Multivariate Brownian motion and quantitative genetics

Friday Harbor Laboratories, 8 June 2017

Joe Felsenstein

Covarying character change along a lineage

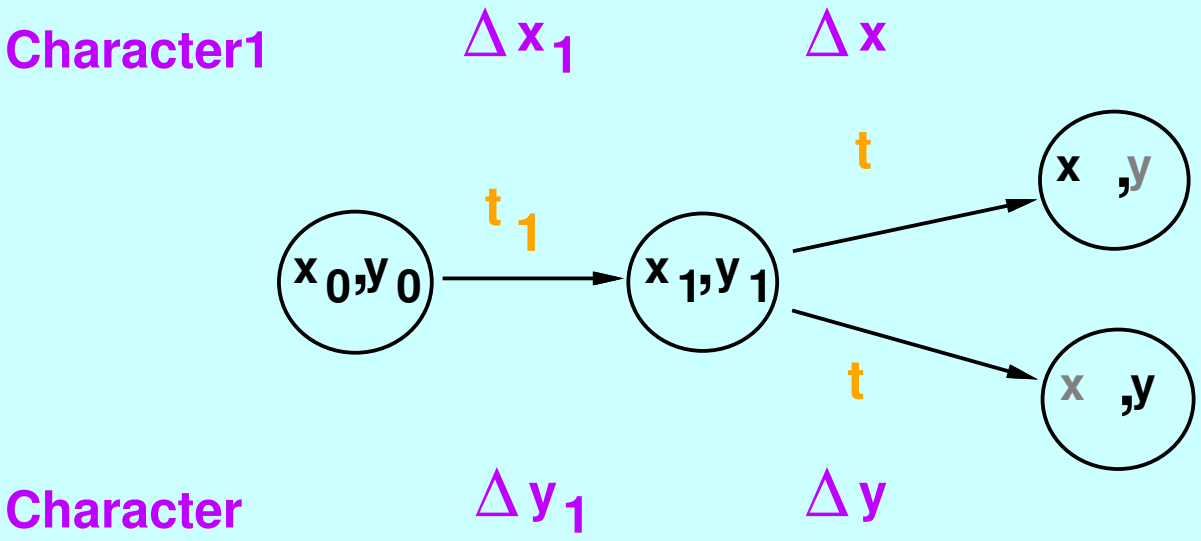
What is the distribution of changes in multiple characters (say p of them) along a lineage? Simply the appropriate multiple of the infinitesimal rate of change per unit branch length.

If a set of characters $\mathbf{x}$, changes under covarying Brownian motion, in time t (or a pseudo-time branch-length t) the change will be distributed as

$$\mathbf{\Delta x} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}t),$$

(where $\mathbf{V}$ is the covariance matrix of the infinitesimal change of the Brownian Motion).

Joint distribution for multiple species, characters



Consider change of two characters, each assessed in a different species. Say character x and character y, the first measured in species 2, the second in species 3. The result will give us the pattern for any two characters measured in any two species.

Seeing that covariances are zero in different branches ...

$$\mathrm{Cov}[\Delta x_1 + \Delta x_2, \;\; \Delta y_1 + \Delta y_3]$$

Given that changes in different branches are independent (whether changes of the same character or of different characters), the only nonzero covariance is between Δx_1 and Δy_1.

$$\mathrm{Cov}[\Delta x_1 + \Delta x_2, \;\; \Delta y_1 + \Delta y_3]$$

$$= \;\; \mathrm{Cov}[\Delta x_1, \;\; \Delta y_1]$$

So the covariance of different characters in different species is the product of the shared evolution to their common ancestor by the (infinitesimal) covariance of the character change per unit branch length.

Joint distribution for many species, many characters

The upshot is that if x_{ik} is character k in species i, and $x_{j\ell}$ is character ℓ in species j, the covariance between them is

$$\mathrm{Cov}[x_{ik}, x_{j\ell}] \;=\; t_{ij}\; v_{k\ell}$$

where t_{ij} is the time (branch length) to the latest common ancestor of species i and species j. $\mathbf{V}$ is the covariance matrix of evolutionary change for the characters.

A stacked vector with a partitioned covariance matrix

We have n species and p characters measured in each. Taking the vector of the p characters for species 1, the vector for species 2, and so on and stacking one on top of the other with the species in order, we get a column vector (shown here transposed).

$$\mathbf{x}^{\mathsf{T}} = (x_{11}, x_{12}, \ldots, x_{1p} \,|\, x_{21}, x_{22}, \ldots, x_{2p} \,|\, \ldots \,|\, x_{n1}, x_{n2}, \ldots, x_{np}),$$

The matrix of covariances of these will be partitioned into n groups of rows and n groups of columns.

For each box in the partitioned matrix, we will get a multiple of the infinitesimal covariance of characters, the multiple being the branch length up to the shared common ancestor of those two species.

The Kronecker product, mostly a bookkeeping device

This is the Kronecker product of $\mathbf{T}$ and $\mathbf{V}$, which is simply the partitioned matrix

$$\mathbf{T} \otimes \mathbf{V} = \begin{bmatrix} t_{11}\mathbf{V} & t_{12}\mathbf{V} & \dots & t_{1p}\mathbf{V} \\ t_{21}\mathbf{V} & t_{22}\mathbf{V} & \dots & t_{2p}\mathbf{V} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1}\mathbf{V} & t_{n2}\mathbf{V} & \dots & t_{np}\mathbf{V} \end{bmatrix}$$

so that this is an $\mathsf{np} \times \mathsf{np}$ matrix.

The algebra

If $\mathbf{T}$ is the covariances of n tips on the tree, and $\mathbf{V}$ is the (unknown) covariances of the Brownian motion of the p characters, the log-likelihood of a set of characters (stacked as a vector) $\mathbf{x}$ is

$$\ln \mathsf{L} \;=\; -(\mathsf{np}/2)\ln(2\pi) - (1/2)\ln|\mathbf{T}\otimes\mathbf{V}| - (1/2)(\mathbf{x}-\boldsymbol{\mu})^{\mathsf{t}}(\mathbf{T}\otimes\mathbf{V})^{-1}(\mathbf{x}-\boldsymbol{\mu})$$

If $\mathbf{C}$ is an $(\mathsf{n}-1)\times\mathsf{n}$ set of contrasts, each orthogonal to the grand mean, we have seen that we can choose them so that each contrast has mean 0 and variance 1. Then $\mathbf{CTC}^t$ is an $\mathsf{n}-1$-dimensional identity matrix.

Taking the density of the transformed data $\mathbf{y}=\mathbf{Cx}$ (and not forgetting the term for the Jacobian of the transformation), this has expectation vector $\mathbf{0}$ so that the density is

$$\mathsf{L} \;=\; \mathsf{K} - (1/2)\ln|\mathbf{I}_{\mathsf{n}-1}\otimes\mathbf{V}| - (1/2)\mathbf{y}^{\mathsf{t}}(\mathbf{I}_{(\mathsf{n}-1)}\otimes\mathbf{V})^{-1}\mathbf{y} \;-\; \frac{1}{2}\sum_{\mathsf{i}=1}^{\mathsf{n}-1}\ln\left(\mathsf{v}_1^{(\mathsf{i})}+\mathsf{v}_2^{(\mathsf{i})}\right)$$

(where K collects the constant stuff and the $\mathsf{v}_1^{(\mathsf{i})}+\mathsf{v}_2^{(\mathsf{i})}$ is the variance of the ith contrast computed from $\mathbf{T}$, not from $\mathbf{V}$).

... simplifying ...

This can also be expressed as

$$\ln \mathsf{L} \;=\; \mathsf{K} - \frac{(\mathsf{n}-1)}{2}\ln|\mathbf{V}| - (1/2)\mathrm{tr}\left(\mathbf{S}\mathbf{V}^{-1}\right)) - \frac{1}{2}\sum_{\mathsf{i}=1}^{\mathsf{n}-1}\ln\left(\mathsf{v}_1^{(\mathsf{i})} + \mathsf{v}_2^{(\mathsf{i})}\right)$$

where

$$\mathbf{S} \;=\; \sum_{\mathsf{i}} \mathbf{y}^{(\mathsf{i})}\left(\mathbf{y}^{(\mathsf{i})}\right)^{\mathsf{t}}$$

is the $\mathsf{p} \times \mathsf{p}$ sum of squares matrix of characters across contrasts. Inferring the Brownian motion phylogenetic covariances by maximum likelihood we find that

$$\widehat{\mathbf{V}} \;=\; \mathbf{S}/(\mathsf{n}-1)$$

which leads to

$$\ln \mathsf{L} \;=\; \mathsf{K}' \;-\; \frac{(\mathsf{n}-1)}{2}\ln|\widehat{\mathbf{V}}|$$

where K' includes the $\ln\left(\mathsf{v}_1^{(\mathsf{i})} + \mathsf{v}_2^{(\mathsf{i})}\right)$ penalty terms, which don't change if the tree stays the same.

What causes change in quantitative characters?

For neutral mutation and genetic drift, can show that for a quantitative character with additive genetic variance V_A and population size N the genetic (additive) value of the population mean is:

$$\mathrm{Var}(\Delta \bar{g}) = V_A/N$$

If mutation and drift are at equilibrium:

$$E\left[V_A^{(t+1)}\right] = V_A^{(t)}\left(1 - \frac{1}{2N}\right) + V_M$$

In neutral traits additive genetic variance rules

so that

$$\mathrm{E}\left[\mathrm{V_A}\right] \;=\; 2\mathrm{NV_M}$$

whereby

$$\mathrm{Var}[\Delta\bar{\mathrm{g}}] \;=\; \left(2\mathrm{NV_M}\right)/\mathrm{N} \;=\; 2\mathrm{V_M}$$

an analog of Kimura's result for neutral mutation.

Thus to transform characters to independent Brownian motions of equal evolutionary variance, we could use the additive genetic variance $\mathrm{V_A}$.

With multiple characters ...

There is a precise analogue of this for multiple characters:

$$\mathrm{E}\left[\mathbf{A}^{(\mathrm{t}+1)}\right] \;=\; \mathbf{A}^{(\mathrm{t})}\left(1-\frac{1}{2\mathrm{N}}\right)+\mathbf{M}$$

where $\mathbf{A}$ is the additive genetic covariances, and $\mathbf{M}$ is the covariance matrix of pleiotropic effects of mutation.

$$\mathrm{E}\left[\mathbf{A}\right] \;=\; 2\mathrm{N}\,\mathbf{M}$$

and

$$\mathrm{Var}[\boldsymbol{\Delta}\bar{\mathbf{g}}] \;=\; \left(2\mathrm{N}\mathbf{M}\right)/\mathrm{N} \;=\; 2\mathbf{M}$$

so as long as mutations cause expected change zero (i.e. they are not near some biological limit), the effect of genetic drift is that the mean phenotype wanders according to the mutational covariances.

With selection ... life is harder

There is the "Breeder's Equation" of Wright and Fisher (1920's)

$$\Delta z \;=\; h^2 S$$

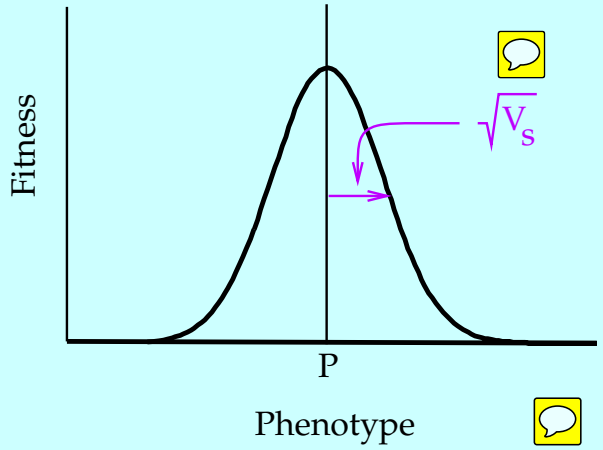
and Russ Lande's (1976) recasting of that in terms of slopes of mean fitness surfaces:

$$S \;=\; V_P \, \frac{d \log(\bar{w})}{d\bar{x}}$$

$$\Delta z \;=\; (V_A/V_P)\, V_P \, \frac{d \log(\bar{w})}{d\bar{x}} \;=\; V_A \, \frac{d \log(\bar{w})}{d\bar{x}}$$

Note – it's heritability times the slope of log of *mean* fitness with respect to *mean* phenotype. There is an exact multivariate analog of this equation.

Selection towards an optimum



If fitness as a function of phenotype is:

$$w(x) = \exp\left[-\frac{(x-p)^2}{2V_s}\right]$$

Then after some completing of squares and integrating, the change of mean phenotype “chases” the optimum:

$$m' - m = \frac{V_A}{V_s + V_P}(p - m)$$

(There is an exact matrix analog of this for multiple characters).

Sources of evolutionary correlation among characters

Variation (and covariation) in change of characters occurs for two reasons:

- **Genetic covariances**. (the same loci affect two or more traits)

Sources of evolutionary correlation among characters

Variation (and covariation) in change of characters occurs for two reasons:

- **Genetic covariances**. (the same loci affect two or more traits) 
- **Selective covariances** (Tedin, 1926; Stebbins 1950). The same environmental conditions select changes in two or more traits – even though they may have no genetic covariance.

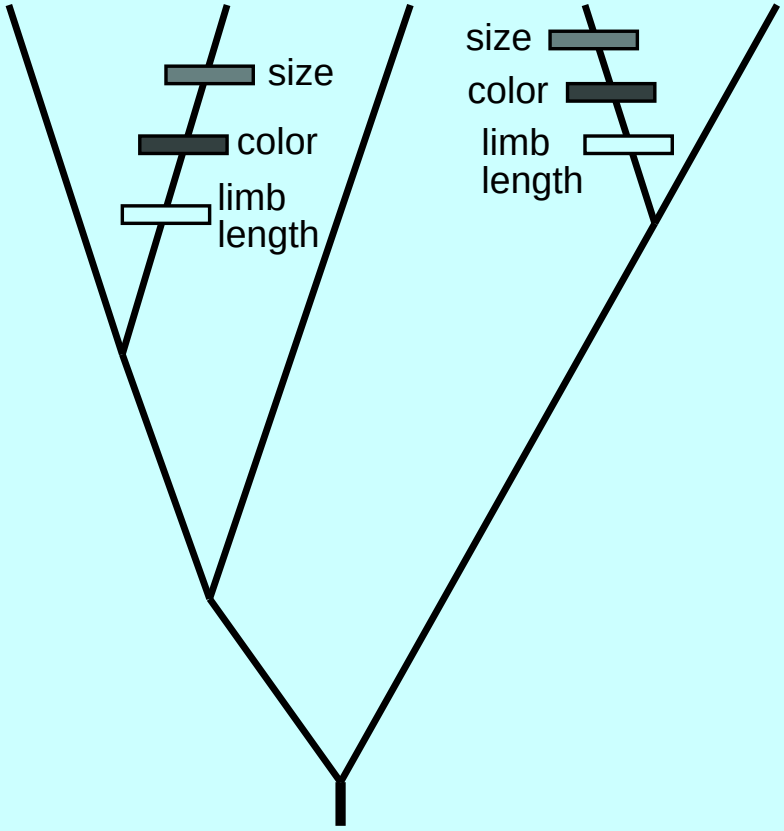
A simple example of selective covariance

covariation due not to genetic correlation
but to covariation of the selection pressure

These are Bergmann's, Allen's and Glogler's Rules
They are presumably not the result of genetic correlations
but result from patterns of selection

a simple example:

(temperate) (arctic) (temperate) (arctic) (temperate)



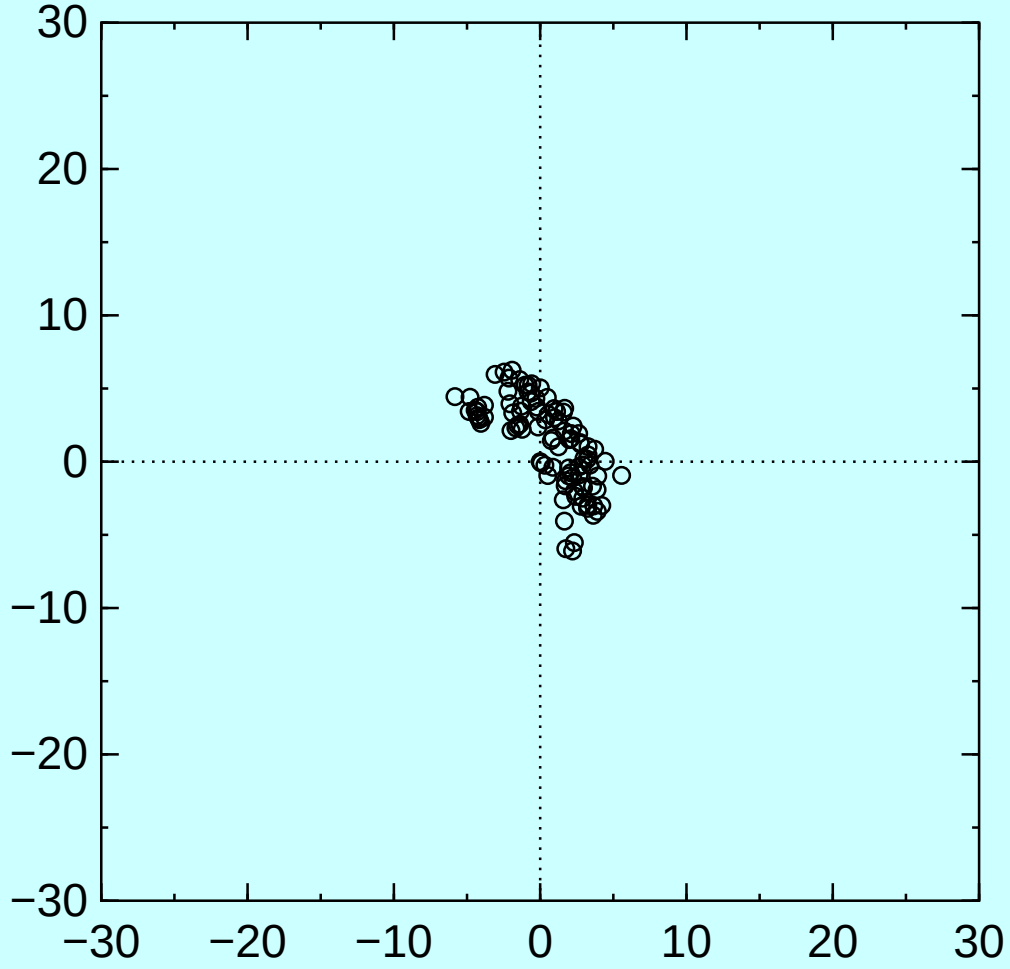
G. L. Stebbins. 1950. *Variation and evolution in plants.* Columbia Univ. Press, New York. page 121

Correcting for correlations among characters

Can we transform the set of characters to remove their correlations and thus end up with independent Brownian motions of equal variance? 

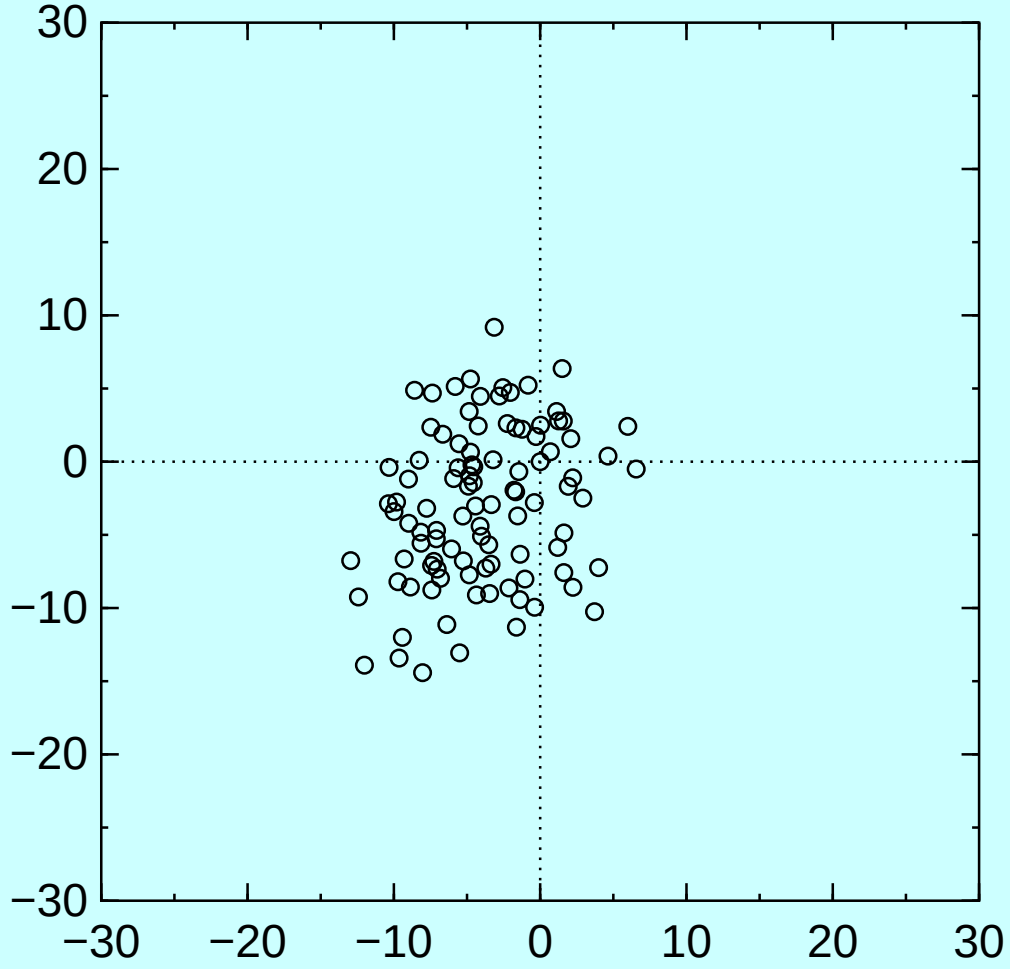
- We might hope to infer additive genetic covariances by doing quantitive genetics breeding experiments to infer them from covariances among relatives.
- There is little or no hope of inferring “selective correlations” without a complete understanding of the functional ecology.

A simulated example with two characters



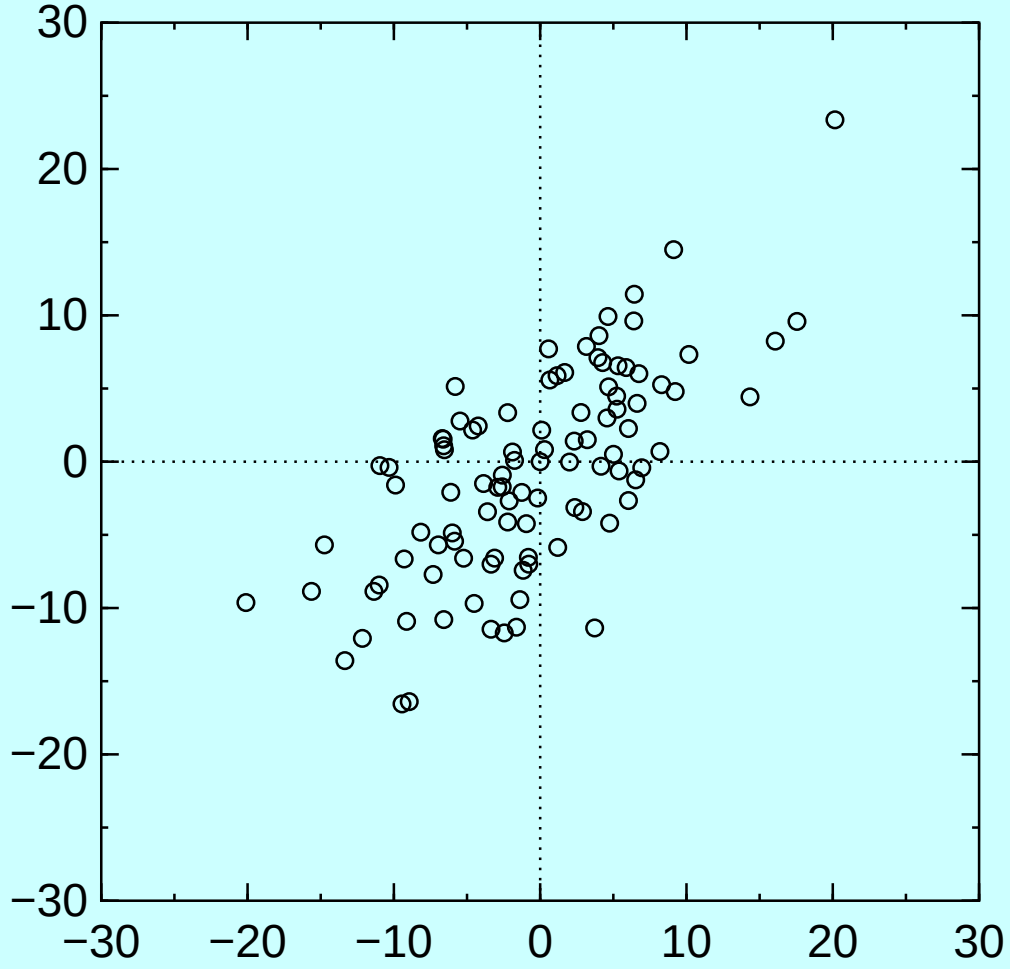
Genetic covariances are negative, but the wanderings of the adaptive peaks are positively correlated. In the first 100 generations the genetic covariances are most influential.

A simulated example with two characters



Genetic covariances are negative, but the wanderings of the adaptive peaks are positively correlated. After a while (every 10th generation up to generation 1000), the wanderings of the peaks start to be influential.

A simulated example with two characters



Genetic covariances are negative, but the wanderings of the adaptive peaks are positively correlated. In the long run (every 100th generation up to generation 10,000) the means go mostly where the peaks go.

A little algebra showing the effect of selective covariance

If we start from the familiar "Breeder's Equation" of quantitative genetics:

$$\Delta z \;=\; h^2\, S$$

it has long been known to have a multivariate version:

$$\boldsymbol{\Delta z} \;=\; \mathbf{G P^{-1} S}$$

Multiplying $\boldsymbol{\Delta z}$ by its transpose:

$$\boldsymbol{\Delta z \Delta z}^{\mathsf{T}} \;=\; \mathbf{G P^{-1} S S^{\mathsf{T}} P^{-1} G}$$

and taking expectations (treating $\mathbf{G}$ and $\mathbf{P}$ as constants) we get for the mean squares:

$$\mathrm{E}[\boldsymbol{\Delta z \Delta z}^{\mathsf{T}}] \;=\; \mathbf{G P^{-1}}\,\mathrm{E}[\mathbf{S S^{\mathsf{T}}}]\,\mathbf{P^{-1} G}$$

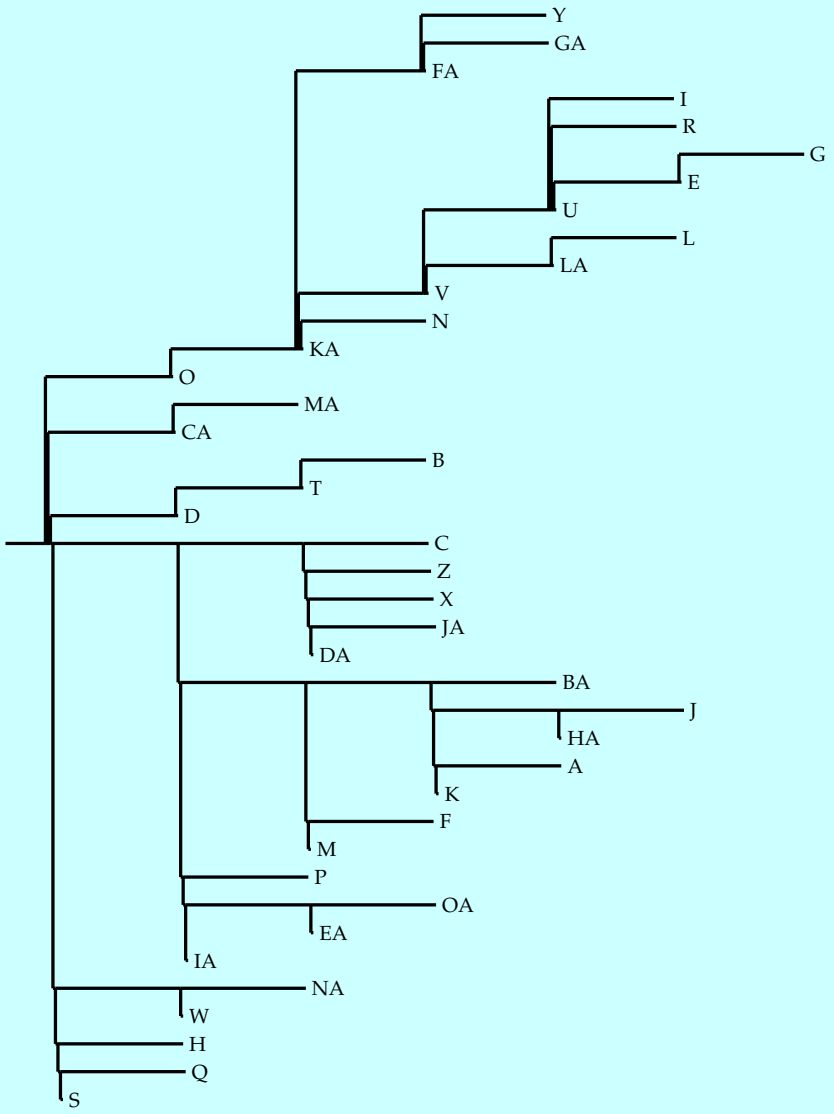
A research program?

What we could imagine doing is:

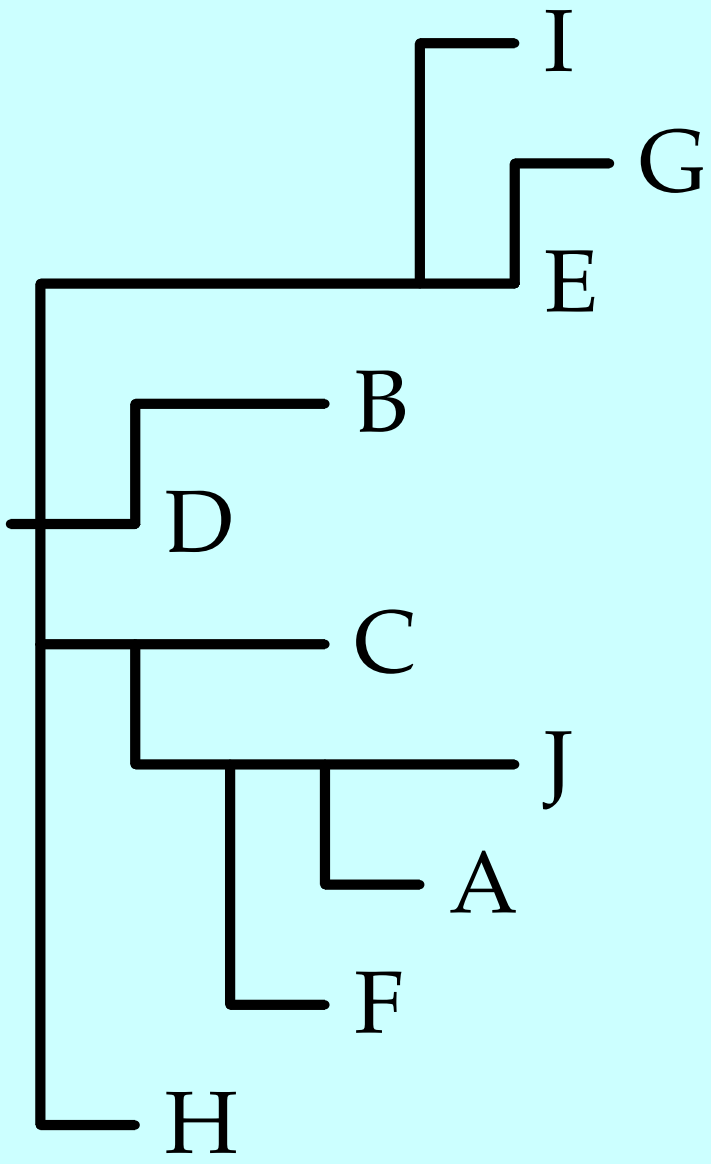
- Infer genetic covariances from a quantitative genetics breeding experiment, perhaps in more than one species.
- Infer the covariances of the changes along the phylogeny.
- From them, back-calculate the selective covariances.

(For the case of chasing peaks, there is also hope of using differences among neighboring species to infer the genetic covariances).

A tree with punctuated equilibrium



The punctuated tree when we sample 10 species



How to use morphometric coordinates on phylogenies?

Is it possible to simply use the coordinates of landmarks $(x_1, y_1), (x_2, y_2), \ldots, (x_p, y_p)$ as continuous phenotypes $x_1, y_1, \ldots, x_p, y_p$ using Brownian motion along a phylogeny?

Yes, but ...

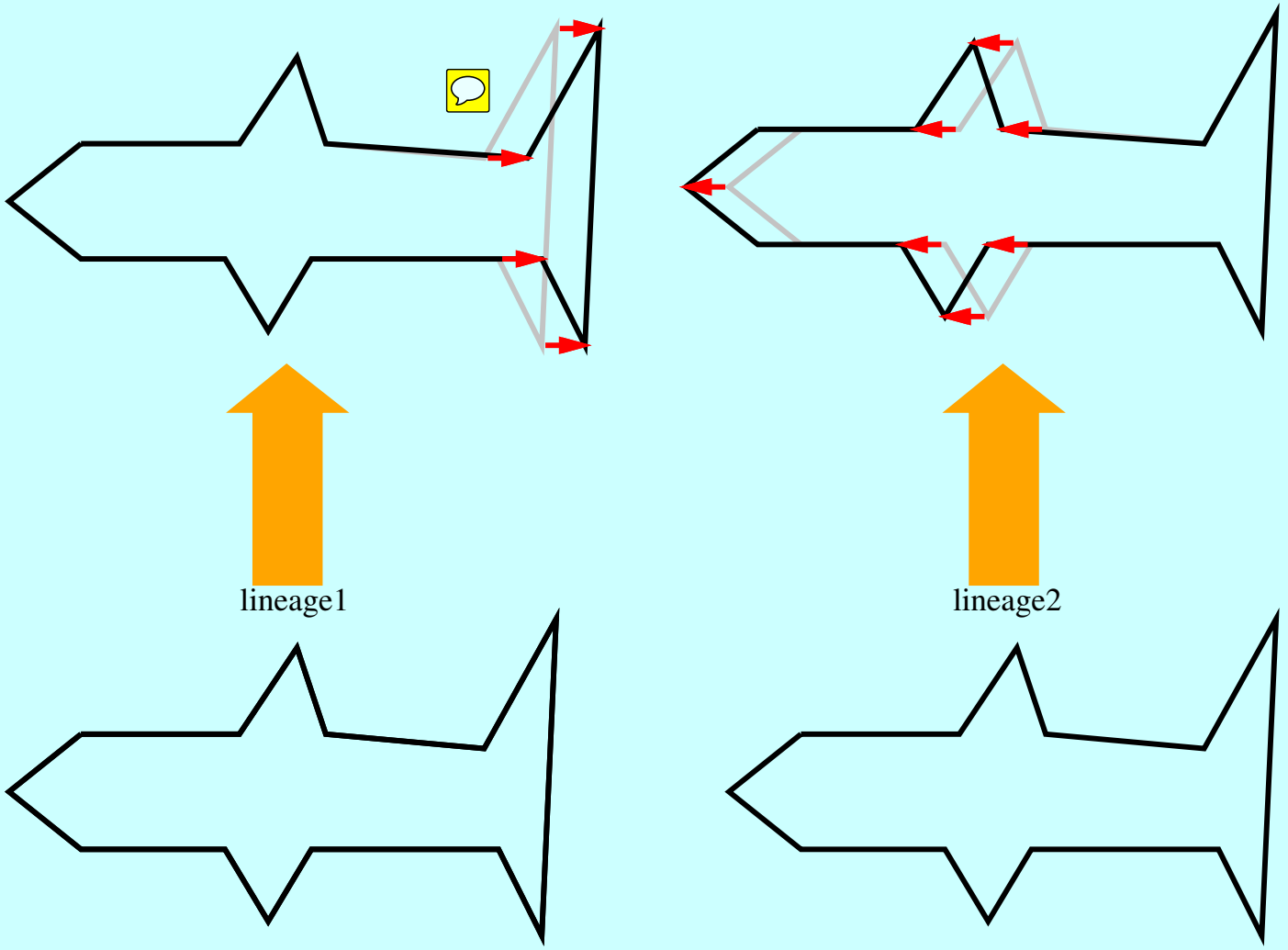
Make sure they are represented in a morphometrically valid way.

Otherwise meaningless translations (shifts) or rotations of the specimens will affect the coordinates.

In effect we are superimposing the specimens properly, although, interestingly, a complete superposition isn't necessary.

Can we superpose specimens?

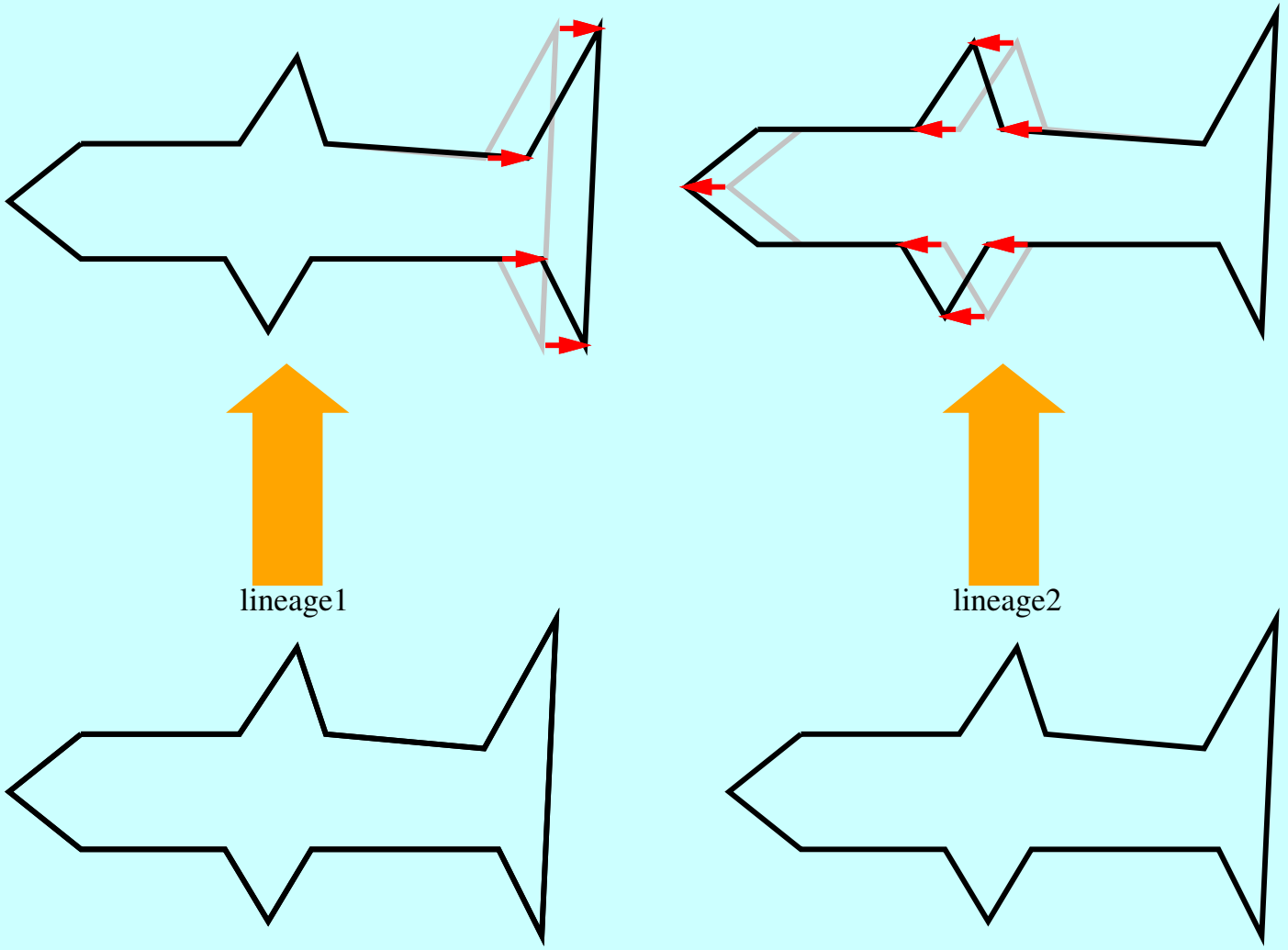
Consider two cases:



Are these different?

Why superposition is in principle impossible

Consider two cases:



Are these different? No!

Dealing with translation

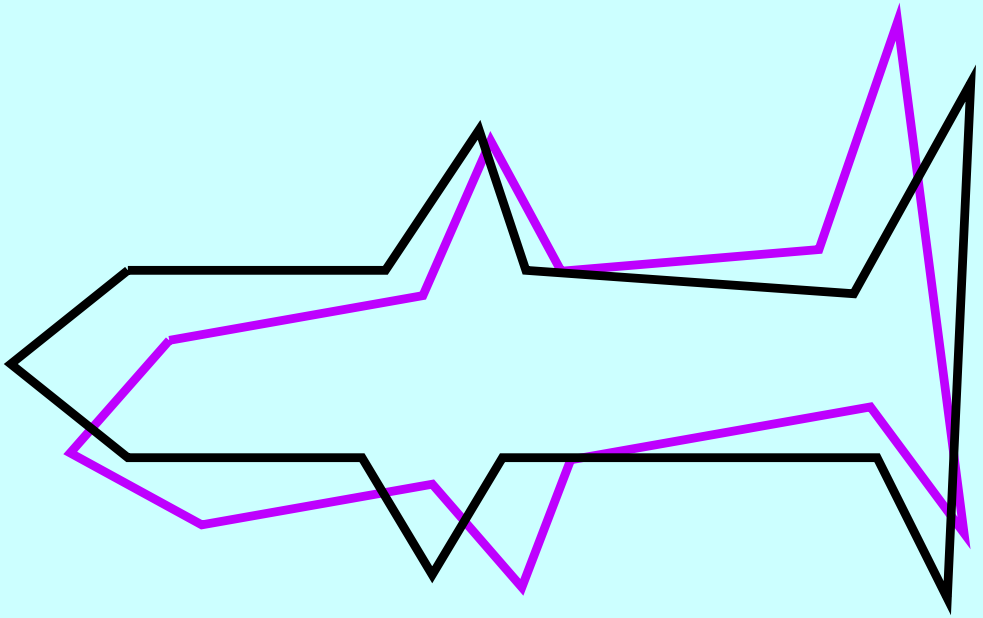
The specimens can be reduced to differences among the `x` coordinates of different points, and differences among the `y` coordinates too, thus losing the grand mean of each specimen.

This amounts to taking contrasts between the different points of one specimen *(a different matter from phylogenetic contrasts, which are for the same coordinate, but between different specimens)*.

In effect one is centering each specimen so that the mean of its points is at $(0, 0)$. (The assumption is that the horizontal and vertical placement of the specimen on the digitizer is not useful information).

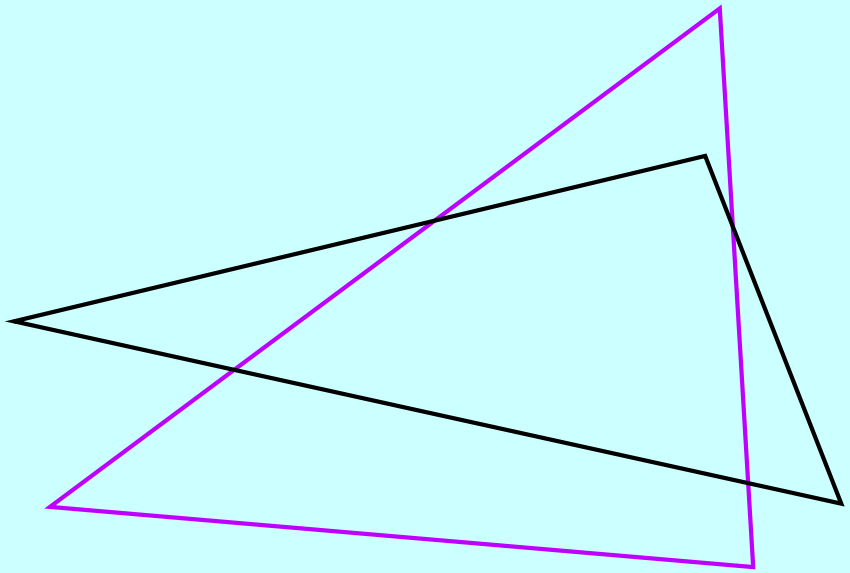
This has the effect of dropping two degrees of freedom so that each specimen now has $2p - 2$ coordinates. It now "lives" in a $(2p - 2)$-dimensional space.

The annoying issue of rotation



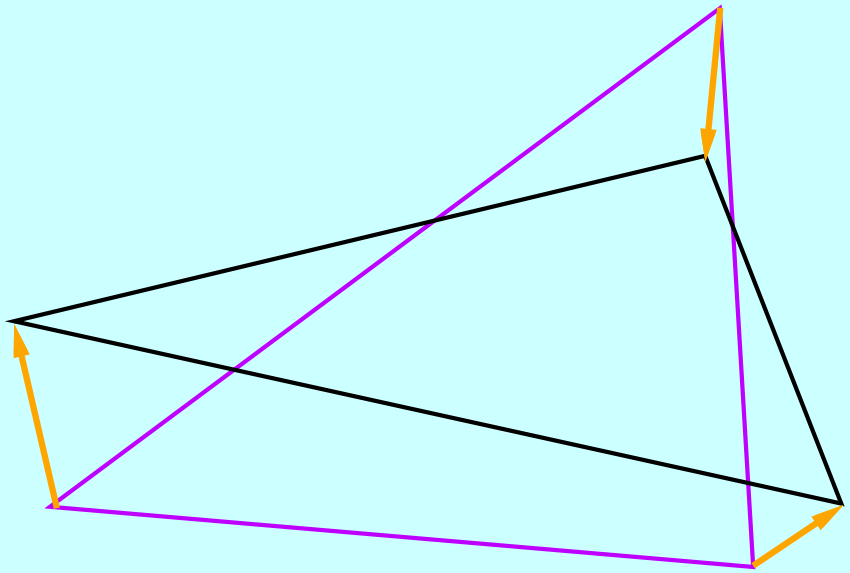
Sadly, there is no corresponding transform that tosses out rotation, as there is for translation.

The Procrustes Transform



Trying to optimally superimpose these forms by translations and rotations so as to minimize some relevant criterion ...

The Procrustes Transform



Achieves a least squares fit by centering and rotating so that the sum of squares of the golden arrows is at a minimum.

Procrustes superposition of multiple forms

One algorithm that works:

- Superpose forms $2, 3, \ldots, n$ each to the first form.

Procrustes superposition of multiple forms

One algorithm that works:

- Superpose forms $2, 3, \ldots, n$ each to the first form.
- Take the averages of the coordinates to get an average form

Procrustes superposition of multiple forms

One algorithm that works:

- Superpose forms $2, 3, \ldots, n$ each to the first form.
- Take the averages of the coordinates to get an average form
- Superpose all n forms to this average.

Procrustes superposition of multiple forms

One algorithm that works:

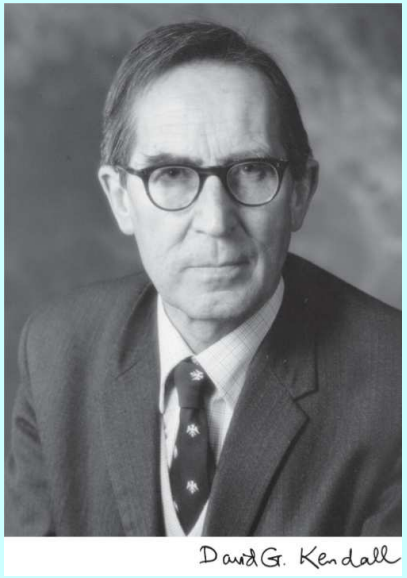
- Superpose forms $2, 3, \ldots, n$ each to the first form.
- Take the averages of the coordinates to get an average form
- Superpose all n forms to this average.
- Now recompute average form and superpose to that.

Procrustes superposition of multiple forms

One algorithm that works:

- Superpose forms $2, 3, \ldots, n$ each to the first form.
- Take the averages of the coordinates to get an average form
- Superpose all n forms to this average.
- Now recompute average form and superpose to that.
- Continue until it converges, which it will, quickly.

The "morphometric consensus"



David Kendall

(1918-2006)

- Superpose by a Procrustes fit
- Compute all pairwise distances between forms
- Do Principal Coordinates (not Components) on these
- These PCA coordinates are the ones you then analyze

John Gower (in 2008)

Problems: ignores phylogeny, implicitly assumes that the model of statistical error is independent, isotropic noise at all landmarks

But ...

(Although we will use Procrustes-superposed data for our morphometric example in the comparative methods exercises) this is *not* the best way to treat these data.

Stay tuned for more tomorrow when we come back to morphometrics.

References for multivariate Brownian motion

Felsenstein, J. 1988. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics* **19:** 445-471. **[Review with mention of Ornstein-Uhlenbeck model]**

Felsenstein, J. 2004. *Inferring Phylogenies*. Sinauer Associates, Sunderland, Massachusetts. **[Mentions this model and also sample size issues in contrasts method]**.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30:** 314-334. **[Lande's classic paper on drift versus optimum selection]**

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

References

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94:** 203-215.

Lynch, M. and W. G. Hill. 1986. Phenotypic evolution by neutral mutation. *Evolution* **40:** 915-935.

Stebbins, G. L. 1950. *Variation and Evolution in Plants*. Columbia University Press, New York. **[Describes selective covariance and cites Tedin (1925) for it]**

Tedin, 0. 1925. Vererbung, Variation, und Systematik der Gattung *Camelina*. *Hereditas* **6:** 275-386.

Armbruster, W. S. 1996. Causes of covariation of phenotypic traits among populations. *Journal of Evolutionary Biology* **9:** 261-276. **[Good exposition of selective covariance]**

References for morphometrics

Bookstein, F. L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, Cambridge, U.K. **[Early statement of the Morphometric Consensus methods]**

Bookstein, F. L. 1997. *Morphometric Tools for Landmark Data*. Cambridge University Press, Cambridge. **[also a bit dated now]**

Dryden, I. L. and K. V. Mardia. 1998. *Statistical Shape Analysis*. John Wiley and Sons, New York. **[Much mathematical machinery for the Morphometric Consensus (MMC)]**

Gower, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53:** 325–338. **[Principal coordinates analysis – principle components for distance matrices, used by MMC]**