# Generalized Model

## Circular Data Analysis

Our previous work reveals that the heading direction $s_m,\ m = 1, 2$ is subject to the normal distribution, that is, $s_m \sim \mathrm{N}(\mu, \sigma^2)$, where $\mu$ this the *mean direction* of $s_m$. The *mean resultant length* $\rho$ is defined as

$$\rho = \mathbb{E}[\cos(s_m - \mu)]. \tag{1}$$

Assuming $\sigma \ll \pi$ ($\kappa > 1$), the *mean resultant length* is calculated to be

$$\rho = \int_{-\pi}^{\pi} \cos(s_m) \frac{\exp[-\frac{(s_m-\mu)^2}{2\sigma^2}]}{\sigma\sqrt{2\pi}} ds_m \approx \int_{-\infty}^{\infty} \cos(s_m) \frac{\exp[-\frac{(s_m-\mu)^2}{2\sigma^2}]}{\sigma\sqrt{2\pi}} ds_m = e^{-\frac{\sigma^2}{2}}. \tag{2}$$

## Definition of Circular Variance

If the *circular variance* $\hat{\sigma}^2$ is defined as

$$\hat{\sigma}^2 \equiv -2\ln\rho, \tag{3}$$

then $\hat{\sigma}^2$ is equivalent to $\sigma^2$. We are using $\frac{1}{\hat{\sigma}^2}$ to represent the *concentration parameter* $\hat{\kappa}$.

However, Zhang et al. assume $s_m$ can be represented by a von Mises distribution, that is, $s_m \sim \mathrm{M}(\hat{s}_m, \hat{\kappa})$. The maximum likelihood of $\hat{\kappa}$ is the solution to

$$\rho = A(\hat{\kappa}), \tag{4}$$

where $A(x) \equiv I_1(x)/I_0(x)$. The solution is not easy to find for a large $\hat{\kappa}$ unless we are using the following approximation (see Fisher (1993) page 88 and Mardia & Jupp (2000) pages 85-86)

$$\hat{\kappa} = \begin{cases} 2\rho + \rho^3 + \frac{5\rho^5}{6}, & \rho < 0.53 \\ -0.4 + 1.39\rho + \frac{0.43}{1-\rho}, & 0.53 \le \rho < 0.85 \\ \frac{1}{3\rho - 4\rho^2 + \rho^3}, & \rho \ge 0.85. \end{cases} \tag{5}$$

Consider the Taylor expansion of $\hat{\kappa}$ centered at $\rho = 1$ for a large $\hat{\kappa}$, then $\hat{\kappa} \approx \frac{1}{3\rho-4\rho^2+\rho^3} \approx \frac{1}{2(1-\rho)} = \frac{1}{2(1-e^{-\sigma^2/2})} \approx \frac{1}{\sigma^2}$ is fine for $\rho \to 1$. That is reason why we still get correct answer while using the Gaussian noise. However, after introducing the output-dependent noise, the concentration parameter is much smaller so that we can no longer use the von Mises distribution to approximate a normal distribution. That is, we shall use Eq. (3) to calculate the variance for circular data.