G Taylor 37

ISBN 0 7325 1208 5

THE UNIVERSITY OF MELBOURNE

SMOOTHNESS CRITERIA FOR MULTI-DIMENSIONAL

WHITTAKER GRADUATION

by
Greg Taylor
The University of Melbourne

RESEARCH PAPER NUMBER 37

October 1996

Centre for Actuarial Studies
Department of Economics
The University of Melbourne
Parkville, Victoria, 3052
Australia.

SMOOTHNESS CRITERIA FOR MULTI-DIMENSIONAL WHITTAKER GRADUATION

Greg Taylor

Consultant, Tillinghast–Towers Perrin,
GPO Box 3279,
SYDNEY NSW 2001, AUSTRALIA

and

Professorial Associate, Centre for Actuarial Studies,
Faculty of Economics and Commerce,
University of Melbourne,
Parkville, Victoria 3052, AUSTRALIA

October 1996

Summary The smoothness term of the multi-dimensional Whittaker graduation objective function is considered. Its form is limited by the fact that a smoothness measure should remain invariant under orthogonal coordinate transformations. A smoothness measure satisfying this requirement is constructed. It treats all polynomials of degree $<m$ in n dimensions as smooth. It is based on m-th order differences, but is otherwise independent of m and n if m is odd. If m is even, it depends on both m and n. The Bayesian interpretations of Whittaker graduation made in recent years have been extended to the multi-dimensional case.

Keywords Multi-dimensional Whittaker graduation, smoothness, Bayesian.

1 Introduction

Whittaker graduation was devised by Whittaker (1923) and introduced into the actuarial literature by Henderson (1932). Since then, it has appeared in a number of standard actuarial texts, e.g. London (1985).

All of the early treatments involved a one-dimensional sequence of observations as the subject of graduation. Typically, a sequence of empirical mortality rates by age was involved.

Two-dimensional Whittaker graduation was introduced by McKay and Wilkin (1977) and extended by Knorr (1984) and Lowrie (1993). Most the extension has been concerned with generalisation of the smoothness criterion:

- to an environment in which the observations subject to graduation lie within a space of arbitrary dimension;
- smoothness is measured in terms of differences (of the graduated response) of arbitrary order, and differences which are mixed in an arbitrary manner in the axes with respect to which they are taken.

The most recent contribution (Broffitt, 1996):

- interprets and discusses the inclusion of such mixed differences; and
- gives matrix formulas for their construction.

The present paper makes use of certain symmetry properties which it will often be useful for a multi-dimensional smoothness criterion to satisfy. These properties remove a large part of the arbitrariness from the selection of difference terms to be included in the criterion, and considerably simplify its matrix construction.

A second strand of development of Whittaker graduation, concerned with its interpretation in a Bayesian context, was begun by Taylor (1992) and developed by Verrall (1993). The findings of the present paper are interpreted in this context.

2 Notation and terminology

2.1 Whittaker graduation

Consider points $x \in R^n$, Euclidean n-space. It will be assumed that the x form a regular lattice L in the sense that

$$x = (x^1, x^2, \ldots, x^n), \tag{2.1}$$

with

$$x^i = r_i, r_i + 1, \ldots, s_i.$$

For reasons which will appear shortly, the contravariant notation x^i is used here in preference to x_i.

Let f: $L \to R$, and call the $f(x)$ **observations.**

The objective is to find Whittaker **smoothed values** $W(x)$ corresponding to $f(x)$.

Define

$$E = \Sigma\, w(x)[f(x) - W(x)]^2, \tag{2.2}$$

for a suitable set of weights $w(x)$. Here, as elsewhere, summations run over $x \in L$ unless otherwise indicated.

Define

$$S = \sum_{x} \sum_{i_1 \ldots i_m} \alpha_{i_1 \ldots i_m} \left[\Delta_{i_1} \Delta_{i_2} \cdots \Delta_{i_m} W(x)\right]^2, \tag{2.3}$$

where the $\alpha_{i_1 \ldots i_m}$ are real constants and Δ_i is the is the forward difference operator taken in the direction of the i-th coordinate of x. The indexes $i_1, \ldots, i_m$ each run over the range 0 to n, with the convention that Δ_0 is the identity operator.

It will sometimes be useful to write this in the form

$$S = \sum_{x} S_x,$$

with

$$S_x = \sum_{i_1 \ldots i_m} \alpha_{i_1 \ldots i_m} \left[\Delta_{i_1} \ldots \Delta_{i_m} W(x)\right]^2.$$

Define

$$F = E + kS, \tag{2.4}$$

where $0 < k < \infty$.

Whittaker graduation is carried out by choosing the $W(x)$ in such a way as to minimise the objective function F. This function consists of a linear combination of the **error criterion** E and the **smoothness criterion** S. The combination is effected by means of the **tuning constant** or **relativity constant** k.

As defined here, there is one redundant degree of freedom in the definition of k and the α's.

2.2 Tensor notation

Tensor notation is the natural vehicle for the multi-dimensional framework considered here.

The Euclidean metric tensor is

$$\delta_{ij} = \delta^{ij} = 1 \text{ if } i = j;$$
$$= 0 \text{ otherwise,}$$

for $i, j = 1, 2, ..., n$.

The **Einstein summation convention** will be used throughout. Thus, a repetition of a suffix within a tensor product indicates summation over that suffix, e.g. for tensors A_{ij} and B^k,

$$A_{ij}\, B^j \equiv \sum_j A_{ij}\, B^j.$$

The metric tensor effects raising and lowering of suffixes, e.g. for a tensor T^{ij}_{klm},

$$\delta_{hi}\, T^{ij}_{klm} = T^{j}_{hklm}. \qquad (2.5)$$

In particular,

$$\delta^i_j = \delta^{ik}\, \delta_{kj}$$
$$= 1 \text{ if } i = j \qquad (2.6)$$
$$= 0 \text{ otherwise.}$$

Define

$$\delta^{i_1 \ldots i_p}_{j_1 \ldots j_p} = \delta^{i_1}_{j_1}\, \delta^{i_2}_{j_2} \ldots \delta^{i_p}_{j_p}. \qquad (2.7)$$

With the summation convention, (2.2) becomes

$$E = \left[f(x^1, ..., x^n) - W(x^1, ..., x^n)\right] w(x^1, ..., x^n, y^1, ..., y^n)$$
$$\times \left[f(y^1, ..., y^n) - W(y^1, ..., y^n)\right], \qquad (2.8)$$

where $w(x^1, ..., y^n)$ is the tensor given by

$$w(x, y) = w(x) \text{ for } x = y,$$
$$= 0, \text{ otherwise.} \qquad (2.9)$$

It is convenient to write (2.8) in the simple form:

$$E = \left[f(x) - W(x)\right] w(x,y) \left[f(y) - W(y)\right], \tag{2.10}$$

where $W(x)$ now assumes the dual role of denoting (1) the evaluation of W at x, and (2) the tensor W with generic argument x.

Let $\Delta_j(y)$ be interpreted as the tensor which, when multiplied by $W(x)$, acts as a forward difference operator at y,

$$\Delta_j(y)\ W(x) = W(y^1, ..., y^j + 1, ..., y^n) - W(y^1, ..., y^j, ..., y^n). \tag{2.11}$$

Note that (2.11) does not involve x on the right and sometimes may be conveniently denoted by just $\Delta_j(y)\,W$.

Write

$$K_{i_1 ... i_m}(y) = \Delta_{i_1}(y) ... \Delta_{i_m}(y) \qquad \text{[no summation over } y\text{]}. \tag{2.12}$$

By (2.5),

$$K^{j_1 ... j_m} = \delta^{j_1 i_1}\, \delta^{j_2 i_2} ... \delta^{j_m i_m}\, K_{i_1 ... i_m}. \tag{2.13}$$

With these understandings, (2.3) may be put in the form:

$$S = \alpha_{i_1 ... i_m}\, K_{i_1 ... i_m}(y)\, W K^{i_1 ... i_m}(y)\, W, \tag{2.14}$$

which may be simplified further, thus:

$$S = \alpha_I\, K_I(y)\, W\, K^I(y)\, W, \tag{2.15}$$

where I denotes the m-fold suffix $i_1 ... i_m$.

A tensor T^I_J will be called **orthogonal** if

$$T^I_J\, T^I_K = \delta^J_K \tag{2.16}$$

If the index I is lowered, and the indexes J and K raised in (2.16), the following alternative form is obtained:

$$T_I^J \, T_I^K = \delta_J^K. \tag{2.17}$$

For a tensor T_J^I, T will denote its **contraction**:

$$T = T_I^I, \tag{2.18}$$

provided that there is no ambiguity. Note that T is a scalar since the double appearance of I in (2.18) implies that all suffixes are summed out.

3 Preliminary results

Proposition 3.1. For fixed m, let $W(x)$ be a polynomial of degree $\le m$ with coefficient λ_i of $(x^i)^m$, and choose α as follows:

$$\begin{aligned} \alpha_{i_1 \dots i_m} &= 1, \text{ if } i_1 = i_2 = \dots = i_m; \\ &= 0, \text{ otherwise.} \end{aligned} \tag{3.1}$$

Then

$$S_x = (m!)^2 \sum_i \lambda_i^2. \tag{3.2}$$

Proof. By direct evaluation from the definition of S_x. □

Proposition 3.2. With $W(x)$ defined as in Proposition 3.1 and $\mu \in L$ chosen arbitrarily, S_μ is invariant under transformations $x^i - \mu^i \to A_j^i (x^j - \mu^j)$ with A_j^i orthogonal.

Proof. See Appendix A. □

Proposition 3.3. Let A_J^I be an orthogonal tensor and let T_K^J be any tensor compatible with A_K^J for multiplication. Define

$$U_L^I = A_J^I \, T_K^J A_K^L. \tag{3.3}$$

Then

$$U = T.$$

Proof. $U = U_I^I = A_J^I \, T_K^J A_K^I = \delta_K^J \, T_K^J \text{ (by orthogonality)} = T_J^J = T.$ □

4 Smoothness

Consider the case $n = 1$, graduation in 1 dimension. In (2.3), $i_1, \ldots, i_m$ are each restricted to the set $\{0, 1\}$, giving

$$S_x = \sum_{p=0}^{m} \beta_p \left[\Delta^p \, W(x)\right]^2. \tag{4.1}$$

Typically, one chooses $\beta_p = 0, p < m$ and $\beta_m = 1$, to obtain

$$S_x = \left[\Delta^m \, W(x)\right]^2. \tag{4.2}$$

Now consider the case $n > 1$. By (4.2), smoothness measured just in the direction of the i -th axis may reasonably be measured by $\left[\Delta_i^m W(x)\right]^2$. A possibility for S_x is therefore some kind of average of these various directional differences over $i = 1, 2, \ldots, n$.

Moreover, it is desirable for S_x to be independent of the (Cartesian) coordinate system chosen at x, i.e. for S_x to be invariant under orthogonal transformations of coordinates with origin taken at x.

The restriction of $W(.)$ to the lattice domain L is inconvenient in the context of orthogonal transformations, since they do not fix L. For the purpose of the present section, it will be assumed that $W: R^n \to R$.

The question of invariance under orthogonal transformation is considered in Appendix B. In particular, Appendix B.2 constructs a smoothness measure as follows.

First choose n orthogonal axes for the n-dimensional space under consideration, with origin at fixed but arbitrary x. Choose ½ (n even) or ½($n-1$) (n odd) pairs of these axes to define planes of rotation. Without loss of generality, suppose that these are the x^1x^2-, x^3x^4-, *etc.* planes.

Now define the **directional difference operator** $\Delta_{q\theta}$ as the forward difference operator along the line obtained by rotating the x^{2q-1}-axis counter-clockwise through an angle θ about x in the - plane, i.e.

$$\Delta_{q\theta}\, W(x) = W(x^1,\ldots,x^{2q-1} + \cos\theta,\, x^{2q} + \sin\theta,\ldots,x^n) - W(x^1,\ldots x^n), \tag{4.3}$$

x^{2q-1} x^{2q} for an arbitrary function $W\colon R^n \to R$.

Take the quantity:

$$\sum_q \int_0^{2\pi} \left[\Delta_{q\theta_q}\, W(x)\right]^2 d\,\theta_q, \tag{4.4}$$

and define S_x as its symmetrization (symmetric sum or average) with respect to the choice of planes of rotation (still holding the choice of n coordinate axes fixed). Appendix B.2 shows this quantity to be invariant under orthogonal transformations of the n- space.

Appendix B.2 also establishes the following key result.

Proposition 4.1. Consider functions $W\colon R^n \to R$ which may be put in the following form by suitable choice of coordinate axes:

$$W(x) = \lambda_1(x^1)^m + \lambda_2(x^2)^m + \ldots + \lambda_n(x^n)^m + \text{other terms}, \tag{4.5}$$

where the "other terms" are of degree $<m$.

Let $y^1, ..., y^n$ be the coordinates obtained from $x^1, ..., x^n$ by an orthogonal transformation about a fixed but arbitrary point x. In the new coordinates, the function (4.5) takes the general form:

$$W(y) = v_{i_1 i_2 \ldots i_m} y^{i_1} y^{i_2} \ldots y^{i_m} + \text{terms of lesser degree,} \tag{4.6}$$

with each index i_r running from 1 to n.

For $m = 2p$, (4.6) may be put in the alternative form:

$$W(y) = v^{i_1 \ldots i_p}_{j_1 \ldots j_p} y^{i_1} \ldots y^{i_p} y^{j_1} \ldots y^{j_p} + \text{terms of lesser order.} \tag{4.7}$$

For such functions $W(.)$, S_x as defined above is (up to constant multipliers):

$$S_x = v_I v_I, \; m \text{ odd;} \tag{4.8}$$

$$= \beta_{mn} (v^I_I)^2 + (1 - \beta_{mn}) \, v^I_J \, v^J_I, \; m \text{ even,} \tag{4.9}$$

where

$$\beta_{mn} = 2^{n/2} \times \frac{(n/2)!}{n!} \binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m}, \; m, n \text{ even;}$$

$$= 2^{(n-1)/2} \times \frac{[(n-1)/2]!}{(n-1)!} \binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m}, \; m \text{ even, } n \text{ odd.} \tag{4.10}$$

The index I is of order m in (4.8) (see (4.6)), and the indexes I and J are each of order ½m in (4.9) (see (4.7)). □

Note that, for $W(.)$ given by (4.6),

$$v_{i_1 \ldots i_m} = \Delta_{i_1} \ldots \Delta_{i_m} W(y). \tag{4.11}$$

i.e.

$$v_I = \Delta_I^m W(y), \tag{4.12}$$

if Δ_I^m is written as an abbreviation for the operator in (4.11).

Similarly, when $W(.)$ takes the form (4.7),

$$v_J^I = \Delta_{IJ}^m W(y). \tag{4.13}$$

Substitution of (4.12) and (4.13) in (4.8) and (4.9) yields:

$$S_x = \left[\Delta_I^m W(x)\right]\left[\Delta_I^m W(x)\right], \; m \text{ odd}; \tag{4.14}$$

$$= \beta_{mn}\left[\Delta_{II}^m W(x)\right]^2 + (1-\beta_{mn})\left[\Delta_{IJ}^m W(x)\right]\left[\Delta_{JI}^m W(x)\right], \; m \text{ even}, \tag{4.15}$$

where the coordinate system y has been replaced by x in the argument of W, since S_x is independent of the choice of (orthogonal) coordinates.

This leads one to adopt (4.14) and (4.15) as the **definition** of the smoothness measure S_x.

Proposition 4.2. Let the smoothness measure S_x at x be defined by (4.14) and (4.15) for any function W: $R^n \to R$. Then, for the particular polynomial functions $W(.)$ discussed in Proposition 4.1, S_x reproduces the measure defined as the symmetrized (4.4). □

Example Consider the case $n = m = 2$. By (4.10), $\beta_{22} \equiv 1/3$, and so by (4.15),

$$S_x = \tfrac{1}{3}\left[\Delta_{11}^2 W(x) + \Delta_{22}^2 W(x)\right]^2$$

$$+ \tfrac{2}{3}\left\{\left[\Delta_{11}^2 W(x)\right]^2 + 2\left[\Delta_{12}^2 W(x)\right]^2 + \left[\Delta_{22}^2 W(x)\right]^2\right\}. \tag{4.16}$$

This value of S_x is inserted into the Whittaker criterion (2.4).

5 Bayesian interpretation

Taylor (1992) provided the following interpretation of Whittaker graduation in the case $n = 1$.

Suppose that

$$f(x) \sim N\left(W(x),\ \sigma^2/w(x)\right), \tag{5.1}$$

$$\Delta^m W(x) \sim N(0,\ \tau^2). \tag{5.2}$$

Then the Bayesian estimate of $W(x)$, conditional on $f(x)$, is given by Whittaker graduation with relativity constant

$$k = \sigma^2/\tau^2. \tag{5.3}$$

The arguments extend to the multi-dimensional case as follows.

In the case $W\colon L \to R$, define $\Delta^m W(x)$ as the vector of differences $\Delta_I^m W(x)$, without repetition.

Proposition 5.1. Let (5.1) hold in the case $W\colon L \to R$. Suppose that, for a suitable choice of orthogonal coordinates,

$$\Delta^m W(x) \sim N(0, \tau^2 I). \tag{5.4}$$

Then the Bayesian estimate of $W(x)$, conditional on $\{f(x)\}$, is given by Whittaker graduation with relativity constant $k = \sigma^2/n^m\ \tau^2$. □

The factor n^m which appears in the denominator of k is the sum of the coefficients of squared differences in (4.14) or (4.15).

Verrall (1993) re-interpreted Taylor's results (5.1) - (5.3) in terms of dynamic linear models (DLMs) and Kalman filtering. These ideas also extend to the multi-dimensional case.

Proposition 5.2. Let (5.1) hold in the case $W: L \to R$. Choose any system of orthogonal coordinates, and let $e_1, \ldots, e_n$ denote the natural basis vectors for that system.

Suppose that the values of $\Delta_j W(x)$ are determined by the following stochastic recursion:

$$\left[\epsilon_{I_1}^{(m)}(x),\ \epsilon_{I_2}^{(m)}(x),\ \ldots\right] \sim N(0,\tau^2 I), \tag{5.5}$$

for each x, with I_1, I_2, etc running over all m-fold indexes $i_1 \ldots i_m$ such that $i_1 \le i_2 \le \ldots, \le i_m$;

$$\epsilon_I^{(k)}(x + e_j) = \epsilon_I^{(k)}(x) + \epsilon_{Ij}^{(k+1)}(x),\ j = 1, \ldots, n;\ k = 1, \ldots, m - 1, \tag{5.6}$$

for each x, and with I denoting any k-fold index;

$$\Delta_j\ W(x) = \epsilon_j^{(1)}(x),\ j = 1, \ldots, n, \tag{5.7}$$

for each x.

Then the Bayesian estimate of $W(x)$ conditional on $\{f(x)\}$, is given by Whittaker graduation with relativity constant:

$$k = \sigma^2 / n^m\ \tau^2$$

□

According to (5.7), the gradients $\Delta_j W$ vary from point to point according to the DLM defined by (5.5) and (5.6).

Appendix A: Proof of Proposition 3.2

Let

$$z^i = \mu^i + A^i_j(y^j - \mu^j), \tag{A.1}$$

$$\bar{W}(y) = W(z). \tag{A.2}$$

Let $\bar{S}$ be defined as for S in (2.15) but with W replaced by $\bar{W}$ and subject to the restriction on α in Proposition 3.1.

By this last restriction,

$$\bar{S} = K_I(y)\,\bar{W}\,K^I(y)\,\bar{W}, \tag{A.3}$$

with I restricted to index sets of the form $ii...i$ (m times).

By (A.2),

$$K_I(y)\,\bar{W} = K_I(z)\,W, \tag{A.4}$$

where

$$W(z) = \lambda_i\left[\mu^i + A^i_j(y^j - \mu^j)\right]^m + p(y) \tag{A.5}$$

with $p(y)$ a polynomial of degree $<m$. Thus

$$W(z) = \lambda_i (A^i_j)^m (y^j)^m + \text{other terms}. \tag{A.6}$$

The other terms here involve powers of y^j all strictly less than m. The m-th differencing operator $K_I(z)$ therefore eliminates these terms. Then (A.4) and (A.6) yield:

$$K_I(y)\,\bar{W} = K_I(z)\left[\lambda_k\,(A^k_j)^m (y^j)^m\right]$$

$$= m!\;\delta^j_i\;\lambda_k (A^k_j)^m. \tag{A.7}$$

By (2.13), **without** summation over y,

$$K_I K^I = K_I \delta^{il} K_L \qquad [L = ll\ldots l]. \tag{A.8}$$

Substitution of (A.7) and (A.8) into (A.3) yields:

$$\begin{aligned} \bar{S}_\mu &= (m!)^2 \delta^j_i \lambda_k (A^k_j)^m \delta^{il} \delta^h_l \lambda_g (A^g_h)^m \\ &= (m!)^2 \delta^{jh} \lambda_k \lambda_g (A^k_j A^g_h)^m \\ &= (m!)^2 \lambda_k \lambda_g (A^k_j A^g_j)^m \\ &= (m!)^2 \lambda_k \lambda_g \delta^k_g \qquad [\text{by orthogonality of A}] \\ &= m! \sum_k \lambda_k^2, \end{aligned}$$

which is the same as (3.2). □

Appendix B: Directional differences

B.1 2-dimensional space

Consider the case $n = 2$. Define $W(x)$ as the following special case of Proposition 3.1:

$$W(x) = \lambda_1 (x^1)^m + \lambda_2 (x^2)^m + \text{other terms}, \tag{B.1}$$

where the "other terms" are of degree $<m$.

Now define the directional difference operator Δ_θ as the difference operator along the line obtained by rotating the x^1-axis counter-clockwise through an angle θ about x. That is,

$$\Delta_\theta W(x) = W(x^1 + \cos\theta, x^2 + \sin\theta) - W(x^1, x^2). \tag{B.2}$$

It may be checked that

$$\begin{aligned} \Delta_\theta^m W(x) &= \Delta_\theta^m \left[\lambda_i (x^1)^m + \lambda_2 (x^2)^m\right] \\ &= m! \left[\lambda_1 \cos^m\theta + \lambda_2 \sin^m\theta\right]. \end{aligned} \tag{B.3}$$

Then

$$\begin{aligned} \left[\Delta_\theta^m \; W(x)\right]^2 &= (m!)^2 \, (\lambda_1^2 \cos^{2m}\theta + \lambda_2^2 \sin^{2m}\theta + 2\lambda_1 \lambda_2 \sin^m\theta \cos^m\theta \\ &= (½)^{2m} (m!)^2 \left[\lambda_1^2 \, (e^{i\theta} + e^{-i\theta})^{2m} + \lambda_2^2 \, (-i)^{2m} \, (e^{i\theta} - e^{-i\theta})^{2m} \right. \\ &\quad \left. + \, 2\lambda_1 \lambda_2 (-i)^m \, (e^{i\theta} + e^{-i\theta})^m \, (e^{i\theta} - e^{-i\theta})^m\right]. \end{aligned} \tag{B.4}$$

If the terms on the right side of this result are expanded, they yield a constant term and a number of periodic terms. Evaluation of the constant term lead to the following results.

If m is odd,

$$[\Delta^m W(x)]^2 = (\tfrac{1}{2})^{2m} (m!)^2 \binom{2m}{m} (\lambda_1^2 + \lambda_2^2) + \text{periodic terms.} \tag{B.5}$$

If m is even,

$$[\Delta_\theta^m W(x)]^2 = (\tfrac{1}{2})^{2m} (m!)^2 \left[\binom{2m}{m} (\lambda_1^2 + \lambda_2^2) + 2\binom{m}{\tfrac{1}{2}m} \lambda_1 \lambda_2\right] + \text{periodic terms.} \tag{B.6}$$

If (B.5) and (B.6) are integrated over θ, the periodic terms have no effect. Then

$$\int_0^{2\pi}[\Delta_\theta^m W(x)]^2 d\theta = \text{const.} \times (\lambda_1^2 + \lambda_2^2), \quad m \text{ odd};$$

$$= \text{const.} \times \left[(\lambda_1^2 + \lambda_2^2) + 2\lambda_1\lambda_2\binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m}\right], \quad m \text{ ev}$$

$$= \text{const.} \times \left\{(\lambda_1 + \lambda_2)^2 \binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m} + (\lambda_1^2 + \lambda_2^2)\left[1 - \binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m}\right]\right\}, \quad m \text{ even.} \tag{B.7}$$

It is necessary to consider how (B.7) changes under orthogonal transformations of the coordinates. This is most easily done by considering separately the cases of m odd and even.

Case I: m odd

First, rewrite (B.1) as:

$$W(x) = \lambda_{i_1 i_2 \ldots i_m} x^{i_1} x^{i_2} \ldots x^{i_m}, \tag{B.8}$$

with each $i_r = 1,2$, and

$$\lambda_{11\ldots1} = \lambda_1, \lambda_{22\ldots2} = \lambda_2,$$

$$\lambda_{i_1 i_2 \ldots i_m} = 0, \text{ otherwise.} \tag{B.9}$$

Thus, in terms of the contraction notation (2.18),

$$\int_0^{2\pi} [\Delta_\theta^m \; W(x)]^2 d\theta = \text{const.} \times \lambda_I \lambda_I. \tag{B.10}$$

Now consider an orthogonal transformation of x to y, with

$$y^i = A_j^i \, x^j. \tag{B.11}$$

Then

$$A_k^i \, y^i = A_k^i \, A_j^i \, x^j = x^k, \tag{B.12}$$

by (2.16).

Substitution of (B.12) in (B.8) gives:

$$W(x) = v_{j_1 \ldots j_m} \, y^{j_1 \ldots j_m}, \tag{B.13}$$

with

$$v_{j_1 \ldots j_m} = \lambda_{i_1 \ldots i_m} \, A_{i_1}^{j_1} \ldots A_{i_m}^{j_m}. \tag{B.14}$$

i.e.

$$v_J = \lambda_I \, A_I^J. \tag{B.15}$$

Then

$$v_J v_K = A_I^J \, \lambda_I \, \lambda_L \, A_L^K. \tag{B.16}$$

Apply Proposition 3.3 to the tensor $\lambda_I \, \lambda^L$ $(= \lambda_I \, \lambda_L)$ to show that

$$v_J v_J = \lambda_I \, \lambda_I \tag{B.17}$$

This shows that, for any $W(y)$ of the form (B.13), obtainable from (B.8) and (B.9) by an orthogonal transformation from x to y, the quantity $v_J\ v_J$ is invariant under orthogonal transformations. Further, by (B.10), this quantity measures the mean square m-th difference appearing there.

Case II: m even

The same reasoning as in Case I can be applied. However, it is more useful to proceed somewhat differently in order to deal with the additional term in (B.7) involving $(\lambda_1 + \lambda_2)$.

Let $m = 2p$, and in place of (B.8) write

$$W(x) = \lambda^{i_1 \ldots i_p}_{j_1 \ldots j_p}\, x^{i_1} \ldots x^{i_p}\, x^{j_1} \ldots x^{j_p}, \tag{B.18}$$

with each $i_r, j_r = 1, 2$, and

$$\lambda^{11\ldots1}_{11\ldots1} = \lambda_1,\ \lambda^{22\ldots2}_{22\ldots2} = \lambda_2,$$

$$\lambda^{i_1 \ldots i_p}_{j_1 \ldots j_p} = 0, \text{ otherwise.} \tag{B.19}$$

Then

$$\int_0^{2\pi} [\Delta_\theta^m\ W(x)]^2 d\theta = \text{const.} \times \left\{ (\lambda_I^I)^2 \binom{m}{\tfrac{1}{2}m} \Big/ \binom{m}{2m} + \lambda_J^I\, \lambda_I^J \left[1 - \binom{m}{\tfrac{1}{2}m} \Big/ \binom{m}{2m} \right] \right\}. \tag{B.20}$$

Parallel to Case I, it is possible to show that, for any $W(y)$ of the form:

$$W(y) = v^{i_1 \ldots i_p}_{j_1 \ldots j_p}\, y^{i_1} \ldots y^{i_p}\, y^{j_1} \ldots y^{j_p}, \tag{B.21}$$

obtainable from (B.18) and (B.19) by an orthogonal transformation from x to y, the quantities v_I^I and $v_J^I\ v_I^J$ are invariant under orthogonal transformations. Further, by (B.20), the quantity:

$$(v_I^J)^2 \binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m} + v_J^I\, v_I^J \left[1 - \binom{m}{\tfrac{1}{2}m} \Big/ \binom{2m}{m}\right] \tag{B.22}$$

measures the mean square m-th difference appearing there.

B.2 n-dimensional space

It is known (e.g. Boerner, 1963, 221-2) that any orthogonal transformation in n-dimensional space is orthogonally similar to (i.e. can, with appropriate orthogonal change of coordinates, be represented as) $\tfrac{1}{2}n$ (n even) or $\tfrac{1}{2}(n-1)$ (n odd) rotations of orthogonal 2-dimensional subspaces.

Consider the following generalization in (B.1):

$$W(x) = \lambda_1 (x^1)^m + \ldots + \lambda_n (x^n)^m + \text{other terms}, \tag{B.23}$$

where the "other terms" are still of degree $<m$.

Note that (B.23) can be written as (B.8) (m odd) or (B.18) (m even) with each $i_r = 1,2,\ldots,n$, and

$$\lambda_{ii\ldots i} = \lambda_i,\ \lambda_{i_1 i_2 \ldots i_n} = 0 \text{ otherwise.} \tag{B.24}$$

Define the directional difference operator $\Delta_{q\theta}$ as the difference operator along the line obtained by rotating the x^{2q-1}-axis counter-clockwise through an angle θ about x in the $x^{2q-1}x^{2q}$-plane.

Consider how the quantity on the left side of (B.7) generalises to n dimensions. One may take

$$\sum_q \int_0^{2\pi} [\Delta^m_{q\theta_q} W(x)]^2 \, d\,\theta_q, \tag{B.25}$$

with q running from 1 to $\tfrac{1}{2}n$ (n even) or $\tfrac{1}{2}(n-1)$ (n odd). Then define $\alpha(W)$ as the average of (B.25) taken over all permutations of the n coordinate axes. This last step symmetrises (B.25) with respect to the choice of planes of rotation.

Following a development parallel to that leading from (B.1) to (B.7), one finds that

$$\alpha(W) = \text{const.} \times (\lambda_1^2 + \ldots + \lambda_n^2), \quad m \text{ odd}; \tag{B.26}$$

$$= \text{const.} \times \left[\frac{n!}{(n/2)!}(\tfrac{1}{2})^{n/2}(\lambda_1^2 + \ldots + \lambda_n^2) + 2\sum_{i>j}\lambda_i\lambda_j\binom{m}{\tfrac{1}{2}m}\bigg/\binom{2m}{m}\right], \quad m, n \text{ even}; \tag{B.27}$$

$$= \text{const.} \times \left[\frac{(n-1)!}{[(n-1)/2]!}(\tfrac{1}{2})^{(n-1)/2}(\lambda_1^2 + \ldots + \lambda_n^2) + 2\sum_{i>j}\lambda_i\lambda_j\binom{m}{\tfrac{1}{2}m}\bigg/\binom{2m}{m}\right], \quad m \text{ even}, n \text{ odd}. \tag{B.28}$$

The factor $n!/(\tfrac{1}{2}n)! \times (\tfrac{1}{2})^{\frac{1}{2}n}$ in (B.27) is the number of choices of $\tfrac{1}{2}n$ planes of rotation in n-dimensional space. The corresponding factor in (B.28) is $n!/[\tfrac{1}{2}(n-1)! \times (\tfrac{1}{2})^{(n-1)/2}$, but each such choice leads to the appearance of only $n-1$ out of n of the λ_i^2 terms. Hence the factor of $(\lambda_1^2 + \ldots + \lambda_n^2)$ in (B.28).

Now (B.26) - (B.28) may be consolidated as follows:

$$\alpha(W) = \text{const.} \times \left[(\lambda_1^2 + \ldots + \lambda_n^2) + 2\beta_{mn}\sum_{i>j}\lambda_i\lambda_j\right], \tag{B.29}$$

with

$$\beta_{mn} = 0,\ m \text{ odd};$$

$$= 2^{n/2} \times \frac{(n/2)!}{n!} \binom{m}{\frac{1}{2}m} \Big/ \binom{2m}{m},\ m\ n \text{ even};$$

$$= 2^{(n-1)/2} \times \frac{[(n-1)/2]!}{(n-1)!} \binom{m}{\frac{1}{2}m} \Big/ \binom{2m}{m},\ m \text{ even, } n \text{ odd.} \tag{B.30}$$

Then $\alpha(W)$ may be put in the form:

$$\alpha(W) = \text{const.} \times \left[\beta_{mn}(\lambda_1 + \ldots + \lambda_n)^2 + (1-\beta_{mn})(\lambda_1^2 + \ldots + \lambda_n^2)\right]. \tag{B.31}$$

As in Appendix B1, consider orthogonal transformations (B.11), converting $W(x)$ to the form (B.13) (m odd) or (B.21) (m even). By the same reasoning as there, the quantities $v_I\ v_I$ (m odd) or v_I^I and $v_J^I\ v_I^J$ (m even) are invariant under orthogonal transformations.

Then the mean square m-th difference $\alpha(W)$ is measured by:

$$v_I v_I,\ m \text{ odd,} \tag{B.32}$$

or

$$\beta_{mn}(v_I^I)^2 + (1-\beta_{mn})\ v_J^I\ v_I^J,\ m \text{ even.} \tag{B.33}$$

References

Boerner, H. (1963). **Representations of groups.** North Holland Publishing Company, Amsterdam.

Broffitt, J.D. (1996). On smoothness terms in multidimensional Whittaker graduation. **Insurance: mathematics and economics**, 18, 13-27.

Henderson, R. (1924). A new method of graduation. **Transactions of the Actuarial Society of America**, 25, 29-40.

Knorr, F.E. (1984). Multidimensional Whittaker-Henderson graduation. **Transactions of the Society of Actuaries**, 36, 213-240.

London, R. (1985). **Graduation: the revision of estimates.** ACTEX, Winsted and Abington, CT.

Lowrie, W.B. (1993). Multidimensional Whittaker-Henderson graduation with constraints and mixed differences. **Transactions of the Society of Actuaries**, 45, 215-255.

McKay, S.F. and Wilkin, J.C. (1977). Derivation of a two-dimensional Whittaker-Henderson type B graduation formula. Appendix to **Experience of Disabled-Worker Benefits Under OASDI, 1965-74.** Actuarial Study No. 74, U.S. Department of Health, Education, and Welfare.

Taylor, G. (1992). A Bayesian interpretation of Whittaker-Henderson graduation. **Insurance: mathematics and economics**, 11, 7-16.

Verrall, R.J. (1993). A state space formulation of Whittaker graduation, with extensions. **Insurance: mathematics and economics**, 13, 7-14.

Whittaker, E.T. (1923). On a new method of graduation. **Proceedings of the Edinburgh Mathematical Society**, 41, 63-75.

RESEARCH PAPER SERIES

No.	Date	Subject	Author
1	MAR 93	AUSTRALIAN SUPERANNUATION : THE FACTS, THE FICTION, THE FUTURE	David M Knox
2	APR 93	AN EXPONENTIAL BOUND FOR RUIN PROBABILITIES	David C M Dickson
3	APR 93	SOME COMMENTS ON THE COMPOUND BINOMIAL MODEL	David C M Dickson
4	AUG 93	RUIN PROBLEMS AND DUAL EVENTS	David CM Dickson Alfredo D Egidio dos Reis
5	SEP 93	CONTEMPORARY ISSUES IN AUSTRALIAN SUPERANNUATION - A CONFERENCE SUMMARY	David M Knox John Piggott
6	SEP 93	AN ANALYSIS OF THE EQUITY INVESTMENTS OF AUSTRALIAN SUPERANNUATION FUNDS	David M Knox
7	OCT 93	A CRITIQUE OF DEFINED CONTRIBUTION USING A SIMULATION APPROACH	David M Knox
8	JAN 94	REINSURANCE AND RUIN	David C M Dickson Howard R Waters
9	MAR 94	LIFETIME INCOME, TAXATION, EXPENDITURE AND SUPERANNUATION (LITES): A LIFE-CYCLE SIMULATION MODEL	Margaret E Atkinson John Creedy David M Knox
10	FEB 94	SUPERANNUATION FUNDS AND THE PROVISION OF DEVELOPMENT/VENTURE CAPITAL: THE PERFECT MATCH? YES OR NO	David M Knox
11	JUNE 94	RUIN PROBLEMS: SIMULATION OR CALCULATION?	David C M Dickson Howard R Waters
12	JUNE 94	THE RELATIONSHIP BETWEEN THE AGE PENSION AND SUPERANNUATION BENEFITS, PARTICULARLY FOR WOMEN	David M Knox
13	JUNE 94	THE COST AND EQUITY IMPLICATIONS OF THE INSTITUTE OF ACTUARIES OF AUSTRALIA PROPOSED RETIREMENT INCOMES STRATEGY	Margaret E Atkinson John Creedy David M Knox Chris Haberecht
14	SEPT 94	PROBLEMS AND PROSPECTS FOR THE LIFE INSURANCE AND PENSIONS SECTOR IN INDONESIA	Catherine Prime David M Knox

15	OCT 94	PRESENT PROBLEMS AND PROSPECTIVE PRESSURES IN AUSTRALIA'S SUPERANNUATION SYSTEM	David M Knox
16	DEC 94	PLANNING RETIREMENT INCOME IN AUSTRALIA: ROUTES THROUGH THE MAZE	Margaret E Atkinson John Creedy David M Knox
17	JAN 95	ON THE DISTRIBUTION OF THE DURATION OF NEGATIVE SURPLUS	David C M Dickson Alfredo D Egidio dos Reis
18	FEB 95	OUTSTANDING CLAIM LIABILITIES: ARE THEY PREDICTABLE?	Ben Zehnwirth
19	MAY 95	SOME STABLE ALGORITHMS IN RUIN THEORY AND THEIR APPLICATIONS	David C M Dickson Alfredo D Egidio dos Reis Howard R Waters
20	JUN 95	SOME FINANCIAL CONSEQUENCES OF THE SIZE OF AUSTRALIA'S SUPERANNUATION INDUSTRY IN THE NEXT THREE DECADES	David M Knox
21	JUN 95	MODELLING OPTIMAL RETIREMENT IN DECISIONS IN AUSTRALIA	Margaret E Atkinson John Creedy
22	JUN 95	AN EQUITY ANALYSIS OF SOME RADICAL SUGGESTIONS FOR AUSTRALIA'S RETIREMENT INCOME SYSTEM	Margaret E Atkinson John Creedy David M Knox
23	SEP 95	EARLY RETIREMENT AND THE OPTIMAL RETIREMENT AGE	Angela Ryan
24	OCT 95	APPROXIMATE CALCULATION OF MOMENTS OF RUIN RELATED DISTRIBUTIONS	David C M Dickson
25	DEC 95	CONTEMPORARY ISSUES IN THE ONGOING REFORM OF THE AUSTRALIAN RETIREMENT INCOME SYSTEM	David M Knox
26	FEB 96	THE CHOICE OF EARLY RETIREMENT AGE AND THE AUSTRALIAN SUPERANNUATION SYSTEM	Margaret E Atkinson John Creedy
27	FEB 96	PREDICTIVE AGGREGATE CLAIMS DISTRIBUTIONS	David C M Dickson Ben Zehnwirth
28	FEB 96	THE AUSTRALIAN GOVERNMENT SUPERANNUATION CO-CONTRIBUTIONS: ANALYSIS AND COMPARISON	Margaret Atkinson
29	MAR 96	A SURVEY OF VALUATION ASSUMPTIONS AND FUNDING METHODS USED BY AUSTRALIAN ACTUARIES IN DEFINED BENEFIT SUPERANNUATION FUND VALUATIONS	Des Welch Shauna Ferris
30	MAR 96	THE EFFECT OF INTEREST ON NEGATIVE SURPLUS	David C M Dickson Alfred D Egídio dos Reis

31	MAR 96	RESERVING CONSECUTIVE LAYERS OF INWARDS EXCESS-OF-LOSS REINSURANCE	Greg Taylor
32	AUG 96	EFFECTIVE AND ETHICAL INSTITUTIONAL INVESTMENT	Anthony Asher
33	AUG 96	STOCHASTIC INVESTMENT MODELS: UNIT ROOTS, COINTEGRATION, STATE SPACE AND GARCH MODELS FOR AUSTRALIA	Michael Sherris Leanna Tedesco Ben Zehnwirth
34	AUG 96	THREE POWERFUL DIAGNOSTIC MODELS FOR LOSS RESERVING	Ben Zehnwirth
35	SEPT 96	KALMAN FILTERS WITH APPLICATIONS TO LOSS RESERVING	Ben Zehnwirth
36	OCT 96	RELATIVE REINSURANCE RETENTION LEVELS	David C M Dickson Howard R Waters
37	OCT 96	SMOOTHNESS CRITERIA FOR MULTI-DIMENSIONAL WHITTAKER GRADUATION	Greg Taylor