Title:

Neighborliness of Randomly-Projected Simplices in High Dimensions

Author(s):

David L. Donoho and Jared Tanner

Technical Report number (Dept. of Statistics, Stanford Univ.):

2005-7

Date:

April 2005

Abstract:

Abstract

Let A be a d by n matrix, $d < n$. Let $T = T^{n-1}$ be the standard regular simplex in $\mathbf{R}^n$. We count the faces of the projected simplex AT in the case where the projection is random, the dimension d is large and n and d are comparable: $d \sim \delta n$, $\delta \in (0, 1)$. The projector A is chosen uniformly at random from the Grassmann manifold of d-dimensional orthoprojectors of $\mathbf{R}^n$. We derive $\rho_N(\delta) > 0$ with the property that, for any $\rho < \rho_N(\delta)$, with overwhelming probability for large d, the number of k-dimensional faces of $P = AT$ is exactly the same as for T, for $0 \leq k \leq \rho d$. This implies that P is $\lfloor \rho d \rfloor$-neighborly, and its skeleton $Skel_{\lfloor \rho d \rfloor}(P)$ is combinatorially equivalent to $Skel_{\lfloor \rho d \rfloor}(T)$. We display graphs of ρ_N.

We also study a weaker notion of neighborliness it asks if the k-faces are all simplicial and if the numbers of k-dimensional faces $f_k(P) \geq f_k(T)(1 - \epsilon)$. This was already considered by Vershik and Sporyshev, who obtained qualitative results about the existence of a threshold $\rho_{VS}(\delta) > 0$ at which phase transition occurs in k/d. We compute and display ρ_{VS} and compare to ρ_N.

Our results imply that the convex hull of n Gaussian samples in R^d, with n large and proportional to d, 'looks like a simplex' in the following sense. In a typical realization of such a high-dimensional Gaussian point cloud $d \sim \delta n$, all points are on the boundary of the convex hull, and all pairwise line segments, triangles, quadrangles, ..., $\lfloor \rho d \rfloor$-angles are on the boundary, for $\rho < \rho_N(d/n)$.

Our results also quantify a precise phase transition in the ability of linear programming to find the sparsest nonnegative solution to typical systems of underdetermined linear equations; when there is a solution with fewer than $\rho_{VS}(d/n)d$ nonzeros, linear programming will find that solution.