Linear Algebra

Review & Recent Progress

Gene Cheung

York University

October 17, 2019

Outline

- What is linear algebra?
- Why should I care?
- System of Linear Equations
- Eigen-decomposition

What is Linear Algebra?

- At the risk of over-simplification, linear algebra solves two problems:

[1] Matrix **A** and vector **y** can also be complex: $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{y} \in \mathbb{C}^n$.

[2] One can also pose the generlized eigenvalue problem: $\mathbf{A}\mathbf{v} = \lambda \mathbf{B}\mathbf{v}$.

What is Linear Algebra?

- At the risk of over-simplification, linear algebra solves two problems:
  1. **System of linear equations**:
     given matrix[1] $\mathbf{A} \in \mathbb{R}^{n \times n}$ and vector $\mathbf{y} \in \mathbb{R}^{n}$, find $\mathbf{x}$

$$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{1}$$

[1] Matrix $\mathbf{A}$ and vector $\mathbf{y}$ can also be complex: $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{y} \in \mathbb{C}^{n}$.

[2] One can also pose the generlized eigenvalue problem: $\mathbf{A}\mathbf{v} = \lambda \mathbf{B}\mathbf{v}$.

What is Linear Algebra?

- At the risk of over-simplification, linear algebra solves two problems:
  1. **System of linear equations**:
     given matrix[1] $\mathbf{A} \in \mathbb{R}^{n \times n}$ and vector $\mathbf{y} \in \mathbb{R}^n$, find $\mathbf{x}$
     $$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{1}$$
  2. **Eigen-decomposition**:
     given matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, find eigen-pair[2] $(\lambda, \mathbf{v})$, where $\lambda \in \mathbb{R}$ and $\mathbf{v} \in \mathbb{R}^n$, such that
     $$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{2}$$

[1] Matrix $\mathbf{A}$ and vector $\mathbf{y}$ can also be complex: $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{y} \in \mathbb{C}^n$.

[2] One can also pose the generlized eigenvalue problem: $\mathbf{A}\mathbf{v} = \lambda\mathbf{B}\mathbf{v}$.

Why should I care? (Part I)

- $\mathbf{Ax} = \mathbf{y}$ yields solution to the least-square problem[3]:

Why should I care? (Part I)

- $\mathbf{Ax} = \mathbf{y}$ yields solution to the least-square problem[3]:
  1. **Overdetermined equation**: For "tall" matrix $\mathbf{H}$, solve

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Hx}\|_2^2 \tag{3}$$

$$\implies \mathbf{H}^\top \mathbf{H} \mathbf{x}^* = \mathbf{H}^\top \mathbf{y} \tag{4}$$

Why should I care? (Part I)

- $\mathbf{Ax} = \mathbf{y}$ yields solution to the least-square problem[3]:
  1. **Overdetermined equation**: For "tall" matrix $\mathbf{H}$, solve

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Hx}\|_2^2 \tag{3}$$

$$\Longrightarrow \quad \mathbf{H}^\top \mathbf{H}\, \mathbf{x}^* = \mathbf{H}^\top \mathbf{y} \tag{4}$$

  2. **Underdetermined equation**: For "fat" matrix $\mathbf{H}$, solve

$$\min_{\mathbf{x}} \|\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{Hx} \tag{5}$$

$$\Longrightarrow \quad \begin{bmatrix} \mathbf{I} & -\frac{1}{2}\mathbf{H}^\top \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{6}$$

Why should I care? (Part I)

- $\mathbf{A}\mathbf{x} = \mathbf{y}$ yields solution to the least-square problem[3]:
  1. **Overdetermined equation**: For "tall" matrix $\mathbf{H}$, solve
$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 \tag{3}$$
$$\Longrightarrow \quad \mathbf{H}^\top \mathbf{H}\, \mathbf{x}^* = \mathbf{H}^\top \mathbf{y} \tag{4}$$
  2. **Underdetermined equation**: For "fat" matrix $\mathbf{H}$, solve
$$\min_{\mathbf{x}} \|\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{H}\mathbf{x} \tag{5}$$
$$\Longrightarrow \quad \begin{bmatrix} \mathbf{I} & -\frac{1}{2}\mathbf{H}^\top \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mu \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{6}$$
  3. **Regularization**:
$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda \|\mathbf{A}\mathbf{x}\|_2^2 \tag{7}$$
$$\Longrightarrow \quad \left(\mathbf{H}^\top \mathbf{H} + \lambda \mathbf{A}^\top \mathbf{A}\right) \mathbf{x}^* = \mathbf{H}^\top \mathbf{y} \tag{8}$$

[3] http://eeweb.poly.edu/iselesni/lecture_notes/least_squares/index.html

Why should I care? (Part II)

- Evaluate stability of $\mathbf{Ax} = \mathbf{y}$.
  - Condition number $\kappa(\mathbf{A})$ of coefficient matrix $\mathbf{A}$:

$$\kappa(\mathbf{A}) = \frac{|\lambda_{\text{max}}(\mathbf{A})|}{|\lambda_{\text{min}}(\mathbf{A})|} \tag{9}$$

Why should I care? (Part II)

- Evaluate stability of $\mathbf{Ax} = \mathbf{y}$.
  - Condition number $\kappa(\mathbf{A})$ of coefficient matrix $\mathbf{A}$:

$$\kappa(\mathbf{A}) = \frac{|\lambda_{\text{max}}(\mathbf{A})|}{|\lambda_{\text{min}}(\mathbf{A})|} \tag{9}$$

  - E-optimality criteria:

$$\max_{\mathbf{A}\in\mathcal{S}} \ \lambda_{\text{min}}(\mathbf{A}) \tag{10}$$

Why should I care? (Part II)

- Evaluate stability of $\mathbf{Ax} = \mathbf{y}$.
  - Condition number $\kappa(\mathbf{A})$ of coefficient matrix $\mathbf{A}$:

$$\kappa(\mathbf{A}) = \frac{|\lambda_{\max}(\mathbf{A})|}{|\lambda_{\min}(\mathbf{A})|} \tag{9}$$

  - E-optimality criteria:

$$\max_{\mathbf{A}\in\mathcal{S}} \ \lambda_{\min}(\mathbf{A}) \tag{10}$$

- Compute graph frequencies [4]:
  1. Define variation operator $\mathbf{\Phi}$ on graph, *e.g.*, *graph Laplacian matrix* $\mathbf{L}$:

$$\mathbf{L} = \mathbf{D} - \mathbf{W} \tag{11}$$

  2. Compute Fourier modes for $\mathbf{L}$ via eigen-decomposition.

$$\mathbf{L} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top \tag{12}$$

where $\mathbf{V}$ contains eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots$ as columns, and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots)$.

[4] A. Ortega et al., "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

Graph Frequencies: Discrete Cosine Transform (GCT)

- Cosines are eigenfunctions of differential operator:
  1. Construct graph Laplacian **L** for line graph with weights 1.
  2. Compute eigenvectors for **L**.

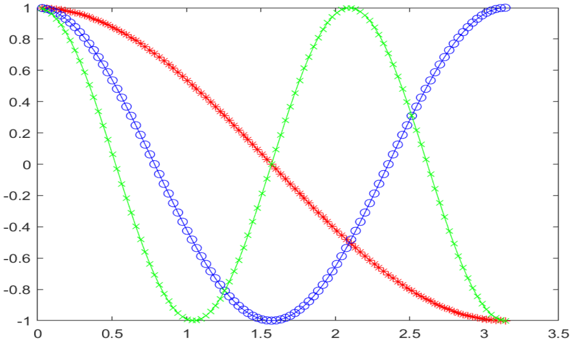

Figure: Eigenvectors of line graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$

Graph Frequencies: Discrete Fourier Transform (GFT)

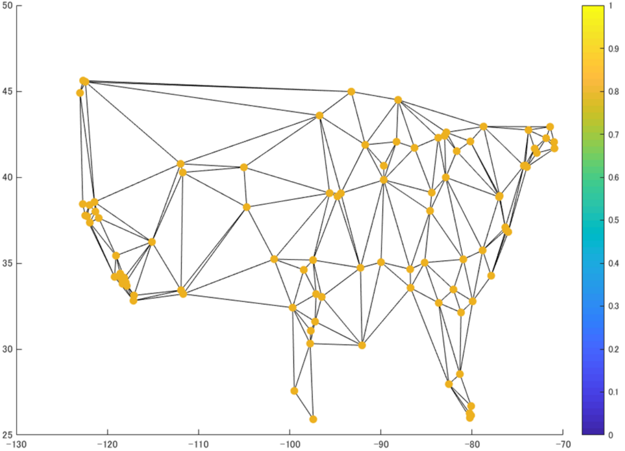

Figure: 1st eigenvector of graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$

Graph Frequencies: Discrete Fourier Transform (GFT)

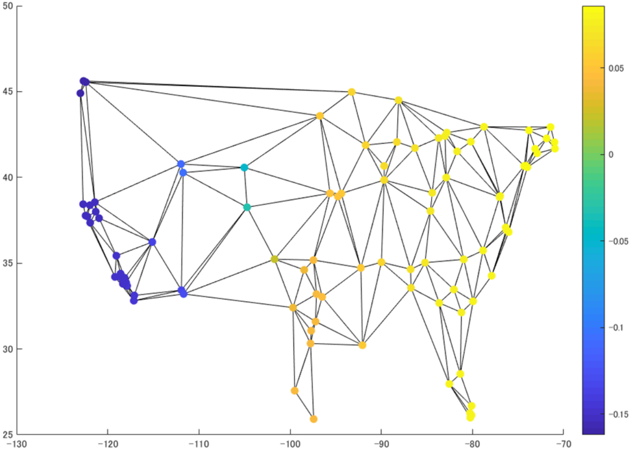

Figure: 2nd eigenvector of graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$

Graph Frequencies: Discrete Fourier Transform (GFT)

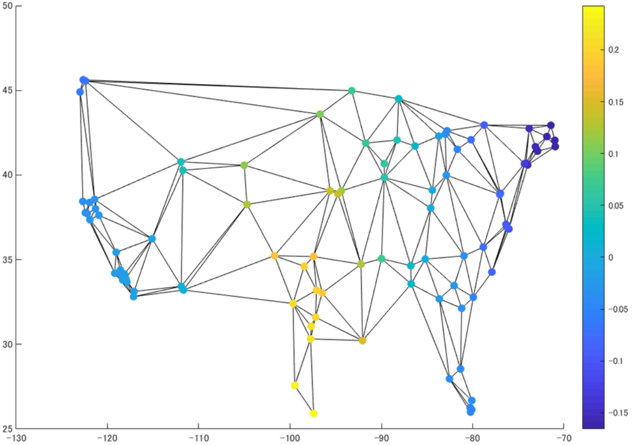

Figure: 3rd eigenvector of graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$

Graph Frequencies: Discrete Fourier Transform (GFT)

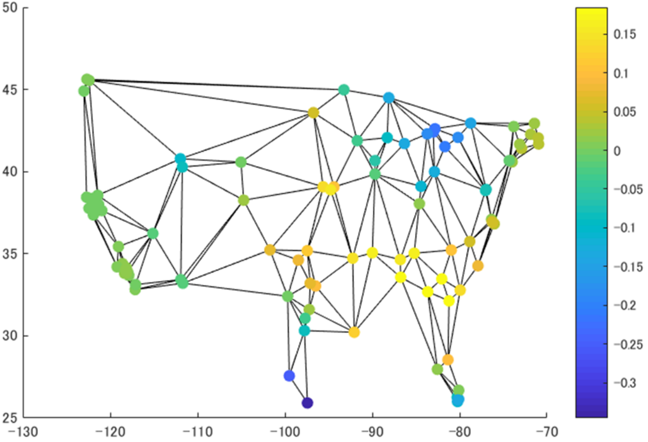

Figure: 9th eigenvector of graph Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$

Solving a System of Linear Equations

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Solving a System of Linear Equations

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

1. **Gaussian elimination**.

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Solving a System of Linear Equations

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

1. **Gaussian elimination**.
2. If $\mathbf{A}$ is symmetric and positive semi-definite (PSD), then **Cholesky decomposition**.

---

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

1. **Gaussian elimination**.
2. If $\mathbf{A}$ is symmetric and positive semi-definite (PSD), then **Cholesky decomposition**.
3. If $\mathbf{A}$ is diagonally dominant, **Jacobi method**.

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Solving a System of Linear Equations

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

1. **Gaussian elimination**.
2. If $\mathbf{A}$ is symmetric and positive semi-definite (PSD), then **Cholesky decomposition**.
3. If $\mathbf{A}$ is diagonally dominant, **Jacobi method**.
4. If $\mathbf{A}$ is diagonally dominant, or symmetric and PD, **Gauss-Seidel method**.

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Solving a System of Linear Equations

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

1. **Gaussian elimination**.
2. If $\mathbf{A}$ is symmetric and positive semi-definite (PSD), then **Cholesky decomposition**.
3. If $\mathbf{A}$ is diagonally dominant, **Jacobi method**.
4. If $\mathbf{A}$ is diagonally dominant, or symmetric and PD, **Gauss-Seidel method**.
5. Successive approximations using **Krylov subspace methods**:

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Solving a System of Linear Equations

Q: How to solve $\mathbf{Ax} = \mathbf{b}$?

A: Many methods.

1. **Gaussian elimination**.
2. If $\mathbf{A}$ is symmetric and positive semi-definite (PSD), then **Cholesky decomposition**.
3. If $\mathbf{A}$ is diagonally dominant, **Jacobi method**.
4. If $\mathbf{A}$ is diagonally dominant, or symmetric and PD, **Gauss-Seidel method**.
5. Successive approximations using **Krylov subspace methods**:
    - If $\mathbf{A}$ symmetric and PD, **Conjugate Gradient**[5].

[5] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, 1952, vol. 49, no. 1.

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{Av} = \lambda\mathbf{v}$?

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

1. **Power method**.

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{Av} = \lambda \mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

1. **Power method**.
2. **QR algorithm**.

Q: How to solve $\mathbf{Av} = \lambda \mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

1. **Power method**.
2. **QR algorithm**.
3. Successive approximations using **Krylov subspace methods**:

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

1. **Power method**.
2. **QR algorithm**.
3. Successive approximations using **Krylov subspace methods**:
    - Arnoldi iteration.

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

1. **Power method**.
2. **QR algorithm**.
3. Successive approximations using **Krylov subspace methods**:
    - Arnoldi iteration.
    - **Divide-and-conquer algorithm**.

Solving the Eigen-decomposition Problem

Q: How to solve $\mathbf{Av} = \lambda \mathbf{v}$?

1. How to compute eigenvalues λ's?
2. How to compute eigenvectors $\mathbf{v}$'s?

A: Many methods.

1. **Power method**.
2. **QR algorithm**.
3. Successive approximations using **Krylov subspace methods**:
   - Arnoldi iteration.
   - **Divide-and-conquer algorithm**.
   - If **A** Hermetian, **Lanczos algorithm**.

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

- Recall E-optimality criteria:

$$\max_{\mathbf{A} \in \mathcal{S}} \lambda_{\min}(\mathbf{A}) \tag{13}$$

[6] R. S. Varga, *Gersgorin and His Circles* Springer, 2004

- Recall E-optimality criteria:

$$\max_{\mathbf{A}\in\mathcal{S}} \lambda_{\min}(\mathbf{A}) \tag{13}$$

Gershgorin Circle Theorem (GCT) [6]

[6] R. S. Varga, *Gersgorin and His Circles* Springer, 2004

- Recall E-optimality criteria:

$$\max_{\mathbf{A} \in \mathcal{S}} \lambda_{\min}(\mathbf{A}) \tag{13}$$

Gershgorin Circle Theorem (GCT) [6]

- Given matrix $\mathbf{A}$, Gershgorin disc i, $\mathbf{\Phi}_i$, has centre $c_i = A_{ii}$ and radius $r_i = \sum_{j \neq i} |A_{ij}|$.

[6] R. S. Varga, *Gersgorin and His Circles* Springer, 2004

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

- Recall E-optimality criteria:

$$\max_{\mathbf{A}\in\mathcal{S}} \lambda_{\min}(\mathbf{A}) \tag{13}$$

Gershgorin Circle Theorem (GCT) [6]

- Given matrix $\mathbf{A}$, Gershgorin disc i, $\mathbf{\Phi}_i$, has centre $c_i = A_{ii}$ and radius $r_i = \sum_{j\neq i} |A_{ij}|$.
- For each eigenvalue λ of $\mathbf{A}$, there exists a disc $\mathbf{\Phi}_i$ such that:

$$c_i - r_i \leq \lambda \leq c_i + r_i \tag{14}$$

[6] R. S. Varga, *Gersgorin and His Circles* Springer, 2004

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

- Recall E-optimality criteria:

$$\max_{\mathbf{A} \in \mathcal{S}} \lambda_{\min}(\mathbf{A}) \tag{13}$$

Gershgorin Circle Theorem (GCT) [6]

- Given matrix $\mathbf{A}$, Gershgorin disc i, $\mathbf{\Phi}_i$, has centre $c_i = A_{ii}$ and radius $r_i = \sum_{j \neq i} |A_{ij}|$.
- For each eigenvalue λ of $\mathbf{A}$, there exists a disc $\mathbf{\Phi}_i$ such that:

$$c_i - r_i \leq \lambda \leq c_i + r_i \tag{14}$$

- Compute a lower bound for $\lambda_{\min}(\mathbf{A})$ without computing eigen-pairs:

$$\lambda^{-}_{\min}(\mathbf{A}) = \min_i c_i - r_i \leq \lambda_{\min}(\mathbf{A}) \tag{15}$$

[6] R. S. Varga, *Gersgorin and His Circles* Springer, 2004

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

Q: How tight is GCT lower bound $\lambda^{-}_{\min}(\mathbf{A})$ for $\lambda_{\min}(\mathbf{A})$?

[7] Y. Bai et al., "Fast Graph Sampling Set Selection Using Gershgorin Disc Alignment," submitted to *IEEE Transactions on Signal Processing*, July 2019.

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

Q: How tight is GCT lower bound $\lambda^-_{\min}(\mathbf{A})$ for $\lambda_{\min}(\mathbf{A})$?
A: Consider instead similar transform[7]:

$$\mathbf{B} = \mathbf{SAS}^{-1} \tag{16}$$

where $\mathbf{S} = \mathrm{diag}(s_1, s_2, \ldots)$.

[7] Y. Bai et al., "Fast Graph Sampling Set Selection Using Gershgorin Disc Alignment," submitted to *IEEE Transactions on Signal Processing*, July 2019.

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

Q: How tight is GCT lower bound $\lambda^{-}_{\min}(\mathbf{A})$ for $\lambda_{\min}(\mathbf{A})$?
A: Consider instead similar transform[7]:

$$\mathbf{B} = \mathbf{S}\mathbf{A}\mathbf{S}^{-1} \tag{16}$$

where $\mathbf{S} = \mathrm{diag}(s_1, s_2, \ldots)$.

- **B** has same eigenvalues as **A**.

[7] Y. Bai et al., "Fast Graph Sampling Set Selection Using Gershgorin Disc Alignment," submitted to *IEEE Transactions on Signal Processing*, July 2019.

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

Q: How tight is GCT lower bound $\lambda^-_{\min}(\mathbf{A})$ for $\lambda_{\min}(\mathbf{A})$?
A: Consider instead similar transform[7]:

$$\mathbf{B} = \mathbf{SAS}^{-1} \tag{16}$$

where $\mathbf{S} = \mathrm{diag}(s_1, s_2, \ldots)$.

- $\mathbf{B}$ has same eigenvalues as $\mathbf{A}$.
- If $\mathbf{A}$ is a *generalized graph Laplacian* with positive edges, $\exists\mathbf{S}$ such that $\lambda^-_{\min}(\mathbf{B}) = \lambda_{\min}(\mathbf{B}) = \lambda_{\min}(\mathbf{A})$.

[7] Y. Bai et al., "Fast Graph Sampling Set Selection Using Gershgorin Disc Alignment," submitted to *IEEE Transactions on Signal Processing*, July 2019.

Lower-bounding Smallest Eigenvalue λ_{min}

- Partition matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22} \end{bmatrix} \tag{17}$$

[8] G. Cheung et al., "Robust Semisupervised Graph Classifier Learning with Negative Edge Weights," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no.4, pp.712-726, December 2018.

Lower-bounding Smallest Eigenvalue λ_{min}

- Partition matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22} \end{bmatrix} \tag{17}$$

- Define Schur Complement:

$$\mathbf{A}/\mathbf{A}_{11} = \mathbf{A}_{22} - \mathbf{A}_{12}^\top \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \tag{18}$$

[8] G. Cheung et al., "Robust Semisupervised Graph Classifier Learning with Negative Edge Weights," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no.4, pp.712-726, December 2018.

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

- Partition matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22} \end{bmatrix} \tag{17}$$

- Define Schur Complement:

$$\mathbf{A}/\mathbf{A}_{11} = \mathbf{A}_{22} - \mathbf{A}_{12}^\top \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \tag{18}$$

- Haynsworth Inertia Additivity formula:

$$\mathrm{In}(\mathbf{A}) = \mathrm{In}(\mathbf{A}_{11}) + \mathrm{In}(\mathbf{A}/\mathbf{A}_{11}) \tag{19}$$

[8] G. Cheung et al., "Robust Semisupervised Graph Classifier Learning with Negative Edge Weights," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no.4, pp.712-726, December 2018.

Lower-bounding Smallest Eigenvalue $\lambda_{\min}$

- Partition matrix $\mathbf{A}$:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^\top & \mathbf{A}_{22} \end{bmatrix} \tag{17}$$

- Define Schur Complement:

$$\mathbf{A}/\mathbf{A}_{11} = \mathbf{A}_{22} - \mathbf{A}_{12}^\top \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \tag{18}$$

- Haynsworth Inertia Additivity formula:

$$\mathrm{In}(\mathbf{A}) = \mathrm{In}(\mathbf{A}_{11}) + \mathrm{In}(\mathbf{A}/\mathbf{A}_{11}) \tag{19}$$

- If $\mathbf{A}$ is real and symmetric, can compute tight lower bound for $\lambda_{\min}$ using a recursive procedure plus shifting $(+\mu \mathbf{I})$ [8].

[8] G. Cheung et al., "Robust Semisupervised Graph Classifier Learning with Negative Edge Weights," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no.4, pp.712-726, December 2018.

Computing Extreme Eigenvectors

locally optimal block preconditioned conjugated gradient (LOBPCG) method [9]:

- Compute extreme eigen-pairs for symmetric, PD matrices.
- Cost per iteration competitive with Lanczos.
- Can directly take advantage of preconditioning.
- Can benefit from warm start.

[9] A. Knyazev, "Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method," *SIAM journal on scientific computing*, 23(2):517–541, 2001.

Summary

- Linear algebra solves:
  1. $\mathbf{Ax} = \mathbf{y}$
  2. $\mathbf{Av} = \lambda \mathbf{v}$.
- Mature algorithms to solve $\mathbf{Ax} = \mathbf{y}$ using Krylov methods.
- Fast algorithms to find extreme eigen-pairs $(\lambda_i, \mathbf{v}_i)$.
- New algorithms to find tight lower bounds for λ_{min}.