This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Smoothing of Time Series

Volume Author/Editor: Frederick R. Macaulay

Volume Publisher: NBER

Volume ISBN: 0-87014-018-3

Volume URL: http://www.nber.org/books/maca31-1

Publication Date: 1931

Chapter Title: Fifth-Degree Parabolic Summation Formulas for Graduation

Chapter Author: Frederick R. Macaulay

Chapter URL: http://www.nber.org/chapters/c9364

Chapter pages in book: (p. 63 - 80)

CHAPTER IV

Fifth-Degree Parabolic Summation Formulas for Graduation.

If the investigator wishes both seasonal elimination and great power to smooth out erratic fluctuations in cyclical data without appreciably dampening the fluctuations of the underlying ideal curve, he must use something more flexible than a third-degree parabolic formula. He will naturally think of fifth-degree formulas. Approximately as close a fit to periodic data such as sine curves can be obtained with weights adapted to fit merely third-degree parabolas, only if the number of weights is radically reduced. However, less numerous weights tend to give a rougher graduation with non-mathematical data.

Sets of weights giving the mid ordinates of fourth (or fifth) [1] degree parabolas fitted by the method of least squares are not difficult to derive. A number are given in Whittaker and Robinson, *The Calculus of Observations*, page 296. The weight diagrams are themselves fourth-degree parabolas, just as the weight diagrams for the mid

[1] Of course, any symmetrical weight diagram which will fit a fourth-degree parabola will also fit a fifth-degree parabola.

ordinates of second (or third) degree parabolas are themselves second-degree parabolas. For example, the fourth (or fifth) degree weights for 13 observations are:

$$+\frac{110}{2431},\quad -\frac{198}{2431},\quad -\frac{135}{2431},\quad +\frac{110}{2431},\quad +\frac{390}{2431},\quad +\frac{600}{2431},$$
$$+\frac{677}{2431},\quad +\frac{600}{2431},\quad +\frac{390}{2431},\quad +\frac{110}{2431},\quad -\frac{135}{2431},\quad -\frac{198}{2431},$$
$$+\frac{110}{2431}.$$

Cyclical data or periodic functions can, of course, be more closely approximated by such fifth-degree parabolic formulas than by corresponding third-degree formulas. More observations can therefore be used. This leads to the possibility of greater smoothness in the graduation. However, the method is open to the same criticisms as third-degree least squares graduation in that it does not eliminate seasonal fluctuations and the fitted curve is extremely laborious to compute.

Fifth-degree parabolic summation formulas with smooth weight diagrams may be constructed almost as readily as third-degree formulas. Such summation formulas are free from a number of the disadvantages of fifth-degree least squares formulas. The amount of computation is much less. The smoothness of the results obtained by fitting to actual data is appreciably greater. Seasonal fluctuations are eliminated if the formula be properly constructed. For example, a smooth weight diagram which, if applied to a fifth-degree or any lower order of parabola, will fall exactly on that parabola and which will eliminate 12-months seasonal fluctuations, may be computed as follows: Take successively a 2-months moving total of the data, then a 3-months moving total, then another 3, then a 4, a 6, an 8, a 10, and a 12-months moving total. To the results apply the following set of weights: + 24,374, − 100,301, + 152,034, − 100,-301, + 24,374. Divide each of the final results by 74,649,600.[1] The 45 weights, to the nearest fifth decimal, are given in column 22 of the table in Appendix IV. Figure 22 of Chart I gives a picture of the weight diagram.

With such a set of fifth-degree parabolic weights it is, of course, possible to obtain a closer fit to data than with a set having the same number of weights but designed to fit merely second (or third) degree parabolas. This particular 45-term fifth-degree parabolic set of weights, if applied to monthly Call Money Rates, gives results resembling those obtained by means of the 43-term curves we have used in the study of interest rates and security prices. The amount of computation

[1] In actual computation this final step is unnecessary. The 5 weights are each divided by 74,649,600 and the results (taken to a sufficient number of decimals) are used instead of the 5 weights above.

is, however, much greater than that required for the 43-term formula.

The reader must remember that, though each of two weight systems contain 45 terms and each be designed to fall on fifth-degree and all lower order of parabolas, it is not necessary that they show the same fit when applied to other curves than parabolas. A 45-term fifth-degree parabolic summation formula which contains fewer moving totals and is therefore easier to compute than the one which has just been described and which also gives a closer fit to sine curves, may be applied as follows: Take successively a 3-months moving total of the data, a 5-months moving total, another 5, an 8, and a 12-months moving total. To the results apply the following set of weights: + 1,331,771, − 1,949,056, 0, 0, 0, 0, 0, 0, + 2,175,370, 0, 0, 0, 0, 0, 0, − 1,949,056, + 1,331,771. Divide each of the final results by 6,773,760,000.[1]

Fifth-degree parabolic weight systems containing a smaller number of terms can, of course, be constructed. The resulting smooth curves will generally follow the data more closely but will also tend to be somewhat less smooth and more affected by merely accidental irregularities and non-typical movements in the data. A smooth 43-term formula which will eliminate 12-months seasonal fluctua-

[1] See note 1, page 65, and Column 21 of the table in Appendix IV and Figure 21, Chart I.

tions and which, if fitted to a fifth degree or lower order of parabola will fall exactly on that parabola, may be applied as follows: Take successively a 5-months moving total of the data, then another 5, an 8, and a 12. To the results apply the following set of weights: + 3,819,893, − 5,598,848, 0, 0, 0, 0, 0, 0, + 6,380,310, 0, 0, 0, 0, 0, 0, − 5,598,848, + 3,819,893. Divide each of the final results by 6,773,760,000.[1]

In a similar manner a 41-term fifth-degree parabolic set of weights, which will eliminate 12-months seasonal fluctuations, may be applied as follows: Take successively a 3-months moving total, a 5, an 8, and a 12-months moving total. To the results apply the following set of weights: + 1,158,703, − 1,703,808, 0, 0, 0, 0, 0, 0, + 2,031,-010, 0, 0, 0, 0, 0, 0, − 1,703,808, + 1,158,703. Divide each of the final results by 1,354,752,000.[2]

A 35-term strictly fifth-degree parabolic set of weights, which will eliminate 12-months seasonal fluctuations, but which has a weight diagram that is only moderately smooth and well shaped, may be applied as follows: Take successively 3, 5, 8 and 12-months moving totals of the data. To the results apply the following set of weights: + 273,632, − 472,175, 0, 0, 0, + 469,086, 0, 0, 0,

[1] See note 1, page 65, and Column 20 of the table in Appendix IV and Figure 20, Chart I.

[2] See note 1, page 65, and Column 19 of the table in Appendix IV and Figure 19, Chart I.

− 472,175, + 273,632. Divide each of the final results by 103,680,000.[1]

Similar strictly fifth-degree formulas may be constructed containing even fewer weights than 35. A 33-term formula which is not strictly fifth-degree (though nearly so) but which is less laborious to use than the preceding strictly parabolic formulas may be applied as follows: Take successively 5, 8 and 12-months moving totals of the data. To the results apply the following simple weights: + 58, − 100, 0, 0, 0, + 100, 0, 0, 0, − 100, + 58. Divide each of the final results by 7680.

The fact that such a formula will not exactly fit a fifth-degree parabola is of negligible significance.[2] When used for graduating such a series as the 97-months of Call Money Rates printed in Appendix VIII, it gives results differing little from those obtained by using the 35-term strictly fifth-degree parabolic formula already described. The graduation is a shade less smooth than that given by the 35-term formula. The formula fits sine curves of short period a shade closer than the 35-term formula. It does not fit those of long period quite so well.

When the number of weights in a fifth-degree

[1] See note 1, page 65, and Column 18 of the Table in Appendix IV and Figure 18, Chart I.

[2] See pages 70, 71.

parabolic, 12-months seasonal eliminating formula is reduced much below 33 or 31, the resulting graduation of such data as early monthly Call Money Rates tends to show a large increase in the number of points of inflection and even in the number of maxima and minima. The graduated curve tends to "weave" among the data. Artificial and arbitrary sinuosities make their appearance. Such peculiar behavior is not difficult to explain. When the number of weights is reduced much below 33 or 31, it becomes impossible to obtain a smooth and well-shaped weight diagram.[1] This impossibility is the result of insisting upon sea- sonal eliminating qualities. The reader will understand how great a restriction this imposes on the shape of the weight diagram when he remembers that the formula is necessarily so constructed

[1] The smallest possible number of weights which will eliminate 12-months seasonal fluctuations and fall on fifth and all lower orders of parabolas is sixteen. However, the points on the smooth curve resulting from the use of such a weight diagram would have to be centered between data dates. With a 17-weight diagram the points on the curve will be centered on data dates. Such a 17-weight diagram necessarily involves the following operations and no others: Take a 12-months moving total of the data. To the results apply the following set of weights:

$$+\frac{3775}{34569},\ -\frac{12201}{34569},\ +\frac{8570}{34569},\ +\frac{8570}{34569},\ -\frac{12201}{34569},\ +\frac{3775}{34569}.$$

There are 17 weights in all. To the first two decimals the weights are: +1.09, −2.43, +.04, +2.52, −1.00, +.08, +.08, +.08, +.08, +.08, +.08, +.08, −1.00, +2.52, +.04, −2.43, +1.09: Total = +1.00. This gives a decidedly bizarre weight diagram, as may be seen from examining the figures.

that it gives the same graduation if applied to the raw data as it does if applied to the data after an adjustment for seasonal fluctuations—*no matter what the seasonal fluctuations may be assumed to be.* Now, if 33 or more terms are taken in the formula, the seasonal fluctuations, *however absurd they may be assumed to be*, are taken care of quite easily by the mere fact of their repetition. However, as the number of terms is reduced, a point is reached at which the repetition of the seasonal fluctuations is not great enough. The formula can then no longer handle the data adequately. The danger line seems to be reached about where it becomes necessary to use a formula having two distinct modes instead of merely one.[1]

While good results can be obtained by fitting strictly fifth-degree parabolic formulas to series such as Call Money Rates, we must not forget that parabolic formulas are purely empirical. There is no logical reason why we should insist upon a formula which will exactly fit a fifth-degree parabola. For most time series an even more important consideration than how closely the formula will fit a fifth-degree parabola is how closely it will fit sine curves of various periods. Moreover, if we do not insist upon an exact fit to a fifth-degree parabola, it is possible easily to develop formulas which

[1] For an example of a third-degree parabolic duo-model formula, see Appendix III.

will give *approximate* fits to such fifth-degree parabolas and also to sine curves of a large range of period, and at the same time be extremely simple to compute. As an example of such a formula, take a 3-months moving total of a 5-months moving total of an 8-months moving total of a 12-months moving total of the data. To the results apply the following extremely simple set of weights: +2, −3, 0, 0, 0, 0, 0, +3, 0, 0, 0, 0, 0, −3, +2. Divide the final results by 1440.[1]

If the 39-term weight system resulting from the above simple procedure be applied to 39 terms of the parabola $y = x^2$ (from $x = -19$ to $x = +19$), the resulting point on the fitted curve will be $-1/6$, in other words the fitted curve will, at this point, fall $1/6$ outside the parabola. If the formula be similarly applied to the parabola $y = x^4$, the resulting point will be $+415\,3/10$. In other words the fitted curve at this point will fall a little more than 415 inside the parabola $y = x^4$. However, as the second-degree parabola above has a range from zero to +361, and the fourth-degree parabola a range from zero to +130,321, the 39-term formula accounts for 100.05 per cent of the term involving the second power of x and 99.68 per cent of the

[1] Actual computation is simplified by using the weights: +10, −15, 0, 0, 0, 0, 0, +15, 0, 0, 0, 0, 0, −15, +10, and dividing by 7,200.

For weights and weight diagram, see Column 23 of the table in Appendix IV and Figure 23, Chart I.

term involving the fourth power of x in the equation $y = A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5$.[1] The formula, of course, fits exactly all the other terms in the equation. This 39-term formula must therefore give an extremely close approximation to any fifth-degree parabola to which it may be fitted. An examination of column 23 of the table in Appendix VII will give the reader an idea of how adequate a fit such a formula will give to a large range of sine curves. From that table, it may be seen that it gives an adequate fit to a larger range of sine curves than any one of the five strictly fifth-degree parabolic formulas described above. Even the 35-term fifth-degree formula gives a closer fit only to sine curves of a shorter period than 19 months,[2] in spite of the fact that the 39-term formula contains enough terms to free it almost entirely from most of the undesirable sinuosities which show a tendency to appear when strictly parabolic formu-

[1] The 33-term formula accounts for 99.84 per cent of the term involving the second power of x and 99.90 per cent of the term involving the fourth power of x.

[2] When fitted to sine curves whose periods are around 40 months, both this 39-term approximately parabolic formula and the 43-term formula discussed immediately below give results falling a fraction of 1 per cent of the amplitude of the sine curve *outside* the sine curve. This is, of course, entirely negligible. It could have been corrected by a slight change in the moving totals used in computation, but such a change does not seem worth making, as any correction of this negligible amount outside the sine curve in the 40 months period range would tend to give an appreciably poorer fit to sine curves of shorter period than 30 months.

las of few terms are used. The results of applying this 39-term formula to 97 months of Call Money Rates may be seen from examining Column 23 of the table in Appendix VIII.

A somewhat similar approximately fifth-degree formula, containing 43 terms, may be applied as follows: Take a 5-months moving total of a 5-months moving total of an 8-months moving total of a 12-months moving total of the data. To the results apply the following extremely simple weights: + 7, − 10, 0, 0, 0, 0, 0, 0, + 10, 0, 0, 0, 0, 0, 0, − 10, + 7. Divide the final results by 9600. There are 43 weights in all.[1]

When such a 43-term formula is applied to actual data, it gives a curve which is slightly smoother and freer from any hint of sinuosity than even the 39-term formula discussed immediately above. The results of applying such a 43-term formula to sine curves of various periods may be seen by examining Column 24 of the table in Appendix VII. The results of applying it to 97 months of Call Money Rates may be seen from the table in Appendix VIII. Like the 39-term formula, this 43-term formula is approximately parabolic. It corrects for 99.96 per cent of the term involving the second power of x and 99.59 per cent of the term involving the fourth power of x in the equa-

[1] For the weights, see Column 24 of the table in Appendix IV. For the weight diagram, see Figure 24, Chart I.

tion $y = A + Bx + Cx^2 + Dx^3 + Ex^4 + Fx^5$, and, of course, 100 per cent of all the other terms. It is the curve we have used throughout the study of interest rates and security prices.

The fact that neither the above 39-term formula nor the above 43-term formula will give a graduation falling exactly on a fifth-degree parabolic curve if fitted to such a curve is a matter of no practical significance. Even if the graduation deviated considerably further from such a fifth-degree parabolic curve than it actually does, such deviation would not necessarily be important. There is nothing magical in parabolic curves. It is true that, if enough terms be taken in a parabolic or power expansion, it may be made to represent almost any curve whatever. However, such representation is purely empirical and, if too many terms be taken in the parabolic expansion, leads to impossible and ridiculous results immediately outside the range of the data.[1]

For most economic time series, approximation to any particular section of the curve is more naturally accomplished by means of sine curves than by means of parabolas. The reader must remember that such series are commonly of an un-

[1] Similar purely empirical fitting may be performed by means of harmonic analysis. In this case, the particular curve is not approximated by means of successive parabolas, but by means of a series of sine curves.

mistakably cyclical type. For smoothing series such as those handled in our study of interest rates and security prices, both the 39-term approximately parabolic formula and the 43-term approximately parabolic formula are to be highly recommended.[1]

Before leaving this section of the discussion, the reader might care to examine Appendix IV, in which the weights are given for each of the formulas just discussed,[2] Chart I in which the weight diagrams are presented, Appendix VII in which the results of applying the various formulas [2] to sine curves of different periods are tabulated, and Appendix VIII, in which are given the graduations which result from applying the various formulas [2] to 97 months of Call Money Rates.

The figures in Chart I are weight diagrams corresponding to the following weight systems:

Fig. 1. A 12-months simple moving average.
Fig. 2. A 2-months moving average of a 12-months moving average.
Fig. 3. An 8-months moving average of a 12-months moving average.
Fig. 4. A 4-months moving average of a 5-months moving average of a 6-months moving average.

[1] Even the 33-term formula may be used with confidence, and it involves one less computation step than either the 39-term or the 43-term. On the other hand, there are some slight advantages in using a formula with more than 33 terms. The graduation is smoother and there is less chance of its ever being more than microscopically affected by the assumption of constant seasonal fluctuations over only a short period.

[2] Except the 33-term formula.

CHART I

GRADUATION WEIGHT DIAGRAMS

A GRAPHIC REPRESENTATION OF THE WEIGHTS IMPLIED IN TWENTY-FOUR SMOOTHING FORMULAS

IN EACH DIAGRAM
THE SUM OF THE ORDINATES EQUALS UNITY

FIG. 1

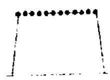


FIG. 2

FIG. 3

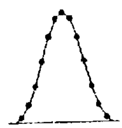


FIG. 4

FIG. 5

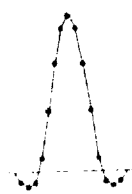


FIG. 6

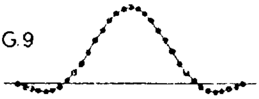
FIG. 7

FIG. 8

CHART I

GRADUATION WEIGHT DIAGRAMS
(CONTINUED)

FIG. 9



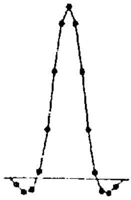
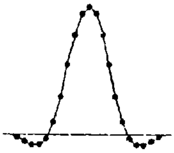


FIG. 10

FIG. 11



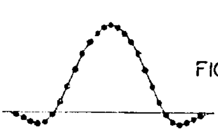
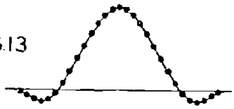


FIG. 12

FIG. 13



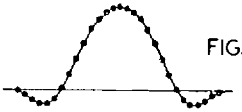


FIG. 14

FIG. 15

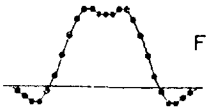


FIG. 16

CHART I

GRADUATION WEIGHT DIAGRAMS
(CONCLUDED)

FIG. 17

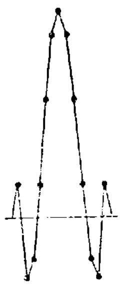


FIG. 18

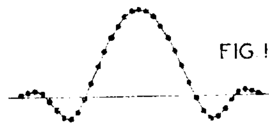


FIG. 19

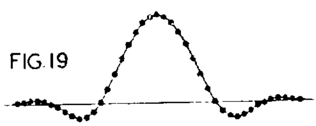


FIG. 20

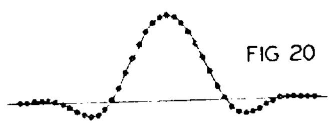


FIG. 21

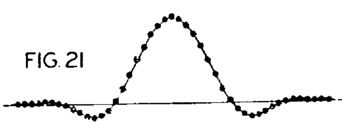


FIG. 22

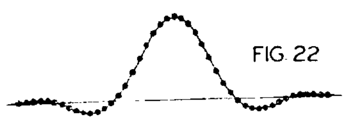


FIG. 23

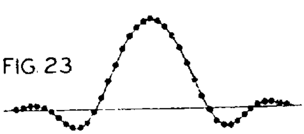
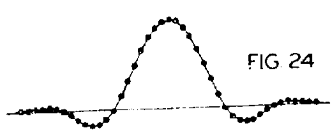


FIG. 24



When examining the preceding weight diagrams, the reader must remember that the weights are infinite in number. To the right and left of the weights represented by black dots, he must think of an infinite number of zero weights (lying on the horizontal straight line). For example, Figure 5 must be considered an extremely rough weight diagram, in spite of the fact that all the real plus or minus weights lie on a second-degree parabola. The *complete* weight diagram (which includes the zero weights) shows four violently disturbing angles—two at the last real weights (represented by black dots) and two at the first zero weights (on the straight line).