Clustering in Markov chains with subdominant eigenvalues close to one

Jane Breen

Joint work with Emanuele Crisostomi, Mahsa Faizrahnemoon, Steve Kirkland, and Robert Shorten.

Dept. of Mathematics,
University of Manitoba

Special Western Canada Linear Algebra Meeting, BIRS 2017

Markov chains

- A discrete-time, time-homogeneous, finite Markov chain can be thought of as a system that undergoes transitions between states, over some finite state space $\{s_1, \ldots, s_n\}$, in discrete time steps.
- For each s_i, s_j, there is some *transition probability* t_{ij} denoting the probability of the system moving from state i to state j in a single time step.
- In this way, the chain is *memoryless*; the movement of the chain depends only on the current state the system is in.

Markov chains

- A discrete-time, time-homogeneous, finite Markov chain can be thought of as a system that undergoes transitions between states, over some finite state space $\{s_1, \ldots, s_n\}$, in discrete time steps.
- For each s_i, s_j, there is some *transition probability* t_{ij} denoting the probability of the system moving from state i to state j in a single time step.
- In this way, the chain is *memoryless*; the movement of the chain depends only on the current state the system is in.

Markov chains

- A discrete-time, time-homogeneous, finite Markov chain can be thought of as a system that undergoes transitions between states, over some finite state space $\{s_1, \ldots, s_n\}$, in discrete time steps.
- For each s_i, s_j, there is some *transition probability* t_{ij} denoting the probability of the system moving from state i to state j in a single time step.
- In this way, the chain is *memoryless*; the movement of the chain depends only on the current state the system is in.

Markov chains

- A Markov chain can be represented entirely by the *probability transition matrix* $T = [t_{ij}]$, a nonnegative row-stochastic matrix.
- Given an initial probability distribution vector $u_0^\top$, the probability distribution after k time-steps is given by $u_k^\top = u_0^\top T^k$.

Markov chains

- A Markov chain can be represented entirely by the *probability transition matrix* $T = [t_{ij}]$, a nonnegative row-stochastic matrix.
- Given an initial probability distribution vector $u_0^\top$, the probability distribution after k time-steps is given by $u_k^\top = u_0^\top T^k$.

Perron-Frobenius theorem

Theorem

Let T be a primitive row-stochastic matrix. Then:

(a) $\rho(T) = 1$, *and* 1 *is an eigenvalue of* T.

(b) *If* $\lambda \neq 1$ *is an eigenvalue of* T, *then* $|\lambda| < 1$.

(c) *There is a positive left eigenvector* $w^\top$ *of* T *corresponding to the eigenvalue* 1.

Stationary vector

Definition

Given that a Markov chain with transition matrix T is ergodic (that is, T is primitive), the left eigenvector $w^\top$ of T corresponding to the eigenvalue 1 is referred to as the **stationary vector** of the chain, and it catalogues the long-term behaviour of the chain.

Mean first passage times

Definition

For states i and j, the **mean first passage time from i to j**, denoted $m_{i,j}$, is the expected length of time it takes for the chain to reach state j for the first time, given that the chain starts in state i.

Road network model

Emanuele Crisostomi, Stephen Kirkland, and Robert Shorten.
A Google-like model of road network dynamics and its application to regulation and control.
International Journal of Control, 84(3):633–651, 2011.

Photo credit: Microsoft

Clustering in Markov chains

Clustering behaviour is usually characterised by the existence of collections of states of the Markov chain for which the system, if starting in a state in a cluster, is unlikely to leave that collection of states in the short term. Also, the expected number of time-steps until the chain is in a state outside of that cluster is relatively large.

Clustering in Markov chains

Example

$$
\begin{bmatrix}
0.0875 & 0.1158 & 0.2665 & 0.2820 & 0.2382 & 0 & 0.0059 & 0.0002 & 0.0022 & 0 & 0 & 0.0017 \\
0.2885 & 0.2870 & 0.1245 & 0 & 0.2900 & 0 & 0.0071 & 0 & 0 & 0.0010 & 0 & 0.0019 \\
0.0295 & 0.2473 & 0.3186 & 0.0610 & 0.3337 & 0.0021 & 0.0047 & 0 & 0.0018 & 0.0013 & 0.0001 & 0 \\
0.3650 & 0.2060 & 0.3579 & 0.0611 & 0 & 0 & 0.0012 & 0.0016 & 0 & 0.0030 & 0.0018 & 0.0024 \\
0.0681 & 0.2789 & 0.2432 & 0.3053 & 0.0946 & 0 & 0.0049 & 0 & 0.0051 & 0 & 0 & 0 \\
0.0046 & 0.0062 & 0.0086 & 0.0006 & 0 & 0 & 0.2443 & 0.0713 & 0.0666 & 0.2911 & 0.2355 & 0.0711 \\
0 & 0.0011 & 0.0052 & 0 & 0.0137 & 0.3158 & 0.1960 & 0.1266 & 0.1098 & 0.2318 & 0 & 0 \\
0.0018 & 0.0133 & 0 & 0.0011 & 0.0037 & 0.1890 & 0.1562 & 0.1396 & 0.1138 & 0.0634 & 0.1436 & 0.1744 \\
0.0024 & 0.0116 & 0.0060 & 0 & 0.0000 & 0.3355 & 0.1983 & 0 & 0.0853 & 0.0364 & 0.0730 & 0.2514 \\
0.0104 & 0 & 0.0061 & 0 & 0.0036 & 0.2141 & 0.1702 & 0.0960 & 0 & 0.1272 & 0.2196 & 0.1529 \\
0 & 0.0107 & 0 & 0.0046 & 0.0047 & 0.0639 & 0.0182 & 0.1455 & 0.2555 & 0.1268 & 0.1469 & 0.2232 \\
0.0200 & 0 & 0 & 0 & 0 & 0.1941 & 0.1876 & 0.0491 & 0.2231 & 0 & 0.1052 & 0.2210
\end{bmatrix}
$$

Clustering in Markov chains

Example

$$
\begin{bmatrix}
0.0875 & 0.1158 & 0.2665 & 0.2820 & 0.2382 & 0 & 0.0059 & 0.0002 & 0.0022 & 0 & 0 & 0.0017 \\
0.2885 & 0.2870 & 0.1245 & 0 & 0.2900 & 0 & 0.0071 & 0 & 0 & 0.0010 & 0 & 0.0019 \\
0.0295 & 0.2473 & 0.3186 & 0.0610 & 0.3337 & 0.0021 & 0.0047 & 0 & 0.0018 & 0.0013 & 0.0001 & 0 \\
0.3650 & 0.2060 & 0.3579 & 0.0611 & 0 & 0 & 0.0012 & 0.0016 & 0 & 0.0030 & 0.0018 & 0.0024 \\
0.0681 & 0.2789 & 0.2432 & 0.3053 & 0.0946 & 0 & 0.0049 & 0 & 0.0051 & 0 & 0 & 0 \\
0.0046 & 0.0062 & 0.0086 & 0.0006 & 0 & 0 & 0.2443 & 0.0713 & 0.0666 & 0.2911 & 0.2355 & 0.0711 \\
0 & 0.0011 & 0.0052 & 0 & 0.0137 & 0.3158 & 0.1960 & 0.1266 & 0.1098 & 0.2318 & 0 & 0 \\
0.0018 & 0.0133 & 0 & 0.0011 & 0.0037 & 0.1890 & 0.1562 & 0.1396 & 0.1138 & 0.0634 & 0.1436 & 0.1744 \\
0.0024 & 0.0116 & 0.0060 & 0 & 0.0000 & 0.3355 & 0.1983 & 0 & 0.0853 & 0.0364 & 0.0730 & 0.2514 \\
0.0104 & 0 & 0.0061 & 0 & 0.0036 & 0.2141 & 0.1702 & 0.0960 & 0 & 0.1272 & 0.2196 & 0.1529 \\
0 & 0.0107 & 0 & 0.0046 & 0.0047 & 0.0639 & 0.0182 & 0.1455 & 0.2555 & 0.1268 & 0.1469 & 0.2232 \\
0.0200 & 0 & 0 & 0 & 0 & 0.1941 & 0.1876 & 0.0491 & 0.2231 & 0 & 0.1052 & 0.2210
\end{bmatrix}
$$

Clustering in Markov chains

Example

$$
\begin{bmatrix}
9.5683 & 7.2856 & 5.9192 & 8.8386 & 6.1623 & 111.4516 & 108.3750 & 134.2927 & 119.9151 & 117.8051 & 123.5211 & 121.5733 \\
7.8461 & 6.2463 & 7.0903 & 11.4373 & 5.5713 & 111.4553 & 108.3307 & 134.3028 & 119.9611 & 117.7897 & 123.5357 & 121.5815 \\
10.4161 & 6.2682 & 5.8823 & 11.2735 & 5.2066 & 111.4228 & 108.3764 & 134.3090 & 119.9243 & 117.7744 & 123.5182 & 121.6233 \\
7.3576 & 6.8711 & 5.5085 & 11.2142 & 7.2498 & 111.4637 & 108.4423 & 134.2491 & 119.9635 & 117.7652 & 123.4718 & 121.5431 \\
9.4234 & 6.2587 & 6.1983 & 8.9797 & 6.9987 & 111.4499 & 108.3828 & 134.3059 & 119.8662 & 117.8042 & 123.5220 & 121.5991 \\
56.4328 & 54.8813 & 55.0335 & 59.7105 & 54.6607 & 16.2745 & 16.3806 & 35.2123 & 24.9775 & 17.7940 & 20.6192 & 23.8023 \\
56.4948 & 54.9064 & 55.0376 & 59.7083 & 54.5837 & 11.6313 & 16.3656 & 33.6425 & 25.2645 & 17.7810 & 25.2785 & 26.2237 \\
56.4292 & 54.8265 & 55.0941 & 59.7095 & 54.6532 & 13.5774 & 17.3639 & 33.5737 & 23.4582 & 22.5940 & 22.7664 & 21.5933 \\
56.4285 & 54.8526 & 55.0460 & 59.7299 & 54.6689 & 11.4165 & 16.2917 & 38.5797 & 24.4123 & 22.4132 & 24.2767 & 20.6083 \\
56.3789 & 54.9226 & 55.0433 & 59.6940 & 54.6523 & 13.5906 & 17.4180 & 34.5322 & 26.1350 & 20.9484 & 20.6825 & 22.1686 \\
56.4234 & 54.8407 & 55.0941 & 59.6727 & 54.6644 & 15.2224 & 19.7046 & 34.1128 & 19.7242 & 22.4636 & 22.7562 & 19.5144 \\
56.2799 & 54.9319 & 55.0728 & 59.6880 & 54.6831 & 13.1262 & 16.5668 & 37.2585 & 20.8087 & 23.8979 & 24.0661 & 20.6803
\end{bmatrix}
$$

Clustering in Markov chains

Example

9.5683	7.2856	5.9192	8.8386	6.1623	111.4516	108.3750	134.2927	119.9151	117.8051	123.5211	121.5733
7.8461	6.2463	7.0903	11.4373	5.5713	111.4553	108.3307	134.3028	119.9611	117.7897	123.5357	121.5815
10.4161	6.2682	5.8823	11.2735	5.2066	111.4228	108.3764	134.3090	119.9243	117.7744	123.5182	121.6233
7.3576	6.8711	5.5085	11.2142	7.2498	111.4637	108.4423	134.2491	119.9635	117.7652	123.4718	121.5431
9.4234	6.2587	6.1983	8.9797	6.9987	111.4499	108.3828	134.3059	119.8662	117.8042	123.5220	121.5991
56.4328	54.8813	55.0335	59.7105	54.6607	16.2745	16.3806	35.2123	24.9775	17.7940	20.6192	23.8023
56.4948	54.9064	55.0376	59.7083	54.5837	11.6313	16.3656	33.6425	25.2645	17.7810	25.2785	26.2237
56.4292	54.8265	55.0941	59.7095	54.6532	13.5774	17.3639	33.5737	23.4582	22.5940	22.7664	21.5933
56.4285	54.8526	55.0460	59.7299	54.6689	11.4165	16.2917	38.5797	24.4123	22.4132	24.2767	20.6083
56.3789	54.9226	55.0433	59.6940	54.6523	13.5906	17.4180	34.5322	26.1350	20.9484	20.6825	22.1686
56.4234	54.8407	55.0941	59.6727	54.6644	15.2224	19.7046	34.1128	19.7242	22.4636	22.7562	19.5144
56.2799	54.9319	55.0728	59.6880	54.6831	13.1262	16.5668	37.2585	20.8087	23.8979	24.0661	20.6803

Result

Theorem (Crisostomi, Kirkland, Shorten, 2011)

Let T be an irreducible stochastic matrix and suppose that $\lambda \in \mathbb{R}$ is an eigenvalue of T. Let $v = [v_1^\top \mid -v_2^\top | \mathbf{0}^\top]^\top$ be a corresponding λ-eigenvector (with $v_1 > 0$ and $v_2 > 0$) and let us partition the matrix T conformally as

$$\left[\begin{array}{c|c|c} T_{11} & T_{12} & T_{13} \\ \hline T_{21} & T_{22} & T_{23} \\ \hline T_{31} & T_{32} & T_{33} \end{array}\right]$$

and label the subsets of the partition as S_1, S_2, and S_0 respectively. Then:

(a) $\rho(T_{11}) > \lambda$ *and* $\rho(T_{22}) > \lambda$.

Result

Theorem(ctd)

(b) *There are subsets* $\tilde{S}_1 \subseteq S_1$, $\tilde{S}_2 \subseteq S_2$, *and positive vectors* $\tilde{w}_1^\top, \tilde{w}_2^\top$ *with supports on* $\tilde{S}_1, \tilde{S}_2$ *respectively such that* $\tilde{w}_1^\top \mathbb{1} = \tilde{w}_2^\top \mathbb{1} = 1$ *and*

$$\sum_{i \in \tilde{S}_1} \tilde{w}_1(i) \sum_{j \notin \tilde{S}_1} t_{ij} = 1 - \rho(T_{11}) \leq 1 - \lambda, \qquad (1)$$

and

$$\sum_{i \in \tilde{S}_2} \tilde{w}_2(i) \sum_{j \notin \tilde{S}_2} t_{ij} = 1 - \rho(T_{22}) \leq 1 - \lambda. \qquad (2)$$

Result

Theorem(ctd)

(c) *For any* $j \in \tilde{S}_2$,

$$\sum_{i \in \tilde{S}_1} \tilde{w}_1(i) m_{ij} \geq \frac{1}{1-\rho(T_{11})} \geq \frac{1}{1-\lambda} \tag{3}$$

and for any $j \in \tilde{S}_1$,

$$\sum_{i \in \tilde{S}_2} \tilde{w}_2(i) m_{ij} \geq \frac{1}{1-\rho(T_{22})} \geq \frac{1}{1-\lambda}, \tag{4}$$

where m_{ij} *are entries of the mean first passage matrix.*

Explanation of result

The existence of a real eigenvalue λ that is 'close' to 1 indicates the existence of clustering behaviour in the Markov chain as follows:

- There are two collections of states, S_1 and S_2, indexed by where the entries of an eigenvector corresponding to λ are positive and negative.
- Some weighted average of the transition probabilities from states in S_1 to states in S_2 (and vice versa) is bounded above by $1 - \lambda$ – i.e. the probability of transitioning from S_1 to S_2 is expected to be small if $\lambda \approx 1$.
- Some weighted average of the mean first passage times from states in S_1 to states in S_2 (and vice versa) is bounded below by $\frac{1}{1-\lambda}$ – i.e. MFP times from S_1 to S_2 are expected to be large if $\lambda \approx 1$.

Explanation of result

The existence of a real eigenvalue λ that is 'close' to 1 indicates the existence of clustering behaviour in the Markov chain as follows:

- There are two collections of states, S_1 and S_2, indexed by where the entries of an eigenvector corresponding to λ are positive and negative.
- Some weighted average of the transition probabilities from states in S_1 to states in S_2 (and vice versa) is bounded above by $1 - \lambda$ – i.e. the probability of transitioning from S_1 to S_2 is expected to be small if $\lambda \approx 1$.
- Some weighted average of the mean first passage times from states in S_1 to states in S_2 (and vice versa) is bounded below by $\frac{1}{1-\lambda}$ – i.e. MFP times from S_1 to S_2 are expected to be large if $\lambda \approx 1$.

Explanation of result

The existence of a real eigenvalue λ that is 'close' to 1 indicates the existence of clustering behaviour in the Markov chain as follows:

- There are two collections of states, S_1 and S_2, indexed by where the entries of an eigenvector corresponding to λ are positive and negative.
- Some weighted average of the transition probabilities from states in S_1 to states in S_2 (and vice versa) is bounded above by $1 - \lambda$ – i.e. the probability of transitioning from S_1 to S_2 is expected to be small if $\lambda \approx 1$.
- Some weighted average of the mean first passage times from states in S_1 to states in S_2 (and vice versa) is bounded below by $\frac{1}{1-\lambda}$ – i.e. MFP times from S_1 to S_2 are expected to be large if $\lambda \approx 1$.

Explanation of result

The existence of a real eigenvalue λ that is ‘close’ to 1 indicates the existence of clustering behaviour in the Markov chain as follows:

- There are two collections of states, S_1 and S_2, indexed by where the entries of an eigenvector corresponding to λ are positive and negative.
- Some weighted average of the transition probabilities from states in S_1 to states in S_2 (and vice versa) is bounded above by $1 - \lambda$ – i.e. the probability of transitioning from S_1 to S_2 is expected to be small if $\lambda \approx 1$.
- Some weighted average of the mean first passage times from states in S_1 to states in S_2 (and vice versa) is bounded below by $\frac{1}{1-\lambda}$ – i.e. MFP times from S_1 to S_2 are expected to be large if $\lambda \approx 1$.

Example

$$
\begin{bmatrix}
0.0875 & 0.1158 & 0.2665 & 0.2820 & 0.2382 & 0 & 0.0059 & 0.0002 & 0.0022 & 0 & 0 & 0.0017 \\
0.2885 & 0.2870 & 0.1245 & 0 & 0.2900 & 0 & 0.0071 & 0 & 0 & 0.0010 & 0 & 0.0019 \\
0.0295 & 0.2473 & 0.3186 & 0.0610 & 0.3337 & 0.0021 & 0.0047 & 0 & 0.0018 & 0.0013 & 0.0001 & 0 \\
0.3650 & 0.2060 & 0.3579 & 0.0611 & 0 & 0 & 0.0012 & 0.0016 & 0 & 0.0030 & 0.0018 & 0.0024 \\
0.0681 & 0.2789 & 0.2432 & 0.3053 & 0.0946 & 0 & 0.0049 & 0 & 0.0051 & 0 & 0 & 0 \\
0.0046 & 0.0062 & 0.0086 & 0.0006 & 0 & 0 & 0.2443 & 0.0713 & 0.0666 & 0.2911 & 0.2355 & 0.0711 \\
0 & 0.0011 & 0.0052 & 0 & 0.0137 & 0.3158 & 0.1960 & 0.1266 & 0.1098 & 0.2318 & 0 & 0 \\
0.0018 & 0.0133 & 0 & 0.0011 & 0.0037 & 0.1890 & 0.1562 & 0.1396 & 0.1138 & 0.0634 & 0.1436 & 0.1744 \\
0.0024 & 0.0116 & 0.0060 & 0 & 0.0000 & 0.3355 & 0.1983 & 0 & 0.0853 & 0.0364 & 0.0730 & 0.2514 \\
0.0104 & 0 & 0.0061 & 0 & 0.0036 & 0.2141 & 0.1702 & 0.0960 & 0 & 0.1272 & 0.2196 & 0.1529 \\
0 & 0.0107 & 0 & 0.0046 & 0.0047 & 0.0639 & 0.0182 & 0.1455 & 0.2555 & 0.1268 & 0.1469 & 0.2232 \\
0.0200 & 0 & 0 & 0 & 0 & 0.1941 & 0.1876 & 0.0491 & 0.2231 & 0 & 0.1052 & 0.2210
\end{bmatrix}
$$

- $\lambda = 0.97$.
- $v^\top = [\ 0.5 \quad 0.5 \quad 0.5 \quad 0.5 \quad 0.5 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1\ ]$

Example

$$
\begin{bmatrix}
0.0875 & 0.1158 & 0.2665 & 0.2820 & 0.2382 & 0 & 0.0059 & 0.0002 & 0.0022 & 0 & 0 & 0.0017 \\
0.2885 & 0.2870 & 0.1245 & 0 & 0.2900 & 0 & 0.0071 & 0 & 0 & 0.0010 & 0 & 0.0019 \\
0.0295 & 0.2473 & 0.3186 & 0.0610 & 0.3337 & 0.0021 & 0.0047 & 0 & 0.0018 & 0.0013 & 0.0001 & 0 \\
0.3650 & 0.2060 & 0.3579 & 0.0611 & 0 & 0 & 0.0012 & 0.0016 & 0 & 0.0030 & 0.0018 & 0.0024 \\
0.0681 & 0.2789 & 0.2432 & 0.3053 & 0.0946 & 0 & 0.0049 & 0 & 0.0051 & 0 & 0 & 0 \\
0.0046 & 0.0062 & 0.0086 & 0.0006 & 0 & 0 & 0.2443 & 0.0713 & 0.0666 & 0.2911 & 0.2355 & 0.0711 \\
0 & 0.0011 & 0.0052 & 0 & 0.0137 & 0.3158 & 0.1960 & 0.1266 & 0.1098 & 0.2318 & 0 & 0 \\
0.0018 & 0.0133 & 0 & 0.0011 & 0.0037 & 0.1890 & 0.1562 & 0.1396 & 0.1138 & 0.0634 & 0.1436 & 0.1744 \\
0.0024 & 0.0116 & 0.0060 & 0 & 0.0000 & 0.3355 & 0.1983 & 0 & 0.0853 & 0.0364 & 0.0730 & 0.2514 \\
0.0104 & 0 & 0.0061 & 0 & 0.0036 & 0.2141 & 0.1702 & 0.0960 & 0 & 0.1272 & 0.2196 & 0.1529 \\
0 & 0.0107 & 0 & 0.0046 & 0.0047 & 0.0639 & 0.0182 & 0.1455 & 0.2555 & 0.1268 & 0.1469 & 0.2232 \\
0.0200 & 0 & 0 & 0 & 0 & 0.1941 & 0.1876 & 0.0491 & 0.2231 & 0 & 0.1052 & 0.2210
\end{bmatrix}
$$

- $\lambda = 0.97$.
- $v^\top = [\ 0.5\ \ 0.5\ \ 0.5\ \ 0.5\ \ 0.5\ \ -1\ \ -1\ \ -1\ \ -1\ \ -1\ \ -1\ \ -1\ ]$

Question: complex eigenvalues?

Can any clustering behaviour be determined from a complex eigenvalue?

That is, given $\lambda \in \mathbb{C}$ an eigenvalue of T where $\lambda = \alpha + i\beta$, can we:

(a) define a conformal partition of a corresponding eigenvector for λ and the matrix T;

(b) determine lower bounds for the spectral radii of T_{11} and T_{22} (the principal submatrices determined by the index set of this partition); and

(c) conclude equivalent statements about the clustering properties of T as in parts (b) and (c) of the theorem.

A brief examination of the proof of the above theorem will determine that (b) and (c) are proven independent of the fact that λ is real; moreover, given lower bounds for $\rho(T_{11}), \rho(T_{22})$, these may be substituted for λ in (1), (2), (3) and (4).

Question: complex eigenvalues?

Can any clustering behaviour be determined from a complex eigenvalue?

That is, given $\lambda \in \mathbb{C}$ an eigenvalue of T where $\lambda = \alpha + i\beta$, can we:

(a) define a conformal partition of a corresponding eigenvector for λ and the matrix T;

(b) determine lower bounds for the spectral radii of T_{11} and T_{22} (the principal submatrices determined by the index set of this partition); and

(c) conclude equivalent statements about the clustering properties of T as in parts (b) and (c) of the theorem.

A brief examination of the proof of the above theorem will determine that (b) and (c) are proven independent of the fact that λ is real; moreover, given lower bounds for $\rho(T_{11}), \rho(T_{22})$, these may be substituted for λ in (1), (2), (3) and (4).

Question: complex eigenvalues?

Can any clustering behaviour be determined from a complex eigenvalue?

That is, given $\lambda \in \mathbb{C}$ an eigenvalue of T where $\lambda = \alpha + i\beta$, can we:

(a) define a conformal partition of a corresponding eigenvector for λ and the matrix T;

(b) determine lower bounds for the spectral radii of T_{11} and T_{22} (the principal submatrices determined by the index set of this partition); and

(c) conclude equivalent statements about the clustering properties of T as in parts (b) and (c) of the theorem.

A brief examination of the proof of the above theorem will determine that (b) and (c) are proven independent of the fact that λ is real; moreover, given lower bounds for $\rho(T_{11}), \rho(T_{22})$, these may be substituted for λ in (1), (2), (3) and (4).

Question: complex eigenvalues?

Can any clustering behaviour be determined from a complex eigenvalue?

That is, given $\lambda \in \mathbb{C}$ an eigenvalue of T where $\lambda = \alpha + i\beta$, can we:

(a) define a conformal partition of a corresponding eigenvector for λ and the matrix T;

(b) determine lower bounds for the spectral radii of T_{11} and T_{22} (the principal submatrices determined by the index set of this partition); and

(c) conclude equivalent statements about the clustering properties of T as in parts (b) and (c) of the theorem.

A brief examination of the proof of the above theorem will determine that (b) and (c) are proven independent of the fact that λ is real; moreover, given lower bounds for $\rho(T_{11}), \rho(T_{22})$, these may be substituted for λ in (1), (2), (3) and (4).

Question: complex eigenvalues?

Can any clustering behaviour be determined from a complex eigenvalue?

That is, given $\lambda \in \mathbb{C}$ an eigenvalue of T where $\lambda = \alpha + i\beta$, can we:

(a) define a conformal partition of a corresponding eigenvector for λ and the matrix T;

(b) determine lower bounds for the spectral radii of T_{11} and T_{22} (the principal submatrices determined by the index set of this partition); and

(c) conclude equivalent statements about the clustering properties of T as in parts (b) and (c) of the theorem.

A brief examination of the proof of the above theorem will determine that (b) and (c) are proven independent of the fact that λ is real; moreover, given lower bounds for $\rho(T_{11}), \rho(T_{22})$, these may be substituted for λ in (1), (2), (3) and (4).

A starting point

- Let T be an irreducible stochastic matrix with an eigenvalue $\lambda = \alpha + i\beta$.
- Let $x + iy$ be an eigenvector of T corresponding to the eigenvalue λ.
- Then since

$$T(x + iy) = (\alpha + i\beta)(x + iy)$$

by equating real and imaginary parts we have

$$Tx = \alpha x - \beta y$$

and

$$Ty = \beta x + \alpha y.$$

A starting point

- Let T be an irreducible stochastic matrix with an eigenvalue $\lambda = \alpha + i\beta$.
- Let $x + iy$ be an eigenvector of T corresponding to the eigenvalue λ.
- Then since

$$T(x + iy) = (\alpha + i\beta)(x + iy)$$

by equating real and imaginary parts we have

$$Tx = \alpha x - \beta y$$

and

$$Ty = \beta x + \alpha y.$$

A starting point

- Let T be an irreducible stochastic matrix with an eigenvalue $\lambda = \alpha + i\beta$.
- Let $x + iy$ be an eigenvector of T corresponding to the eigenvalue λ.
- Then since

$$T(x + iy) = (\alpha + i\beta)(x + iy)$$

by equating real and imaginary parts we have

$$Tx = \alpha x - \beta y$$

and

$$Ty = \beta x + \alpha y.$$

A starting point

- Let T be an irreducible stochastic matrix with an eigenvalue $\lambda = \alpha + i\beta$.
- Let $x + iy$ be an eigenvector of T corresponding to the eigenvalue λ.
- Then since

$$T(x + iy) = (\alpha + i\beta)(x + iy)$$

by equating real and imaginary parts we have

$$Tx = \alpha x - \beta y$$

and

$$Ty = \beta x + \alpha y.$$

- Partition the system (i.e. the matrix T and the vectors x and y) according to where x is positive, negative, and zero.
- Then we have

$$\left[\begin{array}{c|c|c} T_{11} & T_{12} & T_{13} \\ \hline T_{21} & T_{22} & T_{23} \\ \hline T_{31} & T_{32} & T_{33} \end{array}\right] \left[\begin{array}{c} x_1 \\ \hline x_2 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \alpha x_1 - \beta y_1 \\ \hline \alpha x_2 - \beta y_2 \\ \hline -\beta y_3 \end{array}\right]$$

  where $x_1 > 0$ and $x_2 < 0$, entrywise.
- Let S_1, S_2, S_3 denote the index sets of the partition. Note that S_3 may be empty.

- Partition the system (i.e. the matrix T and the vectors x and y) according to where x is positive, negative, and zero.
- Then we have

$$\left[\begin{array}{c|c|c} T_{11} & T_{12} & T_{13} \\ \hline T_{21} & T_{22} & T_{23} \\ \hline T_{31} & T_{32} & T_{33} \end{array}\right] \left[\begin{array}{c} x_1 \\ \hline x_2 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \alpha x_1 - \beta y_1 \\ \hline \alpha x_2 - \beta y_2 \\ \hline -\beta y_3 \end{array}\right]$$

  where $x_1 > 0$ and $x_2 < 0$, entrywise.
- Let S_1, S_2, S_3 denote the index sets of the partition. Note that S_3 may be empty.

This gives:

$$T_{11}x_1 + T_{12}x_2 = \alpha x_1 - \beta y_1,$$

and since $T_{12}x_2$ is entrywise nonpositive,

$$T_{11}x_1 \quad \geq \quad \alpha x_1 - \beta y_1.$$

It follows from this that

$$T_{11}\mathbb{1} \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right) \mathbb{1}.$$

By a well-known result we know that the spectral radius of a nonnegative matrix lies between its minimum and maximum row sums. It follows that

$$\rho(T_{11}) \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right)$$

This gives:

$$T_{11}x_1 + T_{12}x_2 = \alpha x_1 - \beta y_1,$$

and since $T_{12}x_2$ is entrywise nonpositive,

$$T_{11}x_1 \quad \geq \quad \alpha x_1 - \beta y_1.$$

It follows from this that

$$T_{11}\mathbb{1} \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right) \mathbb{1}.$$

By a well-known result we know that the spectral radius of a nonnegative matrix lies between its minimum and maximum row sums. It follows that

$$\rho(T_{11}) \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right)$$

This gives:

$$T_{11}x_1 + T_{12}x_2 = \alpha x_1 - \beta y_1,$$

and since $T_{12}x_2$ is entrywise nonpositive,

$$T_{11}x_1 \quad \geq \quad \alpha x_1 - \beta y_1.$$

It follows from this that

$$T_{11}\mathbb{1} \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right) \mathbb{1}.$$

By a well-known result we know that the spectral radius of a nonnegative matrix lies between its minimum and maximum row sums. It follows that

$$\rho(T_{11}) \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right)$$

This gives:

$$T_{11}x_1 + T_{12}x_2 = \alpha x_1 - \beta y_1,$$

and since $T_{12}x_2$ is entrywise nonpositive,

$$T_{11}x_1 \quad \geq \quad \alpha x_1 - \beta y_1.$$

It follows from this that

$$T_{11}\mathbb{1} \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right) \mathbb{1}.$$

By a well-known result we know that the spectral radius of a nonnegative matrix lies between its minimum and maximum row sums. It follows that

$$\rho(T_{11}) \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right)$$

This gives:

$$T_{11}x_1 + T_{12}x_2 = \alpha x_1 - \beta y_1,$$

and since $T_{12}x_2$ is entrywise nonpositive,

$$T_{11}x_1 \quad \geq \quad \alpha x_1 - \beta y_1.$$

It follows from this that

$$T_{11}\mathbb{1} \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right) \mathbb{1}.$$

By a well-known result we know that the spectral radius of a nonnegative matrix lies between its minimum and maximum row sums. It follows that

$$\rho(T_{11}) \quad \geq \quad \min_j \left(\frac{\alpha x_1(j) - \beta y_1(j)}{x_1(j)} \right)$$

or alternatively,

$$\rho(T_{11}) \quad \geq \quad \alpha - \beta \max_j \left(\frac{y_1(j)}{x_1(j)} \right).$$

Similarly, we may show that

$$\rho(T_{22}) \quad \geq \quad \alpha - \beta \max_j \left(\frac{y_2(j)}{x_2(j)} \right).$$

If $\alpha \approx 1$ and $\beta \approx 0$, then these lower bounds are close to 1, indicating clustering behaviour in the collections of states indexed by S_1 and by S_2.

or alternatively,

$$\rho(T_{11}) \quad \geq \quad \alpha - \beta \max_j \left(\frac{y_1(j)}{x_1(j)} \right).$$

Similarly, we may show that

$$\rho(T_{22}) \quad \geq \quad \alpha - \beta \max_j \left(\frac{y_2(j)}{x_2(j)} \right).$$

If $\alpha \approx 1$ and $\beta \approx 0$, then these lower bounds are close to 1, indicating clustering behaviour in the collections of states indexed by S_1 and by S_2.

or alternatively,

$$\rho(T_{11}) \quad \geq \quad \alpha - \beta \max_j \left(\frac{y_1(j)}{x_1(j)} \right).$$

Similarly, we may show that

$$\rho(T_{22}) \quad \geq \quad \alpha - \beta \max_j \left(\frac{y_2(j)}{x_2(j)} \right).$$

If $\alpha \approx 1$ and $\beta \approx 0$, then these lower bounds are close to 1, indicating clustering behaviour in the collections of states indexed by S_1 and by S_2.

Extra hypotheses

- Note that we need

$$\alpha x_1 - \beta y_1 > 0 \quad \text{and} \quad \alpha x_2 - \beta y_2 < 0$$

in order that these lower bounds are positive.

- We treat the above as hypotheses that must be satisfied in order to conclude anything about clustering behaviour of the chain.

Extra hypotheses

- Note that we need

$$\alpha x_1 - \beta y_1 > 0 \quad \text{and} \quad \alpha x_2 - \beta y_2 < 0$$

in order that these lower bounds are positive.

- We treat the above as hypotheses that must be satisfied in order to conclude anything about clustering behaviour of the chain.

Repartitioning

- Consider the vector $x + ty$, for some $t > 0$, and partition according to where $x + ty$ is positive, negative, or zero, with index sets $\tilde{S}_1, \tilde{S}_2, \tilde{S}_3$.
- Then we have

$$\left[\begin{array}{c|c|c} \widetilde{T}_{11} & \widetilde{T}_{12} & \widetilde{T}_{13} \\ \hline \widetilde{T}_{21} & \widetilde{T}_{22} & \widetilde{T}_{23} \\ \hline \widetilde{T}_{31} & \widetilde{T}_{32} & \widetilde{T}_{33} \end{array}\right] \left[\begin{array}{c} \tilde{x}_1 + t\tilde{y}_1 \\ \hline \tilde{x}_2 + t\tilde{y}_2 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \alpha(\tilde{x}_1 + t\tilde{y}_1) + \beta(t\tilde{x}_1 - \tilde{y}_1) \\ \hline \alpha(\tilde{x}_2 + t\tilde{y}_2) + \beta(t\tilde{x}_2 - \tilde{y}_2) \\ \hline (\alpha t - \beta)\tilde{y}_3 \end{array}\right].$$

- t must be bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left(\frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)}\right) \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left(\frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)}\right).$$

Repartitioning

- Consider the vector $x + ty$, for some $t > 0$, and partition according to where $x + ty$ is positive, negative, or zero, with index sets $\tilde{S}_1, \tilde{S}_2, \tilde{S}_3$.
- Then we have

$$\left[\begin{array}{c|c|c} \widetilde{T}_{11} & \widetilde{T}_{12} & \widetilde{T}_{13} \\ \hline \widetilde{T}_{21} & \widetilde{T}_{22} & \widetilde{T}_{23} \\ \hline \widetilde{T}_{31} & \widetilde{T}_{32} & \widetilde{T}_{33} \end{array}\right] \left[\begin{array}{c} \tilde{x}_1 + t\tilde{y}_1 \\ \hline \tilde{x}_2 + t\tilde{y}_2 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \alpha(\tilde{x}_1 + t\tilde{y}_1) + \beta(t\tilde{x}_1 - \tilde{y}_1) \\ \hline \alpha(\tilde{x}_2 + t\tilde{y}_2) + \beta(t\tilde{x}_2 - \tilde{y}_2) \\ \hline (\alpha t - \beta)\tilde{y}_3 \end{array}\right].$$

- t must be bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left(\frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)}\right) \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left(\frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)}\right).$$

Repartitioning

- Consider the vector $x + ty$, for some $t > 0$, and partition according to where $x + ty$ is positive, negative, or zero, with index sets $\tilde{S}_1, \tilde{S}_2, \tilde{S}_3$.
- Then we have

$$\left[\begin{array}{c|c|c} \widetilde{T}_{11} & \widetilde{T}_{12} & \widetilde{T}_{13} \\ \hline \widetilde{T}_{21} & \widetilde{T}_{22} & \widetilde{T}_{23} \\ \hline \widetilde{T}_{31} & \widetilde{T}_{32} & \widetilde{T}_{33} \end{array}\right] \left[\begin{array}{c} \tilde{x}_1 + t\tilde{y}_1 \\ \hline \tilde{x}_2 + t\tilde{y}_2 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \alpha(\tilde{x}_1 + t\tilde{y}_1) + \beta(t\tilde{x}_1 - \tilde{y}_1) \\ \hline \alpha(\tilde{x}_2 + t\tilde{y}_2) + \beta(t\tilde{x}_2 - \tilde{y}_2) \\ \hline (\alpha t - \beta)\tilde{y}_3 \end{array}\right].$$

- t must be bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left(\frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)}\right) \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left(\frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)}\right).$$

Repartitioning

- Note that this 'repartition' is not substantially different. We simply allow the option of including some extra states in the cluster by including indices *corresponding to positive entries of* y_3 to S_1, and indices *corresponding to negative entries of* y_3 to S_2.
- Moreover, $S_1 \subseteq \tilde{S}_1$ and $S_2 \subseteq \tilde{S}_2$.
- Interpreting in terms of prospective clustering behaviour, we are simply allowing the possible addition of more states into our existing index set to potentially achieve a 'tighter' cluster.

Repartitioning

- Note that this 'repartition' is not substantially different. We simply allow the option of including some extra states in the cluster by including indices *corresponding to positive entries of* y_3 to S_1, and indices *corresponding to negative entries of* y_3 to S_2.
- Moreover, $S_1 \subseteq \tilde{S}_1$ and $S_2 \subseteq \tilde{S}_2$.
- Interpreting in terms of prospective clustering behaviour, we are simply allowing the possible addition of more states into our existing index set to potentially achieve a 'tighter' cluster.

Repartitioning

- Note that this 'repartition' is not substantially different. We simply allow the option of including some extra states in the cluster by including indices *corresponding to positive entries of* y_3 to S_1, and indices *corresponding to negative entries of* y_3 to S_2.
- Moreover, $S_1 \subseteq \tilde{S}_1$ and $S_2 \subseteq \tilde{S}_2$.
- Interpreting in terms of prospective clustering behaviour, we are simply allowing the possible addition of more states into our existing index set to potentially achieve a 'tighter' cluster.

Repartitioning

- Proceeding as before, we have:

$$\tilde{T}_{11}(\tilde{x}_1 + t\tilde{y}_1) \geq \alpha(\tilde{x}_1 + t\tilde{y}_1) + \beta(t\tilde{x}_1 - \tilde{y}_1)$$

$$\Rightarrow \quad \rho(\tilde{T}_{11}) \geq \alpha + \beta \min_j \left(\frac{t\tilde{x}_1(j) - \tilde{y}_1(j)}{\tilde{x}_1(j) + t\tilde{y}_1(j)} \right),$$

- Similarly

$$\rho(\tilde{T}_{22}) \geq \alpha + \beta \min_j \left(\frac{t\tilde{x}_2(j) - \tilde{y}_2(j)}{\tilde{x}_2(j) + t\tilde{y}_2(j)} \right).$$

- These lower bounds are increasing functions of t, and so they are optimized by taking the limit as t approached the minimum of the previous bounds.

Repartitioning

- Proceeding as before, we have:

$$\begin{aligned}\tilde{T}_{11}(\tilde{x}_1 + t\tilde{y}_1) &\geq \alpha(\tilde{x}_1 + t\tilde{y}_1) + \beta(t\tilde{x}_1 - \tilde{y}_1)\\ \Rightarrow \quad \rho(\tilde{T}_{11}) &\geq \alpha + \beta \min_j \left(\frac{t\tilde{x}_1(j) - \tilde{y}_1(j)}{\tilde{x}_1(j) + t\tilde{y}_1(j)} \right),\end{aligned}$$

- Similarly

$$\rho(\tilde{T}_{22}) \geq \alpha + \beta \min_j \left(\frac{t\tilde{x}_2(j) - \tilde{y}_2(j)}{\tilde{x}_2(j) + t\tilde{y}_2(j)} \right).$$

- These lower bounds are increasing functions of t, and so they are optimized by taking the limit as t approached the minimum of the previous bounds.

Repartitioning

- Proceeding as before, we have:

$$\tilde{T}_{11}(\tilde{x}_1 + t\tilde{y}_1) \geq \alpha(\tilde{x}_1 + t\tilde{y}_1) + \beta(t\tilde{x}_1 - \tilde{y}_1)$$

$$\Rightarrow \quad \rho(\tilde{T}_{11}) \geq \alpha + \beta \min_j \left(\frac{t\tilde{x}_1(j) - \tilde{y}_1(j)}{\tilde{x}_1(j) + t\tilde{y}_1(j)} \right),$$

- Similarly

$$\rho(\tilde{T}_{22}) \geq \alpha + \beta \min_j \left(\frac{t\tilde{x}_2(j) - \tilde{y}_2(j)}{\tilde{x}_2(j) + t\tilde{y}_2(j)} \right).$$

- These lower bounds are increasing functions of t, and so they are optimized by taking the limit as t approached the minimum of the previous bounds.

Repartitioning - again!

- Now consider $x + ty$, where t is negative, and partition according to where $x + ty$ is positive, negative and zero, denoting these new index sets $\overline{S}_1, \overline{S}_2, \overline{S}_3$.
- Then there is the possibility of including in the index set S_1 (respectively, S_2) some nodes corresponding to entries of y_3 which are positive (respectively, negative), producing a different partition than before (possibly).
- Since we observed that the expression for the lower bounds were increasing in t, and t is negative, we choose $t \to 0$ to optimise these lower bounds for the spectral radii. This means that we achieve the same lower bounds as in the most basic case.

Repartitioning - again!

- Now consider $x + ty$, where t is negative, and partition according to where $x + ty$ is positive, negative and zero, denoting these new index sets $\overline{S}_1, \overline{S}_2, \overline{S}_3$.
- Then there is the possibility of including in the index set S_1 (respectively, S_2) some nodes corresponding to entries of y_3 which are positive (respectively, negative), producing a different partition than before (possibly).
- Since we observed that the expression for the lower bounds were increasing in t, and t is negative, we choose $t \to 0$ to optimise these lower bounds for the spectral radii. This means that we achieve the same lower bounds as in the most basic case.

Repartitioning - again!

- Now consider $x + ty$, where t is negative, and partition according to where $x + ty$ is positive, negative and zero, denoting these new index sets $\overline{S}_1, \overline{S}_2, \overline{S}_3$.
- Then there is the possibility of including in the index set S_1 (respectively, S_2) some nodes corresponding to entries of y_3 which are positive (respectively, negative), producing a different partition than before (possibly).
- Since we observed that the expression for the lower bounds were increasing in t, and t is negative, we choose $t \to 0$ to optimise these lower bounds for the spectral radii. This means that we achieve the same lower bounds as in the most basic case.

Theorem

Theorem

Let T be an $n \times n$ irreducible and stochastic matrix, let $\lambda = \alpha + i\beta$ be an eigenvalue of T, with $\alpha, \beta > 0$, and let $x + iy$ be a right eigenvector of T corresponding to λ. For i = 1, 2, 3, let S_i, $\widetilde{S}_i$, and $\overline{S}_i$ be the index sets obtained from the partitions described above, let $x_i, y_i, \tilde{x}_i, \tilde{y}_i, \bar{x}_i, \bar{y}_i$ be the subvectors of x and y corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$, and let T_{ii}, $\widetilde{T}_{ii}$ and $\overline{T}_{ii}$ be the principal submatrices of T corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$. Then:

Theorem

Theorem

Let T be an $n \times n$ irreducible and stochastic matrix, let $\lambda = \alpha + i\beta$ be an eigenvalue of T, with $\alpha, \beta > 0$, and let $x + iy$ be a right eigenvector of T corresponding to λ. For i = 1, 2, 3, let S_i, $\widetilde{S}_i$, and $\overline{S}_i$ be the index sets obtained from the partitions described above, let $x_i, y_i, \tilde{x}_i, \tilde{y}_i, \bar{x}_i, \bar{y}_i$ be the subvectors of x and y corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$, and let T_{ii}, $\widetilde{T}_{ii}$ and $\overline{T}_{ii}$ be the principal submatrices of T corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$. Then:

Theorem(ctd)

(a) If $\alpha x_1 - \beta y_1 > 0$,

$$\rho(T_{11}) \geq \alpha - \beta \cdot \max_j \left\{ \frac{y_1(j)}{x_1(j)} \right\}.$$

(b) If $\alpha x_2 - \beta y_2 < 0$,

$$\rho(T_{22}) \geq \alpha - \beta \cdot \max_j \left\{ \frac{y_2(j)}{x_2(j)} \right\}.$$

Theorem(ctd)

(c) If $\alpha\tilde{x}_1 - \beta\tilde{y}_1 > 0$,

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{t\tilde{x}_1(j) - \tilde{y}_1(j)}{\tilde{x}_1(j) + t\tilde{y}_1(j)} \right\},$$

where $t > 0$ and is bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left\{ \frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\} \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left\{ \frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)} \right\}.$$

If $\tilde{y}_1 > 0$ and $\tilde{y}_2 < 0$, then

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\}.$$

Theorem(ctd)

(c) If $\alpha\tilde{x}_1 - \beta\tilde{y}_1 > 0$,

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{t\tilde{x}_1(j) - \tilde{y}_1(j)}{\tilde{x}_1(j) + t\tilde{y}_1(j)} \right\},$$

where $t > 0$ and is bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left\{ \frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\} \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left\{ \frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)} \right\}.$$

If $\tilde{y}_1 > 0$ and $\tilde{y}_2 < 0$, then

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\}.$$

Theorem(ctd)

(c) If $\alpha\tilde{x}_1 - \beta\tilde{y}_1 > 0$,

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{t\tilde{x}_1(j) - \tilde{y}_1(j)}{\tilde{x}_1(j) + t\tilde{y}_1(j)} \right\},$$

where $t > 0$ and is bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left\{ \frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\} \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left\{ \frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)} \right\}.$$

If $\tilde{y}_1 > 0$ and $\tilde{y}_2 < 0$, then

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\}.$$

Theorem(ctd)

(d) If $\alpha\tilde{x}_2 - \beta\tilde{y}_2 < 0$,

$$\rho(\widetilde{T}_{22}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{t\tilde{x}_2(j) - \tilde{y}_2(j)}{\tilde{x}_2(j) + t\tilde{y}_2(j)} \right\},$$

where $t > 0$ and is bounded above by

$$\min_{\substack{j \\ \tilde{y}_1(j)<0}} \left\{ \frac{-\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\} \quad \text{and} \quad \min_{\substack{j \\ \tilde{y}_2(j)>0}} \left\{ \frac{-\tilde{x}_2(j)}{\tilde{y}_2(j)} \right\}.$$

If $\tilde{y}_1 > 0$ and $\tilde{y}_2 < 0$, then

$$\rho(\widetilde{T}_{22}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\tilde{x}_2(j)}{\tilde{y}_2(j)} \right\}.$$

Theorem(ctd)

(e) If $\alpha\bar{x}_1 - \beta\bar{y}_1 > 0$,

$$\rho(\overline{T}_{11}) \geq \alpha - \beta \cdot \min_j \left\{ \frac{\bar{y}_1(j)}{\bar{x}_1(j)} \right\}.$$

(f) If $\alpha\bar{x}_2 - \beta\bar{y}_2 > 0$,

$$\rho(\overline{T}_{22}) \geq \alpha - \beta \cdot \min_j \left\{ \frac{\bar{y}_2(j)}{\bar{x}_2(j)} \right\}.$$

Using the imaginary part of the eigenvector

- Partition the system (i.e. the matrix T and the vectors x and y) according to where y is positive, negative, and zero.
- That is, we have

$$\left[\begin{array}{c|c|c} T_{11} & T_{12} & T_{13} \\ \hline T_{21} & T_{22} & T_{23} \\ \hline T_{31} & T_{32} & T_{33} \end{array}\right] \left[\begin{array}{c} y_1 \\ \hline y_2 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} \beta x_1 + \alpha y_1 \\ \hline \beta x_2 + \alpha y_2 \\ \hline \beta x_3 \end{array}\right]$$

Another theorem

Theorem II

Let T be an $n \times n$ irreducible and stochastic matrix, let $\lambda = \alpha + i\beta$ be an eigenvalue of T, with $\alpha, \beta > 0$, and let $x + iy$ be a right eigenvector of T corresponding to λ. For i = 1, 2, 3, let S_i denote the index sets obtained by partitioning according to where y is positive, negative and zero. Also, let $\widetilde{S}_i$ (respectively, $\overline{S}_i$) be the index sets obtained by partitioning according to where $sx + y$ is positive, negative, and zero, where s is positive (respectively, where s is negative). Let $x_i, y_i, \tilde{x}_i, \tilde{y}_i, \bar{x}_i, \bar{y}_i$ be the subvectors of x and y corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$, and let T_{ii}, $\widetilde{T}_{ii}$ and $\overline{T}_{ii}$ be the principal submatrices of T corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$. Assume that x_i and y_i (resp., $\tilde{x}_i$ and $\tilde{y}_i$, $\bar{x}_i$ and $\bar{y}_i$) are linearly independent, $i = 1, 2$. Then:

Another theorem

Theorem II

Let T be an $n \times n$ irreducible and stochastic matrix, let $\lambda = \alpha + i\beta$ be an eigenvalue of T, with $\alpha, \beta > 0$, and let $x + iy$ be a right eigenvector of T corresponding to λ. For i = 1, 2, 3, let S_i denote the index sets obtained by partitioning according to where y is positive, negative and zero. Also, let $\widetilde{S}_i$ (respectively, $\overline{S}_i$) be the index sets obtained by partitioning according to where $sx + y$ is positive, negative, and zero, where s is positive (respectively, where s is negative). Let $x_i, y_i, \tilde{x}_i, \tilde{y}_i, \bar{x}_i, \bar{y}_i$ be the subvectors of x and y corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$, and let T_{ii}, $\widetilde{T}_{ii}$ and $\overline{T}_{ii}$ be the principal submatrices of T corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$. Assume that x_i and y_i (resp., $\tilde{x}_i$ and $\tilde{y}_i$, $\bar{x}_i$ and $\bar{y}_i$) are linearly independent, $i = 1, 2$. Then:

Another theorem

Theorem II

Let T be an $n \times n$ irreducible and stochastic matrix, let $\lambda = \alpha + i\beta$ be an eigenvalue of T, with $\alpha, \beta > 0$, and let $x + iy$ be a right eigenvector of T corresponding to λ. For i = 1, 2, 3, let S_i denote the index sets obtained by partitioning according to where y is positive, negative and zero. Also, let $\widetilde{S}_i$ (respectively, $\overline{S}_i$) be the index sets obtained by partitioning according to where $sx + y$ is positive, negative, and zero, where s is positive (respectively, where s is negative). Let $x_i, y_i, \tilde{x}_i, \tilde{y}_i, \bar{x}_i, \bar{y}_i$ be the subvectors of x and y corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$, and let T_{ii}, $\widetilde{T}_{ii}$ and $\overline{T}_{ii}$ be the principal submatrices of T corresponding to the index sets S_i, $\widetilde{S}_i$, and $\overline{S}_i$. Assume that x_i and y_i (resp., $\tilde{x}_i$ and $\tilde{y}_i$, $\bar{x}_i$ and $\bar{y}_i$) are linearly independent, $i = 1, 2$. Then:

Theorem II (ctd)

(a) If $\alpha y_1 + \beta x_1 > 0$,

$$\rho(T_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{x_1(j)}{y_1(j)} \right\}.$$

(b) If $\alpha y_2 + \beta x_2 < 0$,

$$\rho(T_{22}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{x_2(j)}{y_2(j)} \right\}.$$

Theorem II (ctd)

(c) If $\alpha\tilde{y}_1 + \beta\tilde{x}_1 > 0$,

$$\rho(\widetilde{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\tilde{x}_1(j)}{\tilde{y}_1(j)} \right\}.$$

(d) If $\alpha\tilde{y}_2 + \beta\tilde{x}_2 > 0$,

$$\rho(\widetilde{T}_{22}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\tilde{x}_2(j)}{\tilde{y}_2(j)} \right\}.$$

Theorem II (ctd)

(e) If $\alpha\bar{y}_1 + \beta\bar{x}_1 > 0$,

$$\rho(\overline{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\bar{x}_1(j) - s\bar{y}_1(j)}{s\bar{x}_1(j) + \bar{y}_1(j)} \right\},$$

where $s < 0$ and is bounded below by

$$\min_{\substack{j \\ \tilde{x}_1(j)>0}} \left\{ \frac{-\tilde{y}_1(j)}{\tilde{x}_1(j)} \right\} \quad \text{and by} \quad \min_{\substack{j \\ \tilde{x}_2(j)<0}} \left\{ \frac{-\tilde{y}_2(j)}{\tilde{x}_2(j)} \right\}.$$

If $\bar{x}_1 < 0$ and $\bar{x}_2 > 0$, then

$$\rho(\overline{T}_{11}) \geq \alpha - \beta \cdot \min_j \left\{ \frac{\bar{y}_1(j)}{\bar{x}_1(j)} \right\}.$$

Theorem II (ctd)

(e) If $\alpha\bar{y}_1 + \beta\bar{x}_1 > 0$,

$$\rho(\overline{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\bar{x}_1(j) - s\bar{y}_1(j)}{s\bar{x}_1(j) + \bar{y}_1(j)} \right\},$$

where $s < 0$ and is bounded below by

$$\min_{\substack{j \\ \tilde{x}_1(j)>0}} \left\{ \frac{-\tilde{y}_1(j)}{\tilde{x}_1(j)} \right\} \quad \text{and by} \quad \min_{\substack{j \\ \tilde{x}_2(j)<0}} \left\{ \frac{-\tilde{y}_2(j)}{\tilde{x}_2(j)} \right\}.$$

If $\bar{x}_1 < 0$ and $\bar{x}_2 > 0$, then

$$\rho(\overline{T}_{11}) \geq \alpha - \beta \cdot \min_j \left\{ \frac{\bar{y}_1(j)}{\bar{x}_1(j)} \right\}.$$

Theorem II (ctd)

(e) If $\alpha\bar{y}_1 + \beta\bar{x}_1 > 0$,

$$\rho(\overline{T}_{11}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\bar{x}_1(j) - s\bar{y}_1(j)}{s\bar{x}_1(j) + \bar{y}_1(j)} \right\},$$

where $s < 0$ and is bounded below by

$$\min_{\substack{j \\ \tilde{x}_1(j)>0}} \left\{ \frac{-\tilde{y}_1(j)}{\tilde{x}_1(j)} \right\} \quad \text{and by} \quad \min_{\substack{j \\ \tilde{x}_2(j)<0}} \left\{ \frac{-\tilde{y}_2(j)}{\tilde{x}_2(j)} \right\}.$$

If $\bar{x}_1 < 0$ and $\bar{x}_2 > 0$, then

$$\rho(\overline{T}_{11}) \geq \alpha - \beta \cdot \min_j \left\{ \frac{\bar{y}_1(j)}{\bar{x}_1(j)} \right\}.$$

Theorem II (ctd)

(f) If $\alpha \bar{y}_2 + \beta \bar{x}_2 < 0$,

$$\rho(\overline{T}_{22}) \geq \alpha + \beta \cdot \min_j \left\{ \frac{\bar{x}_2(j) - s\bar{y}_2(j)}{s\bar{x}_2(j) + \bar{y}_2(j)} \right\},$$

where $s < 0$ and is bounded below by

$$\min_{\substack{j \\ \tilde{x}_1(j)>0}} \left\{ \frac{-\tilde{y}_1(j)}{\tilde{x}_1(j)} \right\} \quad \text{and by} \quad \min_{\substack{j \\ \tilde{x}_2(j)<0}} \left\{ \frac{-\tilde{y}_2(j)}{\tilde{x}_2(j)} \right\}.$$

If $\bar{x}_1 < 0$ and $\bar{x}_2 > 0$, then

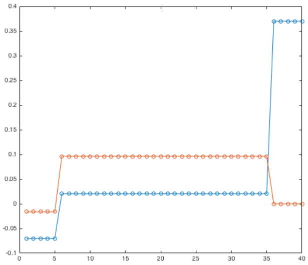
$$\rho(\overline{T}_{22}) \geq \alpha - \beta \cdot \min_j \left\{ \frac{\bar{y}_2(j)}{\bar{x}_2(j)} \right\}.$$

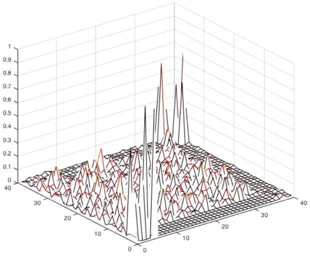
Example

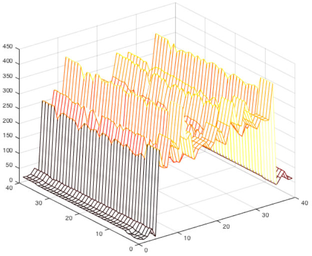
- T a 40×40 irreducible stochastic matrix, with an eigenvalue $\lambda = 0.8188 + 0.0348i$.

Partition wrt:	S_1	LB on $\rho(T_{11})$	S_2	LB on $\rho(T_{22})$
x	6–40	0.6539	1–5	0.8108
$x + ty$, $t > 0$	6–40	0.8115	1–5	0.8027
$x + ty$, $t < 0$	6–40	0.6539	1–5	0.8108
y	1–5	0.9698	6–35	0.8262
$sx + y$, $s > 0$	1–5, 36–40	0.9698	6–35	0.8262
$sx + y$, $s < 0$	1–5	0.8108	6–40	0.8188

Example


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
0
5
10
15
20
25
30
35
40


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
40
30
20
10
0
0
10
20
30
40


450
400
350
300
250
200
150
100
50
0
40
30
20
10
0
0
10
20
30
40

Thank you!