Chapter 4 Notes - LA

John Yang

March 14, 2022

Contents

4 Eigenvalues and Eigenvectors

4.1 Introduction to Eigenvalues and Eigenvectors

- Let A be an $n \times n$ matrix. A scalar λ is called an eigenvalue of A if there is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. Such a vector $\mathbf{x}$ is called an eigenvector of A corresponding to λ.
- Let A be an $n \times n$ matrix and let λ be an eigenvalue of A. The collection of all eigenvectors corresponding to λ, together with the zero vector, is called the eigenspace of λ and is denoted by E_λ.

4.2 Determinants

- Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Then the determinant of A is the scalar

$$\det A = |A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

- We can simplify this equation as:

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$$

$$= \sum_{j=1}^{3} (-1)^{1+j} a_{ij} \det A_{ij}$$

- For any square matrix A, $\det A_{ij}$ is called the (i, j)-minor of A.
- Let $A = [a_{ij}]$ be an $n \times n$ matrix, where $n \geq 2$. Then the determinant of A is the scalar

$$\det A = |A| = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}$$

$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$$

- The (i, j)-cofactor of A is defined as

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

- Thus, the definition of the determinant becomes

$$\det A = \sum_{j=1}^{n} a_{1j} C_{1j}$$

- The Laplace Expansion Theorem: The determinant of an $n \times n$ matrix $A = [a_{ij}]$, where $n \geq 2$, can be computed as

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$
$$\sum_{j=1}^{n} a_{ij} C_{ij}$$

(which is the cofactor expansion along the ith row) and also as

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$
$$= \sum_{i=1}^{n} a_{ij} C_{ij}$$

(the cofactor expansion along the jth column).

- The determiniant of a triangular matrix is the product of the entries on its main diagonal. Specifically, if $A = [a_{ij}]$ is an $n \times n$ triangular matrix, then

$$\det A = a_{11}a_{22} \cdots a_{nn}$$

- Properties of determinants: let A be a square matrix.
  - If A has a zero row (column), then $\det A = 0$
  - If B is obtained by interchanging two rows (columns) of A, then $\det B = -\det A$
  - If A has two identical rows (columns), then $\det A = 0$
  - If B is obtained by multiplying a row (column) of A by k, then $\det B = k \det A$
  - If A, B, and C are identical except that the ith row (column) of C is the sum of the ith rows (columns) of A and B, then $\det C = \det A + \det B$
  - If B is obtained by adding a multiple of one row (column) of A to another row (column), then $\det B = \det A$
- Let E be an $n \times n$ elementary matrix.
  - If E results from interchanging two rows of I_n, then $\det E = -1$
  - If E results from multiplying one row of I_n by k, then $\det E = k$
  - If E results from adding a multiple of one row of I_n to another row, then $\det E = 1$
- Let B be an $n \times n$ matrix and let E be an $n \times n$ elementary matrix. Then

$$\det(EB) = (\det E)(\det B)$$

- A square matrix A is invertible IFF $\det A \neq 0$
- If A is an $n \times n$ matrix, then

$$\det(kA) = k^n \det A$$

- If A and B are $n \times n$ matrices, then

$$\det(AB) = (\det A)(\det B)$$

- If A is invertible, then

$$\det\left(A^{-1}\right) = \frac{1}{\det A}$$

- For any square matrix A,

$$\det A = \det A^T$$

- Cramer's rule: let A be an invertible $n \times n$ matrix and let $\mathbf{b}$ be a vector in $\mathbb{R}^n$. Then the unique solution $\mathbf{x}$ of the system $A\mathbf{x} = \mathbf{b}$ is given by

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det A} \qquad \text{for } i = 1, \cdots, n$$

- Let A be an invertible $n \times n$ matrix. Then

$$A^{-1} = \frac{1}{\det A}\text{adj } A$$

  where the adjoint of A adj A is defined by

$$[C_{ji}] = [C_{ij}]^T = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

- Let A be an $n \times n$ matrix. Then

$$a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} = \det A = a_{11}C_{11} + A_{21}C_{21} + \cdots + a_{n1}C_{n1}$$

- Let A be an $n \times n$ matrix and let B be obtained by interchanging any two rows (columns) of A. Then

$$\det B = -\det A$$

4.3 Eigenvalues and Eigenvectors of $n \times n$ Matrices

- The eigenvalues of a square matrix A are preciesely the solutions λ of the equation

$$\det(A - \lambda I) = 0$$

  .

- Finding the eigenvalues and eigenvectors of a matrix: Let A be an $n \times n$ matrix.
  - Compute the characteristic polynomial $\det(A - \lambda I)$ of A.
  - Find the eigenvalues of A by solving the characteristic equation $\det(A - \lambda I) = 0$ for λ
  - For each eigenvalue λ, find the null space of the matrix $A - \lambda I$. This is the eigenspace E_λ, the nonzero vectors of which are the eigenvectors of A corresponding to λ.
  - Find a basis for each eigenspace.
- The algebraic multiplicity of an eigenvalue is its multiplicity as a root of the characteristic equation. An $n \times n$ matrix will always have n eigenvalues, but some will be duplicates due to algebraic multiplicity.
- The eigenvalues of a triangular matrix are the entries on its main diagonal.

- A square matrix A is invertible IFF 0 is not an eigenvalue of A.
- The fundamental theorem of invertible matrices: version 3. Let A be an $n \times n$ matrix. The following statements are equivalent:
  - A is invertible
  - $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$
  - $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
  - The reduced row echelon form of A is I_n.
  - A is the product of elementary matrices.
  - $\text{rank}(A) = n$
  - $\text{nullity}(A) = 0$
  - The column vectors of A are linearly independent
  - The column vectors of A span $\mathbb{R}^n$
  - The column vectors of A form a basis for $\mathbb{R}^n$
  - The row vectors of A are linearly independent
  - The row vectors of A span $\mathbb{R}^n$
  - The row vectors of A form a basis for $\mathbb{R}^n$
  - $\det A \neq 0$
  - 0 is not an eigenvalue of A
- Let A be a square matrix with eigenvalue λ and corresponding eigenvector $\mathbf{x}$.
  - For any positive integer n, λ^n is an eigenvalue of A^n with corresponding eigenvector $\mathbf{x}$.
  - If A is invertible, then $1/\lambda$ is an eigenvalue of A^{-1} with corresponding eigenvector $\mathbf{x}$.
  - If A is invertible, then for any integer n, λ^n is an eigenvalue of A^n with corresponding eigenvector $\mathbf{x}$.
- Suppose the $n \times n$ matrix A has eigenvectors $\mathbf{v}_1, \mathbf{v}_2 \cdots, \mathbf{v}_m$ with corresponding eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_m$. If $\mathbf{x}$ is a vector in $\mathbb{R}^n$ that can be expressed as a linear combination of these eigenvectors, then for any integer k,
$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_m\lambda_m^k\mathbf{v}_m$$
- Let A be an $n \times n$ matrix and let $\lambda_1, \lambda_2, \cdots, \lambda_m$ be distinct eigenvalues of A with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$. Then $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ are linearly independent.

4.4 Similarity and Diagonalization

- Let A and B be $n \times n$ matrices. We say that A is similar to B if there is an invertible $n \times n$ matrix P such that $P^{-1}AP = B$. If A is similar to B, we write $A \sim B$
- Let A, B, C be $n \times n$ matrices.
  - $A \sim A$
  - If $A \sim B$ then $B \sim A$
  - If $A \sim B$ and $B \sim C$ then $A \sim C$
- Let A and B be $n \times n$ matrices with $A \sim B$. Then
  - $\det A = \det B$
  - A is invertible IFF B is invertible.

- A and B have the same rank
- A and B have the same characteristic polynomial.
- A and B have the same eigenvalues.
- $A^m \sim B^m$ for all integers $m \geq 0$
- If A is invertible, then $A^m \sim B^m$ for all integers m.

- An $n \times n$ matrix A is diagonalizable if there is a diagonal matrix D such that A is similar to D - that is, if there is an invertible $n \times n$ matrix P such that $P^{-1}AP = D$
- Let A be an $n \times n$ matrix. Then A is diagonalizable IFF A has n linearly independent eigenvectors. More precisely, there exist an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ IFF the columns of P are n linearly independent eigenvectors of A and the diagonal entries of D are the eigenvalues of A corresponding to the eigenvectors in P in the same order.
- Let A be an $n \times n$ matrix and let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigenvalues of A. If $\mathcal{B}_i$ is a basis for the eigenspace E_λ, then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ (i.e., the total collection of basis vectors for all the eigenspaces) is linearly independent.
- If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable
- If A is an $n \times n$ matrix, then the geometric multiplicity of each eigenvalue is less than or equal to its algebraic multiplicity.
- The diagonalization theorem: Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \lambda_2, \cdots, \lambda_k$. The following statements are equivalent.
  - A is diagonalizable.
  - The union $\mathcal{B}$ of the bases of the eigenspaces of A (as in theorem 4.24) contains n vectors.
  - The algebraic multiplicity of each eigenvalue equals its geometric multiplicity.

4.5 Iterative Methods for Computing Eigenvalues

- Let A be an $n \times n$ diagonalizable matrix with dominant eigenvalue λ_1. Then there exists a nonzero vector $\mathbf{x}_0$ such that the sequence of vectors $\mathbf{x}_k$ defined by

  $$\mathbf{x}_1 = A\mathbf{x}_0, \mathbf{x}_2 = A\mathbf{x}_1, \mathbf{x}_3 = A\mathbf{x}_2, \cdots, \mathbf{x}_k = A\mathbf{x}_{k-1}, \cdots$$

  approaches a dominant eigenvector of A
- Summarization of the power method: Let A be a diagonalizable $n \times n$ matrix with a corresponding dominant eigenvalue λ_1
  - Let $\mathbf{x}_0 = \mathbf{y}_0$ be any initial vector in $\mathbb{R}^n$ whose largest component is 1.
  - Repeat the following steps for $k = 1, 2, \cdots$:
    * Compute $\mathbf{x}_k = A\mathbf{y}_{k-1}$
    * Let m_k be the component of $\mathbf{x}_k$ with the largest absolute value.
    * Set $\mathbf{y}_k = (1/m_k)\mathbf{x}_k$
- For most choices of $\mathbf{x}_0$, m_k converges to the dominant eigenvalue λ_1 and $\mathbf{y}_k$ converges to a dominant eigenvector.
- Let $A = [a_{ij}]$ be a (real or complex) $n \times n$ matrix, and let r_i denote the sum of the absolute values of the off-diagonal entries in the ith row of A; that is, $r_i = \sum_{j \neq i} |a_{ij}|$. The ith Gerschgorin disk is the circular disk D_i in the complex plane with center a_{ii} and radius r_i. That is,

  $$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}$$
- Gerschgorin's Disk Theorem: Let A be an $n \times n$ (real or complex) matrix. Then every eigenvalue of A is contained within a Gerschgorin disk.

4.6 Applications and the Perron-Frobenius Theorem

- If P is the $n \times n$ transition matrix of a Markov chain, then 1 is an eigenvalue of P.
- Let P be an $n \times n$ transition matrix with eigenvalue λ.
  - $|\lambda| \leq 1$
  - If P is regular and $\lambda \neq 1$, then $|\lambda| < 1$
- Let P be a regular $n \times n$ transition matrix. If P is diagonalizable, then the dominant eigenvalue $\lambda_1 = 1$ has algebraic multiplicity 1
- Let P be a regular $n \times n$ transition matrix. Then as $k \to \infty$, p^k approcahes an $n \times n$ matrix L whose columns are identical, each equal to the same vector $\mathbf{x}$. This vector $\mathbf{x}$ is a steady state probability vector for P.
- Let P be a regular $n \times n$ transition matrix, with $\mathbf{x}$ the steady state probability vector for P, as in the above. Then, for any initial probability vector $\mathbf{x}_0$, the sequence of iterates $\mathbf{x}_k$ approaches $\mathbf{x}$.
- Every Leslie matrix has a unique positive eigenvalue and a corresponding eigenvector with positive components.
- Perron's Theroem: Let A be a positive $n \times n$ matrix. Then A has a real eigenvalue λ_1 with the following properties:
  - $\lambda_1 > 0$
  - λ_1 has a corresponding positive eigenvector.
  - If λ is any other eigenvalue of A, then $|\lambda| \leq \lambda_1$
- The Perron-Frobenius Theorem: Let A be an irreducible nonnegative $n \times n$ matrix. Then A has a real eigenvalue λ_1 with the following properties:
  - $\lambda_1 > 0$
  - λ_1 has a corresponding positive eigenvector.
  - If λ is any other eigenvalue of A, then $|\lambda| \leq \lambda_1$. If A is primitive, then this inequality is strict.
  - If λ is an eigenvalue of A such that $|\lambda| = \lambda_1$, then λ is a (complex) root of the equation $\lambda^n - \lambda_1^n = 0$
  - λ_1 has algebraic multiplicity 1.
- Def: Let $(x_n) = (x_0, x_1, x_2)$ be a sequence of numbers that is defined as follows:
  - $x_0 = a_0, x_1 = a_1, \cdots, x_{k-1} = a_{k-1}$, where $a_0, a_1, \cdots, a_{k-1}$ are scalars.
  - For all $n \geq k, x_n = c_1 x_{n-1} + c_2 x_{n-2} + \cdots + c_k x_{n-k}$, where $c_1, c_2, \cdots, c_k$ are scalars.
- If $c_k \neq 0$, the equation in the second line is called a linear recurrence relation of order k. The equations in the first line are referred to as the initial conditions of the recurrence.
- Let $x_n = ax_{n-1} + bx_{n-2}$ be a recurrence relation that is satisfied by a sequence (x_n). Let λ_1 and λ_2 be the eigenvalues of the associated characteristic equation $\lambda^2 - a\lambda - b = 0$.
  - If $\lambda_1 \neq \lambda_2$, then $x_n = c_1\lambda_1^n + c_2\lambda_2^n$ for some scalars c_1 and c_2.
  - If $\lambda_1 = \lambda_2 = \lambda$, then $x_n = c_1\lambda^n + c_2 n\lambda^n$ for some scalars c_1 and c_2.
- In either case, c_1 and c_2 can be determined using the initial conditions.

- Let $x_n = a_{m-1}x_{n-1} + a_{m-2}x_{n-2} + \cdots + a_0 x_{n-m}$ be a recurrence relation of order m that is satisfied by a sequence (x_n). Suppose the associated characteristic polynomial

$$\lambda^m - a_{m-1}\lambda^{m-1} - a_{m-2}\lambda^{m-2} - \cdots - a_0$$

factors as $(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_k)^{m_k}$, where $m_1 + m_2 + \cdots + m_k = m$. Then x_n has the form

$$x_n = (c_{11}\lambda_1^n + c_{12}n\lambda_1^n + c_{13}n^2\lambda_1^n + \cdots + c_{1m_1}n^{m_1-1}\lambda_1^n) + \cdots$$
$$+(c_{k1}\lambda_k^n + c_{k2}n\lambda_k^n + c_{k3}n^2\lambda_k^n + \cdots + c_{km_k}n^{m_k-1}\lambda_k^n)$$

- Let A be an $n \times n$ diagonalizable matrix and let $P = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$ be such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Then the general solution to the system $\mathbf{x}' = A\mathbf{x}$ is

$$\mathbf{x} = C_1 e^{\lambda_1 t}\mathbf{v}_1 + C_2 e^{\lambda_2 t}\mathbf{v}_2 + \cdots + C_n e^{\lambda_n t}\mathbf{v}_n$$

- Let A be an $n \times n$ diagonalizable matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$. Then the general solution to the system $\mathbf{x}' = A\mathbf{x}$ is $\mathbf{x} = e^{At}\mathbf{c}$, where $\mathbf{c}$ is an arbitrary constant vector. If an initial condition $\mathbf{x}(0)$ is specified, then $\mathbf{c} = \mathbf{x}(0)$.

- Let $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. The eigenvalues of A are $\lambda = a \pm bi$, and if a and b are not both zero, then A can be factored as

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

where $r = |\lambda| = \sqrt{a^2 + b^2}$ and θ is the principal argument of $a + bi$

- Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ (where $b \neq 0$) and corresponding eigenvector $\mathbf{x}$. Then the matrix $P = [\mathrm{Re}\mathbf{x} \quad \mathrm{Im}\mathbf{x}]$ is invertbible and

$$A = P \begin{bmatrix} a & -b \\ b & a \end{bmatrix} P^{-1}$$