Centre Universitaire de Mila	Examen final en « Apprentissage Automatique »
Institut: ST, Département: MI	Durée de l'examen : 1h30 / Mila le : 13-07-2019
Master 1 – STIC	Documents non autorisés

Corrigé type de l'examen final 2018-2019

Exercice 1 : Machines à support de vecteur (04 points)

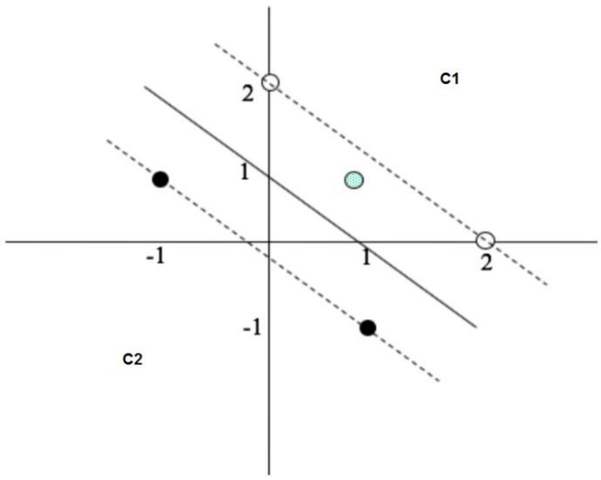
Soient $\boldsymbol{v_1} = (-\mathbf{1},\mathbf{1}), \boldsymbol{v_2} = (\mathbf{2},\mathbf{0}), \boldsymbol{v_3} = (\mathbf{1},-\mathbf{1}), \textbf{and } \boldsymbol{v_3} = (\mathbf{0},\mathbf{2})$ quatre vecteurs de support définissant une ligne de séparation, et soient les coefficients $\boldsymbol{\alpha}$ et le biais $\boldsymbol{w_0}$:

$$\boldsymbol{\alpha_1} = -\mathbf{0.5}, \boldsymbol{\alpha_2} = \mathbf{0.5}, \boldsymbol{\alpha_3} = -\mathbf{0.5}, \boldsymbol{\alpha_4} = \mathbf{0.5} \ \boldsymbol{et} \ \boldsymbol{w_0} = -\mathbf{1}$$

1- Dessiner un graphe des vecteurs de support et tracez l'hyperplan et la marge qu'ils définissent. (Vous pouvez le faire manuellement sans trouver l'équation de l'hyperplan).
2- Donnez la classification SVM résultante de la nouvelle instance $\mathbf{x} = (\mathbf{1},\mathbf{1})$.
3- Trouver le vecteur des poids $\boldsymbol{w}$ associé à l'hyperplan de séparation.
4- En utilisant le vecteur des poids $\boldsymbol{w}$ que vous avez obtenus dans la question 3) et le biais $\boldsymbol{w_0}$, trouvez l'équation de l'hyperplan séparateur : $\boldsymbol{w^T}.\mathbf{x} + \boldsymbol{w_0}$. Quelle est la pente de cette ligne ?

Solution :

1- Graphe des vecteurs de support, hyperplan et de la marge :



2- La classification SVM de la nouvelle instance $\mathbf{x} = (\mathbf{1},\mathbf{1})$:

1ière méthode : D'après le graphe ci-dessus, on peut voir directement que la donnée $\mathbf{x}$ appartient à la classe C1.

2ième méthode : An appliquant la formule suivante :

$$\boldsymbol{f(x) = sgn\left[\sum_{k=1}^{M} \alpha_k (x \cdot x_k) + b\right]}$$

$$\boldsymbol{f}((\mathbf{1},\mathbf{1}) = \textbf{sgn}\,(-\mathbf{0.5}((-\mathbf{1},\mathbf{1})\cdot(\mathbf{1},\mathbf{1})) + \mathbf{0.5}((\mathbf{2},\mathbf{0})\cdot(\mathbf{1},\mathbf{1})) - \mathbf{0.5}((\mathbf{1},-\mathbf{1})\cdot(\mathbf{1},\mathbf{1})) + \mathbf{0.5}((\mathbf{0},\mathbf{2})\cdot(\mathbf{1},\mathbf{1})) - \mathbf{1}) = \textbf{sgn}(\mathbf{0}+\mathbf{1}+\mathbf{0}+\mathbf{1}-\mathbf{1}) = \mathbf{1}$$

3- Le vecteur des poids $\boldsymbol{w}$ associé à l'hyperplan de séparation :

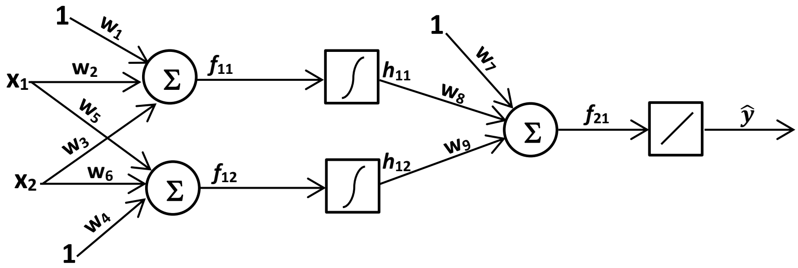
$$-\mathbf{0.5}(-\mathbf{1},\mathbf{1}) + \mathbf{0.5}(\mathbf{2},\mathbf{0}) - \mathbf{0.5}(\mathbf{1},-\mathbf{1}) + \mathbf{0.5}(\mathbf{0},\mathbf{2}) = (\mathbf{1},\mathbf{1})$$

4- Equation de l'hyperplan séparateur : $\boldsymbol{w^T}.\mathbf{x} + \boldsymbol{w_0}$:

- $\boldsymbol{x_2 = -x_1 + 1}$
- **La pente est -1**

Exercice 2 : Réseaux de neurones (08 points)

Soit le réseau de neurones multicouches décrit par le graphe suivant :



1- Donner les formules mathématiques qui déterminent les sorties intermédiaires $\boldsymbol{f_{11}}$, $\boldsymbol{f_{12}}$, $\boldsymbol{h_{11}}$, $\boldsymbol{h_{12}}$, $\boldsymbol{f_{21}}$ ainsi que la sortie finale $\boldsymbol{\hat{y}}$.

2- Soit la fonction d'erreur : $\mathbf{E}(\mathbf{w}) = (\boldsymbol{y} - \boldsymbol{\hat{y}})^2$

En appliquant l'algorithme de propagation en arrière (backpropagation), trouver les expressions des mises à jour des paramètres $\Delta \boldsymbol{w_j}$ pour **j = 1,…,9.**

Solution :

1- Propagation en avant (forward propagartion)

- $f_{11} = w_1 + w_2 x_1 + w_3 x_2$
- $f_{12} = w_4 + w_5 x_1 + w_6 x_2$
- $h_{11} = sigm(f_{11}) = \frac{1}{1+e^{-f_{11}}}$
- $h_{12} = sigm(f_{12}) = \frac{1}{1+e^{-f_{12}}}$
- $\boldsymbol{\hat{y} = f_{21} = w_7 + w_8 h_{11} + w_9 h_{12}}$

2- Propagation en arrière (backpropagartion algorithm)

La fonction d'erreur est donnée par $\mathbf{E}(\mathbf{w}) = (\boldsymbol{y} - \boldsymbol{\hat{y}})^2$

Donc, on aura $\boldsymbol{\frac{\partial E(w)}{\partial w_j} = -2(y - \hat{y})\frac{\partial \hat{y}}{\partial w_j}}$

D'après la propagation en avant, on a : $\boldsymbol{\hat{y} = f_{21} = w_7 + w_8 h_{11} + w_9 h_{12}}$

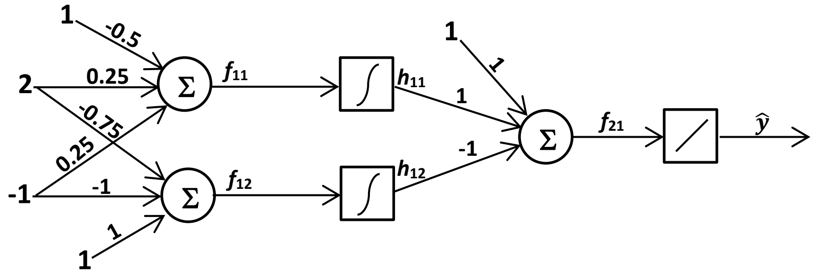
Donc, les dérivées $\frac{\partial \hat{y}}{\partial w_j}$ peuvent être calculées par :

- $\frac{\partial \hat{y}}{\partial w_7} = 1$
- $\frac{\partial \hat{y}}{\partial w_8} = h_{11}$
- $\frac{\partial \hat{y}}{\partial w_9} = h_{12}$
- $\frac{\partial \hat{y}}{\partial w_1} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_1} = w_8 h_{11}(1 - h_{11})$
- $\frac{\partial \hat{y}}{\partial w_2} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_2} = w_8 h_{11}(1 - h_{11})x_1$
- $\frac{\partial \hat{y}}{\partial w_3} = \frac{\partial \hat{y}}{\partial h_{11}}\frac{\partial h_{11}}{\partial f_{11}}\frac{\partial f_{11}}{\partial w_3} = w_8 h_{11}(1 - h_{11})x_2$
- $\frac{\partial \hat{y}}{\partial w_4} = \frac{\partial \hat{y}}{\partial h_{12}}\frac{\partial h_{12}}{\partial f_{12}}\frac{\partial f_{12}}{\partial w_4} = w_9 h_{12}(1 - h_{12})$
- $\frac{\partial \hat{y}}{\partial w_5} = \frac{\partial \hat{y}}{\partial h_{12}}\frac{\partial h_{12}}{\partial f_{12}}\frac{\partial f_{12}}{\partial w_5} = w_9 h_{12}(1 - h_{12})x_1$
- $\frac{\partial \hat{y}}{\partial w_6} = \frac{\partial \hat{y}}{\partial h_{12}}\frac{\partial h_{12}}{\partial f_{12}}\frac{\partial f_{12}}{\partial w_6} = w_9 h_{12}(1 - h_{12})x_2$

En fin, la mise à jour de chaque paramétré est donnée par la formule : $\Delta \boldsymbol{w_j} = \boldsymbol{\alpha(y - \hat{y})\frac{\partial \hat{y}}{\partial w_j}}$

<u>Exercice 3 :</u> Réseaux de neurones (08 points)

- Soit le réseau de neurones multicouches de l'exercice 2 décrit par le graphe suivant :



- Soit la donnée **x = (2, -1), y = 1**

1- Calculer les sorties intermédiaires $\boldsymbol{f_{11}}$, $\boldsymbol{f_{12}}$, $\boldsymbol{h_{11}}$, $\boldsymbol{h_{12}}$, $\boldsymbol{f_{21}}$ ainsi que la sortie finale $\hat{\boldsymbol{y}}$.

2- Calculer les mise à jour $\Delta \boldsymbol{w_j}$ ainsi que les paramètres $\boldsymbol{w_j}$ pour **j = 1,…,9** après une itération de mise à jour (en considérant le paramètre d'apprentissage $\boldsymbol{\alpha} = \mathbf{0.1}$).

NB : La précision des calculs numériques est fixée à 4 chiffres après la virgule.

<u>Solution :</u>

1- Calcul des sorties intermédiaires $\mathbf{f_{11}}, \mathbf{f_{12}}$, $\mathbf{h_{11}}, \mathbf{h_{12}}, \mathbf{f_{21}}, \boldsymbol{et}\ \hat{\mathbf{y}}$**:**

- $\mathbf{f_{11}} = (-\mathbf{0.5}) * \mathbf{1} + \mathbf{0.25} * \mathbf{2} + \mathbf{0.25} * (-\mathbf{1}) = -\mathbf{0.25}$
- $\mathbf{f_{12}} = \mathbf{1} * \mathbf{1} + (-\mathbf{0.75}) * \mathbf{2} + (-\mathbf{1}) * (-\mathbf{1}) = \mathbf{0.50}$
- $\mathbf{h_{11}} = \mathbf{sigm}(\mathbf{f_{11}}) = \frac{\mathbf{1}}{\mathbf{1+e^{-f_{11}}}} = \frac{\mathbf{1}}{\mathbf{1+e^{-(-0.25)}}} = \mathbf{0.4378}$
- $\mathbf{h_{12}} = \mathbf{sigm}(\mathbf{f_{21}}) = \frac{\mathbf{1}}{\mathbf{1+e^{-f_{21}}}} = \frac{\mathbf{1}}{\mathbf{1+e^{-0.5}}} = \mathbf{0.6225}$
- $\mathbf{f_{21}} = \mathbf{1} * \mathbf{1} + \mathbf{1} * \mathbf{0.4378} + (-\mathbf{1}) * \mathbf{0.6225} = \mathbf{0.8154}$
- $\hat{\mathbf{y}} = \mathbf{f_{21}} = \mathbf{0.8154}$

2- Calculer les paramètres $\boldsymbol{w_j}$ pour **j = 1, …, 9 :**

Soit le paramètre d'apprentissage $\boldsymbol{\alpha} = \mathbf{0.1}$, On a $\boldsymbol{w_j} = \boldsymbol{w_j} + \Delta \boldsymbol{w_j}$**.**

2-1- Calcul des $\frac{\partial \hat{y}}{\partial w_j}$ pour j = 1,…,9:

- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_1}} = \mathbf{1} * \mathbf{0.4378} * (\mathbf{1} - \mathbf{0.4378}) = \mathbf{0.2461}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_2}} = \mathbf{1} * \mathbf{0.4378} * (\mathbf{1} - \mathbf{0.4378}) * \mathbf{2} = \mathbf{0.4923}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_3}} = \mathbf{1} * \mathbf{0.4378} * (\mathbf{1} - \mathbf{0.4378}) * (-\mathbf{1}) = -\mathbf{0.2461}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_4}} = -\mathbf{1} * \mathbf{0.6225} * (\mathbf{1} - \mathbf{0.6225}) = -\mathbf{0.2350}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_5}} = -\mathbf{1} * \mathbf{0.6225} * (\mathbf{1} - \mathbf{0.6225}) * \mathbf{2} = -\mathbf{0.4700}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_6}} = -\mathbf{1} * \mathbf{0.6225} * (\mathbf{1} - \mathbf{0.6225}) * (-\mathbf{1}) = \mathbf{0.2350}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_7}} = \mathbf{1}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_8}} = \mathbf{0.4378}$
- $\frac{\partial \hat{\boldsymbol{y}}}{\partial \boldsymbol{w_9}} = \mathbf{0.6225}$

2-2- Calcul des Δw_j pour j = 1,…,9:

- $\Delta w_1 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_1} = 0.1 * (1 - 0.8154) * 0.2461 = 0.0045$
- $\Delta w_2 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_2} = 0.1 * (1 - 0.8154) * 0.4923 = 0.0091$
- $\Delta w_3 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_3} = 0.1 * (1 - 0.8154) * (-0.2461) = -0.0045$
- $\Delta w_4 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_4} = 0.1 * (1 - 0.8154) * (-0.2350) = -0.0043$
- $\Delta w_5 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_5} = 0.1 * (1 - 0.8154) * (-0.4700) = -0.0087$
- $\Delta w_6 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_6} = 0.1 * (1 - 0.8154) * 0.2350 = 0.0043$
- $\Delta w_7 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_7} = 0.1 * (1 - 0.8154) * 1.0000 = 0.0185$
- $\Delta w_8 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_8} = 0.1 * (1 - 0.8154) * 0.4378 = 0.0081$
- $\Delta w_8 = \alpha(y-\hat{y})\frac{\partial\hat{y}}{\partial w_8} = 0.1 * (1 - 0.8154) * 0.6225 = 0.0115$

2-3- Calcul des w_j pour j = 1,…,9:

- $w_1 = w_1 + \Delta w_1 = -0.5 + 0.0045 = -0.4955$
- $w_2 = w_2 + \Delta w_2 = 0.25 + 0.0091 = 0.2591$
- $w_3 = w_3 + \Delta w_3 = 0.25 - 0.0045 = 0.2455$
- $w_4 = w_4 + \Delta w_4 = 1 - 0.0043 = 0.9957$
- $w_5 = w_5 + \Delta w_5 = -0.75 - 0.0087 = -0.7587$
- $w_6 = w_6 + \Delta w_6 = -1 + 0.0043 = -0.9957$
- $w_7 = w_7 + \Delta w_7 = 1 + 0.0185 = 1.0185$
- $w_8 = w_8 + \Delta w_8 = 1 + 0.0081 = 1.0081$
- $w_8 = w_8 + \Delta w_8 = -1 + 0.0115 = -0.9885$