Seuils de troncature optimaux lors d'une sélection à niveaux Indépendants sur une distribution multinormale, avec une application à la sélection chez les bovins laitiers.

Vincent Ducrocq, J Jacques J. J. Colleau

Original article

Optimum truncation points for independent culling level selection on a multivariate normal distribution, with an application to dairy cattle selection

V. Ducrocq and J.J. Colleau

Institut National de la Recherche Agronomique, Station de Génétique Quantitative et Appliquée, Centre de Recherches de Jouy-en-Josas, 78350 Jouy-en-Josas, France

Summary — Independent culling level selection is often practiced in breeding programs because extreme animals for some particular traits are rejected by breeders or because records on which genetic evaluation is based are collected sequentially. Optimizing these selection procedures for a given overall breeding objective is equivalent to finding the combination of truncation thresholds or culling levels which maximizes the expected value of the overall genetic value for selected animals. A general Newton-type algorithm has been derived to perform this maximization for any number of normally distributed traits and when the overall probability of being selected is fixed. Using a powerful method for the computation of multivariate normal probability integrals, it has been possible to undertake the numerical calculation of the optimal truncation points when up to 6 correlated traits or stages of selection are considered simultaneously. The extension of this algorithm to the more complex situation of maximizing annual genetic response subject to nonlinear constraints is demonstrated using a dairy cattle model involving milk production and a secondary trait such as type. Consideration is given to three of the four pathways of selection: dams of bulls; sires of bulls; and sires of cows.

Independent culling level selection – dairy cattle – multistage selection – genetic gain – multivariate normal distribution

Résumé — Seuils de troncature optimaux lors d'une sélection à niveaux indépendants sur une distribution multinormale, avec une application à la sélection chez les bovins laitiers. *Une sélection à niveaux indépendants est souvent pratiquée dans les programmes génétiques, parce que les animaux extrêmes pour certains caractères sont rejetés, ou parce que les données qui servent à l'évaluation génétique des animaux sont recueillies séquentiellement. L'optimisation, pour un objectif donné, de ces règles de sélection équivaut à la recherche des seuils de troncature qui maximisent l'espérance de l'objectif de sélection pour les animaux retenus. Un algorithme général de type Newton est établi pour effectuer cette maximisation pour un nombre quelconque de caractères distribués selon une loi multinormale et lorsque la probabilité finale d'être retenu est fixée. A partir d'une méthode puissante de calcul d'intégrales de lois multinormales, il a été possible d'entreprendre numériquement le calcul des seuils de troncature quand jusqu'à 6 caractères ou étapes de sélection corrélés sont considérés simultanément. L'extension de cet algorithme à des situations plus complexes, comme la maximisation du progrès génétique annuel sous plusieurs contraintes non linéaires, est illustrée à travers le calcul de règles optimales de sélection des mères à taureau, pères à taureau et pères de service pour la production laitière et pour un caractère secondaire tel que le pointage laitier dans un schéma de sélection typique des bovins laitiers.*

sélection à niveaux indépendants – bovins laitiers – sélection par étapes – progrès génétique – distribution multinormale

Introduction

It is often possible to describe selection objectives through linear combinations – aggregate genotype – of breeding values on several (m) traits. Then the optimal selection procedure consists of using a selection index which combines observed values for several (n) sources of information (Hazel and Lush, 1942).

However, this approach is occasionally not used for two main reasons:

a) Practitioners sometimes emphasize the need to cull extreme animals (deviants for some traits) because they are deemed undesirable. Then, the selection objective is implicitly recognized as being nonlinear. The cost of such a practice with respect to a strict application of the optimal linear index may not be justifiable when this nonlinearity is unimportant or questionable.

b) The records required to compute the selection index are not always available simultaneously and/or their cost does not justify their collection for all the candidates for selection. Therefore selection schemes involve different stages which correspond to truncations on the joint distribution of all possible records.

Within these constraints, it is potentially interesting to evaluate and improve the efficiency of practical selection procedures by determining optimal conditions of application of independent culling level selection. As far as we know, the published works on this topic have been greatly limited by the number of variables considered. Generally speaking, studies on the algebraic derivation of optimum truncation thresholds with corresponding numerical computation have dealt with no more than 2 variables (Namkoong, 1970; Evans, 1980; Cotterill and James, 1981; Smith and Quaas, 1982). Recently, Ducrocq and Colleau (1986) treated examples with 3 variables to illustrate potential uses of a numerical method to compute multivariate normal probabilities. In practice, though, the number of traits or selection stages involved may be significantly larger. Moreover, the algorithms previously considered have been specifically developed for a given number n of variables and their extension to any n is not obvious though algebraic conditions which must be verified at the optimum have been reported (Jain and Amble, 1962; Smith and Quaas, 1982).

Indeed, as a consequence of the limitation of the number of variables considered, the optimization problem has been restricted to rather simple types of objective functions, which may not adequately summarize the overall efficiency of selection schemes. In particular, authors have not considered functions such as the annual genetic gain of the selection objective – computed using Rendel and Robertson's (1950) formula – in relatively complex situations (*e.g.* involving the 4 paths of transmission of genetic progress).

This paper presents basic yet general, *i.e.* for any n algorithms which can be used for the computation of optimal truncation points for a broad class of objective functions with one or more constraints. Theory is developed and applications are presented for the general multivariate problem considered by Smith and Quaas (1982). A practical example of application in the dairy cattle context is described. Corresponding numerical results are given.

Solution of Smith and Quaas' problem in the general case

Statement of the problem

Let u, x_1, ...x_n be $n+1$ random variables with joint multivariate normal distribution:

$$N\left(\begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & r_{u1} & r_{u2} & \cdot & r_{un} \\ & 1 & r_{12} & \cdot & r_{1n} \\ & & 1 & \cdot & \cdot \\ \mathrm{sym} & \cdot & \cdot & \cdot & 1 \end{bmatrix}\right)$$

u is the breeding objective and the x_is are the observed variables.

Define $E_P(u) = E(u \mid x_1 > c_1, ... x_n > c_n)$

The problem is to find c_1, c_2,...c_n so as to maximize $E_p(u)$ subject to:

$Pr(x_1 > c_1, ... x_n > c_n) = P$,

where P is given and represents the overall fraction of candidates selected.

Notations

Let $\phi_n(\mathbf{x}; \mathbf{R}_n)$ be the standard multivariate normal density of dimension n with variance covariance matrix $\mathbf{R}_n$.

Let

$$Q(c_1, .. c_n) = \int_{c_1}^{+\infty} ... \int_{c_n}^{+\infty} \phi_n(\mathbf{x}; \mathbf{R}_n)\, d\mathbf{x}$$

We need the following recursive definitions for distributions conditioned on $q \le n-1$ variates.

$$\mathbf{R}_{n-q.k_1..k_q} = \{r_{ij.k_1.k_q}\} \quad 1 \le i, j \le n \quad i, j \ne k_1, ... k_q$$

$$r_{ij.k_1.k_q} = \frac{r_{ij.k_1.k_{q-1}} - r_{ik_q k_1.k_{q-1}}\, r_{jk_q k_1.k_{q-1}}}{\sqrt{(1 - r_{ik_q k_1..k_{q-1}}^2)(1 - r_{jk_q k_1..k_{q-1}}^2)}}$$

$$c_{i.k_1.k_q} = \frac{c_{i.k_1.k_{q-1}} - r_{ik_q k_1..k_{q-1}}\, c_{k_q k_1.k_{q-1}}}{\sqrt{1 - r_{ik_q k_1.k_{q-1}}^2}}$$

$$z_{i.k_1.k_q} = \phi_1(c_{i.k_1.k_q}) = \frac{e^{-\frac{1}{2}c_{ik_1..k_q}^2}}{\sqrt{2\pi}}$$

$$Q_{k_1..k_q} = \int_{c_{i.k_1..k_q}} ... \int \phi_{n-q}(\mathbf{x}; \mathbf{R}_{n-q.k_1.k_q})\, d\mathbf{x}$$
$$i \ne k_1, ... k_q$$

Solution

Using the general result of Jain and Amble (1962) we have:

$$E_P(u) = E(u \mid x_1 > c_1, \ldots x_n > c_n) = \frac{N(c_1, \ldots c_n)}{Q(c_1, \ldots c_n)} \tag{1}$$

with $N(c_1, \ldots c_n) = \sum_i r_{ui} \, z_i \, Q_i$

Since $Q(c_1, \ldots c_n)$ is to be equal to a constant P, the maximization of $E_p(u)$ is tantamount to the maximization of $N(c_1, \ldots, c_n)$. The constraint $Q(c_1, \ldots c_n) = P$ is incorporated using the method of Lagrange multipliers (Bass, 1961, p. 928; Smith and Quaas, 1982): the optimal truncation points are those for which the partial derivatives of the function $f(w) = N(c_1, \ldots c_n) + \lambda(Q(c_1, \ldots c_n) - P)$ with respect to $w' = (c_1, \ldots, c_n, \lambda)$ are 0. λ is called a Lagrange multiplier.

The resulting system of nonlinear equations in $w' = (c_1, \ldots c_n, \lambda)$ is solved iteratively using the multidimensional Newton's method (Dennis and Schnabel, 1983). Denote as $w^{(t)}$ the approximate solution at iteration t ($w^{(0)}$ is a given starting value).

A better estimate $w^{(t+1)}$ is computed from:

$$w^{(t+1)} = w^{(t)} - \left[\left(\frac{\delta^2 f(w)}{\delta w \, \delta w'} \right) \right]^{-1}_{w=w^{(t)}} \left[\frac{\delta f(w)}{\delta w} \right]_{w=w^{(t)}} \tag{2}$$

The final solution $w^{(t)} = w^*$ is obtained when

$$\left\| \frac{\delta f(w)}{\delta w} \right\|_{w=w^{(t)}}$$

is sufficiently small, where $\| \cdot \|$ denotes any norm.

As long as the starting value $w^{(0)}$ is not too far from w^* (generally, $w^{(0)} = 0$ seems to be a robust initial value), convergence is very fast (quadratic convergence: Dennis and Schnabel, 1983). c^* is a local maximum for $E(u)$, provided

$$\left(-\frac{\delta^2 f(w)}{\delta c \, \delta c'} \right)_{w^*}$$

is positive definite, but nothing guarantees that w^* is a global maximum for $f(w)$.

Now, note that:

$$\frac{\delta Q}{\delta c_i} = -z_i \, Q_i \tag{3}$$

So we can write:

$$N = -\sum_i r_{ui} \left(\frac{\delta Q}{\delta c_i} \right) \tag{4}$$

$$\frac{\delta N}{\delta c_j} = -\sum_i r_{ui} \left(\frac{\delta^2 Q}{\delta c_i \, \delta c_j} \right) \tag{5}$$

$$\frac{\delta^2 N}{\delta c_j \, \delta c_j} = -\sum_i r_{ui} \left(\frac{\delta^3 Q}{\delta c_i \, \delta c_j \, \delta c_k} \right) \tag{6}$$

where:

$$\left(\frac{\delta^2 Q}{\delta c_i \, \delta c_j} \right) = \frac{z_i \, z_{j \cdot i} \, Q_{ij}}{\sqrt{1 - r_{ij}^2}} \qquad i \neq j \tag{7}$$

$$\left(\frac{\delta^2 Q}{\delta c_i^2} \right) = -c_i \left(\frac{\delta Q}{\delta c_i} \right) - z_i \left[\sum_{i \neq m} \frac{r_{im}}{\sqrt{1 - r_{im}^2}} \, z_{m \cdot i} \, Q_{im} \right]$$

$$= -c_i \left(\frac{\delta Q}{\delta c_i} \right) - \sum_{i \neq m} r_{im} \left(\frac{\delta^2 Q}{\delta c_i \delta c_m} \right) \tag{8}$$

$$\left(\frac{\delta^3 Q}{\delta c_i \, \delta c_j \, \delta c_k} \right) = \frac{-z_i \, z_{j \cdot i} \, z_{k \cdot ij} \, Q_{ijk}}{\sqrt{(1 - r_{ij}^2)(1 - r_{ik}^2)(1 - r_{jk \cdot i}^2)}} \qquad i \neq j \neq k \tag{9}$$

$$\left(\frac{\delta^3 Q}{\delta c_i^2 \, \delta c_j} \right) = \frac{1}{\sqrt{1 - r_{ij}^2}} \left[-c_i \, z_i \, z_{j \cdot i} \, Q_{ij} + \frac{r_{ij} \, c_{j \cdot i} \, z_i \, z_{j \cdot i} \, Q_{ij}}{\sqrt{1 - r_{ij}^2}} \right.$$

$$\left. -z_i \, z_{j \cdot i} \left[\sum_{m \neq i \neq j} \frac{z_{m \cdot ij} Q_{ijm}}{\sqrt{1 - r_{jm \cdot i}^2}} \cdot \left(\frac{-r_{im}}{1 - r_{im}^2} + \frac{r_{jk \cdot i} \, r_{ij}}{\sqrt{1 - r_{ij}^2}} \right) \right] \right]$$

$$= \left(-c_i + \frac{r_{ij} \, c_{j \cdot i}}{\sqrt{1 - r_{ij}^2}} \right) \left(\frac{\delta^2 Q}{\delta c_i \, \delta c_j} \right) + \sum_{m \neq i \neq j} \left(\frac{r_{jm} \, r_{ij} - r_{im}}{1 - r_{ij}^2} \right) \left(\frac{\delta^3 Q}{\delta c_i \, \delta c_j \, \delta c_m} \right) \quad i \neq j \tag{10}$$

$$\left(\frac{\delta^3 Q}{\delta c_i^3} \right) = -c_i \left(\frac{\delta^2 Q}{\delta c_i^2} \right) - \left(\frac{\delta Q}{\delta c_i} \right) - \sum_{m \neq i} r_{im} \left(\frac{\delta^3 Q}{\delta c_i^2 \, \delta c_j} \right) \tag{11}$$

Another method exists for the derivation of these expressions, first reasoning on derivatives of integrals and then using Jain and Amble's (1962) formula on conditional distributions. This leads to more compact expressions but may be less flexible for considera-

tions on several u (general problem with several constraints on several objectives) (see Appendix).

Expressions (3) to (11) include all the elements required for the computation of the vector $\delta f(\mathbf{w})/\delta \mathbf{w}$ and the matrix $(\delta^2 f(\mathbf{w})/\delta \mathbf{w}\, \delta \mathbf{w}^*)$ in (1). In particular,

$$\frac{\delta f(\mathbf{w})}{\delta c_i} = \frac{\delta N}{\delta c_i} + \lambda \frac{\delta Q}{\delta c_i} \qquad\qquad 1 \le i \le n \quad (12)$$

$$\frac{\delta f(\mathbf{w})}{\delta \lambda} = Q(c_1, \ldots c_n) - P \qquad\qquad (13)$$

$$\frac{\delta^2 f(\mathbf{w})}{\delta c_i \delta c_j} = \frac{\delta^2 N}{\delta c_i \delta c_j} + \lambda \frac{\delta^2 Q}{\delta c_i \delta c_j} \qquad\qquad 1 \le i, j \le n \quad (14)$$

$$\frac{\delta^2 f(\mathbf{w})}{\delta c_i \partial \lambda} = \frac{\delta Q}{\delta c_i} \qquad\qquad 1 \le i, j \le n \quad (15)$$

$$\frac{\delta^2 f(\mathbf{w})}{\delta \lambda^2} = 0 \qquad\qquad (16)$$

It can be observed that the equations $(\delta f(\mathbf{w})/\delta c_i) = 0$ in (12) for $1 \le i \le n$ are linear in λ. The size of the system of equations to be solved can be easily reduced by absorption of the Lagrange multiplier. For example, we have:

$$\frac{\delta f(\mathbf{w})}{\delta c_i} = 0 \implies \lambda = -\frac{\delta N}{\delta c_1} \left(\frac{\delta Q}{\delta c_1}\right)^{-1}$$

and $\dfrac{\delta f(\mathbf{w})}{\delta \mathbf{w}} = 0$ is equivalent to:

$$\begin{cases} \dfrac{\delta N}{\delta c_i} - \dfrac{\delta N}{\delta c_1} \left(\dfrac{\delta Q}{\delta c_1}\right)^{-1} \dfrac{\delta Q}{\delta c_i} = 0 \qquad 2 \le i \le n \\[4mm] Q(c_1, \ldots c_n) - P = 0 \end{cases} \qquad (17)$$

Derivatives with respect to the c_is of the equations in (17) are required for the application of Newton's method as in (2). They are readily derived using (6) to (8).

Numerical applications

Studies on independent culling level selection have been mainly limited to 2-trait selection probably because general and efficient programs to compute the multivariate normal probability integrals in Jain and Amble's formula were not available for dimensions > 2. However, easily programmable algorithms exist. In particular, Dutt (1973, 1975) and Dutt

and Soms (1976) proposed a general method characterized by good precision when correlation coefficients and truncation points are not too extreme. For more details on this method, its precision and computation times, see Ducrocq and Colleau (1986). Dutt's technique is well suited for numerical applications of the optimization algorithm presented in this paper when up to 6 selection stages or traits are considered.

1) • $n = 2$.

In this particular case, expressions (7) to (11) are simpler, since $Q_{ij} = 1$ and $Q_{ijk} = 0$. Algebraic and numerical results are equivalent to those given by Smith and Quaas (1982).

2) • $n = 3$.

In (9), $Q_{ijk} = 1$. The algorithm described here leads to the same results as those presented in Ducrocq and Colleau (1986).

3) • $n = 4$ to 6.

Consider for example, n traits with $r_{ij} = (-1)^j (j/20i)$ for $1 \leq i \leq j \leq n$ and with economic weight $m_i = 1 + i/20$ $1 \leq i \leq n$.

Table I presents the truncation points c_is on these traits which maximize $E_P (u \mid c_1,... c_n)$ when the overall selected fraction is $P = 0.25, 0.025, 0.001$. At iteration 0, the c_is were taken equal to 0. The stopping criterion for the Newton's iteration was:

$$\left(\sum_{i=1}^{n} E_i^2 \right) < 10^{-10}$$

where E_i was the i th left hand side of system (17).

Convergence was fast and depended on how far the initial value of the truncation points was from the solution. Note, however, that in the examples presented in Table I, correlations between variables are not very high ($\mid r_{ij} \mid \leq 0.3$) and the weights of the different traits are of the same magnitude. When this is not the case, the optimal selection procedure may involve no selection at all on one or several of these traits. The same observation applies to small overall selection intensity (Young, 1961; Namkoong, 1970; Smith and Quaas, 1982; Tibau I Font and Ollivier, 1984; Ducrocq and Colleau, 1986). In limiting cases (with very low or very high selection intensity on one or several traits or when correlations are extreme), it should be remembered that the precision of Dutt's algorithm for computation of multivariate probability integrals may be unsatisfactory (Ducrocq and Colleau, 1986). Then alternative methods may have to be used (*e.g.,* Russell *et al.,* 1985).

An application in the dairy cattle context

Assumptions

Dairy cattle selection is performed through a sequence of stages which characterizes the transition from one generation (g) to the next (g+1).

In the additive polygenic situation which is assumed for most of the traits selected in domestic animals, it is possible to describe these stages through truncation selection procedures on different variables (*e.g.,* Smith and Hammond, 1987).

Table I. Optimum truncation points for independent culling level selection on n correlated traits (or stages) ($r_{ij} = 1/20$ $(-1)^j j/i$ for i<j, weighting factors $m_i = 1 + i/20$).

No. of traits (n)	Overal selected fraction (P)	Iterations	Truncation points (first. line) and selected proportion (second line) (in %) at each stage						E(u|cj) (in σ_u)
			1	2	3	4	5	6	
4	0.25	4	−0.648 (74.1)	−0.579 (73.4)	−0.881 (82.2)	−0.054 (55.9)			1.011
	0.025	3	0.206 (41.9)	0.260 (43.4)	−0.066 (54.8)	0.836 (25.1)			1.984
	0.001	6	0.900 (18.4)	0.950 (20.9)	0.557 (31.0)	1.618 (8.4)			2.971
5	0.25	5	−1.098 (86.4)	−0.687 (76.2)	−0.990 (85.3)	0.018 (48.6)	−1.219 (91.6)		0.919
	0.025	3	−0.263 (60.4)	0.068 (49.8)	−0.219 (61.7)	0.803 (19.3)	−0.451 (69.8)		1.843
	0.001	5	0.342 (36.6)	0.680 (28.1)	0.358 (39.5)	1.504 (5.2)	0.071 (47.7)		2.810
6	0.25	6	−1.400 (91.9)	−1.083 (86.5)	−1.481 (93.6)	−0.445 (69.9)	−1.818 (93.9)	0.113 (51.2)	0.968
	0.025	4	−0.11 (65.9)	−0.255 (62.2)	−0.610 (74.7)	0.352 (40.8)	−0.881 (78.0)	0.969 (25.6)	1.889
	0.001	5	0.316 (37.6)	0.378 (39.1)	0.008 (52.0)	0.995 (19.3)	−0.293 (62.1)	1.692 (10.9)	2.841

These first include selection criteria corresponding to the transition between generations g and g+1 (reproductive stage), followed in the course of time by those criteria used during generation g+1, before the next reproductive cycle. Our approach for the optimization of these successive selection stages relies on the assumption of multivariate normality for these 2 criteria when candidates for selection are born. Such an assumption is plausible in the additive polygenic context, especially when heritability values are low (Bulmer, 1980, p. 154; see also Smith and Hammond, 1987, for a discussion on this point). A more strident assumption is that the dispersion parameters of the joint multivariate distribution remain constant through the different selection cycles.

Breeding objective and selection stages

Assume that the selection objective in a dairy cattle breed is a linear combination of 2 traits: "milk production" and a secondary trait such as "type" (both of these traits may be themselves linear combinations of more specific characters). A possible sequence of selection stages which approximates what is often done in practice is the following (Figure 1):

1) Dams of bulls (DB) of generation g are selected based on their estimated breeding values x_1 (for milk) and x_2 (for type), with respective thresholds c_1 and c_2 on the standardized variables. These dams of bulls are mated to sires of bulls (SB) of generation g.

2) The sons of these cows are progeny tested. Sires of cows (SC) and sires of bulls of generation g+1 are then selected according to their estimated breeding values x_3 (for

milk) and x_4 (for type). Truncation thresholds on these 2 variables are different for SC and SB (c_3, c_4 and c_5 c_6, respectively).

Selection of DB can be modelled as if it were performed at birth of the male calves. This is essential in order to be able to invoke the restoring of multivariate normality at each generation. Let RP be the registered (with known pedigree) and recorded population of cows and let γ denote the proportion of these cows which can be potential dams of bulls (e.g. $\gamma = 0.53$ if AI sons are selected from cows with at least 2 known lactations). If it is assumed (as in Ducrocq, 1984) that an average of $n_d = 6$ potential dams must be selected in order to obtain one male calf entering progeny test, it can be considered that DB selection is performed by truncation on the estimated breeding values x_1 and x_2 of the dams of $n_b = (\gamma \, RP)/n_d$ male calves.

The expression of the annual genetic gain given by Rendel and Robertson (1950) is:

$$\Delta G_a = \frac{I_{SB} + I_{SC} + I_{DB} + I_{DC}}{L_{SB} + L_{SC} + L_{DB} + L_{DC}} \tag{18}$$

Selection on the dam of cow path is ignored ($I_{DC} = 0$).

$$I_{DB} = E_{dam}(u \mid x_1 > c_1, x_2 > c_2) = 2 \, E_{son}(u \mid x_1 > c_1, x_2 > c_2) = 2 \, \frac{N(c_1, c_2)}{Q(c_1, c_2)} \tag{19}$$

where d_0 is the fraction of the whole population bred to young sires ($d_0 = T_Y \, RP/T$), i.e.:

$$I_{SC} = (1-d_0) \, [E(u \mid x_1 > c_1, x_2 > c_2, x_3 > c_3, x_4 > c_4) - E(u \mid x_1 > c_1, x_2 > c_2)] \tag{20}$$

$$I_{SC} = (1-d_0) \left(\frac{N(c_1, c_2, c_3, c_4)}{Q(c_1, c_2, c_3, c_4)} - \frac{N(c_1, c_2)}{Q(c_1, c_2)} \right) \tag{21}$$
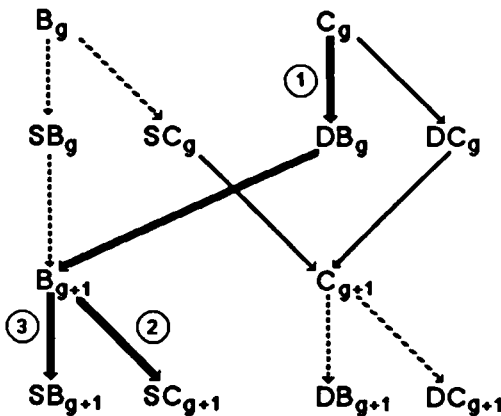
Fig. 1. Selection stages in a typical breeding scheme. B_g, C_g: bull and cow of generation g. SB, SC, DB, DC: sire of bulls, sire of cows, dam of bulls, dam of cows. (1) Selection of dams of bulls on milk (threshold: c_1) and type (c_2) estimated breeding value. (2) Selection of sires of cows on milk (c_3) and type (c_4) estimated breeding value. (3) Selection of sires of bulls on milk (c_5) and type (c_6) estimated breeding value.

Constraints

Three constraints are added here:

1) The fraction T_y of the population bred to young sires is considered as constant, since in practice this is usually the limiting factor for the extension of progeny test. In this example, the number of recorded daughters per young sire $n_{d/y}$ is also assumed constant: then, the number n_y of young sires progeny-tested each year is fixed, as well as their repeatablity.

$n_b \ Q \ (c_1, c_2) = n_y = $ constant

$$\Rightarrow Q \ (c_1, c_2) - d_1 = 0 \tag{22}$$

with $d_1 = n_y/n_b$.

2) The number of sires of cows selected each year is determined by the number of cows ($= (T-RP) + (1-T_y) \ RP$) to be bred to proven sires in the whole population (T) and the total number of doses produced by a given sire during his lifetime (AI).

$$Q \ (c_1, c_2, c_3, c_4) - d_2 = 0 \tag{23}$$

where $d_2 = [\ (T-RP) + (1-T_y) \ RP \ ]/AI$.

3) The number of sires of bulls retained each year is constrained to be equal to n_{SB}, the number below which problems of inbreeding and reduction of genetic variability are feared.

$$Q \ (c_1, c_2, c_5, c_6) - d_3 = 0 \tag{24}$$

where $d_3 = n_{SB}/n_b$.

Numerical methods and results

When constraints (22), (23) and (24) are satisfied, equations (18) to (21) lead to the following result:

$$\Delta G_a = \frac{1}{L} \left(\frac{d_0}{d_1} N(c_1, c_2) + \frac{(1-d_0)}{d_2} \ N(c_1, c_2, c_3, c_4) + \frac{1}{d_3} \ N(c_1, c_2, c_5, c_6) \right) \tag{25}$$

where L, the sum of the generation intervals over the 4 paths, is a constant in our case.

The combination of truncation points c_i, i=1,... 6 which maximizes (25) with the constraints (22), (23) and (24) is obtained by equating to 0 the derivatives of f (**w**) with respect to $\mathbf{w}' = (c_1,..c_n, \lambda, \mu, \nu)$ where:

$$f \ (\mathbf{w}) = \Delta G_a + \lambda \ [Q \ (c_1, c_2) - d_1] + \mu \ [Q \ (c_1, c_2, c_3, c_4) - d_2] + \nu \ [Q \ (c_1, c_2, c_5, c_6) - d_3]$$
$$\tag{26}$$

and λ, μ, ν are Lagrange multipliers.

The first and second derivatives of f (**w**) are readily obtained using the general formulae given in the preceding sections. The 3 Lagrange multipliers are eliminated through trivial absorption. The nonlinear system to solve then involves 6 unknowns: the 6 truncation thresholds. Solutions obtained using Newton's method are presented in Table III, where parameters take the values given in Table II. The stopping criterion for the Newton's iterations was:

$$\left(\sum_{i=1}^{6} E_i^{*2} \right) < 10^{-8}$$

where E_i^* was the ith left hand side of the absorbed system of equations.

Convergence was fast – always less than 8 iterations – as long as the starting value $\mathbf{c}^{(0)}$ of $\mathbf{c} = \{c_i\}_{i=1,6}$ was not too far from the solution. In contrast with the numerical results presented in Table I, $\mathbf{c}^{(0)} = 0$ does not lead to convergence, because the values of $\mathbf{c}$ in the next iterates are found to be extremely large or low; *i.e.*, totally unrealistic and in regions where the precision of Dutt's method is not good. This behavior is, at least indirectly, a consequence of the very large selection intensity of proven sires. To avoid this situation, more realistic starting values must be chosen. For example, $\mathbf{c}^{(0)}$ can be computed as the truncation points for which the conditions (22), (23) and (24) are satisfied with a given distribution of selection efforts between milk and type, *i.e.*, $\mathbf{c}^{(0)}$ is such that $Q(c_1^{(0)}) = d_1/\alpha$, $Q(c_1^{(0)}, c_2^{(0)}) = d_1$, $Q(c_1^{(0)}, c_2^{(0)}, c_3^{(0)}) = d_2/\alpha$, etc... for some α. In most cases presented in Table III, $\alpha = 0.75$. For strongly unbalanced weights for the 2 traits (5:1), convergence is obtained only for even larger values of α (*e.g.*, $\alpha = 0.99$).

It should be noted that selection on type was considered posterior to selection on milk production. This assumption has been chosen because this corresponds to what is done in practice for dams of bulls in France: only the best cows for milk production are evaluated for type. This does not influence the values of optimal thresholds but only their use when results are presented in terms of fraction selected at each stage.

Table II. Variables and assumptions considered in examples 1 and 2.

Variable	Symbol	Value
Heritability of milk production (m)	h^2_m	0.25
of type (t)	h^2_t	0.25
Number of recorded daughters per young sire for milk	N_m	50
Fraction of young sires' daughters recorded for type	N_t/N_m	0.75
Repeatability of genetic evaluation * for sires	$r^2_i = N_i /(N_i - 1) + 4/h^2_i$	
*for dams	$R^2_m =$	0.5
	$R^2_t =$	0.3
Total population	T	400,000
Recorded and registered population	RP	200,000
Fraction of RP bred to young sires	T_y	0.20
Number of cows bred to a young sire to get one recorded daughter	E	5
Number of cows bred to a sire during his career	AI	50,000
Number of potential dams of bulls selected to get one sampled young sire	n_d	6
Fraction of RP which is old enough to be DB	γ	0.53
Number of sires of bulls selected each year	n_{SB}	4
Generation intervals: young sires		2
proven sires		8
dams of cows	L_{DC}	4.5
dams of sires	L_{DS}	5.5

Table III. Optimum truncation points for "milk" and "type" on the different gene transmission pathways (sires of bulls, sires of cows, dams of cows) in example 1 (other parameter values are given in Table II).

No. of young sires/year	Genetic correlation	Economic weights	Proportion of candidates selected at each stage (%)						ΔG_a u	ΔG_a Milk	ΔG_a Type
			Dams of bulls		sires of cows		sires of bulls				
	milk: type	milk: type	milk	type	milk	type	milk	type	(in genetic SD)		
75	−0.15	2:1	0.7	64	14	73	8	64	0.183	0.183	0.018
		3:1	0.5	91	11	93	6	89	0.190	0.195	−0.009
		5:1	0.4	100	10	100	5	100	0.197	0.202	−0.030
	0	2:1	0.8	57	15	70	9	61	0.187	0.183	0.051
		3:1	0.5	82	11	89	6	84	0.192	0.194	0.023
		5:1	0.4	97	10	100	5	100	0.197	0.200	0.002
	0.15	2:1	0.8	51	15	67	9	59	0.190	0.185	0.080
		3:1	0.6	74	12	85	7	80	0.194	0.194	0.057
		5:1	0.5	92	10	100	5	100	0.197	0.200	0.033
150	0.15	2:1	1.5	57	8	58	5	52	0.202	0.196	0.088
		3:1	1.1	78	6	79	4	74	0.206	0.206	0.063
		5:1	0.9	93	5	100	3	100	0.209	0.212	0.035

The results in Table III underline the sensitivity of this type of computation to the economic weights assigned to the 2 traits and their genetic correlation. According to the relative economic weight for type, the optimal situation may correspond to virtually no selection on type (weights 5:1) or to a culling on type evaluation of between a quarter and a half of the candidates (weights 2:1) on each path. When the genetic correlation between milk and type varies from slightly positive (0.15) to slighly negative (−0.15) (thus covering the range of values most often indicated in the literature) optimum selection intensities and overall genetic gain for the objective function are not dramatically modified. But then, the genetic trend for type varies from a very favorable increase (for a positive correlation) even when no selection on type is performed, to a not negligible decline (for a negative correlation and weights 3:1 or 5:1). These features were also pointed out by Ducrocq (1984).

Conclusion

The previous example clearly demonstates that, at least in certain situations, it is possible to algebraically and numerically develop algorithms to compute optimal selection intensities in relatively complex multistage breeding schemes and with several nonlinear constraints. They are not limited by drastic conditions such as 2–stage selection, uncorrelated traits and/or very simple optimization criteria. Indeed, more complex situations than that presented in the example can be envisaged. For example, the overlap of generations of bulls and cows can be accounted for in the determination of genetic superiorities and generation intervals, extending to the multivariate case the techniques described

in Hill (1974), Elsen and Mocquot (1974, 1976) and Ducrocq and Quaas (1988): dividing each population of candidates into homogeneous cohorts of animals of same age, sex and reproductive role, intracohort selection differentials can be computed under the assumption that selection is performed by retaining all the individuals whose estimated genetic values are above a unique set of truncation thresholds for the n traits or stages considered (Ducrocq, 1984). Then, computation of elements in (18) and their derivatives is tedious but perfectly feasible. However, before undertaking optimization studies of complex schemes, it should be remembered that troublesome problems of convergence may occur such as those found in the application described. In the particular case of the dairy cattle context, convergence problems were also found when we tried to relax the assumption of fixed number of young sires sampled (n_y) or the number of daughters per young sire ($n_{d/y}$), *i.e.*, their repeatability *. Obviously, further research is needed in a widening field.

Finally, the determination of an optimal selection policy, which is "optimal" only from a purely genetic standpoint, should be followed by a sensitivity analysis. It appears that the annual genetic gain ΔG_a varies only slightly over a wide range of values for some parameters. Situations close to the optimum may be of greater interest because they are simple, more practical or less expensive to implement. The knowledge of the true optimum is nevertheless required to assess this "suboptimality".

References

Baas J. (1961) *Cours de Mathématique,* Masson, Paris, vol. II

Bulmer M.G. (1980) *The Mathematical Theory of Quantitative Genetics.* Oxford University Press, Oxford

Cotterill P.P. & James J.W. (1981) Optimizing two-stage independent culling selection in tree and animal breeding. *Theor. Appl. Genet.* 59, 67-72

Dennis J.E. & Schnabel R.B. (1983) *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewood Cliffs

Ducrocq, V. (1984) Conséquences sur le progrès génétique laitier d'une sélection sur des caractères secondaires chez les bovins. *Genet. Sel. Evol.* 16, 467-490

Ducrocq V. & Colleau J.J. (1986) Interest in quantitative genetics of Dutt's and Deak's methods for numerical computation of multivariate normal probability integrals. *Genet. Sel. Evol.* 18, 447-474

Ducrocq V. & Quaas R.L. (1988) Prediction of genetic response to truncation selection across generations. *J. Dairy Sci.* 71, 2543-2553

Dutt J.E. (1973) A representation of multivariate probability integrals by integral transforms. *Biometrika* 60, 637-645

Dutt J.E. (1975) On computing the probability integral of a general multivariate t. *Biometrika* 62, 201-205

Dutt J.E. & Soms A.P. (1976) An integral representation technique for calculating general multivariate probabilities with an application to multivariate χ^2. *Commun. Stat.* 5, 377-388

Elsen J.M. & Mocquot J.C. (1974) Recherches pour une rationalisation technique des schémas de sélection des bovins et ovins. *Bull. Tech. Dept. Genet. Anim., INRA* 17

Elsen J.M. & Mocquot J.C. (1976) Optimisation du renouvellement des femelles dans les troupeaux laitiers soumis au croisement terminal. *Ann. Genet. Sel. Anim.* 8, 343-356

Evans R. (1980) Constrained selection by independent culling levels. *Anim. Prod.* 30, 79-83

---

* In the latter case, derivatives of the objective function with respect to $n_{d/y}$ are extremely complex to derive algebraically and a finite-difference approximation of the derivatives is required.

Hazel L.N. & Lush J.L. (1942) The efficiency of three methods of selection. *J. Hered.* 33, 393-399

Hill W.G. (1974) Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.* 13, 37-50

Jain J.P. & Amble V.N. (1962) Improvement through selection at successive stages. *J. Indian Soc. Agric. Stat.* 14, 88-109

Namkoong G. (1970) Optimum allocation of selection intensity in two stages of truncation selection. *Biometrics* 26, 108-115

Rendel J.M. & Robertson A. (1950) Estimation of genetic gain in milk yield by selection in a close herd of dairy cattle. *J. Genet.* 50, 1-8

Russell N.S., Farrier D.R. & Howell J. (1985) Evaluation of multinormal probabilities using Fourier series expansions. *Appl. Statist.* 34, 49-53

Smith S.P. & Quaas R.L. (1982) Optimal truncation points for independent culling level selection involving two traits. *Biometrics* 38, 975-980

Smith S.P. & Hammond K. (1987) Assortative mating and artificial selection: a second appraisal. *Genet Sel. Evol.* 19, 181-186

Tibau I Font J. & Ollivier L. (1984) La sélection en station chez le porc. *Bull. Techn. Dept. Genet. Anim., INRA,* 37

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

Appendix

Another method for calculating the derivatives of N.

Besides the notations defined in the text, let us define:

$$E_{u\,ij} = E(u \mid x_i = c_i, x_j = c_j) = \frac{c_i \, (r_{ui} - r_{uj} \, r_{ij}) + c_j \, (r_{uj} - r_{ui} \, r_{ij})}{1 - r_{ij}^2}$$

$$R_{uij}^2 = \frac{r_{ui}^2 + r_{uj}^2 - 2 \, r_{ij} \, r_{ui} \, r_{uj}}{1 - r_{ij}^2}$$

R_{uij}^2 = square of the multiple correlation coefficient between u and the predictors x_i, x_j.

By first differentiating N and then using Jain and Amble's formula on the resulting conditional distributions, it can be shown that:

$$\frac{\delta N}{\delta c_i} = -z_i \left(r_{ui} \, c_i \, Q_i + \sqrt{1 - r_{ui}^2} \sum_{j \neq i} r_{uj \cdot i} \, z_{j \cdot i} \, Q_{ij} \right)$$

$$\frac{\delta^2 N}{\delta c_i \delta c_j} = \frac{z_i \, z_{j \cdot i}}{\sqrt{1 - r_{ij}^2}} \left(E_{u \cdot ij} \, Q_{ij} + \sqrt{1 - R_{uij}^2} \sum_{k \neq i \neq j} r_{uk \cdot ij} \, z_{k \cdot ij} \, Q_{ijk} \right) \quad i \neq j$$

$$\frac{\delta^2 N}{\delta c_i^2} = -c_i \, \frac{\delta N}{\delta c_i} - z_i \, r_{ui} \, Q_i - \sum_{j \neq i} r_{ij} \, \frac{\delta^2 N}{\delta c_i \delta c_j}$$

This approach was shown to be strictly equivalent both theoretically and computationally to that developed in the text (the same values are found at each step of Newton's algorithm).