Matrix Rank, Tensor Rank, and the Algebraic Statistics of Molecular Evolution

John Rhodes
Bates College

Friday, February 11, 2005
L01 Carson Hall, 11:00am
(Tea 10:30 am Math Lounge)

Abstract

The notion of the *matrix rank* is fundamental in linear algebra and can be easily computed either algorithmically (via Gaussian elimination or SVD) or algebraically (via the vanishing of minors). In addition, an elementary observation shows that rank plays a role in describing what types of 2-dimensional tables might arise from some simple statistical models.

When these statistical models are generalized only slightly, the tables become 3- or more dimensional, and we must deal with the *tensor rank* of higher dimensional arrays. While the concept is an old one, determining tensor rank poses much more difficulty than matrix rank, whether approached algorithmically or algebraically. Many basic questions are still unanswered.

After developing the rank concept as it relates to statistical models, this talk will give recent results that indicate the key role it plays in molecular phylogenetics: Standard statistical models of the evolution of DNA sequences along trees predict high-dimensional data tables that can be characterized in terms of ranks of 'flattened' tables of dimension ≤ 3.

In addition, the viewpoint of statistical models, and specifically phylogenetic ones, has led to new insights in algebraic geometry: We provide construction of new, explicit polynomials that vanish on all tensors of certain ranks and size—which are thus higher-dimensional analogues of matrix minors.

This talk should be accessible to graduate students.