cs224d Notations

mxn

Contents

1 Basic Notations

- Vector: $\boldsymbol{v} \in \mathbb{R}^n$ or $\boldsymbol{v} \in \mathbb{R}^{n \times 1}$ — n rows and 1 column.

    - numpy: `v.shape == (n,)` or `(n,1)`

- Matrix: $\mathbf{M} \in \mathbb{R}^{n \times m}$ — n rows and m columns.

    - numpy: `M.shape == (n, m)`

1.0.1 numpy revision

- creating a range as column vector (n rows, 1 column): `v = np.c_[:n]`

- creating a range as row vector (1 row, m columns): `v = np.r_[:n]`

1.1 Common Network Variable and Dimension Notations

- n: Dimension of embeddings space

- V: Vocabulary with $|V|$ words

- w_i: Word i from vocabulary V

- $\boldsymbol{x}$: Input vectors (usually one-hot vectors)

- $\boldsymbol{y}$: Output vectors (usually one-hot vectors, or vectors of probabilities (after softmax))

- 1 layer network:

    - Input dimension: D_x ($\boldsymbol{x} \in [N, D_x]$)
    - Hidden units: H ($\mathbf{W}_1 \in [D_x, H]$, $\boldsymbol{b}_1 \in [H]$)
    - Ouput dimension: D_y ($\mathbf{W}_2 \in [H, D_y]$, $b_2 \in [D_y]$)

- Continuous Bag of Words (CBOW):

    - $\mathbf{V} \in \mathbb{R}^{n \times |V|}$: Input word matrix
    - $\mathbf{U} \in \mathbb{R}^{|V| \times n}$: Output word matrix
    - $\boldsymbol{v}_i \in \mathbb{R}^{n \times 1}$: i-th column vector of $\mathbf{V}$, the input vector representing word w_i
    - $\boldsymbol{u}_i \in \mathbb{R}^{1 \times n}$: i-th row vector of $\mathbf{U}$, the output vector representing word w_i