

سؤال ۱ (۷ نمره)

صحیح یا غلط بودن گزاره‌های زیر را مشخص کرده و دلیل آن را کامل شرح دهید.

(الف) استفاده از روش‌های منظم سازی به فرار از بهینه‌های محلی در طول آموزش منجر می‌شود. (۱ نمره)

(ب) اگر y بردار مقادیر پیش‌بینی شده در یک مدل رگرسیون خطی^۱ و t بردار مقادیر هدف باشند، آنگاه کمینه خطای جمع مربعات، در واقع متناظر با بردار y ای در فضای ویژگی‌ها است که نزدیکترین به t باشد. در این حالت y تصویر t در فضای ویژگی‌ها است. (۱ نمره)

(ج) در رگرسیون خطی، یک مدل خیلی ساده معمولاً مشکل زیربرازش^۲ و واریانس کم وجود دارد. (۱ نمره)

(د) در تقسیم داده‌ها زیر مجموعه‌های ارزیابی^۳، آموزش و آزمایش داده‌های زیر مجموعه‌های ارزیابی و آزمایش می‌توانند از توزیع‌های مختلفی باشند. (۱ نمره)

(ه) برای تقسیم کردن دیتاست موجود به زیر مجموعه‌های ارزیابی، آموزش و آزمایش نمونه برداری تصادفی از دیتاست به اندازه لازم برای هر زیر مجموعه همواره کفايت می‌کند. (۱ نمره)

(و) در یک مسئله رگرسیون خطی برای داده‌هایی که در آنها مقادیر پیش‌بینی شده و واقعی با یکدیگر برابر می‌شوند استفاده از تابع هزینه MSE بهتر از MAE است. (۱ نمره)

(ز) در روش اعتبارسنجی متقابل K برابر^۴ پس از مشخص شدن پیچیدگی مدل نتایج مدل با بهترین دقت برای گزارش نهایی استفاده می‌شود. (۱ نمره)

(الف) این گزاره غلط است. استفاده از روش‌های منظم سازی (Regularization) هدف آن کاهش اورفیتینگ (Overfitting) و افزایش عمومیت مدل است، نه فرار از بهینه‌های محلی. بهینه‌های محلی به مقادیر کمترین خطای نزدیکترین نقطه‌ها در فضای جستجوی ضرایب مدل اشاره دارند. در حالی که استفاده از روش‌های منظم سازی بهبود عملکرد مدل را در طول آموزش از طریق کاهش اورفیتینگ و کنترل پارامترهای مدل بهینه‌سازی می‌کند. بنابراین، استفاده از روش‌های منظم سازی به منظور فرار از بهینه‌های محلی نیست.

(ب) صحیح است فضایی که توسط ستون‌های X پوشش داده می‌شود هر بردار در این فضا به صورت $W^*X = u$ نشان داده می‌شود و به آن $span(X)$ می‌گویند u بهینه برداری است که برابر نگاشت y در $span(X)$ باشد.

(ج) این گزاره درست است. در رگرسیون خطی، مدل خیلی ساده معمولاً با مشکل زیربرازش (underfitting) و واریانس کم (low variance) مواجه می‌شود. زیربرازش به معنای عدم توانایی مدل در یادگیری روابط پیچیده‌تر و ناهمخوانی با داده‌های آموزش است. در واقع، مدل خیلی ساده قادر نیست رفتار دقیق و پیچیده داده‌ها را توصیف کند و بخشی از اطلاعات موجود در داده‌ها را از دست می‌دهد. نشان دهد.

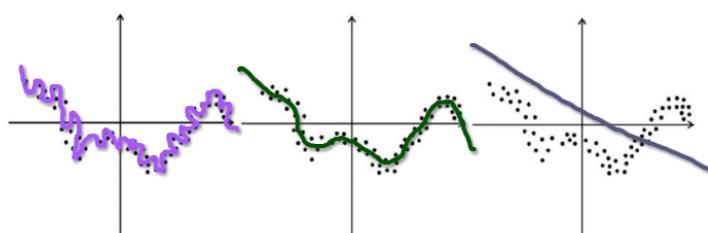
د) به، این امکان وجود دارد که داده های اعتبارسنجی و داده های آزمون دارای توزیع های متفاوتی باشند. در برخی موارد، فرآیند تقسیم داده ها به مجموعه های آموزشی، اعتبارسنجی و آزمون ممکن است منجر به تفاوت هایی در توزیع اعتبارسنجی و داده های آزمون شود. این می تواند به دلیل عوامل مختلفی مانند نمونه گیری تصادفی در طول فرآیند تقسیم داده ها یا تفاوت های ذاتی در ویژگی های داده ها رخدید. به عنوان مثال، اگر داده ها از منابع مختلف یا در زمان های مختلف جمع آوری شده باشند، ممکن است توزیع ها متفاوت باشد. در حالی که برای اعتبارسنجی و داده های آزمایشی ایدهآل است که نماینده توزیع یکسانی باشند، مواجه شدن با تفاوت ها غیرممکن نیست. آگاهی از این امکان و در نظر گرفتن آن هنگام ارزیابی عملکرد یک مدل در اعتبارسنجی و داده های آزمایشی بسیار مهم است.

و) نادرست یکی از دلایل استفاده از MSE در رگرسیون خطی این است که به راحتی قابل مشتق پذیری است. این ویژگی مشتق پذیری برای الگوریتم های بهینه سازی مانند گرادیان کاوهشی بسیار مفید است، زیرا مشتق راحتتر محاسبه می شود و الگوریتم به سرعت به جواب نزدیکتر می شود. از آنجا که MAE (Mean Absolute Error) به جای مربع چگالی خطای خطای (error)، اختلاف مطلق را محاسبه می کند، نقطه ای برای محاسبه مشتق به وجود نمی آورد. بنابراین، در الگوریتم های بهینه سازی که مبتنی بر مشتق هستند، استفاده از MAE نیازمند توجه و پیاده سازی ویژه تری است. اما اگر مقادیر خطای بسیار کوچک و به تقریب صفر نزدیک هستند، تغییر از MSE به MAE تأثیر چندانی نخواهد داشت.

ز) درست نیست. در اعتبار سنجی متقاطع k-fold، دقت نهایی معمولاً با میانگین گیری مقادیر دقت به دست آمده در هر قسمت برآورد می شود. بهترین نتیجه به دست آمده در هر قسمت لزوماً عملکرد کلی مدل را نشان نمی دهد، زیرا ممکن است به دلیل overfit بر روی آن چیز خاص باشد. هدف از اعتبار سنجی متقاطع k-fold بدست آوردن تخمین قوی تر از عملکرد مدل با ارزیابی آن بر روی چندین تقسیم بندی مختلف از داده ها است.

سؤال ۲ (۱۵ نمره)

در شکل سمت چپ یک نمونه‌ی بیش‌پرازش، در شکل وسط یک نمونه‌ی مدل مناسب و در شکل راست یک نمونه‌ی زیرپرازش رسم کنید.



سوال ۳ (۱/۵ نمره)

در صورتی که برای مسئله‌ی رگرسیون خطی، رابطه‌ی مدل به صورت

$$y(x, w) = w^T x, \quad w \in \mathbf{R}^{10} \quad (1)$$

باشد و تابع هزینه برای مسئله‌ی رگرسیون خطی به صورت

$$\tilde{E} = \frac{1}{2} \sum_{n=1}^N \{y(x_n, w) - t_n\}^2 + \frac{\lambda}{2} \|w\|^2 \quad (2)$$

در نظر گرفته شده باشد. مشخص کنید که مقادیر وزن‌های هرستون متعلق به کدام مقادیر 0 , -18 و $-\infty$ برای $\ln \lambda$ است. دلیل خود را کامل توضیح دهید.

	a	b	c
w_0^*	0.35	0.35	0.13
w_1^*	232.37	4.74	-0.05
w_2^*	-5321.83	-0.77	-0.06
w_3^*	48568.31	-31.97	-0.05
w_4^*	-231639.30	-3.89	-0.03
w_5^*	640042.26	55.28	-0.02
w_6^*	-1061800.52	41.32	-0.01
w_7^*	1042400.18	-45.95	-0.00
w_8^*	-557682.99	-91.53	0.00
w_9^*	125201.43	72.68	0.01

$$\ln(\lambda) = 0 \rightarrow \lambda = 1$$

در این حالت لاندا بیشترین مقدار را دارد و پنالتی ترم از بقیه بیشتر است پس مدل به این سمت می‌رود که مقدار وزن‌های خود به کمترین مقدار خود برساند. پس این مقدار $\ln a$ به ستون c مربوط است.

$$\ln(\lambda) = -18 \rightarrow \lambda = e^{-18}$$

در این حالت لاندا مقدار نزدیک به صفر دارد و مقدار پنالتی ترم نزدیک به صفر است پس مقادیر وزن‌ها کنترل شده هستند اما این مقدار لاندا بیشتر از مقدار قبلی اجازه زیاد شدن وزن‌ها را میدهد پس این مقدار $\ln b$ مربوط به ستون b است.

$$\ln(\lambda) = -\infty \rightarrow \lambda = 0$$

در این حالت لاندا تقریباً برابر صفر است و خروجی الگوریتم مثل زمانی است که از رگرسیون ساده استفاده می‌کنیم یعنی روی مقدار وزن‌ها کنترلی وجود ندارد پس این مقدار هم مرتبط با ستون a است.

سؤال ۶ (۲ نمره)

فرض کنید x یک متغیر تصادفی گسسته با احتمال زیر باشد

x	.	۱	۲	۳
$p(x)$	$\frac{\theta}{2}$	$\frac{\theta}{4}$	$\frac{3(1-\theta)}{4}$	$\frac{(1-\theta)}{2}$

اگر مشاهدات حاصل از نمونه برداری از این متغیر به صورت $(0, 3, 1, 2, 1, 0, 2, 1, 3, 2)$ باشد، تخمین حداقل درستنمایی θ را به دست آورید.

$$\text{MLE} = P(x=0) * P(x=3)*P(x=1) * P(x=2)*P(x=1) * P(x=0)*P(x=2) * P(x=1)*P(x=3) * P(x=2) = \\ z/2 * (1-z)/2 * z/4 * 3(1-z)/2 * z/4 * z/2 * 3(1-z)/4 * z/4 * (1-z)/2 * 3(1-z)/4 =$$

برای یافتن برآوردگر حداقل درستنمایی (MLE) برای شواهد داده شده، باید مقدار z را پیدا کنیم که تابع درستنمایی را حداقل می کند.
تابع احتمال با ضرب احتمالات شواهد داده شده تحت توزیع احتمال داده شده محاسبه می شود.

$$L(z) = p(x=0)^2 * p(x=3)^2 * p(x=1)^3 * p(x=2)^3$$

$$L(z) = (z/2)^2 * ((1-z)/2)^2 * (z/4)^3 * (3(1-z)/4)^3$$

برای یافتن MLE، باید مقدار z را پیدا کنیم که $L(z)$ را به حداقل برساند. گرفتن لگاریتم تابع احتمال می تواند محاسبات را ساده کند:

$$\log(L(z)) = 2\log(z/2) + 2\log((1-z)/2) + 3\log(z/4) + 3\log(3(1-z)/4)$$

$$\log(L(z)) = 2\log(z) - 2\log(2) + 2\log(1-z) - 2\log(2) + 3\log(z) - 3\log(4) + 3\log(3) + 3\log(1-z) - 3\log(4)$$

$$\log(L(z)) = 2\log(z) + 2\log(1-z) + 3\log(z) + 3\log(1-z) + c$$

$$\log(L(z)) = 5\log(z) + 5\log(1-z) + c$$

$$d(\log(L(z)))/dz = 5(1/z) - 5(1/(1-z)) = 0 \rightarrow z = 1/2$$

سؤال ۴ (۱۵ نمره)

فرض کنید نمونه‌های x_1, x_2, \dots, x_n از یک توزیع یکنواخت با بازه $[0, \theta]$ نمونه‌برداری شده‌اند و $\theta > 0$ باشد. بیشینه درستنما می‌پارامتر θ را بیابید.

سوال (۴) ۲۷ ■ Apr 2014
حمدانی‌الثانی ۱۴۳۵

- p.d.f for uniform distribution:

$$f(x|\theta) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

The likelihood function:

$$\prod_{i=1}^n f(x_i|\theta) = \frac{1}{\theta^n} \text{ and } \text{درینه Sample En}\text{ درینه}\theta \text{ را دارند.}$$

: Log-Likelihood

$$\Rightarrow \arg\max \log(\theta^n) = -n \log \theta$$

پس مردمی متناسب $\rightarrow -\frac{n}{\theta} \rightarrow \theta \geq n$ باشد

منی‌است رزیانی θ در بینشیم باشد

دلیلی که نیست باز هم دارد زمانی که کمترین داشته باشد

$$\theta = \max(x_1, x_2, x_3, x_4, \dots, x_n)$$

سؤال ۷ (۲ نمره)

برای هریک از مجموعه وزن‌های ستون جدول زیر، تعیین کنید کدام ستون متعلق به تابع هزینه‌ی Ridge و Lasso می‌باشد. دلیل آن را به همراه رسم نمودار در ۲ بعد شرح دهید. برتری هریک بر دیگر را توضیح دهید.

	Column A	Column B
w_1	0.38	0.50
w_2	0.23	0.20
w_3	-0.02	0.00
w_4	0.15	0.09
w_5	0.21	0.00
w_6	0.03	0.00
w_7	0.04	0.00
w_8	0.12	0.05

column A → Ridge regression

column B → Lasso regression

در زمان استفاده از ridge و lasso مقادیری به عنوان وزن انتخاب می‌شوند که خیلی بزرگ نباشند در عین حال مقدار کمی loss را داشته باشند.

اگردو مجموعه رو در نظر بگیریم :

اولی مجموعه کانتور پلات ها که مقادیر لاس رو نشون میدن

دومی مجموعه نقاطی در صفحه که جمع قدرمطلق وزن ها (lasso) یا جمع توان دو آن‌ها (Ridge) از یک مقدار کمتر باشند . برای پیدا کردن بهترین وزن ها باید اشتراک این دو مجموعه رو پیدا کنیم برای lasso چون شکل مربعی میگیرد اشتراک این دو مجموعه فقط یک نقطه روی صفر یا نزدیک صفر است و برای ridge چون اشتراک دو دایره نزدیک صفر است و به یک نقطه محدود نمی‌شود مقادیر وزن ها بزرگ‌تر از صفر است.

:Lasso

1. انتخاب ویژگی: رگرسیون lasso می‌تواند انتخاب ویژگی را انجام دهد و ضرایب را به طور خودکار به صفر کاهش دهد. این به شناسایی و حذف ویژگی‌های نامرتب یا زائد کمک می‌کند و قابلیت تفسیر مدل را بهبود می‌بخشد.

2. چند خطی بودن را کنترل می‌کند: lasso می‌تواند چند خطی بودن را کنترل کند، که زمانی رخ می‌دهد که متغیرهای پیش‌بینی‌کننده همبستگی بالایی دارند.

معایب:

1. فقط زیرمجموعه‌ای از پیش‌بینی‌ها را انتخاب می‌کند: رگرسیون lasso فقط می‌تواند زیرمجموعه‌ای از پیش‌بینی‌ها را انتخاب کند، که ممکن است در شرایطی که انتظار می‌رود همه پیش‌بینی‌ها بر متغیر پاسخ تأثیر بگذارند، یک نقطه ضعف باشد.

رگرسیون ریج:

1. چند خطی بودن را کنترل می‌کند: رگرسیون ریج با کاهش تأثیر پیش‌بینی‌کننده‌های همبسته و کوچک کردن ضرایب آنها به سمت صفر، در مدیریت چند خطی مؤثر است.
2. پایدار و قابل اعتماد: رگرسیون ریج در مقایسه با رگرسیون حداقل مربعات، به ویژه در موقعیت‌هایی با چند خطی یا همخطی بالا، تخمین‌های مدل پایدارتر و قابل اعتمادتری را ایجاد می‌کند.
3. مناسب برای مواردی که دارای پیش‌بینی‌کننده‌های زیادی هستند: رگرسیون ریج به ویژه در هنگام برخورد با مجموعه داده‌هایی که دارای تعداد زیادی پیش‌بینی‌کننده هستند یا زمانی که نسبت پیش‌بینی‌کننده‌ها به مشاهدات زیاد است، مفید است.

معایب:

1. ضرایب به صفر کاهش نیافته اند: برخلاف رگرسیون Ridge، رگرسیون lasso به طور خودکار ویژگی‌های نامربوط را از مدل حذف نمی‌کند.
2. عدم تفسیرپذیری: تخمین ضرایب در رگرسیون ریج به دلیل اثر انقباض اعمال شده بر ضرایب می‌تواند برای تفسیر چالش برانگیز باشد.

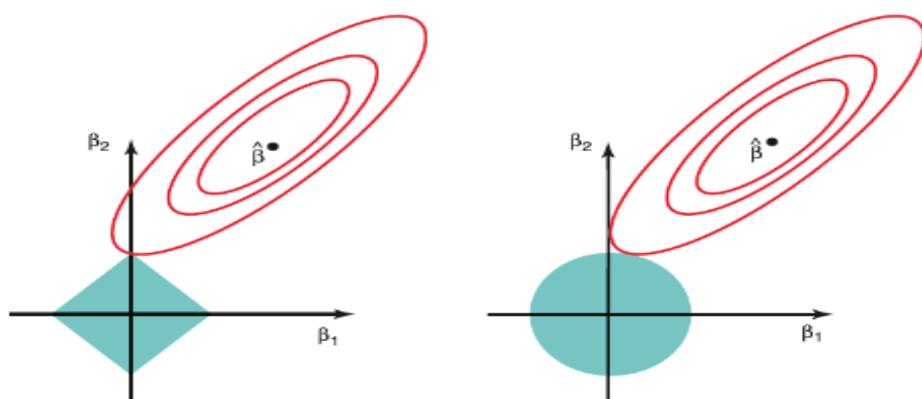


FIGURE 6.7. Contours of the error and constraint functions for the lasso (left) and ridge regression (right). The solid blue areas are the constraint regions, $|\beta_1| + |\beta_2| \leq s$ and $\beta_1^2 + \beta_2^2 \leq s$, while the red ellipses are the contours of the RSS.

سؤال ۸ (۳/۵ نمره)

برای مسئله‌ی رگرسیون خطی، داده‌های زیر موجود هستند. در صورتی که مقادیر وزن‌ها به صورت زیر باشد، به سوال‌های زیر جواب دهید.

$$w_0 + w_1X + w_2Y = Z, \quad w_0 = -59.5, \quad w_1 = -0.15, \quad w_2 = 0.6 \quad (8)$$

$$LearningRate = 0.1$$

X	Y	Z
41	138	37.99
42	153	47.34
37	151	44.38
46	133	28.17

الف) خطای MSE و خطای MAE را در این مرحله محاسبه کنید. (۰/۵ نمره)

ب) به روش SGD وزن‌ها را با MSE تا ۴ مرحله بهروزرسانی کنید. (۱/۵ نمره)

ج) وزن‌ها را با MSE به روش GD یک مرحله بهروزرسانی کنید. (۱ نمره)

د) خطای MSE و وزن‌های حاصل شده برای هریک از روش‌ها مقایسه کنید. (۰/۵ نمره)

$$\text{learning rate} = 0.00001$$

(الف)

X0 * w0	X1 * w1	Y * w2	Pred	Z	e	e^2
-59.5	-6.15	82.8	17.15	37.99	20.84	434.3056
-59.5	-6.3	91.8	26	47.34	21.34	455.3956
-59.5	-5.55	90.6	25.55	44.38	18.83	354.5689
-59.5	-6.9	79.8	13.4	28.17	14.77	218.1529
				MEAN	18.945	365.60575

MAE = 18.94

MSE = 365.6

در ادامه برای راحتی در محاسبات بیشتر از دو رقم اعشار گرد شده است.

$$X \rightarrow \text{few} \times \overset{\text{Big}}{B^T} \quad X^T \rightarrow \overset{\text{Big}}{B} \times \text{few}$$

Subject: _____
Year: _____ Month: _____ Day: _____ ()

page: ()

$$y = w \cdot x \quad e = y - \hat{y} \quad j(0) = \frac{1}{2m} \cdot e^T \cdot e \quad (4.8)$$

$$\frac{\partial J(\Theta)}{\partial \theta} = \frac{1}{m} \cdot x^T \cdot e \rightarrow \text{logistic function output}$$

$$\begin{array}{l}
 \text{d7091} = \boxed{\begin{array}{|c|} \hline 1 \\ \hline 91 \\ \hline 138 \\ \hline \end{array}} \times \boxed{\begin{array}{|c|} \hline e \\ \hline 2984 \\ \hline 1x1 \\ \hline \end{array}} = \boxed{\begin{array}{|c|} \hline 2984 \\ \hline 854 \\ \hline 2875 \\ \hline \end{array}} \xrightarrow[\times 10]{\times 10^{-3}} \boxed{\begin{array}{|c|} \hline 010208 \\ \hline 4854 \\ \hline 2187 \\ \hline \end{array}}
 \end{array}$$

$$\begin{array}{c} \text{w}_0 \\ \hline \left[\begin{array}{c|c|c} -59.5 & 0.02 & -59.52 \\ \hline 0.15 & -0.85 & -1 \\ \hline 0.6 & 2.87 & -2.27 \end{array} \right] \\ \hline \text{w}_{\text{new}} = \end{array}$$

$$x: \begin{array}{|c|c|c|} \hline x_0 & x_1 & x_2 \\ \hline 1 & 42 & 153 \\ \hline \end{array} \quad \begin{array}{|c|} \hline w_1 \\ \hline -59_1 \\ 52 \\ -1 \\ \hline \end{array} = -59,52 - 42 - 347,31 = 448,83 = 7$$

$$\begin{array}{c}
 \text{d7(G)} \\
 \text{dG} \\
 \hline
 = \left[\begin{array}{c|c}
 x_0 & 1 \\
 x_1 & 42 \\
 x_2 & 153 \\
 \hline
 3 \times 1
 \end{array} \right] \quad \boxed{+ 8999995} \quad = \quad \left[\begin{array}{c}
 \sim \rightarrow + 496 \\
 \sim \rightarrow 20932 \\
 \sim \rightarrow 75888 \\
 \hline
 3 \times 1
 \end{array} \right] \quad \times 10^3 \rightarrow \left[\begin{array}{c}
 0/49 \\
 20932 \\
 75888 \\
 \hline
 \text{PASARGAD}
 \end{array} \right]
 \end{array}$$

Subject:

Year : Month : Day : ()

page : ()

$$w_2 = \begin{bmatrix} -59,52 \\ -1 \\ -2,27 \end{bmatrix} - \alpha \frac{\partial f(\theta)}{\partial \theta} \begin{bmatrix} w_1 \\ 0,49 \\ 20,83 \\ 75,88 \end{bmatrix} = \begin{bmatrix} -60 \\ -21,93 \\ -78,15 \end{bmatrix}$$

$$x = \begin{bmatrix} x_0 & x_1 & x_2 \\ 1 & 37 & 151 \end{bmatrix}_{1 \times 3} \cdot w_2: \begin{bmatrix} -60 \\ -21,93 \\ -78,15 \end{bmatrix}_{3 \times 1} = -1260896$$

$$e = y - \hat{y} = 44,38 + 1260896 = 1260940$$

$$\cancel{f(\theta)} = x_0 \begin{bmatrix} x \\ 1 \\ 37 \\ 151 \end{bmatrix} \cdot e: \begin{bmatrix} 1260940 \\ 46654794 \\ 190401985 \end{bmatrix} \xrightarrow{x \cdot r = 10^{-3}}$$

$$\cancel{f'(\theta)} \rightarrow \begin{bmatrix} 126094 \\ 46654179 \\ 19040119 \end{bmatrix} \Rightarrow \begin{bmatrix} -60 \\ -21,93 \\ -78,15 \end{bmatrix} - 1r. \cancel{\frac{f(\theta)}{d\theta}} \approx \begin{bmatrix} -1320 \\ 46633 \\ -19022 \end{bmatrix}$$

21

PASARGAD

Subject:

Year : Month : Day : ()

page : ()

$$\begin{array}{c} \text{X} \\ \text{x} \end{array} \begin{array}{c} x_0 \quad x_1 \quad x_2 \\ | \quad | \quad | \\ 1 \quad 46 \quad 133 \end{array} \cdot w_3 = \begin{array}{c} -1320 \\ | \\ 46600 \\ | \\ -19000 \end{array} = 4,671,920 = \hat{y}$$

: f, h, b, y, e

3

$$e = y - \hat{y} = 28,17 - \hat{y} = +4,671,948$$

$$\begin{array}{c} \text{X} \\ \text{x} \end{array} \begin{array}{c} \cancel{f(G)} \\ \cancel{d_0} \end{array} = \begin{array}{c} x_0 \quad 0 \\ | \quad | \\ x_1 \quad 46 \\ | \\ x_2 \quad 133 \end{array} \cdot e \approx \begin{array}{c} +4,671,948 \\ | \\ 215,000,000 \\ | \\ 621,000,000 \end{array} \xrightarrow{*1r=10^{-3}} \begin{array}{c} 4,671,948 \\ | \\ 215,000 \\ | \\ 621,000 \end{array}$$

$$w_3 \quad w_4$$

$$w_4: \begin{array}{c} -1320 \\ | \\ -46600 \\ | \\ -19000 \end{array} - 18 \cdot \begin{array}{c} \cancel{f(G)} \\ \cancel{d_0} \end{array} = \begin{array}{c} -5992 \\ | \\ -261600 \\ | \\ -640000 \end{array}$$

9

12

$$y = X \cdot w = \begin{array}{c} 17,15 \\ | \\ 26 \\ | \\ 25,5 \\ | \\ 13,4 \end{array} - \begin{array}{c} 37,9 \\ | \\ 47,3 \\ | \\ 44,38 \\ | \\ 28,17 \end{array} = \begin{array}{c} -20,8 \\ | \\ -21,3 \\ | \\ -18,1 \\ | \\ -14,77 \end{array} \quad (E-8)$$

15

$$\begin{array}{c} \cancel{f(B)} \\ \cancel{d_0} \end{array} = \int_m \cdot X^T \cdot e = \begin{array}{c} -35,7 \\ | \\ -3124 \\ | \\ -10938 \end{array} = \begin{array}{c} -18,9 \\ | \\ -781 \\ | \\ -2734 \end{array} \xrightarrow{*1r=10^{-2}} \begin{array}{c} -0,189 \\ | \\ -7,81 \\ | \\ -27,13 \end{array}$$

18

21

$$\begin{array}{c} -59,3 \\ | \\ 7,66 \\ | \\ 27,9 \end{array} = w_2$$

PASARGAD

Subject:

Year : Month : Day : ()

Page : ()

	-159/3	1-	5992	
w-GD :	7166	W-SGD :	-261600	(-8)
3	27/9		64000	

$$MSE-GD = 4229,7$$

$$MSE-SGD: 99039608$$

6

، انت GD میان برآورد نهاد و وزن های میانگین های بودند

12

15

18

21