Model Fitting

Structural Equation Models
Multidimensional Scaling

Basic concepts of iterative fit

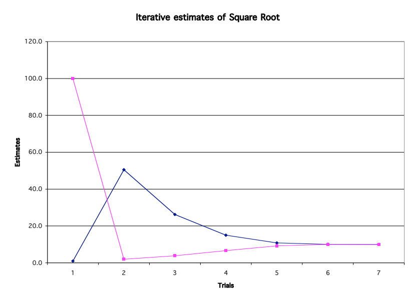
- Most classical statistics (e.g. means, variances, regression slopes) may be found by algebraic solutions of closed form expressions
- More recent statistics are the results of iteratively fitting a model until some criterion is either minimized or maximized.

Simple example: the square root

Target 100

Trial	guess	fit	diff
1	1.0	100.0	-99.0
2	50.5	2.0	48.5
3	26.2	3.8	22.4
4	15.0	6.7	8.4
5	10.8	9.2	1.6
6	10.0	10.0	0.1
7	10.0	10.0	0.0

Iteratively estimating the square root of 100



Applications: Factor Analysis

x1	1.00		
x2	0.70	1.00	
x3	0.60	0.58	1.00

Iterative Fit

x1	x2	x3	fit
1.000	1.000	1.000	0.4264000
1.000	0.800	1.000	0.2180000
1.000	0.800	0.800	0.0536000
0.800	0.800	0.800	0.0088000
0.800	0.800	0.700	0.0056000
0.800	0.800	0.750	0.0040000
0.850	0.800	0.750	0.0022000
0.850	0.800	0.700	0.0008250
0.850	0.800	0.710	0.0005625
0.851	0.823	0.705	0.0000000

Fitted model

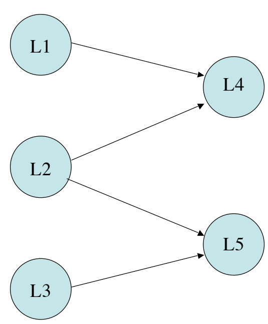
	F1
x1	0.851
x2	0.823
x3	0.705

0.72	0.70	0.60
0.70	0.68	0.58
0.60	0.58	0.50

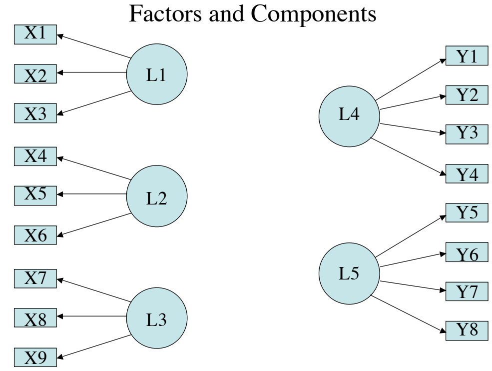
Structural Equation Models

- Basic concept is to apply a measurement model to a structure (regression) model
- Generically known as SEM, particular programs are LISREL, EQS, RAMONA, RAM-path, Mx, sem
- May be used for confirmatory factor analysis as well as sem.

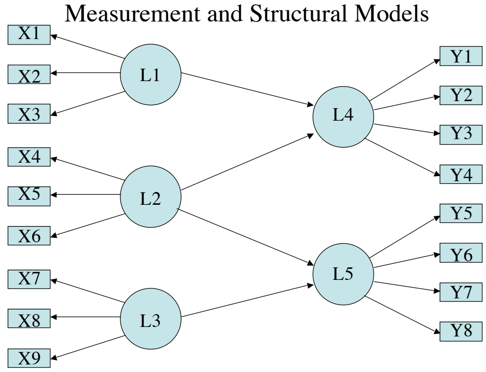
Theory as organization of constructs



Techniques of Data Reduction: Factors and Components



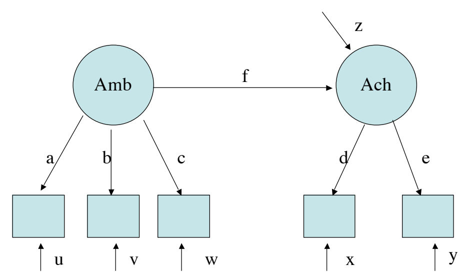
Structural Equation Modeling: Combining Measurement and Structural Models



SEM problem (Loehlin 2.5)

Ach1	Ach2	Amb1	Amb2	Amb3
1	0.6	0.3	0.2	0.2
0.6	1	0.2	0.3	0.1
0.3	0.2	1	0.7	0.6
0.2	0.3	0.7	1	0.5
0.2	0.1	0.6	0.5	1

Ambition and Achievement



R code for sem for Loehlin 2.5

First enter the correlation matrix

```
#Loehlin problem 2.5
 obs.var2.5 = c('Ach1',  'Ach2',  'Amb1',  'Amb2',  'Amb3')
 R.prob2.5 = matrix(c(
  1.00 ,  .60  , .30,  .20,   .20,
   .60,  1.00,   .20,   .30,   .10,
   .30,   .20,  1.00,   .70,   .60 ,
   .20,   .30,   .70,  1.00,   .50,
   .20,   .10,   .60,  .50,  1.00), ncol=5,byrow=TRUE)
```

R code for sem -Ram notation

```
model2.51=matrix(c(
        'Ambit ->  Amb1',      'a', NA,
        'Ambit -> Amb2' ,      'b', NA,
        'Ambit -> Amb3' ,      'c', NA,
        'Achieve -> Ach1',     'd', NA,
        'Achieve -> Ach2',     'e', NA,
        'Ambit -> Achieve',   'f', NA,
        'Amb1 <-> Amb1' ,      'u', NA,
        'Amb2 <-> Amb2' ,      'v', NA,
        'Amb3 <-> Amb3' ,      'w', NA,
        'Ach1 <-> Ach1' ,      'x', NA,
        'Ach2 <-> Ach2' ,      'y', NA,
        'Achieve <-> Achieve',  NA, 1,
        'Ambit <-> Ambit',      NA, 1),
        ncol=3, byrow=TRUE)
```

Run the R code and show results

```
sem2.5= sem(model2.5,R.prob2.5,60, obs.var2.5)
 summary(sem2.5,digits=3)

Model Chisquare =  9.74   Df =  4 Pr(>Chisq) = 0.0450
Goodness-of-fit index =  0.964
Adjusted goodness-of-fit index =  0.865
RMSEA index =  0.120   90 % CI: (0.0164, 0.219)
BIC =  -15.1

Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-5.77e-01 -3.78e-02 -2.04e-06  4.85e-03  3.87e-05  1.13e+00
```

What are the parameters?

```
Parameter Estimates
  Estimate Std Error z value Pr(>|z|)
a   0.920    0.0924   9.966 0.00e+00    Amb1 <--- Ambit
b   0.761    0.0955   7.974 1.55e-15    Amb2 <--- Ambit
c   0.652    0.0965   6.753 1.45e-11    Amb3 <--- Ambit
d   0.879    0.1762   4.986 6.16e-07  Ach1 <--- Achieve
e   0.683    0.1509   4.525 6.03e-06  Ach2 <--- Achieve
f   0.356    0.1138   3.127 1.76e-03 Achieve <--> Ambit
u   0.153    0.0982   1.557 1.20e-01     Amb1 <--> Amb1
v   0.420    0.0898   4.679 2.88e-06     Amb2 <--> Amb2
w   0.575    0.0949   6.061 1.35e-09     Amb3 <--> Amb3
x   0.228    0.2791   0.816 4.15e-01     Ach1 <--> Ach1
y   0.534    0.1837   2.905 3.67e-03     Ach2 <--> Ach2

Iterations =  26
```

Problems in interpretation

- Model fit does not imply best model
- Consider alternative models
  - Reverse the arrows of causality, nothing happens
  - Range of alternative models
  - Nested models can be compared
  - Non-nested alternative models might be better

SEM points to consider

- Goodness of fit statistics
  - Statistical indices of size of residuals as compared to null model are sensitive to sample size
  - Comparisons of nested models
- Fits get better with more parameters -> development of df corrected fits
- Inspect residuals to see what is not being fit
- Avoid temptation to ‘fix’ model based upon results, or, at least be less confident in meaning of good fit

Applications of SEM techniques

- Confirmatory factor analysis
  - Does a particular structure fit the data
- Growth models (growth curve analysis)
- Multiple groups
  - Is the factor structure the same across groups
  - Is the factor structure the same across time

Multidimensional Scaling

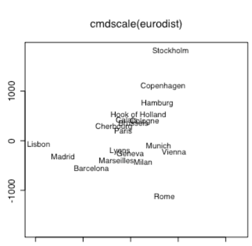
- Application of metric or non-metric scaling
- Metric scaling:
  - Find dimensional representation of observed distances (e.g., latitude and longitude)
  - Strong assumption of data and metric
- Non-metric scaling
  - Scaling to minimize a criterion insensitive to ordinal transformations

Distances between cities

	Athen	Barcelona	Brussels	Calais	Cherburg	Cologne	Copenhage	Geneva	Gilbralter	Hamburg
Barcelona	3313									
Brussels	2963	1318								
Calais	3175	1326	204							
Cherbourg	3339	1294	583	460						
Cologne	2762	1498	206	409	785					
Copenhagen	3276	2218	966	1136	1545	760				
Geneva	2610	803	677	747	853	1662	1418			
Gibralta	4485	1172	2256	2224	2047	2436	3196	1975		
Hamburg	2977	2018	597	714	1115	460	460	1118	2897	
Hook of Holkan	3030	1490	172	330	731	269	269	895	2428	550

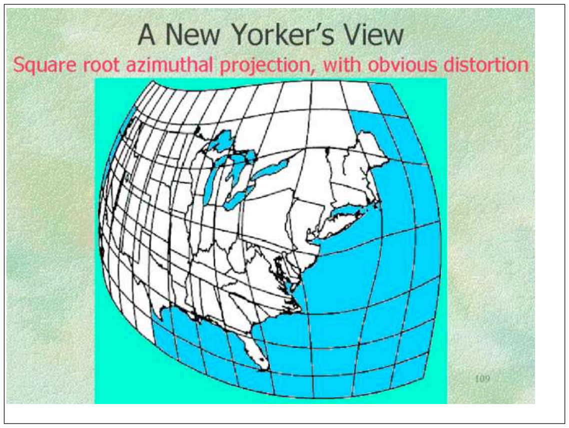
What is the best representation of these distances in a two dimensional space?

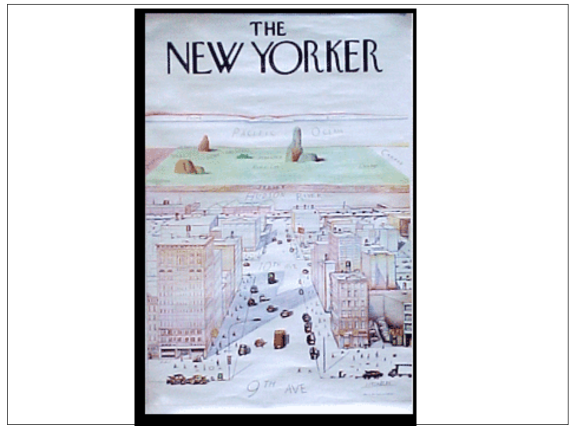
Scaling of European Cities

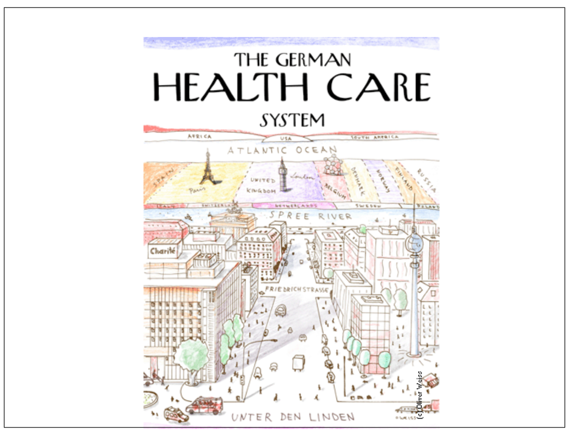


Individual Differences in MDS INDSCAL

- Consider individual differences in MDS
  - Each individual applies a unique weighting to the MDS dimensions
- Solve for Group space as well as individual weights to be applied to the group space


A New Yorker's View
Square root azimuthal projection, with obvious distortion
109


THE
NEW YORKER
9TH AVE


THE GERMAN
HEALTH CARE
SYSTEM
ATLANTIC OCEAN
SPREE RIVER
Charité
FRIEDRICHSTRASSE
UNTER DEN LINDEN