$$\gamma(e) = \limsup_{t \to \infty} \frac{1}{t} \log \| A(t) e \|$$

$$\log \det \left(A(t) e_k \right)$$

$$\dot{x}_t = A(\theta_t \omega) x_t$$

$$x_t = e^{\int A(\theta_t \omega)\, dt} x_0$$

Last time: * Oseledets multiplicative ergodic theorem

* Gradient flow

---

* Proof of OMET
  - Flipped lecture
    M/W 26^{th}, 28^{th}

$\rightarrow 26^{th}$ (Flipped)
$\rightarrow 28^{th}$ (regular)
$\Big\}$ sign up poll for time

Gradient flow

$$\frac{d\varphi^t(x)}{dt} = \underset{\text{differential}}{df}(\varphi^t x)$$

$\rightarrow f : M \to \mathbb{R}$

Time-δt maps

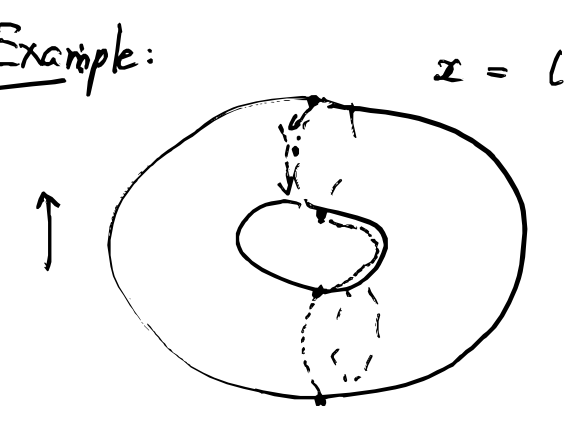
$$F^t(x) = x + \eta \, df(x)$$

η : time-step $\in \mathbb{R}^+$.

Example:

$x = (x_1, x_2, x_3)$

$f(x) = -x_3$



$\rightarrow$ Differential (Differential geometry)

$f : M \to \mathbb{R}$ $\quad f \in C^\infty(M)$

$df : M \to \mathbb{R}^d$ (scientific/engg)

$df_x : T_x M \to \mathbb{R}$

Linear functional on tangent bundle.

$(df)_x \in T_x^* M$ (Dual of tangent bundle : cotangent bundle)

$v \in T_x M$

$$v(f)(x) = \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon v(x)) - f(x)}{\varepsilon}$$

$$= \langle df, v \rangle$$

(Defn applies to Riemannian manifolds)

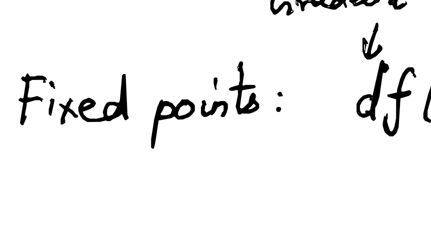
$x = (x_1, x_2, .., x_d)$

$\nabla f_x = (\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_d} f) \in T_x^* M$

$v(f)(x) = df_x \cdot v(x)$, with $df_x \in T_x^* M$, $v(x) \in T_x M$

$v(f)(x)$: Directional derivative of f along v

$$df_x = \underset{v \in T_x M}{\operatorname{argmax}} \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon v) - f(x)}{\varepsilon}$$

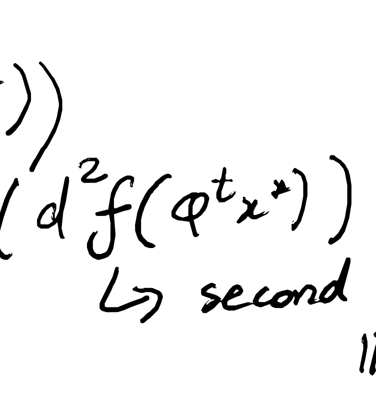


$\mathbb{R}^2$ $\quad f_3 > f_2 > f_1$

$\mathbb{R}^3$

$x = (x_1, x_2, x_3)$

$f(x) = -x_3$



Fixed points: $\underset{\text{Gradient}}{df}(x) = 0$

$$\frac{d\varphi^t(x)}{dt} = v(\varphi^t x) = df(\varphi^t x)$$

$$\operatorname{eig}(dv(\varphi^t x^*)) = \operatorname{eig}(d^2 f(\varphi^t x^*))$$

$\hookrightarrow$ second derivative $\mathbb{R}^{d \times d}$

Example 2 :

Optimization algorithms

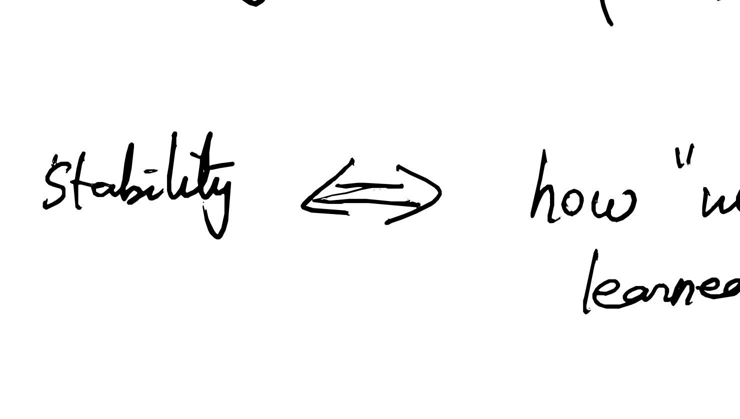
$$f(x) = \frac{1}{n} \sum_{i=1}^n \| y_i - h(z_i; x) \|^2$$

f : Loss function, x : parameter

h : NN $\quad z_i \quad$ x : weights & biases

- $$\frac{d\varphi^t(x)}{dt} = -df(\varphi^t x)$$

- Generalization $\quad$ stability of fixed points



$$d^2 f(\cdot x) = d\left(\frac{2}{n} \sum_{i=1}^n (y_i - h(z_i; x)) \, dh(z_i; x)\right)$$

Stability $\Longleftrightarrow$ how "well" h is learned

Asymptotically, gradient flows converge to fixed points

For any $x \in M$, $\quad X_T := \{\varphi^t(x)\}_{t \geq T}$

$$\lim_{T \to \infty} X_T \subseteq \{\text{critical points of } f\}.$$

Infinite-dimensional gradient flow

$$T^* = \underset{T \in \mathcal{T}}{\operatorname{argmin}} \, C(T_\# \mu, \nu)$$

C : cost; $\mathcal{T}$: class of functions

ν : target measure

μ : known probab dist.

$T_\# \mu = \mu \circ T^{-1}$ $\quad \mu : M_1 \to \mathbb{R}^+$, $\quad \nu : M_2 \to \mathbb{R}^+$

$T : M_1 \to M_2$

$$\frac{d}{dt} \varphi^t(T) = dC(T_\# \mu, \nu)$$

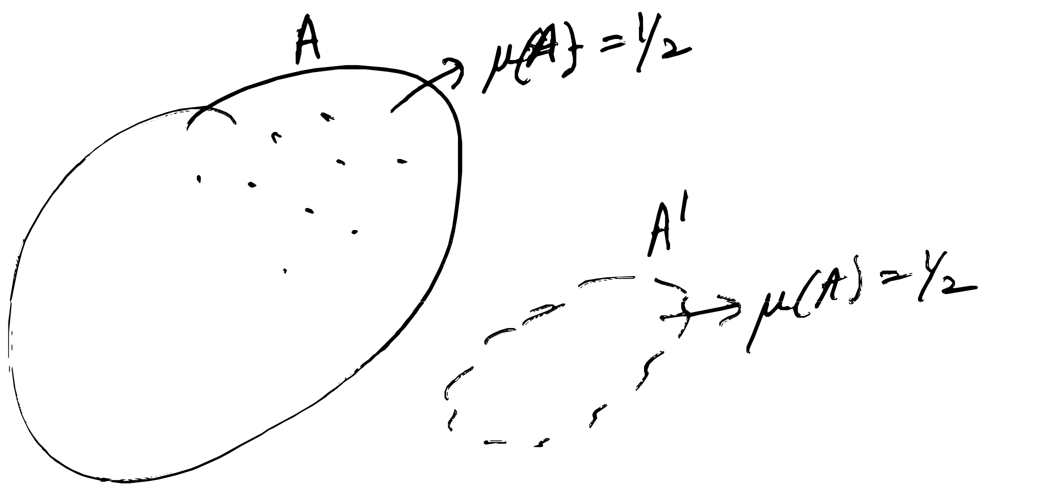
dC : Wasserstein

OMET

→ Ergodic theory : long-time behavior of dynamics ⟺ ensemble/statistical behavior

→ We say that μ is an ergodic distribution for φ if any φ- invariant set has measure 0 or 1.

- A is invariant if $\varphi^{-1}(A) = A$



→ $f \in C^0(M)$, and μ is ergodic for φ, then

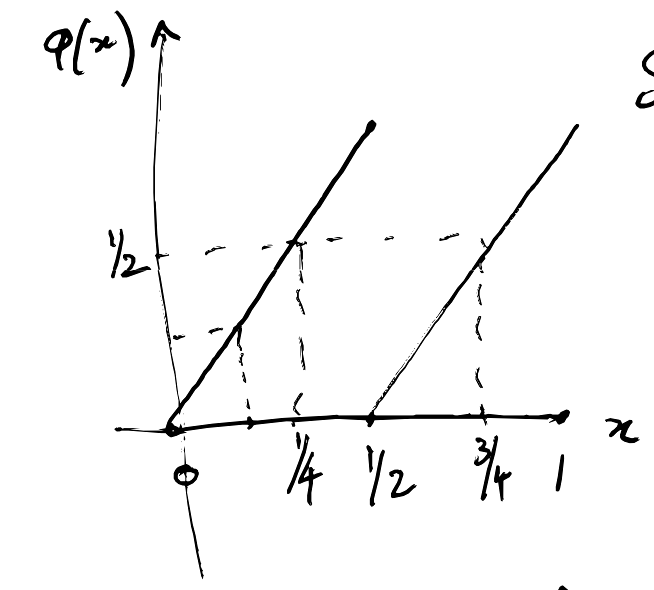
$$\frac{1}{T}\sum_{t=0}^{T-1} f \circ \varphi^t(x) \xrightarrow{T\to\infty} \mathbb{E}_{x\sim\mu} f(x)$$

for μ a.e.

Example

Chaotic system

$\varphi(x) = 2x \,\%\, 1$ $\qquad \mathbb{R}/\mathbb{Z}$.



Sawtooth map

Uniform measure on $[0,1]$.

- $x = \text{rand}()$ → Not Simulating almost every point according to Uniform measure

2^{63} : dyadic

$$\frac{1}{T}\sum_{t\leq T} \varphi^t x \xrightarrow{T\to\infty} \int_0^1 x\,dx = \frac{1}{2}.$$

$$\Big(\mathbb{E}_{x\sim\mu} f(x) = \int f(x)\,d\mu(x)$$
$$\sum_{i=1}^{N} f(x_i)\,\mu(x_i) \Big)$$

→ $\frac{1}{10}$, $\frac{2}{10} = \frac{1}{5}$, $\frac{2}{5}$, $\frac{4}{5}$, $\frac{8}{5} - 1 = \frac{3}{5}$,

$\frac{1}{5}$, $\frac{2}{5}$, $\frac{4}{5}$, $\frac{3}{5}$,

$$\frac{1}{4}\left(\frac{1}{5} + \frac{2}{5} + \frac{4}{5} + \frac{3}{5}\right) = \frac{2}{4}$$