Modern Methods of Decision Making

Solution of Homework 1 Bayan Hendawi

1. $M_X(\theta) = E(e^{\theta X}) = E(e^{\theta X_1 + \theta X_2 + \dots + \theta X_n}) = E(e^{\theta X_1} e^{\theta X_2} \dots e^{\theta X_n})$

$= E(e^{\theta X_1}) E(e^{\theta X_2}) \dots E(e^{\theta X_n}) = E^n(e^{\theta X_1})$ *

now let us put our result* in $\Lambda_X(\theta)$

$\Lambda_X(\theta) = \log M_X(\theta) = \log E^n(e^{\theta X_1}) = n \log E(e^{\theta X_1}) = n \log M_{X_1}($

$\Rightarrow$

$\Lambda^*_X(t) = \sup_{\theta > 0}\{\theta t - \Lambda_X(\theta)\} = \sup_{\theta > 0}\{\theta t - n\Lambda_{X_1}(\theta)\} = n \sup\{\theta \frac{t}{n} -$

$= n \Lambda^*_{X_1}(\frac{t}{n})$

2. (1.) let us use Cramer - Chernoff inequality

$$P(X > t) \leq e^{-\Lambda^*_X(t)}$$

now, we will calculate $\Lambda^*_\xi(t)$ and $M_\xi(\theta)$

- $M_\xi(\theta) = E(e^{\theta X}) = (1-p)e^{\theta \cdot 0} + pe^{\theta \cdot 1} = 1 - p + pe^\theta$
- $\Lambda_\xi(\theta) = \log M_\xi(\theta) = \log(1 - p + pe^\theta)$

$$\Lambda^*_\xi(t) = \sup_{\theta > 0}\{\theta t - \Lambda_\xi(\theta)\}$$

by $f'(\theta)$ let us find critical points of $f(\theta) = \theta t - \Lambda_\xi(\theta) \to$ su

$$f'(\theta) = t - \frac{pe^\theta}{1 - p + pe^\theta} = 0$$

$$pe^{\theta} = t(1-p+pe^{\theta})$$

$$pe^{\theta} = t - tp + tpe^{\theta}$$

$$pe^{\theta} - tpe^{\theta} = t - tp$$

$$e^{\theta} = \frac{t-tp}{p-tp}$$

$$e^{\theta} = \frac{t(1-p)}{p(1-t)}$$

$$\theta^* = \log\frac{t}{p} + \log\frac{1-p}{1-t}$$

now.

$$\Lambda_{\varepsilon}^*(t) = f(\theta^*) = t\log\frac{t}{p} + t\log\frac{1-p}{1-t} - \log\left(1-p+p\cdot\frac{t(1-p)}{p(1-t)}\right)$$

$$= t\log\frac{t}{p} + t\log\frac{1-p}{1-t} - \log\frac{1-p}{1-t}$$

$$= t\log\frac{t}{p} + (1-t)\log\frac{1-t}{1-p}$$

$$= h_p(t)$$

Then,

$$P(\varepsilon - p > t) = P(\varepsilon > t+p) \leq e^{-\Lambda_{\varepsilon}(t+p)} = e^{-h_p(t+p)}$$

② From task 1 we can deduce that

$$\Lambda_{\varepsilon}^*(t) = n\Lambda^*_{\varepsilon_1}\left(\frac{t}{n}\right)$$

$$\Lambda_{\varepsilon_1}^*(t) = h_p(t)$$

$$\Lambda_{\varepsilon}^*(t) = nh_p\left(\frac{t}{n}\right)$$

By Cramér-Chernoff inequality we get

$$P(\xi - np > t) = P(\xi > t + np) \leqslant e^{-\Lambda^*_\xi(t+np)} = e^{-n h_p(\frac{t}{n}+p)}$$

$\boxed{3}$ ① let us say

$$\vec{m} = \frac{\vec{y} - \vec{x}}{\|\vec{y} - \vec{x}\|} \quad , \quad q = \|\vec{y} - \vec{x}\|$$

in the inequality $|f(\vec{y}) - f(\vec{x})| \leqslant L\|\vec{y} - \vec{x}\|$ "we assume to be 2-nor

we can rewrite it as:

[redacted] $|f(\vec{x} + q\vec{m}) - f(\vec{x})| \leqslant L.q$

$$\frac{|f(\vec{x} + q\vec{m}) - f(\vec{x})|}{q} \leqslant L$$

$$\lim_{q \to 0} \frac{|f(\vec{x} + q\vec{m}) - f(\vec{x})|}{q} \leqslant L$$

$$\frac{\partial f(\vec{x})}{\partial \vec{m}} \leqslant L$$

$$(\nabla f(\vec{x}), \vec{m}) \leqslant L$$

Since y can be arbitary, we can consider this case:

$$\vec{m} = \frac{\nabla f(\vec{x})}{\|\nabla f(\vec{x})\|}$$

$$\Rightarrow \frac{(\nabla f(\vec{x}), \nabla f(\vec{x}))}{\|\nabla f(\vec{x})\|} \leqslant L \Rightarrow \|\nabla f(\vec{x})\| \leqslant L$$

② Let us define $g(t) = f(x+t(y-x))$

depends on Fundemental Theorem of Calculus we got

$$g(1) = g(0) + \int_0^1 g'(t)\,dt$$

$$f(y) = f(x) + \int_0^1 (\nabla f(x+t(y-x)))^T (y-x)\,dt =$$

so that.

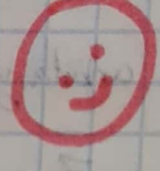
$$|f(y) - f(x) - \nabla f(x)^T (y-x)| = \int_0^1 ((\nabla f(x+t(y-x)) - \nabla f(x))^T (y-x)\,dt$$

$$\leq \int_0^1 |((\nabla f(x+t(y-x)) - \nabla f(x))^T (y-x)|\,dt$$

By Cauchy-Schwarz inequality

$$\int_0^1 |((\nabla f(x+t(y-x)) - \nabla f(x))^T (y-x)|\,dt \leq \|((\nabla f(x+t(y-x)) - \nabla f(x)\|$$

$$\leq \int_0^1 tL\|y-x\|^2\,dt = \frac{\beta}{2}\|y-x\|^2 \quad \square$$



4 we Can obtain

$$f_\alpha(x) - f_\alpha(x^*) \leq \nabla f_\alpha(x)^T (x - x^*)$$

also we have

$$f_\alpha(x) = f(x) - \frac{\alpha}{2}\|x\|^2$$

now we can rewrite inequality as,

$$f(x) - f(x^*) - \frac{\alpha}{2}\|x\|^2 + \frac{\alpha}{2}\|x^*\|^2 \leq \nabla\left(f(x) - \frac{\alpha}{2}\|x\|^2\right)^T (x - x^*)$$

$$f(x) - f(x^*) \leq \nabla\left(f(x) - \frac{\alpha}{2}\|x\|^2\right)^T (x - x^*) + \frac{\alpha}{2}\left(\|x\|^2 - \|x^*\|^2\right)$$

$$= \nabla f(x)^T (x - x^*) + \frac{\alpha}{2}\left(\|x\|^2 - \|x^*\|^2 - \underbrace{\nabla\|x\|^2 (x - x^*)}_{-2x^T x + 2x^T x^*}\right)$$

$$= \nabla f(x)^T (x - x^*) + \frac{\alpha}{2}\left(\|x\|^2 - \|x^*\|^2 - 2\|x\|^2 + 2\|x^*\|^2\right)$$

$$= \nabla f(x)^T (x - x^*) + \frac{\alpha}{2}\left(-\|x\|^2 + \|x^*\|^2\right)$$

$$\leq |\nabla f(x)^T (x - x^*)| - \frac{\alpha}{2}\underbrace{\|x^* - x\|^2}_{r}$$

$$\leq \sup_{r \in R}\left(-\frac{\alpha}{2} r^2 + \|\nabla f(\vec{x})\| r\right)$$

now, let us solve

$$-\alpha r + \|\nabla f(\vec{x})\| = 0$$

$$r^* = \frac{\|\nabla f(\vec{x})\|}{\alpha}$$

$$\sup\left(-\frac{\alpha}{2} r^2 + \|\nabla f(\vec{x})\| r\right) = -\frac{\alpha}{2} \cdot \frac{\|\nabla f(\vec{x})\|^2}{\alpha^2} + \|\nabla f(\vec{x})\| \cdot \frac{\|\nabla f(\vec{x})\|}{\alpha}$$

$$= \frac{\|\nabla f(\vec{x})\|^2}{\alpha} - \frac{\|\nabla f(\vec{x})\|^2}{2\alpha} = \frac{1}{2\alpha}\|\nabla f(\vec{x})\|^2$$

$$f(\vec{x}) - f(\vec{x}_*) \leq \sup_{r \in R}\left(-\frac{\alpha}{2} r^2 + \|\nabla f(\vec{x})\| r\right) = \frac{1}{2\alpha}\|\nabla f(\vec{x})\|^2 \quad \square$$

B.H 26.2.2021