ACTUARIAL RESEARCH CLEARING HOUSE
1992 VOL. 2

AGGREGATE SMOOTHNESS IN MULTI-DIMENSIONAL WHITTAKER-HENDERSON GRADUATION

by Ronald Dubson

abstract

A set of crude rates u'' determined by D independent variables can be thought of as a discrete surface of dimension D. Let W, F, K_i, and S_i have their usual meaning in multi-dimensional Whittaker-Henderson graduation. Let $t_1, \ldots, t_D$ be non-negative smoothness parameters, let

$$A = W + \sum_{i=1}^{D} t_i K_i^T K_i , \qquad u = u(t_i) = A^{-1} W u'' ,$$

let η be the surface determined by $\lim_{t_i \to \infty} (u - \eta)^T W(u - \eta) = 0$, and let $\sigma^2 > 0$ be some desired level of the fit F (e.g., its expected value). Then the point (t_i) at which $(u - \eta)^T W(u - \eta)$ attains a minimum when subject to the constraint $F(u) = \sigma^2$ is determined by the simultaneous equations

$$\frac{\partial F}{\partial t_i} = \lambda S_i \qquad (i = 1, \ldots, D)$$

where λ depends on σ^2.

If C is any smooth curve in the region $\{(t_i): t_i \geqslant 0\}$ with initial point $(0, \ldots, 0)$ and end-point (t_i), then

$$F + \sum_{i=1}^{D} t_i S_i = \int_C \sum_{i=1}^{D} S_i \, dt_i .$$

1. Given a set of crude motality rates u''_x to be graduated by the Whittaker-Henderson method, one may ask under what conditions there exists a statistical estimator for the fit $F = \sum w_x(u_x - u''_x)^2$. If the graduated rates u_x are considered to be random variables with expectation u_x'' and variance σ_x^2, then the expected value of F is

$$E\{F\} = \sum w_x E\{(u_x - u_x)^2\} = \sum w_x \sigma_x^2 \quad .$$

In the case of rates $q_x = \theta_x / E_x$ based on the number of lives or number of policies, it is well known that $\sigma^2_x = q_x(1 - q_x) / E_x = p_x q_x / E_x$.

For the so-called type B Whittaker-Henderson graduation, in which the weights w_x are the exposures E_x, one has

$$E\{F\} = \sum E_x (p_x q_x / E_x) = \sum p_x q_x \quad .$$

In multidimensional WH graduation where the subscript x is an ordered pair, or more generally, a lattice point $(k_1, ..., k_D)$, there is an infinite number of different values for the smoothness parameters which yield different graduated rates but the same fit. These considerations suggest the following problem:

For a given level of fit, how are the smoothness parameters to be chosen so that the graduated rates form, in some natural sense, the smoothest discrete surface?

In the one-dimensional case this is not an issue, as the fit determines the single smoothness parameter, and *vice versa*. Although this is a mathematical rather than an actuarial problem, it should not be without some interest to actuaries, since a solution would provide a standard against which to compare graduated rates that emphasize smoothness along a particular axis or axes.

2. We will use Knorr's notation [1], with the exception that the smoothness parameters will be denoted by t_i ($i = 1, \ldots, D$).

D = dimension of the data set of crude rates,

$N = n_1 n_2 \ldots n_D$, the total number of cells

$u = \{u_{k_1 k_2 \ldots k_D} : k_i = 1, \ldots n_i ; i = 1, \ldots, D\}$ the "unraveled" graduated rates,

$u = (u_1, \ldots, u_N)$ the "raveled" graduated rates,

$u = (u_1'', \ldots, u''_N)$ the raveled crude rates

W = an $N \times N$ diagonal matrix with positive entries $w_1, \ldots, w_N$;

K_i = the $N \times N$ matrix with binomial coefficients needed to determine $\underset{i}{\Delta}^{z_i}$,

K^T_i = the transpose of K_i ,

$$A = W + \sum_{i=1}^{D} t_i K_i^T K_i \quad (t_i > 0);$$

$$F = (u-u'')^T W (u-u'');$$

$$S_i = u^T K^T_i K_i u.$$

Note that $u^T K^T_i K_i u = \sum (\underset{i}{\Delta}^{z_i} u)^2$; e.g., for a two-dimensional data set with $z_1 = 3$,

$$u^T K_1^T K_1 u = \sum_{i,j} (\underset{1}{\Delta}^{3} u_{i\,j})^2$$

$$= \sum_{j=1}^{n_2} \sum_{i=1}^{n_1-3} (u_{i+3,\,j} - 3u_{i+2,\,j} + 3u_{i+1,\,j} - u_{i\,j})^2 .$$

The vector u minimizes $F + \sum t_i S_i$ when $Au = Wu''$. As $t_i \to \infty$, $u \to \eta$, the polynominal function of $(k_1 \ldots, k_D)$ which minimizes F and whose partial derivatives $\partial^{z_i} / \partial k_i$ are zero. The vector η can also be charcterized as the orthogonal projection of u'' onto the intersection of the null spaces of the K_i's. This fact is used in appendix (i).

3. Let σ^2 denote the desired level of fit. As the "smoothest" u (= $A^{-1}Wu^{*}$) for which $F = \sigma^2$, we propose the one which minimizes $(u - \eta)^T W (u - \eta)$, the square of the distance to the ultimate smooth surface η.

In order to derive an equation for the (t_i) at which this condition is satisfied, it will be convenient to use the following notation:

$<u,v> = u^T W v$, $| u |^2 = < u,u >$; hence, $F = | u - u^* |^2$ and

$(u - \eta)^T W (u - \eta) = | u - \eta |^2$.

It is shown in appendix (i) that

$| u - \eta |^2 = 2 < u,u^* > + F - | u^* |^2 - | \eta |^2$,

from which it follows that, when $F = \sigma^2$, $| u - \eta |^2$ and $2 < u, u^* >$ differ by a constant. Hence, $<u,u^*>$ attains a minimum, say μ, at the same point (t_i) as does $| u - \eta |^2$. By Lagrange's Theorem [2], there is a real number λ, depending on (t_i), such that

(1) $\partial <u,u^*>/ \partial t_i = \lambda \ \partial F/\partial t_i$, $i = 1, ..., D$

at the point on the submanifold $\{F = \sigma^2\}$ where $<u,u^*> = \mu$. Geometrically speaking, this says that (t_i) is the point of tangency between the two submanifolds $\{F = \sigma^2\}$ and $\{< u, u^* > = \mu\}$. Replacing λ by $-1/\lambda$ in equation (1) and using the fact that $\partial<u,u^*> /\partial t_i = - S_i$ [appendix (ii)], one obtains

1a) $\partial F / \partial t_i = \lambda S_i$, $i = 1, ..., D$.

Note that $\partial F / \partial ti = 2 <u^* - u, A^{-1} K^T_i K_i u>$
[appendix (iii)]. Since $u = u^*$ at $t_i = 0$,
$\partial F / \partial t_i = 0$ at $t_i = 0$. Hence, if $S_i (u^*) > 0$
for some i, then $\lambda = 0$ at $t_i = 0$.

4. Let $\phi = \begin{vmatrix} \partial_1 F & S_1 \\ \partial_2 F & S_2 \end{vmatrix}$ $\qquad (\partial_i = \partial / \partial t_i)$

In the case D = 2, the equations (1a) are equivalent to the single determinant equation $\phi = 0$, which determines a curve Γ with initial point (0,0). We have thus reduced our problem to locating the intersection of the two curves $F = \sigma^2$ and Γ. An algorithm for approximating this point is the following:

(1) Calculate the slope m_1 of Γ at the origin; it follows from what is shown in section 5 that

$$m_1 = \begin{vmatrix} \partial_{11} F & S_1 \\ \partial_{12} F & S_2 \end{vmatrix} \div \begin{vmatrix} S_1 & \partial_{21} F \\ S_2 & \partial_{22} F \end{vmatrix} ,$$

where all functions are evaluated at (0,0). The straight line through the origin with slope m_1 intersects $F = \sigma^2$ at some point P_1 ;

(2) the tangent line to $F = \sigma^2$ at P_1 has slope $m_2 = - \partial_1 F(P_1) / \partial_2 F(P_1)$ and meets Γ at some point P_2 ; namely, where $\phi(P_2) = 0$;

(3) the straight line through the origin and P_2 intersects $F = \sigma^2$ at a point P_3 ;

(4) repeat step (2) with P_3 in place of P_1 to obtain a point P_4 on Γ. The sequence (P_n) converges to the intersection of $F = \sigma^2$ and Γ (see figure). Finding P_1, P_2, P_3 involves solving for t_1, t_2, t_3 in the equations

$$F(t_1(1, m_1))\ m = \sigma^2 ,$$

$$\phi(P_1 + t_2(1, m_2)) = 0 ,$$

$$F(t_3(1, m_3)) = \sigma^2, \text{ where } m_3 = \frac{m_1 t_1 + m_2 t_2}{t_1 + t_2} .$$

5. When the dimension D> 2 , $F = \sigma^2$ is not a curve but a (D−1) − dimensional surface. This obviates the use of the preceding algorithm. We will show that Γ is an integral curve of a vector field V. Hence, any point of Γ can be approximated by numerical integration of V, starting at the origin. The modified Cauchy-Euler method [3] is an efficient algorithm for this purpose.

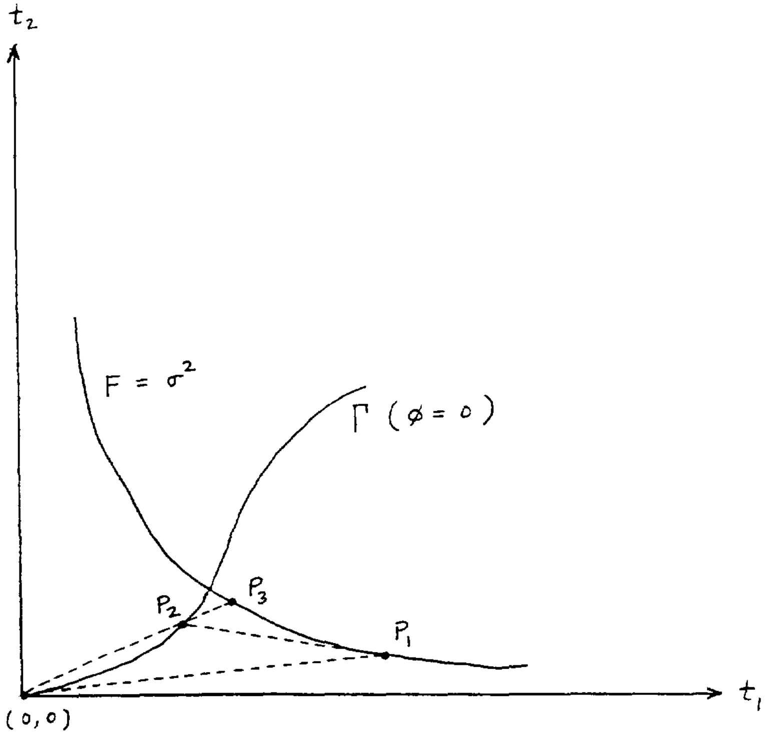

figure: Locating the intersection of $F = \sigma^2$ and $\phi = 0$.
(cf. para. 4)

Assume that the parameter λ in (1a) is monotone in a segment of Γ containing (t_i). Differentiating (1a) with respect to λ gives

(2a) $\sum_j \partial_j(\partial_i F)\frac{dt_j}{d\lambda} = S_i + \lambda \sum_j (\partial_j S_i)\frac{dt_j}{d\lambda}$, (i = 1, ..., D) or

(2b) $(\partial^2 F - \lambda\,\partial S)\frac{d\Gamma}{d\lambda} = S$,

where the matrix $\partial^2 F$ has entries $\partial_j \partial_i F$ or $\partial_{ji} F$,

∂S has entries $\partial_j S_i$,

the column vector $d\Gamma/d\lambda$ has entries $dt_i/d\lambda$,

and S has entries S_i .

The matrix equation (2b) can be solved for $d\Gamma/d\lambda$ by applying Cramer's Rule. Consider first the case D = 2. If Δ is the determinant of $\partial^2 S - \lambda\,\partial S$, then

$$\Delta \frac{dt_1}{d\lambda} = \begin{vmatrix} S_1 & \partial_{21}F - \lambda\partial_2 S_1 \\ S_2 & \partial_{22}F - \lambda\partial_2 S_2 \end{vmatrix}$$

$$= \begin{vmatrix} S_1 & \partial_{21}F \\ S_2 & \partial_{22}F \end{vmatrix} - \begin{vmatrix} S_1 & \lambda\partial_2 S_1 \\ S_2 & \lambda\partial_2 S_2 \end{vmatrix}$$

$$= \begin{vmatrix} S_1 & \partial_{21}F \\ S_2 & \partial_{22}F \end{vmatrix} - \begin{vmatrix} \lambda S_1 & \partial_2 S_1 \\ \lambda S_2 & \partial_2 S_2 \end{vmatrix}$$

(3a) $$= \begin{vmatrix} S_1 & \partial_{21}F \\ S_2 & \partial_{22}F \end{vmatrix} - \begin{vmatrix} \partial_1 F & \partial_2 S_1 \\ \partial_2 F & \partial_2 S_2 \end{vmatrix} = \mathcal{E}_1,$$

and similarly,

(3b) $$\Delta \frac{dt_2}{d\lambda} = \begin{vmatrix} \partial_{11}F & S_1 \\ \partial_{12}F & S_2 \end{vmatrix} - \begin{vmatrix} \partial_1 S_1 & \partial_1 F \\ \partial_1 S_2 & \partial_2 F \end{vmatrix} = \mathcal{E}_2 .$$

The vector field V defined by $(\varepsilon_1, \varepsilon_2)$ depends only on (t_i) and not on λ. On the curve Γ, V is parallel to the tangent vector $d\Gamma/d\lambda$.

At $t_i = 0$, $\partial_i F = 0$ [cf. section 3]; denoting the value of ε_i at $t_i = 0$ by ε_i'', one therefore has

$$\varepsilon_1'' = \begin{vmatrix} S_1 & \partial_{21} F \\ S_2 & \partial_{22} F \end{vmatrix} ,$$

$$\varepsilon_2'' = \begin{vmatrix} \partial_{11} F & S_1 \\ \partial_{12} F & S_2 \end{vmatrix} ,$$

where all functions are evaluated at $t_i = 0$. Now $(\varepsilon_1'', \varepsilon_2'') \neq (0,0)$ if $S_1 \neq 0$ or $S_2 \neq 0$ and

$$\begin{vmatrix} \partial_{11} F & \partial_{21} F \\ \partial_{12} F & \partial_{22} F \end{vmatrix} \neq 0 .$$

If these conditions hold at $t_i = 0$, then one can solve the system of differential equations

$$\frac{dt_1}{d\tau} = \varepsilon_1 , \quad \frac{dt_2}{d\tau} = \varepsilon_2$$

with initial conditions $t_1(0) = 0$, $t_2(0) = 0$; ε_1, ε_2 are given in equations (3a) and (3b). In particular, one can solve for the point (t_i) on Γ where $F = \sigma^2$.

For the case $D > 2$ the equation (2b) cannot be solved so neatly, but the principle is the same. Replacing λ by $\partial_i F/S_i$ in equation (2a), one obtains

$$\text{(2c)} \quad \sum_j \left[\partial_{ji} F - \partial_i F \, \partial_j S_i / S_i \right] \frac{dt_j}{d\lambda} = S_i ,$$

which at $t_i = 0$ reduces to

$$\sum_j \partial_{ji} F \frac{dt_j}{d\lambda} = S_i , \qquad (i = 1, \ldots, D)$$

where $S_i = u''^T K_i^T K_i u''$

and $\partial_{ji} F = 2 u''^T K_j^T K_j W^{-1} K_i^T K_i u''$ [cf. appendix (iii)].

The author wishes to thank Arthur Cragoe for calling his attention to Mr. Knorr's article and for encouraging the work many years ago which led to this paper.

APPENDIX

(i) Since $K_i \eta = 0$, $A\eta = W\eta$ and $\eta = A^{-1} W\eta$. As η is an orthogonal projection of u'', it follows by definition that $<u'' - \eta, \eta> = 0$ or $<u'', \eta> = <\eta, \eta>$; hence
$<u, \eta> = <A^{-1} Wu'', \eta> = u''^T WA^{-1} W\eta = u''^T W\eta = <u'', \eta> = <\eta, \eta>$; hence

$$| u - \eta |^2 = | u |^2 - 2 <u, \eta> + | \eta |^2$$

$$= | u |^2 - | \eta |^2$$

$$= | u - u'' |^2 + 2 < u, u'' > - | u'' |^2 - | \eta |^2 .$$

(ii) Since $Au = Wu'' =$ constant,
$0 = \partial_i Wu'' = \partial_i (Au) = (\partial_i A) u + A \partial_i u$;
$\partial_i A = K^T_i K_i$ now gives
$\partial_i u = - A^{-1} (\partial_i A) u = -A^{-1} K_i^T K_i u$; finally,
$\partial_i < u, u'' > = < u'', \partial_i u > = < u'', -A^{-1} K^T_i K_i u >$
$= -u''^T W A^{-1} K^T_i K_i u = -u^T K^T_i K_i u = -S_i$.

We note a striking identity which follows from this fact. Let C be any smooth curve in D-dimensional space with initial point $(0, ..., 0)$ and endpoint $(t_1, ..., t_D)$. Then,
$Au = Wu''$,

$$u^T Au = u^T Wu'',$$

$$u^T Wu + \sum t_i S_i = u^T Wu'', \quad \text{and}$$

$$\sum t_i S_i = u^T W (u'' - u) = <u, u'' - u>; \quad \text{hence,}$$

$$F + \sum t_i S_i = <u'' - u, u'' - u> + <u, u'' - u>$$

$$= <u'', u'' - u>$$

$$= <u'', u''> - <u, u''>$$

$$= - \int_c \sum \partial_i <u, u''> dt_i$$

$$= \int_c \sum S_i \, dt_i .$$

$$F + \sum t_i S_i = \int \sum S_i \, dt_i$$

(iii) $\partial_i F = \partial_i < u\text{-}u^*, u\text{-}u^* >$

$= 2 <u\text{-}u^*, \partial_i u>$

$= 2 <u^* - u, A^{-1} K^T{}_i K_i u>; \quad \partial_{ji} F =$

$\partial_j (\partial_i F) = 2 < -\partial_j u, A^{-1} K^T{}_i K_i u> + 2 < u^* - u, \dots >;$

since, at $t_i = 0$, $u = u^*$ and $A = W$, we have

$$\partial_{ji} F \Big|_{t_i=0} = 2 <A^{-1} K^T{}_j K_j u, \; A^{-1} K^T{}_i K_i u > \Big|_{t_i=0}$$

$$= 2 \, u^{*T} K^T{}_j K_j W^{-1} K^T{}_i K_i u^*.$$

REFERENCES:

1. Knorr, Frank E., "Multidimensional Whittaker - Henderson Graduation" TSA XXXVI (1985).

2. Bartle, Robert G., Elements of Real Analysis, John Wiley & Sons, Inc., New York, 1964, Chapter 21.

3. Kellison, S.G., Fundamentals of Numerical Analysis, Richard D. Irwin, Homewood, Illinois, 1975, Chapter 10.

Kansas City, Missouri 1991