From Last Meeting

Studying Independent Component Analysis (ICA)

Idea: Find “directions that maximize independence”

Parallel Idea: Find directions that maximize “non-Gaussianity”

References:

Hyvärinen and Oja (1999) Independent Component Analysis: A Tutorial, http://www.cis.hut.fi/projects/ica

Lee, T. W. (1998) Independent Component Analysis: Theory and Applications, Kluwer.

ICA, Last Time (cont.)

Toy Examples:

1. “Cocktail party problem”
2. 2 sine waves
3. Sine wave and noise
4. 2 noise components
5. Long parallel points clouds
6. 2-d discrimination

ICA, Toy Examples (cont.)

6. Crossed X Discrimination

7. Slanted X Discrimination

Some New Toys:

8. Balanced Sine and Noise

Show ICAeg1p1d7Combine.pdf

- Note PCA gives “even split of sine wave”
- ICA gives excellent denoising

ICA, Toy Examples (cont.)

9. Split X Discrimination:

Show HDLSS\HDLSSxd3Raw.ps

- PCA leaves lots of overlap

Show HDLSS\HDLSSxd3ICA.ps

- ICA gives excellent separation

Show HDLSS\HDLSSxd3ICA.ps

- IC1 has “more kurtosis”, but IC2 is best for discrimination
- Useful preprocessing for e.g. CART

ICA, CurvDat Examples

Recall PCA for “Parabs”

Show CurvDat\ParabsCurvDat.ps

- Mean captured “parabola” shape
- PC1 is “vertical shift”
- PC2 is “tilt” (hard to see visually)
- Remaining PCs are “Gaussian noise”

ICA, CurvDat Examples (cont.)

Corresponding ICA for “Parabs”

Show ParabsCurvDatICA.ps

- mean and centered data as before
- sphered data has “no structure” (i.e. this structure is “all in covariance”, i.e. have Gaussian point cloud)
- sphered ICs choose “random non-Gaussian” directions
- sphered ICs seem to find outliers
- Original scale versions capture some “vertical shift”
- Non-orthogonality on original scale $\Rightarrow$ hard to interpret

ICA, CurvDat Examples (cont.)

Recall PCA for “Parabs with 2 outliers”

Show CurvDat\Parabs2outCurvDat.ps

- Mean captured “parabola” shape
- PC1 is “vertical shift affected by hi-freq outlier”
- PC2 is “most of high freq.outlier”
- “low freq outlier” and “tilt” are mixed between PC3 & PC4
- hope ICA can “separate these”???

ICA, CurvDat Examples (cont.)

Corresponding ICA for “Parabs with 2 outliers”

Show Parabs2outCurvDatICA.ps

- ICA finds both outliers well (non-Gaussian direction)
- ICA still misses “shift” and “tilt”
- Since these are elliptical point cloud properties, that are ignored through sphering.
- ∃ analysis which keeps “both kinds of features”????

ICA, CurvDat Examples (cont.)

Recall PCA for “3 bumps, with 2 independent”

Show CurvDat\Bumps2CurvDat.ps

- Finds both sets of bumps in PC1 and PC2
- Slight mixing of clusters

Corresponding ICA for “3 bumps, with 2 independent”

Show CurvDat\Bumps2CurvDatICA.ps

- Bumps not found (since are “Gaussian” features)
- sphering eliminated bumps

ICA, CurvDat Examples (cont.)

Recall PCA for “Parabs Up and Down” (2 clusters)

Show CurvDat\ParabsUpDnCurvDat.ps

- PC1 finds clusters
- Others find usual structure (vertical shift and tilt)

Corresponding ICA for “Parabs Up and Down”

Show ParabsUpDnCurvDatICA.ps

- Clusters not found???? (seems very “non-Gaussian”)
- sphering killed clusters????
- Problem with numerical search algorithm????

ICA, CurvDat Examples (cont.)

Attempted fix 1: Change of “nonlinear function”

Show CurvDat\ParabsUpDnCurvDatICAt5.ps

- similar results
- same happened for other choices

Attempted fix 2: use PCA directions as “starting value”

Show CurvDat\ParabsUpDnCurvDatICAt2.ps

- Gives good solution
- Is this a general problem????
- How generalizable is this solution????

ICA, CurvDat Examples (cont.)

Aapo Hyvärinen comments:

- “Random start” is deliberate choice
- Shows “local minina”, by different answers on replication

ICA for Corpora Collosa Data

Recall: shapes of “window” between brain halves

Show CorpColl\CCFrawAlls3.mpg

Discrimination problem: Schizophrenics vs. Controls

Show CorpColl\CCFrawSs3.mpg & CCFrawCs3.mpg

PCA: poor separation:

Show CorpColl\CCFpcaSCs3PC1.mpg ,CCFpcaSCs3PC2.mpg & CCFpcaSCs3PC3.mpg

ICA for Corpora Collosa Data (cont.)

ICA Problem: HDLSS, $71 = n < d = 80$

Solution: Work only with 1st 20 Principal Components

(Reason for 20 discussed later)

IC1: Seems to find an “outlier”

Show CorpColl\CCFicaSCs3IC1.mpg

- Outlier is Case 50

Show CorpColl\CCFrawAlls3num.mpg, Case 50

ICA for Corpora Collosa Data (cont.)

Similarly for other ICs:

IC	Outlier Case
2	60
3	2 (biggest outlier only 3rd IC? By sphering?)
4	26

Show CorpColl\CCFicaSCs3IC2.mpg, CCFicaSCs3IC3.mpg, CCFicaSCs3IC4.mpg, and CCFrawAlls3num.mpg, Cases 60, 2, 26

Reason:

Outlier distributions have high kurtosis, thus found by ICA

ICA for Corpora Collosa Data (cont.)

Solutions to “ICA driven by outliers” problem?

Sol’n 1: Reduce to only 1st 4 PCs:

Show CorpColl\CCFicaSCs3IC1v2.mpg, CCFicaSCs3IC2v2.mpg, CCFicaSCs3IC3v2.mpg, CCFicaSCs3IC4v2.mpg

- Same as PC space above, not good discrimination

Sol’n 2: Use PCA “starting values” (as for Parabs Up - Dn)

Show CorpColl\CCFicaSCs3IC1v3.mpg, CCFicaSCs3IC2v3.mpg, CCFicaSCs3IC3v3.mpg, CCFicaSCs3IC4v3.mpg

- found some different outliers – Cases 2, 30, *, 22
- found a “bimodal direction”
- but weak discrimination???