Software Tools

PERFORMANCE OF REML ALGORITHMS IN MULTIVARIATE ANALYSES FITTING REDUCED RANK AND FACTOR-ANALYTIC MODELS

Karin Meyer

Animal Genetics and Breeding Unit[1], University of New England, Armidale, NSW 2351

SUMMARY

Convergence behaviour of restricted maximum likelihood algorithms in multivariate analyses imposing a factor-analytic structure on covariance matrices is examined. Results indicate that estimation for such models can entail a more difficult maximisation problem than 'unstructured' estimation. On the other hand, if only factors explaining negligible variation are omitted, convergence can be faster as parameters at the boundaries of the parameter space have been eliminated. The 'parameter expanded' expectation maximisation algorithm tends to require many more iterates than the 'average information' algorithm, but is useful, in particular when combined with the latter.

INTRODUCTION

Factor-analytic (FA) models have been proposed to describe the dispersion among numerous traits or random regression coefficients parsimoniously. This includes reduced rank (RdR) estimation as the special case where specific effects and their variances are assumed to be zero. Thompson et al. (2003) and Meyer and Kirkpatrick (2005) described 'average information' (AI) restricted maximum likelihood (REML) algorithms for such models. Imposing a FA structure on covariance matrices reduces the number of parameters to be estimated and, for RdR estimation, the computational requirements per REML iterate. However, initial applications (Meyer 2005, 2007a) have encountered slow convergence and required more iterates to locate the maximum of the likelihood function ($\log \mathscr{L}$) than comparable analyses estimating full rank, 'unstructured' covariance matrices. Carrying out a few, initial iterates of the 'parameter expanded' expectation maximisation (PX) algorithm (Liu et al. 1998) before switching to the AI algorithm has been found to improve convergence for full rank analyses (Cullis et al. 2004; Meyer 2006a). Recently, Meyer (2007b) outlined an adaptation of the PX algorithm for RdR and FA models. This paper examines convergence rates of the AI and PX algorithms for two practical examples.

MATERIAL AND METHODS

Data. Example 1 comprised records for 6 carcass traits recorded on Angus heifers or steers, as part of a meat quality research project by the Australian Co-operative Research Centre for Cattle and Beef Industry; see Reverter et al. (2000) for details. Example 2 considered weights for calves in the Wokalup selection experiment, recorded on a monthly basis from birth to weaning; see Meyer (2001). Table 1 summarises characteristics of the data structure.

Table 1. Data characteristics

	Set 1	Set 2
No. records	9487	21807
No. animals	1796	3769
No. parents	3105	203
No. parameters	27–42	43–73

Analyses. Analyses for Example 1 fitted a simple animal model. Genetic covariance matrices were estimated fitting increasing numbers of principal components (PC), assuming specific variances were zero. This yielded estimates of rank $1,\ldots,6$. In addition, FA models with

[1] AGBU is a joint venture of NSW Department of Primary Industries and the University of New England

Table 2. Number of iterates required (N) and deviation of $\log\mathscr{L}$ from best value (D $\times 1000$) for different convergence criteria (Δ), for Example 1

Δ		Fit 2		Fit 3		Fit 4		Fit 5		FA 1		FA 2	
		N	D	N	D	N	D	N	D	N	D	N	D
10^{-5}	AI	35	0	17	0	54	0	8	0	16	0	32	0
	PXAI	37	0	18	0	51	0	10	0	17	0	25	-234
	PX	1278	-3	1036	-1	866	-1	715	-2	332	0	6955	-239
0.0001	AI	29	0	15	0	45	0	7	0	13	0	24	0
	PXAI	35	0	16	0	42	0	9	0	16	0	23	-234
	PX	583	-29	761	-12	600	-11	459	-11	242	-3	547	-1199
0.0005	AI	25	-1	13	0	38	-2	6	0	11	0	23	0
	PXAI	26	-1	15	0	35	-2	8	0	15	0	22	-234
	PX	227	-106	574	-58	420	-56	321	-45	199	-13	333	-1248

1 or 2 factors and 6 specific genetic variances were fitted. The residual covariance matrix was assumed to be unstructured and have full rank throughout. Example 2 was analysed fitting a random regression (RR) model, with quadratic B-splines of age at recording (0 to 279 days) and knots at 0, 70, 140, 210 and 280 days as basis functions. This fitted 18 RR coefficients, 6 each for direct additive genetic and direct and maternal permanent environmental effects. Measurement error variances were assumed to change with age according to a step function with 10 classes (0, 1-60, 61-90, ..., 241-279 days). Analyses were carried out fitting 222, 333, 444, 453, 564 and 666 PCs for the 3 random effects, respectively. Furthermore, the genetic covariance matrix was estimated assuming a factor analytic structure with 1 or 2 factors, with the permanent environmental covariance matrices assumed to have rank 5 and 3 (analyses FA1 5 3 and FA2 5 3).

All analyses were performed using our REML program WOMBAT (Meyer 2006b). This implements an AI algorithm which enforces an increase of the likelihood in each iterates, using a backtracking line search to determine optimal step sizes. In addition, the average information matrix is ensured to be 'safely' positive definite and well-conditioned by adding a small constant to its diagonal if the smallest eigenvalue is less than $\min(0.002, 10^{-6} \times \lambda_1)$, with λ_1 the largest eigenvalue. For the same starting values, analyses were carried out using AI and PX steps only, and a combination of 4 initial PX steps followed by AI (PXAI). Iterations ceased when the change in $\log\mathscr{L}$ between iterates was less than 10^{-5} for an AI or less than 10^{-6} for a PX step. No 're-starts' were carried out.

RESULTS

Convergence statistics for Example 1 are summarised in Table 2. D is the deviation from the highest value for $\log\mathscr{L}$ found for a particular analysis, with a value of "0" indicating no difference to the third decimal.

Overall, imposing a structure on the estimated genetic covariance matrix tended to increase the number of iterates required. This trend was most evident when PCs with non-negligible eigenvalues were omitted. In comparison, the full rank, unstructured analysis required only 6 AI iterates to reach $\Delta = 10^{-4}$ with D=0. While the PX algorithm generally yielded larger improvements in $\log\mathscr{L}$ than AI in the

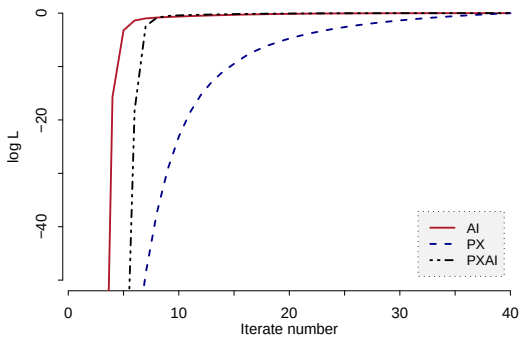

Figure 1. Convergence pattern for Fit 4.

first 2 or 3 iterates, combining PX and AI steps only reduced the total number of iterates in a few cases.

For all analyses, the PX algorithm required a substantial number of iterates to locate the maximum of $\log \mathscr{L}$. Again there was a tendency for the number of iterates to increase as the number of PCs fitted decreased. The typical pattern of convergence is illustrated in Figure 1. Improvements in $\log \mathscr{L}$ for the PX algorithm exhibited a steady decrease over iterates. Though the algorithm tended to reach proximity of the maximum of $\log \mathscr{L}$ relatively quickly, this resulted in a large number of iterates required overall.

The convergence criteria applied were very stringent. For practical purposes, a criterion of $\Delta = 0.0005$ appeared more than adequate for analyses using the AI algorithm. Accurate estimation with the PX algorithm, however, tended to require a criterion as strict as $\Delta = 10^{-5}$ or smaller.

Table 3 shows the convergence behaviour for the AI algorithm for Example 2. For this example, the PX algorithm required excessive numbers of iterates for all cases considered. With multiple random effects and a larger number of parameters to be estimated, the 'unstructured' analysis proved more difficult, though relatively good starting values were used. Analyses employed a parameterisation the Cholesky factor of the covariance matrices to be estimated. Taking logarithms of the diagonal elements in addition, the AI algorithm converged in 10 iterates (Fit 666) and none of the iterates required step size scaling. Without the logarithmic transformation, however, AI failed to locate the maximum of $\log \mathscr{L}$ and comprised numerous iterates which required severe step size modifications.

Results indicated that a RdR analysis with fit 453 would eliminate only PCs with eigenvalues close to zero. Even without log diagonals, convergence for this analysis was swift, while analysis 564 still involved PCs with small eigenvalues and thus required roughly twice the number of iterates. Reducing the numbers of PCs fitted further, the number of iterates required again increased.

DISCUSSION

Similar patterns of convergence have been found in other examples (Meyer 2007b). Results show that RdR estimation fitting the important PCs only can substantially reduce computational requirements, decreasing both the number of iterates required and the operations per iterate. However, care must be

Table 3. Convergence for Example 2 (see Table 2 for notation)

Δ		Fit 222		Fit 333		Fit 444		Fit 453		Fit 564		Fit 666		FA1 53		FA2 53	
		N	D	N	D	N	D	N	D	N	D	N	D	N	D	N	D
0.0001	AI	109	0	52	-1	19	0	11	0	25	0	10	-1	15	0	14	-257
	PXAI	110	0	62	-1	24	0	15	0	25	0	13	0	18	0	22	0
0.0005	AI	102	-1	40	-4	18	0	9	0	21	-1	9	-1	13	1	12	-257
	PXAI	103	-1	49	-4	22	0	13	0	21	-1	12	0	17	0	20	0

taken in selecting the numbers of PCs fitted and the parameterisation employed.

Clearly, structured estimation can represent a more difficult problem than full rank, unstructured analyses. The AI algorithm utilises second derivatives of $\log \mathscr{L}$ in determining its direction of search. This implies a quadratic approximation of the likelihood surface. Failure to perform well for RdR analyses involving few PCs indicates a deviation from the parabolic shape. This is supported by limited investigations of the profile likelihood surface for simple RdR analyses, which have encountered ridges and saddle points (Meyer and Kirkpatrick 2007). Step size modifications in the AI algorithm can induce slow, almost linear rates of convergence, but without strict control the AI algorithm is apt to fail altogether for this kind of analysis.

While the PX algorithm has been found to converge reliably, like most expectation maximisation type algorithms it tends to require many more iterates than the AI algorithm. This seems to hold even if the latter is applied in scenarios where $\log \mathscr{L}$ is not well approximated by a parabolic surface. As suggested by Cullis et al. (2004), the PX algorithm appears most useful in conjunction with the AI algorithm and for 'difficult' problems. For analyses fitting simple models and with good starting values for the variance components to be estimated, however, advantages of PXAI over AI for RdR analyses tend to be limited.

CONCLUSIONS

Results show that RdR estimation can reduce computational requirements of multivariate analyses. On the other hand, imposing a structure which fits too few factors can not only yield biassed estimates, but also increase the difficulty of the optimisation problem to be solved. The 'parameter expanded' expectation maximisation algorithm is a useful addition to our armoury, in particular in combination with other algorithms.

REFERENCES

Cullis, B. R., Smith, A. B. and Thompson, R. (2004) In: *Methods and Models in Statistics in Honour of Professor John Nelder, FRS*, Imperial College Press, London. 53.

Liu, C., Rubin, D. B. and Wu, Y. N. (1998) *Biometrika* **85**:755.

Meyer, K. (2001) *Genet. Select. Evol.* **33**:487.

Meyer, K. (2005) *Anim. Sci.* **81**:337.

Meyer, K. (2006a) *CD-ROM Eighth World Congr. Genet. Appl. Livest. Prod.* Comm. No. 24–15.

Meyer, K. (2006b) *CD-ROM Eighth World Congr. Genet. Appl. Livest. Prod.* Comm. No. 27–14.

Meyer, K. (2007a) *J. Anim. Breed. Genet.* **124**:50.

Meyer, K. (2007b) *Genet. Select. Evol.* (in press).

Meyer, K. and Kirkpatrick, M. (2005) *Genet. Select. Evol.* **37**:1.

Meyer, K. and Kirkpatrick, M. (2007) *Proc. Ass. Advan. Anim. Breed. Genet.* **17**:154.

Reverter, A., Johnston, D. J., Graser, H.-U., Wolcott, M. L. and Upton, W. H. (2000) *J. Anim. Sci.* **78**:1786.

Thompson, R., Cullis, B. R., Smith, A. B. and Gilmour, A. R. (2003) *Austr. New Zeal. J. Stat.* **45**:445.

ACKNOWLEDGEMENTS

This work was supported by grant BFGEN.100 B of Meat and Livestock Australia Ltd (MLA) and funds from the International Livestock Resources and Information Centre (ILRIC).