

A new proof of the Benedetti's inequality and some applications to perturbation to real eigenvalues and singular values

Una nueva prueba de desigualdad de Benedetti y algunas aplicaciones a la perturbación de valores propios reales y valores singulares

Humberto Sarria^{1,a}, Juan Carlos Martínez^{2,b}

Abstract. Using the standard deviation of the real samples $\mu_n \geq \dots \geq \mu_1$ and $\lambda_n \geq \dots \geq \lambda_1$, we refine the Chebyshev's inequality (refer to [5]),

$$\left(\frac{1}{n} \sum_{i=1}^n \mu_i\right) \left(\frac{1}{n} \sum_{i=1}^n \lambda_i\right) \leq \frac{1}{n} \sum_{i=1}^n \mu_i \lambda_i. \quad (1)$$

As a consequence, we obtain a new proof of the Benedetti's inequality (refer to [1], [2] and [4])

$$\frac{1}{n-1} \leq \frac{Cov(\mu, \lambda)}{s(\mu)s(\lambda)}, \quad (2)$$

where $Cov[\mu, \lambda]$, $s(\mu)$ and $s(\lambda)$ denote the covariance, and the standard deviations ($\neq 0$) of the sample vectors $\mu = (\mu_1, \dots, \mu_n)$ and $\lambda = (\lambda_1, \dots, \lambda_n)$, respectively.

We can also get very interesting applications to eigenvalues and singular values perturbation theory. For some kinds of matrices, the result that we present improves the well known Hoffmand-Weiland's inequality.

Keywords: Chebyshev's inequality, Hoffmand-Weiland's inequality, eigenvalues perturbation, singular value perturbation.

Resumen. Usando la desviación estándar de las muestras reales $\mu_n \geq \dots \geq \mu_1$ y $\lambda_n \geq \dots \geq \lambda_1$, refinamos la desigualdad de Chebyshev, ver [5],

$$\left(\frac{1}{n} \sum_{i=1}^n \mu_i\right) \left(\frac{1}{n} \sum_{i=1}^n \lambda_i\right) \leq \frac{1}{n} \sum_{i=1}^n \mu_i \lambda_i. \quad (1)$$

Como consecuencia, obtenemos una nueva prueba de desigualdad de Benedetti, ver [1], [2] y [4]

$$\frac{1}{n-1} \leq \frac{Cov(\mu, \lambda)}{s(\mu)s(\lambda)}, \quad (2)$$

¹Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia

²Departamento de Matemáticas, Universidad del Rosario, Bogotá, Colombia

^ahsarriaz@unal.edu.co

^bjuan.martinez@uexternado.edu.co

dónde $Cov[\mu, \lambda]$, $s(\mu)$ y $s(\lambda)$ denotan la covarianza, y las desviaciones estándar ($\neq 0$) de los vectores de muestra $\mu = (\mu_1, \dots, \mu_n)$ y $\lambda = (\lambda_1, \dots, \lambda_n)$, respectivamente.

También podemos obtener aplicaciones muy interesantes para los valores propios y la teoría de perturbaciones de valores singulares. Para algunos tipos de matrices, el resultado que presentamos mejora la conocida desigualdad de Hoffmand-Weiland.

Palabras claves: Desigualdad de Hoffmand-Weiland, Desigualdad de Chebyshev, perturbación de valores propios, perturbación de valores singulares.

Mathematics Subject Classification: 15A18, 15A42, 15A45, 62D05.

Recibido: enero de 2016

Aceptado: septiembre de 2016

1. Introduction

Carlo Benedetti presented in [1] (1957), one of the most interesting inequalities (2) of the applied statistic, used in measures of correlation. N. Georgescu [2] (1959) got a simpler proof of the Benedetti's theorem, using methods concerning the extrema of symmetrical, homogeneous algebraic forms. In 1994, T. Y Hwang and Ch. Y. Hu [4] rediscover the Benedetti's result, studying bounds of sample correlation coefficient, and in 2001 H. Sarria [6] discovered the inequality (3) in context of the matrix perturbation eigenvalues, that is equivalent to the Benedetti's inequality and is clearly a refinement of the Chebyshev's inequality (1). The inequality (3) was initially obtained using mathematical programming techniques. Once found, however, we remarked that the inequality admitted a shorter proof, using an elementary algebraic process. In this paper, we present this scheme.

For any real vector $\lambda = (\lambda_1, \dots, \lambda_n)$, we define the mean and the standard deviation by

$$m(\lambda) = \frac{1}{n} \sum_{i=1}^n \lambda_i \quad \text{and} \quad s(\lambda)^2 = \frac{1}{n} \sum_{i=1}^n \lambda_i^2 - \left(\frac{1}{n} \sum_{i=1}^n \lambda_i \right)^2,$$

respectively. The covariance of the vectors μ and λ are given by

$$Cov(\mu, \lambda) = \frac{1}{n} \mu^{tr} \lambda - m(\mu)m(\lambda).$$

The main result in this paper is:

Theorem 1.1. *If $\mu_n \geq \dots \geq \mu_1$ and $\lambda_n \geq \dots \geq \lambda_1$, with $n \geq 2$, then*

$$m(\lambda)m(\mu) + \frac{1}{n-1} s(\lambda)s(\mu) \leq \frac{1}{n} \mu^{tr} \lambda. \quad (3)$$

Equality holds, if and only if, any of the following cases is given: (a) $n = 2$, (b) $\mu_1 \leq \mu_2 = \dots = \mu_n$ and $\lambda_1 = \lambda_2 = \dots = \lambda_{n-1} \leq \lambda_n$, (c) $\mu_1 = \mu_2 = \dots = \mu_{n-1} \leq \mu_n$ and $\lambda_1 \leq \lambda_2 = \dots = \lambda_n$.

In order to obtain the proof of the theorem above, we will use the following result:

Lemma 1.2. *Let $a_1, a_2, \dots, a_n \geq 0$ and $b_1, b_2, \dots, b_n \geq 0$, with $n \geq 1$ then*

$$(n - 1)^2 \left(\sum_{i,j=1}^n c_{ij} a_i b_j \right)^2 \geq \sum_{i,j=1}^n c_{ij} a_i a_j \sum_{i,j=1}^n c_{ij} b_i b_j \tag{4}$$

where $c_{ij} = (n - \max\{i, j\} + 1)(\min\{i, j\} - 1)$. Equality holds if, and only if, one of the following holds: **(a)** $n = 1$, **(b)** $n = 2$, **(c)** $a_3 = \dots = a_n = 0$ and $b_2 = \dots = b_{n-1} = 0$ **(d)** $a_2 = \dots = a_{n-1} = 0$ and $b_3 = \dots = b_n = 0$.

Proof. If $n = 1$ or $n = 2$, the statement (4) is evident. Now we suppose $n > 2$. Note that $c_{ij} = c_{ji}$ for $i, j = 1, \dots, n$. Then

$$\left(\sum_{i,j=1}^n c_{ij} a_i b_j \right)^2 = \sum_{i,j=1}^n c_{ij} a_i b_j \sum_{k,t=1}^n c_{kt} a_k b_t = \sum_{i,j,k,t=1}^n c_{ij} c_{kt} a_i a_k b_j b_t.$$

We can also write,

$$\sum_{i,j=1}^n c_{ij} a_i a_j \sum_{i,j=1}^n c_{ij} b_i b_j = \sum_{i,j,k,t=1}^n c_{ik} c_{jt} a_i a_k b_j b_t.$$

Therefore,

$$\begin{aligned} (n - 1)^2 \left(\sum_{i,j=1}^n c_{ij} a_i b_j \right)^2 - \sum_{i,j=1}^n c_{ij} a_i a_j \sum_{i,j=1}^n c_{ij} b_i b_j &= \\ (n - 1)^2 \sum_{i,j,k,t=1}^n c_{ij} c_{kt} a_i a_k b_j b_t - \sum_{i,j,k,t=1}^n c_{ik} c_{jt} a_i a_k b_j b_t &= \\ = \sum_{i,j,k,t=1}^n [(n - 1)^2 c_{ij} c_{kt} - c_{ik} c_{jt}] a_i a_k b_j b_t. &\tag{5} \end{aligned}$$

Note that $c_{i1} = c_{1i} = 0$, so we can assume $2 \leq i \leq n$ and $2 \leq j \leq n$. In this case, $1 \leq n - i + 1 \leq n - 1$ and $1 \leq i - 1 \leq n - 1$, for any $i, 2 \leq i \leq n$. As a consequence, we have $1 \leq (n - j + 1)(i - 1) \leq (n - 1)^2$, for any $2 \leq i, j \leq n$. On the other hand, if $a = \min\{i, j, k, t\}$ and $A = \max\{i, j, k, t\}$, then

$$c_{aA} \leq c_{ij} c_{kt} \leq c_{aA} (n - 1)^2, \tag{6}$$

for $i, j, k, t \in \{2, \dots, n\}$. Therefore

$$(n - 1)^2 c_{ij} c_{kt} - c_{ik} c_{jt} \geq 0, \tag{7}$$

for $2 \leq i, j, k, t \leq n$. Using (5), and (7)

$$(n-1)^2 \left(\sum_{i,j=2}^n c_{ij} a_i b_j \right)^2 - \sum_{i,j=2}^n c_{ij} a_i a_j \sum_{i,j=2}^n c_{ij} b_i b_j \geq 0,$$

and this proves our inequality.

We are going to get the equality conditions. Note that $(n-1)^2 c_{2n}^2 - c_{22} c_{nn} = (n-1)^2 c_{n2}^2 - c_{22} c_{nn} = 0$, so from (5), clearly equality holds if, both $a_3 = \dots = a_n = 0$ and $b_2 = \dots = b_{n-1} = 0$, or if both $a_2 = \dots = a_{n-1} = 0$ and $b_3 = \dots = b_n = 0$. Now, suppose that the equality holds, and there exist $p, q \in \{2, \dots, n\}$, $p \leq q$ such that $a_p, b_q > 0$, from here, (5), and (7)

$$(n-1)^2 c_{pq}^2 - c_{pp} c_{qq} = 0. \quad (8)$$

But from (6), (8) is true if, and only if, both $c_{pq}^2 = c_{aA}$, and $c_{pp} c_{qq} = c_{aA} (n-1)^2$. This implies that $p = 2$ and $q = n$, i.e., if $a_3 = \dots = a_n = 0$ and $b_2 = \dots = b_{n-1} = 0$. The symmetry case $q \leq p$ implies, $a_2 = \dots = a_{n-1} = 0$ and $b_3 = \dots = b_n = 0$. \square

Proof of theorem 1.1. First, observe that (3) is equivalent to the inequality

$$(n-1)^2 \left(n \sum_{i=1}^n \lambda_i \mu_i - \sum_{i=1}^n \lambda_i \sum_{i=1}^n \mu_i \right)^2 \geq \left(n \sum_{i=1}^n \lambda_i^2 - \left(\sum_{i=1}^n \lambda_i \right)^2 \right) \left(n \sum_{i=1}^n \mu_i^2 - \left(\sum_{i=1}^n \mu_i \right)^2 \right).$$

But

$$n \sum_{i=1}^n \lambda_i \mu_i - \sum_{i=1}^n \lambda_i \sum_{i=1}^n \mu_i = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\lambda_j - \lambda_i)(\mu_j - \mu_i),$$

so the inequality (3) is equivalent to

$$(n-1)^2 \left(\sum_{i=1}^n \sum_{j=1}^n (\lambda_j - \lambda_i)(\mu_j - \mu_i) \right)^2 \geq \sum_{i=1}^n \sum_{j=1}^n (\lambda_j - \lambda_i)^2 \cdot \sum_{i=1}^n \sum_{j=1}^n (\mu_j - \mu_i)^2,$$

or to

$$(n-1)^2 \left(\sum_{i < j} (\lambda_j - \lambda_i)(\mu_j - \mu_i) \right)^2 \geq \sum_{i < j} (\lambda_j - \lambda_i)^2 \cdot \sum_{i < j} (\mu_j - \mu_i)^2. \quad (9)$$

If $i < j$, we can write,

$$\begin{aligned} (\lambda_j - \lambda_i)(\mu_j - \mu_i) &= \left[(\lambda_j - \lambda_{j-1}) + (\lambda_{j-1} - \lambda_{j-2}) + \dots + (\lambda_{i+1} - \lambda_i) \right] \cdot \\ &\quad \left[(\mu_j - \mu_{j-1}) + (\mu_{j-1} - \mu_{j-2}) + \dots + (\mu_{i+1} - \mu_i) \right]. \end{aligned} \quad (10)$$

Expanding (10), we get a sum of terms of the form $(\lambda_r - \lambda_{r-1})(\mu_k - \mu_{k-1})$ where $i < r, k \leq j$. Substituting (10) in the sum

$$\sum_{i < j}^n (\lambda_j - \lambda_i)(\mu_j - \mu_i)$$

of the left side sum of (9), we can observe that if $2 \leq r, k \leq n$, the term $(\lambda_r - \lambda_{r-1})(\mu_k - \mu_{k-1})$ appears in this sum, once, each time that the term $(\lambda_j - \lambda_i)(\mu_j - \mu_i)$ appears for both $1 < i \leq \min\{k, r\}$ and $\max\{k, r\} \leq j \leq r$; that is $c_{rk} = (n - \max\{r, k\} + 1)(\min\{r, k\} - 1)$ times. Making a similar analysis in every sum in the right side of (9), we can see that (9) is equivalent to the inequality

$$(n-1)^2 \left(\sum_{i=2}^n \sum_{j=2}^n c_{ij} (\lambda_i - \lambda_{i-1})(\mu_j - \mu_{j-1}) \right)^2 \geq \sum_{i=2}^n \sum_{j=2}^n c_{ij} (\lambda_i - \lambda_{i-1})(\lambda_j - \lambda_{j-1}) \cdot \sum_{i=2}^n \sum_{j=2}^n c_{ij} (\mu_i - \mu_{i-1})(\mu_j - \mu_{j-1}).$$

Finally, we make $a_i = \lambda_i - \lambda_{i-1}$ and $b_j = \mu_j - \mu_{j-1}$, for $i, j = 2, \dots, n$ and apply the Lemma 1.2, in order to get (3). \square

Observation: Benedetti's inequality (2), is clearly equivalent to inequality (3).

2. Some applications to matrix theory

We will use \mathcal{C} to denote the set of complex numbers, and $M_{n,m}(\mathcal{C})$ for the space of $n \times m$ complex matrices. If $n = m$, we will use $M_n(\mathcal{C})$ for the square complex matrix space. If $X \in M_n(\mathcal{C})$, X^* and X^{tr} will denote the conjugate transpose, and the transpose of X , respectively, and $Tr(X)$ its trace.

2.1. Real eigenvalues

One of the problems in classical perturbation theory for matrix eigenvalues consists in providing bounds to the Euclidean distance $\sqrt{\sum_{i=1}^n |\mu_i - \lambda_i|^2}$, where $\mu_1, \mu_2, \dots, \mu_n$ and $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $n \times n$ complex matrices A and B , respectively. The most interesting result in this area is the Hoffman-Weilandt theorem (e.g., refer to [3]):

Theorem 2.1. (Hoffman-Weilandt) *If A and B are normal matrices, then there exists a suitable numbering of the eigenvalues such that*

$$\sum_{i=1}^n |\mu_i - \lambda_i|^2 \leq \|A - B\|_F^2$$

where $\|\cdot\|_F$ denoted the Frobenius' norm.

In this section, we suppose that the matrices A and B have real eigenvalues. Also, we will define the mean and the standard deviation of a matrix X with real eigenvalues respectively by

$$m(X) = \frac{\text{Tr}(X)}{n} \quad \text{and} \quad s(X)^2 = \frac{\text{Tr}(X^2)}{n} - \frac{\text{Tr}(X)^2}{n^2}.$$

The inequalities given here are tightest when all the eigenvalues are real; this happens for example when the matrices are Hermitian. Our main purpose is to give bounds for the Euclidean distance (considering eigenvalues in increasing order) using traces of the matrices A , B , A^2 and B^2 .

Theorem 2.2. *If $\mu_n \geq \dots \geq \mu_1$ and $\lambda_n \geq \dots \geq \lambda_1$, with $n \geq 2$, then*

$$(i) \quad nm(A)m(B) + \frac{n}{n-1}s(A)s(B) \leq \mu^{tr}\lambda,$$

$$(ii) \quad \sum_{i=1}^n (\lambda_i - \mu_i)^2 \leq \text{Tr}(A^2) + \text{Tr}(B^2) - 2nm(A)m(B) - \frac{2n}{n-1}s(A)s(B),$$

where $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)^{tr}$ and $\mu = (\mu_1, \mu_2, \dots, \mu_n)^{tr}$.

Equality holds in (i) and (ii) if, and only if, any the following cases is given:

(a) $n = 2$, (b) $\mu_1 \leq \mu_2 = \dots = \mu_n$ and $\lambda_1 = \lambda_2 = \dots = \lambda_{n-1} \leq \lambda_n$, (c) $\mu_1 = \mu_2 = \dots = \mu_{n-1} \leq \mu_n$ and $\lambda_1 \leq \lambda_2 = \dots = \lambda_n$.

Proof. (i) follows from Theorem 1.1. (ii) is equivalent to (i). □

Observation: Inequalities in Theorem 2.2, are equivalent to the inequality just given in terms of the mean and the standard deviation

$$\sum_{i=1}^n (\lambda_i - \mu_i)^2 \leq (\alpha_1 - \alpha_2)^2 + (n-2)(\alpha_2 - \alpha_3)^2 + (\alpha_3 - \alpha_4)^2, \quad (11)$$

where

$$\begin{aligned} \alpha_1 &= \frac{\text{Tr}(A)}{n} - \sqrt{n-1} s(A) \\ \alpha_2 &= \frac{\text{Tr}(B)}{n} - \sqrt{\frac{1}{n-1}} s(B) \\ \alpha_3 &= \frac{\text{Tr}(A)}{n} + \sqrt{\frac{1}{n-1}} s(A) \\ \alpha_4 &= \frac{\text{Tr}(B)}{n} + \sqrt{n-1} s(B). \end{aligned}$$

The proof is long but straightforward.

2.2. Singular values

We can extend the bounds in Theorem 2.2 to the case of perturbation of singular values. Let be $n + m > 2$, and $A, B \in M_{n,m}(\mathbb{C})$, with singular eigenvalues $\alpha_1, \alpha_2, \dots, \alpha_q$ and $\beta_1, \beta_2, \dots, \beta_q$ respectively, where $q = \min\{n, m\}$. The proof is straightforward using Theorem 2.2 and the following result (refer to [3] Theorem 7.3.7).

Theorem 2.3. *Let $\tilde{A} \in M_{m,n}$ defined by*

$$\tilde{A} = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$$

Let be $\iota_1, \iota_2, \dots, \iota_q$ nonnegative real numbers. The singular values of A are $\iota_1, \iota_2, \dots, \iota_q$ if, and only if, the $m + n$ eigenvalues of \tilde{A} are $\iota_1, \iota_2, \dots, \iota_q, -\iota_1, -\iota_2, \dots, -\iota_q$, and $|m - n|$ additional 0's.

Theorem 2.4. *If $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_q$ and $\beta_1 \leq \beta_2 \leq \dots \leq \beta_q$. Then,*

$$(i) \frac{n}{n+m-1} \sqrt{m(AA^*)m(BB^*)} \leq \alpha^{tr} \beta.$$

$$(ii) \sum_{i=1}^n (\alpha_i - \beta_i)^2 \leq Tr(AA^*) + Tr(BB^*) - \frac{2n}{n+m-1} \sqrt{m(AA^*)m(BB^*)}.$$

Equality holds in (i) and (ii) if, and only if, either $A = 0$ or $B = 0$.

3. Examples

To illustrate the bounds given here, we present some numerical examples.

Example 3.1. Let

$$A = \begin{pmatrix} 6 & 0 & 0 \\ 1 & 3 & 1 \\ 2 & 4 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 5 \end{pmatrix},$$

These matrices have real eigenvalues $(-1, 4, 6)$ and $(0, \frac{1}{2}(7 - \sqrt{65}), \frac{1}{2}(7 + \sqrt{65}))$ respectively. Note that Hoffmand - Weiland's theorem is not applicable here, because A and B are not normal.

Considering order of the eigenvalues, by the Theorem 2.2(i), we get the lower bound:

$$37.2583 \leq \mu^{tr} \lambda \approx 45.7179.$$

Considering order of the eigenvalues, by the Theorem 2.2(ii), we get the upper bound:

$$4.3086 \approx \sqrt{\sum_{i=1}^3 (\mu_i - \lambda_i)^2} \leq 5.9568$$

The following four examples compare Hoffman-Weiland's bound and the bound (ii) in Theorem 2.2:

Example 3.2. (a) Let

$$A = \begin{pmatrix} 1 & 2 & -3 & 1 \\ 2 & 5 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix}.$$

be. We generate B , through the MATLAB's instructions

$$B = \text{rand}(4); \quad B = B' * B;$$

in order to get the following

$\sqrt{\sum_{i=1}^4 (\mu_i - \lambda_i)^2}$	Theorem 2.2	Hoffman-Weiland
4.69	7.23	8.05
5.01	6.91	7.66
5.58	7.15	7.68
4.92	7.01	7.98
3.74	6.74	7.50

Table 1

(b) Now, we make

$$A = \begin{pmatrix} 1 & 2 & -3 & 1 \\ 2 & 5 & 0 & 0 \\ -3 & 0 & 1 & 2 \\ 1 & 0 & 2 & 3 \end{pmatrix}$$

$\sqrt{\sum_{i=1}^4 (\mu_i - \lambda_i)^2}$	Theorem 2.2	Hoffman-Weiland
4.46	7.14	8.01
5.41	7.26	8.20
5.64	7.39	8.13
5.17	7.38	7.41
5.24	7.10	7.89

Table 2

(c)

$$A = \begin{pmatrix} 1 & 2 & -3 & 1 \\ 2 & 5 & 7 & 4 \\ -3 & 7 & 1 & 2 \\ 1 & 4 & 2 & 3 \end{pmatrix}$$

$\sqrt{\sum_{i=1}^4 (\mu_i - \lambda_i)^2}$	Theorem 2.2	Hoffman-Weiland
10.21	12.86	11.25
9.72	12.73	12.03
10.75	13.12	11.34
9.56	12.83	12.80
9.77	12.90	12.14

Table 3

- (d) For small perturbation, Hoffman-Weiland's bound is better than bound (ii) in Theorem 2.2. In the following table we illustrate this aspect. The matrices A and B are generated via the MATLAB's instructions

$$A = \text{rand}(4); A = 10 * A * A';$$

$$B = \text{rand}(4); B = A + 0.1 * B * B';$$

$\sqrt{\sum_{i=1}^4 (\mu_i - \lambda_i)^2}$	Theorem 2.2	Hoffman-Weiland
2.93	50.38	3.20
3.48	39.56	4.08
6.23	51.35	6.62
4.81	37.56	4.93
2.72	59.64	3.62

Table 4

References

- [1] C. Benedetti, *Di alcune disuguaglianze collegate al campo di variazione di indici statistical coefficients*, Metron, International Journal of Statistics **18** (1957), no. 3-4, 102–125.
- [2] N. Georgescu-Roegen, *On the extrema of some statistical*, Metron, International Journal of Statistics **19** (1959), no. 3-4, 38–45.
- [3] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.
- [4] Tea-Yuan Hwang and Chin-Yuan Hu, *The best lower bound of coefficient with ordered sample correlation restrictions*, Statistics & Probability Letters **19** (1994), 195–198.

- [5] D. S. Mitrinović, *Analytic inequalities*, Springer-Verlang, 1970.
- [6] H. Sarria, *Una estimación de la perturbación de autovalores usando trazas y algunos problemas de preservación lineal sobre los espacios de matrices simétricas, antisimétricas y hermitianas*, Ph.D. thesis, Universidad Nacional de Colombia, Diciembre 2001.