Lecture 14: Eigenvalues / Eigenvectors

June 23, 2025

Outline

1. Eigenvalue Problem Definitions
2. Traditional Eigenvalue Problem Review
3. Solving Eigenvalue Problems (Naïve Approach)
4. Eigenvalue/Eigenvector Review Example
5. Rayleigh quotient
6. Power Iteration

Introduction

- So far we have discussed solving linear systems and least-squares problems of the form $Ax = b$.
- We now consider eigenvalue problems, which have the form

$$Ax = \lambda x.$$

- In this lecture, we begin with some definitions and theory about eigenvalue problems.
- Much of the beginning of this lecture should be review from previous courses.

Eigenvalue Problem Definitions

Definition 1.1

Let $A \in \mathbb{R}^{n \times n}$. A non-zero vector $x \in \mathbb{R}^n$ is a (right) **eigenvector** *with corresponding* **eigenvalue** $\lambda \in \mathbb{R}$ *if*

$$Ax = \lambda x.$$

- Note that if x is an eigenvector then so is ax, for $a \neq 0$.
- That is, eigenvectors are unique only up to a multiplicative constant.

Definition 1.2

The set of A's eigenvalues is called its **spectrum**, *denoted $\Lambda(A)$.*

Eigenvalue Problem Definitions

Definition 1.3

The **eigendecomposition** *of a diagonalizable matrix, A, is*

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

where

$$X = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix}, \qquad \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

and $Ax_i = \lambda_i x_i$ for $i = 1, 2, \ldots, n$.

Eigenvalue Problem Definitions

- The columns of X in the eigendecomposition are eigenvectors of A.
- The diagonal matrix Λ has entries that are the eigenvalues corresponding to each eigenvector.
- **Fact:** Real symmetric matrices are diagonalizable by orthogonal matrices.

Eigenvalue Problem Definitions

Equivalently, the eigendecomposition can be written as $AX = X\Lambda$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

- This form corresponds to the form of the eigenvalue problem $Ax = \lambda x$, but stacks all eigenvalue/eigenvector pairs (x_i, λ_i) in one matrix equation (for $i = 1, \ldots, n$).
- See the Lecture Notes for a proof that $AX = X\Lambda$.

Traditional Eigenvalue Problem Review

In introductory linear algebra courses you would normally compute (by hand) eigenvalues and eigenvectors using the characteristic polynomial.

Definition 2.1

The **characteristic polynomial** *of* A, *denoted* $p_A(z)$, *is the degree* n *(monic) polynomial given by*

$$p_A(z) = \det(zI - A).$$

Theorem 1

λ *is an eigenvalue of* A *iff* $p_A(\lambda) = 0$.

Proof.

See Lecture Notes. □

Traditional Eigenvalue Problem Review

- The **fundamental theorem of algebra** tells us that the degree n polynomial $p_A(z)$ has n (possibly complex) roots.
- So A has n (possibly complex) eigenvalues, given by the roots.
- Therefore, we can write

$$p_A(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n).$$

- Given an eigenvalue λ, the corresponding eigenvector(s) are given by solving $(\lambda I - A)x = 0$ for x (i.e., the nullspace of $\lambda I - A$.)
- We will see later why the choice of A will yield real eigenvalues.

Traditional Eigenvalue Problem Review

- Conversely, for every monic polynomial of degree n,

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

there always exists a matrix whose eigenvalues are roots of $p(z)$.

- This matrix is called the **companion matrix**

$$C = \begin{bmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & 0 & & -a_2 \\ & & \ddots & \ddots & \vdots \\ & & & 1 & -a_{n-1} \end{bmatrix}.$$

Eigenvalue Problem Definitions

The following definitions are concerned with the **multiplicity** of eigenvalues.

Definition 2.2

The **algebraic multiplicity** *of λ is the number of times it appears as a root of $p_A(z)$.*

Definition 2.3

The **geometric multiplicity** *of λ is the dimension of the nullspace of the matrix $\lambda I - A$.*

Eigenvalue Problem Definitions

Remarks:

1. Why geometric multiplicity cannot exceed algebraic multiplicity: The geometric multiplicity is the number of linearly independent vectors, and each vector is the solution to one algebraic eigenvector equation, so there must be at least as much algebraic multiplicity.
2. If algebraic multiplicity exceeds the geometric multiplicity then λ is a **defective eigenvalue** (See Lecture Notes examples).
3. The matrix A is then called a **defective matrix**.
4. This is important because only **non**-defective matrices have eigenvalue decompositions.

Eigenvalue Problem Definitions

Multiplicity Example The following example with

$$A = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

is worked out in the Lecture Notes.

Solving Eigenvalue Problems (Naïve Approach)

- Sadly, no *closed form* solutions exist for degree 5 or higher polynomials as the roots cannot be found exactly using a finite number of rational operations.
- Therefore we must use approximations to find their eigenvalues.
- This suggests the application of iterative methods.
    1. Form the characteristic polynomial, $p_A(z)$.
    2. Use the numerical root-finding method to extract the approximate roots/eigenvalues. (e.g bisection, Newton, etc.)
- The problem with this approach, is that root-finding tends to be ill-conditioned, i.e. a small change or error in the input drastically changes the roots.
- More effective strategies based on finding eigendecompositions exist.
- We will explore some of these methods in the following lectures.

Solving Eigenvalue Problems (Naïve Approach)

Some results about bounding eigenvalues based on the Gershgorin circle theorem will be useful later.

Theorem 2

(Gershgorin circle theorem) Let A be any square matrix. The eigenvalues λ of A are located in the union of the n disks (on the complex plane) given by

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|.$$

Disks are denoted by $D(a_{ii}, R_i)$, where $R_i = \sum_{j \neq i} |a_{ij}|$.

Proof.

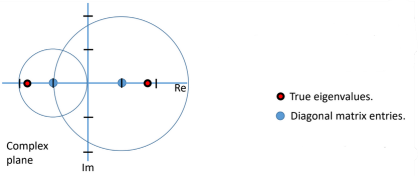
See Lecture Notes. □

Solving Eigenvalue Problems (Naïve Approach)

The Gershgorin circle theorem essentially says the following. If off-diagonal entries in a row are small, then the corresponding eigenvalue must be close to the diagonal entry. For example, with

$$A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}, \text{ which has } \Lambda(A) = \left\{ \sqrt{3}, -\sqrt{3} \right\}.$$

The figure below shows the Gershgorin disks $D(a_{ii}, \sum_{j \neq i} |a_{ij}|)$ for this example.



For this example we have Disk 1 centered at (1,0) with radius 2 or D(1,2). There is also disk 2 centered at (-1,0) with radius 1 or D(-1,1). Note that complex eigenvalues will give circles centred off the real axis.

Solving Eigenvalue Problems (Naïve Approach)

As an exercise, find the Gershgorin disks for

$$A = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 4 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Can we determine any of the eigenvalues exactly for A? What if A was a diagonal matrix, could the eigenvalues be determined exactly?

Eigenvalue/Eigenvector Review Example

We present an example that reviews computing the eigenvalues and eigenvectors using the characteristic polynomial. In this example we find the eigenvalues and eigenvectors for the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 2 \\ 4 & -4 & 5 \end{bmatrix}.$$

Solution: See Lecture Notes.

Eigenvalue/Eigenvector Review Example

- We will generally consider matrices $A \in \mathbb{R}^{n \times n}$ that are symmetric ($A^T = A$).
- Such matrices have useful properties such as real eigenvalues, and a complete set of orthogonal eigenvectors.

$$\{\lambda_1, \lambda_2, \ldots, \lambda_n\}, \{q_1, q_2, \ldots, q_n\}, \text{with } \|q_i\| = 1.$$

Therefore,

$$A = Q\Lambda Q^T,$$

where Q is orthogonal.

Rayleigh quotient

- There are two quantities that must be solved for in eigenvalue problems: the **eigenvalues** and the **eigenvectors**.
- Consider first computing eigenvalues, when given an approximation to an eigenvector.
- An important quantity in this case is the **Rayleigh quotient** defined next.

Definition 5.1

The **Rayleigh quotient** *of a nonzero vector x with respect to A is*

$$r(x) = \frac{x^T A x}{x^T x}.$$

- Note that if x is an eigenvector of A then $r(x) = \lambda$ since

$$\frac{x^T A x}{x^T x} = \frac{x^T(\lambda x)}{x^T x} = \lambda \frac{x^T x}{x^T x} = \lambda.$$

- Otherwise, $r(x)$ gives a scalar α that behaves "most like" an eigenvalue for a given vector x.

Rayleigh quotient

- We can justify the definition of the Rayliegh quotient in another way.
- Consider the following single-variable, $n \times 1$ least squares problem for an unknown $\alpha \in \mathbb{R}$:

$$\min_{\alpha} \left\| \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \alpha - Ax \right\|_2^2,$$

for given A and x.

- Constructing the normal equations for this problem, we have

$$\begin{aligned} (x^T x)\alpha &= x^T(Ax), \\ \Rightarrow \quad \alpha &= \frac{x^T Ax}{x^T x} \\ &= r(x). \end{aligned}$$

Rayleigh quotient

- Consider this Rayleigh quotient example for the following matrix

$$A = \begin{bmatrix} 3 & 2 & 5 \\ 2 & 7 & 5 \\ 0 & 2 & 8 \end{bmatrix},$$

which has an eigenvector near $v \approx [0.7, -0.7, 0.3]^T$.

Rayleigh quotient

The Rayleigh quotient gives

$$
\begin{aligned}
\alpha &= \frac{v^T A v}{v^T v} \\
&= \left(\frac{1}{\begin{bmatrix} \frac{7}{10} & -\frac{7}{10} & \frac{3}{10} \end{bmatrix} \begin{bmatrix} \frac{7}{10} \\ -\frac{7}{10} \\ \frac{3}{10} \end{bmatrix}} \right) \begin{bmatrix} \frac{7}{10} & -\frac{7}{10} & \frac{3}{10} \end{bmatrix} \begin{bmatrix} 3 & 2 & 5 \\ 2 & 7 & 5 \\ 0 & 2 & 8 \end{bmatrix} \begin{bmatrix} \frac{7}{10} \\ -\frac{7}{10} \\ \frac{3}{10} \end{bmatrix} \\
&= \frac{324}{107} \\
&\approx 3.028,
\end{aligned}
$$

which is close to the true value of $\lambda = 3$ (See Lecture Notes for compuation of eigenvalues and eigenvectors).

Rayleigh quotient

- From this example we see that if we had a reasonable guess at an eigenvector, the Rayleigh quotient would be a useful approximation of the eigenvalue.
- In fact, the following theorem (see Trefethen & Bau, Lecture 27) states that the approximation converges quadratically.

Theorem 3

Let q_j be an eigenvector, and $x \approx q_j$. Then

$$r(x) - r(q_j) = O(\|x - q_j\|^2) \text{ as } x \to q_j.$$

- That is, as $x \to q_j$, the error in the estimate of the eigenvalue λ decreases quadratically.

Rayleigh quotient

- To summarize, eigenvalues are approximated using the Rayliegh quotient, given an approximation of an eigenvector.
- We will now look at how to get an approximation of the eigenvectors.
- We will again use **iterative** approaches.

Power Iteration

- The idea of the power iteration is simple.
- Start with an initial vector $v^{(0)}$, then repeatedly multiply by A and normalize:

$$\begin{aligned} v^{(1)} &= \frac{Av^{(0)}}{\left\| Av^{(0)} \right\|}, \\ v^{(2)} &= \frac{Av^{(1)}}{\left\| Av^{(1)} \right\|}, \\ &\vdots \\ v^{(k)} &= \frac{Av^{(k-1)}}{\left\| Av^{(k-1)} \right\|}. \end{aligned}$$

- In the limit, the approximation $v^{(k)}$ approaches q_1, where q_1 is the eigenvector associated with the *largest* magnitude eigenvalue, i.e.,

$$\lim_{k \to \infty} v^{(k)} = q_1.$$

Power Iteration

Remarks:

1. Recall that A must be symmetric, so that there exists an orthogonal basis of eigenvectors of A.
2. If we are lucky, and start with $v^{(1)}$ an eigenvector for eigenvalue λ, depending on $sign(\lambda)$, we will get back the same vector in 1 or possibly 2 steps.

Let's prove that Power Iteration converges.

Proof:

- Let $v^{(0)}$ be an initial guess at the eigenvector q_1.
- Also let $\{q_i\}$ denote the set of orthonormal eigenvectors.
- Then we can write

$$v^{(0)} = c_1 q_1 + c_2 q_2 + \cdots + c_n q_n,$$

for some coefficients c_i, because the eigenvectors span the space.

Power Iteration

- Now since $Aq_i = \lambda_i q_i$ we can write

$$Av^{(0)} = c_1\lambda_1 q_1 + c_2\lambda_2 q_2 + \cdots + c_n\lambda_n q_n.$$

- Further multiplication by A gives

$$\begin{aligned} A^k v^{(0)} &= c_1\lambda_1^k q_1 + c_2\lambda_2^k q_2 + \cdots + c_n\lambda_n^k q_n, \\ &= \lambda_1^k \left(c_1 q_1 + c_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k q_2 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^k q_n \right). \end{aligned}$$

- Now if we have $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$ and $c_1 = q_1^T v^{(0)} \neq 0$, then we observe the following:

$$\left(\frac{\lambda_i}{\lambda_1}\right)^k \to 0 \text{ as } k \to \infty, \text{ for } i > 1.$$

Power Iteration

- Therefore $A^k v^{(0)} \approx c_1 \lambda_1^k q_1$, for large k.
- Since the eigenvectors are orthonormal, the scale factor doesn't matter.
- We can normalize to find q_1 as

$$q_1 \to \frac{A^k v^{(0)}}{\|A^k v^{(0)}\|} \quad \text{as } k \to \infty.$$

- This is the eigenvector for the the largest magnitude eigenvalue. $\square$

The next theorem summarizes this result.

Power Iteration

Theorem 4 (Power iteration convergence)

Suppose $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$, *and* $q_1^T v^{(0)} \neq 0$, *where* q_1 *is the eigenvector for* λ_1. *Then*

$$\left\| v^{(k)} - (\pm q_1) \right\| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right), \text{ and } |\lambda^{(k)} - \lambda_1| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^{2k} \right),$$

as $k \to \infty$.

- This is *linear* convergence for the eigenvector with convergence factor is $\left|\frac{\lambda_2}{\lambda_1}\right|$.
- Convergence is therefore slow if $|\lambda_2| \approx |\lambda_1|$, i.e., if the first two eigenvalues are close in magnitude.
- There will be no convergence at all if $|\lambda_2| = |\lambda_1|$.

Power Iteration

Algorithm 1 gives pseudocode for the power iteration.

Algorithm 1 Power Iteration Algorithm

$v^{(0)} =$ initial guess, s.t. $\left\| v^{(0)} \right\| = 1$
for $k = 1, 2, \ldots$
  $w = Av^{(k-1)}$ ▷ Apply A
  $v^{(k)} = \frac{w}{\|w\|}$ ▷ Normalize
  $\lambda^{(k)} = \left(v^{(k)}\right)^T Av^{(k)}$ ▷ Rayleigh Quotient
end for

Remark: We have not said at all yet how to know which k we might need to make our approximation close enough.