APPLICATIONS OF GRAMIAN TRANSFORMATION FORMULA

Masatoshi Fujii*, Takayuki Furuta** and Ritsuo Nakamoto***

Received October 26, 1999

Dedicated to Professor Masahiro Nakamura on his 80th birthday with respect and affection

Abstract. We point out that the Gramian transformation formula gives us a natural and simple view to some known results, e.g. the translation-invariance of variance of operators, Hadamard theorem and inequalities on Gramian. Moreover we pick up a norm equality which is the essence of a norm inequality closely related to the Bernstein inequality.

1. Introduction. In [2], Björck and Thomee introduced a constant for a (bounded linear) operator T on a Hilbert space H, see also [4,8,11,13,16,17]:

$$\sup\{\|Tx\|^2 - |(Tx,x)|^2; \|x\| = 1\}. \tag{1.1}$$

We denote by M_T the square root of the constant for T. They proved that if T is a normal operator, then M_T concides with the smallest radius of disks containing the spectrum of T, cf.[11,12]. One of properties on M_T is the translation-invariance, i.e., $M_{T-\lambda} = M_T$ for all $\lambda \in \mathbb{C}$. More precisely, the variance of T at a state (i.e., unit vector) $x \in H$

$$\mathrm{Var}_x(T) = \|Tx\|^2 - |(Tx,x)|^2 \tag{1.2}$$

is translation-invariant. Incidentally, it is known that $M_T = \mathrm{d}(T,\mathbb{C})$, the distance of T to $\mathbb{C}$.

On the other hand, related to the Bernstein inequality [1], Furuta [10] and Lin [14] gave the following norm inequality on the difference of the Schwarz inequality

$$\|x\|^2\|y\|^2 - |(x,y)|^2 \le \frac{1}{|\alpha-\beta|}\|x+\alpha y\|^2\|x+\beta y\|^2 \tag{1.3}$$

for all $x, y \in H$ and $\alpha, \beta \in \mathbb{C}$ with $\alpha \ne \beta$. It is clear that the left hand side of (1.3) is the determinant of the Gram matrix

$$G(x,y) = \begin{pmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{pmatrix}.$$

So we recall the Gramian transformation formula, e.g. [3; Lemma 8.7.1]. For a 2×2 matrix $A = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix}$,

$$AG(x,y)A^* = G(\alpha_1 x + \alpha_2 y, \beta_1 x + \beta_2 y) \tag{1.4}$$

1991 *Mathematics Subject Classification.* 47A30 and 15A15.

Key words and phrases. Gramian, Gramian transformation formula, variance of operators, Bernstein inequality, Hadamard theorem and norm inequalities..

and consequently

$$|G(\alpha_1 x + \alpha_2 y, \beta_1 x + \beta_2 y)| = |\det A|^2 |G(x,y)|, \tag{1.5}$$

where both $|X|$ and $\det X$ are the determinant of X.

As a simple application of (1.5), we can explain the translation-invariance of the variance (1.2): Since

$$\mathrm{Var}_x(T) = |G(Tx,x)| \text{ and } \mathrm{Var}_x(T-\lambda) = |G(Tx-\lambda x, x)|,$$

we take $\alpha_1 = 1,\ \alpha_2 = -\lambda,\ \beta_1 = 0$ and $\beta_2 = 1$, i.e., $A = \begin{pmatrix} 1 & -\lambda \\ 0 & 1 \end{pmatrix}$. Then we have

$$AG(Tx,x)A^* = G(Tx - \lambda x, x)$$

and so $|G(Tx,x)| = |G((T-\lambda)x,x)|$ because $\det A = 1$. We here note that

$$\|x\|^2\|y\|^2 - |(x,y)|^2 = \|y\|^2\|x-\lambda y\|^2 - |(y, x-\lambda y)|^2 \tag{1.6}$$

is also showed by the same way as above.

In this note, we give some applications of *the Gramian transformation formula*

$$AG(x_1, \cdots, x_n)A^* = G(\sum_j a_{1j}x_j, \cdots, \sum_j a_{nj}x_j) \tag{1.7}$$

for $n \times n$ matrices $A = (a_{ij})$ and $x_1, \cdots, x_n \in H$. In other wards, we give natural proofs to some known theorems from the viewpoint of the Gramian transformation formula (1.7).

2. Norm inequality. A special case of (1.3) appeared in [15] to show Hua's determinant theorem is as follows:

$$\|x\|^2\|y\|^2 - |(x,y)|^2 \le \frac{1}{4}\|x+y\|^2\|x-y\|^2$$

for all $x, y \in H$. It is the case $\alpha = 1$ and $\beta = -1$ in (1.3) and follows from the norm equality

$$\|x+y\|^2\|x-y\|^2 - |(x+y, x-y)|^2 = 4(\|x\|^2\|y\|^2 - |(x,y)|^2).$$

This suggests us the following norm inequality.

Lemma 1. *The equality*

$$\|x+\alpha y\|^2\|x+\beta y\|^2 - |(x+\alpha y, x+\beta y)|^2 = |\alpha-\beta|^2(\|x\|^2\|y\|^2 - |(x,y)|^2) \tag{2.1}$$

holds for all $x, y \in H$ and $\alpha, \beta \in \mathbb{C}$.

We note that Lemma 1 implies (1.3) obviously and moreover (2.1) is rephrased by

$$|G(x+\alpha y, x+\beta y)| = |\alpha - \beta|^2 |G(x,y)|. \tag{2.2}$$

Namely, by taking $A = \begin{pmatrix} 1 & \alpha \\ 1 & \beta \end{pmatrix}$, we have (2.2) from (1.5) easily.

Incidentally, we can give an alternative proof to (2.1), based on (1.6): Put $u = x + \alpha y$ and $v = x + \beta y$. Then it follows from (1.6) that

$$\begin{aligned} &\|x+\alpha y\|^2\|x+\beta y\|^2 - |(x+\alpha y, x+\beta y)|^2 \\ &= \|v\|^2\|u-v\|^2 - |(v, u-v)|^2 \\ &= |\alpha-\beta|^2(\|v\|^2\|y\|^2 - |(v,y)|^2) \\ &= |\alpha-\beta|^2(\|x\|^2\|y\|^2 - |(x,y)|^2). \end{aligned}$$

Remark. (1) The inequality (1.3) is closely related to the Bernstein inequality. Extensions of the Bernstein inequality are discussed in [6] and [7].

(2) The variance of operators is generalized to the covariance of operators, see [5]. It is defined by

$$\mathrm{Cov}_x(A, B) = (Ax, Bx) - (Ax, x)(x, Bx)$$

for operators A and B on a Hilbert space H, where x is a unit vector in H. (It is the case of vector states.) The covariance is translation-invariant as well as the variance. We now define the covariance for vectors in H as follows:

$$\mathrm{Cov}_x(y, z) = (y, z) - (y, x)(x, z).$$

Clearly $\mathrm{Cov}_x(A, B) = \mathrm{Cov}_x(Ax, Bx)$ for a unit vector x in H. The translation-invariance of it can be explained from the determinantal view, that is,

$$\mathrm{Cov}_x(y, z) = \begin{vmatrix} (x,x) & (y,x) \\ (x,z) & (y,z) \end{vmatrix} = \begin{vmatrix} 1 & (y,x) \\ (x,z) & (y,z) \end{vmatrix}.$$

3. Gram-Schmidt orthogonalization process. In this section, we pay our attention to Gram-Schmidt (orthogonalization) process in order to apply the Gramian transformation formula (1.7).

For the sake of convenience, we cite the following fact in [9], which is easily obtained by (1.7):

Lemma 2. *Let $\{x_1, \cdots, x_n\}$ be a given linearly independent set in H, and $\{e_1, \cdots, e_n\}$ the orthonormal set obtained from $\{x_1, \cdots, x_n\}$ by the Gram-Schmidt process. If A is the matrix correponding to the Gram-Schmidt process, that is, A is triangular and satisfies*

$$\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

then

$$AG(x_1, \cdots, x_n)A^* = G(e_1, \cdots, e_n) = E_n$$

and

$$|\det A|^2 |G(x_1, \cdots, x_n)| = 1.$$

The Hadamard theorem says that

$$|G(x_1, \cdots, x_n)| \le \|x_1\|^2 \cdots \|x_n\|^2 \tag{3.1}$$

for $x_1, \cdots, x_n \in H$. As well-known, it is implied by Lemma 2. Actually

$$|G(x_1, \cdots, x_n)| = \frac{1}{|\det A|^2} = \frac{1}{|a_{11}|^2 \cdots |a_{nn}|^2}.$$

Noting that $\|x_k\|^2 \ge \frac{1}{|a_{kk}|^2}$ for $1 \le k \le n$, we have (3.1).

Following the above argument, we give an elementary proof to the following folk result:

Theorem 3. *Let M be the subspace generated by a linearly independent set $\{x_1, \cdots, x_n\}$. Then $d(x, M)$, the distance of x from M, is expressed as*

$$d^2(x, M) = \frac{|G(x, x_1, \cdots, x_n)|}{|G(x_1, \cdots, x_n)|}. \tag{3.2}$$

Proof. Let A be as in Lemma 2. Put $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$. Then it follows from Lemma 2 that

$$A_1 G(x, x_1, \cdots, x_n) A_1^* = G(x, e_1, \cdots, e_n)$$

and so

$$|\det A_1|^2 |G(x, x_1, \cdots, x_n)| = |G(x, e_1, \cdots, e_n)|.$$

Moreover, noting that $\{e_1, \cdots, e_n\}$ is an orthonormal basis of M, we have

$$|G(x, e_1, \cdots, e_n)| = \|x\|^2 - \sum_{j=1}^{n} |(x, e_j)|^2 = \|x - \sum_{j=1}^{n} (x, e_j) e_j\|^2 = d^2(x, M).$$

On the other hand, since

$$|\det A_1|^2 = |\det A|^2 = \frac{1}{|G(x_1, \cdots, x_n)|}$$

by Lemma 2, we have the required equality.

As a corollary, we show the following inequality on Gramian:

Corollary 4. *For given vectors $x_1, \cdots, x_n$, put $y_i = P x_i$ $(i = 1, \ldots, n)$ for a contraction P. If $y_1, \cdots, y_n$ are linearly independent, then*

$$\frac{|G(y_1, \cdots, y_{n-1})|}{|G(y_1, \cdots, y_n)|} \geq \frac{|G(x_1, \cdots, x_{n-1})|}{|G(x_1, \cdots, x_n)|} \tag{3.3}$$

and

$$|G(x_1, \cdots, x_n)| \geq |G(y_1, \cdots, y_n)| \tag{3.4}$$

Proof. Since P is a contraction, we have

$$d(x_n, [x_1, \cdots, x_{n-1}]) \geq d(P x_n, [P x_1, \cdots, P x_{n-1}]) = d(y_n, [y_1, \cdots, y_{n-1}]),$$

which implies (3.3) by Theorem 3.

The latter is shown by the use of the former (3.3). As a matter of fact, we have

$$\begin{aligned} &\frac{|G(x_1, \cdots, x_n)|}{|G(x_1, \cdots, x_{n-1})|} \frac{|G(x_1, \cdots, x_{n-1})|}{G(x_1, \cdots, x_{n-2})|} \cdots \frac{|G(x_1, x_2)|}{|G(x_1)|} \\ &\geq \frac{|G(y_1, \cdots, y_n)|}{|G(y_1, \cdots, y_{n-1})|} \frac{|G(y_1, \cdots, y_{n-1})|}{|G(y_1, \cdots, y_{n-2})|} \cdots \frac{|G(y_1, y_2)|}{|G(y_1)|} \end{aligned}$$

and so

$$|G(x_1, \cdots, x_n)| \geq |G(y_1, \cdots, y_n)| \frac{\|x_1\|^2}{\|y_1\|^2} \geq |G(y_1, \cdots, y_n)|.$$

Finally we give a simple proof to the following inequality, which is similar to that of Theorem 3.

Theorem 5. *If $x_1, \cdots, x_n$ are linearly indepedent vectors, then for $1 < k < n$*

$$\frac{|G(x_1,\cdots,x_n)|}{|G(x_1,\cdots,x_k)|} \le \frac{|G(x_2,\cdots,x_n)|}{|G(x_2,\cdots,x_k)|} \le \cdots \le |G(x_{k+1},\cdots,x_n)| \tag{3.5}$$

and in particular

$$|G(x_1,\cdots,x_n)| \le |G(x_1,\cdots,x_k)||G(x_{k+1},\cdots,x_n)|. \tag{3.6}$$

Proof. Let A_n be A in Lemma 2. Then we have

$$A_k G(x_1,\cdots,x_k)A_k^* = G(e_1,\cdots,e_k) = E_k$$

and

$$|\det A_k|^2 |G(x_1,\cdots,x_k)| = 1.$$

Therefore, if we put $A_1 = \begin{pmatrix} A_k & 0 \\ 0 & E_{n-k} \end{pmatrix}$, then

$$A_1 G(x_1,\cdots,x_n){A_1}^* = G(e_1,\cdots,e_k,x_{k+1},\cdots,x_n)$$

and

$$|\det A_1|^2 = |\det A|^2 = \frac{1}{|G(x_1,\cdots,x_k)|}.$$

Hence it follows that

$$\frac{|G(x_1,...,x_n)|}{|G(x_1,...,x_k)|} = |G(e_1,...,e_k,x_{k+1},...,x_n)| = \begin{vmatrix} E_k & B_1 \\ B_1^* & D_k \end{vmatrix}$$

and similarly

$$\frac{|G(x_m,...,x_n)|}{|G(x_m,...,x_k)|} = |G(e_m,...,e_k,x_{k+1},...,x_n)| = \begin{vmatrix} E_{k-m+1} & B_m \\ B_m^* & D_k \end{vmatrix}$$

for $1 < m \le k$, where E_j is the $j \times j$ identity matrix, $D_k = G(x_{k+1},\cdots,x_n)$ and

$$B_m = \begin{pmatrix} (e_m,x_{k+1}) & \cdots & (e_m,x_n) \\ \vdots & & \vdots \\ (e_k,x_{k+1}) & \cdots & (e_k,x_n) \end{pmatrix}.$$

We here recall Fisher's inequality: If $\begin{pmatrix} A & B \\ B^* & D \end{pmatrix}$ is positive definite, then $\begin{vmatrix} A & B \\ B^* & D \end{vmatrix} \le |A||D|$. So we have

$$\begin{vmatrix} E_k & B_1 \\ B_1^* & D_k \end{vmatrix} \le \begin{vmatrix} E_{k-1} & B_2 \\ B_2^* & D_k \end{vmatrix} \le \cdots \le \begin{vmatrix} E_1 & B_k \\ B_k^* & D_k \end{vmatrix} \le |D_k| = G(x_{k+1},\cdots,x_n),$$

which is equivalent to the conclusion (3.5).

Acknowledgement. The authors would like to express their thanks to Prof. S.Izumino for his valuable comment.

References

1. H.J.Bernstein, *An inequality for selfadjoint operators on a Hilbert space*, Proc. Amer. Math. Soc., **100** (1987), 319-321.
2. G.Björck and V.Thomee, *A property of bounded normal operators in Hilbert space*, Ark. Mat., **4** (1963), 551-555.
3. P.J.Davis, *Interpolation and Approximation*, Dover, New York, 1963.
4. J.I.Fujii and M.Fujii, *Theorems of Williams and Prasanna*, Math. Japon., **38** (1993), 35-37.
5. M.Fujii, T.Furuta, R.Nakamoto and S.-E.Takahasi, *Operator inequalities and covariance in noncommutative probability*, Math. Japon., **46** (1997), 317-320.
6. M.Fujii, T.Furuta and Y.Seo, *An inequality for some nonnormal operators - Extension to normal approximate eigenvalues*, Proc. Amer. Math. Soc., **118** (1993), 899-902.
7. M.Fujii, R.Nakamoto and Y.Seo, *Covariance in Bernstein's inequality for operators*, Nihonkai Math. J., **8** (1997), 1-6.
8. M.Fujii and S.Prasanna, *Translatable radii for operators*, Math. Japon., **26** (1981), 653-657.
9. T.Furuta, *An elementary proof of Hadamard's theorem*, Math. Vesnik, **23** (1971), 267-269.
10. T.Furuta, *An inequality for some nonnormal operators*, Proc. Amer. Math. Soc., **104** (1988), 1216-1217.
11. T.Furuta, S.Izumino and S.Prasanna, *A characterization of centroid operators*, Math. Japon., **27** (1982), 105-106.
12. G.Garske, *An inequality concerning the smallest disc that contains the spectrum of an operator*, Proc. Amer. Math. Soc., **78** (1980), 529-532.
13. S.Prasanna, *The norm of a derivation and the Björck-Thomee-Istratescu theorem*, Math. Japon., **26** (1981), 585-588.
14. C.-S.Lin, *Operator versions of inequalities and equalities on a Hilbert space*, Linear Algebra and its Appl., **268** (1998), 365-374.
15. M.Marcus, *On a determinantal inequality*, Amer. Math. Monthly, **65** (1958), 266-268.
16. J.G.Stampfli, *The norm of a derivation*, Pacific J. Math., **33** (1970), 737-747.
17. J.P.Williams, *Finite operators*, Proc. Amer. Math. Soc., **26** (1970), 129-136.

* Department of Mathematics, Osaka Kyoiku University, Kashiwara, Osaka 582, Japan

** Department of Applied Mathematics, Faculty of Science, Science University of Tokyo, Kagurazaka, Shinjuku, Tokyo 162, Japan

*** Faculty of Engineering, Ibaraki University, Hitachi, Ibaraki 316, Japan.