17

MODELS AND ESTIMATION IN LINEAR SYSTEMS

17.1 Difference equation

17.1.1 Introduction

Let $z(n), n = 1, 2, ...$ be a description of a system state at a set of consecutive discrete time points $\{n\}$. An nth order linear difference equation is given by the following form:

$$z(k+n) + a_{n-1}(k)z(k+n-1) + ... + a_0(k)z(k) = g(k).$$

Clearly, the linear difference equation specifies a recursive relation between system state. If $g = 0$, then this equation is said to be **homogeneous**, otherwise it is said to be **non-homogeneous**.

A sequence $z(k)$ satisfies the linear difference equation is called a **solution**. Usually, there could be infinitely many solutions satisfying a homogeneous linear difference equation. For example, let $z_0(k)$ be a solution, then for any $\alpha > 0$, $\alpha z_0(k)$ is also a solution.

To further restrict the solutions (which is also required by practical applications), we often need to specify compatible initial conditions. For example, for an nth order homogeneous equation, we can specifies n initial conditions:

$$z(k_0 + i - 1) = c_i, \forall i = 1, 2, ..., n.$$

Example 17.1.1.

- $z(n) = bz(n-1)$ is a first order linear homogeneous difference equation. We can solve $z(n)$ via iteration:

$$z(n) = bz(n-1) = b(bz(n-2)) = b^2 z(n-2) = \cdots = b^n z(0).$$

- $z(n) = b_1 x(n-1) + b_2 x(n-2)$ is a second order homogeneous difference equation.
- $z(n) = az(n-1) + b(n)$ is a first order linear non-homogeneous difference equation, where $b(n)$ is a given sequence and $z(n)$ is unknown. $b(n)$ can take forms like $b(n) = 3n^3 + 1$.

Our objective of studying linear difference equations is to understand the properties of the solution and ultimately to be able to construct solutions. Notably, the existence and uniqueness of solutions are guaranteed under a rather mild condition, as showed by the following theorem.

Theorem 17.1.1 (existence and uniqueness of solutions). *[1, p. 19]Let a difference equation of the form*

$$z(k+n) + f[z(k+n-1), z(k+n-2), ..., z(k), k] = 0$$

where f is an arbitrary real-valued function, be defined over a finite or infinite sequence of consecutive values of k. The equation has one and only one solution corresponding to each arbitrary specification of the n initial values $z(k_0), z(k_0+1)..., z(k_0+n-1)$.

The theorem essentially says as long as there is no future dependence, we can iterate the equation to get the solution, even if f can be arbitrary(except for taking ∞ values).

Example 17.1.2 (Fibonacci sequence). The Fibonacci sequence is a second order linear difference equation given by

$$z(n) = z(n-1) + z(n-2), \quad z(0) = 0, z(1) = 1.$$

It can be showed that it has a solution given by

$$z(n) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

17.1.2 Solution structure of linear difference equations

The solutions to linear difference equations can be characterized and studied using tools of linear algebra.

First, we can show that solutions satisfy the following linearity relationship.

Lemma 17.1.1 (linearity of solutions to homogeneous equation). *If $z_1(k), z_2(k), ..., z_m(k)$ are solutions to the homogeneous linear difference equations (**without considering the initial conditions**), then any linear combination of these solutions is still a solution:*

$$z(k) = \sum_{i=1}^{m} c_i z_i(k), c_i \in \mathbb{R}.$$

We can view each solution as a finite-sized or infinite-sized vector, with each dimension/component associated with an integer time. This allows us to characterize the linear dependence of solutions and the linear structures of the vector space consisting of all the solutions.

Definition 17.1.1 (linear independence of solutions). *Given a set of solutions $z_1, z_2, ..., z_n$ defined for a set of consecutive integers, say $k = 1, 2, ..., N$, to the homogeneous linear difference equation, we say they are **linearly independent** if*

$$\sum_{i=1}^{n} c_i z_i(k) = 0, \forall k$$

only hold when all $c_i = 0$.

Theorem 17.1.2 (the vector space of solutions to homogeneous equation). *Given a nth order linear difference homogeneous equation defined on M consecutive integer, all solutions form a linear space of dimension n.*

Moreover, suppose $z_1(k), z_2(k), ..., z_n(k)$ is a linearly independent set of solutions to the homogeneous equations. Then any solution can be decomposed as

$$z(k) = \sum_{i=1}^{n} c_i z_i.$$

Proof. The linearity of solution has been proved. We can view each solution is the N component vector, and there are $M - n$ equations (each equation is the 1-step sliding version of the other) for each vector to satisfy. Therefore, the null space of the linear equation is n. □

Since we can view the solutions to homogeneous equation as a vector space, then we can study the basis of the vector space. Specifically, the standard basis in the solution space is called **the fundamental set of solutions**.

Definition 17.1.2 (fundamental set of solutions). *[1] Define $\bar{z}_i(k)$ to be the solution to nth order linear homogeneous difference equation with a **special** initial condition:*

$$z(k_0 + i - 1) = 1$$

and

$$z(k_0 + j - 1) = 0, j \in \{1, 2, .., n\}, j \neq i.$$

*Then this set of solutions $\bar{z}_1(k), \bar{z}_2(2), ..., \bar{z}_n(k)$ is called the **fundamental set of solutions**, which is also linearly independent set.*

Calculations of the fundamental set is straight forward: plugging in the special initial condition and iterating forward. Note that because coefficients $a_1, ..., a_n$ are time dependent, we usually cannot derive explicit expression When initial conditions are given, a solution can be immediately constructed from fundamental set. Particularly, we can use following method to construct the solution.

Methodology 17.1.1 (construct solution to arbitrary initial conditions). *Given the fundamental set of solutions to a homogeneous linear difference equation, then the solution to a homogeneous linear difference equation with initial conditions:*

$$z(k_0 + i - 1) = c_i, \forall i = 1, 2, ..., N$$

can be expressed as

$$z = \sum_{i=1}^{n} c_i \bar{z}_i.$$

17.1.3 Solution to non-homogeneous equation

Methodology 17.1.2. *Consider a non-homogeneous equation*

$$z(k+n) + a_{n-1}(k) z(k+n-1) + ... + a_0(k) z(k) = g(k),$$

the procedures to find a solution are:[1]

- *First find a set of n linearly independent solutions $z_1, z_2, ..., z_n$ to the homogeneous equation.*
- *Find a particular solution $\bar{y}$ that satisfies the non-homogeneous solution.*
- *Construct the unique solution:*

$$y = \bar{y} + \sum_{i=1}^{n} c_i z_i$$

where c_i are coefficients make y satisfy the initial conditions.

The coefficients can always be found, its existence can be obtained by setting z_i as the fundamental set, with the initial conditions given by $y - \bar{y}$. Methodology 17.1.1 gives the approach to find c_i.

17.1.4 Linear equations with constant coefficients

17.1.4.1 *Basic case*

Now we are addressing a special case of linear difference equations, whose coefficients are constants. A linear difference equation with **constant coefficients** is defined as

$$z(k+n) + a_{n-1}z(k+n-1) + ... + a_0 z(k) = 0. \tag{8}$$

With coefficients $a_0, ..., a_{n-1}$ being constants, we are able to derive analytical solutions. We first introduce the concept of characteristic equation.

The polynomial equation

$$\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_0 = 0$$

is called **characteristic equation** associated with the linear difference constant coefficients equation Equation 8.

It turns out that characteristic equation can be used to determine if a geometric sequence

$$z(k) = \lambda^k$$

to be a solution to Equation 8.

By plug in $z(k) = \lambda^k$ into Equation 8, we have

$$(\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_0)\lambda^k = 0,$$

which indicates $z(k) = \lambda^k$ is a solution if the constant λ satisfies the associated characteristic equation.

Theorem 17.1.3. *[1, p. 32] A necessary and sufficient condition for the geometric sequence $z(k) = \lambda^k$ to be a solution to the linear difference constant coefficients equation Equation 8 is that the constant λ satisfies the associated characteristic equation.*

Now a viable approach to Equation 8 to seek k different geometric sequences. They will form the basis of the solution space if they are linear independent to each other. Now we introduce the concept of **Casoratian determinant**, which can be used to determine if two sequences are linearly independent.

Definition 17.1.3 (Casoratian determinant). *[2, p. 149] The Casoratian determinant C of n sequence function $z_1(k), ..., z_n(k)$ is defined as*

$$C(k) = det(W)$$

where W is given as

$$W = \begin{pmatrix} z_1(k) & z_2(k) & \dots & z_n(k)) \\ z_1(k+1) & z_2(k+1) & \dots & z_n(k+n)) \\ \vdots & \vdots & & \vdots \\ z_1(k+n) & z_2(k+n) & \dots & z_n(k+n) \end{pmatrix}$$

Lemma 17.1.2 (linear independence criterion). *[2, p. 149] Consider n sequence function $z_1(k), ..., z_n(k)$ defined on consecutive integers $K_1 \leq k \leq K_2$. Then these n sequence functions are linearly independent if and only*

$$C(k) \neq 0, \forall K_1 \leq k \leq K_2 - n.$$

It can be showed that geometric sequences $z(k) = \lambda^k$ with different λ are linearly independent.

Lemma 17.1.3 (linear dependence of geometric sequence). $z_1(k) = \lambda_1^k$ *and* $z_2(k) = \lambda_2^k$ *are linearly independent.*

From fundamental theorem of algebra [Theorem 4.18.3], the nth degree characteristic equation has n roots. Assume these n roots are distinct[1], then the geometric sequences the geometric sequence $\lambda_1^k, \lambda_2^k, ..., \lambda_n^k$ form a linearly independent set. A solution can be constructed via

$$z = c_1\lambda_1^k + c_2\lambda_2^k + \cdots + c_n\lambda_n^k,$$

where $c_1, c_2, ..., c_n$ can be determined by matching initial conditions.

We summarize our approach into the following theorem.

Theorem 17.1.4 (solutions to linear difference equation with constant coefficients). *Let $\lambda_1, \lambda_2, ..., \lambda_n$ be the n distinct eigenvalues to the characteristic equation. It following that*

- *the geometric sequence $\lambda_1^k, \lambda_2^k, ..., \lambda_n^k$ form a linearly independent set*
- *a solution can be constructed via*

$$z = c_1\lambda_1^k + c_2\lambda_2^k + \cdots + c_n\lambda_n^k,$$

1 the case where roots have multiplicity greater than 1 is addressed in the next section.

where $c_1, c_2, ..., c_n$ can be determined by matching initial conditions.

Example 17.1.3 (Fibonacci sequence). [1, p. 35] Consider the difference equation

$$z(k+2) = z(k+1) + z(k)$$

with initial condition $z(1) = z(2) = 1$. The characteristic equation is

$$\lambda^2 - \lambda - 1 = 0$$

we have the solution of $\lambda_1 = 1 + \sqrt{5}/2, \lambda_2 = 1 - \sqrt{5}/2$.
Then we assume the solution will take the form

$$z(k) = A\lambda_1^k + B\lambda_2^k$$

From initial condition we can get

$$z(k) = (\lambda_1^k - \lambda_2^k)\frac{1}{\sqrt{5}}.$$

Example 17.1.4 (roots can be imaginary). Consider the second-order equation

$$z(k+2) + 4z(k) = 0$$

with initial conditions $z(0) = 2, z(1) = -4$. The characteristic equation is

$$\lambda^2 + 4 = 0$$

which has the roots $\lambda = \pm 2i$ (where $i = \sqrt{-1}$). The general solution is given by

$$z(k) = c_1(2i)^k + c_2(-2i)^k.$$

Substitution of the given initial conditions yields the equations

$$c_1 + c_2 = 2$$
$$c_1(i) + c_2(-i) = -4$$

We can solve Thus, $c_1 = 1 + i, c_2 = 1 - i$.

Example 17.1.5 (solving non-homogeneous equation). Now we consider a second order non-homogeneous equation given by

$$z(k+2) - 5z(k+1) + 6z(k) = 5.$$

The roots to the characteristic equation are $\lambda = 2, 3$. So the general solution of the homogeneous equation is

$$z(k) = c_1(3)^k + c_2(2)^k.$$

For a particular equation $z(k+2) - 5z(k+1) + 6z(k-2) = 5$, we can try $z(k) = D \in \mathbb{R}$ and find

$$D - 5D + 6D = 5 \implies D = 5/2.$$

Taken together, the general solution is

$$z(k) = \frac{5}{2} + c_1(3)^k + c_2(2)^k,$$

where the coefficients c_1 and c_2 will be determined by additional initial conditions.

Example 17.1.6 (solving non-homogeneous equation). We consider the difference equation

$$z(k+3) - 7z(k+2) + 14z(k+1) - 8z(k) = 2^k.$$

The roots to the characteristic equation are 1, 2, and 4. So we have the general solution to the homogeneous system solutions

$$z_1(k) = 1, \quad z_2(k) = 2^k, \quad z_3(k) = 4^k.$$

Now we need to find a particular solution, and we can try $z(k) = c2^k$. Plug in and we get

$$c2^{k+3} - 7c2^{k+2} + 14c2^{k+1} - 8c2^k - 2^k.$$

And we can solve c to get the particular solution.

17.1.4.2 *General case*

Now we consider the general case that roots from the characteristic equations have multiplicity greater than 1, which meaning that we have less than n distinct roots. We first introduce some tools needed to develop the approach.

Definition 17.1.4 (forward operator). *We define forward operator E on the function $z(k)$, which is defined on consecutive integers as*

$$Ez(k) = z(k+1)$$

and $E^0 = I$.

Definition 17.1.5 (characteristic polynomial of forward operators). *The linear difference equation with constant coefficients can be rewritten as $\Phi(E)z(k) = (E^n + a_{n-1}E^{n-1} + ... + E^0)z(k) = 0$ where $\Phi(E) = E^n + a_{n-1}E^{n-1} + ... + E^0$ is known as* ***characteristic polynomial of forward operators****.*

An important identity associated with the characteristic polynomial of forward operators is the **key identity**.

Lemma 17.1.4 (key identity).

$$\Phi(E)x^k = x^k\Phi(x).$$

Proof. $\Phi(E)x^k = (x^{n+k} + a_{n-1}x^{n+k-1} + ... + a_0x^k) = x^k\Phi(x)$. □

The key identity allows us to derive linearly independent solutions corresponding to one root with multiplicity greater than 1.

Theorem 17.1.5. *If the polynomial $\Phi(x)$ has a repeated root λ with multiplicity of m. Then for functions defined as*

$$\begin{aligned} z(k) &= \lambda^k, \\ z(k) &= k\lambda^k, \\ z(k) &= k(k-1)\lambda^k, \\ &\cdots \\ z(k) &= k(k-1)\cdots(k-m_i+1)\lambda^k \end{aligned}$$

all satisfies

$$\Phi(E)z(k) = 0.$$

That is, they are solutions to the homogeneous linear equation [Equation 8].

Proof. Using above key identity we have:

$$\begin{aligned} \frac{d}{dx}\Phi(E)x^k &= \frac{d}{dx}x^k\Phi(x) \\ \Phi(E)kx^{k-1} &= kx^{k-1}\Phi(x) + x^k\Phi'(x) \\ &= 0 \end{aligned}$$

Note we use the fact that $\Phi(x) = (x-\lambda)^{m_i}q(x)$ such that $\Phi(\lambda) = 0, \Phi'(\lambda) = 0$. On the other hand, we can show that $\Phi(E)kx^k = \frac{d}{dx}\Phi(E)x^k$, i.e,

$$\Phi(E)k\lambda^k = 0 \implies k\lambda^k \text{ is a solution.}$$

We can generalize this procedure to roots with other multiplicities. □

These roots are linearly independent from each other. And similar to Theorem 17.1.4, we can construct solution to homogeneous and non-homogeneous equations.

Example 17.1.7. Now we consider a second order non-homogeneous equation given by

$$z(k+2) - 2z(k+1) + z(k) = 0.$$

The roots to the characteristic equation are $\lambda = 1$ with multiplicity of 2. So the general solution of the homogeneous equation is

$$z(k) = c_1(1)^k + c_2k(1)^k = c_1 + c_2k.$$

17.2 Differential equations

17.2.1 Linear differential equations

17.2.1.1 *Concepts*

Definition 17.2.1 (order n linear differential equation). *An order n linear differential equation is given by*

$$\frac{\partial^n y}{\partial t^n} + a_{n-1}(t)\frac{\partial^{n-1} y}{\partial t^{n-1}} + ... + a_0(t)y = g(t)$$

If $g(t) = 0$, then it is called homogeneous; otherwise it is called non-homogeneous.

Remark 17.2.1 (caution!). Note that the coefficients $a_0, a_1, ..., a_{n-1}$ can be a constant or a function of t; however, it cannot be a function of y, which is call **nonlinear differential equation**.

Example 17.2.1.

- The ODE
$$y'' - y' - 2y = 0$$
is an second order linear homogeneous differential equation.
- The ODE
$$y'' - y' - 2y = \exp(2t)$$
is an second order linear non-homogeneous differential equation.
- The ODE
$$(1 + t^2)y'' - 2ty' + 2y = 0$$
is an second order linear homogeneous differential equation.

Theorem 17.2.1 (linearity of solutions). *If $z_1(t), z_2(t), ...z_m(t)$ are all solutions of a linear homogeneous differential equation, then the linear combination*

$$z(t) = \sum_{i=1}^{m} c_i z_i(t), c_i \in \mathbb{C},$$

is also the solution.

Note that the coefficients can be complex constants.

Proof. direct plug in to verify. □

17.2.1.2 *Wronskian and linear independence*

Definition 17.2.2 (linear independence of solutions). *[1, p. 40] Given a set of solutions $y_1, y_2, ..., y_n$ on the interval $[t_0, t_1]$ to the homogeneous linear differential equation, we say they are linearly independent if*

$$\sum_{i=1}^{n} c_i y_i(t) = 0, \forall t \in [t_0, t_1]$$

only hold when all $c_i = 0$.

Remark 17.2.2. To check linear independence, we need to check the condition holds for all $t \in [t_0, t_1]$, not just at a single t.

Definition 17.2.3 (Wronskian of a set of solutions). *Given n solutions $z_1(t), z_2(t), ..., z_n(t)$, the* ***Wronskian*** *associated with the n solutions is defined to be a determinant of the fundmental matrix*

$$W[z_1, z_2, ..., z_n](t) = det\ Z = det \begin{pmatrix} z_1(t) & z_2(t) & \cdots & z_n(t) \\ z_1'(t) & z_2'(t) & \cdots & z_n'(t) \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{(n-1)}(t) & z_2^{(n-1)}(t) & \cdots & z_n^{(n-1)}(t) \end{pmatrix}$$

where $Z_{ij} = z_j^{i-1}$(the $i-1$ derivative of jth solution).

Note 17.2.1 (motivation of Wronskians and the initial value problem). Suppose we have n different solutions $z_1, z_2, ..., z_n$ to the order n linear homogeneous ODE on the interval $[t_L, t_R]$. We want the solution $z(t) = \sum_{i=1}^{n} c_i z_i(t)$ to satisfy the initial value conditions given by $z(t_I) = b_0, z'(t_I) = b_1, ..., z^{(n-1)}(t_I) = b_{n-1}, t_I \in [t_L, t_R]$. To solve for the coefficients $c_1, c_2, ..., c_n$, we have

$$\begin{aligned} b_0 &= c_1 z_1(t_I) + c_2 z_2(t_I) + \cdots + c_n z_n(t_I) \\ b_1 &= c_1 z_1'(t_I) + c_2 z_2'(t_I) + \cdots + c_n z_n'(t_I) \\ \vdots &= \vdots \\ b_{n-1} &= c_1 z_1^{(n-1)}(t_I) + c_2 z_2^{(n-1)}(t_I) + \cdots + c_n z_n^{(n-1)}(t_I). \end{aligned}$$

$c_1, c_2, ..., c_n$ can be uniquely solved if and only if

$$det \begin{pmatrix} z_1(t_I) & z_2(t_I) & \cdots & z_n(t_I) \\ z_1'(t_I) & z_2'(t_I) & \cdots & z_n'(t_I) \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{(n-1)}(t_I) & z_2^{(n-1)}(t_I) & \cdots & z_n^{(n-1)}(t_I) \end{pmatrix} \neq 0.$$

Theorem 17.2.2 (Abel's Wronskian theorem and its properties). *[3, p. 32]*

- *If $z_1, z_2, ..., z_n$ are solutions to the order n linear homogeneous ODE, then the Wronskian satisfies the following first order linear system:*

$$\frac{d}{dt} W[z_1, z_2, ..., z_n](t) + a_{n-1}(t) W[z_1, z_2, ..., z_n](t) = 0.$$

- *The Wronskian can be solved to get*

$$W(t) = W(t_0) \exp[-\int_{t_0}^{t} a_{n-1}(s) ds].$$

- *The Wronskian is either always zero or never zero.*

Proof. (1) see reference. (2) solve ODE in (1). (3)Since $\exp[\int_{t_0}^{t} a_{n-1}(s) ds] \neq 0$, the Wronskian has the property of **once zero at a point, all zero; once nonzero at a point, all nonzero**. □

Example 17.2.2. $z_1(t) = \exp(2t)$ and $z_2(t) = \exp(-t)$ are solutions of

$$z'' - z' - 2z = 0.$$

Their Wronskian is given by

$$W[z_1, z_2](t) = det \begin{pmatrix} z_1(t) & z_2(t) \\ z_1'(t) & z_2'(t) \end{pmatrix} = det \begin{pmatrix} \exp(2t) & \exp(-t) \\ 2\exp(2t) & -\exp(-t) \end{pmatrix} = -3\exp(t) \neq 0.$$

Theorem 17.2.3 (linear independence and nonzero Wronskian). *Let $z_1, z_2, ..., z_n$ be the solutions of a linear homogeneous ODE of order n within the interval (t_L, t_R). Then the following properties are equivalent.*

- *the Wronskian $W[z_1, z_2, ..., z_n]$ is nonzero everywhere in (t_L, t_R).*
- *the Wronskian $W[z_1, z_2, ..., z_n]$ is nonzero somewhere in (t_L, t_R).*
- *$z_1, z_2, ..., z_n$ are linearly independent.*

Proof. (1) is equivalent to (2): directly from Theorem 17.2.2. (1) implies (3): suppose at $t_I \in (t_L, t_R)$, $W[z_1, z_2, ..., z_n](t_I) \neq 0$. This implies that the following linear algebraic system

$$\begin{aligned} 0 &= c_1 z_1(t_I) + c_2 z_2(t_I) + \cdots + c_n z_n(t_I) \\ 0 &= c_1 z_1'(t_I) + c_2 z_2'(t_I) + \cdots + c_n z_n'(t_I) \\ \vdots &= \vdots \\ 0 &= c_1 z_1^{(n-1)}(t_I) + c_2 z_2^{(n-1)}(t_I) + \cdots + c_n z_n^{(n-1)}(t_I). \end{aligned}$$

have solution $c_1 = c_2 = ... = c_n = 0$. That is, the equation

$$\sum_{i=1}^{n} c_i z_i(t) = 0, \forall t \in [t_L, t_R]$$

only hold when all $c_i = 0$. (3)implies (1): suppose $z_1, z_2, ..., z_n$ are linearly independent but $W[z_1, z_2, ..., z_n](t_I) = 0$ for some $t_I \in (t_L, t_R)$. This implies that that the following linear algebraic system

$$\begin{aligned} 0 &= c_1 z_1(t_I) + c_2 z_2(t_I) + \cdots + c_n z_n(t_I) \\ 0 &= c_1 z_1'(t_I) + c_2 z_2'(t_I) + \cdots + c_n z_n'(t_I) \\ \vdots &= \vdots \\ 0 &= c_1 z_1^{(n-1)}(t_I) + c_2 z_2^{(n-1)}(t_I) + \cdots + c_n z_n^{(n-1)}(t_I). \end{aligned}$$

have a nonzero solution $c_1, c_2, ..., c_n$ such that

$$z(t) = \sum_{i=1}^{n} c_i z_i(t).$$

The algebraic system also implies that $z(t)$ satisfies the following initial conditions

$$z(t_I) = 0, z'(t_I) = 0, ..., z^{(n-1)}(t_I) = 0.$$

In other words, $z(t)$ is a constant value function of zero. Or equivalently,

$$z(t) = 0 = \sum_{i=1}^{n} c_i z_i(t).$$

This contradicts the fact that $z_1, z_2, ..., z_n$ are linear independent. □

Corollary 17.2.3.1 (linear independence and nonzero Wronskian). *Let $z_1, z_2, ..., z_n$ be the solutions of a linear homogeneous ODE of order n within the interval (t_L, t_R). Then the following properties are equivalent.*

- *the Wronskian $W[z_1, z_2, ..., z_n]$ is zero everywhere in (t_L, t_R).*
- *the Wronskian $W[z_1, z_2, ..., z_n]$ is zero somewhere in (t_L, t_R).*
- *$z_1, z_2, ..., z_n$ are linearly dependent.*

Remark 17.2.3 (Wronskian and linear independence of functions might not hold for general functions

Only when the functions are solutions of the order n linear homogeneous ODE, Theorem 17.2.3 holds.

- For general functions, above theorem might not hold. For example $Y_1(t) = t^2, Y_2(t) = |t|\, t, t \in (-\infty, \infty)$. We can show that

$$\begin{aligned} W[Y_1, Y_2](t) &= det \begin{pmatrix} Y_1(t) & Y_2(t) \\ Y_1'(t) & Y_2'(t) \end{pmatrix} \\ &= det \begin{pmatrix} t^2 & |t|\, t \\ 2t & 2|t| \end{pmatrix} \\ &= 0 \end{aligned}$$

However, it is clearly that $Y_1(t)$ and $Y_2(t)$ are not proportional to each other;i.e., they are linearly independent.

17.2.1.3 *General solution theory*

Definition 17.2.4 (fundamental systems). *[3, p. 30] A collection of n linearly independent solutions $z_1(t), z_2(t), ..., z_n(t)$ is called **fundamental system.***

Theorem 17.2.4 (construct arbitrary solution from fundamental set). *Suppose $z_1(t), z_2(t), ..., z_n(t)$ is a linearly independent set of solutions to the linear homogeneous differential equations of order n, then **any** solution can be decomposed as the linear combination of the*

$$F(t) = [z_1(t)\ z_2(t)\ ...\ z_n(t)].$$

Proof. Consider an arbitrary solution denoted by $Y(t)$. We can generate n initial conditions at t_I by differentiation; that is

$$Z(t_I) = Y(t_I), Z'(t_I) = Y'(t_I), ..., Z^{(n-1)}(t_I) = Y^{(n-1)}(t_I).$$

To show that $Y(t)$ can be actually decomposed via

$$Y(t) = \sum_{i=1}^{n} c_i z_i(t),$$

we want to solve for c from the linear system

$$\begin{aligned}
Y(t_I) &= c_1 z_1(t_I) + c_2 z_2(t_I) + \cdots + c_n z_n(t_I) \\
Y'(t_I) &= c_1 z_1'(t_I) + c_2 z_2'(t_I) + \cdots + c_n z_n'(t_I) \\
\vdots &= \vdots \\
Y(t_I)^{(n-1)} &= c_1 z_1^{(n-1)}(t_I) + c_2 z_2^{(n-1)}(t_I) + \cdots + c_n z_n^{(n-1)}(t_I).
\end{aligned}$$

Because **linear independence implies nonzero Wronskian**(Theorem 17.2.3), then $c_1, c_2, ..., c_n$ can be uniquely solved. □

Remark 17.2.4 (implication for initial value problem). The solution to an initial value problem is just one solution satisfies both the ODE and the initial conditions; therefore it can be constructed via linear combination of the fundamental set.

Theorem 17.2.5 (necessary and sufficient conditions for fundamental system and the solution method). *The necessary and sufficient conditions for $z_1, z_2, ..., z_n$ to be a fundamental system is that there exists a $t_0 \in (r_1, r_2)$, such that $W(t_0) \neq 0$. Note that (t_L, t_R) is the interval where the conditions for existence and uniqueness of solutions are satisfied.*

Proof. (a) forward: use the property of Wronskian that **nonzero Wronskian implies linear independence**.(b) converse: a fundamental system is linearly independent and the Wronskian will be nonzero. □

Theorem 17.2.6 (natural fundamental set of solution). *Let $N_i(t), i = 0, 1, ..., n-1$ be a set of solutions to the linear homogeneous differential equations with special initial conditions given as*

$$\frac{\partial^n y}{\partial t^n} = \begin{cases} 1, n = i-1 \\ 0, otherwise \end{cases}.$$

This set of solutions are called ***natural fundamental set of solutions****, and the solution to linear homogeneous differential equations with arbitrary initial conditions can be constructed via*

$$z = \sum_{i=1}^{n} c_i N_i(t)$$

with $c_i = b_i$.
Moreover, ***the natural fundamental set are linearly independent.***

Proof. directly use the existence and uniqueness theorem plus the vector space nature of the solutions. We can calculate the Wronskian for the natural fundamental set of solutions given by

$$\begin{aligned} W[N_1, N_2, ..., N_n](t_I) &= det \begin{pmatrix} N_1(t_I) & N_2(t_I) & \cdots & N_n(t_I) \\ N_1'(t_I) & N_2'(t_I) & \cdots & N_n'(t_I) \\ \vdots & \vdots & \ddots & \vdots \\ N_1^{(n-1)}(t_I) & N_2^{(n-1)}(t_I) & \cdots & N_n^{(n-1)}(t_I) \end{pmatrix} \\ &= det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = 1 \neq 0. \end{aligned}$$

therefore, the natural fundamental set is linearly independent. □

Example 17.2.3. Consider the ODE

$$y'' - y' - 2y = 0,$$

with initial conditions $y(0) = b_0, y'(0) = b_1$.

The natural fundamental set of solutions are

$$N_0(t) = \frac{\exp(2t) + 2\exp(-t)}{3}, N_1(t) = \frac{\exp(2t) - \exp(-t)}{3}.$$

It can be verify that $N_0(t)$ solves

$$y'' - y' - 2y = 0,$$

with initial conditions $y(0) = b_0, y'(0) = 0$.

And $N_1(t)$ solves

$$y'' - y' - 2y = 0,$$

with initial conditions $y(0) = 0, y'(0) = b_1$.

17.2.1.4 *Existence & uniqueness of solution*

Theorem 17.2.7 (local existence and uniqueness). *[4, p. 91] Let $f(t,x)$ be a piece-wise continuous function and satisfies the Lipschitz condition:*

$$\left|f(t,x) - f(t,y)\right| \leq L\left|x - y\right|$$

$\forall x, y \in B = \{x \in \mathbb{R}^n | \|x - y\| \leq b\}, \forall t \in [t_0, t_1]$. *Then there exist some $\delta > 0$ such that the equation $\dot{x} = f(t,x)$ with $x(0) = x_0$ has a unique solution over $[t_0, t_0 + \delta]$.*

Example 17.2.4.

- the ODE $\dot{x} = x^{1/3}$ with $x(0) = 0$ has non-unique solutions: $x(t) = 0$ and $x(t) = (2t/3)^{3/2}$. It can be verified that at $x = 0$ it does not satisfy above Lipschitz condition.($f(x) = x^{1/3}, f' = 1/3(x^{-2/3})$, therefore $f(x)$ has unbounded derivative at $x = 0$, and therefore it is not locally Lipschitz, since the necessary condition is bounded derivatives.)

Example 17.2.5. Consider the differential equation

$$\frac{dy}{dt} = \frac{4y}{t}, y(1) = 1.$$

Since t cannot equal 0, we first consider the situation of $t \in (0, +\infty)$. By assuming $y \neq 0$ for $t \in (0, +\infty)$, we have

$$\frac{dy}{y} = \frac{4dt}{t} \implies \ln y = 4 \ln t + C \implies y = ke^{4 \ln t} = kt^4.$$

The initial condition implies $k = 1$, also $y \neq 0 \forall t > 0$.

For $t < 0$, we have

$$\frac{dy}{y} = \frac{4d - t}{-t} \implies \ln y = 4 \ln(-t) + C \implies y = ke^{4 \ln(-t)} = k(-t)^4.$$

However, no initial condition can be used to determine k when $t < 0$. That is, when $t < 0$, there is no unique solution.

Note that for any k, y is continuous at $t = 0$, but may take different form for $t \in (-\infty, 0)$. This nonuniqueness is due to the fact that at$t = 0$, dy/dt not well defined.

Theorem 17.2.8 (alternative, for linear differential equations). *[1, p. 40] Given a linear differential equation*

$$\frac{\partial^n y}{\partial t^n} + a_{n-1}(t) \frac{\partial^{n-1} y}{\partial t^{n-1}} + ... + a_0(t) y = g(t)$$

satisfying the initial conditions

$$\frac{\partial^i y}{\partial t^i} = b_i, \forall i = 0, 1, ..., n-1$$

If all a_is and $g(t)$ are continuous on an interval $0 \leq t \leq T$, then there exists an unique solution.

17.2.2 Linear homogeneous differential equations with constant coefficients

17.2.2.1 *The key identity*

Definition 17.2.5 (differential operators and polynomials).

- *The* ***differential operator*** $D : C^\infty \to C^\infty$ *is defined as*

$$D[\phi] = \phi'$$

where $\phi \in C^\infty$.
- *Let*

$$L(D) = a_0 D^n + a_1 D^{n-1} + ... + a_n$$

with polynomial L *defined as*

$$L(x) = a_0 x^n + a_1 x^{n-1} + ... + a_n$$

then $L(D)$ *is called a* ***differential operator polynomial.***

Definition 17.2.6 (characteristic polynomial). *Consider a linear homogeneous differential equation with constant coefficients given by*

Theorem 17.2.9 (key identity for linear homogeneous ODE). *Consider a differential operator polynomial given by*

$$L(D) = a_0 D^n + a_1 D^{n-1} + ... + a_n.$$

Then we have the ***key identity***

$$L(D)e^{rx} = L(r)e^{rx},$$

where

$$L(r) = (a_0 r^n + a_1 r^{n-1} + ... + a_n)$$

is known as the characteristic polynomial.

Proof. Direct differentiate e^{rx} and can verify the correctness. □

Theorem 17.2.10 (characteristic solution to linear homogeneous differential equation). *Let* $\lambda \in \mathbb{C}$ *be a root to the characteristic polynomial* $L(x) = 0$, *then*

$$L(D)e^{\lambda x} = 0.$$

That is, $e^{\lambda x}$ *is one solution to the differential equation* $L(D)e^{rx} = 0$. *Particularly if* $\lambda = 0$, *then constant function* 1 *is the solution.*

Proof. $L(D)e^{\lambda x} = L(\lambda)e^{\lambda x} = 0$ from above lemma. □

17.2.2.2 *The case of real roots*

Theorem 17.2.11 (solutions associated with characteristic polynomial roots with multiplicity). *[3, p. 49] Let λ be a root to polynomial $L(x) = 0$ with multiplicity m, then the function*

$$x^r e^{\lambda x}, r \in \{0, 1, 2, ..., m-1\}$$

satisfies $L(D)x^r e^{\lambda x} = 0$. That is, $x^r e^{\lambda x}$ is the solution to the differential equation $L(D)f(x) = 0$

Proof. Consider the key identity(Theorem 17.2.9)

$$L(D)e^{sx} = e^{sx}L(s)$$

(where D is the differential operator with respect to x). Take the derivative with respect to s we have

$$\begin{aligned} \frac{d}{ds}L(D)e^{sx} &= \frac{d}{ds}e^{sx}L(s) \\ sL(D)e^{st} &= e^{st}(sL(s) + dL(s)/ds) \end{aligned}$$

where we have used the fact that exchange the partial differential operator is legitimate. We can see then if $s = \lambda$ is root to $L(x)$ with multiplicity of 2, then we can write $L(x) = (x-\lambda)^2 Q(x)$, and we have

$$L(\lambda) = 0, \frac{dL}{dx}|_{x=\lambda} = 0.$$

Therefore,

$$L(D)t\exp(\lambda t) = e^{\lambda t}(\lambda L(\lambda) + dL(\lambda)/ds) = 0,$$

and $x\exp(\lambda x)$ is a solution.

More generally, we have

$$\begin{aligned} \frac{d^r(L(D)e^{sx}}{ds^r} &= L(D)x^r e^{sx} = \frac{d^r e^{st} L(s)}{ds^r} \\ &= e^{st}(L^{(r)}(s) + rtL^{(r-1)}(s) + ... + t^r L(s)) \end{aligned}$$

where we have the binomial expression for differentiate products(general leibniz rule).If $s = \lambda$ is root to $L(x)$ with multiplicity greater than r, then $L(D)x^r e^{\lambda x} = 0$. That is, $x^r \exp(\lambda x)$ is a solution. □

17.2.2.3 *The case of complex roots*

Lemma 17.2.1 (complex roots comes as conjugate pairs for real-valued ODE). *Consider a linear homogeneous differential equation given by*

$$L(D)y = 0$$

where $L(x)$ is a polynomial with real-valued coefficients. If $a + bi, a, b \in \mathbb{R}$ is a root to $L(x) = 0$ such that $\exp((a + bi)t)$ is a solution, then $a - bi, a, b \in \mathbb{R}$ is also a root to $L(x) = 0$ such that $\exp((a - bi)t)$ is also a solution.

Proof. Use Theorem 17.2.10 and Theorem 4.18.5 □

Lemma 17.2.2 (conversion between trigonometric functions and complex exponential functions). *Let*

$$y_1(t) = \exp((a + ib)t), y_2 = \exp((a - ib)t), a, b \in \mathbb{R}$$

be the solutions to

$$Ly = 0,$$

then

$$u_1(t) = \Re(y_1(t)) = \exp(at)\cos(bt), u_2(t) = \Im(y_2(t)) = \exp(at)\sin(bt),$$

are also the solutions.

Proof. Use the linearity of solutions in Theorem 17.2.1, we have

$$u_1(t) = \frac{y_1(t) + y_2(t)}{2}, u_2(t) = \frac{y_1(t) - y_2(t)}{2i}.$$

□

17.2.2.4 *The complete solution set*

Theorem 17.2.12 (construct the complete solution set). *[3, p. 49] Let $\lambda_1, \lambda_2, ..., \lambda_k \in \mathbb{C}$,with multiplicity $\mu_1, \mu_2, ..., \mu_k, \sum \mu_i = n$, be the roots to order n polynomial $L(x) = 0$, then the complete linear independent solution set is given by*

$$\begin{aligned}
&x^r e^{\lambda_1 x}, r \in \{0, 1, 2, ..., \mu_1 - 1\} \\
&x^r e^{\lambda_2 x}, r \in \{0, 1, 2, ..., \mu_2 - 1\} \\
&\cdots \\
&x^r e^{\lambda_k x}, r \in \{0, 1, 2, ..., \mu_k - 1\}
\end{aligned}$$

The solution to the differential equation $L(D)y = 0$ can be written by

$$y = \sum_{i=0}^{\mu_1 - 1} c_{1i} x^r \exp(\lambda_1 x) + \sum_{i=0}^{\mu_2 - 1} c_{2i} x^r \exp(\lambda_2 x) + \cdots + \sum_{i=0}^{\mu_k - 1} c_{2i} x^r \exp(\lambda_k x).$$

Example 17.2.6 (two distinct real roots). The general solution of

$$y'' - y' - 2y = 0$$

is given by

$$y = c_1 \exp(-t) + c_2 \exp(2t).$$

This is because the characteristic polynomial is given by

$$p(r) = r^2 - r - 2 = (r+1)(r-2).$$

Its two roots are -1 and 2, with associated solutions $\exp(-t)$ and $\exp(2t)$.

Example 17.2.7 (real roots with multiplicity). The general solution of

$$u'' + 6u + 9 = 0$$

is given by

$$u = c_1 \exp(-3t) + c_2 t \exp(-3t).$$

This is because the characteristic polynomial is given by

$$p(r) = r^2 + 6r + 9 = (r+3)^2.$$

Its root is -3 with multiplicity of 2, with associated solutions $\exp(-3t)$ and $t \exp(-3t)$.

Example 17.2.8 (distinct complex roots). The general solution of

$$u'' + 9u = 0$$

is given by

$$u = c_1 \exp(-3it) + c_2 t \exp(-3it) = c3 \cos(3t) + c4 \sin(3t).$$

This is because the characteristic polynomial is given by

$$p(r) = r^2 + 9 = (r + 3i)(r - 3i).$$

Its root is -3i and 3i, with associated solutions $\exp(-3it)$ and $t \exp(3it)$.

Example 17.2.9 (complex roots with multiplicity). The general solution of

$$u^{(4)} + 8u^{(2)} + 16u = 0$$

is given by

$$\begin{aligned} u &= c_1 \exp(-2it) + c_2 \exp(2it) + c_3 t \exp(-2it) + c_4 t \exp(2it) \\ &= c_5 \cos(-2t) + c_6 \sin(2it) + c_7 t \cos(-2t) + c_8 t \sin(2t) \end{aligned}$$

This is because the characteristic polynomial is given by

$$p(r) = r^4 + 8r^2 + 16 = (r^2 + 4)^2.$$

Its roots are $\pm 2i, \pm 2i$, with associated solutions $\exp(-2t), \exp(2t)$ and $t \exp(2it), t \exp(-2it)$.

Example 17.2.10 (a complete example). The general solution of

$$(D^3 - 2D^2)(D^2 - 2D + 10)^3(D^2 + 4D + 29)y = 0, D \triangleq \frac{d}{dt},$$

is given by

$$\begin{aligned} u = c_1 &+ c_2 t + c_3 \exp(2t) + c_4 \exp(t) \cos(3t) \\ &+ c_5 \exp(t) \sin(3t) + c_6 t \exp(t) \cos(3t) + c_7 t \exp(t) \sin(3t) \\ &+ c_8 t^2 \exp(t) \cos(3t) + c_9 t^2 \exp(t) \sin(3t) \\ &+ c_{10} \exp(-2t) \cos(5t) + c_{11} \exp(-2t) \sin(5t) \end{aligned}$$

This is because the characteristic polynomial is given by

$$\begin{aligned} p(r) &= (r^3 - 2r^2)(r^2 - 2r + 10)^3(r^2 + 4r + 29) \\ &= r^2(r-2)((r-1)^2 + 9)^3((r+2)^2 + 25) \end{aligned}$$

Its 11 roots are $0, 0, 2, 1 \pm 3i, 1 \pm 3i, 1 \pm 3i, -2 \pm i5$.

17.2.3 Solution to non-homogeneous ODEs

17.2.3.1 *General principles*

Theorem 17.2.13 (general principle for linear non-homogeneous ODEs). *Consider a linear differential equation*

$$\frac{\partial^n y}{\partial t^n} + a_{n-1}(t)\frac{\partial^{n-1} y}{\partial t^{n-1}} + ... + a_0(t)y = g(t).$$

The solution space is given by the vector space of the solution $\{c_1y_1 + c_2y_2 + ... + c_ny_n, c_1, ..., c_n \in \mathbb{R}^n\}$ *to the associated homogeneous solution plus a particular solution* y_p *satisfy*

$$\frac{\partial^n y_p}{\partial t^n} + a_{n-1}(t)\frac{\partial^{n-1} y_p}{\partial t^{n-1}} + ... + a_0(t)y_p = g(t).$$

Moreover, when given the initial conditions

$$\frac{\partial^i y}{\partial t^i} = b_i, \forall i = 0, 1, ..., n-1,$$

we can uniquely determine $c_1, ..., c_n$.

Proof. Directly plug in to verify. □

Remark 17.2.5.

- The particular solution y_P does not need to satisfy the initial condition. The coefficients associated with the homogeneous solution needs to satisfy the initial condition.

Lemma 17.2.3 (decomposition of particular solution). *Let $y_{P,1}(t), y_{P,2}(t), ..., y_{P,k}(t)$ be the particular solutions such that*

$$Ly_{P,i} = g_i(t), i = 1, 2, ..., k.$$

Then

$$L(y_{P,1} + y_{P,2} + ... + y_{P,k}) = g_1(t) + g_2(t) + ... + g_k(t).$$

Proof. Using the linearity of the operator L. □

17.2.3.2 *Key identity approach*

Lemma 17.2.4. *If z is a root to $L(x) = 0$ with multiplicity m, then*

$$L[t^m \exp(zt)] = L(z)^{(m)} \exp(zt).$$

and

$$L[t^r \exp(zt)] = 0, \forall r < m.$$

$$L[t^s \exp(st)] = L(z)^{(m)} t^{s-m} \exp(zt), \forall s > m.$$

Proof. If z is a root to $L(x) = 0$ with multiplicity m, then we can write

$$L(x) = (x - z)^m Q(x).$$

For $r = m$, we have

$$L[t^r \exp(zt)] = \frac{d}{dx} L[\exp(xt)]_{x=z} = [\frac{d}{dx} L(x) \exp(xt)]_{x=z} = L(z)^{(m)} \exp(zt).$$

□

Example 17.2.11. Consider

$$y'' - 6y' + 9y = 4\exp(3t).$$

3 is the root to $L(z) = z^2 - 6z + 9$ with multiplicity of 2. We have

$$\begin{aligned}
L[\exp(3t)] &= L(3)\exp(3t) = 0 \\
L[t\exp(3t)] &= L'(3)\exp(3t) + 3L(3)\exp(3t) = 0 \\
L[t^2\exp(3t)] &= L''(3)\exp(3t) + 2tL'(3)\exp(3t) + t^2L(3)\exp(3t) \\
&= L''(3)\exp(3t) = 2\exp(3t) \\
L[t^3\exp(3t)] &= tL''(3)\exp(3t) = 2t\exp(3t)
\end{aligned}$$

Theorem 17.2.14 (key identity approach to particular solution for non-homogeneous linear ODE). *Consider linear non-homogeneous equation given by*

$$L(D)y = f(t),$$

where $f(t)$ has the form of

$$f(t) = (h_0t^d + h_1t^{d-1} + ... + h_d)e^{ut}\cos(vt) + (g_0t^d + g_1t^{d-1} + ... + g_d)e^{ut}\sin(vt).$$

We say f has characteristic form with degree d and characteristic $u + iv$.

Then we can find coefficients $\beta_0, \beta_1, \beta_2, ... \in \mathbb{C}$ such that

$$f(t) = L[\sum_{i=1} \beta_i t^i \exp((u + iv)t)]$$

and the particular solution is given by

$$y_P = \sum_{i=1} \beta_i t^i \exp((u + iv)t).$$

Proof. It is straight forward to see that

$$L[y_P] = f(t).$$

□

Remark 17.2.6. The key identity approach only works for the cases where $f(t)$ take the particular forms.

Lemma 17.2.5 (conversion between trigonometric functions and complex exponential functions). *Let*

$$y_1(t) = \exp((a + ib)t), y_2 = \exp((a - ib)t), a, b \in \mathbb{R}$$

be the solutions to

$$Ly = 0,$$

then

$$u_1(t) = \Re(y_1(t)) = \exp(at)\cos(bt), u_2(t) = \Im(y_2(t)) = \exp(at)\sin(bt),$$

are also the solutions.

Proof. Use the linearity of solutions in Theorem 17.2.1, we have

$$u_1(t) = \frac{y_1(t) + y_2(t)}{2}, u_2(t) = \frac{y_1(t) - y_2(t)}{2i}.$$

$\square$

Example 17.2.12. Consider the ODE given by

$$Ly = y'' - 2y' + 5y = te^t = f(t).$$

The characteristic polynomial roots are 1 ± 2, and therefore we have homogeneous solution

$$y_H(t) = c_1 \exp(t)\cos(2t) + c_2 \exp(t)\sin(2t).$$

We then use key identity approach to find the particular solution. We have

$$\begin{aligned} L(\exp(zt)) &= (z^2 - 2z + 5)\exp(zt) \\ L(t\exp(zt)) &= \frac{d}{dt}[(z^2 - 2z + 5)\exp(zt)] = (z^2 - 2z + 5)t\exp(zt) + (2z - 2)\exp(zt) \end{aligned}$$

Plug in $z = 1$, we have

$$L(\exp(t)) = 4\exp(t), L(t\exp(t)) = 4t\exp(t).$$

Therefore, we can construct a particular solution as

$$y_P(t) = \frac{1}{4}t\exp(t),$$

such that

$$L(\frac{1}{4}t\exp(t)) = t\exp(t) = f(t).$$

Example 17.2.13. Consider the ODE given by

$$Ly = y'' - 2y' + 5y = te^t = f(t).$$

The characteristic polynomial roots are 1 ± 2, and therefore we have homogeneous solution

$$y_H(t) = c_1 \exp(t)\cos(2t) + c_2 \exp(t)\sin(2t).$$

We then use key identity approach to find the particular solution. We have

$$L(\exp(zt)) = (z^2 - 2z + 5)\exp(zt)$$
$$L(t\exp(zt)) = \frac{d}{dt}[(z^2 - 2z + 5)\exp(zt)] = (z^2 - 2z + 5)t\exp(zt) + (2z - 2)\exp(zt)$$

Plug in $z = 1$, we have

$$L(\exp(t)) = 4\exp(t), L(t\exp(t)) = 4t\exp(t).$$

Therefore, we can construct a particular solution as

$$y_P(t) = \frac{1}{4}t\exp(t),$$

such that

$$L(\frac{1}{4}t\exp(t)) = t\exp(t) = f(t).$$

Example 17.2.14. Consider the ODE given by

$$Ly = y'' - 3y' - 10 = te^{-2t} = f(t).$$

The characteristic polynomial roots are $5, -2$, and therefore we have homogeneous solution

$$y_H(t) = c_1 \exp(5t) + c_2 \exp(-2t).$$

We then use key identity approach to find the particular solution. We have

$$\begin{aligned}
L(\exp(zt)) &= (z^2-3z-10)\exp(zt)\\
L(t\exp(zt)) &= \frac{d}{dt}[(z^2-3z-10)\exp(zt)] = (z^2-3z-10)t\exp(zt) + (2z-3)\exp(t)\\
L(t^2\exp(zt)) &= \frac{d}{dt}[(z^2-3z-10)\exp(zt)] = (z^2-3z-10)t^2\exp(zt) + \\
&\quad 2(2z-3)t\exp(t) + 2\exp(t)
\end{aligned}$$

Plug in $z=-2$, we have

$$\begin{aligned}
L(\exp(-2t)) = 0, L(t\exp(-2t)) &= -7\exp(-2t),\\
L(t\exp(-2t)) &= -14t\exp(-2t) + 2\exp(-2t).
\end{aligned}$$

Therefore, we can construct a particular solution y_P such that

$$L(y_P(t)) = L(-\frac{1}{14}t^2\exp(-2t) - \frac{2}{98}t\exp(-2t)) = te^{-2t} = f(t),$$

where

$$y_P(t) = -\frac{1}{14}t^2\exp(-2t) - \frac{2}{98}t\exp(-2t).$$

Corollary 17.2.14.1. *Given*

$$y'' + by' + cy = G(x)$$

with $G(x)$ given by $P_n(x)$, a polynomial of degree n. Denote $L(x) = x^2 + bx + c$. It follows that

- *if 0 is not a root to $L(x)$, then the particular solution will also be a polynomial of degree n. The coefficients of the solution can be determined by substituting.*
- *if 0 is a root to $L(x)$ with multiplicity of 1, then the particular solution will be a polynomial of degree $n+1$. The coefficients of the solution can be determined by substituting.*
- *if 0 is a root to $L(x)$ with multiplicity of 2, then the particular solution will be a polynomial of degree $n+2$. The coefficients of the solution can be determined by substituting.*

Proof. (1) Use Theorem 17.2.14 with characteristic equals 0. (2) Note that □

Example 17.2.15.

- The equation
$$y'' + y' + y = 1$$
has a particular solution of 1.
- The equation
$$y'' + y' + y = x$$
has a particular solution of the form $ax + b$. We can determine a, b by
$$a + b = 0, a = 1.$$
- The equation
$$y'' + y' + y = dx^2 + ex + f$$
has a particular solution of the form $ax^2 + bx + c$. We can determine a, b, c by
$$a = d, 2a + b = e, 2a + b + c = f.$$

Corollary 17.2.14.2. *Given*
$$L(D)y = y'' + by' + cy = G(x)$$
with $G(x)$ given by $e^{kx}P_n(x)$, a polynomial of degree n. Denote $L(x) = x^2 + bx + c$. and k is not a solution of the characteristic equation.It follows that

- *if k is not a root to $L(x)$, then the particular solution will be the form $e^{kx}Q(x)$, $Q(x)$ is a polynomial of degree n. The coefficients of the solution can be determined by substituting.*
- *if k is a root to $L(x)$ with multiplicity 1, then the particular solution will be the form $e^{kx}Q(x)$, $Q(x)$ is a polynomial of degree $n + 1$. The coefficients of the solution can be determined by substituting.*
- *if k is not a root to $L(x)$ with multiplicity 2, then the particular solution will be the form $e^{kx}Q(x)$, $Q(x)$ is a polynomial of degree $n + 2$. The coefficients of the solution can be determined by substituting.*

Example 17.2.16. The equation
$$y'' + 4y = e^{3x}$$

has a particular solution of the form ae^{3x}. We can determine a by

$$9a + 4a = 1, a = 1/13.$$

17.2.4 First order linear differential equation

Theorem 17.2.15 (general solution to first order linear differential equation). *The equation of the form*

$$\frac{dy}{dx} + p(x)y = Q(x), y(0) = y_0$$

has the solution

$$y(x) = \exp(\int_0^x -p(t)dt)y_0 + \int_0^x \exp(\int_t^x -p(\tau)d\tau)Q(t)dt.$$

- *If we introduce* $\Phi(t,\tau) = \exp(-\int_\tau^t p(u)du)$, *then*

$$y = \Phi(x,0)y_0 + \int_0^x \Phi(x,t)Q(t)dt.$$

- *If* $p(x) = p$, *then*

$$y(x) = \exp(-p(t)x)y_0 + \int_0^x \exp(-p(x-t))Q(t)dt.$$

Proof. This is a special case of Theorem 17.3.6. □

Example 17.2.17. Given the example

$$y' + \frac{y}{x} = \frac{1}{x^2}, y(1) = 1,$$

we can calculate the state transition function

$$\Phi(t_1,t_2) = \exp(-\int_{t_1}^{t_2} \frac{1}{x}dx) = \frac{t_1}{t_2}.$$

Then

$$y(x) = \Phi(x,1)y(1) + \int_1^x \Phi(x,\tau)\frac{1}{\tau^2}d\tau = \frac{1}{x} + \frac{\ln x}{x}.$$

Example 17.2.18. Consider the differential equation

$$\frac{dy}{dt} = \frac{4y}{t}, y(1) = 1.$$

Since t cannot equal 0, we first consider the situation of $t \in (0, +\infty)$. By assuming $y \neq 0$ for $t \in (0, +\infty)$, we have

$$\frac{dy}{y} = \frac{4dt}{t} \implies \ln y = 4 \ln t + C \implies y = ke^{4 \ln t} = kt^4.$$

The initial condition implies $k = 1$, also $y \neq 0 \forall t > 0$.

For $t < 0$, we have

$$\frac{dy}{y} = \frac{4d - t}{-t} \implies \ln y = 4 \ln(-t) + C \implies y = ke^{4 \ln(-t)} = k(-t)^4.$$

However, no initial condition can be used to determine k when $t < 0$. That is, when $t < 0$, there is no unique solution.

17.3 Linear system

A discrete-time linear system is given as:

$$x(k+1) = A(k)x(k) + w(k)$$

where $x(k), w(k) \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}$. If $w(k) = 0$, then it is called **homogeneous system**;otherwise it is called **non-homogeneous system**.

17.3.1 Solution space for linear homogeneous system

In this subsection, we discuss the solution property for the linear homogeneous equation of dimension n:

$$\dot{x} = A(t)x, x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n}.$$

Theorem 17.3.1 (Existence and uniqueness of solutions). *[3, p. 8]*

- *Let the equation*
$$\dot{x} = f(t,x), x \in \mathbb{R}^n$$
*be given, where $f(t,x)$ is defined in some domain B in $\mathbb{R}^{n=1}$. Suppose both f and $\partial f/\partial x_i, i = 1,2...,n$ are **defined and continuous** in B. Then for every point $(t_0, x_0) \in B$, there exists a unique solution $x = \phi(t)$ satisfying $\phi(t_0) = x_0$ and defined in some neighborhood of (t_0, x_0).*
- *Particularly for*
$$\dot{x} = f(t,x) = A(t)x(t),, x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n},$$
*if $A(t)$ is **continuous** over some domain B in $\mathbb{R}^{n=1}$,Then for every point $(t_0, x_0) \in B$, there exists a unique solution $x = \phi(t)$ satisfying $\phi(t_0) = x_0$ and defined in some neighborhood of (t_0, x_0).*

Remark 17.3.1 (consequences).

- Suppose we have two solutions satisfying $\dot{x} = f(t,x)$ and initial conditions, then two the solution must be equal. This property can be summarized as: **once agreeing on a point, they will agree on a neighborhood.**
- If we have two solutions **only** satisfying $\dot{x} = f(t,x)$, and the two solutions need not be equal.

Lemma 17.3.1 (linearity of solutions to linear homogeneous systems). *Suppose $z_1(t), z_2(t), ..., z_k(t)$ is a linearly dependent set of solutions to the linear homogeneous differential equations,given by,*

$$\dot{x} = f(t,x) = A(t)x(t), x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n},$$

then

$$z(t) \triangleq \sum_{i=1}^{k} c_i z_i(t)$$

is also a solution.

Proof. Directly from the linearity of the differential operator.

$$\begin{aligned}\frac{d}{dt}z(t) &= \sum_{i=1}^{k} c_i \frac{d}{dt} z_i(t)\\ &= \sum_{i=1}^{k} c_i A z_i(t)\\ &= A\sum_{i=1}^{k} c_i z_i(t)\\ &= Az\end{aligned}$$

□

Theorem 17.3.2 (fundamental solution property of linear homogeneous system). *[3, p. 18]There exists a set F of n linearly independent solution to the linear homogeneous system of dimension n given by*

$$\dot{x} = f(t,x) = A(t)x(t), x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n},$$

Moreover, any solution can be decomposed as the linear combination of F.

Proof. (1) Consider n 1-of-n hot spot initial conditions, each initial condition combined with the differential equation itself will have a unique solution ϕ_i, as guaranteed by the uniqueness and existences theorem(Theorem 17.3.1). Moreover, this set of solution $\{\phi_1, \phi_2, ..., \phi_n\}$ must be linearly independent of each other since they disagree on the point $t = t_0$(i.e. the initial conditions). (2) For any other solution ψ to the equation, let $c = (c_1, c_2, ..., c_n)$ be the value of ψ at t', then we can find a linear combination of $\{\phi_1, \phi_2, .., \phi_n\}$ that agree with ψ at t'. Again the uniqueness theorem(Theorem 17.3.1) will guarantee that this linear combination and ψ are equal to each other. □

Remark 17.3.2 (implications).
This theorem implies that if we can find n linearly independent solutions, we can construct any other solutions to satisfy any initial conditions. In otherwise, once found n linearly independent solutions, we completely solve the problem.

Corollary 17.3.2.1. *The solutions to the linear homogeneous equation form a n dimensional linear space.*

17.3.2 Linear independence and the Wronskian

Definition 17.3.1 (linear independence of solutions). *[1, p. 40] Given a set of solutions $x_1, x_2, ..., x_k \in \mathbb{R}$ on the interval $[t_0, t_1]$ to the homogeneous linear differential equation $\dot{x} = A(t)x(t)$, we say they are linearly independent if*

$$\sum_{i=1}^{k} c_i y_i(t) = 0, \forall t \in [t_0, t_1]$$

only hold when all $c_i = 0$.

Definition 17.3.2 (fundamental system/matrix). *Let $\phi_1, ...\phi_n$ be the solutions to the linear homogeneous system $\dot{x} = A(t)x(t)$. If they are linearly independent , then they are called the **fundamental system**. The matrix $\Phi = [\phi_1, ...\phi_n]$ is called **fundamental matrix**.*

Definition 17.3.3 (The Wronskian). *[3, p. 22]The Wronskian W is defined as a scalar function for the fundamental matrix given by*

$$W[\Phi](t) = det(\Phi(t)).$$

Lemma 17.3.2 (Wronskian and linear independence). *Give a set of solutions $x_1, x_2, ..., x_k \in \mathbb{R}$ on the interval $[t_0, t_1]$ to the homogeneous linear differential equation $\dot{x} = A(t)x(t)$, we say they are linearly independent if and only if its Wronskian $W[x_1, x_2, ..., x_k](t) \neq 0$, for all $t \in [t_0, t_1]$.*

Proof. Based on the definition of linear independence, they are linearly independent if

$$\sum_{i=1}^{k} c_i y_i(t) = 0, \forall t \in [t_0, t_1]$$

only hold when all $c_i = 0$. This is equivalent to say that

$$W \triangleq det[x_1, x_2, ..., x_k] \neq 0$$

for all $t \in [t_0, t_1]$. □

Remark 17.3.3. This theorem is not easy to use; the following Liouvilles' formula provides an easy way to check linear independence among solutions.

Lemma 17.3.3 (Liouvilles' formula for Wronskian). *[3, p. 22][5, p. 481]*

- *Let $W[\Phi]$ be the Wronskian associated with the fundamental matrix of a linear homogeneous system, then*

$$W(t) = W(t_0)\exp[\int_{t_0}^{t} Tr(A(s))]ds.$$

- *The Wronskian has the property of: once zero, always zero; once nonzero, always nonzero.*

Proof. (1) see reference; (2) Since $\exp[\int_{t_0}^{t} Tr(A(s)ds] \neq 0$, the Wronskian has the property of **once zero at a point, all zero; once nonzero at a point, all nonzero**. □

Theorem 17.3.3 (necessary and sufficient conditions for fundamental system). *A necessary and sufficient condition for $\phi_1, \phi_2, ..., \phi_n$ to be a **fundamental system** is that there exists a $t_0 \in (r_1, r_2)$, such that $W(t_0) \neq 0$. where (r_1, r_2) is the interval where the conditions for existence and uniqueness of solutions are satisfied.*

Proof. Use the above two lemma. □

Example 17.3.1. The vector-valued function

$$x_1(t) = \exp(5t)\begin{bmatrix} 1 \\ 1 \end{bmatrix}, x_2(t) = \exp(t)\begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

are solutions of the differential system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The associated Wronskian is given by

$$W(t) = det \begin{bmatrix} \exp(5t) & \exp(t) \\ \exp(5t) & -\exp(t) \end{bmatrix} = -2\exp(6t) \neq 0.$$

Therefore, x_1 and x_2 constitute a fundamental system.

17.3.3 The fundamental system and solution method

Theorem 17.3.4 (solution to initial value problem via fundamental system/matrix property).

- *Let $\Phi(t)$ be the fundamental system, then*

$$\dot{\Phi}(t) = A(t)\Phi(t), A(t) = \dot{\Phi}(t)\Phi(t)^{-1}.$$

- *Further more, any solution $x(t)$ satisfying the initial condition $x(t_0) = x_0$ is given as*

$$x(t) = \Phi(t)\Phi^{-1}(t_0)x_0.$$

- *In particular, if $\Phi(t_0) = I$(i.e., the solution of ϕ_i satisfies initial condition of e_i), then*

$$x(t) = \Phi(t)x_0.$$

Proof. (1) Each column of Φ is a solution satisfies the equation. (2) First, any solution is a linear combination of the basis solutions(Theorem 17.3.2), i.e., $x(t) = \Phi(t)c, c \in \mathbb{R}^n$, note that

$$x_0 = \Phi(t_0)c \Rightarrow c = \Phi^{-1}(t_0)x_0$$

. □

Example 17.3.2. The vector-valued function

$$x_1(t) = \exp(5t)\begin{bmatrix} 1 \\ 1 \end{bmatrix}, x_2(t) = \exp(t)\begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

are solutions of Consider the differential system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

with initial condition $x_0 = (4, -2)^T$. It has the fundamental matrix given by

$$\Phi = \begin{bmatrix} \exp(5t) & \exp(t) \\ \exp(5t) & -\exp(t) \end{bmatrix}.$$

Its solution can be written by

$$\begin{aligned} x(t) &= \Phi(t)\Phi(0)^{-1}x_0 \\ &= \begin{bmatrix} \exp(5t) & \exp(t) \\ \exp(5t) & -\exp(t) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}\begin{bmatrix} 4 \\ -2 \end{bmatrix} \end{aligned}$$

17.3.4 The non-homogeneous linear equation

In this subsection, we discuss the solutions for the linear non-homogeneous equation of dimension n:

$$\dot{x} = A(t)x + B(t), x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n}, B(t) \in \mathbb{R}^n.$$

We first introduce the concept of **state transition matrix**, which plays a critical role in formulating solutions.

Definition 17.3.4 (state transition matrix). *[1, p. 114] The state transition matrix for linear system*

$$\dot{x}(t) = A(t)x(t)$$

is a $n \times n$ matrix function satisfying

$$\frac{d}{dt}\Phi(t,\tau) = A(t)\Phi(t,\tau), \Phi(\tau,\tau) = I.$$

Lemma 17.3.4 (state transition matrix derived from fundamental matrix). *The state transition matrix for linear system*

$$\dot{x}(t) = A(t)x(t)$$

is

$$\Phi(t,\tau) = X(t)X(\tau)^{-1}$$

where $X(t)$ is the fundamental matrix. Moreover, let $x(t)$ be the solution, then

$$x(t) = \Phi(t,\tau)x(\tau).$$

Proof. (1) We can verify that $\Phi(\tau,\tau) = I$ and

$$\frac{d}{dt}\Phi(t,\tau) = \frac{d}{dt}X(t)X(\tau)^{-1} = AX(t)X(\tau)^{-1} = A\Phi(t,\tau).$$

(2) See Theorem 17.3.4. □

Example 17.3.3 (state transition matrix for linear system with constant coefficients). The linear system with constant coefficients

$$\dot{x} = Ax$$

has state transition matrix given as

$$\Phi(t,\tau) = e^{A(t-\tau)}.$$

It can be verified that

$$\Phi(\tau,\tau) = I, \frac{d}{dt}\Phi(t,\tau) = \frac{d}{dt}e^{A(t-\tau)} = Ae^{A(t-\tau)} = A(t)\Phi(t,\tau).$$

Theorem 17.3.5 (general solution principle for linear non-homogeneous systems). *Consider a linear system*

$$\dot{x} = A(t)x + B(t), x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n}, B(t) \in \mathbb{R}^n,$$

The solution is given by the homogeneous solution $y_H(t)$ satisfying

$$\dot{y}_H = A(t)y_H$$

plus a particular solution y_P satisfy

$$\dot{y}_P = A(t)y_P + B(t).$$

Proof. Directly plug in to verify. □

Theorem 17.3.6 (solution to linear non-homogeneous systems via state transition matrix). *The solution $x(t)$ to*

$$\dot{x} = A(t)x + B(t), x \in \mathbb{R}^n, A(t) \in \mathbb{R}^{n\times n}, B(t) \in \mathbb{R}^n,$$

satisfying $x(t) = x_0, r_1 < t_0 < r_2$, is given by

$$x(t) = \Phi(t,0)x_0 + \int_{t_0}^{t} \Phi(t,s)B(s)ds, r_1 < t < r_2$$

where $\Phi(t,\tau)$ is the state transition matrix.

Particularly,

$$y_H(t) = \Phi(t,0)x_0$$

is homogeneous solution, and

$$y_P(t) = \int_{t_0}^{t} \Phi(t,s)B(s)ds.$$

Proof. Take the time derivative, and we get

$$\begin{aligned}
\dot{y}_H(t) &= \frac{d}{dt}\Phi(t,0)x_0 \\
&= A\Phi(t,0)x_0 \\
&= Ax(t)
\end{aligned}$$

and

$$\begin{aligned}
\dot{y}_P(t) &= \frac{d}{dt}(\int_0^t \Phi(t,\tau)B(\tau)d\tau), r_1 < t < r_2 \\
&= \int_0^t \frac{d}{dt}\Phi(t,\tau)B(\tau)d\tau + \Phi(t,t)B(t), r_1 < t < r_2 \\
&= \int_0^t A\Phi(t,\tau)B(\tau)d\tau + \Phi(t,t)B(t), r_1 < t < r_2 \\
&= A(t)\int_0^t \Phi(t,\tau)B(\tau)d\tau + \Phi(t,t)B(t), r_1 < t < r_2 \\
&= A(t)y_P(t) + B(t), r_1 < t < r_2
\end{aligned}$$

where we use the properties of state transition matrix(Definition 17.3.4) that

$$\frac{d}{dt}\Phi(t,0)x_0 = A(t)\Phi(t,0)x_0.$$

□

Example 17.3.4. Consider the differential system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \exp(t) \\ \exp(-t) \end{bmatrix},$$

with initial condition $x_0 = (4, -2)^T$. It has the fundamental matrix given by

$$X(t) = \begin{bmatrix} \exp(t) & \exp(-t) \\ \exp(t) & 3\exp(-t) \end{bmatrix}.$$

Its transition matrix is given by

$$\Phi(t,0) = X(t)X(0)^{-1} = \begin{bmatrix} \exp(t) & \exp(-t) \\ \exp(t) & 3\exp(-t) \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}^{-1}$$

The solution can be written by

$$x(t) = \Phi(t,0)^{-1}x_0 + \int_0^t \Phi(t,\tau)\begin{bmatrix} \exp(\tau) \\ \exp(-\tau) \end{bmatrix} d\tau.$$

Remark 17.3.4 (note on the practical solutions). In practice, the fundamental matrix and state transition matrix can be difficult to obtain when $A(t), B(t)$ take complex function forms. If it is independent of time, then it is possible to solve, as showed in the following sections.

17.3.5 Conversion of linear differential/difference equation to linear systems

We first consider the difference equation example. Given high order linear difference equation, given as $z(k+n) + a_{n-1}y(k+n-1) + ... + a_0(k)z(k) = g(k), k = 0, 1, 2, ...,$we can use the following procedure to convert to linear system. In particular, let

$$x_i(k) = z(k+i-1), i = 1, 2, ..., n-1$$

and

$$x_n = -\sum_{i=0}^{n-1} a_i x_{i+1} + g(k)$$

Then the state vector has n components. See [1, p. 96]. Also, the resulting matrix A is known as *companion matrix*

The matrix A will be

$$\begin{pmatrix} 0 & 1 & 0 & \dots & & 0 \\ 0 & 0 & 1 & \dots & & 0 \\ \vdots & & & & & \\ & & & & & \\ 0 & & & \dots & 0 & 1 \\ -a_n(t)) & -a_{n-1}(t) & & & & -a_1(t)) \end{pmatrix}$$

Now we consider the differential equation example. Given high order linear differential equation, given as

$$\frac{d^n y}{dt^n} + a_1(t)\frac{d^{n-1}y}{dt^{n-1}} + \cdots + a_1(t)\frac{d^{n-1}y}{dt^{n-1}} + a_n(t)y = f(t)$$

,we can use the following procedure to convert to linear system. In particular, let

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots x_n \end{bmatrix} = \begin{bmatrix} y \\ y' \\ \vdots y^{(n-1)} \end{bmatrix}, g(t) = \begin{bmatrix} 0 \\ 0' \\ \vdots f(t) \end{bmatrix}$$

and

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ -a_n(t) & \cdots & -a_3(t) & -a_2(t) & -a_1(t) \end{bmatrix}.$$

then we have

$$\frac{dx}{dt} = A(t)x + g(t).$$

17.3.6 Solution method for discrete system

Consider a linear discrete system given by

$$x(k+1) = A(k)x(k).$$

It can be solved recursively once an initial value of the state is given.

The state-transition matrix of the homogeneous system is

$$\Phi(k,l) = A(k-1)A(k-2)...A(l), k > l,$$

$$\Phi(k,k) = I.$$

With the state-transition matrix, the solution to the homogeneous system can be written as $x(k) = \Phi(k,0)x(0)$.

For a linear discrete time system with external force terms, given by, $x(k+1) = A(k)x(k) + B(k)u(k)$, the general solution can be written as

$$x(k) = \Phi(k,0)x(0) + \sum_{i=1}^{k-1} \Phi(k,i+1)B(i)u(i).$$

It is easy to see that $x(0) = x(0)$, also it can be verified the solution satisfies the original equation by

$$x(k+1) - x(k) = A(k)x(k) - x(k) + B(k)u(k)$$

via subtracting.

17.4 Linear system with constant coefficients

17.4.1 General solutions

The linear system with constant coefficients

$$\dot{x} = Ax$$

has state transition matrix given as

$$\Phi(t,\tau) = e^{A(t-\tau)}.$$

Then the general solution to linear system with constant coefficients

$$\dot{x} = Ax + B(t)$$

It can be showed that

$$\frac{d}{dt}e^{A(t-\tau)} = Ae^{A(t-\tau)}, \Phi(\tau,\tau) = I,$$

which satisfies the definition of state transition matrix(Definition 17.3.4).(2) See Theorem 17.3.6.

Theorem 17.4.1 (general solution to linear system). *The linear system with constant coefficients*

$$\dot{x} = Ax$$

has state transition matrix given as

$$\Phi(t,\tau) = e^{A(t-\tau)}.$$

Then the general solution to linear system with constant coefficients

$$\dot{x} = Ax + B(t)$$

is given as

$$x(t) = \Phi(t,0)x(0) + \int_0^t \Phi(t,\tau)B(\tau)d\tau.$$

17.4.2 System eigenvector method: continuous-time system

17.4.2.1 *Diagonalizable system*

Lemma 17.4.1 (eigenpair and solutions to linear systems). *Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}.$$

If $(\lambda, v), v \in \mathbb{C}^n, \lambda \in \mathbb{C}$ *is an eigenpair of* A*, then*

$$x(t) = \exp(\lambda t)v$$

is a solution to the system.

Proof. Note that

$$\frac{dx}{dt} = \frac{d}{dt}(\exp(\lambda t)v) = \exp(\lambda t)\lambda v = \exp(\lambda t)Av = A\exp(\lambda t)v = Ax,$$

where we use the fact that $\lambda v = Av$. □

Lemma 17.4.2 (conjugate pair for real-value systems). *Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}.$$

- *If* $(\lambda, v), v \in \mathbb{C}^n, \lambda \in \mathbb{C}$ *is an eigenpair of* A*, then* $(\cong \lambda, \cong v)$ *is also an eigenpair of* A*.*
- *Denote* $\lambda = a + bi, a, b \in \mathbb{R}$*, then*

  $$x_1 = \exp(at + bit), x_2 = \exp(at - bit)$$

  are both solutions to the system.
- *Denote* $\lambda = a + bi, a, b \in \mathbb{R}$*, then*

  $$u_1 = \exp(at)\cos(bt), u_2 = \exp(at)\sin(bt)$$

  are both solutions to the system.

Proof. (3) Use linearity of solutions(Lemma 17.3.1). Note that

$$u_1 = \frac{1}{2}(x_1 + x_2), u_1 = \frac{1}{2i}(x_1 - x_2).$$

□

Theorem 17.4.2 (construct complete solution set for diagonalizable system). *Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n},$$

with A being diagonalizable, i.e.,

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

It follows that

- *The fundamental matrix is given by*

$$X(t) = V\exp(t\Lambda).$$

- *The transition matrix is given by*

$$\Phi(t,\tau) = \exp((t-\tau)A) = V\exp((t-\tau)\Lambda)V^{-1},$$

particularly, $\Phi(t,0) = V\exp((t)\Lambda)V^{-1}$.

- *The solution to the system with initial condition $x(\tau)$ is given by*

$$x(t) = \Phi(t,\tau)x(\tau) = V\exp((t-\tau)\Lambda)V^{-1}x(\tau).$$

Proof. (1) Note that

$$\frac{d}{dt}X(t) = \frac{d}{dt}V\exp(t\Lambda) = V\Lambda\exp(t\Lambda) = AV\exp(t\Lambda) = AX(t),$$

plus the fact that $X(t)$ are linearly independent, we have that $X(t)$ is the fundamental matrix. (2) Note that transition matrix is related to fundamental matrix(Lemma 17.3.4) via

$$\Phi(t,\tau) = X(t)X(\tau)^{-1}.$$

(3) Directly from the property of state transition matrix(Lemma 17.3.4). □

Methodology 17.4.1 (the solution recipe summary). *By solving the characteristic equation associated with the linear system $Ly = 0$, we have*

- *each real simple root r yields the solution*

$$\exp(rt).$$

- *each real root r of multiplicity m yields m solutions*

$$\exp(rt), t\exp(rt), t^2\exp(rt), ..., t^{m-1}\exp(rt).$$

- *each conjugate pair of complex simple root $a \pm bi$ yields two solutions*

$$\exp(at)\cos(bt), \exp(at)\sin(bt).$$

- *each conjugate pair of complex root $a \pm bi$ yields $2m$ solutions*

$$\exp(at)\cos(bt), t\exp(at)\cos(bt), \cdots, t^{m-1}exp(at)\cos(bt)$$
$$\exp(at)\sin(bt), t\exp(at)\sin(bt), \cdots, t^{m-1}exp(at)\sin(bt)$$

Example 17.4.1. Consider the linear system

$$\dot{x} = Ax, A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}.$$

We have the following decomposition for A:

$$A = V\Lambda V^{-1}, V = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}, V^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

The transition matrix is given by

$$\Phi(t,0) = \exp(tA) = V\exp(t\Lambda)V^{-1} = \frac{1}{2}\begin{bmatrix} \exp(5t)+\exp(t) & \exp(5t)-\exp(t) \\ \exp(5t)-\exp(t) & \exp(5t)+\exp(t) \end{bmatrix}$$

Example 17.4.2. Consider the linear system

$$\dot{x} = Ax, A = \begin{bmatrix} 3 & 2 \\ -2 & 3 \end{bmatrix}.$$

We have the following decomposition for A:

$$A = V\Lambda V^{-1}, V = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}, \Lambda = \begin{bmatrix} 3+i2 & 0 \\ 0 & 3-i2 \end{bmatrix}, V^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}.$$

Note that

$$\exp(t\Lambda) = \begin{bmatrix} \exp((3+i2)t) & 0 \\ 0 & \exp((3-i2)t) \end{bmatrix} = \exp(3t) \begin{bmatrix} \exp(i2t) & 0 \\ 0 & \exp(i2t) \end{bmatrix}.$$

The transition matrix is given by

$$\begin{aligned}
\Phi(t,0) &= \exp(tA) \\
&= V\exp(t\Lambda)V^{-1} \\
&= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \exp(3t) \begin{bmatrix} \exp(i2t) & 0 \\ 0 & \exp(i2t) \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix} \\
&= \frac{1}{2}\exp(3t) \begin{bmatrix} \exp(i2t)+\exp(-i2t) & -i\exp(i2t)+i\exp(-i2t) \\ i\exp(i2t)-i\exp(-i2t) & \exp(i2t)+\exp(-i2t) \end{bmatrix} \\
&= \exp(3t) \begin{bmatrix} \cos(2t) & \sin(2t) \\ -\sin(2t) & \cos(2t) \end{bmatrix}
\end{aligned}$$

Theorem 17.4.3 (transformation to uncoupled system for homogeneous system). *Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n},$$

with A being diagonalizable, i.e.,

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

Let $y = V^{-1}x$,, then

$$\frac{dy}{dt} = \Lambda y,$$

or explicitly in each component,

$$\begin{aligned}
\dot{y}_1 &= \lambda_1 y_1 \\
\dot{y}_2 &= \lambda_2 y_2 \\
&\vdots \\
\dot{y}_n &= \lambda_n y_n
\end{aligned}$$

The solution to the transformed system is given by

$$y_i(t) = \exp(\lambda_i t) y_i(0), i = 1, 2, ..., n.$$

The solution to the original system is given by

$$x = Vy.$$

Proof. Note that

$$\begin{aligned}\frac{dx}{dt} &= Ax \\ \frac{dx}{dt} &= V\Lambda V^{-1}x \\ V^{-1}\frac{dx}{dt} &= \Lambda V^{-1}x \\ \frac{dV^{-1}x}{dt} &= \Lambda V^{-1}x \\ \frac{dy}{dt} &= \Lambda y\end{aligned}$$

□

Theorem 17.4.4 (confining dynamics in eigenspaces). *[1, p. 136] Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n},$$

with A being diagonalizable, i.e.,

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

- *If the state vector $x(0)$ is initially lying within subspace of $span(v_1, v_2, ..., v_k)$, then it continuous to be the subspace in subsequent time periods.*
- *If the state vector $x(0)$ is initially lying within one-d subspace of $span(v_1)$, then it continuous to be the subspace in subsequent time periods.The solution is given by*

$$x(t) = \exp(\lambda_1 t) y_1(0) v_1.$$

Proof. (1) Note that if $x(0)$ is lying within $span(v_1, v_2, ..., v_k)$, then $x(0) = \sum_{i=1}^k y_i(0)v_i$. And $x(t) = \sum_{i=1}^k y_i(t)v_i$. (2) Note that $x(0) = y_1(0)v_1$. And

$$x(t) = y_1(t)v_1 = \exp(\lambda_1 t) y_1(0) v_1.$$

□

Theorem 17.4.5 (transformation to uncoupled system for non-homogeneous system). *Consider the system*

$$\frac{dx}{dt} = Ax + g(t), x \in \mathbb{R}^n, A \in \mathbb{R}^{n\times n}, g(t) \in \mathbb{R}^n,$$

with A being diagonalizable, i.e.,

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

Let $y = V^{-1}x, h = V^{-1}g$, then

$$\frac{dy}{dt} = \Lambda y,$$

or explicitly in each component,

$$\begin{aligned}
\dot{y}_1 &= \lambda_1 y_1 + h_1(t)\\
\dot{y}_2 &= \lambda_2 y_2 + h_2(t)\\
&\vdots\\
\dot{y}_n &= \lambda_n y_n + h_n(t)
\end{aligned}$$

The solution to the transformed system is given by

$$y_i(t) = \exp(\lambda_i t)y_i(0) + \int_0^t \exp(\lambda_i(t-s))h_i(s)ds.$$

The solution to the original system is given by

$$x = Vy.$$

Proof. Note that

$$\begin{aligned}
\frac{dx}{dt} &= Ax + g(t)\\
\frac{dx}{dt} &= V\Lambda V^{-1}x + g(t)\\
V^{-1}\frac{dx}{dt} &= \Lambda V^{-1}x + V^{-1}g(t)\\
\frac{dV^{-1}x}{dt} &= \Lambda V^{-1}x + h(t)\\
\frac{dy}{dt} &= \Lambda y + h(t).
\end{aligned}$$

To solve the first order ODE, we use Theorem 17.2.15. □

17.4.2.2 *two-by-two non-diagonalizable system*

Now we study a special two-by-two non-diagonalizable system.

Lemma 17.4.3 (two-by-two linear system with Jordan matrix). *The general solution of the system*

$$\frac{dy}{dt} = Jy, J = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

is given by

$$y = c_1 \exp(\lambda t)(tv_1 + v_2) + c_2 \exp(\lambda t) v_2,$$

where

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Note that A is a non-diagonalizable matrix with eigenvalue λ.

Proof. Note that v_1 is the eigenvector associated with the eigenvalue λ. From Lemma 17.4.1, we know that

$$c_1 \exp(\lambda t) v_1$$

will constitute one solution.

To show $c_1 \exp(\lambda t)(tv_1 + v_2)$ is also the solution, we have

$$\begin{aligned}
\frac{d}{dt}(c_1 \exp(\lambda t)(tv_1 + v_2)) &= (c_1 \exp(\lambda t)\lambda(tv_1 + v_2)) + c_1 \exp(\lambda t)\lambda v_1 \\
&= c_1 \exp(\lambda t)(\lambda t v_1 + \lambda v_2 + v_1) \\
&= c_1 \exp(\lambda t)(\lambda t v_1 + A v_2 - A v_1 + \lambda v_1) \\
&= c_1 \exp(\lambda t)(A t v_1 + A v_2) \\
&= A c_1 \exp(\lambda t)(t v_1 + v_2)
\end{aligned}$$

where we use the relation

$$(A - \lambda I)v_2 = v_1 \implies \lambda v_2 = A v_2 - v_1$$

□

Theorem 17.4.6 (two-by-two linear system fundamental matrix). *Consider the system*

$$\frac{dy}{dt} = Ay, A \in \mathbb{R}^{2\times 2}.$$

Suppose we have eigenvalue λ and vector v_1 and v_2 satisfying

$$(A - \lambda I)v_1 = 0$$
$$(A - \lambda I)v_2 = v_1,$$

then the solution is given by

$$y = c_1 \exp(\lambda t)(tv_1 + v_2) + c_2 \exp(\lambda t)v_1.$$

Proof. Note that v_1 is the eigenvector associated with the eigenvalue λ. From Lemma 17.4.1, we know that

$$c_1 \exp(\lambda t)v_1$$

will constitute one solution.

To show $c_1 \exp(\lambda t)(tv_1 + v_2)$ is also the solution, we have

$$\begin{aligned}
\frac{d}{dt}(c_1 \exp(\lambda t)(tv_1 + v_2)) &= (c_1 \exp(\lambda t)\lambda(tv_1 + v_2)) + c_1 \exp(\lambda t)\lambda v_1 \\
&= c_1 \exp(\lambda t)(\lambda t v_1 + \lambda v_2 + v_1) \\
&= c_1 \exp(\lambda t)(\lambda t v_1 + Av_2 - Av_1 + \lambda v_1) \\
&= c_1 \exp(\lambda t)(Atv_1 + Av_2) \\
&= Ac_1 \exp(\lambda t)(tv_1 + v_2)
\end{aligned}$$

where we use the relation

$$(A - \lambda I)v_2 = v_1 \implies \lambda v_2 = Av_2 - v_1$$

□

17.4.2.3 *Non-diagonalizable system*

When A is non-diagonalizable, we have following procedures.

- Find the eigenvalues (we are assuming they are real).
- For each eigenvalue λ_i with multiplicity μ_i, look for μ_i generalized eigenvectors $v_0, v_1, ..., v_{\mu_i}$ by solving $(A - \lambda_i)_i^{\mu} v = 0$. The theory on generalized eigenspace guarantees that μ_i eigenvectors can be found.
- For each generalized eigenvector, we have corresponding solutions of

$$e^{\lambda_i t} v_0, e^{\lambda_i t}(v_0 t + v_1), e^{\lambda_i t}(v_0 t^2/2! + v_1 t + v_2)...$$

- Note that by construction, these n solutions are linearly independent, and therefore will span the solution space.

Also see [1, p. 148].

Theorem 17.4.7 (generalized eigenpair as a solution). *Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n \times n},$$

*where A might not be **diagonalizable**. Let (λ, v) be a generalized eigenpair of degree k such that*

$$(A - \lambda I)^k v = 0.$$

Then

$$x(t) = \exp(\lambda t)(v + t(A - \lambda I)v + \cdots + \frac{t^{k-1}}{(k-1)!}(A - \lambda I)^{k-1} v)$$

is a solution.

Proof. Note that

$$\begin{aligned}(A - \lambda I)x(t) &= \exp(\lambda t)((A - \lambda I)v + t(A - \lambda I)^2 v + \cdots + \frac{t^{k-1}}{(k-1)!}(A - \lambda I)^k v) \\ &= \exp(\lambda t)((A - \lambda I)v + t(A - \lambda I)^2 v + \cdots + \frac{t^{k-2}}{(k-1)!}(A - \lambda I)^{k-1} v)\end{aligned}$$

To show it is the solution, we have

$$\begin{aligned}\frac{dx}{dt} &= \frac{d}{dt}(\exp(\lambda t)((A - \lambda I)v + t(A - \lambda I)^2 v + \cdots + \frac{t^{k-2}}{(k-1)!}(A - \lambda I)^{k-1} v)) \\ &= \lambda \exp(\lambda t)((A - \lambda I)v + t(A - \lambda I)^2 v + \cdots + \frac{t^{k-2}}{(k-1)!}(A - \lambda I)^{k-1} v) \\ &+ \exp(\lambda t)((A - \lambda I)v + t(A - \lambda I)^2 v + \cdots + \frac{t^{k-2}}{(k-1)!}(A - \lambda I)^{k-1} v) \\ &= \lambda x(t) + (A - \lambda I)x(t) \\ &= Ax(t).\end{aligned}$$

□

Theorem 17.4.8 (construct a solution set using a chain of generalized eigenpairs). *Consider the system*

$$\frac{dx}{dt} = Ax, x \in \mathbb{R}^n, A \in \mathbb{R}^{n \times n},$$

where A might not be ***diagonalizable****. Let the vectors $v_1, v_2, ..., v_k, v_1 \neq 0$ be a* ***chain of generated eigenvectors of length k*** *such that*

$$\begin{aligned} v_{r-1} &= (A - \lambda I) v_r \\ v_{r-2} &= (A - \lambda I) v_{r-1} \\ &\vdots \\ v_1 &= (A - \lambda I) v_2 \\ 0 &= (A - \lambda I) v_1. \end{aligned}$$

Then

$$\begin{aligned} x_1(t) &= \exp(\lambda t) v_1 \\ x_2(t) &= \exp(\lambda t)(t v_1 + v_2) \\ &\vdots \\ x_k(t) &= \exp(\lambda t)(v_k + t v_{k-1} + \cdots + \frac{t^{k-1}}{(k-1)!} v_1) \end{aligned}$$

form k linearly independent solutions.

Proof. Theorem 17.4.7 shows that x_i is a solution. And Theorem 4.10.1 shows the independence. □

17.4.3 Equilibrium point

17.4.3.1 *Discrete-time system*

Definition 17.4.1 (equilibrium point). *A vector $\bar{x}$ is an* ***equilibrium point*** *for a dynamic system if*

$$\bar{x}(k+1) = f(\bar{x}(k), k)$$

for discrete time system or

$$d\bar{x}/dt = f(\bar{x}(k), k) = 0$$

Lemma 17.4.4 (equilibrium of linear homogeneous system). *[1, p. 151] The discrete time system $x(k+1) = Ax(k)$ always has the $x = 0$ as an equilibrium point. And if A has eigenvalue of 1, the associated eigenspace are all equilibrium points.*

Proof. (1) $0 = A0$ (2) $Ax = \lambda x, \lambda = 1$. □

Corollary 17.4.8.1 (equilibrium of non-homogeneous system). *[1, p. 151] The dynamical system*

$$x(k+1) = Ax(k) + b$$

has equilibrium point $\bar{x}$ given by

$$\bar{x} = (I - A)^{-1}b.$$

If $I - A$ is non-singular, then the equilibrium point is unique; otherwise, there are infinitely many equilibrium point forming an affine subspace.

Remark 17.4.1. We can easily extend to continuous-time system([1, p. 151]).

17.4.3.2 *Continuous-time system*

Definition 17.4.2 (equilibrium point). *A vector $\bar{x}$ is an* ***equilibrium point*** *for a dynamic system if*

$$\bar{x}(k+1) = f(\bar{x}(k), k)$$

for discrete time system or

$$d\bar{x}/dt = f(\bar{x}(k), k) = 0$$

Lemma 17.4.5 (equilibrium of linear homogeneous system). *[1, p. 151] The discrete time system $x(k+1) = Ax(k)$ always has the $x = 0$ as an equilibrium point. And if A has eigenvalue of 1, the associated eigenspace are all equilibrium points.*

Proof. (1) $0 = A0$ (2) $Ax = \lambda x, \lambda = 1$. □

Corollary 17.4.8.2 (equilibrium of non-homogeneous system). *[1, p. 151] The dynamical system*

$$x(k+1) = Ax(k) + b$$

has equilibrium point $\bar{x}$ given by

$$\bar{x} = (I - A)^{-1}b.$$

If $I - A$ is non-singular, then the equilibrium point is unique; otherwise, there are infinitely many equilibrium point forming an affine subspace.

17.4.4 Stability

Definition 17.4.3 (stability of equilibrium point). *[1, p. 155] Consider a linear time-invariant system*

$$x(k+1) = Ax(k) + b$$

or

$$\dot{x}(t) = Ax(t) + b.$$

- *An equilibrium point $\hat{x}$ of is* ***stable*** *if, for* ***any*** *solution $\phi(t)$ and any $\epsilon > 0$, there exists a $\delta > 0$ satisfying: if*
$$\|\phi(0) - \hat{x}\| < \delta,$$
then for all $t > 0$,
$$\|\phi(t) - \hat{x}\| < \epsilon.$$
- *An equilibrium point $\hat{x}$ of is* ***asymptotically stable*** *if, for* ***any*** *solution $\phi(t)$, there exists a $\delta > 0$ satisfying: if*
$$\lim_{t\to\infty} \|\phi(t) - \hat{x}\| = 0.$$
- *The equilibrium point $\hat{x}$ is* ***unstable*** *if there exists a solution $\phi(t)$ such that*
$$\lim_{t\to\infty} \|\phi(t)\| = \infty.$$

Remark 17.4.2 (interpretation).

- 'stable' means that if a solution starts off close to $\hat{x}$, then stays close to $\hat{x}$ for all positive t.
- 'asymptotic stable' means that if a solution starts off close to $\hat{x}$, then it will eventually stay arbitrarily close to $\hat{x}$ for sufficiently large t.
- 'asymptotic stable' and 'stable' are neither inclusive nor mutual exclusive concepts.

Theorem 17.4.9 (stability condition). *[1, p. 156]A necessary and sufficient condition for an equilibrium point of the system $x(k+1) = Ax(k) + b$ to be asymptotically stable is that the eigenvalues of A all have magnitude less than 1(that is, the eigenvalues must all lie inside the unit circle in the complex plane). If* ***at least one eigenvalue has magnitude greater than 1, the equilibrium point is unstable.***

Proof. Let $\hat{x}$ be the equilibrium point,i.e., $\hat{x} = A\hat{x} + b$. We write the dynamical equation as $x(k+1) - \hat{x} = A(x(k) - \hat{x})$, then $x(k+1) - \hat{x} = A^k(x(0) - \hat{x})$. From Jordan decomposition and its related theorems, we know that if all eigenvalues of A have $|\lambda_i| < 1$, then $A^m \to 0$, as $m \to \infty$. □

Theorem 17.4.10. *[1, p. 157]A necessary and sufficient condition for an equilibrium point of the system $\dot{x}(t) = Ax(t) + b$ to be asymptotically stable is that the eigenvalues of A all have negative real parts. If* ***at least one eigenvalue has positive real part, the equilibrium point is unstable.***

Proof. Same from Jordan decomposition and its related theorems. because e^{tA} can be written as power series of tA, if every $tA^m \to 0$ as $t \to \infty$ (Theorem 4.13.4). □

Theorem 17.4.11 (stability analysis of 2×2 linear homogeneous system). *Consider a 2×2 linear homogeneous system with constant coefficients given by*

$$\dot{x} = Ax.$$

It follows that

- *$\hat{x} = 0$ is the only equilibrium point.*
- *If A has two real eigenvalues r_1, r_2, then $r_1, r_2 < 0$ ensures that $\hat{x} = 0$ be the asymptotically stable point.*
- *If A has two complex eigenvalues $a \pm bi$, then $a < 0$ ensures that $\hat{x} = 0$ be the asymptotically stable point.*
- *If A is non-diagonalizable with a single eigenvalue $r \in \mathbb{R}$, then $r < 0$ ensures that $\hat{x} = 0$ be the asymptotically stable point.*

Proof. (1) (2) The general solution takes the form

$$x(t) = c_1 \exp(r_1 t) v_1 + c_2 \exp(r_2 t) v_2.$$

Only when both $r_1, r_2 < 0$, we have asymptotically stable $\hat{x} = 0$. (3) The general solution takes the form

$$x(t) = c_1 \exp(at) \cos(bt) + c_2 \exp(at) \sin(bt).$$

Only when both $a < 0$, we have asymptotically stable $\hat{x} = 0$. (4) The general solution takes the form

$$x(t) = c_1 \exp(rt) v_1 + c_2 (t \exp(rt) v_1 + \exp(rt) v_2),$$

where v_1 is eigenvector associated with the eigenvalue r, and v_2 is the solution of

$$(A - rI) v_2 = v_1.$$

Only when $r < 0$, we have asymptotically stable $\hat{x} = 0$. □

17.4.5 Complex eigenvalues/eigenvectors

For diagonalizable **real-valued** system, it might happens that there are complex eigenvalues and complex eigenvectors. We can convert them to real-valued basis vectors and 2×2 block matrix. See the linear algebra chapter for more details.

17.4.6 Boundedness of linear systems

Lemma 17.4.6 (boundedness of linear systems). *[6, p. 344] Let $A \in \mathbb{R}^{n\times n}$ be a constant matrix and $x(t), g(t) \in \mathbb{R}^n$ be vector-valued function of time. Then,*

- *The solution to*
$$\frac{dx}{dt} = Ax + g(t), x(0) = x_0$$
is
$$x(t) = \exp(At)x_0 + \int_0^t \exp(A(t-\tau))g(\tau)d\tau.$$
- *If $0 > -a > Re(\lambda_i(A)), \forall i$, then there exists some positive constant scalar c such that*
$$\|x(t)\| \leq ce^{-at}\|x_0\| + c\int_0^t e^{-a(t-\tau)}\|g(\tau)\|\, d\tau.$$
- *If $\|g(\tau)\| \leq \gamma, \forall \tau$ for some scalar constant γ, then $\|x(t)\|$ will be bounded.*
- *If*
$$f(t) = \int_0^t e^{a\tau}\|g(\tau)\|\, d\tau$$
is bounded by a constant, then $\|x(t)\|$ will decay to zero as $t \to \infty$.

Proof. (1) see Theorem 17.4.1. (2)
$$\begin{aligned}
\|x(t)\| &= \left\|\exp(At)x_0 + \int_0^t \exp(A(t-\tau))g(\tau)d\tau\right\| \\
&\leq \|\exp(At)x_0\| + \left\|\int_0^t \exp(A(t-\tau))g(\tau)d\tau\right\| \\
&\leq \|\exp(At)x_0\| + \int_0^t \|\exp(A(t-\tau))\|\|g(\tau)\|\, d\tau \\
&\leq ce^{-at}\|x_0\| + c\int_0^t e^{-a(t-\tau)}\|g(\tau)\|\, d\tau
\end{aligned}$$
where we have use triangle inequality of norm. (3)(4) Straight forward. □

17.4.7 One dimensional Nonlinear dynamical system analysis

Definition 17.4.4 (fixed point, equilibrium point and stability classification). *[7, p. 18] Consider a one-dimensional dynamical system given by*

$$\dot{x} = f(x).$$

*The solution x^* such that $f(x^*) = 0$ is called the **fixed point** or **equilibrium point**.*

*An equilibrium point x^* is **locally stable** if a small deviation from x^* will decay to zero; An equilibrium point x^* is **globally stable** if arbitrary deviation from x^* will decay to zero; An equilibrium point is not locally stable is called **unstable**.*

Lemma 17.4.7 (linear stability analysis). *[7, p. 18] Consider a one-dimensional dynamical system given by*

$$\dot{x} = f(x).$$

Let x^* be an equilibrium point. Assume $f'(x^*) \neq 0$. Then if $f'(x^*) < 0$, x^* is locally stable; if $f'(x^*) > 0$, x^* is unstable.

*The quantity $1/\left|f'(x^*)\right|$ is called **characteristic time scale**, describing how fast a deviation decays when the system being disturbed from x^*.*

Proof. Consider the dynamics of deviation y such that

$$d(x^* + y)/dt = dy/dt = f(x^* + y) = f(x^*) + f'(x^*)y + O(y^2) = f'(x^*)y + O(y^2).$$

When y is small, we have approximation

$$\dot{y} \approx f'(x^*)y,$$

with solution $y(0) = \exp(f'(x^*)t)y(0)$. If $f'(x^*) > 0$, the small deviation will grow; if $f'(x^*) < 0$, vice versa. □

Remark 17.4.3 (the vanishing first order derivative). If the first order derivative is zero, then we need to conduct higher order nonlinear stability analysis to determine the stability of the equilibrium point.

Example 17.4.3. [1, p. 327] Consider the 1D system

$$\dot{x}(t) = ax(t) + cx(t)^2.$$

$x = 0$ is an equilibrium point. Linearize about $x = 0$, we have

$$\dot{y} = ay.$$

Therefore,

- $a < 0$, $x = 0$ is locally stable.
- $a > 0$, $x = 0$ is unstable.
- $a = 0$, cannot tell from linear analysis.

17.5 Least square estimation of constant vectors

17.5.1 linear static estimation from single measurement with no prior information

Considering a random variable X taking values in $\mathbb{R}^n$, the problem of estimating x given measurement Z taking values in $\mathbb{R}^k$ using proposed **linear measurement model**

$$z = Hx + v$$

where H is known, v is the measurement error, is known as linear static estimation problem.

The least square estimation is to obtain $\hat{x}$ via solving the optimization problem

$$\min_{\hat{x}} J = \frac{1}{2}(z - H\hat{x})^T(z - H\hat{x}^T),$$

and its solution is given by the following.

Theorem 17.5.1 (least square estimation without prior information). *Given a linear static estimation problem with no prior information, by minimizing*

$$\min_{\hat{x}} J = \frac{1}{2}(z - H\hat{x})^T(z - H\hat{x}^T)$$

we can obtain the least square estimate of

$$\hat{x} = (H^T H)^{-1} H^T z$$

under the sufficient condition of $\nabla^2 J = H^T H$ being positive definite.

Proof. It is easy to show that first order necessary condition of the optimization problem requires

$$\hat{x} = (H^T H)^{-1} H^T z$$

and $\nabla^2 J = H^T H$. □

Lemma 17.5.1 (condition for $H^T H$ to be positive definite). *H is a $k \times n$ matrix. For the definiteness of $H^T H$, we have following situations:*

- *$k \geq n$ and $rank(H) = n$(H has full column rank)*
- *If $k < n$, then $H^T H$ cannot be positive definite.*

Proof. (1) If $k \geq n$ and $rank(H) = n$, then because $\mathcal{N}(H^TH) = \mathcal{N}(H)$, so we have $rank(H^TH) = rank(H)$ [Theorem 4.4.2]. Therefore H^TH is full rank, H^TH must be positive definite. (2) since $rank(H^TH) = rank(H) \leq k < n$, H^TH must be singular. □

Remark 17.5.1 (interpretation).

- The situation $k < n$ means that the measurement process will cause the lost of information, and therefore the true values cannot be recovered.

17.5.2 linear static estimation from single measurement with prior information

In some cases, we might have prior knowledge regarding the mean and covariance of X, and the distribution of the measurement error. More formally, considering a random variable X taking values in $\mathbb{R}^n$, the problem of estimating x given one measurement z of X taking values in $\mathbb{R}^k$ with prior information about X given as

$$E[X] = x_0, E[(X - EX)(X - EX)^T] = P_0$$

and prior information about measure error

$$E[v] = 0, E[vv^T] = R$$

using proposed linear model

$$z = Hx + v$$

where H is known, v is the measurement error.

The least square estimation is given by the following

Theorem 17.5.2. *Given a linear static estimation problem with prior information, by minimizing*

$$\min_{\hat{x}} J = \frac{1}{2}(x - x_0)^T P_0^{-1}(x - x_0) + \frac{1}{2}(z - Hx)^T R^{-1}(z - Hx)$$

we can obtain the least square estimate of

$$\hat{x} = (H^T R^{-1} H + P_0^{-1})^{-1}(H^T R^{-1} z + P_0^{-1} x_0)$$

under the sufficient condition of $\nabla^2 J = H^T R^{-1} H + P_0^{-1}$ being positive definite.

Proof. It is easy to show that first order necessary condition of the optimization problem gives

$$P_0^{-1}(\hat{x} - x_0) + H^T R^{-1}(z - H\hat{x}) = 0$$

then we can solve
$$\hat{x} = (H^T R^{-1} H + P_0^{-1})^{-1}(H^T R^{-1} z + P_0^{-1} x_0)$$
It can also be showed that $\nabla^2 J = H^T R^{-1} H + P_0^{-1}$. □

Remark 17.5.2 (least square solution as Maximum a posteriori estimation in Bayesian statistics)**.** Let
$$p(x) \propto \exp(-\frac{1}{2}(x - x_0)^T P_0^{-1}(x - x_0))$$
be the prior model and data generation model is given as

$$p(z|x) \propto \exp(-\frac{1}{2}(z - Hx)^T R^{-1}(z - Hx))$$

then
$$\hat{x} = \max_x p(x|z) = \max_x p(x)p(z|x)$$

(using log function)

Moreover, if prior information is not available, which is equivalent to set $p(x)$ as uniform distribution, then least square solution with no prior information gives

$$\hat{x} = \max_x p(x|z) = \max_x p(x)p(z|x) = \max_x p(z|x)$$

17.5.3 Batch and recursive least square estimation with multiple measurements

When we are given $m(m \geq 1)$ measurements $z_i \in \mathbb{R}^k$ of one single $x \in \mathbb{R}^n$, one can concatenate $z = [z_1^T, z_2^T, ..., z_m^T]^T$ and measurement model matrix $H = [H_1; H_2; ...; H_m]$, and then estimate x using H and z. This method is known as **batch least square estimation**.

Batch least square estimation is not flexible and scalable. For example, if new data is available, batch method will re-run the entire data again. Further, for high-dimensional problems, the amount of data also can cause storage issue.

Alternative, one can estimate x using following recursive algorithm:[8, p. 313].

Algorithm 28: Recursive linear least square estimation of dynamical systems.

Input: Given prior on X with $EX = x_0, Cov(X) = P_0$

1 For each new measurement $z_i = H_i x + v_i$, where $V_i \sim N(0, R_i)$ is known

2

$$\hat{x}_i = \hat{x}_{i-1} + K_i(z_i - H_i\hat{x}_{i-1})$$
$$P_i = (P_{i-1}^{-1} + H_i^T R_i^{-1} H_i)^{-1}$$
$$K_i = P_i H_i^T R_i^{-1}$$

(note that update of P_i can be made efficient by using matrix inversion lemma.)

Output: The final estimate x_m

The working mechanism of recursive method can be understood in the following Lemma.

Lemma 17.5.2 (Fundamental lemma of recursive least squares). *[9, p. 160] Let y_1, y_2 be given vectors, and A_1, A_2 be given matrices such that $A_1^T A_1$ is positive definite. Then the vectors*

$$z_1 = \arg\min_{x\in\mathbb{R}^n} \|y_1 - A_1 x\|^2$$

and

$$z_2 = \arg\min_{x\in\mathbb{R}^n} (\|y_1 - A_1 x\|^2 + \|y_2 - A_2 x\|)$$

are also given by

$$z_1 = z_0 + (A_1^T A_1)^{-1} A_1^T (y_1 - A_1 z_0)$$

and

$$z_2 = z_1 + (A_1^T A_1 + A_2^T A_2)^{-1} A_2^T (y_1 - A_1 z_1)$$

Proof. (1) use the result of normal equation $z_1 = (A_1^T A_1)^T A_1^T$ (Theorem 5.4.4) can verify the alternative form of z_1 is equivalent. (2) Again use normal equation and splitting technique, we have

$$z_2 = (A_1^T A_1 + A_2^T A_2)^{-1}(A_1^T y_1 + A_2^T y_2)$$

Use $Ay_1 = A_1^T A_1 z_1$ and we can get the result. □

Remark 17.5.3 (interpretation and significance).

- For a multidimensional least square problem, we can always split into smaller least square problems. For example, we can split

$$y = Ax$$

into

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x$$

- **solutions to smaller least square problems are reusable** by correcting the existing results.(see how z_2 is obtained from z_1.)

Lemma 17.5.3 (Equivalence between batch and recursive least square). *[9, p. 160] Given $m(m \geq 1)$ measurements $z_i\mathbb{R}^k$ of one single $x \in \mathbb{R}^n$ and the measurement model $z_i = H_i x_i + v_i$, one can estimate x using batch least square and recursive least square will give the same result.(Therefore the recursive algorithm is correct.)*

Proof. The recursive method is equivalent to solve the following series of least square problems:

$$\begin{aligned} x_1 &= \arg\min_x \|z_1 - Hx\|^2 + (x - x_0)^T P_0^{-1}(x - x_0) \\ x_2 &= \arg\min_x \|z_1 - Hx\|^2 + \|z_2 - Hx\|^2 + (x - x_0)^T P_0^{-1}(x - x_0) \\ &\cdots\cdots \end{aligned}$$

Then x_n is equivalent to the minimizer of the batch least square problem. □

Recursive methods have the following advantages

- Recursive method has the advantage of avoiding inversion of large matrix $H^T H$.
- Recursive method can have fast convergence rate when the measurement data z_i are 'similar'. (That is, extra measurements does not provide new information.)

Remark 17.5.4. When we are given $m(m \geq 1)$ measurements $z_i \in \mathbb{R}^k$ of m different $x_i \in \mathbb{R}^n$, there are two special cases in the estimation problem:

- Each x_i is independent of the other. In this case, we can individually estimate x_i using single measurement method.
- x_i are generated from a first-order dynamical model, then we can estimate x_i using Kalman filter.

17.5.4 Nonlinear least square estimation

Suppose the measure model is given by nonlinear function, i.e.,

$$z = h(x) + v.$$

The general strategy is to convert to linear estimation problem via

$$z = Hx + v,$$

where $H = \nabla_x h$, and use linear square estimation for multiple cycles.

Algorithm 29: Recursive nonlinear least square of dynamical systems

Input: Given prior on X with $EX = x_0, Cov(X) = P_0$

For each new measurement $z_i = h_i(x) + v_i$, where $V_i \sim N(0, R_i)$ is known

$$\hat{x}_i = \hat{x}_{i-1} + K_i(z_i - H_i\hat{x}_{i-1})$$
$$P_i = (P_{i-1}^{-1} + H_i^T R_i^{-1} H_i)^{-1}$$
$$K_i = P_i H_i^T R_i^{-1}$$

where $H_i = \frac{\partial h_i}{\partial x}|_{x=\hat{x}_{i-1}}$(note that update of P_i can be made efficient by using matrix inversion lemma.)

Output: The estimation $\hat{x}_i, i = 1, 2, ..., N$

Remark 17.5.5 (multiple cycles can improve estimation results)**.**

- For nonlinear system, multiple cycles are needed. For example, the second cycle will use the filtered results from first cycle. Every new cycle can improve the result until cannot be improved. Moreover, multiple cycles cannot guarantee global optimal result can be achieved.
- For linear system, one cycle will guarantee optimal results to be achieved.

17.6 Kalman filer

17.6.1 Preliminary: error propagation in linear systems

17.6.1.1 *Discrete-time system*

Theorem 17.6.1 (Error propagation theorem). *[8, p. 318] Given a discrete-time linear system,*

$$x_k = \Phi_{k-1}x_{k-1} + \Gamma_{k-1}u_{k-1} + \Lambda_{k-1}w_{k-1}$$

where $E[w_k] = 0, E[w_k w_k^T] = Q_k$, and the noise vector is independent. If $x_{k-1} \sim MN(\hat{x}_{k-1}, P_{k-1})$, Then the distribution of x_k is still multivariate normal with mean and covariance given as

$$\hat{x}_k = \Phi_{k-1}\hat{x}_{k-1} + \Gamma_{k-1}u_{k-1}$$

$$P_k = \Phi_{k-1}P_{k-1}\Phi_{k-1}^T + \Lambda_{k-1}Q_{k-1}\Lambda_{k-1}^T$$

Proof. Directly from Theorem 14.1.1 and Corollary 12.1.1.1, the new covariance matrix is the affine transformation of the origin covariance matrix plus the error convariance matrix. □

17.6.1.2 *Continuous-time system*

Lemma 17.6.1 (discrete-time approximation and error propagation). *[8, pp. 326, 336] Given a continuous-time linear system*

$$\dot{x}(t) = F(t)x(t) + G(t)u(t) + L(t)w(t)$$

its discrete-time approximation is given as

$$x(t_k) = \Phi(t_{k-1}, t_k)x(t_{k-1}) + \int_{t_{k-1}}^{t_k} \Phi(t_{k-1}, \tau)(\Gamma(\tau)u(\tau) + L(\tau)w(\tau))d\tau,$$

where $\Phi(t_{k-1}, t_k)$ is the state-transition matrix [Definition 17.3.4].

Note that

$$\hat{x}_k = E[x_k] = \Phi(t_{k-1}, t_k)\hat{x}_{k-1} + \int_{t_{k-1}}^{t_k} \Phi(t_{k-1}, \tau)\Gamma(\tau)u(\tau)d\tau.$$

17.6.2 Batch estimation

Consider the observations $z_i \in \mathbb{R}^k$ are generalized from the following model:

$$x_k = \Phi_{k-1}x_{k-1} + \Gamma_{k-1}u_{k-1} + \Lambda_{k-1}w_{k-1}$$

$$z_k = H_k x_k + n_k$$

where $x_i \in \mathbb{R}^n$ is called the state, $w_k \sim MN(0, Q_k)$ is the state noise, and $n_k \sim MN(0, R_k)$ is the measurement noise. We also assume $w_k, n_k, k = 1, ...$ are independent. Our goal is to infer $x_i, i = 1, ..., n$ from $z_i, i = 1, ..., n$.

We denote $X_t = (x_0, \ldots, x_t)$, $Z_t = (z_1, ...z_t)$. Further, we use notation

$$\begin{aligned} \hat{x}_{t|s} &= E\left[x_t | Z_s\right] \\ \Sigma_{t|s} &= E\left[(x_t - \hat{x}_{t|s})(x_t - \hat{x}_{t|s})\right]^T \end{aligned}$$

We have following theorem summarizing the statistical properties of X_t, Y_t and the inference method.

Theorem 17.6.2 (batch estimation for linear dynamic system).

- *X_t and Z_t are jointly Gaussian.*
- *$x_t | Y_t$ is Gaussian with mean and variance given by*

$$\begin{aligned} \hat{x}_{t|t} &= \bar{x}_t + \Sigma_{x_t Y_t}\Sigma_{Y_t}^{-1}\left(Y_t - \bar{Y}_t\right) \\ P_{t|t} &= Var[x_t] - \Sigma_{x_t Z_t}\Sigma_{Z_t}^{-1}\Sigma_{Z_t x_t} \end{aligned}$$

where $\bar{x}_t$ and $Var[x_t]$ can be obtained from the error propagation theorem [Theorem 17.6.1].

Proof. since X_t and Z_t are linear functions of x_0, w_t, and n_t, we conclude they are all jointly Gaussian. The rest of the results are from conditional distribution formula for multivariate Gaussian distribution [Theorem 14.1.2].

□

Remark 17.6.1 (drawbacks of batch estimation). Although the batch estimation for linear dynamical system seem concise, it is generally impractical:

- The inverse of the matrix, $\Sigma_{Z_t}^{-1}$, has a size grows with t.
- Evaluating $\Sigma_{x_t Z_t}$ is non-trivial.

17.6.3 From batch estimation to Kalman filter

Kalman filter is a smart way to compute $\hat{x}_{t|t}, P_{t|t}$ in a recursive way. It is an iterative algorithm consisting of a prediction step and a correction step.

The first step is the **prediction step**: we are given $x_k|Z_k$ and $P_{k|k}$, the prediction for $x_{k+1}|Z_k$ and $P_{k+1|k}$ can be achieved by noting the propagation equation

$$x_{k+1}|Z_k = \Phi_k x_k|Z_k + \Lambda_k w_k|Z_k = \Phi_k x_k|Z_k + \Lambda_k w_k,$$

we have [Theorem 17.6.1]

$$\hat{x}_{k|k-1} = \Phi_{k-1}\hat{x}_{k-1|k-1}$$
$$P_{k|k-1} = \Phi_{k-1}P_{k-1}\Phi_{k-1}^T + \Lambda_{k-1}Q_{k-1}\Lambda_{k-1}^T.$$

The second step is **correction step**, where the goal is to find $\hat{x}_{k|k}$ and $P_{k|k}$ in terms of the prediction result $\hat{x}_{k|k-1}$ and $P_{k|k-1}$.

Start with $z_k = H_k x_k + n_k$, and condition on Z_{t-1}

$$z_k\,|Z_{k-1} = H_k x_k|\,Z_{t-1} + n_t\,|Z_{t-1} = H_k x_k|\,Z_{k-1} + n_t.$$

Therefore, $x_k|Z_{k-1}$ and $z_k|Z_{k-1}$ are jointly Gaussian with mean and covariance

$$\begin{bmatrix} \hat{x}_{k|k-1} \\ H_k\hat{x}_{k|k-1} \end{bmatrix}, \quad \begin{bmatrix} P_{k|k-1} & P_{t|t-1}H_k^T \\ H_k P_{k|k-1} & H_k P_{k|k-1} H_k^T + R_k \end{bmatrix}.$$

We can get $x_t|Z_t$ via calculating the conditional random variable

$$x_t|Z_{t-1} \mid z_t|Z_{t-1},$$

since $x_t|Z_{t-1} \mid z_t|Z_{t-1} = x_t|z_t, Z_{t-1} = x_t|Z_t$.

Using conditional distribution formula for multivariate Gaussian distribution [Theorem 14.1.2], we have

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1}H_k^T\left(H_k P_{k|k-1}H_k^T + R_k\right)^{-1}\left(z_k - H_k\hat{x}_{k|k-1}\right)$$
$$P_{k|k} = P_{k|k-1} - P_{k|k-1}H_k^T\left(H_k P_{k|k-1}H_k^T + R_k\right)^{-1}H_k P_{k|k-1}$$

this gives us $\hat{x}_{k|k}$ and $P_{k|k}$ in terms of $\hat{x}_{k|k-1}$ and $P_{k|k-1}$

Combining the prediction and correction step, we can obtain $x_{k|k}, P_{k|k}$ from $x_{k-1|k-1}, P_{k-1|k-1}$, which are the key procedures in the Kalman filter algorithm.

In the following, we summarize Kalman filtering algorithm for a more general linear system with external control

$$x_k = \Phi_{k-1}x_{k-1} + \Gamma_{k-1}u_{k-1} + \Lambda_{k-1}w_{k-1}$$

$$z_k = H_k x_k + n_k,$$

where u_k is the external control independent from x_i, w_k.

Algorithm 30: Kalman filtering

Input: Given prior on X with $EX = x_0, Cov(X) = P_0$

1 For each new measurement z_i

2 Prediction:

$$\hat{x}_{k|k-1} = \Phi_{k-1}\hat{x}_{k-1} + \Gamma_{k-1}u_{k-1}$$

$$P_{k|k-1} = \Phi_{k-1}P_{k-1}\Phi_{k-1}^T + \Lambda_{k-1}Q_{k-1}\Lambda_{k-1}^T$$

3 Correction:

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k(z_i - H_k\hat{x}_{k|k-1})$$

$$P_k = (P_{k|k-1}^{-1} + H_k^T R_k^{-1} H_k)^{-1}$$

$$K_k = P_k H_k^T R_k^{-1}$$

4 note that update of P_i can be made efficient by using matrix inversion lemma.

Output: The estimation $\hat{x}_i, i = 1, 2, ..., N$

Remark 17.6.2 (Bayesian interpretation of Kalman filtering)**.** The goal is to obtain

$$p(x_k|z_k, x_{k-1}) \propto p(z_k|x_k)p(x_k|x_{k-1})$$

Since $p(z_k|x_k)$ and $p(x_k|x_{k-1})$ are all multivariate normal distribution, the prediction step is perform estimation of x_k by maximizing $p(x_k|x_{k-1})$ using error propagation theorem(Theorem 17.6.1).
The correction step is maximizing $p(x_k|z_k, x_{k-1})$ with prior information. For a treatment from Bayesian point of view, see [10].

17.6.4 Extended Kalman filter for nonlinear system

Consider the observation $z_i \in \mathbb{R}^k$ is generalized from the following model:

$$x_k = \Phi_{k-1}x_{k-1} + \Gamma_{k-1}u_{k-1} + \Lambda_{k-1}w_{k-1}$$

$$z_k = h_k(x_k) + n_k$$

where

$$x_i \in \mathbb{R}^n, w_k \sim MN(0, Q_k), n_k \sim MN(0, R_k).$$

We can use following Kalman filtering algorithm to infer x_i:

Algorithm 31: Extended Kalman filter

Input: Given prior on X with $EX = x_0, Cov(X) = P_0$

1 For each new measurement z_i

2 Prediction:

$$\hat{x}_{k|k-1} = \Phi_{k-1}\hat{x}_{k-1} + \Gamma_{k-1}u_{k-1}$$

$$P_{k|k-1} = \Phi_{k-1}P_{k-1}\Phi_{k-1}^T + \Lambda_{k-1}Q_{k-1}\Lambda_{k-1}^T$$

3 Correction:

$$\hat{x}_k = \hat{x}_{k|k-1} + K_k(z_i - H_k\hat{x}_{k|k-1})$$

$$P_k = (P_{k|k-1}^{-1} + H_k^T R_k^{-1} H_k)^{-1}$$

$$K_k = P_k H_k^T R_k^{-1}$$

where $H_i = \frac{\partial h_i}{\partial x}|_{x=\hat{x}_{i-1}}$ (note that update of P_i can be made efficient by using matrix inversion lemma.)

Output: The estimation $\hat{x}_i, i = 1, 2, ..., N$

Remark 17.6.3 (multiple cycles can improve estimation results)**.**

- For nonlinear system, multiple cycles are needed. For example, the second cycle will use the filtered results from first cycle. Every new cycle can improve the result until cannot be improved. Moreover, multiple cycles cannot guarantee global optimal result can be achieved.
- For linear system, one cycle will guarantee optimal results to be achieved.

17.7 Notes on bibliography

For an introduction to dynamical system, see [1]. For an intermediate treatment, see [11][12].

For an excellent mathematical treatment on linear differential equation, see[13][14][2].

For dynamical system (with mathematical theory), see[15][16][3].

For high order linear differential equations, see [17]. For advanced level treatment in ordinary differential equations, see [14].

For additional topics in optimization on least square estimation, see chapter 1 of [9].

Sequential Monte Carlo and particle filtering is a generic numerical approach extending Kalman filters. See [18] for more details.

BIBLIOGRAPHY

1. Luenberger, D. *Introduction to dynamic systems: theory, models, and applications* (Wiley, 1979).
2. Jerri, A. J. *Linear difference equations with discrete transform methods* (Springer Science & Business Media, 2013).
3. Sánchez, D. *Ordinary Differential Equations and Stability Theory: An Introduction* ISBN: 9780486638287 (W. H. Freeman, 1968).
4. Khalil, H. K. *Nonlinear systems, 3rd edition* (Prentice Hall, 2002).
5. Meyer, C. D. *Matrix analysis and applied linear algebra* (Siam, 2000).
6. Chirikjian, G. STOCHASTIC MODELS, INFORMATION THEORY, AND LIE GROUPS, Vol 1.
7. Strogatz, S. H. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (Westview press, 2014).
8. Stengel, R. F. *Optimal control and estimation* (Courier Corporation, 2012).
9. Bertsekas, D. *Nonlinear programming* ISBN: 9781886529007 (Athena Scientific, 2016).
10. Särkkä, S. *Bayesian filtering and smoothing* (Cambridge University Press, 2013).
11. Kailath, T., Sayed, A. H. & Hassibi, B. *Linear estimation* (Prentice Hall Upper Saddle River, NJ, 2000).
12. Crassidis, J. L. & Junkins, J. L. *Optimal estimation of dynamic systems* (CRC press, 2011).
13. Hirsch, M., Devaney, R. & Smale, S. *Differential Equations, Dynamical Systems, and Linear Algebra* ISBN: 9780080873763 (Elsevier Science, 1974).
14. Coddington, E. A. & Levinson, N. *Theory of ordinary differential equations* (Tata McGraw-Hill Education, 1955).
15. Hirsch, M. W., Smale, S. & Devaney, R. L. *Differential equations, dynamical systems, and an introduction to chaos* (Academic press, 2012).
16. Perko, L. *Differential equations and dynamical systems* (Springer Science & Business Media, 2013).
17. Coddington, E. A. *An introduction to ordinary differential equations* ISBN: 9780486659428,0486659429,0134913167,9780134913162 (Dover Publications, 1989).
18. Bruno, M. G. Sequential Monte Carlo methods for nonlinear discrete-time filtering. *Synthesis Lectures on Signal Processing* **6,** 1–99 (2013).