Figure 2.2 Perfect Spatial Voting in One Dimension

```
-1---------0--------+1
```

Legislators	X_1		X_2		X_3		X_4		X_5		X_6
Cutpoints		Z_1		Z_2		Z_3		Z_4		Z_5	
1	Y		N		N		N		N		N
2	Y		Y		N		N		N		N
3	Y		Y		Y		N		N		N
4	Y		Y		Y		Y		N		N
5	Y		Y		Y		Y		Y		N

Figure 2.3 Recovering the Legislator Points

	Roll Calls				
Legislators	1	3	5	4	2
Five	N	Y	N	N	N
Six	N	Y	Y	N	N
Four	N	Y	N	Y	N
One	Y	N	N	Y	Y
Two	N	N	N	Y	Y
Three	N	N	N	Y	N

Agreement Scores

1.0					
.8	1.0				
.8	.6	1.0			
.2	.0	.4	1.0		
.4	.2	.6	.8	1.0	
.6	.4	.8	.6	.8	1.0

Squared Distances

.00					
.04	.00				
.04	.16	.00			
.64	1.00	.36	.00		
.36	.64	.16	.04	.00	
.16	.36	.04	.16	.04	.00

Double-Centered Matrix

.09					
.15	.25				
.03	.05	.01			
-.15	-.25	-.05	.25		
-.09	-.15	-.03	.15	.09	
-.03	-.05	-.01	.05	.03	.01

Legislator Points

$X_5 = .3$

$X_6 = .5$

$X_4 = .1$

$X_1 = -.5$

$X_2 = -.3$

$X_3 = -.1$

Appendix: Proof that if Voting Is Perfect in One Dimension, then the First Eigenvector Extracted from the Double-Centered Transformed Agreement Score Matrix has the Same Rank Ordering as the True Data

Notation and Definitions

Let the true ideal points of the p legislators be denoted as $\tilde{X}_1,\ \tilde{X}_2,\ \ldots,\ \tilde{X}_p$. Without loss of generality, let the ordering of the true ideal points of the legislators on the dimension from left to right be:

$$\tilde{X}_1 \leq \tilde{X}_2 \leq \tilde{X}_3 \leq \ldots \leq \tilde{X}_p$$

Let q be the number of *non-unanimous* roll call votes with q > 0 and let the cutting point for the j^{th} roll call be Z_j. Voting is *perfect*. That is, all legislators are sincere voters, and all legislators to the left of a cutting point vote for the same alternative and all legislators to right of a cutting point vote for the opposite alternative. For example, if all legislators to the left of Z_j vote "Nay", then all legislators to the right of Z_j vote "Yea". Without loss of generality we can assume that every legislator to the left of Z_j votes "Yea" and every legislator to the right of Z_j votes "Nay". That is, the "polarity" of the roll call does not affect the analysis below.

Let k_1 be the number of cutting points between legislators 1 and 2, k_2 be the number of cutting points between legislators 2 and 3, and so on, with k_{p-1} being the number of cutting points between legislators p-1 and p. Hence

$$q = \sum_{i=1}^{p-1} k_i > 0 \qquad \text{(A2.1)}$$

The *agreement score* between two legislators is the simple proportion of roll calls in which they vote for the same outcome. Hence the agreement score between legislators 1 and 2 is simply $\frac{q - k_1}{q}$, because 1 and 2 agree on all roll calls except for those with cutting points between them. Similarly, the agreement score between legislators 1 and 3 is $\frac{q - k_1 - k_2}{q}$ and the agreement score between legislators 2 and 3 is $\frac{q - k_2}{q}$. In general, for two legislators X_a and X_b where a≠b, the agreement score is:

$$A_{ab} = \frac{q - \sum_{i=a}^{b-1} k_i}{q} \qquad \text{(A2.2)}$$

The agreement scores can be treated as Euclidean distances by simply subtracting them from 1. That is:

$$d_{ab} = 1 - A_{ab} = 1 - \frac{q - \sum_{i=a}^{b-1} k_i}{q} = \frac{\sum_{i=a}^{b-1} k_i}{q} \qquad \text{(A2.3)}$$

These definitions allow me to state the following theorem:

***Theorem*: If Voting Is Perfect in One Dimension, then the First Eigenvector Extracted from the Double-Centered p by p Matrix of Squared Distances from Equation (A2.3) Has at Least the Same Weak Monotone Rank Ordering as the Legislators.**

***Proof*:** The d's computed from equation (A2.3) satisfy the three axioms of distance: they are non-negative because by (A2.2) $0 \leq A_{ab} \leq 1$ so that $0 \leq d_{ab} \leq 1$; they are symmetric,

$d_{ab} = d_{ba}$; and they satisfy the triangle inequality. To see this, consider any triplet of points $X_a < X_b < X_c$. The distances are:

$$d_{ab} = \frac{\sum_{i=a}^{b-1} k_i}{q} \quad \text{and} \quad d_{bc} = \frac{\sum_{i=b}^{c-1} k_i}{q} \quad \text{and} \quad d_{ac} = \frac{\sum_{i=a}^{c-1} k_i}{q}$$

Hence

$$d_{ac} = d_{ab} + d_{bc} \qquad \text{(A2.4)}$$

Because all the triangle inequalities are *equalities*, in Euclidean geometry this implies that X_a, X_b, and X_c *all lie on a straight line* (Borg and Groenen, 1997, ch. 18).

Because all the triangle inequalities are equalities and all triplets of points lie on a straight line, the distances computed from (A2.2) can be directly written as distances between points:

$$d_{ab} = \frac{\sum_{i=a}^{b-1} k_i}{q} = |X_a - X_b| \qquad \text{(A2.5)}$$

where $d_{aa} = 0$. The p by p matrix of squared distances is:

$$\mathbf{D} = \begin{bmatrix} 0 & (X_2 - X_1)^2 & . & . & . & (X_p - X_1)^2 \\ (X_1 - X_2)^2 & 0 & . & . & . & (X_p - X_2)^2 \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ (X_1 - X_p)^2 & (X_2 - X_p)^2 & . & . & . & 0 \end{bmatrix} \qquad \text{(A2.6)}$$

To recover the X's, simply double-center **D** and perform an eigenvalue-eigenvector decomposition. The first eigenvector is the solution. To see this:

Let the mean of the j^{th} column of **D** be $d_{.j}^2 = \frac{\sum_{i=1}^{p} d_{ij}^2}{p} = X_j^2 - 2X_j\bar{X} + \frac{\sum_{i=1}^{p} X_i^2}{p}$.

Let the mean of the i^th row of **D** be $d_{i.}^2 = \frac{\sum_{j=1}^{p} d_{ij}^2}{p} = X_i^2 - 2X_i\bar{X} + \frac{\sum_{j=1}^{p} X_j^2}{p}$.

Let the mean of the matrix **D** be $d_{..}^2 = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p} d_{ij}^2}{p^2} = \frac{\sum_{j=1}^{p} X_j^2}{p} - 2\bar{X}^2 + \frac{\sum_{i=1}^{p} X_i^2}{p}$.

Where $\bar{X} = \frac{\sum_{i=1}^{p} X_i}{p}$ is the mean of the X_i.

The matrix **D** is double-centered as follows: from each element subtract the row mean, subtract the column mean, add the matrix mean, and divide by –2; that is,

$$y_{ij} = \frac{(d_{ij}^2 - d_{.j}^2 - d_{i.}^2 + d_{..}^2)}{-2} = (X_i - \bar{X})(X_j - \bar{X})$$

This produces the p by p symmetric positive semidefinite matrix **Y**:

$$\mathbf{Y} = \begin{bmatrix} X_1 - \bar{X} \\ X_2 - \bar{X} \\ . \\ . \\ . \\ X_p - \bar{X} \end{bmatrix} \begin{bmatrix} X_1 - \bar{X} & X_2 - \bar{X} & . & . & . & X_p - \bar{X} \end{bmatrix} \qquad \text{(A2.7)}$$

Because **Y** is symmetric with a rank of one, its eigenvalue-eigenvector decomposition is simply:

$$\mathbf{Y} = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_p \end{bmatrix} \lambda_1 \begin{bmatrix} u_1 & u_2 & . & . & . & u_p \end{bmatrix} \qquad \text{(A2.8)}$$

Hence the solution is

$$\begin{bmatrix} X_1 - \bar{X} \\ X_2 - \bar{X} \\ . \\ . \\ . \\ X_p - \bar{X} \end{bmatrix} = \sqrt{\lambda_1} \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_p \end{bmatrix}$$

Because, without loss of generality, the origin can be placed at zero, that is, $\bar{X} = 0$, the solution can also be written as:

$$\begin{bmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_p \end{bmatrix} = \sqrt{\lambda_1} \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_p \end{bmatrix} \qquad \text{(A2.9)}$$

The points from (A2.9) *exactly reproduce* the distances in (A2.4), the agreement scores in (A2.2), and the original roll call votes. In addition, note that the *mirror image* of the points in (A2.9) (a multiplication by -1) also exactly reproduces the original roll call votes. Furthermore, for any pair of true legislator ideal points $\tilde{X}_a$ and $\tilde{X}_b$ with one or more midpoints between them, $\tilde{X}_a < Z_j < \tilde{X}_b$, the recovered legislator ideal points *must have the same ordering*, $X_a < X_b$. If there are no midpoints between $\tilde{X}_a$ and $\tilde{X}_b$ -- that is, their roll call voting pattern is *identical* -- then the recovered legislator ideal points are identical; $X_a = X_b$. Hence, if there are cutting points between every pair of adjacent legislators, that is, $k_i \geq 1$ for i=1, …, p-1, then the rank ordering of the recovered ideal points is the same as the true rank ordering. If some of the $k_i = 0$, then the recovered ideal points have a weak monotone transformation of the true rank ordering (in other words there are ties, some legislators have the same recovered ideal points).

This completes the proof. **QED**.

Discussion

Note that an interval-level set of points is recovered. But *this is an artifact of the distribution of cutting points*. For example, if $k_1 > k_2$, then $d_{12} > d_{23}$ even if the true coordinates $\tilde{X}_1$, $\tilde{X}_2$, $\tilde{X}_3$ were *evenly spaced*. With perfect one-dimensional voting, the legislator configuration is identified only up to a *weak monotone transformation of the true rank ordering*.

The rank ordering can also be recovered directly from the matrix **Y** given in (A2.7) without performing an eigenvalue-eigenvector decomposition. Note that, with the origin at zero, the diagonal elements of **Y** are simply the legislator coordinates squared. The rank ordering can be recovered by taking the square root of the first diagonal element and then dividing through the first row of the matrix. Note that this sets $X_1 > 0$ and the remaining points are identified vis a vis X_1.

Within the field of psychometrics the basic result of this theorem is well known. In Guttman scaling a *perfect simplex* is essentially the same as a perfect roll call matrix in one dimension. However, a perfect simplex has a natural polarity – for example, the "Yeas" are always on the same side of the cutting points. The theorem above is very similar to Schonemann's (1970) solution for the perfect simplex problem. His solution builds upon Guttman's (1954) analysis of the problem. To my knowledge, no one has stated the result in the form that I did above. Namely, as the solution to a roll call voting problem.

1. Compute the p by p agreement score matrix
2. Convert the agreement score matrix into a matrix of squared distances
3. Double-center the matrix of squared distances
4. Perform an eigenvalue-eigenvector decomposition of the Double-centered transformed agreement score matrix. The first eigenvector is the solution.

Figure 2.4 The Effect of the Number of Cutting points

Agreement Scores

1.000					
.875	1.000				
.875	.750	1.000			
.125	.000	.250	1.000		
.625	.500	.750	.500	1.000	
.750	.625	.875	.375	.875	1.000

Squared Distances

0.000000					
0.015625	0.000000				
0.015625	0.062525	0.000000			
0.765625	1.000000	0.562500	0.000000		
0.140625	0.250000	0.062500	0.250000	0.000000	
0.062525	0.140625	0.015625	0.390625	0.015625	0.000000

Double-Centered Matrix

0.062500					
0.093750	0.140625				
0.031250	0.046875	0.015625			
-0.156250	-0.234375	-0.078125	0.390625		
-0.031250	-0.046875	-0.015625	0.078125	0.015625	
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Legislator Points

$X_5 = .250$

$X_6 = .375$

$X_4 = .125$

$X_1 = -.625$

$X_2 = -.125$

$X_3 = .000$

Figure 2.5 Interest Group Ratings in One Dimension

Liberal   Conservative

-1-----------------0----------------+1

Legislators ADA X_1 X_2 X_3 X_4 X_5 X_6 X_7 ... X_{p-1} X_p

Cutpoints Z_1 Z_2 Z_3 Z_4 ...Z_{20}

Figure 2.6 Unfolded Interest Group Ratings

Liberal   Conservative

-1-----------------0----------------+1

Legislators X_3 X_1 ADA X_2 X_4 X_5 X_6 X_7 ... X_{p-1} X_p

Cutpoints Z_1 Z_2 Z_3 Z_4 Z_5 ... Z_{20}

Relationship Between Simple Quadratic Utility Model and the Simple IRT Model

In the quadratic utility model, the utility of the i^{th} legislator for the Yea and Nay alternatives is simply the negative of the squared distance from the legislator's ideal point to the outcomes:

$$U_{iy} = -(X_i - O_{jy})^2 \text{ and } U_{in} = -(X_i - O_{jn})^2$$

In the spatial voting model, if $U_{iy} > U_{in}$ the legislator votes Yea. Stated another way, if the difference, $U_{iy} - U_{in}$, is positive, the legislator votes Yea. Algebraically:

$$U_{iy} - U_{in} = -(X_i - O_{jy})^2 + (X_i - O_{jn})^2 = 2X_iO_{jy} - O_{jy}^2 - 2X_iO_{jn} + O_{jn}^2 =$$

$$2X_i(O_{jy} - O_{jn}) - (O_{jy}^2 - O_{jn}^2) = 2\gamma_j(X_i - Z_j) \qquad (2.5)$$

where $\gamma_j = (O_{jy} - O_{jn})$ and $2Z_j = (O_{jy} + O_{jn})$. With perfect voting the legislator and the chosen outcome are on the same side of the midpoint. Hence:

$$\begin{aligned} &\text{if } 2\gamma_j (X_i - Z_j) > 0 \text{ Vote Yea} \\ &\text{if } 2\gamma_j (X_i - Z_j) < 0 \text{ Vote Nay} \end{aligned} \qquad (2.6)$$

If $O_{jy} > O_{jn}$ this decision rule simplifies to:

$$\text{if } X_i - Z_j > 0 \text{ Vote Yea}$$

$$\text{if } X_i - Z_j < 0 \text{ Vote Nay}$$

The corresponding formulation for the Rasch model is:

$$\begin{aligned} &\text{if } \beta_jX_i - \alpha_j > 0 \text{ correct answer} \\ &\text{if } \beta_jX_i - \alpha_j < 0 \text{ wrong answer} \end{aligned} \qquad (2.7)$$

where β_j is the *item discrimination parameter* and α_j is the *difficulty parameter*. If the test question is clearly stated so that there is no ambiguity, then by convention $\beta_j = 1$. A poorly constructed and ambiguous question would have a β_j near 0. α_j is simply the level of difficulty on the latent dimension.

Let Y_{ij} be the ith individual's response (1 = correct; 0 = incorrect) to the jth question (item). Rasch (1960) showed that for the one parameter IRT model that individual and question parameters can be separately estimated if the errors in the responses are independent. That is, the individual and question parameters "can be distinguished and estimated separately" (van Schuur, 2003). To see this, assume that:

$$\mathbf{P(Y_{ij}} = 1 \mid \mathbf{X_i}, \boldsymbol{\alpha}_j, \boldsymbol{\beta}_j = 1\mathbf{)} = \frac{\mathbf{e}^{(X_i - \alpha_j)}}{1 + \mathbf{e}^{(X_i - \alpha_j)}}$$

Note that this is the standard logit probability formula.

Now, let m be another index for the number of questions (items). Consider the ratio:

$$\frac{\mathbf{P(Y_{ij}} = 1 \cap \mathbf{Y_{im}} = 0\mathbf{)}}{\mathbf{P(Y_{ij}} = 0 \cap \mathbf{Y_{im}} = 1\mathbf{)}} = \frac{\left[\frac{\mathbf{e}^{(X_i - \alpha_j)}}{1 + \mathbf{e}^{(X_i - \alpha_j)}}\right]\left[\frac{1}{1 + \mathbf{e}^{(X_i - \alpha_m)}}\right]}{\left[\frac{1}{1 + \mathbf{e}^{(X_i - \alpha_j)}}\right]\left[\frac{\mathbf{e}^{(X_i - \alpha_m)}}{1 + \mathbf{e}^{(X_i - \alpha_m)}}\right]} = \frac{\mathbf{e}^{(X_i - \alpha_j)}}{\mathbf{e}^{(X_i - \alpha_m)}} = \mathbf{e}^{(\alpha_m - \alpha_j)}$$

Taking the natural log yields:

$\alpha_m - \alpha_j$

Which is independent of individual i.