

Sparse methods for learning on text documents

Francis Bach, SIERRA Project-team

STATWEB Associated team (with Laurent El Ghaoui, U.C. Berkeley)

Matrix factorization problems occur in many situations, like image processing, audio processing, bioinformatics, document processing, multi-task or unsupervised learning. In this talk, I will present recent contributions on sparse methods for matrix factorization, in particular on structured dictionary learning and low-rank decompositions, with applications to topic representations of large corpora of text document.