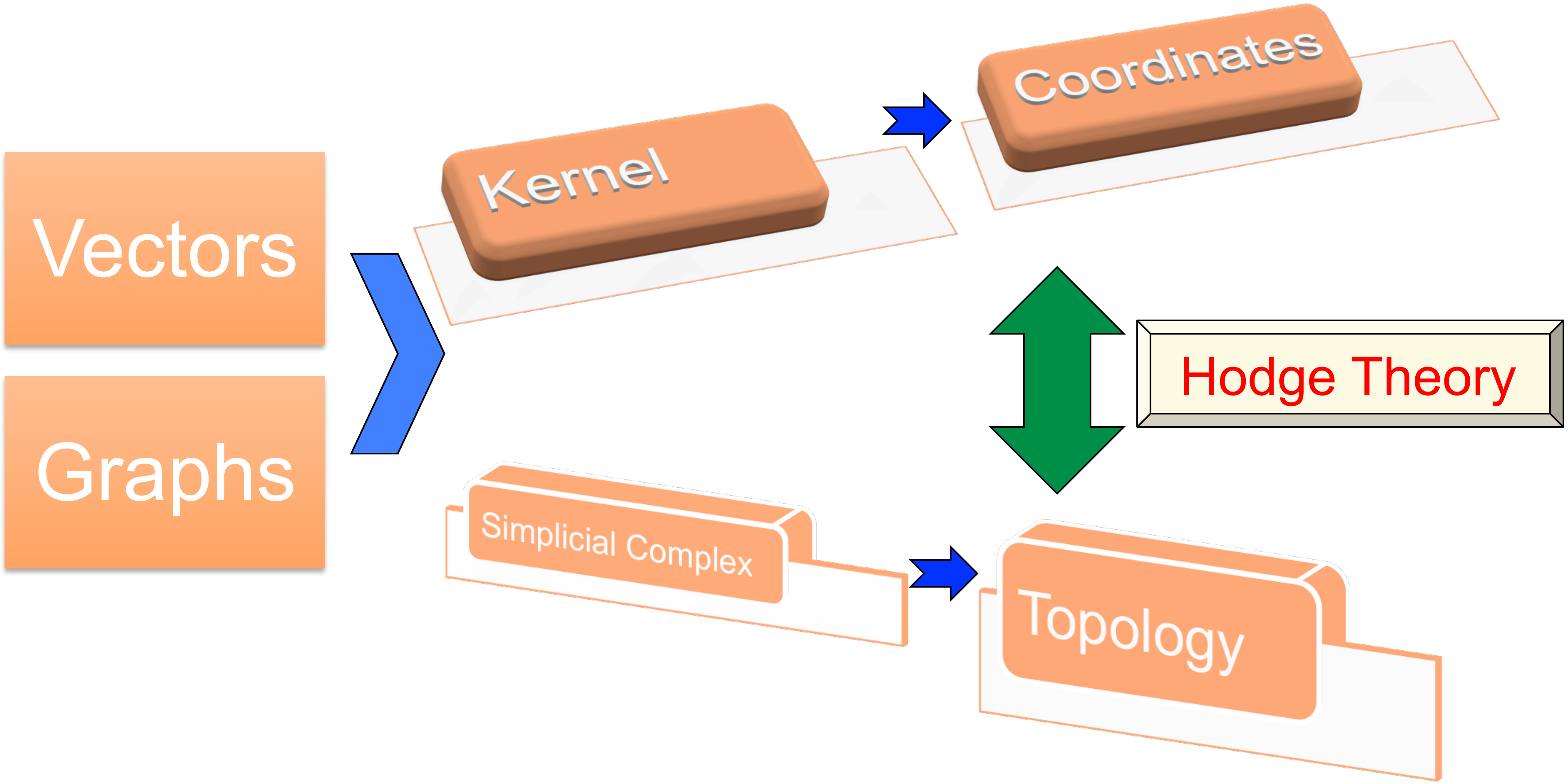
Mathematics of Data
IV
Combinatorial Hodge Theory with Applications

姚 远
北京大学数学科学学院

2011.7.14

Recall



Vectors
Graphs
Kernel
Coordinates
Hodge Theory
Simplicial Complex
Topology

Hodge Decomposition

- **Vector calculus:** Helmholtz's decomposition, ie. vector fields on *nice domains* may be resolved into irrotational (curl-free) and solenoidal (divergence-free) component vector fields

$$\mathbf{F} = -\nabla\varphi + \nabla \times \mathbf{A}$$

  φ scalar potential, $\mathbf{A}$ vector potential.

- **Linear algebra:** additive orthogonal decomposition of a skew-symmetric matrix into three skew-symmetric matrices

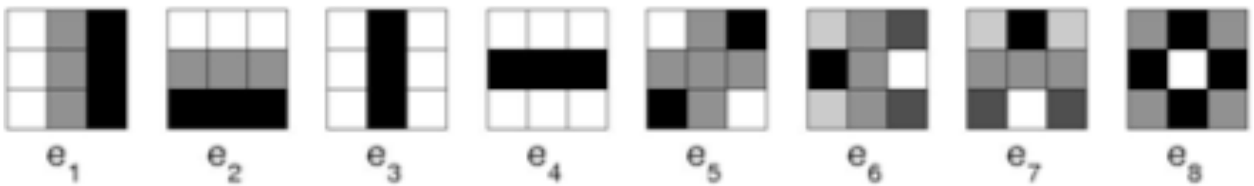
$$W = W_1 + W_2 + W_3$$

  $W_1 = \mathbf{v}\mathbf{e}^T - \mathbf{e}\mathbf{v}^T$, W_2 clique-consistent, W_3 inconsistent.

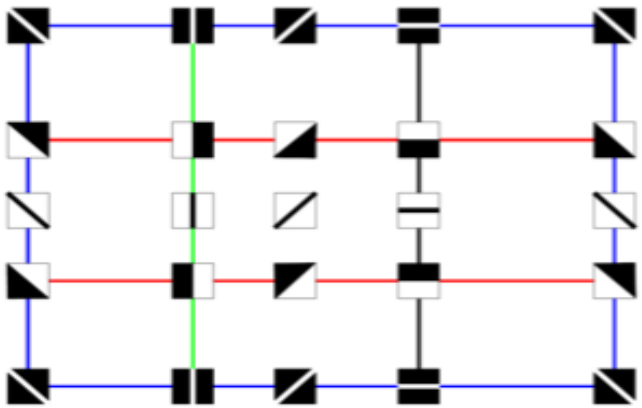
- **Graph theory:** orthogonal decomposition of network flows into acyclic and cyclic components.

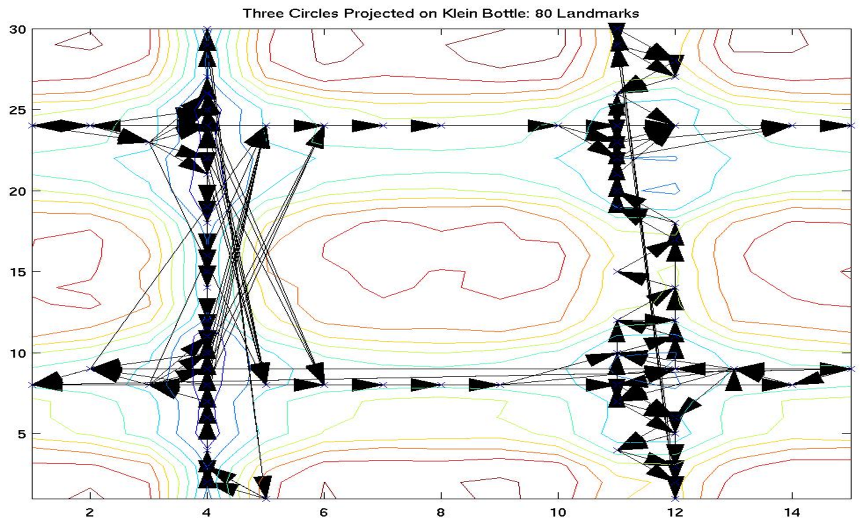
Example I: Visual Image Patches - Point Cloud Data in Metric Spaces

- Ann Lee, Kim Pedersen, David Mumford (2003) studies statistical properties of 3x3 high contrast image patches of natural images (from Van Heteran's database)
- Gunnar Carlsson, Vin de Silva, Tigran Ishkhanov, Afra Zomorodian (2004-present) found those image patches concentrate around a 2-dimensional klein bottle imbedded in 7-sphere
- They build up *simplicial complex* from *point cloud data*
- *1-D Harmonic flows* actually focus on densest region -- 3 major circles



1-D Harmonic Flows on the space of 3x3 Image Patches





- Left Upper：Klein Bottle of 3x3 Image Patch Space (Courtesy of Carlsson-Ishkhanov, 2007)
- Left Lower：Harmonic flows focus on 3 major circles where most of data concentrate

Some new theory: Bartholdi-Schick-Smale-Smale, 2010, “Hodge Theory on Metric Spaces”, preprint.

Here we will focus on "Ranking", or "Preference Aggregation" …

Psychology: L. L. Thurstone (1928) (scaling), et al.

Statistics: M. Kendall (1930s, rank corellation), F. Mosteller, Bradley-Terry,…, P. Diaconis (group theory), et al.

Economics: Condorcet (1785), Borda, K. Arrow, A. Sen (voting and social choice theory, Nobel Laureates) et al.

Computer Science: Google's PageRank, Recommendation in E-commerce, et al.

Had William Hodge met Maurice Kendall

Hodge Theory

Pairwise ranking

The Bridge of Sighs in
Cambridge, St John's College

Thurstone's Crime Scaling in 1928

TABLE 1.

TABLE OF PROPORTIONS p_{k_a} OR $p_{k>a}$

	Abortion	Adultery	Arson	Assault and battery	Bootlegging	Burglary	Counterfeiting	Embezzlement	Forgery	Homicide	Kidnapping	Larceny	Libel	Perjury	Rape	Receiving stolen goods	Seduction	Smuggling	Vagrancy
Abortion	. . . .	.323	.338	.211	.128	.238	.244	.245	.212	.760	.318	.222	.191	.256	.822	.143	.419	.174	.045
Adultery	.677	. . . .	.415	.242	.172	.281	.285	.253	.274	.863	.365	.207	.182	.245	.925	.143	.589	.204	.034
Arson	.662	.585	. . . .	.260	.136	.226	.321	.348	.254	.017	.563	.215	.144	.349	.944	.140	.716	.170	.019
Assault and battery	.789	.757	.740	. . . .	.379	.515	.556	.485	.534	.070	.743	.385	.385	.587	.947	.344	.785	.346	.072
Bootlegging	.872	.828	.864	.621	. . . .	.764	.745	.738	.754	.955	.924	.678	.506	.728	.985	.527	.871	.576	.116
Burglary	.762	.719	.774	.485	.236	. . . .	.593	.605	.580	.981	.856	.333	.322	.478	.981	.221	.769	.284	.027
Counterfeiting	.756	.715	.679	.444	.255	.407	. . . .	.540	.488	.947	.804	.303	.284	.532	.963	.199	.756	.215	.042
Embezzlement	.755	.747	.652	.515	.262	.395	.460	. . . .	.350	.958	.752	.305	.248	.474	.977	.141	.774	.251	.049
Forgery	.788	.726	.746	.466	.246	.420	.512	.650	. . . .	.951	.819	.343	.320	.534	.966	.195	.820	.260	.035
Homicide	.240	.137	.083	.030	.045	.019	.053	.042	.049	. . . .	.083	.030	.034	.079	.441	.027	.181	.026	.011
Kidnapping	.682	.635	.437	.257	.076	.144	.196	.248	.181	.917	. . . .	.170	.106	.288	.902	.098	.595	.086	.026
Larceny	.778	.793	.785	.615	.322	.667	.697	.695	.657	.970	.830	. . . .	.348	.648	.970	.268	.848	.365	.053
Libel	.809	.818	.855	.615	.494	.678	.716	.752	.680	.966	.894	.652	. . . .	.702	.981	.530	.886	.456	.067
Perjury	.744	.755	.651	.413	.272	.522	.467	.526	.466	.921	.712	.352	.298	. . . .	.951	.204	.767	.222	.015
Rape	.178	.075	.056	.053	.015	.019	.037	.023	.034	.559	.098	.030	.019	.049	. . . .	.019	.076	.023	.015
Receiv'g stolen goods	.857	.857	.860	.656	.473	.779	.801	.859	.805	.973	.902	.732	.470	.796	.981	. . . .	.875	.525	.061
Seduction	.581	.411	.284	.215	.129	.231	.244	.226	.180	.819	.405	.152	.114	.233	.924	.125	. . . .	.121	.023
Smuggling	.826	.796	.830	.654	.424	.716	.785	.749	.740	.974	.914	.635	.544	.778	.977	.475	.879	. . . .	.037
Vagrancy	.955	.966	.981	.928	.884	.973	.958	.951	.965	.989	.974	.947	.933	.985	.985	.939	.977	.963	. . . .

- ➢Can we learn a scale for crimes from pairwise comparisons?
- ➢Similar modern problem: wine taste, video quality evaluation, …

Ranking in Economics

STRATEGIES	B Cooperate	B Defect
A Cooperate	(3,3)	(0,5)
A Defect	(5,0)	(1,1)

Prisoner's dilemma in Game Theory, (Flood-Dresher-Tucker 1950)

Voter 1	Voter 2	Voter 3	Voter 4
A>B>C	B>C>A	C>A>B	…

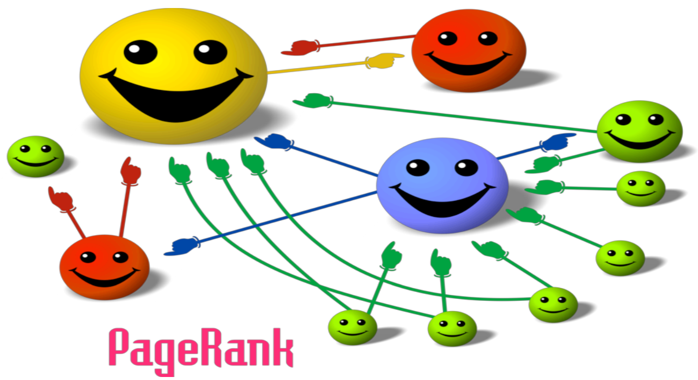
Voting theory and social choice

- Condorcet (1785), Borda (1700s)
- Kenneth Arrow (1972 Nobel Memorial Prize in Economics)
- Amartya Sen (1998 Nobel Memorial Prize in Economics)

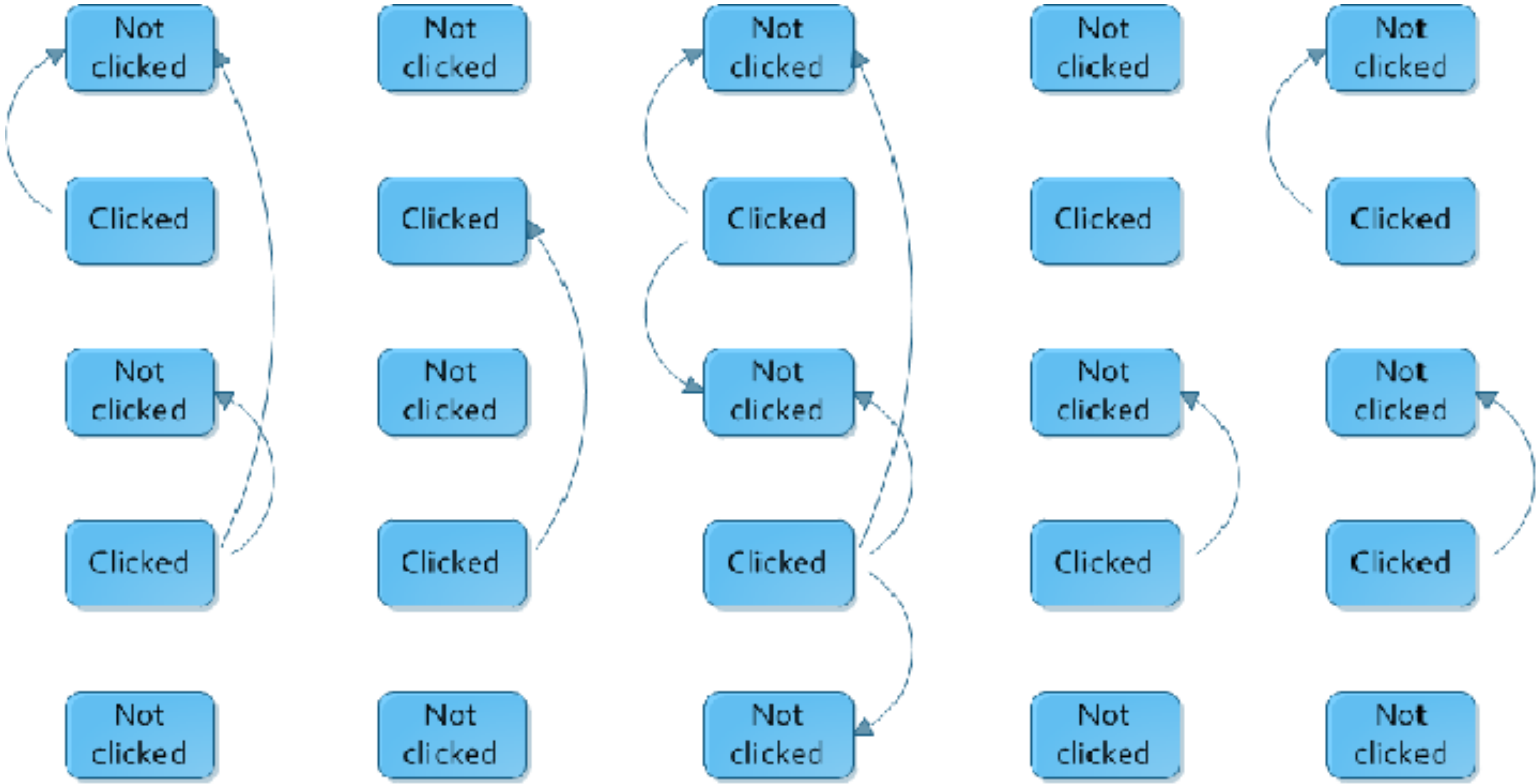
Ranking on Internet

- "Multicriteria" ranking/decision systems
  - Amazon or Netflix's recommendation system (user-product)
  - Interest ranking in social networks (person-interest)
  - S&P index (time-price)
  - Voting (voter-candidate)
- "Peer-review" systems
  - Publication citation systems (paper-paper)
  - Google's webpage ranking (web-web)
  - eBay's reputation system (customer-customer)

	mv1	mv2	mv3
usr1	1	-	-
usr2	2	5	-
usr3	-	-	4
usr4	3	2	5



Clicks implies preference



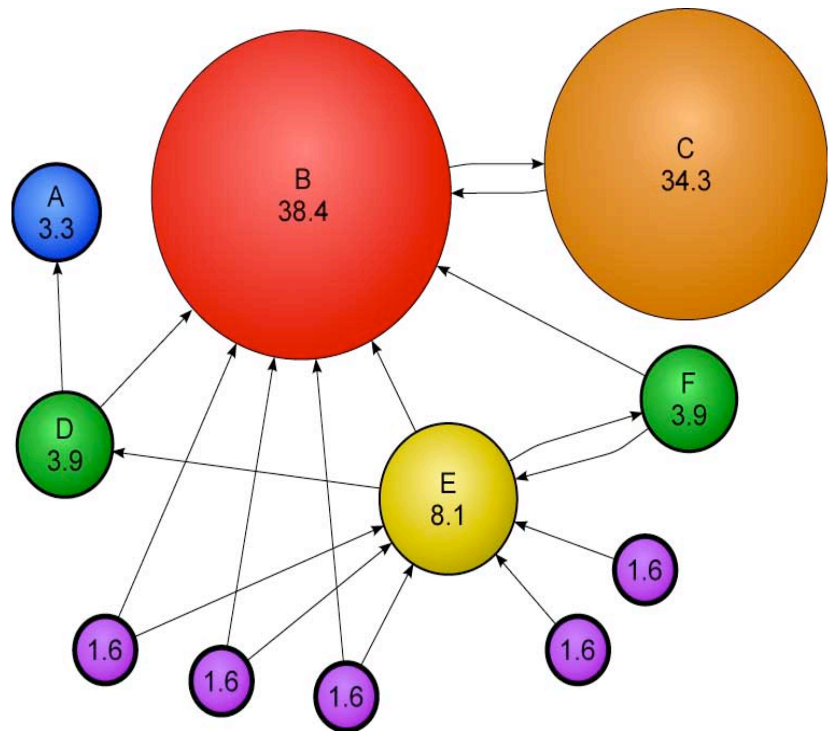
Characteristics

- Aforementioned ranking data are often
  - Incomplete: typically about 1%
  - Imbalanced: heterogeneously distributed votes
  - Cardinal: given in terms of scores or stochastic choice
- Pairwise ranking on graphs: implicitly or explicitly, ranking data may be viewed to live on a simple graph G=(V,E), where
  - V: set of alternatives (webpages, products, etc.) to be ranked
  - E: pairs of alternatives comparable

Look at Hodge decomposition of pairwise comparison edge flows…

Example I: Pagerank

- Model assumption:
  - A Markov chain random walk on networks, subject to the link structure
- Algorithm [Brin-Page'98]
  - Choose Link matrix L, where L(i,j)=# links from i to j.
  - Markov matrix $M=D^{-1} L$, where $D = e^T L$, e is the all-one vector.
  - Random Surfer model: E is all-one matrix
  - PageRank model: $P = c M + (1-c) E/n$, where $c = 0.85$ chosen by Google.
  - Pagerank vector: the primary eigenvector v_0 such that $P^T v_0 = v_0$



Note: SVD decomposition of L gives HITS [Kleinberg'99] algorithm.

Problem: Can we drop Markov Chain model assumption?

Another View on Pagerank

Define pairwise ranking:

$$w_{ij} = \log \frac{P_{ij}}{P_{ji}} = -w_{ji}$$

Where P is the Pagerank Markov matrix.

Claim: if P is a reversible Markov chain, i.e.

$$\pi_i P_{ij} = \pi_j P_{ji}$$

Then

$$w_{ij} = \log \pi_i - \log \pi_j$$

Example II: Netflix Customer-Product Rating

Example (Netflix Customer-Product Rating)

- 480189-by-17770 customer-product rating matrix X
- X is **incomplete**: 98.82% of values missing

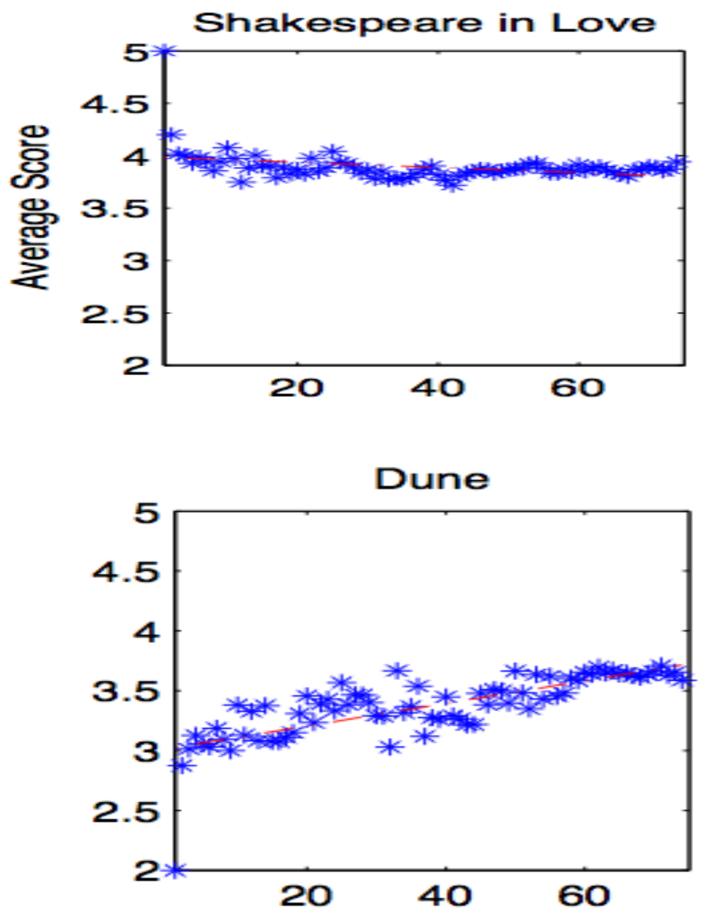
However,

- pairwise comparison graph $G = (V, E)$ is very **dense**!
- only 0.22% edges are missed, **almost a complete graph**
- rank aggregation may be carried out without estimating missing values
- **imbalanced**: number of raters on $e \in E$ varies

Caveat: we are not trying to solve the Netflix prize problem

Rank Aggregation

- The first order statistics, mean score for each product, is often inadequate because of the following:
  - most customers would rate just a very small portion of the products
  - different products might have different raters, whence mean scores involve noise due to arbitrary individual rating scales (right figure)



How about high order statistics?

From 1st order to 2nd order: Pairwise Ranking

- *Linear Model*: average score difference between product i and j over all customers who have rated both of them,

$$w_{ij} = \frac{\sum_k (X_{kj} - X_{ki})}{\#\{k : X_{ki}, X_{kj} \text{ exist}\}}.$$

  Invariant up to translation.

- *Log-linear Model*: when all the scores are positive, the logarithmic average score ratio,

$$w_{ij} = \frac{\sum_k (\log X_{kj} - \log X_{ki})}{\#\{k : X_{ki}, X_{kj} \text{ exist}\}}.$$

  Invariant up to a multiplicative constant.

Pairwise Ranking Continued

- *Linear Probability Model*: the **probability** that product j is preferred to i in excess of a purely random choice,

$$w_{ij} = \Pr\{k : X_{kj} > X_{ki}\} - \frac{1}{2}.$$

  Invariant up to monotone transformation.

- *Bradley-Terry Model*: logarithmic odd ratio (logit)

$$w_{ij} = \log \frac{\Pr\{k : X_{kj} > X_{ki}\}}{\Pr\{k : X_{kj} < X_{ki}\}}.$$

  Invariant up to monotone transformation.

Skew-Symmetric Matrices of Pairwise Ranking

Recall **skew-symmetric** matrices: $W \in \mathbb{R}^{n\times n}$, $W^T = -W$:

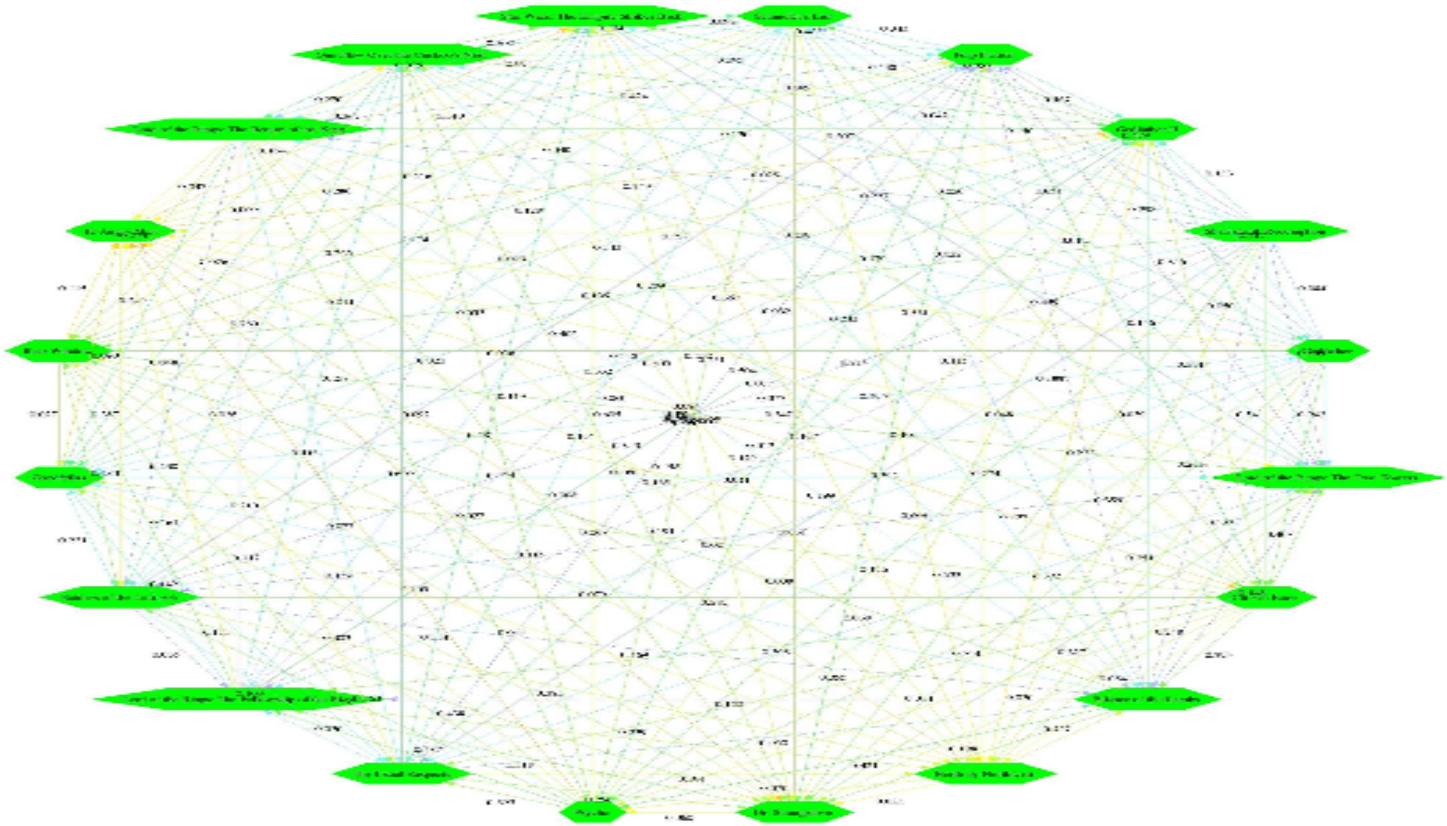
- every $A \in \mathbb{R}^{n\times n}$ decomposable into $A = S + W$, $S = (A + A^T)/2$ symmetric, $W = (A - A^T)/2$ skew-symmetric
- $\mathcal{W} = \{\text{skew-symmetric matrices}\} = \wedge^2(\mathbb{R}) = \mathfrak{o}_n(\mathbb{R})$

All previous models induce (sparse) skew-symmetric matrices of size $|V|$-by-$|V|$

$$w_{ij} = \begin{cases} -w_{ji} & \text{if } \{i,j\} \in E \\ ? & \text{otherwise} \end{cases}$$

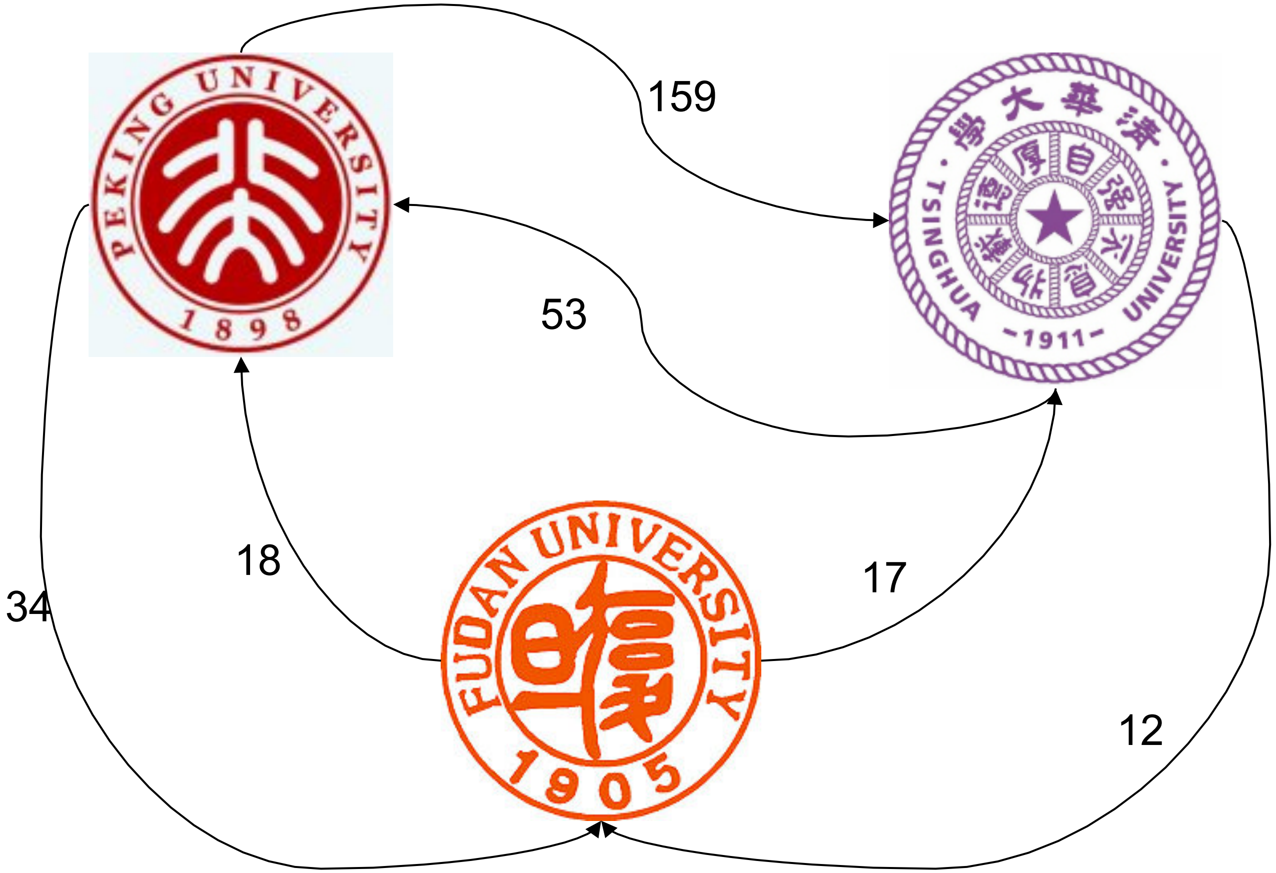
where $G = (V, E)$ is a pairwise comparison graph.
Note: such a skew-symmetric matrix induces a pairwise ranking **flow** on graph G.

Pairwise Ranking of Top 10 IMDB Movies



- Pairwise ranking graph flow among top 10 IMDB movies

Web Link among Chinese Universities



2002, http://cybermetrics.wlv.ac.uk/database/stats/data

Link structure correlated with Research Ranking?

Classical Ordinal Rank Aggregation Problem

- Problem: given a set of partial/total order $\{\succeq_i : i = 1, \ldots, n\}$ on a common set V, find

$$(\succeq_1, \ldots, \succeq_n) \mapsto \succeq^*,$$

as a partial order on V, satisfying certain *optimal* condition.

- Examples:
  - voting
  - Social Choice Theory

Rank Aggregation Problem

Difficulties:

- Arrow's impossibility theorem
- Kemeny-Snell optimal ordering is NP-hard to compute
- Harmonic analysis on $\mathfrak{S}_n$ is impractical for large n since $|\mathfrak{S}_n| = n!$

Our approach:

Problem

Does there exist a global ranking function, $v : V \to \mathbb{R}$, such that

$$w_{ij} = v_j - v_i =: \delta_0(v)(i,j)?$$

Equivalently, does there exists a scalar field $v : V \to \mathbb{R}$ whose gradient field gives the flow w? ie. is w **integrable**?

Utility theory for preference.

Answer: Not Always!

Multivariate calculus: there are non-integrable vector fields; cf. the film *A Beautiful Mind*:

$$A = \{F : \mathbb{R}^3 \setminus X \to \mathbb{R}^3 \mid F \text{ smooth}\}, \quad B = \{F = \nabla g\},$$
$$\dim(A/B) = ?$$

Similarly here,

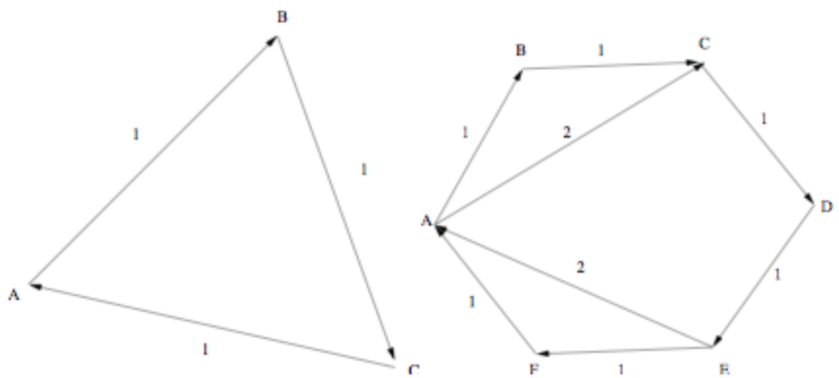


Figure: No global ranking v gives $w_{ij} = v_j - v_i$: (a) triangular cyclic, note $w_{AB} + w_{BC} + w_{CA} \neq 0$; (b) it contains a 4-node cyclic flow $A \to C \to D \to E \to A$, note on a 3-clique $\{A, B, C\}$ (also $\{A, E, F\}$), $w_{AB} + w_{BC} + w_{CA} = 0$

Triangular Transitivity

Fact

$W = [w_{ij}]$ *skew symmetric associated with graph* $G = (V, E)$. *If* $w_{ij} = v_j - v_i$ *for all* $\{i, j\} \in E$, *then* $w_{ij} + w_{jk} + w_{ki} = 0$ *for all 3-cliques* $\{i, j, k\}$.

Transitivity subspace:

$$\{W \text{ skew symmetric} \mid w_{ij} + w_{jk} + w_{ki} = 0 \text{ for all 3-cliques}\}$$

Example in the last slide, (a) lies outside; (b) lies in this subspace, but not a gradient flow.

Ordinal intransitivity: *a>b>c>a*
Cardinal intransitivity: $w_{ab}+w_{bc}+w_{ca} \neq 0$

Hodge Decomposition: Matrix Theoretic

A skew-symmetric matrix W associated with G can be decomposed uniquely

$$W = W_1 + W_2 + W_3$$

where

- W_1 satisfies
  - '*integrable*': $W_1(i,j) = v_j - v_i$ for some $v : V \to \mathbb{R}$.
- W_2 satisfies
  - '*curl free*': $W_2(i,j) + W_2(j,k) + W_2(k,i) = 0$ for all (i,j,k) 3-clique;
  - '*divergence free*': $\sum_{j:(i,j)\in E} W_2(i,j) = 0$
- $W_3 \perp W_1$ and $W_3 \perp W_2$.

Hodge Decomposition: Graph Theoretic

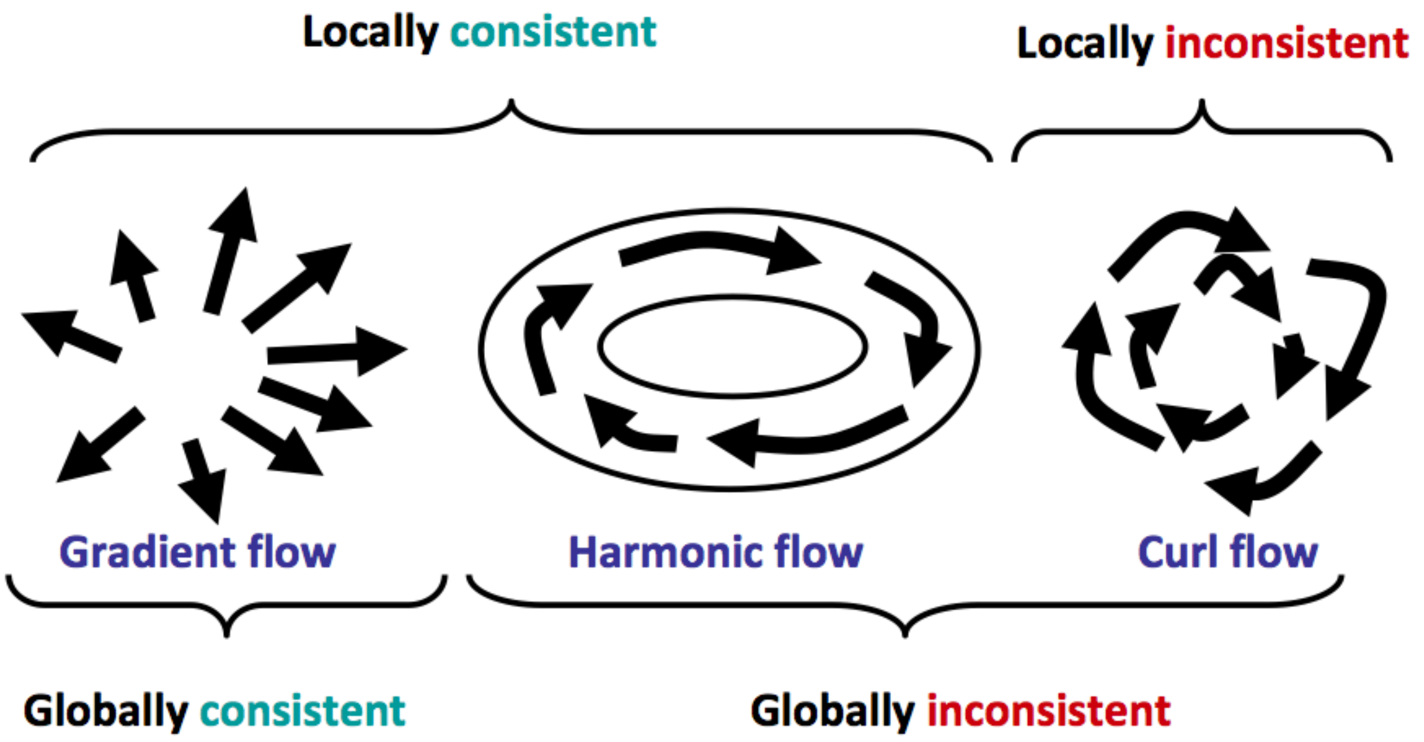
Orthogonal decomposition of network flows on G into

$$\text{gradient flow} + \text{globally cyclic} + \text{locally cyclic}$$

where the first two components make up transitive component and

- gradient flow is integrable to give a global ranking
- example (b) is locally (triangularly) acyclic, but cyclic on large scale
- example (a) is locally (triangularly) cyclic

1-D Hodge (Helmoholz) Decomposition Cartoon



Courtesy of Asu Ozdaglar, MIT

Combinatorial Hodge Theory: Geometric Analysis on Graphs (and Complexes)

图(复形)上的几何分析

Combinatorial Hodge Theory

- Build a simplicial complex (Graph as 1-skeleton) from data
  - Clique complex (Cech, Vietoris-Rips, Witness, …)
- Functions on simplicial complex as alternating forms
  - global ranking as 0-form on vertices
  - pairwise ranking as 1-form on edges
- Boundary/coboundary operators
  - gradient vs. divergence
  - curl vs. triangular boundary
- Combinatorial Laplacians and Hodge Decomposition

Clique Complex of a Graph

Extend graph G to a **simplicial complex** $\mathcal{K}(G)$ by attaching triangles

- 0-simplices $\mathcal{K}_0(G)$: V
- 1-simplices $\mathcal{K}_1(G)$: E
- 2-simplices $\mathcal{K}_2(G)$: triangles $\{i, j, k\}$ such that every edge is in E
- k-simplices $\mathcal{K}_k(G)$: $(k+1)$-cliques $\{i_0, \ldots, i_k\}$ of G

For ranking problems, suffices to construct $\mathcal{K}(G)$ up to dimension **2**!

- **global** ranking $v : V \to \mathbb{R}$, 0-**forms**, ie. vectors
- **pairwise** ranking $w(i, j) = -w(j, i)$ for $(i, j) \in E$, 1-**forms**, ie. skew-symmetric matrices

Discrete Differential Forms

- k-**forms**:

$$C^k(\mathcal{K}(G),\mathbb{R}) = \{u : \mathcal{K}_{k+1}(G) \to \mathbb{R}, u_{i_{\sigma(0)},\dots,i_{\sigma(k)}} = \mathrm{sign}(\sigma) u_{i_0,\dots,i_k}\}$$

  for $(i_0,\dots,i_k) \in \mathcal{K}_{k+1}(G)$, where $\sigma \in \mathfrak{S}_{k+1}$ is a permutation on $(0,\dots,k)$.
- May put metrics/inner products on $C^k(\mathcal{K}(G),\mathbb{R})$.
- The following metric on 1-forms, is useful for the imbalance issue

$$\langle w_{ij}, \omega_{ij} \rangle_D = \sum_{(i,j)\in E} D_{ij} w_{ij} \omega_{ij}$$

  where

$$D_{ij} = |\{\text{customers who rate both } i \text{ and } j\}|.$$

Discrete Exterior Derivatives: coboundary maps

- k-coboundary maps $\delta_k : C^k(\mathcal{K}(G), \mathbb{R}) \to C^{k+1}(\mathcal{K}(G), \mathbb{R})$ are defined as the **alternating difference** operator

$$(\delta_k u)(i_0, \ldots, i_{k+1}) = \sum_{j=0}^{k+1} (-1)^{j+1} u(i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_{k+1})$$

- δ_k plays the role of **differentiation**
- $\delta_{k+1} \circ \delta_k = 0$
- In particular,
  - $(\delta_0 v)(i,j) = v_j - v_i =: (\text{grad } v)(i,j)$
  - $(\delta_1 w)(i,j,k) = (\pm)(w_{ij} + w_{jk} + w_{ki}) =: (\text{curl } w)(i,j,k)$ (triangular-trace of skew-symmetric matrix $[w_{ij}]$)

Curl (旋度) and Divergence (散度)

For each triangle $\{i, j, k\}$, the **curl**

$$(\mathrm{curl}\, w)(i, j, k) = (\delta_1 w)(i, j, k) = w_{ij} + w_{jk} + w_{ki}$$

measures the total flow-sum along the loop $i \to j \to k \to i$.

- $(\delta_1 w)(i, j, k) = 0$ implies the flow is **path-independent**, which defines the **triangular transitivity subspace**.

For each alternative $i \in V$, the **divergence**

$$(\mathrm{div}\, w)(i) := -(\delta_0^T w)(i) := \sum w_{i*}$$

measures the **inflow-outflow sum** at i.

- $(\delta_0^T w)(i) = 0$ implies alternative i is preference-neutral in all pairwise comparisons.
- divergence-free flow $\delta_0^T w = 0$ is **cyclic**

Basic Alg. Top.: Boundary of Boundary is Empty

Fundamental tenet of topology: $\delta_{k+1} \circ \delta_k = 0$.
For $k = 0$,

$$C^0 \xrightarrow{\delta_0} C^1 \xrightarrow{\delta_1} C^2,$$

ie.

$$\text{Global} \xrightarrow{\text{grad}} \text{Pairwise} \xrightarrow{\text{curl}} \text{Triplewise}$$

and so

$$\text{Global} \xleftarrow{\text{grad}^*(=:-\text{div})} \text{Pairwise} \xleftarrow{\text{curl}^*} \text{Triplewise}.$$

We have

$$\text{curl} \circ \text{grad}(\text{Global Rankings}) = 0.$$

This implies

- global rankings are transitive/consistent,
- no need to consider rankings beyond triplewise.

High Dim. Combinatorial Laplacians

- k-dimensional **combinatorial Laplacian**, $\Delta_k : C^k \to C^k$ by

$$\Delta_k = \delta_{k-1}\delta_{k-1}^* + \delta_k^*\delta_k, \qquad k > 0$$

- $k = 0$, **graph Laplacian** or **vertex Laplacian**

$$\Delta_0 = \delta_0^*\delta_0$$

- $k = 1$, **vector Laplcian** (first term is edge Laplacian)

$$\Delta_1 = \delta_0\delta_0^* + \delta_1^*\delta_1 = \mathrm{curl} \circ \mathrm{curl}^* - \mathrm{div} \circ \mathrm{grad}$$

- Important Properties:
  - Δ_k positive semidefinite
  - $\ker(\Delta_k) = \ker(\delta_{k-1}^*) \cap \ker(\delta_k)$ **harmonic forms**
  - Hodge decomposition

Hodge Decomposition Theorem

- Every combinatorial Laplacians Δ_k has an associated Hodge decomposition.
- For $k = 1$, this is the decomposition (of discrete vector fields/skew symmetric matrices/network flows) that we have been discussing.

Theorem (Hodge decomposition for pairwise ranking)

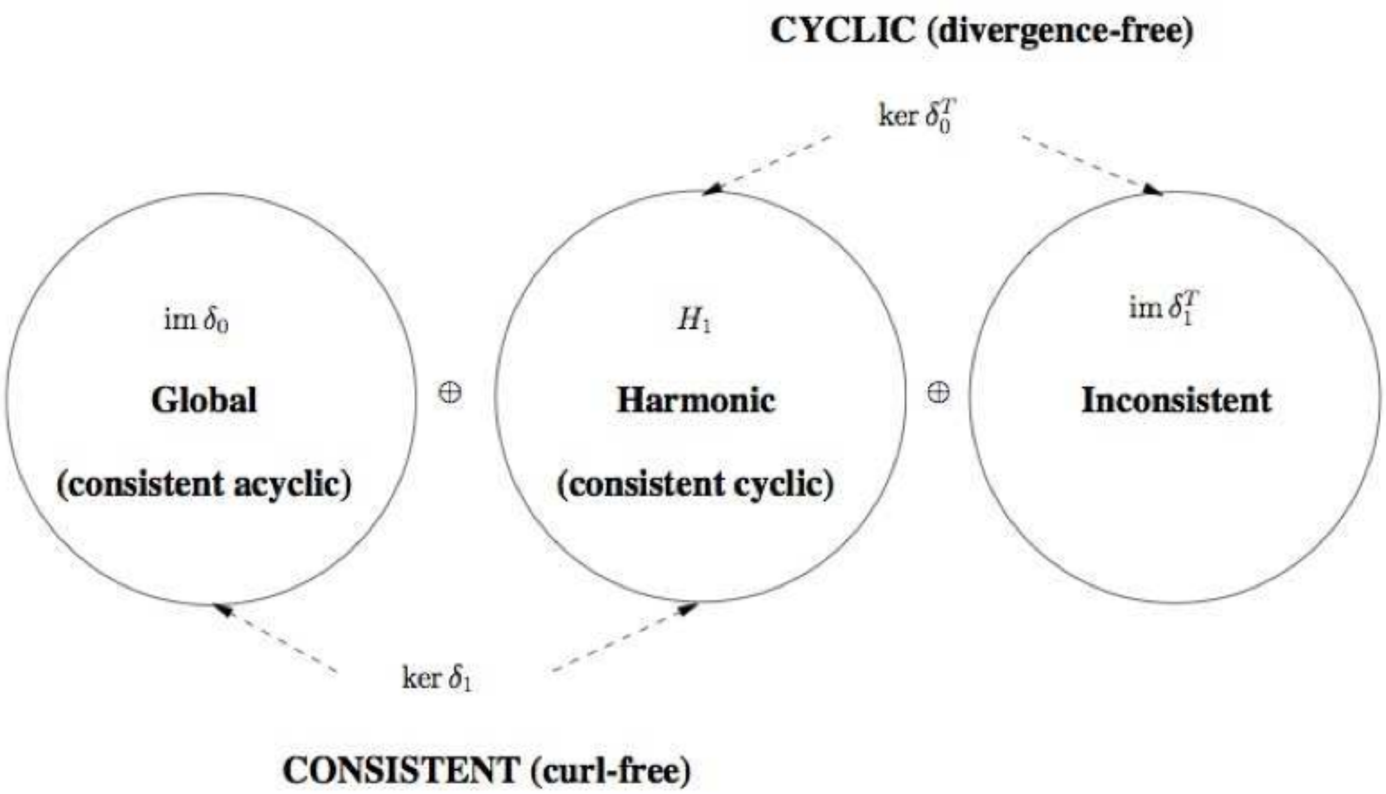
The space of pairwise rankings, $C^1(\mathcal{K}(G), \mathbb{R})$, admits an orthogonal decomposition into three

$$C^1(\mathcal{K}(G), \mathbb{R}) = \operatorname{im}(\delta_0) \oplus H_1 \oplus \operatorname{im}(\delta_1^*)$$

where

$$H_1 = \ker(\delta_1) \cap \ker(\delta_0^*) = \ker(\Delta_1).$$

Hodge Decomposition Theorem



Harmonic rankings: locally consistent but globally inconsistent (circular coordinate)

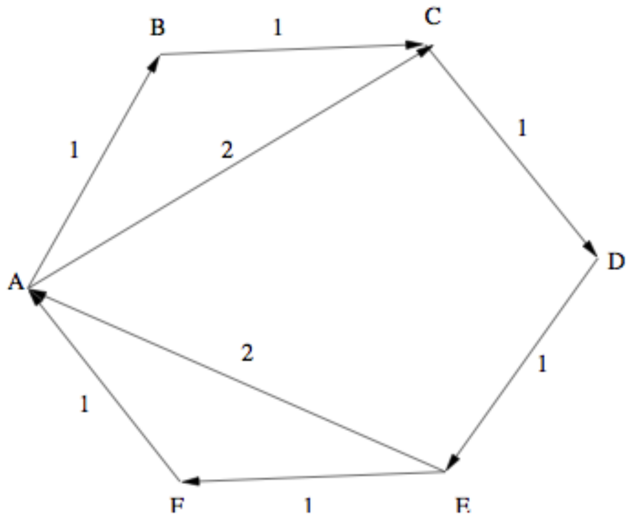


Figure: A locally consistent but globally cyclic harmonic ranking.

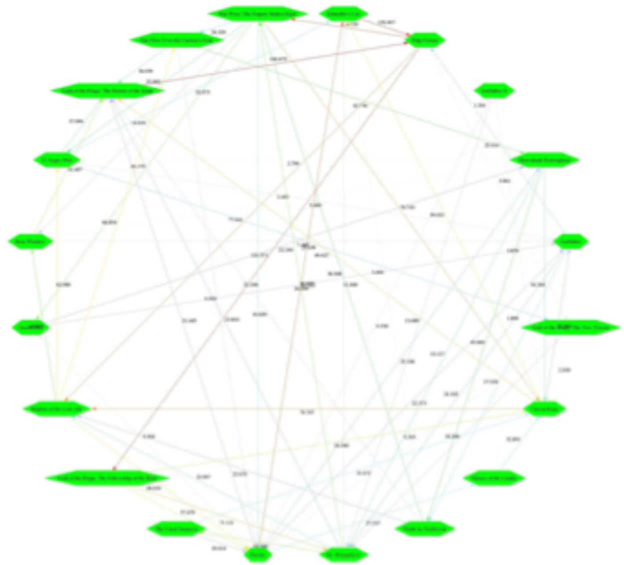


Figure: A harmonic ranking from truncated Netflix movie-movie network

Rank Aggregation as Projection

Rank aggregation problem reduced essentially to linear least squares

Corollary

Every pairwise ranking admits a unique orthogonal decomposition,

$$w = \mathrm{proj}_{\mathrm{im}(\delta_0)} w + \mathrm{proj}_{\ker(\delta_0^*)} w$$

i.e.

$$\textit{pairwise} = \mathrm{grad}(\textit{global}) + \textit{cyclic}$$

Particularly the first projection grad(*global*) *gives a global ranking*

$$x^* = (\delta_0^* \delta_0)^\dagger \delta_0^* w = -(\Delta_0)^\dagger \, \mathrm{div}(w)$$

$O(n^3)$ flops complexity with great algorithms (dense: Givens/Householder QR, Golub-Reinsch SVD; sparse: CGLS, LSQR; sampling: DMM '06)

Don Saari's Geometric Illustration of Different Projections

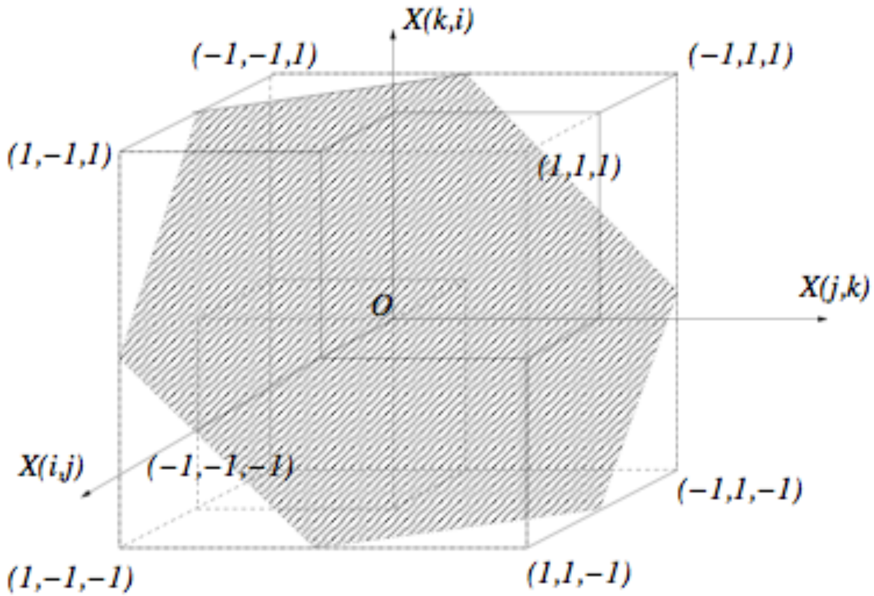


FIGURE 4. The shaded region is the subspace $X_{ij}+X_{jk}+X_{ki}=0$. The transitive region consists of six orthants whose corresponding vertices belong to $\{\pm 1, \pm, 1, \pm 1\} - \{[1,1,1], [-1,-1,-1]\}$. The Borda count or $\min_{X \in \mathcal{M}_G} I(X)$ is the l_2-projection onto the shaded plane while the Kemeny optimization or $\min_{X \in \mathcal{M}_K} I(X)$ is the l_1-projection onto the transitive region.

Measuring Inconsistency by Curls

- Define the cyclicity ratio by

$$C_p := \left(\frac{\left\| proj_{im(curl^*)} w \right\|}{\|w\|} \right)^2 \leq 1$$

This measures the total inconsistency within the data and model w.

- Relative curl

$$c_r(i,j,k) = \frac{|w_{ij} + w_{jk} + w_{ki}|}{|w_{ij}| + |w_{jk}| + |w_{ki}|} \in [0,1]$$

which equals to 1 iff {i,j,k} is a combinatorial intransitive triangle.

Application: 6 Movies with Dynamic Drifts

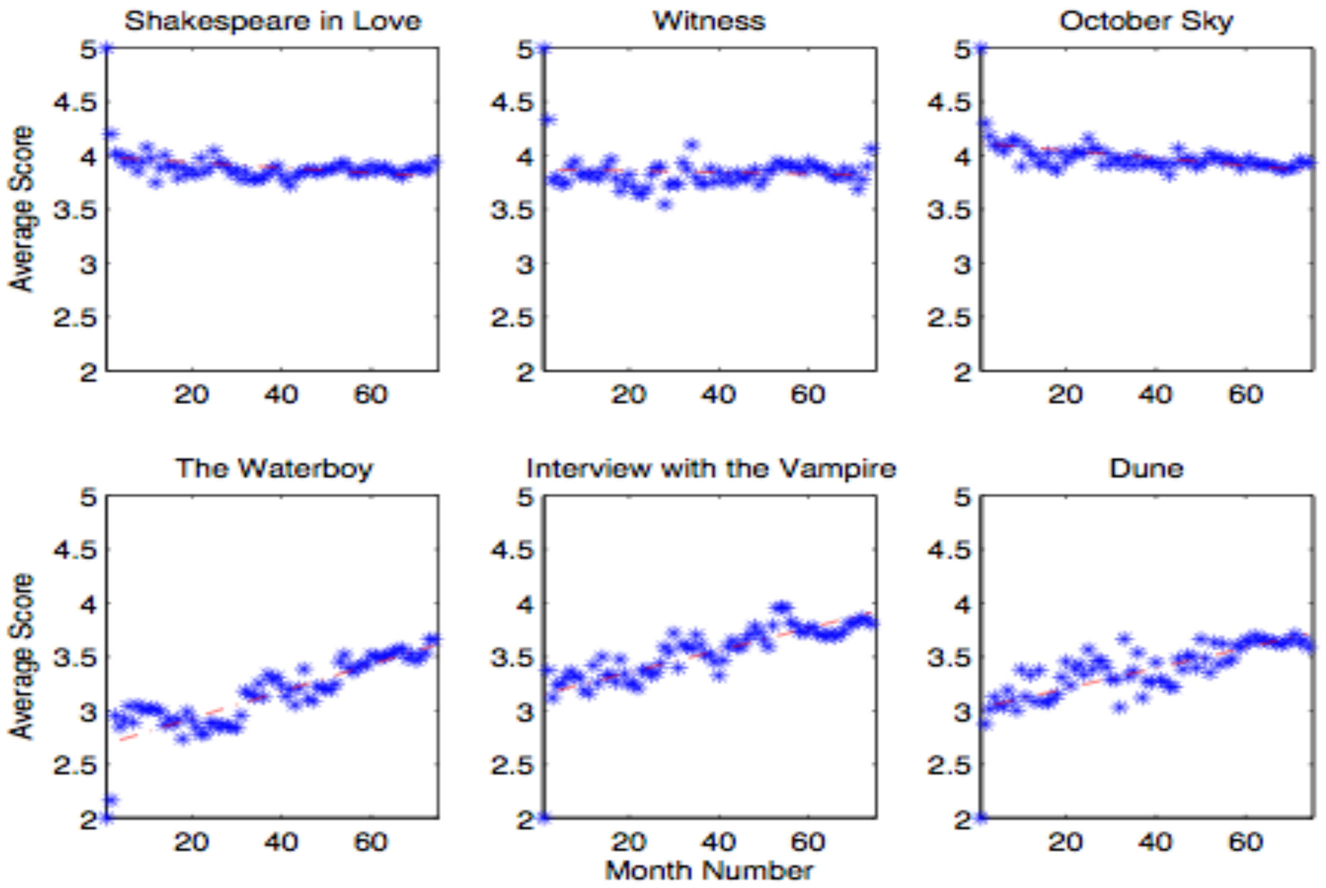


FIGURE 5. Average scores of 6 selected movies over 74 months. The three movies in the top row has a *decreasing* trend in monthly average scores, while in a contrast the other three movies in the bottom row exhibits an *increasing* trend.

Model Selection by Cyclicity Ratio

Movie	Global ranking (Score)				
	MRQE	Mean	Hodge-Difference	Hodge-Ratio	Hodge-Binary
Shakespeare in Love	1 (85)	2 (3.87)	1 (0.247)	2 (0.0781)	1 (0.138)
Witness	2 (77)	3 (3.86)	2 (0.217)	1 (0.0883)	3 (0.107)
October Sky	3 (76)	1 (3.93)	3 (0.213)	3 (0.0775)	2 (0.111)
The Waterboy	4 (66)	6 (3.38)	6 (−0.464)	6 (−0.1624)	6 (−0.252)
Interview with the Vampire	5 (65)	4 (3.71)	4 (−0.031)	4 (−0.0121)	4 (−.012)
Dune	6 (44)	5 (3.49)	5 (−0.183)	5 (−0.0693)	5 (−0.092)
Cyclicity ratio	–	–	0.77	1.15	0.30

TABLE 1. Global ranking of selected six movies via different methods: MRQE, mean score over customers, Hodge decomposition with algorithmic mean score difference, Hodge decomposition with geometric mean score ratio, and Hodge decomposition with binary comparisons. It can be seen that the Hodge decomposition with binary comparisons has the smallest inconsistency in terms of the cyclicity ratio.

MRQE: Movie-Review-Query-Engine (http://www.mrqe.com/)

Just for fun:
Chinese Universities (mainland) Ranking

Rank	Research 2002	Pagerank	HITS Authority	HITS Hub
Pku.edu.cn	1	2	2	1
Tsinghua.edu.cn	1	1	1	6
Fudan.edu.cn	3	7	50	21

DATA: 2002, http://cybermetrics.wlv.ac.uk/database/stats/data

Application II: Hodge Decomposition of Games

- Every strategy profile is a node in a graph
- Two strategy profile is comparable iff only 1 player's strategy changed
- The edge flow is that player's utility difference

STRATEGIES	B Cooperate	B Defect
A Cooperate	(3,3)	(0,5)
A Defect	(5,0)	(1,1)

$$\begin{tikzcd}
(c,c) \arrow[r, "2"] \arrow[d, "2"] & (c,d) \arrow[d, "1"] \\
(d,c) \arrow[r, "1"] & (d,d)
\end{tikzcd}$$

Note: Prisoner's dilemma is a potential game to its Nash equilibrium, not efficient! So we want new way for flow construction...

Candogan-Menache-Ozdaglar-Parrilo, 2010, Flows and Decompositions of Games: Harmonic and Potential Games, arXiv: 1004.2405v1, May 13, 2010.

Hodge Decomposition of Finite Games

- The flows defined above preserve Nash equilibrium
- Every player's utility can be normalized to mean zero, without changing the flow and thus Nash Equilibrium
- The residue of mean-zero utility from the original utility is called non-strategic games

Hodge Decomposition of Finite Games

=

- Non-strategic Games
- Potential Games
- Harmonic Games (zero-sum)

Potential Games

- We consider finite games in strategic form, $\mathcal{G} = \langle \mathcal{M}, \{E^m\}_{m\in\mathcal{M}}, \{u^m\}_{m\in\mathcal{M}} \rangle$.
- $\mathcal{G}$ is an exact potential game if $\exists\, \Phi : E \to \mathbb{R}$ such that

$$\Phi(x^m, x^{-m}) - \Phi(y^m, x^{-m}) = u^m(x^m, x^{-m}) - u^m(y^m, x^{-m}),$$

- Weaker notion: ordinal potential game, if the utility differences above agree only in sign.
- Potential Φ aggregates and explains incentives of all players.
- Examples: congestion games, etc.

Potential Games

- A global maximum of an ordinal potential is a pure Nash equilibrium.
- Every finite potential game has a pure equilibrium.
- Many learning dynamics (such as better-reply dynamics, fictitious play, spatial adaptive play) "converge" to a pure Nash equilibrium [Monderer and Shapley 96], [Young 98], [Marden, Arslan, Shamma 06, 07].

Monderer-Shapley Condition

A path is a collection of strategy profiles $\gamma = (x_0, \ldots, x_N)$ such that x_i and x_{i+1} differ in the strategy of exactly one player where $x_i \in E$ for $i \in \{0, 1, \ldots N\}$. For any path γ, let

$$I(\gamma) = \sum_{i=1}^{N} u^{m_i}(x_i) - u^{m_i}(x_{i-1}),$$

where m_i denotes the player changing its strategy in the ith step of the path.

Theorem ([Monderer and Shapley 96])

A game $\mathcal{G}$ is an exact potential game if and only if for all simple closed paths, γ, $I(\gamma) = 0$. Moreover, it is sufficient to check closed paths of length 4.

Harmonic Games as Orthogoal Complement of Potential Games

Very different properties than potential games.
Agreement between players is never a posibility!

- Simple examples: rock-paper-scissors, cyclic games, etc.
- Essentially, sums of cycles.
- Generically, *never* have pure Nash equilibria.
- Uniformly mixed profile (for all players) is mixed Nash.

Other interesting static and dynamic properties (e.g., correlated equilibria, best-response dynamics, etc.)

Bimatrix Games

For two-player games, simple explicit formulas.
Assume the game is given by matrices (A, B), and (for simplicity), the non-strategic component is zero (i.e., $\mathbf{1}^T A = 0, B\mathbf{1} = 0$). Define

$$S := \frac{1}{2}(A + B), \quad D := \frac{1}{2}(A - B), \quad \Gamma := \frac{1}{2n}(A\mathbf{1}\mathbf{1}^T - \mathbf{1}\mathbf{1}^T B).$$

- Potential component:

$$(S + \Gamma, \quad S - \Gamma)$$

- Harmonic component:

$$(D - \Gamma, \quad -D + \Gamma)$$

Notice that the harmonic component is zero sum.

Reference

- Jiang, Lim, Yao, and Ye (2011). Statistical Ranking and Combinatorial Hodge Theory, Mathematical Programming, 127(1):203-244.
- Candogan, Menache, Ozdaglar, and Parrilo (2011). Flows and Decompositions of Games: Harmonic and Potential Games. Mathematics of Operations Research.
- [Friedman1996] J. Friedman, Computing Betti Numbers via Combinatorial Laplacians. Proc. 28th ACM Symposium on Theory of Computing, 1996. http://www.math.ubc.ca/~jf/pubs/web_stuff/newbetti.html
- [Forman] R. Forman, Combinatorial Differential Topology and Geometry, New perspective in Algebraic Combinatorics, MSRI publications, http://math.rice.edu/~forman/combdiff.ps
- [David1988] David, H.A. (1988). The Method of Paired Comparisons. New York: Oxford University Press.
- [Saari1995] D. Saari, Basic Geometry of Voting, Springer, 1995