A New Method to Find the Eigenvalues of Convex Matrices with Application in Web Page Rating

F. Soleymani

Department of Mathematics, Islamic Azad University
Zahedan branch, Zahedan, Iran
fazl_soley_bsb@yahoo.com

Abstract

Although many factors determine Google's overall ranking of search engine results, Google maintains that the heart of its search engine software is PageRank, the eigenvector corresponds to the maximal eigenvalue in the Google matrix. In this note, we deals with a very special case of this positive matrix; i.e. when it is a positive convex matrix, then we provide a new way which can be used as a direct method, to find the eigenvalues of this type of matrices. Although in practice, the likelihood of this scarce phenomenon is improbable, but it is worth mentioning.

Mathematics Subject Classifications: 15A86, 65F15

Keywords: Eigenvalue, Google matrix, convex matrices, PageRank

1 Introduction

The computation of the eigenvalues a given matrix is one of the most important subjects in applied mathematics. There are several numerical methods to approximate these eigenvalues, such as Power Iteration Method or Rayleigh Quotient Method, etc. that are vastly discussed in literature (see, for example, [6]). Unfortunately, the computations costs of these methods are so much and more precision has been wanted in the real world, (see [4]), like the finding of Google matrix eigenvalues which play an important role in calculation of PageRank (the eigenvector of the Google matrix corresponds to the maximal eigenvalue, i.e., 1).

This paper considers a class of special matrices ($r-$convexity preserving matrices for all $r \leq k$) that are significant in many applications, for instance in computer aided geometric design (see [1,2,5]). As an another instance, in finding the

PageRank. Note that, Google founders Sergey Brin and Larry Page With the financial assistance of a small group of initial investors founded the Web search engine company Google, Inc. in September 1998.

Approximately 94 million American adults use the Internet on a typical day. Search engine use is next in line and continues to increase in popularity. In fact, survey findings indicate that nearly 60 million American adults use search engines on a given day. Even though there are many Internet search engines, Google, Yahoo!, and MSN receive over 81% of all search requests. Despite claims that the quality of search provided by Yahoo! and MSN now equals that of Google, Google continues to thrive as the search engine of choice, receiving over 46% of all search requests, nearly double the volume of Yahoo! and over four times that of MSN [7].

To gain a better understanding of the upcoming propositions, it is better off to provide some fundamental definitions.

Definitition 1.1. Suppose that $k \in Z^+$. A vector $v = (v_1, v_2, \cdots, v_n)^T \in R^n$ is said to be k-convex if $\Delta^k v_i \geq 0$ for all $i \in \{1, 2, \cdots, n - k\}$, where

$$\Delta^k v_i = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} v_{i+j}.$$

So it is crystal clear that a vector is $0-$convex if and only if it is nonnegative and a vector is $1-$convex if and only if it is monotonically increasing. A matrix A is said to be $k - convexity\ preserving$ if for any $k-$convex vector v, the vector Av is also $k-$convex. According to this definition, A is $0-$convexity preserving iff it transforms nonnegative vectors into nonnegative vectors and a matrix A is $1-$convexity preserving iff it is monotonically preserving.

As a result, it would be obvious that v is $k-$concave when $-v$ is $k-$convex.

Lemma 1.1. If we denote the lower triangular matrix as follow

$$E := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & 1 & 0 \\ 1 & \cdots & \ddots & 1 \end{pmatrix},$$

then

$$E^{-1} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -1 & 1 & & & \vdots \\ 0 & -1 & 1 & \ddots & \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

By Lemma 1.1 for each $j \in \{1, \cdots, n\}$, the following result could be concluded, let E_j be the following $n \times n$ matrix: $E_1 := E$ and for $j \geq 2$,

$$E_j := \begin{pmatrix} I_{j-1} & | & 0 \\ - & & - \\ 0 & | & E \end{pmatrix}, \qquad E_j^{-1} = \begin{pmatrix} I_{j-1} & | & 0 \\ - & & - \\ 0 & | & E^{-1} \end{pmatrix},$$

where I_{j-1} is the $(j-1) \times (j-1)$ −identity matrix and E is the $(n-j+1) \times (n-j+1)$ −matrix given by above-mentioned lemma.

The following result corresponds to Corollary 4.4 of [1] and shows an important property of r−convexity preserving matrices for $r = 0,1, \cdots, k$ $(k \geq 1)$.

Proposition 1.1. Let A be a r−convexity preserving matrix for $r = 0,1, \ldots, k$ $(k \geq 1)$. Then

$$(E_1 \ldots E_k)^{-1} A (E_1 \ldots E_k) = \begin{pmatrix} \Lambda_k & | & * \\ - & & - \\ 0 & | & A_k \end{pmatrix},$$

where Λ_k is an upper triangular matrix whose diagonal elements $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k (\geq 0)$ are the largest eigenvalues of A, and A_k is a nonnegative matrix with $\rho(A_k) \leq \lambda_k$.

Observe that, by the previous result, the first k columns of $E_1 \ldots E_k$ form a basis of the invariant subspace corresponding to the k dominant eigenvalues of a matrix A which is r−convexity preserving for $r = 0,1, \cdots, k$ $(k \geq 0)$.

Before presenting the main propositions, it is necessary that mention the following note about the combinatorial numbers.

By induction, it is easy to see that for all $m \geq 0$,

$$\sum_{j=0}^{m} \binom{k+j}{k} = \binom{k+1+m}{k+1}.$$

Given a square matrix A of order n the following results provide, explicit expressions for the columns of $B_k := A E_1 \ldots E_k$ and the rows of $(E_1 \cdots E_k)^{-1} A (E_1 \cdots E_k)$ in terms of the A and of the rows of B_k.

Another main definition which is important in the rest of the paper is the definition of Google matrix. The *Google matrix* is a positive matrix obtained by a special rank-one perturbation of a stochastic matrix which represents the hyperlink structure of the webpages. More technically, if $e = (1,1,\cdots,1)^T$ and v is the probability vector, then the Google matrix is defined as follows:

$$G(\alpha) = \alpha S + (1 - \alpha)ev^T,$$

where α, i.e., the damping factor, is in the real open interval $(0,1)$, S is a stochastic matrix

The $n-$dimensional probability vector v, also called personalization or stochastic vector, is a positive vector which its $1-$norm is equal to 1. That is, $\|v\|_1 = 1$. Consequently v is a personalization vector when it satisfies in the following relation:

$$v_1 + v_2 + \cdots + v_n = 1, \qquad \forall i \quad v_i \geq 0.$$

2 Main Result

Proposition 2.1. Let A be an square matrix of order n with columns $A^1, \cdots, A^n$, $k < n$ and $B_r := AE_1 \cdots E_r$ for $r = 1, \cdots, k$. Let us denote by $B_r^1, B_r^2, \cdots, B_r^n$ the columns of B_r. Then, for each $j \in \{1, \cdots, r - 1\}$,

$$B_r^j = \sum_{i=j}^{n} \binom{i}{j-1} A^i, \tag{1}$$

and, for each $j \in \{r, \cdots, n\}$,

$$B_r^j = \sum_{i=j}^{n} \binom{r-1+i-j}{r-1} A^i. \tag{2}$$

Proof. The simplest way to prove these formulas is the use of induction on $r \in \{1, \cdots, k\}$. For $r = 1$ we have by the definition of E_1 that

$$B_1 = AE_1 = (A^1, A^2, \cdots, A^n)E_1 = \left(\sum_{i=1}^{n} A^i, \sum_{i=2}^{n} A^i, \cdots, \sum_{i=n}^{n} A^i \right).$$

Then $B_1^j = \sum_{i=j}^{n} A^i$ for all $j \in \{1, \dots, n\}$ and therefore formulas (1) and (2) hold for $r = 1$. Let us suppose that (1) and (2) hold for $r \in \{1, \dots, k-1\}$ and let us prove them for $r + 1$. We can write

$$B_{r+1} = AE_1 \cdots E_{r+1} = (AE_1 \cdots E_r)E_{r+1}.$$

Next by the induction hypothesis we have

$$B_{r+1} = (B_r^1, B_r^2, \cdots, B_r^n)E_{r+1},$$

where B_r^j are given by (1) for $j \in \{1, \cdots, r-1\}$ and by (2) for $j \in \{r, \cdots, n\}$. By the definition of E_{r+1} and the induction hypothesis we have that

$$B_{r+1}^j = B_r^j = \sum_{i=j}^{n} \binom{i}{j-1} A^i,$$

for $j \in \{1, \cdots, r\}$. Therefore (1) holds for $j \in \{1, \cdots, r\}$. Analogously, by the definition of E_{r+1} and the induction hypothesis we deduce

$$B_{r+1}^j = \sum_{i=j}^{n} B_r^j = \sum_{i=j}^{n} \sum_{m=i}^{n} \binom{r-1+m-i}{r-1} A^m,$$

for $j \in \{r+1, \cdots, n\}$. Reordering the terms (See for more [3]) in the previous formula it can be written as

$$B_{r+1}^j = \sum_{m=j}^{n} \left[\sum_{i=j}^{m} \binom{r-1+i-j}{r-1} \right] A^m,$$

Changing the index of the inner sum in the previous formula we have that

$$B_{r+1}^j = \sum_{m=j}^{n} \left[\sum_{i=0}^{m-j} \binom{r-1+i}{r-1} \right] A^m,$$

finally, applying the presented note, we deduce

$$B_{r+1}^j = \sum_{m=j}^{n} \binom{r+m-j}{r} A^m.$$

Hence formula (2) also holds for $j \in \{r+1, \cdots, n\}$ and the induction follows. ∎

Proposition 2.2. Let $k < n$, $N_0 := B_k$ be the matrix defined in Proposition 2.1, with rows $B_{k,1}, \cdots, B_{k,n}$, and let $N_r := (E_1 \cdots E_r)^{-1} B_k$ for $r \in \{1, \cdots, k\}$. Let us denote by $N_{r,1}, \cdots, N_{r,n}$ the rows of N_r. Then, for each $i \in \{1, \cdots, r\}$,

$$N_{r,i} = \nabla^{i-1} B_{k,i} = \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} B_{k,i-j}, \tag{3}$$

and, for each $i \in \{r+1, \dots, n\}$,

$$N_{r,i} = \nabla^r B_{k,i} = \sum_{j=0}^{r} (-1)^j \binom{r}{j} B_{k,i-j}, \tag{4}$$

where ∇ is the usual backward difference $\nabla f_i := f_i - f_{i-1}$.

Proof. Performing $E_s^{-1} N_{s-1}$ for all $s \in \{1, \dots, k\}$ consists of subtracting from the rows $s+1, \dots, k$ of N_{s-1} the rows $s, \dots, k-1$ of N_{s-1}, respectively. Therefore, we have that, for each $i \in \{1, \dots, r\}$, $N_{r,i} = \nabla^{i-1} B_{k,i}$, and that, for each $i \in \{r+1, \dots, n\}$, $N_{r,i} = \nabla^r B_{k,i}$. Finally, the combinatorial formulas in (3) and (4) are well-known formulas for the backward difference ∇. ∎

From the last two propositions we derive the following results.

Corollary 2.1. Let k be a positive integer less than n, let A be an r−convexity preserving matrix for $r = 0,1, \dots, k$ $(k \geq 1)$ and let   be the matrix defined in proposition 2.2. Then $m_{11}, m_{22}, \cdots, m_{kk}$ are the largest eigenvalues of A satisfying $m_{11} \geq m_{22} \geq \cdots \geq m_{kk} \geq 0$, and the remaining $n-k$ eigenvalues of A are the $n-k$ eigenvalue of the nonnegative matrix $(m_{ij})_{k+1 \leq i,j \leq n}$.

Google likely will remain the top search engine. But what set Google apart from its competitors in the first place? The answer is PageRank. So it is interesting to improve and discuss about the existed methods and try to improve them. The gist of this work is provided in the following corollary.

Corollary 2.2. If the Google matrix be an r-convexity preserving matrix then the above algorithm can be applied to find the eigenvalues of this matrix.

3 Concluding remarks and discussion

By Corollary 2.1, the explicit formulas of proposition 2.1 and 2.2 allow us to calculate the k largest eigenvalues of a matrix A $r-$convexity preserving for $r = 0, 1, \cdots, k$. If A is an r-convexity preserving matrix for $r = 0, 1, \cdots, n - 1$, then by the previous corollary, all its eigenvalues only depend on the entries of the upper triangular part of A. These formulas include two phases: the first phase corresponds to the calculation of entries above and to right of the first diagonal entries of $B_k = A(E_1 \cdots E_k)$ (phase 1) and the second one corresponds to the calculation of the first diagonal entries of $N_k = (E_1 \cdots E_k)^{-1} B_k$ (phase 2).

We know that PageRank has connections to numerous areas of mathematics and computer science such as matrix theory, numerical analysis, information retrieval, and graph theory. As a result, much research continues to be devoted to explaining and improving PageRank. Here, by Corollary 2.2, it will be easy to find the n eigenvalue of Googe matrix.

References

[1] J.M. Carnicer, M. Garcia-Esnaola, J.M. Pena, Generalized convexity preserving transformations, Comput, Aided Geom. Design 13 (1996) 179-197.

[2] S. Cooper, S. Waldron, The eigenstructure of the Bernstein operator, J. Approx. theory 105 (2000) 133-165.

[3] J. Delgado, J.M. Peña, Computation of the eigenvalues of convexity preserving matrices, Appl. Math. Lett. 22 (2009) 470-479.

[4] J. Ding, A. Zhou, Eigenvalues of rank-one updated matrices with some applications, Appl. Math. Letters (2007) 1223-1226.

[5] N. S. Sapidis, P. D. Kaklis, An algorithm for constructing convexity and monotonicity-preserving splines in tension, Computer Aided Geometric Design, Vol 5, 1988, 127-137

[6] T. Sauer, Numerical Analysis, Addison Wesley Publication, USA, (2005).

[7] S. Serra Capizzano, Google PageRank problem: The model and the analysis, in: Proceedings of the Dagstuhl Conference in Web Retrieval and Numerical Linear Algebra, 2007.