Hodge decomposition in data analysis

Lek-Heng Lim

University of Chicago

January 13, 2014

Thanks: Vin De Silva, Sayan Mukherjee, Yuan Yao

objectives

- Hodge decomposition and cohomology in data applications
- present cohomology in the simplest possible way
- why?
  - everything should be made as simple as possible, but not simpler
  - facilitate creation of novel applications
  - facilitate communication with engineers
- examine two applications
  - ranking: web search, recommendation systems, crowd sourcing
  - game theory: ad auction, happiness index, social networks
- cursory treatment here, detailed treatment in PSAMP lecture notes

applications of Hodge theory

numerical analysis D. Arnold, R. S. Falk, and R. Winther, "Finite element exterior calculus: from Hodge theory to numerical stability," *Bull. Amer. Math. Soc.*, **47** (2010), no. 2, pp. 281–354.

peridynamics Q Du, M. Gunzburger, R. Lehoucq, and K. Zhou, "A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws," *Math. Models Methods Appl. Sci.*, **23** (2013), no. 3, pp. 493–540.

topological data analysis V. De Silva, D. Morozov, and M. Vejdemo-Johansson, "Persistent cohomology and circular coordinates," *Discrete Comput. Geom.*, **45** (2011), no. 4, pp. 737–759.

computational topology J. Friedman, "Computing Betti numbers via combinatorial Laplacians," *Algorithmica*, **21** (1998), no. 4, pp. 331–346.

graphics Y. Tong, S. Lombeyda, A. Hirani, and M. Desbrun, "Discrete multiscale vector field decomposition," *ACM Trans. Graph.*, **22** (2003), no. 3, pp. 445–452.

image processing S. Yu, "Angular embedding: a robust quadratic criterion," *IEEE Trans. Pattern Anal. Mach. Intell.*, **34** (2012), no. 1, pp. 158–173.

applications of Hodge theory

ranking	X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye, "Statistical ranking and combinatorial Hodge theory," *Math. Program.*, **127** (2011), no. 1, pp. 203–244.
game theory	O. Candogan, I. Menache, A. Ozdaglar, and P. Parrilo, "Flows and decompositions of games: harmonic and potential games," *Math. Oper. Res.*, **36** (2011), no. 3, pp. 474–503.
multimedia	Q. Xu, Q. Huang, T. Jiang, B. Yan, Y. Yao, and W. Lin, "HodgeRank on random graphs for subjective video quality assessment," *IEEE Trans. Multimed.*, **14** (2012), no. 3, pp. 844–857.
robotics	P. Kingston and M. Egerstedt, "Distributed-infrastructure multi-robot routing using a Helmholtz-Hodge decomposition," *Proc. IEEE Conf. Decis. Control Eur. Control Conf.*, **50** (2011), pp. 5281–5286.
sensor networks	A. Tahbaz-Salehi and A. Jadbabaie, "Distributed coverage verification in sensor networks without location information," *IEEE Trans. Automat. Control*, **55** (2010), no. 8, pp. 1837–1849.
neuroscience	W. Ma, J.-M. Morel, S. Osher, and A. Chien, "An L_1-based variational model for retinex theory and its applications to medical images," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, (2011), pp. 153–160.

Cohomology for Pedestrians

explaining cohomology

question: how would you explain cohomolgy to an engineer?

my answer: de Rham cohomology and Hodge decomposition

- all engineers know basic fluid mechanics and electrodynamics
- more generally, they know vector calculus
- they probably watch movies too

de Rham cohomology

$$V = \{F : \mathbb{R}^3 \backslash X \to \mathbb{R}^3 \mid \nabla \times F = 0\}; \quad W = \{F = \nabla g\}; \quad \dim(V/W) = ?$$

Laplace equation

- Laplace or homogeneous Poisson equation in $\mathbb{R}^3$:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

- more generally

$$\Delta f = \nabla^2 f = \nabla \cdot \nabla f = \operatorname{div} \operatorname{grad} f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

electrostatics: electric potential in free space with no charge

fluid mechanics: velocity potential of incompressible fluid

thermal conduction: stationary heat equation without a heat source

Laplace equation to cohomology

- usually stated as boundary value problem

$$\begin{cases} \Delta f = 0 & \text{in } \Omega \\ f = g & \text{on } \partial\Omega \end{cases} \quad \text{or} \quad \begin{cases} \Delta f = 0 & \text{in } \Omega \\ \partial f / \partial n = g & \text{on } \partial\Omega \end{cases}$$

- tells us about the topology and geometry of Ω
- solution $f : \Omega \to \mathbb{R}$ called harmonic function
- 0-cohomolgy is the study of solutions to Laplace equation with no boundary conditions
- 0-cohomology class is harmonic function
- 0-cohomology group is set of all harmonic functions

vector Laplace equation

- vector Laplace or homogeneous vector Poisson equation in $\mathbb{R}^3$:

$$\begin{cases} -\operatorname{grad}\operatorname{div} f + \operatorname{curl}\operatorname{curl} f = 0 & \text{in } \Omega \\ f \cdot n = 0, \quad \operatorname{curl} f \times n = 0 & \text{on } \partial\Omega \end{cases}$$

- Helmholtz operator or vector Laplacian in $\mathbb{R}^3$

$$\Delta_1 f = \operatorname{curl}\operatorname{curl} f - \operatorname{grad}\operatorname{div} f = \nabla(\nabla \cdot f) - \nabla \times (\nabla \times f) = \nabla^2 f$$

- solution $f : \Omega \to \mathbb{R}^3$ is vector field on Ω, call this harmonic 1-form
- 1-cohomolgy is the study of solutions to vector Laplace equation with no boundary conditions
- 1-cohomology class is harmonic 1-form
- 1-cohomology group is set of all harmonic 1-form

cohomology for pedestrians

- 0th cohomology classes are solutions to scalar Laplace equation

$$H^0(\Omega) = \ker(\Delta) = \{f : \Delta f = 0\}$$

- 1th cohomology classes are solutions to vector Laplace equation

$$H^1(\Omega) = \ker(\Delta_1) = \{f : \Delta_1 f = 0\}$$

- for $k = 0$

$$\Delta_0 = \operatorname{div} \operatorname{grad}$$

  is our usual Laplace operator or scalar Laplacian Δ
- for $k = 1$

$$\Delta_1 = -\operatorname{grad} \operatorname{div} + \operatorname{curl} \operatorname{curl}$$

  is our usual Helmholtz operator or vector Laplacian
- but we may also define 'higher-order Laplacians'

$$\Delta_k = \delta_{k-1}\delta_{k-1}^* + \delta_k^*\delta_k$$

three approaches

ordinary: $\delta_k : C^k(X) \to C^{k+1}(X)$ coboundary operators,

$$H^k(X) = \ker(\delta_k)/\operatorname{im}(\delta_{k-1})$$

generalized: $\{(E_k, \varepsilon_k) \mid k \in \mathbb{Z}\}$ spectrum,

$$H^k(X) = [X, E_k]$$

Hodge: $\Delta_k = \delta_{k-1}\delta_{k-1}^* + \delta_k^*\delta_k$ combinatorial Laplacian,

$$H^k(X) = \ker(\Delta_k)$$

Hodge approach

cons
- not functorial
- metric dependent
- does not work over rings or fields of positive characteristics

pros
- each cohomology class has unique harmonic representative
- works in noisy setting: eigenfuctions of Δ_k with small eigenvalues
- easily accessible to practioners

easy to visualize

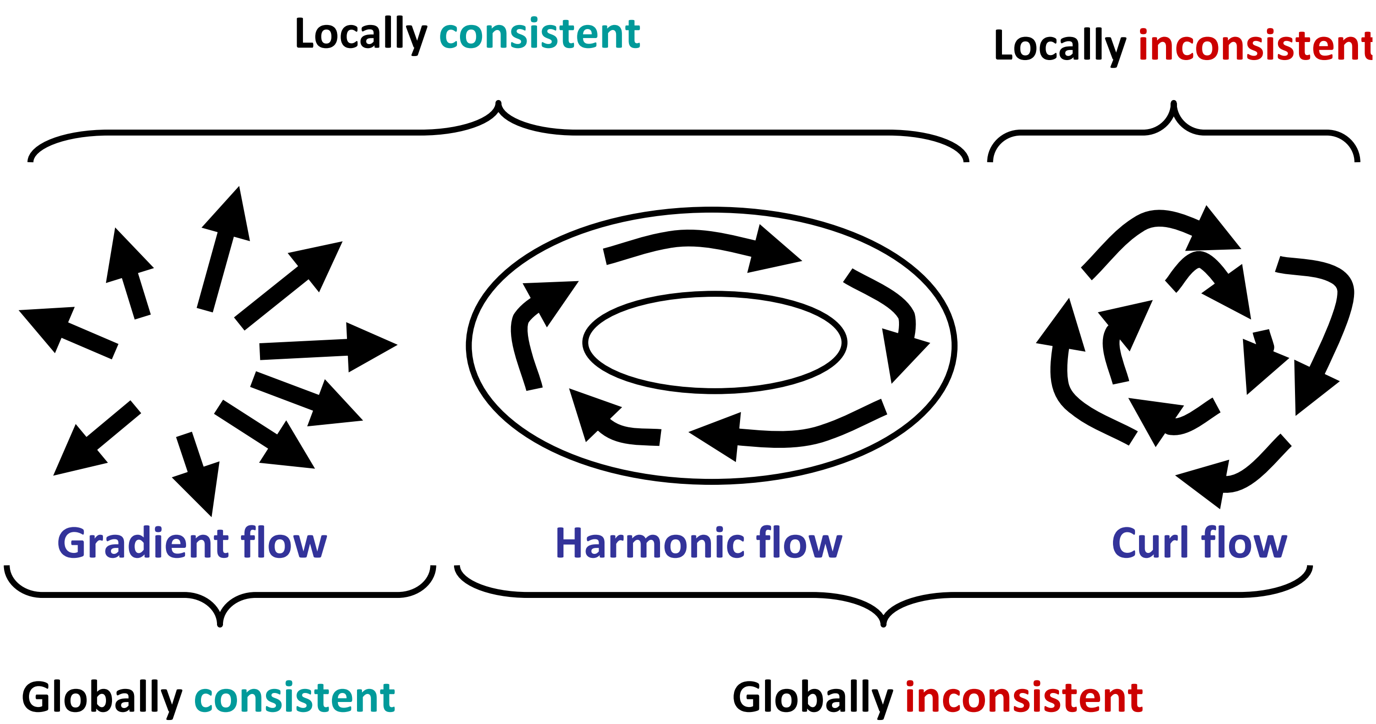


Figure : cartoon of Hodge decomposition (courtesy of Pablo Parrilo)

easy to apply

- fluid mechanics

$$\textit{fluid flow} = \textit{irrotational} \oplus \textit{solenoidal} \oplus \textit{harmonic}$$

- ranking

$$\textit{pairwise ranking} = \textit{consistent} \oplus \textit{locally inconsistent} \oplus \textit{globally inconsistent}$$

- games

$$\textit{multiplayer game} = \textit{potential game} \oplus \textit{nonstrategic game} \oplus \textit{harmonic game}$$

comes in several flavors

differentiable F. W. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, **94**, Springer-Verlag, New York, NY, 1983.

continuous L. Bartholdi, T. Schick, N. Smale, and S. Smale, "Hodge theory on metric spaces," *Found. Comput. Math.*, **12** (2012), no. 1, pp. 1–48.

discrete J. Dodziuk, "Combinatorial and continuous Hodge theories," *Bull. Amer. Math. Soc.*, **80** (1974), no. 5, pp. 1014–1016.

graph theoretic X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye, "Statistical ranking and combinatorial Hodge theory," *Math. Program.*, **127** (2011), no. 1, pp. 203–244.

why important

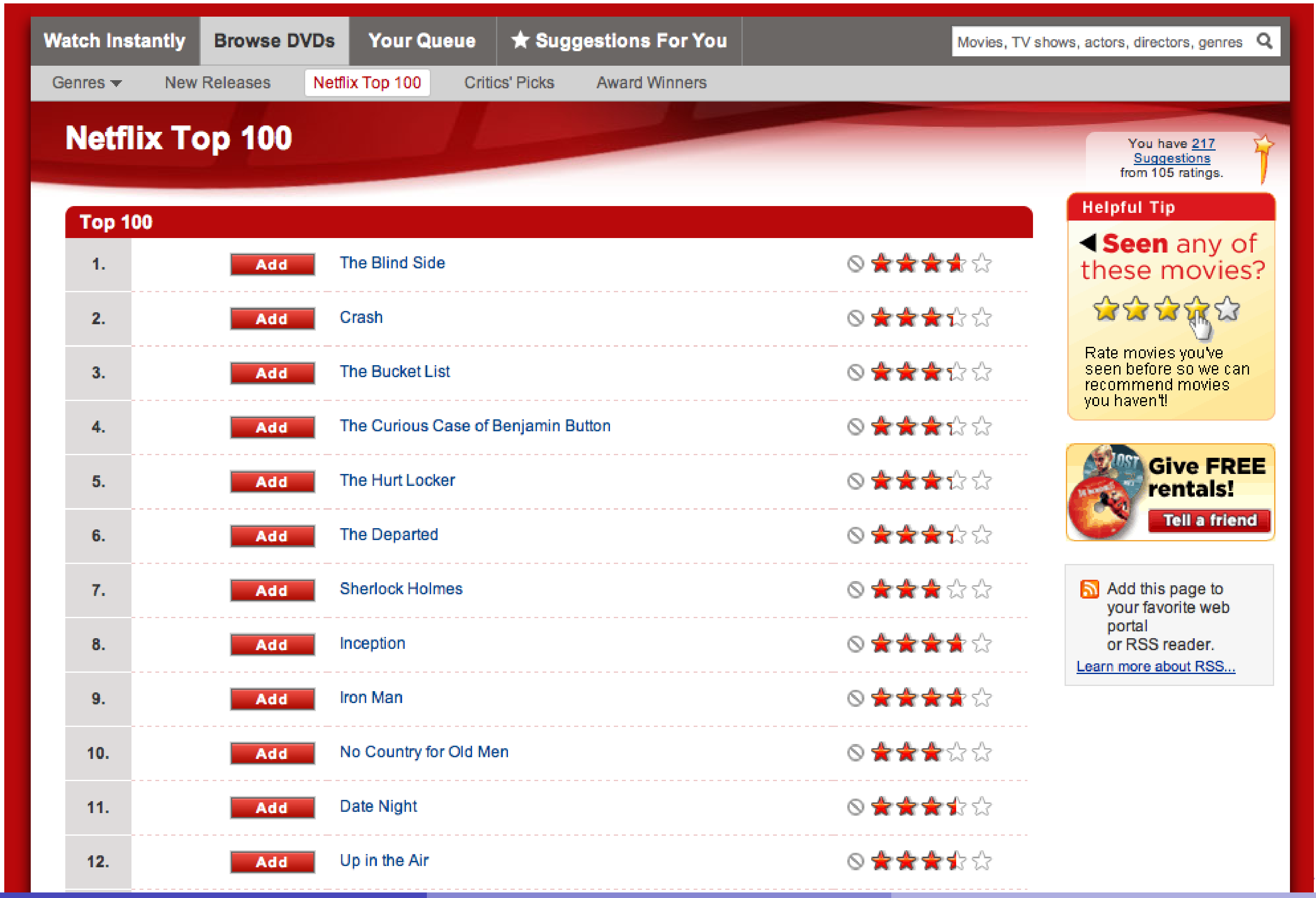
- unlike physical problems (e.g. continuum mechanics or electromagnetics), problems coming from data analytic applications are far less structured
- often one could at most assume (i) some notion of proximity of data points and (ii) some knowledge of the distribution of the data set
- i.e., continuous, discrete, graph theoretic versions usually more useful for data applications

Ranking

ranking problems

- **static ranking** (**Google problem**)
  - alternatives: football teams, websites
  - one voter: entire season of games, hyperlink structure of WWW
  - one ranking: number of matches won by each team, PageRank of each website
  - no paradox, impossibility, chaos, NP-hardness
- **collaborative filtering** (**the better known Netflix problem**)
  - alternatives: movies, drugs
  - many voters: movie viewers, patients
  - many rankings: ideally one for each viewer, patient
  - no paradox, impossibility, chaos, NP-hardness
- **rank aggregation** (**our Netflix problem**)
  - alternatives: colleges, candidates
  - many voters: academics surveyed, electorate
  - one ranking: order all alternatives globally
  - Condorcet's paradox, Arrow's impossibilty, McKelvey's chaos, NP-hard

the Netflix problem in this talk



rank aggregation

- many voters, each rated a few alternatives, want global ranking
- averaging over scores doesn't work: one movie receives one 5☆ and no other ratings, another receives 10,000 5☆ and one 4☆
- should be invariant under monotone transformation:

$$1☆, \ldots, 5☆ \quad \longrightarrow \quad 0☆, \ldots, 4☆$$

- **basic unit of ranking:** pairwise comparison or pairwise ranking
- take average over pairwise rankings instead, get $Y \in \mathbb{R}^{17770 \times 17770}$
- for Netflix data, user-product rating matrix $Z \in \mathbb{R}^{480189 \times 17770}$ has 98.82% missing values, Y has 0.22% missing values

linear model: average score difference between i and j over all who have rated both,

$$y_{ij} = \frac{\sum_h (z_{hj} - z_{hi})}{\#\{h \mid z_{hi}, z_{hj} \text{ exist}\}}$$

invariant under translation

averaging over pairwise rankings

log-linear model: logarithmic average score ratio of positive scores,

$$y_{ij} = \frac{\sum_h (\log z_{hj} - \log z_{hi})}{\#\{h \mid z_{hi}, z_{hj} \text{ exist}\}}$$

invariant up to a multiplicative constant

linear probability model: probability j preferred to i in excess of purely random choice,

$$y_{ij} = \Pr\{h \mid z_{hj} > z_{hi}\} - \frac{1}{2}$$

invariant under monotone transformation

Bradley-Terry model: logarithmic odd ratio (logit),

$$y_{ij} = \log \frac{\Pr\{h \mid z_{hj} > z_{hi}\}}{\Pr\{h \mid z_{hj} < z_{hi}\}}$$

invariant under monotone transformation

difficulties with rank aggregation

- **Condorcet's paradox:** majority vote intransitive $i \succeq j \succeq k \succeq i$ [Condorcet, 1785]
- **Arrow/Sen's impossibility:** any sufficiently sophisticated preference aggregation must exhibit intransitivity [Arrow, 1950], [Sen, 1970]
- **McKelvey/Saari's chaos:** almost every possible ordering can be realized by a clever choice of the order in which decisions are taken [McKelvey, 1979], [Saari, 1989]
- **Kemeny optimal is NP-hard:** even with just 4 voters [Dwork–Kumar–Naor–Sivakumar, 2001], quadratic assignment problem [Cook–Kress, 1984]
- **empirical evidence:** lack of consensus common in group decision making (e.g. US congress)

what we want

ordinal: intransitivity, $i \succeq j \succeq k \succeq i$

cardinal: inconsistency, $X_{ij} + X_{jk} + X_{ki} \neq 0$

- want global ranking of alternatives if a reasonable one exists
- want certificate of reliability to quantify validity of global ranking
- if no meaningful global ranking, analyze nature of inconsistencies

Graph Theoretic Hodge Theory

graph

- $G = (V, E)$ undirected graph
- V vertices
- $E \subseteq \binom{V}{2}$ edges
- $T \subseteq \binom{V}{3}$ triangles or 3-cliques, i.e.,

$$\{i, j, k\} \in T \quad \text{iff} \quad \{i, j\}, \{j, k\}, \{k, i\} \in E$$

- more generally $K_k \subseteq \binom{V}{k}$ k-cliques, i.e.,

$$\{i_1, \dots, i_k\} \in K_k \quad \text{iff} \quad \text{it is a complete subgraph of } G$$

- nonempty family K of finite subsets of a set G is abstract simplicial complex if for every set X in K, every $Y \subseteq X$ also belongs to K
- $K(G)$ clique complex of a graph G is an abstract simplicial complex

functions on graph

function on vertices: $s : V \to \mathbb{R}$

edge flows: $X : V \times V \to \mathbb{R}$, $X(i,j) = 0$ if $\{i,j\} \notin E$,

$$X(i,j) = -X(j,i) \quad \text{for all } i,j$$

triangular flows: $\Phi : V \times V \times V \to \mathbb{R}$, $\Phi(i,j,k) = 0$ if $\{i,j,k\} \notin T$,

$$\begin{aligned}\Phi(i,j,k) &= \Phi(j,k,i) = \Phi(k,i,j) \\ &= -\Phi(j,i,k) = -\Phi(i,k,j) = -\Phi(k,j,i) \quad \text{for all } i,j,k\end{aligned}$$

physics: s, X, Φ potential, alternating vector/tensor field

topology: s, X, Φ 0-, 1-, 2-cochain

ranking: s scores/utility, X pairwise rankings, Φ triplewise rankings

operators on graph

gradient: $\mathrm{grad} : L^2(V) \to L^2(E)$,

$$(\mathrm{grad}\, s)(i,j) = s_j - s_i$$

curl: $\mathrm{curl} : L^2(E) \to L^2(T)$,

$$(\mathrm{curl}\, X)(i,j,k) = X_{ij} + X_{jk} + X_{ki}$$

divergence: $\mathrm{div} : L^2(E) \to L^2(V)$,

$$(\mathrm{div}\, X)(i) = \sum\nolimits_{j:\{i,j\}\in E} w_{ij} X_{ij}$$

graph Laplacian: $\Delta_0 : L^2(V) \to L^2(V)$,

$$\Delta_0 = \mathrm{div}\,\mathrm{grad}$$

graph Helmholtzian: $\Delta_1 : L^2(E) \to L^2(E)$,

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

generalization: cochains

- K abstract simplicial complex with vertex set V
- alternating functions on $k+1$ arguments, i.e. k-forms or k-cochains:

$$C^k(K;\mathbb{R}) = \{u : V^{k+1} \to \mathbb{R} \mid u(i_{\sigma(0)}, \dots, i_{\sigma(k)}) = \operatorname{sign}(\sigma) u(i_0, \dots, i_k)\}$$

  for $(i_0, \dots, i_k) \in K_{k+1}$, where $\sigma \in \mathfrak{S}_{k+1}$
- most interesting for us $K = K(G)$, clique complex of graph G
- may put metrics/inner products on $C^k(K(G);\mathbb{R})$
- e.g. following metric on 1-forms, is useful for imbalanced ranking data:

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

  where

$$D_{ij} = \text{number of voters who rated both } i \text{ and } j$$

generalization: coboundary maps

- k-coboundary maps $\delta_k : C^k(K;\mathbb{R}) \to C^{k+1}(K;\mathbb{R})$ are

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

- **fundamental theorem of topology:** $\delta_{k+1}\delta_k = 0$
- for $k = 0$,

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

$$\text{global} \xrightarrow{\delta_0} \text{pairwise} \xrightarrow{\delta_1} \text{triplewise}$$

$$\text{global} \xleftarrow{\delta_0^*} \text{pairwise} \xleftarrow{\delta_1^*} \text{triplewise}$$

- we have $\delta_1\delta_0(\text{global rankings}) = 0$, i.e.,
  - global rankings are transitive/consistent
  - no need to consider rankings beyond triplewise

combinatorial Laplacian and Hodge theory

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

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

call u a **harmonic form** if $\Delta_k u = 0$

Theorem (Hodge)

1. $H^k(K;\mathbb{R}) = \ker(\delta_k)/\operatorname{im}(\delta_{k-1}) \cong \ker(\Delta_k)$
2. $C^k(K;\mathbb{R}) = \operatorname{im}(\delta_{k-1}) \oplus \ker(\Delta_k) \oplus \operatorname{im}(\delta_k^*)$
3. $\ker(\Delta_k) = \ker(\delta_k) \cap \ker(\delta_{k-1}^*)$

follows from fundamental theorem of topology and Fredholm alternative:

$$\mathbb{R}^n = \ker(A) \oplus \operatorname{im}(A^*), \quad \mathbb{R}^m = \ker(A^*) \oplus \operatorname{im}(A)$$

for $A \in \mathbb{R}^{m\times n}$

special case: Helmholtz decomposition

Theorem (Helmholtz decomposition for graphs)

$G = (V, E)$ graph. The space of edge flows, $C^1(\mathcal{K}(G), \mathbb{R})$, admits an orthogonal decomposition into three subspaces

$$C^1(\mathcal{K}(G), \mathbb{R}) = \mathrm{im}(\mathrm{grad}) \oplus \ker(\Delta_1) \oplus \mathrm{im}(\mathrm{curl}^*)$$

where

$$\ker(\Delta_1) = \ker(\mathrm{curl}) \cap \ker(\mathrm{div}).$$

Application to Ranking

X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye, "Statistical ranking and combinatorial Hodge theory," *Math. Program.*, **127** (2011), no. 1, pp. 203–244.

Helmholtz decomposition applied to rankings

pairwise comparison graph $G = (V, E)$; V: set of alternatives, E: pairs of alternatives compared

Theorem (Helmholtz decomposition for pairwise rankings)

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

$$C^1(\mathcal{K}(G), \mathbb{R}) = \operatorname{im}(\operatorname{grad}) \oplus \ker(\Delta_1) \oplus \operatorname{im}(\operatorname{curl}^*)$$

where

$$\ker(\Delta_1) = \ker(\operatorname{curl}) \cap \ker(\operatorname{div}).$$

our approach: HodgeRank

- Hodge decomposition of ranking:

$$\textit{aggregate pairwise ranking} = \\ \textit{consistent} \oplus \textit{locally inconsistent} \oplus \textit{globally inconsistent}$$

- consistent component gives global ranking
- total size of inconsistent components gives certificate of reliability
- local and global inconsistent components tell us about nature of inconsistencies
- quantifies Condorcet paradox, Arrow's impossibility, McKelvey chaos, etc
- numerical, not combinatorial, so not NP-hard

properties

- im(grad): pairwise rankings that are gradient of score functions, i.e. consistent or *integrable*
- ker(div): $\operatorname{div} X(i)$ measures the inflow-outflow sum at i; $\operatorname{div} X(i) = 0$ implies alternative i is preference-neutral in all pairwise comparisons; i.e. inconsistent rankings of the form $a \succeq b \succeq c \succeq \cdots \succeq a$
- ker(curl): pairwise rankings with zero flow-sum along any triangle
- $\ker(\Delta_1) = \ker(\mathrm{curl}) \cap \ker(\mathrm{div})$: globally inconsistent or *harmonic* rankings; no inconsistencies due to small loops of length 3, i.e. $a \succeq b \succeq c \succeq a$, but inconsistencies along larger loops of lengths > 3
- $\mathrm{im}(\mathrm{curl}^*)$: locally inconsistent rankings; non-zero curls along triangles
- $\operatorname{div}\operatorname{grad}$ is vertex Laplacian, $\operatorname{curl}\operatorname{curl}^*$ is edge Laplacian

analyzing inconsistencies

- locally inconsistent rankings should be acceptable
  - inconsistencies in items ranked closed together but not in items ranked far apart
  - ordering of 4th, 5th, 6th ranked items cannot be trusted but ordering of 4th, 50th, 600th ranked items can
  - e.g. no consensus for hamburgers, hot dogs, pizzas, and no consensus for caviar, foie gras, truffle, but clear preference for latter group
- globally inconsistent rankings ought to be rare

Theorem (Kahle, 2007)

Erdős-Rényi $G(n,p)$, n alternatives, comparisons occur with probability p, clique complex χ_G almost always have zero 1-homology, unless

$$\frac{1}{n^2} \ll p \ll \frac{1}{n}.$$

relates to Kemeny optimum

- ranking data live on **pairwise comparison graph** $G = (V, E)$; V: set of alternatives, E: pairs of alternatives compared
- optimize over model class $\mathcal{M}$

$$\min_{X \in \mathcal{M}} \sum\nolimits_{\alpha,i,j} w_{ij}^{\alpha} (X_{ij} - Y_{ij}^{\alpha})^2$$

- Y_{ij}^{α} measures preference of i over j of voter α. Y^{α} skew-symmetric
- w_{ij}^{α} metric; 1 if α made comparison for $\{i, j\}$, 0 otherwise
- Kemeny optimization:

$$\mathcal{M}_K = \{X \in \mathbb{R}^{n \times n} \mid X_{ij} = \operatorname{sign}(s_j - s_i),\ s : V \to \mathbb{R}\}$$

- relaxed version

$$\mathcal{M}_G = \{X \in \mathbb{R}^{n \times n} \mid X_{ij} = s_j - s_i,\ s : V \to \mathbb{R}\}$$

- rank-constrained least squares with skew-symmetric matrix variables
- solution is precisely consistent component in HodgeRank

comparisons with other methods

- **analytic hierarchy process (AHP):** take

$$a_{ij} = \begin{cases} \exp(y_{ij}) & \text{if } y_{ij} \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

  A reciprocal matrix, i.e. $a_{ji} = 1/a_{ij} > 0$. Principal eigenvector of A gives global scores [Saaty, 1978].
- **tropical AHP:** principal max-plus eigenvector of Y gives global scores [Elsner–Driessche, 2006]
- suppose $n =$ number of alternatives grows with $m =$ number of voters; when does

$$P(\text{recover top } k \text{ rankings}) \to 1 \quad \text{as } m, n \to \infty?$$

Theorem (Tran, 2013)

Under mild assumptions, HodgeRank recovers true ranking of top k items in the above sense. AHP and tropical AHP do not.

online version

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

$$\mathbf{x}_{t+1} = \mathbf{x}_t - \gamma_t(A_t\mathbf{x}_t - \mathbf{b}_t), \quad \mathbb{E}(A_t) = A, \quad \mathbb{E}(\mathbf{b}_t) = \mathbf{b}$$

now consider $\Delta_0 \mathbf{s} = \delta_0^* \hat{Y}$, with new rating $Y_t(i_{t+1}, j_{t+1})$

$$\begin{aligned} \mathbf{s}_{t+1}(i_{t+1}) &= \mathbf{s}_t(i_{t+1}) - \gamma_t[\mathbf{s}_t(i_{t+1}) - \mathbf{s}_t(j_{t+1}) - Y_t(i_{t+1}, j_{t+1})] \\ \mathbf{s}_{t+1}(j_{t+1}) &= \mathbf{s}_t(j_{t+1}) + \gamma_t[\mathbf{s}_t(i_{t+1}) - \mathbf{s}_t(j_{t+1}) - Y_t(i_{t+1}, j_{t+1})] \end{aligned}$$

note:

- updates only occur locally on edge $\{i_{t+1}, j_{t+1}\}$
- initial choice: $\mathbf{s}_0 = \mathbf{0}$ or any vector $\sum_i \mathbf{s}_0(i) = 0$
- step size
  - $\gamma_t = (t + c)^{-\theta}, \quad \theta \in (0, 1]$
  - $\gamma_t = \text{constant}(T)$, e.g. $1/T$ where T is total sample size

averaging process

a second stage averaging process, following $\mathbf{s}_{t+1}$ above

$$\mathbf{z}_{t+1} = \frac{t}{t+1}\mathbf{z}_t + \frac{1}{t+1}\mathbf{s}_{t+1}$$

with $\mathbf{z}_0 = \mathbf{s}_0$
note:

- averaging process speeds up convergence for various choices of γ_t
- one often choose $\gamma_t = c$ to track dynamics
- in this case, $\mathbf{z}_t$ converges to $\hat{\mathbf{s}}$ (population solution), with probability $1-\delta$, in the (optimal) rate

$$\|\mathbf{z}_t - \hat{\mathbf{s}}\| \leq O\left(t^{-1/2} \cdot \kappa(\Delta_0) \cdot \log^{1/2} \frac{1}{\delta}\right)$$

Application to Games

O. Candogan, I. Menache, A. Ozdaglar, and P. Parrilo, "Flows and decompositions of games: harmonic and potential games," *Math. Oper. Res.*, **36** (2011), no. 3, pp. 474–503.

noncooperative strategic-form finite game

- finite set of players $V = \{1, \dots, n\}$
- finite set of strategies E_i, for every $i \in V$
- joint strategy space is $E = \prod_{i \in V} E_i$
- utility function $u_i : E \to \mathbb{R}$, $i \in V$
- a game instance is given by the tuple $(V, \{E_i\}_{i \in V}, \{u_i\}_{i \in V})$

strategy

- $\mathbf{p}_i \in E_i$ denotes strategy of player i
- collection of players' strategies is $\mathbf{p} = \{\mathbf{p}_i\}_{i \in V}$, called strategy profile
- collection of strategies for all players but the ith one denoted by $\mathbf{p}_{-i} \in E_{-i}$
- $h_i = |E_i|$, cardinality of the strategy space of player i
- $|E| = \prod_{i=1}^{n} h_i$ for the overall cardinality of the strategy space
- enumerate the actions of players, so that $E_i = \{1, \ldots, h_i\}$

Nash equilibrium

- Nash equilibrium is strategy profile from which no player can unilaterally deviate and improve its payoff
- formally strategy profile $\mathbf{p} := \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is Nash equilibrium if

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq u_i(\mathbf{q}_i, \mathbf{p}_{-i}), \quad \text{for every } \mathbf{q}_i \in E_i \text{ and } i \in V$$

- strategy profile $\mathbf{p} := \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ is ε-equilibrium if

$$u_i(\mathbf{p}_i, \mathbf{p}_{-i}) \geq u_i(\mathbf{q}_i, \mathbf{p}_{-i}) - \varepsilon \quad \text{for every } \mathbf{q}_i \in E_i \text{ and } i \in V$$

- pure Nash equilibrium is an ε-equilibrium with $\varepsilon = 0$

potential game

- a potential game is a noncooperative game for which there exists a function $\varphi : E \to \mathbb{R}$ satisfying

$$\varphi(\mathbf{p}_i, \mathbf{p}_{-i}) - \varphi(\mathbf{q}_i, \mathbf{p}_{-i}) = u_i(\mathbf{p}_i, \mathbf{p}_{-i}) - u_i(\mathbf{q}_i, \mathbf{p}_{-i}),$$

  for every $i \in V$, $\mathbf{p}_i, \mathbf{q}_i \in E_i$, $\mathbf{p}_{-i} \in E_{-i}$
- φ is referred to as a potential function of the game
- proposed in seminal paper [Monderer–Shapley, 1996]
- widely studied in game theory
- preferences of all players aligned with a global objective
- easy to analyze
- pure Nash equilibrium exists

harmonic games

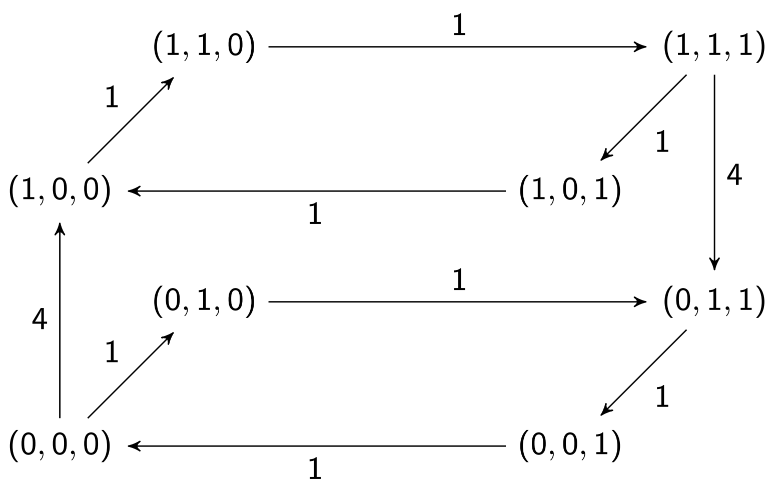
- Helmholtz decompositon applied to space of game flows

$$\textit{game flow} = \textit{potential game} \oplus \textit{nonstrategic game} \oplus \textit{harmonic game}$$

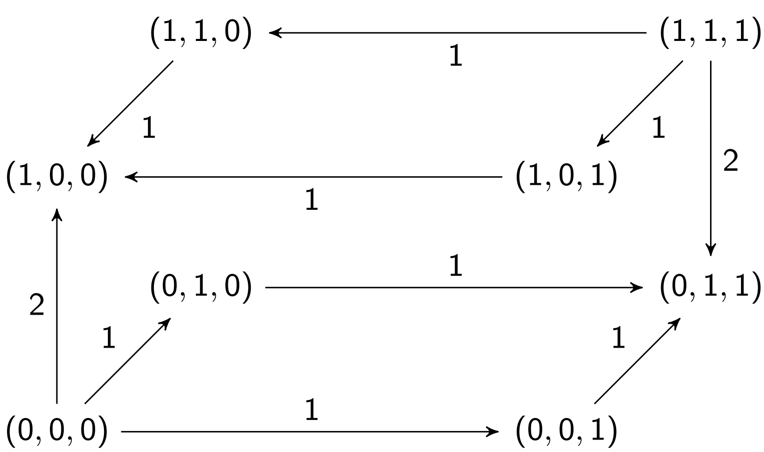
- first defined in [Candogan–Menache–Ozdaglar–Parrilo, 2011] but similar ideas appeared in game theory literature [Hofbauer–Schlag, 2000]
- generically no pure Nash equilibrium
- essentially sums of cycles
- e.g. rock-paper-scissors, cyclic games

example: road sharing game

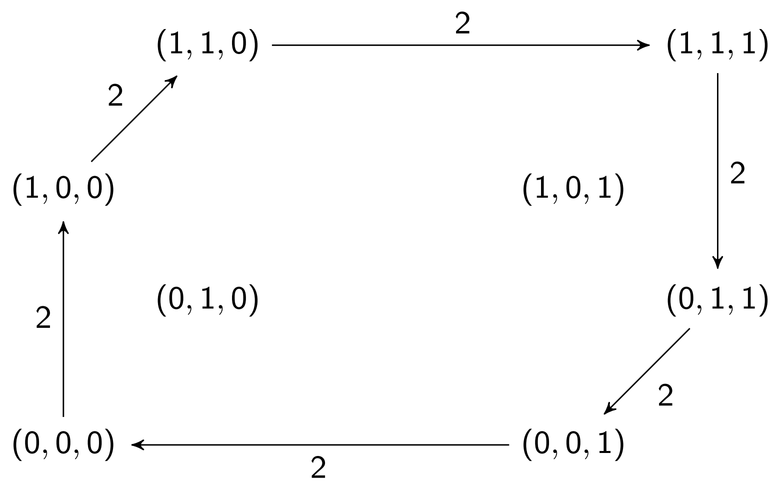
- proposed in [Candogan–Menache–Ozdaglar–Parrilo, 2011]
- three-player game: $V = \{1, 2, 3\}$
- each player choose one of two roads $\{0, 1\}$
- player 3 tries to avoid sharing the road with other players: its payoff decreases by 2 with each other player who shares its road
- player 1 receives a payoff of -1 if player 2 shares its road and 0 otherwise
- payoff of player 2 is equal to negative of the payoff of player 1, i.e., $u_1 + u_2 = 0$
- intuitively, player 1 tries to avoid player 2, whereas player 2 wants to use the same road with player 1



(a) flow of road-sharing game



(b) potential component



(c) harmonic component

Thank You