Applied Hodge Theory: Social Choice, Crowdsourced Ranking, and Game Theory

Yuan Yao

HKUST

April 24, 2017

Topological & Geometric Data Analysis

- Differential Geometric methods: manifolds
  - data manifold: manifold learning/NDR, etc.
  - model manifold: information geometry (high-order efficiency for parametric statistics), Grassmannian, etc.
- Algebraic Geometric methods: polynomials/varieties
  - data: tensor, Sum-Of-Square (MDS, polynom. optim.), etc
  - model: algebraic statistics
- Algebraic Topological methods: complexes (graphs, etc.)
  - persistent homology
  - *Euler calculus
  - Hodge theory (a bridge between geometry and topology via optimization/spectrum)

1. Preference Aggregation and Hodge Theory
   - Social Choice and Impossibility Theorems
   - A Possible: Saari Decomposition and Borda Count
   - HodgeRank: generalized Borda Count
2. Hodge Decomposition of Pairwise Ranking
   - Hodge Decomposition
   - Combinatorial Hodge Theory on Simplicial Complexes
   - Robust Ranking
   - From Social to Personal
3. Random Graphs
   - Phase Transitions in Topology
   - Fiedler Value Asymptotics
4. Game Theory and Others
   - Game Theory: Multiple Utilities
   - Hodge Decomposition of Finite Games

Social Choice Problem

The fundamental problem of preference aggregation:

How to aggregate preferences
which faithfully represent individuals?

Crowdsourcing QoE evaluation of Multimedia

Figure: Crowdsouring subjective Quality of Experience evaluation (Xu-Huang-Y., et al. *ACM-MM* 2011)

Crowdsourced ranking

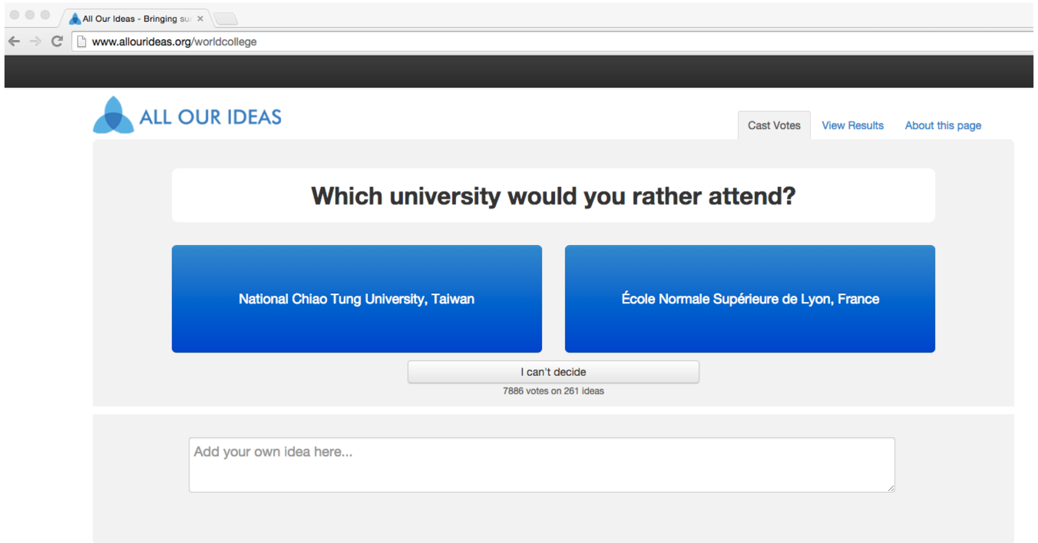


Figure: Left: www.allourideas.org/worldcollege (Prof. Matt Salganik at Princeton); Right: www.crowdrank.net.

Learning relative attributes: age

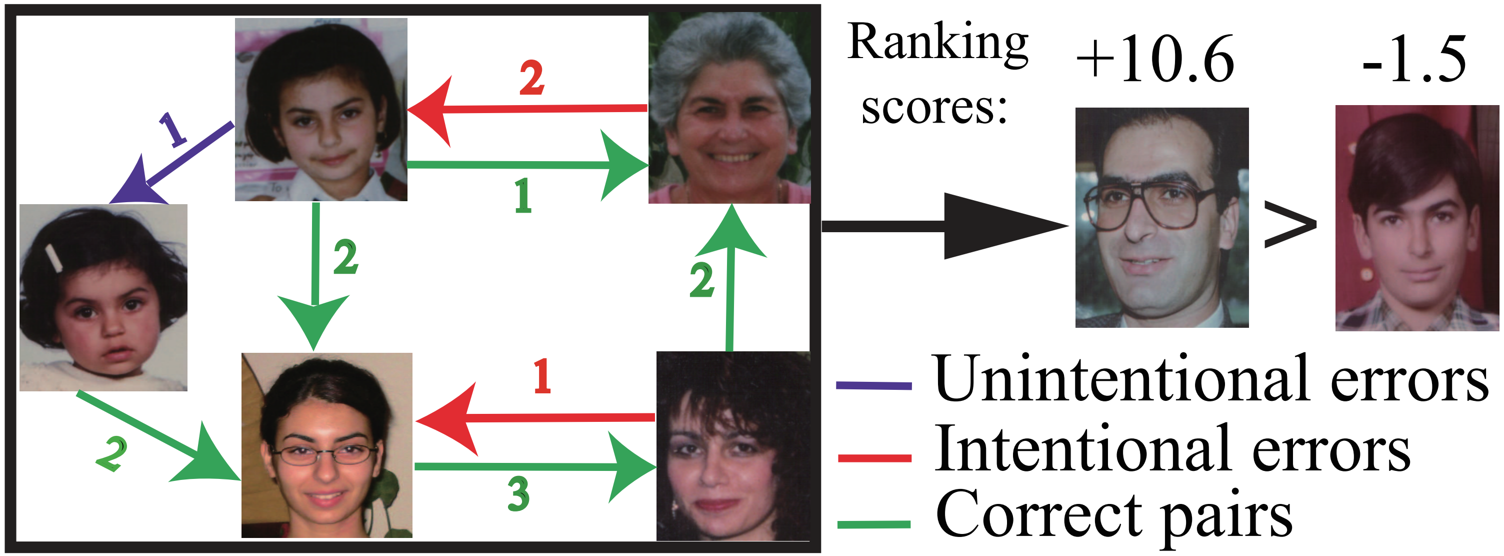


Figure: Age: a relative attribute estimated from paired comparisons (Fu-Hospedales-Xiang-Gong-Y. *ECCV*, 2014)

Netflix Customer-Product Rating

Example (Netflix Customer-Product Rating)

- 480189-by-17770 customer-product 5-star rating matrix X with $X_{ij} = \{1, \ldots, 5\}$
- X contains 98.82% missing values

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

Drug Sensitivity Ranking

Example (Drug Sensitivity Data)

- 300 drugs
- 940 cell lines, with ≈ 1000 genetic features
- sensitivity measurements in terms of IC50 and AUC
- heterogeneous missing values

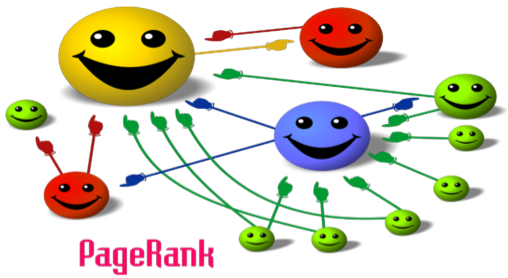
However,

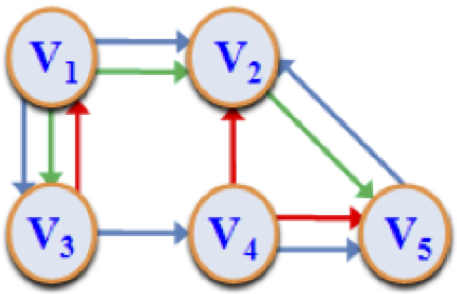
- every two drug d_1 and d_2 has been tested at least in one cell line, hence comparable (which is more sensitive)
- complete graph of paired comparisons: $G = (V, E)$
- imbalanced: number of raters on $e \in E$ varies

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 or Preference Aggregation

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

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

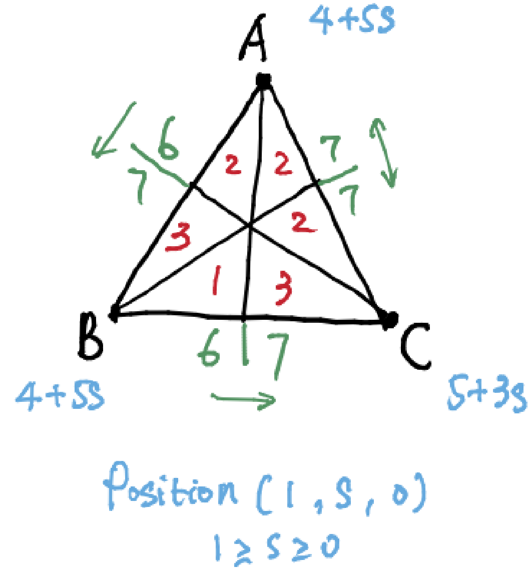
  - Borda Count: $s : V \to \mathbb{R}$ is given by $(n-1, n-2, \ldots, 1, 0)$
  - Vote-for-top-1: $(1, 0, \ldots, 0)$
  - Vote-for-top-2: $(1, 1, 0, \ldots, 0)$

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

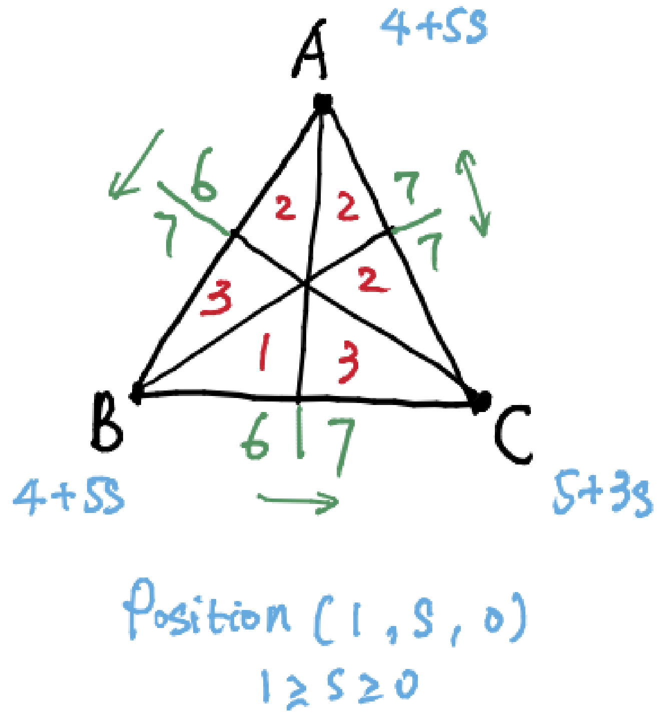
Preference order	Votes
$A \succeq B \succeq C$	2
$B \succeq A \succeq C$	3
$B \succeq C \succeq A$	1
$C \succeq B \succeq A$	3
$C \succeq A \succeq B$	2
$A \succeq C \succeq B$	2



Voting chaos!

- 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

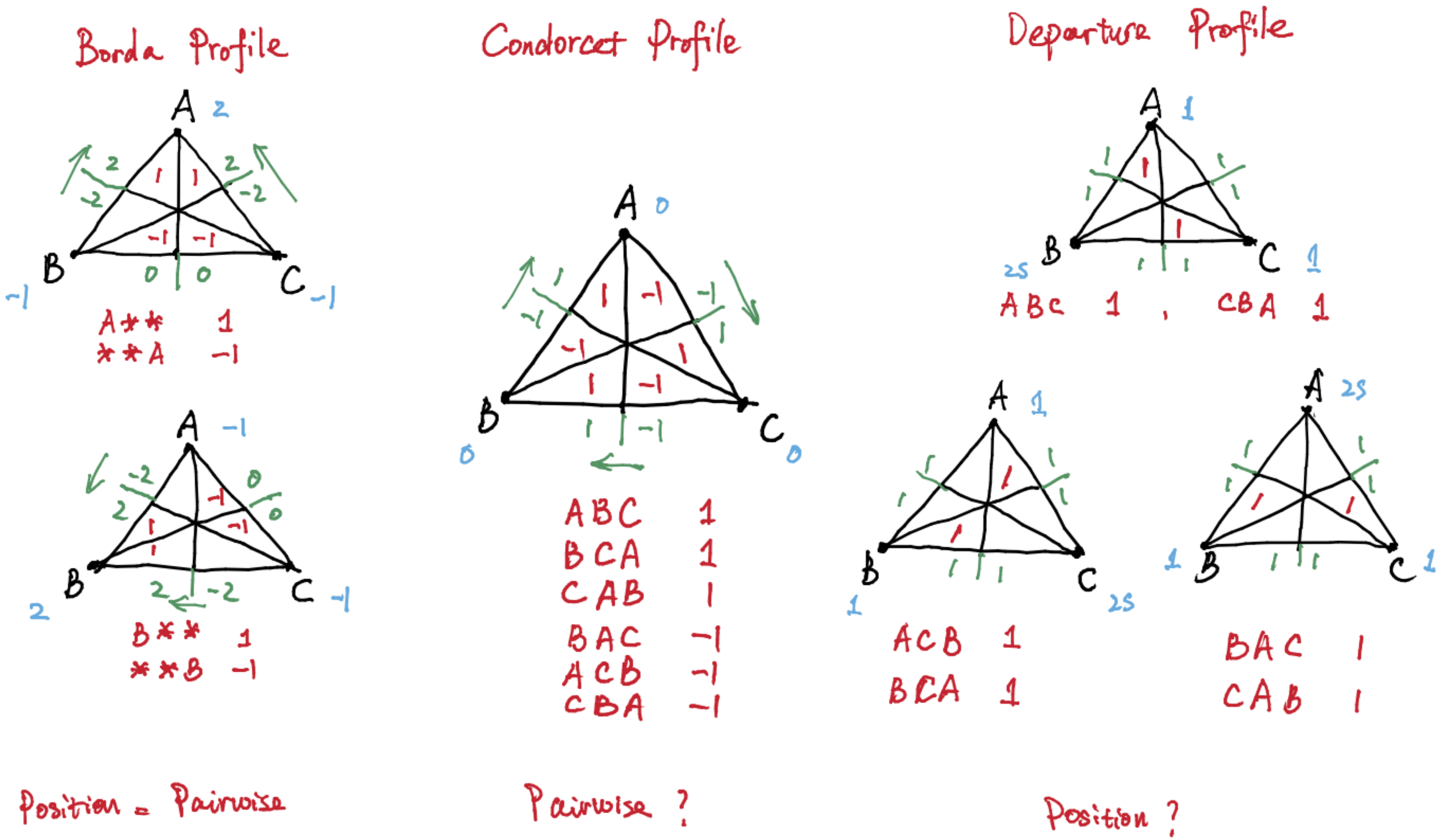
- Pareto: if all voters agree that $A > B$ then such a preference should appear in the social order
- Minimal Liberalism: two distinct voters decide social orders of two distinct pairs respectively

A Possibility: Saari's Profile Decomposition

Every voting profile, as distributions 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

Example: Decomposition of Voting Profile $R^{3!}$



Borda Count: the most consistent rule?

Table: Invariant subspaces of social rules (-)

	Borda Profile	Condorcet	Departure
Borda Count	consistent	-	-
Pairwise	consistent	inconsistent	-
Position (non-Borda)	consistent	-	inconsistent

- So, if you look for a best possibility from impossibility, Borda count is perhaps the choice
- Borda Count is the projection onto the Borda Profile subspace

Equivalently, Borda Count is a Least Square

Borda Count is equivalent to

$$\min_{\beta \in \mathbb{R}^{|V|}} \sum_{\alpha, \{i,j\} \in E} \omega_{ij}^{\alpha} (\beta_i - \beta_j - Y_{ij}^{\alpha})^2,$$

where

- E.g. $Y_{ij}^{\alpha} = 1$, if $i \succeq j$ by voter α, and $Y_{ij}^{\alpha} = -1$, on the opposite.
- Note: NP-hard ($n > 3$) Kemeny Optimization, or Minimimum-Feedback-Arc-Set:

$$\min_{s \in \mathbb{R}^{|V|}} \sum_{\alpha, \{i,j\} \in E} \omega_{ij}^{\alpha} (\mathrm{sign}(\beta_i - \beta_j) - \hat{Y}_{ij}^{\alpha})^2$$

Generalized Borda Count with Incomplete Data

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

$$\Leftrightarrow$$

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

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 $\hat{y} \in l^2_\omega(E)$, 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? Hodge Decomposition in Linear Algebra

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

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

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

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

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

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

Laplacian

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

Hence, in our case

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

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

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

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

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 $\sigma \in \chi(X)$
- k-forms or cochains as alternating functions

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

- 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

- $G = (X, E)$ 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:

  $$z_{ijk} = z_{jki} = z_{kij} = -z_{jik} = -z_{ikj} = -z_{kji}$$
- coboundary operators as alternating difference operators:
  - $(d_0 v)(i,j) = v_j - v_i =: (\mathrm{grad}\, v)(i,j)$
  - $(d_1 w)(i,j,k) = (\pm)(w_{ij} + w_{jk} + w_{ki}) =: (\mathrm{curl}\, w)(i,j,k)$
- $d_1 \circ d_0 = \mathrm{curl}(\mathrm{grad}\, u) = 0$

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 $\Omega^k(X)$
  - $\Omega^k(X) = \mathrm{im}(d_{k-1}) \oplus \ker(\Delta_k) \oplus \mathrm{im}(\delta_k)$
  - $\dim(\ker(\Delta_k)) = \beta_k(\chi(X))$, k-harmonics

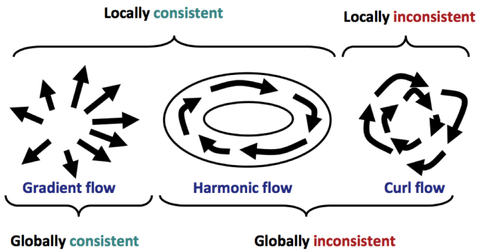


Figure: Courtesy by Asu Ozdaglar