37

Theory of Index Selection

With pit bull terriers it might be deemed that the breed should maintain a specific ratio of ugliness to vicious temperament, and a constrained index could incorporate this requirement along with any other breeding objectives deemed appropriate — Gibson and Kennedy (1990)

Version 26 June 2014

While Chapters 28 and 29 present the basic theory for multivariate response, how, in practice, does one perform artificial selection on multiple traits? One of the commonest schemes is to construct some sort of **index**, wherein the investigator assigns (either explicitly or implicitly) a weighting scheme to each trait, creating a univariate character that becomes the target of selection. For example, if $\mathbf{z}$ is the vector of character values measured in an individual, the most common index is a linear combination $\sum b_i z_i = \mathbf{b}^T\mathbf{z}$ and most of our discussion focuses on such **linear indices**. We start with a general review of the theory of selection on a linear index and then cover in great detail the Smith-Hazel index (the index giving the largest expected response in a specified linear combination of characters) and its extensions . We also discuss a number of other indices for different purposes, such as restricted (constraining changes in specified traits) and desired-gains (specifying how the components, rather than the index, will evolve) indices. We conclude our discussion of index selection by considering how to best handle nonlinear indices. We finish the chapter by examining the other approach for selecting on multiple traits, namely choosing traits *sequentially*. Tandem selection, focusing on a single trait each generation (where the focal trait changes over generations) is one such approach, while the other is to select different traits at different times within the life span of single individuals (independent culling and multistage index selection).

There is a huge literature on the theory and application of selection indices. General reviews can be found in Turner and Young (1969), Lin (1978), Namkoong (1979), Bulmer (1980), James (1982), Baker (1986), and Van Vleck (1993), while specific applications to plant breeding (and other organisms with asexual reproduction and/or selfing) can be found in Wricke and Weber (1986), Baker (1986), and Bernardo (2002).

GENERAL THEORY OF SELECTION ON A LINEAR INDEX

Consider selection on the univariate character defined by the linear index

$$I = \sum b_j z_j = \mathbf{b}^T\mathbf{z} \tag{33.1}$$

where $\mathbf{z}$ is the vector of phenotypic values in an individual and $\mathbf{b}$ a vector of weights. Even though it has multivariate *components*, I is just a univariate trait, so all previous results from Chapters 10-15 apply to the index. For example, if we knew its heritability, the breeder's equation predicts its selection response (Chapter 10). Likewise, if we knew its genetic and phenotypic variance, then we can also predict its change in variance (assuming the infinitesimal model).

Genetic Variance, Heritability, and Response of an Index

What are the variances and the heritability associated with an index? Let $\mathbf{P}$ and $\mathbf{G}$ denote the phenotypic and additive-genetic covariance matrices for the vector $\mathbf{z}$ of component traits. From standard results on the variance of a vector (LW Chapter 8), the phenotypic variance of I is just

$$\sigma_I^2 = \sigma(\mathbf{b}^T\mathbf{z}, \mathbf{b}^T\mathbf{z}) = \mathbf{b}^T\sigma(\mathbf{z},\mathbf{z})\mathbf{b} = \mathbf{b}^T\mathbf{P}\mathbf{b} \tag{33.2a}$$

while its additive genetic variance is given by

$$\sigma_{A_I}^2 = \sigma_A(\mathbf{b}^T\mathbf{z}, \mathbf{b}^T\mathbf{z}) = \mathbf{b}^T\sigma_A(\mathbf{z},\mathbf{z})\mathbf{b} = \mathbf{b}^T\mathbf{G}\mathbf{b} \tag{33.2b}$$

Hence, the heritability of I is

$$h_I^2 = \frac{\sigma_{A_I}^2}{\sigma_I^2} = \frac{\mathbf{b}^T\mathbf{G}\mathbf{b}}{\mathbf{b}^T\mathbf{P}\mathbf{b}} \tag{33.2c}$$

as obtained by Lin and Allaire (1977) and Nordskog (1978). If phenotypes $\mathbf{z}$ and breeding values $\mathbf{g}$ are jointly multivariate normal, linear combinations of each is also normally distributed (LW Chapter 8) and hence the univariate breeders' equation holds for response in I. The selection-intensity version of the breeder's equation (Equation 10.6b) gives the expected response in the index as

$$R_I = \overline{\imath}\, h_I^2\, \sigma_I = \overline{\imath} \cdot \frac{\mathbf{b}^T\mathbf{G}\mathbf{b}}{\mathbf{b}^T\mathbf{P}\mathbf{b}} \sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}} = \overline{\imath} \cdot \frac{\mathbf{b}^T\mathbf{G}\mathbf{b}}{\sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}}} \tag{33.3}$$

Example 33.1. A convenient dataset we use through this chapter is that of Brim et al. (1959), who estimated the genetic and phenotypic covariances for several characters in soybeans. Consider three of these traits, z_1 = oil content, z_2 = protein content, and z_3 = yield. For these characters, Brim et al. estimated the covariance matrices as

$$\mathbf{P} = \begin{pmatrix} 287.5 & 477.4 & 1266 \\ 477.4 & 935 & 2303 \\ 1266 & 2303 & 5951 \end{pmatrix}, \qquad \mathbf{G} = \begin{pmatrix} 128.7 & 160.6 & 492.5 \\ 160.6 & 254.6 & 707.7 \\ 492.5 & 707.7 & 2103 \end{pmatrix}$$

Consider two indices, I_1 which equally weights all three traits, and I_2 which assigns four times the weight to yield as the other triats. The resulting vectors of weight are

$$\mathbf{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}$$

The genetic variance for index one is

$$\sigma_A^2(I_1) = \mathbf{b}_1^T\mathbf{G}\mathbf{b}_1 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 128.7 & 160.6 & 492.5 \\ 160.6 & 254.6 & 707.7 \\ 492.5 & 707.7 & 2103 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 5207.9$$

This index has phenotypic variance $\sigma^2(I_1) = \mathbf{b}_1^T\mathbf{P}\mathbf{b}_1 = 15266.3$, giving a heritability of $h^2(I_1) = 5207.9/15266.3 = 0.34$. Likewise, $\sigma^2(I_2) = 125945.3$, $\sigma_A^2(I_2) = 43954.1$, and $h^2(I_2) = 0.35$.

Suppose truncation selection is practiced where we save the upper 5% of the population based on their index scores. Example 10.10 shows that, provided the index values follow a normal distribution in a large population, $\overline{\imath} = 2.06$. Equation 33.3 gives the response for index one as

$$R(I_1) = \overline{\imath} \cdot \frac{\mathbf{b}_1^T \mathbf{G} \mathbf{b}_1}{\sqrt{\mathbf{b}_1^T \mathbf{P} \mathbf{b}_1}} = 2.06 \, \frac{5207.9}{\sqrt{15266.3}} = 86.8$$

Similarly, the response for index 2 is found to be $R(I_2) = 255.1$.

Response in the Individual Components of the Index

How does selection on this index change the vector of underlying character means? Under the conditions of the multivariate breeder's equation, $\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S}$, so our task is to obtain the vector of directional selection differentials $\mathbf{S}$ given selection on I. Consider S_j, the differential associated with character j. First note that the correlation between relative fitness w and the value of character j can be expressed as $\rho_{z_j,w} = \rho_{z_j,I} \cdot \rho_{I,w}$. Expressed in terms of covariances,

$$\frac{\sigma(z_j, w)}{\sigma_w \, \sigma_{z_j}} = \left(\frac{\sigma(z_j, I)}{\sigma_I \, \sigma_{z_j}} \right) \left(\frac{\sigma(I, w)}{\sigma_I \, \sigma_w} \right) \tag{33.4a}$$

Recalling the Price-Robertson identity (Equation 10.7), $\sigma(z_j, w) = S_j$ and likewise $\sigma(I, w) = S_I = \overline{\imath} \, \sigma_I$ where $\overline{\imath}$ is the selection intensity on the index. Solving for $\sigma(z_j, w)$ and using these identities gives

$$S_j = \sigma(z_j, w) = \frac{\sigma(z_j, I) \cdot \sigma(I, w)}{\sigma_I^2} = \overline{\imath} \cdot \frac{\sigma(z_j, I)}{\sigma_I} \tag{33.4b}$$

Finally, note that $\sigma(z_j, I) = \sigma(z_j, \sum_k b_k z_k) = \sum_k b_k P_{jk}$, where P_{ij} is the ij-th element of the phenotypic covariance matrix $\mathbf{P}$. Hence, the jth selection differential is

$$S_j = \left(\frac{\overline{\imath}}{\sigma_I} \right) \cdot \sum_k b_k P_{jk}, \tag{33.4c}$$

giving the vector of selection differential as

$$\mathbf{S} = \left(\frac{\overline{\imath}}{\sigma_I} \right) \cdot \mathbf{P}\mathbf{b} \tag{33.4d}$$

The vector of responses them becomes

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \left(\frac{\overline{\imath}}{\sigma_I} \right) \cdot \mathbf{G}\mathbf{b} = \overline{\imath} \cdot \frac{\mathbf{G}\mathbf{b}}{\sqrt{\mathbf{b}^T \mathbf{P} \mathbf{b}}} \tag{33.5}$$

Equation 33.5 shows that the vector of responses $\mathbf{R}$ in the components of the index is unchanged if the index weights are rescaled from $\mathbf{b}$ to $c \cdot \mathbf{b}$ as the constant c cancels out. However the response in the univariate index I changes as $\mathbf{b}$ is rescaled. From Equation 33.3 the response in the index using weights $c \cdot \mathbf{b}$ is c times the response expected when the index uses $\mathbf{b}$.

Example 33.2. What are the responses in the component traits of index one from Example 33.1? Here

$$\mathbf{Gb} = \begin{pmatrix} 128.7 & 160.6 & 492.5 \\ 160.6 & 254.6 & 707.7 \\ 492.5 & 707.7 & 2103 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 781.8 \\ 1122.9 \\ 3303.3 \end{pmatrix}$$

Applying Equation 33.5 gives the vector of responses in the index components as

$$\mathbf{R} = \frac{\bar{\imath}}{\sqrt{\mathbf{b}^T\mathbf{Pb}}} \cdot \mathbf{Gb} = \frac{2.06}{\sqrt{15266.3}} \begin{pmatrix} 781.8 \\ 1122.9 \\ 3303.3 \end{pmatrix} = \begin{pmatrix} 13.0 \\ 18.7 \\ 55.1 \end{pmatrix}$$

giving the response in I_1 is $1 \cdot 13.0 + 1 \cdot 18.7 + 1 \cdot 55.1 = 86.8$, as found in Example 33.1.

A related problem is the correlated response in the some other index $J = \mathbf{a}^T\mathbf{z} = \sum a_j z_j$ when selection occurs on $I = \mathbf{b}^T\mathbf{z}$. Applying Equation 33.5, the expected correlated response is

$$R_J = \mathbf{a}^T(\boldsymbol{\mu} + \mathbf{R}) - \mathbf{a}^T\boldsymbol{\mu} = \mathbf{a}^T\mathbf{R} = \bar{\imath} \cdot \frac{\mathbf{a}^T\mathbf{Gb}}{\sqrt{\mathbf{b}^T\mathbf{Pb}}} \tag{33.6}$$

The Retrospective Index

While indices have been presented as the objects of selection, often an investigator observes a vector of selection differentials $\mathbf{S}$ or vector of responses $\mathbf{R}$ and wishes to obtain the linear index that would give the same observed $\mathbf{S}$ and/or $\mathbf{R}$. This approach of constructing a **retrospective index** (or **index in retrospect**) was first suggested by Dickerson et al. (1954). If the vector of selection differentials $\mathbf{S}$ is observed, Equation 33.4d suggests the weights for the retrospective index as

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{S} \tag{33.7}$$

(given our previous remarks we ignore the constant $\bar{\imath}/\sigma_I$). Even when artificial selection occurs using a known index, a retrospective index constructed from the *effective* selection differentials (which measure fertility differences in addition to artificial selection, see Equation 10.8) provides the investigator with a measure of how natural selection interferes with the desired selection scheme. Also note that Equation 33.7 is the same as the selection gradient $\boldsymbol{\beta}$ (Chapter 29). Equation 33.7 allows us to move between the vector $\mathbf{b}$ of index weights (the actual amounts of selection) and $\mathbf{S}$, the observed within-generation change in the component trait means.

An important application of Equation 33.7 arises when using multivariate response data to estimated realized genetic covariances (Equation 30.31). While a target index may have been used, estimates of genetic parameters should be based on the weights provided by the retrospective index (obtained given the observed $\mathbf{S}$), rather than relying on the index weights used in the initial selection (Berger and Harvey 1975, Berger 1977, Gunsett et al. 1982). von Butler et al. (1986) used a retrospective index to show that the actual weights on traits in a mouse experiment estimated from the index were rather different than the values initially selected, due to infertility and differences in offspring number among the selected parents.

Alternatively, the investigator may not know the within-generation change in $\mathbf{S}$ but can observe the between-generation change $\mathbf{R}$. In this case Equation 33.5 (again ignoring the constant $\bar{\imath}/\sigma_I$) suggests the retrospective index

$$\mathbf{b} = \mathbf{G}^{-1}\mathbf{R} \tag{33.8a}$$

Also note that Equation 33.8a corresponds to the realized selection gradient (Chapter 30). Finally, there may be interest in $\mathbf{S}$ directly. Multiplying each side of the multivariate breeder's equation $\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S}$ first by $\mathbf{G}^{-1}$ and then $\mathbf{P}$ recovers

$$\mathbf{S} = \mathbf{P}\mathbf{G}^{-1}\mathbf{R} \tag{33.8b}$$

Humphreys (1995) presents an interesting use of a retrospective index looking at trait response in a population of ryegrass subjected to both artificial and natural selection.

The Selection and Response Indices May Contain Different Traits

In most applications of multivariate selection, $\mathbf{G}$ is symmetric and hence square, as our focus is the response for those traits we selected on ($\mathbf{R}$ and $\mathbf{S}$ refer to the same traits). However, in index selection there are often times where the traits we select on and the traits whose response is of interest do not fully overlap (e.g., Example 33.6 below). In such cases, the multivariate breeder's equation still holds, but now with a more general definition of $\mathbf{G}$. Suppose we select on n traits (the $\mathbf{S}$ vector), but are interested in the response of k traits (the $\mathbf{R}$ vector). One example is that $k < n$, so all of the response traits are found in $\mathbf{S}$, but some traits in $\mathbf{S}$ are not seen in $\mathbf{R}$, as we are not interested in their response. Another example is where there are traits in $\mathbf{R}$ that do not appear in $\mathbf{S}$. For both cases, we still have $\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S}$. As before, $\mathbf{P}$ is the $n \times n$ phenotypic covariance matrix for the traits under selection and $\mathbf{S}$ is their vector of selection differentials. Note that the dimensions of $\mathbf{P}^{-1}\mathbf{S}$ are $n \times 1$, while the dimensions of $\mathbf{R}$ is $k \times 1$. Thus, $\mathbf{G}$ must be of dimension $k \times n$ for the matrix products to conform (LW Chapter 8). [As an aside, a quick check to make sure a matrix product is comformable is an excellent safeguard against errors. For example, is there an error in (say) the matrix product $\mathbf{ABC}$? Recalling that the appropriate dimension of the rows and columns must match for multiplication to be defined, we can write this product as $\mathbf{A}_{i\times j}\mathbf{B}_{j\times k}\mathbf{C}_{k\times l}$ which places constrains on the dimensions of $\mathbf{B}$, given those of $\mathbf{A}$ and $\mathbf{C}$.]

Returning to our expanded definition of $\mathbf{G}$ (which has our standard defination as a spcial case when teh same traits are in $\mathbf{R}$ and $\mathbf{S}$), we can write $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$, where the ijth element (row i and column j) of $\mathbf{G}$ is the covariance between the breeding value g_i of the response trait i and the phenotypic value z_j of selection trait j. Thus, $\mathbf{G}$ need not be square. Further, even if $k = n$, if some of the traits in the selection differential and response vector differ, then $\mathbf{G}$ is not symmetric, as $G_{ij} = \sigma(g_i, z_j) \neq G_{ji} = \sigma(g_j, z_i)$ unless the response and differential vectors index exactly the same traits (See Example 33.6). Thus, when $\mathbf{G}$ is not square, $\mathbf{G}^{-1}$ is not defined. Further, when $\mathbf{G}$ is not symmetric, $\mathbf{G} \neq \mathbf{G}^T$, and need to carefully account for when $\mathbf{G}$ is transposed (when it is symmetric, we ignore this distinction).

For the general case of a nonsymmetric $\mathbf{G}$, how do we estimate a retrospective index for $\mathbf{S}$ since $\mathbf{G}^{-1}$ may not exist, and thus Equation 33.8b is not applicable. To find a unique solution, we minimized the selection intensity required for the observed response, which gives

$$\mathbf{S} = \mathbf{G}^T\left(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\right)^{-1}\mathbf{R} \tag{33.9}$$

as obtained by Xu and Muir (1991). As expected, this collapses to $\mathbf{P}\mathbf{G}^{-1}\mathbf{R}$ when $\mathbf{G}$ is symmetric. Example 33.3 gives the derviation. While a bit involved, this introduces the important tool of Lagrange multipliers (Appendix 5) which appear several times in this chaper.

Example 33.3. We follow Xu and Muir's (1991) derivation of the unique estimate of $\mathbf{S}$ when $\mathbf{G}$ is potentially not symmetric. As will be developed shortly (Equation 33.35), the selection intensity $\overline{i}$ on an index can be expressed as

$$\overline{i} = \sqrt{\mathbf{S}\mathbf{P}^{-1}\mathbf{S}}$$

Note for a single trait that this reduces to S/σ_z, as expected. Thus, we wish to solve $\mathbf{GP}^{-1}\mathbf{S} = \mathbf{R}$ subject to the constraint that $\overline{z}$ is minimized. Appendix 5 introduces the machinery for this sort of problem, namely **Lagrange multipliers**, a powerful tool for finding maximum and minimum of functions subjected to constraints. This approach appears throughout this chapter. First, since we need to take derivatives, it will be much easier to work with $\overline{z}^2/2$ in place of $\overline{z}$, as the former is minimized when the later is minimized. The resulting equation to minimize is thus

$$Q = \frac{\mathbf{SP}^{-1}\mathbf{S}}{2} + \boldsymbol{\lambda}(\mathbf{GP}^{-1}\mathbf{S} - \mathbf{R})$$

Where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers (these are constants introduced to solve the constrained equation, but otherwise are of no interest). Note that the second term is zero at the desired solution, so to minimize Q we need to take derivatives with respect to both $\mathbf{S}$ and $\boldsymbol{\lambda}$ and solve for when these are jointly zero. Again calling on results from Appendix 5, we first have from Equation A5.1b (ignoring terms lacking $\boldsymbol{\lambda}$),

$$\nabla_{\boldsymbol{\lambda}}(Q) = \nabla_{\boldsymbol{\lambda}}\left[\,\boldsymbol{\lambda}(\mathbf{GP}^{-1}\mathbf{S} - \mathbf{R})\,\right] = \mathbf{GP}^{-1}\mathbf{S} - \mathbf{R}$$

where $\nabla_{\mathbf{x}}(f)$ denotes the vector of first partial derivations of f with respect to each element in $\mathbf{x}$ (Appendix 5). Likewise, Equation A5.1c gives

$$\nabla_{\mathbf{S}}(Q) = \nabla_{\mathbf{S}}\left[\frac{\mathbf{SP}^{-1}\mathbf{S}}{2} + \boldsymbol{\lambda}(\mathbf{GP}^{-1}\mathbf{S} - \mathbf{R})\right] = \mathbf{P}^{-1}\mathbf{S} + \mathbf{P}^{-1}\mathbf{G}^T\boldsymbol{\lambda}$$

Both of these vectors equal zero at the solution, giving

$$\mathbf{P}^{-1}\mathbf{S} + \mathbf{P}^{-1}\mathbf{G}^T\boldsymbol{\lambda} = \mathbf{0}, \quad \text{and} \quad \mathbf{GP}^{-1}\mathbf{S} = \mathbf{R}$$

Pre-multiplying both side of the first equation by $\mathbf{G}$ yields

$$\mathbf{GP}^{-1}\mathbf{S} + \mathbf{GP}^{-1}\mathbf{G}^T\boldsymbol{\lambda} = \mathbf{R} + \mathbf{GP}^{-1}\mathbf{G}^T\boldsymbol{\lambda} = \mathbf{0}$$

Solving for $\boldsymbol{\lambda}$,

$$\boldsymbol{\lambda} = -\left(\mathbf{GP}^{-1}\mathbf{G}^T\right)^{-1}\mathbf{R}$$

Substituting this solution for $\boldsymbol{\lambda}$ gives

$$\mathbf{P}^{-1}\left(\mathbf{S} - \mathbf{G}^T\left(\mathbf{GP}^{-1}\mathbf{G}^T\right)^{-1}\mathbf{R}\right) = \mathbf{0}$$

or

$$\mathbf{S} = \mathbf{G}^T\left(\mathbf{GP}^{-1}\mathbf{G}^T\right)^{-1}\mathbf{R}$$

Changes in the Additive Variance of I due to Index Selection

Recall from Chapter 13 that directional selection generates negative disequilibrium, reducing the heritability and hence the effectiveness of selection. This, of course, also holds for a selection index I. Artificial selection on an index usually occurs by truncation selection, where the upper p percent are allowed to reproduce. Equation 13.5 gives the resulting phenotypic variance in I following selection as

$$\sigma^2_{I^*} = (1-\kappa)\sigma^2_I, \quad \text{where} \quad \kappa = \overline{z}\left(\overline{z} - z_{[1-p]}\right) \tag{33.10}$$

Here $z_{[1-p]}$ satisfies $\Pr(U \leq z_{[1-p]}) = 1 - p$ (for U a unit normal), and likewise $\overline{\imath}$ can be expressed as a function of p (Equation 10.26a). Starting from an unselected base population, the initial disequilibrium is assumed zero, $d(0) = 0$. The dynamics of $d(t)$ in subsequent generations are given by iterating Equation 13.12,

$$d(t+1) = \frac{d(t)}{2} - \frac{\kappa}{2}\,\frac{[\,\sigma^2_{A_I} + d(t)\,]^2}{\sigma^2_I + d(t)} \tag{33.11a}$$

From Equation 33.3a and 33.3b, the initial additive variance $\sigma^2_{A_I} = \mathbf{b}^T\mathbf{G}\mathbf{b}$, while the initial phenotypic variance is $\sigma^2_I = \mathbf{b}^T\mathbf{P}\mathbf{b}$. Since disequilibrium changes the additive variance, it changes both the heritability and phenotypic variance (Chapter 13), giving the response in generation t as

$$R_I(t) = \overline{\imath}\, h^2_I(t)\, \sigma_I(t) = \overline{\imath}\, \left(\frac{\sigma^2_{A_I} + d(t)}{\sigma^2_I + d(t)}\right)\, \sqrt{\sigma^2_I + d(t)} \tag{33.11b}$$

Under directional selection, d rapidly approaches its equilibrium value (Chapter 13). The equilibrium additive variance in the index $\widetilde{\sigma}^2_{A_I}$ is given by Equation 13.13, as obtained by several workers (Bruns and Harvey 1976, Bennett and Swiger 1980, Gomez-Raya and Burnside 1990, Villanueva and Kennedy 1993). Since $\widetilde{d} = \widetilde{\sigma}^2_{A_I} - \sigma^2_{A_I}(0)$, the equilibrium phenotypic variance and heritability of the index follow as $\widetilde{\sigma}^2_I = \sigma^2_I(0) + \widetilde{d}$ and $\widetilde{h}^2_I = \widetilde{\sigma}^2_{A_I}/\widetilde{\sigma}^2_I$.

Changes in G and P Under Index Selection

While the behavior of the variance components of the index simply follows from a univariate treatment, the behavior of the covariance matrices of its components is a bit more involved. To examine these changes, we again make the standard assumption of the infinitesimal model so that allele frequency changes can be ignored and all changes genetic variances are due solely to gametic phase disequilibrium. Following Chapter 31, the multivariate extension for disequilibrium (which easily follows from an element -by-element comparison) is to express the additive-genetic and phenotypic covariance matrices at generation t as $\mathbf{G}_t = \mathbf{G} + \mathbf{D}_t$ and $\mathbf{P}_t = \mathbf{P} + \mathbf{D}_t$. The unsubscripted matrices denote their linkage equilibrium values, while $\mathbf{D}_t$ is the matrix of gametic-phase disequilibrium values. As in the univariate case, for unlinked loci, transmission reduces each element of $\mathbf{D}$ by one-half of its previous value each generation, giving $\boldsymbol{\Delta}\mathbf{D}_{t+1} = -(1/2)\,\mathbf{D}_t$. Likewise, selection also generates disequilibrium, $\mathbf{D}^* = \mathbf{G}^* - \mathbf{G}$, where $\mathbf{G}^*$ is the genetic covariance matrix after selection (but before reproduction). Again only half of this new equilibrium is transmitted to the offspring. Putting these together gives the change in the disequilibrium matrix as the sum of both components, or

$$\boldsymbol{\Delta}\mathbf{D}_{t+1} = \frac{1}{2}\,(\mathbf{D}^*_t - \mathbf{D}_t)$$

To follow the change in covariance matrices, we assume $\mathbf{D}_0$ is a matrix of zeros (no initial disequilibrium), and the dynamics of $\mathbf{G}$ and $\mathbf{P}$ follow from the dynamics of $\mathbf{D}_t$, which follow upon specification of $\mathbf{D}^*$, the change induced by selection. Several authors (Zeng 1988, Villanueva and Kennedy 1990, Itoh 1991) have shown that the within-generation change in the phenotypic covariance matrix caused by truncation selection on I is

$$\mathbf{P}^* = \mathbf{P} - \frac{\kappa}{\sigma^2_I}(\mathbf{P}\mathbf{b})(\mathbf{P}\mathbf{b})^T \tag{33.12a}$$

where κ is given by Equation 31.10. Equation 31.10a can also be written as

$$\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1} = -\frac{\kappa}{\sigma^2_I}\,\mathbf{b}\mathbf{b}^T \tag{33.12b}$$

Example 33.4 gives the derivation of both these expressions. The advantage of Equation 33.12b follows from Equation 30.12,

$$\begin{aligned}\mathbf{D}^* = \mathbf{G}^* - \mathbf{G} &= \mathbf{G}\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1}\mathbf{G} = -\frac{\kappa}{\sigma_I^2}\mathbf{G}\mathbf{b}\mathbf{b}^T\mathbf{G} \\ &= -\frac{\kappa}{\sigma_I^2}\left(\mathbf{G}\mathbf{b}\right)\left(\mathbf{G}\mathbf{b}\right)^T \end{aligned} \tag{33.13}$$

The resulting equation for $\mathbf{D}_t$ becomes

$$\Delta\mathbf{D}_t = -\frac{1}{2}\left[\left(\frac{\kappa}{\sigma_I^2(t)}\right)(\mathbf{G}_t\mathbf{b})(\mathbf{G}_t\mathbf{b})^T + \mathbf{D}_t\right] \tag{33.14a}$$

$$= -\frac{1}{2}\left[\left(\frac{\kappa}{\mathbf{b}^T(\mathbf{P}+\mathbf{D}_t)\mathbf{b}}\right)([\mathbf{G}+\mathbf{D}_t]\mathbf{b})([\mathbf{G}+\mathbf{D}_t]\mathbf{b})^T + \mathbf{D}_t\right] \tag{33.14b}$$

where it is generally assumed $\mathbf{D}_0 = \mathbf{0}$. As with the univariate case, we obtain the desired values by iterating Equation 33.14.

As a final check, let's use the above multivariate results to recover to change in d for the index. The phenotypic and genetic variances of the index in generation t are

$$\sigma_I^2(t) = \mathbf{b}^T\mathbf{P}_t\mathbf{b} = \mathbf{b}^T(\mathbf{P}_0 + \mathbf{D}_t)\mathbf{b} = \sigma_I^2(0) + d(t) \tag{33.15a}$$

and

$$\sigma_{A_I}^2(t) = \mathbf{b}^T\mathbf{G}_t\mathbf{b} = \mathbf{b}^T(\mathbf{G}_0 + \mathbf{D}_t)\mathbf{b} = \sigma_{A_I}^2(0) + d(t) \tag{33.15b}$$

where $\mathbf{b}^T\mathbf{D}_t\mathbf{b} = d(t)$ is the disequilibrium in I. Applying Equation 33.14,

$$\begin{aligned}\Delta d(t) = \mathbf{b}^T\Delta\mathbf{D}_t\mathbf{b} &= -\mathbf{b}^T\left(\frac{1}{2}\left[\frac{\kappa}{\sigma_I^2(t)}\mathbf{G}_t\mathbf{b}\mathbf{b}^T\mathbf{G}_t + \mathbf{D}_t\right]\right)\mathbf{b} \\ &= -\frac{1}{2}\left(\frac{\kappa}{\sigma_I^2(t)}\right)\left(\mathbf{b}^T\mathbf{G}_t\mathbf{b}\right)\left(\mathbf{b}^T\mathbf{G}_t\mathbf{b}\right) - \frac{1}{2}\mathbf{b}^T\mathbf{D}_t\mathbf{b} \\ &= -\frac{1}{2}\cdot\frac{\kappa}{\sigma_I^2(t)}\,\sigma_{A_I}^4(t) - \frac{d(t)}{2}\end{aligned} \tag{33.15c}$$

recovering Equation 33.11a, and showing that Bulmer's univariate results also apply to an index under truncation selection, even though the index itself is composed of several different characters whose variances and covariances are changing.

Example 33.4. Let's derive Equations 33.12c and 33.12d. Our starting point is Equation 13.14, which give the changes in variances and covariances for traits responding to correlated selection. Here, selection is on the index $I = \mathbf{b}^T\mathbf{z}$, and we see correlated changes in the component traits z_i. Mentioned breifly in earlier in the chapter, a very useful identity is the covariance between the index and a particular component,

$$\sigma(I, z_i) = \sigma(\mathbf{b}^T\mathbf{z}, z_i) = \sigma\left(\sum_j b_j z_j, z_i\right) = \sum_j b_j\sigma(z_j, z_i) = (\mathbf{P}\mathbf{b})_i$$

where $(\mathbf{P}\mathbf{b})_i$ denotes the i-th element in the vector $\mathbf{P}\mathbf{b}$. Note that this expression accounts for the fact that z_i may also be correlated with other components. From Equation 13.14d, the

change in the phenotypic covariance for component traits i and j from selection on the index I is

$$\Delta P_{ij} = -\kappa \, \frac{\sigma(z_i, I)\sigma(z_j, I)}{\sigma_I^2} = -\frac{\kappa}{\sigma_I^2} \, (\mathbf{Pb})_i \, (\mathbf{Pb})_j$$

Noting that the matrix terms correspond to the ijth element of the matrix multiplication of $(\mathbf{Pb})(\mathbf{Pb})^T$, we recover Equation 31.12a,

$$\Delta \mathbf{P} = \mathbf{P}^* - \mathbf{P} = -\frac{\kappa}{\sigma_I^2} (\mathbf{Pb})(\mathbf{Pb})^T$$

Next, recalling that $(\mathbf{Pb})^T = \mathbf{b}^T\mathbf{P}^T = \mathbf{b}^T\mathbf{P}$, we can rewrite the above equation as

$$\mathbf{P}^* - \mathbf{P} = -\frac{\kappa}{\sigma_I^2} \, \mathbf{Pbb}^T\mathbf{P}$$

pre- and post-multiplying by $\mathbf{P}^{-1}$ recovers Equation 33.12b,

$$\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1} = -\frac{\kappa}{\sigma_I^2} \, \mathbf{bb}^T$$

OPTIMIZING THE EXPECTED RESPONSE OF A LINEAR INDEX

A common goal of multiple character selection is to maximize the response of some overall **merit function** $H(\mathbf{g})$ based on an index of the trait breeding values. The merit is given as a function of breeding values because these are what is passed onto the offspring of the selected parents (Chapter 10), and hence (for a linear index) the response of interest. Typically, the merit function is taken to be a linear index $H(\mathbf{g}) = \mathbf{a}^T\mathbf{g}$, where the vector of **economic weights a** assigns the desirability of relative responses in each character. For example, if a unit response in character one is three times more desirable than a unit response in character two, $a_1/a_2 = 3$. Economic weights are either preset by the investigator or estimated by some prediction of an individual's overall merit as a function of $\mathbf{z}$. An example of this latter approach is the prediction of individual fitness w (merit in this case) from the regression of w on $\mathbf{z}$ (Chapters 29, 30). Other methods for estimating economic weights are reviewed by Harris (1970), Gjedrem (1972), Melton et al. (1979, also see cautions by Thompson 1980 and Goddard 1983 and the reply by Melton et al. 1993), and Cotterhill and Jackson (1985). For example, Bernardo (1991) notes that for some aspects of plant breeding, practitioners use an intuitive weight to either chose or reject lines for future consideration. This leads to 0/1 (reject/include) data, and Bernardo was able to generate economic weights by applying a retrospective index, providing some quantification to an otherwise intuitive process. Note that this is equivalent to survival data under natural selection, and so the machinery of Chapter 29, which is in part based on a retrospective index, can also be used.

The Index of Selection Usually Does Not Equal the Index of Response

Index selection puts together two important concepts from selection theory. The first is from univariate selection: the response depends upon the breeding values of the parents (Chapter 10). Phenotype is one predictor of breeding value, with the correlation between an individual's breeding and phenotypic values given by h^2. If we have additional information, we may be able to improve, often substantially, our ability to predict breeding value, and thus

improve response. Gathering this additional information for an individual into some index, we may find that the correlation between this index (if well chosen) and breeding value significantly exceeds h^2. The second key concept is from multivariate selection: $\mathbf{G}$ rotates and scales an initial selection vector to give a final response vector (Chapter 30). Thus, in order to achieve a specific *response*, we may need to select in a rather different direction.

Under index selection, we are trying to maximize the response in a specific trait, namely the merit, whose component values are known. For now, we assume that merit is a linear combination of these components, and hence the *phenotypic* value for the merit of an individual can be written as

$$\mathbf{a}^T\mathbf{z} = \sum_{i=1}^{k} a_i\, z_i$$

We could easy assign each individual in the population a merit score and then simply select directly on this, choosing those individuals with the largest values of $\mathbf{a}^T\mathbf{z}$ to form the next generation. However, it is *not* the phenotypic value that we are trying to maximize among the parents selected to form the next generation, but rather their *breeding* values,

$$H = \mathbf{a}^T\mathbf{g} = \sum a_i\, g_i \tag{33.16}$$

where g_i is the breeding value for component trait i. The power behind index selection is that if we know the component trait values and their genetic and phenotypic variances, we can almost always improve (or at worst equal) the response by selecting on some appropriately-chosen index versus selecting directly on the trait itself. The reason is that we can find an index, based on the component traits, that is a better predictor of an individual's breeding value for merit H than is their phenotypic value for merit. Put another way, a larger response in the merit can be obtained by selecting on a *correlated character* (another index $I = \mathbf{b}^T\mathbf{z}$) than by directly selecting on the character ($\mathbf{a}^T\mathbf{z}$) itself. We can thus distinguish between a **selection index** of phenotypic values that are used in the selection decision and a **response index** of breeding values that we wish to maximize.

Let's express this a bit more formally. Our goal is to maximize the response of merit, which is done by maximizing the value of H in our selected parents. H is the additive genetic value for merit, and $H - \mu_H = H - \mathbf{a}^T\boldsymbol{\mu}$ its breeding value (scaled to start the population with a mean breeding value of zero). H also goes by many other names in the literature (e.g., **aggregate genetic value**, **genetic merit**, **breeding value for merit**, **profit**, to name a few). Our task is thus to find the linear index based on measurable phenotypic values $I_s = \mathbf{b}_s^T\mathbf{z}$ that has the highest correlation with $H = \mathbf{a}^T\mathbf{g}$. Because of genetic and phenotypic correlations between characters, the best predictor of the additive genetic value of merit H is usually *not* the observed phenotypic merit $\mathbf{a}^T\mathbf{z}$. Expressed in terms of a multivariate response, if we select along the vector $\mathbf{a}^T\mathbf{z}$, $\mathbf{G}$ rotates and scales the response away from this vector, reducing response in the direction desired by selection. However, if we select in an appropriately-chosen direction, then as $\mathbf{G}$ rotates and scales its response, it will align with the maximal change in $\mathbf{a}^T\mathbf{z}$.

Selection and Response Indices With Non-overlapping Traits

Recall that we earlier showed the multivariate breeder's equation can be generalized by allowing the vector of selection differentials to contain different traits from the vector of responses. This result is very useful, as one powerful feature index selection is that index of phenotypic selection I and index of breeding values in merit (i.e., the index of response H) can consist of different traits. We consider the general case where H and I may have at least

some (in the extreme, all) non-overlapping traits,

$$H = \sum_{i=1}^{k} a_i g_i \quad \text{and} \quad I = \sum_{j=1}^{n} b_j z_j$$

Here z_j is the phenotypic value of trait j in the selection index, while g_i is the breeding value for the ith trait in the response index. For this more general case, following our development of the more general version of the multivariate breeder's equation, define $\mathbf{P}$ as the $n \times n$ phenotypic covariance matrix of $\mathbf{z}$. Likewise define $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$ is a $k \times n$ covariance matrix whose elements $G_{ij} = \sigma(g_i, z_j)$ for $1 \leq i \leq k, 1 \leq j \leq n$ are the covariances between the additive genetic values of the k characters comprising H and the phenotypic values of the n characters comprising I. In this case $\mathbf{G}$ is generally *not* symmetric (Example 33.5). When I and H contain the same traits, $\mathbf{G}$ reverts back to a standard $n \times n$ symmetric matrix of genetic covariances. For those cases where H and I contain different elements, it will also be useful to define $\mathbf{G}_g = \sigma(\mathbf{g}, \mathbf{g})$ as the normal genetic covariance matrix for the traits in H. When the traits in both I and H completely overlap, $\mathbf{G} = \mathbf{G}^T = \mathbf{G}_g$.

Example 33.5. Consider the following example from Yamada et al. (1975, with slight corrections from Gibson and Kennedy 1990). The goal was to improve egg production (EP), feed conversion efficiency (FC), and egg weight (EW) in chickens. These are the component traits whose breeding values make up the response index. Phenotypic selection was based on egg weight, adult body weight (BW), and a measure of the individual's average egg production plus that of seven of her full sisters (IEP). Note that egg weight is the only trait present in both the selection and response indices. The resulting vector of selected traits and breeding value of interest for response are

$$\mathbf{z} = \begin{pmatrix} \text{EW} \\ \text{IEP} \\ \text{BW} \end{pmatrix}, \qquad \mathbf{g} = \begin{pmatrix} \text{EP} \\ \text{FC} \\ \text{EW} \end{pmatrix}$$

The phenotypic covariance matrix $\mathbf{P}$ among the selected traits, $P_{ij} = \sigma(z_i, z_j)$, was estimated as

$$\mathbf{P} = \begin{pmatrix} 16 & -1.53 & 28.8 \\ -1.53 & 25.63 & -1.13 \\ 28.8 & -1.13 & 324 \end{pmatrix},$$

while the matrix of covariance between phenotypic and breeding values, $G_{ij} = \sigma(g_i, z_j)$, were estimated as follows:

$$\mathbf{G} = \begin{pmatrix} \sigma[g(EP), z(EW)] & \sigma[g(EP), z(IEP)] & \sigma[g(EP), z(BW)] \\ \sigma[g(FC), z(EW)] & \sigma[g(FC), z(IEP)] & \sigma[g(FC), z(BW)] \\ \sigma[g(EW), z(EW)] & \sigma[g(EW), z(IEP)] & \sigma[g(EW), z(BW)] \end{pmatrix}$$

$$= \begin{pmatrix} -7.59 & 11.71 & 0 \\ -1.02 & -1.38 & -3.05 \\ 8 & 2.61 & 12.88 \end{pmatrix}$$

Here $\mathbf{G}$ is *not* symmetric, as different traits are involved, e.g., $G_{12} = \sigma$[g(EP), z(IEP)] =11.71, while $G_{21} = \sigma$[g(FC), z(EW)] $= -1.02$.

The Smith-Hazel Index

Our treatment of the Smith-Hazel index allows for the general case where H and I may contain different traits. Recall in this case that $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$ and $\mathbf{G}_g$ refers to the genetic covariance matrix associated with the response vector. The goal is to find the weights $\mathbf{b}_s$ for an index $I = \mathbf{b}_s^T\mathbf{z}$ of selection to give it the highest correlation with the index of response $H = \mathbf{a}^T\mathbf{g}$. To do so, first note that

$$\sigma_H^2 = \sigma(\mathbf{a}^T\mathbf{z}, \mathbf{a}^T\mathbf{z}) = \mathbf{a}^T\sigma(\mathbf{g}, \mathbf{g}^T)\mathbf{a} = \mathbf{a}^T\mathbf{G}_g\mathbf{a}$$

and likewise $\sigma_I^2 = \mathbf{b}^T\mathbf{P}\mathbf{b}$. Finally,

$$\sigma_{H,I} = \sigma(\mathbf{a}^T\mathbf{g}, \mathbf{b}^T\mathbf{z})\mathbf{b} = \mathbf{a}^T\sigma(\mathbf{g}, \mathbf{z})\mathbf{b} = \mathbf{a}^T\mathbf{G}\mathbf{b}$$

Putting these together gives the correlation between the breeding value of merit H and a phenotypic index I as

$$\rho_{H,I} = \frac{\sigma_{H,I}}{\sigma_H\sigma_I} = \frac{\mathbf{a}^T\mathbf{G}\mathbf{b}}{\sqrt{\mathbf{a}^T\mathbf{G}_g\mathbf{a}}\ \sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}}} \tag{33.17a}$$

From standard regression theory (LW Chapter 3), the fraction of variation in H accounted for by I is $\rho_{H,I}^2$ so that Equation 33.17a provides a measure of how well I predicts H. To obtain the value of $\mathbf{b}$ maximizing this correlation, first note the since $\mathbf{a}$ is a constant we need only maximize

$$\frac{\mathbf{a}^T\mathbf{G}\mathbf{b}}{\sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}}} \tag{33.17b}$$

Both quadratic products yield a scalar, so that derivatives can be taken by using the standard quotient rule. Taking the derivative with respect to $\mathbf{b}$ (Appendix 5) and denoting solutions giving a derivative of zero by $\widetilde{\mathbf{b}}$, gives

$$\left(\widetilde{\mathbf{b}}^T\mathbf{P}\widetilde{\mathbf{b}}\right)\mathbf{G}^T\mathbf{a} = \left(\mathbf{a}^T\mathbf{G}\widetilde{\mathbf{b}}\right)\mathbf{P}\widetilde{\mathbf{b}} \tag{33.17c}$$

Since both $\widetilde{\mathbf{b}}^T\mathbf{P}\widetilde{\mathbf{b}}$ and $\mathbf{a}^T\mathbf{G}\widetilde{\mathbf{b}}$ are scalars, solutions are of the form $\mathbf{P}\widetilde{\mathbf{b}} = c \cdot \mathbf{G}^T\mathbf{a}$, giving the optimal vector of weights as

$$\mathbf{b}_s = \mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} \tag{33.18a}$$

These weights give the **Smith-Hazel selection index**

$$\begin{aligned} I_s &= \mathbf{b}_s^T\mathbf{z} \\ &= \left(\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}\right)^T\mathbf{z} && (33.18b) \\ &= \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{z} && (33.18c) \end{aligned}$$

Smith (1936), following a suggestion by Fisher to use his recently developed method of discriminant functions (Fisher 1936), obtained this index for the special case of selection on a collection of pure lines (varietal selection, Chapter 20). Hazel (1945) extended Smith's results to outbreeding populations by considered the change in breeding value.

The Smith-Hazel index is the most widely-used set of weights for a linear selection index, as using the weights $\mathbf{b}_s$ for the index of phenotypic selection maximizes response in the index $H = \mathbf{a}^T\mathbf{z}$. The Smith-Hazel index depends *critically* on having good estimates of $\mathbf{G}$ and $\mathbf{P}$. Incorrect estimates result in a less than optimal index.

Example 33.6. Returning to our soybean data (Example 33.1), what are the Smith-Hazel weights to maximize the response in an index that weights yield four times as much as the other two traits? Here, $\mathbf{a} = \begin{pmatrix} 1 & 1 & 4 \end{pmatrix}^T$, giving the Smith-Hazel weights $\mathbf{b}_s$ as

$$\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} = \begin{pmatrix} 287.5 & 477.4 & 1266 \\ 477.4 & 935 & 2303 \\ 1266 & 2303 & 5951 \end{pmatrix}^{-1} \begin{pmatrix} 128.7 & 160.6 & 492.5 \\ 160.6 & 254.6 & 707.7 \\ 492.5 & 707.7 & 2103 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 5.39 \\ -9.28 \\ 4.06 \end{pmatrix}$$

Thus the optimal index to improve $z_1 + z_2 + 4z_3$ is to select on $5.4z_1 - 9.3z_2 + 4.1z_3$.

What is the response? Equation 33.3 gives the response for $I_s = \mathbf{b}_s^T\mathbf{z}$ as 177.3 (assuming $\overline{i}$ = 2.06). The careful reader might recall from Example 34.1 that using the "naive" weights $\mathbf{a}^T = \begin{pmatrix} 1 & 1 & 4 \end{pmatrix}$, we obtained a response in the index of 255.1. What's going on here, as it seems that our "optimal" index gives a much smaller response? The key is that in Example 33.1, we selected on, and predicted the response for, $I = \mathbf{a}^T\mathbf{z}$. However, with a Smith-Hazel index we select on one index $I_s = \mathbf{b}_s^T\mathbf{z}$ and are interested in a response in another (in essence, a correlated trait), $\mathbf{a}^T\mathbf{z}$. Thus, applying Equation 33.3 gives the response in I_s. In order to convert this into a response in I, the index we wish to improve, we apply Equation 33.6, giving

$$R = \overline{i} \cdot \frac{\mathbf{a}^T\mathbf{G}\mathbf{b}_s}{\sqrt{\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s}} = 299.1$$

Thus, basing selection on the index $I_s = \mathbf{S}^T\mathbf{z}$ gave a larger response in the index $I = \mathbf{a}^T\mathbf{z}$ (i.e., gives a larger value of the index $H = \mathbf{a}^T\mathbf{g}$) than occurs by directly selecting on I.

To get some appreciation of how each character is weighted in the Smith-Hazel index, consider the case where there are no phenotypic or genetic correlations ($\mathbf{P}$ and $\mathbf{G}$ are diagonal). Here $b_i = a_i h_i^2$, giving the index as $\sum a_i h_i^2 z_i$ so that characters with both large heritabilities *and* large economic weights receive the most value, while either a large heritability, or economic weight, by itself is not sufficient to insure a large weight. Constructing the Smith-Hazel index requires three sets of parameters – estimates of $\mathbf{P}$, $\mathbf{G}$ and $\mathbf{a}$. Errors in estimating $\mathbf{a}$ appear to have only a small effect on the index weighting (reviewed by James 1982 and Smith 1983), while the consequences of using estimates of $\mathbf{P}$ and $\mathbf{G}$ are considerable and will be discussed shortly.

Properties of the Smith-Hazel Index

1) ***Selection on the Smith-Hazel index provides largest response in merit for a fixed selection intensity*** $\overline{i}$. Equation 33.6 shows that the maximal response in merit given selection on another index $\mathbf{b}^T\mathbf{z}$ is given by maximizing the same expression as Equation 33.18a, namely by using the Smith-Hazel weights $\mathbf{b} = \mathbf{b}_s$. Noting that

$$\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s = \left(\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}\right)^T \mathbf{P}\mathbf{b}_s = \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{P}\mathbf{b}_s = \mathbf{a}^T\mathbf{G}\mathbf{b}_s$$

Equation 33.6, gives the expected response in merit using the Smith-Hazel index as

$$\begin{aligned} R_H &= \overline{i} \cdot \frac{\mathbf{a}^T\mathbf{G}\mathbf{b}_s}{\sqrt{\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s}} = \overline{i} \cdot \sqrt{\mathbf{a}^T\mathbf{G}\mathbf{b}_s} \\ &= \overline{i} \cdot \sqrt{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}} \end{aligned} \tag{33.19}$$

while from Equation 33.5 the change in the vector of character means comprising the response index is

$$\mathbf{R} = \left(\frac{\overline{i}}{\sigma_I}\right) \cdot \mathbf{G}\mathbf{b}_s = \overline{i} \cdot \frac{\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}}{\sqrt{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}}} \tag{33.20}$$

2) ***The Smith-Hazel index is closely related to the least squares regression of breeding value for merit on the vector of phenotypic values*** $\mathbf{z}$. This regression can be written as $E[H \,|\, \mathbf{z}] = a + \mathbf{b}^T\mathbf{z}$, where $\mathbf{b}$ and a are chosen such that the regression accounts for the largest amount of variation in H. This occurs when $\mathbf{b}$ is chosen to maximize the correlation between H and $\mathbf{z}$ and when a satisfies $E[H] = a + \mathbf{b}^T E[\mathbf{z}]$ (LW Chapters 3, 8). Given that $\mathbf{b}_s$ maximizes this correlation, the least squares regression is

$$E[H \,|\, \mathbf{z}] = a + \mathbf{b}_s^T\mathbf{z} = a + I_s \tag{33.21a}$$

Noting that $E[H] = \mathbf{a}^T\boldsymbol{\mu}$ and $E[\mathbf{z}] = \boldsymbol{\mu}$ gives $a = \mathbf{a}^T\boldsymbol{\mu} - \mathbf{b}_s^T\boldsymbol{\mu}$, hence

$$E[\,(H - \mathbf{a}^T\boldsymbol{\mu}) \,|\, \mathbf{z}\,] = \mathbf{b}_s^T(\mathbf{z} - \boldsymbol{\mu}) \tag{33.21b}$$

Since breeding value is the deviation of additive genetic value from the mean, the regression of the breeding value for merit on phenotypic value is

$$\mathbf{b}_s^T(\mathbf{z} - \boldsymbol{\mu}) = I_s - \mathbf{b}_s^T\boldsymbol{\mu} \tag{33.21c}$$

Estimates of breeding value based on least squares regressions are called often best linear predictors (BLPs). Hence, if phenotypic characters are standardized to mean zero, the Smith-Hazel index is the BLP of breeding value for merit. The related method of BLUP (best linear unbiased predictor, LW Chapter 26; Chapters 16, 34), provides the best estimate of breeding value when more general pedigree information is available. Similarities and differences of the Smith-Hazel index BLP estimates and BLUP estimates are examined in more detail in Chapter 34.

Just how much of the variance in breeding values is explained by the regressions given by Equations 33.21a-c? From standard regression theory the fraction of variance explained is $\rho^2(H, a + \mathbf{b}_s^T\mathbf{z}) = \rho^2(H, \mathbf{b}_s^T\mathbf{z}) = \rho^2(H, I_s)$. Substituting $\mathbf{b}_s$ into Equation 33.17 gives

$$\rho_{H,I_s} = \frac{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}}{\sqrt{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}} \cdot \sqrt{\mathbf{a}^T\mathbf{G}_g\mathbf{a}}} = \sqrt{\frac{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}}{\mathbf{a}^T\mathbf{G}_g\mathbf{a}}} \tag{33.22a}$$

Thus the fraction of the additive genetic variance in merit explained by the Smith-Hazel index is

$$\frac{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}}{\mathbf{a}^T\mathbf{G}_g\mathbf{a}} \tag{33.22b}$$

leaving a residual (unexplained) variance of $(1 - \rho^2)\sigma_H^2$ or

$$\left(\frac{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}}{\mathbf{a}^T\mathbf{G}_g\mathbf{a}}\right) \cdot \mathbf{a}^T\mathbf{G}_g\mathbf{a} = \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} \tag{33.22c}$$

The relative performance of any candidate with respect to the optimal Smith-Hazel index can be obtained by comparing the correlation of the candidate index and H with Equation 33.22a.

Equations 33.21 and 33.22 do not require the assumption of multivariate normality, while Equations 33.19 and 33.20 do. Another useful property of the Smith-Hazel index (due to Williams 1962a and Henderson 1963), which also requires multivariate normality, is

3) ***The Smith-Hazel index gives the maximal probability of selecting the individual with the largest breeding value for merit in a sample.***

Other Useful Results for the Smith-Hazel Index

Three popular expressions in the literature relating to the Smith-Hazel index are

$$\sigma(H, I_s) = \sigma^2(I_s), \qquad \rho(H, I_s) = \frac{\sigma(I_s)}{\sigma(H)}, \qquad R_H = \overline{i} \cdot \sigma(I_s) \tag{33.23}$$

These are obtained by first noting for $\mathbf{b}_s = \mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}$ that

$$\sigma^2(I_s) = \mathbf{b}_s^T\mathbf{P}\mathbf{b}_s = \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} = \mathbf{a}^T\mathbf{G}\mathbf{b}_s = \sigma(H, I_s) \tag{33.24}$$

Substitution into Equations 33.22a and 33.19 (respectively) gives the last two identities.

When the phenotypic and genetic covariances are estimated in several populations and/or environments (such as different years), the investigator is faced with a decision as how to combine these estimates when constructing an index. At one extreme the index can be computed using pooled covariance matrices to give an **average index**. At the other extreme, separate indices can be constructed for each population/environment, giving **specific indices**. Hanson and Johnson (1957) develop an approach for the optimal response that differentially weights the covariances matrices, yielding what they refer to as a **general index**. Caldwell and Weber (1965), examining response in soybeans over four populations, found that specific indices gave the best overall performance, but that either general or average indices were reasonable substitutes. Clearly, this is an area for both more theoretical and experimental investigation.

Estimated, Base, Elston, and Other Indices

The Smith-Hazel index requires that both the phenotypic and genetic covariance matrices are known. Since these are usually unknown, the **estimated index**

$$\widehat{I}_s = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}\mathbf{a}^T\mathbf{z} \tag{33.25}$$

is constructed using the estimated phenotypic and genetic covariance matrices ($\widehat{\mathbf{P}}$ and $\widehat{\mathbf{G}}$). Due to the inaccuracies inherent in estimating these matrices (especially $\mathbf{G}$), this index may be quite different from the correct Smith-Hazel index. This lead Panse (1946), Brim et al. (1959), and Williams (1962a,b) to suggest that the **base index**

$$I_b = \sum a_i z_i \tag{33.26}$$

which is independent of $\mathbf{P}$ and $\mathbf{G}$, may in many cases be preferable. Note that the base index is only defined when the traits are H and I are identical (so that $\mathbf{G}$ is now symmetric), and that selecting on the base index is equivalent to direct selection on the phenotypic value of the merit. The base and Smith-Hazel indices are identical when there are no genetic and phenotypic correlations and all characters have the same heritability. More generally, the two indices are equivalent when $\mathbf{P}^{-1}\mathbf{G} = c \cdot \mathbf{I}$ or $\mathbf{P} = c \cdot \mathbf{G}$ for any positive constant c. Heidhues and Henderson (1962) have proposed the **heritability index** with weights $b_i = a_i\, h_i^2$ as an alternative to the base index when confidence in genetic covariance estimates is low. Note

from previous discussion that this reduces to the Smith-Hazel index when there are no genetic or phenotypic correlations between characters. Smith et al. (1981) go a step further, proposing to use the heritabilities as the economic weighs, $a_i = h_i^2$.

The expected response under the base index relative to the true Smith-Hazel index is given by the correlation between these indices,

$$\begin{aligned}\rho_{I_s,I_b} &= \frac{\sigma(I_s, I_b)}{\sigma_{I_s} \cdot \sigma_{I_b}} = \frac{\sigma(\mathbf{b}_s^T\mathbf{z}, \mathbf{a}^T\mathbf{z})}{\sqrt{\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s \cdot \mathbf{a}^T\mathbf{P}\mathbf{a}}} \\ &= \frac{\mathbf{b}_s^T\mathbf{P}\mathbf{a}}{\sqrt{\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s \cdot \mathbf{a}^T\mathbf{P}\mathbf{a}}} = \frac{\mathbf{a}^T\mathbf{G}\mathbf{a}}{\sqrt{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}\mathbf{a} \cdot \mathbf{a}^T\mathbf{P}\mathbf{a}}}\end{aligned} \tag{33.27}$$

Although the base index can be applied without estimates of $\mathbf{G}$ and $\mathbf{P}$, it still requires the assignment of economic weights. If estimates of these covariance matrices are available, the method of desired-gains to be discussed shortly can be use to construct an index without having to specify $\mathbf{a}$. In the extreme case where $\mathbf{P}$, $\mathbf{G}$ and $\mathbf{a}$ are all unknown, Elston (1963) suggests the nonlinear index

$$I_e = (z_1 - m_1)(z_2 - m_2)\cdots(z_n - m_n) = \prod_{j=1}^{n}(z_j - m_j) \tag{33.28}$$

where m_j is the minimal value of character j and each character is scaled to have unit variance (Figure 33.3, at the end of this chapter, shows the form of this index for two characters). In effect, the **Elston index** (occasionally called the **weight-free index**) assumes all characters are equally weighted (Elston 1963, Baker 1974). Theoretical results (Cotterhill 1985) suggest that if the traits in the index are positively correlated (both genetically and phenotypically) the Smith-Hazel, base, and Elston indices give very similar responses. However, if there are negative correlations (genetically or phenotypically) the Smith-Hazel index is significantly superior. Experimental studies comparing base versus estimated indices reviewed later (Table 33.4) show that both often give similar responses. A final nonparametric index in the **rank summation index** of Mulamba and Mock (1978), which ranks all traits and then assigns the sum of an individual's ranks as their index score. Crosbie et al. (1980) noted that the rank summation, base, and Elston indices are not seriously influenced by unequal variances among traits and thus should be consider more often.

Example 33.7. Once again, we return to the soybean data Brim et al. (1959) presented in Example 33.1. The same traits are used for both I and H so that $\mathbf{G} = \mathbf{G}^T$. Assume characters have equal economic weight so that $a_i = 1$. The resulting vector of weights $\widehat{\mathbf{b}}_s$ for the (estimated) Smith-Hazel index is

$$\widehat{\mathbf{b}}_s = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}2.4\\-2.9\\1.2\end{pmatrix}$$

giving the index as $\widehat{I}_s = 2.4z_1 - 2.9z_2 + 1.2z_3$. In contrast, the base index is $I_b = \mathbf{a}^T\mathbf{z} = z_1 + z_2 + z_3$. Suppose two individuals are examined, one of which will be saved. The oil, protein, and yield scores of these individuals are (1, 2, 3) and (1, 1, 1), respectively. Under the estimated index, these individuals have scores of $2.4 \cdot 1 - 2.9 \cdot 2 + 1.2 \cdot 3 = 0.2$ and

$2.4 - 2.9 + 1.2 = 0.7$, while under the base index these individual have scores $1 + 2 + 3 = 6$ and $1 + 1 + 1 = 3$. Hence, individual two is saved under the Smith-Hazel index, while individual one is saved under the base selection. To compare the responses under the base versus Smith-Hazel index, assume that the error from using the estimates of $\mathbf{P}$ and $\mathbf{G}$ is small. Since $(\mathbf{a}^T\widehat{\mathbf{G}}\widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}\mathbf{a})^{1/2} \simeq 49.67$, from Equation 33.19 the response in H to selection on $\widehat{I}_s$ is $\overline{i} \cdot 49.67$. For the base index $\mathbf{a} = (1, 1, 1)^T$, $\sigma_I = \sqrt{\mathbf{a}^T\widehat{\mathbf{P}}\mathbf{a}} \simeq 123.6$ and $\mathbf{a}^T\widehat{\mathbf{G}}\mathbf{a} \simeq 5208$. Substituting these into Equation 33.3 gives

$$\frac{R_I}{\overline{i}} = \frac{\mathbf{a}^T\widehat{\mathbf{G}}\mathbf{a}}{\sqrt{\mathbf{a}^T\widehat{\mathbf{P}}\mathbf{a}}} \simeq \frac{5208}{123.6} \simeq 33.1$$

which is only 85 percent of the expected response under the Smith-Hazel index. Applying Equation 33.17a, the correlation between the two indices (assuming that the estimated index is the correct Smith-Hazel index) is

$$\frac{\mathbf{a}^T\widehat{\mathbf{G}}\mathbf{a}}{\sqrt{\mathbf{a}^T\widehat{\mathbf{G}}\widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}\mathbf{a} \cdot \mathbf{a}^T\widehat{\mathbf{P}}\mathbf{a}}} \simeq \frac{5208}{49.67 \cdot 123.6} \simeq 0.85$$

as expected from the response in the base index relative to the Smith-Hazel index.

While Equation 33.27 gives the relative efficiency of using the base index in place of the true Smith-Hazel index, just is how much error is introduced by using the estimated index $\widehat{\mathbf{b}}_s = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}\mathbf{a}$ in place of the true index $\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}$? As Harris (1963) notes, we need to distinguish the optimal response (that obtained using the true Smith-Hazel index I_s) from the predicted response (which assumes $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{P}}$ are correct) and the achieved response (the expected response using $\widehat{I}_s$), where

$$\frac{R}{\overline{i}} = \begin{cases} \sqrt{\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s} & \text{optimal response} \\ \sqrt{\widehat{\mathbf{b}}_s^T\widehat{\mathbf{P}}\widehat{\mathbf{b}}_s} & \text{predicted response} \\ \dfrac{\mathbf{a}^T\mathbf{G}\widehat{\mathbf{b}}_s}{\sqrt{\widehat{\mathbf{b}}_s^T\mathbf{P}\widehat{\mathbf{b}}_s}} & \text{achieved response} \end{cases} \tag{33.29}$$

The expression for achieved response was obtained by substituting $\mathbf{b} = \widehat{\mathbf{b}}_s$ into Equation 33.6. Thus there are two classes of errors using the estimated index. Errors in estimates of $\mathbf{P}$ and $\mathbf{G}$ not only give incorrect index weighting, they also yield incorrect predictions of the response to selection on this index. Hanson and Johnson (1957) note that the ratio of the estimated response to the optimal response, $\widehat{R}/R$ equals the correlation between the estimated and optimal indices. Hence, the estimated response is always (on average) less than the optimal response.

A number of workers (Nanda 1948; Cochran 1951; Tallis 1960; Harris 1961, 1963, 1964; Heidhues 1961; Williams 1962a,b; Sales and Hill 1976a,b; Hayes and Hill 1980; Tai 1986) have examined the errors using estimated covariance matrices to construct index weights, although the results are often extremely complicated even for two characters and are highly dependent on the particular experimental design used to estimate $\mathbf{G}$ and $\mathbf{P}$. Heidhues (1961)

suggests that one simple way to improve the accuracy of an estimated index is to remove variables that have a low genetic correlation with merit but which are highly correlated with other variables in the phenotypic selection index.

One situation where standard errors for predicted response are easily obtained is when the index parameters are estimated from a parent-offspring regression (Tallis 1960). Assuming the joint distribution of the vector of additive genetic and phenotypic values is multivariate normal, the regression of the vector of additive genetic values $\mathbf{g}$ on the vector of phenotypic values $\mathbf{z}$ is given by Equation 31.9a and can be written as $\mathbf{g} = \mathbf{c} + \mathbf{G}\mathbf{P}^{-1}\mathbf{z} + \mathbf{e}$ where $\mathbf{c}$ is a vector of constants and $\mathbf{e}$ the vector of errors associated with predicting $\mathbf{g}$ from $\mathbf{z}$. Premultiplying by $\mathbf{a}^T$ gives

$$H = \mathbf{a}^T\mathbf{g} = \mathbf{c}^* + \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{z} + \mathbf{e}^* \tag{33.30}$$

Note that Equation 33.30 still holds if $\mathbf{g}$ and $\mathbf{z}$ contain different traits, provided we use $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$. Thus the slope of the regression of offspring merit H on parental phenotypes $\mathbf{z}$ is $\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1} = \mathbf{b}_s^T$ and hence the Smith-Hazel index weights can be directly estimated from such a regression. Standard regression theory can then be used to place error bounds on $\mathbf{b}_s$ (using the standard errors of the slope estimates) and on the expected change in H.

A third class of errors, which occurs even if $\mathbf{P}$ and $\mathbf{G}$ are estimated exactly, is that selection changes the genetic covariance structure (Chapters 13, 24, 31) and hence changes the optimal weighting each generation. Equations 33.14 can be used to compute the change in $\mathbf{D}$ (and hence the changes in $\mathbf{P}$ and $\mathbf{G}$), but this assumes the infinitesiminal model.

A final source of error is that the selection intensity of truncation selection is overestimated in finite populations (Chapter 10), so comparisons of predicted versus actual response should use the empirical selection differential, rather than that expected given the fraction of selection.

Given all these potential sources of error, how well is the use of the Smith-Hazel index supported by experimental data? Table 33.1 summarizes Caballero's (1989) review of experiments from mice, *Drosophila melanogaster*, and *Tribolium castaneum*. The predicted response overestimates, often dramatically, the achieved response.

Table 33.1. Summary of 19 experiments in *Drosophila*, mice, and *Tribolium* examining the relative efficiency of selection on the estimated index, measured as the ratio of achieved to predicted response. Original index estimates refers to the expected response using estimates of the genetic and phenotypic covariances used in constructing the initial index, while improved parameter estimates refers to the predicted response based on covariances estimated from either a larger base population or from estimates during selection. From Caballero (1989).

	Expected response computed using:	
	Original index estimates	Improved parameter estimates
Single-generation response	67%	87%
	(range not available)	(range: 23% - 94%)
Multiple-generation response	37%	50%
	(range: 16% - 95%)	(range: 25% - 69%)

This table highlights two sources of errors that reduce the efficiency of the estimated index: those due to poor estimates and those due to changes in the genetic parameters as

selection proceeds. The data show, as expected, that poor estimates of population parameters results in a loss of efficiency. Achieved single-generation responses averaged 67% of that predicted based on an index constructed using the original parameter values. When improved parameter values were used, the achieved/predicted response ratio increased to an average value 87%. Table 33.1 also shows that changes in genetic parameters as selection proceeds are a significant source of error. Achieved response dropped from 67% to 37% and from 87% to 50% when the response is considered over multiple generations. Caballero (1989) presents evidence that this is due to changes in genetic variances and covariances during the course of selection. Another feature seen when replicated experiments are used is that while the *index* may show a reasonably consistent response over replicates, considerable variation is found in the response of *component traits* making up the index.

One especially interesting situation is **antagonistic index selection**, when the index weights on pairs of traits have opposite sign to the genetic correlations between those characters, e.g., $H = g_1 - g_2$ when $\sigma(A_1, A_2) > 0$ or $H = g_1 + g_2$ when $\sigma(A_1, A_2) > 0$. The experimental results when antagonistic indices occur are mixed. For example, no response was observed for an antagonistic index based on early weight gain and adult weight in mice ($\rho_g \simeq 0.55$; von Bulter et al. 1980), while antagonistic indices based on litter size and body weight in mice ($\rho_g \simeq 0.6$; Eisen 1977a, 1978), plant height and number of leaves in tobacco ($\rho_g \simeq 0.7$; Matzinger et al. 1977), and pupal and adult weight in *Tribolium* ($\rho_g \simeq 0.9$; Campo et al. 1990) showed a reasonable response.

The Hayes-Hill Transformation: Detecting Flaws in the Estimated Index

Suppose that I and H contain the same traits, so that $\mathbf{G}$ is symmetric. One obvious sign that the estimated index is flawed is if $\widehat{\mathbf{G}}$ is not positive-definite and hence not a proper covariance matrix. There is a significant probability of this when sample size is small and/or the number of traits large (Hill and Thompson 1978). Even if $\widehat{\mathbf{G}}$ is positive-definite, it may be inconsistent with estimates of $\widehat{\mathbf{P}}$ as estimated heritabilities can exceed one. While simple inspection of $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{P}}$ may reveal obvious problems such as negative variances or correlations that exceed unity, others (such as partial correlations exceeding unity) can easily be overlooked. Hayes and Hill (1980) note that this problem can be avoided by considering the eigenvalues of $\mathbf{H} = \mathbf{P}^{-1}\mathbf{G}$. Their motivation is as follows. The canonical transformation (Appendix 4) transforms a vector of correlated variables into a new vector whose elements are uncorrelated. Let $\mathbf{U}$ be the matrix giving the canonical transformation of $\mathbf{H} = \mathbf{P}^{-1}\mathbf{G}$, e.g., $\mathbf{U} = (\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n)$ where $\mathbf{e}_i$ are the normalized eigenvectors of $\mathbf{H}$. Hayes and Hill show that the transformed phenotypic and additive genetic values

$$\mathbf{g}_U = \mathbf{U}\mathbf{g} \qquad \text{and} \qquad \mathbf{z}_U = \mathbf{U}\mathbf{z} \tag{33.31a}$$

have covariances matrices

$$\widetilde{\mathbf{P}} = \sigma(\mathbf{z}_U, \mathbf{z}_U) = \sigma(\mathbf{U}\mathbf{z}, \mathbf{U}\mathbf{z}) = \mathbf{U}\mathbf{P}\mathbf{U}^T = \mathbf{I} \tag{33.32b}$$

$$\mathbf{G}_U = \sigma(\mathbf{g}_U, \mathbf{g}_U) = \sigma(\mathbf{U}\mathbf{g}, \mathbf{U}\mathbf{g}) = \mathbf{U}\mathbf{G}\mathbf{U}^T = \boldsymbol{\Lambda} \tag{33.32c}$$

where $\mathbf{I}$ is the identity matrix and $\boldsymbol{\Lambda}$ a diagonal matrix whose diagonal elements are given by the eigenvalues of $\mathbf{H} = \mathbf{P}^{-1}\mathbf{G}$. Hence the transformed characters are uncorrelated and the eigenvalues of $\mathbf{H}$ corresponds to the heritabilities of the transformed characters (as each character as unit phenotypic variance). Under this transformation, the merit function can be written as $\mathbf{a}_U^T \mathbf{g}_U$ where $\mathbf{a}_U = \mathbf{U}^T\mathbf{a}$ is the vector of transformed economic weights. Substituting into Equation 33.19 the response to selection on the Smith-Hazel index can be expressed as

$$\frac{R}{\bar{i}} = \sqrt{\boldsymbol{\alpha}^T \mathbf{G}_U \mathbf{P}_U^{-1} \mathbf{G}_U \boldsymbol{\alpha}} = \sqrt{\boldsymbol{\alpha}^T \boldsymbol{\Lambda} \mathbf{I}^{-1} \boldsymbol{\Lambda} \boldsymbol{\alpha}} = \sqrt{\boldsymbol{\alpha}^T \boldsymbol{\Lambda}^2 \boldsymbol{\alpha}} = \sqrt{\sum_{i=1}^{n} \alpha_i^2 \lambda_I^2} \tag{33.32}$$

Hence for any vector of economic weights $\mathbf{a}$ and (non-singular) covariance matrices $\mathbf{P}$ and $\mathbf{G}$, the Hayes-Hill transformation considerably reduces these $n(n+2)$ parameters ($n(n+1)/2$ for both $\mathbf{P}$ and $\mathbf{G}$ and n for $\mathbf{a}$) to just $2n$ parameters (n transformed economic weights α_i and n heritabilities of the transformed variables λ_i).

In light of this, Hayes and Hill (1980) suggest that the eigenvalues of $\widehat{\mathbf{H}} = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}$ be examined, since these correspond to the heritabilities of the transformed variables and hence should be between zero and one if the estimates of $\widehat{\mathbf{P}}$ and $\widehat{\mathbf{G}}$ are well-behaved. If this is not the case, the estimated covariance matrices can be modified until the estimates are consistent. While one approach is to set negative variances to zero and heritabilities and correlations that exceed unity to unity, the methods of bending and rounding discussed below are preferred.

"Bending" and "Rounding" Corrections of the Estimated Index

Again assume that I and H contain the same traits, and hence $\mathbf{G}$ is symmetric. Hayes and Hill (1981), noting that estimates of eigenvalues tend to biased (with large eigenvalues overestimated and eigenvalues underestimated), suggest a **bending** procedure to improve the efficiency of the estimated index. Their idea is to increase small eigenvalues and decrease large ones while holding the average eigenvalue constant. This is done by computing the eigenvalues of the modified matrix

$$\widehat{\mathbf{H}}^{*} = (1-\gamma)\cdot\widehat{\mathbf{H}} + \gamma\cdot\overline{\lambda}\mathbf{I} \qquad \text{for} \quad 0 \le \gamma \le 1 \tag{33.33}$$

where γ is the bending factor, $\overline{\lambda} = n^{-1}\sum\lambda_i$ is the average eigenvalue of $\widehat{\mathbf{H}} = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}$, and $\mathbf{I}$ the identity matrix. No bending ($\gamma = 0$ and $\widehat{\mathbf{H}}^{*} = \widehat{\mathbf{H}}$) corresponds to the estimated Smith-Hazel index. With complete bending, $\gamma = 1$ giving $\widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}} = \overline{\lambda}\cdot\mathbf{I}$ or $\widehat{\mathbf{P}} = \overline{\lambda}\cdot\widehat{\mathbf{G}}$ and hence weights of

$$\mathbf{b} = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}\mathbf{a} = c\cdot\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{G}}\mathbf{a} = c\cdot\mathbf{a}$$

which recovers the base index. Thus γ can be viewed as scaling the data from the Smith-Hazel ($\gamma = 0$) to the base index ($\gamma = 1$). Figure 33.1 shows how the eigenvalues based on the covariance matrices used in Example 33.1 change during bending.

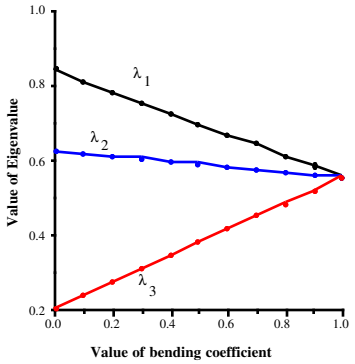

Figure 33.1. The effects of bending. Using the estimates of $\mathbf{P}$ and $\mathbf{G}$ from Example 33.1, the eigenvalues of $\widehat{\mathbf{H}} = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}$ are found to be $\lambda_1 = 0.84$, $\lambda_2 = 0.622$, and $\lambda_3 = 0.205$, for an average value of $\overline{\lambda} = 0.555$. The three eigenvalues for the "bent" matrix $\widehat{\mathbf{H}}^{*} =$

$(1-\gamma)\cdot\widehat{\mathbf{H}}+\gamma\cdot 0.555\cdot\mathbf{I}$ are plotted as a function of the bending coefficient γ. As γ increases towards one, the eigenvalues smoothly converge towards $\overline{\lambda}$.

Simulation studies by Hayes and Hill show bending always improves the estimated index, but the optimal bending factor depends on the unknown parameters. While this is obviously a problem, one common situation where the choice of bending parameter is fairly clear is when eigenvalues of $\widehat{\mathbf{H}}$ are either negative or exceed one. Suppose one eigenvalue is negative. In this case, $\widehat{\mathbf{H}}$ is bend until this eigenvalue increases to zero. Likewise, if an eigenvalue exceeds one, the matrix is bent until this eigenvalue decreases to one. Even if the eigenvalues of $\widehat{\mathbf{H}}$ are between one and zero, Hayes and Hill suggest the sample size alone can be used to obtain the optimal bending parameter, but the theory is not fully developed.

Tai (1989) suggests a different procedure, **rounding**, again based on the canonical transformation of $\widehat{\mathbf{H}}$. Let $\mathbf{U}=(\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_n)$ again denote the transformation matrix associated with $\widehat{\mathbf{H}}$ where $\mathbf{e}_i$ is the normalized eigenvector (of $\mathbf{H}$) corresponding to eigenvalue λ_i. Under this transformation, we can write the index $\widehat{I}_s=\widehat{\mathbf{b}}_s^T\mathbf{z}=\mathbf{d}^T\mathbf{y}$, where $\mathbf{d}=\mathbf{U}\widehat{\mathbf{b}}_s$ and $\mathbf{y}=\mathbf{U}^T\mathbf{z}$. Rounding assigns a vector of zeros for each eigenvector associated with a negative eigenvalue. For example, suppose eigenvalues one through $n-m$ are between zero and one, while the last m eigenvalues are negative. Rounding consists of using the index

$$I=\mathbf{d}^T\mathbf{y}_m \quad \text{where} \qquad \mathbf{y}_m=\widehat{\mathbf{U}}_m^T\mathbf{z} \qquad \text{with} \quad \widehat{\mathbf{U}}_m=(\mathbf{e}_1,\mathbf{e}_2,\cdots,\mathbf{e}_{n-m},\mathbf{0},\cdots,\mathbf{0}) \tag{33.34}$$

While this procedure may seem somewhat less ad hoc that bending, it does not correct the the bias caused by overestimation of large eigenvalues. Rather, it compounds this bias by retaining large eigenvalues while discarding small (and likely underestimated) ones.

Bending is related to the method of **ridge regression** (Hoerl and Kennard 1970, 1981), which was created to handle ill-conditioned design matrices in linear models, with the $(\mathbf{X}^T\mathbf{X})^{-1}$ term in an OLS estimator replaced by $(\mathbf{X}^T\mathbf{X}+\gamma\mathbf{I})^{-1}$. Both Saxton (1986) and Xu and Muir (1990) briefly examined ridge regression of selection indices using $(\mathbf{P}+\gamma\,\text{diag}[\mathbf{P}])^{-1}$ in place of $\mathbf{P}^{-1}$, where diag($\mathbf{P}$) denotes a diagonal matrix consisting of the diagonal elements of $\mathbf{P}$. While ridge regression seems to offer no advantages over bending, it could potentially be very useful in the more general case where $\mathbf{G}$ is not symmetric (and hence bending is not defined), to deal with ill-conditioned $\mathbf{P}$ matrices. However, such matrices can also be handled by simply removing a few of the most highly correlated variables, especially if these have low predictive power for the breeding value of merit.

Constraints on R and S Given a Specified Selection Intensity

In several places, we will need expressions for either all possible responses $\mathbf{R}$ or all selection differentials $\mathbf{S}$ given a specific selection intensity. These are expressed as quadratic products in $\mathbf{R}$ or $\mathbf{S}$. To obtain these, first recall Equation 33.5,

$$\mathbf{R}=\frac{\overline{i}}{\sigma_I}\mathbf{G}\mathbf{b}, \quad \text{implying} \quad \mathbf{b}=\frac{\sigma_I}{\overline{i}}\mathbf{G}^{-1}\mathbf{R} \tag{33.35a}$$

Equation 33.35a shows that our discussion is now restricted to the case where $\mathbf{G}^T=\mathbf{G}$, and hence $\mathbf{G}^{-1}$ is (potentially) defined. Since $\sigma_I^2=\mathbf{b}^T\mathbf{P}\mathbf{b}$ (Equation 33.2a), substituting in the above expression for $\mathbf{b}$ gives

$$\sigma_I^2=\left(\frac{\sigma_I}{\overline{i}}\right)^2\mathbf{R}^T\mathbf{G}^{-1}\mathbf{P}\mathbf{G}^{-1}\mathbf{R}$$

or

$$\bar{\imath}^2 = \mathbf{R}^T\mathbf{G}^{-1}\mathbf{P}\mathbf{G}^{-1}\mathbf{R} \tag{33.35b}$$

This is a quadratic product in terms of the response vector $\mathbf{R}$ and solutions describe a hyper-ellipsoid whose surface corresponds to all possible vectors $\mathbf{R}$ for a given selection intensity $\bar{\imath}$. Likewise, substituting $\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S}$ gives

$$\bar{\imath}^2 = \mathbf{S}^T\mathbf{P}^{-1}\mathbf{G}\mathbf{G}^{-1}\mathbf{P}\mathbf{G}^{-1}\mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \mathbf{S}^T\mathbf{P}^{-1}\mathbf{S} \tag{33.35c}$$

Equation 33.35c also describes a hyper-ellipsoid, now the surface corresponds to all possible vectors of selection differentials yielding the same selection intensity. Recall (Chapter 10) that for a normally-distributed trait, the selection intensity is entirely a function of the fraction p saved in truncation selection. These quadratic products thus correspond to the collection of all possible responses and selection intensities given a specified amount of truncation selection on an index.

RESTRICTED AND DESIRED-GAINS INDICES

While the Smith-Hazel index results in the largest response in a linear combination of characters, often we have different objectives in mind and hence different indices may be required to achieve these. For example, we may wish the largest possible response in some linear combinations of characters while ensuring that another set of characters remains unchanged. This is done by constructing a **restricted index**. Likewise, instead of maximizing the response in the (scalar) merit value, we may instead wish to find a linear combination of characters that gives a prespecified response in the *vector* of characters means. Using such a **desired-gains** index gives us more control over the individual responses in each character. Gibson and Kennedy (1990) forcefully argue that the use of constrained indices (such as restricted and desired-gains) comes at such a cost in terms of decreased response in merit relative to a Smith-Hazel index (e.g., Example 33.8), that they should only be used in highly specialized situations, and not in more general settings where strict economic gains are of concern.

Restricted Indices

Suppose we are interested in the response of k characters, but that characters one to m are at their optimum values and we desire these to remain unchanged. Subject to this constraint we wish to maximize the response of some linear combination $\sum_{i=m+1}^{k} a_i g_i$ of the remaining characters. By defining the vector of weights as $\mathbf{a}^T = (0, \cdots 0, a_{m+1}, \cdots, a_k)$ the index to optimize under the constraint can be written as $\mathbf{a}^T\mathbf{z}$. This problem was first considered by Morely (1955) with a general solution developed by Kempthorne and Nordskog (1959). Since response to selection on an index is proportional to $\mathbf{G}\mathbf{b}$ (Equation 33.5), our constraint of no response can be restated as the first m elements of the vector resulting from this matrix product are zero. In matrix form,

$$\mathbf{C}\mathbf{G}\mathbf{b}_r = \mathbf{0}$$

where $\mathbf{C}$ is an $m \times k$ matrix with ones on the diagonal and all other elements zero. The method of Lagrange multipliers (Example 33.3, Appendix 5) can be used to find the index giving the maximal response in H for a fixed amount of selection subject to this constraint. Using this approach, Kempthorne and Nordskog obtained the vector of weights for the restricted index as

$$\mathbf{b}_r = \left[\mathbf{I} - \mathbf{P}^{-1}\mathbf{G}_r^T\left(\mathbf{G}_r\mathbf{P}^{-1}\mathbf{G}_r^T\right)^{-1}\mathbf{G}_r\right]\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} \tag{33.36a}$$

where $\mathbf{I}$ is the identity matrix and

$$\mathbf{G}_r = \mathbf{CG} \tag{33.36b}$$

Equation 33.36a allows for different traits in I and H, with $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$ as above. We refer to this restriction of no change in m specified characters as the **Kempthorne-Nordskog restriction**. Akbar et al. (1984) extended these results to allow for different traits in the index I and merit H, while Lin (1985) presents a derivation that does not use Lagrange multipliers. For two characters where the goal is to maximize response in character one with no response in character two, the resulting index (after rescaling) becomes

$$I_r = z_1 - \left(\frac{\sigma_{g_1,g_2}}{\sigma^2_{g_2}} \right) \cdot z_2 \tag{33.36c}$$

as obtained by Morely (1955). Note that we have seem similar restrictions in Chapter 30 when the notion of conditional genetic variance and conditional evolvability were examined, and both of these concepts are closely connected to restricted indices.

Example 33.8. Consider the soybean data from Example 33.1. Suppose that character z_1 (oil content) is at its optimal value, but we wish to optimize the sum of protein content and yield, $H = g_2 + g_3$. Here,

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{a} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Applying Equation 33.36a gives

$$\mathbf{b}_r = \begin{pmatrix} -3.3 \\ -0.8 \\ 1.1 \end{pmatrix}$$

Hence $\sigma(I_r) = \sqrt{\mathbf{b}_r^T \mathbf{P} \mathbf{b}_r} \simeq 23.2$ and from Equation 33.5 the expected change in the vector of character means is

$$\mathbf{R} = \frac{\overline{i}}{\sigma(I_r)} \cdot \mathbf{G}\mathbf{b}_r = \overline{i} \cdot \begin{pmatrix} 0 \\ 4.0 \\ 11.2 \end{pmatrix}$$

I The resulting response in merit becomes

$$H = \overline{i}\,(1 \cdot 4.0 + 1 \cdot 11.2) = \overline{i} \cdot 15.2$$

If instead of restricting the change in character one, the Smith-Hazel index is applied with $\mathbf{a}$ as above so that no weight is placed on changes in z_1 (i.e., we do not care what values they take), then $\mathbf{a}$ is as above and

$$\mathbf{b}_s = \mathbf{G}\mathbf{P}^{-1}\mathbf{a} = \begin{pmatrix} 0.38 \\ 0.16 \\ 1.17 \end{pmatrix} \qquad \text{giving} \qquad \mathbf{R} = \frac{\overline{i}}{\sigma(I_s)} \cdot \mathbf{G}\mathbf{b}_s = \overline{i} \cdot \begin{pmatrix} 6.4 \\ 9.2 \\ 27.3 \end{pmatrix}$$

The response in H is now $36.5 \cdot \overline{i}$ under the Smith-Hazel index, as compared with only $15.2 \cdot \overline{i}$ under the restricted index. The price paid for no change in character one is that the expected response in H is only 42 percent of that under no restrictions. This result is fairly typical, in that using a restricted index often results in a rather significant decrease in the expected merit.

A variety of restricted indices handling different classes of constraints have been proposed (e.g., Tallis 1962, 1985, 1986; James 1968; Cunningham et al. 1970; Harville 1975; Niebel and Van Vleck 1982). For example, Tallis (1962) considers the case where a specified response is desired in k linear combinations of characters,

$$L_1 = \sum_{j=1}^{k} c_{j1} R_j = d_1, \quad L_2 = \sum_{j=1}^{k} c_{j2} R_j = d_2, \quad \cdots, \quad L_k = \sum_{j=1}^{k} c_{jk} R_j = d_k \tag{33.37}$$

where R_j is the response in trait j. Here the constraint is $\mathbf{CR} = \mathbf{CGb}_r = \mathbf{d}$, where the $m \times k$ matrix $\mathbf{C}$ has its ij-th element given by c_{ij} and $\mathbf{d}$ is the vector $(d_1, \cdots, d_m)^T$. This **Tallis restriction** is a more general from of the Kempthorne-Nordskog restriction. Again using the method of Lagrange multipliers, response is maximized subject to these constraints by selecting on the index with weights

$$\mathbf{b}_r = \left[\mathbf{I} - \mathbf{P}^{-1}\mathbf{G}_r^T \left(\mathbf{G}_r\mathbf{P}^{-1}\mathbf{G}_r^T\right)^{-1} \mathbf{G}_r\right] \mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} + \mathbf{P}^{-1}\mathbf{G}_r^T \left(\mathbf{G}_r\mathbf{P}^{-1}\mathbf{G}_r^T\right)^{-1} \mathbf{d} \tag{33.38}$$

where $\mathbf{G}_r$ is given by Equation 33.36b (now with $\mathbf{C}$ given as defined above) and hence allows for different traits in I and H. Additional features of restricted indices are reviewed by Brascamp (1984), sampling properties have been considered by Hill and Meyer (1984), and Itoh and Yamada (1987) show the equivalence of several of these indices. Other reviews can be found in Mallard (1972), who gives a geometric interpretation of restricted indices and Harville (1974). Famula (1992) suggests an alternative approach for constraining response in correlated characters based on linear programming. The advantage with this approach is that it gives a smaller mean-squared deviation of the constrained trait from zero relative to restricted index selection, but at a cost of a smaller response in the unconstrained characters. Thus if reducing response is deemed more important that maximal response in the unconstrained characters, linear programming methods should be considered.

Desired-gains Indices

Here the objective is to find the linear index $I_d = \mathbf{b}_d^T\mathbf{z}$ giving a prespecified vector of proportional responses in each character. Besides providing control over how each individual character changes, the desired-gains index does not require specification of economic weights $\mathbf{a}$. It does, however, still require estimates of $\mathbf{P}$ and $\mathbf{G}$ (if these are not available and one is still reluctant to assign values for $\mathbf{a}$, then the Elston index might be considered). Let $\boldsymbol{\Delta}_d$ denote the vector of desired changes, so that the ratio of any two elements, Δ_i/Δ_j is the desired ratio of response in characters i and j. From Equation 33.5, selection on $I = \mathbf{b}^T\mathbf{z}$ gives a vector of response proportional to $\mathbf{Gb}$. Hence $\mathbf{Gb}_d = \boldsymbol{\Delta}_d$, giving the vector of weights for the **desired-gains index** as

$$\mathbf{b}_d = \mathbf{G}^{-1}\boldsymbol{\Delta}_d \tag{33.39}$$

as obtained by Pešek and Baker (1969) and Yamada et al. (1975). In an evolutionary context, this is the same as the realized selection gradient (Chapter 30). If $\mathbf{C}$ in Equation 33.38 is of full rank (n linearly independent restrictions are imposed) then the responses of all characters are completely specified and the Tallis-restricted index reduces to the desired-gain index.

Equation 33.39 assumes that desired response is specified for all n measured traits (i.e., I and H contain the same traits). More generally, if I and H contains different traits, the (as above) $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$. Solutions are still of the form $\mathbf{Gb}_d = \boldsymbol{\Delta}_d$, but this does not yield a unique solution for $\mathbf{b}_d$ as $\mathbf{G}$ may not be square and hence $\mathbf{G}^{-1}$ may not be defined. A unique solution can imposing the additional constraint that selection intensity $\mathbf{b}^T\mathbf{Pb}$ (see Equation 33.35c) is minimized (Itoh and Yamada 1985), giving

$$\mathbf{b}_d = \mathbf{P}^{-1}\mathbf{G}^T(\mathbf{GP}^{-1}\mathbf{G}^T)^{-1}\boldsymbol{\Delta}_d \tag{33.40}$$

This was first obtained by Tai (1977), although in a much more cryptic form. As expected, this reduces to Equation 33.39 when $\mathbf{G}^T = \mathbf{G}$. Itoh and Yamada (1988b) have further modified the desired-gains index to allow for restrictions in the response of specified characters.

Example 33.9. Once again using the soybean data from Example 33.1, suppose we wish to increase $(z_1, z_2, z_3) =$ (oil content, protein content, yield) by relative amounts $(1 : 1 : 1)$ giving the vector of desired gains as

$$\boldsymbol{\Delta}_d = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Using Equation 33.39, and rescaling $\mathbf{b}_d$ so that its first element is one gives

$$\mathbf{b}_d = \begin{pmatrix} 1.0 \\ 0.8 \\ -0.5 \end{pmatrix}$$

Selection on the index $I_d = z_1 + 0.8 \cdot z_2 - 0.5 \cdot z_3$ thus gives the same response in each character. To verify this, first note that $\sigma(I_d) = \sqrt{\mathbf{b}_d^T \mathbf{P} \mathbf{b}_d} \simeq 5.45$, giving

$$\mathbf{R} = \left(\frac{\overline{i}}{\sigma(I_d)} \right) \cdot \mathbf{G}\mathbf{b}_d = 2.22\,\overline{i} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

To distinguish between a desired-gains and Smith-Hazel index, compare this response with that expected from the Smith-Hazel index on the merit function $z_1 + z_2 + z_3$. Under the desired-gains index, the response in merit is $\Delta\mu_1 + \Delta\mu_2 + \Delta\mu_3 = 6.66 \cdot \overline{i}$, only 13 percent of the expected response of $49.67 \cdot \overline{i}$ under the Smith-Hazel index (Example 33.7), illustrating the cost of specifying response in each character as opposed to just being concerned with the maximal response in merit. Now suppose that while all three characters are measured, we are only interested in having equal response in oil and protein content and are unconcerned with yield (z_3). Here

$$\boldsymbol{\Delta}_d = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{G} = \begin{pmatrix} 128.7 & 160.6 & 492.5 \\ 160.6 & 254.6 & 707.7 \end{pmatrix}$$

Applying Equation 33.40,

$$\mathbf{b}_d = \mathbf{P}^{-1}\mathbf{G}^T(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T)^{-1}\boldsymbol{\Delta}_d = \begin{pmatrix} 0.0131 \\ -0.0049 \\ 0.0002 \end{pmatrix}$$

This gives $\sigma^2(I_d) = \mathbf{b}_d^T \mathbf{P} \mathbf{b}_d = 0.01286$, so that the vector of responses is

$$\mathbf{R} = \left(\frac{\overline{i}}{\sigma(I_d)} \right) \cdot \mathbf{G}\mathbf{b}_d = 8.82\,\overline{i} \begin{pmatrix} 1.0 \\ 1.0 \\ 3.4 \end{pmatrix}$$

giving the change in the merit as $\overline{i} \cdot 8.82 \cdot (1 + 1 + 3.4) = 47.63 \cdot \overline{i}$, 96 percent of the response under the Smith-Hazel index.

While it is usually assumed that errors from incorrect estimates have more serious consequences for restricted indices than for unrestricted (Smith-Hazel) indices, the evidence is mixed. Hill and Meyer (1984) extended the Hayes-Hill transformation to restricted indices but, as with unrestricted indices, general statements of the effects of incorrect estimates are difficult to obtain. A second source of error, changes in genetic parameters as selection proceeds, has been examined by Mortimer and James (1987), who found for a four-locus model that the restricted index is particularly sensitive to changes in genetic parameters. If the assumptions of the infinitesimal model holds, changes in the covariance matrices under restricted index selection can be computed using Equation 33.14. By analogy with univariate results under directional selection (Chapter 13), the changes in disequilibrium mainly occur over the first few generations, after which they settle on their equilibrium values.

Experimental Results for Restricted and Desired Gains Indices

Certainly any error in parameter estimates or changes in genetic parameters results in some change in the restricted character, but the efficiencies of restricted indices (measured by actual response in the index versus predicted response) appear similar to those for unrestricted indices (Caballero 1989). Table 33.2 summarizes the results of several experiments on restricted indices. The general conclusions are that the observed response is almost always less (and often considerably so) than the predicted response, and the response under a restricted index is usually less than the response from direct selection on the character of interest, as expected. While significant responses in constrained characters are common, they are usually less (and often considerably so) than the correlated response that occurs when using an unconstrained index. Hence estimated restricted indices are effective at reducing, but not necessarily eliminating, undesirable correlated responses. Likewise, significant change in genetic parameters can occur. For example, Matzinger et al. (1989) performed a restricted selection experiment in tobacco to increase response in total alkaloids (TA) while constraining the response in yield (expected to decrease as a correlated response under unconstrained selection as the additive genetic correlation between these two characters is -0.7). After three cycles of selection, genetic variance in yield was unchanged and additive genetic variance in TA decreased by 60% from its initial value while the genetic correlation was reduced to -0.3.

Table 33.2. Results of representative experiments examining restricted indices.

Garwood and Lowe 1978 3 egg production traits in poultry.	Excellent fit. Nonsignificant response in constrained character while the response of the index of unconstrained characters was as predicted.
Abplanalp et al. 1963 8- and 24-week weight in turkeys.	Poor fit. Response in unconstrained character three times that expected. Negative response in the constrained character.
Scheinberg et al. 1967 Five traits in *Tribolium castaneum*.	Poor fit. Response in unconstrained character far less than expected and constrained characters displayed a significant negative response.
Okada and Hardin 1967, 1970 Larval and adult weight in *Tribolium castaneum*.	Modest fit. Response in unconstrained character far less than expected. Constrained character displayed a negative response, however this response under the restricted index was much smaller

	than the correlated response from direct selection on the unconstrained character. Index response asymmetrical.
Campo and Villanueva 1987 Adult and pupal weight in *Tribolium castaneum*.	Modest fit. Nonsignificant response in constrained characters in two separate sets of experiments, response in unconstrained characters less than half expected value.
Campo and Velasco 1989 Adult and pupal weight in *Tribolium castaneum*.	Good fit using both the Tallis restriction and desired-gain index in spite of very high genetic and phenotypic correlations (both $\rho > 0.9$).
McCarthy and Doolittle 1977 5- and 10-week body weight in mice.	Modest fit. Four different restriction indices applied to two highly correlated characters ($\rho_g \simeq 0.9$) Only two of the four indices gave no response in the constrained character. Responses in the unconstrained characters below expectations.
Eisen 1977a,b Weight gain and feed intake in mice.	Modest-Good fit. Response in unconstrained character close to expected value. No response in the constrained character over first four generations of selection followed by positive correlated response.
Eisen 1992, Eisen et al. 1995 Body fat and body weight in mice.	Poor fit. Significant and asymmetric response in the constrained character.
Matzinger et al. 1989 Total alkaloids (TA) and yield (Y) in tobacco.	Excellent fit. Significance response in TA while nonsignificant response in the constrained character Y despite strong negative genetic correlation ($\rho_g \simeq -0.7$). Response in unconstrained character matched predicted response.
Holbrook et al. 1989 Yield and seed protein in soybeans	Two cycles of restricted index selection resulted in significant increases in yield and total protein, no significant change in protein concentration (the restricted trait).
Atchley et al. 1997 Early vs. late growth in mice	14 generations of restricted selection for increases/decreases in early growth while restricting late growth, and vice versa. Good response in desired trait, little response in restricted trait.

SUMMARY OF LINEAR SELECTION INDICES

A variety of linear indices, which often have very different goals, have been introduced in the last few sections. Table 33.3 reviews the salient features of these.

Table 33.3. Summary of the linear selection indices introduced in this chapter. Response in merit $H = \mathbf{a}^T\mathbf{g}$ is of interest, while selection occurs on the index $I = \mathbf{b}^T\mathbf{z}$. Here $\mathbf{a}$ is the vector of weights for merit and $\mathbf{P}$ the phenotypic covariance matrix. We allow for H and I to include different traits and hence $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$ is not symmetric can be be nonsquare. When the same traits are in both H and I, $\mathbf{G} = \mathbf{G}^T = \sigma(\mathbf{g}, \mathbf{g})$ is the standard genetic covariance matrix. Unless mentioned otherwise, we assume the more general from of $\mathbf{G}$, which has the symmetric genetic variance as a special case.

Smith-Hazel index, $I_s = \mathbf{b}_s^T\mathbf{z}$, where $\mathbf{b}_s = \mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}$.

Selection on I_s maximizes the expected response in $\mathbf{a}^T\mathbf{g}$.

Estimated index, $\widehat{I}_s = \widehat{\mathbf{b}}_s^T \mathbf{z}$, where $\widehat{\mathbf{b}}_s = \widehat{\mathbf{P}}^{-1}\widehat{\mathbf{G}}^T\mathbf{a}$.

The estimate of the Smith-Hazel index using the sample phenotypic and additive genetic covariance matrices.

Base index, $I_b = \mathbf{a}^T\mathbf{z}$.

Suggested as an alternative to the Smith-Hazel index when confidence in the precision of estimates of $\mathbf{P}$ and (especially) $\mathbf{G}$ is low. Only defined when I and H contain the same traits

Restricted indices:

The Kempthorne-Nordskog restriction, $I_r = \mathbf{b}_r^T\mathbf{z}$. Defining $\mathbf{G}_r = \mathbf{CG}$,

$$\mathbf{b}_r = \left[\mathbf{I} - \mathbf{P}^{-1}\mathbf{G}_r^T\left(\mathbf{G}_r\mathbf{P}^{-1}\mathbf{G}_r^T\right)^{-1}\mathbf{G}_r\right]\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}$$

Selection on I_r maximizes the response in $\mathbf{a}^T\mathbf{g}$ for $m+1 \leq z_i \leq k$ while allowing no response in characters $z_1, \cdots, z_m$. $\mathbf{C}$ is an $m \times n$ matrix with ones on the diagonal and all other elements zero.

The Tallis restriction, $I_r = \mathbf{b}_r^T\mathbf{z}$. Defining $\mathbf{G}_r = \mathbf{CG}$,

$$\mathbf{b}_r = \left[\mathbf{I} - \mathbf{P}^{-1}\mathbf{G}_r^T\left(\mathbf{G}_r\mathbf{P}^{-1}\mathbf{G}_r^T\right)^{-1}\mathbf{G}_r\right]\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a} + \mathbf{P}^{-1}\mathbf{G}_r^T\left(\mathbf{G}_r\mathbf{P}^{-1}\mathbf{G}_r^T\right)^{-1}\mathbf{d}$$

Selection on I_r maximizes the response in $\mathbf{a}^T\mathbf{g}$ for $m+1 \leq z_i \leq k$ while specifying the response of m linear constraints of $\mathbf{z}$, such that $\mathbf{CR} = \mathbf{d}$, where $\mathbf{R}$ is the vector of responses and the elements of $\mathbf{C}$ specify the linear combinations. The Kempthorne-Nordskog restriction follows as a special case ($\mathbf{d} = 0$).

Desired-gains indices:

Pešek-Baker index, $I_d = \mathbf{b}_d^T\mathbf{z}$, where $\mathbf{b}_d = \mathbf{G}^{-1}\boldsymbol{\Delta}_d$.

Selection on I_d gives a vector of proportional responses $\boldsymbol{\Delta}_d$, where the ratio of the i-th and j-th elements of this vector gives the ratio of desired responses in these characters. The assumption is that response in all characters in $\mathbf{z}$ is of interest.

Tai-Itoh-Yamada index, $I_d = \mathbf{b}_d^T\mathbf{z}$, where $\mathbf{b}_d = \mathbf{P}^{-1}\mathbf{G}^T(\mathbf{GP}^{-1}\mathbf{G}^T)^{-1}\boldsymbol{\Delta}_d$.

Generalization of the Pešek-Baker index to allow for H and I containing different triats.

Cross-Generational Retrospective index, $I_{rt} = \mathbf{b}_{rt}^T\mathbf{z}$, where $\mathbf{b}_{rt} = \mathbf{G}^{-1}\mathbf{R}$

The difference from the Pešek-Baker index is one of interpretation. Here, the vector of responses $\mathbf{R}$ is *observed* rather than desired and we construct an index that would have accounted for this response. Again assumes H and I contain the same traits. If H and I contain different traits, the Tai-Itoh-Yamada index is used with $\mathbf{R}$ replacing $\boldsymbol{\Delta}_d$.

Within-Generational Retrospective index, $I_{ret} = \mathbf{b}_{ret}^T\mathbf{z}$, where $\mathbf{b}_{ret} = \mathbf{P}^{-1}\mathbf{S}$

Estimation of the coefficients of the index from the observed selection differentials

NONLINEAR SELECTION INDICES

In many cases, merit functions naturally depart from linearity (ratios $H(g_1, g_2) = g_1/g_2$ and products $H(\mathbf{g}) = \prod g_j$ being two examples). Dealing with nonlinearity introduces a number of subtleties, and we consider these first. We then examine the quadratic index as a further introduction into some of these issues. We conclude with a review of various proposed strategies for finding the best *linear* index for approximating a nonlinear merit. In many cases these are related to the Smith-Hazel index, but the weights can change each generation and they can depend on the current population mean, the proposed selection intensity $\bar{\imath}$, and the timeline over which we wish to maximize response.

Specific Issues With Nonlinear Indices

Care is required in considering the improvement goals when using a nonlinear merit function, as apparently subtle differences in the desired outcomes can become critical (Moav and Hill 1966, Goddard 1983, Itoh and Yamada 1988a, Burdon 1990). To see this point, first note that with a linear index, the mean of the index equals the index evaluated at the mean, e.g.,

$$E[H(\mathbf{g})] = \mathbf{a}^T\boldsymbol{\mu} = H(\boldsymbol{\mu}) = H(E[\mathbf{g}])$$

With a nonlinear index, $E[H(\mathbf{g})] \neq H(E[\mathbf{g}])$, and thus we must decide if our the goal to improve the *average merit of all individuals* in the population $E[H]$ or to improve the *merit of the population average* $H(E[z]) = H(\boldsymbol{\mu})$. These goals are equivalent under a linear merit function (as $E[H] = H[\boldsymbol{\mu}]$), but are generally different when nonlinear merit functions are used. For example, suppose $H = g^2$ so that $E[\,g^2\,] = \sigma_g^2 + \mu^2 \geq \mu^2 = (\,E[g]\,)^2$. Goddard (1983) notes that selection response theory predicts changes in means, so that it is convenient to therefore define merit as a function of the mean vector of the component traits, and hence the goal is to maximize merit as a function of the population mean, $H(\boldsymbol{\mu})$, and much of the existing theory has focused on this case. However, Itoh and Yamada (1988) note that in many settings it is more desirable to maximize the overall mean of the merit $E(H)$.

A related concern is whether we wish to maximize the additive genetic value in merit in the parents or in their offspring. Again, with a linear index these are equivalent, as the mean breeding value of the parents $\boldsymbol{\mu}$ equals the mean value in their offspring. This is not the case with nonlinear merit. To see this, again consider the simple merit function $H = g^2$. To further simplify matters, assume that the population mean is initially zero and that phenotypic and additive-genetic values are symmetrically distributed around the mean. Choosing adults with the largest g^2 values generates a group of parents with a mean g value of zero. Mating these parents at random gives offspring with a mean breeding value of zero. While there will be a transient increase in merit due to the transient increase in variance caused by the generation of positive gametic-phase disequilibrium (this merit function is akin to disruptive selection, Chapter 13), this increase decays away after selection stops. A much larger (and permanent) change in the merit occurs by just selecting the largest individuals (select to increase g rather than g^2), which increases the mean and hence increases g^2. Nonrandom mating among selected individuals can also increase response when the merit function is nonlinear (Allaire 1980). In this example, positive assortative mating among selected individuals will further increase response by generating additional positive disequilibrium, although this decays away upon relaxation of selection (Chapter 13).

A final issue, related to the first two, is that essentially all of the theory focuses on changes in means. As we have seen, selection introduces at least transient changes in the covariance matrices $\mathbf{G}$ and $\mathbf{P}$ through the generation of linkage disequilibrium, and these

second-order changes can potentially be quite important in nonlinear indices. Balancing this is the problem that changes in variances (and especially covariances) are extremely hard to predict and tend to be rather unstable.

Quadratic Indices

The quadratic selection provides an excellent introduction into nonlinear indices for several reasons. First, it is the simplest such index, with linearity as a special case. Second, it naturally arises in both breeding and evolutionary biology (recall the best quadratic fit of the fitness surface, Chapter 29). Third, it allows us to highlight some of the subtle issues with nonlinear indices just discussed. Finally, it provides a good introduction to linearization of non-linear indices by using a Smith-Hazel index whose weights change as the mean changes.

Wilton et al. (1968) introduced the quadratic index, and we largely follow their treatment. Consider a quadratic merit function, where H is of the form

$$\begin{aligned} H &= c + \sum_{i=1}^{k} a_i(\mu_i + g_i) + \sum_{i=1}^{k} a_{ii}(\mu_i + g_i)^2 + \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}(\mu_i + g_i)(\mu_j + g_j) \\ &= c + \mathbf{a}^T(\boldsymbol{\mu} + \mathbf{g}) + (\boldsymbol{\mu} + \mathbf{g})^T \mathbf{A} (\boldsymbol{\mu} + \mathbf{g}) \end{aligned} \tag{33.41a}$$

where c is a constant (which we will now ignore, as it does not effect the relative rankings of individuals with different $\mathbf{g}$ values) and $\mathbf{A}$ is a matrix of quadratic weights

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12}/2 & \cdots & a_{1k}/2 \\ a_{12}/2 & a_{22} & \cdots & a_{2k}/2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1k}/2 & a_{2k}/2 & \cdots & a_{kk} \end{pmatrix} \tag{33.41b}$$

Note that we have essentially seen this before, namely Lande-Arnold fitness regression (Chapter 29), where H was relative fitness. The critical different between a linear and a quadratic index is that the relative ranking of individuals is now a function of the population mean. To see this, consider a very simple quadratic index,

$$H = a_1(\mu_1 + g_1) + a_{22}(\mu_2 + g_2)^2$$

when $g_2 = -\mu_2$, there is no contribution from a_{22}, while when g_2 is significantly different from the population mean, being squared, the quadratic term can dominate. Thus an individual with the same g_2 value can have very different merit, depending on the current population mean. Corresponding quadratic phenotypic selection indices for this quadratic merit function have the form

$$I = \alpha + \mathbf{b}^T\mathbf{z} + \mathbf{z}^T\mathbf{B}\mathbf{z} \tag{33.42a}$$

where the matrix $\mathbf{B}$ has the same form as $\mathbf{A}$ except that a_{ij} is replaced by b_{ij}. As with a linear index, it is possible that $\mathbf{z}$ and $\mathbf{g}$ contain different traits and may have different dimensions (n and k, repsectively). As above, define $\mathbf{G} = \sigma(\mathbf{g}, \mathbf{z})$, a $k \times n$ matrix where $G_{ij} = \sigma(g_i, z_j)$. If both $\mathbf{g}$ and $\mathbf{z}$ contain the same traits, this is just the standard $n \times n$ genetic covariance matrix. Wilton et al show that the values of $\mathbf{b}$ and $\mathbf{B}$ that maximize the correlation between I and H are

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}^T(\mathbf{a} + 2\mathbf{A}\boldsymbol{\mu}) \tag{33.42b}$$

and

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{G}^T\mathbf{A}\mathbf{G}\mathbf{P}^{-1} \tag{33.42c}$$

Notice that the vector $\mathbf{b}$ of linear weights has two components: a constant value $\mathbf{P}^{-1}\mathbf{G}^T\mathbf{a}$ that is the same as the Smith-Hazel weight for a linear index with weight $\mathbf{a}$, and a second component $2\mathbf{P}^{-1}\mathbf{G}^T\mathbf{A}\boldsymbol{\mu}$ that is a function of the current population mean μ and the quadratic merit weight $\mathbf{A}$. Thus, the linear weights change with the mean. In constrast, the matrix of quadratic weights $\mathbf{B}$ remains constant.

Example 33.10. Wilton et al. (1968) present the following example of a quadratic index. The value of an angus cattle at weaning is a function of its weaning weight z_1 and type score (z_2). The net return (in 1963) is \$0.111 per pound of weaning weight at a zero type score, and an extra \$0.0049 per pound for each one point increase in the type score, giving

$$H = (\mu_1 + g_1)\,[0.111 + 0.0049(\mu_2 + g_2)]$$

This is a quadratic, being a function of the two direct breeding values g_1 and g_2, and their cross-product $g_1 g_2$. The resulting vector $\mathbf{a}$ and matrix $\mathbf{A}$ of merit weights becomes

$$\mathbf{a} = \begin{pmatrix} 0.111 \\ 0 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0 & 0.00245. \\ 0.00245. & 0 \end{pmatrix}$$

the off-diagonal term following from 0.11 · 0.0049 = 0.00245. The authors give the following covariance matrices

$$\mathbf{P} = \begin{pmatrix} 2649 & 18.49 \\ 18.49 & 1.75 \end{pmatrix}, \qquad \mathbf{G} = \begin{pmatrix} 1452 & 7.20 \\ 7.20 & 1.12 \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} 418.95 \\ 13.35 \end{pmatrix}$$

Noting that $\mathbf{G} = \mathbf{G}^T$ in this case and applying Equation 33.42b yield

$$\mathbf{b} = \mathbf{P}^{-1}\mathbf{G}\mathbf{a} + 2\mathbf{P}^{-1}\mathbf{G}\mathbf{A}\boldsymbol{\mu}$$
$$\begin{pmatrix} 0.622 \\ -0.201 \end{pmatrix} + \begin{pmatrix} -0.0000095 & 0.00274 \\ 0.00322 & -0.00887 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

Likewise from Equation 33.42c, we have

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{G}\mathbf{A}\mathbf{G}\mathbf{P}^{-1} = \begin{pmatrix} -0.0000052 & 0.0009151 \\ 0.0009151 & -0.005855 \end{pmatrix}$$

Notice that although the quadratic merit function put no weight on g_2, the resulting quadratic index does.

Analysis of the quadratic also allows us to consider some of the subtleties with a nonlinear index. The index weights given by Equation 33.42 correspond to optimizing the quadratic function evaluated at the mean, $H(\boldsymbol{\mu})$. Wilton et al. (1968) show that

$$E[H(\mathbf{g})] = H(\boldsymbol{\mu}) + \text{tr}(\mathbf{AP})$$

Thus, for the quadratic the difference between the mean value of the merit and the merit of the mean value is a constant, and hence their optimization is equivalent, *provided* $\mathbf{P}$ remains constant. Recall that tr($\mathbf{AP}$) denotes the trace (sum of the diagonal elements) of the matrix product $\mathbf{AP}$. Thus, if we focus only on changes in the mean, we do not have to worry as to

which specification (optimization of the mean of merit or the merit of the mean) is required. However, selection also changes $\mathbf{G}$, and hence $\mathbf{P}$, by creating disequilibrium. In this setting, the two are *not* equivalent.

The second issue, optimization of the merit in the parents or their offspring, also appears with the quadratic. While the vector of means is unchanged, the genetic variance differs from parent to offspring (due to segregating and decay of disequilibrium), resulting in the two having different values.

Finally, the quadratic allows us to start our discussion of how best to linearize a nonlinear index. Goddard (1983) shows that the maximal response in the merit of the mean can be obtained using a linear index. If the expected change in mean is small (as might occur with a small selection intensity), the Smith-Hazel weighs are given by the best linear approximation of the nonlinear index, namely the first-order Taylor series (see Appendix 5). This approach for a general nonlinear index was first suggested by Moav and Hill (1966) who further noted that it can break down when the change in mean is large, requiring other approaches to find the best linear index (to be discussed shortly). In either case, the best linear approximation changes as the mean (and other factors) change, and thus the Smith-Hazel weights are no longer a constant, but rather must be periodically updated. To see this for the quadratic, consider the merit function again,

$$H = c + \mathbf{a}^T(\boldsymbol{\mu} + \mathbf{g}) + (\boldsymbol{\mu} + \mathbf{g})^T \mathbf{A} (\boldsymbol{\mu} + \mathbf{g})$$

Let's expand H in a first-order Taylor series (Equation A5.6). Taking the vector of first derivatives of the linear term with respect to $\mathbf{g}$ gives

$$\nabla_{\mathbf{g}} \left[\mathbf{a}^T(\boldsymbol{\mu} + \mathbf{g}) \right] = \mathbf{a}$$

which follows from Equation A5.1a. Equation A5.1e gives the derivative for the quadratic term as

$$\nabla_{\mathbf{g}} \left[(\boldsymbol{\mu} + \mathbf{g})^T \mathbf{A} (\boldsymbol{\mu} + \mathbf{g}) \right] = 2\mathbf{A}(\boldsymbol{\mu} + \mathbf{g})$$

When evaluated $\mathbf{g} = 0$, the linearized contribution from the quadratic becomes $2\mathbf{A}\boldsymbol{\mu}$. Hence, the linear approximation of the merit function has the form

$$H(\mathbf{g}) \simeq H(\mathbf{0}) + \nabla_{\mathbf{g}}^T(H)\Big|_{\mathbf{g}=\mathbf{0}} \mathbf{g} = H(\mathbf{0}) + (\mathbf{a} + 2\mathbf{A}\boldsymbol{\mu})^T\mathbf{g} \tag{33.43}$$

and this are the best linear approximation of the weights. Thus the linear term of the quadratic index (Equation 33.42b) is just a Smith-Hazel index now using the best linear approximation for the function, Equation 33.43, as the weights. As the population mean changes under selection, so to do the weights. Note that the best linear approximation is a *local* one, and no longer holds if one considers large differences in $\mathbf{g}$ from around the mean. In such cases, the first-order Taylor series may no longer be a good approximation and other approaches are required to find the best linear weights (as detailed below).

Linear Indices for Nonlinear Merit

With these examples from the analysis of the quadratic in mind, lets now proceed to the general analysis of nonlinear indices. Changes in variance will be ignored, not because they are unimportant but rather because the relevant theory has not yet been fully developed.

While nonlinear indices for specific situations have been proposed (e.g., Wilton et al. 1968, Rønningen 1971, Magnussen 1991), Goddard (1983) suggested when genetic variation is completely additive that the largest response occurs by selecting on a linear, rather than

nonlinear, index. As first sight, this assertion does not seem well-supported by the experimental evidence. Campo and Rodrguez (1990) found selection on a nonlinear index gave a larger response than selection using a linear index when selecting for an increase in the ratio of egg mass to adult body weight in *Tribolium castaneum*. Fairfull et al. (1977) and Campo and de la Blanca (1988) also observed this with another nonlinear trait (total biomass = number of offspring $\times$ offspring weight) in *Tribolium castaneum*. These observations, however, do not necessarily invalidate Goddard's assertion, as there is still some uncertainly in these papers as how to construct the linear index giving the largest response in an nonlinear H. Many of the linear indices considered were somewhat arbitrary and perhaps not surprisingly were outperformed by a specialized non-linear index. Finally, the contribution of nonadditive genetic variation to response in nonlinear indices is an open issue.

How might the best linear index be obtained? The simplest situation is when the nonlinear index can be transformed into a linear index, which can then be maximized using the standard linear theory. For example, if the merit function is a simple polynomial, say $H(\mathbf{g}) = g_1 + g_2^2 + g_3$, by defining $\widetilde{g}_2 = g_2^2$ the index becomes linear, viz. $H(g) = g_1 + \widetilde{g}_2 + g_3$. This approach was first hinted at by Smith (1936) and formally suggested by Kempthorne and Nordskog (1959), but requires the phenotypic and additive genetic covariances of the transformed variables.

A more general approach is to first obtain the best linear approximation of a nonlinear merit using a first-order Taylor series about the population mean,

$$H(g_1, g_2, \cdots g_n) \simeq H(\boldsymbol{\mu}) + \sum_{i=1}^{n} a_i(g_i - \mu_i) \quad \text{where} \quad a_i = \left.\frac{\partial H}{\partial g_i}\right|_{\mathbf{g}=\boldsymbol{\mu}} \tag{33.44a}$$

This provides the (current) economic weights for the best linear approximation of the merit. The resulting optimal index is then given by a Smith-Hazel index using these weight (Moav and Hill 1966, Harris 1970), namely

$$\mathbf{a} = \nabla_{\mathbf{g}}[\, H \,]\,\big|_{\mathbf{g}=\boldsymbol{\mu}} \tag{33.44b}$$

Since $\mathbf{a}$ depends on the population mean, these weights change each generation. Moav and Hill (1966) and Goddard (1983) note that this approximation may be satisfactory when selection intensity is low, but is poor for highly non-linear functions when selection intensity is high (and hence the extremes of the nonlinear function are selected). Burton (1990) shows when the heritability of the characters underlying the index is low that a linear approximation is usually reasonable. Hence, if the heritability of the index obtaining by using Equation 33.44b is low and the selection intensity weak then the Taylor approximation is likely to be reasonable. However, if response is large (as might happen with either high heritability and/or strong selection), then the change in $\mathbf{g}$ may be sufficiently large that the linear Taylor approximation is no longer appropriate. Itoh and Yamada (1988) have developed second-order (quadratic) series approximation in these cases, but exact methods are also available, as we discuss shortly.

Example 33.11. Moav and Hill (1966) introduced the following nonlinear merit function:

$$H = a - bg_1 - \frac{c}{g_2}$$

Here,

$$\nabla_{\mathbf{g}}(H) = \begin{pmatrix} \partial H/\partial g_1 \\ \\ \partial H/\partial g_2 \end{pmatrix} = \begin{pmatrix} -b \\ \\ c/g_2^2 \end{pmatrix}$$

Hence $\nabla_{\mathbf{g}}^T(H)|_{\boldsymbol{\mu}} = (-b \quad c/\mu_2^2)^T$ giving the best linear approximation of H around the current population mean as

$$H \simeq a - bg_1 + \frac{c}{\mu_2} g_2$$

The resulting vector of weights for the Smith-Hazel index become

$$\mathbf{a} = \begin{pmatrix} -b \\ c/\mu_2^2 \end{pmatrix}$$

In the next generation, we would update with the new mean and proceed. In a real experiment, would estimate the new mean each generation and use this for the new value of μ_2. If our goal instead is to predict the expected response after some number of generations, then Equation 33.19 can be used to compute the change in the (linearized) merit, while Equation 33.20 predicts the change in the components (g_1, g_2) of the merit. Thus, in the next generation the weight on g_2 would be $c/(\mu_2 + R_2)^2$ where R_2 is obtained from Equation 33.20. Proceeding in this fashion, one could iterate the expected response out to any desired generation.

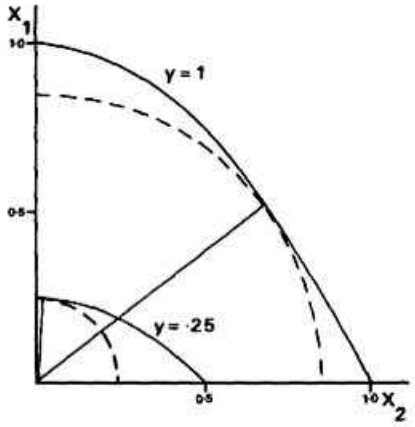

Figure 33.3. Two contours (solid lines) of the merit function $H = g_1 + g_2^2$ are plotted (corresponding to values of 0.25 and 1). The dashed lines are the response surfaces corresponding to two different intensities of selection. The intersection between the merit and response surfaces gives the optimal linear weights on both traits. Note that under weak selection, most of the weight is on g_1, while with stronger selection, there is roughly equal weight on g_1 and g_2. Thus, the optimal index weights are a function of the intensity of selection. After Goddard (1983).

Exact Optimization of Nonlinear Indices

While a Taylor series may provide a good linear approximation to a highly nonlinear merit function when the change in mean is small, it does a poor job when the change in mean is large. An alternative approach to obtaining the best linear index in this case is based on a graphical method suggested by Moav and Hill (1966) for two characters and holds even when selection intensity is strong and heritability high. Itoh and Yamada (1988a) analytically extend this approach to n traits using the method of Lagrange multipliers (Appendix 5, Example 33.3), while Pasternak and Weller (1993), Groen et al (1994), and Dekkers et al. (1995) offer further extensions based on optimal control theory. All of these methods start

from Equation 33.35b, which gives the quadratic response surface (all possible $\mathbf{R}$ vectors) that can be generated at the same selection intensity for a given $\mathbf{G}$ and $\mathbf{P}$,

$$\bar{\imath}^2 = \mathbf{R}^T\mathbf{G}^{-1}\mathbf{P}\mathbf{G}^{-1}\mathbf{R}$$

In the original graphical analysis by Moav and Hill, the response surface is plotted into a series of contours in the merit function, with the optimal response $\mathbf{R}$ in the merit components taken to be those values on the response surface giving the highest merit. Analytically, this is simply the maximal gain in merit $H(\boldsymbol{\mu}+\mathbf{R}) - H(\boldsymbol{\mu})$ subject to the constraints on $\mathbf{R}$ imposed by Equation 33.35b. With the optimal $\mathbf{R}$ in hand, the optimal index weights are given by Equation 33.35a,

$$\mathbf{b} = \frac{\sigma_I}{\bar{\imath}}\mathbf{G}^{-1}\mathbf{R}$$

Figure 33.3 shows an example, and makes the key point that, unlike the Taylor approximation, the optimal index weights are a function of the selection intensity. Goddard (1983) shows this method gives the optimal single-generation response when the breeding goal is to improve merit as a function of the population mean $H[\boldsymbol{\mu}]$.

Optimal Weights Depend on the Lenght of the Experiment

The two above approaches, Taylor series approximation and exact optimization, both apply to predicting a single generation of response. As the mean changes, so does the weight vector of the best linear index. Since $\mathbf{G}$ (and hence $\mathbf{P}$) also change under selection, the response surface used in exact optimization also changes over time (although it likely rapidly approaches an asymptotic value). These changes are generally not accounted for in the literature on nonlinear indices, but is straightforward to use Equations 33.14 to compute the change in $\mathbf{D}$ (and hence $\mathbf{G}$ and $\mathbf{P}$).

When the goal is optimization of the total response in merit over some defined interval, several approaches are possible. First, one could use either Taylor series or exact optimization to update the linear weights $\mathbf{b}$ each generation, resulting in using a new vector of weights each cycle of selection (changes in $\mathbf{G}$ and $\mathbf{P}$ can also be accounted for using this approach). The alternative is exact optimization, but now with the response surface corresponding to the total response by the end of the experiment. The idea is to find the best linear index corresponding to this response, and hence one would use a constant index throughout. Formally, for optimization of the total response after t generations, one would optimize $H(\boldsymbol{\mu} + t\mathbf{R}) - H(\boldsymbol{\mu})$ subject to the constraint of the response surface (Equation 33.35b) with $t\,\bar{\imath}$ replacing the selection intensity. Note, because of nonlinearity, the best *single* generation response $\mathbf{R}$ might not correspond to the best cumulative response after t generations. One advantage of this approach is that a constant index is used (the weights $\mathbf{b}$ do not change). One disadvantage is that it does not account for changes in $\mathbf{G}$ and $\mathbf{P}$. Groen et al. (1994) examined these different approaches, and found that direct optimization had a very slight advantage over methods that updated each generation. Dekkers et al. (1995) expand this problem, using results from optimal control theory when more complex types of optimization (such as the average merit through the experiment) are of interest.

SEQUENTIAL APPROACHES: TANDEM SELECTION AND INDEPENDENT CULLIING

When the goal is to improve performance in a number of traits, one can either select them simultaneously (as we have seen with index selection) or *sequentially*. We conclude this chapter by considering such sequential methods, which could involve selecting different traits in

different generations (**tandem selection**), sequentially selecting traits over a single interval (**independent culling**), or selecting different traits at different times during an individual's life span (**multistage selection**). Throughout, the goal is to maximize $H = \sum a_i g_i$ (or equivalently, the response in $\mathbf{a}^T\mathbf{z}$) for a fixed amount of selection $\overline{\imath}$ applied each generation. Under the assumptions of the multivariate breeder's equation, $\Delta H = \mathbf{a}^T\mathbf{R} = \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{S}$, so that the goal for each method is to obtain the vector of directional selection differentials $\mathbf{S}$ that maximizes ΔH for a fixed amount of selection. It is well established that index selection is optimal for this sort of problem. However, the researcher or breeder also has to worry about the costs (in terms of time and responses) of selection, and thus another metric is to consider the largest rate of *economic* return, defined as the response divided by the cost. Under this framework, index selection may not optimal, as one must score all of the traits in the individuals being tested, which can be quite resource-expensive, both in terms of cost as well as time.

Tandem Selection

Under tandem selection, only a single trait is selected each generation, but the trait chosen changes over time. Selection intensity is usually assumed to be constant over generations, as would occur if truncation selection is used with the same fraction saved, independent of the current trait under selection.

Suppose we seek improvement in n traits. Under tandem selection, in any particular generation only one trait is under direct selection, while other traits can change via correlated responses (Chapter 30). Suppose trait i is currently under direct selection. Assuming a constant selection intensity $\overline{\imath}$, its selection differential is $S_i = \overline{\imath}\,\sigma(z_i)$, giving a selection gradient of $\beta_i = S_i/\sigma^2(z_i) = \overline{\imath}/\sigma(z_i)$. From Equation 30.3a, the selection differential on a (phenotypically) correlated trait j is

$$S_j = \sigma(z_i, z_j)\beta_i = \overline{\imath}\,\frac{\sigma(z_i, z_j)}{\sigma(z_i)}$$

Summing over generations, the expected response to m generations of tandem selection is

$$\Delta H = \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{S}, \quad \text{where} \quad S_i = \overline{\imath}\sum_{j=1}^{n} m_j\,\frac{\sigma(z_i, z_j)}{\sigma(z_i)} \tag{33.45}$$

with m_j the number of generations individuals are chosen solely on character j. The optimal response is obtained by solving for the weights m_j given $\mathbf{P}$, $\mathbf{G}$, and $\mathbf{a}$.

For the special case of no phenotypic or genetic correlations,

$$\Delta H = \sum_{j=1}^{m} a_j\,h_j^2\,S_j = \overline{\imath}\sum_{j=1}^{m} m_j\,\theta_j \qquad \text{with} \qquad \theta_j = a_j\,h_j\,\sigma(g_j) \tag{33.46}$$

which follows by noting that $h_j S_j = h_j\,\overline{\imath}\sigma(z_j) = \overline{\imath}\sigma(g_j)$. Optimal response occurs by selecting the character with the initially largest value of θ_j in the first generation and the character with the largest value of θ_j in each subsequent generation. If the θ_j's remain unchanged as selection proceeds, the optimal strategy is to continue to select only on the character that gave the largest response in the first generation. This is also the optimum strategy for arbitrary $\mathbf{P}$ and $\mathbf{G}$, provided these remain unchanged (Turner and Young 1969). This strategy, however, is rarely used. The breeder usually changes the character being selected after some desired level of performance is reached (in effect, changing the economic weights $\mathbf{a}$ as additional response in the initial character loses some of its desirability relative to response in other characters).

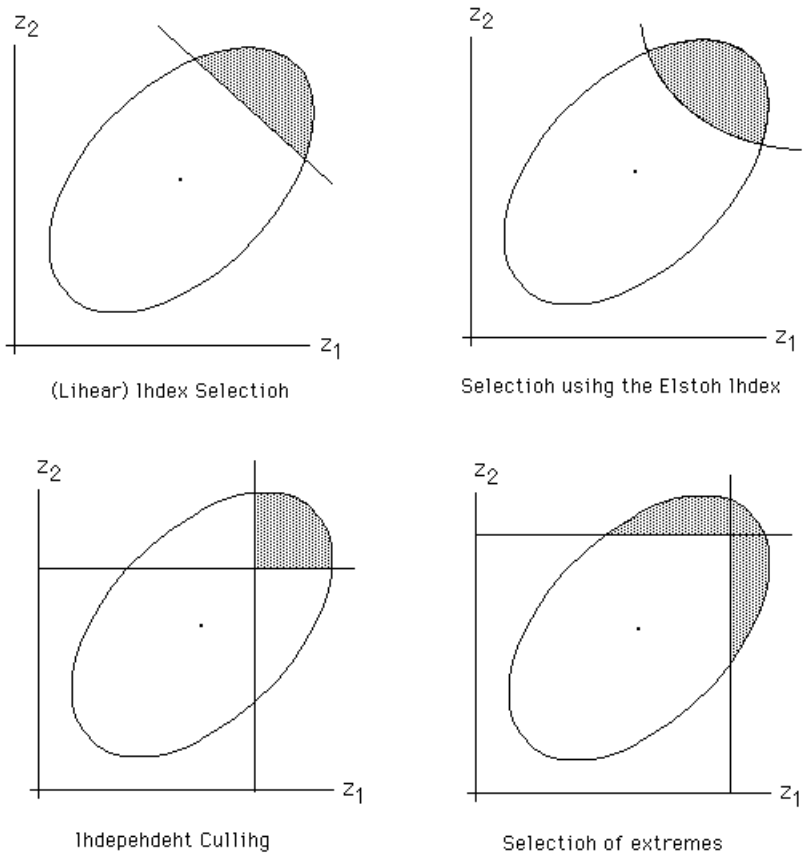

Figure 33.3. Graphical representation of truncation selection on two characters using different multivariate selection methods. The shaded area represents the fraction saved. **Upper left:** under index selection, all individuals above the line $b_1 z_1 + b_2 z_2$ are saved so that the values of the two characters are weighted. **Upper right:** under the Elston index, individuals whose values are above a quadratic curve are saved. **Lower left:** under independent culling an individual must be exceptional in both characters (*both* must exceed their thresholds) to be saved. **Lower right:** under selection of extremes the individual needs only be exceptional in a single character (*either* exceeding its threshold value is sufficient) to be saved.

A second issue is that $\mathbf{G}$ and $\mathbf{P}$ *do* change over time, and hence the optimal trait can change as well. Under the infintesimal framework, if there are no genetic or phenotypic correlations between traits, then selection on one trait reduces only its additive variance (by generating negative disequilibrium, Chapter 13) which can reduces its h^2 value, potentially making other traits more favorable. We can directly see this from Equation 33.13, as with no phenotypic correlations, the vector $\mathbf{b}$ of index weights contains only one non-zero entry (corresponding to the trait under selection), while $\mathbf{G}$ is diagonal. When selections stops on this trait, disequilibrium decays quickly, increasing h^2. Under this model, the optimal approach can be to select on one trait for a few generations, then move onto another to let the selection-generation disequilibrium in that trait decay to zero before restarting selection on it (Villanueva and Kennedy 1993).

An interesting variant of tandem selection is to select for different traits in different populations, and then cross them The F_1 is equivalent to the result from tandem selection,

provided the starting populations are the same. This is typically done with two populations (e.g., Orozco et al. 1980), but in theory is also applies to a synthetic formed by crossing n different populations, each improved for a different trait (Chapter 22). Again, the proviso is all started from the same base stock. The presence on nonadditive variance can also complicate issues if allele frequencies have diverged sufficiently for heterosis to appear (Chapter 22).

Independent Culling

After index selection, some variant of **independent culling** is the most popular approach for selecting on multiple traits. Under independent culling (Figure 33.2), individuals are first screened for one trait (or more generally for an index involving a subset of traits). Individuals above a given threshold level are saved for the next round selection, the rest removed from future consideration. On these surviving individuals a second trait is selected by saving individuals above the assigned threshold for this trait, and so on until all traits are selected. Operationally, if selection occurs on n characters, then for an individual to be saved, it must have $z_i \geq T_i$ for all characters, where T_i is the threshold value for character i. If any character fails to exceed its threshold, the individual is automatically culled without measuring any remaining characters.

Example 33.12. Under the simplest culling approach, a constant fraction f of individuals are saved after each culling. While this is not usually an optimal strategy, it represents the simplest baseline. Suppose p is the desired fraction of the population remaining after all culls. The probability of surviving all n cullings is $f^n = p$, giving $f = p^{1/n}$. For example, if we desire $p = 0.1$ of the population left afer all cullings, then for n = 2 we have $f = 0.1^{1/2} = 0.316$. Values for 3, 4, 5, and 10 cullings are 0.464, 0.562, 0.631, and 0.794. Thus, selection is actually rather weak for any *particular* culling event, with over half the population saved in time for three or more traits (with this p value).

While logistically very straightforward, the technical issues surrounding independent culling are much more complex than those for index selection. Basically, there are two tricky issues. First, culling really is not really $independent$ in that if there are correlations (genetic or phenotypic) among traits, then previous selection changes the means and variances of these traits. This leads to significant computation issues in predicting the response, especially since normality is quickly lost following one (or more) rounds of selection. The second issue is the choice of the optimal culling levels, which can be influenced by such subtleties as the actual order at which traits are culled.

The benefits of independent cullings are savings in time and resources in that not all traits have to be measured. For example, if we have k traits in n individuals, then a total of kn measurements are required for standard index selection. However, under independent culling (assuming a constant fraction f saved each culling), we measure n for the first trait, fn for the second, f^2n for the third and so on, for a total of

$$n\left(1 + f + f^2 + \cdots + f^{k-1}\right) = n\frac{1-f^k}{1-f} = n\frac{1-(p^{1/k})^k}{1-p^{1/k}} = n\left(\frac{1-p}{1-p^{1/k}}\right)$$

measurements. For example, with $p = 0.10$, for k =2, 4, 6, and 10, the total number of measurements under independent culling is $1.3n$, $2.1n$, $2.8n$, and $4.4n$, or 65%, 52%, 47% and 44% of that required under index selection. The savings, however, are typically larger than the above number suggest in that if some traits are very expensive to measure and we save

there cullings for last (or at least later), then instead of n such measurements, we have $f^j n$ measurements if the trait is measured in the $(j+1)$st culling.

As mentioned at the start of this section, while index selection gives a larger response (Hazel and Lush 1942), independent culling may give a larger *economic* return, with a high rate of return per cost. For example, Namkoong (1970) provides tabular solutions to maximize this ratio for two characters, while Xu et al. (1995) outline this maximization for multiple characters.

Selection of Extremes

Abplanalp (1972) has proposed a method related to independent culling, **selection of extremes**. As Figure 33.3 shows this method selects a fixed proportion of the highest ranking individuals for *each character*. Selection of extremes and independent culling are complementary in that selection of extremes for the upper p individuals is equal to independent culling of the lowest $1-p$ individuals.

Under independent culling, an individual must be superior in *all* characters to be selected so that an individual superior in all but one character would still be culled. Index selection considers a weighted average of all characters, so that an individual extremely superior in a few characters and average or inferior in all others can still be selected. Selection of extremes is somewhere in between — it allows for the retention of individuals superior in at least some traits along the advantage of independent culling in that not all characters need to be measured before selection. Abplanalp shows that selection of extremes is superior to independent culling with the proportions of individuals culled in less than half, so that this is method should be considered when selection is weak and there are costs associated with measuring all characters.

RELATIVE EFFICIENCIES OF INDEX SELECTION, INDEPENDENT CULLING, AND TANDEM SELECTION

Theory

While tandem selection is perhaps the most conceptually straightforward approach of artificial selection on multivariate characters and independent culling can have significant cost savings, index selection is theoretically the most efficient method. This was first demonstrated by Hazel and Lush (1942), who examined the simple case of no genetic or phenotypic correlations between characters. When all n characters have equal value (all have the same economic weight a, heritability h^2, and phenotypic variance σ_z^2), the expected response from a single cycle of selection (assuming, as usual, an infinite population) is

$$\frac{\Delta H}{\overline{i} \cdot \sigma_z \cdot a \cdot h^2} = \begin{cases} \sqrt{n}, & \text{Index selection;} \\ 1, & \text{Tandem selection;} \\ n\,\overline{i}_{p,n}/\overline{i} & \text{Independent culling} \end{cases} \tag{33.47}$$

where $\overline{i}$ is the selection intensity for a fraction p culled and $\overline{i}_{p,n}$ the selection intensity for a fraction p^{-n} culled. Under the conditions leading to Equation 33.47, the response under index selection is $\sqrt{n}$ larger than tandem selection, while independent culling is intermediate in efficiency between the other two methods. For example, suppose 5 characters are under selection and the upper ten percent of the population is culled. Here $p_{0.1,5} = 0.1^{1/5} \simeq 0.63$ so that the upper 63 percent of individuals in each character are saved, giving a selection intensity on each character of $\overline{i}_{0.1,5} = 0.60$. For $p = 0.10$, the intensity of selection on all individuals is $\overline{i} = 1.75$ giving the response under index selection relative to independent

culling as $(\sqrt{n}\cdot\bar{i})/(n\cdot\bar{i}_{p,n}) = (\bar{i}/\bar{i}_{p,n})/\sqrt{n} = (1.75\cdot 0.6)/\sqrt{5} \simeq 1.3$, while the ratio of response of index selection to tandem culling is $\sqrt{n} = \sqrt{5} \simeq 2.2$.

Figure 33.4 plots the expected responses of index and tandem selection relative to the response under independent culling for different p and n values. For weak selection (p close to one), tandem selection and independent culling give essentially the same response with index selection being far superior, while with strong selection (p near zero) independent culling and index selection have essentially the same efficiency, both being far superior to tandem selection. Figure 33.4 also shows that the superiority of index selection to independent culling and of independent culling to tandem selection increases with n the number of characters under selection.

Young (1961) and Finney (1962) relaxed the assumption of equal effects used Hazel and Lush, and generalized their result that index selection is at least as good as independent culling, which in turn is at least as good as tandem selection. However, these authors found that in many cases these differences are sufficiently small that the methods are essentially equivalent. When economic weights, heritabilities and phenotypic variances differ, for uncorrelated traits Young (1961) showed that the maximal difference in response between methods occurs when the quantity $\gamma_i = a_i h_i^2/\sigma_{z_i}$ is the same for each character. As these γ_i differ between characters, differences between methods decrease. The conditions examined by Hazel and Lush are those that maximize differences between the three methods and present the best case for index selection. When the characters being selected are phenotypically negatively correlated, Young found index selection is far more efficient than independent culling. When characters are strongly positively correlated, index selection and independent culling are essentially equivalent. This makes sense in that when characters are strong correlated, an extreme value in one character implies extreme values in most/all other characters of interest. Since the methods differ in how the extreme values of character are weighted, as these values become correlated, differences in the methods decrease. Abplanalp (1972) numerically examined selection of extremes and found that while it was always superior to tandem selection and always inferior to index selection, it is inferior to independent culling under strong selection ($p << 0.5$) and superior under weak selection ($p > 0.5$).

The above theoretical examinations are restricted to a single generation of change in an infinite population when all parameters are assumed to be exactly known. Even if the last two conditions hold, as selection proceeds it generates gametic-phase disequilibrium and can change allele frequencies, changing $\mathbf{G}$ and $\mathbf{P}$. This not only changes the weighting of the Smith-Hazel index, but different selection schemes can generate different amounts of disequilibrium and allele frequency change. Thus the asymptotic value of $\mathbf{G}$ under the same selection intensity can be different under independent culling, index, and tandem selection. A small two-trait simulation study (60 loci, 20 affecting each trait independently, 20 jointly affecting both traits) by Bennet and Swiger (1980) found that while index selection gave the largest asymptotic response, the difference in ultimate response were much less that the initial single-generation differences. This was confirmed by Villanueva and Kennedy (1993), who compared the long-term response of index and tandem selection under the assumption of the infinitesimal model (all changes in $\mathbf{G}$ and $\mathbf{P}$ are due to gametic-phase disequilibrium and are given by Equation 33.14). They found that while the largest response occurs when the index is updated (Smith-Hazel weights recomputed each generation using the current values of $\mathbf{G}$ and $\mathbf{P}$), the benefit from updating is small (a maximum of 1.5% for the cases studied).

A final complication is that the estimated Smith-Hazel index always underperforms the true index, and the above comparisons assume that the true index is known. When this is not the case, the difference between methods may become even smaller. Given all of these concerns, what do the data say?

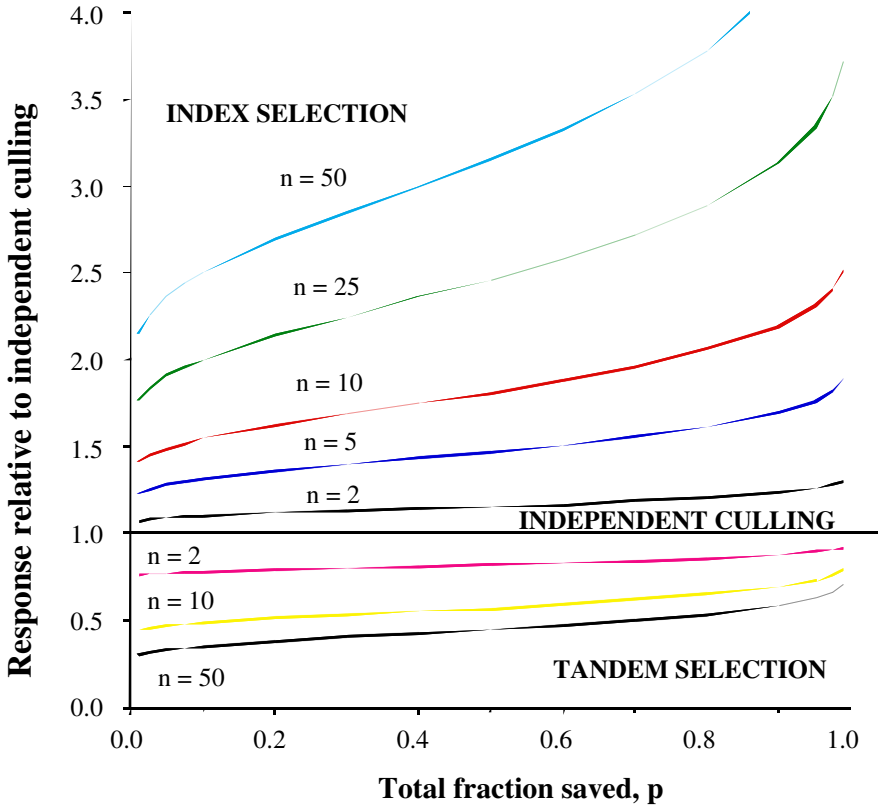

Figure 33.4. Relative efficiencies (the expected response in an infinite population for the same selection intensity) of index selection and tandem selection relative to independent culling for the special case of no genetic or phenotypic correlations when all characters have the same economic weight, heritability and additive genetic variance.

Data

A variety of experiments have examined the relative efficiencies of index selection, independent culling (including multistage selection), and tandem selection (Table 33.4). While all used very small effective population sizes and hence have a high variance in response, the consensus is that response under index selection > response under independent culling > response under tandem culling. It should be noted that in most cases the economic weights were equal, the situation giving the largest differences between methods. In those experiments where both were compared, the base and estimated indices gave essentially identical results. Elgin et al. (1970) found that while the base index had essentially the same performance as Smith-Hazel index, selection on the base index gave a much smoother response to selection. Perhaps the most notable exception to the general trend of index selection being superior was Rasmuson (1964), who examined bristle selection in *Drosophila*. She found independent culling was superior in all up-selected lines, while index selection was superior in all down-selected lines. Rasmuson suggested one reason for this difference is that while index selection allows individuals that are extreme in only one character to be saved, these individuals are lost under independent culling which requires both characters to be extreme. Hence, loci giving an extreme value in one character but not others are favored under index selection, but can be selected against under independent culling.

Table 33.4. Results of experiments examining the relative efficiencies of different methods for simultaneous selection of multiple characters. $\widehat{I}_s$ = selection using the estimated Smith-Hazel index, I_b = selection using the base index, IC = independent culling, TS = tandem selection.

Elgin et al. 1970 Cutting recovery time and four fungal resistance traits in Alfalfa	$\widehat{I}_s$ and I_b were equally effectively, both superior to IC, which in turn was superior to TS.
Sen and Robertson 1964 Abdominal and sternopleural bristles *Drosophila melanogaster*	Considerable heterogeneity between replicate lines, but general trend was $\widehat{I}_s > IC > TS$.
Rasmuson 1964 Abdominal and sternopleural bristles *Drosophila melanogaster*	TS gave poorest response. In all four up-selected lines, IC gave a larger response than $\widehat{I}_s$, while in all four down-selected lines $\widehat{I}_s$ was superior to IC.
Doolittle et al. 1972 Weight gain and litter size in mice	Averaged over replicates, $\widehat{I}_s > IC > TS$, but these differences were not significant.
Eagles and Frey 1974 Grain and straw yield in Oats	$\widehat{I}_s$, I_b, and IC gave similar responses when averaged across different environments and selection intensities.
Orozco et al. 1980 Pupal weight and egg number in *Tribolium castaneum*	$\widehat{I}_s$ gave a larger response that F_1 crosses between lines each selected for a single trait (equivalent to TS).
Campo and Rodriguez 1985 Adult weight and egg number in *Tribolium castaneum*	A modified base-index (adjusted empirically each generation to improve response) gave a larger (but not significant) response than $\widehat{I}_s$.
Campo and Rodriguez 1986 Adult weight and egg number in *Tribolium castaneum*	Replicated single-generation responses were significantly higher under $\widehat{I}_s$ than IC.

There are a couple of important exceptions to the general trend of index beating tandem selection (measured by total gain, not economic rate of return). For disease-resistance traits, there is a significant economic advantage to use independent culling, as a very large number of seedlings can easily be screened for resistance to many diseases (often simply choosing those that survive a disease challenge). Haarmann et al. (1992) selected for increased yield and resistance to leaf blight and stalk rot in maize. They noted that tandem selection was able to offer both resistance to stalk rot and increased yield, while index selection was not. They suggest that disease resistance might be more effectively selected by focusing all effort on this trait in a particular generation.

Finally, like index selection, modifications have been proposed for independent culling to allow for restriction in the response of a character (Evans 1980) and for desired gains (Xu and Muir 1991, 1992). Campo and Villanueva (1987) experimentally compared restriction via independent culling with restriction by index selection for two sets of traits in *Tribolium castaneum*, finding that while both methods restrict response in the constrained character, index selection gave a larger response in the unconstrained characters.

MULTISTAGE SELECTION

While independent culling can save resources and time when the succession of traits is scored over a very short time interval, its real economic power occurs when one is selecting on traits that appear in different life cycle stages. Suppose the goal is to improve some index of both germination time and plant height. Under independent culling, we first select on germination time and then much later select on height at maturity. This is an example of **multistage selection**, where selection occurs over different stages, resulting in fewer individuals to rear through all stages. Under traditional index selection, we need to have measured both traits in all individuals, which requires the much greater expense of growing all the seedlings to maturity.

Another classic example of multistage selection is cattle breeding, where individuals are first chosen based on their phenotypic values. These survivors are then put through much more expensive progeny testing to decide which of the remaining individuals to use for future breeding. Under traditional index selection, culling could only occur after both characters are measured in all individuals, requiring very expensive progeny tests for all cattle.

Note the we have already seen multistage selection in action in our discussion of varietal selection (Chapter 20). Here the goal is to select the best-performing pure line for an initial large collection. The more members we measure from each line, the more precise our estimate. However, we are limited by time and total acreage. Thus, the optimal strategy when a fixed number of plants can be grown each year is a type of multistage selection, where each selection cycle the number of different lines is decreased, while the number of replicates per line is increased, in order to have the highest probability of choosing the true best-performing line from an initially large collection.

Optimal Values for Multistage Cullings

A major problem with applying independent culling has been computing the optimal threshold values when more than a few characters are considered. We assume truncation selection, so that the threshold value is equivalent to the fraction of the population saved. The problem of choosing the optimal values can thus be phrased as follows. The objective is to cull k traits sequentially, where p_i is the fraction for culling cycle i. Subject to the constraint that $p = \prod p_i$ is the total fraction saved, we wish to maximize the response, which is given by $\mathbf{a}^T\mathbf{R}$. The delicate part is that each episode of selection potentially changes the phenotypic and genetic variances and covariances and we must account for this. Further, each cycle of culling further drives the distribution away from normality, further complicating matters (although this is typically ignored).

This problem has a rich history, starting with Cohran (1951). For two-stage selection, Young and Weiler (1960) and Williams and Weiler (1964) give graphs of optimal truncation points while Smith and Quaas (1982) have developed an iterative solution. Jain and Amble (1963) examine more than two stages of selection, while Saxton (1989) and Ducrocq and Colleau (1989) developed programs for more than two characters, but these are extremely slow for more than five characters. The effects of finite population size have been considered by Norell et al. (1991). Approximate solutions assuming either weak selection or low phenotypic correlations between characters (so that we can assume normality holds) have been developed by Hanson and Brim (1963), Namkoong (1970), Cunningham (1975) and Cotterill and James (1981). Xu and Muir (1991, 1992) developed a fairly general approach based on transforming the cullings to a set of orthogonal values, and we discuss this powerful method shortly.

Cotterill and James' Approximately Optimal Two-Stage Selection

Cotterill and James (1981) offer a general approximation for gain after two stages of selection under the assumption that distribution of traits following stage one selection remains roughly normally distributed. Suppose stage one selects on trait x (which may be an index of several traits), stage two selects on trait y (again, this could be an index), and the goal is the optimal improvement of the breeding value for merit g. Cotterill and James assume that $(x, y, g)^T$ is initially trivariate normal, and that the distribution of $(y, g)^T$ remains roughly normal following stage one of selection (x, however, need not remain normal). The new mean in g given selection on x follows from a standard regression of g on x,

$$\mu(g^*) = \mu_g + S_x\sigma(x, g)/\sigma(x)^2 = \mu_g + \overline{i}_1\, \rho_{x,g}\sigma(g) \tag{33.48a}$$

where $\overline{i}_1$ is the intensity for the first stage of selection. The resulting reduction in the mean of x is (Chapter 13)

$$\sigma^2(x^*) = \sigma^2(x)\,(1 - \kappa_1) \tag{33.48b}$$

where κ_1 is given by Equation 33.10 using the values given by $\overline{i}_1$. The variance of y in the selected (stage-one) population is also a classic result (Cochran 1951; Chapters 13, 31) and is

$$\sigma^2(y^*) = \sigma^2(y)\,\left[\,1 - \rho^2_{x,y}\,\kappa_1\,\right] \tag{33.48c}$$

Similarly,

$$\sigma^2(g^*) = \sigma^2(g)\,\left[\,1 - \rho^2_{x,g}\,\kappa_1\,\right] \tag{33.48d}$$

Finally, the covariance between y and g in the selected (stage one) population is

$$\sigma(y^*, g^*) = \sigma(y)\sigma(g)\,[\,\rho_{y,g} - \rho_{x,y}\,\rho_{x,g}\,\kappa_1\,] \tag{33.48e}$$

The mean in g following selection in the second stage (on y^*) is again given by a regression, g^* on y^*, which requires us to use the means, variances, and covariances following selection,

$$\mu(g^{**}) = \mu(g^*) + S_{y^*}\sigma(g^*, y^*)/\sigma(y^*)^2 = \mu(g^*) + \overline{i}_2\,\sigma(g^*, y^*)/\sigma(y^*) \tag{33.49}$$

Write the response in standard deviations, substituting 33.48a - 48e into Equation 33.49, and simplifying yields

$$\frac{\Delta g}{\sigma(g)} = \frac{\mu(g^{**}) - \mu_g}{\sigma(g)} = \overline{i}_1\,\rho_{x,g} + \overline{i}_2\,\left(\frac{\rho_{y,g} - \rho_{y,x}\rho_{x,g}\,\kappa_1}{\sqrt{1 - \rho^2_{x,y}\,\kappa_1}}\right) \tag{33.50}$$

Equation 33.50 is then numerically maximized subject to the constraint that $p_1p_2 = p$. Note that we can simply plot this as a function of p_1, as p_1 determines both $\overline{i}_1$ and κ_1, while $p_2 = p/p_1$ determines $\overline{i}_2$.

Multistage Index Selection

Hanson and Brim (1963) and especially Young (1964) suggested an extension of independent culling when selection occurs at different stages, **multistage index selection**. The first culling takes place as usual, with only individuals with $z_1 \geq T_1$ being saved. However, in subsequent culling, linear combinations of characters are used, so that if z_2 and z_3 are the characters being selected in the next two stages, then individuals are saved at these stages provided $b_{11}z_1 + b_{12}z_2 \geq T_2$ and $b_{21}z_1 + b_{22}z_2 + b_{23}z_3 \geq T_3$ (contrasted with independent culling where individuals are saved if $z_2 \geq T_2$, $z_3 \geq T_3$). By using this additional information, we can obtain a larger response that under simple independent culling. Saxton (1989) has pointed out that high correlations among indices for the different cullings can result in significant biases when

methods assuming normality are used to obtain optimal culling values. Xu and Muir (1991, 1992) offer a way around this by judicious choice of the index weights.

Selection in multiple stages is equivalent to a type of independent culling, even when partial selection indices form the basis of selection in each stage, and hence is expected to be less efficient than single-stage index selection. Wing et al. (1983) estimated that two-stage index selection is slightly less efficient than a single-stage index for a set of Whitehorn chickens, and this general trend in seen in experiments comparing single- and multi-stage selection. For Example, Ayvagari et al. (1985) examined egg number and weight and body weight in White Leghorn chickens, finding that six different two-stage index selection schemes were between 60 and 80 percent as efficient as single-stage index selection. Likewise, Campo and de la Fuente (1991) examined pupal weight and egg number in *Tribolium castaneum*. Single- versus two-stage index selection were compared, with different amounts of selection intensity during the second stage. Two-stage selection was equally effective as standard (one-stage) index selection when second-stage culling was moderate, but much poorer when 2nd-stage culling was stronger.

Xu and Muir's Method of Transformed Culling and Orthogonal Index Selection

A very clever approach to multistage selection was offered by Xu and Muir (1991, 1992), who make use of the **Cholesky decomposition** matrix. Suppose the covariance matrix $\mathbf{P}$ is of full rank (i.e., non-singular). In this case, we can find an upper triangular matrix $\mathbf{T}$ (a matrix will all zeros below the diagonal) such that $\mathbf{T}^T\mathbf{T} = \mathbf{P}$. The general advantage of such a decomposition is as follows. Consider the transformation $\mathbf{y} = (\mathbf{T}^T)^{-1}\mathbf{x}$. The resulting covariance matrix for $\mathbf{y}$ becomes

$$\sigma(\mathbf{y},\mathbf{y}) = (\mathbf{T}^T)^{-1}\sigma(\mathbf{z},\mathbf{z})\mathbf{T}^{-1} = (\mathbf{T}^T)^{-1}\mathbf{P}\mathbf{T}^{-1} = (\mathbf{T}^T)^{-1}\mathbf{T}^T\mathbf{T}\mathbf{T}^{-1} = \mathbf{I}$$

Thus, this transformation rotates the variables in $\mathbf{z}$ to remove any correlation and scales them to unit variance. The specific advantage to multistage selection follows from consideration of $(\mathbf{T}^T)^{-1}$, which is a lower triangular matrix,

$$(\mathbf{T}^T)^{-1} = \begin{pmatrix} t_{11} & 0 & \cdots & 0 \\ t_{21} & t_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{pmatrix}$$

Hence, the vector $\mathbf{y}$ of transformed trait values becomes

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (\mathbf{T}^T)^{-1}\mathbf{x} = \begin{pmatrix} t_{11}z_1 & 0 & \cdots & 0 \\ t_{21}z_1 & t_{22}z_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ t_{n1}z_1 & t_{n2}z_2 & \cdots & t_{nn}z_n \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{pmatrix} \tag{33.51a}$$

where

$$I_k = \sum_{i=1}^{k} t_{k,i}z_i \tag{33.51b}$$

is the index associated with the kth culling. Note that, by construction, $\sigma(I_k, I_j) = 0$ and these indices are orthogonal. Xu and Muir refer to this as **transformed culling**.

One immediate use of this result is for a multistage desired-gains index. Recall from Equation 33.9 that the vector of selection differentials $\mathbf{S}$ required to achieve a particular vector of responses $\mathbf{R}$ is given by

$$\mathbf{S} = \mathbf{G}^T\left(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\right)^{-1}\mathbf{R}$$

If $\mathbf{S}$ is the vector of selection differentials on the untransformed traits $\mathbf{z}$, then

$$\Delta\mathbf{y} = (\mathbf{T}^T)^{-1}\mathbf{S} \tag{33.52a}$$

is the vector of selection differentials on the transformed vector $\mathbf{y}$. Further note that since the phenotypic variance of each y_i is one, $\Delta\mathbf{y}$ corresponds to a vector of selection *intensities*. Thus, with multistage selection, the vector of required selection intensities at each stage to achieve the desired response satisfy

$$\Delta\mathbf{y} = (\mathbf{T}^T)^{-1}\mathbf{G}^T\left(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^T\right)^{-1}\mathbf{R} \tag{33.51b}$$

as obtained by Xu and Muir (1991).

Example 33.13. The following numerical example was presented by Xu and Muir (1991). Suppose we are following three traits, with

$$\mathbf{P} = \begin{pmatrix} 10 & 8 & 3 \\ 8 & 20 & 10 \\ 3 & 10 & 30 \end{pmatrix}, \qquad \mathbf{G} = \begin{pmatrix} 5 & 3 & 4 \\ 3 & 5 & 5 \\ 4 & 5 & 20 \end{pmatrix}$$

Here

$$(\mathbf{T}^T)^{-1} = \begin{pmatrix} 0.3162 & 0 & 0 \\ -0.2169 & 0.2712 & 0 \\ 0.0295 & -0.1121 & 02006 \end{pmatrix}$$

To illustrate how to proceed when $\mathbf{G}$ is not symmetric, suppose that while we select on three traits, our interest is only in the response of the first two, giving

$$\mathbf{G} = \begin{pmatrix} 5 & 3 & 4 \\ 3 & 5 & 5 \end{pmatrix}$$

If our desired gains are $\mathbf{R} = \begin{pmatrix} 0.8 & 1.0 \end{pmatrix}^T$, then Equation 33.51b gives the vector of selection intensties at each of the three stages of selection as

$$\Delta\mathbf{y} = \begin{pmatrix} \overline{\imath}_1 \\ \overline{\imath}_2 \\ \overline{\imath}_3 \end{pmatrix} = \begin{pmatrix} 0.4810 \\ 0.5422 \\ 0.3039 \end{pmatrix}, \quad \text{implying} \quad \mathbf{p} = \begin{pmatrix} 0.708 \\ 0.666 \\ 0.828 \end{pmatrix}$$

where the elements of p are the fraction of the population saved under truncation selection in order to obtained the desired selection intensity (these can easily be obtained by solving Equation 10.26a, $\overline{\imath} = \varphi(z_{[1-p]})/p$ for p given $\overline{\imath}$). The resulting total fraction saved is roughly 39%. Hence, we can achieve the desired gains for the first to traits by selection the upper 71% of trait 1, then on the upper 66.7% of the index $-0.22z_1 + 0.27z_2$ and in the third stage, the upper 83% of the index $0.3z_1 - 0.11z_2 + 0.20z_3$.

Xu and Muir also show how to use this decomposition to obtain the optimal culling values for independent culling. As above, let $H = \mathbf{a}^T\mathbf{g}$ denote the breeding value for the merit function we wish to maximize. First, note that Equation 33.52a implies $\mathbf{S} = \mathbf{T}^T\Delta\mathbf{y}$. Thus we can write the expected change in merit as

$$\Delta H = \mathbf{a}^T\mathbf{R} = \mathbf{a}^T\mathbf{G}\mathbf{p}^{-1}\mathbf{S} = \mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{T}^T\Delta\mathbf{y} = \mathbf{a}^T\mathbf{G}\mathbf{T}^{-1}\Delta\mathbf{y} \tag{33.52a}$$

where the last step follows from

$$\mathbf{P}^{-1}\mathbf{T}^T = (\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T = \mathbf{T}^{-1}(\mathbf{T}^T)^{-1}\mathbf{T}^T = \mathbf{T}^{-1}$$

Letting $\mathbf{t}_i$ denote the ith column of $\mathbf{T}^{-1}$, we can use the triangular nature of this matrix to write the response as

$$\mathbf{a}^T\mathbf{G}\mathbf{T}^{-1}\Delta\mathbf{y} = \mathbf{a}^T\sum_{i=1}^{n}\mathbf{G}\mathbf{t}_i\,\overline{\imath}_i \tag{33.52b}$$

Recalling Equation 10.26a, we can express $\overline{\imath}_i$ as a function of the fraction p_i saved, giving

$$\Delta H = \mathbf{a}^T\sum_{i=1}^{n}\mathbf{G}\mathbf{t}_i\left(\frac{\varphi(z_{[1-p_i]})}{p_i}\right) \tag{33.52c}$$

The goal is to maximize equation 35.52c under the constraint that $\prod p_j = p$. This is a *much* easier problem that finding multiple truncation points for a high-dimensional multivariate normal and can easily (and quickly) be solved using Newton-Raphson iteration (LW Appendix 4), see Xu and Muir (1991) and Xu et al. (1995) for details.

Example 33.14. This example is also do to Xu and Muir (1991). Using the same $\mathbf{G}$ and $\mathbf{P}$ from Example 33.13, obtain the optimal selection fractions if our goal is to maximize H with economic weights of $\mathbf{a}^T = \begin{pmatrix}1 & 2\end{pmatrix}$ (again, only the first two traits are of interest) under an overall fraction saved of 20%. Equation 33.52b gives

$$\Delta H = \mathbf{a}^T\sum_{i=1}^{n}\mathbf{G}\mathbf{t}_i\,\overline{\imath}_i = 3.479\,\overline{\imath}_1 + 1.139\,\overline{\imath}_2 + 1.676\,\overline{\imath}_3$$

Writing the selection intensity $\overline{\imath}_i$ as a function of p_i gives

$$\Delta H = 3.479\left(\frac{\varphi(z_{[1-p_1]})}{p_1}\right) + 1.139\left(\frac{\varphi(z_{[1-p_2]})}{p_2}\right) + 1.676\left(\frac{\varphi(z_{[1-p_3]})}{p_3}\right)$$

as the function of maximize for p_1, p_2, p_3 subject to the constraint the $p_1p_2p_3 = 0.20$. Using Newton-Raphson iteration gives $p_1 = 0.253$, $p_2 = 0.958$, and $p_3 = 0.824$. Hence, almost all of the selection is in stage 1, and the resulting response is $\Delta H = 5.026$. By contrast, if we used a Smith-Hazel index, Equation 33.19 gives the expected responses as 5.635. Thus, multistage selection gives 89% of the response as a standard Smith-Hazel index.

Xu and Muir (1992) extend their approach to the case where the "trait" added each generation can now be an index of trait values not previously scored. The idea is to find the weights that maximizes the correlation between H and the index I_i used in stage i, subject to the constraint that all of the indices remain uncorrelated. Xu and Muir refer to this as selection index updating, but we prefer the term **multistage orthogonal index selection**, as updating usually refers to changing the economic coefficients in a standard Smith-Hazel index. This restriction that the indices remain uncorrelated allows the above machinery to compute the optimal factions saved at any culling cycle. However, it is also akin to a restricted selection index and hence the response is expected to be less than maximizing response without such restrictions. However, under such a case, the indices would be correlated, raising significant computational issues in obtaining the optimal values. Xie and Xu (1997) further extend orthogonal index selection to allow for restricted or desired gains within some subset of the traits.

Literature Cited

Abplanalp, H. 1972. Selection of extremes. *Anim. Prod.* 14: 11-15. [33]

Abplanalp, H., F. X. Ogasawara, and V. S. Asmundson. 1963. Influence of selection from body weight at different ages on growth of turkey. *Brit. Poul. Sci.* 4: 71–82. [33]

Akbar, M. K., C. Y. Lin, N. R. Gyles, J. S. Gavora, and C. J. Brown. 1984. Some aspects of selection indices with constraints. *Poul. Sci.* 63: 1899–1905. [33]

Allaire, F. R. 1980. Mate selection by selection index theory. *Theor. Appl. Genet.* 57: 267–272. [33]

Atchley, W. R., S. Xu, and D. E. Cowley. 1997. Altering developmental trajectories in mice by restricted index selection. *Genetics* 146: 629–640. [33]

Avalos, E. and W. G. Hill. 1981. An experimental check of an index combining individual and family measurements. *Anim. Prod.* 33: 1–5 [33]

Baker, R. J. 1974. Selection indexes without economic weights for animal breeding. *Can. J. Anim. Sci* 54: 1–8. [33]

Baker, R. J. 1986. *Selection indices in plant breeding*. CRC Press, Boca Raton, Florida. [33]

Ayyagari, V., S. C. Mohapatra, G. S. Bisht, T. S. Thiyagasundaram, and D. C. Johari. 1985. Efficiency of one- and two-stage selection indexes for improvement of net economic merit in White Leghorns. *Theor. Appl. Genet* 70: 166–171. [33]

Bennett, G. L., and L. A. Swiger. 1980. Genetic variance and correlation after selection for two traits by index, independent culling levels and extreme selection. *Genetics* 94: 763–775. [33]

Berger, P. J. 1977. Multiple-trait selection experiments: Current status, problem areas and experimental approaches. *In* E. Pollak, O. Kempthorne, and T. B. Bailey, Jr., (eds.), *Proceedings of the international conference on quantitative genetics*, pp. 191–204. Iowa State Univ. Press, Iowa. [33]

Berger, P. J., and W. R. Harvey. 1975. Realized genetic parameters form index selection in mice. *J. Animal Sci.* 40: 38–47. [33]

Bernardo, R. 1991. Retrospective index weights ussd in mutliple trait selection in a maize breeding program. *Crop Sci.* 31: 1174–1179. [33]

Bernardo, R. 2002. *Breeding for quantitative traits in plants.* Stemma Press. [33]

Brim, C. A., H. W. Johnson, and C. C. Cockerham. 1959. Multiple selection criteria in soybeans. *Agron. J.* 51: 42–46. [33]

Brascamp, E. W. 1984. Selection indices with constraints. *Animal Breeding Abstracts* 52: 645–654. [33]

Bulmer, M. G. 1980. *The mathematical theory of quantitative genetics.* Oxford Univ. Press, NY. [33]

Burdon, R. D. 1990. Implications of non-linear economic weights for breeding. *Theor. Appl. Genet.* 79: 65–71. [33]

Bruns, E. and W. R. Harvey. 1976. Effects of varying selection intensity for traits on estimation of realized genetic parameters. *J. Anim. Sci.* 42: 291-298. [33]

Caballero, A. 1989. Efficiency in prediction of response in selection index experiments. *J. Anim. Breed. Genet.* 106: 187–194. [33]

Caldwell, B. E. and C. R. Weber. 1965. General, average, and specific selection indices for yield an F_4 and F_5 soybean populations. *Crop Sci.* 5: 223–226 [33]

Campo, J. L. and A. S. de la Blanca. 1988. Experimental comparison of selection methods to improve a non-linear trait in *Tribolium*. *Theor. Appl. Genet.* 75: 569–574. [33]

Campo, J. L. and M. B. de la Fuente. 1991. Efficiency of two-stage selection indices in *Tribolium*. *J. Heredity* 82: 228–232. [33]

Campo, J. L., M. B. de la Fuente, M. C. Garca Gil, M. C. Rodrguez, and T. Velasco. 1990. Experimental results with linear selection indices in *Tribolium*, *In, Proc. Fourth World Cong. on Genetics Applied to Livestock Production,*pp. 345–348. Edinburgh. [33]

Campo, J. L. and M. C. Rodriquez. 1985. Experimental comparison of methods for simultaneous selection of two correlated traits in *Tribolium*. I. Empirical and theoretical selection indexes. *Thoer. Appl. Genet.* 71: 93–100. [33]

Campo, J. L. and M. C. Rodriquez. 1986. Experimental comparison of methods for simultaneous selection of two correlated traits in *Tribolium*. 2. Index selection and independent culling levels: a replicated single generation test. *Génét. Sél. Evol.* 18: 437–446. [33]

Campo, J. L. and M. Rodrguez. 1990. Relative efficiency of selection methods to improve a ratio of two traits in *Tribolium*. *Theor. Appl. Genet.* 80: 343–348. [33]

Campo, J. L. and B. Villanueva. 1987. Experimental comparison of restricted selection index and restricted independent culling levels in *Tribolium castaneum*. *Genome* 29: 91–96. [33]

Campo, J. L. and T. Velasco. 1989. An experimental test of optimum and desired-gains indexes in *Tribolium*. *J. of Heredity* 80: 48–52. [33]

Cochran, W. G. 1951. Improvement by means of selection. *In* N. Neyman (ed.) *Proc. Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 449-470. University of California Press, Berkeley. [33]

Cotterill, P. P. 1985. On index selection. II. Simple indices which require no genetic parameters or special expertise to construct. *Silvae Genetica* 34: 64–69. [33]

Cotterill, P. P. and J. W. James. 1981. Optimizing two-stage independent culling in tree and animal breeding. *Theor. Appl. Genet.* 59: 67–72. [33]

Cotterill, P. P. and N. Jackson. 1985. On index selection. I. Methods for determining economic weight. *Silvae Genetica* 34: 56–63. [33]

Crosbie, T. M., J. J. Mock, and O. S. Smith. 1980. Comparison of gains predicted by several selection methods for cold tolerance traits of two maize populations. *Crop Sci.* 20: 649–655. [33]

Cunningham, E. P. 1975. Multi-stage index selection. *Theo. Appl. Genet.* 46: 55–61. [33]

Cunningham, E. P., R. A. Moen, and T. Gjedrem. 1970. Restriction of selection indexes. *Biometrics* 26: 67–75. [33]

Dekkers, J. C. M., P. V. Birke, and J. P. Gibson. 1995. Optimum linear selection indexes for multiple generation objectives with non-linear profit functions. *Anim. Sci.* 61: 165–175. [33]

Dickerson, G. E., C. T. Blunn, A. B. Chapman, R. M. Kottman, J. L. Krider, E. J. Warwick, J. A. Whatley, Jr., M. L. Baker, J. L. Lush, and L. M. Winters. 1954. Evaluation of selection in developing inbred lines of swine. *Mo. Agric. Exp. Sta. Res. Bull.* 551. [33]

Doolittle, D. P., S. P. Wilson, and L. L. Hulbert. 1972. A comparison of multiple trait selection methods in the mouse. *J. Heredity* 63: 366–372. [33]

Ducrocq, V. and J. J. Colleau. 1989. Optimum truncation points for independent culling level selection on a multivariate normal distribution, with an application to dairy cattle selection. *Genet. Sel. Evol.* 21: 185–198. [33]

Eagles, H. A. and K. J. Frey. 1974. Expected and actual gains in economic value of oat lines from five selection methods. *Crop Sci.* 14: 861–864. [33]

Eisen, E. J. 1977a. Antagonistic selection index results with mice. *In* E. Pollak, O. Kempthorne, and T. B. Bailey, Jr., (eds.), *Proceedings of the international conference on quantitative genetics*, pp. 117–139. Iowa State Univ. Press, Iowa. [33]

Eisen, E. J. 1977b. Restricted selection index: an approach to selecting for feed efficiency. *J. Anim. Sci.* 44: 958–972. [33]

Eisen, E. J. 1978. Single-trait and antagonistic index selection for litter size and body weight in mice. *Genetics* 88: 781–811. [33]

Eisen, E. J. 1992. Restricted index selection in mice designed to change body fat without changing body weight: direct responses. *Theor. Appl. Genetics* 83: 973–980. [33]

Eisen, E. J., L. S. Benyon, and J. A. Douglas. 1995. Long-term restricted index selection in mice designed to change fat content without changing body size. *Theor. Appl. Genet.* 91: 340–345. [33]

Elgin, J. H., Jr., R. R. Hill, Jr., and K. E. Zeiders. 1970. Comparison of four methods of multiple trait selection for five traits in alfalfa. *Crop Sci.* 10: 190–193. [33]

Elston, R. C. 1963. A weight-free index for the purpose of ranking or selection with respect to several traits at a time. *Biometrics* 19: 85–97. [33]

Evans, R. 1980. Constrained selection by independent culling levels. *Anim. Prod.* 30: 79–83. [33]

Fairfull, R. W., G. W. Friars, and J. W. Wilton. 1977. An emipirical comparison of selection methods for the improvement of biomass. *Theor. Appl. Genet.* 50: 193–198. [33]

Famula, T. R. 1992. A comparison of restricted selection index and linear programing sire selection. *Theor. Appl. Genet.* 84: 384–389. [33]

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

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7: 179–189. [33]

Garwood, V. A., and P. C. Lowe. 1978. A replicated single generation test of a restricted selection index in poultry. *Theor. Appl. Genet.* 52: 227–231. [33]

Gibson J. P., and B. W. Kennedy. 1990. The use of constrained selection indexes in breeding for economic merit. *Theor. Appl. Gene/* 80: 801–805. [33]

Gjedrem, T. 1972. A study on the definition of the aggregate genotype in a selection index. *Acta Agric. Scad.* 22: 11–16. [33]

Goddard, M. E. 1983. Selection indices for non-linear profit functions. *Theor. Appl. Genet.*: 64: 339–344. [33]

Gomez-Raya, L. and E. B. Burnside. 1990. The effect of repeated cycles of selection on genetic variance, heritability, and response. *Theor. Appl. Genet.* 79: 568–574. [33]

Groen, A. F., T. H. E. Meuwissen, A. R. Vollea, and E. W. Brascamp. 1994. A comparison of alternative index procedures for mulitple generation selection on non-linear profit. *ANim. Sci.* 59: 1–9. [33]

Gunsett, F. C., K. N. Andriano, and J. J. Rutledge. 1984. Estimation of genetic parameters from multiple trait selection experiments. *J. Animal Sci.* 58: 591–599. [33]

Haarmann, R. J., D. G. White, and J. W. Dudley. 1993. Index vs. tandem selection for improvement of grain yield, lef blight, and stalk rot resistance in maize. *Maydica* 38: 183–188. [33]

Hanson, W. D., and C. A. Brim. 1963. Optimum allocation of test material with two-stage testing with an application to evaluation of soybean lines. *Crop Sci.* 3: 43–49. [33]

Hanson, W. D. and H. W. Johnson. 1957. Methods for calculating and evaluating a general selection index obtained by pooling information from two or more experiments. *Genetics* 42: 421–432. [33]

Harris, D. L. 1961. A monte carlo study of the influence of errors of parameter estimation upon index selection. *Biometrics* 17: 501–502. [33]

Harris, D. L. 1963. The infleunce of errors of parameter estimation under index selection. *Statistical genetics and plant breeding*, pp. 491–500. Natl. Acad. Sci., Natl. Res. Council Publ. No. 982, Washington, D.C. [33]

Harris, D. L. 1970. Breeding for efficiency in livestock production: defining the economic objectives. *J. Anim. Sci.* 30: 860–865. [33]

Harville, D. A. 1975. Index selection with proportionality constraints. *Biometrics* 31: 223–225. [33]

Harville, D. A. 1974. Optimal procedures for some constrained selection problems. *J. Amer. Sata. Assoc.* 69: 446–452. [33]

Hayes, J. F. and W. G. Hill. 1980. A reparameterization of a genetic selection index to locate its sampling properties. *Biometrics* 36: 237–248. [33]

Hayes, J. F. and W. G. Hill. 1981. Modification of estimates of parameters in the construction of genetic selection indices ('bending'). *Biometrics* 37: 483–493. [33]

Hazel, L. N. 1943. The genetic basis for constructing selection indexes. *Genetics* 28: 476–490. [33]

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

Heidhues, T. 1961. Relative accuracy of selection indices based on estimated genotypic and phenotypic paramters. *Biometrics* 17: 502–503. [33]

Heidhues, T. and C. R. Henderson. 1962. Beitrag zum problem des basisindex. *Z. Tierzüchtg. Züchtgsbiol.* 77: 297–311. [33]

Henderson, C. R. 1963. Selection index and expected genetic advance. *In, Statistical Genetics and Plant Breeding.* NAS-NRC 982: 141-163. [33]

Hill, W. G. and K. Meyer. 1984. Effects of errors in parameter estiames on efficiency of restricted genetic selection indices. *In*, K. Hinkelmann (ed), *Experimental design, statistical models, and genetic statistics: essays in honor of Oscar Kempthorne* pp. 345–367. M. Dekker, New York [33]

Hill, W. G. and R. Thompson. 1978. Probabilities of non-positive definite between group or genetic covariance matrices. *Biometrics* 34: 429–439. [33]

Holbrook, C. C., J. W. Burton, and T. E. Carter, Jr. 1989. Evaluation of recurrent restricted index selection for increasing yield while holding seed protein constant in soybean. *Crop Sci* 29: 324–329. [33]

Hoerl, A. E., and R. W. Kennard. 1970. Ridge regression: biased estimation of nonorthogonal problems. *Technometrics* 2: 55–67. [33]

Hoerl, A. E., and R. W. Kennard. 1981. Ridge regression — 1980. Advances, algorithm, and applications. *Am. J. Math. and Manag. Sci.* 1: 5–83. [33]

Humphreys, M. O. 1995. Multitrait response to selection in *Lolium perenne* L. (perennial ryegrass) populations. *Herediоty* 74: 510–517. [33]

Itoh, Y. 1991. Changes in genetic correlations by index selection. *Genet. Sel. Evol.* 23: 301–308. [33]

Itoh, Y. and Y. Yamada. 1987. Comparisons of selection indices achieving predetermined proportional gains. *Genet. Sel. Evol.* 19: 69–82. [33]

Itoh, Y. and Y. Yamada. 1988a. Linear selection indices for non-linear profit functions. *Theor. Appl. Genet.* 75: 553–560. [33]

Itoh, Y. and Y. Yamada. 1988b. Selection indices for desired relative genetic gains with inequality constraints. *Theor. Appl. Genet.* 75: 731–735. [33]

Jain, J. P., and V. N. Amble. 1962. Improvement through selection at successive stages. *J. Indian Soc. Agric. Stat.* 14: 88–109. [33]

James, J. W. 1968. Index selection with restrictions. *Biometrics* 24: 1015–1018. [33]

James, J. W. 1982. Construction, uses, and problems of multitrait selection indices. *In*, *2nd. World Congress of Genetics Applied to Livestock Production*, Vol 5: 130–139. [33]

Kempthorne, O. and A. W. Nordskog. 1959. Restricted selection indices. *Biometrics* 15: 10–19. [33]

Lin, C. Y. 1978. Index selection for genetic improvement of quantitative characters. *Theo. Appl. Genet.* 52: 49–56. [33]

Lin, C. Y. 1985. A simple stepwise procedure of deriving selection index with restrictions. *Theor. Appl. Genet.* 70: 147–150. [33]

Lin, C. Y., and F. R. Allaire. 1977. Heritability of a linear combination of traits. *Theor. Appl. Genet.* 51: 1–3. [33]

Magnussen, S. 1991. Index selection with nonlinear profit function as a tool to achieve simultaneous genetic gain. *Theor. Appl. Genet.* 82: 305–312. [33]

Mallard, J. 1972. La theorie et le calcu des index de selection avec restrictions: synthese critique. *Biometrics* 28: 713–735. [33]

Matzinger, D. F., C. C. Cockerham, and E. A. Wernsman. 1977. Single character and index mass selection with random mating in a naturally self-fertilizing species.. *In* E. Pollak, O. Kempthorne, and T. B. Bailey, Jr., (eds.), *Proceedings of the international conference on quantitative genetics*, pp. 503–518. Iowa State Univ. Press, Iowa. [33]

Matzinger, D. F., E. A. Wernsman and W. W. Weeks. 1989. Restricted index selection for total alkaloids and yield in tobacco. *Crop. Sic.* 29: 74–77. [33]

McBride, G., and A. Robertson. 1963. Selection using assortative mating in *D. melanogaster*. *Genet. Res.* 4: 356–369. [33]

McCarthy, J. C., and D. P. Doolittle. 1977. Effects of selection for independent changes in two highly correlated body weight traits of mice. *Genet. Res.* 29: 133–145. [33]

Melton, B. E., E. O. Heady, and R. L. Willham. 1979. Estimation of economic values for selection indices. *Anim. Prod.* 28: 279–286. [33]

Melton, B. E., R. L. Willham, and E. O. Heady. 1993. A note on the estimation of economic values for selection indices: a response. *Anim. Prod.* 59: 455–459. [33]

Meuwissen, T. H. E., J. P. Gibson, and M. Quinton. 1995. Genetic improvemnet of production while maintaining fitness. *Theor. Appl. Genet*/ 90: 627–635. [33]

Moav, R., and W. G. Hill. 1966. Specialized sire and dam lines. IV. Selection within lines. *Anim. Prod.* 8: 373–390. [33]

Mortimer, S. I., and J. W. James . 1987. Changes in genetic parameters under restricted index selection. *Genet Res.* 149: 129–134. [33]

Morley, F. H. W. 1955. Selection for economic characters in Australian Merino sheep. V. Further estimates of phenotypic and genetic parameters. *Aust. J. Agric. Res.* 6: 77–90. [33]

Muir, W. M. and S. Xu. 1991. An approximate method for optimum independent culling level selection for n stages of selection with explicit solutions. *Theor. Appl. Genet.* 82: 457–465. [33]

Mulmbda, N. N., and J. J. Mock. 1978. Improvement of yield potential of the Eto Blanco maize (*Zea mays* L.) popualtions by breeding for plant traits. *Egypt. J. Genet. Cytol.* 7: 40–51. [33]

Namkoong, G. 1970. Optimum allocation of selection intensity in two stages of truncation selection. *Biometrics* 26: 465–476. [33]

Namkoong, G. 1979. *Introduction to quantitative genetics in forestry*. U. S. Forest Serivce Technical Bulletin 1588, Washington, D. C. [33]

Nanda, D. N. 1949. The standard errors of discriminant function coefficients in plant-breeding experiments. *J. R. Stat. Soc, B* 11: 283–290. [33]

Niebel, E. and L. D. Van Vleck. 1982. Selection with restriction in cattle. *J. Anim. Sci.* 55: 439–457. [33]

Nordskog, A. W. 1978. Some statistical properties of an index of multiple traits. *Theor. Appl. Genet.* 52: 91–94. [33]

Norell, L., T. Arnason, and K. Hugason. 1991. Multistage index selection in finite populations. *Biometrics* 47: 205–221. [33]

Okada, I and R. T. Hardin. 1967. An experimental examination of restricted selection index, using *Tribolium castaneum*. 1. The results of two-way selection. *Genetics* 57: 227–236. [33]

Okada, I and R. T. Hardin. 1970. An experimental examination of restricted selection index, using *Tribolium castaneum*. 1. The results of long-term one-way selection. *Genetics* 64: 533–539. [33]

Orozco, F., J. L. Campo, M. C. Fuentes, C. Lopz-Fanjul, and P. Tagarro. 1980. A comparison between crossbreeding and index selection for the simultaneous improvement of two traits in *Tribolium castaneum*. *In* A. Roberston, (ed.), *Selection experiments in laboratory and domestic animals*, pp. 37–52. Commonwealth Agricultural Bureaux. [33]

Panse, V. G. 1946. An application of the discriminant function for selection in poultry. *J. Genetics* 47: 242–248. [33]

Pasternak, H. and J. L. Weller. 1993. Optimum linear indices for non-linear profit functions. *Anim. Prod.* 55: 43–50. [33]

Pešek, J. and R. J. and Baker. 1969. Desired improvement in realtion to selection indices. *Can. J. of Plant Science* 49: 803–804. [33]

Rasmuson, M. 1964. Combined selection for two bristle characters in *Drosophila. Hereditas* 51: 231–256. [33]

Rønningen, K. 1971. Selection index for quadratic and cubic models of the aggregate genotype. *Meld. Nor. Landbrukshoøgskole* 50(2): 1-30. [30]

Sales, J. and W. G. Hill. 1976a. Effect of sampling erros on efficiency of selection indices. 1. Use of information from relatives for single trait improvement. *Anim. Prod.* 22: 1–17. [33]

Sales, J. and W. G. Hill. 1976b. Effect of sampling errors on efficiency of selection indices. 2. Use of information on associated traits for improvement of a single important trait. *Anim. Prod.* 23: 1–14. [33]

Saxton, A. M. 1983. A comparison of exact and sequential methods in multi-stage index selection. *Theo. Appl. Genet.* 66: 23–28. [33]

Saxton, A. M. 1986. A comparison of bending and ridge regression in selection index estimation. *Proc. 3rd World Congress on Genetics Applied to Livestock Production* XIII: 449–453. [33]

Saxton, A. M. 1989. INDCULL Version 3.0: Independent culling for two or more traits. *J. Heredity.* 80: 166–167. [33]

Scheinberg, E., A. E. Bell, and V. L. Anderson. 1967. Genetic gain in populations of *Tribolium castaneum* under uni-stage tandem selection and under restrcited selection indices. *Genetics* 55: 69–90. [33]

Sen, B. K. and A. Robertson. 1964. An experimental examination of methods for the simultaneous selection of two characters, using *Drosophila melanogaster*. *Genetics* 50: 199–209. [33]

Smith, C. 1983. Effects of changes in economic weights on the efficiency of index selection. . ANim. Sci. 56: 1057–1064. [33]

Smith, H. F. 1936. A discriminant function for plant selection. *Ann. Eugen.* 7: 240–250. [33]

Smith, O. S., A. R. Hallauer, and W. A. Russell. 1981. Use of index selection in recurrent selection programs in maize. *Euphytica* 30: 611–618. [33]

Smith, S. P. and R. L. Quaas. 1982. Optimal truncation points for independent culling-level selection involving two traits. *Biometrics* 38: 975–980. [33]

Tai, G. C. C. 1977. Index selection with desired gains. *Crop Sci.* 17: 182–183. [33]

Tai, G. C. C. 1989. A proposal to improve the efficiency of index selection by "rounding" . *Theor. Appl. Genet.* 78: 798–800. [33]

Tallis, G. M. 1960. The sampling errors of estimated genetic regression coefficients and the error of predicted genetic gains. *Austral. J. Stat.* 2: 66–77. [33]

Tallis, G. M. 1962. A selection index for optimum genotype. *Biometrics* 18: 120–122. [33]

Tallis, G. M. 1985. Constrained selection. *Jpn. J. Genet.* 60: 151–155. [33]

Tallis, G. M. 1986. Constrained selection — corrigendum and addendum. *Jpn. J. Genet.* 61: 183–184. [33]

Thompson, R. 1980. A note on the estimation of economic values for selection indices. *Anim. Prod.* 31: 115–117. [33]

Turner, H. N. and S. S. Y. Young. 1969. *Quantitative genetics in sheep breeding*. Cornell University Press, Ithaca. [33]

Van Vleck, L. D. 1993. *Selection index and introduction to mixed model methods.* CRC Press. [33]

Villanueva, B. and B. W. Kennedy. 1990. Effects of selection on genetic parameters of correlated traits. *Theor. Appl. Genet.* 80: 746–752. [33]

Villanueva, B. and B. W. Kenndedy. 1993. Index versus tandem selection after repeated generations of selection. *Theor. Appl. Genet.* 85: 706–712. [33]

von Butler, I., J. Aumann, and F. Pirchner. 1980. Preliminary results on antagonistic selection in mice. *In* A. Roberston, (ed.), *Selection experiments in laboratory and domestic animals*, pp. 114–115. Commonwealth Agricultural Bureaux. [33]

von Butler, I., H. Willeke, and F. Pirchner. 1986. Antagonistic selection for correlated body weight traits in different mouse populations. *Theor. Appl. Genet.* 71: 698–702. [33] [33]

Williams, J. M. and H. Weiler. 1964. Further charts for the means of truncated normal bivariate distributions. *Aust. J. Stat.* 6: 117–129. [33]

Williams, J. S. 1962a. Some statistical properties of a genetic selection index. *Biometrika* 49: 325–337. [33]

Williams, J. S. 1962b. The evaluation of a selection index. *Biometrics* 18: 375–393. [33]

Wilton, J. W., D. A. Evans and L. D. Van Vleck. 1968. Selection indices for quadratic models of total merit. *Biometrics* 24: 937–949. [33]

Wing, T. L., H. Singh, and A. W. Nordskog. 1983. Use of individual feed records in a selection program for egg production efficiency. III. Relative effectiveness of two-stage selection. *Poul. Sci.* 62: 721–724. [33]

Wricke, G., and W. E. Weber. 1986. *Quantitative genetics and selection in plant breeding.* Walter de Gruyter and Co., New York. [33]

Xie, C., and S. Xu. 1997. Restricted multistage selection indices. *Genet. Sel. Evol.* 29: 193–203. [33]

Xu, S., T. G. martin, and W. M. Muir. 1995. Multistage selection for maximum economic return with an application to beef cattle breeding. *J. Anim. Sci* 73: 699–710. [33]

Xu, S. and W. M. Muir. 1990. The application of ridge regression to multiple trait selection indices. *J. ANim. Breed. Genet.* 107: 81–88. [33]

Xu, S. and W. M. Muir. 1991. Multistage selection for genetic gain by orthogonal transformation. *Genetics* 129: 963–974. [33]

Xu, S. and W. M. Muir. 1992. Selection index updating. *Theo. Appl. Genet.* 83: 451–458. [33]

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

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. [33]

Young, S. S. Y. 1964. Multi-stage selection for genetic gain. *Heredity* 19: 131–145. [33]

Young, S. S. Y. and H. Weiler. 1960. Selection for two correlated traits by independent culling levels. *J. Genetics* 57: 329–338. [33]

Zeng, Z.-B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. *Evolution 42*: 363-374. [33]