Introduction à la statistique non paramétrique

Catherine MATIAS

CNRS, Laboratoire Statistique & Génome, Évry

http://stat.genopole.cnrs.fr/∼cmatias

Atelier SFDS
27/28 septembre 2012

Partie 4 : Estimation de densité - Compléments

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes . . .

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Estimation de densités multivariées I

Théorie vs pratique

- Tous les estimateurs présentés dans la partie 3 se généralisent à la dimension supérieure,
- En pratique, l'estimation devient plus difficile quand la dimension augmente : c'est le fléau de la dimension.

Estimation de densités multivariées II

Fléau de la dimension (curse of dimensionality)

- Exemple 1 :
  - Dans $\mathbb{R}$
    - Pts uniformément répartis dans $[-1, +1]$
    - 100% de points situées à distance ≤ 1 de l'origine
  - Dans $\mathbb{R}^{10}$
    - Pts uniformément répartis dans $[-1, +1]^{10}$
    - % de points situées à 1 distance ≤ 0.75 de l'origine : 5%
- Exemple 2 : on veut construire un histogramme en s'appuyant sur au moins une moyenne de 10 points par intervalle et 10 classes par variable
  - $\mathbb{R}$: 10 classes $n = 100$
  - $\mathbb{R}^2$: 100 classes $n = 1000$
  - $\mathbb{R}^{10}$: 10^{10} classes $n = 10^{11} = 100 milliards$

Estimation de densités multivariées III

Fléau de la dimension (suite)

- Si p assez grand, l'espace $\mathbb{R}^p$ est pratiquement vide : difficulté dans l'emploi des méthodes avec fenêtres,
- Les points voisins d'un point donné sont tous très loin : difficultés dans l'emploi de méthodes du type k-plus proches voisins.

Représentations graphiques

- Problèmes pour représenter graphiquement les données,
- En dimension 2, on peut encore représenter des densités, au-delà, ça devient compliqué ...

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Estimateurs à noyaux multivariés I

Définition dans $\mathbb{R}^2$

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une densité et $(X_1, Y_1), \ldots, (X_n, Y_n)$ un échantillon de densité f. On utilise un noyau produit et on construit

$$\hat{f}_n(x, y) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right) K\left(\frac{Y_i - y}{h}\right).$$

Généralisation dans $\mathbb{R}^p$

$$\hat{f}_n(x^1, \ldots, x^p) = \frac{1}{nh^p} \sum_{i=1}^{n} \prod_{j=1}^{p} K\left(\frac{X_i^j - x^j}{h}\right).$$

Estimateurs à noyaux multivariés II

Propriétés

- Contrôles similaires du biais et de la variance, pour des classes de régularité généralisées à la dimension $p > 1$.
- Exemple avec $f \in \Sigma_{d,p}(1, L) =$ ensemble des densités sur $\mathbb{R}^p$ qui sont Lipschitziennes
  - le biais ne dépend pas de la dimension de l'espace : $\text{Biais} = O(h)$,
  - Par contre, la variance dépend de p : $\mathbb{V}\text{ar}_f(\hat{f}_n(x)) = O(1/(nh^p))$
  - la fenêtre optimale pour le risque quadratique est $h = cn^{-1/(2+p)}$
  - et la vitesse de convergence du risque quadratique correspondante est $n^{-2/(2+p)}$.
  - Quand la dimension p augmente, cette vitesse est plus lente.

Illustration I

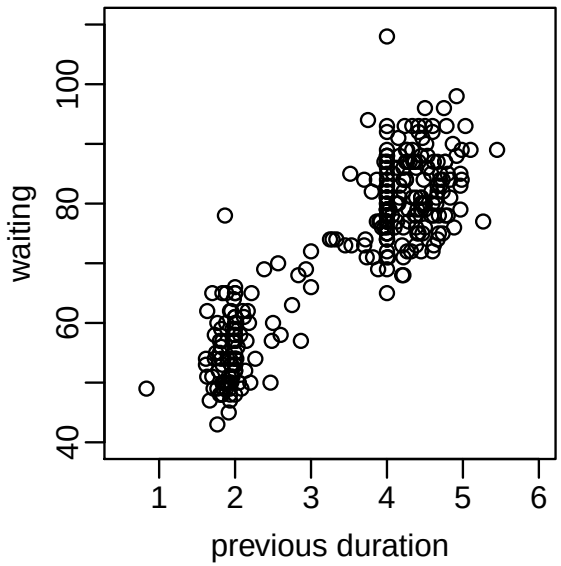
On reprend les données Geyser introduites dans la partie 3 : temps d'attente et durée de chaque eruption. La fonction *kde2d* de la librairie *MASS* implémente un estimateur à noyau en dimension 2 (uniquement).

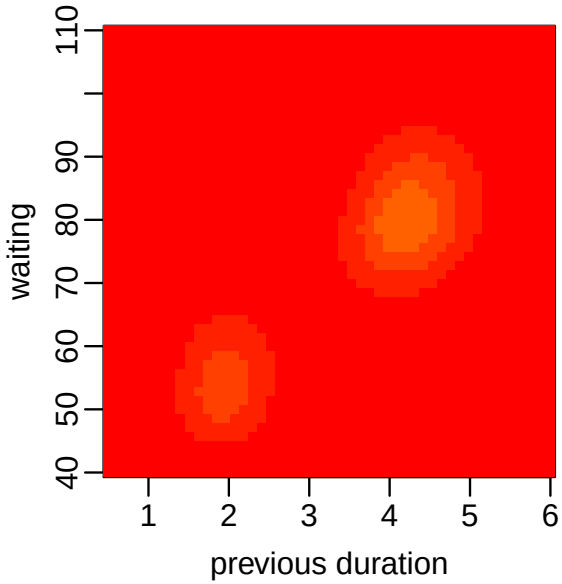
```
> library(MASS)
> geyser2<-data.frame(as.data.frame(geyser)[-1, ],
+                          pduration=geyser$duration[-299])
> attach(geyser2)
> par(mfrow=c(2,2),mgp=c(1.8,0.5,0),mar=c(3,3,3,3))
> plot(pduration,waiting,xlim=c(0.5,6),ylim=c(40,110),
+      xlab="previous duration",ylab = "waiting")
> f1<-kde2d(pduration,waiting,n=50,lims=c(0.5,6,40,110))
> # on a linear scale
> image(f1,zlim=c(0,0.075),xlab="previous duration",
+       ylab ="waiting")
> # to get on a log scale do:
```

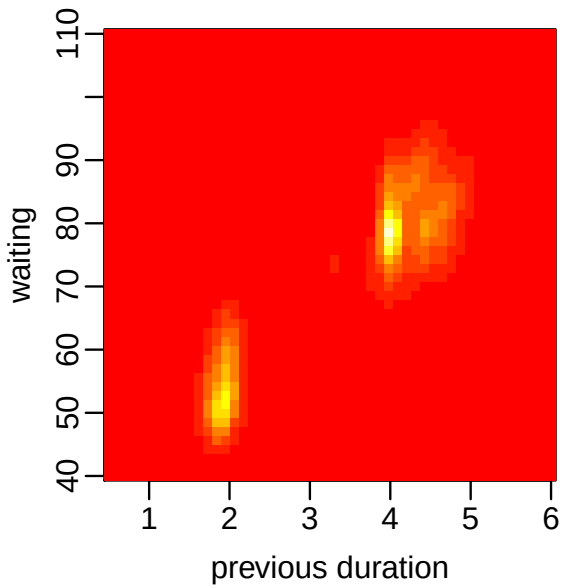
Illustration II

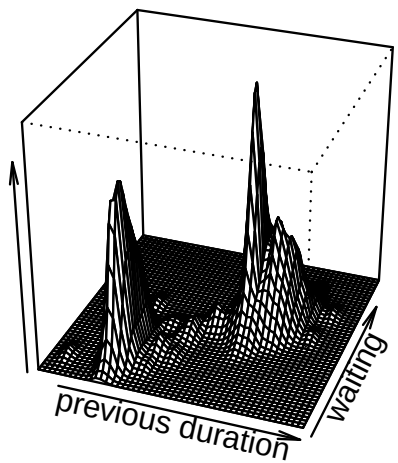
```
> # image(f1$x,f1$y,log(f1$z),zlim =c(-40,0),
> #         xlab ="previous duration",ylab ="waiting")
> f2<-kde2d(pduration,waiting,n=50,lims=c(0.5,6,40,110),
+  h=c(width.SJ(duration),width.SJ(waiting)))
> image(f2,zlim=c(0,0.075),xlab="previous duration",
+       ylab ="waiting")
> persp(f2,phi=30,theta=20,d=5,xlab="previous duration",
+       ylab="waiting",zlab ="")
> detach(geyser2)
```

Illustration III









Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Projection pursuit density estimator I

Principe [Friedman *et al.* 84]

Sélectionner de façon itérative, des directions privilégiées, dans lesquelles on affine l'estimation de la densité.

Construction

- On part d'un estimateur p-dimensionnel initial f_0 de la densité. (Par ex, densité de $\mathcal{N}_p(\bar{x}_n, S_n)$ où $\bar{x}_n, S_n$ moyenne et covariance empiriques),
- À l'étape m,
  - On sélectionne une direction $\theta_m \in \mathbb{R}^p$
  - On choisit une fonction unidimensionnelle ϕ_m, dite fonction d'augmentation, qui va affiner l'estim dans la direction θ_m.

À l'issue de la procédure, on obtient un estimateur de la forme

$$\hat{f}_M^{PPDE}(x) = f_0(x) \prod_{m=1}^{M} \phi_m(\theta_m^\intercal x).$$

Projection pursuit density estimator II

Choix de la direction et de la fonction d'augmentation

θ_m, ϕ_m sont choisis tels qu'ils minimisent la divergence de Kullback-Leibler entre la densité cible f et l'estimateur $\hat{f}_m^{PPDE}$:

$$(\theta_m, \phi_m) = \underset{\theta,\phi}{\text{Argmin}}\, K(\hat{f}_m^{PPDE}, f).$$

En pratique

$$\begin{aligned} K(f, \hat{f}_m^{PPDE}) &= \int_{\mathbb{R}^p} \Big[\log \frac{f(x)}{\hat{f}_m^{PPDE}(x)}\Big] f(x)\, dx \\ &= \int_{\mathbb{R}^p} \log \frac{f(x)}{\phi_m(\theta_m^\intercal x)\hat{f}_{m-1}^{PPDE}(x)} f(x)\, dx, \end{aligned}$$

et en particulier

$$(\theta_m, \phi_m) = \underset{\theta,\phi}{\text{Argmax}} \int_{\mathbb{R}^p} [\log \phi(\theta^\intercal x)] f(x)\, dx. \qquad (1)$$

Projection pursuit density estimator III

- Il faut optimiser la quantité précédente, sous la contrainte $\int \hat{f}_m^{PPDE}(x) = 1$ (pour avoir un estimateur qui soit une densité),
- Pour $\theta_m = \theta$ fixé, lorsque f est connue, la solution en ϕ du problème est donnée par

$$\phi(\theta^\intercal x) = \frac{f^\theta(\theta^\intercal x)}{\hat{f}_{m-1}^{PPDE,\theta}(\theta^\intercal x)},$$

où f^θ et $\hat{f}_{m-1}^{PPDE,\theta}$ sont resp. les marginales uni-dimensionnelles de f et $\hat{f}_{m-1}^{PPDE}$ dans la direction θ.

Projection pursuit density estimator IV

Description de la procédure

À chaque étape m,

- On sélectionne (optimisation numérique) la direction θ qui maximise l'estimateur de (1) pour la valeur courante de ϕ

$$\theta_m = \underset{\theta, \|\theta\|=1}{\text{Argmax}} \frac{1}{n} \sum_{i=1}^{n} \log \phi_{m-1}(\theta^\intercal X_i),$$

  (lorsque $m = 1$, cette quantité vaut $n^{-1} \sum_{i=1}^{n} \log f_0(\theta^\intercal X_i)$). C'est la log-vraisemblance du modèle dans la direction θ.
- On projette les observations X_i dans la direction θ_m, d'où les nouvelles obs $Z_i = \theta_m^\intercal X_i$,
- On estime la densité unidimensionnelle des Z_i, par $\hat{f}^{\theta_m}$,

Projection pursuit density estimator V

- On approche la marginale $\hat{f}_{m-1}^{PPDE,\theta_m}$ dans la direction θ_m de l'estimateur à l'étape précédente $\hat{f}_{m-1}^{PPDE}$ par une approximation de Monte Carlo,
- On construit l'estimateur (unidim) du ratio

$$\hat{\phi}^m(z) = \hat{f}^{\theta_m}(z)/\hat{f}_{m-1}^{PPDE,\theta_m}(z),$$

où $z = \theta_m^\intercal x$.

On obtient ainsi un nouvel estimateur de la densité globale des observations X_i, via

$$\hat{f}_m^{PPDE}(x) = f_0(x) \prod_{k=1}^{m} \hat{\phi}^k(\theta_k^\intercal x).$$

PPDE : compléments

Avantages/Inconvénients

- Méthode qui tend à diminuer le fléau de la dimension, par une sélection des directions importantes,
- L'usage des approximations Monte carlo est numériquement lourd.

Remarques

- En pratique, on commence par transformer les données pour les rendre sphériques (centrées et de matrice de covariance empirique Identité) [Tukey & Tukey 81].

PPDE : illustration [Friedman *et al.* 84] I

1er exemple

- $n = 225$ observations, issues d'un mélange de 3 gaussiennes en dimension 2, centrées en 3 points équidistants, de covariance Identité.

PPDE : illustration [Friedman *et al.* 84] II

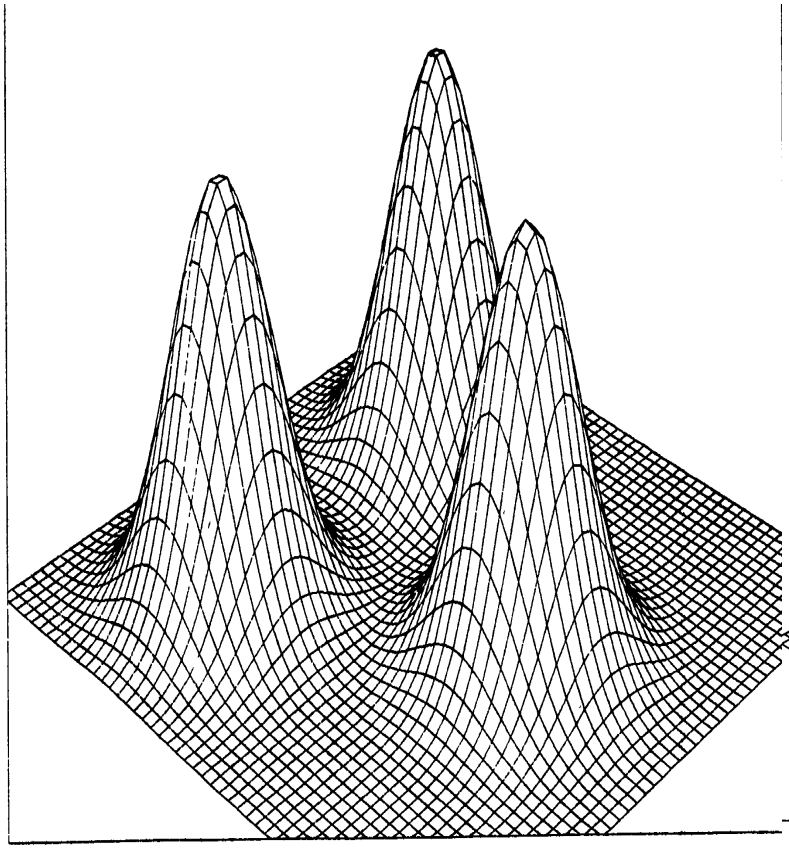


Figure 1. True Density Function—Gaussian Mixture.

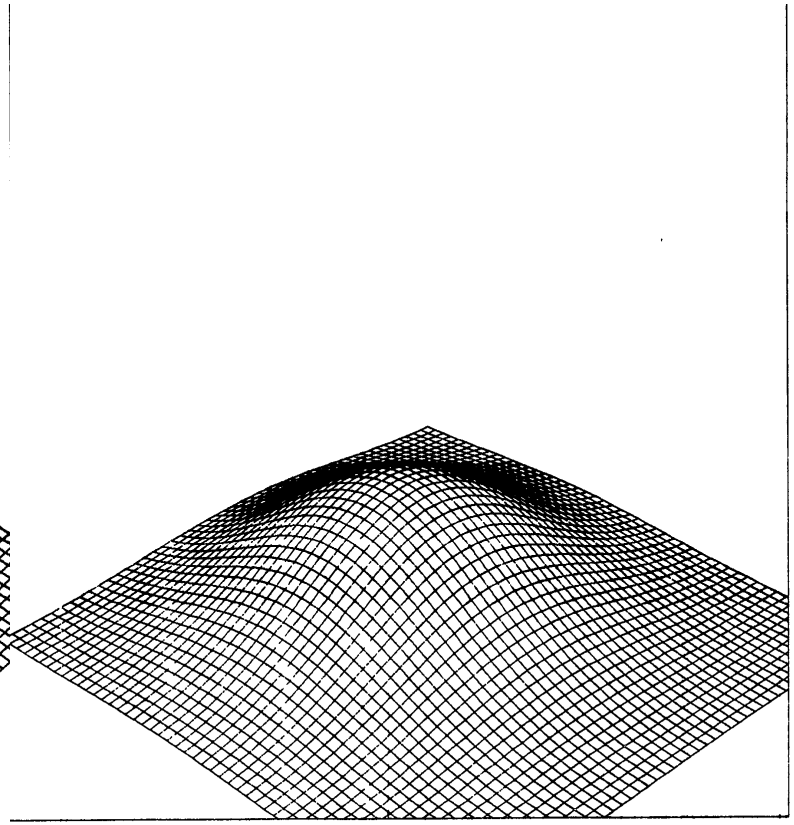


Figure 2. Initial Model $p_0(x)$—Gaussian With Sample Mean and Sample Covariance.

PPDE : illustration [Friedman *et al.* 84] III

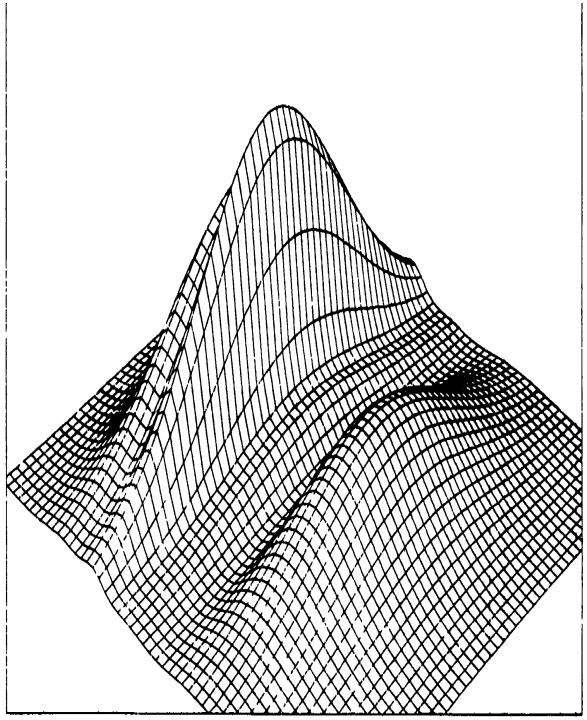


Figure 3. PPDE Estimate—First Iteration. (a) Data (···) and model (X) marginals along θ_1 = (0, 1); (b) first augmenting function $f_1(\theta_1 \cdot x)$; (c) model after the first iteration.

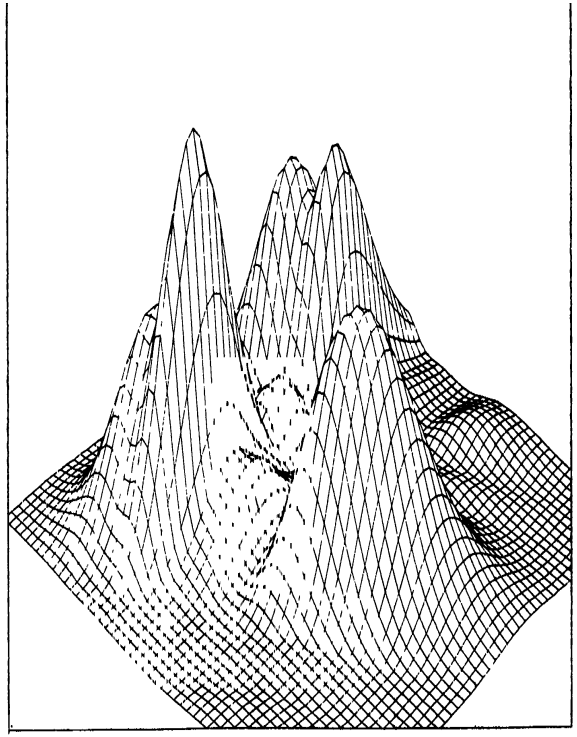


Figure 4. PPDE Estimate—Second Iteration. (a) Data and model marginals along θ_2 = (.87, .49); (b) second augmenting function $f_2(\theta_2 \cdot x)$; (c) model after the second iteration.

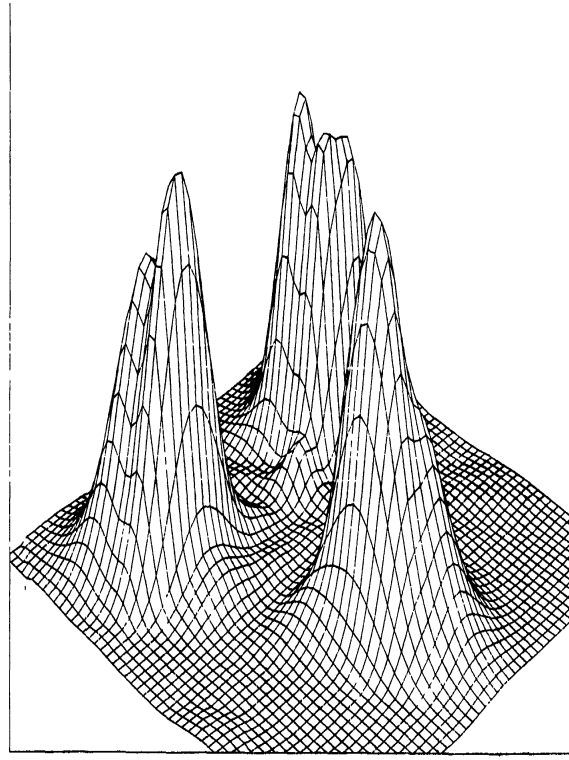


Figure 5. PPDE Estimate—Third Iteration. (a) Data and mod marginals along θ_3 = (.89, −.45); (b) third augmenting function f_3(· x); (c) model after the third iteration.

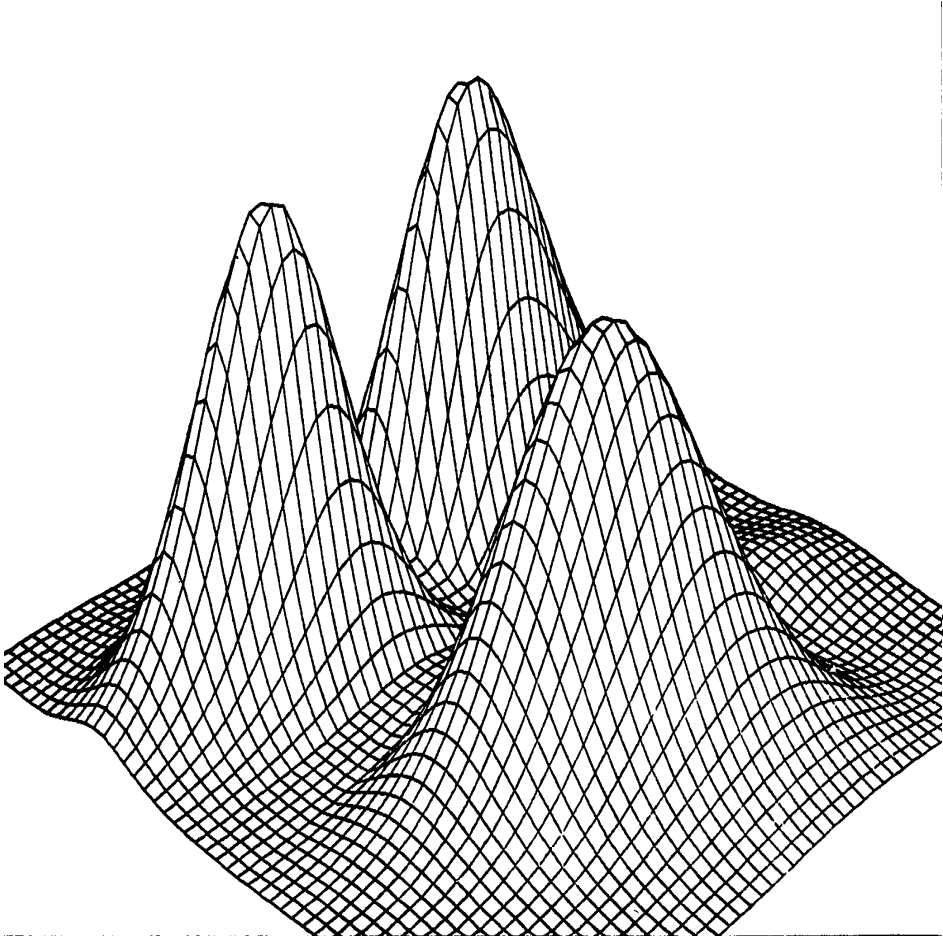
PPDE : illustration [Friedman *et al.* 84] IV

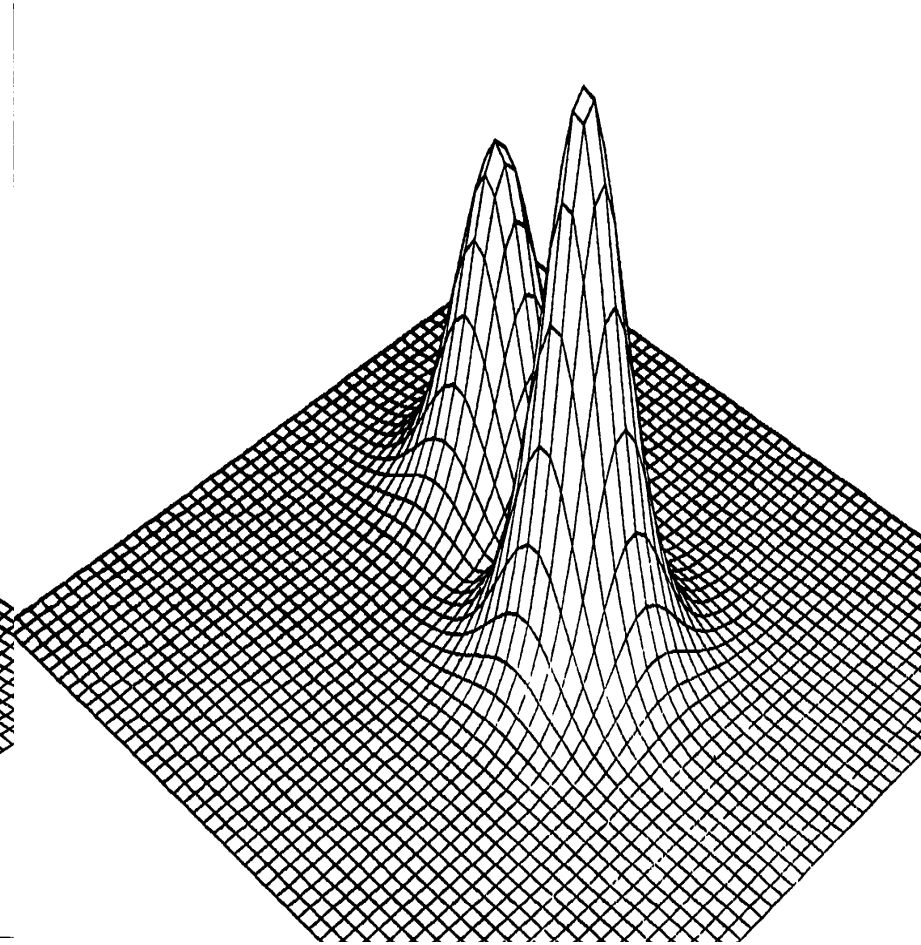
Second exemple : comparaison avec les knn (k plus proches voisins)

- $n = 225$ observations en dimension $p = 10$,
- Dans les 2 premières dimensions, même modèle que pour l'ex 1 : mélange de 3 gaussiennes,
- Les 8 variables restantes sont indépendantes, gaussiennes, centrées, d'écart-type à peu près identique aux deux premières (données sphériques).
- La structure des données se situe donc dans les 2 premières variables, les 8 autres représentent du "bruit".

PPDE : illustration [Friedman *et al.* 84] V

On compare l'estimateur PPDE (4 itérations, à gauche) à l'estimateur KNN (noyau rectangulaire, à droite).





PPDE : illustration [Friedman *et al.* 84] VI

Table 1. Percentage of Variance Explained by PPDE and KNNE Estimators

	PPDE		KNNE	
Dimensions, p	Monte Carlo Estimate (%)	Standard Deviation	Monte Carlo Estimate (%)	Standard Deviation
2	79	3	80	1
5	69	7	42	2
10	63	5	9	2

PPDE : illustration sur données réelles [Friedman *et al.* 84] I

Données

- Il s'agit d'une étude sur le diabète [Reaven & Miller 79]. On observe $n = 145$ sujets et on mesure $p = 5$ variables : poids relatif, tolérance au glucose 1 et 2, sécretion d'insuline, interaction insuline-glucose.
- Les 2 variables de tolérance au glucose étant fortement corrélées ($r = .96$), on ne conserve que la seconde.
- On se demande si la loi des 4 variables restantes s'approche correctement par le produit $p_{ab}(x_a, x_b)p_{cd}(x_c, x_d)$ pour un certain choix de a, b, c, d.
- Intérêt : visualiser les données uniquement à travers les deux nuages induits de points (variables a, b d'une part et c, d d'autre part).

PPDE : illustration sur données réelles [Friedman *et al.* 84] II

Procédure

Pour chaque choix de a, b, c et d,

- Idée : On va construire un estimateur de la densité des observations par PPDE, en partant de f_0 sous forme factorisée $f_0(x^a, x^b, x^c, x^d) = f_{0,ab}(x^a, x^b) f_{0,cd}(x^c, x^d)$. Si la log vraisemblance de l'estimateur PPDE augmente peu au cours des premières itérations, on conclu que les données sont bien décrites par l'estimateur initial factorisé.
- Pour cela, on n'a pas besoin d'un estimateur explicite. Il suffit d'un échantillon de cette loi.
- On va donc générer un échantillon de la loi $f_{ab}(x^a, x^b) f_{cd}(x^c, x^d)$ en permutant au hasard les labels des variables (a, b) et (c, d).

PPDE : illustration sur données réelles [Friedman *et al.* 84] III

Par ex $(a, b, c, d) = (1, 2, 3, 4)$: Ansi, pour $r_1, \ldots, r_n$ une permutation aléatoire des individus $1, \ldots, n$, on crée un échantillon $(x_1^1, x_1^2, x_{r_1}^3, x_{r_1}^4), \ldots, (x_i^1, x_i^2, x_{r_i}^3, x_{r_i}^4), \ldots, (x_n^1, x_n^2, x_{r_n}^3, x_{r_n}^4)$.

Table 2. Increase in (Log) Likelihood of PPDE Solutions From Factored Initial Model $p_o(\mathbf{x}) = p_{ab}(x_a, x_b) p_{cd}(x_c, x_d)$ (third example)

Combination				Number of Iterations	
a	*b*	*c*	*d*	*2*	*4*
1	2	3	4	85.7	124.1
1	3	2	4	47.1	76.3
1	4	2	3	85.8	122.1

PPDE : illustration sur données réelles [Friedman *et al.* 84] IV

Conclusions

- La combinaison $(1,3);(2,4)$ semble la mieux adaptée (accroissement de l'ajustement le plus faible),
- Cependant, le fait que la vraisemblance augmente en partant de cette loi factorisée initiale indique que toute la structure des données n'est pas simplement contenue dans le produit $(1,3)$ versus $(2,4)$.

Plan partie 4

Cas des densités multivariées
- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...
- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Densités monotones

On suppose f densité monotone.

Exemples

- durées de vie $f : [0, +\infty) \to [0, +\infty)$ décroissante,
- distribution des p-values sous l'hypothèse alternative $f : [0, 1] \to [0, +\infty)$ décroissante,

Estimateur de Grenander pour densités monotones I

Log-vraisemblance

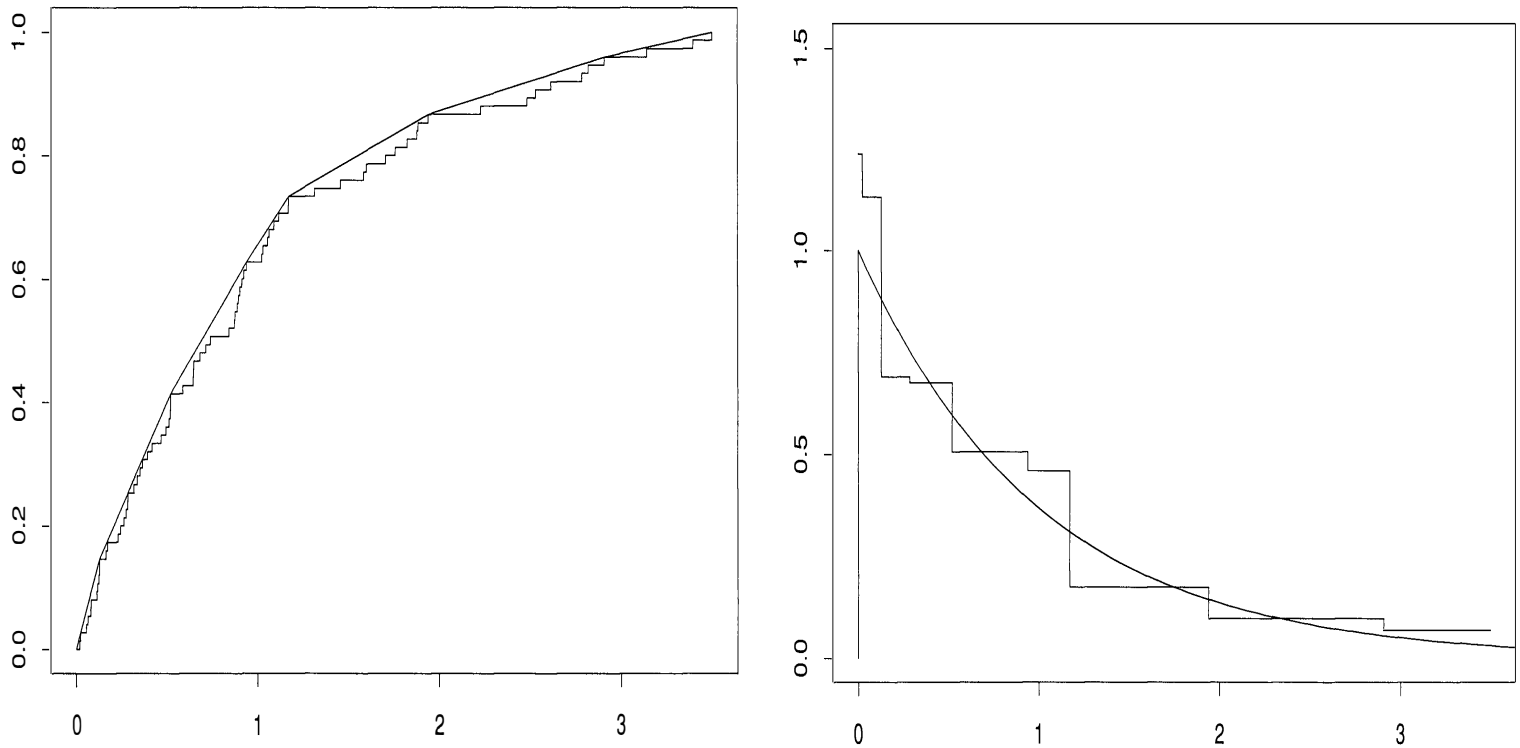
$$\ell_n(f) = \frac{1}{n}\sum_{i=1}^{n} \log f(X_i),$$

log-vraisemblance de la densité f.

- Si f n'est pas contraint (densité quelconque, même supposée régulière) alors $\sup_f \ell_n(f) = +\infty$ et l'e.m.v. n'est pas défini.
- Si f est monotone, alors l'e.m.v. existe et est unique.
- Lorsque f est décroissante, c'est la dérivée à gauche du plus petit majorant concave de la fdr empirique $\mathbb{F}_n$ (Rem : nécessairement décroissante).
- Lorsque f est croissante, c'est la dérivée à droite du plus grand minorant convexe de la fdr empirique $\mathbb{F}_n$ (Rem : nécessairement croissante).

Estimateur de Grenander pour densités monotones II

À gauche : plus petit majorant concave de la fdr empirique d'un échantillon de $n = 75$ variables de loi exponentielle. À droite : dérivée à gauche et vraie densité.



Estimateur de Grenander pour densités monotones III

Autre expression

- On considère les variables ordonnées $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$.
- On note $\mathbb{F}_n$ la fdr empirique et $\hat{F}_n$ son plus petit majorant concave : $\hat{F}_n$ est linéaire par morceaux.
- Sur chaque intervalle $(X_{(i-1)}, X_{(i)}]$, l'estimateur $\hat{f}_n$ vaut la pente de $\hat{F}_n$.

Implémentation

- Sous R, la fonction *grenander* de la bibliothèque *fdrtool* implémente l'estimateur de Grenander.

Propriétés

- L'estimateur de Grenander est également l'estimateur des moindres carrés de f.
- Estimateur consistant : $\forall x, \hat{f}_n(x) \to f(x)$ presque sûrement.
- Vitesse de convergence du risque quadratique $n^{-1/3}$, sans hypothèse de dérivées (à comparer à $n^{-m/(2m+1)}$ lorsqu'on suppose f est m-fois dérivable).
- Cette vitesse est minimax : c'est la meilleure possible si on suppose juste la monotonie.

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Densités unimodales I

Définition

- Si il existe $M_f \in \mathbb{R}$ tel que f est croissante sur $(-\infty, M_f]$ et décroissante sur $[M_f, +\infty)$ alors f est dite unimodale.
- Le mode M_f n'est pas nécessairement unique.

Estimation à mode connu

- Si le mode M_f est connu a priori, alors l'estimation de f se fait sur chaque intervalle $(-\infty, M_f]$ et $[M_f, +\infty)$ via l'estimateur de Grenander.
- Si aucune observation ne prend la valeur du mode, alors on peut montrer que cet estimateur maximise la vraisemblance.

Densités unimodales II

Estimation à mode inconnu

Si M_f n'est pas a priori connu, l'e.m.v. n'existe pas.

- L'idée naïve qui consiste à estimer le mode par une autre méthode, et utiliser l'estimateur de Grenander avec le mode estimé, ne fonctionne pas (sans hyps supplémentaires sur f),
- On peut par contre pour chaque valeur de M fixée, considérer l'estimateur de Grenander $\hat{f}_n^M$ de fdr associée $\hat{F}_n^M$, puis sélectionner le meilleur, par exemple au sens suivant

$$\hat{f} = \underset{\hat{f}_n^M}{\operatorname{Argmin}} \|\hat{F}_n^M - \mathbb{F}_n\|_\infty,$$

où $\mathbb{F}_n$ fdr empirique de l'échantillon.

- On peut mq cet estimateur a les mêmes performances asymptotiques que l'estimateur de Grenander.

Plan partie 4

Cas des densités multivariées

- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes ...

- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Densités convexes

On suppose f densité convexe décroissante,

EMV et MC

- L'e.m.v. et l'estimateur des moindres carrés sont alors bien définis et uniques,
- Par contre, ils ne coïncident pas en général.
- Sous l'hyp supplémentaire f deux fois dérivable, l'e.m.v. converge à la vitesse ponctuelle $n^{-2/5}$

Voir [Groeneboom *et al.* 01] pour plus de détails. Il existe un cadre plus général des fonctions k-monotones [Balabdaoui & Wellner 08].

Densités log-concaves

Définition

Une densité f est dite log-concave si $-\log(f)$ est une fonction convexe sur le support de f (convention $-\log 0 = +\infty$).

Exemples (paramétriques)

Gaussienne, uniforme, Gamma, Beta, Laplace, logistique, ...

Propriétés

- Les fonctions log-concaves sont nécessairement unimodales, mais la réciproque est fausse.
- L'estimateur du max de vraisemblance existe et peut être obtenu par des algos de maximisation sous contrainte.

[Rufibach 06, Rufibach 07] pour plus de détails.

Plan partie 4

Cas des densités multivariées
- Fléau de la dimension
- Généralisations des estimateurs précédents
- Poursuite de projection pour l'estimation de densité

Cas des densités monotones ou unimodales ou convexes . . .
- Densités monotones
- Densités unimodales
- Autres contraintes

Observations bruitées

Déconvolution I

Problème

On observe $X_1, \ldots, X_n$ i.i.d. de loi $X_i = Y_i + \epsilon_i$ où Y_i i.i.d. de densité f inconnue et ϵ_i i.i.d. de densité f_ϵ connue et $\{Y_i\}, \{\epsilon_i\}$ indépendants.

- On veut estimer f à partir des observations bruitées X_i.
- La densité des observations g vérifie $g = f \star f_\epsilon = \int f(\cdot - t) f_\epsilon(t)\, dt$. C'est la convolée entre f et f_ϵ.

Transformation de Fourier

- Si on suppose $f, f_\epsilon \in \mathbb{L}_2(\mathbb{R})$, alors on peut écrire $g^\star(x) = \int e^{itx} g(t)\, dt = f^\star(x) f_\epsilon^\star(x)$.
- Cette relation s'inverse pour donner

$$f(x) = \frac{1}{2\pi} \int e^{-iux} \frac{g^\star(u)}{f_\epsilon^\star(u)}\, du.$$

Déconvolution II

Estimateur à noyau de f

- On définit le noyau k_n via sa transformée de Fourier $k_n^\star$ par

$$k_n^\star(u) = \frac{k(u)}{f_\epsilon^\star(u/h)},$$

où $h = h_n \to 0$ est la fenêtre,

- et l'estimateur à noyau de f

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^{n} k_n \left(\frac{X_i - x}{h} \right).$$

Déconvolution III

Vitesses de convergence

Les résultats dépendent de la régularité de la densité du bruit

- Le bruit est dit super régulier si $f_\epsilon^\star$ décroît exponentiellement vite,
- Le bruit est dit régulier si $f_\epsilon^\star$ décroît polynomialement vite,

et aussi de la régularité de la densité inconnue f.

- Plus la densité du bruit est régulière, plus le pbm est difficile (vitesses lentes, minimax).
- En particulier, pour des bruits super-réguliers, les vitesses de convergence sont logarithmiques !

Voir [Fan 91] pour plus de détails.

Références I


[Balabdaoui & Wellner 08] Balabdaoui and Wellner
Estimation of a k-monotone density : Limit distribution theory and the spline connection.
Ann. Statist. 35 : 2536–2564, 2008.


[Fan 91] J. Fan
On the optimal rate of convergence for nonparametric deconvolution problems.
Annals of Statistics, 19(3) : 1257–1272, 1991.


[Friedman *et al.* 84] J.H. Friedman, W. Stuetzle, A. Schroeder
Projection Pursuit Density Estimation.
JASA, 79(387), 599–608, 1984.

Références II

[Groeneboomn *et al.* 01] Groeneboom & Jongbloed and Wellner

Estimation of a convex function : Characterization and asymptotic theory.

Ann. Statist. 29 : 1653–1698, 2001.

[Reaven & Miller 79] G. M. Reaven and R. G. Miller

An attempt to define the nature of chemical diabetes using a multidimensional analysis

Diabetologia, 16(1), 17–24, 1979.

[Rufibach 06] Rufibach, K.

Log-concave density estimation and bump hunting for I.I.D. observations.

Ph.D. thesis, Universities of Bern and Göttingen, 2006.

Références III


[Rufibach 07] Rufibach, K.
Computing maximum likelihood estimators of a log-concave density function.
J. Statist. Comp. Sim. 77 : 561–574, 2007.


[Tukey & Tukey 81] P.A. Tukey and J.W. Tukey
"Preparation : prechosen sequences of views", in *Interpreting Multivariate Data*
ed. V. Barnett, London : John Wiley, pp.189-213, 1981.