Lecture 8

Monday, February 3, 2025   2:10 PM

A *vector* is a matrix with only one row or only one column. A column vector v is said to be a *stationary vector* of matrix A if $v \neq 0$ and $Av = v$. Vector v is called a *probability vector* if every entry is nonnegative and the sum of all entries is equal to 1.

Example: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

Does A have a stationary vector? If so, does it have a probability stationary vector?

$Av = v = I_2 v$

$(A - I_2)v = 0$

Let $v = \begin{bmatrix} x \\ y \end{bmatrix}$. Then the equation $(A - I_2)v = 0$ is equivalent to a linear system of two equations, whose associated matrix is

$\begin{bmatrix} 0 & 2 & 0 \\ 3 & 3 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

The solution is $v = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, which is not acceptable. Matrix A doesn't have a stationary vector.

Example: $A = \begin{bmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{bmatrix}$

Does A have a stationary vector? If so, does it have a probability stationary vector?

Perron-Frobenius theorem (1907, 1912): Let A be a stochastic matrix (i.e. a matrix whose columns are probability vectors). Suppose that either A or some power A^k has all positive entries. Then A has a unique probability stationary vector.

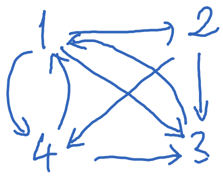
In 1998, while being PhD students at Stanford University, Lawrence Page and Sergey Brin published a paper "The anatomy of a large-scale hypertextual Web search engine" on the journal of *Computer Networks and ISDN systems*. They discovered a profound algorithm to rank web pages. They called it PageRank algorithm, a name that attributes one of the coauthor and also refers to the web *pages*. This algorithm is the basis of the Google search engine.

The overall idea is that web pages are ranked based on their importance. The most important page is ranked first. Each webpage will be given a weight (rank) between 0 and 1. The sum of the weights of all web pages is equally to 1. The following rules are natural:

- Any link from a web page to itself does not count.
- Multiple links from web page A to webpage B are counted only once.
- The importance of a page is determined not only by the number of pages that link to it but also by the importance of the pages that link to it.

Because of the last rule, you may think of the problem of determining the importance of the web pages as a recursive problem.

Example: Consider 4 webpages that link to one another as follows.



At the beginning, all pages are considered equally important. So, the weights of each web page is $1/4 = 0.25$.
Imagine that you are a web-surfer. From web page i, you can "transition" to any webpage that it links to with an equal chance. The following transition matrix incorporate this idea.

$$A = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

The chance you will be in web page 1 will be …

The RREF of $A - I_4$ is

$$A = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2/3 \\ 0 & 0 & 1 & -3/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$