Linear Algebra and Optimization
MIT Fall 2022

Recitation 19

Thursday November 17, 2022

1 Recap

1.1 Non-negative matrices

1. Definition. A non-negative matrix is a **square** matrix with only **non-negative entries**.

2. Perron's Theorem
   Let A be a positive matrix; then A has a **positive eigenvalue** λ such that

   (a) λ has algebraic and geometric **multiplicity 1**

   (b) the value of λ is greater than the absolute value of any other eigenvalue of A (**dominant eigenvalue**)

   (c) the corresponding **eigenvector** v **is positive** (up to a scalar factor)

3. **Markov Matrices**
   A Markov matrix is a **non-negative** matrix where the **sum of each row or column is equal to 1**.

   (a) If the row elements add up to 1, it is called row stochastic. If the column elements add up to 1, then it is called column stochastic.

   (b) It has eigenvalue 1, and the corresponding eigenvector is non-negative.

   (c) All other eigenvalues have absolute value less than or equal to 1.

4. **Applications**

   (a) *Page Rank*
   Suppose we have a system of n pages $T_1, ..., T_n$. Let $B(T_i)$ be the set of links pointing to page T_i, and let $C(T_i)$ be the number of links going out of page T_i. The rank of page T_i is defined as:

$$PR(T_i) = (1-d)\frac{1}{n} + d \sum_{j \in B(T_i)} \frac{PR(T_j)}{|C(T_j)|} \tag{1}$$

   d is a damping parameter between 0 and 1. Smaller d makes the distribution more uniform.
   We want to express the page rank of all pages in matrix form. Let G be an

$n \times n$ matrix where each entry $G_{ij} = \mathbb{1}_{j \in B(T_i)} \times \frac{1}{|C(T_j)|}$, i.e. $G_{ij} = \frac{1}{|C(T_j)|}$ if page T_i points to T_j, otherwise $G_{ij} = 0$. Let $\mathbf{e}$ be a vector of 1's.

$$\begin{bmatrix} PR(T_i) \\ \ldots \\ PR(T_n) \end{bmatrix} = \left(\frac{1-d}{n}\right)\mathbf{e} + d\mathbf{G}\begin{bmatrix} PR(T_i) \\ \ldots \\ PR(T_n) \end{bmatrix} \tag{2}$$

$$\mathbf{p} = \left(\frac{1-d}{n}\mathbf{e}\mathbf{e}^T + d\mathbf{G}\right)\mathbf{p} = \mathbf{M}\mathbf{p} \tag{3}$$

We go from equation (2) to (3) using the fact that $e^T p$ is equal to 1 (the page ranks of all pages add up to 1). Matrix $\mathbf{M}$ in equation (2) is a Markov matrix.

(b) *Probability Transition*
More generally, we can use a Markov matrix to represent the probability transition between states.

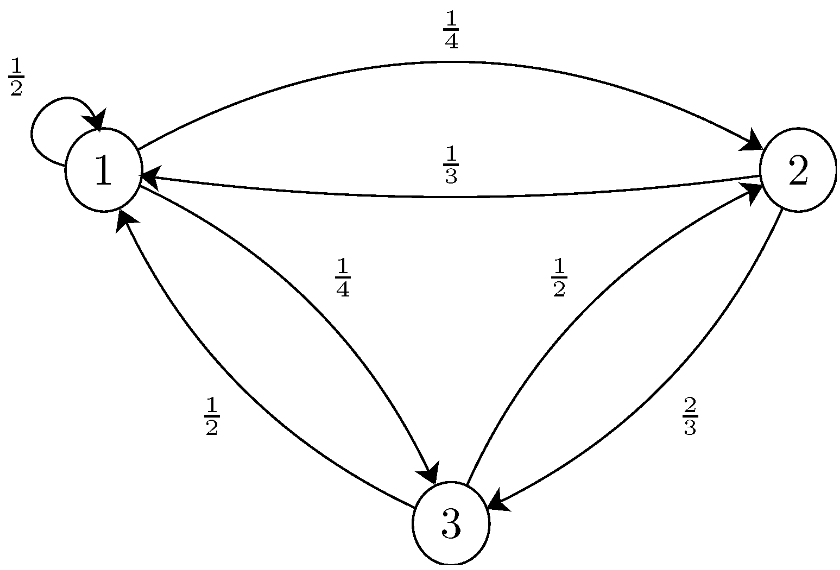
i. Column Stochastic (Hopping Rabbit)
Let A_{ij} be the probability of jumping to state i given that you are currently in state j. In this case, the total transition probability from some state j to any other state i must be 1, i.e. each column adds up to 1. The probability update is $x_{k+1} = Ax_k$.

ii. Row Stochastic
Let A_{ij} be the probability of jumping to state j given that you are currently in state i. Since the total transition probability from state i to all other states must be 1, i.e. each row adds up to 1. The probability update is $x_{k+1}^T = x_k^T A$.

2 Exercises

1. True/False: explain or give a counter example
   (a) A markov matrix can have negative eigenvalues.
   (b) The transpose of a markov matrix can also be a markov matrix.
   (c) The product of two column-stochastic matrices is also column-stochastic.
   (d) The product of two row-stochastic matrices is also row-stochastic.

2. Show that all eigenvalues of a row stochastic matrix are in absolute value smaller or equal to 1.



3. (a) Write the transition matrix A associated with the state diagram above.
   (b) The initial state is $x_0 = (1, 0, 0)$, i.e. we start in state 1. What is the probability vectors for the next time step x_1?
   (c) In the long-term, which state do you think is more probable?
   (d) Find the steady state vector, i.e. what is x such that $x_{k+1} = x_k$? Does this match your expectation in part c? Is it unique?