به نام خوا على قاسم زاده ۲۳۹ ۱۵۱۱ه

ا- دارسمرار،

$$f(y|g,\theta) = \mathcal{N}(y,\theta)$$

$$\Rightarrow L(0) = \prod_{i=1}^{n} \frac{1}{\int Y_{\pi} \theta \, \alpha_{i}} \exp\left(-\frac{(Y_{i} - Y_{9}i)^{T}}{Y_{\theta} \eta_{i}}\right) \xrightarrow{Ln}$$

$$\ln l(0) = -\frac{1}{r} \left(\sum_{i=1}^{n} \ln(r_{\pi}\theta n_{i}) + \frac{(y_{i}-r_{n})^{r}}{\theta n_{i}} \right) \xrightarrow{2p_{0}}$$

$$-\frac{1}{r}\left(\frac{\sum_{i=1}^{n}\frac{1}{\theta}-\frac{(y_i-y_{ni})^r}{\theta^r y_{i}}\right)=0$$

$$\frac{n}{\theta} - \frac{1}{\theta} \sum_{i=1}^{n} \frac{(y_{i} - y_{n_{i}})^{r}}{y_{i}} = 0 \longrightarrow n\theta = \sum_{i=1}^{n} \frac{(y_{i} - y_{n_{i}})^{r}}{y_{i}}$$

$$E\{\{f(n) - E\{f(n)\}\}^r\} + E\{\{f(n) - E\{f(n)\}\}^r\} + E\{\{f(n) - E\{f(n)\}\}^r\}$$

$$F(x) = F(x) + F(x)$$

عمض عارت الا بعنى ترم ما ما مه تر تر عاد أورد بلر مثال إنا الحال الحرف المراح الحرف الحرف المرت و الم

=> Err(x) = Bias (f(n)) + Var (f(n)) + 6

٢-٢. حالا طرسمير،

Yi= Po+ PiXi+ Ei

 $D = \{(\chi_i, \gamma_i), i=1, \longrightarrow, n\}$

 $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i = 0$ $\frac{1}{n} \sum_{i=1}^{n} X_i^T = S_X^T$

 $\hat{f}_i(n) = \frac{1}{n} \sum_{i=1}^n y_i$

fr(n) = Bo + Bon

 $\mathbb{E}\left\{\hat{f}_{i}(y_{i})\right\} = \left\{\frac{1}{n}\sum_{i=1}^{n}y_{i}\right\} = \beta_{0} + \beta_{1}\overline{X} = \beta_{0}$

=> Bids (fi(n)) = Bo + Big - Bo = Big

 $Var(\hat{f}(n)) = Var(\bar{\gamma}) = \frac{\epsilon^r}{n} \rightarrow Err(x) = \beta_1^r n_1^r + \epsilon^r$

in the six is

 $\hat{f}_{\gamma}(x) = \hat{\beta}_{0} + \hat{\beta}_{1}x$, $(\hat{\beta}_{0}, \hat{\beta}_{1}) = \text{digmin} \sum_{i=1}^{n} (y_{i} - \beta_{0} - \beta_{1}x_{i})^{T}$ $\bar{R} = \hat{R}$

~> E{\hat{\beta}_{\beta}}=\beta_{\beta}, E{\hat{\beta}_{\beta}}=\beta_{\beta}, \overline{\beta}_{\beta}, \beta_{\beta})=0

 $\operatorname{Var}(\hat{\beta}_i) = \frac{\delta^T}{n}$, $\operatorname{Var}(\hat{\beta}_i) = \frac{\delta^T}{n s_i^T}$

م الله معلی می المی دیا دی طرد کی درعوض واریان کمتری (ایم) دارد.
کی عمل الم بیسیده تر امرت و بایاس نمارد ولی درعوض واریان سیستری طردند بالا

$$(\hat{\beta}^{\hat{\beta}}, \hat{\beta}^{\hat{\beta}}) = \text{digmin} \sum_{i} (Y_{i} - \beta_{o} - \beta_{i} \times i)^{T} + \beta_{i}^{T} \longrightarrow \beta_{o}, \beta_{i}$$

) هشتی لیری و برابر حمور گذارشتی داریمراتر ۱

$$\hat{\beta}^{\lambda} = \bar{y}$$
, $\hat{\beta}^{\lambda}_{i} = \frac{n s_{x}^{r}}{n s_{x}^{r} + \lambda} \hat{\beta}^{ols}$

م عام في الم عرسون وم يعنى دم ايم الله الرس.

=>
$$\mathbb{E}\left\{\hat{f}_{\eta}(y_1)\right\} = \mathcal{B}_{04} + \frac{\eta s_{\chi}^r}{\eta s_{\chi}^r + \lambda} \mathcal{B}_{1} y_1 \rightarrow bias = \mathcal{B}_{1}^r y_1^r + \frac{\eta^r}{(\eta s_{\chi}^r + \lambda)^r}$$

ridge-ols, Empire oute, 1 7-0 AT Jo

م زيد مرون لر بايس را بسيمتر مسكن ولي طرواس را لممتر مي كذه .

Yi= sin(x91i) + &i , 91i - unif([0,1])

Gen Erry =
$$\mathbb{E}_{x,D,\varepsilon} \left\{ (y-\bar{y})^{r} \right\} = \mathbb{E}_{x}^{\varepsilon} \left(\sin(xx) - \mu^{\varepsilon} \right)$$

 $4 \sqrt{\alpha} r(\bar{y}) + 6^{-r}$

=> @ GenErr =
$$(\frac{1}{7} - \frac{5}{\pi r}) + \frac{1}{n}(\frac{1}{7} - \frac{5}{\pi r} + 6^{-1}) + 6^{-1}$$

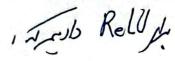
Misspecification est var irreducible noise bias

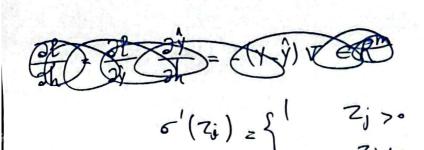
$$\frac{\partial \ell}{\partial c} = \frac{\partial \ell}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial c} = -(y - \hat{y}) \cdot 1 = \sqrt{h} + c - y$$

$$\frac{\partial \ell}{\partial v} = \frac{\partial \ell}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v} = -(y - \hat{y}) \cdot h \in \mathbb{R}^{m}$$

$$\frac{\partial \ell}{\partial v} = \frac{\partial \ell}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v} = -(y - \hat{y}) \cdot h \in \mathbb{R}^{m}$$

$$\frac{\partial \ell}{\partial h} = \frac{\partial \ell}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial h} = -(y - \hat{y}) \cdot \nabla \in \mathbb{R}^{m}$$





$$\int \frac{\partial \ell}{\partial b} = \frac{\partial \ell}{\partial z} = -(Y - \hat{Y}) \vee O \int \{ \sqrt{1 + b} > 0 \}$$

تور مقال مقاصر ميم منه احتمال دينك ن² صفر بارج صفره.

ازطری می تواسیر هر subgradient سراستاه کنیر.

is co dead, il cir in I has I negative Make RelU certisi

وباعث sparsity بستر وeneralization بستر در الله الراسي ها مى مشور

۲- مظرمانی سمال در روس کاهنگ ترامال.

ا بارس المرساد ما عبارت الرساد ،

 $\nabla_{\beta} (\gamma - \chi \beta)^{T} (\gamma - \chi \beta) = 0 \rightarrow \chi^{T} \chi \beta = \chi^{T} \gamma \rightarrow$

 $\beta = (X^T X)^{-1} X^T Y \longrightarrow$ ا مُرْفِي طرول مر بورل XTX

عالا در روش م عش ارادیان طربیرانه

 $B^{(t+1)} = B^{(t)} - \eta \nabla \| y - x \beta^{(t)} \|^2 = B^{(t)} - \eta \chi T (\chi \beta^{(t)} - \gamma) = (I - \eta \chi T \chi) \beta^{(t)}$

$$\beta^{(t)} = (I - \eta A)^{t} \beta^{(0)} + \eta \sum_{K=1}^{t-1} (I - \eta A)^{K} \chi T y$$

$$\chi T \chi \chi T \chi T$$

الله الله العدومال مؤنه بالدواحة الله الله $\frac{\left(1-\left(1-r\eta\lambda_{j}\right)^{t}\right)}{\lambda_{j}} = \frac{1}{\lambda_{+}\lambda_{j}} \longrightarrow \lambda_{+} \frac{\lambda_{j}}{1-\left(1-r\eta\lambda_{j}\right)^{t}}$ ilisably the second of ridge restimated of pt sources the in we - wer batch-size, (sals) ig i (bles), jet ste die Bladain ! المامير عالم به مع عاديد به وابع ساده تر بابذ ، وي ماده تر بابذ ، وي ماده تر بابذ ، وي ماده تر بابذ ، از ماری از ماری موفق کنیز ایس و اکموزی را میلی متوفق کنیز عاد بروز . is to be sidge regression to so to to state overfit / sur · Trus - a $\hat{\theta} = (1 - \frac{C}{\|y\|^r}) y \rightarrow \hat{\theta} - \theta = (y - \theta) - \frac{C}{\|y\|^r} y \Longrightarrow$ 110-011 = 14-011+ Cr 11/11- - 40 (4-0) TY 15 $-\theta \|' = \mathbb{E} \| \mathbb{E} \|' = d$ $\mathbb{E} \{ (y-\theta)^T f(y) \} = \mathbb{E} \{ \text{div } f(y) \}$ stein identity E{110-011'3 = 1E{11E11'} = d P(Y) = 1/11 - Daliv y = 2 3/1 (1/11) = 1/11

$$\Rightarrow \mathbb{E}\left\{\frac{(y-0)^{T}Y}{\|y\|^{T}}\right\} = \mathbb{E}\left\{\frac{d-Y}{\|y\|^{T}}\right\}$$

$$\Rightarrow \mathbb{E}\left\{\|\theta-\theta\|^{T}\right\} = d-1c \mathbb{E}\left\{\frac{d-Y}{\|y\|^{T}}\right\} + C^{T} \mathbb{E}\left\{\frac{1}{\|y\|^{T}}\right\} = d+(c^{T}-Yc(d-T))\mathbb{E}\left\{\frac{1}{\|y\|^{T}}\right\}$$

$$= \frac{1}{2} \frac{d}{d} + \frac{1}{2} \frac{d}{d} + \frac{1}{2} \frac{1}{$$