Eigenvalue Problem

Notes on Linear Algebra

Chia-Ping Chen

Department of Computer Science and Engineering
National Sun Yat-Sen University
Kaohsiung, Taiwan ROC

Introduction

- Following $Ax = b$, we turn to the *eigenvalue problem*: Given a square matrix A, we look for λ's and x's such that

$$Ax = \lambda x.$$

  - λ is called an *eigenvalue*.
  - x is called an *eigenvector*.
- Both left and right sides involve unknown x.
- Only for some $\lambda' s$ can we find non-zero x for the above equation.

$Ax = b$ and $Ax = \lambda x$

- Both problems can be simplified solved by transforming to a diagonal or triangular form.
- Row operations do not change the solution of $Ax = b$. Such cannot be said for $Ax = \lambda x$.
- New operations that leaves the eigenvalues or eigenvectors unchanged will be introduced.

Differential Equations

- Suppose

$$\begin{cases} \frac{dv}{dt} = 4v - 5w, & v = 8 \text{ at } t = 0, \\ \frac{dw}{dt} = 2v - 3w, & w = 5 \text{ at } t = 0. \end{cases}$$

- We can re-write the above in a matrix form

$$\frac{du}{dt} = Au, \text{ with } u = u(0) \text{ at } t = 0,$$

where $u = \begin{bmatrix} v \\ w \end{bmatrix}$ and $A = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$.

An Eigenvalue Problem

- Assume the solution has an exponential form $u = e^{\lambda t} x$.
- Define $x = \begin{bmatrix} y \\ z \end{bmatrix}$. Substituting, we have

$$\begin{cases} 4y - 5z = \lambda y \\ 2y - 3z = \lambda z. \end{cases}$$

- This is an eigenvalue problem

$$Ax = \lambda x!$$

Finding Eigenvalues

- Suppose $Ax = \lambda x$ has a non-zero solution, then

$$(A - \lambda I)x = 0.$$

- $A - \lambda I$ must be singular since its nullspace is not empty. Thus,

$$|A - \lambda I| = 0.$$

- The above is the *characteristic equation*. It is solved for the eigenvalues of A.

Finding Eigenvectors

- Each eigenvalue has its own set of eigenvectors.
- To find the set $E_\lambda = \{x | Ax = \lambda x\}$, simply solve the system of linear equations.
- E_λ is a vector subspace, called eigenspace.
- The dimension of E_λ can be more than 1.
- The multiplicity of a root λ of the characteristic equation is called its algebraic multiplicity. The dimension of the eigenspace corresponding to λ is called geometric multiplicity.
- They do not always agree.

The Example

- For the example, the characteristic equation is

$$|A - \lambda I| = (4-\lambda)(3-\lambda) + 10 = \lambda^2 - \lambda - 2.$$

- The roots are $\lambda_1 = -1$ and $\lambda_2 = 2$. For $\lambda_1 = -1$,

$$\begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = 0 \Rightarrow x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- For $\lambda_2 = 2$,

$$\begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = 0 \Rightarrow x_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Solution

- The general solution is

$$u(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2 = c_1 e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

- The initial condition determines c_1 and c_2,

$$u(0) = c_1 x_1 + c_2 x_2 \Rightarrow \begin{bmatrix} 1 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \end{bmatrix} \Rightarrow \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

- The solution is

$$u = \begin{bmatrix} v \\ w \end{bmatrix} = 3e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5e^{2t} \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

The Example: Summary

- √ $\frac{du}{dt} = Au$ has an exponential solution $u = e^{\lambda t} x$.
- √ The eigenvalue λ gives the rate of growth or decay. The eigenvector x evolves at this rate.
- √ General solutions are mixtures of these pure exponential solutions.
- √ The mixing coefficients are adjusted to fit the initial values.
- √ Each component evolves “independently”.

Diagonal Matrix

- √ The eigenvalue problem for an diagonal matrix is easy to solve.
- √ Let $A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$.

$$A - \lambda I = \begin{bmatrix} 3-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix}$$

$$|A - \lambda I| = 0 \;\Rightarrow\; (3-\lambda)(2-\lambda) = 0$$

$$\lambda_1 = 3,\ x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix};\quad \lambda_2 = 2,\ x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Triangular Matrix

- Let $A = \begin{bmatrix} 1 & 4 & 5 \\ 0 & \frac{3}{4} & 6 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$.

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 4 & 5 \\ 0 & \frac{3}{4}-\lambda & 6 \\ 0 & 0 & \frac{1}{2}-\lambda \end{bmatrix}$$

$$|A - \lambda I| = 0 \;\Rightarrow\; (1-\lambda)(\frac{3}{4}-\lambda)(\frac{1}{2}-\lambda) = 0$$

- We often transform a matrix to a triangular form to find the eigenvalues.

Projection Matrix

- ✓ From $P^2 = P$, a projection matrix always has eigenvalues of 0 or 1!
- ✓ The projection point of an eigenvector of $\lambda = 1$ is itself. It must be in the column space of P.
- ✓ An eigenvector of $\lambda = 0$ is projected to the zero vector, so it must be in the nullspace of P.
- ✓ The column space is the eigenspace of $\lambda = 1$, and the nullspace is the eigenspace for $\lambda = 0$.
- ✓ There is no other eigenvalues as the sum of eigenspaces of $\lambda = 0, 1$ is R^n.

Example

√ Let $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

$$P - \lambda I = \begin{bmatrix} \frac{1}{2} - \lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} - \lambda \end{bmatrix}$$

$$|P - \lambda I| = 0 \;\Rightarrow\; \lambda(\lambda - 1) = 0$$

$$\lambda_1 = 1,\; x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix};\;\; \lambda_2 = 0,\; x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

0 as An Eigenvalue

- Recall that we are only interested in non-zero eigenvectors.
- However, 0, like any other numbers, can be an eigenvalue.
- A zero eigenvalue does not imply that the eigenvector is zero.
- If A is singular, then 0 must be an eigenvalue.
- If A is non-singular (invertible), then 0 cannot be an eigenvalue.

Trace

- The *trace* of a matrix is the sum of diagonal entries
$$tr(A) = a_{11} + \cdots + a_{nn}.$$
- The sum of eigenvalues equals trace,
$$\sum_{i=1}^{n} \lambda_i = tr(A).$$
- The product of eigenvalues equals determinant,
$$\prod_{i=1}^{n} \lambda_i = |A|.$$

Characteristic Equation

- ✓ Let $\{\lambda_i\}_{i=1}^n$ be the eigenvalues, then

$$f(\lambda) = |A - \lambda I| = (-1)^n(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0)$$
$$= \prod_{i=1}^{n}(\lambda_i - \lambda).$$

- ✓ So $|A| = f(\lambda = 0) = \prod_{i=1}^n \lambda_i$.
- ✓ Comparing the coefficients of λ^{n-1}, we have

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

Eigenvector Matrix

- Suppose A has independent eigenvectors $u_1, \ldots, u_n$. Using the eigenvectors as columns of S, then

$$AS = S\Lambda, \text{ where } \Lambda = \operatorname{diag}\begin{bmatrix}\lambda_1 & \lambda_2 & \ldots & \lambda_n\end{bmatrix}$$

- S is the *eigenvector matrix* and Λ is the *eigenvalue matrix*.
- Since S is invertible,

$$AS = S\Lambda \Leftrightarrow A = S\Lambda S^{-1} \Leftrightarrow \Lambda = S^{-1}AS.$$

Diagonalization

- If $AS = S\Lambda$ for diagonal λ, then A is said to be diagonalized by S.
- S is not unique, but it must be an eigenvector matrix. The columns of S must be eigenvectors.
- Not all matrices can be diagonalized. Some matrices do not have enough independent eigenvectors.
- A matrix with no repeated eigenvalues can be diagonalized, as we show next.

Distinct Eigenvalues

√ Eigenvectors of different eigenvalues are linearly independent. This can be proved by induction on the number of eigenvalues. For $n = 2$,

$$A(c_1x_1 + c_2x_2) = 0, \lambda_2(c_1x_1 + c_2x_2) = 0$$

$$\Rightarrow c_1(\lambda_1 - \lambda_2)x_1 = 0 \Rightarrow c_1 = 0 = c_2.$$

Suppose it is true for k. For $k + 1$, since

$$A(\sum_{i=1}^{k+1} c_i x_i) = 0, \text{ and } \lambda_{k+1}(\sum_{i=1}^{k+1} c_i x_i) = 0,$$

$$\sum_{i=1}^{k} c_i(\lambda_i - \lambda_{k+1})x_i = 0 \Rightarrow c_1 = \cdots = c_k = 0 = c_{k+1}.$$

Defective Matrices

- A defective matrix does not have enough linearly independent eigenvectors.
- Consider
$$A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}.$$
- $\lambda = 3$ is a double root to $|A - \lambda I| = 0$, but there is only one independent vector for $\lambda = 3$.
- The existence S leads to contradiction
$$S^{-1}AS = \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \Rightarrow SS^{-1}ASS^{-1} = A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Diagonalization Examples

- ✓ A projection matrix

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \; \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \; S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

- ✓ A 90° rotation matrix

$$K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \; \Lambda = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \; S = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}$$

- ✓ Note that a real matrix may have imaginary (complex) eigenvalues and eigenvectors.

Power of a Matrix

- ✓ An eigenvector of A is an eigenvector of A^2, with the eigenvalue squared.

$$A^2x = A(Ax) = A\lambda x = \lambda Ax = \lambda^2 x.$$

- ✓ This result can be generalized to A^k.
- ✓ A matrix S diagonalizing A also diagonalizes A^k,

$$A^k = (S\Lambda S^{-1})^k = S\Lambda^k S^{-1} \Rightarrow \Lambda^k = S^{-1}A^kS.$$

- ✓ If A is invertible, then

$$Ax = \lambda x \Rightarrow x = \lambda A^{-1}x \Rightarrow A^{-1}x = \frac{1}{\lambda}x.$$

Product AB

- Suppose A has eigenvalue μ and B has eigenvalue λ. Does AB has eigenvalue $\mu\lambda$?
  - ⋆ No in general.
  - ⋆ Yes if they have the same eigenvector.
- Suppose A and B are diagonalizable. A and B have the same eigenvector matrix iff $AB = BA$.

$$(\Rightarrow)\ AB = S\Lambda_A S^{-1} S\Lambda_B S^{-1} = S\Lambda_A\Lambda_B S^{-1}$$
$$= S\Lambda_B\Lambda_A S^{-1} = S\Lambda_B S^{-1} S\Lambda_A S^{-1} = BA$$

$(\Leftarrow)$ For simplicity, suppose all eigenvalues of A are distinct.

$$Ax = \lambda_a x \Rightarrow ABx = BAx = B\lambda_a x = \lambda_a Bx.$$

x and Bx are eigenvectors of A (for λ_a), so $Bx = \lambda_b x$.

Uncertainty Principle

- In quantum mechanics operators are represented by matrices.
- Position P (symmetric) and momentum Q (skew-symmetric) do not commute,

$$QP - PQ = I.$$

- The uncertainty principle is a result of the Schwartz inequality $(Qx)^T(Px) \le |Qx||Px|$. Let x be a wave function,

$$|x|^2 = x^T x = x^T(QP-PQ)x \le 2|Qx||Px| \Rightarrow \frac{|Qx|}{|x|}\frac{|Px|}{|x|} \ge \frac{1}{2}.$$

Discrete and Continuous

- Suppose you invest 1000 at 6% rate for five years.

  compounded once a year $1000(1.06)^5 = 1338$

  once a month $1000(1.005)^{5*12} = 1349$

  once a day $1000(1 + \frac{0.06}{365})^{5*365} = 1349.83$

  every instant $\lim_{N\to\infty} 1000(1 + \frac{0.06}{N})^{5*N} = 1349.87$

- They are not that different. Discrete picture is more concrete.

Difference Equations

- Difference equation evolves in finite steps.
- Difference equations can be used to approximate differential equations.
- Difference equations can also arise when the underlying problem is discrete in nature, as in a sequence.

Fibonacci Sequence

- The Fibonacci sequence is defined recursively by,
  - basis: $F_0 = 0, F_1 = 1.$
  - recursion: $F_{k+2} = F_k + F_{k+1}$.
- First few numbers in the sequence are

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

- This sequence is defined by a difference equation with an initial condition.
- Let's find F_k as a function of k.

Matrix Representation

- √ The recursion can be represented by a matrix. Let

$$u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}, \; A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow \; u_{k+1} = Au_k.$$

- √ Suppose A is diagonalized by S, then

$$u_k = A^k u_0 = S\Lambda^k S^{-1} u_0 = c_1\lambda_1^k x_1 + c_2\lambda_2^k x_2,$$

where λ_i is the ith eigenvalue and x_i is the corresponding eigenvector of A.

$$c = S^{-1}u_0.$$

Solution

- For the eigenvalue λ's, we solve $|A - \lambda I| = 0$.

$$|A - \lambda I| = \lambda^2 - \lambda - 1 = 0 \Rightarrow \lambda = \frac{1 \pm \sqrt{5}}{2}.$$

- The eigenvector for λ is $(\lambda, 1)$ since the second row of $A - \lambda I$ is $(1, -\lambda)$.
- In addition,

$$c = S^{-1} u_0 = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\lambda_1 - \lambda_2} \\ \frac{-1}{\lambda_1 - \lambda_2} \end{bmatrix}.$$

Formula

- From $u_k = c_1\lambda_1^k x_1 + c_2\lambda_2^k x_2$, we have

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = c_1\lambda_1^k \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} + c_2\lambda_2^k \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$$

$$\Rightarrow F_k = \frac{1}{\lambda_1 - \lambda_2}(\lambda_1^k - \lambda_2^k) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right].$$

- Since the second term is always less then $\frac{1}{2}$, we can say F_k is the nearest integer to $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^k$.

- The ratio of $\frac{F_{k+1}}{F_k}$ approaches λ_1.

Comments

- ✓ We can write u_0 as a linear combination of the eigenvectors of A. It follows that u_k is the same combination with $\lambda_i^k x_i$'s. That is,

$$\text{If } u_0 = c_1 x_1 + \cdots + c_n x_n,$$

$$\text{then } u_k = c_1 \lambda_1^k x_1 + \cdots + c_n \lambda_n^k x_n.$$

- ✓ The role of c is to match the initial condition.
- ✓ As a special case, if the initial u_0 happens to be an eigenvector x with eigenvalue λ, then $u_k = \lambda^k x$.

Markov Process

- $\surd$ We introduce another class of problems whose solutions depend on using matrices wisely.
- $\surd$ Suppose for city X every year
  - $\star$ $\frac{1}{10}$ of the people outside move in;
  - $\star$ $\frac{2}{10}$ of the people inside move out.
- $\surd$ This is a *Markov process*. The population at the end of year i, given the population at the end of year $i-1$, does not depend on any populations before that.

Markov Matrix

- √ Let y_i be the people outside and z_i be the people inside at the end of year i. Then

$$\begin{cases} y_{i+1} = 0.9y_i + 0.2z_i \\ z_{i+1} = 0.1y_i + 0.8z_i \end{cases} \quad \Rightarrow x_{i+1} = Ax_i,$$

where $A = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$.

- √ A is a Markov matrix:
  - ⋆ A has only non-negative entries.
  - ⋆ The entries in any column of A sum to 1.

Eigenvalues of Markov Matrix

- The eigenvalues of A are

$$|A - \lambda I| = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = 0.7.$$

- The eigenvectors are

$$x_1 = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \; x_2 = \begin{bmatrix} \frac{1}{3} \\ \frac{-1}{3} \end{bmatrix}.$$

- Note that $A \geq 0 \Rightarrow A^k \geq 0$ (having no negative entries), and we are guaranteed that if the initial x_0 are non-negative, all x_i will be non-negative.

Solution

- √ The solution is a linear combination of two terms.

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} = A^k \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} = S\Lambda^k S^{-1} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

$$= (y_0 + z_0) \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} + (y_0 - 2z_0)(0.7)^k \begin{bmatrix} \frac{1}{3} \\ \frac{-1}{3} \end{bmatrix}$$

- √ The first term in invariant with time, while the second term is exponentially small.
- √ The first term dominates the second term as time goes by.

Steady State

- ✓ Since the column sum is 1, the column sum of $A - I$ is 0. So $\lambda = 1$ is an eigenvalue of A.
- ✓ The eigenvector x for $\lambda = 1$ is a steady state. Once the system is a steady state, it remains in that state,
$$Ax = x.$$

Stability of a Difference Equation

- Suppose we are given $u_{k+1} = Au_k$, and we want to study the behavior of u_k as $k \to \infty$.
- Assuming A can be diagonalized, then

$$u_k = A^k u_0 = S\Lambda^k S^{-1} u_0 = c_1\lambda_1^k x_1 + \cdots + c_n \lambda_n^k x_n.$$

- The growth of u_k depends on the λ_i's.
  - *stable* case: $|\lambda_i| < 1$ for all i. $u^k \to 0$.
  - *neutrally stable* case: $|\lambda_i| = 1$ for some i and $|\lambda_i| < 1$ for others.
  - *unstable* case: $\exists i\ |\lambda_i| > 1$. u_k is unbounded.

Leontief's Input-Output Matrix

- Consumption matrix C: c_{ij} is the amount of product i needed to create one unit of product j.
- Consider the case of steel, service, and labor.
- Can we meet *external* demand of y_1 of steel, y_2 of service and y_3 of labor?
- Let p be total amount. Then

$$p = Cp + y,$$

- p has to be non-negative and

$$p - Cp = y \ \text{ or } \ p = (I - C)^{-1} y.$$

Condition on Eigenvalue

- ✓ Note that $(I - C)^{-1}$ has the same eigenvector as C, with eigenvalue $\frac{1}{1-\lambda}$.

$$(I - C)x = (1 - \lambda)x \Rightarrow \frac{1}{1 - \lambda}x = (I - C)^{-1}x.$$

- ✓ A sufficient condition is $\lambda_1 < 1$,

$$I + C + C^2 + \dots \text{ converges to } (I - C)^{-1},$$

and it is nonnegative.

Prices

- We may be more interested in price than in production. We can use p to represent prices instead of production levels.
- Let p_0 be a vector of prices. Then $p_1 = Cp_0$ is the vector of costs. p_1 is a reasonable vector of prices.
- In equilibrium, the price reflects the cost, so

$$p = Cp.$$

Perron-Frobenius Theorem

- Let A be a positive matrix ($a_{ij} > 0,\ \forall\, i, j$). The largest eigenvalue λ_1 is positive and x_1 is positive,

$$A > 0 \Rightarrow \lambda_1 > 0, x_1 > 0.$$

- Define $T = \{t \mid \exists\, x \geq 0 \text{ s.t. } Ax \geq tx\}$ and let $t_M = \sup_{t \in T} t$. We show $Ax = t_M x$ for some $x \geq 0$.
- Suppose $Ax > t_M x,\ \forall x \geq 0$. Multiplying both sides by A we have $A^2 x > t_M Ax$, or $Ay > t_M y$ where $y = Ax$. This contradicts that t_M is the least upper bound.
- x is positive since Ax is sure to be positive.

A Differential Equation

- ✓ Suppose

$$\frac{du}{dt} = Au = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u.$$

- ✓ The eigenvalues and eigenvectors of A are,

$$\lambda_1 = -1, x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; \lambda_2 = -3, x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

- ✓ So a general solution is

$$u(t) = c_1 e^{\lambda_1 t} x_1 + c_2 e^{\lambda_2 t} x_2 = c_1 e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{-3t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Solution

- ✓ We apply the initial condition to find c_1, c_2,

$$u(t=0) = c_1 x_1 + c_2 x_2, \quad \text{or} \quad Sc = u(0).$$

So

$$c = S^{-1} u(0).$$

- ✓ The solution is

$$u(t) = S \begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix} S^{-1} u(0) = S e^{\Lambda t} S^{-1} u(0).$$

The Exponential of a Matrix

- In the previous slide we introduce a notation using a matrix as exponent.
- Formally, the exponential of a matrix is defined by

$$e^M = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \dots$$

- Specifically, if A can be diagonalized, then

$$e^{At} = I + At + \frac{(At)^2}{2!} + \cdots = I + S\Lambda S^{-1}t + \frac{(S\Lambda S^{-1}t)^2}{2!} + \dots$$

$$= S(I + \Lambda t + \frac{(\Lambda t)^2}{2!} + \dots)S^{-1} = Se^{\Lambda t}S^{-1}$$

Example

- Let

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

- What is e^A? Is it

$$\begin{bmatrix} e^{-2} & e^{1} \\ e^{1} & e^{-2} \end{bmatrix}?$$

- No!

Example

$$e^A = Se^{\Lambda}S^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} e^{\begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-1} & 0 \\ 0 & e^{-3} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}$$

$$= \frac{1}{2} \begin{bmatrix} e^{-1} + e^{-3} & e^{-1} - e^{-3} \\ e^{-1} - e^{-3} & e^{-1} + e^{-3} \end{bmatrix}.$$

Solution of Differential Equations

- ✓ If A can be diagonalized, then the equation $\frac{du}{dt} = Au, \; u(0) = u_0$ has the solution

$$u(t) = Se^{\Lambda t}S^{-1}u_0 = e^{At}u_0.$$

- ✓ The columns of S are eigenvectors of A, so

$$u(t) = c_1 e^{\lambda_1 t}x_1 + \cdots + c_n e^{\lambda_n t}x_n.$$

- ✓ It is a combination of exponentials, with coefficients chosen to match the initial condition $c = S^{-1}u_0$.

e^{At}

- e^{At} is never singular.

$$|e^{At}| = |e^{\Lambda t}| = \prod e^{\lambda_i t} = e^{tr(A)t} \neq 0.$$

- e^{At} sometimes acts like a scalar,

$$(e^{At})(e^{As}) = e^{A(t+s)}$$

$$(e^{At})^{-1} = e^{-At}$$

$$\frac{d}{dt}(e^{At}) = Ae^{At}$$

Linearly Independent Solutions

- ✓ If n solutions are linearly independent at $t = 0$, then they remain linearly independent at all times.
- ✓ This is a result of the non-singularity of e^{At}. If

$$\left| \begin{bmatrix} v_1(0) & v_2(0) & \ldots & v_n(0) \end{bmatrix} \right| \neq 0,$$

then

$$\left| \begin{bmatrix} v_1(t) & \ldots & v_n(t) \end{bmatrix} \right| = |e^{At}| \left| \begin{bmatrix} v_1(0) & \ldots & v_n(0) \end{bmatrix} \right| \neq 0.$$

Stability of a Differential Equation

- What is the asymptotic behavior of $\frac{du}{dt} = Au$ as $t \to \infty$?
- Assuming A can be diagonalized,

$$u(t) = c_1 e^{\lambda_1 t} x_1 + \cdots + c_n e^{\lambda_n t} x_n.$$

- Stability is controlled by $e^{\lambda_i t}$'s. The equation is said to be
  - stable if Re $\lambda_i < 0$ for all i. $e^{At} \to 0$.
  - neutrally stable if Re $\lambda_i = 0$ for some i and Re $\lambda_i \leq 0$ for all i.
  - unstable if Re $\lambda_i > 0$ for some i. $u(t)$ is unbounded.

2×2 **Matrix**

- ✓ Stability is easy to decide for a real 2×2 matrix.

$$\frac{du(t)}{dt} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} u = Au.$$

- ✓ If $tr(A)$ is negative and $|A|$ is positive, then the above equation is stable. Otherwise it cannot be stable.

Rotation Matrix

- √ An interesting example is

$$\frac{du(t)}{dt} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} u = Au.$$

- √ The eigenvalues are i and $-i$, with eigenvectors

$$x_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

$$u(t) = c_1 e^{it} x_1 + c_2 e^{-it} x_2 = c_1 e^{it} \begin{bmatrix} 1 \\ -i \end{bmatrix} + c_2 e^{-it} \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

Rotating Vector

√ Replacing $e^{\pm it} = \cos t \pm i \sin t$,

$$u(t) = \begin{bmatrix} (c_1 + c_2)\cos t + i(c_1 - c_2)\sin t \\ -i(c_1 - c_2)\cos t + (c_1 + c_2)\sin t \end{bmatrix}.$$

√ Let $u(0) = \begin{bmatrix} a \\ b \end{bmatrix}$.

$$\begin{cases} a = c_1 + c_2 \\ b = -i(c_1 - c_2) \end{cases} \quad \Rightarrow u(t) = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}.$$

Real vs. Complex

- In eigenvalue problems, it is no longer possible to work only with real numbers.
- The set of complex numbers includes the set of real numbers. Definitions and properties for complex matrices are valid for real matrices as well.
- Theorems valid for the real set can be extended to the complex set.

Complex Vectors

- A complex vector in $\mathbf{C}^n$ has n complex components,

$$x = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T, \quad x_k = a_k + i b_k.$$

- The length of a complex vector is defined by

$$|x|^2 = |x_1|^2 + \cdots + |x_n|^2, \quad \text{where } |x_k|^2 = a_k^2 + b_k^2.$$

- It agrees with the original definition when x is real.

Inner Product

- The inner product of two vectors is defined by

$$(x, y) = \overline{x}^T y = \overline{x}_1 y_1 + \cdots + \overline{x}_n y_n.$$

- Note that

$$|x|^2 = (x, x)$$
$$(y, x) = (x, y)^*.$$

- Two complex vectors are said to be orthogonal if

$$(x, y) = (y, x)^* = 0.$$

Hermitian of a Matrix

- The *Hermitian* of a matrix A, denoted by A^H, is the *conjugate transpose* of A. That is,

$$A^H = \overline{A}^T, \ (A^H_{ij} = \overline{A}_{ji})$$

- Inner product of two vectors can be written as

$$(x, y) = x^H y.$$

- For any x, A, y,

$$(x, Ay) = x^H Ay = (A^H x)^H y = (A^H x, y).$$

Hermitian Matrix

- A matrix A is said to be *Hermitian* if

$$A = A^H.$$

- The following matrix is Hermitian,

$$\begin{bmatrix} 3 & 2-2i \\ 2+2i & 3 \end{bmatrix}.$$

- Hermitian matrix is the complex counterpart of symmetric matrix.

Property

- √ Let A be Hermitian.
- √ $x^H Ax$ is real for any complex vector x.

$$x^H Ax = (x, Ax) = (x, A^H x) = (Ax, x) = \overline{(x, Ax)}.$$

- √ The eigenvalues are real.

$$(x, Ax) = \lambda(x, x) \Rightarrow \lambda = \frac{(x, Ax)}{(x, x)}.$$

- √ The eigenvectors are orthogonal.

$$(Ax_1, x_2) = (x_1, Ax_2) \Rightarrow (\lambda_1 - \lambda_2)(x_1, x_2) = 0.$$

Real Symmetric Matrices

- Since a real symmetric matrix is Hermitian by definition, the eigenvalues are real and the eigenvectors are orthogonal.
- The eigenvectors can be chosen orthonormal. They go to the columns of an orthogonal matrix Q which diagonalizes A,

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

Spectral Theorem

- √ If we multiply columns by rows (block matrix multiplication) of $A = Q\Lambda Q^T$, then

$$A = \lambda_1 x_1 x_1^T + \cdots + \lambda_n x_n x_n^T.$$

- √ This is also known as the spectral theorem.
- √ A real symmetric matrix is a linear combination of one-dimensional projection matrices (to the eigenvectors). These projection matrices are symmetric with rank 1.

Unitary Matrices

- By definition, a matrix U is *unitary* if

$$U^H U = U U^H = I.$$

- This is a generalization of orthogonal matrices. If U is real, then U is orthogonal.

Property

- Inner product (so is length) is preserved by a unitary transformation,

$$(Ux, Uy) = x^H U^H U y = x^H y.$$

- Eigenvalues of U have unit modulus $|\lambda| = 1$,

$$|x| = |Ux| = |\lambda x| = |\lambda||x|.$$

- Eigenvectors of U corresponding to different eigenvalues are orthogonal, since

$$(Ux_1, Ux_2) = (x_1, U^H U x_2) = (x_1, x_2) \Rightarrow (1 - \overline{\lambda_1}\lambda_2)(x_1, x_2) = 0.$$

Similarity Transformations

- ✓ Suppose A is diagonalizable so $\Lambda = S^{-1}AS$.
  - ⋆ If A is symmetric, then we write Q instead of S since it is orthogonal.
  - ⋆ If A is Hermitian, then we write U instead of Q since it is unitary.
- ✓ The transforming matrix needs not be an eigenvector matrix: We look at matrix $B = M^{-1}AM$ where M is invertible. A and B are said to be *similar*. Going from A to B is a *similarity transformation*.

Example

✓ Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

$$\begin{cases} M_1 = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}, M_1^{-1} = \begin{bmatrix} 1 & -b \\ 0 & 1 \end{bmatrix}, & M_1^{-1} A M_1 = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix} \\ M_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, M_2^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, & M_2^{-1} A M_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \end{cases}.$$

✓ $\begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ are similar to A.

Key Questions

- √ What do A and $M^{-1}AM$ have in common?
- √ The eigenvalues of A and $M^{-1}AM$ are the same. In the example, one can see that A, $M_1^{-1}AM_1$ and $M_2^{-1}AM_2$ have the same eigenvalues 0 and 1.
- √ With a special choice of M, what special form can be achieved by $M^{-1}AM$?
- √ The Jordan form can be achieved with special choice of M.

Eigenvalues of Similar Matrices

- √ A and $B = M^{-1}AM$ have the same eigenvalues. Furthermore, if x is an eigenvector of A, then $M^{-1}x$ is an eigenvector of B.
- √ A and B have the same characteristic equations, so they have the same roots.

$$|B - \lambda I| = |M^{-1}(A - \lambda I)M| = |A - \lambda I|.$$

- √ The eigenvectors are related via M, as

$$Ax = \lambda x \Rightarrow M^{-1}Ax = \lambda M^{-1}x \Rightarrow M^{-1}AMM^{-1}x = \lambda M^{-1}x$$
$$\Rightarrow BM^{-1}x = \lambda M^{-1}x.$$

Change of Basis

- √ Recall that a linear transformation can be represented by a matrix.
- √ The matrix representation is dependent on the basis. When the basis is changed, the matrix representation needs to be changed accordingly.
- √ We will see that similar matrices represent the same linear transformations in different bases.

Two Bases

- Suppose we have a basis $v = \{v_1, \ldots, v_n\}$ and a linear transform defined by

$$Tv_j = a_{1j}v_1 + \cdots + a_{nj}v_n, \;\; j = 1 \ldots n.$$

- Consider another basis $V = \{V_1, \ldots, V_n\}$. In this basis, T is defined by

$$TV_j = b_{1j}V_1 + \cdots + b_{nj}V_n, \;\; j = 1 \ldots n.$$

Matrices

- Since v is a basis, one can write

$$V_j = m_{1j}v_1 + \cdots + m_{nj}v_n, \; j = 1 \ldots n.$$

- Identical vectors are transformed identically, so

$$\begin{cases} TV_j = b_{1j}V_1 + \cdots + b_{nj}V_n = \sum_{p,l} b_{pj}m_{lp}v_l \\ TV_j = T(\sum_k m_{kj}v_k) = \sum_k m_{kj}T(v_k) = \sum_{k,l} m_{kj}a_{lk}v_l \end{cases}$$

- It follows that

$$AM = MB \Rightarrow B = M^{-1}AM.$$

Example

- Suppose that

$$
\begin{aligned}
v &= \{v_1, v_2\} = \{\begin{bmatrix}1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\end{bmatrix}\} \\
V &= \{V_1, V_2\} \\
&= \{\cos\theta\, v_1 + \sin\theta\, v_2, -\sin\theta\, v_1 + \cos\theta\, v_2\}.
\end{aligned}
$$

- Then

$$
M = \begin{bmatrix} c & -s \\ s & c \end{bmatrix},\ M^{-1} = \begin{bmatrix} c & s \\ -s & c \end{bmatrix},
$$

$$
B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \Leftrightarrow A = \begin{bmatrix} c^2 & cs \\ cs & s^2 \end{bmatrix}.
$$

Computation of Eigenvalues

- √ If we apply a sequence of similar transformation, the sequence of similar matrices obtained have the same eigenvalues.
- √ If we can make the matrix form go toward a triangular form, the eigenvalues will be obvious.
- √ This is much better than solving $|A - \lambda I| = 0$ directly.

Schur's Lemma

- Given A, we want to find a M such that $M^{-1}AM$ is triangular.
- The Schur's lemma guarantees the existence of such a matrix.
- For any square matrix A there exists a unitary matrix U such that

$$U^{-1}AU = T$$

  is upper-triangular.
- Note that unitary matrices are used.

Construction of Unitary Matrix

- ✓ A matrix A, say of size 4×4, has at least one eigenvalue λ_1 with a unit-length eigenvector x_1.
- ✓ Place x_1 in the first column of U_1, and fill in the other columns with entries so that U_1 is unitary. Then

$$AU_1 = U_1 \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix} \Rightarrow U_1^{-1} A U_1 = \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}$$

Construction of Unitary Matrix

- Let $B = U_1^{-1}AU_1$ and B_{lr} be the lower-right 3×3 sub-matrix of B.
- Let λ_2, x_2 be an eigenvalue and eigenvector of B_{lr}.
- Let U_2 be a unitary matrix whose lower-right submatrix M_2 is unitary using x_2 as the first column, then

$$BU_2 = \begin{bmatrix} \lambda_1 & v^T \\ 0 & B_{lr} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & M_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & v^T M_2 \\ 0 & B_{lr} M_2 \end{bmatrix}.$$

Construction of Unitary Matrix

- The first column vector of M_2 is an eigenvector of B_{lr}, so

$$
\begin{aligned}
B_{lr}M_2 = B_{lr}\begin{bmatrix} x_2 & * & * \end{bmatrix} &= \begin{bmatrix} \lambda_2 x_2 & * & * \end{bmatrix} \\
&= \begin{bmatrix} x_2 & * & * \end{bmatrix}\begin{bmatrix} \lambda_2 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \\
&= M_2\begin{bmatrix} \lambda_2 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix}.
\end{aligned}
$$

Construction of Unitary Matrix

- It follows that

$$BU_2 = \begin{bmatrix} \lambda_1 & v^T M_2 \\ \mathbf{0} & B_{lr} M_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & M_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & v^T M_2 \\ \mathbf{0} & \begin{bmatrix} \lambda_2 & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \end{bmatrix} = U_2 C.$$

- One more step for $CU_3 = U_3 T$. $U = U_1 U_2 U_3$ is unitary and $U^{-1} A U = T$ is triangular.

Complete Set of Eigenvectors

- √ Any Hermitian matrix (including real symmetric matrices) has a complete set of orthonormal eigenvectors.
- √ It does not matter whether the eigenvalues are distinct or not.
- √ This follows from the Schur's lemma.

Diagonalizing Hermitian Matrices

- ✓ Let A be Hermitian and it is triangularized by U.
- ✓ $U^{-1}AU$ is Hermitian, since

$$(U^{-1}AU)^H = U^H A^H (U^{-1})^H = U^{-1}AU.$$

- ✓ A triangular Hermitian matrix must be diagonal.
- ✓ U is indeed an eigenvector matrix of A, since $U^{-1}AU$ diagonalizes A.
- ✓ The columns of U consist a complete set of orthonormal eigenvectors of A.

Spectral Decomposition

- For a Hermitian matrix A,

$$A = U\Lambda U^H, \text{ where } U \text{ is unitary.}$$

- Every Hermitian matrix with k different eigenvalues can be expressed as a sum of k projection matrices,

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k,$$

where P_i projects to the eigenspace of λ_i. Since the eigenspaces are orthogonal, $P_i P_j = 0,\ i \neq j$.

Normal Matrices

- A matrix N is *normal* if N commutes with N^H,

$$NN^H = N^H N.$$

- As a special case, a Hermitian matrix is a normal matrix.
- A normal matrix has a complete set of orthonormal eigenvectors. We prove this by showing the triangular form is diagonal.

Normal Matrices

- √ If N is normal, then $U^{-1}NU$ is normal.

$$(U^H NU)(U^H NU)^H = U^H NUU^H N^H U = U^H NN^H U$$

$$(U^H NU)^H (U^H NU) = U^H N^H UU^H N^H U = U^H N^H NU$$

- √ Let T be normal and (upper-)triangular.
  - ⋆ Comparing $(TT^H)_{11}$ and $(T^H T)_{11}$, one can see that $T_{1j} = 0, j \neq 1$.
  - ⋆ Comparing $(TT^H)_{22}$ and $(T^H T)_{22}$, one can also see that $T_{2j} = 0, j \neq 2$.
  - ⋆ Continuing, T is shown to be diagonal.

Defective Matrices

- √ If A has a complete set of eigenvectors, we take $M = S$ and $S^{-1}AS = \Lambda$ is diagonal.
- √ For a defective matrix, this is impossible, since there are not enough eigenvectors.
- √ We want to make $M^{-1}AM = J$ as diagonal as possible.
- √ Result: in addition to the eigenvalues on the main diagonal (we do have enough eigenvalues), there will be a 1 just above the main diagonal for every missing eigenvector.

Jordan Form

- √ If A has s independent eigenvectors, it is similar to a matrix with s blocks,

$$J = M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}.$$

- √ J is said to be in Jordan form.
- √ J_i is a triangular matrix with eigenvalue λ_i for the diagonal entries and 1's just above the diagonal.

Examples

✓ Matrices

$$T = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

share the same Jordan form

$$J = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

✓ T, A, B are similar.

Comments on Jordan Form

- ✓ Every missing eigenvector has a "1" just above the main diagonal.
- ✓ J is diagonal iff each block is of size 1×1.
- ✓ Only a repeated eigenvalue may require off-diagonal 1's in J.
- ✓ Multiple blocks may have the same λ.
- ✓ Two matrices may have the same set of eigenvalues but different Jordan forms.