Recitation 6

Alex Dong

CDS, NYU

Fall 2020

Stochastic Processes Rabbit Hole

- Markov Chains are a topic in *stochastic processes*
- Stochastic Processes (&)
    - Finite Markov Chains (Discrete time, Discrete space)
    - Infinite Markov Chains (Discrete time, inf. discrete space)
    - Poisson Process (Continuous time, Discrete space)
    - Brownian Motion (Continuous time, continuous space)
- Main Assumption: Markov Property (Memoryless)
- Lots of Linear Algebra!

Questions: Stochastic Matrices

Let $A, B \in \mathbb{R}^{n \times n}$ be stochastic matrices. True or False for 1,2,3.

1. A is always invertible
2. The eigenvector corresponding to the largest eigenvalue of A is unique
3. A cannot have zero has its eigenvalue
4. Prove that AB is a stochastic matrix.

Hint for 4. Is there a way express the "sum of each column is 1" property as a matrix multiplication?

Solutions 1: Stochastic Matrices

Let A be a stochastic matrix. True or False for 1,2,3.

Solution

1. *A is always invertible*

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

2. *The eigenvector corresponding to the largest eigenvalue of A is unique. False* $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

3. *A cannot have zero has its eigenvalue. False:* $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$

Solutions 2: Stochastic Matrices

4. Prove that the product of two stochastic matrices is a stochastic matrix.

Solution

Let A, B be stochastic matrices in $\mathbb{R}^n$.
Each entry is non-negative:
$AB_{i,j} = \sum_{k=1}^{n} A_{i,k} B_{k,j}$
This summation is a sum of non-negative products, hence it is also non-negative.
Sum of each columns is 1:
Note that the property of each column summing to 1 can be seen as:
A matrix A is stochastic when
$\begin{bmatrix} 1 & \dots & 1 \end{bmatrix} A = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}$.
Then
$\begin{bmatrix} 1 & \dots & 1 \end{bmatrix} AB = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} A = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix}$.
So AB is stochastic.

Change of Basis

- Sometimes, a matrix A 'prefers' certain directions (*eigenvectors*)
- (!!!) These directions act as *anchors* for understanding the action of a matrix.
- These are directions that we will *orient* or *change our basis* to.
- This is related to $P^{-1}DP$, or diagonalization (Lec 7).
  - Defined as $P^{-1}DP$ or PDP^{-1}, depending on which text/notes you reference.

Question: Change of Basis

Let A have eigenvectors $v_1, ..., v_n$ with corresponding eigenvalues $\lambda_1, ..., \lambda_n$. Let $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis of $\mathbb{R}^n$. Let $x = \sum_{i=1}^{n} \alpha_i v_i$

1. Let P be a linear transformation that maps (canoncial basis vectors) e_i to v_i, for all $i \in 1, ..., n$. Write the matrix P.
2. What is $PDP^{-1}x$?
3. Let $k \in \mathbb{N}$. What is $(PDP^{-1})^k x$?
4. If $A = PDP^{-1}$, give an interpretation for the action of A.

Solutions 1: Change of Basis

Let A have eigenvectors $v_1, ..., v_n$ with corresponding eigenvalues $\lambda_1, ..., \lambda_n$. Let $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis of $\mathbb{R}^n$. Let $x = \sum_{i=1}^n \alpha_i v_i$

Solution

1. *Let P be a linear transformation that maps (canoncial basis vectors) e_i to v_i, for all $i \in 1, ..., n$. Write the matrix P.*

$$P = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix}$$

2. *Let $D = Diag(\lambda_1, ..., \lambda_n)$. What is $PDP^{-1}x$?*

$$P^{-1}x = P^{-1}\sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \alpha_i e_i$$
$$DP^{-1}x = D(\sum_{i=1}^n \alpha_i e_i) = \sum_{i=1}^n \lambda_i \alpha_i e_i$$
$$PDP^{-1}x = P(\sum_{i=1}^n \lambda_i \alpha_i e_i) = \sum_{i=1}^n \lambda_i \alpha_i v_i$$

Solutions 2: Change of Basis

Let A have eigenvectors $v_1, ..., v_n$ with corresponding eigenvalues $\lambda_1, ..., \lambda_n$. Let $\mathcal{B} = \{v_1, ..., v_n\}$ be a basis of $\mathbb{R}^n$. Let $x = \sum_{i=1}^{n} \alpha_i v_i$

Solution

3. *Let $k \in \mathbb{N}$ What is $(PDP^{-1})^k x$?*

   $$(PDP^{-1})^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$

   Likewise,

   $$(PDP^{-1})^k = PD^kP^{-1}$$

4. *If $A = PDP^{-1}$, give an interpretation for the action of A on a vector x.*

   First view x as a vector in the coordinates of basis $v_1, ..., v_n$.
   P^{-1} transforms these coordinates into the standard basis.
   D stretches the new coordinate e_i by the eigenvalue λ_i.
   P transforms these coordinates back into the basis of $v_1, ..., v_n$

Spectral Theorem

Theorem (Spectral Theorem (!!!))

Let $A \in \mathbb{R}^{n\times n}$ be a ***symmetric*** *matrix. Then there is a orthonormal basis of $\mathbb{R}^n$ composed of eigenvectors of A.*

- One of most important theorems in Linear Algebra
- Symmetric matrices induce an orthonormal basis of eigenvectors
- Makes analyzing the behavior of a symmetric matrix very easy
- Many examples of 'natural' symmetric matrices
    - Covariance Matrix
- Proved in Homework 5

Questions: Spectral Theorem

1. Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Give a vector v with $\|v\| = 1$ such that $\|Av\|$ is maximized.
2. Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, with eigenvalues $\lambda_1, ..., \lambda_n$, and orthonormal family of associated eigenvectors $u_1, ..., u_n$. Give an orthonormal basis of $Ker(A)$ and $Im(A)$ in terms of the u_i's.

Solutions 1: Spectral Theorem

1. Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Give a vector v with $\|v\| = 1$ such that $\|Av\|$ is maximized.

Solution

Let $A = U \Lambda U^T$, *where* U *is orthogonal with columns* $u_1, ..., u_n$, *and* Λ *is diagonal with entries* $\lambda_1, ..., \lambda_n$.
Let $v = \sum_{i=1}^n \alpha_i u_i$, *where* $\sum_{i=1}^n \alpha_i^2 = 1$
Then,

$$
\begin{aligned}
&U^T v = \textstyle\sum_{i=1}^n \alpha_i e_i. \\
&\Lambda U^T v = \textstyle\sum_{i=1}^n \lambda_i \alpha_i e_i \\
&U \Lambda U^T v = \textstyle\sum_{i=1}^n \lambda_i \alpha_i u_i \\
&\|U \Lambda U^T v\| = \|\textstyle\sum_{i=1}^n \lambda_i \alpha_i u_i\| \\
&\|U \Lambda U^T v\| = \textstyle\sum_{i=1}^n \lambda_i^2 \alpha_i^2 \|u_i\| \qquad \textit{by orthonormality} \\
&\|U \Lambda U^T v\| = \textstyle\sum_{i=1}^n \lambda_i^2 \alpha_i^2
\end{aligned}
$$

Maximize this quantity by setting $\alpha_j = 1$, *where* j *is the index with the largest magnitude eigenvalue.*

Solutions 2 : Spectral Theorem

2. Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, with eigenvalues $\lambda_1, ..., \lambda_n$, and orthonormal family of associated eigenvectors $u_1, ..., u_n$. Give an orthonormal basis of $Ker(A)$ and $Im(A)$ in terms of the u_i's.

Solution

Since A is symmetric, then A induces an orthonormal basis of eigenvectors $u_1, ..., u_n$ with eigenvalues $\lambda_1, ..., \lambda_n$, and $A = U\Lambda U^T$.
Let $v = \sum_{i=1}^{n} \alpha_i u_i$
Then by the previous question,

$$U\Lambda U^T v = \sum_{i=1}^{n} \lambda_i \alpha_i u_i$$

Now, consider all j s.t $\lambda_j = 0$.
$Av = \sum_{i=1}^{n} \lambda_i \langle v, u_i \rangle = 0$ *when* $v \in Span(\{u_j | \lambda_j = 0, \text{ for } j \in 1, ..., n\})$.
So $Ker(A) = Span(\{u_j | \lambda_j = 0, \text{ for } j \in 1, ..., n\})$.
Likewise,
$Im(A) = Span(\{u_k | \lambda_k \neq 0, \text{ for } k \in 1, ..., n\})$