Applications of Chordal Flag Avergages

Nate Mankovich
Joint work with Tolga Birdal

University of Valencia
Image and Signal Processing Group

November 6, 2023

Outline

Introduction

Averages over $\mathbb{R}$

Averages over $\mathbb{R}$

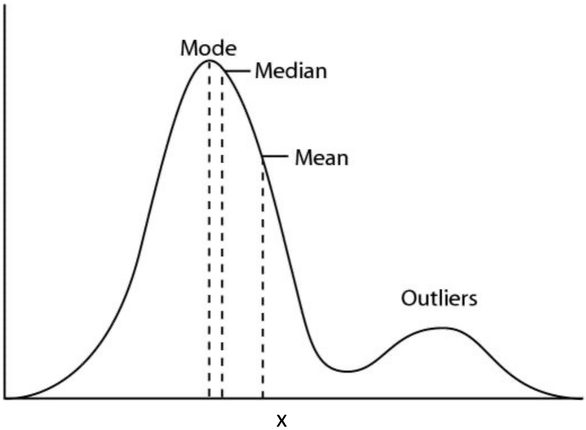


Figure: Hedges et al. 2003

Averages over $\mathbb{R}$

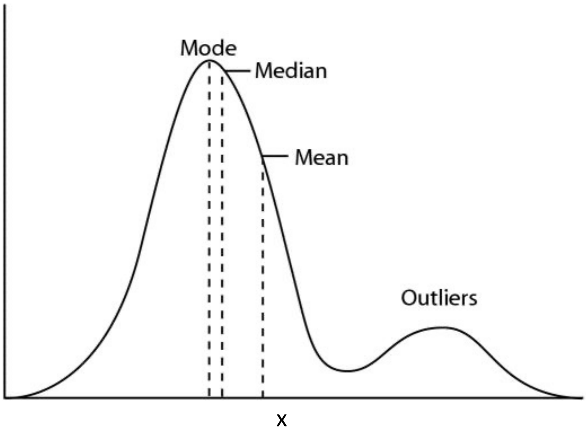


Figure: Hedges et al. 2003

Data:
$\{x^{(1)}, x^{(2)}, \dots, x^{(p)}\} \subset \mathbb{R}$

Averages over $\mathbb{R}$

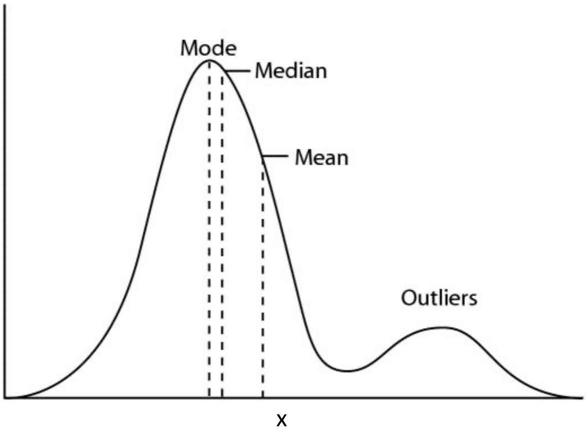


Figure: Hedges et al. 2003

Data:
$\{x^{(1)}, x^{(2)}, \ldots, x^{(p)}\} \subset \mathbb{R}$
The mean

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)^2$$

The geometric median

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)$$

Averages over $\mathbb{R}$

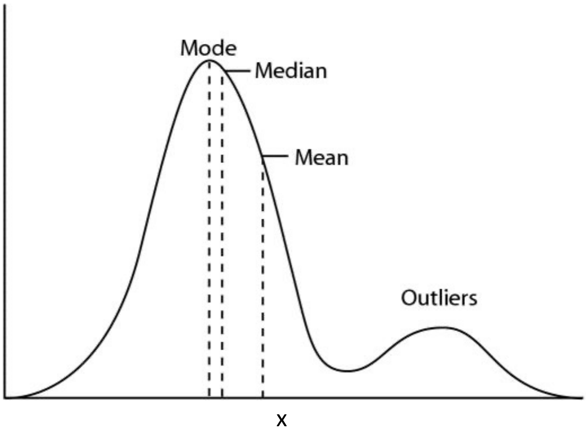


Figure: Hedges et al. 2003

Data:
$\{x^{(1)}, x^{(2)}, \dots, x^{(p)}\} \subset \mathbb{R}$
The mean

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)^2$$

The geometric median

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)$$

Applications

Averages over $\mathbb{R}$

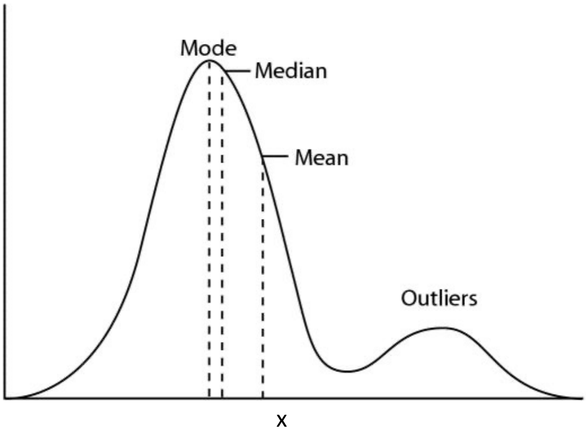


Figure: Hedges et al. 2003

Data:
$\{x^{(1)}, x^{(2)}, \dots, x^{(p)}\} \subset \mathbb{R}$
The mean

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)^2$$

The geometric median

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)$$

Applications

- Clustering algorithms

Averages over $\mathbb{R}$

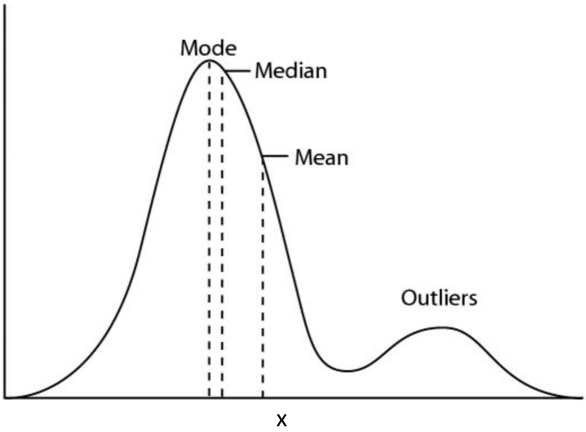


Figure: Hedges et al. 2003

Data:
$\{x^{(1)}, x^{(2)}, \ldots, x^{(p)}\} \subset \mathbb{R}$
The mean

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)^2$$

The geometric median

$$\arg\min_{y \in \mathbb{R}} \sum_{i=1}^{p} d(x^{(i)}, y)$$

Applications

- Clustering algorithms
- Dimensionality reduction

Subspaces in Computer Vision

Subspaces in Computer Vision

Figure: Kirby et al. 2013

Subspaces in Computer Vision

Figure: Kirby et al. 2013

Each row is represented by a subspace.

Subspaces in Computer Vision

Figure: Kirby et al. 2013

Each row is represented by a subspace.

Lowest image: 1^{st} dimension of mean of subspaces (rows)

Subspaces in Computer Vision

Each row is represented by a subspace.

Lowest image: 1^{st} dimension of mean of subspaces (rows)

Figure: Kirby et al. 2013

This mean captures the common face across all three subspaces.

2D Embeddings of Flags vs. Subspaces

2D Embeddings of Flags vs. Subspaces

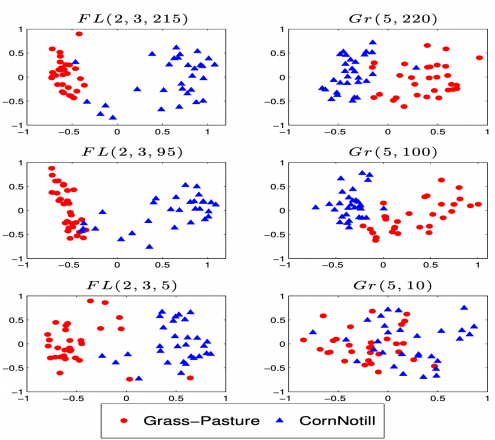


Figure: Indian Pines dataset. Ma et al. 2021

Flag representation (left) and subspace representation (right) for crops.

From Data to Flags

What is a Flag?

What is a Flag?

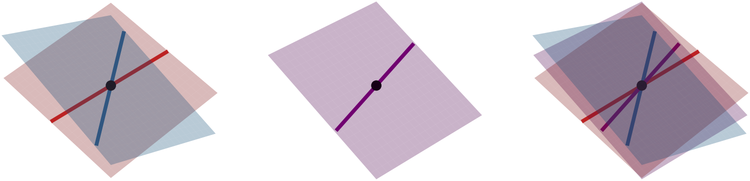


Figure: Data (blue and red) and the chordal flag mean (purple).

What is a Flag?

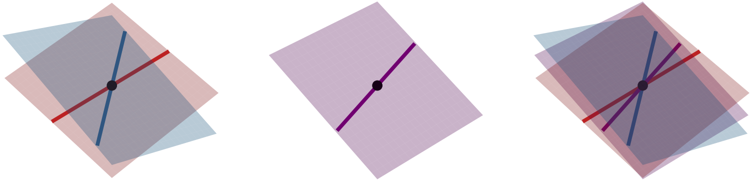


Figure: Data (blue and red) and the chordal flag mean (purple).

$\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ orthogonal unit vectors.

What is a Flag?

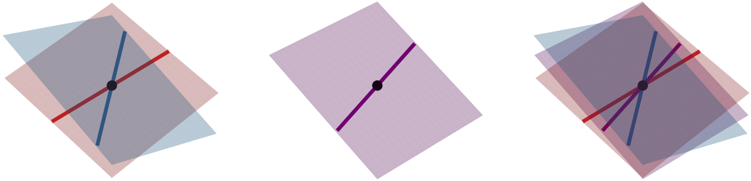


Figure: Data (blue and red) and the chordal flag mean (purple).

$\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ orthogonal unit vectors. $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$

What is a Flag?

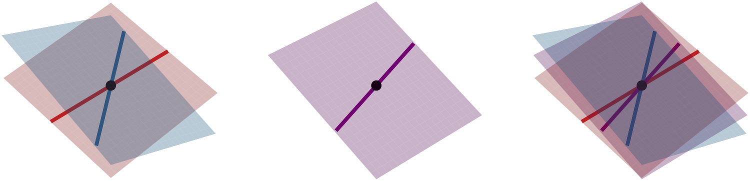


Figure: Data (blue and red) and the chordal flag mean (purple).

$\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ orthogonal unit vectors. $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2]$

Flag manifold of type (1,2,3)

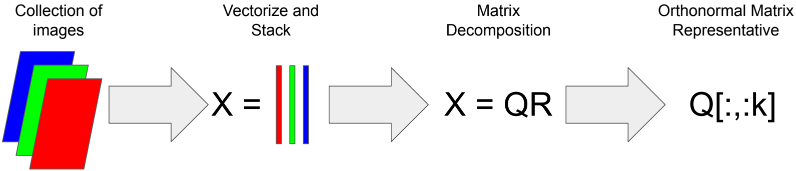
$$\mathrm{FL}(1,2;3) = \left\{ [\![\mathbf{X}]\!] = \mathrm{span}(\mathbf{x}_1) \subset \mathrm{span}(\mathbf{x}_1, \mathbf{x}_2) \subset \mathbb{R}^3 \right\}$$

Representing a Collection of Images

One point (flag or subspace) represents one collection of images.

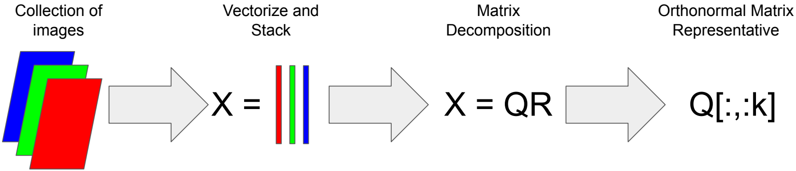
Representing a Collection of Images

One point (flag or subspace) represents one collection of images.



Representing a Collection of Images

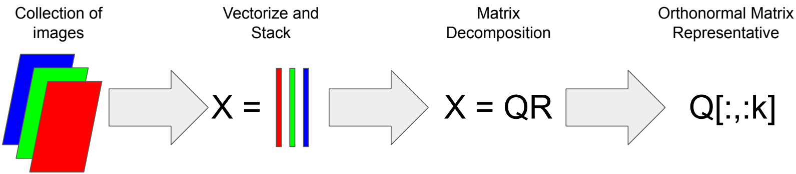
One point (flag or subspace) represents one collection of images.



What should k be?

Representing a Collection of Images

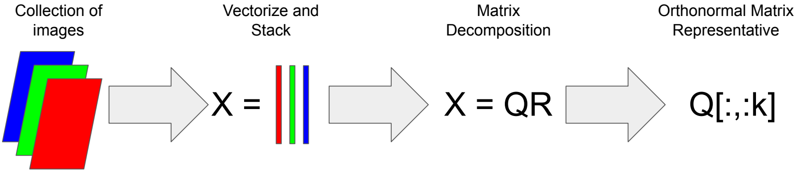
One point (flag or subspace) represents one collection of images.



What should k be? $k = 3$

Representing a Collection of Images

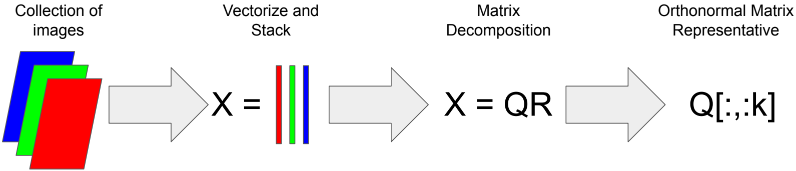
One point (flag or subspace) represents one collection of images.



What should k be? $k = 3$ Other matrix decompositions?

Representing a Collection of Images

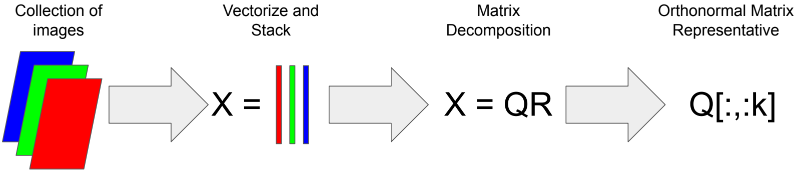
One point (flag or subspace) represents one collection of images.



What should k be? $k = 3$ Other matrix decompositions? SVD

Representing a Collection of Images

One point (flag or subspace) represents one collection of images.

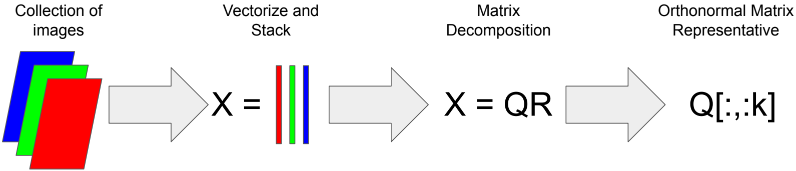


What should k be? $k = 3$ Other matrix decompositions? SVD

Subspace: $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\} \subset \mathbb{R}^n$

Representing a Collection of Images

One point (flag or subspace) represents one collection of images.



What should k be? $k = 3$

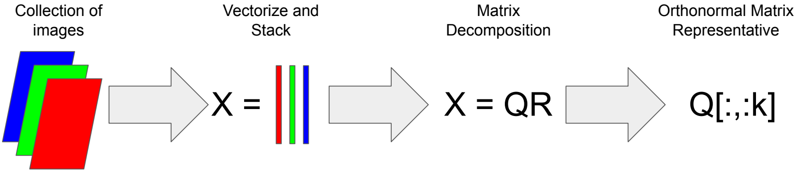
Other matrix decompositions? SVD

Subspace: $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\} \subset \mathbb{R}^n$

Flag: $\text{span}\{\mathbf{q}_1\} \subset \text{span}\{\mathbf{q}_1, \mathbf{q}_2\} \subset \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\} \subset \mathbb{R}^n$

Representing a Collection of Images

One point (flag or subspace) represents one collection of images.



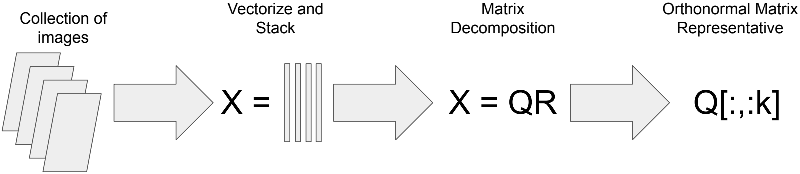
What should k be? $k = 3$ Other matrix decompositions? SVD

Subspace: $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\} \subset \mathbb{R}^n$

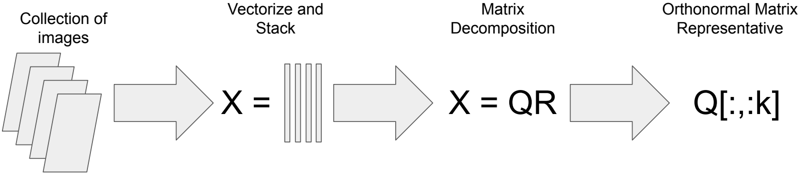
Flag: $\text{span}\{\mathbf{q}_1\} \subset \text{span}\{\mathbf{q}_1, \mathbf{q}_2\} \subset \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\} \subset \mathbb{R}^n$

Flags generalize subspaces by **maintaining the order** of the frames.

Representing Videos and Multispec Images)

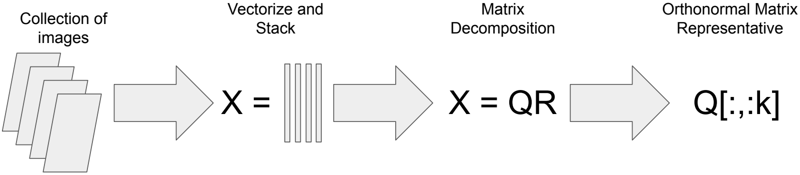


Representing Videos and Multispec Images)



Collection of images = video clip

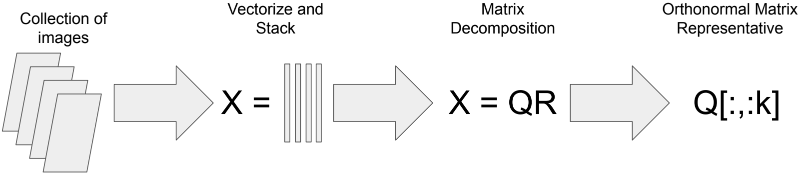
Representing Videos and Multispec Images)



Collection of images = video clip

$\mathbf{x}_i =$ vec(video frame)

Representing Videos and Multispec Images)

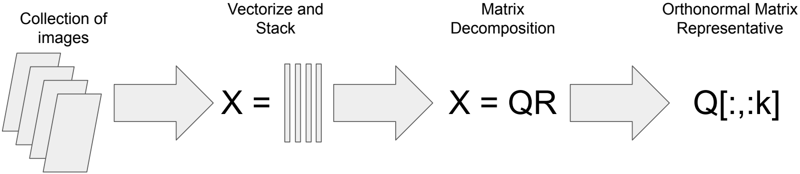


Collection of images = video clip

$\mathbf{x}_i =$ vec(video frame)

Add a frame for each subspace

Representing Videos and Multispec Images)



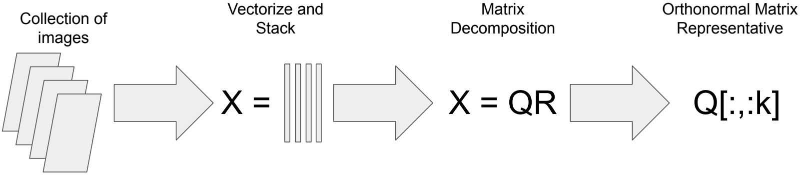
Collection of images = video clip

$\mathbf{x}_i = \text{vec(video frame)}$

Add a frame for each subspace
$\implies$ Total order of subspaces

$$\text{span}\{\mathbf{q}_1\} \subset \cdots \subset \text{span}\{\mathbf{q}_1, \ldots \mathbf{q}_k\} \subset \mathbb{R}^n$$

Representing Videos and Multispec Images)



Collection of images = video clip

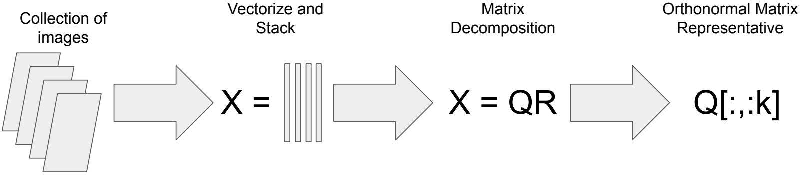
$\mathbf{x}_i = \text{vec(video frame)}$

Add a frame for each subspace
$\implies$ Total order of subspaces

$$\text{span}\{\mathbf{q}_1\} \subset \cdots \subset \text{span}\{\mathbf{q}_1, \dots \mathbf{q}_k\} \subset \mathbb{R}^n$$

Collection of images = multispectral image

Representing Videos and Multispec Images)



Collection of images = video clip

$\mathbf{x}_i = \text{vec(video frame)}$

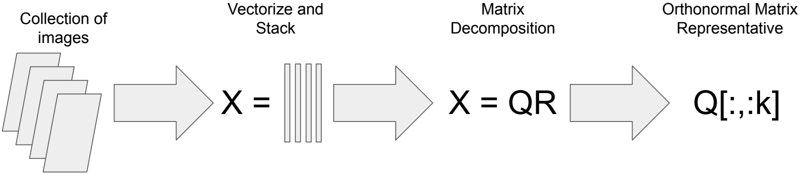
Add a frame for each subspace
$\implies$ Total order of subspaces

$$\text{span}\{\mathbf{q}_1\} \subset \cdots \subset \text{span}\{\mathbf{q}_1, \ldots \mathbf{q}_k\} \subset \mathbb{R}^n$$

Collection of images = multispectral image

$\mathbf{x_i} = \text{vec(1 band)}.$

Representing Videos and Multispec Images)



Collection of images = video clip

$\mathbf{x}_i = \text{vec(video frame)}$

Add a frame for each subspace
$\implies$ Total order of subspaces

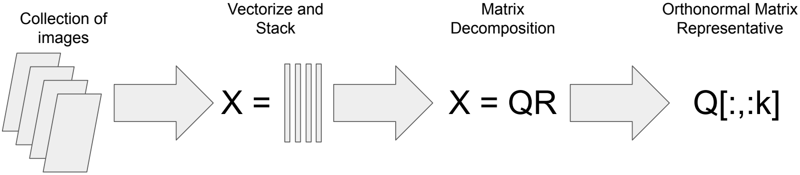
$$\text{span}\{\mathbf{q}_1\} \subset \cdots \subset \text{span}\{\mathbf{q}_1, \ldots \mathbf{q}_k\} \subset \mathbb{R}^n$$

Collection of images = multispectral image

$\mathbf{x_i} = \text{vec(1 band)}$.

Collect bands into groups of size l

Representing Videos and Multispec Images)



Collection of images = video clip

$\mathbf{x}_i = \text{vec(video frame)}$

Add a frame for each subspace
$\implies$ Total order of subspaces

$$\text{span}\{\mathbf{q}_1\} \subset \cdots \subset \text{span}\{\mathbf{q}_1, \dots \mathbf{q}_k\} \subset \mathbb{R}^n$$

Collection of images = multispectral image

$\mathbf{x_i} = \text{vec(1 band)}.$

Collect bands into groups of size l
$\implies$ Partial order of subspaces

$$\text{span}\{\mathbf{q}_1, ..., \mathbf{q}_l\} \subset \cdots \subset \text{span}\{\mathbf{q}_1, ..., \mathbf{q}_k\} \subset \mathbb{R}^n$$

Flags generalize subspaces by **maintaining the order (or partial order)** of the frames.

Chordal Flag Averages

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \text{FL}(d_1, d_2, \dots, d_k; n)$

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \ldots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Chordal Flag Averages

Our space: $\textbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \textbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved.

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved. Flag Mean,

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\operatorname*{arg\,min}_{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved. Flag Mean,
(Draper et al. 2014)

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \ldots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Chordal Flag Median (Mankovich et al. 2023)

Example:
Fl(k;n) solved. Flag Mean,
(Draper et al. 2014)

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Chordal Flag Median (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])$$

Example:
Fl(k;n) solved. Flag Mean,
(Draper et al. 2014)

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved. Flag Mean,
(Draper et al. 2014)

Chordal Flag Median (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])$$

Example:
Fl(k;n)

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved. Flag Mean, (Draper et al. 2014)

Chordal Flag Median (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])$$

Example:
Fl(k;n) FlagIRLS algorithm.

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \text{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved. Flag Mean, (Draper et al. 2014)

Chordal Flag Median (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])$$

Example:
Fl(k;n) FlagIRLS algorithm. Flag Median,

Chordal Flag Averages

Our space: $\mathbf{Fl}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n)$

Data: $\{[\![\mathbf{X}^{(i)}]\!]\}_{i=1}^{p} \subset \mathbf{Fl}(n+1)$

Weights: $\{\alpha_i\}_{i=1}^{p} \subset \mathbb{R}$

Chordal Flag Mean (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])^2$$

Example:
Fl(k;n) solved. Flag Mean, (Draper et al. 2014)

Chordal Flag Median (Mankovich et al. 2023)

$$\underset{[\![\mathbf{Y}]\!] \in \mathbf{Fl}(n+1)}{\arg\min} \sum_{i=1}^{p} \alpha_i d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!])$$

Example:
Fl(k;n) FlagIRLS algorithm. Flag Median, (Mankovich et al. 2022)

Computing Chordal Flag Averages

The problem of averaging flags is largely unstudied.

Computing Chordal Flag Averages

The problem of averaging flags is largely unstudied.

The Chordal Flag Mean (Mankovich et al. 2023)

The chordal flag mean can be simplified to a problem over $\mathbf{X} \in \mathbb{R}^{n \times k}$, $\mathbf{X}^T\mathbf{X} = \mathbf{I}$ and found using Riemannian trust regions.

Computing Chordal Flag Averages

The problem of averaging flags is largely unstudied.

The Chordal Flag Mean (Mankovich et al. 2023)

The chordal flag mean can be simplified to a problem over $\mathbf{X} \in \mathbb{R}^{n \times k}$, $\mathbf{X}^T\mathbf{X} = \mathbf{I}$ and found using Riemannian trust regions.

The Chordal Flag Median (Mankovich et al. 2023)

The chordal flag median can be found using iteratively reweighed chordal flag means.

Computing Chordal Flag Averages

The problem of averaging flags is largely unstudied.

The Chordal Flag Mean (Mankovich et al. 2023)

The chordal flag mean can be simplified to a problem over $\mathbf{X} \in \mathbb{R}^{n \times k}$, $\mathbf{X}^T\mathbf{X} = \mathbf{I}$ and found using Riemannian trust regions.

The Chordal Flag Median (Mankovich et al. 2023)

The chordal flag median can be found using iteratively reweighed chordal flag means.

Chordal Flag Median: Algorithm Convergence (Mankovich et al. 2023)

An iteration of the chordal flag median algorithm does not* increase objective function values.

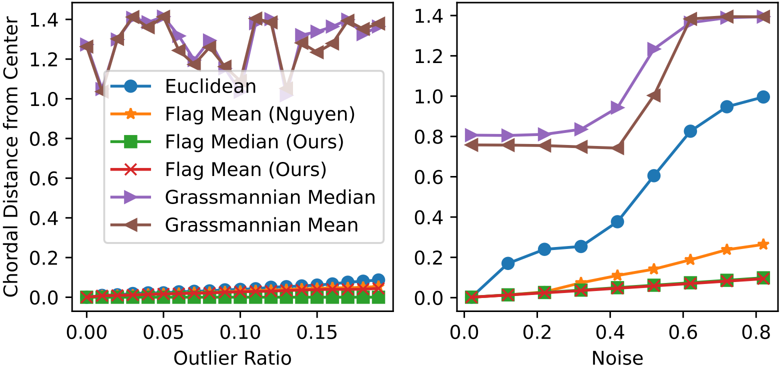
Examples

Flags vs. Subspaces

Flags vs. Subspaces

Flag averaging enforces the order of the faces.

Synthetic Data ($FL(1, 3; 10)$)



- Chordal flag averages- robust to outliers and noise
- Chordal flag mean- more robust estimate of chordal flag mean than Nguyen

Average motions represented as flags

$SE(3) = \{\text{motions in } \mathbb{R}^3\} = \{\text{rotation + translation in } \mathbb{R}^3\}$

Average motions represented as flags

$SE(3) = \{\text{motions in } \mathbb{R}^3\} = \{\text{rotation + translation in } \mathbb{R}^3\}$

Flag representation of motion

$$SE(3) \to SO(4) \to \mathrm{FL}^+(1,2,3;4) \to SO(4) \to SE(3)$$

Average motions represented as flags

$SE(3) = \{\text{motions in } \mathbb{R}^3\} = \{\text{rotation + translation in } \mathbb{R}^3\}$

Flag representation of motion

$$SE(3) \to SO(4) \to \mathrm{FL}^+(1,2,3;4) \to SO(4) \to SE(3)$$

1. Data: cluster of motions using a true mean motion
2. Translate the motions to flags
3. Average the flags
4. Translate average flag to an estimate mean motion
5. Compute error between estimate and true mean

foo

Average motions represented as flags

$SE(3) = \{\text{motions in } \mathbb{R}^3\} = \{\text{rotation + translation in } \mathbb{R}^3\}$

Flag representation of motion

$$SE(3) \to SO(4) \to \text{FL}^+(1,2,3;4) \to SO(4) \to SE(3)$$

1. Data: cluster of motions using a true mean motion
2. Translate the motions to flags
3. Average the flags
4. Translate average flag to an estimate mean motion
5. Compute error between estimate and true mean

foo

Averaging motions represented as flags is robust to noise and outliers compared to classical motion averaging methods.

MNIST Digits and NNs

MNIST Digits and NNs

MNIST Digits = $\{$ images handwritten digits $\}$

MNIST Digits and NNs

MNIST Digits = { images handwritten digits }

1. Dataset: 20 ones and *i* nines

MNIST Digits and NNs

MNIST Digits = { images handwritten digits }

1. Dataset: 20 ones and *i* nines

2. Represent data as planes (GR) or lines inside planes (flags)

MNIST Digits and NNs

MNIST Digits = { images handwritten digits }

1. Dataset: 20 ones and *i* nines
2. Represent data as planes (GR) or lines inside planes (flags)
3. Compute averages

MNIST Digits and NNs

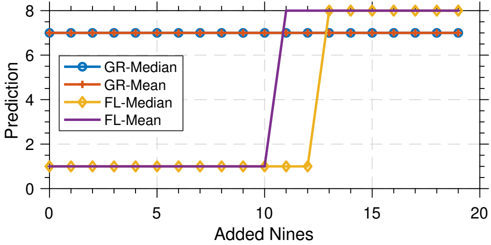
MNIST Digits = { images handwritten digits }

1. Dataset: 20 ones and *i* nines
2. Represent data as planes (GR) or lines inside planes (flags)
3. Compute averages
4. Pass 1^{st} dimension of averages through pretrained NN and predict class

MNIST Digits and NNs

MNIST Digits = { images handwritten digits }

1. Dataset: 20 ones and *i* nines
2. Represent data as planes (GR) or lines inside planes (flags)
3. Compute averages
4. Pass 1st dimension of averages through pretrained NN and predict class

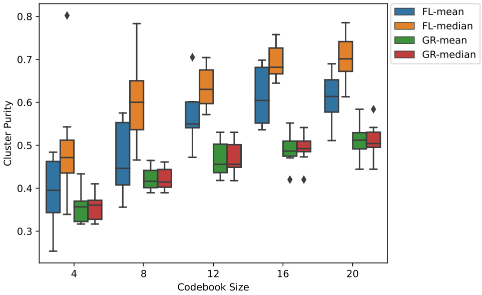


- Chordal flag averages are predicted correctly with $i = 1, 2, \dots, 10$ added nines.
- Median is more robust to outliers than mean.

LBG Clustering (UFC YouTube Dataset)

Linde-Buzo-Gray clustering using

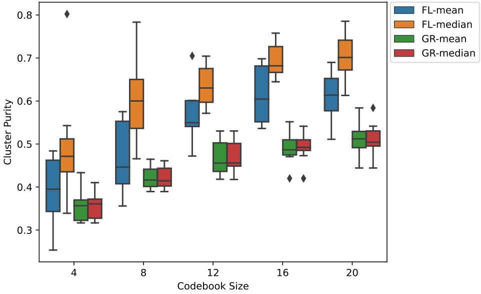
- Distance: between flags (FL) vs. between subspaces (GR)
- Averages: between flags (FL) vs. between subspaces (GR)
- Averages: mean vs. median



LBG Clustering (UFC YouTube Dataset)

Linde-Buzo-Gray clustering using

- Distance: between flags (FL) vs. between subspaces (GR)
- Averages: between flags (FL) vs. between subspaces (GR)
- Averages: mean vs. median



Flag averaging improves cluster purities over averages of subspaces.
Data from Liu et al. 2009.

Batch PCA

Batch PCA

1. Split samples into groups

Batch PCA

1. Split samples into groups
2. Do PCA on each group

Batch PCA

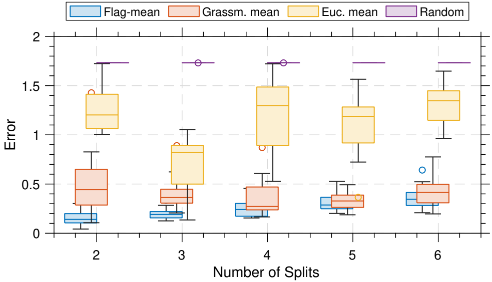
1. Split samples into groups
2. Do PCA on each group
3. Average the PCA weights across groups

Batch PCA

1. Split samples into groups
2. Do PCA on each group
3. Average the PCA weights across groups
4. Compute error: distance between average PCA weights and PCA weights with all subjects

Batch PCA

1. Split samples into groups
2. Do PCA on each group
3. Average the PCA weights across groups
4. Compute error: distance between average PCA weights and PCA weights with all subjects



Chordal flag-means produce lower errors than the other methods.

Conclusion

Contributions

Contributions

- Introduced the novel averages for flag manifolds

Contributions

- Introduced the novel averages for flag manifolds
- Found algorithms for computing these averages

Contributions

- Introduced the novel averages for flag manifolds
- Found algorithms for computing these averages
- Proved convergence of these algorithms

Contributions

- Introduced the novel averages for flag manifolds
- Found algorithms for computing these averages
- Proved convergence of these algorithms
- Improved action clip clustering

Contributions

- Introduced the novel averages for flag manifolds
- Found algorithms for computing these averages
- Proved convergence of these algorithms
- Improved action clip clustering
- Introduced batch PCA

Contributions

- Introduced the novel averages for flag manifolds
- Found algorithms for computing these averages
- Proved convergence of these algorithms
- Improved action clip clustering
- Introduced batch PCA
- Robust Euclidean motion averaging

Contributions

- Introduced the novel averages for flag manifolds
- Found algorithms for computing these averages
- Proved convergence of these algorithms
- Improved action clip clustering
- Introduced batch PCA
- Robust Euclidean motion averaging

For more information see *Chordal Flag Averaging and its Applications*.

Future Work

Future Work

Mathematical guarantees

- Convexity of chordal distance function
- Convergence rates for algorithms

Future Work

Mathematical guarantees

- Convexity of chordal distance function
- Convergence rates for algorithms

Dimensionality reduction

- Finding PCA weights by optimizing over flags (Ye et al. 2023)
- Principal geodesic analysis, PCA on Riemannian manifolds (Fletcher et al. 2004)
- Robust versions of PCA (Tsaskiris et al. 2019 and Neumayer et al. 2019)
- Dynamic mode decomposition, kernels and more (Baddoo et al. 2021)
- Granger Causality and PCA (Varando et al. 2022)

Thank you.

References I

- K. Aftab, R. Hartley, and J. Trumpf, Generalized Weiszfeld algorithms for Lq optimization. IEEE TPAMI, 2014
- D. Bates, B. Davis, M. Kirby, J. Marks, C. Peterson, The max-length-vector line of best fit to a set of vector subspaces and an optimization problem over a set of hyperellipsoids. NLAA, 2015
- S. Chapushtanova, M. Kirby, Classification of hyperspectral imagery on embedded Grassmannians. arXiv, 2015
- M. Chu, L. Watterson, On a multivariate eigenvalue problem, part I: algebraic theory and a power method. SIAM JSC, 1993
- Deng, Li. The MNIST database of handwritten digit images for machine learning research [best of the web]. IEEE signal processing magazine, 2012
- B. Draper, M. Kirby, J. Marks, T. Marrinan, and P. Chris, A flag representation for finite collections of subspaces of mixed dimensions. LAA, 2014
- T. Fletcher, S. Venkatasubramanian and S. Joshi, The geometric median on Riemannian manifolds with application to robust atlas estimation. NeuroImage, 2009,

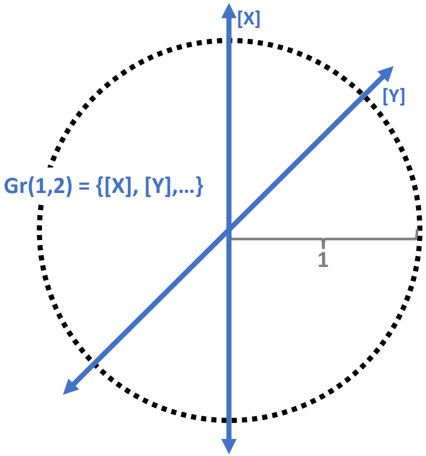
References II

- P. Langfelder and S. Horvath, WGCNA: An R package for weighted correlation network analysis. BMC bioinformatics, 2008
- J. Liu, J. Luo and S. Mubarak, Recognizing realistic actions from videos “in the wild.” CVPR, 2009
- N. Mankovich, E. King, C. Peterson, and M. Kirby, The flag median and FlagIRLS, CVPR, 2022
- N. Mankovich, T. Birdal, Chordal averaging on flag manifolds and its applications, arXiv preprint, 2023
- E. Weiszfeld, Sur le Point Pour Lequel la Somme des Sistances de n Points Donnés est Minimum. Tohoku Mathematical Journal, First Series, 1937

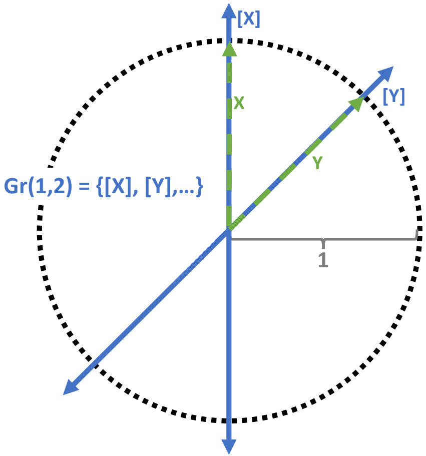
Backup Slides

Backup Slides

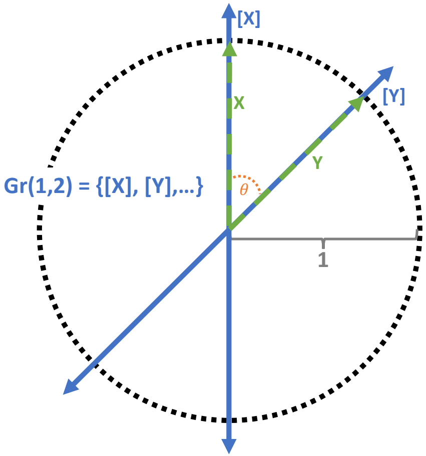
Distances Between Subspaces



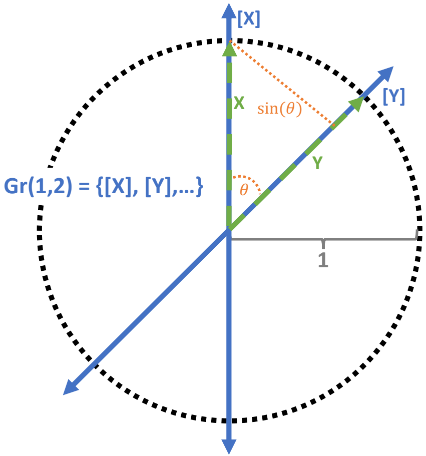
Distances Between 1D Subspaces



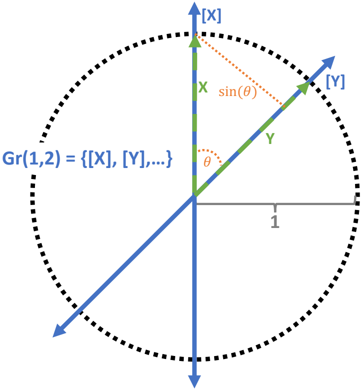
Distances Between 1D Subspaces



Distances Between 1D Subspaces



Distances Between 1D Subspaces



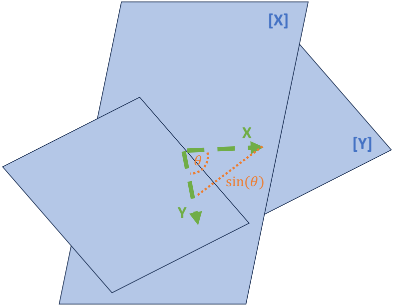
θ is the principal angle between [**X**] and [**Y**].

Geodesic distance is θ

Chordal distance is $\sin\theta$

Distances Between 2D Subspaces

Gr(2,3) = {[X], [Y],...}



There are k principal angles between subspaces of dimension k,

$$\theta([\mathbf{X}], [\mathbf{Y}]) = [\theta_1, \theta_2, \dots, \theta_k]^T.$$

Geodesic distance

$$\|\theta([\mathbf{X}], [\mathbf{Y}])\|_2$$

Chordal distance

$$\|\sin(\theta([\mathbf{X}], [\mathbf{Y}]))\|_2$$

Formalizing Subspace Prototypes

Formalizing Subspace Prototypes

Goal: Find central subspace prototype of $\{[\mathbf{X}_i]\}_{i=1}^{p} \in \mathrm{Gr}(k, n)$

Formalizing Subspace Prototypes

Goal: Find central subspace prototype of $\{[\mathbf{X}_i]\}_{i=1}^{p} \in \mathrm{Gr}(k, n)$

Approach: Optimization problems

Formalizing Subspace Prototypes

Goal: Find central subspace prototype of $\{[\mathbf{X}_i]\}_{i=1}^p \in \mathrm{Gr}(k,n)$

Approach: Optimization problems

Minimize Distance

$$\operatorname*{arg\,min}_{[\mathbf{Y}]\in \mathrm{Gr}(k,n)} \sum_{i=1}^{p} \mathrm{dist}([\mathbf{X}_i],[\mathbf{Y}])$$

Formalizing Subspace Prototypes

Goal: Find central subspace prototype of $\{[\mathbf{X}_i]\}_{i=1}^{p} \in \mathrm{Gr}(k, n)$

Approach: Optimization problems

Minimize Distance

$$\underset{[\mathbf{Y}] \in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \mathrm{dist}([\mathbf{X}_i], [\mathbf{Y}])$$

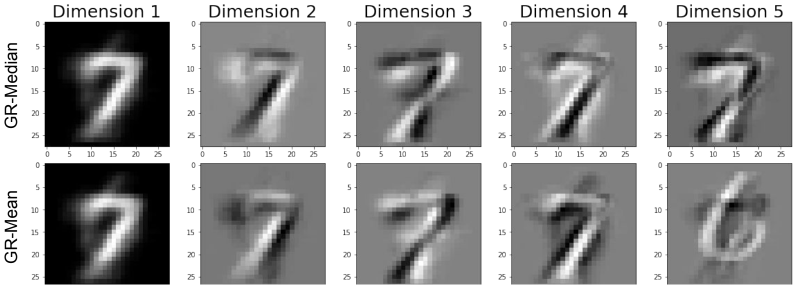
dist involves principal angles between subspaces (e.g., geodesic or chordal distance)!

Chordal Averages of Subspaces

Average 20 examples of 7's and 8 examples of 6's from MNIST digits.

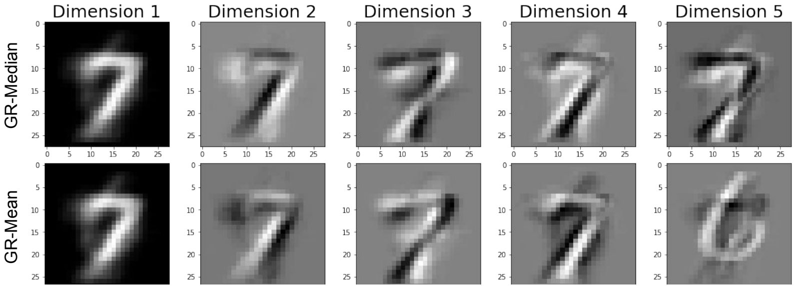
Chordal Averages of Subspaces

Average 20 examples of 7's and 8 examples of 6's from MNIST digits.



Chordal Averages of Subspaces

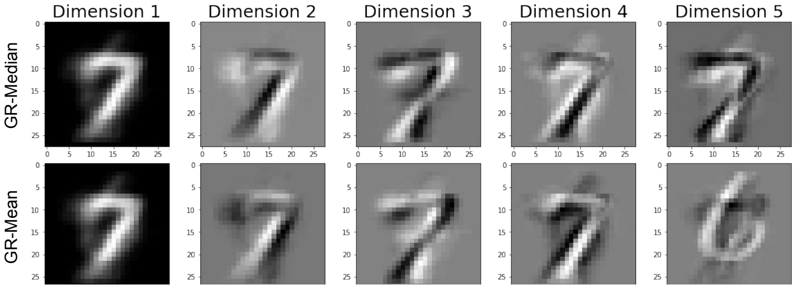
Average 20 examples of 7's and 8 examples of 6's from MNIST digits.



Chordal averages of subspaces are ordered subspaces.

Chordal Averages of Subspaces

Average 20 examples of 7's and 8 examples of 6's from MNIST digits.

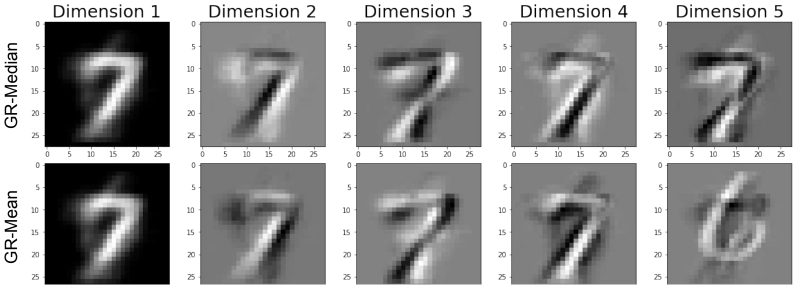


Chordal averages of subspaces are ordered subspaces.

Lower dimensions capture more signal than higher ones.

Chordal Averages of Subspaces

Average 20 examples of 7's and 8 examples of 6's from MNIST digits.



Chordal averages of subspaces are ordered subspaces.

Lower dimensions capture more signal than higher ones.

Median is more robust to outliers in dimension 5.

Algorithm for the Chordal Grassmannian Median

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k, n)$.

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k, n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k, n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y} \in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i], [\mathbf{Y}])\|_2$$

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^p \subset \mathrm{Gr}(k,n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y} \in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i],[\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k,n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y} \in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i],[\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k,n)$.

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k,n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y}\in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin\theta([\mathbf{X}_i],[\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k,n)$.

Repeat until convergence:

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k, n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\operatorname*{arg\,min}_{\mathbf{Y} \in \mathrm{Gr}(k,n)} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i], [\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k, n)$.

Repeat until convergence:

- assign $w_i \leftarrow \frac{1}{\max\{\|\sin \theta([\mathbf{X}_i],[\mathbf{Y}])\|_2, \epsilon\}}$

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^p \subset \mathrm{Gr}(k,n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y}\in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i],[\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k,n)$.

Repeat until convergence:

- assign $w_i \leftarrow \frac{1}{\max\left\{\|\sin\theta([\mathbf{X}_i],[\mathbf{Y}])\|_2,\epsilon\right\}}$
- $\mathbf{Y} \leftarrow$ chordal Grassmannian mean $\left(\left\{\sqrt{w_1}\mathbf{X}_1, \sqrt{w_2}\mathbf{X}_2, \cdots, \sqrt{w_p}\mathbf{X}_p\right\}\right)$

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^{p} \subset \mathrm{Gr}(k, n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y} \in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i], [\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k, n)$.

Repeat until convergence:

- assign $w_i \leftarrow \frac{1}{\max\left\{\|\sin \theta([\mathbf{X}_i],[\mathbf{Y}])\|_2, \epsilon\right\}}$
- $\mathbf{Y} \leftarrow$ chordal Grassmannian mean $\left(\left\{\sqrt{w_1}\mathbf{X}_1, \sqrt{w_2}\mathbf{X}_2, \cdots, \sqrt{w_p}\mathbf{X}_p\right\}\right)$

Proposition (Mankovich 2023)

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^p \subset \mathrm{Gr}(k,n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\arg\min_{\mathbf{Y}\in \mathrm{Gr}(k,n)} \sum_{i=1}^{p} \|\sin \theta([\mathbf{X}_i],[\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k,n)$.

Repeat until convergence:

- assign $w_i \leftarrow \frac{1}{\max\{\|\sin\theta([\mathbf{X}_i],[\mathbf{Y}])\|_2,\epsilon\}}$
- $\mathbf{Y} \leftarrow$ chordal Grassmannian mean $\left(\left\{\sqrt{w_1}\mathbf{X}_1, \sqrt{w_2}\mathbf{X}_2, \cdots, \sqrt{w_p}\mathbf{X}_p\right\}\right)$

Proposition (Mankovich 2023)

$f : \mathrm{Gr}(k,n) \to \mathbb{R}$ *chordal Grassmannian median objective function.*
$T : \mathrm{Gr}(k,n) \to \mathrm{Gr}(k,n)$ *iteration of algorithm.*

Algorithm for the Chordal Grassmannian Median

Data: $\{[\mathbf{X}_i]\}_{i=1}^p \subset \mathrm{Gr}(k,n)$.

Chordal Grassmannian median (Mankovich et al. 2022):

$$\underset{\mathbf{Y}\in \mathrm{Gr}(k,n)}{\arg\min} \sum_{i=1}^{p} \|\sin\theta([\mathbf{X}_i],[\mathbf{Y}])\|_2$$

Algorithm (Mankovich et al. 2022)

Initialize $[\mathbf{Y}] \in \mathrm{Gr}(k,n)$.

Repeat until convergence:

- assign $w_i \leftarrow \frac{1}{\max\left\{\|\sin\theta([\mathbf{X}_i],[\mathbf{Y}])\|_2, \epsilon\right\}}$
- $\mathbf{Y} \leftarrow$ chordal Grassmannian mean $\left(\left\{\sqrt{w_1}\mathbf{X}_1, \sqrt{w_2}\mathbf{X}_2, \cdots, \sqrt{w_p}\mathbf{X}_p\right\}\right)$

Proposition (Mankovich 2023)

$f : \mathrm{Gr}(k,n) \to \mathbb{R}$ chordal Grassmannian median objective function.
$T : \mathrm{Gr}(k,n) \to \mathrm{Gr}(k,n)$ iteration of algorithm. Then

$$f(T([\mathbf{Y}])) \leq f([\mathbf{Y}]) + \frac{p\epsilon}{2}$$

More Examples of Flags

More Examples of Flags

Subspaces vs Flags of Subspaces

$$\text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \subset \mathbb{R}^4 \text{ vs. } \text{span}(\mathbf{x}_1) \subset \text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \subset \mathbb{R}^4$$

More Examples of Flags

Subspaces vs Flags of Subspaces

$$\text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \subset \mathbb{R}^4 \text{ vs. } \text{span}(\mathbf{x}_1) \subset \text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \subset \mathbb{R}^4$$

SVD

$\mathbf{X} \in \mathbb{R}^{n \times k}$, $\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$

$$\text{span}(\mathbf{u}_1) \subset \text{span}(\mathbf{u}_1, \mathbf{u}_2) \subset \cdots \subset \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k) \subset \mathbb{R}^n$$

More Examples of Flags

Subspaces vs Flags of Subspaces

$$\text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \subset \mathbb{R}^4 \text{ vs. } \text{span}(\mathbf{x}_1) \subset \text{span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \subset \mathbb{R}^4$$

SVD

$\mathbf{X} \in \mathbb{R}^{n \times k}$, $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$

$$\text{span}(\mathbf{u}_1) \subset \text{span}(\mathbf{u}_1, \mathbf{u}_2) \subset \cdots \subset \text{span}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k) \subset \mathbb{R}^n$$

Euclidean Motion (Selig 2005 and Ozyesil et al. 2018)

$$SE(3) \to SO(4) \to \text{FL}^+(1,2,3;4) \to SO(4) \to SE(3)$$

The Flag Manifold $\text{FL}(d_1, d_2, \dots, d_k; n)$

The Flag Manifold $\mathrm{FL}(d_1, d_2, \ldots, d_k; n)$

- Let S_i be a subspace. $\dim(S_i) = d_i$

$$\mathrm{FL}(n+1) = \mathrm{FL}(d_1, d_2, \ldots, d_k; n) = \{S_1 \subset S_2 \subset \cdots \subset S_k \subset \mathbb{R}^n\}$$

The Flag Manifold $\text{FL}(d_1, d_2, \ldots, d_k; n)$

- Let S_i be a subspace. $\dim(S_i) = d_i$

$$\text{FL}(n+1) = \text{FL}(d_1, d_2, \ldots, d_k; n) = \{S_1 \subset S_2 \subset \cdots \subset S_k \subset \mathbb{R}^n\}$$

- Formally,

The Flag Manifold $\mathrm{FL}(d_1, d_2, \dots, d_k; n)$

- Let S_i be a subspace. $\dim(S_i) = d_i$

$$\mathrm{FL}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n) = \{S_1 \subset S_2 \subset \cdots \subset S_k \subset \mathbb{R}^n\}$$

- Formally,

$$\frac{SO(n)}{S(O(d_1) \times O(d_2) \times \cdots \times O(d_k) \times O(n - d_k))}$$

The Flag Manifold $\mathrm{FL}(d_1, d_2, \dots, d_k; n)$

- Let S_i be a subspace. $\dim(S_i) = d_i$

$$\mathrm{FL}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n) = \{S_1 \subset S_2 \subset \cdots \subset S_k \subset \mathbb{R}^n\}$$

- Formally,

$$\frac{SO(n)}{S(O(d_1) \times O(d_2) \times \cdots \times O(d_k) \times O(n-d_k))} \cong \frac{\{\mathbf{X} \in \mathbb{R}^{n \times d_k} \ : \ \mathbf{X}^T\mathbf{X} = \mathbf{I}\}}{O(m_1) \times O(m_2) \times \cdots \times O(m_k)}$$

where $m_j = d_j - d_{j-1}$.

The Flag Manifold $\mathrm{FL}(d_1, d_2, \dots, d_k; n)$

- Let S_i be a subspace. $\dim(S_i) = d_i$

$$\mathrm{FL}(n+1) = \mathrm{FL}(d_1, d_2, \dots, d_k; n) = \{S_1 \subset S_2 \subset \cdots \subset S_k \subset \mathbb{R}^n\}$$

- Formally,

$$\frac{SO(n)}{S(O(d_1) \times O(d_2) \times \cdots \times O(d_k) \times O(n - d_k))} \cong \frac{\{\mathbf{X} \in \mathbb{R}^{n \times d_k} \ : \ \mathbf{X}^T\mathbf{X} = \mathbf{I}\}}{O(m_1) \times O(m_2) \times \cdots \times O(m_k)}$$

where $m_j = d_j - d_{j-1}$.

- $[\![\mathbf{X}]\!] \in \mathrm{FL}(n+1)$, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$, and $\mathbf{X}_i \in \mathbb{R}^{n \times m_k}$

$$[\![\mathbf{X}]\!] = \mathrm{span}\{\mathbf{X}_1\} \subset \cdots \subset \mathrm{span}\{\mathbf{X}_1, \dots, \mathbf{X}_k\} \subset \mathbb{R}^n$$

The Flag Manifold $\text{FL}(d_1, d_2, \dots, d_k; n)$

- Let S_i be a subspace. $\dim(S_i) = d_i$

$$\text{FL}(n+1) = \text{FL}(d_1, d_2, \dots, d_k; n) = \{S_1 \subset S_2 \subset \cdots \subset S_k \subset \mathbb{R}^n\}$$

- Formally,

$$\frac{SO(n)}{S(O(d_1) \times O(d_2) \times \cdots \times O(d_k) \times O(n-d_k))} \cong \frac{\{\mathbf{X} \in \mathbb{R}^{n \times d_k} \ : \ \mathbf{X}^T\mathbf{X} = \mathbf{I}\}}{O(m_1) \times O(m_2) \times \cdots \times O(m_k)}$$

where $m_j = d_j - d_{j-1}$.

- $[\![\mathbf{X}]\!] \in \text{FL}(n+1)$, $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$, and $\mathbf{X}_i \in \mathbb{R}^{n \times m_k}$

$$[\![\mathbf{X}]\!] = \text{span}\{\mathbf{X}_1\} \subset \cdots \subset \text{span}\{\mathbf{X}_1, \dots, \mathbf{X}_k\} \subset \mathbb{R}^n$$

See Ye et al. 2022 for more information

Notation... Ugh

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(d_1, d_2, \ldots, d_k; d) = \mathrm{FL}(n+1)$$

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(d_1, d_2, \dots, d_k; d) = \mathrm{FL}(n+1)$$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$

Notation... Ugh

In general

$[\![\mathbf{X}]\!] \in \text{FL}(d_1, d_2, \dots, d_k; d) = \text{FL}(n+1)$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$ and $\mathbf{X}_i \in \mathbb{R}^{d \times m_k}$
  $m_j = d_j - d_{j-1}$

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in FL(d_1, d_2, \dots, d_k; d) = FL(n+1)$$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$ and $\mathbf{X}_i \in \mathbb{R}^{d \times m_k}$
  $m_j = d_j - d_{j-1}$

$$[\![\mathbf{X}]\!] = \text{span}\{\mathbf{X}_1\} \subset \cdots \subset \text{span}\{\mathbf{X}_1, \dots, \mathbf{X}_k\} \subset \mathbb{R}^d$$

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(d_1, d_2, \dots, d_k; d) = \mathrm{FL}(n+1)$$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$ and $\mathbf{X}_i \in \mathbb{R}^{d \times m_k}$
  $m_j = d_j - d_{j-1}$

$$[\![\mathbf{X}]\!] = \mathrm{span}\{\mathbf{X}_1\} \subset \cdots \subset \mathrm{span}\{\mathbf{X}_1, \dots, \mathbf{X}_k\} \subset \mathbb{R}^d$$

Example

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(1, 3; 5)$$

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(d_1, d_2, \dots, d_k; d) = \mathrm{FL}(n+1)$$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$ and $\mathbf{X}_i \in \mathbb{R}^{d \times m_k}$
  $m_j = d_j - d_{j-1}$

$$[\![\mathbf{X}]\!] = \mathrm{span}\{\mathbf{X}_1\} \subset \cdots \subset \mathrm{span}\{\mathbf{X}_1, \dots, \mathbf{X}_k\} \subset \mathbb{R}^d$$

Example

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(1, 3; 5)$$

- $\mathbf{X} \in \mathbb{R}^{4\times 3}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in \text{FL}(d_1, d_2, \dots, d_k; d) = \text{FL}(n+1)$$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k]$ and $\mathbf{X}_i \in \mathbb{R}^{d \times m_k}$
  $m_j = d_j - d_{j-1}$

$$[\![\mathbf{X}]\!] = \text{span}\{\mathbf{X}_1\} \subset \cdots \subset \text{span}\{\mathbf{X}_1, \dots, \mathbf{X}_k\} \subset \mathbb{R}^d$$

Example

$$[\![\mathbf{X}]\!] \in \text{FL}(1, 3; 5)$$

- $\mathbf{X} \in \mathbb{R}^{4 \times 3}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$, $\mathbf{X}_1 \in \mathbb{R}^{4 \times 1}$, $\mathbf{X}_2 \in \mathbb{R}^{4 \times 2}$

Notation... Ugh

In general

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(d_1, d_2, \ldots, d_k; d) = \mathrm{FL}(n+1)$$

- $\mathbf{X} \in \mathbb{R}^{d_k \times d}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k]$ and $\mathbf{X}_i \in \mathbb{R}^{d \times m_k}$
  $m_j = d_j - d_{j-1}$

$$[\![\mathbf{X}]\!] = \mathrm{span}\{\mathbf{X}_1\} \subset \cdots \subset \mathrm{span}\{\mathbf{X}_1, \ldots, \mathbf{X}_k\} \subset \mathbb{R}^d$$

Example

$$[\![\mathbf{X}]\!] \in \mathrm{FL}(1, 3; 5)$$

- $\mathbf{X} \in \mathbb{R}^{4 \times 3}$ and $\mathbf{X}^T\mathbf{X} = \mathbf{I}$
- $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$, $\mathbf{X}_1 \in \mathbb{R}^{4 \times 1}$, $\mathbf{X}_2 \in \mathbb{R}^{4 \times 2}$

$$\mathrm{span}\{\mathbf{X}_1\} \subset \mathrm{span}\{\mathbf{X}_1, \mathbf{X}_2\} \subset \mathbb{R}^5$$

Translating to the Stiefel Manifold

Translating to the Stiefel Manifold

$\mathbf{I}_j$ identity matrix with some zeros on the diagonal.

Translating to the Stiefel Manifold

$\mathbf{I}_j$ identity matrix with some zeros on the diagonal.

$\mathbf{P}_j$ projection matrix is a function of $\mathbf{X}_j^{(i)}$, α_i for $i = 1, 2, \ldots, p$

Translating to the Stiefel Manifold

$\mathbf{I}_j$ identity matrix with some zeros on the diagonal.

$\mathbf{P}_j$ projection matrix is a function of $\mathbf{X}_j^{(i)}$, α_i for $i = 1, 2, \dots, p$

Proposition (Mankoivich, Birdal 2023)

The chordal flag mean optimization problem is equivalent to

$$\min_{\mathbf{Y} \in St(d_k, n)} \sum_{j=1}^{k} m_j - tr\left(\mathbf{I}_j \mathbf{Y}^T \mathbf{P}_j \mathbf{Y}\right).$$

Translating to the Stiefel Manifold

$\mathbf{I}_j$ identity matrix with some zeros on the diagonal.

$\mathbf{P}_j$ projection matrix is a function of $\mathbf{X}_j^{(i)}$, α_i for $i = 1, 2, \dots, p$

Proposition (Mankoivich, Birdal 2023)

The chordal flag mean optimization problem is equivalent to

$$\min_{\mathbf{Y} \in St(d_k, n)} \sum_{j=1}^{k} m_j - tr\left(\mathbf{I}_j \mathbf{Y}^T \mathbf{P}_j \mathbf{Y}\right).$$

Proposition (Mankoivich, Birdal 2023)

The chordal flag median optimization problem can be approximated by a weighted chordal flag mean with weights

$$w_i(\mathbf{Y}) = \sum_{j=1}^{k} \frac{\alpha_i}{\max\left\{d_c([\![\mathbf{X}^{(i)}]\!], [\![\mathbf{Y}]\!]), \epsilon\right\}}.$$

Chordal Distance Between Flags

Let $[\![\mathbf{X}]\!], [\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)$.

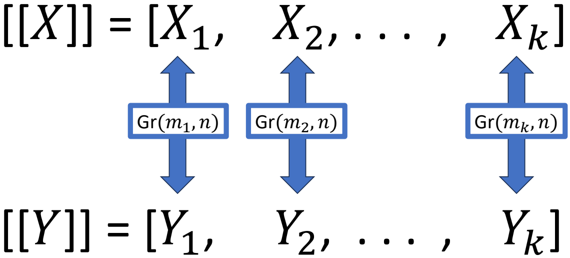
Chordal Distance Between Flags

Let $[\![\mathbf{X}]\!], [\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)$. Measure distance between each $\mathrm{Gr}(m_i, n)$.

$$\begin{tikzcd}
{[[X]] = [X_1,} & {X_2,} & {\dots\ ,} & {X_k]} \\
{[[Y]] = [Y_1,} \arrow[u, leftrightarrow, "\mathrm{Gr}(m_1,n)"] & {Y_2,} \arrow[u, leftrightarrow, "\mathrm{Gr}(m_2,n)"] & {\dots\ ,} & {Y_k]} \arrow[u, leftrightarrow, "\mathrm{Gr}(m_k,n)"]
\end{tikzcd}$$

Chordal Distance Between Flags

Let $[\![\mathbf{X}]\!], [\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)$. Measure distance between each $\mathrm{Gr}(m_i, n)$.



Chordal Distance Between Flags (Pitaval at al. 2013)

$$d_c([\![\mathbf{X}]\!], [\![\mathbf{Y}]\!]) := \sqrt{\sum_{j=1}^{k} \| \sin \theta([\mathbf{X}_j], [\mathbf{Y}_j]) \|_2^2}.$$

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^p \in \mathrm{FL}(n+1)$.

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^{p} \in \mathrm{FL}(n+1)$.

Chordal flag mean

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^{p} \in \mathrm{FL}(n+1)$.

Chordal flag mean

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])^2$$

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^{p} \in \mathrm{FL}(n+1)$.

Chordal flag mean

Chordal flag median

$$\underset{[\![\mathbf{Y}]\!]\in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])^2$$

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^{p} \in \mathrm{FL}(n+1)$.

Chordal flag mean

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])^2$$

Chordal flag median

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])$$

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^{p} \in \mathrm{FL}(n+1)$.

Chordal flag mean

$$\underset{[\![\mathbf{Y}]\!]\in\mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!],[\![\mathbf{Y}]\!])^2$$

Chordal flag median

$$\underset{[\![\mathbf{Y}]\!]\in\mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!],[\![\mathbf{Y}]\!])$$

Algorithm- chordal flag mean (Mankovich and Birdal 2023)

- Translate to a problem over $\mathbf{Y} \in \mathbb{R}^{n\times d_k}$ subject to $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$
- Solve using Stiefel manifold optimization

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^p \in \mathrm{FL}(n+1)$.

Chordal flag mean

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])^2$$

Chordal flag median

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])$$

Algorithm- chordal flag mean (Mankovich and Birdal 2023)

- Translate to a problem over $\mathbf{Y} \in \mathbb{R}^{n \times d_k}$ subject to $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$
- Solve using Stiefel manifold optimization

Algorithm- chordal flag median (Mankovich and Birdal 2023)

Initialize $[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)$. Repeat until convergence:

- assign w_i
- $[\![\mathbf{Y}]\!] \leftarrow$ chordal flag mean $\left(\{\sqrt{w_1}[\![\mathbf{X}_i]\!]\}_{i=1}^p\right)$

Proposition (Convergence of Chordal Flag Mean (Mankovich and Birdal 2023))

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^p \in \mathrm{FL}(n+1)$.

Chordal flag mean

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^p d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])^2$$

Chordal flag median

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^p d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])$$

Algorithm- chordal flag mean (Mankovich and Birdal 2023)

- Translate to a problem over $\mathbf{Y} \in \mathbb{R}^{n \times d_k}$ subject to $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$
- Solve using Stiefel manifold optimization

Algorithm- chordal flag median (Mankovich and Birdal 2023)

Initialize $[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)$. Repeat until convergence:

- assign w_i
- $[\![\mathbf{Y}]\!] \leftarrow$ chordal flag mean $\left(\{\sqrt{w_1}[\![\mathbf{X}_i]\!]\}_{i=1}^p\right)$

Proposition (Convergence of Chordal Flag Mean (Mankovich and Birdal 2023))

$f : \mathrm{Gr}(k,n) \to \mathbb{R}$ *objective function.*

$T : \mathrm{Gr}(k,n) \to \mathrm{Gr}(k,n)$ *iteration of algorithm*

Chordal Flag Mean and Median

Data: $\{[\![\mathbf{X}_i]\!]\}_{i=1}^p \in \mathrm{FL}(n+1)$.

Chordal flag mean

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])^2$$

Chordal flag median

$$\underset{[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)}{\arg\min} \sum_{i=1}^{p} d_c([\![\mathbf{X}_i]\!], [\![\mathbf{Y}]\!])$$

Algorithm- chordal flag mean (Mankovich and Birdal 2023)

- Translate to a problem over $\mathbf{Y} \in \mathbb{R}^{n \times d_k}$ subject to $\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$
- Solve using Stiefel manifold optimization

Algorithm- chordal flag median (Mankovich and Birdal 2023)

Initialize $[\![\mathbf{Y}]\!] \in \mathrm{FL}(n+1)$. Repeat until convergence:

- assign w_i
- $[\![\mathbf{Y}]\!] \leftarrow$ chordal flag mean $\left(\{\sqrt{w_1}[\![\mathbf{X}_i]\!]\}_{i=1}^p\right)$

Proposition (Convergence of Chordal Flag Mean (Mankovich and Birdal 2023))

$f : \mathrm{Gr}(k,n) \to \mathbb{R}$ *objective function.*
$T : \mathrm{Gr}(k,n) \to \mathrm{Gr}(k,n)$ *iteration of algorithm. Then*