Intro to Data Processing and Representation

Ron Benchetrit - 312167554

Nitzan Leshem - 208989400

Theory

1.a. The optimal $\hat{f}_p$ for:

$p=2$: $\hat{f}_2 = \frac{1}{N}\sum_{i=1}^{N} \hat{f}_i$, $\hat{f}_i = N \int_{x\in I_i} f(x)\,dx$

$p=1$: $\hat{f}_1 = \sum_{i=1}^{N} \hat{f}_i \mathbb{1}_{I_i}(x)$, $\hat{f}_i$ s.t. $\exists s$: $\int_{\substack{f(x)<\hat{f}_i \\ x\in I_i}} \mathbb{1}_{I_i}\,dx = \int_{\substack{f(x)>\hat{f}_i \\ x\in I_i}} \mathbb{1}_{I_i}\,dx + s\cdot\int_{\substack{f(x)=\hat{f}_i \\ x\in I_i}} \mathbb{1}_{I_i}\,dx$

1.b. W-MSE: $\varepsilon^2(f,\hat{f}) = \int_0^1 \left(f(x)-\hat{f}(x)\right)^2 w(x)\,dx$ = linearity of the integral

$$= \sum_{i=1}^{N} \int_{x\in I_i} \left(f(x)-\hat{f}_i(x)\right)^2 w(x)\,dx$$

Convex with respect to $\hat{f}_i$. Solve for each i

$$\frac{\partial \varepsilon^2(f,\hat{f})}{\partial \hat{f}_i} = -2\int_{x\in I_i}\left(f(x)-\hat{f}_i(x)\right) w(x)\,dx = -2\left[\int_{x\in I_i} f(x)w(x)\,dx - \int_{x\in I_i}\hat{f}_i(x)w(x)\,dx\right] = 0$$

$$\Rightarrow \hat{f}_i(x)\int_{x\in I_i} w(x)\,dx = \int_{x\in I_i} f(x)w(x)\,dx$$

$$\Rightarrow \hat{f}_i^{*}(x) = \frac{\int_{x\in I_i} f(x)w(x)\,dx}{\int_{x\in I_i} w(x)\,dx}$$

Theory

1

c) finding the optimal $\hat{f}_p$ when $p=1$ and $w>0$ is general:

objective: $\underset{\hat{f}}{\arg\min}\, \mathcal{E}^1(f,\hat{f}) = \underset{\hat{f}}{\arg\min} \int_0^1 |f(x)-\hat{f}(x)|w(x)dx$

- first I'll rewrite the objective using the linearity of the Integral:

$$\int_0^1 |f(x)-\hat{f}(x)|w(x)dx = \int_{I_1} |f(x)-\underset{\hat{f}_{I_1}}{\hat{f}(x)}|w(x)dx + \int_{I_2} |f(x)-\underset{\hat{f}_{I_2}}{\hat{f}(x)}|w(x)dx + \dots + \int_{I_n} |f(x)-\hat{f}(x)|w(x)dx$$

. we will define $\int_{I_i} |f(x)-\underset{\hat{f}_{I_i}}{\hat{f}(x)}|w(x)dx = \underline{\mathcal{E}^1_{I_i}(f,\hat{f})}$

and therefor: $\int_0^1 |f(x)-\hat{f}_i(x)|w(x)dx = \sum_{I_i} \mathcal{E}^1_{I_i}(f,\hat{f})$

now we will find the optimal $\hat{f}_{I_i}$ for $\mathcal{E}^1_{I_i}(f,\hat{f})$:

$$\frac{\partial \mathcal{E}^1(f,\hat{f})}{\partial \hat{f}_{I_i}} = \left(\mathcal{E}^1_I(f,\hat{f})\right)' = \left(\int_{I_i} |f(x)-\hat{f}(x)|w(x)dx\right)'$$

$$= \left(\int_{\hat{f}_i > f(x)}^{I_i} (f(x)-\hat{f}(x))w(x)dx - \int_{\hat{f}_i < f(x)}^{I_i} (\hat{f}(x)-f(x))w(x)dx \right)'$$

$$= \int_{\hat{f}_i > f(x)}^{I_i} w(x)dx - \int_{\hat{f}_i < f(x)}^{I_i} w(x)dx = 0 \Rightarrow \hat{f}$$

looking for minimum

$$\Rightarrow \hat{f}_{opt}(x) = \underbrace{\sum_{i=1}^{N}}_{\text{sum}} \hat{f}_i \cdot \mathbb{1}_{I_i}(x) \text{ s.t } \int_{\hat{f}_i > f(x)}^{I_i} w(x)dx = \int_{\hat{f}_i < f(x)}^{I_i} w(x)dx$$

Theory

1) d) objective: find $\mathcal{E}_i^p(f_i,\hat{f}_i)$ s.t. $\mathcal{E}^p(f,\hat{f}) = \sum_{i=1}^{N} \mathcal{E}_i^p(f,\hat{f}_i)$

$\Rightarrow$ $\mathcal{E}^p(f,\hat{f}) = \int_0^1 |f(x)-\hat{f}(x)|^p w(x)\,dx =$

$=$ (using the linearity of the Integral) $= \sum_{I_i} \int_{I_i} |f(x)-\hat{f}(x)|^p w(x)\,dx$

• we will chose $\mathcal{E}_i^p(f,\hat{f}_i) = \int_{I_i} |f(x)-\hat{f}(x)|^p w(x)\,dx$

as we can see $\mathcal{E}_i^p(f,\hat{f}_i)$ is an independent optimization problem for each i, where $\hat{f}_i$ is restricted to the interval I_i

Additionally, since Min $\mathcal{E}^p(f,\hat{f}_i)$ is trying to minimze the sum of independent functions, it is equivillant to minimizing each of the problems ~~of E~~ by itself

1.e.

i. $$|f_i(x)-\hat{f}_i(x)|^p = w_{f_i,\hat{f}_i}(x)\,(f_i(x)-\hat{f}_i(x))^2$$

$$\Rightarrow \overset{*}{} \quad w_{f_i,\hat{f}_i}(x) = \frac{|f_i(x)-\hat{f}_i(x)|^p}{(f_i(x)-\hat{f}_i(x))^2} = |f_i(x)-\hat{f}_i(x)|^{p-2}$$

* because $f_i(x) \neq \hat{f}_i(x)$ $\forall x \in I_i$ then $(f_i(x)-\hat{f}_i(x))^2 \neq 0$

ii. $$\min_{\hat{f}_i} \varepsilon^p(f_i,\hat{f}_i) = \min_{\hat{f}_i} \int_{x\in I_i} |f_i(x)-\hat{f}_i(x)|^p\, w(x)\,dx =$$

$$= \min_{\hat{f}_i} \int_{x\in I_i} (f_i(x)-\hat{f}_i(x))^2\, w'_{f_i,\hat{f}_i}(x)\, w(x)\,dx$$

iii. Because if $w'_{f_i,\hat{f}_i}$ was independent of $\hat{f}_i$ we could have achieved the optimal solution in the same manner as in clause b.
The problem is with differentiation of $w_{f_i,\hat{f}_i}(x)$.

iv. when we remove the previous assumption we get that the formula is not well defined for $p=1$ where $f_i(x)=\hat{f}_i(x)$. Also, when $\hat{f}_i(x) \to f_i(x)$ we get that $w_{f_i,\hat{f}_i}(x) \to \infty$ which might lead to non-negligible numerical errors.

v. pseudo-code:
Input: f_i, $\hat{f}_i$
~~output:~~
step:

$$f^i_{next} \leftarrow \frac{1}{\int_{I_i} w(x)\,dx} \int_{I_i} f(x)\,w(x)\,dx$$

$$w^i_{next} \leftarrow \min\left\{|f(x)-\hat{f}^i_{next}(x)|^{p-2}, \frac{1}{\varepsilon}\right\}$$

Scanned with CamScanner

f. Input: f, N, p

Initialization:

$w \leftarrow$ positive values (usually 1's)

Loop while stopping condition is not met:

for each i in the domain:

$$\hat{f}^i_{next} \leftarrow \frac{1}{\int_{x \in I_i} w(x)dx} \int_{x \in I_i} f(x)w(x)\,dx$$

$$w^i_{next} \leftarrow \min\left\{|f(x) - f^i_{next}(x)|^{p-2}, \frac{1}{\varepsilon}\right\}$$

$\hat{f} \leftarrow \{f^i_{next} \mid i \text{ in the domain}\}$

$w \leftarrow \{w^i_{next} \mid i \text{ in the domain}\}$

9.9. Iterative Reweighted Least Squares:

2. a

$$\int_{t\in\Delta_i} (t-t_i)^k \, dt = \frac{(t-t_i)^{k+1}}{k+1}\Bigg|_{\frac{i-1}{N}}^{\frac{i}{N}} = \quad \left(t_i = \frac{\frac{i-1}{N} + \frac{i}{N}}{2} = \frac{i-\frac{1}{2}}{N} \right)$$

$$= \frac{1}{k+1}\left[\left(\frac{i}{N} - \left(\frac{i-\frac{1}{2}}{N}\right)\right)^{k+1} - \left(\frac{i-1}{N} - \left(\frac{i-\frac{1}{2}}{N}\right)\right)^{k+1}\right] =$$

$$= \frac{1}{k+1}\left(\frac{1}{N}\right)^{k+1}\left[\left(\frac{1}{2}\right)^{k+1} - \left(-\frac{1}{2}\right)^{k+1}\right] =$$

$$= \frac{\left(\frac{1}{N}\right)^{k+1}}{\cancel{0^k}\,(k+1)}\left[\left(\frac{1}{2}\right)^{k+1} - \left(-\frac{1}{2}\right)^{k+1}\right] = \quad |\Delta_i| = \frac{1}{N}$$

* if k is odd than $k+1$ is even and $\left(-\frac{1}{2}\right)^{k+1} = \frac{1}{2}^{k+1}$

$$= \frac{|\Delta_i|^{k+1}}{(k+1)}\left[\left(\frac{1}{2}\right)^{k+1} - \left(\frac{1}{2}\right)^{k+1}\right] = 0$$

** if k is ~~odd~~ even then $k+1$ is odd and $\left(-\frac{1}{2}\right)^{k+1} = -\left(\frac{1}{2}\right)^{k+1}$

$$= \frac{|\Delta_i|^{k+1}}{k+1}\left[\left(\frac{1}{2}\right)^{k+1} + \left(\frac{1}{2}\right)^{k+1}\right] = \frac{|\Delta_i|^{k+1}}{2^k (k+1)}$$

In conclusion,

$$\int_{t\in\Delta_i} (t-t_i)^k \, dt = \begin{cases} 0, & \text{if } k \text{ is odd} \\ \frac{|\Delta|^{k+1}}{2^k (k+1)}, & \text{if } k \text{ is even} \end{cases}$$

2]

b] we would like to Solve:

$$\underset{c_i, a_i}{\arg\min} \int_{\Delta_i} (\phi(t) - a_i(t-t_i) - c_i)^2 dt = \mathcal{E}_i(c_i, a_i)$$

first we'll find $a_{i_{opt}} = a_i^*$:

$$\cdot \quad \frac{\partial \mathcal{E}}{\partial a_i} = -2\int_{\Delta_i} (\phi(t) - a_i(t-t_i) - c_i)(t-t_i)\,dt = 0 \qquad / \cdot -\tfrac{1}{2}$$

$$\Rightarrow \int_{\Delta_i} \phi(t)\cdot(t-t_i)\,dt - a_i\int_{\Delta_i}(t-t_i)^2 dt - c_i \int_{\Delta_i}(t-t_i)^1 dt = 0$$

$$\overset{(a)}{\Rightarrow} \int_{\Delta_i} \phi(t)\cdot(t-t_i)\,dt - a_i \cdot \frac{|\Delta_i|^3}{2^2\cdot(2+1)} = 0$$

$$\Rightarrow \boxed{a_i^* = \frac{12}{|\Delta_i|^3} \cdot \int_{\Delta_i} \phi(t)(t-t_i)\,dt}$$

Now we'll find $c_{i_{opt}} = c_i^*$:

$$\frac{\partial \mathcal{E}}{\partial c_i} = -2\int_{\Delta_i} (\phi(t) - a_i(t-t_i) - c_i)\,dt = 0 \qquad / \cdot -\tfrac{1}{2}$$

$$\Rightarrow \int_{\Delta_i} \phi(t)\,dt - \cancel{\int_{\Delta_i} a_i(t-t_i)\,dt} - \int_{\Delta_i} c_i\,dt = 0$$

$$\overset{(a)}{\Rightarrow} \boxed{c_i^* = \frac{1}{|\Delta_i|} \cdot \int_{\Delta_i} \phi(t)\,dt}$$

2]

c)

$$MSE^*_{Linear} = \sum_{i=1}^{N} MSE^*_{i,Linear} = \sum_{i=1}^{N} \int_{\Delta_i} (\phi(t) - a_i^*(t-t_i) - c_i^*)^2 dt$$

$$\circledast\ MSE_{i,Linear} = \int_{\Delta_i} (\phi^2(t) - 2\phi(t)\cdot a_i^*(t-t_i) - 2c_i^*\cdot\phi(t) + 2a_i^*(t-t_i)c_i^* + (a_i^*(t-t_i))^2 + c_i^{*2}) dt$$

$$= \int_{\Delta_i} \phi^2(t)dt - 2\int \phi(t)\cdot a_i^*(t-t_i)dt - \int 2c_i\phi(t)dt + 2a_ic_i\underbrace{\int(t-t_i)dt}_{\overset{(a)}{=}0} + \underbrace{\int a_i^2(t-t_i)^2dt}_{\overset{(a)}{=}a_i\cdot\frac{\Delta_i^3}{2^2 3}} + \underbrace{\int c_i^2 dt}_{=c_i^2|\Delta_i|}$$

$$= \int_{\Delta_i} \phi^2(t)dt - 2a_i\int\phi(t)(t-t_i)dt - 2c_i\int\phi(t)dt + a_i^2\cdot\frac{|\Delta_i|^3}{12} + c_i^2|\Delta_i|$$

$$= \int_{\Delta}\phi^2(t)dt \underbrace{- 2\cdot\frac{12}{|\Delta_i|^3}\left(\int\phi(t)(t-t_i)dt\right)^2}_{\mathrm{II}} \underbrace{- 2\cdot\frac{1}{|\Delta_i|}\cdot\left(\int\phi(t)dt\right)^2}_{\mathrm{I}} + \underbrace{\frac{12}{|\Delta_i|^3}\left(\int\phi(t)(t-t_i)dt\right)^2}_{\mathrm{II}} + \underbrace{\frac{1}{|\Delta_i|}\cdot\left(\int\phi(t)dt\right)^2}_{\mathrm{I}}$$

$$= \int_{\Delta_i}\phi^2(t) - \frac{12}{|\Delta_i|^3}\cdot\left(\int_{\Delta_i}\phi(t)(t-t_i)dt\right)^2 - \frac{1}{|\Delta_i|}\left(\int\phi(t)dt\right)^2$$

$$\Rightarrow MSE_{Linear} = \sum_{i=1}^{N}\left(\int_{\Delta_i}\phi^2(t) - \frac{12}{|\Delta_i|^3}\cdot\left(\int_{\Delta_i}\phi(t)(t-t_i)dt\right)^2 - \frac{1}{|\Delta_i|}\cdot\left(\int_{\Delta}\phi(t)dt\right)^2\right)$$

$$= \boxed{\int_0^1\phi^2(t)dt - \sum_{i=1}^{N}\frac{12}{|\Delta_i|^3}\cdot\left(\int_{\Delta_i}\phi(t)(t-t_i)dt\right)^2 - \sum_{i=1}^{N}\frac{1}{|\Delta_i|}\cdot\left(\int_{\Delta_i}\phi(t)dt\right)^2}$$

2.d.

As seen in class. The MSE for using piecewise-constant approximation is:

$$MSE_{const} = \int_0^1 \phi^2(t)\,dt - \frac{1}{N}\sum_{i=1}^{N}(\hat{\phi}_i^*)^2 = \int_0^1 \phi^2(t)\,dt - \sum_{i=1}^{N}\left(\frac{1}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2\right)$$

As seen in part c the MSE for using piecewise-linear approximation is:

$$MSE_{linear} = \int_0^1 \phi^2(t)\,dt - \sum_{i=1}^{N}\frac{12}{|\Delta_i|^3}\left(\int_{\Delta_i}\phi(t)(t-t_i)\right)^2 - \sum_{i=1}^{N}\frac{1}{|\Delta_i|}\left(\int_{\Delta_i}\phi(t)\,dt\right)^2$$

$$MSE_{const} - MSE_{linear} = \underbrace{\sum_{i=1}^{N}\frac{12}{|\Delta_i|^3}\left(\int_{\Delta_i}\phi(t)(t-t_i)\right)^2}_{\geq 0} \geq 0$$

In conclusion we get that

$$MSE_{const} \geq MSE_{linear}$$

Therefore, The MSE for piecewise-linear approximation is lower (or equal) to the MSE for piecewise-constant approximation