Computing the Subdominant Eigenvalues and Eigenvectors

Maria Carmela De Bonis, Ph.D.
Assistant Professor of Numerical Analysis
University of Basilicata
Potenza, ITALY

Introduction

The problem of finding eigenvalues and eigenvectors of a matrix arises in a wide variety of practical applications.

The mathematical models of many engineering problems are systems of differential and difference equations whose solutions are often expressed in terms of the eigenvalues and eigenvectors of the matrices involved in the discretization of these systems.

Furthermore, many important characteristics of physical and engineering systems, such as stability, can often be determined only by knowing the nature and location of the eigenvalues.

In particular, in several applications only the largest or the smallest eigenvalues and the corresponding eigenvectors are needed.

An Example: Convergence of Iterative Methods for solving Linear Systems

The iterative method

$$\mathbf{x}^{(k)} = D^{-1}C\mathbf{x}^{(k-1)} + D^{-1}\mathbf{b}, \quad k = 0, 1, \dots,$$

for solving the linear system

$$A\mathbf{x} = \mathbf{b}, \quad A = D + C,$$

converges to the solution $\mathbf{x}$ for an arbitrary choice of the initial approximation $\mathbf{x}^{(0)}$ if and only if the spectral radius

$$\rho(D^{-1}C) < 1.$$

Thus only the knowledge of the largest eigenvalue of $D^{-1}C$ is needed.

Dominant and Subdominant Eigenvalues

There are other applications where only the first few largest or smallest eigenvalues and the corresponding eigenvectors play an important role.

The largest eigenvalue of a matrix is also called dominant eigenvalue and the corresponding eigenvector is called dominant eigenvector.

The second eigenvalue in magnitude is called subdominant eigenvalue and the corresponding eigenvector is called subdominant eigenvector.

Analogously, the smallest eigenvalue in magnitude and the corresponding eigenvector are called least dominant eigenvalue and least dominant eigenvector, respectively; the second smallest eigenvalue in magnitude and the corresponding eigenvector are called next least dominant eigenvalue and next least dominant eigenvector, respectively.

An example: The Population Study

It is well-known that a Population System can be modeled by

$$\mathbf{p}_{k+1} = A\mathbf{p}_k, \quad k = 0, 1, \ldots,$$

where $\mathbf{p}_k$ is the population vector.

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A in decreasing order and suppose that A has a set of independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

An example: The Population Study

The population $\mathbf{p}_k$ at any time $k > 0$ is given by

$$\mathbf{p}_k = \alpha_1 \lambda_1^k \mathbf{v}_1 + \cdots + \alpha_n \lambda_n^k \mathbf{v}_n,$$

where $\mathbf{p}_0 = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$.

Then

- if $|\lambda_1| < 1$ the population decreases to zero as k becomes large;
- if $|\lambda_1| > 1$
  - there is a long-term growth in the population;
  - the original population approaches a final distribution that is defined by the eigenvector of the dominant eigenvalue;
  - the subdominant eigenvalue determines how fast the original population distribution is approaching the final distribution.
- if $|\lambda_1| = 1$ over the long-term there is neither growth nor decay in the population.

Power Method and Inverse Power Method

You have already studied the Power method for approximating the dominant eigenpair, i.e., the dominant eigenvalue and the corresponding eigenvector.

By strictly modifying the Power Method, it is possible to deduce the Inverse Power method for approximating the least dominant eigenpair, i.e., the least dominant eigenvalue and the corresponding eigenvector.

Moreover, a strictly modification of the Inverse Power method let us to improve a known approximation of one of the eigenvalues and to compute the corresponding eigenvector.

Inverse Power Method

Let A be a diagonalizable $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote its eigenvalues in decreasing order such that

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n|,$$

i.e., the smallest eigenvalue λ_n has algebraic multiplicity 1 and other eigenvalues with the same modulus do not exist. Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ denote the corresponding eigenvectors.

The matrix A^{-1} has eigenvalues $\frac{1}{\lambda_i}, i = 1, \ldots, n$, such that

$$\frac{1}{|\lambda_n|} > \frac{1}{|\lambda_{n-1}|} \geq \cdots \geq \frac{1}{|\lambda_1|}$$

and $\frac{1}{\lambda_n}$ is its largest eigenvalue.

Then, in order to approximate the smallest eigevalue of A it is possible to apply the Power Method to the matrix A^{-1}.This is the reason why the corresponding method is known as Inverse Power Method.

Inverse Power Method

Starting from the Power Method:

$$\begin{cases} \mathbf{u}_k = A\mathbf{t}_{k-1} \\ \\ \mathbf{t}_k = \dfrac{\mathbf{u}_k}{\max\{\mathbf{u}_k\}} \end{cases} \qquad k = 1, 2, \dots$$

where $\mathbf{t}_0 = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_n \mathbf{x}_n$ with $\|\mathbf{t}_0\|_\infty = 1$,

$$\lim_{k\to+\infty} \max\{\mathbf{u}_k\} = \lambda_1, \quad \lim_{k\to+\infty} \mathbf{t}_k = \mathbf{x}_1,$$

the Power Method applied to A^{-1}, becomes

$$\begin{cases} A\mathbf{u}_k = \mathbf{t}_{k-1} \\ \\ \mathbf{t}_k = \dfrac{\mathbf{u}_k}{\max\{\mathbf{u}_k\}} \end{cases} \qquad k = 1, 2, \dots$$

$$\lim_{k\to+\infty} \max\{\mathbf{u}_k\} = \frac{1}{\lambda_n}, \quad \lim_{k\to+\infty} \mathbf{t}_k = \mathbf{x}_1.$$

Inverse Power Method

Note that:

- Since for any $k = 1, 2, \ldots$ it is necessary to solve the linear system $A\mathbf{u}_k = \mathbf{t}_{k-1}$, in order to reduce the computational cost, it is better to perform the factorization $\Pi A = LU$ before starting the iterations and then solve the triangular systems:

$$\begin{cases} L\mathbf{y}_k = \Pi \mathbf{t}_{k-1} \\ \\ U\mathbf{u}_k = \mathbf{y}_k \end{cases} \qquad k = 1, 2, \ldots$$

- If the method requires m iterations, the computational cost is $\mathcal{O}\left(\frac{n^3}{3} + mn^2\right)$.
- The convergence order is $\left|\frac{\lambda_n}{\lambda_{n-1}}\right|^k$, that is the larger the distance between λ_{n-1} and λ_n, the faster the convergence.

Inverse Power Method to improve a known approximation of an eigenvalue and compute its eigenvector

Let A be a diagonalizable $n \times n$ matrix and assume that μ is an approximation of the eigenvalue λ of A. We can write

$$(A - \mu I)\mathbf{x} = A\mathbf{x} - \mu\mathbf{x} = \lambda\mathbf{x} - \mu\mathbf{x} = (\lambda - \mu)\mathbf{x}, \quad \text{being} \quad A\mathbf{x} = \lambda\mathbf{x}.$$

Then

- $\lambda - \mu$ is an eigenvalue of the matrix $A - \mu I$ and $\mathbf{x}$ is the corresponding eigenvector.
- $\eta = \frac{1}{\lambda-\mu}$ is an eigenvalue of the matrix $(A - \mu I)^{-1}$ and $\mathbf{x}$ is the corresponding eigenvector.
- If μ is very close to λ then $\eta = \frac{1}{\lambda-\mu}$ is the dominant eigenvalue of $(A - \mu I)^{-1}$.

Therefore the computation of $\eta = \frac{1}{\lambda-\mu}$ and then of $\lambda = \mu + \frac{1}{\eta}$ can be performed applying the Inverse Power Method to the matrix $(A - \mu I)^{-1}$.

Inverse Power Method to improve a known approximation of an eigenvalue and compute its eigenvector

Note that:

- If the approximation μ of λ is not sufficiently good, the convergence of the method becomes very slow.
- If the approximation μ is very close to λ, the matrix $(A - \mu I)$ is obviously ill conditioned. Consequently, this ill-conditioning might affect the computed approximations of the eigenvector. Fortunately, in practice the ill-conditioning of the matrix $(A - \mu I)$ is exactly what we want: the error at each iteration grows toward the direction of the eigenvector and, it is the direction of the eigenvector in which we are interested.Wilkinson (1965, pp. 620-621) has remarked that in practice $\mathbf{u}_k$ is remarkably close to the solution of

$$(A - \mu I + F)\mathbf{u}_k = \mathbf{t}_{k-1}, \quad F \text{ small.}$$

Matrix Deflation

Once the dominant eigenpair (the dominant eigenvalue and the corresponding eigenvector) or the least dominant eigenpair (the least dominant eigenvalue and the corresponding eigenvector) have been computed, the subdominant eigenvalue or the next least dominant eigenvalue can be computed by using deflation.

The basic idea behind deflation is to replace the original matrix by another matrix of the same or lesser dimension computed by using the dominant or the least dominant eigenpair.

Such matrix called deflated matrix has the same eigenvalues as the original one except the one used to deflate.

Deflation Schemes

Two deflation schemes are commonly used in practice:

- Hotelling Deflation
- Householder Deflation

Hotelling Deflation for approximating the subdominant eigenpair

Let A be a diagonalizable $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote its eigenvalues in decreasing order. Then λ_1 and λ_2 are the dominant and subdominant eigenvalues of A, respectively.

The Hotelling Deflation is a process that replaces the original matrix $A =: A_1$ by another matrix A_2, of the same order of A_1, having as eigenvalues $0, \lambda_2, \ldots, \lambda_n$. Then λ_2 is the dominant eigenvalue of A_2.

The construction of A_2 is done using the spectral decomposition of the matrix A.

In order to introduce it we premise some definitions.

Right and Left eigenvectors

Let A be a diagonalizable $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote its eigenvalues.

A has a complete set of n linearly independent right eigenvectors $\{\mathbf{x}_i\}_{i=1,\ldots,n}$ as well as left eigenvectors $\{\mathbf{y}_i\}_{i=1,\ldots,n}$ such that

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad \mathbf{y}_i^T A = \lambda_i \mathbf{y}_i^T, \quad i = 1, \ldots, n.$$

The name right eigenvector is often abbreviated to just eigenvector, in which case the “right” qualifier is tacitly understood.

The left eigenvectors of A are the right eigenvectors of its transpose A^T, in fact, from the latter equality, we deduce

$$A^T \mathbf{y}_i = \lambda_i \mathbf{y}_i, \quad i = 1, \ldots, n.$$

Biorthonormal Right and Left eigenvectors

A set of left and right eigenvectors $\{\mathbf{x}_i\}_{i=1,\ldots,n}$ and $\{\mathbf{y}_i\}_{i=1,\ldots,n}$ is said to be biorthonormal with respect to A if they verify

$$\mathbf{y}_i^T \mathbf{x}_j = \mathbf{x}_i^T \mathbf{y}_j = \delta_{i,j}, \quad i, j = 1, \ldots, n,$$

where $\delta_{i,j}$ denotes the Kronecker delta.

Consequently, if the right and left eigenvectors $\{\mathbf{x}_i\}_{i=1,\ldots,n}$ and $\{\mathbf{y}_i\}_{i=1,\ldots,n}$ are biorthonormalized, we have

$$\mathbf{y}_i^T A \mathbf{x}_i = \lambda_i, \quad \mathbf{x}_i^T A^T \mathbf{y}_i = \lambda_i, \quad i = 1, \ldots, n.$$

Biorthonormal Right and Left eigenvectors

The biorthonormalized eigenvectors $\{\mathbf{x}_i\}_{i=1,\dots,n}$ and $\{\mathbf{y}_i\}_{i=1,\dots,n}$ can be collected into two $n \times n$ eigenvector matrices built by stacking eigenvectors as columns:

$$Y = [\mathbf{y}_1 \mathbf{y}_2 \cdots \mathbf{y}_n], \quad X = [\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n].$$

These matrices satisfy

$$Y^T X = X^T Y = I, \quad X^{-1} = Y^T, \quad Y^{-1} = X^T.$$

Denoting by Λ the diagonal matrix having as diagonal entries the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A, we get

$$Y^T A X = X^{-1} A X = \Lambda$$

and

$$X^T A^T Y = Y^{-1} A^T Y = \Lambda.$$

Biorthonormal Right and Left eigenvectors

The biorthonormalized eigenvectors $\{\mathbf{x}_i\}_{i=1,\ldots,n}$ and $\{\mathbf{y}_i\}_{i=1,\ldots,n}$ can be collected into two $n \times n$ eigenvector matrices built by stacking eigenvectors as columns:

$$Y^T = [\mathbf{y}_1^T \mathbf{y}_2^T \cdots \mathbf{y}_n^T], \quad X = [\mathbf{x}_1 \mathbf{x}_2 \cdots \mathbf{x}_n].$$

These matrices satisfy

$$Y^T X = X^T Y = I, \quad X^{-1} = Y^T, \quad Y^{-1} = X^T.$$

Denoting by Λ the diagonal matrix having as diagonal entries the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A, we get

$$Y^T A X = X^{-1} A X = \Lambda$$

and

$$X^T A^T Y = Y^{-1} A^T Y = \Lambda.$$

Spectral decomposition of A

Starting from

$$X^{-1}AX = \Lambda,$$

multiplying both sides of it on the left by X we get

$$AX = X\Lambda.$$

Moreover, multiplying both sides of the latter equation on the right by $X^{-1} = Y^T$, we deduce the so-called spectral decomposition of A

$$A = X\Lambda Y^T = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{y}_i^T.$$

Symmetric Real Matrices

Symmetric real matrices are diagonalizable and thus posses complete set linearly independent eigenvectors. Right and left eigenvectors coalesce: $\mathbf{y}_i = \mathbf{x}_i, i = 1, \ldots, n$, and such qualifiers may be omitted.

The eigenvectors $\mathbf{x}_i, i = 1, \ldots, n$, can be orthonormalized so that

$$\mathbf{x}_i^T \mathbf{x}_j = \delta_{i,j}, \quad \mathbf{x}_i^T A \mathbf{x}_i = \lambda_i, \quad i, j = 1, \ldots, n.$$

If eigenvectors are stacked as columns of a matrix X, the foregoing orthonormality conditions can be compactly stated as

$$X^T X = I, \quad X^T A X = \Lambda$$

and the spectral decomposition of A becomes

$$A = X \Lambda X^T = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{x}_i^T.$$

The above properties hold true also for the more general class of Hermitian matrices, replacing the transpose operator by the conjugate transpose.

Hotelling Deflation for approximating the subdominant eigenpair

Suppose that the dominant eigenvalue λ_1 with associated right eigenvector $\mathbf{x}_1$ and left eigenvector $\mathbf{y}_1$ are known.

Taking into account the spectral decomposition of A

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{y}_i^T,$$

the deflated matrix A_2 is computed as follows

$$A_2 = A - \lambda_1 \mathbf{x}_1 \mathbf{y}_1^T.$$

Obviously this has a spectral decomposition identical to the one of A except that λ_1 is replaced by 0. Then $0, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A_2 and λ_2 is its dominant eigenvalue.

If A is symmetric, $\mathbf{y}_1 = \mathbf{x}_1$ and also A_2 is symmetric.

Algorithm: Hotelling Deflation for approximating the subdominant eigenpair

Step 1 Compute the dominant eigenvalue λ_1 and the dominant right eigenvector $\mathbf{x}_1$ of the matrix A using the power method

Step 2 Compute the dominant right eigenvector $\mathbf{y}_1$ of the matrix A^T using the inverse power method. y_1 is also the dominant left eigenvector of A.

Step 3 Compute the biorthonormalized eigenvectors

$$\begin{cases} \mathbf{x}_1 = \frac{\mathbf{x}_1}{\sqrt{\mathbf{y}_1^T \mathbf{x}_1}}, & \mathbf{y}_1 = \frac{\mathbf{y}_1}{\sqrt{\mathbf{y}_1^T \mathbf{x}_1}} \qquad \text{if } \mathbf{y}_1^T \mathbf{x}_1 > 0 \\ \mathbf{x}_1 = -\frac{\mathbf{x}_1}{\sqrt{-\mathbf{y}_1^T \mathbf{x}_1}}, & \mathbf{y}_1 = \frac{\mathbf{y}_1}{\sqrt{-\mathbf{y}_1^T \mathbf{x}_1}} \quad \text{if } \mathbf{y}_1^T \mathbf{x}_1 < 0 \end{cases}$$

Step 4 Compute $A_2 = A - \lambda_1 \mathbf{x}_1 \mathbf{y}_1^T$

Step 5 Compute the dominant eigenvalue λ_2 and the dominant right eigenvector $\mathbf{x}_2$ of the matrix A_2 using the power method. λ_2 and $\mathbf{x}_2$ are the subdominant eigenpair of A.

Computational Cost of Hotelling Deflation for approximating the subdominant eigenpair

Step 1 Compute the dominant eigenvalue λ_1 and the dominant right eigenvector $\mathbf{x}_1$ of the matrix A using the power method. $\mathcal{O}(m_1 n^2)$

Step 2 Compute the dominant right eigenvector $\mathbf{y}_1$ of the matrix A^T using the inverse power method. y_1 is also the dominant left eigenvector of A. $\mathcal{O}\left(\frac{n^3}{3} + m_2 n^2\right)$

Step 3 Compute the biorthonormalized eigenvectors $\mathcal{O}(n)$

$$\begin{cases} \mathbf{x}_1 = \frac{\mathbf{x}_1}{\sqrt{\mathbf{y}_1^T \mathbf{x}_1}}, \quad \mathbf{y}_1 = \frac{\mathbf{y}_1}{\sqrt{\mathbf{y}_1^T \mathbf{x}_1}} & \text{if } \mathbf{y}_1^T \mathbf{x}_1 > 0 \\ \mathbf{x}_1 = -\frac{\mathbf{x}_1}{\sqrt{-\mathbf{y}_1^T \mathbf{x}_1}}, \quad \mathbf{y}_1 = \frac{\mathbf{y}_1}{\sqrt{-\mathbf{y}_1^T \mathbf{x}_1}} & \text{if } \mathbf{y}_1^T \mathbf{x}_1 < 0 \end{cases}$$

Step 4 Compute $A_2 = A - \lambda_1 \mathbf{x}_1 \mathbf{y}_1^T$ $\mathcal{O}(n^2)$

Step 5 Compute the dominant eigenvalue λ_2 and the dominant right eigenvector $\mathbf{x}_2$ of the matrix A_2 using the power method. λ_2 and $\mathbf{x}_2$ are the subdominant eigenpair of A. $\mathcal{O}(m_3 n^2)$

Computational Cost of Hotelling Deflation for approximating the subdominant eigenpair

Summing up the computational cost of the method is

$$\mathcal{O}\left(\frac{n^3}{3} + m_1 n^2 + m_2 n^2 + m_3 n^2\right).$$

If the matrix A is symmetric the computational cost becomes

$$\mathcal{O}\left(m_1 n^2 + m_3 n^2\right).$$

Hotelling Deflation cannot be used for approximating the next least dominant eigenpair

Suppose that the least dominant eigenvalue λ_n with associated right eigenvector $\mathbf{x}_n$ and left eigenvector $\mathbf{y}_n$ are known.

Taking into account the spectral decomposition of A

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{y}_i^T,$$

the deflated matrix A_2 becomes

$$A_2 = A - \lambda_n \mathbf{x}_n \mathbf{y}_n^T.$$

Then $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, 0$ are the eigenvalues of A_2 and 0 is its least dominant eigenvalue!

Householder Deflation for approximating the subdominant eigenpair

Let A be a diagonalizable $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote its eigenvalues in decreasing order. Then λ_1 and λ_2 are the dominant and subdominant eigenvalues of A, respectively.

The Householder Deflation constructs a deflated matrix A_2, of order $n-1$, using a similarity transformation on $A := A_1$.

A_2 has the same eigenvalues of A_1 except for λ_1. Then λ_2 is the dominant eigenvalue of A_2.

The construction of A_2 is done using a Householder matrix.

In order to introduce it we premise some definitions.

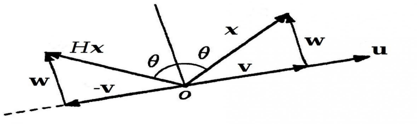
Householder Matrix

A Householder matrix is a matrix of the following form

$$H = I - 2\mathbf{u}\mathbf{u}^T,$$

where $\mathbf{u}$ is a nonzero vector such that $\|\mathbf{u}\|_2 = 1$. It is also known as elementary reflector or Householder transformation.

Next figure gives a geometric interpretation of a Householder transformation.



Let us consider a nonzero vector $\mathbf{x}$ and let us denote by $\mathbf{v}$ its component parallel to $\mathbf{u}$ and by $\mathbf{w}$ its component orthogonal to $\mathbf{u}$. Then

$$\mathbf{v} = \alpha \mathbf{u}, \alpha \in \mathbb{R}, \quad \mathbf{u}^T \mathbf{w} = 0.$$

Letting $\mathbf{x} = \mathbf{v} + \mathbf{w}$, we get

$$H\mathbf{x} = H\mathbf{v} + H\mathbf{w} = \mathbf{v} - 2\mathbf{u}\mathbf{u}^T\mathbf{v} + \mathbf{w} - 2\mathbf{u}\mathbf{u}^T\mathbf{w} = \mathbf{v} + \mathbf{w} - 2\mathbf{u}\mathbf{u}^T\mathbf{v}$$

and, since $\mathbf{u}(\mathbf{u}^T\mathbf{v}) = \alpha\mathbf{u} = \mathbf{v}$, we have

$$H\mathbf{x} = -\mathbf{v} + \mathbf{w},$$

i.e., *H reflects $\mathbf{x}$ with respect to the axis through the origin perpendicular to $\mathbf{u}$.*

Householder Matrix

The Householder matrix H verifies the following properties:

- H is symmetric, then $H = H^T$;
- $H^2 = I$, then H reflects a vector to the other side of the axis perpendicular to $\mathbf{u}$ and H^2 reflects a vector back to itself;
- H is orthogonal, then $HH^T = H^T H = I$.
- $\det(H) = -1$. Letting $P = \{\mathbf{w} \in \mathbb{R}^n \ : \ \mathbf{u}^T \mathbf{w} = 0\}$, P is a $(n-1)$-dimensional subspace of $\mathbb{R}^n$ having $n-1$ linearly independent vectors $\mathbf{y}_1, \ldots, \mathbf{y}_{n-1}$. It is $H\mathbf{y}_i = \mathbf{y}_i, \ i = 1, \ldots, n-1$, then 1 is an $(n-1)$-fold eigenvalue. Moreover, since $H\mathbf{u} = -\mathbf{u}$, -1 is a simple eigevalue of H.

Householder Matrix

The importance of Householder matrices lies in the fact that they can also be used to create zeros in a vector.

Teorema

Given a nonzero vector $\mathbf{x} = (x_1, \ldots, x_n) \neq \mathbf{e}_1 = (1, 0, \ldots, 0)^T$, *the elementary reflector*

$$H = I - \frac{\mathbf{u}\mathbf{u}^T}{\beta},$$

with $\mathbf{u} = \mathbf{x} + \sigma \mathbf{e}_1$, $\beta = \frac{\mathbf{u}^T\mathbf{u}}{2}$ *and* $\sigma = \pm\|\mathbf{x}\|_2$ *is such that*

$$H\mathbf{x} = -\sigma \mathbf{e}_1.$$

Proof

It is easy to verify that

$$H\mathbf{x} = \mathbf{x} - \frac{\mathbf{u}^T\mathbf{x}}{\beta}\mathbf{u}.$$

Since

$$\mathbf{u}^T\mathbf{x} = (\mathbf{x}+\sigma\mathbf{e}_1)^T\mathbf{x} = (\mathbf{x}^T+\sigma\mathbf{e}_1^T)\mathbf{x} = \mathbf{x}^T\mathbf{x}+\sigma\mathbf{e}_1^T\mathbf{x} = \sigma^2+\sigma x_1 = \sigma(\sigma+x_1)$$

and

$$\beta = \frac{\mathbf{u}^T\mathbf{u}}{2} = \frac{1}{2}(\mathbf{x}+\sigma\mathbf{e}_1)^T(\mathbf{x}+\sigma\mathbf{e}_1) = \frac{1}{2}(\mathbf{x}^T\mathbf{x}+2\sigma x_1+\sigma^2) = \sigma(\sigma+x_1),$$

we get

$$H\mathbf{x} = \mathbf{x} - \mathbf{u} = \mathbf{x} - (\mathbf{x} + \sigma\mathbf{e}_1) = -\sigma\mathbf{e}_1. \qquad \square$$

Remarks

- The couple $(\mathbf{u}, \beta)$ of $n+1$ real numbers is sufficient to uniquely determine the matrix H having n^2 entries. The Householder matrix itself does not have to be formed in practice.
- The matrix-vector product with a Householder matrix can be performed just by using the couple $(\mathbf{u}, \beta)$ with only $\mathcal{O}(n)$ flops. In fact, for any nonzero vector $\mathbf{a}$,

$$H\mathbf{a} = (I - \frac{\mathbf{u}\mathbf{u}^T}{\beta})\mathbf{a} = \mathbf{a} - \frac{\mathbf{u}^T\mathbf{a}}{\beta}\mathbf{u}.$$

  The usual matrix-vector product requires $\mathcal{O}(n^2)$ flops.
- The matrix-matrix products with a Householder matrix can be performed just by using the couple $(\mathbf{u}, \beta)$ with only $\mathcal{O}(n^2)$ flops. In fact,

$$HA = H(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = (H\mathbf{a}_1, H\mathbf{a}_2, \ldots, H\mathbf{a}_n)$$

  and

$$AH = (HA^T)^T.$$

  The usual matrix-matrix product requires $\mathcal{O}(n^3)$ flops.

Householder Deflation for approximating the subdominant eigenpair

Suppose that the dominant eigenvalue λ_1 and the corresponding eigenvector $\mathbf{x}_1$ are known.

The method is based upon the following result:

Teorema

Let $(\lambda_1, \mathbf{x}_1)$ be the dominant eigenpair of the $n \times n$ matrix A and let H be the Householder matrix such that $H\mathbf{x}_1 = -\sigma \mathbf{e}_1, \sigma \in \mathbb{R}$. Then

$$HAH = \begin{pmatrix} \lambda_1 & \mathbf{b}^T \\ \mathbf{0} & A_2 \end{pmatrix},$$

where A_2 is a $(n-1) \times (n-1)$ matrix and its eigenvalues are the same as those of A except for λ_1.
In particular the dominant eigenvalue of A_2 is λ_2, which is the subdominant eigenvalue of A.

Proof

Since $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$ and $H^2 = I$, we have

$$HAHH\mathbf{x}_1 = \lambda_1 H\mathbf{x}_1$$

and, taking into account that $H\mathbf{x}_1 = -\sigma\mathbf{e}_1$ we get

$$HAH\sigma\mathbf{e}_1 = \lambda_1\sigma\mathbf{e}_1$$

and, then

$$HAH\mathbf{e}_1 = \lambda_1\mathbf{e}_1.$$

This means that the first column of HAH is λ_1 times the first column of the identity matrix I. Thus HAH must have the form

$$HAH = \begin{pmatrix} \lambda_1 & \mathbf{b}^T \\ \mathbf{0} & A_2 \end{pmatrix}.$$

Moreover, since
$det(A - \lambda I) = \det(HAH - \lambda I) = \det(\lambda_1 - \lambda)\det(A_2 - \lambda I)$, it follows that the eigenvalues of A_2 are the same as those of A minus λ_1. □

Algorithm: Householder Deflation for approximating the subdominant eigenpair

Step 1 Compute the dominant eigenvalue λ_1 and the dominant eigenvector $\mathbf{x}_1$ of the matrix A using the power method

Step 2 Compute the Householder matrix H such that $H\mathbf{x}_1 = -\sigma \mathbf{e}_1$.

Step 3 Compute the matrix HAH

Step 4 Compute the matrix A_2, dropping the first row and the first column of HAH

Step 5 Compute the dominant eigenvalue λ_2 of the matrix A_2 using the power method. λ_2 is the subdominant eigenvalue of A

Step 6 Compute the dominant eigenvector $\mathbf{x}_2$ of the matrix A using the inverse power method.

Computational cost of Householder Deflation for approximating the subdominant eigenpair

Step 1 Compute the dominant eigenvalue λ_1 and the dominant eigenvector $\mathbf{x}_1$ of the matrix A using the power method. $\mathcal{O}\left(m_1 n^2\right)$

Step 2 Compute the Householder matrix H such that $H\mathbf{x}_1 = -\sigma\mathbf{e}_1$. $\mathcal{O}\left(n\right)$

Step 3 Compute the matrix HAH $\mathcal{O}\left(2n^2\right)$

Step 4 Compute the matrix A_2, dropping the first row and the first column of HAH

Step 5 Compute the dominant eigenvalue λ_2 of the matrix A_2 using the power method. λ_2 is the subdominant eigenvalue of A. $\mathcal{O}\left(m_2 n^2\right)$

Step 6 Compute the dominant eigenvector $\mathbf{x}_2$ of the matrix A using the inverse power method. $\mathcal{O}\left(\frac{n^3}{3} + m_3 n^2\right)$

Computational cost of Householder Deflation for approximating the subdominant eigenpair

Summing up the computational cost of the method is

$$\mathcal{O}\left(\frac{n^3}{3} + m_1 n^2 + m_2 n^2 + m_3 n^2\right).$$

Algorithm: Householder Deflation for approximating the next least dominant eigenpair

Step 1 Compute the least dominant eigenvalue λ_n and the least dominant eigenvector $\mathbf{x}_n$ of the matrix A using the inverse power method.

Step 2 Compute the Householder matrix H such that $H\mathbf{x}_n = -\sigma \mathbf{e}_1$.

Step 3 Compute the matrix HAH

Step 4 Compute the matrix A_2, dropping the first row and the first column of HAH

Step 5 Compute the least dominant eigenvalue λ_{n-1} of the matrix A_2 using the inverse power method. λ_{n-1} is the next least dominant eigenvalue of A.

Step 6 Compute the next least dominant eigenvector $\mathbf{x}_{n-1}$ of the matrix A using the inverse power method.

Computational cost of Householder Deflation for approximating the next least dominant eigenpair

Step 1 Compute the least dominant eigenvalue λ_n and the least dominant eigenvector $\mathbf{x}_n$ of the matrix A using the inverse power method. $\mathcal{O}\left(\frac{n^3}{3} + m_1 n^2\right)$

Step 2 Compute the Householder matrix H such that $H\mathbf{x}_n = -\sigma \mathbf{e}_1$. $\mathcal{O}\left(n\right)$

Step 3 Compute the matrix HAH. $\mathcal{O}\left(2n^2\right)$

Step 4 Compute the matrix A_2, dropping the first row and the first column of HAH.

Step 5 Compute the least dominant eigenvalue λ_{n-1} of the matrix A_2 using the inverse power method. λ_{n-1} is the next least dominant eigenvalue of A. $\mathcal{O}\left(\frac{n^3}{3} + m_2 n^2\right)$

Step 6 Compute the next least dominant eigenvector $\mathbf{x}_{n-1}$ of the matrix A using the inverse power method. $\mathcal{O}\left(\frac{n^3}{3} + m_3 n^2\right)$

Computational cost of Householder Deflation for approximating the next least dominant eigenpair

Summing up the computational cost of the method is

$$\mathcal{O}\left(2\frac{n^3}{3} + m_1 n^2 + m_2 n^2 + m_3 n^2\right).$$

Conclusions

- Approximation of the subdominant eigenpair
  - Hotelling Deflation $\mathcal{O}\left(\frac{n^3}{3} + m_1 n^2 + m_2 n^2 + m_3 n^2\right)$ flops. If A is symmetric $\mathcal{O}\left(m_1 n^2 + m_3 n^2\right)$ flops.
  - Householder Deflation $\mathcal{O}\left(\frac{n^3}{3} + m_1 n^2 + m_2 n^2 + m_3 n^2\right)$ flops
- Approximation of the next least dominant eigenpair
  - Householder Deflation $\mathcal{O}\left(2\frac{n^3}{3} + m_1 n^2 + m_2 n^2 + m_3 n^2\right)$

The Hotelling Deflation has rather poor numerical stability (Wilkinson, 1965, pp. 585).