

Statistical Ranking and Combinatorial Hodge Theory

Jiang, Lim, Yao and Ye (2011)

Presenter: Jose Lugo-Martinez
Pedja's Lab Meeting

November 16, 2011



SCHOOL OF INFORMATICS
AND COMPUTING

INDIANA UNIVERSITY
Bloomington



Outline

- Statistical Ranking Problem on Graphs
- Motivation
- HodgeRank Global Ranking Method
 - Hodge and Helmholtz decomposition



Statistical Ranking Problem on Graphs

- Determining a global ranking from a dataset consisting of *alternatives* ranked by *voters*
 - Why rank datasets?
 - Identification of important entries, etc

“Classical” Learning Scenario

Example	Alternatives	Voters
Netflix	movies	viewers
Google’s search engine	websites	other websites
CiteSeer’s citation db	scholarly articles	other articles

- Netflix dataset - instead of ranking, viewers assign opinion scores based on some scale



What's the Problem?

- Why not use traditional ranking methods?
 - Inapplicable or ineffective
 - cardinal scores instead of ordinal orderings
 - incomplete data (99% missing values in Netflix dataset)
 - imbalanced data (blockbuster vs independent movies)
 - data lives on large complex networks
- Goal
 - Novel global ranking method (if possible)
 - Provide 'certificate of reliability' for global ranking



HodgeRank Global Ranking Method

Jiang, Lim, Yao and Ye (2011)

- 1st use of Hodge Theory
 - Hodge decomposition applied to ranking data
- Statistical analysis of inherent inconsistencies
 - harmonic ranking (i.e. feasibility of global ranking)
- Computationally tractable
- Provides certificate of reliability for global ranking

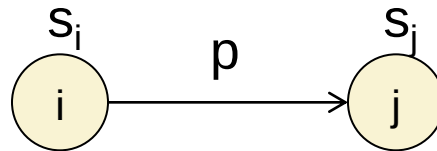


Notation

- $A = \{1, \dots, n\}$ represent set of *alternatives*
- $U = \{1, \dots, m\}$ represent set of *voters*
- For i, j in A , $i \leq_{\alpha} j$ means i preferred over j for α in U

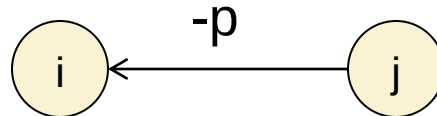
$p :=$ degree of preference

$$Y_{ij}^{\alpha} = p$$



For example, $p = (s_j - s_i)$

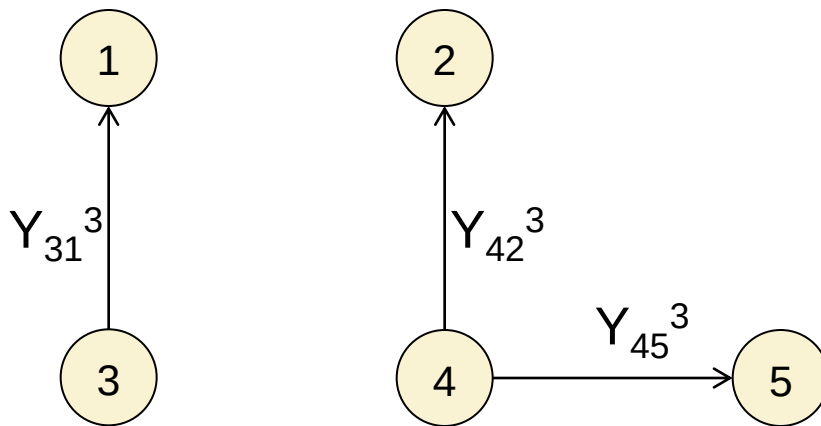
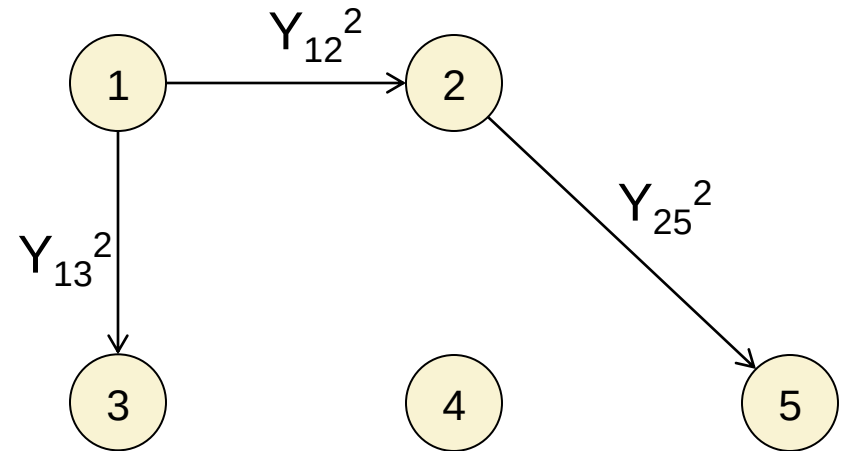
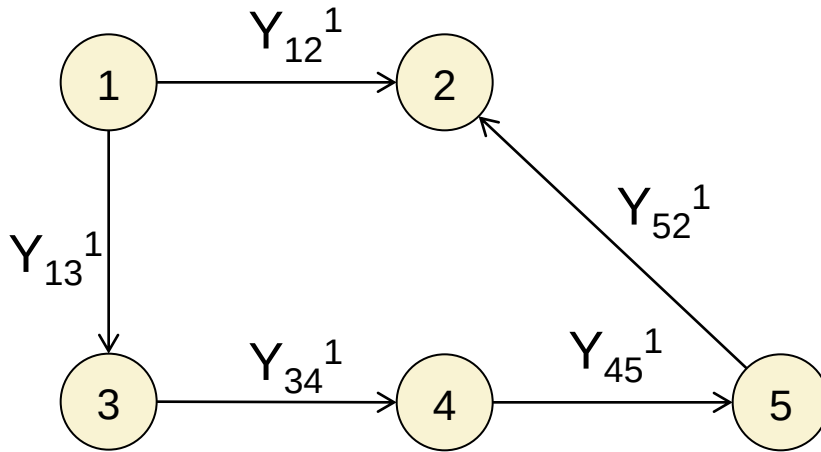
$$Y_{ji}^{\alpha} = -p$$



Y_{ij}^{α} is 1 if rater α prefers i to j and -1 otherwise



Example: $A = \{1,2,3,4,5\}$, $U = \{1,2,3\}$

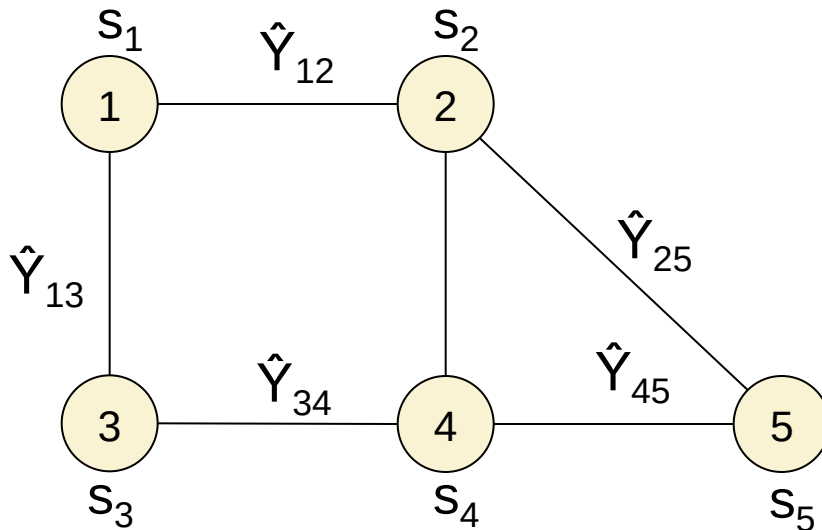




Unifying Learning Scenario

- Graph structures naturally arise from *pairwise comparisons*
 - Pairwise Comparison Graph $G = (V, E)$ where:
 - $V := A$ (i.e. set of alternatives)
 - $E :=$ set of pairs $\{i, j\}$ where pairwise comparisons are made
- Unifies all ranking problems on graph space
 - sparsity structure represents data incompleteness
 - degree distributions represent data imbalance

Example Re-visited as PCG



Objective: Find $\hat{s}$

$$\hat{Y}_{ij} \approx \hat{s}_i - \hat{s}_j \text{ for some } \hat{s} : V \rightarrow \mathbb{R}$$

$$\min_{s \in \mathbb{R}^{|V|}} \sum_{\{i,j\} \in E} \omega_{ij} (s_i - s_j - \hat{Y}_{ij})^2,$$

where

- $\hat{Y}_{ij} = (\sum_{\alpha} \omega_{ij}^{\alpha} Y_{ij}^{\alpha}) / (\sum_{\alpha} \omega_{ij}^{\alpha})$, skew-symmetric matrix
- $\omega_{ij} = \sum_{\alpha} \omega_{ij}^{\alpha}$
- ω_{ij}^{α} is an indicator or confidence weight



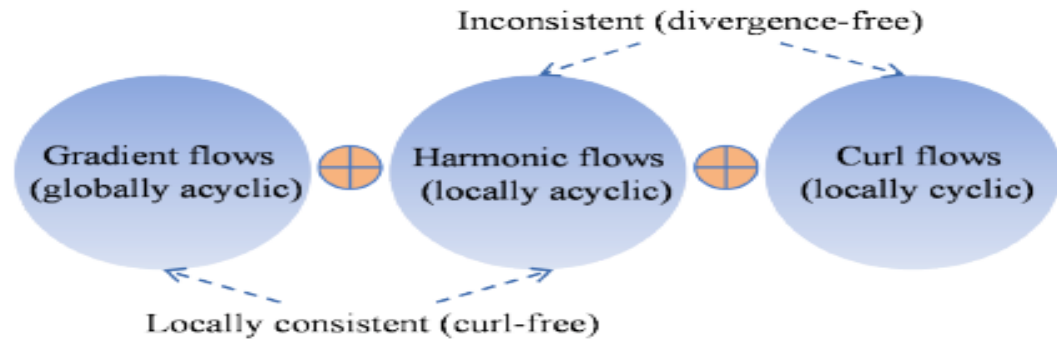
Hodge Theory

- Well-known in pure math (algebraic geometry & topology) and applied math/physics (computational electromagnetics & fluid dynamics)
- Combinatorial Hodge theory
 - simplest possible space (a graph)
 - provides a geometric interpretation of the minimizer and residual

Way to determine a global ranking with a 'certificate of reliability' for validity of such global ranking!

HodgeRank on Graphs

Hodge and Helmholtz
Decomposition Theorem



Every $\hat{Y}$ admits an **orthogonal** decomposition adapted to G ,

$$\hat{Y} = \hat{Y}^{(1)} + \hat{Y}^{(2)} + \hat{Y}^{(3)},$$

where

$$\begin{aligned} \hat{Y}_{ij}^{(1)} &= \hat{s}_i - \hat{s}_j, \quad \text{for some } \hat{s} \in \mathbb{R}^V, \\ \hat{Y}_{ij}^{(2)} + \hat{Y}_{jk}^{(2)} + \hat{Y}_{ki}^{(2)} &= 0, \quad \text{for each } \{i, j, k\} \in T, \\ \sum_{j \sim i} \omega_{ij} \hat{Y}_{ij}^{(2)} &= 0, \quad \text{for each } i \in V. \\ \hat{Y}_{ij}^{(3)} + \hat{Y}_{jk}^{(3)} + \hat{Y}_{ki}^{(3)} &\neq 0, \quad \{i, j, k\} \in T \end{aligned}$$



Harmonic and Triangular Curl Example

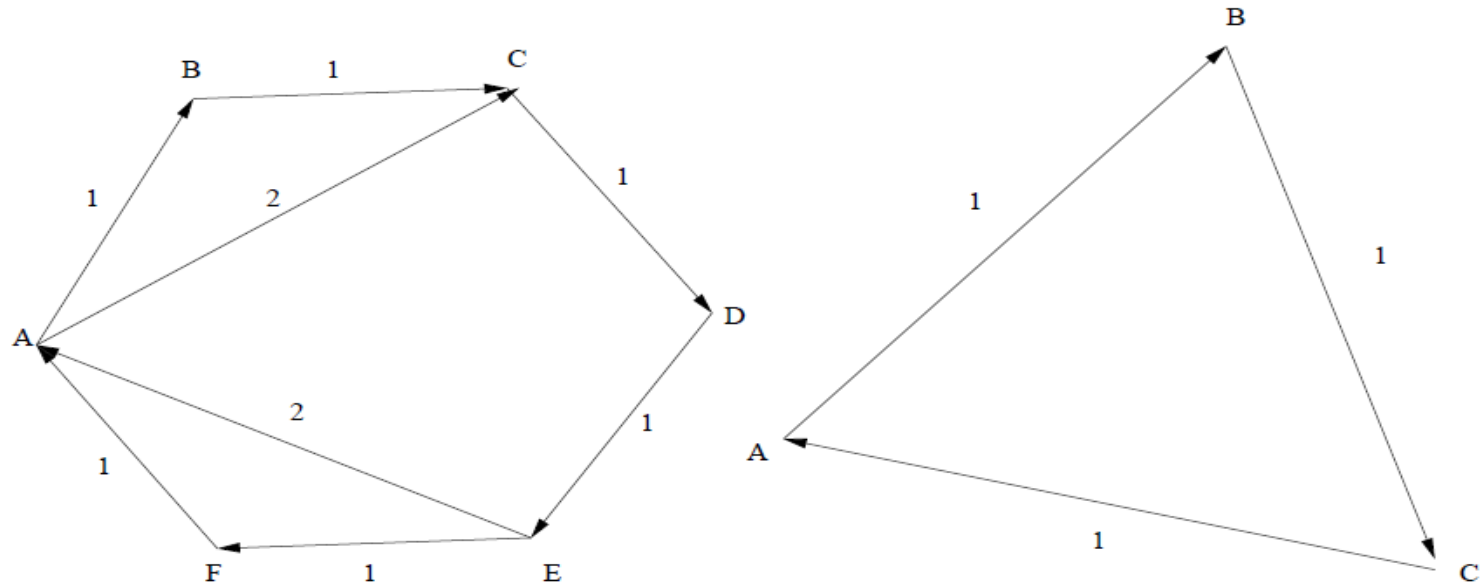


Figure: Left: example of $\hat{Y}^{(2)}$, harmonic; Right: example of $\hat{Y}^{(3)}$, curl.

Left: A harmonic pairwise ranking, which is locally consistent on every triangle but inconsistent along the loop A, B, C, D, E, F, A



Summary

- HT yields orthogonal decomposition
 - *gradient flow* (GF) globally acyclic
 - *harmonic flow* (HF) locally acyclic but globally cyclic
 - *curl flow* (CF) locally cyclic
- Useful because ...
 - GF induces a global ranking of A
 - easily computed via a linear least squares problem
 - L_2 -norm of least squares residual represents contribution from HF and CF, and quantifies validity of global ranking
 - If small, global ranking obtained from GF → majority consensus
 - If large, dataset plagued with cyclic inconsistencies → no reasonable ranking exists



INDIANA UNIVERSITY

Thank You!

QUESTIONS?