In this case G is the graph K_3 (or W_1) consisting of a single triangle. This is the only regular graph satisfying the hypothesis of the theorem, and it also satisfies the conclusion (and it is a windmill). By the first half of the proof, every non-regular graph that possesses the friendship propery has a vertex adjacent to all others, so we have proved the theorem. ☐

The Friendship Theorem is a famous example of the use of matrix and specifically spectral techniques to solve a purely combinatorial problem. The proof here is essentially the original one of Erdös, Rényi and Sós. There are several proofs in the literature, most of which involve consideration of matrix spectra in some way. For many years there was interest in finding a "purely combinatorial" proof. Some do exist now in the literature, see for example "The Friendship Theorem" by Craig Huneke, in the February 2002 volume of the American Mathematical Monthly (available on JSTOR). Another interesting feature of this theorem is that it is no longer true if the condition that G is finite is dropped - there exist examples of infinite "friendship graphs" with no politician.

1.2 Some matrix background

The goal of this section is to fill in some details about matrices that were used implicitly or explicitly in Section 1.1, in particular the assertions about the trace of a square matrix A and its positive integer powers. For this we require the concept and meaning of similarity. Also in the background are the Rank-Nullity Theorem and the concept of an eigenvector, along with its interpretation for graphs.

First we revisit the process of matrix-vector multiplication. Let $A \in M_{m\times n}(\mathbb{R})$, let $v \in \mathbb{R}^n$ and let u in $(\mathbb{R}^m)^T$. Then

- Av is the column in $\mathbb{R}^m$ that is the linear combination of the columns of A with the entries of v as coefficients.
- uA is the row in $\mathbb{R}^n$ that is the linear combination of the rows of A with the entries of u as coefficients.

Definition 1.2.1. *If* A *is* square, $A \in M_n(\mathbb{R})$, *then a non-zero column* v *is an* eigenvector *of* A *if* Av *is a scalar multiple of* v *itself (the scalar that turns up here is the* eigenvalue *of* A *to which* v *corresponds.*

In this context, multiplication on the left by A determines a *function* from $\mathbb{R}^n$ to $\mathbb{R}^n$. An eigenvector of A corresponds to a one-dimensional subspace of $\mathbb{R}^n$ that is mapped into itself by A. In Section 1.1 however, we were considering eigenvectors of the adjacency matrix of a graph. The meaning of this has the following interpretation in graph theory.

Let G be a graph with adjacency matrix A. What is the meaning of an eigenvector of A in terms of the graph G? For example

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

We saw in Section 1.1 that -2 is an eigenvalue of this $A(G)$. A corresponding eigenvector is $u = (0, 1, -1, 1, -1, 1, -1)$.

In the context of graphs and adjacency matrices, we can think of a column vector u as a function on the vertex set, that assigns a number u_i (the ith entry of u) to each vertex v_i. We can think of u_i as the value of u at vertex v_i. In this context the ith entry of the product $A(G)u_i$ is the sum of those u_j for which v_j is a neighbour of v_i, i.e.

$$(A(G)u)_i = \sum_{j:v_j\sim v_i} u_j.$$

If u is an eigenvector of $A(G)$ corresponding to the eigenvalue λ, it means that for every vertex v of G, the sum of the values of u at the neighbours of v is the value at v itself multiplied by λ. In the example above the function corresponding to u is

and it is easily checked for each vertex in this picture that the sum of the values labelling the neighbouring vertices is -2 multiplied by the value at the vertex itself. Thus the vector $u = (0, 1, -1, 1, -1, 1, -1)$ is an eigenvector of this graph corresponding to the eigenvalue -2.

Definition 1.2.2. *Let $\mathbb{R}^n$ denote the vector space of column vectors of length n over $\mathbb{R}$. A linear transformation of $\mathbb{R}^n$ is a function $T : \mathbb{R}^n \to \mathbb{R}^n$ such that*

- $T(u + v) = T(u) + T(v)\ \forall\, u, v \in \mathbb{R}^n$;
- $T(ku) = kT(u)\ \forall\, u \in \mathbb{R}^n,\ k \in \mathbb{R}$.

Let $\mathcal{C} = \{c_1, \dots, c_n\}$ be a basis of $\mathbb{R}^n$. Then every element of $\mathbb{R}^n$ has a unique set of $\mathcal{C}$-coordinates, namely the coefficients of its expression as a linear combination of the elements of $\mathcal{C}$. Let T be a linear transformation of $\mathbb{R}^n$.

Definition 1.2.3. *The* matrix of T with respect to $\mathcal{C}$ *is the $n \times n$ matrix $A_{\mathcal{C}}$ whose jth column has the $\mathcal{C}$-coordinates of b_j as its entries.*

If $v \in \mathbb{R}^n$ and $v_{\mathcal{C}}$ is the column vector whose entries are the $\mathcal{C}$-coordinates of v, then the matrix-vector product $A_{\mathcal{C}} v_{\mathcal{C}}$ is the column vector whose entries are the $\mathcal{C}$-coordinates of $T(v)$. This follows from the observation that this product is nothing but the linear combination of the columns of A with the entries of $v_{\mathcal{C}}$ as coefficients.

Example 1.2.4. *Let T be the linear transformation of $\mathbb{R}^2$ with $T(e_1) = \binom{1}{1}$ and $T(e_2) = \binom{-2}{4}$ (where $e_1 = \binom{1}{0}$ and $e_2 = \binom{0}{1}$ are the elements of the standard basis). The matrix of T with respect to the standard basis is*

$$A = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix},$$

and if $v = \binom{a}{b}$ is any element of $\mathbb{R}^2$, then

$$T(v) = T(ae_1 + be_2) = aT(e_1) + bT(e_2) = a\begin{pmatrix}1\\1\end{pmatrix} + b\begin{pmatrix}-2\\4\end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}.$$

Now let $\mathbb{C}$ be the basis of $\mathbb{R}^2$ consisting of $c_1 = \binom{-2}{1}$ and $c_2 = \binom{1}{-1}$. Then

$$\begin{aligned} T(c_1) &= \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix} = 2c_1 \Longrightarrow [T(c_1)]_{\mathcal{C}} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \\ T(c_2) &= \begin{pmatrix} 1 & -2 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} = 3c_2 \Longrightarrow [T(c_2)]_{\mathcal{C}} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} \end{aligned}$$

The matrix $A_{\mathcal{C}}$ of T with respect to $\mathcal{C}$ is $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

Definition 1.2.5. *Two matrices in $M_n(\mathbb{R})$ are* similar *if they represent the same linear transformation with respect to different bases of $\mathbb{R}^n$.*

Two similar matrices can look quite different. When dealing with linear transformations, a general goal is to try to find a basis with respect to which the transformation is easy to describe. The best you can hope for is that there might be a basis consisting entirely of *eigenvectors* for the transformation, i.e. non-zero vectors v for which $T(v) = \lambda v$ for a real number λ (the corresponding eigenvalue of T). The matrix of T with respect to such a basis is diagonal; however such bases might not exist.

Definition 1.2.5 is the essential meaning of similarity, but it has a meaning in terms of matrix algebra also, which is equivalent and often useful. To figure this out, let $\mathcal{B} = \{b_1, \dots, b_n\}$ and $\mathcal{C} = \{c_1, \dots, c_n\}$ be different bases of $\mathbb{R}^n$, and let the matrix of T with respect to $\mathcal{B}$ be B.

Let P be the matrix whose jth column contains the $\mathcal{B}$-coordinates of the vector c_j. Then P is invertible, and the $\mathcal{B}$-coordinates of any vector $v \in \mathbb{R}^n$ are given by the matrix-vector product $Pv_{\mathcal{C}}$, where $v_{\mathcal{C}}$ is the column consisting of the $\mathcal{C}$-coordinates of v. Thus, for any vector v in $\mathbb{R}^n$, $v_{\mathcal{B}} = Pv_{\mathcal{C}}$ and equivalently $v_{\mathcal{C}} = P^{-1}v_{\mathcal{B}}$.

Now the $\mathcal{C}$-coordinates of $T(v)$ may be found by first multiplying the column vector $v_{\mathcal{C}}$ on the left by P (to convert to $\mathcal{B}$-coordinates), then multiplying on the left by the matrix B (this gives the $\mathcal{B}$-coordinates of $T(v)$), then multiplying on the left by P^{-1} (to get back to $\mathcal{C}$-coordinates). This means that the matrix of T with respect to $\mathcal{B}$ is $P^{-1}BP$, prompting the following arithmetic definition of similarity.

Definition 1.2.6. *Two matrices* B *and* C *in* $M_n(\mathbb{R})$ *are* similar *to each other if*

$$C = P^{-1}BP,$$

for some invertible $P \in M_n(\mathbb{R})$.

Definitions 1.2.5 and 1.2.6 are entirely equivalent to each other and it useful to be able to think of both of them together. The arithmetic version is particularly useful for proving some of the shared properties of similar matrices.

To connect to the content of Section 1.1, we first show that similar matrices have the same trace, which is a consequence of the fact that while the matrix products AB and BA are generally different, they always have the same trace.

Lemma 1.2.7. *Let* A *and* B *be matrices in* $M_n(\mathbb{R})$. *Then* $\text{trace}(AB) = \text{trace}(BA)$.

Proof. The trace of AB is the sum of the diagonal entries of AB. The entry in the (i,i) position of AB is the scalar product of Row i of A with Column i of B. Thus

$$\begin{aligned}\text{trace}(AB) &= \sum_{i=1}^{n}(AB)_{ii} \\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} A_{ik}B_{ki} \\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} B_{ki}A_{ik} \\ &= \sum_{k=1}^{n}(BA)_{kk} \\ &= \text{trace}(BA).\end{aligned}$$

□

Corollary 1.2.8. *Suppose that* A *and* B *are similar matrices in* $M_n(\mathbb{R})$. *Then* $\text{trace}(A) = \text{trace}(B)$.

Proof. Since A and B are similar, $B = P^{-1}AP$ for some invertible $P \in M_n(\mathbb{R})$. Then

$$\text{trace}(B) = \text{trace}(P^{-1}AP) = \text{trace}(APP^{-1}) = \text{trace}(A),$$

by Lemma 1.2.7. □

Our next goal is to show that the trace of a square matrix is the sum of its eigenvalues, something that we used in Section 1.1. We recall a few details about eigenvalues first.

1. The (possibly complex) number λ is an eigenvalue of the square matrix $A \in M_n(\mathbb{R})$ if and only if there exists a non-zero column vector v with $Av = \lambda v$, which means $(\lambda I - A)v = 0$.

2. This means that some non-zero linear combination of the columns of $\lambda I - A$ is the zero vector, which means exactly that these columns are linearly dependent, which occurs exactly if $\det(\lambda I - A) = 0$. So the eigenvalues of A are the roots of the polynominal $\det(\lambda I - A)$ (the characteristic polynomial of A). So similar matrices have the same spectrum.

3. The determinant is a *multiplicative* function on $M_n(\mathbb{R})$, which means that $\det(AB) = \det(A)\det(B)$ for $A, B \in M_n(\mathbb{R})$ (this is the Cauchy-Binet formula, definitely not an obvious thing). If P is an invertible matrix, then it follows from the Cauchy-Binet formula $\det(P)\det(P^{-1}) = \det I_n = 1$ and that A and $P^{-1}AP$ have the same characteristic polynomial for all $A \in M_n(\mathbb{R})$, i.e.

$$\begin{aligned}\det(\lambda I - P^{-1}AP) &= \det(\lambda P^{-1}IP - P^{-1}AP) = \det(P^{-1}(\lambda I - A)P) \\ &= \det(P^{-1})\det(\lambda I - A)\det(P) = \det(\lambda I - A).\end{aligned}$$

4. Finally suppose that A is an upper triangular matrix (i.e. $A_{ij} = 0$ whenever $i > j$, all entries of A below the main diagonal are zero). Then the determinant of A is just the product of the entries on the main diagonal, and the characteristic polynomial of A is just the product over i of $(\lambda - A_{ii})$ - so the spectrum of A consists of the entries on the main diagonal.

For the next theorem we consider matrices over $\mathbb{C}$, the reason being that the field $\mathbb{C}$ of complex numbers is *algebraically closed*, which means that every polynomial with complex coefficients has a full set of roots in $\mathbb{C}$. Note that every matrix in $M_n(\mathbb{R})$ is also in $M_n(\mathbb{C})$. All eigenvalues of a real matrix are complex, they are not necessarily all real.

Theorem 1.2.9. *Let $A \in M_n(\mathbb{C})$. Then A is similar in $M_n(\mathbb{C})$ to an upper triangular matrix in $M_n(\mathbb{C})$.*

Proof. By induction on n. The case $n = 1$ is clear, since every 1×1 matrix is upper triangular. Let T be the linear transformation of $\mathbb{C}^n$ determined by $T(v) = Av$, for $v \in \mathbb{C}^n$. Let λ_1 be an eigenvalue of A in $\mathbb{C}$ with corresponding eigenvector v_1. Expand $\{v_1\}$ to a basis $\{v_1, v_2, \dots, v_n\}$ of $\mathbb{C}^n$. Then the matrix A_1 of T with respect to this basis has λ_1 in its top left entry and zeros otherwise in its first column, write this matrix as

$$A_1 = \left(\begin{array}{c|ccc} \lambda_1 & * & \dots & * \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right),$$

where $A' \in M_{n-1}(\mathbb{C})$. By the induction hypothesis, there exists an invertible matrix $Q \in M_{n-1}(\mathbb{C})$ for which $T' = Q^{-1}AQ$ is upper triangular. Write

$$P = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right).$$

Then

$$P^{-1} = \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right),$$

and

$$\begin{aligned}
P^{-1}A_1P &= \left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q^{-1} & \\ 0 & & & \end{array}\right)\left(\begin{array}{c|ccc} \lambda_1 & * & \dots & * \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right)\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right) \\
&= \left(\begin{array}{c|ccc} \lambda_1 & \square & \dots & \square \\ \hline 0 & & & \\ \vdots & & Q^{-1}A'Q & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} \lambda_1 & \square & \dots & \square \\ \hline 0 & & & \\ \vdots & & T' & \\ 0 & & & \end{array}\right).
\end{aligned}$$

Thus A_1, and hence A, is similar to an upper triangular matrix. ☐

Finally we are in a position to prove the following statement.

Theorem 1.2.10. *Let* $A \in M_n(\mathbb{R})$*, and let* $\operatorname{spec}(A) = [\lambda_1, \dots, \lambda_n]$ *(a multiset of complex numbers). For every positive integer* k*,*

$$\operatorname{trace}(A^k) = \sum_{i=1}^{n} \lambda_i^k.$$

Proof. Let T be an upper triangular matrix similar to A in $M_n(\mathbb{C})$. Since A and T have the same spectrum, the diagonal entries of T are $\lambda_1, \dots, \lambda_n$ (in some order). Since A and T have the same trace we have $\operatorname{trace}(A) = \sum \lambda_i$.

Since $A = P^{-1}TP$ for some invertible $P \in M_n(\mathbb{C})$ we have

$$A^k = (PP^{-1}TP)(P^{-1}TP)\dots(P^{-1}TP) = P^{-1}T^kP,$$

so A^k is similar to T^k. Using the mechanism of matrix multiplication and the fact that T is upper triangular, it is straightforward to see that the diagonal entries of T^K are the kth powers of the corresponding diagonal entries of T. Thus

$$\operatorname{trace}(A^k) = \operatorname{trace}(T^k) = \sum_{i=1}^{k} \lambda_i^k.$$

☐