Thèse de doctorat (2015–2018)

Décomposition conjointe par approximation parcimonieuse en imagerie multispectrale astronomique

Contexte

L'étude de la cinématique des gaz au sein des galaxies est une clé pour comprendre l'histoire et l'évolution de l'Univers. Pour cela, les télescopes actuels sont capables de fournir des images multispectrales, c'est-à-dire des images 3D dont la troisième dimension correspond à la longueur d'onde : chaque pixel de l'image multispectrale est un spectre de raies. Ainsi, le spectre des galaxies est mesuré afin d'y analyser le décalage des raies (dû à l'effet Doppler) induit par la cinématique des galaxies.

L'analyse des observations est, encore maintenant, effectuée visuellement ou en utilisant des procédures interactives bas niveau. Une analyse plus fine et moins fastidieuse requiert des algorithmes performants de traitement du signal et des images.

Le problème revient à décomposer chaque spectre en une somme de raies décrites par des paramètres de forme (centre, largeur, amplitude). Ceux-ci évoluent lentement entre deux spectres voisins. Par ailleurs, le rapport signal-sur-bruit est généralement peu favorable. Pour ces raisons, il est important d'utiliser la redondance spatiale pour permettre un traitement fiable et robuste au bruit. En d'autres termes, la décomposition doit être effectuée conjointement sur tous les spectres.

Sujet

La décomposition spectroscopique sera considérée comme un problème inverse. Une régularisation parcimonieuse est par ailleurs naturelle puisque l'on cherche à décomposer les spectres en un faible nombre de raies. Il n'existe pas, à notre connaissance, d'approche pour effectuer la décomposition conjointe de spectres, hormis celle que nous avons développé depuis quelques années sur des séquences temporelles [1]. Celle-ci utilise l'approche bayésienne couplée à un algorithme MCMC. Le problème majeur de ce genre d'approche est la difficulté d'explorer convenablement l'espace des solutions en un temps de calcul raisonnable, notamment parce que les inconnues sont très corrélées.

Pour ces raisons, nous souhaitons explorer les techniques récentes d'approximation parcimonieuse (voir par exemple [2]). L'idée est de modéliser chaque spectre $\boldsymbol{y}_s$ comme le produit d'un « dictionnaire » sur-dimensionné $\boldsymbol{A}$ et d'un vecteur $\boldsymbol{x}_s$ dont seuls quelques éléments sont non nuls. Chaque colonne de $\boldsymbol{A}$ représente une raie potentiellement présente dans le spectre. L'estimation de $\boldsymbol{x}_s$ est obtenue en minimisant un critère composé d'un terme d'adéquation aux données et d'un terme favorisant la parcimonie ($\|\boldsymbol{x}_s\|_0$ ou $\|\boldsymbol{x}_s\|_1$ par exemple). Dans notre contexte, il est indispensable d'ajouter deux a priori.

D'une part, les amplitudes des raies $\boldsymbol{x}_s$ sont positives (raies d'émission). Cette contrainte de non-négativité oriente le choix des algorithmes d'optimisation. On peut citer par exemple la version non-négative de OMP [3] ou les algorithmes proximaux dont la forme générique permet naturellement d'imposer la contrainte de positivité (algorithmes ADMM [4] ou de seuillage itératif [5, 6, 7]). Un premier objectif de cette thèse est de mettre en concurrence ces algorithmes dans le contexte des problèmes inverses mal conditionnés.

Le deuxième a priori concerne le problème ouvert de la décomposition conjointe de spectres à travers la prise en compte du voisinage spatial (évolution lente des raies). Les méthodes permettant de lier les atomes entre eux (parcimonie par groupe ou structurée [8]) ne peuvent pas directement s'appliquer car l'évolution des raies est inconnue et variable d'un spectre à l'autre. Par ailleurs, il existe de nombreux travaux récents (parcimonie simultanée [9]) visant à décomposer conjointement un ensemble de données sur un dictionnaire. Cependant, l'hypothèse sous-jacente est trop forte pour pouvoir être appliquée ici. Il faut donc proposer de nouvelles solutions pour lier les vecteurs $\boldsymbol{x}_s$ deux à deux de façon plus flexible.

Enfin, les méthodes développées devront être appliquées à des images astronomiques réelles, dont les tailles peuvent comporter jusqu'à presque 4000 bandes dans le cas de MUSE. Il faudra donc mettre en œuvre des implémentations efficaces pour effectuer la décomposition en un temps raisonnable.

Compétences requises

Le candidat, titulaire d'un master 2 et/ou d'un diplôme d'ingénieur, devra avoir des connaissances solides en mathématiques et en traitement du signal et des images. Il devra savoir développer en programmation scientifique (Matlab, Python, ...) et maîtrisera l'anglais scientifique. Le candidat devra avoir l'ouverture scientifique et la capacité de dialoguer dans un contexte pluridisciplinaire.

Financement

Cette thèse bénéficie d'un financement de l'Agence nationale de la recherche dans le cadre du projet DSIM (dsim.unistra.fr) ; elle débutera au deuxième semestre 2015.

Encadrement

L'encadrement sera effectué par Vincent MAZET (ICube), Charles SOUSSEN (CRAN, co-directeur de thèse), et Christophe COLLET (ICube, directeur de thèse). Elle se déroulera au sein de l'équipe MIV du laboratoire ICube. L'inscription s'effectuera à l'école doctorale MS2I de l'Université de Strasbourg.

Candidature

Adressez votre candidature à Vincent MAZET (vincent.mazet@unistra.fr) en fournissant un CV, une lettre de motivation, une lettre de recommandation (au minimum) et les résultats universitaires des dernières années (avec classements, le cas échéant).

Références

[1] V. Mazet, S. Faisan, S. Awali, M.-A. Gaveau, L. Poisson, « Unsupervised Joint Decomposition of a Spectroscopic Signal Sequence » *Signal Processing*, vol. 109, p. 193–205, 2015.

[2] M. Elad, *Sparse and Redundant Representations*, Springer, 2010.

[3] A. Bruckstein, M. Elad, and M. Zibulevsky, « On the uniqueness of non-negative sparse solutions to underdetermined systems of equations », *IEEE Transactions on Information Theory*, vol. 54, p. 4813–4820, 2008.

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, « Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers », *Foundations and Trends in Machine Learning*, vol. 3, p. 1–122, 2010.

[5] A. Beck, M. Teboulle, « A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems », *SIAM Journal on Imaging Sciences*, vol. 2, p. 183–202, 2009.

[6] T. Blumensath, M. Davies, « Iterative Thresholding for Sparse Approximations », *The Journal of Fourier Analysis and Applications*, vol. 14, p. 629–654, 2008.

[7] P. L. Combettes and J.-C. Pesquet, « Proximal splitting methods in signal processing », *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer, p. 185–212, 2011.

[8] F. Bach, R. Jenatton, J. Mairal, G. Obozinski, « Optimization with Sparsity-Inducing Penalties », *Foundations and Trends in Machine Learning*, vol. 4, p. 1–106, 2012.

[9] M.E. Davies , Y.C. Eldar, « Rank Awareness in Joint Sparse Recovery », *IEEE Transactions on Information Theory*, vol. 58, p. 1135–1146, 2012.