# —NOTES—

## EXPERIMENTS IN THE SMOOTHING OF DATA*

By MARK LOTKIN (*Avco Research and Advanced Development Division*)

**1. Introduction.** When using sequences of observed data for the evaluation of certain desired functions of these data it is frequently necessary to apply some smoothing process in order to avoid unreasonable results. In such cases the points of paramount concern are then not only what smoothing process to choose, but also when to apply it. The following remarks deal mainly with the "when" rather than the "how." In particular, the question of argument smoothing versus function smoothing is discussed here, by means of a few simple examples. Both functional values and derivatives are considered.

While the results obtained may not be of the greatest generality they are, however, considered sufficiently indicative to be of some practical utility.

**2. Definitions and specifications.** Let the exact values of a function $u - f(t)$ for a sequence $t_i$, $i = 1, 2, \cdots, m$, of the independent variable $t$ be denoted by $u_i$, and let us assume that there are also available observed values of $u$, denoted by $u_i'$. Of interest are the magnitudes of the errors $e_i(u')$, $e_i(u') = u_i' - u_i$, as well as the length $\| e(u') \|$ of the error vector:

$$\| e(u') \|^2 = \sum_{i=1}^{m} e_i^2(u'). \tag{1}$$

A process which replaces the $u_i'$ by values $u_i^*$ that differ as little as possible from the corresponding $u_i'$ while diminishing the vector $\| e(u^*) \|$, $e_i(u^*) = u_i^* - u_i$, will be called a "smoothing" process. The particular smoothing process to be considered here is based on a moving least squares polynomial of fixed degree [1]:

$$p(t) = a_0 + a_1 t + \cdots + a_r t^r.$$

The smoothing consists in putting $u_i^* = p(t_i)$, $U_i^* = (dp/dt)_i$. For the sake of simplicity we assume that $m = 2n + 1$ is odd, and that the sequence $t_i$ is equidistant: $t_i = t_1 + (i - 1)\Delta t$, $i = 1, 2, \cdots, 2n + 1$.

Then the usual transformations

$$\tau_i = (t_i - t_{n+1})/\Delta t,$$

$$v_{i-n-1} = u_i', \qquad i = 1, 2, \cdots, 2n + 1,$$

move the midpoint from $t_{n+1}$ to $\tau_{n+1} = 0$, change the $t_i$ to $\tau_i = i - n - 1$, and center the $u_i'$ about $v_0$. Further, $p(t)$ transforms into a polynomial $\pi(\tau)$, so that

$$p(t_i) = \pi(\tau_i) = \sum_{\rho=0}^{r} \alpha_\rho \tau_i^\rho ,$$

and, consequently,

$$u_{n+1}^* \equiv \pi(0) = \alpha_0 , \tag{2}$$

*Received March 5, 1957.

$$U^*_{n+1} \equiv \alpha_1/\Delta t. \tag{3}$$

The coefficients $\alpha_\rho$, $\rho = 0, 1, \cdots, r$ of the least squares polynomial $\pi(\tau)$ are obtained in the well known manner from the system of conditional equations $C\alpha = v$, by solving the associated normal equations $N\alpha = w$, with $N = C^TC$, $w = C^Tv$, where the superscript $T$ denotes transposition. Thus, if we put

$$\alpha_0 = a_1 v, \qquad \alpha_1 = a_2 v, \tag{4}$$

it is seen that, for example, for $r = 3, n = 2$,

$$a_1 = (1/35)(-3, 12, 17, 12, -3),$$

$$a_2 = (1/12)(+1, -8, 0, +8, -1).$$

Let us now consider the following special smoothing problem. Suppose there are given two sequences $x'_i$, $y'_i$, of observed values, and a function $f'_i \equiv f(x'_i, y'_i)$. In applying smoothing procedures to the calculation of $f_i$ a number of possibilities present themselves:

I. *Function smoothing*. Smooth the functional values $f'_i$ only.

II. *Argument smoothimg*. Smooth the arguments $x'_i$, $y'_i$, into $x^*_i$, $y^*_i$ then calculate

$$f^*_i \equiv f(x^*_i, y^*_i).$$

III. *Argument and function smoothing*. Smooth the $x'_i$, $y'_i$, and subsequently calculate and smooth also the $f^*_i$.

Since the amount of computational labor involved in each case may vary greatly, appropriate indications as to which of these possibilities may be the most economical would be of considerable value. While the experiments described below do not provide a general answer to this question, they do permit certain interesting conclusions.

**3. Outline of procedure.** In order to investigate this problem we conducted the following experiments. First exact mathematical functions $x(t)$, $y(t)$, $f(t) \equiv f(x(t), y(t))$ were chosen, and their values, as well as those of $X(t) \equiv dx/dt$, $Y(t)$, $F(t)$ were computed for a sequence $t_i$ of equispaced values of $t$.

Next the $x_i$ and $y_i$ were subjected to certain well defined, preferably pseudo-random perturbations which produced corresponding sequences $x'$, $y'$, $f' = f(x', y')$.

In using Method I, Eq. (4) was applied to $v = f'$, so that the $\alpha_0$ produced smoothed quantities $(f')^* \equiv g$ of the function. On the other hand, in using Method II, Eq. (4a) was applied in succession to $v = x'$, $v = y'$, resulting in smoothed arguments $x^*$, $y^*$. From these there was subsequently calculated $f^* = f(x^*, y^*)$. In applying III, finally, the smoothing process was imposed also on $v = f^*$, now leading to "doubly" smoothed values $(f^*)^* \equiv h$ of the function.

In computing derivatives by means of $\alpha_1$ of Eq. (4) the smoothing of $f'$ led to the set $\varphi$ of derivatives of $f'$, while the smoothing of $f^*$ resulted in quantities $\psi$ of derivatives of $f^*$. A comparison of items $\varphi$, $\psi$ with $F$ then indicated the relative accuracies obtained.

The computation of rates of change of $f$ may obviously be accomplished also in a number of different ways. Thus, for example, quantities $X^*$, $Y^*$ may be obtained first, by the application of the process $\alpha_1$ to $x^*$, $y^*$, and then a derivative of $f$ calculated by means of

$$F^* = \frac{\partial f}{\partial x}(x^*, y^*)X^* + \frac{\partial f}{\partial y}(x^*, y^*)Y^*.$$

**4. Examples.** *A. Example* 1. We took

$$x(t) = \sin(t/10), \qquad y(t) = (1/10)\exp(t/10)$$
$$f(x, y) = x + y,$$

for $t = 0, 1, \cdots, 30$.

Now $x$ and $y$ were perturbed as follows. Denote the third digit of $x$ by $d_3$. Then we put

$$t' = (t/10) + \epsilon(-1)^{d_3}d_3, \qquad \epsilon = .001,$$
$$x' = \sin t', \qquad y' = (1/10)\exp t'.$$
$$f' = x' + y'.$$

The calculations showed that, for $2 \leq t \leq 28$,

$$\| e(t') \|^2 = 915.10^{-6}$$
$$\| e(x') \|^2 = 450.10^{-6}, \qquad \| e(y') \|^2 = 520.10^{-6}$$
$$\| e(f') \|^2 = 480.10^{-6}.$$

Smoothing of $t'$ by Eq. (4) resulted in $\| e(t^*) \|^2 = 395.10^{-6}$, so that the smoothing effect on $t'$ itself was very pronounced indeed.

Smoothing of $x', y', f'$ into sequences $x^*, y^*, f^* = x^* + y^*$ had the following effect:

$$\| e(x^*) \|^2 = 170.10^{-6}, \qquad \| e(y^*) \|^2 = 130.10^{-6}, \qquad \| e(f^*) \|^2 = 240.10^{-6}.$$

Since $f = x + y$ is a linear function, $g = f^*$. Further smoothing of $f^*$ led to $h$, and it was found that $\| e(h) \|^2 = 160.10^{-6}$, to be compared with a value of $\| e(f^*) \|^2 = 229.10^{-6}$ for the corresponding entries. For the derivatives we found

$$\| e(\varphi) \|^2 = 326.10^{-6}, \qquad \| e(\psi) \|^2 = 154.10^{-6}.$$

Due to the linearity of $f$, $F^* = \psi$.

*B. Example* 2. Next, we let

$$x(t) = (1/2)(t/10)^2, \qquad y(t) = (1/10)(t/10)^3 + (t/10)$$
$$f(t) = y - x,$$

for $t = 0, 1, \cdots, 30$. For the disturbances $\xi, \eta$ of $x, y$, respectively, we chose the following functions:

$$x_i' = x_i + \xi_i, \qquad y_i' = y_i + \eta_i,$$
$$\xi_i = \epsilon(-1)^{\sigma_{7i}}\sigma_{7i}, \qquad \eta_i = (-1)^{\tau_{7i}} \cdot \tau_{7i},$$

with $\epsilon = .001$, and $\sigma_{7i}, \tau_{7i}$ denoting, respectively, the seventh decimal of $\sin[(60 + i)\pi/180]$, $\tan[(50 + i)\pi/180]$, for $i = 0, 1, 2, \cdots, 30$.

It was seen that

$$\| e(x') \|^2 = 715.10^{-6}, \qquad \| e(y') \|^2 = 708.10^{-6}, \qquad \| e(f') \|^2 = 1862.10^{-6}$$

between $i = 2$ and $i = 28$.

Smoothing by Eq. (4), leading to $x^*$, $y^*$ and $f^*$ had this effect:

$$\| e(x^*) \|^2 = 358.10^{-6}, \qquad \| e(y^*) \|^2 = 468.10^{-6}, \qquad \| e(f^*) \|^2 = 1428.10^{-6}.$$

For $4 \leq i \leq 26$, $\| e(f^*) \|^2 = 938.10^{-6}$, $\| e(h) \|^2 = 864.10^{-6}$. As to derivatives, the computations led to

$$\| e(\psi) \|^2 = 437.10^{-6}, \qquad \| e(\varphi) \|^2 = 1351.10^{-6}.$$

*C. Example* 3. Next we considered the following example:

$$x(t) = (t - 1)/(t + 1), \qquad y(t) = 2(t + 1)/(2 + 1), \qquad f(x, y) = x \cdot y = 2(t - 1)/2t + 1),$$

to be computed for $t = 3, 4, \cdots, 30$. The "roughing up" of $x$, $y$ was done in the following manner: Let $\sigma_4$, $\tau_4$ denote the fourth decimal of $x$, $y$, respectively. Then we put

$$x' = \begin{cases} x + \epsilon(-1)^{\sigma_4} & \text{for} \quad \sigma_4 \neq 0 \\ x & \text{for} \quad \sigma_4 = 0, \end{cases}$$

with $\epsilon = .0005$. Similarly for $y'$. It turned out that for $x'$, $y'$, $f' = x'y'$, between $i = 5$ and $i = 28$,

$$\| e(x') \|^2 = 375.10^{-8}, \qquad \| e(y') \|^2 = 525.10^{-8}, \qquad \| e(f') \|^2 = 714.10^{-8}.$$

Smoothing of $x'$ and $y'$ led to $x^*$, $y^*$, and $f^*$. It was seen that

$$\| e(x^*) \|^2 = 284.10^{-8}, \qquad \| e(y^*) \|^2 = 420.10^{-8},$$

so that the smoothing did diminish somewhat the length of the error vectors. Further, $\| e(f^*) \|^2 = 526.10^{-8}$.

The values of $g$, obtained by smoothing the functional values $f'$ showed very close agreement with those of $f^*$. Further smoothing of $f^*$ resulted in values $h$, for which $\| e(h) \|^2 = 268.10^{-8}$, which was to be compared with $305.10^{-8}$ for the corresponding entries in $f^*$. The derivatives of $f$, computed in the three different ways described in Sec. 3, led to

$$\| e(\psi) \|^2 = 236.10^{-8}, \qquad \| e(\varphi) \|^2 = 540.10^{-8}, \qquad \| e(F^*) \|^2 = 248.10^{-8}.$$

*D. Example* 4. Finally, we took

$$x(t) = (1/10)((t/10) - 3)^2((t/10) + 1)^3$$
$$y(t) = -(1/10)((t/10) - 3)^3((t/10) + 1)$$
$$f(t) = x(t)/y(t),$$

for $t = -15, -14, \cdots, +15$. A perturbation was introduced into $x$, $y$ in the following "sinusoidal" manner:

$$x_i' = x_i + \xi_i, \qquad y_i' = y_i + \eta_i,$$

$$\xi_i = \epsilon \sin(\pi i/3)/\sin(\pi/3), \qquad \eta_i = \epsilon \sin[(\pi/3)(i + 1)]/\sin(\pi/3),$$

for $i = 0, \pm 1, \pm 2, \cdots$, and $\epsilon = .0005$. This resulted in the perturbations

$$\xi_i = 0, \epsilon, \epsilon, 0, -\epsilon, -\epsilon, 0, \cdots, \eta_i = \epsilon, \epsilon, 0, -\epsilon, -\epsilon, 0, \cdots$$

for $i = 0, 1, 2, \cdots$. We found that, for $-13, \leq i \leq 13$,

$$\| e(x') \|^2 = 450.10^{-8}, \qquad \| e(y') \|^2 = 450.10^{-8}, \qquad \| e(f') \|^2 = 386.10^{-8}.$$

In roughing up $x'$, $y'$, $f'$ at $i = -10$, as well as in smoothing these quantities, the values had been left unchanged at zero, in order to avoid absurd entries there.

The smoothing again did diminish the error vectors somewhat:

$$\| e(x^*) \|^2 = 383.10^{-8}, \qquad \| e(y^*) \|^2 = 379.10^{-8}, \qquad \| e(f^*) \|^2 = 341.10^{-8}.$$

On the other hand, $\| e(g) \|^2 = 370.10^{-8}$. Further, $\| e(h) \|^2 = 99.10^{-8}$, while $\| e(f^*) \|^2 = 128.10^{-8}$ for the corresponding entries between $i = -11$ and $i = 11$. For the derivatives the following results were obtained:

$$\| e(\psi) \|^2 = 118.10^{-8}, \qquad \| e(\varphi) \|^2 = 161.10^{-8}, \qquad \| e(F^*) \|^2 = 132.10^{-8}.$$

**5. Conclusion.** In view of the results obtained above the following conclusions, at least for the examples discussed, seem to be justified:

(a) $f^*$ is not significantly better than $g$;

(b) the improvement $h$ due to the further smoothing of $f^*$ may be worthwhile;

(c) $F^*$ is about as good as $\psi$, while $\varphi$ is considerably worse than either of the foregoing.

In summary, then, it may be stated that if functional values $f$ alone are desired, possibility I of Sec. 2 would seem to be adequate. However, if rates of change $df/dt$ also are of interest, then possibility II should be utilized.

## Reference

1. See, for example, Whittaker-Robinson, *The calculus of observations*, 4th ed., 1944, p. 291

# TRANSVERSE VELOCITIES IN FULLY DEVELOPED FLOWS*

By STEPHEN H. MASLEN

By STEPHEN H. MASLEN (*Lewis Flight Propulsion Laboratory, Cleveland, Ohio*)

We consider an incompressible flow in a channel whose generators are parallel and lie in the $x$-direction. The driving force, either a pressure gradient or some kind of a body force, is also in the $x$-direction. If the driving force and the flow velocities are independent of $x$, then the flow is said to be fully developed. In such a case, it is usually assumed that the velocity components normal to the $x$-direction vanish identically. However, in some cases it is not clear that this must be true. One such example is that of a free convective flow where the velocity ($x$-direction) shows a cellular structure [1], being sometimes in one direction (up, say) and sometimes in the other (down). It is the purpose of the present note to show that regardless of these circumstances the transverse velocity components in a fully developed incompressible flow must vanish everywhere.

Consider a fully developed incompressible flow, where the velocity components $u$, $v$, $w$, are in the (cartesian) $x$, $y$, $z$ directions. If $P$, $\rho$, $\nu$, and $F$ are pressure, density,

*Received March 15, 1957.