Exercícios Computacionais 1 – ECp1

Rodrigo Seiji Piubeli Hirao (186837)

16 de dezembro de 2021

Conteúdo

1 Questão 01

A função analítica consegue encontrar a solução usando apenas uma fórmula, porém tal fórmula começa a ficar computacionalmente complexa com a adição de mais variáveis, assim o treinamento iterativo é uma opção melhor por sua simplicidade computacional mesmo com múltiplas variáveis, o ponto contra o uso da solução iterativa é que ela procura por mínimos mais próximos, ou seja, minimos locais e não necessariamente globais, o que pode não ser tão otimizado quando a solução analítica.

2 Questão 02

a) Pelo fato da regularização LASSO usar a norma L1, os parâmetros do sistema se aproximam mais rapidamente de 0, podendo também se igualar a zero, assim encontrando pontos ótimos do valor do coeficiente onde prâmetros se cancelam, podendo assim ser eliminados.

b) A Elastic Net nada mais é que uma combinação entre as regressões LASSO e Ridge, de forma que ela pode se comportar como qualquer uma delas ou até ter combinações do comportamento de ambas, assim ele tem aproxima os coeficientes de 0 como a LASSO, mas não some com coeficientes altamente correlacionas, apenas mantendo-os pequenos, o ponto negativo da Elastic Net é o fato de ter de ser calculado 2 parâmetros α e λ, nas outras regressões com regularização é apenas necessário calcular λ.

c) Como cada pixel da imagem possuia um parametro, então rodar o LASSO iria fazer com que alguns pixels fossem ignorados.

d) Como na ELM possuíamos 1000 neurônios todos com uma distribuição de participação de cada pixel, não seria possível eliminar totalmente a participação de um pixel, apenas de alguns neurônios.

3 Questão 03

Sabendo que um neurônio parte de um anterior com a equação 1

$$x^{[q]} = \tanh(w^{[q]} x_i^{[q-1]} + b^{[q-1]}) \tag{1}$$

Mas assumindo que a distribuição é idênticamente distribuída temos que $E(w_i) = 0$, $E(x_i^{[q-1]}) = 0$ e $\sum_{i=1}^{k} \operatorname{Var}(x_i^{[q-1]}) \operatorname{Var}(w_i) = n^{[q-1]} \operatorname{Var}(w) \operatorname{Var}(x^{[q-1]})$ como pode ser visto na resolução da equação 2

$$\begin{aligned}
\operatorname{Var}(x^{[q]}) &= \operatorname{Var}(\sum_{i=1}^{k}(w_i^{[q]} x_i^{[q-1]})) \\
&= \sum_{i=1}^{k} \operatorname{Var}(w_i^{[q]} x_i^{[q-1]}) \\
&= \sum_{i=1}^{k} (E(w_i^{[q]}))^2 \operatorname{Var}(x_i^{[q-1]}) + (E(x_i^{[q-1]}))^2 \operatorname{Var}(w_i^{[q]}) + \operatorname{Var}(x_i^{[q-1]}) \operatorname{Var}(w_i^{[q]}) \\
&= \sum_{i=1}^{k} \operatorname{Var}(x_i^{[q-1]}) \operatorname{Var}(w_i^{[q]}) \\
&= n^{[q-1]} \operatorname{Var}(w^{[q]}) \operatorname{Var}(x^{[q-1]})
\end{aligned} \tag{2}$$

Sabendo disso temos a relação da equação 3, pois $\operatorname{Var}(x^{[q]}) = \operatorname{Var}(x^{[q-1]})$

$$\operatorname{Var}(w^{[q]}) = \frac{1}{n^{[q-1]}} \tag{3}$$

E como temos que $\operatorname{Var}(x) = \frac{(b-a)^2}{12}$ onde, no nosso caso, $a = -b$

$$\operatorname{Var}(w^{[q]}) = \frac{(2b)^2}{12} = \frac{4b^2}{12} = \frac{b^2}{3} \tag{4}$$

E, por conclusão, temos o resultado da equação 5

$$\mathrm{Var}(w^{[q]}) = \frac{1}{n^{[q-1]}} = \frac{b^2}{3}$$
$$b = \sqrt{\frac{3}{n^{[q-1]}}} \tag{5}$$

4 Questão 04

A convolução é uma operação que pega um contorno de um ponto e adiciona valores ponderados de seus arredores, em uma imagem isso pode ser visto como uma multiplicação ponderada de todos os pixels em volta, sendo que a imagem possui 3 camadas (vermelho, verde e azul), mas podem ser usados diferentes números de camadas.

Logo, temos uma matriz de entrada $m \times n \times p$ e um filtro $i \times j \times p$, tal que devemos fazer a multiplicação do filtro ponto a ponto com a matriz de entrada, assim tendo nosso resultado como uma matrix $(m-i+1) \times (n-j+1) \times 1$ considerando que i e j são ímpares.

Uma convolução separável é o ato de separar a convolução em 2 etapas, uma em que multiplicamos os canais individualmente por matrizes $i \times j \times 1$, e colocamos cada resultado um em cima do outro novamente, tendo a matriz resultante $(m-i+1) \times (n-j+1) \times p$ (depth wise convolution) e a segunda etapa, onde multiplicamos ponto a ponto por uma matriz $1 \times 1 \times p$ para ter a convolução das camadas ponto a ponto, tendo novamente o resultado final $(m-i+1) \times (n-j+1) \times 1$ (point wise convolution).

Sua vantagem é o fato de ser uma operação paralelizável no número de camadas e possuir mais operações que são mais simples.

5 Questão 05

Sendo x o valor a ser mutado, y o resultado, $N(x, \frac{x}{10})$ a função normal com média x e variância $\sqrt{\frac{x}{10}}$ e rand() uma função de geração de números aleatórios uniformes de 0 a 1, temos a formulação da equação 6

$$\begin{cases} y = N(x, \frac{x}{10})(rand()) \mod (v_i^{[max]}) & , y \geq v_i^{[min]} \\ y = v_i^{[min]} + (N(x, \frac{x}{10})(rand()) \mod (v_i^{[max]})) & , y < v_i^{[min]} \end{cases} \tag{6}$$

6 Questão 06

Sejam C constantes de configuração e f o valor de cada parâmetro, temos a função de fitness da equação 7

$$f_{fitness} = \frac{1}{C_{NC} f_{NC} + C_{NMD} f_{NMD} + C_{NDA} f_{NDA} + 1} \tag{7}$$

7 Questão 07

a) Simulated annealing é quando não decidimos mudar para um outro ponto apenas se a perda é menor, mas de acordo com a probabilidade resultante da equação 8, que depende da temperatura T, uma variável que começa alta e diminui com as execuções, assim dando cada mez meis importância à perda, se comportando cada vez mais como uma otimização normal em busca de máximos locais. Porém pelo fato de esta buscar em lugares com maior perda no começo ela tem uma chance muito mais elevada de cair num máximo global.

$$p = e^{-\frac{perda}{T}} \tag{8}$$

b) Considerando que já há um grafo de feromônios populados de acordo com a quantidade de formigas que passou por aquelas arestas recentemente (considerando feromonização e evaporação) denominadas τ e um grafo com os custos de cada aresta invertidos $\eta = \frac{1}{custo}$, temos que a probabilidade de ser escolhido a aresta i entre n arestas está definido na equação 9. α e β são constantes configurais para a influência de cada variável.

$$p = \frac{\tau_i^\alpha \eta_i^\beta}{\sum_{j=0}^{n} \tau_j^\alpha \eta_j^\beta} \tag{9}$$

c) Por meio de operadores evolutivos como mutação e seleção é gerado novos participantes, e dentre esses (ou dentre esses e os anteriores) são escolhidos membros para recomeçar o processo. Os membros são escolhidos de forma a maximizar o fitness, uma função calculada a partir das características dos modelos produzidos.

8 Questão 08

a) $\binom{10}{7} = 120$

b) $\binom{10}{7} - \binom{5}{2} - \binom{5}{1} = 105$

9 Questão 09

a) Esolhendo os dados da dabela a seguit, onde o carro apenas faz curva quando está prestes a colidir com uma parede e vira para um dos lados possíveis (ou direita quando os 2 casos são iguais) temos os resultados de $P \cdot PP \cdot G = NS$ e $P \cdot MP \cdot G = NB$

d_1, d_3 / d_2	MP	PP	ME	MD	PG	MG
P, P	–	–	Z	Z	Z	Z
M, P	PB	PS	Z	Z	Z	Z
G, P	PB	PS	Z	Z	Z	Z
P, M	PB	PS	Z	Z	Z	Z
M, M	PB	PS	Z	Z	Z	Z
G, M	NB	NS	Z	Z	Z	Z
P, G	NB	NS	Z	Z	Z	Z
M, G	NB	NS	Z	Z	Z	Z
G, G	PB	PS	Z	Z	Z	Z

b)

- $d_1 = 0.7$ - [1, 0, 0] - **P**
- $d_2 = 1.2$ - [0.8, 0.2, 0, 0, 0, 0] - **0.8 MP**
- $d_3 = 4.1$ - [0, 0, 1] - **G**

Logo temos $\min(1, 0.8, 1) = 0.8$ pelo método de Mandani, ou seja **0.8 PB** que resulta no ângulo de **28**