Multidimensional Scaling:
DATA & Distance (see TUG A2.1)

Distance measure:

$$d(i,i) >= 0 \quad \text{Non-negativity}$$

$$d(i,j) = d(j,i) \quad \text{Symmetry}$$

$$\text{if } d(i,j) > d(k,l) \text{ and } d(k,l) > d(m,n)$$

$$\text{then } d(i,j) > d(m,n) \quad \text{Transitivity}$$

$$d(i,k) + d(j,k) >= d(i,j) \quad \text{Triangle Inequality}$$

Multidimensional Scaling:

DATA: Dis/similarity Measures

- Proximity or Similarity measure (***s***)

matched by <u>closeness</u>

the higher the value, the more similar

- Dissimilarity measure (δ)

matched by <u>distance</u>

The higher the value, the more different

- Conversion:

$$\text{Has the form}: x = [\max(x) - x] + 1$$

$$s = \max(\delta) - \delta$$

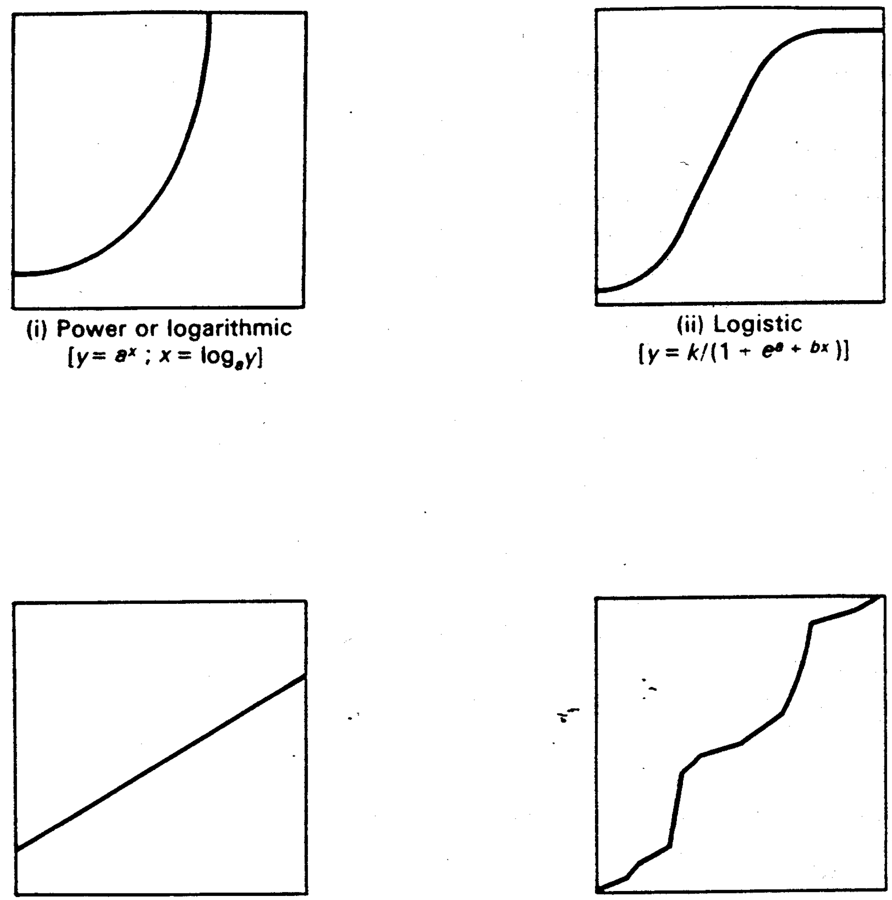
$$\delta = \max(s) - s$$

Multidimensional Scaling:

Transformations & Levels of Measurement (TUG, ch2)

	Name	*Permissible Transformat*	*Preserves*	*Direct Data*	*Aggregate measures*
	Absolute	none	All (true 0)	--	
M	**Ratio**	similarity	(a/b) (Unit change)	Counts, rates, **distances**	
	Interval	linear	(a-b) ..& origin	Rating	PM Correln. ρ Covariance σ
NM	**Ordinal**	**monotonic**	**order**	**Ranking**	**Kendall's** τ **G & K 's** γ
	Nominal (inc dichot.)	**isotonic**	**categories**	**Sorting**	χ**-sq based Pairbonds**

Multidimensional Scaling: <u>Monotonic</u> Transformations



(i) Power or logarithmic
[$y = a^x$; $x = \log_a y$]

(ii) Logistic
[$y = k/(1 + e^{a + bx})$]

(iii) Linear
[$y = a + bx$]

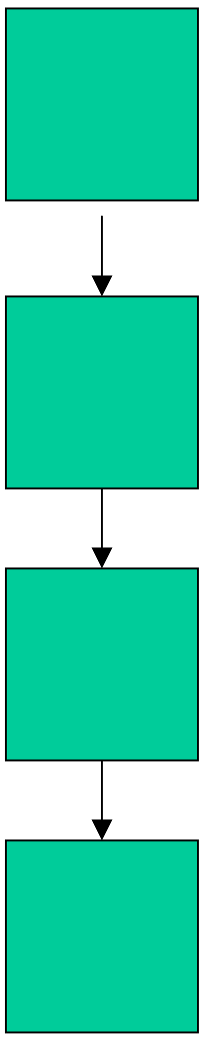
(iv) Irregular monotonic

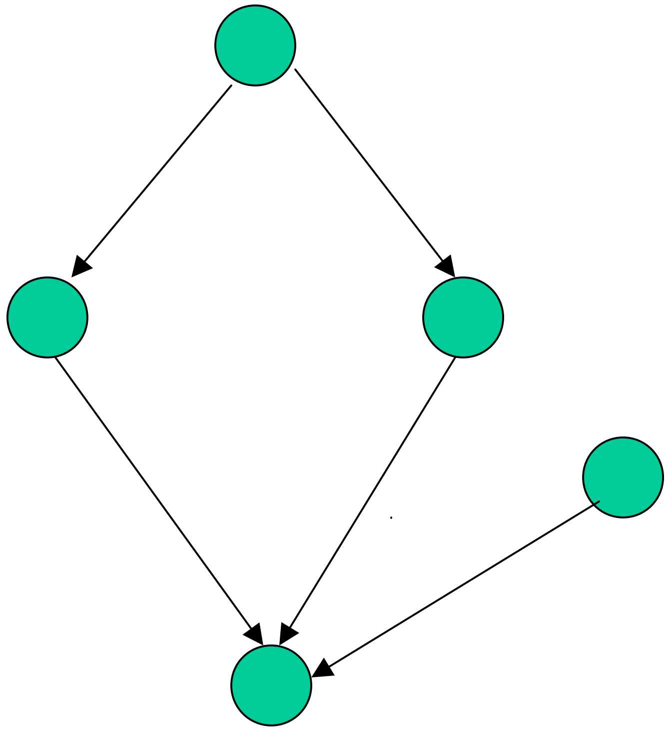
Figure 5.1 *Monotone curves*

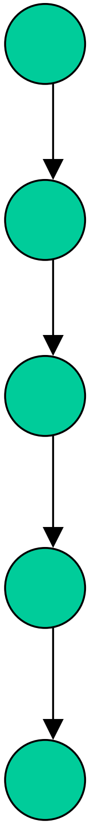
Types of Order:

Weak **Partial** **Strict**

Types of Order







Multidimensional Scaling:

Distances & Maps

Two (related) problems ..

- **Given a configuration/map: find the distances (X->D)**

  (Easy: Given a co-ordinate system, calculate pairwise distances according to formula; scaling factor/unit is arbitrary)

And a more difficult one: <u>the scaling problem…</u>

- **Given a set of distances, find the map/configuration that generates them**

  (D -> X)

  … Rather more difficult …

Two Solutions …

1. **"Classic" Metric MDS:** $\delta \cong d \Rightarrow X$

   i.e. Data are treated as if they are/approximate distances (see TUG A5.2)

It can be done … more or less … under certain circumstances … like

- **Error-free data**
- **Good estimate of the "additive constant"**
- **Recovery up to a similarity transformation (free rotation, location of axes, scaling factor)**

- **BUT: The solution is NOT ROBUST!**

Solution (2): **δ . M(d) Y X**

(equals: Non-metric ordinal scaling, 1962 onwards)

But …   ***Can this be done*****?**

< **NO** (Torgerson initially, and rest of the world)

< **Perhaps,** (Coombs, Abelson)

< **YES!** (Shepard, Kruskal, Hayashi, Guttman**)**

... ***what, 'recover' quantitative data from ordinal input?***

- ***What is this, a magic quantifier?***

- ***What's the trick, then?***

▸ **... here's how and here's why**, Shepard 1962

"Ordinal constraints, if imposed in sufficient number, act as metric constraints and guarantee in the limit a metric solution"

... plus a program to do it! Shepard 1962

... and here's how you justify it as really monotonic and in LS terms (Kruskal 1964)

... so that even statisticians love it (MLS, 1977, Ramsay)

AND SO ... ! MDS WAS BORN!!