On weighted means, geometric and arithmetic.

Notation and nomenclature

Function application is denoted by an infix dot ".", function composition by an infix ringlet "$\circ$" [i.e., $g.(h.x) = (g \circ h).x$].

Dummy variable i ranges over some finite, non-empty domain $\mathcal{D}$.

Function p is of type $\mathcal{D} \to \mathbb{R}$ and satisfies

(0) $\langle \forall i :: 0 \le p.i \rangle \wedge \langle \Sigma i :: p.i \rangle = 1$.

For any function q of type $\mathcal{D} \to \mathbb{R}$, the weighted average $w.q$ is defined by

(1) $w.q = \langle \Sigma i :: p.i * q.i \rangle$.

Finally, function a of type $\mathcal{D} \to \mathbb{R}$ satisfies $\langle \forall i :: 0 < a.i \rangle$.

The theorem

The well-known fact that the geometric mean is at most the arithmetic mean, i.e., with N equal to the number of elements in $\mathcal{D}$,

(2) $\sqrt[N]{\langle \Pi i :: a.i \rangle} \le \langle \Sigma i :: a.i \rangle / N$,

is a special case of

(3) $\langle \Pi i :: a.i^{\,p.i} \rangle \leq \langle \Sigma i :: p.i * a.i \rangle$.

[Take $p.i = 1/N$, and (3) reduces to (2).] Our target is to prove (3) . Taking logarithms of both sides we are led to the relation

(4) $w.(\log \circ a) \leq \log.(w.a)$.

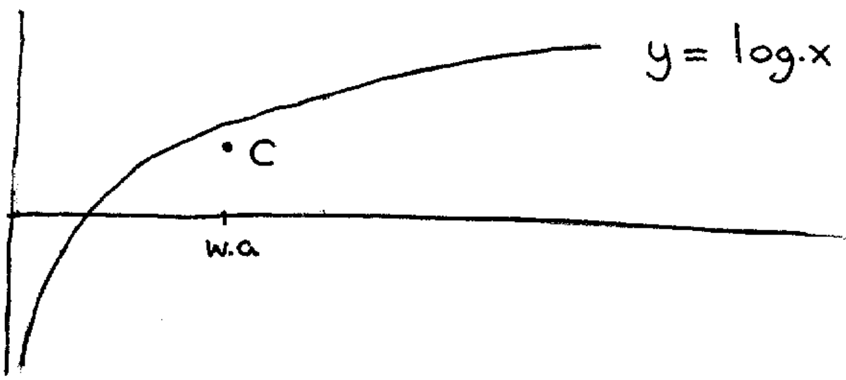
Because of the monotonicity of the exponential function, (4) implies (3) ; our target is now to prove (4) .

The proof

Our proof uses two facts:

(i) When weights $p.i$ are placed at points belonging to a convex region, their centre of gravity belongs to that region as well. [Note that this is a possible definition of convexity.]

(ii) In the Cartesian plane, the points on or below the curve $y = \log.x$ [i.e., the points (x,y) such that $0 < x \wedge y \leq \log.x$] form a convex region. [This is a direct consequence of the fact that the 2nd derivative of the logarithm is negative.]



Let the weights p.i be placed at (a.i , log.(a.i)) — i.e., all the weights are on the curve y = log.x — . Then their centre of gravity C is at

(5) $\qquad (w.a,\ w.(\log{}^\circ a))$.

Because of facts (i) and (ii) , C lies on or below the curve y = log.x , i.e. is a point (x, y) such that $y \leq \log.x$; in terms of (5) this means

$$w.(\log{}^\circ a) \leq \log.(w.a)$$

which is what had to be proved.

In retrospect

In my initial endeavours to prove (3), I tried mathematical induction over the size of D . I noticed (i) that I got formulae that reminded me of the centre of gravity and (ii) that the exponentiation ~~in particular~~ $\cancel{(a+a')^{p+p'}}$ gave rise to ugly formulae. The step from the centre of gravity to convexity is very small. [I felt that I was benefiting from my past as a physicist.] To get rid of the exponentiation, I decided to take logarithms, feeling free to do so because (3) and (4) are as a matter of (simple) fact equivalent. But (4) is so much more beautiful than (3) that I suddenly knew that I was on a very promising track. Please note that we can even write (4) as

$$(4') \qquad w.(\log^{\circ} a) \leq (\log^{\circ} w).a \quad ,$$

in which both sides contain the same collection of symbols. [During this play with symbols I felt I was profiting from my familiarity with functional programming.]

Connecting convexity with the logarithm was the last, in fact simple and obvious step. In order to appreciate how well we disentangled the argument, the reader is invited to prove (under the same conditions)

$$\langle \Sigma i :: p.i * a.i \rangle \leq \sqrt[2]{\langle \Sigma i :: p.i * (a.i)^2 \rangle} \quad .$$

<u>Acknowledgement</u>

I would like to express my gratitude to

the members of the Austin Tuesday Afternoon Club who joined me in the first exploration of the theorem and whose comments on my draft helped me with the proof's presentation. Finally I would like to express my gratitude to Dr. Martin Peters and Ruth Allewelt of Springer-Verlag for making a present to me of [0], where I found the theorem. [Needless to say, I prefer my proof.]

[0] Aigner, Martin & Ziegler, Günter M., "Proofs from THE BOOK", Springer-Verlag Berlin, Heidelberg, New York, 1998 [ISBN 3-540-63698-6]

Austin, 9 February 2000

prof. dr. Edsger W. Dijkstra
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712-1188
USA