Recall that if S is a set and $k \geq 0$, then $\binom{S}{k}$ is the set of all $A \subseteq S$ with $|A| = k$. In particular, $\binom{S}{2}$ is the set of all subsets $\{i, j\} \subseteq S$ with $i \neq j$.

Example. If $S = \{1, 2, 3\}$, then

$$\binom{S}{2} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

Definition. A **graph** G is a pair (V, E), where

- V is a finite set.
- $E \subseteq \binom{V}{2}$.

The elements of V are the **vertices** of G, and the elements of E are the **edges** of G. We typically illustrate a graph by drawing a dot for each element of V, and for each $\{u, v\} \in E$, a line between the dots representing u and v.

ILLUSTRATION

Example. If S is a finite set with some relation $\sim$ between its elements, we get a graph $G = (S, E)$, letting

$$E = \{\{u, v\} \mid u \sim v\}.$$

For instance, taking $S = \{1, \ldots, 8\}$ and letting $i \sim j$ be the relation $i \mid j$ or $j \mid i$, gives the graph

ILLUSTRATION

Example. Applications of graphs outside of pure maths are plentiful:

- Social networks, with vertices representing people and edges representing friendship or interaction, are used e.g. for studying the spread of information or disease.
- Graphs model networks of roads or public transit. Here one can take intersections or stations as nodes, and roads or transit connections as edges. Algorithms (like "finding the shortest path") help optimize routing and reduce travel time.
- In neuroscience, graphs can represent the connectivity in the brain, with vertices corresponding to neurons and edges corresponding to electrical connections between them, or vertices corresponding to brain regions and edges corresponding to statistical correlation between them.
- The internet is a graph (with servers and routers as vertices and physical or wireless links as edges).

Remark. It will be convenient to write uv rather that $\{u, v\}$ for edges.

Remark. The definition of a graph varies in the literature. In particular, other authors may (but we do not) allow

- the sets V and E to be infinite
- an edge from a vertex to itself (a "loop")
- multiple edges between two edges.

Example. The **complete graph** on n vertices is denoted K_n. We take $|V| = n$ and $E = \binom{V}{2}$, that is every pair of vertices has an edge between them.

Definition. We say a graph $G = (V, E)$ is **bipartite** if we can write V as a disjoint union $V = S \cup T$, in such a way that every edge connects an element of S to an element of T.

The **complete bipartite graph** $K_{m,n}$ is obtained by taking $|S| = m$ and $|T| = n$, and letting $E = \{\{u, v\} \mid u \in S, v \in T\}$. In other words, take an edge from every element of S to every element of T.

Example. The **hypercube** graph Q_n has vertices labelled by strings of length n on letters 0 and 1, and an edge between strings u and v if they differ in at most one letter.

Example. The **path graph** P_n has $V = \{u_1, \ldots, u_n\}$, and $E = \{u_1u_2, u_2u_3, \ldots, u_{n-1}u_n\}$.
The **circuit graph** C_n has $V = \{u_1, \ldots, u_n\}$, and $E = \{u_1u_2, u_2u_3, \ldots, u_{n-1}u_n, u_nu_1\}$.

Example. The **Petersen graph** is this one.

In analysing the structure of graphs, it is worthwhile understanding when two graphs are "the same". This is captured by the following definition:

Definition. Two graphs $G = (V, E)$ and $G' = (V', E')$ are called **isomorphic**, and we write $G \cong G'$, if there is a bijection $\phi\colon V \to V'$ such that $uv \in E$ if and only if $\phi(u)\phi(v) \in E'$.

Example. We have $K_{2,2} \cong C_4 \cong Q_2$.

If we want to show that two graphs G and G' are isomorphic, we need to provide an suitable bijection between their vertices. If we want to show that they are *not* isomorphic, we typically specify some structural aspect of G and G' which makes them differ.

Example. TWO GRAPHS OF ORDER 8 AND SIZE 14

Graph theory terms.

Definition (Relations between vertices and edges)**.** We say

- vertices u, v are **adjacent** or **neighbours** if $uv \in E$.
- an edge k and a vertex u are **incident** if $k = uv$ for some v.
- two edges k, j are **incident** if they are both incident to the same vertex.

Definition (Neigbour set and degree of vertex)**.** Let $u \in V$. The set of vertices adjacent to u is denoted $N(u)$, and we write $d(u) = |N(u)|$, the **degree** of u.
A vertex with no neighbours is called **isolated**.

Definition. If $G = (V, E)$, with $|V| = n$, we can order the vertices $u_1, \ldots, u_n$ such that

$$d(u_1) \geq d(u_2) \geq \cdots \geq d(u_n).$$

The resulting sequence of integers is the **degree sequence** of G.

Definition (Size terms)**.** For a graph $G = (V, E)$, we usually write $n = |V|$ and $q = |E|$, and call n the **order** of G and q the **size** of G.

Example. The graphs K_n, C_n, P_n have order n.
The graph Q_n has order 2^n.

Example. If G and G' are isomorphic, then clearly their orders and sizes must agree, as must their degree sequences.
In the example above, these three invariants agree for both graph, we have $n = 8$, $q = 14$ and degree sequence

$$4, 4, 4, 4, 3, 3, 3, 3.$$

But we can notice that in one of them, there are two degree 3 vertices which are neigbours, while in the other, none of the degree 3 vertices are neighbours. This means that the two graphs are non-isomorphic.

The following gives a simple constraint on the degree sequence.

Theorem. *For any graph $G = (V, E)$, we have $\sum_{u \in V} d(u) = 2|E|$.*

Proof. For every u, $d(u)$ is the number of pairs (u, k) with k incident to u. On the other hand, for every $k \in E$, there are two vertices u and v such that (u, k) and (v, k) form an incident pair. So

$$\sum_{u \in V} d(u) = |\{(u, k) \mid u \in V, k \in E, u \text{ incident to } k\}| = \sum_{k \in E} 2 = 2|E|.$$

□

Representing graphs in matrices. Both for analysing graphs mathematically, and for encoding them in a computer, it is useful to package the data of a graph in a matrix.

We let $G = (V, E)$ be a graph, and assume that we have labelled the vertices $V = \{u_1, \ldots, u_n\}$, and the edges $E = \{k_1, \ldots, k_q\}$.

Definition. The **adjacency matrix** of G, denoted A, is the $(n \times n)$-matrix given by

$$A = (a_{ij})_{i,j=1,\ldots,n}$$

where

$$a_{ij} = \begin{cases} 1 \text{ if } u_i u_j \in E \\ 0 \text{ otherwise.} \end{cases}$$

Remark. The matrix A is always symmetric, and always has 0's along the diagonal.

Example. For K_3, the adjacency matrix is

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

For $K_{2,2}$, the adjacency matrix is

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

Example. If A is the adjacency matrix of G, then

$$A \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} d(u_1) \\ \vdots \\ d(u_n) \end{pmatrix}$$

Given vertices u, v of G, let's say a **walk of length** l from u to v is a sequence of edges

$$k_1, k_2, \ldots, k_l$$

such that k_1 is incident to u, k_l is incident to v, and for $i = 1, \ldots, l-1$, k_i is incident to k_{i+1}. Let $W_l(u, v)$ be the number of walks from u to v of length l, that is the number of

Theorem. *For $l \geq 1$, the matrix A^l has as (i,j)-th entry $W_l(u_i, u_j)$.*

Proof. For $l = 1$, we clearly have $W^1(u,v) = 1$ if u and v are adjacent, and $W^1(u,v) = 0$ otherwise, so $A_{ij} = W^1(u_i, u_j)$. Moreover, the number of walks of lengths l from u to v is the same as the number of walks of length $l-1$ from u to a neighbour of v. Thus

$$W_l(u_i, u_j) = \sum_{u_m \in N(u_j)} W_{l-1}(u_i, u_m) = \sum_{u_m \in V} W_{l-1}(u_i, u_m) W_1(u_m, u_j) = \sum_{m=1}^{n} (A^{l-1})_{im} A_{mj} = (A^l)_{ij}.$$

□

An alternative encoding of a graph in a matrix is the following.

Definition. The **incidence matrix** of a graph $G = (V, E)$ is the $(n \times q)$-matrix

$$B = (b_{ij})$$

with

$$b_{ij} = \begin{cases} 1 \text{ if } u_i \text{ is incident to } k_j \\ 0 \text{ otherwise.} \end{cases}$$

Remark. Since every edge is incident to 2 vertices, every column of B has two 1s and $n-2$ 0s.

Theorem. *If a graph $G = (E, V)$ has adjacency matrix A and incidence matrix B, then*

$$BB^T = A + D$$

where D is the diagonal matrix with entries $d(u_1), d(u_2), \ldots, d(u_n)$.

Proof. If $1 \leq i, j \leq n$, then

$$(BB^T)_{ij} = \sum_{k=1}^{q} B_{ik} B_{kj}^T = \sum_{k=1}^{q} B_{ik} B_{jk}$$

If $i \neq j$, then

$$B_{ik} B_{jk} = \begin{cases} 1 \text{ if edge } k \text{ is incident to vertices } i \text{ and } j \\ 0 \text{ otherwise.} \end{cases}$$

Thus $(B^T B)_{ij} = \sum_{k=1}^{q} B_{ik} B_{jk}$ is 1 if i and j are neighbours, and 0 otherwise. This means that $(B^T B)_{ij} = A_{ij}$.

If $i = j$, then

$$B_{ik} B_{jk} = B_{ik}^2 = \begin{cases} 1 \text{ if edge } k \text{ is incident to vertex } i \\ 0 \text{ otherwise} \end{cases}$$

This means $(BB^T)_{ii}$ equals the number of edges incident to u_i, which is $d(u_i)$. This gives the matrix D in the statement of the theorem. □

Exercise: Show that $C = B^T B - 2I_q$ is the adjacency matrix of a graph, and describe this graph concretely.

Paths and circuits.

Definition. Let $G = (V, E)$ be a graph. A graph $H = (V', E')$ is a **subgraph** of G if $V' \subseteq V$ and $E' \subseteq E$. We say H is an **induced subgraph** if $E' = \binom{V'}{2} \cap E$.

Remark. In words, an induced subgraph of G is obtained by replacing the vertex set V with some subset V', and replacing the edge set with the set of all edges connecting vertices in V'.

Example.
- Every graph of order n can be thought of as a subgraph of K_n.
- The induced subgraphs of K_n are all isomorphic to K_m for some $0 \leq m \leq n$.

We get special graphs by deleting vertices or edges:

Definition. Let $G = (V, E)$ be a graph, and let $A \subseteq V$, $B \subseteq E$. We then get graphs
- $G \setminus A$, the induced subgraph of G with vertex set $V \setminus A$.
- $G \setminus B = (V, E \setminus B)$.

Definition. Given vertices u, v, a u, v**-path** is a path in G with start vertex u and end vertex v.

If there exists a u, v-path, we say that u is reachable from v. A graph G is **connected** if every vertex is reachable from every other vertex. If G is not connected, we can partition V into $V_1, \ldots, V_m$ in such a way that u, v lie in the same V_i if and only if u is reachable from v. The induced subgraphs on $V_1, \ldots, V_m$ are the **components** of G.

Example. All the graphs we have given names are connected.
The graph $G = (V, \varnothing)$ with $E = \varnothing$ has components $(\{u\}, \varnothing)_{u \in V}$.

An edge $k \in E$ is a **cut edge** (or **bridge**) if $G \setminus \{k\}$ has more connected components than G.

Example. The path graph P_n has $n - 2$ cut edges.
The Peterson graph has no cut edges.

Definition. Given vertices u, v, the **distance** $d(u, v)$ between u and v is the smallest possible length of a u, v-path. We set $d(u, u) = 0$ and $d(u, v) = \infty$ if no u, v-path exists
The **diameter** $d(G)$ of a graph $G = (V, E)$ is $\max_{u,v \in V} d(u, v)$.

Example. In K_n, if $u \neq v$, we have $d(u, v) = 1$, so $d(K_n) = 1$.

Example. In the hypercube graph Q_n, the distance between vertices $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ equals the number of i for which $a_i \neq b_i$. We get $d((0, \ldots, 0), (1, \ldots, 1)) = n = d(Q_n)$.

Example. Let G be the graph where V is the set of mathematicians, and $uv \in E$ if mathematician u has written and published a paper together with mathematician v. This graph has a vertex called Paul Erdős with an exceptionally high degree, ≥ 500. The **Erdős number** of mathematician x is then $d(x, \text{Paul Erdős})$. So if you have an Erdős number of 0, then you are Paul Erdős, if you have an Erdős number of 1, then you have written a paper with him (but you are not him), and so on.

Theorem (Theorem 6.4.)**.** *A graph G is bipartite if and only if all its circuits have even length, and in particular if it has no circuits.*

Note the mistake in the book, the criterion is all circuits have even length, not "the same length".

Proof. Assume G is bipartite. Then we can partition the vertex set into S and T, with every edge connecting a vertex of S with a vertex of T. So, travelling along an edge will move you from S to T or vice versa. Starting from a vertex in S and travelling along k edges, you will therefore end at a vertex in S if k is even, and end at a vertex in T if k is odd. Since a circuit starts and ends in the same vertex, it must contain an even number of edges.

Conversely, assume that every circuit of G has even length. We assume without loss of generality that G is connected. We first claim that for any $u \in V$, any walk W beginning and ending in u has *even* length $l(W)$.

We prove this by induction on the lengths of walks, more precisely we prove by induction on l the claim that "if W is a walk of length $\leq l$ starting and ending in the same vertex u then W has even length".

The case $l = 0$ is clear. Let now $W = uu_1 \dots u_k u$ be a walk of length $\leq l + 1$ beginning and ending in u. If W does not pass through any vertex more than once, then it is a circuit, and so by assumption it has even length. Otherwise, we can find some vertex $v \in V$ such that W is of the form

$$uu_1 \dots u_a v v_1 \dots v_m v u_b \dots u$$

We can split this into two walks

$$uu_1 \dots u_a v u_b \dots u$$

and

$$v v_1 \dots v_m v.$$

Since these have smaller length than W, they both have even length, by the induction assumption. Since the length of W is the sum of the lengths of these two walks, we get that W has even length, completing the proof of our claim.

We now fix some $u \in V$. We declare that $u \in S$, and for any other $v \in V$, we say $v \in S$ if $d(u, v)$ is even, and $v \in T$ if $d(u, v)$ is odd. It remains to prove that this gives a bipartite graph, that is, if vw is an edge, then v and w do not both lie in S or both in T. This in turn means that $d(u, v)$ and $d(u, w)$ have different parities. There is a walk W_1 of length $d(u, v)$ from u to v, and a walk W_2 of length $d(u, w)$ from w to u. Let W be the walk $W_1(vw)W_2$, we then get

$$l(W) = l(W_1) + 1 + l(W_2) = 1 + d(u, v) + d(u, w).$$

Since $l(W)$ is even, we must have $d(u, v) + d(u, w)$ is odd, and so $d(u, v)$ and $d(u, w)$ have different parities. □