MULTIDIMENSIONAL α-NOMINATE

by

Keith T. Poole

Department of Political Science

School of Public and International Affairs

University of Georgia

ABSTRACT

In this paper I show a method to generalize α-NOMINATE to more than one dimension. In one dimension α-NOMINATE is a mixture model that can distinguish between Quadratic ($\alpha=0$) and Gaussian ($\alpha=1$) deterministic utility. The widely used IRT model is isomorphic with a random utility model with Quadratic deterministic utility and normal error. However, as shown in Poole (2001) if the correct geometry is used the QN model can be identified in any number of dimensions. In the α-NOMINATE framework I switch the geometry to that used by optimal classification (2000, 2005) and QN (2001). This allows α to be estimated in any number of dimensions.

1. Introduction

In this paper I combine the geometry of Optimal Classification (Poole, 2000; 2005) with the α-NOMINATE model developed by Carroll, et al. (2013). This change in geometry generalizes the α parameter so that it is identified in any number of dimensions.

2. The Alpha-NOMINATE Model

Suppose there are n individuals, q stimuli, and s dimensions indexed by i=1,..,n, j=1,...,q, and k=1,…,s, respectively. To simplify the exposition I assume that each stimulus is a binary choice between two outcomes which I will denote as "y" or "n". Let X denote individuals and O denote outcomes. Specifically, X_{ik} is the i^{th} individual's ideal point on dimension k, O_{jky} and O_{jkn} are the outcome locations for the j^{th} stimulus on the k^{th} dimension, and $\mathbf{d}^2_{ijky}$ and $\mathbf{d}^2_{ijkn}$ are the squared distances of individual i to outcomes "y" and "n", respectively, on the k^{th} dimension for stimulus j:

$$\begin{aligned} d^2_{ijky} &= (X_{ik} - O_{jky})^2 \\ d^2_{ijkn} &= (X_{ik} - O_{jkn})^2 \end{aligned} \qquad (1)$$

The major probabilistic models of binary choice are based on the ***random utility model*** developed by McFadden (1976). In the random utility model, an individual's overall utility for choosing an outcome is the sum of a deterministic utility and a random error. Hence, individual i's utility for the "y" and "n" outcomes on stimulus j is:

$$
\begin{aligned}
U_{ijy} &= u_{ijy} + \varepsilon_{ijy} \\
U_{ijn} &= u_{ijn} + \varepsilon_{ijn}
\end{aligned}
\qquad (2)
$$

where u_{ijy} and u_{ijn} are the deterministic portion of the utility function and ε_{ijy} and ε_{ijn} are the stochastic or random portion of the utility function. I assume that ε_{ijy} and ε_{ijn} are a random sample of size two from a normal distribution with mean zero and variance one-half. The difference between the two errors has a standard normal distribution; that is

$$\varepsilon_{ijn} - \varepsilon_{ijy} \sim N(0, 1)$$

and the distribution of the difference between the overall utilities is

$$U_{ijy} - U_{ijn} \sim N(u_{ijy} - u_{ijn}, 1)$$

Hence the probability that individual i votes "y" on the j^{th} stimulus can be rewritten as:

$$\mathbf{P}_{ijy} = \mathbf{P}(U_{ijy} > U_{ijn}) = \mathbf{P}(\varepsilon_{ijn} - \varepsilon_{ijy} < u_{ijy} - u_{ijn}) = \Phi[u_{ijy} - u_{ijn}] \quad (3)$$

I follow Carroll et al. (2013) in the specification of the deterministic utility function. They show a simple method of nesting quadratic deterministic utility within Gaussian deterministic utility. Specifically, let the deterministic utility be the NOMINATE model of Poole and Rosenthal (1985; 1997):

$$u_{ijy} = \beta e^{\left(-\frac{1}{2}\sum_{k=1}^{s}\left(X_{ik}-O_{jky}\right)^2\right)} = \beta e^{\left(-\frac{1}{2}d_{ijy}^2\right)}$$
$$u_{ijn} = \beta e^{\left(-\frac{1}{2}\sum_{k=1}^{s}\left(X_{ik}-O_{jkn}\right)^2\right)} = \beta e^{\left(-\frac{1}{2}d_{ijn}^2\right)} \quad (4)$$

Now, use the exponential series to obtain:

$$u_{ijy} = \beta e^{\left(-\frac{1}{2}d_{ijy}^2\right)} = \beta\sum_{m=0}^{\infty}\left(\frac{\left(-\frac{1}{2}d_{ijy}^2\right)^m}{m!}\right) = \beta\left(1 - \frac{d_{ijy}^2}{2} + \frac{d_{ijy}^4}{2^2 * 2!} - \frac{d_{ijy}^6}{2^3 * 3!} + \frac{d_{ijy}^8}{2^4 * 4!} - \ldots\right)$$
$$u_{ijn} = \beta e^{\left(-\frac{1}{2}d_{ijn}^2\right)} = \beta\sum_{m=0}^{\infty}\left(\frac{\left(-\frac{1}{2}d_{ijn}^2\right)^m}{m!}\right) = \beta\left(1 - \frac{d_{ijn}^2}{2} + \frac{d_{ijn}^4}{2^2 * 2!} - \frac{d_{ijn}^6}{2^3 * 3!} + \frac{d_{ijn}^8}{2^4 * 4!} - \ldots\right)$$

The difference between the two utilities simplifies to:

$$\boldsymbol{u}_{ijy}-\boldsymbol{u}_{ijn}=\boldsymbol{\beta}\left\{-\left(\frac{\boldsymbol{d}_{ijy}^2-\boldsymbol{d}_{ijn}^2}{2}\right)+\boldsymbol{\alpha}\left[\left(\frac{\boldsymbol{d}_{ijy}^4-\boldsymbol{d}_{ijn}^4}{2^2*2!}\right)-\left(\frac{\boldsymbol{d}_{ijy}^6-\boldsymbol{d}_{ijn}^6}{2^3*3!}\right)+\left(\frac{\boldsymbol{d}_{ijy}^8-\boldsymbol{d}_{ijn}^8}{2^4*4!}\right)-\dots\right]\right\}=$$

$$\boldsymbol{\beta}\left\{-\left(\frac{\boldsymbol{d}_{ijy}^2-\boldsymbol{d}_{ijn}^2}{2}\right)+\boldsymbol{\alpha}\left[\left(\boldsymbol{e}^{\left(-\frac{1}{2}\boldsymbol{d}_{ijy}^2\right)}-\boldsymbol{e}^{\left(-\frac{1}{2}\boldsymbol{d}_{ijn}^2\right)}\right)+\left(\frac{\boldsymbol{d}_{ijy}^2-\boldsymbol{d}_{ijn}^2}{2}\right)\right]\right\} \quad (5)$$

If α=0 we have quadratic deterministic utility and if α=1 we have Gaussian deterministic utility (Carroll et al. 2013).

Unfortunately, using the formulation in (5) does not allow the estimation of α in more than one dimension. To see this, note that the difference between the deterministic utilities when α=0 is:

$$-\left(\frac{d_{ijy}^2-d_{ijn}^2}{2}\right)=-\frac{1}{2}\sum_{k=1}^{s}(X_{ik}-O_{jky})^2+\frac{1}{2}\sum_{k=1}^{s}(X_{ik}-O_{jkn})^2$$

$$=-\frac{1}{2}\sum_{k=1}^{s}X_{ik}^2+\sum_{k=1}^{s}X_{ik}O_{jky}-\frac{1}{2}\sum_{k=1}^{s}O_{jky}^2+\frac{1}{2}\sum_{k=1}^{s}X_{ik}^2-\sum_{k=1}^{s}X_{ik}O_{jkn}+\frac{1}{2}\sum_{k=1}^{s}O_{jkn}^2$$

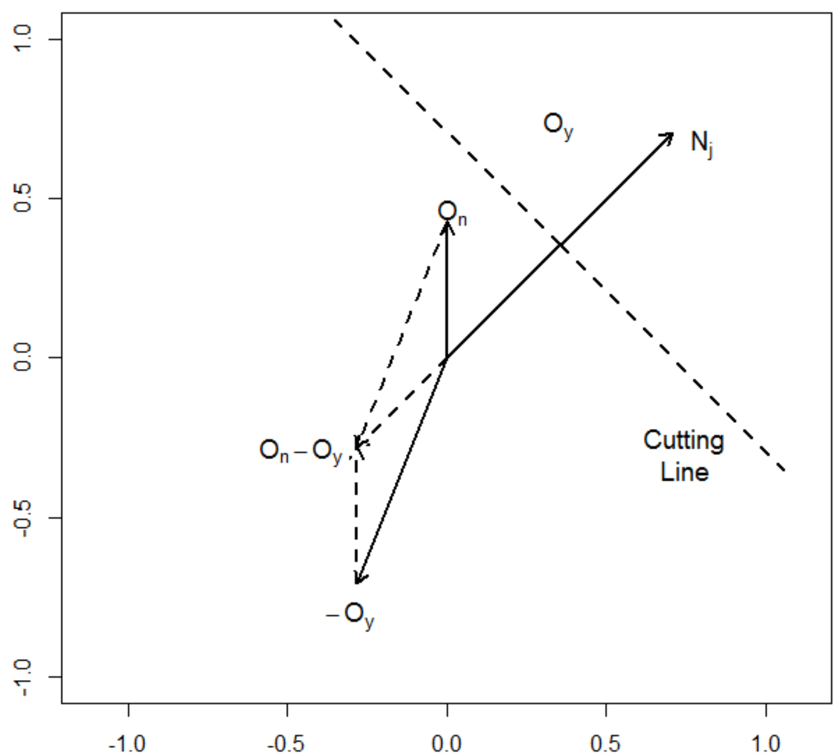
$$-\sum_{k=1}^{s}X_{ik}\left(O_{jkn}-O_{jky}\right)+\frac{1}{2}\sum_{k=1}^{s}\left(O_{jkn}-O_{jky}\right)\left(O_{jkn}+O_{jky}\right) \quad (6)$$

Let $\boldsymbol{N}_j$ be the unit length normal vector to the plane ($\boldsymbol{N}_j'\boldsymbol{N}_j=1$) that separates the two outcome points at the midpoint of the line that joins them (see Figure 1). This plane is known as the ***cutting plane*** between the two outcomes. The difference between the two outcome points, is the s by 1 vector:

$$\mathbf{O}_{jn} - \mathbf{O}_{jy} = \begin{bmatrix} O_{j1n} - O_{j1y} \\ O_{j2n} - O_{j2y} \\ \vdots \\ O_{jsn} - O_{jsy} \end{bmatrix}$$

This vector is equal to a constant times the normal vector as shown in Figure 1:

Figure 1: Vector Difference of Outcome Coordinates



(see Poole, 2001; 2005). Specifically,

$$\gamma_j \mathbf{N}_j = \mathbf{O}_{jn} - \mathbf{O}_{jy} \qquad (7)$$

Where γ_j is the directional distance between the two outcome points:

$$\gamma_j = \left[\sum_{k=1}^{s}(\boldsymbol{O}_{jkn} - \boldsymbol{O}_{jky})^2\right]^{\frac{1}{2}} \;\; \textbf{\textit{if}}\; \boldsymbol{O}'_{jn}\boldsymbol{N}_j > \boldsymbol{O}'_{jy}\boldsymbol{N}_j$$
$$\gamma_j = -\left[\sum_{k=1}^{s}(\boldsymbol{O}_{jkn} - \boldsymbol{O}_{jky})^2\right]^{\frac{1}{2}} \;\; \textbf{\textit{if}}\; \boldsymbol{O}'_{jn}\boldsymbol{N}_j < \boldsymbol{O}'_{jy}\boldsymbol{N}_j \qquad (8)$$

Now, the midpoint between the Yea and Nay outcomes for roll call j is the s by 1 vector:

$$\boldsymbol{M}_j = \frac{\left(\boldsymbol{O}_{jy} + \boldsymbol{O}_{jn}\right)}{2}$$

This allows equation (6) to be rewritten as the vector equation:

$$-\left(\frac{\boldsymbol{d}_{ijy}^2 - \boldsymbol{d}_{ijn}^2}{2}\right) = \gamma_j \boldsymbol{M}_j' \boldsymbol{N}_j - \gamma_j \boldsymbol{X}_i' \boldsymbol{N}_j = \gamma_j \left(\boldsymbol{M}_j' - \boldsymbol{X}_i'\right)\boldsymbol{N}_j \qquad \textbf{(9)}$$

This shows that equation (6) collapses to a scalar expression. If $\alpha=0$ we have quadratic deterministic utility but it is only identified in one dimension because, as shown in equation (9), equation (6) is a scalar.

However, as shown in Poole (2001), if the correct geometry is used the Quadratic deterministic utility function with Normal error (QN) model shown in equation (9) can be identified in more than one dimension. In more than one dimension only the cutting line and the distance between the two outcome points are identified; that is, the directional distance γ_j, the normal vector $\boldsymbol{N_j}$, and the projection of the midpoint $\boldsymbol{M_j}$ on the normal vector, $\boldsymbol{M}_j'\boldsymbol{N}_j = \boldsymbol{c}_j$, where $\boldsymbol{c_j}$ is a scalar. Although it appears that the number of parameters for a roll call is s+2, given the fact that the normal vector has unit length, that is, $\boldsymbol{N_j}'\boldsymbol{N_j}=1$, the normal vector is completely determined by the s-1 angles of the vector from the coordinate axes so that the actual number of parameters is s+1. In a polar coordinate system the s-1 angles produce s coordinates provided that the vector is of fixed length. For example, in two dimensions the normal vector can always be written as $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ where θ is the angle of the vector from the horizontal axis and $0 \leq \theta \leq 2\pi$; in three dimensions the normal vector can always be written as $\begin{bmatrix} \sin\varphi\cos\theta \\ \sin\varphi\sin\theta \\ \cos\varphi \end{bmatrix}$ where θ is the angle in the plane of the first and second dimensions, $0 \leq \theta \leq 2\pi$, and φ is the angle from the third dimension to the plane, $0 \leq \varphi \leq \pi$.

In sum, in the quadratic utility model each roll call is determined by s+1 parameters – γ_j, $\boldsymbol{c_j}$, and the s-1 angles. As a practical problem, however, it is easier simply to estimate the normal vector, $\boldsymbol{N_j}$, directly rather than parameterize the problem in terms of the underlying angles.

In more than one dimension the fact that each roll call is determined by only s+1 parameters means that the outcome coordinates are identified only up to ***parallel tracks through the space***. Note that the absolute value of the directional distance, $|\gamma_j|$, is the width of the parallel tracks. Figure 2 shows this geometry.

Figure 2: Identification of Outcome Coordinates

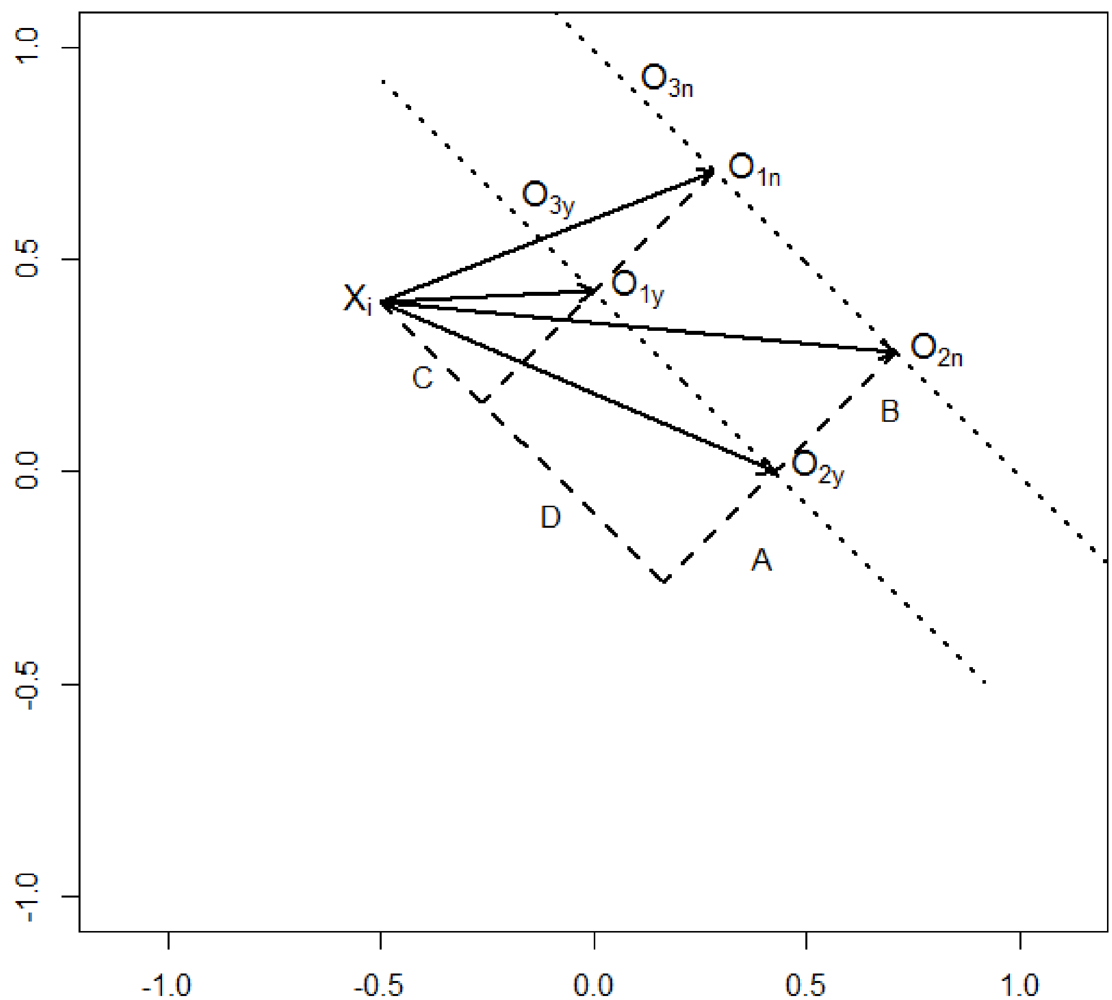


Figure 2 shows one legislator, $\boldsymbol{X_i}$, and her distances to two pairs of outcomes (the vectors in the figure). The parallel dotted lines pass through the two pairs of outcomes. By the Pythagorean Theorem, the squared distance between $\boldsymbol{X_i}$ and $\boldsymbol{O_{1y}}$ is $C^2 + A^2$ (see the indicated line segments in the figure). Similarly, the squared distance between $\boldsymbol{X_i}$ and $\boldsymbol{O_{1n}}$ is $C^2 + (A+B)^2$. Therefore,

$$u_{iy} - u_{in} = -d_{iy}^2 + d_{in}^2 = -(C^2 + A^2) + [C^2 + (A+B)^2] = 2AB + B^2$$

Similarly, the squared distance between $\boldsymbol{X_i}$ and $\boldsymbol{O_{2y}}$ is $(C+D)^2 + A^2$ and the squared distance between $\boldsymbol{X_i}$ and $\boldsymbol{O_{2n}}$ is $(C+D)^2 + (A+B)^2$. Therefore,

$$u_{iy} - u_{in} = -d_{iy}^2 + d_{in}^2 = -[(C+D)^2 + A^2] + [(C+D)^2 + (A+B)^2] = 2AB + B^2$$

In sum, in more than one dimension the outcome coordinates in the quadratic utility model are identified only up to the parallel tracks shown in Figure 2. The disadvantage is that the outcome points can never be definitively estimated. The advantage is the simplicity of the geometry.

In contrast, in the Gaussian deterministic utility function the outcome coordinates are identified, but the price is 2s parameters – the two outcome points $\boldsymbol{O_{jn}}$ and $\boldsymbol{O_{jy}}$. Holding the distance between the two outcomes fixed as in Figure 2, $u_{iy} - u_{in}$ will vary as the outcomes are moved up and down the tracks, and it is a maximum when the outcomes are located at $\boldsymbol{O_{3y}}$ and $\boldsymbol{O_{3n}}$. This identification is due to the non-linearity of the utility difference. For any legislator with the distance between the two outcomes held fixed, the utility difference is maximized when the legislator's ideal point and the outcomes lie on a line parallel to the normal vector for the roll call cutting plane. In the example shown in Figure 2, the utility difference for the O_{3y} and O_{3n} outcome points is:

$$u_{iy} - u_{in} = e^{-A^2} - e^{-(A+B)^2} = e^{-A^2}\left(1 - e^{-(2AB+B^2)}\right)$$

The utility difference for the O_{1y} and O_{1n} outcome points is:

$$u_{iy} - u_{in} = e^{-(A^2+C^2)} - e^{-[C^2+(A+B)^2]} = e^{-(A^2+C^2)}\left(1 - e^{-(2AB+B^2)}\right)$$

Now, because $e^{-A^2} > e^{-(A^2+C^2)}$ the legislator's utility difference is always maximized when her ideal point and the outcomes lie on a line parallel to the normal vector.

Not every legislator can be on the line running through the outcome points! Consequently, ***if we hold the distance between the outcomes fixed***, the outcomes will be positioned on the tracks so that the total of the utility differences of all the legislators vis a vis their chosen outcomes is maximized. In contrast, the total of the utility differences of all the legislators in the quadratic model is the same regardless of the location of the outcome points on the tracks.

This presents some thorny estimation problems. I turn to these practical matters after showing the posterior distribution.

3. The Posterior Distribution

Let **Y** be the n by q matrix of observed "y"/"n" choices and let **Y*** be the n by q matrix of unobserved latent utility differences. From a classical perspective the joint probability distribution of the sample for **Y*** is:

$$f(Y^*|u_{ijy} - u_{ijn}) = \prod_{i=1}^{n}\prod_{j=1}^{q}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(y_{ij}^{*}-(u_{ijy}-u_{ijn}))^2} \qquad (10)$$

It is easy to show that (10) is a probability distribution and integrates to one. Unfortunately, the latent utility differences are not observed and we do not have any simple expression for the joint probability distribution for the sample of discrete choices -- $\boldsymbol{f}(\boldsymbol{Y}\,|\,\boldsymbol{u}_y - \boldsymbol{u}_n)$. However, it is easy to write down the distribution corresponding to any particular choice, that is, $\boldsymbol{f}_{ij}\left(\boldsymbol{y}_{ij}\,|\,\boldsymbol{u}_{ijy} - \boldsymbol{u}_{ijn}\right)$. The product of these nq distributions is proportional to the joint probability distribution of the sample of discrete choices ($\boldsymbol{f}\left(\boldsymbol{Y}\,|\,\boldsymbol{u}_y-\boldsymbol{u}_n\right)$) and the corresponding likelihood function ($\boldsymbol{L}\left(\boldsymbol{u}_y-\boldsymbol{u}_n\,|\,\boldsymbol{Y}\right)$). Specifically, let:

$$y_{ij} = \begin{cases} \text{"y" if } y^*_{ij} > 0 \\ \text{"n" if } y^*_{ij} \le 0 \end{cases} \text{ so that } \begin{cases} P(y^*_{ij} > 0) = \Phi(u_{ijy} - u_{ijn}) \\ P(y^*_{ij} \le 0) = 1 - \Phi(u_{ijy} - u_{ijn}) \end{cases} \tag{11}$$

If the y_{ij} are independent Bernoulli random variables, that is:

$$f_{ij}(y_{ij}|u_{ijy} - u_{ijn}) \sim \text{Bernoulli}(\Phi(u_{ijy} - u_{ijn}))$$

then

$$\boldsymbol{f}(\boldsymbol{Y}\,|\,\boldsymbol{u}_y-\boldsymbol{u}_n) \;\propto \prod_{i=1}^{n}\prod_{j=1}^{q} f_{ij}(y_{ij}\,|\,u_{ijy} - u_{ijn}) = \boldsymbol{L}\left(\boldsymbol{u}_y-\boldsymbol{u}_n\,|\,\boldsymbol{Y}\right) =$$

$$\prod_{\boldsymbol{i=1}}^{\boldsymbol{n}}\prod_{\boldsymbol{j=1}}^{\boldsymbol{q}}[\Phi(\boldsymbol{u_{ijy}} - \boldsymbol{u_{ijn}})]^{y_{ij}}[1-\Phi(\boldsymbol{u_{ijy}} - \boldsymbol{u_{ijn}})]^{(1-y_{ij})} = \prod_{\boldsymbol{i}=1}^{\boldsymbol{n}}\prod_{\boldsymbol{j}=1}^{\boldsymbol{q}}\prod_{\tau=1}^{2}\boldsymbol{P}_{\boldsymbol{ij}\tau}^{\boldsymbol{C}_{\boldsymbol{ij}\tau}} \tag{12}$$

where τ is the index for "y" and "n", $P_{ij\tau}$ is the probability of voting for choice τ, and $C_{ij\tau}$= 1 if the individual's actual choice is τ and zero otherwise. (This representation is convenient for working with the derivatives.)

To implement a Bayesian model requires that prior distributions be stated for all of the parameters. I assume that the prior distribution for a legislator ideal point is an s-dimensional normal distribution with variance-covariance matrix $\zeta^2\boldsymbol{I}_s$ (where $\boldsymbol{I}_s$ is an s by s identity matrix):

$$\xi\left(X_i\right)=\frac{1}{(2\pi\zeta)^{\frac{s}{2}}}e^{-\frac{1}{2\zeta^2}(X_{i1}^2+X_{i2}^2+...+X_{is}^2)} \qquad (13)$$

Similarly, I assume that the prior distributions for the outcome points are also an s-dimensional normal distributions with variance-covariance matrices $\boldsymbol{\kappa}^2\boldsymbol{I}_s$:

$$\xi(O_{jy})=\frac{1}{(2\pi\kappa)^{\frac{s}{2}}}e^{-\frac{1}{2\kappa^2}(O_{j1y}^2+O_{j2y}^2+...+O_{jsy}^2)}$$

and (14)

$$\xi(O_{jn})=\frac{1}{(2\pi\kappa)^{\frac{s}{2}}}e^{-\frac{1}{2\kappa^2}(O_{j1n}^2+O_{j2n}^2+...+O_{jsn}^2)}$$

I assume an uniform prior for β:

$$\xi(\beta)=\frac{1}{c},\ \ 0< c < b \qquad (15)$$

where, empirically, b is no greater than 20; and an uniform 0,1 prior for α.

$$\xi(\alpha)=1,\ \ 0< \alpha < 1 \qquad (16)$$

The posterior distribution is therefore:

$$\xi\left(X,O_a,O_b,\beta,\alpha\,|\,Y\right)\propto\prod_{i=1}^{n}\prod_{j=1}^{q}\left\{f_{ij}(y_{ij}\,|\,u_{ijy}-u_{ijn})\xi(X_i)\xi(O_{ja})\xi(O_{jb})\xi\left(\beta\right)\xi\left(\alpha\right)\right\}=$$

$$\prod_{i=1}^{n}\prod_{j=1}^{q}\left\{[\Phi(u_{ijy}-u_{ijn})]^{y_{ij}}[1-\Phi(u_{ijy}-u_{ijn})]^{(1-y_{ij})}\xi(X_i)\xi(O_{ja})\xi(O_{jb})\xi(\beta)\xi(\alpha)\right\} =$$

$$\prod_{i=1}^{n}\prod_{j=1}^{q}\xi(X_i)\xi(O_{ja})\xi(O_{jb})\xi(\beta)\xi(\alpha)\prod_{\tau=1}^{2}P_{ij\tau}^{C_{ij\tau}} \qquad (17)$$

The natural log of the posterior is:

$$\ell n\xi \propto \sum_{i=1}^{n}\sum_{j=1}^{q}\sum_{\tau=1}^{2}C_{ij\tau}\ln P_{ij\tau}-\frac{1}{2\zeta^2}\left(\sum_{i=1}^{n}\sum_{k=1}^{s}X_{ik}^2\right)-\frac{1}{2\kappa^2}\left(\sum_{j=1}^{q}\sum_{k=1}^{s}O_{jka}^2\right)-\frac{1}{2\kappa^2}\left(\sum_{j=1}^{q}\sum_{k=1}^{s}O_{jkb}^2\right)-\ln(c)=$$

$$\ell n\xi^* = \sum_{i=1}^{n}\sum_{j=1}^{q}\sum_{\tau=1}^{2}C_{ij\tau}\ln P_{ij\tau}-\frac{1}{2\zeta^2}\left(\sum_{i=1}^{n}\sum_{k=1}^{s}X_{ik}^2\right)-\frac{1}{\kappa^2}\left[\sum_{j=1}^{q}\sum_{k=1}^{s}M_{jk}\left(O_{jka}-O_{jkb}\right)\right]-\ln(c) \qquad (18)$$

An alternative way of writing the log of the posterior is to focus on the observed choices; that is:

$$\ell n\xi^* = \sum_{i=1}^{n}\sum_{j=1}^{q}\left\{\ln\left[\Phi\left(u_{ija}-u_{ijb}\right)\right]\right\}-\frac{1}{2\zeta^2}\left(\sum_{i=1}^{n}\sum_{k=1}^{s}X_{ik}^2\right)-\frac{1}{\kappa^2}\left[\sum_{j=1}^{q}\sum_{k=1}^{s}M_{jk}\left(O_{jka}-O_{jkb}\right)\right]-\ln(c)=$$

$$\sum_{i=1}^{n}\sum_{j=1}^{q}\left\{\ln\left[\Phi\left(\beta\psi_{ija}\right)\right]\right\}-\frac{1}{2\zeta^2}\left(\sum_{i=1}^{n}\sum_{k=1}^{s}X_{ik}^2\right)-\frac{1}{\kappa^2}\left[\sum_{j=1}^{q}\gamma_j\sum_{k=1}^{s}M_{jk}N_{jk}\right]-\ln(c) \qquad (15)$$

Where

$$\boldsymbol{\psi}_{ija} = -\left(\frac{\boldsymbol{d}_{ija}^2-\boldsymbol{d}_{ijb}^2}{2}\right)+\boldsymbol{\alpha}\left[\left(\boldsymbol{e}^{\left(-\frac{1}{2}\boldsymbol{d}_{ija}^2\right)}-\boldsymbol{e}^{\left(-\frac{1}{2}\boldsymbol{d}_{ijb}^2\right)}\right)+\left(\frac{\boldsymbol{d}_{ija}^2-\boldsymbol{d}_{ijb}^2}{2}\right)\right]$$

$$\boldsymbol{\psi}_{ija} = -\left(\gamma_j\left(\boldsymbol{M}_j^{'}-\boldsymbol{X}_i^{'}\right)\boldsymbol{N}_j\right)+\boldsymbol{\alpha}\left[\left(\boldsymbol{e}^{\left(-\frac{1}{2}\boldsymbol{d}_{ija}^2\right)}-\boldsymbol{e}^{\left(-\frac{1}{2}\boldsymbol{d}_{ijb}^2\right)}\right)+\left(\gamma_j\left(\boldsymbol{M}_j^{'}-\boldsymbol{X}_i^{'}\right)\boldsymbol{N}_j\right)\right]$$

Intuitively, if outcome "a" is above the cutting plane and "b" is below the cutting plane then the directional distance, γ_j, is positive. Otherwise it is negative. Also if $\boldsymbol{X_i}$ is above the cutting plane then $-\left(\gamma_j\left(\boldsymbol{M}_j^{'}-\boldsymbol{X}_i^{'}\right)\boldsymbol{N}_j\right)>0$ and if ***α=0*** then $\Phi\left(\beta\psi_{ija}\right)>\frac{1}{2}$.

4. Estimation

I use slice sampling (Neal, 2003) to obtain estimates of the parameters in equation (15). Estimation is complicated by the fact that any pair of outcomes along the "track" as shown in Figure 2 will have the same value in the quadratic terms of equation (15). Hence, the normal vector, $\boldsymbol{N_j}$, sets the position through the s-space of the tracks, and the width of the tracks, $\left|\gamma_j\right|$, jointly determine the solution to the quadratic portion of equation (15). I will only deal with the s=2 case:

Step 0: Get starts for the parameters from the double-centered dissimilaries matrix (formed from the Agreement Scores). Set ***=1/2*** and find optimal using the Brent local minimization algorithm (Brent, 2002).

Step 1: draw θ_j *from the region* $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$, $N_{j1}=cos\,\theta_j$ *and* $N_{j2}=sin\,\theta_j$

Step 2: draw γ_j *from the region* $\left[-2.0,2.0\right]$

Step 3: Given $\boldsymbol{N_j}$ and $\boldsymbol{\gamma_j}$ from Steps 1 and 2 find optimal $\boldsymbol{O_{jy}}$ and $\boldsymbol{O_{jn}}$ along the "tracks" that maximize the Gaussian deterministic utility in equation (15).

Step 4: Given $\boldsymbol{N_j}$ and $\boldsymbol{\gamma_j}$ from Steps 1 and 2 and $\boldsymbol{O_{jy}}$ and $\boldsymbol{O_{jn}}$

Draw $\boldsymbol{X_{i1}}$ and $\boldsymbol{X_{i2}}$ from the region [-2.0, 2.0]

Step 5: Go to Step 1 10 (or more) times.

Step 6: draw $\boldsymbol{\beta}$ from the region [1, 20].

Step 7: draw $\boldsymbol{\alpha}$ from the region [0, 1]

Step 8: Go to Step 1 and skip Step 5.

5. Results

The algorithm does work but $\boldsymbol{\alpha}$ is always near 1.0 for matrices that have an $\boldsymbol{\alpha}$ near 1.0 for the one-dimensional alpha-NOMINATE (Carroll et al., 2013; R package **anominate**). Unfortunately I ran out of time due to health reasons to work on this program. I need some Monte Carlo matrices in two dimensions that are clearly the product of quadratic deterministic utility with normal error. This has turned out to be a hard problem so I end here.

References

Bakker, Ryan and Keith T. Poole. 2013a. "Bayesian Metric Multidimensional Scaling." *Political Analysis*

Bakker, Ryan and Keith T. Poole. 2013b. "Appendix to Bayesian Metric Multidimensional Scaling." Available at:

Brent, Richard. 2002. *Algorithms for Minimization Without Derivatives*. New York: Dover).

Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. 2013. "The Structure of Utility in Spatial Models of Voting." *American Journal of Political Science*, 57(4):1008-1028.

Clinton, Joshua D., Simon D. Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data: A Unified Approach." *American Political Science Review*, 98:355-370.

Ladha, Krishna K. 1991. "A Spatial Model of Legislative Voting With Perceptual Error." *Public Choice*, 68:151-174.

Londregan, John B. 2000a. "Estimating Legislators' Preferred Points." *Political Analysis*, 8(1), 35-56.

McFadden, Daniel. 1976. "Quantal Choice Analysis: A Survey." *Annals of Economic and Social Measurement*, 5:363-390.

Martin, Andrew D. and Kevin M. Quinn. 2002. “Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999.” *Political Analysis*, 10:134-153.

Neal, Radford M. 2003. “Slice Sampling.” ***The Annals of Statistics*** 31(3): 705-767.

Poole, Keith T. 2000. "Non-Parametric Unfolding of Binary Choice Data." ***Political Analysis***, 8:211-237.

Poole, Keith T. 2001. "The Geometry of Multidimensional Quadratic Utility in Models of Parliamentary Roll Call Voting." ***Political Analysis***, 9:211-226.

Poole, Keith T. 2005. ***Spatial Models of Parliamentary Voting.*** New York: Cambridge University Press, 2005

Poole, Keith T. and Howard Rosenthal. 1985. “A Spatial Model for Legislative Roll Call Analysis.” *American Journal of Political Science*, 29:357-384.

Poole, Keith T. and Howard Rosenthal. 1987. "Analysis of Congressional Coalition Patterns: A Unidimensional Spatial Model." ***Legislative Studies Quarterly***, 12, 1987, 55-75.

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* New York: Oxford University Press.