ECE504: Lecture 5

D. Richard Brown III

Worcester Polytechnic Institute

06-Oct-2009

Lecture 5 Major Topics

We are still in Part II of ECE504: **Quantitative and qualitative analysis of systems**

mathematical description → results about behavior of system

Our focus is going to turn to LTI systems for a while. Recall that, when $\boldsymbol{A}$ is not a function of time, the STMs become

$$\begin{aligned} \boldsymbol{\Phi}[k,j] &= \boldsymbol{A}^{k-j} \text{ (discrete time)} \\ \boldsymbol{\Phi}(t,s) &= \exp\{(t-s)\boldsymbol{A}\} \text{ (continuous time)} \end{aligned}$$

Also recall that these are not computed element-by-element. Today, we will develop techniques for efficient computation of $\boldsymbol{A}^{k-j}$ and $\exp\{(t-s)\boldsymbol{A}\}$.

You should be reading Chen Chapter 4 now and referring back to Chapter 3 for the necessary linear algebraic concepts.

Basic Properties of $\boldsymbol{A}^k$ and $\exp\{t\boldsymbol{A}\}$

Recall that the matrix exponential $\exp\{t\boldsymbol{A}\}$ is not performed element-by-element. Nevertheless, the matrix exponential has many of the same properties as the usual scalar exponential. Specifically:

1. For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$

$$\lim_{t \to 0} \exp\{t\boldsymbol{A}\} = \boldsymbol{I}_n$$

   This can be seen directly from the definition of $\exp\{t\boldsymbol{A}\}$.

2. For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$

$$\exp\{(t_1 + t_2)\boldsymbol{A}\} = \exp\{t_1\boldsymbol{A}\} \exp\{t_2\boldsymbol{A}\}$$

   This is a consequence of the semigroup property of $\boldsymbol{\Phi}(t, s)$.

3. Given $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and $\tilde{\boldsymbol{A}} \in \mathbb{R}^{n \times n}$, does

$$\exp\{t(\boldsymbol{A} + \tilde{\boldsymbol{A}})\} = \exp\{t\boldsymbol{A}\} \exp\{t\tilde{\boldsymbol{A}}\}?$$

Basic Properties of $\boldsymbol{A}^k$ and $\exp\{t\boldsymbol{A}\}$ (cont.)

4. Given any $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ such that

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a_{nn} \end{bmatrix} \text{ is diagonal, then } \exp\{t\boldsymbol{A}\} = \begin{bmatrix} e^{a_{11}t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{a_{nn}t} \end{bmatrix}$$

   This can be seen directly from the definition of $\exp\{t\boldsymbol{A}\}$.

5. Given any $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ such that we can find some invertible $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ satisfying

$$\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_{nn} \end{bmatrix}$$

   then

$$\boldsymbol{A}^k = \underbrace{(\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1})\dots(\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1})}_{k-\text{fold product}} = \boldsymbol{V}\boldsymbol{\Lambda}^k\boldsymbol{V}^{-1}$$

Basic Properties of $\boldsymbol{A}^k$ and $\exp\{t\boldsymbol{A}\}$ (cont.)

6. Given any $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ such that we can find some invertible $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ satisfying

$$\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_{nn} \end{bmatrix}$$

then

$$\begin{aligned} \exp\{t\boldsymbol{A}\} &= \sum_{k=0}^{\infty} \frac{t^k}{k!}(\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1})^k = \boldsymbol{V}\left[\sum_{k=0}^{\infty} \frac{t^k}{k!}\boldsymbol{\Lambda}^k\right]\boldsymbol{V}^{-1} \\ &= \boldsymbol{V}\exp\{t\boldsymbol{\Lambda}\}\boldsymbol{V}^{-1} \\ &= \boldsymbol{V}\begin{bmatrix} e^{\lambda_{11}t} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{\lambda_{nn}t} \end{bmatrix}\boldsymbol{V}^{-1} \end{aligned}$$

Diagonalizability of Square Matrices

Diagonalizability makes the computation of $\boldsymbol{A}^k$ and $\exp\{t\boldsymbol{A}\}$ easy!

But there are lots of questions:

1. Is every matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ diagonalizable? In other words, can we always find an invertible $\boldsymbol{V}$ such that $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}$ is a diagonal matrix?
2. What procedure can we use to diagonalize square matrices?
3. What should we do if there are some $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ that are not diagonalizable?

To answer these questions, we need to understand some linear algebra:

1. Subspaces
2. Nullspace and range
3. Rank
4. Matrix invertibility
5. Eigenvalues and eigenvectors

Sets and Subspaces

Let $\mathcal{A}$ and $\mathcal{B}$ be sets.

- $\mathcal{A} \subset \mathcal{B}$ means that all elements of the set $\mathcal{A}$ are also in the set $\mathcal{B}$.
- $x \in \mathcal{A}$ to mean that x is an element of the set $\mathcal{A}$.
- $\mathcal{A} \subset \mathcal{B}$ and $x \in \mathcal{A}$ implies that $x \in \mathcal{B}$.

Definition

$\mathcal{S} \subset \mathbb{R}^n$ is a **subspace** if and only if $\mathcal{S}$ is closed under addition and scalar multiplication, i.e.

$$\boldsymbol{x} \in \mathcal{S} \text{ and } \boldsymbol{y} \in \mathcal{S} \quad \Rightarrow \quad \boldsymbol{x} + \boldsymbol{y} \in \mathcal{S}$$

and

$$\boldsymbol{x} \in \mathcal{S} \text{ and } \alpha \in \mathbb{R} \quad \Rightarrow \quad \alpha \boldsymbol{x} \in \mathcal{S}.$$

Note that subspaces must always include the zero vector.

Spanning Set of a Subspace

Definition

A spanning set for the subspace $\mathcal{S} \subset \mathbb{R}^n$ is a set of vectors $\boldsymbol{s}_1, \ldots \boldsymbol{s}_p$, each in $\mathcal{S}$, such that every element of $\mathcal{S}$ can be expressed as a linear combination of the vectors $\boldsymbol{s}_1, \ldots \boldsymbol{s}_p$, i.e.

$$\boldsymbol{x} \in \mathcal{S} \quad \Rightarrow \quad \text{there exists } \alpha_1, \ldots, \alpha_p \text{ such that } \boldsymbol{x} = \alpha_1 \boldsymbol{s}_1 + \cdots + \alpha_p \boldsymbol{s}_p$$

where $\alpha_i \in \mathbb{R}$ for $i = 1, \ldots, p$.

Example: Suppose $\mathcal{S}$ is the xy plane in $\mathbb{R}^3$. Which of the following are spanning sets?

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \text{ or } \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \text{ or } \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ or } \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\} \text{ or } \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Some Facts About Subspaces, Spanning Sets, and Bases

1. Every $\mathcal{S} \subset \mathbb{R}^n$ possesses a linearly independent spanning set. Such a set is called a **basis** for $\mathcal{S}$. This basis is not unique, of course.
2. The number of vectors in any basis for $\mathcal{S}$ is the same. This number is called the **dimension** of $\mathcal{S}$. We use the notation $\dim(\mathcal{S})$ to denote the dimension of a subspace.
3. If $\mathcal{S}$ is a subspace of $\mathbb{R}^n$, then $\dim(\mathcal{S}) \leq n$ with equality if and only if $\mathcal{S} = \mathbb{R}^n$.
4. Any spanning set for $\mathcal{S}$ contains at least $\dim(\mathcal{S})$ vectors.
5. Any set with elements from $\mathcal{S}$ containing more than $\dim(\mathcal{S})$ vectors is linearly dependent.
6. A basis is a minimally-sized spanning set of $\mathcal{S}$.
7. A basis is a maximally-sized linear independent set of vectors in $\mathcal{S}$.

Nullspace and Range

Given $\boldsymbol{W} \in \mathbb{R}^{m \times n}$ (not necessarily square), there are two important subspaces related to this matrix.

Definition

The **nullspace** of $\boldsymbol{W}$ is defined as the set of all $\boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{0}$. We denote this subspace of $\mathbb{R}^n$ as $\mathrm{null}(\boldsymbol{W})$.

Definition

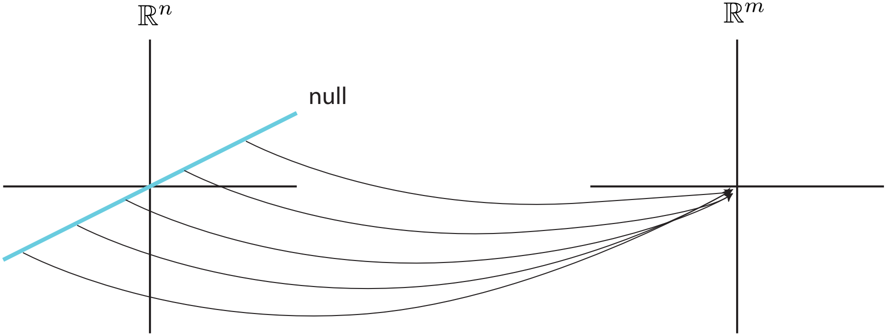
The **range** of $\boldsymbol{W}$ is defined as the set of all $\boldsymbol{y} \in \mathbb{R}^m$ such that there exists an $\boldsymbol{x}$ satisfying $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{y}$. We denote this subspace of $\mathbb{R}^m$ as $\mathrm{range}(\boldsymbol{W})$.

The range is also sometimes called the “column space” because it is the subspace generated by linear combinations of the columns of $\boldsymbol{W}$.

Note that both subspaces always include the zero vector.

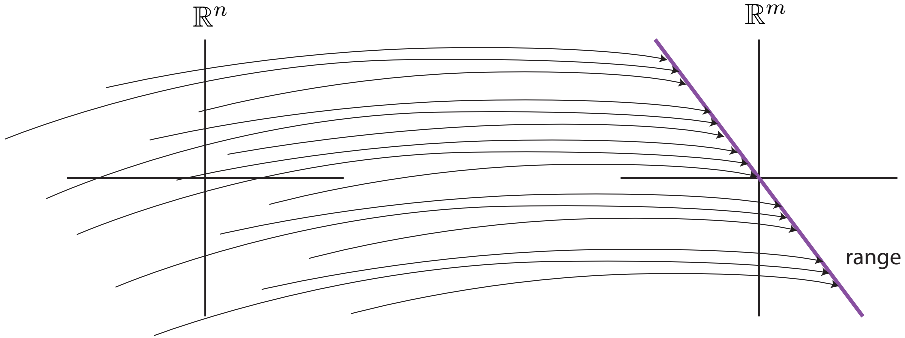
Nullspace

The matrix $\boldsymbol{W} \in \mathbb{R}^{m \times n}$ maps vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$. The nullspace of $\boldsymbol{W}$ is a subspace of $\mathbb{R}^n$.



Range

The matrix $\boldsymbol{W} \in \mathbb{R}^{m \times n}$ maps vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$. The range of $\boldsymbol{W}$ is a subspace of $\mathbb{R}^m$.



Nullspace and Range Examples

Suppose

$$\boldsymbol{W} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (1)$$

$$\boldsymbol{W} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (2)$$

$$\boldsymbol{W} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (3)$$

What is the nullspace and range of $\boldsymbol{W}$ in each case?

Existence and Uniqueness of Solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

For any $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$, a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ exists if and only if $\boldsymbol{b} \in \mathrm{range}(\boldsymbol{A})$.

For any $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$, a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is unique if and only if $\dim(\mathrm{null}(\boldsymbol{A})) = 0$.

Gaussian Elimination and Echelon Form

- GE is an algorithm for reducing a matrix to **echelon form**.
- Once you have a matrix in echelon form, you can easily determine its range and the dimension of its nullspace.
- This allows you to easily answer questions about the existence and uniqueness of solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

1. Form "augmented matrix" $\boldsymbol{U} = [\boldsymbol{A} \mid \boldsymbol{b}] \in \mathbb{R}^{m \times n+1}$.
2. Notation $\boldsymbol{U}(k,:)$ is the k^{th} row of $\boldsymbol{U}$ and $\boldsymbol{U}(k,j)$ is the k,j^{th} element of $\boldsymbol{U}$.
3. Force $\boldsymbol{U}(2,1) = 0$ by forming an appropriate combination of other rows and subtracting this combination from $\boldsymbol{U}(2,:)$.
4. Force $\boldsymbol{U}(3,1) = \boldsymbol{U}(3,2) = 0$ using the same technique.
5. Keep doing this until you have an upper triangular matrix.
6. You can now solve the last row since it has only one unknown.
7. Back substitute your answer and solve the second last row.
8. Keep doing this until you solve the top row.

Gaussian Elimination and Echelon Form Examples

Using the Echelon Form to Determine a Basis for $\mathrm{range}(\boldsymbol{A})$

The pivot columns of $\boldsymbol{A}$ for a basis for the range of $\boldsymbol{A}$. Note that the echelon form matrix tells you which columns to pick from $\boldsymbol{A}$. Example:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -1 & 3 & 0 \\ -1 & -2 & 2 & -2 & -1 \\ 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 2 & 2 & -1 \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 6 \\ 7 \end{bmatrix} \tag{4}$$

After reduction to echelon form of $\boldsymbol{U} = [\boldsymbol{A} \mid \boldsymbol{b}]$, we have

$$\begin{bmatrix} 1 & 2 & -1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{5}$$

The pivot columns here are the first, third, and fifth. Continued...

Using the Echelon Form to Determine a Basis for $\mathrm{range}(\boldsymbol{A})$

Hence a basis for the range of $\boldsymbol{A}$ is

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ -1 \end{bmatrix} \right\}$$

You should be able to verify that these vectors are linearly independent.

The range of $\boldsymbol{A}$ is the subspace formed by all linear combinations of these vectors. Solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ exist only when $\boldsymbol{b} \in \mathrm{range}(\boldsymbol{A})$.

Using the Echelon Form to Determine $\dim(\text{null}(\boldsymbol{A}))$

The dimension of the nullspace of $\boldsymbol{A}$ (also called the **nullity**) is simply the number of non-pivot columns in the echelon form.

You can use the echelon form to also find a basis for $\text{null}(\boldsymbol{A})$ (see any good linear algebra textbook for the details). In our example, a basis for the nullspace is

$$\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

You should be able to verify that these vectors are linearly independent and that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ if $\boldsymbol{x}$ is any linear combination of these basis vectors.

Most importantly, solutions to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ are unique only when $\dim(\text{null}(\boldsymbol{A})) = 0$.

Rank

Definition

The rank of $\boldsymbol{W}$ is defined as the dimension of the range of $\boldsymbol{W}$, i.e.

$$\mathrm{rank}(\boldsymbol{W}) \quad := \quad \dim(\mathrm{range}(\boldsymbol{W})).$$

Some useful facts:

- For $\boldsymbol{W} \in \mathbb{R}^{m \times n}$, $0 \leq \mathrm{rank}(\boldsymbol{W}) \leq \min\{m, n\}$.
- $\mathrm{rank}(\boldsymbol{W})$ is equal to the number of pivot columns in the echelon form of $\boldsymbol{W}$.
- Since $\dim(\mathrm{null}(\boldsymbol{W}))$ is equal to the number of non-pivot columns in the echelon form, rank + nullity must equal n.
- $0 \leq \mathrm{rank}(\boldsymbol{U}\boldsymbol{W}) \leq \min\{\mathrm{rank}(\boldsymbol{U}), \mathrm{rank}(\boldsymbol{W})\}$. In other words, matrix multiplication can only decrease rank.

Matrix Transpose

Definition

Given

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

the transpose of $\boldsymbol{W}$ is given as

$$\boldsymbol{W}^\top = \begin{bmatrix} w_{11} & w_{21} & \dots & w_{m1} \\ w_{12} & w_{22} & \dots & w_{m2} \\ \vdots & \vdots & & \vdots \\ w_{1n} & w_{2n} & \dots & w_{mn} \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

An Important Property of the Matrix Transpose

For any $\boldsymbol{A} \in \mathbb{R}^{m \times p}$ and $\boldsymbol{B} \in \mathbb{R}^{p \times n}$, the product $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ is an $m \times n$ real-valued matrix. The transpose of $\boldsymbol{C}$ is

$$\boldsymbol{C}^\top = (\boldsymbol{A}\boldsymbol{B})^\top = \boldsymbol{B}^\top \boldsymbol{A}^\top \in \mathbb{R}^{n \times m}$$

Note that the order of the matrix product has been changed by the transpose. Do the matrix dimensions agree?

Invertibility of Square Matrices

Definition

Given $\boldsymbol{W} \in \mathbb{R}^{n \times n}$, we say that $\boldsymbol{W}$ is invertible if there exists $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ such that $\boldsymbol{V}\boldsymbol{W} = \boldsymbol{W}\boldsymbol{V} = \boldsymbol{I}_n$. The quantity $\boldsymbol{V}$ is called the matrix inverse for $\boldsymbol{W}$ and we use the notation: $\boldsymbol{V} = \boldsymbol{W}^{-1}$.

The matrix inverse does not always exist, but when it does, it is unique.

Fact: If $\boldsymbol{W}$ is invertible, then $\boldsymbol{W}^\top$ is also invertible. To see this, just use what you know about the matrix inverse and the matrix transpose

$$\begin{aligned} \boldsymbol{W}\boldsymbol{W}^{-1} &= \boldsymbol{I}_n \\ (\boldsymbol{W}\boldsymbol{W}^{-1})^\top &= \boldsymbol{I}_n^\top \\ (\boldsymbol{W}^{-1})^\top \boldsymbol{W}^\top &= \boldsymbol{I}_n \end{aligned}$$

and, by the definition, $(\boldsymbol{W}^{-1})^\top = (\boldsymbol{W}^\top)^{-1}$.

Invertibility of Square Matrices: Equivalences

The following statements are equivalent:

1. $\boldsymbol{W}$ is invertible.
2. The only $\boldsymbol{x} \in \mathbb{R}^n$ satisfying $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{0}$ is $\boldsymbol{x} = 0$.
3. For every $\boldsymbol{b} \in \mathbb{R}^n$, there exists a unique $\boldsymbol{x} \in \mathbb{R}^n$ solving $\boldsymbol{W}\boldsymbol{x} = \boldsymbol{b}$.
4. The echelon form of $\boldsymbol{W}$ has no rows composed of all zeros.
5. $\det(\boldsymbol{W}) \neq 0$.
6. $\mathrm{rank}(\boldsymbol{W}) = n$.

Proofs...

Eigenvalues of Square Matrices

Definition

Given $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, λ_0 is an eigenvalue of $\boldsymbol{A}$ if and only if $(\boldsymbol{A} - \boldsymbol{I}_n \lambda_0)$ is not invertible.

Equivalently, based on what we already know about invertibility, we can say

$$\lambda_0 \text{ is an eigenvalue of } \boldsymbol{A} \Leftrightarrow \det(\boldsymbol{A} - \boldsymbol{I}_n \lambda_0) = 0$$

Definition

The characteristic polynomial of $\boldsymbol{A}$ is $\det(\lambda \boldsymbol{I}_n - \boldsymbol{A})$ where λ is a variable.

What is $\deg(\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}))$?

It is not too hard to show that the eigenvalues of $\boldsymbol{A}$ are equivalent to the roots of the characteristic polynomial of $\boldsymbol{A}$.

Some Consequences of What We Know About Eigenvalues

1. There can be at most n different eigenvalues of $\boldsymbol{A}$.
2. Even if $\boldsymbol{A}$ is real, its eigenvalues can be complex.
3. If $\boldsymbol{A}$ is real and λ_0 is a complex eigenvalue of $\boldsymbol{A}$, then λ_0^* is also an eigenvalue of $\boldsymbol{A}$ where the notation $()^*$ means complex conjugate, i.e.

$$z = a + jb \quad \Leftrightarrow \quad z^* = a - jb$$

4. If λ_0 is an eigenvalue of $\boldsymbol{A}$, then $\dim(\text{null}(\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n)) \geq 1$ since $\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n$ is not invertible.

This last consequence is of particular importance. Let $\boldsymbol{v}_0$ be any vector (except for the zero vector) in the nullspace of $\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n$. Then we can say that

$$(\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n)\boldsymbol{v}_0 = \boldsymbol{0}$$

which can be rewritten as $\boldsymbol{A}\boldsymbol{v}_0 = \lambda_0 \boldsymbol{v}_0$.

Eigenvectors of Square Matrices

Definition

Given $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{v}_0$ is an eigenvector of $\boldsymbol{A}$ corresponding to the eigenvalue λ_0 if and only if $\boldsymbol{A}\boldsymbol{v}_0 = \lambda_0 \boldsymbol{v}_0$.

Intuition: The matrix $\boldsymbol{A}$ scales its eigenvectors by its eigenvalues.

Note that there is nothing unique about an eigenvector. If $\boldsymbol{v}_0$ is an eigenvector of $\boldsymbol{A}$ corresponding to the eigenvalue λ_0, then so is $\alpha \boldsymbol{v}_0$ for any $\alpha \neq 0$ since

$$\boldsymbol{A}(\alpha \boldsymbol{v}_0) = \alpha(\boldsymbol{A}\boldsymbol{v}_0) = \alpha(\lambda_0 \boldsymbol{v}_0) = \lambda_0(\alpha \boldsymbol{v}_0)$$

Linear Independence of Eigenvectors

Fact: If $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_s \in \mathbb{R}^n$ are eigenvectors of $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ corresponding respectively with different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_s \in \mathbb{C}$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_s\}$ is a linearly independent set.

Intuitively: Eigenvectors corresponding to different eigenvalues are linearly independent.

Geometrically: Let $\mathcal{V}_j = \mathrm{null}(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)$. If $\lambda_j \neq \lambda_m$, then $\mathcal{V}_j \bigcap \mathcal{V}_m = \{\boldsymbol{0}\}$.

Important special case: If the matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ has n distinct eigenvalues, then there must exist n linearly independent eigenvectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\} \in \mathbb{C}^n$. Let

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \dots & \boldsymbol{v}_n \end{bmatrix}.$$

What can we say about the invertibility of $\boldsymbol{V}$?

When is $\boldsymbol{A}$ is Diagonalizable?

We now know that, when $\boldsymbol{A}$ has n distinct eigenvalues, $\boldsymbol{A}$ is diagonalizable and we can write

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$$

since $\boldsymbol{V} = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \dots \quad \boldsymbol{v}_n]$ is invertible in this case.

In this case, computing $\exp\{t\boldsymbol{A}\}$ and $\boldsymbol{A}^k$ is "easy". The main difficulty is finding the eigenvalues (finding the roots of a degree n polynomial).

What if $\boldsymbol{A}$ does not have n distinct eigenvalues? Does this mean that $\boldsymbol{A}$ is not diagonalizable?

Example

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

To determine the eigenvalues, we need to compute the roots of the characteristic polynomial, i.e. solve $\det(\lambda \boldsymbol{I}_3 - \boldsymbol{A}) = 0...$

Algebraic Multiplicity of an Eigenvalue

We can always write the characteristic polynomial of $\boldsymbol{A}$ in terms of its roots, i.e.

$$\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s}$$

where $\{\lambda_1, \dots, \lambda_s\}$ is the set of distinct eigenvalues of $\boldsymbol{A}$ with $1 \leq s \leq n$.

Definition

The algebraic multiplicity of the eigenvalue λ_j of the matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is the number of times the root λ_j appears in the characteristic polynomial of $\boldsymbol{A}$ and is denoted as r_j.

Eigenspace and Geometric Multiplicity of an Eigenvalue

Definition

Given $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, if $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $\boldsymbol{A}$, then the eigenspace corresponding with λ_0, denoted as $\mathcal{E}(\lambda_0)$, is the subspace of $\mathbb{C}^n$ spanned by the eigenvectors corresponding to the eigenvalue λ_0, i.e.

$$\mathcal{E}(\lambda_0) = \text{null}(\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n)$$

Definition

The geometric multiplicity of the eigenvalue λ_j of the matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is the dimension of the eigenspace of λ_j and is denoted as m_j, i.e.

$$m_j = \dim(\text{null}(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n))$$

Fact: For each $j \in \{1, \dots, s\}$, $1 \leq m_j \leq r_j$.

When is A Diagonalizable?

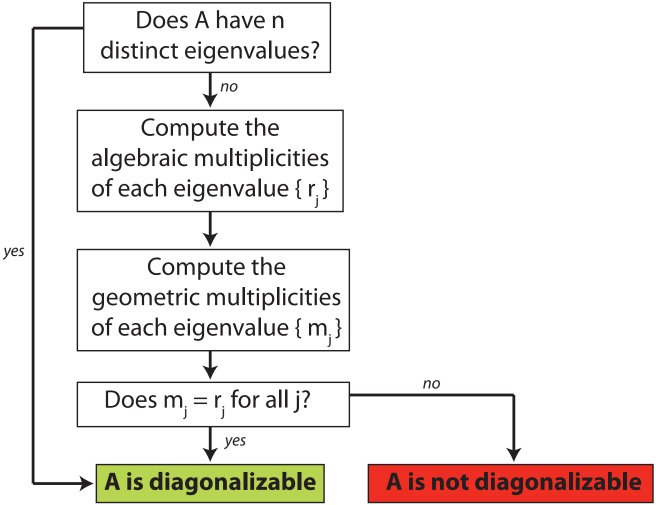
Theorem

If, for each $j \in \{1, \dots, s\}$, $m_j = r_j$, then A is diagonalizable.

This should be obvious when A has distinct eigenvalues since $m_j = r_j = 1$ for all j.

Proof sketch for the case when A does not have distinct eigenvalues:

A Procedure to Know When A is Diagonalizable



Summary of Diagonalization

1. Compute the eigenvalues of $\boldsymbol{A}$ and denote the distinct values as $\{\lambda_1, \dots, \lambda_s\}$.
2. If $\boldsymbol{A}$ is diagonalizable (see procedure on previous slide), then for each $j \in \{1, \dots, s\}$, find a basis for the eigenspace $\mathcal{E}(\lambda_j) = \mathrm{null}(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)$. You can do this with Gaussian elimination and echelon form. Let

$$B_j = \{\boldsymbol{v}_{j1}, \boldsymbol{v}_{j2}, \dots, \boldsymbol{v}_{jr_j}\}$$

   be a basis for $\mathcal{E}(\lambda_j)$.
3. Form $\boldsymbol{V}$ by stringing bases together. Note that $\boldsymbol{V}$ will be a square matrix since $\sum_{j=1}^{s} r_j = \sum_{j=1}^{s} m_j = n$.
4. Now $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$ (you should check it to be sure).

What Can We Do If $\boldsymbol{A}$ is Not Diagonalizable?

Some options:

1. It might be possible to just compute $\exp\{t\boldsymbol{A}\}$ by the definition, e.g.

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

2. You can use the fundamental matrix method to compute the CT-STM $\boldsymbol{\Phi}(t, s)$, and hence compute $\exp\{(t-s)\boldsymbol{A}\}$.
3. You can still use the eigenvalue/eigenvector method except you have to work with "generalized eigenvectors".

Generalized Eigenvectors

Definition

Given $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and $\lambda_0 \in \mathbb{C}$ an eigenvalue of $\boldsymbol{A}$, we say that $\boldsymbol{v}_0 \in \mathbb{C}^n$ is a generalized eigenvector corresponding with λ_0 if $\boldsymbol{v}_0 \neq \boldsymbol{0}$ and

$$(\boldsymbol{A} - \lambda_0 \boldsymbol{I}_n)^k \boldsymbol{v}_0 = \boldsymbol{0}$$

for some integer $k \geq 1$.

Question: Are all regular eigenvectors also generalized eigenvectors?

Question: Are all generalized eigenvectors also regular eigenvectors?

Generalized Eigenspace

Definition

The generalized eigenspace of the eigenvalue λ_0 is the subspace of $\boldsymbol{C}^n$ spanned by all of the generalized eigenvectors corresponding to λ_0.

We will use the notation $\mathcal{F}(\lambda_0)$ to denote the generalized eigenspace of the eigenvector λ_0.

Examples...

Some Basic Properties of Generalized Eigenspaces

1. $\mathcal{E}(\lambda_0) \subset \mathcal{F}(\lambda_0)$, i.e., the regular eigenspace of the eigenvalue λ_0 is a subset of the generalized eigenspace of the eigenvalue λ_0. Why?
2. If $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is a set of generalized eigenvectors corresponding to different eigenvalues, then $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is a linearly independent set.
3. If $\boldsymbol{v}_0 \in \mathcal{F}(\lambda_0)$, then $\boldsymbol{A}\boldsymbol{v}_0 \in \mathcal{F}(\lambda_0)$. In other words, the subspace $\mathcal{F}(\lambda_0)$ is invariant under $\boldsymbol{A}$.
4. Given the characteristic polynomial of $\boldsymbol{A}$

$$\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s}$$

where $\lambda_1, \dots, \lambda_s$ are all distinct and $r_1, \dots, r_s$ are the respective algebraic multiplicities, it can be shown that

$$\dim(\mathcal{F}(\lambda_j)) = r_j.$$

When combined with property #1, this implies that

$$1 \leq \dim(\mathcal{E}(\lambda_j)) \leq \dim(\mathcal{F}(\lambda_j)) = r_j.$$

Some Basic Properties of Generalized Eigenspaces (cont.)

5. A consequence of properties #2 and #4. If

$$\begin{array}{ccc} \{\boldsymbol{v}_{11},\dots,\boldsymbol{v}_{1r_1}\} & \text{is a basis for} & \mathcal{F}(\lambda_1) \\ \vdots & \vdots & \vdots \\ \{\boldsymbol{v}_{s1},\dots,\boldsymbol{v}_{sr_s}\} & \text{is a basis for} & \mathcal{F}(\lambda_s) \end{array}$$

then we can string all of these sets of generalized eigenvectors together into a big set $\{\boldsymbol{v}_{11},\dots,\boldsymbol{v}_{1r_1},\dots,\boldsymbol{v}_{s1},\dots,\boldsymbol{v}_{sr_s}\}$.

How many vectors are in this set? ______

This set of generalized eigenvectors is a basis for $\boldsymbol{C}^n$. Why?

Some Basic Properties of Generalized Eigenspaces (cont.)

6. From property #3, if $\boldsymbol{v}_{jk} \in \mathcal{F}(\lambda_j)$, then $\boldsymbol{A}\boldsymbol{v}_{jk} \in \mathcal{F}(\lambda_j)$.

   $\Leftrightarrow$ $\boldsymbol{A}\boldsymbol{v}_{jk}$ can be expressed as a linear combination of the vectors comprising a basis for $\mathcal{F}(\lambda_j)$.

   $\Leftrightarrow$ If the basis for $\mathcal{F}(\lambda_j)$ is $\{\boldsymbol{v}_{j1}, \dots, \boldsymbol{v}_{jr_j}\}$ then

$$\begin{aligned} \boldsymbol{A}\boldsymbol{v}_{j1} &= \alpha_{11}\boldsymbol{v}_{j1} + \cdots + \alpha_{1r_1}\boldsymbol{v}_{jr_j} \\ &\vdots \\ \boldsymbol{A}\boldsymbol{v}_{jr_j} &= \alpha_{j1}\boldsymbol{v}_{j1} + \cdots + \alpha_{jr_1}\boldsymbol{v}_{jr_j} \end{aligned}$$

   $\Leftrightarrow$ We can rewrite these r_j equations as one big matrix equation:

$$\boldsymbol{A}\underbrace{\begin{bmatrix}\boldsymbol{v}_{j1} & \dots & \boldsymbol{v}_{jr_j}\end{bmatrix}}_{\boldsymbol{V}_j} = \underbrace{\begin{bmatrix}\boldsymbol{v}_{j1} & \dots & \boldsymbol{v}_{jr_j}\end{bmatrix}}_{\boldsymbol{V}_j}\underbrace{\begin{bmatrix}\alpha_{11} & \alpha_{21} & \dots & \alpha_{r_j1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{r_j2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1r_j} & \alpha_{2r_j} & \dots & \alpha_{r_jr_j}\end{bmatrix}}_{\boldsymbol{Q}_j}$$

Some Basic Properties of Generalized Eigenspaces (cont.)

Property #6 continued...

We now have $\boldsymbol{A}\boldsymbol{V}_j = \boldsymbol{V}_j\boldsymbol{Q}_j$. What are the dimensions of $\boldsymbol{A}$, $\boldsymbol{V}_j$, and $\boldsymbol{Q}_j$?

Let $\boldsymbol{V} = [\boldsymbol{V}_1 \quad \boldsymbol{V}_2 \quad \dots \quad \boldsymbol{V}_s]$ and

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{Q}_1 & & \\ & \ddots & \\ & & \boldsymbol{Q}_s \end{bmatrix} \text{ (block diagonal form)}$$

What are the dimensions of $\boldsymbol{V}$ and $\boldsymbol{Q}$?

We now have $\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{Q}$. From property #5, what can we say about the invertibility of $\boldsymbol{V}$?

Hence, we can write $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{Q}\boldsymbol{V}^{-1}$. Note that $\boldsymbol{Q}$ is not diagonal, but block diagonal.

Some Basic Properties of Generalized Eigenspaces (cont.)

7. By the definition of generalized eigenvectors and generalized eigenspaces, the statement $\boldsymbol{v} \in \mathcal{F}(\lambda_j)$ is equivalent to

$$(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)^k \boldsymbol{v} = \boldsymbol{0} \qquad (6)$$

for some integer $k \geq 1$. Note that, if (6) is true when $k = k_0$, then it is also true for all $k \geq k_0$.

This implies that

$$\mathcal{F}(\lambda_j) \;=\; \mathrm{null}\left((\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)^{r_j}\right).$$

In other words, to determine the generalized eigenspace for the eigenvalue λ_j, you don't need to compute the nullspace of $(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)^k$ for $k = 1, 2, \dots, r_j$.

You can just compute the nullspace of the matrix $(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)^{r_j}$ by doing the standard Gaussian elimination and putting the result in echelon form.

Some Basic Properties of Generalized Eigenspaces (cont.)

8. From properties #6 and #7, we can say that

$$(\boldsymbol{Q}_j - \lambda_j \boldsymbol{I}_{r_j})^{r_j} = \boldsymbol{0}$$

Why?

Nilpotent Matrices

Definition

A nilpotent matrix is a square matrix $\hat{\boldsymbol{N}}$ with the property that $\hat{\boldsymbol{N}}^m = \boldsymbol{0}$ for some positive integer m.

An equivalent definition for a nilpotent matrix is a square matrix with eigenvalues all equal to zero.

Examples: Which of these matrices are

$$\boldsymbol{A}_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \boldsymbol{A}_2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad \boldsymbol{A}_3 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

nilpotent?

Some Basic Properties of Generalized Eigenspaces (cont.)

9. From property #8 we know that

$$\begin{bmatrix} \boldsymbol{Q}_1 - \lambda_1 \boldsymbol{I}_{r_1} & & \\ & \ddots & \\ & & \boldsymbol{Q}_s - \lambda_s \boldsymbol{I}_{r_s} \end{bmatrix} = \hat{\boldsymbol{N}}$$

is nilpotent. If we let

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 \boldsymbol{I}_{r_1} & & \\ & \ddots & \\ & & \lambda_s \boldsymbol{I}_{r_s} \end{bmatrix}$$

then we have $\boldsymbol{Q} = \boldsymbol{\Lambda} + \hat{\boldsymbol{N}}$ where Λ is diagonal and $\hat{\boldsymbol{N}}$ is nilpotent.

We can show that

$$\boldsymbol{\Lambda}\hat{\boldsymbol{N}} = \hat{\boldsymbol{N}}\boldsymbol{\Lambda}$$

In other words, $\hat{\boldsymbol{N}}$ and $\boldsymbol{\Lambda}$ commute.

Computing $\boldsymbol{A}^k$ When $\boldsymbol{A}$ is Not Diagonalizable

We now know everything we need to compute $\boldsymbol{A}^k$ when $\boldsymbol{A}$ is not diagonalizable. In general, we can always write

$$\boldsymbol{A} = \boldsymbol{V}(\boldsymbol{\Lambda} + \hat{\mathbf{N}})\boldsymbol{V}^{-1}$$

This implies that

$$\boldsymbol{A}^k = \boldsymbol{V}(\boldsymbol{\Lambda} + \hat{\mathbf{N}})^k\boldsymbol{V}^{-1}$$

By the binomial expansion theorem and property #9, we can write

$$\boldsymbol{A}^k = \boldsymbol{V}\left[\sum_{j=0}^{k}\binom{k}{j}\boldsymbol{\Lambda}^{k-j}\hat{\boldsymbol{N}}^j\right]\boldsymbol{V}^{-1}$$

But $\hat{\boldsymbol{N}}$ is nilpotent. Hence, for $j \geq \max\{r_1, \dots, r_s\}$, $\hat{\boldsymbol{N}}^j = 0$.

Computing $\exp\{t\boldsymbol{A}\}$ When $\boldsymbol{A}$ is Not Diagonalizable

By property #9, we can write

$$\begin{aligned}\exp\{t\boldsymbol{A}\} &= \boldsymbol{V}\exp\{t(\boldsymbol{\Lambda}+\hat{\boldsymbol{N}})\}\boldsymbol{V}^{-1}\\ &= \boldsymbol{V}\exp\{t\boldsymbol{\Lambda}\}\exp\{t\hat{\boldsymbol{N}}\}\boldsymbol{V}^{-1}\end{aligned}$$

The term $\exp\{t\boldsymbol{\Lambda}\}$ is easy to compute because $\boldsymbol{\Lambda}$ is diagonal.

What about the term $\exp\{t\hat{\boldsymbol{N}}\}$? Look at the definition:

$$\exp\{t\hat{\boldsymbol{N}}\} = \sum_{k=0}^{\infty}\frac{t^k}{k!}\hat{\boldsymbol{N}}^k.$$

But $\hat{\boldsymbol{N}}$ is nilpotent. So the sum will only have a finite number of terms:

$$\exp\{t\hat{\boldsymbol{N}}\} = \sum_{k=0}^{\max\{r_1,\dots,r_s\}-1}\frac{t^k}{k!}\hat{\boldsymbol{N}}^k.$$

In typical cases, there are only a few terms to compute.

Examples

Putting it All Together

A procedure for finding $\boldsymbol{A}^k$ and/or $\exp\{t\boldsymbol{A}\}$ for arbitrary $\boldsymbol{A} \in \mathbb{R}^{n\times n}$:

1. Find all of the eigenvalues of $\boldsymbol{A}$. Usually you do this by computing the roots of the characteristic polynomial, i.e.

$$\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) \quad = \quad (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_s)^{r_s} = 0$$

2. For each $j \in \{1, \dots, s\}$, find a basis for $\mathcal{E}(\lambda_j) = \mathrm{null}(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)$.
   - If $\dim(\mathcal{E}(\lambda_j)) = r_j$ then good! Move on to next eigenvalue.
   - If $\dim(\mathcal{E}(\lambda_j)) > r_j$ then you've done something wrong.
   - If $\dim(\mathcal{E}(\lambda_j)) < r_j$ then you need to find a basis for the generalized eigenspace $\mathcal{F}(\lambda_j) = \mathrm{null}(\boldsymbol{A} - \lambda_j \boldsymbol{I}_n)^{r_j}$. This basis must contain r_j linearly independent vectors.
3. Form $\boldsymbol{V} \in \mathbb{C}^{n\times n}$ by stringing together all of the bases.
4. Compute $\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V} = \boldsymbol{Q} = \boldsymbol{\Lambda} + \hat{\boldsymbol{N}}$ where $\boldsymbol{\Lambda}$ is diagonal and $\hat{\boldsymbol{N}}$ is nilpotent. Note that $\hat{\boldsymbol{N}} = \boldsymbol{0}$ when $\boldsymbol{A}$ is diagonalizable.

Putting it All Together (cont.)

5. Compute $\boldsymbol{A}^k$ via

$$\boldsymbol{A}^k \quad = \quad \boldsymbol{V} \left[\sum_{j=0}^{k} \binom{k}{j} \boldsymbol{\Lambda}^{k-j} \hat{\boldsymbol{N}}^j \right] \boldsymbol{V}^{-1}$$

   where the nilpotent property of $\hat{\boldsymbol{N}}$ implies that the sum will have at most $\max\{r_1, \dots, r_s\} - 1$ terms for any k.

6. Compute $\exp\{t\boldsymbol{A}\}$ via

$$\exp\{t\boldsymbol{A}\}\boldsymbol{V} \exp\{t\boldsymbol{\Lambda}\} \exp\{t\hat{\boldsymbol{N}}\}\boldsymbol{V}^{-1}$$

   where the term $\exp\{t\boldsymbol{\Lambda}\}$ is easy to compute because $\boldsymbol{\Lambda}$ is diagonal and the term

$$\exp\{t\hat{\boldsymbol{N}}\} \quad = \quad \sum_{k=0}^{\max\{r_1,\dots,r_s\}-1} \frac{t^k}{k!} \hat{\boldsymbol{N}}^k.$$

   is also not too difficult since the sum is finite.

Remarks

Note that $\exp\{t\mathbf{\Lambda}\}$ will have elements that look like $e^{\lambda_1 t}, e^{\lambda_2 t}$,.... What will the elements of $\exp\{t\hat{\boldsymbol{N}}\}$ look like?

Hence, when $\boldsymbol{A}$ is diagonalizable, $\exp\{t\boldsymbol{A}\}$ will only have terms that look like $e^{\lambda t}$. When $\boldsymbol{A}$ is not diagonalizable, $\exp\{t\boldsymbol{A}\}$ will also have terms that look like $t^m e^{\lambda t}$.

Conclusions

1. You now know how to solve LTV-CT, LTV-DT, LTI-CT, and LTI-DT systems with state-space representations.
2. State transition matrix (STM)
   2.1 Peano-Baker series
   2.2 Fundamental matrix method
3. STM computation for LTI systems: $\boldsymbol{A}^{k-j}$ or $\exp\{(t-s)\boldsymbol{A}\}$.
4. Linear algebraic tools:
   - Subspaces
   - Nullspace and range
   - Rank
   - Invertibility
   - Eigenvalues and eigenvectors
   - Diagonalizability
5. A procedure for computing $\boldsymbol{A}^{k-j}$ or $\exp\{(t-s)\boldsymbol{A}\}$ when $\boldsymbol{A}$ is or isn't diagonalizable.