Dynamical System for Optimization (Guest Lecture by Ziang Chen)

Brief intro to stability analysis of dynamical system

Deterministic dynamical system

$$\varphi: T \times X \to X$$
$$(t, x) \mapsto \varphi(t, x)$$

Example: $T = \{0, 1, 2, 3, \dots\}$ $\quad \min f(x)$

$$x_{t+1} = x_t - \alpha \nabla f(x_t)$$
$$g = I - \alpha \nabla f$$
$$\varphi(t, x) = \underbrace{g \circ g \circ \cdots \circ g}_{t\text{-fold}}(x)$$

Linearization near $x_* = 0$: $\nabla f(x) = \nabla f(x) - \nabla f(x_*) \approx \underbrace{H}_{\text{Hessian}}(x - x_*)$

Linearization gives $\Phi(t, x) = A^t x$

Assume $x_* = 0$ $\qquad = (I - \alpha H)^t x$ $\quad$ power t is an integer

Equilibrium

$\varphi(t, \bar{x}) = \bar{x}$, $\forall t \in T$

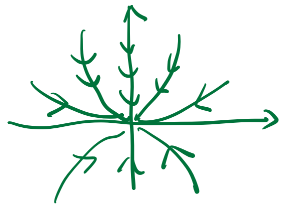
Example: $\bar{x} - \alpha \nabla f(\bar{x}) = \bar{x} \iff \nabla f(\bar{x}) = 0$

Stability near Equilibrium

Example: $x_t = A^t x_0$, $A \in \mathbb{R}^{2\times 2}$, $x_0 \in \mathbb{R}^2$

① Let λ_1, λ_2 be eigenvalues of A

$\lambda_1, \lambda_2 \in \mathbb{R}$ $\quad |\lambda_1| < 1$, $|\lambda_2| < 1$, then stable meaning $x_t \to \bar{x} = 0$ for any $x_0 \neq 0$

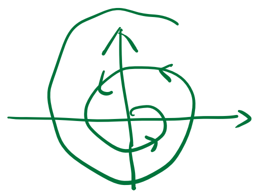


sink

② $\lambda_1, \lambda_2 \in \mathbb{R}$ $|\lambda_1| > 1$, $|\lambda_2| > 1$, source

unstable

③ $\lambda_1, \lambda_2 = re^{\pm i\theta}$, $r < 1$

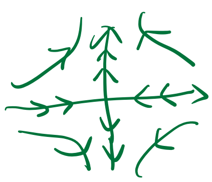


spiral sink

④ $\lambda_1, \lambda_2 = re^{\pm i\theta}$, $r > 1$

spiral source

⑤ $\lambda_1, \lambda_2 \in \mathbb{R}$ $|\lambda_1| < 1$, $|\lambda_2| > 1$



saddle

Reference Micheal Scharb 1987

Global stability of dynamical system

Theorem $g: X \to X$, $\varphi(t, x) = \underbrace{g \circ \cdots \circ g}_{t\text{-fold}}(x) = g^t(x)$

↳ finite dim linear space

If ① $g(0) = 0$, $g(0) = A$

② $X = X_1 \oplus X_2$ (A-invariant decomposition)

$A(X_1) \subseteq X_1 \quad A(X_2) \subseteq X_2$

③ $\|(A|_{X_1})^{-t}\| \leq c\lambda^t$, $\lambda \in (0,1)$

$\|(A|_{X_2})^{t}\| \leq c\lambda^t$, $\forall t \in T = \mathbb{N}$

$\left.\begin{array}{l} A^{-1} \text{ on } X_1 \\ A \text{ on } X_2 \end{array}\right\}$ is a contraction

A on X_1 is an expansion

(Example: $A = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ $X_1 = \mathrm{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$
$X_2 = \mathrm{Span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$)

then Locally near 0, there exists local manifolds M_1, M_2 which are tangent to X_1, X_2 s.t.

① $x \in M_1 \cap \underset{\substack{\downarrow \\ \text{Local ball in } X}}{B} \iff \exists\, x_t \to 0$, s.t. $g^t(x_t) = x$

$x_t = (g^{-1})^t(x) \to 0$

② $x \in M_2 \cap B \iff x_t = g^t(x) \to 0$ ← exponential rate

③ If $x \notin M_2 \cap B$, $\exists t$ s.t. $g^t(x) \notin B$

Remark: ① Only a Local result.

② If $X = X_s \oplus X_c \oplus X_u$

$|\lambda|<1 \quad |\lambda|=1 \quad |\lambda|>1$

$M_s \quad M_c \quad M_u$

For center manifold M_c, then

1) $g^t(X)$ escapes B (subexponential rate)

2) $g^t(X) \to 0$ (subexponential rate)

3) $g^t(X) \in B$, but not converging to 0.

Random Dynamical System

Def Metric Dynamical System

Given probability space $(\Omega, \mathcal{F}, \mathbb{P})$

$\{\theta(t) : \Omega \to \Omega\}$

① $\theta(0) = Id_\Omega$

② $\theta(t+s) = \theta(t) \cdot \theta(s)$

③ $\theta(t)$ is $\mathbb{P}$-preserving

$\mathbb{P}(\theta(t)^{-1} B) = \mathbb{P}(B), \quad \forall B \in \mathcal{F}$

Def (Random Dynamical System) $X = \mathbb{R}^d$

$\varphi : T \times \Omega \times X \to X$

$(t, \omega, x) \mapsto \varphi(t, \omega, x)$

cocycle property

① $\varphi(0, \omega, x) = x$

② $\varphi(t+s, \omega, x) = \varphi(s, \theta(t)\omega, \varphi(t, \omega, x))$

Example: Randomized Coordinate GD

$$X_{t+1} = X_t - \alpha e_{i_t} \partial_{i_t} f(X_t)$$

$$i_t \sim U(\{1, 2, \cdots, d\})$$

$\omega = (i_0, i_1, i_2, \cdots)$

$g(\omega, x) = x - \alpha e_i \partial_i f(x) \quad i = \pi_0(\omega)$

$\theta(t) = \theta^t$