This section of the chapter talks about probabilistic matrix factorization for collaborative filtering. Collaborative filtering, also known as CF, requires predicting entries in a matrix, such as the rating that a user gives to a movie. In this way, we can see that CF is a kind of relational learning problem, and one with a particular level of commercial importance.

In fact, much of the work in this area makes use of the data that Netflix made available in their competition. The technique used here is called probabilistic matrix factorization, also known as PMF. The intuition behind this method is that each use and each movie get embedded into the same low-dimensional continuous space. If a user is close to a movie in space, they are likely to rate it highly.

PMF is closely related to the SVD. In particular, if there is no missing data, then computing the MLD for us and vs is equivalent to finding a rank K approximation to R. However, as soon as there is missing data, the problem becomes non-convex, and standard SVD methods cannot be applied. However, just maximizing the likelihood results in overfitting, so we have to regularize this by imposing Gaussian priors.