Lecture
Tips for Homework 1

IE590 DEEP LEARNING IN MACHINE VISION

Tips for KNN

- Let $A^{2\times 2}$ and $X^{3\times 2}$ be two matrices. We want to compute the pairwise distance between each row in A to each row in X.
- Let $D^{2\times 3}$ be the distance matrix, where $D(i,j)$ represents the pairwise distance between i^{th} row in A to j^{th} row in X

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \qquad X = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}$$

Tips for KNN

$$A$$

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

$$X$$

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}$$

Tips for KNN

$$A \qquad\qquad X$$

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \qquad \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}$$

$$D(A,X) = \begin{bmatrix} (a_1 - x_1)^2 + (b_1 - y_1)^2 & (a_1 - x_2)^2 + (b_1 - y_2)^2 & (a_1 - x_3)^2 + (b_1 - y_3)^2 \\ (a_2 - x_1)^2 + (b_2 - y_1)^2 & (a_2 - x_2)^2 + (b_2 - y_2)^2 & (a_2 - x_3)^2 + (b_2 - y_3)^2 \end{bmatrix}$$

Tips for KNN

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \qquad X = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}$$

$$D(A,X) = \begin{bmatrix} (a_1 - x_1)^2 + (b_1 - y_1)^2 & (a_1 - x_2)^2 + (b_1 - y_2)^2 & (a_1 - x_3)^2 + (b_1 - y_3)^2 \\ (a_2 - x_1)^2 + (b_2 - y_1)^2 & (a_2 - x_2)^2 + (b_2 - y_2)^2 & (a_2 - x_3)^2 + (b_2 - y_3)^2 \end{bmatrix}$$

Don't use this equation literally !!!

$$D(A,X) = \sqrt{\Sigma A^2 \ - 2\,(A.X) + \Sigma X^2}$$

https://arxiv.org/pdf/1502.07541.pdf

SVM Loss

Naïve SVM Loss

$$L_i = \sum_{j \neq y_i} \max(0, w_j^T x_i - w_{y_i}^T x_i + \Delta)$$

Example

$$s = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix} \qquad y_0 = 0,\ y_1 = 1$$

$$l_0 = \max(0, -7 - 13 + 1) + \max(0, 14 - 13 + 1) = 2$$

$$l_1 = \max(0, -2 - 8 + 1) + \max(0, 10 - 8 + 1) = 3$$

3 classes
2 data points

$$Loss = 2 + 3 = 5$$

Example

Vectorized SVM Loss

$$s = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix}$$

$$y_0 = 0, \; y_1 = 1$$

$$L_i = \sum_{j \neq y_i} \max(0, w_j^T x_i - w_{y_i}^T x_i + \Delta)$$

Example

Vectorized SVM Loss

$$s = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix} \qquad y_0 = 0,\ y_1 = 1$$

$$L_i = \sum_{j \neq y_i} \max(0, w_j^T x_i - w_{y_i}^T x_i + \Delta)$$

$$s = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix} \longrightarrow c = \begin{bmatrix} 13 \\ 8 \end{bmatrix}$$

Example

Vectorized SVM Loss

$$s = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix}$$

$y_0 = 0,\ y_1 = 1$

$$L_i = \sum_{j \neq y_i} \max(0, w_j^T x_i - w_{y_i}^T x_i + \Delta)$$

$$s = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix} \longrightarrow c = \begin{bmatrix} 13 \\ 8 \end{bmatrix}$$

$$s_1 = s - c + 1$$

$$s_1 = \begin{bmatrix} 1 & -19 & 2 \\ -9 & 1 & 3 \end{bmatrix}$$

$$s_2 = \begin{bmatrix} 0 & -19 & 2 \\ -9 & 0 & 3 \end{bmatrix}$$

Element wise $\max(s, 0)$

$$s_2 = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

$Loss = sum\ of\ all\ elements = 5$

SoftMax Loss

Example

SoftMax Classifier

- Say there are 3 classes and two data points
- Let S be the score function i.e. $S = X\,W$

Example

Vectorized SoftMax Loss

$$S = \begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix} \qquad y_0 = 0, \; y_1 = 1$$

$$L_s = -\frac{1}{M}\sum_{i=1}^{M} \log \frac{\exp\left(W_{y_i}^T f(x_i) + b_{y_i}\right)}{\sum_{j=1}^{C} \exp\left(W_j^T f(x_i + b_j)\right)}$$

Take exponential

$$S = e^{\begin{bmatrix} 13 & -7 & 14 \\ -2 & 8 & 10 \end{bmatrix}} = \begin{bmatrix} e^{13} & e^{-7} & e^{14} \\ e^{-2} & e^{8} & e^{10} \end{bmatrix}$$

$$S = 1e + 06 \; * \begin{bmatrix} 0.44 & 0.00 & 1.20 \\ 0.00 & 0.00 & 0.02 \end{bmatrix}$$

Example

Vectorized SoftMax Loss

$$S = 1e+06 \ * \begin{bmatrix} 0.44 & 0.00 & 1.20 \\ 0.00 & 0.00 & 0.02 \end{bmatrix}$$

$$L_s = -\frac{1}{M}\sum_{i=1}^{M} \log \frac{\exp\left(W_{y_i}^T f(x_i) + b_{y_i}\right)}{\sum_{j=1}^{C} \exp\left(W_j^T f(x_i + b_j)\right)}$$

Find max in each row

$$Z = 1e+06 \ * \begin{bmatrix} 0.44 \\ 0.02 \end{bmatrix}$$

Subtract max from each row

$$S = S \ - Z = 1e+06 \ * \begin{bmatrix} -0.76 & -1.20 & 0.00 \\ -0.02 & -0.02 & 0.00 \end{bmatrix}$$

Example

Vectorized SoftMax Loss

$$S = 1e+06 \ * \begin{bmatrix} -0.76 & -1.20 & 0.00 \\ -0.02 & -0.02 & 0.00 \end{bmatrix}$$

$$L_s = -\frac{1}{M}\sum_{i=1}^{M} \log \frac{\exp\left(W_{y_i}^T f(x_i) + b_{y_i}\right)}{\sum_{j=1}^{C} \exp\left(W_j^T f(x_i + b_j)\right)}$$

Find sum of each row

$$Q = 1e+06 \ * \begin{bmatrix} -1.96 \\ -0.04 \end{bmatrix}$$

Normalize each row

$$S = \frac{S}{Z} = \begin{bmatrix} 0.39 & 0.61 & 0.00 \\ 0.54 & 0.46 & 0.00 \end{bmatrix}$$

Example

Vectorized SoftMax Loss

$$S = \begin{bmatrix} 0.39 & 0.61 & 0.00 \\ 0.54 & 0.46 & 0.00 \end{bmatrix}$$

$$S = \begin{bmatrix} 0.39 & 0.61 & 0.00 \\ 0.54 & 0.46 & 0.00 \end{bmatrix} \longrightarrow c = \begin{bmatrix} 0.39 \\ 0.46 \end{bmatrix}$$

$$L = -0.5 * (\log(0.39) + \log(0.46))$$

$$L = -1.44$$

Math

SoftMax Classifier

- No. of data points - n
- Dimension of the data - d
- No. of classes - c
- Subscript i indicates i^{th} row
- Superscript j indicates j^{th} column
- Input data matrix - $X^{n \times d}$
- Weight matrix - $W^{d \times c}$
- Scores ($S^{n \times c}$)

Math

SoftMax Classifier

$$S = X\,W$$

- S_i is i^{th} row in S
- S_i^j is the element in the i^{th} row and j^{th} column

$$S_i = X_i^T\,W$$

$$\Rightarrow\; S_i = X_i^T\,[W^1\,|\,W^2\,|\,\ldots\,|\,W^c]$$