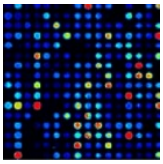
Séminaire – Math For Genomics

Séance du mercredi 3 avril 2019. 10h30.
Evry. IBGBI. LaMME.

Changepoint detection with kernels



Morgane Perre-Jean (CEA, Evry)

Title : Statistical methods for DNA copy number segmentation in cancer studies

Normal cells have two copies of DNA, inherited from each biological parent of the individual. Changes in DNA copy numbers are a hallmark of cancer cells. Therefore, the accurate detection and interpretation of such changes are two important steps toward improved diagnosis and treatment. In tumor cells, parts of a chromosome of various sizes (from kilobases to a chromosome arm) may be deleted or copied several times. The analysis of copy number profiles measured from high-throughput technologies such as array-comparative genomic hybridization (array-CGH), Single Nucleotide Polymorphism array (SNP array) or high-throughput DNA sequencing data (WGS and WES) raises a number of statistical and bioinformatic challenges.

DNA copy numbers in tumor cells are piecewise constant along the genome. SNPs arrays and sequencing techniques provide both the DNA copy number and the heterozygosity at a large number of position along the genome. As a result, the signal is composed of two dimensions. We first present a method to simulate realistic datasets of DNA copy number profiles. Then, we present an algorithm using the two dimensions of the signal. We show the performance of change-point detection is improved if we use the two dimensions. We implement both methods and simulation framework into an R package named `jointseg`.

This is joint work with Pierre Neuvial and Guillem Rigaill.

Alain Celisse (Université de Lille. Laboratoire Painlevé)

Coming soon !