

Nom et prénom	Etablissement	Grade	Qualité
OUASSOU Idir	École Nationale des Sciences Appliquées, Marrakech	PES	Président
KIOUACH Driss	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Rapporteur
BENMOUMEN Mohammed	Faculté des Sciences, Oujda	PES	Rapporteur
EL MELHAOUI Said	Faculté des Sciences Juridiques, Économiques et Sociales, Oujda	PES	Rapporteur
RACHDI Mustapha	Université Pierre-Mendès-France, Grenoble	PES	Examineur
RHOMARI Nouredine	Faculté des Sciences, Oujda	PES	Examineur
EL MATOUAT Abdelaziz	Ecole Normale Supérieure, Fès	PES	Invité
HAMZAOUI Hassania	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Directeur de thèse



Résumé :

Le sujet de cette thèse est d'étudier le problème de sélection de modèles de mélange par les critères d'information. Nous nous intéressons principalement aux critères d'Akaike AIC et de Schwarz BIC. Le critère de Schwarz BIC est convergent mais il peut conduire à une sous-paramétrisation du modèle lorsque la taille de l'échantillon est petite. Nous avons élaboré un critère de Schwarz modifié, noté BICr, en utilisant une technique de rééchantillonnage pour la sélection d'un modèle paramétrique quelconque tout en préservant sa propriété de convergence. La performance du nouveau critère par rapport aux critères AIC et BIC a été montrée sur des données simulées de modèles de mélange gaussien. Le critère d'Akaike AIC n'est pas convergent, il conduit asymptotiquement à une surparamétrisation stricte du modèle, et lorsque la taille de l'échantillon est petite, il peut conduire à une surestimation du nombre de composantes et du nombre de variables explicatives pour les modèles de mélange de régressions. Nous avons dérivé un nouveau critère, noté MRCv, pour l'estimation conjointe du nombre de composantes et du nombre de variables explicatives pour les modèles de mélange de régressions multivariées à partir de la log-vraisemblance des données complètes et de l'information de Kullback-Leibler. Avec des données simulées de modèles de mélange de régressions multivariées, nous avons montré l'intérêt du critère MRCv par rapport aux critères AIC et BIC. Ensuite, dans le même contexte, nous avons proposé un critère d'information, noté MRCsdv, basé sur la divergence symétrique de Kullback-Leibler et la log-vraisemblance des données complètes. Nous avons mis en évidence les performances du critère MRCsdv par rapport au critère MRCv sur des données simulées de modèles de mélange de régressions multivariées.

En présence de données manquantes, nous avons considéré la probabilité a posteriori des modèles candidats et l'algorithme EM pour développer un critère de type BIC, noté BICcd. La performance de ce critère par rapport au critère AICcd élaboré par Cavanaugh et al. et les critères AIC et BIC a été illustrée sur des données simulées de modèles autorégressifs et de modèles de mélange de régressions univariées.

Mots clés :

Sélection de modèles, Critères d'information, Algorithme EM, Modèles de mélange de lois gaussiennes, Modèles de mélange de régressions

MIXTURE MODELS SELECTION BY INFORMATION CRITERIA

Abstract:

The subject of this thesis is the study of the problem of a mixture model selection using information criteria. We are mainly interested in the Akaike criterion AIC and the Schwarz

Model selection, Information criteria, Resampling technique, EM algorithm, Gaussian mixture models, Regression mixture models