Collective attention and ranking methods

Gabrielle Demange
Paris School of Economics

September 13, 2010

- Increasing importance of rankings
    - on Internet to rank pages, or to provide an index of trust in e-commerce,
      a search engine determines a ranking through various criteria, clicks, hyperlink structure (PageRank on Google)
    - in the academic world (universities, journals, researchers)

  rankings based on 'experts' statements': citations, links, votes
- Justification: aggregates relevant information $\rightarrow$ helps decision making and saves on search cost (public good aspect)
- When there are many alternatives and the situation is recurrent, a ranking alters 'collective' attention
  What is the long run impact on behaviors and statements ?
  Convergence ?

Related literature

- influence of opinions channelled by 'neighbors' in a network (e.g. Goyal's book 05, deMarzo et al. 03, Golub-Jackson 08)
  Here information is public. Differential impact on statements because experts differ in their preferences.
- concerns about the influence of search engines among computer scientists.
  Main criticism : biased towards already popular webpages
  Proposals to introduce some randomness or to account of the date of creation of a page (Cho et al. 2005, Pandey et al. 05).
- axiomatization of methods (Palacios-Huerta Volij 04 Slutzki Volij 06 Altman Tennenholtz 06)
  Here I introduce a new method and provide an axiomatization

Outline

1. Ranking problems, Ranking methods, properties
2. The handicap-based method
3. Dynamics under the influence of rankings

   Contrasting results for two classes:

   - the 'generalized handicap methods'
   - the 'peers' methods'

Ranking problems

- $N = \{1, ..., n\}$ be a set of n 'individuals' to rank
  $M = \{1, ..., m\}$ be a set of m 'experts'
- $n \times m$ matrix $\Pi = (\pi_{ij})$, $\pi_{ij} \geq 0$
  j's column= j's statements on N
  assume first Π strictly positive
- ranking vector $r = (r_i) \geq 0, \sum r_i = 1$
  r_i = score of i (cardinal up to a factor)
- A RANKING METHOD ASSIGNS A RANKING VECTOR r, TO EACH Π
- In a journal ranking problem : π_{ij}= average number of references of an article from j to articles in i.
  Web, (perturbed) incidence matrix
  In an apportionment problem, $N = \{parties\}$, $M = \{regions\}$,
  π_{ij}= number of votes from region j to party i.

Some properties

- axioms
  intensity-invariance
  uniformity and exactness
  homogeneity
- property of supporting weights

Invariance

- let $[\Pi]$ be obtained from Π by scaling columns' sums to 1

  F is intensity invariant if

  $$F(\Pi) = F([\Pi]) \text{ each } \Pi$$

  important property when statements - number of references per article, number of links - are not controlled
- the intensity invariant version G of F:

  $$G(\Pi) = F([\Pi]) \text{ each } \Pi.$$

Uniformity, exactness

- Benchmark : balanced matrices.

$$\sum_j \pi_{ij} = a \text{ each } i \quad \sum_i \pi_{ij} = b \text{ each } j \ (na = mb)$$

  Each row receives the same total as well as each column

- A method is uniform if it assigns equal scores to balanced matrices.
  A method is exact if the reverse is true for normalized matrices: $F([\Pi]) = (\frac{1}{n})$ implies that $[\Pi]$ is balanced.

Homogeneity

- A method is homogeneous if multiplying row i by $\lambda > 0$ multiplies its rank relative to other rows by the same λ.
- Not preserved by factoring out intensities
  ex: The counting method
  scores (r_i) proportional to totals $(\sum_{j \in M} \pi_{i,j})$
  is homogeneous but not the intensity invariant version

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \text{ equal scores } \begin{pmatrix} 4 & 2 \\ 1 & 2 \end{pmatrix} \text{ scores prop to } (13, 7)$$

Two examples : the invariant and the Hits methods

$N = M$

- The invariant method

$$r_i = \sum_{j \in N} [\pi_{i,j}] r_j \text{ for each } i.$$

r principal eigenvector of $[\Pi]$ (its largest eigenvalue $= 1$)

- The Hits method (Kleinberg 99) distinguishes between the ability as an individual (authorities) as an expert (hubs) define r and q such that

$$r_i = \sum_j \pi_{i,j} q_j \text{ each } i \text{ and } q_j = \lambda \sum_i \pi_{i,j} r_i \text{ each } j$$

r eigenvector of $\Pi\tilde{\Pi}$

- I 'll qualify the invariant method as a peers' method but not the Hits one

Equilibrium supporting weights

- Method F is supported by weights $Q^F = (Q_j^F) \geq 0$ if

$$r_i \quad = \sum_j [\pi_{i,j}] q_j \text{ each } i$$

$$\text{where } r \quad = F(\Pi) \; q = Q^F(\Pi).$$

equilibrium relationships
– counting method: $Q^F = 1/m$,
– invariant method: $F_i(\Pi) = Q_i^F(\Pi)$ each i.
– Hits method $Q^F(\Pi) \propto \tilde{\Pi} F(\Pi)$

- Useful:
  – to define new methods through equilibrium relationships
  – to transform a method by adjusting the weights
  – to give a precise meaning to what a 'peers' method means
  – to study dynamics

The handicap-based method

Handicaps and scores may be seen as inversely related

- Given Π, there is a unique ranking vector $r = (r_i)$ such that

$$1 = \frac{1}{r_i} \sum_j \pi_{i,j} q_j \ \text{ each } i$$

$$1 = \frac{m}{n} q_j \sum_i \frac{\pi_{i,j}}{r_i} \ \text{ each } j$$

  handicaps equalize weighted counts
  experts' weights equalize the distributed handicaps

- The handicap-based method H assigns the unique ranking r.
  It is supported by weights.
- equivalent to $P = (\pi_{i,j} \frac{q_j}{r_i})$ is a balanced matrix
  related to scaling matrix problems (e.g. Balinski-Demange 89)

Characterization of the handicap-based method

- **Theorem** *H* IS THE ONLY METHOD THAT IS INTENSITY INVARIANT, HOMOGENEOUS, AND EXACT.
- generalized handicap-based methods: adjust the weights

$$G_i(\Pi) = \sum_{j \in M} [\pi_{i,j}] \frac{q_j^{1-\gamma}}{\sum_{k \in M} q_k^{1-\gamma}} \text{ for each } i \in N \text{ where } q = Q^H(\Pi).$$

  family indexed by γ:
  $\gamma = 0$: handicap-based, $\gamma = 1$ counting method

Influence model

- Experts' statements depend
(1) on preferences (2) on attention intensities
- Interpret Π as the 'true' preferences:
$\pi_{i,j}$ =proba for j to state a positive vote on i when j evaluates each i with equal attention.
- Attention intensities (b_i): b_i represents the intensity spent on assessing i in N. Results in statements proportional to $(b_i \pi_{i,j})$
- Influence of rankings : (b_i) proportional to $B(r_i)$
start with 'linear' influence : $B(x) = x$ then $B(x) = x^\alpha$.

$$\rightarrow \pi_{i,j}^{(t)} = r_i^{(t)} \pi_{i,j}$$

$$r^{(t+1)} = F(\Pi^{(t)})$$

Rest points

$$\Pi^{(t)} = dg(r^{(t)})\Pi$$

where $dg(r)$= diagonal matrix with r on the diagonal

- For F supported by weights, write $\mathbf{q}_j(r) = Q_j^F(dg(r)\Pi)$

$$r_i^{(t+1)} = \sum_j \frac{\pi_{i,j} r_i^{(t)}}{\sum_{\ell \in N} \pi_{\ell,j} r_\ell^{(t)}} \mathbf{q}_j(r^{(t)}) \text{ each } i.$$

  Self-enforcing mechanism: $r_i = 0$ possible for a fixed point.
  Stability ? Need some continuity assumptions

- A rest point

$$\sum_j \frac{\pi_{i,j}}{\sum_{\ell \in N} \pi_{\ell,j} r_\ell^*} \mathbf{q}_j(r^*) \leq 1 \text{ each } i \text{ with } = \text{ if } r_i^* > 0$$

  Necessary conditions for r^* to be stable for the dynamics.

Continuity

- Beware: in general, F and Q cannot be extended by continuity over all matrices ≥ 0
  ex: a non-negative matrix may admit multiple positive eigenvectors
- Continuity assumptions
  (a) F and Q continuous over the set of positive matrices.
  (b) $\mathbf{q}(r^*) = \lim_{\rho \to r^*} \mathbf{q}(\rho)$ well defined for any ranking $r^* \geq 0$
  where $\mathbf{q}_j(\rho) = Q_j^F(dg(\rho)\Pi)$

  satisfied by all current methods

Convergence for generalized handicap-based methods

- Proposition Consider a generalized handicap-based method with γ strictly positive. Under the linear influence dynamics, there is a unique rest point, which is furthermore globally stable: the rankings converge to it for any initial value of r.
  Lyapounov function
- applies to the counting method
  does not apply for the handicap-based method, but generically a unique stable rest point
- Convergence also under influence function with diminishing returns r^{α}, $\alpha < 1$

Peers' method: definition

- Let $N = M$. Minimal requirement: an individual who receives a small score is also assigned a small expert's weight.
  F supported by Q is a PEERS' METHOD if there is a positive k such that $Q_i(\Pi) < kF_i(\Pi)$ for each Π.
- invariant method is a peers' method, the counting and the Hits methods are not,
- In (Demange 09) I analyze the influence of the invariant method and compare with a search mechanism. Some results extend to ANY peers' method

Peers' methods

- Proposition Consider a peers' method. Given Π, subset I of N is the support of a rest point if and only if

  $$\text{there is } x \text{ in } \Re^I, x \gg 0,\ \Pi_{I\times I}x = \mathbb{1}_I,\ \Pi_{N-I\times I}x \leq \mathbb{1}_{N-I}$$

  Characterization independent of the peers' method
  conditions on citations from I towards I and towards $N - I$
- Proposition Consider a peers' method. There are matrices Π for which the dynamics admit several locally stable points.

  strong self-enforcing mechanism

Concluding remarks

- Rankings induce a coordination on attention
  The interplay between preferences and the ranking method results in a variety of different outcomes
  Self enforcing mechanism for peers' methods
- Analyze the support of the rest points
  Consider several rankings ?