ENGG 5781 Matrix Analysis and Computations Lecture 3: Eigenvalues and Eigenvectors

Wing-Kin (Ken) Ma

2024–25 First Term

Department of Electronic Engineering The Chinese University of Hong Kong

Lecture 3: Eigenvalues and Eigenvectors

- facts about eigenvalues and eigenvectors
- eigendecomposition, the case of Hermitian and real symmetric matrices
- power method
- Schur decomposition
- PageRank: a case study

Notation and Conventions

- a square matrix \mathbf{A} is said to be symmetric if $a_{ij} = a_{ji}$ for all i, j with $i \neq j$, or equivalently, if $\mathbf{A}^T = \mathbf{A}$
 - example:

$$\mathbf{A} = \begin{bmatrix} 1 & -0.5 & 3 \\ -0.5 & -2 & 0.9 \\ 3 & 0.9 & 0.1 \end{bmatrix}$$

- a square matrix A is said to be Hermitian if $a_{ij} = a_{ji}^*$ for all i, j with $i \neq j$, or equivalently, if $A^H = A$
- we denote the set of all $n \times n$ real symmetric matrices by \mathbb{S}^n
- we denote the set of all $n \times n$ complex Hermitian matrices by \mathbb{H}^n

Notation and Conventions

- note the following subtleties:
 - by definition, a real symmetric matrix is also Hermitian
 - when we say that a matrix is Hermitian, we often imply that the matrix may be complex (at least for this course); a real Hermitian matrix is simply real symmetric
 - we can have a complex symmetric matrix, though we will not study it

Main Results

A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) is said to admit an eigendecomposition if there exists a nonsingular $\mathbf{V} \in \mathbb{C}^{n \times n}$ and a collection of scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1},$$

where $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n)$.

- the above $(\mathbf{V}, \mathbf{\Lambda})$ satisfies $\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for $i = 1, \dots, n$, which are eigen-equations
- $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are required to be linearly independent
- eigendecomposition *does not* always exist

Main Results

A real symmetric matrix $\mathbf{A} \in \mathbb{S}^n$ always admits an eigendecomposition

 $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$

where $\mathbf{V} \in \mathbb{R}^{n \times n}$ is orthogonal; $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_i \in \mathbb{R}$ for all i.

A Hermitian matrix $\mathbf{A} \in \mathbb{H}^n$ always admits an eigendecomposition

 $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$

where $\mathbf{V} \in \mathbb{C}^{n \times n}$ is unitary; $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_i \in \mathbb{R}$ for all i.

- differences: a Hermitian or real symmetric matrix always has
 - an eigendecomposition
 - real λ_i 's
 - a ${\bf V}$ that is not only nonsingular but also unitary

We start with the basic definition of eigenvalues and eigenvectors.

Problem: given a $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$), find a vector $\mathbf{v} \in \mathbb{C}^n$ with $\mathbf{v} \neq \mathbf{0}$ such that

$$\mathbf{Av} = \lambda \mathbf{v}, \quad \text{for some } \lambda \in \mathbb{C}$$
 (*)

- (*) is called an eigenvalue problem or eigen-equation
- let (\mathbf{v}, λ) be a solution to (*). We call
 - (\mathbf{v}, λ) an eigen-pair of \mathbf{A}
 - λ an eigenvalue of A; v an eigenvector of A associated with λ
- if (\mathbf{v}, λ) is an eigen-pair of \mathbf{A} , $(\alpha \mathbf{v}, \lambda)$ is also an eigen-pair for any $\alpha \in \mathbb{C}, \alpha \neq 0$
- unless specified, we will assume $\|\mathbf{v}\|_2 = 1$ in the sequel

Fact: Every $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) has n eigenvalues.

• from the eigenvalue problem we see that

$$\begin{aligned} \mathbf{A}\mathbf{v} &= \lambda\mathbf{v} \text{ for some } \mathbf{v} \neq \mathbf{0} &\iff (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \text{ for some } \mathbf{v} \neq \mathbf{0} \\ &\iff \det(\mathbf{A} - \lambda\mathbf{I}) = 0 \end{aligned}$$

- let $p(\lambda) = \det(\mathbf{A} \lambda \mathbf{I})$, called the characteristic polynomial of \mathbf{A}
- from the determinant def., it can be shown that $p(\lambda)$ is a polynomial of degree n, viz., $p(\lambda) = \alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2 + \ldots + \alpha_n \lambda^n$ where α_i 's depend on \mathbf{A}
- as $p(\lambda)$ is a polynomial of degree n, it can be factored as $p(\lambda) = \prod_{i=1}^{n} (\lambda_i \lambda)$ where $\lambda_1, \ldots, \lambda_n$ are the roots of $p(\lambda)$
- we have $det(\mathbf{A} \lambda \mathbf{I}) = 0 \iff \lambda \in \{\lambda_1, \dots, \lambda_n\}$

Let $\lambda_1, \ldots, \lambda_n$ denote the *n* eigenvalues of **A**. We write

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i, \qquad i = 1, \dots, n,$$

where \mathbf{v}_i denotes an eigenvector of \mathbf{A} associated with λ_i .

- we should be careful about the meaning of n eigenvalues: they are defined as the n roots of the characteristic polynomial $p(\lambda) = \det(\mathbf{A} \lambda \mathbf{I})$
- example: consider

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

- from the original definition $Av = \lambda v$, one can verify that $\lambda = 1$ is the only eigenvalue of A
- from the characteristic polynomial, which is $p(\lambda) = (1 \lambda)^2$, we see two roots $\lambda_1 = \lambda_2 = 1$ as two eigenvalues

Fact: an eigenvalue can be complex even if \mathbf{A} is real.

- a polynomial $p(\lambda) = \alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2 + \ldots + \alpha_n \lambda^n$ with real coefficients α_i 's can have complex roots
- example: consider

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

– we have
$$p(\lambda) = \lambda^2 + 1$$
, so $\lambda_1 = oldsymbol{j}$, $\lambda_2 = -oldsymbol{j}$

Fact: if A is real and there exists a real eigenvalue λ of A, the associated eigenvector v can be taken as real.

- obviously, when $\mathbf{A} \lambda \mathbf{I}$ is real we can define $\mathcal{N}(\mathbf{A} \lambda \mathbf{I})$ on \mathbb{R}^n
- or, if v is a complex eigenvector of a real A associated with a real λ , we can write $\mathbf{v} = \mathbf{v}_{\mathrm{R}} + \boldsymbol{j}\mathbf{v}_{\mathrm{I}}$, where $\mathbf{v}_{\mathrm{R}}, \mathbf{v}_{\mathrm{I}} \in \mathbb{R}^{n}$. It is easy to verify that \mathbf{v}_{R} and \mathbf{v}_{I} are eigenvectors associated with λ

Further Discussion: Repeated Eigenvalues

- w.l.o.g., order $\lambda_1, \ldots, \lambda_n$ such that $\{\lambda_1, \ldots, \lambda_k\}$, $k \leq n$, is the set of all distinct eigenvalues of **A**; i.e., $\lambda_i \neq \lambda_j$ for all $i, j \in \{1, \ldots, k\}$, $i \neq j$; $\lambda_i \in \{\lambda_1, \ldots, \lambda_k\}$ for all $i \in \{1, \ldots, n\}$
- denote μ_i as the number of repeated eigenvalues of λ_i , $i = 1, \ldots, k$
 - μ_i is called the algebraic multiplicity of the eigenvalue λ_i
- every λ_i can have more than one eigenvector (scaling not counted)
 - if $\dim \mathcal{N}(\mathbf{A} \lambda_i \mathbf{I}) = r$, we can find r linearly independent \mathbf{v}_i 's
 - denote $\gamma_i = \dim \mathcal{N}(\mathbf{A} \lambda_i \mathbf{I})$, $i = 1, \dots, k$
 - γ_i is called the geometric multiplicity of the eigenvalue λ_i

Property 3.1. We have $\mu_i \ge \gamma_i$ for all i = 1, ..., k (not trivial, requires a proof)

- Implication: no. of repeated eigenvalues \geq no. of linearly indep. eigenvectors

A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) is said to be diagonalizable, or admit an eigendecomposition, if there exists a nonsingular $\mathbf{V} \in \mathbb{C}^{n \times n}$ such that

$$\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1},$$

where $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n)$.

in defining diagonalizability, we didn't say that (v_i, λ_i) has to be an eigen-pair of
 A. But

 $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1} \iff \mathbf{A} \mathbf{V} = \mathbf{V} \mathbf{\Lambda}, \ \mathbf{V} \text{ nonsingular}$ $\iff \mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{v}_i, \ i = 1, \dots, n, \ \mathbf{V} \text{ nonsingular}$

Also, $\lambda_1, \ldots, \lambda_n$ must be the *n* eigenvalues of **A**; this can be seen from the characteristic polynomial $\det(\mathbf{A} - \lambda \mathbf{I}) = \det(\mathbf{\Lambda} - \lambda \mathbf{I}) = \prod_{i=1}^n (\lambda_i - \lambda)$

• the non-trivial part lies in finding n linearly independent eigenvectors

If A admits an eigendecomposition, the following properties can be shown (easily):

•
$$\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$$

•
$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$$

- the eigenvalues of \mathbf{A}^k are $\lambda_1^k, \dots, \lambda_n^k$
- $rank(\mathbf{A}) = number of nonzero eigenvalues of \mathbf{A}$
- suppose that A is also nonsingular. Then, $\mathbf{A}^{-1} = \mathbf{V} \mathbf{\Lambda}^{-1} \mathbf{V}^{-1}$

Note: the first three properties can be shown to be valid for any A; the fourth property may not be valid when A does not admit an eigendecomposition

Question: Does every $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) admit an eigendecomposition?

- the answer is no.
- counter example: consider

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- the characteristic polynomial is $p(\lambda) = -\lambda^3$, so $\lambda_1 = \lambda_2 = \lambda_3 = 0$
- it is easy to see that

$$\mathcal{N}(\mathbf{A} - \lambda_1 \mathbf{I}) = \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{A}^T)^{\perp} = \operatorname{span} \left\{ \begin{bmatrix} 1\\0\\0 \end{bmatrix}, \begin{bmatrix} 0\\1\\0 \end{bmatrix} \right\}$$

– any selection of $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3\in\mathcal{N}(\mathbf{A})$ is linearly dependent

Question: under which conditions can a matrix admit an eigendcomposition?

- there exist matrix subclasses in which eigendecomposition is guaranteed to exist
 - one example is the circulant matrix subclass, as seen in the last lecture
 - another example is the Hermitian matrix subclass, as we will see
- there exist simple sufficient conditions under which eigendec. exists

Property 3.2. Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$), and suppose that λ_i 's are ordered such that $\{\lambda_1, \ldots, \lambda_k\}$ is the set of all distinct eigenvalues of \mathbf{A} . Also, let \mathbf{v}_i be any eigenvector associated with λ_i . Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ must be linearly independent.

Implications:

• if all the eigenvalues of A are distinct, i.e.,

$$\lambda_i \neq \lambda_j$$
, for all $i, j \in \{1, \dots, n\}$ with $i \neq j$,

then \mathbf{A} admits an eigendecomposition

- to have all the eigenvalues to be distinct is not that hard, as we will see later
- A admits an eigendcomposition if and only if $\mu_i = \gamma_i$ for all i

Eigendecomposition for Hermitian & Real Symmetric Matrices

Consider the Hermitian matrix subclass.

Property 3.3. Let $\mathbf{A} \in \mathbb{H}^n$.

- 1. the eigenvalues $\lambda_1, \ldots, \lambda_n$ of **A** are real
- 2. suppose that λ_i 's are ordered such that $\{\lambda_1, \ldots, \lambda_k\}$ is the set of all distinct eigenvalues of **A**. Also, let \mathbf{v}_i be *any* eigenvector associated with λ_i . Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ must be orthonormal.
- the above results apply to real symmetric matrices; recall $\mathbf{A} \in \mathbb{S}^n \Longrightarrow \mathbf{A} \in \mathbb{H}^n$
- Corollary: for a real symmetric matrix, all eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ can be chosen as real

Eigendecomposition for Real Symmetric & Hermitian Matrices

Theorem 3.1. Every $\mathbf{A} \in \mathbb{H}^n$ admits an eigendecomposition

 $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H,$

where $\mathbf{V} \in \mathbb{C}^{n \times n}$ is unitary; $\mathbf{\Lambda} = \text{Diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_i \in \mathbb{R}$ for all *i*. Also, if $\mathbf{A} \in \mathbb{S}^n$, \mathbf{V} can be taken as real orthogonal.

- a consequence of a more powerful decomposition, namely, the Schur decomposition; we will go through it later
- does not require the assumption of distinct eigenvalues
- Corollary: if A is Hermitian or real symmetric, $\mu_i = \gamma_i$ for all i (no. of repeated eigenvalues = no. of linearly indep. eigenvectors)

Power Method

- a method of numerically computing an eigenvector of a given matrix
- simple
- not the best in convergence speed
 - a comprehensive coverage of various computational methods for the eigenvalue problem can be found in the textbook [Golub-Van Loan'12]
- suitable for large-scale sparse problems, e.g., PageRank

Power Method

- assumptions:
 - ${\bf A}$ admits an eigendecomposition
 - λ_i 's are ordered such that $|\lambda_1| \ge |\lambda_2| \ge \ldots \ge |\lambda_n|$
 - $|\lambda_1| > |\lambda_2|$
 - we have an initial guess x that satisfies $[\mathbf{V}^{-1}\mathbf{x}]_1 \neq 0$ (random guess should do)
- consider $\mathbf{A}^k \mathbf{x}$. Let $\boldsymbol{\alpha} = \mathbf{V}^{-1} \mathbf{x}$, and observe

$$\mathbf{A}^{k}\mathbf{x} = \mathbf{V}\mathbf{\Lambda}^{k}\mathbf{V}^{-1}\mathbf{x} = \sum_{i=1}^{n} \alpha_{i}\lambda_{i}^{k}\mathbf{v}_{i} = \alpha_{1}\lambda_{1}^{k}\left(\mathbf{v}_{1} + \sum_{i=2}^{n} \frac{\alpha_{i}}{\alpha_{1}}\left(\frac{\lambda_{i}}{\lambda_{1}}\right)^{k}\mathbf{v}_{i}\right)$$

where \mathbf{r}_k is a residual and has

$$\|\mathbf{r}_k\|_2 \le \sum_{i=2}^n \left|\frac{\alpha_i}{\alpha_1}\right| \left|\frac{\lambda_i}{\lambda_1}\right|^k \|\mathbf{v}_i\|_2 \le \left|\frac{\lambda_2}{\lambda_1}\right|^k \sum_{i=2}^n \left|\frac{\alpha_i}{\alpha_1}\right|$$

• convergence: let
$$c_k = \frac{|\alpha_1||\lambda|^k}{\alpha_1 \lambda_1^k}$$
 (note $|c_k| = 1$). We have
$$\lim_{k \to \infty} c_k \frac{\mathbf{A}^k \mathbf{x}}{\|\mathbf{A}^k \mathbf{x}\|_2} = \mathbf{v}_1$$

Power Method

```
Algorithm: Power Method

input: \mathbf{A} \in \mathbb{C}^{n \times n} and a starting point \mathbf{v}^{(0)} \in \mathbb{C}^n

k = 0

repeat

\tilde{\mathbf{v}}^{(k+1)} = \mathbf{A}\mathbf{v}^{(k)}

\mathbf{v}^{(k+1)} = \tilde{\mathbf{v}}^{(k+1)} / \|\tilde{\mathbf{v}}^{(k+1)}\|_2

k := k + 1

until a stopping rule is satisfied

output: \mathbf{v}^{(k)}
```

• it can be verified that
$$\mathbf{v}^{(k)} = rac{\mathbf{A}^k \mathbf{v}^{(0)}}{\|\mathbf{A}^k \mathbf{v}^{(0)}\|_2}$$

- complexity per iteration: $\mathcal{O}(n^2)$, or $\mathcal{O}(\mathrm{nnz}(\mathbf{A}))$ for sparse \mathbf{A}
- convergence rate depends on $\left|\frac{\lambda_2}{\lambda_1}\right|$; slower if $|\lambda_2|$ is closer to $|\lambda_1|$

Deflation

- the power method finds the largest eigenvalue (in modulus) only
- how can we compute all the eigenvalues and eigenvectors?
- there are many ways and let's consider a simple method called deflation
- consider a Hermitian A with $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$, and note the outer-product representation n

$$\mathbf{A} = \sum_{i=1}^{H} \lambda_i \mathbf{v}_i \mathbf{v}_i^H.$$

• Deflation: use the power method to obtain \mathbf{v}_1, λ_1 , do the subtraction

$$\mathbf{A} := \mathbf{A} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^H = \sum_{i=2}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^H,$$

and repeat until all the eigenvectors and eigenvalues are found

- if we want the first k eigen-pairs only, deflation can also do that

Schur Decomposition

Theorem 3.2. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$, and let $\lambda_1, \ldots, \lambda_n$ be its eigenvalues. The matrix \mathbf{A} admits a decomposition

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H,$$

for some unitary $\mathbf{U} \in \mathbb{C}^{n \times n}$ and for some upper triangular $\mathbf{T} \in \mathbb{C}^{n \times n}$ with $t_{ii} = \lambda_i$ for all *i*. If **A** is real and $\lambda_1, \ldots, \lambda_n$ are all real, **U** and **T** can be taken as real.

- we will call the above decomposition the Schur decomposition in the sequel
- some insight: Suppose A can be written as $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$ for some unitary U and upper triangular T, but it's not known if $t_{ii} = \lambda_i$. Then

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \det(\mathbf{T} - \lambda \mathbf{I}) = \prod_{i=1}^{n} (t_{ii} - \lambda)$$

This implies that t_{11}, \ldots, t_{nn} are the eigenvalues of **A**

• see the accompanying note for the proof of Theorem 3.2

Schur Decomposition

- the Schur decomposition is a powerful tool
- e.g., we can use it to show that for *any* square A (with or without eigendec.),
 - det(\mathbf{A}) = $\prod_{i=1}^{n} \lambda_i$
 - $-\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i$
 - the eigenvalues of \mathbf{A}^k are $\lambda_1^k, \dots, \lambda_n^k$
- we may use it to prove the convergence of the power method when eigendecomposition does not exist
- the Jordan canonical form, which we will not teach, requires the Schur decomposition as the first key step

Implications of the Schur Decomposition

- proof of Theorem 3.1:
 - let A be Hermitian, and let $A = UTU^H$ be its Schur decomposition. Observe

$$\mathbf{0} = \mathbf{A} - \mathbf{A}^{H} = \mathbf{U}\mathbf{T}\mathbf{U}^{H} - \mathbf{U}\mathbf{T}^{H}\mathbf{U}^{H} = \mathbf{U}(\mathbf{T} - \mathbf{T}^{H})\mathbf{U}^{H} \quad \Longleftrightarrow \quad \mathbf{0} = \mathbf{T} - \mathbf{T}^{H}$$

- since T is upper triangular and T^H is lower triangular, $T = T^H$ implies that T is diagonal; thus, the Schur decomposition is also the eigendecomposition
- similar results apply to real symmetric ${\bf A},$ except that we use real ${\bf T}, {\bf U}$
- note: $\mathbf{T} = \mathbf{T}^H$ also implies that t_{ii} 's are real; so the proof also confirms that λ_i 's are real
- skew-Hermitian matrices: $\mathbf{A} \in \mathbb{C}^{n \times n}$ is said to be skew-Hermitian if $\mathbf{A}^H = -\mathbf{A}$
 - by the Schur decomposition, we can show that any skew-Hermitian A admits an eigendecomposition with unitary V and the eigenvalues are purely imaginary

Implications of Schur Decomposition

• another result from the Schur decomposition:

Proposition 3.1. Let $\mathbf{A} \in \mathbb{C}^{n \times n}$. For every $\varepsilon > 0$, there exists a matrix $\tilde{\mathbf{A}} \in \mathbb{C}^{n \times n}$ such that the *n* eigenvalues of $\tilde{\mathbf{A}}$ are distinct and

$$\|\mathbf{A} - \tilde{\mathbf{A}}\|_F \le \varepsilon.$$

- Implication: for any square A, we can always find an \tilde{A} that is arbitrarily close to A and admits an eigendecomposition
- proof:

- let
$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$$
 be the Schur dec. of \mathbf{A}

- let $\tilde{\mathbf{A}} = \mathbf{U}(\mathbf{T} + \mathbf{D})\mathbf{U}^{H}$, where $\mathbf{D} = \text{Diag}(d_{1}, \dots, d_{n})$ is such that $|d_{i}| \leq \left(\frac{\varepsilon}{n}\right)^{1/2}$ for all i, and $t_{11} + d_{1}, \dots, t_{nn} + d_{n}$ are distinct

-
$$\|\mathbf{A} - \tilde{\mathbf{A}}\|_F^2 = \|\mathbf{D}\|_F^2 \le \varepsilon$$
, the eigenvalues of $\tilde{\mathbf{A}}$ are $t_{11} + d_1, \dots, t_{nn} + d_n$

– by Property 3.2, $\tilde{\mathbf{A}}$ admits an eigendecomposition

PageRank: A Case Study

- PageRank is an algorithm used by Google to rank the pages of a search result.
- the idea is to use counts of links of various pages to determine pages' importance.



• further reading: [Bryan-Tanya2006]

PageRank Model

• Model:

$$\sum_{j \in \mathcal{L}_i} \frac{v_j}{c_j} = v_i, \quad i = 1, \dots, n,$$

where c_j is the number of outgoing links from page j; \mathcal{L}_i is the set of pages with a link to page i; v_i is the importance score of page i.

• example:



PageRank Problem

- let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a matrix such that $a_{ij} = 1/c_j$ if $j \in \mathcal{L}_i$ and $a_{ij} = 0$ if $j \notin \mathcal{L}_i$
- \bullet Problem: find a non-negative ${\bf v}$ such that ${\bf A}{\bf v}={\bf v}$
 - \mathbf{A} is extremely large and sparse, and we want to use the power method
- Questions:
 - does a solution to Av = v exist? Or, is $\lambda = 1$ an eigenvalue of A?
 - does Av = v have a non-negative solution? Or, does a non-negative eigenvector associated with $\lambda = 1$ exist?
 - is the solution to Av = v unique? Or, would there exist more than one eigenvector associated with $\lambda = 1$?
 - * a unique solution is desired for this problem
 - is $\lambda = 1$ the only eigenvalue that is the largest in modulus?
 - $\ast\,$ this is required for the power method

Some Notation and Conventions

- notation:
 - $\mathbf{x} \geq \mathbf{y}$ means that $x_i \geq y_i$ for all i
 - $\mathbf{x} > \mathbf{y}$ means that $x_i > y_i$ for all i
 - $\mathbf{x} \nsucceq \mathbf{y}$ means that $\mathbf{x} \ge \mathbf{y}$ does not hold
 - the same notations apply to matrices
- conventions:
 - ${\bf x}$ is said to be non-negative if ${\bf x} \geq {\bf 0},$ and non-positive if $-{\bf x} \geq {\bf 0}$
 - ${\bf x}$ is said to be positive if ${\bf x}>{\bf 0},$ and negative if $-{\bf x}>{\bf 0}$
 - the same conventions apply to matrices
 - a square A is said to be column-stochastic if $\mathbf{A} \geq \mathbf{0}$ and $\mathbf{A}^T \mathbf{1} = \mathbf{1}$
 - * a column-stochastic A has every column \mathbf{a}_i satisfying $\mathbf{a}_i^T \mathbf{1} = \sum_{j=1}^n a_{jj} = 1$

PageRank Matrix Properties

- $\bullet\,$ in PageRank, ${\bf A}$ is column-stochastic if all pages have outgoing links
 - see the literature to see how to deal with cases where some pages do not have outgoing links (dangling nodes)

Property 3.4. Let \mathbf{A} be column-stochastic. Then,

- 1. $\lambda = 1$ is an eigenvalue of ${f A}$
- 2. $|\lambda| \leq 1$ for any eigenvalue λ of \mathbf{A}
- Implications:
 - a solution to $\mathbf{A}\mathbf{v}=\mathbf{v}$ does exist, though it doesn't say if $\mathbf{v}\geq \mathbf{0}$ or not
 - $\lambda=1$ is an eigenvalue that has the largest modulus, but we don't know if it is the only eigenvalue that has the largest modulus
- we resort to non-negative matrix theory to answer the rest of the questions

Non-Negative Matrix Theory

Theorem 3.3 (Perron-Frobenius). Let A be square positive. There exists an eigenvalue ρ of A such that

- 1. ρ is real and $\rho > 0$
- 2. $|\lambda| < \rho$ for any eigenvalue λ of **A** with $\lambda \neq \rho$
- 3. there exists a positive eigenvector associated with ρ
- 4. the algebraic multiplicity of ρ is 1 (so the geometric multiplicity of ρ is also 1)

A weaker result for general non-negative matrices:

Theorem 3.4. Let A be square non-negative. There exists an eigenvalue ρ of A such that

- 1. ρ is real and $\rho \geq 0$
- 2. $|\lambda| \leq \rho$ for any eigenvalue λ of \mathbf{A}
- 3. there exists a non-negative eigenvector associated with ρ

PageRank Matrix Properties

- further implication by Theorem 3.4:
 - a non-negative solution to Av = v exists, though it doesn't say if there exists another solution
 - even worse, it is not known if there exists another solution ${f v}$ such that ${f v}
 eq 0$

PageRank Matrix Properties

 $\bullet\,$ PageRank actually considers a modified version of ${\bf A}$

$$\tilde{\mathbf{A}} = (1 - \beta)\mathbf{A} + \beta \begin{bmatrix} 1/n & \dots & 1/n \\ \vdots & & \vdots \\ 1/n & \dots & 1/n \end{bmatrix}$$

where $0 < \beta < 1$ (typical value is $\beta = 0.15$)

- $\tilde{\mathbf{A}}$ is positive
- further implications by Theorem 3.3:
 - $\lambda=1$ is the only eigenvalue that has the largest modulus
 - there exists only one eigenvector associated with $\lambda=1;$ that eigenvector is either positive or negative
 - so the power method should work

References

[Golub-Van Loan'12] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd edition, JHU Press, 2012.

[Bryan-Tanya2006] K. Bryan and L. Tanya, "The 25,000,000,000 eigenvector: The linear algebra behind Google," *SIAM Review*, vol. 48, no. 3, pp. 569–581, 2006.