SAPP
Statistical Analysis of Point Processes with R

統計数理研究所

Jun 1, 2010

1 はじめに

SASeis(Statistical Analysis of Seismicity) は，統計数理研究所で開発された地震活動などの統計的解析とモデリングのためのプログラムパッケージである．これは TIMSAC84-SASE Version 2, SASeis-W（Windows 版 SASeis）, IASPEI-SASeis-VB（SASeis Visual Basic）と SASeis2006 の総称である．中でも TIMSAC84-SASE Version 2 と SASeis2006 は Fortran で書かれたプログラムであり，特にその計算結果の出力ファイルはフリーソフト R を使ってグラフィックス表示が可能な形式になっている．一方 R はフリーな統計処理言語・環境であり，豊富なグラフィック関数と Fortran や C のサブルーチンを簡単に呼び出せるインタフェースも備えている．そこで Fortran で書かれた TIMSAC84-SASE Version 2 と SASeis2006 の計算処理機能をライブラリ化することにより，必要であればその解析結果を直接グラフィック 表示できるようにした関数群が R パッケージ SAPP である．

TIMSAC84-SASE Version 2 と SASeis2006 は，それぞれ Computer Science Monographs No.32[1] , No.33[2] に発表された．TIMSAC84-SASE Version 2 は，TIMSAC84 Part 2 (Computer Science Monographs No.23[3]) に発表された点過程の統計解析プログラムの GPSL によるグラフ出力部分を削除し単純化した改良版である．また，SASeis2006 では大森・宇津の公式と点過程 ETAS(Epidemic Type Aftershock Sequence) モデルを扱っている．ETAS モデルは大森・宇津の公式（余震の頻度関数）に地域性を持たせ拡張・発展させた地震活動の標準化モデルであり，世界各国で利用されている．この標準モデルによる予測と実際の地震活動との残差から，その地震活動の静穏化や活発化など変化の特徴が捉えられる．

なお，本研究の一部は，情報・システム研究機構 新領域融合研究センターの融合研究プロジェクト「機能と帰納：情報化時代にめざす科学的推論の形」の一環として実施されている．

[1] Y. Ogata, K. Katsura and J. Zhuang (2006) Statistical Analysis of Series of Events (TIMSAC84-SASE) Version 2, *Computer Science Monographs, No.32*, The Institute of Statistical Mathematics, Tokyo.

[2] Y. Ogata, (2006) Statistical Analysis of Seismicity - updated version (SASeis2006), *Computer Science Monographs, No.33*, The Institute of Statistical Mathematics, Tokyo.

[3] H. Akaike, T. Ozaki, M. Ishiguro, Y. Ogata, G. Kitagawa, Y. Tamura, E. Arahata, K. Katsura and R. Tamura, (1985) TIMSAC-84 Part 2,*Computer Science Monographs, No.23*, The Institute of Statistical Mathematics, Tokyo.

2 本パッケージの関数について

ここでは，発表された年次別にパッケージの関数とそこで仮定される公式やモデルについて簡単に説明する． 詳しくは，R Help を参照のこと．

2-1 TIMSAC84-SASE Version 2 を利用する R の関数

ptspec() ：
定常ポアソン過程の周期成分検索するにおける重要な帯域の点過程スペクトル解析を行う

linlin() ：
他の点過程入力, 周期的成分とトレンド成分をもつ自己励起型点過程の線形強度モデルの最尤推定値を求める

simbvh() ：
相互励起型二値の Hawkes 点過程のシミュレーション

linsim() ：
自己励起型点過程のシミュレーション

* eptren.f と pgraph.f は，SASeis2006 に含まれている改良版を利用した．

2-2 SASeis2006 を利用する R の関数

momori() ：
余震発生率の時間的減衰を表す大森・宇津の公式のパラメータの最尤推定値を求める

eptren() ：
強度関数が指数多項式または指数フーリエ級数である非定常ポアソン過程におけるパラメータの最尤推定値を求める

etasap() ：
ETAS モデルのパラメータの最尤推定値を求める

etasim() ：
ETAS モデルに基づいた地震データのシミュレーション

pgraph() ：
点過程データまたはその残差データの基本的な統計上の性質をグラフィック表示する

respoi() ：
momori() によって求められた最尤推定値をもつ大森・宇津の公式を使って地震の発生時刻を変換した残差を求める

etarpp()：

etasap() を使って求められた最尤推定値をもつ ETAS モデルを使って地震の発生時刻を変換した残差を求める

2-3 指数多項式と指数フーリエ級数 (eptren)

地震の条件付き強度関数のトレンド成分を指数多項式

$$f(t) = exp\{A_1 + A_2 t + A_3 t^2 + ...\}$$

で表わし，周期性成分を指数フーリエ級数

$$f(t) = exp\{A_1 + A_2 cos(2\pi t/P) + B_2 sin(2\pi t/P) + A_3 cos(4\pi t/P) + B_3 sin(4\pi t/P) + ...\}$$

で表わすとすれば，$A_1, A_2, ..., A_n$ および $A_1, A_2, B_2, ..., A_N, B_N$ は最尤法を用いて求められる．これら二つのモデルは非定常ポアソン過程に属し，最適な次数は最小 AIC(Akaike Information Criterion) によって決定される．非定常ポアソン過程であることが前提なので，余震や前震などの地震群や群発地震を含むデータに当てはめた場合には，結果の解釈に注意が必要である．

2-4 自己励起型点過程の線形強度モデル (linlin, linsim)

他の点過程入力, 周期的成分とトレンド成分をもつ自己励起型点過程の線形強度関数は

$$f(t) = P(t) + Q(t) + \Sigma g(t - t_i) + \Sigma h(t - u_j)$$

で表わされる．ただし，和 Σ はそれぞれ $t_i < t$ を満たすすべての i および $u_j < t$ を満たすすべての j についての合計である．ここで，周期的成分 $P(t)$ はフーリエ級数

$$P(t) = C(1)cos(2\pi t/P) + C(2)sin(2\pi t/P) + ...$$
$$+C(K_C - 1)cos(K_C \pi t/P) + C(K_C)sin(K_C \pi t/P),$$

トレンド成分 $Q(t)$ は多項式

$$Q(t) = T(1) + T(2)t + ... + T(K_T)t^{K_T - 1},$$

自己励起 (ある地域の時刻 t における地震活動が時刻 $t + s$ の地震活動に及ぼす影響) はラゲールの多項式

$$g(s) = exp(-Cs)\{A(1) + A(2)s + ... + A(K_A)s^{K_A - 1}\},$$

外部入力 (例えば他の地域の地震活動の影響など）はラゲールの多項式

$$h(s) = exp(-Ds)\{B(1) + B(2)s + ... + B(K_B)s^{K_B - 1}\}$$

で与えられる．最適な K_C, K_T, K_A, K_B は最小 AIC によって決定され，残りのパラメータ $C(1), ..., C(K_C), T(1), ..., T(K_T), A(1), ..., A(K_A), B(1), ..., B(K_B), C, D$ は最尤法により求められる．

2-5 大森・宇津の公式 (momori, respoi)

大森・宇津の公式は，単位時間あたりの余震頻度の減衰を表す．t を本震の発生からの経過時間とすれば，単位時間当たりの余震発生率 f(t) は

$$f(t) = \mu + K/(t+c)^p$$

に従って減衰する．ただし，μ, K, c, p は負でない定数で，μ は一定の割合で余震に含まれるかもしれない常時地震活動を表わす．これらのパラメータは，上式を条件付き強度関数と考え，余震発生を非定常ポアソン過程とみなし最尤法を用いて求められる．また，時刻 t_0 から t までに発生した余震の累積度数は，$f(t)$ の積分によって

$$F(t) = \mu(t-t_0) + K\{c^{1-p} - (t-t_i+c)^{1-p}\}/(p-1),$$

で与えられる．ただし，余震総数を N とすれば $i = 1, ..., N$ である．

2-6 ETAS モデル (etasap, etarpp)

ETAS モデルは，ある地域である期間に発生したマグニチュード M_z 以上の地震の時刻 t における発生率（点過程の強度関数）表す点過程モデルである．時刻 t における i 番めの地震 (時刻 t_i ¡ t, マグニチュード M_z) に伴う余震の発生率は，大森・宇津の公式を使って

$$n_i(t) = Kexp[\alpha(M_i - M_z)]/(t-t_i+c)^p$$

で与えられる．ここで，K, α, c, p はその地域に発生したすべての余震系列に共通の定数である．地震活動を一定の発生率 μ をもつ常時地震活動（ポアソン過程）と (別の地震の余震を含む) 各地震に伴う余震活動を足し合わせたものと考えれば，時間 t に起こった地震活動全体の発生率は

$$\lambda(t) = \mu + \Sigma_i n_i(t)$$

で表わされ，５つのパラメータ　μ, K, c, α, p によって地域性が特徴づけられる．ただし，和 Σ は $t_i < t$ を満たす全ての i についての合計である．