
1 The Power Method and the Dominant Eigenvalue

We know eigenvalues as the roots of a characteristic polynomial of a matrix. In practice, the problem of finding eigenvalues is virtually never solved by finding roots of the characteristic polynomial, as this task is computationally prohibitive for large matrices. Iterative methods are used instead. In this section we describe the Power Method and some variations. In [QR Factorization](#) we will touch upon a better iterative method for approximating eigenvalues of a non-singular matrix called the QR-algorithm. The methods we present in this text are a mere introduction to the iterative methods used to compute eigenvalues and eigenvectors.

Finding the Dominant Eigenvalue using the Power Method

In Exploration ?? of [Diagonalizable Matrices and Multiplicity](#), our initial rationale for diagonalizing matrices was to be able to compute the powers of a square matrix, and the eigenvalues were needed to do this. In this section, we are interested in efficiently computing eigenvalues, and it may come as no surprise that the first method we discuss uses the powers of a matrix.

Definition 1. An eigenvalue λ of an $n \times n$ matrix A is called a *dominant eigenvalue* if λ has multiplicity 1, and

$$|\lambda| > |\mu| \quad \text{for all eigenvalues } \mu \neq \lambda$$

Any corresponding eigenvector is called a *dominant eigenvector* of A .

When such an eigenvalue exists, one technique for finding it is as follows: Let $\mathbf{x}_0$ in $\mathbb{R}^n$ be a first approximation to a dominant eigenvector λ , and compute successive approximations $\mathbf{x}_1, \mathbf{x}_2, \dots$ as follows:

$$\mathbf{x}_1 = A\mathbf{x}_0 \quad \mathbf{x}_2 = A\mathbf{x}_1 \quad \mathbf{x}_3 = A\mathbf{x}_2 \quad \dots$$

In general, we define

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{for each } k \geq 0$$

We will see in this section that if the first estimate $\mathbf{x}_0$ is good enough, these vectors $\mathbf{x}_n$ will approximate the dominant eigenvector λ . This technique is called the *Power Method* (because $\mathbf{x}_k = A^k \mathbf{x}_0$ for each $k \geq 1$). Observe that if $\mathbf{z}$ is any eigenvector corresponding to λ , then

$$\frac{\mathbf{z} \cdot (A\mathbf{z})}{\|\mathbf{z}\|^2} = \frac{\mathbf{z} \cdot (\lambda\mathbf{z})}{\|\mathbf{z}\|^2} = \lambda$$

Because the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n, \dots$ approximate dominant eigenvectors, this suggests that we define the *Rayleigh quotients* as follows:

$$r_k = \frac{\mathbf{x}_k \cdot \mathbf{x}_{k+1}}{\|\mathbf{x}_k\|^2} \quad \text{for } k \geq 1$$

Then the numbers r_k approximate the dominant eigenvalue λ .

Exploration 1 Use the power method to approximate a dominant eigenvector and eigenvalue of $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$.

The eigenvalues of A are 2 and -1 , with eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$, as we can verify using Octave.

```
A=[1 1; 2 0]
[EIGENVECTORS,EIGENVALUES] = eig(A)
```

[Link to code.](#)

For more information on using Octave to find eigenvalues, please refer to the [Octave Tutorial](#) section of this text.

If we take $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as the first approximation and compute $\mathbf{x}_1, \mathbf{x}_2, \dots$, successively, from $\mathbf{x}_1 = A\mathbf{x}_0, \mathbf{x}_2 = A\mathbf{x}_1, \dots$, the result is

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 11 \\ 10 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 21 \\ 22 \end{bmatrix}, \dots$$

These vectors are approaching scalar multiples of the dominant eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Moreover, we will find that the Rayleigh quotients are

$$r_1 = \frac{7}{5}, r_2 = \frac{27}{13}, r_3 = \frac{115}{61}, r_4 = \frac{451}{221}, \dots$$

and these are approaching the dominant eigenvalue 2.

1. Verify the above calculations using the following Octave code.
2. Try a different starting vector. Do we still get convergence?
3. What happens if the starting vector is changed to $\mathbf{x}_0 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$? Why doesn't this method work in this case?
4. Try using $\mathbf{x}_0 = \begin{bmatrix} 1 \\ -2.01 \end{bmatrix}$. You may need to increase the number of iterations to see convergence.

```
% The Power Method
A=[1 1; 2 0]; %Input a square matrix
x_0=[1;0]; %Input an initial vector
maxit =5; %Choose number iterations
x = x_0;
count = 0;
r=0;
while count < maxit
    y=x;
    x = A*y;
    if count > 0;
        r=(transpose(y)*x)/(transpose(y)*y);
        RQUOTIENTS(count)=r;
    end
    count = count + 1;
    VECTORS(:,count)=x;
end
VECTORS
RQUOTIENTS
```

[Link to code.](#)

To see why the power method works, let $\lambda_1, \lambda_2, \dots, \lambda_m$ be eigenvalues of A with λ_1 dominant and let $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m$ be corresponding eigenvectors. What is required is that the first approximation $\mathbf{x}_0$ be a linear combination of these eigenvectors:

$$\mathbf{x}_0 = a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2 + \dots + a_m \mathbf{y}_m \quad \text{with } a_1 \neq 0$$

If $k \geq 1$, the fact that $\mathbf{x}_k = A^k \mathbf{x}_0$ and $A^k \mathbf{y}_i = \lambda_i^k \mathbf{y}_i$ for each i gives

$$\mathbf{x}_k = a_1 \lambda_1^k \mathbf{y}_1 + a_2 \lambda_2^k \mathbf{y}_2 + \dots + a_m \lambda_m^k \mathbf{y}_m \quad \text{for } k \geq 1$$

Hence

$$\frac{1}{\lambda_1^k} \mathbf{x}_k = a_1 \mathbf{y}_1 + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{y}_2 + \dots + a_m \left(\frac{\lambda_m}{\lambda_1} \right)^k \mathbf{y}_m$$

The right side approaches $a_1 \mathbf{y}_1$ as k increases because λ_1 is dominant

$\left(\left| \frac{\lambda_i}{\lambda_1} \right| < 1 \text{ for each } i > 1 \right)$. Because $a_1 \neq 0$, this means that $\mathbf{x}_k$ approximates the dominant eigenvector $a_1 \lambda_1^k \mathbf{y}_1$.

The power method requires that the first approximation $\mathbf{x}_0$ be a linear combination of eigenvectors. In Example ??, the eigenvectors form a basis of $\mathbb{R}^2$. But even in this case, the method fails if $a_1 = 0$, where a_1 is the coefficient of the dominant eigenvector (as we saw in part ??). In general, the rate of convergence is quite slow if any of the ratios $\left| \frac{\lambda_i}{\lambda_1} \right|$ is near 1 (as we saw in part ??).

Exploration 2 Use the Power Method to approximate a dominant eigenvector and eigenvalue of $A = \begin{bmatrix} 9 & 2 & 8 \\ 2 & -6 & -2 \\ -8 & 2 & -5 \end{bmatrix}$

You can use Octave to employ the power method.

```
% The Power Method
A=[9 2 8; 2 -6 -2; -8 2 -5]; %Input a square matrix
xinit=[1;1;1];
%Input an initial vector
maxit =15;
%Choose number iterations
x = xinit;
count = 0;
r=0;
while count < maxit
    y=x;
    x = A*y;
    if count > 0;
        r=(transpose(y)*x)/(transpose(y)*y);
        RQUOTIENTS(count)=r;
    end
    count = count + 1;
    VECTORS(:,count)=x;
```

```

end
VECTORS
RQUOTIENTS
[EIGENVECTORS, EIGENVALUES]=eig(A)
VECTORS(:,maxit)/norm(VECTORS(:,maxit))

```

[Link to code.](#)

We observe that the Rayleigh quotients approach -3 , which is the dominant eigenvalue. We can confirm this using the command 'eig'. To see that our vectors are approaching the dominant eigenvector, in the last step we divide our final approximate vector by its norm to obtain a unit vector. Note that we get the same result as the 'eig' command. In practice, we achieve greater accuracy if we normalize the approximate eigenvector at each iteration rather than waiting until the end.

Finding Other Eigenvalues

If the Power Method converges to the dominant eigenvalue, how can we determine other eigenvalues of a matrix? This section is devoted to some variations of the Power Method that address this problem. Many of the ideas in this section can be employed in other settings.

The Inverse Power Method

The Inverse Power Method is based upon the following theorem.

Theorem 1. Let M be an invertible $n \times n$ matrix. Then the eigenvalues of M^{-1} are reciprocals of the eigenvalues of M .

Proof The proof is left as an exercise. See Practice Problem ??.

■

We return to the matrix in Exploration ??. Let $M = \begin{bmatrix} 9 & 2 & 8 \\ 2 & -6 & -2 \\ -8 & 2 & -5 \end{bmatrix}$ and let $A = M^{-1}$.

Now use Octave to verify that $A = \begin{bmatrix} 17/3 & 13/3 & 22/3 \\ 13/3 & 19/6 & 17/3 \\ -22/3 & -17/3 & -29/3 \end{bmatrix}$ and that the eigenvalues of A are indeed the reciprocals of the eigenvalues of M .

```

% Eigenvalues of the inverse are reciprocals of eigenvalues of the matrix
M=[9 2 8; 2 -6 -2; -8 2 -5];
eig(M)
A=inv(M)
eig(A)

```

[Link to code.](#)

If we now apply the Power Method to A , the Rayleigh quotients converge toward the dominant eigenvalue of A , which is -1 . If we take the reciprocal of the dominant eigenvalue of A , we will have found the “least dominant” eigenvalue of M , i.e., the eigenvalue closest to zero. Sure enough, the reciprocal of -1 is -1 , which is the eigenvalue of M closest to zero.

The Shifted Power Method

Again, we begin with a theorem.

Theorem 2. Let M be an invertible $n \times n$ matrix and let z be any number (real or complex). If λ is an eigenvalue of M , then $\lambda - z$ is an eigenvalue of $M - zI$.

Proof The proof is left as an exercise. See Practice Problem ??.

Once again, let $M = \begin{bmatrix} 9 & 2 & 8 \\ 2 & -6 & -2 \\ -8 & 2 & -5 \end{bmatrix}$ and let $A = M - (-3)I$.

Using Octave, you can verify that $A = \begin{bmatrix} 12 & 2 & 8 \\ 2 & -3 & -2 \\ -8 & 2 & -2 \end{bmatrix}$ and that the eigenvalues of A are obtained by subtracting

$c = -3$ from (i.e. by adding 3 to) each eigenvalue of M . The dominant eigenvalue of M gives us an eigenvalue of zero for A (you may find it gives a number very close to zero, due to rounding error). So if we apply the Power Method to A to find its dominant eigenvalue (which is 5), we can add -3 to get the “second largest” eigenvalue of M (i.e., the eigenvalue second-farthest from zero), which is 2.

To see this, we will use Octave.

```
% The shifted power method
M=[9 2 8;
 2 -6 -2;
 -8 2 -5];
eig(M)
c=-3;
A=M-c*eye(3)
eig(A)
xinit=[1;1;1]; %Input an initial vector
maxit =15; %Choose number iterations
x = xinit;
count = 0;
r=0;
while count < maxit
  y=x;
  x = A*y;
  if count > 0;
    r=(transpose(y)*x)/(transpose(y)*y);
    RQUOTIENTS(count)=r;
```

```

        end
        count = count + 1;
        VECTORS(:,count)=x;
    end
    VECTORS
    RQUOTIENTS
    RQUOTIENTS(maxit-1)+c

```

[Link to code.](#)

The Shifted Inverse Power Method

The Shifted Inverse Power method combines the two ideas we just studied. Using the Inverse Power method, we know how to find the eigenvalue closest to zero. We also know that we can change which eigenvalue is the dominant eigenvalue using the Shifted Power method. With the Shifted Inverse Power method, we are able to find the eigenvalue closest to a target τ , where τ can be any number, real or complex. We begin with the following theorem.

Theorem 3. Let M be an $n \times n$ matrix and let λ be an eigenvalue of M . If τ is any number (real or complex) other than an eigenvalue of M , then $\frac{1}{\lambda - \tau}$ is an eigenvalue of $(M - \tau I)^{-1}$.

Proof Suppose $M\mathbf{x} = \lambda\mathbf{x}$. By Theorem ?? we have

$$(M - \tau I)\mathbf{x} = (\lambda - \tau)\mathbf{x}.$$

Since τ is not an eigenvalue, $\det(M - \tau I) \neq 0$, so $M - \tau I$ is invertible. Also, since τ is not an eigenvalue, $\lambda \neq \tau$. We multiply both sides on the left by $(M - \tau I)^{-1}$ and divide both sides by the nonzero scalar $\lambda - \tau$. We arrive at

$$\frac{1}{\lambda - \tau}\mathbf{x} = (M - \tau I)^{-1}\mathbf{x}$$

as desired. ■

Again we let $M = \begin{bmatrix} 9 & 2 & 8 \\ 2 & -6 & -2 \\ -8 & 2 & -5 \end{bmatrix}$. We will use the Shifted Inverse Power Method to compute the eigenvalue of M closest to the target $\tau = 1$.

Using Octave, you can verify that $A = \begin{bmatrix} 5.75 & 3.5 & 6.5 \\ 3.5 & 2 & 4 \\ -6.5 & -4 & -7.5 \end{bmatrix}$. When we apply the Power Method to A to find its dominant eigenvalue, our Rayleigh quotients approach 1. This means that the eigenvalue of M closest to our target $\tau = 1$ is $1 + \tau$, or 2.

We will use Octave.

```

% The shifted inverse power method
M=[9 2 8; 2 -6 -2; -8 2 -5];

```

```

eig(M)
tau=1;
A=inv(M-tau*eye(3)) %I love the fact that eye(n) is the command for the identity matrix
eig(A)
xinit=[1;1;1]; %Input an initial vector
maxit =15; %Choose number iterations
x = xinit;
count = 0;
r=0;
    while count < maxit
        y=x;
        x = A*y;
        if count > 0;
            r=(transpose(y)*x)/(transpose(y)*y);
            RQUOTIENTS(count)=r;
        end
        count = count + 1;
        VECTORS(:,count)=x;
    end
VECTORS
RQUOTIENTS
RQUOTIENTS(maxit-1)+tau

```

[Link to code.](#)

Practice Problems

Problem 3 1. Prove Theorem ??.

2. Prove Theorem ??.

Problem 4 In each case, find the exact eigenvalues and determine corresponding eigenvectors. Then start with

$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and compute $\mathbf{x}_4$ and r_3 using the power method.

1. $A = \begin{bmatrix} 2 & -4 \\ -3 & 3 \end{bmatrix}$

2. $A = \begin{bmatrix} 5 & 2 \\ -3 & -2 \end{bmatrix}$

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

4. $A = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}$

Problem 5 Apply the power method to

$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, starting at $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Does it converge? Explain.

Text Source

This section was adapted from Section 8.5 of Keith Nicholson's *Linear Algebra with Applications*. (CC-BY-NC-SA)

W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, pp. 441–445.