Lecture 5:
FA, PCA, MCA

GEOG413/613
Dr. Anthony Jjumba

1

FA, PCA, MCA

- Factor Analysis
- Principal Component Analysis
- Multi-Criteria Analysis

2

FA, PCA, MCA

- Factor Analysis and Principal Component Analysis
  - Classified under multivariate analysis.
    - Multivariate analysis involves the application of a set of statistical methods to examine the independence and interdependence of a number of variables.
  - In spatial analysis, they are used to detect and identify the groups of closely related variables.

3

FA, PCA, MCA

- Factor Analysis and Principal Component Analysis
  - Both are variable reduction methods
  - They share some similarities but there are also differences between them
  - Broadly speaking, a correlation matrix describing relations among a number of variables is used to generate a smaller number of dimensions (factors, components) that represent most the information contained in the variables.

4

Principal Component Analysis

- Useful where the observed variables are highly correlated
- Reduces the number of observed variables to a smaller number of principal components.
- The principal components account for most of the variance in the observed variables
- PCA is essentially a linear recombination of original variables into a new set of variables, each of which is orthogonal to one another. All of the variance in the original dataset is re-allocated among the new measures.

5

Factor Analysis

- Factor analysis is a method of data reduction
- Seeks latent variables that are reflected in the observed variables
- Latent variables are constructs that estimate the factors which influence the observed variables
- Assumes that the original variable is influenced by various determinants. One part shared by other variables is called *common variance.* Other part is known as a *unique variance*.

6

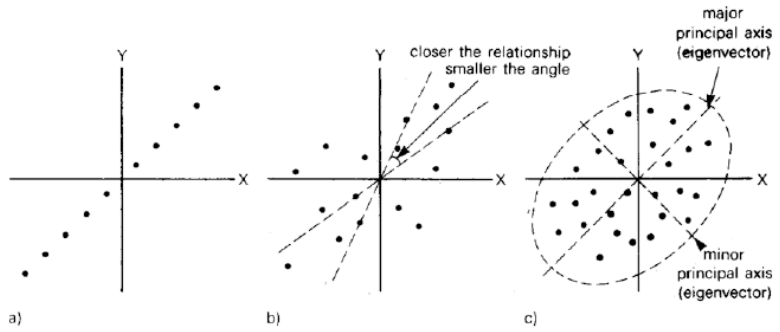


Figure 15.2 Bivariate distributions and types of relationships: (a) perfect correlation, (b) moderate correlation, (c) no correlation

Shaw G., Wheeler D.,(1985) Statistical techniques in Geographical analysis

7

- The factors will be uncorrelated with one another because they are orthogonal to each other.
- If the data are characterized by an *extremely elongated ellipse*, it implies that the data could be well represented by a single factor. This factor would capture *all of the variability* of the data.
- If all data fall in a *circular ellipse* where there is no dominant axis, all factors explain *an equal amount of the variability* in the original data.

8

Multi-criteria Evaluation

- MCE
  - The implementation of decision-making rules to identify and enable the combination of many criteria, in the form of data layers in a GIS, into a single map (Lopez-Marrero et al, 2011)
  - The combination of information from several criteria to form a single index of evaluation
  - AKA
    - Multi Criteria Analysis
    - Multi Criteria Preference Analysis
    - Multi Criteria Decision Making
    - Multi Objective Evaluation

9

Multi-criteria Evaluation

- Two basic types of criteria
  - Boolean Criteria
    - Boolean Logic (application of logical AND (intersection); logical OR (union))
    - Referred to as Constraints
  - Quantitative Criteria
    - Continuous variables (e.g. 0–1, 0–100, 0–255)
    - Express varying degrees of suitability for the decision under consideration
    - Referred to as Factors

10

Binary Criterion

Continuous Criterion

11

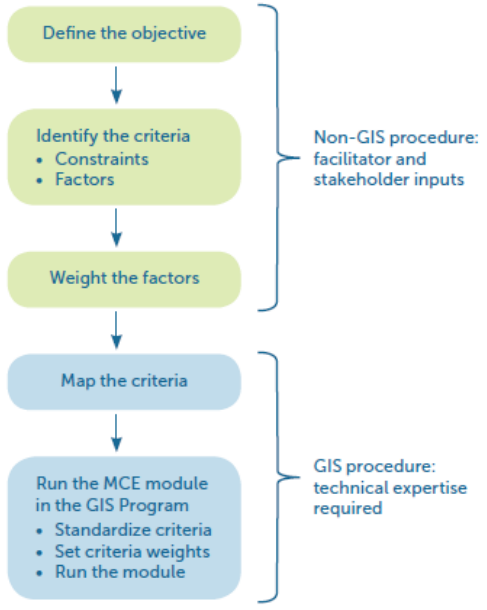
Lopez-Marrero et al, 2011

MCE Process

What	How	Who
Decision Making	Criteria	Stakeholders Government Policy Makers Researchers
	Weighing Criteria	

12

MCE Process



Lopez-Marrero et al, 2011

3

MCE Process

- Decisions:
  - a choice between alternatives
- Criteria:
  - Factors: *enhances* or *detracts* from the suitability of a land use alternative (e.g. distance from a road)
  - Constraint: limits the alternatives
- Objective:
  - some characteristic that the solution must possess (a positive constraint)

14

MCE Process

- Determine criteria
  - Oversimplification of the decision problem could lead to too few criteria being used
  - Using a large number of criteria reduces the influence of any one criteria
  - Often proxies must be used since the criteria of interest may not be determinable

15

MCE Process

- Determine the weights
  - A decision is the result of a comparison of one or more alternatives with respect to one or more criteria that we consider relevant for the task at hand.
  - Among the relevant criteria we consider some as more important and some as less important; this is equivalent to assigning weights to the criterion according to their relative importance.

16

Six factors were identified as being important to include in the analysis and production of the map to identify areas to protect from urban expansion around El Yunque: distance from rivers (D River), distance from El Yunque National Forest boundaries (D EYNF), forest land cover (Forest), northeast portion of the study area (NE Quad), landscape connectivity (Land Conn), and wetland cover (Wetland). Among pairs of factors, the factor in each cell was considered more important by participants in the indentification of such areas.

The final matrix showing the factors' pairwise comparisons and the resulting score for each factor is shown below:

FACTOR	D RIVER	D EYNF	FOREST	NE QUAD	LAND CONN	WETLAND	SCORE
D RIVER		D River	D River	D River	D River	Wetland	4
D EYNF			D EYNF	NE Quad	D EYNF	D EYNF	3
FOREST				NE Quad	Land Conn	Forest	1
NE QUAD					Land Conn	NE Quad	3
LAND CONN						Land Conn	3
WETLAND							1

17

Lopez-Marrero et al, 2011

The weight for a factor was calculated by dividing the score of that factor by the total score of all factors. In this study, the sum of all the scores was 15. To calculate the weight for the factor "distance from river," the score of that factor (4) was divided by 15. The weights for all other factors were calculated by using the same prodecure. All weights should add up to 1. The calculated weight for each factor was as follows:

FACTOR	WEIGHT
Distance from rivers	0.2667
Distance from El Yunque	0.2000
Forest land cover	0.0667
Northeast quadrant	0.2000
Landscape connectivity	0.2000
Wetland cover	0.0667
Total	**1.00**

18

Lopez-Marrero et al, 2011

MCE Process

- Sensitivity Analysis
  - Adjust the weights of the factors to determine the sensitivity of the solution to minor changes
    - Consider the choice of the criteria (e.g. why included?)
  - Assesses the reliability of data: how stable is the final result / solution?
    - Often the choice for weighting factors is subjective
    - Will the overall solution change if you use other weighing factors?

19

Brief Exercise

20

References

- References
  - López-Marrero, T., González-Toro, A., Heartsill-Scalley, T., Hermansen-Báez, L.A. (2011) Multi-Criteria Evaluation and Geographic Information Systems for Land-Use Planning and Decision Making. [Guide]. Gainesville, FL: USDA Forest Service, Southern Research Station. 8 p.

- Further Reading
  - R. Eastman (1999) Multi-criteria evaluation and GIS. P. Longley, M.F. Goodchild, D.J. Maguire, D. Rhind (Eds.), Geographical information systems, pp. 493–502
  - Urška Demšar, Paul Harris, Chris Brunsdon, A. Stewart Fotheringham & Sean McLoone (2012): Principal Component Analysis on Spatial Data: An Overview, Annals of the Association of American Geographers