$$f_\lambda(x) = \frac{1}{1+e^{-\lambda x}}$$

$$\frac{\partial f_\lambda(x)}{\partial x} = \lambda f_\lambda(x)\left(1 - f_\lambda(x)\right)$$

λ pente au point d'inflexion
si $\lambda >> 100$, on a +/- heaviside U, mais dérivable.

On fixe lambda tq $\lambda >> 100$ et on note $f_\lambda(x) = Sig(x)$

$$(M\underline{\underline{x}})^+ = \begin{pmatrix} \sum_{i=1}^{2M} M_{1,i}\, x_i \\ \vdots \\ \sum_{i=1}^{2M} M_{K,i}\, x_i \end{pmatrix} \otimes \begin{pmatrix} Sig\left(\sum_{i=1}^{2M} M_{1,i}\, x_i\right) \\ \vdots \\ Sig\left(\sum_{i=1}^{2M} M_{K,i}\, x_i\right) \end{pmatrix}$$

$$= diag(M\underline{\underline{x}})\, Sig(M\underline{\underline{x}})$$

$$\left(\frac{\partial (M\underline{\underline{x}})^+}{\partial \underline{\underline{x}}}\right)_{i,j} = \frac{\partial}{\partial x_j}\left(\sum_{k=1}^{2M} M_{i,k}\, x_k\right) \times Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right) + \sum_{k=1}^{2M} M_{i,k} x_k \frac{\partial}{\partial x_j}\left(Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right)\right)$$

$$= M_{i,j}\, Sig\left(\sum_{k=1}^{2M} M_{i,k}\, x_k\right) + \sum_{k=1}^{2M} M_{i,k}\, x_k \times \frac{\partial}{\partial x_j} Sig\left(\sum_{k=1}^{2M} M_{i,k}\, x_k\right)$$

$$\left(\frac{\partial}{\partial x_j} Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right) = \frac{\partial}{\partial x_j} Sig(a(x)) = a'(x)\, Sig'(a(x)) \right.$$
$$= M_{i,j}\, \lambda Sig(a(x))\left(1 - Sig(a(x))\right)$$
$$\left. = M_{i,j}\, \lambda\, Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right)\left(1 - Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right)\right) \right)$$

$$= \underbrace{M_{i,j}\, Sig\left(\sum_{k=1}^{2M} M_{i,k}\, x_k\right)}_{\text{gauche}} + \underbrace{\sum_{k=1}^{2M} M_{i,k}\, x_k \times M_{i,j}\, \lambda\, Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right)\left(1 - Sig\left(\sum_{k=1}^{2M} M_{i,k} x_k\right)\right)}_{\text{droite}}$$

$$\sum_{k=1}^{2M} M_{i,k}x_k \times M_{ij}\,\lambda\,\mathrm{Sig}\left(\sum_{k=1}^{2M} M_{i,k}x_k\right)\left(1-\mathrm{Sig}\left(\sum_{k=1}^{2M} M_{i,k}x_k\right)\right)$$

$$\frac{\partial (M\underline{x})^{+}_{\text{droite}}}{\partial \underline{x}}$$

$$\begin{array}{c} \downarrow j \\ \begin{pmatrix} \\ \vec{i} \quad \bullet \\ \\ \end{pmatrix} \end{array} = \begin{pmatrix} M_{i,1} & \cdots & M_{i,2M} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{2M} \end{pmatrix} \times \lambda \begin{pmatrix} M_{i,j} \end{pmatrix} \mathrm{Sig}\left[\begin{pmatrix} M_{i,1} & \cdots & M_{1,2M} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{2M} \end{pmatrix}\right] \times \left(1-\mathrm{Sig}(\dots)\right)$$

$$\frac{\partial (M\underline{x})^{+}_{\text{droite}}}{\partial \underline{x}} = \lambda \begin{pmatrix} (M\underline{x})_1 & \cdots & (M\underline{x})_1 \\ \vdots & & \vdots \\ (M\underline{x})_K & \cdots & (M\underline{x})_K \end{pmatrix} \otimes \begin{pmatrix} M_{1,1} & \cdots & M_{1,2M} \\ \vdots & & \vdots \\ M_{K,1} & & M_{K,2M} \end{pmatrix} \otimes \left[\mathrm{SIG}\left(\mathbb{1}-\mathrm{SIG}\right)\right]$$

$$\mathbb{1} = \begin{array}{c} \overset{2M}{\longleftrightarrow} \\ \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \end{array} \updownarrow K \qquad \mathrm{SIG} = \begin{array}{c} \overset{2M}{\longleftrightarrow} \\ \begin{pmatrix} \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{1,m}x_m\right) & \cdots & \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{1,m}x_m\right) \\ \vdots & & \vdots \\ \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{K,m}x_m\right) & & \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{K,m}x_m\right) \end{pmatrix} \end{array} \updownarrow K$$

$$\lambda (M\underline{x})_i \times M_{i,j} \times \mathrm{Sig}\left(\sum_{k=1}^{2M} M_{i,1}x_k\right)\left(1-\mathrm{Sig}\left(\sum_{k=1}^{2M} M_{i,1}x_k\right)\right)$$

$$\hookrightarrow (M\underline{x})_i = \sum_{m=1}^{2M} M_{i,m}x_m$$

ou

$$\frac{\partial (M\underline{x})^{+}_{\text{droite}}}{\partial \underline{x}} = \lambda\, \mathrm{Diag}(M\underline{x})\, \mathrm{Diag}(\mathrm{SIG}_{:,j})\,(I_K - \mathrm{SIG}_{:,j})\, M$$

$$\frac{\partial (M\underline{x})^{+}_{\text{gauche}}}{\partial \underline{x}}$$

$$\begin{array}{c} \downarrow j \\ \begin{pmatrix} \vec{i} \quad \bullet \end{pmatrix} \end{array} = \begin{pmatrix} M_{1,1} & \cdots & M_{1,2M} \\ \vdots & & \vdots \\ M_{K,1} & \cdots & M_{K,2M} \end{pmatrix} \otimes \mathrm{SIG} = \begin{pmatrix} \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{1,m}x_m\right) & & 0 \\ & \ddots & \\ 0 & & \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{K,m}x_m\right) \end{pmatrix} M$$

$$M_{i,j} \times \mathrm{Sig}\left(\sum_{m=1}^{2M} M_{i,m}x_m\right)$$

$$\frac{\partial (M\underline{x})^+}{\partial \underline{x}} = M \otimes \begin{array}{c} \overset{2M}{\longleftrightarrow} \\ \begin{pmatrix} Sig\left(\sum_{l=1}^{2M} M_{1,l}\, x_l\right) & \cdots & Sig\left(\sum_{l=1}^{2M} M_{1,l}\, x_l\right) \\ \vdots & & \vdots \\ Sig\left(\sum_{l=1}^{2M} M_{n,l}\, x_l\right) & \cdots & Sig\left(\sum_{l=1}^{2M} M_{n,l}\, x_l\right) \end{pmatrix} \end{array} \updownarrow n$$

$$+ \; Dupl(M\underline{x}, 2M) \otimes \lambda M \otimes \left[SIG\,(\mathbb{1} - SIG) \right]$$

on

$\boxed{Diag(SIG_{:,j})\, M}$

$\boxed{+ \lambda\, Diag(M\underline{x})}$

$\boxed{\times Diag(SIG_{:,j})\, Diag(SIG_{:,j})}$

$\times M$

avec $Dupl(A, K)$:

$\mathcal{M}_{n\times 1}(\mathbb{C}) \to \mathcal{M}_{n\times K}(\mathbb{C})$

duplique K fois le vecteur colonne A

$$\text{avec} \quad SIG = \begin{array}{c} \overset{2M}{\longleftrightarrow} \\ \begin{pmatrix} Sig(M\underline{x}) & \cdots & Sig(M\underline{x}) \end{pmatrix} \end{array} \updownarrow n$$

$$\text{et} \quad \mathbb{1} = \begin{array}{c} \overset{2M}{\longleftrightarrow} \\ \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \end{array} \updownarrow n$$

$$\frac{\partial (M\underline{x})^+}{\partial \underline{x}} = Diag(SIG)\, M + \lambda\, Diag(SIG)\, Diag(M\underline{x})\,(I_n - Diag(SIG))\, M$$

$$= Diag(SIG)\left[I_n + \lambda\, Diag(M\underline{x})(I_n - Diag(SIG)) \right] M$$

$$f(x) = \| S(M\underline{x})^+ - I_n A \underline{x} \|^2$$

$$= \sum_{i=1}^{n} \overline{\left(S(M\underline{x})^+ - I_n A \underline{x} \right)_i} \left(S(M\underline{x})^+ - I_n A \underline{x} \right)_i$$

$$\left(S(M\underline{x})^+ - I_n A \underline{x} \right)_i = s_i \sum_{n=1}^{2M} M_{i,n}\, \underline{x}_n \times Sig\left(\sum_{n=1}^{2M} M_{i,n}\, \underline{x}_n \right) - \sum_{n=1}^{2M} (I_n A)_{i,n}\, \underline{x}_n$$

$i \in [\![1; n]\!]$

$$f(x) = \| S(M\underline{x})^+ - I_n A \underline{x} \|^2 = \sum_{i=1}^{n} \left| \left(S(M\underline{x})^+ - I_n A \underline{x} \right)_i \right|^2$$

$$= \sum_{i=1}^{n} \underbrace{\left| s_i \sum_{n=1}^{2M} M_{i,n}\, \underline{x}_n\, Sig\left(\sum_{n=1}^{2M} M_{i,n}\, \underline{x}_n \right) - \sum_{n=1}^{2M} (I_n A)_{i,n}\, \underline{x}_n \right|^2}_{a_i}$$

$$\frac{\partial}{\partial \underline{x}_l} \| S(M\underline{x})^+ - I_n A \underline{x} \|^2 = \sum_{i=1}^{K} \frac{\partial}{\partial \underline{x}_l} (\bar{a}_i\, a_i) = \sum_{i=1}^{K} \bar{a}_i \frac{\partial a}{\partial \underline{x}_l} + a_i \frac{\partial \bar{a}_i}{\partial \underline{x}_l}$$

$\forall\, l \in [\![1; 2M]\!]$

$$\frac{\partial a_i}{\partial \underline{\underline{x}}_l} = \frac{\partial}{\partial \underline{\underline{x}}_l}\left(s_i \sum_{n=1}^{2M} M_{i,n} \underline{\underline{x}}_n \, Sig\left(\sum_{n=1}^{2M} M_{i,n} \underline{\underline{x}}_n\right) - \sum_{n=1}^{2M} (I_K A)_{i,n} \underline{\underline{x}}_n \right)$$

$$= s_i \left(M_{i,l} \, Sig\left(\sum_{n=1}^{2M} M_{i,n} \underline{\underline{x}}_n\right) + \sum_{n=1}^{2M} M_{i,n} x_n M_{i,j} \, \lambda \, Sig\left(\sum_{n=1}^{2M} M_{i,n} x_n\right)\left(1 - Sig\left(\sum_{n=1}^{2M} M_{i,n} x_n\right)\right)\right.$$

$$- (I_K A)_{i,l}$$

$$= \left(S \; Diag\,(SIG)\, M\right)_{i,l} + \lambda\left(S \; Diag\,(M\underline{\underline{x}})\, Diag\,(SIG)\left(I_K - Diag\,(SIG)\right) M\right)_{i,l} - (I_K A)_{i,l}$$

(1)

avec $SIG = \begin{pmatrix} Sig\left(\sum_{n=1}^{2M} M_{1,n} x_n\right) \\ \vdots \\ Sig\left(\sum_{n=1}^{2M} M_{K,n} x_n\right) \end{pmatrix}$

$$\sum_{i=1}^{K} \overline{a_i} \frac{\partial a_i}{\partial \underline{\underline{x}}_l} = \sum_{i=1}^{K} a_i \overline{\frac{\partial a_i}{\partial \underline{\underline{x}}_l}} = \sum_{i=1}^{K} (1)_{i,l} \left(S(M\underline{\underline{x}})^+ - I_K A \underline{\underline{x}}\right)_{i,1}$$

$$= \sum_{i=1}^{K} (1)^{T}_{l,i} \left(S(M\underline{\underline{x}})^+ - I_K A \underline{\underline{x}}\right)_{i,1}$$

Donc $$\frac{\partial \| S(M\underline{\underline{x}})^+ - I_K A \underline{\underline{x}} \|^2}{\partial \underline{\underline{x}}_l} = \left(S \, Diag\,(SIG)\left[I_K + \lambda \, Diag\,(M\underline{\underline{x}})\left(I_K - Diag\,(SIG)\right)\right] M - I_K A\right)^{T}$$

$$\times \overline{\left(S(M\underline{\underline{x}})^+ - I_K A \underline{\underline{x}}\right)}$$

$$+ \overline{\left(S \, Diag\,(SIG)\left[I_K + \lambda \, Diag\,(M\underline{\underline{x}})\left(I_K - Diag\,(SIG)\right)\right] M - I_K A\right)}^{T} \left(S(M\underline{\underline{x}})^+ - I_K A \underline{\underline{x}}\right)$$