Random perturbation of low rank matrices: Improving classical bounds

Sean O'Rourke, Yale University

Computing the singular values and singular vectors of a large matrix is a basic task in high dimensional data analysis with many applications in computer science and statistics. In practice, however, data is often perturbed by noise. A natural question is the following. How much does a small perturbation to the matrix change the singular values and vectors?

Classical (deterministic) theorems, such as those by Davis-Kahan, Wedin, and Weyl, give tight estimates for the worst-case scenario. In this talk, I will consider the case when the perturbation is random. In this setting, better estimates can be achieved when our matrix has low rank. As an application, I will discuss a number of matrix reconstruction problems including hidden clique, hidden coloring, clustering, and Netflix-type problems.

This talk is based on joint work with Van Vu and Ke Wang.