Modified on Wednesday 7th October, 2020, 11:11:5111:11

Oct 7, The power method. [5.4] Fundamental spaces of a matrix, rank. [7.3, 7.4, 7.5]

- For any $n \times n$ matrix A, $\{I_n, A, A^2, \dots\}$ is linearly dependent. There is a polynomial $p(t)$ such that $p(A) = \mathbf{0}$. Let $m_A(t)$ be the monic polynomial in t of the smallest degree that satisfies $m_A(A) = \mathbf{0}$. $m_A(t)$ is called the **minimal polynomial** of A. Note that $m_A(t)$ divides the characteristic polynomial $p_A(t)$, by the Cayley-Hamilton theorem.
- Do the following matrices have the same minimal polynomials?

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The above matrices have the same characteristic polynomials, $(t-1)^2$. But their minimial polynomials are $t-1$ and $(t-1)^2$, respectively.

- For a square matrix A, we can define

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

If P is an invertible matrix, we have $e^{P^{-1}AP} = P^{-1}e^A P$. Hence for a diagonalizable matrix of A, $P^{-1}AP = D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, we have a very simple formula

$$e^A = e^{PDP^{-1}} = Pe^D P^{-1} = P \operatorname{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}) P^{-1}$$

- Let A be a real symmetric matrix of size $n \times n$, with $n = 10^5$. We know that it is orthogonally diagonalizable and has a spectral decomposition

$$A = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T + \cdots + \lambda_n \mathbf{v}_n \mathbf{v}_n^T.$$

How can we find the eigenvalues and eigenvectors?

In application, we may just want the largest eigenvalue and its eigenvectors.

- Dominant eigenvalue
Let A be a symmetric matrix. Let $\lambda_1, \dots, \lambda_k$ be all distinct eigenvalues of A arranged in the order of absolute values,

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_k|.$$

If $|\lambda_1| > |\lambda_2|$, then λ_1 is called a **dominant eigenvalue** of A and any eigenvector corresponding to a dominant eigenvalue is called a **dominant eigenvector** of A.

$$|\lambda_1| > \lambda_2| \geq \cdots \geq |\lambda_k|.$$

- **The power method.** Let A be a symmetric $n \times n$ matrix and $\lambda > 0$ is a dominant eigenvalue. If a unit vector $\mathbf{x}_0 \in \mathbb{R}^n$ is not orthogonal to the eigenspace corresponding to λ, then the sequence

$$\mathbf{x}_0,\ \mathbf{x}_1 = \frac{A\mathbf{x}_0}{\|A\mathbf{x}_0\|},\ \mathbf{x}_2 = \frac{A\mathbf{x}_1}{\|A\mathbf{x}_1\|},\ \dots,\ \mathbf{x}_k = \frac{A\mathbf{x}_{k-1}}{\|A\mathbf{x}_{k-1}\|},\ \dots$$

converges to a unit dominant eigenvector, and the sequence

$$\lambda^{(1)} = A\mathbf{x}_1 \cdot \mathbf{x}_1,\ \lambda^{(2)} = A\mathbf{x}_2 \cdot \mathbf{x}_2,\ \dots,\ \lambda^{(k)} = A\mathbf{x}_k \cdot \mathbf{x}_k, \dots$$

converges to the dominant eigenvalue λ.

- Relative error: $\left|\frac{\lambda - \lambda^{(k)}}{\lambda}\right|$, estimated relative error: $\left|\frac{\lambda^{(k)} - \lambda^{(k-1)}}{\lambda^{(k)}}\right|$,

- Application of the power method to internet searches: Consider an adjacency matrix

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

  where $a_{13} = a_{14} = 1$ implies that site 1 refereces sites 3 and 4. The site may be a hub, meaning that it refereces many other sites, or it may be an authority, meaning that it is refereced by many other sites. The initial hub and authority vectors for the adjacency matrix A is

$$\mathbf{h}_0 = A \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \ \mathbf{a}_0 = A^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \\ 2 \end{bmatrix}.$$

  How can we further incorporate the weights of the sites?

$$\mathbf{h}_1 = \frac{A\mathbf{a}_0}{||A\mathbf{a}_0||} \approx \begin{bmatrix} 0.43 \\ 0.32 \\ 0.54 \\ 0.65 \end{bmatrix}, \ \mathbf{a}_1 = \frac{A^T\mathbf{h}_1}{||A^T\mathbf{h}_1||} \approx \begin{bmatrix} 0.69 \\ 0.30 \\ 0.49 \\ 0.44 \end{bmatrix}, \ \mathbf{h}_2 = \frac{A\mathbf{a}_1}{||A\mathbf{a}_1||}, \mathbf{a}_2 = \frac{A^T\mathbf{h}_2}{||A^T\mathbf{h}_2||}, \dots,$$

$$\dots, \mathbf{h}_k = \frac{(AA^T)\mathbf{h}_{k-1}}{||(AA^T)\mathbf{h}_{k-1}||}, \ \mathbf{a}_k = \frac{(A^TA)\mathbf{a}_{k-1}}{||(A^TA)\mathbf{a}_{k-1}||}, \dots$$

- The fundamental spaces of an $m \times n$ matrix A
  - The **row space** of A, row(A), is the subspace of $\mathbb{R}^n$ spanned by the rows of A.
  - The **column space** of A, col(A), is the subspace of $\mathbb{R}^m$ spanned by the columns of A.
  - The **null space** of A, null(A), is the subspace of $\mathbb{R}^n$ consisting of solutions to $A\mathbf{x} = \mathbf{0}$.
  - The **null space** of A^T, null(A^T), is the subspace of $\mathbb{R}^m$ consisting of solutions to $A^T\mathbf{x} = \mathbf{0}$.
- The dimension of the row space of A is called the **rank** of A, denoted by $\mathrm{rank}(A) = \dim \mathrm{row}(A)$.
- The dimension of the null space of A is called the **nullity** of A, denoted by nullity(A).
- For a nonempty subset S of $\mathbb{R}^n$, the **orthogonal complement** of S, denoted by $S^\perp$, is defined to be the set of all vectors in $\mathbb{R}^n$ that are orthogonal to every element of S.
- For any nonempty subset S of $\mathbb{R}^n$, $S^\perp$ is a subspace of $\mathbb{R}^n$.
- $\{\mathbf{0}\}^\perp = \mathbb{R}^n$, $(\mathbb{R}^n)^\perp = \{\mathbf{0}\}$.
- Theorem
  - If W is a subspace of $\mathbb{R}^n$, then $W^\perp \cap W = \{\mathbf{0}\}$ and $(W^\perp)^\perp = W$.
  - If S is a nonempty subset of $\mathbb{R}^n$, then $S^\perp = \mathrm{span}(S)^\perp$.
- Relations between fundamental spaces of an $m \times n$ matrix A
  - $\mathrm{row}(A)^\perp = \mathrm{null}(A)$, $\mathrm{null}(A)^\perp = \mathrm{row}(A)$
  - $\mathrm{col}(A)^\perp = \mathrm{null}(A^T)$, $\mathrm{null}(A^T)^\perp = \mathrm{col}(A)$
- Theorem
  - If $A \sim_r B$, then $\mathrm{row}(A) = \mathrm{row}(B)$ and $\mathrm{null}(A) = \mathrm{null}(B)$.

  - If $A \sim_r B$ and B is in row echelon form, then nonzero rows of B form a basis for $\text{row}(A)$.

- Find bases of fundamental spaces of A.

$$A = \begin{bmatrix} 2 & 2 & 4 & 4 \\ 3 & 3 & 6 & 6 \\ 3 & 3 & 6 & 8 \\ 0 & 2 & 2 & 2 \end{bmatrix} \to \quad \cdots \quad \to \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Theorem (Dimension theorem for matrices) If A is an $m \times n$ matrix, then

$$\text{rank}(A) + \text{nullity}(A) = n.$$

- Theorem (Dimension theorem for subspaces) If W is a subspace of $\mathbb{R}^n$, then

$$\dim(W) + \dim(W^{\perp}) = n.$$

- If $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$ are nonzero vectors, then $A = \mathbf{u}\mathbf{v}^T$ is an $m \times n$ matrix of rank 1. The converse holds as well.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

- For any nonzero $\mathbf{v} \in \mathbb{R}^n$, $\mathbf{v}\mathbf{v}^T$ is symmetric and of rank 1. If an $n \times n$ matrix is symmetric and of rank 1, then $A = \mathbf{v}\mathbf{v}^T$ or $A = -\mathbf{v}\mathbf{v}^T$ for some $\mathbf{v} \in \mathbb{R}^n$.

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad -\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}^T = -\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = -\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

- Theorem (The Rank Theorem) For any matrix A, $\dim \text{row}(A) = \dim \text{col}(A)$.
- A, $A^T A$ and AA^T all have the same rank.