On Powers of a Random Orthogonal Matrix

Morris L. Eaton & Robb J. Muirhead

University of Minnesota               Pfizer Global R&D

SEA’06 @ MIT
July 11, 2006
Outline

• Motivation
• Background
• Some new (we think ...) distributional results
Motivation

• Talk by Tom Marzetta (Bell Labs) last October at SEA ’05.
• Take a unit sphere (centered at the origin) in 3 dimensions, and transform the North Pole using a random (Haar) orthogonal matrix. Result: a point that is uniformly distributed on the sphere.
• Transform this point using the same orthogonal matrix. Result?
Orthogonal Transformations of the North Pole of a 3-dimensional Unit Sphere

Start at North Pole and apply random orthogonal transformation: uniformly random position

Apply the same transformation a second time. Bias towards the northern hemisphere.

T. L. Marzetta (Bell Labs): Presentation at SEA’05
Some Background Material

• Let $O(p)$ be the group of $p \times p$ orthogonal matrices $H$ \quad ($H'H = I_p$).

• Let $\mathcal{D} \subset O(p)$. There is a unique probability measure $\mu$ on $O(p)$ satisfying

$\mu(\Gamma \mathcal{D}) = \mu(\mathcal{D} \Gamma) = \mu(\mathcal{D})$, for all $\Gamma \in O(p)$.

• This is called the (invariant) Haar distribution on $O(p)$; also the uniform distribution on $O(p)$. 
Background (cont.)

To generate a uniformly distributed $H$:

• Take a $p \times p$ matrix of iid $N(0, 1)$ variables.

• Now do **Gram-Schmidt** on the columns to get $p$ orthonormal vectors.

• The $p \times p$ matrix $H$ with these as columns then has the uniform distribution.

• **Key fact:** If $H$ is uniform on $O(p)$ then $H$ and $PHQ$ have the same distribution, for all $p \times p$ orthogonal matrices $P$ and $Q$. 
Distributional Results

Throughout this talk, we assume $p \geq 3$.

On the unit $p$-sphere $S_p = \{ x \mid x \in \mathbb{R}^p, x'x = 1 \}$, let

$$x_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{(the "North Pole")}. $$

Let $H$ be a uniform orthogonal $p \times p$ matrix.

After $k$ transformations by $H$, the North Pole is mapped to the (random) point $U_k = H^k x_0$ on $S_p$. 
Distributional Results (cont.)

• We are interested in the first (i.e., the “north-pointing”) coordinate of $U_k$. This is the random variable

\[ V_k = x'_0 U_k = x'_0 H^k x_0. \]

• What can we say about the distribution of $V_k$? (Tom Marzetta’s simulations suggest that when $p=3$, $P(V_2 > 0) > \frac{1}{2}$.)
Distributional Results (cont.)

• Note that for each vector $x \in S_p$, $x'H^kx$ has the same distribution as $V_k = x_0'H^kx_0$.

Proof: For each $\Gamma \in O(p)$,

$$x'H^kx = (\Gamma x)'(\Gamma H \Gamma')(\Gamma H \Gamma')(\Gamma H \Gamma')(\Gamma x)$$

Since $\Gamma H \Gamma'$ has the same distribution as $H$, it follows that $x'H^kx$ has the same distribution as $(\Gamma x)'H^k(\Gamma x)$. Now pick $\Gamma$ so that $\Gamma x = x_0$. 
Distributional Results (cont.)

- Objective of this talk: Describe the distributions of $V_1$, $V_2$, and $V_3$.
- Partition $H$ as

\[
H = \begin{pmatrix}
  h_{11} & h_{12} \\
  h_{21} & H_{22}
\end{pmatrix},
\]

where

\[
\begin{align*}
h_{11} & \in (-1, 1), \\
h_{12} & : 1 \times (p - 1), \\
h_{21} & : (p - 1) \times 1, \\
H_{22} & : (p - 1) \times (p - 1).
\end{align*}
\]
Distribution of $V_1$

Now, $V_1 = x'Hx_0 = h_{11}$, and it is well-known that $h_{11}^2$ has a Beta distribution with parameters $\alpha = \frac{1}{2}$ and $\beta = \frac{p-1}{2}$. Because $h_{11}$ and $-h_{11}$ have the same distribution, it follows that the density function of $h_{11}$ is

$$f(t \mid p) = \frac{\Gamma\left(\frac{p}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p-1}{2}\right)}(1-t^2)^{(p-3)/2}, \quad -1 < t < 1.$$ 

Note: When $p=3$, $h_{11}$ is uniform on $(-1, 1)$, a result that dates back (at least) to Poincaré (1854-1912).
Distribution of $V_2$

We have

$$V_2 = x_0' H^2 x_0 = h_{11}^2 + h_{12} h_{21}.$$  

Write this as

$$V_2 = h_{11}^2 + (1 - h_{11}^2) \frac{h_{12}}{(1 - h_{11}^2)^{1/2}} \frac{h_{21}}{(1 - h_{11}^2)^{1/2}}.$$
Distribution of $V_2$ (cont.)

$$V_2 = h_{11}^2 + (1 - h_{11}^2) \frac{h_{12}}{(1 - h_{11}^2)^{1/2}} \frac{h_{21}}{(1 - h_{11}^2)^{1/2}}.$$ 

Now, let $\Gamma$ and $\Delta$ be fixed (for the moment) $p \times p$ orthogonal matrices of the form

$$\Gamma = \begin{pmatrix} 1 & 0 \\ 0 & \Gamma_1 \end{pmatrix}, \quad \Delta = \begin{pmatrix} 1 & 0 \\ 0 & \Delta_1 \end{pmatrix}.$$ 

where $\Gamma_1$ and $\Delta_1$ are $(p-1) \times (p-1)$ orthogonal matrices.
Then, because $H$ and $\Gamma H \Delta$ have the same distribution, and

$$
\Gamma H \Delta = \begin{pmatrix} 1 & 0 \\ 0 & \Gamma_1 \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & H_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \Delta_1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \Delta_1 \\ \Gamma_1 h_{21} & \Gamma_1 H_{22} \Delta_1 \end{pmatrix},
$$

we have

$$
V_2^d = x'_0 \Gamma H \Delta \Gamma H \Delta x_0 = h_{11}^2 + (1 - h_{11}^2) \frac{h_{12}}{(1 - h_{11}^2)^{1/2}} \Delta_1 \Gamma_1 \frac{h_{21}}{(1 - h_{11}^2)^{1/2}}.
$$
Distribution of $V_2$ (cont.)

\[ V_2 = h_{11}^2 + (1-h_{11}^2) \frac{h_{12}}{(1-h_{11}^2)^{1/2}} \Delta_1 \Gamma_1 \frac{h_{21}}{(1-h_{11}^2)^{1/2}} \]

Now, note that this holds for all fixed $\Delta_1$ and $\Gamma_1$, so it also hold for any random $\Delta_1$ and $\Gamma_1$ independent of $H$ (because $H$ and $\Gamma H A$ have the same distribution). So, pick $\Delta_1$ and $\Gamma_1$ to be independent uniform on $O(p-1)$, so that $\Delta_1 \Gamma_1$ is uniform on $O(p-1)$ (independent of $H$).
Distribution of $V_2$ (cont.)

\[ V_2 \overset{d}{=} h_{11}^2 + (1 - h_{11}^2) \frac{h_{12}}{(1 - h_{11}^2)^{1/2}} \Delta_1 \Gamma_1 \frac{h_{21}}{(1 - h_{11}^2)^{1/2}} \]

The next step involves conditioning on $H$, so that

\[ u = \frac{h_{12}}{(1 - h_{11}^2)^{1/2}} \quad \text{and} \quad v = \frac{h_{21}}{(1 - h_{11}^2)^{1/2}} \]

are fixed (as is $h_{11}$), with $u, v \in S_{p-1}$.
Distribution of $V_2$ (cont.)

We now use the following result:

**Lemma:** If $u$ and $v$ are $r \times 1$ fixed vectors of length 1, and $Q$ is uniform on $O(r)$, then $u'Qv$ has the density function $f(\cdot|\mathbf{r})$.

**Proof:** Rotate both $u$ and $v$ into the first coordinate vector, and use the invariance property of the uniform distribution on $O(r)$. Then $u'Qv$ has the same distribution as the $(1, 1)$ element of $Q$, which is $f(\cdot|\mathbf{r})$. 
It then follows that, conditional on $H$, 

$$ \frac{h_{12}}{(1-h_{11}^2)^{1/2}} \Delta_1 \Gamma_1 \frac{h_{21}}{(1-h_{11}^2)^{1/2}} $$

has the density $f(\cdot \mid p-1)$. Since this does not depend on $H$, it is also the unconditional distribution. Thus we have:
Theorem 1

\[ V_2^d = T + (1 - T)Y \]

where \( T \) and \( Y \) are independent, \( T \sim \text{Beta}\left(\frac{1}{2}, \frac{p-1}{2}\right) \) and \( Y \) has density function

\[ f_Y(y) = f(y \mid p-1) = \frac{\Gamma\left(\frac{p-1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{p}{2} - 1\right)}(1 - y^2)^{(p-4)/2}, \quad -1 < y < 1. \]
Distribution of $V_2$ (cont.)

From this, it follows easily that

$$P(V_2 > 0) > \frac{1}{2}, \text{ and } P(V_2 > 0) \rightarrow \frac{1}{2} \text{ as } p \rightarrow \infty.$$

\[
\begin{array}{|c|c|c|c|c|c|c|c|c|}
\hline
p & 3 & 4 & 5 & 10 & 20 & 50 & 100 & 500 \\
\hline
P(V_2 > 0) & .71 & .68 & .66 & .62 & .59 & .56 & .54 & .52 \\
\hline
\end{array}
\]
Distribution of $V_2$ (cont.)

- The distribution of $V_2$ just given is also a consequence of the following result, whose proof requires (we think) a substantial amount of group invariance theory.

**Theorem 2:** If

$$H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & H_{22} \end{pmatrix}$$

is uniform, then

$$h_{11}, \quad W_1 = \frac{h_{21}}{\left(1 - h_{11}^2\right)^{1/2}}, \quad W_2 = \frac{h'_{12}}{\left(1 - h_{11}^2\right)^{1/2}}$$

are independent, with $W_1$ and $W_2$ being uniformly distributed on $S_{p-1}$ and $h_{11}$ having the density $f(\cdot | p)$. 21
Distribution of $V_2$ (cont.)

- In terms of $h_{11}, W_1, W_2$,

$$V_2 = h_{11}^2 + (1 - h_{11}^2) W_2' W_1.$$  

Lemma: $Y \equiv W_2' W_1$ has density $f(\cdot | p - 1)$.

Proof: The conditional distribution of $Y$ given $W_2$ has density $f(\cdot | p - 1)$ by the orthogonal invariance of the distribution of $W_1$. 


Distribution of $V_3$

- The description of the distribution of $V_1$ involves a single random variable (namely $h_{11}$) with density $f(\cdot | p)$.
- Our description of the distribution of $V_2$ involves two independent random variables: $h_{11}$ with density $f(\cdot | p)$ and $Y$ with density $f(\cdot | p-1)$.
- Claim: The distribution of $V_3$ may be described using three independent random variables: $h_{11}$ with density $f(\cdot | p)$, $Y$ with density $f(\cdot | p-1)$, and $Z$ with density $f(\cdot | p-2)$.

$$V_3 = x_0' H^3 x_0 = h_{11}^3 + 2 h_{11} h_{12} h_{21} + h_{12} H_{22} h_{21}$$
**Theorem 3**

$V_3$ has the same distribution as

$$h_{11}^3 + 2h_{11}(1 - h_{11}^2)Y + (1 - h_{11}^2)[-h_{11}Y^2 + (1 - Y^2)Z],$$

where $h_{11}$, $Y$ and $Z$ are independent with respective density functions $f(.|p)$, $f(.|p-1)$, and $f(.|p-2)$.

Note: When $p=3$, $f(.|1)$ is not a density. In this case, its interpretation is as a discrete distribution, taking the values 1 and –1 with probability $\frac{1}{2}$ each.
The distribution of $V_3$ is symmetric; i.e., $V_3$ and $-V_3$ have the same distribution. This is true of $V_k$ for any odd integer $k$:

$$V_k = x_0' H^k x_0$$

$$-V_k = x_0' (-H)^k x_0 = x_0' H^k x_0 = V_k.$$
Open Problems & Conjectures

1. What are the distributions of $V_k$, for $k \geq 4$?

2. Conjecture: $V_2$ is stochastically bigger than $V_1$: i.e.,

$$P(V_2 > t) \geq P(V_1 > t), \text{ for all } t \in (-1, 1).$$

(We know this holds for $t = 0$.)

3. Conjecture: $V_3$ is has “more mass” near +1 and −1 than $V_1$. One possibility is that $V_3^2$ is stochastically bigger than $V_1^2$. 
