Homework 2

Vladimir Frants, Yunhua Zhao, Mohamed Ben Zid

September 2020

1 Exercise 2.4

1.1 (a)

because

$$x_n(t) = \vartheta^\epsilon(t_n + t)$$

and

$$\vartheta^\epsilon(t) = \theta_{m(t)}$$

so

$$x_n(t) = \theta_{m(t+t_n)}$$

$$x_n(t+s) = \theta_{m(t+s+t_n)}$$

then

$$x_n(t+s) - x_n(t) = \theta_{m(t+s+t_n)} - \theta_{m(t+t_n)}$$

which

$$= \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i G(\theta_i)$$

Because $X_\epsilon(.)$ is piecewise point, $G(X_\epsilon(.))$ is also piecewise constant and its jump times are given by $t_n = \sum_{k=1}^{n} \epsilon_k$. Thus the definite integral on $[t_n+t, t_n+t+s]$ of $G(X_\epsilon(.))$ is a sum that can be approximation expressed as

$$\int_{t_n+t}^{t_n+t+s} G[x_\epsilon(u)]du$$

together

$$\int_{t_n+t}^{t_n+t+s} G[x_\epsilon(u)]du \approx \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i G(\theta_i)$$

$$\int_{t_n+t}^{t_n+t+s} G[x_\epsilon(u)]du = \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i G(\theta_i) + \rho_(\epsilon), (2.1)$$

1.2 (b)

formula

$$x_n(t+s) - x_n(t) = \theta_{m(t+s+t_n)} - \theta_{m(t+t_n)} = \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i G(\theta_i)$$

contains $m(q) - m(r) - 1$ terms. For ϵ sufficiently small, set the ϵ_b is the biggest ϵ and the ϵ_s is the smallest ϵ in interval (r,q) so that the number of terms is bounded by $(\frac{q-r}{\epsilon_b}, \frac{q-r}{\epsilon_s})$. This yields, for small ϵ,

$$\|(x_\epsilon(q) - x_\epsilon(r))\|_\infty = \sum_{i=m(r)}^{m(q)-1} \epsilon_i G(\theta_i)$$

Because G is bounded, let use L to represent G's bounder, so

$$\|x_\epsilon(q) - x_\epsilon(r)\|_\infty = L \sum_{i=m(r)}^{m(q)-1} \epsilon_i =< \epsilon_b L(q-r)/\epsilon_s, (2.2)$$

To summarize, for ϵ sufficiently small, we have shown that for any $\eta > 0$, we may let $\delta_\eta = \frac{\eta}{L(\epsilon_b/\epsilon_s)}/$ so that it follows that $\|x_\epsilon(q) - x_\epsilon(r)\|_\infty =< \eta$ wherever $\|q-r\| =< \delta_\eta(\epsilon_b/\epsilon_s)$. This establishes equicontinuity in the extended sense.

1.3 (c)

Let $a < t$ and $b > t+s$ and consider $x_{\epsilon_k}(.)$ on (a,b). Set $x_n(0) = \theta_0$ for all k. Therefor, for ϵ sufficiently small, by (b) formula 2.2,

$$|x_{\epsilon_k}(r)|_\infty =< |\theta_0|_\infty + rL\frac{\epsilon_b}{\epsilon_s}$$

for all $r > 0$, which suffices to show that x_ϵ is uniformly bounder in (a,b). This together with equicontinuity of x_{ϵ_k} implies by the Ascoli-Arzela Theorem 2.2 that any infinite subsequence of x_{ϵ_k} has a convergent subsequence with a continuous limit on (a, b). Consider a convergent subsequence along $\epsilon_r \to 0$, so that $\hat{x}(.) = \lim_{\epsilon_r \to 0} x_{\epsilon_r}(.)$ (in the sup norm) and continuous. Then

$$\lim_{\epsilon_r \to 0}(x_{\epsilon_r}(t+s) - x_{\epsilon_r}(t)) = \lim_{\epsilon_r \to 0} \int_t^{t+s} G(x_{\epsilon_r}(u))du$$

$$= \int_t^{t+s} \lim_{\epsilon_r \to 0} G(x_{\epsilon_r}(u))du$$

$$= \int_t^{t+s} G(\hat{x}(u))du$$

Where the first formula follows from the fact that $\rho(\epsilon)$ in (2.1) is bounded by $L(\epsilon_b + \epsilon_s)$ and thus of order $\mathcal{O}(\epsilon)$, the second formula follows from Lebesgue Dominated Convergence Theorem, and the third formula is a consequence of the continuity of $G(\hat{x}(.))$ on (a,b). We arrive for $s > 0$ at

$$\frac{\hat{x}(t+s) - \hat{x}(t)}{s} = \frac{1}{s}\int_t^{t+s} G(\hat{x}(u))du$$

By continuity of $G(\hat{x}(.))$, taking the limit as s goes to zero, the above right-hand side converges to $G(\hat{x}(t))$, which establishes the ODE in the question for $\hat{x}(.)$. Because G is continuous and bounded on the trajectory $\hat{x}$, it follows from Theorem 2.1 that the ODE has a unique solution for each initial condition,establishing that all accumulation points have the same limit, proving the claim for the unbiased case.

1.4 (d)

if the bias is not 0, then first:

$$\int_{t_n+t}^{t_n+t+s} G[x_\epsilon(u)]du = \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i G(\theta_i) + \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i \beta_n(\theta_i) + \rho_(\epsilon)$$

using the bound on the perturbations, for any $r < q$,

$$\sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_i \beta_n(\theta_i) =< \sum_{i=m(t_n+t)}^{m(t_n+t+s)-1} \epsilon_b \beta_n(\theta_i) = (q-r)\mathcal{O}(\epsilon_b)$$

so this term can be added to the approximation error $\rho_(\epsilon)$ and the proof follows directly.

2 Excercise 2.6

2.1 a)

Theorem 2.3 Fix T and let $G : \mathbb{R}^d \to \mathbb{R}^d$ be a Lipschitz-contininous function. Consider the (biased version) of recursion (2.8) up to the T.

$$\theta_{n+1} = \theta_n + \epsilon(G(\theta_n) + \beta_\epsilon(\theta_n)), n\epsilon \leq T$$

with constant step size $\epsilon > 0$.

Let $x_\epsilon = \vartheta^\epsilon(t)$, $0 \leq t \leq T$ denote the interpolation process of $\{\theta_n\}$ on $[0, T]$, for $T > 0$ with $x_\epsilon(0) = \theta_0$.

If $\sup_\theta \|\beta_\epsilon(\theta)\| = \mathcal{O}(\epsilon)$, the the ϵ-indexed sequence of process $\{x_\epsilon((t); 0 \leq t \leq T : \epsilon > 0\}$ converges as $\epsilon \to 0$ in the sup norm) to the solution of the ODE:

$$\frac{dx(t)}{dt} = G(x(t)), 0 \leq t \leq T$$

$$\theta_{n+1} = \theta_n - \epsilon J^{'}(\theta_n)$$

Show that as $\epsilon \to 0$ the interpolation process converges to the ODE:

$$\frac{dx(t)}{dt} = \frac{\sin(\alpha_2(x(t))}{v_2} - \frac{\sin(\alpha_1(x(t)))}{v_1}$$

$$G(x(t)) = \frac{\sin(\alpha_2(\theta))}{v_2} - \frac{\sin(\alpha_1(\theta))}{v_1} = -\frac{1}{v_1} \cdot \sin(\alpha_1(\theta)) + \frac{1}{v_2} \cdot \sin(\alpha_2(\theta))$$

$$G' = \frac{dG(x(t))}{dt} = -\frac{d\alpha_1(\theta)}{dt} \cdot \frac{\cos(\alpha_1(\theta))}{v_1} + \frac{d\alpha_2(\theta)}{dt} \cdot \frac{\cos(\alpha_2(\theta))}{v_2}$$

to apply theorem 2.3 we need to prove that $G(x(t))$ is Lipschitz function; then we can apply 2.3 and prove the convergence at $\epsilon \to 0$.

Using the mean value theorem: if G is continuous in closed interval $[a, b]$ and differentiable on the open (a, b) then $\exists c \in (a, b)$ such that $G'(c) = \frac{G(b)-G(a)}{b-a}$ so:

$$\|G(b) - G(a)\| = |G'(c)| \cdot \|b - a\|$$

is the case $G'(x)$ is bounded in (a, b). Then we will have

$$\|G(b) - G(a)\| \leq L\|b - a\|$$

which means that G is Lipschitz function.

Let

$$G(x(t)) = \frac{dx(t)}{dt} = \frac{\sin(\alpha_2(x(t))}{v_2} - \frac{\sin(\alpha_1(x(t))}{v_1}$$

$$\Rightarrow$$

$$G'(x(t)) = \frac{d\alpha_2(\theta)}{dt} \cdot \frac{\cos(\alpha_2(\theta)}{v_2} - \frac{d(\alpha_1(\theta))}{dt} \cdot \frac{\cos(\alpha_1(\theta))}{v_1}$$

Refer to the book page 16:

$$tan(\alpha_1(\theta)) = \frac{\theta}{a}$$

$$tan(\alpha_2(\theta)) = \frac{d - \theta}{b}$$

$$a = -1, b = 1$$

$$\begin{aligned}\frac{d\alpha_1(\theta)}{d\theta} &= \frac{1}{\theta}\sin(\alpha_1(\theta))\cos(\alpha_1(\theta))\\ &= -\frac{1}{\tan(\alpha_1(\theta))}\sin(\alpha_1(\theta))\cos(\alpha_1(\theta))\\ &= -\cos^2(\alpha_1(\theta))\end{aligned}$$

$$\frac{d\alpha_1(\theta)}{d\theta} = -\cos^2(\alpha_1(\theta)) \tag{1}$$

$$\begin{aligned}\frac{d\alpha_2(\theta)}{d\theta} &= -\frac{1}{d-\theta}\sin(\alpha_2(\theta))\cos(\alpha_2(\theta))\\ &= -\frac{1}{tan(\alpha_2(\theta))}\sin(\alpha_2(\theta))\cos(\alpha_2(\theta))\\ &= -\cos^2(\alpha_2(\theta))\end{aligned}$$

$$\frac{d\alpha_2(\theta)}{d\theta} = -\cos^2(\alpha_2(\theta)) \tag{2}$$

substitute **1** and **2**:

$$\begin{aligned}G'(x(t)) &= -\frac{d\alpha_1(\theta)}{dt}\cdot\frac{\cos(\alpha_1(\theta))}{v_1} + \frac{d\alpha_2(\theta)}{dt}\cdot\frac{\cos\alpha_2(\theta)}{v_2}\\ &= \cos^2(\alpha_1(\theta))\cdot\frac{\cos(\alpha_1(\theta)}{v_1} - \cos^2(\alpha_2(\theta))\cdot\frac{\cos(\alpha_2(\theta))}{v_2}\\ &= \frac{\cos^3(\alpha_1(\theta))}{v_1} - \frac{\cos^3(\alpha_2(\theta))}{v_2}\\ &\le |\frac{\cos^3(\alpha_1(\theta))}{v1}| + |\frac{\cos^3(\alpha_2(\theta))}{v_2}|\\ &\le |\frac{1}{v_1}| + |\frac{1}{v_2}|\end{aligned}$$

Let $L = |\frac{1}{v_1}| + |\frac{1}{v_2}|$

then $G'(x(t))$ is bounded by $L = |\frac{1}{v_1}| + |\frac{1}{v_2}|$ since $G(x(t))$ is continuous and differentiable, then $G(x(t))$ is Lipschitz. Applying theorem 2.3 to $G(x(t))$ then as $\epsilon \to 0$ the interpolation process converge to the ODE.

2.2 b)

Show that θ^* is stable and argue that a solution $x(t)$ of the above ODE must then satisfy $\lim_{t\to\infty} x(t) = \theta^*$

$$G(x(t)) = \frac{\sin(\alpha_2(\theta))}{v_2} - \frac{\sin(\alpha_1(\theta)}{v_1}$$

$$G(x(t)) = -\frac{\sin(\alpha_1(\theta))}{v_1} + \frac{\sin(\alpha_2(\theta))}{v_2}$$

constant $v_2 < v_1$ then:

$$G(x(t)) \leq |-\frac{\sin(\alpha_1(\theta))}{v_2}| + |\frac{\sin(\alpha_2(\theta))}{v_2}|$$

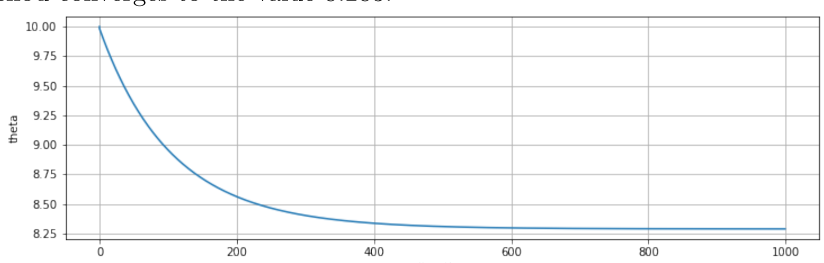
$$G(x(t)) \leq \frac{2}{v_2} \Rightarrow G(x(t))$$

is bounded we can apply theorem 2.6.

Any accumulation point θ^* is an asymptotically stable point θ^* is an asymptotically stable point of $G(x(t))$ so then $G(\theta^*) = 0$ then $\lim_{t\to\infty}(x(t)) = \theta^*$

2.3 c)

Here is our simulation for $a = 2$, $b = 5$, $d = 10$, $v_1 = 3$, $v_2 = 1$, and $\epsilon = 0.05$. We directly programmed the gradient search method and used $\theta_0 = 10.0$. The method converges to the value 8.283.



We use this python script to generate the plot:

Listing 1: Resolve contention

```
import math
import matplotlib.pyplot as plt
%matplotlib inline

a = 2.0
b = 5.0
d = 10.0
v1 = 3.0
v2 = 1.0
epsilon = 0.05
```

```
theta0 = 0.0

def J_div(theta):
  return (1/v1)*(theta/math.sqrt(theta**2 + a**2)) - \
  (1/v2)*(d-theta)/math.sqrt((d-theta)**2 + b**2)


theta = theta0
thetas = [theta0]
iter_nums = [0]
for i in range(1000):
  theta = theta - epsilon*J_div(theta)
  thetas.append(theta)
  iter_nums.append(i)

plt.figure(figsize=(12, 8))
plt.subplot(2, 1, 1)
plt.plot(iter_nums, thetas)
plt.ylabel('theta')
plt.xlabel('iteration')
plt.grid()
```