Projet Informatique Individuelle
ENSC – Semestre 8

Théorie du perceptron multi-couches

Plan du cours

1. Problème

Soit le perceptron multi-couches dont les couches sont liées entre elles par des relations de type :

$$f_B(B * W) = A \Leftrightarrow f_B\left(\begin{bmatrix} B_1 & \cdots & B_{n_B} \end{bmatrix} * \begin{bmatrix} W_{1,1} & \cdots & W_{1,n_A} \\ \vdots & & \vdots \\ W_{n_B,1} & \cdots & W_{n_B,n_A} \end{bmatrix}\right) = \begin{bmatrix} A_1 & \cdots & A_{n_A} \end{bmatrix}$$

avec

- n_B le nombre d'entrées dans la couche B
- n_A le nombre de sorties dans la couche A
- f_B la fonction d'activation de la couche B ($f_B : \mathbb{R} \to [-1, 1]$)
- $B \in \mathcal{M}_{1,n_B}$ la matrice contenant les valeurs d'entrée
- $\{B_b, \forall\ b \in [\![1, n_B]\!]\}$ les valeurs des entrées
- $A \in \mathcal{M}_{1,n_A}$ la matrice contenant les valeurs de sortie
- $\left\{A_j, \forall\ a \in [\![1, n_A]\!]\right\}$ les valeurs des sorties
- $W \in \mathcal{M}_{n_B,n_A}$ la matrice contenant les poids
- $\{W_{b,a}, \forall\ (b,a) \in [\![1, n_B]\!] \times [\![1, n_A]\!]\}$ les poids entre les B_b et les A_a

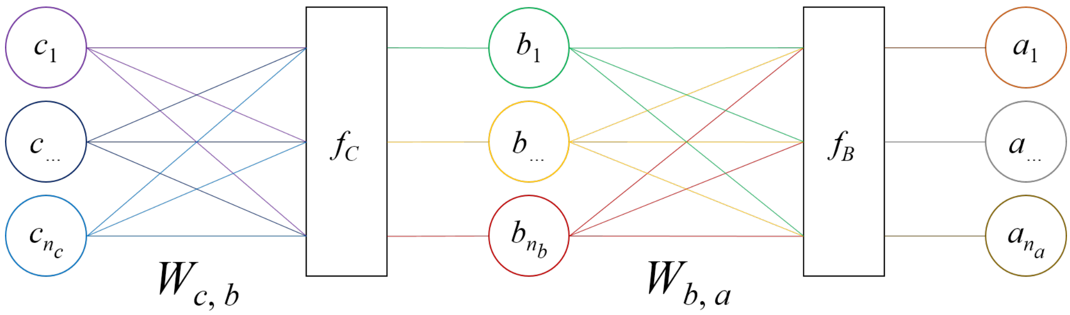

FIGURE 1 – Représentation des relations entre quelques couches de neurones

Dans le reste du document, $*$ désignera la multiplication matricielle et $\times$ la multiplication terme à terme.
On notera ${}^{\mathrm{t}}M$ la transposée de la matrice M.

2. Correction des poids de la dernière couche

Supposons que pour tout $a \in [\![1, n_A]\!]$ la sortie théorique soit α_a et non A_a.

L'erreur quadratique $\mathcal{E}$ commise sur un tel réseau est : $\mathcal{E} = \frac{1}{2} \sum\limits_{\ell=1}^{n_A} (\alpha_\ell - A_\ell)^2$.

On s'intéresse aux variations qu'apporteraient les modifications de $W_{b,a}$ ($(b, a) \in [\![1, n_B]\!] \times [\![1, n_A]\!]$) sur l'erreur quadratique.

Il s'agit donc de faire varier les $W_{b,a}$ pour chaque couple $(b, a) \in [\![1, n_B]\!] \times [\![1, n_A]\!]$ de telle façon à ce que réduire au maximum $\mathcal{E}$.

On va donc calculer pour tout $(b, a) \in [\![1, n_B]\!] \times [\![1, n_A]\!]$ les dérivées partielles $\varepsilon_{b,a}$ de $\mathcal{E}$ par rapport aux $W_{b,a}$.

Pour un b et un a donnés :

$$
\begin{aligned}
\varepsilon_{b,a} &= \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} \mathcal{E} \\
&= \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} \frac{1}{2} \sum_{\ell=1}^{n_A} (\alpha_\ell - A_\ell)^2 \\
&= \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} \frac{1}{2} \sum_{\ell=1}^{n_A} \left(\alpha_\ell - f_B \left(\sum_{k=1}^{n_B} B_k W_{\ell,k} \right) \right)^2 \\
&= \frac{1}{2} \sum_{\ell=1}^{n_A} \left[\frac{\mathrm{d}}{\mathrm{d}W_{b,a}} \left(\alpha_\ell - f_B \left(\sum_{k=1}^{n_B} B_k W_{\ell,k} \right) \right)^2 \right] \\
&= \sum_{\ell=1}^{n_A} \left[\left(\alpha_\ell - f_B \left(\sum_{j=1}^{n_B} B_j W_{\ell,j} \right) \right) \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} \left(\alpha_\ell - f_B \left(\sum_{k=1}^{n_B} B_k W_{\ell,k} \right) \right) \right] \\
&= - \sum_{\ell=1}^{n_A} \left[\left(\alpha_\ell - f_B \left(\sum_{j=1}^{n_B} B_j W_{\ell,j} \right) \right) \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} f_B \left(\sum_{k=1}^{n_B} B_k W_{\ell,k} \right) \right] \\
&= - \sum_{\ell=1}^{n_A} \left[\left(\alpha_\ell - f_B \left(\sum_{j=1}^{n_B} B_j W_{\ell,j} \right) \right) f'_B \left(\sum_{i=1}^{n_B} B_i W_{\ell,i} \right) \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} \sum_{k=1}^{n_B} B_k W_{\ell,k} \right] \\
&= - \sum_{\ell=1}^{n_A} \left[\left(\alpha_\ell - f_B \left(\sum_{j=1}^{n_B} B_j W_{\ell,j} \right) \right) f'_B \left(\sum_{i=1}^{n_B} B_i W_{\ell,i} \right) \sum_{k=1}^{n_B} \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} B_k W_{\ell,k} \right] \\
&= - \sum_{\ell=1}^{n_A} \left[\left(\alpha_\ell - f_B \left(\sum_{j=1}^{n_B} B_j W_{\ell,j} \right) \right) f'_B \left(\sum_{i=1}^{n_B} B_i W_{\ell,i} \right) \sum_{k=1}^{n_B} \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} B_k W_{\ell,k} \right] \\
&= - \left(\alpha_a - f_B \left(\sum_{j=1}^{n_B} B_j W_{a,j} \right) \right) f'_B \left(\sum_{i=1}^{n_B} B_i W_{a,i} \right) \frac{\mathrm{d}}{\mathrm{d}W_{b,a}} B_b W_{a,b} \\
&= - \left(\alpha_a - f_B \left(\sum_{j=1}^{n_B} B_j W_{a,j} \right) \right) f'_B \left(\sum_{i=1}^{n_B} B_i W_{a,i} \right) B_b \\
&= - (\alpha_a - A_a) \, f'_B \left(\sum_{i=1}^{n_B} B_i W_{a,i} \right) B_b
\end{aligned}
$$

En notant $\varepsilon_a = \begin{bmatrix} \varepsilon_{1,a} & \cdots & \varepsilon_{n_B,a} \end{bmatrix} \in \mathcal{M}_{1,n_B}$, on a donc :

$$\varepsilon_a = -\left(\alpha_a - A_a\right) f_B'\left(\sum_{i=1}^{n_B} B_i W_{a,i}\right)\begin{bmatrix} B_1 & \cdots & B_{n_B} \end{bmatrix} = -\left(\alpha_a - A_a\right) f_B'\left(\sum_{i=1}^{n_B} B_i W_{a,i}\right) B$$

Puis en notant $\varepsilon = \begin{bmatrix} \varepsilon_{1,1} & \cdots & \varepsilon_{1,n_A} \\ \vdots & & \vdots \\ \varepsilon_{n_B,1} & \cdots & \varepsilon_{n_B,n_A} \end{bmatrix} = \begin{bmatrix} {}^{\mathbf{t}}\varepsilon_1 & \cdots & {}^{\mathbf{t}}\varepsilon_{n_B} \end{bmatrix} \in \mathcal{M}_{n_B,n_A}$, on en déduit que :

$$\begin{aligned}
\varepsilon &= \left[-\left(\alpha_1 - A_1\right) f_B'\left(\sum_{i=1}^{n_B} B_i W_{1,i}\right) {}^{\mathbf{t}}B \quad \cdots \quad -\left(\alpha_{n_A} - A_{n_A}\right) f_B'\left(\sum_{i=1}^{n_B} B_i W_{n_A,i}\right) {}^{\mathbf{t}}B\right] \\
&= -\,{}^{\mathbf{t}}B * \left[\left(\alpha_1 - A_1\right) f_B'\left(\sum_{i=1}^{n_B} B_i W_{1,i}\right) \quad \cdots \quad \left(\alpha_{n_A} - A_{n_A}\right) f_B'\left(\sum_{i=1}^{n_B} B_i W_{n_A,i}\right)\right] \\
&= -\,{}^{\mathbf{t}}B * \left(\begin{bmatrix} (\alpha_1 - A_1) & \cdots & (\alpha_{n_A} - A_{n_A}) \end{bmatrix} \times \left[f_B'\left(\sum_{i=1}^{n_B} B_i W_{1,i}\right) \quad \cdots \quad f_B'\left(\sum_{i=1}^{n_B} B_i W_{n_A,i}\right)\right]\right) \\
&= -\,{}^{\mathbf{t}}B * \left[\left(\begin{bmatrix} \alpha_1 & \cdots & \alpha_{n_A} \end{bmatrix} - \begin{bmatrix} A_1 & \cdots & A_{n_A} \end{bmatrix}\right) \times f_B'\left(\left[\sum_{i=1}^{n_B} B_i W_{1,i} \quad \cdots \quad \sum_{i=1}^{n_B} B_i W_{n_A,i}\right]\right)\right] \\
&= -\,{}^{\mathbf{t}}B * \left[(\alpha - A) \times f_B'\left(B * W\right)\right]
\end{aligned}$$

On posant $\delta_B = \left[(\alpha - A) \times f_B'\left(B * W\right)\right]$, on a $\varepsilon = -\,{}^{\mathbf{t}}B * \delta_B$.

3. Correction des poids de l'avant-dernière précédentes

Maintenant qu'on a calculé les variations à appliquer sur les poids entre la dernière et l'avant-dernière couche, il s'agit désormais de faire de même pour les poids entre les couches précédentes.

L'erreur quadratique $\mathcal{E}$ commise est toujours : $\mathcal{E} = \frac{1}{2}\sum_{\ell=1}^{n_A}(\alpha_\ell - A_\ell)^2$.

On s'intéresse maintenant aux variations qu'apporteraient les modifications de $W'_{c,b}$ $((c,b) \in [\![1,n_C]\!] \times [\![1,n_B]\!])$ sur l'erreur quadratique.

Il s'agit donc de faire varier les $W'_{c,b}$ pour chaque couple $(c,b) \in [\![1,n_C]\!] \times [\![1,n_B]\!]$ de telle façon à ce que réduire au maximum $\mathcal{E}$.

On va donc calculer pour tout $(c,b) \in [\![1,n_C]\!] \times [\![1,n_B]\!]$ les dérivées partielles $\varepsilon'_{c,b}$ de $\mathcal{E}$ par rapport aux $W'_{c,b}$.

Pour un c et un b donnés :

$$
\begin{aligned}
\varepsilon'_{c,b} &= \frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\mathcal{E} \\
&= \frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\frac{1}{2}\sum_{\ell=1}^{n_A}(\alpha_\ell - A_\ell)^2 \\
&= \frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\frac{1}{2}\sum_{\ell=1}^{n_A}\left(\alpha_\ell - f_B\left(\sum_{k=1}^{n_B} B_k W_{k,\ell}\right)\right)^2 \\
&= \frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\frac{1}{2}\sum_{\ell=1}^{n_A}\left(\alpha_\ell - f_B\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right)^2 \\
&= \frac{1}{2}\sum_{\ell=1}^{n_A}\left[\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(\alpha_\ell - f_B\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right)^2\right] \\
&= \sum_{\ell=1}^{n_A}\left[\left(\alpha_\ell - f_B\left(\sum_{i=1}^{n_B} f_C\left(\sum_{h=1}^{n_c} C_h W'_{h,i}\right) W_{i,\ell}\right)\right)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(\alpha_\ell - f_B\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right)\right] \\
&= \sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(\alpha_\ell - f_B\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right)\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}} f_B\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell) f'_B\left(\sum_{i=1}^{n_B} f_C\left(\sum_{h=1}^{n_c} C_h W'_{h,i}\right) W_{i,\ell}\right)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell) f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(\sum_{k=1}^{n_B} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right) W_{k,\ell}\right)\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell) f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right)\sum_{k=1}^{n_B}\left(W_{k,\ell}\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}} f_C\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right)\right)\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell) f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right)\sum_{k=1}^{n_B}\left(W_{k,\ell}\, f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,k}\right)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(\sum_{j=1}^{n_c} C_j W'_{j,k}\right)\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
\varepsilon'_{c,b} &= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{b,\ell}\ f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,b}\right)\frac{\mathrm{d}}{\mathrm{d}W'_{c,b}}\left(C_c W'_{c,b}\right)\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{b,\ell}\ f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,b}\right) C_c\right] \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{b,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,b}\right) C_c
\end{aligned}
$$

En notant $\varepsilon'_b = \begin{bmatrix}\varepsilon_{1,b} & \cdots & \varepsilon_{n_C,b}\end{bmatrix} \in \mathcal{M}_{1,n_C}$, on a donc :

$$
\begin{aligned}
\varepsilon'_b &= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{b,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,b}\right)\begin{bmatrix}C_1 & \cdots & C_{n_C}\end{bmatrix} \\
&= -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{b,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,b}\right) C
\end{aligned}
$$

Puis en notant $\varepsilon' = \begin{bmatrix}\varepsilon'_{1,1} & \cdots & \varepsilon'_{1,n_B} \\ \vdots & & \vdots \\ \varepsilon'_{n_C,1} & \cdots & \varepsilon'_{n_C,n_B}\end{bmatrix} = \begin{bmatrix}{}^{\mathbf{t}}\varepsilon'_1 & \cdots & {}^{\mathbf{t}}\varepsilon'_b\end{bmatrix} \in \mathcal{M}_{n_C,n_B}$, on en déduit que :

$$
\begin{aligned}
\varepsilon' &= \left[-\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{1,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,1}\right){}^{\mathbf{t}}C \quad \cdots \quad -\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{n_B,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,n_B}\right){}^{\mathbf{t}}C\right] \\
&= -{}^{\mathbf{t}}C * \left[\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{1,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,1}\right) \quad \cdots \quad \sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{n_B,\ell}\right] f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,n_B}\right)\right] \\
&= -{}^{\mathbf{t}}C * \left(\left[\sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{1,\ell}\right] \quad \cdots \quad \sum_{\ell=1}^{n_A}\left[(\alpha_\ell - A_\ell)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,\ell}\right) W_{n_B,\ell}\right]\right] \times \left[f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,1}\right) \quad \cdots \quad f'_C\left(\sum_{h=1}^{n_c} C_h W'_{h,n_B}\right)\right]\right) \\
&= -{}^{\mathbf{t}}C * \left[\left(\left[(\alpha_1 - A_1)\, f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,1}\right) \quad \cdots \quad \left(\alpha_{n_A} - A_{n_A}\right) f'_B\left(\sum_{i=1}^{n_B} B_i W_{i,n_A}\right)\right] * \begin{bmatrix}W_{1,1} & \cdots & W_{n_B,1} \\ \vdots & & \vdots \\ W_{1,n_A} & \cdots & W_{n_B,n_A}\end{bmatrix}\right) \times f'_C\left(\left[\sum_{h=1}^{n_c} C_h W'_{h,1} \quad \cdots \quad \sum_{h=1}^{n_c} C_h W'_{h,n_B}\right]\right)\right] \\
&= -{}^{\mathbf{t}}C * \left[\left(\left[(\alpha - A) \times f'_B\,(B * W)\right] * {}^{\mathbf{t}}W\right) \times f'_C\,(C * W')\right] \\
&= -{}^{\mathbf{t}}C * \left[\left(\delta_B * {}^{\mathbf{t}}W\right) \times f'_C\,(C * W')\right]
\end{aligned}
$$

On posant $\delta_C = \left[\left(\delta_B * {}^{\mathbf{t}}W\right) \times f'_C\,(C * W')\right]$, on a $\varepsilon' = -\,{}^{\mathbf{t}}C * \delta_C$.

4. Correction des poids des précédentes couches

Pour une couche D ayant n_D neurones, la fonction d'activation f_D et les poids $W''_{d,c}$ $((d,c) \in [\![1, n_D]\!] \times [\![1, n_C]\!])$ entre les couches D et C, on peut montrer que :

$$\varepsilon'' = -\,{}^{\mathbf{t}}D * \left[\left(\delta_C * {}^{\mathbf{t}}W'\right) \times f'_D\,(D * W'')\right]$$

On posant $\delta_D = \left[\left(\delta_C * {}^{\mathbf{t}}W'\right) \times f'_D\,(D * W'')\right]$, on a $\varepsilon'' = -\,{}^{\mathbf{t}}D * \delta_D$.
De même, pour une couche E ayant n_E neurones, la fonction d'activation f_E et les poids $W'''_{e,d}$ $((e,d) \in [\![1, n_E]\!] \times [\![1, n_D]\!])$ entre les couches E et D, on peut montrer que :

$$\varepsilon''' = -\,{}^{\mathbf{t}}E * \left[\left(\delta_D * {}^{\mathbf{t}}W''\right) \times f'_E\,(E * W''')\right]$$

On posant $\delta_E = \left[\left(\delta_D * {}^{\mathbf{t}}W''\right) \times f'_E\,(E * W''')\right]$, on a $\varepsilon''' = -\,{}^{\mathbf{t}}E * \delta_E$.