

Министерство образования и науки Российской Федерации  
Федеральное государственное автономное образовательное учреждение высшего  
образования

САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ ИТМО

Факультет систем управления и робототехники

**Лабораторная работа № 5**  
**"Спектральная теория графов"**

по дисциплине Практическая линейная алгебра

Выполнила: студентка гр. **R3238**

**Нечаева А. А.**

Преподаватель: *Перегудин Алексей Алексеевич*

Санкт-Петербург, 2023-2024

# 1 Кластеризация социальной сети

Для начала построим модель небольшой социальной сети, где каждый пользователь обозначен одной из 18 вершин графа, а ребра показывают дружбу между людьми.

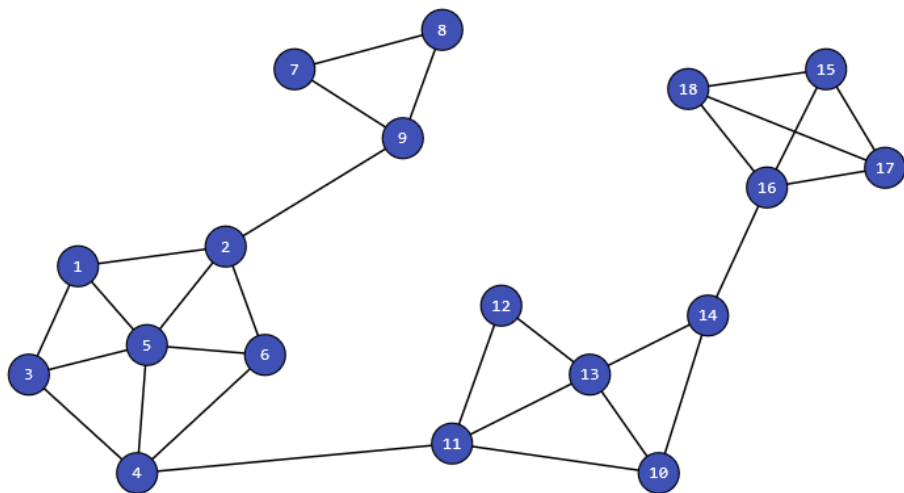


Рис. 1. Модель социальной небольшой сети

Соответствующая графу на рисунке 1 матрица Лапласа: Все собственные числа и соответствующие им собственные векторы приведены в разделе *Приложение* в конце документа.

$$\begin{pmatrix}
 3 & -1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 -1 & 4 & 0 & 0 & -1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 -1 & 0 & 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & -1 & 4 & -1 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 -1 & -1 & -1 & -1 & 5 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & -1 & 0 & -1 & -1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & -1 & 0 & 0 & 0 & 0 & -1 & -1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & -1 & 0 & -1 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 4 & -1 & -1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 4 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 3 & 0 & -1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & -1 & -1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 4 & -1 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 3 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1
 \end{pmatrix}$$

Рис. 2. Матрица Лапласа

Выберем  $k = 4$  компоненты кластеризации графа. И составим матрицу из 4 собственных векторов матрицы Лапласа, соответствующих самым маленьким собственным числам.

$$V = \begin{pmatrix} 1 & -0,541 & -0,569 & 1,766 \\ 1 & -0,609 & -0,184 & 1,141 \\ 1 & -0,470 & -0,736 & 1,718 \\ 1 & -0,329 & -0,803 & 0,976 \\ 1 & -0,497 & -0,614 & 1,569 \\ 1 & -0,493 & -0,594 & 1,470 \\ 1 & -0,933 & 1,570 & -1,573 \\ 1 & -0,933 & 1,570 & -1,573 \\ 1 & -0,850 & 1,096 & -0,799 \\ 1 & 0,375 & -0,885 & -1,742 \\ 1 & 0,173 & -1,026 & -1,333 \\ 1 & 0,276 & -1,160 & -2,090 \\ 1 & 0,355 & -0,944 & -1,819 \\ 1 & 0,564 & -0,419 & -1,217 \\ 1 & 1 & 1 & 1 \\ 1 & 0,911 & 0,698 & 0,508 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad (1)$$

Рассмотрим строки составленной матрицы  $V$  как точки пространства  $R^4$ . Применим к этим точкам метод кластеризации  $k - means$ , реализованный на языке *Python*.

```
kmeans = KMeans(n_clusters=4)
```

```
d = np.array([[1, -0.541, -0.569, 1.766],
               [1, -0.609, -0.184, 1.141],
               [1, -0.470, -0.736, 1.718],
               [1, -0.329, -0.803, 0.976],
               [1, -0.497, -0.614, 1.569],
               [1, -0.493, -0.594, 1.470],
               [1, -0.933, 1.570, -1.573],
               [1, -0.933, 1.570, -1.573],
               [1, -0.850, 1.096, -0.799],
               [1, 0.375, -0.885, -1.742],
               [1, 0.173, -1.026, -1.333],
```

```

[1, 0.276, -1.160, -2.090],
[1, 0.355, -0.944, -1.819],
[1, 0.564, -0.419, -1.217],
[1, 1, 1, 1],
[1, 0.911, 0.698, 0.508],
[1, 1, 1, 1],
[1, 1, 1, 1]
])
kmeans.fit(d)
print(kmeans.labels_)

```

*Листинг 1. Фрагмент кода кластеризации.*

В результате выполнения программы получили:

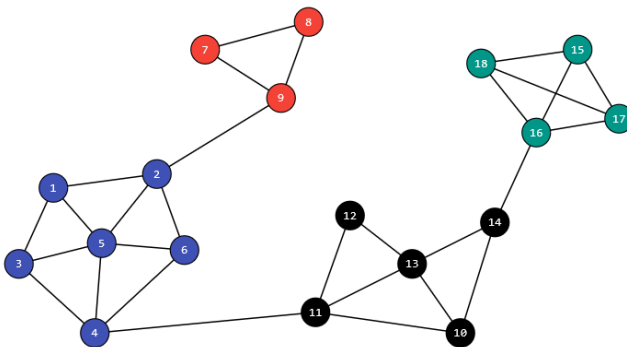
```

[0 0 0 0 0 0 1 1 1 2 2 2 2 2 3 3 3]
Process finished with exit code 0

```

*Рис. 3. Результат кластеризации*

Соответствующий вычисленной кластеризации граф изображен на рисунке 4.



*Рис. 4. Граф раскрашенный в соответствии с полученной кластеризацией  $k = 4$*

Так же проведена кластеризация для значений  $k = 5$  и  $k = 6$ , результаты представлены на рисунках 5 и 6 соответственно.

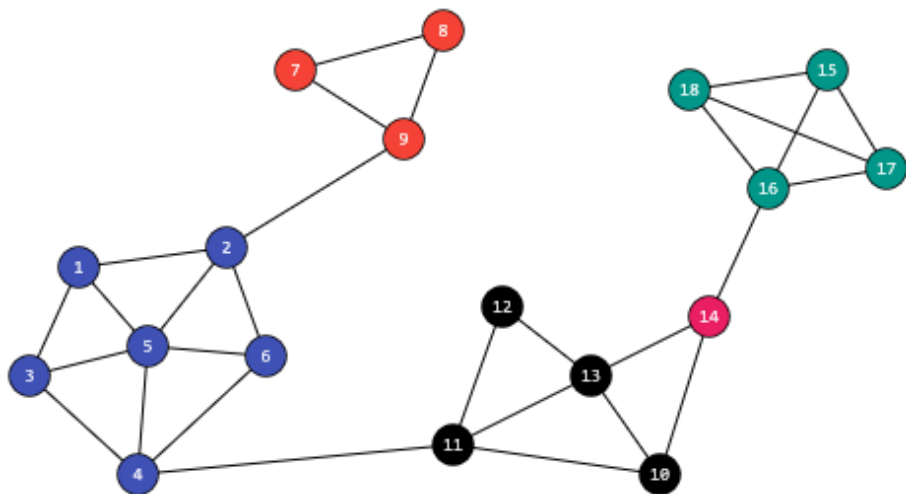


Рис. 5. Граф раскрашенный в соответствии с полученной кластеризацией  $k = 5$

1. Результат, полученный для кластеризации при  $k = 4$  иллюстрирует изначально заложенное разделение вершин графа на 4 сообщества.

2. При  $k = 5$  глобально остаются те же 4 больших кластера, что и в предыдущем случае, но еще выделяется 1 кластер представленный вершиной номер 14, расположенной как бы на границе двух кластеров. Заметим, что у этой вершины всего 3 связи: 2 связывают с черным кластером, 1 с зеленым, то есть у нее достаточно слабая связь с каждым из больших смежных кластеров. Для сравнения рассмотрим, например, вершину под номером 11 также расположенную на границе крупных кластеров: синего и черного, однако она "крепче" связана с черным кластером, благодаря 3 связям, поэтому не была выделена в отдельный кластер на данном этапе.

3. При  $k = 6$  к предыдущему результату добавляется еще 1 кластер, в этот раз представленный вершиной 9 расположенной на границе коричневого и синего кластеров. Заметим, что она так же имеет 3 связи: 2 с коричневым кластером и 1 с синим.

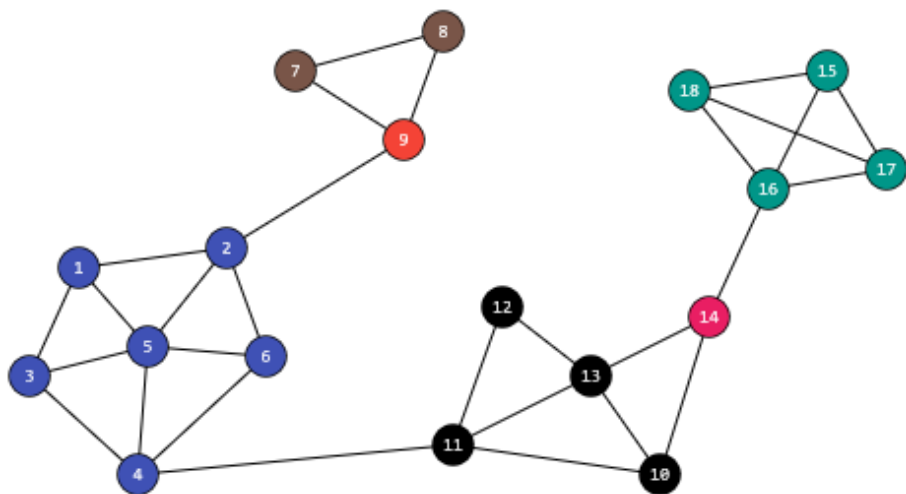


Рис. 6. Граф раскрашенный в соответствии с полученной кластеризацией  $k = 6$

**4. Почему это работает?** Вообще, суть кластеризации состоит в группировке элементов множества по какому-то критерию сходства, в данном случае по степени связанности вершин друг с другом. Матрица Лапласа отражает информацию о связях вершин между собой (матрицу смежности) и о количестве вершин, с которыми связана каждая (матрица степеней). Собственные числа матрицы Лапласа всегда вещественные и неотрицательные, их можно отсортировать в порядке неубывания. Наибольший интерес представляют минимальные из них: например, кратность числа  $\lambda = 0$  отражает количество компонент связности графа, последующие числа – насколько связи в графе прочны. Метод  $k$  – *means* основывается на поиске центроид (центральных элементов для каждого кластера) и сравнении расстояний от каждого элемента до определенной центроиды. Обратим внимание на матрицу  $V$  из собственных векторов (уравнение 1), заметим, что если воспринимать строки матрицы, как некие координаты вершины в пространстве размерности  $k$ , можно разделить эти точки на  $k$  групп, расстояния в которых между каждой парой точек будет минимальным. Поэтому этот довольно успешный метод так распространен в анализе данных.

## 2 Google PageRank алгоритм

Придумаем связный ориентированный граф из 10 вершин и 25 стрелок (рисунок 7)

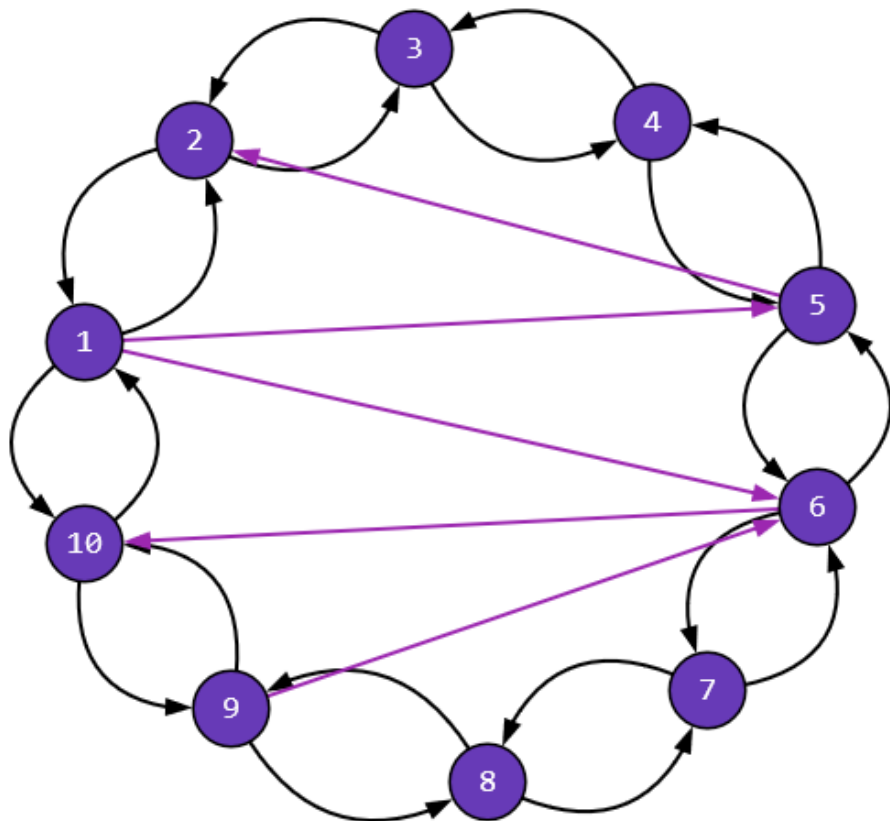


Рис. 7. Граф связей веб-страниц

Составим матрицу

$$M = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1\ 10} \\ m_{21} & m_{22} & \dots & m_{2\ 10} \\ \dots & \dots & \dots & \dots \\ m_{10\ 1} & m_{10\ 2} & \dots & m_{10\ 10} \end{pmatrix},$$



где  $m_{ij}$  – отношение числа ссылок на  $j$ -й странице, которые ведут на  $i$ -ю страницу, к общему числу ссылок на  $j$ -й странице.

$$M = \begin{pmatrix} 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0.25 & 0 & 0.5 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0.5 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0.5 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0.25 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \end{pmatrix}$$

Собственный вектор матрицы  $M$  соответствующий наибольшему собственному числу  $\lambda = 1$ :

$$v = \begin{pmatrix} \frac{54}{49} \\ \frac{59}{49} \\ \frac{209}{196} \\ \frac{13}{14} \\ \frac{465}{392} \\ \frac{75}{56} \\ \frac{153}{196} \\ \frac{131}{196} \\ \frac{327}{392} \\ 1 \end{pmatrix} \approx \begin{pmatrix} 1.1020 \\ 1.2041 \\ 1.0663 \\ 0.9286 \\ 1.1862 \\ 1.3393 \\ 0.7806 \\ 0.6684 \\ 0.8342 \\ 1 \end{pmatrix}$$

Ранжируем веб-страницы (вершины графа) в соответствии с *PageRank*-алгоритм при отсутствии затухания ( $d = 1$ ).

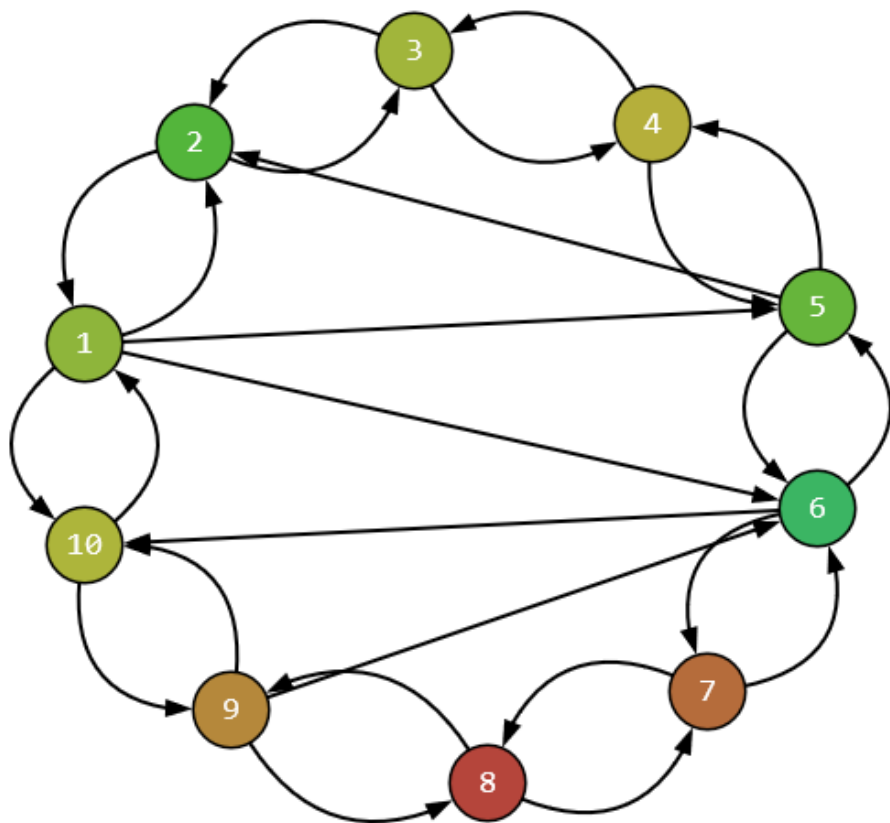


Рис. 8. Граф ранжированных веб-страниц

Рисунок 8 иллюстрирует полученный результат: ярко-зеленый – самый высокий приоритет, красный – самый низкий. Соответствующий список вершин:

- |    |   |     |    |
|----|---|-----|----|
| 1. | 6 | 6.  | 10 |
| 2. | 2 | 7.  | 4  |
| 3. | 5 | 8.  | 9  |
| 4. | 1 | 9.  | 7  |
| 5. | 3 | 10. | 8  |

*Почему это работает?*

### 3 Приложение

Все собственные числа соответствующие им собственные векторы для матрицы Лапласа из задания 1.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \lambda_1 = 0
 \quad
 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \lambda_2 = 3
 \quad
 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \lambda_3 = 4
 \quad
 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \lambda_4 = 4$$

$$\begin{pmatrix} (-0,541) \\ (-0,609) \\ (-0,470) \\ (-0,329) \\ (-0,497) \\ (-0,493) \\ (-0,933) \\ (-0,933) \\ (-0,850) \\ (0,375) \\ (0,173) \\ (0,276) \\ (0,355) \\ (0,564) \\ 1 \\ (0,911) \\ 1 \\ 1 \end{pmatrix}, \lambda_5 = 0.089
 \quad
 \begin{pmatrix} (-0,569) \\ (-0,184) \\ (-0,736) \\ (-0,803) \\ (-0,614) \\ (-0,594) \\ (1,570) \\ (1,570) \\ (1,096) \\ (-0,885) \\ (-1,026) \\ (-1,160) \\ (-0,944) \\ (-0,419) \\ 1 \\ (0,698) \\ 1 \\ 1 \end{pmatrix}, \lambda_6 = 0.302
 \quad
 \begin{pmatrix} (1,766) \\ (1,141) \\ (1,718) \\ (0,976) \\ (1,569) \\ (1,470) \\ (-1,573) \\ (-1,573) \\ (-0,799) \\ (-1,742) \\ (-1,333) \\ (-2,090) \\ (-1,819) \\ (-1,217) \\ 1 \\ (0,508) \\ 1 \\ 1 \end{pmatrix}, \lambda_7 = 0.492$$

$$\begin{pmatrix} (-0,483) \\ (-0,304) \\ (-0,109) \\ (0,480) \\ (-0,117) \\ (0,055) \\ (0,102) \\ (0,102) \\ (-0,092) \\ (-3,902) \\ (1,177) \\ (6,448) \\ (-0,559) \\ (-4,895) \\ 1 \\ (-0,904) \\ 1 \\ 1 \end{pmatrix}, \lambda_8 = 1.904
 \quad
 \begin{pmatrix} (-112,670) \\ (51,599) \\ (-122,421) \\ (36,220) \\ (5,800) \\ (162,222) \\ (-18,290) \\ (-18,290) \\ (26,025) \\ (2,964) \\ (11,522) \\ (-16,445) \\ (-4,567) \\ (-5,244) \\ 1 \\ (-1,423) \\ 1 \\ 1 \end{pmatrix}, \lambda_9 = 2.423
 \quad
 \begin{pmatrix} (-15,810) \\ (-19,115) \\ (13,664) \\ (19,505) \\ (0,545) \\ (3,014) \\ (7,572) \\ (7,572) \\ (-12,795) \\ (2,437) \\ (8,335) \\ (-8,656) \\ (-2,365) \\ (-5,214) \\ 1 \\ (-1,690) \\ 1 \\ 1 \end{pmatrix}, \lambda_{10} = 2.690$$

$$\begin{pmatrix} (0, 661) \\ (0, 984) \\ (-0, 783) \\ (0, 071) \\ (-0, 445) \\ (-1, 660) \\ (-0, 872) \\ (-0, 872) \\ (2, 065) \\ (3, 889) \\ (2, 932) \\ (-2, 241) \\ (0, 133) \\ (-4, 497) \\ 1 \\ (-2, 368) \\ 1 \\ 1 \end{pmatrix}, \lambda_{11} = 3.368
\begin{pmatrix} (9, 817) \\ (-0, 245) \\ (-8, 229) \\ (-5, 286) \\ (1, 399) \\ (5, 733) \\ (6, 255) \\ (6, 255) \\ (-17, 018) \\ (4, 618) \\ (-0, 381) \\ (-0, 251) \\ (0, 812) \\ (-3, 760) \\ 1 \\ (-2, 721) \\ 1 \\ 1 \end{pmatrix}, \lambda_{12} = 3.721
\begin{pmatrix} (-3, 665) \\ (-0, 297) \\ (5, 560) \\ (-4, 796) \\ (0, 190) \\ (3, 295) \\ (-0, 093) \\ (-0, 093) \\ (0, 324) \\ (5, 299) \\ (-6, 706) \\ (2, 647) \\ (0, 121) \\ (-1, 299) \\ 1 \\ (-3, 488) \\ 1 \\ 1 \end{pmatrix}, \lambda_{13} = 4.488$$

$$\begin{pmatrix} (-0, 288) \\ (0, 067) \\ (0, 517) \\ (-0, 629) \\ (-0, 047) \\ (0, 327) \\ (0, 013) \\ (0, 013) \\ (-0, 050) \\ (-3, 826) \\ (-0, 251) \\ (-2, 370) \\ (7, 043) \\ (0, 348) \\ 1 \\ (-3, 866) \\ 1 \\ 1 \end{pmatrix}, \lambda_{14} = 4.866
\begin{pmatrix} (1, 758) \\ (-2, 275) \\ (-1, 818) \\ (2, 390) \\ (0, 041) \\ (-0, 068) \\ (-0, 287) \\ (-0, 287) \\ (1, 236) \\ (0, 292) \\ (-1, 273) \\ (0, 995) \\ (-2, 016) \\ (2, 616) \\ 1 \\ (-4, 305) \\ 1 \\ 1 \end{pmatrix}, \lambda_{15} = 5.305
\begin{pmatrix} (-6, 448) \\ (10, 339) \\ (-0, 417) \\ (2, 122) \\ (5, 312) \\ (-7, 523) \\ (1, 245) \\ (1, 245) \\ (-5, 430) \\ (-0, 064) \\ (-0, 263) \\ (0, 831) \\ (-2, 531) \\ (2, 945) \\ 1 \\ (-4, 363) \\ 1 \\ 1 \end{pmatrix}, \lambda_{16} = 5.363$$

$$\begin{pmatrix} (-0, 141) \\ (2, 993) \\ (5, 426) \\ (-6, 661) \\ (-8, 035) \\ (4, 280) \\ (0, 273) \\ (0, 273) \\ (-1, 295) \\ (-3, 217) \\ (9, 878) \\ (-0, 961) \\ (-6, 291) \\ (5, 209) \\ 1 \\ (-4, 734) \\ 1 \\ 1 \end{pmatrix}, \lambda_{17} = 5.734
 \quad
 \begin{pmatrix} (-13, 211) \\ (-42, 779) \\ (-7, 706) \\ (-55, 201) \\ (93, 510) \\ (1, 373) \\ (-2, 829) \\ (-2, 829) \\ (14, 873) \\ (-10, 039) \\ (37, 402) \\ (-5, 598) \\ (-13, 572) \\ (8, 864) \\ 1 \\ (-5, 257) \\ 1 \\ 1 \end{pmatrix}, \lambda_{18} = 6.257$$