Perturbing Doubly Stochastic Matrices

Philip A. Knight[1]   Daniel Ruiz[2]

[1]Department of Mathematics and Statistics
University of Strathclyde

[2]ENSEEIHT

May 2019

Scaling to doubly stochastic form

- $\mathbf{A} \geq \mathbf{0}$.
- Seek doubly stochastic matrix $\mathbf{P} = \mathbf{DAF}$ where $\mathbf{D}$ and $\mathbf{F}$ are diagonal matrices.
- Unique scaling (up to scalar multiple) if $\mathbf{A}$ is **fully indecomposable** (FI).
- Multiple scalings exist if $\mathbf{A}$ has **total support**.
- Use scaling as a tool to discover hidden block structure.

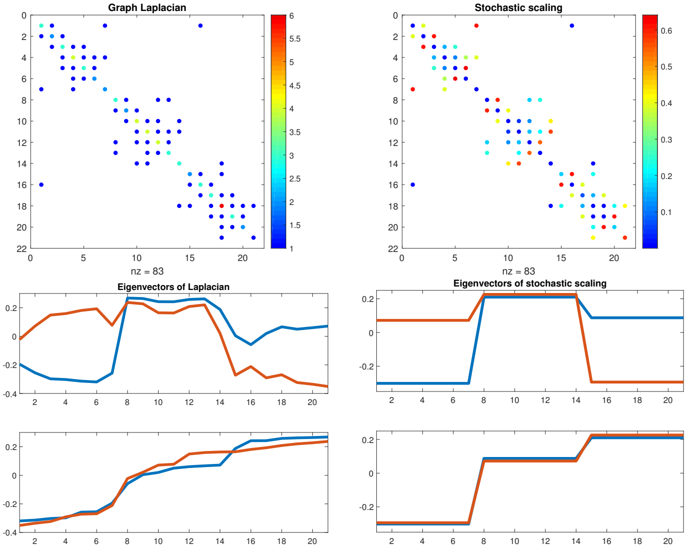
Spectral Properties of Doubly Stochastic Matrices

- Suppose that $\mathbf{S} \in \mathbb{R}^{n \times n}$ is FI, and doubly stochastic.
- Principal singular value of $\mathbf{S}$ is $\mathbf{1}$, with multiplicity $\mathbf{1}$.
- Associated singular vector is equal to a multiple of $\mathbf{e}$.
- Suppose $\mathbf{S}$ is a permutation of a block diagonal structure with k blocks.
- Singular value $\mathbf{1}$ has multiplicity k.
- Partition $\mathbf{S}$ according to the blocks and partitioned components of singular vectors are constant.

Simple Block Detection Scheme

- Compute a principal singular vector $\mathbf{x}$.
- $\mathbf{x} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k$.
- Characterise the partitions exactly using the set $\{a_1, \ldots, a_k\}$.
- If the matrix has a structure that is close to block diagonal, then we hope the leading singular vectors have a similar structure.
- Look for steps in the computed vectors to reveal an underlying near block structure.
- Akin to Fiedler vectors.

Example



Graph Laplacian
nz = 83
Stochastic scaling
nz = 83
Eigenvectors of Laplacian
Eigenvectors of stochastic scaling

Perturbation theory

- Let $\mathbf{P} = \mathbf{RAC}$.
- Consider a small perturbation to $\mathbf{A}$, namely $\widetilde{\mathbf{A}} = \mathbf{A} + \epsilon\mathbf{H}$.
- How close is the scaling of $\widetilde{\mathbf{A}}$ to $\mathbf{P}$?
- We'd like to assume that $\mathbf{A}$ is block diagonal matrix.
- Start by assuming that $\mathbf{A}$ is FI.

Theorem

- $\mathbf{A}, \widetilde{\mathbf{A}}, \mathbf{R}, \mathbf{C}, \mathbf{H}$ and $\mathbf{P}$ as described above.
- $\widetilde{\mathbf{P}} = \mathbf{P} + \epsilon\mathbf{Q}$ where $\mathbf{Q} = \mathbf{RHC}$.
- There exist vectors $\mathbf{f}$ and $\mathbf{g}$ such that $\|\mathbf{f}\|, \|\mathbf{g}\| = O(\epsilon)$.
- If $\mathbf{F} = \mathcal{D}(\mathbf{e} + \mathbf{f})$, $\mathbf{G} = \mathcal{D}(\mathbf{e} + \mathbf{g})$, then

$$\begin{bmatrix} \mathbf{F}\widetilde{\mathbf{P}}\mathbf{G}\mathbf{e} \\ \mathbf{G}\widetilde{\mathbf{P}}^T\mathbf{F}\mathbf{e} \end{bmatrix} = \mathbf{e} + \mathbf{m}.$$

- $\|\mathbf{m}\| = O(\epsilon^2)$.

Proof

- Key is to find $\mathbf{f}$ and $\mathbf{g}$ so that

$$\mathbf{P}_{sym}\begin{bmatrix}\mathbf{f}\\ \mathbf{g}\end{bmatrix} = \begin{bmatrix}\mathbf{I} & \mathbf{P}\\ \mathbf{P}^T & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{f}\\ \mathbf{g}\end{bmatrix} = -\epsilon\begin{bmatrix}\mathbf{Qe}\\ \mathbf{Q}^T\mathbf{e}\end{bmatrix}.$$

- Not automatic as we know that $\mathbf{P}_{sym}$ is singular.
- Since $\mathbf{A}$ is FI, kernel is 1D with basis $\begin{bmatrix}\mathbf{e}\\ -\mathbf{e}\end{bmatrix}$.
- $\begin{bmatrix}\mathbf{e}\\ -\mathbf{e}\end{bmatrix}^T\begin{bmatrix}\mathbf{Qe}\\ \mathbf{Q}^T\mathbf{e}\end{bmatrix} = \mathbf{e}^T\mathbf{Qe} - \mathbf{e}^T\mathbf{Q}^T\mathbf{e} = \mathbf{0}$.
- Choice of $\mathbf{f}$ and $\mathbf{g}$ is motivated by the result of applying Newton method to $\widetilde{\mathbf{P}}$ with initial vector $\mathbf{e}$.

Is $\mathbf{P}$ close to $\widetilde{\mathbf{P}}$?

- Use theorem to give precise conditions under which $\mathbf{A}$ and $\widetilde{\mathbf{A}}$ have nearby doubly stochastic scalings.
- $\|\mathbf{F}\widetilde{\mathbf{P}}\mathbf{G} - \mathbf{P}\|$ is small when $\|\mathbf{f}\|$ and $\|\mathbf{g}\|$ are small.
- This is true if $\|\mathbf{Q}\|$ is constrained and the smallest nonzero eigenvalue of $\mathbf{P}_{sym}$ is bounded away from zero.
- These are reasonable expectations in the context of uncovering block structure.

Extending The Result

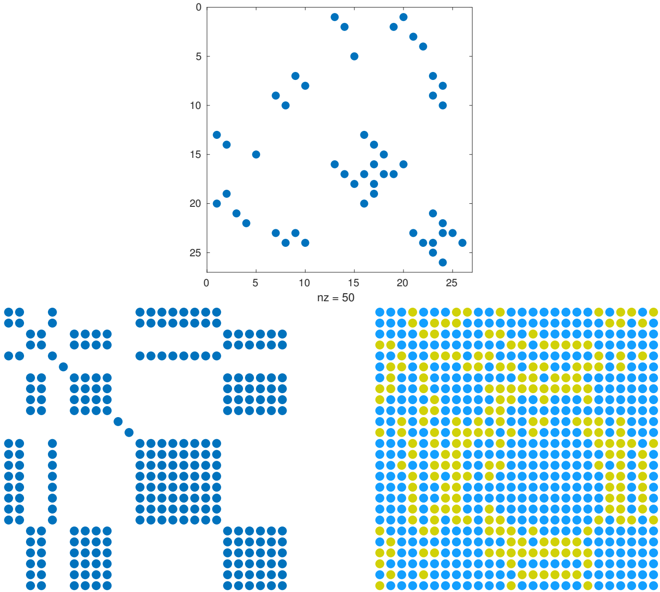
- We need to extend our analysis to cover the case when $\mathbf{A}$ is not FI but has total support.
- This occurs observe when $\mathbf{A}$ has perfect block structure.
- Unit singular value has multiplicity equal to number of blocks.
- Extra analysis needed to show system involving $\mathbf{P}_{sym}$ is consistent.

Two Blocks

- $\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_2 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix}.$
- Kernel spanned by $\mathbf{e}_x = \begin{bmatrix} \mathbf{e}_1 \\ 0 \\ -\mathbf{e}_1 \\ 0 \end{bmatrix}, \ \mathbf{e}_y = \begin{bmatrix} 0 \\ \mathbf{e}_2 \\ 0 \\ -\mathbf{e}_2 \end{bmatrix}.$
- $\begin{bmatrix} \mathbf{e}_x^T \\ \mathbf{e}_y^T \end{bmatrix} \begin{bmatrix} \mathbf{Q}\mathbf{e} \\ \mathbf{Q}^T\mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1^T\mathbf{Q}_{12}\mathbf{e}_2 - \mathbf{e}_1^T\mathbf{Q}_{21}^T\mathbf{e}_2 \\ \mathbf{e}_2^T\mathbf{Q}_{21}\mathbf{e}_1 - \mathbf{e}_2^T\mathbf{Q}_{12}^T\mathbf{e}_1 \end{bmatrix}.$
- Diagonal factors for $\mathbf{P}_1$ and $\mathbf{P}_2$ are completely decoupled.
- Gives a degree of freedom which allows us to make inner product zero.
- If there are more than two blocks then we can establish consistency recursively a block at a time.

Summary

- If we make a small perturbation to a matrix with block structure then it is reasonable to assume that the associated doubly stochastic matrices are also close together.
- We would like to draw corresponding conclusions about the corresponding singular vectors but so far we can only provide empirical evidence of this.


0
5
10
15
20
25
0
5
10
15
20
25
nz = 50