Appendix to

BAYESIAN METRIC MULTIDIMENSIONAL SCALING

This Appendix shows a set of proofs of the existence of full rank Hessians for metric MDS in section A1; WINBUGS code for the Bayesian dissimilarities scaling of the 90^{th} Senate disagreement scores in A2; and the first and second derivatives for the similarities and unfolding models corresponding to equations (11) and (16) respectively, in A3.

In section A1 we show how to identify solutions for metric MDS problems. By identification what we mean is estimating the smallest number of parameters such that the Hessian matrix corresponding to a solution is full rank. If too many parameters are estimated the Hessian is singular. If too few are estimated then the log-posterior is distorted and a sub-optimal solution will be the result.

Assume that our dissimilarities data are squared distances between pairs of stimuli (the analysis of unfolding data is essentially the same). Our q by q symmetric matrix of data has q(q-1)/2 unique entries (we ignore the diagonal of zeroes). Suppose there is an exact solution; that is, a set of q points in s dimensions that exactly reproduces the squared distances. Clearly, given that we only observe the distances, it does not matter what origin and the rotation around that origin we select as long as the configuration of points vis a vis one another is not altered.

With q points in s dimensions we have to solve for q*s parameters. However, we can set any point to the origin – (0,0,…,0) – so this leaves us with q*s – s= (q-1)*s parameters. To pin down the configuration we need to set the rotation. In general a rotation matrix is determined by s-1 angles from the origin and sign flips on each dimension. For example, in two dimensions the general form of the rotation matrix is:

$$\Gamma = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \; 0 \le \theta \le 2\pi$$

However, note that given a *specific* θ we have *four* rotation matrices:

$$\Gamma_1 = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \Gamma_2 = \begin{bmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \Gamma_3 = \begin{bmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{bmatrix} \quad \Gamma_4 = \begin{bmatrix} -\cos\theta & -\sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

Or

$$\tilde{\Gamma} = \Delta\Gamma \text{ where } \Delta = \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$$

That is, given a specific θ, there are 2^s sign flips corresponding to the s columns of the rotation matrix. With *s*=2, suppose that we have a solution $\tilde{Z}$ such that it reproduces our matrix of squared distances, **D**. Then there are three more solutions corresponding to the above rotation matrices that also exactly reproduce **D**. In general, in *s* dimensions, with one point set to the origin with s-1 fixed angles in the rotation matrix, if we have a solution $\tilde{Z}$ that exactly reproduces the matrix of squared distances then there an additional 2^s-1 solutions that exactly reproduce **D**.

This identification problem is very similar to that discussed by Rivers (2003). He discusses the identification of the classical maximum likelihood factor analysis problem and shows the number of restrictions necessary to get identification (these include fixing the origin and sign flips). However, his main concern is the identification of the multidimensional IRT model where the data are indicators and he shows that fixing s+1 points (or s(s+1) parameters) fully identifies the model. Our result is different because we assume that we observe (noisy) ratio scale data. Identification is somewhat different in this setting.

With these preliminaries we turn to our existence proofs.

A1: Existence Proofs for the Hessian

We have a total of (s*q)+1 parameters – the q points plus σ^2 (κ^2 is a fixed constant) for the similarities problem. For the unfolding problem we have (s*(n+q))+1 parameters. Because only distances, the $\boldsymbol{d}_{jm}^*$ and the d_{jm}, are used in the log-posterior, we impose the constraints that $\ln(d_{jm}^*) > 0, \forall j \neq m$ (or $\ln(d_{ij}^*) > 0, \forall i, j$) and $\ln(d_{jm}) > 0, \forall j \neq m$ (or $\ln(d_{ij}) > 0, \forall i, j$) for our proofs below. If any $\boldsymbol{d}_{jm}^*$ or d_{jm} is equal to zero for j≠m then equation (11) is equal to -∞. As a practical matter, this is not a problem for the observed data, $\boldsymbol{d}_{jm}^*$, because it can be rescaled or the corresponding j^{th} column and j^{th} row can be dropped on the assumption that the underlying j^{th} and m^{th} stimuli are the same.

For the unfolding problem if a $\boldsymbol{d}_{ij}^*$ is zero then equation (16) is equal to -∞. Again, as a practical matter the offending $\boldsymbol{d}_{ij}^*$ can be rescaled (e.g., set to a small distance greater than zero) or treated as missing data.

In our proofs below we analyze only the similarities problem because the unfolding problem is a subset of the similarities problem albeit with missing data. That is, we could set $\boldsymbol{W} = \left[\frac{\boldsymbol{Z}}{\boldsymbol{X}}\right]$ where **W** is a (q+n) by s matrix and all the proofs would hold using **W** instead of **Z**.

In our proofs we assume that all the points are *distinct*; that is,

Definition: A set of points is distinct if $\boldsymbol{d_{jm}} > 0\ \forall \boldsymbol{j} \neq \boldsymbol{m}$, or equivalently, $\forall$ j,m=1,…,q, and j≠m, $Z_j \neq Z_m$.

In practice distinctness is not a serious problem because if two points were the same, that is, $\mathbf{Z_j} = \mathbf{Z_m}$, then there is a "pinhole" that goes down to -∞ in the surface of equation (11). Such a "pinhole" cannot be a maximum in any event. We simply avoid the problem by always "moving around" them.

Let $\ell n \xi^*$ denote the right hand side of equation (11). For any configuration of points in s dimensions, $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_{q-1}, \boldsymbol{Z}_q$, there is a unique $\hat{\sigma}^2$ which is simply the mean of the q(q-1)/2 squared differences between $\ln(d_{jm}^*)$ and $\ln(d_{jm})$ (see equation (A3) and (A12) in Appendix A3). Hence we will ignore it in the notation below and simply assume that it is computed from the configuration; that is, $\hat{\sigma}^2$ (**Z**) or $\hat{\sigma}^2$ **(Z,X)**.

Given a configuration of points in s dimensions, there are an infinite number of configurations that produce the same $\ell n\xi^*$ by adding a constant and rotating the original configuration. Let **Ω** be the set

$$\Omega = \left\{ \boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_{q-1}, \boldsymbol{Z}_q, \boldsymbol{\alpha}, \Gamma \right\} \qquad \text{(A1)}$$

where **α** is an s-length vector of additive constants and **Γ** is an s by s rotation matrix. Let $\ell n\xi^*(\Omega)$ be the function value for the set **Ω**. This allows us to state a simple non-existence theorem for the Hessian.

Theorem 1: Given Ω such that all Z are *distinct*, then the Hessian for any Ω that maximizes $\ell n\xi^*$ will be singular.

Proof: Given that there are an infinite number of configurations of points, there are an infinite number of **Ω**. However, since every possible configuration is a member of some **Ω** we can compute all possible values of $\ell n\xi^*(\Omega)$. Hence, it must be the case that for some **Ω***, $\ell n\xi^*(\Omega^*) \geq \ell n\xi^*(\Omega) \forall \Omega \neq \Omega^*$. However, no member of **Ω*** can be an inflection point because there are an infinite number of configurations in **Ω*** within any arbitrary distance from any selected configuration. Therefore the Hessian is singular for all members of **Ω***. Q.E.D.

Note that because $\ell n\xi^*\left(\Omega^*\right)$ is the value for every element of **Ω*** then this results in a uniform distribution over a subspace of the real q*s hyperplane of the parameters (much like a "mesa" but infinitely long). The same is true for other **Ω ≠ Ω*.** Geometrically, there are an infinite number of stacked uniform "mesas" over the q*s hyperplane of the parameters with **Ω*** having the highest "altitude" $\ell n\xi^*\left(\Omega^*\right)$. If one point is set to the origin then we still have an infinite number of stacked uniform "mesas" but now the radius of each "mesa" is finite with a value of $\sqrt{\sum_{j=1}^{q}\sum_{k=1}^{s}\boldsymbol{Z}_{jk}^2}$ where one of the $\mathbf{z}_j = \mathbf{0}$.

We now show that with q distinct points and s(s+1)/2 hard constraints the Hessian is full rank. Without loss of generality, we can pick **α** and **Γ** so that the q by s coordinate matrix, $\tilde{\boldsymbol{Z}}$, has the following form:

$$\tilde{\boldsymbol{Z}} = \begin{bmatrix} \boldsymbol{Z}_{11} & \boldsymbol{Z}_{12} & \boldsymbol{Z}_{13} & \cdots & \boldsymbol{Z}_{1,\boldsymbol{q}-1} & \boldsymbol{Z}_{1\boldsymbol{q}} \\ \boldsymbol{Z}_{21} & \boldsymbol{Z}_{22} & \boldsymbol{Z}_{23} & \cdots & \boldsymbol{Z}_{2,\boldsymbol{q}-1} & \boldsymbol{Z}_{2\boldsymbol{q}} \\ & & \vdots & & & \\ \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s}+1,1} & \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s}+1,2} & \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s}+1,3} & \cdots & \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s}+1,\boldsymbol{q}-1} & 0 \\ \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s},1} & \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s},2} & \boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s},3} & \cdots & 0 & 0 \\ & & \vdots & & & \\ \boldsymbol{Z}_{\boldsymbol{q}-2,1} & \boldsymbol{Z}_{\boldsymbol{q}-2,2} & 0 & \cdots & 0 & 0 \\ \boldsymbol{Z}_{\boldsymbol{q}-1,1} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad \text{(A2)}$$

That is, we set $\mathbf{Z_q}$ to the origin and then pick s-1 angles for $\mathbf{\Gamma}$ such that all but one of the coordinates for $\mathbf{Z_{q-1}}$ are equal to zero, all but two of the coordinates for $\mathbf{Z_{q-2}}$ are equal to zero, and so on. As we explained above, we have the sign flips, $\tilde{\Gamma} = \Delta\Gamma$, where $\mathbf{\Delta}$ is an s by s diagonal matrix with plus or minus ones on the diagonal. Given s-1 specific angles, θ_1 , θ_2 ,..., θ_{s-1} , then there are 2^s sign flips corresponding to the s columns of the rotation matrix. This allows us to define

$$\Omega = \left\{\tilde{Z}, \Delta\right\} \tag{A3}$$

Theorem 2: Given Ω as in (A3) such that all $\tilde{Z}$ are *distinct*, then the Hessian for any Ω that maximizes $\ell n\xi^*$ will be rank q*s-(s*(s+1)/2).

Proof: Every configuration of points, **Z**, can be transformed into $\tilde{Z}$ as in equation (A2) by choice of origin and rotation without changing the inter-point distances. Given that there are an infinite number of configurations of points, **Z**, and each one can be transformed into a $\tilde{Z}$ that satisfy equation (A2), there are an infinite number of **Ω** that satisfy equation (A3). However, since every possible $\tilde{Z}$ configuration is a member of some **Ω**, then we can compute all possible values of $\ell n\xi^*(\Omega)$.

Hence, it must be the case that for some **Ω***,

$\ell n\xi^*\left(\Omega^*\right) \geq \ell n\xi^*\left(\Omega\right) \forall \Omega \neq \Omega^*$. By construction, given s-1 specific angles, θ_1 , θ_2 ,..., θ_{s-1} , then there are 2^s sign flips corresponding to the s columns **Δ**. Hence, **Ω*** has 2^s members and each is separated from the others by a distance of at least $\boldsymbol{\delta} = \min\left\{2\sqrt{\sum_{j=1}^{q-1}\tilde{\boldsymbol{Z}}_{j1}^2}, 2\sqrt{\sum_{j=1}^{q-1}\tilde{\boldsymbol{Z}}_{j2}^2}, ..., 2\sqrt{\sum_{j=1}^{q-1}\tilde{\boldsymbol{Z}}_{js}^2}\right\} > 0$. Denote these configurations as $\tilde{\boldsymbol{Z}}^*$ so that $\Omega^* = \left\{\tilde{\boldsymbol{Z}}^*, \Delta\right\}$. Consider any nearby configuration of distinct points, $\tilde{\boldsymbol{Z}}$, $\Omega = \left\{\tilde{\boldsymbol{Z}}, \Delta\right\}$, within an infinitesimal distance of **Ω***; that is, $\gamma = \sqrt{\sum_{j=1}^{q}\sum_{k=1}^{s}\left(\tilde{\boldsymbol{Z}}_{jk}^* - \tilde{\boldsymbol{Z}}_{jk}\right)^2} > 0$. Hence, by construction $\ell n\xi^*\left(\Omega^*\right) > \ell n\xi^*\left(\Omega\right)$. Because this is true for an infinitesimal distance on the q*s-(s*(s+1)/2) dimensional hyperplane in any direction from $\tilde{\boldsymbol{Z}}^*$, the 2^s members of **Ω*** are inflection points with corresponding full rank Hessians. Q.E.D.

Note that the key difference between Theorems 1 and 2 is that **Ω*** in Theorem 1 had an infinite number of members and in Theorem 2 **Ω*** had 2^s members. In Theorem 1 this meant that no member of **Ω*** could be an inflection point because there are an infinite number of members within an infinitesimal distance of any selected member (the "mesa"). In contrast, the 2^s members of

Ω* from Theorem 2 are separated from each other by non-zero distances. Hence, it is easy to show using a standard argument from mathematical analysis that any configuration not in **Ω*** that is an infinitesimal distance from one of the 2^s members of **Ω*** must be, by construction, less than the maximum; that is, $\ell n\xi^*\left(\Omega^*\right) > \ell n\xi^*\left(\Omega\right)$.

We now show two corollaries: first, if the number of hard constraints is less than s(s+1)/2, then the Hessian is singular; and second, if the number of hard constraints is greater than s(s+1)/2 then the solution is inferior in the sense that $\ell n\xi^*\left(\Omega^*\right) > \ell n\xi^*\left(\Omega\right)$.

Corollary 1: Let the number of hard constraints be less than s(s+1)/2. Given Ω such that all Z are *distinct*, then the Hessian for any Ω that maximizes $\ell n\xi^*$ will be singular.

Proof: Suppose that the number of hard constraints is (s(s+1)/2)-1. Without loss of generality modify $\tilde{\boldsymbol{Z}}$ so that $-\infty < Z_{q-s+1,q} < +\infty$, that is, coordinate $Z_{q-s+1,q}$ is not constrained to be zero. Denote this modified configuration as $\tilde{\boldsymbol{Z}}_{(-1)}$. There are an infinite number of **Ω** that satisfy $\Omega = \left\{\tilde{\boldsymbol{Z}}_{(-1)}, \Delta\right\}$. However, since every possible $\tilde{\boldsymbol{Z}}_{(-1)}$ configuration is a member of some **Ω,** then we can compute all possible values of $\ell n\xi^*(\Omega)$. Hence, it must be the case that for some **Ω***, $\ell n\xi^*(\Omega^*) \geq \ell n\xi^*(\Omega) \forall \Omega \neq \Omega^*$. **Ω*** has an infinite number of members because, by construction, $-\infty < Z_{q-s+1,q} < +\infty$. However, no member of **Ω*** can be an inflection point because there are an infinite number of configurations in **Ω*** within any arbitrary infinitesimal distance from any selected configuration in the direction of $Z_{q-s+1,q}$. Therefore the Hessian is singular for all members of **Ω***. Finally, it is easy to construct similar arguments for $\tilde{\boldsymbol{Z}}_{(-2)}$, $\tilde{\boldsymbol{Z}}_{(-3)}$, etc. Q.E.D.

Corollary 2: Let the number of hard constraints be greater than s(s+1)/2. Given Ω such that all Z are *distinct*, then the Ω that maximizes $\ell n\xi^*$ will be less than $\ell n\xi^*$ for a Z with (s+1)/2 hard constraints as in equation (A2).

Proof: Suppose that the number of hard constraints is (s(s+1)/2)+1. Without loss of generality modify $\tilde{\boldsymbol{Z}}$ so that an additional coordinate is constrained to be a constant; for example, let $\boldsymbol{Z}_{q-s+1,q-1} = \boldsymbol{C}_{q-s+1,q-1}$. Denote this modified configuration as $\tilde{\boldsymbol{Z}}_{(+1)}$. However this modified configuration is a member of some **Ω** used in Theorem 2. From Theorem 2 we have $\ell n\xi^*(\Omega^*) \geq \ell n\xi^*(\Omega) \forall \Omega \neq \Omega^*$. Hence, unless $\boldsymbol{C}_{q-s+1,q-1}$ is exactly equal to

$\boldsymbol{Z}_{\boldsymbol{q}-\boldsymbol{s}+1,\boldsymbol{q}-1}$ in **Ω*** $\ell n\xi^{*}\left(\Omega^{*}\right) > \ell n\xi^{*}\left(\Omega\right)$. Finally, it is easy to construct similar arguments for any subset of additionally constrained coordinates. Q.E.D.

In one dimension, setting one point to the origin results in two solutions with the same $\ell n\xi^{*}$ values. For purposes of characterizing the distributions of the parameters with MCMC methods, setting the sign -- a soft constraint -- on a second point (typically a point that is distant from the origin) isolates one log-posterior. So we get a unique log-posterior with one hard and one soft constraint. However, note that, if we use two hard constraints by fixing two points we get an inferior result because we have fixed one of the distances.

A2: WINBUGS SIMILARITIES MODEL

```
#
#  MDS Model for 90th Senate--over constrained
#
model{

#  Fix one point
#
        x[8,1] <- -0.626000480
………………  x[8,2] <-  0.46524749
#
# llh and sumllh monitor the log-likelihood
#
for (i in 1:101){
    llh[i,i] <- 0.0
    for (j in i+1:102){
#
#  Read in Distances rather than the similarities (makes handling missing data easier)
#
         dstar[i,j] ~ dlnorm(mu[i,j],tau)
         mu[i,j] <- log(sqrt((x[i,1]-x[j,1])*(x[i,1]-x[j,1])+(x[i,2]-x[j,2])*(x[i,2]-x[j,2])))
         llh[i,j] <- (log(dstar[i,j])-mu[i,j])*(log(dstar[i,j])-mu[i,j])
         llh[j,i] <- (log(dstar[i,j])-mu[i,j])*(log(dstar[i,j])-mu[i,j])
    }
}

   llh[102,102] <- 0.0
   sumllh <- sum(llh[,])
#
  ## priors
  tau ~ dgamma(1,1)

#
# Informed priors placed below (not all shown)
#
  x[1,1] ~ dnorm(0,.1) I(0,)
  x[1,2] ~ dnorm(0,.1) I(,0)
  x[2,1] ~ dnorm(0,.1) I(,0)
  x[2,2] ~ dnorm(0,.1) I(0,)

...etc. etc.

  x[98,1] ~ dnorm(0,.1) I(,0)
  x[98,2] ~ dnorm(0,.1) I(0,)
  x[99,1] ~ dnorm(0,.1) I(,0)
  x[99,2] ~ dnorm(0,.1) I(, -0.5)
  x[100,1] ~ dnorm(0,.1) I(,0)
  x[100,2] ~ dnorm(0,.1) I(,-0.5)
  x[101,1] ~ dnorm(0,.1) I(0.5,)
  x[101,2] ~ dnorm(0,.1) I(0.2,)
  x[102,1] ~ dnorm(0,.1) I(,-0.2)
  x[102,2] ~ dnorm(0,.1) I(,0)

}
```

A3: The Derivatives for the Log-Normal Bayesian Model

Similarities: The first derivatives for the similarities problem are:

$$\frac{\partial \ell n\xi}{\partial Z_{jk}} = -2\frac{1}{2\sigma^2}\sum_{j\neq m}^{q}\left\{\left(\ln\left(d^*_{jm}\right)-\ln\left(d_{jm}\right)\right)\left(-\frac{1}{d_{jm}}\right)\left(\frac{1}{2}\right)\left[\sum_{k=1}^{s}(Z_{jk} - Z_{mk})^2\right]^{-\frac{1}{2}}\left(2\left[Z_{jk} - Z_{mk}\right]\right)\right\} - \frac{Z_{jk}}{\kappa^2}$$

which simplifies to

$$\frac{\partial \ell n\xi}{\partial Z_{jk}} = \frac{1}{\sigma^2}\sum_{j\neq m}^{q}\left\{\frac{\left(\ln\left(d^*_{jm}\right)-\ln\left(d_{jm}\right)\right)}{d^2_{jm}}\left(Z_{jk} - Z_{mk}\right)\right\} - \frac{Z_{jk}}{\kappa^2} \qquad \textbf{(A1)}$$

and

$$\frac{\partial \ell n\xi}{\partial \sigma^2} = -\frac{q(q-1)}{4\sigma^2} + \frac{1}{2\sigma^4}\sum_{j=1}^{q-1}\sum_{m=j+1}^{q}\left(\ln\left(d^*_{jm}\right)-\ln\left(d_{jm}\right)\right)^2 \qquad \textbf{(A2)}$$

Hence, we get the usual result for the variance term:

$$\hat{\sigma}^2 = \frac{2}{q(q-1)}\sum_{j=1}^{q-1}\sum_{m=j+1}^{q}\left(ln\left(d^*_{jm}\right)-ln\left(d_{jm}\right)\right)^2 \qquad \textbf{(A3)}$$

Note that if κ^2 is a vague prior, the practical effect is that at an inflection point we have $\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\sigma^2} \approx \frac{\partial \ell n\xi}{\partial Z_{jk}} = 0$. Numerically, this is a handy result because it makes computing the inverse Hessian much easier to accomplish.

The second derivative for the variance is:

$$\frac{\partial^2 \ell n\xi}{\partial\sigma^2\partial\sigma^2} = \frac{q(q-1)}{4\sigma^4} - \frac{1}{\sigma^6}\sum_{j=1}^{q-1}\sum_{m=j+1}^{q}\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right)^2 \qquad \textbf{(A4)}$$

Substituting (A3) into (A4) it is easy to show that $\frac{\partial^2 \ell n\xi}{\partial\sigma^2\partial\sigma^2} < 0$ so that when the global maximum for the Z_{jk} is found σ^2 will be a maximum as well.

The second derivatives for the coordinates are:

$$\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{jk}} = -4\sum_{j\neq m}^{q}\frac{\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right)}{d^4_{jm}}\left(Z_{jk} - Z_{mk}\right)^2 - 2\sum_{j\neq m}^{q}\left[\frac{\left(Z_{jk} - Z_{mk}\right)^2}{d^4_{jm}}\right] + 2\sum_{j\neq m}^{q}\left[\frac{\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right)}{d^2_{jm}}\right] - \frac{1}{\kappa^2}$$
(A5)

$$\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{mk}} = 4\frac{\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right)}{d^4_{jm}}\left(Z_{jk} - Z_{mk}\right)^2 + 2\frac{\left(Z_{jk} - Z_{mk}\right)^2}{d^4_{jm}} - 2\frac{\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right)}{d^2_{jm}} \qquad \textbf{(A6)}$$

In more than one dimension

$$\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{j\ell}} = -2\sum_{j\neq m}^{q}\left\{\left[\frac{\left(Z_{jk} - Z_{mk}\right)\left(Z_{j\ell} - Z_{m\ell}\right)}{d^4_{jm}}\right]\left[2\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right) - 1\right]\right\} \qquad \textbf{(A7)}$$

$$\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{m\ell}} = 2\frac{\left(Z_{jk} - Z_{mk}\right)\left(Z_{j\ell} - Z_{m\ell}\right)}{d^4_{jm}}\left[2\left(\ln\left(d^*_{jm}\right) - \ln\left(d_{jm}\right)\right) + 1\right] \qquad \textbf{(A8)}$$

where $\ell = 1,\ldots,s$ and $\ell \neq k$.

Unfolding: The first derivatives for the unfolding problem are:

$$\frac{\partial \ell n\xi}{\partial X_{ik}} = \frac{1}{\sigma^2}\sum_{j=1}^{q}\left\{\frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^2}\left(X_{ik}\ \text{-}\ Z_{jk}\right)\right\}-\frac{X_{ik}}{\zeta^2} \qquad \textbf{(A9)}$$

$$\frac{\partial \ell n\xi}{\partial Z_{jk}} = -\frac{1}{\sigma^2}\sum_{i=1}^{n}\left\{\frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^2}\left(X_{ik}\ \text{-}\ Z_{jk}\right)\right\}-\frac{Z_{jk}}{\kappa^2} \qquad \textbf{(A10)}$$

and

$$\frac{\partial \ell n\xi}{\partial \sigma^2} = -\frac{nq}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}\sum_{j=1}^{q}\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)^2 \qquad \textbf{(A11)}$$

Hence, we get the usual result for the variance term for the unfolding model:

$$\hat{\sigma}^2 = \frac{1}{nq}\sum_{i=1}^{n}\sum_{j=1}^{q}\left(ln\left(d_{ij}^*\right)-ln\left(d_{ij}\right)\right)^2 \qquad \textbf{(A12)}$$

Note that if ζ^2 and κ^2 are vague priors, the practical effect is that at an inflection point we have $\frac{\partial^2 \ell n\xi}{\partial X_{ik}\sigma^2} \approx \frac{\partial \ell n\xi}{\partial X_{ik}} = 0$.

The second derivative for the variance is:

$$\frac{\partial^2 \ell n\xi}{\partial \sigma^2 \partial \sigma^2} = \frac{nq}{2\sigma^4}-\frac{1}{\sigma^6}\sum_{i=1}^{n}\sum_{j=1}^{q}\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)^2 \qquad \textbf{(A13)}$$

Substituting (A12) into (A13) it is easy to show that $\frac{\partial^2 \ell n\xi}{\partial\sigma^2\partial\sigma^2} < 0$ so that when the global maximum for the X_{ik} and Z_{jk} is found σ^2 will be a maximum as well.

The second derivatives for the coordinates are:

$$\frac{\partial^2 \ell n\xi}{\partial X_{ik}\partial X_{ik}} = -2\sum_{j=1}^{q}\frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^4}\left(X_{ik}-Z_{jk}\right)^2 - \sum_{j=1}^{q}\left[\frac{\left(X_{ik}-Z_{jk}\right)^2}{d_{ij}^4}\right] + \sum_{j=1}^{q}\left[\frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^2}\right] - \frac{1}{\zeta^2} \quad \textbf{(A14)}$$

$$\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{jk}} = 2\sum_{i=1}^{n}\frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^4}\left(X_{ik}-Z_{jk}\right)^2 + \sum_{i=1}^{n}\left[\frac{\left(X_{ik}-Z_{jk}\right)^2}{d_{ij}^4}\right] - \sum_{i=1}^{n}\left[\frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^2}\right] - \frac{1}{\kappa^2} \quad \textbf{(A15)}$$

$$\frac{\partial^2 \ell n\xi}{\partial X_{ik}\partial Z_{jk}} = \left[\frac{\left(X_{ik}-Z_{jk}\right)^2}{d_{ij}^4}\right]\left[2\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)+1\right] - \frac{\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)}{d_{ij}^2} \quad \textbf{(A16)}$$

$$\frac{\partial^2 \ell n\xi}{\partial X_{ik}\partial X_{hk}} = \frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{mk}} = 0 \quad \textbf{(A17)}$$

Where h=1,...,n and h≠i. In more than one dimension

$$\frac{\partial^2 \ell n\xi}{\partial X_{ik}\partial X_{i\ell}} = -\sum_{j=1}^{q}\left\{\left[\frac{\left(X_{ik}-Z_{jk}\right)\left(X_{i\ell}-Z_{j\ell}\right)}{d_{ij}^4}\right]\left[2\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)+1\right]\right\} \quad \textbf{(A18)}$$

$$\frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{j\ell}} = -\sum_{i=1}^{n}\left\{\left[\frac{\left(X_{ik}-Z_{jk}\right)\left(X_{i\ell}-Z_{j\ell}\right)}{d_{ij}^4}\right]\left[2\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)+1\right]\right\} \quad \textbf{(A19)}$$

$$\frac{\partial^2 \ell n\xi}{\partial X_{ik}\partial Z_{j\ell}} = \left[\frac{\left(X_{ik}-Z_{jk}\right)\left(X_{i\ell}-Z_{j\ell}\right)}{d_{ij}^4}\right]\left[2\left(\ln\left(d_{ij}^*\right)-\ln\left(d_{ij}\right)\right)+1\right] \quad \textbf{(A20)}$$

$$\frac{\partial^2 \ell n\xi}{\partial X_{ik}\partial X_{h\ell}} = \frac{\partial^2 \ell n\xi}{\partial Z_{jk}\partial Z_{m\ell}} = 0 \qquad \textbf{(A21)}$$

APPENDIX REFERENCES

Rivers, Douglas. 2003. "Identification of Multidimensional Spatial Voting Models." Working Paper, Stanford University (http://polmeth.wustl.edu/media/Paper/river03.pdf).