On the eigenvalues of specially low-rank perturbed matrices

Yunkai Zhou*

April 12, 2011

Abstract

We study the eigenvalues of a matrix A perturbed by a few special low-rank matrices. The perturbation is constructed from certain basis vectors of an invariant subspace of A, such as eigenvectors, Jordan vectors, or Schur vectors. We show that most of the eigenvalues of the low-rank perturbed matrix stayed unchanged from the eigenvalues of A; the perturbation can only change the eigenvalues of A that are related to the invariant subspace. Existing results mostly studied using eigenvectors with full column rank for perturbations, we generalize the results to more general settings. Applications of our results to a few interesting problems including the Google's second eigenvalue problem are presented.

Key words: eigenvalue, low-rank, Jordan, Schur, canonical form, invariant subspace, perturbation, Google matrix

1 Introduction

In recent years, much attention has been directed at the second eigenvalue problem of a Google matrix (e.g. [7, 8, 6, 10, 11]), where a Google matrix is a special rank-1 perturbation of a row stochastic matrix. This "second eigenvalue" problem is critical for the convergence rate of the power-related methods used for Google's PageRank computation (see [9] and references therein), moreover, the problem has its own theoretic interests.

The original study of the problem in [7] utilizes properties of Markov chain and ergodic theorem. The result is generalized using purely algebraic derivation in [8, 6, 10], where all (including the second) of the eigenvalues of a Google matrix can be explicitly expressed.

A Google matrix can be considered as a special case of the following specially low rank perturbed matrix

$$A + UV^{\mathrm{H}}, \text{ where } A \in \mathbb{C}^{n\times n}, \text{ and } U, V \in \mathbb{C}^{n\times k} \tag{1}$$

where A is a general $n \times n$ matrix, it can be real or complex, and $k \geq 1$. The superscript $()^{\mathrm{H}}$ denotes the conjugate transpose.

The U matrix is usually restricted to k linearly independent eigenvectors of A. This case is of considerable interests and has been studied in [2, 4, 12, 13, 5]. Besides including the Google matrix as its special case, the eigenproblem of (1) has applications in various other situations (see the many applications in e.g. [5]).

*Department of Mathematics, Southern Methodist University, Dallas, TX 75275 (yzhou@smu.edu). Supported in part by the National Science Foundation under Grant No. CMMI-0727194 and OCI-0749074.

We generalize U from eigenvectors of A to a basis of an invariant subspace of A, and also extend to the cases where U is not linearly independent. The conclusion on the eigenvalues of the perturbed matrix is quite similar to those perturbed by explicit eigenvectors. More specifically, the nice feature that most eigenvalues remain the same under the perturbation is kept; and the eigenvalues that are changed can be readily computed from a special $k \times k$ matrix.

We use a proof technique which is based on expressing a similarity transformation, not in the form of $A = XBX^{-1}$, but in the more convenient invariant (sub)space form of $AX = XB$, or more generally $A\tilde{X} = \tilde{X}B$, where B is of upper triangular form and $\tilde{X}$ contains columns of X. This simple technique is applicable to all the results in Section 2, and it simplifies the proofs in [2, 12, 13] when we restrict U in (1) to eigenvectors of A.

2 Main results

The statement of each theorem in this section starts with a canonical form of A. The availability of a canonical form makes the proof straightforward, since it makes a desired similarity transformation readily available. If one starts only with the basis of an invariant subspace, then the proof will begin with adding linearly independent vectors to the basis to construct a similarity transformation, and then to show that the two-sided projection of A on the added vectors preserve certain eigenvalues. The approach we take is simpler and can quickly reveal the essence of the change of eigenvalue due to a given perturbation.

Lemma 2.1. *Let the Jordan canonical form of $A \in \mathbb{C}^{n\times n}$ be $A = XJX^{-1}$. Assume that J contains two diagonal blocks $J = \begin{bmatrix} J_1 & \\ & J_2 \end{bmatrix}$, where $J_1 \in \mathbb{C}^{k\times k}$ with $k < n$. Partition X accordingly as $X = [X_1, X_2]$ such that $AX_i = X_iJ_i, i = 1, 2$. Let V be any matrix in $\mathbb{C}^{n\times k}$. Then the eigenvalues of $A + X_1V^H$ are the union of the eigenvalues of $J_1 + V^H X_1$ and the diagonal elements of J_2.*

Proof: Note that

$$\begin{aligned} (A + X_1V^{\mathrm{H}})[X_1, X_2] &= [X_1J_1, X_2J_2] + X_1V^{\mathrm{H}}[X_1, X_2] \\ &= [X_1(J_1 + V^{\mathrm{H}}X_1), X_2J_2 + X_1V^{\mathrm{H}}X_2] \\ &= [X_1, X_2] \begin{bmatrix} (J_1 + V^{\mathrm{H}}X_1), & V^{\mathrm{H}}X_2 \\ & J_2 \end{bmatrix}. \end{aligned}$$

Therefore $A + X_1V^{\mathrm{H}}$ is similar to the block upper triangular matrix $\begin{bmatrix} (J_1 + V^{\mathrm{H}}X_1), & V^{\mathrm{H}}X_2 \\ & J_2 \end{bmatrix}$, which proves the result. ■

In fact, there is no need to assume that the Jordan form contains two (disjoint) diagonal blocks. We can obtain the same result on the eigenvalues of the perturbed matrix by dropping this assumption. In Theorem 2.1, clearly, if $span\{X_2\}$ contains a dimension $n-k$ invariant subspace of A, then J_{12} is a k by $n-k$ zero matrix.

Theorem 2.1. *Let the Jordan canonical form of $A \in \mathbb{C}^{n\times n}$ be $A = XJX^{-1}$. Partition J into four subblocks as $J = \begin{bmatrix} J_1 & J_{12} \\ & J_2 \end{bmatrix}$, where $J_1 \in \mathbb{C}^{k\times k}$ with $k < n$. Partition X accordingly as*

$X = [X_1, X_2]$ such that $AX_1 = X_1 J_1$ and $AX_2 = X_1 J_{12} + X_2 J_2$. Let V be any matrix in $\mathbb{C}^{n\times k}$. Then the eigenvalues of $A+X_1V^H$ are the union of the eigenvalues of $J_1+V^H X_1$ and the diagonal elements of J_2.

Proof: Using the same trick as in Lemma 2.1 ,

$$\begin{aligned}(A + X_1V^{\mathrm{H}})[X_1, X_2] &= [X_1J_1,\ X_1J_{12} + X_2J_2] + X_1V^{\mathrm{H}}[X_1, X_2] \\ &= [X_1(J_1 + V^{\mathrm{H}}X_1),\quad X_1J_{12} + X_2J_2 + X_1V^{\mathrm{H}}X_2] \\ &= [X_1, X_2]\begin{bmatrix}(J_1 + V^{\mathrm{H}}X_1), & J_{12} + V^{\mathrm{H}}X_2 \\ & J_2\end{bmatrix}.\end{aligned}$$

Therefore $A+X_1V^{\mathrm{H}}$ is similar to the block upper triangular matrix $\begin{bmatrix}(J_1 + V^{\mathrm{H}}X_1), & J_{12} + V^{\mathrm{H}}X_2 \\ & J_2\end{bmatrix}$, whose eigenvalues are the union of the eigenvalues of the two diagonal blocks $J_1 + V^{\mathrm{H}}X_1$ and J_2. ■

When Jordan blocks are diagonal, the Jordan vectors become eigenvectors. Therefore, Lemma 2.1 and Theorem 2.1 may be considered as a generalization of the results in [3, 2, 4, 12, 13, 5] where eigenvectors are used to construct the low-rank perturbation matrix.

There is no particular need to restrict ourselves to Jordan vectors, we can further generalize the results using Schur vectors which are numerically more stable to compute, as stated in Theorem 2.2.

Theorem 2.2. *Let the Schur canonical form of $A \in \mathbb{C}^{n\times n}$ be $A = QSQ^H$, where S is upper triangular and Q is unitary. Partition S into subblocks $S = \begin{bmatrix}S_{11} & S_{12} \\ & S_{22}\end{bmatrix}$, with $S_{11} \in \mathbb{C}^{k\times k}$, $k < n$. Partition Q accordingly as $Q = [Q_1, Q_2]$ such that $AQ_1 = Q_1S_{11}$ and $AQ_2 = Q_1S_{12} + Q_2S_{22}$. Let V be any matrix in $\mathbb{C}^{n\times k}$, and $\hat{Q}_1$ be any matrix in $\mathbb{C}^{n\times k}$ such that $span\{\hat{Q}_1\} \subseteq span\{Q_1\}$, (i.e., $\hat{Q}_1 = Q_1M$ for an $M \in \mathbb{C}^{k\times k}$). Then the eigenvalues of $A + \hat{Q}_1V^H$ are the union of the eigenvalues of $S_{11} + MV^HQ_1$ and the diagonal elements of S_{22}.*

Proof: Again, we construct a similarity transformation expressed in an invariant space form:

$$\begin{aligned}(A + \hat{Q}_1V^{\mathrm{H}})[Q_1, Q_2] &= [Q_1S_{11}, Q_1S_{12} + Q_2S_{22}] + Q_1MV^{\mathrm{H}}[Q_1, Q_2] \\ &= [Q_1(S_{11} + MV^{\mathrm{H}}Q_1),\quad Q_1S_{12} + Q_2S_{22} + Q_1MV^{\mathrm{H}}Q_2] \\ &= [Q_1, Q_2]\begin{bmatrix}(S_{11} + MV^{\mathrm{H}}Q_1), & S_{12} + MV^{\mathrm{H}}Q_2 \\ & S_{22}\end{bmatrix}.\end{aligned}$$

Therefore the block upper triangular matrix $\begin{bmatrix}(S_{11} + MV^{\mathrm{H}}Q_1), & S_{12} + MV^{\mathrm{H}}Q_2 \\ & S_{22}\end{bmatrix}$ is similar to $A + \hat{Q}_1V^{\mathrm{H}}$, which proves the result. ■

Besides using Schur vectors, another feature of Theorem 2.2 distinguishable from existing results is that we do not need full rank-k Schur vectors to construct the perturbation matrix $\hat{Q}_1V^{\mathrm{H}}$. The $\hat{Q}_1$ can be rank deficient.

Similar result holds when A is perturbed by a low-rank matrix constructed from the left eigen-/Jordan-/Schur- vectors of A. We only list in Theorem 2.3 the one that uses Schur vectors and skip the other cases which can be constructed similarly. Note that the perturbation vectors need to be associated with an invariant subspace of A, in the left Schur vectors case, they are the last (instead of the first) k Schur vectors.

Theorem 2.3. *Let the Schur decomposition of $A \in \mathbb{C}^{n\times n}$ be $Q^H A = SQ^H$, where S is upper triangular and Q is unitary. Partition S into subblocks $S = \begin{bmatrix} S_{11} & S_{12} \\ & S_{22} \end{bmatrix}$, with $S_{22} \in \mathbb{C}^{k\times k}$, $k < n$. Partition Q^H accordingly as $Q^H = [Q_1, Q_2]^H$ such that $Q_2^H A = S_{22} Q_2^H$ and $Q_1^H A = S_{11} Q_1^H + S_{12} Q_2^H$. Let V be any matrix in $\mathbb{C}^{n\times k}$, and $\hat{Q}_2$ be any matrix in $\mathbb{C}^{n\times k}$ such that $span\{\hat{Q}_2\} \subseteq span\{Q_2\}$. Then the eigenvalues of $A + V\hat{Q}_2^H$ are the union of the diagonal elements of S_{11} and the eigenvalues of $S_{22} + Q_2^H V M^H$, where $M \in \mathbb{C}^{k\times k}$ satisfies $\hat{Q}_2 = Q_2 M$.*

The proof is similar to that of Theorem 2.2, which can be obtained by a similarity transformation of $A + V\hat{Q}_2^{\mathrm{H}}$ to a block upper triangular matrix. Again, the transformation can be succinctly expressed in an invariant (sub)space form.

Proof:

$$
\begin{aligned}
\begin{bmatrix} Q_1^{\mathrm{H}} \\ Q_2^{\mathrm{H}} \end{bmatrix} (A + V\hat{Q}_2^{\mathrm{H}}) &= \begin{bmatrix} S_{11} Q_1^{\mathrm{H}} + S_{12} Q_2^{\mathrm{H}} \\ S_{22} Q_2^{\mathrm{H}} \end{bmatrix} + \begin{bmatrix} Q_1^{\mathrm{H}} V M^{\mathrm{H}} Q_2^{\mathrm{H}} \\ Q_2^{\mathrm{H}} V M^{\mathrm{H}} Q_2^{\mathrm{H}} \end{bmatrix} \\
&= \begin{bmatrix} S_{11} & S_{12} + Q_1^{\mathrm{H}} V M^{\mathrm{H}} \\ & S_{22} + Q_2^{\mathrm{H}} V M^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} Q_1^{\mathrm{H}} \\ Q_2^{\mathrm{H}} \end{bmatrix}
\end{aligned}
$$

Therefore the eigenvalues of $A + V\hat{Q}_2^{\mathrm{H}}$ are the eigenvalues of S_{11} and $S_{22} + Q_2^{\mathrm{H}} V M^{\mathrm{H}}$. ∎

Comment: The result in Theorem 2.3 may be obtained directly by applying Theorem 2.2 to A^{H}, noticing that A^{H} and A have the same eigenvalues.

We note that the arbitrariness of the V matrix may be used to absorb the scaling factor M for Q_i. Therefore the M matrix in Theorems 2.2 and 2.3 is not essential. We use it to emphasize that the full rank assumption of $\hat{Q}_i$ is not necessary. This M also implies that the columns in $\hat{Q}_i$ need not be of unit length. With an explicit M addressing the possible rank deficiency in $\hat{Q}_i$, the V can be any *fixed* given matrix in $\mathbb{C}^{n\times k}$.

The same proof extends to the case where Q is not unitary. In this case we construct Q^{-1} instead of Q^{H} for the similarity transformation. This is essentially Theorem 2.1 for the case where A is perturbed by $span\{\hat{X}_1\} \subseteq span\{X_1\}$. We state it as a corollary.

Corollary 2.1. *Assume the same notations and conditions as in Theorem 2.1. Then for any $\hat{X}_1 = X_1 M$, where $M \in \mathbb{C}^{k\times k}$, the eigenvalues of $A + \hat{X}_1 V^H$ are the union of the eigenvalues of $J_1 + MV^H X_1$ and the diagonal elements of J_2.*

Replacing Schur vectors by Jordan vectors in Theorem 2.3, we can readily obtain similar results for the case where partial left Jordan vectors are used for perturbation, we skip the details for this case.

3 Some applications

In this section we apply the results in the previous section to a few interesting problems.

First, we look at the Brauer's Theorem [1, 11].

Theorem 3.1. *Let the eigenvalues of $A \in \mathbb{C}^{n\times n}$ be $\lambda_i, i = 1, ..., n$. Let x_1 be the right eigenvector of A associated with λ_1. Then for any $v \in \mathbb{C}^n$, the eigenvalues of $A + x_1 v^H$ are $\lambda_1 + v^H x_1$ and $\lambda_i, i = 2, ..., n$.*

This result can be proved directly in a very simple and elementary manner, it is also straightforward to see it as a special case ($k = 1$) of Theorem 2.1 or Theorem 2.2.

Applying Theorem 2.3 we also see a corresponding Brauer's Theorem where the perturbation is by a left eigenvector.

Corollary 3.1. *Let the eigenvalues of $A \in \mathbb{C}^{n\times n}$ be $\lambda_i, i = 1, ..., n$, with y_1 the left eigenvector of A associated with λ_1. Then for any $v \in \mathbb{C}^n$, the eigenvalues of $A + v y_1^H$ are $\lambda_1 + y_1^H v$ and $\lambda_i, i = 2, ..., n$.*

We now look at the well-known Google's second eigenvalue problem, first addressed in [7]. The result is generalized in [8, 6] to cover all eigenvalues. The generalized result can be summarized as the following.

Theorem 3.2. *Let $H \in \mathbb{C}^{n\times n}$ be a row stochastic matrix with eigenvalues $\{1, \lambda_2, \lambda_3, \ldots, \lambda_n\}$. Let $\mathbf{e} = [1, 1, \ldots, 1]^T \in \mathbb{R}^n$ and let $v \in \mathbb{R}^n$ be a probability vector. Then the eigenvalues of the Google matrix $G = \alpha H + (1-\alpha)\mathbf{e}v^T$ are $\{1, \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_n\}$.*

Row stochasticity of a matrix means that each row of the matrix sums to 1. Therefore $H\mathbf{e} = \mathbf{e}$, which means $\mathbf{e}$ is a right eigenvector of H associated with the eigenvalue 1. Clearly, the eigenvalues of αH are α and $\alpha\lambda_i, i = 2, ..., n$. Applying Theorem 2.1 for the special case $k = 1$, we see that the perturbed matrix $\alpha H + (1-\alpha)\mathbf{e}v^{\mathrm{T}}$ changes only the eigenvalue of αH associated with $\mathbf{e}$ from α into $\alpha + (1-\alpha)v^{\mathrm{T}}\mathbf{e}$, which is 1 since a probability vector v satisfies $v^{\mathrm{T}}\mathbf{e} = 1$. This neat deduction however is not new, similar derivation using the Brauer's Theorem already appeared in [11, 3].

Since the magnitude of all eigenvalues of a stochastic matrix is upper bounded by 1, and the α is related to a certain probability, normally $\alpha \in (0, 1)$, we see that the eigenvalues of G satisfy $|\lambda_i(G)| = \alpha|\lambda_i(H)| \le \alpha < 1$ for $i > 1$. Theorem 3.2 readily reveals that the second eigenvalue equals to α only when 1 is a multiple eigenvalue of H.

Next we look at a generalization of Brauer's Theorem. In [2, 12, 13] the theorem is generalized to the higher rank perturbation case. We summarize the result in the following.

Theorem 3.3. *Let the eigenvalues of $A \in \mathbb{C}^{n\times n}$ be $\lambda_i, i = 1, ..., n$. Let x_i be a right eigenvector of A associated with λ_i, $(i = 1, ..., k)$. Assume that $X = [x_1, ..., x_k]$ is of rank-k. Then for any $V \in \mathbb{C}^{n\times k}$, the eigenvalues of $A + XV^H$ are the union of $\{\lambda_i\}_{i=k+1,\ldots,n}$ and the eigenvalues of $\mathrm{diag}(\lambda_1, \ldots, \lambda_k) + \mathrm{V}^H\mathrm{X}$.*

The result is a special case of Lemma 2.1: under the given conditions of Theorem 3.3, the Jordan block J_1 in Lemma 2.1 turns into the diagonal matrix $\mathrm{diag}(\lambda_1, ..., \lambda_k)$, and the corresponding Jordan vectors in X_1 are here denoted as the X matrix. Therefore Theorem 3.3 follows immediately from Lemma 2.1, which simplifies the proofs in [2, 12, 13].

We note that the full rank condition on $X = [x_1, ..., x_k]$ is not necessary for the result in Theorem 3.3 to hold. A common technique is to use the eigenvalue continuity argument [1, 11] to address the rank deficiency. Another algebraic approach to bypass the possible rank deficiency in

X is to resort to various properties of determinants, which is nicely done in [4] (see also [5, p.25]). A more straightforward way is to utilize the same technique as used in Theorem 2.2. In fact, replacing $\mathrm{diag}(\lambda_1, ..., \lambda_\mathrm{k})$ by a Schur block easily removes the condition of linear independence on X. Although the final $k \times k$ matrix appears more complicated than $\mathrm{diag}(\lambda_1, ..., \lambda_\mathrm{k}) + \mathrm{V^H X}$, the computational cost for computing the eigenvalues of a $k \times k$ matrix remains the same, unless V is specifically chosen to make $V^\mathrm{H} X$ diagonal.

Finally we look at a generalization of Theorem 3.2 which may be called the Google's k-th eigenvalue problem.

Theorem 3.4. *Let $H \in \mathbb{C}^{n\times n}$ be a row stochastic matrix with eigenvalues $\{\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n\}$, let the right eigenvectors of H associated with $\{\lambda_i\}_{i=1,\ldots,k}$ be $X = [x_1, x_2, \ldots, x_k]$, where $\lambda_1 = 1$ and $x_1 = [1, 1, \ldots, 1]^T$. Let $V = [v_1, \ldots, v_k] \in \mathbb{C}^{n\times k}$. Then the eigenvalues of the matrix $G = \alpha_0 H + \sum_{i=1}^{k} \alpha_i x_i v_i^T$ are the union of $\{\alpha_0 \lambda_i\}_{i=k+1,\ldots,n}$ and the eigenvalues of $\alpha_0 \mathrm{diag}(\lambda_1, \ldots, \lambda_\mathrm{k}) + \mathrm{diag}(\alpha_1, \ldots, \alpha_\mathrm{k}) \mathrm{V}^H \mathrm{X}$.*

The proof is straightforward by Corollary 2.1 if we notice that $G = \alpha_0 H + XMV^\mathrm{H}$ where $M = \mathrm{diag}(\alpha_1, \ldots, \alpha_\mathrm{k})$.

There are various ways to narrow the choice of V and the α_i's to make the final result resemble the succinct form of the original Google's second eigenvalue problem. However we stop here since this generalized problem does not seem to have practical value: The H in real applications is often of enormous dimension, computing the eigenvectors of H other than the known one $x_1 = [1, 1, \ldots, 1]^\mathrm{T}$ can quickly become prohibitive.

Acknowledgement: The author thanks the two anonymous referees for their constructive comments that improved the paper. He thanks one referee particularly for the comment made right after Theorem 2.3.

References

[1] A Brauer. Limits for the characteristic roots of a matrix. IV: Applications to stochastic matrices. *Duke Math. J.*, 19(1):75–91, 1952.

[2] J. Ding and Y. Gao. The eigenvalue problem of a specially updated matrix. *Appl. Math. Comput.*, 185:415–420, 2007.

[3] J. Ding and A. Zhou. Eigenvalues of rank-1 updated matrices with some applications. *Appl. Math. Lett.*, 20:1223–1226, 2007.

[4] J. Ding and A. Zhou. Characteristic polynomials of some perturbed matrices. *Appl. Math. Comput.*, 199(2):631–636, 2008.

[5] J. Ding and A. Zhou. *Nonnegative matrices, positive operators, and applications.* World Scientific Press, 2009.

[6] L. Eldén. The eigenvalues of the Google matrix. Technical Report LiTH-MAT-R-04-01, Linköping University, 2004.

[7] T. Haveliwala and S. Kamvar. The second eigenvalue of the Google matrix. Technical Report 2003-20, Stanford InfoLab, 2003.

[8] A. N. Langville and C. D. Meyer. Deeper inside PageRank. *Internet Mathematics*, I(3):335–380, 2004.

[9] A. N. Langville and C. D. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2006.

[10] S. Serra-Capizzano. Jordan canonical form of the Google matrix: A potential contribution to the PageRank computation. *SIAM J. Matrix Anal. Appl.*, 27:305–312, 2005.

[11] S. Serra-Capizzano. Google PageRanking problem: the model and the analysis. In *Proceedings of the Dagstuhl Conference in Web Retrieval and Numerical Linear Algebra*, 2007.

[12] G. Wu. Eigenvalues and Jordan canonical form of a successively rank-one updated complex matrix with applications to Google's PageRank problem. *J. Comput. Appl. Math.*, 216(2):364 – 370, 2008.

[13] Z. Xiong and B. Zheng. On the eigenvalues of a specially rank-k updated matrix. *Comput. Math. Appl.*, 57:1645–1650, 2009.