Formalisation
$a_{1..n} \in [-1;1]$ une action à composantes continues
$s_{1..m} \in [-1;1]$ un état continue
$h \in \mathbb{N}$ nombre de neurones cachées
$v_{1..h} \in \mathbb{R}$ vecteur poids de la couche de sortie
$w_{1..h\times 1..(n+m)} \in \mathbb{R}$ matrice poids de la couche caché

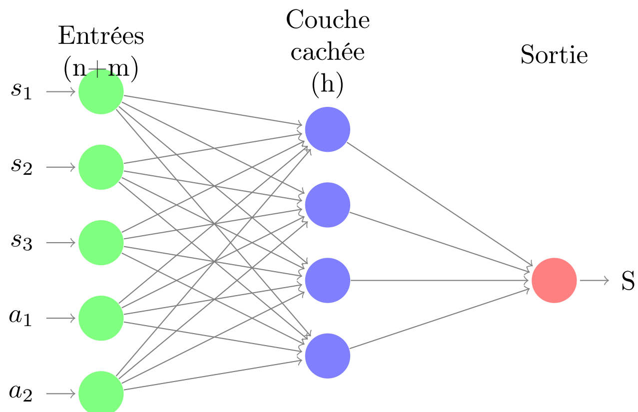
La sortie du réseau sans biais peut s'écrire :

$$S = \tanh[\sum_i^h (v_i \times \tanh[\sum_j^n (w_{i,j} \times s_j) + \sum_j^m (w_{i,(j+n)} \times a_j)])]$$

Le but du réseau est d'approximer $Q(s,a)$:

$$S \to r + \gamma \times \max_a Q(s_{t+1}, a)$$

Exemple



Optimisation Trouver les actions a_i qui maximisent la sortie S lorsque les s_i sont fixés.

$$a^*_{1..n} = \arg\max_{a_{1..n}}(\tanh[\sum_i^h (v_i \times \tanh[\sum_j^n (w_{i,j} \times s_j) + \sum_j^m (w_{i,(j+n)} \times a_j)])])$$
$$\text{sous contraintes}$$
$$\forall i \in \{1,..,n\}, a_i \geq -1$$
$$\forall i \in \{1,..,n\}, a_i \leq 1$$

tanh est une bijection strictement croissante, $\sum_j^n (w_{i,j} \times s_j)$ est constant, on peut donc réécrire :

$$a^*_{1..n} = \arg\max_{a_{1..n}}(\sum_i^h (v_i \times \tanh[C_{i,j} + \sum_j^m (w_{i,(j+n)} \times a_j)]))$$
$$\text{sous contraintes}$$
$$\forall i \in \{1,..,n\}, a_i + 1 \geq 0$$
$$\forall i \in \{1,..,n\}, -a_i + 1 \geq 0$$

Pour remplir les conditions de Kuhn-Tucker, il faut au minimum que les contraintes soient quasi-concaves (elles sont linéaires donc c'est bon) et que la fonction objective soit pseudo-concave (est-ce le cas ?).

Posons

$$
\begin{aligned}
f(a_{1..n}) &= (\sum_{i}^{h}(v_i \times \tanh[C_{i,j} + \sum_{j}^{m}(w_{i,(j+n)} \times a_j)])) \\
g(a_i) &= a_i + 1 \\
h(a_i) &= -a_i + 1
\end{aligned}
$$

Le lagrangien s'écrit alors :

$$\mathcal{L}(a_{1..n}, \lambda_{1..2\times n}) = f(a_{1..n}) + \sum_{i=1}^{n}(\lambda_i \times g(a_i)) + \sum_{i=n}^{2\times n}(\lambda_i \times h(a_i))$$

On cherche donc à résoudre

$$a^*, \lambda^* = \arg\max_{a_{1..n}} \mathcal{L}(a_{1..n}, \lambda_{1..2\times n})$$

Les contraintes sont linéaires donc les conditions de qualification des contraintes sont vérifiées, on peut donc passer aux conditions du premier ordre :
Si le vecteur a^* est solution alors il existe un unique vecteur λ^* tel que ces vecteurs vérifient les $(n + 3(2 * n)) = 7n$ conditions suivantes :

$$
\begin{aligned}
\forall i \in \{1..n\}, &\forall j \in \{1..(2\times n)\} : \\
\frac{\partial \mathcal{L}}{\partial a_i} &= 0 \\
\frac{\partial \mathcal{L}}{\partial \lambda_j} &\geq 0 \\
\lambda_j &\geq 0 \\
\lambda_i \times g(a_i) &= 0 \\
\lambda_{n+i} \times h(a_i) &= 0
\end{aligned}
$$

Exemple Espace