CSE 840: Computational Foundations of Artificial Intelligence September 11, 2023

Eigenvalues, Characteristic Polynomial, Trace of Matrix

Instructor: Vishnu Boddeti *Scribe: Abhinay Devapatla, Yangru Zhou, Yihua Zhang*

1 Eigenvalues

Definition 1 *Let T: $V \to V$. A scalar $\lambda \in F$ is called an eigenvalue if there exists a $v \in V$, $v \neq 0$, such that $Tv = \lambda v$. A vector $v \neq 0$ with this property is called an eigenvector corresponding to the eigenvalue λ. The set of all eigenvalues of λ is called the eigenspace. $E(\lambda, T) = Ker(T - \lambda I)$*

Remark 2 *The following are remarks.*

- *Eigenvalue/eigenvectors realizes a "scaling", $v \to \lambda v$ direction of vector v does not change.*
- *Many linear mappings do not have eigenvectors, e.g. rotation. This is not true for algebraically closed fields like $\mathbb{C}$. $\mathbb{R}$ is not algebraically closed.*
- *If λ is an eigenvevalue, it has many eigenvectors, e.g. if v is an eigenvector, then any $a \cdot v$ ($a \in K$) is an eigenvector.*
  $$T(a \cdot v) = a \cdot Tv = a \cdot \lambda v = \lambda(a \cdot v) \implies T(a \cdot v) = \lambda(a \cdot v)$$
- *Eigenvectors corresponding to distinct eigenvalues are linearly independent. The intuition behind that is as follows. Suppose there are two distinct eigenvalues λ_1, λ_2 and $\lambda_1 \neq \lambda_2$. Assume v_1, v_2 are eigenvectors that are not linearly independent i.e. $v_2 = c \cdot v_1$.*
  $$Tv_1 = \lambda_1 v_1$$
  $$Tv_2 = \lambda_2 v_2 = \lambda_2(c \cdot v_1)$$
  $$Tv_2 = T(c \cdot v_1) = c \cdot Tv_1 = c \cdot \lambda_1 v_1$$
  $$Tv_2 \neq Tv_2, \text{ since } \lambda_2(c \cdot v_1) \neq c \cdot \lambda_1 v_1$$
  This contradiction shows that the eigenvectors are linearly independent.
- *Eigenvectors that correspond to the same eigenvalue do not need to be independent, e.g. v eigenvector $\to c \cdot v$ is also an eigenvector, but v & $c \cdot v$ are not linearly independent.*
- *They can be linearly independent: e.g. $A = I$, all eigenvalues are 1. $I \cdot v = 1 \cdot v$. But eigenvectors v can be linearly independent.*
- *The eigenspace $E(\lambda, T)$ is always a linear subspace of V.*

Proposition 3 *For finite-dimensional vector space, the following statements are equivalent:*

(i) *λ eigenvalue of T*

(ii) *$T - \lambda I$ not injective*

(iii) *$T - \lambda I$ not subjective*

(iv) *$T - \lambda I$ not bijective*

Proposition 4 *Suppose V is a finite-dimensional vector space, $T \in L(V)$, and $\lambda_1 ... \lambda_m$ are distinct eigenvalues of T. Then a sum of eigenspaces $E(\lambda_1,T) + E(\lambda_2,T) + ... + E(\lambda_m,T)$ is a direct sum. In particular $dim(E(\lambda_1,T)) + ... + dim(E(\lambda_m,T)) \leq dim\ V$.*

Theorem 5 *Every operator T: $V \to V$ on a finite-dimensional vector space with a complex field has at least one eigenvalue.*

Proof of Theorem 5: Let n = dim V. Choose a vector v $\in$ V, v $\neq$ 0. Then the set

$$v,\ Tv,\ T^2v, ..., T^nv$$

has to be linearly independent, since it consists of n+1 vectors in a n-dim vector space.

Find coefficients a_0, a_1,...,a_n such that

$$a_0v + a_1Tv + a_2T^2v + ... + a_nT^nv = 0.$$

Now consider a polynomial on $\mathbb{C}$ with the same coefficients: $P(z) := a_0 + a_1z + ... + a_nz^n$.

Over $\mathbb{C}$, we can factorize polynomial as: $P(z) = c \cdot (z - \lambda_1)((z - \lambda_2)....(z - \lambda_m)$ where $m \leq n$.

Consider again: $a_0v + a_1Tv + a_2T^2v + ... + a_nT^nv = 0$

$$\implies (a_0 + a_1T + a_2T^2 + ... + a_nT^n)v = 0$$
$$\implies c \cdot (T - \lambda_1)((T - \lambda_2)....(T - \lambda_m)v = 0$$
(factorization of polynomial $\to$ factorization of operator)
$$\implies v \in Ker(*) \text{ (injective } \leftrightarrow \text{ Ker0)}$$
$$\implies * \text{ is not an injective function}$$
$$\implies \text{there exists } i \in 1,...., m \text{ such that } (T - \lambda_iI \text{ is not injective}$$
$$\implies \lambda_i \text{ is an eigenvalue of T!}$$
'v' is not necessarily the eigenvector of T.

□

2 Characteristic Polynomial

Definition 6 *$Av = \lambda v$, where A is a nxn matrix and $v \neq 0$*

$$(A - \lambda I)v = 0$$
$$\implies v \in Ker(A - \lambda I)$$
$$\implies rank(A - \lambda I) < n)$$
$$dim(V) = dim(Ker(T)) + dim(image(T))$$
$$dim(Ker(A - \lambda I)) \geq 1$$
$$\implies det(A - \lambda I) = 0$$

Definition 7 *The characteristic polynomial of a nxn matrix A is defined as $P_A(t) := det(A - tI)$*

Example: $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

$det(A - tI) = det(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} - t\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$

$= det(\begin{pmatrix} a_{11} - t & a_{12} \\ a_{21} & a_{22} - t \end{pmatrix}) = (a_{11}-t)(a_{22}-t) - a_{12}a_{21} = t^2 - t(a_{11} + a_{22}) + a_{11}a_{22} - a_{12}a_{21}$

Observation 8 *The following is an observation.*

- $P_A(t)$ *is a polynomial with degree n if A is a nxn matrix*
- *Characteristic polynomials do not depend on the choice of basis*

  Proof:
  Consider A, basis transformation matrix U. Want to check if the characteristic polynomial for matrix A and UAU^{-1} *are the same.*
  $= det(UAU^{-1} - tI)$
  $= det(UAU^{-1} - tUU^{-1})$
  $= det(U(A-tI)U^{-1})$
  $= det(U) \cdot det(A-tI) \cdot det(U^{-1})$
  $= det(A - tI)$ ◇

- *The roots of the characteristic polynomial correspond exactly to the eigenvalues of A.*
- *Over* $\mathbb{C}$*, the characteristic polynomial always has n roots, so the matrix has "n eigenvalues" (not necessarily unique).*
- *A is invertible* $\iff$ *0 is not an eigenvalue.*

  If 0 is an eigenvalue, $Av = 0 \cdot v = 0$, $v \neq 0$

  $\implies$ *Ker(A) is non-trivial* $\iff$ *A not invertible*
- *Let* $A \in L(V)$, λ *is a eigenvalue of A. Then* λ^R *is an eigenvalue of* A^R*. Geometrically applying A twice will stretch the eigenvector twice:* $\lambda \cdot \lambda = \lambda^2$
- *Let A be invertible,* λ *be eigenvalue of A. Then* Y_λ *is an eigenvalue of* A^{-1}*. Geometrically inverse is unscaling* $\rightarrow Y_\lambda$

Definition 9 *For an operator A with eigenvalue* λ*, we define its geometric multiplicity as the dimension of the corresponding eigenspace* $E\lambda$, A.

Definition 10 *The algebraic multiplicity is the multiplicity of the root* λ *in the characteristic polynomial.*

Remark 11 *In general, geometric multiplicity and algebraic multiplicity are NOT the same.*

Remark 12 *Computing Eigenvalues and Eigenvectors*

- *Write down the characteristic polynomial, find the roots* $\rightarrow$ *eigenvalues.*
- *To compute eigenvectors, solve the linear system:* $Ax = \lambda x$

3 Trace of a Matrix

Definition 13 *The track of a square matrix $A \in F^{nxn}$ is the sum of its diagonal elements:*

$tr(A) = \sum_{i=1}^{n} a_{ii}$

Remark 14 *The following are remarks.*

- *tr:* $\mathbb{R}^{nxn} \to \mathbb{R}$ *is a linear operator in particular,* $tr(A+B) = tr(A) + tr(B)$
- $tr(A \cdot B) = tr(B \cdot A)$. *NOTE:* $tr(A \cdot B) \neq tr(A) \cdot tr(B)$
- *trace does not depend on the choice of basis*

  Let $T \in L(V)$*, and U and W be two bases of V. Then:* $tr(M(T,U)) = tr(M(T,W))$
- *The trace of an operator equal to the sum of its complex eigenvalues.*

  $\tilde{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ *w.r.t some basis* $v_1, v_2, ..., v_n$

  $tr(\tilde{A}) = \sum_{i=1}^{n} \lambda_i$

  Over $\mathbb{C}$*, we can always find basis of eigenvectors:* $A \in \mathbb{R}^{nxn}$

  Over $\mathbb{C}$ *I can find the representation:* $\tilde{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_n \end{pmatrix}$, $\lambda_i \in \mathbb{C}$

  $tr(\tilde{A}) = \sum \lambda_i = \sum_{i=1}^{n} a_{ii} = tr(A) \implies \sum \lambda_i \in \mathbb{R}$
- *trace equals the negative of the coefficient corresponding to the (n-1) degree term of the characteristic polynomial.*

  $P_A(t) = t^n + a_{n-1}t^{n-1} + ...$
- $tr(A)$ *= sum of its eigenvalues (if exists)*
- $det(A)$ *= product of its eigenvalues (if exists)*

Example: Consider a rotation matrix

- $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$
- $R(\theta)$ does not have any real eigenvalues
- The trace of $R(\theta)$ is $2\cos\theta$
- The characteristic polynomial of $R(\theta)$ is: $P_{R(\theta)}(t) := \det(R(\theta) - tI)$

  $= \det\begin{pmatrix} \cos\theta - t & -\sin\theta \\ \sin\theta & \cos\theta - t \end{pmatrix}$

  $= (\cos\theta - t)^2 + \sin^2\theta$

  $= t^2 = 2\cos\theta \cdot t + \cos^2\theta + \sin^2\theta$

  $= t^2 - 2\cos\theta \cdot t + 1$

- The roots of the characteristic polynomial

  $\lambda_{1/2} = \frac{2\cos\theta \pm \sqrt{(2\cos\theta)^2 - 4}}{2} = \cos\theta \pm \Im \sin\theta$
- The matrix has a diagonal representation

  $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

  $\mathrm{tr}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}) = \cos\theta + \Im \sin\theta + \cos\theta - \Im \sin\theta = 2\cos\theta$