Mathematical
Inequalities
& Applications
Volume 20, Number 1 (2017), 1–12 doi:10.7153/mia-20-01

SOME GENERALIZATIONS AND PROBABILITY VERSIONS OF SAMUELSON'S INEQUALITY

HONGWEI JIN AND JULIO BENÍTEZ

(*Communicated by C. P. Niculescu*)

Abstract. Several generalizations of Samuelson's inequality are given, including complex data and inequalities concerning random variables in a probability space. The proofs of these generalizations need only a well known result from inner product spaces, namely, Bessel's inequality. Finally we apply these generalizations to locate the eigenvalues of certain matrices and tensors, as well as the complex roots of polynomials.

1. Preliminaries

Samuelson's inequality establishes that for $x_1,\dots,x_n \in \mathbb{R}$, one has

$$[x_j - m(x)]^2 \leqslant (n-1)s^2, \tag{1}$$

where $m(x)$ is the *arithmetic mean* and s is the *sample variance* of the real data $x_1,\dots,x_n$, i.e.,

$$m(x) = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m(x))^2.$$

Samuelson's inequality was proven in [7]. Several proofs of this inequality can be found in the literature (see [3]). We present some generalizations of Samuelson's inequality with simple proofs which require Bessel's inequality.

Bessel's inequality is a well known fact of inner product spaces. If $\{\widehat{\mathbf{v}}_1,\dots,\widehat{\mathbf{v}}_n\}$ is an orthonormal set in a complex Euclidean space E and $\mathbf{x} \in E$, then

$$|\langle \mathbf{x}, \widehat{\mathbf{v}}_1\rangle|^2 + \cdots + |\langle \mathbf{x}, \widehat{\mathbf{v}}_n\rangle|^2 \leqslant \|\mathbf{x}\|^2. \tag{2}$$

Furthermore, the inequality (2) becomes an equality if and only if $\mathbf{x} \in \operatorname{span}\{\widehat{\mathbf{v}}_1,\dots,\widehat{\mathbf{v}}_n\}$ (see, e.g., the proof of Bessel's inequality given in [1, p. 441]).

Mathematics subject classification (2010): 60E15, 15A18, 30C15.
Keywords and phrases: Samuelson's inequality, Bessel's inequality, random variables, location of eigenvalues of matrices and tensors, location of the complex roots of polynomials.

Paper MIA-20-01

2. Generalizations of Samuelson's inequality

In this section, we will give several generalizations of Samuelson's inequality.

We shall review the definition of the expectation and variance of complex random variables since usually the real random variables are studied. Let $(\Omega,\Sigma,\mathbb{p})$ be a probability space and let $Z:\Omega\to\mathbb{C}$ be a complex random variable, i.e., $\operatorname{Re}Z:\Omega\to\mathbb{R}$ and $\operatorname{Im}Z:\Omega\to\mathbb{R}$ are random variables. For any measurable $A\subset\mathbb{C}$ one has

$$\mathbb{p}(Z\in A)=\mathbb{p}(Z^{-1}(A))=\int_{Z^{-1}(A)}\mathrm{d}\mathbb{p}.$$

The *expectation* of Z is defined by

$$\mathrm{E}(Z)=\int_\Omega Z\,\mathrm{d}\mathbb{p}=\int_\Omega \operatorname{Re}Z\,\mathrm{d}\mathbb{p}+\mathrm{i}\int_\Omega \operatorname{Im}Z\,\mathrm{d}\mathbb{p},$$

provided the two integrals in the right hand side exist. The *variance* of Z is defined by $\operatorname{Var}(Z)=\mathrm{E}\left(|Z-\mathrm{E}(Z)|^2\right)$. Observe that although Z is complex, $\operatorname{Var}(Z)$ is always real (and non-negative). Since $\overline{\mathrm{E}(Z)}=\mathrm{E}(\overline{Z})$, after few manipulations, one gets $\operatorname{Var}(Z)=\mathrm{E}(|Z|^2)-|\mathrm{E}(Z)|^2$.

Let $S\in\Sigma$ be such that $\mathbb{p}(S)\neq 0$. Define $\Sigma_{|S}=\{R\cap S: R\in\Sigma\}$ and $\mathbb{p}_{|S}:\Sigma_{|S}\to\mathbb{R}$ given by $\mathbb{p}_{|S}(R)=\mathbb{p}(R)/\mathbb{p}(S)$. It is easily checked that $(S,\Sigma_{|S},\mathbb{p}_{|S})$ is a probability space. Therefore, if $Z:\Omega\to\mathbb{C}$ is a complex random variable and A is a measurable subset of $\mathbb{C}$ such that $\mathbb{p}(Z\in A)\neq 0$, then the restriction of Z to $Z^{-1}(A)$ is a random variable defined in the probability space $(Z^{-1}(A),\Sigma_{|Z^{-1}(A)},\mathbb{p}_{|Z^{-1}(A)})$. This new random variable is called *the A-truncated distribution of Z* and this distribution will be denoted by $Z|A$ (intuitively speaking, Z is truncated so that only the values in A are observed). The expectation of $Z|A$ is

$$\mathrm{E}(Z|A)=\int_{Z^{-1}(A)}Z\,\mathrm{d}\mathbb{p}_{|Z^{-1}(A)}=\frac{1}{\mathbb{p}(Z\in A)}\int_{Z^{-1}(A)}Z\,\mathrm{d}\mathbb{p}.$$

THEOREM 1. *Let $(\Omega,\Sigma,\mathbb{p})$ be a probability space and let $Z:\Omega\to\mathbb{C}$ be a random variable whose expectation and variance are finite. If A is a measurable subset of $\mathbb{C}$ such that $\mathbb{p}(Z\in A)\neq 0$, then*

$$\frac{\mathbb{p}(Z\notin A)}{\mathbb{p}(Z\in A)}\operatorname{Var}(Z)\geqslant|\mathrm{E}(Z)-\mathrm{E}(Z|A)|^2. \tag{3}$$

This inequality becomes an equality if and only if $\mathbb{p}(Z\in A)=1$ or Z takes two values α and β such that $\alpha\in A$ and $\beta\notin A$.

Proof. If $\mathbb{p}(Z\in A)=1$, then $\mathbb{p}(Z\notin A)=0$ and $\mathrm{E}(Z|A)=\int_{Z^{-1}(A)}Z\,\mathrm{d}\mathbb{p}=\int_\Omega Z\,\mathrm{d}\mathbb{p}=E(Z)$ because $\Omega\setminus Z^{-1}(A)$ has measure zero. Therefore, if $\mathbb{p}(Z\in A)=1$, the inequality becomes an equality. Thus, in the rest of the proof we can assume that $\mathbb{p}(Z\in A)<1$.

We will consider the Hilbert space $\mathscr{L}^2(\Omega,\mathbb{p})=\{u:\Omega\to\mathbb{C}:\int_\Omega|u|^2\,\mathrm{d}\mathbb{p}<\infty\}$ endowed with the inner product $\langle u,v\rangle=\int_\Omega u\overline{v}\,\mathrm{d}\mathbb{p}$. For $S\subset\Omega$, we denote its indicator function by $\mathbb{1}_S$ (i.e., $\mathbb{1}_S$ has value 1 at points of S and 0 at points of $\Omega \setminus S$). We denote $p = \mathbb{p}(Z \in A)$ and $q = 1 - p = \mathbb{p}(Z \notin A)$.

It is simple to prove that the functions $\mathbb{1}_\Omega$ and $-q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)}$ are orthogonal:

$$\left\langle \mathbb{1}_\Omega, -q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)} \right\rangle = -q\mathbb{p}(Z \in A) + p\mathbb{p}(Z \notin A) = -qp + pq = 0.$$

The following equalities are trivial: $\|Z\|^2 = \int_\Omega |Z|^2 \, d\mathbb{p} = \mathrm{E}(|Z|^2)$, $\langle Z, \mathbb{1}_\Omega \rangle = \int_\Omega Z \, d\mathbb{p} = \mathrm{E}(Z)$, and $\|\mathbb{1}_\Omega\|^2 = \int_\Omega d\mathbb{p} = 1$. Also we have that

$$\begin{aligned}\left\langle Z, -q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)} \right\rangle &= -q \int_{Z^{-1}(A)} Z \, d\mathbb{p} + p \int_{\Omega\setminus Z^{-1}(A)} Z \, d\mathbb{p} \\ &= (-q-p) \int_{Z^{-1}(A)} Z \, d\mathbb{p} + p \int_\Omega Z \, d\mathbb{p} \\ &= p\,[\mathrm{E}(Z) - \mathrm{E}(Z|A)],\end{aligned}$$

and by noticing that $\mathbb{1}_{Z^{-1}(A)}$ and $\mathbb{1}_{\Omega\setminus Z^{-1}(A)}$ are orthogonal,

$$\left\| -q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)} \right\|^2 = q^2 \left\| \mathbb{1}_{Z^{-1}(A)} \right\|^2 + p^2 \left\| \mathbb{1}_{\Omega\setminus Z^{-1}(A)} \right\|^2 = q^2 p + p^2 q = pq \neq 0.$$

Now, it is sufficient to apply Bessel's inequality (2) when $\mathbf{x} = Z$, $\widehat{\mathbf{v}}_1 = \mathbb{1}_\Omega$, and $\widehat{\mathbf{v}}_2 = \| -q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)} \|^{-1} (-q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)})$ to get the inequality (3).

If the inequality (3) is an equality, then $Z \in \operatorname{span}\{\mathbb{1}_\Omega, -q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)}\}$. Therefore, there exist $\lambda, \mu \in \mathbb{C}$ such that

$$Z = \lambda \mathbb{1}_\Omega + \mu \left(-q\mathbb{1}_{Z^{-1}(A)} + p\mathbb{1}_{\Omega\setminus Z^{-1}(A)} \right). \tag{4}$$

This equality implies that

$$Z(\omega) = \begin{cases} \lambda - \mu q & \text{if } \omega \in Z^{-1}(A), \\ \lambda + \mu p & \text{if } \omega \notin Z^{-1}(A). \end{cases}$$

Observe that $Z^{-1}(A)$ and $\Omega \setminus Z^{-1}(A)$ are not empty sets because $\mathbb{p}(Z \in A) \neq 0$ and $\mathbb{p}(Z \notin A) \neq 0$. Therefore, exist $\omega_1 \in Z^{-1}(A)$ and $\omega_2 \notin Z^{-1}(A)$, and thus, $Z(\omega_1) = \lambda - \mu q$ and $Z(\omega_2) = \lambda + \mu p$. If $\lambda + \mu p \in A$, then $w_2 \in Z^{-1}(\lambda + \mu p) \subseteq Z^{-1}(A)$, which is not possible, hence $\lambda + \mu p \notin A$. Similarly, one gets $\lambda - \mu q \in A$.

If Z takes two values α and β such that $\alpha \in A$ and $\beta \notin A$, then $Z = \alpha\mathbb{1}_{Z^{-1}(A)} + \beta\mathbb{1}_{\Omega\setminus Z^{-1}(A)}$. Define $\lambda = \alpha p + \beta q$ and $\mu = \beta - \alpha$. It is not difficult to see that $\lambda - \mu q = \alpha$ and $\lambda + \mu p = \beta$. Therefore, Z can be written as in (4), and by Bessel's inequality, the inequality (3) becomes an equality. $\square$

REMARK 1. Let $(\Omega, \Sigma, \mathbb{p})$ be a probability space and let A be a measurable subset of $\mathbb{C}$. We denote by A^c the complementary of A in $\mathbb{C}$, i.e., $A^c = \mathbb{C} \setminus A$. If Z is a complex random variable such that $\mathbb{p}(Z \in A) \neq 0$ and $\mathbb{p}(Z \in A^c) \neq 0$, then

$$\mathbb{p}(Z \in A^c)\,[\mathrm{E}(Z|A^c) - \mathrm{E}(Z|A)] = \mathrm{E}(Z) - \mathrm{E}(Z|A). \tag{5}$$

In fact,

$$\begin{aligned}\mathbb{p}(Z\in A^{\mathrm{c}})\left[\mathrm{E}(Z|A^{\mathrm{c}})-\mathrm{E}(Z|A)\right] &= \int_{Z^{-1}(A^{\mathrm{c}})} Z\,\mathrm{d}\mathbb{p} - \frac{\mathbb{p}(Z\in A^{\mathrm{c}})}{\mathbb{p}(Z\in A)}\int_{Z^{-1}(A)} Z\,\mathrm{d}\mathbb{p}\\ &= \int_{\Omega} Z\,\mathrm{d}\mathbb{p} - \left(1+\frac{\mathbb{p}(Z\in A^{\mathrm{c}})}{\mathbb{p}(Z\in A)}\right)\int_{Z^{-1}(A)} Z\,\mathrm{d}\mathbb{p}\\ &= \int_{\Omega} Z\,\mathrm{d}\mathbb{p} - \frac{1}{\mathbb{p}(Z\in A)}\int_{Z^{-1}(A)} Z\,\mathrm{d}\mathbb{p},\end{aligned}$$

which proves (5).

REMARK 2. In view of the equality (5), under the assumptions of Theorem 1 and in addition $\mathbb{p}(Z\notin A)\neq 0$, we have

$$\mathrm{Var}(Z)\geqslant \mathbb{p}(Z\in A)\mathbb{p}(Z\notin A)\left|\mathrm{E}(Z|A^{c})-\mathrm{E}(Z|A)\right|^{2}.$$

To the best knowledge of the authors, the following result is not known in the literature. It generalizes Samuelson's inequality (1) in two ways: it deals with complex numbers and more than one data is compared with the arithmetic mean. The sample variance of the complex data $\{z_1,\dots,z_n\}$ is defined by $\frac{1}{n}\sum_{k=1}^{n}|z_k-m(z)|^2$.

THEOREM 2. *For* $z_1,\dots,z_n\in\mathbb{C}$ *and* $p\in\{1,\dots,n\}$, *one has*

$$|m(z)-m(y)|^2\leqslant\frac{n-p}{p}s^2, \tag{6}$$

where $m(z)$ *and* $m(y)$ *are the arithmetic mean of* $\{z_1,\dots,z_n\}$ *and* $\{z_1,\dots,z_p\}$, *respectively, and* s^2 *is the sample variance of* $\{z_1,\dots,z_n\}$.

Proof. Let $\Omega=\{1,\dots,n\}$. If we set $\mathbb{p}(R)=\mathrm{card}(R)/n$ for $R\subseteq\Omega$, then it is clear that $(\Omega,\mathscr{P}(\Omega),\mathbb{p})$ is a probability space. Let $\varepsilon>0$ and for $k=1,\dots,n$, let $w_k\in\mathbb{C}$ be such that $|w_k-z_k|<\varepsilon$ and $w_1,\dots,w_n$ are pairwise distinct. Let $W:\Omega\to\mathbb{C}$ be the complex random variable defined by $W(k)=w_k$. Finally, let $A=\{w_1,\dots,w_p\}$. It is clear that $\mathrm{E}(W)=\sum_{k=1}^{n}w_k\mathbb{p}(W=w_k)=(w_1+\cdots+w_n)/n$, $\mathrm{E}(W|A)=(w_1+\cdots+w_p)/p$, and $\mathrm{Var}(W)=\mathrm{E}(|W-\mathrm{E}(W)|^2)$ is the sample variance of $w_1,\dots,w_n$. It is enough to apply Theorem 1 and let $\varepsilon\to 0$ to prove this theorem. □

REMARK 3. The inequality (6) can be proved by using Bessel's inequality without proving Theorem 2. The idea is to use (2) for $\mathbf{x}=(z_1,\dots,z_n)$, the subspace spanned by $\mathbb{1}_n=(1,\dots,1)$ and $\mathbf{a}=(p-n,\dots,p-n,p,\dots,p)$, where $p-n$ is repeated p times, and $\mathbb{C}^n$ is endowed with the standard inner product, i.e., if $\mathbf{u}=(u_k)_{k=1}^{n}\in\mathbb{C}^n$ and $\mathbf{v}=(v_k)_{k=1}^{n}\in\mathbb{C}^n$, then $\langle\mathbf{u},\mathbf{v}\rangle=\sum_{k=1}^{n}u_k\overline{v_k}$.

REMARK 4. The procedure established in the previous remark can be considered to prove that the inequality (6) becomes an equality if and only if $\mathbf{x}\in\mathrm{span}\{\mathbb{1}_n,\mathbf{a}\}$, and this is equivalent to say that $\{z_1,\dots,z_p\}$ and $\{z_{p+1},\dots,z_n\}$ are both singletons.

REMARK 5. Observe that the case $p=1$ and $z_1,\dots,z_n\in\mathbb{C}$ reduces to

$$|m(z)-z_i|^2\leqslant(n-1)s^2,\quad i\in\{1,\dots,n\},$$

which is the complex version of Samuelson's inequality, and $p=1$, $z_1,\dots,z_n\in\mathbb{R}$ reduces to the classical Samuelson's inequality.

REMARK 6. The inequality of Theorem 2 can be written in another way. Under the notation of this theorem, the next equality can be easily checked, $(n-p)[m(y)-m(x)]=n[m(y)-m(z)]$, being $m(x)$ the arithmetic mean of $\{z_{p+1},\dots,z_n\}$. Hence,

$$|m(y)-m(x)|^2\leqslant\frac{n^2}{p(n-p)}s^2.$$

Also, we can compare the arithmetic mean of a whole sample with the mean of two subsets of the sample.

THEOREM 3. *For $z_1,\dots,z_n\in\mathbb{C}$ and $p,q\in\{1,\dots,n\}$, we denote by $m(z)$, $m(y)$, and $m(x)$ the arithmetic mean of $\{z_1,\dots,z_n\}$, $\{z_1,\dots,z_p\}$, and $\{z_{p+1},\dots,z_{p+q}\}$, respectively, and by s^2 the sample variance of $\{z_1,\dots,z_n\}$. If $p+q<n$, then*

$$ns^2\geqslant\frac{|p(1-\alpha)[m(z)-m(y)]+q(1-\beta)[m(z)-m(x)]|^2}{p\alpha(\alpha-1)+q\beta(\beta-1)},\tag{7}$$

where $\alpha,\beta\in\mathbb{C}$ satisfy $n=p(1-\alpha)+q(1-\beta)$.

Moreover, the inequality (7) is an equality if and only if exist $A,B,C\in\mathbb{C}$ such that $\{A\}=\{z_1,\dots,z_p\}$, $\{B\}=\{z_{p+1},\dots,z_{p+q}\}$, $\{C\}=\{z_{p+q+1},\dots,z_n\}$, and $A(1-\beta)+B(\alpha-1)+C(\beta-\alpha)=0$.

Proof. The symbols $\mathbb{1}_k$ and $\mathbf{0}_k$ will denote the vectors of $\mathbb{R}^k$ all of whose components are 1 and 0, respectively. Set $\mathbf{x}=(z_1,\dots,z_n)$ and $\mathbf{a}=(\alpha\mathbb{1}_p\mid\beta\mathbb{1}_q\mid\mathbb{1}_{n-(p+q)})$. Since $p+q<n$ we have $\mathbf{a}\neq\mathbf{0}_n$. Since $n=p(1-\alpha)+q(1-\beta)$ we get that $\mathbb{1}_n$ and $\mathbf{a}$ are orthogonal. Now we have

$$\begin{aligned}\langle\mathbf{x},\mathbf{a}\rangle&=\alpha(z_1+\cdots+z_p)+\beta(z_{p+1}+\cdots+z_{p+q})+z_{p+q+1}+\cdots+z_n\\&=(\alpha-1)(z_1+\cdots+z_p)+(\beta-1)(z_{p+1}+\cdots+z_{p+q})+z_1+\cdots+z_n\\&=p(\alpha-1)m(y)+q(\beta-1)m(x)+n\cdot m(z)\\&=p(\alpha-1)m(y)+q(\beta-1)m(x)+[p(1-\alpha)+q(1-\beta)]m(z)\\&=p(1-\alpha)[m(z)-m(y)]+q(1-\beta)[m(z)-m(x)]\end{aligned}$$

and

$$\|\mathbf{a}\|^2=p\alpha^2+q\beta^2+n-(p+q)=p\alpha^2+q\beta^2-p\alpha-q\beta=p\alpha(\alpha-1)+q\beta(\beta-1).$$

From Bessel's inequality we get (7).

The inequality (7) becomes an equality if and only if $\mathbf{x} \in \operatorname{span}\{\mathbb{1}_n, \mathbf{a}\}$, i.e., if and only if exist $\lambda, \mu \in \mathbb{C}$ such that $\mathbf{x} = \lambda \mathbb{1}_n + \mu \mathbf{a}$. Therefore, the inequality (7) is an equality if and only if exist $\lambda, \mu \in \mathbb{C}$ such that

$$z_i = \lambda + \mu\alpha \ \ (i = 1, \ldots, p), \qquad z_j = \lambda + \mu\beta \ \ (j = p+1, \ldots, p+q),$$

and

$$z_k = \lambda + \mu \ \ (k = p+q+1, \ldots, n).$$

This is equivalent to say that the sets $\{z_i\}_{i=1}^{p}$, $\{z_j\}_{j=p+1}^{p+q}$, and $\{z_k\}_{k=p+q+1}^{n}$, are singletons, and if we denote by A, B, C their unique elements, respectively, then exist $\lambda, \mu \in \mathbb{C}$ such that $A = \lambda + \mu\alpha$, $B = \lambda + \mu\beta$, $C = \lambda + \mu$ (equivalently, the rank of $\begin{bmatrix} 1 & \alpha & A \\ 1 & \beta & B \\ 1 & 1 & C \end{bmatrix}$ is two). □

REMARK 7. Theorem 3 is a generalization of Theorem 2. If we choose $\alpha = \beta$ or $\alpha = 1$, or $\beta = 1$, in Theorem 3, one gets Theorem 2.

Until now, we have used Bessel's inequality with two terms. We will use this inequality with more terms to study the arithmetic mean of several nested samples.

THEOREM 4. *Let $z_1, \ldots, z_n$ be complex numbers and $p_1, \ldots, p_k \in \{1, \ldots, n\}$ such that $1 \leqslant p_k < \cdots < p_1 < n$. Let m and s^2 denote the arithmetic mean and the sample variance of $z_1, \ldots, z_n$, respectively. If m_i denotes the arithmetic mean of $z_1, \ldots, z_{p_i}$, then*

$$s^2 \geqslant \frac{p_1}{n - p_1}|m - m_1|^2 + \frac{p_1}{n}\frac{p_2}{p_1 - p_2}|m_1 - m_2|^2 + \cdots + \frac{p_{k-1}}{n}\frac{p_k}{p_{k-1} - p_k}|m_{k-1} - m_k|^2.$$

This inequality becomes an equality if and only if $\{z_1, \ldots, z_{p_k}\}$, $\{z_{p_k+1}, \ldots, z_{p_{k-1}}\}$, ..., and $\{z_{p_1+1}, \ldots, z_n\}$ are singletons.

Proof. Define $\mathbf{x} = (z_1, \ldots, z_n)$ and let the symbols $\mathbb{1}_k$ and $\mathbf{0}_k$ have the same meaning as in the proof of Theorem 3. Define the vectors $\mathbf{a}_1 = ((p_1 - n)\mathbb{1}_{p_1} \mid p_1 \mathbb{1}_{n-p_1})$, $\mathbf{a}_2 = ((p_2 - p_1)\mathbb{1}_{p_2} \mid p_2 \mathbb{1}_{p_1 - p_2} \mid \mathbf{0}_{n-p_1})$, $\mathbf{a}_3 = \left((p_3 - p_2)\mathbb{1}_{p_3} \mid p_3 \mathbb{1}_{p_2 - p_3} \mid \mathbf{0}_{n-p_2}\right)$, and so on, until $\mathbf{a}_k$. Now, it is sufficient to apply Bessel's inequality to $\mathbf{x}$ and the subspace spanned by $\{\mathbb{1}_n, \mathbf{a}_1, \ldots, \mathbf{a}_k\}$. □

3. Some applications

We will show some applications of the previous results.

APPLICATION 1. We shall use Theorem 2 to locate the eigenvalues of a complex matrix. Recall that the spectrum of a square matrix A is the set of the eigenvalues of A and usually is denoted by $\sigma(A)$. If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of a matrix and $f : \sigma(A) \to \mathbb{C}$, we denote by $m(f(\lambda))$ the arithmetic mean of $f(\lambda_1), \ldots, f(\lambda_n)$.

In [5, 8] the authors proved that for a definite positive matrix A of order n and $\lambda_0 \in \sigma(A)$, one has (8). We extend this inequality to matrices whose spectrum is contained in $\mathbb{R}$.

THEOREM 5. *Let A be a complex $n\times n$ matrix such that $\sigma(A)\subset\mathbb{R}$. If $\lambda_0\in\sigma(A)$, then*

$$\frac{\operatorname{tr}(A)}{n}-\xi\leqslant\lambda_0\leqslant\frac{\operatorname{tr}(A)}{n}+\xi, \tag{8}$$

where

$$\xi=\frac{\sqrt{(n-1)(n\operatorname{tr}(A^2)-\operatorname{tr}(A)^2)}}{n}.$$

Moreover, $\lambda_0\in\{\operatorname{tr}(A)/n-\xi,\operatorname{tr}(A)/n+\xi\}$ if and only if $\sigma(A)\setminus\{\lambda_0\}$ is a singleton.

Proof. Let $\lambda_1,\dots,\lambda_n$ be the eigenvalues of A, counting algebraic multiplicities. It is known that $\operatorname{tr}(A)=\lambda_1+\cdots+\lambda_n$ and $\operatorname{tr}(A^2)=\lambda_1^2+\cdots+\lambda_n^2$ (this last equality follows from Schur's triangularization theorem). Now it is sufficient to apply Samuelson's inequality to the real data $\lambda_1,\dots,\lambda_n$ and observe that

$$m(\lambda)=\frac{\lambda_1+\cdots+\lambda_n}{n}=\frac{\operatorname{tr}(A)}{n}$$

and if s^2 denotes the sample variance of $\lambda_1,\dots,\lambda_n$, then

$$s^2=m(\lambda^2)-m(\lambda)^2=\frac{\lambda_1^2+\cdots+\lambda_n^2}{n}-\left(\frac{\lambda_1+\cdots+\lambda_n}{n}\right)^2=\frac{\operatorname{tr}(A^2)}{n}-\frac{\operatorname{tr}(A)^2}{n^2}. \quad\square$$

Recall that the spectrum of any Hermitian matrix is always contained in $\mathbb{R}$, which shows that the hypotheses of the previous theorem are not artificial.

Also, Theorem 2 permits to locate the spectrum of any complex matrix. It is known (see e.g., [9, Th. 9.1]) that $\operatorname{tr}(AA^*)\geqslant|\lambda_1|^2+\cdots+|\lambda_n|^2$, where A is a complex $n\times n$ matrix and $\lambda_1,\dots,\lambda_n$ are the eigenvalues of A. Moreover, $\operatorname{tr}(AA^*)=|\lambda_1|^2+\cdots+|\lambda_n|^2$ if and only if A is normal, i.e., $AA^*=A^*A$.

THEOREM 6. *Let A be a complex $n\times n$ matrix. If $\lambda_0\in\sigma(A)$, then*

$$\left|\lambda_0-\frac{\operatorname{tr}(A)}{n}\right|\leqslant\frac{\sqrt{(n-1)(n\operatorname{tr}(AA^*)-|\operatorname{tr}(A)|^2)}}{n}.$$

This inequality becomes an equality if and only if A is normal and $\sigma(A)\setminus\{\lambda_0\}$ is a singleton.

Proof. As in the previous theorem, we have $\operatorname{tr}(A)=n\cdot m(\lambda)$. By the paragraph preceding this theorem, $\operatorname{tr}(AA^*)\geqslant n\cdot m(|\lambda|^2)$. An obvious application of Theorem 2 finishes the proof. $\square$

EXAMPLE 1. Let

$$A=\begin{bmatrix}1&2&4\\2&4&5\\4&5&6\end{bmatrix},\qquad B=\begin{bmatrix}1&0&\mathrm{i}/\sqrt{2}\\0&1&-\mathrm{i}/\sqrt{2}\\\mathrm{i}/\sqrt{2}&-\mathrm{i}/\sqrt{2}&1\end{bmatrix}.$$

Since A is Hermitian, $\sigma(A) \subset \mathbb{R}$. By using Theorem 5, we have that $\sigma(A)$ is contained in $[-4.61, 11.94]$. Observe that by Gershgorin's disk theorem, we have $\sigma(A) \subset [1-6, 1+6] \cup [4-7, 4+7] \cup [6-9, 6+9] = [-5, 15]$, a poorer location than the previous one. By using a numerical software, we get $\sigma(A) = \{-1.40, 0.54, 11.863\}$.

The matrix B is normal. By using Theorem 6, we get $\sigma(B) \subset \{\lambda \in \mathbb{C} : |\lambda - 1| \leqslant 2\sqrt{3}/3\}$. By Gershgorin's disk theorem, we get $\sigma(B) \subset \{\lambda \in \mathbb{C} : |\lambda - 1| \leqslant \sqrt{2}\}$. Since $2\sqrt{3}/3 \simeq 1.154$ and $\sqrt{2} \simeq 1.414$, the bound obtained by Theorem 6 is better than the bound obtained by Gershgorin's disk theorem. A computation shows that $\sigma(B) = \{1, 1+\mathrm{i}, 1-\mathrm{i}\}$.

The spectrum is usually seen as part of the closed disk $D_{r(A)}(0)$ centered at 0, being $r(A)$ the spectral radius of the matrix A. It is known that $r(A) \leqslant \|A\|$, where $\|\cdot\|$ is any matrix norm. We can choose $\|\cdot\|_1$ or $\|\cdot\|_\infty$ in view of the easiness of computing these norms. But for the above examples we obtain $\|A\|_1 = \|A\|_\infty = 15$ and $\|B\|_1 = \|B\|_\infty = 1 + \sqrt{2} \simeq 2.414$. Therefore, $\sigma(A) \subset [-15, 15]$ (recall that A is Hermitian, and thus, $\sigma(A) \in \mathbb{R}$) and $\sigma(B) \subset D_{1+\sqrt{2}}(0)$.

We can improve the bounds of the previous paragraph by using other matrix norms. In particular, if the matrix is normal, then its spectral radius equals the Euclidean norm ($\|\cdot\|_2$) of the matrix. Let us note that the aforementioned examples are normal matrices and we have $\|A\|_2 \simeq 11.863$ and $\|B\|_2 = \sqrt{2} \simeq 1.414$. In general, for a normal matrix X, it must exist $\lambda \in \sigma(X)$ such that $|\lambda| = \|X\|_2$.

However, let us note that the Euclidean norm of a matrix is harder to compute than the bounds obtained by Theorems 5 and 6 since the computation of $\operatorname{tr}(A)$ and $\operatorname{tr}(A^2)$ – or $\operatorname{tr}(AA^*)$ – are simpler than the computation of $\|A\|_2$, since $\|A\|_2$ is the largest singular value of A. Precisely, the purpose of Theorems 5 and 6 is locate the eigenvalues by means of easily computable terms.

The different circles containing the eigenvalues of the matrix B can be seen in Figure 1.

COROLLARY 1. *Let A be a complex $n \times n$ matrix.*

(i) *If $\sigma(A) \subset \mathbb{R}$ and $(n-1)\operatorname{tr}(A^2) < \operatorname{tr}(A)^2$, then A is nonsingular.*

(ii) *If $(n-1)\operatorname{tr}(AA^*) < |\operatorname{tr}(A)|^2$, then A is nonsingular.*

Proof. We will prove the first item, since the proof of the second one is similar. If A is singular, then $0 \in \sigma(A)$. By Theorem 5, $-\xi \leqslant \operatorname{tr}(A)/n \leqslant \xi$, where the meaning of ξ is also given in Theorem 5. Therefore, $\operatorname{tr}(A)^2/n^2 \leqslant \xi^2$, which is equivalent to $\operatorname{tr}(A)^2 \leqslant (n-1)\operatorname{tr}(A^2)$. □

REMARK 8. The reverse implications of the previous corollary are not true. Take the nonsingular matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Observe that $\sigma(A) = \{-1, 1\} \subset \mathbb{R}$.

APPLICATION 2. Next, we will use Theorem 2 to locate the complex roots of a complex polynomial. Laguerre focused ([4]) on n-th degree polynomials with all its

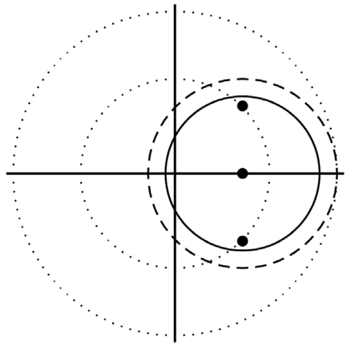

Figure 1: *The location of the eigenvalues of the matrix B. The solid points correspond to the three eigenvalues. The solid circle corresponds to the bound obtained by Theorem 6. The dashed circle corresponds to the bound obtained by Gershgorin's disk theorem. The inner dotted circle corresponds to the bound $r(B) \leqslant \|B\|_2$, and the outer dotted circle corresponds to he bound $r(B) \leqslant \|B\|_1$.*

roots real. Let $t_1,\dots,t_n$ denote the roots, all of which we will assume to be real, of the n-th degree polynomial equation with $n \geqslant 2$:

$$a_0 + a_1 t + a_2 t^2 + \cdots + a_{n-1} t^{n-1} + t^n = 0.$$

Lagerre proved that

$$-\frac{a_{n-1}}{n} - b\sqrt{n-1} \leqslant t_j \leqslant -\frac{a_{n-1}}{n} + b\sqrt{n-1}, \qquad j = 1,2,\dots,n, \tag{9}$$

where

$$b = \sqrt{\frac{(n-1)a_{n-1}^2}{n^2} - \frac{2a_{n-2}}{n}}.$$

In [3] it is shown that $-a_{n-1}/n$ is the arithmetic mean of $t_1,\dots,t_n$ and b is the sample variance of $t_1,\dots,t_n$. Therefore, (9) reduces to Samuelson's inequality.

Given the complex polynomial $p(t) = c_0 + c_1 t + \cdots + c_{n-1} t^{n-1} + t^n$, the matrix

$$C(p) = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{bmatrix}$$

is called the *companion matrix* of the polynomial p. It is known that the eigenvalues of $C(p)$ are the roots of the polynomial p. An application of Theorem 6 permits locate in an easy manner the roots of p.

THEOREM 7. *Let* $p(t) = c_0 + c_1 t + \cdots + c_{n-1} t^{n-1} + t^n$ *be a complex polynomial and let* $z_1, \ldots, z_n$ *be the roots of* $p(t)$*. Then*

$$\left| z_j + \frac{c_{n-1}}{n} \right| \leqslant \frac{1}{n} \sqrt{(n-1)(n\alpha + |c_{n-1}|^2)}, \qquad j = 1, \ldots, n,$$

where $\alpha = n - 1 + \sum_{j=0}^{n-1} |c_j|^2$.

APPLICATION 3. Next we will give some bounds for the eigenvalues of a tensor. Firstly, let us recall some concepts concerning tensors. A tensor can be regarded as a higher order generalization of a matrix, which takes the form

$$\mathscr{A} = (a_{i_1,\ldots,i_m}), \;\; a_{i_1,\ldots,i_m} \in \mathbb{C}, \;\; i_j \in \{1, \ldots, n\}, \;\; j \in \{1, \ldots, m\}.$$

Such a multidimensional array is called an *m-order n-dimensional complex tensor* and the class of m-order n-dimensional real tensors is denoted by $\mathscr{T}(\mathbb{C}^n, m)$. The m-order n-dimensional identity tensor, denoted by $\mathscr{I}$, is the tensor with entries

$$a_{i_1,\ldots,i_m} = \begin{cases} 1 \text{ if } i_1 = \cdots = i_m, \\ 0 \text{ otherwise.} \end{cases}$$

Given a column vector $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$, we define $\mathscr{A}\mathbf{x}^{m-1}$ to be the vector in $\mathbb{C}^n$ whose i-th coordinate is the scalar

$$(\mathscr{A}\mathbf{x}^{m-1})_i = \sum_{i_2,\ldots,i_m=1}^{n} a_{i,i_2,\ldots,i_m} x_{i_2} \cdots x_{i_m}.$$

The following definition, borrowed from Qi [6], extends the classical concept of eigenvalues of square matrices.

DEFINITION 1. Let $\mathscr{A} \in \mathscr{T}(\mathbb{C}^n, m)$, $\mathscr{I} \in \mathscr{T}(\mathbb{C}^n, m)$ be the identity tensor, and $\lambda \in \mathbb{C}$. If exists $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ such that $(\mathscr{A} - \lambda \mathscr{I})\mathbf{x}^{m-1} = \mathbf{0}$, we say that λ is an *eigenvalue* of the tensor $\mathscr{A}$ and $\mathbf{x}$ an *eigenvector* of $\mathscr{A}$ associated with λ.

Let the differential operators $\widehat{g}_i$ be defined by

$$\widehat{g}_i = \sum_{i_2=1}^{n} \cdots \sum_{i_m=1}^{n} a_{i,i_2,\ldots,i_m} \frac{\partial}{\partial a_{i,i_2}} \cdots \frac{\partial}{\partial a_{i,i_m}}, \quad i \in \{1, \ldots, n\},$$

where A is an auxiliary $n \times n$ matrix consists of indeterminate variables a_{ij}'s. Hu et al. [2] defined the d-th order trace of the tensor $\mathscr{A}$ by

$$\operatorname{tr}_d(\mathscr{A}) = (m-1)^{n-1} \left[\sum_{\sum_{i=1}^{n} k_i = d} \prod \frac{(\widehat{g}_i)^{k_i}}{((m-1)k_i)!} \right] \operatorname{tr}(A^{(m-1)d}).$$

Next result can be found in [2, Corollary 6.5].

LEMMA 1. *Let* $\mathscr{A} \in \mathscr{T}(\mathbb{C}^n, m)$. *Then*

(i) $\operatorname{tr}_1(\mathscr{A}) = \sum_{\lambda_i \in \sigma(\mathscr{A})} m_i \lambda_i$,

(ii) $\operatorname{tr}_2(\mathscr{A}) = \sum_{\lambda_i \in \sigma(\mathscr{A})} m_i \lambda_i^2$,

where $\sigma(\mathscr{A})$ *is the set of all the eigenvalues of the tensor* $\mathscr{A}$ *and* m_i *is the algebraic multiplicity of the eigenvalue* λ_i.

And next result can be found in [6, Theorem 1 (b)].

LEMMA 2. *The number of eigenvalues of an* m*-order* n*-dimensional tensor* $\mathscr{A}$ *is* $n(m-1)^{n-1}$.

Now, we give a new estimation of the eigenvalues of a certain class of complex tensors by using the first order and second order trace of $\mathscr{A}$.

THEOREM 8. *Let* $\mathscr{A} \in \mathscr{T}(\mathbb{C}^n, m)$. *If* $\sigma(\mathscr{A}) \subset \mathbb{R}$ *and* $\lambda \in \sigma(\mathscr{A})$, *then*

$$\left|\frac{\operatorname{tr}_1(\mathscr{A})}{N} - \lambda\right| \leqslant (N-1)\left[\frac{\operatorname{tr}_2(\mathscr{A})}{N} - \frac{\operatorname{tr}_1(\mathscr{A})^2}{N^2}\right], \tag{10}$$

where $N = n(m-1)^{n-1}$.

Proof. Let $\lambda_1, \ldots, \lambda_N$ be the eigenvalues of $\mathscr{A}$ and let s^2 be the sample variance of these eigenvalues. By Lemma 1 and Lemma 2, we have

$$s^2 = \frac{\operatorname{tr}_2(\mathscr{A})}{N} - \frac{\operatorname{tr}_1(\mathscr{A})^2}{N^2}.$$

Applying Theorem 2 to $\lambda_1, \ldots \lambda_N$, we get (10). □

Acknowledgement. This work was supported by Hunan Provincial Innovation Foundation For Postgraduate under Grant [number CX2016B073]. The first author is thankful to China Scholarship Council for giving him a grant to have a further study in Universidad Politécnica de Valencia, Spain.

REFERENCES

[1] A. D. R. CHOUDARY, C. P. NICULESCU, *Real Analysis on Intervals*, Springer, New Delhi, 2014.

[2] S. HU, Z. HUANG, C. LING, L. QI, *On determinants and eigenvalue theory of tensors*, J. Symbolic Comput., **50**, (2013), 508–531.

[3] S. T. JENSEN, *The Laguerre-Samuelson inequality with extensions and applications in statistics and matrix theory*, M. Sc. Thesis.

[4] H. LAGUERRE, *Sur une méthod pour obtenir par approximation les racines d'une équation algébrique qui a toutes ses racines réeles* [in French]. Nouvelles Annales de Mathématiques (Paris), 2e Série 19 (1880), 161–171 & 193–202 [JFM 12:71]. [Reprinted in Œuvres de Laguerre (publiées sous les auspices de l'Académie des Sciences par MM. Ch. Hermite, H. Poincaré et E. Rouché), Gauthier-Villars, Paris, vol. 1, pp. 87–103 (1898)].

[5] A. D. MADEN, S. BÜYÜKKÖSE, *Bounds for Laplacian Graph Eigenvalues*, Math. Inequal. Appl., **15**, (2012), 529–536.

[6] L. QI, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Computation, **40**, (2005), 1302–1324.

[7] P. A. SAMUELSON, *How deviant can you be?*, J. Amer. Statist. Asoc., **63**, (1968), 1522–1525.

[8] D. TASCI, S. BÜYÜKKÖSE, *Bounds for singular values using traces of matrices*, Selçuk University Art and Science Faculty, Journal of Science, Konya, **17**, (2000), 137–146.
[9] F. ZHANG, *Matrix Theory, Basic Results and Techniques* (2nd edition), Springer.

(Received December 7, 2015)

Hongwei Jin
College of Mathematics and Econometrics, Hunan University
410082, Changsha, P. R. China
and
Instituto de Matemática Multidisciplinar
Universidad Politécnica de Valencia
Camino de Vera s/n, 46022, Valencia, Spain
e-mail: hw-jin@hotmail.com

Julio Benítez
Instituto de Matemática Multidisciplinar
Universidad Politécnica de Valencia
Camino de Vera s/n, 46022, Valencia, Spain
e-mail: jbenitez@mat.upv.es