Computational Statistics

SpringerLink

Search 🔍 🛒 Log in

Computational Statistics

Search within journal

Volume 29, issue 6, December 2014

20 articles in this issue

A generalized Newton algorithm for quantile regression models

Songfeng Zheng

Original Paper | Published: 30 April 2014 | Pages: 1403 - 1426

Row–column interaction models, with an R implementation

Thomas W. Yee & Alfian F. Hadi

Original Paper | Published: 27 May 2014 | Pages: 1427 - 1445
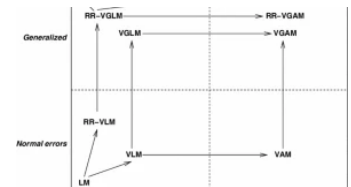
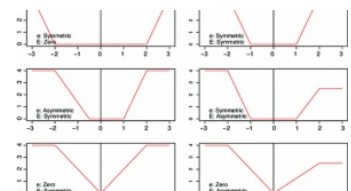### An e–E-insensitive support vector regression machine

Amir Safari

Original Paper | Published: 11 May 2014 | Pages: 1447 - 1468

Efficient computation for the Poisson binomial distribution
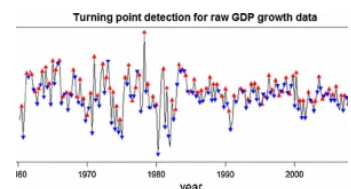
Bruce E. Barrett & J. Brian Gray

Original Paper | Published: 20 May 2014 | Pages: 1469 - 1479
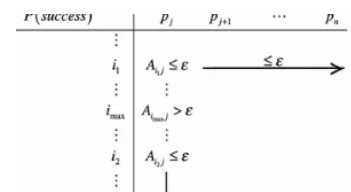
Wavelet improvement in turning point detection using a hidden Markov model: from the aspects of cyclical identification and outlier correction
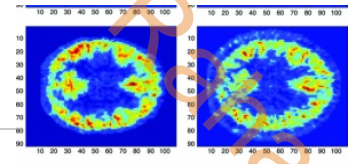
Yushu Li & Simon Reese

Original Paper | Published: 28 May 2014 | Pages: 1481 - 1496

Functional data classification in a wavelet approach

Original Paper | Published: 04 June 2014 | Pages: 1497 - 1513

A VIF-based optimization model to alleviate collinearity problems in multiple linear regression
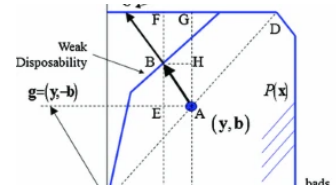
Yow-Jen Jou, Chien-Chia Liäm Huang & Hsun-Jung Cho

Original Paper | Published: 05 June 2014 | Pages: 1515 - 1541

Dynamic activity analysis model-based win-win development forecasting under environment regulations in China

Shiyi Chen & Wolfgang K. Härdle

Original Paper | Published: 01 June 2014 | Pages: 1543 - 1570

Quantile regression of right-censored length-biased data using the Buckley–James-type method

Jung-Yu Cheng & Shinn-Jia Tzeng

Original Paper | Published: 31 May 2014 | Pages: 1571 - 1592
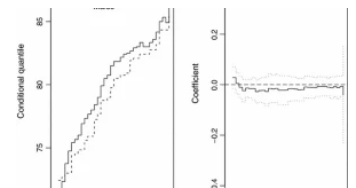
A Levene-type test of homogeneity of variances against ordered alternatives

Philip Pallmann, Ludwig A. Hothorn & Gemechis D. Djira

Original Paper | Published: 06 June 2014 | Pages: 1593 - 1608
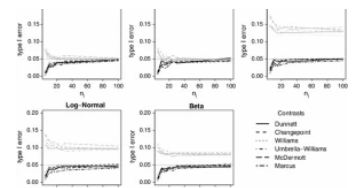
Multivariate nonlinear least squares: robustness and efficiency of standard versus Beauchamp and Cornell methodologies

Renato Guseo & Cinzia Mortarino

Original Paper | Published: 07 June 2014 | Pages: 1609 - 1636
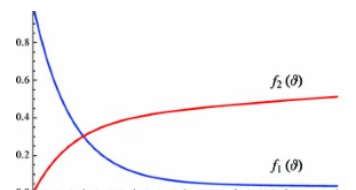
Recursions on the marginals and exact computation of the normalizing constant for Gibbs processes

Cécile Hardouin & Xavier Guyon

Original Paper | Published: 06 June 2014 | Pages: 1637 - 1650

Composite support vector quantile regression estimation

Jooyong Shim, Changha Hwang & Kyungha Seok

Original Paper | Published: 26 July 2014 | Pages: 1651 - 1665

✓ You have access to our articles

For authors

[Submission guidelines](#)

[Manuscript editing services](#)

[Ethics & disclosures](#)

[Open Access fees and funding](#)

[Contact the journal](#)

Submit manuscript

Working on a manuscript?

Avoid the most common mistakes and prepare your manuscript for journal editors.

[Learn more →](#)

Explore

[Online first articles](#)

[Volumes and issues](#)

Sign up for alerts

Row–column interaction models, with an R implementation

Thomas W. Yee · Alfian F. Hadi

Abstract We propose a family of models called row–column interaction models (RCIMs) for two-way table responses. RCIMs apply some link function to a parameter (such as the cell mean) to equal a row effect plus a column effect plus an optional interaction modelled as a reduced-rank regression. What sets this work apart from others is that our framework incorporates a very wide range of statistical models, e.g., (1) log-link with Poisson counts is Goodman's RC model, (2) identity-link with a double exponential distribution is median polish, (3) logit-link with Bernoulli responses is a Rasch model, (4) identity-link with normal errors is two-way ANOVA with one observation per cell but allowing semi-complex modelling of interactions of the form $\mathbf{A}\mathbf{C}^T$, (5) exponential-link with normal responses are quasi-variances. Proposed here also is a least significant difference plot augmentation of quasi-variances. Being a special case of RCIMs, quasi-variances are naturally extended from the $M = 1$ linear/additive predictor η case (within the exponential family) to the $M > 1$ case (vector generalized linear model families). A rank-1 Goodman's RC model is also shown to estimate the site scores and optimums of an equal-tolerances Poisson unconstrained quadratic ordination. New functions within the VGAM R package are described with examples. Altogether, RCIMs facilitate the analysis of matrix responses of many data types, therefore are potentially useful to many areas of applied statistics.

T. W. Yee (✉)
Department of Statistics, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
e-mail: t.yee@auckland.ac.nz

A. F. Hadi
Department of Mathematics, University of Jember, Jalan Kalimantan 37, Jember 68 121, Indonesia
e-mail: afhadi@unej.ac.id

1 Introduction

Two-way tables are a common form of data. These include contingency tables of counts, and tables with the presence of outliers. This article proposes a new class of models called *row–column interaction models* (RCIMs) to analyze general two-way tables where the matrix response may be continuous, counts, proportions, etc. As a specific example, consider Goodman's RC association model (Goodman 1981) applied to a $n \times M$ matrix of counts $\mathbf{Y} = [(y_{ij})]$:

$$\log \mu_{ij} = \mu + \alpha_i + \gamma_j + \sum_{r=1}^{R} c_{ir} a_{jr}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, M, \quad (1)$$

where it is assumed Y_{ij} has a Poisson distribution. Here, $\mu_{ij} = E(Y_{ij})$ is the mean of the i-j cell and R is known as the *rank* (usually a small integer such as 0, 1 or 2). The LHS of (1) takes on the form of a generalized linear model (GLM)—it is $g(\theta)$ where g is a link function and θ is a parameter. The RHS consists of an intercept, and row and column main effects plus optional interaction terms in the form of a reduced-rank regression (RRR; see below).

Put succinctly, RCIMs generalize (1) so that, for the LHS, (1) the link function can be any reasonable monotonic and twice-differentiable function, and (2) the parameter is general and not necessarily a mean. Of course, the link function is chosen to match the parameter being modelled. RCIMs are defined for a matrix $\mathbf{Y}$ response.

As an example of a RCIM, consider the number of alcohol offenders caught in New Zealand during 2009 whose blood-alcohol level was above the legal limit while driving (Table 1). Of interest is how the counts vary by day and by time. For example, one would expect drink-driving to occur more in the evenings, especially during the weekends and Friday. If this were true then one would like to quantify the effect and provide some graphical display showing these features.

It will be shown in this article that RCIMs are a subfamily within the *reduced-rank vector generalized linear model* (RR-VGLM) family, which is itself a variant of the VGLM class (Yee and Hastie 2003). Figure 1 gives an overview of all the major classes of models in the statistical framework considered by this article. They are detailed later. We believe RCIMs are surprisingly useful, for example, they allow quasi-variances (Firth and de Menezes 2004; Firth 2000) to be conveniently computed as a special case, and furthermore, the models need not necessarily belong to the classical exponential family. Quasi-variances are developed as RCIMs in Sect. 5.3.

In this article the following novel results are established.

(i) The RCIM class is proposed.

Table 1 The R data frame `alcoff` in the **VGAM** package, which is dimension 24 × 7 (the 24-h clock time by day of the week)

t							
Hour	Mon	Tue	Wed	Thu	Fri	Sat	Sun
0	121	98	165	324	827	1,379	1,332
1	97	92	157	278	619	1,327	1,356
2	60	69	107	229	410	979	1,011
3	55	60	75	238	401	693	718
4	25	38	48	145	223	346	410
5	19	10	20	56	139	188	287
6	13	9	9	55	70	155	213
7	20	6	9	42	40	160	200
8	5	8	12	29	40	79	96
9	7	8	14	28	38	44	58
10	7	13	20	36	38	52	69
11	8	10	20	32	27	44	39
12	10	13	8	39	37	73	59
13	9	26	32	37	27	41	45
14	22	41	31	46	42	58	53
15	39	48	62	69	59	75	70
16	28	48	71	85	55	96	95
17	46	59	98	141	136	154	130
18	53	100	117	185	223	236	121
19	74	119	155	289	335	337	146
20	74	135	283	508	591	490	166
21	84	154	326	610	866	754	131
22	90	143	345	765	976	1,026	114
23	110	169	363	899	1,265	1,179	159

The counts are the number of alcohol offenders caught from breath screening drivers, accumulated over 2009 in New Zealand

(ii) It is shown that quasi-variances can be computed as a RCIM. Consequently, quasi-variances are now available for models outside the classical exponential family. In particular, they can be calculated for models with more than one linear predictor or parameter.

(iii) A least significant difference (LSD) plot innovation is proposed for (ii).

(iv) An easy-to-use implementation of the RCIM class is described based on the first author's **VGAM** R package.

An outline of this paper is as follows. It is firstly needful to summarize VGLMs and RR-VGLMs in the next section. RCIMs are defined in Sect. 3 with examples. Estimation and software details are given in Sect. 4. Some real-life examples are given in Sect. 5, and the article concludes with a discussion.
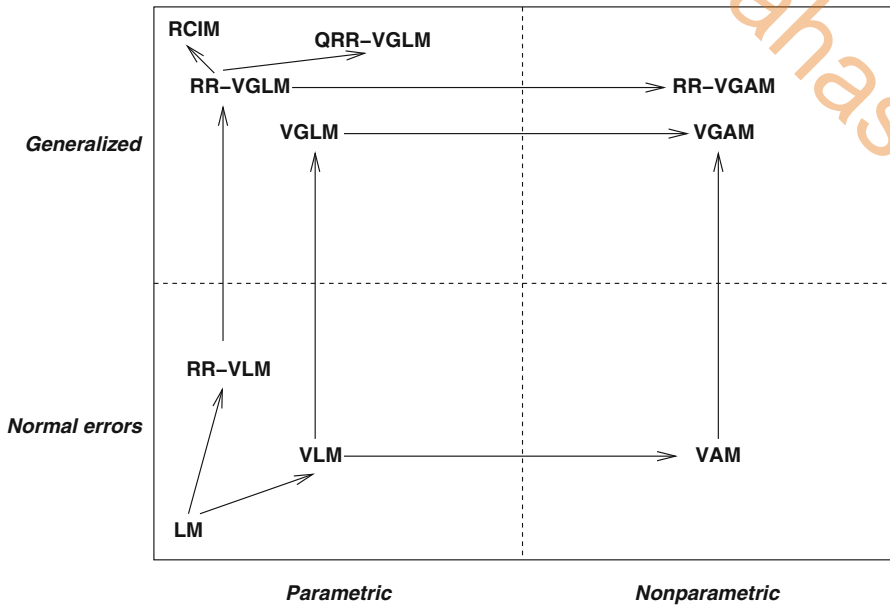
Fig. 1 Flowchart for different classes of models relevant to this article. Legend: *LM* linear model, *V* vector, *generalized*, *A* additive, *RR* reduced-rank, *Q* quadratic, *RCIM* row–column interaction model

2 VGLMs and RR-VGLMs

RCIMs are built upon the VGLM/RR-VGLM infrastructure, therefore it is necessary to briefly summarize them. The general VGLM/RR-VGLM statistical framework is described formally in Yee and Hastie (2003) and informally in Yee (2008). The reader is directed to these for further details. A recent article illustrating the versatility of RR-VGLMs is Yee (2014). Many examples here pertain to categorical data analysis, therefore Yee (2010) is also useful. We now very briefly describe the essential details of VGLMs and RR-VGLMs.

Let the data be $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ for $i = 1, \ldots, n$, and write $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})^\top$ with $x_{i1} = 1$ if there is an intercept. The $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ are explanatory and response respectively. Sometimes we drop the subscript i and write $\boldsymbol{x} = (x_1, \ldots, x_p)^T$ for simplicity. Consider the one observation case where the observed $\boldsymbol{y}$ is a q-dimensional vector. VGLMs are defined as a model where

$$f(\boldsymbol{y}|\boldsymbol{x}; \mathbf{B}) = f(\boldsymbol{y}, \eta_1, \ldots, \eta_M) \tag{2}$$

for some known function $f(\cdot)$, $\mathbf{B} = (\boldsymbol{\beta}_1\, \boldsymbol{\beta}_2 \cdots \boldsymbol{\beta}_M)$ is a $p \times M$ matrix of unknown regression coefficients. This means that the model's parameters are modelled as linear predictors. In (2) we say the distribution of $\boldsymbol{y}$, given $\boldsymbol{x}$, depends on a matrix of parameters $\mathbf{B}$. The jth linear predictor is

$$\eta_j(\boldsymbol{x}) = \boldsymbol{\beta}_j^T \boldsymbol{x} = \sum_{k=1}^{p} \beta_{(j)k}\, x_k, \quad j = 1, \ldots, M. \tag{3}$$

The η_j of VGLMs may be applied directly to parameters of a distribution rather than just to means as for GLMs, e.g., a univariate distribution with location parameter a and scale parameter $b > 0$. Then we might take $\eta_1 = a$ (identity link) and $\eta_2 = \log b$. In general,

$$\eta_j(\boldsymbol{x}) = g_j(\theta_j), \quad j = 1, \ldots, M, \tag{4}$$

and we say g_j is a *parameter link function* because θ_j is any parameter and not restricted to be a mean as with GLMs. Equation (4) therefore offers more scope than GLM modelling.

Write

$$\boldsymbol{\eta}_i = \begin{pmatrix} \eta_1(\boldsymbol{x}_i) \\ \vdots \\ \eta_M(\boldsymbol{x}_i) \end{pmatrix} = \mathbf{B}^T \boldsymbol{x}_i = \begin{pmatrix} \boldsymbol{\beta}_1^T \boldsymbol{x}_i \\ \vdots \\ \boldsymbol{\beta}_M^T \boldsymbol{x}_i \end{pmatrix}. \tag{5}$$

In practice we may wish to constrain the effect of a covariate to be the same for some of the η_j and to have no effect for others. For example,

$$\eta_1(\boldsymbol{x}_i) = \beta_{(1)1}^* + \beta_{(1)2}^* x_{i2} + \beta_{(1)3}^* x_{i3},$$
$$\eta_2(\boldsymbol{x}_i) = \beta_{(2)1}^* + \beta_{(1)2}^* x_{i2},$$

so that $\beta_{(1)2} \equiv \beta_{(2)2}$ and $\beta_{(2)3} \equiv 0$. The superscript "*" denotes regression coefficients that are actually estimated. For VGLMs, we can represent these models using

$$\boldsymbol{\eta}(\boldsymbol{x}_i) = \sum_{k=1}^p \boldsymbol{\beta}_{(k)} \, x_{ik} = \sum_{k=1}^p \mathbf{H}_k \, \boldsymbol{\beta}_{(k)}^* \, x_{ik} \tag{6}$$

where $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_p$ are *known constraint matrices* of full column-rank [i.e., their rank equals `ncol`$(\mathbf{H}_k)$], $\boldsymbol{\beta}_{(k)}^*$ is a vector containing a possibly reduced set of regression coefficients. With no constraints at all, all $\mathbf{H}_k = \mathbf{I}_M$ and $\boldsymbol{\beta}_{(k)}^* = \boldsymbol{\beta}_{(k)}$. Then

$$\mathbf{B}^T = \begin{pmatrix} \mathbf{H}_1 \boldsymbol{\beta}_{(1)}^* & \mathbf{H}_2 \boldsymbol{\beta}_{(2)}^* & \ldots & \mathbf{H}_p \boldsymbol{\beta}_{(p)}^* \end{pmatrix} \tag{7}$$

is an expression of (5) concentrating on columns rather than rows. We need both (5) and (7) because sometimes we focus on the η_j and at other times on the variables x_k.

2.1 Reduced-rank vector generalized linear models

Partition $\boldsymbol{x}_i$ into $(\boldsymbol{x}_{1i}^T, \boldsymbol{x}_{2i}^T)^T$ and $\mathbf{B} = (\mathbf{B}_1^T \ \mathbf{B}_2^T)^T$. In general, $\mathbf{B}$ is a dense matrix of full rank, i.e., min(M, p). The $M \times p$ regression coefficients to be estimated means, for some models and data sets, this is "too" large and is susceptible to overfitting. A dimension reduction method is warranted.

A simple solution is to replace $\mathbf{B}_2$ by a RRR: $\mathbf{B}_2 = \mathbf{A}\, \mathbf{C}^T$ where $\mathbf{A}$ and $\mathbf{C}$ are two thin matrices each of rank R. They are 'thin' because often R is low, e.g., 0 or 1 or 2, hence the number of columns is much less than the number of rows. If R is low then

the number of regression coefficients can be reduced enormously. When $R = 2$ the $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{C}}$ may be biplotted (Yee and Hastie 2003). The RRR is applied to $\mathbf{B}_2$ because we want to make provision the variables in $\mathbf{x}_{1i}$ to be left alone, e.g., the intercepts.

An important fact is that RR-VGLMs are VGLMs where some of the constraint matrices are estimated. For RR-VGLMs,

$$\boldsymbol{\eta}_i = \mathbf{B}_1^T \mathbf{x}_{1i} + \mathbf{A}\, \mathbf{C}^T \mathbf{x}_{2i} = \mathbf{B}_1^T \mathbf{x}_{1i} + \mathbf{A}\, \boldsymbol{v}_i, \tag{8}$$

where $\mathbf{C} = (\boldsymbol{c}_{(1)}\, \boldsymbol{c}_{(2)} \ldots \boldsymbol{c}_{(R)})$ is $p_2 \times R$, $\mathbf{A} = (\boldsymbol{a}_{(1)}\, \boldsymbol{a}_{(2)} \ldots \boldsymbol{a}_{(R)}) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_M)^T$ is $M \times R$. Both $\mathbf{A}$ and $\mathbf{C}$ are of full column-rank. Of course, $R \leq \min(M, p_2)$ but ideally we want $R \ll \min(M, p_2)$ because a more parsimonious model is often desired. One can think of (8) as a RRR of the coefficients of $\mathbf{x}_{2i}$ after having adjusted for the variables in $\mathbf{x}_{1i}$. For identifiability, $\mathbf{A}$ may be restricted to the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_R \\ \widetilde{\mathbf{A}} \end{pmatrix}, \quad \text{say,} \tag{9}$$

called *corner* constraints: only $\widetilde{\mathbf{A}}$ need be estimated.

3 Row–column interaction models

Define RCIMs as a RR-VGLM applied to $\mathbf{Y}$, with

$$g_1(\theta_1) \equiv \eta_{1ij} = \mu + \alpha_i + \gamma_j + \sum_{r=1}^{R} c_{ir}\, a_{jr}, \tag{10}$$

[cf. (1)], where $R \leq \min(M, p_2)$. This means that the first parameter of a statistical model relating to a response matrix is, after a suitable transformation, equal to the sum of an intercept, a row and column effect plus optional interaction term of the form $\mathbf{A}^T \mathbf{C}$.

Note that (10) applies to the *first* (default) linear/additive predictor. For models with $M > 1$ parameters one can leave $\eta_2, \ldots, \eta_M$ in (4) unchanged because these are functions of nuisance parameters (e.g., scale and shape parameters) that are best left alone (probably as intercept-only: $\eta_j = \beta_{(j)1}$). Of course, choosing η_1 for (10) is only for convenience and one could choose some other η_j to assign the RHS of (10).

3.1 Examples

Table 2 lists several RCIM examples that are currently implemented in software. Here are some supplementary comments.

1. The median polish example (Mosteller and Tukey 1977), which is a method similar to two-way ANOVA that uses medians $\widetilde{\mu}$ instead of means, can be fitted by esti- mating the location parameter of an asymmetric Laplace distribution (Kotz et al.

Table 2 Some VGAM family functions useful in conjunction with `rcim()` (they normally operate with `vglm(..., family = )`.) Some models are linked to equations found in the article. "GRC" stands for Goodman's RC model

Family function name	Comments
`alaplace2(0.5)`	Median polish when rank-0, Eq. (11)
`binomialff(mv = TRUE)`	Rasch conditional (fixed-effects) model, Eq. (12)
`negbinomial()`	GRC with overdispersion wrt Poisson
`poissonff()`	GRC model, Eq. (1)
`uninormal()`	2-way ANOVA (1 observation per cell)
`uninormal("explink")`	Quasi-variances when rank-0
`zipoissonff()`	GRC with surplus 0s wrt Poisson

2001). (Actually, medians correspond to the case where the asymmetry parameter makes the distribution symmetric.) The formula is

$$\widetilde{\mu}_{ij} = \mu + \alpha_i + \gamma_j + \sum_{r=1}^{R} c_{ir} \, a_{jr}. \tag{11}$$

It is therefore to be expected that median polish is more robust to outliers than two-way ANOVA.

2. Another RCIM is the simple Rasch model (Powers and Xie 2008)

$$\text{logit } p_{ij} = \theta_i - b_j, \tag{12}$$

where p_{ij} is the ith individual's probability of a correct response on item j, θ_i denotes an individual's ability parameter, and b_j is the difficulty of the jth item.

3. Another RCIM is a Goodman's RC model extended to handle excess 0s, hence a zero-inflated Poisson model could be a possibility:

$$P(Y = y) = I(y = 0) \, (1 - \phi^*) + \phi^* \, e^{-\lambda} \lambda^y / y! \tag{13}$$

where $I(\cdot)$ is the indicator function, $\phi = 1 - \phi^*$ is the probability of a structural zero, and

$$\boldsymbol{\eta} = \left(\log \lambda, \text{logit } \phi^* \right)^T. \tag{14}$$

4. A final example is the quasi-variance methodology of Firth and de Menezes (2004) applied to an explanatory factor variable X with L levels of which the first is baseline. These are similar to 'floating absolute risks' (Easton et al. 1991), however quasi-variances are more reliably accurate because they minimize the relative error of an approximation whereas floating absolute risks minimize the absolute error. Quasi-variances are an especially interesting special case, therefore we devote the next section to its further development.

3.2 Quasi-variances

For an explanatory factor variable X with L levels of which the first is baseline, the usual output is $\widehat{\beta}_j$, $j = 2, \ldots, L$, and its variance-covariance matrix $\widetilde{\mathrm{Var}}(\widehat{\beta})$. Plotting the estimates with their standard errors means that the reference level is simply a point without any standard errors, e.g., Fig. 4a. While all other levels can be compared to the first, it is impractical to compare between two other levels. The purpose of quasi-variances is to be able to quickly make inferences for contrasts of the parameters without having to re-evaluate standard errors after reparameterization. The methodology attempts to assign standard errors to all levels such that *all* pairwise levels may be compared. The L quasi-standard errors $\mathrm{qse}(\widehat{\beta}_j)$, $j = 1, \ldots, L$, are treated as uncorrelated, and their derivation is based on an approximation. When plotted (e.g., Fig. 4b) the overall result does not treat the reference level as different from the others.

Without divulging unnecessary details, the quantities

$$v_{ij} \equiv \mathrm{Var}\left(\widehat{\beta}_i - \widehat{\beta}_j\right) \approx q_i + q_j, \quad i = 1, \ldots, L, \quad j = 1, \ldots, L, \qquad (15)$$

are approximated as best as possible with respect to the quasi-variances q_i. By considering a relative error, the q_i may be estimated (Firth 2000) by minimizing

$$\sum_{i<j} \left\{ \log v_{ij} - \log \left(q_i + q_j \right) \right\}^2. \qquad (16)$$

This can be achieved within the RCIM framework by letting $y_{ij} = \log v_{ij}$ be the 'response' and choosing

$$g(\mu_{ij}) = \eta_{ij} = \exp(\mu_{ij}) \qquad (17)$$

so that the link function is an exponential [the opposite to a log-link]. If it is assumed that $y_{ij} \sim N(\mu_{ij}, \sigma^2)$ then fitting a rank-0 RCIM to $\mathbf{Y} = [(y_{ij})]$ with exponential link function and normal errors will result in minimizing (16) by least squares because $\eta_{ij} = q_i + q_j$.

The function explink() has been written to facilitate these calculations. Care is required to keep $\eta_{ij} > 0$ to avoid range violations. By choosing $\eta_{ij}^{(a)} = \max(\eta_{ij}^{(a)}, \varepsilon)$ at iteratively reweighted least squares (IRLS) iteration a for suitably small positive ε, any value beyond the boundary of the parameter space is handled during the early iterations. Our experience is that the computations are generally reliable given this tweak. The VGAM family function uninormal() has an argument that allows ε to be inputted; its current default value is 1×10^{-5} but any value in $(0, 0.1]$ is accepted. In practice, there might only be a problem if the model converged or was stopped at an IRLS iteration that this adjustment was needed, however this has never occurred in the author's experience.

There are several benefits from calculating quasi-variances as a RCIM. Because RCIMs reside within the VGLM/RR-VGLM framework the methodology automatically becomes available to a wide range of models and distributions. These include models with $M > 1$ linear predictors, i.e., the entire VGLM class. Specialized software in another package not being needed is another benefit.

For quasi-variances we now propose a small innovation: adding least significant difference (LSD) intervals to the plot (Andrews et al. 1980; Schenker and Gentleman 2001). We wish users to be able to visually determine instantaneously whether $H_0 : \beta_i = \beta_j$ is rejected or not, i.e., are any two levels significantly different? Such methodology already exists in the form of the LSD idea. Ideally one would want LSD intervals that do not overlap if and only if they are statistically significantly different. This feature is not available in the usual quasi-SE interval presentation. The LSD addition allows the reader to make immediate decisions by making comparisons between pairs of factor levels easier. Another alternative would be based on honestly significant difference (HSD) intervals. In our implementation, LSD intervals are represented as arrows inside the usual quasi-SE line segments (e.g., Fig. 4b). The LSD innovation uses the result that, given $\widehat{\theta}_1$, $\widehat{\theta}_2$, $\mathrm{se}(\widehat{\theta}_1)$ and $\mathrm{se}(\widehat{\theta}_2)$, then $\widehat{\theta}_1 - \widehat{\theta}_2 \pm 1.96\,\mathrm{se}(\widehat{\theta})/\sqrt{2}$ is an approximate 5 % LSD interval for the difference between two parameters, $\theta_1 - \theta_2$, provided $\mathrm{se}(\widehat{\theta}_1) \approx \mathrm{se}(\widehat{\theta}_2)$ ($= \mathrm{se}(\widehat{\theta})$, say). The general formula for the 100α % LSD interval is

$$\widehat{\beta}_i \pm \frac{z(\alpha/2)}{\sqrt{2}} \sqrt{q_i}. \tag{18}$$

The proof is as follows. If $q_i \approx q_j$ ($= q$, say) then a $100(1-\alpha)$ % confidence interval for $\beta_i - \beta_j$ is $\widehat{\beta}_i - \widehat{\beta}_j \pm z(\alpha/2)\sqrt{2q}$. Suppose $\widehat{\beta}_i > \widehat{\beta}_j$. At the point of overlap we have λ such that $\widehat{\beta}_i - \lambda\sqrt{q} = \widehat{\beta}_j + \lambda\sqrt{q}$. The $100(1-\alpha)$ % confidence interval touches 0 when $\widehat{\beta}_i - \widehat{\beta}_j = z(\alpha/2)\sqrt{2q}$. Equating the two equations shows that $\lambda = z(\alpha/2)/\sqrt{2}$, hence (18).

4 Estimation and software

We now sketch the estimation details of RCIMs. They are related to the basics given in Sect. 2 on VGLMs/RR-VGLMs, and Yee and Hastie (2003).

In (10) the parameters α_i and γ_j are called the *row* and *column scores* (or *effects*) respectively. Identifiability constraints are needed for these, such as corner constraints, e.g., $\alpha_1 = \gamma_1 = 0$. The parameters a_{ir} and c_{jr} also need constraints, e.g., $a_{1r} = c_{1r} = 0$ for $r = 1, \ldots, R$. Following Yee and Hastie (2003), RCIMs may be fitted as RR-VGLMs by setting up the appropriate indicator variables as

$$\boldsymbol{\eta}_i = \mathbf{B}_1^T \boldsymbol{x}_{1i} + \mathbf{A}\mathbf{C}^T \boldsymbol{x}_{2i} \tag{19}$$

$$= \left\{ \mu\mathbf{1}_M + \alpha_i\mathbf{1}_M + \sum_{j=2}^{M} \gamma_j\,\boldsymbol{e}_j \right\} + \begin{pmatrix} \mathbf{0}^T \\ \mathbf{A}_{[-1]} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{C}_{[-1]}^T \end{pmatrix} \boldsymbol{e}_i \tag{20}$$

where a subscript "$[-1]$" means the first element or row is removed from the vector or matrix, and $\boldsymbol{e}_i$ is the ith column of $\mathbf{I}_M$. The first rows of $\mathbf{A}$ and $\mathbf{C}$ are $\mathbf{0}^T$ for identifiability—they correspond to structural zeros.

Since $\mathbf{A}$ takes on the role of the constraint matrices of a VGLM, it is possible to alternate between estimating $\mathbf{A}$ given $\mathbf{C}$ and estimating $\mathbf{C}$ given $\mathbf{A}$. Both operations entail fitting a VLM (Fig. 1). These are conducted within the IRLS algorithm that performs Fisher scoring iterations.

In our experience the choice of initial values does sometimes play a large role whether successful convergence is obtained or not. By 'sometimes', it depends on the distribution or model being fitted. For example, distributions with 3 or 4 or more parameters can pose much more difficulty than 2 parameter ones, e.g., the 2-parameter Lomax distribution is the 4-parameter generalized beta II distribution with shape parameters $a = p = 1$. There are tricks that can be used to fit the more complex models such as fitting a simpler one first and use it for initial values. Some of these tricks are described in Yee (2015). In general, because the expected information matrices (EIMs) are used, the parameter space where the working weight matrices are positive-definite is much larger than for the Newton–Raphson algorithm (which uses the observed information matrices). Thus often the starting values need not be very close to the solution. Consequently, successful convergence is often insensitive to the starting values, which is a good thing. All VGAM family functions are self-starting. Many have a variety of methods implemented to select the starting values; such family functions have the argument imethod (which stands for 'initialization method') that always has the default value 1 for the method deemed most likely to provide convergence, e.g., the method of moments. If this fails then the user can invoke imethod = 2, and if that fails, then imethod = 3, etc.

For most of the 150+ VGAM family functions their EIMs are well-conditioned. But there do exist distributions whose parameterization must be chosen with care, e.g., some generalized hyperbolic distributions fitted in GeneralizedHyperbolic (Scott 2012), however the present version of VGAM does not implement any such distributions whose EIM is so ill-conditioned that this might be an issue.

In this vein, the use of suitable link functions helps with the overall stability of the parameterization and circumvents range restriction problems, e.g., parameters in (0, 1) are best estimated using a link such as the logit. Reasonable default link functions are chosen for all parameters estimated within VGAM, therefore most users need not concern themselves with problems relating to parameter range violations.

In the case of the asymmetric Laplace distribution, the first derivatives are not continuous at the location parameter, therefore scoring is not a suitable algorithm. Hence alaplace2() in Table 2 is not entirely reliable. However, the half-stepping feature of vglm() can help ameliorate this problem. Here, half-stepping allows a partial step (e.g., $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$...) to be taken from the current estimate $\boldsymbol{\beta}^{(a)}$ in the direction of $\boldsymbol{\beta}^{(a+1)}$ if $\boldsymbol{\beta}^{(a+1)}$ is not an improvement. Half-stepping ensures that an improvement is made at each iteration because the step directions are ascending.

VGLMs and RR-VGLMs are illustrated in Fig. 1. RR models operate on $\boldsymbol{v}_i = (v_{1i}, \ldots, v_{iR})^T$, which may be interpreted as a vector of latent variables. They are optimal linear combinations of the $\boldsymbol{x}_{2i}$—optimal in the sense that they are estimated by maximizing the likelihood. Then the $\boldsymbol{v}_i$ can be considered as taking the role of $\boldsymbol{x}_{2i}$, i.e., as a smaller set of explanatory variables. Models VLMs and RR-VLMs are computational building blocks of the other methods but do not have any significant practical utility; they are fitted to pseudo-responses within the IRLS algorithm.

For the quasi-variance calculations (17), the function explink() has been written to facilitate these calculations. Care is required to keep $\eta_{ij} > 0$ to avoid range violations. By choosing $\eta_{ij} = \max(\eta_{ij}, \varepsilon)$ for suitably small positive ε, any value

beyond the boundary of the parameter space is handled during the early iterations. Our experience is that the computations are generally reliable given this tweak.

4.1 Modelling function `rcim()`

The modelling function `rcim()` sets up the dummy variables and constraint matrices according to (20) as part of fitting the RR-VGLM. Equation (20) is the central formula for fitting RCIMs; it shows how the indicator variables are set up for the row and column effects. The `family` argument of `rcim()` may be assigned a **VGAM** family function which can handle multiple responses because each column of **Y** is treated as a response and the basic unit of the computations is $\boldsymbol{\eta}_i$.

The argument `Rank = 0` by default means no interaction terms are fitted. The argument `Index.corner` specifies which rows of **A** are to receive corner constraints, cf. (9). For more information about these and other arguments, see the online help files.

Setting the argument `which.linpred = t` of `rcim()` corresponds to replacing η_{1ij} in (10) by η_{tij}, e.g., `zipoisson()` has $\boldsymbol{\eta} = (\text{logit } p_{\text{structural 0}}, \log \lambda)^T$, hence it would be appropriate to set `which.linpred = 2` because the second parameter is of central interest.

Given a fitted `rcim()` object, S4 generic functions such as `coef()`, `fitted()`, `predict()` and `summary()` are available. For rank-0 models `plot()` can produce a plot of the main effects.

For quasi-variance calculations, functions `rcim()`, `Qvar()`, `qvar()`, `explink()` and `uninormal()` are used collectively to calculate them for a specified factor explanatory variable. The online help gives examples, as well as below. The methods function `plotqvar()` is available for plotting the quasi-variances.

5 Examples

Several RCIMs are illustrated in this section. R scripts are available from http://www.stat.auckland.ac.nz/~yee/updates, and these use the **VGAM** and **VGAMdata** packages (versions 0.9-4 or later is assumed for both). A practical overview of **VGAM** is Yee (2008) and currently Yee (2015) is being written for practitioners.

5.1 Goodman's RC fitted to alcoholic offences

The first example concerns `alcoff`, a 24×7 table of alcohol offenders caught from breath screening drivers, accumulated over 2009 in New Zealand. Rows are hour of the day starting from midnight, and columns are the week days starting with Monday.

We fit an independence (rank-0 Goodman's RC) model to `alcoff`. The data are preprocessed by offsetting the data with respect to the hour. We say the *effective day* starts at 6am, say, since partying late at night often spills over to the early morning. Hence effective Mon starts at 6am and finishes on Tuesday at the same hour. The function `moffset()` (for matrix offset) enables us to accomplish this.
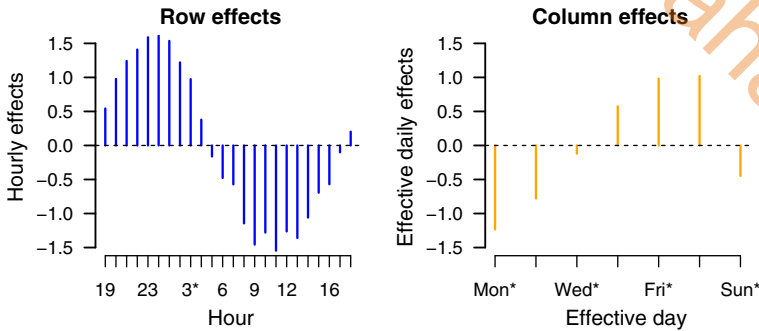
Fig. 2 Hourly and effective daily effects of a Goodman's RC model fitted to `alcoff`. This is output from `plot()`

The GRC model is fitted by

```
grc0.alcoff <- rcim(moffset(alcoff, "6", postfix = "*"),
                    rprefix = "Hour.24.", cprefix = "Day.")
```

Then applying

```
plot(grc0.alcoff, rcol = "blue", ccol = "orange",
     rfirst = 14, cfirst = 1, rtype = "h", ctype = "h",
     lwd = 2, ylim = c(-1.5, 1.5),
     cylab = "Effective daily effects",  # "c" for column, "xlab" for x-label
     rylab = "Hourly effects",           # "r" for row, "ylab" for y-label
     rxlab = "Hour", cxlab = "Effective day")
```

the methods `plot()` function to the fitted RCIM gives Fig. 2 which displays the fitted main effects. It may be seen that the daily effects are comparable in size to the hourly effects. The results agree with what is expected: the greatest number of alcohol-related offences occur on Friday and Saturday nights (and their following morning), and there is a gradual increase from Sunday/Monday to these peak days. Also, they are at their lowest in the late morning to lunchtime period.

5.2 Median polish fitted to crash injuries

Of the several crash data sets in **VGAM** the data frame `crashi` (crash injuries) have the largest counts in general, therefore we apply median polish to obtain a robust additive fit to the frequencies.

```
mp0.crashi <- rcim(crashi, alaplace2(tau = 0.5, intparloc = TRUE))
plot.mp0.crashi <-
 plot(mp0.crashi, lwd = 2, hlwd = 2,
      hcol = "purple", rcol = "blue", ccol = "darkorange",
      rtype = "p", ctype = "p", cylab = "Daily effects",
      rylab = "Hourly effects", rxlab = "Hour", cxlab = "Day")
```
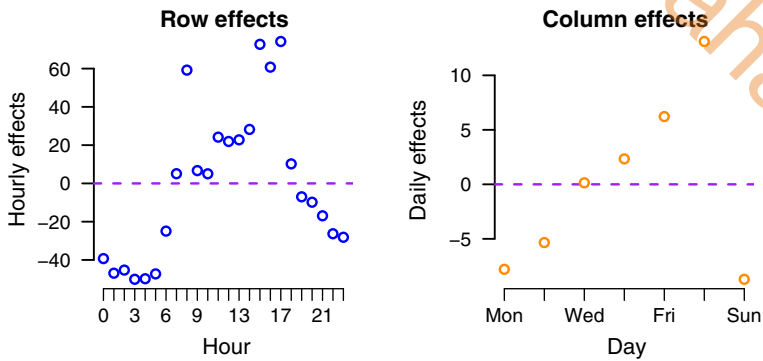
Fig. 3 Median polish fit to crash injuries, `crashi`. This is output from `plot()`. The *y* axes differ in scale

This gives Fig. 3. An identity link is effective, and the fit involves medians. It can be seen that hourly effects are more dominant than the daily effect. The largest hourly effects are from 0800–0900 and 1500–1700 which corresponds to rush hour traffic and includes the start and finish of school. It appears that Sunday and Monday are approximately equal, then a linearly increasing effect from Monday to Saturday. Not surprisingly, the least injuries occur during the very early hours of the morning.

5.3 Quasi-variances example

We illustrate quasi-variances here for a simple $M = 1$ Poisson model on the ships data (McCullagh and Nelder 1989) where the damage rate on a log scale is regressed upon 5 ship types (A–E). Internally, the preprocessing function `Qvar()` coerces a fitted `vglm()` Poisson model into a square matrix as input for `uninormal()` (the 'responses' mentioned in Sect. 3.2).

From a user's viewpoint, the relevant code is of the form

```
Shipmodel <- vglm(incidents ~ type + year + period, quasipoissonff,
                  data = ships, subset = (service > 0),
                  offset = log(service))  # A fitted VGLM
fit1 <- rcim(Qvar(Shipmodel, "type"), family = uninormal("explink"))
quasiVar <- qvar(fit1)  # The quasi-variances
plotqvar(fit1)  # A simple plot of the quasi-variances
```

This results in a plot similar to Fig. 4b. Here, it may be seen that the 5 % LSD intervals for the first two levels are somewhat close to crossing, therefore we expect their *p* value to be somewhat close to 0.05. Indeed, testing the second coefficient from the original fit, one has a Wald statistic of -1.96, which borders on exactly the 5 % level of significance. Visually, two other pairs appear significantly different, viz. B versus E, and C versus E.
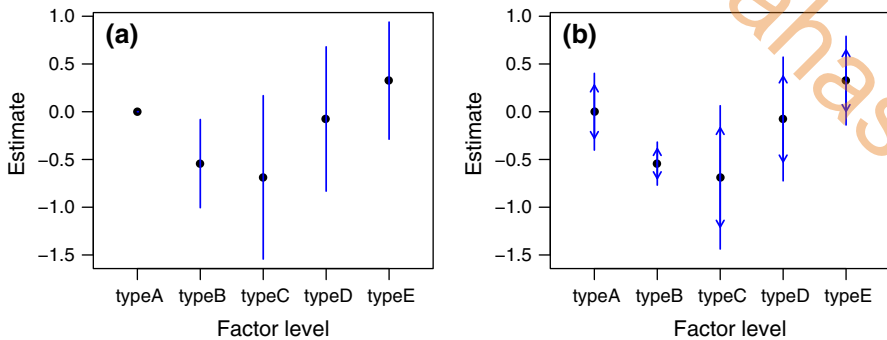
Fig. 4 a Standard plot for estimates of a factor effect whose first level is baseline; the *lines* denote ± 2 SEs. **b** 5 % LSD intervals (*arrows*) based on quasi-variances; these are $\widehat{\beta}_i \pm z(0.025)\sqrt{q_i/2}$ and $\widehat{\beta}_i \pm 2\sqrt{q_i}$ respectively. The fitted model is a quasi-Poisson-GLM fitted to the `ships` data

5.4 Zero-inflated Poisson example

Here we fit quasi-variances to a $M = 2$ model. We fit a zero-inflated Poisson to a subset of 2,845 women from a large cross-sectional study of New Zealanders collected during the early 1990s. The data frame is `xs.nz` in **VGAMdata**. The response is the number of births given, and the factor of interest is ethnicity, which has values 0 for European (NZ European or British or other European), 1 for Maori and 2 for Pacific Island Polynesian. For homogeneity, we exclude the "Other" group (Chinese, Indian and other ethnicities). It was noted that a ZIP model is more reasonable than a Poisson because a simple barchart (not given) revealed a surplus of 0s relative to a Poisson distribution (so that there is bimodal behaviour).

The model

$$\eta_1 = \log\lambda \;\; = \beta_{(1)1} + f_{(1)2}(\text{age}) + \text{ethnicity}_1 \cdot \beta_{(1)3} + \text{ethnicity}_2 \cdot \beta_{(1)4}, \quad (21)$$
$$\eta_2 = \text{logit}\,\phi^* = \beta_{(2)1} + f_{(2)2}(\text{age}) + \text{ethnicity}_1 \cdot \beta_{(2)3} + \text{ethnicity}_2 \cdot \beta_{(2)4} \quad (22)$$

is fitted since the first level of a factor is zero by default. This model may be estimated using regression splines by placing the appropriate constraint matrices $\mathbf{H}_k$ [see (6)] into `clist`:

```
clist <- list("bs(age, df = 4)"  = rbind(1, 0),
              "bs(age, df = 3)"  = rbind(0, 1),
              "ethnic"           = diag(2),
              "(Intercept)"      = diag(2))
fit1 <- vglm(babies ~ bs(age, df = 4) + bs(age, df = 3) + ethnic,
             zipoissonff(zero = NULL), data = xs.nz.f,
             constraints = clist)
```

Here, we provide a little more flexibility to the function of λ as opposed to ϕ.

Now the quasi-SE plots can be obtained by

```
Fit1 <- rcim(Qvar(fit1, "ethnic", which = 1),  # For log(lambda)
```
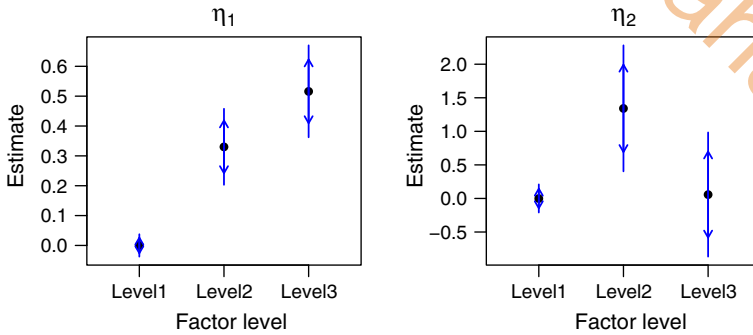
Fig. 5 Quasi-SE plots for $\widehat{\eta}_1$ and $\widehat{\eta}_2$ for a zero-inflated Poisson model (21)–(22) fitted to a subset of females from the dataframe `xs.nz` with `babies` as the response

```
                 uninormal("explink"), maxit = 99)
Fit2 <- rcim(Qvar(fit1, "ethnic", which = 2),  # For logit(phi)
                 uninormal("explink"), maxit = 99)
plotqvar(Fit1, scol = "blue", pch = 16,  # Most are graphical parameters
         main = expression(eta[1]),
         slwd = 1.5, las = 1, length.arrows = 0.07)
plotqvar(Fit2, scol = "blue", pch = 16,
         main = expression(eta[2]),
         slwd = 1.5, las = 1, length.arrows = 0.07)
```

This results in Fig. 5. One might interpret a structural zero as the inability or decision not to have children. Adjusting for age, the RHS plot indicates that the probability of a structural zero is lowest for Maori, and the European and Polynesians are indistinguishable. That is, Maori are most available to have the possibility of children.

The LHS plot shows that, amongst those who could possibly have children, Europeans have the least number of children on average, followed by Maori, and Polynesians have the most. But while the Europeans are significantly lower than the other two ethnicities, the other two groups may not be significantly different.

5.5 Goodman's RC and unconstrained quadratic ordination

This example applies the Goodman's RC model to multispecies counts to solve an important problem in statistical ordination in ecology: the Goodman's RC model estimates the site scores and optimums of an equal-tolerances Poisson *unconstrained quadratic ordination* (UQO).

To see this, as background, suppose the complete data are **Y** and **X** where the latter is a $n \times p$ matrix of p environmental variables measured at n sites. The $n \times q$ matrix **Y** are species counts for q species at the n sites. Suppose the true model generating the abundances is

$$\log \mu_{ij} = \eta_{ij} = A_j - \frac{1}{2} \left(\frac{v_i - u_j}{t_j} \right)^2 \tag{23}$$

where $v = \mathbf{c}^T \mathbf{x}$ is the latent variable or underlying gradient. The quadratic term means a bell-shaped response curve with respect to v—this is biologically sensible because a species typically dies when the gradient (e.g., temperature) is too low or too high. Parameters u_j and t_j are the jth species' *optimum* and *tolerance* respectively. The latter is a measure of niche width, which is a biologically meaningful quantity. The species thrives about its optimum environment.

The above scenario fits into the problem of ordination in statistical ecology, where species (columns) and sites (rows) are ordered or arranged so that their interrelationships can be displayed in a special type of plot called a biplot (Gower et al. 2011). Equation (23) is known as a *constrained quadratic ordination* (CQO) because the site scores v_i are constrained to be a linear combination of the environmental variables $\mathbf{x}_i$. The unconstrained case occurs when the data is solely $\mathbf{Y}$, and hence the v_i are treated as free parameters. The ordination problem involves estimating the parameters v_i, u_j, A_j and t_j for $i = 1, \ldots, n$ and $j = 1, \ldots, q$.

Suppose that all species' tolerances are equal in the unconstrained case. Without loss of generality set $t_j = 1$ for all j. Then (23) can be expanded to

$$\eta_{ij} = -\frac{1}{2}v_i^2 + \left(A_j - \frac{1}{2}u_j^2\right) + v_i\, u_j. \tag{24}$$

This can be compared to a rank-1 Goodman's RC model

$$\log \mu_{ij} = \eta_{ij} = \beta_0 + \alpha_i + \gamma_j + c_i\, a_j, \tag{25}$$

and the parameters matched as

$$a_j = u_j \text{ and } c_i = v_i \tag{26}$$

(actually, the RCIM is slightly overparameterized but we focus on the cross-product term.) Thus a rank-1 Goodman's RC model may be used to fit an equal-tolerances Poisson UQO. More generally, a rank-1 RCIM might be used to estimate the site scores and optimums of an equal-tolerances UQO for a range of data types since both models share a common η_{ij}.

Here is the methodology illustrated using simulated data.

```
set.seed(123)  # For reproducibility
n <- 100; p <- 5; S <- 5  # S is the number of species
pdata <- rcqo(n, p, S, es.opt = FALSE, eq.max = FALSE, eq.tol = FALSE,
              sd.tolerances = 1, sd.latvar = 0.75)  # Random X and Y
true.nu <- attr(pdata, "latvar")  # The 'truth'
Y <- pdata[, paste("y", 1:S, sep = "")]  # Y matrix (n x S)
uqo.rcim1 <- rcim(Y, Rank = 1, str0 = NULL,  # Delta covers entire n x M matrix
                  iindex = 1:nrow(Y),  # RRR covers the entire Y
                  has.intercept = FALSE)  # Suppress the intercept
```

This fit uses a slightly new but equivalent parameterization: it has $\beta_0 \equiv 0$ and no structural zeros so that the rank-1 approximation $\boldsymbol{\Delta}$ covers the entire $n \times M$ matrix of η_{ij}s.

The estimated species' optimums can be plotted against the true optimums as follows (and ditto for the site scores).

```
# Plot "(a)" for the optimums
plot(attr(pdata, "optima"), Coef(uqo.rcim1)@A,
     col = "blue", type = "p", main = "(a)")
mylm <- lm(Coef(uqo.rcim1)@A ~ attr(pdata, "optima"))
abline(coef = coef(mylm), col = "orange", lty = "dashed")

# Plot "(b)" for the site scores
fill.val <- NULL  # Choose this for the new parameterization
plot(attr(pdata, "latvar"), c(fill.val, concoef(uqo.rcim1)),
     las = 1, col = "blue", type = "p", main = "(b)")
mylm <- lm(c(fill.val, concoef(uqo.rcim1)) ~ attr(pdata, "latvar"))
abline(coef = coef(mylm), col = "orange", lty = "dashed")
```

This gives Fig. 6a, b. It may be seen that there is close correspondence between the fitted RCIM and the truth. The estimates are linearly related to the exact values because of some scaling that has not yet been performed. In Fig. 6b the correlation between the true site scores and the estimated UQO site scores is 0.873 which indicates good agreement.

Given a UQO fit, one can then fit a CQO on the $\widehat{v_i}$ as explanatory variables without making an equal tolerances assumption. This can be done here with

```
myform <- attr(pdata, "formula")
p1ut <- cqo(myform, family = poissonff,
            eq.tol = FALSE, trace = FALSE, data = pdata)
c1ut <- cqo(cbind(y1, y2, y3, y4, y5) ~ scale(latvar(uqo.rcim1)),
            family = poissonff, eq.tol = FALSE, trace = FALSE, data = pdata)
lvplot(p1ut, lcol = 1:S, y=TRUE, pcol = 1:S, pch = 1:S, pcex = 0.5, main = "(c)")
lvplot(c1ut, lcol = 1:S, y=TRUE, pcol = 1:S, pch = 1:S, pcex = 0.5, main = "(d)")
```

Here, `p1ut` is a CQO fit of the original data, and `c1ut` is a CQO fit of the UQO fit. The two plots are very similar, which is to be expected with such 'nice' data.

6 Discussion

Two way tables are a very common form of data, therefore it is not surprising that some overlapping methodology has been proposed already. In particular, the *generalized additive main effects and multiplicative interaction models*, or GAMMI models (see, e.g., a description in Turner and Firth 2007). These also comprise the row and column main effects plus one or more components of the multiplicative interaction. The singular value corresponding to each multiplicative component is often factored out, as a measure of the strength of association between the row and column scores, indicating the importance of the component, or axis. For cell means μ_{ij} a GAMMI-R model has the form

$$g(\mu_{ij}) = \alpha_i + \beta_j + \sum_{r=1}^{R} \sigma_r \, \gamma_{ri} \, \delta_{rj}. \tag{27}$$
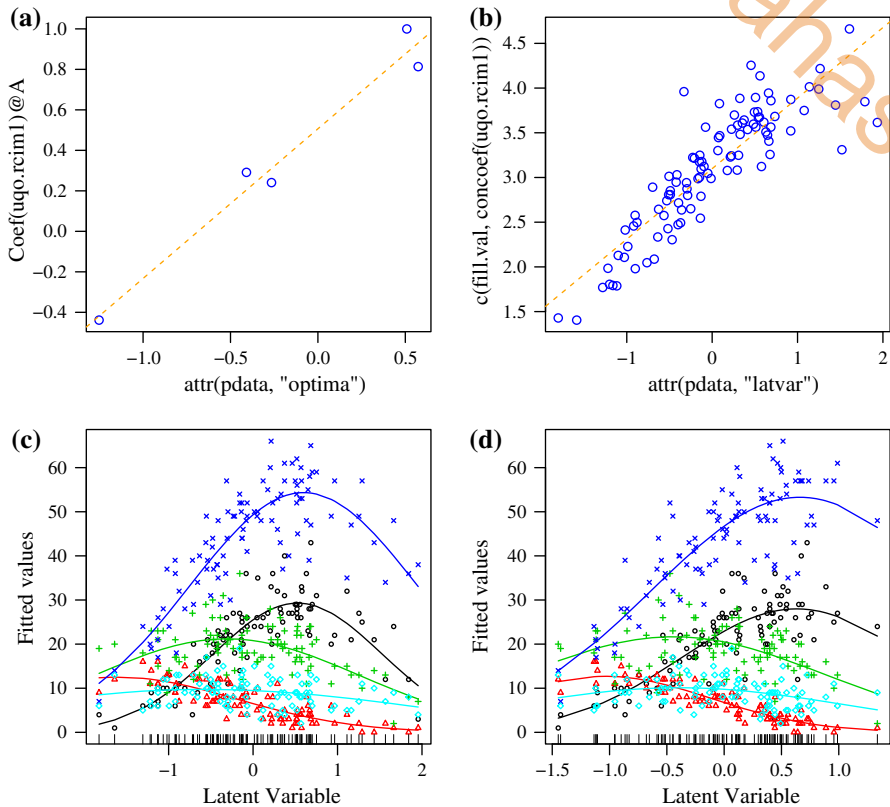
Fig. 6 **a**, **b** Compares the UQO solution with the truth for a rank-1 Poisson simulated data set. **a** Estimated optima $\widehat{u}_j$ versus u_j. **b** Estimated site scores $\widehat{v}_i$ versus v_i. **c** CQO fitted to the original data. **d** CQO fitted to the scaled UQO site scores. In **a**, **b** the *dashed orange line* is a simple linear regression through the points (colour figure online)

Based on (27) GAMMI models appear to be identical to RCIMs. In (27) their interaction term is a singular value decomposition of our $\mathbf{AC}^T$. While our interaction term uses corner constraints, their parameterization is quite interpretable and is related to some of the other parameterizations described in Yee and Hastie (2003).

Another class of models with (10) applied to an error distribution from the exponential family has been called *generalized biadditive models* (GBMs; see e.g., de Rooij 2007). But like GAMMIs, the class of VGLMs is much larger than the classical exponential family of GBMs.

In summary, RCIMs are a variant of RR-VGLMs tailored specifically towards two-way table responses. Each row of the response is modelled simultaneously via $\boldsymbol{\eta}_i$, and the appropriate row and column dummy variables (20) are set up. Thus RCIMs are RR-VGLMs where one of the linear predictors is modelled as the sum of a row effect, a column effect, and an optional interaction effect which is expressed multiplicatively as a RRR. Graphical techniques, when there are no interactions, are available to display the main effects on the transformed scale (4).

The VGLM/VGAM framework is very large and covers many models and distributions from survival analysis, extreme value analysis, quantile and expectile regression, time series, bioassay data, nonlinear least-squares models, and scores of standard and nonstandard univariate and continuous distributions. RCIMs are a particularly useful subclass of VGLMs that facilitates the analysis of two-way tables of many data types. Consequently it is hoped that RCIMs prove useful in many areas of applied statistics.

References

Andrews HP, Snee RD, Sarner MH (1980) Graphical display of means. Am Stat 34:195–199

de Rooij M (2007) The distance perspective of generalized biadditive models: scalings and transformations. J Comput Graph Stat 16:210–227

Easton DF, Peto J, Babiker AGAG (1991) Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat Med 10:1025–1035

Firth D (2000) Quasi-variances in Xlisp-Stat and on the web. J Stat Softw 5:1–13, http://www.jstatsoft.org/v05/i04

Firth D, de Menezes RX (2004) Quasi-variances. Biometrika 91:65–80

Goodman LA (1981) Association models and canonical correlation in the analysis of cross-classifications having ordered categories. J Am Stat Assoc 76:320–334

Gower JC, Lubbe SG, Le Roux NJ (2011) Understanding biplots. Wiley, Chichester

Kotz S, Kozubowski TJ, Podgórski K (2001) The laplace distribution and generalizations: a revisit with applications to communications, economics, engineering, and finance. Birkhäuser, Boston

McCullagh P, Nelder JA (1989) Generalized linear models, 2nd edn. Chapman & Hall, London

Mosteller F, Tukey JW (1977) Data analysis and regression. Addison-Wesley, Reading, MA

Powers DA, Xie Y (2008) Statistical methods for categorical data analysis, 2nd edn. Bingley, Emerald

Schenker N, Gentleman JF (2001) On judging the significance of differences by examining the overlap between confidence intervals. Am Stat 55:182–186

Scott D (2012) GeneralizedHyperbolic: the generalized hyperbolic distribution. http://CRAN.R-project.org/package=GeneralizedHyperbolic, R package version 0.8-1

Turner H, Firth D (2007) g*nm*: a package for generalized nonlinear models. R News 7:8–12, http://CRAN.R-project.org/doc/Rnews/

Yee TW (2008) The *VGAM* package. R News 8:28–39, http://CRAN.R-project.org/doc/Rnews/

Yee TW (2010) The *VGAM* package for categorical data analysis. J Stat Softw 32:1–34, http://www.jstatsoft.org/v32/i10/

Yee TW (2014) Reduced-rank vector generalized linear models with two linear predictors. Comput Stat Data Anal 71:889–902

Yee TW (2015) Vector generalized linear and additive models. Springer, NY (in preparation)

Yee TW, Hastie TJ (2003) Reduced-rank vector generalized linear models. Stat Model 3:15–41

SJR   Scimago Journal & Country Rank

Enter Journal Title, ISSN or Publisher Name

Home   Journal Rankings   Country Rankings   Viz Tools   Help   About Us

Computational Statistics

COUNTRY

Germany

Universities and research institutions in Germany

Media Ranking in Germany

SUBJECT AREA AND CATEGORY

Decision Sciences
  Statistics, Probability and Uncertainty

Mathematics
  Computational Mathematics
  Statistics and Probability

PUBLISHER

Springer Verlag

H-INDEX

46

Ad clo
Go

PUBLICATION TYPE

Journals

ISSN

09434062

COVERAGE

1996-2021

INFORMATION

Homepage

How to publish in this journal

w-sakamoto@okayama-u.ac.jp

Rahasia

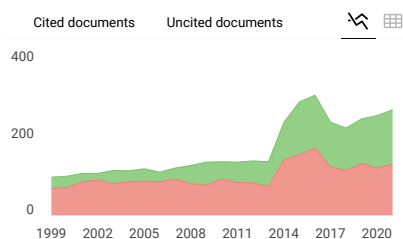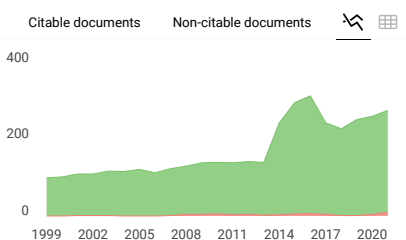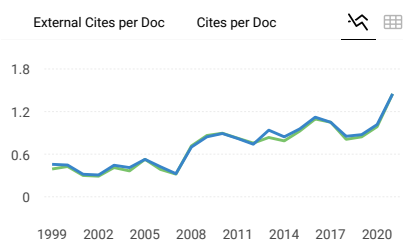| 1	2	3	4	5
Statistical Methods and Applications	**Journal of Applied Statistics**	**AStA Advances in Statistical Analysis**	**Journal of Statistical Computation and Simulation**	**Statistica Neerlandica**
DEU	GBR	DEU	GBR	GBR
85%	**77%**	**76%**	**75%**	**74%**
similarity	similarity	similarity	similarity	similarity

SJR

Total Documents

Total Cites   Self-Cites

Citations per document

External Cites per Doc   Cites per Doc

% International Collaboration

Citable documents   Non-citable documents

Cited documents   Uncited documents

- Cites / Doc. (4 years)
- Cites / Doc. (3 years)
- Cites / Doc. (2 years)

Computational Statistics

Q2   Computational Mathematics

best quartile

SJR 2021
0.54

powered by scimagojr.com

← Show this widget in your own website

Just copy the code below and paste within your html code:

`<a href="https://www.scimag`

G SCImago Graphica

Explore, visually communicate and make sense of data with our **new data visualization tool**.