Theor Appl Genet (1988) 75:553–560

Linear selection indices for non-linear profit functions

Y. Itoh and Y. Yamada

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

Received May 20, 1987; Accepted September 3, 1987
Communicated by J. S. F. Barker

Summary. Two methods of deriving linear selection indices for non-linear profit functions have been proposed. One is by linear approximation of profit, and another is the graphical method of Moav and Hill (1966). When profit is defined as the function of population means, the graphical method is optimal. In this paper, profit is defined as the function of the phenotypic values of individual animals; it is then shown that the graphical method is not generally optimal. We propose new methods for constructing selection indices. First, a numerical method equivalent to the graphical method is proposed. Furthermore, we propose two other methods using quadratic approximation of profit: one is based on Taylor series about means before selection, and the other is based on Tayler series about means after selection. Among these different methods, it is shown that the method using quadratic approximation based on Taylor series about means after selection is the most efficient.

Key words: Selection index – Non-linear profit – Newton-Raphson method – Taylor series

Introduction

Selection is practiced to improve a certain quantity related to the genetic properties of a population. Using the terminology of Goddard (1983), the quantity is called "profit". Selection index theory was originally formulated based on a profit function defined as a linear function of traits. However, in some cases, profit may be better characterized by a non-linear function of traits. Articles discussing selection for efficiency of animal production have been published (Elsen et al. 1986), and many of them consider non-linear profit functions. Proposed methods of constructing selection indices to improve non-linear profit may be classified into two groups: a group of linear selection indices and a group of non-linear ones. The latter group includes one of the indices of Wilton et al. (1968), one of the indices of Harris (1970), the indices of Rönningen (1971), and the indices of Van Vleck (1983). Goddard (1983), however, pointed out that if component traits are inherited additively, genetic progress of profit based on linear selection indices is always greater than that based on non-linear ones. Thus, if the optimal selection index for a certain non-linear profit function exists, it should exist among linear selection indices. This paper, therefore, considers only linear selection indices.

Two different methods for deriving linear selection indices for non-linear profit functions have been proposed. The first is to approximate the profit function by a linear function, and use the partial derivatives of the profit function evaluated at the means before selection as economic weights in conventional method of constructing an index. In this method, the economic weight vector is expressed as

$$\boldsymbol{t} = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}_{\boldsymbol{x}=\boldsymbol{\mu}} \tag{1}$$

where f is a profit function, and $\boldsymbol{x}$ and $\boldsymbol{\mu}$ are phenotypic and mean vectors of component traits, respectively. Using (1), the index weight vector may be expressed as

$$\boldsymbol{b}_0 = \boldsymbol{P}^{-1} \boldsymbol{G} \boldsymbol{t} \tag{2}$$

where P and G are phenotypic and genotypic variance covariance matrices, respectively. This method was introduced by Moav and Hill (1966), Harris (1970), Melton et al. (1979), James (1982) and Brascamp et al. (1985). The defect of this method, however, is that the error caused by linear approximation of profit is not always negligible, and it becomes greater as deviations from the means of the component traits increase. Hence, the efficiency of the index decreases as the variance of the component traits and the genetic gains increase.

The second method is the graphical method proposed by Moav and Hill (1966). [Originally this method was not intended as a formal way of finding the index weights, rather as a way of illustrating the principles. Later, Goddard (1983) discussed it as a way of finding the index weights.] It derives an index from a diagram assuming two traits whose values have been transformed to have equal variances, equal heritabilities, and no correlation. Under a given selection intensity, the resulting linear selection index gives a response curve that forms a circle as index weights vary. If the profit contours are plotted on the same diagram, then the point on the response curve with the maximum profit can be located, and the index weights correspondingly deduced. The major differences of this method from the previous one are that this method evaluates the means after selection, and that the index weights vary as the selection intensity varies. However, this method also is defective, because it is not applicable to profit involving more than two traits and it is difficult to draw non-linear profit contours precisely.

Non-linear profit functions could be defined in two ways: as the functions of phenotypic values of individual animals, or as the functions of population means. Goddard (1983) assumed the latter type of profit functions, and showed that the graphical method of Moav and Hill (1966) is optimal. However, if the former type of profit functions are assumed, the graphical method is no longer optimal: it maximizes the profit function of the means of component traits, instead of maximizing the mean of the profit which is what is required, as Elsen et al. (1986) pointed out. To optimize the profit of the whole population, we should maximize the mean of the profit.

In the present paper, we restrict ourselves to the profit functions on an individual basis and propose new methods. The first one is a numerical method equivalent to the graphical method of Moav and Hill (1966). This method is not optimal in general, but it gives approximate solution close to optimal. With a computer, it is conveniently applicable to profit functions of more than two traits. Other methods we propose are based on a quadratic approximation of profit. They are not optimal either, but we found that one of them is always more efficient than those which had been previously proposed.

1 A numerical method equivalent to the graphical method of Moav and Hill (1966)

1.1 Notations and assumptions

We use the following notations to describe the method: n the number of component traits; $\boldsymbol{\mu}$ a vector of population means before selection; $\boldsymbol{d}$ a vector of expected selection responses; $\boldsymbol{x}$ a vector of phenotypic values of progeny after selection with expectation $\mathrm{E}(\boldsymbol{x}) = \boldsymbol{\mu} + \boldsymbol{d}$; $f(\boldsymbol{x})$ non-linear profit function; i selection intensity; $\boldsymbol{P}$ phenotypic variance covariance matrix; $\boldsymbol{G}$ genetic variance covariance matrix; $\boldsymbol{b}$ a vector of index weights; σ_I^2 variance of the index, i.e. $\sigma_I^2 = \boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}$.

We assume that all components traits are inherited additively. Further, we assume that matrices $\boldsymbol{P}$ and $\boldsymbol{G}$ are known and not changed by selection. Although these assumptions may be unrealistic, they approximate the reality sufficiently, can simplify the theory to a great degree, and will not pose a serious problem in application. We will consider only the case where the traits included in the index are the same as those included in the profit function.

1.2 General formulation

The graphical method of Moav and Hill (1966) finds the means with the maximum profit among all attainable means after selection under a given selection intensity. Under a given selection intensity i, as the direction of selection varies, expected selection responses $\boldsymbol{d}$ forms an ellipsoid defined by

$$\boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d} = i^2 \tag{3}$$

which can be obtained by substituting $\boldsymbol{b} = (\sigma_I/i)\,\boldsymbol{G}^{-1}\boldsymbol{d}$ into $\sigma_I^2 = \boldsymbol{b}'\boldsymbol{P}\boldsymbol{b}$. Now our problem is to find $\boldsymbol{d}$ which maximizes

$$f(\mathrm{E}(\boldsymbol{x})) = f(\boldsymbol{\mu} + \boldsymbol{d})$$

among all $\boldsymbol{d}$'s which satisfy (3). We can get such a vector $\boldsymbol{d}$ by solving the following non-linear equation system with respect to $\boldsymbol{d}$.

$$\begin{cases} \boldsymbol{h}(\boldsymbol{d}) + 2\lambda\,\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d} = 0 \\ \boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d} - i^2 = 0 \end{cases} \tag{4}$$

where λ represents a Lagrange multiplier and

$$\boldsymbol{h}(\boldsymbol{d}) = \begin{bmatrix} \dfrac{\partial f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_1} \\ \dfrac{\partial f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_2} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_n} \end{bmatrix} \tag{5}$$

If the profit function is linear and its economic weight vector is $\boldsymbol{a}$, we can solve (4) directly. The solution is $\boldsymbol{d} = i\,\boldsymbol{GP}^{-1}\boldsymbol{G}\,\boldsymbol{a}(\boldsymbol{a}'\boldsymbol{GP}^{-1}\boldsymbol{G}\,\boldsymbol{a})^{-1/2}$, which agrees with the expected selection response based on the conventional selection index. In general, if $f(\boldsymbol{\mu}+\boldsymbol{d})$ is non-linear, we cannot solve (4) directly, but we can solve it iteratively by the Newton-Raphson method. Then, between iterations, a vector of corrections of $\boldsymbol{d}$, $\Delta\boldsymbol{d}$, is obtained as a solution of

$$\begin{bmatrix} \boldsymbol{H}(\boldsymbol{d}) + 2\lambda\,\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1} & 2\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d} \\ 2\,\boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1} & 0 \end{bmatrix}\begin{bmatrix}\Delta\boldsymbol{d}\\ \lambda\end{bmatrix} = -\begin{bmatrix}\boldsymbol{h}(\boldsymbol{d})\\ \boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d} - i^2\end{bmatrix} \tag{6}$$

where

$$\boldsymbol{H}(\boldsymbol{d}) = \begin{bmatrix} \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_1^2} & \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_1\,\partial d_2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_1\,\partial d_1} \\ \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_2\,\partial d_1} & \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_2^2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_2\,\partial d_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_n\,\partial d_1} & \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_n\,\partial d_2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{\mu}+\boldsymbol{d})}{\partial d_n^2} \end{bmatrix} \tag{7}$$

After obtaining $\boldsymbol{d}$ by iteration, we can get index weights by

$$\boldsymbol{b} = \boldsymbol{G}^{-1}\boldsymbol{d} \tag{8}$$

using the result of Pešek and Baker (1969) or Yamada et al. (1975). These index weights give an index which has standard deviation equal to the selection intensity.

Being iterative, this method needs a proper initial value of $\boldsymbol{d}$. If an inappropriate initial value is used, there is a risk of obtaining an inappropriate solution which does not maximize $f(\mathrm{E}(\boldsymbol{x}))$. The approximate method based on (1) gives values close to the optimal solution, therefore it can be used as the initial value. It is expressed as

$$\boldsymbol{d}_0 = i\,\boldsymbol{G}\boldsymbol{b}_0(\boldsymbol{b}_0'\boldsymbol{P}\boldsymbol{b}_0)^{-1/2} \tag{9}$$

where $\boldsymbol{b}_0$ is given by (2). If $\Delta\boldsymbol{d}$ in (6) is assumed to be negligible, we can obtain the initial value of λ as

$$2\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d}_0\lambda + \boldsymbol{e} = \boldsymbol{h}(\boldsymbol{d}) \tag{10}$$

from the first equation of (6), where $\boldsymbol{e}$ is a vector of errors caused by assuming $\Delta\boldsymbol{d} = 0$. Applying the least squares method to (10), λ becomes

$$\lambda_0 = -\tfrac{1}{2}\,[(\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1})(\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d}_0)]^{-1} \cdot (\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1})\,\boldsymbol{h}(\boldsymbol{d}_0). \tag{11}$$

It is sufficient to use this λ_0 as the initial value of λ.

This method gives solutions equivalent to that of the graphical method of Moav and Hill (1966), and has the advantage of being applicable to profit functions involving more than two traits and being conveniently computed with a computer. However, this method maximizes $f(\mathrm{E}(\boldsymbol{x}))$ instead of $\mathrm{E}(f(\boldsymbol{x}))$, and therefore does not give the optimal solution. Nevertheless, it gives an approximate solution close to optimal and gives the exact optimal solution when the profit is quadratic, as shall be described in the next subsection.

1.3 The optimal index for quadratic profit

Now we assume that a profit function is quadratic and defined by

$$f(\boldsymbol{x}) = \boldsymbol{a}'\boldsymbol{x} + \boldsymbol{x}'\boldsymbol{A}\,\boldsymbol{x} \tag{12}$$

as did Wilton et al. (1968), where $\boldsymbol{A}$ is a symmetric constant matrix and $\boldsymbol{a}$ is a constant vector. Then the expectation of the profit after selection becomes

$$\mathrm{E}(f(\boldsymbol{x})) = \boldsymbol{a}'\,\mathrm{E}(\boldsymbol{x}) + \mathrm{E}(\boldsymbol{x}')\,\boldsymbol{A}\,\mathrm{E}(\boldsymbol{x}) + \mathrm{tr}(\boldsymbol{AP}) = f(\mathrm{E}(\boldsymbol{x})) + \mathrm{tr}(\boldsymbol{AP})$$

using the result for expectations of quadratic forms (Searle 1971). If we assume that $\boldsymbol{P}$ is not changed by selection, the difference between $\mathrm{E}(f(\boldsymbol{x}))$ and $f(\mathrm{E}(\boldsymbol{x}))$ is constant, so maximization of $f(\mathrm{E}(\boldsymbol{x}))$ becomes equivalent to maximization of $\mathrm{E}(f(\boldsymbol{x}))$. Hence, the method described in the previous subsection is optimal when the profit is quadratic.

From (12), we get

$$f(\boldsymbol{\mu}+\boldsymbol{d}) = \boldsymbol{d}'(2\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{a}) + \boldsymbol{d}'\boldsymbol{A}\,\boldsymbol{d} + \text{Constant}.$$

Then we get $\boldsymbol{H}(\boldsymbol{d}) = 2\boldsymbol{A}$ and $\boldsymbol{h}(\boldsymbol{d}) = 2\boldsymbol{A}\,\boldsymbol{d} + 2\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{a}$, so (6) becomes

$$2\begin{bmatrix}\boldsymbol{A} + \lambda\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1} & \boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d}\\ \boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1} & 0\end{bmatrix}\begin{bmatrix}\Delta\boldsymbol{d}\\ \lambda\end{bmatrix} = -\begin{bmatrix}2\boldsymbol{A}\,\boldsymbol{d} + 2\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{a}\\ \boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d} - i^2\end{bmatrix}. \tag{13}$$

After iteration, index weights can be obtained by $\boldsymbol{b} = \boldsymbol{G}^{-1}\boldsymbol{d}$. From (9) and (11), the initial values for $\boldsymbol{d}$ and λ may be expressed as

$$\boldsymbol{d}_0 = i\,\boldsymbol{G}\,\boldsymbol{b}_0(\boldsymbol{b}_0'\,\boldsymbol{P}\boldsymbol{b}_0)^{-1/2},$$

$$\lambda_0 = -\tfrac{1}{2}\,[(\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1})(\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1}\boldsymbol{d}_0)]^{-1} \cdot (\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{PG}^{-1})(2\boldsymbol{A}\,\boldsymbol{d}_0 + 2\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{a})$$

where

$$\boldsymbol{b}_0 = \boldsymbol{P}^{-1}\boldsymbol{G}(2\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{a}) \tag{14}$$

which is identical to the index weights of Wilton et al. (1968), as pointed out by Goddard (1968).

Furthermore, by substituting $\boldsymbol{d} = \boldsymbol{G}\boldsymbol{b}$ and $\Delta\boldsymbol{d} = \boldsymbol{G}\,\Delta\boldsymbol{b}$ into (13) and modifying it, we finally get

$$2\begin{bmatrix}\boldsymbol{GAG} + \lambda\boldsymbol{P} & \boldsymbol{P}\boldsymbol{b}\\ \boldsymbol{b}'\boldsymbol{P} & 0\end{bmatrix}\begin{bmatrix}\Delta\boldsymbol{b}\\ \lambda\end{bmatrix} = -\begin{bmatrix}2\boldsymbol{GAGb} + 2\boldsymbol{G}\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{G}\,\boldsymbol{a}\\ \boldsymbol{b}'\boldsymbol{P}\boldsymbol{b} - i^2\end{bmatrix}. \tag{15}$$

Table 1. Index weights and desired selection responses for the quadratic profit under various selection intensities

Selection intensity	Index weights			Desired selection responses		
	b_1	b_2	b_2/b_1	d_1	d_2	d_2/d_1
0.001	0.00001752	0.0001908	10.89	0.02681	0.0003398	0.01268
0.01	0.0001751	0.001909	10.90	0.2681	0.003399	0.01268
0.1	0.001750	0.01920	10.97	2.679	0.03411	0.01273
0.5	0.008718	0.09842	11.29	13.37	0.1730	0.01294
1	0.01736	0.2027	11.68	26.66	0.3520	0.01320
2	0.03441	0.4277	12.43	53.04	0.7268	0.01370
3	0.05116	0.6730	13.15	79.13	1.1222	0.01418
4	0.06765	0.9367	13.85	104.96	1.5361	0.01463
5	0.08387	1.2169	14.51	130.54	1.9668	0.01507

If we use this equation, we can compute the optimum $\boldsymbol{b}$ without computing $\boldsymbol{d}$. The index weights given by using (15) also give an index with standard deviation equal to the selection intensity. In (15), $\boldsymbol{b}_0$ of (14) can be used as the initial value of $\boldsymbol{b}$, but the scale of $\boldsymbol{b}_0$ has to be changed to make the standard deviation of the index equal to the selection intensity; so the initial value of $\boldsymbol{b}$ is given by

$$\boldsymbol{b}_* = i\,\boldsymbol{b}_0(\boldsymbol{b}_0'\,\boldsymbol{P}\,\boldsymbol{b}_0)^{-1/2}. \tag{16}$$

The initial value of λ is given by

$$\lambda_0 = -\tfrac{1}{2}(\boldsymbol{b}_*'\,\boldsymbol{P}\boldsymbol{P}\,\boldsymbol{b}_*)^{-1}\,\boldsymbol{b}_*'\,\boldsymbol{P}(2\boldsymbol{G}\boldsymbol{A}\boldsymbol{G}\,\boldsymbol{b}_* + 2\boldsymbol{G}\boldsymbol{A}\,\boldsymbol{\mu} + \boldsymbol{G}\,\boldsymbol{a}).$$

1.4 A numerical example for a quadratic profit

We will use the same example as Wilton et al. (1968):

$$f(\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{A}\,\boldsymbol{x} + \boldsymbol{a}'\,\boldsymbol{x},$$

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0.00245 \\ 0.00245 & 0 \end{bmatrix},\quad \boldsymbol{a} = \begin{bmatrix} 0.0950 \\ 1.0354 \end{bmatrix},\quad \boldsymbol{\mu} = \begin{bmatrix} 418.95 \\ 13.35 \end{bmatrix},$$

$$\boldsymbol{G} = \begin{bmatrix} 1452.00 & 7.20 \\ 7.20 & 1.12 \end{bmatrix},\quad \boldsymbol{P} = \begin{bmatrix} 2649.00 & 18.49 \\ 18.49 & 1.75 \end{bmatrix},$$

Index weights and desired selection responses for various selection intensities are given in Table 1. Convergences of solutions were very fast: for example, under selection intensity of unity, corrections became less than 10^{-10} at the 4th iteration using either (13) or (15). Relative index weights (b_2/b_1) and desired direction of improvement (d_2/d_1) change as the selection intensity varies. When the selection intensity approaches 0, b_2/b_1 and d_2/d_1 approach 10.89 and 0.01268, respectively, which are the values based on (14).

2 Numerical methods based on quadratic approximation

In general, if a profit function is not quadratic, the difference between $\mathrm{E}(f(\boldsymbol{x}))$ and $f(\mathrm{E}(\boldsymbol{x}))$ is not constant, and the method described in Section 1 is not optimal. While we have no idea of how to get the optimal index, we can consider other approximate methods.

2.1 Quadratic approximation based on Taylor series about means before selection

Harris (1970) considered (1) as a linear approximation of profit based on Taylor series, and approximated the profit by

$$f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}) + \boldsymbol{t}'(\boldsymbol{x} - \boldsymbol{\mu}) \tag{17}$$

where $\boldsymbol{t}$ is a constant vector defined by (1). However, this approximation is not precise enough. An alternative way is to approximate the profit by a quadratic function based on Taylor series, i.e.,

$$f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}) + \boldsymbol{t}'(\boldsymbol{x} - \boldsymbol{\mu}) + \tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})'\,\boldsymbol{T}(\boldsymbol{x} - \boldsymbol{\mu}) \tag{18}$$

where $\boldsymbol{T}$ is a constant matrix defined as

$$\boldsymbol{T} = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1\,\partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2\,\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x_n\,\partial x_1} & \dfrac{\partial^2 f}{\partial x_n\,\partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}_{\boldsymbol{x}=\boldsymbol{\mu}}$$

Generally, this approximation should be more precise than that based on (17). Now we consider a selection index which maximizes the expectation of profit approximated by (18). Since (18) is quadratic, the method described in 1.3 may be used to derive the selection index. The expectation of (18) becomes

$$\begin{aligned} \mathrm{E}(f(\boldsymbol{x})) &\fallingdotseq f(\boldsymbol{\mu}) + \boldsymbol{t}'\,\boldsymbol{d} + \tfrac{1}{2}\,\boldsymbol{d}'\,\boldsymbol{T}\boldsymbol{d} + \tfrac{1}{2}\,\mathrm{tr}\,(\boldsymbol{T}\boldsymbol{P}) \\ &= \boldsymbol{t}'\,\boldsymbol{d} + \tfrac{1}{2}\,\boldsymbol{d}'\,\boldsymbol{T}\boldsymbol{d} + \text{Constant}. \end{aligned}$$

Then the desired selection responses $\boldsymbol{d}$ can be obtained iteratively from

$$\begin{bmatrix} \boldsymbol{T}+2\lambda \boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1} & 2\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d} \\ 2\boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1} & 0 \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{d} \\ \lambda \end{bmatrix} = -\begin{bmatrix} \boldsymbol{T}\boldsymbol{d}+\boldsymbol{t} \\ \boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d}-i^2 \end{bmatrix}. \qquad (19)$$

The initial value of λ is given by

$$\lambda_0 = -\tfrac{1}{2}[(\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1})(\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d}_0)]^{-1} \cdot (\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1})(\boldsymbol{T}\boldsymbol{d}_0+\boldsymbol{t})$$

where $\boldsymbol{d}_0$ is the initial value of $\boldsymbol{d}$ given by (9). After iteration, index weights can be obtained by $\boldsymbol{b} = \boldsymbol{G}^{-1}\boldsymbol{d}$. Vector $\boldsymbol{b}$ can also be obtained directly from iteration if equation (19) is modified by substituting $\boldsymbol{d} = \boldsymbol{G}\boldsymbol{b}$ and $\Delta\boldsymbol{d} = \Delta\boldsymbol{G}\boldsymbol{b}$ as in 1.3, which we will not describe in detail here.

2.2 Quadratic approximation based on Taylor series about means after selection

The defect of the method described in the previous subsection is that approximation by (18) is not precise enough when $\boldsymbol{d}$ is large. To overcome this, we can consider quadratic approximation based on Taylor series about means after selection instead of that before selection. It is expressed as

$$f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}+\boldsymbol{d}) + \boldsymbol{h}(\boldsymbol{d})'(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d}) + \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d})'\boldsymbol{H}(\boldsymbol{d})(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d}), \qquad (20)$$

where $\boldsymbol{h}(\boldsymbol{d})$ and $\boldsymbol{H}(\boldsymbol{d})$ are given by (5) and (7), respectively. The expectation of (20) becomes

$$\mathrm{E}(f(\boldsymbol{x})) \fallingdotseq f(\boldsymbol{\mu}+\boldsymbol{d}) + \tfrac{1}{2}\,\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d})).$$

As in (6), we get the following equation from which we can get the desired $\boldsymbol{d}$ iteratively:

$$\begin{bmatrix} \boldsymbol{H}(\boldsymbol{d})+\boldsymbol{W}(\boldsymbol{d})+2\lambda \boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1} & 2\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d} \\ 2\boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1} & 0 \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{d} \\ \lambda \end{bmatrix} = -\begin{bmatrix} \boldsymbol{h}(\boldsymbol{d})+\boldsymbol{w}(\boldsymbol{d}) \\ \boldsymbol{d}'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d}-i^2 \end{bmatrix} \qquad (21)$$

where

$$\boldsymbol{w}(\boldsymbol{d}) = \frac{1}{2}\begin{bmatrix} \dfrac{\partial\,\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_1} \\ \dfrac{\partial\,\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_2} \\ \vdots \\ \dfrac{\partial\,\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_n} \end{bmatrix}$$

and

$$\boldsymbol{W}(\boldsymbol{d}) = \frac{1}{2}\begin{bmatrix} \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_1^2} & \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_1\partial d_2} & \cdots & \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_1\partial d_n} \\ \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_2\partial d_1} & \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_2^2} & \cdots & \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_2\partial d_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_n\partial d_1} & \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_n\partial d_2} & \cdots & \dfrac{\partial^2\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}(\boldsymbol{d}))}{\partial d_n^2} \end{bmatrix}$$

The initial value of λ used in (21) is given by

$$\lambda_0 = -\tfrac{1}{2}[(\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1})(\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1}\boldsymbol{d}_0)]^{-1} \cdot (\boldsymbol{d}_0'\boldsymbol{G}^{-1}\boldsymbol{P}\boldsymbol{G}^{-1})(\boldsymbol{h}(\boldsymbol{d}_0)+\boldsymbol{w}(\boldsymbol{d}_0))$$

where $\boldsymbol{d}_0$ is the initial value of $\boldsymbol{d}$ which is given by (9).

3 Comparison of the four methods

The methods of computing linear selection indices for non-linear profit functions are summarized as follows:
Method 1: Deriving economic weights using partial derivatives evaluated at means before selection. This is also considered as the method maximizing an expectation of a linear approximate function of profit based on Taylor series about means before selection.
Method 2: Numerical evaluation equivalent to the graphical method of Moav and Hill (1966).
Method 3: Maximizing an expectation of a quadratic approximate function based on Taylor series about means before selection.
Method 4: Maximizing an expectation of a quadratic approximate function based on Taylor series about means after selection.

A primary difference among these methods lies in the approximation of the profit function. The approximate functions on which these methods are based are summarized as follows:
Method 1: $f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}) + \boldsymbol{t}'(\boldsymbol{x}-\boldsymbol{\mu})$
Method 2: $f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}+\boldsymbol{d})$
Method 3: $f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}) + \boldsymbol{t}'(\boldsymbol{x}-\boldsymbol{\mu}) + \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{T}(\boldsymbol{x}-\boldsymbol{\mu})$
Method 4: $f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}+\boldsymbol{d}) + \boldsymbol{h}(\boldsymbol{d})'(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d}) + \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d})'\boldsymbol{H}(\boldsymbol{d})(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d})$.

Except for *Method 2,* each method maximizes expectation of its approximate function of profit. However, because $\mathrm{E}(f(\boldsymbol{\mu}+\boldsymbol{d}) + \boldsymbol{h}(\boldsymbol{d})'(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d})) = f(\boldsymbol{\mu}+\boldsymbol{d})$, *Method 2* also can be considered as the method that maximizes the expectation of the approximate function which is expressed as

$$f(\boldsymbol{x}) \fallingdotseq f(\boldsymbol{\mu}+\boldsymbol{d}) + \boldsymbol{h}(\boldsymbol{d})'(\boldsymbol{x}-\boldsymbol{\mu}-\boldsymbol{d}).$$

This function can be considered as a linear approximation based on Taylor series about means after selection.

Therefore, all four methods are considered as those which maximize the expectations of the approximate functions of the profit based on Taylor series.

These four methods may be classified into two groups, depending on whether the approximate function is linear or quadratic, and depending on whether Taylor series is expanded about $\boldsymbol{\mu}$ or about $\boldsymbol{\mu} + \boldsymbol{d}$. These classifications are summarized in Table 2.

In general, a quadratic approximation is more precise than a linear one, so we can say that *Method 3* is more efficient than *Method 1*, and *Method 4* is more efficient than *Method 2*. On the other hand, approximation by Taylor series is more precise expanded about means after selection than before selection, so we can say that *Method 2* is more efficient than *Method 1*, and *Method 4* is more efficient than *Method 3*. From these facts, we conclude that *Method 4* is always the most efficient and *Method 1* is the least efficient. It is not clear which is more efficient between *Method 2* and *Method 3*; it might depend on circumstances.

When non-linear profit functions are defined on a population mean basis, the graphical method of Moav and Hill (1966) (which is equivalent to *Method 2*) is optimal, as shown by Goddard (1983). However, when non-linear profit functions are defined on an individual animal basis, this is not true except for quadratic profit, and more efficient methods could exist.

We have only considered the situation where the traits included in the selection index are the same as those included in the profit function. When not all traits are included in the index, the four methods need some modifications which are described in the Appendix.

4 A numerical illustration of the methods for non-linear profit

We consider the following hypothetical problem to illustrate the methods. Let

$$f(\boldsymbol{x}) = \frac{x_1}{x_2}, \quad \boldsymbol{\mu} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{G} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

where profit function $f(\boldsymbol{x})$ represents the efficiency of animal production, and x_1 and x_2 are traits associated with return and cost of production, respectively; they are assumed to be uncorrelated for simplicity. Vector and matrices required to compute the indices are given in the following:

$$\boldsymbol{t} = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix}, \quad \boldsymbol{T} = \begin{bmatrix} 0 & -0.01 \\ -0.01 & 0.02 \end{bmatrix},$$

$$\boldsymbol{h}(\boldsymbol{d}) = \begin{bmatrix} \dfrac{1}{10 + d_2} \\[2ex] -\dfrac{10 + d_1}{(10 + d_2)^2} \end{bmatrix},$$

$$\boldsymbol{H}(\boldsymbol{d}) = \begin{bmatrix} 0 & -\dfrac{1}{(10 + d_2)^2} \\[2ex] -\dfrac{1}{(10 + d_2)^2} & \dfrac{2(10 + d_1)}{(10 + d_2)^3} \end{bmatrix},$$

$$\operatorname{tr}(\boldsymbol{PH}(\boldsymbol{d})) = \frac{2(10 + d_1)}{(10 + d_2)^3}, \quad \boldsymbol{w}(\boldsymbol{d}) = \begin{bmatrix} \dfrac{1}{(10 + d_2)^3} \\[2ex] -\dfrac{3(10 + d_1)}{(10 + d_2)^4} \end{bmatrix},$$

$$\boldsymbol{W}(\boldsymbol{d}) = \begin{bmatrix} 0 & -\dfrac{3}{(10 + d_2)^4} \\[2ex] -\dfrac{3}{(10 + d_2)^4} & \dfrac{12(10 + d_1)}{(10 + d_2)^5} \end{bmatrix}.$$

Desired selection responses and increases of mean profit based on the four methods for various selection intensities are given in Table 3. In all the methods (2, 3 and 4), convergences of the desired selection responses were very fast: for example, under selection intensity of unity, corrections became less than 10^{-10} at the 4th iteration in all methods. The increases of mean profit were computed by numerical integration assuming bivariate normal distribution. We assumed that variances and covarinces are invariant to selection. Differences in increases of mean profit among methods are generally not very large: they are very small when the selection intensity is small, but increase as the selection intensity becomes greater. For all selection intensities, *Method 4* is the best and *Method 1* is the worst, as was expected. The second is *Method 2* and the third is *Method 3*, but their differences are very small when the selection intensity is small.

In this example, the advantage of *Method 4* over the other methods, especially *Method 2*, was small. However, this may not always be true. This example is hypothetical and used only for illustrating, hence it is very simple and involves only two traits. In practical application, the profit function may be more complicated and involve many traits, and there would be a possibility of greater differences among methods.

Table 2. Classification of the four methods depending on how to approximate the profit

Is the approximate function linear or quadratic?	Is Taylor series expanded about $\boldsymbol{\mu}$ or $\boldsymbol{\mu} + \boldsymbol{d}$?	
	$\boldsymbol{\mu}$	$\boldsymbol{\mu} + \boldsymbol{d}$
Linear	Method 1	Method 2
Quadratic	Method 3	Method 4

Table 3. Comparison of the four methods for various selection intensities

Selection intensity	Method	Desired selection responses		Mean profit after selection	Increase of mean profit
		d_1	d_2		
0.01	1	0.0035355	−0.0035355	1.01103835	0.00072219
	2	0.0035343	−0.0035368	1.01103835	0.00072220
	3	0.0035343	−0.0035368	1.01103835	0.00072220
	4	0.0034994	−0.0035713	1.01103839	0.00072224
0.1	1	0.0353553	−0.0353553	1.01756178	0.00724562
	2	0.0352301	−0.0354801	1.01756210	0.00724594
	3	0.0352310	−0.0354792	1.01756209	0.00724594
	4	0.0348782	−0.0358262	1.01756250	0.00724634
0.5	1	0.176777	−0.176777	1.0470802	0.0367640
	2	0.173624	−0.179874	1.0470932	0.0367771
	3	0.173734	−0.179768	1.0470930	0.0367768
	4	0.171789	−0.181627	1.0470954	0.0367792
1	1	0.353553	−0.353553	1.0852300	0.0749139
	2	0.340832	−0.365832	1.0853096	0.0749934
	3	0.341712	−0.365011	1.0853072	0.0749911
	4	0.336966	−0.369397	1.0853145	0.0749983
2	1	0.70711	−0.70711	1.166017	0.155701
	2	0.65534	−0.75534	1.166583	0.156267
	3	0.66228	−0.74925	1.166557	0.156241
	4	0.64675	−0.76270	1.166596	0.156280
3	1	1.06066	−1.06066	1.253406	0.243090
	2	0.94218	−1.16718	1.255341	0.245024
	3	0.96522	−1.14820	1.255213	0.244897
	4	0.92786	−1.17859	1.255365	0.245049
4	1	1.41421	−1.41421	1.348255	0.337939
	2	1.20000	−1.60000	1.353097	0.342780
	3	1.25349	−1.55845	1.352664	0.342347
	4	1.17875	−1.61572	1.353139	0.342823
5	1	1.76777	−1.76777	1.451578	0.441262
	2	1.42743	−2.05243	1.461793	0.451477
	3	1.52959	−1.97746	1.460611	0.450295
	4	1.39780	−2.07271	1.461864	0.451548

Acknowledgements. We wish to thank Dr. W. G. Hill for his critical reading and valuable comments. This study was supported in part by Grants-in-Aid for Scientific Research No. 61304029 from the Ministry of Education, Science and Culture.

References

Brascamp EW, Smith C, Guy DR (1985) Derivation of economic weights from profit equations. Anim Prod 40:175–179

Elsen JM, Bibe B, Landais E, Ricordeau G (1986) Twenty remarks on economic evaluation of selection goals, vol 12. In: Proc 3rd World Congr Genet Appl Livest Prod. Lincoln, Nebraska, pp 321–327

Goddard ME (1983) Selection indices for non-linear profit functions. Theor Appl Genet 64:339–344

Harris DL (1970) Breeding for efficiency in livestock production: defining the economic objectives. J Anim Sci 30:860–865

James JW (1982) Economic aspects of developing breding objectives: general considerations. In: Baker JSF, Hammond K, McClintock AE (eds) Future developments in the genetic improvement of animals. Academic Press, London, pp 107–116

Melton BE, Heady EO, Willham RL (1979) Estimation of economic values for selection indices. Anim Prod 28:279–286

Moav R, Hill WG (1966) Specialized sire and dam lines. IV. Selection within lines. Anim Prod 8:373–390

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

Rönningen K (1971) Selection index for quadratic and cubic models of the aggregate genotype. Meld Nor Landbrukshoegsk 50:1–30

Searle SR (1971) Linear models. Wiley and Sons

Van Vleck LD (1983) Non-linear economic value and restricted selection. In: Notes on the theory and application of selection principles for the genetic improvement of animals. Cornell Univ 14:171–180

Wilton JW, Evans DA, Van Vleck LD (1968) Selection indices for quadratic models of total merit. Biometrics 24:937–949

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

Appendix

In the text, it was assumed that the traits included in the profit function were the same as those included in the selection index. However, when traits are difficult and expensive to measure, they will be excluded from the selection index in spite of their economic importance. Thus, not all traits in the profit function can be included in the index. In such cases, the four methods described in the text cannot be used to construct the index directly, and some modifications are necessary.

Now traits are divided into two groups: one is the group of m traits included in the index and another is the group of $n-m$ traits not included in the index. According to this division, the vectors and the matrices defined in 1.1 are partitioned as follows:

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \boldsymbol{d} = \begin{bmatrix} \boldsymbol{d}_1 \\ \boldsymbol{d}_2 \end{bmatrix}, \quad \boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ \boldsymbol{P}'_{12} & \boldsymbol{P}_{22} \end{bmatrix} \quad \text{and}$$

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_{11} & \boldsymbol{G}_{12} \\ \boldsymbol{G}'_{12} & \boldsymbol{G}_{22} \end{bmatrix}$$

where $\boldsymbol{\mu}_1$ is the mean vector of the m traits included in the index and $\boldsymbol{\mu}_2$ is the mean vector of the $n-m$ traits not included in the index, etc. When the selection based on m traits causes the expected selection responses $\boldsymbol{d}_1$ in these m traits, the expected selection responses in the remaining $n-m$ traits will become

$$\boldsymbol{d}_2 = \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_2. \tag{A.1}$$

Thus, the expected selection responses of all n traits caused by selection on m traits can be expressed as

$$\boldsymbol{d} = \boldsymbol{S}\boldsymbol{d}_1 \quad \text{where} \quad \boldsymbol{S} = \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1} \end{bmatrix}. \tag{A.2}$$

Using (A.1) or (A.2), the four methods for constructing the selection indices can be rewritten.

Method 1

The expectation of a linear approximate profit function for Method 1 can be rewritten as

$$\mathrm{E}(f(\boldsymbol{x})) \fallingdotseq \boldsymbol{t}'\boldsymbol{S}\boldsymbol{d}_1 + \text{Constant.}$$

Then the economic weight vector becomes $\boldsymbol{S}'\boldsymbol{t}$, and the index weight vector of m traits is found to be

$$\boldsymbol{b} = \boldsymbol{P}_{11}^{-1}\boldsymbol{G}_{11}\boldsymbol{S}'\boldsymbol{t} = \boldsymbol{P}_{11}^{-1}[\boldsymbol{G}_{11}|\boldsymbol{G}_{12}]\boldsymbol{t}.$$

Method 2

The expectation of a linear approximate profit function for Method 2 can be rewritten as

$$\mathrm{E}(f(\boldsymbol{x})) = f(\boldsymbol{\mu}_1 + \boldsymbol{d}_1, \boldsymbol{\mu}_2 + \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1).$$

Now our problem is to find $\boldsymbol{d}_1$ which maximizes

$$f(\boldsymbol{\mu}_1 + \boldsymbol{d}_1, \boldsymbol{\mu}_2 + \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1)$$

subject to

$$\boldsymbol{d}'_1\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 = i^2.$$

Such a vector $\boldsymbol{d}_1$ can be obtained iteratively from

$$\begin{bmatrix} \boldsymbol{H}_1(\boldsymbol{d}_1) + 2\lambda\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1} & 2\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 \\ 2\boldsymbol{d}'_1\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1} & 0 \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{d}_1 \\ \lambda \end{bmatrix} = -\begin{bmatrix} \boldsymbol{h}_1(\boldsymbol{d}_1) \\ \boldsymbol{d}'_1\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 - i^2 \end{bmatrix}$$

where

$$\boldsymbol{h}_1(\boldsymbol{d}_1) = \left\{\frac{\partial}{\partial d_{1i}} f(\boldsymbol{\mu}_1 + \boldsymbol{d}_1, \boldsymbol{\mu}_2 + \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1)\right\}$$

$$\boldsymbol{H}_1(\boldsymbol{d}_1) = \left\{\frac{\partial^2}{\partial d_{1i}\partial d_{1j}} f(\boldsymbol{\mu}_1 + \boldsymbol{d}_1, \boldsymbol{\mu}_2 + \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1)\right\}$$

and d_{1i} $(i = 1, \ldots, m)$ is the i-th element of $\boldsymbol{d}_1$. After obtaining $\boldsymbol{d}$ by iteration, we can get the index weight vector from

$$\boldsymbol{b} = \boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1.$$

Method 3

The expectation of a quadratic approximate profit function for Method 3 can be rewritten as

$$\mathrm{E}(f(\boldsymbol{x})) \fallingdotseq \boldsymbol{t}'\boldsymbol{S}\boldsymbol{d}_1 + \tfrac{1}{2}\boldsymbol{d}'_1\boldsymbol{S}'\boldsymbol{T}\boldsymbol{S}\boldsymbol{d}_1 + \text{Constant.}$$

Thus the desirable $\boldsymbol{d}_1$ can be obtained iteratively from

$$\begin{bmatrix} \boldsymbol{S}'\boldsymbol{T}\boldsymbol{S} + 2\lambda\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1} & 2\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 \\ 2\boldsymbol{d}'_2\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1} & 0 \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{d}_1 \\ \lambda \end{bmatrix} = -\begin{bmatrix} \boldsymbol{S}'\boldsymbol{T}\boldsymbol{S}\boldsymbol{d}_1 + \boldsymbol{S}'\boldsymbol{t} \\ \boldsymbol{d}'_1\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_1\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 - i^2 \end{bmatrix}.$$

Method 4

The expectation of a quadratic approximate profit function for Method 4 can be rewritten as

$$\mathrm{E}(f(\boldsymbol{x})) \fallingdotseq f(\boldsymbol{\mu}_1 + \boldsymbol{d}_1, \boldsymbol{\mu}_2 + \boldsymbol{G}'_{12}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1) + \tfrac{1}{2}\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}_1(\boldsymbol{d}_1)).$$

Thus the desirable $\boldsymbol{d}_1$ can be obtained iteratively from

$$\begin{bmatrix} \boldsymbol{H}_1(\boldsymbol{d}_1) + \boldsymbol{W}_1(\boldsymbol{d}_1) + 2\lambda\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1} & 2\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 \\ 2\boldsymbol{d}'_1\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1} & 0 \end{bmatrix} \cdot \begin{bmatrix} \Delta\boldsymbol{d}_1 \\ \lambda \end{bmatrix} = -\begin{bmatrix} \boldsymbol{h}_1(\boldsymbol{d}_1) + \boldsymbol{w}_1(\boldsymbol{d}_1) \\ \boldsymbol{d}'_1\boldsymbol{G}_{11}^{-1}\boldsymbol{P}_{11}\boldsymbol{G}_{11}^{-1}\boldsymbol{d}_1 - i^2 \end{bmatrix}$$

where

$$\boldsymbol{w}_1(\boldsymbol{d}_1) = \left\{\frac{1}{2}\frac{\partial\,\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}_1(\boldsymbol{d}_1))}{\partial d_{1i}}\right\} \quad \text{and}$$

$$\boldsymbol{W}_1(\boldsymbol{d}_1) = \left\{\frac{1}{2}\frac{\partial^2\,\mathrm{tr}(\boldsymbol{P}\boldsymbol{H}_1(\boldsymbol{d}_1))}{\partial d_{1i}\partial d_{1j}}\right\}.$$