Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 5, Issue 1, Article 16, 2004

LOWER BOUNDS ON PRODUCTS OF CORRELATION COEFFICIENTS

FRANK HANSEN

INSTITUTE OF ECONOMICS,
UNIVERSITY OF COPENHAGEN,
STUDIESTRAEDE 6, DK-1455 COPENHAGEN K, DENMARK.
Frank.Hansen@econ.ku.dk

Received 23 June, 2003; accepted 17 February, 2004
Communicated by N.S. Barnett

ABSTRACT. We consider square integrable stochastic variables $X_1, \dots, X_n$ without imposing any further conditions on their distributions. If $r_{i,j}$ denotes the correlation coefficient between X_i and X_j then the product $r_{1,2}r_{2,3}\cdots r_{(n-1),n}r_{n,1}$ is bounded from below by $-\cos^n(\pi/n)$. The configuration of stochastic variables attaining the minimum value is essentially unique.

Key words and phrases: Correlation coefficient, Bessis-Moussa-Villani conjecture, Robust portfolio.

2000 *Mathematics Subject Classification.* 46C05, 26D15.

The main result in this note is the inequality

$$-\cos^n\left(\frac{\pi}{n}\right) \le (x_1 \mid x_2)(x_2 \mid x_3)\cdots(x_{n-1} \mid x_n)(x_n \mid x_1) \tag{1}$$

valid for arbitrary unit vectors $x_1, \dots, x_n$ in a real Hilbert space. The inequality is of intrinsic interest as it provides more information than can be gleaned by simply using the Cauchy-Schwartz' inequality. The inequality grew out of a study of the Bessis-Moussa-Villani conjecture [1, 7, 8], which states that the function $t \to \operatorname{Tr}\exp(A - tB)$ is the Laplace transform of a positive measure, when A and B are self-adjoint, positive semi-definite matrices. The conjecture can be reformulated to provide conditions of sign for the derivatives of arbitrary order of the function where these derivatives can be written as sums of particular functions with coefficients as given by the right hand side of (1). Subsequently it has appeared that the inequality (1) and in particular the optimal configuration of the vectors given rise to the equality, is related to the notion of robust portfolio in finance theory. Finally the inequality gives not always obvious constraints for correlation coefficients of random variables, especially in the important case $n = 3$.

Lemma 1. *Let x and z be unit vectors in a real Hilbert space H and consider the function*

$$f(y) = (x \mid y)(y \mid z) \qquad y \in H.$$

ISSN (electronic): 1443-5756

The author wishes to thank the referees for carefully reading the manuscript and for having pointed out a now corrected calculation error in the proof of Proposition 3.

087-03

The supremum of f on the unit sphere H_1 in H is given by

$$\sup_{y\in H_1} f(y) = \frac{1+(x \mid z)}{2}\,.$$

If $x = z$ the supremum is attained only in $y = \pm\, x$. If $x = -z$ the supremum is attained in any unit vector y orthogonal to x. In all other cases the supremum is attained only in $\pm\, y_0$, where $y_0 \in U = \operatorname{span}\{x, z\}$ is the unit vector such that the angle between x and y_0 equals the angle between y_0 and z, thus $(x \mid y_0) = (y_0 \mid z)$.

Proof. Apart from the trivial cases, $\dim U = 2$ and we may choose an orthonormal basis (e_1, e_2) for U such that, with respect to this basis, $x = (1, 0)$ and $z = (\cos\beta, \sin\beta)$ for some $\beta \in]0, \pi[$. We set $y_0 = (\cos(\beta/2), \sin(\beta/2))$ and calculate

$$f(y_0) = \cos^2\left(\frac{\beta}{2}\right) = \frac{1+\cos\beta}{2} = \frac{1+(x \mid z)}{2}.$$

Let y be an arbitrary unit vector in U and write it on the form $y = (\cos\alpha, \sin\alpha)$ for some $\alpha \in [0, 2\pi[$. The difference

$$\begin{aligned}
f(y_0) - f(y) &= \frac{1+\cos\beta}{2} - \cos\alpha\,(\cos\alpha\cos\beta + \sin\alpha\sin\beta)\\
&= \frac{1+\cos\beta}{2} - \frac{1+\cos 2\alpha}{2}\cos\beta - \frac{1}{2}\sin 2\alpha\sin\beta\\
&= \frac{1}{2}\,(1 - \cos 2\alpha\cos\beta - \sin 2\alpha\sin\beta)\\
&= \frac{1}{2}\,(1 - \cos(2\alpha - \beta)) \ge 0
\end{aligned}$$

with equality only for $\alpha = \beta/2$ or $\alpha = \beta/2 + \pi$. Finally, we must show $f(y_0) > f(y)$ for arbitrary unit vectors $y \notin U$. But since $f(y_0) > 0$, we only need to consider unit vectors $y \notin U$ such that $f(y) > 0$. Let y_1 denote the orthogonal projection on U of such a vector, then $0 < \|y_1\| < 1$ and

$$0 < f(y) = f(y_1) < \frac{f(y_1)}{\|y_1\|^2} = f\left(\frac{y_1}{\|y_1\|}\right) \le f(y_0),$$

where the last inequality follows since $\|y_1\|^{-1}y_1$ is a unit vector in U. $\square$

Lemma 2. *Let H be a real Hilbert space of dimension greater than or equal to two. Then there exists, for each $n \ge 2$, unit vectors $x_1, \ldots, x_n$ in H such that*

$$(x_1 \mid x_2)(x_2 \mid x_3)\cdots(x_{n-1} \mid x_n)(x_n \mid x_1) = -\cos^n\left(\frac{\pi}{n}\right). \tag{2}$$

Proof. Let U be a two-dimensional subspace of H and choose an orthonormal basis (e_1, e_2) for U. Relative to this basis we set

$$x_i = \left(\cos\left(\frac{(i-1)\pi}{n}\right), \sin\left(\frac{(i-1)\pi}{n}\right)\right) \qquad i = 1, \ldots, n.$$

The angle between consecutive vectors in the sequence $x_1, x_2, \ldots, x_n, -x_1$ is equal to π/n, therefore

$$(x_1 \mid x_2)(x_2 \mid x_3)\cdots(x_{n-1} \mid x_n)(x_n \mid -x_1) = \cos^n\left(\frac{\pi}{n}\right)$$

and the statement follows. $\square$

We notice that the solution in Lemma 2 above constitutes a *fan of vectors* dividing the radian interval $[0,\pi]$ into n slices, and that the angle π/n between consecutive vectors is acute for $n \geq 3$. The expression in (2) is indifferent to a change of sign of some of the vectors, but after such an inversion the angle between consecutive vectors is no longer acute, except in the case when all the vectors are inverted. But then we are back to the original construction for the vectors $-x_1, -x_2, \ldots, -x_n$.

Proposition 3. *The inequality*

$$\cos^{n-1}\left(\frac{\pi}{n-1}\right) < \cos^n\left(\frac{\pi}{n}\right)$$

is valid for $n = 2, 3, \ldots$. Furthermore, $\cos^n(\pi/n) \nearrow 1$ as n tends to infinity.

Proof. The inequality is trivial for $n = 2$. We introduce the function $f(t) = \cos^t(\pi/t)$ for $t > 2$. Since $\log f(t) = t \log\cos(\pi/t)$, we have

$$\frac{f'(t)}{f(t)} = \log\cos\left(\frac{\pi}{t}\right) - t\frac{\sin(\pi/t)}{\cos(\pi/t)}\frac{(-\pi)}{t^2}$$

or

$$f'(t) = (\cos\theta\cdot\log\cos\theta + \theta\sin\theta)\frac{f(t)}{\cos\theta} \quad \text{where} \quad 0 < \theta = \frac{\pi}{t} < \frac{\pi}{2}.$$

Setting $g(\theta) = \cos\theta\cdot\log\cos\theta + \theta\sin\theta$ for $0 < \theta < \pi/2$ we obtain

$$g'(\theta) = -\sin\theta\cdot\log\cos\theta + \theta\cos\theta > 0,$$

showing that g is strictly increasing, and since $g(\theta) \to 0$ for $\theta \to 0$ we obtain that both g and f' are strictly positive. This proves the inequality for $n \geq 3$. We then use the mean value theorem to write

$$\cos\left(\frac{\pi}{n}\right) - 1 = \frac{\pi}{n}(-1)\sin\left(\frac{\pi\theta}{n}\right) \geq -\frac{\pi^2}{n^2}$$

where $0 < \theta < 1$. To each $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that $\pi^2 n^{-1} < \varepsilon$ and consequently

$$\cos\left(\frac{\pi}{n}\right) \geq 1 - \frac{\pi^2}{n^2} \geq 1 - \frac{\varepsilon}{n}$$

for $n \geq n_0$. Hence

$$\lim_{n\to\infty}\cos^n\left(\frac{\pi}{n}\right) \geq \lim_{n\to\infty}\left(1 - \frac{\varepsilon}{n}\right)^n = \exp(-\varepsilon)$$

and since $\varepsilon > 0$ is arbitrary, the statement follows. □

Theorem 4. *Let $x_1, \ldots, x_n$ for $n \geq 2$ be unit vectors in a real Hilbert space H of dimension greater than or equal to two. Then*

$$-\cos^n\left(\frac{\pi}{n}\right) \leq (x_1 \mid x_2)(x_2 \mid x_3)\cdots(x_{n-1} \mid x_n)(x_n \mid x_1)$$

with equality only for the configuration in Lemma 2 together with configurations that are derived from this by multiplying some of the vectors by -1.

Proof. We prove the theorem by induction and notice that the statement is obvious for $n = 2$. We then consider, for $n \geq 3$, the function

$$f(y_1, \ldots, y_n) = (y_1 \mid y_2)(y_2 \mid y_3)(y_3 \mid y_4)\cdots(y_{n-1} \mid y_n)(y_n \mid -y_1)$$

for arbitrary vectors $y_1, \ldots, y_n$ in H_1. We equip H with the weak topology and notice that f is continuous and the unit ball compact in this topology, hence f attains its maximum on H_1 in some n-tuple $(x_1, \ldots, x_n)$ of unit vectors. It follows from Lemma 2 that

$$f(x_1, \ldots, x_n) = (x_1 \mid x_2)(x_2 \mid x_3)(x_3 \mid x_4)\cdots(x_{n-1} \mid x_n)(x_n \mid -x_1) > 0.$$

Each vector appears twice in the expression of $f(x_1,\dots,x_n)$, so the value of f is left unchanged by multiplication of one or more of the vectors by -1. Possibly by multiplying x_2 by -1 we may thus assume $(x_1 \mid x_2) > 0$. Possibly by multiplying x_3 by -1 we may next assume $(x_2 \mid x_3) > 0$ and so forth, until possibly by multiplying x_n by -1, we realize that we may assume $(x_{n-1} \mid x_n) > 0$. After these rearrangements which leave the value of f unchanged and since $f(x_1,\dots,x_n) > 0$, we finally realize that also $(x_n \mid -x_1) > 0$. The angle between any two consecutive vectors in the sequence $x_1, x_2, x_3, \dots, x_n, -x_1$ is thus acute. None of these angles can be zero, since if any two consecutive vectors are identical, say $x_2 = x_1$, then

$$f(x_1,\dots,x_n) = (x_2 \mid x_3)(x_3 \mid x_4)\cdots(x_{n-1} \mid x_n)(x_n \mid -x_2) = f(x_2,\dots,x_n).$$

By the induction hypothesis and Proposition 3 we thus have

$$f(x_1,\dots,x_n) \le \cos^{n-1}\left(\frac{\pi}{n-1}\right) < \cos^n\left(\frac{\pi}{n}\right)$$

which contradicts the optimality of $(x_1,\dots,x_n)$, cf. Lemma 2. We may therefore assume that each angle between consecutive vectors in the sequence $x_1, x_2, \dots, x_n, -x_1$ is acute but non-zero.

Since all the n factors in $f(x_1,\dots,x_n)$ are positive, we could potentially obtain a larger value of f by maximizing $(x_1 \mid x_2)(x_2 \mid x_3)$ as a function of $x_2 \in H_1$. However, since f already is optimal in the point $(x_1,\dots,x_n)$, we derive that also $(x_1 \mid x_2)(x_2 \mid x_3)$ is optimal as a function of x_2. According to Lemma 1, this implies that $x_2 \in U = \operatorname{span}\{x_1, x_3\}$ and that the angle between x_1 and x_2 equals the angle between x_2 and x_3. Potentially, $-x_2$ could also be a solution, but this case is excluded by the positivity of each inner product in the expression of $f(x_1,\dots,x_n)$. We may choose an orthonormal basis (e_1, e_2) for U such that $x_1 = e_1$ and the angle between x_1 and x_2 is positive, thus $x_2 = (\cos\theta, \sin\theta)$ and consequently $x_3 = (\cos 2\theta, \sin 2\theta)$ for some $\theta \in]0, \pi/2[$ with respect to this basis. We similarly obtain $x_4 \in U$ and that the angle, θ, between x_2 and x_3 is equal to the angle between x_3 and x_4, thus $x_4 = (\cos 3\theta, \sin 3\theta)$. We continue in this way until we obtain $x_n \in U$ with the representation $x_n = (\cos(n-1)\theta, \sin(n-1)\theta)$ and that the angle between x_n and $-x_1$ is θ. We conclude that $n\theta = \pi + k2\pi$ or $\theta = (2k+1)\pi/n$ for some $k = 0, 1, 2, \dots$. However, since θ is acute we obtain

$$0 < \cos\theta = \cos\left(\frac{(2k+1)\pi}{n}\right) \le \cos\left(\frac{\pi}{n}\right),$$

and this inequality contradicts the optimality of $(x_1,\dots,x_n)$ unless $k = 0$, thus $\theta = \pi/n$. We have derived that the vectors $(x_1,\dots,x_n)$ have the same configuration as in Lemma 2 and that $f(x_1,\dots,x_n) = \cos^n(\pi/n)$. □

If $X_1,\dots,X_n$ are non-constant square-integrable stochastic variables, then the correlation coefficient $r_{i,j}$ between X_i and X_j is defined by

$$r_{i,j} = \frac{\operatorname{Cov}(X_i, X_j)}{\|X_i\|_2 \cdot \|X_j\|_2} \qquad i, j = 1, \dots, n,$$

where $\|X\|_2 = \operatorname{Var}[X]^{1/2}$. Theorem 4 then states that

$$-\cos^n\left(\frac{\pi}{n}\right) \le r_{1,2} r_{2,3} \cdots r_{(n-1),n} r_{n,1}.$$

Notice that for the optimal configuration in Lemma 2, we can calculate all possible correlation coefficients, not only the coefficients between neighbours in the loop $X_1, X_2, \dots, X_n, X_1$.

For $n = 2$ the inequality reduces to $0 \le r_{1,2}^2$ with equality, when the stochastic variables are uncorrelated. The most striking case is probably $n = 3$ where $\cos^n(\pi/n) = 1/8$ and thus

$$-\frac{1}{8} \le r_{1,2} r_{2,3} r_{3,1}.$$

This is the only case where each correlation coefficient is represented exactly once in the product. For $n = 4$ we obtain

$$-\frac{1}{4} \le r_{1,2} r_{2,3} r_{3,4} r_{4,1}$$

and so forth.

REFERENCES

[1] D. BESSIS, P. MOUSSA AND M. VILLANI, Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics, *J. Math. Phys.*, **16** (1975), 2318–2325.

[2] T.E. COPELAND AND J.F. WESTON, *Financial Theory and Corporate Policy*, Addison-Wesley, Reading, Massachusetts, 1992.

[3] F.R. GANTMACHER, *Matrix Theory*, Volume 1, Chelsea, New York, 1959.

[4] D. GOLDFARB AND G. IYENGAR, Robust portfolio selection problems, *Mathematics of Operations Research*, **28** (2003), 1–38.

[5] F. HANSEN AND M.N. OLESEN, *Lineær Algebra*, Akademisk Forlag, Copenhagen, 1999.

[6] R. HORN AND C.R. JOHNSON, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[7] C.R. JOHNSON AND C.J. HILLAR, Eigenvalues of words in two positive definite letters, *SIAM J. Matrix Anal. Appl.*, **23** (2002), 916–928.

[8] P. MOUSSA, On the representation of $\mathrm{Tr}(e^{(A-\lambda B)})$ as a Laplace transform, *Reviews in Mathematical Physics*, **12** (2000), 621–655.