A NOTE ON OSTROWSKI'S INEQUALITY

AURELIA FLOREA AND CONSTANTIN P. NICULESCU

This paper deals with the problem of estimating the deviation of the values of a function from its mean value. We consider the following special cases: (i) the case of random variables (attached to arbitrary probability fields); (ii) the comparison is performed additively or multiplicatively; (iii) the mean value is attached to a multiplicative averaging process.

1. Introduction

The inequality of Ostrowski [7] gives us an estimate for the deviation of the values of a smooth function from its mean value. More precisely, *if $f : [a,b] \to \mathbb{R}$ is a differentiable function with bounded derivative, then*

$$\left| f(x) - \frac{1}{b-a} \int_a^b f(t)dt \right| \le \left[\frac{1}{4} + \frac{(x-(a+b)/2)^2}{(b-a)^2} \right] (b-a)\|f'\|_\infty \qquad \text{(O)}$$

for every $x \in [a,b]$. Moreover the constant $1/4$ is the best possible.

The proof is an application of Lagrangian's mean value theorem:

$$
\begin{aligned}
\left| f(x) - \frac{1}{b-a} \int_a^b f(t)dt \right| &= \left| \frac{1}{b-a} \int_a^b (f(x) - f(t))dt \right| \\
&\le \frac{1}{b-a} \int_a^b |f(x) - f(t)|\,dt \\
&\le \frac{\|f'\|_\infty}{b-a} \int_a^b |x-t|\,dt \\
&= \left[\frac{(x-a)^2 + (b-x)^2}{2(b-a)} \right] \|f'\|_\infty \\
&= \left[\frac{1}{4} + \left(\frac{x-(a+b)/2}{b-a} \right)^2 \right] (b-a)\|f'\|_\infty.
\end{aligned}
\qquad (1.1)
$$

The optimality of the constant 1/4 is also immediate, checking the inequality for the family of functions $f_\alpha(t) = |x - t|^\alpha \cdot (b - a)$ $(t \in [a,b], \ \alpha > 1)$ and then passing to the limit as $\alpha \to 1+$.

It is worth to notice that the smoothness condition can be relaxed. In fact, the Lipschitz class suffices as well, by replacing $\|f'\|_\infty$ with the Lipschitz constant of f, that is,

$$\|f\|_L = \sup_{x \neq y} \left| \frac{f(x) - f(y)}{x - y} \right|. \tag{1.2}$$

The extension to the context of vector-valued functions, with values in a Banach space, is straightforward.

Since a Lipschitz function on $[a,b]$ is absolutely continuous, a natural direction of generalization of the Ostrowski inequality was its investigation within this larger class of functions (with refinements for $f' \in L^p([a,b])$, $1 \leq p < \infty$). See Fink [2]. Also, several Ostrowski type inequalities are known within the framework of Hölder functions as well as for functions of bounded variation.

The problem to estimate the deviation of a function from its mean value can be investigated from many other points of view:

  (i) by considering random variables (attached to arbitrary probability fields);
 (ii) by changing the algebraic nature of the comparison (e.g., switching to the multiplicative framework);
(iii) by considering other means (e.g., the geometric mean);
(iv) by estimating the deviation via other norms (the classical case refers to the sup norm, but L^p-norms are better motivated in other situations).

The aim of this paper is to present a number of examples giving support to this program.

2. Ostrowski type inequalities for random variables

In what follows, X will denote a locally compact metric space and E a Banach space.

THEOREM 2.1. *The following two assertions are equivalent for $f : X \to E$ a continuous mapping:*

  (i) *f is Lipschitz that is, $\|f\|_L = \sup_{x \neq y}(\|f(x) - f(y)\|/d(x,y)) < \infty$;*
 (ii) *for every $x \in X$ and every Borel probability measure μ on X such that $f \in \mathcal{L}^1(\mu)$ we have*

$$\left\| f(x) - \int_X f \, d\mu \right\| \leq \|f\|_L \int_X^* d(x,y) d\mu, \tag{2.1}$$

here $*$ marks the upper integral.

Proof. (i)$\Rightarrow$(ii). As $d(x, \cdot)$ is continuous it is also Borel measurable, so being nonnegative its upper integral is perfectly motivated. Then we can proceed as in the classical case, described in Section 1.

(ii)$\Rightarrow$(i). Consider the particular case of Dirac measure δ_y (concentrated at y). Then

$$\|f(x) - f(y)\| \leq \|f\|_L d(x, y) \tag{2.2}$$

which shows that f must be Lipschitz.   □

If X is a bounded metric space, then the above theorem works for all continuous mappings. In fact, if $\|f\|_L < \infty$, then f is necessarily bounded (and thus it belongs to $\mathscr{L}^\infty(\mu) \subset \mathscr{L}^1(\mu)$). Also, the mappings $d(x, \cdot)$ are μ-integrable (which makes $*$ unnecessary).

The condition $f \in \mathscr{L}^1(\mu)$ is automatically fulfilled by all continuous bounded functions regardless what Borel probability measure μ we consider on X; in fact, they are in $\mathscr{L}^\infty(\mu) \subset \mathscr{L}^1(\mu)$. In general, not every continuous function f is μ-integrable. For, think at the case where $X = \mathbb{R}$, $f(x) = x$, and $\mu = (1/\pi(1 + x^2))dx$.

We will illustrate Theorem 2.1 in a number of particular situations. The first one concerns the case of classical probability fields.

COROLLARY 2.2. *Let E be a normed vector space and let $x_1, \ldots, x_n$ be n vectors in E. Then, for $i = 1, \ldots, n$,*

$$\left\| x_i - \frac{1}{n} \sum_{k=1}^n x_k \right\| \leq \frac{1}{n} \left[\left(i - \frac{n+1}{2} \right)^2 + \frac{n^2 - 1}{4} \right] \cdot \sup_{1 \leq k \leq n-1} \|x_{k+1} - x_k\|. \tag{2.3}$$

Proof. We consider the measure space (X, Σ, μ), where $X = \{1, \ldots, n\}$, $\Sigma = \mathscr{P}(X)$ and $\mu(A) = |A|$ for every $A \subset X$.

X has a natural structure of metric subspace of $\mathbb{R}$. The function

$$f : X \longrightarrow E, \quad f(i) = x_i, \tag{2.4}$$

is Lipschitz, with Lipschitz constant

$$L = \sup_{1 \leq k \leq n-1} \|x_{k+1} - x_k\|. \tag{L}$$

In fact, if $i < j$, then

$$\begin{aligned}
\|f(i) - f(j)\| &= \|x_i - x_j\| \\
&\leq \|x_i - x_{i+1}\| + \cdots + \|x_{j-1} - x_j\| \\
&\leq (j - i) \cdot \sup_{1 \leq k \leq n-1} \|x_{k+1} - x_k\|,
\end{aligned} \tag{2.5}$$

which proves the inequality $\leq$ in (L). The other inequality is clear.

According to Theorem 2.1,

$$\left\| f(i) - \frac{1}{\mu(X)} \int_X f(k) d\mu(k) \right\| \leq \frac{L}{\mu(X)} \int_X |i - k| d\mu(k), \tag{2.6}$$

which can be easily shown to be equivalent to the inequality in the statement of Corollary 2.2 because

$$\int_X |i-k| \, d\mu(k) = \sum_{k=1}^{n} |i-k| = \left(i - \frac{n+1}{2}\right)^2 + \frac{n^2-1}{4}. \tag{2.7}$$

$\square$

Notice that the right-hand side of the inequality in Corollary 2.2 is $\geq \sqrt{\mathrm{var}(f)}$, where

$$\mathrm{var}(f) = \frac{1}{\mu(X)} \int_X \left| f(x) - \frac{1}{\mu(X)} \int_X f(t) d\mu(t) \right|^2 d\mu(x) \tag{2.8}$$

represents the *variance* of f. According to the classical Chebyshev inequality,

$$\mu\left\{ \left| f - \frac{1}{\mu(X)} \int_X f \, d\mu \right| \leq \varepsilon \right\} \geq 1 - \frac{\mathrm{var}\, f}{\varepsilon^2}, \tag{2.9}$$

and the discussion above shows that the range of interest in this inequality is precisely

$$\mathrm{var}\, f < \varepsilon \leq \frac{1}{n}\left[\left(i - \frac{n+1}{2}\right)^2 + \frac{n^2-1}{4} \right] \cdot \sup_{1 \leq k \leq n-1} \|x_{k+1} - x_k\|. \tag{2.10}$$

As well known, convolution by smooth kernels leads to good approximation schemes. Theorem 2.1 allows us to estimate the speed of convergence. Here is an example.

COROLLARY 2.3. *Let $f : \mathbb{R} \to E$ be a Lipschitz mapping. Then*

$$\left\| f(x) - \frac{n}{(2\pi)^{1/2}} \int_{\mathbb{R}} f(t) e^{-n^2(x-t)^2/2} dt \right\| \leq \frac{2\|f\|_L}{n(2\pi)^{1/2}} \tag{2.11}$$

for every $x \in \mathbb{R}$ and every positive integer n. Particularly, f is the uniform limit of a sequence $(f_n)_n$ of Lipschitz functions of class C^∞, with $\|f_n\|_L \leq \|f\|_L$ for every n.

We end this section with the case of functions of several variables.

COROLLARY 2.4. *Let $f = f(x, y)$ be a differentiable function defined on a compact 2-dimensional interval $R = [a, b] \times [c, d]$ such that $|\partial f/\partial x| \leq L$ and $|\partial f/\partial y| \leq M$ on R. Then*

$$\left\| f(x, y) - \frac{1}{\mathrm{Area}\, R} \iint_R f(u, v) du \, dv \right\|$$

$$\leq L\left[\frac{1}{4} + \left(\frac{x - (a+b)/2}{b-a} \right)^2 \right] \mathrm{Area}\, R + M\left[\frac{1}{4} + \left(\frac{y - (c+d)/2}{d-c} \right)^2 \right] \mathrm{Area}\, R. \tag{2.12}$$

Proof. Clearly, we may assume that $L, M > 0$. Then f is a Lipschitz function (with Lipschitz constant 1) provided that R is endowed with the metric

$$d((x, y), (u, v)) = L|x - u| + M|y - v|. \tag{2.13}$$

Now the conclusion follows from Theorem 2.1.

$\square$

3. Comparison of arithmetic means

Suppose that X is a locally compact bounded metric space on which there are given two Borel probability measures μ and ν. We are interested to estimate the difference

$$\int_X f\,d\mu - \int_X f\,d\nu \tag{3.1}$$

for f a Lipschitz function on X, with values in a Banach space E. For $\nu = \delta_x$, this reduces to the classical Ostrowski inequality.

Following the ideas in the preceding section we are led to

$$\left\| \int_X f\,d\mu - \int_X f\,d\nu \right\| \le \|f\|_L \int_X \int_X d(x,y)\,d\mu(x)\,d\nu(y) \tag{3.2}$$

but this is not always the best result.

For example, for $\mu = dx/(b-a)$ and $\nu = (\delta_a + \delta_b)/2$ (on $X = [a,b]$) it yields the trapezoid inequality

$$\left\| \frac{1}{b-a}\int_a^b f(x)dx - \frac{f(a)+f(b)}{2} \right\| \le C\|f\|_L(b-a), \tag{3.3}$$

with $C = 1/2$. This can be improved up to $C = 1/4$ within the Ostrowski theory (by applying (O) to $f\,|\,[a,(a+b)/2]$ and $f\,|\,[(a+b)/2,b]$, for $x = a$ and $x = b$, resp.). However, the Iyengar inequality gives us a better upper bound:

$$\left| \frac{1}{b-a}\int_a^b f(x)dx - \frac{f(a)+f(b)}{2} \right| \le \frac{\|f\|_L}{4}(b-a) - \frac{(f(b)-f(a))^2}{4(b-a)\|f\|_L}. \tag{Iy}$$

See [3, 4], or [6] for details. We can combine (O) and (Iy) to get a more general result:

$$
\begin{aligned}
&\left| \frac{1}{b-a}\int_a^b f(x)dx - \left[\lambda f\left(\frac{a+b}{2}\right) + (1-\lambda)\frac{f(a)+f(b)}{2} \right] \right| \\
&\quad \le \frac{\|f\|_L}{4}(b-a) - (1-\lambda)\frac{(f(b)-f(a))^2}{4(b-a)\|f\|_L},
\end{aligned}
\tag{3.4}
$$

for every Lipschitz function f and every $\lambda \in [0,1]$. This is optimal in the Lipschitz class, but better results are known for smooth functions. See [4].

There is a large activity (motivated by the problems in numerical integration) concerning the approximation of probability measures by convex combinations of Dirac measures. However, sharp general formulas remain to be found.

4. The multiplicative setting

According to [5], the *multiplicative mean value* of a continuous function $f : [a,b] \to (0,\infty)$ (where $0 < a < b$) is defined by the formula

$$
\begin{aligned}
M_\star(f) &= \exp\left(\frac{1}{\log b - \log a} \int_{\log a}^{\log b} \log f(e^t)\, dt \right) \\
&= \exp\left(\frac{1}{\log b - \log a} \int_a^b \log f(t)\, \frac{dt}{t} \right).
\end{aligned}
\tag{4.1}
$$

Thus $M_\star(f)$ represents the geometric mean of f with respect to the measure dt/t. The main properties of the multiplicative mean are listed below:

$$
\begin{aligned}
M_\star(1) &= 1, \\
m \le f \le M &\Longrightarrow m \le M_*(f) \le M, \\
M_*(fg) &= M_*(f) M_*(g).
\end{aligned}
\tag{4.2}
$$

Given a function $f : I \to (0,\infty)$ (with $I \subset (0,\infty)$) we will say that f is *multiplicatively Lipschitz* provided there exists a constant $L > 0$ such that

$$
\max\left\{ \frac{f(x)}{f(y)}, \frac{f(y)}{f(x)} \right\} \le \left(\frac{y}{x} \right)^L
\tag{4.3}
$$

for all $x < y$ in I; the smallest L for which the above inequality holds constitutes the *multiplicative Lipschitz* constant of f and it will be denoted by $\|f\|_{\star\,\mathrm{Lip}}$.

Remark 4.1. Though the family of multiplicatively Lipschitz functions is large enough (to deserve attention in its own), we know the exact value of the multiplicative Lipschitz constant only in a few cases.
   (i) If f is of the form $f(x) = x^\alpha$, then $\|f\|_{\star\,\mathrm{Lip}} = \alpha$.
   (ii) If $f = \exp | [a,b]$ (where $0 < a < b$), then $\|f\|_{\star\,\mathrm{Lip}} = b$.
   (iii) Clearly, $\|f\|_{\star\,\mathrm{Lip}} \le 1$ for every nondecreasing functions f such that $f(x)/x$ is nonincreasing. For example, this is the case of the functions sin and sec on $(0,\pi/2)$ and the quasi-concave functions on $(0,\infty)$. The latter class of functions plays an important role in interpolation theory. See [1].
   (iv) If f and g are two multiplicatively Lipschitz functions (defined on the same interval) and $\alpha, \beta \in \mathbb{R}$, then $f^\alpha g^\beta$ is multiplicatively Lipschitz too. Moreover,

$$
\|f^\alpha g^\beta\|_{\star\,\mathrm{Lip}} \le |\alpha| \cdot \|f\|_{\star\,\mathrm{Lip}} + |\beta| \cdot \|g\|_{\star\,\mathrm{Lip}}.
\tag{4.4}
$$

The following result represents the multiplicative counterpart of the classical Ostrowski inequality.

Theorem 4.2. *Let $f : [a,b] \to (0,\infty)$ be a multiplicatively Lipschitz function with $\|f\|_{*\,\mathrm{Lip}} = L$. Then*

$$\max\left\{\frac{f(x)}{M_*(f)}, \frac{M_*(f)}{f(x)}\right\} \le \left(\frac{b}{a}\right)^{L(1/4+\log^2(x/\sqrt{ab})/\log^2(b/a))}. \tag{4.5}$$

Proof. In fact,

$$\frac{M_*(f)}{f(x)} = \exp\left(\frac{1}{\log b - \log a}\int_a^b \log\frac{f(t)}{f(x)}\frac{dt}{t}\right)$$

$$\le \exp\left(\frac{L}{\log(b/a)}\left(\int_a^x \log\frac{x}{t}\frac{dt}{t} + \int_x^b \log\frac{t}{x}\frac{dt}{t}\right)\right)$$

$$= \exp\left(\frac{L}{\log(b/a)}\left(\log x(2\log x - \log a - \log b) - \log^2 x + \frac{\log^2 a}{2} + \frac{\log^2 b}{2}\right)\right)$$

$$= \exp\left(\frac{L}{\log b/a} \cdot \frac{(\log x - \log a)^2 + (\log b - \log x)^2}{2}\right)$$

$$= \exp\left(\frac{L}{\log b/a} \cdot \frac{(\log b - \log a)^2 + 4(\log x - (\log a + \log b)/2)^2}{4}\right)$$

$$= \left(\frac{b}{a}\right)^{L(1/4+\log^2(x/\sqrt{ab})/\log^2(b/a))} \tag{4.6}$$

and a similar estimate is valid for $f(x)/M_*(f)$. □

For $f = \exp \mid [a,b]$ (where $0 < a < b$), we have $M_*(f) = \exp((b-a)/(\log b - \log a))$ and $\|f\|_{*\,\mathrm{Lip}} = b$. By Theorem 4.2, we infer the inequality

$$\exp\left|\frac{b-a}{\log b - \log a} - x\right| \le \left(\frac{b}{a}\right)^{b(1/4+\log^2(x/\sqrt{ab})/\log^2(b/a))}, \tag{4.7}$$

that is,

$$\left|\frac{b-a}{\log b - \log a} - x\right| \le b\left(1/4 + \frac{\log^2(x/\sqrt{ab})}{(\log b - \log a)^2}\right)(\log b - \log a) \tag{4.8}$$

for every $x \in [a,b]$. Particularly,

$$\left|\frac{b-a}{\log b - \log a} - \sqrt{ab}\right| \le \frac{b}{4}(\log b - \log a). \tag{4.9}$$

5. Open problems

The deviation of the values of a function from its mean value can be estimated via a variety of norms. For example, the Ostrowski inequality yields

$$\left\| f(x) - \frac{1}{b-a} \int_a^b f(t)dt \right\|_\infty \leq \frac{b-a}{2} \|f'\|_\infty \tag{O'}$$

for every $f \in C^1([a,b])$.

In some instances, the L^p-norms are more suitable. An old result in this direction is the following inequality due to Stekloff (see [4, 10, 11]),

$$\left(\int_a^b \left| f(x) - \frac{1}{b-a} \int_a^b f(t)dt \right|^2 dx \right)^{1/2} \leq \frac{b-a}{\pi} \left(\int_a^b |f'(x)|^2 dx \right)^{1/2} \tag{S}$$

that works also for every $f \in C^1([a,b])$. In terms of variance, (S) may be read as

$$\operatorname{var}(f) \leq \frac{b-a}{\pi^2} \int_a^b |f'(t)|^2 dt, \tag{5.1}$$

so that, combined with the Schwarz inequality, it yields the following estimate for the covariance of two random variables (of class C^1):

$$\begin{aligned}
\operatorname{cov}(f,g) &\leq \operatorname{var}^{1/2}(f) \cdot \operatorname{var}^{1/2}(g) \\
&\leq \frac{b-a}{\pi^2} \left(\int_a^b |f'(t)|^2 dt \right)^{1/2} \left(\int_a^b |g'(t)|^2 dt \right)^{1/2}.
\end{aligned} \tag{5.2}$$

There is a large literature in this area, including deep results in the higher dimensional case. See [4].

A natural way to pack together (O') and (S) is as follows.

Consider a separable Banach lattice E. Then E contains quasi-interior points $u > 0$ and admits strictly positive functionals $x' \in E'$. This means that

$$\lim_{n \to \infty} \|x - x \wedge nu\| = 0 \quad \text{for every } x \in E, \ x \geq 0,$$
$$x \geq 0, \quad x'(x) = 0 \quad \text{implies } x = 0. \tag{5.3}$$

See Schaefer [8], for details.

To each such a pair (u, x') with $x'(u) = 1$, one can associate a positive linear projection,

$$M : E \longrightarrow E, \quad M(x) = x'(x) \cdot u, \tag{5.4}$$

whose image is $\mathbb{R} \cdot u$. Clearly, this projection provides an analogue of the integral mean.

Problem 5.1. Characterize all the pairs (u, x') as above, for which there exist a densely defined linear operator $D : \operatorname{dom} D \subset E \to E$ and a positive constant $C = C(x', u)$ such that

$$\|x - M(x)\| \leq C \|D(x)\| \quad \text{for every } x \in \operatorname{dom} D. \tag{5.5}$$

In the examples at the beginning of this section, E is one of the spaces $C([a,b])$ or $L^2([a,b])$, u is the function identically 1, x' is the normalized Lebesgue measure, and D is the differential. The same picture follows from Corollary 2.4 above. The problem is how general could be the existence of a differential like operator in the context of separable Banach lattices.

Another open problem refers to the probabilistic approach of measuring the deviation (and it is related to the concentration of measure phenomenon). Assume that K is the unit ball of $\mathbb{R}^n$ associated to the norm

$$\|x\|_1 = \sum_{k=1}^{n} |x_k| \tag{5.6}$$

and denote by m the normalized Lebesgue measure restricted to K. According to a result due to Schechtman and Zinn [9], there exist two absolute constants $C > 0$ and $\alpha > 0$ such that

$$m\left(\left\{x \mid \left| f(x) - \int_K f\, dm \right| > t\right\}\right) \le Ce^{-\alpha nt} \tag{SZ}$$

for all functions $f : K \to \mathbb{R}$ which are Lipschitz when K is endowed with the Euclidean metric. A similar result, for m replaced by the Gaussian probability measure on $\mathbb{R}^n$ ($d\gamma(x) = (2\pi)^{-n/2}e^{-|x|^2/2}dx$), was proved by Talagrand [12].

Problem 5.2. How general are the Borel probability spaces $(K, \mathcal{B}(K), m)$ for which estimates of the type (SZ) work for all Lipschitz functions $f : K \to \mathbb{R}$?

Acknowledgments

References

[1]   J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer, Berlin, 1976.

[2]   A. M. Fink, *Bounds on the deviation of a function from its averages*, Czechoslovak Math. J. **42(117)** (1992), no. 2, 289–310.

[3]   K. S. K. Iyengar, *Note on an inequality*, Math. Student **6** (1938), 75–76.

[4]   D. S. Mitrinović, J. E. Pečarić, and A. M. Fink, *Inequalities Involving Functions and Their Integrals and Derivatives*, Mathematics and Its Applications (East European Series), vol. 53, Kluwer Academic, Dordrecht, 1991.

[5]   C. P. Niculescu, *A multiplicative mean value and its applications*, Inequality Theory and Applications. Vol. I (Y. J. Cho, J. K. Kim, and S. S. Dragomir, eds.), Nova Science, New York, 2001, pp. 243–255.

[6]   C. P. Niculescu and F. Popovici, *A note on the Denjoy-Bourbaki theorem*, Real Anal. Exchange **29** (2003/04), no. 2, 639–646.

[7]   A. Ostrowski, *Über die Absolutabweichung einer differenzierbaren Funktion von ihrem Integralmittelwert*, Comment. Math. Helv. **10** (1938), 226–227 (German).

[8]   H. H. Schaefer, *Banach Lattices and Positive Operators*, Die Grundlehren der mathematischen Wissenschaften, vol. 215, Springer, New York, 1974.

[9]   G. Schechtman and J. Zinn, *Concentration on the ℓ_p^n ball*, Geometric Aspects of Functional Analysis. Proc. Israel Seminar, 1996–2000, Lecture Notes in Math., vol. 1745, Springer, Berlin, 2000, pp. 245–256.

[10]   W. Stekloff, *Sur le problème de refroidissment d'une barre hétérogène*, C. R. Sci. Paris **126** (1898), 215–218 (French).

[11]   ———, *Problème de refroidissment d'une barre hétérogène*, Ann. Fac. Sci. Toulouse **3** (1901), no. 2, 281–313 (French).

[12]   M. Talagrand, *A new isoperimetric inequality and the concentration of measure phenomenon*, Geometric Aspects of Functional Analysis. Proc. Israel Seminar, 1989–90, Lecture Notes in Math., vol. 1469, Springer, Berlin, 1991, pp. 94–124.

Aurelia Florea: Department of Philology, University of Craiova, Craiova 200585, Romania
E-mail address: aurelia_florea@yahoo.com

Constantin P. Niculescu: Department of Mathematics, University of Craiova, Craiova 200585, Romania
E-mail address: cniculescu@central.ucv.ro