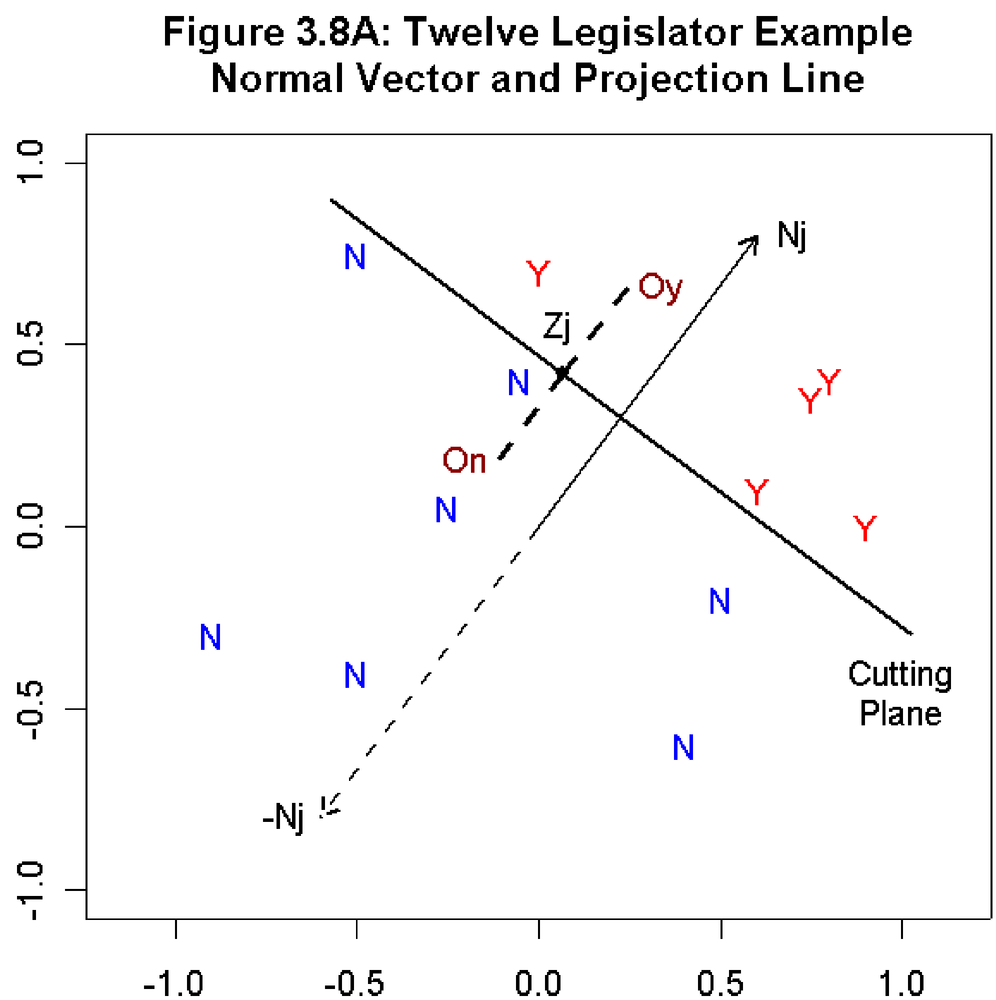
Supplement to Chapter 3 of *Spatial Models of Parliamentary Voting*

By

Keith T. Poole

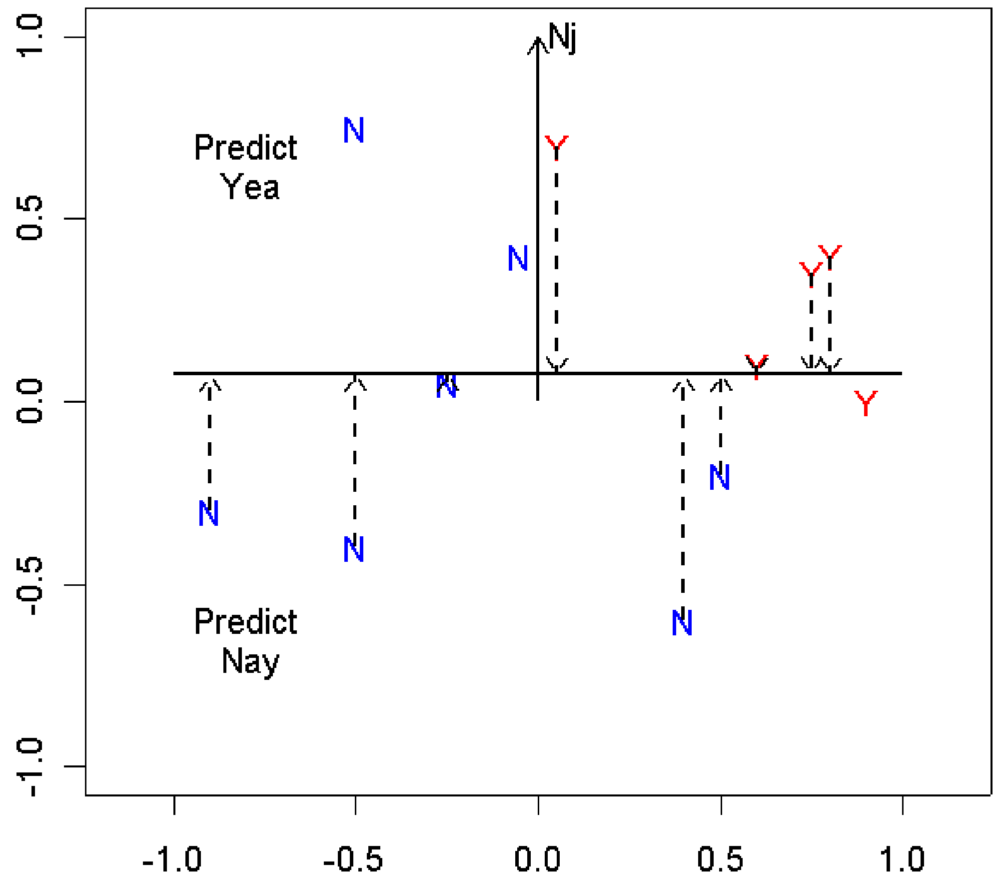
The purpose of this supplement is to show more detail about the *cutting plane procedure* detailed in Chapter 3. Specifically, to more fully characterize the use of the principle of least squares to move the cutting plane through the space of the legislators in order to maximize correct classification.

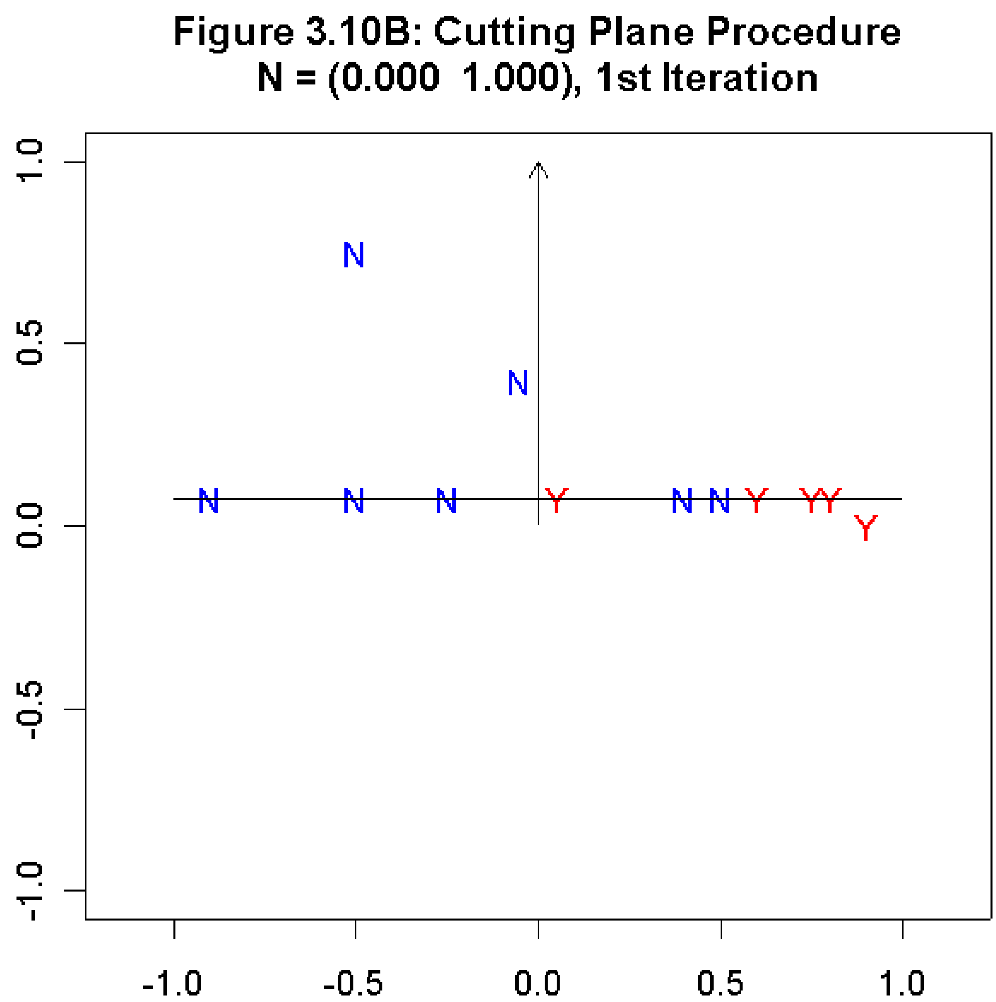
The voting example shown in Figure 3.8A of Chapter 3 is reproduced below. There are 12 legislators with 7 voting Nay and 5 voting Yea and the cutting line perfectly classifies the roll call vote. Given the legislator ideal points and their choices on the roll call, the purpose of the cutting plane procedure is to find a cutting line that maximizes the correct classification on the roll call. In terms of Figure 3.8A, the purpose is to find the cutting plane shown in the figure.



The first step of the cutting plane procedure is to project the legislator points onto the current estimate of the cutting plane as illustrated in Figure 3.10A. This produces the set of points shown in Figure 3.10B. The next step is to find the least squares line through the set of points shown in Figure 3.10B. This has the effect of rotating the cutting plane towards the classification errors.

Figure 3.10A: Cutting Plane Procedure
Projecting Points onto Cutting Line





Let **Ψ** be the p by s matrix of points projected onto a cutting plane as illustrated in Figure 3.10B. Technically, $p > s \geq 2$ and the rank of **Ψ** is s. The least squares problem is to find a plane of rank s-1 through the points in **Ψ** such that the sum of the squared distances from the points in **Ψ** to their orthogonal projections on the plane – the p by s matrix **B** of rank s-1 – is minimized. Following the notation of Chapter 3, the equation of this plane can be written as:

$$\mathbf{N_j'Y} = \alpha \qquad \textbf{(1)}$$

Where $\mathbf{N_j}$ is the s by 1 normal vector to the plane such that $\mathbf{N_j}'\mathbf{N_j} = 1$, $\mathbf{Y}$ is an s by 1 vector of *any point* on the plane, and α is a constant. In Figure 3.10C "Nj (new)" is the normal vector, $\mathbf{N_j}$, for the least squares line. Hence, for any point in $\mathbf{B}$ (row of $\mathbf{B}$), $\mathbf{B_i}$;

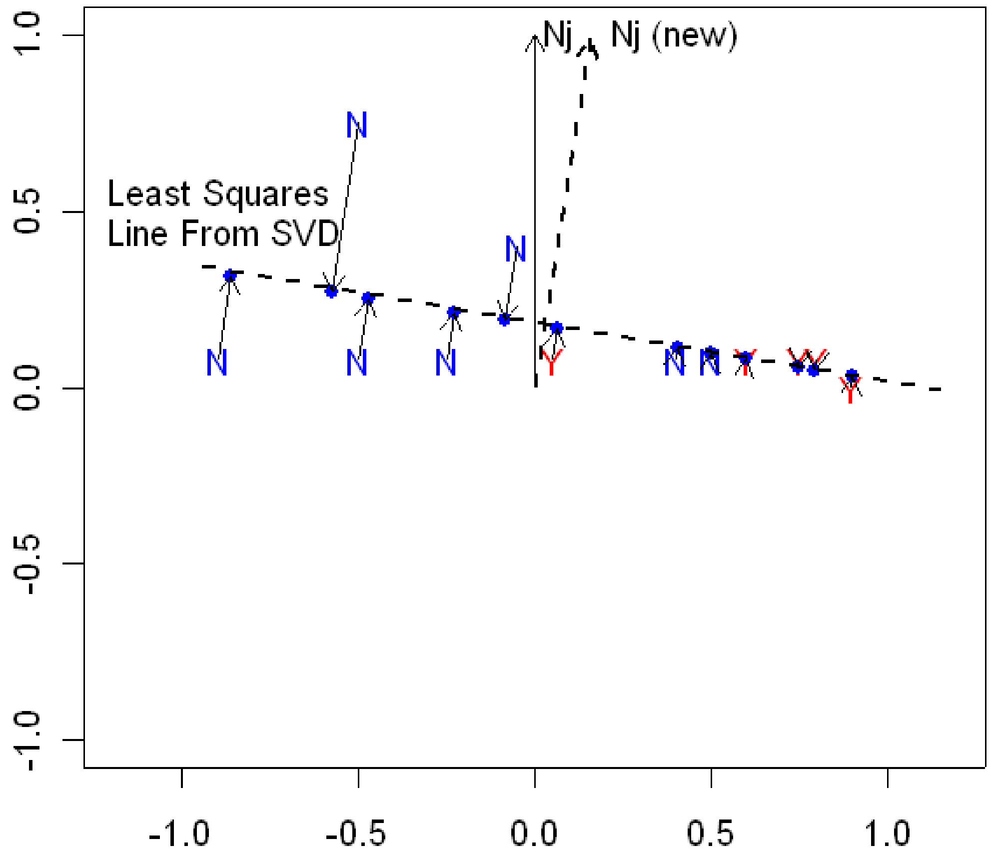
$$\mathbf{N_j}'\mathbf{B_i} = \alpha \qquad \textbf{(2)}$$

The points along the least squares line in Figure 3.10C are the $\mathbf{B_i}$'s. The orthogonal projection of a point in $\mathbf{\Psi}$ (row of $\mathbf{\Psi}$), $\mathbf{\Psi_i}$ (the "N" and "Y" tokens in Figure 3.10C), onto the plane produces the point $\mathbf{B_i}$. The equation for this projection is:

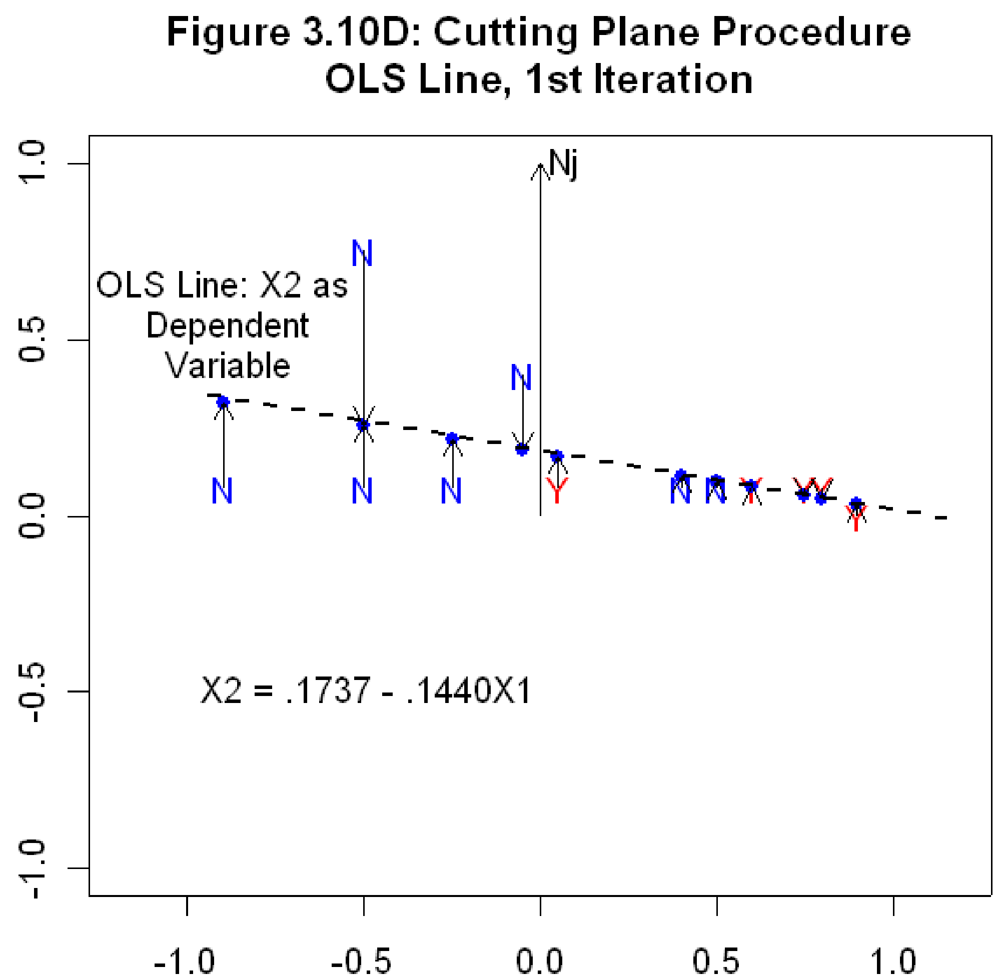
$$\mathbf{B_i} = \mathbf{\Psi_i} + (\alpha - w_i)\mathbf{N_j} \qquad \textbf{(3)}$$

Where $\mathbf{N_j}'\mathbf{\Psi_i} = w_i$.

Figure 3.10C: Cutting Plane Procedure Least Squares Line, 1st Iteration



The $\mathbf{\Psi}_i$ points in Figure 3.10C are projected orthogonally onto the least squares line. In contrast, in a simple linear regression problem one of the dimensions is the dependent variable and the projection to the regression line – quite literally the residual – is *parallel* to the dimension representing the dependent variable. For example, in Figure 3.10D (not shown in Chapter 3) the second dimension – denoted as X2 – is the dependent variable and the first dimension is the independent variable. In a simple OLS the sum of squared error is equal to the sum of squared distances from each observation to the regression line. The projection of the point representing an observation is parallel to the dimension representing the dependent variable. The sum of squared distances from the OLS projection shown in Figure 3.10D is *not the same* as the sum of squared distances in Figure 3.10C. To restate the basic difference, there is no “dependent” variable but it is still a *least squares problem*.



The sum of the squared orthogonal projections in Figure 3.10C is:

$$\sum_{i=1}^{p}\sum_{k=1}^{s}(\psi_{ik}-b_{ik})^2 = \sum_{i=1}^{p}\sum_{k=1}^{s}\left[(\alpha - w_i)N_{jk}\right]^2 =$$

$$\left(\sum_{k=1}^{s} N_{jk}^2\right)\sum_{i=1}^{p}(\alpha - w_i)^2 = \sum_{i=1}^{p}(\alpha - w_i)^2 = \sum_{i=1}^{p}(w_i - \alpha)^2 \qquad (4)$$

Hence for the sum of squared distances to be a minimum it must be the case that $\alpha = \bar{\mathbf{w}}$, where $\bar{\mathbf{w}}$ is the mean of the w_i. However, geometrically, this is equivalent to the least squares line/plane passing through the means of the columns of $\mathbf{\Psi}$; namely

$$\alpha = \bar{w} = \frac{\sum_{i=1}^{p} w_i}{p} = \frac{\sum_{i=1}^{p}\sum_{k=1}^{s}\psi_{ik}N_{jk}}{p} = \sum_{k=1}^{s}\left(N_{jk}\frac{\sum_{i=1}^{p}\psi_{ik}}{p}\right) = \mathbf{N_j'\mu} \qquad \mathbf{(5)}$$

where $\mathbf{\mu}$ is the s by 1 vector of the column means of $\mathbf{\Psi}$.

Substituting equation (5) into equation (4) and factoring out the normal vector yields:

$$\sum_{i=1}^{p}\sum_{k=1}^{s} N_{jk}^2\left(\psi_{ik} - \mu_k\right)^2 \qquad (6)$$

This expression can be rewritten as the matrix equation:

$$\mathbf{N_j'}\left[\mathbf{\Psi}\text{-}\mathbf{J_p\mu'}\right]'\left[\mathbf{\Psi}\text{-}\mathbf{J_p\mu'}\right]\mathbf{N_j} = \mathbf{N_j'\Psi^{*\prime}\Psi^{*}N_j} \qquad (7)$$

where $\mathbf{\Psi^* = \Psi - J_p\mu'}$, $\mathbf{\Psi}$ and $\mathbf{\Psi^*}$ are p by s matrices, $\mathbf{J_p}$ is a p by 1 vector of ones, and $\mathbf{\mu}$ is the s by 1 vector of column means of $\mathbf{\Psi}$.

Hence, by equation (7) the least squares problem is to find an s by 1 vector $\mathbf{N_j}$ that minimizes (7) subject to the constraint that $\mathbf{N_j'N_j} = 1$. The solution for $\mathbf{N_j}$ is a straightforward application of the Eckart-Young (1936) theorem. To see this, let the singular value decomposition of $\mathbf{\Psi^*}$ be $\mathbf{U\Lambda V'}$ where $\mathbf{U}$ is a p by s orthogonal matrix, $\mathbf{\Lambda}$ is an s by s diagonal matrix with the singular values in decreasing order on the diagonal, $\mathbf{V}$ is an s by s orthogonal matrix, $\mathbf{U'U} = \mathbf{V'V} = \mathbf{VV'} = \mathbf{I_s}$, and $\mathbf{I_s}$ is an s by s identity matrix. Substituting into equation (7):

$$\mathbf{N}_j'\mathbf{\Psi}^{*\prime}\mathbf{\Psi}^*\mathbf{N}_j = \mathbf{N}_j'\mathbf{V\Lambda U'U\Lambda V'}\mathbf{N}_j = \mathbf{N}_j'\mathbf{V\Lambda}^2\mathbf{V'}\mathbf{N}_j \qquad (8)$$

Note that setting $\mathbf{N_j}$ equal to the sth column of **V**, $\mathbf{V_s}$ produces:

$$\mathbf{V}_s'\mathbf{V\Lambda}^2\mathbf{V'}\mathbf{V}_s = \lambda_s^2 \qquad (9)$$

that is, setting $\mathbf{N_j} = \mathbf{V_s}$ produces a sum of squared distances from the orthogonal projections in Figure 3.10C equal to the square of the sth singular value of **Ψ*.** This is the solution from the Eckart-Young theorem. To see why this is the solution consider the relationship between the points in **Ψ** and their projections on the least squares line – the matrix **B** shown in equation (3). Subtracting the column means from **Ψ** to produce **Ψ*** and subtracting these same means from the corresponding columns of **B** produces a graph identical to Figure 3.10C except that the center of the coordinate axes is moved. This cannot affect the sum of the squared orthogonal projections. To see this, let $\mathbf{B^* = B - J_p\mu'}$ and let $w_i^* = \mathbf{N_j'\Psi_i^*}$. Because α=0, equation (3) becomes:

$$\mathbf{B_i^*} = \mathbf{\Psi_i^*} - w_i^*\mathbf{N_j} \qquad (10)$$

Equation (4) becomes:

$$\sum_{i=1}^{p}\sum_{k=1}^{s}\left(\psi_{ik}^* - b_{ik}^*\right)^2 = \sum_{i=1}^{p}\left(w_i^*\right)^2 = \mathbf{N}_j'\mathbf{\Psi}^{*\prime}\mathbf{\Psi}^*\mathbf{N}_j \qquad (11)$$

which is equivalent to equation (8).

Equation (11) is an instance of the problem considered by Eckart and Young (1936). Namely, given a p by s matrix **Ψ*** of rank s, find a p by s matrix **B*** of rank s-1 such that equation (11) is minimized. The solution is to compute the singular value decomposition of **Ψ***, **UΛV′,** and then set the sth (smallest) singular value on the diagonal of **Λ,** $\boldsymbol{\lambda_s}$ **,** equal to zero and then remultiplying. That is, let $\Lambda^\#$ be the same as Λ expect for the substitution of $\boldsymbol{\lambda_s}$ with 0, then the solution is:

$$\mathbf{B^* = U\Lambda^{\#}V'} \qquad (12)$$

Therefore, the least squares line is:

$$\mathbf{B = B^* + J_p \mu'} \qquad (13)$$

and the normal vector is:

$$\mathbf{N_j} = \mathbf{V_s} \qquad (14)$$

Where $\mathbf{V_s}$ is the sth column of **V.**