Decomposing the rating matrix

The user-item rating matrix R of rank n is approximated by a matrix $\hat{R} = PQ^T$ of rank $k < n$, where P is a $|U| \times k$ matrix of users factors and Q a $|I| \times k$ matrix of item factors. Intuitively, the u^{th} row of P, $p_u \in R_k$, represents the coordinates of user u projected in the k-dimensional latent space. Likewise, the i^{th} row of Q, $q_i \in R_k$, can be seen as the coordinates of item i in this latent space. Matrices P and Q are normally found by minimizing the reconstruction error defined with the squared Frobenius norm:

$$\begin{aligned} \operatorname{err}(P,\, Q) &= \left\| R - PQ^T \right\|_F^2 \\ &= \sum_{u,\, i} \left(r_{ui} - \boldsymbol{p}_u \boldsymbol{q}_i^T \right)^2 \end{aligned}$$

Minimizing this error is equivalent to finding the *Singular Value Decomposition* (SVD) of R

$$R = U \Sigma V^T$$

Denote by Σ_k, U_k and V_k the matrices obtained by selecting the subset containing the k highest singular values and their corresponding singular vectors, the user and item factor matrices correspond to $P = U_k \Sigma_k^{1/2}$ and $Q = V_k \Sigma_k^{1/2}$.

Although it is possible to assign a default value to rui, as mentioned above, this would introduce a bias in the data. More importantly, this would make the large matrix R dense and, consequently, render impractical the SVD decomposition of R. The common solution to this problem is to learn P and Q using only the known ratings

$$\operatorname{err}(P,\, Q) = \sum_{u,\, i} \left(r_{ui} - \boldsymbol{p}_u \boldsymbol{q}_i^T \right)^2 + \lambda \left(\|\boldsymbol{p_u}\|^2 + \|\boldsymbol{q_i}\|^2 \right)$$

where λ is a parameter that controls the level of regularization