独立性及び複帯域性に基づいた自然電磁場データからのノイズ除去の試み

佐藤 真也 [1]; 後藤 忠徳 [2]
[1] 京大・工; [2] 京都大学

Challenge in noise reduction from natural EM data based on independence and multiband

Shinya Sato[1]; Tada-nori Goto[2]
[1] Engineering, Kyoto Univ.; [2] Kyoto Univ.

It is important to conduct a MT survey near a population area when surveying an active fault for evaluating disaster risk. However, it is difficult to visualize the subsurface resistivity structure because electromagnetic (EM) noises contaminate the EM data acquired at such area. Near a population area, two conditions are expected: a) most time-series of EM data are contaminated by large EM noises, and b) the number of EM signals and noises is more than the number of observed components (i.e., underdetermined system). Many methods for noise reduction of MT data have been suggested and have succeeded to derive MT responses with high accuracy from noisy data (e.g., Egbert, 1997; Chave and Thomson, 2004; Weckmann et al., 2005). However, these are not assumed to be used under such a great noise-level environment. The noise reduction method based on Independent Component Analysis (ICA) at Frequency Domain was suggested by Sato et al. (2017; 2019, submitted), and can decompose noisy data into EM signals and EM noises with high accuracy. However, the accuracy of ICA-based method degrades when three or more large noises are included in MT data at a station. We therefore must remove EM noises directly from EM data under an underdetermined system in order to derive MT responses at a population area with high accuracy. We newly propose noise reduction method maintaining its accuracy under an underdetermined system focusing on their statistical independence and characteristics of multiband. The proposal method is named Independent Multiband EM-Signal Analysis (IMEMSA).

Here, we introduce the algorithm of IMEMSA shortly. IMEMSA is designed combining ICA and Nonnegative Matrix Factorization (NMF: Kameoka et al., 2009 or Multi-Channel NMF: Sawada et al., 2013; Sato and Goto, 2018). ICA can estimate source signals with high accuracy under overdetermined and determined systems. However, it cannot be applied theoretically under an underdetermined system. NMF can estimate the components Basis constituting source signals focusing on signals' characteristics of multiband theoretically even under an underdetermined system. However, NMF has its instability to work properly. Additionally, NMF's cost function focuses on only large amplitude, and this is the reason for its difficulty to apply to EM data including large noises. For example, when electric field data and magnetic field data have S/N below 0.1 and 0.5 respectively, EM signals cannot possibly be reconstructed by NMF. At first in IMEMSA, we apply ICA to EM data and obtain independent separated components. Using magnetic field data at a remote site, we divide the separated components into two class: a) including EM signals much (defined here as Y_s) and b) including noises much (Y_n). We make the spectrogram of Y_s combining Y_s obtained at each frequency. Applying NMF to the spectrograms, NMF can start near the EM signals and derive Basis. We apply the same process to Y_n, and NMF can start near the noises and derive Basis. These obtained Basis are classified into noises and EM signals using reference data. As a result, we can separate and extract EM signals and noises from the underdetermined system.

In the presentation, we introduce the detail algorithm of IMEMSA and the application example to synthetic data.

災害リスク評価などのために，MT 法が活断層調査に用いられているが，大都市近郊での実施例は限定的である．都市域ではノイズの電磁（EM）データへの混入により，地下構造解析が困難となる場合が多いためである．この時，時系列の大部分に大きな振幅を持ったノイズが混入していること，かつ，観測データの数よりも EM 信号・ノイズの数の方が多いこと (劣決定問題) が予想される．様々なノイズ除去法がこれまで提案されてきたが (例えば，Egbert, 1997; Chave and Thomson, 2004; Weckmann et al., 2005)，このような環境下での解析は想定されていない．また，独立成分分析 (ICA) に基づき直接ノイズを EM データから分離する手法 (Sato et al., 2017; Sato et al., submitted, 2019) も提案されたが，1 箇所の EM データに対して大きなノイズが 3 成分以上混合した場合，分離精度は低下する．それ故，都市域での EM データから高い精度で MT 応答関数を導出するためには，劣決定条件下で直接 EM データからノイズを抽出・除去する必要がある．本研究では，信号とノイズ間における 1) 独立性および 2) 複帯域性に着目することで，劣決定条件下においても，EM データからノイズを除去できる手法，すなわち独立複帯域 EM 信号解析 (Independent Multiband EM-Signal Analysis: IMEMSA) を新たに提案する．

ここでは IMEMSA のアルゴリズムについて，簡単に紹介する．IMEMSA は ICA と非負値行列因子分析 (NMF: Kameoka et al., 2009; Sawada et al., 2013; Sato and Goto, 2018) の両アルゴリズムを用いたものである．ICA は優決定条件下では，原信号の推定を高い精度で可能である一方，劣決定条件下では適用することが難しい．NMF では，信号の複帯域性に着目し，理論上では劣決定条件下においても，原信号を構成する要素“ Basis ”を推定可能な手法であるが，信号分離の安定性に問題が残る．NMF の目的関数は，大きな振幅の信号・ノイズへの重みが大きく，(音声データとは異なり）様々な信号・ノイズを含む EM データに適用は困難である．例えば，電場の S/N が 0.1，磁場の S/N が 0.5 程度の場合，EM 信号を NMF では表現できない可能性がある．そこで IMEMSA では，まず EM データに対し ICA を適用し独立な分離信号を得る．分離信号に対して，参照観測点で取得された磁場データを基準に，EM 信号を多く含んでいるもの (Y_s) と，ノイズ

を多く含んでいるもの (Y_n) とに分類する．各周波数で得られた Y_s を結合し，Y_s に関するスペクトログラムを生成し，NMF を適用することで，EM 信号の近傍で Basis を探索することが可能である．Y_n に関しても同じ処理を施すことで，ノイズの近傍で Basis を探索することが出来る．こうして得られた Basis から，Reference データを用いて，EM 信号・ノイズに相当するものを判別する．その結果，劣決定条件下においても，EM 信号・ノイズの分離・抽出が可能となる．

発表では，IMEMSA の詳細なアルゴリズムと，仮想データへの適用例を紹介する．