AN EVALUATION OF 'DEFLATION' TO IMPROVE CONVERGENCE RATES FOR SINGLE-STEP GENOMIC EVALUATION WITH THE HYBRID MODEL

K. Meyer and A.A. Swan

Animal Genetics & Breeding Unit*, University of New England, Armidale, NSW, 2351 Australia

SUMMARY

Single step genomic evaluation fitting a 'hybrid' model which combines marker effects for individuals with genotypes with breeding values for non-genotyped animals can readily accommodate large numbers of genotyped animals. However, iterative solution of the pertaining mixed model equations via a preconditioned gradient scheme has been reported to be afflicted by much slower convergence rates than the standard breeding value model. 'Deflation' of the coefficient matrix has been proposed as a second preconditioning step and shown to dramatically reduce numbers of iterations and computing time required. We describe its application for a set of sheep data. Results indicate that assignment of marker effects to subdomains in moderately sized chunks together with a separate treatment of genetic group effects could reduce total computing times by about a third.

INTRODUCTION

The single-step procedure for joint genetic evaluation of genotyped and non-genotyped animals has become routine in many livestock improvement schemes. Many implementations rely on extending the classic breeding value model (BVM) by combining the pedigree-based relationship matrix with estimates of genomic relationships. An equivalent alternative is the so-called hybrid model (HM) which fits marker effects instead of breeding values for genotyped animals (Fernando *et al.* 2016). This does not require the inverse of the genomic relationship matrix and thus readily accommodates large numbers of genotyped animals. However, initial experience with a preconditioned conjugate gradient (PCG) algorithm to solve the pertaining mixed model equations (MME) has been that convergence rates tended to be slow and that many iterates could be required. Recently, Vandenplas *et al.* (2018) showed that a second level of preconditioning – through a 'deflation' of the coefficient matrix in the MME – could dramatically improve convergence rates and demonstrated its effectiveness for a large, multi-trait analysis of dairy field data. This paper examines the scope of the resulting, deflated preconditioned gradient (DPCG) solver for a practical sheep data set.

BACKGROUND

Let $\mathbf{Cx} = \mathbf{r}$ represent the MME to be solved, with $\mathbf{C}$ (of size $N \times N$) the coefficient matrix, $\mathbf{x}$ the vector of effects and $\mathbf{r}$ the vector of right hand sides. A widely used iterative method to solve for $\mathbf{x}$ is the conjugate gradient (CG) algorithm. Its convergence rate is heavily influenced by the condition number, of $\mathbf{C}$, $\kappa(\mathbf{C})$, i.e. the ratio of its largest to its smallest eigenvalue. Convergence rates can be improved if $\kappa(\mathbf{C})$ can be reduced. An extensively used method to achieve this is to 'pre-condition' the MME, i.e. to solve $\mathbf{M}^{-1}\mathbf{Cx} = \mathbf{M}^{-1}\mathbf{r}$ instead. Choice of the preconditioning matrix $\mathbf{M}$ usually represents a compromise between $\mathbf{M}$ being close to $\mathbf{C}$ (so that $\mathbf{M}^{-1}\mathbf{C}$ is close to an identity matrix) and requirements for storing or inverting $\mathbf{M}$. Simple, effective choices are (block-) diagonal matrices where $\mathbf{M}$ contains the diagonals (or small diagonal blocks) of $\mathbf{C}$.

Deflation has been advocated as a method to eliminate 'unfavourable' eigenvalues of a matrix by projection on a suitable subspace. Let $\mathbf{P}$ denote a matrix comprised of S linearly independent columns (of size N) which form a subspace of $\mathbf{C}$ so that $\mathbf{CP} = \mathbf{PT}$ and $\mathbf{T}$ is a non-singular matrix of order S . For $\mathbf{VP} = \mathbf{I}$ (where $\mathbf{I}$ is an identity matrix), Householder (1961) showed that the deflated matrix $\mathbf{B} = \mathbf{C} - \mathbf{PTV}$ has S zero eigenvalues and the remaining eigenvalues of $\mathbf{B}$ are those of $\mathbf{C}$ that are not

* A joint venture of NSW Department of Primary Industries and the University of New England

eigenvalues of **T**. Hence, assuming **C** is non-singular, **B** has rank $N - S$. Similarly, the eigenvectors of **B** are those of **C** that correspond to their common eigenvalues. In other words, "deflation of an eigenspace cancels the eigenvalues without affecting the rest of the spectrum" (Frank and Vuik 2001).

Use of a deflation preconditioner for CG and PCG algorithms has been considered by various authors in a range of fields (e.g Tang *et al.* 2009; Jönsthövel *et al.* 2012). Combining deflation with the 'standard' preconditioner yields the DPCG, a two-level preconditioning scheme particularly suited to ill-conditioned systems of equations. It involves solving $\mathbf{M}^{-1}\mathbf{PCx} = \mathbf{M}^{-1}\mathbf{Pr}$ with $\mathbf{P} = \mathbf{I} - \mathbf{CS}(\mathbf{S}'\mathbf{CS})^{-1}\mathbf{S}'$ aimed at reducing $\kappa(\mathbf{C})$ and **S** a matrix of size $N \times S$ which defines the deflation subspace (Frank and Vuik 2001). This requires the choice of **S**. Loosely speaking, the closer the deflation vectors (i.e. columns of **S**) approximate the 'unfavourable' eigenvectors of **C** the more effective deflation is likely to be. However, as for **M** it involves trade-offs between improvements in convergence and extra computational requirements. A simple strategy is to divide the space of **C** in correspondence to non-overlapping subsets of equations, referred to as subdomains (Frank and Vuik 2001). Let the $i-$th element of **x** belong to the $j-$th domain ($j = 1$ to S). This gives a matrix **S** with $ij-$th element equal to unity while the remaining elements are equal to zero, i.e. each row of **S** has only one non-zero element. At the extreme, fitting subdomains for individual, single effects is analogous to 'absorbing' the pertaining equations in the mixed model.

MATERIAL AND METHODS

Data consisted of 1,206,908 measurements for eye muscle depth recorded for Australian terminal sire sheep breeds between 1990 and 2018. Data were pre-corrected for fixed effects other than contemporary groups. There were 1,698,838 animals in the pedigree and genotype information, comprised of marker counts for 48,599 SNPs, was available for 23,040 animals. Invoking the HM, additive genetic effects were fitted for non-genotyped animals and marker effects modelled those of genotyped individuals. For simplicity, additional polygenic effects were assumed to be absent. In addition, the model included 54,094 contemporary groups (fixed), 93 genetic groups (random) and 56,212 sire × flock-year (random) effects.

MME were built and solved using either PCG or DPCG with independent subdomains as described above, using a diagonal preconditioner, $\mathbf{M} = \text{Diag}\{\mathbf{C}\}$, throughout. Solutions were assumed to have converged when $\alpha\sqrt{(\mathbf{x}_k - \mathbf{x}_{k-1})'(\mathbf{x}_k - \mathbf{x}_{k-1})/\mathbf{x}_k'\mathbf{x}_k} < 10^{-7}$, with $\mathbf{x}_k$ denoting the vector of solutions from the $k-$th iterate and α the step size parameter in the (D)PCG algorithm. Analyses were carried out considering all markers and reduced marker panels. To select the latter a simple GWAS was performed fitting markers as fixed covariables, one at a time. Subsets, of size m, were then selected to include those with p-values less than 0.5, 0.2, 0.1 and 0.05. Following Vandenplas *et al.* (2018), single

Table 1. Numbers of iterates required to solve the mixed model equations for different deflation subdomain ('chunk') sizes and marker subsets

p[a]	m[b]	No. of iterates							Correlation[c]	
		–	200[d]	100	50	20	10	5	NOG[e]	GEN
–	48599	3961	2722	2222	1741	1188	859	612	–	–
0.50	28875	3348	2525	2091	1682	1167	840	599	1.000	0.995
0.20	13318	2565	2118	1871	1559	1126	832	598	0.999	0.977
0.10	7858	2159	1833	1654	1416	1085	824	606	0.998	0.962
0.05	4680	1756	1560	1461	1293	1025	820	619	0.997	0.943

[a] Minimum p value for marker subset selection [b] Number of markers [c] Correlation of total breeding values from analyses using all and a subset of markers [d] Number of markers per 'chunk' [e] NOG non-genotyped, GEN genotyped

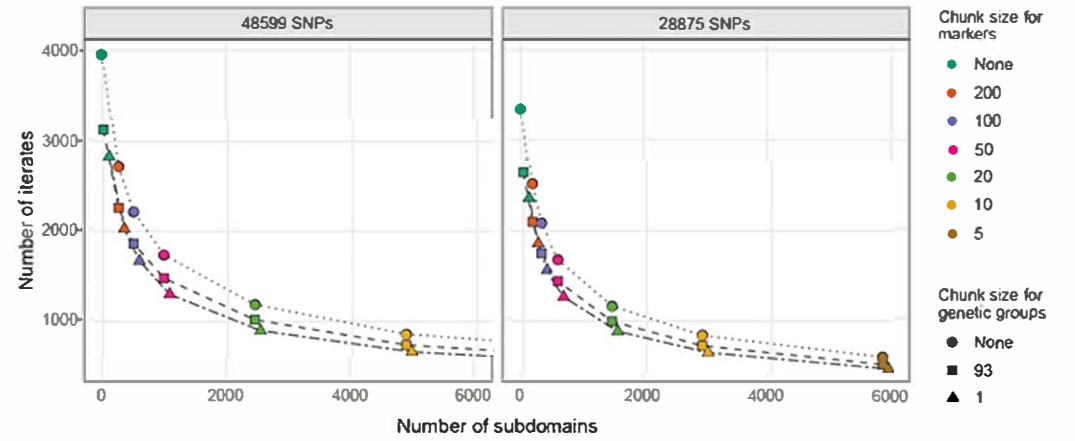

Figure 1. Numbers of iterates required for different deflation schemes and chunk sizes

domains were allocated to fixed effects and to all random effects other than marker effects. Equations for marker effects were divided into subdomains by selecting subsequent chunks (of equations) of size 5, 10, 20, 50, 100 or 200 to investigate the effect of chunk size on efficacy of deflation. This is referred to as scheme A. Scheme B was similar, but fitted separate subdomains for genetic group effects, with chunk sizes of 1 or 93. Computations were carried out under Linux on a shared machine with 512GB of RAM and 28 Intel Xeon CPU E5-2697 cores, rated at 2.6Gh using up to 28 threads.

RESULTS AND DISCUSSION

Numbers of iterates required to solve the MME for deflation scheme A are summarised in Table 1. For comparison, a corresponding analysis fitting the BVM and standard PCG (not shown) converged in 691 iterates. As reported by Vandenplas *et al.* (2018), deflation dramatically improved convergence, but small chunk sizes – and thus many subdomains – were required to achieve rates similar to those fitting the BVM. Reducing the number of markers decreased the number of iterates, especially for the larger chunk sizes (or no deflation), as well as reducing computations per iterate that were proportional to the number of markers. While correlations between predicted breeding values from analyses using the full and reduced marker sets for genotyped animals were less than 0.99 when markers with p-values less than 0.5 were eliminated, marker selection often affects the accuracy of evaluation considerably less, i.e. there is likely more scope for marker reduction than these correlations suggest. E.g., Saatchi and Garrick (2016) proposed a reduced panel for beef cattle comprising about 2,300 markers to capture most of the predictive performance of the full 50K panel.

Figure 1 illustrates the relationship between numbers of iterates required and deflation subdomains. Patterns for the other marker subsets were similar. Clearly, as emphasized by Frank and Vuik (2001), the efficacy of deflation increases with the number of subdomains employed. However, as S increases additional reductions in numbers of iterates achieved decrease. Our model of analysis fitted genetic groups as an additional random effect. This is known to affect convergence rates unfavourably – investigations for the BVM found that it almost doubled the number of iterates needed (Meyer *et al.* 2015). Additional analyses (not shown) identified a similar pattern for our data for the HM with standard PCG. Hence, scheme B attempted to counteract the detrimental effects of fitting genetic groups by defining additional subdomains. As shown in Figure 1 this yielded further reductions in the number of iterates required, the more so the larger chunk size for deflation of equations for marker effects. Even adding a single subdomain for all genetic groups (chunk size of 93) proved highly effective. Similarly, applying DPCG for the BVM, fitting a single subdomain for genetic groups (in

addition to two subdomains comprising all fixed and all other random effects, respectively) reduced the number of iterates required from 691 to 536.

While DPCG has the scope to dramatically improve convergence rates and its implementation is straightforward, deflation incurs additional computational cost per iterate and for set-up steps which need to be balanced against reductions in numbers of iterates and additional memory requirements. Figure 2 shows total, elapsed computing times for different analyses. Matrices $\mathbf{CS}$ and $(\mathbf{S'CS})^{-1}$ only need to be computed once but the computational burden increases with S and S^2, respectively, and storage for large numbers may become prohibitive. For our data, values of S greater than about 2,000 (using all markers) tended to increase total computing times, primarily due to these overheads. Overall, moderate deflation for markers, involving chunks of 20 to 100 SNPs, paired with assigning genetic groups to individual subdomains appeared to yield a reasonable compromise between improvements in convergence behaviour and additional computations for deflation. Our implementation relied on in-core storage of $\mathbf{CS}$ and $(\mathbf{S'CS})^{-1}$ and the data part of $\mathbf{C}$, but involved only limited optimisation of the computations associated with deflation. Values for 'iteration on data', out of core storage or improved parallel processing may differ; see Vandenplas *et al.* (2018) for some timings and discussion.

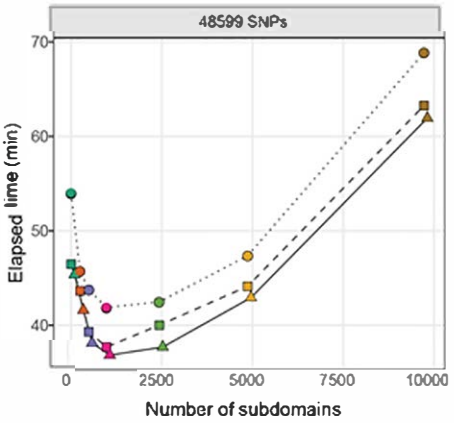

Figure 2. Total computing times[a]

[a] See Figure 1 for legend

As demonstrated for genetic groups, deflation assigning additional, separate subdomains to random effects other than markers was found to be advantageous. Further analyses (not shown) identified extra improvements in convergence rates when defining subdomains for groups of additive genetic effects for non-genotyped animals. Moreover, deflation also proved capable of improving convergence rates for the BVM. Further work will need to examine the efficacy of DPCG for multivariate analyses involving many traits and models fitting maternal effects, and to improve its implementation.

CONCLUSIONS

Deflation of the coefficient matrix in the mixed model equation reduces its condition number and thus improves convergence rates of an iterative solution scheme employing a conjugate gradient algorithm. It appears to be a valuable addition to our toolkit for genomic evaluation.

ACKNOWLEDGEMENTS

This work was supported by Meat and Livestock Australia grant L.GEN.1704.

REFERENCES

Fernando R.L., Cheng H., Golden B.L. and Garrick D.J. (2016) *Genet. Sel. Evol.* **48**:96.
Frank J. and Vuik C. (2001) *SIAM J. Sci. Comput.* **23**:442.
Householder A.S. (1961) *J. Soc. Indust. Appl. Math.* **9**:89.
Jönsthövel T.B., Van Gijzen M.B., MacLachlan S., Vuik C. and Scarpas A. (2012) *Comput. Mech.* **50**:321.
Meyer K., Swan A.A. and Tier B. (2015) *Proc. Ass. Advan. Anim. Breed. Genet.* **21**: Paper no. 18.
Saatchi M. and Garrick D.J. (2016) In *Plant and Animal Genome XXIV Conference*. San Diego, CA, January 9–13, 2016. Abstr.
Tang J.M., Nabben R., Vuik C. and Erlangga Y.A. (2009) *J. Sci. Comput.* **39**:340.
Vandenplas J., Eding H., Calus M.P.L. and Vuik C. (2018) *Genet. Sel. Evol.* **50**:51.