統計数理(2013)
第 61 巻 第 1 号 147–166

［総合報告］

アイゲンファクターを知る

増田 直紀[†]

（受付 2012 年 10 月 21 日；改訂 12 月 11 日；採択 12 月 11 日）

要　　旨

インパクトファクターは，論文誌の価値を定量化する有名な指標である．その是非については度々議論されていつつも，国内外を問わず広く使われている．2007 年に，論文の引用–被引用のネットワークを用いた新しい論文誌評価指標であるアイゲンファクターが提案された．2009 年からは，インパクトファクターを発表するトムソン・ロイター社の Journal Citation Reports でもアイゲンファクターの値が発表されるようになり，注目が高まっている．指標の値がひとり歩きしないためにも，これらの指標を参照する研究者や評価者は，指標の意味を正しく理解する価値がある．本稿は，アイゲンファクターの仕組みを，インパクトファクターと対比しながら説明する．アイゲンファクターは，グーグルがウェブサイトを順位づけするために検索エンジンに用いているページランクとほぼ同じである．したがって，本稿は，アイゲンファクターより有名であると思われるページランクを丁寧に解説することを通じて，インパクトファクターやアイゲンファクターについての理解を深めるという道筋をとる．インパクトファクターやアイゲンファクターについて誤解されがちな点についても議論する．

キーワード： 有向グラフ，ネットワーク，ページランク，インパクトファクター，中心性，計量書誌学.

1. はじめに

Impact Factor（インパクトファクター．以下では IF と記す）は，大多数の研究者が知る，論文誌（「論文」誌でないものも含みうるが，以下では単に論文誌と記す）の評価指標である．トムソン・ロイター社が提供する商品の 1 つである Journal Citation Reports®で，毎年，値が公表されている．IF を考案したのは，Institute for Scientific Information（現トムソン・ロイター）の創始者である Eugene Garfield である（1955）．IF の考え方の起源は，1920 年代にまで遡る（Gross and Gross, 1927; Pinski and Narin, 1976; ウェスト, 2008）．

IF は，大学図書館が行う論文誌を購読するか否かの意思決定から，大学のランキング，研究費助成対象の選定材料，研究者の雇用や昇進に際する資料としてまで，国内外を問わず広く用いられている．IF は，論文誌の格付けであり，個々の論文や研究者を評価する指標ではないにも関わらずである．

3 節で述べるように，IF についてはいくつかの短所が指摘されている（角田・小野寺, 2007 の序論なども参考になる）．そこで，IF に代わる論文誌の格付け指標が提案され，異なる指標間の比較も行われている（Bollen et al., 2009; 佐藤, 2009）．その中で，Eigenfactor$^{\mathrm{TM}}$（アイゲンファクター．以下 EF）という指標が 2007 年に提案され（eigenfactor.org, 2012），2009 年 1 月か

[†] 東京大学大学院 情報理工学系研究科：〒113–8656 東京都文京区本郷 7–3–1

ら，Journal Citation Reports でも各論文誌の EF が提供されるようになった．

EF をなるべくわかりやすく，かつ，正確に解説するのが本稿の目的である．

EF の解説は，日本語 (中西印刷株式会社, 2009, 2011; ウェスト, 2008; 佐藤, 2009) も英語 (zeigenfactor.org, 2012) もあるが，その数理面や，IF との本質的な違いなどの解説について，必ずしも十分でない印象を受ける．EF が Journal Citation Reports でも提供されるようになり，提案者の一人が日本で講演を行った (ウェスト, 2008) 昨今，日本国内でも EF への注目が高まっているように見受ける．そういうわけで，本稿を執筆しようと思った．

なお，本稿は，EF に興味がある読者全般を想定する．読者の中には，数式にはあまり興味のない方もいるかもしれない．そのような読者は，細かな数式は飛ばして数式以外の文章と例を追えば，それなりの理解ができるだろう．

2. ページランク

2.1 設計指針

EF は，グーグルが検索エンジンに用いているページランク (以下，PageRank) とほとんど同じである．PageRank は，ネットワークに基づいてウェブページに順位をつける方法であり，ネットワーク科学のビジネス応用が最も成功した事例の 1 つである．重要なウェブページの決定だけでなく，病気の原因遺伝子の推定 (Chen et al., 2009; Iván and Grolmusz, 2011)，特許の順位づけ (Wu et al., 2011)，テニス選手のランキング (Radicchi, 2011) などにも応用されている (増田, 2012a の第 4 章に，これらの応用についての解説がある)．

以下で見るように，PageRank には，細かな調整箇所がいくつかあることに応じて，幾つかの変種がある．EF は PageRank の一種であると言える．PageRank の本質を理解できれば，EF を理解できる．そして，PageRank の方が EF よりも有名，かつ，直感的にも少しわかりやすい．したがって，本節では PageRank を丁寧に説明する．PageRank の他の解説には，Baldi et al. (2007) の 5 章，Langville and Meyer (2009)，増田・今野 (2010) の 9.2.1，9.2.2 節，増田 (2010, 2012a) などがある．

検索エンジンとは，検索語に関係するウェブページに得点をつけて，得点順に並べる方法である．ウェブ (World Wide Web, WWW) は，ウェブページを頂点，クリックしてウェブページからウェブページへと飛ぶことができるハイパーリンクを枝とするネットワークである．図 1 はネットワークの例であり，黒丸が頂点，矢印が枝である．ネットワーク科学では，ネットワーク構造に基づいて頂点に得点をつける仕組みのことを中心性と呼ぶ (増田・今野, 2010; 増田, 2012a)．例えば，頂点に集まる枝の数をその頂点の得点と考えることは，中心性の定義の 1 つである．ネットワークをひとつ固定しても，基準に応じて様々な中心性の指標がありうる．PageRank は中心性の一例である．IF や EF も，中心性の例である．

グーグルの共同設立者である Brin と Page は，1998 年に PageRank を考案し，グーグル検索の理論的な土台を作った (Brin and Page, 1998)．PageRank のページとは，ウェブページの意味であるとともに，設立者の 1 人で 2012 年現在の CEO である Lawrence Page の苗字でもある．PageRank では，重要なウェブページは以下の基準を満たすページであると規定して，各ページに得点をつける．

- 基準 1：多くのページからリンクされること．
- 基準 2：重要なページからリンクされること．
- 基準 3：厳選されたリンクを受けとること．

図 2 において，丸の大きさが PageRank を表すとすると，頂点 1 の PageRank は頂点 2 の PageRank より大きく (基準 1)，頂点 3 の PageRank は頂点 4 の PageRank より大きく (基準

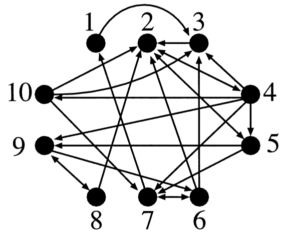

図 1.　ネットワークの例.

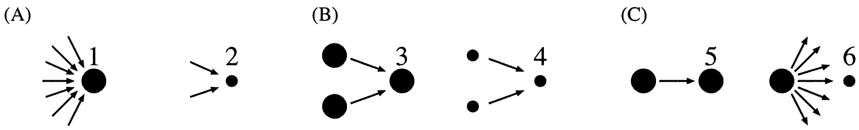

図 2.　PageRank の 3 つの基準.（A）基準 1,（B）基準 2,（C）基準 3. ●の大きさは頂点の PageRank を表す.

2)，頂点 5 の PageRank は頂点 6 の PageRank より大きい(基準 3).

基準 1 は，ウェブページの重要性の指標としていかにもふさわしそうだ．しかし，仮に基準 1 だけで PageRank が定義されているなら，自分のページの PageRank を上げることは容易になってしまう．たくさんのゴミページを作り，それらのゴミページから順位を上げたい自分の本物のページにリンクすればよい．そのような方法でページの重要性が左右されては困るので，基準 2 が必要である．ゴミページは重要なページでないので，そのようなページからいくら多くリンクしても，自分の本物のページの PageRank はほとんど上がらない.

基準 3 も必要である．もし自分のページが有名なページからリンクされていても，その有名ページがポータルサイトや一覧表示的なまとめサイトならば，そのうちのリンク 1 本を受け取ることにさしたる価値はないだろう．有名ページとはいえ，リンクを多く振る舞いすぎているのである．そうではなくて，少数のリンクしか出していないような有名サイトから，そのうちの 1 本のリンクを受け取ると，自分のページの価値が大きく上がる.

2.2　PageRank の定義と例

PageRank の数理を説明するには行列が必要である．与えられたネットワークの頂点数を N とし，隣接行列と呼ばれる $N\times N$ 行列 A の i 行 j 列要素 A_{ij} $(1\le i,j\le N)$ を,

$$A_{ij}=\begin{cases}1 & (\text{頂点 } i \text{ から頂点 } j \text{ に枝がある}),\\ 0 & (\text{頂点 } i \text{ から頂点 } j \text{ に枝がない}),\end{cases} \tag{2.1}$$

で定める．$A_{ii}=0$ とする．1 つのネットワークは，その隣接行列 A と等価である．例えば図 1 のネットワークの隣接行列は

$$A = \begin{pmatrix} 0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&1&1&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0 \\ 0&1&1&0&1&0&1&0&1&1 \\ 0&1&0&0&0&0&1&0&1&0 \\ 0&1&1&0&0&0&1&0&0&0 \\ 1&0&0&0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&1&0&1&0&0 \\ 0&1&1&0&0&0&1&0&0&0 \end{pmatrix} \tag{2.2}$$

である．

ネットワーク科学では，リンクに方向がない場合，すなわち，i から j にリンクがあれば j から i にもリンクがある無向ネットワークを多く扱う．この場合は $A_{ij}=A_{ji}$ なので，A は対称行列である．しかし，ウェブにおいては，A は一般に非対称である．自分が相手のページにリンクをはっても，相手は自分のページにリンクをはってくれるとは限らない．

本当は細かな配慮が必要なのだが，まずは，PageRank の本質を理解するために，ページ i の PageRank $x_i\,(1\le i\le N)$ を

$$x_i=\sum_{j=1}^{N} B_{ji}x_j, \tag{2.3}$$

$$B_{ji}=\frac{A_{ji}}{\sum_{\ell=1}^{N} A_{j\ell}}, \tag{2.4}$$

$$\sum_{i=1}^{N} x_i=1 \tag{2.5}$$

で定義する．式 (2.5) は，全ウェブページの PageRank の和が 1 になるようにする規格化条件である．

この定義が基準 1~3 を満たすことを確認しよう．まず，ページ i がたくさんのページから枝を受けとるとき，式 (2.3) の右辺にたくさんの 0 でない項がある．右辺の各項の値の大小にはよるが，左辺の x_i は，右辺の項が多い分だけ大きくなりやすい (基準 1)．次に，ページ i が PageRank の大きいページ j から枝を受けとるとき，式 (2.3) の右辺に大きな値の x_j が現れる．したがって，左辺の x_i は大きくなる (基準 2)．最後に，ページ i が j から貴重な枝を受けとる状況を考える．j からの枝が貴重であるとは，j から出るリンク数 $\sum_{\ell=1}^{N} A_{j\ell}$ が小さいことである．すると，式 (2.3) の右辺に現れる B_{ji}，すなわち $A_{ji}/\sum_{\ell=1}^{N} A_{j\ell}$ は，分母が小さい分だけ大きくなる．今，j から i にリンクがある状況を考えているので $A_{ji}=1$ であることに注意する．すると，左辺の x_i も大きくなりやすい (基準 3)．結局，式 (2.3), (2.4), (2.5) は，基準 1～3 を満たす定義である．

式 (2.3) の行列 B は，式 (2.2)，つまり図 1 の例に対しては

$$B = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \end{pmatrix} \tag{2.6}$$

となる．PageRank の 3 つの基準の意味を確認するために，式 (2.3) の $i=2$ の場合の式，すなわち，

$$x_2 = x_3 + \frac{x_4}{6} + \frac{x_5}{3} + \frac{x_6}{3} + \frac{x_8}{2} + \frac{x_{10}}{3} \tag{2.7}$$

について考えよう．式 (2.7) の右辺の項が比較的多いことは，頂点 2 が受け取るリンクの数 (入(いり) 次数という) が大きいことを反映している (基準 1)．図 1 や式 (2.2) からもわかるように，頂点 2 の入次数は 6 であり，他のどの頂点の入次数よりも大きい．そして，x_3，x_4，x_5，x_6，x_8，x_{10} が平均的に大きければ x_2 も大きくなる (基準 2)．また，頂点 2 は，頂点 3 が出す唯一の枝を受けとっている (基準 3)．式 (2.7) の右辺の各項の分母は，対応する頂点が出す枝の数 (出(で) 次数という) である．出次数が小さいほど，その頂点が出している各枝は貴重である．

この例に対して，2.5 節で説明する方法で各頂点の PageRank を計算すると，表 1 となる．各頂点の入次数，出次数も記した．PageRank は入次数 (基準 1) と相関している．例えば，頂点 2 は，PageRank も入次数も，全頂点の中で最大である．しかし，PageRank と入次数は本質的に異なる．例えば頂点 1 と 4 は入次数が共に 1 だが，x_4 は x_1 の 2 倍以上大きい．また，頂点 4 と 6 を比べると，PageRank では頂点 4 の方が大きいが，入次数では頂点 6 の方が大きい．また，PageRank と出次数は特に関係がない．例えば，頂点 5 と 6 は，入次数も出次数も等しいが，$x_5 \neq x_6$ である．また，出次数が最大である頂点 4 は，特に PageRank が大きいわけではない．頂点 4 は，入次数が 1 しかない割には PageRank が大きく，それは出次数が大きいからであると思われるかもしれない．しかし，頂点 1 と 10 を比べると，入次数は同じで，出次数は頂点 10 の方が頂点 1 よりも 3 倍大きいが，PageRank は頂点 1 の方が倍以上大きい．

なお，枝に方向のない場合 ($A_{ij} = A_{ji}$) は，直接の代入により $x_i^* = k_i / \sum_{\ell=1}^{N} k_\ell$ が式 (2.3) の

表 1． 図 1 のネットワークにおける各点の PageRank.

頂点	x_i	入次数	出次数
1	0.0518	1	1
2	0.2350	6	2
3	0.1097	4	1
4	0.1175	1	6
5	0.1371	2	3
6	0.0953	2	3
7	0.1036	4	2
8	0.0435	1	2
9	0.0870	3	2
10	0.0196	1	3

解であることがわかる(増田・今野, 2010 の 9.2.1 節).ここで,$k_i = \sum_{\ell=1}^{N} A_{i\ell} = \sum_{\ell=1}^{N} A_{\ell i}$ は頂点 i の次数(直接つながっている枝の数)である.つまり,基準 2 と 3 も考慮して PageRank を定義したはずなのに,この場合は,基準 2 と 3 が効かず,基準 1 だけで求めたのと同じことになってしまう.PageRank が効果を発揮するのは,枝に方向があるネットワークの場合のみである.

2.3 ランダム・ウォークで見る PageRank

PageRank をランダム・ウォークの言葉で理解することは,特に,2.5 節で説明する PageRank の微調整を理解するために有益である.そして,EF は本質的に PageRank でありながらも PageRank のとある変種を用いていて,EF を細かに理解するためには PageRank の微調整を理解することが必要になる.したがって,本節では,ランダム・ウォークの立場から見た PageRank を説明する.PageRank と EF は大体同じである,という大雑把な理解で十分だという読者は,次の段落を読んだ上で,本節を飛ばしてもよい.

今いるウェブページ上のハイパーリンクの 1 つを等確率で選び,クリックする.行き先のページで,やはり等確率で 1 つのハイパーリンクを選んでクリックし,また次のページへ飛ぶ.このような「ランダム・サーフ」を繰り返すユーザを考える.実は,x_i は,ランダム・サーファーがページ i を訪れる確率である.このようなランダム・サーフのことを,ネットワーク上の単純ランダム・ウォークと呼ぶ.

頂点(ウェブページ) i にいるランダム・ウォーカーが次の時刻に j に行く確率について考える.(a) i から j へ枝がなければ,一歩で i から j に行けないので,この確率は 0 である.(b) i から j へ枝があれば,この移動確率は $A_{ij}/\sum_{\ell=1}^{N} A_{i\ell}$ である.(b) のとき,$A_{ij}=1$ であること,および,$\sum_{\ell=1}^{N} A_{i\ell}$ は i の出次数であることに注意する.実は,(a) の場合も,移動確率は $B_{ij} \equiv A_{ij}/\sum_{\ell=1}^{N} A_{i\ell}$ でよい(この B の定義は,式 (2.3) で与えられた B の定義と合致する).なぜなら,(a) の場合には $A_{ij}=0$ となるからである.$N \times N$ 行列 $B=(B_{ij})$ は,ランダム・ウォークの推移確率行列と呼ばれる.

$P_i(t)$ を,ウォーカーが時刻 t で i にいる確率とする.$P_i(t)$ の時間発展は,

$$P_i(t+1) = \sum_{j=1}^{N} P_j(t) B_{ji} \quad (1 \le i \le N) \tag{2.8}$$

で表される.

十分に時間がたったときにウォーカーが頂点 i にいる確率(定常密度)を P_i^* と書く.式 (2.8) で $t \to \infty$ と思って $P_i^* = P_i(t) = P_i(t+1)$ とすると

$$P_i^* = \sum_{j=1}^{N} P_j^* B_{ji} \quad (1 \le i \le N), \tag{2.9}$$

$$\sum_{i=1}^{N} P_i^* = 1. \tag{2.10}$$

式 (2.3), (2.5) と比べると,P_i^* は x_i と同一であることがわかる.したがって,PageRank は,ランダム・サーファーがページ i を訪れる頻度に等しい.

実際のウェブでは,画面の上の方にあるハイパーリンクほどクリックされやすかったり,表示方法によってリンクがクリックされやすくなったりされにくくなったりする.PageRank は,そういう事情を無視してすべてのクリック活動が無作為に起こると見なす近似に対応する.

2.4 強連結でないネットワークの PageRank

どの頂点からどの頂点へも枝の方向に沿って歩けばいずれ到達できるとき,そのネットワー

クは強連結であるという．図 1 は，強連結なネットワークの例である．強連結なネットワークに対しては，Perron–Frobenius の定理という線形代数の定理（齋藤, 1966; 伊理, 2003; 木村, 2003）により，全ての頂点の PageRank x_i が正であることが保証されている．また，PageRank の定義も式 (2.3) で問題ない．

図 3 は，強連結でないネットワークの例である．このようなネットワークに対して PageRank を求めるとき，以下の 2 つの問題が発生する．

第 1 に，強連結でないネットワークには，自分から出る枝が 1 本もない行き止まり（ぶら下がり頂点，dangling node と呼ばれる）がしばしば存在する．図 3 の頂点 8 がその例である．ウェブでは，ハイパーリンクを持たないページがぶら下がり頂点であり，画像のページや pdf 文書のページが典型例である．ぶら下がり頂点 j に対しては $\sum_{\ell=1}^{N} A_{j\ell} = 0$ なので，式 (2.4) と式 (2.3) の右辺が未定義になってしまう．ランダム・ウォークでは，j からの出口がないので，j にウォーカーが到達した場合に次の行き先を決められないことに対応する．

第 2 に，図 3 の頂点 1 のように，どの頂点からも枝を受けとらないページがしばしばある．式 (2.3) で $i=1$ とすると，（ぶら下がり頂点の問題をさて置けば）全ての j について $A_{ji}=0$ なので $x_1=0$ となる．ランダム・ウォークでは，長時間後には頂点 1 にウォーカーが訪れないことに対応する．よって，そのような頂点 1 から出る枝を受け取ることが，受け手（図 3 では頂点 2 と 6）の PageRank の上昇に全く寄与しないことになる．

頂点 1 は，単一頂点が最上流に位置する場合だが，いくつかの頂点からなる集合が「最下流というわけではない」位置にある場合も同様である．例えば，頂点 2，3，4 は，そのようなグループを成す．頂点 2，3，4 は他の頂点からの枝を受け取るが，実は $x_2 = x_3 = x_4 = 0$ である．この 3 頂点のグループは，それより上流に頂点 1 という「グループ」が存在するため，最上流に位置するわけではない．ただ，より下流に 5，6，7 の 3 頂点から成るグループが存在することが理由で，$x_2 = x_3 = x_4 = 0$ となる．すると，これらの頂点から枝を受け取ることが，受け手（図 3 では頂点 7 と 8）の PageRank の上昇に全く寄与しない．要するに，式 (2.3) に基づくと，ぶら下がり頂点が存在しないとしても，最下流グループ以外の頂点の PageRank は 0 になってしまう．

ウェブを含めて，世の中のたいていの有向ネットワークは強連結でない．したがって，一般の有向ネットワークに対して PageRank を定義するためには，式 (2.3) を修正する必要がある．

修正は 2 項目からなる．まず，第 2 の点に対処するために，ランダム・ウォーカーは，単純ランダム・ウォークを確率 $1-q$ で行い，残りの確率 q でネットワーク構造にとらわれないある種の移動を行うことにする．確率 q の移動にはテレポーテーションという奥ゆかしい名前がついている．図 3 の頂点 5，6，7 のような最下流の集団にウォーカーが到達しても，テレポーテーションによって，他の頂点へと脱出できるようになる．そうなれば，この修正されたランダム・ウォークが長時間行われるとき，ウォーカーは，全ての頂点を正の割合で訪れることになる．

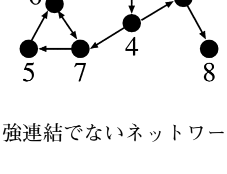

図 3.　強連結でないネットワークの例.

数式としては，ランダム･ウォークの推移確率行列 $B=(B_{ij})$ を，式 (2.4) の $B_{ij}=A_{ij}/\sum_{\ell=1}^{N}A_{i\ell}$ から

$$B_{ij}=q\mathbf{1}\boldsymbol{p}^{\top}+(1-q)\frac{A_{ij}}{\sum_{\ell=1}^{N}A_{i\ell}} \tag{2.11}$$

へと変更した上で，式 (2.3) を適用する．$\top$ は転置を表す．ここで，$\mathbf{1}$ は，全ての要素が 1 である N 次元列ベクトル (= 縦ベクトル) である．$\boldsymbol{p}=(p_1\,p_2\,\cdots\,p_N)^{\top}$ は，$\sum_{\ell=1}^{N}p_\ell=1$ を満たす列ベクトルである．したがって，$\mathbf{1}\boldsymbol{p}^{\top}$ は $N\times N$ 行列である．式 (2.11) の右辺第 1 項の (i,j) 要素は，qp_j である．これは，出発頂点 i に関わらず，確率 q で，元のネットワーク構造に従わずに確率 p_j で頂点 j に移動する，と読む．$\boldsymbol{p}$ は personalization vector あるいは teleportation vector と呼ばれる (Langville and Meyer, 2009; Franceschet, 2011)．

最も簡単な $\boldsymbol{p}$ は，$p_j=1/N\,(1\le j\le N)$ で与えられる (増田・今野，2010 ではこの場合のみを紹介した)．この場合，確率 q で，ウォーカーは全ウェブページのうちの等確率で選ばれた 1 つに飛ぶ．こうせずに，何らかの事前知識を用いて $\boldsymbol{p}$ を決めることもある．例えば，p_j を頂点 j の入次数に比例するように定めて，確率 q のテレポーテーションではハブに行きやすくするとよいという結果がある (Lambiotte and Rosvall, 2012)．

第 1 のぶら下がり頂点の問題に対応するためには，式 (2.11) をさらに改変する必要がある．ぶら下がり頂点 i に対しては，式 (2.11) の右辺第 2 項に現れる行列 $A_{ij}/\sum_{\ell=1}^{N}A_{i\ell}$ の第 i 行 (すなわち，$(i,1),\ldots,(i,N)$ 要素) が未定義である．これらを，第 $(i,1),\ldots,(i,N)$ 要素の和が 1 になるように改変する．頂点 i に到達したウォーカーの，次の時刻での行き先分布を決めることに対応する．

この行き先分布としては，2 つの決め方がよく使われる．1 つ目は，テレポーテーションの時と同様に，ベクトル $\boldsymbol{p}$ で行き先分布を決める方法である．テレポーテーションの時と同様に，$\boldsymbol{p}$ は，一様分布でも，何らかの事前知識に基づく非一様な分布でもよい．また，テレポーテーションで使われる $\boldsymbol{p}$ とぶら下がり頂点に対して使われる $\boldsymbol{p}$ は，異なっていてもよい．

2 つ目は，ぶら下がり頂点 i に到達したら，次の時刻でも i にとどまることである．$p_i=1$, $p_j=0\ (j\ne i)$ となるように $\boldsymbol{p}$ を選ぶことに相当する．この場合の $\boldsymbol{p}$ は，ぶら下がり頂点ごとに異なることになる．この 2 つ目の方式を用いる場合，確率 q でテレポーテーションが起こる場合にのみ，ランダム・ウォーカーはぶら下がり頂点から脱出できる．

ともかく，このように修正した B (これもなお，とあるランダム・ウォークの推移確率行列である) に対して式 (2.3) を適用することにより，やっと，全ての頂点の PageRank が求まる．PageRank の定義は，q の値，テレポーテーションの行き先分布 $\boldsymbol{p}$，ぶら下がり頂点到達時の扱い，の 3 箇所に任意性がある．$q>0$ である限り，強連結でないネットワークに対しても $x_i>0$ $(1\le i\le N)$ となる．

2.5 計算方法

修正前 (強連結である場合) にしろ修正後 (強連結でない場合) にしろ，実は，行列 B の最大固有値は 1 であり，特殊な場合を除けば，他の固有値の絶対値は 1 未満である．式 (2.3) を行列で書くと

$$(x_1\,x_2\,\cdots\,x_N)=(x_1\,x_2\,\cdots\,x_N)B \tag{2.12}$$

となるから，$(x_1\,x_2\,\cdots\,x_N)$ は B の最大固有値に対応する左固有ベクトルである．ネットワーク研究の文脈では，PageRank は固有ベクトル中心性などの名で呼ばれる中心性指標 (Katz, 1953; Bonacich, 1972) の仲間に属する．

この仲間に属する中心性指標に共通することとして，絶対値が最大の固有値の場合に使える

べき乗法という計算法を用いれば，$(x_1\,x_2\,\cdots\,x_N)$ を高速に求めることができる．式 (2.12) の連立一次方程式を真面目に解く必要はない．べき乗法では，適当な初期行ベクトル（＝横ベクトル）から始めて，B を右から何回も掛け算する．掛け算を施すたびに，得られる行ベクトルを規格化する．掛け算を繰り返すと，行ベクトルは $(x_1\,x_2\,\cdots\,x_N)$ に収束する．

q が大きいほど，べき乗法の収束が速く，数値計算誤差が少ない（Langville and Meyer, 2009 の 4.6 節）．しかし，q が大きいと，テレポーテーションが使われる確率が大きい．テレポーテーションは元のネットワークを無視しているので，q が大きいと PageRank が元のネットワークの構造を反映しなくなってしまう．$q=1$ という極端な場合には，$x_i=p_i\,(1\le i\le N)$ となる．ページの重要性を測るためには小さい q が望ましく，数値計算のためには大きい q が望ましい．折衷案として，経験上 $q=0.15$ がよく使われる．

3. IF（インパクトファクター）

IF は，PageRank の基準 1 のみに基づく順位付けである．基準 2 や 3 を無視していることが主な理由で，批判がなされているのである．

引用関係は，論文誌を頂点，論文誌 i が論文誌 j を引用することを i から j への有向枝とするネットワークと見なされる．このネットワーク構造に基づいて論文誌の順位付けを行う場合，順位付け方式を 1 つ決めることは，このネットワークにおける中心性指標を 1 つ決めることに等しい．

論文誌は論文の集まりであり，引用行動を行うのは，論文誌ではなく個々の論文である．したがって，i から j への枝の定義は，正確には，論文誌 i から出版された論文 I が，論文誌 j から出版された論文 J を引用することである．論文誌 i，j を固定しても，引用–被引用関係にある I と J の組は複数ありうる．なので，隣接行列 A は重み付きであると考える．すなわち，PageRank の場合と異なり，A_{ij} が 0 または 1 でない場合も許す．実際には，$A_{ij}\ge 0$ $(1\le i,j\le N)$ である限り，PageRank の定義はそのまま通用する．

A_{ij} の定め方の原則は，IF と EF に共通している．

$$A_{ij}=(\text{時間窓 } W_1 \text{ に論文誌 } i \text{ から発行された論文が, 時間窓 } W_2 \text{ に論文誌 } j \text{ から発行された論文を引用する回数}) \tag{3.1}$$

とする (図 4)．IF の計算においては，例えば 2012 年の IF を求めるなら，$W_1=2012$ 年の 1 年間，$W_2=2010\sim2011$ 年の 2 年間とする．すなわち，過去 2 年間に論文誌 j から発行された論文が，今年に論文誌 i から発行された論文に引用される回数を数える．

論文誌 i の IF は

$$\mathrm{IF}_i=\frac{\sum_{\ell=1}^{N}A_{\ell i}}{a_i} \tag{3.2}$$

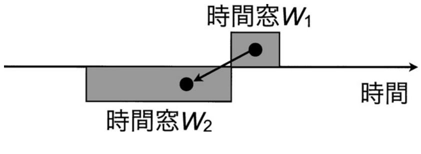

図 4．引用と被引用の時間窓．

で定義される．ここで，

$$a_i \equiv (\text{論文誌 } i \text{ が時間窓 } W_2 \text{ に発行した論文数}) \tag{3.3}$$

である．

式 (3.2) の分子は，時間窓 W_1 から時間窓 W_2 への引用の意味 (図 4) で，論文誌 i が受け取る引用の総数である．この数を分母の量で規格化することによって，論文誌 i から発行された論文 1 本の平均的な価値を定義したのが IF である．

IF はネットワーク的というほどの指標ではないが，引用関係ネットワークを用いた中心性指標の 1 つだと言える．そして，基準 1，すなわち，入次数 $\sum_{\ell=1}^{N} A_{\ell i}$ だけで得点を決めている．引用元の論文がどれだけ重要な論文誌に掲載されているか，という基準 2 は無視されている．被引用数が多い論文を読んでみて質の低さに失望することは多々ある．そのような論文を引用した論文のリスト (トムソン・ロイター社の Web of Science®ですぐに出力できる) を見てみると，なるほど，自分自身の引用や，身内 (特に国籍でわかる) 同士の引用で被引用数を稼いでいることがある．この場合，引用する側の論文は大した論文誌に載っていないことが常なので，こうした引用は，基準 2 に照らせば，貴重な引用ではない．また，IF は，引用元の論文が八方美人的にやたら多くの文献を引用して，そのうち 1 つの引用を得ただけかどうか，という基準 3 も無視している．

4.　EF (アイゲンファクター)

4.1　EF の定義

EF はほぼ PageRank なので，引用元の論文が発行された論文誌の重要性を考慮する (基準 2)．ただ，PageRank を論文誌の格付けに用いるという発想は，実は EF の新規性ではない (Pinski and Narin, 1976; eigenfactor.org, 2012; Franceschet, 2011)．EF の強みは，細かな数学的扱いや実装を詰めて実用的なランキングシステムを確立し，普及活動も熱心に行っていることなのかもしれない．EF の哲学はウェスト (2008) に詳しい．

EF と IF では，用いるネットワークが若干異なる．EF では，A_{ij} の定義に際して，引用する側 i の時間窓 W_1 が 2012 年の 1 年間のとき，引用される側 j の時間窓 W_2 を 2007∼2011 年の 5 年間とする．IF の場合の W_2 は，2010∼2011 年の 2 年間である．次に，EF では $A_{ii} \equiv 0$ と定義する．すなわち，同じ論文誌内の引用関係は数えない．IF では，A_{ii} を特に 0 にはしない．

ネットワークが定まった．このネットワークに対して PageRank を求める．テレポーテーションの実装の仕方はいくつかあるが (2.4 節)，personalization vector として $\boldsymbol{p} = (a_1\, a_2\, \cdots\, a_N)^\top$ を用いる．a_i は，その論文誌が 5 年間の時間窓 W_2 の間に発行した論文数の，その 5 年間に出た全ての論文誌の総論文数に占める割合である．すると，テレポーテーションの行き先が論文誌 i である確率は，i が発行する論文数に比例する．そして，ぶら下がり頂点に到達して確率 $1-q$ で通常のランダム・ウォークが起こる場合の処置も，確率 a_i で論文誌 i に行くことにする．ぶら下がり頂点に到達すれば，確率 q か $1-q$ かに関わらず，同じテレポーテーションが起こる．最後に，PageRank と同じく $q = 0.15$ とする．

このようにして定めた $B = (B_{ij})$，

$$B_{ij} = \begin{cases} q\mathbf{1}(a_1\, a_2\, \cdots\, a_N) + (1-q)\dfrac{A_{ij}}{\sum_{\ell=1}^{N} A_{i\ell}} & (\sum_{\ell=1}^{N} A_{i\ell} \geq 1), \\ \mathbf{1}(a_1\, a_2\, \cdots\, a_N) & (\sum_{\ell=1}^{N} A_{i\ell} = 0), \end{cases} \tag{4.1}$$

に対して，式 (2.3) を解いて，各論文誌の PageRank $x_i\,(1 \leq i \leq N)$ を求める．PageRank の一種として完成である．

しかし，EF は x_i を少し変更した指標である．まず，各 i に対して

$$\frac{\sum_{j=1}^{N} \frac{A_{ji}}{\sum_{\ell=1}^{N} A_{j\ell}} x_j}{\sum_{\ell'=1}^{N} \sum_{j=1}^{N} \frac{A_{j\ell'}}{\sum_{\ell=1}^{N} A_{j\ell}} x_j} \tag{4.2}$$

を求める．分子の $A_{ji}/\sum_{\ell=1}^{N} A_{j\ell}$ は，テレポーテーションなどを実装する前のランダム・ウォークで，ウォーカーが j から i に一歩で移動する確率である．j がぶら下がり頂点の場合は，$A_{ji}/\sum_{\ell=1}^{N} A_{j\ell}$ の分母が 0 で値が不定になってしまうので，そのような j については，$A_{ji}/\sum_{\ell=1}^{N} A_{j\ell}$ の代わりに 0 とする（このことは，eigenfactor.org の methods ページにある公式 pdf 文書には細かく書かれていないが，このサイトで公開されている Mathematica のプログラムを読むとわかる）．すると，式 (4.2) の分子は，正確にはランダム・ウォークの一歩になっていない．なぜなら，ぶら下がり頂点である j から一歩進んだウォーカーはどこかに消えてしまうからである．この最後の操作は，テレポーテーションの影響を減らす効果がある（開発者の 1 人である Martin Rosvall 氏から教えてもらった）．ともかくも，分母で規格化することによって，式 (4.2) の値の $1 \le i \le N$ に渡る和が 1 になるようにしている．

最後に，全論文誌の合計値が 100 になるように式 (4.2) を 100 倍したものが，論文誌 i の EF である．

4.2 Article Influence（AI）

質が同じなら，発行する論文数に比例して，論文誌 i の EF は大きくなる．A_{ji} $(j \neq i)$ は，論文誌 j の論文が論文誌 i の論文を引用した総数なので，i が発行する論文数に比例するからである．したがって，発行する論文数が大きいという意味で大規模な論文誌ほど，大きい EF 値を持ちやすい．EF は論文誌 i の総価値を表す，と言われる．

一方，小規模でも論文一本当たりの価値が高いこともまた，論文誌の価値である．そこで，eigenfactor.org は，EF の他に Article Influence™（以下 AI）も提供している．論文誌 i の AI は，EF を 0.01 倍して，さらに，式 (3.3) で与えられる a_i（$W_2 = 5$ 年間であることに再度言及しておく）で割った値である．a_i で割るので，論文 1 本当たりの価値の指標になっている．100 で割ったので，全ての論文誌の AI の和が約 1 になる．細かいことだが，AI の和はちょうど 1 ではない．各論文の価値の和，つまり，雑誌 i の AI を a_i で重みづけした上での $1 \le i \le N$ に渡る和，が 1 になる．

IF も，論文 1 本当たりの価値を表す指標なので（式 (3.2)），IF は，EF でなく AI と比べられるべきである．IF も AI も，個別の論文の価値を表すわけではないことに注意する．あくまで，1 つの論文誌に含まれる様々な論文の平均値である．IF も AI も，個別の論文でなく論文誌 1 つに対して 1 つ定まる値である．EF は，論文誌 i が発行する論文数に渡って和をとった量なので，論文誌 i が時間窓 W_2 の間に発行した全ての論文の被引用数の和と比べられるべきである．指標の分類を表 2 に示す．

4.3 Pinski and Narin（1976）との関係

Pinski and Narin（1976）は，PageRank や EF を遥かに先取って論文誌の評価指標を提案した（eigenfactor.org, 2012; Franceschet, 2011）．Pinski と Narin の定義では，式 (2.3) を論文誌の評価指標として用いる．ただし，B として，式 (2.4) の代わりに

$$B_{ji} = \frac{A_{ji}}{\sum_{\ell=1}^{N} A_{i\ell}} \tag{4.3}$$

を用いる．式 (2.4) と式 (4.3) は，右辺の分母が異なる．

表 2. 論文誌評価指標の分類. IF: Impact Factor, EF: Eigenfactor, AI: Article Influence.

測っている価値	基準 1 のみ	基準 1〜3
論文誌の総価値	被引用数の総計	EF
論文 1 本当たりの価値	IF	AI

$$y_i \equiv x_i \sum_{\ell=1}^{N} A_{i\ell} \tag{4.4}$$

と置くと, 式 (2.3), (4.3) より

$$y_i = \sum_{j=1}^{N} \frac{A_{ji}}{\sum_{\ell=1}^{N} A_{j\ell}} y_j \tag{4.5}$$

となる. 式 (4.5) は, 式 (2.3) に式 (4.3) ではなく式 (2.4) を代入したものと同じ形である. したがって, y_i は, ネットワークが強連結である場合に使える PageRank の原型(すなわち, 式 (2.3), (2.4) で定義される値) である. Pinski and Narin (1976) の指標は, y_i を $\sum_{\ell=1}^{N} A_{i\ell}$ で割って (式 (4.4)), 論文誌 i の論文が他の論文を引用する回数に対して罰を課している. 1 つの論文が引用する論文数平均が論文誌 i に依存しないならば, $\sum_{\ell=1}^{N} A_{i\ell}$ は, 論文誌が発行する論文数 a_i に比例する. したがって, Pinski and Narin の指標は, 本質的に AI と同じになる. 実際には, 1 つの論文が引用する論文数は分野ごとにかなり異なるので, Pinski and Narin の指標は AI に等しくない.

5. IF, EF, AI の比較

2011 年の IF, EF, AI の上位 20 誌は, それぞれ, 表 3, 4, 5 となる. 各表には各誌の発行論文数も付した.

IF の上位には総説誌が多く含まれる. Nature, Science といういわゆる二大有名誌は, それぞれ 11 位と 18 位である. EF では, Nature と Science はそれぞれ 1 位と 3 位だ. 2 位の PNAS も, 科学の広い分野で有名である. EF の上位 20 位以内には, 総説誌が 1 つも入っていない (Physical Review Letters は, Review という単語を名前に含むが, 総説誌ではない).

これだけを見ると, IF より EF の方が「良い」ように見えなくもないが, 注意を要する. EF は雑誌の規模に比例する. 2〜8 位の雑誌は, Nature と Science を除けば, 表 4 の中で発行論文数が多い部類に入る. Nature と Science 以外の 2〜8 位の論文誌が 2〜8 位である部分的な理由は, それらが大規模だからである.

表 4 の中で, 大規模であることの恩恵を最も受けているのは, 12 位の PLOS ONE だ. PLOS ONE は, それまでの雑誌と異なる方針で 2006 年に創刊された. 内容が正しければ, その論文の新規性や価値 (それは, しばしば査読者の主観に左右される) は問わずに受理することになっている. 論文の価値は, 発行後 (オープン・アクセスなので, 誰でも読める) のダウンロード回数や被引用回数などで事後的に定まる, という考え方だと私は理解している. 研究者の間では, PLOS ONE は玉石混合というのが通説だ. そして, 発行論文数がとても多い. PLOS ONE が一流誌だと考える研究者はまずいないが, EF では 12 位に入る.

むしろ, 表 4 から読み取るべきことは, Nature や Science, あるいは 9 位の New England Journal of Medicine や 10 位の Cell が, 発行論文数が少ない割に EF の上位に食い込んでいることである.

表 3. 2011 年の IF 上位 20 誌. Journal Citation Reports のデータを用いた. 論文誌名は正式名が長くない場合は正式名で表示し, 長い場合は Journal Citation Reports の略記を用いた. 表内で略記されている論文誌の正式名は, 17: Energy Education Science and Technology, 20: JAMA-Journal of the American Medical Association. 論文数は 2011 年の値. 本表, および以降の表の数値は, トムソン・ロイター社の許可を得て掲載.

順位	論文誌	IF	論文数
1	CA: A Cancer Journal for Clinicians	101.780	19
2	New England Journal of Medicine	53.298	349
3	Annual Review of Immunology	52.761	23
4	Review of Modern Physics	43.933	38
5	Chemical Reviews	40.197	196
6	Nature Reviews Molecular Cell Biology	39.123	66
7	Lancet	38.278	276
8	Nature Reviews Genetics	38.075	71
9	Nature Reviews Cancer	37.545	68
10	Advances in Physics	37.000	9
11	Nature	36.280	841
12	Nature Genetics	35.532	196
13	Annual Review of Biochemistry	34.317	41
14	Nature Reviews Immunology	33.287	69
15	Nature Materials	32.841	134
16	Cell	32.403	338
17	ENERGY EDUC SCI TECH	31.677	174
18	Science	31.201	871
19	Nature Reviews Neuroscience	30.445	47
20	JAMA-J AM MED ASSOC	30.026	220

AI (表 5) では, IF の場合と同様に, 多くの総説誌が上位に並ぶ. IF と AI を比べてどちらが「自然か」については, 筆者は何とも判断できない.

6. 議論

本節では, IF, EF, AI についてしばしば誤解される点を中心に, いくつかの事項について議論する.

6.1 小規模な論文誌は有利か?

式 (3.2) の分母を小さくすれば IF は大きくなるので, 規模が小さい (= 年ごとの発行論文数が少ない) 論文誌ほど IF で有利である, としばしば言われる. 実際, 表 3 には, 発行論文数が少ない論文誌が多い. しかし, これは論理的には正しくない.

表 3 に小規模な雑誌が多く現れる本当の理由は 2 つある. 1 つ目は, 総説論文誌の場合だ. 表 3 には総説誌が多く (IF 1 位の CA 誌も, 名前からはわかりにくいが, 基本的に総説誌だ), その 2011 年の発行論文数は高々71 であり少ない. 総説は分野を概観するしばしば濃厚な解説なので, 分野 1 つについていくつもの論文が短期間に出るわけではなく, 通常の論文誌と異なる. なので, 総説論文の数は限られる. そして, 総説は, 他の論文から引用されやすい. しかし, このことをもってして, 総説誌は「分母が小さいから」IF が大きい, とは言えない.

Nature Reviews シリーズには, 表 6 に示すように分野ごとに 15 の総説誌がある. それぞれの総説誌の発行論文数は少なめで 47~71 である (2011 年). 仮に, これらが 1 つの Nature

表 4. 2011 年の EF 上位 20 誌. 正式名は, 2: Proceedings of the National Academy of Sciences of the United States of America, 11: Angewandte Chemie International Edition.

順位	論文誌	EF	論文数
1	Nature	1.65524	841
2	P NATL ACAD SCI USA	1.60168	3614
3	Science	1.41162	871
4	Physical Review Letters	1.14457	3229
5	Journal of the American Chemical Society	0.81730	3176
6	Physical Review B	0.75604	6121
7	Journal of Biological Chemistry	0.74213	4382
8	Applied Physics Letters	0.67575	4419
9	New England Journal of Medicine	0.66383	349
10	Cell	0.66082	338
11	ANGEW CHEM INT EDIT	0.51421	2002
12	PLOS ONE	0.50162	13781
13	Journal of Neuroscience	0.44963	1790
14	Astrophysical Journal	0.42962	2472
15	Blood	0.42794	1568
16	Journal of Clinical Oncology	0.39463	729
17	Journal of Immunology	0.36989	1438
18	Cancer Research	0.36273	749
19	Lancet	0.36095	276
20	Nano Letters	0.34591	955

表 5. 2011 年の AI 上位 20 誌.

順位	論文誌	AI	論文数
1	Review of Modern Physics	28.900	38
2	CA: A Cancer Journal for Clinicians	24.536	19
3	Nature Reviews Molecular Cell Biology	23.861	66
4	Annual Review of Immunology	23.427	23
5	New England Journal of Medicine	21.304	349
6	Cell	20.554	338
7	Nature	20.373	841
8	Annual Review of Biochemistry	19.755	41
9	Advances in Physics	17.985	9
10	Nature Reviews Cancer	17.949	68
11	Nature Materials	17.925	134
12	Nature Genetics	17.584	196
13	Science	17.525	871
14	Annual Review of Astronomy and Astrophysics	17.157	13
15	Nature Reviews Genetics	16.961	71
16	Nature Reviews Immunology	16.820	69
17	Annual Review of Neuroscience	16.169	24
18	Nature Photonics	16.130	96
19	Nature Reviews Neuroscience	16.117	47
20	Cell Stem Cell	15.981	96

表 6. Nature Reviews シリーズ. IF，論文数は 2011 年の値.

論文誌	IF	論文数
Nature Reviews Molecular Cell Biology	39.123	66
Nature Reviews Genetics	38.075	71
Nature Reviews Cancer	37.545	68
Nature Reviews Immunology	33.287	69
Nature Reviews Neuroscience	30.445	47
Nature Reviews Drug Discovery	29.008	61
Nature Reviews Microbiology	21.182	70
Nature Reviews Neurology	12.461	55
Nature Reviews Clinical Oncology	11.963	65
Nature Reviews Encocrinology	9.971	60
Nature Reviews Cardiology	8.833	53
Nature Reviews Rheumatology	8.388	66
Nature Reviews Gastroenterology & Hepatology	8.102	53
Nature Reviews Nephrology	7.092	62
Nature Reviews Urology	4.415	61
（仮）Nature Reviews	21.656	927

Reviews という総説誌に統合されたとする．すると，論文数は全部で 927 本となり，IF の分母が小さくはなくなる．それでも，そのようにして統合された（仮）Nature Reviews 誌は高い IF 値 21.656 を持つ（表 6）．1 本の論文が平均的にたくさん引用されるから，IF が大きいのだ．

実際には，そのような仮想的な総説誌は存在せず，分野ごとに総説誌が出ている．そして，分野ごとならば，1 年に出る総説論文数は限られている．なので，分母が結果的に小さい．総説誌の IF が高いのは，総説論文は一般的によく引用されるからである．

2 つ目の理由は，単に受理基準が厳しい，という理由である．総説誌でない論文誌 i が，ある年に 1000 本の論文を発行しているとする．これを，受理した論文のうち，仮想的に「でたらめに選んだ」50 本を発行して残りの 950 本は受理したはずなのに実際には発行しなかったとする（ひどい話だ！）．すると，式 (3.2) の分母は小さくなるが，IF は細かな変動を除けば変わらない．論文 1 本ごとの平均被引用回数はこの操作の前後で変わらないからである．だから，単に分母を小さくしても，論文誌 i の IF は大きくならない．

1000 本の中からよりすぐりの 50 本だけを発行するならば，論文誌 i の IF は上昇する．そして，式 (3.2) の分母は小さくなる．しかし，この場合の IF 上昇は，発行本数を減らしたからというよりも，受理基準を厳しく絞り込んだからと解釈するのが適切だろう．論文誌 i の編集委員会が，論文誌 i をよりよくするために，査読基準を今までより厳しくして，程度の低い（と彼らが思う）論文は受理しない方針を打ち出すのは，至って健全な行為だ．この行為をもってして，「分母を小さくして IF を高めるための操作」とは私は思わない．

まとめると，分母が小さい論文誌は，確かに大きい IF を持つ傾向がある．ただ，総説論文誌である，または，受理基準が厳しい，ということが，分母を小さくすることと IF を高めることの両方に同時に寄与している．したがって，分母を小さくした「から」IF が大きくなったかのように見える．実際には，分母をどのようにして小さくするかが肝要だ．

なお，分母が小さい論文誌ほど値が（結果的に）大きくなりやすいことは，AI にも共通して言える（表 5）．

6.2 総説誌が多くの論文を引用することの意味

IF は総説誌の価値を過大評価している，と言われることがある．また，総説誌は，多くの論文を引用することが普通だ．しかし，この 2 つの間に直接の因果関係はない．

IF が EF や AI に比べて過大評価しているのは，総説誌の論文が行う 1 つ 1 つの引用の価値である．総説論文は他の多くの論文を引用しやすいので，PageRank の基準 3 によると，個々の引用の価値は低いと見なされるべきである．しかし，IF は，そのようにせず，1 つの総説論文について 100 を超えることも珍しくない 1 つ 1 つの引用が，10 本しか他の論文を引用しない論文の 1 引用と同じ価値だと見なす．

総説論文が多くの論文を引用することは，IF でも EF（や AI）でも，その総説誌の価値を下げることにはならない．あくまで，総説誌の価値は，含まれる総説論文の被引用数（基準 1），引用してくれる相手の価値（基準 2），引用してくれる相手が多すぎる引用をしていないか（基準 3）で決まる．PageRank は出次数（引用数に対応）とは特に関係しない（表 1）ことを思い出そう．

6.3 分野によって論文発行の速さが異なること

IF 値は，分野によって大きく異なる．

数学の論文誌は，概して IF が低い．数学では，論文 1 本あたりの引用論文が少ない．また，そもそも，論文数が他の多くの分野よりも少ない．数学の論文はやはり数学の論文から最も引用されやすいだろうから，数学の論文が引用される回数は，どうしても少なくなる．それに加えて，査読が概して遅い．最初の査読結果を受け取るために 1 年待たされることはざらである．なので，IF の時間窓である直近 2 年間のうちに，数学の論文は他の論文から引用されにくい．一方，生命科学では，論文の引用文献数が多く，論文数も多く，査読や発行も速い．すると，IF の意味では数学よりも生命科学の方が圧倒的に有利だ．

EF，AI の時間窓は 5 年なので，分野間の偏りを是正する効果があると言われる（eigenfactor.org, 2012）．しかし，これは時間窓の長さの問題である．IF の設計指針と EF（または AI）の設計指針のどちらがより妥当か，という議論とはあまり関係ない．

Journal Citation Reports は，5 年 IF という指標も毎年発表している．これは，近年 5 年間の IF の平均値とかいう意味ではない．引用数を数える時間窓を 5 年にしたという意味である．つまり，時間窓 W_1 が 2012 年なら，時間窓 W_2 を 2007∼2011 年にする．もちろん，a_i もこの 5 年間に発行された論文数とする．5 年 IF の時間窓と EF の時間窓は同じである．2 年ではまだ引用されにくく，5 年程度で引用が増えてくるような分野に光を当てることは，IF でもできるのだ．逆に，EF を 2 年の時間窓で定義し直すこともやればできる．

したがって，査読速度，論文刊行速度の分野間格差に関する限りは，IF と EF の違いは，設計の本質的な違いではなく時間窓の違いから来ている．

6.4 EF の問題点

EF の問題点として，総被引用数と強い相関があることが指摘されている（Davis, 2008; Fersht, 2009; 佐藤, 2009）．論文誌 i の総被引用数とは $\sum_{\ell=1}^{N} A_{\ell i}$ のことである．したがって，同じ論文誌内の引用 A_{ii} を含めるか否かという細かいことを除けば，これは，EF は基準 1∼3 で設計されているにも関わらず，実際には基準 1 だけで大体すんでしまう，という身も蓋もない指摘である．

同じことは，論文 1 本あたりの価値の指標についてもなされている．AI は IF と比べられる指標である（表 2）．AI は，その基準 1 だけからなる簡単版と言える IF と強く相関するという観察結果がある（中西, 2011）．実際に，表 3 と表 5 は結構似ている．

ウェブの PageRank についても，基準 1 に相当する入次数と基準 1∼3 を考慮する PageRank が強く相関しているという証拠がいくつかある（Fortunato et al., 2008; Ghoshal and Barabási,

2011)．ただし，逆の結果も存在する (Donato et al., 2004)．

もちろん，総被引用数と EF，IF と AI は，相関が強くても同じではない．今後のさらなる比較，考察が必要だろう．

6.5 個々の論文の評価指標

論文誌だけでなく個々の論文を同様の手法で評価しよう，という発想は自然である．基準 1 だけで済ませることは，自分の論文の被引用数を数えることに相当する．この量は，Google Scholar やトムソン・ロイター社の Web of Science で容易に入手できることも理由で，研究者の間でもよく話題になる．日本でも海外でも，公募や研究費助成の申請書に，自分の論文の被引用数を書くことはしばしばある．被引用数は基準 2 と 3 を無視した指標なので，あまり価値のない論文に引用されて被引用数を伸ばしている場合などを排除できない．なので，PageRank 化，EF 化しようという発想である．

ところが，論文の評価指標を PageRank 化することは，論文誌の場合よりも難しい．論文は自分より古い論文しか引用できない，という時間の一方向性があるからである．論文のネットワークは強連結でないどころか，強連結から程遠く，引用可能な方向が 1 つの時間軸で強く規定されている．論文誌 i が論文誌 j を引用し逆の引用もあることは普通だ．しかし，個々の論文の場合には，論文 i が論文 j を引用し，逆の引用もあるという状況は基本的には起こらない．

ただ，この時間軸を考慮して PageRank 風に論文の価値を定量化する研究はある (Walker et al., 2007)．時間軸に沿ってイベント (ここでは，論文 i が論文 j を引用すること) が起こり，その総体がネットワークを成す，というテンポラル・ネットワーク (Holme and Saramäki, 2012; 増田, 2012b) の観点も，今後の指標提案に役立つかもしれない．

7. おわりに

細かな数理的な面を含めて，IF と EF，その違いなどを解説した．評価指標の良し悪しはある程度主観的な議論であり，1 つの基準に絞りこむことはできないだろう (Bollen et al., 2009; Fersht, 2009)．とはいえ，IF が猛威をふるい，EF という新しい指標が Journal Citation Reports にも掲載され始めた昨今である．各指標について正しく理解を深め，異なる指標について議論を行うことは，IF や EF が用いられているどの研究分野にとっても有益であろう．本稿がその一助となれば幸いである．

謝　辞

原稿を読んでコメントを下さった高口太朗氏 (東京大学) に御礼申し上げる．本原稿は科研費 (23681033) の助成を受けたものである．

参　考　文　献

Baldi, P., Frasconi, P. and Smyth, P. (2007)．『確率モデルによる Web データ解析法 —データマイニング技法から e–コマースまで』(水田正弘，南弘征，小宮由里子 訳)，森北出版，東京.

Bollen, J., Van de Sompel, H., Hagberg, A. and Chute, R. (2009)． A principal component analysis of 39 scientific impact measures, *PLOS ONE*, **4**, e6022.

Bonacich, P. (1972)． Factoring and weighting approaches to status scores and clique identification, *Journal of Mathematical Sociology*, **2**, 113–120.

Brin, S. and Page, L. (1998)． Anatomy of a large-scale hypertextual web search engine, *Proceedings* *of the 7th International World Wide Web Conference*, 107–117.

Chen, J., Aronow, B. J. and Jegga, A. G. (2009). Disease candidate gene identification and prioritization using protein interaction networks, *BMC Bioinformatics*, **10**, 73.

Davis, P. M. (2008). Eigenfactor: Does the principle of repeated improvement result in better estimates than raw citation counts?, *Journal of the American Society for Information Science and Technology*, **59**, 2186–2188.

Donato, D., Laura, L., Leonardi, S. and Millozzi, S. (2004). Large scale properties of the Webgraph, *European Physical Journal B*, **38**, 239–243.

eigenfactor.org (2012). Eigenfactor, http://www.eigenfactor.org/.

Fersht, A. (2009). The most influential journals: Impact Factor and Eigenfactor, *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 6883–6884.

Fortunato, S., Boguñá, M., Flammini, A. and Menczer, F. (2008). Approximating PageRank from in-degree, *Lecture Notes in Computer Science*, **4936**, 59–71.

Franceschet, M. (2011). PageRank: Standing on the shoulders of giants, *Communications of the ACM*, **54**, 92–101.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas, *Science*, **122**, 108–111.

Ghoshal, G. and Barabási, A.-L. (2011). Ranking stability and super-stable nodes in complex networks, *Nature Communications*, **2**, 394.

Gross, P. L. K. and Gross, E. M. (1927). College libraries and chemical education, *Science*, **66**, 385–389.

Holme, P. and Saramäki, J. (2012). Temporal networks, *Physics Reports*, **519**, 97–125.

伊理正夫(2003).『一般線形代数』, 岩波書店, 東京.

Iván, G. and Grolmusz, V. (2011). When the Web meets the cell: Using personalized PageRank for analyzing protein interaction networks, *Bioinformatics*, **27**, 405–407.

Katz, L. (1953). A new status index derived from sociometric analysis, *Psychometrika*, **18**, 39–43.

木村英紀(2003).『線形代数—数理科学の基礎』, 東京大学出版会, 東京.

Lambiotte, R. and Rosvall, M. (2012). Ranking and clustering of nodes in networks with smart teleportation, *Physical Review E*, **85**, 056107.

Langville, A. N. and Meyer, C. D. (2009).『Google PageRank の数理—最強検索エンジンのランキング手法を求めて』(岩野和生, 黒川利明, 黒川洋 訳), 共立出版, 東京.

増田直紀(2010). 検索エンジンのネットワーク科学, 数学セミナー, 7 月号, 94–95.

増田直紀(2012a).『なぜ 3 人いると噂が広まるのか』, 日経プレミアシリーズ, 日本経済新聞出版社, 東京.

増田直紀(2012b). テンポラルネットワーク, 人工知能学会誌, **27**(4), 432–436.

増田直紀, 今野紀雄(2010).『複雑ネットワーク—基礎から応用まで』, 近代科学社, 東京.

中西印刷株式会社(2009). Eigenfactor™(アイゲンファクター)の数学的基礎, http://www.nacos.com/pdf/abouteigen.pdf.

中西印刷株式会社(2011). Eigenfactor™ (アイゲンファクター)とは, http://www.nacos.com/pdf/eigenfactor.pdf.

Pinski, G. and Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to literature of physics, *Information Processing & Management*, **12**, 297–312.

Radicchi, F. (2011). Who is the best player ever? A complex network analysis of the history of professional tennis, *PLOS ONE*, **6**, e17249.

齋藤正彦(1966).『線型代数入門』, 東京大学出版会, 東京.

佐藤翔(2009). 学術情報をめぐる新たな評価指標: Impact Factor, *h*-index, Eigenfactor, Article Influence, Usage Factor, 薬学図書館，**54**(2), 121–132.

角田裕之，小野寺夏生(2007). 論文と研究者のインパクトに対する新しい計算書誌学的指標，情報メディア研究，**5**(1), 1–20.

Walker, D., Xie, H., Yan, K.-K. and Maslov, S. (2007). Ranking scientific publications using a model of network traffic, *Journal of Statistical Mechanics*, P06010.

ウェスト，ジェヴィン(2008). アイゲンファクター(Eigenfactor)：学術情報のランキングおよびマッピング，第6回 SPARC Japan セミナー 2008「IF を超えて — さらなる研究評価の在り方を考える」，講演要旨，http://www.nii.ac.jp/sparc/event/2008/pdf/112508/document/Jevin_West_document_jp.pdf，http://www.nii.ac.jp/sparc/event/2008/20081125.html.

Wu, J., Wang, S. and Pan, D. (2011). Evaluation of technological influence power of enterprises through the enterprise citation network, *International Journal Knowledge and Systems Science*, **2**, 32–42.

Introduction to Eigenfactor

Naoki Masuda

Graduate School of Information Science and Technology, The University of Tokyo

A widely known index called Impact Factor quantifies the value of individual academic journals. Many people working in academy related fields refer to Impact Factor, although its plausibility is often disputed. In 2007, Eigenfactor$^{\mathrm{TM}}$, which takes advantage of the structure of a citation network among journals, was proposed. Since 2009, Eigenfactor$^{\mathrm{TM}}$ has been included in Journal Citation Reports by Thomson Reuters, which issues Impact Factor. Those who refer to these indices should be aware of their correct meanings and implications. In this article, I explain how Eigenfactor$^{\mathrm{TM}}$ works, in particular in comparison to Impact Factor. In fact, Eigenfactor$^{\mathrm{TM}}$ is essentially the same as PageRank, which Google uses for its search engine to rank websites. Therefore, I start with an extended explanation of PageRank, which is presumably more popular than Eigenfactor$^{\mathrm{TM}}$, to guide readers' understanding of Impact Factor and Eigenfactor$^{\mathrm{TM}}$. I also discuss some possible misunderstandings regarding these different indices.

Key words: Directed graph, network, PageRank, Impact Factor, centrality, bibliometrics.