34

Multivariate Response: Changes in Means

The whole organism is so tied together than when slight variations in one part occur, and are accumulated through natural selection, other parts become modified. This is a very important subject, most imperfectly understood. — Charles Darwin (1859)

Version 26 June 2014

The multivariate breeder's equation for predicting the change in a vector of means is one of the most powerful results in quantitative genetics. We have already informally discussed this equation in several previous chapters (10, 16, 29), and here we start its full development. Reviews of multivariate response are given by Arnold (1992), Arnold et al. (2001), and McGuigan (2006).

We begin with a brief overview of the multivariate breeder's, reviewing its key features before discussing two slight generalizations of the breeder's equation: the multivariate secondary theorem of natural selection and the response under linear (but not necessarily normal) parent-offspring regressions. We then return to the historical roots of the multivariate breeder's equation, reviewing (in a bivariate framework) correlated responses and estimation of realized genetic correlations. We then move on to what is the core of this chapter, the detailed analysis of the breeder's equation from an evolutionary biology perspective, as this field has driven much of the recent work on the analysis of multivariate selection. We first examine schemes to standardized response and apply these to the analysis of realized selection gradients – attempts to infer historical patterns of natural selection from the patterns of population divergence. We then turn to a detailed geometric analysis of the $\mathbf{G}$ matrix of genetic variances and covariances. The geometry of this matrix imposes constraints on direction of evolution and we review some central concepts, such as how the eigenstructure of $\mathbf{G}$ imposes constraints, the evolution of mean fitness, evolution along genetic lines of least resistance, subspace projection of response, and the multivariate extensions of evolvability and conditional evolvability.

A key assumption through, which is relaxed in Chapter 31, is the constancy of $\mathbf{G}$. Given our previous discussions about change in genetic variances under selection (Chapters 13, 24-26), we of course expect $\mathbf{G}$ is change, in particular because genetic covariances are much more sensitive to changes from selection than are variances. The justification for a constant $\mathbf{G}$ that appears in the evolutionary literature is that if selection has been proceeding for a long (i.e. evolutionary) time, the resulting covariance matrix may be in mutation-selection-drift equilibrium and hence be (relatively) constant (Chapter 27). While this may be true if selection has been occurring in a particular direction for a long period of time, such is clearly not the case when a natural population is subjected to selection in a novel direction (either through artificial selection or a sudden change in the environment), in which cases rapid changes in $\mathbf{G}$ can occur (Chapter 31).

THE MULTIVARIATE BREEDER'S EQUATION

Overview of Key Features and Concepts

One of the central, and most elegant, expressions in quantitative genetics is the multivariate breeder's equation,

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} \tag{30.1a}$$

which relates the vector $\mathbf{R}$ of responses (the *between*-generation change in the means for a vector of traits) with the vector $\mathbf{S}$ of selection differentials (*within*-generation changes in the mean) and the phenotypic $\mathbf{P}$ and genetic $\mathbf{G}$ covariances matrices for these traits, where

$$P_{ij} = \sigma(z_i, z_j), \qquad G_{ij} = \sigma(A_i, A_j) \tag{30.1b}$$

where z_i and A_i denote the phenotypic and breeding values for trait i. This result was anticipated by Hazel (1943), and was more formally developed (in various forms) by Young and Weiler (1960), Harvey and Bearden (1962), Magee (1965), and Lande (1979). The vector of responses $\mathbf{R}$ is also given as $\boldsymbol{\Delta\mu}$ and $\boldsymbol{\Delta\bar{z}}$ in the literature and we will occasionally use these alternative forms in our discussion. The most compact representation, the **Lande Equation**, is based on the selection gradient $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S}$,

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} \tag{30.2}$$

As was discussed in Chapter 29, if the distribution of phenotypes in multivariate normal, there is an intimate connection between $\boldsymbol{\beta}$ and the mean fitness surface $\overline{W}(\boldsymbol{\mu})$, namely $\boldsymbol{\beta} = \nabla_{\boldsymbol{\mu}}[\ln \overline{W}(\boldsymbol{\mu})]$, the gradient of (log) mean fitness. Thus $\boldsymbol{\beta}$ represents the optimal direction of change in the vector of means to increase fitness. The Lande equation highlights the importance of **genetic constraints** in selection response. The optimal direction to change the means is $\boldsymbol{\beta}$, but the *actual* direction of response $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$ is a *rotation* and *scaling* away from $\boldsymbol{\beta}$ (Appendix 4). These constraints are imposed by the geometry inherent in the genetic variance-covariance structure $\mathbf{G}$.

The Lande equation is quite profound, providing a connection between evolution, ecology, genetics, and development. The response $\mathbf{R}$ is the actual evolution that occurs for a vector of traits. This depends on the nature of phenotypic selection on those traits ($\boldsymbol{\beta}$), which is determined by ecological interactions of an individual with its biotic and abiotic environments. The mean fitness surface $\overline{W}(\boldsymbol{\mu})$ describes an **adaptive topography** or **surface** (Wright 1931, Simpson 1944, Lande 1979, Arnold et al. 2001), so $\boldsymbol{\beta}$ gives the optimal direction for adaptation (the direction of steepest climb on the topography). Ecology along is not sufficient for evolution, as there also has to be *heritable variation* in order for response to occur, no matter how strong the selection. As mentioned above, constraints in this variation, which are ultimately functions of genetics and development, are encapsulated in $\mathbf{G}$ (Cheverud 1984). The Lande equation also provides connections to **macroevolution** (evolution above the species level), a topic discussed in Chapter 43 (also see Lande 1980, Arnold 2001, Polly 2008).

Equation 30.2 also nicely demonstrates the relationship between the two main complications with selection on multiple characters: the *within-generation* change due to *phenotypic* correlations $\boldsymbol{\beta}$ and the *between-generation* change $\mathbf{R}$ due to *additive genetic* correlations. $\boldsymbol{\beta}$ represents the amount of direct selection on a trait, while the *observed* within-generation change is given by $\mathbf{S} = \mathbf{P}\boldsymbol{\beta}$. Hence, the observed selection differential on trait j is

$$S_j = P_{jj}\beta_j + \sum_{i \neq j} P_{ij}\,\beta_i = \sigma^2(z_j)\,\beta_j + \sum_{i \neq j} \sigma(z_j, z_i)\,\beta_i \tag{30.3a}$$

Likewise, the response (between-generation change) in trait j is

$$R_j = G_{jj}\beta_j + \sum_{i \neq j} G_{ij}\,\beta_i = \sigma^2(A_j)\,\beta_j + \sum_{i \neq j} \sigma(A_j, A_i)\,\beta_i \tag{30.3b}$$

In both Equations 30.3a and b, the first term is the change due to *direct selection* on trait j, while the sum is the *indirect contributions* from the correlated response from selection on other traits. Again, these equation stress the key points that within-generation changes are influenced by *phenotypic* correlations, between-generation changes by *genetic* correlations. Even if direct selection only occurs on character i, other characters genetically correlated with i also change,

$$R_j = G_{ij}\beta_j = \sigma(A_j, A_i)\,\beta_i \tag{30.4a}$$

Thus, the ratio of the expected change in two characters when only one (trait i) is under direction selection is

$$\frac{R_j}{R_i} = \frac{G_{ij}\beta_i}{G_{jj}\beta_i} = \frac{G_{ij}}{G_{jj}} = \frac{\sigma(A_j, A_i)}{\sigma^2(A_j)} = r_A(ij)\sqrt{\frac{\sigma^2(A_j)}{\sigma^2(A_i)}} \tag{30.4b}$$

If both characters have the same additive genetic variance ($\sigma^2(A_i) = \sigma^2(A_j)$), the ratio of response simply reduces to $r_A(ij)$, the correlation between their additive genetic values.

Example 30.1. Conner and Via (1992) investigated the effects of pupal weight, wing length, and thorax width on female reproductive success in a laboratory population of the flour beetle *Tribolium castaneum*. They also examined the effects of these same traits on male reproductive success, which we ignore here. The estimated selection gradient in females for the three traits (in the order of pupal weight, wing length, and width) was

$$\boldsymbol{\beta} = \begin{pmatrix} 0.22 \\ -0.01 \\ -0.16 \end{pmatrix}$$

Using the estimated $\mathbf{G}$ matrix

$$\mathbf{G} = \frac{1}{1000}\begin{pmatrix} 1.25 & -0.01 & 0.36 \\ -0.01 & 0.02 & 0.14 \\ 0.36 & 0.14 & 0.20 \end{pmatrix}$$

gives the predicted response as

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \frac{1}{1000}\begin{pmatrix} 0.216 \\ -0.025 \\ 0.046 \end{pmatrix}$$

Differences from univariate predictions are immediately apparent. The optimal direction $\boldsymbol{\beta}$ to increase fitness is to increase pupal weight while decreasing thorax width by roughly the same amount ($+0.22$ vs. -0.16). However, the response shows that *both* traits are increased. To obtain a better understanding of these results, we can further break the response into direct ($G_{ii}\beta_i$) and correlated ($R_i - G_{ii}\beta_i$) components (all scaled by 1/1000),

Trait	Direct	Correlated
Pupal weight	0.2705	-0.0575
Wing lenght	-0.0002	-0.0246
Thorax width	-0.0320	0.0778

There is a (predicted) correlated response in the trait not under selection (wing length), and a response in the opposite direction of that favored by strong selection for thorax width, due to the correlated response from pupal weight overpowering the direct response on width. (Note that our results differ from those published by Conner and Via, who considered an analysis using all six traits simultaneously, while we focused on just the three in females.)

The multivariate breeder's equation is such a powerful tool because numerous important problems in quantitative genetics can be treated in a multiple trait framework. For example, a trait expressed over several environments can be treated as a series of correlated traits (Falconer 1952). When traits in the mother and traits in the offspring both influence offspring fitness, we can treat these direct and maternal effects as potentially correlated (Chapter 35). When the goal is to optimize the response from a vector of traits, **index selection** is often used (Chapters 32, 33), selecting on those individuals with the best index I value,

$$I = \sum_i a_i\, z_i = \mathbf{a}^T\mathbf{z}$$

Examples include within- and between-family selection (Chapter 17), selecting on optimal performance over a series of environments (Chapter 37), **marker-assisted** and **genomic** selection (Chapter 35) wherein molecular marker information is used in the selection decisions, and **BLUP selection** (Chapter 34) whereby individuals with the highest breeding values are selected using information from their relatives. Analysis of all of these problems eventually reverts back to the multivariate breeder's equation.

Derivation of the Multivariate Breeder's Equation

The critical assumption leading to the multivariate breeder's equation is that the vectors of phenotypic and breeding values are jointly multivariate normal. We (somewhat) relax this assumption at the end of this section. The general view is that the breeder's equation is *somewhat* robust to gentle relaxation of these assumptions. The derivation of this equation highlights the importance of normality. Of course, we have already seen a long list (Table 10.2) of other concerns for the univariate breeder's equation, and these (of course) apply to its multivariate extension. The assumption of normality takes on additional importance with multiple characters, wherein the nature of association between characters is critical. If the distribution we are examining (phenotypic or genotypic) is Gaussian, then the covariance matrix completely describes the association between all sets of characters. This is by no means guaranteed when distributions are non-Gaussian. Since almost all of the theory of genetic constraints to multivariate response is build around analysis of the $\mathbf{G}$ matrix, and hence the assumption that the breeding values are MVN, the robustness of these results remains unclear.

Recall from LW Chapter 8 that conditional distributions of subvectors from a multivariate normal are also multivariate normal. In particular, partition a MVN distributed n-dimensional column vector $\mathbf{x}$ into two components, an m-dimensional column vector $\mathbf{x_1}$ and an $(n-m)$-dimensional column vector $\mathbf{x_2}$ of the remaining variables, e.g.,

$$\mathbf{x} = \begin{pmatrix} \mathbf{x_1} \\ \mathbf{x_2} \end{pmatrix}$$

where the mean vector and variance-covariance matrix are similarly be partitioned as

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu_1} \\ \\ \boldsymbol{\mu_2} \end{pmatrix} \qquad \text{and} \qquad \mathbf{V} = \begin{pmatrix} \mathbf{V}_{\mathbf{x_1}\mathbf{x_1}} & \mathbf{V}_{\mathbf{x_1}\mathbf{x_2}} \\ \\ \mathbf{V}^T_{\mathbf{x_1}\mathbf{x_2}} & \mathbf{V}_{\mathbf{x_2}\mathbf{x_2}} \end{pmatrix} \tag{30.5}$$

The conditional random variable $\mathbf{x_1}|\mathbf{x_2}$ is distributed MVN with (m-dimensional) mean vector

$$\boldsymbol{\mu}_{\mathbf{x}_1|\mathbf{x}_2} = \boldsymbol{\mu}_1 + \mathbf{V}_{\mathbf{x}_1\mathbf{x}_2}\mathbf{V}_{\mathbf{x}_2\mathbf{x}_2}^{-1}(\mathbf{x_2} - \boldsymbol{\mu}_2) \tag{30.6a}$$

and ($m \times m$) variance-covariance matrix

$$\mathbf{V}_{\mathbf{x}_1|\mathbf{x}_2} = \mathbf{V}_{\mathbf{x}_1\mathbf{x}_1} - \mathbf{V}_{\mathbf{x}_1\mathbf{x}_2}\mathbf{V}_{\mathbf{x}_2\mathbf{x}_2}^{-1}\mathbf{V}_{\mathbf{x}_1\mathbf{x}_2}^{T} \tag{30.6b}$$

Likewise (LW Equation 8.27), the regression of $\mathbf{x_1}$ on $\mathbf{x_2}$ is given by

$$\mathbf{x_1} = \boldsymbol{\mu}_1 + \mathbf{V}_{\mathbf{x}_1\mathbf{x}_2}\mathbf{V}_{\mathbf{x}_2\mathbf{x}_2}^{-1}(\mathbf{x_2} - \boldsymbol{\mu}_2) + \boldsymbol{\epsilon} \tag{30.7a}$$

where the vector $\boldsymbol{\epsilon}$ of residuals (prediction errors) is also MVN,

$$\boldsymbol{\epsilon} \sim \text{MVN}_m\left(\mathbf{0}, \mathbf{V}_{\mathbf{x}_1|\mathbf{x}_2}\right) \tag{30.7b}$$

A direct application of these results is the multivariate breeder's equation, $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$. Assume the vector $\mathbf{z} = (z_1, z_2, \cdots, z_n)^T$ of phenotypic values of characters in an individual can be written as $\mathbf{z} = \mathbf{g} + \mathbf{e}$, the sum of a vector of additive genetic (breeding) values $\mathbf{g}$ plus an independent vector of environmental (and nonadditive genetic) values $\mathbf{e}$. Assuming $\mathbf{x} \sim$ MVN($\boldsymbol{\mu}$, $\mathbf{G}$) and $\mathbf{e} \sim$ MVN($\mathbf{0}$, $\mathbf{E}$), then $\mathbf{z} \sim$ MVN($\boldsymbol{\mu}$, $\mathbf{P}$) where $\mathbf{P} = \mathbf{G} + \mathbf{E}$.

In order to compute the expected change in $\mathbf{z}$ due to selection, consider the distribution of breeding values conditioned on the observed phenotypic value. Since we assume $\mathbf{g}$ and $\mathbf{e}$ are independent,

$$\boldsymbol{\sigma}(\mathbf{g}, \mathbf{z}) = \boldsymbol{\sigma}(\mathbf{g}, \mathbf{g} + \mathbf{e}) = \boldsymbol{\sigma}(\mathbf{g}, \mathbf{g}) = \mathbf{G}$$

Hence the joint distribution of $\mathbf{g}$ and $\mathbf{z}$ becomes

$$\begin{pmatrix}\mathbf{g}\\ \mathbf{z}\end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix}\boldsymbol{\mu}\\ \boldsymbol{\mu}\end{pmatrix}, \begin{pmatrix}\mathbf{G} & \mathbf{G}\\ \mathbf{G} & \mathbf{P}\end{pmatrix}\right) \tag{30.8}$$

In the notation of Equation 30.5, $\mathbf{V}_{\mathbf{g}\,\mathbf{g}} = \mathbf{V}_{\mathbf{g}\,\mathbf{z}} = \mathbf{G}$ and $\mathbf{V}_{\mathbf{z}\,\mathbf{z}} = \mathbf{P}$. From Equations 30.6a/b, the conditional distribution of $\mathbf{g}$ given $\mathbf{z}$ is MVN with mean

$$\boldsymbol{\mu}_{\mathbf{g}\,|\,\mathbf{z}} = \boldsymbol{\mu} + \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}) \tag{30.9a}$$

and variance-covariance matrix

$$\mathbf{V} = \mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{G} \tag{30.9b}$$

Alternatively, this can be restated as the regression of the vector of breeding values on the vector of phenotypic values,

$$\mathbf{g} - \boldsymbol{\mu} = \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}) + \boldsymbol{\epsilon} \tag{30.10a}$$

where the vector of prediction errors is MVN,

$$\boldsymbol{\epsilon} \sim \text{MVN}(\mathbf{0}, \mathbf{V}) \tag{30.10b}$$

Given a vector of phenotypic observations $\mathbf{z}$, the expected vector of breeding values is $\boldsymbol{\mu} + \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu})$, while the actual vector of breeding values is distributed about this mean vector as a Gaussian with covariance matrix $\mathbf{V}$. The variance-covariance matrix of the residual vector $\boldsymbol{\epsilon}$ is independent of the actual value of $\mathbf{z}$, and hence the regression of $\mathbf{G}$ on $\mathbf{z}$ is both

linear (from Equation 30.9a) and homoscedastic (Equation 30.9b). In univariate terms, $g = A$ and $\mathbf{G} = \sigma_A^2$,

$$A - \mu = \sigma_A^2 \sigma_z^{-2}(z - \mu) + \epsilon = h^2(z - \mu) + \epsilon \tag{30.10c}$$

where

$$\sigma_\epsilon^2 = \sigma_A^2 - \sigma_A^2 \sigma_z^{-2} \sigma_A^2 = \sigma_A^2(1 - h^2) \tag{30.10d}$$

Taking expectations over all selected individuals, and assuming that all between-generation changes in character value are due to changes in breeding value,

$$\begin{aligned} \mathbf{R} &= E[\mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}) + \boldsymbol{\epsilon}] \\ &= \mathbf{G}\mathbf{P}^{-1}E[(\mathbf{z} - \boldsymbol{\mu})] + E(\boldsymbol{\epsilon}) \\ &= \mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \mathbf{G}\boldsymbol{\beta} \end{aligned} \tag{30.11}$$

as obtained (in various forms) by Young and Weiler (1960), Harvey and Bearden (1962), Magee (1965), and Lande (1979). It is important to note that all the caveats of the univariate breeder's equation (Table 10.1) also apply to the multivariate breeder's equation.

Example 30.2. What is $\mathbf{G}^*$, the variance-covariance matrix of breeding values after selection (but before recombination and random mating) under the assumptions leading to the multivariate breeder's equation? From the definition of a covariance matrix,

$$\mathbf{G}^* = E\left((\mathbf{g} - \boldsymbol{\mu}^*)(\mathbf{g} - \boldsymbol{\mu}^*)^T \right)$$

where $\boldsymbol{\mu}^*$ is the vector of phenotypic means following selection. Using, respectively, Equation 30.10a, the matrix identity $(\mathbf{ABc})^T = \mathbf{c}^T\mathbf{B}^T\mathbf{A}^T$ (recalling that $\mathbf{G}$ and $\mathbf{P}^{-1}$ are symmetric), and expanding gives

$$\begin{aligned} \mathbf{G}^* &= E\Big(\left[\mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}^*) + \boldsymbol{\epsilon}\right] \left[\mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}^*) + \boldsymbol{\epsilon}\right]^T \Big) \\ &= E\Big(\left[\mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}^*) + \boldsymbol{\epsilon}\right] \left[(\mathbf{z} - \boldsymbol{\mu}^*)^T\mathbf{P}^{-1}\mathbf{G} + \boldsymbol{\epsilon}^T\right] \Big) \\ &= E\Big(\mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}^*)(\mathbf{z} - \boldsymbol{\mu}^*)^T\mathbf{P}^{-1}\mathbf{G} \Big) + E\Big(\mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}^*)\boldsymbol{\epsilon}^T \Big) \\ &\qquad + E\Big(\boldsymbol{\epsilon}(\mathbf{z} - \boldsymbol{\mu}^*)^T\mathbf{P}^{-1}\mathbf{G} \Big) + E\big(\boldsymbol{\epsilon}\,\boldsymbol{\epsilon}^T \big) \end{aligned}$$

Using LW Equation 8.16a and the independence of $\boldsymbol{\epsilon}$ and $\mathbf{z}$, this reduces to

$$\begin{aligned} \mathbf{G}^* = \mathbf{G}\mathbf{P}^{-1}E\big((\mathbf{z} - \boldsymbol{\mu}^*)(\mathbf{z} - \boldsymbol{\mu}^*)^T \big)\mathbf{P}^{-1}\mathbf{G} + \mathbf{G}\mathbf{P}^{-1}E\big(\mathbf{z} - \boldsymbol{\mu}^* \big)E\big(\boldsymbol{\epsilon}^T \big) \\ + E\big(\boldsymbol{\epsilon} \big)E\big((\mathbf{z} - \boldsymbol{\mu}^*)^T \big)\mathbf{P}^{-1}\mathbf{G} + E\big(\boldsymbol{\epsilon}\,\boldsymbol{\epsilon}^T \big) \end{aligned}$$

This can be further simplified by noting that $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and that $E[(\mathbf{z} - \boldsymbol{\mu}^*)(\mathbf{z} - \boldsymbol{\mu}^*)^T] = \mathbf{P}^*$ is the phenotypic variance-covariance matrix after selection. Finally, from Equation 30.9b we have $E(\boldsymbol{\epsilon}\,\boldsymbol{\epsilon}^T) = \mathbf{V}$, giving

$$\mathbf{G}^* = \mathbf{G}\mathbf{P}^{-1}\mathbf{P}^*\mathbf{P}^{-1}\mathbf{G} + \mathbf{0} + \mathbf{0} + (\mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{G})$$

Writing $\mathbf{G}\mathbf{P}^{-1}\mathbf{G} = \mathbf{G}\mathbf{P}^{-1}\mathbf{P}\mathbf{P}^{-1}\mathbf{G}$ and factoring gives the within-generation change in the variance-covariance matrix of breeding values, obtained by Lande and Arnold (1983), as

$$\begin{aligned}\mathbf{G}^* - \mathbf{G} &= \mathbf{G}\mathbf{P}^{-1}\mathbf{P}^*\mathbf{P}^{-1}\mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{P}\mathbf{P}^{-1}\mathbf{G} \\ &= \mathbf{G}\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1}\mathbf{G} \end{aligned} \tag{30.12}$$

While there is little theory for predicting multivariate response in the absence of normality, we can make some general statements under two different sets of assumptions: when all genetic variation is additive and when the multivariate parent-offspring regression is linear. We consider these in turn.

The Multivariate Secondary Theorem of Natural Selection.

One of the more general statements on univariate selection response is Robertson's secondary theorem of natural selection — the rate of change in a character equals the additive genetic covariance between the character and relative fitness (Chapter 5). While this does not hold in general, it is true if all genetic variance in the character is additive, and the error introduced by the presence of dominance is generally small if there are a large number of loci each of small effect (see Equation 5.21). Following Rausher (1992), we can extend Robertson's theorem to selection on multiple characters as follows. Consider the entire collection of multilocus genotypes that influence the vector of characters. Denote the vector of additive genetic values and *relative* fitness of the kth multilocus genotype by $\mathbf{g}(k)$ and $w(k)$, respectively. If $f(k)$ is the frequency of genotype k before selection, then its frequency after selection is $w(k)\, f(k)$. Hence, the within-generation change in the mean additive genetic value is

$$\boldsymbol{\mu}^*_{\mathbf{g}} - \boldsymbol{\mu} = \sum_k \mathbf{g}(k)\, w(k)\, f(k) - \sum_k \mathbf{g}(k)\, f(k)$$

Now consider $\mathbf{c} = \sigma(w, \mathbf{g})$, the vector of covariances between relative fitness and the breeding values of the traits being considered. By definition, this is

$$\begin{aligned}\mathbf{c} = \sigma(w, \mathbf{g}) &= E\left[\, (w - E[w]\,)\, (\mathbf{g} - E[\mathbf{g}]\,)\,\right] \\ &= \sum_k f(k)\left[\, (w(k) - 1)(\mathbf{g}(k) - \boldsymbol{\mu})\,\right] \\ &= \sum_k f(k)\, w(k)\, \mathbf{g}(k) - \boldsymbol{\mu} - \sum_k f(k)\, \mathbf{g}(k) + \boldsymbol{\mu} \\ &= \boldsymbol{\mu}^*_{\mathbf{g}} - \boldsymbol{\mu}\end{aligned}$$

where we have used the fact that, for relative fitness, $E[w] = \sum_k w(k)\, f(k) = 1$. This is an analogue of the Robertson-Price identity $\mathbf{S} = \sigma(w, \mathbf{z})$, except that we are examining the within-generation change in additive genetic, rather than phenotypic, values. If all genetic variation is additive and we can ignore the effects of linkage, then the expected mean of the offspring from these selected parents is their average additive genetic value. In this case, we recover the multivariate version of Robertson's theorem,

$$\mathbf{R} = \mathbf{c} = \sigma(w, \mathbf{g}) \tag{30.13}$$

Equation 30.13 fails when the additive genetic values of parents are not sufficient to predict offspring value (as could occur when epistasis, maternal effects, or genotype $\times$ environment interactions are present).

When Equation 30.13 holds, Rausher (1992) suggested a clever way to decouple the contribution for direct selection on a character from the correlated response due to selection on other (additive) genetically correlated characters. Since $\mathbf{G}\mathbf{G}^{-1} = \mathbf{I}$, rewrite the change in mean as

$$\mathbf{R} = \mathbf{G}\mathbf{G}^{-1}\mathbf{c} = \mathbf{G}\mathbf{b} \tag{30.14a}$$

where

$$\mathbf{b} = \mathbf{G}^{-1}\mathbf{c} \tag{30.14b}$$

Denoting the additive genetic covariance between characters i and j by g_{ij}, the net change in character i is

$$R_i = b_i\, G_{ii} + \sum_{j \neq i} b_j\, G_{ij} \tag{30.15}$$

where b_i denotes the ith element of $\mathbf{b}$. Thus $b_i\, G_{ii}$ is the change due to direct selection on character i, while the sum is the net change generated by selection on genetic correlated characters. $\mathbf{b}$ is the additive-genetic analogue of $\boldsymbol{\beta}$ (when the conditions leading to the breeder's equation hold they are in fact equal). Since $\mathbf{c}$ is the vector of covariances between additive genetic values of the characters and relative fitness and $\mathbf{G}$ the covariance matrix of these additive genetic values, $\mathbf{b}$ is the vector of partial regression coefficients of fitness on additive genetic (breeding) values. Hence, if $\mathbf{g}$ is the vector of additive genetic values of an individual, $\mathbf{b}$ are the coefficients of the best-fit linear regression of relative fitness on genotype value,

$$w(\mathbf{g}) = 1 + \sum b_i(g_i - \mu_i) \tag{30.16}$$

where g_i is the additive genetic value in that individual for character i.

Response When the Parent-offspring Regression is Multivariate Linear

When the parent-offspring regression is linear, we also obtain a simple form for the single-generation response to selection. First, we define what we mean by a linear regression when multiple characters are considered in the obvious fashion: if n characters are measured and y_i is the mid-parental value of the ith character, we say that the parent-offspring regression is (multivariate) linear if the value of character j in the offspring is given by

$$\mu_j + h_{1j}(y_1 - \mu_1) + h_{2j}(y_2 - \mu_2) + \cdots + h_{nj}(y_n - \mu_n) + e_j$$

where the residual e_j has expected value zero. In matrix form, if $\mathbf{z}$ is the vector of offspring means and $\mathbf{y}$ the vector of midparent values, then $\mathbf{z} = \boldsymbol{\mu} + \mathbf{H}(\mathbf{y} - \boldsymbol{\mu}) + \mathbf{e}$ where $\mathbf{H}$ is the (symmetric) matrix of partial regression coefficients. Taking expectations over the selected parents gives the expected change in the offspring mean as

$$\boldsymbol{\Delta\mu} = E[\mathbf{H}(\mathbf{y} - \boldsymbol{\mu}) + \mathbf{e}] = \mathbf{H}\,(E[\mathbf{y}] - \boldsymbol{\mu}) = \mathbf{HS} \tag{30.17}$$

provided that the selection of parental values does not influence the matrix of regression coefficients $\mathbf{H}$. If the parent-offspring regression is nonlinear, the response is not necessarily predictable from $\mathbf{S}$, but rather can depend on additional phenotypic moments of the selected population (see Figure 10.1). When all genetic variation is additive, **Equations 30.38** and **30.42** imply $\mathbf{HS} = \mathbf{c}$, or

$$\mathbf{S} = \mathbf{H}^{-1}\mathbf{c} \tag{30.18}$$

When the joint distribution of phenotypic and breeding (additive genetic) values is multivariate normal, Equation 31.9a shows that the regression of additive genetic values $\mathbf{g}$ for an individual given its observed vector $\mathbf{z}$ of phenotypes is $\boldsymbol{\mu}_{\mathbf{g}\,|\,\mathbf{z}} = \boldsymbol{\mu} + \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu})$. Under

the standard genetic assumptions of the breeder's equation (no epistasis, maternal effects, genotype $\times$ environment interactions, etc.) parents only pass on additive genetic effects to their offspring, so that the expected mean value of an offspring is the mean breeding value of its parents. In this case, we can decompose the matrix of parent-offspring regression coefficients into separate genetic and phenotypic components, viz., $\mathbf{H} = \mathbf{GP}^{-1}$. This decomposition allows us to consider the effects of genetic and phenotype correlations independently (in the general setting, $\mathbf{H}$ confounds genetic and phenotypic effects). Further, Equation 31.9b shows that in this case the regression is not only linear, it is also homoscedastic, so that selection of any set of individuals does not influence the parent-offspring regression, as first noticed by Pearson (1903).

BIVARIATE SELECTION

We now restrict attention to selection on just two traits. This allows us to both develop several important concepts for multi-trait selection, as well as to present several common expressions found in the literature. We start with the simplest case of selection on just one trait before considering the more general case of jointly selecting on both.

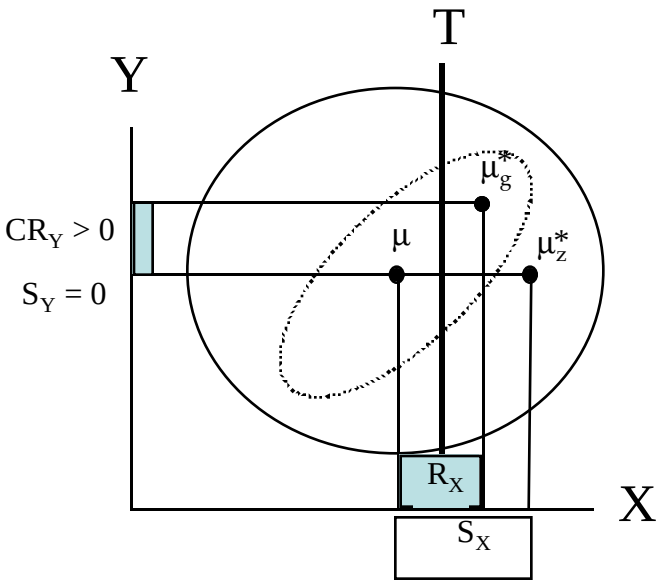


Figure 30.1. The differential effects of phenotypic and genetic correlations on selection response. The solid circle is the joint distribution of the phenotypes for traits X and Y, the dotted ellipse the distribution of their breeding values. Truncation selection occurs only on X, with all individuals with $X > T$ saved, with μ the mean before selection, while μ_z^* and μ_g^* are the mean phenotypic and breeding values after selection. Here X and Y are phenotypically uncorrelated, so that while there is a selection differential on X, the differential on Y is zero. Compare this with Figure 29.1, which shows different selection differentials introduced when the traits are phenotypically correlated. For our case, although there is no *phenotypic* correlation, there is a *genetic* correlation. Following selection, the mean breeding values of *both* X and Y have changed, resulting in a selection response in both. Thus, the selection differential on Y is zero, but we still see a response, a result entirely due to the *genetic* correlation between the two traits.

Correlated Response to Selection

Suppose that direct selection is occurring only on trait X. What happens to an unselected, but correlated, trait Y? We follow historical (e.g., Falconer 1960a) notation and let R_X denote the direct response in X, namely selection is directly on X. Likewise denote the correlated response by CR_Y, the response in Y given direct selection on X. Again, a key point is that *phenotypic and genetic covariances have very different consequences.* With selection only on X, its response is just $R_X = h^2_X S_X$. What about the response in Y? If we ignore the multivariate nature of the problem, given the selection differential S_Y, one might think this is just $h^2_Y S_Y$. This expression, however, is incorrect. Selection *is not* occurring on Y, rather any selection differential S_Y arises from a phenotypic correlation with the trait under direct selection. The actual *response* in Y does *not* simply depend upon its heritability — no matter the magnitude of S_Y or the size of h^2_Y, there is *no* response in Y unless there is a **genetic correlation** between X and Y. Figure 30.1 attempts to make this point, showing a situation where the selection differential on Y is zero, yet we still see a response in Y.

Example 30.3. Godfray and colleagues (Fellowes et al. 1998, 1999a, 1999b; Kraaijeveld and Godfray 1997; Kraaijeveld et al. 2001) examined correlated responses in fitness following directional selection for parasitoid resistance in *Drosophila melanogaster.* Two hymenopterian parasitoids, the Braconid wasp *Asobara tabida* and the Eucoilid wasp *Leptopilina boulardi*, both attack *melanogaster* eggs. In their founding population, only a small fraction of *melanogaster* eggs (5% for *tabida* and 0.4% for *boulardi*) were able to survive attacks by these parasitoids. However, five generations of selection for increased resistance jumped these survival probabilities to 45% (for *boulardi*) and 60% (for *tabida*). Interestingly, there was an asymmetry in the transference of this resistance. The line selected for *boulardi* resistance also showed good resistance against *tabida*. However, the line selected for *tabida* resistance showed only a slight increase in survival against *boulardi*. Both selected lines generally showed no differences in fitness (relative to the unselected control), with the important exception of both lines showing lower competitive ability (due to reduced larval feeding) under high stress (high larval density) conditions. Thus, there was a tradeoff between parasitoid survival on one hand and larval competition on the other, which might explain the generally low natural levels of resistance.

When characters are genetically correlated (i.e., trait breeding values are correlated within an individual), selection solely on one trait results in a **correlated response** in the other. In the above example, beneficial increases in the resistance to other parasitoids arose as a correlated response, but so did detrimental reductions in competitive ability.

The expected correlated response can be obtained as follows. The response to selection of character X (the mean value of offspring from the selected parents) is (by definition) the mean breeding value of the selected parents (Chapter 10). Thus the change in character Y equals the change in its mean breeding value caused by selection on X, which is given from the regression of A_Y on A_X. For standard regression theory (LW Chapter 3), the slope of this regression is given by

$$b_{A_Y|A_X} = \frac{\sigma(A_X, A_Y)}{\sigma^2_{A_X}} = \frac{r_A\, \sigma_{A_X}\, \sigma_{A_Y}}{\sigma^2_{A_X}} = r_A\, \frac{\sigma_{A_Y}}{\sigma_{A_X}} \tag{30.19a}$$

where r_A is the correlation among breeding values (the genetic correlation). Recall that a regression passes through the mean of the dependent (A_Y) and independent (A_X) variables (LW Chapter 3), both of which (by definition) are zero before selection. Hence, our regression

passes through the origin, and has expected value

$$A_Y = b_{A_Y|A_X}\, A_X = r_A\, \frac{\sigma_{A_Y}}{\sigma_{A_X}}\, A_X \tag{30.19b}$$

For making comparisons, it will often prove useful to express the direct response $R_X = h_X^2 S_X$ in terms of the selection intensity $\overline{\imath}_X = S_X/\sigma_X$, giving

$$R_X = \overline{\imath}_X\, h_X^2\, \sigma_{P_X} = \overline{\imath}_X\, h_X\, \sigma_{A_X} \tag{30.20}$$

where last expression follows from $h_X \sigma_{P_X} = \sigma_{A_X}$. Recalling our comment above that R_X is the change in the breeding value of X, the correlated response of character Y is just its change in mean breeding value given the change R_X,

$$\begin{aligned} CR_Y &= b_{A_Y|A_X} R_X \\ &= (r_A\, \sigma_{A_Y}/\sigma_{A_X})\, (\overline{\imath}_X\, h_X\, \sigma_{A_X}) \\ &= r_A\, \sigma_{A_Y}\, \overline{\imath}_X\, h_X \qquad (30.21a) \\ &= \overline{\imath}_X\, h_X\, h_Y\, r_A\, \sigma_{P_Y} \qquad (30.21b) \end{aligned}$$

A result obtained by Falconer (1952). Equation 30.21b follows from $\sigma_{A_Y} = h_Y \sigma_{P_Y}$. Noting that the direct response on X is $R_X = \overline{\imath}_X\, h_X^2\, \sigma_{P_X}$, we see that h_X^2 and $h_X\, h_Y\, r_A$ play similar roles, resulting in the later occasionally being called the **co-heritability** (Janssens 1979). An important feature to note from Equation 30.21b is that the phenotypic correlation between X and Y *never enters* into the response in Y. If X and Y are *genetically* correlated, we observe a response in Y, even (in the extreme case) when $S_Y = 0$ (e.g., Figure 30.1).

Example 30.4. A classic application of the theory of correlated response is Lande's 1979 paper on the evolution of the brain-body size relationship in mammals. These show an **allometric** relationship (LW Chapter 11), with a linear regression of log brain size on log body size,

$$\ln(\text{brain size}) = \alpha \ln(\text{body size}) + b$$

The slope α of this regression is typically 0.2 to 0.4 when based on closely-related species, but much larger (around 0.67) when the regression uses more distantly-related taxa. Given that there is often selection for increased body size (see our discussion on Cope's law in Chapter 29), Lande was interested in whether this allometric pattern could simply be a consequence of a correlated response in brain size from direct selection on body size. The notion of allometric patterns arising for other traits as a result of a correlated response to selection on size was apparently first suggested by Reeve (1950). Transforming these two traits to a log scale, Equations 30.21b and 30.20 give the ratio of correlated responses in (log) brain (lbr) to direct selection on (log) body size (lbd) as

$$\frac{CR_{lbr}}{R_{lbd}} = \frac{\overline{\imath}_{lbd}\, r_A\, h_{lbr}\, h_{lbd}\, \sigma_{lbr}}{\overline{\imath}_{lbd}\, \sigma_{lbd}\, h_{lbd}^2} = r_A\, \frac{\sigma_{lbr}\, h_{lbr}}{\sigma_{lbd}\, h_{lbd}}$$

This also follows directly from Equation 30.4b. Note that this ratio corresponds to the slope α, as both traits have been log-transformed. Using values form the mouse literature, Lande found

$$r_A\, \frac{\sigma_{lbr}\, h_{lbr}}{\sigma_{lbd}\, h_{lbd}} = 0.68\, \frac{0.058 \cdot \sqrt{0.64}}{0.145 \cdot \sqrt{0.37}} = 0.36$$

Lande thus conclude that the allometric relationship within closely-related taxa (α 0.2 to 0.4) could be accounted for by a correlated response to selection on body size. He also noted that random drift would also give this result, as the means of the two traits, being correlated, with also drift together (Chapters 6, 31). However, Lande noted that this α value is much smaller than seen when the regression is computed using more distantly-related taxa. He suggested that over short time scales (those seen within closely-related taxa), any selection on brain size might be indirect, but over longer time scales selection might occur directly on brain size.

Riska and Atchley (1985) provided some additional clarification on this issue. They note that increases in size result from two different developmental mechanisms: increases in cell number (**hyperplastic responses**) and increases in cell size (**hypertrophic responses**). Most early growth is hyperplastic, while most later growth is hypertrophic. They conjectured that hyperplastic responses result in larger brain-body size correlations than hypertrophic responses. If their idea is correct, brain-body genetic correlations should be stronger in early life than in later life, and this was indeed observed in their mouse experiments. They suggest that much of the divergence among closely-related taxa occurs via changes in cell volume, and hence would have a lower genetic correlation, while changes among higher taxonomic groups occurs via changes in cell numbers. (They mention the fun fact that elephant cells are only twice as large as mouse cells, showing that most of the difference in size is due to differences in numbers.) Under their model, it is possible that *both* patterns of allometry seen across the two different taxonomic scales could be consistent with a simple correlated response to selection on body size, which was shown in later simulation studies by Riska (1989).

Example 30.5. Of course, Equations 30.20 and 30.21 can easily be recovered using the multivariate breeder's equation. For notation ease, let 1 denote the trait under direct selection and 2 the correlated trait, so that $\beta_1 \neq 0, \beta_2 = 0$. From Equation 30.3a, the selection differentials become

$$S_1 = P_{11}\beta_1 + P_{12}\beta_2 = P_{11}\beta_1$$

$$S_2 = P_{21}\beta_1 + P_{22}\beta_2 = P_{21}\beta_1$$

Thus $\beta_1 = S_1/P_{11}$, giving the (correlated) selection differential trait on 2 as

$$S_2 = S_1\frac{P_{21}}{P_{11}} = S_1\frac{\sigma(z_1, z_2)}{\sigma^2(z_1)}$$

Likewise, Equation 30.3b gives the response as

$$R_j = G_{1j}\beta_1 + G_{2j}\beta_2 = G_{1j}\beta_1$$

Hence

$$R_1 = G_{11}\beta_1 = G_{11}\frac{S_1}{P_{11}} = S_1 h_1^2$$

and we recover the breeder's equation for the trait under direct selection. Recalling that $S_1 = \overline{i}_1\sigma(z_1)$ recovers Equation 30.20. Turning to the correlated response,

$$R_2 = G_{21}\beta_1 = G_{21}\frac{S_1}{P_{11}} = S_1\frac{\sigma(A_1, A_2)}{\sigma^2(z_1)}$$

A little rearrangment gives

$$R_2 = \frac{S_1}{\sigma(z_1)}\frac{\sigma(A_1, A_2)}{\sigma(z_1)} = \overline{i}_1\frac{\rho_A\sigma(A_1)\sigma(A_2)}{\sigma(z_1)}\frac{\sigma(z_2)}{\sigma(z_2)} = \overline{i}_1\,\rho_A\,h_1\,h_2\,\sigma(z_2)$$

recovering Equation 30.21b.

Indirect Selection May Give a Larger Response Than Direct Selection

In applied breeding, one usually focuses on directly selecting on the trait of interest to improve its value. However, Equation 30.21b suggests that, in some settings, we can actually

get a *greater* response by instead selecting on a trait *correlated* with X, rather than on X itself (Lerner and Cruden 1948, Falconer 1952; Turner 1959; Searle 1965, 1978). The ratio of response R_X from directly selecting on trait X (with selection intensity $\overline{\imath}_X$) to the correlated response CR_X in X from selection on Y (with selection intensity $\overline{\imath}_Y$) is

$$\frac{CR_X}{R_X} = \frac{\overline{\imath}_Y\, r_A\, \sigma_{A_X} h_Y}{\overline{\imath}_X\, h_X\, \sigma_{A_X}} = \frac{\overline{\imath}_Y\, r_A\, h_Y}{\overline{\imath}_X\, h_X} \tag{30.22a}$$

a result due to Lerner and Cruden (1948) and Falconer (1952). Lerner and Cruden defined the measure

$$r_A\, \frac{h_Y}{h_X} \tag{30.22b}$$

as the **relative efficiency of indirect selection**, and Searle (1965) provides an approximation for its standard error. Note that the correlated "trait" Y might actually be an index of traits, potentially even including X, e.g. $Y = aZ + bX$ where Z is a genetically correlated trait. Searle examined this case in some detail, and we return to it in Chapter 33.

From Equation 30.22a, the correlated response of X to selection for Y will be greater than direct response to selection for X when $\overline{\imath}_Y\, r_A\, h_Y > \overline{\imath}_X\, h_X$, or when

- character Y has a greater heritability than X, and the genetic correlation between X and Y is high. This could occur if X is difficult to measure with precision but Y is not.
- the selection intensity is much greater for Y than X. This would be true if Y were measurable in both sexes but X measurable in only one sex.

Example 30.6. Young et al. (1965) examined body weight at weaning and at 15-16 months of age in Australian merino sheep. While merinos are known for their outstanding wool characteristics, the authors were interested in their potential for improved meat production, as measured by weight at weaning. Unfortunately, estimates of the heritability of body weight at weaning were very low (around 0.10). However, estimates for body weight at 15-16 months were considerably higher (0.6), and the estimated genetic correlation between these two traits was above 0.90. Assuming the same selection intensity, the predicted ratio of correlated response (weaning weight response when selecting for weight at 15-16 months) to direct response (weaning weight when weaning weight is directly selected) is

$$\frac{CR}{R} = \frac{0.9\,\sqrt{0.6}}{\sqrt{0.1}} = 2.2$$

For the same selection intensity, selecting on 15-16 month weight is expected to yield over twice the response versus directly selecting on weaning weight. Of course there are tradeoffs with waiting until an age of 15-16 months, but the larger expected gain may offset these.

Realized Genetic Correlations

Recall (Chapter 14) that we used selection experiments to compute a realized heritability of a trait, in the simplest form by $h^2 = R/S$. Likewise, inspection of Equations 30.20 and 30.21 suggests we can similarly obtain a selection-based estimate of the genetic correlation r_A. Consider the ratio of CR_X/R_X, the correlated response in X (given direct selection on Y) to the direct response in X. From Equations 30.20 and 30.21a,

$$\frac{CR_X}{R_X} = \frac{\overline{\imath}_Y}{\overline{\imath}_X}\,\frac{h_Y}{h_X}\, r_A \tag{30.23a}$$

Rearranging to solve for the genetic correlation yields

$$r_A = \frac{CR_X}{R_X}\,\frac{\overline{\imath}_X}{\overline{\imath}_Y}\,\frac{h_X}{h_Y} \tag{30.23b}$$

A similar expression follows based on CR_Y/R_Y, the correlated to direct response in Y. To apply Equation 30.23b, we need at least two selected lines, one the result of directly selecting on X, the other directly selecting on Y (**reciprocal-selection**). Such data yield two separate estimates of r_A, one based on X (CR_X/R_X) and the other on Y (CR_Y/R_Y). In theory, both estimates should return similar values (subject, of course, to sampling error), provided selection is for a single generation from an unselected base population. When data from more than a single generation of response is used, we might expect these estimates to differ, as they are based on populations likely to diverge over time. Specifically, a population directly selected for X will likely have differences in the allele frequencies and linkage disequilibria for loci controlling X relative to a population directly selected for Y. Comparison of these two estimates provides a test of consistency of the realization correlation (Example 30.7).

Often, we do not have reciprocal-selection data, but instead have results from only a single selected line: a direct response R_X in X and a correlated response CR_Y in Y. Applying Equation 30.20 and 30.21 to the ratio CR_Y/R_X and rearranging, we find that the common selection intensity $\overline{\imath}_X$ cancels, but the different phenotypic variances do not, yielding

$$r_A = \frac{CR_Y}{R_X}\,\frac{h_X\sigma_{P_X}}{h_Y\sigma_{P_Y}} \tag{30.23c}$$

Finally, use of Equations 30.20 and 30.21 also shows that

$$\frac{CR_X}{R_X}\,\frac{CR_Y}{R_Y} = \left(\frac{\overline{\imath}_Y}{\overline{\imath}_X}\,\frac{h_Y}{h_X}\,r_A\right)\left(\frac{\overline{\imath}_X}{\overline{\imath}_Y}\,\frac{h_X}{h_Y}\,r_A\right) = r_A^2 \tag{30.24a}$$

suggesting an alternative (or **aggregate**) estimator

$$r_A = \sqrt{\frac{CR_X\,CR_Y}{R_X\,R_Y}} \tag{30.24b}$$

The advantage of the aggregate estimator is that we do not need the heritabilities nor selection intensities. The downside is that this estimator can disguise the fact that estimators based on reciprocal data (Equation 30.23b) can yield very different values from each other.

Example 30.7. Falconer (1960b) examined mice growth rate (measured as change in weight from week three to week six) in two different environments: high (H) and low (L) nutrition. Directional selection for both increased and decreased growth was practiced in both environments. Correlated responses were measured by rearing a second litter in the opposite environment. The results from the first 13 generations of selection are summarized below, using the divergence between the up- and down-selected lines in each environment.

Generations	0 – 4		4 – 13		0 – 13	
	H	L	H	L	H	L
Selection intensity, $\overline{\imath}$	1.66	1.40	1.65	1.52	1.66	1.48
Realized heritability, h_r^2	0.41	0.36	0.22	0.13	0.30	0.20
Direct response, R	0.90	1.20	0.53	0.46	0.69	0.79
Correlated response, CR	0.48	0.98	0.46	-0.01	0.38	0.38
Realized correlation, r_A	0.67	0.69	1.25	-0.02	0.74	0.39

Realized genetic correlations in the table were calculated using Equation 30.23b. For example, using responses (direct and correlated) for H as our reference, the response over the first four generations gives

$$r_A = \frac{CR_H}{R_H}\frac{\overline{i}_H}{\overline{i}_L}\frac{h_H}{h_L} = \frac{0.48}{0.90}\frac{1.66}{1.49}\sqrt{\frac{0.41}{0.36}} = 0.67$$

Using L as our reference gives

$$r_A = \frac{CR_L}{R_L}\frac{\overline{i}_L}{\overline{i}_H}\frac{h_L}{h_H} = \frac{0.98}{1.20}\frac{1.49}{1.66}\sqrt{\frac{0.36}{0.41}} = 0.69$$

These two estimates of the genetic correlation between H and L are in excellent agreement. Note, however, that this is not the case for estimates based on using data after generation four. If we used Equation 30.24b to estimate the genetic correlation (say using the generation 0-13 data), we would obtain

$$r_A = \sqrt{\frac{CR_L}{R_H}\frac{CR_H}{R_L}} = \sqrt{\frac{0.38}{0.69}\frac{0.38}{0.79}} = 0.52$$

In reality, the reciprocal estimates yield very different estimates (0.74 and 0.39), something we would have missed by just using the aggregate estimator.

Strictly speaking, the estimators for r_A given by Equations 30.23b,c and 30.24a assume only a single generation of selection from an unselected base population. When using data from more than one generation (as in Example 30.6), we replaced single generation responses and selection intensities with cumulative responses and intensities. Recall our extensive discussion in Chapter 14 on how to best incorporate multi-generation information into an estimate of the realized heritability. While cumulative values could be used, one could also use regression-based estimators. For the case of correlated responses, these are the regression of the cumulative correlated response CR_Y on S_X, the cumulative selection differential in X. See Atkins and Thompson (1986a,b) for a worked example.

Despite our ability to compute realized correlations, the general suggestion in the literature is to treat such estimates with great caution, with relative-based estimates generally preferred over realized estimates. As we detail in Chapter 31, genetic correlations are much more unstable than genetic variances in that the later are somewhat robust to small changes in allele frequencies, while the former can be quite sensitive to them (Bohren et al. 1966). Thus, while realized heritability estimates seem to be somewhat robust, the same cannot be said about realized genetic correlations.

General Bivariate Selection

The above discussion concerns the special case of direct selection on just a single trait. From the standpoint of the multivariate breeder's equation $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$, this implies $\beta_X \neq 0$, $\beta_Y = 0$. We now consider the general case where both traits are under direct selection. Traits are labeled as z_1 and z_2 to emphasize that neither is more special than the other (in contrast to above where we highlighted the selected trait).

For the bivariate case, there are two correlations: r_A and r_z, the correlations among breeding and phenotypic values. The genetic covariance thus becomes $G_{12} = r_A\sigma_{A_1}\sigma_{A_2}$, and Equation 30.3b gives the response as

$$R_1 = \sigma^2_{A_1}\beta_1 + r_A\sigma_{A_1}\sigma_{A_2}\beta_2 \tag{30.25a}$$

$$R_2 = r_A\sigma_{A_1}\sigma_{A_2}\beta_1 + \sigma^2_{A_2}\beta_2 \tag{30.25b}$$

To express $\boldsymbol{\beta}$ in terms of the *observed* selection differentials $\mathbf{S}$, recall that $P_{12} = r_z \sigma_{z_1} \sigma_{z_2}$, giving

$$\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S} = \begin{pmatrix} \sigma_{z_1}^2 & r_z \sigma_{z_1} \sigma_{z_2} \\ r_z \sigma_{z_1} \sigma_{z_2} & \sigma_{z_2}^2 \end{pmatrix}^{-1} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \tag{30.26a}$$

Noting that the determinant of the $\mathbf{P}$ matrix is just

$$\det(\mathbf{P}) = \sigma_{z_1}^2 \sigma_{z_2}^2 - r_z^2 \sigma_{z_1}^2 \sigma_{z_2}^2 = \sigma_{z_1}^2 \sigma_{z_2}^2 \left(1 - r_z^2\right)$$

and using the standard expression (LW Equation 8.11) for the inverse of a 2×2 matrix,

$$\begin{pmatrix} \sigma_{z_1}^2 & r_z \sigma_{z_1} \sigma_{z_2} \\ r_z \sigma_{z_1} \sigma_{z_2} & \sigma_{z_2}^2 \end{pmatrix}^{-1} = \frac{1}{\det(\mathbf{P})} \begin{pmatrix} \sigma_{z_2}^2 & -r_z \sigma_{z_1} \sigma_{z_2} \\ -r_z \sigma_{z_1} \sigma_{z_2} & \sigma_{z_1}^2 \end{pmatrix}$$

which, upon some algebraic simplification (see LW Example 8.3), gives

$$\beta_1 = \frac{\overline{\imath}_1 - \overline{\imath}_2 r_z}{1 - r_z^2}\, \sigma_{z_1}^{-1}, \qquad \text{and} \qquad \beta_2 = \frac{\overline{\imath}_2 - \overline{\imath}_1 r_z}{1 - r_z^2}\, \sigma_{z_2}^{-1} \tag{30.26b}$$

where $\overline{\imath}_i = S_i / \sigma_{z_i}$ is the standardized selection intensity on z_i. Substitution of 30.26b into 30.25b gives the general expression for bivariate response under any genetic and phenotypic correlation structure. Let's examine a few special cases.

First, suppose there are no phenotypic correlations ($r_z = 0$), giving $\beta_i = \overline{\imath}_i / \sigma_{z_i}$ and

$$R_1 = h_1^2\, S_1 + r_A \sigma_{A_1} \sigma_{A_2} \overline{\imath}_2 / \sigma_{z_2} \tag{30.27a}$$

$$R_2 = h_2^2\, S_2 + r_A \sigma_{A_1} \sigma_{A_2} \overline{\imath}_1 / \sigma_{z_1} \tag{30.27b}$$

showing the corrections to the univariate breeder's equation when genetic correlations exist.

Conversely, when there is no correlation between the additive genetic (breeding) values of the two traits within an individual ($r_A = 0$),

$$R_i = h_i^2\, \sigma_{z_i}^2\, \beta_i \tag{30.28}$$

If $\overline{\imath}_2 = 0$ (there is no *apparent* selection on z_2), this reduces to

$$R_1 = \frac{\overline{\imath}_1\, h_1^2 \sigma_{z_1}}{1 - r_z^2} \geq \overline{\imath}_1\, h_1^2 \sigma_{z_1} \tag{30.29}$$

Note that the right hand side is just the response under the univariate breeder's equation (Equation 10.6b), so the effect of phenotypic correlation ($r_z \neq 0$) is to *increase* the response relative to the case of no correlation. At first, this seems rather odd, but makes sense when considered in terms of $\boldsymbol{\beta}$. From Equation 30.26b, $\overline{\imath}_2 = 0 = r_z \sigma_{z_1} \beta_1 + \sigma_{z_2} \beta_2$, or $\beta_2 = -\beta_1 r_z \sigma_{z_1} / \sigma_{z_2}$. Thus even though $\overline{\imath}_2 = 0$, there is direct selection on z_2 which is hidden by the phenotypic correlation between z_1 and z_2. When there is no direct selection on z_2, then $\boldsymbol{\beta}_2 = a_i / \sigma_{z_i} = 0$, and (from Equation 30.26b) $\overline{\imath}_2 = r_z \overline{\imath}_1$.

Finally, writing $\sigma_{A_i} = h_i \sigma_{z_i}$, response for the general case with both genetic *and* phenotypic correlations can be written as

$$R_1 / \sigma_{z_1} = h_1^2\, a_1 + r_A h_1 h_2\, a_2 \tag{30.30a}$$

$$R_2 / \sigma_{z_2} = h_2^2\, a_2 + r_A h_1 h_2\, a_1 \tag{30.30b}$$

where $a_i = \beta_i \sigma_{z_i}$, viz.,

$$a_1 = \frac{\overline{i}_1 - r_z \overline{i}_2}{1 - r_z^2} \qquad \text{and} \qquad a_2 = \frac{\overline{i}_2 - r_z \overline{i}_1}{1 - r_z^2} \tag{30.30c}$$

as obtained by Harvey and Bearden (1962) and Bell and Burris (1973).

Realized Correlations with Bivariate Selection

We have previous seen how to estimate realized genetic correlations when selection is acting on only a *single* trait and a correlated response is observed (Equations 30.23 and 30.25). How do we handle cases where we select on *both* traits (for example, by including each in a selection index)? The basic approach was developed by Harvey (1972), Berger and Harvey (1975), and Berger (1977) for the case of index selection, but we can easily generalized their results using Equation 30.30.

Suppose we perform two (or more) selection experiments from the same base populations. For experiment i, let a_{ij} denote the a value for trait j as given by Equation 30.30c and let R'_{ij} denote its (scaled) response, i.e. $R'_{ij} = R_{ij}/\sigma(z_j)$. If the same base population is used, then the values of h_1^2, h_2^2 and (most importantly) r_A are the same for both experiments. Again, recall our comments above that this applies only strictly in the first generation of selection from a common, and unselected, base population. Define the matrices

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{11} & a_{12} \\ a_{21} & a_{22} & 0 \\ 0 & a_{21} & a_{22} \end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} h_1^2 \\ r_A\, h_1\, h_2 \\ h_2^2 \end{pmatrix}, \quad \mathbf{R}' = \begin{pmatrix} R'_{11} \\ R'_{12} \\ R'_{21} \\ R'_{22} \end{pmatrix} \tag{30.31a}$$

The predicted response to both experiments can be compactly written in matrix form using Equation 30.30a/b, yielding

$$\mathbf{Ag} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{11} & a_{12} \\ a_{21} & a_{22} & 0 \\ 0 & a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} h_1^2 \\ r_A\, h_1\, h_2 \\ h_2^2 \end{pmatrix} = \begin{pmatrix} R'_{11} \\ R'_{12} \\ R'_{21} \\ R'_{22} \end{pmatrix} = \mathbf{R}' \tag{30.31b}$$

The reader can easily verify that each line of this equation recovers Equation 30.30a/b. Since there are fewer unknowns than observations, we solve for the vector of genetic parameters $\mathbf{g}$ using least-squares to give

$$\mathbf{g} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1} \mathbf{A}^T\mathbf{R}' \tag{30.31c}$$

This follows since $\mathbf{A}$ plays the same role as the design matrix $\mathbf{X}$ in a linear model and the observed vector of responses $\mathbf{R}'$ plays the same role as the vector of observations $\mathbf{y}$.

Bruns and Harvey (1976) and Gunsett et al. (1984) examined how the choice of a values influences the estimates, noting that equal selection both traits is most effective. Gunsett et al. frame this problem in an experimental design context, given the parallels of $\mathbf{A}$ with the design matrix $\mathbf{X}$, and find more general solutions. When index selection is used, Berger and Harvey (1975) and Berger (1977) note that care must be taken to correctly assign the a_{ij}, which should be estimated after the fact from the actual selection differentials (as opposed to the values used in the original index), see their papers (and Chapter 32) for details. A major deficiency of this approach is obtaining standard errors for the estimates. Rutledge et al. (1973) used replicated selection lines, and a start at some theory was presented by Gunsett et al. (1982).

This approach easily extends any number of traits. Let R_{ij} and β_{ij} be the response and selection gradient for trait j in experiment i. From Equation 30.3b, the response is just

$$R_{ij} = G_{jj}\beta_{ij} + \sum_{k \neq j}^{n} G_{jk}\beta_{ik}$$

As above, we can gather up all of the responses from multiple experiments in matrix form (again assuming all experiments start from the same unselected based population). For example, with three traits and two experiments,

$$\begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} & 0 & 0 & 0 \\ 0 & \beta_{11} & 0 & \beta_{12} & \beta_{13} & 0 \\ 0 & 0 & \beta_{11} & 0 & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} & 0 & 0 & 0 \\ 0 & \beta_{21} & 0 & \beta_{22} & \beta_{23} & 0 \\ 0 & 0 & \beta_{21} & 0 & \beta_{22} & \beta_{23} \end{pmatrix} \begin{pmatrix} G_{11} \\ G_{12} \\ G_{13} \\ G_{22} \\ G_{23} \\ G_{33} \end{pmatrix} = \begin{pmatrix} R_{11} \\ R_{12} \\ R_{13} \\ R_{21} \\ R_{22} \\ R_{23} \end{pmatrix}$$

Here β_{ij} and R_{ij} refer to trait j in experiment i, while $G_{k\ell}$ is the genetic covariance between traits k and ℓ, which is assumed common over all experiments. As above, we solve for the vector of common genetic variances/covariances using least squares. With n traits, there are $n(n+1)/2$ genetic parameters to estimate. The number of independent selection experiments thus needed is $kn \geq n(n+1)/2$ or $k \geq (n+1)/2$. Hence, for three traits, two experiments are sufficient, but with 10 traits, six separate selection experiments (independent sets of $\boldsymbol{\beta}$) are needed.

Asymmetric Correlated Responses are Frequently Seen

Recall from Chapter 14 the notion of an asymmetric selection response, wherein the realized heritabilities differed in up- versus down-selected lines. The same phenomena is also seen, and more frequently, with correlated responses. Falconer's (1960) mouse experiments (Example 30.7) showed very different estimates of the realized genetic correlation occur when we change which trait is under direct selection. Even if we restrict selection to a single trait, asymmetric correlated responses can occur. For example, Clayton et al. (1957) observed no correlated responses in sternopleural bristle number when *Drosophila* lines were selected for decreased sternital bristles, but a strong correlated response when the lines were up-selected. Gromko et al (1991) also observed unpredictability in correlated responses in *Drosophila*. Male copulation duration is genetically correlated with courtship vigor ($r_A = -0.41$) and fertility ($r_A - 0.27$). Correlated responses in these last two traits were followed in eight lines selected for copulation duration (four up, four down). While all eight replicate lines showed consistent behavior in the response to direct selection, only 3/8 had the predicted correlated responses. The general picture from the literature seems to be that correlated responses are far less predictable than direct responses. As will be discussed in Chapter 31, this is not unexpected, as there are several genetic reasons why correlations would be more fragile than genetic variances (Bohren et al. 1966).

COMPARISON OF MULTIVARIATE RESPONSES

Returning now to the full breeder's equation, we are often interested in comparing responses, both across characters and across different populations/species. Different traits are usually measured in different units, and may also vary dramatically between populations/species.

Thus, standardization of these measurements is required in order for meaningful comparisons. Recall we briefly discussed standardization of $\boldsymbol{\beta}$ values in Chapter 29. We develop this a bit more fomrally here before turning our focus to estimating realized selection gradients.

Standarization of Response

The idea behind standardization is to place very different sets of measurements (from different traits or different populations) on a common scale for comparison. There is no single standardization that is uniformly best (Houle 1992, Hansen et al. 2003a, Hansen and Houle 2008), as standardization depends on both the nature of the trait and the type of comparison desired. For example, in quantitative genetics, some traits have a natural zero value (height, weight, etc.), whereas other (calendar date of first flowering) can have a rather arbitrary zero value. This affects the choice of standardization. Likewise, the desired type of comparison is also critical in setting the scale. In quantitative genetics, the raw data is typically standardized for comparisons on one of three scales: interval, log-interval or ratio. For comparisons on an interval scale, variance-based standardizations are appropriate, but they are not for comparisons on a ratio or log-interval scale (Hansen and Houle 2008). Mean-based standardizations are appropriate for ratio and log-interval scales, but also require that the trait has a natural zero value. Depending on the types of comparisons, one type of standardization may not be appropriate to all, in which case analysis should be restricted to comparisons within the same standardization class. See Hansen and Houle (2008) for more detailed discussion of these (and other) issues.

The most common standarization in the literature is the **variance standardization**, dividing each trait by the square root of its phenotypic variance, giving all re-scaled traits a phenotypic variance of one. **P standardization** goes a step further, not only scaling to give all traits unit variance, but also rotating them to remove any correlations. While potentially appealing, this transform creates new sets of variables (traits) corresponding to the principle components of the $phenotypic$ space. This can often hinder, rather than elucidative, the analysis (Chapter 29). Thus, we generally do not recommend this standardization unless it is specifically motivated by the particular problem under consideration. A final approach is **means standardization**, dividing the trait by its phenotypic mean (Houle 1992, Hereford et al. 2006, Hansen and Houle 2008). We discuss the potential appeal of mean standardization in more detail in the next section.

We use the subscripts σ, $\mathbf{P}$, and μ to denote that a quantify was been variance-, $\mathbf{P}$-, or mean-standardized. For example $\boldsymbol{\beta}_\mu$ refers to the mean-standardized selection gradients, while $\mathbf{G}_\sigma$ refers to the (phenotypic) variance-standardized genetic covariance matrix. How do we obtain expressions for these? All three of these standarizations can be written as a simple matrix multiplication of the vector of traits on the original scale $\mathbf{z}$ by

$$\mathbf{z}_T = \mathbf{T}\mathbf{z} \tag{30.32a}$$

where $\mathbf{T}$ is a symmetric matrix. For example, the scaling matrix for a means standardization is given by

$$\mathbf{T}_\mu = \begin{pmatrix} 1/\mu_1 & 0 & \cdots & 0 \\ 0 & 1/\mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\mu_n \end{pmatrix} \tag{30.32b}$$

while the scaling matrix for variance standarization is

$$\mathbf{T}_\sigma = \begin{pmatrix} 1/\sqrt{P_{11}} & 0 & \cdots & 0 \\ 0 & 1/\sqrt{P_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sqrt{P_{kk}} \end{pmatrix} \tag{30.32c}$$

as suggested (in matrix form) by Rogers and Mukherjee (1992) and Boulding and Hay (1992). Under variance standardization, the ijth element of $\mathbf{P}_\sigma$ is $P_{ij}/\sqrt{P_{ii}\,P_{jj}}$, or one on the diagonals and the phenotypic covariances on the off-diagonals. Finally, the scaling matrix for $\mathbf{P}$-standardization is given by

$$\mathbf{T}_P = \mathbf{P}^{-1/2} \tag{30.32d}$$

where $\mathbf{P}^{1/2}$ is the **square root** of the matrix $\mathbf{P}$, and satisfies $\mathbf{P}^{1/2}\mathbf{P}^{1/2} = \mathbf{P}$, as well as $\mathbf{P}^{-1/2}\mathbf{P} = \mathbf{P}^{1/2}$. The reason for this choice follows from covariance matrix of the transformed vector of traits. Recalling LW Equation 8.21b, and the fact that $\mathbf{T}$ is symmetric, so that $\mathbf{T}^T = \mathbf{T}$,

$$\boldsymbol{\sigma}(\mathbf{z}_T, \mathbf{z}_T) = \boldsymbol{\sigma}(\mathbf{T}\mathbf{z}, \mathbf{T}\mathbf{z}) = \mathbf{T}\boldsymbol{\sigma}(\mathbf{z}, \mathbf{z})\mathbf{T}^T = \mathbf{T}\mathbf{P}\mathbf{T}^T = \mathbf{T}\mathbf{P}\mathbf{T}$$

Thus, under $\mathbf{P}$-standarization, the phenotypic covariance matrix becomes

$$\boldsymbol{\sigma}(\mathbf{z}_T, \mathbf{z}_T) = \mathbf{P}^{-1/2}\mathbf{P}\mathbf{P}^{-1/2} = \mathbf{P}^{1/2}\mathbf{P}^{-1/2} = \mathbf{I}$$

namely, the identity matrix.

Given all of these standardizations have the common form given by Equation 30.32a, we can develop some general expressions between standardized and unstandardized components of the breeder's equation. The phenotypic covariance matrix on the standardized scale becomes

$$\mathbf{P}_T = \mathbf{T}\mathbf{P}\mathbf{T}, \quad \text{with inverse} \quad \mathbf{P}_T^{-1} = \mathbf{T}^{-1}\mathbf{P}^{-1}\mathbf{T}^{-1} \tag{30.33a}$$

Likewise, since

$$\mathbf{z}_T = \mathbf{T}\mathbf{z} = \mathbf{T}(\mathbf{g} + \mathbf{e})$$

the vectors of standarized ($\mathbf{g}_T$) and unstandarized ($\mathbf{g}$) breeding values are related by $\mathbf{g}_T = \mathbf{T}\mathbf{g}$, giving

$$\mathbf{G}_T = \mathbf{T}\mathbf{G}\mathbf{T} \tag{30.33b}$$

Under mean standardization, the ijth element of $\mathbf{G}_\mu$ is $G_{ij}/(\mu_i\,\mu_j)$, while variance standardization this is $G_{ij}/\sqrt{P_{ii}\,P_{jj}}$. Thus the diagonal elements of $\mathbf{G}_\sigma$ correspond to h_i^2 and the off-diagonal elements to $h_i h_j r_{ij}$, namely the heritabilities and co-heritabilities, respectively. Recalling the Robertson-Price dentity (Equation 10.7), the selection differential is the covariance with relative fitness,

$$\mathbf{S}_T = \boldsymbol{\sigma}(\mathbf{T}\mathbf{z}, w) = \mathbf{T}\boldsymbol{\sigma}(\mathbf{z}, w) = \mathbf{T}\mathbf{S} \tag{30.33c}$$

Finally, the standarized selection gradient becomes

$$\boldsymbol{\beta}_T = \mathbf{P}_T^{-1}\mathbf{S}_T = \left(\mathbf{T}^{-1}\mathbf{P}^{-1}\mathbf{T}^{-1}\right)(\mathbf{T}\mathbf{S}) = \mathbf{T}^{-1}\boldsymbol{\beta} \tag{30.33d}$$

Putting all of these pieces together, the relationship between the standardized and unstandardized response is simply

$$\mathbf{R}_T = \mathbf{G}_T\mathbf{P}_T^{-1}\mathbf{S} = (\mathbf{T}\mathbf{G}\mathbf{T})\left(\mathbf{T}^{-1}\mathbf{P}^{-1}\mathbf{T}^{-1}\right)(\mathbf{T}\mathbf{S}) = \mathbf{T}\mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \mathbf{T}\mathbf{R} \tag{30.34a}$$

Likewise, Lande's equation is given by

$$\mathbf{R}_T = \mathbf{G}_T\boldsymbol{\beta}_T = (\mathbf{T}\mathbf{G}\mathbf{T})\left(\mathbf{T}^{-1}\boldsymbol{\beta}\right) = \mathbf{T}\mathbf{G}\boldsymbol{\beta} \tag{30.34b}$$

For example, if the response and gradient are in standarized units, but $\mathbf{G}$ is in unstandard units, then the realized gradient is given by

$$\mathbf{R}_T = \mathbf{G}_T\boldsymbol{\beta}_T = (\mathbf{T}\mathbf{G}\mathbf{T})\,\boldsymbol{\beta}_T$$

giving

$$\boldsymbol{\beta}_T = (\mathbf{TGT})^{-1}\,\mathbf{R}_T = \mathbf{T}^{-1}\mathbf{G}^{-1}\mathbf{T}^{-1}\mathbf{R}_T \tag{30.35}$$

Mean Standarization and Evolvability

While variance standardization may seem the most intuitive, Haldane (1949) noted that a natural scale for many comparisons is the amount of proportional, rather than absolute, change. A give amount of change in standard deviations in two different populations could result in very different proportional amounts of change. For example, a four standard deviation increase in (say) size might correspond to a one percent, or a 200 percent, overall increase in size. Which of these comparisons is more important is a function of the question being asked, but both are relevant. Following on this observation, Houle (1992) noted that in many settings a more natural measure of the ability of populations to evolve is not their heritability, which provides a measure of the rate of absolute change on an interval, but rather their coefficient of additive genetic variance, $CV_A = \sigma_A/\mu_z$, for the trait of interest. He called this measure **evolvability**. We could also use the mean-standardized additive variance,

$$I_A = CV_A^2 = \frac{\sigma_A^2}{\mu_z^2} \tag{30.36}$$

Houle's motivation follows from the univariate breeder's equation, $R = \sigma_A^2\beta$. The *proportional* rate of change is given by

$$\frac{R}{\mu_z} = \frac{\sigma_A^2}{\mu_z}\beta = \frac{\sigma_A^2}{\mu_z^2}\,(\mu_z\beta) = I_A\,\beta_\mu \tag{30.37}$$

Thus, the proportional rate of change is the product of I_A and the mean-scaled gradient $\beta_\mu = \mu_z\beta$ (see Equation 30.33d). While this expression may initially appear less transparent than the simple breeder's equation, there is a very powerful interpretation for β_μ. Namely, it corresponds the **elasticity** of relative fitness with respect to that trait, the percentage change in relative fitness per percentage change in the trait (van Tiendeeren 2000, Hansen et al. 2003a, Hereford et al 2004). Elasticity offers a natural measure of the strenght of selection, as an elasticity of one implies that the selection on the trait is the same as selection directly on fitness. Hence, a value of β_μ of (say) 0.3 means that the strength of selection on that trait is 30 percent of the strength of selection on fitness itself (Chapter 29).

Thus, *where appropriate*, mean-standardization has a powerful interpretative appeal. It requires that the trait have a natural zero point (otherwise the mean can be arbitrarily set). Further, both CV_A, and I_A are overly sensitive to very small mean trait values (Polak and Starmer 2001). The obvious asymmetry in response in the up versus down direction for a treat near zero is not captured by this measure.

Realized Selection Gradients

With a suitably standardized vector of responses in hand, such as the differences between two populations with an initial founder, it is of considerable interest to estimate that nature of selection that could have resulted in their observed differences.

Provided we know (or can estimate) the vector of responses $\mathbf{R}$ and the genetic covariance matrix $\mathbf{G}$, then we can solve $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$ for the amount of phenotypic selection required for the observed change,

$$\boldsymbol{\beta} = \mathbf{G}^{-1}\mathbf{R} \tag{30.38}$$

This was first suggested by Lande (1979), who termed this the **net selection gradient**. Equally appropriates terms are **realized selection gradient** or **retrospective selection gradient.** Strictly speaking, Equation 30.38 holds for only a single generation of response from

an unselected base population. After one generation of selection, linkage disequilibrium will change $\mathbf{G}$ (Chapters 13, 24, 31) and after a sufficient number of generations, allele frequency changes will also impact $\mathbf{G}$. Chapters 24 - 27 detail the changes in genetic variances, which become increasingly unpredictable as selection proceeds and allele frequency change becomes important. Changes in the genetic *covariances* are likely to be even more fickle (Chapter 31).

The thought of being able to estimate the net long-term nature of selection on a suite of divergent traits is just to much for many evolutionary biologists to resist, despite these serious concerns about $\mathbf{G}$ remaining (somewhat) constant. Thus, a typical application of Equation 30.38 is not to estimate $\boldsymbol{\beta}$ over a few generations of selection, but rather over hundreds to thousands of generations. For such cases, let i index generations. The net divergence and net gradient over t generations are given by

$$\mathbf{R}_t = \overline{\mathbf{z}}_n - \overline{\mathbf{z}}_0 = \sum_{i=0}^{t-1} \mathbf{G}_i \boldsymbol{\beta}_i \tag{30.39a}$$

Provided $\mathbf{G}$ remains constant,

$$\mathbf{R}_t = \mathbf{G}\boldsymbol{\beta}_t, \quad \text{where} \quad \boldsymbol{\beta}_t = \sum_{i=0}^{t-1} \boldsymbol{\beta}_i \tag{30.39b}$$

Thus, the net gradient is estimated by

$$\boldsymbol{\beta}_t = \mathbf{G}^{-1}\mathbf{R}_t \tag{30.39c}$$

Even if the assumption of a constant $\mathbf{G}$ is appropriate, there is still need for caution. If $\mathbf{G}$ is ill-conditioned (has one, or more, very small eigenvalues), these become very large eigenvalues for $\mathbf{G}^{-1}$, and small errors in their estimation result in *large* errors for $\mathbf{G}^{-1}$. Recalling from Appendix 4 that $\mathbf{G}$ and $\mathbf{G}^{-1}$ have the same eigenvectors, so that those linear combinations of traits showing the smallest variance for $\mathbf{G}$ give the largest effect for $\mathbf{G}^{-1}$. This makes sense, in that even small changes along these directions (given their reduced variation) require very large amounts of selection, but significant caution is in order in that some of the largest estimates for individual trait gradients may, in fact, be the most unstable.

What happens if $\mathbf{G}$ is changing? Turelli (1988) offers a few useful pointers. In this case, we can write the total response as

$$\mathbf{R}_t = \overline{\mathbf{G}}\boldsymbol{\beta}_t + \sum_{i=0}^{t-1} \left(\mathbf{G}_i - \overline{\mathbf{G}}\right)\boldsymbol{\beta}_i \tag{30.40a}$$

Here

$$\overline{\mathbf{G}} = \frac{1}{t}\sum_{i=0}^{t-1} \mathbf{G}_i \tag{30.40b}$$

is the average covariance matrix over this time span. Turelli notes that if the fluctuations in $\mathbf{G}_i$ and $\boldsymbol{\beta}_i$ are uncorrelated, the sum cancels and Equation 30.39c is an unbiased estimate *provided* we have a good estimate of $\overline{\mathbf{G}}$. While this sounds promising, Turelli notes two reasons why we expect $\mathbf{G}$ and $\boldsymbol{\beta}$ to be correlated. First $\boldsymbol{\beta}_i = \mathbf{P}_i^{-1}\mathbf{S}_i$ where $\mathbf{P}$ is, in part, a function of $\mathbf{G}$. Second, and perhaps more serious, is that large values of $\boldsymbol{\beta}$ correspond to strong selection, which is likely to impact specific elements of $\mathbf{G}$ more heavily than others. Solving Equation 20.28a for the net selection gradient yields

$$\boldsymbol{\beta}_t = \overline{\mathbf{G}}^{-1}\mathbf{R}_t - \overline{\mathbf{G}}^{-1}\left(\sum_{i=0}^{t-1}\left(\mathbf{G}_i - \overline{\mathbf{G}}\right)\boldsymbol{\beta}_i\right) \tag{30.41}$$

The first term of Equation 30.41 is our typical estimator (again, provided that we can obtain a good estimator for $\overline{\mathbf{G}}$), while the second term is the error arising from joint fluctuations in $\mathbf{G}_i$ and $\boldsymbol{\beta}_i$. Turelli makes the key point that if there are large fluctuations in the direction and intensity of selection $\boldsymbol{\beta}_i$, then this second term can dominant the equation, even if fluctuations in $\mathbf{G}_i$ tend to be relatively small.

Again, despite these concerns, the temptation to estimate a retrospective gradient is often too great to ignore. The usual application involves a vector of trait means from each of two divergent populations, be they geographic isolates, experimental populations, or even recent species. The assumption is that both populations started from the same mean value, in which case $\mathbf{R} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$, where $\boldsymbol{\mu}_j$ is the mean of population j. If there are estimates of the covariance matrix from both populations, their average is usually taken as the ancestral $\overline{\mathbf{G}}$. In many applications, however, only one population has an estimate of $\mathbf{G}$, which can introduce additional error if this is not a good estimator of $\overline{\mathbf{G}}$.

Example 30.8. Arnold (1981) examined various aspects of feeding behavior in coastal versus inland populations of garter snakes (*Thamnophis elegnans*) in Northern California. Coastal snakes are terrestrial foragers and largely feed on slugs, while inland snakes are often aquatic foragers, mainly feeding on fish and amphibians. Feeding preference for a variety of potential prey were assayed by number to tongue flicks during a one minute time interval. Coastal and inland populations showed considerable divergence for slug and leech preference, with the vector of variance-standardized means for the two populations as follows:

$$\overline{\mathbf{z}}_{coastal} = \begin{pmatrix} \overline{z}_{slug} \\ \overline{z}_{leach} \end{pmatrix} = \begin{pmatrix} 3.26 \\ 2.29 \end{pmatrix}, \qquad \overline{\mathbf{z}}_{inland} = \begin{pmatrix} 0.40 \\ 0.95 \end{pmatrix}$$

These two traits showed a very high genetic correlation ($r_A = 0.89$) and had individual heritabilities of 0.17 (slug) and 0.59 (leach). Hence, the standarized response between populations was

$$\mathbf{R}_\sigma = \boldsymbol{\Delta}\overline{\mathbf{z}}_\sigma = \begin{pmatrix} 3.26 - 0.40 \\ 2.29 - 0.95 \end{pmatrix} = \begin{pmatrix} 2.86 \\ 1.24 \end{pmatrix}$$

Since we are working with standardized variables, we have $\mathbf{R}_\sigma = \mathbf{G}_\sigma \boldsymbol{\beta}_\sigma$, where $\mathbf{R}_\sigma$ is the variance-standarized response, $\boldsymbol{\beta}_\sigma$ the standardized selection gradients, and $\mathbf{G}_\sigma$ the standarized genetic covariance matrix. From Equation 30.36b

$$\mathbf{G}_\sigma = \begin{pmatrix} h_1^2 & r_A h_1 h_2 \\ r_A h_1 h_2 & h_2^2 \end{pmatrix} = \begin{pmatrix} 0.17 & 0.28 \\ 0.28 & 0.59 \end{pmatrix}$$

Solving gives the realized (variance-standardized) selection gradient to obtain this response as

$$\boldsymbol{\beta}_\sigma = \begin{pmatrix} 0.17 & 0.28 \\ 0.28 & 0.59 \end{pmatrix}^{-1} \begin{pmatrix} 2.86 \\ 1.24 \end{pmatrix} = \begin{pmatrix} 61.2 \\ -26.9 \end{pmatrix}$$

Thus the observed response is very misleading about the actual nature of selection on these traits. Despite both traits responding in the same direction, the actual selection was **antagonistic**, namely against the direction of the genetic correlation, with strong selection *against* leach preference and strong selection for slug preference.

Example 30.9. A more complicated analysis by Merilä et al. (1994) shows the power, and problems (even when assuming a constant $\mathbf{G}$) with realized selection gradients. They examined morphological divergence in seven traits between to recent sister species of European

flycatchers (*Ficedula albicollis* and *F. hypoleuca*). The authors assumed that *albicollis* was the parental species (the covariance matrix before divergence). The genetic correlations among these seven traits were modest, with a mean around 0.22, so at first blush one might expect their effects to be small. Using the variance-standardized differences in mean trait values $\Delta\overline{z}_\sigma$, Merilä et al. computed two different measures of the amount of selection required to account for the observed divergence (of course, drift is also a possibility as well, Chapter 6). The first measure was to treat each mean divergence as a univariate problem, computing the required total selection intensity $\overline{i}$ needed to account for the change given the heritability of the trait ($\overline{i} = \Delta z_\sigma / h^2$). They also computed realized (standardized) selection gradients $\boldsymbol{\beta}_\sigma$ using the $\mathbf{G}$ matrix from *albicollis*. Their results were as follows:

Trait	$\Delta\overline{z}_\sigma$	$\overline{i}$	β_σ
Beak width	0.00	0.00	4.69
Beak depth	0.00	0.00	-3.29
Beak lenght	0.00	0.00	-0.11
Tarsus lenght	0.00	0.00	2.11
Wing lenght	-1.88	-3.67	-7.19
Tail lenght	0.51	1.13	4.46
lenght 1st primary	-0.37	-0.69	-0.09

The authors made several observations about these results. First, the amount of required selection is much greater when genetic correlations are taken into account (based by the lengths of the gradient versus intensity vectors, 10.43 vs. 3.90). They also note the considerable amount of selection required to keep the first four traits unchanged in the face of selection on the other correlated traits.

While these results are striking, the authors caution that care must be taken in their interpretation. In particular, gradients are *estimates*, subjected to errors in the estimates of $\mathbf{G}$. In an attempt to account for these uncertainties, they bootstrapped the original $\mathbf{G}$ matrix to obtain mean bootstrapped values ($\widehat{\beta_\sigma}$) and 95% bootstrap confidence intervals (given as the lower 2.5% and upper 97.5% of the empirically-generated bootstrapped distribution). This was done by using families as the sampling units, drawing 280 families (at random, and with replacement) from the original data set, then computing the resulting $\mathbf{G}$ matrix and realized gradient. This procedure was repeated 1000 times to generate the bootstrap values shown below:

Trait	β_σ	$\widehat{\beta_\sigma}$	2.5%	97.5%
Beak width	4.69	-0.8	-16.1	38.4
Beak depth	-3.29	1.2	-21.0	24.5
Beak lenght	-0.11	-0.8	-8.1	16.2
Tarsus lenght	2.11	-0.9	-5.2	9.5
Wing lenght	-7.19	-7.8	-22.1	-1.7
Tail lenght	4.46	9.6	0.004	35.0
lenght 1st primary	-0.09	11.11	-6.7	42.7

The while observation of strong selection required with correlations holds with the bootstrapped data, there is no support for realized gradients on the three beak traits and tarsus length (as all of their confidence intervals span zero). Thus the notion of strong selection required to keep them in place while other traits change is not supported. Removing these four traits from consideration, Merilä et al. re-estimated $\boldsymbol{\beta}$ focusing on the three remaining traits, finding

$$\boldsymbol{\beta}_\sigma = \begin{pmatrix} -6.9 \\ 8.3 \\ 8.1 \end{pmatrix}, \quad \boldsymbol{\Delta}\overline{\mathbf{z}}_\sigma = \begin{pmatrix} -1.88 \\ 0.51 \\ -0.37 \end{pmatrix}$$

One measure of the constraints imposed by $\mathbf{G}$ is to consider the angle θ between the direction preferred by selection ($\boldsymbol{\beta}_\sigma$) and the actual direction of response ($\boldsymbol{\Delta}\overline{\mathbf{z}}_\sigma$). From Equation A4.2 this is given by

$$\cos(\theta) = \frac{\left(\boldsymbol{\Delta}\overline{\mathbf{z}}_\sigma\right)^T \boldsymbol{\beta}_\sigma}{|| \boldsymbol{\Delta}\overline{\mathbf{z}}_\sigma || \, || \boldsymbol{\beta}_\sigma ||} = 0.559, \quad \text{or} \quad \theta = 56 \text{ degrees}$$

we suppressed the σ subscript for ease of presentation. As with the previous example, there is considerable divergence between the actual and optimal directions.

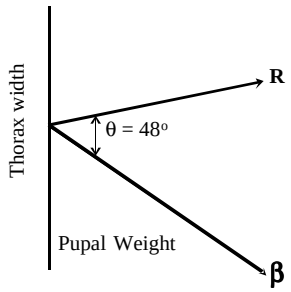
EVOLUTIONARY CONSTRAINTS IMPOSED BY GENETIC CORRELATIONS

As we have mentioned numerous times, the response $\mathbf{R}$ does not track with the direction $\boldsymbol{\beta}$ giving largest increase in fitness. In all but very unusual settings, $\mathbf{G}$ rotates the response vector away for $\boldsymbol{\beta}$. If this rotation is slight, $\mathbf{G}$ imposes relatively minimal constraints. If the rotation is significant, the evolution can be slowed to a crawl, with some traits showing non-adaptive responses (changes in their means in the opposite direction of that favored by selection). In addition to potentially rotating the response from $\boldsymbol{\beta}$, $\mathbf{G}$ also *scales* the response. If the scaling factor is very small, the response can be very small even with no rotation, i.e. all of the response is in the direction of $\boldsymbol{\beta}$ but it is very small (see Figure 30.2b). In the extreme case, this scaling can have a value of zero if $\boldsymbol{\beta}$ is in the same direction as an eigenvalue of $\mathbf{G}$ with a corresponding zero eigenvalue, giving no response for any strength of phenotypic selection. Thus it is only by understanding the geometric structure of any particular $\mathbf{G}$ matrix that we can fully understand its inherent constraints.

Example 30.10. Let's reconsider the data (Example 30.1) of Conner and Via (1992) on the effects of pupal weight, wing length, and thorax width on female reproductive success in a laboratory population of the flour beetle *Tribolium castaneum*. We ignore wing length, as there was essentially no selection ($\beta = -0.01$). The resulting (scaled) $\mathbf{G}$ matrix for pupal weight and thorax width and the expected response is

$$\mathbf{G} = \begin{pmatrix} 1.25 & 0.36 \\ 0.36 & 0.20 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} 0.22 \\ -0.16 \end{pmatrix}, \qquad \mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \begin{pmatrix} 0.217 \\ 0.047 \end{pmatrix}$$

Selection ($\boldsymbol{\beta}$) strongly favors an increase in weight and a roughly equal decrease in width. However, the actual response to selection is to increase both traits. As shown below, the vector of actual responses has been rotated away from the direction for optimal response $\boldsymbol{\beta}$, with angle θ between these two vector being 48 degrees.



The response shows a strong constraint, with selection $\boldsymbol{\beta}$ favoring an increase in pupal weight and a decrease in thorax width, while the response $\mathbf{R}$ gives an increase in each trait.

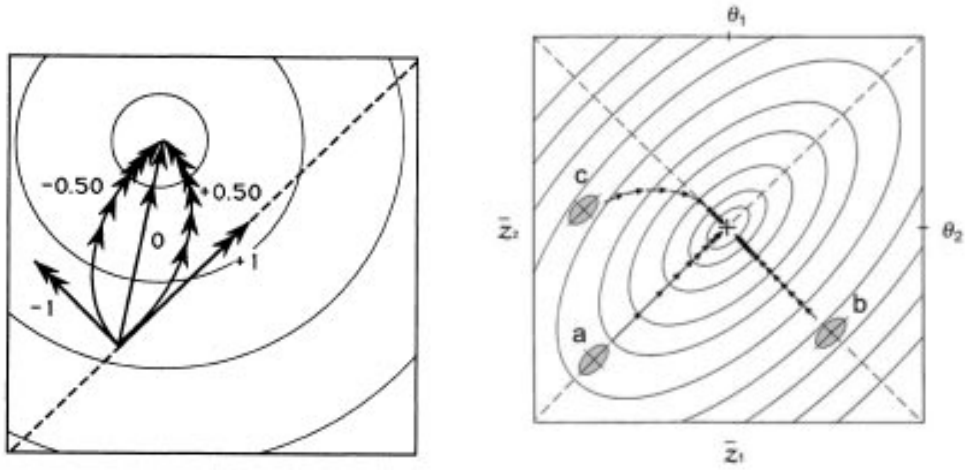


Figure 30.2. The effects of genetic correlations on the direction and speed of approach to a local optimum. Both figures show mean fitness surfaces $\overline{W}(\boldsymbol{\mu})$, with contours representing trait combinations having identical fitness. Both surfaces have a single optimum value to which the population mean tries to evolve. **Left:** The genetic variances for both traits are assumed equal, with trajectories plotted for different genetic correlations ($r_A = 0, \pm 0.5, \pm 1.0$). An $r_A = \pm 1$ is a complete genetic constraint, resulting in one trait reaching its optimum value while the other is constrained. When there is no genetic correlations, since the variances are equal, the trajectory is straight. With correlation of ± 0.5, the trajectories are curved, and the time to reach a local neighborhood of the optimum is greatly increased (arrows show 50-generation intervals of change). **Right.** The nature of constraints in $\mathbf{G}$ are a *joint* function of $\boldsymbol{\beta}$ and the eigenvectors of $\mathbf{G}$. Three populations with the same $\mathbf{G}$ matrix but different means are placed an the adaptive landscape. Population **a** has its principal axis of variation aligned with the selection gradient and the trajectory towards the optimum is a straight line. Population **b** has $\boldsymbol{\beta}$ aligned with the minor axis of $\boldsymbol{\beta}$. It also evolves along a straight line, but in a much slower fashion (arrows indicate equal time intervals). Finally, population **c** has neither axis aligned with $\boldsymbol{\beta}$. The result is a curved trajectory. After Arnold (1991 simulations by Via and Lande 1985) and Arnold et al. (2001).

Dynamics of Quantitative Traits on an Adaptive Topography

Given this rotation and scaling of the response by $\mathbf{G}$, how do traits evolve? Some traits do indeed show strong adaptive changes (their means changing in the direction favored by selection). Evolutionary change is most rapid along the directions of maximal genetic variance (weighted by the β), so these traits change most quickly. As these traits approach their equilibrium values (as would occur if there is a peak in the fitness surface), their evolution slows down, with other traits now showing the greater amounts of change. Eventually, all traits will reach this local fitness peak (provided $\mathbf{G}$ is not singular), but their paths from their current position to the fitness peak do not follow straight trajectories. Rather theses are curved (Figure 30.2), resulting in much longer times to approach the optimum. If the environment is changing sufficiently rapidly, then the constraints imposed by $\mathbf{G}$ may not

allow the population to evolve rapidly enough to avoid extinction (**Chapter xx**). Finally, an important point is made by Figure 30.2b in that the nature of the constraints is a function of *both* $\boldsymbol{\beta}$ and $\mathbf{G}$. In this figure, three populations with the same $\mathbf{G}$ matrix, all heading for the same optimal peak, start off at different locations on the fitness surface, with the amount of constraints depending upon their location. The actual dynamics can easily be much more complicated that suggested by Figure 30.2. For example, if selection is occurring on multiple traits, but we simply follow the trajectory of two traits, oscillations and reversals can often be seen (see figures in Björklund 1996).

What Happens to Mean Fitness $\overline{W}$?

If the response is rotated away from the direction giving the largest local increase in mean population fitness, what happens to mean fitness $\overline{W}(\boldsymbol{\mu})$? Assuming frequency-independent selection (individual fitness is not a function of the trait means), mean fitness still increases, just not at the maximal rate (Lande 1979). Lande's argument is as follows: $\boldsymbol{\beta} = \nabla_{\boldsymbol{\mu}}[\ln \overline{W}(\boldsymbol{\mu})]$ is expressed in terms of log mean fitness, so we will work on that scale. Since log is a monotonic function, if $\ln \overline{W}$ increases, then so must $\overline{W}$. We wish to detemined how a change in the mean $\mathbf{R} = \boldsymbol{\Delta\mu}$ translates into a change in ln mean fitness, $\Delta \ln \overline{W}(\boldsymbol{\mu})$. To do so, we expand the log of mean fitness in a Taylor series and assume that second and higher-order terms can be neglected (as would occur with weak selection and the population mean away from an equilibrium point). As shown in Example A5.7, this gives

$$\Delta \ln \overline{W}(\boldsymbol{\mu}) \simeq \left(\nabla_{\boldsymbol{\mu}}[\ln \overline{W}(\boldsymbol{\mu})] \right)^T \mathbf{R} \tag{30.42a}$$

Assuming that the joint distribution of phenotypes and additive genetic values is MVN, then $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$, or $\nabla_{\boldsymbol{\mu}}[\ln \overline{W}(\boldsymbol{\mu})] = \boldsymbol{\beta} = \mathbf{G}^{-1}\mathbf{R}$. Substituting gives

$$\Delta \ln \overline{W}(\boldsymbol{\mu}) \simeq \left(\mathbf{G}^{-1}\mathbf{R}\right)^T \mathbf{R} = \mathbf{R}^T \mathbf{G}^{-1} \mathbf{R} \geq 0 \tag{30.42b}$$

The inequality follows since $\mathbf{G}$ is a variance-covariance matrix and hence is non-negative definite (all the eigenvalues of $\mathbf{G}$ are ≥ 0, see Appendix 4), with the inequality strictly greater than zero when $\mathbf{G}$ is non-singular (no zero eigenvalues). Thus under these conditions, mean population fitness always increases, although since $\mathbf{R} \neq \nabla_{\boldsymbol{\mu}}[\ln \overline{W}(\boldsymbol{\mu})]$ fitness does not increase in the fastest possible manner. If selection is frequency-dependent, average fitness need not increase and the equilibrium phenotypic mean can lie off the fitness optimum (Lande 1976, 1980).

Constraints are Given by the Eigenstructure of G

Much of the background material for this next section is developed in Appendix 4, and the reader with a modest knowledge matrix algebra might wish to review this before proceeding.

How do constrains such as that seen in Example 30.10 arise? The answer is in the geometry of $\mathbf{G}$, namely the directions of $\mathbf{G}$ that show the most (and least) variation. These are given by the **eigenstructure** of $\mathbf{G}$, with the eigenvalues λ_i of $\mathbf{G}$ indicating how much variation resides along certain directions (their associated eigenvectors $\mathbf{e}_i$). These also determine the rotation (eigenvectors) and scaling (eigenvalues) of vectors as we will see shortly (Equation 30.46b).

The mean vector changes in the direction most favored by selection if and only if

$$\mathbf{G}\boldsymbol{\beta} = \lambda\boldsymbol{\beta} \tag{30.43}$$

which only occurs when $\boldsymbol{\beta}$ is an eigenvector of $\mathbf{G}$. Note that even if $\mathbf{G}$ is a diagonal matrix (no genetic correlations), Equation 30.43 is usually not satisfied. Only when $\mathbf{G} = \sigma_A^2\mathbf{I}$ is

Equation 30.43 satisfied for arbitrary $\boldsymbol{\beta}$. Thus, only when both (i) all characters have the same additive genetic variance and (ii) there no additive genetic covariance between characters is the response to selection in the direction most favored by natural selection. Thus differences in the amounts of additive genetic variances between characters, in addition to non-zero additive-genetic covariances, also impose constraints on character evolution.

Recall (Appendix 4) that a symmetric matrix (such as a covariance matrix) can be decomposed into its **spectral decomposition**, writing it as a function of its eigenvalues and associated eigenvectors (Equation A4.9b), giving

$$\mathbf{G} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^T + \cdots + \lambda_n \mathbf{e}_n \mathbf{e}_n^T \tag{30.44}$$

The importance of this decompositon becomes apparent when we consider the **projection** of one vector onto another. The projection of $\boldsymbol{\beta}$ on (say) $\mathbf{e}_i$ is denoted Proj($\boldsymbol{\beta}$ on $\mathbf{e}_i$). Recalling Equation A4.3b (and using the fact that $\mathbf{e}_i$ is of unit lenght),

$$\text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_i) = \left(\mathbf{e}_i^T \boldsymbol{\beta}\right) \mathbf{e}_i \tag{30.46a}$$

Note that the inner product $\mathbf{e}_i^T \boldsymbol{\beta}$ is a scalar, while the actual *direction* of the projection is along the vector $\mathbf{e}_i$, so that the projection onto $\mathbf{e}_i$ is a scaled (by $\mathbf{e}_i^T \boldsymbol{\beta}$) vector pointing in the direction of $\mathbf{e}_i$. When we multiply $\mathbf{G}$ and $\boldsymbol{\beta}$, Equations 30.44 and 30.45a imply

$$\begin{aligned} \mathbf{R} = \mathbf{G}\boldsymbol{\beta} &= \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T \boldsymbol{\beta} + \lambda_2 \mathbf{e}_2 \mathbf{e}_2^T \boldsymbol{\beta} + \cdots + \lambda_n \mathbf{e}_n \mathbf{e}_n^T \boldsymbol{\beta} \\ &= \lambda_1 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_1) + \lambda_2 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_2) + \cdots + \lambda_n \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_n) \end{aligned} \tag{30.45b}$$

Thus we can break the response vector $\mathbf{R}$ into the sum of scaled projections of $\boldsymbol{\beta}$ onto each of the eigenvectors of $\mathbf{G}$. If the majority of $\boldsymbol{\beta}$ lies along the direction of an eigenvector with a very small eigenvalue (i.e., this direction accounts for very little of the total genetic variance in $\mathbf{G}$), then the response will also be small.

Example 30.11. Let's return to Example 30.10 and see how $\mathbf{G}$ helps explain the significant depature of the response vector $\mathbf{R}$ from the optimal direction $\boldsymbol{\beta}$ favored by phenotypic selection. Here

$$\mathbf{G} = \begin{pmatrix} 1.25 & 0.36 \\ 0.36 & 0.20 \end{pmatrix}$$

The eigenvalues (Appendix 4) of $\mathbf{G}$ are

$$\lambda_1 = 1.36, \quad \lambda_2 = 0.09$$

with their associated (unit-lenght) eigenvectors

$$\mathbf{e}_1 \simeq \begin{pmatrix} 0.955 \\ 0.296 \end{pmatrix} \qquad \mathbf{e}_2 \simeq \begin{pmatrix} 0.296 \\ -0.955 \end{pmatrix}$$

Note that 94%, 1.36/(1.36+0.09), of the variation in $\mathbf{G}$ is along the direction given by $\mathbf{e}_1$. Consider the angle between $\boldsymbol{\beta}$ and each of these eigenvectors. Here $||\boldsymbol{\beta}|| = 0.272$, $\mathbf{e}_1^T \boldsymbol{\beta} = 0.163$, $\mathbf{e}_2^T \boldsymbol{\beta} = 0.218$, and (by construction) $||\mathbf{e}_1|| = ||\mathbf{e}_2|| = 1$. Applying Equation A4.2, the angles between $\boldsymbol{\beta}$ and each of two eigenvectors are given by

$$\cos(\theta|\mathbf{e}_1, \boldsymbol{\beta}) \simeq \frac{0.163}{0.272 \cdot 1} \simeq 0.600, \qquad \text{and} \qquad \cos(\theta|\mathbf{e}_2, \boldsymbol{\beta}) \simeq \frac{0.218}{0.272 \cdot 1} \simeq 0.801$$

The resulting angle between $\mathbf{e}_1$ and $\boldsymbol{\beta}$ is $\theta(\mathbf{e}_1, \boldsymbol{\beta}) \simeq 53.2^\circ$, so that selection gradient is 52 degrees away from most of the usable variation in $\mathbf{G}$. On the other hand, $\theta(\mathbf{e}_1, \boldsymbol{\beta}) \simeq 36.8^\circ$ so that $\boldsymbol{\beta}$ is much closer in direction to the eigenvalue that only accounts for around 6% of the total variation in $\mathbf{G}$. This is our first glimpse into the nature of constraints on this $\mathbf{G}$.

Equation 30.45b breaks the total response into the projection of $\boldsymbol{\beta}$ onto each of the eigenvectors of $\mathbf{G}$,

$$\mathbf{R} = \lambda_1 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_1) + \lambda_2 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_2)$$

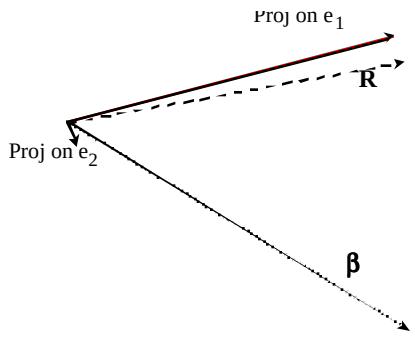
The scaled projections of $\boldsymbol{\beta}$ on these eigenvectors are

$$\lambda_1 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_1) = 1.36(\boldsymbol{\beta}^T \mathbf{e}_1)\, \mathbf{e}_1 \simeq \begin{pmatrix} 0.211 \\ 0.066 \end{pmatrix}, \quad \lambda_2 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_2) \simeq \begin{pmatrix} 0.006 \\ -0.019 \end{pmatrix}$$

giving the total response as

$$\mathbf{R} = \begin{pmatrix} 0.211 \\ 0.066 \end{pmatrix} + \begin{pmatrix} 0.006 \\ -0.019 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.047 \end{pmatrix}$$

as obtained in Example 30.10. As the figure belows shows, the eigenstructure of $\mathbf{G}$ explains the unusual behavior of response seen in Example 30.10. The eigenvector associated with the **leading eigenvalue** λ_1 accounts for most (94%) of the variation inherent in $\mathbf{G}$, and this eigenvector corresponds to an index with positive values for each trait. Hence, even though $\boldsymbol{\beta}$ points in very much the same direction as $\mathbf{e}_2$, because $\lambda_1 >> \lambda_2$ the projection of $\boldsymbol{\beta}$ on $\mathbf{e}_1$ gives a vector of greater length than the projection on $\mathbf{e}_2$, and it is the sum of these two projection vectors that result in the increase in width despite selection to decrease it.



Equation 30.45a provides an exact treatment of our comment that response occurs along the directions of greatest variation when weighted for selection. Equation A4.3c gives the length of the projection of $\boldsymbol{\beta}$ onto $\mathbf{e}_i$ as

$$||\text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_i)|| = |\cos(\theta_i)|\ ||\boldsymbol{\beta}|| \tag{30.46a}$$

where θ is the angle between $\boldsymbol{\beta}$ and $\mathbf{e}_i$. Thus, the length of response from the genetic variation associated with the ith eigenvalue is just

$$\frac{||R||_i}{||\boldsymbol{\beta}||} = \lambda_i\, |\cos(\theta_i)| \tag{30.46b}$$

showing how the amount of variation (λ_i) and the angle θ_i between this direction of variation and $\boldsymbol{\beta}$ determine the contribution to response.

Distribution of eigenvalues provides some information on the constraints in any direction of $\mathbf{G}$. The more "even" the distribution, the less the constraints. A quick measure of this is the variance of the eigenvalues (Wagner 1984). If all eigenvalues are the same, the variance is zero, while if there are a few large eigenvalues and many small ones, the variance can

be high. Hence, the smaller the variance, the fewer the constraints in any direction. This measure has been related to the concept of **modularity** (Wagner 1994, Hansen 2003), and this is breifly examined in Chapter 31 and in more detail in Chapter 44.

Trade-offs, Developmental Constraints, Genetic Correlations, and the Lande Equation

Evolution is widely-regarded as optimization of fitness given trade-offs and constraints. Indeed, there often appears to be a schizophrenic nature in how evolution is viewed. On one hand, organisms are typically regarded as being highly fine-tuned to their environment, and yet on the other it is widely suggested that there are limited evolutionary degrees of freedom because of lack of sufficient genetic variation and other constraints (Björklund 1996). This raises the issue as to how organisms are able to adopt to changing environments (biotic and abiotic).

What is the nature of the potential roadblocks to efficient evolution? Ecologists and evolutionary biologists are often concerned with **tradeoffs**, wherein a trait that does well at one stage of the life cycle may do poorly at another (e.g. Example 30.12). Organism are thus stuck with trait values that are not optimally adaptive at either stage, but rather a compromise. Developmental biologists, on the other hand, are concerned with **developmental constraints** — combinations of trait values that are generally not (or at least not easily) accessible given the nature of development. There can be confusion in reading the literature, as the notion of a trade-off and the notion of a constraint are often used synonymously, when in fact they can be quite distinct. One can have trade-offs and yet no constraints, and likewise have constraints and yet no trade-offs. The Lande equation helps us to place these in the proper context.

Example 30.12. An interesting example of a trade-off was seen in male crickets (*Gryllus firmus*) by Crnokrak and Roff (1998). Males in this species show two distinct wing morphs, one winged with functional flight muscles, the other has reduced wings and is flightless. These forms are denoted by LW (long wing) and SW (short wing), respectively. Crnokrak and Roff examined the phenotypic correlation between wing morph and male calling rate, the later being positively correlated with acquisition of mates. A significant difference in both calling rate and number of attracted females existed between the LW and SW morphs, with the flightless individuals being superior on both counts. The authors note that calling is energetically very expensive, and that SW males might be able to allocate more energy to calling versus expending it on flying. This is an example of a *phenotypic* tradeoff. However, these phenotypic correlations also translated into negative genetic correlations (-0.5 and -0.7 for between wing morph and calling rate in the offspring from SW and LW males, respectively) making this also a *genetic* tradeoff, as selection to increase both traits goes against the direction of the genetic correlation.

Example 30.12 makes the key point that tradeoffs and constraints can operate at two levels, phenotypic and genetic. Placed in the context of the Lande equation, these amount to constraints at the level of selection on the phenotype (constraints on β), and constraints on how phenotypes are transmitted to their offspring (constraints inherit in the geometry of a particular $\mathbf{G}$ matrix).

In the recent literature, discussions of constraints at the phenotypic level are typically focused on conflicting selective pressures, where the β for a particular episode shows much stronger values that for the β when the entire lifetime fitness is considered. Schluter et al. (1991) reviews examples of such **selective tradeoffs**, making the important comment that

such conflicting pressures on life-history traits may be more common than observed. Their reasoning is our old friend, a hidden environmental variable correlated with traits and fitness (Chapter 16), in particular, nutrition. Nutrition can influence fitness components in positive directions, masking any potential tradeoffs that would normally be present if nutrition was also not inflating fitness.

Constraints imposed by phenotypes are not limited to selective tradeoffs. Much of the historical discussion on development constraints was based on observations of phenotypic correlation among traits. A classic pattern is Olson and Miller's concept of **morphological integration**, wherein high phenotypic correlations are associated with functional or developmental parts (i.e., the bones in a wing are more highly correlated with each other than they are to foot bones in the same individual). The idea is that the developmental program (or trajectory) results in integration of component parts to make a whole trait more functional. However, Cheverud (1984) makes the critical point that since $\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S}$, the effects of any such phenotypic correlation on selection are removed from the response to selection, and all that matters for response is the *genetic* correlation among these components. Of course, it is still possible to have limited genetic correlations between components and still observe morphological integration. Feedback among the developing parts within an individual during the ontology of that trait can result in *environmental* correlations that override any genetic correlations to generate the observed pattern of phenotypic correlations.

A final route by which phenotypic correlation can enter is if $\mathbf{P}$ is singular. If this is true, than $\mathbf{G}$ is also singular (Pease and Bull 1988), although the directions of null variance may be different in genotypic and phenotypic space. Thus, a singular $\mathbf{P}$ may result in selection being unable to act independently on component traits, even if these have usable genetic variation. If $\mathbf{P}$ is singular, then generalized inverses (LW Appendix 3) can be used, with $\boldsymbol{\beta} = \mathbf{P}^{-}\mathbf{S}$. Instead of returning a unique solution, some of the individual β will be expressed as a linear constraint of the others (i.e., solutions are lines or planes, not points). Cheverud (1988) suggested that phenotypic correlations may generally mirror genetic one, so that as a first approximation when nothing else is available, one might use suitably scaled phenotypic correlations. Willis et al. (1991) pointed out problems with this suggestion, and these are also reflected in our discussion of phenotypic constraints. Nevertheless, constraints at the phenotypic level can be significant and have a major impact on selection response.

Finally lets consider genetic constraints more carefully. From the standpoint of the Lande equation, these are eigenvectors of $\mathbf{G}$ that have very small eigenvalues. In the extreme of a zero eignevalue, the direction of its associated eigenvector is a **forbidden evolutionary trajectory** (Kirkpatrick and Lofsvold 1992), an evolutionary path that will not response to selection without a change in $\mathbf{G}$. At the genetic-level, high negative genetic correlations are often called trade-offs, especially in the life-history evolution literature (**refs here**). For example, if individuals with large breeding values for mating success tend to have small breeding values for fecundity, then these two traits show a negative genetic correlation, which is often phrased as a trade-off between mating success and fecundity per mating. Since these are selected in the both direction, the negative correlation retards the simultaneous rapid increase in both traits. In the extreme, neither might evolve despite strong selection on each. Viewed in terms of the geometry of $\mathbf{G}$, these negative correlations result in eigenvectors along the direction of increase in both traits having smaller eigenvalues, and hence limited response. However, the sign of a genetic correlation does not guarantee a constraint, as the following example illustrates.

Example 30.13. Consider the following gradient of selection differentials and three different

covariance matrices,

$$\boldsymbol{\beta} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}, \quad \mathbf{G}_2 = \begin{pmatrix} 10 & 8 \\ 8 & 10 \end{pmatrix}, \quad \mathbf{G}_3 = \begin{pmatrix} 10 & -8 \\ -8 & 10 \end{pmatrix}$$

The three resulting response vectors are

$$\mathbf{R}_1 = \begin{pmatrix} 20 \\ 20 \end{pmatrix}, \quad \mathbf{R}_2 = \begin{pmatrix} 36 \\ 36 \end{pmatrix}, \quad \mathbf{R}_3 = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

Notice in all three cases that the *direction* of response is optimal, but that the *scaling* of the response differs. Taking the length of the three response vectors gives (in order) 6.32, 8.46, and 2.82. The presence of strong negative correlation resulted in a greatly reduced response relative to the uncorrelated case ($\mathbf{R}_1$ vs. $\mathbf{R}_3$). Conversely, the presence of positive genetic correlations *increases* response. Thus, the *lack* of positive correlation in this case could indeed be viewed as a constraint (Björklund 1996). As Gould (1989) and Björklund (1996) point out, constraints have a dual nature, depending on the direction of selection. For example, if we replace $\boldsymbol{\beta}$ by $(2, -2)^T$, now $\mathbf{G}_2$ constrains response relative to $\mathbf{G}_1$, while $\mathbf{G}_3$ enhances it. While constraints can be absolute (no evolution along a trajectory corresponding to a zero eigenvalue), they also have a relative quality depending on the direction selection is attempting to move the population.

Cheverud (1984) has forcefully argued that $\mathbf{G}$ mirrors any internal developmental constraints, as it looks at the available variation that can be passed from parent to offspring. This statement is both fundamentally correct and fundamentally flawed. The breeder's equation is essentially a *local predictor* of response and constraints (e.g. Rice 2008). $\mathbf{G}$ changes under selection (Chapter 31), and a key (and unresolved) issue is whether eigenvectors of $\mathbf{G}$ that have small associated eigenvalues are the result of developmental constraints, or from past selection eroding existing variation, or perhaps both. As $\mathbf{G}$ changes, the prediction of response (and its constraints) also change. Phrased another way, one could treat true developmental constraints as those that exist no matter what the possible pattern of genetic variance. As development unfolds, certain combination of trait values may be very difficult to obtain, even in the absence of removal of genetic variation by selection. Within an organism, traits need to adapt to other traits, not just an external environment, and the accessible combinations of $\mathbf{G}$ are themselves product of selection and evolution for developmental pathways. A final comment is that we have been assuming genotypic (and phenotypic) values line on a continuous space with no gaps. Clearly, development may potentially result in such gaps, in which case one could have very significant developmental constraints (holes in the accessible genotypic space), and yet still no genetic correlations (i.e., a full-rank $\mathbf{G}$ matrix). However, we saw with threshold traits (Chapter 10, LW Chapter 25) that quantitative-genetic models can easily handle phenotypic discontinuities, *provided* they result from a characterizable mapping from a continuous space (on which the breeder's equation operates) to the phenotypic space we observed. See Polly (2008) for further discussion.

Multivariate Measures of Evolvability

As the last several sections have shown, the nature of any potential constraints in $\mathbf{G}$ is a function of *both* $\mathbf{G}$ and a particular $\boldsymbol{\beta}$. Hansen and Houle (2008) suggest that the actual response $\mathbf{R}$ is a measure of **respondability**, namely how much the population changes, but what is of real interest is the amount of response along the direction of $\boldsymbol{\beta}$, which they call the **evolvability**, as the latter is the actual change that increases fitness (Figure 30.5). These two terms were introduced by Houle (1992), but in a different context for univariate selection.

Hansen and Houle formally define respondability as the ratio of lengths of the response vector to the selection gradient, both of which are suitably standardized (typically either by variance- or mean-standardization). Respondability is thus a function of $\boldsymbol{\beta}$, and they represent this by $r(\boldsymbol{\beta})$, where

$$r(\boldsymbol{\beta}) = \frac{||\mathbf{R}||}{||\boldsymbol{\beta}||} = \sqrt{\frac{\mathbf{R}^T\mathbf{R}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}} \tag{30.47a}$$

Recall that $||\mathbf{x}|| = \sqrt{\mathbf{x}^T\mathbf{x}}$ is the length (or norm) of the vector $\mathbf{x}$ (Appendix 4). Using the breeder's equation $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$, and the fact that $\mathbf{G}^T = \mathbf{G}$, we can alternatively write Equation 30.47a as

$$r(\boldsymbol{\beta}) = \sqrt{\frac{\boldsymbol{\beta}^T\mathbf{G}^2\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}}} \tag{30.47b}$$

Hansen and Houle suggest that natural a measure of evolvability is given by the lenght of the projection of $\mathbf{R}$ onto $\boldsymbol{\beta}$, which from Equation A4.3c is given by

$$||\text{Proj}(\mathbf{R}\,\text{on}\,\boldsymbol{\beta})|| = \cos(\theta)\,||\mathbf{R}|| \tag{30.48a}$$

where θ is the angle between $\boldsymbol{\beta}$ and $\mathbf{R}$. Note that we have left this as a signed lenght, with a negative value implying the projection vector is in the complete opposite direction (180 degrees) from $\boldsymbol{\beta}$ (a reflection about the origin, see Appendix 4). From Equation A4.2,

$$\cos(\theta) = \frac{\boldsymbol{\beta}^T\mathbf{R}}{||\mathbf{R}||\,||\boldsymbol{\beta}||} = \frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{||\mathbf{R}||\,||\boldsymbol{\beta}||}$$

giving

$$||\text{Proj}(\mathbf{R}\,\text{on}\,\boldsymbol{\beta})|| = \frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{||\mathbf{R}||\,||\boldsymbol{\beta}||}\;||\mathbf{R}|| = \frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{||\boldsymbol{\beta}||} \tag{30.48b}$$

Hansen and Houle use this as the basis for their measure of evolvability $e(\boldsymbol{\beta})$, scaling the length of the projection vector by the length of $\boldsymbol{\beta}$, giving

$$e(\boldsymbol{\beta}) = \frac{||\text{Proj}(\mathbf{R}\,\text{on}\,\boldsymbol{\beta})||}{||\boldsymbol{\beta}||} = \frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{||\boldsymbol{\beta}||^2} = \frac{\boldsymbol{\beta}^T\mathbf{G}\boldsymbol{\beta}}{\boldsymbol{\beta}^T\boldsymbol{\beta}} \tag{30.48c}$$

Finally, recalling the definition of respondability (Equation 30.47a), we see a simple relationship with evolvability,

$$e(\boldsymbol{\beta}) = \cos(\theta)\,\frac{||\mathbf{R}||}{||\boldsymbol{\beta}||} = \cos(\theta)\,r(\boldsymbol{\beta}) \tag{30.48d}$$

Hence, the ratio of evolvability to respondability is simple the cosine of the angle θ between $\mathbf{R}$ and $\boldsymbol{\beta}$.

Schluter's Genetic Line of Least Resistance, $\mathbf{g_{max}}$

An important concept, linking the constraints of the breeder's equation (though $\mathbf{G}$) with patterns of divergence among closely-related species, was suggested by Schluter (1996), who argued that populations often diverge in the direction of the most usable genetic variation. Specifically, he defined the **genetic line of least resistance** $\mathbf{g}_{max}$ as the eigenvector corresponding to the lead (largest) eigenvalue of $\mathbf{G}$ (i.e., first principal component of $\mathbf{G}$) and argued that $\mathbf{g}_{max}$ might be informative on patterns of divergence among distant populations.

Schluter argued that even though the breeder's equation likely only holds for a few generations and that $\mathbf{G}$ can change over time, $\mathbf{G}$ may still allow us make general statements

about divergence over modest evolutionary time spans. His idea is that evolution may initially be constrained along those directions with the maximal genetic variance. In particular, he focused on just the first such direction, $\mathbf{g}_{max}$. Figure 30.3 depicts his three predictions. In this figure, the current population is under selective pressure to move towards an optimum. However, much of the genetic variation for this population is along $\mathbf{g}_{max}$, as shown by the initial angle of the ellipse depicting the distribution of genotypic values. Early on, much of the evolution towards to optimum is along this line, with θ_1 measuring the angle between the early divergence and the initial $\mathbf{g}_{max}$. Over time, the departure from $\mathbf{g}_{max}$ will increase as the population approaches its equilibrium. Schluter's first prediction is that, early on, θ should closely track $\mathbf{g}_{max}$. His second prediction is that this association should decay over time, resulting in a negative relationship between θ and divergence time. His final prediction is that evolution should be rather slow if in directions other than $\mathbf{g}_{max}$, and hence θ and amount of divergence should be negatively associated (large *angular* deviations away for $\mathbf{g}_{max}$ should result in little *absolute* divergence, at least early on). How well do these predictions hold up?

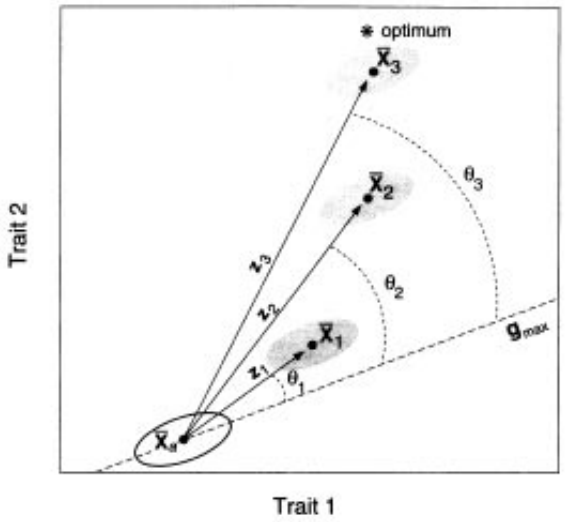


Figure 30.3. Schluter's predictions about population divergence if most of the initial genetic variation is along the first principal component, $\mathbf{g}_{max}$. An ancestral population starts at $\overline{x}_a$ and is under selection pressure to move towards an optimal value. The initial orientation of the breeding values along $\mathbf{g}_{max}$ is shown by the angle of the ellipse depicting the distribution of breeding values, while θ denotes the angle between the vector of divergence and $\mathbf{g}_{max}$ at various points in time. As time progresses, and the population mvoes closer to the opitmal, θ will increase. Compare with Figure 30.2, which shows that evolutionary trajectories usually follow curved paths to reach an optimum. Early on, the trajectory is best approximated by a line along the direction of maximal variation, namely $\mathbf{g}_{max}$. After Schluter (1996).

Schluter looked at morphological divergence data in a small set of vertebrates (stickleback fish, mice, and three species of birds). Let $\mathbf{d}$ denote the vector of differences between the species means, which we can scale to unit length by setting $\mathbf{d}' = \mathbf{d}/\|\,\mathbf{d}\,\|$. Using this scaled divergence vector, we can look at the angle θ between $\mathbf{d}$ and $\mathbf{g}_{max}$, namely the angle between the vector of divergence and the vector (or direction) of maximal additive genetic variation, where

$$\theta = \cos^{-1}(\mathbf{g}_{max}^T \mathbf{d}') \tag{30.49}$$

(A technical aside is that, for his stickleback populations where no clear phylogeny, and hence no ordering of the populations, was evident, Schluter replaced the vector of scaled trait differences $\mathbf{d}$ with the first eigenvector of the covariance matrix of the population means.)

As shown in Figure 30.4, Schluter did indeed observe his three predictions in these data. The left part of this figure shows his first two predictions hold. First, the smallest values of θ occurred between the most recently diverged populations. Second, θ tended to increase with time. The right part of this figure shows that the third prediction also holds: the greater

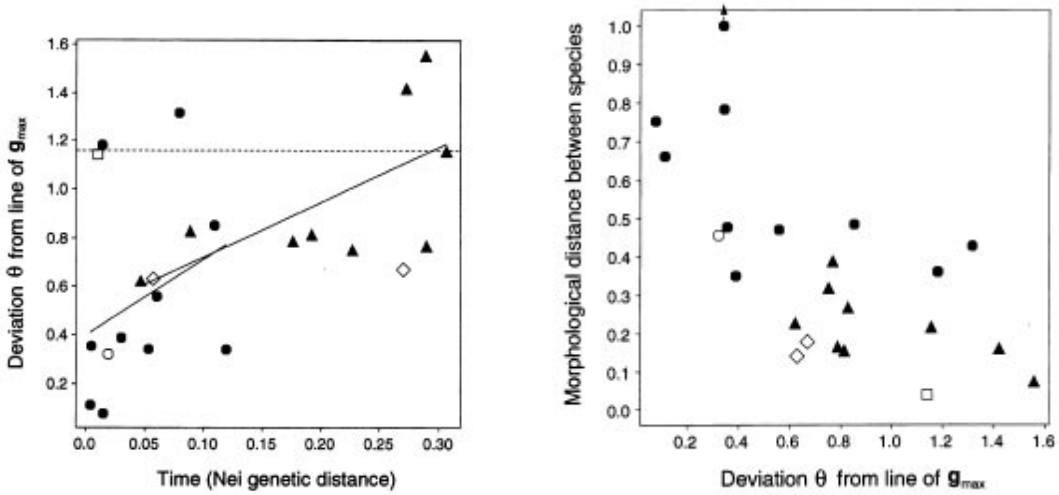


Figure 30.4. Observations on the angle θ between the amount of divergence and $\mathbf{g}_{max}$ (the first eigenvector of $\mathbf{G}$). **Left:** θ is positively correlated with estimated divergence time, so the departures from $\mathbf{g}_{max}$ increase over time. The horizontal dotted line is the random expectation for θ, while the two sloped lines correspond to the regressions using the two largest clades in the study (finches and sparrows). **Right:** Total amount of divergence is inversely related to θ. Thus, when the divergence is away from $\mathbf{g}_{max}$, it tends to be small. Open circles correspond to stickleback fish populations, filled circles to Galapagos finches, squares to flycatchers, triangles to sparrows, and diamonds to mice. θ is reported in radians. After Schluter (1996).

the value of θ, the smaller the total amount of divergence. Surprisingly, Schluter noted that the effect $\mathbf{g}_{max}$ on the absolute amount of divergence showed no tendency to weaken with time (out to at least 4 million years).

Thus there is strong empirical evidence that at least some populations tend to evolve along lines of least genetic resistance (i.e., lines of maximal genetic variance). There are two ways to interpret this observation. The first is that such lines constraint selection. The second is that such lines are also the lines upon which maximal genetic drift occurs (the between-mean variance being proportion to the total amount of genetic variation), see Chapter 6.

Example 30.14. McGuigan et al. (2005) present an interesting study that offered both some support and some counterexamples to Schulter's general findings. They looked at divergence in two species of Australian rainbow fish (genus *Melanotaenia*) that each have populations differentially adapted to lake vs. stream hydrodynamic environments. Divergence between species, as well as divergence within replicate hydrodynamic populations within each species, followed Schluter's results (small angular departures between $\mathbf{d}$ and $\mathbf{g}_{max}$). However, hydrodynamic divergence between lake vs. stream populations in each species were along directions

quite removed from $\mathbf{g}_{max}$ (as well as the other eigenvectors of $\mathbf{G}$ that described most of the genetic variation). Thus, the between- and within-species divergence within the same hydrodynamic environment are consistent with drift, while hydrodynamic divergence had to occur against a gradient of very little genetic variation. Of course, one cannot rule out that the adaptation to these environments resulted in a depletion of genetic variation along these directions.

Is There Genetic Variation in the Direction of Selection?

Schulter's work suggests that, where possible, populations tend to evolve along directions of significant genetic variation. On the other hand, selection tends to erode genetic variation over time (e.g., Chapters 5, 25, 26, 31). Thus, if a population has been selected in a consistent direction for a sufficient amount of time, most of the initial variation has likely been exploited. Further response requires the generation of new variation by mutation or migration (Chapters 26, 27, 31). Thus, we are also left with an empirical question. Is there genetic variation in the direction of selection? Simply looking at heritabilities is insufficient. One subtle, but extremely important, feature of multivariate response is that there can be genetic variation (i.e., non-zero heritabilities) in all traits under selection, but little or no genetics variation along direction that selection is trying to move the population (Dickerson 1955). To see this point, consider the following example.

Example 30.15. Suppose the $\mathbf{G}$ matrix is:

$$\mathbf{G} = \begin{pmatrix} 10 & 20 \\ 20 & 40 \end{pmatrix}$$

Further suppose that selection is optimized by increasing trait one by two units for every unit trait two is decreased, so that

$$\boldsymbol{\beta} = c \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

The resulting response is

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Thus, even though there is considerable additive genetic variation in both traits 1 and 2, there is no response. This occurs because $\mathbf{G}$ as a zero eigenvalue, whose corresponding eigenvector exactly corresponds to our $\boldsymbol{\beta}$ (the angle between these two vectors is zero). Hence, there is *no* additive genetic variance along this particular direction, and hence no response. Likewise, if $\boldsymbol{\beta}$ were only a few degrees away from this eigenvector, the resulting response would be very small. Note that the genetic correlation for this example is positive, re-emphasizing our comments that significant constraints can easily arise in the absence of negative correlations.

Example 30.16. While the above was a cartoon example to emphasize a point, real data does indeed suggest this can occur. Blows et al. (2004) examined 8 cuticular hydrocarbons (CHC) in the fly *Drosophila serata*, which are important cues in mate choice. The first two eigenvalues account for 78% of the original additive genetic variation. Blows also estimated the selection gradient on these traits, which play a role in sexual selection. The resulting two

leading eigenvectors and $\boldsymbol{\beta}$ were as follows:

$$\mathbf{e}_1 = \begin{pmatrix} 0.232 \\ 0.132 \\ 0.255 \\ 0.536 \\ 0.449 \\ 0.363 \\ 0.430 \\ 0.239 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0.319 \\ 0.182 \\ 0.213 \\ -0.436 \\ 0.642 \\ -0.362 \\ -0.014 \\ -0.293 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} -0.099 \\ -0.055 \\ 0.133 \\ -0.186 \\ -0.133 \\ 0.779 \\ 0.306 \\ -0.465 \end{pmatrix}$$

Consider the angle θ between the direction of maximal genetic variation ($\mathbf{e}_1$) and and the optimal direction favored by selection ($\boldsymbol{\beta}$). From Equation A4.2,

$$\cos(\theta) = \frac{\mathbf{e}_1^T \boldsymbol{\beta}}{||\mathbf{e}_1||\ ||\boldsymbol{\beta}||} = \frac{\mathbf{e}_1^T \boldsymbol{\beta}}{\sqrt{\mathbf{e}_1^T\ \mathbf{e}_1}\sqrt{\boldsymbol{\beta}^T \boldsymbol{\beta}}} = \frac{0.147496}{\sqrt{0.99896 \cdot 0.999502}} = 0.1476$$

Giving $\theta = \cos^{-1}(0.1476) = 81.5$ degrees. Thus, the vector of maximal genetic variation and the vector of optimal response are almost at right angles. Likewise, the angle between $\mathbf{e}_2$ and $\boldsymbol{\beta}$ is $\theta = 99.65$ degrees. Thus, very little of the standing additive genetic variation is in the direction of the optimal selection response. While all of the CHC traits showed significant variation, and indeed responded to artifical selection, there is very little useable genetic variation in the direction that sexual selection is trying to move the population.

Blow's Matrix Subspace Projection

How does one quantify both the amount of constraint, as well as the usable volume of genetic variation, implied by the geometry of $\mathbf{G}$? As will be discussed in Chapter 31, one approach for quantifying constraints is to estimate the dimension of $\mathbf{G}$ (the number of non-zero eigenvalues), indicating the number of independent directions in which variability exists. At the other end of the spectrum, Schulter suggested the use of $\mathbf{g}_{max}$ to indicate the primary direction where genetic variation resides. If the leading eigenvector dominates all of the others (and hence accounts for most of the variance), then $\mathbf{g}_{max}$ may indeed be a sufficient descriptor of the usable variation. However, in many cases the first few eigenvalues together may account for most of the variation, so that focusing only on the largest may miss a significant fraction of the variation. Recall that the fraction of variation residing along the direction of the k-th eigenvector is just $\lambda_k / \sum \lambda_i = \lambda_k / \text{trace}(\mathbf{G})$, providing a quick method for accounting for variation. In cases where more than a single eigenvector is required to account for variation, we can consider the usable genetic variation in the **matrix subspace** spanned by those eigenvectors and similarly consider the projection of $\boldsymbol{\beta}$ into that space (in essence, considering the first k terms of Equation 30.45b).

This approach was suggested by Blows et al. (2004), and is further motivated by the common problem that the $\mathbf{G}$ matrix often **ill-conditioned**, in that $\lambda_{max}/\lambda_{min}$ is large. In such cases (as well as others!) estimation of the $\mathbf{G}$ matrix may result in estimates of eigenvalues that are very close to zero or even negative. Negative estimates arise due to sampling, but values near zero may reflect the true biology in that although n traits may be measured, there is very little variation in certain dimensions. In such cases, one might extract a subset of $\mathbf{G}$, for example by taking the leading k eigenvectors. This set forms a **subspace** of the full genetic variance described by $\mathbf{G}$. It is usually the case the $\mathbf{G}$ contains several (perhaps most!) eigenvalues that account for almost no variation (i.e., $\lambda_i / \text{tr}(\mathbf{G}) \simeq 0$). In such cases, most of the genetic variation residues on a lower-dimensional subspace.

We can examine the genetic constraints on this subspace by looking at the **projection** of the full space into this subspace (this is just the matrix extension to the projection of one vector onto another that was discussed earlier). Suppose we have included the first k eigenvectors in our analysis. We can use these to form a **projection matrix** by first defining the matrix $\mathbf{A}$, where

$$\mathbf{A} = (\, \mathbf{e}_1, \quad \mathbf{e}_2, \quad \cdots, \mathbf{e}_k \,) \tag{30.50}$$

so that the $\mathbf{A}$ matrix consists of the first k eigenvectors of $\mathbf{G}$. The resulting projection matrix becomes

$$\mathbf{P}_{roj} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \tag{30.51a}$$

and in particular, the projection $\boldsymbol{\beta}$ onto this subspace of $\mathbf{G}$ (the subspace that essentially contains all of the usable additive variation) is given by

$$\mathbf{p} = \mathbf{P}_{roj}\boldsymbol{\beta} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \boldsymbol{\beta} \tag{30.51b}$$

Example 30.17. Let's reconsider Blow's CHC data. The first two eigenvalues account for roughly 80% of the total variation in $\mathbf{G}$, i.e., $(\lambda_1 + \lambda_2)/\sum \lambda_i = 0.78$. The resulting $\mathbf{A}$ matrix becomes

$$\mathbf{A} = (\, \mathbf{e}_1, \mathbf{e}_2 \,) = \begin{pmatrix} 0.232 & 0.319 \\ 0.132 & 0.182 \\ 0.255 & 0.213 \\ 0.536 & -0.436 \\ 0.449 & 0.642 \\ 0.363 & -0.362 \\ 0.430 & -0.014 \\ 0.239 & -0.293 \end{pmatrix}$$

Applying Equation 30.51a gives an 8×8 projection matrix (not show here), and Equation 30.51b gives the projection vector $\mathbf{b}$ of $\boldsymbol{\beta}$ onto the subspace given by $\mathbf{A}$ as

$$\mathbf{p} = \mathbf{P}_{roj}\boldsymbol{\beta} = \begin{pmatrix} -0.0192 \\ -0.0110 \\ 0.0019 \\ 0.1522 \\ -0.0413 \\ 0.1142 \\ 0.0658 \\ 0.0844 \end{pmatrix}$$

The angle θ between $\boldsymbol{\beta}$ and the projection of $\boldsymbol{\beta}$ into the subspace of the genetic variance is given by

$$\theta = \cos^{-1} \left(\frac{\mathbf{p}^T \boldsymbol{\beta}}{\sqrt{\mathbf{p}^T \mathbf{p}} \sqrt{\boldsymbol{\beta}^T \boldsymbol{\beta}}} \right) = \cos^{-1} (0.223) = 77.1 \text{degrees}$$

Thus the direction of optimal response is 77 degrees away from the total genetic variation (78%) described by this subspace.

Matrix subspace project thus provides a very compact single descriptor of the amount of constraints given $\mathbf{G}$ and $\boldsymbol{\beta}$, returning the angle θ between $\boldsymbol{\beta}$ and the amount of significant genetic variation residing in $\mathbf{G}$. Furthermore (as discussed in length in Chapter 31), it can be much easier to estimate this space where most of the usable genetic variation lies with more confidence than one can estimate a $\mathbf{G}$ matrix. Since much of the evolutionary action is likely to occur in this space, this is a significant advantage. On the other hand, if past (and potentially on-going) selection has eroded away usable genetic variation in certain directions, those are directions of considerable interest.

Conditional Genetic Variance and Conditional Evolvability

A final useful metric of genetic constraints is Hansen's notion of **conditional evolvability** (Hansen et al. 2003b, Hansen 2003, Hansen and Houle 2008). The motivation is simple: how much genetic variation for a given trait is *independent* of the other traits? This concept was first suggested by Zhu (1995) under a slightly different framework. Partitioning our vector of traits into two components, $\mathbf{z}^T = (\mathbf{y}^T, \mathbf{x}^T)$, Equation 30.6b gives the condition genetic variance in $\mathbf{y}$ given $\mathbf{x}$ as

$$\mathbf{G}_{\mathbf{y}\,|\,\mathbf{x}} = \mathbf{G}_{\mathbf{yy}} - \mathbf{G}_{\mathbf{yx}}\,\mathbf{G}_{\mathbf{xx}}^{-1}\,\mathbf{G}_{\mathbf{yx}}^{T} \tag{30.52a}$$

where $\mathbf{G}_{\mathbf{yy}}$ is the covariance matrix for $\mathbf{y}$, $\mathbf{G}_{\mathbf{xx}}$ for $\mathbf{x}$, and $\mathbf{G}_{\mathbf{yx}}$ the covariance matrix between $\mathbf{y}$ and $\mathbf{x}$. Of special interest is the conditional genetic variation for a single trait, so that $\mathbf{y}$ is now a scalar, our focal trait, whose genetic variance is denoted by σ_G^2. In this case, we can simplify Equation 30.52a to

$$\sigma_G(y\,|\,\mathbf{x}) = \sigma_G^2\left(1 - \sigma_G^{-2}\mathbf{G}_{y\mathbf{x}}\mathbf{G}_{\mathbf{xx}}^{-1}\mathbf{G}_{y\mathbf{x}}^{T}\right) \tag{30.52b}$$

where $\mathbf{G}_{y\mathbf{x}}$ is the row vector of the genetic covariances between y and $\mathbf{x}$. Recalling (e.g., Anderson 1984) that the scalar

$$r_{y,\mathbf{x}}^2 = \sigma_G^{-2}\,\mathbf{G}_{y\mathbf{x}}\,\mathbf{G}_{\mathbf{xx}}^{-1}\,\mathbf{G}_{y\mathbf{x}}^{T} \tag{30.52c}$$

is the squared multiple correlation coefficient between y and $\mathbf{x}$, we have

$$\sigma_G(y\,|\,\mathbf{x}) = \sigma_G^2\left(1 - r_{y,\mathbf{x}}^2\right) \tag{30.52d}$$

Notice that this measure, while very useful, is also almost always an *overestimate* of the amount of genetic variance in our focal trait that is unconstrained by the remaining traits in the organism. We can see this directly from Equation 30.52c, in that as we add more traits to $\mathbf{x}$, $r_{y,\mathbf{x}}^2$ never decreases, and most likely increases, even if very slightly.

Example 30.18. Compute the conditional genetic variances for *Tribolium castaneum* pupal weight and thorax width given the (scaled) covariance matrix from Example 30.10. Here

$$\mathbf{G} = \begin{pmatrix} 1.25 & 0.36 \\ 0.36 & 0.20 \end{pmatrix}$$

Since there are only two traits here, both have the same correlation coefficient. We will use the matrix form (Equation 30.52c) to compute this, using pupal weight as our focal trait, giving

$$r_{y\mathbf{x}}^2 = \sigma_G^{-2}\,\mathbf{G}_{y\mathbf{x}}\,\mathbf{G}_{\mathbf{xx}}^{-1}\,\mathbf{G}_{y\mathbf{x}}^{T} = \frac{1}{1.25}\,0.36\,\frac{1}{0.20}\,0.36 = 0.52$$

Thus the unconditional variance in either trait is only 48% of the actual variance, giving an unconditional additive variance of 0.60 for pupal weight and 0.096 for thorax width.

$\mathbf{G}_{\mathbf{y}\,|\,\mathbf{x}}$ is a measure of how much of the genetic variance in $\mathbf{y}$ is independent of the genetic variance in other traits. Suppose there $\mathbf{x}$ is under a constraint for no change, such as would occur if these components were under strong stabilizing selection and at (or very near) their optimal values. The conditional genetic variance is the variance in $\mathbf{y}$ that is available given these constraints. In this case, the expected responsse in $\mathbf{y}$ is given by

$$\mathbf{R}_{\mathbf{y}} = \mathbf{G}_{\mathbf{y}\,|\,\mathbf{x}}\boldsymbol{\beta}_{\mathbf{y}} \tag{30.5x}$$

where $\boldsymbol{\beta}_{\mathbf{y}}$ is the vector of selection gradients for the traits in $\mathbf{y}$ (Hansen 2003). Hansen and Houle (2008) refer to the projection of $\mathbf{R}_{\mathbf{y}}$ onto $\boldsymbol{\beta}_{\mathbf{y}}$ as the **conditional evolvability** (Figure 30.5).

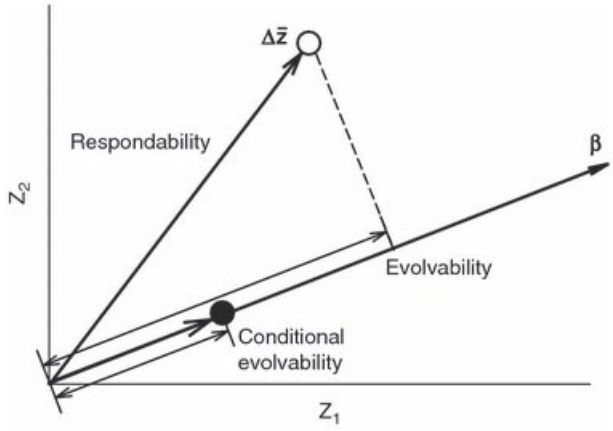


Figure 30.5. The relationshipship between respondability (the scaled lenght of the response vector $\mathbf{R}$), evolvability (the projection of the response vector onto $\boldsymbol{\beta}$) and conditional evolvablity (the projection of the response vector using the conditional genetic variance onto $\boldsymbol{\beta}$). Respondability is the ability for the population to change, evolvability the ability to increase along the direction that optimally increases fitness, and conditional evolvability the ability to along $\boldsymbol{\beta}$ given constraints imposed by stabilizing (or other) selection on genetically correlated traits. After Hansen and Houle (2008).

Literature Cited

Anderson, T. W. 1984. *An introduction to multivariate statistical analysis*, 2nd ed. Wiley, New York. [30]

Arnold, S. J. 1981. Behavioral variation in natural populations. I. Phenotypic, genetic and environmental correlations between chemoreceptive responses to prey in the garter snake, *Thamnophis elegans. Evolution* 35: 489–509. [30]

Arnold, S. J. 1992. Constraints on phenotypic evolution. *Amer. Natl.* 140: S85–S107. [30]

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112-113. 9–32. [30]

Bell, A. E., and M. J. Burris. 1973. Simultaneous selection for two correlated traits in si Tribolium. *Genet. Res. Camb.* 21: 29–46. [30]

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. [32]

Berger, P. J., and W. R. Harvey. 1975. Realized genetic parameters form index selection in mice. *J. Animal Sci.* 40: 38–47. [30]

Björklund, M. 1996. The importance of evolutionary constraints in ecological time scales. *Evol. Ecology* 10: 423–431. [30]

Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and fitness surface for multiple male sexually selected traits. *Amer. Natl.* 163: 329–340. [30]

Bohren, B. B., W. G. Hill, and A. Robertson. 1966. Some observations on asymmetrical correlated responses to selection. *Genet. Res.* 7: 44–57. [30]

Boulding, E. G., and T. K. Hay. 1993. Quantitative genetics of shell form of an intertidal snail: constrains on short-term response to selection. *Evolution* 47: 576–592. [30]

Bruns, E., and W. R. Harvey. 1976. Effects of varying selection intensity for two traits on estimation of realized genetic parameters. *J. Animal Sci.* 42: 291–298. [30]

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* 110: 155–171. [30]

Cheverud, J. M. 1988. A comparison of genetic and phenotypic correlations. *Evolution* 42: 958–968. [30]

Clayton, G. A., G. R. Knight, J. A. Morris, and A. Robertson. 1957. An experimental check of quantitative genetical theory. III. Correlated responses. *J. Genetics* 55: 171–180. [30]

Conner, J., and S. Via. 1992. Natural selection on body size in *Tribolium:* possible genetic constraints on adapative evolution. *Heredity* 69: 73–83. [30]

Crnokrak, P., and D A. Roff. 1998. The genetic basis of the trade-off between calling and wing morph in males of the cricket *Gryllus firmus. Evolution* 52: 1111–1118. [30]

Dickerson, G. E. 1955. Genetic slippage in response to selection for multiple objectives. *Cold Spring Harb Symp Quant Biol.* 20: 213–224. [30]

Falconer, D. S. 1952. The problem of environment and selection. *Amer. Natl.* 86: 293–298. [30]

Falconer, D. S. 1960a. *Introduction of quantitative genetics.* Oliver and Boyd, London. [30]

Falconer, D. S. 1960b. Selection of mice for growth on high and low planes of nutrition. *Genet. Res. Camb.* 1: 91–113. [30]

Fellowes, M. D. E., A. R. Kraaijeveld, and H. C. J. Godfray. 1998. Trade-off assoicated with selection for increased ability to resist parasitoid attack in *Drosophila melaogaster. Proc. R. Soc. Lond. B* 265: 1553–1558. [30]

Fellowes, M. D. E., A. R. Kraaijeveld, and H. C. J. Godfray. 1999a. Cross-resistance following artificial selection for increased defense against parasitoids in *Drosophila melanogaster. Evolution* 53: 966–972. [30]

Fellowes, M. D. E., A. R. Kraaijeveld, and H. C. J. Godfray. 1999b. Association between feeding rate and parasitoid resistance in *Drosophila melanogaster. Evolution* 53: 1302–1305. [30]

Gromko, M. H., A. Briot, S. C. Jensen, and H. H. Fukui. 1991. Selection for copulation duration in *Drosophila melanogaster:* predictability of direct response versus unpredictability of correlated response. *Evolution* 45: 69–81. [30]

Gunsett, F. C., K. N. Andriano, and J. J. Rutledge. 1982. Estimating the precision of estimates of genetic parameters realized from multiple-trait selection experiments. *Biometrics* 38: 981–989. [30]

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. [30]

Janssens, M. J. J. 1979. Co-heritability: its relation to correlated response, linkage, and pleiotropy in cases of polygenic inheritance. *Euphytica* 28: 601–608. [30]

Haldane, J. B. S. 1949. Suggestions as to quantitative measurement of rates of evolution. *Evolution* 3: 51–56. [30]

Hansen, T. F. 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evovlability. *BioSys.* 69: 83–94. [30]

Hansen, T. F., W. S. Armbruster, M. L. Carlson, and C. Pélabon. 2003b. Evolvability and genetic constraints in *Dalechampia* blossoms: genetic correlations and conditional evolvability. *J. Exp. Zoo.* 296B: 23–29. [30]

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. *J. Evol. Bio.* 21: 1201–1219. [30]

Hansen, T. F., C. Pélabon, W. S. Armbruster, and M. L. Carlson. 2003a. Evolvability and genetic constraints in *Dalechampia* blossoms: components of variance and measures of evolvability. *J. Evol. Biol.* 16: 754–766. [30]

Harvey, W. R. 1972. Direct and indirect response in two-trait selection experiments in mice. *Proc. 21st Annual Session, National Breeders Round Table.* Kansas City, Mo.

Harvey, W. R., and G. D. Bearden. 1962. Tables of expected genetic progress in each of two taits. *USDA Agricultural Research Service Publication* no. 20-12. [30]

Hazel, L. N. 1943. The genetic basis for constructing selection indices. *Genetics* 28: 476–490. [30]

Hereford, J., T. F. Hansen, and D. Houle. 2004. Comparing strenghts of directional selection: how strong is strong? *Evolution* 58: 2133–2143. [30]

Houle, D. 1992. Comparising evolvability and variability of quantitative traits. *Genetics* 130: 195–204. [30]

Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. *Evolution* 46: 954–971. [30]

Kraaijeveld, A. R., and H. C. J. Godfray. 1997. Trade-off between parasitioid resistance and larval competitive ablity in *Drosophila melanogaster. Nature* 389: 278–280 [30]

Kraaijeveld, A. R., E. C. Limentani and H. C. J. Godfray. 2001. Basis of the trade-off between parasitioid resistance and larval competitive ability in *Drosophila melanogaster. Proc. R. Soc. Lond. B* 268: 259–261. [30]

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33: 402–416. [30]

Lande, R. 1980. Microevolution in relation to macroevolution. *Paelbio.* 6: 233–238. [30]

Magee, W. T. 1965. Estimating response to selection. *J. Animal Sci.* 24: 242–247. [30]

McGuigan, K. 2006. Studying phenotypic evolution using multivariate quantitative genetics. *Mol. Ecology* 15: 883–869. [30]

McGuigan, K., S. F. Chenoweth, and M. W. Blows. 2005. Phenotypic divergence along lines of genetic variance. *Amer. Nat.* 165: 32-43. [30]

Merilä, J., M. Björklund, and L. Gustafsson. 1994. Evolution of morpholoigcal differences with moderate genetic correlatiosn among traists as exemplified by two flycatcher species (*Ficedula;* Musciapidae). *Biol. J. Linn. Soc.* 52: 19–30. [30]

Nagylaki, T. 1991. Error bounds for the fundamental and secondary theorems of natural selection. *Proc. Natl. Acad. Sci.* 88: 2402–2406. [30]

Olson, E. C., and R. L. Miller. 1958. *Morphological integration.* University of Chicago Press. [30]

Pease, C. M., and J. J. Bull. 1988. A critique of methods for measuring life history trade-off. *J. Evol. Biol.* 1: 293–303. [30]

Pearson, K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Phil. Trans. Roy. Soc. Lond.* A 200: 1–66. [30]

Polak, M., and W. T. Starmer. 2001. The quantitative genetics of fluctuating asymmetry. *Evolution* 55: 498–511. [30]

Polly, P. D. 2008. Developmental dyamics and G-matrices: can morphometric spaces be used to model phenotypic evolution? *Evol. Bio.* 35: 83–96. [30]

Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* 46: 616–626. [30]

Reeve, E. C. R. 1950. Genetical aspects of size allometry. *Proc. Roy. Soc. Lond. B*: 137: 515–518. [30]

Rice, S. H. 2008. The **G**-matrix as one piece of the phenotypic evolution puzzle. *Evol. Biol.* xx: xx-xx. [30]

Riska, B. 1989. Composite traits, selection response, and evolution. *Evolution* 43: 1172–1191. [30]

Riska, B., and W. R. Atchley. 1985. Genetics of growth predict patterns of brain-size evolution. *Science* 299: 668–671. [30]

Rutledge, J. J., E. J. Eisen, and J. E. Legates. 1973. An experimental evalaution of genetic correlations. *Genetics* 75: 709–726. [30]

Rogers, A., and A. Mukherjee. 1992. Quantitative genetics of sexual dimorphism in human body size. *Evolution* 46: 226–234. [30]

Searle, S. R. 1965. The value of indirect selection. I. mass selection. *Biometrics* 21: 682–707 [30]

Searle, S. R. 1978. The value of indirect selection. II. progeny testing. *Theor. Appl. Gent.* 51: 289–296. [30]

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* 50: 1766–1774. [30]

Schluter, D., T. D. Price, and L. Rowe. 1991. Conflicting selection pressures and life hisotry trade-offs. *Proc. R. Soc. Lond. B* 246: 1–17. [30]

Simpson, G. G. 1944. *Tempo and mode in evolution.* Columbia University Press, New York. [30]

Turner, H. N. 1959. Ratios as criteria for selection in animal and plant breeding, with particular refernce to efficiency of food conversion in sheep. *Aust. J. Agri. Res.* 10: 565–580. [30]

Via, S., and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* 39: 505–522. [30]

van Tienderen, P. H. 2000. Elasticities and the link between demographic and evolutionary dynamics. *Ecology* 81: 666–679. [30]

Wagner, G. P. 1984. On the eigenvalue distribution of genetic and phenotypic dispersion matrices: evidence for a nonrandom organization of quantitative character variation. *J. Math. Biol.* 21: 77–95. [30]

Willis, J. H., J. A. Coyne, and M. Kirkpatrick. 1991. Can one predict the evolution of quantitative characters without genetics? *Evolution* 45: 441–444. [30]

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159. [30]

Young, S. S. Y., G. H. Brown, H. N. Turner, and C. H. S. Dolling. 1965. Genetic and phenotypic parameters for body weight and greasy fleece weight at weaning in Australan merino shepp. *Aust. J. Agri. Res.* 16: 997–1009. [30]

Zhu, J. 1995. Analysis of conditional genetic effects and variance components in developmental genetics. *Genetics* 141: 1663–1639. [30]