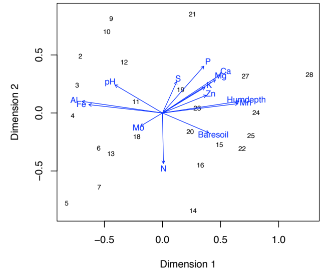
Ordination and environment

We take granted that vegetation is controlled by environment, so

1. Two sites close to each other in ordination have similar vegetation, and
2. If two sites have similar vegetation, they have similar environment; moreover
3. Two sites far away from each other in ordination have dissimilar vegetation, and perhaps
4. If two sites have different vegetation, they have different environment

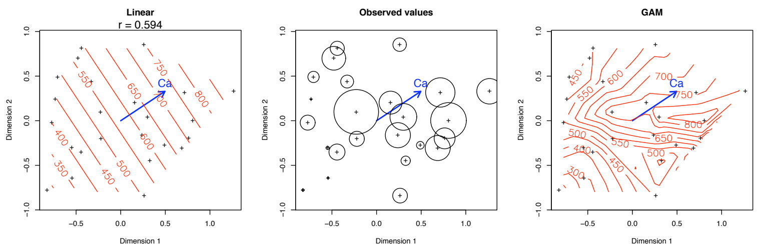
Fitted vectors

- **Direction** of fitted vector shows the gradient, **length** shows its importance.
- For every arrow, there is an equally long arrow into opposite direction: Decreasing direction of the gradient.
- Implies a linear model: Project sample plots onto the vector for expected value.
- Class values as weighted averages.

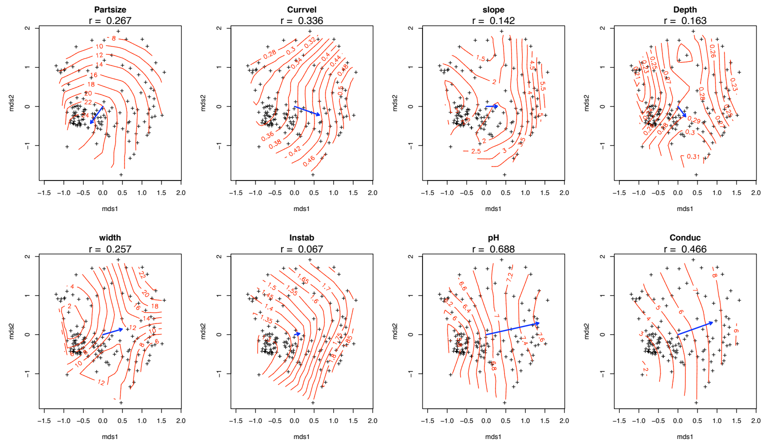


Alternatives to vectors

- Fitted vectors natural in constrained ordination, since these have linear constraints.
- Distant sites are different, but may be different in various ways: Environmental variables may have a non-linear relation to ordination.



Example: River bryophytes

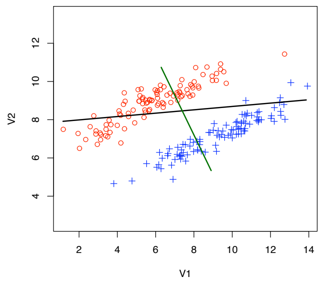


Lessons from environmental interpretation

- Environmental variables need not be parallel to ordination axes.
- Axes cannot be taken as gradients, but gradients are oblique to axes: You cannot tear off an axis from an ordination.
- **Never** calculate a correlation between an axis and an environmental variable.
- Environmental variables need not be linearly correlated with the ordination, but locations in ordination can be exceptional.

Constrained vs. unconstrained aims

- Unconstrained ordination tries to display the variation in data.
- Constrained ordination tries to display only the variation that can be explained with constraining variables.
- You can observe only things that you have measured.



The constraining toolbox

- **Linear tools** based on PCA framework:
  - Discriminant analysis, Canonical Correlations
  - Redundancy Analysis (RDA).
  - Only RDA useful in community ecology – if linear model is adequate.
- **Unimodal tools** based on CA framework:
  - Constrained or 'Canonical' Correspondence Analysis (CCA).
  - Absolutely the most important constrained ordination in ecology: The only one dealt with in these lectures.

Constrained Correspondence Analysis (CCA)

Ordinary Correspondence analysis gives:

1. **Site scores** which may be regarded as describing the gradients.
2. **Species scores** which may be taken as location of species optima in the space spanned by site scores.

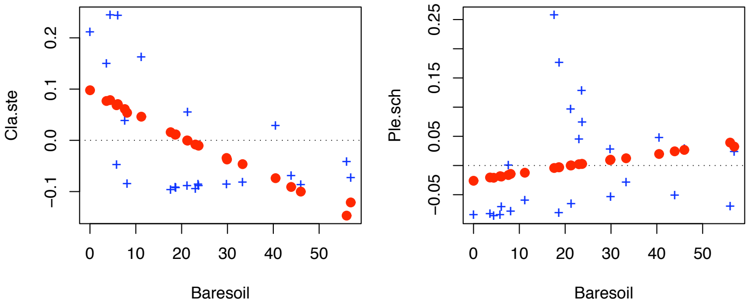
Constrained or 'Canonical' Correspondence Analysis gives in addition:

3. **Environmental scores** which define the gradient space.

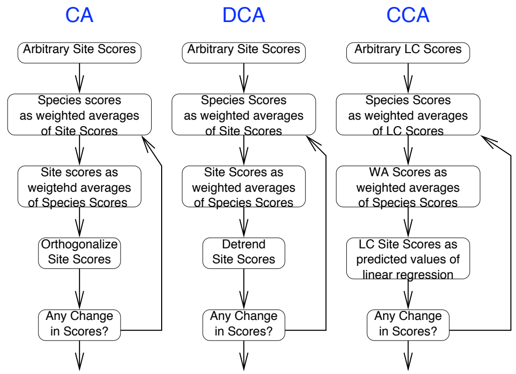
And optimizes the interpretability of results.

CCA: Algorithm

1. Fit weighted linear regression to all species individually using all constraints as explonatory variables.
2. Analyse fitted values using CA



CCA: Alternating regression algorithm



Two kind of site scores:

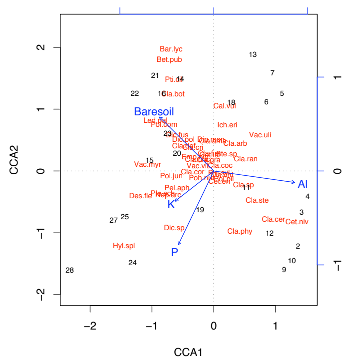
1. **LC scores** are predicted values of multiple regression with constraining variables = the constraints.
2. **WA Scores** are weighted averages of species scores.

Those numbers. . .

- Eigenvalues exactly like in CA.
  - CCA eigenvalue *should* be lower than in CA – or constraining may have been useless.
  - Eigenvalue has nothing to do with variance, so there is neither 'variance explained'.
- Species – Environment correlation: Multiple correlation from constraining regression: Usually high even with poor models.
- Pointwise goodness of fit can be expressed either as residual distance from the ordination space or as proportion of projection from thet total Chi-squared distance – exactly like in PCA.

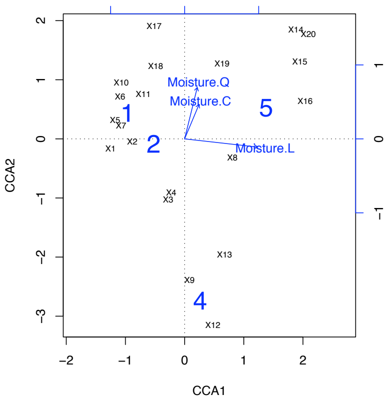
CCA plot

- Like CA plot, but now a **triplot**: Vectors for *linear* constraints.
- Classes as weighted averages.
- Most use LC scores: These are the constraints.
- Popular to scale species relative to eigenvalues, but keep sites unscaled.
- Sites do not display the configuration, but their projections onto environmental vectors are the estimated values.



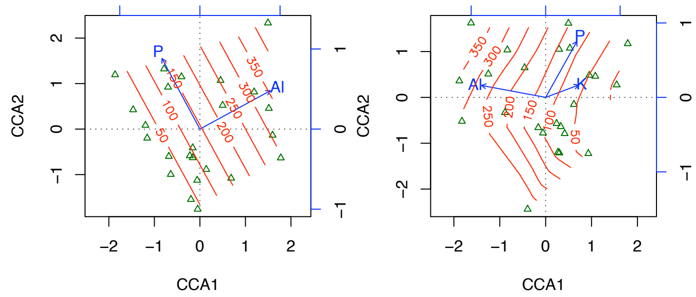
Class constraints

- Class variables usually as 'dummy' variates: Make $m-1$ indicator variables out of m levels
- Indicator scoring: 1 if site belongs the the class, 0 otherwise
- One dummy less than levels, because all are redundant
- Ordered factors may be better expressed with polynomial constraints



Predicted values of constraints

- Project a site point onto environmental arrow: Prediction
- Exact with two constraints: Multidimensional space warped



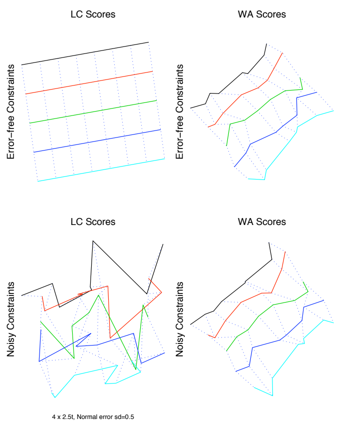
LC or WA Scores?

MIKE PALMER:

- Use LC scores, because they give the best fit with the environment, and WA scores are a step from CCA towards CA.

BRUCE MCCUNE:

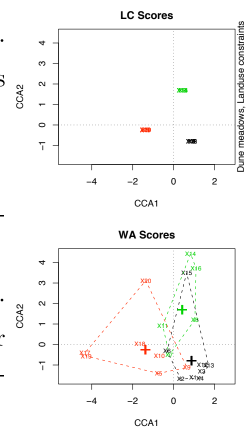
- LC scores are excellent, if you have no error in constraining variables. Even with small error, LC scores become miserable, but WA scores are good even in noisy data.



WA and LC scores with class constraints

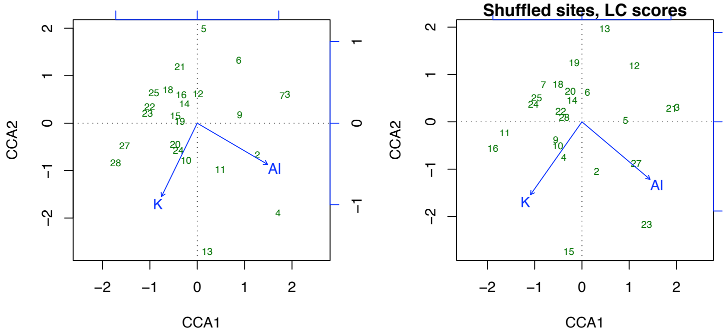
1. Make class centroids as distinct as possible.
2. Make clouds about centroids as compact as possible.

- Success ≈ λ.
- LC scores are the class centroids: The expected locations.
- If high λ, WA scores are close to LC scores.
- With several class variables, or together with continuous variables, the simple structure becomess blurred.

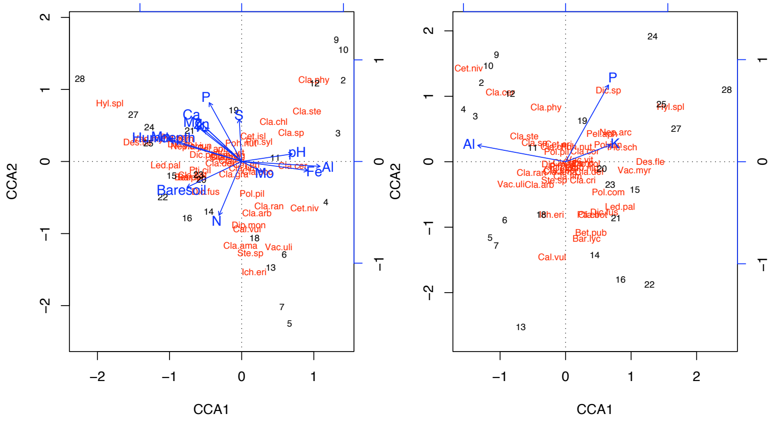


LC scores are the constraints

LC scores *are* the scaled and weighted constraints and shuffling the order of sites of the community data does not change the configuration.

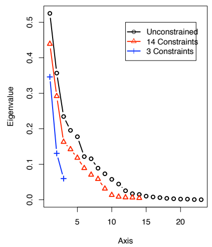


Number of constraints and the plot



Number of constraints and curvature

- Curvature cured because forced to linear constraints.
- High number of constraints = no constraint.
- Absolute limit: Number of constraints = $\min(S, N) - 1$, but release from the constraints can begin much earlier.
- Reduce environmental variables so that only the important remain: Heuristic value better than statistics.
- Reduces multicollinearity as well.



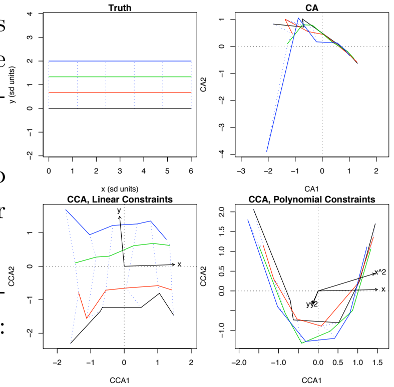
DECORANA in Disguise

Constrained Correspondence Analysis replaced DECORANA as the canonical method — and indeed, it is DECORANA in disguise

- **Detrending:** Based on fitted values from linear regression
- **Rescaling:** Linear combinations of environmental variables – scaled similarly
- **Downweighting:** Rare species fit poorly

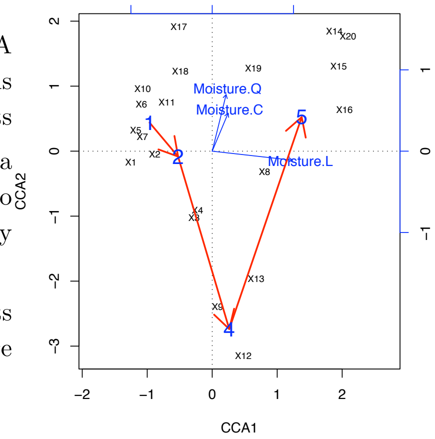
Polynomial Constraints: A Bad Idea

- Unconstrained CA produces curves, because species have non-linear responses to gradients
- Constrained CA straightens up curves, because it forces linear species responses
- Polynomial constraints produce quadratic fitted values: Ordination will be quadratic.

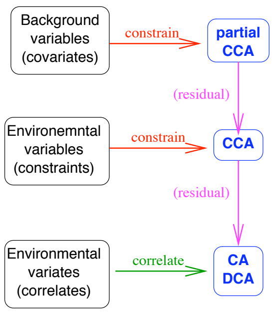


Constrained horseshoe

- Curve is removed in CCA because the solution is forced to linear constraints
- If contraints have a quadratic relation to each other, a curve may re-appear
- Polynomial constraints and interactions are generally a bad idea



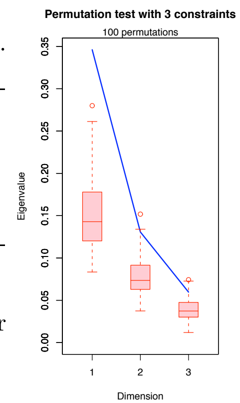
Levels of environmental intervention



- **Partial CCA** removes the effect of background variables before proper (C)CA: 'random' or 'nuisance' variables.
- Residual ordinations may be analysed at all level: Partitioning of variation.
- Constraints are linear: If levels of environmental variables are not orthogonal, this may result in negative 'components of variation'.
- Information of lower levels mixed with upper.

Significance of constraints

- CCA maximizes eigenvalue with constraints.
- Permutation tests can be used to assess significance:
  - Permute lines of environmental data.
  - Repeat CCA with permuted data.
  - If observed λ higher than (most) permutations, regarded as significant.
- Many constraints = much opportunity for optimizing: Significance usually lower.



Permutation statistic

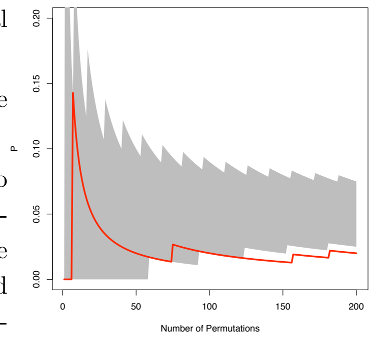
- Without constraints, sum of all eigenvalues is a natural choice
- Testing of first eigenvalue has an unclear meaning
- In partial models, use "pseudo-F":

$$F_{p,q} = \left(\sum_i^p \lambda_i^{(c)} / p \right) / \left(\sum_i^q \lambda_i^{(r)} / q \right)$$

  with constrained (c) and residual (r) eigenvalues λ and respective number of axes p and q.
- Not at all distributed like real F, but used in permutation tests

Number of permutations

- Too few permutations: Cannot detect "significant" response when it is close to a critical limit
- Too many permutations waste time
- Sequential testing: Permute so many times that assessed significance is "certainly" outside the grey zone where it could be either "significant" or "non-significant"



What is permuted?

- No conditioning variables: Community data or constraints can be permuted
- In partial models ith conditioning variables:
  - Community data cannot be permuted, because it is dependent on conditions
  - Constraints cannot be permuted, because they correlate with conditions
- Residuals are exchangeable *if they are independent and identically distributed...*
- **Reduced model** permutes residuals after conditions, **Full model** residuals after conditions and constraints

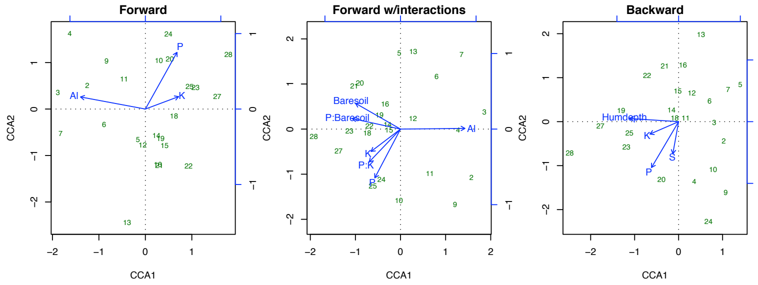
Selecting constraining variables

- Small number of variables means stricter constraints, reduced curvature, improve interpretation, increases signficance: Try to end up with one or two constraints for each independent factor.
- Significance tests and visual inspection *help* in selecting environmental variables.
- Automated selection dangerous: Small changes in data set can change the whole selection history, and omission of a variable does not mean it is unimportant.
- Final selection must be made with heuristic criteria.

The purpose of computation is insight, not numbers

Automatic stepping is dangerous

Automatic model selection may give different results depending on stepping direction, scope or small changes in the data set

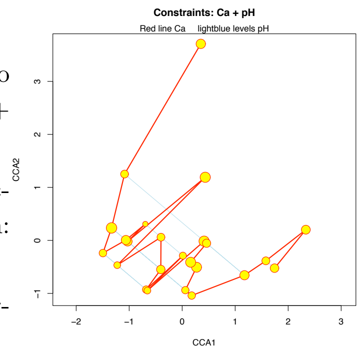


Components of Variation

- Take a partial model CCA($Y \sim X|Z$)
- The explained Inertia can be decomposed into two components:
  1. Explained by X in a simple model CCA($Y \sim X$)
  2. The residual effect of X after removing the variation caused by the conditioning variable Z
- After conditioning by Z, the eigenvalue of X decreases by the amount of shared component of variation

Negative Components of Variation

- CCA($Y \sim X|Z$) is equal to CCA(Res(CCA($Y \sim Z$)) $\sim X + Z$)
- If variables are better predictors together than in isolation: $\lambda_{X+Z} > \lambda_X + \lambda_Z$
- Constraints allow the reappearance of the curve



$$\lambda_{\mathrm{Ca}} = 0.157$$
$$\lambda_{\mathrm{Ca|pH}} = 0.183$$