Selection indices for desired relative genetic gains with inequality constraints

Y. Itoh and Y. Yamada

Department of Animal Science, College of Agriculture, Kyoto University, Kyoto 606, Japan

Summary. A selection index which maximizes genetic gains in a desired direction has been previously suggested. We extend this index to the case where desired relative genetic gains are constrained to be not less or not greater than pre-specified levels. Further, we suggest an index for desired relative genetic gains constrained to be between certain levels. All these indices are obtained using quadratic programming techniques.

Key words: Selection index — Desired relative genetic gains — Inequality constraints — Non-linear programming

Introduction

In multiple trait selection for genetic improvement of animal and plant populations, Smith (1936) and Hazel (1943) presented a selection index which maximized the improvement of an overall genetic-economic value. Later, Kempthorne and Nordskog (1959) introduced the idea of a restricted selection index which restricts the genetic changes of certain traits. Subsequently, many other types of constrained selection indices were proposed by Tallis (1962, 1985), Harville (1974, 1975), Rao (1962, 1965), and others. All of these indices are based on the relative economic values of component traits.

However, the assignment of the relative economic values is not always easy; reasons for difficulties were described by Yamada et al. (1975) as: (1) in some traits, no standard for assigning relative economic values is readily available, e.g., body conformation; (2) in some traits, the assumption that the relative economic values remain constant over the range of varia-

tion is not satisfied; and (3) when some restrictions were assigned, the relative economic values were no longer the same as those of the unconstrained selection index.

Lin (1978) also summarized the difficulties of assigning relative economic values: (1) when the relative economic values are derived from direct economic analysis of a production system, some traits are hard to define objectively; (2) when the relative economic values are derived from multiple regression analysis, they vary depending on how profit is defined, the number of traits considered in multiple regression equation and sampling variability; and (3) the relative economic values may change from time to time or vary from one location to another.

For these reasons, index selection has not been used extensively in practice until recently, despite its theoretical superiority in efficiency.

On the other hand, Pešek and Baker (1969) suggested a selection index to attain predetermined desired genetic gains. Their index is fundamentally different from that of Smith (1936) and Hazel (1943): it does not require the definition of aggregate economic values in the derivations. Yamada et al. (1975) proposed a selection index that attains predetermined breeding goals at a minimum number of generations of selection. The breeding goals are defined as ultimate levels of the traits of interest, expressed as deviations from present population means. (In general, the breeding goals do not mean desired genetic gains in one generation, but mean desired genetic gains in an uncertain number of generations.) Brascamp (1979) and Itoh and Yamada (1986) discussed the theoretical basis of this index in more detail. Similar indices were derived by Harville (1975), Rouvier (1977), Essl (1981) and Tallis (1985). The advantage of this index is that it does not require relative economic value for each component trait.

One may argue that the difficulty with this index in practice is to choose a set of proper levels of component traits. It is, however, not difficult for an experienced breeder, because he must know the relative merits and demerits of his strain as a result of the Random Sample Test or a critical comparison with competitors' stocks (Yamada et al. 1975). For these reasons, we conclude that the selection indices based on the relative economic values are ideal from a theoretical point of view, but the selection indices for attaining the breeding goals often more useful from a practical point of view.

There are some cases, however, where it is difficult or unnecessary to precisely determine breeding goals of certain traits. In such cases, it may be more convenient and efficient to improve the population in a direction within a permissible range, than to improve it in a direction of strictly fixed goals. Furthermore, there are cases where one does not know, or does not want to specify, the breeding goals for certain traits, but simply wants to avoid their deterioration. An inequality constraint forcing the genetic changes for such traits to be merely non-negative seems then much more adequate than the arbitrary choice of a direction of genetic improvement.

Within the conventional selection index framework, Rao (1962) suggested an index constraining genetic gains for some traits to be non-negative. An easier procedure was proposed by Mallard (1972), Rao (1965) and Harville (1974) further discussed more general indices with constraints. All of these indices are computed using non-linear programming techniques.

In the present paper, similar techniques are applied to the index for selection in a direction within a acceptable range, for which an *a priori* knowledge of relative economic values is not required.

Methods

Selection index for improvement in a desired direction

A brief description of the index of Yamada et al. (1975) for attaining pre-specified breeding goals is first necessary, because this index provides the basic ideas for its extension in this paper.

To describe it, we use the following notations:

b an $n \times 1$ vector of index weights
p an $n \times 1$ vector of phenotypic values expressed as deviations from their respective means
g_1 an $m \times 1$ vector of additive genetic values
P an $n \times n$ phenotypic variance covariance matrix, i.e., $P = \text{Var}(p)$
G_1 an $n \times m$ covariance matrix between p and g_1, i.e., $G_1 = \text{Cov}(p, g_1)$
k_1 an $m \times 1$ vector of desired relative genetic gains
I a selection index, i.e., $I = b' p$
i the intensity of selection
σ_I the standard deviation of the index, i.e., $\sigma_I = \sqrt{b' P b}$.

If q is a vector of breeding goals expressed as the deviations from present means, then the vector k_1 can be defined as $k_1 = \alpha q$, where α is an arbitrary positive scalar. A simple example of k_1 is $k_1' = [1 \mid 2]$, which means that, in one generation of selection, the expected genetic gain of the second trait is desired to be twice as large as the expected genetic gain of the first one, i.e., $\text{E}(\varDelta g_2) = 2\,\text{E}(\varDelta g_1)$, where $\text{E}(\varDelta g_1)$ and $\text{E}(\varDelta g_2)$ represent the expected genetic gains of trait 1 and trait 2, respectively (see Fig. 1). Vector k_1 can have elements equal to zero, which means that no genetic change of the corresponding trait is desired. Note that no relative economic value nor aggregate genotypic value is assumed here.

After one generation of selection, we have the expected genetic gains as:

$$\text{E}(\varDelta g_1) = i\,\text{Cov}(g_1, I)/\sigma_I = i\,G_1'\,b/\sigma_I. \tag{1}$$

To make $\text{E}(\varDelta g_1)$ proportional to pre-specified k_1, b must satisfy the following condition:

$$i\,G_1'\,b/\sigma_I = \theta\,k_1 \tag{2}$$

where θ is an arbitrary positive scalar. Let us assume $\theta = i/\sigma_I$, then b is a solution of

$$G_1'\,b = k_1. \tag{3}$$

Substituting (3) into (1), we get

$$\text{E}(\varDelta g_1) = i k_1/\sigma_I. \tag{4}$$

If $n = m$, equation (3) has a unique solution as

$$b = (G_1')^{-1} k_1.$$

If $n > m$, no unique solution exists. But $\text{E}(\varDelta g_1)$ is inversely proportional to σ_I, therefore the best choice among all solutions is the vector b corresponding to the minimum value of σ_I, thus maximizing the absolute values of the elements of $\text{E}(\varDelta g_1)$ in (4). Then the optimal b should satisfy

$$\min_{\{b \mid G_1' b = k_1\}} b'\,P\,b.$$

We denote the problem of finding this b as:

$$\begin{cases} G_1'\,b = k_1 \\ b'\,P\,b \quad \text{minimum}. \end{cases} \tag{5}$$
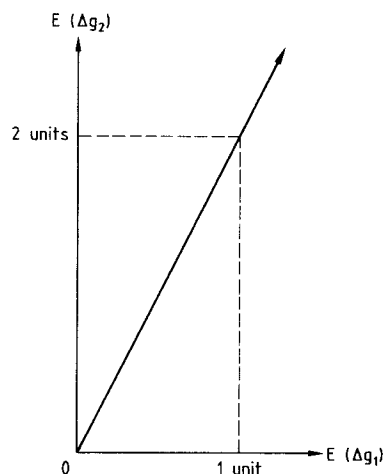
Fig. 1. Diagram illustrating the desired relative genetic gains such that $\text{E}(\varDelta g_2) = 2\,\text{E}(\varDelta g_1)$. A *sloping arrow* represents the desired direction of genetic improvement

The solution of (5) can be written explicitly as:

$$b = P^{-1} G_1 (G_1' P^{-1} G_1)^{-1} k_1 . \tag{6}$$

This process of deriving (6) was described by Itoh and Yamada (1985).

The index for desired relative genetic gains not less than or not greater than certain levels

Partition the traits into two groups, and let the desired relative genetic gains of the first group be equal to k_1 and those of the second group be equal to or larger than k_2. For example, for simplicity consider only two traits. If the genetic gain for trait 2 is desired to be at least twice the genetic gain for trait 1 (i.e., $E(\Delta g_2) \geq 2E(\Delta g_1)$) (see Fig. 2), then take $k_1 = 1$ as a reference and $k_2 = 2$. If the genetic change for trait 2 is desired to be simply non-negative ($E(\Delta g_2) \geq 0$), we set $k_1 = 1$, again as a reference, and $k_2 = 0$.

Let g_1 and g_2 be the vectors of additive genetic values of the first and second groups, respectively, and let $G_1 = \text{Cov}(p, g_1)$ and $G_2 = \text{Cov}(p, g_2)$. Then, in the same way as (3) was derived, it can be shown that b is a solution of:

$$G_1' b = k_1 \quad \text{and} \quad G_2' b \geq k_2 .$$

If no unique solution exists, the optimal b is such that

$$\begin{cases} G_1' b = k_1 \\ G_2' b \geq k_2 \\ b' P b \quad \text{minimum} . \end{cases} \tag{7}$$

This is problem of quadratic programming. Note that the problem dealing with only inequality constraints has no practical meaning, because the equality constraints provide the necessary standards for the relative gains. Therefore, k_1 should involve at least one non-zero element. Then we can interpret (7) as the problem of obtaining b which maximizes the absolute values of the elements of $E(\Delta g_1) = i k_1/\sigma_I$ subject to $G_1' b = k_1$ and $G_2' b \geq k_2$.

Conversely, we can consider the problem that the desired relative genetic gains of a group of traits are equal to k_1, and those of another group of traits are equal to or less than k_3. For example, if only two traits are considered and $k_1 = 1$ and $k_3 = 2$, then the genetic gain of trait 2 is desired to be equal to or less than twice that of trait 1 ($E(\Delta g_2) \leq 2E(\Delta g_1)$). A simpler example is the case where a certain element of k_3 equals zero, which means that the genetic change of the corresponding trait is desired to be zero or negative.

Let g_3 be a vector of additive genetic values of the traits of the latter group and put $G_3 = \text{Cov}(p, g_3)$; the optimal b of this problem is such that

$$\begin{cases} G_1' b = k_1 \\ G_3' b \leq k_3 \\ b' P b \quad \text{minimum} . \end{cases} \tag{8}$$

Further, it is also possible to combine these two problems (7) and (8); the combined problem is such that

$$\begin{cases} G_1' b = k_1 \\ G_2' b \geq k_2 \\ G_3' b \leq k_3 \\ b' P b \quad \text{minimum} . \end{cases} \tag{9}$$

The index for desired relative genetic gains between certain levels

Relative genetic changes can also be constrained to be between two sets of levels. This constraint is useful when one can determine only approximate desired relative genetic gains of some traits.

We assume that desired relative genetic gains of a group of traits are k_1, and those of another group of traits are between k_4 and k_5, where k_4 and k_5 represent the lower and upper limits of the permissible ranges of the desired relative genetic gains. For example, if only two traits are considered and $k_1 = 1$, $k_4 = 1.5$ and $k_5 = 2$, then we want $1.5\,E(\Delta g_1) \leq E(\Delta g_2) \leq 2E(\Delta g_1)$ (see Fig. 3).
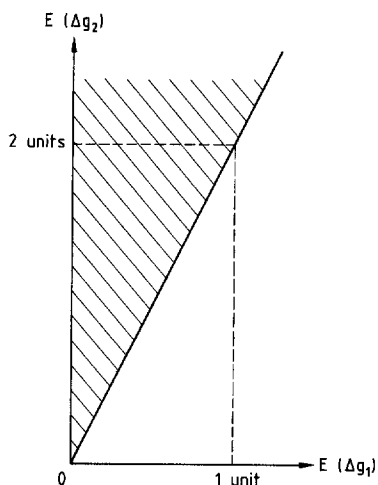

Fig. 2. Diagram illustrating the desired relative genetic gains such that $E(\Delta g_2) \geq 2E(\Delta g_1)$. A *shaded area* represents the desired range of genetic improvement
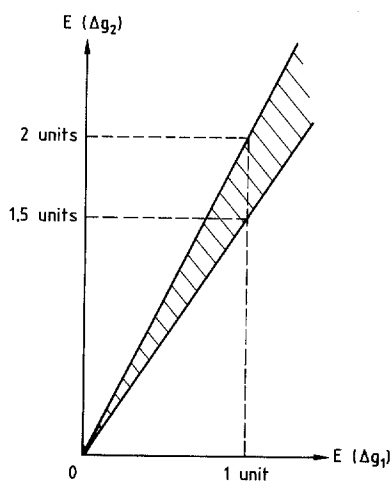

Fig. 3. Diagram illustrating the desired relative genetic gains such that $1.5\,E(\Delta g_1) \leq E(\Delta g_2) \leq 2E(\Delta g_1)$. A *shaded area* represents the desired range of genetic improvement

Let g_4 be a vector of additive genetic values of the traits of the latter group and put $G_4 = \mathrm{Cov}(p, g_4)$, then the optimal b of this problem is such that

$$\begin{cases} G'_1 b = k_1 \\ k_4 \leqq G'_4 b \leqq k_5 \\ b' P b \quad \text{minimum}. \end{cases} \tag{10}$$

This problem also can be solved by quadratic programming.

It is possible to make any combination with the other types of constraints mentioned, and the combined problem with the constraints of all four types is such that

$$\begin{cases} G'_1 b = k_1 \\ G'_2 b \geqq k_2 \\ G'_3 b \leqq k_3 \\ k_4 \leqq G'_4 b \leqq k_5 \\ b' P b \quad \text{minimum}. \end{cases} \tag{11}$$

Computing techniques

All the indices described in this paper can be computed using quadratic programming techniques. In our problem, the optimal solution is always found at one of the boundaries of a region constructed with the equality and inequality con-straints. These boundaries are determined by all possible combinations of equations. For example, the boundaries of the problem (11) are determined by all possible combinations of the following equations:

$$G'_1 b = k_1, \quad G'_2 b = k_2, \quad G'_3 b = k_3, \quad G'_4 b = k_4 \quad \text{and} \quad G'_4 b = k_5. \tag{12}$$

Therefore, we can obtain the optimal solution of this problem as follows:

1) Calculate indices as in (6) for each possible combination of the equality constraints of (12).
2) Eliminate those that do not satisfy the inequalities in (11).
3) Select the index with a minimum variance from the ones that are left.

This procedure resembles that of Mallard (1972). The major difference is that we attempt to minimize the variance of the index and not to maximize its efficiency.

Numerical examples

To illustrate the indices, we will use a flock of poultry as an example. Traits considered are egg weight (EW), egg production rate (EP) and body weight (BW). Their phenotypic and genotypic variances and covariances are given in Table 1.

Examples of indices for various desired relative genetic gains are given in Table 2. In Index 1, the genetic gain of EP is desired to be 0.5 times as large as that of EW, and that of BW is desired to be −0.1 times as large as that of EW. On the other hand, in Index 2, the genetic gain of EP is desired to be greater than or equal to 0.5 times that of EW. This index has a smaller variance than Index 1, and gives larger expected genetic gains than Index 1, not only in EP but also in EW and BW. In Index 3, the same relative genetic gains in EW and EP as those of Index 1 are desired, but the genetic change of BW is desired to be zero or negative. This index has a much smaller variance, and gives much larger expected genetic gains in EW and EP than Index 1. In Index 4, the relative genetic gains in EP and BW are desired to be between certain levels. This

Table 1. Phenotypic and genotypic variances and covariances. Phenotypic variances and covariances are above the diagonal and genotypic variances and covariances are below the diagonal

	EW (g^2)	EP $(\%^2)$	BW (kg^2)
EW	13.3949 / 6.6974	−1.8371	0.2417
EP	2.1345	100.7805 / 30.2341	−0.1326
BW	0.1376	0.	0.0698 / 0.0314

Table 2. Examples of indices for various relative desired genetic gains

No.	Relative desired genetic gains			Index weights			Variance of the index	Expected genetic gains[a]		
	EW	EP	BW	EW	EP	BW		EW	EP	BW
1	= 1	= 0.5	= −0.1	0.2360	0.0001	−4.2191	1.5073	0.8145	0.4073	−0.0815
2	= 1	≧ 0.5	= −0.1	0.2326	0.0096	−4.2042	1.4976	0.8172	0.6424	−0.0817
3	= 1	= 0.5	≦ 0	0.1623	0.0051	−0.7713	0.3329	1.7332	0.8666	0.
4	= 1	≧ 0.4 ≦ 0.6	≧ −0.08 ≦ −0.12	0.2201	0.0045	−3.5123	1.1387	0.9371	0.5623	−0.0750

[a] The selection intensity is assumed to be 1

Table 3. Process of calculation of the last example in Table 2

Equality constraints of desired relative genetic gains			Satisfies the inequalities?	Variance of the index
EW	EP	BW		
= 1			No	0.2901
= 1	= 0.4		No	0.2938
= 1	= 0.6		No	0.2901
= 1		= − 0.12	No	1.9235
= 1		= − 0.08	No	1.1359
= 1	= 0.4	= − 0.12	Yes	1.9444
= 1	= 0.4	= − 0.08	Yes	1.9292
= 1	= 0.6	= − 0.12	Yes	1.1506
= 1	= 0.6	= − 0.08	Yes	1.1387

index has a smaller variance and gives larger expected genetic gains in EW and EP than Index 1.

The examples in Table 2 were computed using the procedure described in *"Computing techniques."* Intermediate results illustrating the computations for the last example in Table 2 are presented in Table 3. Nine indices for nine combinations of equality constraints of desired relative genetic gains are calculated, and four of them satisfy the inequalities; among them the 7th index has the minimum variance, 4.3583, so this is the optimum index.

References

Brascamp EW (1979) Selection index for desired gains. 30th Annual Meeting of the European Association of Animal Production, Harrogate, G 6.5

Essl A (1981) Index selection with proportional restriction: another view point. Z Tierz Züchtungsbiol 98:125–131

Harville DA (1974) Optimal procedures for some constrained selection problems. J Am Stat Assoc 69:446–452

Harville DA (1975) Index selection with proportionality constraints. Biometrics 31:223–225

Hazel LN (1943) The genetic basis for constructing selection indexes. Genetics 28:476–490

Itoh Y, Yamada Y (1986) Re-examination of selection index for desired gains. Genet Sel Evol 18:499–504

Kempthorne O, Nordskog AW (1959) Restricted selection indices. Biometrics 15:10–19

Lin CY (1978) Index selection for genetic improvement of quantitative characters. Theor Appl Genet 52:49–56

Mallard J (1972) La theorie et le calcul des index de selection avec restrictions: synthese critique. Biometrics 28:713–735

Pešek J, Baker RJ (1969) Desired improvement in relation to selection indices. Can J Plant Sci 49:803–804

Rao CR (1962) Problems of selection with restrictions. J R Stat Soc Ser B 24:401–405

Rao CR (1965) Problem of selection involving programming techniques. Proc IBM Sci Comp Symp Stat, White Plains, New York. IBM Data Processing Division, pp 29–51

Rouvier R (1977) Mise au point sur le modèle classique d'estimation de la valeur génétique. Ann Genet Sel Anim 9: 17–26

Smith HF (1936) A discriminant function for plant selection. Ann Eugen 7:240–250

Tallis GM (1962) A selection index for optimum genotype. Biometrics 18:120–122

Tallis GM (1985) Constrained selection. Jpn J Genet 60: 151–155; Corrigendum and addendum, Jpn J Genet 61: 183–184

Yamada Y, Yokouchi K, Nishida A (1975) Selection index when genetic gains of individual traits are of primary concern. Jpn J Genet 50:33–41