Constraints on multivariate evolution

Bruce Walsh

Departments of Ecology & Evolutionary Biology, Animal Science, Biostatistics, Plant Science

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

-- Charles Darwin

Overview

- Review of univariate selection theory
- Multivariate traits: genetic correlations
- Multivariate traits: selection response
- Constraints in multivariate response
- Fundamental role of G
- Schuler's lines of Genetic least resistance
- How constrained is multivariate selection in nature?

Abundant Genetic Variation + Strong Selection = Multivariate Genetic Constraints: A Geometric View of Adaptation

Bruce Walsh[1,2] and Mark W. Blows[1]

[1] School of Biological Sciences, The University of Queensland, Brisbane, Qld 4072 Australia; email: m.blows@uq.edu.au

[2] Permanent address: Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721; email: jbwalsh@u.arizona.edu

Univariate (single-trait) selection

- What is the expected response to selection in a single trait?

- The key is to consider the slope of the parent-offspring regression (Galton)

- Here one plots the average trait value of a family's parents on the horizontal axis and the mean trait value in their offspring on the vertical axis

- The slope of this regression is h^2, the heritability of the trait

Parent-offspring regressions

Mean value of Trait in offspring

Trait in midparent
(the average of both parents)

Slope = h^2
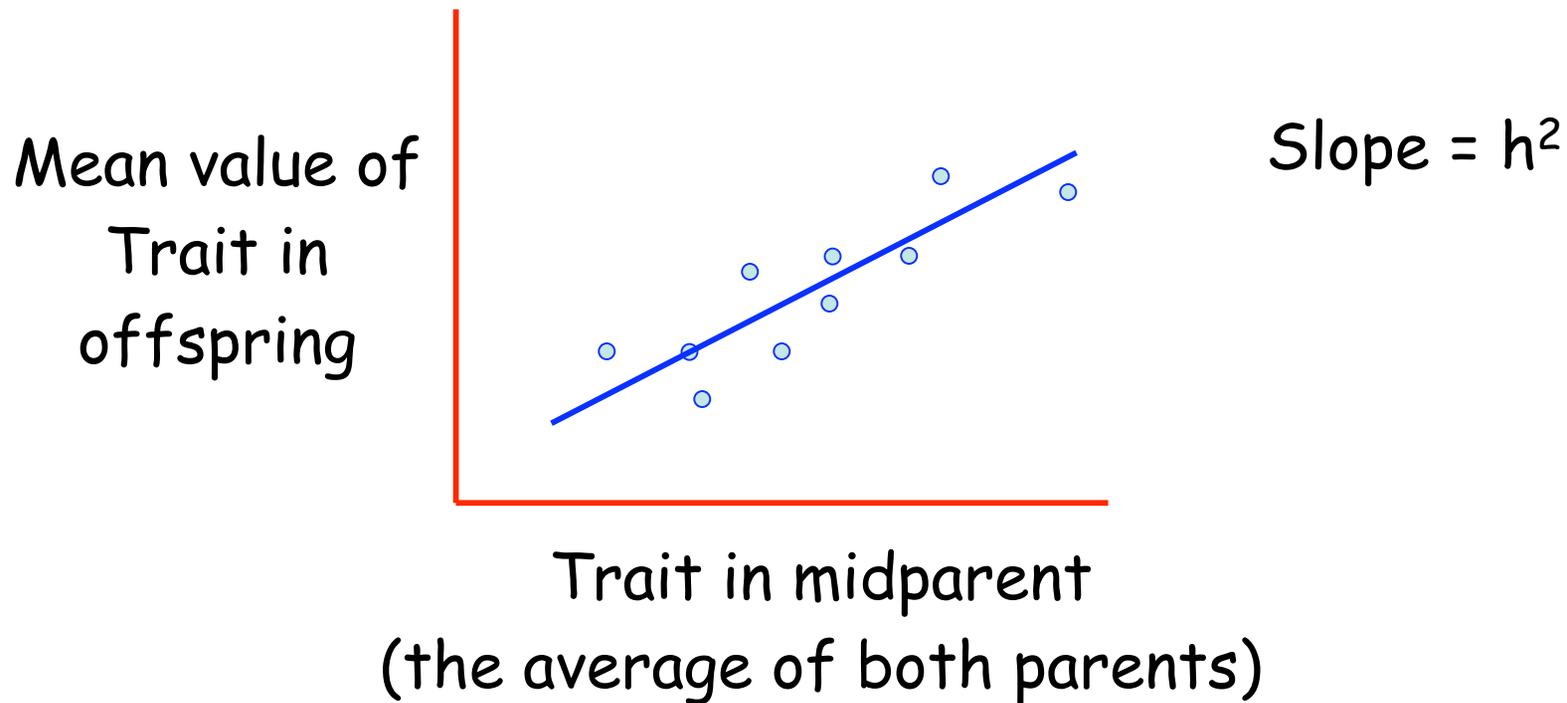
If slope is close to one, offspring closely resemble their parents. If slope is close to zero, offspring do not resemble their parents. Looks at the heritable component of variation.

The Breeder's Equation

- The expected response R (change in mean in the offspring, the between-generation change) is related to the selection differential S (the within-generation change in trait mean from selection) by the breeder's equation

- $R = h^2 S$

- Smaller the heritability value h^2 for a trait, the slower the response to selection

- Most traits have h^2 values of around 0.2 to 0.5

Natural selection is not Evolution (Fisher 1930)

- Natural selection is S -- within-generation change from selection

- Evolution is R -- the between-generation change, the response to selection

- The amount of usable genetic variation constrains the amount of evolution (R) given the amount of selection (S)

Heritability and additive genetic variance

- h^2 is the ratio of two variances,
  - Var(A), the additive genetic variance (the variance of that part of trait value that a parent passes along to its offspring)
  - Var(P), the total phenotypic (trait) variance
  - h^2 = Var(A)/Var(P)
- With constant selection on a trait, fixation of favorable genes drives Var(A) to near zero. The input of new additive variation (via mutation and/or migration) can partly counter this decline
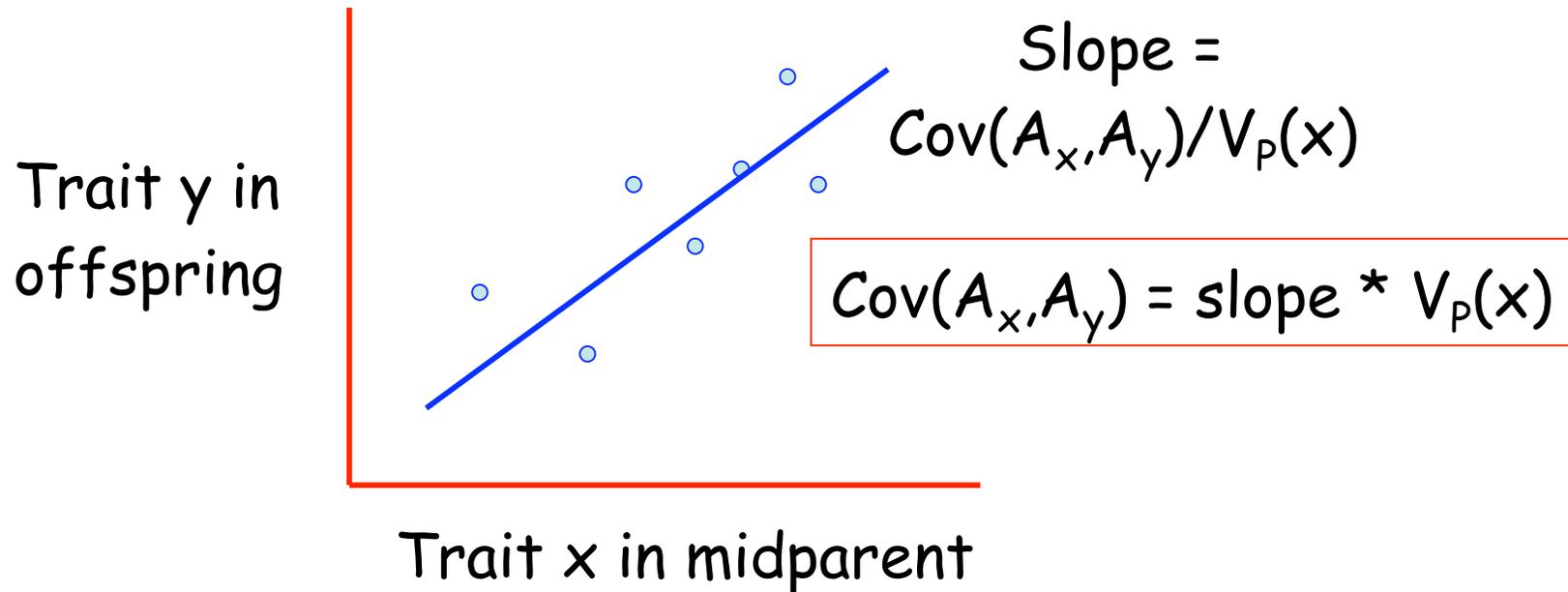
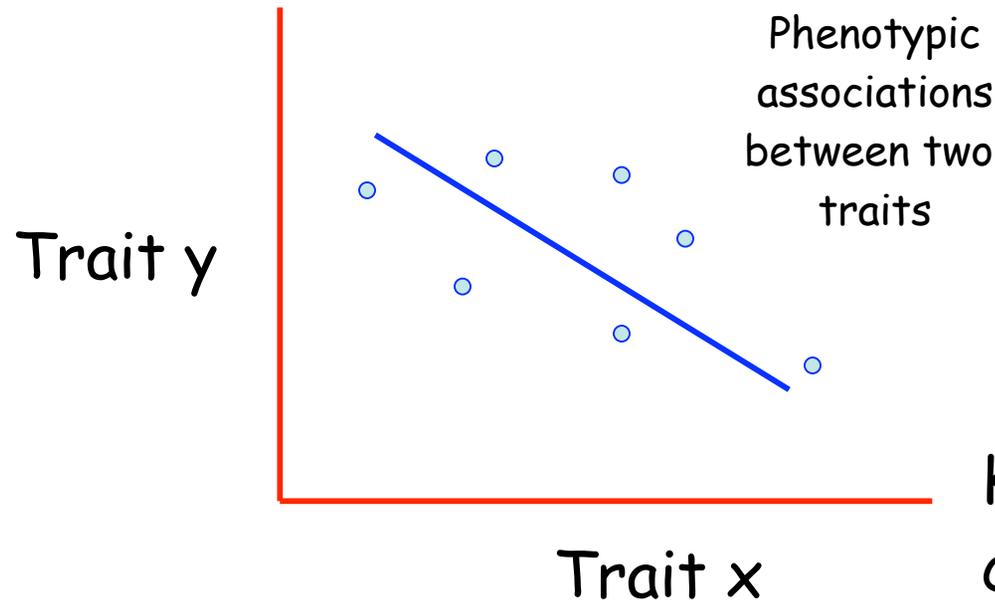# Response to artificial vs. natural selection

- Essentially all traits show response to strong artificial selection

- However, many traits with non-zero heritabilities under apparently constant selection in natural populations show stasis --- lack of selection response.

- One potential reason is that if selection is multivariate, using **univariate models gives a VERY misleading picture** (as we will see).

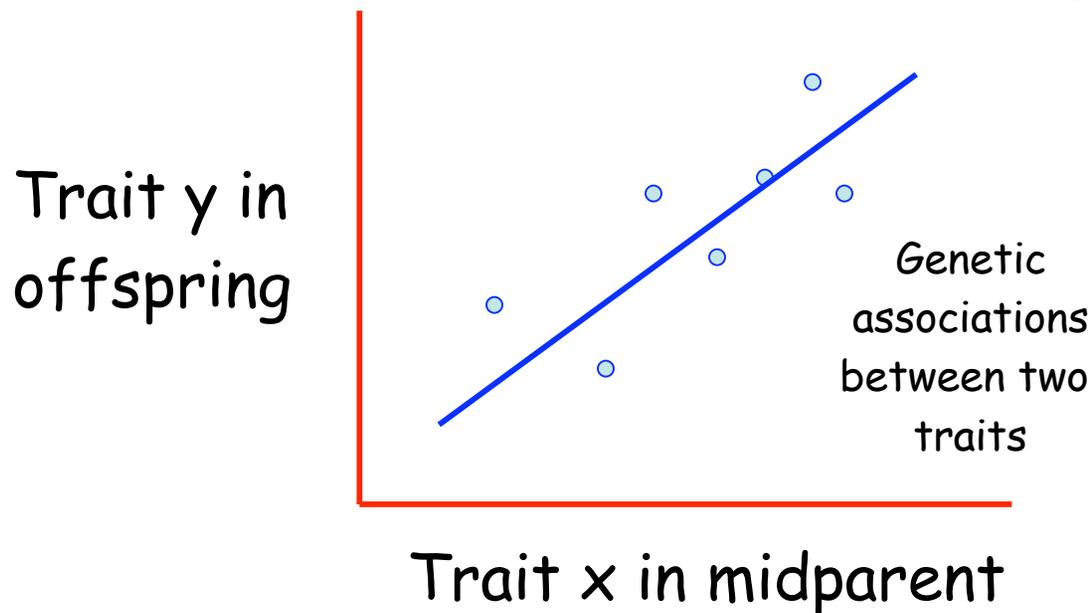# Multivariate traits: Genetic vs. phenotypic correlations

- Within an individual, multiple traits can show phenotypic correlations

- However, what matters for selection response is genetic correlations

- Genetic correlations (covariances) can be estimated with a simple modification of the parent-offspring regression

We can estimate $Cov(A_x, A_y)$, the genetic covariance for traits x and y, (the covariance of their breeding values A) by the regression of trait x in the parent and trait y in the offspring

Trait y in offspring

Slope = $Cov(A_x, A_y)/V_P(x)$

$Cov(A_x, A_y) = $ slope $* V_P(x)$

Trait x in midparent

Trait y

Phenotypic associations between two traits

Trait x

KEY: Phenotypic associations can be very poor predictors of genetic associations

Trait y in offspring

Genetic associations between two traits

Trait x in midparent

The multivariate breeder's equation

The selection response for multiple traits is handled by using matrices. Let R be the vector of responses in the means of interest, S the vector of their selection differentials, and G and P genetic and phenotypic covariance matrices. Then

$$R = G\, P^{-1}\, S$$

$$R = h^2 S = (V_A/V_P)\, S$$

Natural parallels with univariate Breeder's equation

Multivariate trait selection

$$\mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix}$$

Vector of selection differentials   Vector of responses

P = phenotypic covariance matrix. P_{ij} = Cov(P_i,P_j)

G = Genetic covariance matrix. G_{ij} = Cov(A_i,A_j) = V_A(Pi,pj)

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

Selection gradients and fitness regressions

- $\mathbf{P^{-1}S}$ is usually written as β, the vector of selection gradients

- Hence, we can write the selection response as $\mathbf{R = G}$ β

- β **is the direction that selection is trying to move** the trait means to maximize the gain in mean fitness

- Why? The regression of relative fitness w on the vector of trait values $z_1, .., z_n$ is given by $w = 1 + \Sigma \, \beta_i(z_i - \mu_i)$

Response in trait j under the multivariate
Breeder's equation

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

Response under
univariate
breeder's eq
(direct response)

Response from
genetically
correlated traits
(correlated response)

Multivariate Constraints to Response

Consider the following G and β:

$$\mathbf{G} = \begin{pmatrix} 10 & 20 \\ 20 & 40 \end{pmatrix}, \qquad \boldsymbol{\beta} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$
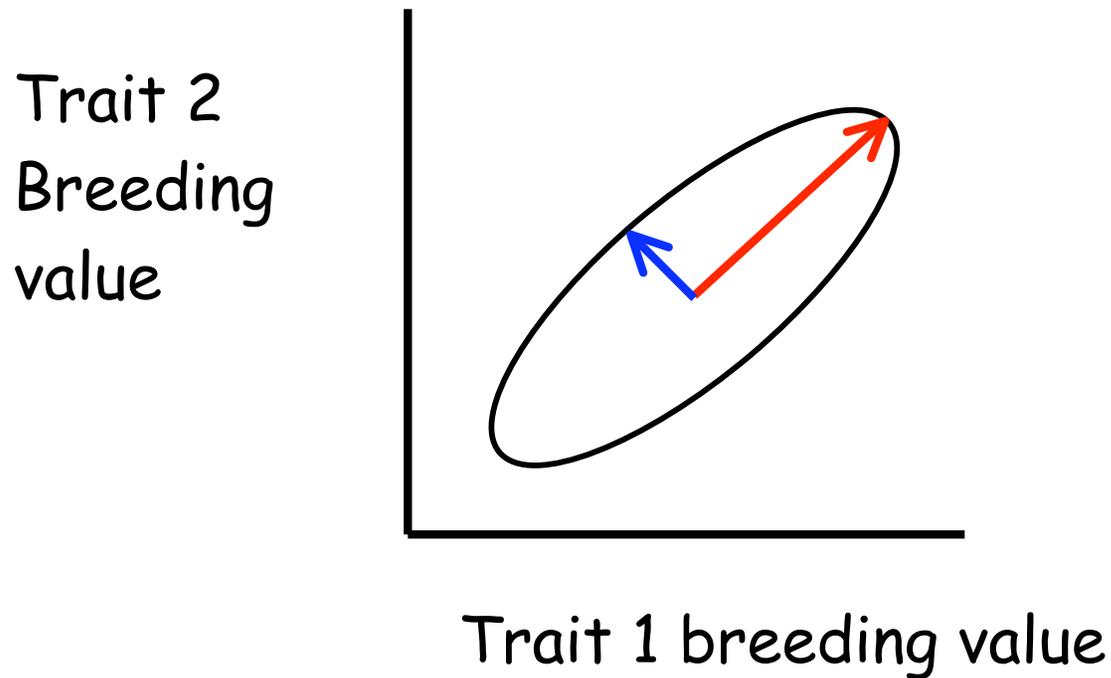
Taken one trait at a time, we might expect $R_i = G_{ii}\beta_i$

giving $R_1 = 20$, $R_2 = -40$.

What is the actual response?

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

Hence, despite considerable variation in both traits, there is no response to selection. Why?
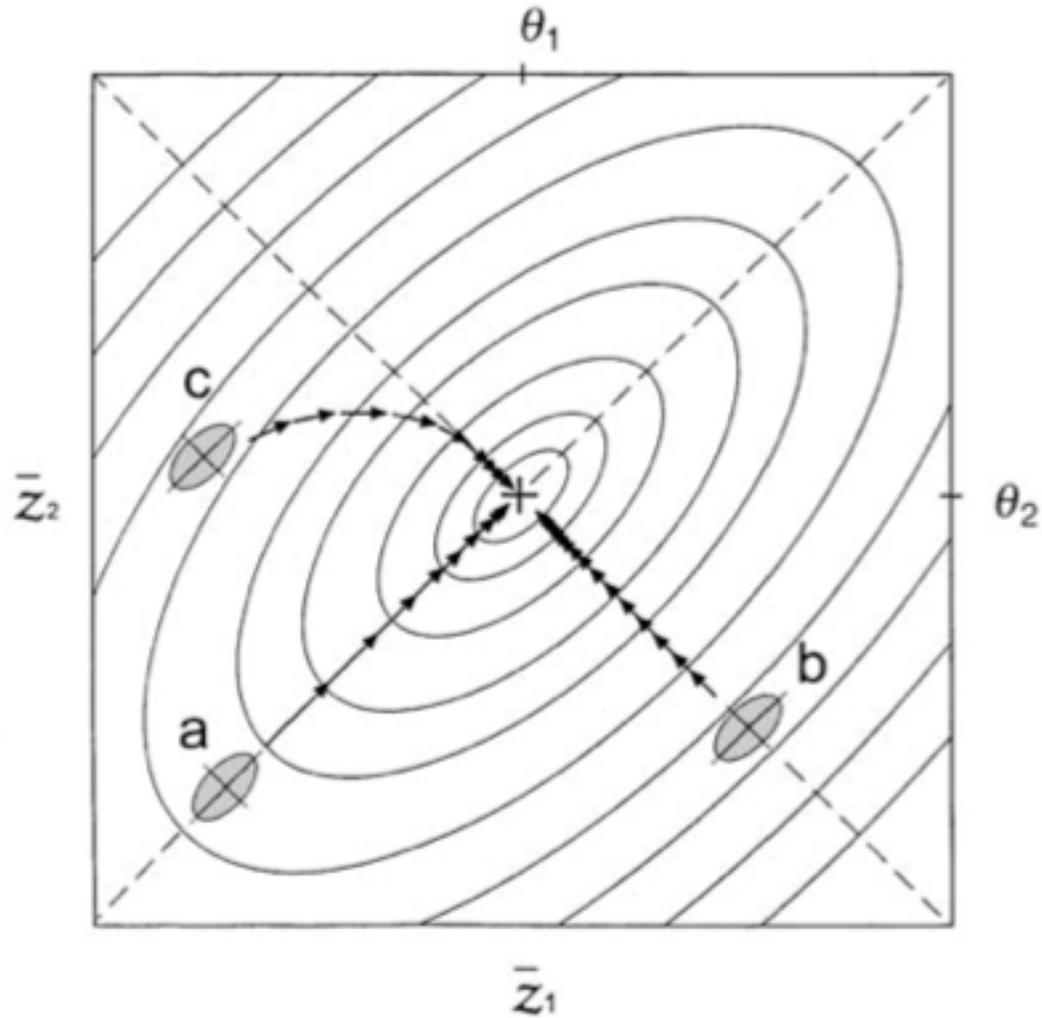
Trait 2 Breeding value

Trait 1 breeding value

Both traits show considerable variation in breeding value (ability to response to selection)

However, much more variation along the red direction (vector) than along the blue direction (vector)

Examples of increased constraints (along the blue direction)
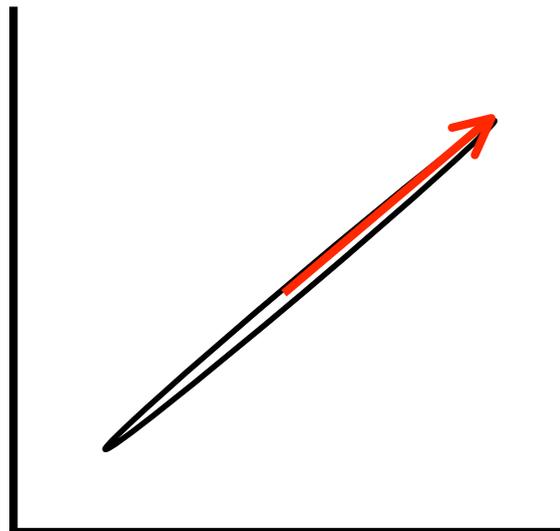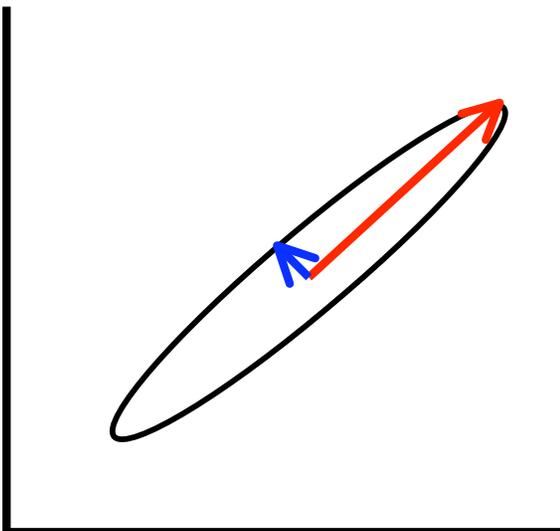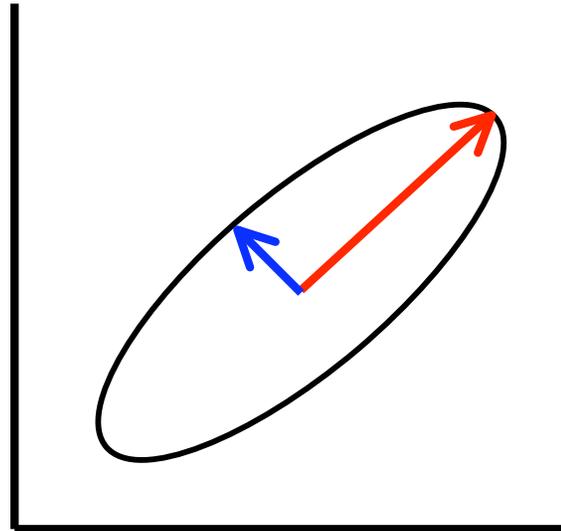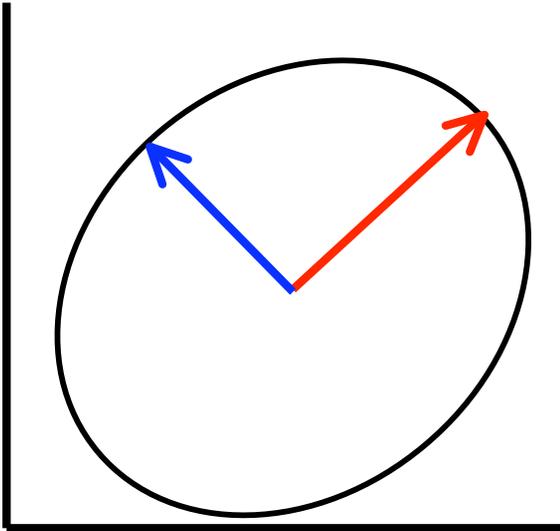
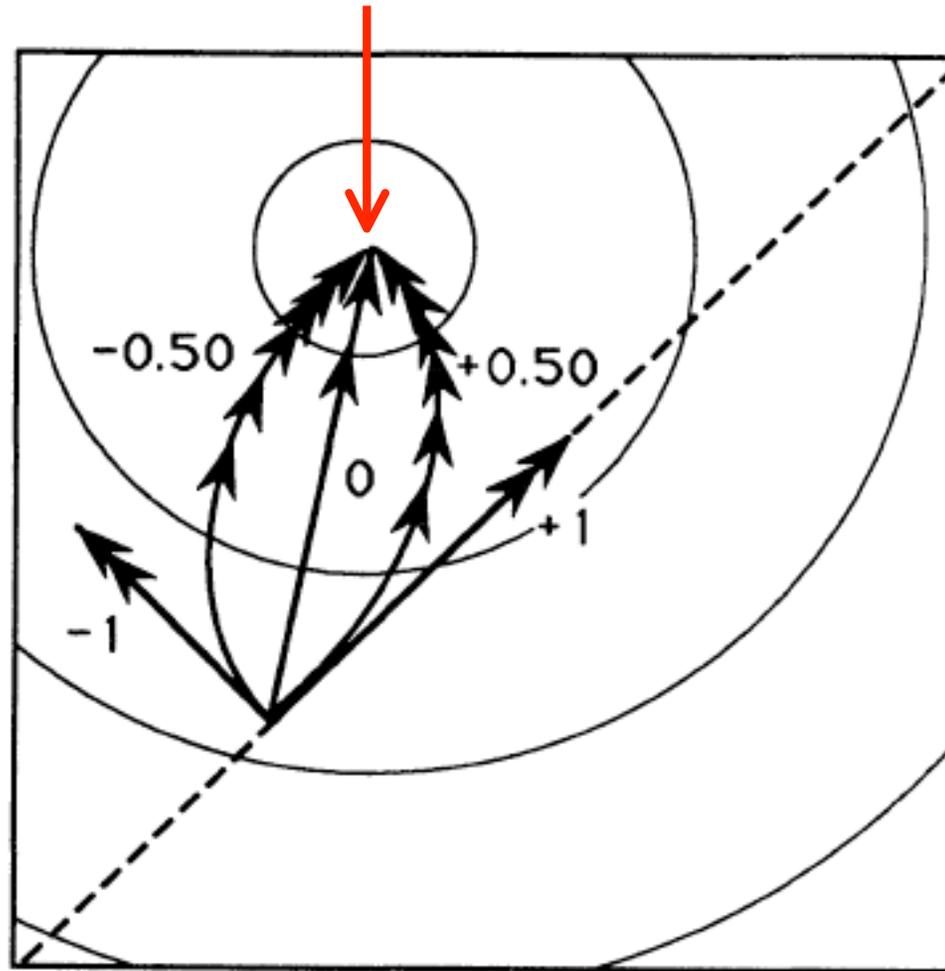# Trait evolution on a fitness surface

Fitness optimum

Trait 2

Trait 1
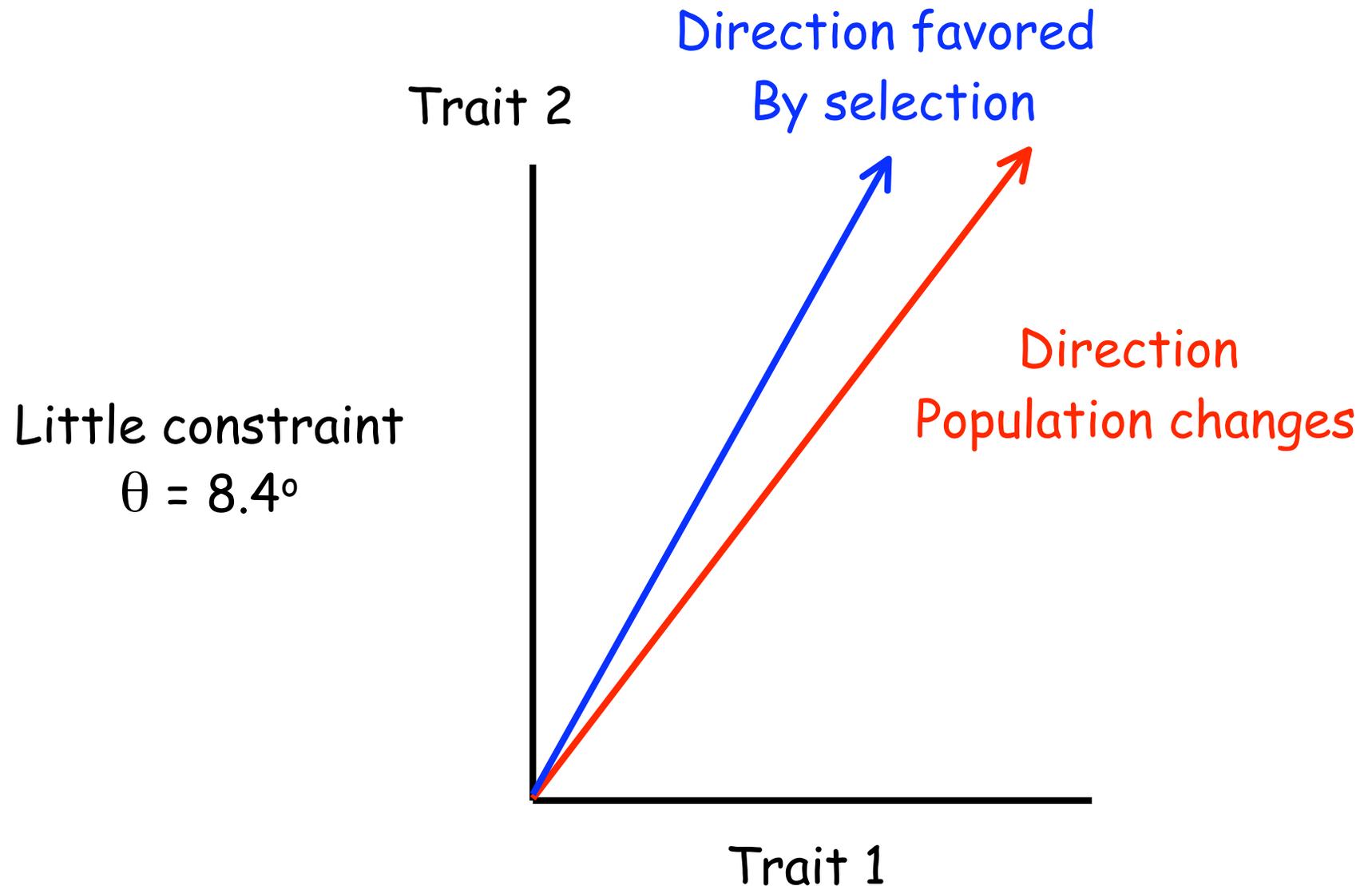
-0.50   +0.50

0   +1

-1

Constraints imposed by the geometry of G

- Selection is trying to move the vector of traits means in the direction given by β

- The response vector is $R = G\beta$, so that the genetic covariance matrix G rotates and scales the response away from the direction most favored by selection

- One measure of constraint is the angle θ between β and R

$$\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \mathbf{G}_1 = \begin{pmatrix} 6 & 4 \\ 4 & 8 \end{pmatrix} \qquad \mathbf{R}_1 = \mathbf{G}_1 \boldsymbol{\beta} = \begin{pmatrix} 6 & 4 \\ 4 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 14 \\ 20 \end{pmatrix}$$

Direction favored
By selection

Trait 2

Direction
Population changes

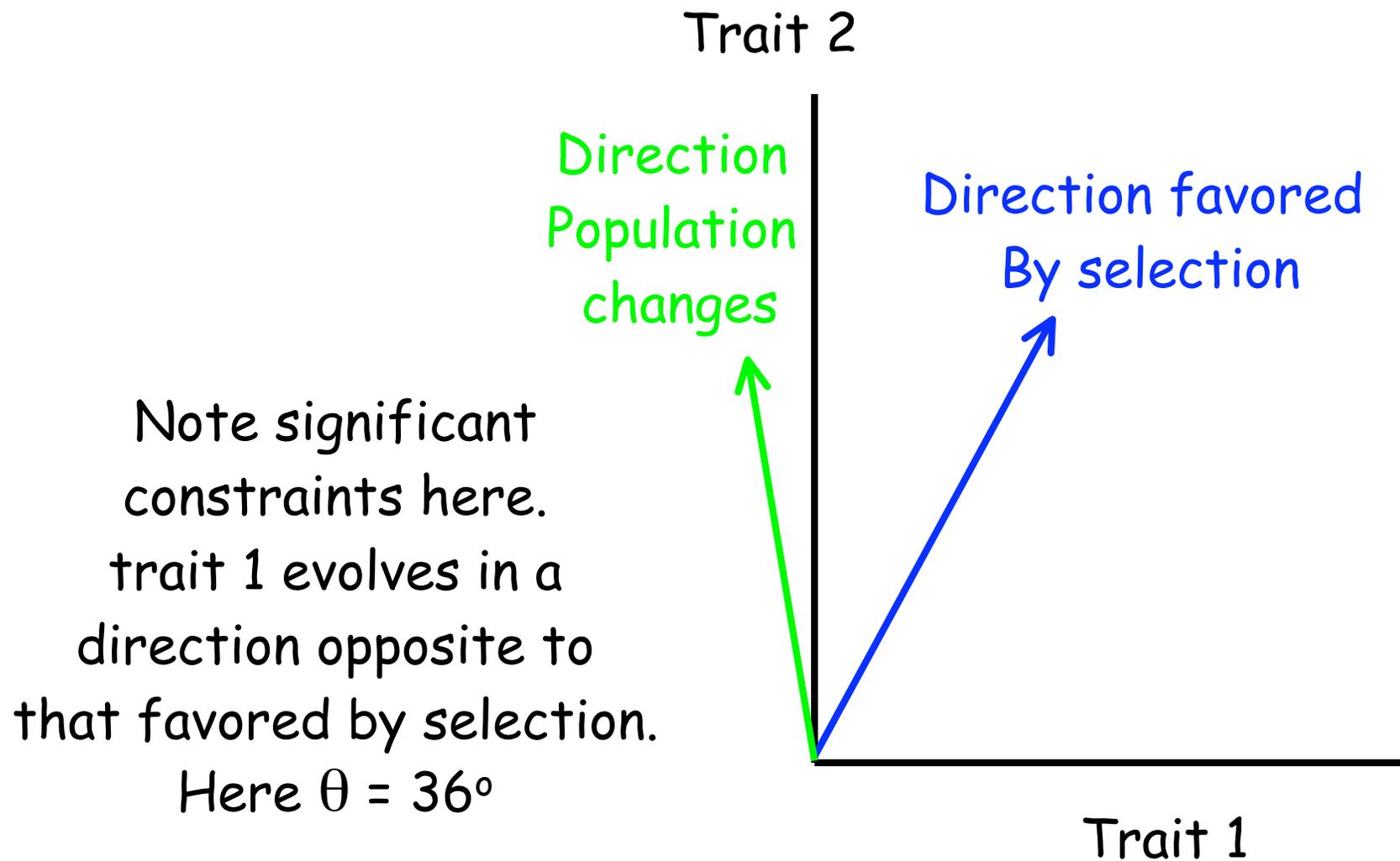Little constraint
$\theta = 8.4^{\circ}$

Trait 1

$$\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \mathbf{G}_2 = \begin{pmatrix} 6 & -4 \\ -4 & 8 \end{pmatrix} \qquad \mathbf{R}_2 = \mathbf{G}_2 \boldsymbol{\beta} = \begin{pmatrix} 6 & -4 \\ -4 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 12 \end{pmatrix}$$

Trait 2

Direction Population changes

Direction favored By selection

Note significant constraints here. trait 1 evolves in a direction opposite to that favored by selection. Here θ = 36°

Trait 1

The geometry of a matrix

- Multiplying a vector by a matrix results in a rotation and scaling of the original vector to obtain a new one

- The eigenvectors of G determine the axes of variation (the rotation)

- The eigenvector associated with each eigenvalue is the amount of variation along that direction (the scaling)

Eigenvalue (scaling)
for vector e_i

Response along vector e_i

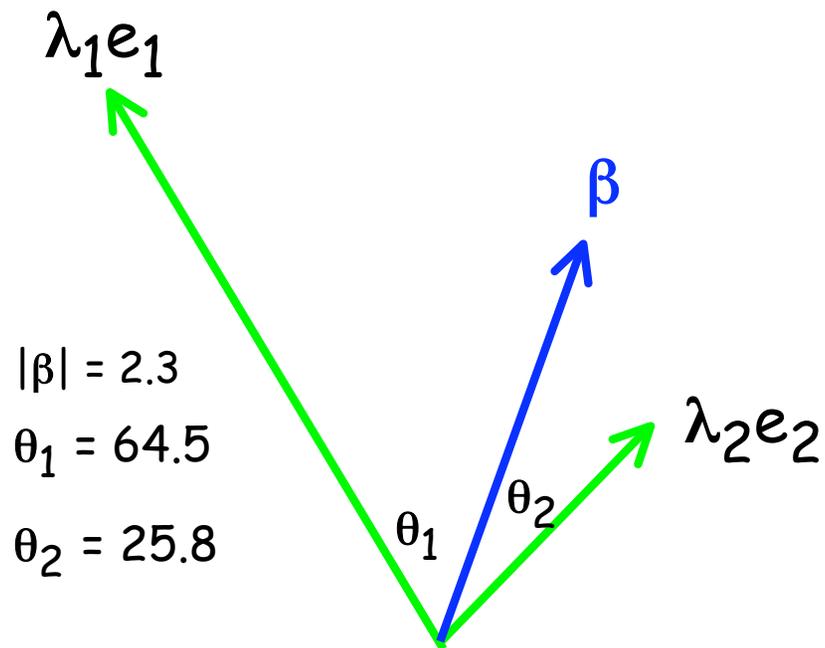$$\mathbf{R} = \sum_{i=1}^{n} \lambda_i \mathrm{Proj}(\boldsymbol{\beta} \, \mathrm{on} \, \mathbf{e}_i)$$

$$= \sum_{i=1}^{n} \lambda_i \|\boldsymbol{\beta}\| \cos(\theta_i) \mathbf{e}_i$$

Length of β
(strength of
Total selection)

Angle between e_i and β

$$\mathbf{G}_2 = \begin{pmatrix} 6 & -4 \\ -4 & 8 \end{pmatrix} \quad \boldsymbol{\beta} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$R = 11*e_1 + 6*e_2$

$\lambda_1 e_1$

β

$|\beta| = 2.3$

$\theta_1 = 64.5$

$\theta_2 = 25.8$

$\lambda_2 e_2$

θ_1

θ_2

$11\ e_1$

β

$6\ e_1$

$\lambda_1\ |\beta|\ \cos(\theta_1)\ e_1 = 11\ e_1$

$\lambda_2\ |\beta|\ \cos(\theta_2)\ e_2 = 6\ e_2$
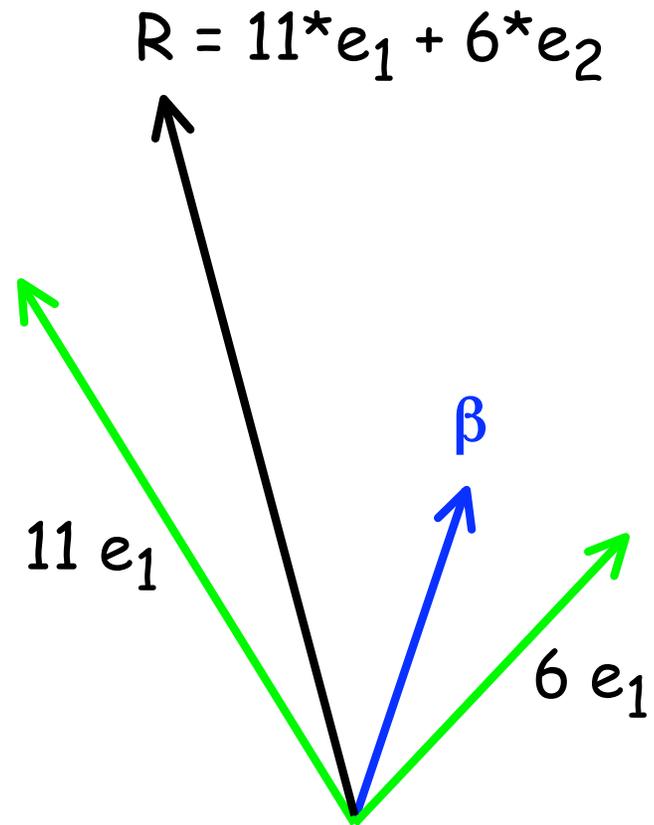
Theory vs. reality

- Theory: Genetic variation in individual traits not sufficient to ensure selection response

  - If little genetic variation in the direction of β, little expected response

- How likely is β to be near an eigenvector of G associated with a small eigenvalue?

- This is an empirical question

  - While such an association might seen unlikely, recall that constant selection erodes away genetic variation along that direction

Current data

- Many well-studies cases of natural populations of vertebrates with strong selection on a heritable trait, but no observed response (stasis)

  - Potentially due to considering a multivariate problem in a univariate setting (ignoring other traits under selection can give very misleading results)

- What is known about the geometry of G in cases where β is known and assumed relatively constant?

Geometry of G and β

- Few studies where both G and β are estimated for a large (>5) number of traits
- Blows et al. (2004) looked at 8 cuticular hydrocarbons in Drosophila which are used in mate choice by females.
  - Mate-choice experiments allows β to be estimated in the lab with high precision
  - The first two eigenvalues of G account for 78% of the total variation
  - The angle θ between β and these two eigenvectors are 82 and 99 degrees
 - Hence, most usable variation in directions other than β

G and the direction of Evolution

- Schluter (1996) examined several small data sets of divergent vertebrate populations (fishes, mammals) with estimated divergence times up to 4 MY

- He found that the divergence in a vector of morphological traits tended to be very close to the first eigenvector e_1 of **G** (the axis of most variation)

  – Schluter looked at the angle θ between the vector of divergence (changes in means) vs. e_1

  – Schluter denoted e_1 by g_{max}

Lines of least genetic resistance

- Schluter observed that
  - Smallest values of θ occurred between most recently diverged populations
  - The greater the value of θ, the smaller the amount of divergence
- Schluter suggests that populations tend to evolve along the lines of least genetic resistance (directions with the largest genetic variation).
- Lack of variation may constrain response in other directions
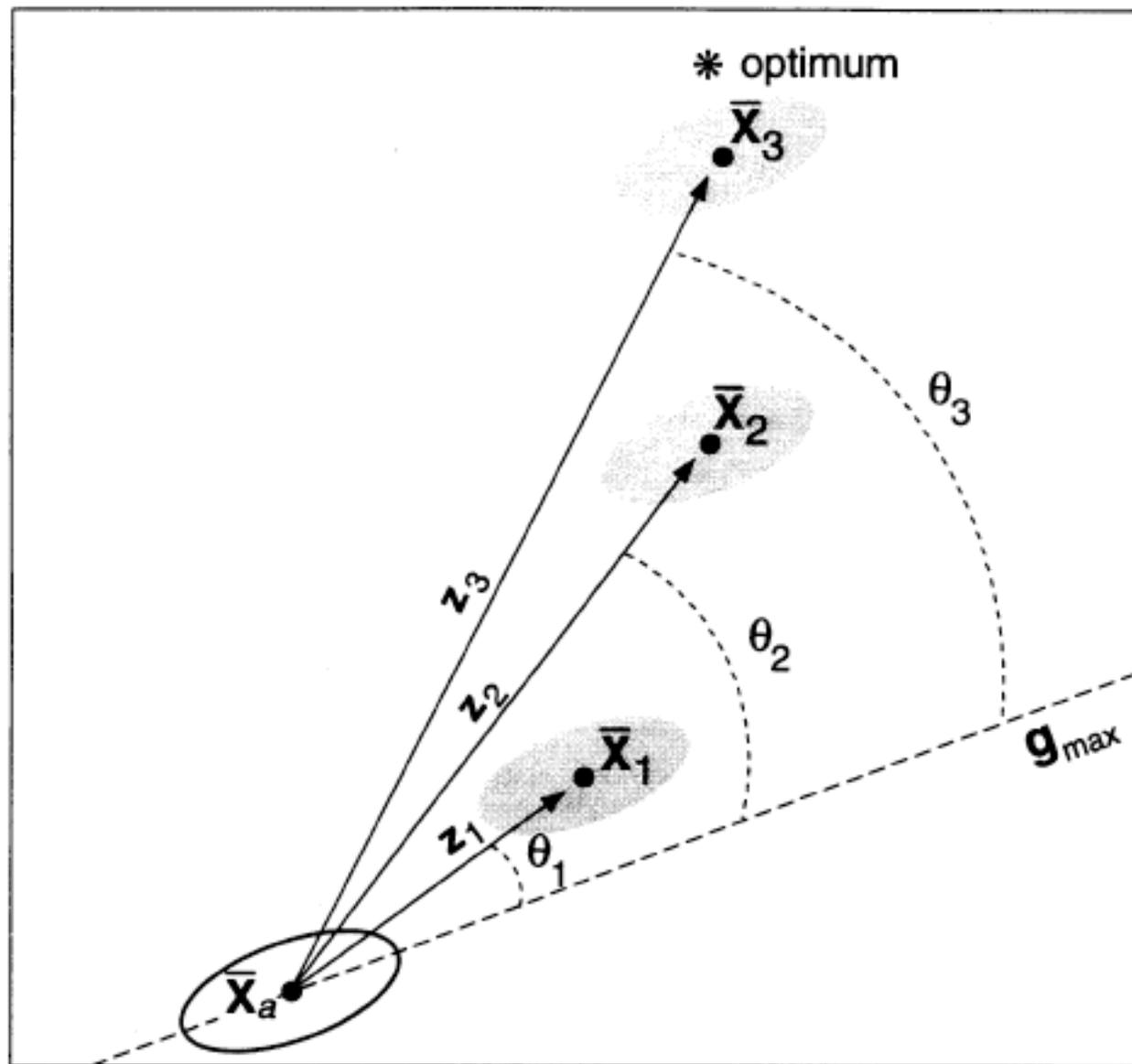
$$\mathbf{R} = \sum_{i=1}^{n} \lambda_i ||\boldsymbol{\beta}|| \cos(\theta_i) \mathbf{e}_i$$

If λ_1 is sufficiently large, θ must be very close to 90 degrees, else $\lambda_1 \cos(\theta_1)$ will dominate other $\lambda_i \cos(\theta_i)$ terms

In such cases, most of the response to selection will be along the direction $\mathbf{e}_1$.

However, it is also the case that under strict genetic drift, most of the divergence also occurs along the axis with the most variation

Under drift, expected vector of divergence is MVN with mean vector 0 and covariance (dispersion) matrix (t/2Ne)$\boldsymbol{G}$

Support (and counterexamples): McGuigan et al. 2005

Examined two species of Australian rainbow fish in a nested design of stream and lake populations for both species

- Divergence between species, as well as divergence between replicate population of the same species in the same hydrodynamic environments (lake vs. stream) followed g_{max}
  - Between-species and within species divergence in the same hydrodynamic environments consistent with drift
- However, populations of the same species in different hydrodynamic environments were at directions quite removed from g_{max}, as well as the other major eigenvectors of G
  - Within-species adaptation to different hydrodynamic environments occurred against a gradient of little variation

Conclusions

- The geometry of G constrains multivariate evolution.

- Treating evolution as a series of univariate responses is highly misleading.

- Recent evidence suggests that multivariate evolution might be more constrained that previously thought.

- How important constraints actually are in natural populations remains a significant open question.