

カーネル密度推定と混合正規分布モデルによる地震動波形の包絡線近似

岐阜大学流域圏科学研究センター 正会員 ○久世益充

岐阜大学工学部 正会員 能島暢呂

1. はじめに

地震動特性の評価は地震工学における基本的な課題であり、地震動ごとに異なる特徴を的確に捉えることが重要である。筆者らは正規化加速度累積パワー曲線(Husid plot)に基づいた地震動の経時特性を評価し、その評価方法を検討した^{1),2)}。本研究では、Husid Plot より算出した特徴ベクトルを用いて、地震動の包絡線を確率密度分布で近似する。近似では、ノンパラメトリックモデルであるカーネル密度推定と、セミパラメトリックモデルである混合正規分布モデルをそれぞれ適用し、地震動の包絡線近似について、より少ない次元で特徴ベクトルを縮約する手法について考察する。

2. 経時特性の特徴ベクトルとカーネル密度推定

2.1 特徴ベクトルの算出

地震動経時特性は、水平1成分ごとの地震動加速度波形 $A(t)$ の全パワーで正規化された累積パワー曲線(Husid plot) $P_c(\tau)$ より算出する。筆者ら¹⁾は $P_c(\tau)$ を分割数 $n=100$ 、つまり1%刻みで離散化し、99個のパーセンタイル値 $t_i = P_c^{-1}(\tau)$ ($i=1, \dots, 99$) を99次元の特徴ベクトル $\mathbf{t}=\{t_i\}$ とした。本研究では表1に示すように、分割数 n を200, 500, 1000に拡張し、それぞれ、199, 499, 999次元の特徴ベクトルを算出して比較する。

表1 特徴ベクトルの一覧

分割数 n	次元	$P_c(\tau)$ の範囲(%)	刻み(%)
100	99	1.0~99.0	1.0
200	199	0.5~99.5	0.5
500	499	0.2~99.8	0.2
1000	999	0.1~99.9	0.1

2.2 カーネル密度推定

表1に示す99~999次元の特徴ベクトルを対象に、カーネル密度推定により地震動経時特性を包絡線近似する。特徴量 $\mathbf{t}=\{t_i\}$ が与えられたときの密度関数 $p(t)$ は、次式で表される³⁾。

$$p(t) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{t-t_i}{h}\right) \quad K(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \quad (1), (2)$$

$$h = 0.9\sigma n^{-0.2} \quad \sigma = \sqrt{\sum_{i=1}^n (t_i - \bar{t})^2 / n} \quad (3), (4)$$

ここに、 K はカーネル関数(ここではガウス分布)、 n はデータ数である。 h はバンド幅(平滑化パラメータ)であ

り、Silvermanの方法⁴⁾では、標準偏差 σ (または四分位範囲 $\div 1.34$ の小さい方)とデータ数 n より与える。

2.3 カーネル密度推定による経時特性の包絡線近似

2011年東北地方太平洋沖地震において、防災科学技術研究所 K-NET⁵⁾で観測された加速度記録(EW成分)を用いた。図1に示す加速度波形は、(a)地震動が長く標準的な包絡形状を有した波形、(b)振幅が鋭く単峰形の波形、(c)複数の大振幅を持つ双峰形の波形である。

カーネル密度推定結果を図2に示す。図中下端の赤い分布は、分割数が最も多い999次元($n=1000$)における特徴ベクトルの分布である。2.2に前述したように、バンド幅はデータ数とばらつきに依存する。表2に各波形のバンド幅を示す。特徴ベクトルの標準偏差および次元(分割数)に依存してバンド幅が異なることがわかる。例えば、同じ波形でも、99次元($n=100$)に対して、999次元($n=1000$)では0.63倍の幅となる。

図2中赤線で示した、真値に最も近い999次元($n=1000$)の包絡線近似と比較すると、双峰形の包絡線

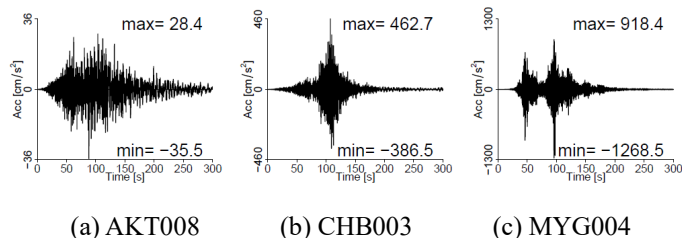


図1 加速度波形の比較

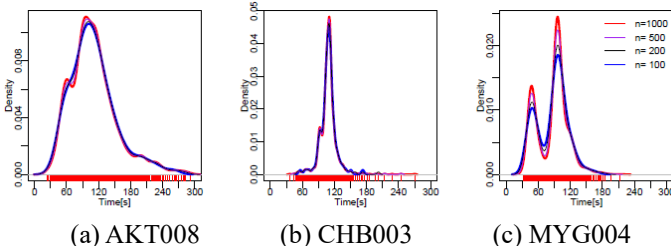


図2 カーネル密度推定による包絡線近似の比較

表2 バンド幅 h の比較

分割数 n	次元	バンド幅 h [s]		
		AKT008	CHB003	MYG004
100	99	12.84	2.91	10.09
200	199	11.39	2.57	8.91
500	499	9.50	2.17	7.50
1000	999	8.28	1.91	6.56

キーワード 地震動、経時特性、特徴ベクトル、カーネル密度推定、混合正規分布モデル

連絡先 〒501-1193 岐阜市柳戸1-1 岐阜大学流域圏科学研究センター TEL 058-293-2427

形状である図 2(c)では、次元が少ないと包絡線のピークを捉えていない傾向も見られるが、図 1(c)の加速度波形と比較すると、次元によらず、概ね近似できていると言える。図 2(a), (b)でも同様に近似できており、次元の違いは図 2(c)ほど顕著ではない。以上のように、次元数や波形による多少の違いは見られるが、カーネル密度推定により、地震動の包絡線を概ね近似していることが確認できた。

3. 混合正規分布モデルによる包絡線近似

3.1 混合正規分布モデル

2.で前述した特徴ベクトルに対して、混合正規分布モデルを適用し、包絡線近似を行う。混合正規分布とは、複数の確率分布を加重平均によって組み合わせたもので、単純な分布を混合して、より複雑な分布を記述するためのモデルである³⁾。

特徴ベクトル $\mathbf{t}$ について、平均 μ 、標準偏差 σ の正規分布の確率密度関数を、

$$\phi(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} \quad (5)$$

とするとき、 M 個の正規分布を要素モデルとする混合正規分布の確率密度関数は、

$$p(t; \theta) = \sum_{m=1}^M \pi_m \phi(t; \mu_m, \sigma_m) \quad (6)$$

で表される。ただし、 $\pi_m (m=1, \dots, M)$ は

$$\pi_m \geq 0, \sum_{m=1}^M \pi_m = 1 \quad (7)$$

を満たし、 θ は全てのパラメータを表したものである。

$$\theta = \{\pi_m, \mu_m, \sigma_m; m=1, \dots, M\} \quad (8)$$

混合数 M は、BIC(Bayesian Information Criterion)に基づいて決定する。BIC は最尤推定で得られたモデルのパラメータを $\hat{\theta}$ として、次式より算出する。

$$\text{BIC}(\hat{\theta}) = -2 \sum_{i=1}^n \ln p(t_i; \hat{\theta}) + k \ln n \quad (9)$$

ただし k はモデルの自由度、すなわちパラメータ数であり、本研究の場合は、式(8)より $k=3M-1$ となる。

3.2 混合正規分布モデルの適用と考察

特徴ベクトル $\mathbf{t}=\{t_i\}$ を用いて混合正規分布モデルを適用した。図 3 に 99 次元($n=100$)の場合の混合数と BIC を示す。式 (9) の定義では BIC 最小のモデルが最適モデルとなるが、図 3 では BIC 最大が最適となっている⁶⁾。波形・次元によって最適な混合数は異なるが、単峰形の AKT008, CHB003 は混合数 2、双峰形のやや複雑な MYG004 は混合数 3 が最適であった。それぞれの最適な混合数で近似した包絡線を図 4 に示す。AKT008,

CHB003 は 99 次元の特徴ベクトルによる包絡線とほぼ一致している。MYG004 ではピークの鋭さがかなり強調され、図 2(c) の $n=1000$ にやや近くなっている。いずれにしても、 M 組の (π, μ, σ) 、つまり $3M-1$ 次元に縮約された情報によって、包絡線を近似できている。

なお、 $n=200, 500, 1000$ で混合正規分布モデルをあてはめたところ、より高次のモデルが最適となる傾向が見られた。ただし、混合数を固定すれば近似結果は n に依存せず、ほとんど差が見られなかった。

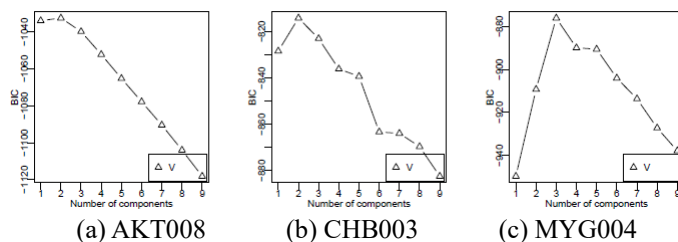


図 3 混合正規分布モデルの混合数 M と BIC

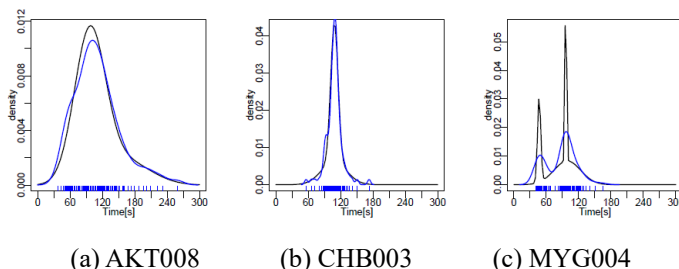


図 4 混合正規分布モデルによる包絡線近似
(青線は 99 次元、黒線は混合正規分布モデル)

4. おわりに

本研究では地震動の経時特性の特徴ベクトルを用いて、カーネル密度推定および混合正規分布モデルを用いた包絡線近似を行った。混合正規分布モデルによって特徴ベクトルが $M=2$ ないし 3 の正規分布で縮約表現可能であることが確認できた。非定常スペクトルについても同様の検討を進めており、震源過程との関連性の分析など、今後、手法の適用性の拡大を図る方針である。

謝辞: 本研究では(国研)防災科学技術研究所 K-NET の強震記録を使用しました。記して謝意を表します。

参考文献

- 1) 能島暢呂・久世益充・高島拓也：地震動の経時特性の特徴抽出と階層的クラスター分析による分類，日本地震工学会論文集 2017. (登載決定)
- 2) 久世益充・能島暢呂・高島拓也：地震動経時特性の特徴抽出と自己組織化マップによる評価，土木学会論文集 A1S, Vol.73, No.4, 2017. (登載決定)
- 3) 金森敬文・竹之内高志・村田昇：パターン認識，R で学ぶデータサイエンス 5，共立出版，2009.
- 4) R Documentation, <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/density.html>
- 5) (国研)防災科学技術研究所，強震観測網(K-NET, KiK-net), <http://www.kyoshin.bosai.go.jp/kyoshin/>
- 6) 金明哲：R によるデータサイエンス データ解析の基礎から最新手法まで，森北出版，2007.