

TD réseau de neurones

Gaston LENCZNER, Javiera CASTILLO NAVARRO,
Guillaume VAUDAUX RUTH, Adrien CHAN-HON-TONG

Notations et rappels :

- on note T la transposition matricielle
- $relu(x) = \max(x, 0) = [x]_+$ (par composante pour les vecteurs)
- les vecteurs colonnes dans \mathbb{R}^J sont considérés comme des matrices $J \times 1$,
et, $\forall (A, B) \in \mathbb{R}^{I \times J} \times \mathbb{R}^{J \times K}$ 2 matrices, AB est leur produit dans $\mathbb{R}^{I \times K}$
- un multi layer perceptron de profondeur P est une fonction qui peut
s'écrire $W_P relu(W_{P-1} relu(\dots (W_2 relu(W_1 x + b_1) + b_2) \dots) + b_{P-1}) + b_P$
avec W_p des matrices et b_p des vecteurs

Partie 1 réseau préappris

Q1

On considère la fonction $f(x) = f((x_1 \ x_2)^T) = x_2 - relu(x_1 - x_2)$

Q1.1 : déterminez les zones où f est positive vs négative.

- déjà relu est toujours positif, donc si $x_2 < 0$, f est forcément négative
- maintenant même si $x_2 > 0$, f pourrait être négative en fonction de
 $x_1 - x_2$:
- soit $x_1 < 2x_2$ et $f(x) > 0$
- soit $x_1 > 2x_2$, et $f(x) < 0$

Q1.2 : écrivez cette fonction comme un réseau de neurones.

aide : $x = relu(x) - relu(-x)$ et si $x = (x_1 \ x_2)^T$ alors $x_1 = (1 \ 0)x$

$$f(x) = (1 \ -1 \ -1) relu \left(\begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 1 & -1 \end{pmatrix} x \right)$$

Q2 même questions avec $g((x_1 \ x_2)^T) = x_2 + relu(x_1 - x_2)$ et $h((x_1 \ x_2)^T) = x_1 + relu(x_2 - x_1)$, que remarquez vous ?

Si $x_1 > x_2$, $g(x) = x_1$ et si $x_1 < x_2$, $g(x) = x_2$ donc $g(x) = \max(x_1, x_2)$,
donc $g(x) < 0 \Leftrightarrow x_1 < 0$ et $x_2 < 0$.

$$g(x) = (1 \ -1 \ 1) relu \left(\begin{pmatrix} 0 & 1 \\ 0 & -1 \\ 1 & -1 \end{pmatrix} x \right)$$

$g = h$ alors qu'ils s'écrivent différemment - c'est une *symétrie* cachée.

*Notez au passage qu'on peut construire le max de 2 neurones avec des relu, il
faudra vous en rappeler pour le cours suivant, pour relativiser les nouveautés
introduites dans les CNN !*

Partie 2 réseau à déterminer (chercher des poids triviaux)

1 neurone

Q3.1 : Montrez qu'il est possible d'apprendre par coeur la base de données $((1 \ 1)^T, 1)$, $((-1 \ -1)^T, -1)$ avec 1 neurone sans biais (et sans activation puisque les activations concernent les couches cachées).

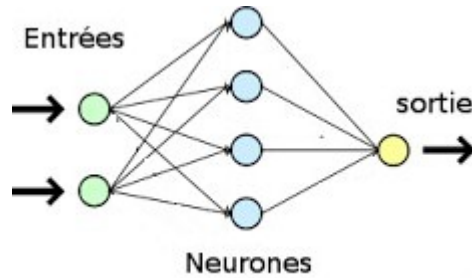
$$f(x) = (1 \ 1)x \text{ car } (1 \ 1)(1 \ 1)^T = 2 > 0 \text{ et } (1 \ 1)(-1 \ -1)^T = -2 < 0$$

Q3.2 : Est-il possible d'apprendre par coeur la base de données $((0 \ 1)^T, 1)$, $((0 \ -1)^T, 1)$, $((1 \ 0)^T, -1)$, $((-1 \ 0)^T, -1)$ avec 1 neurone sans biais ?

Non : notons $w = (\alpha, \beta)$, alors $w(1, 0)^T < 0$ et $w(-1, 0)^T < 0$ implique $\alpha < 0$ et $-\alpha < 0$!

2 couches de neurones

Q4.1 : Montrez qu'il possible d'apprendre par coeur la base de données $((0 \ 1)^T, 1)$, $((0 \ -1)^T, 1)$, $((1 \ 0)^T, -1)$, $((-1 \ 0)^T, -1)$ avec le réseau ci dessous (sans biais et avec activation relu).



$$f(x) = \text{relu}((0 \ 1)x) + \text{relu}((0 \ -1)x) - \text{relu}((1 \ 0)x) - \text{relu}((-1 \ 0)x)$$

car :

$$\Rightarrow \text{relu}((0 \ 1)(0 \ 1)^T) + \text{relu}((0 \ -1)(0 \ 1)^T) - \text{relu}((1 \ 0)(0 \ 1)^T) - \text{relu}((-1 \ 0)(0 \ 1)^T) = \text{relu}(1) + \text{relu}(-1) - 0 - 0 = 1 > 0$$

$$\Rightarrow \text{relu}((0 \ 1)(0 \ -1)^T) + \text{relu}((0 \ -1)(0 \ -1)^T) - \text{relu}((1 \ 0)(0 \ -1)^T) - \text{relu}((-1 \ 0)(0 \ -1)^T) = \text{relu}(-1) + \text{relu}(1) - 0 - 0 = 1 > 0$$

$$\Rightarrow \text{relu}((0 \ 1)(1 \ 0)^T) + \text{relu}((0 \ -1)(1 \ 0)^T) - \text{relu}((1 \ 0)(1 \ 0)^T) - \text{relu}((-1 \ 0)(1 \ 0)^T) = 0 + 0 - \text{relu}(1) - \text{relu}(-1) = -1 < 0$$

$$\Rightarrow \text{relu}((0 \ 1)(-1 \ 0)^T) + \text{relu}((0 \ -1)(-1 \ 0)^T) - \text{relu}((1 \ 0)(-1 \ 0)^T) - \text{relu}((-1 \ 0)(-1 \ 0)^T) = 0 + 0 - \text{relu}(-1) - \text{relu}(1) = -1 < 0$$

Q4.2 : estimez les zones $f(x) > 0$ et $f(x) < 0$.

l'espace est partitionné en 4 par les 4 diagonales ($\pm x_1 \pm x_2 = 0$)

Q4.3 : Est-il possible d'apprendre avec le même réseau (mais d'autres poids) la base $((0 \ 2)^T, 1), ((0 \ -2)^T, 1), ((2 \ 0)^T, 1), ((-2 \ 0)^T, 1), ((0 \ 0)^T, -1)$?

pas de biais implique $f((0 \ 0)^T) = 0$, il est impossible d'avoir $f((0 \ 0)^T) < 0$

2 couches de neurones avec biais

Q5 : Considérons encore même la base de données $((0 \ 2)^T, 1), ((0 \ -2)^T, 1), ((2 \ 0)^T, 1), ((-2 \ 0)^T, 1), ((0 \ 0)^T, -1)$, ainsi que les 2 réseaux

— $\psi(x) = [(0 \ 1)x]_+ + [(0 \ -1)x]_+ + [(1 \ 0)x]_+ + [(-1 \ 0)x]_+ - 1$

— $\phi(x) = 2\text{relu}((-1 \ 1)x - 1) + 2\text{relu}((1 \ -1)x - 1) - 1$

Q5.1 : Montrez qu'ils apprennent la base par coeur.

faire comme en Q4.1 - tester chaque point

Q5.2 : Donnez la structure de chaque réseau.

réseau à 2 couches dans les 2 cas, 4 puis 1 neurones pour le premier, 2 puis 1 neurones pour le second.

Q5.3 : Dessinez les zones positives et négatives.

la frontière forme un losange pour le premier réseau mais 2 droites parallèles formant un bandeau *diagonal* pour le second.

Pour votre culture : cette base est intéressant car il est possible de l'apprendre asymétriquement avec un réseau de 3 neurones. Mais pour obtenir une solution symétrique et bornée, il faut 5 neurones. Ainsi, dans cet exemple précis, plus de paramètres permet d'obtenir une solution plus élégante. Attention c'est plutôt faux en général !