol{u}_j}_{=\delta_{ij}} && (24)\\
&= \sum_i \alpha_i^2 \lambda_i && (25)
\end{aligned}
$$

   Since this is minimized (or maximized) subject to $\sum_i \alpha_i^2 = 1$ it is quite obvious that $\alpha_i = \delta_{1i}$ is an optimal solution if the eigenvalues are ordered like $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_1$ (or $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_1$). Thus $\boldsymbol{u}_1$ is the optimal solution.

   Extra question: If for some reason $\boldsymbol{u}_1$ is not wanted as a solution and the solution should be orthogonal to it wrt a metric induced by $\boldsymbol{B}$, i.e. $\boldsymbol{y}\boldsymbol{B}\boldsymbol{u}_1 \overset{!}{=} 0$, what would then be the optimal solution?

3.4.3 Exercise: Eigenvectors of a graph with six nodes

Consider a graph with six nodes arranged in a 2×3 lattice with edges of weight 1 between direct neighbors, like the one shown in the figure.

Make an educated guess how the six eigenvectors of the corresponding Laplacian matrix might look.

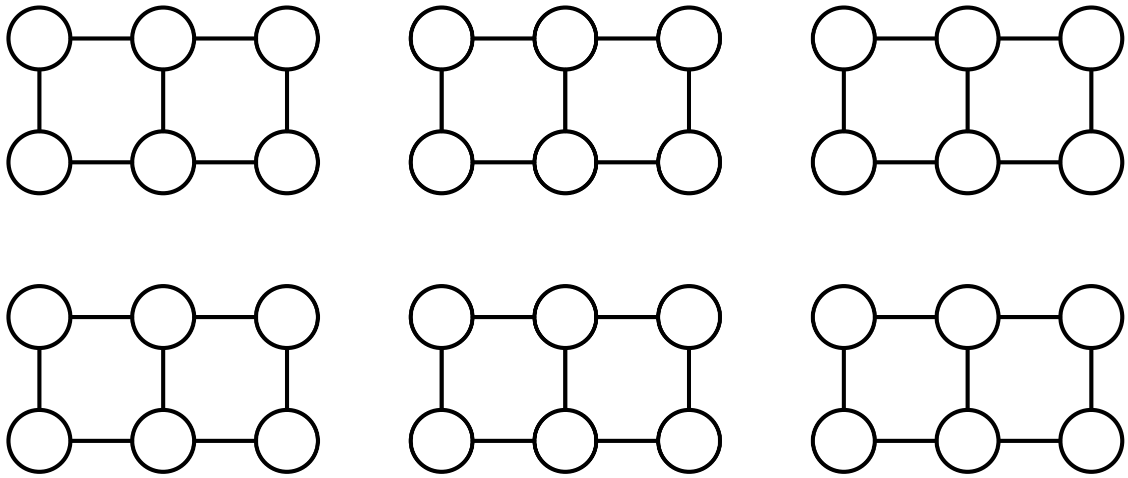

Figure: Graph with six nodes. Color the nodes red/blue or mark them with +/- according to the sign you give them. Nodes with value near zero stay empty.

Solution: Since

$$\boldsymbol{u}^T \boldsymbol{L} \boldsymbol{u} = \frac{1}{2} \sum_{ij} (u_i - u_j)^2 W_{ij} \tag{1}$$

the first eigenvector $\boldsymbol{u}_1$ minimizes $\sum_{ij}(u_{1,i} - u_{1,j})^2 W_{ij}$ under a fixed norm constraint. It is obvious that a vector with identical components does that. Higher eigenvectors $\boldsymbol{u}_i, 1 < i$ minimize the same function, but must also be orthogonal to the first and any other earlier eigenvector. Thus one looks for values on the nodes that are similar for neighboring nodes and, at the same time, vary enough to obey the orthogonality constraint. One possible proposal for eigenvectors is given in the figure below.

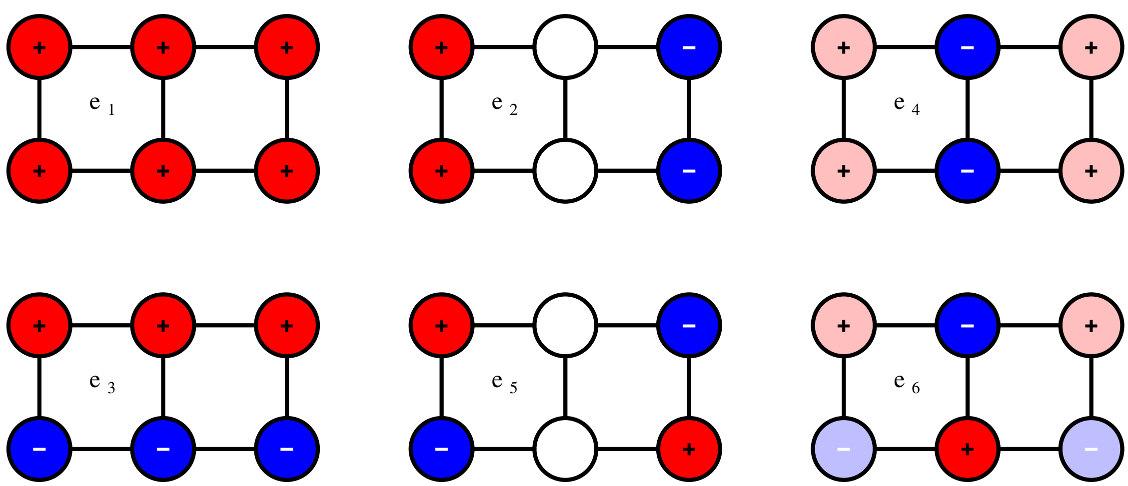

Figure: Hypothesis about six possible eigenvectors. Fainter colors indicate smaller values compared to full colors.

The next figure shows concrete values calculated numerically for this graph.

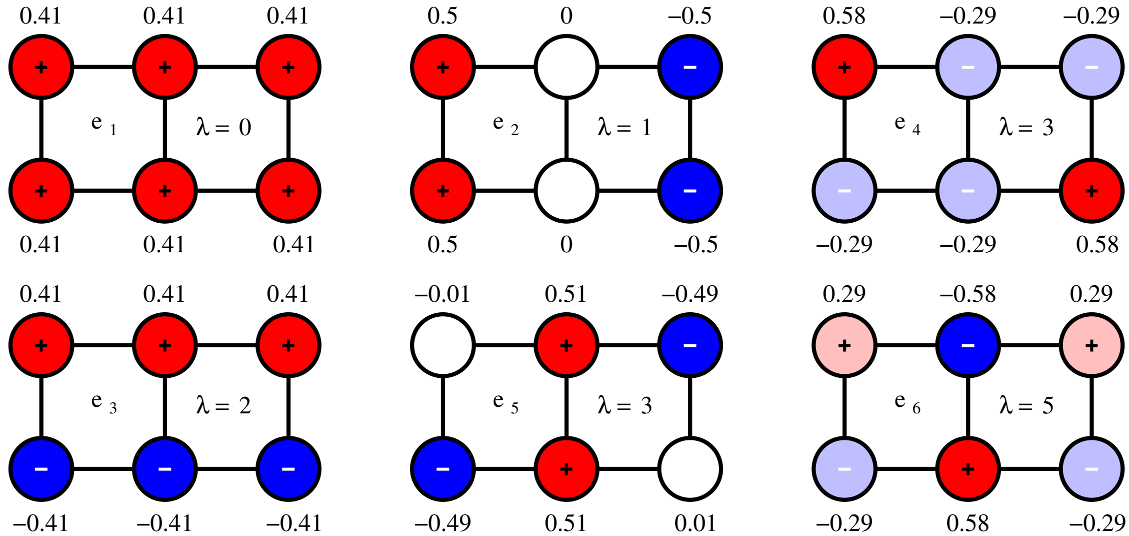

Figure: Six eigenvectors numerically calculated by Sebastian Gallon (WS'18).
Extra question: How do you explain the discrepancy between the hypothesized and the numerically determined eigenvectors.

3.4.4 Exercise: Example of Laplacian eigenmaps with three nodes

Given a connected graph with vertices v_i and undirected edges $e_k = (v_i, v_j)$ with symmetric positive weights W_{ij}, the goal of the Laplacian eigenmaps algorithm is to asign a value w_i to each vertex v_i to

$$\text{minimize} \quad \frac{1}{2}\sum_{i,j}(w_i - w_j)^2 W_{ij} \tag{1}$$

under the constraints

$$\begin{aligned} \boldsymbol{w}^T\boldsymbol{D}\mathbf{1} &= 0 \quad \text{(weighted zero mean)} & (2)\\ \text{and} \quad \boldsymbol{w}^T\boldsymbol{D}\boldsymbol{w} &= 1 \quad \text{(weighted unit variance)} & (3) \end{aligned}$$

where $\boldsymbol{D}$ is a diagonal matrix with

$$D_{ii} := \sum_j W_{ij} = \sum_j W_{ji} \quad \text{(since } \boldsymbol{W} \text{ is symmetric)} \tag{4}$$

One can show that this is solved by the second eigenvector of the generalized eigenvalue equation

$$\boldsymbol{L}\boldsymbol{w} = \lambda \boldsymbol{D}\boldsymbol{w} \tag{5}$$

with the Laplacian matrix

$$\begin{aligned} \boldsymbol{L} &:= \boldsymbol{D} - \boldsymbol{W} & (6)\\ \iff \quad L_{ij} &= D_{ii}\delta_{ij} - W_{ij} & (7) \end{aligned}$$

Verify this statement for the graph with the following Laplacian matrix:

$$\boldsymbol{L} = \begin{pmatrix} a & -a & 0 \\ -a & a+b & -b \\ 0 & -b & b \end{pmatrix} \tag{8}$$

with $0 < a, b$. Proceed as follows.

1. Sketch the graph of the Laplacian matrix.

   Solution: The graph has three nodes, the first two are connected with a weight a, the last two with a weight b. Sketch not available.

2. Solve the generalized eigenvalue problem $\boldsymbol{L}\tilde{\boldsymbol{w}}_i = \lambda_i \boldsymbol{D}\tilde{\boldsymbol{w}}_i$ for $i = 1, 2, 3$. The eigenvectors do not need to be normalized yet, that is done in the next step.

   Solution: A solution exists for values of λ for which $|\boldsymbol{L} - \lambda\boldsymbol{D}| = 0$. With $\zeta := 1 - \lambda$ we have

$$
\begin{aligned}
0 &\overset{!}{=} |\boldsymbol{L} - \lambda\boldsymbol{D}| && (9)\\
&= \begin{vmatrix} a - \lambda a & -a & 0 \\ -a & (a+b) - \lambda(a+b) & -b \\ 0 & -b & b - \lambda b \end{vmatrix} && (10)\\
&= \begin{vmatrix} \zeta a & -a & 0 \\ -a & \zeta(a+b) & -b \\ 0 & -b & \zeta b \end{vmatrix} && (11)\\
&= \zeta^3 a(a+b)b - \zeta a^2 b - \zeta a b^2 && (12)\\
&= (\zeta^3 - \zeta)(a^2 b + a b^2) && (13)\\
\iff \quad 0 &= (\zeta^3 - \zeta) \quad (\text{since } 0 < a, b) && (14)\\
\Longleftarrow \quad \zeta_1 &= 1 && (15)\\
\vee \quad \zeta_2 &= 0 && (16)\\
\vee \quad \zeta_3 &= -1 && (17)\\
\iff \quad \lambda_1 &= 0 && (18)\\
\vee \quad \lambda_2 &= 1 && (19)\\
\vee \quad \lambda_3 &= 2 && (20)
\end{aligned}
$$

   For $\lambda_1 = 0$ we can chose $\tilde{\boldsymbol{w}}_1 = (1, 1, 1)^T$, since

$$
\mathbf{0} \overset{?}{=} (\boldsymbol{L} - \lambda_1\boldsymbol{D})\tilde{\boldsymbol{w}}_1 = \begin{pmatrix} a & -a & 0 \\ -a & (a+b) & -b \\ 0 & -b & b \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (21)
$$

   For $\lambda_2 = 1$ we can chose $\tilde{\boldsymbol{w}}_2 = (-b, 0, a)^T$, since

$$
\mathbf{0} \overset{?}{=} (\boldsymbol{L} - \lambda_2\boldsymbol{D})\tilde{\boldsymbol{w}}_2 = \begin{pmatrix} 0 & -a & 0 \\ -a & 0 & -b \\ 0 & -b & 0 \end{pmatrix} \begin{pmatrix} -b \\ 0 \\ a \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (22)
$$

   For $\lambda_3 = 2$ we can chose $\tilde{\boldsymbol{w}}_3 = (1, -1, 1)^T$, since

$$
\mathbf{0} \overset{?}{=} (\boldsymbol{L} - \lambda_3\boldsymbol{D})\tilde{\boldsymbol{w}}_3 = \begin{pmatrix} -a & -a & 0 \\ -a & -(a+b) & -b \\ 0 & -b & -b \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \overset{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (23)
$$

3. Check the eigenvectors whether they are consistent with constraint (2).

Solution: For the first eigenvector we find

$$
\begin{aligned}
0 &\stackrel{?}{=} \tilde{\boldsymbol{w}}_1^T \boldsymbol{D}\mathbf{1} && (24)\\
&= (1,1,1)\begin{pmatrix} a & 0 & 0 \\ 0 & (a+b) & 0 \\ 0 & 0 & b \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} && (25)\\
&= (1,1,1)\begin{pmatrix} a \\ (a+b) \\ b \end{pmatrix} && (26)\\
&= 2(a+b) && (27)\\
&\stackrel{!}{\neq} 0 && (28)
\end{aligned}
$$

which is not true. This is not surprising, since for $0 < a, b$ the expression $\tilde{\boldsymbol{w}}_i^T \boldsymbol{D}\mathbf{1}$ can be interpreted as an inner product of $\tilde{\boldsymbol{w}}_i$ with $\mathbf{1}$ with a metric induced by $\boldsymbol{D}$. Since $\tilde{\boldsymbol{w}}_1 = \mathbf{1}$ the expression $\tilde{\boldsymbol{w}}_1^T \boldsymbol{D}\mathbf{1}$ becomes the norm of $\mathbf{1}$ squared, which should not be zero for non-zero vectors.

For the second eigenvector we find

$$
\begin{aligned}
0 &\stackrel{?}{=} \tilde{\boldsymbol{w}}_2^T \boldsymbol{D}\mathbf{1} && (29)\\
&= (-b,0,a)\begin{pmatrix} a \\ (a+b) \\ b \end{pmatrix} && (30)\\
&= -ab + ab && (31)\\
&\stackrel{!}{=} 0 && (32)
\end{aligned}
$$

For the third eigenvector we find

$$
\begin{aligned}
0 &\stackrel{?}{=} \tilde{\boldsymbol{w}}_3^T \boldsymbol{D}\mathbf{1} && (33)\\
&= (1,-1,1)\begin{pmatrix} a \\ (a+b) \\ b \end{pmatrix} && (34)\\
&= a - (a+b) + b && (35)\\
&\stackrel{!}{=} 0 && (36)
\end{aligned}
$$

This is again not surprising, since generalized eigenvectors for symmetric matrices obey the relation $\tilde{\boldsymbol{w}}_i^T \boldsymbol{D}\tilde{\boldsymbol{w}}_j = 0$ for $i \neq j$ and the first eigenvector is $\tilde{\boldsymbol{w}}_1 = \mathbf{1}$.

4. Scale the eigenvectors such that they become consistent with constraint (3).

   Hint: Use the notation $\boldsymbol{w}_i := \sigma_i \tilde{\boldsymbol{w}}_i$ with appropriate scaling factors σ_i (with $0 < \sigma_i$ to make it unique).

Solution: We calculate the scaling factors by inserting the $\boldsymbol{w}_i$ in equation (3).

$$
\begin{aligned}
1 &\overset{!}{=} \boldsymbol{w}_1^T \boldsymbol{D} \boldsymbol{w}_1 && (37)\\
&= \sigma_1^2 \tilde{\boldsymbol{w}}_1^T \boldsymbol{D} \tilde{\boldsymbol{w}}_1 && (38)\\
&= \sigma_1^2 (1,1,1) \begin{pmatrix} a & 0 & 0 \\ 0 & (a+b) & 0 \\ 0 & 0 & b \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} && (39)\\
&= \sigma_1^2 (1,1,1) \begin{pmatrix} a \\ (a+b) \\ b \end{pmatrix} && (40)\\
&= \sigma_1^2 \, 2\,(a+b) && (41)\\
\Longleftrightarrow \quad \sigma_1 &= 1/\sqrt{2\,(a+b)} \quad (\text{since } 0 < \sigma_i) && (42)
\end{aligned}
$$

$$
\begin{aligned}
1 &\overset{!}{=} \boldsymbol{w}_2^T \boldsymbol{D} \boldsymbol{w}_2 && (43)\\
&= \sigma_2^2 \tilde{\boldsymbol{w}}_2^T \boldsymbol{D} \tilde{\boldsymbol{w}}_2 && (44)\\
&= \sigma_2^2 (-b,0,a) \begin{pmatrix} a & 0 & 0 \\ 0 & (a+b) & 0 \\ 0 & 0 & b \end{pmatrix} \begin{pmatrix} -b \\ 0 \\ a \end{pmatrix} && (45)\\
&= \sigma_2^2 (-b,0,a) \begin{pmatrix} -ab \\ 0 \\ ab \end{pmatrix} && (46)\\
&= \sigma_2^2 (ab^2 + a^2 b) && (47)\\
\Longleftrightarrow \quad \sigma_2 &= 1/\sqrt{ab^2 + a^2 b} \quad (\text{since } 0 < \sigma_i) && (48)
\end{aligned}
$$

$$
\begin{aligned}
1 &\overset{!}{=} \boldsymbol{w}_3^T \boldsymbol{D} \boldsymbol{w}_3 && (49)\\
&= \sigma_3^2 \tilde{\boldsymbol{w}}_3^T \boldsymbol{D} \tilde{\boldsymbol{w}}_3 && (50)\\
&= \sigma_3^2 (1,-1,1) \begin{pmatrix} a & 0 & 0 \\ 0 & (a+b) & 0 \\ 0 & 0 & b \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} && (51)\\
&= \sigma_3^2 (1,-1,1) \begin{pmatrix} a \\ -(a+b) \\ b \end{pmatrix} && (52)\\
&= \sigma_3^2 \, 2\,(a+b) && (53)\\
\Longleftrightarrow \quad \sigma_3 &= 1/\sqrt{2\,(a+b)} \quad (\text{since } 0 < \sigma_i) && (54)
\end{aligned}
$$

Note that since $\boldsymbol{D}$ is diagonal with strictly positive diagonal elements, it is positive definite and $\boldsymbol{w}_i^T \boldsymbol{D} \boldsymbol{w}_i$ can be interpreted as a norm with a metric induced by $\boldsymbol{D}$. Not surprisingly, $\sigma_3 = \sigma_1$ because $\tilde{\boldsymbol{w}}_1$ and $\tilde{\boldsymbol{w}}_3$ differ only in a sign, which does not matter for the norm.

5. Any vector in $\boldsymbol{y} \in \mathbb{R}^3$ can be written as a linear combination of the three normalized eigenvectors, i.e.

$$
\boldsymbol{y} = \sum_{i=1}^{3} \alpha_i \boldsymbol{w}_i \qquad (55)
$$

Derive constraints on the α_i that follow if we impose constraints (2,3) on $\boldsymbol{y}$.

Solution: Constraint (2) yields

$$
\begin{aligned}
0 &\overset{!}{=} \boldsymbol{y}^T \boldsymbol{D}\boldsymbol{1} &(56)\\
&\overset{(55)}{=} \sum_{i=1}^{3} \alpha_i \underbrace{\boldsymbol{w}_i^T \boldsymbol{D}\boldsymbol{1}}_{\overset{(29,33,37)}{=} \delta_{i1}} &(57)\\
&= \alpha_1 &(58)
\end{aligned}
$$

Constraint (3) yields

$$
\begin{aligned}
1 &\overset{!}{=} \boldsymbol{y}^T \boldsymbol{D}\boldsymbol{y} &(59)\\
&\overset{(55)}{=} \sum_{i,j=1}^{3} \alpha_i \alpha_j \underbrace{\boldsymbol{w}_i^T \boldsymbol{D}\boldsymbol{w}_j}_{=\delta_{ij}} &(60)\\
&= \sum_{i=1}^{3} \alpha_i^2 &(61)
\end{aligned}
$$

6. Find weights α_i that are consistent with constraints (2,3) and that minimize the objective function.

   Solution: From the previous step we know that $\alpha_1 = 0$ and $\alpha_2^2 + \alpha_3^2 = 1$. We also know the value of the objective function is $\boldsymbol{y}^T \boldsymbol{L}\boldsymbol{y}$. Combining this with the ansatz (55) yields

$$
\begin{aligned}
\text{minimize} \quad \boldsymbol{y}^T \boldsymbol{L}\boldsymbol{y} &\overset{(55)}{=} \sum_{i,j=2}^{3} \alpha_i \alpha_j \boldsymbol{w}_i^T \boldsymbol{L}\boldsymbol{w}_j &(62)\\
&= \sum_{i,j=2}^{3} \alpha_i \alpha_j \lambda_j \underbrace{\boldsymbol{w}_i^T \boldsymbol{D}\boldsymbol{w}_j}_{=\delta_{ij}} \quad \text{(since } \boldsymbol{w}_j \text{ are eigenvectors)} &(63)\\
&= \sum_{i=2}^{3} \alpha_i^2 \lambda_i &(64)
\end{aligned}
$$

   This is obviously solved by setting $\alpha_2 = 1, \alpha_3 = 0$ and with that we get $\boldsymbol{y}^T \boldsymbol{L}\boldsymbol{y} = \lambda_2$.

   Extra question: What is the value of the objective function for eigenvectors in general?

7. Plot the result of the Laplacian eigenmaps algorithm for this simple example, i.e. visualize the components of $\boldsymbol{w}_2$ and $\boldsymbol{w}_3$ in a 2D plot. Discuss, how the plot changes with a and b.

   Solution: The two eigenvectors can be written as

$$
\begin{aligned}
\boldsymbol{w}_2 &= \frac{1}{\sqrt{ab^2 + a^2 b}} \begin{pmatrix} -b \\ 0 \\ a \end{pmatrix} &(65)\\
&= \frac{\sqrt{2}}{\sqrt{ab}} \frac{1}{\sqrt{2\,(a+b)}} \begin{pmatrix} -b \\ 0 \\ a \end{pmatrix} &(66)\\
&= \frac{1}{\sqrt{2\,(a+b)}} \begin{pmatrix} -\sqrt{\frac{2\,b}{a}} \\ 0 \\ \sqrt{\frac{2\,a}{b}} \end{pmatrix} &(67)\\
& &(68)
\end{aligned}
$$

$$\boldsymbol{w}_3 = \frac{1}{\sqrt{2\,(a+b)}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \tag{69}$$

so that they have a common prefactor. Keeping the prefactor constant one, the vectors can be visualized like in the figure. Changing the prefactor only scales all points.

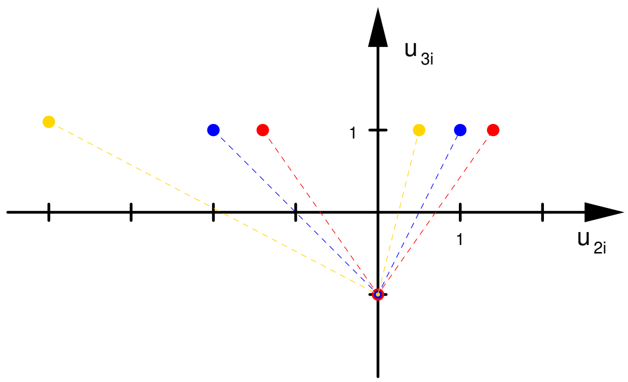

Figure: Laplacian eigenmaps example with three nodes and values for a and b such that $(a+b) = 1/2$ and $b/a = 1$ (red), $b/a = 2$ (blue), and $b/a = 16$ (yellow).

I find it interesting that the components of $\boldsymbol{w}_3$ do not depend at all on the ratio of a and b and that the components of $\boldsymbol{w}_3$ change so little. The direction of change is reasonable; nodes less connected move farther apart.

It is a question whether it is particularly useful in this example to plot the nodes in the 2D space of the first two non-zero eigenvectors. A plot in 1D might be more appropriate.

Extra question: How do the results change if the graph is diconnected, i.e. $0 = a < b$ or even $0 = a = b$?

3.4.5 Exercise: Constraints of the Laplacian eigenmaps

In the Laplacian eigenmap algorithm, each vertex v_i of a graph with symmetric edge weights W_{ij} gets asigned a value w_i with the goal to

$$\text{minimize} \quad \frac{1}{2}\sum_{ij}(w_i - w_j)^2 W_{ij} \tag{1}$$

under the constraints

$$\begin{aligned} \boldsymbol{w}^T\boldsymbol{D}\mathbf{1} &= 0 \quad \text{(weighted zero mean)} &(2)\\ \text{and} \quad \boldsymbol{w}^T\boldsymbol{D}\boldsymbol{w} &= 1 \quad \text{(weighted unit variance)} &(3)\end{aligned}$$

One can show that

$$\frac{1}{2}\sum_{ij}(w_i - w_j)^2 W_{ij} = \boldsymbol{w}^T\boldsymbol{L}\boldsymbol{w} \tag{4}$$

where

$$\begin{aligned} L_{ij} &:= D_{ii}\delta_{ij} - W_{ij} &(5)\\ \text{with} \quad D_{ii} &:= \sum_j W_{ij} &(6)\end{aligned}$$

Show whether it is possibly to find an invertible linear coordinate transformation

$$\tilde{\boldsymbol{w}} = \boldsymbol{T}\boldsymbol{w} \tag{7}$$

such that the optimization problem simplifies to

$$\text{minimize} \quad \tilde{\boldsymbol{w}}^T \hat{\boldsymbol{L}} \tilde{\boldsymbol{w}} \tag{8}$$

under the constraints

$$\tilde{\boldsymbol{w}}^T \mathbf{1} = 0 \quad \text{(zero mean)} \tag{9}$$

$$\text{and} \quad \tilde{\boldsymbol{w}}^T \tilde{\boldsymbol{w}} = 1 \quad \text{(unit variance)} \tag{10}$$

Solution: No, that is not possible. With (7) we would get the following relation between the two optimization problems:

$$\text{minimize} \quad \tilde{\boldsymbol{w}}^T \hat{\boldsymbol{L}} \tilde{\boldsymbol{w}} = \boldsymbol{w}^T \underbrace{\boldsymbol{T}^T \hat{\boldsymbol{L}} \boldsymbol{T}}_{\stackrel{?}{=} \boldsymbol{L}} \boldsymbol{w} \tag{11}$$

under the constraints

$$\tilde{\boldsymbol{w}}^T \mathbf{1} = \boldsymbol{w}^T \underbrace{\boldsymbol{T}^T}_{\stackrel{?}{=} \boldsymbol{D}} \mathbf{1} = 0 \quad \text{(zero mean)} \tag{12}$$

$$\text{and} \quad \tilde{\boldsymbol{w}}^T \tilde{\boldsymbol{w}} = \boldsymbol{w}^T \underbrace{\boldsymbol{T}^T \boldsymbol{T}}_{\stackrel{?}{=} \boldsymbol{D}} \boldsymbol{w} = 1 \quad \text{(unit variance)} \tag{13}$$

$\boldsymbol{T}^T \hat{\boldsymbol{L}} \boldsymbol{T} \stackrel{?}{=} \boldsymbol{L}$ would be ok, but $\boldsymbol{T}^T \stackrel{?}{=} \boldsymbol{D}$ and $\boldsymbol{T}^T \boldsymbol{T} \stackrel{?}{=} \boldsymbol{D}$ cannot be simultaneously fulfilled.

If one would set $\boldsymbol{T} := \boldsymbol{D}^{1/2} = \text{diag}(\sqrt{D_{ii}})$ then the two optimization problems would be almost identical, just in two different coordinate systems. To make them really identical, one would only have to change the original constraint (2) to

$$\boldsymbol{w}^T \boldsymbol{D}^{1/2} \mathbf{1} = 0 \quad \text{(weighted zero mean)} \tag{14}$$

See the symmetric normalized Laplacian matrix in the lecture notes.

Extra question: How would the solutions of the original optimization problem change if one did that.

3.5 The role of the weighted normalization constraint

3.6 Symmetric normalized Laplacian matrix

3.6.1 Exercise: Eigenvectors and -values of the symmetric normalized Laplacian matrix

Consider the Laplacian matrix

$$\boldsymbol{L} = \begin{pmatrix} a & -a & 0 \\ -a & a+b & -b \\ 0 & -b & b \end{pmatrix} \tag{1}$$

with $0 < a, b$.

1. Calculate the symmetric normalized Laplacian matrix.

   Solution: The symmetric normalized Laplacian is

$$\hat{\boldsymbol{L}} := \underline{\boldsymbol{D}}^T \boldsymbol{L} \underline{\boldsymbol{D}} \tag{2}$$
$$= \begin{pmatrix} 1/\sqrt{a} & 0 & 0 \\ 0 & 1/\sqrt{a+b} & 0 \\ 0 & 0 & 1/\sqrt{b} \end{pmatrix} \begin{pmatrix} a & -a & 0 \\ -a & a+b & -b \\ 0 & -b & b \end{pmatrix} \begin{pmatrix} 1/\sqrt{a} & 0 & 0 \\ 0 & 1/\sqrt{a+b} & 0 \\ 0 & 0 & 1/\sqrt{b} \end{pmatrix} \tag{3}$$
$$= \begin{pmatrix} 1 & \frac{-a}{\sqrt{a(a+b)}} & 0 \\ \frac{-a}{\sqrt{a(a+b)}} & 1 & \frac{-b}{\sqrt{b(a+b)}} \\ 0 & \frac{-b}{\sqrt{b(a+b)}} & 1 \end{pmatrix} \tag{4}$$
$$= \begin{pmatrix} 1 & -\sqrt{\frac{a}{a+b}} & 0 \\ -\sqrt{\frac{a}{a+b}} & 1 & -\sqrt{\frac{b}{a+b}} \\ 0 & -\sqrt{\frac{b}{a+b}} & 1 \end{pmatrix} \tag{5}$$

2. Solve the ordinary eigenvalue problem $\hat{\boldsymbol{L}}\hat{\boldsymbol{w}}_i = \lambda_i \hat{\boldsymbol{w}}_i$ for $i = 1, 2, 3$.

   Solution: The roots of the characteristic polynomial are

$$0 \stackrel{!}{=} |\hat{\boldsymbol{L}} - \lambda \boldsymbol{I}| \tag{6}$$
$$= \begin{vmatrix} 1-\lambda & -\sqrt{\frac{a}{a+b}} & 0 \\ -\sqrt{\frac{a}{a+b}} & 1-\lambda & -\sqrt{\frac{b}{a+b}} \\ 0 & -\sqrt{\frac{b}{a+b}} & 1-\lambda \end{vmatrix} \tag{7}$$
$$= (1-\lambda)^3 - (1-\lambda)\frac{b}{a+b} - (1-\lambda)\frac{a}{a+b} \tag{8}$$
$$= (1-\lambda) \cdot \left[(1-\lambda)^2 - \frac{b}{(a+b)} - \frac{a}{(a+b)}\right] \tag{9}$$
$$= (1-\lambda) \cdot \left[(1-\lambda)^2 - 1\right] \tag{10}$$
$$\Longleftarrow \quad \lambda_1 = 0 \tag{11}$$
$$\vee \quad \lambda_2 = 1 \tag{12}$$
$$\vee \quad \lambda_3 = 2 \tag{13}$$

   For $\lambda_1 = 0$ we can chose $\hat{\boldsymbol{w}}_1 = (\sqrt{a}, \sqrt{a+b}, \sqrt{b})^T$, since

$$\boldsymbol{0} \stackrel{?}{=} (\hat{\boldsymbol{L}} - \lambda_1)\hat{\boldsymbol{w}}_1 = \begin{pmatrix} 1 & -\sqrt{\frac{a}{a+b}} & 0 \\ -\sqrt{\frac{a}{a+b}} & 1 & -\sqrt{\frac{b}{a+b}} \\ 0 & -\sqrt{\frac{b}{a+b}} & 1 \end{pmatrix} \begin{pmatrix} \sqrt{a} \\ \sqrt{a+b} \\ \sqrt{b} \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{14}$$

   For $\lambda_2 = 1$ we can chose $\hat{\boldsymbol{w}}_2 = (-b\sqrt{a}, 0, a\sqrt{b})^T$, since

$$\boldsymbol{0} \stackrel{?}{=} (\boldsymbol{L} - \lambda_2)\hat{\boldsymbol{w}}_2 = \begin{pmatrix} 0 & -\sqrt{\frac{a}{a+b}} & 0 \\ -\sqrt{\frac{a}{a+b}} & 0 & -\sqrt{\frac{b}{a+b}} \\ 0 & -\sqrt{\frac{b}{a+b}} & 0 \end{pmatrix} \begin{pmatrix} -b\sqrt{a} \\ 0 \\ a\sqrt{b} \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{15}$$

For $\lambda_3 = 2$ we can chose $\hat{\boldsymbol{w}}_1 = (\sqrt{a}, -\sqrt{a+b}, \sqrt{b})^T$, since

$$\mathbf{0} \stackrel{?}{=} (\boldsymbol{L} - \lambda_3 \boldsymbol{D})\hat{\boldsymbol{w}}_3 = \begin{pmatrix} -1 & -\sqrt{\frac{a}{a+b}} & 0 \\ -\sqrt{\frac{a}{a+b}} & -1 & -\sqrt{\frac{b}{a+b}} \\ 0 & -\sqrt{\frac{b}{a+b}} & -1 \end{pmatrix} \begin{pmatrix} \sqrt{a} \\ -\sqrt{a+b} \\ \sqrt{b} \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{16}$$

Extra question: How, do you think, did I get the eigenvectors of this eigenvalue equation?

3.7 Random walk normalized Laplacian matrix +

3.8 Summary of mathematical properties

4 Algorithms

4.1 Similarity graphs

4.2 Laplacian eigenmaps (LEM)

4.2.1 Motivation

4.2.2 Objective

4.2.3 Algorithm

4.2.4 Sample applications

4.3 Locality preserving projections (LPP)

4.3.1 Linear LPP

4.3.2 Sample application

4.3.3 Nonlinear LPP

4.4 Spectral clustering

4.4.1 Objective

4.4.2 Algorithm

4.4.3 Sample application