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

$$x = (x^1, x^2, ..., x^n),$$

(2.1)

with

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

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

Let $f: L \rightarrow \mathbb{R}$, and call the $f(x)$ observations.

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

Define

$$E = \sum w(x)[f(x) - W(x)]^2,$$

(2.2)

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

$$S = \sum_{x} \sum_{i_{1}, \ldots , i_{m}} \alpha_{i_{1}, \ldots , i_{m}} \left[ \Delta_{i_{1}} \Delta_{i_{2}} \ldots \Delta_{i_{m}} W(x) \right]^{2},$$

(2.3)

where the $\alpha_{i_{1}, \ldots , i_{m}}$ are real constants and $\Delta_{i}$ is the $i$-th 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 $\Delta_{0}$ 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

$$F = E + kS,$$  

(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 $\alpha$'s.

2.2 Tensor notation

Tensor notation is the natural vehicle for the multi-dimensional framework considered here.
The Euclidean metric tensor is

\[ \delta_{ij} = \begin{cases} 
1 & \text{if } i = j; \\
0 & \text{otherwise},
\end{cases} \]

for \( i, j = 1, 2, \ldots, n \).

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 \( B^i \),

\[ A_{ij} B^i = \sum_j A_{ij} B^i. \]

The metric tensor effects raising and lowering of suffixes, e.g., for a tensor \( T_{klm} \),

\[ \delta_{hi} T_{klm} = T_{klm}. \tag{2.5} \]

In particular,

\[ \delta_{ij}^{i'} = \delta_{ij}^{ik} \delta_{kj}^{i'} \\
= 1 \text{ if } i = j \tag{2.6} \\
= 0 \text{ otherwise.} \]

Define

\[ \delta_{i_1 \ldots i_p}^{i_1' \ldots i_p'} = \delta_{i_1}^{i_1'} \delta_{i_2}^{i_2'} \ldots \delta_{i_p}^{i_p'}. \tag{2.7} \]

With the summation convention, (2.2) becomes

\[ E = \left[ f(x^1, \ldots, x^n) - W(x^1, \ldots, x^n) \right] w(x^1, \ldots, x^n, y^1, \ldots, y^n) \\
\times \left[ f(y^1, \ldots, y^n) - W(y^1, \ldots, y^n) \right], \tag{2.8} \]

where \( w(x^1, \ldots, y^n) \) is the tensor given by

\[ w(x, y) = w(x) \text{ for } x = y, \]
\[ = 0, \text{ otherwise.} \tag{2.9} \]
It is convenient to write (2.8) in the simple form:

\[ E = \left[f(x) - W(x)\right] w(x, y) \left[f(y) - W(y)\right], \]  

(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 \( y \),

\[ \Delta_j(y) \ W(x) = W(y^1, ..., y^{i+1}, ..., y^n) - W(y^1, ..., y^i, ..., y^n). \]  

(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

\[ K_{i_1...i_m}(y) = \Delta_{i_1}(y) ... \Delta_{i_m}(y) \quad [\text{no summation over } y]. \]  

(2.12)

By (2.5),

\[ K^{i_1...i_m} = \delta^{i_1}_{j_1} \delta^{i_2}_{j_2} ... \delta^{i_m}_{j_m} K_{j_1...j_m}. \]  

(2.13)

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

\[ S = \alpha_{i_1...i_m} K_{i_1...i_m}(y) W K^{i_1...i_m}(y) W, \]  

(2.14)

which may be simplified further, thus:

\[ S = \alpha_j K_j(y) W K^j(y) W, \]  

(2.15)

where \( I \) denotes the \( m \)-fold suffix \( i_1...i_m \).

A tensor \( T^I_j \) will be called orthogonal if

\[ T^I_j T^K_l = \delta^K_l. \]  

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

$$T^I_I T^K_K = \delta^K_J. \quad (2.17)$$

For a tensor $T^J_J$, $T$ will denote its contraction:

$$T = T^J_J, \quad (2.18)$$

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 $\leq m$ with coefficient $\lambda_i$ of $(x^i)^m$, and choose $\alpha$ as follows:

$$\alpha_{i_1 \ldots i_m} = 1, \text{ if } i_1 = i_2 = \ldots = i_m;$$

$$= 0, \text{ otherwise.} \quad (3.1)$$

Then

$$S_x = (m!)^2 \sum \lambda_i^2. \quad (3.2)$$

Proof. By direct evaluation from the definition of $S_x$. \hfill \Box

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

Proof. See Appendix A. \hfill \Box

Proposition 3.3. Let $A^I_J$ be an orthogonal tensor and let $T^K_K$ be any tensor compatible
with $A^L_K$ for multiplication. Define

$$U^I_L = A^I_J T^J_K A^L_K.$$ (3.3)

Then

$$U = T.$$

Proof. \[ U = U^I_L = A^I_J T^J_K A^L_K = \delta^I_K T^I_K \text{ (by orthogonality)} = T^I_I = T. \]

\[ \square \]

4 Smoothness

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

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

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

$$S_x = [\Delta^m W(x)]^2.$$ (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 $[\Delta^m_i W(x)]^2$. A possibility for $S_x$ is therefore some kind of average of these various directional differences over $i = 1, 2, ..., n$.

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 $W: \mathbb{R}^n \to \mathbb{R}$. 

30/09/96 04:48PM S:\GAMSONM\GENERAL\G17.66
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 \( \frac{1}{2} (n \text{ even}) \) or \( \frac{1}{2} (n-1) \) (odd) pairs of these axes to define planes of rotation. Without loss of generality, suppose that these are the \( x^1 x^2 \), \( x^3 x^4 \), etc. planes.

Now define the directional difference operator \( \Delta_{\theta_0} \) as the forward difference operator along the line obtained by rotating the \( x^{2q-1} \)-axis counter-clockwise through an angle \( \theta \) about \( x \) in the \( x^q \) plane, i.e.

\[
\Delta_{\theta_0} W(x) = W(x^1, ..., x^{2q-1} + \cos \theta, x^{2q} + \sin \theta, ..., x^n) - W(x^1, ..., x^n),
\]

\( x^{2q-1}, x^{2q} \) for an arbitrary function \( W: \mathbb{R}^n \rightarrow \mathbb{R} \).

Take the quantity:

\[
\sum_q \int_0^{2\pi} |\Delta_{\theta_0} W(x)|^2 \ d \theta
\]

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 \( W: \mathbb{R}^n \rightarrow \mathbb{R} \) which may be put in the following form by suitable choice of coordinate axes:

\[
W(x) = \lambda_1 (x^1)^m + \lambda_2 (x^2)^m + ... + \lambda_n (x^n)^m + \text{other terms},
\]
where the "other terms" are of degree \(< m\).

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

\[
W(y) = v_{i_1i_2\ldots i_m} y^{i_1}y^{i_2}\ldots y^{i_m} + \text{terms of lesser degree},
\]

(4.6)

with each index \(i_r\) running from 1 to \(n\).

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

\[
W(y) = v_{i_j\ldots i_j} y^{i_1}y^{i_2}\ldots y^{i_p}y^{j_1}\ldots y^{j_p} + \text{terms of lesser order}.
\]

(4.7)

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

\[
S_x = v_I v_J m \text{ odd};
\]

(4.8)

\[
= \beta_{mn} (v_I^J)^2 + (1 - \beta_{mn}) v_I^J v_J^I m \text{ even},
\]

(4.9)

where

\[
\beta_{mn} = 2^{n/2} \times \frac{(n/2)!}{n!} \left( \begin{array}{c} m \\ m/2 \end{array} \right)/\left( \begin{array}{c} 2m \\ m \end{array} \right), m, n \text{ even};
\]

\[
= 2^{(n-1)/2} \times \frac{[(n-1)/2]!}{(n-1)!} \left( \begin{array}{c} m \\ m/2 \end{array} \right)/\left( \begin{array}{c} 2m \\ m \end{array} \right), m \text{ even, } n \text{ odd}.
\]

(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 \(\sqrt{m}\) in (4.9) (see (4.7)).

Note that, for \(W()\) given by (4.6),

\[
v_{i_1\ldots i_m} = \Delta_{i_1} \ldots \Delta_{i_m} W(y).
\]

(4.11)

i.e.
\[ v_I = \Delta_I^m W(y), \quad (4.12) \]

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

Similarly, when \( W(\cdot) \) takes the form (4.7),
\[ v_j^I = \Delta_I^m W(y). \quad (4.13) \]

Substitution of (4.12) and (4.13) in (4.8) and (4.9) yields:
\[ S_x = [\Delta_I^m W(x)] [\Delta_I^m W(x)], \quad m \text{ odd}; \quad (4.14) \]
\[ = \beta_{mm} [\Delta_H^m W(x)]^2 + (1 - \beta_{mm}) [\Delta_H^m W(x)] [\Delta_H^m W(x)], \quad m \text{ even}, \quad (4.15) \]

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 \( W: \mathbb{R}^n \to \mathbb{R} \). Then, for the particular polynomial functions \( W(\cdot) \) discussed in Proposition 4.1, \( S_x \) reproduces the measure defined as the symmetrized (4.4). \( \square \)

**Example** Consider the case \( n = m = 2 \). By (4.10), \( \beta_{22} = \frac{1}{2} \), and so by (4.15),
\[ S_x = \frac{1}{2} [\Delta_{11}^2 W(x) + \Delta_{22}^2 W(x)]^2 \]
\[ + \frac{1}{2} \left( [\Delta_{11}^2 W(x)]^2 + 2[\Delta_{12}^2 W(x)]^2 + [\Delta_{22}^2 W(x)]^2 \right). \quad (4.16) \]

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
\[
  f(x) \sim N(W(x), \sigma^2/w(x)),
\]
\[
  \Delta^m W(x) \sim N(0, \tau^2).
\]

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

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

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

**Proposition 5.1.** Let (5.1) hold in the case $W: L \to R$. Suppose that, for a suitable choice of orthogonal coordinates,
\[
  \Delta^m W(x) \sim N(0, \tau^2 I).
\]

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 \( W: L \rightarrow R \). Choose any system of orthogonal coordinates, and let \( \epsilon_1, ..., \epsilon_n \) denote the natural basis vectors for that system.

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

\[
[\epsilon^{(m)}_{I_1}(x), \epsilon^{(m)}_{I_2}(x), ...] \sim N(0, \tau^2 I),
\]

(5.5)

for each \( x \), with \( I_1, I_2, \) etc running over all \( m \)-fold indexes \( i_1, ..., i_m \) such that \( i_1 \leq i_2 \leq ..., \leq i_m \);

\[
\epsilon^{(k)}_{I}(x + \epsilon_j) = \epsilon^{(k)}_{I}(x) + \epsilon^{(k+1)}_{I}(x), \quad j = 1, ..., n; \quad k = 1, ..., m - 1,
\]

(5.6)

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

\[
\Delta_j W(x) = \epsilon_j^{(n)}(x), \quad j = 1, ..., n,
\]

(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
\[ z^i = \mu^i + A^i_j (y^j - \mu^j), \]  \hfill (A.1)
\[ \bar{W}(y) = W(z). \]  \hfill (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 \( \alpha \) in Proposition 3.1.

By this last restriction,
\[ \bar{S} = K_I(y) \bar{W} K^I(y) \bar{W}, \]  \hfill (A.3)
with \( I \) restricted to index sets of the form \( ii \ldots i \) (\( m \) times).

By (A.2),
\[ K_I(y) \bar{W} = K_I(z) W, \]  \hfill (A.4)
where
\[ W(z) = \lambda_i [\mu^i + A^i_j (y^j - \mu^j)]^m + \rho(y) \]  \hfill (A.5)
with \( \rho(y) \) a polynomial of degree <\( m \). Thus
\[ W(z) = \lambda_i (A^i_j)^m (y^j)^m + \text{other terms}. \]  \hfill (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:
\[ K_I(y) \bar{W} = K_I(z) \left[ \lambda_i (A^i_j)^m (y^j)^m \right] \]
\[ = m! \delta_i^i \lambda_i (A^i_j)^m. \]  \hfill (A.7)
By (2.13), without summation over $y$,

$$K_I K^I = K_I \delta^{ij} K_L \quad [L = II...I] \quad (A.8)$$

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

$$\tilde{S}_\mu = (m!)^2 \delta^i_j \lambda_k \lambda_g (A^i_j)^m \delta^{ij} \delta^h_j \lambda_g (A^g_h)^m$$

$$= (m!)^2 \delta^{ih} \lambda_k \lambda_g (A^i_j A^g_h)^m$$

$$= (m!)^2 \lambda_k \lambda_g (A^i_j A^g_h)^m$$

$$= (m!)^2 \lambda_k \lambda_g \delta^i_i \quad [\text{by orthogonality of } A]$$

$$= m! \sum_k \lambda_k^2,$$

which is the same as (3.2). \hfill \Box
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:

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

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

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

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

It may be checked that

$$\Delta_\theta^m W(x) = \Delta_\theta^m \left[ \lambda_1 (x^1)^m + \lambda_2 (x^2)^m \right]$$

$$= m! \left[ \lambda_1 \cos^m \theta + \lambda_2 \sin^m \theta \right].$$  \hspace{1cm} (B.3)

Then

$$\left[ \Delta_\theta^m W(x) \right]^2 = (m!)^2 \left( \lambda_1^2 \cos^{2m} \theta + \lambda_2^2 \sin^{2m} \theta + 2\lambda_1 \lambda_2 \sin^m \theta \cos^m \theta \right)$$

$$= (m!)^2 \frac{1}{2} \left[ \lambda_1^2 (e^{i\theta} + e^{-i\theta})^{2m} + \lambda_2^2 (-i)^m (e^{i\theta} - e^{-i\theta})^{2m} \right.$$  

$$\left. + 2\lambda_1 \lambda_2 (-i)^m (e^{i\theta} + e^{-i\theta})^m (e^{i\theta} - e^{-i\theta})^m \right].$$  \hspace{1cm} (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,

\[
[\Delta^m W(x)]^2 = \left( \frac{1}{2} \right)^{2m} (m!)^2 \left( \begin{array}{c} 2m \\ m \end{array} \right) (\lambda_1^2 + \lambda_2^2) + \text{periodic terms.} \tag{B.5}
\]

If \( m \) is even,

\[
[\Delta_0^m W(x)]^2 = \left( \frac{1}{2}\right)^{2m} (m!)^2 \left[ \left( \begin{array}{c} 2m \\ m \end{array} \right) (\lambda_1^2 + \lambda_2^2) + 2 \left( \begin{array}{c} m \\ \sqrt{2m} \end{array} \right) \lambda_1 \lambda_2 \right] + \text{periodic terms.} \tag{B.6}
\]

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

\[
\int_0^{2\pi} [\Delta_0^m W(x)]^2 d\theta = \text{const.} \times (\lambda_1^2 + \lambda_2^2), \quad m \text{ odd;}
\]

\[
= \text{const.} \times \left[ (\lambda_1^2 + \lambda_2^2) + 2\lambda_1 \lambda_2 \left( \begin{array}{c} m \\ \sqrt{2m} \end{array} \right) / \left( \begin{array}{c} 2m \\ m \end{array} \right) \right], \quad m \text{ ev}
\]

\[
= \text{const.} \times \left\{ (\lambda_1 + \lambda_2)^2 \left( \begin{array}{c} m \\ \sqrt{2m} \end{array} \right) / \left( \begin{array}{c} 2m \\ m \end{array} \right) \right\} + (\lambda_1^2 + \lambda_2^2) \left[ 1 - \left( \begin{array}{c} m \\ \sqrt{2m} \end{array} \right) / \left( \begin{array}{c} 2m \\ m \end{array} \right) \right], \quad m \text{ even.} \tag{B.7}
\]

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

\[
\lambda_{11 \ldots 1} = \lambda_1, \quad \lambda_{22 \ldots 2} = \lambda_2.
\]
\[ \lambda_{i_1 \ldots i_m} = 0, \text{ otherwise.} \quad (B.9) \]

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

\[ \int_0^{2\pi} \left[ \Delta_0^m W(x) \right]^2 d\theta = \text{const.} \times \lambda_I \lambda_J. \quad (B.10) \]

Now consider an orthogonal transformation of \( x \) to \( y \), with

\[ y^i = A^i_j x^j. \quad (B.11) \]

Then

\[ A^i_k y^i = A^i_k A^k_j x^j = x^k, \quad (B.12) \]

by (2.16).

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

\[ W(x) = \nu_{i_1 \ldots i_m} y^{i_1 \ldots i_m}, \quad (B.13) \]

with

\[ \nu_{i_1 \ldots i_m} = \lambda_{i_1 \ldots i_m} A^i_{j_1} \ldots A^i_{j_m}. \quad (B.14) \]

i.e.

\[ \nu_J = \lambda_J A^J_i. \quad (B.15) \]

Then

\[ \nu_J \nu_K = A^J_i \lambda_J \lambda_K A^K_i. \quad (B.16) \]

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

\[ \nu_J \nu_J = \lambda_J \lambda_I. \quad (B.17) \]
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

$$W(x) = \lambda^{i_j}_{j_i} x^{i_1} ... x^{i_p} y^{j_1} ... y^{j_p},$$

(B.18)

with each $i_p, j_p = 1, 2,$ and

$$\lambda_{i1}^{11} = \lambda_1, \lambda_{i2}^{22} = \lambda_2,$$

$$\lambda_{i1}^{i_1} = 0, \text{ otherwise.}$$

(B.19)

Then

$$\int_0^{2\pi} [\Delta_0^m W(x)]^2 d\theta = \text{const.} \times \left\{ \left( \lambda_2 \right)^{2} \left( \frac{m}{\sqrt{2m}} \right) \left( \frac{m}{2m} \right) \right\}$$

$$+ \lambda_{i1}^{i_1} \lambda_{j1}^{j_1} \left[ 1 - \left( \frac{m}{\sqrt{2m}} \right) \left( \frac{m}{2m} \right) \right]$$

(B.20)

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

$$W(y) = v^{i_1}_{j_1} ... y^{i_p} y^{j_1} ... y^{j_p},$$

(B.21)

obtainable from (B.18) and (B.19) by an orthogonal transformation from $x$ to $y$, the quantities $v_{i}^{T}$ and $v_{j}^{T}$ are invariant under orthogonal transformations. Further, by (B.20), the quantity:
\[(\nu^2 I^2 \left( \begin{array}{c} m \\ \nu \cdot m \end{array} \right) / \left( \begin{array}{c} 2m \\ m \end{array} \right) + \nu^2 \nu I \left[ 1 - \left( \begin{array}{c} m \\ \nu \cdot m \end{array} \right) / \left( \begin{array}{c} 2m \\ m \end{array} \right) \right] \]  

(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) \( \frac{1}{2} n \) \((n \text{ even})\) or \( \frac{1}{2} (n-1) \) \((n \text{ odd})\) rotations of orthogonal 2-dimensional subspaces.

Consider the following generalization in (B.1):

\[ W(x) = \lambda_1 (x^1)^m + \ldots + \lambda_n (x^n)^m + \text{other terms,} \]  

(B.23)

where the "other terms" are still of degree \(< m \).

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

\[ \lambda_{\ldots i} = \lambda_i, \lambda_{i_1 i_2 \ldots i} = 0 \text{ otherwise.} \]  

(B.24)

Define the directional difference operator \( \Delta_{\theta} \) as the difference operator along the line obtained by rotating the \( x^{2q-1} \)-axis counter-clockwise through an angle \( \theta \) 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

\[ \Sigma_{i} \int_{0}^{2\pi} \left[ \Delta_{\theta}^m W(x) \right]^2 d \theta \]  

(B.25)

with \( q \) running from 1 to \( \frac{1}{2} n \) \((n \text{ even})\) or \( \frac{1}{2} (n-1) \) \((n \text{ odd})\). 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

\[
\alpha(W) = \text{const.} \times (\lambda_1^2 + \ldots + \lambda_n^2), \quad m \text{ odd;}
\]

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

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

The factor \( n!/\left(\frac{1}{2}n\right)! \times \left(\frac{1}{2}\right)^{\frac{1}{2}n} \) in (B.27) is the number of choices of \( \frac{1}{2}n \) planes of rotation in \( n \)-dimensional space. The corresponding factor in (B.28) is \( n!/[\frac{1}{2}(n-1)! \times \left(\frac{1}{2}\right)^{(n-1)/2}] \), but each such choice leads to the appearance of only \( n-1 \) out of \( n \) of the \( \lambda_i^2 \) 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:

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

with
\[ \beta_{mn} = 0, \text{ } m \text{ odd;} \]
\[ = 2^{n/2} \times \frac{(n/2)!}{n!} \left( \frac{m}{2m} \right) \left( \frac{m}{m} \right), \text{ } m \text{ even;} \]
\[ = 2^{(n-1)/2} \times \frac{[(n-1)/2]!}{(n-1)!} \left( \frac{m}{2m} \right) \left( \frac{m}{m} \right), \text{ } m \text{ even, } n \text{ odd.} \]  
(B.30)

Then \( \alpha(W) \) may be put in the form:
\[ \alpha(W) = \text{const.} \times \left[ \beta_{mn} (\lambda_1 + \ldots + \lambda_m)^2 + (1 - \beta_{mn}) (\lambda_1^2 + \ldots + \lambda_m^2) \right]. \]  
(B.31)

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

Then the mean square \( m \)-th difference \( \alpha(W) \) is measured by:
\[ v_I v_J, \text{ } m \text{ odd,} \]  
(B.32)

or
\[ \beta_{mn} (v_I^J)^2 + (1 - \beta_{mn}) v_I^J v_J^I, \text{ } m \text{ even.} \]  
(B.33)
References


<table>
<thead>
<tr>
<th>No.</th>
<th>Date</th>
<th>Subject</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>APR 93</td>
<td>AN EXPONENTIAL BOUND FOR RUIN PROBABILITIES</td>
<td>David C M Dickson</td>
</tr>
<tr>
<td>3</td>
<td>APR 93</td>
<td>SOME COMMENTS ON THE COMPOUND BINOMIAL MODEL</td>
<td>David C M Dickson</td>
</tr>
<tr>
<td>4</td>
<td>AUG 93</td>
<td>RUIN PROBLEMS AND DUAL EVENTS</td>
<td>David CM Dickson Alfredo D Egidio dos Reis</td>
</tr>
<tr>
<td>5</td>
<td>SEP 93</td>
<td>CONTEMPORARY ISSUES IN AUSTRALIAN SUPERANNUATION - A CONFERENCE SUMMARY</td>
<td>David M Knox John Piggott</td>
</tr>
<tr>
<td>6</td>
<td>SEP 93</td>
<td>AN ANALYSIS OF THE EQUITY INVESTMENTS OF AUSTRALIAN SUPERANNUATION FUNDS</td>
<td>David M Knox</td>
</tr>
<tr>
<td>7</td>
<td>OCT 93</td>
<td>A CRITIQUE OF DEFINED CONTRIBUTION USING A SIMULATION APPROACH</td>
<td>David M Knox</td>
</tr>
<tr>
<td>8</td>
<td>JAN 94</td>
<td>REINSURANCE AND RUIN</td>
<td>David C M Dickson Howard R Waters</td>
</tr>
<tr>
<td>9</td>
<td>MAR 94</td>
<td>LIFETIME INCOME, TAXATION, EXPENDITURE AND SUPERANNUATION (LITES): A LIFE-CYCLE SIMULATION MODEL</td>
<td>Margaret E Atkinson John Creedy David M Knox</td>
</tr>
<tr>
<td>10</td>
<td>FEB 94</td>
<td>SUPERANNUATION FUNDS AND THE PROVISION OF DEVELOPMENT/VENTURE CAPITAL: THE PERFECT MATCH? YES OR NO</td>
<td>David M Knox</td>
</tr>
<tr>
<td>11</td>
<td>JUNE 94</td>
<td>RUIN PROBLEMS: SIMULATION OR CALCULATION?</td>
<td>David C M Dickson Howard R Waters</td>
</tr>
<tr>
<td>12</td>
<td>JUNE 94</td>
<td>THE RELATIONSHIP BETWEEN THE AGE PENSION AND SUPERANNUATION BENEFITS, PARTICULARLY FOR WOMEN</td>
<td>David M Knox</td>
</tr>
<tr>
<td>13</td>
<td>JUNE 94</td>
<td>THE COST AND EQUITY IMPLICATIONS OF THE INSTITUTE OF ACTUARIES OF AUSTRALIA PROPOSED RETIREMENT INCOMES STRATEGY</td>
<td>Margaret E Atkinson John Creedy David M Knox Chris Haberecht</td>
</tr>
<tr>
<td>14</td>
<td>SEPT 94</td>
<td>PROBLEMS AND PROSPECTS FOR THE LIFE INSURANCE AND PENSIONS SECTOR IN INDONESIA</td>
<td>Catherine Prime David M Knox</td>
</tr>
<tr>
<td>Date</td>
<td>Authors</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>OCT 94</td>
<td>David M Knox</td>
<td>PRESENT PROBLEMS AND PROSPECTIVE PRESSURES IN AUSTRALIA'S SUPERANNUATION SYSTEM</td>
<td></td>
</tr>
<tr>
<td>DEC 94</td>
<td>Margaret E Atkinson, John Creedy, David M Knox</td>
<td>PLANNING RETIREMENT INCOME IN AUSTRALIA: ROUTES THROUGH THE MAZE</td>
<td></td>
</tr>
<tr>
<td>JAN 95</td>
<td>David C M Dickson, Alfredo D Egidio dos Reis</td>
<td>ON THE DISTRIBUTION OF THE DURATION OF NEGATIVE SURPLUS</td>
<td></td>
</tr>
<tr>
<td>FEB 95</td>
<td>Ben Zehnwirth</td>
<td>OUTSTANDING CLAIM LIABILITIES: ARE THEY PREDICTABLE?</td>
<td></td>
</tr>
<tr>
<td>MAY 95</td>
<td>David C M Dickson, Alfredo D Egidio dos Reis, Howard R Waters</td>
<td>SOME STABLE ALGORITHMS IN RUIN THEORY AND THEIR APPLICATIONS</td>
<td></td>
</tr>
<tr>
<td>JUN 95</td>
<td>David M Knox</td>
<td>SOME FINANCIAL CONSEQUENCES OF THE SIZE OF AUSTRALIA'S SUPERANNUATION INDUSTRY IN THE NEXT THREE DECADES</td>
<td></td>
</tr>
<tr>
<td>JUN 95</td>
<td>Margaret E Atkinson, John Creedy</td>
<td>MODELLING OPTIMAL RETIREMENT IN DECISIONS IN AUSTRALIA</td>
<td></td>
</tr>
<tr>
<td>JUN 95</td>
<td>Margaret E Atkinson, John Creedy, David M Knox</td>
<td>AN EQUITY ANALYSIS OF SOME RADICAL SUGGESTIONS FOR AUSTRALIA'S RETIREMENT INCOME SYSTEM</td>
<td></td>
</tr>
<tr>
<td>SEP 95</td>
<td>Angela Ryan</td>
<td>EARLY RETIREMENT AND THE OPTIMAL RETIREMENT AGE</td>
<td></td>
</tr>
<tr>
<td>OCT 95</td>
<td>David C M Dickson</td>
<td>APPROXIMATE CALCULATION OF MOMENTS OF RUIN RELATED DISTRIBUTIONS</td>
<td></td>
</tr>
<tr>
<td>DEC 95</td>
<td>David M Knox</td>
<td>CONTEMPORARY ISSUES IN THE ONGOING REFORM OF THE AUSTRALIAN RETIREMENT INCOME SYSTEM</td>
<td></td>
</tr>
<tr>
<td>FEB 96</td>
<td>Margaret E Atkinson, John Creedy</td>
<td>THE CHOICE OF EARLY RETIREMENT AGE AND THE AUSTRALIAN SUPERANNUATION SYSTEM</td>
<td></td>
</tr>
<tr>
<td>FEB 96</td>
<td>David C M Dickson, Ben Zehnwirth</td>
<td>PREDICTIVE AGGREGATE CLAIMS DISTRIBUTIONS</td>
<td></td>
</tr>
<tr>
<td>FEB 96</td>
<td>Margaret Atkinson</td>
<td>THE AUSTRALIAN GOVERNMENT SUPERANNUATION CO-CONTRIBUTIONS: ANALYSIS AND COMPARISON</td>
<td></td>
</tr>
<tr>
<td>MAR 96</td>
<td>Des Welch, Shauna Ferris</td>
<td>A SURVEY OF VALUATION ASSUMPTIONS AND FUNDING METHODS USED BY AUSTRALIAN ACTUARIES IN DEFINED BENEFIT SUPERANNUATION FUND VALUATIONS</td>
<td></td>
</tr>
<tr>
<td>MAR 96</td>
<td>David C M Dickson, Alfredo D Egidio dos Reis</td>
<td>THE EFFECT OF INTEREST ON NEGATIVE SURPLUS</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Author(s)</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>MAR 96</td>
<td>Greg Taylor</td>
<td>RESERVING CONSECUTIVE LAYERS OF INWARDS EXCESS-OF-LOSS REINSURANCE</td>
<td></td>
</tr>
<tr>
<td>AUG 96</td>
<td>Anthony Asher</td>
<td>EFFECTIVE AND ETHICAL INSTITUTIONAL INVESTMENT</td>
<td></td>
</tr>
<tr>
<td>AUG 96</td>
<td>Michael Sherris, Leanna Tedesco, Ben Zehnwirth</td>
<td>STOCHASTIC INVESTMENT MODELS: UNIT ROOTS, COINTEGRATION, STATE SPACE AND GARCH MODELS FOR AUSTRALIA</td>
<td></td>
</tr>
<tr>
<td>AUG 96</td>
<td>Ben Zehnwirth</td>
<td>THREE POWERFUL DIAGNOSTIC MODELS FOR LOSS RESERVING</td>
<td></td>
</tr>
<tr>
<td>SEPT 96</td>
<td>Ben Zehnwirth</td>
<td>KALMAN FILTERS WITH APPLICATIONS TO LOSS RESERVING</td>
<td></td>
</tr>
<tr>
<td>OCT 96</td>
<td>David C M Dickson, Howard R Waters</td>
<td>RELATIVE REINSURANCE RETENTION LEVELS</td>
<td></td>
</tr>
<tr>
<td>OCT 96</td>
<td>Greg Taylor</td>
<td>SMOOTHNESS CRITERIA FOR MULTIDIMENSIONAL WHITTAKER GRADUATION</td>
<td></td>
</tr>
</tbody>
</table>