ARITHMETIC MEAN-GEOMETIRC MEAN INEQUALITY
: A New Perspective and Significant Applications

***Khosrow Moshirvaziri**, Information Systems Dept., California State University, Long Beach, Long Beach, CA 90840, 562-985-7965, moshir@csulb.edu*
***Mahyar A. Amouzegar**, College of Engineering, California State Polytechnic University, Pomona, CA 91768, 909-869-2472, mahyar@cpp.edu*
***Fahimeh Rezayat**, Information Systems & Operations Management, California State University, Dominguez Hills, Carson, CA 90747, 310-243-3484, frezayat@csudh.edu*

ABSTRACT[1]

We present both theoretical as well as simulated bounds on the difference between the two common economic measures, namely Arithmetic and Geometric means[2]. Additionally, we derive a novel proof of the Arithmetic-Geometric Mean Inequality (A-G) by a unique use of global optimization theory and provide a brief survey of significant application areas of A-G in Engineering Design, Mathematics, Business, and Economics.

BACKGROUND

The classical inequality of arithmetic and geometric means, or more simply the AM–GM or even briefer, A–G inequality, states that the arithmetic mean (A) of a series of non-negative real numbers is at least as large as the geometric mean (G) of the same series; with the two means being equal if and only if every member in the series is the same.

The inequality is a fundamental relationship in mathematics and statistical sciences. It is a powerful utility in computational analysis, for problem solving, and for establishing relationships with other mathematical concepts. The arithmetic- geometric mean was first discovered by Lagrange and rediscovered by Gauss a few years later [1]. Much of its current popularity is due to its new applications in Engineering Design, Signal Processing, Telecommunication, Motion picture and Animation Editing, using Geometric Mean Filter, and of course, its tremendous applications in Investment Science and Financial Engineering, as an instrument of economic analysis, amongst others. [2] [3]

Computer scientists, electrical and chemical engineers, biologists, and other advanced manufacturers and scientists are often concerned with identifying a precise measurement of certain concentrations (e.g., bacteria population) or particles in their production environments such as in water or in air. Since levels of concentration in such data may vary significantly, computational algorithms to determine a safe threshold must be based on geometric means. Additionally, the use of geometric mean methods in improving the convergence rate of matrix decomposition procedures has received significant attention. In the same context, its extension of the means from a positive real number to a positive semidefinite matrix operator has seen extensive developments [25]

[1] Proceedings of the Forty-Fifth Annual Meeting of Western DSI Conference, Las Vegas, NV, April 5th – April 9th, 2016
[2] The full version of this paper including computational bounds and simulation results is in preparation for submission elsewhere.

INTRODUCTION

There are several inductive proofs of (1) below, making use of either Cauchy-Shwarz, Jensen's Inequality, or Polya's Inequality (for a detailed discussion see, [4], [5], [6], [20], [21], [22], [23]), including one of the original paper by Muirhead [7] in 1903.

Our proof is not inductive and unique in the sense that it will for the first time makes use of results of optimization technique of nonconvex programs.

1. The Arithmetic-Geometric Inequality

Consider vector $x \in \mathbb{R}^n$, $x_i \geq 0, \ \forall\, i = 1, \cdots, n,$ then, the A–G inequality is given by equation (1).

$$G = \left[\prod_{i=1}^{n} x_i\right]^{1/n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n} = A, \qquad (A - G) \qquad (1)$$

Where equality is attained when all$x_i, \ i = 1, \cdots, n.$, are equal. Observing that $x_i = e^{\ln x_i}$, we can write the inequality in a more compact form.

$$G = exp\left[\frac{1}{n}\sum_{i=1}^{n} \ln x_i\right] \leq \frac{1}{n}\sum_{i=1}^{n} x_i = A, \qquad (2)$$

Therefore, the geometric mean may also be calculated by computing the arithmetic mean of the logarithms of the data values and taking the inverse logarithm of the result.

2. Geometric Programming (GP)

To pave the way to present our new proof of (1), we outline below some of the relevant results and preliminaries from Geometric Programming.
For $c > 0$ and numbers $a_i \in \mathbb{R}^1$, $a_i \geq 0, \ \forall\, i = 1, \cdots, n.$ the function $f(x)$ on $x \in \mathbb{R}^n$ defined below, is called monomial function:

$$f(x) = c\, x_1^{a_1}\, x_2^{a_2} \cdots x_n^{a_n}, \qquad (A) \qquad (3)$$

Monomials are closed under multiplication and division, in the sense that, if f and g are two monomials then it can be shown that both fg and f/g are monomials as well (see [10].) This is also true when scaling the monomials by any positive constant. A monomial raised to any power α is also a monomial:

$$(f(x))^{\alpha} = c^{\alpha}\, x_1^{\alpha\, a_1}\, x_2^{\alpha\, a_2} \cdots x_n^{\alpha\, a_n}, \qquad (4)$$

For constants $c_j > 0$, the sum of one or more monomials of the form (4) is called Posynomial:

$$f(x) = \sum_{j=1}^{J} c_j\, x_1^{a_{1j}}\, x_2^{a_{2j}} \cdots x_n^{a_{nj}}, \qquad \text{(PS)} \qquad (5)$$

Clearly, any monomial is a posynomial. Posynomial functions are known to be closed under addition, multiplication, and positive scaling. For example, the result of dividing posynomials by monomials is a posynomial.

Generally, if *f* is a posynomial and *g* is a monomial, then *fg* is a posynomial. For a nonnegative integer power *α* and *f* is a posynomial, then *fα*, a product of *α* posynomials, is also a posynomial.

In the sequel, we make use of results presented here to solve a related optimization problem of the form:

$$\max \{ f(x) \mid e^{\top} x \leq a, \; x \geq 0. \}, \qquad \text{(P)} \qquad (6)$$

Where,

$$f(x) = \left[\prod_{i=1}^{n} x_i \right]^{1/n},$$

and $e \in \mathbb{R}^n$ and $e = (1, \cdots, 1)^{\top}$ is a vector of all ones and *a,* a constant. Note that $f(x)$ is a monomial, a simple case of equation (4) with $c = 1$ and all $a_i, \; i = 1, \cdots n,$ being equal to $1/n$ (equivalently, a posynomial, a simple case of equation (5) with $J = 1, \; a_{ij} = 1/n, \; i = 1, \; j = 1, \cdots, n.$ Furthermore, since we are maximizing f, the equality will be achieved on the constraint (6).

Theorem 1 *Problem (6) above attains its unique optimal solution with an optimal function of* $(a/n)^{1/n}$ *and solution vector* $\{x^{\star} = a/n \,, \; i = 1, \cdots, n.\}$

Proof: We note that the objective function $f(x)$ being a simple case of a posynomial is amenable to a positive constant scaling. This means, without loss of generality, we can assume $a = 1$, which implies that if A-G inequality holds for some vector $x \in \mathbb{R}^n$, it will also hold for βx, for any constant $\beta > 0$. Clearly a choice of $\beta = 1/\sum x$ results in $a = 1$. Alternatively, one may substitute $\beta x_i \,, \; i = 1, \cdots, n,$ into (1) to verify that the inequality still holds after scaling by β. Next, if $x^{\star}$ is not an optimal solution for (P), then we assume there exists an optimal solution, say $\hat{x}$ which is different from $x^{\star}$ by at least two components, say and $\hat{x}_j < a/n$ and $\hat{x}_k > a/n$ for k and j between a and n for otherwise, the sum would not add up to a. And of course $x_i^{\star} = \hat{x}_i, \; i = \{1, \cdots, n\} \setminus \{j, k\}$. By using an algebraic identity one can show:

$$(\hat{x}_j - \hat{x}_k)^2 \geq 0, \qquad \Longrightarrow \quad \hat{x}_j \hat{x}_k \leq \frac{1}{4}\left(\hat{x}_j + \hat{x}_k\right)^2 = \left(\frac{\hat{x}_j + \hat{x}_k}{2}\right)\left(\frac{\hat{x}_j + \hat{x}_k}{2}\right).$$

However, the later expression simply suggests that we can form another solution, say $\bar{x}$ which is the same as $\hat{x}$ except for the two components $\bar{x}_j$ and $\bar{x}_k$.

$$\bar{x}_j = \bar{x}_k = \left(\frac{\hat{x}_j + \hat{x}_k}{2}\right) \quad \Longrightarrow \quad f(\hat{x}) \leq f(\bar{x}), \qquad (7)$$

Yet a simple comparison of the corresponding objective values implies that we have a better solution and thus a contradiction to the assumption that $\hat{x}$ is an optimal solution. Hence, x^{*} is a unique optimal solution to (6). Consequently, for any other solution vector, $x \in \mathbb{R}^n$, we have:

$$f(x) = \left[\prod_{i=1}^{n} x_i \right]^{1/n} \leq \; f(x^{*}) = \left(\frac{a}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad (8)$$

This completes the proof. It simply states that for any nonnegative vector $x \in \mathbb{R}^n$, (Data array of length *n* of positive numbers), we have arrived at the A–G inequality:

$$G = \left[\prod_{i=1}^{n} x_i\right]^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n} x_i = A, \tag{9}$$

Note that in the preceding argument, function *f* is defined on the domain $\mathbb{R}_+^n$, $f: \mathbb{R}_+^n \longmapsto \mathbb{R}^1$. Thus, due to concavity of *f*, problem (P) given by (6) is a convex program and thus attains a unique global solution. Concavity of *f* is guaranteed since its Hessian matrix $\nabla^2 f(x)$ may be written as:

$$\frac{\partial^2 f(x)}{\partial x_k^2} = (1-n)\,\frac{f(x)}{n^2 x_k^2}, \qquad \frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{f(x)}{n^2 x_k \partial x_l} \quad \text{for} \quad k \neq l,$$

This implies that:

$$\nabla^2 f(x) = -\,\frac{\prod_{i=1}^{n} x^{1/n}}{n^2}\,(\,n\,\mathbf{diag}\,(1/x_1^2 \,\cdots\, 1/x_n^2) - qq^{\intercal})$$

Where elements of vector q is given by $\{q_i = 1/x_i^2,\ i = 1, \cdots, n\}$ and **dig** is a diagonal matrix with the diagonal elements as given. To show the concavity of *f* it remains to show the Hessian $\nabla^2 f(x) \preccurlyeq 0$. However, we can observe that

$$u^{\intercal}\,\nabla^2 f(x)\,u = -\,\frac{\prod_{i=1}^{n} x^{1/n}}{n^2}\left(n\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} u_i^2/x_i^2 - \left(\sum_{i=1}^{n} u_i/x_i\right)^2\right) \leq 0.$$

This is true for all such vector $u \in \mathbb{R}^n$. Hence, the Hessian is a Negative Semidefinite (NSD) matrix and thus *f* (*x*) as given by equation (8) above is concave. The last inequality follows the Cauchy-Schwarz inequality which states that for any two vectors *a* and *b* in $\mathbb{R}^n$, ||*a*||||*b*|| ≥ ||*ab*|| [11].

3. Applications areas

IEEE Xplore Digital Library [26] lists over 2000 scientific articles involving the use of this inequality in diverse areas in Science, Technology, Engineering and Mathematics (STEM). Here we briefly list a few exciting and newly developed application areas.

Matrix Decomposition - For a complex matrix H, a particular decomposition of *H* is constructed such that *H* = *QRP∗* where *Q* and *P* have orthonormal columns, and *R* is a real upper triangular matrix with diagonal elements equal to the geometric mean of the positive singular values of *H* and * denotes conjugate transpose (see ([2] and references therein for details.) This technique is known as the Geometric Mean Decomposition (GMD) and has major applications in signal processing and in the design of telecommunication networks. A unitary matrix is a matrix that has an inverse and a transpose whose corresponding elements are pairs of conjugate complex numbers. As it turns out, the unitary matrices correspond to information lossless filters applied to transmitted and received signals that minimize the maximum error rate of the network [2].

Geometric Mean Decomposition (GMD) algorithm is considered an efficient pre-coding scheme in joint Multiple Inputs Multiple Outputs (MIMO) transceiver designs capable of facilitating asymptotically equivalent performance of maximum likelihood detector (MLD) (See [13] and [14] for more details).

Test Matrices - Another application of GMD is in the construction of test matrices with user pre-specified singular values [2] utilized in the solution of a class of minimax problems [15].

Aspect Ratio - Compromising aspect ratio in video and multi-media industry is a challenging problem for smart television and monitor manufacturers. Geometric mean of various aspect ratios provides a compromise between them, distorting or cropping both in some sense equally [16]. For example, in compromising between the two aspect ratios of 4:3 and 2.35 (2.35 ≅ 33:14) the Society of Motion Picture and Television Engineers (SMPTE) adopted 16:9 (1.77777777) which is the closest to:

$$\left((2.35) \times \left(\frac{4}{3}\right)\right)^{1/2} = 1.770122406$$

There is also growing needs in the smart phone and computer industry to design optimum display devices. In depth description of the techniques used and related US Patents are available in [17] and [18].

Subspace Selection - Subspace selection approaches are powerful tools in pattern classification and data visualization. One of the most important subspace approaches is the linear dimensionality reduction step in the Fisher's linear discriminant analysis (FLDA), which has been successfully employed in many fields such as biometrics, bioinformatics, and multimedia information management. However, the linear dimensionality reduction step in FLDA has some critical drawback. The geometric mean for subspace selection has shown superior results [19].

CONCLUDING REMARKS

We have used the opportunity of exploring the significance of AM-GM Inequality to also stress the use of optimization techniques and theories as means of analyzing non trivial inequalities and signal the use of classical inequalities such as AM–GM to fine-tune objective function and constraints of difficult to solve optimization problems (e.g., nonconvex optimization). The latter may be achieved by using inequalities to obtain tighter bounds on constraints or on the objective function value in an optimization model. Such techniques may lead to faster optimization or better pruning of branches; say in an Integer Programming Branch-and-Bound approach or in a Branch-and-Cut technique in global optimization algorithms. This unique proof should provide a new vehicle for other operations researchers to pay closer attention to AM–GM Inequality, given its wide application. For ex- ample, the Cobb-Douglas production function, well known in economic theory and management science, is a posynomial of the form described in an earlier section. Thus, this fundamental inequality continues to generate a vast amount of research in a variety of engineering and sciences disciplines.

REFERENCES

1. Almkvist, G. and Berndt, B., (1988) Gauss, Landen, Ramanujan, the Arithmetic-Geometric Mean, Ellipses, p, and the Ladies Diary, The American Mathematical Monthly, 95:7, 585– 608.
2. Jiang, Y., Hager, W., and Li, J., (2005), The Geometric Mean Decomposition, Linear Algebra and its Applications 396: 373-384
3. Kim, J.K., Oh, C.H., Chung, C.H., Hwang, J., and Lee, W., (2003), Lecture Notes in Computer Science, 2668, 411–418

4. Hardy, G., Littlewood, J.E., and Polya, G., (1952), Inequalities, Second edition, reprinted in paperback in 1999.
5. Chong, K.M., (1976), An Inductive Proof of A.M.–G.M. Inequality, The American Mathematical Monthly, 83:5, 369.
6. Chong, K.M., (1976), The Arithmetic Mean-Geometric Mean Inequality: A New Proof, Mathematics Magazine, 49(2), 87–88.
7. Muirhead, R.F., (1903), Proof That the Arithmetic Mean Is Greater than the Geometric Mean The Mathematical Gazette, 2:39, 283–287.
8. Spizman, L., and Weinstein, M. A., (2008), A Note on Utilizing the Geometric Mean: When, Why and How the Forensic Economist Should Employ the Geometric Mean. Journal of Legal Economics 15(1): 43–55.
9. Costa, K., J., Calculating Geometric Means, Buzzards Bay National Estuary Program http://www.buzzardsbay.org/geomean.htm
10. Boyd, S., Kim, S. J., Vandenberghe, L., and Hassibi, A., (2007), A tutorial on geometric programming, Optim. Eng. 8: 67-127.
11. Boyd, S. and Vandenberghe, L. (2004) Convex Optimization, Cambridge University Press New York, ISBN:0521833787
12. Bjorn Poonen, B., (1999), Inequalities. Berkeley Math Circle, 6–7.
13. Hwang, Y.T., Chen, W-D., and Hong, C-R., (2014), A Low Complexity Geometric Mean Decomposition Computing Scheme and Its High Throughput VLSI Implementation IEEE Transactions on Circuits and Systems-I,a 61:4, 1170–1182.
14. Chen, W-D. and Hwang, Y-T., (2013, A Constant Throughput Geometric Mean Decom- position Scheme Design for Wireless MIMO Precoding, IEEE Transactions on Vehicular Technology, 62:5, 2080–2090.
15. Kosowski, P. and Smoktunowicz, A., (2000), On constructing unit triangular matrices with prescribed singular values, Computing 64: 279-285.
16. Rowley, E. E., (1987), The Financial System Today. Manchester University Press, ISBN 0719014875
17. Technical Bulletin, (2001), Understanding Aspect Ratios. The Cinema Source Press, Retrieved 2009-10-24.
18. US 5956091, (1999), "Method of showing 16:9 pictures on 4:3 displays", issued September 21, 1999
19. Tao, D., Li, X., Wu, X., and Maybank, S., (2009) Geometric Mean for Subspace Selection, IEEE Transactions on Pattern Analysis Machine Intelligence, 31:2.
20. Beckenbach, E.F. and Bellman, R.E., (1965), Inequalities, Springer-Verlag.
21. Vu, T., (2011), California's Ports Trade Volume Increases in a Slow, Delicate Recovery, California Chamber of Commerce, P.O. Box 1736, Sacramento, CA 95812.
22. Diananda P.H., (1960), A simple proof of the arithmetic mean is greater than geometric mean, Amer. Math. Monthly 67:1007.
23. Chrystal, G., (1900), Algebra II, Edition 2, London.
24. Freeman, W., (2012), Minimizing Areas and Volumes and a Generalized AMGM Inequality, Mathematics Magazine, 85:2, 116–123.
25. Rassouli, M. and Leazizi, F., (2003) Continued fraction expansion of the geometric matrix mean and applications, Linear Algebra and its Applications 359:3757.
26. IEEE Xplore Digital Library. Retrieved from: http://ieeexplore.ieee.org/xpl/periodicals.jsp