Recitation #6

Irina Espejo (iem244@nyu.edu)

Center for Data Science

DS-GA 1014: Optimization and Computational Linear Algebra for Data Science

recall: Markov chain

Markov chain

A sequence of random variables $(X_0, X_1, ...)$ is a Markov chain with state space E and transition matrix P if for all $t \geq 0$,

$$\mathbb{P}(X_{t+1} = y | X = x_0, ..., X_t = x_t) = P(x_t, y)$$

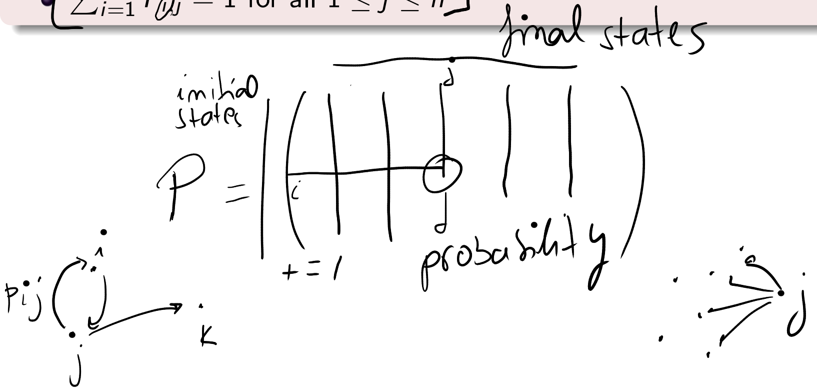
for all $x_0, ..., x_t$ such that $\mathbb{P}(X_0 = x_0, .., X_t = x_t) > 0$
Intuitively, "if the future only depends on the present and not the past"

recall: Markov chain

Stochastic matrix

Let $P \in \mathbb{R}^{n\times n}$ be a matrix, we say P is stochastic if:

- $P_{i,j} \geq 0$ for all $1 \leq i, j \leq n$
- $\sum_{i=1}^{n} P_{i,j} = 1$ for all $1 \leq j \leq n$



Exercise 1

Let $A, B \in \mathbb{R}^{n \times n}$ be stochastic matrices then

1. A is invertible ✓True ✗False
2. The eigenvector corresponding to the largest eigenvalue of A is unique ✓True ✗False
3. A does not have $\lambda = 0$ as eigenvalue ✓True ✗False

1) Counterexample $\begin{pmatrix} 1 & \boxed{1} \\ 0 & \boxed{0} \end{pmatrix}$ Not invertible

2) Counterexample $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ has $\lambda = 1$ twice

3) Counterexample $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ has $\lambda = 0$ because $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Exercise 2

Let $A, B \in \mathbb{R}^{n\times n}$ then be stochastic matrices then AB is also a stochastic matrix.

Hint: express the condition "sum of each column = 1" as a matrix multiplication.

$C = AB$ Goal: C stochastic

① $c_{ij} \geq 0$ ✓ $a_{ij}, b_{ij} \geq 0$

② Sum of the colums of $C = 1$

$(1\dots1)C = (1\dots1)AB$ (A stochastic)

$= (1\dots1)B$ (B stochastic)

$= (1\dots1)$ ✓

A has the condition if and only if

$(\underbrace{1\dots1}_{n})A = (\underbrace{1\dots1}_{n})$

Why? because of matrix product

$$(1\,1\dots1)\overset{A}{\begin{pmatrix} \bullet & \bullet & \cdots & \bullet \\ \vdots & \vdots & & \vdots \\ \bullet & \bullet & \cdots & \bullet \end{pmatrix}} = (1\,1\dots1)$$

$=1 \quad =1 \quad =1$

recall: Markov chain

Proposition

For a Markov chain with the notation above, for all $t \geq 0$

$$\boxed{x^{(t+1)} = Px^{(t)}} \quad \text{and consequently,} \quad \boxed{x^{(t)} = P^t x^{(0)}}$$

$\mu = P\mu$

and recall that the limit $t \to \infty$ is $\boxed{x^{(t)} \to \mu}$ for some $\mu \in \Delta_n$ (probability vector).

Perron-Frobenius Theorem

Let P be a stochastic matrix such that exists $k \geq 1$ such that all the entries of P^k are strictly positive. Then,

- $\lambda = 1$ is an eigenvalues of P with μ its an eigenvector.
- The eigenvalue $\lambda = 1$ has multiplicity equal to $ker(P - Id) = \boxed{Span(\mu)}$

$(P - Id)\mu = 0$

$P\mu = \mu \cdot 1$ $\quad$ $P\mu - \mu = 0$

- For all probability vectors $x \in \Delta_n$ we have $P^t x \to \mu$ in the limit $t \to \infty$

practice: change of basis and diagonalization

Exercise 3

Let $A \in \mathbb{R}^{n \times n}$ be a matrix with eigenvectors $v_1, .., v_n$ and associated eigenvalues $\lambda_1, .., \lambda_n$. Let $\underline{x = \alpha_1 v_1 + ... + \alpha_n v_n}$ be a vector in $\in \mathbb{R}^n$. Show

1. Let P be a linear transformation that maps the canonical basis $e_1, .., e_n$ of $\mathbb{R}^n$ to the eigenvector basis of A: $v_1, .., v_n$. Write P explicitly.
2. What is PDP^{-1}? ($D = diag(\lambda_1, .., \lambda_n)$)
3. Simplify $(PDP^{-1})^k$ for $k \in \mathbb{N}$
4. If $A = PDP^{-1}$, give an interpretation of the action of A

Exercise 3

Let P be a linear transformation that maps the canonical basis $e_1, .., e_n$ of $\mathbb{R}^n$ to the eigenvector basis of A: $v_1, .., v_n$.
Write P explicitly.

$$P : e_i \longrightarrow v_i$$

$$P = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$$

$$e_1 \rightarrow P(e_1)$$

Exercise 3

What is PDP^{-1}? ($D = diag(\lambda_1, .., \lambda_n)$)
Simplify $(PDP^{-1})^k$ for $k \in \mathbb{N}$

$x = \sum_i \alpha_i v_i$

- $A = PDP^{-1}$? $Ax = PDP^{-1}x \stackrel{\sim}{=} = PDP^{-1}\left(\sum_i \alpha_i v_i\right)$

  $P: e_i \to v_i \quad P^{-1}: v_i \to e_i$

  $= PD\sum_i \alpha_i e_i$

  $= P\left(\sum_i \alpha_i \lambda_i e_i\right)$

  $= \sum \lambda_i \alpha_i v_i \quad \square$

- $\underbrace{(PDP^{-1})^k}_{A^k} = PDP^{-1}\,PDP^{-1} \ldots PDP^{-1}$

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

Exercise 3

If $A = PDP^{-1}$, give an interpretation of the action of A

P^{-1} : takes eigenvectors to canonical basis

D : expands the coordinate i by λ_i for all $i = 1, \dots, n$
or shrinks

P : takes the canonical basis to eigenvector basis

$A = PDP^{-1}$: does all the above in order

recall: Spectral theorem

Spectral theorem !!

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

Exercise 4

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

$v = (\; . \; . \; . \;)?$

$v = (v_1 \ldots v_n)$

$A = P D P^{-1}$ P is orthonormal

eigenbasis $\mu_1 \ldots \mu_n$ $P^{-1} = P^T$

$$\|Av\| = \| P D P^{-1} \| = \| \sum_{i=1} \lambda_i v_i \| \leq |\lambda_{max}| \, v_j$$

$\square$

$v = \sum_i v_i \mu_i$ $\qquad \lambda_{max} = \max_i |\lambda_i|$

$v = (0 \ldots \underset{j}{1} \ldots)$ $\quad \lambda_j = \lambda_{max}$

practice: Spectral theorem

Exercise 5

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenvalues $\lambda_1, .., \lambda_n$ and orthonormal eigenvectors $v_1, .., v_n$.
Give an orthonormal basis of $Ker(A)$ and $Im(A)$ in terms of $v_1, .., v_n$ only.

$$Ker(A) = \{v \in \mathbb{R}^n \mid Av = 0\} = \{v \in \mathbb{R}^n \mid Av = \lambda v \text{ with } \underbrace{\lambda = 0}_{\text{eigenvalue}}\}$$

therefore, $Ker(A) = Span(\{v_j \in \mathbb{R}^n \mid \lambda_j = 0\})$

$$Im(A) = \{w \in \mathbb{R}^n \mid \text{exists } v \in \mathbb{R}^n \ Av = w\}$$

Appy rank-nullity theorem: $d(Im(A)) + dim(Ker(A)) = n$

$$\left.\begin{array}{l} dim(Im(A)) = n - dim(Ker(A)) \ + \ v_1, .., v_n \text{ is orthonormal basis of } \mathbb{R}^n \end{array}\right\} Im(A) = Span(\{v_k \mid \lambda_k \neq 0\})$$

Exercise 6(*)

Let $A, B \in \mathbb{R}^{n\times n}$ be symmetric matrices. Show that $AB = BA$ if and only if A and B diagonalize in the same basis. (Does the same hold if we just assume that A, B are diagonalizable?)

$\boxed{\Rightarrow}$ $AB = BA$ show A, B have same eigenvectors

Let $v \in \mathbb{R}^n$ be an eigenvector of A : $Av = \lambda v$

$A(Bv) = B(Av) = B\lambda v = \lambda(Bv) \rightarrow Bv$ is also an eigenvector of A

But $\lambda_1 \neq \dots \neq \lambda_n$ so the multiplicity is 1 therefore $u = \alpha Bu$ $\alpha \in \mathbb{R}$

So u is also an eigenvalue of B and viceversa.

$\boxed{\Leftarrow}$ Let $A = PD_AP^{-1}$ and $B = PD_BP^{-1}$ show $AB = BA$

$AB = (PD_A\cancel{P^{-1}})(\cancel{P}D_BP^{-1}) = PD_AD_BP^{-1} = PD_BD_AP^{-1} = \underbrace{PD_BP^{-1}}\underbrace{PD_AP^{-1}} = BA$