JAPAN. J. GENETICS Vol. 50, No. 1: 33-41 (1975)

SELECTION INDEX WHEN GENETIC GAINS OF INDIVIDUAL TRAITS ARE OF PRIMARY CONCERN

YUKIO YAMADA, KUNIO YOKOUCHI AND AKIRA NISHIDA

Department of Breeding and Genetics, National Institute of Animal Industry, 959 Aoba-cho, Chiba 280

Received July 10, 1974

INTRODUCTION

Since the method of selection index for improving multiple objectives was first introduced by Hazel in 1943 to animal breeding work, it has been shown by several authors, e.g. Hazel and Lush (1942), Young (1961), and Finney (1962), that the method is more efficient relative to independent culling levels and tandem selection. The extension of the theory to cover the restriction of no change in some traits was made by Kempthorne and Nordskog (1959). Henderson (1963) and more recently Mallard (1972) have presented a thorough review of the subject. The selection index is now finding an increasing use in modern animal breeding practice due to the relative ease with complicated calculation by computers.

In constructing a selection index it is necessary to have the following information: An $n \times n$ phenotypic variance-covariance matrix; an $n \times m$ genetic variance-covariance matrix; and a vector of relative economic weights of the m traits in the aggregate genotype.

The essential part of the procedure for constructing a selection index is to obtain a vector of weighting factors to be used in the index by maximizing the correlation between the aggregate genotype and the index.

The definition of the aggregate genotype is however quite ambiguous and thus it is common in practice to include only the most important economic traits conveniently chosen by each worker. Recently, Gjedrem (1972) questioned how aggregate ought to be defined and concluded that it should include all traits of economic improtance, each trait being weighted by its relative economic value. He has proposed further that economically important traits which are not recorded, e.g. feed efficiency, should as a rule be included also.

The assessment of the relative economic value to each component genotype in the aggregate genotype is not easy in some cases, because no standard for assessing the economical importance of the trait is readily available and in some traits, *e.g.* egg weight in poultry and fat percentage in dairy cattle, the assumption that relative economic weights remain constant over the range of variation is not satisfied (Cunningham, 1969).

Consequently, some adjustments of relative economic weights are often made after constructing a *tentative* index, by taking the accuracy of the index and expected correlated gains of important traits into consideration. This indicates that, in practice, the economic importance of some traits included in the aggregate genotype was, to some extent, assessed subjectively by breeders rather than derived by detailed economical analyses.

From the breeder's point of view, these economic weights are often not of his primary concern, while he is interested in the change in each trait of economic importance. Then, a question has arisen if the selection index which satisfies the breeder's intention to change these traits to the level he desires is obtainable.

The present report concerns the selection index constructed from the breeder's point of view, based on his intended genetic change of individual traits when the relative economic values of these traits are unknown or even ignored.

THEORY

We shall suppose that the breeder wants to change the means of m traits by the amount of $Q_j, j=1, 2, \ldots, m$. Then, $\boldsymbol{Q}$ which is defined as an $m\times 1$ vector of intended genetic changes for m traits assigned by the breeder is

$$\boldsymbol{Q}'=[Q_1, Q_2, \ldots, Q_m] \tag{1}$$

where the prime denotes the transpose of the original vector or matrix throughout in this paper.

In order to attain the assigned breeding goal or breeder's objective selection is made on the basis of the following index,

$$I=\boldsymbol{b}'\mathrm{X} \tag{2}$$

where $\boldsymbol{X}$ is an $n\times 1$ vector of sources of information, usually phenotipic measurements on the candidate for selection or its relatives, and $\boldsymbol{b}$ is an $n\times 1$ vector of weighting factors.

Since the expected genetic gain per generation in the jth trait, defined as $\Delta G_{j\cdot I}$, when selection is made on the index is expressed as the regression of the breeding value of the jth trait on the index,

$$\Delta G_{j\cdot I}=\frac{i_I}{\sigma_I}\mathrm{Cov}\,(G_j, I) \tag{3}$$

where i_I is the standardized selection differential on the index, σ_I is the standard deviation of the index, and $\mathrm{Cov}\,(G_j, I)$ is the covariance of the breeding value of the jth trait and the index. This covariance is rewritten as

$$\begin{aligned}\mathrm{Cov}\,(G_j, I)&=\mathrm{Cov}\,(G_j, \boldsymbol{b}'\boldsymbol{X})\\&=[\boldsymbol{G}'\boldsymbol{Rb}]_j\end{aligned} \tag{4a}$$

where $[\boldsymbol{G}'\boldsymbol{Rb}]_j$ is the jth element of $[\boldsymbol{G}'\boldsymbol{Rb}]$.

Substituting (4a) into (3), we have

$$\Delta G_{j \cdot I}=\frac{i_I}{\sigma_I}[\boldsymbol{G}'\boldsymbol{R}\boldsymbol{b}]_j. \tag{4b}$$

The $m \times 1$ vector of the expected genetic gains per generation in m traits, denoted as $\Delta\boldsymbol{G}$, is

$$\Delta\boldsymbol{G}=\frac{i_I}{\sigma_I}[\boldsymbol{G}'\boldsymbol{R}\boldsymbol{b}] \tag{4c}$$

where $\boldsymbol{G}$ is an $n \times m$ matrix of genetic covariances between the elements of $\boldsymbol{X}$ and those in $\boldsymbol{Q}$ on the individual basis, $\boldsymbol{R}$ is an $n \times n$ diagonal matrix of r_i, which is Wright's coefficient of relationship between the candidate and the relative(s) who provides the information of X_i.

If the goal assigned by the breeder is attained by q generations of selection, under the assumption of no changes in the population parameters during the coures of selection, the vector $\boldsymbol{Q}$ is

$$\boldsymbol{Q}=q\Delta\boldsymbol{G} \tag{5a}$$

$$=q\left(\frac{i_I}{\sigma_I}\right)[\boldsymbol{G}'\boldsymbol{R}\boldsymbol{b}]. \tag{5b}$$

In order to solve the equation (5b) with respect to $\boldsymbol{b}$, we set

$$q\left(\frac{i_I}{\sigma_I}\right)=1 \tag{6}$$

because this quantity is merely a constant multiplier to $\boldsymbol{b}$.

The solution is therefore

$$\boldsymbol{b}=(\boldsymbol{G}'\boldsymbol{R})^{-1}\boldsymbol{Q}. \tag{7}$$

The standard deviation of σ_I, is

$$\sigma_I=\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}} \tag{8}$$

where $\boldsymbol{P}$ is an $n \times n$ matrix of covariances between elements of $\boldsymbol{X}$.

Equation (7) can be solved in the case that the matrix $\boldsymbol{G}'\boldsymbol{R}$ is non-singular. The necessary conditions for this are that $n=m$ and that the rows and columns of the matrix are all independent each other.

Next, consider how many generations are required to attain the breeder's objective, $\boldsymbol{Q}$, by selecting for the derived index. This is obtained from the equation (6). That is, for example,

$$q=\frac{\sigma_I}{i_I}=\frac{\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}}}{i_I}. \tag{9}$$

If y is the average generation interval in years, the breeder's goal is expected to be attained by $t=qy$ years.

If the breeder wants to know how strongly he should select for the index so as to attain the goal within a given period of years, say t years, this is obtained from

$$i_I = y\frac{\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}}}{t}. \tag{10}$$

It is assumed of course that the population parameters do not change during the course of selection. This assumption may not be realistic but would be suffice to get equations (9) and (10).

The traits in $\boldsymbol{Q}$ are not necessarily the same to those which are included in the index as its component traits. The situation may be illustrated in the manner shown in Fig. 1, in which U is a universal set which consists of infinite number of traits and traits 1, 2, 4 and 5 are in the subset $\boldsymbol{Q}$ and traits 2, 3, 5 and 6 are in the subset $\boldsymbol{I}$, while trait 7 belongs to neither $\boldsymbol{Q}$ nor $\boldsymbol{I}$. This means that traits 2, 3, 5 and 6 are included in the index so as to improve traits 1, 2, 4 and 5.

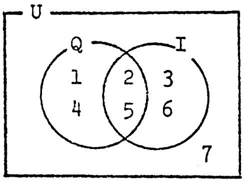

Fig. 1. Identifications of traits.
$\boldsymbol{U}$ is an universal set of the traits,
$\boldsymbol{Q}$ is the subset of $\boldsymbol{U}$ in which the traits are of breeder's concern,
$\boldsymbol{I}$ is the subset of $\boldsymbol{U}$ consisted of the traits taken into the selection index.

The genetic gains in all traits of our concern which includes of course the assigned m traits are obtained from

$$\Delta \boldsymbol{G}^* = \frac{i_I}{\sigma_I}\boldsymbol{G}^{*\prime}\boldsymbol{R}\boldsymbol{b} \tag{11}$$

where $\Delta \boldsymbol{G}^*$ is the vector whose elements are consisted of $\Delta G_{k\cdot I}$, $k=1, 2, \ldots, m, l$, and $\boldsymbol{G}^*$ is an $n\times 1$ matrix of genetic covariance (elements $\sigma_{G_{ik}}$). After q generations of selection, the total changes of these traits are

$$\boldsymbol{Q}^* = q\Delta \boldsymbol{G}^* = \boldsymbol{G}^{*\prime}\boldsymbol{R}\boldsymbol{b}. \tag{12}$$

The limitation which is encountered in solving equation (7) with respect to $\boldsymbol{b}$, *i.e.* $n=m$, can be removed by the use of equation (13) instead of equation (7).

$$\boldsymbol{b} = \boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{G}[\boldsymbol{G}'\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{G}]^{-1}\boldsymbol{Q}. \tag{13}$$

If the number of information is equal to that of the traits in $\boldsymbol{Q}$, *i.e.* $n=m$, we see that

$$[(\boldsymbol{G}'\boldsymbol{R})(\boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{G})]^{-1} = (\boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{G})^{-1}(\boldsymbol{G}'\boldsymbol{R})^{-1}. \tag{14}$$

Substituting equation (14) into (13), we have

$$\begin{aligned}\boldsymbol{b} &= (\boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{G})(\boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{G})^{-1}(\boldsymbol{G}'\boldsymbol{R})^{-1}\boldsymbol{Q} \\ &= (\boldsymbol{G}'\boldsymbol{R})^{-1}\boldsymbol{Q}\end{aligned}$$

which reduces to (5b).

The equation (13) is therefore a more general form applicable to any case in which the number of information is not necessarily equal to that of the traits in breeder's objective and more than two sources of information are available for the same trait, except that any phenotypic correlation between the traits in $\boldsymbol{P}$ is unity.

The theoretical basis for deriving equation (13) and a discussion on the association between economic weights and restrictions will be presented in our later paper.

DISCUSSION

The characteristic feature of this selection index is to have a unique application to the case where relative economic values of individual traits are unknown or otherwise difficult to assess, provided that the breeder has a definite breeding objective for improving individual traits of his concern.

The breeder's intention may vary with the level of performance and also among strains. The breeder has to set up the most efficient breeding plans by constructing selection indexes, each fits to his breeding objective in each strain. In some strains, the level of a certain trait is already so high that the breeder wants to maintain the level of that trait or even decrease it to the level to be balanced with other traits. In this case, conditional selection index with some restriction derived by Kempthorne and Nordskog (1959) should be obtained. Nevertheless, as they stated, the economic weight associated with the trait subjected to the restriction becomes irrelevant by imposing the restriction. This implies that the aggregate genotype after imposing such constraints is no longer the same to the one originally defined. This makes it nonsense to calculate ρ_{HI} or the efficiency of a restricted index relative to the unrestricted one.

The problems of restricted index have been considered by Rao (1962), Tallis (1962), James (1968) and Cunningham *et al.* (1970). In particular the procedure presented by Tallis (1962) and its extension by James (1968) are very similar to our method presented in this paper, except that the economic weights are completely eliminated in our formulae (7) and (13).

The efficiency of a restricted index proposed here can be compared in terms of q, the number of generations needed to attain the goal, $\boldsymbol{Q}$. A particular example is included in the numerical example.

One may argue the difficulty with our procedure in practice to choose a set of proper levels of objective traits. It is however not difficult to find $\boldsymbol{Q}$ for an experienced breeder because he must know the relative merit and demerit of his strain from appropriate information such as the Random Sample Test or a critical comparison of his with competitor's stocks. If hybridization of pure strains is common in practice, the $\boldsymbol{Q}$ in the pure strain can be adjusted to the level of average heterosis in each component trait.

By the use of this index, it is also not necessary to introduce a quadratic term in the aggregate genotype as Rao (1962) and Kempthorne and Nordskog (1959) have proposed. Because Q_j can be assigned to maintain the optimum if it is desired, without manipulating the aggregate genotype. It should be noticed that the introduction of a quadratic term for the optimum genotype is a different problem from setting the correlated gain of the trait be zero or equal to the assigned level.

A NUMERICAL EXAMPLE

Suppose a breeder wants to improve one of his flock of poultry from the present levels of performance, 65% egg production rate, 2.8 feed requirement and 58 g egg weight to the levels of 73%, 2.5 and 5.8 g, respectively. Since it takes two years to use 500-day egg production and it is quite laborious to measure feed intake of individual birds, the breeder prefers to use an index which consists of egg weight, egg production rate to 275 days of age and adult body weight. Egg and body weights are recorded individually, while the rate of egg production is based on 8 full sisters' average. The correlated gains in sexual maturity and adult viability are also of the breeder's interest. The basic parameters of a flock are given in Table 1.

Table 1. A hypothetical parameters of various traits

Trait number	Code*	Unit	σ_P	h^2	Trait number 1	2	3	4	5	6	7
1	EP_{500}	%	10	.20	r_P \ r_G	−.60	−.60	.85	0	−.50	0
2	FC	1/10	2	.20	−.80		−.40	−.50	.30	.20	0
3	EW	g	4	.50	−.10	−.30		.30	.40	.30	0
4	EF_{275}	%	10	.30	.80	−.40	−.10		0	−.50	.05
5	BW	10 g	18	.40	.10	.40	.40	−.05		.10	.10
6	SM	days	12	.30	−.30	.20	.20	−.40	.20		0
7	AV	%	30	.05	0	0	0	0	.10	0	

* EP_{500}=egg rate to 500 days, EP_{275}=egg rate to 275 days, FC=feed requirement, EW=egg weight, BW=body weight, SM=sexual maturity, and AV=adult viability.

$$\boldsymbol{Q}'=[8 \quad -3 \quad 0], \qquad m=3.$$

X_1=EW, X_2=Average of 8 full sisters of EP_{275}, and X_3=BW, $n=3$.
$r_1=1$, $r_2=.50$, and $r_3=1$.

$$\boldsymbol{G}=\begin{bmatrix} \sigma_{G31} & \sigma_{G32} & \sigma_{G3}^2 \\ \sigma_{G41} & \sigma_{G42} & \sigma_{G43} \\ \sigma_{G51} & \sigma_{G52} & \sigma_{G53} \end{bmatrix}=\begin{bmatrix} -7.5895 & -1.0119 & 8 \\ 20.8207 & -2.4495 & 4.6476 \\ 0 & 3.0547 & 12.8798 \end{bmatrix}$$

$$\boldsymbol{R}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$(\boldsymbol{G}'\boldsymbol{R})^{-1}=\begin{bmatrix} -7.5896 & 10.4104 & 0 \\ -1.0119 & -1.2248 & 3.0547 \\ 8 & 2.3238 & 12.8798 \end{bmatrix}^{-1}$$

$$=\begin{bmatrix} -.04058 & -.2379 & .05642 \\ .06647 & -.1734 & .04113 \\ .01321 & .1790 & .03518 \end{bmatrix}$$

$$\boldsymbol{b}'=[.3891 \quad 1.0520 \quad -.4313].$$

The index is therefore,

$$I=.389X_1+1.052X_2-.431X_3 \quad \text{or}$$
$$I=.389P_3+1.052\bar{P}_{4(F)}-.431P_5.$$

$$\boldsymbol{P}=\begin{bmatrix} \sigma_{P_3} & \sigma_{P_3\bar{P}_4} & \sigma_{P_3P_5} \\ \sigma_{\bar{P}_4P_3} & \sigma^2_{\bar{P}_4} & \sigma_{\bar{P}_4P_5} \\ \sigma_{P_5P_3} & \sigma_{P_5\bar{P}_4} & \sigma^2_{P_5} \end{bmatrix}=\begin{bmatrix} 16. & 1.5999 & 28.8 \\ 1.5333 & 25.625 & -1.125 \\ 28.8 & -1.125 & 324 \end{bmatrix}$$

$$\sigma_I=\sqrt{\boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}}=\sqrt{83.6651}=9.1469.$$

$$q=9.1469/i_I \quad \text{generations.}$$

$$\boldsymbol{G}^*=\begin{bmatrix} \sigma_{G_{31}} & \sigma_{G_{32}} & \sigma^2_{G_3} & \sigma_{G_{34}} & \sigma_{G_{35}} & \sigma_{G_{36}} & \sigma_{G_{37}} \\ \sigma_{G_{41}} & \sigma_{G_{42}} & \sigma_{G_{43}} & \sigma^2_{G_4} & \sigma_{G_{45}} & \sigma_{G_{46}} & \sigma_{G_{47}} \\ \sigma_{G_{51}} & \sigma_{G_{52}} & \sigma_{G_{53}} & \sigma_{G_{54}} & \sigma^2_{G_5} & \sigma_{G_{56}} & \sigma_{G_{57}} \end{bmatrix}$$

$$=\begin{bmatrix} -7.5895 & -1.0119 & 8 & 4.6476 & 12.6798 & 5.5771 & 0 \\ 20.8207 & -2.4495 & 4.6476 & 30 & 0 & -18 & 1.8371 \\ 0 & 3.0547 & 12.8798 & 0 & 129.6 & 7.4825 & 1.08 \end{bmatrix}$$

$$\Delta\boldsymbol{G}^{*\prime}=(i_I/\sigma_I)\boldsymbol{b}'\boldsymbol{R}\boldsymbol{G}^*$$
$$=[.8746 \quad -.3280 \quad 0 \quad 1.9228 \quad -5.5634 \quad -1.1507 \quad .0547]$$

$$\boldsymbol{Q}^{*\prime}=[8.0 \quad -3.0 \quad 0 \quad 17.59 \quad -50.89 \quad -10.53 \quad .50]$$

Correlated gains	Per generation ($\times i_I$)	After q generations
EP_{500}	.87%	8%
FC	−.03	−.3
EW	0	0
EP_{275}	1.92%	17.6%
BW	−.06	−.5 kg
SM	−1.15	−10.5 days
AV	.05%	.5%

In order to utilize further information on individual records for egg production rate, the condition is $m=3$ and $n=4$, which leads us to use equation (13). The efficiency of the index which is consisted from more information relative to the former index reflects to the decrease in the number of generations to attain the goal. The derived index is

$$I=.389P_3+.268P_4+.516\bar{P}_4(F)-.431P_5$$

$$q=8.7448/i_I$$

The efficiency is therefore $E=9.1469/8.7448=1.05$, 5% more efficient than the original index.

Correlated gains	Per generation $(\times i_I)$	After q generations
EP_{500}	.91%	8%
FC	−.03	−.3
EW	0 g	0 g
EP_{275}	2.01%	17.6%
BW	−.06 kg	−.5 kg
SM	−1.2 days	−10.5 days
AV	.06%	.5%

SUMMARY

A new procedure for constructing economic weight free selection index, based on exclusively breeder's intention, is presented. In its derivation the aggregate genotype which is absolutely required in the conventional method was ignored. In this sense this method is completely new and is applicable to the case in which the relative economic weight of each trait is unknown or difficult to assess because of no available standard for the assessment.

The use of the technique to evaluate the breeder's objective whether it is attainable within a given period of time with a given selection intensity is described.

This selection index has a wide application to practical breeding operations, since it includes several conditions of constraints proposed by other workers. The formula for predicting genetic gains in those traits which are not of the breeder's primary objective is also presented.

LITERATURE CITED

Cunningham, E. P., 1969 Animal Breeding Theory. Agr. College of Norway, Oslo.

Cunningham, E. P., R. A. Moen, and T. Gjedrem, 1970 Restrictions of selection indexes. Biometrics **26**: 67–74.

Finney, D. J., 1962 Genetic gains under three methods of selection. Genet. Res. **3**: 417–423.

Gjedrem, T., 1972 A study of the definition of the aggregate genotype in a selection index. Acta Agr. Scand. **22**: 11–16.

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

Hazel, L. N., and J. L. Lush, 1942 The efficiency of three methods of selection. J. Herd. **33**: 393–399.

Henderson, C. R., 1963 Selection index and expected genetic advance. In Statistical Genetics and Plant Breeding. NAS-NRC **982**: 141–163.

James, J. W., 1968 Index selection with restrictions. Biometrics **24**: 1015–1018.

Kempthorne, O., and A. W. Nordskog, 1959 Restricted selection indices. Biometrics **15**: 10–19.

Mallard, J., 1972 La theorie et le calcul des index de selection avec restricitions: Synthese critique. Biometics **28**: 713–735.

Rao, C. R., 1962 Problems of selection with restrictions. J. R. Statist. Soc. **B24**: 401–405.

Tallis, G. M., 1962 A selection index for optimum genotype. Biometrics **18**: 120–122.

Young, S. S. Y., 1961 A further examination of the relative efficiency of three methods of selection for genetic gains under less-restricted conditions. Genet. Res. **2**: 106–121.