Eigenvalues

Abstract
We consider when and how a matrix can be shown to be equivalent to a diagonal matrix. The fundamental notions of "eigenvalue" and "eigenvector" are developed. The power method and variations are used to find the eigenvalues and eigenvectors.

MATLAB Commands
eig, plot, inv, poly, roots, rank, norm

Linear Algebra Concepts
Eigenvalue, Eigenvector, Eigenspace, Diagonalizable Matrix, Characteristic Polynomial, Power Method

Background
We say \( \lambda \) is an eigenvalue and \( v \neq \vec{0} \) is an eigenvector if \( Av = \lambda v \). A matrix \( A \) is diagonalizable if there is an invertible matrix \( P \) and a diagonal matrix \( D \) such that \( P^{-1}AP = D \). If we let \( D = \text{diag}(\lambda_1, \ldots, \lambda_n) \) and \( P = [v_1, \ldots, v_n] \), then

\[
AP = [Av_1, \ldots, Av_n] = [\lambda_1 v_1, \ldots, \lambda_n v_n] = PD.
\]

Thus \( Av_i = \lambda_i v_i \) for \( i = 1, \ldots, n \).

From these observations the diagonal entries of \( D \) are seen to be eigenvalues and the columns of \( P \) are eigenvectors. The eigenvectors and the eigenvalues are important notions independently of diagonalization. The eigenvectors give "direction" to a matrix, in the sense that \( \lim_{n \to \infty} A^n v \) will frequently tend to a dominant eigenvector.

Suppose that we have \( Ax = \lambda x \), then \( Ax - \lambda x = \vec{0} = (A - \lambda I_n)x \) and so \( x \) is in the Null Space of \( A - \lambda I_n \). This null space is called the eigenspace of \( \lambda \).

Try the following

\[
B = \text{rand}(2); \quad A = 2*B' * B; \\
[P, D] = \text{eig}(A)
\]

The diagonal entries of \( D \) are the eigenvalues and the columns of \( P \) are the eigenvectors. Now try

\[
A = [2, 2; 0, 2]; \\
[P, D] = \text{eig}(A)
\]

This time you will notice that the columns of \( P \) are linearly dependent. Check \( \text{rank}(P) \)

This last example is important since \( A \) cannot be diagonalized.

Theorem 1. \( A \) is diagonalizable if and only if there is a basis for \( \mathbb{R}^n \) of eigenvectors.

Our example \( A = [2, 2; 0, 2] \) is not diagonalizable. If we try to solve \( Ax = \lambda x \) for \( \lambda \) we get only one eigenvalue \( \lambda = 2 \), but the Null Space of \( A - 2I \) has dimension 1, and so there can be no basis for \( \mathbb{R}^2 \) of eigenvectors. The problem of finding a basis of eigenvectors rests exactly on the repeated eigenvalues.

Theorem 2. If the eigenvalues \( \lambda_1, \ldots, \lambda_n \) for \( A \) are distinct, then \( A \) is diagonalizable.

Suppose that we try

\[
A = \text{rand}(5); \quad \text{eig}(A)
\]

The eigenvalues are distinct and so by Theorem 2 \( A \) is diagonalizable. Be careful in using this theorem. It does not say that if a matrix is diagonalizable, then the eigenvalues are distinct. Do you believe that the following matrix is diagonalizable?

\[
A = 3*\text{eye}(4)
\]
Please say “yes.” Please. When we have \((A - \lambda I)v = 0\) and \(v \neq 0\), \(A - \lambda I\) is singular, yielding \(\det(A - \lambda I) = 0\).
If we replace \(\lambda\) by \(x\), \(\det(A - xI)\) is a polynomial called the characteristic polynomial of \(A\). The eigenvalues of \(A\) are the roots of the characteristic polynomial. The MATLAB function \(\text{poly}(A)\) will find the characteristic polynomial of \(A\). Follow this with \(\text{roots}(\text{poly}(A))\), and you will have the eigenvalues of \(A\). Of course, \(\text{eig}(A)\) is much easier.

**Theorem 3.** If \(A\) is a triangular matrix, then the eigenvalues of \(A\) are the diagonal entries of \(A\).

The multiplicity of \(\lambda\) in the characteristic polynomial is \(m\) if \((x - \lambda)^m\) divides the characteristic polynomial but \((x - \lambda)^{m+1}\) does not. In other words, \(x - \lambda\) occurs as a factor exactly \(m\) times. There is an important connection between the multiplicity of \(\lambda\) and the eigenspace of \(\lambda\). First try this in MATLAB

\[
A=[[2*eye(2);zeros(2)],[ones(4,2)]]
\]
\[
[P,D]=\text{eig}(A)
\]

We see that the multiplicity of the eigenvalue \(\lambda = 2\) is 3. We also note that the eigenvectors appearing in \(P\) are not linearly independent, so \(A\) is not diagonalizable. We can check this easily with

\[
\text{rank}(\text{P})
\]

Finally we can compute the dimension of the eigenspace for \(\lambda = 2\)

\[
\text{4-rank}(\text{A-2*eye(4)})
\]

The next theorem includes a proof since the proof is simple and rarely appears in textbooks.

**Theorem 4.** If \(\lambda\) has multiplicity \(m\), then the dimension of the eigenspace of \(\lambda\) is \(\leq m\).
Proof: Let \(v_1, \ldots, v_k\) be a basis for the Null Space of \(\lambda I\). We want to show that \(m \geq k\). Expand \(v_1, \ldots, v_k\) to a basis for \(R^n\), \(v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\), and let \(B = [v_1, \ldots, v_n]\). Let \(C = B^{-1}AB\).

We will now get a closer look at the first \(k\) columns of \(C\).

\[
B^{-1}[Av_1, \ldots, Av_k] = [B^{-1}\lambda v_1, \ldots, B^{-1}\lambda v_k]
\]

\[
= [\lambda B^{-1}v_1, \ldots, \lambda B^{-1}v_k]
\]

\[
= [\lambda e_1, \ldots, \lambda e_k]
\]

Thus the first \(k\) columns look like

\[
\begin{pmatrix}
\lambda & 0 & \ldots & 0 \\
0 & \lambda & \ddots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \ldots & 0 & \lambda \\
0 & \ldots & \ldots & 0 \\
0 & \ldots & \ldots & 0
\end{pmatrix}
\]

But now we see that \((x - \lambda)^k\) divides the characteristic polynomial of \(C\) which is the same as the characteristic polynomial of \(A\). Thus \(m \geq k\).

This offers a way of definitively checking if a matrix is diagonalizable.

**Theorem 5.** \(A\) is diagonalizable if and only if for each eigenvalue \(\lambda\) the multiplicity of \(\lambda\) is the dimension of the eigenspace of \(\lambda\).

The matrix \(A\) constructed above is not diagonalizable since the multiplicity of the eigenvalue \(\lambda = 2\) is 3, while the dimension of the eigenspace is 2. When a matrix is diagonalizable, we would like to treat it just like a diagonal matrix. Unfortunately, some translation is necessary to get the desired results. The next theorem shows how to do this when computing powers.
Theorem 6. Suppose that $A$ is diagonalizable with $P^{-1}AP = D$, $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and $P = [v_1, \ldots, v_n]$. If $x$ is any vector and $c$ is the vector of coordinates of $x$ with respect to $v_1, \ldots, v_n$, that is, $x = Pc = \sum_{i=1}^{n} c_i v_i$, then for all $m$

$$A^m x = PD^m c = \sum_{i=1}^{n} c_i \lambda_i^m v_i$$

We now turn to the problem of finding eigenvalues and eigenvectors. We will look at the easiest method of finding an eigenvalue, the power method. Given a square matrix $A$ and a vector $v$ consider the sequence given by $v_0 = v$ and $v_{n+1} = Av_n$. This sequence of vectors will tend towards an eigenvector, but we need to worry about the terms in the sequence growing toward overflow or underflow. Thus, at each stage we will re-scale the vector, that is we assign

$$v_n = \frac{v_n}{\max(\text{abs}(v_n))}$$

Suppose that after re-scaling we arrive at a stage where $v_n \approx v_{n+1}$ where $\approx$ means $\|v_n - v_{n+1}\|_2 < \text{tol}$, and $\text{tol}$ is some small value we have chosen. Then we have found an eigenvector. This is easy to do in MATLAB.

```matlab
A = ones(5) + 5*eye(5); eig(A)
```

We can see the largest eigenvalue. Now use the power method to find it.

```matlab
v = rand(5,1);
for i = 1:20, y = v; v = A*y; v = v/max(abs(v)); norm(v-y), pause, end
```

The norm has been inserted so that we can watch the vectors converge. The `pause` lets you see the value. To continue the iteration, hit any key. When this is done we will look for the eigenvalue.

$$A^*v, /v$$

Theorem 7. If $A$ is diagonalizable with a unique largest (in absolute value) eigenvalue $\lambda$, then $v_n$ converges to an eigenvector $u$ for $\lambda$.

Proof: Let $x = c_1 \lambda u + \sum_{i=2}^{n} c_i \lambda_i u_i$ where $|\lambda| > |\lambda_i|$ for $i = 2, \ldots, n$. Then by Theorem 6 $A^m x = c_1 \lambda^m u + \sum_{i=2}^{n} c_i \lambda_i^m u_i$. Now divide both sides by $\lambda^m$ to get

$$(1/\lambda^m)A^m x = c_1 u + \sum_{i=2}^{n} c_i (\lambda_i/\lambda)^m u_i.$$ 

Since $|\lambda| > |\lambda_i|$, $|\lambda/\lambda_i| < 1$ and so the terms in the sum tend to 0, and thus for large $m$ we get

$$A^m x \approx c_1 \lambda^m u.$$ 

So for large $m$, $A^m x$ and $A^{m+1} x$ are approximately scalar multiples. After rescaling $A^m x \approx v_m$ and $A^{m+1} x \approx v_{m+1}$ so that $v_m$ and $v_{m+1}$ are positive scalar multiples each with a 1 in some coordinate, and hence $v_m \approx v_{m+1}$.

The hypothesis of the theorem that there be a unique largest eigenvalue is needed. Try

```matlab
A = ones(5)-5*eye(5); v = rand(5,1); eig(A)
for i = 1:20, y = v; v = A*y; v = v/max(abs(v)); norm(v-y), pause, end
```

In this case you note that $\text{norm}(v-y)$ does not go to 0.

The proof indicates that the starting vector $x$ is not so critical as long as the coefficient $c_1 \neq 0$. Typically we would not know this coefficient, but for an arbitrary $x$ it is unlikely that $c_1 = 0$.

With this approximate eigenvector, $v$, in hand, we need to find $\lambda$. A reasonable approach would be to look at $Av = \lambda v$. Since it may not be the case that $Av$ is an exact multiple of $v$, this approach is problematic.
In fact, viewing \( Av = \lambda v \) as an \( m \times 1 \) system with variable \( \lambda \), this system is likely to be inconsistent. We will use a least squares solution to get the eigenvalue.

**Theorem 8.** The least squares solution \( \lambda \) to \( Av = \lambda v \) is given by \( \lambda = \frac{v^T Av}{v^T v} \). The quotient \( \frac{v^T Av}{v^T v} \) is called the Rayleigh Quotient.

Proof: Look at the equation \( v \lambda = Av \). The normal equation is \( v^T v \lambda = v^T Av \), which has the solution \( \lambda = \frac{v^T Av}{v^T v} \).

We can use this in MATLAB.

```matlab
A=ones(5)+5*eye(5);
v=rand(5,1)
for i=1:20, y=v; v=A*y; v=v/max(abs(v)); end
```

We have an approximation to an eigenvector with \( v \). The Rayleigh Quotient will approximate the eigenvalue.

\[
\rho = \frac{(v^T A v)(v^T v)}{(v^T v)}
\]

Compare with `eig(A)`

Suppose that \( A \) is invertible and that \( \lambda \neq 0 \) is an eigenvalue of \( A \) with eigenvector \( v \). We have \( Av = \lambda v \), and multiplying by inverses \( A^{-1}v = (1/\lambda)v \). If \( 1/\lambda \) is the largest (in absolute value) eigenvalue of \( A^{-1} \), then \( \lambda \) is the smallest (in absolute value) eigenvalue of \( A \). This method, the inverse method, will find the smallest eigenvalue of \( A \) by applying the power method the inverse of \( A \). Notice that \( v \) is an eigenvector for both \( A \) and \( A^{-1} \). To get an eigenvalue for \( A \) from \( v \) compute the Rayleigh Quotient on \( A \). Try

```matlab
A=ones(5)+diag(1:5); B=inv(A);
for i=1:20, y=v; v=B*y; v=v/max(abs(v)); end
norm(y-v)
r=(v^T A v)/(v^T v)
min(eig(A))
```

Another variation is the shift method. Here you choose a scalar \( s \) and apply the power method to the shifted matrix \( A - sI \). Notice that if \( v \) is an eigenvector of \( A \) with eigenvalue \( \lambda \), then \( (A - sI)v = Av - sv = (\lambda - s)v \), so that \( \lambda - s \) is an eigenvalue of the shifted matrix and \( v \) is an eigenvector of both \( A \) and \( A - sI \).

By combining the inverse method with the shift method we get an effective iteration, the inverse shift method. If you choose \( s \) close enough to \( \lambda \), then \( \lambda - s \) is the smallest eigenvalue of the shifted matrix \( A - sI \), thus applying the power method to the inverse of \( A - sI \) will converge to an eigenvector. The problem is choosing the shift \( s \). Let’s try this in MATLAB, but we will cheat a bit.

```matlab
A=ones(3)+diag(1:3), eig(A)
```

Now choose \( s=2 \) and invert \( B=A-2*eye(3) \)

```matlab
B=inv(A-2*eye(3)), v=rand(3,1);
for i=1:20, v=B*v; v=v/max(abs(v)); end
```

\( v \) is an eigenvector for the original \( A \) so the Rayleigh Quotient gives an eigenvalue for \( A \).

\[
r = \frac{(v^T A v)(v^T v)}{(v^T v)}
\]

\( \text{eig}(A) \)

One way to choose a shift is the Rayleigh Quotient \( s = \frac{x^T Ax}{x^T x} \) of some vector \( x \). This presupposes that you have some vector \( x \), in mind for the eigenvector. A less obvious choice is based on Gershgorin’s Theorem, which suggests using the diagonal entries \( a_{ii} \) as shift values.
Problems

1. Determine which of the following matrices are diagonalizable.
   (1) pascal(5)
   (2) list(5)
   (3) jord(5,5)
   (4) ones(5)
   (5) magic(6)

2. When $A$ is invertible $Ax = b$ can be solved with $x = A^{-1}b$. If $A$ is an invertible diagonal matrix, we can go further to say $x(i) = b(i)/A(i,i)$. Suppose that $A$ is diagonalizable, so that $P^{-1}AP = D$ where $D$ is diagonal and $Ax = PD^{-1}P^{-1}x = b$. Let $z = P^{-1}x$ be new variables. Then to solve $PDz = b$, multiply both sides by $P^{-1}$, so $Dz = P^{-1}b$, a diagonal system, which is easily solved. Now $x$ is obtained by $x = Pz$. Let $A = hilb(5)$ and find an invertible matrix $P$ and a diagonal matrix $D$ such that
   \[ \text{inv}(P)*A*P = D. \]
   Solve the matrix equation $Ax = b$ where $b = (1:5)'$ using the technique described above.

3. Let $A = \text{rand}(5,3)$; $P = A*(A'*A) \setminus A'$ so that $P$ is the projection matrix into the Column Space of $A$. What are the eigenvalues of $P$? One of the eigenspaces of $P$ is $\text{Col}(A)$. What is the other eigenspace?

4. Let $A = \text{rand}(5,3)$; $P = (A*(A'*A) \setminus A')$; $R = -\text{eye}(5)*2*P$; so that $R$ is the reflection matrix across the Column Space of $A$. What are the eigenvalues of $R$? Describe the eigenspaces of $R$. Hint: See problem 3.

5. For each of the matrices in problem 1 compute $\text{prod(eig(A))}$, $\text{det(A)}$, and $\text{poly(A)}$. What can you say about these answers? Can you explain why this is so?

6. Write a MATLAB function $b = \text{diagonal}(A)$ which returns $b = 1$ if $A$ is diagonalizable and $b = 0$ otherwise.

7. Test your $\text{diagonal}$ on the matrices in problem 1.

8. Let $A = \text{rand}(5)$ and $B = \text{rref}(A)$. Is there any relationship between the eigenvalues of $A$ and the eigenvalues of $B$?

9. Use the power method to find an eigenvalue for each of the following; you can check for accuracy by comparing with MATLAB's $\text{eig}$.
   (1) list(5)
   (2) pascal(5)
   (3) hilb(5)
   (4) jord(5,3)
   (5) house(rand(5,1))
   What happens with $\text{jord}(5,3)$ and $\text{house}(\text{rand}(5,1))$?

10. Write a MATLAB function $[v, r] = \text{powsmth}(A, x, n)$ which locates an eigenvector $v$ and an eigenvalue $r$ for $A$ using the power method. $n$ is the maximum number of iterations, and $x$ is the seed vector.

11. Test your $\text{powsmth}$ on each of the matrices in problem 9.

12. Use the inverse shift method to find all of the eigenvalues of $h = \text{hilb}(4)$ by carefully choosing the shift $s$. First try using the diagonal entries of $h$ for the shift $s$. If you do not succeed in getting all of the eigenvalues this way, you may use $\text{eig(h)}$ as a guide to choosing $s$.

13. Write a MATLAB function $[v, r] = \text{invsmt}(A, x, s, n)$ which finds an eigenvector $v$ and an eigenvalue $r$ for $A$ using the inverse shift method. The shift is given by $s$, while the other parameters have the same meaning as in problem 10.