Nicholas Massad 23 Janvier 2023

Masn2601

#1

a) Vrai

$$X = X^T \rightarrow \; x_{ij} = x_{ji} \quad symétrique$$

$$A = \begin{pmatrix} a_{11} & \cdots & a_{N1} \\ \vdots & \ddots & \vdots \\ a_{1N} & \cdots & a_{NN} \end{pmatrix} \quad A^T = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix}$$

$$X = A^T A = \begin{pmatrix} \sum_{j=1}^{N} a_{j1}^T a_{1j} & \cdots & \sum_{j=1}^{N} a_{j1}^T a_{Nj} \\ \vdots & \ddots & \vdots \\ \sum_{j=1}^{N} a_{jN} a_{1j} & \cdots & \sum_{j=1}^{N} a_{jN}^T a_{Nj} \end{pmatrix}$$

$$\sum_{j=1}^{N} a_{jN}^T a_{1j} = \sum_{j=1}^{N} a_{j1}^T a_{Nj} \rightarrow x_{ij} = x_{ji} \rightarrow X = X^T$$

b) Vrai

$$\alpha^T X^T X \alpha \geq 0$$

$$\alpha^T X^T X \alpha = \sum_{j=1}^{N} \sum_{i=1}^{N} \alpha_j x_{ji} x_{ij} \alpha_i = \sum_{j=1}^{N} \sum_{i=1}^{N} \alpha_i^2 x_{ij}^2 \; \geq 0$$

c) Faux, sauf cas particulier ou det(A) = 1 --> ab-cd = 1

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad adj(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$A^{-1} = \frac{1}{\det(A)} adj(A) = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Donc le numéro est faux dans un cas général et vrai dans le cas ou le det(A) est 1

d) Faux

$$A^{-1} \neq \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 2 & 0 & 0 & | & 1 & 0 & 0 \\ 0 & -2 & 0 & | & 0 & 1 & 0 \\ 0 & 0 & 3 & | & 0 & 0 & 1 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & | & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & | & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & | & 0 & 0 & \frac{1}{3} \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \neq \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

#2

a) Calculez l'entropie d'une variable aléatoire uniforme discrète X prenant des valeurs dans {1, 2, ..., m}.

$$P(x) = \frac{1}{m}$$

$$H(X) = -\sum_{x}^{m} \frac{1}{m} \log\left(\frac{1}{m}\right) = -m\left(\frac{1}{m}\log\left(\frac{1}{m}\right)\right)$$

$$= -\log\left(\frac{1}{m}\right) = \log(m)$$

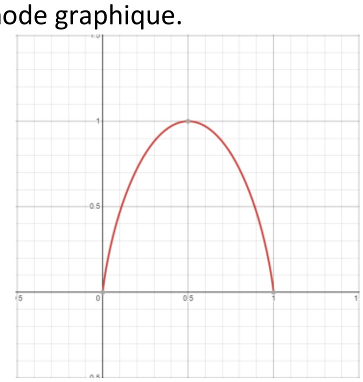
b) Calculez l'entropie d'une variable al´eatoire de Bernoulli X avec le param`etre ϕ. Rappelons que si X ∼ Bernoulli(ϕ) alors p(x) = ϕ x (1 − ϕ) 1−x , x = 0, 1.

$$P(X) = \phi^x(1-\phi)^{1-x}$$

$$H(X) = -\sum_{x=0}^{1} \phi^x(1-\phi)^{1-x} \log(\phi^x(1-\phi)^{1-x}), \quad \phi = [0,1]$$

$$= -[(1-\phi)\log(1-\phi) + \phi\log(\phi)]$$

c) Puisque ϕ va de 0 à 1 et que l'entropie va de 0 a un on observe 0 à 0 et 1 et 1 à 0.5. Le maximum est donc à 0.5. On pourrait utiliser le dérivé a 0 pour obtenir cette valeur mais puisque ce n'est pas spécifier j'ai utilisé une méthode graphique.



d) X et Y indépendant p(x,y) = p(x)p(y)

$$H(X,Y) = -\sum_x \sum_y p(x,y) \log\big(p(x,y)\big)$$
$$= -\sum_x \sum_y p(x)p(y) \log(p(x)p(y))$$
$$= -\sum_x \sum_y p(x)p(y) \left(\log(p(x)) + \log(p(y))\right)$$

$$= -\sum_x p(x)\log(p(x)) \sum_y p(y) - \sum_y p(y)\log(p(y)) \sum_x p(x)$$
$$= -\sum_x p(x)\log\big(p(x)\big) * 1 - \sum_y p(y)\log\big(p(y)\big) * 1$$
$$= H(X) + H(Y)$$

#3

a)

$$E(Y) = \sum_{i=1}^{m} y^{(i)} \ \phi^{y^{(i)}}(1-\phi)^{1-y^{(i)}}$$
$$VAR(Y) = \frac{\sum_{i=1}^{m}\left(y^i - E(y)\right)^2}{m}$$

b) Calculer l'estimateur de maximum de vraisemblance de ϕ

$$l(\phi) = \log\left(\prod_{i=1}^{m} p\big(y^i\big)\right) = \log\left(\prod_{i=1}^{m} \phi^{y^{(i)}}(1-\phi)^{1-y^{(i)}}\right)$$
$$= \sum_{i=1}^{m} \log(\phi)^{y^i} + \log(1-\phi)^{1-y^i} = \sum_{i=1}^{m} \mathrm{y^i}\log(\phi) + (1-y^i)\log(1-\phi)$$

Afin d'obtenir la plus petite valeur de loss selon la fonction l(ϕ) on cherche ϕ lorsque la dérivé est a 0

$$\frac{d}{d\phi}\sum_{i=1}^{m} \mathrm{y^i}\log(\phi) + (1-y^i)\log(1-\phi) = 0$$
$$\phi = arg\min_{\phi} \sum_{i=1}^{m} \frac{y^i}{\phi ln2} + \frac{1-y^i}{1-\phi ln2}$$
$$\phi = 0.5$$

#4

a) On cherche a démontrer que d/dx(σ(x)) = σ(x)*(1- σ(x))

$$\frac{d}{dx}\sigma(\mathrm{x}) = \frac{d}{dx}\frac{1}{1+e^{-x}} = \frac{d}{dx}\left(1+e^{-x}\right)^{-1}$$
$$= -\frac{1}{(1+e^{-x})^{-2}}\frac{d}{dx}\left(1+e^{-x}\right) = -\frac{1}{(1+e^{-x})^{-2}}\left(-e^{-x}\right)$$
$$= \frac{1}{1+e^{-x}} * \frac{e^{-x}}{1+e^{-x}} = \frac{1}{1+e^{-x}} * \frac{1+e^{-x}-1}{1+e^{-x}} = \frac{1}{1+e^{-x}} * \frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}$$

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

b) Soient x = (x1, x2)^T ∈ R^2 , w = (w1, w2)^T ∈ R^2 et g(x) = x^T w = x.w. Montrez que le gradient ∇_xg(x) = w, ou ∇_xg(x) = (∂g(x)/∂x1 , ∂g(x)/∂x2)^T

$$g(x) = x_1 w_1 + x_2 w_2$$

$$\nabla_x g(x) = \left(\frac{\partial g(x)}{\partial x_1}, \frac{\partial g(x)}{\partial x_2}\right), \quad \frac{\partial g(x)}{\partial x_1} = w_1, \quad \frac{\partial g(x)}{\partial x_2} = w_2$$

$$\nabla_x g(x) = (w_1, w_2)^T = \boldsymbol{w}$$

c) Soient x = (x1, x2)^T ∈ R^2 , A = $\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$, et Q(x) = x^TAx. Calculez ∇_xQ(x) et comparez-le à (A + AT)x.

$$Q(x) = (x_1 a_{11} + x_2 a_{12}, x_1 a_{21} + x_2 a_{22})(x_1, x_2)^T = x_1^2 + 3x_1x_2 + x_2^2$$

$$(\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{x} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (2x_1 + 3x_2, 3x_1 + 2x_2)$$

$$\nabla_x g(x) = \left(\frac{\partial g(x)}{\partial x_1}, \frac{\partial g(x)}{\partial x_2}\right) = (2x_1 + 3x_2, 3x_1 + 2x_2$$

$$\nabla_x g(x) = (\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{x}$$

d) En utilisant le résultat de (c) déduisez ∇x $\|x\|_2^2$, ou $\|x\|_2^2$= x^T x = x.x

$$\|x\|_2^2 = (x_1^2 + x_2^2)$$

$$\nabla_x \|x\|_2^2 = (2x_1, 2x_2)$$