Recitation 4/28

Sungwoo Jeong Tuesday 10AM, 11AM

April 28, 2020

Markov Matrix

$M \in \mathbb{R}^{n\times n}$ is called *Markov matrix* if it has following properties

- All the entries are positive or zero
- Sum of any column is 1.

Markov matrix is used to express a probabilistic behaviors. For example, we have two cities A, B. A resident of A moves to B a year later with probability 0.2 and a resident in B moves to city A a year later with probability 0.1. Then the matrix equation of population of year n can be expressed as,

$$\begin{pmatrix} \text{pop(A)}_n \\ \text{pop(B)}_n \end{pmatrix} = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix} \begin{pmatrix} \text{pop(A)}_{n-1} \\ \text{pop(B)}_{n-1} \end{pmatrix}$$

An important property of Markov matrix is that they always have eigenvalue one. (Why?)

An eigenvector π corresponding to $\lambda = 1$ is called startionary vector, as they satisfy $M\pi = \pi$.

(Fact 1) All the eigenvalues of M have absolute value smaller or equal to 1.
(Fact 2) If all the entries are positive, the eigenvalue 1 has multiplicity one.
(Fact 3) If all the entries are positive and M has linearly independent eigenvectors then $M^n u$ approaches the direction of π, for any vector u.

Differential Equations

We will discuss a differential equations,

$$\frac{d}{dt}x(t) = Ax(t)$$

where $A \in \mathbb{R}^{n\times n}$ and $x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, a vector valued function of $t \in \mathbb{R}$. So in other words, the derivatives of x_i can be expressed as linear combination of $x_1, x_2, \ldots, x_n$.

Here, a matrix exponential is used. Recall we defined matrix exponential of matrices who has eigendecomposition(i.e. diagonalizable matrices) as

$$e^A = X \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} X^{-1}$$

where $A = X\Lambda X^{-1}$ is the eigendecomposition of A. Another definition of matrix exponential can be defined as

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \ldots$$

For differential equations above, a matrix valued function e^{tA} is used. Assume $x(t) = e^{tA}y$ where $y \in \mathbb{R}^n$ is just a vector. We have

$$\frac{d}{dt}x(t) = Ae^{tA}y = Ax(t)$$

and we need to use the initial condition to find y.

Symmetric Matrix

$A \in \mathbb{R}^n$ be symmetric matrix, i.e. $A^T = A$. Then we have following important identities.

1. All eigenvalues are in $\mathbb{R}$.
   - Recall that eigenvalues are usually complex - try eigvals(randn(10, 10)) in Julia
2. Eigenvectors are orthogonal to each other(except for colinear ones belong to same eigenvalue)(Why?)
3. A is diagonalizable

Thus we have a new eigendecomposition when we select orthonormal eigenvectors

$$A = X\Lambda X^T$$

Symmetric Positive Definite Matrix

(Formal Definition) Symmetric matrix $A \in \mathbb{R}^{n\times n}$ is called *Positive Definite* if for any vector $x \in \mathbb{R}^n$ we have $x^T Ax > 0$. Similarly, A is *Positive Semidefinite* if $x^T Ax \geq 0$ and *Negative Definite* if $x^T Ax < 0$ for all vectors x.

(Easier Definition) Symmetric matrix A is Positive definite if all eigenvalues are positive.

- Easy way to construct a positive (semi)definite matrix : $A^T A$ for any matrix A. (Why?)

Similar Matrices

Matrices $A, B \in \mathbb{R}^{n\times n}$ are called similar if there exists an invertible matrix P such that $A = PBP^{-1}$.

Diagonalization in terms of similar? How can we define Diagonalizable?

If A, B are similar, then they have same
- Eigenvalues
- Trace(Sum of diagonal)
- Determinants
- Rank

Problems

1. Assume A, B are similar matrices. What is the trace of $A - B$?

2. (a) For symmetric positive definite matrix A, we always have a symmetric matrix $B = \sqrt{A}$ by applying square roots to eigenvalues on eigendecomposition. Then we have $B^2 = B^T B = A$. Can we always find non-symmetric B such that $A = B^T B$?

(b) For symmetric positive definite A, Cholesky decomposition is defined as $A = LL^T$ where L is lower triangular matrix. Prove it exists, by using (a) and QR decomposition.

ANSWERS

1. Trace is a linear function so trace of $A - B$ is zero.

2.(a) We can use SVD. Let eigendecomposition of $A = V\Lambda V^T$. We can take $B = USV^T$ with some random orthogonal matrix U, and S being the diagonal matrix with diagonals being square root values of diagonals of Λ. Then $B^TB = A$ and B is not a symmetric matrix.

(b) From (a) we have B such that $B^TB = A$. Let QR decomposition of B be $B = QR$. Then $B^TB = R^TQ^TQR = R^TR$ and if we take $L = R^T$ a lower triangular matrix, we have $A = LL^T$ so it exists by existence of QR decomposition.