



Hitotsubashi ICS-FS Working Paper Series

FS-2017-J-001

Hawkes グラフによる倒産リスクの伝播構造の推定

監物 輝夫

一橋大学大学院 国際企業戦略研究科 修了生

中川 秀敏

一橋大学大学院国際企業戦略研究科

First version: 2017年5月24日

All the papers in this Discussion Paper Series are presented in the draft form. The papers are not intended to circulate to many and unspecified persons. For that reason any paper can not be reproduced or redistributed without the authors' written consent.

Hawkes グラフによる倒産リスクの伝播構造の推定*

監物 輝夫[†] 中川 秀敏[‡]

2017年5月24日

概要

Hawkes 過程を用いたイベント発生モデル化はファイナンス分野でも近年注目されている。本研究では、日本における比較的長期の倒産履歴データを対象に、業種および企業規模に基づいて倒産発生イベントを複数のイベントタイプに分類し、そのうえで、多次元 Hawkes 過程の強度過程に指数減衰型カーネル関数を仮定して最尤法で推定する従来からのモデルに加えて、Embrechts and Kirchner [3] が提唱した Hawkes グラフ・モデルを用いて分析を行い、カテゴリ間での倒産発生伝播構造を定量的に把握するだけでなく視覚化を試み、分析結果の比較・考察を通じて、Hawkes グラフを用いた信用リスク依存関係のモデル化が十分応用可能性があることを確認する。

1 はじめに

一般に、Hawkes 過程とは（単一あるいは複数のタイプの）イベントの累積発生件数をカウントする確率過程であり、イベント発生確率を特徴づける「強度過程」に自己励起性（あるイベント発生によって同タイプの強度がジャンプする）や相互励起性（あるイベント発生によって別のタイプの強度がジャンプする）といった性質を持っていることが特徴である。Hawkes 過程は、Hawkes [5] が集団間の伝染病の伝染メカニズムのモデル化などの応用を念頭において提唱したとされている。日本では、尾形が1980年代に Hawkes 過程に関連する理論的な研究（[7]）に貢献した一方で、地震（余震）発生メカニズムのモデル化への応用研究（[8]）でも貢献した実績が認められている。

ファイナンス関連という点では、Hawkes 過程を信用リスク計量に応用した先行研究の一つとして Errais et al. [4] が挙げられる。同論文では、信用ポートフォリオ内のデフォルトイベント伝播を対象にして、Hawkes 過程やその拡張として拡散項を含むジャンプ・アフィン強度過程に基づくモデルに対する実証分析を行っている。また、中川 [10] や Yamanaka et al. [9] などでは、主に日本企業の R&I の格付変更アナウンスを対象イベントとして Hawkes 過程による格付変更モデルの分析を行っている。さらに、最近では株式や先物市場での取引や価格変化をイベントとして、高頻度取引データの適用を意図した Hawkes 過程の取引モデルへの応用が増えてきている（Bacry et al. [2] を参照のこと）。

このように、Hawkes 過程によるイベント発生モデル化はファイナンス分野でも近年注目されている中で、Embrechts and Kirchner [3] は、イベントタイプが多数ある場合を念頭にして Hawkes 過程がもつ自己励起性や相互励起性の様子を、Hawkes グラフと名付けられた「重み付き有向グラフ」を用いて視覚化するアイデアを示し、同時に高次元でのモデル・パラメータ推定における計算負荷が軽減されるような推定アルゴリズムを提唱している。特に、従来はイベント発生強度をモデル化する際にパラ

* 本研究は JSPS 科研費 JP26330026 の助成を受けて行われたものです。

[†] 一橋大学大学院 国際企業戦略研究科 修士

[‡] 一橋大学大学院 国際企業戦略研究科

メトリックな関数を具体的に与えてパラメータの最尤推定を行うことが多かったが、同論文が提案する Hawkes グラフに対しては INAR モデル^{*1}と呼ばれる整数値の AR モデルを取り入れており、イベント発生強度をノンパラメトリックに推定しているのが特徴である。

本研究では、日本における比較的長期の倒産履歴データを対象に、倒産発生イベントを業種および企業規模に基づき複数カテゴリに分類することで複数のイベントタイプを作り、その上で Embrechts and Kirchner [3] が提案している Hawkes グラフの推定手法を適用し、業種あるいは企業規模に関するカテゴリ間での倒産発生の伝播構造を視覚化することを試みる。同論文が想定するほどの多数のイベントタイプや大きなサイズのデータを扱うわけではないが、Hawkes グラフの推定アルゴリズムを実際のデータに対して適用する際は、推定に必要なチューニング・パラメータの設定方法に注意が必要となることを確認する。また、日本企業の倒産履歴データを対象に、Hawkes [5] でも例示されている指数減衰型カーネル関数をもつ強度過程で特徴付けられる Hawkes 過程モデルを最尤推定した結果と、論文 [3] による Hawkes グラフの推定結果を比較・考察し、Hawkes グラフを用いた信用リスク依存関係のモデル化が十分応用可能性があることを確認する。

2 分析データについて

本研究における事例研究の対象データは、(株)東京商工リサーチが集計した 2003 年 1 月 1 日から 2015 年 12 月 31 日まで (4,748 日分) の間に発生した倒産事例データである。ただし、倒産時負債総額 10 億円以上という条件でスクリーニングを依頼してデータを入手している。

次に、以下のような条件に沿って元データを整理した。

- 「関東 (東京・神奈川・千葉・埼玉)」「関西 (大阪・京都・兵庫・奈良)」「その他 (の地域)」に区分する
- 同じ倒産発生日の関連企業 (= 代表者が同じ企業) の倒産事例は、負債総額などを一つのサンプルとしてまとめる^{*2}
- 「資本金」「従業員数」「負債総額」「倒産原因」「倒産形態」がエラー値のものを除外する^{*3}

さらに、このデータに含まれる倒産企業を、以下のように業種および企業規模に関してカテゴリ区分をする。

まず、業種に関しては、東京商工リサーチが与えた業種区分に基づいて、倒産企業を「製造関係」「インフラ・流通関係」「金融関係」「サービス関係」の 4 つの業種カテゴリに区分している^{*4}。

また、企業規模に関しては、最初に「中小企業基本法第二条第五項」における小規模企業者の定義に当てはまる倒産企業群を「小企業」として区分し、次に「中小企業基本法第二条第一項」に当てはまる企業群から先に区分した小企業群を除いた企業群を「中企業」として区分し、これらの定義に該当しない残るその他の企業を「大企業」と区分して、3 つの企業規模カテゴリを作っている^{*5}。

^{*1} 「INteger-valued AutoRegression モデル」の略。

^{*2} 例えば、2008 年 9 月 16 日に倒産発生と記録された「リーマン・ブラザーズ証券」「リーマン・ブラザーズ・ホールディングス」「リーマン・ブラザーズ・コマーシャル・モーゲージ」の場合や、2010 年 1 月 19 日に倒産発生と記録された「日本航空」と「日本航空インターナショナル」の場合などは 1 件の倒産事例としてまとめた。

^{*3} 倒産時負債総額 10 億円以上という条件でスクリーニングを依頼したが「負債総額」が欠損しているものも含まれていた。

^{*4} 「製造」には農林業・漁業、「インフラ・流通」には情報通信業、運輸業、卸売業、小売業、「金融」には不動産業、といった業種が含まれている。なお、元データの段階では、業種カテゴリ別の倒産発生件数は、製造関係が 1,798 件、インフラ・流通関係が 2,098 件、金融関係が 1,366 件、サービス関係が 1,866 件であった。

^{*5} なお、元データの段階では、その結果、企業規模カテゴリ別の倒産発生件数は、小企業が 2,633 件、中企業が 3,431 件、

表1 「関東」「関西」「その他」の各地域区分ごとの業種カテゴリおよび企業規模カテゴリにおける倒産発生件数

	関東				関西				その他			
	小企業	中企業	大企業	(計)	小	中	大	(計)	小	中	大	(計)
製造	149	471	102	722	113	363	37	513	350	1104	83	1537
インフラ・流通	214	293	208	715	98	184	96	378	245	352	202	799
金融	323	83	65	471	152	28	18	198	282	55	18	355
サービス	136	93	177	406	62	58	81	201	354	276	326	956
(計)	822	940	552	2314	425	633	232	1290	1231	1787	629	3647

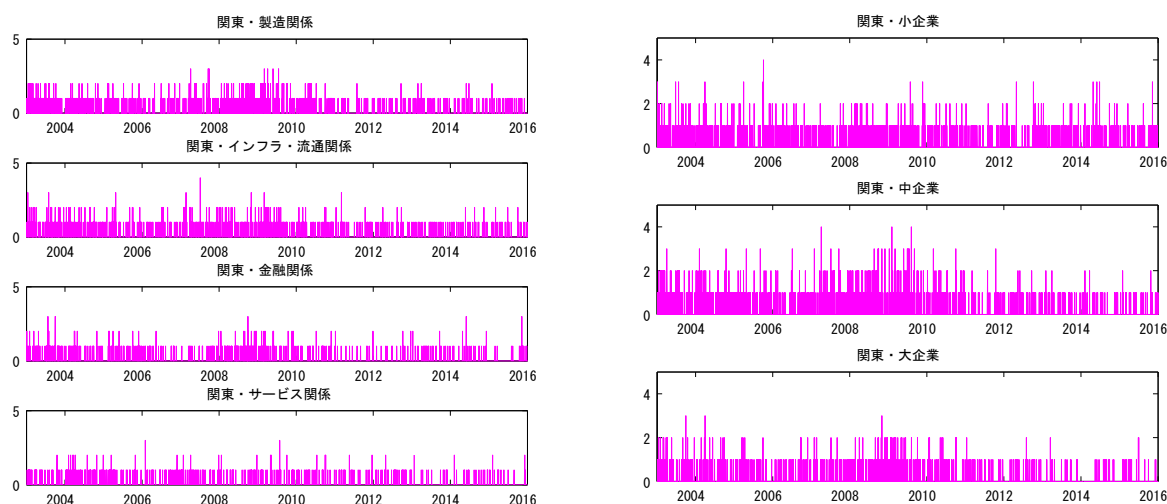


図1 (株)東京商工リサーチの倒産履歴データを整理した後の、カレンダーベースでの関東地区の日次での倒産発生件数の推移。左は業種別カテゴリ、右は企業規模別カテゴリによるもの。横軸は2003年1月1日を起点にカレンダーベースでの経過日数(計4748日分)

以上の条件で元データを整理し、「関東」「関西」「その他」という各地域区分ごとに業種カテゴリおよび企業規模カテゴリで倒産発生件数のクロス集計を行った結果を表1で表している。

本稿では、事例研究の端緒として、地域別の中ではサンプル数が多い関東地区のデータに限定して倒産リスクの依存構造の分析を行うことにする。図1は、(株)東京商工リサーチの元データを整理した後の、カレンダーベースでの関東地区の日次での倒産発生件数の推移(左は業種別カテゴリ、右は企業規模別カテゴリ)を表したものである。また、付録A節には、関東地区の倒産企業のうち負債総額が上位20の企業一覧をまとめておいた。

これを概観する限りでは、業種や企業規模によっては倒産が集中している時期が確認でき、倒産リスクの自己励起性が存在するかどうかを調査する意義があると考え。また、異なる業種カテゴリ間あるいは異なる企業規模カテゴリ間での倒産リスクの伝播の存在については明確に断定はできないが、分析を通じて調査する意義は十分にあると考える。

なお、連続時間モデルによる分析を行うことになるため、倒産発生時点を連続時間上にマッピングする必要がある。ただし、倒産発生は発生年月日しか特定できないことと、同じ日に複数の倒産発生が観測されることがある(実際にあった)ことに注意しなければならない。

本研究では営業日ベースで倒産発生が起こっているという想定のもとで、1年間を一律1/365単位とする時間幅に分割して、データ起点の2003年1月1日を $t = 0$ 、終点の2015年12月31日が

大企業が1,445件であった。

$t = T := 13.00548$ となるように定めた。その上で、倒産発生日を起点日からの経過日数 k (整数) に対して $k/365$ として連続時間にマッピングした。また、複数の倒産が (同じカテゴリ内で) 発生した場合は同時点での発生件数を「マーク変数」として扱うことにした。

3 一般的な多次元 Hawkes 過程とその典型例：指数減衰型強度モデル

当節では、まず一般的な多次元 Hawkes 過程の定義を確認し、次に複数タイプの倒産発生を特に Hawkes [5] のオリジナル論文でも例として挙げられている「指数減衰型」の性質をもちあわせたパラメトリックな強度過程を用いたモデルの定式化を紹介する。そして、この強度モデルのパラメータを最尤推定する方法を中川 [10] を参考にしてまとめる。

3.1 多次元 Hawkes 過程

適当なフィルター付き完備確率空間 $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ を与える。 $m \in \mathbb{N}$ 種類のイベントを対象にする。ここで 1 から自然数 n までの集合を $[n] = \{1, 2, \dots, n\}$ のように表すことにする。

各イベント $j \in [m]$ に対して、 \mathbb{F} -適当な点過程 (停止時刻の増大列) $\{\tau_k^j\}_{k \in \mathbb{N}}$ を考える。すなわち、 τ_k^j をタイプ j のイベントが k 回目に発生する \mathbb{F} -停止時刻と見なし、 \mathbf{P} -a.s. に $0 (= \tau_0^j) < \tau_1^j < \tau_2^j < \dots$ が成り立つとする。

さらに、各イベント $j \in [m]$ に対して、点過程 $\{\tau_k^j\}_{k \in \mathbb{N}}$ に付随する計数過程を $N_t^j := \sum_{k \geq 1} \mathbf{1}_{\{\tau_k^j \leq t\}}$ と定義し、まとめて $\mathbf{N}_t = (N_t^1, \dots, N_t^m)^\top$ と表す。

このとき、一般に、多次元 Hawkes 過程は次のような強度過程で特徴づけられる計数過程として定義される (Bacry et al. [2] の Definition 1 などを参照)。

定義 (多次元 Hawkes 過程). m -次元 計数過程のベクトル \mathbf{N}_t が m -次元 Hawkes 過程であるとは、各 $j \in [m]$ に対して N_t^j に付随する強度過程^{*6}が

$$\lambda_t^j = \mu^j + \sum_{i=1}^m \int_{-\infty}^t h_{i \rightarrow j}(t-s) dN_s^i \quad (1)$$

で与えられるものをいう。ただし、 $\mu^j > 0$ は外生的な強度パラメータ、被積分関数 $h_{i \rightarrow j}(u)$ は任意の $(i, j) \in [m]^2$ に対して非負値の確定的な可積分関数で、 $u < 0$ のときは $h_{i \rightarrow j}(u) \equiv 0$ を満たすものとする。

(1) 式の被積分関数 $h_{i \rightarrow j}(u)$ は「カーネル関数」とも呼ばれ、 $h_{i \rightarrow j}(t-s)$ は具体的には「時点 s で発生したタイプ i のイベントが、タイプ j のイベント発生強度に与えたインパクトが時点 t では (すなわち $t-s$ だけ経過後に) どの程度残っているか」を表していると解釈できる。したがって、(1) から分かるように、Hawkes 過程は強度過程のカーネル関数の形をどのように与えるかによって特徴づけられる。

5 節においては、倒産発生の依存関係をモデル化する際に Hawkes 過程の特徴づけとして次の 2 つのアプローチを採用し、2 節で述べた日本の関東地区の倒産履歴データに対する両アプローチの分析結果 (推定された依存関係および事後的なイベント発生強度の推移) を比較・考察する。

1. $(i, j) \in [m]^2$ に対して、指数減衰型カーネル関数 $h_{i \rightarrow j}(u) = \xi^{i \rightarrow j} e^{-\kappa^j u}$ (ただし、 $\xi^{i \rightarrow j} \geq 0, \kappa^j > 0$)

^{*6} 非負値 (発展的可測) 過程 λ_t^j が計数過程 N_t^j の強度過程であるとは、 $N_t^j - \int_0^t \lambda_s^j ds$ が \mathbb{F} - (局所) マルチンゲールになることと定義される。

とする)による強度過程をもつ Hawkes 過程を仮定し、「倒産時点」をデータとしてパラメータを最尤推定する。

2. カーネル関数の形は特定しない。「単位期間ごとの倒産件数」をデータとし、次の単位期間における倒産件数を整数値の AR モデル(「INAR モデル」と呼ばれる)で近似的に表現できると仮定して、単位期間を区切る「離散時点上のカーネル関数」のノンパラメトリックな推定値を線形代数の演算によって計算する。

次小節では、中川 [10] を参考にして、前者の指数減衰型カーネル強度過程で特徴づけられる倒産イベント発生依存関係モデルを示して、パラメータの最尤推定の方法についてまとめておく。

なお、倒産発生強度モデルとしては、上記の Hawkes 過程では不十分であるという議論が当然ある。なぜなら、Hawkes 過程モデルにおいては、あるカテゴリ内の倒産発生リスクを上昇させるのは、実際に倒産がどこかのカテゴリで発生したことによる自己励起性あるいは相互励起性だけであるということ仮定することになるからである。

例えば、Azizpour et al. [1] によると、米国の社債市場のデフォルト・クラスタリングの要因説明のための強度モデルとして、以下のような強度モデルを用いて実証分析を行っている。Azizpour et al. [1] では(業種などでのカテゴリ分けによる分析を行っているわけではなく)負債を Moody's 社に格付されている企業群に対する倒産発生強度 λ_t として、以下のような定式化のものを考えている:

$$\lambda_t = \exp\left(a_0 + \sum_{i=1}^d a_i X_{i,t}\right) + Y_t + Z_t.$$

ただし、 $(X_{1,t}, \dots, X_{d,t})$ は、 d 種類の観測可能な(マクロの)説明変数群であり、 Y_t は伝播(contagion)効果を現す項でマーク付きで指数減衰カーネル関数をもつ Hawkes 過程として

$$Y_t = b \sum_{n:T_n \leq t} e^{-\kappa(t-T_n)} \max\{0, \log u_n\},$$

と表されるもの(T_n, u_n はそれぞれ n 番目のデフォルト発生時とその時点のデフォルト負債合計額面) また Z_t は潜在変数(frailty)項で、平均回帰性をもつ金利の CIR モデルと同じように

$$dZ_t = k(z - Z_t)dt + \sigma\sqrt{Z_t}dW_t,$$

という確率微分方程式を満たすもの、と仮定している。

Azizpour et al. [1] の結論を要約すると、米国の社債市場のデフォルト・クラスタリングの要因説明には、観測可能な説明変数と frailty 項だけでは不十分であるが、contagion 項まで加えると十分に説明可能である、ということになる。

これに対し、本研究では複数カテゴリを対象にしたモデル化に取り組むという点では Azizpour et al. [1] より進展はあると言えるが、採用した強度モデルに関しては Azizpour et al. [1] のモデルにおける観測可能な説明変数による項および frailty 項は含めずに、contagion 項だけを(1)のようにより素朴な形で用いることになる。これは、Hawkes グラフを提唱する Embrechts and Kirchner [3] の強度モデルに整合させることで、推定法をそのまま踏襲できるようにすることを選択したためである。

当然、倒産発生リスクをとらえるモデルを作ることを目的とするのであれば、少なくとも観測可能なマクロ経済説明変数を含む強度モデルを用いた分析が必要となることは間違いない。本来マクロ経済変数や潜在変数の動きによって捕捉されるはずの倒産リスクの伝播性が、(1)式の Hawkes 過程の強度モデルの初期強度 μ^j やカーネル関数 $h_{i \rightarrow j}(u)$ の推定値として無理な形で表出される可能性は大きい。

しかし、今回の研究では、Hawkes グラフという手法によって、倒産リスクの依存関係がどのように計量化・視覚化されるかを把握すること主眼とする。そのためには、(1) 式で与えられる強度モデルに基づいて分析しても十分に意義があると考え、Embrechts and Kirchner [3] が提唱した推定手法を修正・改良する必要が生じるような強度モデルの拡張には着手しない。その意味では、後述する分析結果およびその考察は、あくまでもマクロ経済変数や潜在変数の動きを明示的に織り込まない場合に可視化されるリスクの伝播構造の様子であることを注意する必要がある。

3.2 指数減衰型カーネル強度過程による倒産イベント発生モデル

前節をふまえると、本研究で用いられるデータは、各カテゴリ $j \in [m]$ について、 $t = 0$ 時点から数えて ℓ 回目の倒産発生時点 $\tilde{\tau}_\ell^j$ とその時の同時発生件数 $\tilde{\eta}_\ell^j$ の組の集合 $(\tilde{\tau}^j, \tilde{\eta}^j) := \{(\tilde{\tau}_\ell^j, \tilde{\eta}_\ell^j)\}_{\ell=1, \dots, \tilde{N}_T^j}$ という形式で与えられることに注意する。

同時点での倒産発生件数 $\{\eta_k^j\}_{k=1, \dots, N_t^j}$ をマーク変数とするマーク付き (\mathcal{F}_t) -適な純ジャンプ過程 L_t^j を $L_t^j := \sum_{k=1}^{N_t^j} \eta_k^j$ とすることで、カテゴリ j の実際の累積倒産発生件数が表される。カテゴリ j の倒産強度過程を λ_t^j と表すと、指数減衰型カーネル関数 $h_{i \rightarrow j}(u) = \xi^{i \rightarrow j} e^{-\kappa^j u}$ ($\xi^{i \rightarrow j} \geq 0, \kappa^j > 0$) を用いることで

$$\lambda_t^j = \mu^j + \sum_{i=1}^m \xi^{i \rightarrow j} \int_0^t e^{-\kappa^j(t-s)} dL_s^i \quad (2)$$

と表される。

この場合、カテゴリごとに(ただし、全サンプルを用いて)強度過程に含まれるパラメータを最尤推定することができる。具体的には、カテゴリ j に対して、パラメータ集合を $\theta^j := (\mu^j, \kappa^j, \{\xi^{i \rightarrow j}\}_{i \in [m]})$ とすると、対数尤度関数 $\log \mathcal{L}(\theta^j | (\tilde{\tau}^j, \tilde{\eta}^j))$ の本質的な部分は、

$$\int_0^T \log(\tilde{\lambda}_{s-}^{j, \theta^j}) d\tilde{L}_s^j - \int_0^T \tilde{\lambda}_s^{j, \theta^j} ds$$

と表されることが知られている。より具体的には以下の目的関数を最大化するパラメータ集合 θ^j の値を数値的に求めることになる。

$$\sum_{k=1}^{\tilde{N}_T^j} \log \left\{ \mu^j + \sum_{i=1}^m \xi^{i \rightarrow j} \sum_{\tilde{\tau}_\ell^i < \tilde{\tau}_k^j} \tilde{\eta}_\ell^i e^{-\kappa^j(\tilde{\tau}_k^j - \tilde{\tau}_\ell^i)} \right\} - \mu^j T - \frac{1}{\kappa^j} \sum_{i=1}^m \xi^{i \rightarrow j} \sum_{k=1}^{\tilde{N}_T^i} \tilde{\eta}_k^i (1 - e^{-\kappa^j(T - \tilde{\tau}_k^i)}) \quad (3)$$

4 Hawkes グラフ

当節では、Embrechts and Kirchner [3] によって導入された Hawkes スケルトンと Hawkes グラフを紹介する*7。

Hawkes グラフとは、カテゴリの集合を頂点集合とし、イベント発生による自己励起性や相互励起性が認められる関係を辺集合とする有向グラフで、さらに強度過程を特徴付けるカーネル関数の推定値に基づいて得られる自己励起性や相互励起性の大きさを「辺の重み」の情報を各辺に付加したものである。Hawkes スケルトンとは、Hawkes グラフを求める前段階で得ることになる「辺の重み」を付加しない有向グラフのことを指す。

*7 ここでは、特に倒産イベントに限定する必要がないので一般的なイベントに対する話題として説明しておく。

さらに、Hawkes スケルトンを推定する際には、Hawkes 推定量と名付けられる「(分析上定める) 単位期間ごとのカーネル関数のノンパラメトリックな推定値」からなる行列を、イベント発生データから求める必要がある。この Hawkes 推定量は、最終的に得られる Hawkes グラフに整合的な強度過程の事後的なパスを描く際にも用いられ、この事後的なパスと実際の倒産発生状況の推移と最尤推定された指数減衰型カーネル強度モデルの事後パスとの比較を通じて、Hawkes グラフの推定アルゴリズムの有効性を議論することが可能である。

4.1 Hawkes 推定量

Hawkes グラフの構築に際しては、式 (1) にみられる強度過程のカーネル関数 $h_{i \rightarrow j}(u)$ に特定の関数形を与えず、Kirchner [6] で詳しく紹介されている INAR(p) モデルという p -次のラグをもつ整数値 AR モデルを適用して求める。

まず (十分に小さい) ある $\Delta > 0$ を分析上定める単位期間として固定する。 $k \in [n]$ (ただし $n := \lfloor T/\Delta \rfloor$) に対して、時間区間 $((k-1)\Delta, k\Delta]$ 内で観測されるタイプ j のイベント数を $X_k^{(j,\Delta)}$ と表すことにする*8。こうして $\mathbf{X}_k^{(\Delta)} = (X_k^{(1,\Delta)}, X_k^{(2,\Delta)}, \dots, X_k^{(m,\Delta)})^\top$ を時間区間 $((k-1)\Delta, k\Delta]$ 内のタイプ別イベント発生件数の m -次元ベクトルとして定義する。

$p \in \mathbb{N}$ を固定し、パラメータである m -次元ベクトル $\alpha_0 := (\alpha_0^1, \dots, \alpha_0^m)^\top$ と行列 $(\alpha_k^{i,j})_{1 \leq i,j \leq m} \in \mathbb{R}_{\geq 0}^{m \times m}$ を用いると、INAR(p) モデルは以下の確率差分方程式系で定式化される：

$$u \in \{p+1, \dots, n\} \quad \mathbf{X}_u^{(\Delta)} = \sum_{k=1}^p \begin{pmatrix} \sum_{j=1}^m X_{u-k}^{(1,\Delta)} \sum_{\ell=1}^m \xi_\ell^{(\alpha_k^{1,j})} \\ \vdots \\ \sum_{j=1}^m X_{u-k}^{(m,\Delta)} \sum_{\ell=1}^m \xi_\ell^{(\alpha_k^{m,j})} \end{pmatrix} + \varepsilon_u.$$

ただし、 $\xi_k^{(\alpha)}$ は平均 α のポアソン分布 $Po(\alpha)$ に従う確率変数を表すものとする。また、 ε_u は、各成分が i.i.d. で、各 $j \in [m]$ に対して $\varepsilon_u^j \sim Po(\alpha_0^j)$ となる m -次元確率変数の列ベクトルとする。

なお、パラメータ α_0 および $(\alpha_k^{i,j})_{(i,j) \in [m]^2, k \in [p]}$ は「条件付き最小二乗法 (CLS)」を適用して算出される*9。

INAR(p) モデルのパラメータと、強度過程の式 (1) に含まれる定数強度 μ^j およびカーネル関数 $h_{i \rightarrow j}(t)$ の間には、 $\alpha_0^j \approx \mu^j \Delta$ ($j \in [m]$) および $\alpha_k^{i,j} \approx h_{i \rightarrow j}(k\Delta) \Delta$ ($(i,j) \in [m]^2, k \in [n]$) という近似関係が成り立つことに注意する。

Kirchner [6] では、 p -期間分の強度過程 (1) に関するノンパラメトリック推定量の行列 $\hat{H}^{(\Delta,p)} := \left(\hat{H}_1^\top \dots \hat{H}_p^\top \hat{\mu} \right)^\top \in \mathbb{R}^{(mp+1) \times m}$ を Hawkes 推定量と呼んでいる。ただし $k \in [p]$ に対して $\hat{H}_k := \left(\hat{h}_{i \rightarrow j}(k\Delta) \right)_{(i,j) \in [m]^2} \in \mathbb{R}^{m \times m}$ とする。

*8 このようにある区間内でのイベント発生件数をデータとして用いることになるので、倒産データのように同時点に複数のイベント発生がある場合でも、データの技術的な扱いという点では、指数減衰型強度過程のときのような特別な処理をする必要はなくなる。

*9 これらのパラメータ推定には、最尤法や Yule-Walker (YW) 法によるモーメント法的な推定法も考えられる。Kirchner (2016) では、パラメータ数が多くなるとナイーブな最尤法は難しく、YW 型は単位根に近い場合には推定バイアスが大きくなる、といったことがコメントされている。

論文 [6] では、Hawkes 推定量 $\hat{H}^{(\Delta,p)}$ が次式で与えられることを示している。Hawkes 推定量が (4) 式で得られる背景のアイデアについては 付録 B 節にまとめている。

$$\hat{H}^{(\Delta,p)} = \frac{1}{\Delta} (\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{Y}. \quad (4)$$

ただし、

$$\mathbf{Y} := (\mathbf{X}_{p+1}^{(\Delta)}, \mathbf{X}_{p+2}^{(\Delta)}, \dots, \mathbf{X}_n^{(\Delta)})^\top \in \mathbb{N}^{(n-p) \times m}, \quad (5)$$

$$\mathbf{Z} := \begin{pmatrix} (\mathbf{X}_p^{(\Delta)})^\top & (\mathbf{X}_{p-1}^{(\Delta)})^\top & \dots & (\mathbf{X}_1^{(\Delta)})^\top & 1 \\ (\mathbf{X}_{p+1}^{(\Delta)})^\top & (\mathbf{X}_p^{(\Delta)})^\top & \dots & (\mathbf{X}_2^{(\Delta)})^\top & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (\mathbf{X}_{n-1}^{(\Delta)})^\top & (\mathbf{X}_{n-2}^{(\Delta)})^\top & \dots & (\mathbf{X}_{n-p}^{(\Delta)})^\top & 1 \end{pmatrix} \in \mathbb{N}^{(n-p) \times (mp+1)}. \quad (6)$$

4.2 Hawkes スケルトン

以下、ある $\Delta_{skel} > 0$ を固定し、 $\hat{H}^{(\Delta_{skel},p)}$ を Hawkes 推定量の式 (4) において、 $\Delta = \Delta_{skel}$ とおいたものとする。

Hawkes グラフを作成する際に最も重要なことは、最終的なグラフの「辺の重み」に相当する $a_{i,j} := \int_0^\infty h_{i \rightarrow j}(t) dt$ の推定量 $\hat{a}_{i,j}$ およびその標準誤差の 2 乗に相当する $\hat{\sigma}_{i,j}^2$ を求めることである。理論的には行列 $(\hat{a}_{i,j})_{(i,j) \in [m]^2}$ および $(\hat{\sigma}_{i,j}^2)_{(i,j) \in [m]^2}$ は、Hawkes 推定量 $\hat{H}^{(\Delta_{skel},p)}$ および Hawkes 推定量の分散共分散行列 \widehat{S}^2 を用いて、それぞれ

$$(\hat{a}_{i,j})_{(i,j) \in [m]^2} = \Delta_{skel} \mathbf{B} \hat{H}^{(\Delta_{skel},p)}, \quad (\hat{\sigma}_{i,j}^2)_{(i,j) \in [m]^2} = \Delta_{skel}^2 \mathbf{E}_{(i-1)m+j}^\top \widehat{S}^2 \mathbf{E}_{(i-1)m+j} \quad (7)$$

で与えられることが分かっている。ただし、 $\mathbf{B} := (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m)^\top$ とし、各 $j \in [m]$ に対して、 \mathbf{b}_j は $((k-1)m+j)$ -成分 ($k \in [p]$) だけ 1 で、残りは 0 となる $(mp+1)$ -次元ベクトル、 $\mathbf{E}_{(i-1)m+j}$ は $((k-1)m^2+(i-1)m+j)$ -成分 ($k \in [p]$) だけ 1 で、残りは 0 となる (m^2p+m) -次元ベクトル。

これらを元に、まずグラフの骨組みにあたる Hawkes スケルトン が以下で定義される。

定義 (Hawkes スケルトン). (1) で与えられる強度過程をもつ m -次元 Hawkes 過程 N_t に対する Hawkes スケルトンとは、有向グラフ $\mathcal{G}_S := (\mathcal{V}_S, \mathcal{E}_S)$ で、頂点の集合が $\mathcal{V}_S = [m]$ 、辺の集合が $\mathcal{E}_S = \{(i,j) \in [m]^2 \mid a_{i,j} > 0\}$ で与えられたものをいう。

ただし、実際の Hawkes スケルトンの辺集合は $\widehat{\mathcal{E}}_S = \{(i,j) \in [m]^2 \mid \hat{a}_{i,j} > \hat{\sigma}_{i,j} z_{1-\alpha_{skel}}^{-1}\}$ と推定される。ただし、 $z_{1-\alpha_{skel}}^{-1}$ は、信頼区間の水準を定めるチューニング・パラメータ $\alpha_{skel} \in (0, 0.5)$ に対して、標準正規分布の $(1 - \alpha_{skel}) \times 100\%$ 点分位点の値とする。したがって、信頼区間 $(\hat{a}_{i,j} - \hat{\sigma}_{i,j} z_{1-\alpha_{skel}}^{-1}, \hat{a}_{i,j} + \hat{\sigma}_{i,j} z_{1-\alpha_{skel}}^{-1})$ は、 $(1 - 2\alpha_{skel}) \times 100\%$ 両側信頼区間を表すことになる*¹⁰。

論文 [3] では、Hawkes スケルトンを求める際に最も負荷が高いのが Hawkes 推定量の分散共分散行列 \widehat{S}^2 の推定と説明されている。 \widehat{S}^2 の計算式を具体的に与えることもできるが、同論文では $(\hat{\sigma}_{i,j}^2)_{(i,j) \in [m]^2}$ を行列演算で算出できるアルゴリズムを示している (付録 C 節を参照のこと)。

*¹⁰ つまり $\alpha_{skel} = 0.05$ であれば、 $(\hat{a}_{i,j} - \hat{\sigma}_{i,j} z_{0.95}^{-1}, \hat{a}_{i,j} + \hat{\sigma}_{i,j} z_{0.95}^{-1})$ は 90% 両側信頼区間となる。

4.3 Hawkes グラフ

前小節で導入した Hawkes スケルトン $\mathcal{G}_S = (\mathcal{V}_S, \mathcal{E}_S)$ の頂点と辺それぞれにイベント発生強度のカーネル関数から派生する「重み」情報を付加したものが Hawkes グラフである。

定義 (Hawkes グラフ). *Hawkes* グラフとは、以下で与えられる重み付き有向グラフ $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ のことである。

$$\begin{aligned} \mathcal{V} &:= \{(j; \mu^j) \mid j \in \mathcal{V}_S, \mu^j \text{ は外生的な定数強度}\}, \\ \mathcal{E} &:= \left\{ (i, j; a_{i,j}) \mid (i, j) \in \mathcal{E}_S, a_{i,j} = \int_0^\infty h_{i \rightarrow j}(t) dt \right\}. \end{aligned}$$

しかしながら、実際にデータから Hawkes グラフを推定する際には、前小節のアルゴリズムを通して推定された Hawkes スケルトン $\widehat{\mathcal{G}}_S = (\widehat{\mathcal{V}}_S, \widehat{\mathcal{E}}_S)$ にただ「重み」の情報を加えればよいというわけではない。実は Hawkes スケルトンを推定する段階ではイベントタイプ間の依存関係の判定を少し甘めにしておいて、ある程度関係性を絞ってから、より本格的な依存関係の推定をし直すことで、最終的な Hawkes グラフを推定するという 2 段階の推定を行うことが Embrechts and Kirchner [3] の提唱するアルゴリズムとなる。

そのため、まず $j \in [m]$ に対して「親集合」 $PA(j) := \{i \mid (i, j) \in \mathcal{E}_S\}$ を定義し、要素数を $m_j := |PA(j)|$ とする。また、 $\Delta_{graph} > 0$ を $\Delta_{graph} \ll \Delta_{skel}$ となるようにとり固定する。したがって、以下では取り直した Δ_{graph} に対して再度 Hawkes 推定量を計算することになり、記号上は特に区別しないが、 n といった推定期間の分割数なども更新されたものになることに留意されたい。

その上で、推定 Hawkes グラフ $\widehat{\mathcal{G}} = (\widehat{\mathcal{V}}, \widehat{\mathcal{E}})$ を $\widehat{\mathcal{V}} := \{(j; \hat{\mu}^j) \mid j \in [m]\}$ ($\hat{\mu}^j$ は適当に推定された外生的な定数強度) および

$$\widehat{\mathcal{E}} := \bigcup_{j \in [m]} \left\{ (i_\ell, j; \hat{a}_{i_\ell, j}) \mid \{i_1, i_2, \dots, i_{m_j}\} = \widehat{PA}(j), \hat{a}_{i_\ell, j} = \mathbf{b}_{\ell, j}^\top \widehat{\mathbf{H}}_j^{(\Delta_{graph}, p)} \right\}$$

として与えることにする。ただし、 $\mathbf{b}_{\ell, j}$ は $((k-1)m_j + \ell)$ -成分 ($k \in [p]$) だけ 1 で、残りは 0 となる $(m_j p + 1)$ -次元ベクトルである。

また、Hawkes グラフで与える「辺の重み」は、信頼区間の水準を決めるチューニング・パラメータ $\alpha_{graph} \in (0, 0.5)$ に対して、 $[\hat{a}_{i_\ell, j} \pm \hat{\sigma}_{i_\ell, j} z_{1-\alpha_{graph}}^{-1}]$ という信頼区間の形で与えられ、「頂点の重み」となる定数強度 μ^j に対する信頼区間は、 $[\hat{\mu}^j \pm \hat{\sigma}_j z_{1-\alpha_{graph}}^{-1}]$ という形で与えられる。ここでも「標準誤差の 2 乗」 $\hat{\sigma}_{i_\ell, j}^2$ の算出がポイントとなり、加えて Hawkes グラフでは、「頂点の重み」に対する「標準誤差の 2 乗」 $\hat{\sigma}_j^2$ についても算出することとなる。いずれについても、論文 [3] では以下のアルゴリズムで算出することが提案されている。

4.4 チューニング・パラメータの決め方

前節で述べた Hawkes グラフによる分析を行う前に決めなくてはならないチューニング・パラメータは $(p, \Delta_{skel}, \Delta_{graph}, \alpha_{skel}, \alpha_{graph})$ である。

今回の分析におけるチューニング・パラメータの基本的設定は以下のように与える。ただし、分析において、チューニング・パラメータの違いによって Hawkes グラフの推定結果がどのように変わるかに

についても簡単な比較分析も行う。

チューニング・パラメータのうち、倒産発生之余波の残る期間をコントロールする p については、Kirchner [6] が提案しているように、次式で与えられる AIC の最小化を基準に決定する方法が提案されている：

$$\text{AIC}^{\Delta_0}(p) := \log \left(\det \hat{\Sigma}^{(\Delta_0)}(p) \right) + \frac{2pm^2}{n_0 - p}, \quad n_0 = \lfloor T/\Delta_0 \rfloor, \quad p \in \{1, \dots, p_0\},$$

ただし、 U_k を Hawkes スケルトン推定の過程で得られる $\text{INAR}(p)$ のホワイトノイズ項ベクトル（付録 C 節を参照）として、 $\hat{\Sigma}^{(\Delta_0)}(p)$ は次で定義されるものとする：

$$\hat{\Sigma}^{(\Delta_0)}(p) := \frac{1}{n_0 - p} \sum_{k=p+1}^{n_0} U_k U_k^T.$$

今回の関東地区の倒産サンプルに対しては、 $\Delta_0 = 21/375, p_0 = 8$ としたうえで AIC 最小化をしたところ、業種カテゴリ別では $p = 4$ 、企業規模カテゴリ別では $p = 2$ が得られた。

また、 $(\Delta_{skel}, \alpha_{skel})$ については、恣意的ではあるが、最終的に Hawkes グラフ推定時には 1 週間を、その前提となる Hawkes スケルトン推定時には 3 週間を単位として、 $\Delta_{skel} = 21/365, \Delta_{graph} = 7/365$ と設定する。倒産報告件数が平日に比べて非常に少ない土曜日と日曜日も含むカレンダー通りの時間軸を用いていることや、平日の中でも比較的金曜日の報告件数が多くなっていることなどをふまえると、1 週間という単位で考えることにはある程度合理性があると判断する^{*11}。

残る $(\alpha_{skel}, \alpha_{graph})$ については、論文 [3] の実験例を参考にして、いずれも 0.05 に設定した（90% 信頼区間で倒産リスクの伝播を表す辺を選択することになる）。

5 分析結果

本節では、2 節に述べた 2003 年初～2015 年末の関東地区の倒産履歴データを業種カテゴリおよび企業規模カテゴリに分類したものをもとに、Hawkes 過程に基づく倒産強度モデルを 2 種類の特徴づけ（指数減衰型カーネル強度の最尤推定、および Hawkes グラフ推定）によって分析を行い、特に推定された伝播構造をネットワーク・グラフとして視覚化し考察を行う。

まず、従来からよく行われていた指数減衰型カーネル強度過程に対する最尤推定の結果を示す。次に、4.4 節のチューニング・パラメータに基づき全期間データを用いて推定した Hawkes グラフの推定結果を示す。

さらに、チューニング・パラメータを変更した場合、および推定に用いるデータ期間を変更した場合に、推定される Hawkes グラフがどの程度当初の推定結果と変わりうるかを調べる。実際に条件を変えた設定に対する推定結果もネットワーク・グラフで表して、当初の推定結果と比較するという単純な検証ではあるが、いくつかのケースを調べて考察を行う。

5.1 指数減衰型カーネル強度過程に対する最尤推定の結果

まず、指数減衰型カーネル強度過程 (2) に対する最尤法によるパラメータ推定結果は表 2, 3 である。なお、目的関数 (3) を最大化する際には、数値的最適化の初期値への依存性を考慮して、 $\xi^{i \rightarrow j}$ について

^{*11} 予備分析として、カレンダーベースの 1 日単位 $\Delta_{graph} = 1/365$ でも推定を行ったが、週単位 $\Delta_{graph} = 7/365$ での推定結果とほとんど同じであった。

表 2 業種カテゴリ別の指数減衰型カーネル強度過程のパラメータ推定結果。製造関係を $j = 1$ 、インフラ・流通関係を $j = 2$ 、金融関係を $j = 3$ 、サービス関係を $j = 4$ 。カッコ内は Hesse 行列から算出される標準誤差。水色は 10% 有意、緑色は 5% 有意

j	$\xi^{1 \rightarrow j}$	$\xi^{2 \rightarrow j}$	$\xi^{3 \rightarrow j}$	$\xi^{4 \rightarrow j}$	μ^j	κ^j
製造 (1)	1.62 (0.93)	1.67 (1.14)	1.96 (1.12)	0.00 (1.21)	15.88 (4.56)	8.07 (1.96)
イン・流通 (2)	1.65 (0.62)	0.00 (1.17)	0.07 (0.70)	0.09 (0.72)	17.97 (5.14)	3.35 (1.87)
金融 (3)	1.48 (0.66)	0.00 (0.93)	1.71 (0.82)	0.00 (0.80)	4.94 (3.10)	5.32 (1.82)
サービス (4)	0.00 (1.02)	2.00 (1.14)	0.00 (0.80)	0.48 (0.99)	13.98 (4.55)	8.67 (7.30)

表 3 企業規模別カテゴリにした際の指数減衰型カーネル強度過程のパラメータ推定結果。小企業を $j = 1$ 、中企業を $j = 2$ 、大企業を $j = 3$ とする。カッコ内は Hesse 行列から算出される標準誤差。水色は 10% 有意、緑色は 5% 有意

j	$\xi^{1 \rightarrow j}$	$\xi^{2 \rightarrow j}$	$\xi^{3 \rightarrow j}$	μ^j	κ^j
小企業 (1)	1.18 (0.71)	0.32 (0.63)	1.33 (0.84)	28.64 (5.97)	6.12 (2.09)
中企業 (2)	0.00 (0.69)	2.00 (0.83)	1.40 (0.82)	21.04 (5.83)	5.47 (1.59)
大企業 (3)	0.00 (0.79)	1.66 (0.64)	1.87 (0.94)	7.89 (4.44)	6.95 (2.43)

は $\{0, 0.3, 0.5\}$ を、 μ^j, κ^j については $\{1, 3, 5, 10\}$ を最適化の初期値の候補とし、計 12 通りの初期値の組合せについて、それぞれ (3) の最大化を行って、最終的に (3) を最大化したパラメータの組を最尤推定値とする。

表 2 から、業種カテゴリ別では、金融関係の自己励起性 ($\xi^{3 \rightarrow 3}$) や製造関係からインフラ・流通関係および金融関係への相互励起性 ($\xi^{1 \rightarrow 2}, \xi^{1 \rightarrow 3}$) が 5% で有意に推定されている。また、有意水準を 10% と少し緩めると、製造関係の自己励起性 ($\xi^{1 \rightarrow 1}$) や、インフラ・流通関係からサービス関係、および金融関係から製造関係という相互励起性 ($\xi^{2 \rightarrow 4}, \xi^{3 \rightarrow 1}$) も許容される。さらに、初期強度に関しては金融関係 (μ^3) だけが 10% でも有意でない。

この推定結果からは、例えば一つの解釈として「金融関係の倒産は製造関係の倒産の伝播によるのみ起因し、その後は両者の自己励起性および相互依存性によって製造関係と金融関係の両者において連鎖的な倒産の可能性も高まる」といった見方ができる。

一方で、企業規模カテゴリ別では、中企業および大企業の自己励起性 ($\xi^{2 \rightarrow 2}, \xi^{3 \rightarrow 3}$) および中企業から大企業への相互励起性 ($\xi^{2 \rightarrow 3}$) が 5% で有意に推定されている。有意水準を 10% と少し緩めると、小企業の自己励起性 ($\xi^{1 \rightarrow 1}$) や大企業から中企業への相互励起性 ($\xi^{3 \rightarrow 2}$) も許容される。この推定結果からは、例えば一つの解釈として「小企業と中・大企業間における倒産リスクの依存関係は無視できるが、中企業と大企業との間は双方向に比較強い依存関係が存在する」といった見方ができる。

後で示す Hawkes グラフと比較できるように、先の指数減衰型カーネル関数に対する最尤推定の結果をもとにグラフ表示したものが図 2, 3 である。

Hawkes グラフでは「辺の重み」として $a_{i,j} := \int_0^\infty h_{i \rightarrow j}(t) dt$ の推定量 $\hat{a}_{i,j}$ を考えるので、それに整合するように、指数減衰型カーネル関数に関しては $\xi_{i \ell, j} / \kappa^j$ で、分子の推定値が有意と見なされるものを辺の重みとして付記している*12。

*12 Hawkes グラフの「辺の重み」 $\hat{a}_{i \ell, j}$ はカーネル関数の時間積分 $\int_0^\infty h_{i \ell \rightarrow j}(t) dt$ の推定値である。一方で、指数減衰型カーネル強度 $h_{i \rightarrow j}(t) = \xi^{i \rightarrow j} e^{-\kappa^j t}$ に対しては $\int_0^\infty h_{i \ell \rightarrow j}(t) dt = \xi^{i \ell \rightarrow j} / \kappa^{i \ell}$ となることに注意する。したがって、Hawkes グラフの強度と指数減衰型カーネル強度を比較するにあたっては、 $\hat{a}_{i \ell, j}$ と $\hat{\xi}^{i \ell \rightarrow j} / \hat{\kappa}^{i \ell}$ とを比較することに少なくとも意味があると考えられる。

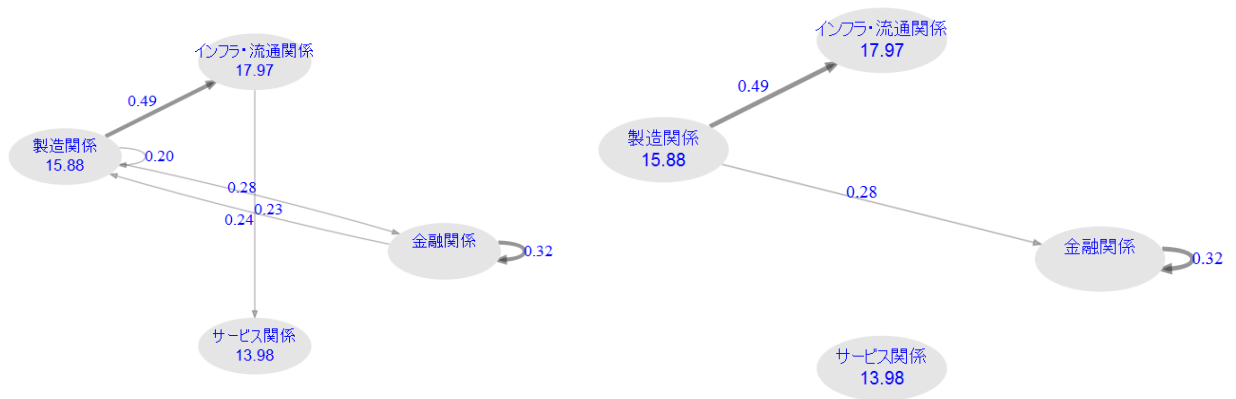


図2 業種カテゴリ別に関して指数減衰型カーネルの結果をグラフ表示したもの。 $\hat{\xi}_{i_\ell, j}$ および $\hat{\mu}^j$ の推定値が有意な辺と定数強度だけを表示。左は10% 有意、右は5% 有意のもの。辺の矢印に付随する数値は $\hat{\xi}_{i_\ell, j} / \hat{\kappa}^j$ である。各カテゴリの楕円内の数値は定数強度の推定値 $\hat{\mu}^j$ の推定値。

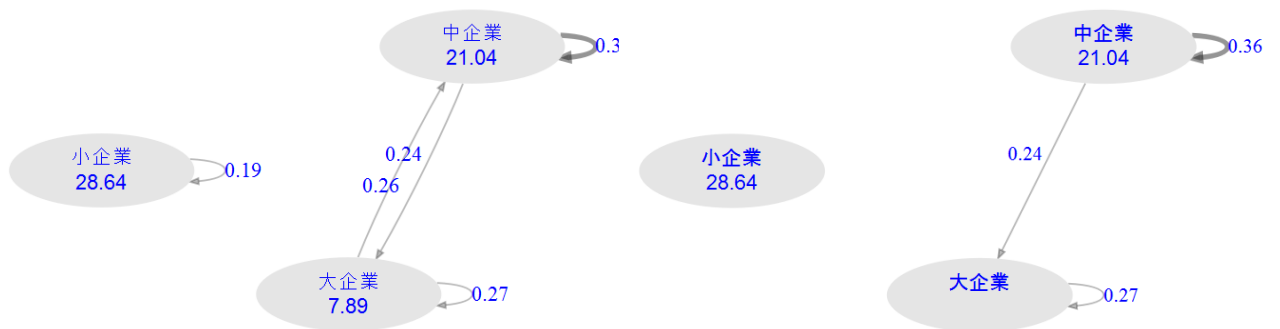


図3 企業規模カテゴリ別に関して指数減衰型カーネルの結果をグラフ表示したもの。 $\hat{\xi}_{i_\ell, j}$ および $\hat{\mu}^j$ の推定値が有意な辺と定数強度だけを表示。左は10% 有意、右は5% 有意のもの。辺の矢印に付随する数値は $\hat{\xi}_{i_\ell, j} / \hat{\kappa}^j$ である。各カテゴリの楕円内の数値は定数強度の推定値 $\hat{\mu}^j$ の推定値。

5.2 Hawkes グラフの推定結果

次に、Hawkes グラフの推定結果を示す。4.4 節で触れたように、分析における基本的設定としてのチューニング・パラメータは、 p については業種カテゴリでは $p = 4$ 、企業規模カテゴリ別では $p = 2$ とする。それ以外は、共通して $(\Delta_{skel}, \Delta_{graph}, \alpha_{skel}, \alpha_{graph}) = (21/365, 7/365, 0.05, 0.05)$ を選択する。

これを受けて、4.1 ~ 4.3 節で述べた方法で推定された Hawkes グラフを図 4 に示している*13。矢印に付随する区間は 4.3 節で言及した「辺の重み」の推定値 $\hat{a}_{i_\ell, j}$ の ($\alpha_{graph} = 0.05$ に対する) 90% 両側信頼区間である。

なお、今回の Hawkes グラフの推定においては Hawkes スケルトンの段階で選ばれた「辺」は全て Hawkes グラフの推定時にも選ばれており、スケルトン推定を省いても推定には影響がなかった可能性

*13 Hawkes グラフの描画に際しては、「辺の重み」 $\hat{a}_{i_\ell, j}$ の推定値の水準を「0 以上 0.3 未満」「0.3 以上 0.5 未満」「0.5 以上」の 3 段階に分けて、矢印の太さを水準の大きさに応じて変えている。

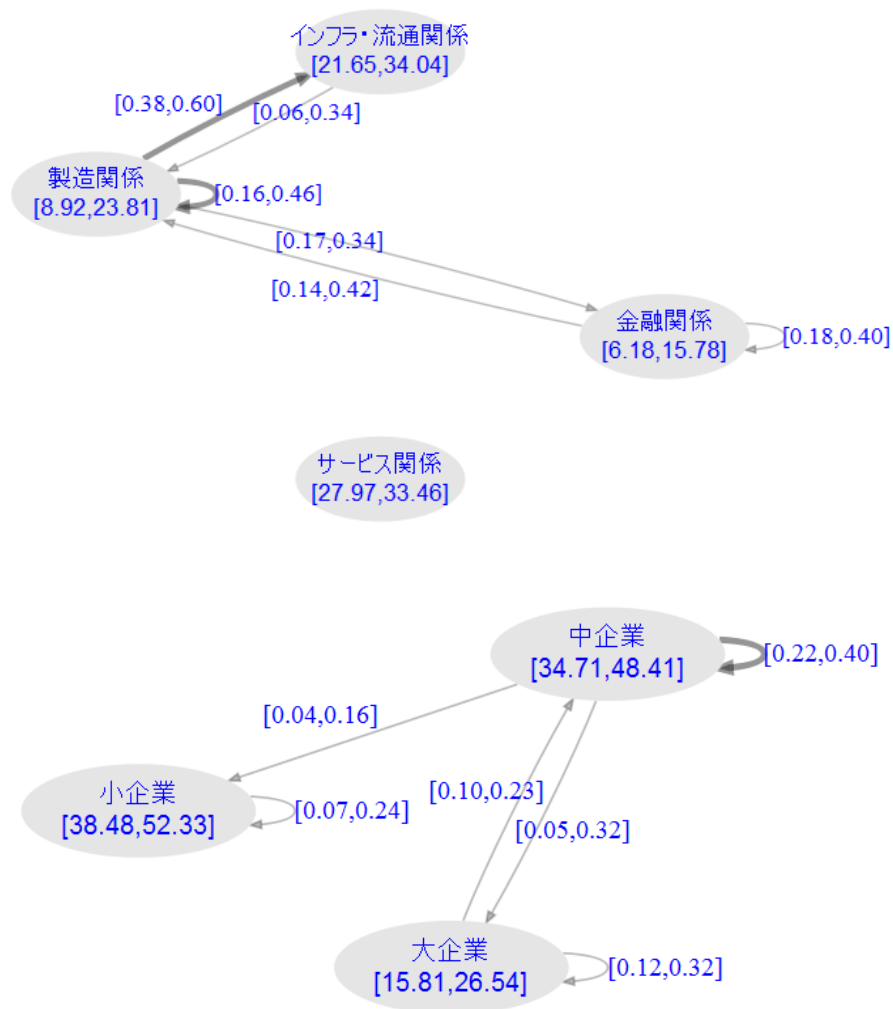


図4 上：業種カテゴリ別に関して推定された Hawkes グラフ ($p = 4, \Delta_{skel} = 21/365, \Delta_{graph} = 7/365, \alpha_{skel} = \alpha_{graph} = 0.05$) 下：規模カテゴリ別に関して推定された Hawkes グラフ ($p = 2$ 後は業種カテゴリ別と同じ)。矢印に付随する区間は $\hat{a}_{i_\ell, j}$ の 90% 信頼区間。各カテゴリの楕円内の区間は「頂点の重み」に当たる定数強度 μ^j の信頼区間。

もある。ただし、一般にスケルトンとグラフという 2 段階の推定が必要ないかどうかについては別途検証の必要があると考える。

業種カテゴリ別の結果 (図 4 左) では、指数減衰型カーネル強度の最尤推定の結果と同様に、サービス関係の自己励起性が認められる一方で、指数減衰型カーネル強度モデルではほとんど認められない「サービス関係以外の自己励起性」や「金融関係から製造関係およびインフラ・流通関係への相互励起性」が Hawkes グラフでは認められるなど、両者の推定結果が示唆する依存関係は異なっている。

一方の企業規模カテゴリ別の結果 (図 4 右) においては、最尤推定で認められた自己励起性に加え、「大企業の自己励起性」が認められる。

そのため、システミック・リスクが大きいと考えられる「金融関係」や「大企業」の影響が捕捉できている点で、Hawkes グラフの推定結果の方がより直感に近いと考えられる。

次に、指数減衰型カーネル強度の最尤推定の結果と、Hawkes グラフの推定結果の乖離度合いを別の観点から見るために、それぞれの推定結果に整合的な強度の事後パスを描いて比較する。

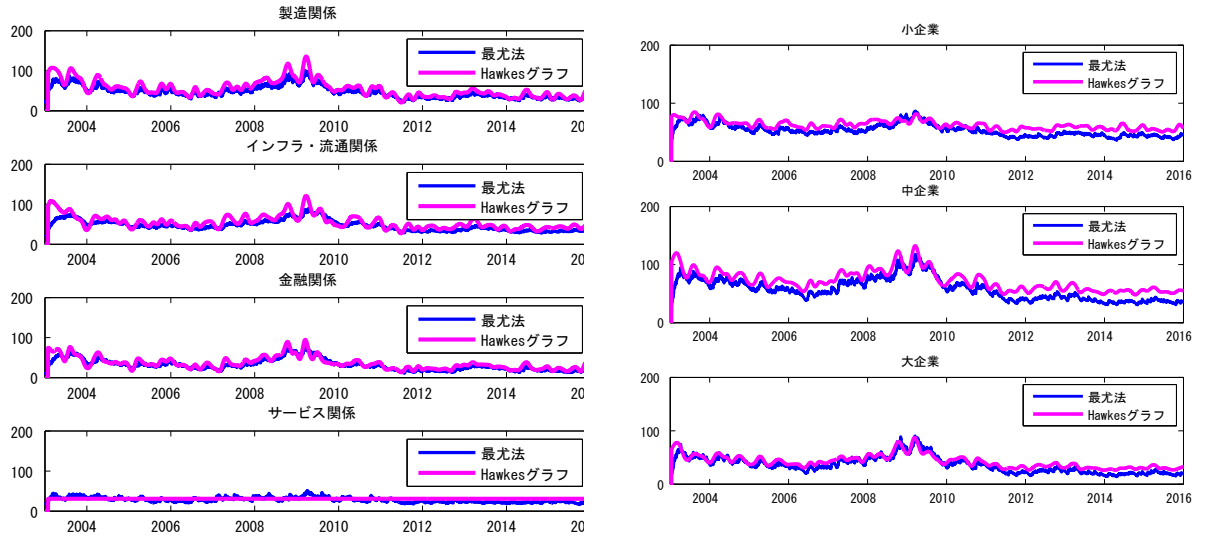


図5 業種カテゴリ別（左）と企業規模カテゴリ別（右）それぞれで、「指数減衰型カーネル+最尤法」と「Hawkes グラフ」それぞれの推定結果をもとに事後的に算出した倒産発生強度 $\hat{\lambda}_t^j$ のパス。横軸は2003年1月1日を起点に表示。Hawkes グラフの結果は、Hawkes 推定量の結果をスプライン補間して表示している。

指数減衰型カーネル強度モデルに関しては、以下のようにパラメータ推定値と倒産発生時点のデータを組み合わせることで、 $\hat{\lambda}_t^j$ の事後強度パスを描くことができる：

$$\hat{\lambda}_{k\Delta_{graph}}^j = \hat{\mu}^j + \sum_{i=1}^m \hat{\xi}^{i \rightarrow j} \sum_{\tilde{\tau}_\ell^i < k\Delta_{graph}} \tilde{\eta}_\ell^i e^{-\hat{\kappa}^j (k\Delta_{graph} - \tilde{\tau}_\ell^i)}.$$

一方で、Hawkes グラフモデルでは、 Δ_{graph} 時点ごとの強度過程の推定値を、Hawkes 推定量と実際の時間区間ごとの観測件数データを用いて計算する。具体的には、 $k = p + 1, \dots, n$ に対して

$$\begin{aligned} \hat{\lambda}_{k\Delta_{graph}}^j &\approx \frac{1}{\Delta_{graph}} \mathbf{E} \left[X_k^{(j, \Delta_{graph})} \mid \{X_\ell^{(\Delta_{graph})}\}_{\ell=k-1, \dots, k-p} \right] \\ &= \left(\mathbf{Z}_j \hat{\mathbf{H}}_j^{(\Delta_{graph}, p)} \right) \text{ の第 } (k-p) \text{ 成分.} \end{aligned}$$

とする。ただし、 $k\Delta_{graph}$ と表示される時点以外の Hawkes グラフの強度の値は、R の spline 関数を用いてスプライン補間している。

図5が、業種カテゴリ別および企業規模カテゴリ別の、指数減衰型カーネル強度モデルと Hawkes グラフの強度の事後パスの推移を比較したものである。

いずれのカテゴリにおいても両者の推移は、振幅の程度に差がある時期もあるが、強度の水準としてはほぼ同程度に推移していると思なせることが確認できる。

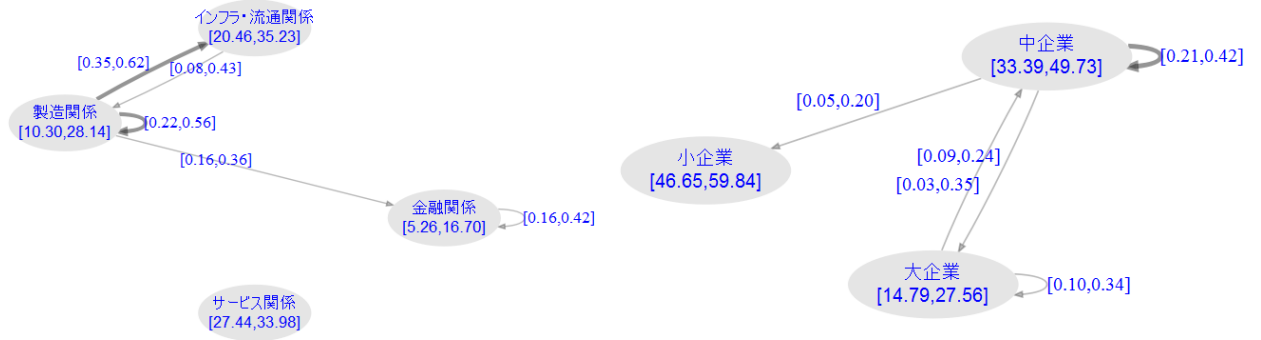


図 6 $\alpha_{skel} = \alpha_{graph} = 0.025$ として推定された Hawkes グラフ (その他のパラメータは最初の分析と同じ)。矢印に付随する区間は $\hat{a}_{i\ell, j}$ の 95% 信頼区間。左は業種別カテゴリ ($p = 4$)、右は企業規模別カテゴリ ($p = 2$)。

5.3 チューニングパラメータの変化に対する Hawkes グラフの推定結果への影響

チューニング・パラメータが変わると、Hawkes グラフ推定の推定結果はどのくらい変わってくるか、というロバスト性を調査するために、以下のようにパラメータごとに基本的設定とは異なるパラメータ値を設定したうえで、Hawkes グラフを最推定する。

以下のような変更パターンを試行し、推定結果を見る。

- $\alpha_{skel}, \alpha_{graph}$: 0.05 (当初の値) 0.025
- p : 2 ~ 10 を総当たりで試す (当初の値は、業種別が 4、企業規模別が 2)
- Δ_{skel} : 21/365 (当初の値) 28/365, 35/365

まず、 $\alpha_{skel}, \alpha_{graph}$ の違いに関しては、図 6 を参照のこと。業種カテゴリ別でも企業規模カテゴリ別でも、AIC 最小化で得た p に対しては、 α_{skel} および α_{graph} を 0.05(90% 信頼区間) から 0.025(95% 信頼区間) としたことで、それぞれグラフの辺が 1 本ずつ消えたが、推定結果に大きな変化が生じたとはまでは言えない^{*14}。

次に、倒産発生の伝播の残存期間をコントロールする p の違いに関して、業種カテゴリ別の結果は図 7 を、企業規模カテゴリ別の結果は図 8 をそれぞれ参照のこと。業種カテゴリ別では、 $p = 3$ および $p = 5$ のときは、AIC 最小化で得られた $p = 4$ と同様のグラフが推定された。 $p = 2$ のときの推定結果は、指数減衰型カーネルの 10% 有意な推定値に基づいて描いたグラフ (図 2 の左図) と同様になった。また、 $p \geq 6$ のときは「製造業」からの作用についての結果が特に不安定となる傾向が見られた。

企業規模カテゴリ別では、 $p = 3, 4, 5$ のときは、AIC 最小化 $p = 2$ と同様のグラフが推定された。 $p \geq 6$ のときは「小企業 中企業」という相互励起性が棄却された以外は同様のグラフが推定された (ただし $p \geq 8$ では「大企業の初期強度」が棄却された)。

最後に、 Δ_{skel} の違いに関して、業種カテゴリ別の結果を図 9 に示す。 $\Delta_{skel} = 28/365, 35/365$ においては、辺の重みは若干違うところもあるが、伝播構造のグラフの形状は同じ推定結果である。しかし、これらは図 4 の右で示される $\Delta_{skel} = 21/365$ の場合の推定結果と比較的大きく異なっていることに注意が必要である。一方、企業規模カテゴリ別では $\Delta_{skel} = 28/365, 35/365$ としても $\Delta_{skel} = 21/365$

^{*14} $\alpha_{skel} = 0.05$ および $\alpha_{graph} = 0.025$ という組み合わせも試したが、ほとんど $\alpha_{skel} = \alpha_{graph} = 0.05$ の場合とほぼ同じ結果であった。

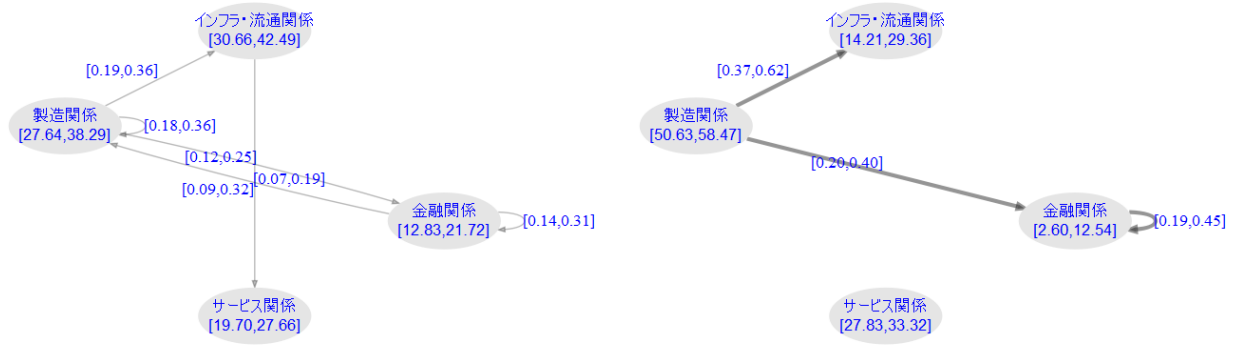


図7 p を変えて業種別カテゴリに対して推定された Hawkes グラフ (その他のパラメータは最初の分析と同じ), 左は $p = 2$ 、右は $p = 6$ の場合

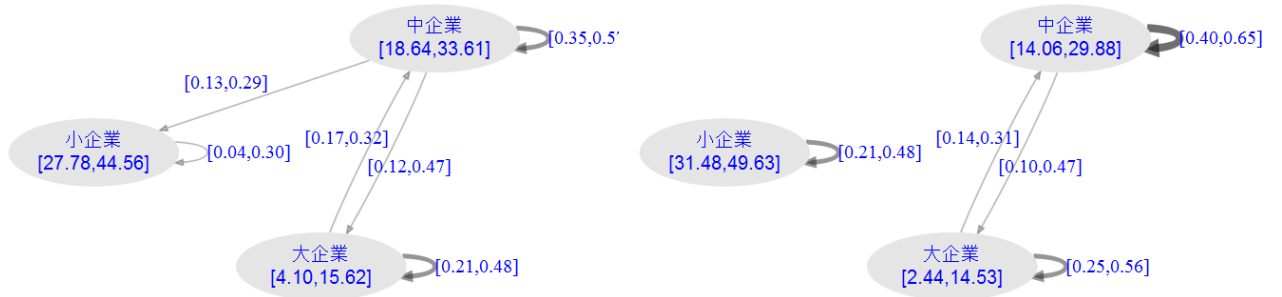


図8 p を変えて企業規模別カテゴリに対して推定された Hawkes グラフ (その他のパラメータは最初の分析と同じ), 左は $p = 4$ 、右は $p = 6$ の場合

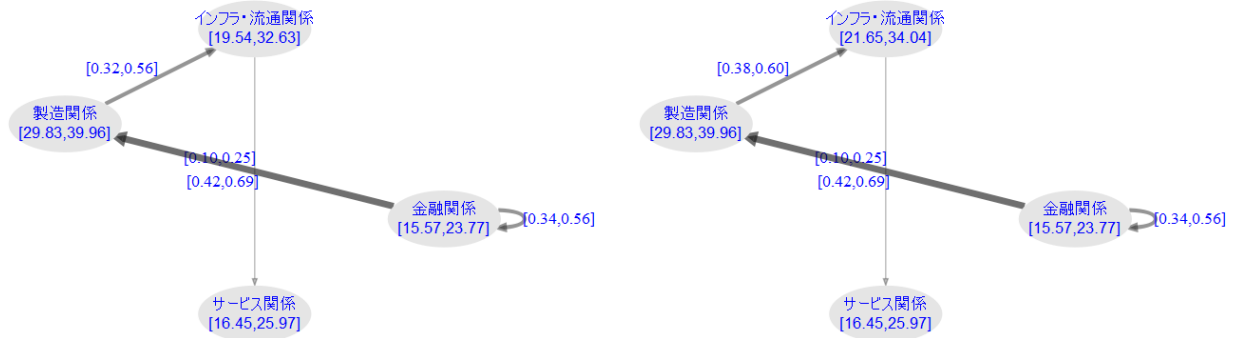


図9 Δ_{skel} を変えて業種別カテゴリに対して推定された Hawkes グラフ (その他のパラメータは最初の分析と同じ), 左は $\Delta_{skel} = 28/365$ 、右は $\Delta_{skel} = 35/365$ の場合

と全く同じ Hawkes グラフが推定された。

5.4 推定期間の違いによる Hawkes グラフの推定結果への影響

次に、Hawkes グラフの推定に用いるデータの利用期間を変えた場合に Hawkes グラフの推定結果がどのように変わるかを見ていく。

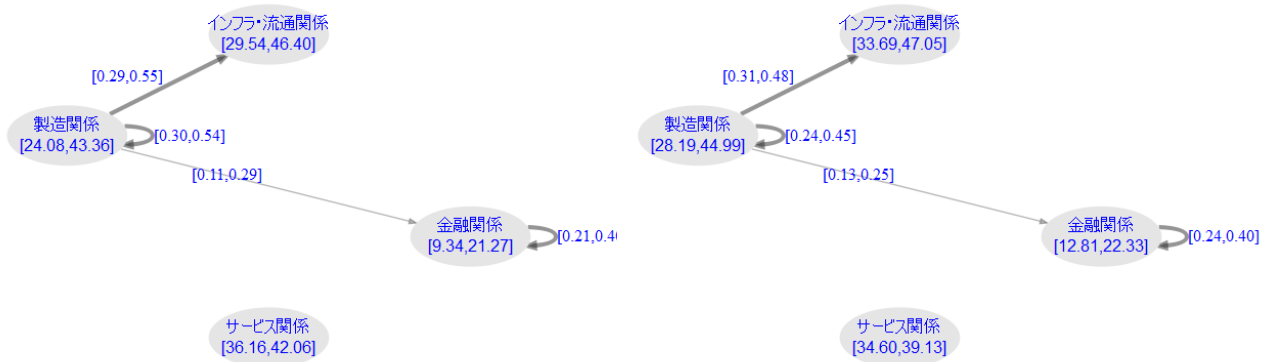
具体的には、業種カテゴリ別、企業規模カテゴリ別それぞれについて (1) 2003 年初～2007 年末 (2) 2003 年初～2010 年末 (3) 2008 年初～2015 年末 (4) 2011 年初～2015 年末 の 4 つの期間のデータを用いて Hawkes グラフを推定する (チューニング・パラメータは 4.4 節で設定した当初のものとする)。

なお、2008 年から 2010 年にかけては、リーマン・ブラザーズ、日本航空、日本振興銀行、富士などの大型の倒産が比較的多い時期ということになる (図 1 の中企業と大企業の倒産件数もこの時期にクラスタリングしているように見える)。

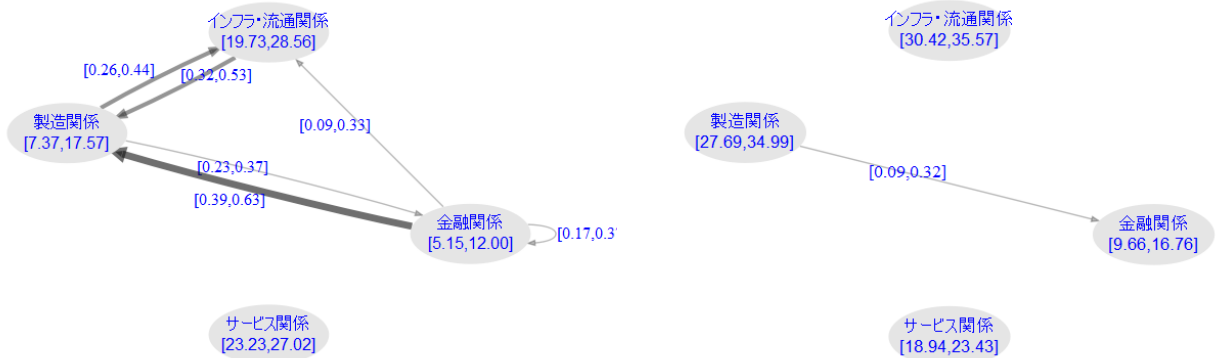
推定結果は、業種カテゴリ別については図 10 に、企業規模カテゴリ別については図 11 にそれぞれ示す。いずれの場合も、(3) 2008 年初～2015 年末に対して推定された Hawkes グラフが最も多くの依存関係を示唆し、特に「金融関係から製造関係」および「大企業から中企業」への相互励起性が比較的高く表れている。一方で、(4) 2011 年初～2015 年末に対しては、業種カテゴリ別では製造関係から金融関係への弱い相互励起性が一つだけ得られ、企業規模カテゴリ別では自己励起性も相互励起性も得られていない。

特に 2008 年～2010 年の時期を含む期間について、自己励起性・相互励起性が比較的多く示唆されていることに注意する必要があると考える。この期間は、背景となるマクロ経済が低調だったために比較的大きな企業の倒産件数が多く、そのことが Hawkes グラフの豊富な自己励起性・相互励起性という形で表出しただけという可能性も大いに考えられる。それゆえ、3.1 節でも議論したように、マクロ経済変数を取り込んだ倒産発生強度モデルを用いた分析を行い、マクロ経済の影響を除いた後でも残る倒産リスクの伝播構造を把握するということが必要になってくると考える。

なお、推定期間の違いによる Hawkes グラフの推定結果の違いについて 付録 D 節に調査してまとめておく。



推定期間：2003年1月1日～2007年12月31日 推定期間：2003年1月1日～2010年12月31日

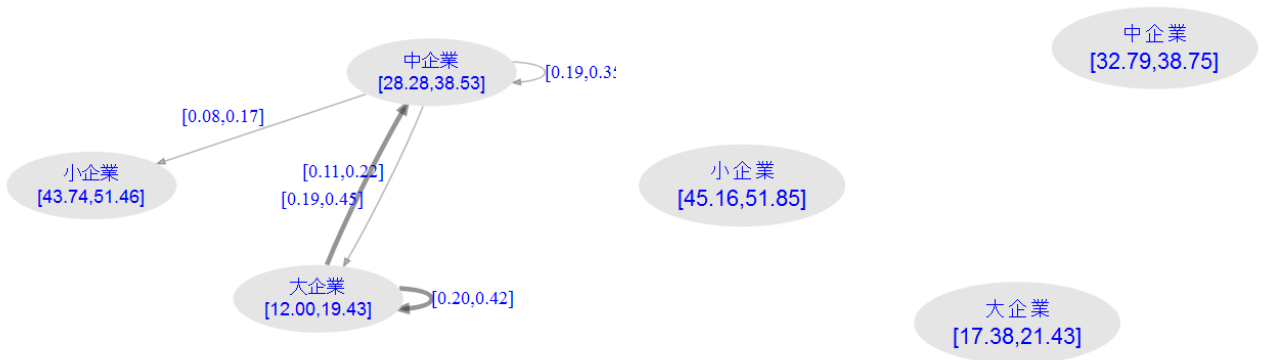


推定期間：2008年1月1日～2015年12月31日 推定期間：2011年1月1日～2015年12月31日

図 10 業種別カテゴリに対して推定された Hawkes グラフ。推定に用いたデータ期間は (左上) 2003年1月1日～2007年12月31日、(右上) 2003年1月1日～2010年12月31日、(左下) 2008年1月1日～2015年12月31日 (右下) 2011年1月1日～2015年12月31日



推定期間：2003年1月1日～2007年12月31日 推定期間：2003年1月1日～2010年12月31日



推定期間：2008年1月1日～2015年12月31日 推定期間：2011年1月1日～2015年12月31日

図 11 企業規模別カテゴリに対して推定された Hawkes グラフ。推定に用いたデータ期間は (左上) 2003年1月1日～2007年12月31日、(右上) 2003年1月1日～2010年12月31日、(左下) 2008年1月1日～2015年12月31日 (右下) 2011年1月1日～2015年12月31日

6 まとめ

本研究では、特に関東地区企業の比較的長期の倒産履歴データを対象に、業種および企業規模に基づいて複数カテゴリに区分して、倒産リスクの依存関係を Hawkes 過程で分析した。特に、Embrechts and Kirchner [3] で提案された Hawkes グラフ推定を試みて、従来から用いられてきた指数減衰型カーネル関数で特徴付けられる強度モデルを最尤推定した結果との比較・考察することを通じて、Hawkes グラフによる倒産リスク依存関係のモデル化に対する応用可能性を検討した。

まず、倒産リスクの伝播構造を示すネットワーク・グラフ図に関しては、Hawkes グラフ・モデルの推定結果の方が、指数減衰型カーネル関数強度モデルよりも多くの依存関係の存在を示唆する傾向が見られた。あくまでも今回のデータセットからの傾向の可能性であることに留意すべきであるが、企業規模カテゴリ別では、指数減衰型カーネル関数強度モデルの依存関係は、Hawkes グラフ・モデルのそれに包含されていた。その一方で、業種カテゴリ別では、指数減衰型カーネル関数強度モデルが示唆した一部の依存関係が、Hawkes グラフ・モデルでは現れなかった。

また、事後強度パスに関しては、いずれのカテゴリで見ても、観測された期間全体で、両モデルは比較的似た水準で推移していた。

以上、あくまでも限定されたデータセットに対する分析結果からの考察であるが、チューニング・パラメータをどのように決定するかという実務的課題はあるが、カテゴリ数が多くなるにつれて、指数減衰型カーネル関数強度モデルよりも、Hawkes グラフの方がリスクの伝播構造を簡便に把握する分析方法として十分に応用可能性があると考えられる。

もちろん、倒産リスク依存関係のモデルとして真に応用可能性を主張するためには、他の期間や地域のデータセットに対する結果を考察したり、別の切り口によるカテゴリ分けに対する推定結果なども十分吟味していく必要がある。くわえて、3.1 節の後半に論じたように、倒産リスクに対する強度モデルとしての有効性を主張するためには、Azizpour et al. [1] などで示唆されている結果をふまえて、倒産発生の強度モデルをマクロ経済変数などの共変量によっても説明可能なように拡張することが不可欠であると考えられる。

そのためには、Hawkes グラフの推定法自体を、マクロ経済変数などを含むように Hawkes 過程を拡張した強度モデルにも適用できるように改良する必要がある。これらの事項については今後の課題としたい。

参考文献

- [1] S. Azizpour, K. Giesecke, G. Schwenkler (2016), “Exploring the Sources of Default Clustering”, *Working paper*, <http://people.bu.edu/gas/papers/exploring.pdf> (accessed on April, 26, 2017).
- [2] E. Bacry, I. Mastromatteo and J.-F. Muzy (2015), “Hawkes processes in finance”, *Working paper*, arXiv:1502.04592.
- [3] P. Embrechts and M. Kirchner (2016), “Hawkes graphs”, *Theory of Probability & Its Applications*, 62(1), 163–193.
- [4] E. Errais, K. Giesecke and L. R. Goldberg (2010), “Pricing credit from the top down with affine point processes”, *SIAM Journal on Financial Mathematics*, 1, 642–665.
- [5] A.G. Hawkes (1971), “Spectra of self-exciting and mutually exciting point processes”,

Biometrika, 58, 83–90.

- [6] M. Kirchner (2017), “An estimation procedure for the Hawkes process”, *Quantitative Finance*, 17(4), 571–595.
- [7] Y. Ogata (1981), “On Lewis’ simulation method for point processes”, *IEEE Transactions on Information Theory*, 27(1), 23–31.
- [8] Y. Ogata (1988), “Statistical models for earthquake occurrences and residual analysis for point processes”, *Journal of American Statistical Association*, 83(401), 9–27.
- [9] S. Yamanaka, M. Sugihara and H. Nakagawa, “Modeling of contagious credit events and risk analysis of credit portfolios,” *Asia-Pacific Financial Markets*, 19(1), 43–62.
- [10] 中川 秀敏 (2010), 「相互作用型の格付変更強度モデルによる格付変更履歴データの分析」, 日本応用数学会論文誌, 20(3), 183–202.

付録 A 対象データ内の負債総額トップ 20

2 節で説明した (株) 東京商工リサーチの倒産履歴データ (2003 年 ~ 2015 年) の中で、分析対象とした企業所在都道府県が関東地区の倒産企業のうち負債総額が上位 20 の企業一覧を表 4 にまとめた。

表 4 (株) 東京商工リサーチの倒産履歴において、企業所在都道府県が関東地区の倒産企業のうち負債総額が上位 20 の企業一覧 (代表者が同一の複数の企業が同じ日に倒産していた場合は負債総額が最も大きい企業に含めた。この中では、(株) 日本航空インターナショナル (2)、松栄不動産 (株)、東京臨海副都心建設 (株)、大洋緑化 (株)、麹町土地建物 (株) がこのケースに該当する)。

倒産発生日	漢字商号	所在地	負債総額 (百万円)	倒産形態	業種カテゴリ
20080916	リーマン・ブラザーズ証券 (株)	東京都	3,431,400	民事再生法	金融関係
20100119	(株) 日本航空インターナショナル	東京都	2,199,497	会社更生法	インフラ・流通関係
20100910	日本振興銀行 (株)	東京都	680,563	民事再生法	金融関係
20070625	麻布建物 (株)	東京都	564,800	会社更生法	金融関係
20120227	エルピーダメモリ (株)	東京都	448,033	会社更生法	製造関係
20100928	(株) 武富士	東京都	433,608	会社更生法	金融関係
20051115	(株) エー・シー・リアルエステート	東京都	352,676	民事再生法	金融関係
20090223	(株) S F C G	東京都	338,040	民事再生法	金融関係
20030217	(株) エスティティコーポレーション	東京都	313,100	民事再生法	サービス関係
20030319	ジャパン石油開発 (株)	東京都	307,600	民事再生法	製造関係
20081010	大和生命保険 (株)	東京都	269,500	会社更生法	金融関係
20030106	松栄不動産 (株)	埼玉県	265,300	破産	金融関係
20060512	東京臨海副都心建設 (株)	東京都	261,000	民事再生法	金融関係
20040210	大洋緑化 (株)	東京都	253,722	会社更生法	サービス関係
20031113	麹町土地建物 (株)	東京都	237,000	破産	金融関係
20070111	(株) エフ・アール・イー	東京都	228,500	破産	金融関係
20030430	(株) アール・エス・ティー	東京都	212,400	破産	金融関係
20100218	(株) ウィルコム	東京都	206,000	会社更生法	インフラ・流通関係
20090205	日本綜合地所 (株)	東京都	197,549	会社更生法	金融関係
20080404	ケイアール不動産 (株)	東京都	167,763	特別清算	金融関係

付録 B Hawkes 推定量による強度推定量近似のアイデア

ここでは、4.1 節で説明している「(小さい) 時間区間におけるイベント発生件数」の時系列データに基づいて算出される Hawkes 推定量 (4) によって、時間軸のグリッド上での強度の推定値 $\lambda_{k\Delta}^j$ が近似的に得られることについてのアイデアを、Kirchner [6] の議論を参考にしておいてまとめる。

なお、あくまでもアイデアに言及する目的であるので、3 節、および 4.1 節の設定以上には、数学的設定を厳密に与えずに論じることをお断りしておく。

まず、計数過程の一般的な性質として $N_t^j - \int_0^t \lambda_s^j ds$ はマルチンゲールになることから、

$$X_k^{(j,\Delta)} := N_{k\Delta}^j - N_{(k-1)\Delta}^j = \int_{(k-1)\Delta}^{k\Delta} \lambda_s^j ds + \text{あるマルチンゲールの時間差}$$

と表すことができる。ここで、過去の履歴 $X_{k-1}^{(\Delta)}, X_{k-2}^{(\Delta)}, \dots$ に関する条件付き期待値をとると、以下

が成り立つことになる：

$$\mathbf{E} \left[X_k^{(j,\Delta)} \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] = \int_{(k-1)\Delta}^{k\Delta} \mathbf{E} \left[\lambda_t^j \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] dt.$$

ここで、強度過程 λ_t^j が (1) 式で与えられているとすると、(十分小さい) $\Delta > 0$ および(十分大きい) $p \in \mathbb{N}$ に対しては、以下のような近似の議論が可能となる^{*15}をすることで、カーネル関数 $h_{i \rightarrow j}(t)$ の離散時間の各時点の値を近似しようというのが Hawkes 推定量の基本的なアイデア： $k = p + 1, \dots, n$ に対して

$$\begin{aligned} \mathbf{E} \left[X_k^{(j,\Delta)} \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] &= \int_{(k-1)\Delta}^{k\Delta} \mathbf{E} \left[\lambda_t^j \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] dt \\ &= \int_{(k-1)\Delta}^{k\Delta} \mathbf{E} \left[\mu^j + \sum_{i=1}^m \int_{-\infty}^t h_{i \rightarrow j}(t-s) dN_s^i \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] dt \\ &\approx \Delta \cdot \mu^j + \sum_{i=1}^m \int_{(k-1)\Delta}^{k\Delta} \mathbf{E} \left[\int_{-\infty}^{(k-1)\Delta} h_{i \rightarrow j}(t-s) dN_s^i \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] dt \\ &\quad (\text{\textit{s} についての積分区間を } (k-1)\Delta \text{ までとする}) \\ &\approx \Delta \cdot \mu^j + \Delta \sum_{i=1}^m \int_{-\infty}^{(k-1)\Delta} h_{i \rightarrow j}(k\Delta - s) \mathbf{E} \left[dN_s^i \mid \mathbf{X}_{k-1}^{(\Delta)}, \mathbf{X}_{k-2}^{(\Delta)}, \dots \right] \\ &\quad (\text{\textit{dt} に関する積分の長方形近似}) \\ &\approx \Delta \cdot \mu^j + \Delta \sum_{i=1}^m \int_{(k-p-1)\Delta}^{(k-1)\Delta} h_{i \rightarrow j}(k\Delta - s) dN_s^i \quad (\text{過去の影響を適当にカットオフ}) \\ &\approx \Delta \cdot \mu^j + \Delta \sum_{i=1}^m \sum_{\ell=1}^p \int_{(k-\ell-1)\Delta}^{(k-\ell)\Delta} h_{i \rightarrow j}(\ell\Delta) dN_s^i \quad (\text{カーネル関数を区分的定数と見なす}) \\ &= \Delta \cdot \mu^j + \Delta \sum_{i=1}^m \sum_{\ell=1}^p h_{i \rightarrow j}(\ell\Delta) \left(N_{(k-\ell)\Delta}^i - N_{(k-\ell-1)\Delta}^i \right) \\ &= \Delta \cdot \mu^j + \Delta \sum_{i=1}^m \sum_{\ell=1}^p h_{i \rightarrow j}(\ell\Delta) X_{k-\ell}^{(i,\Delta)}. \end{aligned}$$

よって、 $X_k^{(j,\Delta)}$ は INAR(p) モデルを満たす形になっており、グリッド上のカーネル関数の値に Δ をかけた $(\Delta h_{i \rightarrow j}(\Delta), \dots, \Delta h_{i \rightarrow j}(p\Delta))$ はまさに線形の重回帰式の回帰係数に相当している。

したがって、グリッド上のカーネル関数の値を行列形式にまとめた Hawkes 推定量が、線形の重回帰モデルにおける OLS 推定量の行列演算式と同じ (4) のような行列演算で近似的に得られることが説明できる。

付録 C Hawkes グラフ推定のアルゴリズム

以下では、Embrechts and Kirchner [3] で提示されている Hawkes グラフ推定において、ネットワークのグラフの辺を残すかどうかを選ぶ際に重要となる「辺の重み」の推定値 $\hat{a}_{i_\ell, j}$ の標準誤差 $\hat{\sigma}_{i_\ell, j}^2$ の計算アルゴリズムを掲載しておく。

^{*15} 下から 2 つ目の近似は、カーネル関数 $h_{i \rightarrow j}(t)$ の台が有限で $[0, p\Delta]$ がその台も含むのであれば、真の等号として成立する。

Hawkes スケルトン推定における $\hat{\sigma}_{i,j}^2$ の計算アルゴリズム

1. $E \in \{0, 1\}^{m^2 \times (m^2 p + m)}$ をベクトル E_1, \dots, E_{m^2} をすべて転置して行ベクトルとして、行方向に順に並べた行列とする。
2. (6) 式で求めた Z を用いて、 $E(Z^\top Z)^{-1} Z^\top \otimes 1_{m \times m} \in \mathbb{R}^{m^2 \times (m(n-p))}$ という行列計算を行い、得られた行列の各行ベクトルを転置して列ベクトルに変換し、それを順に列方向に積み重ねて、 $m^3(n-p)$ -次元ベクトルを作る。さらにこのベクトルを順に m 個の $m^2(n-p)$ -次元ベクトルに分割し、それらのベクトルを行ベクトルに転置して、行方向に順に並べて $\mathbb{R}^{m^2(n-p) \times m}$ -行列を作る。この行列を C とする。
3. $(U_{p+1}, U_{p+2}, \dots, U_n)^\top := U = (Y - \Delta Z \hat{H}^{(\Delta, \text{skel}, p)}) \in \mathbb{R}^{(n-p) \times m}$ を計算する。このとき U の m 次元の行ベクトルは $U_k = \left(X_k^{(\Delta)} - \Delta \hat{\mu} - \sum_{\ell=1}^p \Delta \hat{H}_\ell^\top X_{k-\ell}^{(\Delta)} \right)$ ($k = p+1, p+2, \dots, n$) と表される。さらに、この U を m^2 個行方向に並べて得られる行列を $U^{(rep)} \in \mathbb{R}^{m^2(n-p) \times m}$ とする。
4. 同じ成分どうしをかける行列のアダマール積 $C \odot U^{(rep)}$ を求め、得られた行列の各行の m 個の成分の和を 2 乗して、 $m^2(n-p)$ -次元ベクトルを作る。次に、このベクトルを順に m^2 個の $(n-p)$ -次元ベクトルに分割し、それらのベクトルを行ベクトルに転置して、行方向に順に並べて $\mathbb{R}^{m^2 \times (n-p)}$ -行列を作る。さらに、得られた行列の各行の $(n-p)$ 個の成分の和を計算することで m^2 -次元ベクトルを得る。
5. 上で得られた m^2 -次元ベクトルを順に m 個の m -次元ベクトルに分割し、それらのベクトルを行ベクトルに転置して、行方向に順に並べて得られる $\mathbb{R}^{m \times m}$ -行列が求める $(\hat{\sigma}_{i,j})_{(i,j) \in [m]^2}$ となる。

Hawkes グラフ推定における $\hat{\sigma}_{i_\ell, j}^2$ および $\hat{\sigma}_j^2$ の計算アルゴリズム

まず時間単位を Δ_{graph} に変更することで選び直された k 番目の時間区間に対するカーネル関数の推定値にあたる m_j -次元ベクトルを

$$\hat{H}_{PA(j)}(k\Delta_{graph}) := (\hat{h}_{i_1 \rightarrow j}(k\Delta_{graph}), \hat{h}_{i_2 \rightarrow j}(k\Delta_{graph}), \dots, \hat{h}_{i_{m_j} \rightarrow j}(k\Delta_{graph}))$$

と与える。ただし、 $k \in [p]$, $PA(j) = \{i_1, i_2, \dots, i_{m_j}\}$ とする。また、これら p 個の m_j -次元ベクトルを順に並べ、最後に $\hat{\mu}_j$ を加えて作られる $(pm_j + 1)$ -次元ベクトルを $\hat{H}_j^{(\Delta_{graph}, p)}$ とする。

推定量ベクトル $\hat{H}_j^{(\Delta_{graph}, p)}$ は以下で推定できる^{*16}。

$$\hat{H}_j^{(\Delta_{graph}, p)} = \frac{1}{\Delta_{graph}} (Z_j^\top Z_j)^{-1} Z_j^\top Y_j.$$

ただし、この Y_j, Z_j はそれぞれ (5), (6) 式において、 $\Delta = \Delta_{graph}$ および $m = m_j$ として、いずれも j に対しての推定結果のみを含んだ行列としている。

前のアルゴリズムとほぼ同様の以下の手順で、 (i_ℓ, j) (ただし $\ell \in [m_j], j \in [m]$) に対して $\hat{\sigma}_{i_\ell, j}^2$ が得られ、これから信頼区間を作り出すことで最終的に Hawkes グラフが得られる。

1. (i_ℓ, j) に対し、 $e(i_\ell, j)$ を $(k-1)m_j + \ell$ -成分^{*17} ($k \in [p]$) だけ 1 で、残りは 0 となる $(m_j p + 1)$ -次元ベクトルとする。

^{*16} 推定された Hawkes グラフに整合的な強度過程の事後パスは、この推定量に基づいて得ることができる。

^{*17} 原論文では $(k-1)m_j + (\ell-1)$ -成分となっていたが、 $k=1, \ell=1$ で意味をなさない。

2. $(n-p)$ -次元ベクトル $C_{l,j} := \left((Z_j^\top Z_j)^{-1} Z_j^\top \right)^\top e(i_{l,j})$ を計算する。
3. $(n-p)$ -次元ベクトル $U_j = Y_j - \Delta_{graph} Z_j \hat{H}_j^{(\Delta_{graph}, P)}$ を計算する。
4. アダマール積 $C_{l,j} \odot U_j$ として得られるベクトルの各成分を 2 乗して各成分を足すと $\hat{\sigma}_{i_{l,j}}^2$ が得られる。
5. $(n-p)$ -次元ベクトル Z_{Last} を $((Z_j^\top Z_j)^{-1} Z_j^\top)$ の最終行とすると、アダマール積 $Z_{Last} \odot U_j$ として得られるベクトルの各成分を 2 乗して各成分を足した後に Δ_{graph}^2 を乗ずることにより、 $\hat{\sigma}_j^2$ が得られる。

付録 D Hawkes グラフの各辺が有意と推定された期間

推定期間の Hawkes グラフの推定結果への影響を調査するために、業種カテゴリおよび企業規模カテゴリそれぞれに対して次のような調査を行った。なお、この分析では $\Delta_{graph} = 1/365$ としている。

1. 最初に、2003/1/1 ~ 2005/12/31 (1096 日分) を用いて Hawkes グラフを推定し、その後は推定に用いるサンプル日数は 1096 日分を固定しながら (moving-window 方式) 1 日ずつ期間をずらして Hawkes グラフを推定していく
2. 最初に、2003/1/1 ~ 2005/12/31 (1096 日分) を用いて Hawkes グラフを推定し、その後は推定に用いるサンプル数を増やしながら (サンプル累積方式) 1 日ずつ期間をずらして Hawkes グラフを推定していく

図 12 は、業種カテゴリおよび企業規模カテゴリそれぞれに対して「moving-window 方式」と「サンプル累積方式」で推定した Hawkes グラフにおいて、それぞれの自己励起性・相互励起性について有意と選択された期間を表している。

「moving-window 方式」の結果をみると、有意となる自己励起性・相互励起性は、推定に用いるサンプル期間によって大きく異なることが確認される。業種カテゴリにおける「製造関係 金融関係」という相互励起性が認められる期間が比較的多いものの、期間全体を通じて「moving-window 方式」による推定で有意となる伝播の自己励起性・相互励起性は無い。「moving-window 方式」は直近 3 年分の伝播関係を捕捉するとも考えられ、推定の結果有意となった自己励起性・相互励起性は、むしろその時点におけるマクロ経済の状況で説明されていしまう可能性が十分にある。

一方の「サンプル累積方式」は「moving-window 方式」の結果と大きく異なっている。「サンプル累積方式」では、ある時点で自己励起性・相互励起性が有意と認められると、その伝播性はそれ以降の比較的長い期間続けて有意に現れるケースが多い。両者を比較したとき、直近に認められる伝播関係を捕捉する「moving-window 方式」に比べて、「サンプル累積方式」は伝播関係を過大評価しているとも考えられる。

いずれにしても、本文でも触れたようにマクロ経済変数等の要因を除外したうえで、それでもなお倒産リスクの自己励起性・相互励起性が残るかどうかを分析することが次の課題である。

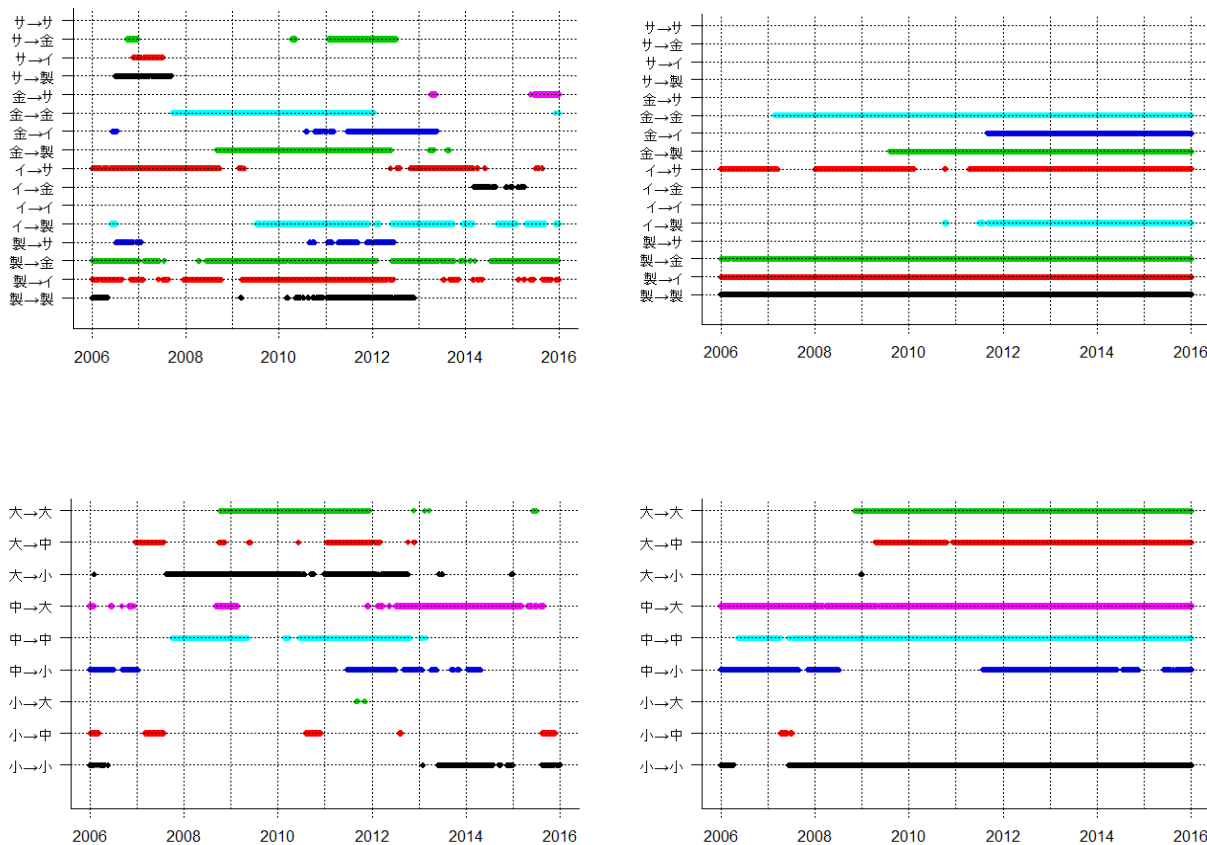


図 12 2006 年初から 2015 年末の各日で推定された Hawkes グラフのうち、それぞれの自己励起性・相互励起性について有意と選択された期間。左上が「業種カテゴリの moving-window 方式」に対する結果、右上が「業種カテゴリのサンプル累積方式」に対する結果。「製」は製造関係、「イ」はインフラ・流通関係、「金」は金融関係、「サ」はサービス関係のカテゴリをそれぞれ示す。また、左下が「企業規模カテゴリの moving-window 方式」に対する結果、右下が「企業規模カテゴリのサンプル累積方式」に対する結果。