Multiplicative mixed models for the analysis of sensory evaluation data

Brian Cullis[1], Alison Smith[1], Robin Thompson[2]

[1] NSW Agriculture, Wagga Wagga, Australia

[2] IACR-Rothamsted, UK

Wine tasting experiment

- Project conducted in Germany to compare ecologically and conventionally produced Riesling wines. Data kindly provided by Ulrich Fischer.
- 56 wines (28 of each style) evaluated by 22 judges for sensory traits.

Wine tasting experiment

- A 2-phase experiment:
  - Phase I: wines of each style selected
  - Phase II: wines evaluated by judges
- Each phase characterised by a design - should be accounted for in the analysis.

Wine tasting experiment

- Phase I: stratified sampling of wines
  - Wines sampled from 30 wineries
  - Wineries located in one of 4 areas
- Each winery is classified as ecological or conventional so experimental unit for style treatment is winery (not wine). Inclusion of winery stratum in analysis is crucial.

Wine tasting experiment

- Phase II: wines allocated to judges with some balance but details unknown.
  - Not all judges assessed all wines: 809 (66%) of the possible 56*22 combinations are present.
  - Only 2 judges assessed all wines; 1 judge assessed 8 wines only. Each judge assessed both styles of wine (roughly equal proportions)
  - Each wine assessed by 14 or 15 judges
  - 2 or 3 replicates (glasses?).

Basic 2-phase ANOVA

source	decomposition		fixed/random
judge			R
wine			
	area		R
	winery (within area)		
		style	F
		residual	R
	residual wine		R
judge.wine			
	judge.area		R
	judge.winery		
		judge.style	R
		residual	R
	residual judge.wine		R
residual			R

Basic 2-phase ANOVA

- A linear mixed model

  $$y = X\beta + Zu + e$$

- β = vector of fixed effects
- u = vector of random effects comprising 8 independent sub-vectors, u_s ($s=1\ldots8$), with

  $$\text{var}(u_s) = \sigma_s^2 I$$

- e = vector of residuals with

  $$\text{var}(e) = \sigma^2 I$$

Basic 2-phase model

- Trait analysed: "artificialfruit.pineapple"
- Scored on a continuous scale from 0 to 10. Concerned about Normality. Diagnostics suggested use of

  logafp = log(artificialfruit.pineapple+0.01)
- Define factors: judge (22), wine (56), area(4), winery (30), style (2)

Basic 2-phase model

- logafp ~ style *judge area winery wine judge.area judge.style judge.winery judge.wine*
- no significant difference between styles:
  style 1 mean = -0.934
  style 2 mean = -0.808
  sed = 0.141

Basic 2-phase model

source	decomposition		variance comp
judge			0.353
wine			
	area		0.047
	winery (within area)		
		style	F
		residual	0.038
	residual wine		0.119
judge.wine			
	judge.area		0.381
	judge.winery		
		judge.style	0
		residual	0
	residual judge.wine		0.317
residual			1.402

Multiplicative models for 2-way random effects

- Large judge.area and residual judge.wine variation so investigate further - use multiplicative models – as an example consider the wine.judge term
- Let $U^{(mxp)} = \{u_{ij}\}$ denote matrix of random wine.judge effects, m=56 and p=22
- For i=1..m, j=1..p we write

$$u_{ij} = \lambda_{1i} f_{1j} + \lambda_{2i} f_{2j} + \dots + \lambda_{ki} f_{kj} + \delta_{ij}$$

Multiplicative mixed models

$$u_{ij} = \lambda_{1i} f_{1j} + \lambda_{2i} f_{2j} + \dots + \lambda_{ki} f_{kj} + \delta_{ij}$$

- Like a regression on k x-variables $\lambda_1 \dots \lambda_k$ with separate coefficients $f_1 \dots f_k$ for each level of factor A.
- But both coefficients and x-variables are unknown so must be estimated from data. Thus a multiplicative model of "loadings" (λ_{ri}) and "scores" (f_{rj}). Residual or lack of fit of multiplicative model given by δ_{ij}

Multiplicative mixed models

- In matrix notation, let $u^{(mp \times 1)} = vec(U)$:

  $$u = (\Lambda \otimes I_m) f + \delta$$

  where $\Lambda^{(p \times k)}$ is matrix of loadings, $f^{(mk \times 1)}$ and $\delta^{(mp \times 1)}$ are vectors of scores, residuals.
- Assume f and δ independent random effects

  $$var(f) = I_k \otimes I_m$$

  $$var(\delta) = diag(\psi_1 \ldots \psi_p) \otimes I_m$$
- Let $\Psi = diag(\psi_1 \ldots \psi_p)$ – specific variances for each judge
- then …

Multiplicative mixed models

$$\text{var}(u) = (\Lambda \Lambda^T + \Psi) \otimes I_m$$

- This is Factor Analytic (FA-k) variance model used extensively for modelling variety x environment interaction (Piepho;Smith et al)
- Models the variance/covariance matrix between judges – allows for agreement/disagreement, scaling (see later)

Multiplicative mixed models

- Estimate parameters of multiplicative model as part of overall mixed model analysis (one-stage analysis)
- Obtain REML estimates of variance parameters (including L and Y)
- Obtain BLUEs of fixed effects
- Obtain BLUPs of random effects (including f)

Multiplicative mixed models

- Goodness of fit (number of multiplicative terms needed) easily tested using REMLRT
- Interpretation aided with use of bi-plots of scores and loadings
- Provides parsimonious variance model
  - Variance for judge i

    $\lambda_{1i}^2 + \ldots + \lambda_{ki}^2 + \psi_i$
  - Covariance between judge i and judge l

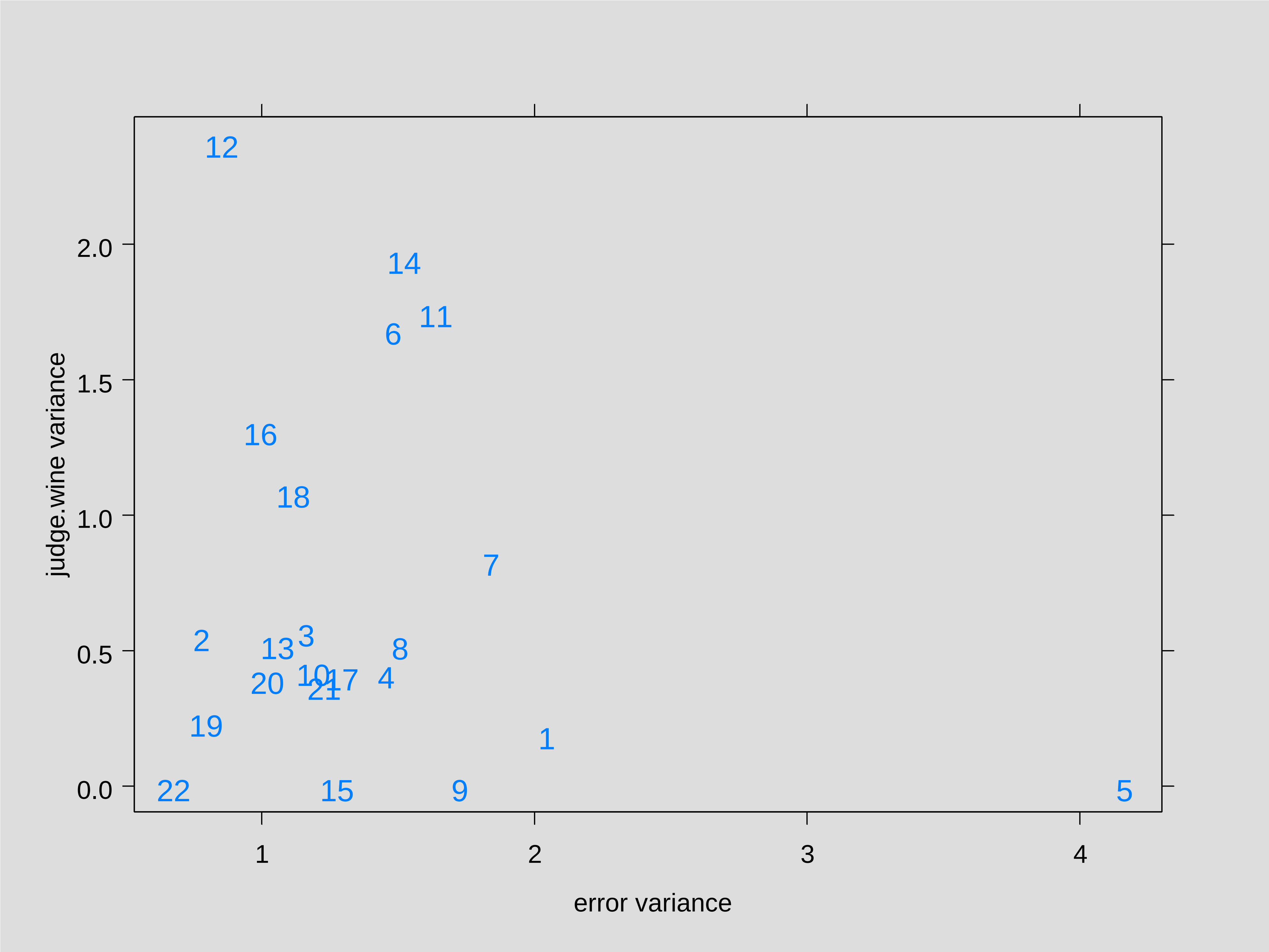
    $\lambda_{1i}\,\lambda_{1l} + \ldots + \lambda_{ki}\,\lambda_{kl}$

Wine tasting experiment: exploring sources of variation

- Large error variance: examine heterogeneity between judges (judge reliability)
- Large judge.area: multiplicative model with FA variance structure for area dimension
- Large residual judge.wine: multiplicative model with FA variance structure for judge dimension

Exploring sources of variation

- Begin by checking for judges with no wine variance, ie. no discriminating ability. Difficult to model these so omit from final analysis.
- Fit model which is analogous to a separate analysis for each judge:
  logafp ~ judge *judge.wine*
  - Separate error variance for each judge
  - Separate judge.wine variance for each judge



judge.wine variance
error variance
0.0
0.5
1.0
1.5
2.0
1
2
3
4
12
14
11
6
16
18
7
2
13
3
8
20
10
21
17
4
19
1
22
15
9
5

Exploring sources of variation

- Judges 5, 9, 15 and 22 do not discriminate between wines they assessed.
  - judge 5: large error variance
  - judge 22: only assessed 8 wines
  - judges 9, 15?
- Omit 4 judges from analysis. Re-fit of basic 2-phase ANOVA gave almost identical variance components

Exploring sources of variation

- Error variance heterogeneity between judges was significant: range from 0.715 (judge 19) to 2.584 (judge 11)

model	logl	REMLRT	df	p-value
basic	-1208.31			
error heterogeneity	-1177.66	61.30	17	<0.001

Exploring sources of variation

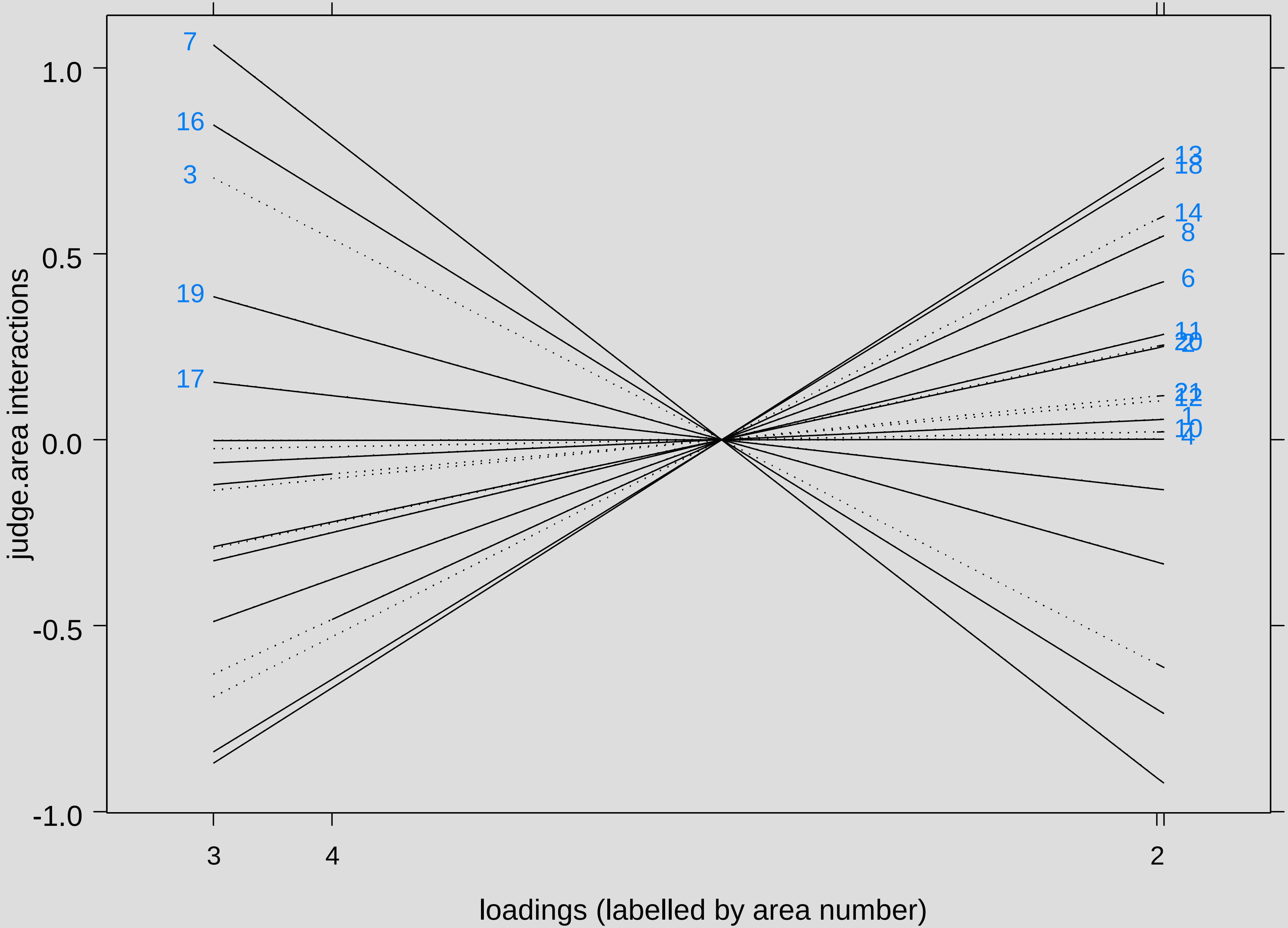
- Multiplicative model for area.judge interaction was significant

model	logl	REMLRT	df	p-value
basic	-1208.31			
error heterogeneity	-1177.66	61.30	17	<0.001
FA(1) area.judge	-1157.96	39.40	7	<0.001

Multiplicative model for area.judge

- loadings contrast areas 1,2 with 3,4
- area.judge interaction almost entirely explained by multiplicative model (only one non-zero lack of fit variance)

area	loading	lof variance
1	0.56	0
2	0.55	0
3	-0.64	0.38
4	-0.49	0



judge.area interactions
1.0
0.5
0.0
-0.5
-1.0
3
4
2
loadings (labelled by area number)
7
16
3
19
17
13
18
14
8
6
11
20
2
21
12
1
10
4

Exploring sources of variation

- Multiplicative model for residual judge.wine interaction was significant

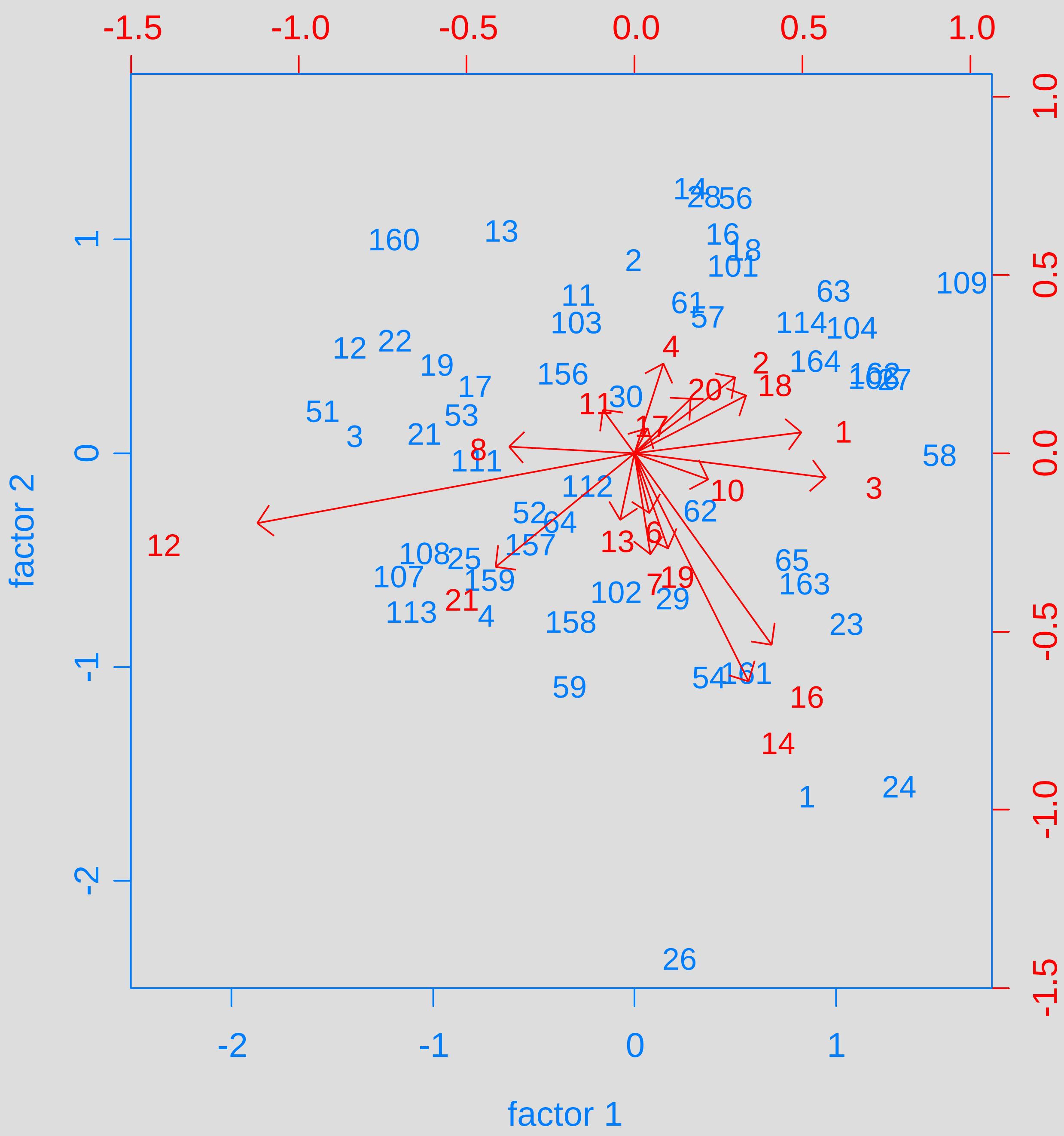
model	logl	REMLRT	df	p-value
basic	-1208.31			
error heterogeneity	-1177.66	61.30	17	<0.001
FA(1) area.judge	-1157.96	39.40	7	<0.001
FA(1) judge.wine	-1120.59	74.74	35	<0.001

Exploring sources of variation

- Addition of second multiplicative term for residual judge.wine interaction was significant

model	logl	REMLRT	df	p-value
basic	-1208.31			
error heterogeneity	-1177.66	61.30	17	<0.001
FA(1) area.judge	-1157.96	39.40	7	<0.001
FA(1) judge.wine	-1120.59	74.74	35	<0.001
FA(2) judge.wine	-1103.83	33.52	17	0.010

judge	load1	load2	lof variance
1	0.62		0
2			0.19
3	0.71		0
4			0.33
6			0.66
7			0
8	-0.47		0
10			0
11			1.33
12	-1.41		0
13			0
14	0.42	-0.80	0
16	0.51	-0.67	0
17			0.21
18	0.42		0
19			0
20			0
21	-0.52		0



-1.5
-1.0
-0.5
0.0
0.5
1.0
factor 2
1
0
-1
-2
-2
-1
0
1
factor 1
1.0
0.5
0.0
-0.5
-1.0
-1.5

Mixed model analysis of wine data

- Style treatment differences tested. Inclusion of appropriate strata important
- Magnitude of sources of variation quantified: large error variation, judge.area and residual judge.wine interactions
- Error variance heterogeneity between judges modelled
- Structure in judge.area and residual judge.wine interactions identified using multiplicative models

Mixed model analysis of wine data

- Multiplicative model for residual judge.wine interactions is a form of “assessor” model
- Assessor models (eg. Brockhoff and Skovgaard, 1994) are designed to investigate:
  - reliability of each assessor
  - discriminating ability of each assessor
  - agreement between assessors

Mixed model analysis of wine data

- In terms of our mixed model:
  - assessor reliability: heterogeneity of error variance between judges
  - assessor discriminating ability: heterogeneity of judge.wine variance between judges
  - assessor agreement: heterogeneity of judge.wine correlations between pairs of judges
- Last two accommodated in Factor Analytic variance structure that results from the multiplicative model

Summary

- Multiplicative mixed model (MMM) appears useful for analysis of sensory evaluation data (eg. for product by assessor studies)
- The MMM is a random effects analogue of PCA (with error). Advantages:
  - unbalanced data and complex design structure easily handled
  - significance testing straightforward (including tests for number of multiplicative terms)

Summary

- Extensions to 3-way models (eg. for product by assessor by attribute) subject of current research
- Software – models can be fitted in GENSTAT, functions within Splus or ASReml