# Dynamical systems

Discrete time / maps

$$F : M \to M$$

(exponential notation) $F^t = F \circ F \circ \cdots \circ F \quad t \in \mathbb{Z}^+$

$$F^t = \underbrace{F^{-1} \circ F^{-1} \circ \cdots \circ F^{-1}}_{|t| \text{ times}} \quad t \in \mathbb{Z}^-$$

Continuous time / flows

ODEs

$$\frac{dx}{dt} = v(x, t)$$

$$x \in \mathbb{R}^d$$

$$\frac{d F^t(x)}{dt} = v(F^t(x), t)$$

$$t \in \mathbb{R}^+$$

State: $x$

$M$: state space
phase space

In this class: $M \leq \mathbb{R}^d$

All possible "solutions" / states $\{F^t(x)\}$

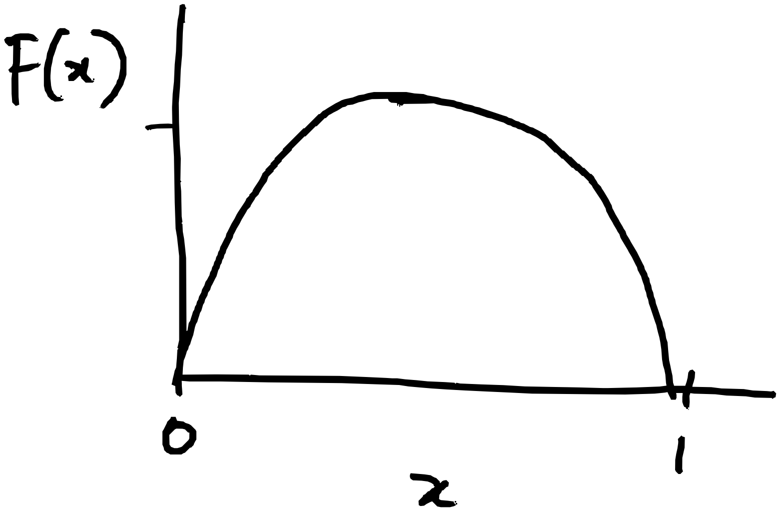
$x \sim \rho$

Questions:

1. How do you reduce dimensionality ?

2. How do you get guarantees for learning dynamics?

$$F(x) = 4x(1-x)$$



$$\text{orbit}(x) := \{F^t(x)\}_{t \in \mathbb{Z}}$$

<u>observable</u>:

$$f : \underset{\substack{\uparrow \\ \text{phase} \\ \text{space}}}{M} \to \mathbb{C}$$

(Scalar functions of state).

Homework 0 : Review

Linear Algebra / functional analysis

$\rightarrow v = \sum_i a_i e_i$ V: vector space

$e_i$ : basis elements

linearly independent vectors

spans

$\rightarrow$ Norm : $\| \cdot \| : V \rightarrow \mathbb{R}^+$

- positive definiteness $\|x\| = 0$ iff $x = 0$
- Triangle inequality:
  $\|x+y\| < \|x\| + \|y\|$
- Absolute homogeneity
  $\|\alpha x\| = |\alpha| \|x\|$
  $\alpha \in \mathbb{C}$.

Ex: $\| \cdot \|$ is a continuous function

$\rightarrow$ Inner product:
$\langle , \rangle : V \rightarrow \mathbb{C}$

(definition)

$\|x\|^2 = |\langle x, x \rangle|$

$\rightarrow$ Separable Hilbert space
$\downarrow$
complete inner product

$\rightarrow$ Operators : $V \rightarrow V$

Matrix : linear operator on finite-dimensional vector spaces

- $A \in \mathbb{R}^{m \times n}$ $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$
  $A_{ij} = \langle A b_i^n , b_j^m \rangle$

- A has m datapoints in $\mathbb{R}^n$.

- $A = \underset{m \times m}{U} \underset{m \times n}{\Sigma} \underset{n \times n}{V^T}$
  $U^* = U^{-1}$ $\Sigma$ : diagonal

- Cols of $V$ $\equiv$ right singular vectors $\equiv$ eigenvectors of $A^TA$

$$\frac{1}{m} \underset{n \times n}{(A^TA)_{ij}} = \frac{1}{m} \sum_{k=1}^{m} x_k^{(i)} x_k^{(j)} \quad \succeq \text{SPSD}$$

$$A = \begin{bmatrix} x_1 \\ \hline x_2 \\ \hline \vdots \\ \hline x_m \end{bmatrix}$$

* QR factorization

Induced norms

* $\|A\| = \sup_{\|x\|=1} \|Ax\|$

* $\ell^2$ norm: $\|x\| := \left( \sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$

Ex: All $\ell^p$ norms in finite-dimensional vector space are equivalent.

$( C_1 \|x\|_p \leq \|x\|_2 \leq C_2 \|x\|_p )$

# Probability

$(\Omega, \Sigma, \mathbb{P})$ $\rightarrow$ measure

sample space

sigma algebra
"all possible events"

$$\mathbb{P} : \Sigma \Rightarrow \mathbb{R}^{+}.$$

- $X_1, X_2, \dots, X_n$ : sequence from $(\Omega, \Sigma)$. For now, functions from $\Omega$ to $\mathbb{R}$.
- Convergence in distribution

$$X_n \xrightarrow{d} X$$

$$\lim_{n \Rightarrow \infty} \mathrm{CDF}(X_n)(x) = \mathrm{CDF}(X)(x)$$

wherever $\mathrm{CDF}(X)$ is continuous.

- Central limit theorem

$$Y_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mathbb{E}X_1)$$

$$Y_n \xrightarrow{d} \mathcal{N}(0,1)$$

- Characteristic function

$$\varphi_X(t) = \mathbb{E}[e^{itX}]$$

Levy's continuity theorem:

- Convergence in distribution $\Longleftrightarrow$ pointwise convergence of $\varphi_{X_n}(t)$ at each $t \in \mathbb{R}$. (to $\varphi_X(t)$).
  - Here, $X_n$ need not be from the same probability space
  - Used in proof of CLT.