

## Derivatives in Spectroscopy

### Part III — Computing the Derivative

Howard Mark and Jerome Workman Jr.

Our previous two columns in this series (1, 2) discussed the theoretical aspects of using derivatives in the analysis of spectroscopic data. Here we consider some of the practical aspects. First, in the presence of some arbitrary but (presumably) constant amount of noise, what is the optimum spacing of data at which to compute a difference to give the highest signal-to-noise ratio (S/N)? In the face of constant noise, this obviously reduces to the question: what is the spacing (for a normal distribution) that gives the largest value for the numerator term? Note that the criterion for “best” has changed from our previous discussions, where “best” was considered to be the closest approximation to the true derivative.

We have noted that the largest value of the true first derivative occurs when  $X - \mu = \sigma$ . Therefore the largest differences between two points will occur when they are varied from  $\mu + \sigma$  (or  $\mu$

$-\sigma$ ) by some amount  $\delta$ , the spacing, which we need to determine. Therefore we need to determine the largest difference of

$$e^{\left(\frac{\mu + \sigma + \delta/2}{\sigma}\right)^2} - e^{\left(\frac{\mu + \sigma - \delta/2}{\sigma}\right)^2} \quad [21]$$

The first question we need to ask is whether there is, in fact, a maximum value. That there is one can be seen from noting that the normal absorbance band approaches zero as  $X$  approaches infinity in both directions. Therefore if

$$\delta \longrightarrow \infty$$

the difference will approach zero. At small values of  $\delta$  the difference will be finite, while as

$$\delta \longrightarrow 0$$

the difference will again approach zero; therefore, there must be a maximum somewhere between 0 and  $\infty$ . To get some idea of where that maximum is, in Figure 9 we show a plot of the difference as a function of  $\delta$  for the normal absorbance band of 20-nm bandwidth we've shown in Figure 1. For a more precise result we must solve equation 21, but because it is transcendental, we must solve it by successive approximations. The result of doing so is  $\delta_{\max} = 34.28$  nm. Because the bandwidth of the underlying absorbance band is 20 nm, the spacing needed for maximizing the first derivative S/N for any normal absorbance band is therefore  $34.28/20 = 1.714$  times the bandwidth. However, this analysis is based on considering a single peak in isolation; as we will see for the second derivative, at some point it becomes necessary to take into account the presence and nature of whatever other materials exist in the sample.

The second derivative is both simpler and more complicated to deal with. As we saw, the second derivative is maximum at the wavelength of the peak of the underlying absorbance curve, and we noted previously that the numerator term at that point increases monotonically with the spacing (see Figures 4 and 5 in reference 1). Therefore we expect the signal-to-noise ratio of the second derivative to improve continually as the spacing becomes larger and larger.

#### Howard Mark

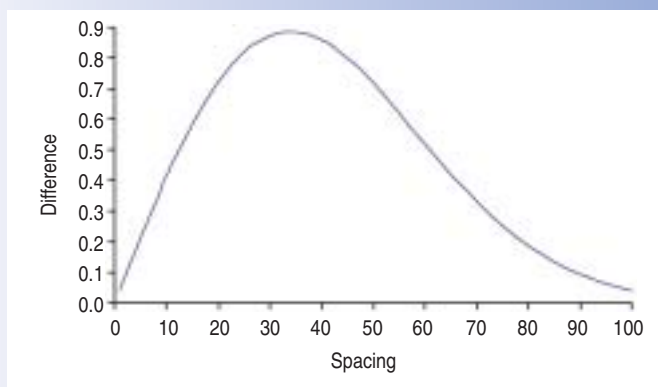
serves on the Editorial Advisory Board of *Spectroscopy* and runs a consulting service, Mark Electronics (69 Jamie Court, Suffern, NY 10901). He provides assistance, training, and consultation in NIR spectroscopy, chemometrics, and statistical data analysis, as well as custom software and hardware development. He can be reached via e-mail at [hmark@prodigy.net](mailto:hmark@prodigy.net).



#### Jerome Workman Jr.

serves on the Editorial Advisory Board of *Spectroscopy* and is chief technical officer and vice president of research and engineering for Argose (Waltham, MA). He may be reached by e-mail at [jworkman@argose.com](mailto:jworkman@argose.com).





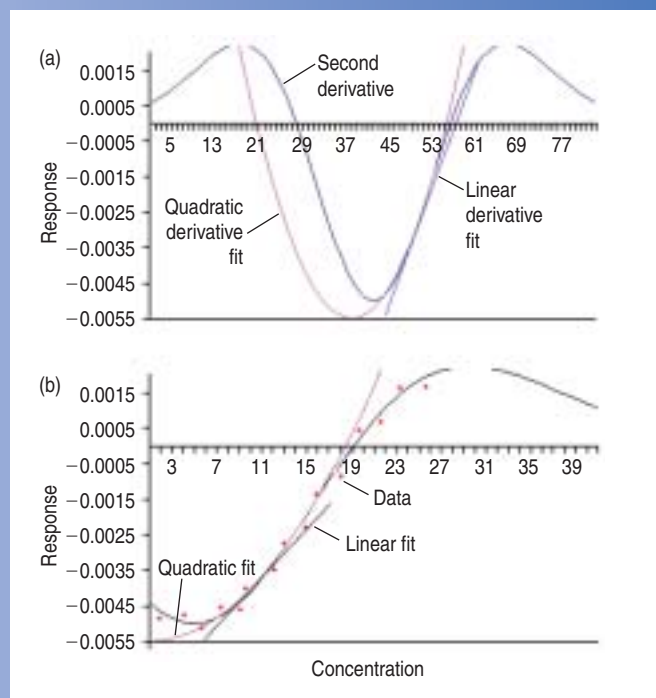
**Figure 9.** The difference between the ordinates of two points equally spaced around  $\mu + \sigma$  as a function of the spacing. In this figure, the underlying absorbance curve has a bandwidth of 20 nm.

Although the signal part of the second derivative increases with the spacing used, the noise of the computed second derivative is independent of the spacing. It is, however, larger than the noise of the underlying spectrum. As we have shown (3), from elementary statistical considerations multiplying a random variable  $X$  by a constant  $A$  causes the variance of the product  $AX$  to be multiplied by  $A^2$  compared to the variance of  $X$  itself. Now, regardless of the spacing of the terms used to compute the second derivative, the operative multipliers for the data at the three wavelengths used are 1,  $-2$ , 1. Therefore the multiplier for the variance of the derivative is  $1^2 + 2^2 + 1^2 = 6$ , and the standard deviation of the derivative is therefore  $6^{1/2}$  times the standard deviation of the spectrum, but nevertheless independent of the derivative spacing. The signal-to-noise ratio of the second derivative is therefore determined solely by the magnitude of the computed numerator value, which, as we have seen, increases with spacing.

In real samples, however, the wider the spacing, the more likely it becomes that one of the points used for the derivative computation will be affected by the presence of other constituents in the sample, and the question of the optimum spacing for the derivative computation becomes dependent on the nature of the sample in which it is contained.

### Methods of Computing the Derivative

The method we have used until now for estimating the derivative, simply calculating the difference between absorbance values of two data points spaced some distance apart (and dividing that by  $\Delta X$ , of course), is probably the simplest method available. As we discussed in our previous column (1), however, there is a disadvantage associated with this method. This method causes a decrease in the signal-to-noise ratio as compared with the underlying absorbance band, and this decrease has two sources. The lesser source is the increase in the noise level due to the addition of variances that occurs when numbers are added or subtracted. The far larger effect is due to the fact that the derivatives are much smaller than the absorbance, and the second derivative



**Figure 10.** The Savitzky-Golay method of computing derivatives is based on a least-squares fit of a polynomial to the data of interest. In both parts of this figure, the underlying second derivative curve is shown as the black line, while the linear (first degree) and quadratic (second degree) polynomials are shown as mauve and blue lines, respectively. (a) Linear and quadratic fits to a normal spectral curve. (b) An expansion of (a) that shows how the polynomials are determined using a least-squares fit to the actual data in the region where the derivative is computed when the data are contaminated with noise. Red dots represent the actual data.

is much smaller (by an order of magnitude) than the first. The net result is that, the closer the theoretical approximation to the true derivative is, the noisier the actual computed derivative becomes.

Several methods have been devised to circumvent this characteristic of the process of taking derivatives. One of the very common methods is to reduce the initial noise of the spectrum by computing averages: averaging the spectral data over some number of wavelengths before estimating the derivative by calculating the difference between the resulting averages. This process is sometimes called “smoothing” because it smooths out the noise of the spectrum. However, because we are not discussing smoothing, we will not consider this any further here.

The next common method of computing derivatives is the use of Savitzky-Golay convolution functions. The application to spectroscopy is based on what is one of the most often cited papers in the literature (4).

This paper is a classic, but turns out to be not as original as is sometimes thought based on that reputation. In reality it is a (relatively) recent addition to a long line of development. Savitzky and Golay themselves point out: “The bases

for the methods to be discussed have been reported previously, mostly in the mathematical literature” and then give a list of references going back to 1941.

However, the mathematical community has known about that type of analysis for much longer. We were put on the track of these earlier developments by a letter to the editor of *C&E News* (5). This letter to the editor pointed us to articles going back to at least 1913 (6, 7). Those papers clearly show that fitting polynomials to data using least squares was already a well-known technique, and that the idea of computing coefficients to describe the fitting function was well under development by then. The 1913 paper refers to an earlier paper by the same author that performed the same functions using a slightly different approach, and it is clear that the underlying concepts had already been known in the mathematical community for a long time. The main limitation those authors encountered was the lack of an easy way to actually perform the required calculations, because even (what we would now consider) primitive computers were at least 30 years in their future.

So the theoretical foundations for Savitzky and Golay's work were well-laid when they did it. However, they did make several contributions that resulted in the current widespread acceptance of the methodology and the now-classic reputation that their paper now enjoys. First of all, they brought it out of the arcane world of mathematicians and into the real world of analytical chemists. Secondly, they did it at a time when scientists were just starting to gain access to computers (albeit mainframe computers, back then) to implement the algorithm locally, and apply it to their own data. Thirdly (and synergistically with point number two) they provided computer code (in FORTRAN, the leading scientific computer language of that time) to assist other scientists in implementing the concepts. So while Savitzky and Golay did not make a fundamental breakthrough, what they did was arguably even more important: they brought the concept to the attention of the world at just the time when the world was ready for it, and in just the way that allowed it to become the widespread method of data fitting that it still is today.

This now-classic paper presents the concept underlying this method for computing derivatives (including the zero-order derivative, which reduces to what is basically a weighted smoothing operation); Figure 10a shows this diagrammatically. The assumption is that the mathematical nature of the underlying spectral curve is unknown, but can be represented over some finite region by a polynomial; *polynomial* in this sense in general and includes straight lines. If the equation for the polynomial is known, then the derivative of the spectrum can be calculated from the properties of the fitted polynomial.

The key to all this is the fact that the nature of the polynomial can be calculated from the spectral data by doing a least-squares fit of the polynomial to the data in the region of interest, as shown in Figure 10b. Figure 10a shows that various polynomials may be used to approximate the derivative curve at the point of interest, and Figure 10b shows that when the derivative curve is based on data that has error, the

polynomials can be computed using a least-squares fit to the data. At the point for which the derivative is computed, all three lines in Figure 10 are tangent to each other. The Savitzky–Golay approach provides for the use of varying numbers of data points to be used in the computation of the fitting polynomial. We will discuss the effect of changing the number of data points shortly.

So the steps that Savitzky and Golay took to create their classic paper were:

1. Fit a curve (polynomial) of the desired type and degree to the data.
2. Compute the desired order of derivative of that polynomial.
3. Evaluate the expression for the derivative of that polynomial at the point for which the derivative is to be computed. In the Savitzky–Golay paper, this is the central point of the set used to fit the data. As we shall see, in general this need not be the case, although doing so simplifies the formulas and computations.
4. Convert those formulas into a set of coefficients that can be used to multiply the data spectrum by, to produce the value of the derivative according to the specified polynomial fit, at the point of the center of the set of data. As we shall see, however, their paper ignores some key points.

And finally, although this work was all of very important theoretical interest, Savitzky and Golay took one more step that turned the theory into a form that could be easily put to practical use:

5. For a good number of sets of derivative orders, fitting polynomials and numbers of data points, they calculated and printed in their paper tables of the coefficients needed for the cases considered. Thus the practicing chemist needed to be neither a heavy-duty theoretician nor more than a minimal computer programmer to make use of the results produced.

Unfortunately there are also several caveats that have to go along with the use of the Savitzky–Golay results. The most important and also the best-known caveat is that there are errors in the tables in their paper. This was pointed out by Steinier (8) in a paper that is invariably cited along with the original Savitzky–Golay paper, and which should be considered a “must read” along with the original paper by anyone taking an interest in the Savitzky–Golay approach to computation of derivatives.

The Savitzky–Golay coefficients provide a simplified form of computation for the derivative of the desired order at a single point. To produce a derivative spectrum the coefficients must be applied successively to sets of spectral data, each set offset from the previous one by a single wavelength increment. This is known as the convolution of the two functions. Having done that, the result of all the theoretical development and computation is that the derivative spectrum so produced simultaneously is based on a smoothed version



of the spectrum. The amount of smoothing depends on the number of data points used to compute the least-squares fit of the polynomial to the data; use of more data points is equivalent to performing more smoothing. Using higher-degree polynomials as the fitting function, on the other hand, is equivalent to using less smoothing, because high-order polynomials can twist and turn more to follow the details of the data.

### Limitations of the Savitzky–Golay Method

The publication of the Savitzky–Golay paper (augmented by the Steinier paper) was a major breakthrough in data analysis of chemical and spectroscopic data. Nevertheless it does have some limitations, and some more caveats that need to be considered when using this approach.

One limitation is that the method as originally described is applicable only to computations using odd numbers of data points. This was implied earlier when we discussed the fact that a derivative (of any order) is computed at the central point (wavelength) of the set used.

Another limitation is that, also because of the computation being applicable to the central data point, there is an end effect to using the Savitzky–Golay approach: it does not provide for the computation of derivatives that are too close to the end of the spectrum. The reason is that at the end of the spectrum, there is no spectral data to match up to the coefficients on one side or the other of the central point of the set of coefficients; therefore, the computation at or near the ends of the spectrum cannot be performed.

Of course, an inherent limitation is the fact that only those combinations of parameters (derivative order, polynomial degree, and number of data points) that are listed in the Savitzky–Golay/Steinier tables are available for use. Although those cover what are likely to be the most common needs, anyone wanting to use a set of parameters beyond those supplied is out of luck.

A caveat to the use of the Savitzky–Golay tables is that, even after Steinier's corrections, they apply only to a special case of data, and do not, in general, produce the correct value of the true derivative. The reason for this is similar to the problem we pointed out in our first column dealing with computation of derivatives (1): applying the Savitzky–Golay coefficients to a set of spectral data is equivalent to assuming that the data is separated by unit  $X$  distance, and therefore is equivalent to computing only the numerator term of a finite difference computation, without taking into account the  $\Delta X$  (spacing) to which the computed  $\Delta Y$  corresponds. Therefore, to compute the Savitzky–Golay estimate of a true derivative, the value computed using the Savitzky–Golay coefficients must be divided by  $(\Delta X)^n$ , where  $n$  is the order of the derivative.

Another limitation is perhaps not so much a limitation as, perhaps, a strange characteristic, albeit one that can catch the unwary. To demonstrate, we consider the simplest Savitzky–Golay derivative function, that for the first derivative using a five-point quadratic fitting function. The convolution coefficients (after including the normalization factor) are  $-0.2, -0.1, 0, 0.1, \text{ and } 0.2$ .

Suppose we compute a second derivative by applying this first derivative function twice. The effect is easily shown to be equivalent to applying the convolution coefficients:

$$0.04, 0.04, 0.01, -0.04, -0.1, -0.04, 0.01, 0.04, 0.04.$$

This is a collection of nine coefficients that produces a second derivative, based on the Savitzky–Golay first-derivative coefficients.

However, this collection of convolution coefficients appears nowhere in the Savitzky–Golay tables. The nine-point Savitzky–Golay second derivative with a quadratic or cubic polynomial fit has the following coefficients:

$$0.0606, 0.0152, -0.0173, -0.0368, -0.0433, -0.0368, -0.0173, 0.0152, 0.0606.$$

And the nine-point Savitzky–Golay second derivative with a quartic or quintic polynomial fit has the following coefficients:

$$-0.8811, 2.5944, 1.0559, -1.4755, -2.5874, -1.4755, 1.0559, 2.5944, -0.8811.$$

The original Savitzky–Golay paper (4) describes how to compute other Savitzky–Golay convolution coefficients from given ones; these other coefficients are also functions that follow the basic concepts of the Savitzky–Golay procedure: the derivative of a least-squares, best-fitting polynomial function. Because they do not produce the convolution coefficients we generated by applying the Savitzky–Golay first derivative coefficients twice, we are forced to the conclusion that even though the coefficients for the first derivative follow the Savitzky–Golay concepts, applying them two (or multiple) times in succession does not produce a set of convolution coefficients that is part of the Savitzky–Golay collection of convolution functions. This seems to be generally true for the Savitzky–Golay convolution coefficients as a whole.

### Extensions to the Savitzky–Golay Method

Several extensions have been developed to the original concept. First we'll consider those that don't change the fundamental structure of the Savitzky–Golay approach, but simply make it easier to use. The main development along this line is the elimination of the tables. On the one hand, tables of coefficients are easy to deal with conceptually because they can be applied mechanically — just copy down the entries and use them to multiply the data by. In fact, our initial foray into the world of Savitzky–Golay involved writing just such a program. The task was tedious, but having done it once and having verified the numbers it should never be necessary to do it again. However, as noted above this approach has the inherent limitation of including only those conditions that are listed in the Savitzky–Golay tables, extensions to the derivative order, polynomial degree, or number of data points used are excluded.

An extension of this idea was presented in a paper by Madden (9). Instead of presenting the already-worked-out numbers, Madden derived formulas from which the coeffi-

cients could be computed and presented a table of those formulas in this paper. This is definitely a step up because it confers several advantages:

1. Through the use of these formulas, Savitzky–Golay convolution coefficients could be computed for a convolution function using any odd number of data points for the convolution.
2. Because the coefficients are being computed by the computer, there is no chance for typographical errors occurring in the coefficients.

Madden's paper, however, also has limitations:

1. The paper contains formulas for only those derivative orders and degrees of polynomials that are contained in the original Savitzky–Golay paper; therefore, we are still limited to those derivative order and polynomial degrees.
2. The coefficients produced still contain the implicit assumption that the value of  $\Delta X = 1$ . Therefore, to produce correct derivatives, it will still be necessary to divide the results from the formulas by  $(\Delta X)^n$ , as above.
3. The formulas are at least as complicated, difficult, and tedious to enter as the tables they replace, and as fraught with the possibility of typos during their entry. This is exacerbated by the fact that, being a formula in a computer program, everything must be just so, and all the parentheses and so forth must be placed correctly, which, for formulas as complicated as those are, is not easy to do. Nevertheless, as with the tables, once it is done correctly it need not be done again (but make sure you back up your work!). However, for the real kick in the pants, see the next item on this list.
4. There is an error in one of the formulas! While writing the program to implement the formulas in Madden's paper, despite the tedium, most of the formulas (10 of the 11 given) in the program were working correctly in fairly short order — "correctly" in this case meaning that the coefficients agree with those of Savitzky–Golay or of Steinier, as appropriate. There was a problem with one of the formulas, however: the one for the third derivative using a quintic (fifth degree) polynomial fitting function. The coefficients produced were completely unreasonable, as well as being wrong. The coding of the formula was checked a couple of ways. First that formula was rewritten, starting from scratch and using a different scheme to convert the printed formula to computer code, and the same wrong answers were obtained both times. Then our buddy Dave Hopkins checked the coding; he reported not finding any discrepancies between the printed formula and what was coded. This left two possibilities: either the printed formula was wrong, or the corresponding Steinier table was wrong.

We first tried to contact Hannibal Madden because the paper gave his affiliation as Sandia National Laboratory, but he was no longer there and the human resources department had no information as to his current whereabouts. Finally the problem was posted to an on-line discussion group (the

discussion group for the International Chemometrics Society), asking if anybody had information relating to this problem. Fortunately, Premek Lubal (one of the members of the group) had run into this problem previously, while checking the derivations in Madden's paper and knew the solution (10). To save grief on the part of anybody who might want to code these formulas for themselves, here is the solution: in the formula for the case involved, the quintic fitting function for the third derivative, the term  $(50 * m)$  has the wrong sign. The sign in the printed formula is negative ( $-$ ), and it should be positive ( $+$ ). After changing the sign of that term, the program produced the correct coefficients.

So now the question presents itself: is there a more general method of computing coefficients for any arbitrary set of combinations of derivative order, polynomial degree and number of data points to fit? That is, is there an automated method for computing Madden's formulas, or at least the Savitzky–Golay convolution coefficients? The answer turns out to be yes. In the same on-line discussion that produced the solution to the problem in the Madden paper, Chris Brown pointed out some pertinent literature citations (11, 12) and summarized them in the general solution we discuss below (13). Is the solution as simple as the tables in Savitzky–Golay/Steinier or the formulas in Madden? This is a matter of perception. If this general solution was presented to the chemical/spectroscopic community in 1964 (at the time of the original Savitzky–Golay paper) it would have been considered far beyond what most chemists would be expected to know, and would never have gained the popularity it currently enjoys.

With the advent of modern software tools, however — tools such as MATLAB and even the older language, APL — matrix operations can be coded directly from the matrix-math expressions, and then it becomes near-trivial to create and solve the matrix equations on-the-fly, so to speak, and calculate the coefficients for any derivative using any desired polynomial, and computed over any odd number of data points.

Wentzell et al. (12) presented this scheme in a very clear way, the same way that Chris Brown gave it to me: We start by creating a matrix. This matrix is based on the index of coefficients that are to be ultimately produced. Savitzky and Golay labeled the coefficients in relation to the central data point of the convolution, therefore the coefficients a three-term set of coefficients are labeled  $-1, 0, 1$ . A five-term set is labeled  $-2, -1, 0, 1, 2$ , and so forth.

The matrix (**M**) is set up like this table (this, of course, is only one example, for expository purposes):

$$\mathbf{M} = \begin{matrix} & 1 & -3 & 9 & -27 \\ & 1 & -2 & 4 & -8 \\ & 1 & -1 & 1 & -1 \\ \mathbf{M} = & 1 & 0 & 0 & 0 \\ & 1 & 1 & 1 & 1 \\ & 1 & 2 & 4 & 8 \\ & 1 & 3 & 9 & 27 \end{matrix} \quad [22]$$

What are the key characteristics that we need to know about this matrix? The first one is that each column contains the set of index numbers raised to the  $n - 1$  power, where  $n$  is the column number in the table. Thus the first column contains the zero power, which is all ones; the second column contains the first power, which is the set of index numbers themselves; and the rest of the columns are the second and third powers of the index numbers.

What determines the number of rows and columns? The number of rows is determined by the number of coefficients that are to be calculated. In this example, therefore, we will compute a set (sets, actually, as we will see) of seven coefficients. The number of columns is determined by the degree of the polynomial that will be used as the fitting function. The number of columns also determines the maximum order of derivative that can be computed. In our example we will use a third-power fitting function and we can produce up to a third derivative. As we shall see, coefficients for lower-order derivatives are also computed simultaneously.

The matrix  $\mathbf{M}$  is then used as the argument for the following matrix equation:

$$\text{Coefficients} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \quad [23]$$

where, by convention, the boldface  $\mathbf{M}$  refers to the matrix we produced, the superscript T refers to the transpose of the matrix, and the superscript  $-1$  means the matrix inverse of the argument.

Let us evaluate this expression. The matrix  $\mathbf{M}$  is given earlier as equation 22. The transpose, then, is:

$$\mathbf{M}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 9 & 4 & 1 & 0 & 1 & 4 & 9 \\ -27 & -8 & -1 & 0 & 1 & 8 & 27 \end{bmatrix} \quad [24]$$

We then need to multiply these two matrices together to form  $\mathbf{M}^T \mathbf{M}$  (rules for matrix multiplication are given in many books, including reference 14):

$$\mathbf{M}^T \mathbf{M} = \begin{bmatrix} 7 & 0 & 28 & 0 \\ 0 & 28 & 0 & 196 \\ 28 & 0 & 196 & 0 \\ 0 & 196 & 0 & 1588 \end{bmatrix} \quad [25]$$

Then we compute the matrix inverse of equation 25 (in MATLAB, this is just: `inv [m]`):

$$(\mathbf{M}^T \mathbf{M})^{-1} = \begin{bmatrix} 0.333333 & 0 & -0.0476190 & 0 \\ 0 & -0.262566 & 0 & -0.0324074 \\ -0.0476190 & 0 & 0.01190476 & 0 \\ 0 & -0.0324074 & 0 & 0.00462962 \end{bmatrix} \quad [26]$$

Finally, multiplying equation 26 by equation 24 gives:

$$(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T = \begin{bmatrix} -0.09523 & 0.142857 & 0.285714 & 0.333333 & 0.285714 & 0.142857 & -0.09523 \\ 0.087301 & 0.2658730 & -0.230158 & 0 & 0.230158 & 0.2658730 & -0.087301 \\ 0.0595238 & 0 & 0.0357142 & 0.0476190 & 0.0357142 & 0 & 0.0595238 \\ 0.0277777 & 0.0277777 & 0.0277777 & 0 & 0.0277777 & 0.0277777 & 0.0277777 \end{bmatrix} \quad [27]$$

Equation 27 contains scaled coefficients for the zero through third derivative convolution functions, using a third degree polynomial fitting function. The first row of equation 27 contains the coefficients for smoothing, the second row contains the coefficients for the first derivative, and so forth.

Equation 27 gives the coefficients, but a scaling factor is missing. Therefore one more final computation must be performed to create the correct coefficients; each row must be multiplied by the scaling factor. The scaling factor is  $(p - 1)!$  where  $p$  is the row number. Therefore the scaling factors for the first two rows are unity, because  $0!$  and  $1!$  are both unity, the scaling factor for the third row is two and for the fourth row is six. The final set of coefficients, therefore, is:

$$(\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \text{ (corrected for scaling)} = \begin{bmatrix} -0.09523 & 0.142857 & 0.285714 & 0.333333 & 0.285714 & 0.142857 & -0.09523 \\ 0.087301 & -0.2658730 & -0.230158 & 0 & 0.230158 & 0.2658301 & -0.087301 \\ 0.119047 & 0 & 0.0714285 & -0.095238 & 0.1714285 & 0 & 0.119076 \\ -0.166666 & 0.166666 & 0.166666 & 0 & -0.166666 & -0.166666 & 0.166666 \end{bmatrix} \quad [28]$$

Finally, for those who are facile with the matrix math, Bialkowski (11) also shows how the "end effect" can be obviated, as well as allowing the use of even numbers of data points, but the advanced considerations involved are beyond the scope of our column. ■

## References

1. H. Mark and J. Workman, *Spectroscopy* 18(4), 32–37 (2003).
2. H. Mark and J. Workman, *Spectroscopy* 18(9), 25–28 (2003).
3. H. Mark and J. Workman, *Spectroscopy* 3(8), 13–15 (1988).
4. A. Savitzky and M. Golay, *Anal. Chem.* 36(8), 1627–1639 (1964).
5. R. deLevie, letter to the editor, *C&E News*, September 8, 2003, p.10.
6. C.W.M. Sherriff, *Proc. Roy. Soc. Edinburgh* 40, 112–128 (1920).
7. W.F. Sheppard, *Proc. London Math. Soc.* 13(1196), 97–108 (1913).
8. J. Steinier, Y. Termonia, and J. Deltour, *Anal. Chem.* 44(11), 1906–1909 (1971).
9. H.H. Madden, *Anal. Chem.* 50(9), 383–1386 (1978).
10. P. Lupal, private communication (2003).
11. S.E. Bialkowski, *Anal. Chem.* 61(11), 1308–1310 (1989).
12. P.D. Wentzell, and C.D. Brown, "Signal Processing in Analytical Chemistry," in *Encyclopedia of Analytical Chemistry* (John Wiley & Sons, Chichester, 2000).
13. C. Brown, Private Communication (2002).
14. H. Mark and J. Workman, *Statistics in Spectroscopy* (Academic Press, New York, 1991). ■