tut10

Monday, July 26, 2021 5:46 PM

Conceptual

1. A ________ is a matrix which contains very few non-zero elements.
   (a) confusion matrix
   (b) sparse matrix ✓
   (c) utility matrix
   (d) similarity matrix
2. What is a content-based recommender system?
   (a) Content-based recommender systems tries to recommend items to users based on their profile built upon preferences and taste. ✓
   (b) Content-based recommender systems tries to recommend items based on similarity among items.
   (c) Content-based recommender systems tries to recommend items based on the similarity of users when buying, watching, or enjoying something.
   (d) None of the above.
3. Which of the following is true of collaborative filtering systems?
   (a) Suppose you are writing a recommender system to predict a user's book preferences. In order to build such a system, you need that user to rate all the other books in your training set.
   (b) To use collaborative filtering, you need to manually design a feature vector for every item (e.g. movie) in your dataset, that describes that item's most important properties.
   (c) If you have a dataset of users ratings' on some products, you can use these to predict one user's preferences on products he has not rated. ✓
4. What is the meaning of "cold start" in collaborative filtering?
   (a) The difficulty in recommendation when we do not have enough ratings in the user-item dataset.
   (b) The difficulty in recommendation when we have new user and we cannot create a profile for him/her, or when we have a new item which has not received any rating yet. ✓
   (c) The difficulty in recommendation when the number of users or items increases and the amount of data expands, so algorithms will begin to suffer drops in performance.
5. Three computers, A, B and C, have the numerical features listed below:

Feature	A	B	C
Processor speed	3.06	2.68	2.92
Disk size	500	320	640
Main-memory size	6	4	6

We may imagine these values as defining a vector for each computer; for instance, A's vector is $[3.06, 500, 6]$.

A certain user has rated the three computers as follows: A: 4 stars, B: 2 stars, C: 5 stars.

(a) Normalise the ratings for this user. That is, compute the average rating and subtract it from individual ratings.

(b) Compute a user profile for the user, with components for processor speed, disk size, and main memory size.

Ans:

(a) The average rating is $(4+2+5)/3 = 11/3$. Therefore, the mean-centred ratings for each item are:
$A: 4 - 11/3 = 1/3$
$B: 2 - 11/3 = -5/3$
$C: 5 - 11/3 = 4/3$

(b) The user profile of a component is computed as the weighted average of the component's values, with weights given by the user's ratings.
Processor speed: $3.06 \times 1/3 - 2.68 \times 5/3 + 2.92 \times 4/3 = 0.4467$
Disk size: $500 \times 1/3 - 320 \times 5/3 + 640 \times 4/3 = 486.6667$
Main memory size: $6 \times 1/3 - 4 \times 5/3 + 6 \times 4/3 = 3.3333$

6. (optional) In this question, we will generalise Example 4 and develop the general formula for optimising optimising an arbitrary element in the matrix U or V. Recall that a UV-decomposition factorise an n-by-m matrix M into two matrices U and V of dimensions n-by-d and d-by-m, for some d:

$$\underbrace{\begin{bmatrix} u_{11} & \cdots & u_{1d} \\ \vdots & & \vdots \\ u_{n1} & \cdots & u_{nd} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{d1} & \cdots & v_{dm} \end{bmatrix}}_{V} = \underbrace{\begin{bmatrix} m_{11} & \cdots & m_{1m} \\ \vdots & & \vdots \\ m_{n1} & \cdots & m_{nm} \end{bmatrix}}_{M}.$$

We shall use m_{ij}, u_{ij} and v_{ij} for entries in row i and column j of M, U and V.

Suppose we want to vary u_{rs}. Show that the value of this element that minimises the sum of squared error between M and UV is given by:

$$u_{rs} = \frac{\sum_j v_{sj}(m_{rj} - \sum_{k\neq s} u_{rk}v_{kj})}{\sum_j v_{sj}^2}.$$

Ans:

6. Let P denote the product of U and V, i.e. $P = UV$. Note that u_{rs} affects only the elements in row r of P. By the definition of matrix multiplication, the j^{th} element in row r of P equals to:

$$p_{rj} = \sum_{k=1}^{d} u_{rk}v_{kj} = \sum_{k\neq s}^{d} u_{rk}v_{kj} + u_{rs}v_{sj}.$$

Therefore, the sum of squared error (SSE) that is affected by $x = u_{rs}$ is given by:

$$\sum_j (m_{rj} - p_{rj})^2 = \sum_j (m_{rj} - \sum_{k\neq s}^{d} u_{rk}v_{kj} - xv_{sj})^2.$$

To find the value of x that minimises the SSE, we take the first derivative of the above expression and set it to 0:

$$\sum_j -2v_{sj}(m_{rj} - \sum_{k\neq s}^{d} u_{rk}v_{kj} - xv_{sj})$$

$$x = \frac{\sum_j v_{sj}(m_{rj} - \sum_{k\neq s} u_{rk}v_{kj})}{\sum_j v_{sj}^2}$$

In a similar way, we can derive that the value of v_{rs} that minimises the SSE is given by:

$$v_{rs} = \frac{\sum_i u_{ir}(m_{is} - \sum_{k\neq r} u_{ik}v_{ks})}{\sum_i u_{ir}^2}.$$

We could use these two formulae to update the elements of U and V.