APRML: Assignment 1

Anshuman Suri (2014021)

2. (A) *L1*, *L2*, *L_elastic norm*:

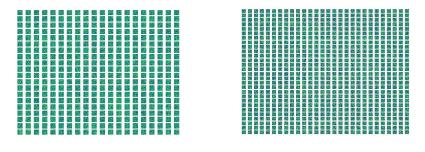
(·) = ·, = <u>-</u> · · · · · · · · · · · · · · · · · · ·
• Z = 6(W1X+b1) X = W2 2 +b2
Standard 2789 0 - 1 11 x - x 11 + 21 (1/41/41/11)
+ negularger 2 + \(\lambda_2 \left(\reft(\left(\reft(\left(\left(\left(\left(\reft(\left(\left(\reft(\
Standard error , $e = 1 x - x ^2 + \lambda_1(w_1 + w_2)$ + regularizer a + $\lambda_1(w_1 ^2 + w_2 ^2)$ Assuming regularization for both role of weights
[12] can be desired using elastic net (relling 2,100,2,00)
For Loslic:
2e - 2.1 (1-x) (7T)+ 2, sign(11)
2W2 2 m
$\frac{\partial e}{\partial w_2} = \frac{2.1}{2} (x^2 - x).(z^T) + \frac{1}{2} sign(w_2)$ $\frac{\partial e}{\partial w_2} = \frac{2.1}{2} (x^2 - x).(z^T) + \frac{1}{2} sign(w_2)$ $\frac{\partial e}{\partial w_2} = \frac{2.1}{2} (x^2 - x).(z^T) + \frac{1}{2} sign(w_2)$ $\frac{\partial e}{\partial w_2} = \frac{2.1}{2} (x^2 - x).(z^T) + \frac{1}{2} sign(w_2)$
de = 2.1 (x-x) W2. Z(1-z).X + 21 sign(W1)
2W1 2 ~ [964.
$\frac{\partial e}{\partial w_1} = 2.1 (x^2 - x) w_2 . Z(1-z).X + \lambda_1 sign(w_1)$ $\frac{\partial w_1}{\partial w_1} = 2.1 (x^2 - x) w_2 . Z(1-z).X + \lambda_1 sign(w_1)$ $\frac{\partial w_1}{\partial w_1} = 2.1 (x^2 - x) w_2 . Z(1-z).X + \lambda_1 sign(w_1)$ $\frac{\partial w_1}{\partial w_1} = 2.1 (x^2 - x) w_2 . Z(1-z).X + \lambda_1 sign(w_1)$ $\frac{\partial w_1}{\partial w_1} = 2.1 (x^2 - x) w_2 . Z(1-z).X + \lambda_1 sign(w_1)$ $\frac{\partial w_1}{\partial w_1} = 2.1 (x^2 - x) w_2 . Z(1-z).X + \lambda_1 sign(w_1)$
Since fagulosizes is not on biases. Le e de ase unoffedel.
there rules con therefore be used for back- prop.
Food L, norm, not 21=12 /2=0 in whose
For Ly nown, let 4=0 & 21=1 in alone
& /

L_trace norm:

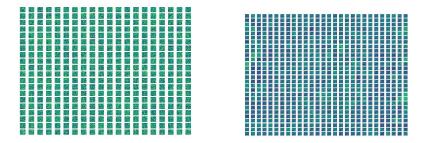
7	= 6(W1X +b1), x - W2Z+b2
	e=1 11 × -x112+ sum(eigenvaluen (W1)) +sum (eigenvaluen (W))
S	um(eigenvalues(x)) = +9(2) where X=UEVT
_	K= UZVT XV= UZ VV= UZ
	UTX V= UTU S = 2 .: tx(2) = tx (UTXV)
A	ccollidating to motorix cookbook, 2 (AXB) = (BA)
f	tese we need a ta (VIXV)
1	lanever, the eigenflux = D's eigenvectors in Al B will not affect the derivative. So, we consider only the part of UEV with 2>0 is,
0	only the post of USV with 2>0 is,

X= U' S'VT, -> Cropped outstaintly	
4 Egynous >0	
Chapter 1748 (N. 4711) 210	_ T.
ousdinger : 2ta (U: 3xV) = 2ta	(NXA)
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	7-x
{on 4x(4,)= 4x (2')}	
: 24x(Uxv) = (1,U,T) = U,V,T	02 1 2
Using this 28-21/x x) 2T + U2'V3",	
20 2 20 =2.1 W, Z(+z)x
JW1 2 4 U	17,1

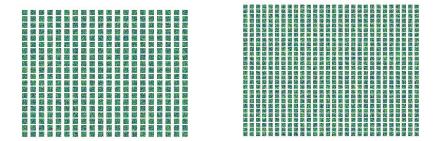
(B) Visualized weights L1 norm (83.17% testing accuracy)



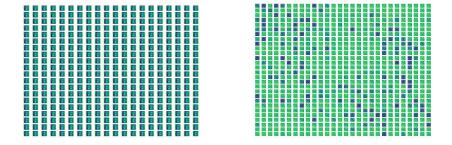
L2 norm (84.13% testing accuracy)

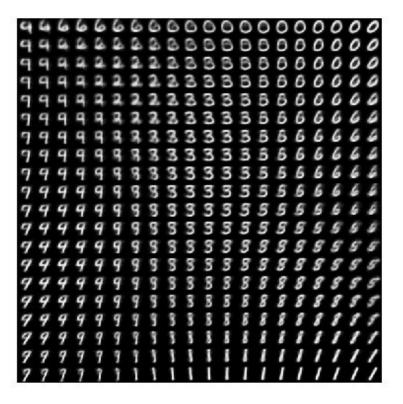


L-elastic (0.5, 0.5) norm (78.8% testing accuracy)



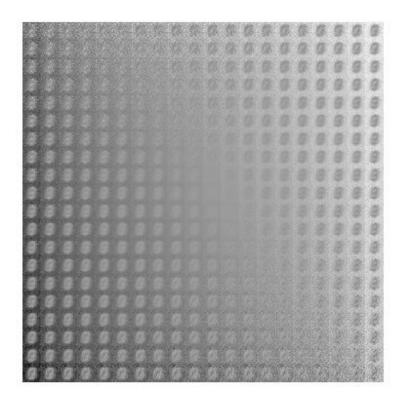
L-trace norm (46.47% testing accuracy)





4	4	4	6	4	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
4	4	4	4	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	5
9	4	4	4	4	6	4	6	6	6	6	G	G	6	6	6	6	6	6	5
9	4	4	4	4	4	6	6	G	6	G	G	G	G	G	G	6	Ø	5	5
9	9	4	9	q	G	Q	G	0	G	G	G	G	G	G	G	5	5	5	5
9	9	4	4	4	q	a	G	G	G	G	G	G	G	\mathcal{G}	G	5	\mathcal{S}	5	3
9	9	9	4	q	\mathcal{Q}	q	\mathcal{G}	g	G	G	G	S	G	S	5	5	5	5	3
9	9	4	4	q	q	q	\mathcal{Q}	8	G	B	G	G	G	\mathcal{G}	g	S	S	5	3
9	9	4	4	q	q	q	q	g	g	G	g	g	g	S	S	5	8	3	3
9	9	9	4	q	q	q	q	Q_{i}	B	B	B	g	g	g	S	S	S	3	3
9	4	q	q	q	q	q	q	8	8	g	g	g	8	g	S	8	8	3	3
9	9	q	4	9	q	q	q	Q_3	g	8	g	g	g	g	S	8	8	3	3
9	9	9	4	q	q	q	Q	8	8	8	8	8	8	g	8	8	8	3	3
9	9	q	9	9	Q.	9	Q.	g	8	8	8	8	8	8	8	8	8	8	3
9	4	q	9	9	Q.	q	ā	g	8	00	8	00	8	8	8	8	8	8	3
9	9	9	4	9	9	9	8	8	8	8	8	80	8	8	8	8	8	8	3
7	7	9	7	7	7	à	3	3	8	8	8	8	8	8	8	8	8	8	3
7	7	7	7	7	7	7	2	3	3	5	8	8	8	8	8	8	8	8	3
7	7	7	7	7	7	?	?	3	5	3	5	5	5	8	8	8	8	8	8
7	7	7	7	7	7	7	7	7	2	2	2	2	2	2	2	2	2	2	2

Gaussian, 2D (Final loss 45.27)



Gaussian, 5D (Final loss:)

