COMP229: Introduction to Data Science
Lecture 10: counting cycles in graphs

Vitaliy Kurlin, vitaliy.kurlin@liverpool.ac.uk
Autumn 2018, Computer Science department
University of Liverpool, United Kingdom

Cycles of friends on Facebook

Certain subgraphs such as triangular cycles of friends on Facebook help targeted advertising.

Even if you hide personal preferences, Facebook learns a lot about you from your friends.

Cycles and subgraphs are important features of any network. What's the maximum number of subgraphs in a graph on n vertices?

The number of proper subgraphs

Claim 10.1. The complete graph K_n on n vertices has $2^n - 2$ *proper* subgraphs (not empty, not K_n).

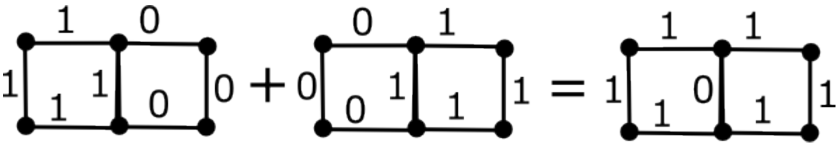
Proof. We count all possible subsets of n vertices. For each of n vertices, we have exactly two choices (include or not include into a subgraph).

Since these choices are independent, we multiply them and get the product 2^n of n factors 2.

It remains to remove from the set of 2^n subsets the two trivial subgraphs (empty and K_n itself). $\square$

How many cycles in total?

The length (number of edges) of a cycle isn't a topological invariant. From a topological point of view we can count only cycles of all lengths.



How many can you count in the above graph? If we allow cycles with repeated edges, there are many of them, but they can be well-organised.

Even and odd numbers

Definition 10.2. An integer number m is *even* if m is *divisible* by 2, i.e. there is another integer k such that $m = 2k$, e.g. 0 is even since $0 = 2 \times 0$.

An integer m is *odd* if m has the form (or can be written as) $m = 2k + 1$ for an integer k. For example, -1 is odd, because $-1 = 2 \times (-1) + 1$.

Always write brackets around negative numbers, because 2×-1 looks like $2x - 1$ (two x minus 1).

All integers split two classes: even and odd.

Addition of even and odd numbers

Claim 10.3. Here are the rules modulo 2:

even + even = even, e.g. $2 + 4 = 6$;

even + odd = odd, e.g. $2 + 3 = 5$;

odd + odd = even, e.g. $1 + (-1) = 0$.

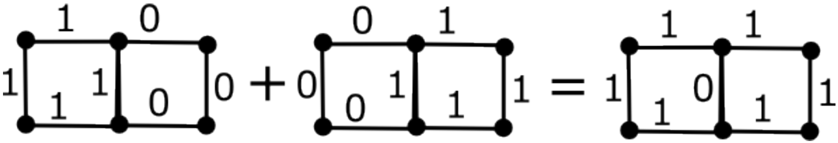
Proof. The sum of any even numbers $2m$ and $2k$ is also even, because $2m + 2k = 2(m + k)$.

even+odd : $2m + (2k + 1) = 2(m + k) + 1$.

odd+odd : $(2m + 1) + (2k + 1) = 2(m + k + 1)$. $\square$

Algebraic sum of cycles

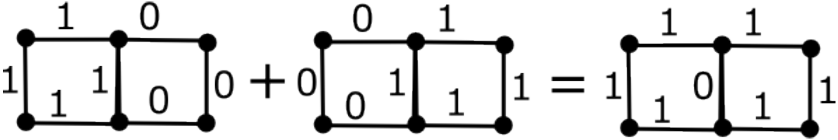
Definition 10.4. Let $e_1, \dots, e_k$ be all edges of a graph G. Then any cycle in G can be represented as a *formal sum* of edges (with coefficients 1 for edges in the cycle, coefficients 0 for all others).



The sum of two cycles is obtained by adding all edges from both cycles (with all coefficients considered modulo 2), i.e. any double edge is 0.

The first Betti number

Definition 10.5. The *first Betti* number of a graph G is the minimum number of basis cycles such that any other cycle of G can be algebraically written as a linear combination of basis cycles.



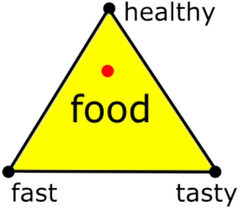
The long cycle of 6 edges on the right is the sum of two 4-edge cycles on the left. This graph (any figure-eight graph) has first Betti number $b_1 = 2$.

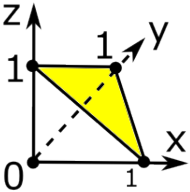
Claim 10.6. $b_1(G) = \#\text{components} - \chi(G)$.

A 2D meaning of a triple

Any graph represents only pairwise connections between features. A relationship between three features is represented by a triangle (simplex).

A point inside the triangle can be represented by 3 barycentric (fractional) coordinates (ratios) $x, y, z \in [0, 1]$ whose sum is $x + y + z = 1$.

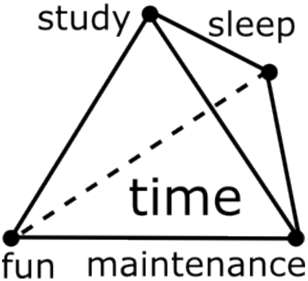




All meals can form a point cloud (per day or person) in the triangle.

Multiple features of data

2D cycles are formed by triangles, e.g. the boundary of a tetrahedron. High-dimensional simplices (and their cycles) represent relationships between various features.



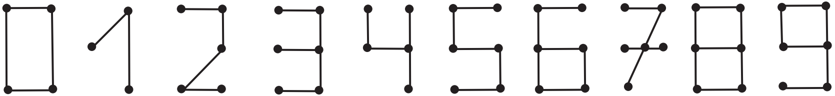
A student can be described by point in the time tetrahedron (how you divide 24 hours per day between your different activities).

Your questions and the quiz

To benefit from the lecture, now you could

- ask or submit your anonymous questions to the COMP229 folder after the lecture;
- write down your summary in 2–3 phrases, e.g. list key concepts you have learned;
- talk to your classmates to revise the lecture.

Q. Find the first Betti numbers of the graphs.



Answer to the quiz and summary

Answer. The graphs $1, 2, 3, 4, 5, 7$ have $b_1 = 0$, the graphs $0, 6, 9$ have $b_1 = 1$, finally $b_1(8) = 2$.

- K_n has $2^n - 2$ proper subgraphs (no $\emptyset$, K_n).
- Cycles of a graph can be added modulo 2.
- $b_1(G) = \#\text{components} - \chi(G)$ equals the number of basis cycles under addition mod 2.
- Relationships between several features are represented by high-dimensional simplices.