

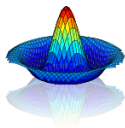
Exploration de modèles par méthodes de surface de réponse

École Chercheurs Mexico

Robert Faivre

INRA - Biométrie et Intelligence Artificielle, MIA Toulouse

Giens, le 10 Juin 2010



MEXICO
WEXICO

1 Introduction

- Objectifs et Enjeux
- Questions

2 Méthodes de régression paramétrique

- Ecriture du modèle et estimation des paramètres
- Critères d'ajustement et comparaison de modèles
- Cas multidimensionnel
- Retour sur l'Analyse de Sensibilité
- Plans d'expérience optimaux

3 Méthodes de régression non-paramétrique

- Exemples de modèles non paramétriques
- Critère d'ajustement
- Extension multidimensionnelle
- Plans d'expérience

4 Références

Plan

1 Introduction

- Objectifs et Enjeux
- Questions

2 Méthodes de régression paramétrique

- Ecriture du modèle et estimation des paramètres
- Critères d'ajustement et comparaison de modèles
- Cas multidimensionnel
- Retour sur l'Analyse de Sensibilité
- Plans d'expérience optimaux

3 Méthodes de régression non-paramétrique

- Exemples de modèles non paramétriques
- Critère d'ajustement
- Extension multidimensionnelle
- Plans d'expérience

4 Références

Objectifs et Enjeux

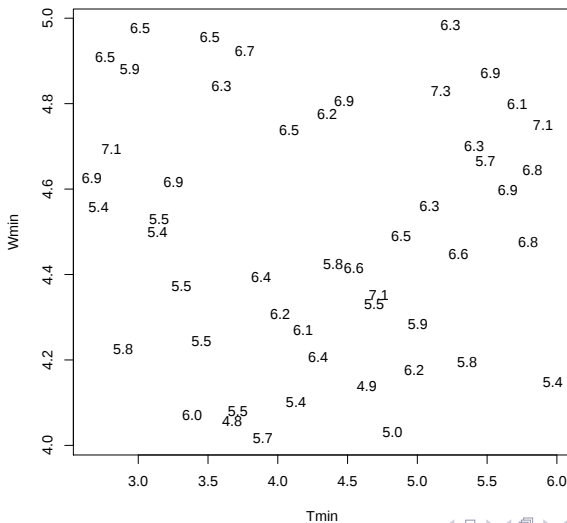
On dispose de **valeurs simulées en certains points** (ici 2D) échantillonnés sur un domaine (espace continu). On cherche à avoir une estimation (prédiction) de la réponse sur tout le plan

- Donner à voir (souvent en 2D)
 - courbes de niveau, images, perspective 3D, en dynamique 3D
- Prédire avec un modèle plus manipulable (méta-modèle)
 - fournir de l'intelligibilité (écriture analytique ou plus compréhensible)
 - réduire les temps de calcul (pour les prédictions et réutilisation)
 - permettre les calculs des indices de sensibilité (voir présentation de B. Iooss)
 - simplifier (choix de modèle)

$$y = \mathcal{G}(x^{(1)}, x^{(2)}, \dots, x^{(K)})$$

$$y \approx f(x^{(1)}, x^{(2)}, \dots, x^{(K)})$$

Visualisation de Résultats [▶ Back](#)



Objectifs et Enjeux

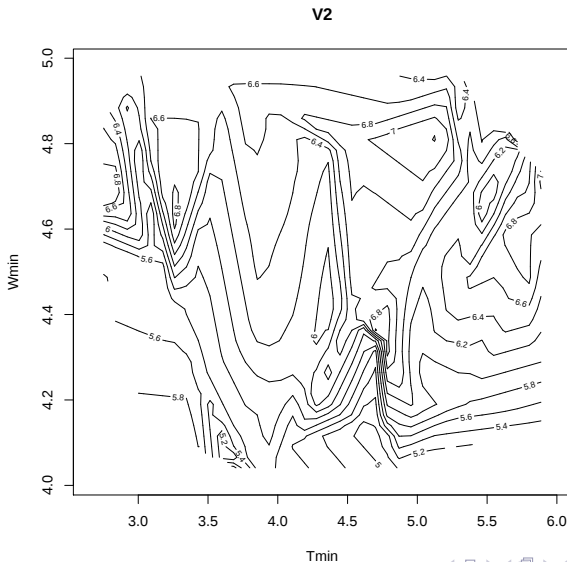
On dispose de **valeurs simulées en certains points** (ici 2D) échantillonnés sur un domaine (espace continu). On cherche à avoir une estimation (prédiction) de la réponse sur tout le plan

- Donner à voir (souvent en 2D)
 - courbes de niveau, images, perspective 3D, en dynamique 3D
- Prédire avec un modèle plus manipulable (méta-modèle)
 - fournir de l'intelligibilité (écriture analytique ou plus compréhensible)
 - réduire les temps de calcul (pour les prédictions et réutilisation)
 - permettre les calculs des indices de sensibilité (voir présentation de B. Iooss)
 - simplifier (choix de modèle)

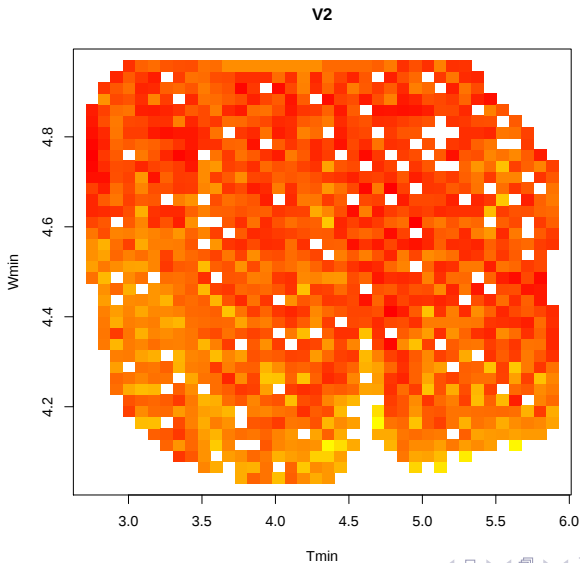
$$y = \mathcal{G}(x^{(1)}, x^{(2)}, \dots, x^{(K)})$$

$$y \approx f(x^{(1)}, x^{(2)}, \dots, x^{(K)})$$

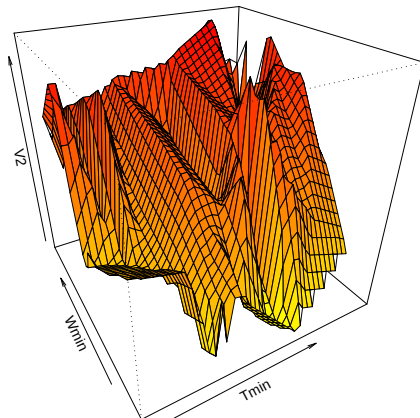
Visualisation de Résultats

[► Back](#)

Visualisation de Résultats

[► Back](#)

Visualisation de Résultats

[► Back](#)

Objectifs et Enjeux

On dispose de **valeurs simulées en certains points** (ici 2D) échantillonnés sur un domaine (espace continu). On cherche à avoir une estimation (prédiction) de la réponse sur tout le plan

- Donner à voir (souvent en 2D)
 - courbes de niveau, images, perspective 3D, en dynamique 3D
- Prédire avec un modèle plus manipulable (méta-modèle)
 - fournir de l'intelligibilité (écriture analytique ou plus compréhensible)
 - réduire les temps de calcul (pour les prédictions et réutilisation)
 - permettre les calculs des indices de sensibilité (voir présentation de B. Iooss)
 - simplifier (choix de modèle)

$$y = \mathcal{G}(x^{(1)}, x^{(2)}, \dots, x^{(K)})$$

$$y \approx f(x^{(1)}, x^{(2)}, \dots, x^{(K)})$$

Questions

Quelles sont les valeurs entre les points ? Comment les obtenir "facilement" ? Trois questions liées comme d'habitude :

- Quel plan d'expérience ?
- Quel modèle d'interpolation, d'estimation, de prédiction ?
- Quel critère de comparaison de modèles ?

⇒ Problématique de modélisation statistique "standard"

Basée principalement sur l'article de **Storlie et Helton, 2008**.
Multiple predictor smoothing methods for sensitivity analysis : Description of techniques. *Reliability Engineering and Safety*, **93**, 28-54.

Plan

1 Introduction

- Objectifs et Enjeux
- Questions

2 Méthodes de régression paramétrique

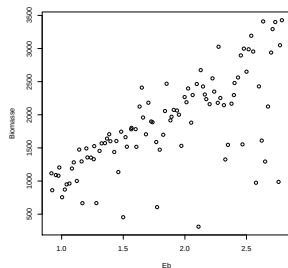
- Ecriture du modèle et estimation des paramètres
- Critères d'ajustement et comparaison de modèles
- Cas multidimensionnel
- Retour sur l'Analyse de Sensibilité
- Plans d'expérience optimaux

3 Méthodes de régression non-paramétrique

- Exemples de modèles non paramétriques
- Critère d'ajustement
- Extension multidimensionnelle
- Plans d'expérience

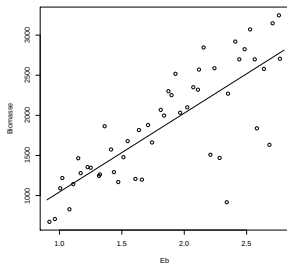
4 Références

Éléments principaux



- écriture du modèle de régression
- estimateurs des paramètres
- critère d'ajustement
- critère de comparaison de modèles (test sur les paramètres)
- extension multidimensionnelle
- plans d'expérience optimaux

Ecriture de modèle et estimation des paramètres



$$y = \mu + \beta x + \varepsilon$$

`lm( y ~ x )`

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

`lm( y ~ x + I(x^2) )`

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p + \varepsilon$$

`lm( y ~ poly(x,p) )`

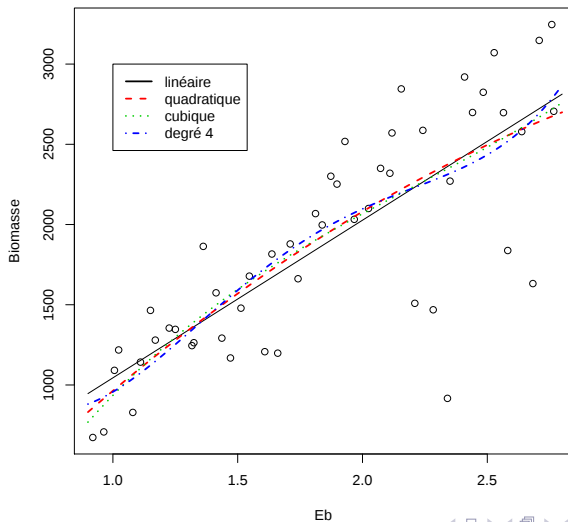
$$y = X\Theta + \varepsilon$$

$$\hat{\Theta} = (X'X)^{-1}X'Y$$

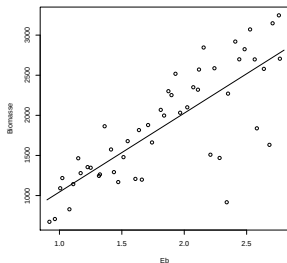
`coef( modèle estimé )`

fitted (aux x observés) ; predict (pour les nouveaux x)

Modèle wwdm et régressions linéaires

[Back](#)

Ecriture de modèle et estimation des paramètres



$$y = \mu + \beta x + \varepsilon$$

`lm( y ~ x )`

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

`lm( y ~ x + I(x^2) )`

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_p x^p + \varepsilon$$

`lm( y ~ poly(x,p) )`

$$y = X\Theta + \varepsilon$$

$$\hat{\Theta} = (X'X)^{-1}X'Y$$

`coef( modèle estimé )`

fitted (aux x observés) ; predict (pour les nouveaux x)

Critères d'ajustement

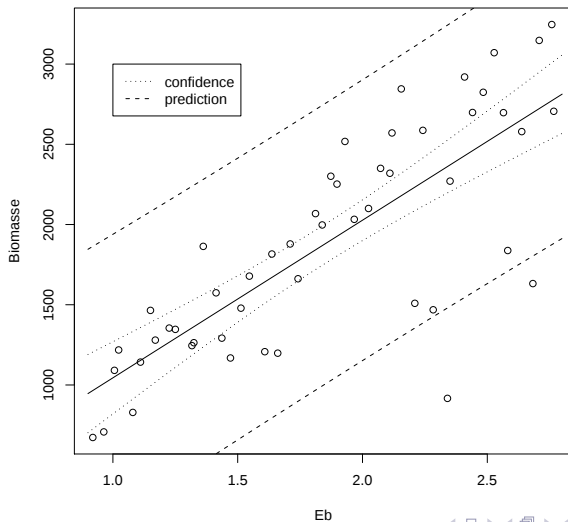
- Qualité d'ajustement sur un *a priori* sur les "critères"

Variance d'erreur estimée : $\widehat{\sigma^2} = \frac{SSE}{n-p} = \frac{\sum_i (y_i - \hat{y}_i)^2}{n-p}$

Coefficient de détermination : $R^2 = \frac{\sum_i (\hat{y}_i - \bar{y})^2}{\sum_i (y_i - \bar{y})^2}$

- Analyse des résidus
- Intervalles de confiance : `predict(  $\widehat{modele}$ , newdata, interval = "prediction")`

Modèle wwdm et intervalles de confiances

[▶ Back](#)

Comparaison de modèles

Analogie avec la méthodologie statistique "classique" bien que nous soyons dans le cas de l'expérimentation numérique.

On part du modèle le plus "paramétré" (complet) et on le simplifie par un modèle (1) plus grossier ($p_1 < p$)

Perte d'explication : $SSE^{(1)} - SSE^{complet}$ $p - p_1$

Base de la variabilité : $SSE^{complet}$ $n - p$

Test :
$$\frac{(SSE^{(1)} - SSE^{complet}) / (p - p_1)}{SSE^{complet} / (n - p)} \quad F(p - p_1; n - p)$$

summary ; anova ; step ; update

Cas multidimensionnel

Mêmes principes que pour l'unidimensionnel : les estimateurs des paramètres, le critère d'ajustement et les stratégies de tests sont les mêmes.

$$y = \beta_0 + \sum_{j=1}^K (\beta_j x_j + \beta_2 x_j^2) + \sum_{j=1}^K \sum_{k=j+1}^K \beta_{jk} x_j x_k + \varepsilon$$

$$\hat{\Theta} = (X'X)^{-1}X'Y$$

lm (Biomasse ~ polym(Eb,Lmax,B,degree=2))

Modèle wwdm : modèle multiplicatif

[► Back](#)

Call:

`lm(formula = Biomasse ~ polym(Eb, Lmax, B, degree = 2))`

Residuals:

Min	1Q	Median	3Q	Max
-1086.74	-147.33	25.77	208.71	622.26

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	1854.24	51.32	36.130	< 2e-16	***
polym(Eb, Lmax, B, degree = 2)1.0.0	3392.80	362.36	9.363	1.24e-11	***
polym(Eb, Lmax, B, degree = 2)2.0.0	-231.37	402.85	-0.574	0.56896	
polym(Eb, Lmax, B, degree = 2)0.1.0	1296.96	361.88	3.584	0.00091	***
polym(Eb, Lmax, B, degree = 2)1.1.0	1736.60	2387.38	0.727	0.47121	
polym(Eb, Lmax, B, degree = 2)0.2.0	-741.90	387.95	-1.912	0.06301	.
polym(Eb, Lmax, B, degree = 2)0.0.1	-1125.12	367.55	-3.061	0.00393	**
polym(Eb, Lmax, B, degree = 2)1.0.1	-4862.94	2799.03	-1.737	0.09002	.
polym(Eb, Lmax, B, degree = 2)0.1.1	1126.25	2963.05	0.380	0.70588	
polym(Eb, Lmax, B, degree = 2)0.0.2	-547.67	419.38	-1.306	0.19904	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 347 on 40 degrees of freedom

Multiple R-squared: 0.7953, Adjusted R-squared: 0.7492

F-statistic: 17.26 on 9 and 40 DF, p-value: 3.488e-11

Cas multidimensionnel

Mêmes principes que pour l'unidimensionnel : les estimateurs des paramètres, le critère d'ajustement et les stratégies de tests sont les mêmes.

$$y = \beta_0 + \sum_{j=1}^K (\beta_j x_j + \beta_2 x_j^2) + \sum_{j=1}^K \sum_{k=j+1}^K \beta_{jk} x_j x_k + \varepsilon$$

$$\hat{\Theta} = (X'X)^{-1}X'Y$$

lm (Biomasse ~ polym(Eb,Lmax,B,degree=2))

Ensuite, **méta** modèle sélectionné.

Modèle wwdm : méta modèle

[▶ Back](#)

Call:

```
lm(formula = Biomasse ~ Eb + Eimax + Eb:Eimax + K + Eb:Lmax + Eb:I(Lmax^2) + Eb:A + Eb:Lmax:A + I(A^2) +
+ Eb:I(A^2) + I(A^3) + Eb:B + A:B + Eb:A:B + I(A^2):B + Eb:I(B^2) + TI + Eimax:TI + Eb:A:TI + I(A^2):TI +
+ B:TI +
Eb:B:TI + A:B:TI + I(TI^2) + B:I(TI^2))
```

Residuals:

Min	1Q	Median	3Q	Max
-520.8318	-93.7253	-0.7456	93.0730	514.5766

Biomasse ~ polym(Eb, Eimax, K, Lmax, A, B, TI, degree=3)

Residual standard error: 142.9 on 974 degrees of freedom
 Multiple R-squared: 0.9602, Adjusted R-squared: 0.9591
 F-statistic: 938.9 on 25 and 974 DF, p-value: < 2.2e-16

141.2 on 880 degrees of freedom
0.9648, 0.9601
202.9 on 119 and 880 DF, p: < 2.2e-16

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-9.734e+03	1.465e+03	-6.645	5.03e-11	***
Eb	-1.148e+03	3.679e+02	-3.121	0.001856	**
Eimax	4.081e+03	1.480e+03	2.758	0.005926	**
K	4.498e+02	7.922e+01	5.678	1.80e-08	***
I(A^2)	-1.202e+08	9.965e+06	-12.064	< 2e-16	***
I(A^3)	9.261e+09	6.379e+08	14.518	< 2e-16	***
TI	2.113e+01	1.927e+00	10.965	< 2e-16	***
I(TI^2)	-9.356e-03	8.477e-04	-11.036	< 2e-16	***
Eb:Eimax	1.434e+03	3.196e+02	4.486	8.11e-06	***
Eb:Lmax	1.181e+02	6.816e+00	17.322	< 2e-16	***
Eb:I(Lmax^2)	-4.141e+00	3.876e-01	-10.683	< 2e-16	***
Eb:A	4.469e+05	4.647e+04	9.617	< 2e-16	***
Eb:I(A^2)	-2.714e+07	2.430e+06	-11.170	< 2e-16	***
Eb:B	1.267e+06	1.372e+05	-9.233	< 2e-16	***

R. Faivre

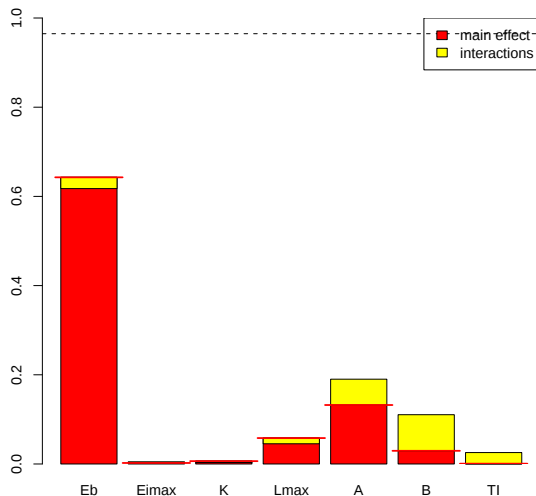
26 de 26 de réponse

Retour sur l'Analyse de Sensibilité

Calculer les contributions des facteurs

- part de variance expliquée par le modèle complet
 $lm(\text{Biomasse} \sim \text{polym}(\text{Eb}, \text{Eimax}, \text{K}, \text{Lmax}, \text{A}, \text{B}, \text{TI}, \text{degree} = 3))$
Multiple R-squared : 0.9648, Adjusted R-squared : 0.96
- part de variance expliquée par une seule variable
 $lm(\text{Biomasse} \sim \text{poly}(\text{Eb}, \text{degree} = 3))$
Multiple R-squared : 0.6428, Adjusted R-squared : 0.6417
- part de variance expliquée par tous sauf la variable
 $lm(\text{Biomasse} \sim \text{polym}(\text{Eimax}, \text{K}, \text{Lmax}, \text{A}, \text{B}, \text{TI}, \text{degree} = 3))$
Multiple R-squared : 0.3473, Adjusted R-squared : 0.2881

Modèle wwdm et contributions des facteurs



Plans d'expérience optimaux

Optimalité = estimabilité de tous les paramètres + régularité de la précision sur toute la zone en utilisant un minimum de points pour une qualité donnée.

Connu : nombre de points par facteur $>$ degré du polynome (3 points pour un polynome de degré 2)

- plans factoriels complets $\longrightarrow$ fractionnaire (même principe que pour l'AOV)
- Box-Behnken, ...
- Voir DOCUMENTATION/Surface_de_reponse sur la clé ; Govaerts par exemple

Plan

1 Introduction

- Objectifs et Enjeux
- Questions

2 Méthodes de régression paramétrique

- Ecriture du modèle et estimation des paramètres
- Critères d'ajustement et comparaison de modèles
- Cas multidimensionnel
- Retour sur l'Analyse de Sensibilité
- Plans d'expérience optimaux

3 Méthodes de régression non-paramétrique

- Exemples de modèles non paramétriques
- Critère d'ajustement
- Extension multidimensionnelle
- Plans d'expérience

4 Références

Principes généraux

- exemples de modèles non paramétriques
- critère d'ajustement
- critère de comparaison de modèles
- extension multidimensionnelle
- plans d'expérience

Fenêtre glissante (ksmooth) et régression locale (loess)

Fenêtre glissante : on se déplace le long de l'axe x et on estime avec les données autour (bandwidth d) selon un modèle k particulier

$$\hat{f}(x) = \sum_i^n k(x_i - x; h) y_i / \sum_i^n k(x_i - x; h)$$

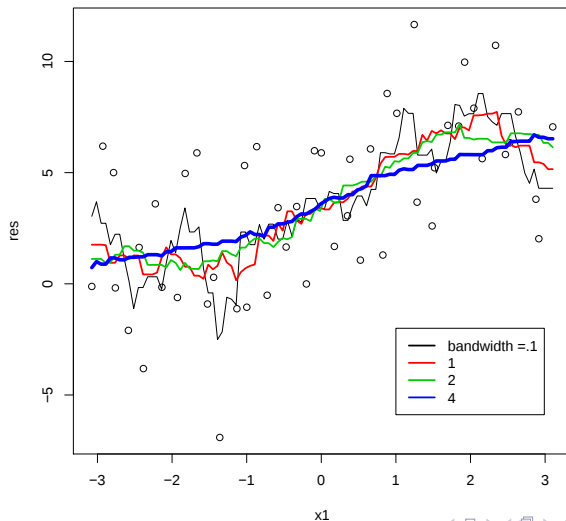
$$k(z, h) = \begin{cases} \mathbf{1}_{[0,d]}(|z|) & \text{moyenne glissante} \\ \frac{1}{h\sqrt{2\pi}} \exp(-\frac{z^2}{2h^2}) & \text{moyenne locale pondérée noyau gaussien} \end{cases}$$

Extension à la **régression polynomiale** ($\rightarrow 2$) locale et au **multidimensionnel**.

Avantage : suit bien les données

Inconvénient : il faut en avoir beaucoup, surtout en multidimensionnel

Modèle ishigami : ksmooth

[► Back](#)

Fenêtre glissante (ksmooth) et régression locale (loess)

Fenêtre glissante : on se déplace le long de l'axe x et on estime avec les données autour (bandwidth d) selon un modèle k particulier

$$\hat{f}(x) = \sum_i^n k(x_i - x; h) y_i / \sum_i^n k(x_i - x; h)$$

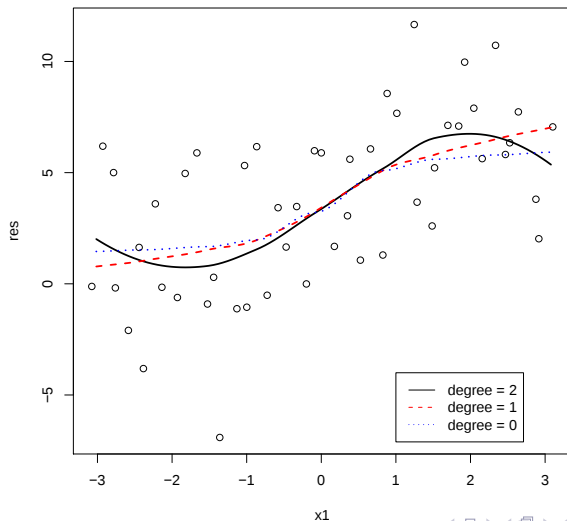
$$k(z, h) = \begin{cases} \mathbf{1}_{[0,d]}(|z|) & \text{moyenne glissante} \\ \frac{1}{h\sqrt{2\pi}} \exp(-\frac{z^2}{2h^2}) & \text{moyenne locale pondérée noyau gaussien} \end{cases}$$

Extension à la **régression polynomiale** ($\rightarrow 2$) locale et au **multidimensionnel**.

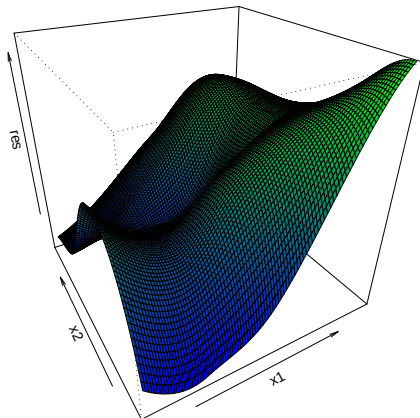
Avantage : suit bien les données

Inconvénient : il faut en avoir beaucoup, surtout en multidimensionnel

Modèle ishigami : loess (res $\sim$ x1, degree = 2)

[▶ Back](#)

Modèle ishigami : $\text{loess}(\text{res} \sim \text{loess}(x_1, x_2))$

[▶ Back](#)

Fenêtre glissante (ksmooth) et régression locale (loess)

Fenêtre glissante : on se déplace le long de l'axe x et on estime avec les données autour (bandwidth d) selon un modèle k particulier

$$\hat{f}(x) = \sum_i^n k(x_i - x; h) y_i / \sum_i^n k(x_i - x; h)$$

$$k(z, h) = \begin{cases} \mathbf{1}_{[0,d]}(|z|) & \text{moyenne glissante} \\ \frac{1}{h\sqrt{2\pi}} \exp(-\frac{z^2}{2h^2}) & \text{moyenne locale pondérée noyau gaussien} \end{cases}$$

Extension à la **régression polynomiale** ($\rightarrow 2$) locale et au **multidimensionnel**.

Avantage : suit bien les données

Inconvénient : il faut en avoir beaucoup, surtout en multidimensionnel

Splines de lissage

Critère à minimiser

$$\sum_i^n [y_i - \hat{f}(x_i)]^2 + \lambda \int_a^b [d^2 \hat{f}(x)/dx^2]^2 dx$$

(fidélité aux données + régularité)

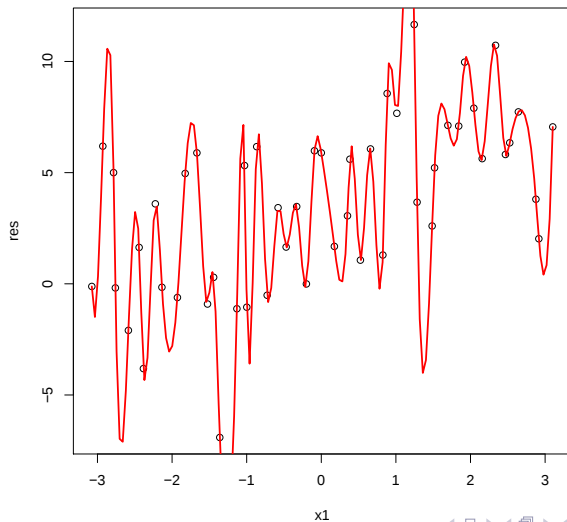
Ajout d'une pénalité qui lisse la fonction pour éviter les fluctuations trop grandes $\rightarrow$ paramètre de lissage λ

Décomposition sur une base de fonctions : **splines cubiques** et **thin plate**.

$$\hat{f}(x_i) = \sum_{j=1}^{p+q} \theta_j S_j(x_i)$$

$\Rightarrow$ proche de la régression.

Modèle ishigami : spline d'interpolation

[▶ Back](#)

Splines de lissage

Critère à minimiser

$$\sum_i^n [y_i - \hat{f}(x_i)]^2 + \lambda \int_a^b [d^2 \hat{f}(x)/dx^2]^2 dx$$

(fidélité aux données + régularité)

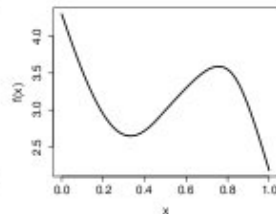
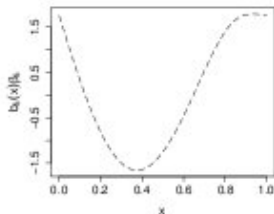
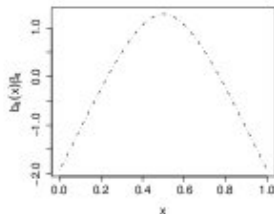
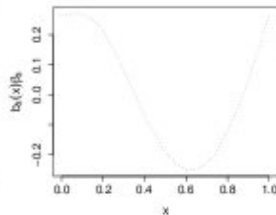
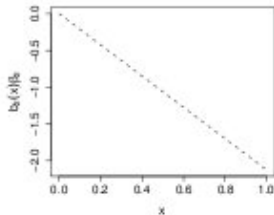
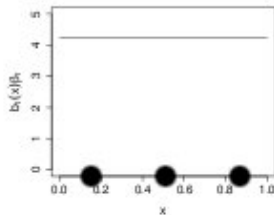
Ajout d'une pénalité qui lisse la fonction pour éviter les fluctuations trop grandes $\rightarrow$ paramètre de lissage λ

Décomposition sur une base de fonctions : **splines cubiques** et **thin plate**.

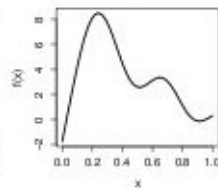
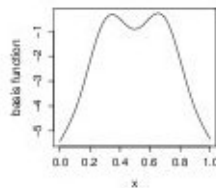
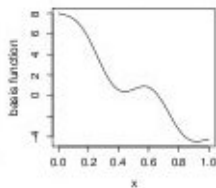
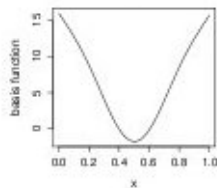
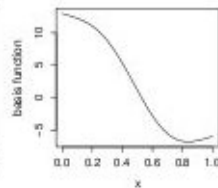
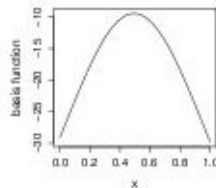
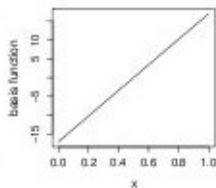
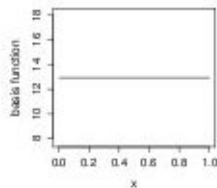
$$\hat{f}(x_i) = \sum_{j=1}^{p+q} \theta_j S_j(x_i)$$

$\Rightarrow$ proche de la régression.

Splines cubiques de lissage $s(\cdot, bs = "cs")$

[► Back](#)


Splines thin-plate de lissage $s(\cdot, bs = "tp")$

[▶ Back](#)


Splines de lissage

Critère à minimiser

$$\sum_i^n [y_i - \hat{f}(x_i)]^2 + \lambda \int_a^b [d^2 \hat{f}(x)/dx^2]^2 dx$$

(fidélité aux données + régularité)

Ajout d'une pénalité qui lisse la fonction pour éviter les fluctuations trop grandes $\rightarrow$ paramètre de lissage λ

Décomposition sur une base de fonctions : **splines cubiques** et **thin plate**.

$$\hat{f}(x_i) = \sum_{j=1}^{p+q} \theta_j S_j(x_i)$$

$\Rightarrow$ proche de la régression.

Critère d'ajustement

Degrés de liberté (en régression $n - p$) $\rightarrow$ on cherche à s'en rapprocher
 moyenne générale ($p=1$) ; interpolation, passe par tous les points ($p = n$)
 $df = n - \text{tr}(\text{projecteur } H)$; $\hat{y} = Sy$; $H = S(S'S)^{-1}S'$

- Utilisation de la valisation croisée (jackknife ou leave-one-out) et du PRESS (predicted sum of squares) pour la sélection du paramètre de lissage
- critère de comparaison de modèles (test sur les modèles) : analogie avec le paramétrique

Extension multidimensionnelle

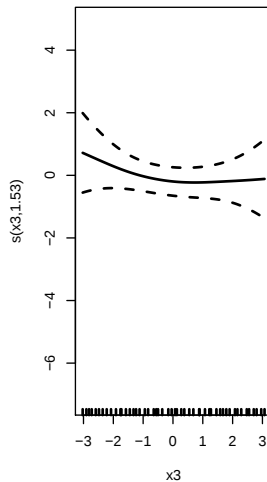
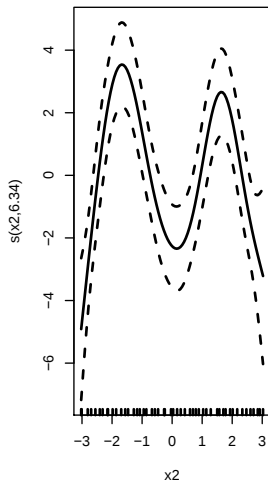
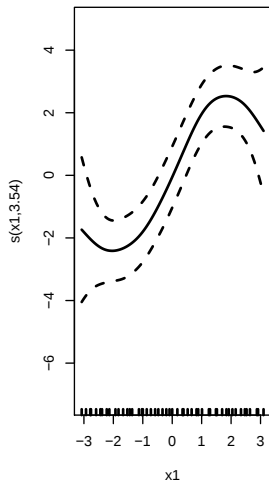
Le modèle additif généralisé (gam = generalized additif model)

$$y_i = \sum_i^k f^{(j)}(x_{ij}) + \varepsilon_i$$

`gam( res ~ s(x1) + s(x2) + s(x3) )`

Choix de modèle, tests : similarité avec lm. Utilisation de anova, step.

Modèle Ishigami : modèle multiplicatif

[▶ Back](#)

Extension multidimensionnelle

Le modèle additif généralisé (gam = generalized additive model)

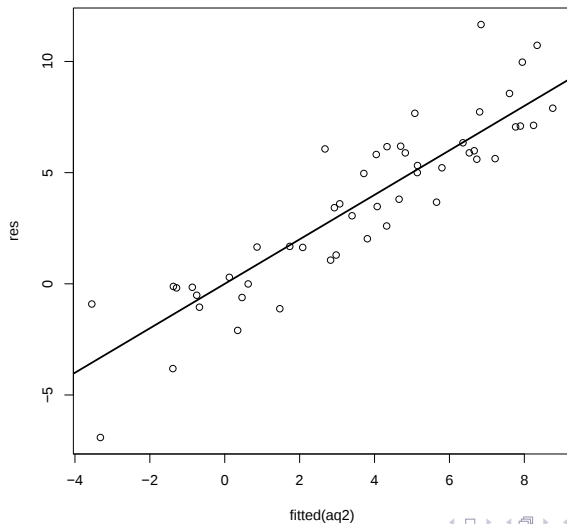
$$y_i = \sum_i^k f^{(j)}(x_{ij}) + \varepsilon_i$$

`gam( res ~ s(x1) + s(x2) + s(x3) )`

Choix de modèle, tests : similarité avec lm. Utilisation de anova, step.

`fitted(res) vs res`

Modèle Ishigami : ajustés vs observés

[▶ Back](#)

Extension multidimensionnelle

Le modèle additif généralisé (gam = generalized additif model)

$$y_i = \sum_i^k f^{(j)}(x_{ij}) + \varepsilon_i$$

`gam( res ~ s(x1) + s(x2) + s(x3) )`

Choix de modèle, tests : similarité avec lm. Utilisation de anova, step.

`fitted(res)` vs `res`

Modèles collent aux données : importance si possible de valider les qualités prédictives sur un nouveau jeu de données

Plans d'expérience

Hyper-cube latin (à faible discrédance et un peu d'aléatoire) pour identifier (répartition des points sur chaque facteur)

Echantillonnage par méthode de quasi-Monte Carlo utilisant une suite de Sobol à discrédance faible.

Plan

1 Introduction

- Objectifs et Enjeux
- Questions

2 Méthodes de régression paramétrique

- Ecriture du modèle et estimation des paramètres
- Critères d'ajustement et comparaison de modèles
- Cas multidimensionnel
- Retour sur l'Analyse de Sensibilité
- Plans d'expérience optimaux

3 Méthodes de régression non-paramétrique

- Exemples de modèles non paramétriques
- Critère d'ajustement
- Extension multidimensionnelle
- Plans d'expérience

4 Références

Références

- Storlie et Helton (2008). Multiple predictor smoothing methods for sensitivity analysis : Description of techniques. *Reliability Engineering and Safety*, **93**, 28-54.
- Myers, Montgomery et Anderson-Cook (2009). Response Surface Methodology : Process and Product Optimization Using Designed Experiments, Wiley, 704p.
- Simon Wood (2006). Generalized Additive Models : an introduction with R. CRC/Chapman & Hall.
- Cours de B Govaerts sur les Plans d'expérience pour l'estimation de surfaces de réponse (sur la clé et le site).
- Présentation de Ch. Thomas-Agnan sur les Estimateurs splines (sur la clé et sur le site).

Licence

Copyrights MEXICO 2010 ©

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation ; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".
see [http ://www.gnu.org/licenses/fdl.html](http://www.gnu.org/licenses/fdl.html)