THE PROBABILITY THAT A SUBSPACE CONTAINS A POSITIVE VECTOR

KENT E. MORRISON

ABSTRACT. We determine the probability that a random k-dimensional subspace of $\mathbf{R}^n$ contains a positive vector.

For positive integers k and n with $k \leq n$, let $p(n,k)$ denote the probability that a random k-dimensional subspace of $\mathbf{R}^n$ contains a positive vector. The aim of this article is to prove

$$p(n,k) = \frac{1}{2^{n-1}} \sum_{j=0}^{k-1} \binom{n-1}{j}. \tag{1}$$

First we make the definitions precise. A vector $t \in \mathbf{R}^n$ is *positive* if $t_i \geq 0$ for all i and $t_i > 0$ for at least one i, and a *random subspace* is a point in the Grassmann manifold $G(n,k)$ with its natural $\mathrm{O}(n)$-invariant probability measure. This measure is constructed by starting with Haar measure on the orthogonal group $\mathrm{O}(n)$, which is bi-invariant and has total mass 1, and then pushing Haar measure down to $G(n,k)$ using the natural projection $\mathrm{O}(n) \twoheadrightarrow G(n,k)$. We also call this Haar measure.

To prove (1) we use a result of J. G. Wendel [2] showing that $p(n,d)$ is the probability that n random points in $\mathbf{R}^d$ lie in a half-space or, equivalently, that the convex hull of the points does not contain the origin. Let $d = n-k$ be the complementary dimension for our random subspaces. Given points $z_1, \ldots, z_n \in \mathbf{R}^d$ we define the linear map

$$\hat{z} \colon \mathbf{R}^n \to \mathbf{R}^d : (t_1, \ldots, t_n) \mapsto \sum t_i z_i.$$

Then the convex hull of the z_i contains the origin if and only if $\ker \hat{z}$ contains a positive vector. (The forward implication is immediate. For the converse, suppose $t \in \ker \hat{z}$ is a positive vector. Thus, $\sum t_i z_i = 0$. Then $\sum_i (t_i/T) z_i = 0$ is a convex combination of the z_i, where $T = \sum_i t_i$.)

Date: April 2, 2010.
2000 *Mathematics Subject Classification.* Primary 60D05.
Key words and phrases. Geometric probability, random subspace.

If the points z_i are random, then with only mild restrictions on their distribution, $\hat{z}$ has maximal rank, and so the kernel of $\hat{z}$ has dimension k. This holds, for example, if the z_i are iid with a distribution absolutely continuous with respect to Lebesgue measure. But if we further assume that the z_i are drawn from the probability distribution on $\mathbf{R}^d$ for which the components are iid standard normal variables, then $\ker \hat{z}$ is Haar distributed in $G(n, k)$.

To prove this we note that the distribution of $\hat{z}$ is $\mathrm{O}(n)$-invariant and that the kernel map from the subset of $d \times n$ matrices of maximal rank to the Grassmannian $G(n, k)$ is $\mathrm{O}(n)$-equivariant. In particular, for a $d \times n$ matrix A and an orthogonal matrix $g \in \mathrm{O}(n)$, we have $\ker(Ag^{-1}) = g(\ker A)$. It follows that the induced probability measure on $G(n, k)$ is $\mathrm{O}(n)$-invariant and must be Haar measure.

Then the probability that $\ker \hat{z}$ contains a positive vector is the probability that the origin is in the convex hull of the z_i, which is $1 - p(n, d)$. Finally, $1 - p(n, d) = p(n, k)$, which follows from the identity

$$2^{n-1} = \sum_{j=0}^{n-1} \binom{n-1}{j}.$$

This completes the proof of (1).

(The identity $p(n, k) + p(n, d) = 1$ says that almost surely for V in $G(n, k)$ exactly one of the subspaces V and $V^{\perp}$ contains a positive vector. This is a probabilistic version of the theorem stating that a subspace contains a positive vector if and only if its orthogonal complement does not contain a strictly positive vector, i.e., a vector all of whose components are positive [1].)

References

[1] S. Roman, *Advanced Linear Algebra*, 3rd ed., Springer, New York, 2008.

[2] J. G. Wendel, A problem in geometric probability, *Math. Scand.* 11 (1962) 109–111.

California Polytechnic State University, San Luis Obispo, CA 93407

Current address: American Institute of Mathematics, Palo Alto, CA 94306

E-mail address: kmorriso@calpoly.edu