J. of Inequal. & Appl., 1998, Vol. 2, pp. 275–283
Reprints available directly from the publisher
Photocopying permitted by license only

An Integral Analogue of the Ostrowski Inequality

C.E.M. PEARCE[a,*], J. PEČARIĆ[b] and S. VAROŠANEC[c]

[a] *Department of Applied Mathematics, The University of Adelaide, Adelaide SA 5005, Australia;* [b] *Faculty of Textile Technology, University of Zagreb, Pierottijeva 6, 41000 Zagreb, Croatia;* [c] *Department of Mathematics, University of Zagreb, Bijenička cesta 30, 41000 Zagreb, Croatia*

(Received 2 April 1997)

We give an integral analogue of the Ostrowski inequality and several extensions, allowing in particular for multiple linear constraints.

Keywords: Ostrowski inequality; Integral inequality; Multiple linear constraints

AMS 1991 Subject Classification: Primary 26D15; Secondary 60E15

1 INTRODUCTION

The result known as Ostrowski's inequality [6] is as follows.

THEOREM A *Let a, b and z be real n-tuples with $a \neq 0$ and such that*

$$\sum_{i=1}^{n} a_i z_i = 0 \quad and \quad \sum_{i=1}^{n} b_i z_i = 1. \tag{1}$$

Then

$$\sum_{i=1}^{n} z_i^2 \geq \frac{\sum_{i=1}^{n} a_i^2}{(\sum_{i=1}^{n} a_i^2)(\sum_{i=1}^{n} b_i^2) - (\sum_{i=1}^{n} a_i b_i)^2}. \tag{2}$$

* Corresponding author.

Equality occurs if and only if

$$z_j = \frac{b_j \sum a_i^2 - a_j \sum a_i b_i}{(\sum a_i^2)(\sum b_i^2) - (\sum a_i b_i)^2} \quad (1 \le j \le n).$$

We remark that (1) entails that the sequences (a_i), (b_i) are not proportional, so that by the condition for equality in Cauchy's theorem the common denominator of the expressions on the right-hand sides of the last two relations is nonzero.

Ostrowski's inequality has been extended by Alić and Pečarić [1], who established Theorem B below.

THEOREM B *Suppose the conditions of Theorem A hold and $p \ge 1$ is a real number. Then*

$$\left(\sum z_i^2\right)^p \ge \frac{(\sum a_i^2)^p}{(\sum a_i^2)^p (\sum b_i^2)^p - (\sum a_i b_i)^{2p}}.$$

This extended an earlier result of Madevski [3]. Alić and Pečarić used Theorem B to derive a number of applications.

The aim of this paper is to carry these ideas somewhat further. First we present an integral analogue to Ostrowski's inequality. In fact both Theorems A and B can be so extended. This is the substance of Section 2.

In Section 3 we note briefly how this may be used to derive some results for moments of probability distributions. We then turn to extensions of the discrete formulation. In Section 4 we note that the results of [1] generalize to the case of nonuniform weighting and in Section 5 we obtain a higher-dimensional discrete version of Theorem A, allowing for variables which are subject to a general number of linear constraints.

We conclude Section 5 with a corresponding extension to the integral analogue allowing a general number of linear constraints.

2 AN INTEGRAL OSTROWSKI INEQUALITY

It will be convenient to first derive an integral version of Theorem A and then extend this to provide an integral analogue of Theorem B.

THEOREM 1 *Let σ be a nonnegative measure on the real line* $\mathbf{R}$ *and* $f, g, h : \mathbf{R} \to \mathbf{R}$ *be functions with g not identically zero and such that* $f^2, g^2, h^2 \in \mathcal{L}_1(\mathbf{R}, \sigma)$, *with*

$$\int g(x)f(x)\,\mathrm{d}\sigma = 0 \quad \textit{and} \quad \int h(x)f(x)\,\mathrm{d}\sigma = 1. \tag{3}$$

Then

$$\int f^2(x)\,\mathrm{d}\sigma \geq \frac{\int g^2(x)\,\mathrm{d}\sigma}{(\int g^2(x)\,\mathrm{d}\sigma)(\int h^2(x)\,\mathrm{d}\sigma) - (\int g(x)h(x)\,\mathrm{d}\sigma)^2}, \tag{4}$$

with equality if and only if

$$f(x) = \frac{h(x)\int g^2(x)\,\mathrm{d}\sigma - g(x)\int h^2(x)\,\mathrm{d}\sigma}{(\int g^2(x)\,\mathrm{d}\sigma)(\int h^2(x)\,\mathrm{d}\sigma) - (\int g(x)h(x)\,\mathrm{d}\sigma)^2}.$$

Proof Set $A := \int g^2(x)\,\mathrm{d}\sigma$, $B := \int h^2(x)\,\mathrm{d}\sigma$, $C := \int g(x)h(x)\,\mathrm{d}\sigma$ and define function $w : \mathbf{R} \to \mathbf{R}$ by

$$w(x) = \frac{Ah(x) - Cg(x)}{AB - C^2}.$$

As with our comments following the enunciation of Theorem A, the denominator in this last expression is nonvanishing. It is easy to check that

$$\int g(x)w(x)\,\mathrm{d}\sigma = 0, \quad \int h(x)w(x)\,\mathrm{d}\sigma = 1,$$
$$\int f(x)w(x)\,\mathrm{d}\sigma = \frac{A}{AB - C^2}, \quad \int w^2(x)\,\mathrm{d}\sigma = \frac{A}{AB - C^2}.$$

Hence we have

$$\begin{aligned}
0 &\leq \int (f(x) - w(x))^2\,\mathrm{d}\sigma \\
&= \int f^2(x)\,\mathrm{d}\sigma - 2\int f(x)w(x)\,\mathrm{d}\sigma + \int w^2(x)\,\mathrm{d}\sigma \\
&= \int f^2(x)\,\mathrm{d}\sigma - \frac{2A}{AB - C^2} + \frac{A}{AB - C^2} \\
&= \int f^2(x)\,\mathrm{d}x - \frac{A}{AB - C^2},
\end{aligned}$$

giving the desired result.

THEOREM 2 *Assume the conditions of Theorem* 1 *hold and let* $p \geq 1$ *be a real number. Then*

$$\left(\int f^2(x)\,\mathrm{d}\sigma\right)^p \geq \frac{(\int g^2(x)\,\mathrm{d}\sigma)^p}{(\int g^2(x)\,\mathrm{d}\sigma)^p(\int h^2(x)\,\mathrm{d}\sigma)^p - (\int g(x)h(x)\,\mathrm{d}\sigma)^{2p}}.$$

Proof For $u \geq v \geq 0$, the inequality between power sums of orders $p \geq 1$ and 1 provides

$$((u-v)^p + v^p)^{1/p} \leq (u-v) + v = u,$$

that is, $(u-v)^p \leq u^p - v^p$. Hence by (4)

$$\begin{aligned}\left(\int g^2(x)\,\mathrm{d}\sigma\right)^p\left(\int h^2(x)\,\mathrm{d}\sigma\right)^p - \left(\int g(x)h(x)\,\mathrm{d}\sigma\right)^{2p} \\ \geq \left(\int g^2(x)\,\mathrm{d}\sigma \int h^2(x)\,\mathrm{d}\sigma - \left(\int g(x)h(x)\,\mathrm{d}\sigma\right)^2\right)^p \\ \geq \left(\frac{\int g^2(x)\,\mathrm{d}\sigma}{\int f^2(x)\,\mathrm{d}\sigma}\right)^p,\end{aligned}$$

which gives the stated result.

If $\int h(x)\tilde{f}(x)\,\mathrm{d}\sigma \neq 0$, then from the substitution

$$f(x) = \frac{\tilde{f}(x)}{\int h(x)\tilde{f}(x)\,\mathrm{d}\sigma}$$

we obtain the following result.

THEOREM 3 *Suppose* g, h *and* $\tilde{f}$ *are functions such that* $g^2, h^2, \tilde{f}^2 \in \mathcal{L}_1(\mathbf{R}, \sigma)$,

$$\int g(x)\tilde{f}(x)\,\mathrm{d}\sigma = 0 \quad \textit{and} \quad \int \tilde{f}^2(x)\,\mathrm{d}\sigma \neq 0.$$

Then

$$\begin{aligned}\left(\int g^2(x)\,\mathrm{d}\sigma\right)^p\left(\int h^2(x)\,\mathrm{d}\sigma\right)^p - \left(\int g(x)h(x)\,\mathrm{d}\sigma\right)^{2p} \\ \geq \frac{(\int g^2(x)\,\mathrm{d}\sigma)^p(\int h(x)\tilde{f}(x)\,\mathrm{d}\sigma)^{2p}}{(\int \tilde{f}^2(x)\,\mathrm{d}\sigma)^p}.\end{aligned} \tag{5}$$

Remark 1 The result of Theorem 2 can be improved. Suppose that g, h and f are as in Theorem 2 and G is a nondecreasing, superadditive function. Then

$$G\left(\int g^2(x)\,\mathrm{d}\sigma \int h^2(x)\,\mathrm{d}\sigma\right) - G\left(\left(\int g(x)h(x)\,\mathrm{d}\sigma\right)^2\right) \geq G\left(\frac{\int g^2(x)\,\mathrm{d}\sigma}{\int f^2(x)\,\mathrm{d}\sigma}\right).$$

In particular, this inequality holds for any nondecreasing, convex function G.

3 APPLICATIONS TO MOMENTS

Let $F\colon \mathbf{R} \to \mathbf{R}$ be a probability distribution function and suppose that the corresponding mean $a = \int_{\mathbf{R}} x\,\mathrm{d}F(x)$ exists. The rth central moment of F, when the integral exists, is defined by

$$\mu_r = \int_{\mathbf{R}} (x-a)^r\,\mathrm{d}F(x).$$

We have trivially that $\mu_1 = 0$.

Suppose the distribution has variance unity, so that $\mu_2 = 1$. On setting $\tilde{f}(x) = 1$ and $g(x) = x-a$ in (5) we obtain since $\int \mathrm{d}F(x) = 1$ and $\mu_1 = 1$ that

$$\left(\int h^2(x)\,\mathrm{d}F(x)\right)^p - \left(\int (x-a)h(x)\,\mathrm{d}F(x)\right)^{2p} \geq \left(\int h(x)\,\mathrm{d}F(x)\right)^{2p}. \tag{6}$$

By using substitutions of the form

$$h(x) = \sum_{k\in J} c(x-a)^k, \quad J \subseteq \mathbf{Z}$$

we can get different inequalities for the central moments.

Thus on putting $h(x)=(x-a)^r+\lambda(x-a)^s+\mu$ in (6), where $\lambda,\mu\in\mathbf{R}$ and $r,s\in\mathbf{Z}$, we get

$$(\mu_{2r}+\lambda^2\mu_{2s}+\mu^2+2\lambda\mu_{r+s}+2\mu\mu_r+2\lambda\mu\mu_s)^p \geq (\mu_{r+1}+\lambda\mu_{s+1})^{2p}+(\mu_r+\lambda\mu_s+\mu)^{2p}.$$

So in particular for $r=2$, $s=1$ we have

$$(\mu_4+2\lambda\mu_3+\lambda^2+\mu^2+2\mu)^p\geq(\mu_3+\lambda)^{2p}+(1+\mu)^{2p}$$

and for $\lambda=\mu=0$ we have

$$(\mu_{2r})^p\geq(\mu_{r+1})^{2p}+(\mu_r)^{2p}$$

(cf. [1,3]).

4 NONUNIFORM WEIGHTS

In [1], Alić and Pečarić used the substitutions $z_i=1/\sum_{i=1}^n b_i(1\leq i\leq n)$ to give a useful corollary to Theorem B.

If (y_i) *is an n-tuple such that* $\sum y_i=0$ *and* $\sum y_i^2=n$, *then*

$$\left(\frac{1}{n}\sum b_i^2\right)^p\geq\left(\frac{1}{n}\sum y_ib_i\right)^{2p}+\left(\frac{1}{n}\sum b_i\right)^{2p}. \tag{7}$$

Using substitutions of the form

$$b_i=\sum_{k\in J}c_iy_i^k,\quad J\subseteq\mathbf{Z},\ i=1,\ldots,n$$

and the notation $\alpha_r:=(1/n)\sum_{i=1}^n y_i^r$ they obtained many improvements and generalizations of known statistical inequalities given in [2,3,7,8]. See also [4, pp. 339–340]. We show that the uniform weighting $1/n$ can be replaced by a general probabilistic weighting p_i with $\sum_{i=1}^n p_i=1$.

Let F be the probability distribution function of the discrete random variable X with $P\{X=x_k\}=p_k$, $k\in N$, so that X has expectation $a=\sum_k x_kp_k$. If the variance of X is equal to unity, that is,

$\sum_k (x_k - a)^2 p_k = 1$, then (6) assumes the form

$$\left(\sum_i p_i b_i^2\right)^p \geq \left(\sum_i p_i b_i\right)^{2p} + \left(\sum_i p_i y_i b_i\right)^{2p},$$

where $y_i := x_i - a$. In the case $p_i = 1/n$ $(1 \leq i \leq n)$ this reduces to (7).

5 MULTIPLE LINEAR CONSTRAINTS

We now proceed to higher-dimensional versions of Theorems A and 1. We start with the former, replacing (1) with sets of constraints

$$\sum_{i=1}^{n} z_i a_{i,j} = 0 \quad (1 \leq j \leq m)$$

and

$$\sum_{i=1}^{n} z_i b_{i,j} = 1 \quad (1 \leq j \leq r).$$

Typically we expect $m + r < n$ in applications.

We shall assume that the columns of the matrix $A_0 = (a_{i,j})$ are linearly independent, which by Gram's inequality (see, for example, [5, Ch. 20 Theorem 1] implies that the matrix $A := A_0^T A_0$ be invertible.

THEOREM 4 *Let A_0, B_0 be respectively $n \times m$ and $n \times r$ real matrices and let z be a real column n-vector satisfying*

$$z^{\mathrm{T}} A_0 = 0 \quad and \quad z B_0 = e_r^{\mathrm{T}}, \tag{8}$$

where e_t represents the column t-vector $(1, 1, \ldots, 1)^{\mathrm{T}}$. We suppose that the columns of A_0 are linearly independent, so that $A := A_0^T A_0$ is invertible. We define $B := B_0^{\mathrm{T}} B_0$, $C := A_0^{\mathrm{T}} B_0$ and suppose that B_0 is such that $B - C^{\mathrm{T}} A^{-1} C$ is also invertible. We denote its inverse by K. Then

$$z^{\mathrm{T}} z = \sum_{i=1}^{n} z_i^2 \geq e_r^{\mathrm{T}} K e_r,$$

with equality if and only if

$$z = (B_0 - A_0 A^{-1} C) K e_r. \tag{9}$$

Proof The vector y given by the right-hand side of (9) satisfies

$$y^{\mathrm{T}} A_0 = e_r^{\mathrm{T}} K^{\mathrm{T}} (B_0^{\mathrm{T}} A_0 - C^{\mathrm{T}} A^{-1} A_0^{\mathrm{T}} A_0) = e_r^{\mathrm{T}} K^{\mathrm{T}} (C^{\mathrm{T}} - C^{\mathrm{T}} A^{-1} A) = 0$$

and

$$y^{\mathrm{T}} B_0 = e_r^{\mathrm{T}} K^{\mathrm{T}} (B_0^{\mathrm{T}} B_0 - C^{\mathrm{T}} A^{-1} A_0^{\mathrm{T}} B_0) = e_r^{\mathrm{T}} K (B - C^{\mathrm{T}} A^{-1} C) = e_r^{\mathrm{T}},$$

and so meets the conditions of the enunciation. Also, if z is any solution to (8), then

$$z^{\mathrm{T}} y = z^{\mathrm{T}} (B_0 - A_0 A^{-1} C) K e_r = e_r^{\mathrm{T}} K e_r,$$

and in particular

$$y^{\mathrm{T}} y = \sum_{i=1}^{n} y_i^2 = e_r^{\mathrm{T}} K e_r.$$

Any vector z subject to (8) therefore satisfies

$$z^{\mathrm{T}} z - y^{\mathrm{T}} y = \sum_{i=1}^{n} z_i^2 - \sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} (z_i - y_i)^2,$$

which gives the stated result.

For the integral result, we replace (3) by the set of constraints

$$\int g_j(x) f(x) \, \mathrm{d}\sigma = 0 \quad (1 \leq j \leq m)$$

and

$$\int h_j(x) f(x) \, \mathrm{d}\sigma = 1 \quad (1 \leq j \leq r).$$

We assume the functions g_j are linearly independent.

THEOREM 5 *Let σ be a nonnegative measure on* $\mathbf{R}$ *and* f, $g = (g_j)$, $h = (h_j)$ *respectively scalar, column m-vector and column r-vector valued functions from* $\mathbf{R}$ *to* $\mathbf{R}$ *with square-integrable components with respect to* σ *with*

$$\int g(x)f(x)\,\mathrm{d}\sigma = 0 \quad \textit{and} \quad \int h(x)f(x)\,\mathrm{d}\sigma = e_r.$$

Define matrices A, B, C *by*

$$A_{i,j} = \int g_i(x)g_j(x)\,\mathrm{d}\sigma,$$

$$B_{i,j} = \int h_i(x)h_j(x)\,\mathrm{d}\sigma,$$

$$C_{i,j} = \int g_i(x)h_j(x)\mathrm{d}\sigma.$$

Let (g_i) *be a linearly independent set, so that* A *is invertible, and suppose that the matrix* $B - C^{\mathrm{T}}A^{-1}C$ *is invertible, with inverse* K, *say. Then*

$$\int f^2(x)\,\mathrm{d}\sigma \geq e_r^{\mathrm{T}} K e_r,$$

with equality if and only if

$$f = e_r^{\mathrm{T}} K(h - C^{\mathrm{T}}A^{-1}g).$$

The proof parallels that of the previous theorem, *mutatis mutandis.*

References

[1] M. Alić and J. Pečarić, On some inequalities of Madevski and Ostrowski, *Rad. Hrvatske Akad. Znan. Umjet.* (to appear).

[2] M. Lakshmanamurti, On the upper bound of $\sum_{i=1}^{n} x_i^m$ subject to the conditions $\sum_{i=1}^{n} x = 0$ and $\sum_{i=1}^{n} x_i^2 = n$, *Math. Student* **18** (1950), 111–116.

[3] Ž. Madevski, Quelques conséquences d'une inégalité d'Ostrowski, *Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat. Fiz.* **412–460** (1973), 167–170.

[4] D.S. Mitrinović, *Analytic Inequalities*, Springer-Verlag, Heidelberg (1970).

[5] D.S. Mitrinović, J.E. Pečarić and A.M. Fink, *Classical and New Inequalities in Analysis*, Kluwer, Dordrecht (1993).

[6] A.M. Ostrowski, *Vorlesungen über Differential und Integralrechnung II*, Birkhauser, Basel (1951).

[7] K. Pearson, Mathematical contributions to the theory of evolution. XIX: second supplement to a memoir on skew variation, *Phil. Trans. Roy. Soc. A* **216** (1916), 432.

[8] J.E. Wilkins, A note on skewness and kurtosis, *Ann. Math. Statist.* **15** (1944), 333–335.