

УДК 515.146.27

МЕТОД СТАЛИХ ГОМОЛОГІЙ ТОПОЛОГІЧНОГО АНАЛІЗУ ДАНИХ

І. А. Юрчук, канд. фіз.-мат. наук

Національний авіаційний університет

iyurch@ukr.net

На основі схеми аналізу даних методом сталих гомологій з фільтрацією комплексами Вісторіса-Ріпа розроблено блок-схеми алгоритмів побудови межової та зведеної матриць фільтрації.

Ключові слова: метод сталих гомологій; комплекс Вісторіса-Ріпа.

Based on the scheme of persistent homology method with the Vietoris-Rips complexes filtration, the flowcharts of construction of a boundary matrix and its reduced one are obtained.

Keywords: persistent homology method; Vietoris-Rips complex.

Вступ

Під час обробки даних використовують різні методи аналізу отриманої вибірки, кожен з яких дає можливість зробити певний висновок або про існування різних типів залежностей між величинами, або про структуру вибірки. Зокрема, кластерний аналіз дозволяє розбити її на групи зі схожими характеристиками. Проте його методи мають істотні недоліки у практичному застосуванні. Як приклад, число кластерів залежить від однорідності вибірки та метрики, яку зазвичай вибирають покладаючись на досвід.

З розвитком топології з'явився сучасний напрям у аналізі даних — топологічний аналіз, що є аналогом кластерного і має істотні переваги в застосуванні, оскільки дозволяє зробити висновки про глобальну структуру об'єкта, виходячи з його локальних характеристик. Ефективність застосування топологічних методів до розв'язання задач аналізу даних обґрунтовано в серії робіт американських математиків А. Зомородяна [1; 2], Р. Гріста [3; 4], Е. К. Дея [5], Г. Еделсбрунера [6; 7], Г. Карлсона [8] та ін.

Якщо в результаті експерименту чи обробки інформації отримано скінченну вибірку з деякого невідомого простору, то, припускаючи, що він є топологічним, можна встановити його структуру. Таке припущення є більш загальним та значно слабшим, ніж існування структури многовида, метрики чи наявності кривизни на досліджуваному просторі. Основні топологічні методи аналізу даних, поняття та теореми, що їх стосуються, наведені в працях [1; 2; 7; 8].

Метод сталих гомологій є одним з важливих методів топологічного аналізу даних. Зокрема, його ефективність доведено при обробці зображень [3], розв'язанні задачі про втечу з обмеженої території, що має недостатнє покриття сенсорами руху [4], побудові макро- та мікроскопічних молекулярних структур [7] та структури потенціалу електростатичного акумулятора [1].

Постановка проблеми

Нехай задано скінченну вибірку S , $S \subseteq X$, де X — деякий невідомий простір. Топологічний аналіз вибірки складається з двох основних етапів: апроксимації простору X комбінаторною структурою K та обчислення топологічного інваріанта для K , за яким роблять висновок про структуру X .

Апроксимація простору відбувається за допомогою фільтрації, складовими якої є симплекціальні комплекси (Чеха, Делоне, Вісторіса-Ріпа, кубічні та ін.), побудовані на елементах вибірки S .

Другий етап — вибір топологічного інваріанта (комбінаторного чи алгебричного) та його обчислення.

Мета статті — побудова блок-схем алгоритмів дослідження структури скінченної вибірки, використовуючи її апроксимацію комплексами Вісторіса-Ріпа та обчислення сталих гомологій за допомогою їх матричного представлення.

Основні поняття теорії сталих гомологій

Визначення 1. Симплексом або n -симплексом називають опуклу множину $(n+1)$ точок v_0, v_1, \dots, v_n евклідового простору R^m , що не лежать в одній n -вимірній гіперплощині.

Точки v_i назвемо вершинами симплекса, а кожен n -симплекс, побудований на вершинах v_i , будемо позначати як набором його вершин $[v_0, \dots, v_n]$, так і літерою Δ_n .

У малих розмірностях симплексами є точка (0-симплекс), відрізок (1-симплекс), трикутник (2-симплекс) та тетраедр (3-симплекс).

Якщо з множини $(n+1)$ вершин деякого n -симплекса видалити одну, то отриманий n -симплекс називається його *гранню*.

Визначення 2. Межею $\text{bd } \Delta^n$ n -симплекса Δ^n називається об'єднання всіх його граней.

Знаходження межі симплекса можна формально описати за допомогою межового оператора ∂ , що означений за таким правилом

$$\begin{aligned} \partial \Delta^n &= \partial([v_0, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n]) = \\ &= \sum_i (-1)^i [v_0, \dots, v_{i-1}, v_{i+1}, \dots, v_n]. \end{aligned}$$

Визначення 3. Симплекціальним комплексом K називається скінченний набір симплексів, які задовольняють умову: якщо $\alpha \in K$ та $\beta \subseteq \alpha$, то $\beta \in K$.

Означимо комплекс Вісторіса-Ріпа.

Нехай задана деяка множина точок $X = \{x_1, x_2, \dots, x_k\} \subseteq R^n$. Зафіксуємо $\varepsilon > 0$. Припустимо, що існує підмножина $\tilde{X} = \{x_{i_1}, x_{i_2}, \dots, x_{i_m}\}$, де $m \leq k$, така, що для довільної пари точок x_{i_i} та x_{i_n} справедливо $d(x_{i_i}, x_{i_n}) \leq \varepsilon$, де d — евклідова відстань між точками. На точках $x_{i_1}, x_{i_2}, \dots, x_{i_m}$ підмножини \tilde{X} побудуємо $(m-1)$ -вимірний симплекс $[x_{i_1}, x_{i_2}, \dots, x_{i_m}]$, що відповідає ε .

Комплекс Вісторіса-Ріпа (ВР комплекс) — це об'єднання всіх симплексів, що побудовані за описаною вище схемою. Зрозуміло, що ВР комплекс залежить від вибору числа ε . Максимальна розмірність ВР комплексу, який побудовано на k точках, дорівнює числу $(k-1)$.

На рис. 1 зображено ВР комплекси, які побудовані на фіксованих чотирьох точках та відповідають $\varepsilon_0 = 0,5$, $\varepsilon_1 = 1$, $\varepsilon_2 = 2$ та $\varepsilon_3 = 3$. Видно, що при $\varepsilon_0 = 0,5$ структура ВР комплексу залишається дискретною і не відрізняється від вибірки, а при $\varepsilon_3 = 3$ комплекс має одну зв'язну компоненту.

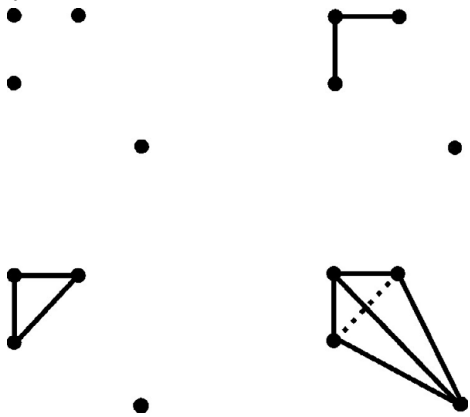


Рис. 1. ВР комплекси: зверху ліворуч $\varepsilon = 0,5$, праворуч $\varepsilon = 1$, знизу ліворуч $\varepsilon = 2$ та праворуч $\varepsilon = 3$.

За працями [9; 10] доступне програмне забезпечення для побудови комплексу Вісторіса-Ріпа.

Визначення 4. P -ланцюгом називається формальна сума $\sum a_i \Delta_i^p$, де Δ_i^p — p -симплекс і a_i набуває значення 0 чи 1.

Визначення 5. Межею p -ланцюга називається сума меж симплексів, з яких він складається, тобто $\text{bd}(\sum a_i \Delta_i^p) = \sum a_i \text{bd}(\Delta_i^p)$.

У термінах межового оператора ∂ межа p -ланцюга c має запис $\partial c = \sum_i a_i \partial \Delta_i^p$.

На множині p -ланцюгів легко означити операцію \oplus — додавання за модулем 2, причому елемент $0 = \sum 0 \Delta_i^p$ є нейтральним відносно даної операції. Для кожного p -ланцюга c існує протилежний $-c$ такий, що $c + (-c) = 0$. Тому, p -ланцюги утворюють групу, яку позначають C_p .

Кожен межовий оператор здійснює відображення (гомоморфізм) між групами C_p та C_{p-1} . Надалі, оператор ∂ будемо позначати з індексом p , тобто $\partial_p : C_p \rightarrow C_{p-1}$. Об'єднаємо групи ланцюгів та межові оператори у наступний ланцюговий комплекс:

$$\begin{aligned} 0 \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} \dots \\ \dots \xrightarrow{\partial_{p+1}} C_p \xrightarrow{\partial_p} C_{p-1} \xrightarrow{\partial_{p-1}} \dots \\ \dots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0. \end{aligned}$$

Нагадаємо, що образом відображення $f : X \rightarrow Y$ називають підмножину $B \subseteq Y$ таку, що $B = \{f(x) : x \in X\}$, та позначають $\text{Im } f$. Ядром відображення $f : X \rightarrow Y$ називають підмножину $A \subseteq X$ таку, що $A = \{x : f(x) = 0\}$, та позначають $\text{Ker } f$.

Фактор-простір — простір класів еквівалентності деякого простору X за відношенням еквівалентності \sim , який позначається X/\sim .

Кожному ланцюговому комплексу поставимо у відповідність групу гомологій H_p , яку означимо за таким правилом:

$$H_p = \text{Ker } \partial_p / \text{Im } \partial_{p+1}.$$

Якщо $z \in \text{Ker } \partial_p$, то z назвемо p -циклом. Якщо $z \in \text{Im } \partial_{p+1}$, то z назвемо p -межею. На далі, $\text{Ker } \partial_p$ будемо позначати через Z_p , а $\text{Im } \partial_{p+1}$ через β_p .

Числом Бетті β_p називається ранг групи гомологій H_p , тобто $\beta_p = \text{rank } H_p$. Числа Бетті мають таку геометричну інтерпретацію: β_0 —

кількість зв'язних компонент простору, β_1 — кількість «розривів» простору, які можна зробити при цьому не змінивши число компонент зв'язності і β_3 — кількість «пустот». Для прикладу, у випадку тора $\beta_0 = 1$, $\beta_1 = 2$ та $\beta_3 = 1$.

Перейдемо до визначень, що стосуються сталих гомологій [1; 2; 7; 8].

Визначення 6. Послідовність вкладених симплекціальних підкомплексів $\emptyset = K_0 \subset K_1 \subset \dots$

$\dots \subset K_n = K$ деякого комплексу K називається його фільтрацією.

Оскільки при $i < j$ комплекс K_j утворений додаванням до K_i деякої кількості «граней», то така фільтрація породжує частковий порядок ($K_i < K_j$, якщо $i < j$) на множині $\{K_i\}_{i=0}^n$.

Функція $f(x) = x$ задає відображення $f_* : H_p(K_{i-1}) \rightarrow H_p(K_p)$, яке для кожного індексу p породжує таку послідовність груп гомологій: $0 = H_p(K_0) \rightarrow H_p(K_1) \rightarrow \dots$

$$\dots \rightarrow H_p(K_n) = H_p(K).$$

Введемо позначення $H_p^i = H_p(K_i)$ та означимо відображення $f_p^{i,j} : H_p^i \rightarrow H_p^j$ як таке, що його образ складається з усіх p -вимірних класів гомологій, які з'являються «в» чи «після» K_i та зникають після K_j .

Визначення 7. p -вимірними сталими групами гомологій $H_p^{i,j}$ називаються образи відображень

$$f_p^{i,j}, H_p^{i,j} = \text{Im } f_p^{i,j}, \text{ для } 0 \leq i < j \leq n+1.$$

Визначення 8. p -вимірними сталими числами Бетті $v_p^{i,j}$ називаються ранги груп $H_p^{i,j}$, $v_p^{i,j} = \text{rank } H_p^{i,j}$.

Сталі групи гомологій складаються з усіх тих гомологій K_i , які залишаються «живими» у K_j , або, більш формально, $H_p^{i,j} = Z_p^i / (B_p^j \cap Z_p^i)$, де Z_p^i — p -цикли в K_i і B_p^j — p -межі у K_j .

За допомогою сталих чисел Бетті $v_p^{i,j}$ означимо числа $\mu_p^{i,j} = (\beta_p^{i,j-1} - \beta_p^{i,j}) - (\beta_p^{i-1,j-1} - \beta_p^{i-1,j})$, що мають таку інтерпретацію: перша різниця вказує кількість класів у K_{j-1} , що з'явилися в або до K_i та зникли в K_j , а друга — кількість класів в K_{j-1} , що з'явилися в або до K_{i-1} та зникли в K_j .

Звідки, $\mu_p^{i,j}$ — кількість класів гомологій, що з'явилися в K_i та зникли в K_j .

Визначення 8. Діаграмою сталості $\text{Dgm}_p(f)$ фільтрації, що породжена функцією f , називається сукупність точок з координатами (i, j) та кратністю $\mu_p^{i,j}$. Додамо до діаграми $\text{Dgm}_p(f)$ діагональ $i - j = 0$. Оскільки $i < j$, то всі точки (i, j) лежать вище діагоналі. Число $\beta_p^{i,j}$ — це кількість точок, що містяться верхньому лівому квадранті, вершина якого в точці з координатами (i, j) . Клас, що «народжується» в K_i і «помирає» в K_j врахований у $\beta_p^{i,j}$ тоді і лише тоді, коли $i \leq k$ і $j > l$. Тому такий квадрат є замкненим справа та відкритим зверху.

Зрозуміло, що $\text{Dgm}_p(f)$ залежить від функції f . Як зміниться $\text{Dgm}_p(f)$, якщо скінченний набір точок той самий, а функцію f замінити деякою функцією g ?

Визначення 9. Нехай $B = \{b\}$ — множина всіх біекцій $b : \text{Dgm}_p(f) \rightarrow \text{Dgm}_p(g)$. Розглянемо відстань

$$\|x - y\|_\infty = \max \{|x_1 - x_2|, |y_1 - y_2|\}, \text{ де } x = (x_1, x_2) \text{ і } y = (y_1, y_2). \text{ Тоді відстань } W_\infty(\text{Dgm}_p(f), \text{Dgm}_p(g)) \text{ між діаграмами } \text{Dgm}_p(f) \text{ та } \text{Dgm}_p(g) \text{ означимо таким чином: } W_\infty(\text{Dgm}_p(f), \text{Dgm}_p(g)) = \inf_{b \in B} \sup_{x \in \text{Dgm}_p(f)} \|x - b(x)\|_\infty.$$

Теорема про стійкість діаграм сталості. Нехай $f, g : K \rightarrow R$ дві монотонні функції, визначені на симплекціальному комплексі K . Тоді, для $\forall p \geq 0$, $W_\infty(\text{Dgm}_p(f), \text{Dgm}_p(g)) \leq \|f - g\|_\infty$.

Матричне представлення сталих гомологій та блок-схеми алгоритмів

Нехай задано симплекціальний комплекс K . Для кожної розмірності p побудуємо межову

$$\text{матрицю } \partial_p = \begin{pmatrix} a_{11} & \dots & a_{1n_p} \\ \dots & \dots & \dots \\ a_{n_{p-1}1} & \dots & a_{n_{p-1}n_p} \end{pmatrix}, \text{ рядками якої}$$

є $(p-1)$ -сиплекси, стовпцями p -сиплекси і $a_{ij} = 1$ тоді і лише тоді, коли i -й $(p-1)$ -сиплекс є межею j -го p -сиплексу.

Операції над рядками та стовпцями межової матриці:

- перестановка місцями i -го та j -го стовпців (рядків);
- додавання за модулем 2 відповідних елементів i -го до елементів j -го стовпців (рядків).

За допомогою вказаних вище операцій межову матрицю ∂_p можна звести до смітової нормальної форми: діагональ матриці складається з одиниць, а все решта нулі.

Утворимо матрицю $D = [d_{ij}]$, записавши межові матриці всіх розмірностей ∂_p в одну. Позначимо через σ_i симплекс, що належить $K_i - K_{i-1}$, тоді $d_{ij} = \begin{cases} 1, & \sigma_i \in \partial \sigma_j; \\ 0, & \sigma_i \notin \partial \sigma_j. \end{cases}$ У матриці D

рядки і стовпці впорядковані по мірі їх появи під час фільтрації, а межа симплексу записана в його стовпці.

Приклад. Побудуємо межову матрицю для тетраедра, зображеного на рис. 1. Позначимо вершини (0-симплекси) літерами a, b, c, d , ребра (1-симплекси) e, f, g, h, i, j та грані (2-симплекси) k, l, o, m так, що $e = [a, b]$ (ребро e сполучає вершини a та b), $f = [b, c]$, $g = [a, c]$, $h = [c, d]$, $i = [a, d]$, $j = [b, d]$, $k = [e, f, g]$ (межа грані k складається із ребер e, f та g), $l = [e, i, j]$, $o = [g, h, i]$, $m = [f, h, j]$. Межова матриця D для розмірності 2 має вигляд

$$D_2 = \begin{pmatrix} & k & l & o & m \\ e & 1 & 1 & 0 & 0 \\ f & 1 & 0 & 0 & 1 \\ g & 1 & 0 & 1 & 0 \\ h & 0 & 0 & 1 & 1 \\ i & 0 & 1 & 1 & 0 \\ j & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Позначимо через $l_M[j]$ номер рядка, що містить останню 1 в j -му стовпці матриці M . У прикладі, наведеному вище, $l_{D_2}[1] = 3$.

Визначення 10. Скажімо, що матриця M зведена, якщо $l_M[j] \neq l_M[j']$ для довільних $j \neq j'$.

Твердження 1. Нехай $R = DV$, де R — зведена матриця, а V — трикутна. Симплекс σ_i належить деякому симплексу σ , який зникає в симплексі σ_j , де $i < j$, тоді і лише тоді, коли $l_R[j] = i$.

Оскільки $R = DV$, то j -й стовпець матриці R відповідає $(p-1)$ -ланцюгу ∂c_j .

Твердження 2. Нехай c_j та c_j' — p - та $(p-1)$ -ланцюги, що відповідають j -му стовпцеві матриці V та j -му стовпцеві матриці R відповідно. Тоді $c_j' = \partial_p c_j$.

Схема обробки даних за допомогою методу сталих гомологій:

1. Зафіксуємо скінченний набір чисел $\{\epsilon_0, \epsilon_1, \dots, \epsilon_n\}$, $\epsilon_i \in R_+$, $\epsilon_i < \epsilon_j$ при $i < j$ та побудуємо відповідну йому фільтрацію $\{K_i\}_{i=0}^n$, де K_i — ВР комплекс, що відповідає ϵ_i . Для кожного i , $i = \overline{0, n}$ отримаємо набір 0-симплексів $\{c_{0,j}^i\}_{j=1}^{m_0^i}$, 1-симплексів $\{c_{1,j}^i\}_{j=1}^{m_1^i}$, 2-симплексів $\{c_{2,j}^i\}_{j=1}^{m_2^i}$ і т.д. Загалом, для кожних p та i отримаємо набір p -симплексів $\{c_{p,j}^i\}_{j=1}^{m_p^i}$, де $i = \overline{0, n}$, $m_k^i \in N$. Згідно з визначеннями 4 і 5, для довільних i та $p \exists s, t$ такі, що $c_{p-1,s}^i \in \partial c_{p,t}^i$. Надалі, той факт, що $c_{p-1,s}^i \in \partial c_{p,t}^i$ будемо позначати через $\text{pr}(c_{p-1,s}^i; c_{p,t}^i)$.

Упорядкуємо всі p -симплекси по мірі їх виникнення у фільтрації, утворимо послідовність

$$\{c_{p,1}^0, \dots, c_{p,m_p^0}^0, c_{p,1}^1, \dots, c_{p,m_p^1}^1, \dots, c_{p,m_p^n}^n\} \quad (1)$$

та перенумеруємо їх, записавши послідовність (1) у вигляді

$$\{c'_{p,1}, c'_{p,2}, \dots, c'_{p,N_p}\} = \{c'_{p,j}\}_{j=1}^{N_p}, \quad (2)$$

де $N_p = \sum_{i=0}^n m_p^i$ — кількість p -симплексів у комплексі K . Якщо у $(p-1)$ - та p -симплексах елементам $c_{p-1,s}^i$ та $c_{p,t}^i$ послідовності типу (1) відповідають елементи $c'_{p-1,u}$ та $c'_{p,v}$ відповідно, послідовності типу (2) і $\text{pr}(c_{p-1,s}^i; c_{p,t}^i)$, то $\text{pr}(c'_{p-1,u}; c'_{p,v})$. Іншими словами, відношення pr зберігається при зміні нумерації симплексів.

2. Для комплексу $K = \bigcup_{i=0}^n K_i$ побудуємо межову матрицю D .

На рис. 2 зображена блок-схема алгоритму такої побудови, де $1 = \langle \text{початок} \rangle$, $2 = \langle \text{pr}(c'_{i,u}; c'_{i+1,v}); n; N_0; \dots; N_n \rangle$, $3 = \langle N_{-1} = 0 \rangle$, $4 = \langle l = 1, i = 1 \rangle$, $5 = \langle l = l + 1 \rangle$, $6 = \langle k = 1 \rangle$,

$7 = \left\langle s = \sum_{r=1}^{i-2} N_r + k, t = \sum_{r=0}^{i-1} N_r + l \right\rangle,$
 $8 = \left\langle k < \sum_{r=0}^n N_r + 1 \right\rangle, \quad 9 = \left\langle l < \sum_{r=0}^i N_r + 1 \right\rangle,$
 $10 = \langle i = i + 1 \rangle, \quad 11 = \langle \text{pr}(c'_{i-1,k}; c'_{i,l}) \rangle, \quad 12 = \langle d_{st} = 0 \rangle,$
 $13 = \langle i < n + 1 \rangle, \quad 14 = \langle d_{st} = 1 \rangle, \quad 15 = \langle d_{st} = 0, t = \overline{1, N_0},$
 $s = 1, \sum_{r=0}^n N_r \rangle, \quad 16 = \langle \text{друк } d_{st} \rangle, \quad 17 = \langle k = k + 1 \rangle,$
 $18 = \langle \text{кінець} \rangle.$

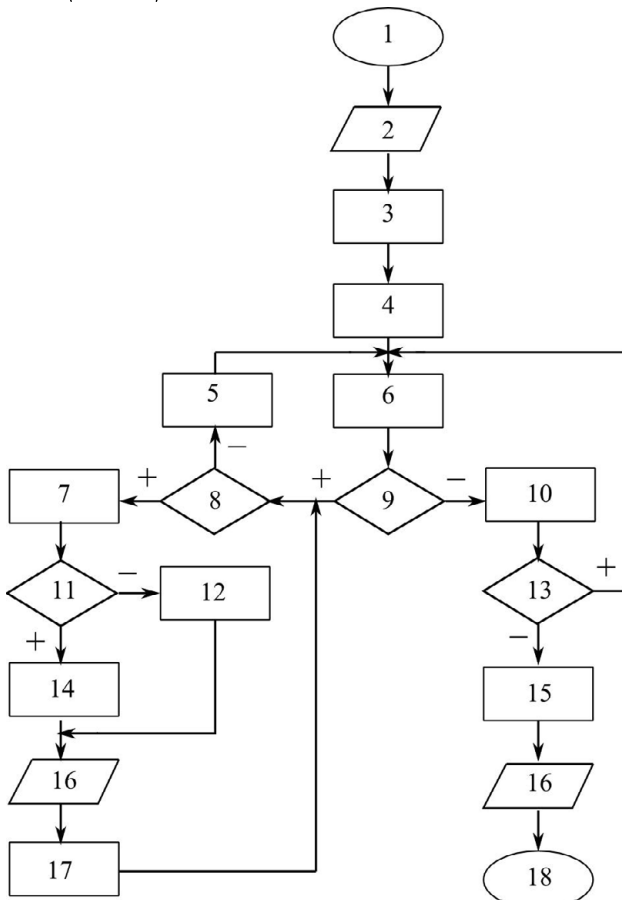


Рис. 2. Блок-схема алгоритму побудови матриці D

Перевіримо коректність алгоритму:

Ініціалізація. При $l=1$ та $k=1$ елемент $d_{1(N_0+1)} = \text{pr}(c'_{0,1}; c'_{1,1})$.

Збереження. В тілі зовнішнього циклу відбувається побудова елементів матриці доти, доки не заповняться всі рядки стовпця з номером l , який залежить від i , що є показником розмірності.

Завершення. При $i=n$ зовнішній цикл завершується порядковою побудовою елементів у l -му стовпці, де $l = \sum_{r=0}^n N_r$. Якщо $i=n+1$, всі елементи матриці виводяться на екран.

Час роботи даного алгоритму $\Theta(n^4)$, де $n = \left(\sum_{i=0}^n N_i \right)^{\sum_{i=0}^n N_i}$.

3. Нехай $D = [d_{ij}]$, $i, j = \overline{1, N}$, — межова матриця отримана в п. 2. Для кожного j запишемо число $l_D[j]$, де $j = \overline{1, N}$. За допомогою операції \oplus отримаємо зведену матрицю $R = [r_{ij}]$, $i, j = \overline{1, N}$.

На рис. 3 зображена блок-схема даного алгоритму, де $1 = \langle \text{початок} \rangle$, $2 = \langle \text{ввести } D = [d_{ij}]; l_D[j]; N \rangle$, $3 = \langle j = 1 \rangle$, $4 = \langle j_0 = 1 \rangle$, $5 = \langle i = 1 \rangle$, $6 = \langle l_D[j] = l_D[j_0] \rangle$, $7 = \langle j_0 < j \rangle$, $8 = \langle d_{ij} = d_{ij} \oplus d_{i j_0} \rangle$, $9 = \langle j = N + 1 \rangle$, $10 = \langle j = j + 1 \rangle$, $11 = \langle i < N + 1 \rangle$, $12 = \langle j_0 = j_0 + 1 \rangle$, $13 = \langle r_{ij} = d_{ij} \rangle$, $14 = \langle i = i + 1 \rangle$, $15 = \langle \text{друк } r_{ij} \rangle$, $16 = \langle \text{кінець} \rangle$.

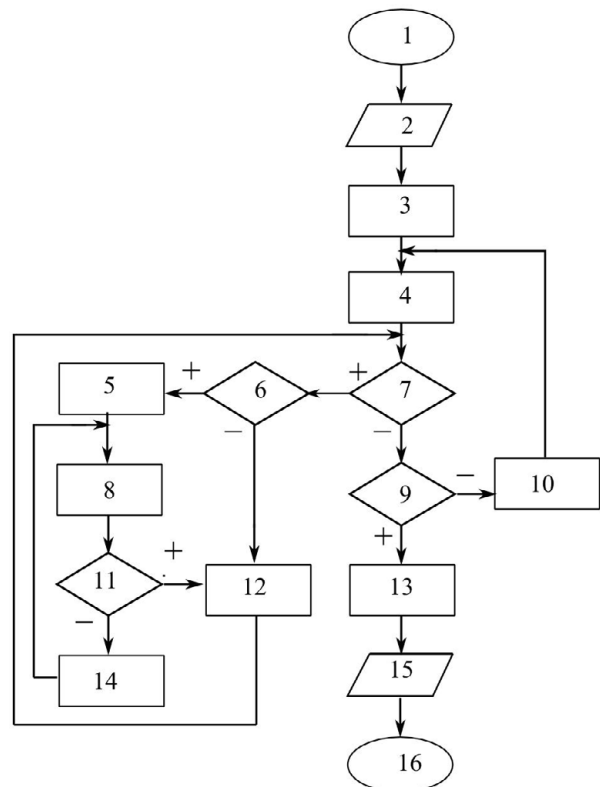


Рис. 3. Блок-схема алгоритму побудови зведеної матриці R

Ініціалізація. При $j_0=1$ та $j=1$ елементи першого стовпця залишаються тими самими, а елементи другого стовпця змінюються залежно від виконання умови $l_D[j] = l_D[j_0]$.

Збереження. В тілі зовнішнього циклу відбувається побудова елементів матриці до тих пір, поки не заповняться всі рядки стовпця з номером j .

Завершення. При $j = N$ зовнішній цикл завершується порядковою побудовою елементів N -го стовпця. Якщо $j = N + 1$, всі елементи матриці виводяться на екран.

Час роботи даного алгоритму $\Theta(n^4)$, де $n = N^2$.

4. Обчислюємо значення чисел $\mu_p^{i,j}$: якщо для деякого номеру j матриці R існує s таке, що $l_R[j] = s$, то $\mu_x^{s,j} = 1$, де

$$x = \begin{cases} 0, & i < N_0, \\ p, & i = \sum_{k=0}^p N_k, \\ p+1, & \sum_{k=0}^p N_k < i < \sum_{k=0}^{p+1} N_k. \end{cases}$$

5. Для кожного p будуємо діаграму сталості $\text{Dgm}_p(f)$.

Висновок

У даній роботі запропоновано схему обробки даних методом сталих гомологій, яка використовує побудову фільтрації комплексами Вісторіа-Ріпа. Вперше розроблено блок-схеми алгоритмів побудови межової та зведеної матриць, що дозволяють створити програмний продукт на будь якій мові програмування.

У перспективі розвитку даної тематики існує необхідність оптимізувати дані алгоритми з метою зменшення часу їх роботи.

Як приклад, для реалізації алгоритму побудови зведеної матриці фільтрації, яка складається зі 100 елементів, необхідно витрати $\Theta(10^{16})$ часу.

ЛІТЕРАТУРА

1. Zomorodian A. Computing persistent homology / A. Zomorodian, G. Carlsson // *Disc. Comput. Geom.* — 2005. — V. 33. — No 2. — P. 249–274.
2. Zomorodian A. *Advances in Applied and Computational Topology* / Afra Zomorodian. — NY: AMS, 2012. — P. 232. — (AMS Short Course on Computational Topology; V. 70).
3. Ghrist R. Barcodes: The persistent topology of data // *Bull. Amer. Math. Soc.* — 2008. — V. 45. — № 1. — P. 61–75.
4. Silva V. Homological sensor networks / V. Silva, R. Ghrist // *Not. Amer. Math. Soc.* — 2007. — V. 54. — № 1. — P. 10–17.
5. Dey T. Optimal homologous cycles, total unimodularity and linear programming / T. Dey, A. N. Hirani, B. Krishnamoorthy // *SIAM Jour. Comp.* — 2011. — V. 40. — № 4. — P. 1026–1040.
6. Cohen-Steiner D. Vines and vineyards by updating persistence in linear time / D. Cohen-Steiner, H. Edelsbrunner, D. Morozov // *Proc. 22nd Annu. Sympos. Comput. Geom.* — 2006. — P. 119–134.
7. Edelsbrunner H. Topological persistence and simplification / H. Edelsbrunner, D. Letscher, A. Zomorodian // *Disc. Comput. Geom.* — 2002. — V. 28. — № 4. — P. 511–533.
8. Carlsson G. Topology and data // *Bull. Amer. Math. Soc.* — 2009. — V. 46. — № 2. — P. 255–308.
9. Morozov D. Dionysus. [Electronic Resource]. — Mode of access: URL: <http://www.mrzv.org/software/dionysus/>
10. Sexton H. JPlex / H. Sexton, M.V. Johansson // [Electronic Resource]. — Mode of access: URL: <http://www.comptop.stanford.edu/programs/jplex/>.

Стаття надійшла до редакції 28.05.2014.