A FUNCTIONAL CHARACTERIZATION OF A CLASS OF
COVARIANCE INEQUALITIES

by
George Kimeldorf and Allan R. Sampson

FSU Statistics Report M108
A FUNCTIONAL CHARACTERIZATION OF A CLASS OF
COVARIANCE INEQUALITIES

by
George Kimeldorf and Allan R. Sampson

1. Introduction and Summary. In this paper we consider the class
of inequalities of the form

(1)
\[ \text{cov} \ (X,Y) \leq \text{var} \ (X,Y). \]

The functions \( g \) having continuous first partial derivatives and sat-
istifying (1) for all suitable random variables \( X \) and \( Y \) are shown to be
characterized by a functional inequality and an equivalent partial dif-
ferential inequality.

Since \( \text{var} \ g(X,Y) = \text{cov} \ [g(X,Y), g(X,Y)] \), we can view \( g \) as a function that
combines the individual random variables \( X \) and \( Y \) to create a new ran-
don variable \( Z = g(X,Y) \) which has the property that \( \text{cov} \ (Z,Z) \geq \text{cov} \ (X,Y) \).

Thus \( g \) can be considered as a covariance-increasing function. Koop [2]
considered the inequality \( \text{cov} \ (U, ZU^{-1}) \leq \text{var} \ (Z^2) \) for positive \( U \) and \( Z \).

It is easily seen that his inequality is a special case of (1) restricted
to the positive quadrant with \( X = U, Y = ZU^{-1} \), and \( g(x,y) = (xy)^2 \).

2. The main results. Let \( A \) be an open connected subset of the
plane \( \mathbb{R}^2 \) and define \( \mathcal{A} \) to be the class of all pairs \( (X,Y) \) of real ran-
don variables having finite variances and jointly taking values in \( A \).

Assume \( g \) is a function with domain \( \mathcal{A} \) and with continuous first
partial derivatives.

**Theorem 1.** A necessary and sufficient condition that \( \text{cov} \ (X,Y) \leq \text{var} \ g(X,Y) \) for all \((X,Y) \in \mathcal{A}\) is that

(2)
\[
\begin{pmatrix}
\frac{\partial g}{\partial x} \\
\frac{\partial g}{\partial y}
\end{pmatrix}
\begin{pmatrix}
\frac{\partial g}{\partial x} \\
\frac{\partial g}{\partial y}
\end{pmatrix} \geq h^2
\]

for all points \((x,y) \in A\).
The proof of Theorem 1 follows immediately from the next two lemmas.

Lemma 1. Inequality (1) holds for all \((x, y) \in A\) if, and only if,

\[
(x - x_1)(y - y_1) \leq \left( g(x_2, y_2) - g(x_1, y_1) \right)^2.
\]

Proof. Suppose (1) holds for all \((x, y) \in A\). If \((x, y)\) only takes values \((x_1, y_1)\) and \((x_2, y_2)\), then (1) reduces immediately to (3).

Conversely, if (1) is violated, then it is violated by some pair \((x, y)\) of random variables jointly assuming only a finite number of values with equal probability. Let \((x, y)\) take values \((x_1, y_1), \ldots, (x_n, y_n)\), each with probability \(n^{-1}\). If (3) holds on \(A\), then

\[
\text{cov}(x, y) = n^{-1} \sum x_i y_i - n^{-2} \left( \sum x_i \right) \left( \sum y_i \right)
\]

\[
= n^{-2} \sum_{i \neq j} (x_i - x_j)(y_i - y_j)
\]

\[
\leq n^{-2} \sum_{i \neq j} \left( g(x_i, y_i) - g(x_j, y_j) \right)^2.
\]

This latter expression equals

\[
n^{-1} \left[ g(x_1, y_1) \right]^2 - \left[ n^{-1} \sum g(x_i, y_i) \right]^2 = \text{var}(g(x, y)).
\]

Lemma 2. Inequality (3) holds on \(A\) if, and only if, (2) holds on \(A\).

Proof. We first show that (2) is equivalent to the condition

\[
0 \leq \left( \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y} \right)^2 - \alpha
\]

for all real \(\alpha\),

where \(\frac{\partial g}{\partial x}\) and \(\frac{\partial g}{\partial y}\) are the partial derivatives of \(g\) with respect to \(x\) and \(y\), respectively. To see the equivalence, note that for any fixed \((x, y)\)
the right hand side of (4), a quadratic polynomial in \( z \), attains a minimum value of \( (e_x e_y - \lambda)/e_y^2 \), which is non-negative if, and only if, (2) holds.

Now suppose (3) holds on \( A \). Setting \( x_2 = x_1 + \delta \) and \( y_2 = y_1 + \delta \) and taking limits as \( \delta \to 0 \), we derive (4). To show the converse implication, we fix \( x_1, x_2, y_1, \) and \( y_2 \) and set \( a = (y_2 - y_1)/(x_2 - x_1) \). Because \( A \) is connected, there exists by the mean value theorem a point \((\bar{x}, \bar{y}) \in A\) for which \( g(x_2, y_2) - g(x_1, y_1) = (x_2 - x_1)(ag_y + e_x) \) where the partials are evaluated at \((\bar{x}, \bar{y})\). Hence, (4) implies
\[
|g(x_2, y_2) - g(x_1, y_1)|^2 \leq (x_2 - x_1)^2 a = (y_2 - y_1)(x_2 - x_1).
\]

3. "tightness" considerations. It is immediately obvious that if \( g \) satisfies (1), then for \( |a| \geq 1 \) every linear transform \( ag + b \) also satisfies (1). This situation arises because we have not imposed the condition that (1) be a "tight" inequality, i.e., an equality for some particular set of random variables \( X \) and \( Y \). There are many such conditions that can be imposed. In numerous cases of interest, \( A \) contains an open interval of the line \( (y = x) \). Hence, meaningful and reasonable conditions are that \( \text{var} g(X,X) = \text{cov} (X,X) \) and \( g(0,0) = 0 \). It can be readily shown that these "tightness" conditions hold if, and only if,
\[
(5) \quad g(x,x) = x \quad \text{for all} \ (x,x) \in A.
\]

Another justification of (5), albeit highly heuristic, is to consider \( g(x,y) \) as a quasilinear weighted mean (e.g., see Aczel [1, pp. 240ff]) of \( x \) and \( y \). Thus, we require reflexivity, i.e., \( g(x,x) = x \). Clearly, if \( A \) contains an open interval of the line \( (y = x) \) and if \( g \) satisfies
(5), then no nontrivial linear function of $g$ also satisfies (5).

4. Examples. One interesting example of inequality (1) is the case where $A = \mathbb{R}^2$. Under condition (5), we show that the only function $g$ for which (1) is true for all $(X, Y) \in A$ is $g(x, y) = (x + y)/2$.

**Corollary 1.** If $A = \mathbb{R}^2$ and $g$ satisfies (5), then (1) holds for all $(X, Y) \in A$ if, and only if, $g(x, y) = (x + y)/2$.

**Proof.** By Lemma 2, it is sufficient to show (3) implies $g(x, y) = (x + y)/2$, as the converse implication is immediate.

Let $x_1 = y_1 = z$, so that $g(x_1, y_1) = g(z, z) = z$. Inequality (3) now becomes after simplification

$$x_2y_2 - g^2(x_2, y_2) \leq z \left[ x_2 + y_2 - 2g(x_2, y_2) \right].$$

For any fixed $x_2$ and $y_2$, we can always choose $z$ to violate inequality (6) unless the expression in brackets is zero, and hence $g(x, y) = (x + y)/2$.

Another example is to take $A$ to be the positive quadrant of $\mathbb{R}^2$, and thus $X$ and $Y$ to be positive random variables. Assuming that (5) holds, we demonstrate that there is no unique function $g$ satisfying (1), although some necessary bounds on $g$ are obtained.

**Corollary 2.** If $A$ is the positive quadrant of $\mathbb{R}^2$ and $g$ satisfies (5), then a necessary condition that (1) hold for all $(X, Y) \in A$ is

$$0 \leq g(x, y) \leq (x + y)/2$$

for all $x > 0$ and $y > 0$.

**Proof.** If $A$ is the positive quadrant and $g$ satisfies (1) and (5), then the argument in the proof of Corollary 1 implies that (6)
holds for all \( z > 0 \). If \( g(x, y) < (xy)^{\frac{1}{3}} \), then the left hand side of (6) is positive, so that the inequality can be violated for sufficiently small \( z \). Similarly, if \( g(x, y) > (x + y)/2 \), the right hand side of (6) is negative, so that (6) is violated for sufficiently large \( z \).

It is clear that (7) is not sufficient, for there exist \( g \) satisfying (7) but violating (2) at some point. Also, using the differential inequality (2), we can easily verify that \( g(x, y) = a(xy)^{\frac{1}{3}} + (1 - a)(x + y)/2 \) satisfies (1) for \( 0 \leq a \leq 1 \).

5. Remarks. Observe that \( (xy)^{\frac{1}{3}} \) and \( (x + y)/2 \) are both means, being the geometric and arithmetic means, respectively. However, another mean, the harmonic mean, does not satisfy (2) for sufficiently small \( x \) or \( y \). The class of quasilinear weighted means may provide further functions that satisfy (1) for specific regions \( A \).

Although we have not pursued the question here, the covariance inequalities we have considered should have interesting multivariate analogs.

We thank Dr. Christopher Hunter for a helpful conversation.
REFERENCES
