Decomposition and interpretation of genomic breeding values from a unified one-step national evaluation

T.J. Lawlor[#1] .I. Misztal[*] , S. Tsuruta[*], I. Aguilar[*†], and A. Legarra[§]

Introduction

Incorporating genomic information through the relationship matrix leads to improved genetic evaluations (VanRaden, P.M., 2008). For animals without progeny, the genomic EBV can be decomposed into its parent average (PA) and a part that measures the differences between its pedigree-based and genomic relationships. Aguilar et al. (2010) has presented formulas utilizing the information contained in the inverses of these matrices, G^{-1} and A^{-1}. The primary objective of this paper is to show how the decomposition of EBVs can provide valuable information on showing how changes in relationships influence our genetic evaluations.

Material and methods

This study included an analysis of 6,232,548 phenotypes, and BovineSNP50 genotypes on 17,187 bulls, of which 6,508 had progeny information. A full description of the data and analyses used for final score of US Holsteins are provided in Aguilar et al. (2010). A key component of their analysis was the use of a pedigree-genomic relationship matrix H, which is the joint distribution of the BVs for all animals, genotyped and non-genotyped. Choice of the scaling factor λ, to combine the genomic and pedigree-based relationships together, has been shown to be very influential in determining the accuracy and unbiasness of the evaluations (Tsuruta et al., 2010).

The decomposition of EBVs (u_i), for an animal without progeny, is given below. The adjustment in an animal's EBV due to differences between its pedigree-based and genomic relationships (AR) is defined as follows. Percent variation in AR explained by most influential animals = Variance in AR $_{\text{subset of animals}}$ / Variance in AR $_{\text{all animals}}$.

$$u_i = \frac{u_{sire} + u_{dam} + \lambda \sum_{j,\, j \neq i} (a_{22}^{ij} - g^{ij}) u_j}{2 + \lambda (g^{ii} - a_{22}^{ii})}$$

$$\mathbf{AR} = \frac{\sum_{j,\, j \neq i} (a_{22}^{ij} - g^{ij}) u_j}{2 + (g^{ii} - a_{22}^{ii})}$$

$$\textbf{\% PA Contribution} = \frac{2}{2 + (g^{ii} - a_{22}^{ii})}$$

$$\textbf{\% AR} = \frac{\sum_{j,\, j \neq i} (a_{22}^{ij} - g^{ij})}{2 + (g^{ii} - a_{22}^{ii})}$$

---

[1] [#] Holstein Association USA Inc., [*] University of Georgia, Athens, GA, USA
[†] Instituto Nacional de Investigación Agropecuaria, Uruguay, [§] INRA, France

The diagonal elements of G^{-1} and A_{22}^{-1} reflect the amount of information variation explained by its parents, inbreeding and its relationship to other animals. The off-diagonal elements convey information between its parents and their mates. Positive covariances come from its direct ancestors, progeny and ancillary relatives. Negative covariances come from the mates of these animals. Although A_{22}^{-1} tends to be sparse, G^{-1} is much more dense.

Results and discussion

The impact of genomic information on EBVs is quite important, offering dairy cattle breeders an opportunity to dramatically improve genetic progress. However, to be fully embraced by practitioners, explanations of how genomic enhanced EBVs are calculated will be required.

A higher genomic relationship with another animal, as compared to the pedigree-based value, indicates that more DNA, and presumably more genes, is shared between those two individuals than normal gene segregation would have predicted. Intuitively, if two individuals are found to be genetically more similar, then more emphasis should be placed on that relative's EBV. If more DNA comes from one side of a pedigree then less DNA must have come from the other side. Necessitating a reduction in the emphasis that the EBV of a mate of a relative would receive.

For illustrative purposes, the parts of the genomic EBVs of several young genomic tested bulls with no progeny will be provided. They include two high EBV bulls, an average EBV bull with three complete generations of genomic tested ancestors, and two bulls with either a large increase or decrease due to the addition of genomic information.

Essentially all of our 17,187 genomic tested bulls contribute in some small way to a young bull's EBV. However, often there are a limited number of individuals whose genomic covariance differs significantly from the pedigree-based values. Providing us with an understanding of how the genomic relationships are changing an animal's EBV. A good example of this is the bull Java. He's a bull who, due to the addition of genomic information, has an EBV that is much higher than his parent average. Most of the changes in his genetic covariances occur amongst his close relatives or their mates.

The two individuals with the highest AR, for this bull, are his sire Reggie and maternal grandsire Aaron. Their genomic relationships with Java are 0.87 and 0.70, respectively. The third highest AR bull, Airraid, is the mate of Java's half-sister. They have a son, Fakir, who has a genomic relationship with Java of 0.62. Thus the maternal side of the pedigree has become more important. Unlike the sire and mgs, Airraid's EBV will be subtracted from Java's EBV. Since Airraid's DNA is not being transmitted to Java, a larger negative covariance will help ensure that less of Airraid's genetic merit, through Fakir, will be credited to Java.

Braxton, the top genomic tested young bull, better illustrates how certain individuals contribute more or less. He has a large maternal family and many of the high AR individuals are also close relatives and their mates. His dam, Barbie, has with several genomic tested sons. For seven of the sibs, the genomic relationship with Braxton ranges from 0.55 to 0.93.

These genomic relationships are much higher than the pedigree-based values. A positive increase in the genetic covariances of these maternal siblings with Braxton means that a larger genetic contribution will be coming from the dam’s offspring.

Below is a table of the changes in inverse elements ($a^{ij} - g^{ij}$) for Braxton. The maternal siblings with a high genomic relationship with Braxton will receive more weight and contribute directly to Braxton’s genomic EBV. The mates of Barbie all contribute less, with the exception of Shottle the sire of Braxton. Barbie also has a daughter, named Cinderella, who also has several genomic tested sons. The EBVs of these bulls will provide more of a contribution to the EBV of Braxton than a traditional evaluation using a pedigree-based relationship matrix.

Maternal brothers out of **Barbie**			Sires of the bulls **Mates** of **Barbie**			**Sons** of Cinderella	
	$a^{ij} - g^{ij}$			$a^{ij} - g^{ij}$			$a^{ij} - g^{ij}$
Baltimor	+2.56		Shottle **	+0.84		Cade	+0.54
Bedford	+2.28		Mac	-0.27		Clinto	+0.27
Birk	+1.53		Toystory	-0.97		Capture	+0.09
Blackjack	+1.23		Dundee	-0.81		Casino	+0.47
Bossman	+1.10		Mac	-0.27		Chancellor	+0.72
Chilton	+0.89		Champion	-1.46		Casey	+0.63
Bonzai	+0.53		Dundee	-0.81		Carriage	+0.36

** Sire of Braxton

Often a large proportion of variation in AR can be explained by a small fraction of the 18,187 bulls with genomic information. For example, ten bulls explain 38.4% of the change in AR for Atwood. An average bull, O-Style, with lots of genomic tested relatives, has 21.4% of the variation explained by 10 bulls. Of these five bulls, the least explainable change in EBV, due to genomic information, is the one with a large decrease. Perhaps, grouping the AR values of family members together would help in explaining this change.

	g^{ii}	a^{ii}	Var	Percent Variation in AR explained by			
				top 10	top 20	top 50	top 100
High GPTAT-Braxton	21.08	1.97	2900	34.2	39.4	46.3	52.7
High GPTAT-Atwood	18.08	2.03	2718	38.4	42.5	48.8	54.6
Average bull - O-Style	16.45	1.81	1379	21.4	23.6	28.2	33.3
Large increase - Java	14.58	1.87	203	30.2	33.4	37.7	41.4
Large decrease -Mako	8.25	1.47	692	12.4	16.1	22.7	28.6

The PA and AR for these five bulls are listed below. In all cases, the adjustment due to differences between its pedigree-based and genomic relationships is 77.0 to 90.5% of the animal’s EBV.

	PA	% PA	% AR		PA contribution	AR		GEBV
High GPTAT - Braxton	6.93	9.5	90.5		0.62	7.44		8.06
High GPTAT - Atwood	5.90	10.6	89.4		0.65	7.04		7.69
Average bull - O-Style	1.33	12.0	88.0		0.16	1.40		1.56
Large increase - Java	2.37	13.6	86.4		0.33	3.87		4.21
Large decrease - Mako	3.16	23.0	77.0		0.73	0.50		1.23

One way to reduce the contribution of the genomic relationships to the EBVs is by reducing the scaling factor. The proper choice of a scaling factor, may differ by trait and population, and has been shown to influence the accuracy and unbiasness of the evaluations. A change in λ from 1.0 to 0.7 on our five example bulls is given below. A decrease in λ causes the PA of the animal to change, as well as, placing more emphasis on the PA.

	PA λ=1.0	%PA λ=1.0	% AR λ=1.0	**GEBV λ=1.0**		PA λ=0.7	%PA λ=0.7	%AR λ=0.7	**GEBV λ=0.7**
Braxton	6.93	9.5	90.5	**8.06**		6.59	13.0	87.0	**7.09**
Atwood	5.90	10.6	89.4	**7.69**		5.70	14.5	85.5	**7.36**
O-Style	1.33	12.0	88.0	**1.56**		1.57	16.3	83.7	**2.00**
Java	2.37	13.6	86.4	**4.21**		2.54	18.3	81.7	**3.95**
Mako	3.16	23.0	77.0	**1.23**		3.56	29.9	70.1	**2.09**

Conclusion

For animals without progeny, the genomic EBV can be decomposed into its parent average (PA) and a part that measures the differences between its pedigree-based and genomic relationships. The EBV of a direct relative with a higher genomic relationship, as compared to the pedigree-based value, will contribute more to a young animal's EBV than a less related relative. This change will also cause a larger negative covariance between the mate of the direct relative and the animal of interest. Often a large proportion of the variation, caused by these differences in genomic versus pedigree-based relationships, can be explained by a small fraction of the animals with genomic information. This leads to some straightforward explanations on why genetic evaluations change when genomic information is added. The opportunity to control the impact of genomic relationships on the EBVs can be accomplished by scaling the combination of G^{-1} and A^{-1}. This adjustment will change the evaluations of all animals with progeny, as well as, the emphasis that the genomic relationships will have on the EBVs of the animals without progeny.

References

Aguilar, I., Misztal, I., Johnson, D.L. *et al.* (2010). *J. Dairy Sci.*, 93:743-752.
Tsuruta, S., Aguilar, I., Misztal, *et al.* (2010). In *Proc 9th WCGALP.*
VanRaden, P. M. (2008). *J. Dairy Sci.*, 91:4414-4423.