第35回信号処理シンポジウム
2020年11月4日～6日（リモート開催）

マルチチャンネルベータダイバージェンスの導出

Introduction of multi-channel beta-divergence

太刀岡勇気
デンソーアイティーラボラトリ

Yuuki TACHIOKA
Denso IT Laboratory

アブストラクト　β-divergence は音源分離に使われる種々の距離尺度を統合した距離尺度である．ここではマルチチャンネル版の β-divergence を提案し，それが既に提案されているユークリッド距離，マルチチャンネル板倉斎藤距離を含む一般化になっていることを確認した．

1　はじめに

β-divergence は音源分離に使われる種々の距離尺度を統合した距離尺度である [1]．スペクトル x と y の距離を測るのに使われ，ユークリッド距離，Kullback-Leibler(KL) divergence，板倉斎藤距離の一般化となっていることが知られている．これを利用した非負値行列因子分解のアルゴリズムも提案されている [2], [3]．一方で文献 [4]–[6] においては，マルチチャンネルの信号処理のため，n チャンネル分の x，y の空間相関行列 $\boldsymbol{X}$，$\boldsymbol{Y}$ を対象として，ユークリッド距離，板倉斎藤距離が提案されているが，これの一般化はなされていない．ここではこれらを含むマルチチャンネルにおける β-divergence を提案し，それが，マルチチャンネルに拡張したユークリッド距離，マルチチャンネル板倉斎藤距離を含むことを示したうえで，新たにマルチチャンネル KL divergence を導出する．

2　ベータダイバージェンス (β-divergence)

β-divergence は 2 つのスカラー x と y の距離を測る場合，

$$d_\beta = \frac{y^\beta}{\beta(\beta-1)} + \frac{x^\beta}{\beta} - \frac{yx^{\beta-1}}{\beta-1} \tag{1}$$

のようにあらわされる．例えば，$\beta = 2$ のとき，ユークリッド距離

$$d_2 = \frac{1}{2}(x-y)^2 \tag{2}$$

$\beta = 1$ のとき，KL divergence

$$d_1 = y\log\frac{y}{x} + (x-y) \tag{3}$$

$\beta = 0$ のとき，板倉斎藤距離

$$d_0 = \frac{y}{x} - \log\frac{y}{x} - 1 \tag{4}$$

となる．音源分離の応用においては，単一チャンネルのスペクトル (スカラー)x と y の距離を測るのに使われる．

3　マルチチャンネル β-divergence の提案

ここでは $\boldsymbol{X}$,$\boldsymbol{Y} \in \mathbb{C}^{n\times n}$ が Hermite 行列のときのマルチチャンネル β-divergence

$$d_\beta = \frac{tr(\boldsymbol{Y}^\beta)}{\beta(\beta-1)} + \frac{tr(\boldsymbol{X}^\beta)}{\beta} - \frac{tr(\boldsymbol{Y}\boldsymbol{X}^{\beta-1})}{\beta-1} \tag{5}$$

を提案する[1]．ただし，$tr(\boldsymbol{Y}^\beta) = \sum_n (\lambda_{\boldsymbol{Y}})^\beta$ で，$\lambda_{\boldsymbol{Y}}$ は $\boldsymbol{Y}$ の固有値である．以下，この提案 (5) が，すでに [6] で提案されているマルチチャンネルに拡張したユークリッド距離，マルチチャンネル板倉斎藤距離を含むことを示す．また，新たにマルチチャンネル KL divergence を導出する．$\boldsymbol{X}$ と $\boldsymbol{Y}$ は任意の Hermite 行列に適応できるが，例えば n チャンネルで観測した信号の空間相関行列に対して用いることができる．

3.1　マルチチャンネルユークリッド距離

$\beta = 2$ のときは，

$$\begin{aligned} d_2 &= \frac{tr(\boldsymbol{Y}^2)}{2} + \frac{tr(\boldsymbol{X}^2)}{2} - \frac{tr(\boldsymbol{X}\boldsymbol{Y})}{2} - \frac{tr(\boldsymbol{Y}\boldsymbol{X})}{2} \\ &= \frac{1}{2}tr(\boldsymbol{X}^H\boldsymbol{X} - \boldsymbol{X}^H\boldsymbol{Y} - \boldsymbol{Y}^H\boldsymbol{X} + \boldsymbol{Y}^H\boldsymbol{Y}) \\ &= \frac{1}{2}tr((\boldsymbol{X}-\boldsymbol{Y})^H(\boldsymbol{X}-\boldsymbol{Y})) = \frac{1}{2}||\boldsymbol{X}-\boldsymbol{Y}||_F \end{aligned} \tag{6}$$

$tr(\boldsymbol{X}\boldsymbol{Y}) = tr(\boldsymbol{Y}\boldsymbol{X})$ であることと，$\boldsymbol{X}$,$\boldsymbol{Y}$ が Hermite 行列のときは $\boldsymbol{Y}^2 = \boldsymbol{Y}^H\boldsymbol{Y}$，$\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{X}^H\boldsymbol{Y}$ であることを利用している．H は Hermite 転置．$||\cdot||_F$ はフロベニウスノルムであり，マルチチャンネルに拡張したユークリッド距離 d_{Eu}[6] に一致する．

3.2　マルチチャンネル板倉斎藤距離

$\beta = 0$ のときは，

$$d_0 = \lim_{\beta\to 0}\left[-\frac{tr(\boldsymbol{Y}^\beta - \boldsymbol{X}^\beta)}{\beta} - tr(\boldsymbol{Y}^\beta)\right] + tr(\boldsymbol{Y}\boldsymbol{X}^{-1}) \tag{7}$$

[1] $\boldsymbol{X}$，$\boldsymbol{Y}$ が Hermite 行列なら d_β は非負の実数

と変形できるので，

$$d_0 = -\log\det \boldsymbol{Y}\boldsymbol{X}^{-1} - N + tr(\boldsymbol{Y}\boldsymbol{X}^{-1}) \tag{8}$$

となる．ここで

$$\begin{aligned}\lim_{\beta\to 0}\frac{tr(\boldsymbol{Y}^{\beta}-\mathbb{I})}{\beta} &= \sum_n \log\lambda_{\boldsymbol{Y}} \\ &= \log\prod_n \lambda_{\boldsymbol{Y}} = \log\det\boldsymbol{Y}\end{aligned} \tag{9}$$

また，$\lim_{\beta\to 0}\boldsymbol{Y}^{\beta} = \mathbb{I}$ を利用した ($\mathbb{I}\in\mathbb{R}^{n\times n}$ は単位行列)．これはマルチチャンネル板倉斎藤距離 d_{IS}[6] に一致する．

3.3 マルチチャンネル KL divergence の導出

$\beta = 1$ のときは，

$$d_1 = \lim_{\beta\to 1}\frac{tr(\boldsymbol{Y}(\boldsymbol{Y}^{\beta-1}-\boldsymbol{X}^{\beta-1}))}{\beta-1} + tr(\boldsymbol{X}-\boldsymbol{Y}) \tag{10}$$

と変形できるので，

$$\begin{aligned}d_1 = &\sum_n \lambda_{\boldsymbol{Y}}\log\lambda_{\boldsymbol{Y}} \\ &- tr(\boldsymbol{Y}\boldsymbol{U}\log[\boldsymbol{\Lambda}]\boldsymbol{U}^{-1}) + tr(\boldsymbol{X}-\boldsymbol{Y})\end{aligned} \tag{11}$$

となる．ここで，

$$\begin{aligned}\lim_{\beta\to 0}\frac{tr(\boldsymbol{Y}(\boldsymbol{Y}^{\beta}-\mathbb{I}))}{\beta} &= \sum_n \lambda_{\boldsymbol{Y}}\log\lambda_{\boldsymbol{Y}} \\ \lim_{\beta\to 0}\frac{tr(\boldsymbol{Y}(\boldsymbol{X}^{\beta}-\mathbb{I}))}{\beta} &= tr(\boldsymbol{Y}\boldsymbol{U}\log[\boldsymbol{\Lambda}]\boldsymbol{U}^{-1})\end{aligned} \tag{12}$$

を利用した．ここで $\log[\cdot]$ は非負の行列要素に対する対数演算を示す．また $\boldsymbol{X}$ の固有値分解を $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1}$ とする．著者らの知見によれば，マルチチャンネルに拡張した KL divergence は初出である．

4 まとめ

マルチチャンネル版の β-divergence を提案し，それが既に提案されているユークリッド距離，マルチチャンネル板倉斎藤距離を含む一般化になっていることを確認した．また，新たにマルチチャンネル KL divergence を導出した．

参考文献

[1] S. Eguchi and Y. Kano, “Robustifying maximum likelihood estimation,” Tokyo Institute of Statistical Mathematics Technical report (2001).

[2] M. Minami and S. Eguchi, “Robust blind source separation by beta divergence,” Neural Computation, **14**, 1859–1886 (2002).

[3] M. Nakano, H. Kameoka, J. Le Roux, Y. Kitano, N. Ono, and S. Sagayama, “Convergence-guaranteed multiplicative algorithms for non-negative matrix factorization with beta-divergence,” Proceedings of MLSP, pp.283–288 (2010).

[4] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, “Efficient algorithms for multichannel extensions of Itakura-Saito nonnegative matrix factorization,” Proc. ICASSP, pp.261–264 (2012).

[5] 澤田宏，亀岡弘和，荒木章子，上田修功，“非負値行列因子分解 NMF の多チャンネル拡張,” 信号処理シンポジウム，pp.1–6 (2012).

[6] H. Sawada, H. Kameoka, S. Araki, and N. Ueda, “Multichannel extensions of non-negative matrix factorization with complex-valued data,” IEEE Transactions on Audio, Speech, and Language Processing, **21**, 971–982 (2013).