

# THÈSE

présentée à

L'UNIVERSITÉ PARIS IX

pour obtenir

LE DIPLÔME DE DOCTEUR DE 3<sup>ème</sup> CYCLE

par

Christian CHARLES

Spécialité : Mathématiques Appliquées

Mention : Analyse des Données

## REGRESSION TYPOLOGIQUE ET RECONNAISSANCE DES FORMES

Soutenu le 10 juin 1977 devant la Commission composée de :

|     |                      |            |
|-----|----------------------|------------|
| MM. | <b>J. C. SIMON</b>   | Président  |
|     | <b>A. BENSOUSSAN</b> | Examineurs |
|     | <b>P. CAZES</b>      |            |
|     | <b>E. DIDAY</b>      | Invité     |
|     | <b>P. SOLETY</b>     |            |

BIBLIOTHEQUE DU CERIST

# THÈSE

présentée à

L'UNIVERSITÉ PARIS IX

pour obtenir

LE DIPLÔME DE DOCTEUR DE 3<sup>ème</sup> CYCLE

par

Christian CHARLES

Spécialité : Mathématiques Appliquées

Mention : Analyse des Données

## REGRESSION TYPOLOGIQUE ET RECONNAISSANCE DES FORMES

Soutenu le 10 juin 1977 devant la Commission composée de :

|     |               |            |
|-----|---------------|------------|
| MM. | J. C. SIMON   | Président  |
|     | A. BENSOUSSAN | Examineurs |
|     | P. CAZES      |            |
|     | E. DIDAY      | Invité     |
|     | P. SOLETY     |            |

BIBLIOTHEQUE DU CERIST

123456789

1011

*Je remercie,*

*Monsieur le Professeur SIMON qui me fait l'honneur de présider le jury de ma thèse.*

*Monsieur le Professeur BENSOUSSAN, Directeur de l'U.E.R. de Mathématiques de la Décision dont la richesse de l'enseignement m'a permis d'aborder cette étude.*

*Monsieur DIDAY, Maître de Conférences, qui m'a donné le sujet de la thèse, de m'avoir accueilli dans son équipe de l'IRIA, pour les conseils et les encouragements qu'il m'a prodigués tout au long de ce travail.*

*Monsieur CAZES, Maître Assistant, pour l'intérêt bienveillant avec lequel il a examiné ce travail.*

*Monsieur SOLETY du BRGM qui m'a permis d'appliquer ces algorithmes au domaine géologique.*

*Je tiens également à remercier, Mademoiselle ULMO et Monsieur COHEN, de Paris IX, qui par de judicieuses remarques m'ont apporté une aide efficace.*

*Monsieur SITBON, Directeur de l'UER d'Informatique de Gestion de Paris IX, pour l'aide matérielle efficace que son secrétariat m'a apporté une aide efficace.*

*Tous mes camarades de travail et tout particulièrement Mesdames DUPONT et SCHROEDER, Messieurs BOCHI, FEVRE et LECHEVALLIER.*



"Le seul trait distinctif des objets d'une loi statistique peut être leur appartenance à un certain groupe. Ils ne se distinguent que des objets extérieurs à leur groupe, ils ne se distinguent pas des objets intérieurs. La loi est établie sur la supposition qu'un membre d'un groupe est aussi propre que n'importe quel autre à satisfaire certaines conditions. Tous les caractères individualisants sont effacés par l'introduction de l'individu dans le groupe. Sa définition en tant qu'individu est une définition en tant que membre d'un groupe. On touche désormais le réel par son appartenance à une classe. C'est au niveau de la classe qu'il faudra chercher les propriétés du réel".

Joli!

Gaston BACHELARD

*Le nouvel esprit scientifique*



## TABLE DES MATIERES

### I INTRODUCTION

### II METHODE LOCALE DES MOINDRES CARRES ORDINAIRES

- 1 Rappels théoriques et inverses généralisées
- 2 L'algorithme
  - 2.1 Les notations employées
  - 2.2 Le critère
  - 2.3 Les fonctions  $f$  et  $g$
  - 2.4 Les suites  $u_n$  et  $v_n$
  - 2.5 La convergence de l'algorithme
- 3 Protection de la régression
  - 3.1 Comparaisons entre inverses généralisées et ridge régression
  - 3.2 Schéma général de l'algorithme
  - 3.3 Le programme informatique REGTYPO
- 4 Elimination de l'effet de cylindre
- 5 Choix optimal du nombre de classes
- 6 Méthode de prévision associée aux modèles linéaires locaux

### III METHODE LOCALE DES MOINDRES CARRES GENERALISES

- 1 Les estimateurs des moindres carrés généralisés
- 2 La méthode des moindres carrés généralisés locale
- 3 Cas particulier

### IV METHODE DE REGRESSION MULTIDIMENSIONNELLE LOCALE

- 1 Le modèle classique
- 2 Le modèle local

### V METHODE LOCALE DE REGRESSION SIMULTANEE

### VI METHODE DE REGRESSION ORTHOGONALE LOCALE

- 1 Régression orthogonale et composantes principales
- 2 La méthode de régression orthogonale locale
  - 2.1 Les notations
  - 2.2 Le critère à optimiser
  - 2.3 La fonction de représentation
  - 2.4 La fonction d'identification

- 2.5 Construction des suites  $u_n$ ,  $v_n$
- 2.6 Liens avec l'analyse factorielle typologique

#### VII METHODE DE REGRESSION PONDEREE LOCALE

- 1 le nouveau modèle
- 2 Le modèle de régression pondérée locale
  - 2.1 Les notations
  - 2.2 Le critère à optimiser
  - 2.3 La fonction de représentation
  - 2.4 La fonction d'identification
  - 2.5 la convergence de l'algorithme

#### VIII METHODE LOCALE DES MOINDRES VALEURS ABSOLUES

- 1 Estimation par la méthode des moindres valeurs absolues
  - 1.1 Problème de programmation linéaire
  - 1.2 Schéma général de la méthode des moindres valeurs absolues
- 2 la méthode locale des moindres valeurs absolues
  - 2.1 le critère
  - 2.2 la fonction d'identification
  - 2.3 la fonction de représentation
  - 2.4 la convergence de l'algorithme

#### IX REGRESSION LOCALE ET TEST DE CHOW

- 1 le test de Chow
- 2 Lien entre le critère et le test de Chow

#### X NOTIONS SUR L'ANALYSE DES SERIES TEMPORELLES DISCRETES

- 1 Modèle autorégressif d'ordre  $p$
- 2 Modèle stationnaire
  - 2.1 Autocovariance et autocorrélation théorique
  - 2.2 Estimateurs de l'autocovariance et de l'autocorrélation
  - 2.3 Fonction d'autocorrélation et équations de Yule-Walker
  - 2.4 Fonction d'autocorrélation partielle théorique
  - 2.5 Fonction d'autocorrélation partielle estimée

## XI ETUDE LOCALE DES SERIES TEMPORELLES

- 1 La méthode de lissage pour le modèle I
- 2 L'estimation du trend par un polynome
- 3 La méthode de régression polynomiale locale ( modèle [I] )
  - 3.1 Le principe
  - 3.2 Construction de la base de polynomes orthogonaux
  - 3.3 Le formalisme
  - 3.4 L'algorithme
  - 3.5 La convergence de l'algorithme
  - 3.6 Le programme REGPI
- 4 La méthode de régression polynomiale locale ( modèle [II] )
  - 4.1 Le problème
  - 4.2 L'estimation des coefficients de l'équation [I]
  - 4.3 La démarche générale
  - 4.4.1 L'espace des noyaux
  - 4.4.2 L'espace des partitions
  - 4.4.3 Notations
  - 4.4.4 Le critère à optimiser
  - 4.4.5 La fonction de représentation
  - 4.4.6 Obtention des parties contigues
  - 4.4.7 La construction de l'algorithme et sa convergence
  - 4.5 Le programme REGP3

## XII RECONNAISSANCE DES FORMES ET MEMORISATION CONDENSEE DE COURBES

- 1 Introduction
- 2 Le modèle
  - 2.1 Obtention des parties contigues
  - 2.2 Obtention des polynomes locaux
  - 2.3 Le formalisme

- 2.4 Le critère
- 2.5 L'algorithme et sa convergence
- 2.6 Le programme REGP2
- 3 Le modèle modifié
  - 3.1 La nouvelle fonction de représentation
  - 3.2 Le formalisme et la convergence
  - 3.3 Généralisation du modèles sous contraintes
  - 3.4 Le nouveau programme REGP2
- 4 Nouvel algorithme fournissant la partition en classes contigues
- 5 Généralisations
  - 5.1 Cas d'une fonction de  $R^2$ 
    - 5.1.1 Le formalisme
    - 5.1.2 Le programme Fortran REGP2B
  - 5.2 Cas de la minimisation de valeurs absolues

### XIII MODELE LINEAIRE TYPOLOGIQUE SUR VARIABLES HETEROGENES

- 1 Notations et hypotheses
- 2 Le problème
  - 2.1 L'espace des partitions et des fonctions de codage
  - 2.2 L'espace des noyaux
  - 2.3 Le critère à optimiser
  - 2.4 La fonction d'identification f
  - 2.5 La fonction de représentation et les opérateurs de codage
    - 2.5.1 Les opérateurs de codage
    - 2.5.2 La matrice  $\tilde{X}$  et le vecteur  $\tilde{Y}$
    - 2.5.3 L'algorithme de calcul des codages optimaux et des coefficients de régression
      - 2.5.3.1 Calcul des coefficients de régression
      - 2.5.3.2 Obtention des codages optimaux
      - 2.5.3.3 L'algorithme de codage et sa convergence
- 3 La convergence de l'algorithme de régression typologique sur variables hétérogènes

3.1 Les suites  $u_n$  et  $v_n$

3.2 La convergence

4 Applications et organigramme général

XIV APPLICATIONS DES DIVERS ALGORITHMES

1 Application de l'algorithme de régression typologique REGTYP0

2 Applications des programmes de mémorisation et de reconnaissance  
de courbes

2.1 Application du programme REGP 2

2.2 Application du programme SPLINE

2.3 Application du programme REGP 2 B

XV CONCLUSION

XVI BIBLIOGRAPHIE

XVII ANNEXES