

# THE EQUATIONS OF SELECTION THEORY

**Tom J. M. Van Dooren and Inez Demon**

*version 04/12/02*

Matrices describe linear transformations of coordinates  $\mathbf{x}$  with respect to a certain basis into coordinates  $\mathbf{y}$  with respect to a different basis (or to the same basis again, then  $\mathbf{A}$  has to be square). Axler (1997) gives a nice introduction to matrices and their connection to linear mappings. In this document, we use linear algebra to derive the most often used equations in selection theory.

## 1. VARIANCE-COVARIANCE MATRICES, SYMMETRIC MATRICES

Suppose  $x_1, x_2, \dots, x_n$  are random variables with means  $\mu_1, \mu_2, \dots$ , variances  $\sigma_1^2, \sigma_2^2, \dots$  (or  $\sigma_{11}, \sigma_{22}, \dots, \sigma_{nn}$ ) and covariances  $\sigma_{12}, \sigma_{13}, \dots, \sigma_{n-1,n}$ . We can collect the random variables in a vector  $\mathbf{x}$ , write the means as a vector  $\boldsymbol{\mu}$ , and the variances and covariances in a matrix, the “variance-covariance” matrix  $\boldsymbol{\Sigma}_x$  of  $\mathbf{x}$ .

$$\text{var}(\mathbf{x}) = \boldsymbol{\Sigma}_x = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & & & \vdots \\ \sigma_{n1} & \dots & \dots & \sigma_{nn} \end{pmatrix}$$

1.1. Variance-covariance matrices are symmetric,  $\sigma_{ji} = \sigma_{ij}$ , and these symmetric matrices (say  $\mathbf{A}$ ) are non-negative definite (Searle 1982, Horn and Johnson 1985).

Positive definite (p.d.)	Positive semidefinite (p.s.d.)	
$\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$	$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x}$ and $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$ for some $\mathbf{x} \neq \mathbf{0}$	e.g., Searle 1982
$\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$	$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ for all $\mathbf{x}$	e.g., Horn and Johnson 1985

The two classes together, p.d. and p.s.d. are called non-negative definite (n.n.d.). This classification applies to symmetric matrices. One can similarly define negative definite, negative semidefinite and non-positive matrices. Matrices that do not fit in any of the categories, are called indefinite.

1.2. Symmetric matrices have the following useful properties (Searle 1982, Walsh & Lynch manuscript):

1) If  $A^{-1}$  exists for a symmetric matrix  $A$ , then  $A^{-1}$  is symmetric as well.

2) The eigenvalues and eigenvectors of a symmetric matrix are real

3) A symmetric matrix  $A$  is diagonalizable as follows:

$A = U\Lambda U^T$  with  $\Lambda$  a diagonal matrix and  $U$  an orthonormal matrix (i.e.,  $U^{-1} = U^T$ ). The entries on the diagonal of  $\Lambda$  are the eigenvalues of  $A$  and the columns of matrix  $U$  are the eigenvectors of  $A$ . Then one can write  $A = U^T A U$ .

1.3. Since variance-covariance matrices are symmetric, they are diagonalizable. This is done by choosing a specific rotation of the original coordinate system. Orthonormal matrices  $U$  describe such rotations. We rotate a vector  $\mathbf{x}$  of stochastic variables with variance-covariance matrix  $\Sigma_x$  as follows towards a vector  $\mathbf{y}$ :  $\mathbf{y} = U^T \mathbf{x}$ . Matrix  $U$  has the normalized eigenvectors of  $\Sigma_x$  as columns. This specific rotation is called the 'canonical transformation' of  $\mathbf{x}$ . The variance-covariance matrix of  $\mathbf{y}$  becomes  $\Sigma_y = U^T \Sigma_x U$ . This is a diagonal matrix  $\Lambda_y$ , all covariances between  $y_i$  become equal to zero.

1.4. A quadratic form of  $A$ ,  $\mathbf{x}^T A \mathbf{x}$  has the following property:

$$\mathbf{x}^T A \mathbf{x} = (U\mathbf{y})^T A U \mathbf{y} = \mathbf{y}^T U^T A U \mathbf{y} = \mathbf{y}^T \Lambda \mathbf{y} = \sum_{i=1}^n \lambda_i y_i^2, \text{ with } n \text{ the number of rows in } A.$$

This shows that, for positive definite matrices, all eigenvalues must be positive, for non-negative matrices, they cannot be negative.

1.5. If  $A^{-1}$  exists, then it has the same eigenvectors as matrix  $A$ . The eigenvalues of  $A^{-1}$  each are the inverse of an eigenvalue of  $A$ .

## 2. THE MULTIVARIATE NORMAL DISTRIBUTION $MVN(\boldsymbol{\mu}, \boldsymbol{\Sigma}_x)$ .

The probability density function  $f(\mathbf{x})$  of a multivariate normal distribution with vector of means  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}_x$  equals

$$f(\mathbf{x}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}_x|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \quad (2.1)$$

Volumes that satisfy the condition  $f(\mathbf{x}) > c^2$  are delimited by  $n$ -dimensional ellipsoids specified by the equation

$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1}(\mathbf{x} - \boldsymbol{\mu}) = c^2$$

The axes of such an ellipsoid are the eigenvalues  $\boldsymbol{u}_i$  of the matrix  $\boldsymbol{\Sigma}_x$ , with  $i$  the indexation of the eigenvalues and eigenvectors.

The length of the  $i$ -th axis equals  $c\sqrt{\lambda_i}$ , with  $\lambda_i$  the eigenvalue that belongs to eigenvector  $\boldsymbol{u}_i$ .

A slice in dimensions  $i$  and  $j$  through the ellipsoid produces an ellipse in the  $x_i - x_j$  plane.

One can apply a so-called canonical transformation in order to remove the covariances between different variables in  $\mathbf{x}$ . This diagonalization procedure is described in section 1. With  $\mathbf{y} = \mathbf{U}^T(\mathbf{x} - \boldsymbol{\mu})$  one finds that  $\mathbf{y} \sim MVN(\mathbf{0}, \mathbf{A}_x)$ .  $\mathbf{U}$  is the matrix that diagonalizes  $\boldsymbol{\Sigma}_x$ . If we instead apply the transformation  $\mathbf{y}' = \mathbf{A}^{-1/2} \mathbf{U}^T(\mathbf{x} - \boldsymbol{\mu})$ , we find the standard multivariate normal distribution  $\mathbf{y}' \sim MVN(\mathbf{0}, \mathbf{I})$ .

We can split a MVN vector  $\mathbf{x}$  into a pair of sub-vectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , with  $\mathbf{x}_1$  the first  $m$  entries and  $\mathbf{x}_2$  the last  $n - m$  entries of  $\mathbf{x}$ . Then one obtains

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \text{ and } \boldsymbol{\Sigma}_x = \begin{pmatrix} \boldsymbol{\Sigma}_{x_1} & \boldsymbol{\Sigma}_{x_1, x_2} \\ \boldsymbol{\Sigma}_{x_1, x_2}^T & \boldsymbol{\Sigma}_{x_2} \end{pmatrix} \text{ with } \boldsymbol{\Sigma}_{x_1, x_2} \text{ the covariance matrix of elements in } \mathbf{x}_1 \text{ and } \mathbf{x}_2.$$

The conditional random variable  $\mathbf{x}_1 | \mathbf{x}_2$  is also MVN distributed (Searle 1971) with vector of means

$$\boldsymbol{\mu}_{x_1|x_2} = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{x_1, x_2} \boldsymbol{\Sigma}_{x_2}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2) \quad (2.2)$$

and variance-covariance matrix

$$\boldsymbol{\Sigma}_{x_1|x_2} = \boldsymbol{\Sigma}_{x_1} - \boldsymbol{\Sigma}_{x_1, x_2} \boldsymbol{\Sigma}_{x_2}^{-1} \boldsymbol{\Sigma}_{x_1, x_2}^T \quad (2.3)$$

The linear regression of variable  $\mathbf{x}_1$  on  $\mathbf{x}_2$  is given by

$$\mathbf{x}_1 = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{x_1, x_2} \boldsymbol{\Sigma}_{x_2}^{-1} (x_2 - \mu_2) + \boldsymbol{\varepsilon} \quad (2.4)$$

with  $\boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \boldsymbol{\Sigma}_{x_1|x_2})$

See also Tallis (1987).

### 3. THE RESPONSE TO SELECTION

The *breeder's equation* (section 9) gives the expected change across one generation in the averages of a number of phenotypic traits, when a population of individuals is subjected to selection. In this appendix we present the first part necessary to construct this equation: the response to selection. This is the difference in expected *breeding value* or in phenotypic means that we obtain when selecting a batch of individuals (this will be the stock from which offspring are bred) from a base population.

We assume that trait vector  $\mathbf{z}$  is the sum of a genetic component  $\mathbf{g}$  and an independent environmental part  $\mathbf{e}$ . We assume that they are MVN distributed in the base population:  $\mathbf{g} \sim \text{MVN}(\boldsymbol{\mu}_g, \mathbf{G})$  en  $\mathbf{e} \sim \text{MVN}(\mathbf{0}, \mathbf{E})$ . Phenotype  $\mathbf{z}$  then is  $\text{MVN}(\boldsymbol{\mu}_z, \mathbf{P})$  with  $\mathbf{P} = \mathbf{G} + \mathbf{E}$  and  $\boldsymbol{\mu}_z = \boldsymbol{\mu}_g$ .

We now calculate the regression of genetic component vector  $\mathbf{g}$  on a selected phenotype vector  $\mathbf{z}$ . For that, we use results from section 2.

We obtain a MVN vector  $\mathbf{x}$  by concatenating vectors  $\mathbf{g}$  and  $\mathbf{z}$ . Vector  $\mathbf{g}$  provides the first  $n$  elements of  $\mathbf{x}$  and  $\mathbf{z}$  the last  $n$  ones, such that  $\mathbf{x} = (\mathbf{g}, \mathbf{z})^\top$ . The probability distribution of  $\mathbf{x}$  is given by the vector of means and the variance covariance matrix:

$$\mathbf{x} = \begin{pmatrix} \mathbf{g} \\ \mathbf{z} \end{pmatrix} = \text{MVN} \left( \begin{pmatrix} \boldsymbol{\mu}_z \\ \boldsymbol{\mu}_g \end{pmatrix}, \begin{pmatrix} \mathbf{G} & \mathbf{G} \\ \mathbf{G} & \mathbf{P} \end{pmatrix} \right)$$

The covariance matrix of genotypic values  $\mathbf{g}$  and phenotype  $\mathbf{z}$  equals  $\mathbf{G}$  because genotype and environment are independently distributed:

$$\begin{aligned} \sigma(\mathbf{g}, \mathbf{z}) &= \sigma(\mathbf{g}, \mathbf{g} + \mathbf{e}) = \sigma(\mathbf{g}, \mathbf{g}) + \sigma(\mathbf{g}, \mathbf{e}) = \sigma(\mathbf{g}, \mathbf{g}) = \mathbf{G} \\ \sigma(\mathbf{g}, \mathbf{g}) &= \mathbf{G} \\ \sigma(\mathbf{z}, \mathbf{z}) &= \mathbf{P} = \mathbf{G} + \mathbf{E} \end{aligned}$$

Using results in section 2 we then obtain

$$\boldsymbol{\mu}_g|_z = \boldsymbol{\mu}_g + \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z) \quad (3.1)$$

and

$$\mathbf{V}_e = \mathbf{V}_g|_z = \mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{G} \quad (3.2)$$

The linear regression of  $\mathbf{g}$  on  $\mathbf{z}$  becomes

$$\mathbf{g} = \boldsymbol{\mu}_g|_z + \boldsymbol{\varepsilon} = \boldsymbol{\mu}_g + \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z) + \boldsymbol{\varepsilon} \quad (3.3)$$

$$\mathbf{g} - \boldsymbol{\mu}_g = \mathbf{G}\mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z) + \boldsymbol{\varepsilon}, \quad \text{and} \quad \boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \mathbf{V}_e)$$

The change in genotypic value  $\mathbf{g} - \boldsymbol{\mu}_g$  in selected individuals with phenotype  $\mathbf{z}$ , depends on the difference between the selected phenotype  $\mathbf{z}$  and the expected phenotype  $\boldsymbol{\mu}_z$  in the base population. We see that this phenotypic difference is mapped through matrix  $\mathbf{GP}^{-1}$  onto the space of genotypic values.

Independence holds between  $\boldsymbol{\varepsilon}$  and  $\mathbf{z}$ , so the regression of  $\mathbf{g}$  on  $\mathbf{z}$  is linear and homoscedastic. The value  $\mathbf{g} - \boldsymbol{\mu}_g$  is often called the *breeding value* of phenotype  $\mathbf{z}$ .

For a univariate phenotype we find, with  $\mathbf{g} = g$ ,  $\boldsymbol{\mu}_g = \mu_z = \mu$ ,  $\mathbf{G} = \sigma_g^2$  and  $\mathbf{P} = \sigma_z^2$

$$g - \mu = \sigma_g^2 \sigma_z^{-2} (z - \mu) + \boldsymbol{\varepsilon} = h^2(z - \mu) + \boldsymbol{\varepsilon}$$

$$\text{and } \sigma_{\boldsymbol{\varepsilon}}^2 = \sigma_{g/z}^2 = \sigma_g^2 - \sigma_g^2 \sigma_z^{-2} \sigma_g^2 = \sigma_g^2(1 - h^2)$$

When all changes in  $\mathbf{z}$  between base and selected population are due to changes in breeding value (and not environmental trends), one can calculate the expectation over the breeding values per phenotype  $\mathbf{z}$ , as given by expression (3.3). To this end, we average over all selected phenotypes. The mean genotypic value in our selected population is written as  $\boldsymbol{\mu}_g^*$ . The change in the expected value of  $\mathbf{g}$  after selection becomes:

$$\begin{aligned} \Delta \boldsymbol{\mu}_g &= \boldsymbol{\mu}_g^* - \boldsymbol{\mu}_g \\ &= \mathbf{E}[\mathbf{GP}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z) + \boldsymbol{\varepsilon}] \\ &= \mathbf{GP}^{-1}\mathbf{E}[(\mathbf{z} - \boldsymbol{\mu}_z)] + \mathbf{E}[\boldsymbol{\varepsilon}] \\ &= \mathbf{GP}^{-1}\mathbf{E}[(\mathbf{z} - \boldsymbol{\mu}_z)] \\ &= \mathbf{GP}^{-1}\mathbf{s} \\ &= \mathbf{G}\boldsymbol{\beta} \end{aligned} \tag{3.4}$$

The vector  $\mathbf{s}$  is usually called the *selection differential*, and vector  $\boldsymbol{\beta}$  the *selection gradient*.

$\mathbf{E}$  signifies calculating the expected value in the population after selection. We assume the absence of environmental trends, i.e.,  $\mathbf{E}[\boldsymbol{\varepsilon}] = \mathbf{0}$ .

We now calculate  $\mathbf{G}^*$ , the genetic variance-covariance matrix after selection.

$$\begin{aligned} \mathbf{G}^* &= \mathbf{E}[(\mathbf{g} - \boldsymbol{\mu}_g^*)(\mathbf{g} - \boldsymbol{\mu}_g^*)^T] \\ &= \mathbf{E}[(\mathbf{g} - \boldsymbol{\mu}_g - \Delta \boldsymbol{\mu}_g)(\mathbf{g} - \boldsymbol{\mu}_g - \Delta \boldsymbol{\mu}_g)^T] \\ &= \mathbf{E}[(\mathbf{GP}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z^*) + \boldsymbol{\varepsilon})(\mathbf{GP}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z^*) + \boldsymbol{\varepsilon})^T] \\ &= \mathbf{E}[(\mathbf{GP}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z^*) + \boldsymbol{\varepsilon})(\mathbf{z} - \boldsymbol{\mu}_z^*)^T \mathbf{P}^{-1} \mathbf{G} + \boldsymbol{\varepsilon}^T]) \\ &= \mathbf{E}[\mathbf{GP}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z^*)(\mathbf{z} - \boldsymbol{\mu}_z^*)^T \mathbf{P}^{-1} \mathbf{G}] + \mathbf{E}[\mathbf{GP}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z^*) \boldsymbol{\varepsilon}^T] + \mathbf{E}[\boldsymbol{\varepsilon}(\mathbf{z} - \boldsymbol{\mu}_z^*)^T \mathbf{P}^{-1} \mathbf{G}] + \\ &\quad \mathbf{E}[\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T] \\ &= \mathbf{GP}^{-1} \mathbf{E}[(\mathbf{z} - \boldsymbol{\mu}_z^*)(\mathbf{z} - \boldsymbol{\mu}_z^*)^T] \mathbf{P}^{-1} \mathbf{G} + \mathbf{GP}^{-1} \mathbf{E}[(\mathbf{z} - \boldsymbol{\mu}_z^*)] \mathbf{E}[\boldsymbol{\varepsilon}^T] + \mathbf{E}[\boldsymbol{\varepsilon}] \mathbf{E}[(\mathbf{z} - \boldsymbol{\mu}_z^*)^T] \mathbf{P}^{-1} \mathbf{G} + \\ &\quad \mathbf{E}[\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}^T] \\ &= \mathbf{GP}^{-1} \mathbf{P}^* \mathbf{P}^{-1} \mathbf{G} + \mathbf{G} - \mathbf{GP}^{-1} \mathbf{G} \end{aligned} \tag{3.5}$$

$\mathbf{P}^*$  equals the phenotypic variance covariance matrix after selection. The change in the genetic variance covariance matrix becomes (Walsh and Lynch manuscript, Lande and Arnold 1983),

$$\Delta\mathbf{G} = \mathbf{G}^* - \mathbf{G} = \mathbf{G}\mathbf{P}^{-1}(\mathbf{P}^* - \mathbf{P})\mathbf{P}^{-1}\mathbf{G} \quad (3.6)$$

#### 4. MULTIVARIATE SELECTION WITH GAUSSIAN SELECTION/FITNESS FUNCTIONS

The Gaussian fitness function  $W(\mathbf{z})$  is often used in selection models. This is because all probability distributions of genotypic values and phenotypes remain normal after selection.

$$W(\mathbf{z}) = \exp(\boldsymbol{\alpha}^T \mathbf{z} - \frac{1}{2}(\mathbf{z} - \boldsymbol{\theta})^T \mathbf{W}(\mathbf{z} - \boldsymbol{\theta})) = \sum_i \alpha_i z_i - \frac{1}{2} \sum_i \sum_j (z_i - \theta_i)(z_j - \theta_j) W_{ij} \quad (4.1)$$

Please do not confuse function  $W(\mathbf{z})$  with the matrix  $\mathbf{W}$ . The Gaussian fitness function allows for directional, and stabilising or as disruptive selection.

Some examples:

- (1) Directional selection occurs when  $\boldsymbol{\alpha} \neq 0$  and/or  $\boldsymbol{\mu}_z \neq \boldsymbol{\theta}$
- (2) Stabilising selection occurs, for example, when  $\mathbf{W}$  is a diagonal matrix and  $W_{ii} > 0$
- (3) Disruptive selection occurs when  $\mathbf{W}$  is a diagonal matrix and  $W_{ii} < 0$

We can usually diagonalize matrix  $\mathbf{W}$  by transforming to a different basis (Section 1). The signs of eigenvalues with respect to that basis show in which directions selection is stabilising or disruptive (see section 7 below).

When  $p(\mathbf{z})$  represents the probability distribution of phenotypes before selection, the probability distribution of phenotypes after selection becomes  $p^*(\mathbf{z})$

$$p^*(\mathbf{z}) = \frac{p(\mathbf{z})W(\mathbf{z})}{\int d\mathbf{z} p(\mathbf{z})W(\mathbf{z})} = cp(\mathbf{z})W(\mathbf{z}) \quad (4.2)$$

Equation (4.2) does not require Gaussian fitness functions. When  $p(\mathbf{z})$  is the MVN( $\boldsymbol{\mu}_z, \mathbf{P}$ ) probability distribution,  $p^*(\mathbf{z})$  becomes

$$\begin{aligned} p^*(\mathbf{z}) &= cp(\mathbf{z})W(\mathbf{z}) \\ &= c(2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu}_z)^T \mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z)\right) \cdot \exp\left(\boldsymbol{\alpha}^T \mathbf{z} - \frac{1}{2}(\mathbf{z} - \boldsymbol{\theta})^T \mathbf{W}(\mathbf{z} - \boldsymbol{\theta})\right) \quad (4.3) \\ &= c(2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(\boldsymbol{\alpha}^T \mathbf{z} - \frac{1}{2}\left((\mathbf{z} - \boldsymbol{\mu}_z)^T \mathbf{P}^{-1}(\mathbf{z} - \boldsymbol{\mu}_z) + (\mathbf{z} - \boldsymbol{\theta})^T \mathbf{W}(\mathbf{z} - \boldsymbol{\theta})\right)\right) \end{aligned}$$

If  $p^*(\mathbf{z})$  is MVN, then we can write it as

$$p^*(\mathbf{z}) = (2\pi)^{-n/2} |\mathbf{P}^*|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{z} - \boldsymbol{\mu}_z^*)^T \mathbf{P}^{*-1}(\mathbf{z} - \boldsymbol{\mu}_z^*)\right) \quad (4.4)$$



Note that we use symbol  $\mu_z^*$  in this section for the mean phenotype after selection. Collecting the terms in (4.3) and (4.4) that are quadratic in  $z$ , we observe that the following equality needs to hold true for obtaining MVN normality after selection:

$$\begin{aligned}\frac{1}{2}z^T P^{*-1}z &= \frac{1}{2}z^T (W + P^{-1})z \\ \Leftrightarrow P^* &= (W + P^{-1})^{-1}\end{aligned}\tag{4.5a}$$

This will be the expression for the phenotypic variance after selection. Alternatively, we can write this variance  $P^*$  as

$$P^* = (W + P^{-1})^{-1} = P(P + W^{-1})^{-1}W^{-1}\tag{4.5b}$$

Proof:

$$\begin{aligned}I &= P^{-1}(W + P^{-1})^{-1}WW^{-1}(W + P^{-1})P \\ &= P^{-1}(W + P^{-1})^{-1}W(I + W^{-1}P^{-1})P \\ &= P^{-1}(W + P^{-1})^{-1}W(P + W^{-1}) \\ \Leftrightarrow (W + P^{-1})^{-1} &= P(P + W^{-1})^{-1}W^{-1}\end{aligned}$$

Similarly, we can derive an expression for the expected value  $\mu_z^*$  from the cross terms of  $\mu_z^*$  en  $z$ :

$$\begin{aligned}\mu_z^{*T} P^{*-1}z &= \theta^T Wz + \alpha^T z + \mu_z^T P^{-1}z \\ \Leftrightarrow \mu_z^{*T} P^{*-1} &= (\theta^T W + \alpha^T + \mu_z^T P^{-1})P^* P^{*-1} \\ \Leftrightarrow \mu_z^{*T} &= (\theta^T W + \alpha^T + \mu_z^T P^{-1})P^* \\ \Leftrightarrow \mu_z^* &= P^*(W\theta + \alpha + P^{-1}\mu_z)\end{aligned}\tag{4.6}$$

This implies that  $p^*$  will be a MVN probability distribution, described by (4.5) and (4.6). All that remains to be done is deriving an expression for the normalization constant  $c$ . We can just as well derive an expression for the mean fitness  $\bar{W}(z)$  in the original base population, since

$$\bar{W}(z) = \int dz p(z) W(z) = c^{-1}$$

$$\begin{aligned}
& \overline{W}(z) \\
&= (2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \int dz \exp\left(-\frac{1}{2}(z - \boldsymbol{\mu}_z)^T \mathbf{P}^{-1}(z - \boldsymbol{\mu}_z)\right) \cdot \exp\left(\boldsymbol{\alpha}^T z - \frac{1}{2}(z - \boldsymbol{\theta})^T \mathbf{W}(z - \boldsymbol{\theta})\right) \\
&= (2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z - \frac{1}{2}\boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta}\right) \int dz \exp\left(-\frac{1}{2}z^T (\mathbf{P}^{-1} + \mathbf{W})z + z^T (\boldsymbol{\alpha} + \mathbf{W}\boldsymbol{\theta} + \mathbf{P}^{-1}\boldsymbol{\mu}_z)\right) \\
&= (2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z - \frac{1}{2}\boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta}\right) \int dz \exp\left(-\frac{1}{2}z^T \mathbf{P}^{*-1}z + z^T \mathbf{P}^{*-1}\boldsymbol{\mu}_z^*\right) \\
&= (2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z - \frac{1}{2}\boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} + \frac{1}{2}\boldsymbol{\mu}_z^{*T} \mathbf{P}^{*-1}\boldsymbol{\mu}_z^*\right) \int dz \exp\left(-\frac{1}{2}(z - \boldsymbol{\mu}_z^*)^T \mathbf{P}^{*-1}(z - \boldsymbol{\mu}_z^*)\right) \\
&= (2\pi)^{-n/2} |\mathbf{P}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z - \frac{1}{2}\boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} + \frac{1}{2}\boldsymbol{\mu}_z^{*T} \mathbf{P}^{*-1}\boldsymbol{\mu}_z^*\right) (2\pi)^{n/2} |\mathbf{P}^*|^{1/2} \\
&= \sqrt{\frac{|\mathbf{P}^*|}{|\mathbf{P}|}} \exp\left(-\frac{1}{2}(\boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z + \boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} - \boldsymbol{\mu}_z^{*T} \mathbf{P}^{*-1}\boldsymbol{\mu}_z^*)\right)
\end{aligned} \tag{4.7}$$

In the last derivation, we made use of the equality:

$$(2\pi)^{-n/2} |\mathbf{P}^*|^{-1/2} \int dz \exp\left(-\frac{1}{2}(z - \boldsymbol{\mu}_z^*)^T \mathbf{P}^{*-1}(z - \boldsymbol{\mu}_z^*)\right) = 1$$

Looking at  $\overline{W}(z)$ , we see that the polynomial in the exponential function can be rewritten as follows, such that  $\boldsymbol{\mu}_z^*$  can be eliminated from the expression:

$$\begin{aligned}
& \boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z + \boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} - \boldsymbol{\mu}_z^{*T} \mathbf{P}^{*-1}\boldsymbol{\mu}_z^* \\
&= \boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z + \boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} - (\mathbf{P}^* (\mathbf{P}^{-1}\boldsymbol{\mu}_z + \mathbf{b}))^T \mathbf{P}^{*-1} (\mathbf{P}^* (\mathbf{P}^{-1}\boldsymbol{\mu}_z + \mathbf{b})) \\
&= \boldsymbol{\mu}_z^T \mathbf{P}^{-1}\boldsymbol{\mu}_z + \boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} - (\boldsymbol{\mu}_z^T \mathbf{P}^{-1} + \mathbf{b}^T) \mathbf{P}^* (\mathbf{P}^{-1}\boldsymbol{\mu}_z + \mathbf{b}) \\
&= \boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta} + \boldsymbol{\mu}_z^T \mathbf{P} (\mathbf{I} - \mathbf{P}^* \mathbf{P}^{-1}) \boldsymbol{\mu}_z - 2\mathbf{b}^T \mathbf{P}^* \mathbf{P}^{-1} \boldsymbol{\mu}_z - \mathbf{b}^T \mathbf{P}^* \mathbf{b}
\end{aligned} \tag{4.8}$$

with  $\mathbf{b} = \mathbf{W}\boldsymbol{\theta} + \boldsymbol{\alpha}$ .

We now know that the probability distribution of phenotypes stays MVN, but what about the breeding values  $\mathbf{g}$ ?

We define the fitness of a genotype  $w(\mathbf{g})$  as

$$w(\mathbf{g}) = \int dz p(z | \mathbf{g}) w(z) \tag{4.9}$$

and the probability distribution of genotypes after selection is  $p^*(\mathbf{g})$ , with

$$p^*(\mathbf{g}) = \frac{w(\mathbf{g})p(\mathbf{g})}{\int d\mathbf{g}p(\mathbf{g})w(\mathbf{g})} \quad (4.10)$$

The conditional distribution of  $\mathbf{z}$  given  $\mathbf{g}$  is  $\text{MVN} \sim (\mathbf{g}, \mathbf{E})$ . We can show that  $w(\mathbf{g})$  becomes a "Gaussian" exponential function.

$$\begin{aligned} w(\mathbf{g}) &= (2\pi)^{-n/2} |\mathbf{E}|^{-1/2} \int d\mathbf{z} \exp\left(-\frac{1}{2}(\mathbf{z} - \mathbf{g})^T \mathbf{E}^{-1}(\mathbf{z} - \mathbf{g})\right) \cdot \exp\left(\boldsymbol{\alpha}^T \mathbf{z} - \frac{1}{2}(\mathbf{z} - \boldsymbol{\theta})^T \mathbf{W}(\mathbf{z} - \boldsymbol{\theta})\right) \\ &= (2\pi)^{-n/2} |\mathbf{E}|^{-1/2} \int d\mathbf{e} \exp\left(-\frac{1}{2}(\mathbf{e})^T \mathbf{E}^{-1}(\mathbf{e})\right) \cdot \exp\left(\boldsymbol{\alpha}^T (\mathbf{g} + \mathbf{e}) - \frac{1}{2}(\mathbf{g} + \mathbf{e} - \boldsymbol{\theta})^T \mathbf{W}(\mathbf{g} + \mathbf{e} - \boldsymbol{\theta})\right) \\ &= (2\pi)^{-n/2} |\mathbf{E}|^{-1/2} \exp\left(\mathbf{g}^T (\boldsymbol{\alpha} + \mathbf{W}\boldsymbol{\theta}) - \frac{1}{2}\mathbf{g}^T \mathbf{W}\mathbf{g} - \frac{1}{2}\boldsymbol{\theta}^T \mathbf{W}\boldsymbol{\theta}\right) \int d\mathbf{z} \exp\left(-\frac{1}{2}\mathbf{e}^T (\mathbf{E}^{-1} + \mathbf{W})\mathbf{e} + \mathbf{e}^T (\boldsymbol{\alpha} + \mathbf{W}\boldsymbol{\theta} - \mathbf{W}\mathbf{g})\right) \\ &= c_2 \exp\left(\mathbf{g}^T \boldsymbol{\alpha} - \frac{1}{2}(\mathbf{g} - \boldsymbol{\theta})^T \mathbf{W}(\mathbf{g} - \boldsymbol{\theta})\right) \end{aligned} \quad (4.11)$$

The probability distribution of breeding values after selection  $p^*(\mathbf{g})$  then equals a gaussian times a normal distribution times a constant, which becomes a new multivariate normal distribution (compare with Eqn. 4.2).

To conclude, we now derive the selection differential  $s = \boldsymbol{\mu}_z^* - \boldsymbol{\mu}_z$  in the context of gaussian fitness functions. Substituting  $\boldsymbol{\mu}_z^*$  yields

$$\begin{aligned} s &= \mathbf{P}^*(\mathbf{W}\boldsymbol{\theta} + \boldsymbol{\alpha} + \mathbf{P}^{-1}\boldsymbol{\mu}_z) - \boldsymbol{\mu}_z \\ &= \mathbf{P}^*(\mathbf{W}\boldsymbol{\theta} + \boldsymbol{\alpha} + \mathbf{P}^{-1}\boldsymbol{\mu}_z - \mathbf{P}^{*-1}\boldsymbol{\mu}_z) \\ &= \mathbf{P}^*(\mathbf{W}\boldsymbol{\theta} + \boldsymbol{\alpha} + \mathbf{P}^{-1}\boldsymbol{\mu}_z - \mathbf{W}\boldsymbol{\mu}_z - \mathbf{P}^{-1}\boldsymbol{\mu}_z) \\ &= \mathbf{P}^*(\mathbf{W}(\boldsymbol{\theta} - \boldsymbol{\mu}_z) + \boldsymbol{\alpha}) \end{aligned} \quad (4.12)$$

and

$$\begin{aligned} \mathbf{G}^* &= \mathbf{G}\mathbf{P}^{-1}\mathbf{P}^*\mathbf{P}^{-1}\mathbf{G} + \mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{G} \\ &= \mathbf{G}\mathbf{P}^{-1}\mathbf{P}(\mathbf{P} + \mathbf{W}^{-1})^{-1}\mathbf{W}^{-1}\mathbf{P}^{-1}\mathbf{G} + \mathbf{G} - \mathbf{G}\mathbf{P}^{-1}\mathbf{G} \\ &= \mathbf{G} - \mathbf{G}((\mathbf{P} + \mathbf{W}^{-1})^{-1}\mathbf{W}^{-1} - \mathbf{I})\mathbf{P}^{-1}\mathbf{G} \\ &= \mathbf{G} - \mathbf{G}(\mathbf{P} + \mathbf{W}^{-1})^{-1}(\mathbf{W}^{-1} - \mathbf{P} - \mathbf{W}^{-1})\mathbf{P}^{-1}\mathbf{G} \\ &= \mathbf{G} - \mathbf{G}(\mathbf{P} + \mathbf{W}^{-1})^{-1}\mathbf{G} \end{aligned} \quad (4.13)$$

At an equilibrium for a selection process  $s_i$  becomes zero, such that either

$$\boldsymbol{\mu}_i = \boldsymbol{\theta} + \mathbf{W}^{-1}\boldsymbol{\alpha} \quad (4.14)$$

or  $\mathbf{W}(\boldsymbol{\theta} - \boldsymbol{\mu}_z) + \boldsymbol{\alpha}$  lies in the null space of  $\mathbf{P}^*$ .

## 5. GRADIENTS AND VECTORS

The gradient of a function  $f$  with respect to a vector  $\mathbf{x}$  is written as

$$\nabla_{\mathbf{x}} f = \frac{\partial f}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \quad (5.1)$$

A number of useful equalities are easy to derive (e.g. Searle 1982). For a vector  $\mathbf{a}$  or a matrix  $\mathbf{A}$  one finds:

$$\nabla_{\mathbf{x}} (\mathbf{a}^T \mathbf{x}) = \nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{a}) = \mathbf{a}$$

$$\nabla_{\mathbf{x}} (\mathbf{A}\mathbf{x}) = \mathbf{A}^T$$

$$\nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{x}) = 2\mathbf{x}$$

$$\nabla_{\mathbf{x}} \exp(f(\mathbf{x})) = \exp(f(\mathbf{x})) \cdot \nabla_{\mathbf{x}} f(\mathbf{x})$$

$$\nabla_{\mathbf{x}} \ln(f(\mathbf{x})) = \frac{1}{f(\mathbf{x})} \cdot \nabla_{\mathbf{x}} f(\mathbf{x})$$

With  $\mathbf{A}$  a symmetric matrix, derivatives of quadratic forms become

$$\nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{A}\mathbf{x}) = 2\mathbf{A}\mathbf{x}$$

$$\nabla_{\mathbf{x}} ((\mathbf{x} - \mathbf{y})^T \mathbf{A}(\mathbf{x} - \mathbf{y})) = 2\mathbf{A}(\mathbf{x} - \mathbf{y}) \quad \text{or} \quad \nabla_{\mathbf{x}} ((\mathbf{y} - \mathbf{x})^T \mathbf{A}(\mathbf{y} - \mathbf{x})) = 2\mathbf{A}(\mathbf{y} - \mathbf{x})$$

## 6. THE GRADIENT OF FITNESS $W(z)$

When  $f(x)$  is a MVN probability distribution with

$$f(x) = (2\pi)^{-n/2} |\Sigma_x|^{-1/2} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma_x^{-1}(x - \mu)\right) \quad (6.1)$$

then  $\nabla_x f(x)$  becomes (Section 5)

$$\begin{aligned} \nabla_x f(x) &= f(x) \cdot \nabla_x \left(-\frac{1}{2}(x - \mu)^T \Sigma_x^{-1}(x - \mu)\right) \\ &= -f(x) \Sigma_x^{-1}(x - \mu) \end{aligned} \quad (6.2)$$

If we consider  $f(x)$  as a function of the mean  $\mu$ , we get instead the partial derivative

$$\begin{aligned} \nabla_\mu f(x, \mu) &= f(x, \mu) \cdot \nabla_\mu \left(-\frac{1}{2}(x - \mu)^T \Sigma_x^{-1}(x - \mu)\right) \\ &= f(x, \mu) \Sigma_x^{-1}(x - \mu) \end{aligned} \quad (6.3)$$

We can use this to show that  $\nabla_{\mu_z} \ln(\bar{W}(z)) = \beta$  when the probability distribution of phenotypes in the base population is MVN (Lande 1979). Note that, in this manner, the selection gradient is determined by the vector of phenotypic mean trait values in the population.

Proof:

$$\begin{aligned} \nabla_{\mu_z} \ln(\bar{W}(z)) &= \bar{W}^{-1} \nabla_{\mu_z} \bar{W}(z) \\ &= \nabla_{\mu_z} \int dz f(z, \mu_z) \frac{W(z)}{\bar{W}} \\ &= \nabla_{\mu_z} \int dz f(z, \mu_z) w(z) \\ &= \int dz w(z) f(z) P^{-1}(z - \mu_z) \\ &= P^{-1} \left( \int dz w(z) f(z) z - \mu_z \int dz w(z) f(z) \right) \\ &= P^{-1}(\mu_z^* - \mu_z) = P^{-1}s = \beta \end{aligned} \quad (6.4)$$

This result does not depend on the fitness function assumed. When the fitness function is Gaussian, we can carry out the following calculation, with obvious result (Lynch and Walsh manuscript, see the results in Section 4):

$$\begin{aligned}
\nabla_{\boldsymbol{\mu}_z} \ln(\overline{W}(z)) &= -\frac{1}{2} \nabla_{\boldsymbol{\mu}_z} (\boldsymbol{\theta}^T \mathbf{W} \boldsymbol{\theta} + \boldsymbol{\mu}_z^T \mathbf{P}^{-1} (\mathbf{I} - \mathbf{P}^* \mathbf{P}^{-1}) \boldsymbol{\mu}_z - 2\mathbf{b}^T \mathbf{P}^* \mathbf{P}^{-1} \boldsymbol{\mu}_z - \mathbf{b}^T \mathbf{P}^* \mathbf{b}) \\
&= -\frac{1}{2} \nabla_{\boldsymbol{\mu}_z} (\boldsymbol{\mu}_z^T \mathbf{P}^{-1} (\mathbf{I} - \mathbf{P}^* \mathbf{P}^{-1}) \boldsymbol{\mu}_z - 2\mathbf{b}^T \mathbf{P}^* \mathbf{P}^{-1} \boldsymbol{\mu}_z) \\
&= \mathbf{P}^{-1} (\mathbf{P}^* \mathbf{P}^{-1} - \mathbf{I}) \boldsymbol{\mu}_z + \mathbf{P}^{-1} \mathbf{P}^* \mathbf{b} \\
&= \mathbf{P}^{-1} \mathbf{P}^* ((\mathbf{P}^{-1} - \mathbf{P}^{*-1}) \boldsymbol{\mu}_z + \mathbf{b}) \\
&= \mathbf{P}^{-1} \mathbf{P}^* (-\mathbf{W} \boldsymbol{\mu}_z + \mathbf{W} \boldsymbol{\theta} + \boldsymbol{\alpha}) \\
&= \mathbf{P}^{-1} \mathbf{P}^* (\mathbf{W} (\boldsymbol{\theta} - \boldsymbol{\mu}_z) + \boldsymbol{\alpha}) \\
&= \mathbf{P}^{-1} \mathbf{s} = \boldsymbol{\beta}
\end{aligned} \tag{6.5}$$

Here we again used a result following from (4.5a):  $\mathbf{P}^{*-1} = \mathbf{W} + \mathbf{P}^{-1}$ . We now demonstrate that the average fitness in a population increases (Lande 1979), or rather, that it never decreases.

To this end, we write out a first-order approximation for the change in the logarithm of average fitness. We treat mean fitness as a function of the population mean this time.

$$\begin{aligned}
\Delta \ln(\overline{W}(\boldsymbol{\mu}_z)) &= \ln(\overline{W}(\boldsymbol{\mu}_z + \Delta \boldsymbol{\mu}_z)) - \ln(\overline{W}(\boldsymbol{\mu}_z)) \\
&\cong \left( \nabla_{\boldsymbol{\mu}_z} \ln(\overline{W}(\boldsymbol{\mu}_z)) \right)^T \Delta \boldsymbol{\mu}_z \\
&= \boldsymbol{\beta} \mathbf{G} \boldsymbol{\beta} \geq 0
\end{aligned} \tag{6.6}$$

We find that this quadratic form cannot be negative, since matrix  $\mathbf{G}$  is a non-negative definite variance - covariance matrix (Section 1).

## 7. STABILIZING AND DISRUPTIVE SELECTION

To investigate the properties of an extremum of a function, one makes use of the *Hessian*  $\mathbf{H}(f(\mathbf{x}))$ , a matrix of second-order derivatives (7.1).

$$\mathbf{H}_x f(\mathbf{x}) = \nabla_x (\nabla_x f(\mathbf{x}))^T = \frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}^T} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \quad (7.1)$$

Under regular conditions, this matrix is symmetric (consult an analysis textbook to find out about these conditions). When the Hessian  $\mathbf{H}$ , evaluated at a point  $\mathbf{x}_0$ , is positive definite, then  $\mathbf{x}_0$  is a local minimum. with  $\mathbf{H}$  negative definite,  $\mathbf{x}_0$  is a local maximum. When some eigenvalues of  $\mathbf{H}$  are negative and some positive, the hessian  $\mathbf{H}$  corresponds to a saddle.

For example, the Hessian of the MVN probability distribution, seen as a a function of vector  $\mathbf{x}$  equals

$$\begin{aligned} H_x f(\mathbf{x}) &= \nabla_x (\nabla_x f(\mathbf{x}))^T = f(\mathbf{x}) \cdot \nabla_x \left( -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \\ &= -\nabla_x \left( f(\mathbf{x}) (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1} \right) \\ &= -\nabla_x f(\mathbf{x}) \left( (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1} \right) - f(\mathbf{x}) \nabla_x \left( (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1} \right) \\ &= f(\mathbf{x}) \left( \boldsymbol{\Sigma}_x^{-1} (\mathbf{x} - \boldsymbol{\mu}) (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}_x^{-1} - \boldsymbol{\Sigma}_x^{-1} \right) \end{aligned} \quad (7.2)$$

We now calculate the Hessian of mean fitness, assuming a gaussian fitness function and MVN distributed phenotypes with expected value  $\hat{\boldsymbol{\mu}}_z$ .

$$\begin{aligned} \mathbf{H}_\mu \ln(\overline{W}(\hat{\boldsymbol{\mu}}_z)) &= \nabla_\mu \left( \nabla_\mu \ln(\overline{W}(\hat{\boldsymbol{\mu}}_z)) \right)^T \\ &= \nabla_{\boldsymbol{\mu}_z} \left( \mathbf{P}^{-1} (\mathbf{P}^* \mathbf{P}^{-1} - \mathbf{I}) \boldsymbol{\mu}_z + \mathbf{P}^* \mathbf{W} \boldsymbol{\theta} + \mathbf{P}^* \boldsymbol{\alpha} \right) \Big|_{\boldsymbol{\mu}_z = \hat{\boldsymbol{\mu}}_z} \\ &= \mathbf{P}^{-1} (\mathbf{P}^* \mathbf{P}^{-1} - \mathbf{I}) \\ &= \mathbf{P}^{-1} \mathbf{P}^* (\mathbf{P}^{-1} - \mathbf{P}^{*-1}) \\ &= \mathbf{P}^{-1} \mathbf{P}^* (\mathbf{P}^{-1} - \mathbf{W} - \mathbf{P}^{-1}) \\ &= -\mathbf{P}^{-1} \mathbf{P}^* \mathbf{W} \\ &= -(\mathbf{P} + \mathbf{W}^{-1})^{-1} \end{aligned} \quad (7.3)$$

If all eigenvalues of  $(\mathbf{P} + \mathbf{W}^{-1})$  are positive, then the eigenvalues of  $\mathbf{H}_\mu$  are negative at the point  $\hat{\mu}$ . In that case  $(\mathbf{P} + \mathbf{W}^{-1})$  is positive definite. For every  $\mathbf{x}$  it holds true that

$$\begin{aligned} \mathbf{x}^T (\mathbf{P} + \mathbf{W}^{-1}) \mathbf{x} &= \mathbf{x}^T \mathbf{P} \mathbf{x} + \mathbf{x}^T \mathbf{W}^{-1} \mathbf{x} \\ &= \sum_{i=1}^n y_i^2 \lambda_i + \sum_{j=1}^n u_j^2 \gamma_j^{-1} > 0 \end{aligned} \tag{7.4}$$

We diagonalized  $\mathbf{P}$  and  $\mathbf{W}^{-1}$  in order to obtain this expression. We can suppose that the eigenvalues of  $\mathbf{P}$  are all positive. Provided that the eigenvalues of  $\mathbf{W}$  are positive as well, which implies stabilizing selection on all traits, since  $(\mathbf{P} + \mathbf{W}^{-1})$  is positive definite. We can apply Theorem 7.63 from Horn and Johnson (1985) to derive a somewhat more general result. This theorem states that, if we multiply a symmetric "Hermitian" matrix (i.e., *the transpose of the complex conjugate of such a matrix, is equal to the matrix itself*) by a positive definite matrix, the resulting matrix has as many negative, zero, and positive eigenvalues as the original matrix. Matrix  $\mathbf{W}$  is Hermitian. From Equation (4.5b) we know that

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

When  $\mathbf{P}^{-1}$  and  $\mathbf{P}^*$  are both positive definite, which we can assume in most cases, their product will be positive definite as well. This implies that matrix  $\mathbf{W}$  has as many positive eigenvalues as the  $\mathbf{H}_\mu$  will have negative ones and vice versa. In the case of disruptive selection in some directions, the matrix  $\mathbf{W}$  will have some negative eigenvalues and the Hessian becomes a local minimum of mean fitness in a number of directions. When the product of  $\mathbf{P}^{-1}$  and  $\mathbf{P}^*$  is semi-positive definite, the Hessian does not have more positive (negative) eigenvalues than matrix  $\mathbf{W}$  has negative (positive) eigenvalues (Horn and Johnson 1985).



## 8. DIFFERENCE EQUATIONS FOR MEANS AND VARIANCES: AUTOSOMAL GENES

We can also track means and variances of quantitative traits across generations. In order to achieve this, we need expression that couple the traits of parents  $z_s$  (sire) and  $z_d$  (dam) to the trait values of their offspring. Assuming that parental and offspring phenotypic traits  $z_o$  are MVN, we can make use of the results in Section 2 on conditional distributions.

The joint distribution of parental and offspring phenotypes is

$$\begin{pmatrix} z_o \\ z_s \\ z_d \end{pmatrix} = MVN \left( \begin{pmatrix} \mu_o \\ \mu_s \\ \mu_d \end{pmatrix}, \begin{pmatrix} P_o & G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_o, z_s}^T & P_s & G_{z_s, z_d} \\ G_{z_o, z_d}^T & G_{z_s, z_d}^T & P_d \end{pmatrix} \right) \quad (8.1)$$

where the  $G$  matrix blocks stand for genetic covariances between parents and offspring. With random mating among selected parents, the genetic covariance between parental phenotypes becomes zero. The conditional distribution of  $z_o | (z_s, z_d)$  can be calculated as follows. Conditional on a pair of selected parents, we get as expected value of the offspring phenotypes

$$\mu_{z_o | (z_s, z_d)} = \mu_o + \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \end{pmatrix} \begin{pmatrix} P_s & G_{z_s, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \begin{pmatrix} z_s - \mu_s \\ z_d - \mu_d \end{pmatrix} + \epsilon \quad (8.2)$$

and variance-covariance matrix of the error term  $\epsilon$ :

$$P_{z_o | (z_s, z_d)} = P_o - \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \end{pmatrix} \begin{pmatrix} P_s & G_{z_s, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \begin{pmatrix} G_{z_o, z_s}^T \\ G_{z_o, z_d}^T \end{pmatrix} \quad (8.3)$$

Taking expectations over selected parents, (8.2) leads to the following equation for the expected change in the mean over a generation:

$$\mu_o^* - \mu_o = \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \end{pmatrix} \begin{pmatrix} P_s & G_{z_s, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \begin{pmatrix} s_s \\ s_d \end{pmatrix} \quad (8.4)$$

Recall that selection differentials  $s$  represent the difference between expected trait values in selected parents and base population. The expected breeding value of offspring can be calculated similarly, and is equal to this expression.

This equation is used to track phenotypic and genetic changes in means over generations. It is easy to figure out that, with equal phenotypic and genetic (co)variances in sires and dams, random mating among parents, and without sex differences in selection, that (8.4) reduces to the well-known *breeder's equation*  $\Delta\mu_z = G\beta$ . It gives the expected change

across one generation in the averages of a number of phenotypic traits  $\Delta\mu_z$  when a population of individuals is subjected to a selection gradient  $\beta$ , with  $G$  the genotypic variance-covariance matrix of the involved traits.

The phenotypic variance in the offspring of selected parents becomes

$$\begin{aligned} P_o^* &= E[(z_o - \mu_o^*)(z_o - \mu_o^*)] \\ &= E \left[ \left( \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix} \begin{pmatrix} P_s & G_{z_s, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \begin{pmatrix} z_s - \mu_s^* \\ z_d - \mu_d^* \end{pmatrix} + \varepsilon \right) \left( \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \begin{pmatrix} z_s - \mu_s^* \\ z_d - \mu_d^* \end{pmatrix} + \varepsilon \right)^T \right] \end{aligned} \quad (8.5a)$$

This equals:

$$P_o^* = \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} E \left[ \begin{pmatrix} z_s - \mu_s^* \\ z_d - \mu_d^* \end{pmatrix} \begin{pmatrix} z_s - \mu_s^* \\ z_d - \mu_d^* \end{pmatrix}^T \right] \begin{pmatrix} P_s & G_{z_s, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \end{pmatrix}^T + E[\varepsilon\varepsilon^T] \quad (8.5b)$$

If we mate parents randomly, also from the selected pool, no phenotypic covariance will be introduced among parents:

$$E \left[ \begin{pmatrix} (z_s - \mu_s^*)(z_s - \mu_s^*)^T \\ (z_d - \mu_d^*)(z_d - \mu_d^*)^T \end{pmatrix} \right] = \begin{pmatrix} P_s^* & 0 \\ 0 & P_d^* \end{pmatrix}$$

and we can rewrite (8.5b) as

$$P_o^* = P_o + \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & P_d \end{pmatrix}^{-1} \left[ \begin{pmatrix} P_s^* & 0 \\ 0 & P_d^* \end{pmatrix} - \begin{pmatrix} P_s & 0 \\ 0 & P_d \end{pmatrix} \right] \begin{pmatrix} P_s & 0 \\ 0 & P_d \end{pmatrix}^{-1} \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \end{pmatrix}^T \quad (8.5c)$$

Similarly, we can write for the offspring genetic variance:

$$\begin{aligned} G_o^* &= E[(g_o - \mu_{g,o}^*)(g_o - \mu_{g,o}^*)] \\ &= E \left[ \left( \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & G_d \end{pmatrix} \begin{pmatrix} G_s & G_{z_s, z_d} \\ G_{z_s, z_d}^T & G_d \end{pmatrix}^{-1} \begin{pmatrix} g_s - \mu_{g,s}^* \\ g_d - \mu_{g,d}^* \end{pmatrix} + \varepsilon \right) \left( \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & G_d \end{pmatrix}^{-1} \begin{pmatrix} g_s - \mu_{g,s}^* \\ g_d - \mu_{g,d}^* \end{pmatrix} + \varepsilon \right)^T \right] \end{aligned} \quad (8.6a)$$

Assuming random mating again, we obtain

$$G_o^* = G_o + \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \\ G_{z_s, z_d}^T & G_d \end{pmatrix}^{-1} \begin{bmatrix} G_s^* - G_s & 0 \\ 0 & G_d^* - G_d \end{bmatrix} \begin{pmatrix} G_s & 0 \\ 0 & G_d \end{pmatrix}^{-1} \begin{pmatrix} G_{z_o, z_s} & G_{z_o, z_d} \end{pmatrix}^T$$

(8.6b)

Let's now focus on some simple example calculations, following Lynch and Walsh (1998, manuscript) and assuming fully additive genetics, random mating and equal selection on dams and sires. We obtain:

$$\begin{pmatrix} z_o \\ z_s \\ z_d \end{pmatrix} = MVN \left( \begin{pmatrix} \mu_z \\ \mu_z \\ \mu_z \end{pmatrix}, \begin{pmatrix} P & \frac{1}{2}G & \frac{1}{2}G \\ \frac{1}{2}G^T & P & 0 \\ \frac{1}{2}G^T & 0 & P \end{pmatrix} \right)$$

.as average response the breeder's equation is recovered:

$$\mu_o^* - \mu_z = \begin{pmatrix} \frac{1}{2}G & \frac{1}{2}G \\ \frac{1}{2}G & \frac{1}{2}G \end{pmatrix} \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}^{-1} \begin{pmatrix} s \\ s \end{pmatrix} = GP^{-1}s = G\beta \quad (8.7a)$$

Using equation (8.5c) we obtain the phenotypic variance among offspring:

$$P_o^* = P + \frac{1}{2}GP^{-1}(P_s^* - P)P^{-1}G \quad (8.7b)$$

and as genetic variance, we find using (8.6b),

$$G_o^* = G + \frac{1}{4} \begin{pmatrix} I & I \\ I & I \end{pmatrix} \begin{pmatrix} G^*-G & 0 \\ 0 & G^*-G \end{pmatrix} \begin{pmatrix} I \\ I \end{pmatrix} = G + \frac{1}{2}(G^*-G) \quad (8.7c)$$

9. DIFFERENCE EQUATIONS FOR MEANS AND VARIANCES OF SEX-DEPENDENT TRAITS:  
AUTOSOMAL GENES OR GENES ON SEX CHROMOSOMES

This derivation follows the appendix of Lande (1980) closely.

In the derivations, we will assume additive genetics and we also make use of the so-called additive effects of gametes  $\mathbf{x}_m$  in males and  $\mathbf{x}_f$  in females.

Mean phenotypes before selection are

$$\boldsymbol{\mu}_{z_m} = nE(\mathbf{x}_m) = n\boldsymbol{\mu}_{x_m}$$

in males and

$$\boldsymbol{\mu}_{z_f} = nE(\mathbf{x}_f) = 2\boldsymbol{\mu}_{x_f}$$

in females. We use  $n$  here to allow for autosomal or sex-linked determination. For autosomal genes,  $n$  equals 2, for sex-linked trait determination,  $n$  equals 1 and we assume that the male is the heterogametic sex.

In order to derive next-generation averages, we need to track to parental origins of genetic contributions, hence we use the notation  $f(\mathbf{x}_m)$ , for genes with gametic effects on males that reside in females, and similarly  $m(\mathbf{x}_f)$ , for (hidden) with effects on females but residing in males.

Next generation averages become, with this notation:

$$\begin{aligned}\boldsymbol{\mu}_{x_{o,m}} &= \frac{1}{n}\boldsymbol{\mu}_{f(x_m)}^* + \left(1 - \frac{1}{n}\right)\boldsymbol{\mu}_{x_m}^* \\ \boldsymbol{\mu}_{f(x_{o,m})} &= \frac{\boldsymbol{\mu}_{x_m}^* + \boldsymbol{\mu}_{f(x_m)}^*}{2} \\ \boldsymbol{\mu}_{x_{o,f}} &= \frac{\boldsymbol{\mu}_{x_f}^* + \boldsymbol{\mu}_{m(x_f)}^*}{2} \\ \boldsymbol{\mu}_{m(x_{o,f})} &= \left(1 - \frac{1}{n}\right)\boldsymbol{\mu}_{m(x_f)}^* + \frac{1}{n}\boldsymbol{\mu}_{x_f}^*\end{aligned}\tag{9.1}$$

Using our, by now standard, trick of conditional MVN distributions, we can figure out that, provided that mating is at random with respect to the traits,

$$\begin{aligned}\boldsymbol{\mu}_{x_m}^* &= \boldsymbol{\mu}_{x_m} + \mathbf{G}_{x,m}\mathbf{P}_m^{-1}\mathbf{s}_m \\ \boldsymbol{\mu}_{f(x_m)}^* &= \boldsymbol{\mu}_{f(x_m)} + \mathbf{B}_x\mathbf{P}_f^{-1}\mathbf{s}_f \\ \boldsymbol{\mu}_{x_f}^* &= \boldsymbol{\mu}_{x_f} + \mathbf{G}_{x,f}\mathbf{P}_f^{-1}\mathbf{s}_f \\ \boldsymbol{\mu}_{m(x_f)}^* &= \boldsymbol{\mu}_{m(x_f)} + \mathbf{B}^T_x\mathbf{P}_m^{-1}\mathbf{s}_m\end{aligned}\tag{9.2}$$

with  $\mathbf{G}_{x,m}$ ,  $\mathbf{G}_{x,f}$ , genetic variance covariance matrixes of gametic effects and  $\mathbf{B}_x$  the covariance matrix of effects hidden in one sex and apparent in the other. It is assumed that the matrix for both types of hidden effects is approximately equal, up to transposition (Lande 1980).

For autosomal inheritance, the relationship with "zygotic" variances and covariances is

$$\mathbf{G}_{x,m} = \frac{1}{2}\mathbf{G}_m, \mathbf{G}_{x,f} = \frac{1}{2}\mathbf{G}_f, \text{ and } \mathbf{B}_x = \frac{1}{2}\mathbf{B};$$

and the distributions of effects are the same in zygotes of both sexes, such that

$$\boldsymbol{\mu}_{f(x_m)} = \boldsymbol{\mu}_{x_m} \text{ and } \boldsymbol{\mu}_{m(x_f)} = \boldsymbol{\mu}_{x_f}.$$

We obtain for the phenotypic response in autosomes:

$$\begin{pmatrix} \boldsymbol{\mu}_{z_{o,m}} - \boldsymbol{\mu}_{z_m} \\ \boldsymbol{\mu}_{z_{o,f}} - \boldsymbol{\mu}_{z_f} \end{pmatrix} = 2 \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}}^* - \boldsymbol{\mu}_{x_m} \\ \boldsymbol{\mu}_{x_{o,f}}^* - \boldsymbol{\mu}_{x_f} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \mathbf{G}_m & \mathbf{B}_x \\ \mathbf{B}_x^T & \mathbf{G}_f \end{pmatrix} \begin{pmatrix} \mathbf{P}_m^{-1}\mathbf{s}_m \\ \mathbf{P}_f^{-1}\mathbf{s}_f \end{pmatrix} \quad (9.3)$$

Substitution of (9.2) into (9.1) gives for sex-linked inheritance

$$\begin{aligned} \Delta \begin{pmatrix} \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}} - \boldsymbol{\mu}_{f(x_{o,m})} \\ \boldsymbol{\mu}_{m(x_{o,f})} - \boldsymbol{\mu}_{x_{o,f}} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \boldsymbol{\mu}_{f(x_m)}^* - \boldsymbol{\mu}_{x_m}^* \\ \boldsymbol{\mu}_{x_f}^* - \boldsymbol{\mu}_{m(x_f)}^* \end{pmatrix} - \begin{pmatrix} \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \mathbf{B}_x \mathbf{P}_f^{-1} \mathbf{s}_f - \mathbf{G}_{x,m} \mathbf{P}_m^{-1} \mathbf{s}_m + \boldsymbol{\mu}_{f(x_m)} - \boldsymbol{\mu}_{x_m} \\ \mathbf{G}_{x,f} \mathbf{P}_f^{-1} \mathbf{s}_f - \mathbf{B}_x^T \mathbf{P}_m^{-1} \mathbf{s}_m + \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{m(x_f)} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} \mathbf{G}_{x,m} & \mathbf{B}_x \\ \mathbf{B}_x^T & \mathbf{G}_{x,f} \end{pmatrix} \begin{pmatrix} -\mathbf{P}_m^{-1} \mathbf{s}_m \\ \mathbf{P}_f^{-1} \mathbf{s}_f \end{pmatrix} - \frac{3}{2} \begin{pmatrix} \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f} \end{pmatrix} \end{aligned} \quad (9.4)$$

One can see that the difference between effects in both sexes has an equilibrium value

$$\begin{pmatrix} \boldsymbol{\mu}_{x_m} - \boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} \mathbf{G}_{x,m} & \mathbf{B}_x \\ \mathbf{B}_x^T & \mathbf{G}_{x,f} \end{pmatrix} \begin{pmatrix} -\mathbf{P}_m^{-1} \mathbf{s}_m \\ \mathbf{P}_f^{-1} \mathbf{s}_f \end{pmatrix} \quad (9.5)$$

Substitution of (9.2) into (9.1) also yields

$$\begin{aligned}
\Delta \begin{pmatrix} \boldsymbol{\mu}_{x_m} + 2\boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} + 2\boldsymbol{\mu}_{x_f} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}} + 2\boldsymbol{\mu}_{f(x_{o,m})} \\ \boldsymbol{\mu}_{m(x_{o,f})} + 2\boldsymbol{\mu}_{x_{o,f}} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\mu}_{x_m} + 2\boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} + 2\boldsymbol{\mu}_{x_f} \end{pmatrix} \\
&= \begin{pmatrix} 2\boldsymbol{\mu}_{f(x_m)} + 2\mathbf{B}_x \mathbf{P}_f^{-1} \mathbf{s}_f + \boldsymbol{\mu}_{x_m} + \mathbf{G}_{x,m} \mathbf{P}_m^{-1} \mathbf{s}_m \\ 2\boldsymbol{\mu}_{x_f} + 2\mathbf{G}_{x,f} \mathbf{P}_f^{-1} \mathbf{s}_f + \boldsymbol{\mu}_{m(x_f)} + \mathbf{B}_x^T \mathbf{P}_m^{-1} \mathbf{s}_m \end{pmatrix} - \begin{pmatrix} \boldsymbol{\mu}_{x_m} + 2\boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} + 2\boldsymbol{\mu}_{x_f} \end{pmatrix} \\
&= \begin{pmatrix} 2\mathbf{B}_x \mathbf{P}_f^{-1} \mathbf{s}_f + \mathbf{G}_{x,m} \mathbf{P}_m^{-1} \mathbf{s}_m \\ 2\mathbf{G}_{x,f} \mathbf{P}_f^{-1} \mathbf{s}_f + \mathbf{B}_x^T \mathbf{P}_m^{-1} \mathbf{s}_m \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} \mathbf{G}_{x,m} & \mathbf{B}_x \\ \mathbf{B}_x^T & \mathbf{G}_{x,f} \end{pmatrix} \begin{pmatrix} \mathbf{P}_m^{-1} \mathbf{s}_m \\ 2\mathbf{P}_f^{-1} \mathbf{s}_f \end{pmatrix}
\end{aligned} \tag{9.6}$$

Using the equilibrium condition (9.5) one finds from (9.6) that

$$\begin{aligned}
\frac{1}{3} \Delta \begin{pmatrix} \boldsymbol{\mu}_{x_m} + 2\boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_f)} + 2\boldsymbol{\mu}_{x_f} \end{pmatrix} &= \frac{1}{3} \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}} + 2\boldsymbol{\mu}_{f(x_{o,m})} - \boldsymbol{\mu}_{x_m} - 2\boldsymbol{\mu}_{f(x_m)} \\ \boldsymbol{\mu}_{m(x_{o,f})} + 2\boldsymbol{\mu}_{x_{o,f}} - \boldsymbol{\mu}_{m(x_f)} - 2\boldsymbol{\mu}_{x_f} \end{pmatrix} \\
&= \frac{1}{3} \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}} + 2(\boldsymbol{\mu}_{x_{o,m}} + \boldsymbol{\mu}_{f(x_m)} - \boldsymbol{\mu}_{x_m}) - \boldsymbol{\mu}_{x_m} - 2\boldsymbol{\mu}_{f(x_m)} \\ (\boldsymbol{\mu}_{x_{o,f}} + \boldsymbol{\mu}_{m(x_f)} - \boldsymbol{\mu}_{x_f}) + 2\boldsymbol{\mu}_{x_{o,f}} - \boldsymbol{\mu}_{m(x_f)} - 2\boldsymbol{\mu}_{x_f} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}} - \boldsymbol{\mu}_{x_m} \\ \boldsymbol{\mu}_{x_{o,f}} - \boldsymbol{\mu}_{x_f} \end{pmatrix} \\
&= \frac{1}{3} \begin{pmatrix} 2\mathbf{B}_x \mathbf{P}_f^{-1} \mathbf{s}_f + \mathbf{G}_{x,m} \mathbf{P}_m^{-1} \mathbf{s}_m \\ 2\mathbf{G}_{x,f} \mathbf{P}_f^{-1} \mathbf{s}_f + \mathbf{B}_x^T \mathbf{P}_m^{-1} \mathbf{s}_m \end{pmatrix} \\
&= \frac{1}{3} \begin{pmatrix} \mathbf{G}_{x,m} & 2\mathbf{B}_x \\ \mathbf{B}_x^T & 2\mathbf{G}_{x,f} \end{pmatrix} \begin{pmatrix} \mathbf{P}_m^{-1} \mathbf{s}_m \\ \mathbf{P}_f^{-1} \mathbf{s}_f \end{pmatrix}
\end{aligned} \tag{9.7}$$

this gives, finally,

$$\begin{aligned}
\begin{pmatrix} \Delta \boldsymbol{\mu}_{z_m} \\ \Delta \boldsymbol{\mu}_{z_f} \end{pmatrix} &= \begin{pmatrix} \boldsymbol{\mu}_{x_{o,m}} - \boldsymbol{\mu}_{x_m} \\ 2\boldsymbol{\mu}_{x_{o,f}} - 2\boldsymbol{\mu}_{x_f} \end{pmatrix} \\
&= \frac{1}{3} \begin{pmatrix} \mathbf{G}_{x,m} & 2\mathbf{B}_x \\ 2\mathbf{B}_x^T & 4\mathbf{G}_{x,f} \end{pmatrix} \begin{pmatrix} \mathbf{P}_m^{-1} \mathbf{s}_m \\ \mathbf{P}_f^{-1} \mathbf{s}_f \end{pmatrix}
\end{aligned} \tag{9.8}$$

## EPILOGUE

The equations presented, are based on a set of assumptions that is usually referred to as the *infinitesimal model*. It is assumed that a practically infinite number of loci contributes to the distribution of phenotypes and genetic variance. The infinitesimal model assures the Gaussian distributions we made use of. Now that genes are not virtual entities anymore, such as in the early days of population genetics, models need to be developed that account for the finite number of loci that make up most traits, and that tend to be flexible in the way genetic contributions to phenotype are specified. Such models are still in development, and there is much scope for relevant contributions. See Lande (1976) and, for instance, Chevalet (1988) and Hospital and Chevalet (1996) for models that develop approximations for traits determined by a finite number of loci. Walsh (2001) makes it clear that the approaches elaborated in this document don't all of a sudden become obsolete in the age of genomics.

## REFERENCES

Axler, S. (1997) *Linear Algebra Done Right*. Springer-Verlag.

Chevalet, C. (1988) Control of genetic drift in selected populations. In: Proc. 2nd Int. Conf. Quant. Genet. Raleigh NC (eds. Weir, B. S. , E. J. Eisen, M. M. Goodman and G. Namkoong), pp. 379-394. Sinauer Ass. Sunderland MA.

Horn, R. A. and C. R. Johnson (1985) *Matrix Analysis*. Cambridge University Press.

Hospital F., C. Chevalet (1996) Interactions of selection, linkage and drift in the dynamics of polygenic characters. *Genet. Res.* 67: 77-87.

Lande (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30: 314-334.

Lande, R. (1979) Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33: 402-416.

Lande, R. (1980) Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34: 292-305.

Lande, R and S. J. Arnold (1983) The measurement of selection on correlated characters. *Evolution* 37: 1210-1226.

Lynch, M. and B. Walsh (1998) *Genetics and Analysis of Quantitative Characters*. Sinauer Associates, Sunderland MA.

Lynch and Walsh (in press) Evolution and Selection of Quantitative Traits. Manuscript chapters available from [http://nitro.biosci.arizona.edu/zbook/volume\\_2/vol2.html](http://nitro.biosci.arizona.edu/zbook/volume_2/vol2.html)

Searle, S. R. (1971) Linear Models. John Wiley and Sons, Inc. New York.

Searle, S. R. (1982) Matrix Algebra Useful for Statistics. John Wiley and Sons, Inc., New York.

Tallis, G. M. (1987) Ancestral covariance and the Bulmer effect. *Theor. Appl. Genet.* 73: 815-820.

Walsh, B. (2001) Quantitative Genetics in the Age of Genomics. *Theor. Pop. Biol.* 59: 175-184.