Spatial Summarization of Image Collections

Diego A. Ballesteros Villamizar

ETH Zürich

February 29th, 2016

Outline

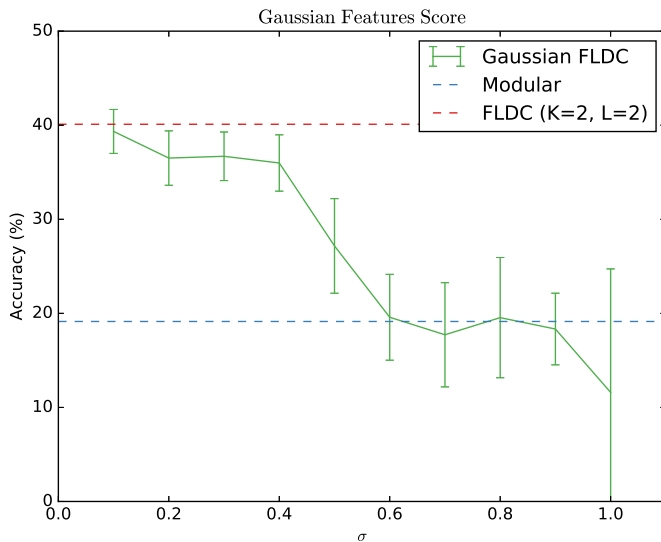1. Featurized Model

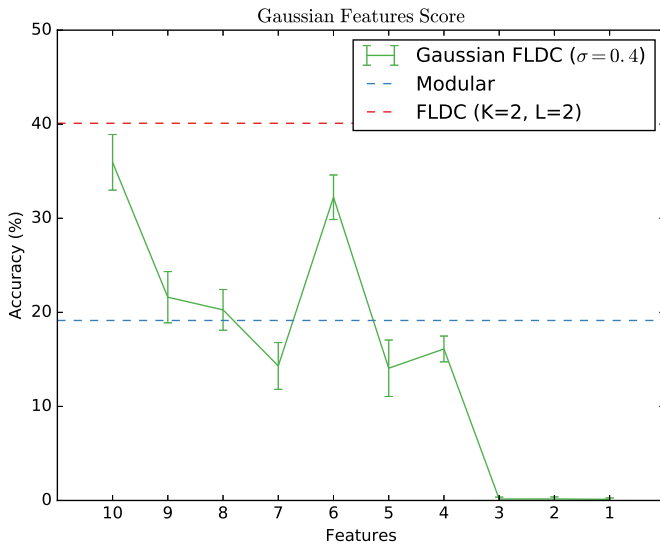2. Synthetic Datasets

Gaussian features

- Features matrix

$$\mathbf{X} \in \mathbb{R}^{|V| \times |V|}$$

$$x_{ij} = \exp\left(-\frac{(lat_j - lat_i)^2 + (lon_j - lon_i)^2}{2\sigma^2}\right)$$

- For $\sigma = 0$, $\mathbf{X} = \mathbb{I}$.
- Latitudes and longitudes are scaled to [0,1] on the set of items.

Effect number of features

Outline

One-dimension model

- FLID model with 3 items.

$$\mathbf{u} = [2, 1, 1]$$
$$\mathbf{W} = [0, 100, 100]^{\mathsf{T}}$$

- Distribution is:
  - {0}: 13.94%
  - {1}, {2}: 5.13%
  - {0,1}, {0,2}: 37.90%
  - {1,2}, {0,1,2}: 0.00%

Learning accuracy

- Dataset generated with 10k samples.
- Modular and FLID model trained on this dataset.
- When conditioning on 1 or 2, both models get the prediction right.
- When conditioning on 0, there is no information. Both 1 and 2 are equally likely, only in half of the cases the models get it right.

Model	Accuracy (%)
Modular	75.23 ± 0.82
FLID (L=1)	74.96 ± 2.04

Learned distribution

- Scores are equal but learned distributions are different.
- For modular (one of the folds):
    - {0}: 30.17%
    - {1}, {2}: 2.64%, 2.74%
    - {0,1}, {0,2}: 22.26%, 23.11%
    - {1,2}, {0,1,2}: 2.02%, 17.05%
- For FLID (one of the folds):
    - {0}: 16.07%
    - {1}, {2}: 5.59%, 5.17%
    - {0,1}, {0,2}: 37.36%, 34.59%
    - {1,2}, {0,1,2}: 0.16%, 1.07%
- Learned FLID model (one of the folds):

$$\mathbf{u} = [1.92, 0.87, 0.79]$$
$$\mathbf{W} = [0.02, 4.34, 4.34]^\intercal$$

FLDC dataset

- FLDC model with 4 items.

$$\mathbf{u} = [1, 1, 1, 1]$$

$$\mathbf{W}_D = \begin{pmatrix} 100 & 100 & 0 & 0 \\ 0 & 0 & 100 & 100 \\ 100 & 0 & 100 & 0 \\ 0 & 100 & 0 & 100 \end{pmatrix}$$

$$\mathbf{W}_C = \begin{pmatrix} 5 & 0 \\ 5 & 0 \\ 0 & 5 \\ 0 & 5 \end{pmatrix}$$

- Distribution is:
  - {0}, {1}, {2}, {3}: 10.60%
  - {0,1}, {2,3}: 28.81%
  - {0,2}, {0,3}, . . . : 0.00%
  - {0,1,2}, {0,1,3}, . . . : 0.00%

Model	Accuracy (%)
Modular	50.61 ± 1.72
FLID (L=1)	49.39 ± 1.72
FLID (L=2)	69.37 ± 25.02
FLID (L=3)	79.54 ± 25.10
FLID (L=4)	84.79 ± 23.25
FLDC (L=4, K=2)	100.0 ± 0.00

- 2 dimensions are enough to achieve 100% accuracy. But only a few of the folds learn properly.
- As the number of parameters increases, it's more likely that all folds will learn correctly.
- A good $L = 2$ model is:

$$\mathbf{u} = [0.18, 7.46, 3.68, 3.69]$$

$$\mathbf{W}_D = \begin{pmatrix} 0.02 & 0.02 \\ 8.46 & 8.47 \\ 8.46 & 0.02 \\ 0.02 & 8.47 \end{pmatrix}$$

Learned model

- For the same number of parameters as the original model.
- The learned FLDC model for $L = 4$ and $K = 2$ is:

$$\mathbf{u} = [0.01, -0.16, 1.89, 1.89]$$

$$\mathbf{W}_D = \begin{pmatrix} 5.25 & 5.25 & 0.05 & 0.05 \\ 0.039 & 0.039 & 5.22 & 5.22 \\ 5.25 & 0.034 & 0.034 & 5.22 \\ 0.035 & 5.25 & 5.22 & 0.035 \end{pmatrix}$$

$$\mathbf{W}_C = \begin{pmatrix} 3.48 & 3.48 \\ 3.50 & 3.50 \\ 1.39 & 1.36 \\ 1.36 & 1.39 \end{pmatrix}$$