Optimizing multiple trait selection

J.H.J. van der Werf[1], R.G. Banks[2] and J.C.M Dekkers [3]

[1] *School of Environm. and Rural Science, Univ. of New England, Armidale, NSW, Australia*
jvanderw@une.edu.au (Corresponding Author)
[2] *Animal Breeding and Genetic Unit, University of New England, Armidale, NSW, Australia*
[3] *Department of Animal Science, Iowa State University, Ames, Iowa, USA*

Summary

Multiple trait selection is commonly based on Hazel's selection index theory. In spite of this solid theory, multi-trait selection is not always fully understood or optimized, and in practice, selection decisions frequently deviate from the selection index optimum. We explore a two trait selection example with traits that have an unfavourable correlation, e.g. productivity and fitness, as this relates to the case where multiple trait selection is most challenging because response is most sensitive to changes in both economic value and the accuracy of the estimated breeding values. This is also the case where measurement strategies, including genomic selection, have the largest effect on the direction of trait improvement. Two drawbacks of the optimal strategy provided by selection index theory are also mostly pronounced in this case, the first being the assumption of linearity of the breeding objective (and the index), and the other that selection response is optimized conditional on the information available. Consequently, outcomes are always leaning towards more selection response for traits that are easy to improve, i.e. easy to measure and higher heritability, and to the detriment of traits that are unfavourably correlated but have limited information available on them to base selection on. We suggest a simple criterion that can indicate where multiple trait selection is most sensitive to changes in economic values or additional trait measurement and provide a more rational approach to indexes that provide undesirable responses.

Keywords: Multiple Trait Selection, Economic Values

Introduction

Sustainable genetic improvement requires balanced selection on multiple traits. The common framework to optimize multiple trait selection is based on selection index theory (Hazel, 1943), where optimal weights are derived to maximize selection response, assuming genetic parameters are known and conditional on a certain set of information available. Index selection is a well-accepted framework in animal breeding and is generally quite robust, e.g. with respect to assumptions about linearity of the relationship of genetic values for traits with profit, and estimation errors in genetic parameters. However, index selection can give less satisfactory results in case of non-favourable correlations among traits that are under selection, i.e. when the sign of the genetic correlation between two traits is opposite to the sign of the product of their economic values. The optimized selection response then becomes more sensitive to economic values and is much more likely leading to suboptimal results if genetic parameters or economic values are not known without error. Also, such cases often lead to undesirable predicted responses, and economic values are then 'adjusted' to obtain more desirable responses. A common example is selection for milk yield and fertility in dairy cows. More generally, the problem occurs when selecting on productivity and fitness, or on production (e.g. growth rate) and product quality (e.g. eating quality of meat).

Breeders have been aware of undesirable correlated responses and have invented 'solutions' such as desired gain indexes. For example, the weighting on fertility is artificially increased to avoid a decrease of fertility that would have occurred with a rational economic approach where derived economic values would have been used as weights, because from an economic perspective the reduction in fertility is compensated by an increase in milk yield that has more economic value (e.g. Pryce et al, 2009). One could argue that the original economic values were incorrectly derived by not accounting for a possible non-financial aspect of fertility decrease. Another view could be that in the longer term, the breeding objective is not linear, and extremely negative values for one trait are not compensated by extreme positive values for the other trait.

Another problem in the case of fertility is also that the trait is more poorly measured. Increasing information about estimated breeding values (EBVs) for these traits will also help to drive them in a more desirable direction. Dekkers and Van der Werf (2014) discussed consequences of increasing the accuracy of trait EBV, e.g. via genomic selection, and its effect on selection responses to traits in the breeding objective. A better understanding of how EBV accuracy affects trait response is useful when making investment decisions in breeding programs. Dekkers and Van der Werf (2014) also showed graphically that the optimal response can become more sensitive to errors in economic values when genomic selection is applied. In this paper, we will further generalize this point and present a simple relationship between accuracy of EBV and economic value of traits and its effect on optimal response

The purpose of this paper is to provide more insight in multiple trait selection, particularly when traits have unfavourable correlations. We discuss and present expressions that clarify when responses are most sensitive to changes in economic values and changes in trait measurement. We also discuss strategies to overcome situations where trait responses are undesirable and question the optimality of Hazel's selection index approach to long term multiple trait genetic improvement. We use a two trait example such that a visual representation of response ellipses (Moav and Hill, 1966) can be used to illustrate our points.

Multiple Trait Selection with unfavourable correlations

The ellipse representation

A classical breeding problem is when two traits have an unfavourable genetic correlation This is the case when the correlation is negative, and both traits have an economic value of equal sign, or when the correlation between the traits is positive, and the economic values have an opposite sign. The best way to illustrate the problem is to draw an ellipse of possible responses (Moav and Hill, 1966), which provides all possible combinations of response in the two traits that can be achieved for a given selection intensity across the full range of relative economic values for the two traits, assuming a certain accuracy for each of the EBVs for the two traits. The ellipse represents the bivariate distribution of the EBVs for the traits. The optimal response according to selection index theory is at the tangent of the ellipse and iso-economic lines (Figure 1), which are straight lines of equal profit if the breeding objective is defined as a linear function, as is typically the case in the selection index approach.

In the example, both traits are standardized to have a phenotypic variance of 1 and a heritability of 0.3. The genetic correlation is -0.5 and the phenotypic correlation is -0.1. Consider 30 and 2 progeny measured for traits 1 and 2, such that multivariate EBV accuracies are 0.84 and 0.52, respectively. The shape of the ellipse is determined by the correlation between EBVs for the two traits. Note that this only represents the genetic correlation between the traits if the EBVs are highly accurate. In the example case, however, the correlation between the EBVs is -0.8, as information from the correlated trait is used to obtain EBV. The higher the EBV correlation, the more difficult it is to improve both traits in the desired direction, i.e. into the top-right quadrant in Figure 1.

Figure 1. Sensitivity of optimal response when the economic value for trait 1 changes from 0.5 (left) to 0.75 (right), while the economic value for trait 2 is equal to 1.

Sensitivity to economic values

In the example in Figure 1 the optimal response is an improvement for one trait, while the other trait has zero improvement. A slight change in the economic vales causes a slight change in the slope of the iso-economic line. In the example, the economic value for trait 2 is 1, whereas it is 0.50 and 0.75 for trait 1 in the left and right frame, respectively. This change causes a shift in response from trait 2 to trait 1 (from the left frame to the right frame in Figure 1). Therefore, this relatively small change in economic value causes a large shift in trait responses. From a biological point of view, this represents strong sensitivity to the economic value. From an economic point of view, the sensitivity is much smaller. Assuming a 'true' economic value of 0.625 for trait 1, an upward or downward change of 20% in the estimate of the economic value for trait 1 used for the selection index, results in a change in profit of 6.5% and 4.4%, respectively (Figure 2, left frame). A similar change in the economic value of trait 1 would cause a very small change in profit response when the economic value for trait 2 is much smaller, as illustrated in the left frame of Figure 2.

The case in Figure 3 (left) represents the typical example of selection for production (trait 1) versus 'fitness' (trait 2). Due to the limited information about trait 2, the selection index seeks an optimal solution by improving the trait with more information and with more economic value (Figure 3, left frame). Changing the weight on fitness has limited impact on the optimal response and one would have to at least quadruple the economic weight to avoid a decline in fitness.

Figure 2. Sensitivity of response in profit to errors in the economic value for trait 1. True economic values are 0.625 for trait 1 and for trait 2 they are 1.0 in the left and 0.2 in the right frame.

Figure 3. Optimal responses when the economic values are 0.625 and 0.4 for traits 1 and 2, respectively, when the accuracy of the EBV for trait 1 is 0.84 and for trait 2 is 0.52 (left) or 0.84 (right).

The sensitivity of trait response to economic values is highest when the iso-economic line runs parallel to the major axis of the ellipse. This is the case when the economic value per standard deviation of the EBV is equal for the two traits. In other words, the largest sensitivity occurs when $r_1.\sigma_{a1}.v_1 = r_2.\sigma_{a2}.v_2$ if there is a negative correlation between the traits and it occurs when $r_1.\sigma_{a1}.v_1 = - r_2.\sigma_{a2}.v_2$ in case of a positive correlation, where r_i is the accuracy of the (multi-trait) EBV of trait i, σ_{ai} is the genetic standard deviation of trait i and v_i is the economic value. This shows that both economic values and the accuracies of the EBVs impact optimal responses and index sensitivity. This expression is also useful to identify cases where this sensitivity occurs. For example, when considering economic values per unit of genetic standard deviation (i.e. $v\sigma_{ai} = \sigma_{ai}.v_i$), then, in the case of a negative genetic correlation, the most sensitive case is when $r_1 = r_2.(v\sigma_{a2}/v\sigma_{a1})$. If there is a large difference in economic value per genetic standard deviation between the traits, then trait response to selection will not be sensitive to changes in economic value, provided the traits have similar accuracy. Equally, one can conclude that the most sensitive case occurs when $v\sigma_{a1} = v\sigma_{a2}$ (r_2/r_1), i.e. when there is a large difference in EBV accuracy between the traits, trait responses will not be sensitive to changes in economic value, unless the least accurate trait has a much higher economic value per genetic standard deviation.

Sensitivity to increasing trait information.

Note that the point where $r_1.\sigma_{a1}.v_1 = r_2.\sigma_{a2}.v_2$ is not only the point where the index solution is most sensitive, it is also where both traits are improved in the desirable direction. Therefore, this expression can also be used to predict the outcome of increasing the accuracy of the EBV of traits. Figure 3 illustrates that changing the accuracy of the EBV of trait 2 results in the optimal response such that trait 2 is not declining (Figure 3, right). Again, one can predict whether this will happen from the expression above. The most sensitive point is when $r_2 = r_1.(v\sigma_{a1}/v\sigma_{a2})$. Hence, if the economic value per genetic standard deviation ($v\sigma_{ai}$) differs substantially between the traits, increasing the accuracy of traits has limited effect on trait responses to selection. However, if these economic values are similar, the optimal response more likely improves both traits if the accuracy of the EBVs are more similar.

Conclusions

Multiple trait selection responses are sensitive to economic values and accuracies of EBVs but only when traits have unfavourable correlations and the economic value per standard deviation of the EBV is similar between the traits. It is also in that case that trait responses to index selection are more likely in a desired direction for both traits. In other cases, the effect of changing index weights or increasing the accuracy of the EBV of traits will have limited effect on the direction of genetic change and selection index methodology may not lead to desirable improvement for all traits on the long term.

List of References

Dekkers, J.C.M, and J.H.J. van der Werf. 2014. Breeding Goals and Phenotyping Programs for Multi-Trait Improvement in the Genomics Era. Proceedings, 10th World Congress of Genetics Applied to Livestock Production

Hazel , L.N. 1943. The genetic basis for constructing selection indexes. Genetics 28: 476-490

Moav, R. and Hill, W.G. 1966. Specialised sire and dam lines. IV. Selection within lines. Anim. Prod. 8: 375-390.

Pryce, J.E., J.H.J. van der Werf, M. Haile-Mariam, B. Malcolm and M.E. Goddard. 2009. Updated index weights for the Australian Profit Ranking in dairy cattle. Proc. Assoc. Advmt. Anim. Breed. Genet. 18: 143-146.