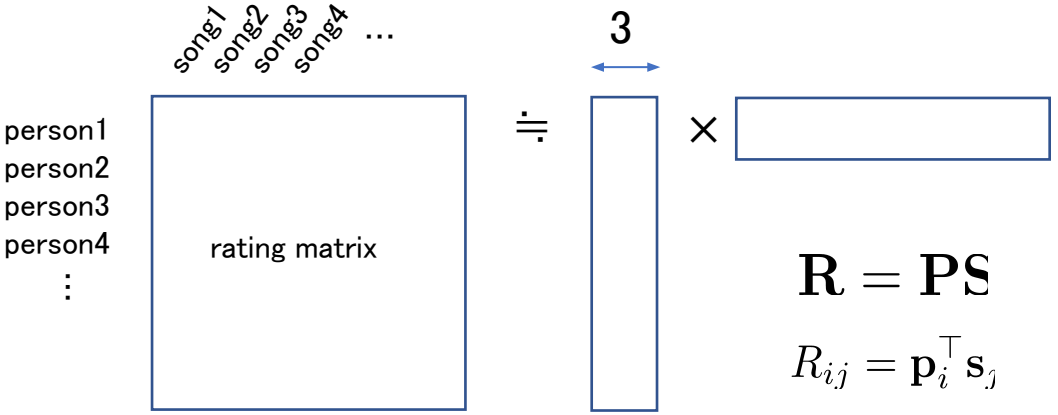
Exercise 10.1

- We wish to predict how a person rates songs

- Some people have similar tastes about like/dislike of music
  - That said, there will be no two persons having exactly the same taste
  - This kind of problems is known as *collaborative filtering*
- We approximate the rating matrix by a matrix of rank=3



$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \vdots \\ \mathbf{p}_{15}^\top \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_{20} \end{bmatrix}$$

$$\mathbf{R} = \mathbf{PS}$$

$$R_{ij} = \mathbf{p}_i^\top \mathbf{s}_j$$

Exercise 10.1

- Ratings of 20 songs are available (rating1.txt by 5 persons, rating2.txt by 15 persons)
  - Download `rating1.txt` from the course page and read into `R` by

```
>> load('rating1.txt')
```

  - Rating is represented by an integer in the range of [1,5]
  - R(2,4)=3 means "person2 gave rating=3 for song4"

- Suppose a new (i.e., 16th) person gives ratings for three songs
  - song1=4, song3=2, song7=3, i.e., $R_{16,1} = 4,\ R_{16,3} = 2,\ R_{16,7} = 3$

- Estimate ratings by this person for other songs
  - The following steps should be performed for each rating date (rating1.txt and rating2.txt)
  - First, **find a rank-3 approximation of R**, i.e., obtain 5x3 P and 3x20 S
  - Second**, find $\mathbf{p}_{16}$ that satisfies the following equations using S**:

$$R_{16,1} = \mathbf{p}_{16}^{\top}\mathbf{s}_1$$
$$R_{16,3} = \mathbf{p}_{16}^{\top}\mathbf{s}_3$$
$$R_{16,7} = \mathbf{p}_{16}^{\top}\mathbf{s}_7$$

  - Finally, **calculate prediction of ratings** by $R_{16,j} = \mathbf{p}_{16}^{\top}\mathbf{s}_j$
  - True ratings of R_{16} are:

4 3 2 2 3 3 3 2 3 1 2 3 2 2 3 4 3 3 3 3