Université de Mila
Département: informatique

Apprentissage Automatique :Travaux dirigées

Année universitaire 2025-2026

exercice01

- Exercice 01 : Montrer que la fonction sigmoïde satisfait la propriété suivante :

$$\mathbf{sigm}(-\mathbf{a}) = \mathbf{1} - \mathbf{sigm}(\mathbf{a})$$

- et que son inverse est donné par
-

$$\mathbf{sigm}^{-1}(\mathbf{y}) = \mathbf{ln}\{y/(1-y)\}$$

Exercice01-solution

$$\mathbf{sigm(-a) \; = \; 1 - sigm(a)}$$

La fonction sigmoïde est définie par :

$$\sigma(a) = \frac{1}{1+e^{-a}}$$

- **1. Vérification de la propriété**

$$\sigma(-a) = \frac{1}{1+e^{a}}$$

On veut montrer que : $\sigma(-a) = 1 - \sigma(a)$

On calcule :

$$1 - \sigma(a) = 1 - \frac{1}{1+e^{-a}} = \frac{(1+e^{-a}) - 1}{1+e^{-a}} = \frac{e^{-a}}{1+e^{-a}}$$

Puis, en multipliant numérateur et dénominateur par e^a, on obtient :

Donc : $\frac{1}{1+e^{a}} = \sigma(-a)$

$$\boxed{\sigma(-a) = 1 - \sigma(a)}$$

Exercice01-solution

- 2. Détermination de l'inverse

$$\boxed{\mathbf{sigm^{-1}(y) = ln\{y/(1-y)\}}}$$

- On a: $$y = \frac{1}{1+e^{-a}}$$
- Ensuite: $$y(1+e^{-a}) = 1$$
$$y + ye^{-a} = 1$$
$$ye^{-a} = 1 - y$$
$$e^{-a} = \frac{1-y}{y}$$
- On applique le logarithme naturel :
- $$-a = \ln\left(\frac{1-y}{y}\right)$$
$$a = -\ln\left(\frac{1-y}{y}\right)$$

D'où l'inverse : $$\sigma^{-1}(y) = \ln\left(\frac{y}{1-y}\right)$$

exercice02

- Soit la fonction sigmoïde définie par : $$\sigma(a) = \frac{1}{1 + \exp(-a)}$$

- Vérifier que la relation pour la dérivée de la fonction sigmoïde définie par :

$$\frac{d\sigma}{da} = \sigma(1 - \sigma)$$

Exercice02-solution

$$\boxed{\frac{d\sigma}{da} = \sigma(1-\sigma)}$$

- La fonction sigmoïde est donnée par : $\sigma(a) = \frac{1}{1+\exp(-a)}.$
- Exprimer σ(a) sous une forme différente $\sigma(a) = (1+e^{-a})^{-1}$
- Calculer la dérivée dσ/da

$$\frac{d\sigma}{da} = \frac{d}{da}\left((1+e^{-a})^{-1}\right) = -1\cdot(1+e^{-a})^{-2}\cdot\frac{d}{da}(1+e^{-a})$$

$$= -(1+e^{-a})^{-2}\cdot(-e^{-a}) = \frac{e^{-a}}{(1+e^{-a})^2}$$

- On a: $1-\sigma(a) = 1 - \frac{1}{1+e^{-a}} = \frac{e^{-a}}{1+e^{-a}}$
- Multiplions maintenant σ(a) et 1−σ(a):

$$\sigma(a)(1-\sigma(a)) = \left(\frac{1}{1+e^{-a}}\right)\left(\frac{e^{-a}}{1+e^{-a}}\right) = \frac{e^{-a}}{(1+e^{-a})^2}$$

- On voit que :

$$\boxed{\frac{d\sigma}{da} = \frac{e^{-a}}{(1+e^{-a})^2} = \sigma(a)(1-\sigma(a))}$$

Exercice03

Pour un ensemble $\{\phi_n, t_n\}$, où $t_n \in \{0,1\}$ et $\phi_n = \phi(\mathrm{x}_n)$, avec $n = 1, \dots, N$, la fonction de vraisemblance peut écrite sous la forme :

$$p(\mathbf{t}|\boldsymbol{w}) = \prod_{n=1}^{N} y_n^{t_n}\{1 - y_n\}^{1-t_n}$$

avec $t = (t_1, .., t_N)^T$ et $y_n = p(C_1|\phi_n)$. On peut définir une fonction d'erreur par prendre le logarithm negative de la fonction de vraisemblence, qui va donner la fonction d'erreur (cross entropy) de la forme :

$$E(\boldsymbol{w}) = -\ln p(\mathbf{t}|\boldsymbol{w}) = -\sum_{n=1}^{N}\{t_n \ln y_n + (1 - t_n)\ln(1 - y_n)\}$$

En utilisant le résultat de l'exercice 02 pour la dérivée du sigmoïde logistique, montrer que la dérivée de la fonction d'erreur pour le modèle de régression logistique est donnée par

$$\nabla E(\boldsymbol{w}) = \sum_{n=1}^{N}(1 - t_n)\phi_n$$

Avec $y_n = \sigma(a_n)$ et $a_n = w^T\phi_n$

Exercice 03-solution

- La fonction d'erreur peut être écrite comme :

$$E(\mathbf{w}) = -\sum_{n=1}^{N} \{t_n \ln y_n + (1 - t_n) \ln(1 - y_n)\} .$$

Pour un échantillon *n* donné, la contribution à l'erreur est :

$$E_n = -\left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n)\right] .$$

1. Nous commençons par calculer la dérivée partielle de E_n par rapport à y_n :

$$\frac{\partial E_n}{\partial y_n} = \frac{\partial}{\partial y_n} \left(-t_n \ln y_n - (1 - t_n) \ln(1 - y_n)\right) .$$

Exercice 03-solution

$$\frac{\partial E_n}{\partial y_n} = \frac{\partial}{\partial y_n}\left(-t_n \ln y_n - (1-t_n)\ln(1-y_n)\right).$$

- En appliquant la règle des dérivées, nous obtenons :

$$\frac{\partial E_n}{\partial y_n} = -t_n \cdot \frac{1}{y_n} - (1-t_n)\cdot\frac{-1}{1-y_n} \qquad = -\frac{t_n}{y_n} + \frac{1-t_n}{1-y_n}$$

$$= \frac{(1-t_n)y_n - t_n(1-y_n)}{y_n(1-y_n)} \quad \boxed{= \frac{y_n - t_n}{y_n(1-y_n)}.}$$

Equ:1

2. Calcul de $\frac{\partial y_n}{\partial a_n}$

La probabilité y_n est définie comme : $$y_n = \sigma(a_n) = \frac{1}{1+\exp(-a_n)}.$$

Exercice 03-solution

- Nous savons que la dérivée de la fonction sigmoïde est

$$y_n = \sigma(a_n) = \frac{1}{1+\exp(-a_n)}. \qquad \frac{\partial \sigma(a_n)}{\partial a_n} = \sigma(a_n)(1-\sigma(a_n)).$$

$$\boxed{\frac{\partial y_n}{\partial a_n} = \frac{\partial \sigma(a_n)}{\partial a_n} = y_n(1-y_n).}$$

Equation: 02

- Calcul de ∇a_n :
- Le score linéaire a_n est défini comme : $a_n = \mathbf{w}^T \phi_n,$
- où ϕ_n est le vecteur de caractéristiques associé à l'échantillon *n*. La dérivée de a_n par rapport à **w** est :

$$\boxed{\nabla a_n = \frac{\partial a_n}{\partial \mathbf{w}} = \phi_n.}$$

Equation: 03

Exercice 03-solution

- En combinant (1), (2) et (3) utilisant la règle de chaine, on obtient :

$$\nabla E = \sum_{n=1}^{N} \frac{\partial E}{\partial y_n} \cdot \frac{\partial y_n}{\partial a_n} \cdot \nabla a_n.$$

$$\nabla E = \sum_{n=1}^{N} \left(\frac{y_n - t_n}{y_n(1 - y_n)} \cdot y_n(1 - y_n) \cdot \phi_n \right).$$

$$\nabla E = \sum_{n=1}^{N} (y_n - t_n)\phi_n.$$

Rappelle

Appliquer la règle de la chaîne((dérivation des fonctions composées).)

La règle de la chaîne dit que :

$$\frac{d}{da}\left[(u(a))^n\right] = n \cdot (u(a))^{n-1} \cdot \frac{du}{da}.$$

rappelle

$$\frac{d}{dy_n}\ln(1-y_n) = \frac{1}{1-y_n}\cdot\frac{d}{dy_n}(-y_n)$$

Dérivée de $-y_n$:

La dérivée de $-y_n$ par rapport à y_n est simplement -1.

Résultat final :

En combinant tout cela, nous obtenons la dérivée de $\ln(1-y_n)$ par rapport à y_n :

$$\frac{d}{dy_n}\ln(1-y_n) = \frac{-1}{1-y_n}$$