NOM : Prénom :

Traitement d'image avancé

Examen Module M6, parcours SI, Master EEAP

Durée 2h00,

Documents autorisés, appareils communiquant interdits.

I – Level-set (~20 min)

1) Quelles sont les propriétés intéressantes des méthodes de segmentation dites « level-set » ?

2) Afin de simplifier les illustrations, nous allons traiter un problème en 1D (une dimension). Le but de cet exercice est de segmenter (***sans l'utilisation de pré-traitements***) l'image 1D dont le profil d'intensité est donné ci-dessous (les points correspondent au résultat de segmentation souhaité) :

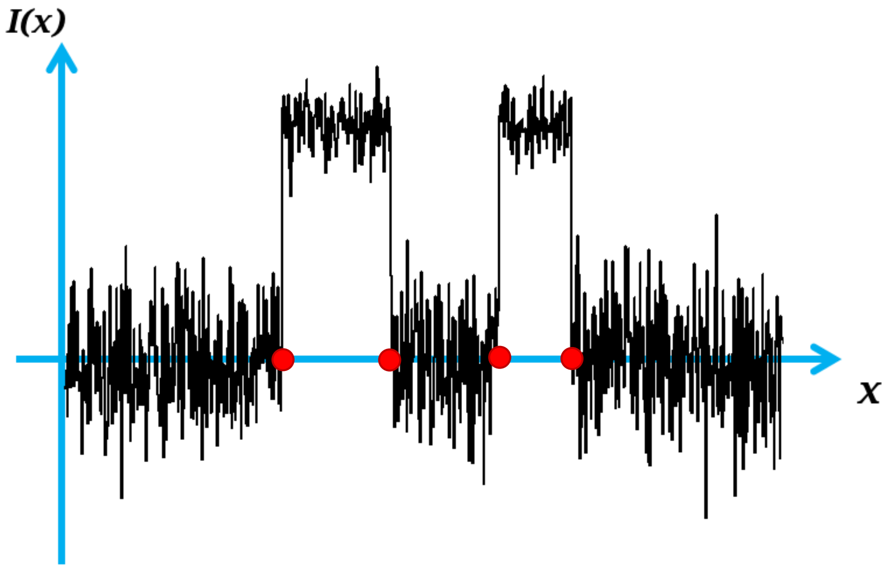


Figure 1 : Image 1D à segmenter

a. Donner l'expression du critère d'énergie que vous souhaitez utiliser pour segmenter une telle image. Vous devez argumenter ce choix en 3 lignes ***maximum***.

b. Afin de segmenter une telle image, nous avons décidé de faire évoluer qu'un seul level-set $\Phi(x)$ dont l'initialisation est donnée dans la figure ci-dessous.

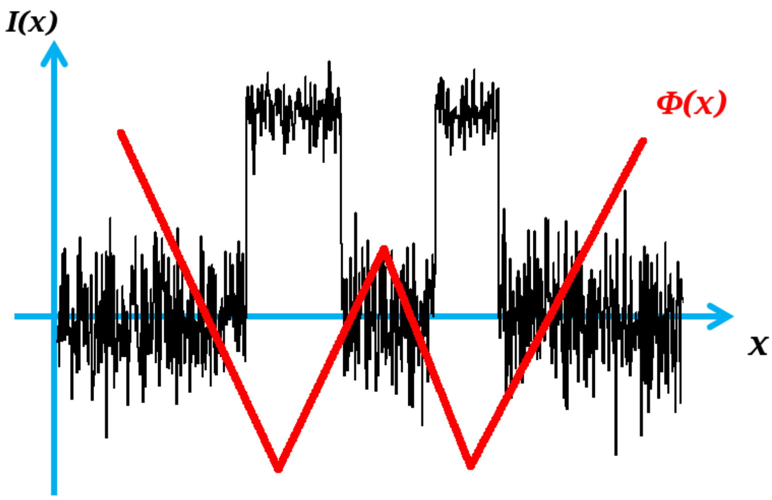


Figure 2 : Level-set initial

c. Expliquez en 2 lignes ***maximum*** comment un tel level-set permet de segmenter l'image 1D donnée dans la figure 1.

d. Quel type d'évolution de level-set choisissez-vous (sur toute l'image, en bande étroite) ? Vous devez justifier votre réponse de façon claire et précise.

II - Régression non-paramétrique (~90 min)

Un brin d'explications

Les méthodes de régression permettent d'analyser ou d'expliquer statistiquement le comportement d'une variable aléatoire (Y) par rapport à une ou plusieurs autres (X).

Dans la pratique, cela correspond d'abord à obtenir le couple de mesures ($\mathbf{x}_i$,$\mathbf{y}_i$) pour chacun des n individus d'un échantillon ($i = 1..n$) ; puis de déterminer la fonction f telle que $\mathbf{y}=f(\mathbf{x})$. Le but étant que la fonction f explique au mieux la valeur $\mathbf{y}$ à partir de $\mathbf{x}$ en se basant sur les observations ($\mathbf{x}_i$,$\mathbf{y}_i$).

Mais quelle fonction f ? On peut fixer un modèle de fonction ($a.cos(b.\mathbf{x}+c)$, un polynôme de degré 3, …) dont il faudra uniquement déterminer les paramètres (les coefficients du polynôme par exemple). Il s'agit d'un a priori sur le comportement des données (donc il n'y a rien à expliquer puisqu'on connait déjà le comportement ;)). Mais quand on n'a pas d'a priori, ou qu'on ne veut pas en faire, il faut avoir recourt aux techniques non-paramétriques, comme l'est la méthode de Parzen pour l'estimation de densité de probabilité. Il s'agit alors de la régression non-paramétrique.

La figure suivante illustre une régression non-paramétrique d'une fonction de R dans R.

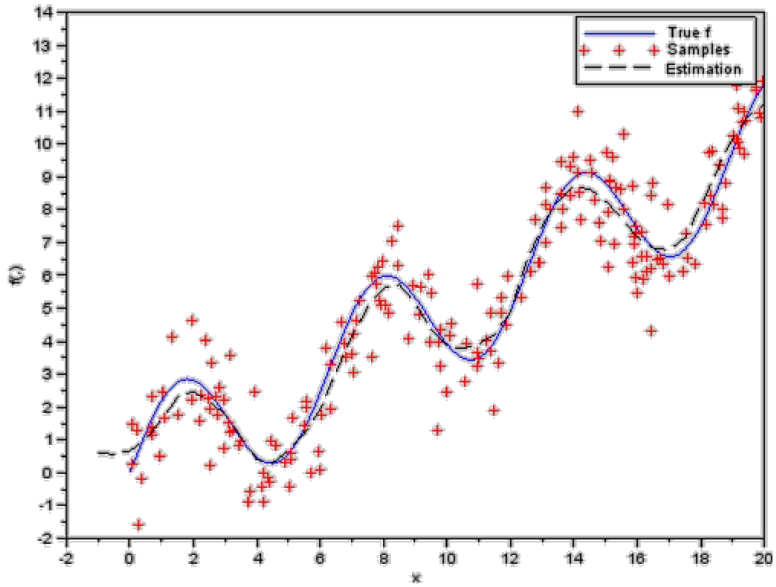
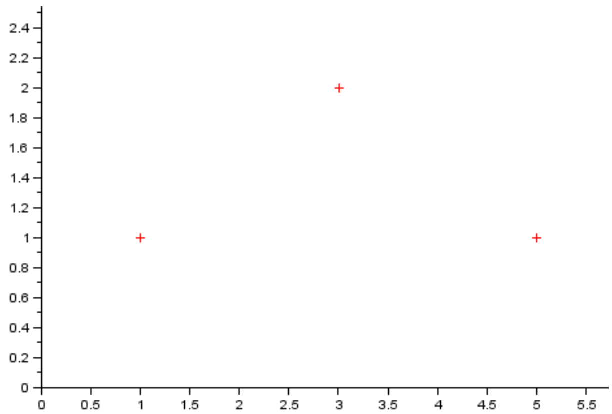


Figure 1: Régression (ligne pointillée) effectuée à partir de 200 mesures bruitées (échantillons, représentés par les croix). La vraie fonction est représentée par la ligne continue.

Un peu de travail

Lire toutes les questions, répondre dans l'ordre qui vous convient

1) Mise en pratique simple
    a. Illustrer le fonctionnement et donner les valeurs obtenues par la régression sur les échantillons de la figure suivante, pour les valeurs $f(2)$ et $f(3)$ en utilisant un noyau rectangle et respectivement une échelle $h=1$ et $h=2$



    b. De quel type de filtre se rapproche la régression ?

2) Formulation de la régression de *f*

Modifier l'estimation de Parzen pour réaliser une régression non paramétrique d'une fonction f en un point x à partir des n mesures (x_i,y_i) scalaires.

Remarques : il faut bien dissocier le rôle des 2 caractéristiques (x et y). Ce processus n'est pas itératif. Penser à normaliser par la somme des pondérations…

3) Relation entre n et h
   a. Quelle doit être la forme de la relation entre n, le nombre d'échantillons, et h l'échelle ?
   b. Faire une proposition.

4) Mesure d'erreur et résidus
   a. Proposer une mesure de la fiabilité de l'estimation au point x.
   b. Exprimer une distance entre les mesures y_i et l'estimation $f(x_i)$.

5) Outliers (points aberrants)
   a. A partir des mesures de la question précédente, proposer une méthode permettant de détecter les points aberrants.
   b. Toujours en vous inspirant de la question 4, proposer une amélioration de votre méthode de régression vis-à-vis des points aberrants. Préciser mathématiquement le sens de 'améliorer' (par rapport à quel critère ou quelle mesure) ?

Remarque : question ouverte. Plusieurs solutions possibles. Vous devez montrer la cohérence de votre proposition par rapport au critère amélioré.

Tout ça pour ça : Classification supervisée

On propose d'utiliser la régression pour séparer deux ensembles de points A et B. Ce problème est un problème classique de classification supervisée où le but est de déterminer la fonction séparant au mieux (suivant un critère donné) 2 ou plus ensembles de points.

6) Sur la figure suivante,
   a. Définir votre critère de 'bonne' séparation des 2 ensembles A et B
   b. Tracer sur la figure la fonction séparant les 2 ensembles en respectant votre critère.

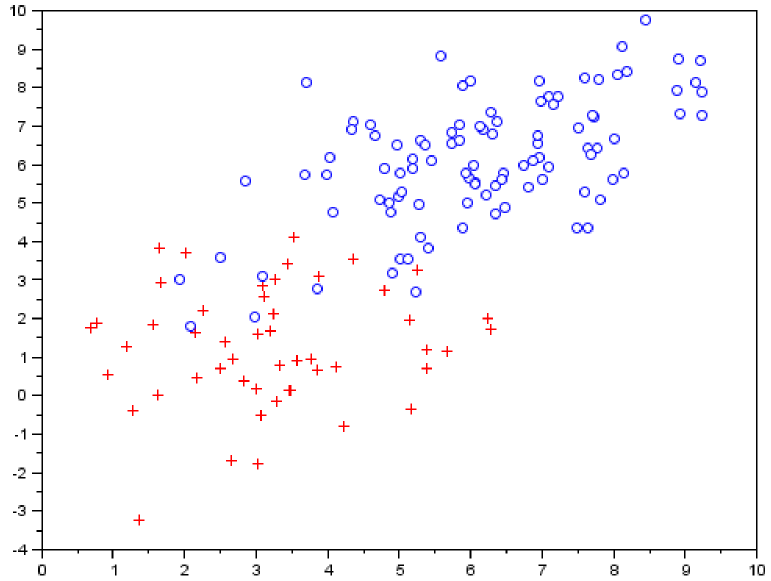


Figure 2:les 2 ensembles (A : les '+' et B les 'o') à séparer

7) **Proposer une méthode pour déterminer la fonction *f* séparant au mieux, selon votre critère, les 2 ensembles.**

Pour simplifier, nous ferons les hypothèses suivantes

- La fonction *f* sera toujours bijective
- Il n'y a toujours que 2 ensembles de points

Remarque : Sans écrire de code, il faudra détailler et illustrer votre proposition de manière suffisante pour pouvoir la tester (et à minima que je la comprenne !)

III – Exercice : EQM (20 min)

On synthétise une image bruitée G par l'équation suivante :

$$G(i) = I(i) + \eta(i)$$

Avec η un bruit gaussien de moyenne nulle et d'écart type σ ; et *I* l'image de référence sans bruit **qui est connue**.

Dans la suite, on filtre l'image *G* par convolution avec un filtre *F* et on s'intéresse à divers résultats de l'Erreur Quadratique Moyenne. L'expression de l'EQM entre deux images G_1 et G_2 de *N* pixels est donnée par :

$$EQM(G_1, G_2) = \frac{1}{N}\sum_{i=1}^{N}\left\|G_1(i) - G_2(i)\right\|^2$$

1- Dans le cadre du débruitage d'image (on cherche à retrouver *I* à partir de *G*), quelle est la valeur de l'EQM correspondant au cas le plus favorable ?

2- Hors conditions de la question précédente, que vaut l'EQM entre *I* et *G* (non filtrée) ?

3- Quelles hypothèses sur *I* et η sont nécessaires pour définir un filtre *F* permettant, lors de la convolution avec *G*, d'obtenir *I* ? Est-ce toujours possible ? De quel type de filtre est F (passe bas, passe bande, …) ?

4- Montrer que si le filtre *F* est tel que la convolution $F * \eta \cong 0$, alors l'erreur quadratique moyenne entre *I* et *G* filtrée par *F*, est au plus égale à la variance de l'image *I*.

Fin.