Annales Univ. Sci. Budapest., Sect. Comp. 13 (1992) 21-24

A PROOF OF THE GENERALIZED HADAMARD INEQUALITY VIA INFORMATION THEORY

E. Klafszky and P. Kas (Budapest, Hungary)

The generalized Hadamard inequality is considered in this paper. The new proof presented here is motivated by the idea of mutual information of random variables. For a strict algebraic proof see e.g. [3].

Let T be and $m \times n$ matrix of real numbers and denote its column vectors with $t_1, t_2, \ldots, t_n$. Let the set of column indices $J = \{1, 2, \ldots, n\}$ partitioned into sets $J_1, J_2, \ldots, J_\rho$, that is $J_i \cap J_j = \emptyset$ for $i \neq j$, $\bigcup\limits_{k=1}^{\rho} J_k = J$. Let $T_k = \{t_j : j \in J_k\}$. Assume that the vectors $t_j \in T_k$ are linearly independent for $k = 1, 2, \ldots, \rho$. Now the generalized Hadamard inequality can be stated as follows.

Theorem.

$$\det(T^*T) \leq \det(T_1^*T_1)\det(T_2^*T_2)\ldots\det(T_\rho^*T_\rho)$$

and equality holds iff $T_i^*T_j = 0$ for all $i \neq j$. (A^* stands for the transpose of the matrix A).

Note that if $\rho = n$ and $T_1 = \{t_1\}, T_2 = \{t_2\}, \ldots, T_n = \{t_n\}$ then the assumption says that $t_j \neq 0$ for all j and the theorem states the well-known version of the Hadamard inequality, that is

$$\det(T^*T) \leq t_1^2 t_2^2 \ldots t_n^2$$

and equality holds iff $t_i^* t_j = 0$ for all $i \neq j$.

Our proof goes through a couple of remarks.

Remark 1. Consider a real-valued function $f(x), x = (x_1, x_2, \ldots, x_n)$. Let the set of indices $J = \{1, 2, \ldots, n\}$ be partitioned into sets $J_1, \ldots, J_\rho$. After a proper rearranging of the components $x_1, x_2, \ldots, x_n$ we can split the vector x into subvectors $x_1, x_2, \ldots, x_\rho$ such that x_k contains the components $x_j, j \in J_k$. With this notation $f(x) = f(x_1, x_2, \ldots, x_\rho)$. Assume that

(i) $\quad f(x_1, x_2, \ldots, x_\rho) \geq 0 \quad (x_1, x_2, \ldots, x_\rho) \in I\!R^n$,

(ii) $\quad \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_\rho) dx_1 dx_2 \ldots dx_\rho = 1.$

Assumptions (i), (ii) obviously hold for a density function of an n-dimensional random variable. Introducing the functions

$$f_k(x_k) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_\rho) dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_\rho, \quad k = 1, 2, \dots, \rho$$

we have

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_k(x_k) dx_k = 1 \quad k = 1, 2, \dots, \rho$$

and

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_\rho) \log \frac{f(x_1, x_2, \dots, x_\rho)}{f_1(x_1) f_2(x_2) \dots f_\rho(x_\rho)} dx_1 dx_2 \dots dx_\rho \geq 0.$$

Here equality holds iff $f(x_1, x_2, \dots, x_\rho) = f_1(x_1) f_2(x_2) \dots f_\rho(x_\rho)$. The latter inequality plays an important role in our discussion and is well-known in information theory (Shannon lemma). It can be proved from the fact that

$$\log t \geq 1 - \frac{1}{t} \quad \text{for} \quad t \geq 0$$

and equality holds iff $t = 1$.

Remark 2. Consider a matrix T for which all assumptions incurred in the theorem are valid. Then the $n \times n$ matrix T^*T is positive definite, symmetric and can be partitioned as follows.

$$T^*T = \begin{array}{|c|c|c|c|} \hline T_1^*T_1 & T_1^*T_2 & \dots & T_1^*T_\rho \\ \hline T_2^*T_1 & T_2^*T_2 & \dots & T_2^*T_\rho \\ \hline \vdots & \vdots & & \vdots \\ \hline T_\rho^*T_1 & T_\rho^*T_2 & \dots & T_\rho^*T_\rho \\ \hline \end{array}$$

With this T^*T let us consider the function

$$g(x) = \frac{1}{\pi^{n/2}} \frac{1}{(\det(T^*T))^{1/2}} \exp\Big(- x^*(T^*T)^{-1} x \Big).$$

$g(x)$ is the density function of an n-dimensional random variable with Gauss distribution. Obviously

$$g(x) \geq 0 \qquad x \in I\!R^n,$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x)dx = 1.$$

Let us partition the components of x into subvectors $x_1, x_2, \ldots, x_\rho$ according to the partition $J_1, J_2, \ldots, J_\rho$ of the column indices $\{1, 2, \ldots, n\}$. One can show by evaluating the integral (see e.g.[1]) that

$$g_k(x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_\rho) dx_1 \ldots dx_{k-1} dx_{k+1} \ldots dx_\rho =$$

$$= \frac{1}{\pi^{|J_k|/2}} \frac{1}{(\det(T_k^* T_k))^{1/2}} \exp\Big(- x_k^* (T_k^* T_k)^{-1} x_k \Big)$$

where $\mid J_k \mid$ denotes the cardinality of the set J_k.

Remark 3. Evaluating the integrals one can easily check that

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_\rho) \log \frac{g(x_1, x_2, \ldots, x_\rho)}{g_1(x_1) g_2(x_2) \ldots g_\rho(x_\rho)} dx_1 dx_2 \ldots dx_\rho =$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \ldots, x_\rho) \log g(x_1, x_2, \ldots, x_\rho) dx_1 dx_2 \ldots dx_\rho -$$

$$- \sum_{k=1}^{\rho} \Big(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_k(x_k) \log g_k(x_k) dx_k \Big) =$$

$$= \frac{1}{2} \log \frac{\det(T_1^* T_1) \det(T_2^* T_2) \ldots \det(T_\rho^* T_\rho)}{\det(T^* T)}.$$

Now applying the inequality of Remark 1 to $g(x)$ states that the integral above is non-negative, and equals zero iff

$$\begin{array}{|c|c|c|c|}\hline T_1^*T_1 & T_1^*T_2 & \ldots & T_1^*T_\rho \\ \hline T_2^*T_1 & T_2^*T_2 & \ldots & T_2^*T_\rho \\ \hline \vdots & \vdots & & \vdots \\ \hline T_\rho^*T_1 & T_\rho^*T_2 & \ldots & T_\rho^*T_\rho \\ \hline \end{array} = \begin{array}{|c|c|c|c|}\hline T_1^*T_1 & 0 & \ldots & 0 \\ \hline 0 & T_2^*T_2 & \ldots & 0 \\ \hline \vdots & \vdots & & \vdots \\ \hline 0 & 0 & \ldots & T_\rho^*T_\rho \\ \hline \end{array}$$

Hence the theorem is proved.

Corollary. Let us consider a positive definite symmetric matrix C of size $n \times n$. Take a partition of indices $J = \{1, 2, \dots, n\}$ into the sets $J_1, J_2, \dots, J_\rho$ and denote the submatrix of elements c_{ij} $i \in J_k, j \in J_l$ with C_{kl}. Then

$$\det(C) \leq \det(C_{11})\det(C_{22}) \dots \det(C_{\rho\rho})$$

and equality holds iff $C_{kl} = 0$ for all $k \neq l$. To prove the corollary it should be noted that according to Cholesky's factorization theorem for every positive definite, symmetric matrix $C = T^*T$ with a suitable non-singular matrix T. Then the theorem directly applies.

References

[1] **Anderson T.W.,** *An Introduction to Multivariate Statistical Analysis,* John Wiley & Sons Inc., New York, 1957.

[2] **Csiszár I. and Körner J.,** *Information Theory,* Academic Press, 1981.

[3] **Gantmacher F.R.,** *The theory of matrices,* Chelsea Publishing Company, New York, N.Y., 1959.

[4] **Klafszky E. and Kas P.,** Megjegyzés egy többdimenziós Cauchy eloszlásról, *Alkalmazott Matematikai Lapok,* **13**(1987-88), 145-161.

(Received November 1, 1989)

E. Klafszky
Department for Building Construction and Management
Technical University, Budapest, Hungary

P. Kas
Operations Research Department
Eötvös Loránd University, Budapest, Hungary