Dynamic content monitoring and exploration using vector spaces

Benyou Wang
University of Padua
Supervised by Massimo Melucci and Emanuele Di Buccio
SIGIR DC, Paris, France, 07/2019
Dynamics

- Sequential data
- Language model and Language generation
- Historical corpora (Google book, ArXiv paper collections)
- Conversation/dialogue
- Recommendation with historical interactions
- Video tracking
- Electrical Healthy records

Focusing on **textual** problem
Taxonomy in textual applications

Why vector spaces?

✓ Well-defined properties of vector space and it is naturally used in IR (VSM in IR).

✓ Representing words in vector space is a commonly-used paradigm in textual problems [1]

Currently, there is some shortage in modelling dynamic aspects in vector space

Methods

- Inner dynamics
  - In-between dynamics
  - high-dimensional dynamics

- Sequence-sensitive vector space
- Evolved vector space
- Extended high-dimensional vector space
Inner dynamics - Encoding order for a single vector

- Problems: position-sensitive vector without recurrent architectures
**Inner dynamics - Encoding order for a single vector**

Previous methods: Position Embedding (PE) + Word Embedding (WE)

\[ f: (N, N) \rightarrow R^k, \text{ e.g., } PE(pos_i) + WE(w_j) \text{ or } [PE(pos_i); WE(w_j)] \]

Extending embedding from a vector to a **continuous function** over variable the position (pos)

Technically, \[ f: (N, N) \rightarrow R^k \] To \[ f: (N) \rightarrow G\{g; g: N \rightarrow R^k\} \]

- Word index
- position index

Now the question becomes *how to decide the function*
Properties for $f$

Now, for a specific word $w$, we have to get its embedding over all the positions, namely a function $g_w: \mathbb{N} \rightarrow \mathbb{R}^k$

**Property 1: Invariant relative-distance transformation**

With the any linear transformation $f_{ij}$, s.t. $f_{ij}(g_w(pos_i)) = g_w(pos_j)$, one should know whether $pos_i$ is in front of $pos_j$, as well as how far is between $pos_i$ and $pos_j$, no matter how big $pos_i$ and $pos_j$ is.

**Property 2: Boundedness**

The function $g_w$ should be bounded, in order to model long-enough sentence
Encoding order in complex embedding

The **only** solution of $g_w: \mathbb{N} \to \mathbb{R}^k$ to meet the previous properties is

$$g_w(pos) = \text{WE}_w \odot e^{i \cdot \text{pos} \cdot \lambda}$$

For the $k$-th dimension, $[g_w(pos)]_k = \text{WE}_{w,k} \cdot e^{i \cdot \text{period}_k \cdot \text{pos}}$

With Euler’ formula, we can get

$$[g_w(pos)]_k = \text{WE}_{w,k} \cdot [\cos(\text{period}_k \cdot \text{pos}) + i \sin(\text{period}_k \cdot \text{pos})]$$
Methods

- Inner dynamics
- **In-between dynamics**
- high-dimensional dynamics
- ✓ Sequence-sensitive vector space
- ✓ **Evolved vector space**
- ✓ Extended high-dimensional vector space
Quantum probability

- **Classical probability**
  - Set-based probability theory
  - Events are limited to be discrete and mutually-exclusive

- **Quantum probability**
  - Projective geometry based probability theory
  - Events, which is defined in a complex continuous vector space, can be represented as arbitrary vector
Density matrix in vector space

Evolved density matrix

\[ \rho_{t_{k+1}} = u \rho_{t_k} u^*, \text{ where } u \text{ is unitary} \]

\[ \frac{d \rho_{t_k}}{dt} = f(\rho, t_k) \]
Advantages of Evolved density matrix

• Better interpretability from quantum probability [1]
• Better optimization based on unitary transformation [2]
• Linking with neural (Schrödinger) differential equation [3]

Future works

- Investigating the dynamic aspects in high-dimensional vector space, i.e., tensor space

---

Correspondence between languages of Tensor Analysis and Deep Learning.

<table>
<thead>
<tr>
<th>Tensor Decompositions</th>
<th>Deep Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP-decomposition</td>
<td>shallow network</td>
</tr>
<tr>
<td>TT-decomposition</td>
<td>RNN</td>
</tr>
<tr>
<td>HT-decomposition</td>
<td>CNN</td>
</tr>
<tr>
<td>rank of the decomposition</td>
<td>width of the network</td>
</tr>
</tbody>
</table>

Publications


5. Wei Zhao, Wang Benyou, Jianbo Ye, Yongqiang Gao, Min Yang, Xiaojun Chen, PLASTIC: Prioritize Long and Short-term Information in Top-n Recommendation using Adversarial Training, IJCAI 2018

6. Wei Zhao, Wang Benyou, Jianbo Ye, Min Yang, Zhou Zhao, Ruotian Luo, Yu Qiao A Multi-task Learning Approach for Image Captioning, IJCAI 2018


Thanks