Spectral Methods for Data Science

AP

An equaliser lecture

Motivations (encore)

Activity tables show how users *map* their choices or, viceversa, how available products *map* onto their adopters.

	Matrix	Alien	Star Wars	Casablanca	Titanic
Joe	1	1	1	0	0
Jim	3	3	3	0	0
John	4	4	4	0	0
Jack	5	5	5	0	0
Jill	0	0	0	4	4
Jenny	0	0	0	5	5
Jane	0	0	0	2	2

Figure 11.6: Ratings of movies by users

Essentially, a weighted, binary relationship between users and films…

Spectral Methods

What is it?

Most activity matrices represent the connections between n entities, e.g., users and m entities, such as films, $n <> m$.

Sometimes the connections is between the same entities, such as endorsement, teams defeating other teams, friends or followers on social networks etc.

In such cases, the matrix is square.

Hence, standard Geometry holds and we can extracts the **Eigenpairs**.

Eigenpairs

Matrix A has a real λ and a vector $\mathbf{e}$ s.t.

$$A\mathbf{e} = \lambda\mathbf{e}$$

λ is an *eigenvalue* and $\mathbf{e}$ an *eigenvector* of A.

. . .

If A has rank n, then there could be up to n eigenpairs. In practice,

- they might not be real, nor $\neq 0$, and
- are always *costly* (at least quadratic time in the size of the m., $\Omega(n^2)$) to find.

Interesting square matrices

A is called *symmetric* when $A = A^T$

Also called *positive semidefinite* when for any $\mathbf{x}$ we have

$$\mathbf{x}^T A\mathbf{x} \geq 0$$

In such case its eigenvalues are non-negative: $\lambda_i \geq 0$.

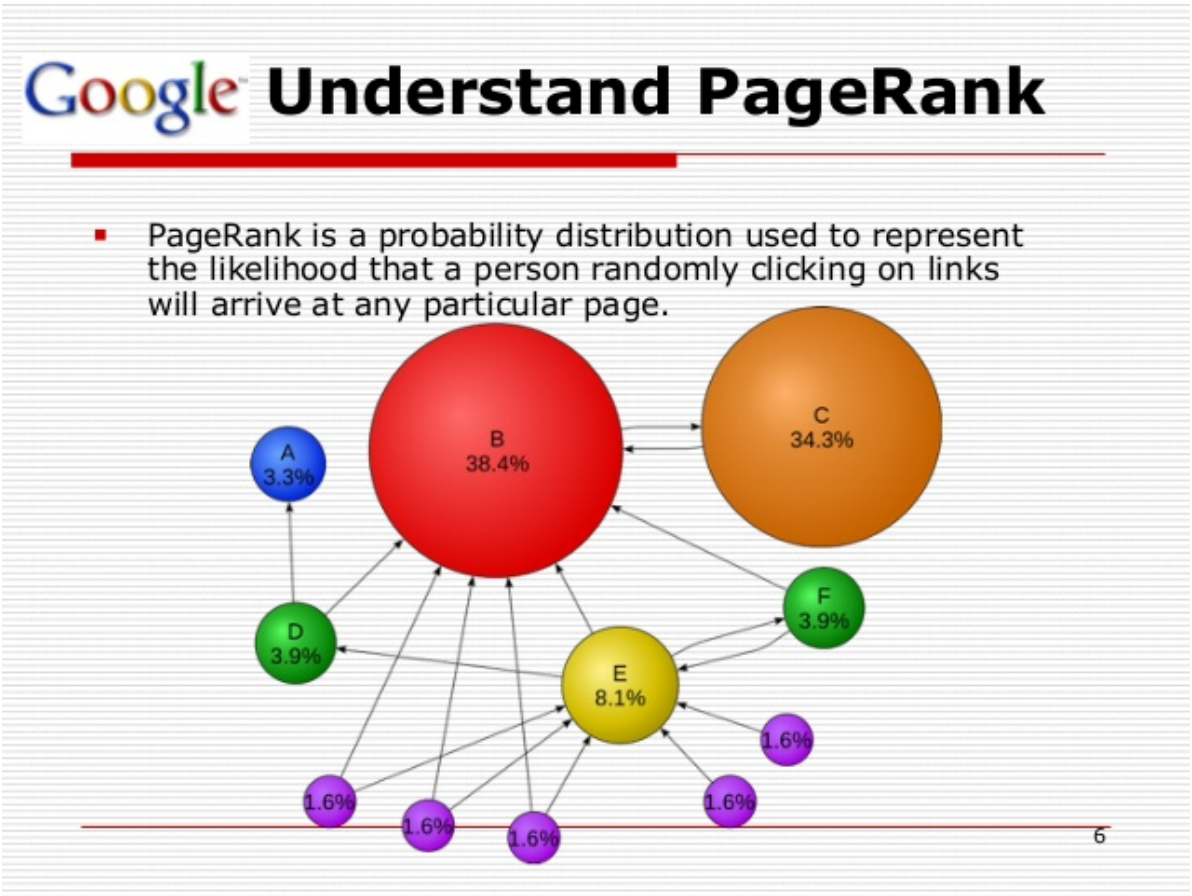
Applications of Spectral analysis

Spectral properties

Adjacency matrices represent connections between entities in a network (graph), e,g., the Web.

The eigenvalues of adjacency matrices provide bounds for several network features.

The Google PageRank algorithm *is* spectral network analysis.



Early applications in Psychology, Social science, Bibliometrics, Economy, and Choice theory (seriously).

Spectral ranking

Given a matrix representing preference or likeability between people, can we rank the participants (from best to worst) on the basis of their general, intrinsic likeability?

. . .

[Seely, 1949] created an index of likeability based on the ideas of *diffusion:* it is important to be liked by people who in turn are well-liked and so on.

Let M be a square matrix where m_{ij} represents *approval* or *endorsement* (negative values represent *disapproval)*

. . .

my *likeability index* should be equal to the weighted sum of of the indices of the people who like me.

---

my *likeability index* should be equal to the weighted sum of of the indices of the people who like me.

But their likeability is turn will depend on mine…

Let's use row vectors $\mathbf{r} = [r_1, r_2, \ldots r_n]$:

$$\mathbf{r} = \mathbf{r}M$$

i.e., $\mathbf{r}$ is a left eigenvector of M.

This formula might have no solution, but matrix preprocessing can assure that one exists.

Study plan

Background study

Ian Goodfellow, Yoshua Bengio and Aaron Courville: Deep Learning, MIT Press, 2016.

available in HTML and PDF from the module; it is *a refresher* of notation and properties: no examples and no exercises. It can be read in the background of our classes.

- Phase 1: read §§ 2.1—2.7, then § 2.11.
- Phase 2: read §§ 2.8—2.10