Pescara, Italy, July 2019



DIGRAPHS I Mathematical Background: Perron-Frobenius, Jordan Normal Form, Cauchy-Binet, Jacobi's Formula

Based on various sources.

J. J. P. Veerman, Math/Stat, Portland State Univ., Portland, OR 97201, USA. email: veerman@pdx.edu

> Conference Website: www.sci.unich.it/mmcs2019

SUMMARY:

* This is a review of four theorems from linear algebra that are important for the development of the algebraic theory of directed graphs. These theorems are the Perron-Frobenius theorem, the Cauchy-Binet formula, the Jordan Normal Form, and Jacobi's Formula.

OUTLINE:

The headings of this talk are color-coded as follows:

Graph Theory Definitions

Perron-Frobenius

Jordan Normal Form



Jacobi's Formula



•

Definitions: Digraphs

Definition: A directed graph (or **digraph**) is a set $V = \{1, \dots, n\}$ of **vertices** together with set of ordered pairs $E \subseteq V \times V$ (the **edges**).



A directed edge $j \rightarrow i$, also written as ji. A directed path from j to i is written as $j \rightsquigarrow i$.

Digraphs are everywhere: models of the internet [6], social networks [7], food webs [11], epidemics [10], chemical reaction networks [12], databases [5], communication networks [4], and networks of autonomous agents in control theory [8], to name but a few.

A BIG topic: Much of mathematics can be translated into graph theory (discretization, triangulation, etc). In addition, many topics in graph theory that do not translate back to *continuous* mathematics.

Definitions: Connectedness of digraphs

Undirected graphs are connected or not. But...



Definition:

* A digraph G is **strongly connected** if for every ordered pair of vertices (i, j), there is a path $i \rightsquigarrow j$. **SCC**!

* A digraph G is **unilaterally connected** if for every ordered pair of vertices (i, j), there is a path $i \rightsquigarrow j$ or a path $j \rightsquigarrow i$.

* A digraph G is **weakly connected** if the **underlying UNdirected graph** is connected.

* A digraph G is **not connected:** if it is not weakly connected.

Definition:

Multilaterally connected: weakly connected but not unilaterally connected.

The Adjacency Matrix

Definition: The combinatorial adjacency matrix Q of the graph G is the matrix whose entry $Q_{ij} = 1$ if there is an edge ji and equals 0 otherwise.

Interpretation: We think of $Q_{ij} = 1$ as information going from j to i. Or: i "sees" j. In the graph below, both 2 and 6 "see" 1. So $Q_{21} = Q_{61} = 1$.





•

Non-Negative Matrices

Definition: A non-negative matrix Q is **irreducible** if for every i, j, there is a k such that $(Q^k)_{ij} > 0$.

OR: for all i, j, there is path from j to $i: j \rightsquigarrow i$.

Definition: A non-negative matrix Q is **primitive** if there is a k such that for every i, j, we have $(Q^k)_{ij} > 0$.

OR: $\exists k$ such that for all i, j, there is $j \rightsquigarrow i$ of length k.

Q is adjacency matrix of graph G. Both imply that G is SCC.

Irreducible but not primitive: any cyclic permutation.



Perron-Frobenius

The single most important theorem in algebraic graph theory!! Gives leading eigenpair of many important matrices. 1st order description of dynamical processes on graphs. More details in [1] and [13].

Theorem 1A: Let $A \ge 0$ be irreducible. Then: (a) Its spectral radius $\rho(A)$ is a simple eval of A. (b) Its associated evec is the only strictly positive evec.

Thus its largest eval is simple, real, and positive. But there may be other evals of the same modulus.

Theorem 1B: Let $A \ge 0$ be primitive. Then also: All other evals have modulus strictly smaller than $\rho(A)$.

(Note 3-fold rotational symmetry in irreducible case.)



Irreducible Has Period p

In the irreducible case, the matrix A has a **period** p > 1. That is: after permutation of vertices, A is **block cyclic**. Example: p = 3:

$$A = \left(\begin{array}{rrrr} 0 & A_1 & 0 \\ 0 & 0 & A_2 \\ A_3 & 0 & 0 \end{array}\right)$$

In this **cyclic block form**, the A_i are **rectangular**!

Exercise 1: Show that

$$A^{3} = \begin{pmatrix} A_{1}A_{2}A_{3} & 0 & 0 \\ 0 & A_{2}A_{3}A_{1} \\ 0 & 0 & A_{3}A_{1}A_{2} \end{pmatrix}$$

Now, the diagonal blocks are primitive.

By **Cauchy-Binet** (later): each diagonal block D of A^3 has same non-zero spectrum. Suppose non-zero spectrum D is: $\{\lambda_i\}_{i=1}^s$.

The non-zero spectrum of A consists of **all 3rd roots** of these.





So, Q is block-triangular and thus *not* irreducible. But: The two non-trivial blocks are **irreducible but** *not* **primitive**. Notice the grouping of the evals.

The spectrum is $\{0, 0, 1, e^{2\pi i/3}, e^{-2\pi i/3}, 1, -1\}.$

Other Eigenvectors

Theorem 1C: Let A be irreducible. Any other evec but the leading <u>cannot</u> be real and non-negative.

This is clear if the eigenvalue is non-real. So only needs proof for real evecs.

This is the beginning of the study of **Nodal Domains**. A classical problem in analysis (since Courant): count the

number of nodal domains of e.fns to the Laplace operator. See Figure.



For undirected graphs there are many results. But for digraphs very little is known. (After all, evecs may not be real!)

JORDAN NORMAL FORM

•

Spectral Theorem

From now: A is $n \times n$ matrix with real or complex coeff's: **real symmetric** \subset **self-adjoint** \subset **normal.** (A is normal if $A^*A = AA^*$.)

Theorem 2 (spectral): A has orthonormal basis of evecs $\{v_i\}_{i=1}^n$ iff A normal.

These evals are real, if A is self-adjoint.

Computations simplify (e.g. quantum mechanics and statistical physics):

Let A a (normal) matrix with e.pairs $\{\lambda_i, v_i\}$. Suppose $\dot{x} = Ax$ with initial condition $x(0) = x_0$. Then:

$$x(t) = \sum_i (v_i, x_0) e^{\lambda_i t} v_i$$

where (., .) is real or Hermitian inner product. (v_i, x_0) is the orthogonal projection of x_0 onto v_i .

Exercise 2: The matrix norm $||A|| \equiv \sup_x \{Ax \mid |x| = 1\}$ equals norm of its largest eval if A is normal. (*Hints: a*) Show $\sum (v_i, x)^2 = 1$; b) Show that $Ax = \sum \lambda_i (v_i, x)$; c) Show that (Ax, Ax) is a weighted mean of λ_i^2 .)

Life in a Non-normal Universe



Let $\dot{x} = Ax$. Sps evecs v_1 and v_2 nearly parallel.

$$x(t) = A_1 e^{\lambda_1 t} v_1 + A_2 e^{\lambda_2 t} v_2$$

Example: $\lambda_i = \{-0.1, -1.0\}$ and init. condin x(0) as indicated.

Large **transient**! Stable system may initially "look" unstable. Below we plot |x(t)|.



Exercise 3: Define a 2-dim. system of ODE plus initial condition that exhibits this type of behavior.

Case I: *n* **Eigenvectors**

Let A be $n \times n$ matrix.

In general, it may have real and/or complex epairs.

Evals are the solutions $\{\lambda_i\}_{i=1}^k$ (with $k \leq n$) of

 $\det(A - \lambda I) = 0$

Case I: *n* linearly independent evecs $\{v_i\}_{i=1}^n$. Given λ_i , then $\{v_i\}$ is the solution of

$(A-\lambda_i I)v=0$

Let H the matrix whose *i*th column equals v_i . Then A is **diagonalizable**, or:

$D = H^{-1}AH$

with D diagonal with $D_{ii} = \lambda_i$ (real if A is self-adjoint).

Application: Suppose $\dot{x} = Ax$ with init. cond. x_0 . Then:

$$x(t) = \sum_i lpha_i e^{-\lambda_i t} v_i$$

But the α_i are less simple to calculate. Set t = 0, you get:

$$Hlpha=x_0$$

Case II: Less than n Eigenvectors

Let A be $n \times n$ matrix.

Case II: less than *n* **linearly independent** evecs $\{v_i\}_{i=1}^n$.

This happens when for some i, λ_i is a root of order \underline{k} of

$$\det(A - \lambda I) = 0$$

but

$$(A - \lambda_i I)v = 0$$

has <u>less than k</u> linearly independent solutions for v.

Definition: The <u>algebraic</u> multiplicity of an eigenvalue λ_i of A is the **order** of the root λ_i of det $(A - \lambda I)$. The <u>geometric</u> multiplicity of λ_i is the **number** of linearly independent evecs associated with λ_i .

In this case A is not diagonalizable but **block diagonaliz**able. There is matrix H so that

$$J = H^{-1}AH$$

Exercise 4: *J* has diagonal **Jordan blocks** (or JB), all of the form:

$$B_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots \\ 0 & \lambda_i & 1 & \dots \\ \dots & \dots & \dots & 1 \\ \dots & \dots & 0 & \lambda_i \end{pmatrix}$$

Case II: Not Enough LI Eigenvectors

Find all evals λ satisfying

 $\det(A - \lambda I) = 0$

For each eval λ_i , find its evecs:

 $(A - \lambda_i I)v = 0$

These vectors span the **eigenspace** of λ_i . For simplicity: assume there is only one: v_i .

If geom mult (λ_i) < alg mult (λ_i) : Start with evec v_i . Find vector w_{i1} such that

$$(A-\lambda_i I)w_{i1}=v_i$$

Find w_{i2} such that

 $(A-\lambda_i I)w_{i2}=w_{i1}$

Etc. The v_i together with w_{ij} are **generalized eigenvec**tors. They span the **generalized eigenspace** of λ_i .

Thus there are exactly n linearly independent generalized eigenvectors v_i .

Case II: Construction of the Matrix H

H is the matrix whose columns are:

 $\{v_1, w_{11}, \cdots, w_{1n_1}, v_2, w_{21}, \cdots, w_{2n_2}, \cdots, v_k, w_{k1}, \cdots, w_{kn_k}\}$ equals v_i . Then

$J = H^{-1}AH$

and J has non-trivial Jordan blocks.

Example: If 1st block has dim ≥ 3 (or $n_1 \geq 2$):

$$\lambda_1 e_1 \xleftarrow{H^{-1}} \lambda_1 v_1 \xleftarrow{A} v_1 \xleftarrow{H} e_1$$

$$\lambda_1 e_2 + e_1 \stackrel{H^{-1}}{\longleftarrow} \lambda_1 w_{11} + v_1 \stackrel{A}{\longleftarrow} w_{11} \stackrel{H}{\longleftarrow} e_2$$

$$\lambda_1 e_3 + e_2 \stackrel{H^{-1}}{\longleftarrow} \lambda_1 w_{12} + w_{11} \stackrel{A}{\longleftarrow} w_{12} \stackrel{H}{\longleftarrow} e_3$$

Definition: Thus J becomes:

λ_1	1	0	• • •)
0	λ_1	1	• • •	
0	0	λ_1	• • •	
(• • •	•••	• • •	· · · <i>J</i>

This is called **Jordan normal form**.

$\dot{x} = Ax$, General Case

Exercise 1: Let I be the identity and

$$N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad J = \lambda I + N = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

a) Compute e^{Jt} via the usual expansion.

(*Hint:* $e^{\lambda t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$.) b) Use a) to give solutions of $\dot{x} = Jx$, where $x(0) = (a_1, a_2)^T$. (*Hint:* $e^{\lambda t} \begin{pmatrix} a_1 + a_2 t \\ a_2 \end{pmatrix}$.)

The expansion of e^{Jt} in the exercise

$$e^{Jt} = I + Jt + \frac{J^2t^2}{2} + \frac{J^3t^3}{3!} + \cdots$$

simplifies because $J = \lambda I + N$ and $N^2 = 0$.

Back to the general problem $\dot{x} = Ax$, $x(0) = x_0$. Step 1: Write init. cond as sum of gener. evecs.

$$x_0 = \sum \alpha_i v_i$$
 where $H\alpha = x_0$

Step 2: Suppose $x_0 = \alpha_{12}w_{12}$. Then

$$x(t) = \alpha_{12} e^{\lambda t} \left(\frac{t^2}{2} v_1 + t w_{11} + w_{12} \right)$$

Step 3: Sum those contributions.



Two digraphs. The first has adjacency matrix

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

with spectrum $\{1.68, -1.03 \pm 0.74i, 0.37\}$ (approximately). The second has adjacency matrix

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

with spectrum $\{0^{(2)}, \pm \sqrt{2}\}$. The eigenvalue 0 has an associated 2-dimensional Jordan block.

Additional Exercises

Exercise 2: Show that the matrix

$$\left(\begin{array}{cc} a-b & c\\ -cd & a+b \end{array}\right)$$

has a non-trivial Jordan block (JB) if $b^2 = c^2 d$ and $c \neq 0$ and $d \neq 0$.

Exercise 3: So you may think JB's are rare (co-dimension one). But symmetries can change that. Show that

a) Newton's equation $\ddot{x} = 0$ gives rise to a JB.

b) That JB explains why two bodies without forcing separate linearly in time (Newton's first law).



•

Generalized Cauchy-Binet

A is a $n \times e$ matrix and B is a $e \times m$ matrix.



Notation: $k \leq n, m \leq e$. (See figure). Let $I \subseteq \{1, \dots, n\}$, $J \subseteq \{1, \dots, m\}$, and $K \subseteq \{1, \dots, e\}$. All subsets have the <u>same</u> cardinality k.

Definition: The matrix consisting of the entries of A in $I \times K$ is called a **minor** of A. **Principal minor** if I = K. It is denoted by A[I, K].

Theorem 3 (generalized Cauchy-Binet):

$$\det\left((AB)[I,J]\right) = \sum_{K} \, \det(A[I,K]) \det(B[K,J])$$

where the sum is over all $K \subseteq \{1, \dots e\}$ with |K| = k.

Corollaries

A and B as depicted, where $n \leq e$. Now $I = J = \{1, \dots, n\}$



Corollary (Cauchy-Binet): We have

 $\det (AB) = \sum \det(A[J, K]) \det(B[K, J])$ where the sum is over all $K \subseteq \{1, \dots e\}$ with |K| = n. If X is $n \times n$, by standard matrix computation $\det(X + z Id) = \dots + z^{n-k} \sum_{|K|=k} \det X[K, K] + \dots$ By generalized C-B, we also have for $k \leq n$: $\sum_{|K|=k} \sum_{|L|=k} \det A[K, L] \det B[L, K]$ equals $\sum_{|K|=k} \det(AB)[K, K]$ and $\sum_{|L|=k} \det(BA)[L, L]$. Corollary: We have

 $\det(BA + z Id) = z^{e-n} \det(AB + z Id)$

Sketch of Proof of Cauchy-Binet

Inspired by Gessel-Viennot [9].



Crossing paths give canceling contributions.

For the crossing as pictured (right figure):

$$(AB)_{11}(AB)_{22}(AB)_{3\sigma(3)}\cdots = (AB)_{12}(AB)_{21}(AB)_{3\sigma(3)}\cdots$$

All other terms equal. But σ changes by 1 transpos.: $1 \leftrightarrow 2$. **Thus:**

$$\prod_{i \in I} \sum_{\ell \in E} A_{i\ell} B_{\ell\sigma(i)} = \left(\sum_{\ell_1} A_{1\ell_1} B_{\ell_1\sigma(1)} \right) \left(\sum_{\ell_2} A_{2\ell_2} B_{\ell_2\sigma(2)} \right) \cdots \left(\sum_{\ell_n} A_{n\ell_n} B_{\ell_n\sigma(n)} \right)$$

Where now $\overline{\ell} \equiv (\ell_1, \dots \ell_n)$ runs over <u>ALL</u> *n*-tuples of <u>distinct</u> members of *E*. (No two ℓ_i 's are the same.)

Sketch of Proof Continued

Idea: Regroup the sums so that for any given $K \subset E$ of size n, the permutations of K are grouped together. Then:

$$\prod_{i \in I} \sum_{\ell \in E} A_{i\ell} B_{\ell\sigma(i)} = \sum_{K,|K|=n} \left(\sum_{\ell_1} A_{1\ell_1} B_{\ell_1\sigma(1)} \right) \cdots \left(\sum_{\ell_n} A_{n\ell_n} B_{\ell_n\sigma(n)} \right)$$

For each K, the ℓ_i run over the permutations of K.

For fixed K, choose permutations ρ and τ :

$$I \xrightarrow{\tau} K \xrightarrow{\rho} J$$
 so that $\rho(\tau) = \sigma$:
 $\dots = \sum_{|K|=4} \prod_i \sum_{\tau} A_{i\tau(i)} B_{\tau(i)\rho(\tau(i))}$

We obtain:

 $\det AB = \sum_{|K|=4} \sum_{\sigma} \operatorname{sgn} \sigma \prod_{i} \sum_{\tau} A_{i\tau(i)} B_{\tau(i)\rho(\tau(i))}$

For fixed K, this is determ. of product of square matrices: $\cdots = \sum_{|K|=4} \det(A[I,K]) \det(B[K,J])$



•

The Formula and Its Corollaries

A a square matrix, $\mathbf{adj}(A)$ its $\mathbf{adjugate}$:

$$A \operatorname{adj}(A) = \operatorname{adj}(A) A = \det(A) I$$

Suppose A depends (differentiably) on a parameter t.

Theorem 4:
$$\frac{d}{dt} \det(A) = \operatorname{Tr}\left(\operatorname{adj}(A) \frac{dA}{dt}\right).$$

We give some common corollaries as easy exercises.

Replace
$$\frac{dA}{dt}$$
 by B whose only non-zero entry is $B_{k\ell} = 1$:
Exercise 4: Show $\frac{d}{dA_{k\ell}} \det(A) = (\operatorname{adj}(A))_{\ell k}$.

Instead, replace A by e^{Bt} and so $\operatorname{adj}(A)$ by $e^{-Bt} \operatorname{det}(e^{Bt})$: **Exercise 5:** Show $\frac{d}{dt} \operatorname{det}(e^{tB}) = \operatorname{Tr}(B) \operatorname{det}(e^{tB})$.

The latter gives an ODE. Solve it: **Exercise 6:** Show the latter implies: $det(e^{tB}) = e^{Tr(Bt)}$.

Sketch of Proof

B has evals λ_i (with mult.). Then $I + \epsilon B$ has evals $1 + \epsilon \lambda_i$: $\det(I + \epsilon B) = \prod_i (1 + \epsilon \lambda_i)$

Thus

$$\begin{split} \lim_{\epsilon \to 0} \frac{\det(I + \epsilon B) - \det(I)}{\epsilon} &= \sum_{i} \lambda_{i} = \operatorname{Tr}(B) \\ \text{For an invertible } A: \\ \lim_{\epsilon \to 0} \frac{\det(A + \epsilon B) - \det(A)}{\epsilon} &= \\ \lim_{\epsilon \to 0} \frac{\det(A) \left[\det(I + \epsilon A^{-1}B) - \det(I)\right]}{\epsilon} &= \\ \det(A) \operatorname{Tr}(A^{-1}B) \end{split}$$

Extend to non-invertible: replace $det(A) A^{-1}$ by adj(A):

 $\cdots = \operatorname{Tr} (\operatorname{adj}(A) B)$

... And replace B by $\frac{dA}{dt}$:

$$\cdots = \operatorname{Tr}\left(\operatorname{adj}(A) \, rac{dA}{dt}
ight)$$



References

- [1] M. Boyle, Notes of the Perron-Frobenius Theory of Nonnegative Matrices, https://www.math.umd.edu/~mboyle/courses/475sp05
- [2] J. S. Caughman, J. J. P. Veerman, Kernels of Directed Graph Laplacians, Electronic Journal of Combinatorics, 13, No 1, 2006.
- [3] J. J. P. Veerman, E. Kummel, Diffusion and Consensus on Weakly Connected Directed Graphs, Linear Algebra and Its Applications, accepted, 2019.
- [4] R. Ahlswede et al., Network Information Flow, IEEE Transactions on Information Theory, Vol. 46, No. 4, pp. 1204-1216, 2000.
- [5] R. Angles, C. Guiterrez, Survey of Graph Database Models, ACM Computing Surveys, Vol. 40, No. 1, pp. 1-39, 2008.
- [6] A. Broder et al., Graph Structure of the Web, Computer Networks, 33, pp. 309-322, 2000.
- [7] P. Carrington, J. Scott, S. Wasserman, Models and Methods in Social Network Analysis, Cambridge University Press, 2005.
- [8] J. Fax, R Murray, Information Flow and Cooperative Control of Vehicle Formations, IEEE Transactions on Automatic Control, Vol. 49, No. 9, 2004.

- [9] I. Gessel, X. Viennot, Binomial determinants, paths, and hook length formulae, Adv. Math., 58 (1985), pp. 300-321
- [10] T. Jombert et al., Reconstructing disease outbreaks from genetic data: a graph approach, Heredity 106, 383-390, 2011.
- [11] Robert M. May, Qualitative Stability in Model Ecosystems, Ecology, Vol. 54, No. 3. (May, 1973), pp. 638-641.
- [12] S. Rao, A. van der Schaft, B. Jayawardhana, A graphtheoretical approach for the analysis and model reduction of complex-balanced chemical reaction networks, J. Math. Chem., Vol. 51, No. 9, pp. 2401-2422, 2013.
- [13] S. Sternberg, Dynamical Systems, Dover Publications, Mineola, NY, 2010, revised edition 2013.