§9.2 Orthogonal Matrices and Similarity Transformations

Def: A matrix $Q \in \mathbb{R}^{n\times n}$ is said to be orthogonal if its columns $\left\{\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \cdots, \mathbf{q}^{(n)}\right\}$ form an orthonormal set in $\mathbb{R}^n$.

§9.2 Orthogonal Matrices and Similarity Transformations

Def: A matrix $Q \in \mathbb{R}^{n\times n}$ is said to be orthogonal if its columns $\left\{\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \cdots, \mathbf{q}^{(n)}\right\}$ form an orthonormal set in $\mathbb{R}^n$.

Thm: Suppose matrix $Q \in \mathbb{R}^{n\times n}$ is orthogonal. Then

- Q is invertible with $Q^{-1} = Q^T$.
- For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\,\mathbf{x})^T (Q\,\mathbf{y}) = \mathbf{x}^T \mathbf{y}$.
- For any $\mathbf{x} \in \mathbb{R}^n$, $\|Q\,\mathbf{x}\|_2 = \|\mathbf{x}\|_2$.

§9.2 Orthogonal Matrices and Similarity Transformations

Def: A matrix $Q \in \mathbb{R}^{n\times n}$ is said to be orthogonal if its columns $\left\{\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \cdots, \mathbf{q}^{(n)}\right\}$ form an orthonormal set in $\mathbb{R}^n$.

Thm: Suppose matrix $Q \in \mathbb{R}^{n\times n}$ is orthogonal. Then

- Q is invertible with $Q^{-1} = Q^T$.
- For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $(Q\,\mathbf{x})^T (Q\,\mathbf{y}) = \mathbf{x}^T \mathbf{y}$.
- For any $\mathbf{x} \in \mathbb{R}^n$, $\|Q\,\mathbf{x}\|_2 = \|\mathbf{x}\|_2$.

Ex

$$H = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad H^T H = I.$$

Def: Two matrices A and B are SIMILAR if a nonsingular matrix S exists with $A = S^{-1} B S$.

Def: Two matrices A and B are SIMILAR if a nonsingular matrix S exists with $A = S^{-1}\,B\,S$.

Thm: Suppose A and B are similar matrices with $A = S^{-1}\,B\,S$ and λ is an eigenvalue of A with associated eigenvector $\mathbf{x}$. Then λ is an eigenvalue of B with associated eigenvector $S\,\mathbf{x}$.

Def: Two matrices A and B are SIMILAR if a nonsingular matrix S exists with $A = S^{-1} B S$.

Thm: Suppose A and B are similar matrices with $A = S^{-1} B S$ and λ is an eigenvalue of A with associated eigenvector $\mathbf{x}$. Then λ is an eigenvalue of B with associated eigenvector $S\,\mathbf{x}$.

Proof: Let $\mathbf{x} \neq \mathbf{0}$ be such that

$$A\,\mathbf{x} = \left(S^{-1}\,B\,S\right)\,\mathbf{x} = \lambda\,\mathbf{x}.$$

It follows that $B\,(S\,\mathbf{x}) = \lambda\,(S\,\mathbf{x})$

Thm: $A \in \mathbb{R}^n$ is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors.

Thm: $A \in \mathbb{R}^n$ is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors.

Proof:

$$A = S\,D\,S^{-1}, \ \left(S = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right), \ D = \textbf{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)\right)$$

$$\Longleftrightarrow \quad A\left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right) = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right) \textbf{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right),$$

$$\Longleftrightarrow \quad A\,\mathbf{v}^{(j)} = \lambda_j\,\mathbf{v}^{(j)}, \quad j = 1, 2, \cdots, n. \quad \mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)} \text{ L.I.D.}$$

Thm: $A \in \mathbb{R}^n$ is similar to a diagonal matrix D if and only if A has n linearly independent eigenvectors.

Proof:

$$A = S\,D\,S^{-1}, \;\left(S = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right),\; D = \mathbf{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)\right)$$

$$\iff \quad A\left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right) = \left(\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)}\right)\,\mathbf{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right),$$

$$\iff \quad A\,\mathbf{v}^{(j)} = \lambda_j\,\mathbf{v}^{(j)}, \quad j = 1, 2, \cdots, n. \quad \mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \cdots, \mathbf{v}^{(n)} \text{ L.I.D.}$$

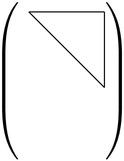
Cor: $A \in \mathbb{R}^n$ with n distinct eigenvalues is similar to diagonal matrix.

Def: *A matrix* $U \in \mathcal{C}^{n\times n}$ *is* UNITARY *if* $\|U\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ *for any vector* $\mathbf{x}$.

Def: *A matrix $U \in \mathcal{C}^{n\times n}$ is* UNITARY *if $\|U\,\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ for any vector $\mathbf{x}$.*

Schur Thm: *Let $A \in \mathbb{R}^n$. A* UNITARY *matrix U exists such that*

$$T = U^{-1}\,A\,U =$$



is upper-triangular.

The diagonal entries of T are the eigenvalues of A.

Def: The COMPLEX CONJUGATE of a complex vector $\mathbf{u} = \mathbf{a} + \sqrt{-1}\,\mathbf{b} \in \mathcal{C}^n$ is $\overline{\mathbf{u}} = \mathbf{a} - \sqrt{-1}\,\mathbf{b}$.

Def: The COMPLEX CONJUGATE of a complex vector $\mathbf{u} = \mathbf{a} + \sqrt{-1}\,\mathbf{b} \in \mathcal{C}^n$ is $\overline{\mathbf{u}} = \mathbf{a} - \sqrt{-1}\,\mathbf{b}$.

Def: $\|\mathbf{u}\|_2 = \sqrt{\mathbf{a}^T\mathbf{a} + \mathbf{b}^T\mathbf{b}} = \sqrt{\overline{\mathbf{u}}^T\mathbf{u}}$.

Def: The COMPLEX CONJUGATE of a complex vector $\mathbf{u} = \mathbf{a} + \sqrt{-1}\,\mathbf{b} \in \mathcal{C}^n$ is $\overline{\mathbf{u}} = \mathbf{a} - \sqrt{-1}\,\mathbf{b}$.

Def: $\|\mathbf{u}\|_2 = \sqrt{\mathbf{a}^T\mathbf{a} + \mathbf{b}^T\mathbf{b}} = \sqrt{\overline{\mathbf{u}}^T\mathbf{u}}$.

Thm: Let $A = A^T \in \mathbb{R}^n$ be symmetric. Then all eigenvalues of A are real.

Def: The COMPLEX CONJUGATE of a complex vector $\mathbf{u} = \mathbf{a} + \sqrt{-1}\,\mathbf{b} \in \mathcal{C}^n$ is $\overline{\mathbf{u}} = \mathbf{a} - \sqrt{-1}\,\mathbf{b}$.

Def: $\|\mathbf{u}\|_2 = \sqrt{\mathbf{a}^T\mathbf{a} + \mathbf{b}^T\mathbf{b}} = \sqrt{\overline{\mathbf{u}}^T\mathbf{u}}$.

Thm: Let $A = A^T \in \mathbb{R}^n$ be symmetric. Then all eigenvalues of A are real.

Proof: Let λ be an eigenvalue of A with eigenvector $\mathbf{u}$. Then $\overline{\lambda}$ is eigenvalue,

$$A\,\mathbf{u} = \lambda\,\mathbf{u}, \quad \longrightarrow A\,\overline{\mathbf{u}} = \overline{\lambda}\,\overline{\mathbf{u}}.$$

$$\begin{aligned} \lambda\left(\overline{\mathbf{u}}^T\mathbf{u}\right) &= \overline{\mathbf{u}}^T(A\,\mathbf{u}) \\ &= (A\,\overline{\mathbf{u}})^T\,\mathbf{u} = \overline{\lambda}\left(\overline{\mathbf{u}}^T\mathbf{u}\right). \end{aligned}$$

Therefore $\lambda = \overline{\lambda} \in \mathbb{R}$.

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & \ddots & \\ & & \ddots \end{pmatrix} Q^T.$$

Thm: A matrix $A \in \mathbb{R}^n$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & & \\ & & \end{pmatrix} Q^T.$$

Proof:

- By induction on n. Assume theorem true for $n-1$.
- Let λ be eigenvalue of A with UNIT eigenvector $\mathbf{u}$: $A\,\mathbf{u} = \lambda\,\mathbf{u}$.
- We extend $\mathbf{u}$ into an orthonormal basis for $\mathbb{R}^n$: $\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are unit, mutually orthogonal vectors.
- $$U \stackrel{def}{=} (\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n) \stackrel{def}{=} \left(\mathbf{u}, \widehat{U}\right) \in \mathbb{R}^{n\times n} \quad \text{is orthogonal.}$$

$$U^T A\,U = \begin{pmatrix} \mathbf{u}^T \\ \widehat{U}^T \end{pmatrix} \left(A\,\mathbf{u}, A\,\widehat{U}\right) = \begin{pmatrix} \mathbf{u}^T (A\,\mathbf{u}) & \mathbf{u}^T \left(A\,\widehat{U}\right) \\ \widehat{U}^T (A\,\mathbf{u}) & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}$$

$$= \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}.$$

- Matrix $\widehat{U}^T \left(A\,\widehat{U}\right) \in \mathbb{R}^{(n-1)\times(n-1)}$ is symmetric.

Thm: A matrix $A \in \mathbb{R}^{n \times n}$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & \ddots & \\ & & \ddots \end{pmatrix} Q^T.$$

Thm: A matrix $A \in \mathbb{R}^{n \times n}$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q so that $A = Q\,D\,Q^T = Q \begin{pmatrix} \diagdown \end{pmatrix} Q^T$.

Proof:

- $$U^T A\, U = \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}.$$
- By induction, there exist diagonal matrix $\widehat{D}$ and orthogonal matrix $\widehat{Q} \in \mathbb{R}^{(n-1)\times(n-1)}$,
$$\widehat{U}^T \left(A\,\widehat{U}\right) = \widehat{Q}\,\widehat{D}\,\widehat{Q}^T.$$
- therefore

$$U^T A\, U = \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{Q}\,\widehat{D}\,\widehat{Q}^T \end{pmatrix}.$$

$$A = \left(U \begin{pmatrix} 1 & \\ & \widehat{Q} \end{pmatrix}\right) \begin{pmatrix} \lambda & \\ & \widehat{D} \end{pmatrix} \left(U \begin{pmatrix} 1 & \\ & \widehat{Q} \end{pmatrix}\right)^T \stackrel{def}{=} Q\,D\,Q^T.$$

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric. Then A is positive definite if and only if all eigenvalues of A are positive.

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric. Then A is positive definite if and only if all eigenvalues of A are positive.

Proof:

- Let the diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q be so that $A = Q\, D\, Q^T$.
- $D = \mathbf{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$. $\lambda_1, \lambda_2, \cdots, \lambda_n$ eigenvalues of A.

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric. Then A is positive definite if and only if all eigenvalues of A are positive.

Proof:

- Let the diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q be so that $A = Q\,D\,Q^T$.
- $D = \mathbf{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$. $\lambda_1, \lambda_2, \cdots, \lambda_n$ eigenvalues of A.

$$
\begin{aligned}
& A \text{ is positive definite} \\
\iff \quad & \mathbf{x}^T A\,\mathbf{x} > 0 \quad \text{for any non-zero } \mathbf{x}
\end{aligned}
$$

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric. Then A is positive definite if and only if all eigenvalues of A are positive.

Proof:

- Let the diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q be so that $A = Q\,D\,Q^T$.
- $D = \mathbf{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$. $\lambda_1, \lambda_2, \cdots, \lambda_n$ eigenvalues of A.

$$\begin{aligned} & A \text{ is positive definite} \\ \Longleftrightarrow\quad & \mathbf{x}^T A\,\mathbf{x} > 0 \quad \text{for any non-zero } \mathbf{x} \\ \Longleftrightarrow\quad & \left(Q^T \mathbf{x}\right)^T D \left(Q^T \mathbf{x}\right) > 0 \quad \text{for any non-zero } \mathbf{x} \end{aligned}$$

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric. Then A is positive definite if and only if all eigenvalues of A are positive.

Proof:

- Let the diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q be so that $A = Q\,D\,Q^T$.
- $D = \textbf{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$. $\lambda_1, \lambda_2, \cdots, \lambda_n$ eigenvalues of A.

$$
\begin{aligned}
& A \text{ is positive definite} \\
\Longleftrightarrow \quad & \mathbf{x}^T A\,\mathbf{x} > 0 \quad \text{for any non-zero } \mathbf{x} \\
\Longleftrightarrow \quad & \left(Q^T \mathbf{x}\right)^T D \left(Q^T \mathbf{x}\right) > 0 \quad \text{for any non-zero } \mathbf{x} \\
\Longleftrightarrow \quad & \mathbf{y}^T D\,\mathbf{y} > 0 \quad \text{for any non-zero } \mathbf{y}
\end{aligned}
$$

Thm: Let matrix $A \in \mathbb{R}^{n \times n}$ be symmetric. Then A is positive definite if and only if all eigenvalues of A are positive.

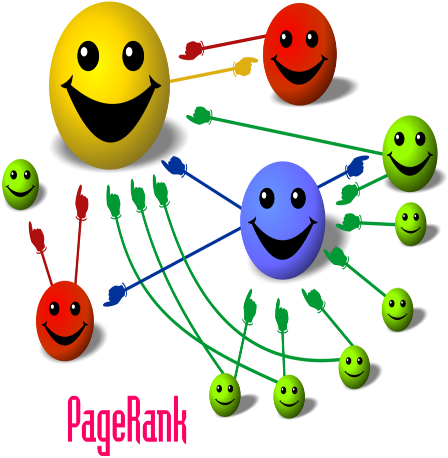
Proof:

- Let the diagonal matrix $D \in \mathbb{R}^{n \times n}$ and an orthogonal matrix Q be so that $A = Q\,D\,Q^T$.
- $D = \mathbf{diag}\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$. $\lambda_1, \lambda_2, \cdots, \lambda_n$ eigenvalues of A.

$$
\begin{aligned}
& A \text{ is positive definite} \\
\Longleftrightarrow\quad & \mathbf{x}^T A\,\mathbf{x} > 0 \quad \text{for any non-zero } \mathbf{x} \\
\Longleftrightarrow\quad & \left(Q^T\mathbf{x}\right)^T D\,\left(Q^T\mathbf{x}\right) > 0 \quad \text{for any non-zero } \mathbf{x} \\
\Longleftrightarrow\quad & \mathbf{y}^T D\,\mathbf{y} > 0 \quad \text{for any non-zero } \mathbf{y} \\
\Longleftrightarrow\quad & \text{diagonal entries of } D \text{ are positive.}
\end{aligned}
$$

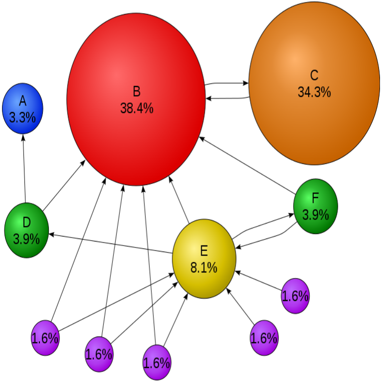
§9.3 The Power Method for Google PageRank (I)

- The PageRank Principle: The importance of each Webpage is proportional to the total size of the other Webpages which are pointing to it.



§9.3 The Power Method for Google PageRank (II)

- RANDOM SURF WITH JUMP: A Websurfer surfs the next Webpage
  - either jumping to a page chosen at random from the entire Web at 15% likelihood,
  - or choosing a random link from the Webpage at 85% likelihood.



§9.3 The Power Method for Google PageRank (III)

- Google Matrix G: each row/column represents a webpage, each G entry models Web connectivity and Web user surf patterns,
- PageRank vector $\mathbf{x}$ is eigenvector for G:

$$G\,\mathbf{x} = 1 \cdot \mathbf{x},$$

  where 1 is always a simple eigenvalue of G.
- Power Method for iteratively computing $\mathbf{x}$, given $\mathbf{x}^{(0)}$,

$$\mathbf{x}^{(k+1)} = G\,\mathbf{x}^{(k)}, \quad k = 0, 1, \cdots,$$

The Power Method, in general

Given: Matrix $A \in \mathbb{R}^{n\times n}$, with n eigenvalues

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| .$$

(A has precisely one eigenvalue, λ_1, that is largest in magnitude.)

The Power Method, in general

Given: Matrix $A \in \mathbb{R}^{n \times n}$, with n eigenvalues

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| .$$

(A has precisely one eigenvalue, λ_1, that is largest in magnitude.)

Task: Compute λ_1 and corresponding eigenvector $\mathbf{v}_1$.

The Power Method, in general

Given: Matrix $A \in \mathbb{R}^{n \times n}$, with n eigenvalues

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|.$$

(A has precisely one eigenvalue, λ_1, that is largest in magnitude.)

Task: Compute λ_1 and corresponding eigenvector $\mathbf{v}_1$.

Despite condition on λ_1, PM usually first method to try.

The Power Method

- Assume $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are eigenvectors pertaining to $\lambda_1, \lambda_2, \cdots, \lambda_n$.
- Given initial vector $\mathbf{x} \neq \mathbf{0}$. Then

$$\mathbf{x} = \sum_{j=1}^{n} \beta_j \, \mathbf{v}_j$$

for some coefficients $\beta_1, \cdots, \beta_n$. $\boxed{\text{Assume } \beta_1 \neq 0.}$

The Power Method

- Assume $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are eigenvectors pertaining to $\lambda_1, \lambda_2, \cdots, \lambda_n$.
- Given initial vector $\mathbf{x} \neq \mathbf{0}$. Then

$$\mathbf{x} = \sum_{j=1}^{n} \beta_j \, \mathbf{v}_j$$

  for some coefficients $\beta_1, \cdots, \beta_n$. $\boxed{\text{Assume } \beta_1 \neq 0.}$
- For any $k > 0$

$$\begin{aligned} A^k \, \mathbf{x} &= \sum_{j=1}^{n} \beta_j \, \lambda_j^k \, \mathbf{v}_j \\ &= \beta_1 \lambda_1^k \left(\mathbf{v}_1 + \sum_{j=2}^{n} \left(\frac{\lambda_j}{\lambda_1} \right)^k \left(\frac{\beta_j}{\beta_1} \mathbf{v}_j \right) \right) \\ &= \beta_1 \lambda_1^k \left(\mathbf{v}_1 + O\left(\left(\frac{\lambda_2}{\lambda_1} \right)^k \right) \right) \end{aligned}$$

The Power Method

- Assume $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are eigenvectors pertaining to $\lambda_1, \lambda_2, \cdots, \lambda_n$.
- Given initial vector $\mathbf{x} \neq \mathbf{0}$. Then

$$\mathbf{x} = \sum_{j=1}^{n} \beta_j \, \mathbf{v}_j$$

  for some coefficients $\beta_1, \cdots, \beta_n$. $\boxed{\text{Assume } \beta_1 \neq 0.}$
- For any $k > 0$

$$\begin{aligned} A^k \, \mathbf{x} &= \sum_{j=1}^{n} \beta_j \, \lambda_j^k \, \mathbf{v}_j \\ &= \beta_1 \lambda_1^k \left(\mathbf{v}_1 + \sum_{j=2}^{n} \left(\frac{\lambda_j}{\lambda_1} \right)^k \left(\frac{\beta_j}{\beta_1} \mathbf{v}_j \right) \right) \\ &= \beta_1 \lambda_1^k \left(\mathbf{v}_1 + O\left(\left(\frac{\lambda_2}{\lambda_1} \right)^k \right) \right) \end{aligned}$$

Rayleigh quotient

Given: *Approximate eigenvector* $\mathbf{x}$.

Rayleigh quotient

Given: *Approximate eigenvector* $\mathbf{x}$.
Task: *Find approximate eigenvalue* λ.
LS for λ: *Choose* λ *in LS sense*

$$\mathbf{min}_{\lambda} \|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 \,.$$

LS Solution:

$$\lambda = \frac{\mathbf{x}^T (A\,\mathbf{x})}{\mathbf{x}^T\,\mathbf{x}}$$

Algorithm 1 The Power Method

Input: Matrix $A \in \mathbb{R}^{n \times n}$,
initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^n$, and tolerance $\tau > 0$.

Output: Approximate eigenvalue λ, eigenvector $\mathbf{x}$.

Algorithm:

Normalize: $\mathbf{x}^{(0)} = \mathbf{x}^{(0)} / \left\|\mathbf{x}^{(0)}\right\|_2$, $\mathbf{y}^{(0)} = A\,\mathbf{x}^{(0)}$, $k = 0$.

$\lambda = \left(\mathbf{x}^{(0)}\right)^T \mathbf{y}^{(0)}$.

while $\left\|\mathbf{y}^{(k)} - \lambda\,\mathbf{x}^{(k)}\right\|_2 \geq \tau$ **do**

  $\mathbf{x}^{(k+1)} = \mathbf{y}^{(k)} / \left\|\mathbf{y}^{(k)}\right\|_2$, $\mathbf{y}^{(k+1)} = A\,\mathbf{x}^{(k+1)}$.

  $\lambda = \left(\mathbf{x}^{(k+1)}\right)^T \mathbf{y}^{(k+1)}$.

  $k = k + 1$.

end while

Algorithm 2 The Symmetric Power Method

Input: Symmetric matrix $A \in \mathbb{R}^{n \times n}$,
initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^n$, and tolerance $\tau > 0$.

Output: Approximate eigenvalue λ, eigenvector $\mathbf{x}$.

Algorithm:

Normalize: $\mathbf{x}^{(0)} = \mathbf{x}^{(0)} \Big/ \left\|\mathbf{x}^{(0)}\right\|_2$, $\mathbf{y}^{(0)} = A\,\mathbf{x}^{(0)}$, $k = 0$.

$\lambda = \left(\mathbf{x}^{(0)}\right)^T \mathbf{y}^{(0)}$.

while $\left\|\mathbf{y}^{(k)} - \lambda\,\mathbf{x}^{(k)}\right\|_2 \geq \tau$ **do**

  $\mathbf{x}^{(k+1)} = \mathbf{y}^{(k)} \Big/ \left\|\mathbf{y}^{(k)}\right\|_2$, $\mathbf{y}^{(k+1)} = A\,\mathbf{x}^{(k+1)}$.

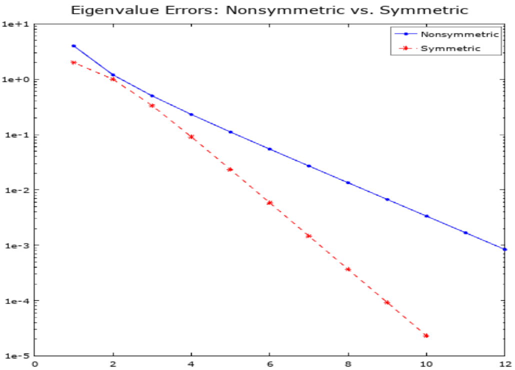
  $\lambda = \left(\mathbf{x}^{(k+1)}\right)^T \mathbf{y}^{(k+1)}$.

  $k = k + 1$.

end while

Same PM, but Symmetric PM converges much faster.

- **Ex 1:** $A = \begin{pmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{pmatrix}$ with $\mathbf{x}^{0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ for $\lambda_1 = 6$.
- **Ex 2:** $A = \begin{pmatrix} -4 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 3 \end{pmatrix}$ with $\mathbf{x}^{0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ for $\lambda_1 = 6$.



Thm: Let $A \in \mathbb{R}^{n\times n}$ is symmetric with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. If we have $\|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 \le \tau$ for some real number λ and unit vector $\mathbf{x}$, then

$$\mathbf{min}_{1\le j\le n} |\lambda - \lambda_j| \le \tau.$$

Thm: Let $A \in \mathbb{R}^{n\times n}$ is symmetric with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. If we have $\|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 \le \tau$ for some real number λ and unit vector $\mathbf{x}$, then

$$\mathbf{min}_{1\le j\le n} |\lambda - \lambda_j| \le \tau.$$

Proof: Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ form an orthonormal set of A eigenvectors associated with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then the matrix $Q \stackrel{def}{=} (\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)$ is orthogonal, and

$$\mathbf{x} = \beta_1\,\mathbf{v}_1 + \beta_2\,\mathbf{v}_2 + \cdots + \beta_n\,\mathbf{v}_n$$

$$\text{with} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \stackrel{def}{=} Q^T\,\mathbf{x} \quad \text{unit vector.}$$

Thm: Let $A \in \mathbb{R}^{n\times n}$ is symmetric with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. If we have $\|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 \le \tau$ for some real number λ and unit vector $\mathbf{x}$, then

$$\mathbf{min}_{1\le j\le n} |\lambda - \lambda_j| \le \tau.$$

Proof: Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ form an orthonormal set of A eigenvectors associated with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then the matrix $Q \stackrel{def}{=} (\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)$ is orthogonal, and

$$\mathbf{x} = \beta_1\,\mathbf{v}_1 + \beta_2\,\mathbf{v}_2 + \cdots + \beta_n\,\mathbf{v}_n$$

$$\text{with} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \stackrel{def}{=} Q^T\,\mathbf{x} \quad \text{unit vector.}$$

$$\begin{aligned} \|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 &= \|\beta_1\,(\lambda_1 - \lambda)\,\mathbf{v}_1 + \cdots + \beta_n\,(\lambda_n - \lambda)\,\mathbf{v}_n\|_2 \\ &= \sqrt{\beta_1^2\,(\lambda_1 - \lambda)^2 + \cdots + \beta_n^2\,(\lambda_n - \lambda)^2} \\ &\ge \left(\mathbf{min}_{1\le j\le n} |\lambda - \lambda_j|\right) \sqrt{\beta_1^2 + \cdots + \beta_n^2} \end{aligned}$$

Thm: Let $A \in \mathbb{R}^{n\times n}$ is symmetric with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. If we have $\|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 \le \tau$ for some real number λ and unit vector $\mathbf{x}$, then

$$\mathbf{min}_{1\le j\le n} |\lambda - \lambda_j| \le \tau.$$

Proof: Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ form an orthonormal set of A eigenvectors associated with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then the matrix $Q \stackrel{def}{=} (\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)$ is orthogonal, and

$$\mathbf{x} = \beta_1\,\mathbf{v}_1 + \beta_2\,\mathbf{v}_2 + \cdots + \beta_n\,\mathbf{v}_n$$

$$\text{with} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \stackrel{def}{=} Q^T\,\mathbf{x} \quad \text{unit vector.}$$

$$\begin{aligned}
\|A\,\mathbf{x} - \lambda\,\mathbf{x}\|_2 &= \|\beta_1\,(\lambda_1 - \lambda)\,\mathbf{v}_1 + \cdots + \beta_n\,(\lambda_n - \lambda)\,\mathbf{v}_n\|_2 \\
&= \sqrt{\beta_1^2\,(\lambda_1 - \lambda)^2 + \cdots + \beta_n^2\,(\lambda_n - \lambda)^2} \\
&\ge \left(\mathbf{min}_{1\le j\le n} |\lambda - \lambda_j|\right)\sqrt{\beta_1^2 + \cdots + \beta_n^2} = \mathbf{min}_{1\le j\le n} |\lambda - \lambda_j|
\end{aligned}$$

The Inverse Power Method (I)

Given: Matrix $A \in \mathbb{R}^{n\times n}$, with n eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$; and given SHIFT q.

Task: Compute λ_i that is closest to q, and corresponding eigenvector $\mathbf{v}_i$.

The Inverse Power Method (I)

Given: Matrix $A \in \mathbb{R}^{n \times n}$, with n eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$; and given SHIFT q.

Task: Compute λ_i that is closest to q, and corresponding eigenvector $\mathbf{v}_i$.

Apply Power Method to $(A - q\,I)^{-1}$.

- Matrix $(A - q\,I)^{-1}$ has eigenvalues

$$\frac{1}{\lambda_1 - q}, \frac{1}{\lambda_2 - q}, \cdots, \frac{1}{\lambda_n - q}.$$

- Matrix $(A - q\,I)^{-1}$ has eigenvalues

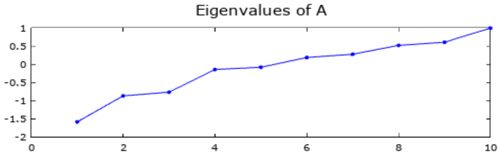
$$\frac{1}{\lambda_1 - q}, \frac{1}{\lambda_2 - q}, \cdots, \frac{1}{\lambda_n - q}.$$

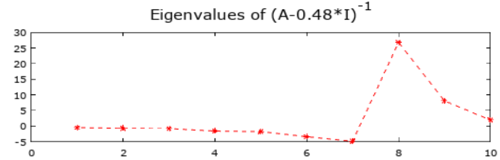
- Assume q closest to λ_i and λ_k, but closer to λ_i.

- Matrix $(A - q\,I)^{-1}$ has eigenvalues

$$\frac{1}{\lambda_1 - q}, \frac{1}{\lambda_2 - q}, \dots, \frac{1}{\lambda_n - q}.$$

- Assume q closest to λ_i and λ_k, but closer to λ_i.
- IPM converges to λ_i at order $\left(\frac{\lambda_i - q}{\lambda_k - q}\right)^k$.





Algorithm 3 The Inverse Power Method

Input: Matrix $A \in \mathbb{R}^{n \times n}$, shift q,
initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^n$, and tolerance $\tau > 0$.

Output: Approximate eigenvalue λ, eigenvector $\mathbf{x}$.

Algorithm:

Normalize: $\mathbf{x}^{(0)} = \mathbf{x}^{(0)} / \left\| \mathbf{x}^{(0)} \right\|_2$, $\mathbf{y}^{(0)} = (A - q\,I)^{-1}\,\mathbf{x}^{(0)}$.

$\lambda = \left(\mathbf{x}^{(0)}\right)^T \mathbf{y}^{(0)}$, $k = 0$.

while $\left\| \mathbf{y}^{(k)} - \lambda\,\mathbf{x}^{(k)} \right\|_2 \geq \tau$ **do**

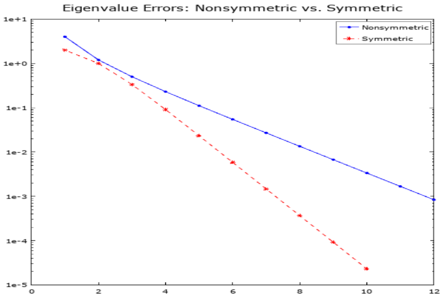
$\mathbf{x}^{(k+1)} = \mathbf{y}^{(k)} / \left\| \mathbf{y}^{(k)} \right\|_2$, $\mathbf{y}^{(k+1)} = (A - q\,I)^{-1}\,\mathbf{x}^{(k+1)}$.

$\lambda = \left(\mathbf{x}^{(k+1)}\right)^T \mathbf{y}^{(k+1)}$.

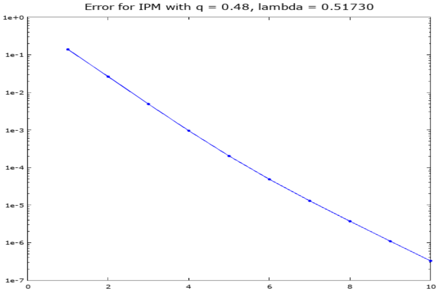
$k = k + 1$.

end while

- Symmetric/Non-symmetric PM Errors



- Symmetric IPM Errors



Review

Thm: A matrix $A \in \mathbb{R}^{n \times n}$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^n$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \ddots & & \\ & & \\ & & \end{pmatrix} Q^T.$$

Review

Thm: A matrix $A \in \mathbb{R}^{n\times n}$ is symmetric if and only if there exists a diagonal matrix $D \in \mathbb{R}^{n}$ and an orthogonal matrix Q so that

$$A = Q\,D\,Q^T = Q \begin{pmatrix} \diagdown & \\ & \end{pmatrix} Q^T.$$

Proof:

- Let λ be eigenvalue of A with UNIT eigenvector $\mathbf{u}$: $A\,\mathbf{u} = \lambda\,\mathbf{u}$.
- We extend $\mathbf{u}$ into an orthonormal basis for $\mathbb{R}^n$: $\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are unit, mutually orthogonal vectors.
- $U \stackrel{def}{=} (\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n) \stackrel{def}{=} \left(\mathbf{u}, \widehat{U}\right) \in \mathbb{R}^{n\times n}$ is orthogonal.

$$U^T A\, U = \begin{pmatrix} \mathbf{u}^T \\ \widehat{U}^T \end{pmatrix} \left(A\,\mathbf{u}, A\,\widehat{U}\right) = \begin{pmatrix} \mathbf{u}^T (A\,\mathbf{u}) & \mathbf{u}^T \left(A\,\widehat{U}\right) \\ \widehat{U}^T (A\,\mathbf{u}) & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}$$

$$= \begin{pmatrix} \lambda & \mathbf{0}^T \\ \mathbf{0} & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}.$$

- Repeat on symmetric matrix $\widehat{U}^T \left(A\,\widehat{U}\right) \in \mathbb{R}^{(n-1)\times(n-1)}$.

Computing all eigenvalues of matrix $A \in \mathbb{R}^{n \times n}$

Computing all eigenvalues of matrix $A \in \mathbb{R}^{n\times n}$

- Compute one approximate eigenvalue λ of A with UNIT eigenvector $\mathbf{u}$: $A\,\mathbf{u} = \lambda\,\mathbf{u}$.
- Extend $\mathbf{u}$ into an orthonormal basis for $\mathbb{R}^n$: $\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n$ are unit, mutually orthogonal vectors.
- $$U \overset{def}{=} (\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n) \overset{def}{=} \left(\mathbf{u}, \widehat{U}\right) \in \mathbb{R}^{n\times n} \quad \text{is orthogonal.}$$

$$\begin{aligned} U^T A\, U &= \begin{pmatrix} \mathbf{u}^T \\ \widehat{U}^T \end{pmatrix} \left(A\,\mathbf{u}, A\,\widehat{U}\right) = \begin{pmatrix} \mathbf{u}^T (A\,\mathbf{u}) & \mathbf{u}^T \left(A\,\widehat{U}\right) \\ \widehat{U}^T (A\,\mathbf{u}) & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix} \\ &= \begin{pmatrix} \lambda & \mathbf{u}^T \left(A\,\widehat{U}\right) \\ \mathbf{0} & \widehat{U}^T \left(A\,\widehat{U}\right) \end{pmatrix}. \quad (\text{DEFLATION}) \end{aligned}$$

- Continue on matrix $\widehat{A} \overset{def}{=} \widehat{U}^T \left(A\,\widehat{U}\right) \in \mathbb{R}^{(n-1)\times(n-1)}$.

Householder Reflection

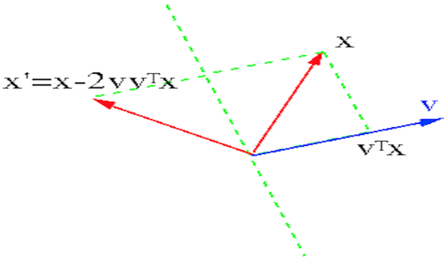
Let $\mathbf{v} \in \mathbb{R}^n$ be a unit vector. Define Householder Reflection matrix

$$H = I - 2\mathbf{v}\,\mathbf{v}^T \in \mathbb{R}^{n\times n}.$$

- H is symmetric and orthogonal

$$H = H^T, \quad H^2 = I - 4\mathbf{v}\,\mathbf{v}^T + 4\mathbf{v}\,\mathbf{v}^T = I.$$

- For any vector $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^\dagger \stackrel{def}{=} H\mathbf{x} = \mathbf{x} - 2\mathbf{v}\,\mathbf{v}^T\mathbf{x}$ reflects $\mathbf{x}$ in the direction $\mathbf{v}^\perp$:



Deflation with Householder Reflection (I)

- Given eigenvalue λ of A with UNIT eigenvector $\mathbf{u}$: $A\,\mathbf{u} = \lambda\,\mathbf{u}$.
- Extend $\mathbf{u}$ into an orthonormal basis with a Householder reflection
- $$U = I - 2\mathbf{v}\,\mathbf{v}^T \stackrel{def}{=} (\mathbf{u}, \mathbf{u}_2, \cdots, \mathbf{u}_n) \stackrel{def}{=} \left(\mathbf{u}, \widehat{U}\right)$$

$$U^T A\,U \;=\; \begin{pmatrix} \lambda & \mathbf{u}^T\left(A\,\widehat{U}\right) \\ \mathbf{0} & \widehat{U}^T\left(A\,\widehat{U}\right) \end{pmatrix}.$$

Find unit vector $\mathbf{v}$ so first column of $I - 2\mathbf{v}\,\mathbf{v}^T$ is $\mathbf{u}$.

Deflation with Householder Reflection (II)

- Partition
$$\mathbf{u} = \begin{pmatrix} \mu \\ \widehat{\mathbf{u}} \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} \nu \\ \widehat{\mathbf{v}} \end{pmatrix}.$$
- First column of $I - 2\mathbf{v}\,\mathbf{v}^T$ is
$$\begin{pmatrix} \mu \\ \widehat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} - 2\begin{pmatrix} \nu \\ \widehat{\mathbf{v}} \end{pmatrix} \nu.$$
- If $\mu \leq 0$, then
$$\nu = \sqrt{\frac{1-\mu}{2}}, \quad \widehat{\mathbf{v}} = -\frac{\widehat{\mathbf{u}}}{2\,\nu}, \quad U = I - 2\mathbf{v}\,\mathbf{v}^T = \left(\mathbf{u}, \widehat{U}\right). \quad (1)$$
- If $\mu > 0$, then $-\mathbf{u}$ is also unit eigenvector. Compute $\mathbf{v}$ with equation (1) on $-\mathbf{u}$:
$$\nu = \sqrt{\frac{1+\mu}{2}}, \quad \widehat{\mathbf{v}} = \frac{\widehat{\mathbf{u}}}{2\,\nu}, \quad U = I - 2\mathbf{v}\,\mathbf{v}^T = \left(-\mathbf{u}, \widehat{U}\right). \quad (2)$$

Equations (1) and (2) ensure numerical stability

Householder Reflection (II)

Let $\mathbf{v} \in \mathbb{R}^n$ be a unit vector. Define Householder Reflection matrix

$$H = I - 2\,\mathbf{v}\,\mathbf{v}^T \in \mathbb{R}^{n \times n}.$$

- For any vector $\mathbf{x} \in \mathbb{R}^n$, choose $\mathbf{v}$ so that

  $$H\,\mathbf{x} = \begin{pmatrix} \pm\,\|\mathbf{x}\|_2 \\ \mathbf{0} \end{pmatrix}, \quad \text{(sign to be chosen for numerical stability.)}$$

- Partition $\mathbf{x} = \begin{pmatrix} \xi \\ \widehat{\mathbf{x}} \end{pmatrix}$, $\begin{pmatrix} \pm\,\|\mathbf{x}\|_2 \\ \mathbf{0} \end{pmatrix} = H\,\mathbf{x} = \begin{pmatrix} \xi \\ \widehat{\mathbf{x}} \end{pmatrix} - 2\,\mathbf{v}\,\mathbf{v}^T\mathbf{x}$,

  $$\mathbf{u} \stackrel{def}{=} \begin{pmatrix} \pm\,\|\mathbf{x}\|_2 - \xi \\ -\widehat{\mathbf{x}} \end{pmatrix} \stackrel{choose}{=\!=\!=} -\begin{pmatrix} \mathbf{sign}\,(\xi)\,(\|\mathbf{x}\|_2 + |\xi|) \\ \widehat{\mathbf{x}} \end{pmatrix}$$

  and $\mathbf{v} = \mathbf{u}\,/\|\mathbf{u}\|_2$