PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 147, Number 2, February 2019, Pages 481–486
http://dx.doi.org/10.1090/proc/14023
Article electronically published on November 5, 2018

LOGARITHMIC INEQUALITIES UNDER A SYMMETRIC POLYNOMIAL DOMINANCE ORDER

SUVRIT SRA

(Communicated by Walter Van Assche)

ABSTRACT. We consider a dominance order on positive vectors induced by the elementary symmetric polynomials. Under this dominance order we provide conditions that yield simple proofs of several monotonicity questions. Notably, our approach yields a quick (4 line) proof of the so-called *"sum-of-squared-logarithms"* inequality conjectured in (Bîrsan, Neff, and Lankeit, J. Inequalities and Applications (2013); P. Neff, Y. Nakatsukasa, and A. Fischle; SIMAX, 35, 2014). This inequality has been the subject of several recent articles, and only recently it received a full proof, albeit via a more elaborate complex-analytic approach. We provide an elementary proof, which, moreover, extends to yield simple proofs of both old and new inequalities for Rényi entropy, subentropy, and quantum Rényi entropy.

1. INTRODUCTION

Let x be a real vector with n components. Let e_k denote the k-th elementary symmetric polynomial defined by

$$e_k(x) := \sum_{1\le i_1<\cdots<i_k\le n} \prod_{j=1}^{k} x_{i_j}.$$

For nonnegative vectors x, y in $\mathbb{R}^n_+$, we consider the dominance order $\prec_E$ induced by the elementary symmetric polynomials. More precisely, we say $x \prec_E y$ if

$$e_k(x) \le e_k(y), \quad k = 1, \dots, n-1, \quad \text{and} \quad e_n(x) = e_n(y). \tag{1.1}$$

If the last equality is just an inequality $e_n(x) \le e_n(y)$, we write $x \preceq_E y$. We consider functions that are monotonic under the partial order $\prec_E$. Specifically, we say a function $F : \mathbb{R}^n_+ \to \mathbb{R}$ is *E-monotone* if

$$x \prec_E y \quad \Longrightarrow \quad F(x) \le F(y). \tag{1.2}$$

This paper is motivated by a body of recent papers that study E-monotonicity of a specific function: the so-called *"sum-of-squared-logarithms"* $L_n(x)=\sum_{i=1}^n(\log x_i)^2$. Indeed, $L_n(x)$ has been the focus of several works [3, 9, 10, 12], wherein the key open question was establishing its E-monotonicity. The works [3, 9, 12] establish E-monotonicity for $n = 2, 3, 4$; the authors of [10] also highlighted the powerful implications of the general case towards solving certain nonconvex optimization problems to global optimality. Only very recently, a full solution was obtained via a complex analysis [4, 8]. While preparing this paper, it was brought to our notice [7] that [14] has obtained a characterization of E-monotone functions via the

Received by the editors November 8, 2016, and, in revised form, November 7, 2017.
2010 *Mathematics Subject Classification.* Primary 15A42; Secondary 65F60, 26D07, 15AF5.

theory of Pick functions.[1] Our work offers a complementary, and in our view, perhaps the simplest perspective, which yields a short (4 line) proof of E-monotonicity of L_n as a byproduct.

2. E-monotonicity

We introduce now our elementary approach, which leads to a short proof of the E-monotonicity of L_n as well as similar results for related entropy and subentropy inequalities of [5]. Our proof technique should generalize to monotonicity induced by other symmetric polynomials (e.g., Schur polynomials); we leave such an exploration to the interested reader.

Our main result is the following simple, albeit powerful sufficient condition:

Proposition 2.1. *Let ψ be a real-valued function admitting the representation*

$$\psi(s) = \int_0^a \log(t+s)d\mu(t) \quad \text{or} \quad \psi(s) = \int_0^a \log(1+ts)d\mu(t),$$

where $a > 0$, $s \geq 0$, and μ is a nonnegative measure. Then, $\sum_{i=1}^n \psi(x_i)$ is E-monotone.

Proof. Recall first the generating functions for elementary symmetric polynomials

$$\sum\nolimits_{k=0}^n t^k e_k(x) = \prod\nolimits_{i=1}^n (1+tx_i),$$
$$\sum\nolimits_{k=0}^n t^k e_{n-k}(x) = \prod\nolimits_{i=1}^n (t+x_i).$$

Let $x, y \in \mathbb{R}_+^n$, and suppose $x \preceq_E y$. Then using the above generating function representation under this hypothesis we immediately obtain

$$\prod\nolimits_{i=1}^n (1+tx_i) \leq \prod\nolimits_{i=1}^n (1+ty_i) \qquad \forall t \geq 0, \tag{2.1}$$

$$\prod\nolimits_{i=1}^n (t+x_i) \ \leq \prod\nolimits_{i=1}^n (t+y_i) \qquad \forall t \geq 0. \tag{2.2}$$

Taking logarithms, multiplying by $d\mu(t)$, and integrating, it then follows that

$$\sum_{i=1}^n \int_0^a \log(1+tx_i)d\mu(t) \leq \sum_{i=1}^n \int_0^a \log(1+ty_i)d\mu(t)$$
$$\implies \quad F(x) = \sum\nolimits_i \psi(x_i) \leq \sum\nolimits_i \psi(y_i) = F(y).$$

Similarly, with (2.2) we again obtain $F(x) = \sum_i \psi(x_i) \leq \sum_i \psi(y_i) = F(y)$. $\square$

Remark. Observe that the E-monotonicity relation is *weaker* than the usual majorization order. Indeed, if $x \prec y$ (i.e., $\sum_{i=1}^k x_i^\downarrow \leq \sum_{i=1}^k y_i^\downarrow$ for $1 \leq k < n$, and $x^T 1 = y^T 1$), then $e_k(x) \geq e_k(y)$ because e_k is Schur-concave [6].

[1] E-monotonicity of L_n has additional interesting history. P. Neff offered a reward of one ounce of fine gold for its proof, a conjecture that he also announced on the MathOverflow platform [8]. Shortly thereafter, the first full proof was sketched by L. Borisov using contour integration [8]. Approximately two weeks after Borisov's proof, Šilhavý independently characterized E-monotone functions [7]. His results are based on the theory of Pick functions, a natural and elegant approach to study E-monotonicity, which was foreshadowed in the remarkable work of Josza and Mitchison [5].

2.1. **Proof of the SSLI.** As an immediate corollary to Proposition 2.1 we obtain the announced E-monotonicity of $L_n(x) = \sum_{i=1}^n (\log x_i)^2$.

Corollary 2.2. *Let $x, y \in \mathbb{R}_+^n$ such that $x \prec_E y$. Then, $L_n(x) \leq L_n(y)$.*

Proof. The key is to rewrite $(\log x)^2$ so that Proposition 2.1 applies. We notice that

$$(\log x)^2 = \int_0^\infty [\log(1+tx) + \log(t+x) - 2\log(x(1+t))] \frac{dt}{t}. \tag{2.3}$$

Using the integrand of (2.3) with inequalities (2.1), (2.2) and the assumption $e_n(x) = e_n(y)$ (whereby $\sum_i \log(rx_i) = \sum_i \log(ry_i)$ for $r > 0$), we thus obtain

$$\sum\nolimits_i \log(1+tx_i)(t+x_i) - \log\big((1+t)^2 x_i\big) \leq \sum\nolimits_i \log(1+ty_i)(t+y_i) - \log\big((1+t)^2 y_i\big).$$

Integrating this inequality under $d\mu(t) = \frac{dt}{t}$ and using (2.3) the claim follows. □

2.2. **Entropy.** Now we consider application of Proposition 2.1 to obtain entropy inequalities. Recall that for a probability vector x (i.e., x lies in the unit simplex), the Rényi entropy of order α, where $\alpha \geq 0$ and $\alpha \neq 1$, is defined as

$$H_\alpha(x) := \frac{1}{1-\alpha} \log\left(\sum\nolimits_{i=1}^n x_i^\alpha\right). \tag{2.4}$$

The limiting value $\lim_{\alpha \to 1} H_\alpha$ yields the usual (Shannon) entropy $-\sum_i x_i \log x_i$.

Theorem 2.3. *Suppose x and y lie in the unit simplex. Then,*

$$x \prec_E y \implies H_\alpha(x) \leq H_\alpha(y) \quad \textit{for } \alpha \in [0, 2].$$

Proof. Since log is monotonic, to analyze E-monotonicity of H_α, it suffices to consider the following three special cases:

$$\sum\nolimits_{i=1}^n x_i^\alpha \leq \sum\nolimits_{i=1}^n y_i^\alpha \qquad \text{if } 0 < \alpha < 1 \quad \text{and} \quad e_1(x) = e_1(y), \tag{2.5a}$$

$$\sum\nolimits_{i=1}^n x_i^\alpha \geq \sum\nolimits_{i=1}^n y_i^\alpha \qquad \text{if } 1 < \alpha < 2 \quad \text{and } e_1(x) = e_1(y), \tag{2.5b}$$

$$-\sum\nolimits_{i=1}^n x_i \log x_i \leq -\sum\nolimits_{i=1}^n y_i \log y_i \quad \text{and} \quad e_1(x) = e_1(y). \tag{2.5c}$$

For $0 < \alpha < 1$ and $s \geq 0$, we notice the integral representation (see also [13, Ch. 8])

$$s^\alpha = \frac{\alpha \sin(\alpha\pi)}{\pi} \int_0^\infty \log(1+ts) t^{-\alpha-1} dt. \tag{2.6}$$

Given (2.6), an application of Proposition 2.1 immediately yields (2.5a).

For (2.5b), we notice a different representation (notice the extra ts term):

$$s^\alpha = \frac{\alpha \sin(\alpha\pi)}{\pi} \int_0^\infty (\log(1+ts) - ts)\, t^{-\alpha-1} dt. \tag{2.7}$$

This integral converges for $1 < \alpha < 2$ and $s \geq 0$. Since $x \prec_E y$ and we assumed $e_1(x) = e_1(y)$, it follows that $\sum_i \big(\log(1+tx_i) - tx_i\big) \leq \sum_i \big(\log(1+ty_i) - ty_i\big)$. Thus, using (2.7) and noting that $\sin(\alpha\pi) < 0$ for $1 < \alpha < 2$, we obtain (2.5b).

To obtain (2.5c) we apply a limiting argument to (2.5b). In particular, recall that

$$\lim_{\alpha \to 1} \frac{x_i^\alpha - x_i}{\alpha - 1} = x_i \log x_i,$$

so that upon using $\sum_i x_i = \sum_i y_i$ in (2.5b), dividing by $\alpha - 1$, and taking limits as $\alpha \to 1$, we obtain (2.5c). □

2.3. **Inequalities for positive definite matrices.** We note below some inequalities on (Hermitian) positive definite matrices that follow from the above discussion. We write $A > 0$ to indicate that A is positive definite. We extend the definition (1.1) to such matrices in the usual way. In particular, let $A, B > 0$. We say

$$A \prec_E B \quad \Longleftrightarrow \quad \lambda(A) \prec_E \lambda(B), \tag{2.8}$$

where $\lambda(\cdot)$ denotes the vector of eigenvalues. Recalling that $e_k(\lambda(A)) = \operatorname{tr}(\wedge^k A)$, where $\wedge$ is the exterior product [1, Ch. 1], we obtain the following result.

Proposition 2.4. *Let A, B be positive definite matrices. Then,*

$$\operatorname{tr}(\wedge^k A) \leq \operatorname{tr}(\wedge^k B) \quad \textit{for } k = 1, \ldots, n \implies \log\det(I + A) \leq \log\det(I + B).$$

Remark. A classic result in eigenvalue majorization states that if $\log\lambda(A) \prec \log\lambda(B)$ (the usual dominance order), then we have $\log\det(I + A) \leq \log\det(I + B)$. Proposition 2.4 presents an *alternative* condition that implies the same determinantal inequality.

Let us now state two other notable consequences of the order (2.8). To that end, we recall the Riemannian distance on the manifold of positive definite matrices (see, e.g., [2, Ch. 6]) as well as the S-divergence [15]:

$$\delta_R(A, B) := \|\log B^{-1/2} A B^{-1/2}\|_{\mathrm{F}}, \tag{2.9}$$

$$\delta_S(A, B) := \log\det\left(\tfrac{A+B}{2}\right) - \tfrac{1}{2}\log\det(AB). \tag{2.10}$$

Proposition 2.5. *If $A, B, C > 0$ and $AC^{-1} \prec_E BC^{-1}$, then*

$$\delta_R(A, C) \leq \delta_R(B, C), \tag{2.11}$$

$$\delta_S(A, C) \leq \delta_S(B, C). \tag{2.12}$$

Proof. Inequality (2.11) (for $C = I$) was also noted in [4, 11]. It follows readily from Corollary 2.2 once we use definition (2.9) and observe that

$$\delta_R^2(A, C) = \|\log C^{-1/2} A C^{-1/2}\|_{\mathrm{F}}^2 = \sum\nolimits_{i=1}^n (\log\lambda_i(AC^{-1}))^2.$$

To obtain inequality (2.12), first observe that

$$\det(A + C) = \det(C)\det(I + AC^{-1}) = \det(C)\prod\nolimits_{i=1}^n (1 + \lambda_i(AC^{-1})).$$

Thus, we have

$$\begin{aligned}
\delta_S(A, C) &= \log\det(C) + \log\prod\nolimits_{i=1}^n \tfrac{1+\lambda_i(AC^{-1})}{2} - \tfrac{1}{2}\log\det(AC) \\
&\leq \log\det(C) + \log\prod\nolimits_{i=1}^n \tfrac{1+\lambda_i(BC^{-1})}{2} - \tfrac{1}{2}\log\det(AC) \\
&= \log\det\left(\tfrac{B+C}{2}\right) - \tfrac{1}{2}\log\det(BC) \\
&= \delta_S(B, C),
\end{aligned}$$

where the inequality holds due to the hypothesis $\lambda(A) \prec_E \lambda(B)$, which is also used to conclude the second equality by using $\det(A) = \det(B)$. $\square$

2.4. **Quantum entropy.** The entropy inequalities (2.5a)-(2.5c) also extend to their counterparts in quantum information theory. Specifically, recall that the quantum Rényi entropy of order $\alpha \in (0,1) \cup (1,\infty)$ is given by

$$H_\alpha(X) := \frac{1}{1-\alpha} \log \frac{\operatorname{tr}(X^\alpha)}{\operatorname{tr} X}, \tag{2.13}$$

where X is positive definite; moreover, one typically assumes the normalization $\operatorname{tr} X = 1$. Using an argument of the same form as used to prove Theorem 2.3 we can obtain the following result for the Rényi entropy; we omit the details for brevity.

Theorem 2.6. *Let X and Y be positive definite matrices with unit trace. Then,*

$$X \prec_E Y \quad \implies H_\alpha(X) \leq H_\alpha(Y) \quad \text{for } \alpha \in [0,2].$$

3. Subentropy

Next, we briefly discuss an important extension, namely, E-monotonicity of *subentropy*, a quantity that has found use in physics [5]. Formally,

$$Q(x_1, \dots, x_n) := -\sum_{i=1}^{n} \frac{x_i^n}{\prod_{j \neq i}(x_i - x_j)} \log x_i, \tag{3.1}$$

defines a natural entropy-like quantity that characterizes a quantum state with eigenvalues $x_1, \dots, x_n$ (thus $x \geq 0$ and $e_1(x) = 1$). A main result in the work [5] is the following monotonicity theorem for subentropy (rephrased in our notation):

Theorem 3.1 ([5])**.** *If $x \preceq_E y$ and $e_1(x) = e_1(y) = 1$, then $Q(x) \leq Q(y)$.*

Josza and Mitchison [5] prove Theorem 3.1 by appealing to an argument based on contour integration. We note below how a key identity derived by Josza and Mitchison already implies this theorem. Instead of the logarithmic representation of Proposition 2.1, the key idea is to consider the representation

$$\psi(x_1, \dots, x_n) = \int_0^\infty h\left(\prod\nolimits_{i=1}^{n} (t + x_i)\right) d\mu(t), \tag{3.2}$$

where h is any monotonically increasing function and μ is a nonnegative measure. Clearly, if $x \preceq_E y$, then $h(\prod_i (t + x_i)) \leq h(\prod_i (t + y_i))$, whereby $\psi(x) \leq \psi(y)$.

Therefore, to prove $Q(x) \leq Q(y)$, we just need to find a function h such that Q can be expressed as (3.2). Such a representation was already obtained in [5], wherein it is shown that for $x > 0$ such that $e_1(x) = 1$, we have

$$Q(x_1, \dots, x_n) = -\int_0^\infty \Big[\frac{t^n}{\prod_{j=1}^{n}(t + x_j)} - \frac{t}{1+t}\Big] dt. \tag{3.3}$$

Thus, using $h(s) = -1/s$ and $d\mu(t) = t^n dt$, and adding $\frac{-t}{1+t}$ to ensure convergence (the constraint $e_1(x) = e_1(y)$ is needed to cancel out the effect of this term), we obtain $Q(x) \leq Q(y)$ whenever $x \preceq_E y$ and $e_1(x) = e_1(y)$.

A similar argument yields the following inequality, which is otherwise not obvious:

$$x \prec_E y \implies \sum_{i=1}^{n} \frac{(-1)^{i+1} x_i^\alpha}{\prod_{j \neq i}(x_i - x_j)} \quad \geq \quad \sum_{i=1}^{n} \frac{(-1)^{i+1} y_i^\alpha}{\prod_{j \neq i}(y_i - y_j)} \quad \text{for } 0 < \alpha < 1. \tag{3.4}$$

References

[1] Rajendra Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, New York, 1997. MR1477662

[2] Rajendra Bhatia, *Positive definite matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007. MR2284176

[3] Mircea Bîrsan, Patrizio Neff, and Johannes Lankeit, *Sum of squared logarithms—an inequality relating positive definite matrices and their matrix logarithm*, J. Inequal. Appl. (2013), 2013:168, 17. MR3057239

[4] Lev Borisov, Patrizio Neff, Suvrit Sra, and Christian Thiel, *The sum of squared logarithms inequality in arbitrary dimensions*, Linear Algebra Appl. **528** (2017), 124–146. MR3652841

[5] Richard Jozsa and Graeme Mitchison, *Symmetric polynomials in information theory: entropy and subentropy*, J. Math. Phys. **56** (2015), no. 6, 062201, 17. MR3369891

[6] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold, *Inequalities: theory of majorization and its applications*, 2nd ed., Springer Series in Statistics, Springer, New York, 2011. MR2759813

[7] Patrizio Neff. Private communication. Sep. 2015. 18/9/2015.

[8] Patrizio Neff. The sum of squared logarithms conjecture. MathOverflow, 2015.

[9] Patrizio Neff, Bernhard Eidel, Frank Osterbrink, and Robert Martin. The Hencky strain energy $\|\log U\|^2$ measures the geodesic distance of the deformation gradient to $\mathrm{SO}(n)$ in the canonical left-invariant Riemannian metric on $\mathrm{GL}(n)$. *Proceedings in Applied Mathematics and Mechanics*, 13(1):369–370, 2013.

[10] Patrizio Neff, Yuji Nakatsukasa, and Andreas Fischle, *A logarithmic minimization property of the unitary polar factor in the spectral and Frobenius norms*, SIAM J. Matrix Anal. Appl. **35** (2014), no. 3, 1132–1154. MR3252811

[11] Patrizio Neff, Bernhard Eidel, and Robert J. Martin, *Geometry of logarithmic strain measures in solid mechanics*, Arch. Ration. Mech. Anal. **222** (2016), no. 2, 507–572. MR3544313

[12] Waldemar Pompe and Patrizio Neff, *On the generalised sum of squared logarithms inequality*, J. Inequal. Appl. (2015), 2015:101, 17. MR3336366

[13] René L. Schilling, Renming Song, and Zoran Vondraček, *Bernstein functions*, 2nd ed., De Gruyter Studies in Mathematics, vol. 37, Walter de Gruyter & Co., Berlin, 2012. Theory and applications. MR2978140

[14] Miroslav Šilhavý. A functional inequality related to analytic continuation. 2015. `www.math.cas.cz/fichier/preprints/IM_20150623102729_44.pdf`.

[15] Suvrit Sra, *Positive definite matrices and the S-divergence*, Proc. Amer. Math. Soc. **144** (2016), no. 7, 2787–2797. MR3487214

Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Email address: `suvrit@mit.edu`