

チュートリアル：非負値行列因子分解

亀岡 弘 和^{†1}

本チュートリアルでは、近年音響信号処理分野で広く用いられるようになった非負値行列因子分解について説明する。音楽音響信号の任意の短時間フレームにおけるスペクトルには、一般に複数の楽音スペクトルが混在しており、個々のスペクトルを分離抽出することは難しい。ここでもし、全フレームの観測スペクトルが比較的少数の楽音スペクトルだけで構成されていて、各フレームにおける観測スペクトルがこれらを適当な「ミキシングバランス」で混ぜ合わせたものと仮定できるならばどうだろうか。各楽音スペクトルの推定値が既に得られているなら、各フレームの観測スペクトルに当てはまるように、それぞれの楽音スペクトルのミキシングバランスを推定してやることで各楽音に帰属するスペクトルへの分解が可能になるし、このとき観測スペクトルに十分うまく当てはまらないのであれば、うまく当てはまるように各楽音スペクトル推定値を再度求め直してやることで分解精度をさらに高めることができるだろう。実は、観測スペクトログラムを行列と見なして非負値行列因子分解の反復アルゴリズムを適用することは、これらの処理ステップを繰り返していることに相当している。実際、音楽信号は、限られた種類の音高および楽器音のスペクトルしか出現しえない、という仮定がある程度成り立つ場合が多く、このことが音楽スペクトログラムの分解方法として非負値行列因子分解が有効である理由の一つである。本チュートリアルでは、非負値行列因子分解の基本性質、非負値制約によってもたらされる効果、分解表現を得るための反復アルゴリズムの導出方法などを概説した上で、非負値行列因子分解による具体的な応用例を紹介する。

Tutorial: Non-negative Matrix Factorization

HIROKAZU KAMEOKA^{†1}

In this tutorial, I will give a brief introduction to non-negative matrix factorization (NMF), which has attracted a lot of attention in the field of audio signal processing in recent years. I will mention the basic properties of NMF, the effects induced by the non-negative constraint, how to derive the iterative algorithm for NMF, and some of the attempts that have been made to apply NMF to audio processing problems.

参考文献

- 1) D. D. Lee and H. S. Seung, “Learning the parts of objects with nonnegative matrix factorization,” *Nature*, vol. 401, pp. 788–791, 1999.
- 2) P. Paatero, U. Tapper, “Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values,” *Environmetrics* 5: 111–126, 1994.
- 3) P. Smaragdis and J.C. Brown, “Non-negative matrix factorization for music transcription,” *Proc. WASPAA 2003*, pp. 177–180, 2003.
- 4) D.D. Lee and H.S. Seung, “Algorithms for nonnegative matrix factorization,” In *Adv. Neural Information Processing Systems 13 (NIPS*2000)*, pp. 556–562, 2000.
- 5) P. Smaragdis, B. Raj, M.V. Shashanka, “Supervised and semi-supervised separation of sounds from single-channel mixtures,” In *Proc. 7th International Conference on Independent Component Analysis and Signal Separation*, pp. 414–421, 2007.
- 6) P. Smaragdis and B. Raj, “Example-driven bandwidth expansion,” In *Proc. 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2007)*, pp. 135–138, 2007.
- 7) 亀岡, 後藤, 嵯峨山, “スペクトル制御エンベロープによる混合音中の周期および非周期成分の選択的イコライザ,” *情報処理学会研究報告*, 2006-MUS-66-13, pp. 77–84, 2006.
- 8) C. Févotte, N. Bertin and J.-L. Durrieu, “Nonnegative matrix factorization with the Itakura-Saito divergence with application g pp to music analysis,” *Technical Report TELECOM ParisTech 2008D006*, 2008.
- 9) S. Eguchi and Y. Kano, “Robustifying maximum likelihood estimation,” *Technical report, Institute of Statistical Mathematics, Research Memo. 802*, 2001.
- 10) M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono and S. Sagayama, “Convergence-guaranteed multiplicative algorithms for nonnegative matrix factorization with betadivergence,” In *Proc. 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010)*, in CD-ROM, 2010.
- 11) B.A. Olshausen, and D.J. Field, “Emergence of simple-cell receptive field properties by learning a sparse code for natural images,” *Nature*, vol. 381, pp. 607–609, 1996.
- 12) T. Hofmann, “Probabilistic latent semantic analysis,” In *Proc. Uncertainty in Artificial Intelligence (UAI’99)*, pp. 289–296, 1999.
- 13) 亀岡, 小野, 柏野, 嵯峨山, “複素 NMF: 新しいスパース信号分解表現と基底系学習アルゴリズム,” *日本音響学会 2008 年秋季研究発表会講演論文集 2-8-13*, pp. 657–660, 2008.

^{†1} 東京大学 大学院情報理工学系研究科 / NTT コミュニケーション科学基礎研究所