UnSupervised Tensor Factorization on dnations data

Zhuoyan Xu

1 Unsupervised Tensor factorization

The general model is:

$$logit\left\{\mathbb{E}\left[\mathcal{X}^{d_1 d_2 d_3}\right]\right\} = \mathcal{G}^{r_1 r_2 r_3} \times_1 A^{d_1 r_1} \times_2 B^{d_2 r_2} \times_3 C^{d_3 r_3}$$

Where d_1, d_2, d_3 is dimension of tensor. The r_1, r_2, r_3 is the dimension of low rank tensor.

For each of the factor matrix and core tensor, we can unfold or vectorize it to deduce a matrix form product, then use GLM to compute.

1.1 Matrix form GLM

Before we deduce, we first introduce a matrix form GLM. Consider

$$logit[E(Y))_{n*p}] = U_{n*p} = X_{n*R} \times \beta_{R*p}$$

Where $U_{n*p} = (u_1, \ldots, u_p)$, $\beta_{R*p} = (\beta_1, \ldots, \beta_p)$

We have:

$$\begin{gathered} u_1^{N*1} = X^{n*R} \times \beta_1^{R*1} \\ u_2^{N*1} = X^{n*R} \times \beta_2^{R*1} \\ \vdots \\ u_p^{N*1} = X^{n*R} \times \beta_P^{R*1} \end{gathered}$$

Then we implement GLM.

1.2 Vectorization

Then we define a vectorization of the tensor in this scenario: When we say vectorize a tensor G (3 dimension), we extract all its mode-3 fiber and then list it according first mode-2 then according to mode-1. In other words, when we count the vector elements, the element's third mode index change first, then second mode index, finally first mode index.

Here is a small example for understanding:

$$Vec(G^{2*2*2}) = \begin{bmatrix} G_{111} \\ G_{211} \\ G_{121} \\ G_{221} \\ G_{112} \\ G_{222} \\ G_{112} \\ G_{222} \end{bmatrix}$$

The reason why we define like this is this is how as.vector() in r works.
In R: Vec(G) = as.vector(G).

1.3 Optimization/Updating method

Consider the general model, we can deduce a matrix form equation, then apply a matrix form GLM.

In this section use Y to denote the logits $logit(\mathbb{E}(Y))$.

1.3.1 The update of factor matrix A,B,C

$$Y^{d_1 d_1 d_3} = G^{r_1 r_2 r_3} \times_1 A^{d_1 r_1} \times_2 B^{d_2 r_2} \times_3 C^{d_3 r_3} = G_{AB}^{r_1 r_2 d_3} \times_3 C^{d_3 r_3}$$

Update A,B,C through:

$$Y_{(3)}^T = G_{AB(3)}^T C^T$$

Since for distinct modes in a series of multiplications, the order of the multiplication is irrelevant. Thus, A,B,C the same update way.

1.3.2 The update of core tensor $\mathcal{G}$

Consider

$$Y^{d_1 * d_2 * d_3} = G^{r_1 * r_2 * r_3} \times_1 A^{d_1 * r_1} \times_2 B^{d_2 * r_2} \times_3 C^{d_3 * r_3}$$

We have:

$$Y_{ijk} = \sum_{k'=1}^{r_3} \sum_{j'=1}^{r_2} \sum_{i'=1}^{r_1} G_{i'j'k'} A_{ii'} B_{jj'} C_{kk'}$$

Let

$$[M^{ijk}]_{i'j'k'} = A_{ii'} B_{jj'} C_{kk'}$$

$i = 1, \ldots, d_1; j = 1, \ldots, d_2; k = 1, \ldots, d_3$ There are $d_1 * d_2 * d_3$ M^{ijk} tensors totally.The dimension of M^{ijk} is $r_1 * r_2 * r_3$.

Then we have:

$$Y_{ijk} = \sum_{k'=1}^{r_3} \sum_{j'=1}^{r_2} \sum_{i'=1}^{r_1} G_{i'j'k'} M_{i'j'k'}^{ijk}$$

Since we have:

$$\text{Vec}\left(M^{ijk}\right) = C_{k:} \otimes B_{j:} \otimes A_{i:}$$

Thus:

$$\begin{aligned} Y_{ijk} &= \sum_{k'=1}^{r_3}\sum_{j'=1}^{r_2}\sum_{i'=1}^{r_1} G_{i'j'k'} M^{ijk}_{i'jk'} \\ &= Vec(G)^T \, \text{Vec}\left(M^{ijk}\right) = \text{Vec}\left(M^{ijk}\right)^T_{1*r_1r_2r_3} Vec(G)_{r_1r_2r_3*1} \end{aligned}$$

First step:

$$Y_{:jk} = \begin{bmatrix} y_{1jk} \\ y_{2jk} \\ \vdots \\ y_{d_1jk} \end{bmatrix}_{d_1*1} = \begin{bmatrix} Vec(M^{1jk})^T \\ Vec(M^{2jk})^T \\ \vdots \\ Vec(M^{d_1jk})^T \end{bmatrix}_{d_1*r_1r_2r_3} \times Vec(G)_{r_1r_2r_3*1} = Vec(M_{d_1jk}) \times Vec(G)$$

Second step:

$$Y_{::k} = \begin{bmatrix} y_{:1k} \\ y_{:2k} \\ \vdots \\ y_{:d_2k} \end{bmatrix}_{d_1d_2*1} = \begin{bmatrix} Vec(M_{d_11k}) \\ Vec(M_{d_12k}) \\ \vdots \\ Vec(M_{d_1d_2k}) \end{bmatrix}_{d_1d_2*r_1r_2r_3} \times Vec(G)_{r_1r_2r_3*1} = Vec(M_{d_1d_2k}) \times Vec(G)$$

Final step:

$$Y_{:::} = \begin{bmatrix} y_{::1} \\ y_{::2} \\ \vdots \\ y_{::d_3} \end{bmatrix}_{d_1d_2d_3*1} = \begin{bmatrix} Vec(M_{d_1d_21}) \\ Vec(M_{d_1d_22}) \\ \vdots \\ Vec(M_{d_1d_2d_3}) \end{bmatrix}_{d_1d_2d_3*r_1r_2r_3} \times Vec(G)_{r_1r_2r_3*1} = M^{long}_{d_1d_2d_3*r_1r_2r_3} \times Vec(G)_{r_1r_2r_3*1}$$

Where $Y_{:::}$ is the Vec(Y) vectorization of Y we defined brfore.
,