Bounds on elementary symmetric functions

Andrew L. Rukhin

National Institute of Standards and Technology, 100 Bureau Dr., Gaithersburg, MD 20899, USA

ARTICLE INFO

Article history:
Received 20 December 2013
Accepted 8 January 2014
Available online 31 January 2014
Submitted by A. Böttcher

MSC:
26C10
26E60
26D05

Keywords:
Discriminant
Hankel matrices
Moments
Newton identities
Vandermonde matrix

ABSTRACT

Tight bounds on an elementary symmetric function are established under the assumption that the values of the elementary symmetric functions of lower orders are given. The explicit form of the inverse Hankel moment matrix leads to inequalities for moments and for elementary symmetric polynomials.

1. Setting of the problem

The paper concerns a bound that arises in a statistical meta-analysis model [12]. Specifically, let \( S = \{s_1, \ldots, s_n\} \) be a given set of real numbers, and let \( E_m = E_m(S) \) be the \( m \)-th elementary symmetric polynomial in \( s_1, \ldots, s_n \); i.e., \( E_m \) is the coefficient of \( x^{n-m} \) in the polynomial \( P(x) = \prod_{i=1}^{n} (x + s_i) \). The problem in [12] (in the case of positive distinct \( s_i \)'s) consists in obtaining tight bounds on \( E_n \) if all other elementary symmetric...
functions $E_1, \ldots, E_{n-1}$ are given. We consider the somewhat more general problem of obtaining bounds on $E_k$, $2 \leq k \leq n$, for fixed values $E_1, \ldots, E_{k-1}$.

An equivalent formulation of this problem is as follows. Let the monic polynomial $P(x)$ have only real roots $-s_1, \ldots, -s_n$. Given the roots $-t_1, \ldots, -t_{n-1}$ of the derivative, $P'(x) = \frac{d}{dx} \prod_1^n (x + s_i) = n \prod_1^{n-1} (x + t_j)$, establish the range of possible values $P(0) = E_n(S)$. Since for $0 \leq k \leq n - 1$, $(n - k)E_k(S) = nE_k(t_1, \ldots, t_{n-1})$, elementary symmetric functions of $t_1, \ldots, t_{n-1}$ determine $E_1 = E_1(S), \ldots, E_{n-1} = E_{n-1}(S)$.

There are some general results on the extremal values of linear combinations of elementary symmetric functions over the real variety $\{E_1 = e_1, \ldots, E_{k-1} = e_{k-1}\}$ [7]. For example each of the points where $E_k$ attains local extrema has at most $k$ different coordinates.

Our approach is based on the well known Newton identities [6] relating the elementary symmetric functions to the classical power sums,

$$M_k = \sum_i s_i^k, \quad k = 1, \ldots, n.$$ 

Indeed the functions $E_1, \ldots, E_k$ define $M_1, \ldots, M_k$ and vice versa. For example,

$$M_k = \det \begin{pmatrix} E_1 & 1 & 0 & \cdots & 0 \\ 2E_2 & E_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ kE_k & E_{k-1} & E_{k-2} & \cdots & E_1 \end{pmatrix} = \sum_{j_1+2j_2+\cdots+kj_k = k} \nu_{j_1\ldots j_k}^{(k)} E_1^{j_1} \cdots E_k^{j_k}, \quad (1)$$

with integer coefficients $\nu_{j_1\ldots j_k}^{(k)}$. In particular, $\nu_{00\ldots 01}^{(k)} = (-1)^{k+1}$. Thus our problem reduces to that of specifying the range of $M_k$ for the given values $M_1, \ldots, M_{k-1}$. The latter problem has a known solution given in terms of canonical moments [3, Section 1.4]. Indeed the theory of canonical moments provides bounds on the $k$-th moment of a measure on an interval $[a, b]$ when the first $k - 1$ moments are given.

The next section takes advantage of positive definiteness of Hankel matrices and provides the inequalities for moments and for elementary symmetric polynomials. Some examples are discussed in Section 3. A lower bound for weighted moments is given in Section 4. Numerical comparison with the known inequalities in Section 5 concludes this paper.
2. Main result

Let \( a = \min s_i < \max s_i = b \), and define the \textit{Hankel} matrices,

\[
H_{2m} = \{ M_{k+\ell} \}_{k,\ell=0}^{m-1}, \quad 1 \leq m \leq n/2, \\
H_{2m+1} = \{ M_{k+\ell+1} - aM_{k+\ell} \}_{k,\ell=0}^{m-1}, \quad 1 \leq m \leq (n-1)/2, \\
H_{2m} = \{ -abM_{k+\ell} + (a + b)M_{k+\ell+1} - M_{k+\ell+2} \}_{k,\ell=0}^{m-2}, \quad 1 \leq m \leq n/2, \\
H_{2m+1} = \{ bM_{k+\ell} - M_{k+\ell+1} \}_{k,\ell=0}^{m-1}, \quad 1 \leq m \leq (n-1)/2.
\]

Note that \( H_{2m} \) is of order \( m - 1 \), but the other three matrices are of order \( m \). It is assumed that \( H_1 = 1, H_1 = 0 \).

According to the well known solution of the Hausdorff moment problem, see e.g., [3, Theorem 1.4.3], all matrices above are nonnegative definite. They are positive definite under the conditions of the following Theorem 1.

Put

\[
h_{2m} = (M_m, \ldots, M_{2m-1})^T, \\
h_{2m+1} = (M_{m+1} - aM_m, \ldots, M_{2m} - aM_{2m-1})^T, \\
\]

and

\[
\bar{h}_{2m+1} = (bM_m - M_{m+1}, \ldots, bM_{2m-1} - M_{2m})^T,
\]

which are column-vectors of dimension \( m \). The last needed vector, \( \bar{h}_{2m} = (-abM_{m-1} + (a + b)M_m - M_{m+1}, \ldots, -abM_{2m-3} + (a + b)M_{2m-2} - M_{2m-1})^T \), has dimension \( m - 1 \).

If \( I_n = \{i_1, \ldots, i_n\} \), \( m \leq n \), is a subset of \( \{1, \ldots, n\} \), then \( S_{I_m} \) will represent the corresponding subset of \( S \), \( S_{I_m} = \{s_{i_1}, \ldots, s_{i_m}\} \). Let \( V(S_{I_m}) = \prod_{i,j \in I_m, i > j} (s_i - s_j)^2 \) denote the \textit{discriminant} corresponding to this subset of \( S \). We put \( V^a(S_{I_m}) = V(S_{I_m}) \prod_{j \in I_m} (s_j - a)(b - s_j) \) for notational convenience.

**Theorem 1.** If the number of distinct \( s \)-values is at least \( m \) then the first inequality in (2) holds, while the second is valid provided there are at least \( m - 1 \) distinct \( s \)-values none of which is equal to \( a \) or \( b \),

\[
\frac{\sum_{I_m} V(S_{I_m-1}) \sum_{i} s_i^{m-1} \prod_{j \in I_m-1} (s_i - s_j)^2}{\sum_{I_n} V(S_{I_m})} \leq M_{2m} \\
\leq (a + b)M_{2m-1} - abM_{2m-2} \\
\leq \frac{\sum_{I_m-2} V^b(S_{I_m-2}) \sum_{i} s_i^{m-2} (s_i - a)(b - s_i) \prod_{j \in I_m-2} (s_i - s_j)^2}{\sum_{I_m-1} V^b(S_{I_m-1})}.
\]

(2)
Similarly, in the inequality

\[
\sum_{I_m^{-1}} V(S_{I_m^{-1}}) \prod_{j \in I_m^{-1}} (s_j - a) \left[ \sum_{i} s_i^m (s_i - a) \prod_{j \in I_m} (s_i - s_j) \right]^2 
\leq M_{2m+1} 
\leq bM_{2m} - \sum_{I_m^{-1}} V(S_{I_m^{-1}}) \prod_{j \in I_m^{-1}} (b - s_j) \left[ \sum_{i} s_i^m (b - s_i) \prod_{j \in I_m} (s_i - s_j) \right]^2 
\sum_{I_m} V(S_{I_m}) \prod_{j \in I_m} (b - s_j),
\] (3)

the left-hand bound holds if the number of distinct \(s\)-values different from \(a\) is at least \(m\), with \(b\) replacing \(a\) in the condition for the right-hand bound validity.

**Proof.** If \(H_{2m}\) is a nonsingular matrix, i.e., if \(\det(H_{2m}) > 0\),

\[
h_{2m}^T H_{2m}^{-1} h_{2m} \leq M_{2m} \leq (a + b)M_{2m-1} - abM_{2m-2} - \frac{h_{2m}^T \Pi_{2m}^{-1} h_{2m}}{M_{2m}^2 + \ldots + M_{2m-2}^2},
\] (4)

Since

\[
\det(H_{2m+2}) = \det(H_{2m}) \left( M_{2m} - h_{2m}^T H_{2m}^{-1} h_{2m} \right),
\]

the first inequality in (8) follows. The second inequality holds provided that \(\det(H_{2m}) > 0\), as

\[
\det(H_{2m+2}) = \det(H_{2m}) \left( -abM_{2m-2} + (a + b)M_{2m-1} - M_{2m} - \frac{h_{2m}^T \Pi_{2m}^{-1} h_{2m}}{M_{2m}^2 + \ldots + M_{2m-2}^2} \right).
\]

Similarly, if the determinants of \(H_{2m+1}\) and \(\Pi_{2m+1}\) are strictly positive,

\[
aM_{2m} + \frac{h_{2m+1}^T H_{2m+1}^{-1} h_{2m+1}}{M_{2m+1}^2 + \ldots + M_{2m+2}^2} \leq M_{2m+1} \leq bM_{2m} - \frac{h_{2m+1}^T \Pi_{2m+1}^{-1} h_{2m+1}}{M_{2m+1}^2 + \ldots + M_{2m+2}^2}.
\] (5)

Now we evaluate \(H_{2m}^{-1}\), as well as the inverses of other related matrices starting with their determinants.

Consider the Vandermonde-type \(m \times n\) matrix

\[
V = V_{mn} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ s_1^{m-1} & \cdots & s_n^{m-1} \end{pmatrix}.
\]

As is well known and easy to check,

\[
H_{2m} = V V^T,
\]

so that by the Binet–Cauchy formula,

\[
\det(H_{2m}) = \sum_{I_m} V(S_{I_m}),
\] (6)
Indeed in our notation the determinant of the $m \times m$ submatrix of $V$ corresponding to the set $I_m$ is $\prod_{i,j \in I_m, i \geq j} (s_i - s_j)$. According to (6), $H_{2m}$ is a singular matrix if and only if for any $m, \prod_{i,j \in I_m, i \geq j} (s_i - s_j)^2 = 0$, which means that the number of distinct $s$-values is less than $m$.

Similar formulas hold for $\det(H_{2m})$, $\det(H_{2m+1})$ and $\det(H_{2m+1})$. For example,

$$H_{2m} = V_{m-1} W V_{m-1}^T$$

with the diagonal matrix $W = \text{diag}((s_1 - a)(b - s_1), \ldots, (s_n - a)(b - s_n))$. Thus

$$\det(H_{2m}) = \sum_{I_{m-1}} V_a^{S_{I_{m-1}}} > 0$$

when and only when there are at least $m - 1$ distinct $s$-values none of which is equal to $a$ or $b$.

For $H_{2m+1}$ the diagonal matrix $W$ is to be taken as $\text{diag}(s_1 - a, \ldots, s_n - a)$, and for $H_{2m+1}$, $W = \text{diag}(b - s_1, \ldots, b - s_n)$. In the first case a necessary and sufficient condition for positivity of $\det(H_{2m+1}) = \prod_{I_m} V(S_{I_m}) \prod_{j \in I_m} (s_j - a)$ is that the number of distinct $s$-values exceeding $a$ is at least $m$. In the second case the diagonal matrix $W$ is $\text{diag}(b - s_1, \ldots, b - s_n)$, and $\det(H_{2m+1}) > 0$ if and only if there are $m$ or more distinct $s$-values which are strictly smaller than $b$.

Thus the conditions of Theorem 1 guarantee that the considered matrices are positive definite and the inequalities (4) and (5) are valid. When the number of distinct $s$-values is exactly $m$, so that $\det(H_{2m+2}) = 0$, but $\det(H_{2m}) > 0$, the lower inequality in (4) reduces to an equality. Similar numbers of distinct $s$-values in the open interval $(a, b)$ (or in a semi-closed interval with the end points $a, b$) show that the remaining inequalities are also sharp.

To find $H_{2m}^{-1}$ we determine the adjugate matrix, $\text{adj}(H_{2m})$, via its elements $\text{adj}(H_{2m})_{k\ell}, 0 \leq k, \ell \leq m - 1$. According to the already used Binet–Cauchy theorem,

$$\text{adj}(H_{2m})_{k\ell} = (-1)^{k+\ell} \sum_{I_{m-1}} \det \left( \begin{array}{ccc} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ s_{i_1}^{m-1} & \cdots & s_{i_m}^{m-1} \end{array} \right) \det \left( \begin{array}{ccc} 1 & \cdots & s_{i_1}^{m-1} \\ \vdots & \ddots & \vdots \\ 1 & \cdots & s_{i_m}^{m-1} \end{array} \right),$$

where the $k$-th row of the first matrix and the $\ell$-th column of the second matrix in the right-hand side of this formula are deleted.

The first determinant is known to be $E_{m-1-k}(S_{I_{m-1}}) \prod_{i,j \in I_{m-1}, i \geq j} (s_i - s_j)$, and the second is $E_{m-1-\ell}(S_{I_{m-1}}) \prod_{i,j \in I_{m-1}, i \geq j} (s_i - s_j)$ [2, p. 36]. Therefore,

$$\text{adj}(H_{2m})_{k\ell} = (-1)^{k+\ell} \sum_{I_{m-1}} E_{m-1-k}(S_{I_{m-1}}) E_{m-1-\ell}(S_{I_{m-1}}) V(S_{I_{m-1}}).$$

Since for any fixed $j$,

$$\sum_k (-1)^k s_j^k E_{m-1-k}(S_{I_{m-1}}) = \prod_{i \in I_{m-1}} (s_i - s_j),$$
one has

\[ h_{2m}^T \text{adj}(H_{2m}) h_{2m} \]

= \sum_{l_m} \sum_{k,\ell} (-1)^{k+\ell} E_{m-1-k} (S_{l_m-1}) E_{m-1-\ell} (S_{l_m-1}) V(S_{l_m-1}) M_{m+k} M_{m+\ell}

= \sum_{l_m} V(S_{l_m-1}) \sum_s m_s \prod_{j \in l_m-1} (s_i - s_j) (s_j - s_j)

= \sum_{l_m} V(S_{l_m-1}) \left[ \sum_i s_i^m \prod_{j \in l_m-1} (s_i - s_j) \right]^2.

Of course the last summation can be performed only for \( i \notin I_{m-1} \).

The remaining formulas needed to establish (2) and (3) are proven similarly. \( \square \)

The bounds (2) lead to the range of values for \( E_k \) that can be expressed in terms of \( E_1, \ldots, E_{k-1}, a \) and \( b \). We formulate the corresponding inequalities according to the parity of \( k \). For even \( k \) the upper bound does not involve \( a \) or \( b \) while the lower bound depends on these quantities in a symmetric fashion. If \( k = 2m + 1 \), the lower bound (2) can be obtained from the upper bound with \( b \) replacing \( a \).

**Theorem 2.** Under the conditions of Theorem 1 ensuring (2), the following inequalities for \( E_{2m} \) are valid:

\[
\sum_{\ell_0+\ell_1+k_1+\cdots+(2m-1)k_{2m} = m(m+1)} \beta^{(2m)+1}_{\ell_0+\ell_1+k_1+\cdots+(2m-1)k_{2m}} a^{\ell_0} b^{\ell_1} E_1^{k_1} \cdots E_{2m-2}^{k_{2m-2}} \\
\leq 2m E_{2m} \leq \sum_{\ell_0+\ell_1+k_1+\cdots+(2m-1)k_{2m} = m(m+1)} \gamma^{(2m)+1}_{\ell_0+\ell_1+k_1+\cdots+(2m-1)k_{2m}} a^{\ell_0} b^{\ell_1} E_1^{k_1} \cdots E_{2m-2}^{k_{2m-2}}.
\]

(8)

The integer coefficients \( \beta^{(2m)}_{\ell_0+\ell_1+k_1+\cdots+k_{2m}} \), \( 1 \leq m \leq n/2 \), are defined by (11) with \( \beta^{(2m)}_{\ell_0+\ell_1+k_1+\cdots+k_{2m}} \) satisfying the recurrent formula (12). The integer coefficients \( \gamma^{(2m)}_{\ell_0+\ell_1+k_1+\cdots+k_{2m}} \) are defined by (9) with \( \gamma^{(2m)}_{\ell_0+\ell_1+k_1+\cdots+k_{2m}} \) satisfying (10).

**Proof.** Clearly \( \det(H_{2m}) \), \( m \leq n/2 \), is a symmetric homogeneous polynomial in the variables \( s_1, \ldots, s_n \) of degree \( m(m-1) \). Therefore,

\[
\sum_{l_m} V(S_{l_m}) = \sum_{k_1+\cdots+(2m-2)k_{2m-2} = m(m-1)} \gamma^{(2m-2)}_{k_1+\cdots+k_{2m-2}} E_1^{k_1} \cdots E_{2m-2}^{k_{2m-2}},
\]

(9)

with integer coefficients \( \gamma^{(2m-2)}_{k_1+\cdots+k_{2m-2}} \).

The leading term in lexicographic order is \( s_1^{2m-2} s_2^{2m-4} \cdots s_{m-1}^2 \) with the coefficient \( n - m + 1 \), so that \( \gamma^{(2m-2)}_{2\cdots2m\cdots0} = n - m + 1 \). Because of (1),
$$\sum_{I_{m+1}} \mathcal{V}(S_{I_{m+1}}) = -2mE_{2m} \sum_{I_{m}} \mathcal{V}(S_{I_{m}})$$

$$+ \sum_{k_1 + \cdots + (2m-1)k_{2m-1} = m(m+1)} \gamma_{k_1 \cdots k_{2m-1}}^{(2m)} \beta_{k_1 \cdots k_{2m-1}}^{0} E_{1}^{k_{1}} \cdots E_{2m-1}^{k_{2m-1}},$$

leading to the recurrence,

$$\gamma_{k_1 \cdots k_{2m-2}01}^{(2m)} = -2m\gamma_{k_1 \cdots k_{2m-2}}^{(2m-2)}. \quad (10)$$

One has

$$2mE_{2m} \leq \frac{\sum_{I_{m+1}} \mathcal{V}(S_{I_{m+1}}) + 2mE_{2m} \sum_{I_{m}} \mathcal{V}(S_{I_{m}})}{\sum_{I_{m}} \mathcal{V}(S_{I_{m}})},$$

$$= \frac{\sum_{k_1 + \cdots + (2m-1)k_{2m-1} = m(m+1)} \gamma_{k_1 \cdots k_{2m-1}}^{(2m)} \beta_{k_1 \cdots k_{2m-1}}^{0} E_{1}^{k_{1}} \cdots E_{2m-1}^{k_{2m-1}}}{\sum_{I_{m}} \mathcal{V}(S_{I_{m}})},$$

which establishes the second inequality in (8) not involving \(a\) or \(b\).

To prove the inequality from below (which symmetrically depends on \(a\) and \(b\)), one can use a similar representation,

$$\sum_{I_{m-1}} \mathcal{V}_{a}^{b}(S_{I_{m-1}})$$

$$= \sum_{\ell_{0} + \ell_{1} + k_{1} + \cdots + (2m-2)k_{2m-2} = m(m-1)} \beta_{\ell_{0} \ell_{1} k_{1} \cdots k_{2m-2}}^{(2m)} a_{\ell_{0}} b_{\ell_{1}} E_{1}^{k_{1}} \cdots E_{2m-1}^{k_{2m-2}},$$

$$= 2(m-1)E_{2m-2} \sum_{I_{m-2}} \mathcal{V}_{a}^{b}(S_{I_{m-2}})$$

$$+ \sum_{\ell_{0} + \ell_{1} + k_{1} + \cdots + (2m-3)k_{2m-3} = m(m-1)} \beta_{\ell_{0} \ell_{1} k_{1} \cdots k_{2m-3}}^{(2m)} a_{\ell_{0}} b_{\ell_{1}} E_{1}^{k_{1}} \cdots E_{2m-3}^{k_{2m-3}}. \quad (11)$$

Then \(\beta_{\ell_{0} \ell_{1} k_{1} \cdots k_{2m-2}}^{(2m)} = \beta_{\ell_{0} k_{1} \cdots k_{2m-2}}^{(2m)}, \beta_{\ell_{0} \ell_{1} k_{1} \cdots k_{2m-2}}^{(2m)} = 0\) if \(\ell_{0} + \ell_{1} > m\), and

$$\beta_{\ell_{0} \ell_{1} k_{1} \cdots k_{2m-4}01}^{(2m)} = 2(m-1)\beta_{\ell_{0} \ell_{1} k_{1} \cdots k_{2m-4}}^{(2m-2)}. \quad (12)$$

The same argument demonstrates the validity of the first inequality in (8). \(\square\)

To formulate the inequality for odd values of \(m\), define integer coefficients \(a_{k_0 k_1 \cdots k_{2m-1}}^{(2m)}\) via the representation

$$\sum_{I_{m}} \mathcal{V}(S_{I_{m}}) \prod_{j \in I_{m}} (s_{j} - a)$$

$$= \sum_{k_{0} + k_{1} + \cdots + (2m-1)k_{2m-1} = m^{2}} a_{k_0 k_1 \cdots k_{2m-1}}^{(2m)} a_{k_0} E_{1}^{k_{1}} \cdots E_{2m-1}^{k_{2m-1}}. \quad (13)$$
It follows that
\[
\sum_{I_m} \mathcal{V}(S_{I_m}) \prod_{j \in I_m} (b - s_j) = (-1)^m \sum_{k_0 + k_1 + \ldots + (2m-1)k_{2m-1} = m^2} \alpha^{(2m)}_{k_0k_1\ldots k_{2m-1}} b^{k_0} E_1^{k_1} \ldots E_{2m-1}^{k_{2m-1}},
\]
and similarly to (10) or (12),
\[
\alpha^{(2m)}_{k_0k_1\ldots k_{2m-1}01} = (2m - 1)\alpha^{(2m-2)}_{k_0k_1\ldots k_{2m-3}1}.
\]
As a polynomial in \(a\), (13) has the degree \(m\), so that \(\alpha^{(2m)}_{k_0k_1\ldots k_{2m-1}} = 0\) if \(k_0 > m\).

**Theorem 3.** Under the conditions of Theorem 1 which guarantee that (3) holds,

\[
\sum_{k_0 + k_1 + \ldots + 2mk_{2m} = (m+1)^2} \alpha^{(2m+2)}_{k_0k_1\ldots k_{2m}} a^{k_0} E_1^{k_1} \ldots E_{2m}^{k_{2m}}
\]

\[
\sum_{k_0 + k_1 + \ldots + (2m-1)k_{2m-1} = m^2} \alpha^{(2m)}_{k_0k_1\ldots k_{2m-1}} a^{k_0} E_1^{k_1} \ldots E_{2m-1}^{k_{2m-1}}
\]

\[
\leq (2m + 1)E_{2m+1} \leq \frac{\sum_{k_0 + k_1 + \ldots + 2mk_{2m} = (m+1)^2} \alpha^{(2m+2)}_{k_0k_1\ldots k_{2m}} b^{k_0} E_1^{k_1} \ldots E_{2m}^{k_{2m}}}{\sum_{k_0 + k_1 + \ldots + (2m-1)k_{2m-1} = m^2} \alpha^{(2m)}_{k_0k_1\ldots k_{2m-1}} b^{k_0} E_1^{k_1} \ldots E_{2m-1}^{k_{2m-1}}} (14)
\]

with integer coefficients \(\alpha^{(2m)}_{k_0k_1\ldots k_{2m-1}}\), \(1 \leq m \leq n/2\), which are defined by (13) and which do not depend on \(a\) or \(b\).

Thus the upper bound in (14) is formally obtained from the lower bound if \(a\) is replaced by \(b\) and vice versa. The proof of Theorem 3 is omitted.

According to (4) and (5), the bounds for \(E_m\) are tight since for the indicated number of distinct \(s\)-values (different from \(a\) and/or \(b\)), (8) or (14) reduce to identities. The bounds in Theorem 3 are valid for any \(a < \min s_i\) and \(b > \max s_i\), but then they are weaker than (14).

3. Examples

When \(m = 1\), we obtain the bounds for \(M_2\),

\[
n^{-1}M^2_1 \leq M_2 \leq (a + b)M_1 - nab,
\]
or

\[
n\left[E_1^2 - (a + b)E_1 + nab\right] \leq 2nE_2 \leq (n - 1)E_1^2.
\]

The bounds for \(M_3\) are

\[
aM_2 + \frac{(M_2 - aM_1)^2}{M_1 - na} \leq M_3 \leq bM_2 - \frac{(bM_1 - M_2)^2}{nb - M_1},
\]
which shows that
\[
\frac{4E_2^2 - E_1^2E_2 + (n-1)aE_1^3 - (3n-2)aE_1E_2 - (n-1)a^2E_1^2 + 2na^2E_2}{3(E_1 - na)} \leq E_3 \leq \frac{E_1^2E_2 - 4E_2^3 - (n-1)bE_1^3 + (3n-2)bE_1E_2 + (n-1)b^2E_1^2 - 2nb^2E_2}{3(nb - E_1)}.
\]
For \(M_4\), (2) becomes
\[
\frac{nM_3^2 + M_2^3 - 2M_1M_2M_3}{nM_2 - M_1^2} \leq M_4 \leq (a + b)M_3 - abM_2 - \frac{(abM_1 - (a + b)M_2 + M_3)^2}{|abM_0 - (a + b)M_1 + M_2|}
\]
or
\[
|nab - (a + b)E_1 + E_1^2 - 2E_2|^{-1}[-(n-1)abE_1^4 + (n-1)ab(a + b)E_1^3 + (a + b)E_1^3E_2 + 2E_1^3E_3 - (n-1)a^2b^2E_1^2 - (a^2 + b^2 - 4(n-1)ab)E_1^2E_2 - E_1^2E_2^2 + (a + b)E_1^2E_3 - (3n-2)ab(a + b)E_1E_2 - 4(a + b)E_1E_2^2 - 10E_1E_2E_3 - (3a^2 + 3b^2 + 4nab)E_1E_3 + 2na^2b^2E_2 + [4a^2 + 4b^2 - 2(n-2)ab]E_2^2 + 6(a + b)E_2E_3 + 4E_3^2 + 3nab(a + b)E_3 + 9E_3^2] \leq 4E_4 \leq \frac{-2(n-1)E_1^2E_3 + (n-2)E_1^2E_2^2 + (10n - 12)E_1E_2E_3 - 4(n-2)E_2^3 - 9nE_3^2}{(n-1)E_1^2 - 2nE_2}.
\]
For example, the denominator of the upper inequality follows from the identity
\[
\sum_{I_2} \mathcal{V}(S_{I_2}) = \sum_{i>j} (s_i - s_j)^2 = (n-1)E_1^2 - 2nE_2,
\]
and its numerator can be verified by putting \(E_4 = 0\) in the formula
\[
\sum_{I_3} \mathcal{V}(S_{I_3}) = -2(n-1)E_1^3E_3 + (n-2)E_1^2E_2^2 - 4(n-1)E_1^2E_4 + (10n - 12)E_1E_2E_3 + 8nE_2E_4 - 4(n-2)E_2^3 - 9nE_3^2.
\]

4. Weighted moments

Here we give a lower bound in the spirit of (2) for weighted moments,
\[
\mu_m = \sum_i w_is_i^m, \quad m = 1, \ldots, n,
\]
where \(w_i \geq 0\).
The key facts are that the $m \times m$ symmetric Hankel matrix

$$H_{2m} = \begin{pmatrix} \mu_0 & \cdots & \mu_{m-1} \\ \vdots & \ddots & \vdots \\ \mu_{m-1} & \cdots & \mu_{2m-2} \end{pmatrix},$$

is nonnegative definite, and that it admits a factorization like (7) with the diagonal matrix $W = \text{diag}(w_1, \ldots, w_n)$.

Since inverting general Hankel matrices is difficult both theoretically [5] and numerically [14], we formulate the following result whose proof is similar to that of Theorem 1.

**Lemma 1.** In the notation of Section 2, the matrix $H_{2m}$ has the determinant

$$\det(H_{2m}) = \sum_{I_m} V(S_{I_m}) \prod_{j \in I_m} w_j,$$

which is positive if and only if there are $m$ distinct values $s_{i_1}, \ldots, s_{i_m}$ such that $\prod_{1}^{m} w_{i_k} > 0$. The entries of its inverse are

$$(H_{2m}^{-1})^k_\ell = \frac{(-1)^{k+\ell} \sum_{I_{m-1}} E_{m-1-k}(S_{I_{m-1}}) E_{m-1-\ell}(S_{I_{m-1}}) V(S_{I_{m-1}}) \prod_{j \in I_{m-1}} w_j}{\det(H_{2m})},$$

$k, \ell = 0, \ldots, m - 1$. One has with $h_{2m} = (\mu_m, \ldots, \mu_{2m-1})^T$,

$$h_{2m}^T H_{2m}^{-1} h_{2m} = \frac{\sum_{I_{m-1}} V(S_{I_{m-1}}) \prod_{j \in I_{m-1}} w_j [\sum_i w_i s_i^m \prod_{j \in I_{m-1}} (s_i - s_j)]^2}{\sum_{I_m} V(S_{I_m}) \prod_{j \in I_m} w_j}.$$

Since

$$\det(H_{2m+2}) = \det(H_{2m})(\mu_{2m} - h_{2m}^T H_{2m}^{-1} h_{2m}),$$

the desired inequality follows from Lemma 1,

$$\mu_{2m} \geq \frac{\sum_{I_{m-1}} V(S_{I_{m-1}}) \prod_{j \in I_{m-1}} w_j [\sum_i w_i s_i^m \prod_{j \in I_{m-1}} (s_i - s_j)]^2}{\sum_{I_m} V(S_{I_m}) \prod_{j \in I_m} w_j} \tag{15}$$

provided that $\det(H_{2m}) > 0$.

When all $s_i$’s are positive, one can take $w_i = s_i^r$ to get from (15) a lower bound for the moment of any order, $M_{2m+r} = \mu_{2m}$.

Simic [13] obtained an extension of Newton’s classical inequality, $E_{m-2} E_m \leq (m-1) \times (n-m+1) E_{m-1}^2 / [m(n-m+2)]$ [1], for weighted combinations of elementary symmetric polynomials, $E_m^{(c)} = \sum_i w_i E_m(S \setminus \{s_i\})$, $\sum w_i = 1$. The inequality (15) can be used to get an extension of Theorem 1 for such combinations if all $s_i$’s are positive.
5. Numerical comparisons

The classical nature of elementary symmetric polynomials led to a body of work related to their interrelationship beyond the classical Newton–Maclaurin inequalities (see e.g. [7,8]). This section contains some numerical comparisons with the latest work in [4,9–11].

Namely, we here present the results of numerical comparison of the accuracy of inequalities (8) and (14) as given in Section 2 for $E_4$, $n = 4$, against the following bounds:

1. The classical Newton–Maclaurin bound [1],

$$E_4 \leq \frac{3E_3^2}{8E_2}. \quad (16)$$

2. The Pierce–Foregger–Li bound [4,7],

$$E_4 \geq \frac{E_1E_3}{4} - \frac{E_2E_1^2}{32}. \quad (17)$$

3. The Rosset bounds [11],

$$E_4^2 - E_4 \left( \frac{3(n - 3)E_2E_3}{(n - 1)E_1} - \frac{4(n - 2)(n - 3)E_3^2}{3(n - 1)^2E_1^2} \right) + \frac{3(n - 3)^2E_3^3}{(n - 1)(n - 2)E_1} - \frac{3(n - 3)^2E_3^2E_2^2}{4(n - 1)^2E_1^2} \leq 0. \quad (18)$$

This quadratic in $E_4$ inequality (18) delivers both an upper bound and a lower bound.

4. One of Niculescu’s bounds [10, p. 8],

$$E_4 \leq \frac{1}{4} \left( \frac{3E_4^4}{16} - E_1^2E_2 + E_1E_3 + E_2^2 \right). \quad (19)$$

5. One of Mitev’s bounds [9, p. 8],

$$E_4 \leq \frac{1}{16} \left( E_1^4 - 4E_1^2E_2 + 9E_1E_3 \right). \quad (20)$$

We performed a Monte Carlo experiment (with 50,000 runs) in which random $s_1, \ldots, s_4$ were taken to be uniformly distributed on the interval [1, 2] (to prevent very small values for $E_4$ which bring numerical instability). Their elementary symmetric functions were evaluated along with all bounds for $E_4$.

Table 1 presents the relative errors of the bounds.

Some histograms of the logarithms of errors of all bounds normalized by $E_4$ are portrayed in Figs. 1–3. The bound (17) is compared in Fig. 1 to the Newton bound (16). Fig. 2 shows histograms of (18).
Table 1
The average square root errors relative to Newton’s bound (16).

<table>
<thead>
<tr>
<th></th>
<th>(16)</th>
<th>(17)</th>
<th>(20)</th>
<th>(18) lower</th>
<th>(18) upper</th>
<th>(19)</th>
<th>(8) lower</th>
<th>(8) upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.13</td>
<td>3.27</td>
<td>0.31</td>
<td>0.26</td>
<td>0.12</td>
<td>0.12</td>
<td>0.02</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Fig. 1. Histograms for the bound from (17) (left panel), and of Newton’s bound (16) (right panel).

Fig. 2. Histograms for the lower bound (18) (left panel), and for the upper bound (18) (right panel).

The performance of bounds (8) is depicted in Fig. 3. These two approximations seem to be the best over all. Mitev’s bound (20) and the Pierce–Foregger–Li bound (17) (admittedly derived for another purpose) performed worse in this situation than Newton’s
upper bound (16). The Rosset upper bound in (18) looks to be more accurate than the lower bound, which is not true for (8). Niculescu’s bound (19) is superior to Rosset’s bound but not to (8).

These pattern holds in simulations performed for other tractable values of $n$ (like $n = 3, 5, 6$).

Acknowledgement

The author is grateful to Dr. G.W. Stewart for his continual help, in particular for his penetrating comments on the draft of this paper.

References


