Applied Hodge Theory

Yuan Yao

School of Mathematical Sciences
Peking University

May 8, 2015

Helmholtz-Hodge Decomposition

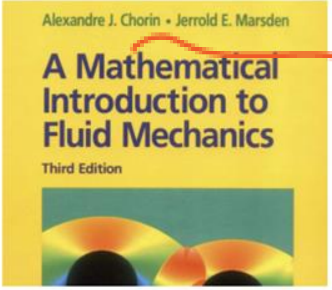
Theorem (c.f. Marsden-Chorin 1992)

A vector field $\mathbf{w}$ *on a simply-connected* D *can be uniquely decomposed in the form*

$$\mathbf{w} = \mathbf{u} + \operatorname{grad}\phi$$

where $\mathbf{u}$ *has zero divergence and is parallel to* ∂D.

vector field = curl-free + div-free



Hodge Theory in Linear Algebra

For inner product spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, consider

$$\mathcal{X} \xrightarrow{A} \mathcal{Y} \xrightarrow{B} \mathcal{Z}.$$

Hodge Laplacian

and $\Delta = AA^* + B^*B : \mathcal{Y} \to \mathcal{Y}$ where $(\cdot)^*$ is adjoint operator of $(\cdot)$. If

$$B \circ A = 0,$$

then $\ker(\Delta) = \ker(A^*) \cap \ker(B)$ and "*orthogonal*" decomposition

$$\mathcal{Y} = \mathrm{im}(A) + \ker(\Delta) + \mathrm{im}(B^*)$$

Note: $\ker(B)/\mathrm{im}(A) \simeq \ker(\Delta)$ is the (real) (co)-homology group ($\mathbb{R} \to$ rings; vector spaces$\to$module).

Hodge Decomposition=Rank-Nullity Theorem

Take product space $V = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, define

Dirac
$$D = \begin{pmatrix} 0 & 0 & 0 \\ A & 0 & 0 \\ 0 & B & 0 \end{pmatrix}, \quad BA = 0,$$

Laplacian

$$L = (D+D^*)^2 = \mathrm{diag}(A^*A, AA^*+B^*B, BB^*) = \mathrm{diag}(L_0, L_1, L_2^{(down)})$$

Rank-nullity Theorem: $\mathrm{im}(D) + \ker(D^*) = V$, in particular

$$\begin{aligned} \mathcal{Y} &= \mathrm{im}(A) + \ker(A^*) \\ &= \mathrm{im}(A) + \ker(A^*)/\,\mathrm{im}(B^*) + \mathrm{im}(B^*), \text{since } \mathrm{im}(A) \subseteq \ker(B) \\ &= \mathrm{im}(A) + \ker(A^*) \cap \ker(B) + \mathrm{im}(B^*) \end{aligned}$$

Terminology

- coboundary maps: $A : \mathcal{X} \to \mathcal{Y}$, $B : \mathcal{Y} \to \mathcal{Z}$
- cochains: elements in $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$
- cochain complex: $\mathcal{X} \xrightarrow{A} \mathcal{Y} \xrightarrow{B} \mathcal{Z}$.
- cocycles: elements of $\ker(A)$
- coboundaries: elements of $\operatorname{im}(B)$
- cohomology classes: elements of $\ker(A)/\operatorname{im}(B)$
- harmonic cochains: elements of $\ker(A^*A + BB^*)$
- Betti number: $\dim\ker(A^*A + BB^*)$
- closed: $Ax = 0$
- exact: $x = Bz$

Classical Hodge Theory on Riemannian Manifolds

- (W.V.D. Hodge, 1903-1975) de Rham complex:

$$0 \to \Omega^0(M) \xrightarrow{d_0} \Omega^1(M) \xrightarrow{d_1} \cdots \xrightarrow{d_{n-1}} \Omega^n(M) \xrightarrow{d_n} 0$$

- M: compact Riemannian manifold
- $\Omega^k(M)$: with k-differential forms
- d: the exterior derivative

$$d^2 = d_k \circ d_{k-1} = 0$$

Hodge Theory on Metric Spaces

- (Alexander-Spanier, Bartholdi-Schick-Smale-Smale, 2011)

$$0 \to L^2(X) \xrightarrow{d_0} L^2(X^2) \xrightarrow{d_1} \cdots \xrightarrow{d_{n-1}} L^2(X^n) \xrightarrow{d_n} \cdot$$

- X: metric space
- $L^2(X)$: square integral functions on X
- $d : L^2(X^k) \to L^2(X^{k+1})$, finite difference

$$(df)(x_0, \ldots, x_k) = \sum_{i=1}^{k} (-1)^i \prod_{j \neq i} \sqrt{K(x_i, x_j)} f(x_{-i})$$

- adjoint operator $\delta : L^2(X^{k+1}) \to L^2(X^k)$

$$\delta g(x) = \sum_{i=0}^{k} (-1)^i \int_X \prod_{j=0}^{k-1} \sqrt{K(t, x_j)} g(x_0, \ldots, x_{i-1}, t, x_i, \ldots, x_{k-1}) dt$$

Combinatorial Hodge Theory on Simplicial Complexes

$$0 \to \Omega^0(X) \xrightarrow{d_0} \Omega^1(X) \xrightarrow{d_1} \cdots \xrightarrow{d_{n-1}} \Omega^n(X) \xrightarrow{d_n} \cdots$$

- X is finite
- $\chi(X) \subseteq 2^X$: simplicial complex formed by $X \Leftrightarrow$ if $\tau \in \chi(X)$ and $\sigma \subseteq \tau$, then $\sigma \in \chi(X)$

X

- k-forms or cochains as alternating functions

$$\Omega^k(X) = \{u : \chi_{k+1}(X) \to \mathbb{R}, u_{i_{\sigma(0)},\ldots,i_{\sigma(k)}} = \mathrm{sign}(\sigma) u_{i_0,\ldots,i_k}\}$$

A

- coboundary maps $d_k : \Omega^k(X) \to \Omega^{k+1}(X)$ alternating difference

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

- $d_k \circ d_{k-1} = 0$

Example: graph and clique complex

x Y Z

- $G = (X, E)$ is a undirected but oriented graph
- Clique complex $\chi_G \subseteq 2^X$ collects all complete subgraph of G
- k-forms or cochains $\Omega^k(\chi_G)$ as alternating functions:
  - 0-forms: $v : V \to \mathbb{R} \cong \mathbb{R}^n$
  - 1-forms as skew-symmetric functions: $w_{ij} = -w_{ji}$
  - 2-forms as triangular-curl:

$$z_{ijk} = z_{jki} = z_{kij} = -z_{jik} = -z_{ikj} = -z_{kji}$$

A, B

- coboundary operators as alternating difference operators:
  - $(d_0 v)(i,j) = v_j - v_i =: (\text{grad } v)(i,j)$
  - $(d_1 w)(i,j,k) = (\pm)(w_{ij} + w_{jk} + w_{ki}) =: (\text{curl } w)(i,j,k)$
- $d_1 \circ d_0 = \text{curl}(\text{grad } u) = 0$

Hodge Laplacian

- combinatorial Laplacian $\Delta = d_{k-1}d_{k-1}^* + d_k^* d_k$
  - $k = 0$, $\Delta_0 = d_0^* d_0$ is the (unnormalized) graph Laplacian
  - $k = 1$, 1-Hodge Laplacian (Helmholtzian)

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

- Hodge decomposition holds for $\Omega^k(X)$
  - $\Omega^k(X) = \mathrm{im}(d_{k-1}) \oplus \ker(\Delta_k) \oplus \mathrm{im}(\delta_k)$
  - $\dim(\Delta_k) = \beta_k(\chi(X))$

$\ker(\Delta_k)$

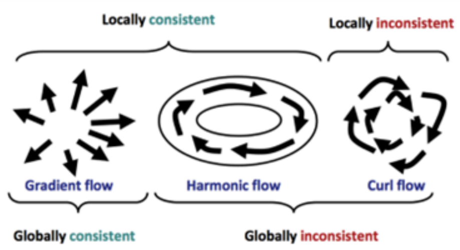


Figure: Courtesy by Asu Ozdaglar

Social Choice Problem

How to aggregate preferences
which faithfully represent individuals?

Crowdsourcing QoE evaluation of Multimedia

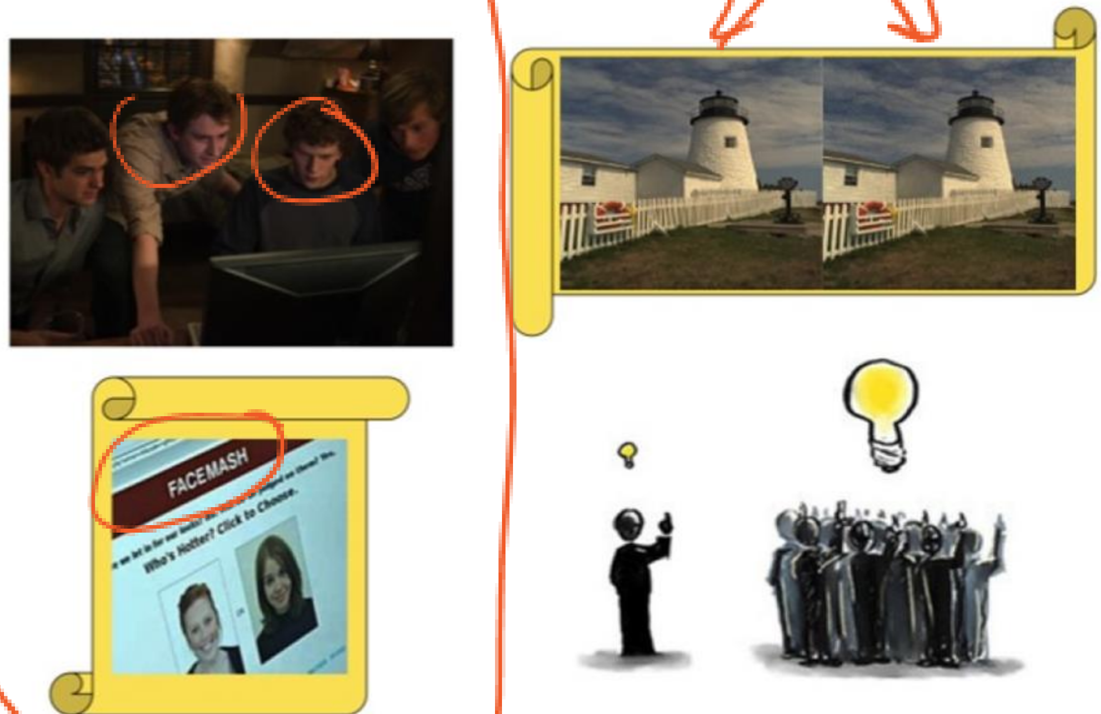


Figure: (Xu-Huang-Y., et al. 11) Crowdsouring subjective Quality of Experience evaluation

Crowdsourcing ranking

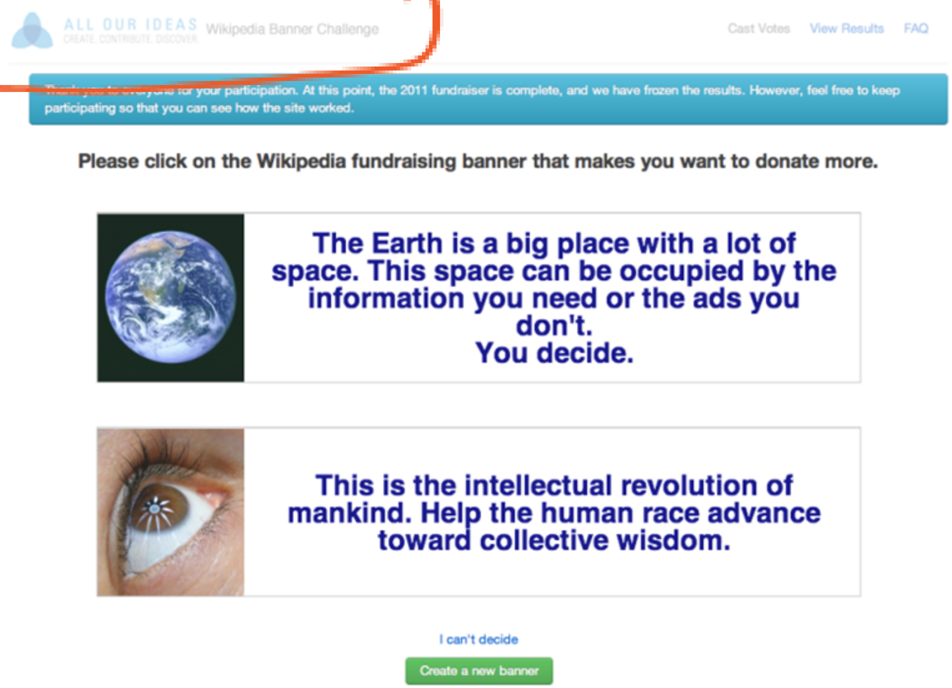


Figure: Left: www.allourideas.org/wikipedia-banner-challenge, by Prof. Matt Salganik at Princeton; Right: www.crowdrank.net

Learning relative attributes: age

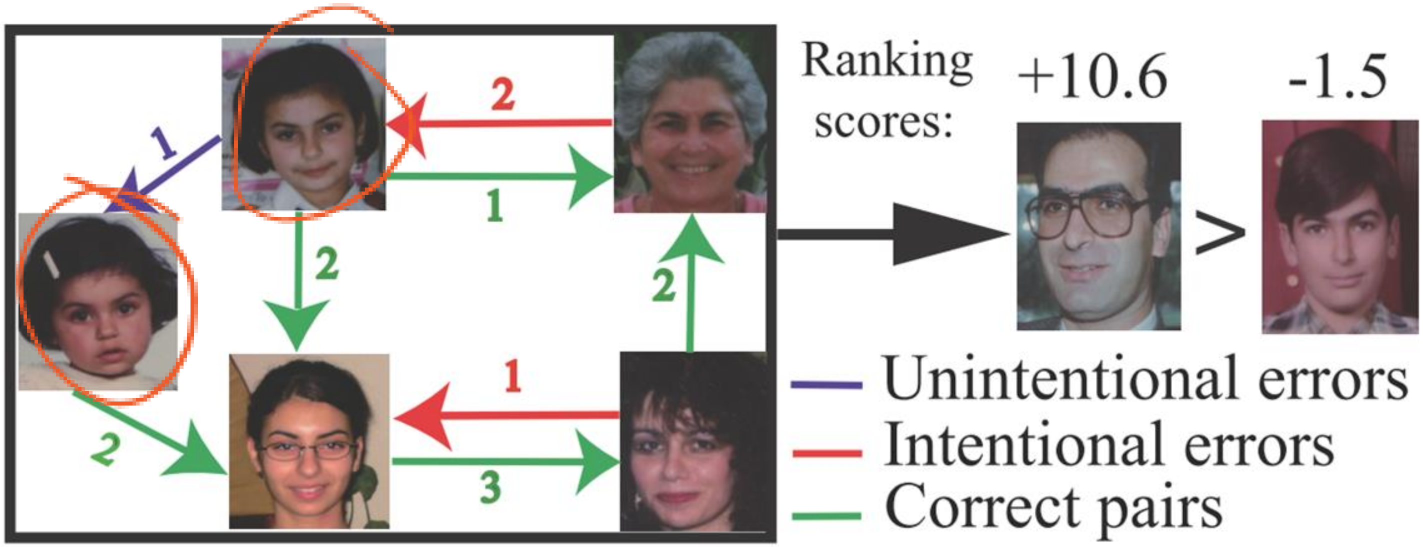
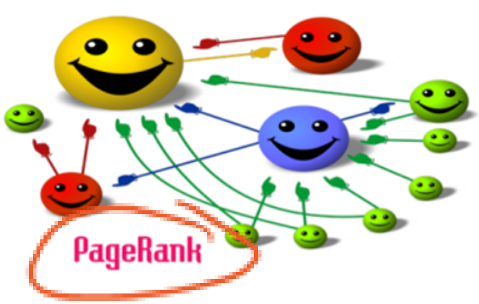


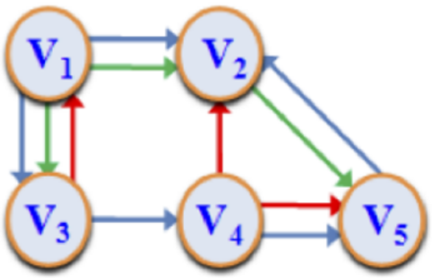
Figure: Age: a relative attribute estimated from paired comparisons (Fu-Y.-Xiang et al. 2014)

Paired comparison data on graphs

Graph $G = (V, E)$

- V: alternatives to be ranked or rated
- $(i_\alpha, j_\alpha) \in E$ a pair of alternatives
- $y_{ij}^\alpha \in \mathbb{R}$ degree of preference by rater α
- $\omega_{ij}^\alpha \in \mathbb{R}_+$ confidence weight of rater α
- Examples: relative attributes, subjective QoE assessment, perception of illuminance intensity, sports, wine taste, etc.





Modern settings

Modern ranking data are

- distributive on networks
- incomplete with missing values
- imbalanced
- even adaptive to dynamic and random settings?

Here we introduce:

Hodge Theory approach to Social Choice

History

Classical social choice theory origins from Voting Theory

- *Borda* 1770, B. Count against plurality vote
- *Condorcet* 1785, C. Winner who wins all paired elections
- Impossibility theorems: *Kenneth Arrow* 1963, *Amartya Sen* 1973
- Resolving conflicts: *Kemeny*, *Saari* ...
- In these settings, we study complete ranking orders from voters.

Classical Social Choice or Voting Theory

Problem

Given m voters whose preferences are total orders (permutation) $\{\succeq_i : i = 1, \ldots, m\}$ *on a candidate set* V*, find a social choice mapping*

$$f : (\succeq_1, \ldots, \succeq_m) \mapsto \succeq^*,$$

as a total order on V*, which "best" represents voter's will.*

Example: 3 candidates ABC

Preference order	Votes
$A \succeq B \succeq C$	2
$B \succeq A \succeq C$	3
$B \succeq C \succeq A$	1
$C \succeq B \succeq A$	3
$C \succeq A \succeq B$	2
$A \succeq C \succeq B$	2

What we did in practice I: Position rules

There are two important classes of social mapping in realities:

- I. Position rules: assign a score $s : V \to \mathbb{R}$, such that for each voter's order(permutation) $\sigma_i \in S_n$ $(i = 1, \ldots, m)$, $s_{\sigma_i(k)} \geq s_{\sigma_i(k+1)}$. Define the social order by the descending order of total score over raters, i.e. the score for k-th candidate

$$f(k) = \sum_{i=1}^{m} s_{\sigma_i}(k).$$

  - Borda Count: $s : V \to \mathbb{R}$ is given by $(n-1, n-2, \ldots, 1, 0)$
  - Vote-for-top-1: $(1, 0, \ldots, 0)$
  - Vote-for-top-2: $(1, 1, 0, \ldots, 0)$

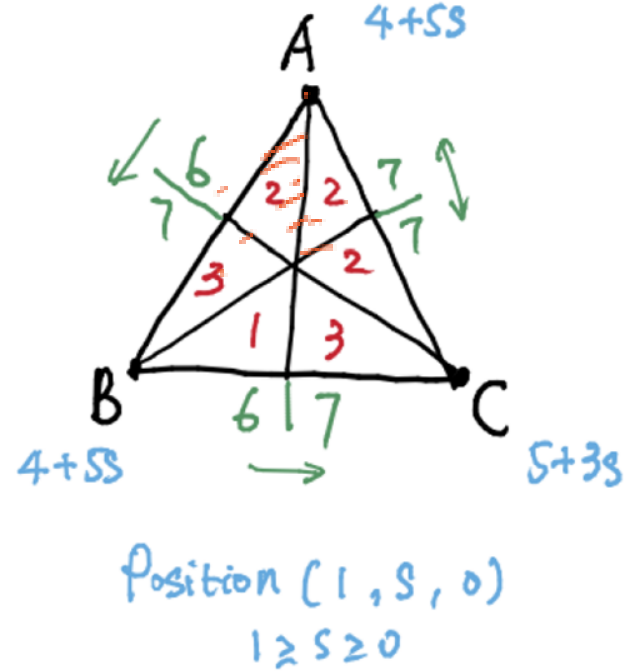
What we did in practice II: pairwise rules

- II. Pairwise rules: convert the voting profile, a (distribution) function on $n!$ set S_n, into paired comparison matrix $X \in \mathbb{R}^{n \times n}$ where $X(i,j)$ is the number (distribution) of voters that $i \succ j$; define the social order based on paired comparison data X.
  - Kemeny Optimization: minimizes the number of pairwise mismatches to X over S_n (NP-hard)
  - Pluarity: the number of wins in paired comparisons (tournaments) – equivalent to Borda count in complete Round-Robin tournaments

Revisit the ABC-Example

- Position:
  - $s < 1/2$, C wins
  - $s = 1/2$, ties
  - $s > 1/2$, A/B wins
- Pairwise:
  - A, B: 13 wins
  - C: 14 wins
  - ~~Kemeny~~ winner C

so completely in chaos!



Arrow's Impossibility Theorem

(Arrow'1963)

Consider the Unrestricted Domain, i.e. voters may have all complete and transitive preferences. The only social choice rule satisfying the following conditions is the dictator rule

- Pareto (Unanimity): if all voters agree that $A \succeq B$ then such a preference should appear in the social order
- Independence of Irrelevant Alternative (IIA): the social order of any pair only depends on voter's relative rankings of that pair

Sen's Impossibility Theorem

(Sen'1970)

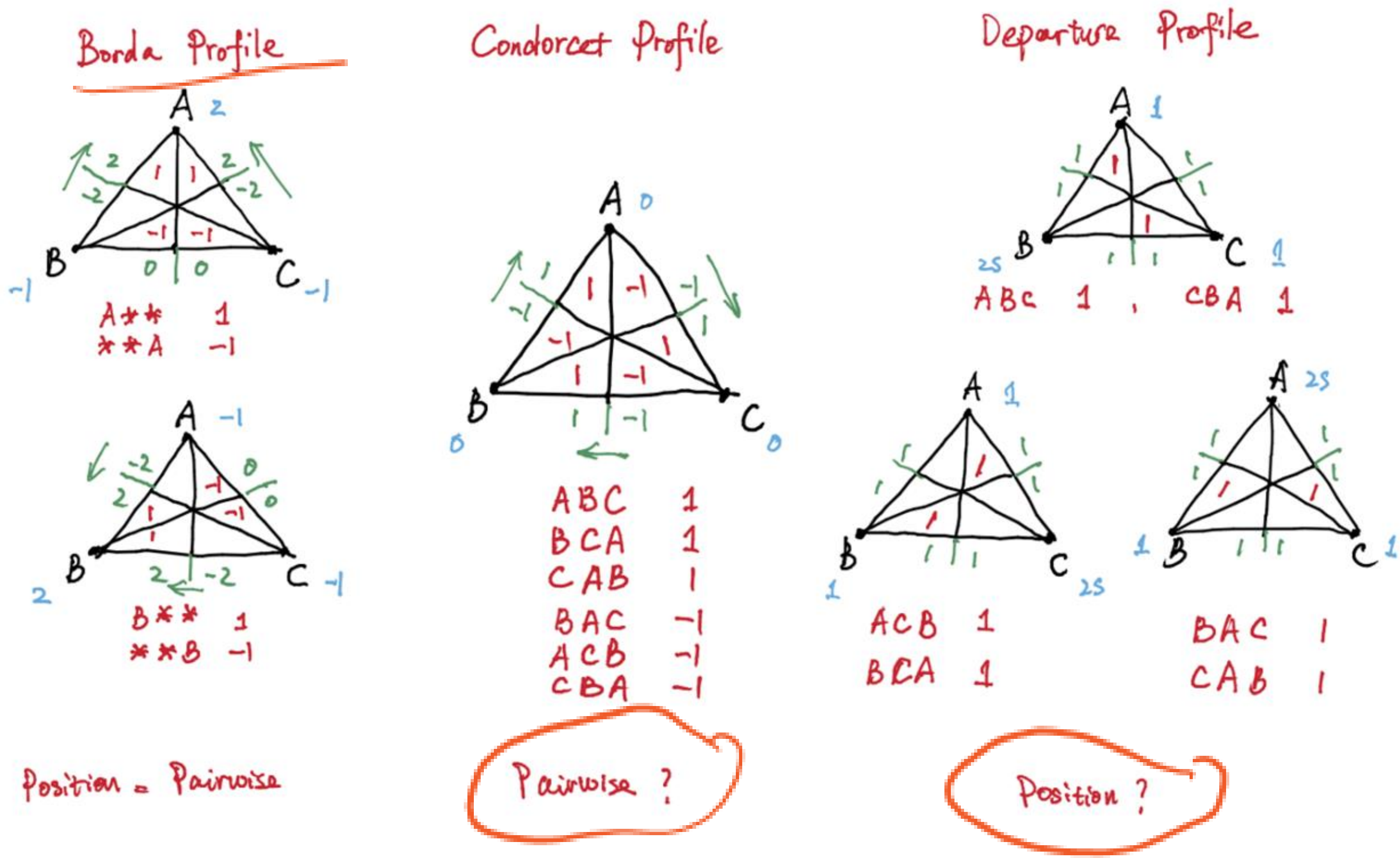
With Unrestricted Domain, there are cases with voting data that no social choice mapping,

$$f : (\succeq_1, \ldots, \succeq_m) \mapsto 2^V,$$

exists under the following conditions

- Pareto: if all voters agree that $A > B$ then such a preference should appear in the social order
- Minimal Liberalism: two distinct voters decide social orders of two distinct pairs respectively

A Decomposition of Voting Profile $R^{3!}$



Saari Decomposition of Complete Voting Profile

Every profile, or distribution function on symmetric group S_n, can be decomposed into the following components:

- Universal kernel: all ranking methods induce a complete tie on any subset of V
  - dimension: $n! - 2^{n-1}(n-2) - 2$
- Borda profile: all ranking methods give the same result
  - dimension: $n-1$
  - basis: $\{1(\sigma(1) = i, *) - 1(*, \sigma(n) = i) : i = 1, \ldots, n\}$
- Condorcet profile: all positional rules give the same result
  - dimension: $\frac{(n-1)!}{2}$
  - basis: sum of Z_n orbit of σ minus their reversals
- Departure profile: all pairwise rules give the same result

Borda Count: the most faithful representation?

Borda count is the most consistent ranking method, since

- for full set ranking, it only depends on Borda profile
- for subset ranking it depends on both Borda and Condorcet profiles

while

- Pairwise rules depend on both Borda and Condorcet profiles
- Position rules depend on both Borda and Departure profiles (except Borda)

So, if you look for best possibility from impossibility, perhaps Borda count is the choice.

Borda Count as a Pairwise Rule: Least Square

Borda Count is equivalent to

$$\min_{x \in \mathbb{R}^{|V|}} \sum_{\alpha, \{i,j\} \in E} (x_i - x_j - y_{ij}^{\alpha})^2,$$

where

- $y_{ij}^{\alpha} > 0$ (e.g. 1), if $i \succeq j$ by voter α, and $y_{ij}^{\alpha} < 0$, on the opposite (e.g. -1).

Note: NP-hard Kemeny Optimization, or Minimimum-Feedback-Arc-Set:

$$\min_{x \in \mathbb{R}^{|V|}} \sum_{\alpha, \{i,j\} \in E} (\mathrm{sign}(x_i - x_j) - y_{ij}^{\alpha})^2,$$

Generalized Borda Count with Incomplete Data

$$\min_{x \in \mathbb{R}^{|V|}} \sum_{\alpha, \{i,j\} \in E} \omega_{ij}^{\alpha} (x_i - x_j - y_{ij}^{\alpha})^2,$$

$$\Leftrightarrow$$

$$\min_{x \in \mathbb{R}^{|V|}} \omega_{ij} \sum_{\{i,j\} \in E} ((x_i - x_j) - \hat{y}_{ij})^2,$$

where $\hat{y}_{ij} = \hat{\mathbb{E}}_{\alpha} y_{ij}^{\alpha} = (\sum_{\alpha} \omega_{ij}^{\alpha} y_{ij}^{\alpha}) / \omega_{ij} = -\hat{y}_{ji}$, $\quad \omega_{ij} = \sum_{\alpha} \omega_{ij}^{\alpha}$

So $\hat{y} \in l_{\omega}^2(E)$, inner product space with $\langle u, v \rangle_{\omega} = \sum u_{ij} v_{ij} \omega_{ij}$, u, v skew-symmetric

Statistical Majority Voting: $l^2(E)$

- $\hat{y}_{ij} = (\sum_\alpha \omega_{ij}^\alpha y_{ij}^\alpha)/(\sum_\alpha \omega_{ij}^\alpha) = -\hat{y}_{ji}$, $\omega_{ij} = \sum_\alpha \omega_{ij}^\alpha$
- $\hat{y}$ from generalized linear models:
  - [1] *Uniform* model: $\hat{y}_{ij} = 2\hat{\pi}_{ij} - 1$.
  - [2] *Bradley-Terry* model: $\hat{y}_{ij} = \log \frac{\hat{\pi}_{ij}}{1-\hat{\pi}_{ij}}$.
  - [3] *Thurstone-Mosteller* model: $\hat{y}_{ij} = \Phi^{-1}(\hat{\pi}_{ij})$, $\Phi(x)$ is Gaussian CDF
  - [4] *Angular transform* model: $\hat{y}_{ij} = \arcsin(2\hat{\pi}_{ij} - 1)$.

Hodge Decomposition of Pairwise Ranking

$\hat{y}_{ij} = -\hat{y}_{ji} \in l^2_\omega(E)$ admits an orthogonal decomposition,

$$\hat{y} = Ax + B^T z + w, \tag{1}$$

where

$$(Ax)(i,j) := x_i - x_j, \quad \text{gradient, as "Borda profile",} \tag{2a}$$

$$(B\hat{y})(i,j,k) := \hat{y}_{ij} + \hat{y}_{jk} + \hat{y}_{ki}, \quad \text{trianglar cycle/curl, Condorcet} \tag{2b}$$

$$w \in \ker(A^T) \cap \ker(B), \quad \text{harmonic, Condorcet.} \tag{2c}$$

In other words

$$\mathrm{im}(A) \oplus \ker(AA^T + B^T B) \oplus \mathrm{im}(B^T)$$

Why?

Note $B \circ A = 0$ since

$$(B \circ Ax)(i,j,k) = (x_i - x_j) + (x_j - x_k) + (x_k - x_i) = 0.$$

Hence

$$A^T \hat{y} = A^T(Ax + B^T z + w) = A^T Ax \Rightarrow x = (A^T A)^{\dagger} A^T \hat{y}$$

$$B\hat{y} = B(Ax + B^T z + w) = BB^T z \Rightarrow z = (BB^T)^{\dagger} B\hat{y}$$

$$A^T w = Bw = 0 \Rightarrow w \in \ker(\Delta_1), \quad \Delta_1 = AA^T + B^T B.$$

Generalized Borda Count estimator

Gradient flow $\hat{y}^{(g)} := (Ax)(i,j) = x_i - x_j$ gives the generalized Borda count score, x which solves Graph Laplacian equation

$$\min_{x \in \mathbb{R}^{|V|}} \sum_{\alpha,(i,j)\in E} \omega_{ij}^{\alpha}(x_i - x_j - y_{ij}^{\alpha})^2 \Leftrightarrow \Delta_0 x = A^T \hat{y}$$

where $\Delta_0 = A^T A$ is the unnormalized graph Laplacian of G.

- In theory, nearly linear algorithms for such equations, e.g. Spielman-Teng'04, Koutis-Miller-Peng'12, etc.
- But in practice? ...

Online HodgeRank as Stochastic Approximations

Robbins-Monro (1951) algorithm for $\bar{A}x = \bar{b}$

$$x_{t+1} = x_t - \gamma_t(A_t x_t - b_t), \quad \mathbb{E}(A_t) = \bar{A}, \; \mathbb{E}(b_t) = b$$

Now consider $\Delta_0 x = \delta_0^* \hat{y}$, with new rating $y_t(i_{t+1}, j_{t+1})$

$$\begin{aligned} x_{t+1}(i_{t+1}) &= x_t(i_{t+1}) - \gamma_t[x_t(i_{t+1}) - x_t(j_{t+1}) - y_t(i_{t+1}, j_{t+1})] \\ x_{t+1}(j_{t+1}) &= x_t(j_{t+1}) + \gamma_t[x_t(i_{t+1}) - x_t(j_{t+1}) - y_t(i_{t+1}, j_{t+1})] \end{aligned}$$

Note:

- updates only occur locally on edge $\{i_{t+1}, j_{t+1}\}$
- initial choice: $s_0 = 0$ or any vector $\sum_i x_0(i) = 0$
- step size
  - $\gamma_t = a(t+b)^{-\theta}$ $(\theta \in (0, 1])$
  - $\gamma_t = const(T)$, .e.g. $1/T$ where T is total sample size

Minimax Optimal Convergence Rates

- Choose $\gamma_t \sim t^{-1/2}$ (e.g. a=$1/\lambda_1(\Delta_0)$ and b large enough)
- In this case, s_t converges to s^* (population solution) in the (optimal) rate of t

$$\mathbb{E}\|x_t - x^*\|^2 \leq O\left(t^{-1} \cdot \lambda_2^{-2}(\Delta_0)\right)$$

where $\lambda_2(\Delta_0)$ is the Fiedler Value of graph Laplacian

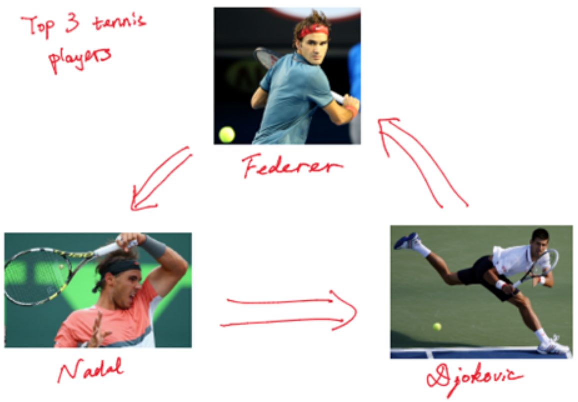
- Using Tong Zhang's stochastic variance reduction gradient (SVRG)

$$\mathbb{E}\|x_t - x^*\|^2 \leq O\left(t^{-1} + \lambda_2^{-2}(\Delta_0) t^{-2}\right)$$

Condorcet Profile splits into Local vs. Global Cycles

Residues $\hat{y}^{(c)} = B^T z$ and $\hat{y}^{(h)} = w$ are cyclic rankings, accounting for conflicts of interests:

- $\hat{y}^{(c)}$, the local/triangular inconsistency, triangular curls (Z_3-invariant)
  - $\hat{y}^{(c)}_{ij} + \hat{y}^{(c)}_{jk} + \hat{y}^{(c)}_{ki} \neq 0$, $\{i,j,k\} \in T$



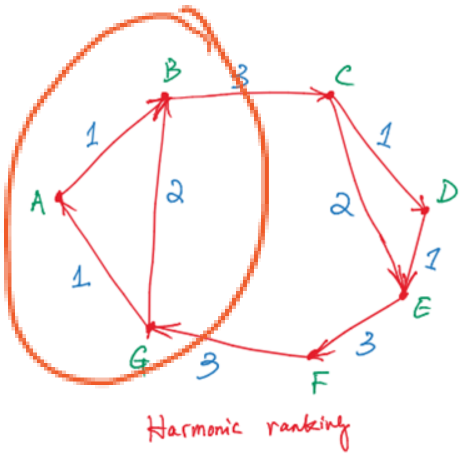
Condorcet Profile in Harmonic Ranking

- $\hat{y}^{(h)} = w$, the global inconsistency, harmonic ranking (Z_n-invariant)

$$\hat{y}_{ij}^{(h)} + \hat{y}_{jk}^{(h)} + \hat{y}_{ki}^{(h)} = 0, \quad \text{for each } \{i,j,k\} \in T, \tag{3a}$$

$$\sum_{j \sim i} \omega_{ij} \hat{y}_{ij}^{(h)} = 0, \quad \text{for each } i \in V. \tag{3b}$$

- voting chaos: *circular coordinates on* $V \Rightarrow$ *fixed tournament* issue



Cyclic Ranking and Outliers [Xu-Xiong-Huang-Y.'13]

- Robust ranking can be formulated as a Huber's LASSO problem (Gannaz'07, She-Owen'09, Fan-Tang-Shi'12)
- Sparse outliers are sparse approximation of cyclic rankings (curl+harmonic)
- Bregman ISS or linearized Bregman iterations: efficient algorithms
- Exact recovery is possible without Gaussian noise
- Outlier detection is possible against Gaussian noise, provided
  - Irrepresentable condition (e.g. random graph)
  - Outliers have large enough magnitudes

Topological Obstructions

Two topological conditions are important:

- Connectivity:
  - G is connected $\Rightarrow$ unique global ranking is possible;
- Loop-free:
  - for cyclic rankings, consider clique complex $\chi_G^2 = (V, E, T)$ by attaching triangles $T = \{(i, j, k)\}$
  - $\dim(\ker(\Delta_1)) = \beta_1(\chi_G^2)$, so harmonic ranking $w = 0$ if χ_G^2 is loop-free, here topology plays a role of obstruction of fixed-tournament
  - "Triangular arbitrage-free implies arbitrage-free"

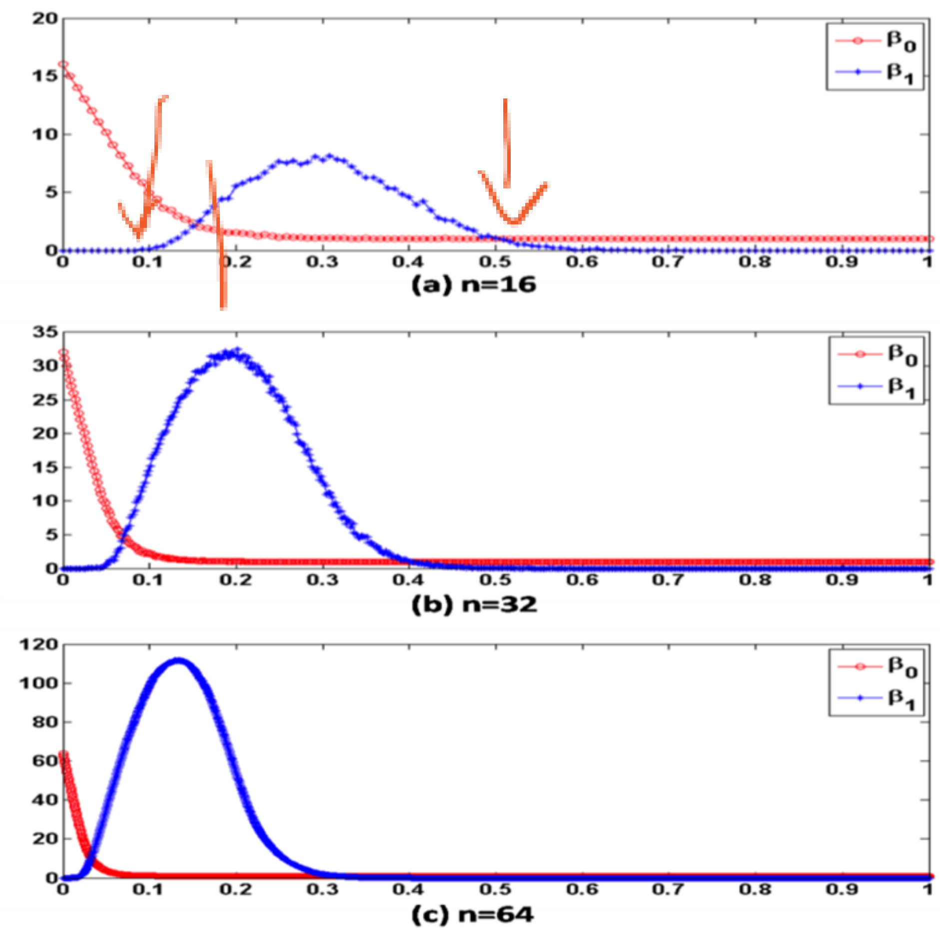
Random Graph Models for Crowdsourcing

- Recall that in crowdsourcing ranking on internet,
  - unspecified raters compare item pairs randomly
  - online, or sequentially sampling
- random graph models for experimental designs
  - P a distribution on random graphs, invariant under permutations (relabeling)
  - Generalized de Finetti's Theorem [Aldous 1983, Kallenberg 2005]: $P(i,j)$ (P ergodic) is an uniform mixture of

$$h(u,v) = h(v,u) : [0,1]^2 \to [0,1],$$

h unique up to sets of zero-measure
  - Erdös-Rényi: $P(i,j) = P(edge) = \int_0^1 \int_0^1 h(u,v)dudv =: p$
  - edge-independent process (Chung-Lu'06)

Phase Transitions in Erdös-Rényi Random Graphs



Phase Transitions of Large Random Graphs

For an Erdos-Renyi random graph $G(n,p)$ with n vertices and each edge independently emerging with probability $p(n)$,

- (Erdös-Rényi 1959) One phase-transition for β_0
  - $p << 1/n^{1+\epsilon}$ $(\forall \epsilon > 0)$, almost always disconnected
  - $p >> log(n)/n$, almost always connected
- (Kahle 2009) Two phase-transitions for β_k $(k \geq 1)$
  - $p << n^{-1/k}$ or $p >> n^{-1/(k+1)}$, almost always β_k vanishes;
  - $n^{-1/k} << p << n^{-1/(k+1)}$, almost always β_k is nontrivial

For example: with $n = 16$, 75% distinct edges included in G, then χ_G with high probability is connected and loop-free. In general, $O(n\log(n))$ samples for connectivity and $O(n^{3/2})$ for loop-free.

Some reference

- Random graph sampling models: Erdös-Rényi and beyond
  - Xu, Jiang, Yao, Huang, Yan, and Lin, *ACM Multimedia*, 2011, IEEE Trans Multimedia, 2012
- Online algorithms
  - Xu, Huang, and Yao, *ACM Multimedia* 2012
- l_1-norm ranking
  - Osting, Darbon, and Osher, 2012
- Robust ranking: Huber's Lasso
  - Xiong, Cheng, and Yao, 2013
  - Xu, Xiong, Huang, and Yao, *ACM Multimedia* 2013
- Active sampling
  - Osting, Brune, and Osher, *ICML* 2013
  - Osting, Xiong, Xu, and Yao, 2014

Multiple Utility Flows for Games

	O	F
O	3, 2	0, 0
F	0, 0	2, 3

(a) Battle of the sexes

$$\begin{tikzcd} (O,O) & (O,F) \arrow[l, "2"'] \arrow[d, "2"] \\ (F,O) \arrow[u, "3"] \arrow[r, "3"] & (F,F) \end{tikzcd}$$

Extension to multiplayer games: $G = (V, E)$

- $V = \{(x_1, \ldots, x_n) =: (x_i, x_{-i})\} = \prod_{i=1}^{n} S_i$, n person game;
- undirected edge: $\{(x_i, x_{-i}), (x_i', x_{-i})\} = E$
- each player has utility function $u_i(x_i, x_{-i})$;
- Edge flow (1-form): $u_i(x_i, x_{-i}) - u_i(x_i', x_{-i})$

Nash and Correlated Equilibrium

$\pi(x_i, x_{-i})$, a joint distribution tensor on $\prod_i S_i$, satisfies $\forall x_i, x_i'$,

$$\sum_{x_{-i}} \pi(x_i, x_{-i})(u_i(x_i, x_{-i}) - u_i(x_i', x_{-i})) \geq 0,$$

i.e. expected flow ($\mathbb{E}[\cdot|x_i]$) is nonnegative. Then,

- tensor π is a correlated equilibrium (CE, Aumann 1974);
- if π is a rank-one tensor,

$$\pi(x) = \prod_i \mu(x_i),$$

  then it is a Nash equilibrium (NE, Nash 1951);
- fully decided by the edge flow data.

Hodge Decomposition of Finite Games

Theorem (Candogan-Menache-Ozdaglar-Parrilo,2011)

Every finite game admits a unique decomposition:

$$\textit{Potential Games} \oplus \textit{Harmonic Games} \oplus \textit{Neutral Games}$$

Furthermore:

- Shapley-Monderer Condition: Potential games $\equiv$ quadrangular-curl free
- Extending $G = (V, E)$ to complex by adding quadrangular cells, harmonic games can be further decomposed into (quadrangular) curl games

Bimatrix Games

For bi-matrix game (A, B),

- potential game is decided by $((A + A')/2, (B + B')/2)$
- harmonic game is zero-sum $((A - A')/2, (B - B')/2)$
- Computation of Nash Equilibrium:
  - each of them is tractable
  - however direct sum is NP-hard
  - approximate potential game leads to approximate NE

What Does Hodge Decomposition Tell Us?

Christos Papadimitriou: best response players might experience

transient potential games + periodic equilibrium

Summary

Hodge Decomposition for Social Choice:

- Generalized Borda Count
- Borda profile in gradient flow $\Rightarrow$ global ranking or utility function
- Condorcet profile in cyclic ranking, triangular cyclic or harmonic rankings

for Game theory with multiple utility functions:

- Potential games in gradient flow
- Harmonic games in cycles
- CE and NE are preserved, tractable in some settings

in Computer Vision: optical flow decomposition, subjective visual attributes, and more are coming ...