

Joint segmentation using mixed linear models: application to the analysis of multiple CGH arrays and climatic data

Emilie Lebarbier
AgroParisTech/INRIA

Resumen We consider the joint segmentation of multiple series in the sense that each series has its own segmentation which is achieved jointly with others. We use a mixed linear model with breakpoints which allows us to take into account for both covariates and correlations between series. First, for a fixed number of segments, we obtained the maximum likelihood estimators by using the EM algorithm. Then the number of segments is estimated by using a model selection strategy. In a computational point of view, we propose a new Dynamic Programming (DP) algorithm for the segmentation parameter. This method is applied in two different fields: first for the analysis of microarray CGH data where the objective is to detect chromosomal aberrations (within chromosome) and second for the analysis of harvest dates data where the aim is to detect changes in grapes harvest dates due to changes in agricultural practices and not to variations in climate.