Frobenius Normal Form

August 30, 2026

Abstract

Walkthrough of a delightful connection between matrices and directed graphs which allows an elegant proof of the existence of a matrix normal form called the Frobenius Normal Form.

Keywords. Linear Algebra · Directed Graphs

1 Introduction

We are going to prove the **Frobenius Normal Form Theorem** [1] using an elegant connection between matrices and directed graphs. I got the idea of this walkthrough after watching "The Single Most Undervalued Fact of Linear Algebra" [2] (a slight exaggeration 🙂). We state the theorem here in the introduction as a goal and will describe the mathematical objects from the theorem and their properties in more detail in the rest of this walkthrough.

Theorem 1.1. ***(Frobenius Normal Form; see [3] and [4] for classical treatments).*** *Let $A \in \mathbb{F}_{n,n}$ be a square matrix over a field $\mathbb{F}$. Then there exists a permutation matrix $P \in \mathbb{F}_{n,n}$ such that PAP^T is in block upper triangular form:*

$$PAP^T = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ 0 & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{kk} \end{pmatrix} \tag{1.1}$$

where:

*(i) Each diagonal block $A_{ii} \in \mathbb{F}_{n_i,n_i}$ (with $\sum_{i=1}^{k} n_i = n$) is **irreducible**—that is, it cannot be further transformed into a block triangular form by any permutation similarity.*

(ii) Each diagonal block A_{ii} corresponds to a strongly connected component of the directed graph $G(A)$ associated with A.

(iii) The block upper triangular structure corresponds to a topological ordering of the condensation DAG of $G(A)$.

Moreover, the diagonal blocks $A_{11}, \dots, A_{kk}$ are uniquely determined up to simultaneous permutation within each block and the choice of topological sorting among incomparable components.

We start with permutation matrices in the next section, followed by sections on directed graphs, graph matrices and relabeling. This all will seem a little mysterious right now, but I promise it will pay off later in the last section. After we collect all the necessary properties of these mathematical objects, the proof of the theorem will fall out of this upfront effort in one short swoop. Daisuke Oyama's slide deck [5] guided me on this walkthrough so I hope I didn't veer off too far from the happy path towards the proof.

2 Permutation Matrices

Let $\pi \in S_n$ be a permutation of $\{1, 2, \ldots, n\}$. We construct a matrix $M \in \mathbb{F}_{n,n}$ that corresponds to π in the following way: row i of M has zeros in all columns except in column $\pi(i)$ where it has a one. $\mathbb{F}$ can be any field, it doesn't really matter. All we need are the additive and multiplicative neutral elements (zero and one). We call M the permutation matrix of π.

Let $\Phi : S_n \to \mathbb{F}_{n,n}$ be the function that maps a permutation π to its permutation matrix M: $\Phi(\pi) = M$.

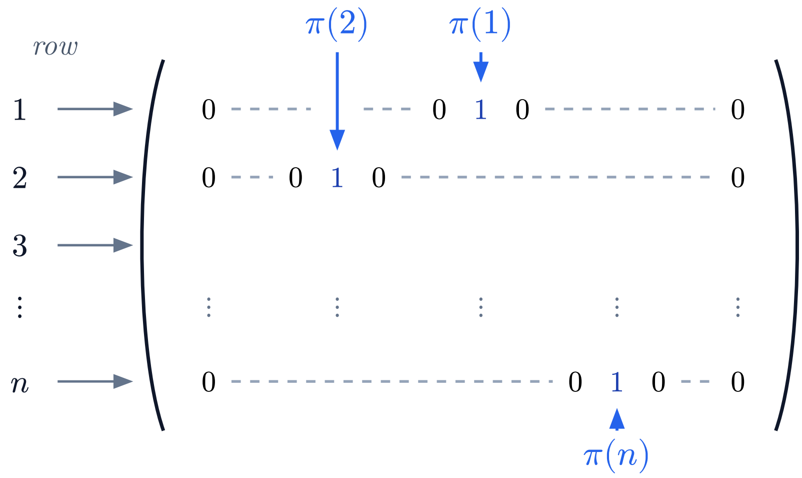


Figure 1: Visual representation of the permutation matrix $M = \Phi(\pi)$.

We can express a permutation matrix algebraically using the Kronecker delta function $\delta : \{1, \ldots, n\} \times \{1, \ldots, n\} \to \{0, 1\}$:

$$\delta(i, j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{2.1}$$

Then for $M = \Phi(\pi)$ we have $M_{i,j} = \delta(\pi(i), j)$, which says the same thing as the figure Figure 1: in row i, column j is zero unless it is the column equal to $\pi(i)$ in which case it is one.

Let's explore some of the properties of permutation matrices and of this mapping Φ.

Theorem 2.1. *Each row and each column in a permutation matrix has exactly one 1 and the rest zeros. Also, any matrix with exactly one 1 and the rest zeros in each row and column defines a permutation.*

Proof. The statement is already true by construction for rows. Assume that one column has more than one non-zero value. That means that the corresponding permutation π has two indices i, j such that $\pi(i) = \pi(j)$ which means π is not injective, a contradiction. If there is a column with only zeros, again it would be a contradiction to π being surjective because it would mean that no value in the domain $\{1, \ldots, n\}$ maps to that column number.

The inverse is also clear. Given a permutation matrix, we retrieve the corresponding permutation value for each $i \in \{1, \ldots, n\}$ by looking at the column number which has a one. □

It's pretty clear now that Φ is bijective, but we can go further:

Theorem 2.2. *Permutation matrices are closed under matrix multiplication.*

Proof. Let P and Q be two permutation matrices. Let's first treat P as a list of column vectors:

$$P = (p_1 \,|\, p_2 \,|\, \dots \; p_n) \tag{2.2}$$

Then the column j of the matrix product PQ is the linear combination

$$\sum_{k=1}^{n} Q_{k,j} p_k \tag{2.3}$$

but only one of the $Q_{k,j}$ is not zero, so the column j of PQ is one of the columns of P. This column of P already has the desired property: one 1 and rest zeros.

We do similarly for rows, which will be linear combinations of rows of Q but only one of the rows in the combination gets the one factor, so it's one of the rows of Q. A row of Q already has the desired property. □

Theorem 2.3. *Let π and φ be two permutations in S_n. Then*

$$\Phi(\varphi \circ \pi) = \Phi(\pi)\Phi(\varphi) \tag{2.4}$$

Here $\varphi \circ \pi$ is function composition, so $(\varphi \circ \pi)(i) = \varphi(\pi(i))$.

Proof. We already know from Theorem 2.2 that $\Phi(\pi)\Phi(\varphi)$ is a permutation matrix. We have to show that the corresponding permutation is indeed $\varphi \circ \pi$. We achieve this by following an index pair (i, j) through the expression chains. We can do this purely algebraically:

$$\begin{aligned}
\Phi(\pi)\Phi(\varphi)(i,j) &= \sum_{k=1}^{n} \Phi(\pi)_{i,k} \Phi(\varphi)_{k,j} \\
&= \sum_{k=1}^{n} \delta(\pi(i), k)\delta(\varphi(k), j) \\
&= \delta(\varphi(\pi(i)), j) = \delta((\varphi \circ \pi)(i), j) \\
&= \Phi(\varphi \circ \pi)_{i,j}
\end{aligned} \tag{2.5}$$

□

We now want to explore what effects permutation matrices have on general matrices when multiplied. We start with very simple permutations, namely transpositions. A transposition swaps two elements and leaves all the other elements untouched. The usual notation for a transposition π is cycle notation, which for a transposition is a pair $\pi = (i, j)$. This means $\pi(i) = j, \pi(j) = i, \text{ and } \pi(k) = k \quad \text{for all } k \neq i, j$. What does the permutation matrix of a transposition look like? We start with the identity matrix and just swap rows i and j.

Let $A \in \mathbb{F}_{n,n}$ be an arbitrary square matrix and let's multiply it on the right with the permutation matrix $\Phi((i, j))$. Since we are multiplying from the right, we are dealing with linear combinations of columns of A:

$$A = (a_1 \,|\, a_2 \,|\, \dots \; a_n) \tag{2.6}$$

Each column of $A\Phi((i, j))$ will be a column of A because the columns of $\Phi((i, j))$ are all zero except one which is 1. The linear combination effectively chooses a column of A. For most of the columns of the result it chooses the column of A with the same column number (because most columns of $\Phi((i, j))$ are identical to those of the identity matrix). But for two column numbers, namely i and j, the resulting columns of the product are the columns of A swapped.

Similarly, if we multiply A with $\Phi((i,j))$ on the left, we operate on rows using the exact same linear combination reasoning, resulting in rows i and j being swapped while the rest remain unchanged. In summary:

- **Left** multiplication $\Phi((i,j))A$ swaps **rows** i and j of A.
- **Right** multiplication $A\Phi((i,j))$ swaps **columns** i and j of A.

Now the interesting part: what happens if we multiply A from left and from right with $\Phi((i,j))$?

$$\Phi((i,j))A\Phi((i,j)) \tag{2.7}$$

We already know that rows and columns that are not numbered i or j will stay unchanged.

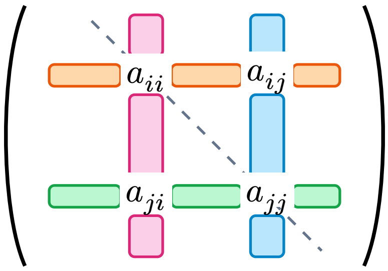


Figure 2: Partition of matrix A along rows i, j and columns i, j, showing the four intersection entries and the main diagonal.

Multiplying from the left by $\Phi((i,j))$ swaps rows i and j, while multiplying from the right by $\Phi((i,j))$ swaps columns i and j. The resulting matrix has both rows i, j and columns i, j interchanged:

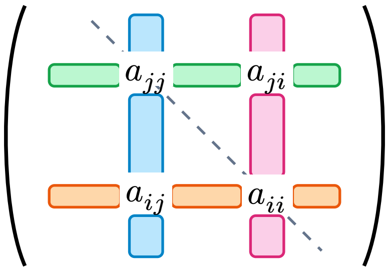


Figure 3: Resulting matrix $\Phi((i,j))A\Phi((i,j))$: the diagonal entries a_{ii} and a_{jj} are swapped, the off-diagonal entries a_{ij} and a_{ji} are swapped, and the row and column blocks are interchanged.

Notice something interesting that will become important later. Pretend that the row and column of a given matrix element encodes some information related to that element. So for Figure 2 element a_{ii} the orange and pink blocks contain information that is related to a_{ii}. Now observe in Figure 3 a_{ii} travelled to a new location and this relevant information (orange and pink) travelled with the element. If you follow the colors you can see that this is also true for a_{ij}, a_{ji} and a_{jj}. It is a direct and trivial consequence of swapping whole rows and columns.

We will explore more effects of permutation matrices on other matrices later on but first we have to switch topics and talk about directed graphs.

3 Directed Graphs

Given is a set of vertices $V = \{v_1, ..., v_n\}$. A directed graph is a tuple $G = (V, E \subset V \times V)$. The set E represents the edges of the graph. If $(v_i, v_j) \in E$ then the edge (v_i, v_j) is outgoing from v_i and incoming into v_j (this is the "directed" part of the graph). Vertex v_i is connected to vertex v_j if $v_i = v_j$ (the trivial, empty path), or if there are zero or more vertices $v_{k_1}, v_{k_2}, ..., v_{k_l}$

such that $(v_i, v_{k_1}), (v_{k_1}, v_{k_2}), ..., (v_{k_l}, v_j) \in E$. We say $(v_i, v_{k_1}, ..., v_{k_l}, v_j)$ is a path from v_i to v_j and v_i is connected to v_j. We use the notation $v_i \rightsquigarrow v_j$ if v_i is connected to v_j. We say v_i and v_j are strongly connected iff $v_i \rightsquigarrow v_j$ and $v_j \rightsquigarrow v_i$ and we use the notation $v_i \leftrightsquigarrow v_j$.

Theorem 3.1. *The binary relation $\leftrightsquigarrow \subset V \times V$ of being strongly connected is an equivalence relation on set V.*

Proof. We need to show the three properties
(i) reflexive
(ii) symmetric
(iii) transitive

All three properties are pretty straightforward to establish. □

Then $\leftrightsquigarrow$ induces a quotient set $V' = V/\leftrightsquigarrow$. Vertices of V that are members of a an equivalence class $v' \in V'$ are all strongly connected with each other. Vertices from two different equivalence classes on the other hand can only have at most one path connecting them. We can define a “collapsed” graph $G' = (V', E')$ which has an edge between two equivalence class v'_i and v'_j iff there is there exists an edge in V connecting a member of v'_i with a member of v'_j. $G' = G/\leftrightsquigarrow$ is the quotient graph or condensation graph.

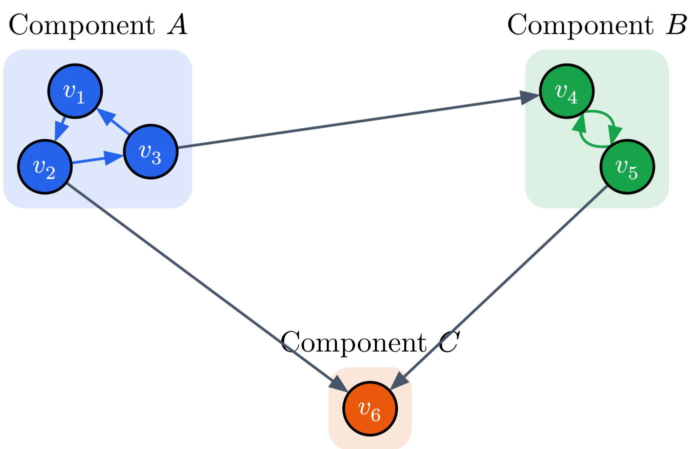


Figure 4: A directed graph G with three strongly connected components. v_1, v_2, v_3 (blue) form a cycle, v_4, v_5 (green) form a cycle, and v_6 (orange) is a trivial component on its own. The remaining edges $v_3 \to v_4$, $v_2 \to v_6$ and $v_5 \to v_6$ only ever point **between** components, never back.

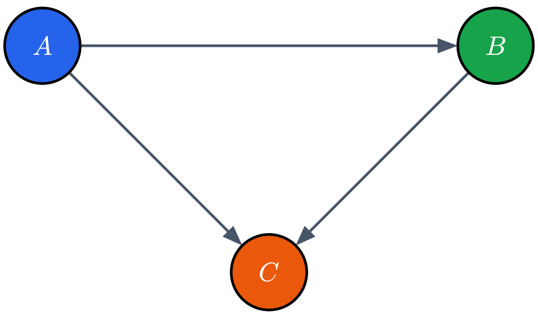


Figure 5: The quotient (condensation) graph $G' = G/\leftrightsquigarrow$ of Figure 4. Each strongly connected component collapses to a single vertex, and the inter-component edges $v_3 \to v_4$, $v_2 \to v_6$ and $v_5 \to v_6$ collapse to $A \to B$, $A \to C$ and $B \to C$. G' has no directed cycles.

Figure 4 shows an example directed graph together with its three strongly connected components, and Figure 5 shows the corresponding quotient graph G'. Notice that G' is necessarily a DAG ("directed acyclic graph"): if it contained a directed cycle, every component on that cycle could reach every other one in both directions, so they would all have been strongly connected in G already and would have collapsed into a single component in the first place. That's why we call it the condensation DAG (as mentioned in Theorem 1.1).

We're almost ready to attack the Frobenius Normal Form theorem but before that we need two more things: matrices connected to directed graphs and then what happens to a graph when we multiply a graph matrix with a permutation matrix from the left and from the right.

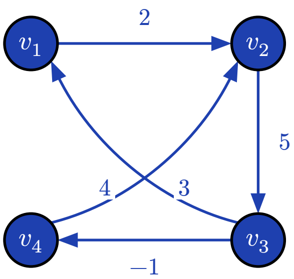
4 Graph Matrices

In this section we construct a matrix representation of a directed graph. Assume the graph is also weighted so that we have values to put into the matrix. Weighted means that each edge has a value associated with it, $G = (V, E)$ becomes the weighted graph $G = (V, \omega : E \to \mathbb{F})$ where ω is the weight function. Unfortunately we need one value in $\mathbb{F}$ that will represent "no edge". We usually pick the additive zero of the field for this so the weight cannot be zero then. $G = (V, \omega : E \to \mathbb{F} \setminus \{0\})$.

Then for a graph $G = (V = \{v_1, ..., v_n\}, \omega : E \to \mathbb{F} \setminus \{0\})$ we define a matrix $A \in \mathbb{F}_{nn}$:

$$A_{ij} = \begin{cases} \omega((v_i, v_j)) \text{ if } (v_i, v_j) \in E \\ 0 \text{ otherwise} \end{cases} \tag{4.1}$$

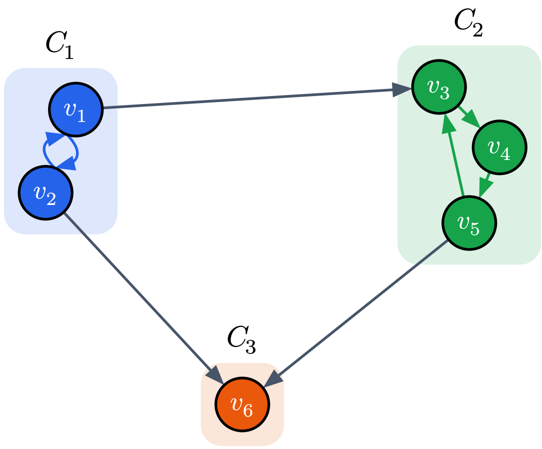
Conversely, given any matrix $A \in \mathbb{F}_{nn}$, we can recover a directed graph $G(A) = (V, E)$ from it: take $V = \{v_1, ..., v_n\}$ and $(v_i, v_j) \in E$ iff $A_{ij} \neq 0$.



$$A = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 4 & 0 & 0 & -1 \\ 0 & 3 & 0 & 0 \end{pmatrix} \tag{4.2}$$

Figure 6: A directed graph G on four vertices with edge weights (left) and the corresponding matrix A (right). Row i of A records the outgoing edges of v_i together with their weights: row 3, for instance, has a 4 in column 1 and a -1 in column 4 because $v_3 \to v_1$ has weight 4 and $v_3 \to v_4$ has weight -1.

Here is a preview of where all of this is heading. Suppose the vertices of the directed graph are numbered (labeled) such that in the condensation graph later condensed vertices never are connected back to earlier condensed vertices (this is called topological order). Since a topological order guarantees there is no edge running from a later component back to an earlier one, every matrix entry below the resulting diagonal blocks is forced to be zero. Figure 7 shows this for a small graph with two nontrivial strongly connected components and one trivial one.



$$A = \left(\begin{array}{cc|ccc|c} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{4.3}$$

Figure 7: A directed graph with vertices numbered strongly-connected-component by strongly-connected-component, in topological order: $C_1 = \{v_1, v_2\}$, $C_2 = \{v_3, v_4, v_5\}$, $C_3 = \{v_6\}$ (left). Because no edge points from a later component back to an earlier one, every entry of A below the block diagonal is zero (right) — already a hint of the shape in Theorem 1.1.

5 Relabeling

This is the last section before we prove the theorem. Assume we have a graph G with associated matrix A, and we also have a permutation matrix $P = \Phi(\rho)$ for some $\rho \in S_n$—not necessarily a transposition this time, any permutation at all. What does $G(PAP^T)$ look like, compared to $G(A)$? You can probably guess from the name of this section.

The simplest case is a good warm-up: if $\rho = (i, j)$ is a single transposition, Figure 3 already tells us that conjugating A by $P = \Phi(\rho)$ swaps rows i, j and columns i, j simultaneously, and since $\Phi(\rho)$ is symmetric here, $PAP^T = PAP$. The adjacency information of a matrix entry travels with the entry when the swap happens, so the graph itself is preserved—it's just that two vertices have been relabeled: what used to be vertex i is now vertex j, and vice versa.

Does the same thing happen for an **arbitrary** permutation, not just a swap of two labels? Figure 8 shows a genuine three-way rotation, $\rho = (2, 3, 4)$ in cycle notation (so $\rho(1) = 1$, $\rho(2) = 3$, $\rho(3) = 4$, $\rho(4) = 2$), applied to the same little graph. It is, again, the exact same picture —only the labels have moved, this time three of them instead of two. In general, $G(PAP^T)$ is always $G(A)$ with vertex $v_{\rho(p)}$ now called v_p, for **every** permutation ρ, not just transpositions. We will prove this properly—and put it to work—in the next section.

$G(A)$ $\qquad\qquad\qquad\qquad$ $G(PAP^T)$

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.1}$$

$$PAP^T = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \tag{5.2}$$

Figure 8: $G(A)$ (left) and $G(PAP^T)$ for the rotation $\rho = (2, 3, 4)$ (right), with the corresponding matrices below. Same vertices, same places, same edges—only the labels have been permuted (orange = moved; vertex 1, fixed by ρ, stays put). Read off one edge: $2 \to 3$ in $G(A)$ becomes $4 \to 2$ in $G(PAP^T)$, since vertex 2 is now called 4 and vertex 3 is now called 2—and indeed $A_{2,3} = 1$ while $(PAP^T)_{4,2} = 1$.

6 The Proof

We now have collected all the ingredients to prove Theorem 1.1. We will do so in a constructive way by providing the recipe to transform a given matrix to its Frobenius Normal Form.

Here's the recipe in broad strokes:

We start with a matrix $A \in \mathbb{F}_{nn}$. We consider the graph $G(A)$ defined by A as an adjacency matrix. We condense it using the equivalence relation $\leftrightsquigarrow$ to a condensed DAG. We relabel the vertices from this condensed DAG in topological order using a permutation matrix P. Then PAP^T is in Frobenius Normal Form.

Three things we haven't shown yet: that a topological order of the condensed DAG always exists in the first place, exactly what relabeling by an arbitrary permutation matrix P (not just a single transposition) does to A and $G(A)$, and that the resulting diagonal blocks are irreducible. Let's close all three gaps and then assemble the proof.

Theorem 6.1. *Let $H = (W, F)$ be a finite directed acyclic graph (a directed graph with no directed cycles). Then there is a labeling of W by $1, ..., |W|$ such that every edge of H points from a lower-labeled vertex to a higher-labeled one. We call such a labeling a* ***topological order*** *of H.*

This is a standard fact; we won't prove it here, but the idea is simple enough: a finite DAG always has a **source** (a vertex with no incoming edges—otherwise you could walk backward along incoming edges forever, and with finitely many vertices you'd eventually repeat one, giving a cycle), so strong induction on $|W|$ works by peeling off a source, labeling it first, and topologically ordering the rest.

Theorem 6.2. *Let $\rho \in S_n$ and $P = \Phi(\rho)$. Then for every $A \in \mathbb{F}_{n,n}$ and all $p, q \in \{1, ..., n\}$,*

$$(PAP^T)_{pq} = A_{\rho(p),\rho(q)}. \tag{6.1}$$

In particular, $G(PAP^T)$ is exactly $G(A)$ with the vertex $v_{\rho(p)}$ relabeled v_p, for every p.

Proof. We do this algebraically using the fact from Section 2 that $P_{a,b} = \Phi(\rho)_{a,b} = \delta(\rho(a), b)$. Then

$$(PA)_{p,b} = \sum_a P_{p,a} A_{a,b} = \sum_a \delta(\rho(p), a) A_{a,b} = A_{\rho(p),b}, \tag{6.2}$$

since only the term $a = \rho(p)$ survives, and, since $P^T_{b,q} = P_{q,b} = \delta(\rho(q), b)$,

$$\left(PAP^T\right)_{p,q} = \sum_b (PA)_{p,b} P^T_{b,q} = \sum_b A_{\rho(p),b} \delta(\rho(q), b) = A_{\rho(p),\rho(q)}, \tag{6.3}$$

again because only $b = \rho(q)$ survives. Consequently $\left(PAP^T\right)_{p,q} \neq 0$ exactly when $A_{\rho(p),\rho(q)} \neq 0$, i.e. exactly when $\left(v_{\rho(p)}, v_{\rho(q)}\right) \in E(G(A))$—which is exactly the claimed relabeling of $G(A)$. □

For a single transposition $\rho = (i, j)$, $\Phi(\rho)$ is symmetric, so $PAP^T = PAP$ and we get exactly the swap discussed at the start of Section 5; Figure 8 shows this for a more general permutation that shuffles more than two vertices.

We used the word **irreducible** in Theorem 1.1 so let's define it precisely.

Definition 6.3. *A matrix $M \in \mathbb{F}_{m,m}$ is* ***irreducible*** *if there is no permutation matrix R for which RMR^T is block upper triangular with two or more diagonal blocks. Otherwise M is* ***reducible****.*

In other words, M is irreducible exactly when the construction of Theorem 1.1 applied to M itself can't split it any further—it's already as condensed as a diagonal block can get. The next theorem shows that this is the same thing as $G(M)$ being strongly connected.

Theorem 6.4. *Let $M \in \mathbb{F}_{m,m}$, let $B_1, ..., B_r$ be the strongly connected components of $G(M)$, listed in a topological order of the condensation graph $G(M)/\leftrightsquigarrow$ (which exists by Theorem 6.1, since a condensation graph is always a DAG—as we saw discussing Figure 5), and let $\sigma \in S_m$ enumerate the vertices of B_1 first, then those of B_2, and so on. Then for $Q = \Phi(\sigma)$, the matrix QMQ^T is block upper triangular with diagonal blocks $M_{11}, ..., M_{rr}$ of sizes $|B_1|, ..., |B_r|$, and each M_{ii} is irreducible.*

Proof. **Block upper triangular.** By Theorem 6.2, $\left(QMQ^T\right)_{p,q} = M_{\sigma(p),\sigma(q)}$, so $G\left(QMQ^T\right)$ is $G(M)$ with vertices relabeled by σ. Suppose p falls in the range assigned to B_a and q in the range assigned to B_b with $a > b$, and suppose toward a contradiction that $\left(QMQ^T\right)_{p,q} \neq 0$. Then $\left(v_{\sigma(p)}, v_{\sigma(q)}\right) \in E(G(M))$ with $v_{\sigma(p)} \in B_a$ and $v_{\sigma(q)} \in B_b$, which by definition of the condensation graph gives an edge $B_a \to B_b$ in $G(M)/\leftrightsquigarrow$. Since $B_1, ..., B_r$ is a topological order, this forces $a < b$, contradicting $a > b$. So $\left(QMQ^T\right)_{p,q} = 0$ whenever p's block index exceeds q's—that is, QMQ^T is block upper triangular, with block (i, i) consisting of exactly the entries of M among the vertices of B_i, by construction of σ.

Diagonal blocks irreducible. Fix i and write $B = B_i$. First, any path in $G(M)$ between two vertices of B stays entirely inside B: if $u, v \in B$ and $u \to w_1 \to ... \to w_l \to v$ is a path, then for each w_t we have $u \rightsquigarrow w_t$ (the path so far) and $w_t \rightsquigarrow v \leftrightsquigarrow u$ (the rest of the path, then $v \leftrightsquigarrow u$ since both are in B), so $w_t \leftrightsquigarrow u$ too, i.e. $w_t \in B$. Hence the graph of the diagonal block M_{ii} —which is exactly $G(M)$ restricted to B—is strongly connected: B is a strongly connected component of the whole graph, and we just saw that no connecting path ever needs to leave B.

Assume that M_{ii} is reducible: some permutation matrix R makes $RM_{ii}R^T$ block upper triangular with $s \geq 2$ diagonal blocks $E_1, ..., E_s$. Pick a vertex u in the **last** block E_s and a vertex v in the **first** block E_1. Since $G(M_{ii})$ is strongly connected there is a path from u to v, and by Theorem 6.2 this path carries over unchanged (just relabeled) to a path between the corresponding vertices of $G(RM_{ii}R^T)$. But in a block upper triangular matrix every edge goes from a block to an equal-or-later one, so no path can ever move from a later block to an earlier one—yet we would need one from block s to block 1, and $s \geq 2$ means these are genuinely different blocks. Contradiction. So M_{ii} is irreducible. $\square$

Proof. **(of Theorem 1.1)**. Apply Theorem 6.4 to $M = A$: writing $C_1, ..., C_k$ for the strongly connected components of $G(A)$ in a topological order of the condensation, $n_i = |C_i|$ (so $\sum_{i=1}^{k} n_i = n$), and $P = \Phi(\sigma)$ for the corresponding enumeration σ, the matrix PAP^T is block upper triangular with diagonal blocks $A_{11}, ..., A_{kk}$ of sizes $n_1, ..., n_k$, each irreducible, each corresponding exactly to the strongly connected component C_i, and the block order is by construction a topological order of the condensation DAG of $G(A)$. This is exactly the claimed Frobenius Normal Form, proving existence.

Uniqueness. Before the formal argument, here is the idea: relabeling doesn't change the graph, only what we call its vertices (Theorem 6.2), so the strongly connected components $C_1, ..., C_k$ and the condensation DAG they form are already uniquely determined by A itself, independent of any choice of P. Any valid block upper triangular decomposition can therefore only be reading off this same, unique condensation—as diagonal blocks in some topological order of it. What follows just makes that precise.

Suppose Q is any permutation matrix such that QAQ^T is block upper triangular with irreducible diagonal blocks $D_1, ..., D_r$, and let $D_1, ..., D_r$ also denote the corresponding vertex sets of $G(A)$ (via Theorem 6.2). We show $\{D_1, ..., D_r\} = \{C_1, ..., C_k\}$ as sets.

Fix i and let $C = C_i$. Since block upper triangularity forbids any edge from a later block back to an earlier one (the same argument as in Theorem 6.4), and C is strongly connected, no two vertices of C can lie in different blocks: if $u, v \in C$ were in different blocks with u's earlier than v's, the required path $v \rightsquigarrow u$ would have to move from a later block to an earlier one, which is impossible. So C lies entirely inside a single block, say D_j.

Conversely, since D_j is irreducible, it is strongly connected: if it were not, applying Theorem 6.4 to $M = D_j$ itself would exhibit a further nontrivial block triangularization of D_j, contradicting its irreducibility. A strongly connected D_j containing C, together with C being a **maximal** strongly connected set of vertices of $G(A)$, forces $D_j = C$.

So every C_i equals some D_j and, by the same argument in reverse, every D_j equals some C_i: the two lists of blocks agree as sets of vertices, so in particular $r = k$.

Finally, the order in which the blocks may be listed is constrained only by: whenever the condensation has an edge $C_a \to C_b$, block upper triangularity forces $a < b$ (again by the no-backward-edges argument)—that is, the order must be **some** topological order of the condensation DAG, and by Theorem 6.4 any topological order is achievable. Components with no directed path between them in the condensation are free to appear in either relative order. And within a single block C_i, any ordering of its n_i vertices leaves the block upper triangular structure intact and simply permutes A_{ii}'s rows and columns simultaneously—exactly the freedom claimed in the theorem. $\square$

Bibliography

[1] Wikipedia contributors, "Perron–Frobenius theorem — Wikipedia, The Free Encyclopedia." [Online]. Available: https://en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem

[2] T. Danka, "The Single Most Undervalued Fact of Linear Algebra." [Online]. Available: https://youtu.be/8gNKE5jVqYY

[3] J. L. Stuart, "Digraphs and Matrices," in *Handbook of Linear Algebra*, L. Hogben, Ed., Chapman, Hall/CRC, 2006, ch. 29.

[4] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory.* in Encyclopedia of Mathematics and its Applications, No. 39. Cambridge: Cambridge University Press, 1991.

[5] D. Oyama, "Nonnegative Matrices I." [Online]. Available: https://www.oyama.e.u-tokyo.ac.jp/theory17/theory17nonneg_mat01R.pdf