拓扑数据处理编程作业报告

黎治圻 陈羽 吴佳昱 郯许和

2022 年 12 月 31 日

摘要

本次前沿数学专题讨论课程我们小组主要进行学习的是拓扑数据处理。本次编程作业我们选取的是同调理论以及持续同调进行实现。主要参考的是 E 组论文中的 c 部分，程序实现了流形生成过程的可视化，同时绘制了多个不同实例的 persistent barcode(PB) 和 persistent diagram(PD), 可以很好地反映输入实例的拓扑性质。代码主要实现了 Vietoris-Rips 复形和 Witness 复形的构造，同时具有较强的计算稀疏矩阵的能力。

目录

1 简介

持续同调相关的函数库在黎治圻同学的 github 上 ZhiqiLi-CG (Zhiqi Li(黎治圻)), 具体而言:

计算代码库:

https://github.com/ZhiqiLi-CG/zqBasicMath

homology 计算实例:

https://github.com/ZhiqiLi-CG/zqBasicMath/tree/main/examples/math/homology

可视化代码库:

https://github.com/ZhiqiLi-CG/zqVisualization

persistent homology 实例:

https://github.com/ZhiqiLi-CG/zqBasicMath/tree/main/examples

persistent homology 的实例打包在 zqVisualization 中，其具体使用说明和运行结果见报告余下内容。

组员的具体分工为:

代码编写: 黎治圻、陈羽

实验报告: 吴佳昱、郏许和

2 Homology 相关计算

2.1 预备知识

定义 2.1. *Let $v_0, \cdots, v_k$ be linearly independent points in R^n. The k-simplex denoted by $[v_0, \cdots, v_k]$ is the topological space given by the set*

$$\{\sum_{i=0}^{k} t_i v_i | t_0 + t_1 + \cdots + t_k = 1, t_i \geq 0\}$$

with the topology induced by the Euclidian metric. The numbers t_i are the coordinates of the point $x = \sum_i t_i v_i \in [v_0, \cdots, v_k]$.

定义 2.2. *Set X be a finite set. A simplicial complex K on X is a set of subsets of X such that:*

1.$\{x\} \in K$ for any $x \in X$.

2.if $\sigma \in K$ and $\tau \subset \sigma$, then $\tau \in K$

定义 2.3. *We use the symbol K_i to denote the set of simplices in K of dimension i, for example $K_0 = X$.*

定义 2.4. *Let R be a commutative ring, and X be a set and K be a simplicial complex on X. Define $C_n(K) := \bigoplus_{\sigma \in K_n} R$. ∂_n is a map between free $R-$module $C_n(K)$ and $C_{n-1}(K)$:*

$$\partial_n(e_\sigma) = e_{d_0\sigma} - e_{d_1\sigma} + \cdots + (-1)^n e_{d_n\sigma}$$

定义 2.5. *The quotient $Ker\partial_n / Im\partial_{n+1}$ is called the homology group of K with coefficients in R and is denoted by $H_n(K, R)$. The isomorphism type of the $R-$module $H_n(K, R)$ does not depend on the ordering we have chosen on X. For our purposes we are mainly interested in the case R being a finites field F_p.*

定义 2.6. *The dimension of the $R-$vector space denoted by $H_n(K, R)$ is called the $n-th$ Betti number of K with respect to R and is denoted by $\beta_n(K, R)$.*

2.2 库的使用

2.2.1 单纯复形的表示

```
    std::vector<std::vector<int>> complex_index{
                {0},{1},{2},{3},
                {0,1},{1,2},{0,2},{2,3},{1,3},
```

```
                {0,1,2}
        };
        std::vector<int> index{ 0,1,2,3 };
        zq::Simplical_Complex<int> sc(&index[0], 4,
            complex_index);
```

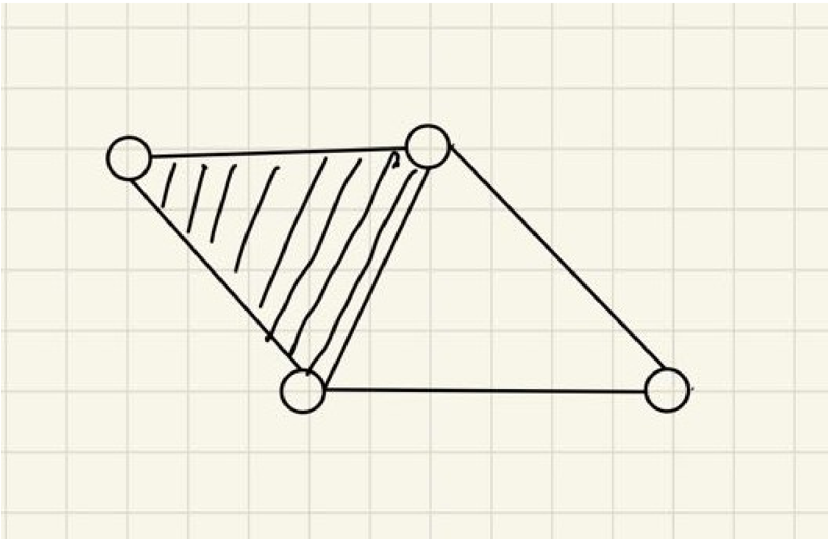

图 1: 单纯复形的表示

2.2.2 K_n 的计算

```
    std::vector<zq::Simplex<int>> result;
        sc.NChain(i, result);
```

2.2.3 计算 $i-1$ 和 i 之间的边缘矩阵

```
    zq::DenseMatrix<int> bounadry_m = sc.
        BoundaryMatrix(i-1, i );
```

2.2.4 边缘矩阵约化为 Smith 标准型

```
zq::DenseMatrix<int> bounadry_m = sc.
    ReducedBoundaryMatrix(i-1, i, rank, nullity);
```

2.2.5 计算维数小于 3 的 Betti 数

```
std::vector<int> betti=sc.BettiNumber(3);
```

3 持续同调相关计算

3.1 简介

持续同调 (persistent homology) 作为拓扑数据处理 (topological data analysis, TDA) 领域的重要内容，是利用拓扑学对数据进行处理研究的方法。基于持续同调, 小组编写了相关的函数库, 并以数个复形为实例考察了它们的 persistent barcode(PB) 和 persistent diagram(PD).

3.2 预备知识

定义 3.1 (Victoris-Rips complex.). *Define $V(X,\epsilon)$ to be a simplicial complex on set X consisting of these subsets $\sigma \subset X$ where $d(x,y) < \epsilon$ for any x, y, σ.*

定义 3.2 (Persistence pairing.). *If, in addition to the Betti numbers, we wish to compute the persisitent pairing we need to be sensitive to the ordering of the simplices. We begin with a filtration of the complex, $\varnothing = K_0 \subset K_1 \subset \cdots \subset K_m = K$, and sort the simplices to get a compatible total ordering of the simplices in $K, \sigma_1, \sigma_2, \cdots, \sigma_n$. Compatible means that:*

-the simplices in each complex K_i in the filtration precede the ones in $K - K_i$;

-the faces of a simplex precede the simplex.

Instead of parceling out the face relation we do the computations wholesale on the combined boundary matrix define by $D[i, j] = 1$ if σ_i is a codimension 1 face of σ_j and $D[i, j] = 0$ otherwise. We restrict oueselves to column addition.

3.3 库的使用

3.3.1 VR 复形的构建

```
std::vector<std::vector<int>> results;
            zq::VRComplexConstruct(
                    epsilon_list[i],
                    points.value,
                    simplex_dimension,
                    results
            );
            complex_list.push_back(zq::
               Simplical_Complex<zq::Vec3f>(&(
               points.value[0]), points.Dim(),
               results));
```

3.3.2 计算 Persistent Data

```
zq::CalculatePersistentData(
            epsilon_list,
            max_epsilon,
```

```
                    complex_list ,
                    epsilon_interval ,
                    feture_type
```

3.3.3 绘制 barcode

```
    zq::homology::DrawPersistentBarCode(
        epsilon_interval ,
        feture_type ,
        max_epsilon ,
        display_max_epsilon ,
        win_x,
        win_y,
        win_ox,
        win_oy
    );
```

3.3.4 绘制 persistent diagram

```
    zq::homology::DrawPersistentDiagram(
        epsilon_interval ,
        feture_type ,
        max_epsilon ,
        display_max_epsilon ,
        max_epsilon*0.1
        win_x,
        win_y,
        win_ox,
```

```
            win_oy
      );
```

4 程序使用方法和实例

4.1 程序使用方法说明

1. 在解压得到的文件夹 **zqVisualization** 中,
运行脚本 **make_project.py**。

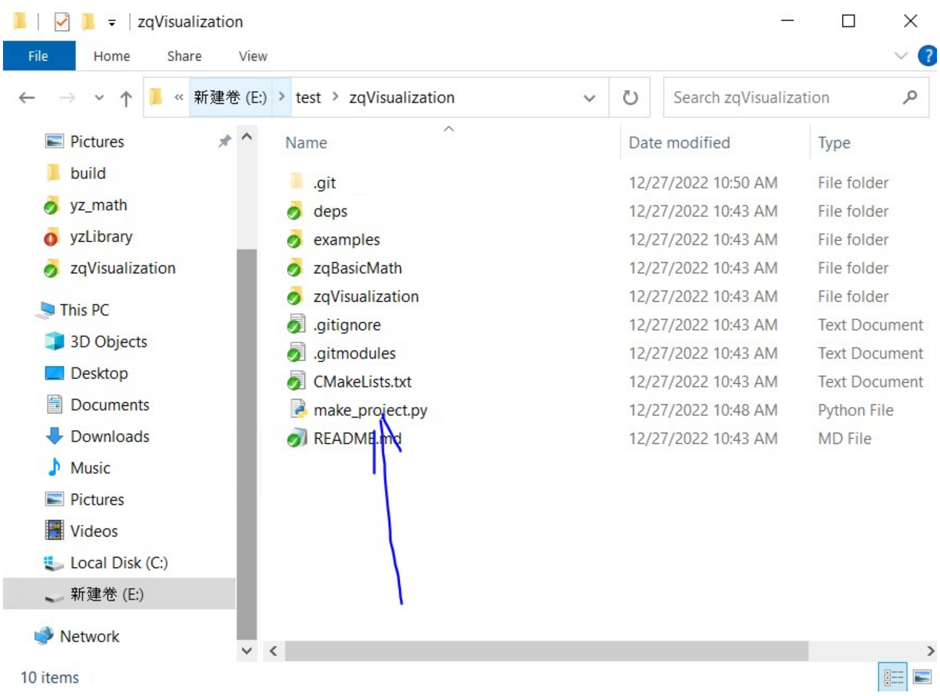

图 2:

2. 运行脚本后, 文件夹 **zqVisualization** 中出现新的文件夹 **build**, 打开后看到如下内容:

图 3:

3. 使用 **Visual Studio** 打开 **ALL_BUILD.vcxproj** 并将 **example** 下的 **persistent homology** 设为启动项目。

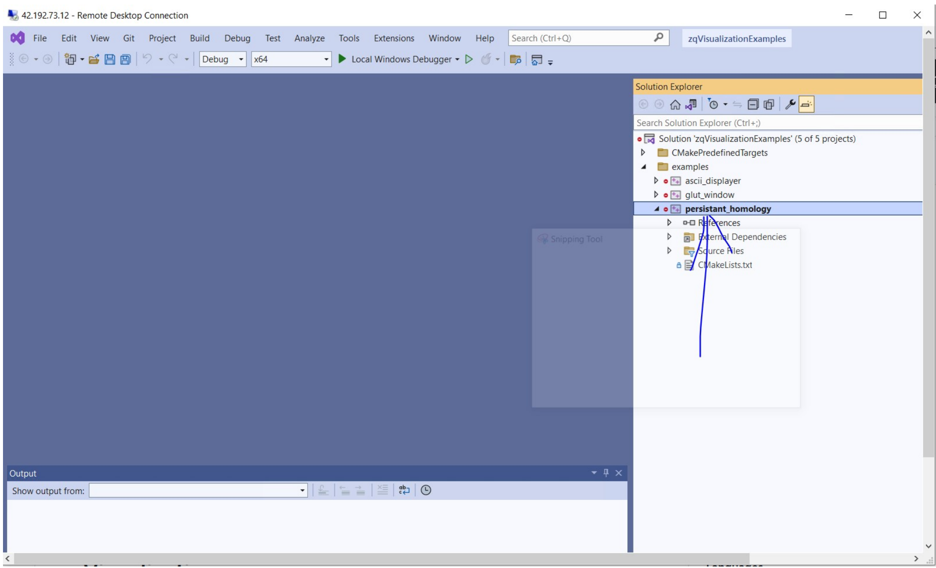

图 4:

4. 这时运行程序并输入 0,1 或 2,3 即可看到对应的复形例子, 同时可以看到其 **PB(persistent barcode)** 和 **PD(persistent diagram)**. 通过 A

键和 D 键可以调整 ϵ 的大小, 观察对应 ϵ 下复形集合的变化和 PB 的情况。

4.2 具体实例及输入

输入: 0,1,2,3 分别对应如下四个准备好地复形例子, 其中前三个是利用 **Vietoris-Rips complex** 给出的构造, 第四个是利用 **Witness** 版本的 **VR complex** 给出的构造:

4.2.1 例 1(输入 0): 圆周上均匀分布的离散点集

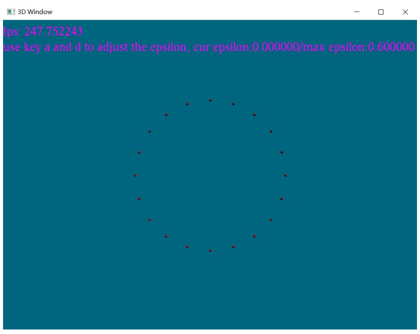

图 5: 正面.

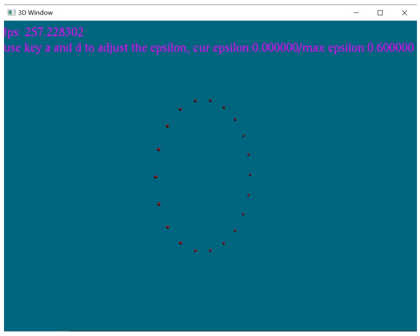

图 6: 侧面.

4.2.2 例 2(输入 1): 对圆周上均匀分布的离散点集施加扰动

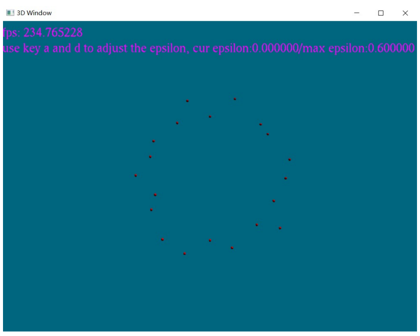

图 7: 正面.

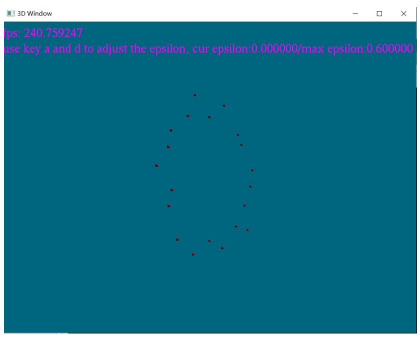

图 8: 侧面.

4.2.3 例 3(输入 2): 点集基本分布在两个垂直的平面上

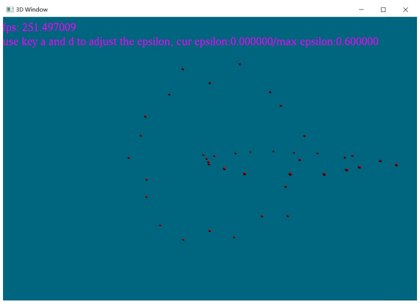

图 9: 正面.

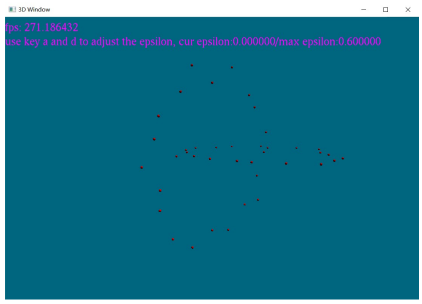

图 10: 侧面.

4.2.4 例 4(输入 3):Witness 版本

使用 witness 版本的复型构造, 其中比较鲜艳的红色是 witness 点集, 暗红色的是 witness 点集外的点。

图 11:

4.3 输出

4.3.1 例 1: 圆周上均匀分布的离散点集

图 12: persistent barcode.

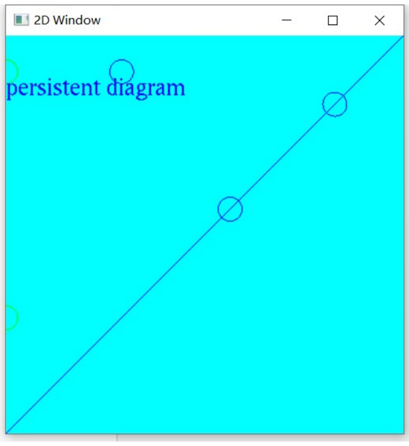

图 13: persistent diagram.

运用 A,D 键调整 ϵ 的大小:(PB 中的红线代表对应的 ϵ 值).

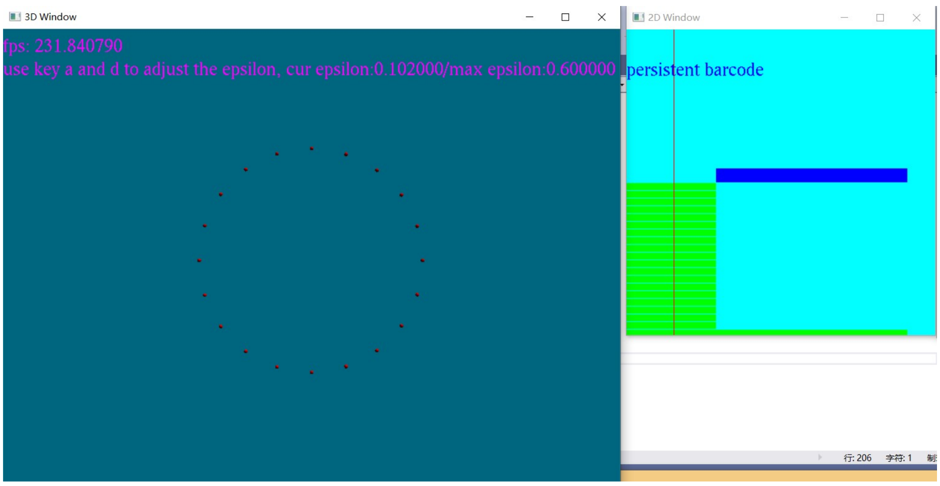

图 14: 例 1

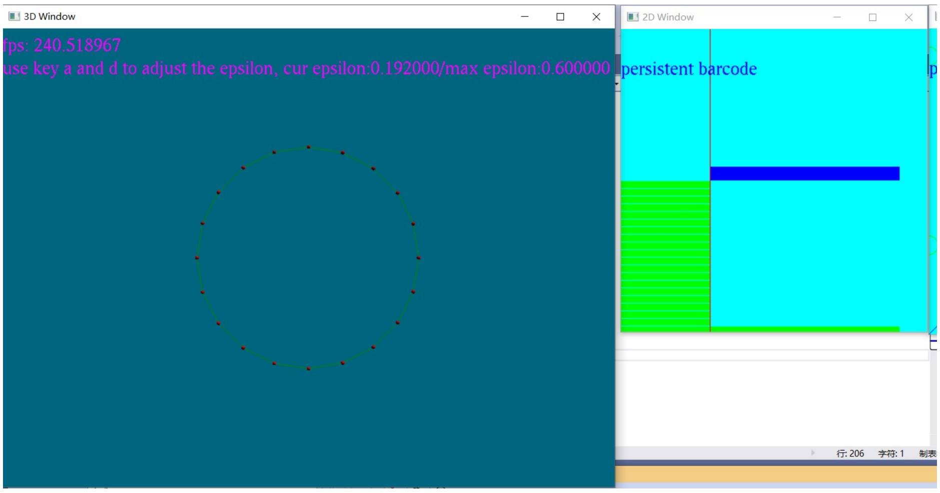

图 15: 例 1

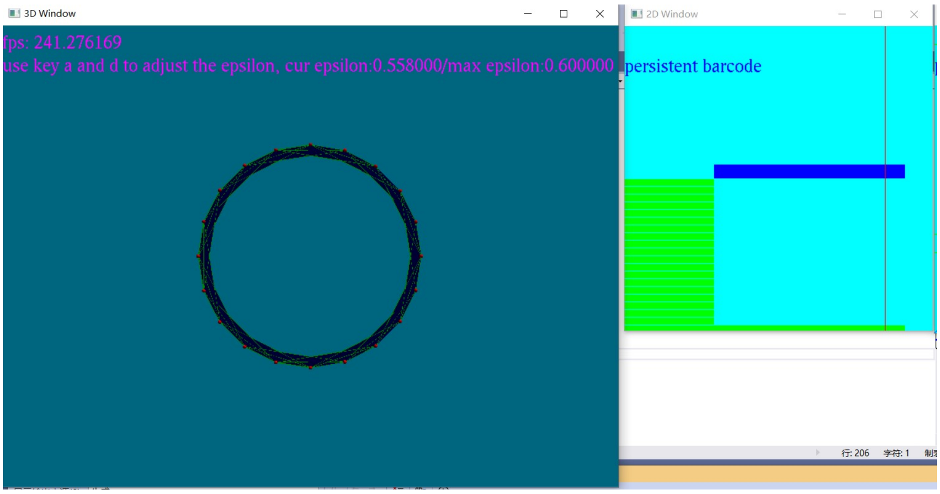

图 16: 例 1

4.3.2 例 2: 对圆周上均匀分布的离散点集施加扰动

运用 A,D 键调整 ϵ 的大小:(PB 中的红线代表对应的 ϵ 值).

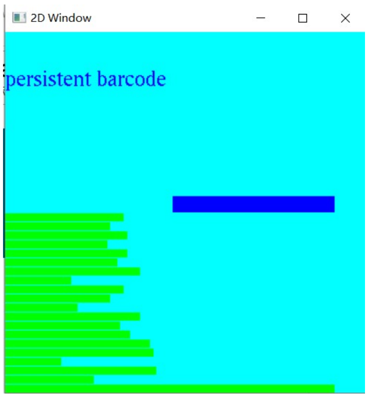

图 17: persistent barcode.

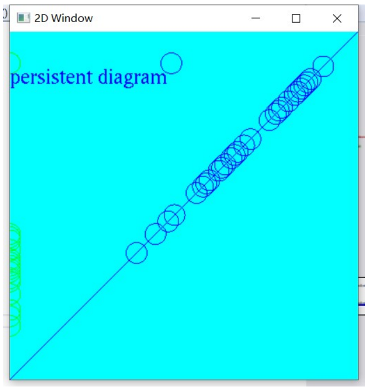

图 18: persistent diagram.

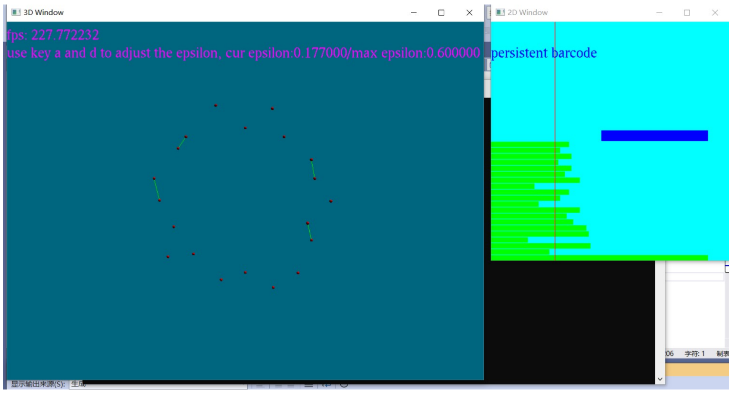

图 19: 例 2

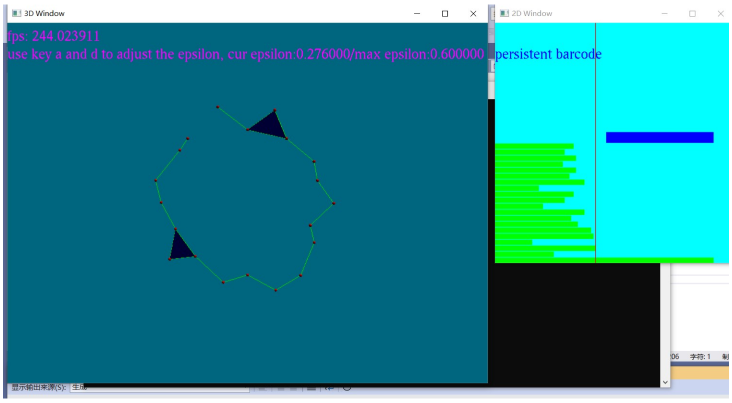

图 20: 例 2

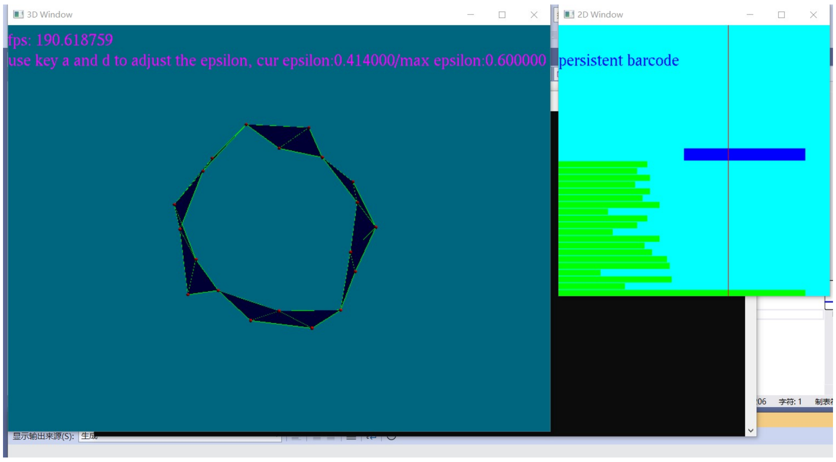

图 21: 例 2

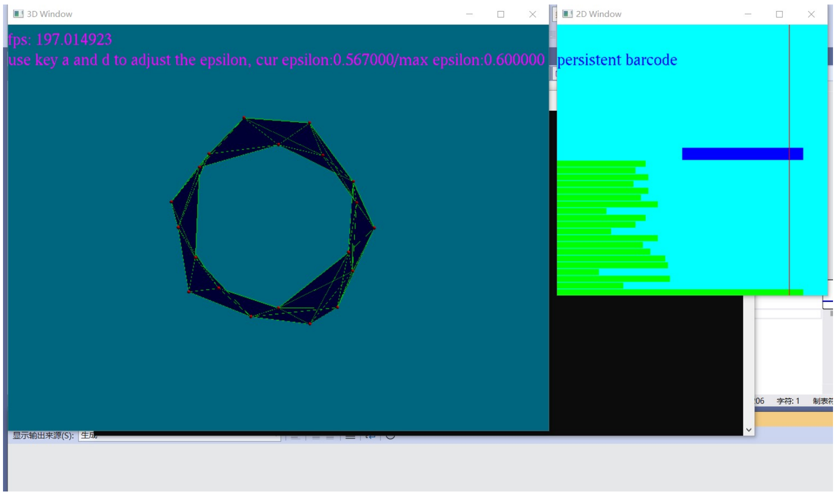

图 22: 例 2

4.3.3 例 3: 点集基本分布在两个垂直的平面上

运用 A,D 键调整 ϵ 的大小:(PB 中的红线代表对应的 ϵ 值).

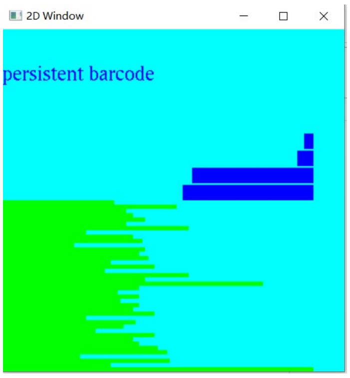

图 23: persistent barcode.

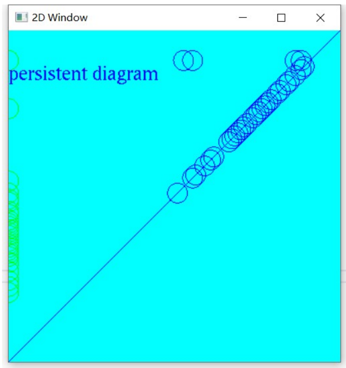

图 24: persistent diagram.

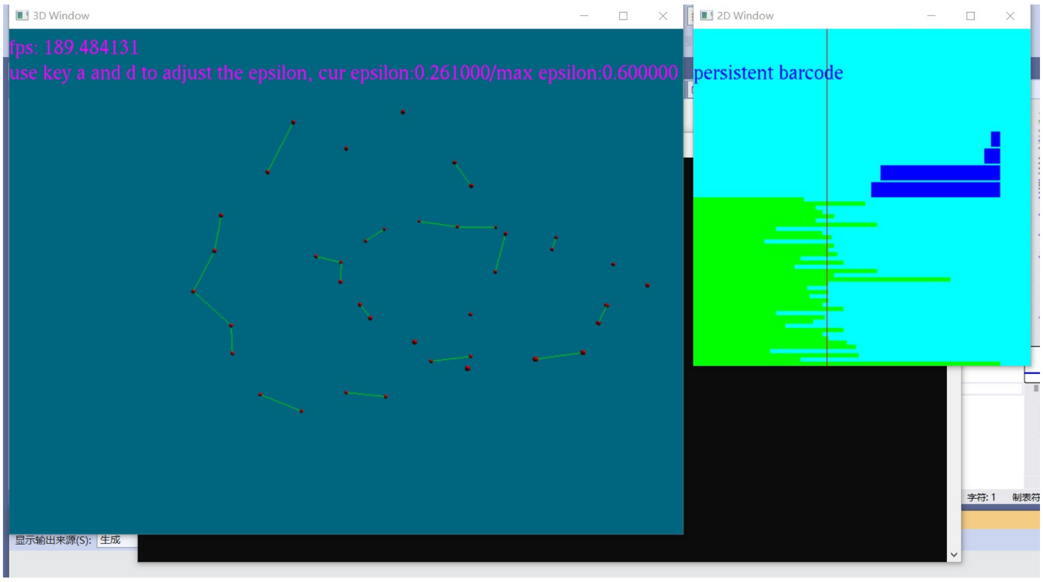

图 25: 例 3

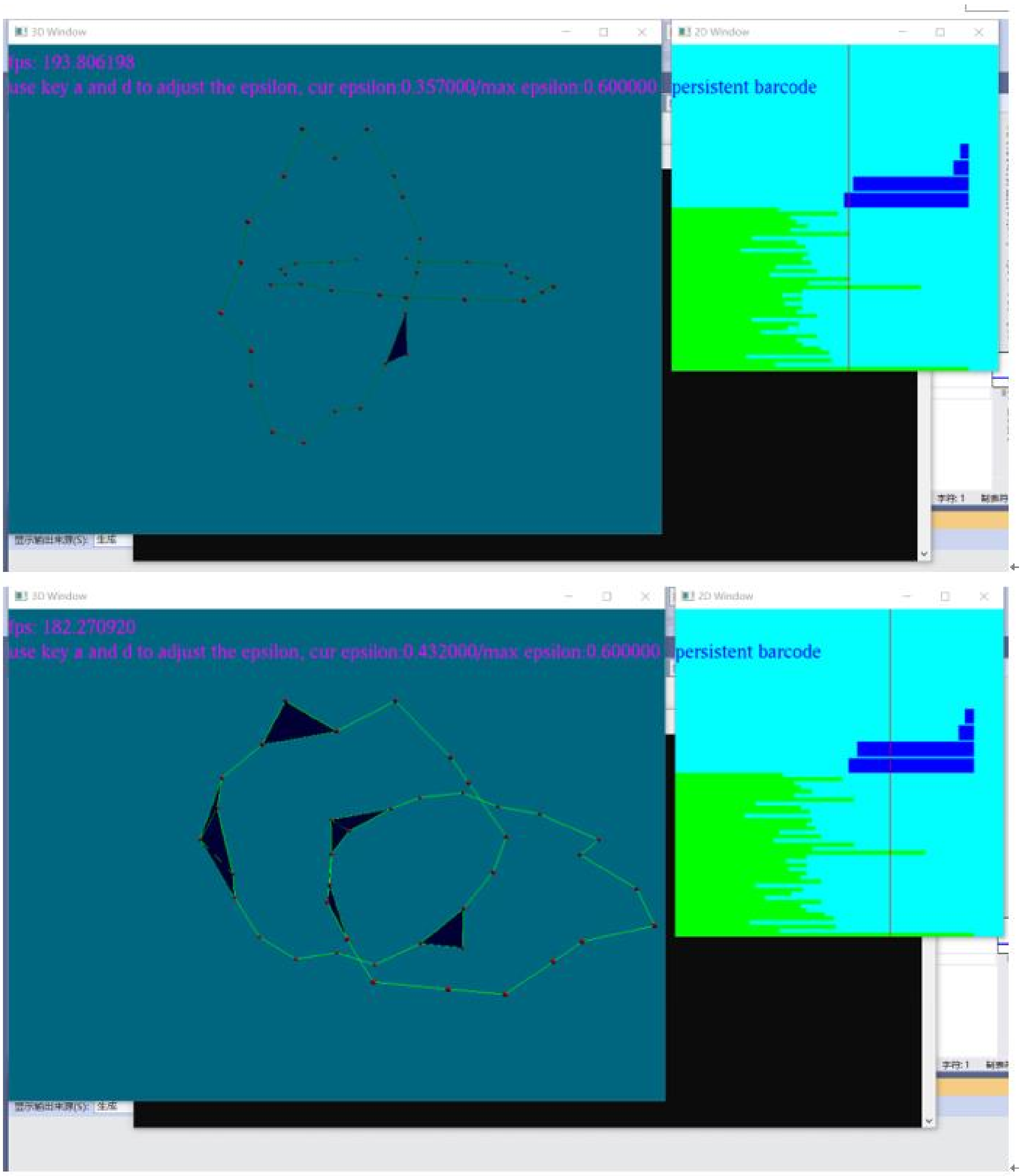
3D Window
fps: 193.806198
use key a and d to adjust the epsilon, cur epsilon:0.357000/max epsilon:0.600000
2D Window
persistent barcode
3D Window
fps: 182.270920
use key a and d to adjust the epsilon, cur epsilon:0.432000/max epsilon:0.600000
2D Window
persistent barcode

图 26: 例 3

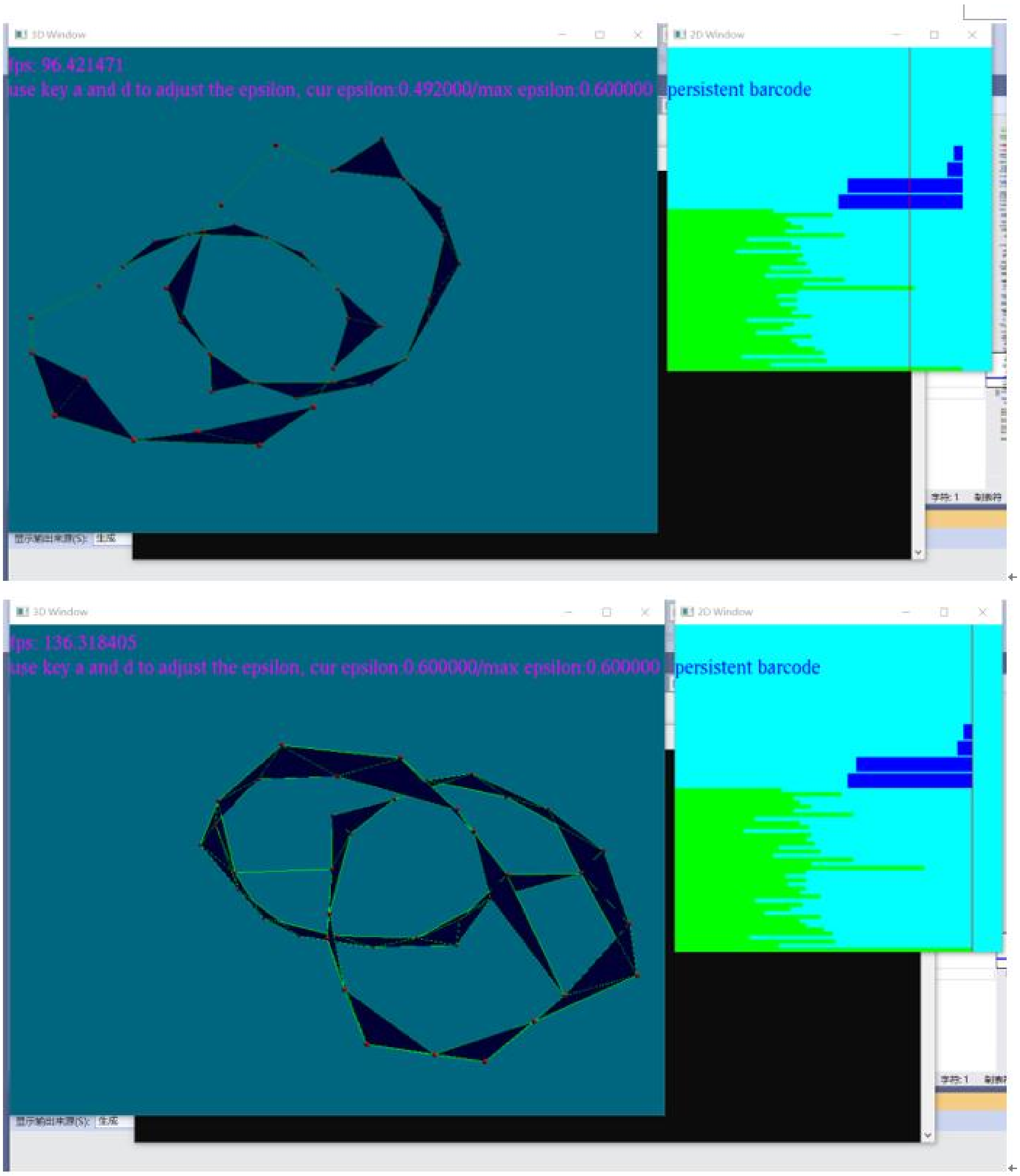

图 27: 例 3

4.3.4 例 4:Witness 版本

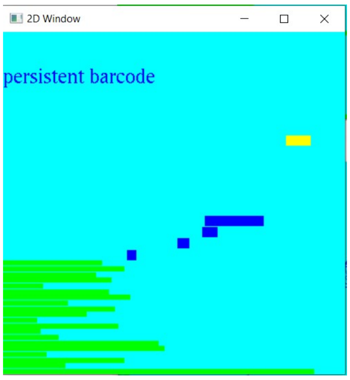

图 28: persistent barcode.

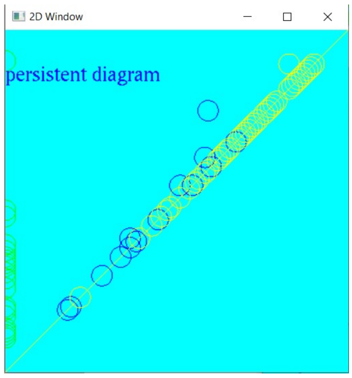

图 29: persistent diagram.

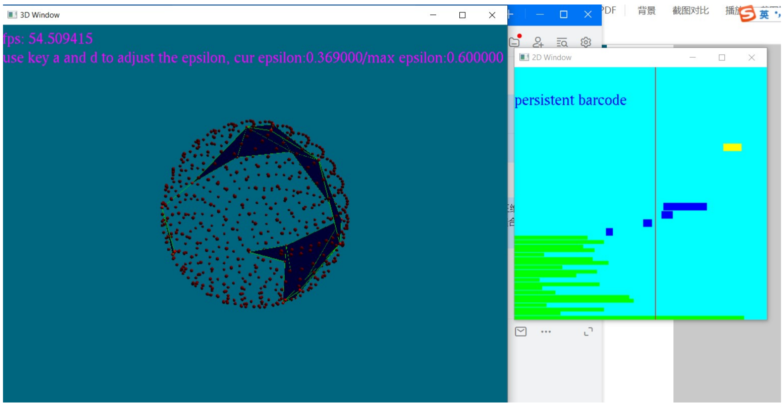

图 30: 例 4

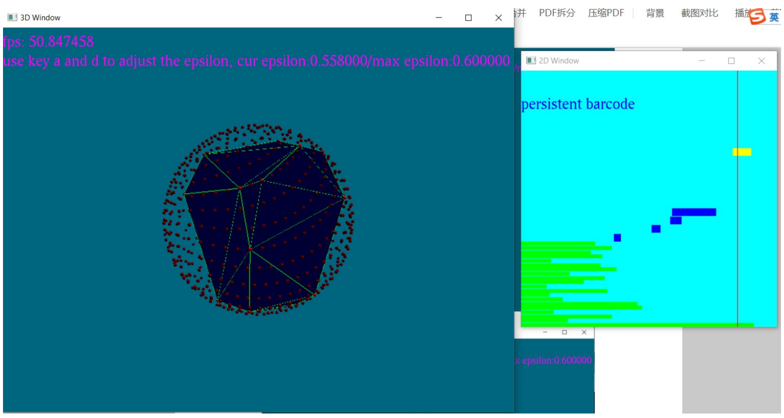

图 31: 例 4

4.4 注意事项

程序所使用的需为 **Visual Studio2022**(一定要确保使用的 VS 版本支持 C++17),

如版本不同需要修改脚本 **make_project.py** 中的对应内容:

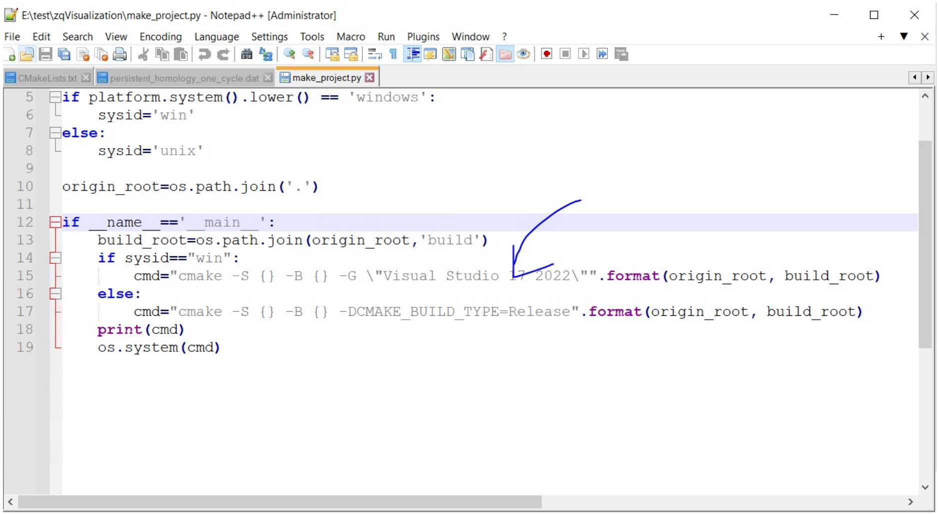

图 32: 注意事项

参考文献

[1] Herbert Edelsbrunner and John Harer. *Persistent Homology —a Survey*[M]. Stanford University,2008.

[2] Zomorodian, A., Carlsson, G. *Computing Persistent Homology*[M].Discrete Comput Geom,2005.

[3] Herbert Edelsbrunner and Dmitriy Morozov. *Persistent Homology: Theory and Practice*[M]. EPJ Data Sci,2013.

[4] Edelsbrunner, Letscher and Zomorodian . *Topological Persistence and Simplification*[M]. Discrete Comput Geom,2008.

A 持续同调复形构造

a.Vietoris-Rips complex： Define $V(X,\epsilon)$ to be a simplicial complex on the set X consisting of these subsets $\sigma \subset X$ where $d(x,y) < \epsilon$ for any $x, y \in \sigma$.

b.Witness version: Choose a subset $X_W \subset X$ of witnesses of X. Define$V_W(X,\epsilon)$, the Witness version of the Vietoris-Rips complex, to be a complex on X_W that consists of these $\sigma \subset X_W$, where for any $x, y \in \sigma$, there is $z \in X$ such that $d(x,y) < \epsilon$ and $d(y,z) < \epsilon$.