Practical Session on Embeddings Applied to Zero-shot Learning

Yongqin Xian

Max-Planck Institute for Informatics

12 September 2016

Outline

- Zero-shot Learning
- Structural Joint Embedding
- Demos
- Latent Embedding
- Demos

Please download the code and data at:
`https://github.com/yqxian/GCPR_Tutorial`

Zero-shot Learning(ZSL) Problem

Training phase:

1. Observed data set

$$S = \{(x, y) | x \in X, y \in Y\}$$

Zero-shot Learning(ZSL) Problem

Training phase:

1. Observed data set

$$S = \{(x, y) | x \in X, y \in Y\}$$

2. Extract image and class embeddings

$$S_{\theta,\varphi} = \{(\theta(x), \varphi(y)) | \theta(x) \in \mathcal{X}, \varphi(y) \in \mathcal{Y}\}$$

$\theta(x)$: image embedding, $\varphi(y)$: class embedding

Zero-shot Learning(ZSL) Problem

Training phase:

1. Observed data set

$$S = \{(x, y) | x \in X, y \in Y\}$$

2. Extract image and class embeddings

$$S_{\theta,\varphi} = \{(\theta(x), \varphi(y)) | \theta(x) \in \mathcal{X}, \varphi(y) \in \mathcal{Y}\}$$

   $\theta(x)$: image embedding, $\varphi(y)$: class embedding

3. Using $S_{\theta,\varphi}$, learn function

$$f : X \to Y$$

Zero-shot Learning(ZSL) Problem

Training phase:

1. Observed data set

$$S = \{(x, y) | x \in X, y \in Y\}$$

2. Extract image and class embeddings

$$S_{\theta,\varphi} = \{(\theta(x), \varphi(y)) | \theta(x) \in \mathcal{X}, \varphi(y) \in \mathcal{Y}\}$$

   $\theta(x)$: image embedding, $\varphi(y)$: class embedding

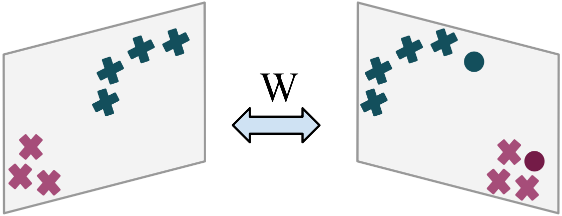
3. Using $S_{\theta,\varphi}$, learn function

$$f : X \to Y$$

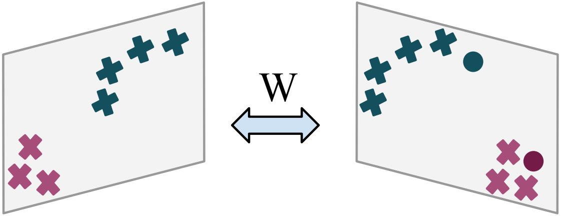
Testing phase:

1. Predict x^* of unseen classes using $f(x^*)$

Structural Joint Embedding



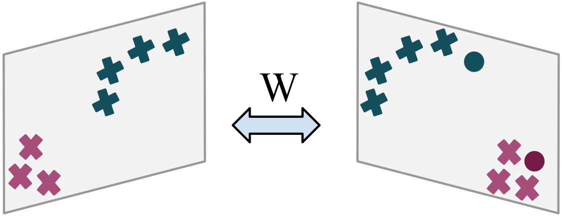
Structural Joint Embedding



Prediction function

$$f(x; W) = \arg\max_{y \in Y} \theta(x)^T W \varphi(y),$$

Structural Joint Embedding



Prediction function

$$f(x; W) = \arg\max_{y \in Y} \theta(x)^T W \varphi(y),$$

Objective function

$$\min_W \frac{1}{N} \sum_{n=1}^{N} \max_{y \in Y} \{0, \Delta(y_n, y) + \theta(x_n)^T W \varphi(y) - \theta(x_n)^T W \varphi(y_n)\},$$

Structural Joint Embedding

Algorithm: SGD optimization of SJE

1: Given $\mathcal{T} = \{(x, y) | x \in \mathbb{R}^{d_x}, y \in \mathbb{R}^{d_y}\}$
2: Initialize W randomly

Structural Joint Embedding

Algorithm: SGD optimization of SJE

```
Given $\mathcal{T} = \{(x, y) | x \in \mathbb{R}^{d_x}, y \in \mathbb{R}^{d_y}\}$
Initialize $W$ randomly
for $t = 1$ to $T$ do
    for $n = 1$ to $|\mathcal{T}|$ do
        Draw $(x_n, y_n) \in \mathcal{T}$
```

Structural Joint Embedding

Algorithm: SGD optimization of SJE

Given $\mathcal{T} = \{(x, y) | x \in \mathbb{R}^{d_x}, y \in \mathbb{R}^{d_y}\}$
Initialize W randomly
for $t = 1$ to T **do**
    for $n = 1$ to $|\mathcal{T}|$ **do**
        Draw $(x_n, y_n) \in \mathcal{T}$
        $y^* \leftarrow \underset{y \in \mathcal{Y} \setminus \{y_n\}}{\arg\max}\, x_n^\top W y$

Structural Joint Embedding

Algorithm: SGD optimization of SJE

Given $\mathcal{T} = \{(x, y) | x \in \mathbb{R}^{d_x}, y \in \mathbb{R}^{d_y}\}$
Initialize W randomly
for $t = 1$ to T **do**
  for $n = 1$ to $|\mathcal{T}|$ **do**
    Draw $(x_n, y_n) \in \mathcal{T}$
    $y^* \leftarrow \underset{y \in \mathcal{Y} \setminus \{y_n\}}{\arg\max}\, x_n^\top W y$
    if $x_n^\top W y^* + 1 > x_n^\top W y_n$ **then**
      $W \leftarrow W - \eta x_n (y^* - y_n)^\top$
    end if
  end for
end for

Demo

Dataset: Animal with Attributes, train 40 cls, test 10 cls

- Task 1: Image Classification

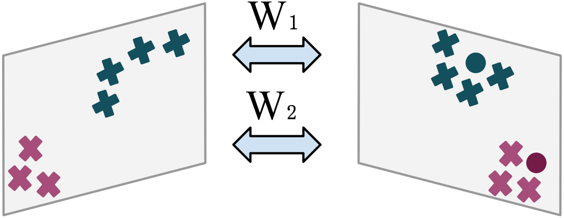
Demo

Dataset: Animal with Attributes, train 40 cls, test 10 cls

- Task 1: Image Classification
- Task 2: Image Retrieval

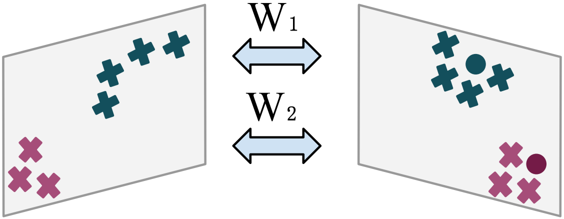
$$\arg\max_{x \in X} \theta(x)^\top W \varphi(y)$$

Latent Embeddings Method (LatEm)



$$F(x, y) = \max_{1 \leq i \leq K} \theta(x)^\top W_i \varphi(y)$$

Latent Embeddings Method (LatEm)



$$F(x, y) = \max_{1 \le i \le K} \theta(x)^\top W_i \varphi(y)$$

- Learn a collection of matrices
- Selection of which one to use is latent

Latent Embeddings Method (LatEm)

Loss function

$$L(x_n, y_n) = \sum_{y \in \mathcal{Y}} \left[\Delta(y_n, y) + F(x_n, y) - F(x_n, y_n)\right]_+ ,$$

Latent Embeddings Method (LatEm)

Loss function

$$L(x_n, y_n) = \sum_{y \in \mathcal{Y}} \left[\Delta(y_n, y) + F(x_n, y) - F(x_n, y_n)\right]_+ ,$$

Objective function

$$\frac{1}{N} \sum_{n=1}^{N} L(x_n, y_n).$$

Latent Embeddings Method (LatEm)

Algorithm: SGD optimization of LatEm

for all $t = 1$ to T **do**
    for all $n = 1$ to $|\mathcal{T}|$ **do**
        Draw $(x_n, y_n) \in \mathcal{T}$ and $y \in \mathcal{Y} \setminus \{y_n\}$

Latent Embeddings Method (LatEm)

Algorithm: SGD optimization of LatEm

for all $t = 1$ to T **do**
    for all $n = 1$ to $|\mathcal{T}|$ **do**
        Draw $(x_n, y_n) \in \mathcal{T}$ and $y \in \mathcal{Y} \setminus \{y_n\}$
        if $F(x_n, y) + 1 > F(x_n, y_n)$ **then**

Latent Embeddings Method (LatEm)

Algorithm: SGD optimization of LatEm

for all $t = 1$ to T **do**
    for all $n = 1$ to $|\mathcal{T}|$ **do**
        Draw $(x_n, y_n) \in \mathcal{T}$ and $y \in \mathcal{Y} \setminus \{y_n\}$
        if $F(x_n, y) + 1 > F(x_n, y_n)$ **then**
            $i^* \leftarrow \underset{1 \le k \le K}{\arg\max}\, x_n^\top W_k y$
            $j^* \leftarrow \underset{1 \le k \le K}{\arg\max}\, x_n^\top W_k y_n$

Latent Embeddings Method (LatEm)

Algorithm: SGD optimization of LatEm

for all $t = 1$ to T **do**
  for all $n = 1$ to $|\mathcal{T}|$ **do**
    Draw $(x_n, y_n) \in \mathcal{T}$ and $y \in \mathcal{Y} \setminus \{y_n\}$
    if $F(x_n, y) + 1 > F(x_n, y_n)$ **then**
      $i^* \leftarrow \underset{1 \leq k \leq K}{\arg\max}\, x_n^\top W_k y$
      $j^* \leftarrow \underset{1 \leq k \leq K}{\arg\max}\, x_n^\top W_k y_n$
      if $i^* = j^*$ **then**
        $W_{i^*}^{t+1} \leftarrow W_{i^*}^{t} - \eta_t x_n (y - y_n)^\top$
      end if

Latent Embeddings Method (LatEm)

Algorithm: SGD optimization of LatEm

for all $t = 1$ to T **do**
  for all $n = 1$ to $|\mathcal{T}|$ **do**
    Draw $(x_n, y_n) \in \mathcal{T}$ and $y \in \mathcal{Y} \setminus \{y_n\}$
    if $F(x_n, y) + 1 > F(x_n, y_n)$ **then**
      $i^* \leftarrow \underset{1 \leq k \leq K}{\arg\max}\, x_n^\top W_k y$
      $j^* \leftarrow \underset{1 \leq k \leq K}{\arg\max}\, x_n^\top W_k y_n$
      if $i^* = j^*$ **then**
        $W_{i^*}^{t+1} \leftarrow W_{i^*}^t - \eta_t x_n (y - y_n)^\top$
      end if
      if $i^* \neq j^*$ **then**
        $W_{i^*}^{t+1} \leftarrow W_{i^*}^t - \eta_t x_n y^\top$
        $W_{j^*}^{t+1} \leftarrow W_{j^*}^t + \eta_t x_n y_n^\top$
      end if
    end if
  end for
end for

Demo

Dataset: Animal with Attributes, train 40 cls, test 10 cls

- Task 1: Image Classification

Demo

Dataset: Animal with Attributes, train 40 cls, test 10 cls

- Task 1: Image Classification
- Task 2: Image Retrieval

$$\arg\max_{x \in X, y \in Y} \theta(x) W_i \varphi(y)$$

Reference

Akata, Zeynep, et al. "Evaluation of output embeddings for fine-grained image classification." IEEE CVPR 2015.

Xian, Yongqin, et al. "Latent Embedding for Zero-shot Classification." IEEE CVPR 2016.

Thank you!