Computer Science Department, University of Bonn

Knowledge Graph Analysis

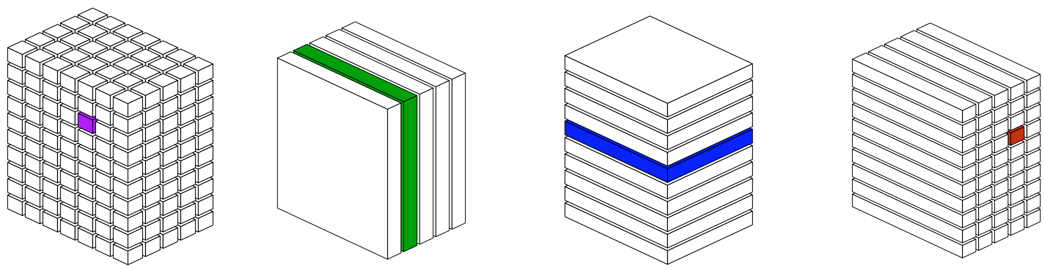
Solutions to Exercise Sheet 4

Dr. Asja Fischer, Prof. Jens Lehmann

February 6, 2018

1 In Class

1. **Tensors: notation and terminology**



(a) (b) (c) (d)

Figure 1.1: (a) - the tensor entry $\mathcal{T}_{3,5,1}$, (b) - 2nd frontal slice $\mathcal{T}_{:,:,2}$, (c) - the fifth horizontal slice $\mathcal{T}_{5,:,:}$, and (d) - the mode-2 fiber $\mathcal{T}_{4,:,4}$

2. **Adjacency tensors**

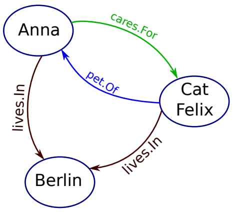


Figure 1.2: Knowledge graph described by the adjacency tensor.

3. **Convert Tensor into Matrix**

$$\boldsymbol{X}_{(1)} = \begin{bmatrix} 1 & 5 & 3 & 7 \\ 2 & 6 & 4 & 8 \end{bmatrix} \tag{1.1}$$

$$\boldsymbol{X}_{(2)} = \begin{bmatrix} 1 & 2 & 5 & 6 \\ 3 & 4 & 7 & 8 \end{bmatrix} \tag{1.2}$$

$$\boldsymbol{X}_{(3)} = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 7 & 6 & 8 \end{bmatrix} \tag{1.3}$$

for more details check `https://sewiki.iai.uni-bonn.de/_media/teaching/lectures/kga/2017/tensor_ex_3-9.jpg`.

4. **Matrix factorization**
For

$$\boldsymbol{X} = \begin{bmatrix} 5 & 3 & 0 & 1 \\ 4 & 0 & 0 & 1 \\ 1 & 1 & 0 & 5 \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 5 & 4 \end{bmatrix}$$

the prediction based on the factorization into 2 latent dimensions is given by

$$\hat{\boldsymbol{X}} = \begin{bmatrix} 5.01 & 2.59 & 5.44 & 1.02 \\ 3.88 & 1.97 & 4.23 & 0.99 \\ 2.11 & 1.08 & 3.94 & 4.65 \\ 1.66 & 0.85 & 3.13 & 3.74 \\ 3.28 & 1.67 & 4.83 & 3.99 \end{bmatrix} = \begin{bmatrix} 2.25 & 0.97 \\ 1.68 & 0.82 \\ 0.26 & 1.97 \\ 0.19 & 1.58 \\ 0.92 & 1.86 \end{bmatrix} \begin{bmatrix} 1.91 & 0.97 & 1.65 & -0.6 \\ 0.82 & 0.42 & 1.78 & 2.44 \end{bmatrix}$$

5. **Loss function for matrix factorization**

The Mean Squared Error (MSE) is given by

$$
\begin{aligned}
||\boldsymbol{X} - \hat{\boldsymbol{X}}||^2 &= \sum_{ij}(X_{ij} - \hat{X}_{ij})^2 \\
&= 0.1^2 + 0.2^2 + 0.1^2 + 0.3^2 + 0.1^2 + 0.1^2 \\
&= 0.17
\end{aligned}
$$

6. **Outer product**
   Given are the following three vectors

$$
\begin{aligned}
\boldsymbol{a} &= \begin{bmatrix} 1 & 3 & 2 \end{bmatrix} \\
\boldsymbol{b} &= \begin{bmatrix} 4 & 2 & 1 \end{bmatrix} \\
\boldsymbol{c} &= \begin{bmatrix} 1 & 3 & 4 \end{bmatrix}
\end{aligned}
$$

   The $\boldsymbol{\mathcal{T}}_{;,1,1}$ fiber of the rank-a tensor $\boldsymbol{\mathcal{T}}$ resulting from the outer product of the vectors

$$\mathcal{T} = \boldsymbol{a} \circ \boldsymbol{b} \circ \boldsymbol{c} \ .$$

   is given by

$$\begin{bmatrix} 4 & 12 & 8 \end{bmatrix}^{\mathrm{T}} \ ,$$

   since the ijk-th element of $\boldsymbol{\mathcal{T}}$ is given by $\mathcal{T}_{ijk} = a_i * b_j * c_k$ and

$$\boldsymbol{\mathcal{T}}_{:,1,1} = \begin{bmatrix} a_1 * b_1 * c_1 & a_2 * b_1 * c_1 & a_3 * b_1 * c_1 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} a_1 * 4 & a_2 * 4 & a_3 * 4 \end{bmatrix}^{\mathrm{T}} \ .$$

7. **Collective learning in bipartide model**
   In a bipartide model (like CP) for each entity two latent representations are learned, one describing its role as a subject the other its role as an object of relations. In a unipartide model (Like RESCAL) an unique latent representation is learned for each entity. This enables information flow and improves the ability for collective learning.

8. **RESCAL**
   A) Given the entity matrix

$$\boldsymbol{A} = \begin{bmatrix} 0.1 & 0.8 & 0.1 \\ 0.7 & 0.2 & 0.1 \end{bmatrix} \tag{1.4}$$

   and the k-th frontal slice of the relation tensor

$$\boldsymbol{R}_{:,:,k} = \begin{bmatrix} 1 & 0 & 1.2 \\ 1 & 1.2 & 0 \\ 0.1 & 0 & 1 \end{bmatrix} \tag{1.5}$$

   the k-th slice of approximated tensor is give by

$$\boldsymbol{A}\boldsymbol{R}_{:,:,k}\boldsymbol{A}^T = \begin{bmatrix} 0.881 & 0.851 \\ 0.377 & 0.779 \end{bmatrix} \ . \tag{1.6}$$

The result describes the confidence of the model that the two entities are connected via the k-th relation (entries close to 1/0 indicate that a relation exists/not exists).

B)

$$\boldsymbol{A} = \begin{bmatrix} latent - representations - of - CatFelix \\ latent - representations - of - Anna \\ latent - representations - of - Berlin \end{bmatrix}. \tag{1.7}$$

$$\boldsymbol{R}_{caresFor} \tag{1.8}$$

$$\boldsymbol{R}_{petOf} \tag{1.9}$$

$$\boldsymbol{R}_{livesIn} \tag{1.10}$$

$$\mathrm{P}(X|A,\mathbf{R}) = \prod_{i=1}^{n}\prod_{j=1}^{n}\prod_{k=1}^{m} \mathrm{P}(x_{ijk}|\, \boldsymbol{a}_i^T R_k \boldsymbol{a}_j).$$

9. **Applications of RESCAL**

a) **Link Prediction**: Let

$$\boldsymbol{\mathcal{T}}_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{1.11}$$

be the 1-mode presentation of a tensor and

$$\hat{\boldsymbol{\mathcal{T}}}_{(1)} = \begin{bmatrix} 1.2 & 0.02 & 0.03 & 0.89 & 0.9 & 0.01 & 0.1 & 1.04 & 0.03 \\ 0 & 0.95 & 0.04 & 0.07 & 0.11 & 1.1 & 0.07 & 1.05 & 0.0 \\ 1.1 & 1 & 0.01 & 0.2 & 1.2 & 1 & 0.1 & 0.08 & 0.09 \end{bmatrix} \tag{1.12}$$

its approximation gained by a tensor decomposition technique.

check here for details how to fold back the tensor and get the predicted links `https://sewiki.iai.uni-bonn.de/_media/teaching/lectures/kga/2017/tensor_ex_3-9.jpg`

New links in the knowledge graph are (e_1, r_2, e_1) (corresponding to $\hat{\mathcal{T}}_{112} = 0.89$), (e_2, r_1, e_2) (corresponding to $\hat{\mathcal{T}}_{221} = 0.95$), and (e_2, r_2, e_2) (corresponding to $\hat{\mathcal{T}}_{322} = 1.2$).

Links that got removed are (e_2, r_1, e_3) (corresponding to $\hat{\mathcal{T}}_{231} = 0.04$) and (e_2, r_3, e_1) (corresponding to $\hat{\mathcal{T}}_{213} = 0.07$).

b) **Entity Similarity**: Let

$$\boldsymbol{A} = \begin{bmatrix} 0.8 & 0.3 & 1.7 \\ 0.9 & 0.4 & 1.6 \\ 0.2 & 0.2 & 0.3 \\ 1.0 & 0.3 & 1.5 \end{bmatrix} \tag{1.13}$$

be the entity matrix gained by RESCAL. Entities 1,2 and 4 corresponding to the latent features $\boldsymbol{a}_1 = (0.8, 0.3, 1.7), \boldsymbol{a}_2 = (0.9, 0.4, 1.6)$ and $\boldsymbol{a}_3 = (1.0, 0.3, 1.5)$ respectively are similar to each other (since similarity of entities can be measured based on the similarity of their latent representations).

c) **Relation Similarity**: Given are the k-th, the l-th, and the m-th frontal slice of the relation tensor of RESCAL

$$\boldsymbol{R}_{:,:,k} = \begin{bmatrix} 1.1 & 0.9 & 0.9 \\ 1.3 & 1.2 & 0.8 \\ 0.9 & 1.1 & 1.1 \end{bmatrix} \tag{1.14}$$

$$\boldsymbol{R}_{:,:,l} = \begin{bmatrix} 2 & 5 & 9 \\ 3 & 0.5 & 1 \\ 0.3 & 4 & 7 \end{bmatrix} \tag{1.15}$$

$$\boldsymbol{R}_{:,:,m} = \begin{bmatrix} 1.2 & 1.1 & 0.8 \\ 1.0 & 0.8 & 0.8 \\ 0.8 & 0.9 & 0.9 \end{bmatrix} \tag{1.16}$$

The m-th antd the k-th relation are similar (since similarity of relations can be measured based on the similarity of their latent representations).

2 At Home

1. **Singular value decomposition (SVD)**

$$\lambda_1 u_1 \circ v_1 = 9.03 \times \begin{bmatrix} -0.44 \\ -0.3 \\ -0.52 \\ -0.4 \\ -0.52 \end{bmatrix} \times \begin{bmatrix} -0.47 & -0.78 & 0.17 & 0.37 \end{bmatrix} =$$

$$\begin{bmatrix} 1.867404 & 3.099096 & -0.675444 & -1.470084 \\ 1.27323 & 2.11302 & -0.46053 & -1.00233 \\ 2.206932 & 3.662568 & -0.798252 & -1.737372 \\ 1.69764 & 2.81736 & -0.61404 & -1.33644 \\ 2.291814 & 3.803436 & -0.828954 & -1.804194 \end{bmatrix}$$

$$\lambda_2 u_2 \circ v_2 = 6.23 \times \begin{bmatrix} -0.67 \\ -0.44 \\ 0.14 \\ 0.11 \\ 0.57 \end{bmatrix} \times \begin{bmatrix} -0.26 & -0.21 & 0.25 & -0.91 \end{bmatrix} =$$

$$\begin{bmatrix} 1.085266 & 0.876561 & -1.043525 & 3.798431 \\ 0.712712 & 0.575652 & -0.6853 & 2.494492 \\ -0.226772 & -0.183162 & 0.21805 & -0.793702 \\ -0.178178 & -0.143913 & 0.171325 & -0.623623 \\ -0.923286 & -0.745731 & 0.887775 & -3.231501 \end{bmatrix}$$

$$\lambda_3 u_3 \circ v_3 = 3.77 \times \begin{bmatrix} -0.3 \\ -0.05 \\ -0.05 \\ -0.48 \\ 0.61 \end{bmatrix} \times \begin{bmatrix} -0.3 & 0.46 & 0.81 & 0.21 \end{bmatrix} =$$

$$\begin{bmatrix} 0.3393 & -0.52026 & -0.91611 & -0.23751 \\ 0.05655 & -0.08671 & -0.152685 & -0.039585 \\ 0.05655 & -0.08671 & -0.152685 & -0.039585 \\ 0.54288 & -0.832416 & -1.465776 & -0.380016 \\ -0.68991 & 1.057862 & 1.862757 & 0.482937 \end{bmatrix}$$

$$\lambda_4 u_4 \circ v_4 = 1.84 \times \begin{bmatrix} -0.49 \\ 0.8 \\ -0.29 \\ 0.21 \\ 0.08 \end{bmatrix} \times \begin{bmatrix} -0.78 & 0.37 & -0.5 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} 0.703248 & -0.333592 & 0.4508 & 0 \\ -1.14816 & 0.54464 & -0.736 & 0 \\ 0.416208 & -0.197432 & 0.2668 & 0 \\ -0.301392 & 0.142968 & -0.1932 & 0 \\ -0.114816 & 0.054464 & -0.0736 & 0 \end{bmatrix}$$

$$X = \sum_{i=1}^{4} \lambda_i u_i \circ v_i = \begin{bmatrix} 3.995218 & 3.121805 & -2.184279 & 2.090837 \\ 0.894332 & 3.146602 & -2.034515 & 1.452577 \\ 2.452918 & 3.195264 & -0.466087 & -2.570659 \\ 1.76095 & 1.983999 & -2.101691 & -2.340079 \\ 0.563802 & 4.170031 & 1.847978 & -4.552758 \end{bmatrix}$$