Annexe B

Présentation du calcul du gradient de la fonction de coût J dans le cas d'un réseau d'ondelettes d'état avec possibilité de choisir la sortie comme variable d'état

Nous présentons ici le calcul du gradient de J pour un réseau d'état où la sortie peut être choisie comme une des variables d'état.

Le nombre total des états sera N_s avec :

$$N_s = N_{sy} + N_{ss} \tag{1}$$

Si la sortie est une variable d'état alors on a : N_{sy} =1. Sinon il vaut zéro. N_{ss} est donc le nombre des variables d'état différentes de la sortie.

1. Notations.

Pour pouvoir indicer les neurones d'état, on introduit une variable logique associée à N_{sy} ,définie de la façon suivante :

$$A_{sy} = \begin{cases} 1 \; si \; N_{sy}=0 \\ 0 \; si \; N_{sy}=1 \end{cases} \tag{2}$$

Dans un cas général, les neurones d'état seront indicés de $N_e+N_s+N_w+A_{sy}+1$ à $N_e+N_s+N_w+N_{ss}+1$. Les paramètres du réseau sont donc :

- les translations m_{jk} et les dilatations d_{jk} avec k=1 , ... , N_e+N_s et j=1 , ... , N_w ;
- les pondérations et les coefficients directs : on note c_{kj} le paramètre associé à la connexion entre la fonction (ou le neurone d'entrée) j et le neurone de sortie (ou le neurone d'état) k. Pour les pondérations nous avons $j= N_e+N_s+1$, ... , $N_e+N_s+N_w$ et $k=N_e+N_s+N_w+1$, ... ,$N_e+N_s+N_w+N_{ss}+1$; pour les coefficients directs nous avons j=1 , ... , N_e+N_s et $k=N_e+N_s+N_w+1$, ... ,$N_e+N_s+N_w+N_{ss}+1$;
- un terme constant sur le neurone linéaire de sortie, noté c_0 ;

Le nombre de composantes du vecteur θ est alors $2N_w(N_e+N_s)+(N_{ss}+1)(N_e+N_s+N_w)+1$.

La sortie y ainsi que l'expression des variables d'état en sortie sont identiques à celles données par les relations (59) et (60) du chapitre III.

Pour chaque copie du réseau (n=2, ... , N), les variables d'état en entrée sont calculées à partir de la relation suivante :

$$x_k^n = x_{k+N_s+N_w+A_{sy}}^{n\pm1} \quad \text{avec } k = N_e+1 , \ldots , N_e+N_{sy}+N_{ss} \tag{3}$$

Le cas particulier de la première copie est discuté au paragraphe VI.4.1 du chapitre III.

2. Calcul du gradient de *J* par rapport aux états par rétropropagation.

Pour la copie N, nous avons :

Pour la sortie :

$$\frac{\partial J}{\partial y^N} = \pm e^N \tag{4}$$

Pour les variables d'état en sortie, $k=N_e+N_s+N_w+2, \ldots, N_e+N_s+N_w+N_{ss}+1$:

$$\frac{\partial J}{\partial x_k^N} = 0 \tag{5}$$

Pour les variables d'état en entrée, $k=N_e+1, \ldots, N_e+N_s$:

$$\frac{\partial J}{\partial x_k^N} = \frac{\partial J}{\partial y^N}\frac{\partial y^N}{\partial x_k^N} = \pm e^N\left(c_{\alpha,k} + \sum_{j=1}^{N_w}\frac{c_{\alpha,N_e+N_s+j}}{d_{jk}}\frac{\partial \Phi_j(x)}{\partial z_{jk}}\right) \tag{6}$$

avec $\alpha = N_e+N_s+N_w+1$.

Pour les copies de $n=N-1$ à 2, nous avons :

Pour la sortie :

$$\frac{\partial J}{\partial y^n} = \frac{\partial J}{\partial x^n_{N_e+N_s+N_w+1}} = \begin{cases} \pm e^n \quad si\ N_{sy}=0 \\ \pm e^n + \dfrac{\partial J}{\partial x^{n+1}_{N_e+1}} \quad sinon \end{cases} \tag{7}$$

Pour les variables d'état en sortie, $k=N_e+N_s+N_w+2, \ldots, N_e+N_s+N_w+N_{ss}+1$:

$$\frac{\partial J}{\partial x_k^n} = \frac{\partial J}{\partial S^{n+1}_{k\pm N_s\pm N_w\pm A_{sy}}} \tag{8}$$

Pour les variables d'état en entrée, $k=N_e+1, \ldots, N_e+N_s$:

$$\frac{\partial J}{\partial x_k^n} = \frac{\partial J}{\partial y^n}\frac{\partial y^n}{\partial x_k^n} = \pm e^n\left(c_{\alpha,k} + \sum_{j=1}^{N_w}\frac{c_{\alpha,N_e+N_s+j}}{d_{jk}}\frac{\partial \Phi_j(x)}{\partial z_{jk}}\right) + \sum_{j=N_e+N_s+N_w+2}^{N_e+N_s+N_w+N_{ss}+1} c_{j,k}\frac{\partial J}{\partial x_j^n} \tag{9}$$

Pour la copie $n=1$, nous avons :

Pour la sortie :

$$\frac{\partial J}{\partial y^1} = \frac{\partial J}{\partial x^1_{N_e+N_s+N_w+1}} = \begin{cases} \pm e^1 \text{ si } N_{sy} \\ \pm e^1 + \dfrac{\partial J}{\partial x^2_{N_e+N_s+N_w 1}} \text{ sinon} \end{cases} \tag{10}$$

Pour les variables d'état en sortie, k=N_e+N_s+N_w+2, ... ,N_e+N_s+N_w+N_{ss}+1 :

$$\frac{\partial J}{\partial x^1_k} = \frac{\partial J}{\partial x^2_{k \pm N_s \pm N_w \pm A_{sy}}} \tag{11}$$

Pour les variables d'état en entrée, k=N_e+1, ... ,N_e+N_s : le calcul des $\frac{\partial J}{\partial x^1_k}$ n'est pas utile.

3. Calcul du gradient de *J* par rapport aux paramètres du réseau.

Pour les coefficients directs sur la sortie :

$$\frac{\partial J}{\partial c_{\alpha j}} = \sum_{n=1}^{N} \frac{\partial J}{\partial y^n} \frac{\partial y^n}{\partial c^n_{\alpha j}} = \pm \sum_{n=1}^{N} e^n x^n_j \text{ avec } j = 1, \dots, N_e+N_s \text{ et } \alpha = N_e+N_s+N_w+1 \tag{12}$$

Pour les coefficients directs sur les états :

$$\frac{\partial J}{\partial c_{k,j}} = \sum_{n=1}^{N} \frac{\partial J}{\partial x^n_k} \frac{\partial x^n_k}{\partial c^n_{k,j}} = \sum_{n=1}^{N} \frac{\partial J}{\partial x^n_k} x^n_j \tag{13}$$

avec $j = 1, \dots, N_e+N_s$ et $k = N_e+N_s+N_w+2, \dots, N_e+N_s+N_w+N_{ss}+1$

Pour les pondérations sur la sortie :

$$\frac{\partial J}{\partial c_{\alpha, N_e+N_s+j}} = \sum_{n=1}^{N} \frac{\partial J}{\partial y^n} \frac{\partial y^n}{\partial c^n_{\alpha, N_e+N_s+j}} = \pm \sum_{n=1}^{N} e^n \Phi_j(x^n) \tag{14}$$

avec $j = 1, \dots, N_w$ et $a = N_e+N_s+N_w+1$

Pour les pondérations sur les états :

$$\frac{\partial J}{\partial c_{k, N_e+N_s+j}} = \sum_{n=1}^{N} \frac{\partial J}{\partial x^n_k} \frac{\partial x^n_k}{\partial c^n_{k, N_e+N_s+j}} = \sum_{n=1}^{N} \frac{\partial J}{\partial x^n_k} \Phi_j(x^n) \tag{15}$$

avec $j = 1, \dots, N_w$ et $k = N_e+N_s+N_w+2, \dots, N_e+N_s+N_w+N_{ss}+1$

Pour le terme constant sur le neurone de sortie :

$$\frac{\partial J}{\partial c_0} = \sum_{n=1}^{N} \frac{\partial J}{\partial y^n} \frac{\partial y^n}{\partial c^n_0} = \pm \sum_{n=1}^{N} e^n \tag{16}$$

Pour les translations :

$$\frac{\partial J}{\partial m_{jk}} = \sum_{n=1}^{N} \frac{\partial J}{\partial m_{jk}^{n}} = \sum_{n=1}^{N} \frac{1}{d_{jk}} \frac{\partial \Phi_j(x^n)}{\partial z_{jk}} \left(c_{\alpha, N_e+N_s+j}\, e^n \pm \sum_{l=N_e+N_s+N_w+2}^{N_e+N_s+N_w+N_{ss}+1} c_{l, N_e+N_s+j} \frac{\partial J}{\partial x_l^n} \right) \quad (17)$$

Pour les dilatations :

$$\frac{\partial J}{\partial d_{jk}} = \sum_{n=1}^{N} \frac{\partial J}{\partial d_{jk}^{n}} = \sum_{n=1}^{N} \frac{z_{jk}^{n}}{d_{jk}} \frac{\partial \Phi_j(x^n)}{\partial z_{jk}} \left(c_{\alpha, N_e+N_s+j}\, e^n \pm \sum_{l=N_e+N_s+N_w+2}^{N_e+N_s+N_w+N_{ss}+1} c_{l, N_e+N_s+j} \frac{\partial J}{\partial x_l^n} \right) \quad (18)$$