# Research Student's Research Diary – Contents

# September 20

## 1 Citation Test

$$\mathbb{E}[x] + \nabla \boldsymbol{f} + \mathbb{R}^d.$$

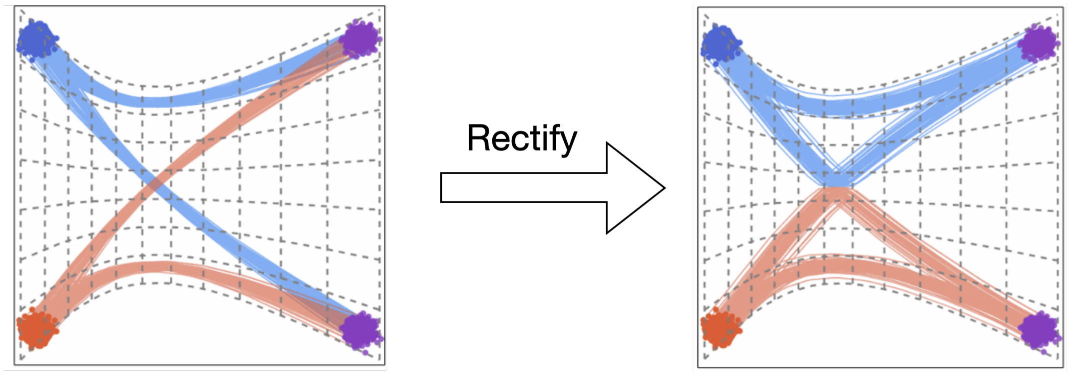
$$\boldsymbol{\theta} = \arg\min L(\theta, \mathcal{D}).$$

We have a citation [2] here. Solve this problem:

# September 20

## 2 Rectified Flow

We have a figure here.




We have another citation here [1].

# September 21

## 3 Momentum + Cautious WD

$$\dot{x} = -m - \mathbb{I}(mx \geq 0)x \tag{1}$$
$$\dot{m} = \nabla f(x) - m. \tag{2}$$

$$H(x, m) = f(x) + \frac{1}{2}m^2 + (mx)_+$$

Let us use $s \in [0, 1]$ to denote the subgradient of $(mx)_+$ within $H$. Let $p \in [0, 1]$ be the Flippiov variable of $\mathbb{I}(mx \geq 0)$ inside the velocity field. Note that $s$ and $p$ are different. We need to keep track of it in the derivation.

$$\begin{aligned}\dot{H}(s, p) &= g\dot{x} + m\dot{m} + s(m\dot{x} + x\dot{m}) \\ &= -g(m + px) + m(g - m) + s(-m(m + px) + x(g - m)) \\ &= -pxg - m^2 + s(-m^2 + xg - (1 + p)xm)\end{aligned}$$

If $\alpha$ belongs to $\dot{H}$, we want that for each $s$, there exists $p$, such that

$$-pxg - m^2 + s(-m^2 + xg - (1 + p)xm) = -\alpha.$$

So we have

$$p(-xg - sxm) - m^2 + s(-m^2 + xg - xm) = \alpha.$$
$$p = \frac{\alpha - (1 + s)m^2 + s(xg - xm)}{xg + sxm}.$$

At the non-smooth point, we have $xm = 0$, and it gives

$$p = \frac{\alpha - (1 + s)m^2 + s(xg)}{xg}.$$

So $p$ is in range of

$$\frac{\alpha - 2m^2 + xg}{xg}, \quad \frac{\alpha - m^2}{xg}.$$

If $xg > 0$, we have

$$m^2 \leq \alpha \leq m^2 + xg$$
$$\alpha \leq 2m^2 \leq \alpha + xg$$

So

$$2m^2 - xg \leq \alpha \leq 2m^2$$

$$\cap_s \cup_p \dot{H}(s, p).$$

For know the flippov system, the set valued Lie-derivation is defined as:

$$\mathcal{L}_v H = \{\alpha \colon \exists p, \quad s.t. \quad \dot{H}(s, p) = \alpha \quad \forall s \in \partial H\}.$$

So we want to find the $p$ such that the $\dot{H}(s,p)$ is independent with the choice of $s$.

So to make it independent on $s$, we want

$$-m^2 + xg - (1+p)xm = 0.$$

This gives

$$\begin{aligned}\dot{H} &= -p(m^2 + (1+p)xm) - m^2\\ &= -(1+p)m^2 - p(1+p)xm.\end{aligned}$$

If $xm < 0$, we have $p = 0$, and hence $\dot{H} = -m^2$. Invariant set is $\{xm < 0, \quad m = 0\}$.

If $xm > 0$, we have $p = 1$, and hence, $\dot{H} = -(2m^2 + 2(xm)_+)$. Invariant set is $\{xm = 0, \quad m = 0\}$.

If $xm = 0$, we have $\dot{H} = -(1+p)m^2$. Invariant set is $\{xm = 0, m = 0\}$.

Hence, the invariant set is included in

$$\{xm \leq 0, \quad m = 0\}.$$

But since $m = 0$, we have $xm \leq 0$ anyway.

## 4 General Fillipov From Scratch

**Theorem 1.** Assume we have non-smooth function $V(x)$ with clark generalized gradient, and let $d$ be an update direction. We have

**Proof.**

$$\lim_{\epsilon \to 0^+} \frac{V(x+\epsilon d) - V(x)}{\epsilon} = ...$$

□

**Theorem 4.1** (Great Theorem)**.** *Yes, it is.*

*Proof.* Yes. □

**Theorem 4.2** (Great Theorem)**.** *This is the second.*

*Proof.* This is the best. □

# September 21

## 5 Momentum + Cautious WD

$$\dot{x} + \mathbb{R}^d + \nabla f + \mathbb{E}[x].$$

$$\widehat{f}$$

$$\dot{x} = -m - \mathbb{I}(mx \geq 0)x \tag{3}$$
$$\dot{m} = \nabla f(x) - m. \tag{4}$$

$$H(x, m) = f(x) + \frac{1}{2}m^2 + (mx)_+$$

Let us use $s \in [0, 1]$ to denote the subgradient of $(mx)_+$ within $H$. Let $p \in [0, 1]$ be the Flippiov variable of $\mathbb{I}(mx \geq 0)$ inside the velocity field. Note that $s$ and $p$ are different. We need to keep track of it in the derivation.

$$\begin{aligned}
\dot{H}(s, p) &= g\dot{x} + m\dot{m} + s(m\dot{x} + x\dot{m}) \\
&= -g(m + px) + m(g - m) + s(-m(m + px) + x(g - m)) \\
&= -pxg - m^2 + s(-m^2 + xg - (1 + p)xm)
\end{aligned}$$

If $\alpha$ belongs to $\dot{H}$, we want that for each $s$, there exists $p$, such that

$$-pxg - m^2 + s(-m^2 + xg - (1 + p)xm) = -\alpha.$$

So we have

$$p(-xg - sxm) - m^2 + s(-m^2 + xg - xm) = \alpha.$$

$$p = \frac{\alpha - (1 + s)m^2 + s(xg - xm)}{xg + sxm}.$$

At the non-smooth point, we have $xm = 0$, and it gives

$$p = \frac{\alpha - (1 + s)m^2 + s(xg)}{xg}.$$

So $p$ is in range of

$$\frac{\alpha - 2m^2 + xg}{xg}, \qquad \frac{\alpha - m^2}{xg}.$$

If $xg > 0$, we have

$$m^2 \leq \alpha \leq m^2 + xg$$

$$\alpha \leq 2m^2 \leq \alpha + xg$$

So

$$2m^2 - xg \leq \alpha \leq 2m^2$$

$$\cap_s \cup_p \dot{H}(s,p).$$

For know the flippov system, the set valued Lie-derivation is defined as:

$$\mathcal{L}_v H = \{\alpha \colon \exists p, \quad s.t. \quad \dot{H}(s,p) = \alpha \quad \forall s \in \partial H\}.$$

So we want to find the $p$ such that the $\dot{H}(s,p)$ is independent with the choice of $s$.

So to make it independent on $s$, we want

$$-m^2 + xg - (1+p)xm = 0.$$

This gives

$$\begin{aligned}\dot{H} &= -p(m^2 + (1+p)xm) - m^2 \\ &= -(1+p)m^2 - p(1+p)xm.\end{aligned}$$

If $xm < 0$, we have $p = 0$, and hence $\dot{H} = -m^2$. Invariant set is $\{xm < 0, \quad m = 0\}$.

If $xm > 0$, we have $p = 1$, and hence, $\dot{H} = -(2m^2 + 2(xm)_+)$. Invariant set is $\{xm = 0, \quad m = 0\}$.

If $xm = 0$, we have $\dot{H} = -(1+p)m^2$. Invariant set is $\{xm = 0, m = 0\}$.

Hence, the invariant set is included in

$$\{xm \leq 0, \quad m = 0\}.$$

But since $m = 0$, we have $xm \leq 0$ anyway.

## 6 General Fillipov From Scratch

**Theorem 2.** Assume we have non-smooth function $V(x)$ with clark generalized gradient, and let $d$ be an update direction. We have

**Proof.**

$$\lim_{\epsilon \to 0^+} \frac{V(x + \epsilon d) - V(x)}{\epsilon} = \ldots$$

□

**Theorem 6.1** (Great Theorem)**.** *Yes, it is.*

*Proof.* Yes. □

**Theorem 6.2** (Great Theorem)**.** *This is the second.*

*Proof.* This is the best. □

# September 21

## 7 Momentum + Cautious WD

$$\dot{x} = -m - \mathbb{I}(mx \geq 0)x \tag{5}$$
$$\dot{m} = \nabla f(x) - m. \tag{6}$$

$$H(x, m) = f(x) + \frac{1}{2}m^2 + (mx)_+$$

Let us use $s \in [0, 1]$ to denote the subgradient of $(mx)_+$ within $H$. Let $p \in [0, 1]$ be the Flippiov variable of $\mathbb{I}(mx \geq 0)$ inside the velocity field. Note that $s$ and $p$ are different. We need to keep track of it in the derivation.

$$\begin{aligned}\dot{H}(s, p) &= g\dot{x} + m\dot{m} + s(m\dot{x} + x\dot{m}) \\ &= -g(m + px) + m(g - m) + s(-m(m + px) + x(g - m)) \\ &= -pxg - m^2 + s(-m^2 + xg - (1 + p)xm)\end{aligned}$$

If $\alpha$ belongs to $\dot{H}$, we want that for each $s$, there exists $p$, such that

$$-pxg - m^2 + s(-m^2 + xg - (1 + p)xm) = -\alpha.$$

So we have

$$p(-xg - sxm) - m^2 + s(-m^2 + xg - xm) = \alpha.$$
$$p = \frac{\alpha - (1 + s)m^2 + s(xg - xm)}{xg + sxm}.$$

At the non-smooth point, we have $xm = 0$, and it gives

$$p = \frac{\alpha - (1 + s)m^2 + s(xg)}{xg}.$$

So $p$ is in range of

$$\frac{\alpha - 2m^2 + xg}{xg}, \quad \frac{\alpha - m^2}{xg}.$$

If $xg > 0$, we have

$$m^2 \leq \alpha \leq m^2 + xg$$
$$\alpha \leq 2m^2 \leq \alpha + xg$$

So

$$2m^2 - xg \leq \alpha \leq 2m^2$$

$$\cap_s \cup_p \dot{H}(s, p).$$

For know the flippov system, the set valued Lie-derivation is defined as:

$$\mathcal{L}_v H = \{\alpha \colon \exists p, \quad s.t. \quad \dot{H}(s, p) = \alpha \quad \forall s \in \partial H\}.$$

So we want to find the $p$ such that the $\dot{H}(s,p)$ is independent with the choice of $s$.

So to make it independent on $s$, we want

$$-m^2 + xg - (1+p)xm = 0.$$

This gives

$$\begin{aligned}\dot{H} &= -p(m^2 + (1+p)xm) - m^2 \\ &= -(1+p)m^2 - p(1+p)xm.\end{aligned}$$

If $xm < 0$, we have $p = 0$, and hence $\dot{H} = -m^2$. Invariant set is $\{xm < 0, \quad m = 0\}$.

If $xm > 0$, we have $p = 1$, and hence, $\dot{H} = -(2m^2 + 2(xm)_+)$. Invariant set is $\{xm = 0, \quad m = 0\}$.

If $xm = 0$, we have $\dot{H} = -(1+p)m^2$. Invariant set is $\{xm = 0, m = 0\}$.

Hence, the invariant set is included in

$$\{xm \leq 0, \quad m = 0\}.$$

But since $m = 0$, we have $xm \leq 0$ anyway.

## 8 General Fillipov From Scratch

**Theorem 3.** Assume we have non-smooth function $V(x)$ with clark generalized gradient, and let $d$ be an update direction. We have

**Proof.**

$$\lim_{\epsilon \to 0^+} \frac{V(x + \epsilon d) - V(x)}{\epsilon} = ...$$

□

**Theorem 8.1** (Great Theorem)**.** *Yes, it is.*

*Proof.* Yes. □

**Theorem 8.2** (Great Theorem)**.** *This is the second.*

*Proof.* This is the best. □

# September 21

## 9 Today is a New Day

Dealing with math, code, people, self.

# September 22

## 10 Comprehensive LaTeX Command Test

This document tests all custom commands defined in `diary_commands.sty` to verify the automatic command parsing system.

### 10.1 Mathematical Sets and Spaces

$$\mathbb{R}^d, \mathbb{C}^n, \mathbb{N}, \mathbb{P}, \mathbb{E}[X], \mathbb{D}$$

$$\mathcal{P}(\mathcal{A}), \mathcal{H}, \mathcal{X}, \mathcal{B}, \mathcal{F}, \mathbb{M}$$

### 10.2 Statistical Functions

$$\operatorname{var}(X), \operatorname{Var}(X), \operatorname{cov}(X,Y), \operatorname{Cov}(X,Y)$$

$$\operatorname{corr}(X,Y), \operatorname{Corr}(X,Y), \Pr(A), \Pr(B)$$

$$X \sim \mathcal{N}(\mu, \sigma^2), \mathrm{MSE}, \mathrm{KL}(P \parallel Q)$$

### 10.3 Math Operators

$$x^* = \arg\max_{x \in \mathbb{R}} f(x)$$

$$y^* = \arg\min_{y \in \mathbb{C}} g(y)$$

$$\hat{\mu} = \operatorname{med}\{x_1, \ldots, x_n\}$$

$$\hat{x} = \operatorname{mod}\{x_1, \ldots, x_n\}$$

### 10.4 Vectors and Bold Symbols

$$\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\beta}$$

$$\mathbf{A}, \mathbf{B}, \boldsymbol{x}, \boldsymbol{z}$$

### 10.5 Fractions and Derivatives

$$\frac{a+b}{c+d} = \frac{a+b}{c+d}$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}$$

### 10.6 Special Symbols

$$a := b + c$$

### 10.7 Calculus and Analysis

$$\nabla f(x), \mathbb{I}(A), \mathrm{trace}(A), \mathrm{diag}(A)$$

### 10.8 Color Commands (for text)

Some red text, blue text, green text, magenta text, gray text.

TODO: This is a todo item

### 10.9 Complex Mathematical Expression

Combining multiple commands:

$$\mathbb{E}\left[\frac{\boldsymbol{x}^T \mathbf{A} \boldsymbol{x}}{\mathrm{trace}(\mathbf{B})}\right] = \arg\max_{\boldsymbol{\theta} \in \mathbb{R}^d} \Pr\left(\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})\right)$$

Where we use the derivative:

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \nabla L(\boldsymbol{\theta})$$

And the indicator function:

$$\mathbb{I}\left(\boldsymbol{x} \in \mathcal{X}\right) \cdot \mathrm{var}(Y)$$

### 10.10 Delimited Commands Test

Testing the `\bb...\ee` commands:

$$\boldsymbol{x} = \arg\max_{\boldsymbol{w}} \mathbb{E}\left[\frac{\boldsymbol{w}^T \boldsymbol{x}}{\mathrm{trace}(\mathbf{K})}\right]$$
$$\text{subject to } \boldsymbol{w} \in \mathbb{R}^d$$

# September 22

## 11 Testing Custom Commands

Testing `\dd` and `\ind` commands:

$$\dot{m} = \nabla f(x) - m$$

$$\mathbb{I}(mx \geq 0) = \mathbb{I}(mx \geq 0)$$

Testing `\argmax` and `\argmin`:

$$x^* = \arg\max_{x \in \mathbb{R}} f(x)$$

$$y^* = \arg\min_{y \in \mathbb{C}} g(y)$$

Testing `\bb...\ee` delimited commands:

$$\begin{aligned} x + y &= z \\ a + b &= c \end{aligned}$$

Testing `\bba...\eea`:

$$\sum_{i=1}^{n} x_i = 0 \qquad \text{eq:test} \quad (7)$$

Testing `\vv{x}` and `\V{y}`:

$$\boldsymbol{x} = \mathbf{y} + \boldsymbol{x}$$

# September 22

## 12 Today

$$\mathrm{Cov}(x), \mathrm{cov}(x), \mathrm{Pr}(x), \mathcal{N}(x).$$

It is the best.

## References

[1] Li, X. L., Thickstun, J., Gulrajani, I., Liang, P., and Hashimoto, T. B. (2022). Diffusion-lm improves controllable text generation. *arXiv preprint arXiv:2205.14217*.

[2] Smith, J. and Doe, J. (2024). *Research Methods in Computer Science*. Academic Press.