A Conjugate Inequality for General Means with Applications to Extremum Problems

CARL D. EBEN and JOHN R. FERRON

University of Delaware, Newark, Delaware

Inequalities are derived for a general class of convex-functional means which may contain negative weights. Specific cases include the arithmetic-mean–geometric-mean inequality and other power-mean inequalities. These inequalities make possible the solution of a wider class of extremum problems than are susceptible to the classical means with positive weights. In particular it is shown that the geometric programming algorithm may in some cases be extended to functions with negative coefficients. A weaker result is derived for application without restrictions on the signs of the coefficients or variables. This leads to a computational scheme which is useful in the solution of certain classes of nonlinear programming and inequality-constrained multistage optimization problems.

There has long been considerable interest in the mathematics of inequalities as a tool in analysis. Recently convenient computational procedures based on inequalities have been used for the solution of certain classes of optimization problems (*1* to *8*). This paper develops methods which extend the applicability of these schemes and provide the basis for new computational algorithms.

Engineering optimization problems are usually formulated in the context of the methods of differential and variational calculus or dynamic programming. This approach is by no means mandatory; for example, it has long been recognized that algebraic and integral inequalities can be established to provide bounds for a variety of mathematical functions. Until very recently, however, engineers have made little use of the vast body of inequality theory. One reason for this is that before the development of geometric programming by Duffin, Zener, and Peterson (*8*) there were no techniques based upon inequalities which did not have to be specifically adapted to fit each application. The geometric programming procedure uses the arithmetic-mean–geometric-mean inequality to provide an elegant algorithm applicable to a wide variety of maximization problems. Two important limitations of the method are its restriction to generalized polynomials with positive coefficients and variables and the potentially difficult computations required to solve the dual program for large systems.

By deriving a conjugate form of the arithmetic-mean–geometric-mean inequality, we will show that in certain cases the geometric programming algorithm can be applied to functions which contain negative terms. We then will derive a weaker result and use it to develop an algorithm which applies to functions containing arbitrary mixtures of positive and negative coefficients and variables. Finally, we will illustrate a useful new computational scheme. This scheme has been found to be especially efficient in the optimization of inequality-constrained multistage decision processes. An important property of these procedures is that they do not depend on the existence of derivatives.

A form of the arithmetic-mean–geometric-mean inequality with one positive and any number of negative weighting functions will first be derived. This result will later be shown to be a particular case of a conjugate inequality between arbitrary convex-functional means. It will then be shown that no more general conjugate inequalities are possible.

Carl D. Eben is with Monsanto Company, Springfield, Massachusetts.

A CONJUGATE FORM OF THE ARITHMETIC-MEAN–GEOMETRIC-MEAN INEQUALITY

Of fundamental importance in the theory of inequalities are the generalized Bernoulli inequalities. From these the usual arithmetic-mean–geometric mean (A-G) inequality and the basic inequalities of Holder and Minkowski can be derived (*9*, *10*). We are concerned with the following form of Bernoulli's inequality:

$$x^{\alpha} - \alpha x + \alpha - 1 \geqq 0 \tag{1}$$

which applies for $x > 0$ and $\alpha > 1$, or $\alpha < 0$. There is equality only when $x = 1$.

Following Beckenbach and Bellman (*10*), letting $x = x_1/x_2$ and $\alpha = \alpha_1$, we obtain

$$x_1^{\alpha_1} x_2^{-\alpha_2} \geqq \alpha_1 x_1 - \alpha_2 x_2 \tag{2}$$

with $\alpha_1 - \alpha_2 = 1$ and $\alpha_1 > 1$. There is equality only when $x_2 = x_2$. We shall call this the conjugate form of the weighted A-G inequality for two variables. Note that the direction of the inequality is the reverse of that found in the classical form with positive coefficients.

Proceeding inductively we let $x_1 = y_1^{\gamma_1/\alpha_1} y_2^{-\gamma_2/\alpha_1}$, $\alpha_1 = \gamma_1 - \gamma_2$, $x_2 = y_3$ and $\alpha_2 = \gamma_3$ with y_i and $\gamma_i > 0$. Evidently, $\gamma_1 - \gamma_2 - \gamma_3 = 1$. Making use of Inequality (2) gives

$$\begin{aligned} y_1^{\gamma_1} y_2^{-\gamma_2} y_3^{-\gamma_3} &= x_1^{\alpha_1} x_2^{-\alpha_2} \\ &\geqq \alpha_1 x_1 - \alpha_2 x_2 \\ &= (\gamma_1 - \gamma_2)\, y_1^{\gamma_1/\alpha_1} y_2^{-\gamma_2/\alpha_1} - \gamma_3 y_3 \\ &\geqq \gamma_1 y_1 - \gamma_2 y_2 - \gamma_3 y_3 \end{aligned} \tag{3}$$

Equality applies only if $y_1 = y_2 = y_3$. Continuing in this fashion, one can see that

$$x_1^{\gamma_1} \prod_{i=2}^{n} x_i^{-\gamma_i} \geqq \gamma_1 x_1 - \sum_{i=2}^{n} \gamma_i x_i \tag{4}$$

with $\gamma_1 - \sum_{i=2}^{n} \gamma_i = 1$ and γ_i, $x_i > 0$. There is equality only when all the x_i's are equal.

As will be shown, this inequality can be incorporated directly into the geometric programming algorithm for generalized polynomials with one positive and any number of negative coefficients. It will also be shown that this result cannot be extended to conjugate means with more than one positive coefficient; that is, Inequality (4) is the most general conjugate A-G inequality. Before proceeding, we will derive a conjugate inequality for general means.

INEQUALITIES FOR FUNCTIONAL AND POWER MEANS

Following Hardy, Littlewood and Polya (9), we define the functional means

$$M_{\psi}(\gamma, x) \equiv \psi^{-1}\left\{\sum_{i=1}^{n} \gamma_i \psi(x_i)\right\} \quad (5)$$

where the function $\psi(x)$ and its inverse function $\psi^{-1}(x)$ are continuous and strictly monotonic in the interval of interest, and $\sum_{i=1}^{n} \gamma_i = 1$. The geometric mean results by taking $\psi(x) = \ln x$; the arithmetic mean results when $\psi(x) = x$. [These give the respective inverse functions $\psi^{-1}(x) = e^x$ and $\psi^{-1}(x) = x$.]

If it is required that $x_i, \gamma_i > 0$ and if $\psi(x)$ and $\chi(x)$ are continuous and strictly monotonic with $\chi(x)$ increasing, then (9)

$$M_{\psi}(\gamma, x) \leqq M_{\chi}(\gamma, x) \quad (6)$$

provided that the function $\Phi = \chi(\psi^{-1})$ is convex. There is equality if all the x_i's are equal. Note that Inequality (6) includes the ordinary A-G inequality; that is, taking $\psi(x) = \ln x$ and $\chi(x) = x$ gives the convex function $\Phi(x) = e^x$.

For the case $n = 2$ we propose to show that if one of the coefficients is negative, Inequality (6) is reversed; that is, for $\gamma_1 - \gamma_2 = 1$,

$$\psi^{-1}\{\gamma_1 \psi(x_1) - \gamma_2 \psi(x_2)\} \geqq \chi^{-1}\{\gamma_1 \chi(x_1) - \gamma_2 \chi(x_2)\} \quad (7)$$

The other conditions in Inequality (6) are unchanged.

We begin the proof by using the property of a convex function,

$$\Phi(\gamma_1 x_1 + \gamma_2 x_2) \leqq \gamma_1 \Phi(x_1) + \gamma_2 \Phi(x_2) \quad (8)$$

This states that the chord of a convex function lies above or on the curve as shown in Figure 1. If we now take the weights $\gamma_1 > 1$ and $\gamma_2 = \gamma_1 - 1$ and if $x_1 \leqq x_2$ then

$$\gamma_1(x_2 - x_1) \geqq (x_2 - x_1) \quad (9)$$

or

$$\gamma_1 x_1 - \gamma_2 x_2 \leqq x_1$$

Similarly, if $x_1 \geqq x_2$,

$$\gamma_1 x_1 - \gamma_2 x_2 \geqq x_1 \quad (10)$$

These inequalities indicate the regions in which the chord is extended and given by $\gamma_1 \Phi(x_1) - \gamma_2 \Phi(x_2)$. In these regions it is clear that if Φ is convex, the chord lies below or on the curve; that is,

$$\Phi(\gamma_1 x_1 - \gamma_2 x_2) \geqq \gamma_1 \Phi(x_1) - \gamma_2 \Phi(x_2) \quad (11)$$

Inequality (7) follows directly from this by taking $\Phi = \chi\psi^{-1}$ and replacing x_i by $\psi(x_i)$. Inequality (2) is a specific form of Inequality (7) in the same way that the ordinary A-G inequality derives from Inequality (6).

To extend Inequality (7) to more than two variables, we shall make use of mathematical rather than geometrical arguments, although the latter will continue to be useful for visualizing the results. We choose to deal with convex functions which are twice differentiable (this is an unrestrictive assumption for our purposes) and propose to find the maximum values of $F = \sum_{i=1}^{n} \sigma_i \gamma_i \Phi(x_i)$ subject to the constraint $\Phi\left[\sum_{i=1}^{n} \sigma_i \gamma_i x_i\right] = \mu$, a constant. In these expressions $\sigma_i = \pm 1$, and $\gamma_i, x_i > 0$. Eliminating x_n, we obtain

$$F = \sum_{i=1}^{n-1} \sigma_i \gamma_i \Phi(x_i) + \sigma_n \gamma_n \Phi\left\{\frac{\Phi^{-1}(\mu) - \sum_{i=1}^{n-1} \sigma_i \gamma_i x_i}{\sigma_n \gamma_n}\right\} \quad (12)$$

The stationary points of F are given by

$$\frac{\partial F}{\partial x_i} = \sigma_i \gamma_i \Phi'(x_i) - \sigma_i \gamma_i \Phi'\left\{\frac{\Phi^{-1}(\mu) - \sum_{i=1}^{n-1} \sigma_i \gamma_i x_i}{\sigma_n \gamma_n}\right\} = 0 \quad (13)$$

for $i = 1, 2, \ldots, n-1$. This is satisfied only if all the x_i's are equal (invariance condition) or if Φ is linear in a specified region. The first condition arises from the fact that if Φ is strictly convex, Equation (13) is satisfied only if the arguments are equal. The second case is unimportant for our needs.

To determine the nature of the stationary points, we define

$$\Delta_j = \begin{vmatrix} F_{11} & F_{12} & \cdots & F_{1j} \\ F_{21} & F_{22} & \cdots & F_{2j} \\ \vdots & \vdots & & \vdots \\ F_{j1} & F_{j2} & \cdots & F_{jj} \end{vmatrix} (\Phi'')^{-j} \quad (14)$$

where $F_{ij} = \dfrac{\partial^2 F}{\partial x_j \, \partial x_i}$ for $j = 1, 2, \ldots, n-1$. Since $\Phi'' > 0$, Δ_j will have the same signs as the discriminants of the Hessian of F. From Equation (14),

$$\Delta_j = \frac{\prod_{i=1}^{j} \sigma_i \gamma_i}{\sigma_n \gamma_n} \left[\sigma_n \gamma_n + \sum_{i=1}^{j} \sigma_i \gamma_i\right] \quad (15)$$

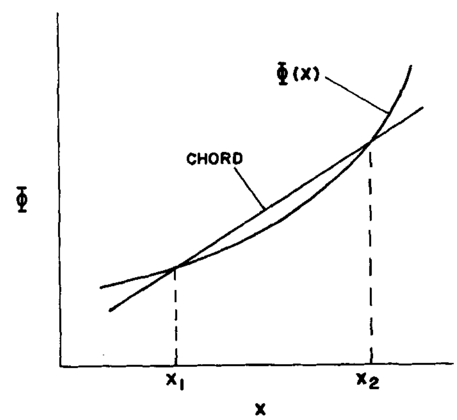

Fig. 1. A convex function and an extended chord.

A sufficient condition for F to be maximum is that Δ_{n-1} be negative definite; that is,

$$(-1)^j \Delta_j > 0 \tag{16a}$$

for $j = 1, 2, \ldots, n-1$. If Δ_{n-1} is positive definite, that is

$$\Delta_j > 0 \tag{16b}$$

then F is minimum. If neither Equation (16a) nor Equation (16b) is satisfied and none of the Δ_j are zero, there is a saddle point. The special case when Δ_{n-1} is semidefinite is difficult and is not worthy of systematic treatment.

Note that Equation (16b) is satisfied only if all the $\sigma_i = +1$; Equation (16a) requires that all except one of the σ_i's equal minus one. It follows that if Φ is convex and twice differentiable, $\gamma_1 - \sum_{i=2}^{n} \gamma_i = 1$ and $\gamma_i, x_i > 0$ then

$$\gamma_1 \Phi(x_1) - \sum_{i=2}^{n} \gamma_i \Phi(x_i) \leq \Phi \left[\gamma_1 x_1 - \sum_{i=2}^{n} \gamma_i x_i \right] \tag{17a}$$

and if $\sum_{i=1}^{n} \gamma_i = 1$,

$$\sum_{i=1}^{n} \gamma_i \Phi(x_i) \geq \Phi \left[\sum_{i=1}^{n} \gamma_i x_i \right] \tag{17b}$$

Further, no similar inequalities are possible for any other choice of signs.

These inequalities are the desired extensions of Inequalities (11) and (8), respectively. Inequality as shown in Equation (6) follows directly from Inequality (17b) and Equation (5). Similarly, Inequality (17a) gives Inequality (7) and its extension to any number of variables,

$$\psi^{-1}\left\{ \gamma_1 \psi(x_1) - \sum_{i=2}^{n} \gamma_i \psi(x_i) \right\} \geq \chi^{-1} \left\{ \gamma_1 \chi(x_1) - \sum_{i=2}^{n} \gamma_1 \chi(x_i) \right\} \tag{18}$$

This result may be applied, for example, to the power means which are obtained from Equation (5) by taking $\psi = x^r$,

$$M_r(\gamma, x) = \left\{ \sum_{i=1}^{n} \gamma_i x_i^r \right\}^{1/r} \tag{19}$$

($r = 1$ is the arithmetic mean, $r = 0$ is the geometric mean, $r = -1$ is the harmonic mean, $r = 2$ is the root mean square). If we take $\chi = x^s$ then $\Phi = \chi\psi^{-1} = x^{s/r}$. If $r < s$ then Φ is everywhere strictly convex ($\Phi'' > 0$) and if $\gamma_1 - \sum_{i=2}^{n} \gamma_i = 1$ it follows from Inequality (18) that

$$\left[\gamma_1 x_1^r - \sum_{i=2}^{n} \gamma_i x_i^r \right]^{1/r} \geq \left[\gamma_1 x_1^s - \sum_{i=2}^{n} \gamma_i x_i^s \right]^{1/s} \tag{20}$$

This is the conjugate power-means inequality. If $r = 0$ and $s = 1$ this reduces to the conjugate A-G inequality.

A CONJUGATE FORM OF HOLDER'S INEQUALITY

The A-G inequality is the most important of the familiar inequalities; consequently, its conjugate form is particularly useful. For example, it may be used to derive a conjugate form of Holder's inequality. This follows by taking (a), (b), $\ldots$, (l) each to be sets of n positive numbers and $\alpha, \beta, \ldots, \lambda$ to be positive constants with $\alpha - \beta - \ldots - \lambda = 1$; then, suppressing the index,

$$\frac{\Sigma a^\alpha b^{-\beta} \ldots l^{-\lambda}}{(\Sigma a)^\alpha (\Sigma b)^{-\beta} \ldots (\Sigma l)^{-\lambda}} = \sum \left(\frac{a}{\Sigma a}\right)^\alpha \left(\frac{b}{\Sigma b}\right)^{-\beta} \ldots \left(\frac{l}{\Sigma l}\right)^{-\lambda}$$
$$\geq \sum \left(\frac{\alpha a}{\Sigma a} - \frac{\beta b}{\Sigma b} - \ldots - \frac{\lambda l}{\Sigma l} \right) = \alpha - \beta - \ldots - \lambda = 1 \tag{21}$$

Therefore

$$\Sigma a^\alpha b^{-\beta} \ldots l^{-\lambda} \geq (\Sigma a)^\alpha (\Sigma b)^{-\beta} \ldots (\Sigma l)^{-\lambda} \tag{22}$$

There is equality only if the sets (a), (b), $\ldots$, (l) are proportional or if one set is null.

Inequality (22) is easily extended to

$$M_r(\gamma, a\, b \ldots l) \geq M_{r/\alpha}(\gamma, a)\, M_{-r/\beta}(\gamma, b) \ldots M_{-r/\lambda}(\gamma, l) \tag{23}$$

with $r > 0$. There is equality only when $(a^{1/\alpha})$, $(b^{-1/\beta})$, $\ldots$, $(l^{-1/\lambda})$ are proportional or one set is null. If $r < 0$ the inequality is reversed. This inequality is proved by Hardy (9) for two sets; it can be used to prove Minkowski's inequality.

A WEAKER RESULT

Returning to the main discussion, we have seen that for values of σ_i aside from those which satisfy Equations (16a) or (16b) no inequality is possible between convex-functional means. However, Equation (13) applies for any choice of the σ's. It follows that all the *stationary* points of F are given by the condition that all the x_i's are equal. Denoting the stationary values by a bar results in

$$\overline{M}_\psi(\sigma\gamma, x) = \overline{M}_\chi(\sigma\gamma, x) \tag{24}$$

The utility of this arises from the fact that the stationary values $\overline{M}_\psi$ can be obtained from $\overline{M}_\chi$ and, further, that these points are given by the invariance condition $\overline{x}_i$ = constant.

This result is quite analogous to the weak form of the discrete maximum principle (*11, 12*). Just as a stationary value of the Hamiltonian for a discrete system may lead to an extreme of the objective function, so a stationary value of $\overline{M}_\chi$ may lead to an extreme of $\overline{M}_\psi$; the converse is, of course, also true.

We are particularly interested in the special A-G form of Equation (24) obtained by replacing x_i by x_i/γ_i. Taking $\psi\left(\frac{x}{\gamma}\right) = \frac{x}{\gamma}$ and $\chi\left(\frac{x}{\gamma}\right) = \ln\left(\frac{x}{\gamma}\right)$ in Equation (24) and denoting the variables at the stationary points by $\overline{x}_i$ results in

$$\sum_{i=1}^{n} \sigma_i \overline{x}_i = \prod_{i=1}^{n} \left(\frac{\overline{x}_i}{\gamma_i}\right)^{\sigma_i\gamma_i} \tag{25}$$

As before, $\sigma_i = \pm 1$ and $\sum_{i=1}^{n} \sigma_i\gamma_i = 1$. The invariance condition is that $\overline{x}_i/\gamma_i$ is constant.

Without the normality condition, Equation (25) becomes

$$\sum_{i=1}^{n} \sigma_i \overline{x}_i = \prod_{i=1}^{n} \left(\frac{\overline{x}_i}{\gamma_i}\right)^{\frac{\sigma_i\gamma_i}{\sum_i \sigma_i\gamma_i}} \left[\sum_{i=1}^{n} \sigma_i\gamma_i \right] \tag{26}$$

Other specific forms of Equation (24) analogous to Inequalities (20), (22), etc., can be developed similarly.

The usefulness of these results will be illustrated by several simple nonlinear programming problems.

APPLICATIONS TO EXTREMUM PROBLEMS

When Inequality (4) is written analogous to Equation (25) in the form

$$x_1 - \sum_{i=2}^{n} x_i \leq \left(\frac{x_1}{\gamma_1}\right)^{\gamma_1} \prod_{i=2}^{n} \left(\frac{x_i}{\gamma_i}\right)^{-\gamma_i} \qquad (27)$$

with $\gamma_1 - \sum_{i=2}^{n} \gamma_i = 1$, or, without the normality restriction,

$$x_1 - \sum_{i=2}^{n} x_i \leq \left\{ \left(\frac{x_1}{\gamma_1}\right)^{\gamma_1} \prod_{i=2}^{n} \left(\frac{x_i}{\gamma_i}\right)^{-\gamma_i} \right\}^{\frac{1}{\gamma_1 - \sum_{i \neq 1} \gamma_i}} \left[\gamma_1 - \sum_{i=2}^{n} \gamma_i \right] \qquad (28)$$

it is directly applicable to the geometric programming algorithm for generalized polynomials of the form

$$U = c_1\, y_1^{a_{11}}\, y_2^{a_{12}} \ldots y_m^{a_{1m}} - \sum_{i=2}^{n} c_i\, y_1^{a_{i1}}\, y_2^{a_{i2}} \ldots y_m^{a_{im}} \qquad (29)$$

This satisfies Equation (16*a*), and consequently the duality property of geometric programming is retained. Also, note that the objective function is maximized rather than minimized as in conventional geometric programming. For example, suppose we wish to maximize the quantity

$$Z(u, v, w) = \frac{4\sqrt{u}}{v} - \frac{3\,u^2\sqrt{v}}{w^2} \qquad (30)$$

subject to the constraint

$$\frac{4}{w} - \frac{u^{1/4}}{v^2\sqrt{w}} \geq 1 \qquad (31)$$

Applying Inequality (27) to Equation (30) and Inequality (28) to Equation (31) and multiplying the resulting inequalities gives

$$Z \leq \left(\frac{4\,u^{1/2}}{\gamma_1 v}\right)^{\gamma_1} \left(\frac{3\,v^{1/2}\,u^2}{\gamma_2 w^2}\right)^{-\gamma_2} \cdot \left(\frac{4}{\gamma_3 w}\right)^{\gamma_3} \left(\frac{u^{1/4}}{\gamma_4 v^2 w^{1/2}}\right)^{-\gamma_4} \cdot (\gamma_3 - \gamma_4)^{(\gamma_3 - \gamma_4)} \qquad (32)$$

Following Duffin we eliminate u, v, and w from this inequality by choosing

$$\begin{aligned} \tfrac{1}{2}\gamma_1 - 2\gamma_2 - \tfrac{1}{4}\gamma_4 &= 0 \\ -\gamma_1 - \tfrac{1}{2}\gamma_2 + 2\gamma_4 &= 0 \\ 2\gamma_2 - \gamma_3 + \tfrac{1}{2}\gamma_4 &= 0 \end{aligned} \qquad (33)$$

These are the orthogonality condtions. The normality condition is

$$\gamma_1 - \gamma_2 = 1 \qquad (34)$$

Solving Equations (33) and (34) gives

$$\gamma_1 = \frac{11}{9},\ \gamma_2 = \frac{2}{9},\ \gamma_3 = \frac{7}{9},\ \gamma_4 = \frac{2}{3} \qquad (35)$$

Then

$$Z \leq \left(\frac{4}{\gamma_1}\right)^{\gamma_1} \left(\frac{3}{\gamma_2}\right)^{-\gamma_2} \left(\frac{4}{\gamma_3}\right)^{\gamma_3} \left(\frac{1}{\gamma_4}\right)^{-\gamma_4} (\gamma_3 - \gamma_4)^{(\gamma_3 - \gamma_4)} = 5.104 \qquad (36)$$

Duffin shows that the values $\bar{u}$, $\bar{v}$ and $\bar{w}$ may be recovered by using the following relationships, which result from the invariance conditions:

$$\frac{4\sqrt{\bar{u}}}{\bar{v}} = \gamma_1 \bar{Z}, \quad \frac{3\,\bar{u}^2\sqrt{\bar{v}}}{\bar{w}^2} = \gamma_2 \bar{Z}$$

$$\frac{4}{\bar{w}} = \frac{\gamma_3}{\gamma_3 - \gamma_4}, \quad \frac{\bar{u}^{1/4}}{\bar{v}^2\sqrt{\bar{w}}} = \frac{\gamma_4}{\gamma_3 - \gamma_4} \qquad (37)$$

These give $\bar{u} = 0.4356$, $\bar{v} = 0.4232$ and $\bar{w} = 0.5174$.

In this example both the objective function and the constraint were in the form of Equation (29). Had one or both contained all positive (or all negative) coefficients then the ordinary A-G inequality would have been used. Also, the fact that we wished to maximize Z rather than minimize it is no real restriction since Min(Z) can be obtained from Max(−Z).

The geometric programming computations in this problem were greatly facilitated by the fact that the values for all the weighting factors could be calculated from Equations (33) and (34). More commonly, some of the weighting factors will be left unspecified and must be chosen to minimize the geometric mean in Inequality (36). Also, if the coefficients in the objective function and the constraints did not have the signs required by Equations (16*a*) or (16*b*) we have seen that the strong result would no longer apply and the weaker hypothesis [Equation (24)] would be used. That is, any weighting factors which were left unspecified would be chosen to make the geometric mean stationary, not necessarily minimum.

These points are illustrated by the problem of locating the extremes of the function

$$Z = -\frac{1}{x^3} + \frac{1}{x^2} - \frac{x}{y^2} + \frac{1}{y} \qquad (38)$$

Applying Equation (25) to this gives

$$\bar{Z} = \left(\frac{1}{\gamma_1 \bar{x}^3}\right)^{-\gamma_1} \left(\frac{1}{\gamma_2 \bar{x}^2}\right)^{\gamma_2} \left(\frac{\bar{x}}{\gamma_3 \bar{y}^2}\right)^{-\gamma_3} \left(\frac{1}{\gamma_4 \bar{y}}\right)^{\gamma_4} \qquad (39)$$

The orthogonality and normality conditions are

$$\begin{aligned} 3\gamma_1 - 2\gamma_2 - \gamma_3 &= 0 \\ 2\gamma_3 - \gamma_4 &= 0 \end{aligned} \qquad (40)$$

and

$$-\gamma_1 + \gamma_2 - \gamma_3 + \gamma_4 = 1 \qquad (41)$$

Therefore

$$\bar{Z} = \left[\frac{1}{2 - \frac{\gamma_4}{2}}\right]^{-2 + \frac{\gamma_4}{2}} \left(\frac{1}{3 - \gamma_4}\right)^{3 - \gamma_4} \left(\frac{2}{\gamma_4}\right)^{-\frac{\gamma_4}{2}} \left(\frac{1}{\gamma_4}\right)^{\gamma_4} \qquad (42)$$

The stationary points can be obtained from

$$\frac{d \ln \bar{Z}}{d\gamma_4} = 0 \qquad (43)$$

which has the solutions $\gamma_4 = 1.177$ and $\gamma_4 = 3.823$.

Substituting the first root into Equation (42) gives $\bar{Z} = 0.3289$. Equations (40) and (41) give $\gamma_1 = 1.412$, $\gamma_2 = 1.823$ and $\gamma_3 = 0.589$. As in the previous example the invariance conditions can be used to obtain $\bar{x} = 1.292$ and $\bar{y} = 2.583$. Perturbing Equation (38) about $(\bar{x}, \bar{y})$ shows

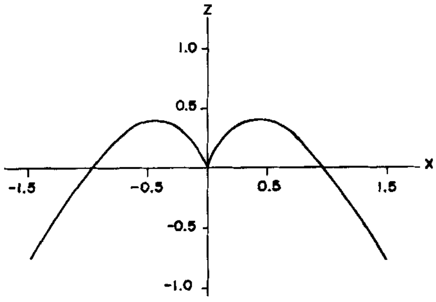

Fig. 2. Graph of Equation (46).

$\overline{Z}$ to be maximum. It can be calculated that $d^2 \ln \overline{Z}/d\, \gamma_4^2 < 0$, which indicates that the geometric mean is also maximized. This demonstrates the weakened hypothesis of Equation (24) and shows that the duality property of geometric programming is not retained.

For the root $\gamma_4 = 3.823$ Equation (42) cannot be evaluated. This indicates that $\overline{Z}$ is not positive. Further, Equations (40) and (41) give $\gamma_1 = 0.090$, $\gamma_2 = -0.821$ and $\gamma_3 = 1.912$, which violate the positivity condition. This indicates that the point $(\overline{x}, \overline{y})$ is not contained in the positive quadrant.

These problems can be formally circumvented by redefining the variables in Equation (38) to suitably reorient the axes. Alternatively, since it can be seen that invariance does not depend on positivity, the invariance conditions can be used directly to obtain $\overline{x}$ and $\overline{y}$. In this example the invariance conditions are

$$\frac{1}{\gamma_1 \overline{x}^3} = \frac{1}{\gamma_2 \overline{x}^2} = \frac{\overline{x}}{\gamma_3 \overline{y}^2} = \frac{1}{\gamma_4 \overline{y}} \tag{44}$$

Substituting the γ_i's into Equation (44) and solving gives $\overline{x} = -9.122$ and $\overline{y} = -18.24$. From Equation (38) it is found that $\overline{Z} = -0.0141$, which is found to be a minimum. This completes the solution.

It is frequently advantageous to avoid entirely manipulation of the geometric mean by making use only of the invariance, orthogonality, and normality conditions. In the second example the values of $\overline{x}$ and $\overline{y}$ could have been obtained by simultaneous solution of Equations (40), (41), and (44). One approach to the solution would be first to eliminate the γ_i's, leaving the two equations

$$\frac{3}{\overline{x}^4} - \frac{2}{\overline{x}^3} - \frac{1}{\overline{y}^2} = 0$$

and

$$\frac{2\overline{x}}{\overline{y}^3} - \frac{1}{\overline{y}^2} = 0 \tag{45}$$

These are exactly the equations that would result from equating the partial derivatives of Equation (38) to zero. There are, of course, numerous other approaches to the solution of these simultaneous equations. It is significant that in multistage decision problems this procedure leads to a split-boundary-value problem.

This method is especially attractive when a large number of weighting factors (γ_i's) are left unspecified by the orthogonality and normality conditions. It is easily seen that the number of unspecified γ_i's is equal to the number of terms in the objective function and constraints minus one more than the number of variables. Duffin calls this the degree of difficulty.

Finally, the fact that the inequality solution does not depend on the existence of derivatives is illustrated by the local maxima and minima of the function

$$Z = x^{2/3} - x^2 \tag{46}$$

This is shown in Figure 2. Applying Equation (28) gives

$$\overline{Z} = \left(\frac{\overline{x}^{2/3}}{\gamma_1}\right)^{\gamma_1} \left(\frac{\overline{x}^2}{\gamma_2}\right)^{-\gamma_2} \tag{47}$$

The orthogonality and normality conditions are

$$\frac{2}{3}\gamma_1 - 2\gamma_2 = 0 \tag{48}$$

and

$$\gamma_1 - \gamma_2 = 1 \tag{49}$$

which give $\gamma_1 = 3/2$ and $\gamma_2 = 1/2$

The invariance condition is

$$\frac{\overline{x}^{2/3}}{3} = \overline{x}^2 \tag{50}$$

which gives $\overline{x} = 0, \pm 0.44$.

Geometric programming using the conjugate A-G inequality would locate only the positive root. The application of ordinary calculus would be embarrassed at $x = 0$. This property is of obvious importance in treating poorly behaved functions.

CONCLUSIONS

Starting with Bernoulli's inequality we have derived a conjugate form of the arithmetic-mean–geometric-mean inequality for means which contain one positive and any number of negative weights. Consideration of general convex-functional means showed that similar conjugate inequalities are possible for a broad class of functions which includes the arithmetic and geometric means. The conjugate A-G inequality was shown to be applicable to the geometric programming algorithm.

A weaker result was derived for application to generalized means without regard to the signs of the coefficients or variables. Applying the A-G form of this result to the problem of finding extreme solutions of generalized polynomials revealed that it is sometimes useful to include the equality conditions in the computations.

The natural inclusion of equality and inequality constraints and the absence of dependence on the derivative are especially important characteristics of these schemes. It is to be expected that further results can be obtained from the general-means inequality which may be useful in treating problems that are not in the generalized polynomial form. To date it has been found necessary to tailor these inequalities to the problem at hand; only the A-G inequality has proven to have broad applicability. The particular usefulness of the A-G inequality derives from the fact that many physical processes can be described by general polynomials.

Interesting and useful computational schemes can be expected to derive from the equality conditions. Algorithms which incorporate these conditions have been found to provide straightforward, rapid solutions to large multistage decision processes (*13*).

ACKNOWLEDGMENT

During a portion of this work Carl D. Eben was a National Science Foundation Cooperative Fellow.

NOTATION

(a), (b), (l) = sets of positive numbers
a_{ij} = constant exponent in generalized polynomial

c_i = positive constant coefficient in generalized polynomial
F = weighted sum of convex functions
F_{ij} = $\partial^2 F/\partial x_j \partial x_i$
M_r = power mean
M_χ, M_ψ = functional means
r, s = constants
U = generalized polynomial
u, v, w, x, y = independent variables
Z = dependent variable

Greek Letters

α, β, λ = positive constants
α_i, γ_i = weighting factors
Δ_j = discriminants of modified Hessian matrix
μ = a constant
σ = ± 1
Φ = convex function
χ, ψ = continuous, strictly monotonic functions
χ^{-1}, ψ^{-1} = inverse functions

Superscript

$-$ = value at stationary point

LITERATURE CITED

1. Zener, C. M., *Proc. Nat. Acad. Sci.*, **47**, 537 (1961).
2. ———, Scientific Paper 65-1HO-Opdes-PL, Westinghouse Research Laboratories, Pittsburgh, Pa. (January 20, 1965).
3. Duffin, R. J., *Operations Res.*, **10**, 668 (1962).
4. ———, *J. Soc. Ind. Appl. Math.*, **10**, 119 (1962).
5. ———, and E. L. Peterson, Scientific Paper 64-158-129-P5, Westinghouse Research Laboratories, Pittsburgh, Pa. (December 11, 1964).
6. Ferron, J. R., *Chem. Eng. Progr. Symp. Ser. No. 50*, **60**, 60 (1964).
7. ———, lectures prepared for the NSF-AIChE sponsored Conference on Optimization Theory, Stanford Univ., Stanford, Calif. (August, 1965).
8. Duffin, R. J., E. L. Peterson, and C. M. Zener, "Geometric Programming," Wiley, New York (1967).
9. Hardy, G. H., J. E. Littlewood, and G. Polya, "Inequalities," 2 ed., pp. 12-101, Cambridge Univ. Press, London (1951).
10. Beckenbach, Edwin F., and Richard Bellman, " Inequalities," second revised printing, pp. 1-30, Springer-Verlag, New York (1965).
11. Denn, M. M., and Rutherford Aris, *Ind. Eng. Chem. Fundamentals*, **4**, 7 (1965).
12. Jackson, R., and F. Horn, *ibid.*, 110 (1965).
13. Eben, C. D., Ph.D. thesis, Univ. of Delaware, Newark, in preparation.

Manuscript received June 6, 1967; revision received September 25, 1967; paper accepted September 29, 1967.

Nonlinear Feedforward Control of Chemical Reactors

WILLIAM L. LUYBEN

E. I. du Pont de Nemours and Company, Wilmington, Delaware

This paper illustrates the synthesis of nonlinear feedforward controllers for chemical reactors. In most of the theoretical development and application of feedforward control only linear systems have been considered. There are, however, no inherent linear limitations in feedforward control. Since chemical reactors are usually nonlinear, the effectiveness of control should be improved by including nonlinearities in the design of feedforward controllers. This is particularly true for batch reactors because of the large changes in variables during a batch cycle. Continuous stirred-tank reactors are studied with single and consecutive reactions of first and higher order. Effectiveness of linear and nonlinear feedforward controllers is compared for disturbances of various magnitude and direction. Feedforward control of batch and tubular reactors is also discussed.

Feedforward control of chemical processes has received an increasing amount of attention in recent years. This interest is a result of the recognition of the advantages of feedforward control in many chemical engineering applications; improved knowledge and appreciation of the dynamics of chemical processes; and the availability of computers to permit realistic analysis, synthesis, and evaluation of control systems, both on line and off line.

Distillation columns and chemical reactors have received most of the attention. The general theory of feedforward controller synthesis was developed by Bollinger and Lamb

William L. Luyben is at Lehigh University, Bethlehem, Pennsylvania.