Abstract

The increased availability of the multi-view data (data on the same samples from multiple sources) has led to strong interest in models based on low-rank matrix factorizations. These models represent each data view via shared and individual components, and have been successfully applied for exploratory dimension reduction, association analysis between the views, and further learning tasks such as consensus clustering. Despite these advances, the existing methods have two limitations: they focus on continuous data and do not allow for partial sharing of components between the views. In this talk, I will focus on addressing the first limitation by using exponential family distributions to model heterogeneous (continuous/binary/count) data types. Time permitting, I will also discuss a new structural decomposition of continuous multi-view data that allows for partially-shared components. Based on joint work with Gen Li.