COST BENEFIT ANALYSIS; ECONOMICS 581: LECTURE NOTES

By W. Erwin Diewert   March, 2011.
Email: diewert@econ.ubc.ca
Website: http://www.econ.ubc.ca/diewert/hmpgdie.htm

CHAPTER 2: Inequalities

1. Introduction

Inequalities play an important role in many areas of economics. Unfortunately, this topic is not usually covered in the typical Mathematics for Economists course so we will give an introduction to this topic in this chapter, deriving the most important inequalities that are used in applied economics.

In section 2, we provide some proofs of the *Cauchy Schwarz Inequality* while section 3 provides a proof of the *Theorem of the Arithmetic and Geometric Means*.

Section 4 introduces the *mean of order r*, which is a special case of the Constant Elasticity of Substitution (or CES) functional form for a utility or production function. Means of order r are required in order to state *Schlömilch's Inequality*, which is a generalization of the Theorem of the Arithmetic and Geometric Means. Schlömilch's Inequality will be proven in section 5.

Section 6 introduces a type of mean or average that plays a prominent role in index number theory: the *logarithmic mean* of two positive numbers.

Section 7 establishes a few more properties of the means of order r. In particular, we look at limiting cases as r tends to plus or minus infinity.

Finally, section 8 concludes with a brief summary of methods that are used to establish inequalities.

2. The Cauchy-Schwarz Inequality

Proposition 1: *Cauchy* (1821; 373) - *Schwarz* (1885) *Inequality*.

Let x and y be N dimensional vectors. Then[1]

(1) $(x^T y)^2 \le (x^T x)(y^T y)$.

Proof: Define the N by 2 matrix A as follows:

---

[1] This proof may be found in Hardy, Littlewood and Polya (1934; 16).

(2) $A \equiv [x, y]$.

Define the 2 by 2 matrix B as follows:

(3) $B \equiv A^T A = [x,y]^T [x,y] = \begin{bmatrix} x^T x & x^T y \\ y^T x & y^T y \end{bmatrix}$.

It is easily seen that B is a positive semidefinite matrix, since

(4) $z^T B z = z^T A^T A z = (Az)^T (Az) = u^T u \geq 0$

where $z^T \equiv [z_1, z_2]$ and the N dimensional vector u is defined as $Az = xz_1 + yz_2$. The determinantal conditions for B to be positive semidefinite imply:

(5) $0 \leq |B| = \begin{vmatrix} x^T x & x^T y \\ y^T x & y^T y \end{vmatrix} = x^T x \, y^T y - (x^T y)^2$,

and (5) simplifies to (1).   Q.E.D.

Note that for (1) to be a strict inequality, (5) must be a strict inequality and hence B must be positive definite. This in turn implies that we must have:

(6) $0_N \neq u = xz_1 + yz_2$ for $(z_1, z_2) \neq (0, 0)$,

and (6) in turn implies that both x and y must be nonzero and nonproportional. Thus to obtain a strict inequality in (1), we cannot have $x = ky$ or $y = kx$ for any scalar k.

The problem below provides an alternative proof for (1). The problems below and the material in the following sections will provide many applications of the Cauchy-Schwarz inequality.

Problems:

1. Assume $x \neq 0_N$ and $y \neq 0_N$ (the inequality (1) is trivially true if either x or y equals 0_N), and for each real number λ, define $f(\lambda)$ as

(i) $f(\lambda) \equiv (x + \lambda y)^T (x + \lambda y) = \lambda^2 \, y^T y + 2\lambda x^T y + x^T x \geq 0$.

The inequality in (i) is true because $(x + \lambda y)^T (x + \lambda y) = \sum_{i=1}^N (x_i + \lambda y_i)^2$ is a sum of squares. Now use standard calculus techniques and minimize $f(\lambda)$ with respect to λ. Let the minimizing λ be denoted as λ^*. Now calculate $f(\lambda^*)$ and it will turn out that the inequality

(ii) $f(\lambda^*) \geq 0$

is equivalent to the Cauchy-Schwarz Inequality (1) above. (It is not necessary to check the second order conditions for the minimization problem associated with minimizing $f(\lambda)$.)

2. Let Y and X be two N dimensional vectors; i.e., define $Y^T \equiv [Y_1, \ldots Y_N]$; $X^T \equiv [X_1, \ldots X_N]$. Define the arithmetic mean of the Y_n and X_n as $Y^* \equiv (1/N)\sum_{n=1}^{N} Y_n$ and $X^* \equiv (1/N)\sum_{n=1}^{N} X_n$ respectively. Now define the vectors y and x as Y and X except we subtract the respective means from each vector; i.e., define:

(i) $y \equiv Y - Y^* 1_N$; $x \equiv X - X^* 1_N$

where 1_N is a vector of ones of dimension N. Now consider the following regression models of y on x and x on y:

(ii) $y = \alpha x + u$;
(iii) $x = \beta y + v$

where u and v are error vectors and α and β are unknown parameters. We assume that x $\neq 0_N$ and y $\neq 0_N$. The *least squares estimator for* α is the α^* which solves the unconstrained minimization problem:

(iv) $\min_{\alpha} f(\alpha)$

where f is defined as

(v) $f(\alpha) \equiv u^T u = (y - \alpha x)^T (y - \alpha x)$.

The *least squares estimator for* β is the β^* which solves the unconstrained minimization problem:

(iv) $\min_{\beta} g(\beta)$

where g is defined as

(v) $g(\beta) \equiv v^T v = (x - \beta y)^T (x - \beta y)$.

(a) Find the least squares estimators for α and β, α^* and β^*. Check the second order conditions for your solutions.

The *variances* for Y and X and the *covariance* between Y and X are defined as follows:

(vi) $\text{Var}(Y) \equiv y^T y/N$; $\text{Var}(X) \equiv x^T x/N$; $\text{Cov}(Y,X) \equiv x^T y/N$.

The *correlation coefficient* ρ between Y and X is defined as follows:

(vii) $\rho \equiv \text{Cov}(Y,X)/[\text{Var}(Y)\text{Var}(X)]^{1/2} = x^T y/(x^T x)^{1/2}(y^T y)^{1/2}$.

Note that ρ is well defined since we have assumed that $x \neq 0_N$ and $y \neq 0_N$ and hence $(x^T x) > 0$ and $(y^T y) > 0$ and so the positive square roots, $(x^T x)^{1/2}$ and $(y^T y)$ are well defined positive numbers.

(b) Prove that the correlation coefficient is bounded from below by minus one and from above by plus one; i.e., show that:

(viii) $-1 \leq \rho \leq 1$.

(c) Assume that the correlation coefficient between Y and X is positive; i.e., assume that $\rho > 0$. Prove that:

(ix) $\alpha^* \leq 1/\beta^*$.

(d) Under what conditions will (ix) hold as an equality?

(e) Assume that the correlation coefficient between Y and X is negative and derive a counterpart inequality involving α^* and β^* to (ix) above.

Comment: The result (ix) is reasonably well known in the literature; e.g., see Kendall and Stuart (1967; 380) or Bartelsman (1995; 60). However, the implications of the inequality are rather important for applied economists. In many applications, the magnitude of α or β is very important. Hence if ρ is positive and a client wants an applied economist to obtain a small estimate for the parameter α, then the applied economist will be tempted to run a regression of Y on X but if the client wants a large estimate for α and hence a small estimate for β, then the applied economist will be tempted to run a regression of X on Y in order to please the client.

3. The *Triangle Inequality*. The (Euclidean) *distance* (or norm) of an N dimensional vector x from the origin is defined as

(i) $d(x) \equiv (x^T x)^{1/2}$

Let x and y be two N dimensional vectors. Show that the following inequality is satisfied:

(ii) $d(x+y) \leq d(x) + d(y)$.

Comment: This inequality dates back to Euclid.

4. Let A be an N by N positive semidefinite symmetric matrix and let x and y be two N dimensional vectors. Show that the following inequality is true.

(i) $(x^TAy)^2 \le (x^TAx)(y^TAy)$.

Hint: Since A is symmetric, there exists an orthonormal matrix U such that:

(ii) $U^TAU = \Lambda$;
(iii) $U^TU = I_N$

where Λ is a diagonal matrix which has the nonnegative eigenvalues of A, $\lambda_1,...,\lambda_N$, running down its main diagonal and I_N is the N by N identity matrix. Thus A can be written as:

(iv) $A = U\Lambda U^T = U\Lambda^{1/2}\Lambda^{1/2}U^T = U\Lambda^{1/2}U^TU\Lambda^{1/2}U^T = SS$

where $\Lambda^{1/2}$ is a diagonal matrix with the positive square roots of the eigenvalues $\lambda_1,...,\lambda_N$ running down its main diagonal and S is the *symmetric square root matrix* for A.

3. The Theorem of the Arithmetic and Geometric Mean

Let $x \equiv [x_1,...,x_N]$ be a vector of nonnegative numbers.[2] The *ordinary geometric mean* of the N numbers contained in the vector x is defined as $(x_1x_2...x_N)^{1/N}$ and the *ordinary arithmetic mean* of these numbers is defined as $(x_1+x_2+...+x_N)/N$.

In this section, we will deal with *generalized or weighted geometric and arithmetic means* of the N nonnegative numbers x_n; n = 1,...,N. In order to define these weighted means, we first define a vector of *positive weights* $\alpha \equiv [\alpha_1,...,\alpha_N]$; i.e. define the components of the α vector to satisfy the following restrictions:[3]

(7) $\alpha \gg 0_N$; $1_N^T\alpha \equiv \sum_{n=1}^{N} \alpha_n = 1$.

Now we are ready to define the *weighted geometric mean* $M_0(x)$ as follows:

(8) $M_0(x) \equiv \prod_{n=1}^{N} x_n^{\alpha_n}$.

In a similar fashion, we define the *weighted arithmetic mean* $M_1(x)$ as follows:[4]

---

[2] For consistency, we should define the column vector x as $[x_1,...,x_N]^T$. When it is important to be precise, we will consider all vectors to be column vectors and transpose them when required but when casually defining vectors of variables, we will often define the vector as a row vector.

[3] Notation: $\alpha \gg 0_N$ means that each component of the vector α is positive, $\alpha \ge 0_N$ means that each component of α is nonnegative and $\alpha > 0_N$ means $\alpha \ge 0_N$ but $\alpha \ne 0_N$.

[4] The functions $M_0(x)$ and $M_1(x)$ should be defined as $M_0(x,\alpha)$ and $M_1(x,\alpha)$ since these means depend on the vector of weights α as well as the vector of nonnegative variables x that is being averaged. However, in all of our applications in this chapter, we will hold the weighting vector α constant when comparing various means and so for simplicity, we have followed the example of Hardy, Littlewood and Polya (1934; 12) and suppressed the vector α from the notation. The subscripts 0 and 1 that appear in $M_0(x)$ and $M_1(x)$ will be explained later: it will turn out that $M_0(x)$ and $M_1(x)$ are special cases of the *means of order r*, $M_r(x)$, where r is equal to 0 and 1 respectively.

(9) $M_1(x) \equiv \sum_{n=1}^{N} \alpha_n x_n$.

Of course, if each α_n equals $1/N$, then the weighted means defined by (8) and (9) reduce to the ordinary geometric and arithmetic means of the x_n.

Before we prove the main result in this section, we require a preliminary result.

(10) *Proposition 2*: Let the vector α satisfy the restrictions (7). Define the N by N matrix A by

(11) $A \equiv - \hat{\alpha} + \alpha\alpha^T$

where $\hat{\alpha}$ is an N by N diagonal matrix with nnth element α_n for n = 1,2,...,N. Then A is a negative semidefinite matrix.

Proof: It can readily be verified that A is symmetric. We need to show that for all $z \neq 0_N$, we have:

(12) $z^T A z = z^T[-\hat{\alpha} + \alpha\alpha^T] z \leq 0$ or
       $z^T\alpha\alpha^T z \leq z^T \hat{\alpha} z$ or
(13) $(\alpha^T z)^2 \leq z^T \hat{\alpha} z$ since $z^T\alpha = \alpha^T z$.

Since $\alpha \gg 0_N$, we can take the positive square root of each α_n. Let $\hat{\alpha}^{1/2}$ denote the diagonal N by N matrix which has nnth element $\alpha_n^{1/2}$ for n = 1, 2, . . ., N. Now define the N dimensional vectors x and y as follows:

(14) $x \equiv \hat{\alpha}^{1/2} 1_N$; $y \equiv \hat{\alpha}^{1/2} z$

where 1_N is an N dimensional vector of ones. Recall the Cauchy-Schwarz inequality (1). Substituting (14) into (1) yields:

(15) $(1_N^T \hat{\alpha}^{1/2} \hat{\alpha}^{1/2} z)^2 \leq (1_N^T \hat{\alpha}^{1/2} \hat{\alpha}^{1/2} 1_N)(z^T \hat{\alpha}^{1/2} \hat{\alpha}^{1/2} z)$ or
       $(1_N^T \hat{\alpha} z)^2 \leq (1_N^T \hat{\alpha} 1_N)(z^T \hat{\alpha} z)$ or
       $(\alpha^T z)^2 \leq (1_N^T \alpha)(z^T \hat{\alpha} z)$ or
       $(\alpha^T z)^2 \leq (z^T \hat{\alpha} z)$ using (7)

which is (13).   Q.E.D.

We note that to get a strict inequality in (13), we require $z \neq 0_N$ and x not proportional to y or using (14), we require $z \neq k 1_N$ for any scalar k.

Proposition 3: *Theorem of the Arithmetic and Geometric Means*:[5]

For every $x \gg 0_N$ and positive vector of weights α which satisfies (7), we have:

(16) $M_0(x) \leq M_1(x)$.

The strict inequality in (16) holds unless $x = k1_N$ for some $k > 0$ in which case (16) becomes:

(17) $M_0(k1_N) = M_1(k1_N) = k$;

i.e., the weighted geometric mean of N positive numbers is always less than the corresponding weighted arithmetic mean, unless all of the numbers are equal, in which case the means are equal.

Proof: Define the function of N variables $f(x)$ for $x \geq 0_N$ as follows:

(18) $f(x) \equiv M_0(x) - M_1(x) = \prod_{n=1}^{N} x_n^{\alpha_n} - \sum_{n=1}^{N} \alpha_n x_n$.

We wish to show that for every $x \gg 0_N$,

(19) $f(x) \leq 0$.

One way to establish (16) or (19) is to solve the following maximization problem and show that maximizing values of the objective function are equal to or less than 0:

(20) $\max_x \{f(x) : x \geq 0_N\}$.

To begin our proof, we show that points x^0 which satisfy the first order necessary conditions for maximizing the $f(x)$ defined by (18) (ignoring for now the nonnegativity restrictions $x \geq 0_N$) are such that $f(x^0) = 0$.

Partially differentiating f defined by (18) and setting the resulting partial derivatives equal to zero yields the following system of equations:

(21) $\partial f(x)/\partial x_n = \alpha_n x_n^{-1} M_0(x) - \alpha_n = 0$; $n = 1,...,N$ or
(22) $x_n = M_0(x)$; $n = 1,...,N$.

Thus if each x_n^0 equals a positive constant, $k > 0$ say, we will satisfy the first order necessary conditions (21) for maximizing $f(x)$ in the interior of the feasible region. Thus x^0 of the form:

---

[5] The equal weights case of this Theorem, where α is equal to $(1/N)1_N$, can be traced back to Euclid and Cauchy (1821; 375) according to Hardy, Littlewood and Polya (1934; 17). For alternative proofs of the general Theorem, see Hardy, Littlewood and Polya (1934; 17-21).

(23) $x^0 \equiv k1_N$; $k > 0$

are such that:

(24) $\nabla_x f(x^0) = 0_N$ and
(25) $f(x^0) = M_0(k1_N) - M_1(k1_N) = k - k = 0$.

where we have used the restrictions in (7) to derive (25).

We now calculate the matrix of second order partial derivatives of f defined by (18):

(26) $\partial^2 f(x)/\partial x_n^2 = -\alpha_n x_n^{-2} M_0(x) + \alpha_n^2 x_n^{-2} M_0(x)$; $n = 1,...,N$;
(27) $\partial^2 f(x)/\partial x_i \partial x_j = \alpha_i \alpha_j x_i^{-1} x_j^{-1} M_0(x)$; $i \neq j$.

Thus the matrix of second order partial derivatives of f evaluated at $x \gg 0_N$ can be written as follows:

(28) $\nabla^2_{xx} f(x) = \hat{x}^{-1}[-\hat{\alpha} + \alpha\alpha^T]\hat{x}^{-1} M_0(x)$

where   and   are the vectors x and a diagonalized into matrices. Note also that $M_0(x) > 0$ for any $x \gg 0_N$.

To determine the definiteness properties of the $\nabla^2_{xx} f(x)$ defined by (28), look at:

(29) $M_0(x)^{-1} z^T \nabla^2_{xx} f(x) z = z^T \hat{x}^{-1}[-\hat{\alpha} + \alpha\alpha^T]\hat{x}^{-1} z$
$= y^T[-\hat{\alpha} + \alpha\alpha^T] y$ where $y \equiv \hat{x}^{-1} z$
≤ 0

where the inequality follows using Proposition 2. The inequality in (29) will be strict provided that $y \neq 0_N$ and $y \neq k1_N$ for any k.

Now let $x^1 \gg 0_N$ be an arbitrary positive vector which is not on the equal component ray; i.e.,

(30) $x^1 \gg 0_N$ but $x^1 \neq k1_N$ for any k.

Recall that if $x^0 = k1_N$ for $k > 0$, then (25) implies $f(x^0) = 0$. Hence to establish our result, we need only show $f(x^1) < 0$.

Recall Taylor's Theorem for $n = 2$. The multivariate version of this Theorem yields the following relationship between $f(x^0)$ and $f(x^1)$ where x^0 is defined by (23) and x^1 is defined by (30): there exists a t such that $0 < t < 1$ and

(31) $f(x^1) = f(x^0) + \nabla_x f(x^0)^T (x^1 - x^0) + (1/2)(x^1 - x^0)^T \nabla^2_{xx} f((1-t)x^0 + tx^1)(x^1 - x^0)$
$= 0 + 0^T(x^1 - x^0) + (1/2)(x^1 - x^0)^T \nabla^2_{xx} f((1-t)x^0 + tx^1)(x^1 - x^0)$ using (24) and (25)

≤ 0 using $M_0(x) > 0$ and (29) for $x = x^1$ and $z = x^1 - x^0$.

In order for the inequality (31) to be strict, we require that:

(32) $\hat{x}^{-1}(x^1 - x^0) \neq k1_N$ for any k where $x \equiv (1-t)x^0 + tx^1$

or equivalently, that

(33) $x^1 - x^0 \neq k[(1-t)x^0 + tx^1]$ for any k.

Using the facts that $x^0 \neq x^1$ and $0 < t < 1$, it can be verified that (32) is true and hence the inequality in (31) is strict. Thus we have proven (16). Q.E.D.

The geometry associated with the inequalities in (33) is illustrated in Figure 1 below.

Figure 1
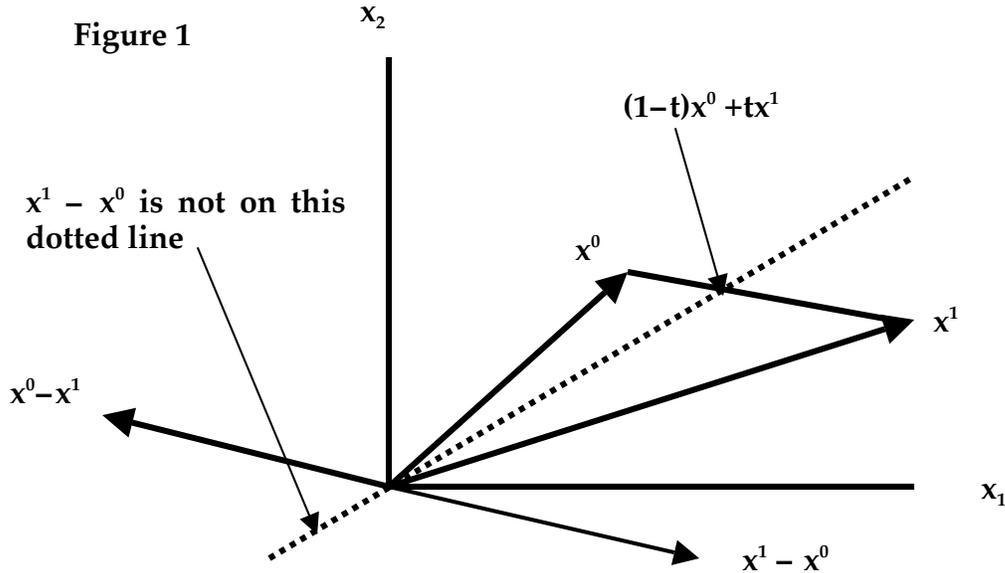
We have established that the weighted geometric mean $M_0(x)$ is strictly less than the corresponding weighted arithmetic mean $M_1(x)$ for *strictly positive* $x \gg 0_N$, unless x has all components equal, in which case the two means coincide and are equal to the common component. It is useful to extend the Theorem to cover the case where x is *nonnegative*; i.e., to cover the case where one or more components of the x vector are equal to zero. But this is easily done. In this case, $M_0(x)$ is equal to zero and $M_1(x)$ is equal to or greater than 0 (and strictly greater than 0 if $x > 0_N$). Thus we have:

(34) $0 = M_0(x) < M_1(x)$ if $x > 0_N$ and one or more components of x are equal to 0.

4. Means of Order r

As in the previous section, we again assume that the vector of weights α has positive components which sum to one; i.e., we assume α satisfies conditions (7). We assume initially that the number r is not equal to zero and the vector x has positive components and define the *weighted mean of order r* of the N numbers in x as follows:[6]

$$(33) \quad M_r(x) \equiv \left[\sum_{n=1}^{N} \alpha_n x_n^r\right]^{1/r}.$$

It can be seen that the *mean of order 1* is the weighted arithmetic mean defined earlier by (9). It is easy to verify that the means of order r are (positively) linearly homogeneous in the x variables; i.e.,[7]

$$(34) \quad M_r(\lambda x) = \lambda M_r(x) \quad\quad\quad \text{for every } x \gg 0_N \text{ and scalar } \lambda > 0.$$

The functional form defined by (33) occurs frequently in the economics literature. If we multiply $M_r(x)$ by a constant, then we obtain the CES (constant elasticity of substitution) functional form popularized by Arrow, Chenery, Minhas and Solow (1961) in the context of production theory. This functional form is also widely used as a utility function and it also used extensively when measures of income inequality are constructed.

Three other properties of the means of order r which are useful are the following ones (we assume $x \gg 0_N$ and $r \neq 0$):[8]

$$(35) \quad M_r(x_1,...,x_N) = [M_1(x_1^r,...,x_N^r)]^{1/r} \; ;$$
$$(36) \quad M_0(x_1,...,x_N) = \exp[M_1(\ln x_1,...,\ln x_N)] \; ;$$
$$(37) \quad M_{-r}(x_1,...,x_N) = 1/M_r(x_1^{-1},...,x_N^{-1})] \; .$$

Problem 5: Prove (35), (36) and (37).

We now consider the problems associated with extending the definition of $M_r(x)$ from the positive orthant (the set of x such that $x \gg 0_N$) to the nonnegative orthant (the set of x such that $x \geq 0_N$). If $r \geq 0$, there is no problem with making this extension since in this case, x_n^r tends to 0 as x_n tends to zero and $M_r(x)$ turns out to be a nice continuous function over the nonnegative orthant. But if $r < 0$, there is a problem since x_n^r tends to $+\infty$ as x_n tends to zero in this case. However, in this case, we define $M_r(x)$ to equal zero:

$$(38) \quad M_r(x) \equiv 0 \quad\quad\quad\quad \text{if } r < 0 \text{ and any component of x is 0.}$$

It turns out that with definition (38), the means of order r are continuous functions over the nonnegative orthant even if r is less than 0. To see why this is the case, consider the

---

[6] Hardy, Littlewood and Polya (1934; 12-13) refer to this family of means or averages as elementary weighted mean values and study their properties in great detail. When they consider the case where the weights are equal, they refer to the family of means as ordinary mean values.
[7] This is property (2.2.13) noted in Hardy, Littlewood and Polya (1934; 14).
[8] These properties may be found in Hardy, Littlewood and Polya (1934; 14).

case where $r = -1$, $N = 2$, $\alpha_1 = \frac{1}{2}$, $\alpha_2 = \frac{1}{2}$ and x_1 tends to 0 with $x_2 > 0$. In this case, we have for $x_1 > 0$:

$$(39)\ M_{-1}(x_1,x_2) = [(\tfrac{1}{2})x_1^{-1} + (\tfrac{1}{2})x_2^{-1}]^{-1}$$
$$= 1/[(\tfrac{1}{2})(1/x_1) + (\tfrac{1}{2})(1/x_2)]$$
$$= x_1/[(\tfrac{1}{2}) + (\tfrac{1}{2})(x_1/x_2)].$$

Taking the limit of the right hand side of (39) as x_1 approaches 0 gives us the limiting value of 0.

We will now calculate the vector of first order derivatives of $M_r(x)$ and the matrix of second order derivatives of $M_r(x)$ for $r \neq 0$ and $x \gg 0_N$.[9]

Proposition 4: The matrix of second order partial derivatives of $M_r(x)$ with respect to the components of the vector x, $\nabla^2_{xx}M_r(x)$, is negative semidefinite for $r \leq 1$ and positive semidefinite for $r \geq 1$ for $x \gg 0_N$ and $r \neq 0$.

Proof: Differentiating $M_r(x)$ with respect to x_i yields:

$$(40)\ \partial M_r(x)/\partial x_i = (1/r)[\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-1}\alpha_i\, r\, x_i^{\ r-1} = [\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-1}\alpha_i\, x_i^{\ r-1}\ ;\quad i = 1,\dots,N.$$

Differentiating (40) again with respect to x_i yields:

$$(41)\ \partial^2 M_r(x)/\partial x_i^{\ 2} = [(1/r)-1][\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-2}\alpha_i\, r x_i^{\ r-1}\alpha_i\, x_i^{\ r-1}$$
$$+ [\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-1}\alpha_i\,(r-1)x_i^{\ r-2}\ ;\qquad\qquad i = 1,\dots,N$$
$$= [r-1][\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-2}\{[\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]\alpha_i\, x_i^{\ r-2} - \alpha_i^{\ 2} x_i^{\ 2r-2}\}.$$

Differentiating (40) with respect to x_j for $j \neq i$ yields:

$$(42)\ \partial^2 M_r(x)/\partial x_i \partial x_j = [(1/r)-1][\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-2}\alpha_j\, r x_j^{\ r-1}\alpha_i\, x_i^{\ r-1}$$
$$= -(r-1)[\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-2}\, \alpha_i\alpha_j x_i^{\ r-1}x_j^{\ r-1}.$$

Using (41) and (42), we can write the matrix of second order partial derivatives of $M_r(x)$ as follows:

$$(43)\ \nabla^2_{xx}M_r(x) = (r-1)[\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]^{(1/r)-2}\{[\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]\,\hat{x}^{(r/2)-1}\hat{\alpha}\hat{x}^{(r/2)-1} - \hat{x}^{r-1}\alpha\alpha^T\hat{x}^{r-1}\}$$

where $\hat{x}^{r-1}$ is a diagonal matrix which has nnth element equal to $x_n^{\ r-1}$ and $\hat{x}^{(r/2)-1}$ is a diagonal matrix which has diagonal elements equal to $x_n^{(r/2)-1}$ for $n = 1,\dots,N$.

We now want to show that the matrix A defined as

$$(44)\ A \equiv [\textstyle\sum_{n=1}^N \alpha_n x_n^{\ r}]\,\hat{x}^{(r/2)-1}\hat{\alpha}\hat{x}^{(r/2)-1} - \hat{x}^{r-1}\alpha\alpha^T\hat{x}^{r-1}$$

---

[9] We have already calculated these derivatives for $M_0(x)$ in Proposition 3.

is positive semidefinite. A will be positive definite if for every vector z, we have $z^T A z \geq 0$ or

$$(45) \; [\textstyle\sum_{n=1}^{N} \alpha_n x_n^{\;r}] z^T \hat{x}^{(r/2)-1} \hat{\alpha} \hat{x}^{(r/2)-1} z \geq z^T \hat{x}^{r-1} \alpha \alpha^T \hat{x}^{r-1} z \qquad \text{or}$$
$$(\alpha^T \hat{x}^{r-1} z)^2 \leq [\textstyle\sum_{n=1}^{N} \alpha_n x_n^{\;r}] z^T \hat{x}^{(r/2)-1} \hat{\alpha} \hat{x}^{(r/2)-1} z.$$

In order to establish (45), note that:

$$
\begin{aligned}
(46) \; (\alpha^T \hat{x}^{r-1} z)^2 &= (1_N^T \hat{\alpha} \hat{x}^{r-1} z)^2 && \text{using } \alpha = \hat{\alpha} 1_N \\
&= (1_N^T \hat{\alpha}^{1/2} \hat{\alpha}^{1/2} \hat{x}^{r/2} \hat{x}^{(r/2)-1} z)^2 && \\
&= (1_N^T \hat{\alpha}^{1/2} \hat{x}^{r/2} \hat{x}^{(r/2)-1} \hat{\alpha}^{1/2} z)^2 && \text{since diagonal matrices commute} \\
&= (u^T v)^2 && \text{with } u^T \equiv 1_N^T \hat{\alpha}^{1/2} \hat{x}^{r/2} \text{ and } v \equiv \hat{x}^{(r/2)-1} \hat{\alpha}^{1/2} z \\
&\leq (u^T u)(v^T v) && \text{using the Cauchy-Schwarz inequality} \\
&= (1_N^T \hat{\alpha}^{1/2} \hat{x}^{r/2} \hat{x}^{r/2} \hat{\alpha}^{1/2} 1_N)(z^T \hat{\alpha}^{1/2} \hat{x}^{(r/2)-1} \hat{x}^{(r/2)-1} \hat{\alpha}^{1/2} z) && \\
&= (1_N^T \hat{\alpha} \hat{x}^r 1_N)(z^T \hat{x}^{(r/2)-1} \hat{\alpha} \hat{x}^{(r/2)-1} z) && \text{since diagonal matrices commute} \\
&= (\alpha^T \hat{x}^r 1_N)(z^T \hat{x}^{(r/2)-1} \hat{\alpha} \hat{x}^{(r/2)-1} z) && \text{since } 1_N^T \hat{\alpha} = \alpha^T \\
&= [\textstyle\sum_{n=1}^{N} \alpha_n x_n^{\;r}] z^T \hat{x}^{(r/2)-1} \hat{\alpha} \hat{x}^{(r/2)-1} z &&
\end{aligned}
$$

which establishes (44); i.e., A is positive semidefinite. Returning to (43), we have:

$$(47) \; \nabla^2_{xx} M_r(x) = (r-1)[\textstyle\sum_{n=1}^{N} \alpha_n x_n^{\;r}]^{(1/r)-2} A.$$

Since A is positive semidefinite and $[\sum_{n=1}^{N} \alpha_n x_n^{\;r}]^{(1/r)-2}$ is positive since we have assumed that $x \gg 0_N$, we see that $\nabla^2_{xx} M_r(x)$ is positive semidefinite if $r \geq 1$ and is negative semidefinite if $r \leq 1$. \hfill Q.E.D.

The above Proposition shows that $M_r(x)$ is a *concave function* of x over the positive orthant if $r \leq 1$ and a *convex function* of x if $r \geq 1$.

5. Schlömilch's Inequality

In this section, we show that if $x \neq k 1_N$, then $M_r(x)$ increases as the parameter r increases. In order to do this, we require a preliminary inequality.

Proposition 5: Let $\alpha \gg 0_N$, $\alpha^T 1_N = 1$ and $y \gg 0_N$. Then

$$(48) \; f(y) \equiv \alpha^T y \ln(\alpha^T y) - \textstyle\sum_{n=1}^{N} \alpha_n y_n \ln y_n \leq 0$$

and the inequality is strict if $y \neq k 1_N$.

Proof: We use the same technique of proof that we used in proving the Theorem of the Arithmetic and Geometric Mean. We start out by attempting to maximize f(y) over the

positive orthant.   The first order necessary conditions for solving this maximization problem are:

(49) $\partial f(y)/\partial y_n = \alpha_n \ln(\alpha^T y) + (\alpha^T y)(\alpha^T y)^{-1}\alpha_n - \alpha_n \ln y_n - \alpha_n y_n /y_n$; $\quad\quad\quad$ n = 1,...,N

$\quad\quad = \alpha_n \ln(\alpha^T y) - \alpha_n \ln y_n$

$\quad\quad = 0.$

Equations (49) imply that $\ln y_n = \ln(\alpha^T y)$ for n = 1, . . ., N. Thus solutions to (49) have the form:

(50) $y^0 = k1_N;$ $\quad\quad$ k > 0.

Note that

(51) $\nabla_y f(y^0) = 0_N$ $\quad\quad$ and

(52) $f(y^0) = \alpha^T k1_N \ln(\alpha^T k1_N) - \sum_{n=1}^N \alpha_n k \ln k = k \ln k - k \ln k = 0$

where we have used $\alpha^T 1_N = 1$. Now differentiate equations (49) again in order to obtain the following second order partial derivatives of f:

(53) $f_{ii}(y) = \alpha_i(\alpha^T y)^{-1}\alpha_i - \alpha_i y_i^{-1}$; $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ i = 1,...,N ;

(54) $f_{ij}(y) = \alpha_i(\alpha^T y)^{-1}\alpha_j$; $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ i ≠ j.

Equations (53) and (54) can be rewritten in matrix form as follows:

(55) $\nabla^2 f(y) = - \hat{y}^{-1/2} \hat{\alpha} \, \hat{y}^{-1/2} + (\alpha^T y)^{-1}\alpha\alpha^T$

$y_i^{-1/2}$ for i = 1, 2, . . ., N. We now show that $\nabla^2 f(y)$ is negative semidefinite; i.e., we want to show that for all z:

(56) $- z^T \hat{y}^{-1/2} \hat{\alpha} \, \hat{y}^{-1/2} z + z^T(\alpha^T y)^{-1}\alpha\alpha^T z \le 0$ $\quad\quad\quad\quad\quad\quad$ or

(57) $(\alpha^T z)^2 \le (\alpha^T y) z^T \hat{y}^{-1/2} \hat{\alpha} \, \hat{y}^{-1/2} z$ $\quad\quad\quad\quad\quad\quad\quad\quad$ for all z ≠ 0_N.

To prove (57), we will use the Cauchy-Schwarz inequality:

(58) $(\alpha^T z)^2 = (z^T \hat{y}^{-\frac{1}{2}} \hat{y}^{\frac{1}{2}} \hat{\alpha} 1_N)^2$ $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad$ since $\alpha = \hat{\alpha} 1_N$

$\quad\quad = (z^T \hat{y}^{-\frac{1}{2}} \hat{y}^{\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} 1_N)^2$

$\quad\quad = (z^T \hat{\alpha}^{\frac{1}{2}} \hat{y}^{-\frac{1}{2}} \hat{y}^{\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} 1_N)^2$ $\quad\quad\quad\quad\quad$ since diagonal matrices commute

$\quad\quad \le (z^T \hat{\alpha}^{\frac{1}{2}} \hat{y}^{-\frac{1}{2}} \hat{y}^{-\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} z) (1_N^T \hat{\alpha}^{\frac{1}{2}} \hat{y}^{\frac{1}{2}} \hat{y}^{\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} 1_N)$

$\quad\quad$ using the Cauchy-Schwarz inequality with x ≡ $\hat{y}^{-1/2} \hat{\alpha}^{1/2} z$ and y ≡ $\hat{y}^{1/2} \hat{\alpha}^{1/2} 1_N$

$\quad\quad = (z^T \hat{y}^{-\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} \hat{y}^{-\frac{1}{2}} z) (1_N^T \hat{\alpha}^{\frac{1}{2}} \hat{\alpha}^{\frac{1}{2}} \hat{y}^{\frac{1}{2}} \hat{y}^{\frac{1}{2}} 1_N)$

$$= (z^T \hat{y}^{-\frac{1}{2}} \hat{\alpha} \, \hat{y}^{-\frac{1}{2}} z) \, (1_N^T \hat{\alpha} \, \hat{y} \, 1_N)$$

$$= (z^T \hat{y}^{-\frac{1}{2}} \hat{\alpha} \, \hat{y}^{-\frac{1}{2}} z) \, (\alpha^T y)$$

$$\text{since } \alpha = \hat{\alpha} \, 1_N \text{ and } y = \hat{y} \, 1_N$$

which is (57). The inequality (58) will be strict provided that $\hat{y}^{-1/2} \hat{\alpha}^{1/2} z$ and $y \equiv \hat{y}^{-1/2} \hat{\alpha}^{1/2} 1_N$ are not proportional or $\hat{\alpha}^{1/2} z$ and $\hat{\alpha}^{1/2} 1_N$ are not proportional or z and 1_N are not proportional or provided that $z \neq k1_N$.

Now let y^1 be a positive vector that does not have all components equal; i.e.,

(59) $y^1 >> 0_N$ but $y^1 \neq k1_N$ for any k.

We need only show $f(y^1) < 0$ to complete the proof. Apply Taylor's Theorem to the f defined by (48) and the y^0 and y^1 defined by (50) and (59). Thus there exists a t such that $0 < t < 1$ and

(60) $f(y^1) = f(y^0) + \nabla_y f(y^0)^T (y^1 - y^0) + (1/2)(y^1 - y^0)^T \nabla^2_{yy} f((1-t)y^0 + ty^1)(y^1 - y^0)$

$= 0 + 0^T (y^1 - y^0) + (1/2)(y^1 - y^0)^T \nabla^2_{yy} f((1-t)y^0 + ty^1)(y^1 - y^0)$ using (51) and (52)

≤ 0

where the inequality follows using (58) which implies that $\nabla^2_{yy} f((1-t)y^0 + ty^1)$ is negative semidefinite.

In order for the inequality (60) to be strict, we require that $y^1 - y^0 = y^1 - k1_N$ is not proportional to 1_N. This is true using definition (59); i.e., if $y^1 - k1_N$ were proportional to 1_N, then we would deduce that y^1 was also proportional to 1_N, which contradicts definition (59). Q.E.D.

Now we are ready for the main result in this section.

Proposition 6: *Schlömilch's* (1858) *Inequality*:[10] Let $x >> 0_N$ but $x \neq k1_N$ for any $k > 0$ and let $r < s$. As usual, we assume the weighting vector α satisfies (7). Then

(61) $M_r(x) < M_s(x)$.

If $x = k1_N$ for some $k > 0$, then $M_r(x) = M_s(x) = k$.

Proof: The second part of the theorem is easily verified. The first part of the theorem, (61), will be true if we can show that $M_r(x)$ is a monotonically increasing function of r or equivalently if we can show that:[11]

---

[10] See Hardy, Littlewood and Polya (1934; 26) for alternative proofs of this result.

[11] This is not quite equivalent to the desired result: we still have to deal with the cases where r or s are equal to 0; i.e., we cannot differentiate $M_r(x)$ defined by (33) with respect to r when $r = 0$.

(62) $\partial \ln M_r(x)/\partial r > 0$ for all $r \neq 0$, $x >> 0_N$, $x \neq k1_N$ for any $k > 0$, $\alpha >> 0_N$ and $\alpha^T 1_N = 1$.

Recall that $\partial c^r/\partial r = \partial e^{r\ln c}/\partial r = e^{r\ln c} \ln c = c^r \ln c$ so that using definition (33), it can be verified that the inequality (62) is equivalent to:

(63) $\partial \ln M_r(x)/\partial r = - r^{-2}\ln[\sum_{n=1}^{N} \alpha_n x_n^r] + r^{-1}[\sum_{n=1}^{N} \alpha_n x_n^r]^{-1}[\sum_{n=1}^{N} \alpha_n x_n^r \ln x_n] > 0$ or

(64) $r^{-1}[\sum_{n=1}^{N} \alpha_n x_n^r \ln x_n] > r^{-2}[\sum_{n=1}^{N} \alpha_n x_n^r] \ln[\sum_{n=1}^{N} \alpha_n x_n^r]$.

Since $r \neq 0$, $r^2 > 0$ and $r^{-2} > 0$. Thus

(65) $r^{-2}[\sum_{n=1}^{N} \alpha_n x_n^r] \ln[\sum_{n=1}^{N} \alpha_n x_n^r] \leq r^{-2}[\sum_{n=1}^{N} \alpha_n x_n^r \ln x_n^r]$ using (48) with $y_n \equiv x_n^r$

$\qquad\qquad\qquad\qquad = r^{-2}[\sum_{n=1}^{N} \alpha_n x_n^r\, r\, \ln x_n]$ using $\ln x_n^r = r \ln x_n$

$\qquad\qquad\qquad\qquad = r^{-1}[\sum_{n=1}^{N} \alpha_n x_n^r \ln x_n]$

and (65) is a weak version of (64). But the inequality (65) is strict provided that the $y_n = x_n^r$ are not all equal. This is the case since we have assumed $x \neq k1_N$ and thus (65) is a strict inequality.

We still need to establish (61) when r or s equal 0. We first consider the case where $s > r = 0$. Let $x >> 0_N$ with $x \neq k1_N$. Then we have:

(66) $[M_0(x)]^s = [\prod_{n=1}^{N} x_n^{\alpha_n}]^s$ using definition (8)

$\qquad\qquad = M_0(x_1^s,...,x_N^s)$

$\qquad\qquad < M_1(x_1^s,...,x_N^s)$ by Proposition 3 since not all of the x_n^s are equal

$\qquad\qquad = M_s(x_1,...,x_N)^s$ using definition (33).

Since $s > 0$, taking the 1/s root of both sides of (66) will preserve the inequality which establishes (61) for $r = 0 < s$.

We now consider (61) when $-r < 0 = s$. Again, let $x >> 0_N$ with $x \neq k1_N$. Then we have:

(67) $M_{-r}(x) = 1/M_r(x_1^{-1},...,x_N^{-1})$ using property (37)

$\qquad\qquad < 1/M_0(x_1^{-1},...,x_N^{-1})$ using $r > 0$, $x \neq k1_N$ and (66) which implies that

$\qquad\qquad\qquad\qquad\qquad M_r(x_1^{-1},...,x_N^{-1}) > M_0(x_1^{-1},...,x_N^{-1})$

$\qquad\qquad = M_0(x)$ using definition (8).   Q.E.D.

The above Theorem shows that the *weighted harmonic mean* of N positive numbers, $x_1,...,x_N$, will always be equal to or less than the corresponding weighted arithmetic mean; i.e., we have for $x >> 0_N$:

(68) $M_{-1}(x) \leq M_1(x)$ or

(69) $[\sum_{n=1}^{N} \alpha_n x_n^{-1}]^{-1} \leq \sum_{n=1}^{N} \alpha_n x_n$

and the inequality (69) is strict provided that the x_n are not all equal to the same positive number.

What happens if one or more of the components of the x vector are equal to 0? Using the continuity of the functions $M_r(x)$ over the nonnegative orthant, it can be seen that (61) will still hold as a weak inequality. It should be kept in mind that if $r \leq 0$ and any component of x is 0, then (38) implies that $M_r(x)$ is equal to 0.

Problems:

6. Suppose a Statistical Agency collects price quotes on a "homogeneous" commodity (e.g. red potatoes) from N outlets during periods 0 and 1. Denote the vector of period t price quotes by $p^t \equiv [p_1^t,...,p_N^t]$ for $t = 0, 1$. An *elementary price index* $P(p^0, p^1)$ is a function of 2N variables that aggregates this micro information on potatoes into an aggregate price index for potatoes that will be a component of the overall consumer price index (CPI). Examples of widely used functional forms for P are the *Carli* (1764) and *Jevons* (1865) formulae defined by (i) and (ii) below:

(i) $P_C(p^0,p^1) \equiv \sum_{n=1}^{N} (1/N)(p_n^1/p_n^0)$

which is the equally weighted *arithmetic* mean of the N price ratios;

(ii) $P_J(p^0,p^1) \equiv [\prod_{n=1}^{N} (p_n^1/p_n^0)]^{1/N}$

which is the equally weighted *geometric* mean of the N price ratios.

A very useful property for an elementary price index to satisfy is the *time reversal test*:

(iii) $P(p^0,p^1) \, P(p^1,p^0) = 1$;

i.e., suppose prices in period 1 reverted back to the base period prices p^0. Under these conditions, we should end up at our starting point.

(a) Show that $P_J(p^0,p^1)$ satisfies the time reversal test.
(b) Show that $P_C(p^0,p^1)$ has an upward bias; i.e., show that if $p^1 \neq kp^0$, then

(iv) $P_C(p^0,p^1) \, P_C(p^1,p^0) > 1$.

Hint: You may find (69) useful.

Comment: Many Statistical Agencies are still using the biased Carli formula to aggregate their price quotes at the lowest level of aggregation. However, in the past decade, several countries (Canada, the U.S. and the member countries of the EU for their harmonized indexes) have switched to the Jevons formula. The use of P_C rather than P_J is thought to

have generated an upward bias in the CPI in the 0.1- 0.4% per year range. Fisher (1922; 66 and 383) seems to have been the first to establish the upward bias of the Carli index and he made the following observations on its use by statistical agencies:

"In fields other than index numbers it is often the best form of average to use. But we shall see that the simple arithmetic average produces one of the very worst of index numbers. And if this book has no other effect than to lead to the total abandonment of the simple arithmetic type of index number, it will have served a useful purpose." Irving Fisher (1922; 29-30).

7. A *general mean function*, $M(x)$, is a function of N variables, defined for $x \gg 0_N$ that has the following three properties:

(i) $M(k1_N) = k$ for $k > 0$ (*mean value property*);
(ii) $M(x)$ is a *continuous* function; and
(iii) $M(x)$ is *increasing* in its components; i.e., if $x^1 < x^2$, then $M(x^1) < M(x^2)$.

It is easy to see that the weighted means of order r, $M_r(x)$ defined by (33), satisfy properties (i) and (ii). Show that they also satisfy property (iii).

Hint: Show that $\partial M_r(x)/\partial x_n > 0$ for $r \neq 0$.

8. $M(x)$ is a *symmetric mean* if M is a mean and has the following property:

(iv) $M(Px) = M(x)$ where Px is a permutation of the components of x. Are the means of order r symmetric means? If not, what conditions on α will make $M_r(x)$ a symmetric mean?

9. $M(x)$ is a *homogeneous mean* if it is a mean and satisfies the following additional property:

(v) $M(\lambda x) = \lambda M(x)$ for all $\lambda > 0$, $x \gg 0_N$.

If $M(x)$ is a homogeneous mean, show that it also satisfies the following property:

(vi) $\alpha \equiv \min_n \{x_n : n = 1,...,N\} \leq M(x) \leq \max_n \{x_n : n = 1,...,N\} \equiv \beta$.

This result is due to Eichhorn and Voeller (1976; 10). *Hint*: $\alpha 1_N \leq x \leq \beta 1_N$. Note that properties (ii) and (iii) for a mean $M(x)$ imply that the following property also holds for M:

(vii) $M(x^1) \leq M(x^2)$ if $x^1 \leq x^2$.

6. L'Hospital's Rule and Logarithmic Means

In this section, we show that the weighted geometric mean, $M_0(x)$, is a limiting case of the corresponding weighted mean of order r, $M_r(x)$, as r tends to zero. Before we do this, we require a preliminary result.

Proposition 7: *L'Hospital's* (1696) *Rule*:[12] Suppose f(z) and g(z) are once continuously differentiable functions of one variable z around an interval including z = b. In addition, suppose f(b) = g(b) = 0 but g'(b) ≠ 0. Then

(70) $\lim_{z \to b} f(z)/g(z) = f'(b)/g'(b)$.

Proof: Let z be close to b but z ≠ b. Then by the Mean Value Theorem, there exist z^* and z^{**} between z and b such that:

$$(71)\ f(z) = f(b) + f'(z^*)(z - b)$$
$$= f'(z^*)(z - b) \qquad\qquad \text{since } f(b) = 0;$$
$$(72)\ g(z) = g(b) + g'(z^{**})(z - b)$$
$$= g'(z^{**})(z - b) \qquad\qquad \text{since } g(b) = 0.$$

Taking the ratio of (71) to (72) and using the assumptions that g'(b) ≠ 0 and that the derivative function g'(z) is continuous, we can deduce that $g'(z^{**}) \neq 0$ using if z is close enough to b and hence for z − b ≠ 0 and z close to b, we get:

(73) $f(z)/g(z) = f'(z^*)/g'(z^{**})$.

Now take limits on both sides of (73) as z approaches b. Since z^* and z^{**} are between z and b, z^* and z^{**} will tend to b and thus (70) follows, since both f' and g' are assumed to be continuous functions. Q.E.D.

The following problems illustrate a few of the uses of L'Hospital's Rule.

Problems:

10. If x > 0, show that $\lim_{r \to 0} (x^r - 1)/r = \ln x$.

Hint: Use L'Hospital's Rule with $f(r) \equiv x^r - 1$ and $g(r) \equiv r$. Note that if $h(r) = x^r = e^{r\ln x}$, then $h'(r) = e^{r\ln x} \ln x = x^r \ln x$.

Comment: The function $(x^r - 1)/r$ is known as the *Box-Cox transformation* and it is widely used in statistics and econometrics as well as in the study of choice under uncertainty.

11. The *logarithmic mean*, $L(x_1, x_2)$ of two positive numbers $x_1 > 0$ and $x_2 > 0$, is defined as follows:

---

[12] See Rudin (1953; 82) for a proof of this result.

(i) $L(x_1,x_2) \equiv [x_1 - x_2]/[\ln x_1 - \ln x_2]$ if $x_1 \neq x_2$;

$\qquad \equiv x_2$ if $x_1 = x_2$.

Show that if $0 < x_1 < x_2$, then

(ii) $\lim x_1 \rightarrow x_2 \, L(x_1,x_2) = x_2$.

Hint: Define $f(x_1) \equiv x_1 - x_2$ and $g(x_1) \equiv \ln x_1 - \ln x_2$ and apply L'Hospital's Rule.

Comment: This result establishes the continuity of $L(x_1,x_2)$ over the positive orthant.

12. Refer to problems 7-11 above and show that $L(x_1, x_2)$ defined in Problem 11 above is a homogeneous symmetric mean.

Hint: The definition of $L(x_1,x_2)$ in Problem 11 establishes property (i) in Problem 7. Problem 11 establishes the validity of property (ii) in Problem 7. To prove property (iii), just show $\partial L(x_1, x_2)/\partial x_n > 0$ for $n = 1,2$ (you can assume $x_1 \neq x_2$). In the case of only two variables, the symmetry property (iv) is just $L(x_1, x_2) = L(x_2, x_1)$ which you can verify. Finally, verify the homogeneity property, (v), that was defined in problem 9.

Comment: The logarithmic mean (sometimes called the Vartia mean) plays a key role in index number theory; see Vartia (1976) and Diewert (1978).

7. Additional Properties of Means of Order r

Proposition 8 below justifies our notation, $M_0(x)$, for the weighted geometric mean since this Proposition shows that $M_0(x)$ is a limiting case of $M_r(x)$ as r tends to 0.

Proposition 8:[13] The limiting case of the weighted mean of order r, $M_r(x)$, as r tends to 0 is the weighted geometric mean, $M_0(x)$; i.e., for $x \gg 0_N$, $\alpha \gg 0_N$, $\alpha^T 1_N = 1$:

(74) $\lim_{r \rightarrow 0} M_r(x) = M_0(x)$.

Proof: Proposition 6 above showed that $M_r(x)$ is a nondecreasing function of r. Since $M_r(x)$ is a homogeneous mean, Problem 9 above shows that $M_r(x)$ is bounded from above and below; i.e., for all $r \neq 0$;

(75) $\min_n \{x_n : n = 1,...,N\} \leq M_r(x) \leq \max_n \{x_n : n = 1,...,N\}$.

The fact that $M_r(x)$ is a nondecreasing function of r and is also bounded from above and below is sufficient to imply the existence of $\lim_{r \rightarrow 0} M_r(x)$ and also that

(76) $\lim_{r \rightarrow 0} \ln M_r(x) = \ln [\lim_{r \rightarrow 0} M_r(x)]$.

---

[13] See Hardy, Littlewood and Polya (1934; 15) for a proof of this result.

We now compute $\ln M_r(x)$ for $r \neq 0$:

(77) $\ln M_r(x) = (1/r)\ln[\sum_{n=1}^{N} \alpha_n x_n^{r}] \equiv f(r)/g(r)$

where $g(r) \equiv r$ and $f(r) \equiv \ln[\sum_{n=1}^{N} \alpha_n x_n^{r}]$. Note that:

(78) $g(0) = 0$;
(79) $f(0) = \ln[\sum_{n=1}^{N} \alpha_n (x_n)^0] = \ln[\sum_{n=1}^{N} \alpha_n 1] = \ln 1 = 0$ using $\sum_{n=1}^{N} \alpha_n = 1$.

Now calculate the derivatives of $f(r)$ and $g(r)$ and evaluate them at $r = 0$:

(80) $g'(r) = 1$ and hence
(81) $g'(0) = 1$.
(82) $f'(r) = [\sum_{n=1}^{N} \alpha_n x_n^{r}]^{-1}\sum_{n=1}^{N} \alpha_n x_n^{r} \ln x_n$ and hence
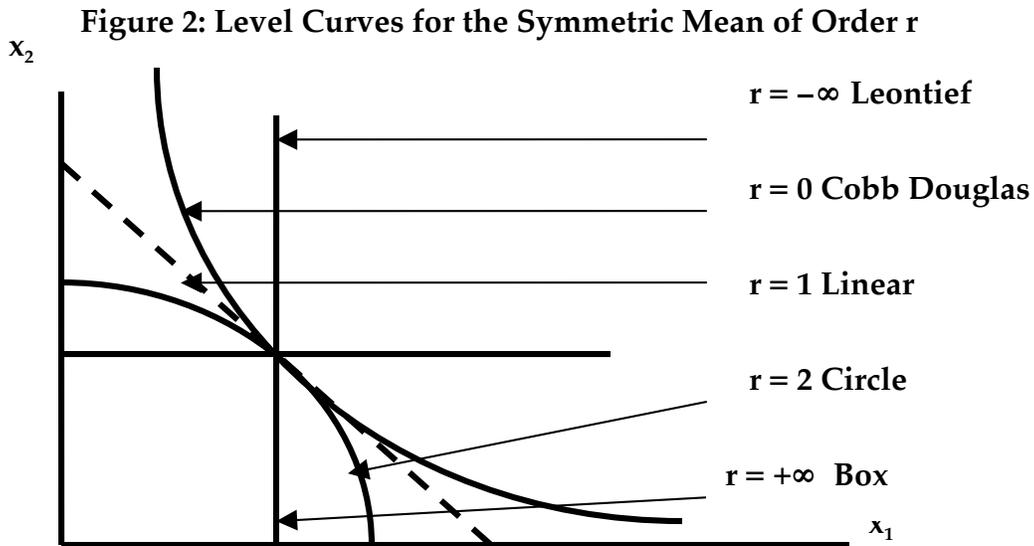(83) $f'(0) = [\sum_{n=1}^{N} \alpha_n]^{-1}\sum_{n=1}^{N} \alpha_n \ln x_n$
       $= \sum_{n=1}^{N} \alpha_n \ln x_n$ using $\sum_{n=1}^{N} \alpha_n = 1$.

Now apply L'Hospital's Rule to (77) when $r = 0$. The resulting equation is:

(84) $\lim_{r \to 0} \ln M_r(x) = f'(0)/g'(0) = \sum_{n=1}^{N} \alpha_n \ln x_n$ using (81) and (83).

We can exponentiate both sides of (84) and deduce that (74) holds.   Q.E.D.

When $N = 2$ and $\alpha_1 = \alpha_2 = 1/2$, we can graph the level curves $\{(x_1, x_2): M_r(x_1,x_2) = 1\}$ for various values of r; see Figure 2 below.

Figure 2: Level Curves for the Symmetric Mean of Order r

We conclude with some results on limiting cases of $M_r(x)$ as r tends to plus or minus infinity. The results in Proposition 9 are used in Figure 2.

Proposition 9: Hardy, Littlewood and Polya (1934; 15): The limits of $M_r(x)$ as r tends to plus or minus infinity are as follows:

(85) $\lim_{r\to\infty} M_r(x) = \max_n \{x_n : n = 1,\ldots,N\}$;
(86) $\lim_{r\to -\infty} M_r(x) = \min_n \{x_n : n = 1,\ldots,N\}$.

Proof: Let $x > 0_N$ and let $x_k = \max_n \{x_n : n = 1,\ldots,N\}$. Then using the results in Problem 9, we have:

(87) $M_r(x) \le x_k$.

Since the x_n are nonnegative and the α_n are positive, we have:

(88) $\alpha_k x_k^r \le \sum_{n=1}^N \alpha_n x_n^r$.

Now take the rth root of both sides of (88). If $r > 0$, the inequality is preserved and so we have in this case:

(89) $(\alpha_k)^{1/r} x_k \le [\sum_{n=1}^N \alpha_n x_n^r]^{1/r} = M_r(x)$.

Now take the limit of both sides of (89) as r tends to plus infinity and since $(\alpha_k)^{1/r}$ tends to $(\alpha_k)^0 = 1$, we find that

(90) $x_k \le \lim_{r\to\infty} M_r(x)$.

It can be seen that (87) and (90) imply (85).

Now consider (86). If one or more of the x_n are zero, then $M_r(x)$ equals 0 for all $r < 0$; recall (38) above. Hence if one or more of the x_n are 0, then it is easy to verify that (86) holds. Thus we consider the case where $x \gg 0_N$ and let $x_k = \min_n \{x_n : n = 1,\ldots,N\}$. By Problem 9, we have:

(91) $x_k \le M_r(x)$.

Since the x_n are positive and the α_n are positive, we again have (88) but now we assume that $r < 0$, so that when we take the rth root of each side of (88), the inequality is reversed and so we have:

(92) $(\alpha_k)^{1/r} x_k \ge [\sum_{n=1}^N \alpha_n x_n^r]^{1/r} = M_r(x)$.

Now take the limit of both sides of (92) as r tends to minus infinity and since $(\alpha_k)^{1/r}$ tends to $(\alpha_k)^0 = 1$, we find that

(93) $x_k \geq \lim_{r \to -\infty} M_r(x) \geq x_k$

where the last inequality follows using (91). It can be seen that (91) and (93) imply (86).

<div align="right">Q.E.D.</div>

8. Summary of Methods used to Establish Inequalities

A careful look at the methods of proof that we have used to establish the validity of various inequalities will show that we have basically used 3 methods:

- Transform the given inequality into a known inequality using ordinary algebra.
- Transform the given inequality into the form $f(x) \leq 0$ for the domain of definition for the inequality, say $x \in S$, and show that x^* which solve max $_x$ $\{f(x) : x \in S\}$ are such that $f(x^*) \leq 0$.
- Consider the case where the last method leads to a twice continuously differentiable objective function $f(x)$ which has the following properties: (a) there exist points x^* such that $f(x^*) = 0$ and $\nabla f(x^*) = 0_N$; (b) the domain of definition set S is convex and (c) $\nabla^2 f(x)$ is negative semidefinite for each $x \in S$. Then in this case, we can use Taylor's Theorem for $n = 2$ and establish the desired result, $f(x) \leq f(x^*) = 0$ for all $x \in S$.

It turns out that the three main inequalities that were established in this chapter (the Cauchy Schwarz Inequality, the Theorem of the Arithmetic and Geometric Means and Schlömilch's Inequality) have many applications in all branches of applied economics.

Problems

13. Let $\phi(z)$ be a monotonically increasing, continuous function of one variable that is defined for $z > 0$ so that the inverse function for ϕ, $\phi^{-1}(y)$, is also a monotonically increasing, continuous function of y for all y's belonging to the range of ϕ. As usual, define the vector of weights $\alpha \equiv [\alpha_1,...,\alpha_N]$ which satisfies:

(i) $\alpha \gg 0_N$ and (ii) $1^T \alpha = 1$.

We use the function ϕ in order to define the following *quasilinear mean* for all $x \gg 0_N$:[14]

(iii) $M_\phi(x) \equiv \phi^{-1}[\sum_{n=1}^{N} \alpha_n \phi(x_n)]$.

---

[14] Eichhorn (1978; 32) used this terminology. The axiomatic properties of this type of mean were first explored by Nagumo (1930), Kolmogoroff (1930) and Hardy, Littlewood and Polya (1934; 65-69). Kolmogoroff used the term "regular mean" while Hardy, Littlewood and Polya (1934; 65) used the awkward term "mean value with an arbitrary function". Diewert (1993; 358-359) used the term "separable mean". Kolmogoroff, Nagumo and Diewert studied only the equally weighted case.

Show that $M_\phi(x)$ defined by (iii) is a general mean; i.e., it satisfies properties (i)-(iii) listed in Problem 7 above.

Hint: You do not have to prove part (ii), continuity, which is obvious.

Comment: Note that if $\phi(z) \equiv z^r$ for $r > 0$, then $M_\phi(x)$ reduces to the weighted mean of order r, $M_r(x)$ and if $\phi(z) \equiv \ln z$, then $M_\phi(x)$ reduces to the weighted geometric mean, $M_0(x)$. Hardy, Littlewood and Polya (1934; 68) show that if we require $M_\phi(x)$ to be a homogeneous mean [15], then essentially, $M_\phi(x)$ must be a mean of order r.[16]

14. Find a general mean function, $M(x_1, x_2)$, which is *not* a quasilinear mean of the type defined in Problem 13.

References

Balk, B.M. (1995), "Axiomatic Price Index Theory: A Survey", *International Statistical Review* 63, 69-93.

Arrow, K.J., H.B. Chenery, B.S. Minhas and R.M. Solow, (1961), "Capital-Labour Substitution and Economic Efficiency", *Review of Economics and Statistics* 63, 225-250.

Bartelsman, E. J. (1995), "Of Empty Boxes: Returns to Scale Revisited," *Economics Letters* 49, 59-67.

Cauchy, A.L., (1821), Cours d'analyse de l'École Royal Polytechnique: Analyse algébrique: Paris.

Diewert, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation", *Econometrica* 46, 883-900.

Diewert, W.E. (1993), "Symmetric Means and Choice under Uncertainty", pp. 355-433 in *Essays in Index Number Theory*, Volume 1 (W.E. Diewert and A.O. Nakamura editors), Amsterdam: North-Holland.

Diewert, W.E. (1998), "Index Number Issues in the Consumer Price Index", *Journal of Economic Perspectives* 12:1 (Winter), 47-58.

Eichhorn, W. (1978), *Functional Equations in Economics*, Reading, MA: Addison-Wesley Publishing Company.

---

[15] Recall Property (v) in Problem 9.
[16] For proofs of this result in the case of equally weighted or symmetric separable means, see Nagumo (1930) and Diewert (1993; 381).

Eichhorn, W. and J. Voeller (1976), *Theory of the Price Index*, Lecture Notes in Economics and Mathematical Systems, Vol. 140, Berlin: Springer-Verlag.

Fisher, I. (1922), *The Making of Index Numbers*, Houghton-Mifflin, Boston.

Hardy, G.H., J.E. Littlewood and G. Polya, (1934), *Inequalities*, Cambridge, England: Cambridge University Press.

Kendall, M.G. and A.S. Stuart (1967), *The Advanced Theory of Statistics: Volume 2: Inference and Relationship*, Second Edition, New York: Hafner Publishing Co.

Kolmogoroff, A. (1930), "Sur la notion de la moyenne", *Atti della Reale Academia nazionale dei Lincei* 12(6), 388-391.

L'Hospital (1696), *L'analyse des infiniment petits pour l'intelligence des lignes courbes*, Paris.

Nagumo, M. (1930), "Űber eine Klasse der Mittelwerte", *Japanese Journal of Mathematics* 7, 71-79.

Rudin, W., (1953), Principles of Mathematical Analysis, New York: McGraw-Hill Book Co.

Schlömilch, O., (1858), "Über Mïttelgrössen verschiedener Ordnungen", *Zeitschrift für Mathematik und Physik* 3, 308-310.

Schwarz, H.A., (1885), "Über ein die Flächen Kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung", *Acta Societatis scientiarum Fennicae* 15, 315-362.

Vartia, Y.O. (1976), "Ideal Log-Change Index Numbers", *Scandinavian Journal of Statistics* 3, 121-126.