MATH:7450 (22M:305) Topics in Topology: Scientific and Engineering Applications of Algebraic Topology

Oct 21, 2013: Cohomology

Fall 2013 course offered through the University of Iowa Division of Continuing Education

Isabel K. Darcy, Department of Mathematics, Applied Mathematical and Computational Sciences, University of Iowa

http://www.math.uiowa.edu/~idarcy/AppliedTopology.html
from dionysus import Simplex, ZigzagPersistence, \
    vertex_cmp, data_cmp \\
#, enable_log

complex = {
    Simplex((0,), 0): None,  # A
    Simplex((1,), 1): None,  # B
    Simplex((2,), 2): None,  # C
    Simplex((0,1), 2.5): None,  # AB
    Simplex((1,2), 2.9): None,  # BC
    Simplex((0,2), 3.5): None,  # CA
    Simplex((0,1,2), 5): None}  # ABC
print "Complex:"
for s in sorted(complex.keys()):
  print s
print

triangle idarcy$ python2.7 triangle-zigzag.py
Complex:
<0>
<1>
<2>
<0, 1> 2.500000
<1, 2> 2.900000
<0, 2> 3.500000
<0, 1, 2>
#enable_log("topology/persistence")

zz = ZigzagPersistence()

# Add all the simplices
b = 1
for s in sorted(complex.keys(), data_cmp):
    print "%d: Adding %s" % (b, s)

1: Adding <0>
2: Adding <1>
3: Adding <2>
# Add all the simplices

```python
b = 1
for s in sorted(complex.keys(), data_cmp):
    print "%d: Adding %s" % (b, s)
    i, d = zz.add([complex[ss] for ss in s.boundary], b)
    complex[s] = i
    if d:
        print "Interval (%d, %d)" % (d, b-1)
    b += 1
```
# Add all the simplices

\[ b = 1 \]

for s in sorted(complex.keys(), data_cmp):
    print "\%d: Adding %s" % (b, s)
    i, d = zz.add([complex[ss] for ss in s.boundary], b)
    complex[s] = i
    if d:
        print "Interval (%d, %d)" % (d, b-1)
    b += 1

4: Adding \langle 0, 1 \rangle 2.500000
Interval (2, 3)
5: Adding \langle 1, 2 \rangle 2.900000
Interval (3, 4)
6: Adding \langle 0, 2 \rangle 3.500000
7: Adding \langle 0, 1, 2 \rangle
Interval (6, 6)
# Remove all the simplices
for s in sorted(complex.keys(), data_cmp, reverse = True):
    print "%d: Removing %s" % (b, s)
    d = zz.remove(complex[s], b)
    del complex[s]
    if d: print "Interval (%d, %d)" % (d, b-1)
    b += 1

8: Removing <0, 1, 2>  
9: Removing <0, 2> 3.500000
   Interval (8, 8)  
10: Removing <1, 2> 2.900000
11: Removing <0, 1> 2.500000
12: Removing <2>
Interval (10, 11)
13: Removing <1>
Interval (11, 12)
14: Removing <0>
Interval (1, 13)
Persistent Cohomology and Circular Coordinates

Vin de Silva · Dmitriy Morozov · Mikael Vejdemo-Johansson

NKI (2002) breast cancer data:
gene expression levels of 24,000 from 272 tumors

→ Data = 272 points living in $\mathbb{R}^{24000}$
NKI (2002) breast cancer data:
gene expression levels of 24,000 from 272 tumors

→ Data = 272 points living in $\mathbb{R}^{24000}$

In reality one cleans the data first, removing unreliable data, etc.

One may also remove dimensions (genes) that appear irrelevant.
NKI (2002) breast cancer data: gene expression levels of 24,000 from 272 tumors

→ Data = 272 points living in $\mathbb{R}^{24000}$

Since our dataset consists of only 272 points, it can’t be 24000-dimensional.
For example, suppose our very fake data is

\[
\{ (9, 8, 7, 6, 5, 4, 3, 2, 3) t_i \mid i = 1, 2, 3, \ldots 1000 \} = \\
\{ (9, 8, 7, 6, 5, 4, 3, 2, 3), (18, 16, 14, 12, 10, 8, 6, 4, 6), \ldots \}
\]
Dimensionality Reduction:

Given dataset $D \subseteq \mathbb{R}^N$

Want: embedding $f: D \to \mathbb{R}^n$ where $n \ll N$

which “preserves” the structure of the data.

Example:

$$\{ (9, 8, 7, 6, 5, 4, 3, 2, 3) t_i \mid i = 1, 2, 3, \ldots 1000 \} \subseteq \mathbb{R}^9$$

$f: D \to \mathbb{R}, \quad f((9, 8, 7, 6, 5, 4, 3, 2, 3) t_i) = t_i$
Example: Principle component analysis (PCA)
Many reduction methods:

\[ f_1: D \rightarrow \mathbb{R}, \ f_2: D \rightarrow \mathbb{R}, \ldots \ f_n: D \rightarrow \mathbb{R} \]

\[ (f_1, f_2, \ldots f_n): D \rightarrow \mathbb{R}^n \]

Many are linear, \( M: \mathbb{R}^N \rightarrow \mathbb{R}^n, \ Mx = y \)

But there are also non-linear dimensionality reduction algorithms.
$f_1: D \rightarrow \mathbb{R}, f_2: D \rightarrow \mathbb{R}, \ldots f_n: D \rightarrow \mathbb{R}$

$(f_1, f_2, \ldots f_n): D \rightarrow \mathbb{R}^n$
Goal

\( f_1: D \rightarrow S^1, f_2: D \rightarrow S^1, \ldots f_n: D \rightarrow S^1 \)
Note $S^1$ is a ring:
Cohomology and circular coordinates:

Let \( X \) = finite simplicial complex.

\( X^0 = \) set of 0-simplices = vertices.
\( X^1 = \) set of 1-simplices = edges.
\( X^2 = \) set of 2-simplices = faces.
Homology

Let $X$ = finite simplicial complex.

$C_0$ = set of 0-chains = linear combinations of vertices

$C_1$ = set of 1-chains = linear combinations of edges.

$C_2$ = set of 2-chains = linear combinations of faces.
Cohomology and circular coordinates:

Let $X$ = finite simplicial complex, $R$ a ring.

$C^0 = \text{set of 0-cochains} = \{ f : C_0 \rightarrow R \mid f \text{ homomorphism}\}$

$C^1 = \text{set of 1-cochains} = \{ f : C_1 \rightarrow R \mid f \text{ homomorphism}\}$

$C^2 = \text{set of 2-cochains} = \{ f : C_2 \rightarrow R \mid f \text{ homomorphism}\}$
Cohomology and circular coordinates:

Let \( X \) = finite simplicial complex, \( R \) a ring.

\[ C^0(X, R) = \text{set of 0-cochains} = \{ f : X^0 \rightarrow R \} \]

\[ C^1(X, R) = \text{set of 1-cochains} = \{ f : X^1 \rightarrow R \} \]

\[ C^2(X, R) = \text{set of 2-cochains} = \{ f : X^2 \rightarrow R \} \]

. . .
Proposition 1: Let $\alpha \in C^1(X;\mathbb{Z})$ be a cocycle. Then there exists a continuous function

$$\theta : X \rightarrow S^1$$

which maps each vertex to 0, and each edge $ab$ around the entire circle with winding number $\alpha(ab)$. 
Proposition 2: Let $\alpha \in C^1(X;\mathbb{R})$ be a cocycle. Suppose there exists $\alpha \in C^1(X;\mathbb{Z})$ and $f \in C^0(X;\mathbb{R})$ such that $\alpha = \alpha + d_0 f$. Then there exists a continuous function

$$\theta : X \to S^1$$

which maps each edge $ab$ linearly to an interval of length $\alpha(ab)$, measured with sign.
$\theta : X \rightarrow \mathbb{R}/\mathbb{Z} = S^1$
coboundary maps

\[ d_0 : \quad C^0 = \{ f : X^0 \to \mathbb{R} \} \quad \to \quad C^1 = \{ f : X^1 \to \mathbb{R} \} \]

\[ (d_0 f)(ab) = f(b) - f(a) \]

\[ d_1 : \quad C^1 = \{ f : X^1 \to \mathbb{R} \} \quad \to \quad C^2 = \{ f : X^2 \to \mathbb{R} \} \]

\[ (d_1 \alpha)(abc) = \alpha(bc) - \alpha(ac) + \alpha(ab). \]
Let $\alpha \in C^i$, 
\[ d_i : C^i = \{ f : X^i \to R \} \to C^{i+1} = \{ f : X^{i+1} \to R \} \]

$\alpha$ is a **cocycle** if $d_i \alpha = 0$.

$\alpha$ is a **coboundary** if there exists $f$ in $C^{i-1}$ s.t. $d_{i-1} f = \alpha$

Note $d_1 d_0 f = 0$ for any $f \in C^0$.

\[(d_0 f)(ab) = f(b) - f(a)\]

\[(d_1 d_0 f)(abc) = d_0 f(bc) - d_0 f(ac) + d_0 f(ab) = d_0 f(c) - d_0 f(b) - d_0 f(c) + d_0 f(a) + d_0 f(b) - d_0 f(a)\]
\[ d_1 d_0 f = 0 \text{ implies } \Im (d_0) \subseteq \Ker (d_1). \]

\[ H^1 (X;A) = \Ker (d_1)/ \Im (d_0). \]

Two cocycles \( \alpha, \beta \) are \textit{cohomologous} if \( \alpha - \beta \) is a coboundary.
$d_1 d_0 f = 0$ implies $\text{Im} \ (d_0) \subseteq \text{Ker} \ (d_1)$.

$H^1 (X;A) = \text{Ker} \ (d_1)/\text{Im} \ (d_0)$. 
\[ d_1 d_0 f = 0 \text{ implies } \text{Im} \ (d_0) \subseteq \text{Ker} \ (d_1). \]

\[ H^1 (X;A) = \text{Ker} \ (d_1)/ \text{Im} \ (d_0). \]
$C^0(X, R) = \text{set of } 0\text{-cochains} = \{ f : X^0 \rightarrow R \}$
\[ C^0(X, R) = \text{set of 0-cochains} = \{ f : X^0 \to R \} \]

\[ = \{ f \mid f(v_1) = r_1, f(v_2) = r_2, f(v_3) = r_3; (r_1, r_2, r_3) \in \mathbb{R}^3 \} \]

\[ C^1(X, R) = \text{set of 1-cochains} = \{ f : X^1 \to R \} \]

\[ = \{ f \mid f(e_1) = r_1, f(e_2) = r_2, f(e_3) = r_3; (r_1, r_2, r_3) \in \mathbb{R}^3 \} \]
\( C^0(X, R) = \text{set of 0-cochains} = \{ f : X^0 \to R \} \)

\[ = \{ f \mid f(v_1) = r_1, f(v_2) = r_2, f(v_3) = r_3; (r_1, r_2, r_3) \text{ in } R^3 \} \]

\( C^1(X, R) = \text{set of 1-cochains} = \{ f : X^1 \to R \} \)

\[ = \{ f \mid f(e_1) = r_1, f(e_2) = r_2, f(e_3) = r_3; (r_1, r_2, r_3) \text{ in } R^3 \} \]

\[ H^i (X; A) = \text{Ker } (d_i) / \text{Im } (d_{i-1}) \]

\[ 0 \to C^0 \to C^1 \to 0 \]