# 1 Method 1

Suppose we take time series data points, N (indexed by a) from each of S species.

$$X_i^a \rightarrow X^1 = \begin{bmatrix} x_1^1 \\ x_2^1 \\ \vdots \\ x_S^1 \end{bmatrix}, X^2 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_S^2 \end{bmatrix} ...., X^N = \begin{bmatrix} x_1^N \\ x_2^N \\ \vdots \\ x_S^N \end{bmatrix}$$

Goal: Maximizing$P(X|\mu, \sigma^2, d)$

Assumptions:

1). A is a symmetric S by S matrix
2). $A_{ij} = \frac{\mu}{S} + B_{ij}$, where $B_{ij}\ N(0, \frac{\sigma^2}{S}), i.e. A_{ij}\ N(\frac{\mu}{S}, \frac{\sigma^2}{S})$ 3). $A_{ii} = -d + \frac{\mu}{S}$
4). For A sufficiently large, the determinants of A for different random entries are similar so we use the average determinat of A in place of A. $< detA >= D(\mu, \sigma^2, d)$
Since, $x_i$ are independent, the average probability of observing $x_i$ given $\mu, \sigma, d$:

$$P(x|\mu, \sigma^2, d) = \int dA \prod (P(x^a|A)P(A|\mu, \sigma^2, d)$$

we can pull the product out and distribute it to each term

$$= \int (\prod_{i<j}(dA_{ij})\frac{(detA)^{N/2}}{(2\pi)^{NS/2}} exp(\frac{1}{2}\sum_{a=1}^{N}\sum_{ij}^{S} x_i^a A_{ij} x_j^a) P(A|\mu, \sigma^2, d)$$

Given our assumptions:

$$= \frac{< detA >^{N/2}}{(2\pi)^{NS/2}} \int (\prod_{i<j} dB_{ij}) exp(\frac{1}{2}\sum_{a=1}^{N}(-d\sum(x_i^a)^2 + \frac{\mu}{S}\sum_{ij} x_i^a x_j^a + \sum_{i<j} x_i B_{ij} x_j) P(B|\sigma^2))$$

pulling out the terms seperate from $B_{ij}$

$$= \frac{< detA >^{N/2}}{(2\pi)^{NS/2}} exp(\frac{-d}{2}\sum_{a=1}^{N}\sum_{i=1}^{S}(x_i^a)^2 + \frac{\mu}{2S}\sum_{a=1}^{N}(\sum_{i=1}^{S} x_i^a)^2 \times \int ((\prod_{ij} dB_{ij}) exp(\sum_{a=1}^{N}\sum_{i<j} x_i^a B_{ij} x_j^a) exp(-\sum_{i<j}\frac{B_{ij}^2}{2\sigma^2/S})$$

$$\times (\sqrt{\frac{S}{2\pi}}\frac{1}{\sigma}))$$

The firt part of the integral becomes

$$\prod_{i<j} dB_{ij} \rightarrow \prod_{i<j} exp(\frac{\sigma^2}{2S}(\sum_{a=1}^{N} x_i^a x_j^a)^2) = exp(\frac{\sigma^2}{4S}\sum_{ij}\sum_{a} x_i^a x_j^a \sum_{b} x_i^b x_j^b)$$

Then we can condense the summations and obtain:

$$exp(\frac{\sigma^2}{4S}\sum_{ab}(\sum_i x_i^a x_i^b)^2)$$

The portion in the exponent becomes

$$\prod_{i<j} exp(-\frac{B_{ij}^2}{2\sigma^2/S} + \sum_{a=1}^{N} x_i^a B_{ij} x_j^a)\sqrt{\frac{S}{2\pi}}\frac{1}{\sigma}$$

After rearranging the terms we get

$$\prod_{i<j} dB_{ij}\sqrt{\frac{S}{2\pi}}\frac{1}{\sigma} exp(-\frac{B_{ij}^2}{2\sigma^2/S} + \sum_{a=1}^{N} x_i^a B_{ij} x_j^a)$$

Let $y = B_{ij}, \tau = \frac{\sigma}{\sqrt{S}}, z = \sum_{a=1} x_i^a x_j^a$

Then our integral becomes :

$$\int dy\sqrt{\frac{1}{2\pi}}\frac{1}{\tau} exp(\frac{-y^2}{2\tau^2} + yz)$$

Which when solved, simplifies to:

$$exp(\frac{\tau^2 z^2}{2})$$

Now let us denote $P(x|\mu, \sigma, d)$ as P:

$$P = D \times exp(\frac{-d}{2}\sum_{a=1}^{N}\sum_{i=1}^{S}(x_i^a)^2 + \frac{\mu}{2S}\sum_{a=1}^{N}(\sum_{i=1}^{S} x_i^a)^2 + \frac{\sigma^2}{4S}\sum_{ab}(\sum_i x_i^a x_i^b)^2)$$

$$\frac{2}{N} logP = logD - \frac{d}{N}\sum_{a=1}^{N}\sum_{i=1}^{S}(x_i^a)^2 - \frac{\mu}{SN}\sum_{a=1}^{N}(\sum_{i=1}^{S} x_i^a)^2 + \frac{\sigma^2}{2SN}\sum_{ab}(\sum_i x_i^a x_i^b)^2$$

## 1.1 Optimization

To optimize P we take the partial derivatives wrt $\mu, \sigma$, and d and set them all equal to 0

1). $\frac{2}{N}\frac{\partial logP}{\partial d} = \frac{1}{D}\frac{\partial D}{\partial d} - \frac{1}{N}\sum_{a=1}^{N}\sum_{i=1}^{S}(x_i^a)^2 = 0$

2). $\frac{2}{N}\frac{\partial logP}{\partial \sigma^2} = \frac{1}{D}\frac{\partial D}{\partial \sigma^2} + \frac{1}{2SN}\sum_{ab}(\sum_i x_i^a x_i^b)^2 = 0$

3). $\frac{2}{N}\frac{\partial logP}{\partial \mu} = \frac{1}{D}\frac{\partial D}{\partial \mu} - \frac{1}{SN}\sum_{a=1}^{N}(\sum_{i=1}^{S} x_i^a)^2 = 0$

## 1.2 Determinant of A

The average determinant as given to us by the semicircular law in Random Matrix Theory-Robert Wiegner

$$< logD >= \sum_i log(\lambda_i) = log(d+\mu) + (S-1)\int d\lambda \frac{\sqrt{4\sigma^2-(\lambda+d)^2}}{2\pi\sigma} log(\lambda)$$

We evaluate this integral from $-d-2\sigma : -d+2\sigma$ as that is the radius in which all of our eigenvalues lie according to the semicircular law. Let $z = \frac{\lambda - d}{2\sigma}, b = \frac{2d}{\sigma}$ so $\lambda \to 2\sigma z + d, d\lambda \to 2\sigma dz$. Now we divide everything by $2\sigma$ and the integral:

$$\int_{-d/2\sigma-1}^{-d/2\sigma+1} d\lambda \frac{\sqrt{1-(\frac{\lambda-d}{2\sigma})^2}}{2\pi\sigma} log(\frac{\lambda}{2\sigma}) + log(2\sigma)$$

becomes:

$$I = \frac{2\sigma}{\pi}[log(2\sigma)\int_{-b-1}^{-b+1} dz\sqrt{1-z^2} + \int_{-b-1}^{-b+1} dz log(z+b)\sqrt{1-z^2}]$$

By using mathematica we solve the integral, where $F_{3,1}$ is the Hypergeometric function $HypegeometricPFQ[(1,1,\frac{3}{2}),(2,3),\frac{1}{b^2}]$:

$$I = log(d) - \frac{\sigma^2}{2d^2}[F_{3,11}(\frac{4\sigma^2}{d^2})]$$

Now we can rewrite the average log determinant as:

$$\frac{1}{S} < logD >= \frac{1}{S} log(d+\mu) + (1-\frac{1}{S})I$$

## 1.3 Solving the Optimization Equations

Now that we have the $< logD >$ we can solve optimization equations 1)., 2). and 3).

So first let $g(x) = \frac{1}{2}xF_{3,1}(4x)$, Then $g'(x) = \frac{1}{1+\sqrt{1-4x}-2x}$

**Optimization equation 3).** becomes

$$\frac{2}{N}\frac{\partial logP}{\partial \mu} = \frac{1}{d+\mu} - \frac{1}{SN}\sum_{a=1}^{N}(\sum_{i=1}^{S} x_i^a)^2 = 0$$

If we solve optimization equation 3). for $d+\mu$ we get $d+\mu = \frac{SN}{\sum_{a=1}^N(\sum_{i=1}^S x_i^a)^2}$. We replace $d+\mu$ in optimization equation 1). with this.

**Optimization equation 1).** becomes:

$$\frac{2}{N}\frac{\partial logP}{\partial d} = \frac{1}{d} + \frac{2\sigma^2}{d^3}g'(\frac{\sigma^2}{d^2}) - \frac{1}{SN}\sum_{a=1}^{N}\sum_{i=1}^{S}(x_i^a)^2 + \frac{1}{S^2N}\sum_{a=1}^{N}(\sum_{i=1}^{S}x_i^a)^2 = 0$$

$$\rightarrow \frac{1}{d} + \frac{2\sigma^2}{d^3}g'(\frac{\sigma^2}{d^2}) = \frac{1}{SN}\sum_{a=1}^{N}\sum_{i=1}^{S}(x_i^a)^2 - \frac{1}{S^2N}\sum_{a=1}^{N}(\sum_{i=1}^{S}x_i^a)^2$$

**Optimization equation 2).** becomes:

$$\frac{2}{N}\frac{\partial logP}{\partial \mu} = -\frac{1}{d^2}g'(\frac{\sigma^2}{d^2}) + \frac{1}{2SN}\sum_{ab}(\sum_{i}x_i^a x_i^b)^2 = 0$$

$$\rightarrow \frac{1}{S}\frac{1}{2SN}\sum_{ab}(\sum_{i}x_i^a x_i^b)^2 = \frac{1}{d^2}g'(\frac{\sigma^2}{d^2})$$

To solve for $\mu, \sigma$ and d we use rootsolvers in R...