# Introduction to learning of text representations

Prakhar Gupta

Department of Computer and Communication Sciences
EPFL

Advanced Topics in Machine Learning

# Outline

# Outline

# Matrix Factorization Methods

# Matrix Factorization Methods

- Models that can be described as optimization problems of the form

$$\min_{U,V} F(UV^T) \tag{1}$$

# GloVe: Global Vectors for Word Representation

## Some important definitions and notation

# GloVe: Global Vectors for Word Representation

## Some important definitions and notation

1. Context window is defined as the neighbourhood of a word and its length can be chosen as desired.

# GloVe: Global Vectors for Word Representation

## Some important definitions and notation

1. Context window is defined as the neighbourhood of a word and its length can be chosen as desired.
2. Word co-occurence matrix is denoted by $X$

# GloVe: Global Vectors for Word Representation

## Some important definitions and notation

1. Context window is defined as the neighbourhood of a word and its length can be chosen as desired.
2. Word co-occurence matrix is denoted by $X$
    1. where $X_{ij}$ = Number of times word $j$ appears in context of word $i$

# GloVe: Global Vectors for Word Representation

# GloVe: Global Vectors for Word Representation

1. The aim is to capture $X_ij$ using the source embeddings $u_i$ and target embeddings $v_j$

# GloVe: Global Vectors for Word Representation

1. The aim is to capture $X_i j$ using the source embeddings $u_i$ and target embeddings $v_j$
2. The GloVe problem is thus formulated as

$$\min_{U,V} \sum i, j \in W f(X_{ij})(u_i^T v_j + b_i + c_j + log(X_i j)) \tag{2}$$

where $f(X_{ij})$ is the weight assigned to the source-target pair, $b_i$ and $c_j$ are the biases associated with $u_i$ and $v_j$ respectively and $W$ is the vocabulary.

# GloVe: Global Vectors for Word Representation

1. The aim is to capture $X_i j$ using the source embeddings $u_i$ and target embeddings $v_j$
2. The GloVe problem is thus formulated as

$$\min_{U,V} \sum i, j \in W f(X_{ij})(u_i^T v_j + b_i + c_j + log(X_i j)) \tag{2}$$

where $f(X_{ij})$ is the weight assigned to the source-target pair, $b_i$ and $c_j$ are the biases associated with $u_i$ and $v_j$ respectively and $W$ is the vocabulary.

3. $f(X_{ij})$ is often chosen to be $(\frac{X_{ij}}{Y})^{\alpha}$ where $Y = \max_{kl} X_{kl}$.

# GloVe: Global Vectors for Word Representation

1. The aim is to capture $X_i j$ using the source embeddings $u_i$ and target embeddings $v_j$
2. The GloVe problem is thus formulated as

$$\min_{U,V} \sum i, j \in W f(X_{ij})(u_i^T v_j + b_i + c_j + log(X_i j)) \tag{2}$$

   where $f(X_{ij})$ is the weight assigned to the source-target pair, $b_i$ and $c_j$ are the biases associated with $u_i$ and $v_j$ respectively and $W$ is the vocabulary.
3. $f(X_{ij})$ is often chosen to be $(\frac{X_{ij}}{Y})^{\alpha}$ where $Y = \max_{kl} X_{kl}$.
4. Empirically $\alpha = \frac{3}{4}$ gives the best performance.

# Outline

# Second Slide Title

- First item.

# Second Slide Title

- First item.
- Second item.

# Second Slide Title

- First item.
- Second item.
- Third item.

# Second Slide Title

- First item.
- Second item.
- Third item.
- Fourth item.

# Second Slide Title

- First item.
- Second item.
- Third item.
- Fourth item.
- Fifth item.

# Second Slide Title

- First item.
- Second item.
- Third item.
- Fourth item.
- Fifth item. Extra text in the fifth item.

# Outline

# Blocks

## Block Title

You can also highlight sections of your presentation in a block, with it's own title

## Theorem

*There are separate environments for theorems, examples, definitions and proofs.*

## Example

Here is an example of an example block.

# Summary

- The first main message of your talk in one or two lines.
- The second main message of your talk in one or two lines.
- Perhaps a third message, but not more than that.

- Outlook
  - Something you haven't solved.
  - Something else you haven't solved.

# For Further Reading I


A. Author.
*Handbook of Everything*.
Some Press, 1990.

S. Someone.
On this and that.
*Journal of This and That*, 2(1):50–100, 2000.