



## TD 1 – Optimisation

### Formulation mathématique

#### ▷ Exercice 1. Régression linéaire simple

Soit  $n$  points expérimentaux  $M_i = (x_i, y_i)$  pour  $i = 1, \dots, n$ . On considère le modèle suivant :  $y(x) = \beta_0 + \beta_1 x$ . On désire estimer les paramètres par la méthode des moindres carrés.

**1.1.** On veut estimer les paramètres par les moindres carrés. Écrire le problème sous la forme :

$$\begin{cases} \min f(\beta) = \frac{1}{2} \|r(\beta)\|^2 = \frac{1}{2} \|y - X\beta\|^2 \\ \beta \in \mathbb{R}^p \end{cases}$$

On donnera les valeurs de  $X$  et de  $y$  et à quoi correspond  $\beta$ .

**1.2.** On souhaite maintenant trouver la meilleure droite au sens des moindres carrés qui passe par l'origine. Écrire le problème d'optimisation.

▷ **Exercice 2** (Courbe étalon). La première étape d'un dosage radio-immunologique consiste à établir une courbe étalon. Un dosage repose sur l'hypothèse qu'une hormone et son *isotope marqué* se comportent de façon équivalente vis-à-vis de leur anticorps spécifique : lorsque l'on met en présence une quantité déterminée d'anticorps, une quantité déterminée d'hormone radioactive et une quantité variable d'hormone froide, la dose de complexe anticorps-hormone marquée en fin de réaction est d'autant plus faible que la quantité d'hormone froide est importante. Néanmoins, la relation qui existe entre la dose d'hormone froide mise en réaction et la radioactivité de complexe extrait n'est pas stable et doit être appréciée dans chaque situation expérimentale. C'est l'objet de l'établissement de la courbe étalon, à partir d'une gamme de dilutions connues d'une quantité déterminée de l'hormone à doser. La table 1 donne les données recueillies pour une telle courbe dans le cas d'un dosage du cortisol : on a mesuré la radioactivité du complexe (en coups par minute ou cpm). On considère le modèle suivant :

$$y(x) = \beta_2 + \frac{\beta_1 - \beta_2}{(1 + \exp(\beta_3 + \beta_4 x))^{\beta_5}}. \quad (1)$$

On désire estimer les paramètres par les moindres carrés (attention, il y a pour chaque dose 4 observations de  $y$ ). On notera  $(x_i)_{i=1,\dots,16}$  (respectivement  $(y_{i,j})_{i=1,\dots,16; j=1,\dots,4}$ ) les éléments de la première colonne (respectivement des 4 dernières colonnes) de la table 1 et  $r_{i,j}(\beta)$  le résidu liés au point  $(x_i, y_{i,j})$ .

**2.1.** Écrire le résidu lié au point (0.04, 2378).

**2.2.** 1. Quelle est la dimension du vecteur des paramètres  $\beta$ .

2. Quel est le nombre de points  $n$  ?

**2.3.** Écrire le problème d'optimisation des paramètres par les moindres carrés.

| Dose en ng/.1 ml | Réponse en c.p.m. |      |      |      |
|------------------|-------------------|------|------|------|
| 0                | 2868              | 2785 | 2849 | 2805 |
| 0                | 2779              | 2588 | 2701 | 2752 |
| 0.02             | 2615              | 2651 | 2506 | 2498 |
| 0.04             | 2474              | 2573 | 2378 | 2494 |
| 0.06             | 2152              | 2307 | 2101 | 2216 |
| 0.08             | 2114              | 2052 | 2016 | 2030 |
| 0.1              | 1862              | 1935 | 1800 | 1871 |
| 0.2              | 1364              | 1412 | 1377 | 1304 |
| 0.4              | 910               | 919  | 855  | 875  |
| 0.6              | 702               | 701  | 689  | 696  |
| 0.8              | 586               | 596  | 561  | 562  |
| 1                | 501               | 495  | 478  | 493  |
| 1.5              | 392               | 358  | 399  | 394  |
| 2                | 330               | 351  | 343  | 333  |
| 4                | 250               | 261  | 244  | 242  |
| 100              | 131               | 135  | 134  | 133  |

TABLE 1 – Données pour un dosage de Cortisol

|        |         |         |         |         |         |         |         |
|--------|---------|---------|---------|---------|---------|---------|---------|
| années | 1900    | 1910    | 1920    | 1930    | 1940    | 1950    | 1960    |
| pop.   | 75.995  | 91.972  | 105.711 | 123.203 | 131.669 | 150.697 | 179.323 |
| années | 1970    | 1980    | 1990    | 2000    |         |         |         |
| pop.   | 203.212 | 226.505 | 249.633 | 281.422 |         |         |         |

TABLE 2 – Données provenant de "U.S. Census"

▷ **Exercice 3.** On donne ci-dessous la population des États Unis pour les années 1900 à 2000<sup>1</sup>.

On désire ajuster ces données sur le modèle

$$y(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$$

On utilise la méthode des moindres carrés.

**3.1.** Donner la formulation mathématique de ce problème et visualiser la fonction à minimiser.

**3.2.** Écrire le problème sous la forme :

$$\begin{cases} \text{Min} & f(\beta) = \frac{1}{2} \|r(\beta)\|^2 = \frac{1}{2} \|y - X\beta\|^2 \\ \beta \in \mathbb{R}^p \end{cases}$$

On donnera les valeurs de  $X$  et de  $y$  et à quoi correspond  $\beta$ .

1. Exemple provenant du cours de Cleve Moler page 4 chap. 5 "Numerical computing with Matlab"

▷ **Exercice 4. Maintenance d'un réseau de distribution**<sup>2</sup>

Un ingénieur responsable de la maintenance d'un réseau de distributeurs de boissons aimerait prédire le temps nécessaire pour l'approvisionner. Le service d'approvisionnement comprend le remplissage des machines et leurs réglages éventuels. Deux variables influencent ce temps : le nombre de caisses à charger et la distance parcourue par l'employé pour approvisionner l'ensemble des machines. Le responsable dispose de 25 observations, qui sont résumées dans le tableau suivant :

| Obs. | Temps | Nb caisses | Dist. | Obs. | Temps | Nb. caisses | Dist. |
|------|-------|------------|-------|------|-------|-------------|-------|
| 1    | 16.68 | 7          | 560   | 13   | 13.50 | 4           | 255   |
| 2    | 11.50 | 3          | 220   | 14   | 19.75 | 6           | 462   |
| 3    | 12.03 | 3          | 340   | 15   | 24.00 | 9           | 448   |
| 4    | 14.88 | 4          | 80    | 16   | 29.00 | 10          | 776   |
| 5    | 13.75 | 6          | 150   | 17   | 13.35 | 6           | 200   |
| 6    | 18.11 | 7          | 330   | 18   | 19.00 | 7           | 132   |
| 7    | 8.00  | 2          | 110   | 19   | 9.50  | 3           | 36    |
| 8    | 17.83 | 7          | 210   | 20   | 35.10 | 17          | 770   |
| 9    | 79.24 | 30         | 1460  | 21   | 17.90 | 10          | 140   |
| 10   | 21.50 | 5          | 605   | 22   | 52.32 | 26          | 810   |
| 11   | 40.33 | 16         | 688   | 23   | 18.75 | 9           | 450   |
| 12   | 21.00 | 10         | 215   | 24   | 19.83 | 8           | 685   |
|      |       |            |       | 25   | 10.75 | 4           | 150   |

On désire ajuster à cet ensemble de données un modèle de régression multiple

$$y(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

où  $y$  est le temps requis,  $x_1$  est le nombre de caisses utilisées et  $x_2$  est la distance parcourue.

**4.1.** Donner la formulation mathématique de ce problème.

**4.2.** Écrire le problème sous la forme :

$$\begin{cases} \text{Min} & f(\beta) = \frac{1}{2} \|r(\beta)\|^2 = \frac{1}{2} \|y - X\beta\|^2 \\ & \beta \in \mathbb{R}^p \end{cases}$$

On donnera les valeurs de  $X$  et de  $y$  et à quoi correspond  $\beta$ .

▷ **Exercice 5.** La maquette d'un nouveau type d'éolienne est testé en soufflerie. 20 mesures sont réalisées entre 1 et 20  $m/s$ . L'allure de la réponse suggère un modèle à rupture (cf. la figure 1)

La production  $y$  est modélisée en fonction du vent généré dans la soufflerie  $x$  de la façon suivante : entre 1 et 5  $m/s$ , la réponse est supposée constante, elle augmente linéairement entre 5 et 15  $m/s$ , avant de saturer (redevenir constante) au delà de 15  $m/s$ . Il y a bien sur continuité de la réponse aux points 5 et 15  $m/s$

<sup>2</sup>. Exemple provenant du livre d'A. Antoniadis, J. Berruyer et R. Carmona, régression non linéaire et applications", éditions Economica, p. 45

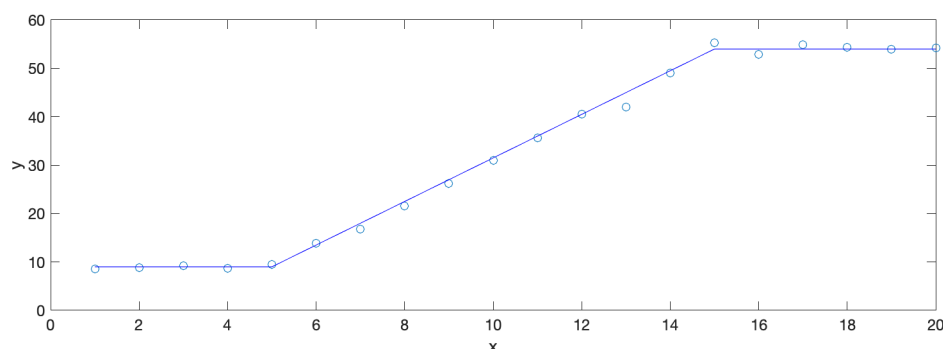


FIGURE 1 – Données et modèle pour une éolienne.

**5.1.** Écrire le modèle  $y(x, \beta)$  en fonction des plages des valeurs de  $x$ . Quelle est la dimension de  $\beta$ .

**5.2.** Écrire le problème aux moindres carrés d'estimation des paramètres  $\beta$ . Ce problème est-il linéaire? Si oui on donnera le vecteur  $\mathbf{y}$  et la matrice  $\mathbf{X}$  permettant d'écrire le problème sous la forme

$$(P) \begin{cases} \min \frac{1}{2} \|\mathbf{y} - \mathbf{X}\beta\|^2 \\ \beta \in \mathbb{R}^p. \end{cases}$$

▷ **Exercice 6. Géoréférence d'une image satellite**<sup>3</sup>

On dispose d'une image satellite que l'on désire recaler par rapport à une carte géographique que l'on a à notre disposition. Pour cela on définit  $n$  points, appelés points d'amer, que l'on peut parfaitement faire correspondre sur la carte et sur l'image satellite. On prend par exemple un croisement de route, un point particulier sur une rivière... Concrètement on a donc à notre disposition  $n$  coordonnées  $(x_i, y_i)$  des  $n$  points d'amer sur la carte et  $n$  coordonnées  $(x'_i, y'_i)$  de ces mêmes points sur l'image satellite. On choisit d'exprimer ces coordonnées :

- en pixels pour les  $(x'_i, y'_i)$  (coordonnées (0,0) pour le coin inférieur gauche) ;
- en mètres relativement à un référentiel géodésique particulier pour les  $(x_i, y_i)$ , via une carte IGN par exemple.

On a par exemple les données suivantes :

| Numéros  | $x_i$    | $y_i$    | $x'_i$   | $y'_i$   |
|----------|----------|----------|----------|----------|
| 1        | 252      | 2661     | 458805   | 1831634  |
| 2        | 235      | 2603     | 458157   | 1830577  |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 23       | 1021     | 2254     | 471301   | 1819574  |

En pratique l'image satellite est déformée par rapport à la réalité. Cette déformation a plusieurs origines : satellite non vertical par rapport à la prise de vue, présence de nuages dans

3. Voir cours de C. Monteil

l'atmosphère... En conséquence on suppose que l'on peut écrire :

$$\begin{cases} x = \gamma_0 + \gamma_1 x' + \gamma_2 y' + \gamma_3 x'^2 + \gamma_4 x' y' + \gamma_5 y'^2 \\ y = \delta_0 + \delta_1 x' + \delta_2 y' + \delta_3 x'^2 + \delta_4 x' y' + \delta_5 y'^2 \end{cases}$$

On désire estimer les paramètres par les moindres carrés.

**6.1.** Pour l'estimation des paramètres  $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_5)$  quelles sont les données ?

**6.2.** Écrire le problème d'estimation par les moindres carrés linéaires de  $\gamma$ .

**6.3.** Mêmes questions pour  $\delta$ .

▷ **Exercice 7.** On désire construire un réservoir de forme cylindrique de volume maximum dont la surface latérale est inférieure à  $S_{lat}$  et la surface totale est inférieure à  $S_{tot}$ .

**7.1.** Formaliser le problème.

▷ **Exercice 8.** Soit  $B(a, \delta)$  la boule de centre  $a$  et de rayon  $\delta > 0$  fixée. Soient  $p$  frères ennemis. On désire enfermer ces frères ennemis dans  $B(a, \delta)$  en maximisant la distance minimale  $\xi$  qui les sépare deux à deux.

**8.1.** Formaliser le problème.

▷ **Exercice 9.** On s'intéresse ici à la modélisation via un neurone formel.

**Définition 1.** Un neurone formel est une fonction paramétrée par  $n+1$  paramètres  $w_1, \dots, w_n, \theta$  :

$$\begin{aligned} g: \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} &\longrightarrow \mathbb{R} \\ (x, w, b) &\longmapsto g(x, w, b) := \sigma(\sum_{i=1}^n w_i x_i + b) \end{aligned}$$

où  $\sigma$  est une fonction donnée qui s'appelle une fonction d'activation. Chaque paramètre  $w_i$  s'appelle le poids synaptique associé au signal d'entrée  $x_i$ .

On prendra dans la suite, sauf mention contraire, comme fonction  $\sigma$  la fonction tangente hyperbolique :

$$\begin{aligned} \sigma: \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \sigma(x) := \frac{1 - e^{-x}}{1 + e^{-x}}. \end{aligned}$$

La figure 2 schématise un neurone formel.

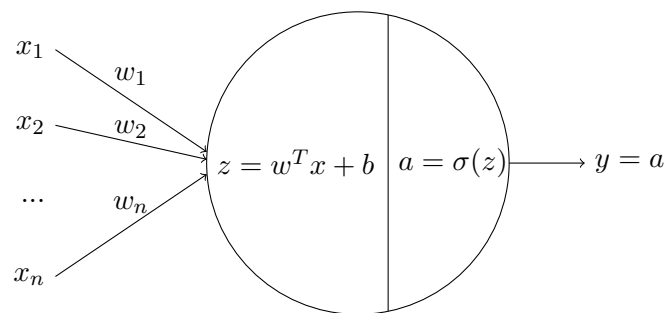
1. produit scalaire entre les entrées  $x$  et les poids synaptiques  $w : w^T x$  ;
2. ajout d'une valeur de référence (biais  $b$ ) :  $z = w^T x + b$
3. application de la fonction d'activation à la valeur obtenue  $z : a = \sigma(z)$

**Définition 2.** On a à notre disposition  $K$  points  $x^k \in \mathbb{R}^n$  et  $y^k \in \mathbb{R}$ , on appelle apprentissage du neurone l'estimation par les moindres carrés des paramètres du neurone.

**9.1.** Écrire le problème aux moindres carrés qui définit l'apprentissage. On donnera en particulier la fonction résidu  $r$  en précisant clairement l'espace de départ et l'espace d'arrivée.

**9.2.** Ce problème est-il un problème aux moindres carrés linéaires ? Si oui, on donnera la matrice  $X$ .

**9.3.** Si on prend comme fonction d'activation  $\sigma$  l'identité, le problème aux moindres carrés devient-il linéaire ? Si oui, on donnera la matrice  $X$ .

FIGURE 2 – *Représentation schématique d'un neurone.*