Lecture 13
Genetic Correlations and Multivariate Selection Response

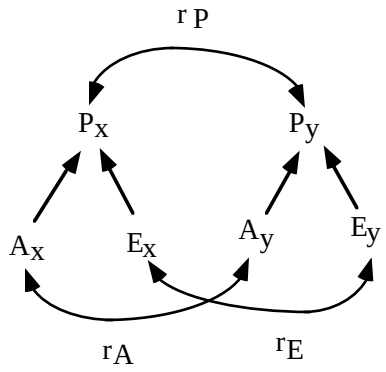
ECOL 519A, March 2007. University of Arizona

Selection rarely (if ever!) simply acts on a single trait. In reality, selection occurs simultaneously on an entire suite of characters. These next two lectures deal with the theory of selection response and the measurement of selection (13 and 14, respectively) on multiple traits. The key concerns required for this theory are the phenotype and genetic (breeding values) correlations/covariances, and we start by discussing how to estimate these.

PHENOTYPIC AND GENETIC CORRELATIONS

Many characters are positively or negatively correlated at the level of phenotype (e.g., height and weight, forearm length and digit span) and we can directly measure the phenotypic correlation, r_P, between two traits X and Y.

As the following (path) diagram indicates, the phenotypic correlation r_P between two traits is generated by correlations between the genetic (r_A) and/or environmental (r_E) values of X and Y. In the figure, double-headed arrows imply possible correlations between variables.



r_A = correlation of breeding values arises from two sources

- pleiotropic effects of loci on both traits; correlation from pleiotropy indicates the extent to which the character is controlled by the same genes
- linkage disequilibrium, which will decay over time

r_E = correlation of environmental deviations

- includes non-additive genetic effects
- arises from exposure of the two traits to the same individual environment

To express r_P in terms of r_A and r_E, first note that

$$cov_P = r_P\,\sigma_{XP}\,\sigma_{YP}, \quad cov_A = r_A\,\sigma_{XA}\,\sigma_{YA}, \quad cov_E = r_E\,\sigma_{XE}\,\sigma_{YE}$$

The covariance of phenotypic values = the sum of the genetic and environmental covariances:

$$cov_P = cov_A + cov_E$$

so that

$$r_P\,\sigma_{XP}\,\sigma_{YP} = r_A\,\sigma_{XA}\,\sigma_{YA} + r_E\,\sigma_{XE}\,\sigma_{YE} \tag{13.1a}$$

describes the relationship between phenotypic, genetic and environmental correlations.

This expression can be simplified somewhat. First, define h_X and h_Y as the square roots of the heritabilities of characters X and Y, with

$$h = \frac{\sigma_A}{\sigma_P}, \quad \sigma_A = h\sigma_P$$

Likewise define $e^2 = 1 - h^2$

$$e^2 = \frac{\sigma_E^2}{\sigma_P^2}, \quad e = \sqrt{1-h^2} = \frac{\sigma_E}{\sigma_P}, \quad \sigma_E = e\sigma_P$$

Substituting $\sigma_A = h\sigma_P$ and $\sigma_E = e\sigma_P$ into equation (13.1a) gives

$$\begin{aligned} r_P &= h_X\,h_Y\,r_A + e_X\,e_Y\,r_E \\ &= r_A\,h_X\,h_Y + r_E\,\sqrt{(1-h_x^2)(1-h_Y^2)} \end{aligned} \tag{13.1b}$$

Hence, the phenotypic correlation is a function of the heritabilities of the traits and the genetic and environmental correlations. If heritabilities are high, the genetic correlation is more important; if heritabilities are low, the environmental correlation is more important. In practice, phenotypic and genetic correlations often have the same sign and are of similar magnitude, but this is not always the case and hence should not be assumed.

Estimating the Genetic Correlation

Methods for estimating r_A and r_E are analogous to estimating heritabilities from resemblance between relatives. The difference is that with a single trait, the covariance between the trait value in two relatives provides an estimate of the additive genetic variance of that trait, while with two traits, the covariance of trait X in one relative and trait Y in the other provides an estimate of the additive genetic covariance between the two traits.

Half sibs

The covariance of traits X and Y between sires = $cov_{XY} = (1/4)cov_A$

The variance between sires of trait $X = \sigma^2_{SX} = (1/4)\sigma^2_{AX}$

The variance between sire of trait $Y = \sigma^2_{SY} = (1/4)\sigma^2_{AY}$

Therefore, $r_A = cov_{XY}/(\sigma_{SX}\sigma_{SY})$

Parent-Offspring

The covariance of trait X in the offspring with trait Y of the parents

$$cov(X_O, Y_P) = (1/2)cov_{A_X,A_Y}$$

The covariance of trait X in the offspring with trait X of the parents

$$cov(X_O, X_P) = (1/2)\sigma^2_{A_X}$$

The covariance of trait Y in the offspring with trait Y of the parents

$$cov(Y_O, Y_P) = (1/2)\sigma^2_{A_Y}$$

Hence,

$$r_A = \frac{cov(X_O, Y_P)}{\sqrt{cov(X_O, X_P)\cdot cov(Y_O, Y_P)}}$$

Estimates of genetic correlation have very large sampling errors, typically requiring extremely large experiments for precise estimates.

Correlated Response to Selection

When characters are genetically correlated (i.e., breeding values for those traits are correlated), selection solely on one will result in a correlated change in the second. Such a change in the unselected character is called a **correlated response**.

We compute the expected correlated response in Y given selection on X as follows. The response to selection of character X (the mean value of offspring of selected parents) is (by definition) the mean breeding value of the selected group. Thus the change in character Y in response to selection on X is the regression of the breeding value of Y on the breeding value of X. For standard regression theory (Lecture 1), the slope of this regression is given by

$$b_{A_Y|A_X} = \frac{cov_A}{\sigma^2_{A_X}} = \frac{r_A\,\sigma_{A_X}\,\sigma_{A_Y}}{\sigma^2_{A_X}} = r_A\,\frac{\sigma_{A_Y}}{\sigma_{A_X}} \tag{13.2a}$$

Recalling first that a regression passes through the mean of both variables (with $y-\mu_y = b_{y\,|\,x}[x-\mu_x]$) and second that the breeding values have mean zero ($\mu_{A_X} = \mu_{A_Y} = 0$), the regression of the breeding values of Y on the breeding values of X is just

$$Y = b_{A_Y|A_X}\,X = r_A\,\frac{\sigma_{A_Y}}{\sigma_{A_X}}\,X \tag{13.2b}$$

The response of the directly selected character X is

$$R_X = i_X\,h^2_X\,\sigma_{PX} = i\,(\sigma^2_{A_X}/\sigma^2_{P_X})\,\sigma_{P_X} = i_X\,\sigma^2_{A_X}/\sigma_{P_X} = i_X\,h_x\,\sigma_{AX}$$

where $i_X = S_x/\sigma_X$. Recalling our comment above that R_X is the change in the breeding value of X, the correlated response of character Y is

$$\begin{aligned} CR_Y &= b_{A_Y|A_X}R_X \\ &= (r_A\,\sigma_{A_Y}/\sigma_{A_X})\,(i_X\,h_X\,\sigma_{A_X}) \\ &= r_A\,\sigma_{A_Y}\,i_X\,h_X \end{aligned} \tag{13.3a}$$

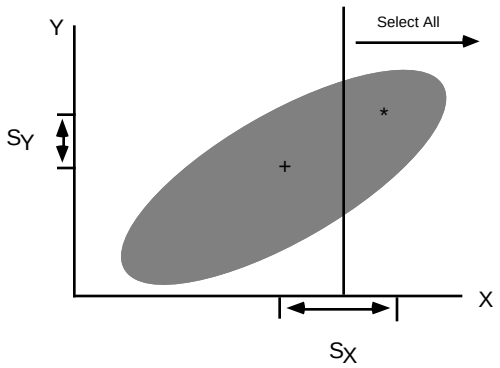
Substituting $\sigma_{A_Y} = h_Y\,\sigma_{PY}$ gives

$$CR_Y = i_X\,h_X\,h_Y\,r_A\,\sigma_{P_Y} \tag{13.3b}$$

Noting that the direct response on X is $R_X = i_X\,h^2_X\,\sigma_{P_X}$, we see that h^2_X and $h_X\,h_Y\,r_A$ play similar roles, resulting in the later being called the **co-heritability**.

Correlated Selection Differentials

When selection is applied to character X, character Y will show a correlated selection differential when Y is phenotypically correlated with X. The figure below shows the change in the bivariate mean before ($+$) and after ($*$) selection. In this case, there is truncation selection only on trait X, but there is a correlated within-generation change in Y, with the mean of Y in the selected parents differing from the mean before selection.



Note that the correlated selection differential simply measures the *within-generation* change. Whether this translates into a *between-generation* change (i.e., a response to selection) depends on whether there is any genetic correlation between the characters. With only phenotypic correlations, none of the correlated within-generation change is passed on to the offspring.

The reason for this can be illustrated by a hypothetical example, with truncation selection only on X. Suppose that $r_A = 0$ and $r_E > 0$; therefore r_P and S_Y will be > 0. The selected individuals, however, will not have higher than average breeding values for Y, in spite of their higher than average phenotypic values. Therefore the correlated response will be zero, regardless of the magnitude of h^2_Y and S_Y.

Indirect Selection

There are two ways the mean of a character X can change by selection:

- as a direct response to selection for trait $X (R_X)$
- as a correlated, or indirect, response to direct selection for trait $Y (CR_X)$

The relative magnitudes of the change in mean is given by the ratio

$$\frac{CR_X}{R_X} = \frac{i_Y\, r_A\, \sigma_{AX} h_Y}{i_X\, h_X\, \sigma_{AX}} = \frac{i_Y\, r_A\, h_Y}{i_X\, h_X} \tag{13.4}$$

It follows that the correlated response of X to selection for Y will be greater than direct response to selection for X when $i_Y\, r_A\, h_Y > i_X\, h_X$, 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.

General Multivariate Selection Response

The multivariate breeders equation partitions the roles of phenotypic and genetic correlations in selection response. We will first introduce the equation and then give its derivation a little bit later.

Suppose there are n traits under selection and we place the n selection differentials for each trait into a column vector $\mathbf{S}$,

$$\mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{pmatrix}$$

Likewise, define the **phenotypic** and **additive genetic covariance matrices**, $\mathbf{P}$ and $\mathbf{G}$, respectively, as matrices whose ij-th element is the covariance (phenotype or additive genetic) between traits i and j. Note that the diagonal elements are the variances. For example, for two characters

$$\mathbf{P} = \begin{pmatrix} \sigma^2(P_1) & \sigma(P_1, P_2) \\ \sigma(P_1, P_2) & \sigma^2(P_2) \end{pmatrix}, \quad \text{and} \quad \mathbf{G} = \begin{pmatrix} \sigma^2(A_1) & \sigma(A_1, A_2) \\ \sigma(A_1, A_2) & \sigma^2(A_2) \end{pmatrix}$$

Let $\mathbf{R}$ denote the column vector of selection responses, so that the ith element in the list is R_i, the change in the mean of character i following one generation of selection. The response to selection becomes

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} \tag{13.5}$$

This equation is often referred to as the **multidimensional breeders' equation**. Recall that the response for a single character under selection is $R = h^2 S = \sigma_A^2(\sigma_P^2)^{-1}S$. In the multidimensional case, the genetic and phenotypic variance are replaced by variance-covariance matrices and we use matrix inversion and multiplication.

The Directional Selection Gradient

The multivariate breeder's equation can also be written as

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} \tag{13.6a}$$

where

$$\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S} \tag{13.6b}$$

is called the **directional selection gradient**. Recall the univariate version of β developed in Lecture 5, $\beta = S/\sigma_z^2 = (1/\sigma_z^2)S$. The multivariate version follows by replacing the inverse of the phenotypic variance with the inverse of the phenotypic covariance matrix. As we show in the next lecture, β_i, the ith element of $\boldsymbol{\beta}$, measures the amount of direct selection on trait i (i.e., the effects of correlated selection differentials are removed). Note that $\mathbf{P}\boldsymbol{\beta} = \mathbf{S}$, so that the selection differential on trait j can be written as

$$S_j = \sigma^2(P_j)\,\beta_j + \sum_{i \neq j} \sigma(P_j, P_i)\,\beta_i \tag{13.7}$$

The first term represents the contribution from direct selection on trait j and the sum term the contribution to the within-generation change due to direct selection on phenotypically correlated traits. The observed within-generation change in the mean of trait j thus represents the contributions from both any direct selection on trait j ($\beta_j \neq 0$) plus a correlated response to direct selection on other traits ($\beta_k \neq 0$) that are phenotypically-correlated with trait j ($\sigma(P_j, P_i) \neq 0$). Thus a non-zero selection differential on trait i is NOT sufficient to imply that there is direct selection on i. Likewise, a zero selection differential in i is also NOT sufficient to imply that there is no direct selection on i.

In a similar fashion, note that we can (from the rules of matrix multiplication) write the response in trait j as

$$R_j = \sigma^2(A_j)\,\beta_j + \sum_{i \neq j} \sigma(A_j, A_i)\,\beta_i \tag{13.8}$$

where the first term is the change due to direct selection on trait j and the sum is the indirect contributions from the correlated response of selection on other traits. Here, as with Equation 13.7, a non-negative response is NOT sufficient to assume direct selection on trait i nor is a zero response sufficient to assume no direct selection.

Derviation of the Multivariate Breeders' Equation

The derviation of the multivariate breeders' equation follows from features of the multivariate normal (Lecture 2). Recall that conditional distributions of subvectors from a multivariate normal are also multivariate normal. In particular, if we partition a MVN distributed vector n-dimensional column $\mathbf{x}$ 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

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu_1} \\ \\ \boldsymbol{\mu_2} \end{pmatrix} \qquad \text{and} \qquad \mathbf{V} = \begin{pmatrix} \mathbf{V}_{\mathbf{x_1}\mathbf{x_1}} & \mathbf{V}_{\mathbf{x_1}\mathbf{x_2}} \\ \\ \mathbf{V}^T_{\mathbf{x_1}\mathbf{x_2}} & \mathbf{V}_{\mathbf{x_2}\mathbf{x_2}} \end{pmatrix} \tag{13.9}$$

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

$$\boldsymbol{\mu}_{\mathbf{x_1}|\mathbf{x_2}} = \boldsymbol{\mu_1} + \mathbf{V}_{\mathbf{x_1x_2}}\mathbf{V}^{-1}_{\mathbf{x_2x_2}}(\mathbf{x_2} - \boldsymbol{\mu_2}) \tag{13.10a}$$

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

$$\mathbf{V}_{\mathbf{x_1}|\mathbf{x_2}} = \mathbf{V}_{\mathbf{x_1x_1}} - \mathbf{V}_{\mathbf{x_1x_2}}\mathbf{V}^{-1}_{\mathbf{x_2x_2}}\mathbf{V}^{T}_{\mathbf{x_1x_2}} \tag{13.10b}$$

Likewise, the regression of $\mathbf{x_1}$ on $\mathbf{x_2}$ is given by

$$\mathbf{x_1} = \boldsymbol{\mu_1} + \mathbf{V}_{\mathbf{x_1x_2}}\mathbf{V}^{-1}_{\mathbf{x_2x_2}}(\mathbf{x_2} - \boldsymbol{\mu_2}) + \mathbf{e} \tag{13.11a}$$

where

$$\mathbf{e} \sim \text{MVN}_m\left(\mathbf{0}, \mathbf{V}_{\mathbf{x_1}|\mathbf{x_2}}\right) \tag{13.11a}$$

A direct application of these results is the multivariate breeders' 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 $\mathbf{g} \sim$ 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,

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

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

$$\begin{pmatrix} \mathbf{g} \\ \mathbf{z} \end{pmatrix} \sim \text{MVN}\left(\begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \mathbf{G} & \mathbf{G} \\ \mathbf{G} & \mathbf{P} \end{pmatrix}\right) \tag{13.12}$$

From Equation 13.10a and 13.10b, the conditional distribution of $\mathbf{g}$ given $\mathbf{z}$ is MVN with mean

$$\boldsymbol{\mu}_{\mathbf{g}\,|\,\mathbf{z}} = \boldsymbol{\mu} + \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}) \tag{13.13a}$$

and variance-covariance matrix

$$\mathbf{V}_{\mathbf{e}} = \mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{G} \tag{13.13b}$$

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

$$\mathbf{g} - \boldsymbol{\mu} = \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}) + \mathbf{e} \tag{13.14a}$$

where

$$\mathbf{e} \sim \text{MVN}(\mathbf{0}, \mathbf{V}_{\mathbf{e}}) \tag{13.14b}$$

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 $\mathbf{V}_{\mathbf{e}}$. The variance-covariance matrix of the residual vector $\mathbf{e}$ is independent of the actual value of $\mathbf{z}$, and hence the regression of $\mathbf{G}$ on $\mathbf{z}$ is both linear (from Equation 13.13a) and homoscedastic (Equation 13.13b). For example, with only a single character, $g = A$ and $\mathbf{G} = \sigma^2_A$, and Equation 13.14a (recalling the $\mu_A = 0$) implies

$$A = \sigma^2_A\sigma_z^{-2}(z - \mu) + e = h^2(z - \mu) + e \tag{13.14d}$$

where

$$\sigma^2_e = \sigma^2_A - \sigma^2_A\sigma_z^{-2}\sigma^2_A = \sigma^2_A(1 - h^2) \tag{13.14e}$$

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

$$\begin{aligned}\mathbf{R} &= E[\mathbf{G}\mathbf{P}^{-1}(\mathbf{z}-\boldsymbol{\mu})+\mathbf{e}] \\ &= \mathbf{G}\mathbf{P}^{-1}E[(\mathbf{z}-\boldsymbol{\mu})]+E(\mathbf{e}) \\ &= \mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \mathbf{G}\boldsymbol{\beta}\end{aligned} \tag{13.15}$$

as obtained (in various forms) by Young and Weiler (1960), Harvey and Bearden (1962) and Lande (1979). It is important to note that all the caveats of the univariate breeder's equation (it is compromised if epistasis, G × E, maternal effects, etc. are present) also apply to the multivariate breeders' equation.

Realized Selection Gradients

The normal use of the multivariate breeder's equation is to predict response $\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$ given the vector of directional selection gradients $\boldsymbol{\beta}$ and the additive-genetic covariance matrix $\mathbf{G}$. However, we can also use this equation to estimate the long-term selection gradient on a trait given some observed amount of population differentation, with $\mathbf{R} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$, where $\boldsymbol{\mu}_i$ is the mean of population i. Pre-multiplying both sides of the breeder's equation by $\mathbf{G}^{-1}$ recovers

$$\boldsymbol{\beta} = \mathbf{G}^{-1}\mathbf{R} \tag{13.16}$$

Equation 13.16 requires the strong assumption of constancy of $\mathbf{G}$ over long periods of time.

Example 13.1. Suppose you are looking at nose lenght (z_1) and head size (z_2) between a population of mice on the mainland and a divergent population on a nearby island. On the mainland, $\mu_1 = 20$ and $\mu_2 = 30$, while on the island $\mu_1 = 10$ and $\mu_2 = 35$, so that

$$\mathbf{R} = \begin{pmatrix} 20-10 \\ 30-35 \end{pmatrix} = \begin{pmatrix} 10 \\ -5 \end{pmatrix}$$

Suppose the genetic covariance matrix for these two traits has been stable over time and is

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

The realized selection gradient to obtain this response is estimated as

$$\boldsymbol{\beta} = \begin{pmatrix} 20 & -10 \\ -10 & 40 \end{pmatrix}^{-1} \begin{pmatrix} 10 \\ -5 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0 \end{pmatrix}$$

Thus, even though character 2 decreased on the island, there was no direct selection on this trait, rather this change is entirely due to a correlated response to direct selection to increase trait 1.

The Effects of Genetic Correlations: Direct and Correlated Responses

While the use of $\boldsymbol{\beta}$ removes any further evolutionary effect of phenotypic correlations, additive genetic correlations strongly influence the response to selection. If n characters are under selection, and $\sigma(A_j, A_i)$ is the additive genetic covariance between traits i and j, then the response in character i to a single generation of selection is

$$R_i = \sum_{j=1}^{n} \sigma(A_j, A_i)\,\beta_j = \sigma^2(A_i)\,\beta_i + \sum_{j \neq i} \sigma(A_j, A_i)\,\beta_j \tag{13.17}$$

so that response has a component due to direct selection on that character ($\sigma^2(A_i)\,\beta_i$) plus an addition component due to selection on all other genetically correlated characters.

Even if direct selection only occurs on character i, other characters genetically correlated with i also change,

$$R_j = \sigma(A_j, A_i)\,\beta_i$$

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

$$\frac{R_j}{R_i} = \frac{\sigma(A_j, A_i)}{\sigma^2(A_j)} = \rho_A(ij)\sqrt{\frac{\sigma^2(A_j)}{\sigma^2(A_i)}} \tag{13.18}$$

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

Realized Genetic Correlations

Equation 13.18 suggests that we can estimate genetic correlations by comparing direct and corrected responses. These *realized genetic correlations* are the analogue of the realized heritabilities introduced for univariate responses in Lecture 10. Suppose direct selection is applied to a single character (for example by truncation selection on character i). Denote the direct response as R_i and the correlated response in character j due to selection on character i as $CR_{j\,|\,i}$. By jointly (i) applying direct selection to i and looking at the correlated response in j and (ii) applying direct selection to j and looking at the correlated response in i, Equation 13.18 implies

$$\left(\frac{\mathrm{CR}_{i\,|\,j}}{\mathrm{R}_j}\right)\left(\frac{\mathrm{CR}_{j\,|\,i}}{\mathrm{R}_i}\right) = \left(\frac{\sigma(A_j, A_i)}{\sigma^2(A_j)}\right)\left(\frac{\sigma(A_j, A_i)}{\sigma^2(A_i)}\right) = \rho_A^2(ij) \tag{13.19}$$

Evolutionary Constraints Imposed by Genetic Correlations

One immediate consequence of the Breeders' Equation (13.6a) is that a character under selection need not necessarily change in the direction most favored by natural selection if the correlated response on other characters under selection is stronger. For example, fitness may maximally increase if μ_2 decreases, so that $\beta_2 < 0$. However, if the sum of correlated responses is positive, then μ_2 may increase. Thus, a character may be dragged off its optimal value by correlated responses on other traits. However, once these responses (driven by strong directional selection) reach their equilibrium, then $\beta_i \simeq 0$, at which point the response from β_2 dominates.

In general, $\mathbf{R} \neq \boldsymbol{\beta}$, the direction of change that results in the largest increase in mean population fitness. Thus, the effect of the additive-genetic covariance matrix $\mathbf{G}$ is to constrain the selection response from its optimal value. The mean vector changes in the direction most favored by selection if and only if

$$\mathbf{G}\boldsymbol{\beta} = \lambda\boldsymbol{\beta} \tag{13.20}$$

which only occurs when $\boldsymbol{\beta}$ is an eigenvector of $\mathbf{G}$ (as discussed below). Note that even if $\mathbf{G}$ is a diagonal matrix (there is no correlation between the additive genetic values of the characters under selection) Equation 13.20 is usually not satisfied. In fact, only when we can write $\mathbf{G} = \sigma_A^2\mathbf{I}$ is Equation 13.20 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 directional most favored by natural selection. Differences in the amounts of additive genetic variances between characters, in addition to non-zero additive-genetic covariances, also impose constraints on character evolution.

Example 13.2 Consider the following phenotypic and genetic covariance matrices and vector of selection differentials,

$$\mathbf{S} = \begin{pmatrix} 10 \\ -10 \end{pmatrix}, \qquad \mathbf{P} = \begin{pmatrix} 20 & -10 \\ -10 & 40 \end{pmatrix}, \qquad \mathbf{G} = \begin{pmatrix} 20 & 5 \\ 5 & 10 \end{pmatrix}$$

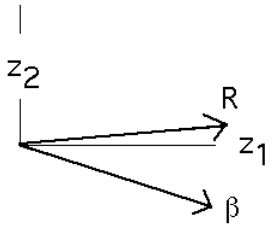
First, assuming no selection on other traits, what is the true nature of selection on these two traits?

$$\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S} = \mathbf{P} = \begin{pmatrix} 20 & -10 \\ -10 & 40 \end{pmatrix}^{-1} \begin{pmatrix} 10 \\ -10 \end{pmatrix} = \begin{pmatrix} 0.43 \\ -0.14 \end{pmatrix}$$

Thus, mean population fitness is maximized by increasing trait one and decreasing trait two. What does the actual response look like?

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \begin{pmatrix} 20 & 5 \\ 5 & 10 \end{pmatrix} \begin{pmatrix} 0.43 \\ -0.14 \end{pmatrix} = \begin{pmatrix} 7.86 \\ 0.71 \end{pmatrix}$$

Thus, the actual response to selection is to increase both traits. The figure below shows the optimal response $\boldsymbol{\beta}$ and the actual response $\mathbf{R}$.



A Short Diversion: The Geometry of a Matrix

As the above example highlights, we can think about the response to selection in terms of geometry, here the difference between the optimal vector and the actual response. Hence, we start with a slight digression on the geometry of matrices and vectors.

Comparing Vectors: Lengths and Angles

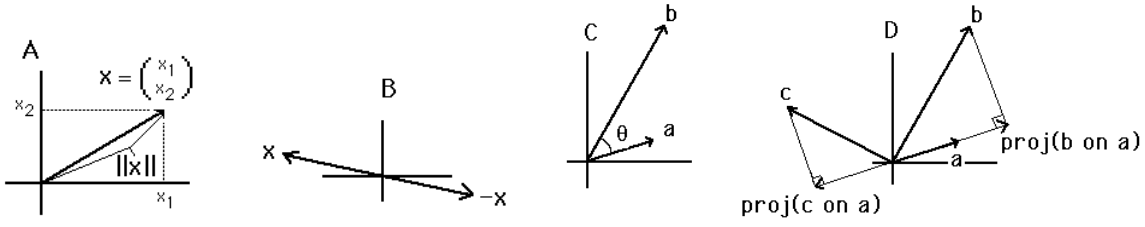


Figure 13.1. Some basic geometric concepts of vectors. While we use examples from two dimensions, these concepts easily extend to n dimensions. **A**: A vector $\mathbf{x}$ can be thought of as an arrow from the origin to a point in space whose coordinates are given by the elements of $\mathbf{x}$. **B**: Multiplying a vector by -1 results in a *reflection* about the origin. **C**: One measure of the difference in direction between two vectors is the angle θ between them. **D**: Proj$(\mathbf{b}$ on $\mathbf{a})$ is the vector resulting from the projection of $\mathbf{b}$ on $\mathbf{a}$. Note that the resulting projection vector is either in the same direction as $\mathbf{a}$ or in the direction of the reflection of $\mathbf{a}$, as seen for Proj$(\mathbf{c}$ on $\mathbf{a})$.

As Figure 13.1 shows, a vector $\mathbf{x}$ can be treated as a geometric object, an arrow leading from the origin to the n dimensional point whose coordinates are given by the elements of $\mathbf{x}$. By changing coordinate systems, we change the resulting vector, potentially changing both its direction (**rotating** the vector) and length (**scaling** the vector). This geometric interpretation suggests several ways for comparing vectors, such as the *angle* between two vectors and the *projection* of one vector onto another.

Consider first the length (or **norm**) of a vector. The most common length measure is the Euclidean distance of the vector from the origin, $||\mathbf{x}||$, which is defined by

$$||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\mathbf{x}^T\mathbf{x}} \tag{13.21a}$$

Hence for any scalar a, $||a\,\mathbf{x}|| = |a|\,||\mathbf{x}||$. If $a < 0$, the vector $a\mathbf{x}$ is scaled by $|a|$ and reflected about the origin as is shown in Figure 13.1. Similarly, the Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$ is

$$||\mathbf{x} - \mathbf{y}||^2 = \sum_{i=1}^{n}(x_i - y_i)^2 = (\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y}) = (\mathbf{y} - \mathbf{x})^T(\mathbf{y} - \mathbf{x}) \tag{13.21b}$$

Vectors can differ by length, direction, or both. The angle θ between two vectors $\mathbf{x}$ and $\mathbf{y}$ provides a measure of how much they differ in direction (Figure 13.1). If the vectors satisfy $a\mathbf{x} = \mathbf{y}$ (where $a > 0$) they point in exactly the same direction, and they are defined to be zero degrees apart. If $a < 0$, they are exactly 180 degrees apart and differ in direction only by a reflection of the axes about the origin. At the other extreme, two vectors can be at right angles to each other ($\theta = 90^\circ$ or 270°), in which case the vectors are said to be **orthogonal**. Orthogonal vectors of unit length are further said to be **orthonormal**. For any two n dimensional vectors, θ satisfies

$$\cos(\theta) = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||\;||\mathbf{y}||} = \frac{\mathbf{y}^T\mathbf{x}}{||\mathbf{x}||\;||\mathbf{y}||} \tag{13.22}$$

Note that since $\cos(90^\circ) = \cos(270^\circ) = 0$, two vectors are orthogonal if and only if their inner product is zero, $\mathbf{x}^T\mathbf{y} = 0$.

Another way to compare vectors, illustrated in Figure 13.1, is to consider the **projection** of one vector on another. For any two n dimensional vectors, the projection of $\mathbf{x}$ on $\mathbf{y}$ generates a vector defined by

$$\text{Proj}(\mathbf{x}\,\text{on}\,\mathbf{y}) = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{y}}\,\mathbf{y} = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{y}||^2}\,\mathbf{y} = \left(\cos(\theta)\,\frac{||\mathbf{x}||}{||\mathbf{y}||}\right)\,\mathbf{y} \tag{13.23a}$$

If $||\mathbf{y}|| = 1$, then

$$\text{Proj}(\mathbf{x}\,\text{on}\,\mathbf{y}) = (\mathbf{x}^T\mathbf{y})\,\mathbf{y} = (\cos(\theta)\,||\mathbf{x}||\,)\,\mathbf{y} \tag{13.23b}$$

Note that since the term involving cosines in Equations 13.23a/b is a scalar, the vector resulting from the projection of $\mathbf{x}$ on $\mathbf{y}$ is in the same direction as $\mathbf{y}$, unless $90^\circ < \theta < 270^\circ$ in which case $\cos(\theta) < 0$ and the projection vector is in exactly the opposite direction (the reflection of $\mathbf{y}$ about the origin). The length of the projection vector is

$$||\text{Proj}(\mathbf{x}\,\text{on}\,\mathbf{y})|| = |\cos(\theta)|\;||\mathbf{x}|| \tag{13.23c}$$

If two vectors lie in exactly the same direction, the projection of one on the other just recovers the original vector (as $\text{Proj}(\mathbf{x}\,\text{on}\,\mathbf{y}) = \mathbf{x}$). Conversely, if two vectors are orthogonal, then the projection of one on the other yields a vector of length zero. An important use of projection vectors is that if $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n$ is any set of mutually orthogonal n dimensional vectors, then any n dimensional vector $\mathbf{x}$ can be represented as the sum of projections of $\mathbf{x}$ onto the members of this set,

$$\mathbf{x} = \sum_{i=1}^{n}\text{Proj}(\mathbf{x}\,\text{on}\,\mathbf{y}_i) \tag{13.24}$$

Matrices Describe Vector Transformations

When we multiply a vector $\mathbf{x}$ by a matrix $\mathbf{A}$ to create a new vector $\mathbf{y} = \mathbf{A}\mathbf{x}$, $\mathbf{A}$ rotates and scales the original vector $\mathbf{x}$ to give $\mathbf{y}$. Thus $\mathbf{A}$ describes a *transformation* of the original coordinate system of $\mathbf{x}$ into a new coordinate system $\mathbf{y}$ (which has different dimensions from $\mathbf{x}$ unless $\mathbf{A}$ is square).

Example 13.3. Suppose

$$\mathbf{G} = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} \qquad \boldsymbol{\beta} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

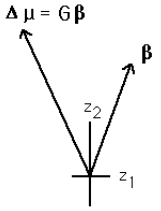
hence

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

The resulting change in character means are different from those most favored by natural selection. Selection favors an increase in z_1, but when the genetic variance-covariance structure is taken into account, the resulting change in z_1 is negative. Taking the appropriate inner products, we find $||\boldsymbol{\beta}|| = \sqrt{10}$, $||\mathbf{R}|| = \sqrt{20}$, and $\boldsymbol{\beta}^T\mathbf{R} = 10$. Applying Equation 13.22,

$$\cos\theta = \frac{\boldsymbol{\beta}^T\mathbf{R}}{||\mathbf{R}||\ ||\boldsymbol{\beta}||} = \frac{1}{\sqrt{2}}$$

Thus the response vector is $\cos^{-1}(1/\sqrt{2}) = 45^\circ$ from the selection gradient, implying that the constraints introduced by the genetic variance-covariance matrix moves the response vector considerably away from the direction most favored by natural selection, as the following figure shows :



Eigenvalues and Eigenvectors

The **eigenvalues** and their associated **eigenvectors** of a square matrix describe the geometry of the transformation induced by that matrix. Eigenvalues describe how the original coordinate axes are scaled in the new coordinate system while eigenvectors describe how the original axes are rotated.

Suppose that the vector $\mathbf{y}$ satisfies the matrix equation

$$\mathbf{A}\mathbf{y} = \lambda\mathbf{y} \tag{13.25}$$

for some scalar value λ. Geometrically, this means that the new vector resulting from transformation of $\mathbf{y}$ by $\mathbf{A}$ points in the same direction (or is exactly reflected about the origin if $\lambda < 0$) as $\mathbf{y}$. For such vectors, the only action of the matrix transformation is to scale them by some amount λ. Hence, it is natural that the system of such vectors along with their corresponding scalar multipliers completely describes the geometry of the transformation associated with $\mathbf{A}$. Vectors satisfying Equation 13.25 are referred to as *eigenvectors* and their associated scaling factors are *eigenvalues*. If $\mathbf{y}$ is an eigenvector, then $a\mathbf{y}$ is also an eigenvector as $\mathbf{A}(a\mathbf{y}) = a(\mathbf{A}\mathbf{y}) = \lambda(a\mathbf{y})$. Note, however, that the associated eigenvalue remains unchanged. Hence, we typically scale eigenvectors to be of unit length to give *unit* or *normalized* eigenvectors. In particular, if $\mathbf{u}_i$ is the eigenvector associated with the ith eigenvalue, then the associated normalized eigenvector is $\mathbf{e}_i = \mathbf{u}_i/||\mathbf{u}_i||$.

The eigenvalues of square matrix $\mathbf{A}$ of dimension n are solutions of Equation 13.25, which is usually expressed as the **characteristic equation** $|\mathbf{A} - \lambda\mathbf{I}| = 0$. This can be also be expressed using the **Laplace expansion** as

$$|\mathbf{A} - \lambda\mathbf{I}| = (-\lambda)^n + S_1(-\lambda)^{n-1} + \cdots + S_{n-1}(-\lambda)^1 + S_n = 0 \tag{13.26}$$

where S_i is the sum of all **principal minors** (minors including diagonal elements of the original matrix) of order i. Minors are defined in Chapter 8 of Lynch and Walsh. Finding the eigenvalues thus requires solving a polynominal equation of order n. In practice, for $n > 2$ this is usually done numerically, and most statistical and numerical analysis packages offer routines to accomplish this task.

Two of these principal minors are easily obtained and provide some information on the nature of the eigenvalues. The only principal minor having the same order of the matrix is the full matrix itself, so that $S_n = |\,\mathbf{A}\,|$, the determinant of $\mathbf{A}$. S_1 is also related to an important matrix quantity, the **trace**. This is denoted by $\mathrm{tr}(\mathbf{A})$ and is the sum of the diagonal elements of the matrix,

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$$

Observe that $S_1 = \mathrm{tr}(\mathbf{A})$ as the only principal minors of order one are the diagonal elements themselves, the sum of which equals the trace. The trace and determinant can be expressed as functions of the eigenvalues,

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i \qquad \text{and} \qquad |\mathbf{A}| = \prod_{i=1}^{n} \lambda_i \tag{13.27}$$

Hence $\mathbf{A}$ is singular ($|\mathbf{A}| = 0$) if and only if at least one eigenvalue is zero.

Let $\mathbf{e}_i$ be the (unit-length) eigenvector associated with eigenvalue λ_i. If the eigenvectors of $\mathbf{A}$ can be chosen to be mutually orthogonal, e.g., $\mathbf{e}_i^T\mathbf{e}_j = 0$ for $i \neq j$, we can express $\mathbf{A}$ as

$$\mathbf{A} = \lambda_1\mathbf{e}_1\mathbf{e}_1^T + \lambda_2\mathbf{e}_2\mathbf{e}_2^T + \cdots + \lambda_n\mathbf{e}_n\mathbf{e}_n^T \tag{13.28a}$$

This decomposition is called the **spectral decomposition** of $\mathbf{A}$. Hence,

$$\begin{aligned}\mathbf{A}\mathbf{x} &= \lambda_1\mathbf{e}_1\mathbf{e}_1^T x + \lambda_2\mathbf{e}_2\mathbf{e}_2^T x + \cdots + \lambda_n\mathbf{e}_n\mathbf{e}_n^T x \\ &= \lambda_1\mathrm{Proj}(\mathbf{x}\,\mathrm{on}\,\mathbf{e}_1) + \lambda_2\mathrm{Proj}(\mathbf{x}\,\mathrm{on}\,\mathbf{e}_2) + \cdots + \lambda_n\mathrm{Proj}(\mathbf{x}\,\mathrm{on}\,\mathbf{e}_n)\end{aligned} \tag{13.28b}$$

The last equality follows since $\mathbf{e}_i(\mathbf{e}_i^T\mathbf{x}) = (\mathbf{e}_i^T\mathbf{x})\mathbf{e}_i$ as ${\mathbf{e}_i}^T\mathbf{x}$ is a scalar.

Example 13.4. Determine the **eigenstructure** (the set of eigenvalues and their associated unit eigenvectors) of the genetic variance-covariance matrix $\mathbf{G}$ given in Example 13.3. Writing the characteristic equation,

$$\begin{aligned}|\mathbf{G} - \lambda\mathbf{I}| &= \left|\begin{pmatrix} 4-\lambda & -2 \\ -2 & 2-\lambda \end{pmatrix}\right| \\ &= (4-\lambda)(2-\lambda) - (-2)^2 = \lambda^2 - 6\lambda + 4 = 0\end{aligned}$$

Alternatively, using the Laplace expansion, $\mathrm{tr}(\mathbf{G}) = 4 + 2 = 6$ and $|\mathbf{G}| = 4 * 2 - (-2)^2 = 4$ also recovers the characteristic equation, which has solutions

$$\lambda_1 = 3 + \sqrt{5} \simeq 5.236 \qquad \lambda_2 = 3 - \sqrt{5} \simeq 0.764$$

The associated unit eigenvectors are

$$\mathbf{e}_1 \simeq \begin{pmatrix} -0.851 \\ 0.526 \end{pmatrix} \qquad \mathbf{e}_2 \simeq \begin{pmatrix} 0.526 \\ 0.851 \end{pmatrix}$$

These are orthogonal as ${\mathbf{e}_1}^T\mathbf{e}_2 = 0$. From Example 13.2, $||\boldsymbol{\beta}|| = \sqrt{10}$, while ${\mathbf{e}_1}^T\boldsymbol{\beta} \simeq 0.727$ and ${\mathbf{e}_2}^T\boldsymbol{\beta} \simeq 3.079$. Since $||\mathbf{e}_1|| = ||\mathbf{e}_2|| = 1$,

$$\cos(\theta|\mathbf{e}_1, \boldsymbol{\beta}) \simeq \frac{0.727}{\sqrt{10}} \simeq 0.201 \qquad \text{and} \qquad \cos(\theta|\mathbf{e}_2, \boldsymbol{\beta}) \simeq \frac{3.079}{\sqrt{10}} \simeq 0.974$$

giving the angle between $\mathbf{e}_1$ and $\boldsymbol{\beta}$ as $\theta(\mathbf{e}_1, \boldsymbol{\beta}) \simeq 78.4°$, while $\theta(\mathbf{e}_2, \boldsymbol{\beta}) \simeq 13.2°$. The corresponding scaled projections of $\boldsymbol{\beta}$ on these eigenvectors are

$$\lambda_1 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_1) \simeq \begin{pmatrix} -3.236 \\ 2 \end{pmatrix} \qquad \text{and} \qquad \lambda_2 \text{Proj}(\boldsymbol{\beta} \text{ on } \mathbf{e}_2) \simeq \begin{pmatrix} 1.236 \\ 2 \end{pmatrix}$$

From Equation 13.28b, $\boldsymbol{\beta}$ is the sum of these two projections. As Figure 13.2 shows, the eigenstructure of $\mathbf{G}$ explains the unusual behavior of response seen in Example 13.3. The eigenvector associated with the **leading eigenvalue** λ_1 accounts for most of the variation inherent in $\mathbf{G}$, and this eigenvector corresponds to a strong negative correlation between the additive genetic values of z_1 and z_2. 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 this projection vector that results in the decrease in μ_{z_1}.

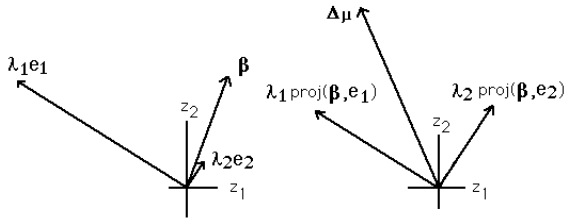


Figure 13.2. **Left**: The scaled eigenvectors associated with the variance-covariance matrix $\mathbf{G}$ from Example 13.2. Note that $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal and hence can be thought of as describing a new coordinate system. **Right**: Since $\lambda_1 >> \lambda_2$, the leading eigenvector largely dominates the transformation. This is shown by taking the projections of $\boldsymbol{\beta}$ on each of these eigenvectors. Even though $\boldsymbol{\beta}$ is nearly orthogonal to $\mathbf{e}_2$, the projection of $\boldsymbol{\beta}$ on $\mathbf{e}_1$ yields a vector of greater length than the projection of $\boldsymbol{\beta}$ on $\mathbf{e}_2$. From Equation 13.28b, the response to selection $\mathbf{R}$ is the sum of these two projections.

Quantifying Multivariate Constraints to Response

Is There Genetic Variation in the Multivariate Direction of Selection?

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. To see this point, consider the following example.

Example 13.5. Suppose the $\mathbf{G}$ matrix is:

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

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, with corresponding eigenvector that exactly corresponds to our $\boldsymbol{\beta}$. Hence, there is NO additive genetic variance along this particular direction, and hence no response. Likewise, if our $\boldsymbol{\beta}$ was only a few degress away from the eigenvector, the resulting response would be very small.

While Example 13.5 is hypothetical, a real-world example is offered from the work of Blows.

Example 13.6. 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:

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

Let's look at 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 13.22,

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

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.

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

The notion of multivariate constraints to response, even when all of the component traits show additive variation, traces back to Dickerson (1955). It is somewhat surprising then that it took over forty years for the first attempts to describe the potential geometry of these constraints. Schluter (1996) defined the **genetic line of least resistance** $\mathbf{g}_{max}$ as the eigenvector corresponding to the lead (largest) eigenvalue of $\mathbf{G}$ (this is just the **first principal component** of $\mathbf{G}$). 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

$$\mathbf{d}' = \frac{\mathbf{d}}{\sqrt{\mathbf{d}^T\mathbf{d}}} \tag{13.29}$$

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

$$\theta = \cos^{-1}(\mathbf{g}_{max}^T\mathbf{d}') \tag{13.30}$$

Schluter observed the following in his (small) data set

(i). The smallest values of θ (departures from the direction of maximal genetic variation) occurred between the most recently diverged species.

(ii). The greater the value of θ (angle between the maximal genetic variation and direction of actual divergence), the smaller the total amount of divergence.

(iii). 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 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).

McGuigan et al. (2005) offered a very 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 from $\mathbf{d}$ and $\mathbf{g}_{max}$). However, hydrodynamic divergence between lake vs. stream populations in each species were along directions that we 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.

Blow's Matrix Subspace Projection

Schluter's approach only deals with the leading eigenvector of $\mathbf{G}$. If the leading eigenvector dominates all of the others (and hence accounts for most of the variance), then this makes sense. (Recall that the fraction of variation accounted from along the direction of the k-th eigenvalue is just $\lambda_k/\sum\lambda_i = \lambda_k/\text{trace}(\mathbf{G})$). 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. Blows et al. (2004) suggested a matrix subspace projection approach to consider more of the eigenvectors of $\mathbf{G}$.

A common problem is that the $\mathbf{G}$ matrix is **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 constrainsts 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 eariler). 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

$$\mathbf{A} = (\, \mathbf{e}_1, \quad \mathbf{e}_2, \quad \cdots, \mathbf{e}_k \,) \tag{13.31}$$

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

$$\mathbf{P}_{roj} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \tag{13.32a}$$

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

$$\mathbf{p} = \mathbf{P}_{roj}\boldsymbol{\beta} = \mathbf{A} \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \boldsymbol{\beta} \tag{13.32b}$$

Example 13.7. 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 13.32a gives an 8×8 projection matrix (not show here), and Equation 13.32b 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.

General Theory of Selection on a Linear Index

Our final topic is selection on an **index** of trait values. This is typically how artifical selection on multiple traits occurs, in both laboratory experiments and in animal and plant breeding.

Consider selection on the univariate character defined by the linear index $I = \sum b_j z_j = \mathbf{b}^T\mathbf{z}$ where $\mathbf{z}$ is the vector of phenotypic values in an individual and $\mathbf{b}$ a vector of weights. As before, let $\mathbf{P}$ and $\mathbf{G}$ denote the phenotypic and additive-genetic covariance matrices. Applying Equation 2.20, the phenotypic variance of this new synthetic character I is

$$\sigma_I^2 = \sigma(\mathbf{b}^T\mathbf{z}, \mathbf{b}^T\mathbf{z}) = \mathbf{b}^T\sigma(\mathbf{z},\mathbf{z})\mathbf{b} = \mathbf{b}^T\mathbf{P}\mathbf{b} \tag{13.33a}$$

Likewise, the additive genetic variance of I becomes

$$\sigma_{A_I}^2 = \sigma_A(\mathbf{b}^T\mathbf{z}, \mathbf{b}^T\mathbf{z}) = \mathbf{b}^T\sigma_A(\mathbf{z},\mathbf{z})\mathbf{b} = \mathbf{b}^T\mathbf{G}\mathbf{b} \tag{13.33b}$$

giving the heritability of I as

$$h_I^2 = \frac{\sigma_{A_I}^2}{\sigma_I^2} = \frac{\mathbf{b}^T\mathbf{G}\mathbf{b}}{\mathbf{b}^T\mathbf{P}\mathbf{b}} \tag{13.33c}$$

as obtained by Lin and Allaire (1977) and Nordskog (1978). If phenotypes $\mathbf{z}$ and breeding values $\mathbf{g}$ are jointly multivariate normal, linear combinations of each is also normally distributed and hence the univariate breeders' equation holds for response in I. If i is the selection intensity on I, the response in I is

$$R_I = \overline{i}\, h_I^2\, \sigma_I = \overline{i} \cdot \frac{\mathbf{b}^T\mathbf{G}\mathbf{b}}{\mathbf{b}^T\mathbf{P}\mathbf{b}}\sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}} = \overline{i} \cdot \frac{\mathbf{b}^T\mathbf{G}\mathbf{b}}{\sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}}} \tag{13.34}$$

At the risk of mentioning the obvious, we note that the index I is a standard univariate character and hence all univariate results from previous lectures apply to selection on the trait I. For example, the realized heritability of the index can be obtained by the methods discussed in Lecture 10.

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

$$\frac{\sigma(z_j,w)}{\sigma_w\,\sigma_{z_j}} = \left(\frac{\sigma(z_j,I)}{\sigma_I\,\sigma_{z_j}}\right)\left(\frac{\sigma(I,w)}{\sigma_I\,\sigma_w}\right) \tag{13.35a}$$

Recalling the Price-Robertson Equation (12.8), $\sigma(z_j,w) = S_j$ and likewise $\sigma(I,w) = S_I = i\,\sigma_I$ where i is the selection intensity on the index. Solving for $\sigma(z_j,w)$ and using these identities gives

$$S_j = \sigma(z_j,w) = \frac{\sigma(z_j,I)\cdot\sigma(I,w)}{\sigma_I^2} = i \cdot \frac{\sigma(z_j,I)}{\sigma_I} \tag{13.35b}$$

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

$$S_j = \left(\frac{i}{\sigma_I}\right)\cdot\sum_k b_k P_{jk} \tag{13.36a}$$

giving the vector of selection differential as

$$\mathbf{S} = \left(\frac{i}{\sigma_I}\right)\cdot\mathbf{P}\mathbf{b} \tag{13.36b}$$

and the resulting vector of responses as

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \left(\frac{i}{\sigma_I}\right) \cdot \mathbf{G}\mathbf{b} = i \cdot \frac{\mathbf{G}\mathbf{b}}{\sqrt{\mathbf{b}^T\mathbf{P}\mathbf{b}}} \tag{13.36c}$$

Equation 13.36c shows that the vector of responses $\mathbf{R}$ is unchanged if the index weights are rescaled from $\mathbf{b}$ to $c \cdot \mathbf{b}$ as the constant c cancels out. However the response in the index itself changes as $\mathbf{b}$ is rescaled. From Equation 13.34 the response in the index using weights $c \cdot \mathbf{b}$ is c times the response expected when the index uses $\mathbf{b}$.

While indices have been presented as the objects of selection, often an investigator observes a vector of selection differentials $\mathbf{S}$ or vector of responses $\mathbf{R}$ and wishes to obtain the linear index that would give the same $\mathbf{S}$ and/or $\mathbf{R}$. This approach of constructing a **retrospective index** (or *index in retrospect*) was first suggested by Dickerson et al. (1954). If the vector of selection differentials $\mathbf{S}$ is observed, Equation 13.36a suggests the retrospective index $\mathbf{b} = \mathbf{P}^{-1}\mathbf{S}$ (given our previous remarks we ignore the constant i/σ_I). Even when artificial selection occurs using a known index, a retrospective index constructed from the *effective* selection differentials (which measure fertility differences in addition to artificial selection, see Problem 1 from Lecture 9) provides the investigator with a measure of how natural selection interferes with the desired artificial selection scheme. Alternatively, the investigator may not know the within-generation change in $\mathbf{S}$ but can observe the between-generation change $\mathbf{R}$. In this case Equation 13.36c (again ignoring constants) suggests the retrospective index $\mathbf{b} = \mathbf{G}^{-1}\mathbf{R}$.

The Smith-Hazel Index

The standard goal of multiple character selection is to maximize the response in some overall merit based on an index of phenotypic values. Typically, the merit function is taken to be a linear index $J = \mathbf{a}^T\mathbf{z}$ where the vector of *economic weights* $\mathbf{a}$ assigns the desirability of relative responses in each character. What index weights $\mathbf{b}$ should we use so that the index $I = \mathbf{b}^T\mathbf{z}$ maximizes the response in the index J? The solution, due to Smith (1936) and Hazel (1945), gives the **Smith-Hazel selection index**

$$I_s = \mathbf{b}_s^T\mathbf{z} = \left(\mathbf{P}^{-1}\mathbf{G}\mathbf{a}\right)^T\mathbf{z} \qquad \text{where} \qquad \mathbf{b}_s = \mathbf{P}^{-1}\mathbf{G}\mathbf{a} \tag{13.37a}$$

Noting that $(\mathbf{ABC})^T = \mathbf{C}^T\mathbf{B}^T\mathbf{A}^T$ and that both $\mathbf{P}$ and $\mathbf{G}$, being symmetric matrices, equal their transposes, it follows that

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

To obtain the response, note that the covariance in breeding values between the two indices is given by

$$\sigma_A(I, J) = \sigma_A(\mathbf{a}^T\mathbf{z}, \mathbf{b}^T\mathbf{z}) = \mathbf{a}^T\sigma_A(\mathbf{z}, \mathbf{z})\mathbf{b} = \mathbf{a}^T\mathbf{G}\mathbf{b}$$

Along the lines of Equation 13.34, the response is the selection intensity times the additive genetic covariance between the selected and response index divided by σ_I^2, giving

$$R_H = \overline{i} \cdot \frac{\mathbf{a}^T\mathbf{G}\mathbf{b}_s}{\sqrt{\mathbf{b}_s^T\mathbf{P}\mathbf{b}_s}} = \overline{i} \cdot \frac{\mathbf{a}^T\mathbf{G}\mathbf{b}_s}{\sqrt{\mathbf{a}^T\mathbf{G}\mathbf{b}_s}} = \overline{i} \cdot \sqrt{\mathbf{a}^T\mathbf{G}\mathbf{b}_s} = \overline{i} \cdot \sqrt{\mathbf{a}^T\mathbf{G}\mathbf{P}^{-1}\mathbf{G}\mathbf{a}} \tag{13.37b}$$

MAS – Marker Assisted Selection

As will be discussed extensively in Lectures 16 and 17, the recent ability to obtain polymorphic genetic markers spanning the genome at almost arbitrarily small map distances has profound implications for quantitative genetics. Markers linked to QTLs (quantitative trait loci) can provide

information on the genotype, and hence the breeding value, of an individual. Thus an individual's marker genotype can be viewed as a character (or set of characters) that can be used to improve selection response. While one might think that investigation of such **marker-assisted selection** (or **MAS**) schemes is recent area of research, they were first considered by Neimann-Sørensen and Robertson (1961) and Smith (1967). These authors developed selection indices for a single locus that directly influences the character of interest. Our treatment follows the more general approach Lande and Thompson (1990), who consider a selection index that allows for multiple marker loci linked to QTLs. Related approaches based on BLUP (best linear unbiased predictors) have been developed (Fernando and Grossman 1989, Fernando 1990, Cantet and Smith 1991, Goddard 1992), wherein marker information is used to construct the BLUP estimate for each individual's breeding value. Selection occurs by choosing those individuals with highest estimated values breeding values. These BLUP approaches are computationally very intense.

Assume a large number of markers have been scored and that at least some of these are in gametic-phase disequilibrium with QTLs underlying the trait of interest and hence provide some information about an individual's breeding value. How is this marker information translated into usable characters for index selection? To motivate the approach, consider the simplest case of a single diallelic locus (alleles Q and q) that is a QTL for the trait of interest. The contribution of this locus to the additive genetic value is $a+m\alpha$ for an individual with $m = 0, 1, 2$ copies of Q. The slope α of this regression of additive genetic value on number of copies of Q is the difference in average effects between alleles Q and q (Lecture 3). This regression predicting additive genetic value as a function of the number of copies of a particular allele also holds more generally for a marker loci in gametic phase disequilibrium with a QTL (as opposed to the locus being the QTL itself). In this case, α decays to zero as the gametic phase disequilibrium between marker and QTL decays, so that a marker provides no information in the absence of disequilibrium with a QTL. Building on this single-locus regression, suppose now that a large number of diallelic marker loci are scored. We restrict analysis to each linkage group (operationally, each chromosome) under the assumption that disequilibrium between unlinked groups is very small. Suppose the linkage group being considered has n diallelic marker loci, the ith of which has alleles B_i and b_i. Letting $m_i = 0, 1, 2$ be the variable indicating the number of copies of allele B_i in the individual being examined, the average effects associated with the n markers are given by the slope coefficients of the least-squares regression of additive genetic value on the vector of marker genotypes $\mathbf{m}^T = (m_1, \cdots, m_n)$,

$$g = a + \sum_{i=1}^{n} \alpha_i \cdot m_i = a + \boldsymbol{\alpha}^T \mathbf{m} \tag{13.38a}$$

Multiple regression is used instead of n separate univariate regressions for each marker loci because the genotypes of marker loci may be correlated with each other (marker loci in gametic phase disequilibrium) and/or a QTL may influence several markers at once (again generating correlations between loci). Multiple regression accounts for these correlated effects, with the resulting regression coefficients being the contribution from that locus with all others held constant (Lecture 2). Hence, $\alpha_i \cdot m_i$ is the estimated contribution towards additive genetic value from the ith marker locus and we define m, the **marker score** of an individual, as the sum of the additive effects associated with these marker loci,

$$m = \sum_{i=1}^{n} \alpha_i \cdot m_i = \boldsymbol{\alpha}^T \mathbf{m} \tag{13.38b}$$

Here m_i = 0, 1, or 2 is the number of copies of the marker allele at locus i. Marker score is the best linear predictor of additive genetic value given the vector of marker genotypes. Assuming no genotype-environment interactions, we use the observed phenotypic value z in place of the unobserved additive genetic value giving the estimate of $\boldsymbol{\alpha}$ from standard regression theory (Lecture 2) as

$$\widehat{\boldsymbol{\alpha}} = \mathbf{S}_{\mathbf{m}}^{-1} \boldsymbol{\sigma}(z, \mathbf{m}) \tag{13.39}$$

where $\mathbf{S_m}$ is the estimated covariance matrix for $\mathbf{m}$ (with $S_{m_{ij}} = \sigma(m_i, m_j)$) and $\boldsymbol{\sigma}(z, \mathbf{m})$ is the vector of estimated covariances between phenotype value and each single-locus marker genotype ($\boldsymbol{\sigma}(z, \mathbf{m})_i = \sigma(z, m_i)$). The number of scored individuals must exceed the number of marker loci in order to uniquely estimate the marker effects for each loci. One practical point about constructing this regression is that only some of the markers are likely to be useful so that only a subset of the initially scored markers usually need to be followed. Care, however, is required in choosing the appropriate set of markers. The standard procedure is to choose markers by a stepwise multiple regression. Here the marker accounting for the most variation is used as the first variable in the regression. Next the marker accounting for the largest fraction of the remaining variance is examined and this marker is added to the regression if it explains a significant fraction of this variation. This procedure is repeated until the addition of a new marker does not account for a significant fraction of the remaining variance. The problem with this approach is that it produces biased overestimates of α_i. Lande and Thompson suggest a two-step approach where first a set of markers is chosen using stepwise regression on data from a previous generation and then a new regression computed using these chosen markers on the data from the current generation. The initial regression from the previous generation picks those markers that are potentially useful while the second regression using these markers on new data produces unbiased estimates.

Lande and Thompson show in many cases that the univariate marker score m is as informative as the entire vector of genotypic values, so that all marker information can be collapsed into the single variable m. Taking the vector of characters as $\mathbf{z} = (z, m)^T$, the vector of associated economic weights is $\mathbf{a} = (1, 0)^T$ as our concern is only response in the phenotypic value z. Since our concern here is to maximize the breeding value, it can be shown that the vector of optimal weights in the Smth-Hazel index are

$$\mathbf{b}_s = \mathbf{P}^{-1}\mathbf{g}$$

where $\mathbf{g}$ is vector of additive genetic covariances associated with z. Let $\sigma^2_m = \rho\sigma^2_A$ denote the variance in marker score, which accounts for ρ of the total variance in additive genetic value. Since m is a subset of genotypic values, $\sigma(m, g) = \sigma^2_m$, giving

$$\mathbf{g} = \begin{pmatrix} \sigma^2_A \\ \\ \sigma^2_m \end{pmatrix} = \sigma^2_A \cdot \begin{pmatrix} 1 \\ \\ \rho \end{pmatrix} \tag{13.40a}$$

Since $\sigma(z, m) = \sigma(g + e, m) = \sigma(g, m) = \sigma^2_m = \rho\sigma^2_A = \rho h^2 \sigma^2_z$, the associated phenotypic covariance matrix becomes

$$\mathbf{P} = \begin{pmatrix} \sigma^2_z & \sigma^2_m \\ \\ \sigma^2_m & \sigma^2_m \end{pmatrix} = \sigma^2_z \cdot \begin{pmatrix} 1 & h^2\,\rho \\ \\ h^2\,\rho & h^2\,\rho \end{pmatrix} \tag{13.40b}$$

The Smith-Hazel index becomes

$$I_s = \mathbf{b}_s^T\mathbf{z} = \mathbf{g}^T\mathbf{P}^{-1}\mathbf{z} = h^2\left(\frac{1-\rho}{1-h^2\,\rho}\right)\cdot z + \left(\frac{1-h^2}{1-h^2\,\rho}\right)\cdot m \tag{13.41a}$$

which can be rescaled as

$$I_s = z + \left(\frac{1-h^2}{h^2(1-\rho)}\right)\cdot m \tag{13.41b}$$

Recall from Equation 13.37b that the response to selection is a function of $\mathbf{a}^T\mathbf{G}\mathbf{b}_s$. Noting that $\mathbf{G}\mathbf{a} = \mathbf{g}$, it follows that $\mathbf{a}^T\mathbf{G}\mathbf{b}_s = \mathbf{g}^T\mathbf{b}_s = \mathbf{g}^T\mathbf{P}^{-1}\mathbf{g}$. From 13.40a and b, this becomes

$$\mathbf{g}^T\mathbf{P}^{-1}\mathbf{g} = \sigma^2_A\,h^2\;\frac{1+\rho/h^2-2\rho}{1-h^2\rho} = \sigma^2_A\,h^2\cdot\left(1+\frac{(1-h^2)^2\,\rho}{h^2(1-h^2\rho)}\right) \tag{13.42}$$

Thus the response under marker assisted selection is

$$\frac{R}{i} = \sqrt{\mathbf{a}^T \mathbf{G} \mathbf{b}_s} = \sqrt{\mathbf{g}^T \mathbf{P}^{-1} \mathbf{g}} = \sigma_A \, h \cdot \sqrt{1 + \frac{(1-h^2)^2 \, \rho}{h^2 (1 - h^2 \rho)}} \tag{13.43}$$

This was obtained by Smith (1967) for a single marker locus and in the above more general form by Lande and Thompson (1990).

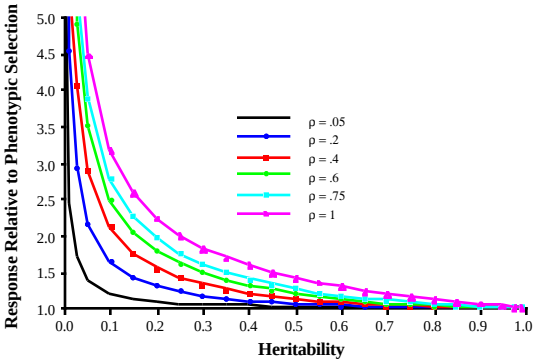


Figure 13.3. Ratio of the expected (single-generation) response of the optimal index for marker-assisted selection relative to selection solely on phenotypic value z.

Figure 13.3 plots the response as a function of h^2 and ρ. As expected, the response using marker information is more efficient that simple phenotypic selection (which has response $R/i = \sigma_g \, h$). For a given heritability, as ρ approaches one, the increase in response under MAS approaches it maximum value of $1/h$, so that the increase under MAS can be considerable if heritability is small. Conversely, as heritability increases towards one the response under MAS is not significantly different from simple phenotypic selection.

Indirect Selection on Marker Score: Applications to Sex-limited Traits

Marker score itself can be used as the sole basis of selection. The expected response in z to selection on m is

$$\frac{R_z}{i} = \frac{\sigma_{m,z}}{\sigma_m} = \sigma_m^2 / \sigma_m = \sigma_m = \sigma_g \sqrt{\rho} \tag{13.44a}$$

giving an efficiency relative to selection on z of

$$\frac{\sigma_m}{h \cdot \sigma_g} = \sqrt{\frac{\rho}{h^2}} \tag{13.44b}$$

Thus as expected, marker selection gives a larger response than phenotypic selection when marker score accounts for more of the additive genetic variation that phenotypic value ($p > h^2$).

Marker-assisted selection can be used to enhance the response of a sex-limited character. Marker loci, by providing information on breeding value, can be used to select on the sex not expressing the character. Assume the sex-limited character is only expressed in females. Lande and Thompson (1990) consider the situation where the optimal index of phenotypic value and marker score is used

to select females while marker scores are used to select males. Combining Equations 13.43 and 13.44a, the expected response is

$$R = h\,\sigma_g\left(i_f \cdot \sqrt{1 + \frac{\rho\,(1-h^2)^2}{h^2\,(1-h^2\rho)}} \;+\; i_m \cdot \sqrt{\frac{\rho}{h^2}}\right) \tag{13.45a}$$

where i_f and i_m is the selection intensity on females and male respectively. The expected response when selection occurs only on character value in females is $R = h\,\sigma_g\,i_f$, showing that MAS increases response by

$$\sqrt{1 + \frac{\rho\,(1-h^2)^2}{h^2\,(1-h^2\rho)}} \;+\; \frac{i_m}{i_f} \cdot \sqrt{\frac{\rho}{h^2}} \tag{13.45b}$$

Lecture 13 Problems

1. Consider selection acting on two traits (1 and 2). The phenotypic variances and covariances are $\sigma^2(P_1) = \sigma^2(P_2) = 10, \sigma(P_1, P_2) = -5$, while the additive genetic variances and covariances are $\sigma^2(A_1) = 4, \sigma^2(A_2) = 9, \sigma(A_1, A_2) = 3$. Compute the response on both characters when:

a: We select directly on trait 2, with $S_2 = 10$

b: We select directly on trait 1, with $S_1 = 10$

c: $S_1 = 5, S_2 = 5$ (use the multivariate breeders equation)

2. The following covariances of performance in randomly chosen dam and daughter pairs of dairy cattle were obtained from an analysis. Estimate the heritabilities of milk yield and fat % and the genetic, phenotypic, and environmental correlations between them.

	Dam's yield	Dam's fat %
Dam's yield (in 100 kg units)	68	-0.55
Dam's fat %	-0.55	0.11
Daughter's yield	7.8	-.20
Daughter's fat	-0.18	0.035

3. The heritability of growth rate in pine seedlings is 0.5 when they are grown in the greenhouse, and 0.2 when grown in the field. The genetic correlation between growth rate in the two environments is 0.8. Suppose you wish to select for increased growth in the field. In which environment should you do the selection? How would your conclusions change if the genetic correlations between growth rate in the two environments was 0.5?

4. Consider the following values for $\mathbf{G}$ and $\boldsymbol{\beta}$

$$\mathbf{G} = \begin{pmatrix} 100 & -50 & -50 \\ -50 & 55 & 40 \\ -50 & 40 & 33 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

a: For this $\mathbf{G}$ matrix, what is $\mathbf{g}_{max}$? What fraction of the total additive variance does this direction account for?

b: What is the angle θ between $\boldsymbol{\beta}$ and $\mathbf{g}_{max}$?

c: Consider the subspace of $\mathbf{G}$ consisting of only the two largest eigenvalues. What fraction of additive variance is accounted for by the subspace? What is the matrix $\mathbf{A}$?

d: What is the vector for the projection of $\boldsymbol{\beta}$ into this subspace?

e: What is the angle θ between this projection and $\boldsymbol{\beta}$?

Solutions Lecture 13 Problems

1. Here

$$h_X = \sqrt{\frac{4}{10}} = 0.63, \quad h_Y = \sqrt{\frac{9}{10}} = 0.95, \quad r_A = \frac{3}{\sqrt{4 \cdot 9}} = 0.5$$

a) $S_2 = 10$ implies $i = 10/\sqrt{10} = 3.16$

$$R_2 = i\, h_2^2\, \sigma_{P_w} = 3.16 \cdot 0.9 \cdot \sqrt{10} = 8.99$$
$$CR_1 = r_A\, \sigma_{A_1}\, i\, h_2 = 0.5 \cdot \sqrt{4} \cdot 3.16 \cdot 0.95 = 3.00$$

b) $S_1 = 10$ implies $i = 10/\sqrt{10} = 3.16$

$$R_1 = i\, h_1^2\, \sigma_{P_1} = 3.16 \cdot 0.4 \cdot \sqrt{10} = 4.00$$
$$CR_2 = r_A\, \sigma_{A_2}\, i\, h_1 = 0.5 \cdot \sqrt{9} \cdot 3.16 \cdot 0.63 = 3.58$$

c) $S_1 = S_2 = 5$. Using the multivariate breeder's equation,

$$\mathbf{S} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}, \quad \mathbf{P} = \begin{pmatrix} 10 & -5 \\ -5 & 10 \end{pmatrix},$$

First note that

$$\mathbf{P}^{-1} = \frac{1}{15}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \text{and} \quad \mathbf{P}^{-1}\mathbf{S} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Hence

$$\mathbf{R} = \mathbf{G}\mathbf{P}^{-1}\mathbf{S} = \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 12 \end{pmatrix}$$

2. Let character 1 denote yield, 2 denote fat %. First, using the covariance of a trait with itself to estimate the phenotypic variances, we have

$$\sigma^2_{P_1} = 68, \quad \sigma^2_{P_2} = 0.11$$

Second, the additive genetic variance in the trait is twice the covariance (for the same character) between parent and offspring, giving

$$\sigma^2_{A_1} = 2 \cdot 7.8 = 15.6, \quad \sigma^2_{A_2} = 2 \cdot 0.035 = 0.07$$

The phenotypic covariance is the covariance between characters 1 and 2 in the parent,

$$\sigma(P_1, P_2) = -0.55$$

Finally, the additive genetic covariance is twice the covariance for trait 1 in the parent and trait 2 in the offspring. Likewise trait 2 in parent and 1 in offspring also estimates this covariance, so we take the average of these two,

$$\sigma(A_1, A_2) = 2(1/2)(-0.20 - 0.18) = -0.38$$

Thus

$$h_1^2 = \frac{15.6}{68} = 0.23, \quad \text{and} \quad h_2^2 = \frac{0.07}{0.11} = 0.64$$

$$r_p = \frac{-0.55}{\sqrt{68 \cdot 0.11}} = -0.20$$

$$r_A = \frac{-0.38}{\sqrt{15.6 \cdot 0.07}} = -0.36$$

Finally, since $V_E = V_P - V_A$, and $Cov(P_1, P_2) = Cov(A_1 + A_2) + Cov(E_1, E_2)$

$V_{E_1} = 68 - 15.6 = 52.4, \quad V_{E_2} = 0.11 - 0.07 = 0.4, \quad Cov(E_1, E_2) = -0.55 - (-0.38) = -0.17$

giving

$$r_e = \frac{-0.17}{\sqrt{52.4 \cdot 0.4}} = -0.04$$

3. Compare the ratio of the direct and correlated response. Let X denote the trait value in the field and Y the value in the lab. Assuming the same amount of selection in either setting, the ratio of the correlated field response CR_X (based on selection in the lab) to the direct response R_X if selection is in the field is

$$\frac{CR_X}{R_X} = \frac{i_Y r_A \sigma_{A_X} h_Y}{i_X h_X \sigma_{A_X}} = \frac{r_A h_Y}{h_X} = \frac{0.8\sqrt{0.5}}{\sqrt{0.2}} = 1.265$$

So that a larger response in the field is given by selecting in the lab than by selecting directly in the field.

If the correlation between environments is 0.5, then

$$\frac{CR_X}{R_X} == \frac{0.5\sqrt{0.5}}{\sqrt{0.2}} = 0.791,$$

in which case direct selection in the field is more efficient.

4. Use the **eigen** function **R** here.

a: For this $\mathbf{G}$ matrix, what is $\mathbf{g}_{max}$? What fraction of the total additive variance does this direction account for?

$$\mathbf{g}_{max} = \mathbf{e}_1 = \begin{pmatrix} -0.743 \\ 0.505 \\ 0.439 \end{pmatrix}$$

$\lambda_1/\text{tr}(\mathbf{G}) = 163.50/188 = 0.870$. Hence, this direction accounts for 87% of the total variance.

b: What is the angle θ between $\boldsymbol{\beta}$ and $\mathbf{g}_{max}$?

$$\theta = \cos^{-1}\left(\frac{\mathbf{g}_{max}^T \boldsymbol{\beta}}{\sqrt{\boldsymbol{\beta}^T \boldsymbol{\beta}}}\right) = \cos^{-1}(-0.208) = 77.8 \text{ degrees}$$

c: What fraction of additive variance is accounted for by the subspace? What is the matrix $\mathbf{A}$?

Total variation = (163.5 + 24.1)/188 = 0.998, or 99.8%. From $\mathbf{R}$, $\mathbf{e}_2$ is

$$\mathbf{e}_2 = \begin{pmatrix} 0.633 \\ 0.743 \\ 0.218 \end{pmatrix}, \quad \text{implying} \quad \mathbf{A} = \begin{pmatrix} -0.743 & 0.633 \\ 0.505 & 0.743 \\ 0.439 & 0.218 \end{pmatrix}$$

d: What is the vector for the projection of $\boldsymbol{\beta}$ into this subspace?

$$\mathbf{p} = \mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\boldsymbol{\beta} = \begin{pmatrix} 2.86 \\ 2.28 \\ 0.44 \end{pmatrix}$$

e: What is the angle θ between this projection and $\boldsymbol{\beta}$?

$$\theta = \cos^{-1}\left(\frac{\mathbf{p}^T\boldsymbol{\beta}}{\sqrt{\mathbf{p}^T\mathbf{p}}\sqrt{\boldsymbol{\beta}^T\boldsymbol{\beta}}}\right) = \cos^{-1}(0.985) = 9.8 \text{ degrees}$$