

NB les notations sont celles du cours. Le prêt de tout objet est strictement interdit.

Problème 1

A) Régression linéaire simple

On a relevé sur 12 individus les observations d'un couple de caractères (X, Y) données ci-dessous.

x_i	154	145	165	163	148	161	151	176	159	159	151	155
y_i	67	66	69	69	68	69	70	74	70	73	69	71

- 1) Coefficient de corrélation linéaire entre X et Y ?
- 2) Droite de régression de Y en X ? et y_{25} si $x_{25} = 171$?
- 3) Soit M_i le point du plan $\mathbb{R}^2$ de coordonnées (x_i, y_i) et $\Delta : y = \alpha + \beta x$ une droite du plan. Soit O l'origine du repère de $\mathbb{R}^2$ et P_i la projection orthogonale de M_i sur Δ . On note par $d(M, P)$ la distance euclidienne entre deux points M et P et on souhaite déterminer la droite Δ qui rend minimum la somme

$$S = \sum_{i=1}^n d^2(M_i, P_i).$$

a) Montrer que

$$\min \sum_{i=1}^n \|M_i P_i\|^2 \iff \max \sum_{i=1}^n \|O P_i\|^2$$

et que

$$\max \sum_{i=1}^n \|O P_i\|^2 \iff \max_{u' u = 1} u' V u$$

où u est un vecteur directeur de la droite Δ et V est une matrice carrée à déterminer.

b) A.N. Trouver l'équation de Δ dans le cas des observations ci-dessus.

B) Régression linéaire multiple

On considère 5 observations de X_1, X_2, Y données par

X_1	0.12	0.14	0.31	0.37	0.49
X_2	29	33	17	21	12
Y	21	24.3	12.7	15.6	9

- 1) Rappeler la formule donnant l'estimateur $\hat{\beta}$ de β du modèle linéaire multiple

$$Y = X\beta + \varepsilon$$

où X, Y, β et ε sont à préciser.

2) Rappeler la formule de l'erreur empirique MSE du modèle ajusté $\hat{Y} = X\hat{\beta}$ avec la fonction de coût $l(y, y') = (y - y')^2$.

3) A.N. En utilisant les observations ci-dessus, donner l'expression du modèle d'apprentissage à l'aide de l'estimateur $\hat{\beta}$ et la valeur du MSE .

C) Régression linéaire régularisée

1) Rappeler la formule de la régularisation avec la fonction de coût précédente. On pourra utiliser la fonction Argmin ou Argmax.

1) En déduire celles de la régression Ridge, Lasso et Elastic Net.

Problème 2

A) Réseaux de neurones

1) On donne $x_1 = (0, 0, 1, 1)$ et $x_2 = (0, 1, 0, 1)$.

1) Calculer $\hat{y} = f(x_1, x_2)$ et le comparer à $y = (0, 0, 0, 1)$ à l'aide d'un perceptron avec les poids $w_0 = -1.5$, $w_1 = w_2 = 1$ et la fonction d'activation sign ($\text{sign}(x) = 1$ si $x > 0$, $\text{sign}(x) = -1$ si $x < 0$).

2) Calculer $\hat{y} = f(x_1, x_2)$ et le comparer à $y = (0, 1, 1, 0)$ à l'aide d'un perceptron avec une couche cachée à deux neurones avec les poids $w_{01}^1 = -0.5$, $w_{02}^1 = -0.5$, $w_{11}^1 = 1$, $w_{12}^1 = -1$, $w_{21}^1 = -1$, $w_{22}^1 = 1$, $w_0^2 = -0.5$, $w_1^2 = 1$, $w_2^2 = 1$ et la fonction d'activation sign.

3) On considère un réseau de neurones décrit par le graphe ci-dessous.

a) Donner les expressions de a_1^1 , a_2^1 et $\hat{y}$.

b) On prend pour fonction d'activation la sigmoïde. Donner la valeur de $\hat{y}$ pour $x = 2$, $w_1 = 0.5$, $w_2 = -1$, $w_3 = -2$, $w_4 = 1.5$, $w_5 = 1$, $w_6 = 1$, $w_7 = -1$. L'entrée de w_5 est 1.

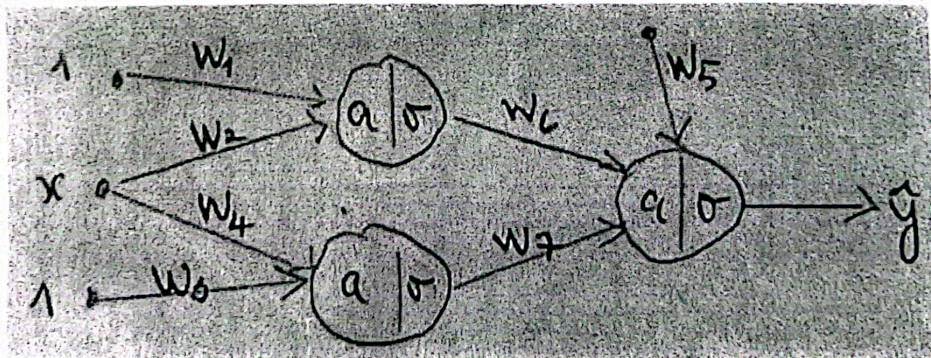


FIGURE 1 – Graphe du réseau

B) Support vector machine

Soit $\mathcal{D} = \{(x_i, y_i), i = 1 \dots n\}$ un jeu de données avec $x_i \in \mathbb{R}^p$, $y_i \in \{-1, 1\}$. Soit

$$H : \langle w, x \rangle + b = 0$$

l'hyperplan séparateur optimal de $\mathcal{D}$.

- 1) Rappeler la formulation primale de la SVM à marge rigide de $\mathcal{D}$.
- 2) Rappeler la formulation duale de la SVM à marge souple de $\mathcal{D}$.