Lesson 7: Indirect Ordination

1. Indirect gradient analysis
2. Similarity/dissimilarity indices
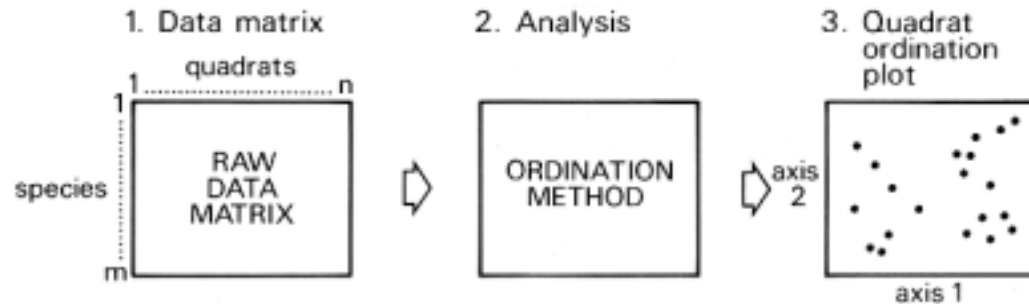3. Polar ordination (Bray and Curtis ordination)

Indirect gradient analysis

- **Indirect gradient analysis is a multidimensional picture of the relationships between stands of vegetation based on their floristic similarity to each other (in the case of a plot ordination), or between species based on their concurrence in vegetation plots (in the case of a species ordination).**

- The first step is to calculate the similarity of each stand to all the others. This can be done using a variety of similarity indices. This information is then arranged into a similarity/dissimilarity matrix.

- The coordinates of stands or species in the multidimensional space are then correlated with environmental information to detect possible environmental gradients within the data set.

Goals of ordination

- **The goals of ordination are**:

    - Show floristic relationships between stands of vegetation or between species. The distances between points on the ordination are measures of their floristic degree of similarity.

    - Reduce noise (unexplained variation that masks the similarity relationships between species and/or plots)

    - Discover the underlying structure of the vegetation data that is due to redundancy. The redundant nature of vegetation data is caused by sampling similar stands of vegetation and is due to the coordinated species responses in similar environments. Many vegetation samples have similar species composition, presumably due to their occurrence in similar environments.
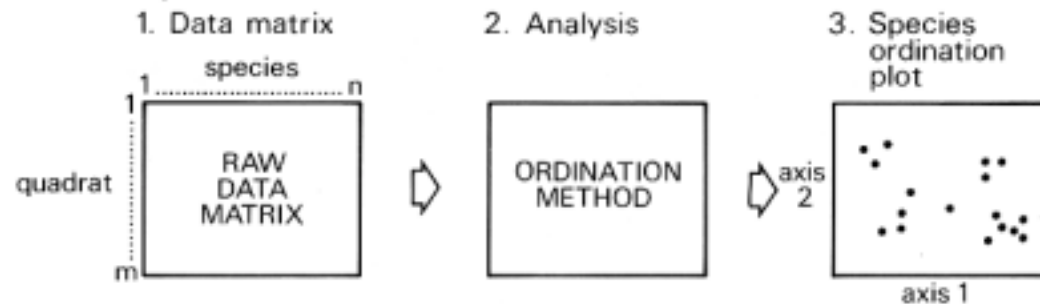
Plot and species ordinations

Quadrat Ordination

1. Data matrix

quadrats
1..........................n

species
1

m

RAW
DATA
MATRIX

2. Analysis

ORDINATION
METHOD

3. Quadrat
ordination
plot

axis
2

axis 1

On the ordination plot, each point represents a **quadrat** and the greater the distance between any two points, the greater the difference in floristic composition of the quadrats which they represent.
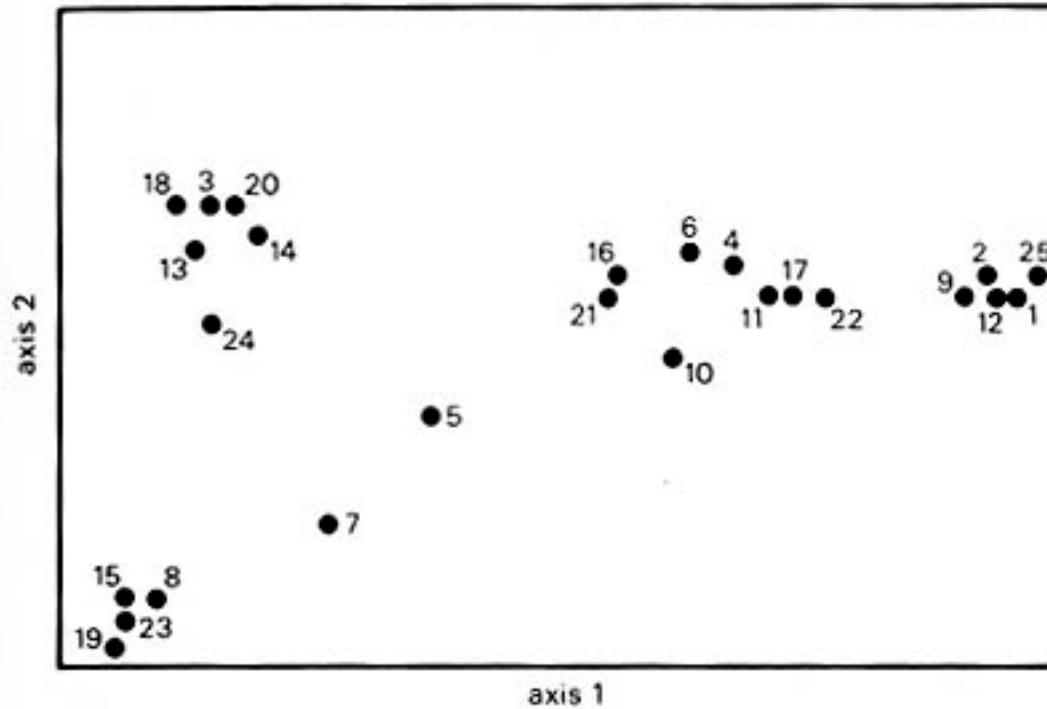
Species Ordination

1. Data matrix

species
1..........................n

quadrat
1

m

RAW
DATA
MATRIX

2. Analysis

ORDINATION
METHOD

3. Species
ordination
plot

axis
2

axis 1

On the ordination plot, each point represents a **species** and the greater the distance between any two points, the greater the difference in the distribution of the species across all the quadrats

Plot ordination

Detrended correspondenc analysis of Gutter Tor data, Kent and Coker, 1992

Species ordination

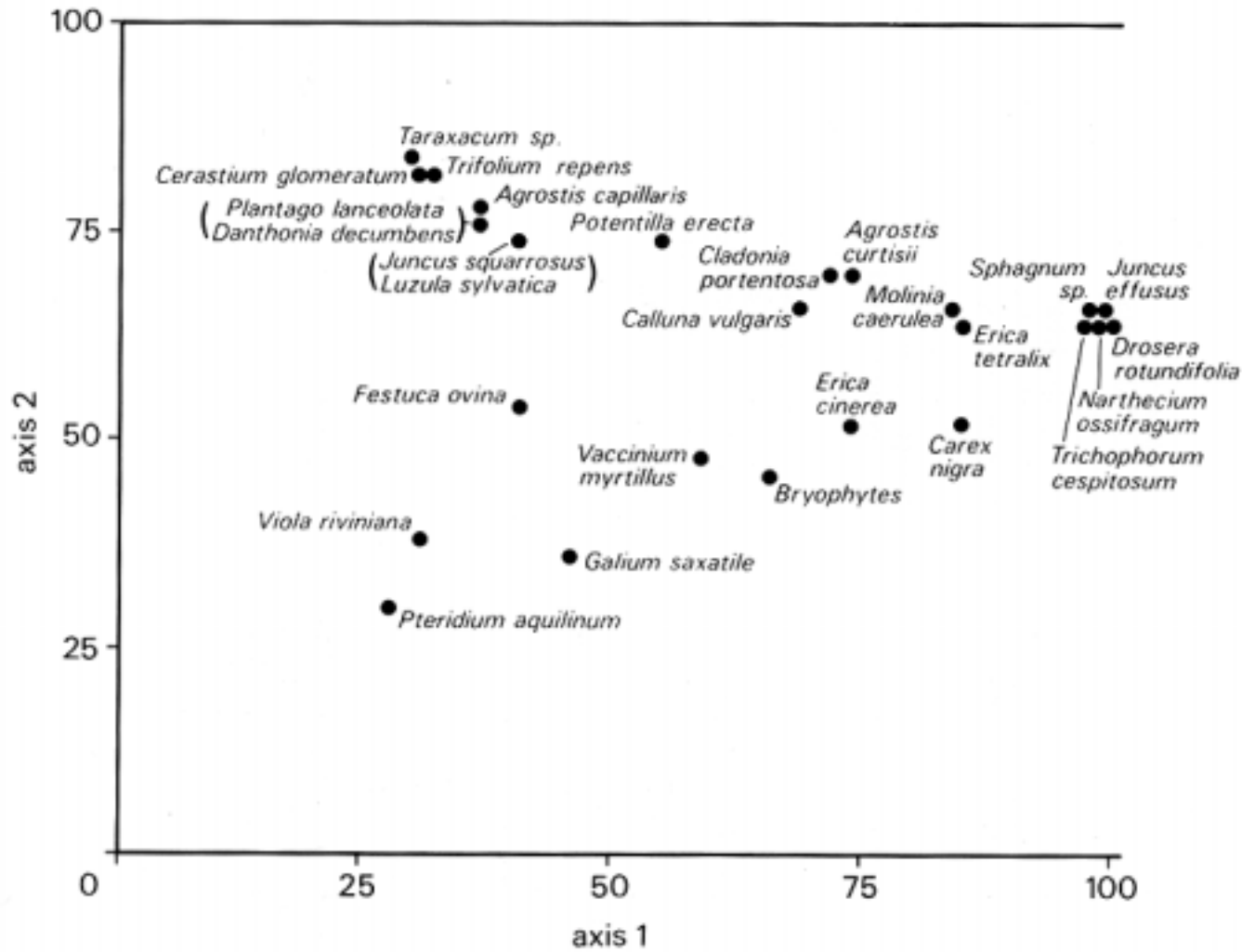**Figure 5.11** A two-dimensional species ordination plot derived from detrended correspondence analysis of the Gutter Tor data (Table 3.4)

Kent and Coker, 1992

Interpreting the ordination diagram

- The axes of the ordination are *gradients of floristic similarity* (in the case of quadrat ordinations) or *gradients of plot-occurrence similarity* (in the case of species ordinations).

- The results of ordination can be displayed one, two or three dimensions which define the *ordination space*.

- In general, the axes produced by ordination analysis come out in descending order of importance. The first axis is most important and describes the most variation in the floristic data.

- A powerful attribute of plot ordinations is that the plot numbers can be replaced with species cover values or environmental data to portray trends of environmental variables or species cover within the ordination space.

Relating the plot ordination to environmental data

Figure 5.9 (a) A two-dimensional quadrat ordination plot derived from detrended correspondence analysis for the Gutter Tor data (Table 3.4). (b) % soil moisture (gravimetric method) plotted on the quadrat ordination diagram. (c) Soil pH plotted on the quadrat ordination diagram. (d) Grazing intensity plotted on the quadrat ordination diagram

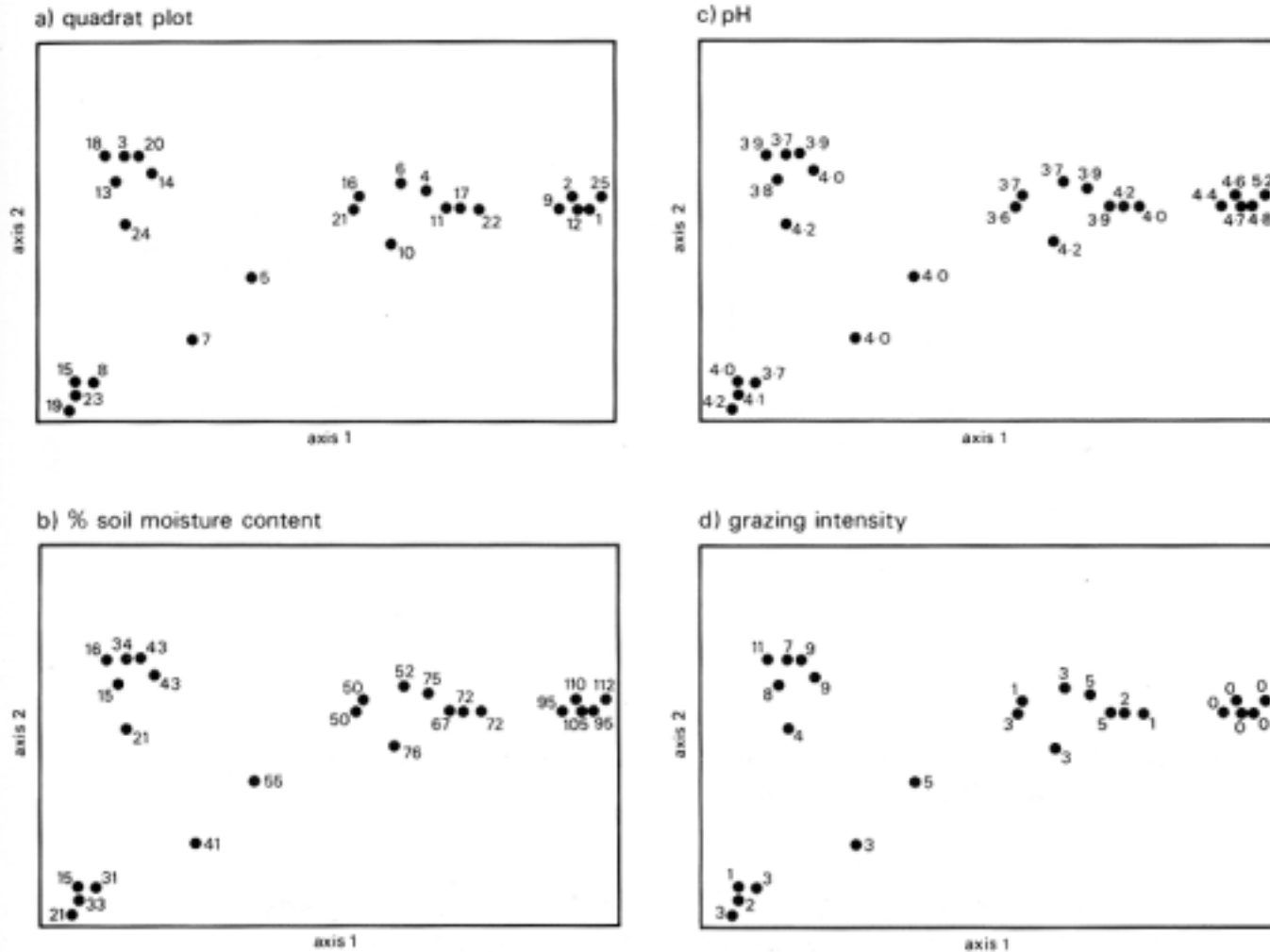# Relating the plot ordination to environmental data

Figure 5.9 (a) A two-dimensional quadrat ordination plot derived from detrended correspondence analysis for the Gutter Tor data (Table 3.4). (b) % soil moisture (gravimetric method) plotted on the quadrat ordination diagram. (c) Soil pH plotted on the quadrat ordination diagram. (d) Grazing intensity plotted on the quadrat ordination diagram

Environmental gradients and the species ordination
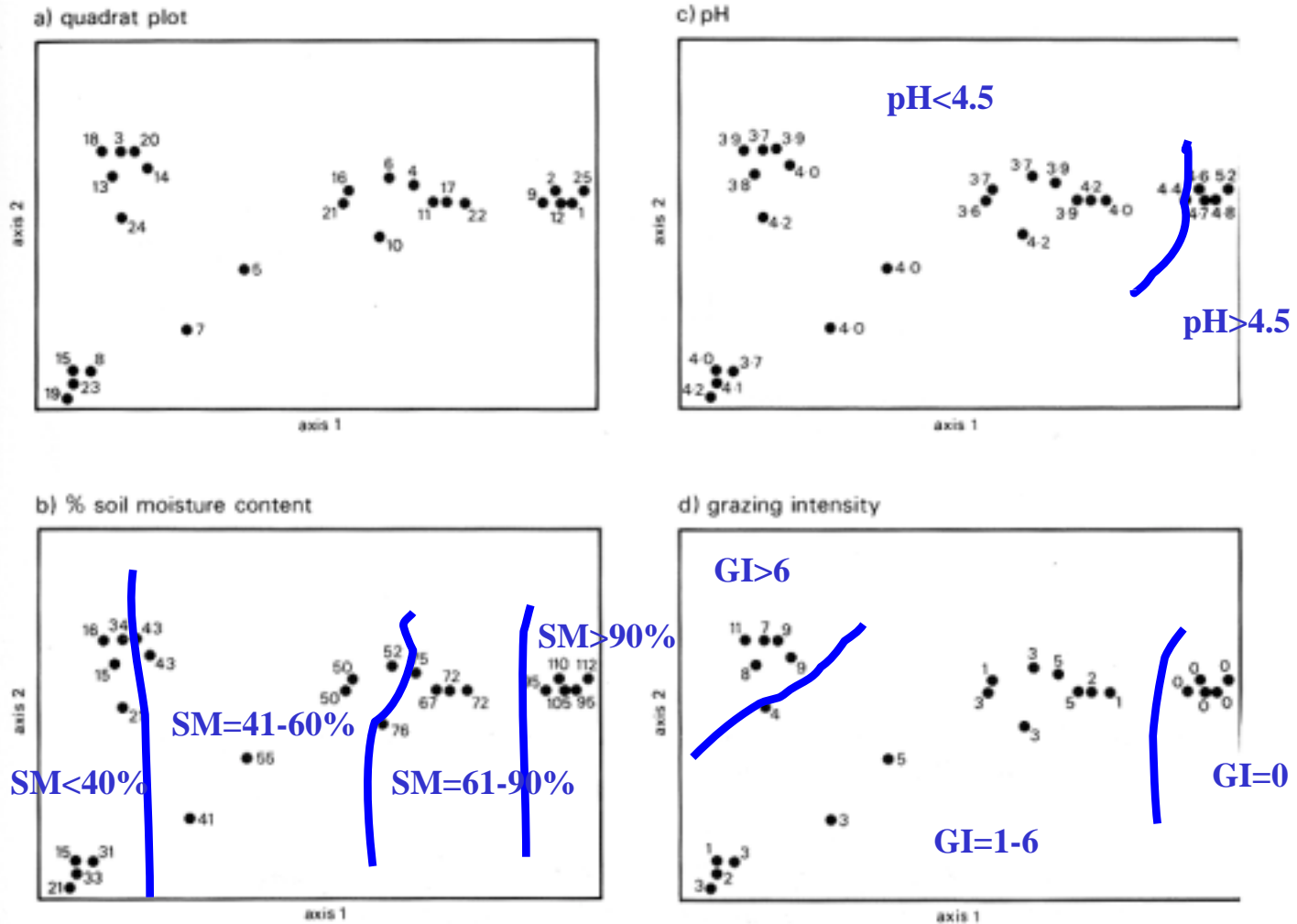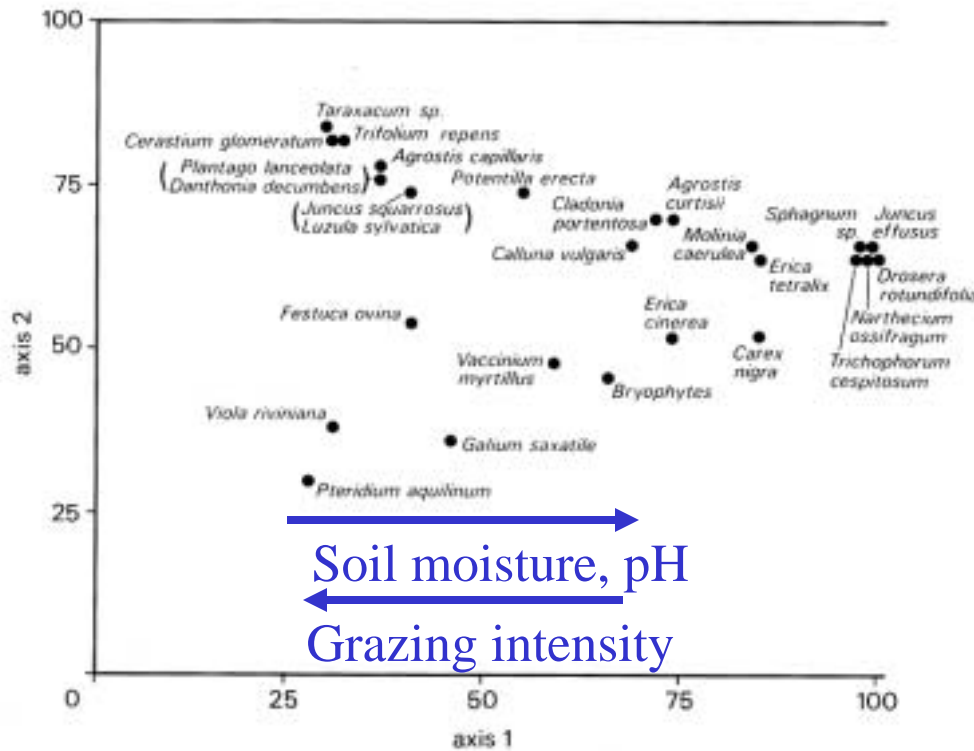

Figure 5.11 A two-dimensional species ordination plot derived from detrended correspondence analysis of the Gutter Tor data (Table 3.4)

• In Detrended Correspondence Analysis, the plot and the species ordinations are plotted in the same space.

• Thus, the same environmental gradients noted on the plot ordination also apply to the species ordination.

Kent and Coker, 1992

Polar ordination (Bray and Curtis ordination)

- Developed in 1957 by Bray and Curtis in the upland forests of Wisconsin.
- Followed a fixed set of rules that could be done by computer.
- Based on Euclidean geometry.
- Intuitive method
- Good for teaching.

Steps in calculation of the ordination

1. Begin with the primary data matrix (rows=species, columns=plot numbers).

2. Calculate the similarity and dissimilarity matrices.

3. Compute the first axis and position all plots along it.

- The first reference stand is selected that has the least similarity to all other plots.
- The second reference stand is the one with the least similarity to the first. The maximum possible dissimilarity between these plots is 100%. The dissimilarity between the two stands defines the length of the first axis.
- Locate all the other plots with respect to these two reference stands based on their geometric relationships:

Wisconsin example data set

Plots

Species	1	2	3	4	5	6	7	8	9	10
Quercus macrocarpa	9	8	3	5	6		5			
Quercus velutina	8	9	8	7						
Carya ovata	6	6	2	7		1				
Prunus serotina	3	5	6	6	6	4	5		4	1
Quercus alba	5	4	9	9	7	7	4	6		2
Juglans nigra	2				3	5	6	7	3	
Quercus rubra	3	4		6	9	8	7	9	4	3
Juglans cinerea			5		2			2		2
Ulmus americana	2	2	4	5	6		5		2	5
Tilia americana					2	7	6	6	7	6
Ulmus rubra	4		2	2	5	7	8	8	8	7
Carya cordiformis						5	6	4		3
Ostrya virginiana							7	4	6	5
Acer saccharum						5	4	8	8	9

Steps in calculation of the ordination

1. Begin with the primary data matrix (rows=species, columns=plot numbers).

2. Calculate the similarity and dissimilarity matrices.

3. Compute the first axis and position all plots along it.

- The first reference stand is selected that has the least similarity to all other plots.
- The second reference stand is the one with the least similariy to first. The maximum possible dissimilarity between these plots is 100%. The dissimilarity between the two stands defines the length of the first axis.
- Locate all the other plots with respect to these two reference stands based on their geometric relationships:
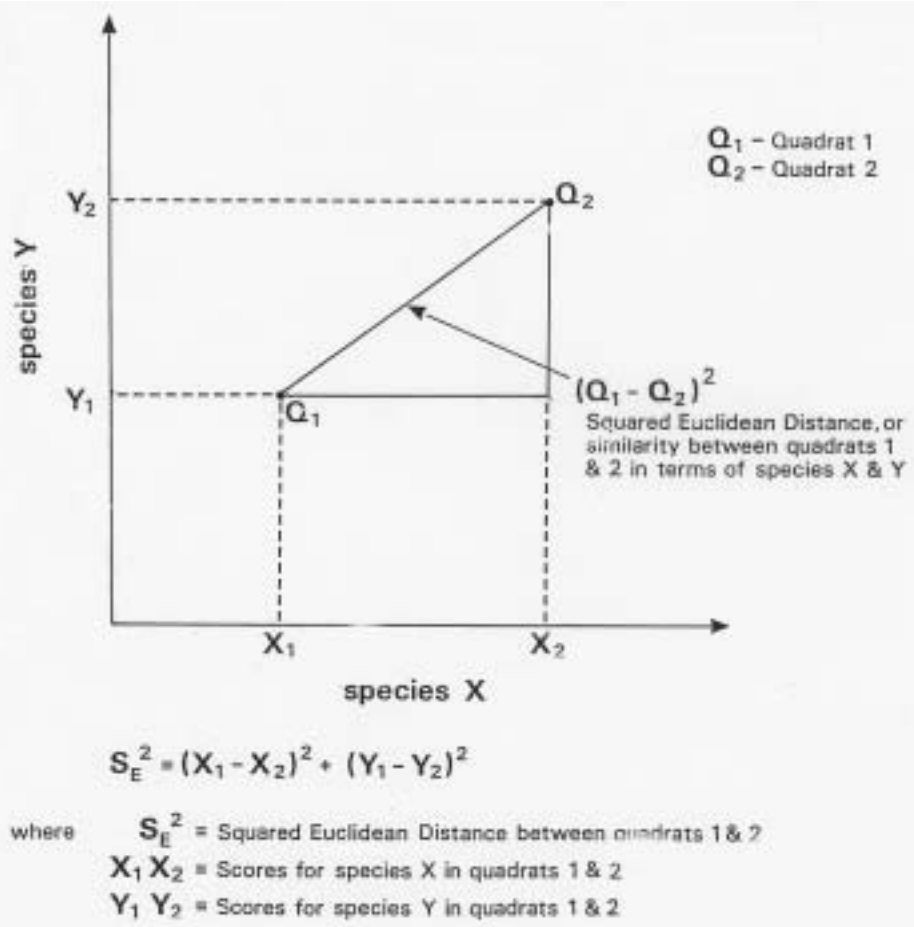
Calculating Euclidian distance

*The calculation of similarity is fairly intuitive for the situation where **two plots** have only* ***two species***. *In this case, the* ***Euclidian distance*** *can be calculated according to the Pathagorean theorem:*

$$S_E^2 = (X_1 - X_2)^2 + (Y_1 - Y_2)^2$$

where S_E^2 = Squared Euclidean Distance between quadrats 1 & 2
$X_1 X_2$ = Scores for species X in quadrats 1 & 2
$Y_1 Y_2$ = Scores for species Y in quadrats 1 & 2

Jaccard's and Sorenson's similarity indices

*For situations with many plots and many species, a variety of similarity indices have been developed. **Jaccard's and Sorenson's indices** are based on **presence/absence** of species that are shared between samples of vegetation **and** species that are unique to each sample.*

Jaccard's similarity index:

$$IS_J = \frac{a}{a + b + c}$$

a = number of species in common between the stands
b = number of species unique to the first stand,
c = number of species unique to the second stand.

Sorenson's similarity index:

$$IS_S = \frac{2a}{2a + b + c}$$

Places more emphasis on the similarities.

Czenkanowski's index of similarity

Czekanowski's coefficient (IS_c) is a modification of Sorenson's similarity index that considers the **abundance** (cover) of species:

$$IS_c = \frac{2\sum_{i=1}^{m} \min(x_i, y_i)}{\sum_{i=1}^{m} x_i + \sum_{i=1}^{m} y_i}$$

Where, in each pair of stands *(x* and *y)*, the numerator is the twice the sum of the lowest cover values for each species (*i*) in the stands; *m* is the total number of species.

The denominator is the sum of the cover values for all species in stand *x* plus the sum of the cover values in stand *y*.

Wisconsin example data set
(values are log cover values, 1-9)

Plots

Species	1	2	3	4	5	6	7	8	9	10
Quercus macrocarpa	9	8	3	5	6		5			
Quercus velutina	8	9	8	7						
Carya ovata	6	6	2	7		1				
Prunus serotina	3	5	6	6	6	4	5		4	1
Quercus alba	5	4	9	9	7	7	4	6		2
Juglans nigra	2				3	5	6	7	3	
Quercus rubra	3	4		6	9	8	7	9	4	3
Juglans cinerea			5		2			2		2
Ulmus americana	2	2	4	5	6		5		2	5
Tilia americana					2	7	6	6	7	6
Ulmus rubra	4		2	2	5	7	8	8	8	7
Carya cordiformis						5	6	4		3
Ostrya virginiana							7	4	6	5
Acer saccharum						5	4	8	8	9

Calculation of similarity and dissimilarity for plots 1 and 2 using Czenkanowski's index of similarity

Species	Plots									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Quercus macrocarpa	9	8	3	5	6		5			
Quercus velutina	8	9	8	7						
Carya ovata	6	6	2	7		1				
Prunus serotina	3	5	6	6	6	4	5		4	1
Quercus alba	5	4	9	9	7	7	4	6		2
Juglans nigra	2				3	5	6	7	3	
Quercus rubra	3	4		6	9	8	7	9	4	3
Juglans cinerea			5		2			2		2
Ulmus americana	2	2	4	5	6		5		2	5
Tilia americana					2	7	6	6	7	6
Ulmus rubra	4		2	2	5	7	8	8	8	7
Carya cordiformis						5	6	4		3
Ostrya virginiana							7	4	6	5
Acer saccharum						5	4	8	8	9
Total	42	38								

Similarity:

$$IS_c = \frac{2\sum_{i=1}^{m}\min(x_i, y_i)}{\sum_{i=1}^{m}x_i + \sum_{i=1}^{m}y_i}$$

$$IS_{1,2} = \frac{2(8+8+6+3+4+0+3+0+2+0+0+0+0+0)}{42+38}$$

$$= 2(34)/80 = 68/80 = 0.85 = 85\%$$

Dissimiliaity = 1 - Similarity

$$= 1 - 0.85 = 0.15 = 15\%$$

Matrices of similarities and dissimilarities
(Wisconsin Upland forest, Czenkanowski's index)

Matrix of similarities

Sample number	1	2	3	4	5	6	7	8	9	10
1		85	62	74	57	40	44	29	33	28
2	85		62	78	50	30	40	17	25	20
3	62	62		38	28	32	35	22	20	27
4	74	78	38		67	42	49	28	27	29
5	57	50	28	67		61	66	54	45	45
6	40	30	32	42	61		75	80	66	59
7	44	40	35	49	66	75		74	71	68
8	29	17	22	28	54	80	74		69	72
9	33	25	20	27	45	66	71	69		75
10	28	20	27	29	45	59	68	72	75	

Matrix of dissimilarities

Sample number	1	2	3	4	5	6	7	8	9	10
1		15	38	26	43	60	56	71	67	72
2	15		38	22	50	70	60	83	75	80
3	38	38		62	72	68	65	78	80	73
4	26	22	62		33	58	51	72	73	71
5	43	50	72	33		39	34	46	55	55
6	60	70	68	58	39		25	20	34	41
7	56	60	65	51	34	25		27	30	32
8	71	82	79	72	46	20	27		31	28
9	67	75	80	73	55	34	30	31		25
10	72	80	73	71	55	41	32	28	25	

Fill in matrix by calculating similarities and dissimilarities for all pairs of plots.

Steps in calculation of the ordination

1. Begin with the primary data matrix (rows=species, columns=plot numbers).

2. Calculate the similarity and dissimilarity matrices.

3. Compute the first axis and position all plots along it.

- The first reference stand is selected that has the least similarity to all other plots.
- The second reference stand is the one with the least similariy to first. The maximum possible dissimilarity between these plots is 100%. The dissimilarity between the two stands defines the length of the first axis.
- Locate all the other plots with respect to these two reference stands based on their geometric relationships.

Determination of end points of the first axis
(Wisconsin Upland forest, Czenkanowski's index)

| Matrix of similarities										
Sample number	1	2	3	4	5	6	7	8	9	10
1		85	62	74	57	40	44	29	33	28
2	85		62	78	50	30	40	17	25	20
3	62	62		38	28	32	35	22	20	27
4	74	78	38		67	42	49	28	27	29
5	57	50	28	67		61	66	54	45	45
6	40	30	32	42	61		75	80	66	59
7	44	40	35	49	66	75		74	71	68
8	29	17	22	28	54	80	74		69	72
9	33	25	20	27	45	66	71	69		75
10	28	20	27	29	45	59	68	72	75	
Matrix of dissimilarities										
Sample number	1	2	3	4	5	6	7	8	9	10
1		15	38	26	43	60	56	71	67	72
2	15		38	22	50	70	60	83	75	80
3	38	38		62	72	68	65	78	80	73
4	26	22	62		33	58	51	72	73	71
5	43	50	72	33		39	34	46	55	55
6	60	70	68	58	39		25	20	34	41
7	56	60	65	51	34	25		27	30	32
8	71	82	79	72	46	20	27		31	28
9	67	75	80	73	55	34	30	31		25
10	72	80	73	71	55	41	32	28	25	

• First sum all the similarity values for each plot.
• Find the plot with the lowest sum of similarities. This plot is the end of the first axis.
• Then determine the plot that is least similar to the first plot. This is the other end of the first axis.

Sum of similarities:

Plot 1	2	3	4	5	6	7	8	9	10
452	407	326	431	473	483	521	444	431	423

Stand 3 is least similar to all others. Plot 9 is most dissimilar to Plot 3

Calculation of x values along the 1st axis,
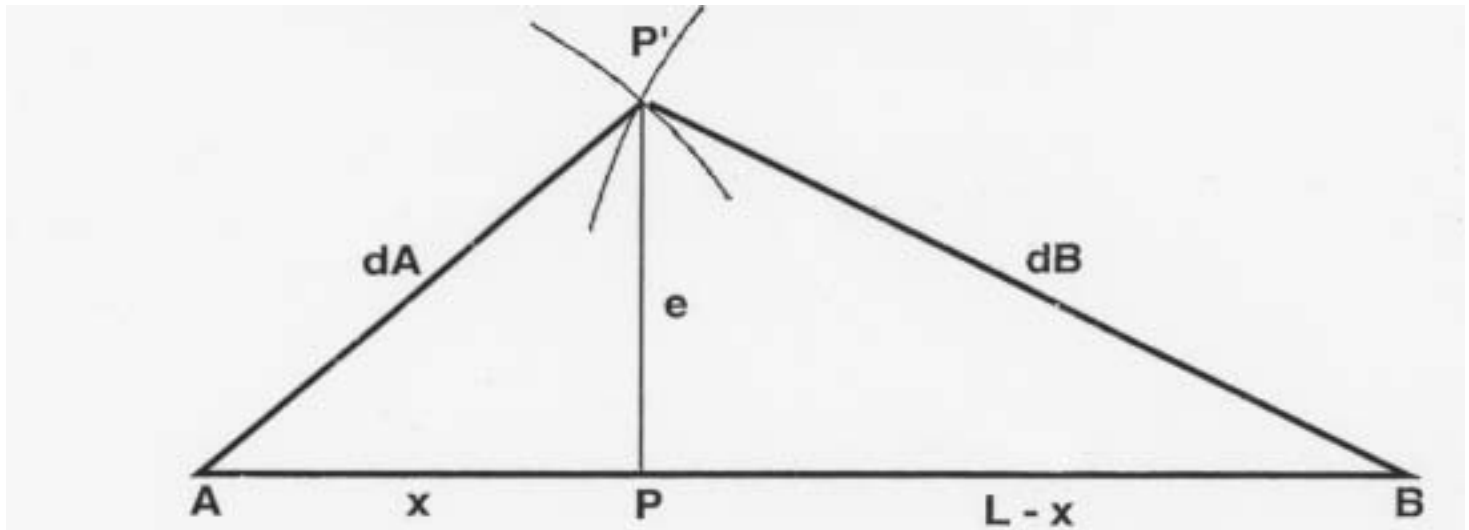
A and B = reference stands
L = Length of x axis (dissimilarity between A and B)
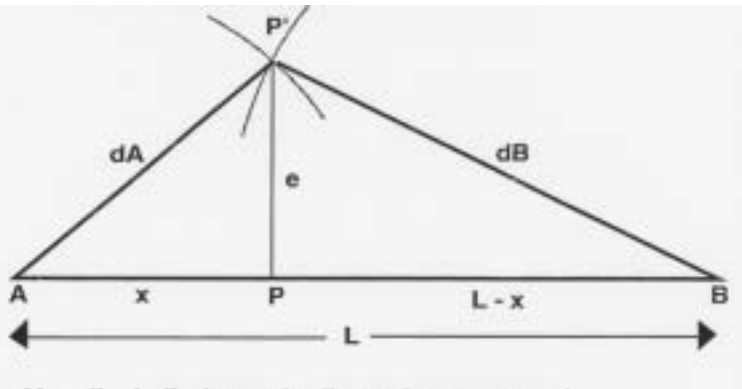P' = plot to be positioned along AB.
P = Position of P' along 1st axis.
dA = Dissimilarity of P' to A.
dB = Dissimilarity of P' to B.
x = coordinate of P along the 1st axis.

Solving for x as a function of known values (dA, dB, and L).

A and B = reference stands
L = Length of x axis (dissimilarity between A and B)
P' = plot to be positioned along AB.
P = Position of P' along 1st axis.
dA = Dissimilarity of P' to A.
dB = Dissimilarity of P' to B.
x = coordinate of P along the 1st axis.

By the Pythagorean theorem:
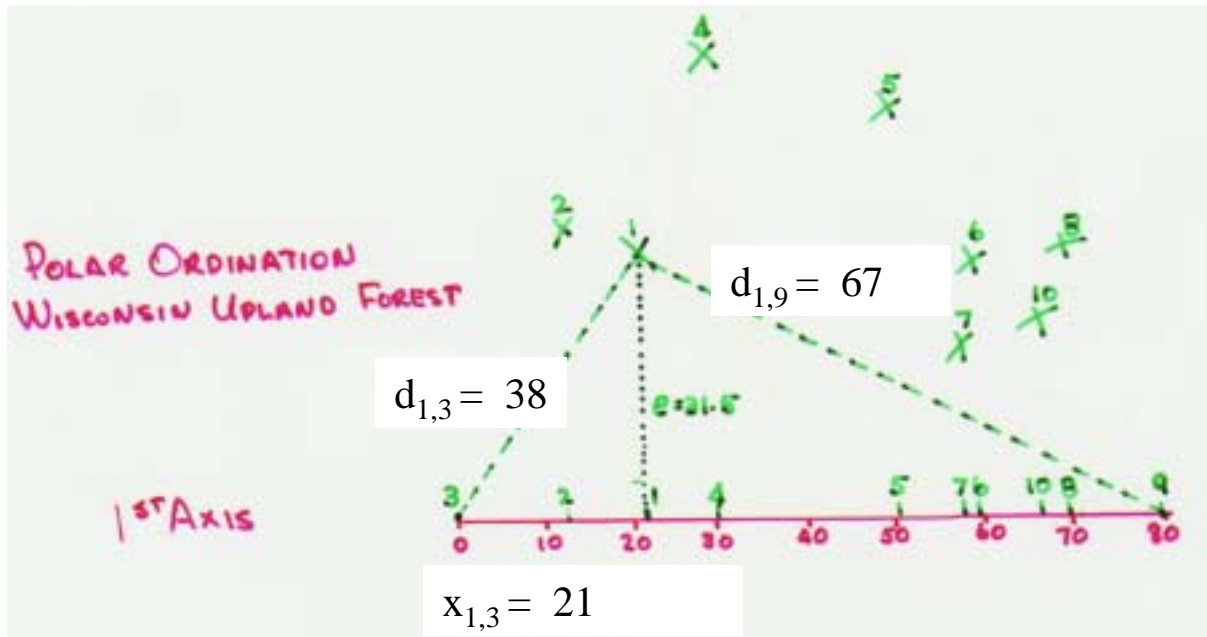
$$e^2 + x^2 = (dA)^2, \; e^2 = (dA)^2 - x^2$$
$$e^2 + (L - x)^2 = (dB)^2$$
$$(dA)^2 - x^2 + (L - x)^2 = (dB)^2$$
$$(dA)^2 - x^2 + L^2 - 2Lx + x^2 = (dB)^2$$
$$x = (L^2 + (dA)^2 - (dB)^2)/2L$$

$$x_{1,3} = (L_{3,9}^2 + (d_{1,3})^2 - (d_{1.9})^2)/2L_{3,9}$$
$$x_{1,3} = (80^2 + (38)^2 - (67)^2)/(2(80))$$
$$x_{1,3} = (6400 + 1444 - 4489)/160$$
$$x_{1,3} = 3355/160 = 21$$

Distances on the x axis

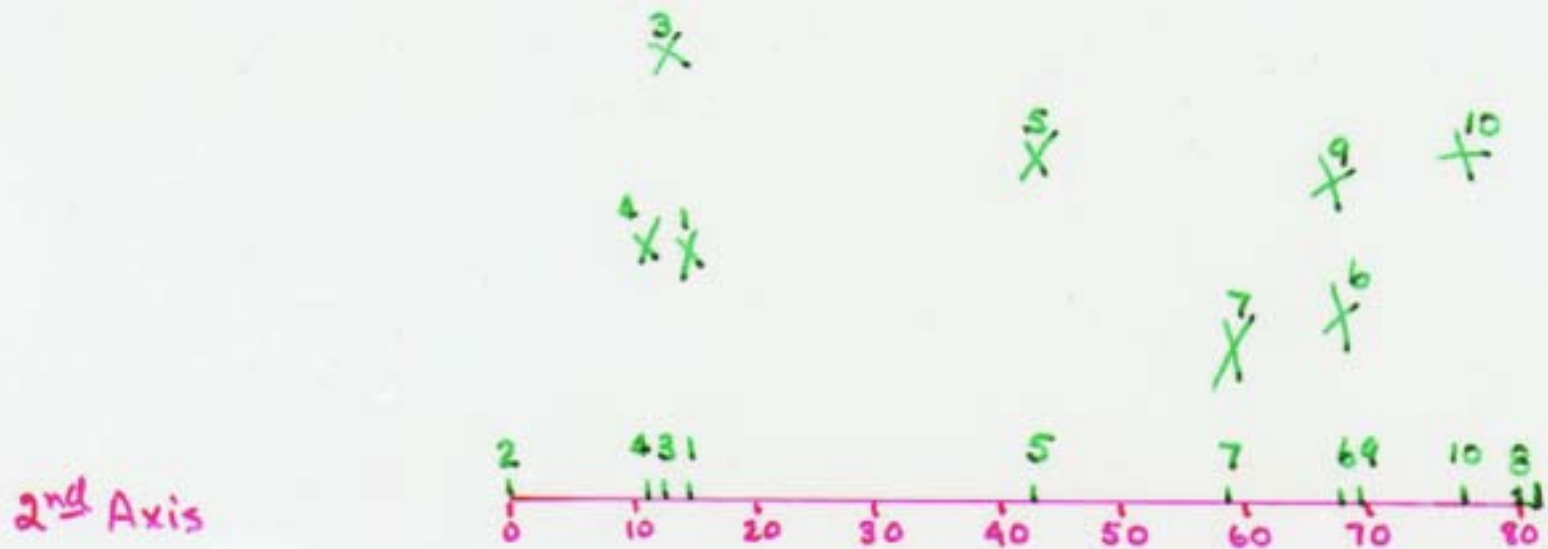| Distances on the first axis:					
Calculate the distance along the x-axis of all stands in reference to stand 3 (first reference stand):					
Use the geometric solution (K&C, p. 185) to calculate distance from stand 3:					
Stand	dA (dissim to 3)	dB (dissim to 9)	x		
1	38.2	66.6	21.62		
2	37.6	75	13.93		
3	0	80.3	0.00		
4	61.6	73	30.60		
5	71.8	54.6	53.69		
6	68.2	34.1	61.87		
7	64.7	29.5	60.80		
8	78.5	31.2	72.46		
9	80.3	0	80.30		
10	73.2	24.7	69.72		

$x = (L^2 + (dA)^2 - (dB)^2)/2L$

Calculation of the ordination (cont')

Define the second axis and position all plots along it.

- – **Select a pair of reference stands that are near to the center of the first axis and which which are close together but have a high dissimilarity between them.**
- – **Locate the other stands in exactly the same fashion as on the first axis.**

# Geometric determination of plots along the second axis

Distances on the x and y axes

Distances on the first axis:					
Calculate the distance along the x-axis of all stands in reference to stand 3 (first reference stand):					
Use the geometric solution (K&C, p. 185) to calculate distance from stand 3:					
Stand	dA (dissim to 3)	dB (dissim to 9)	x		
1	38.2	66.6	21.62		
2	37.6	75	13.93		
3	0	80.3	0.00		
4	61.6	73	30.60		
5	71.8	54.6	53.69		
6	68.2	34.1	61.87		
7	64.7	29.5	60.80		
8	78.5	31.2	72.46		
9	80.3	0	80.30		
10	73.2	24.7	69.72		
Distances on the second axis:					
Reference stands are 4 and 7. Distances (y) from stand 4:					
stand	dA (dissim to 4)	dB (dissim to 7)	y		
1	25.8	56.2	-20.78		
2	22.4	60.4	-30.59		
3	61.6	64.7	10.77		
4	0	50.9	-22.25		
5	33.3	33.9	16.04		
6	58.3	25	58.30		
7	50.9	0	55.55		
8	72.3	26.5	84.59		
9	73	29.5	83.60		
10	71.1	32.1	77.08		

$$x = (L^2 + (dA)^2 - (dB)^2)/2L$$

Plotting the ordination

x	y
21.6	-20.8
13.9	-30.6
0.0 1	0.8
30.6	-22.3
53.7	16.0
61.9	58.3
60.8	55.6
72.5	84.6
80.3	83.6
69.7	77.1

Bray and Curtis Wisconsin Upland Forest
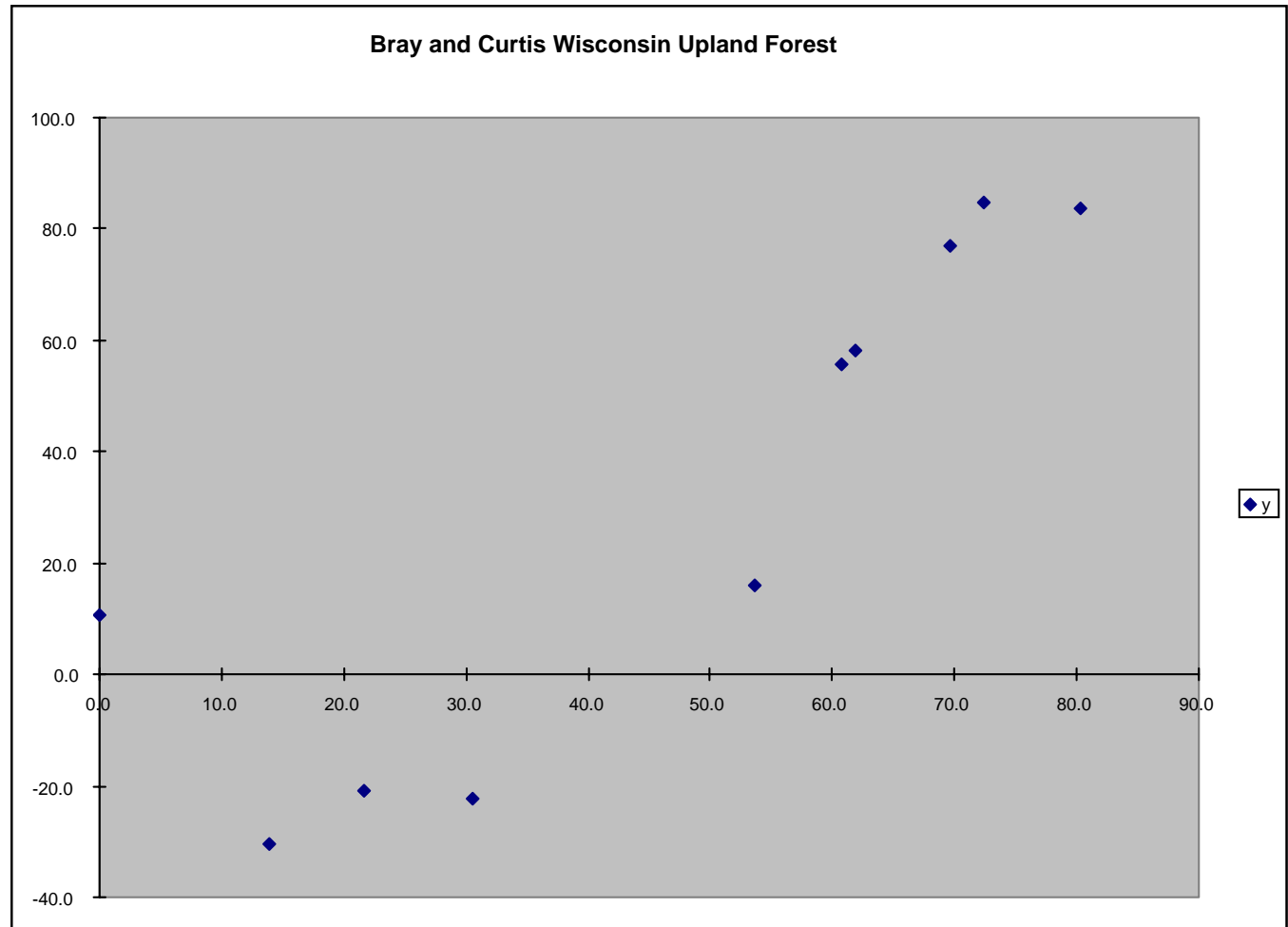
Note: The negative y *values for the for several of the plots can cause difficulties for interpretation. Selecting two other endpoints for the axis could improve this.*

Interpreting the ordination

- We have formed an "ordination space" where the axes are of gradients of floristic similarity. Stands closest together are most similar.

- At this point, we have described a coenocline, but we have no idea what the axes mean ecologically.

- If we wanted to create a "species space", we would use the same method except begin with species similarity matrix instead of samples similarity matrix. The species and the quadrat ordinations cannot be related.

- The x and y axes of the Wisconsin upland forest ordination are obviously strongly correlated and there is little new information displayed on the 2nd axis.

- The x and y axes can be correlated with environmental data for each plot with the x and y coordinates using three primary methods:

  - (1) isloline diagram's (as shown earlier in this lesson),
  - (2) Spearman's rank correlation. The variable with highest correlation with the x values best represents the x axis.
  - (3) Biplot diagrams. A group of correlation arrows centrally located in the ordination space to show direction and strength of correlations between environmental variables and the ordination space.
  - We will use the latter two methods in a later lesson.

Polar ordination strengths and weaknesses

Advantages

- Simple, easy to understand geometric method, easily taught.

Weaknesses

- Axes are not orthogonal. With large data sets, it may be difficult to get a consistent ordination.

- Not completely objective, won't always get the same answer. However, this is a function of the decision regarding reference stands, and is really amounts to viewing the ordination from different angles, although the problem of nonorthogonal axes can cause considerable distortion to the ordination space. Some of this problem can be overcome by using rules to define the reference stands.

- Distances are not metric (i.e., they are relative only)

- No explicit statement of underlying model.