DSA5206 Advanced Topics in Data Science

AY2024/25 Sem2 By Zhao Peiduo

Part 1

A dynamical process is a sequence of states indexed by time:

$$\{x(t) \in \mathcal{X} : t \in \mathcal{T}\},$$

where $\mathcal{T}$ is a set of time indices, which can be subsets of either $\mathbb{Z}$ (discrete) or $\mathbb{R}$ (continuous).
Model of a Dynamical Process: A mathematical description of how the states $x(t)$ depend on time t. Examples include:

- Explicit formula: $x(t) = \sin(t)$
- Differential equation: $\dot{x}(t) = f(t, x(t)) \quad (t \in \mathbb{R})$
- Difference equation: $x(t+1) = g(t, x(t)) \quad (t \in \mathbb{Z})$

Input-Output Systems:

$$\mathbf{x} = \{x(t) : t \in \mathbb{R}\}, \quad \mathbf{y} = F(\mathbf{x}), \quad y(t) = F_t(x), \quad t \in \mathcal{T}$$

Examples:

- Convolutional model: $y(t) = \int_{\mathbb{R}} \rho(t-s)x(s)ds$
- Time-delay model: $y(t) = x(t-\tau)$

First-Principles vs Empirical Models:

- First-principles: derived from physical laws (e.g., Newton's law: $F = G\frac{Mm}{r^2}$)
- Empirical: fit to data, e.g., $z(t) = \sum_{j=1}^{n} a_j e^{i\omega_j t}$

Empirical Model Classification:

- *Non-parametric:* no fixed model form, e.g., $y(t) = \sum_{s=0}^{\infty} \rho(s)x(t-s)$
- *Parametric:* specified structure, e.g.,

$$y(t) = \sum_{s=1}^{n} a(s)y(t-s) + \sum_{r=0}^{m} b(r)x(t-r)$$

- *Black-box:* no interpretation, only predictions
- *Grey-box:* incorporates some physical knowledge

Temporal Index Sets:

- Discrete: $\mathcal{T} \subset \mathbb{Z}$
- Continuous: $\mathcal{T} \subset \mathbb{R}$

$$y(t) = \int_{-\infty}^{t} \rho(t-s)x(s)ds \quad \rightarrow \quad y(t) = \sum_{s \le t-\delta} \rho(t-s\delta)x(s\delta)\delta$$

Linear Functions:

$$H(\alpha x_1 + \beta x_2) = \alpha H(x_1) + \beta H(x_2) \quad \forall \alpha, \beta \in \mathbb{R},\ x_1, x_2 \in \mathcal{X}$$

Linear vs Nonlinear Models:

- Linear differential: $\dot{x}(t) = Ax(t) + b(t)$
- Linear difference: $x(t+1) = Ax(t) + b(t)$
- Linear PDE: $\partial_t u(t,x) = \partial_x^2 u(t,x)$
- Convolution: $y(t) = \int_{\mathbb{R}} \rho(t-s)x(s)ds$

Time-Varying vs Time-Invariant:

$$(T_\tau x)(t) = x(t-\tau)$$

System is time-invariant if: $T_\tau F(x) = F(T_\tau x)$

Time-Invariant Systems

- $y(t) = \int_{-\infty}^{t} e^{-w(t-s)}x(s)\,ds$
  Reason: The kernel $e^{-w(t-s)}$ depends only on the time difference $t-s$.
- $y(t) = x(t-\tau)x(t-\tau)$
  Reason: The output depends only on delayed versions of the input; the delay is constant.
- $y(t) = x(t-\tau) + 3x(t-2\tau)$
  Reason: The output depends only on delayed versions of the input; the delay is constant.
- $y(t) = \alpha y(t-1) + \beta x(t-1)$
  Reason: Recursive equation with fixed delay; time-shifted input leads to time-shifted output.

Time-Variant Systems

- $y(t) = \int_{-\infty}^{t} \sin(ts)x(s)\,ds$
  Reason: The kernel $\sin(ts)$ explicitly depends on t, not only on $t-s$.

Other Model Classifications

- Single-scale vs multi-scale
- Deterministic vs stochastic
- Markovian vs non-Markovian
- Time-domain vs frequency-domain

System Identification

- Given data: $\{(x_i, y_i)\}_{i=1}^{N}$
- Objective: Find a function F such that $y_i \approx F(x_i)$ for all i

Focus: Discrete Linear Time-Invariant (LTI) Systems

- Discrete: $\mathcal{T} \subset \mathbb{Z}$
- Linear: $F(\alpha x_1 + \beta x_2) = \alpha F(x_1) + \beta F(x_2)$
- Time-invariant: $T_\tau F(x) = F(T_\tau x)$

Convolution Model (SISO, Discrete LTI)

$$y(t) = \sum_{s=-\infty}^{\infty} \rho(t-s)x(s) = \sum_{s=-\infty}^{\infty} \rho(s)x(t-s)$$

- $\rho(t)$: impulse response (IR) coefficients
- Interpretation: weighted sum over all past, present, and future values of $x(t)$

Why is ρ called IR Coefficients?

- Impulse input: $x(t) = \delta_{t,0} \Rightarrow y(t) = \rho(t)$

Delta Functions

- Kronecker delta: $\delta_{t,s} = \begin{cases} 1 & t = s \\ 0 & t \ne s \end{cases},\ t, s \in \mathbb{Z}$
- Dirac delta: $\delta(t) = \begin{cases} \infty & t = 0 \\ 0 & t \ne 0 \end{cases},\ t \in \mathbb{R}$

Causality for Discrete LTI Systems

- Causal if $\rho(t) = 0 \quad \forall t < 0$
- Strictly causal if $\rho(t) = 0 \quad \forall t \le 0$

Example Causality Table

- $y(t) = \sum_{s=-\infty}^{t-1} x(s)$: Causal: True, Strictly Causal True
- $y(t) = x(t) - x(t-1)$: Causal True, Strictly Causal False
- $y(t) = x(t+1) - 2x(t) + x(t-1)$: Causal False, Strictly Causal False
- $y(t) = \sum_{s=-\infty}^{t-1} \left(\frac{1}{5}\right)^{t-s} x(s)$: Causal True, Strictly Causal True
- $y(t) = \frac{1}{2}y(t-1) + x(t-1)$: Causal True, Strictly Causal True

Stability – BIBO Stability

- $\|x\| := \sup_{t \in \mathcal{T}} |x(t)|$
- BIBO stable if:

$$\sup_{\|x\| \le 1} \|F(x)\| < \infty$$

BIBO Stability Theorem

- A discrete LTI system is BIBO stable if and only if:

$$\sum_{t \in \mathbb{Z}} |\rho(t)| < \infty$$

Difference Equations

- Difference equations provide a recursive relation between current output and past inputs/outputs.
- General form:

$$y(t) = \mathrm{SomeFunction}[y(t-1), \ldots, y(t-\tau_1); x(t), x(t-1), \ldots, x(t-\tau_2)]$$

Causal LTI Systems as Difference Equations

$$y(t) = \sum_{s=1}^{\tau_1} a(s)y(t-s) + \sum_{s=0}^{\tau_2} b(s)x(t-s)$$

Terminology

- **Order:** τ_1, assuming $a(\tau_1) \neq 0$
- **Delay:** Smallest s such that $b(s) \neq 0$
- **Memory:** τ_2, assuming $b(\tau_2) \neq 0$

Difference Equation as Parametric Convolution

- Example: $y(t) = \alpha y(t-1) + \beta x(t-1),\ y(0) = 0$
- Equivalent convolution:

$$y(t) = \sum_{s=0}^{t-1} \beta\alpha^{t-s-1} x(s), \quad \rho(t) = \beta\alpha^{t-1}$$

Asymptotic Stability

- With $x(t) = 0\, \forall t$, the system:

$$y(t) = \sum_{s=1}^{\tau_1} a(s) y(t-s)$$

  is asymptotically stable if

$$\lim_{t\to\infty} y(t) = 0 \quad \forall y(0) \in \mathbb{R}$$

State-Space Models (SSMs)

- Represent input-output relationships via internal (hidden) states.
- Avoid some limitations of convolution and difference models.
- General form (discrete):

$$\begin{aligned} h(t+1) &= Ah(t) + Bx(t) \\ y(t) &= Ch(t) + Dx(t) \end{aligned}$$

- Dimensions:

$$x(t) \in \mathbb{R}^n, \quad y(t) \in \mathbb{R}^p, \quad h(t) \in \mathbb{R}^m$$
$$A \in \mathbb{R}^{m\times m}, B \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{p\times m}, D \in \mathbb{R}^{p\times n}$$

- Block matrix: $G = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$

Continuous-Time SSM

$$\begin{aligned} \dot{h}(t) &= Ah(t) + Bx(t) \\ y(t) &= Ch(t) + Dx(t) \end{aligned}$$

- Continuous-time version of SSMs with $\mathcal{T} \subset \mathbb{R}$
- Properties: causal, linear, time-invariant

Why Use State-Space Models?

- Handle multi-dimensional inputs and outputs
- Explicitly model internal processes
- Support filtering, control, and identification
- Universality: can approximate any LTI system if $m \to \infty$

Limitations of State-Space Models

- Non-uniqueness of representation
- Performance depends on chosen representation
- Potential structural bias in modeling

Forms of State-Space Models (SSMs)

- SSMs are preserved under similarity transformation:

$$G = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \longrightarrow \quad \tilde{G} = \begin{pmatrix} T^{-1}AT & T^{-1}B \\ CT & D \end{pmatrix}$$

- Canonical forms for SISO systems ($n = p = 1$):
  - Diagonal canonical form: A is diagonal
  - Observer canonical form: $C^\top = (1, 0, 0, \ldots, 0)$
  - Controller canonical form: $B^\top = (1, 0, 0, \ldots, 0)$

Definition (Minimal Realisation)

- A realization (A, B, C, D) is *minimal* if any other realization $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ must satisfy:

$$\dim(\tilde{A}) \geq \dim(A)$$

Theorem (Stability of SSMs)

- An SSM is asymptotically stable if and only if all eigenvalues of A satisfy:

$$|\lambda| < 1$$