Statistics on Stiefel Manifolds.

The uniform distribution on $V_k(\mathbb{R}^n)$ is given by the usual (invariant) metric.

Given observations $X_1, \dots, X_n$ on $V_k(\mathbb{R}^n)$, we can regard them as $n \times k$ matrices and compute

$$\overline{X} = \frac{1}{n} \sum X_i$$

We can define a Rayleigh test for uniformity by measuring the (Frobenius norm) size of this matrix.

$$S = pn \; tr(\overline{X}^T \overline{X})$$

$$= pn \sum_{i,j} \overline{X}_{ij}^2$$

Under uniformity, the entries in a random ~~matrix~~ matrix in $V_k(\mathbb{R}^n)$ are ~ normal, so for large k,n we have

$$S \sim \chi^2_{nk}$$

As for the circle, we can compute the modified statistic

$$S^* = S\left(1 - \frac{1}{2n}\left(1 - \frac{S}{nk+2}\right)\right).$$

which converges faster to χ^2_{nk}.

Another simple test is to test the first vector in the frame for uniformity on S^{n-1}.

We can test uniformity on $G_k(\mathbb{R}^n)$ with the Bingham test: Let $\Psi_1, \dots, \Psi_m$ be (matrix) observations of $n \times n$ projection matrices encoding elements in $G_k(\mathbb{R}^n)$.

We expect the mean over $G_k(\mathbb{R}^n)$ of such matrices to be $(k/n) I_n$. So we compute

$$S = \frac{(n-1)n(n+2)}{2k(n-k)} m \left(\operatorname{tr}(\bar{\Psi}^2) - \frac{k^2}{n} \right)$$

Under uniformity, this converges to

$$\chi^2_{(n+1)(n+2)/2}$$

The Bingham test on the Grassmann can be corrected by

$$S^* = S\left(1 - \frac{1}{m}\left(B_0 + B_1 S + B_2 S^2\right)\right)$$

where

$$\cancel{B_0 = p^2(p^2+p-p}$$

$$B_0 = \frac{n^2(n^2+n-2) + 2K(n-K)(n^2+4n-20)}{12K(n-K)(n-2)(n-1)(n+4)(n+2)}$$

$$B_1 = -\frac{n^2(n^2+n-2) - K(n-K)(n^2-2n+16)}{3K(n-K)(n^2+n+2)(n-2)(n+4)}$$

$$B_2 = \frac{(n-2K)^2(n-1)(n+2)}{3K(n-K)(n-2)(n+4)(n^2+n+2)(n^2+n+6)}$$

to converge to $\chi^2_{(n-1)(n+2)/2}$ at rate $1/m$ instead of $1/\sqrt{m}$.

The simplest non-uniform distribution on a Stiefel manifold $V_K(\mathbb{R}^n)$ is the matrix Fisher (or Langevin) distribution given by

$$f(X;F) = \left({}_0F_1\left(\tfrac{n}{2};\tfrac{1}{4}F^TF\right)\right)^{-1} e^{\operatorname{tr} F^TX}$$

where F is an $n\times K$ parameter matrix, and

$${}_0F_1\left(\tfrac{n}{2};\tfrac{1}{4}F^TF\right) = \int_{V_K(\mathbb{R}^n)} e^{\operatorname{tr}(F^TX)}\, d\mathrm{Vol}.$$

This distribution has a mode given by the polar part of F.

We can use this ~~test~~ ~~to~~ distribution to define a mean on the Stiefel manifold.

Suppose that we have observations $X_1, \dots, X_n$ from a matrix Fisher distribution based on F. Since the X_i are matrices, we can average them (entrywise) to obtain $\bar{X} = \frac{1}{n} \sum X_i$.

Now take the polar decomposition of $\bar{X}$. The maximum likelihood estimate of the polar part of F is the polar part of $\bar{X}$, and the "elliptical part" ~~is~~ K in the polar decomp $F = MK$ can be estimated in a complicated way from

$$\bar{R} = (\bar{X}^T \bar{X})^{1/2}.$$

~~and equivariance~~

In general, if

$$\bar{R} = \mathcal{U}^T \operatorname{diag}(g_1, \dots, g_K)\, \mathcal{U}$$

is the SVD of $\bar{R}$, then the MLE of K

$$\hat{K} = \mathcal{U}^T \operatorname{diag}(\hat{\varphi}_1, \dots, \hat{\varphi}_K)\, \mathcal{U}$$

where

$$g_i = \frac{\partial}{\partial \varphi_i} {}_0F_1\left(\frac{n}{2}; \frac{1}{4} \operatorname{diag}(\varphi_1^2, \dots, \varphi_K^2)\right)\Bigg|_{(\varphi_1, \dots, \varphi_K) = (\hat{\varphi}_1, \dots, \hat{\varphi}_K)}.$$

Generally for small φ_i,

$$\hat{\varphi}_i \simeq n g_i.$$

and for large φ_i, we can solve the system of equations

$$\left(n-K-\frac{1}{2}\right)\hat{\varphi}_i^{-1}+\sum_{j=1}^{K}\left(\hat{\varphi}_i+\hat{\varphi}_j\right)^{-1}\simeq 2(1-g_i)$$

for the MLE estimates $\hat{\varphi}_i$.