Eigenvalue Bounds for the Finite-State Birth-Death Process Intensity Matrix

Renzo Roel P. Tan

renzopereztan@gmail.com

Len Patrick Dominic M. Garces

lgarces@ateneo.edu

Department of Mathematics
Ateneo de Manila University
Quezon City, Philippines

May 18, 2018

ATENEO

Introduction and Preliminaries

ATENEO

Continuous-Time Markov Chain

Definition (Continuous-Time Markov Chain)

A stochastic process $\{X(t) : t \geq 0\}$ with a discrete state space S is a **continuous-time Markov chain** (CTMC) if for all $t \geq 0$, $s \geq 0$, and $i, j \in S$, it holds that

$$\begin{aligned} &\mathbb{P}\left[X(t+s) = j | X(s) = i, \{X(u) : 0 \leq u < s\}\right] \\ &= \mathbb{P}\left[X(t+s) = j | X(s) = i\right]. \end{aligned} \tag{1}$$

ATENEO

Continuous-Time Markov Chain

- The quantity $P_{ij} = \mathbb{P}\left[X(t+s) = j | X(s) = i\right]$ is the probability of transitioning to state j from state i.
- For a CTMC, the future realization $X(s+t)$ is only dependent on the current state $X(s)$ and not on information occurring prior to time s.
- The matrix $\mathbf{P} = [P_{ij}]$ corresponding to a CTMC is known as the matrix of **transition probabilities**.

ATENEO

Birth-Death Process

Definition (Birth-Death Process)

A **birth-death process** is a continuous-time Markov chain for which

$$P_{ij} = \mathbb{P}\left[X(t+s) = j | X(s) = i\right] = 0$$

whenever $|i-j| > 1$.

Intensity Matrices

- Aside from transition probability matrices, information on CTMCs can also be stored in **transition rate matrices** or **intensity matrices**.
- The **intensity matrix** describes the rate at which a CTMC transitions or moves between states in its state space, thereby fully characterizing the process.
- The intensity matrix can also be used to determine the steady-state of the system.

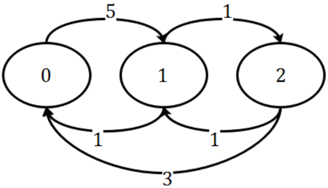
ATENEO

Intensity Matrices

Intensity matrix

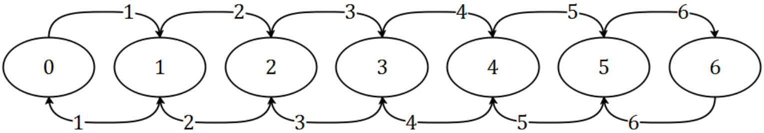
$$\mathbf{Q} = \begin{bmatrix} -5 & 5 & 0 \\ 1 & -2 & 1 \\ 3 & 1 & -4 \end{bmatrix}$$

Diagram



Intensity Matrix of a Birth-Death Process

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -3 & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & -5 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & -7 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & -9 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 & -11 & 6 \\ 0 & 0 & 0 & 0 & 0 & 6 & -6 \end{bmatrix}$$



ATENEO

Intensity Matrices: Issues on Size

- Calculations involving the intensity matrix may prove inefficient as the state space becomes larger.
- Instead of directly dealing with the intensity matrix, we may investigate its eigenvalues.
- However, larger orders of the intensity matrix implies higher degrees of the characteristic polynomial for the eigenvalues whose roots may not be easily determined.

ATENEO

Statement of the Problem

- Determine upper and lower bounds for the eigenvalues of the intensity matrix of a finite-state birth-death process using principal sub-matrices of the intensity matrix.
- Examine how the bounds can be used to numerically compute for the eigenvalues of the given intensity matrix.

ATENEO

Existing Techniques on Eigenvalue Bounding

Bounds via Matrix Elements

- A bound using matrix order and matrix elements by Hirsch; [17]
- The Gerschgorin circle theorem; [13]
- The ovals of Cassini by Brauer; and [2]
- Another inequality giving regions for the eigenvalues by Ostrowski. [25]

Bounds via Matrix Norms

- A relation between eigenvalues and powers of the matrix by Householder; [19]
- An upper bound for the modulus of the largest eigenvalue by the same author; [19]
- A more precise bound by Lorch; and [22]
- Matrix partition norm inequality by Feingold. [8]

ATENEO

Bounds through Related Matrices

- One through the complex conjugate by Hirsch; [17]
- A result by Bodewig; and [1]
- Several theorems by Wittmeyer. [32]

ATENEO

Interlacing Property of Eigenvalues

Interlacing Sequences

Definition (Interlacing Sequences)

Consider two sequences of real numbers: $\lambda_1 \geq \cdots \geq \lambda_n$ and $\mu_1 \geq \cdots \geq \mu_m$ with $m < n$. The second sequence is said to interlace the first whenever

$$\lambda_i \geq \mu_i \geq \lambda_{n-m+i}, \qquad i = 1, \ldots, m \tag{2}$$

If $m = n - 1$, the interlacing inequalities become

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_m \geq \lambda_n.$$

ATENEO

Interlacing Property for the FSBD Intensity Matrix

Let Q be a transition rate matrix for the seven-state birth-death process:

$$\mathbf{Q} = \begin{bmatrix} -a_0 & \lambda_0 & 0 & 0 & 0 & 0 & 0 \\ \mu_1 & -a_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_2 & -a_2 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \mu_3 & -a_3 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & \mu_4 & -a_4 & \lambda_4 & 0 \\ 0 & 0 & 0 & 0 & \mu_5 & -a_5 & \lambda_5 \\ 0 & 0 & 0 & 0 & 0 & \mu_6 & -a_6 \end{bmatrix} \tag{3}$$

List the eigenvalues of $\mathbf{Q}$ as $\mathcal{Q} = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6\}$, an increasing sequence with distinct and negative elements.

ATENEO

Interlacing Property for the FSBD Intensity Matrix

Consider the principal submatrices

$$\mathbf{P} = \begin{bmatrix} -a_0 & \lambda_0 & 0 \\ \mu_1 & -a_1 & \lambda_1 \\ 0 & \mu_2 & -a_2 \end{bmatrix} \quad \text{and} \quad \mathbf{R} = \begin{bmatrix} -a_4 & \lambda_4 & 0 \\ \mu_5 & -a_5 & \lambda_5 \\ 0 & \mu_6 & -a_6 \end{bmatrix}$$

obtained upon omission of row and column four in $\mathbf{Q}$. Correspondingly, list the eigenvalues of $\mathbf{P}$ and $\mathbf{R}$ as increasing sequences $\mathcal{P} = \{p_0, p_1, p_2\}$ and $\mathcal{R} = \{r_4, r_5, r_6\}$, respectively.

Interlacing Property for the FSBD Intensity Matrix

Proposition

Given the preceding set-up and notation, the possible outcomes are as follows:

1. *The sequence of eigenvalues* $\mathcal{Q} \setminus (\mathcal{P} \cap \mathcal{R})$ *from* $\mathbf{Q}$ *interlace with the sequence of eigenvalues* $\mathcal{P} \cup \mathcal{R}$ *from* $\mathbf{P}$ *and* $\mathbf{R}$.
2. *If* $\mathcal{P} \cap \mathcal{R}$ *is non-empty, then* $\mathcal{P} \cap \mathcal{R} \subseteq \mathcal{Q}$. *In other words, every eigenvalue of* $\mathbf{P}$ *that is equal to an eigenvalue of* $\mathbf{R}$ *is also an eigenvalue of* $\mathbf{Q}$.

ATENEO

Interlacing Property for the FSBD Intensity Matrix

For the first item, assuming all are distinct, for example, and listing both sets of eigenvalues from P and R as a single increasing sequence

$$\mathcal{U} = \{u_0, u_1, u_2, u_3, u_4, u_5\},$$

the interlacing property for the eigenvalues may be explicitly expressed as the inequality series,

$$q_0 \leq u_0 \leq q_1 \leq u_1 \leq q_2 \leq u_2 \leq q_3 \leq u_3 \leq q_4 \leq u_4 \leq q_5 \leq u_5 \leq q_6.$$

ATENEO

Outline of the Proof

1. Consider a set of probabilities across states
2. Write one as a sum of the convolutions of the others
3. Take the Laplace transform of the outcome
4. Rewrite some elements in the equation as entries of Q
5. Further simplify the expression through eigenvalues of Q
6. Acquire the eigenpolynomial of generator Q
7. Examine cases and summarize results

Proof of the Interlacing Property (Overview)

We consider the following probabilities:

$$y_{3,3}(t) = \mathbb{P}(\{Q(s) = 3 \ \forall \ s \in [0,t] \mid Q(0) = 3\} \tag{4}$$

$$\begin{aligned} y_{3,4}(t)dt = \mathbb{P}(\{Q(s) = 3 \ \forall \ s \in [0,t] \ \wedge Q(u) = 4 \\ \exists \ u \in (t, t+dt) \mid Q(0) = 3\}) \end{aligned} \tag{5}$$

$$\begin{aligned} y_{3,2}(t)dt = \mathbb{P}(\{Q(s) = 3 \ \forall \ s \in [0,t] \ \wedge Q(u) = 2 \\ \exists \ u \in (t, t+dt) \mid Q(0) = 3\}) \end{aligned} \tag{6}$$

$$\begin{aligned} y_{4}(t)dt = \mathbb{P}(\{Q(s) \geq 4 \ \forall \ s \in [0,t] \ \wedge Q(u) = 3 \\ \exists \ u \in (t, t+dt) \mid Q(0) = 4\}) \end{aligned} \tag{7}$$

$$\begin{aligned} y_{2}(t)dt = \mathbb{P}(\{Q(s) \leq 2 \ \forall \ s \in [0,t] \ \wedge Q(u) = 3 \\ \exists \ u \in (t, t+dt) \mid Q(0) = 2\}) \end{aligned} \tag{8}$$

ATENEO

Proof of the Interlacing Property (Overview)

We can therefore write $x_{3,3}(t)$ as

$$\begin{aligned} x_{3,3}(t) = y_{3,3}(t) &+ y_{3,4}(t) * y_4(t) * x_{3,3}(t) \\ &+ y_{3,2}(t) * y_2(t) * x_{3,3}(t) \end{aligned} \tag{9}$$

where $f(t) * g(t)$ represents the convolution of $f(t)$ and $g(t)$.

Taking the Laplace transform of both sides, we obtain

$$x^{\mathcal{L}}_{3,3}(s) = \frac{y^{\mathcal{L}}_{3,3}(s)}{1 - y^{\mathcal{L}}_{3,4}(s) \cdot y^{\mathcal{L}}_4(s) - y^{\mathcal{L}}_{3,2}(s) \cdot y^{\mathcal{L}}_2(s)} \tag{10}$$

where $h^{\mathcal{L}}(s)$ denotes the Laplace transform of $h(t)$.

Proof of the Interlacing Property (Overview)

From Keilson [20], we obtain the following representations:

$$y^{\mathcal{L}}_{3,3}(s) = \frac{1}{s + \lambda_3 + \mu_3} \qquad y^{\mathcal{L}}_{3,4}(s) = \frac{\lambda_3}{s + \lambda_3 + \mu_3}$$

$$y^{\mathcal{L}}_{3,2}(s) = \frac{\mu_3}{s + \lambda_3 + \mu_3}$$

and

$$y^{\mathcal{L}}_{4}(s) = -\sum_{i=4}^{6} \frac{c_i r_i}{s - r_i} \qquad y^{\mathcal{L}}_{2}(s) = -\sum_{i=0}^{2} \frac{c_i p_i}{s - p_i},$$

where $c_i \geq 0$ for all $i = 0, 1, \ldots, 6$, $c_0 + c_1 + c_2 \leq 1$, and $c_4 + c_5 + c_6 \leq 1$.

Proof of the Interlacing Property (Overview)

From the representations made prior, we can thus write

$$x_{3,3}^{\mathcal{L}}(s) = \frac{\delta(s)}{\epsilon(s)},$$

where

$$\begin{aligned} \delta(s) &= (s-p_0)(s-p_1)(s-p_2)(s-r_4)(s-r_5)(s-r_6) \\ \epsilon(s) &= \delta(s)(s+\lambda_3+\mu_3) + \mu_3 \sum_{i=0}^{2} \frac{c_i p_i \delta(s)}{s-p_i} + \lambda_3 \sum_{i=4}^{6} \frac{c_i r_i \delta(s)}{s-r_i}. \end{aligned} \tag{11}$$

According to [20], $\epsilon(s)$ is the polynomial whose roots are the eigenvalues of the transition rate matrix $\mathbf{Q}$.

ATENEO

Proof of the Interlacing Property (Overview)

CASE 1: If $p_i = r_j$ for some $i = 0, 1, 2$ and $j = 4, 5, 6$, $\epsilon(p_i) = \epsilon(r_j) = 0$, then $p_i = r_j$ is an eigenvalue of $\mathbf{Q}$.

- Write the equal eigenvalues as $q'_1, q'_2, \ldots, q'_k$. Let the excess eigenvalues from both $\mathbf{P}$ and $\mathbf{R}$ be $b_1, b_2, \ldots, b_{6-k}$, in increasing order.
- An alternative form for $\epsilon(s)$ is attained:

$$\epsilon(s) = (s - q'_1) \cdot (s - q'_2) \cdots (s - q'_{k-1}) \cdot (s - q'_k) \cdot \epsilon_E(s).$$

ATENEO

Proof of the Interlacing Property (Overview)

- Since all eigenvalues from P and R are negative and $c_i > 0$ for all i, $(-1)^i \cdot \epsilon_E(b_i) > 0$ for $i = 1, \ldots, 6 - k$.
- This means that the zeros of $\epsilon_E(s)$, the eigenvalues of $\mathbf{Q}$ that are not among $q'_1, \ldots, q'_k$, interlace with $b_1, b_2, \ldots, b_{k-1}$.

ATENEO

Proof of the Interlacing Property (Overview)

CASE 2: If all eigenvalues from P and R are distinct, then one has $b_1, b_2, b_3, b_4, b_5, b_6$, in sequence as well.

- Similar to what was justified earlier, $(-1)^i \cdot \epsilon(b_i) > 0$ for $i = 1, \ldots, 6$.
- This means that the roots of $\epsilon(s)$, the eigenvalues of $\mathbf{Q}$, interlace with $b_1, b_2, \ldots, b_6$.

ATENEO

Extensions

- The same result also holds for the intensity matrix $\mathbf{Q}$ of a finite-state birth-death process with $n+1$ states.
- One may also select any state q, $0 < q < n$, around which the principal submatrices $\mathbf{P}$ and $\mathbf{R}$ are formed.
- The proof of the extended case follows the same arguments as the proof presented prior.

ATENEO

Numerical Approximation of Eigenvalues Using the Interlacing Property—An Example

Example

We consider the transition rate matrix for the $M/M/1$ queue with $\lambda_0 = \mu_0 = 2$, signifying equal arrival and service rates:

$$\mathbf{Q} = \begin{bmatrix} -2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -4 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & -4 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -4 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -4 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -4 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & -4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}.$$

ATENEO

Example

We extract the 3×3 principal submatrix in the upper left corner and 5×5 principal submatrix in the lower right:

$$\mathbf{P} = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -4 & 2 \\ 0 & 2 & -4 \end{bmatrix} \quad \text{and} \quad \mathbf{R} = \begin{bmatrix} -4 & 2 & 0 & 0 & 0 \\ 2 & -4 & 2 & 0 & 0 \\ 0 & 2 & -4 & 2 & 0 \\ 0 & 0 & 2 & -4 & 2 \\ 0 & 0 & 0 & 2 & -2 \end{bmatrix}.$$

Outline of the Implementation

1. Implement a recursive strategy to bring forth the characteristic polynomial of the involved matrices;
2. Solve for the eigenvalues of the the chosen 3×3 principal submatrix, preferably analytically;
3. (Quasi-recursion) Further extract two 2×2 submatrices from the 5×5 principal submatrix and compute for the eigenvalues;
4. Numerically find the eigenvalues of the 5×5 principal submatrix based on the pooled eigenvalues of the 2×2 matrices; and
5. Follow a numerical method to find the eigenvalues of the 9×9 matrix through the completed set of bounds.

ATENEO

Numerically Solving the Characteristic Polynomial

- The characteristic polynomial of $\mathbf{Q}$ may be obtained following the routine of White [31] for tridiagonal matrices.
- In our study, we used the **bisection method** to numerically solve for the roots of the characteristic polynomial of $\mathbf{Q}$ as informed by the upper and lower bounds obtained from the interlacing property.

ATENEO

Concluding Remarks

Conclusion

- The proved interlacing property can be used to construct upper and lower bounds for each eigenvalue of a given intensity matrix.
- Given upper and lower bounds, numerical root-finding methods can then be performed more efficiently to approximate the actual eigenvalues of the intensity matrix.

ATENEO

Recommendations

Considerations for future work are:

1. To identify the pattern in which the pooled eigenvalues interlace
2. To attempt a different submatrix selection — consider the case of overlapping submatrices
3. To examine the implications of the theorem graphically
4. To further investigate how the eigenvalue bounds can be used for calculations using intensity matrices (i.e. finding transition probabilities, steady states, etc.)
5. To check for variations of the property for other Markov chains

ATENEO

Acknowledgements

The authors would like to thank Ateneo de Manila University **Department of Mathematics** and the **Vice President of the Loyola Schools** for their support and assistance in making this presentation possible.

ATENEO

References

References I

Bodewig, E. (1959).
Matrix Calculus.

Brauer, A. (1947).
Limits for the Characteristic Roots of a Matrix.

Brauer, A. (1961).
On the Characteristic Roots of Power-Positive Matrices.

Browne, E. (1939).
Limits to the Characteristic Roots of a Matrix.

Burden, R. and J. Faires (1985).
The Bisection Algorithm.

Crawford, F. (2012).
General Birth-Death Processes: Probabilities, Inference, and Applications.

References II

Diaconis, P. and D. Stroock (1991).
Geometric Bounds for Eigenvalues of Markov Chains.

Feingold, D. (1962).
Block Diagonally Dominant Matrices and Generalizations of the Gerschgorin Circle Theorem.

Frobenius, G. (1908).
Über Matrizen aus Positiven Elementen.

Frobenius, G. (1912).
Über Matrizen aus nicht Negativen Elementen.

Garren, K. (1968).
Bounds for the Eigenvalues of a Matrix.

Gaver, D., *et. al.* (1984).
Finite Birth-and-Death Models in Randomly Changing Environments.

ATENEO

References III

Gerschgorin, S. (1931).
Über die Abgrenzung der Eigenwerte einer Matrix.

Ghahramani, S. (2005).
Fundamentals of Probability with Stochastic Processes.

Godsil, C. (1993).
Algebraic Combinatorics.

Halmos, P. (1957).
Introduction to Hilbert Space and the Theory of Spectral Multiplicity.

Hirsch, M. (1902).
Sur les Racines d'une Équation Fondamentale.

Hohn, F. (1966).
Elementary Matrix Algebra.

ATENEO

References IV

- Householder, A. (1955).
  On the Convergence of Matrix Iterations.
- Keilson, J. (1979).
  Markov Chain Models — Rarity and Exponentiality.
- Latouche, G. and V. Ramaswami (1999).
  Introduction to Matrix Analytic Methods in Stochastic Modeling.
- Lorch, E. (1962).
  Spectral Theory.
- Marcus, M. (1960).
  Basic Theorems in Matrix Theory.
- Novozhilov, A., *et. al.* (2005).
  Biological Applications of the Theory of Birth-and-Death Processes.

ATENEO

References V

Ostrowski, A. (1951).
Ueber das Nichtverschwinden Einer Klasse von Determinanten und die Lokalisierung der Charakteristischen Wurzeln von Matrizen.

Parker, W. (1951).
Characteristic Roots and Field of Values of a Matrix.

Ross, S. (1996).
Stochastic Processes.

Sigman, K. (2009).
Continuous-Time Markov Chains.

Taylor, H. and S. Karlin (1998).
An Introduction to Stochastic Modeling.

Virtamo, J. (2005).
S-38.3143 Queueing Theory Lecture Notes.

ATENEO

References VI

White, P. (1958).
The Computation of Eigenvalues and Eigenvectors of a Matrix.

Wittmeyer, H. (1936).
Einfluss der Änderung Einer Matrix auf die Lösung des Zugehörigen Gleichungssystems, Sowie auf die Charakteristischen Zahlen und die Eigenvektoren.

ATENEO