Supplementary: Nonparametric Fisher Geometry with Application to Density Estimation

Babak Shahbaba, Shiwei Lan, Jeffrey D. Streets, and Andrew J. Holbrook

1 Some basic results

In this section, we present a few basic results regarding the nonparametric Fisher geometry, working with the L^2 sphere model and transferring results to the traditional Fisher geometry. We note that investigations of the nonparameteric Fisher information have independently appeared in (Srivastava, Jermyn, and Joshi, 2007; Chen, Streets, and Shahbaba, 2015; Itoh and Satoh, 2015; Kurtek and Bharath, 2015; Srivastava and Klassen, 2016; Peter, Rangarajan, and Moyou, 2017). We reproduce some fundamental aspects of this geometry relevant to Theorem 1 for convenience. The proofs are provided later.

Consider the infinite-dimensional L^2 sphere. This space is isometric to the space of density functions equipped with the nonparameteric Fisher metric.

Proposition 1. *The map $S : (\mathcal{P}, g_F) \to (\mathcal{Q}, \langle\cdot,\cdot\rangle_{L^2})$ defined by $S(p) := 2\sqrt{p}$ is a Riemannian isometry.*

We describe the tangent space to $\mathcal{Q}$ as follows:

Proposition 2. *Given $q \in \mathcal{Q}$, one has that*

$$T_q\mathcal{Q} := \left\{ f : \mathcal{D} \to \mathbb{R} \mid \int_{\mathcal{D}} q(x) f(x)\, \mu(dx) = 0 \right\}. \tag{1}$$

We next solve one version of the geodesic problem on $\mathcal{P}$. In particular we consider an initial point and velocity and solve for continuing the geodesic in that direction. This result is relevant to our discussion of the relationship between Riemannian HMC and infinite-dimensional spherical HMC in Section 4.1. We will exploit the isometry between $\mathcal{P}$ and $\mathcal{Q}$ and solve first in $\mathcal{Q}$.

Lemma 1. *Given $q_0 \in \mathcal{Q}$ and $f \in T_q\mathcal{Q}$ a unit vector, the geodesic with initial condition q_0 and velocity f exists on $(-\infty, \infty)$ and takes the form*

$$q_t = q_0 \cos t + f \sin t. \tag{2}$$

We now translate this result into a corresponding one for geodesics in $\mathcal{P}$.

Lemma 2. *Given $p_0 \in \mathcal{P}$ and $f \in T_p\mathcal{P}$ a unit vector, the geodesic with initial condition p_0 and initial velocity f exists on $(-\infty, \infty)$, and takes the form*

$$p_t = \left(\sqrt{p_0}\cos t + \frac{f}{2\sqrt{p_0}}\sin t\right)^2. \tag{3}$$

Lemma 3. *Geodesic flows on the finite sphere $\mathcal{S}^{I-1}$ converge to geodesic flows on the infinite-dimensional sphere $\mathcal{S}^\infty$ as $I \to \infty$.*

These basic lemmas show the advantage of working in $\mathcal{Q}$, yielding a conceptual derivation of the geodesic equation. These lemmas will be used to prove Theorem 1.

2 Proofs

Proposition 1. *The map $S : (\mathcal{P}, g_F) \to (\mathcal{Q}, \langle\cdot,\cdot\rangle_{L^2})$ defined by $S(p) := 2\sqrt{p}$ is a Riemannian isometry.*

Proof. We must show that $\langle S_*\psi, S_*\phi\rangle_{L^2} = g_F(\psi,\phi)_p$, where S_* is the pushforward (or Jacobian) of S:

$$S_* = \frac{dS}{dp}(p) = \frac{d(2\sqrt{p})}{dp} = \frac{1}{\sqrt{p}}. \tag{4}$$

By direct computation,

$$\begin{aligned}\langle S_*\psi, S_*\phi\rangle_{L^2} &= \int_{\mathcal{D}} (S_*\psi)(x)\,(S_*\phi)(x)\,\mu(dx) = \int_{\mathcal{D}} \frac{\psi(x)}{\sqrt{p(x)}}\,\frac{\phi(x)}{\sqrt{p(x)}}\mu(dx) \\ &= \int_{\mathcal{D}} \frac{\psi(x)\phi(x)}{p(x)}\mu(dx) = g_F(\psi,\phi)_p\,.\end{aligned} \tag{5}$$

□

Proposition 2. *Given $q \in \mathcal{Q}$, one has that*

$$T_q\mathcal{Q} := \left\{ f : \mathcal{D} \to \mathbb{R} \mid \int_{\mathcal{D}} q(x)f(x)\,\mu(dx) = 0 \right\}. \tag{1}$$

Proof. If $q_t : (-\epsilon, \epsilon) \to \mathcal{Q}$ denotes a path in Q satisfying $dq_t/dt|_{t=0} = f$, then the unit integration constraint on $p = q^2$ means

$$0 = \frac{d}{dt}\int_{\mathcal{D}} q_t(x)^2\mu(dx)\Big|_{t=0} = 2\int_{\mathcal{D}} q_0(x)\frac{dq}{dt}(x)\Big|_{t=0}\mu(dx) = 2\int_{\mathcal{D}} q_0(x)f(x)\mu(dx)\,. \tag{6}$$

□

Lemma 1. *Given $q_0 \in \mathcal{Q}$ and $f \in T_q\mathcal{Q}$ a unit vector, the geodesic with initial condition q_0 and velocity f exists on $(-\infty, \infty)$ and takes the form*

$$q_t = q_0 \cos t + f \sin t. \tag{2}$$

Proof. First we derive the geodesic equation in $\mathcal{Q}$. One conceptual method for obtaining this, exploiting the spherical structure of $\mathcal{Q}$, is to first observe that if q_t is a path in $\mathcal{Q}$ and $a_t \in T_{q(t)}\mathcal{Q}$ is a tangent vector along the curve, the Fisher geometry induces a covariant derivative along the path via

$$\frac{D}{\partial t} a = \dot{a} - q \int_{\mathcal{D}} \dot{a} q, \tag{7}$$

which is manifestly the time derivative of the family a_t projected to the tangent space at q_t, as expected. For a curve q_t to be a geodesic, it should have zero acceleration, i.e.

$$0 = \frac{D}{\partial t} \dot{q} = \ddot{q} - q \int_{\mathcal{D}} \ddot{q} q. \tag{8}$$

However, using that $\int_{\mathcal{D}} q_t(x)^2 \, \mu(dx) = 1$ for all q and differentiating twice in t, one sees that this is equivalent to

$$\ddot{q} + q \int_{\mathcal{D}} \dot{q}^2 = 0, \tag{9}$$

which we now take as the geodesic equation in $\mathcal{Q}$. Another method for deriving this equation is to solve for which curves are critical points for the length functional with fixed endpoints.

Now, to solve this equation in our setting, first let us observe that since $f \in T_{q_0}\mathcal{Q}$, by Lemma 2 we have

$$\int_{\mathcal{D}} q_0 f = 0. \tag{10}$$

Using this and the fact that f is a unit vector we compute

$$\begin{aligned} \frac{d}{dt} \int_{\mathcal{D}} \dot{q}^2 &= 2 \int_{\mathcal{D}} \ddot{q} \dot{q} = 2 \int_{\mathcal{D}} \left[-q_0 \cos t - f \sin t\right] \left[-q_0 \sin t + f \cos t\right] \\ &= 2 \int_{\mathcal{D}} \left[q_0^2 - f^2\right] \cos t \sin t = 0 \end{aligned} \tag{11}$$

Thus $\int_M \dot{q}^2 = \int_{\mathcal{D}} f^2 = 1$. We then simply observe the ODE

$$\ddot{q} = -q, \tag{12}$$

and it is clear that q satisfies (9), and so the lemma follows. $\square$

Lemma 2. *Given $p_0 \in \mathcal{P}$ and $f \in T_p\mathcal{P}$ a unit vector, the geodesic with initial condition p_0 and initial velocity f exists on $(-\infty, \infty)$, and takes the form*

$$p_t = \left(\sqrt{p_0}\cos t + \frac{f}{2\sqrt{p_0}}\sin t\right)^2. \tag{3}$$

Proof. We use Lemma 1 and reinterpret the geodesic equation in terms of square-roots. In this formalism the initial condition is $q_0 = \sqrt{p_0}$ and the initial velocity is

$$\frac{d}{dt}q = \frac{d}{dt}\sqrt{p} = \frac{f}{2\sqrt{p_0}} = \frac{f}{2q_0}.$$

$\square$

Lemma 3. *Geodesic flows on the finite sphere $\mathcal{S}^{I-1}$ converge to geodesic flows on the infinite-dimensional sphere $\mathcal{S}^\infty$ as $I \to \infty$.*

Proof. For any point $q \in \mathcal{S}^\infty$, let $q^I \in \mathcal{S}^{I-1}$ be vector obtained by applying the truncation operator to q and then normalizing:

$$q^I = \frac{T_I(q)}{\|T_I(q)\|} = \frac{(q_1, \ldots, q_I)^T}{\sqrt{(q_1, \ldots, q_I)(q_1, \ldots, q_I)^T}}. \tag{13}$$

Similarly, for any vector in the tangent space to S^∞

$$v \in T_q\mathcal{S}^\infty = \left\{ v \in \ell^2 \middle| \langle v, q\rangle_{\ell^2} = \sum_{i=1}^{\infty} q_i v_i = 0 \right\} \tag{14}$$

let $v^I \in T_{q^I}\mathcal{S}^{I-1}$ be the I-dimensional vector obtained by truncating v, projecting onto the tangent space $T_{q^I}\mathcal{S}^{I-1}$, and scaling such that $\|v\|_{\ell^2} = \|v^I\|$ (where $\|\cdot\|$ is the Euclidean norm):

$$\tilde{v}^I = T_I(v) - q^I \langle q^I, T_I(v)\rangle_{\ell^2}, \quad \text{and} \quad v^I = \tilde{v}^I \frac{\|v\|_{\ell^2}}{\|\tilde{v}^I\|}. \tag{15}$$

It follows from the definition of truncation that $q^I \to q$ and $v^I \to v$ with respect to $\langle \cdot, \cdot \rangle_{\ell^2}$ as $I \to \infty$.

Next, let $t \mapsto (q(t), v(t))$ be the geodesic flow on $\mathcal{S}^\infty$ with initial position $q_0 = q(0)$ and initial velocity $v_0 = v(0) \in T_{q_0}\mathcal{S}^\infty$. Let $t \mapsto (q^I(t), v^I(t))$ be the analogous flow on the tangent bundle $T\mathcal{S}^{I-1}$, where q_0^I and v_0^I are obtained from q_0 and v_0 following Formulas (14) and (15), respectively. Denote the distance between flows at time t

$$f(t) = \|q_t - q_t^I\|_{\ell^2}^2 + \|\dot{q}_t - \dot{q}_t^I\|_{\ell^2}^2 . \tag{16}$$

Our goal is to show that

$$\lim_{I\to\infty} \int_0^T f(t)\, dt = 0 , \tag{17}$$

for any finite T, and hence that geodesic flows on the finite sphere converge to those on $\mathcal{S}^\infty$. Begin by bounding $\dot{f}(t)$ by a constant times $f(t)$:

$$\begin{aligned} \frac{d}{dt} f(t) &= 2\left(\langle q_t - q_t^I, \dot{q}_t - \dot{q}_t^I\rangle_{\ell^2} + \langle \dot{q}_t - \dot{q}_t^I, \ddot{q}_t - \ddot{q}_t^I\rangle_{\ell^2}\right) \\ &= 2\left(\langle q_t - q_t^I, \dot{q}_t - \dot{q}_t^I\rangle_{\ell^2} + \langle \dot{q}_t - \dot{q}_t^I, -q_t\|\dot{q}_t\|_{\ell^2}^2 + q_t^I\|\dot{q}_t^I\|^2\rangle_{\ell^2}\right). \end{aligned} \tag{18}$$

Here, the second line follows from the geodesic formula. Noting that $\|\dot{q}_t\|_{\ell^2}^2 = \|\dot{q}_0\|_{\ell^2}^2$, $\|\dot{q}_t^I\|^2 =$

$\|\dot{q}_0^I\|^2$, and that (by Equation (15)) $\|\dot{q}_0^I\|^2 = \|\dot{q}_0\|_{\ell^2}^2$, we get

$$\frac{d}{dt} f(t) = 2\left(\langle q_t - q_t^I, \dot{q}_t - \dot{q}_t^I\rangle_{\ell^2} - \langle \dot{q}_t - \dot{q}_t^I, q_t - q_t^I\rangle_{\ell^2} \|\dot{q}_0\|_{\ell^2}^2\right) \tag{19}$$

$$= 2\big(1 - \|\dot{q}_0\|_{\ell^2}^2\big)\langle q_t - q_t^I, \dot{q}_t - \dot{q}_t^I\rangle_{\ell^2} \,.$$

We obtain our bounds by noting that

$$0 \leq \|q_t - q_t^I\|_{\ell^2}^2 - 2\langle q_t - q_t^I, \dot{q}_t - \dot{q}_t^I\rangle_{\ell^2} + \|\dot{q}_t - \dot{q}_t^I\|_{\ell^2}^2 \tag{20}$$

$$= f(t) - 2\langle q_t - q_t^I, \dot{q}_t - \dot{q}_t^I\rangle_{\ell^2} \tag{21}$$

$$= f(t) - \frac{\dot{f}(t)}{1 - \|\dot{q}_0\|_{\ell^2}^2} \,,$$

and hence that

$$\frac{d}{dt} f(t) \leq \big(1 - \|\dot{q}_0\|_{\ell^2}^2\big)\, f(t) \,. \tag{22}$$

Integrating gives

$$f(t) \leq f(0)\, e^{t\,(1 - \|\dot{q}_0\|_{\ell^2}^2)} \,. \tag{23}$$

Since, by definition, $f(0) \to 0$ as $I \to \infty$, we have

$$\int_0^T f(t)\, dt \leq f(0) \int_0^T e^{t\,(1 - \|\dot{q}_0\|_{\ell^2}^2)}\, dt \tag{24}$$

$$= c\, f(0) \longrightarrow 0 \,.$$

Thus we have proven the convergence of geodesic flows on the finite sphere to those on $\mathcal{S}^\infty$. $\square$

Theorem 1. *Let $q(\cdot) = \sqrt{p(\cdot)} \in \mathcal{Q}$ be a square-root density function with expansion satisfying*

$$q(\cdot) = \sum_{i=1}^{\infty} q_i \phi_i(\cdot)\,, \text{ and}$$
$$1 = \int_{\mathcal{D}} q(x)^2 \, \mu(dx) = \sum_{i=1}^{\infty} q_i^2\,,$$

with random, real-valued coefficients $q_i, i = 1, \ldots, \infty$. Then, in the infinite-dimensional limit, spherical HMC follows the nonparametric Fisher metric's geodesic flows in the same way that Riemannian HMC follows the Fisher metric's geodesic flows over the parametric family of distributions $\mathcal{P}_\theta$.

Proof. Each of these algorithms relies on a split Hamiltonian (Shahbaba et al., 2014) integration scheme, wherein the Hamiltonian of interest (H) is split into two Hamiltonians $(H^1 + H^2)$ that are then iteratively simulated. The formal Hamiltonian for spherical HMC on $lim_{I\to\infty}\mathcal{S}^{I-1} = \mathcal{S}^\infty$ has the same form as before, but in this case the velocity v is restricted to the tangent space to $\mathcal{S}^\infty$ at q, $T_q\mathcal{S}^\infty$. The Hamiltonian corresponding to Riemannian HMC is also split in the following way (Byrne and Girolami, 2013):

$$\begin{aligned} H(\theta,\xi) &= -\log p(\theta) + \frac{1}{2}\log \mathcal{I}(\theta) + \frac{1}{2}\xi^T \mathcal{I}^{-1}(\theta)\xi \\ H^1(\theta,\xi) &= -\log p(\theta) + \frac{1}{2}\log \mathcal{I}(\theta) \\ H^2(\theta,\xi) &= \frac{1}{2}\xi^T \mathcal{I}^{-1}(\theta)\xi\,, \end{aligned} \tag{25}$$

where $\mathcal{I}(\theta)$ is the Fisher information, and ξ is the auxiliary momentum variable.

Switching out $\xi(t)$ for $\nabla_\theta \ell(\theta(t))$, it follows that the solutions to the Hamilton's equations for Hamiltonian $H^2(\theta,\xi)$ are the geodesics on the Riemannian manifold $(\mathcal{P}_\theta, g_F)$. This is because the Hamiltonian flow $\theta(t)$ preserves $H^2(\theta, \xi)$:

$$\frac{d}{ds}E(\theta) = \frac{d}{ds}\frac{1}{2}\int_a^b \xi(t)^T \mathcal{I}(\theta(t))^{-1}\xi(t) = \frac{d}{ds}\frac{b-a}{2}\xi(a)^T \mathcal{I}(\theta(a))^{-1}\xi(a) = 0\,. \tag{26}$$

Thus, Riemannian HMC steps around the state space by minimizing the parametric Fisher energy.

In the same way, by exchanging $v(t)$ for $\dot{q}(t)$, it follows that the solutions to the Hamilton's equations for Hamiltonian $H^2(q, v)$ are geodesics on the Riemannian manifold $(\mathcal{S}^\infty, \langle\cdot,\cdot\rangle_{\ell^2})$ and, by Lemmas 1 and 2, correspond to geodesics on $(\mathcal{P}, g_F(\cdot,\cdot))$. Hence, both formal algorithms move around the state space by iteratively perturbing the velocity (H^1) and travelling the geodesics corresponding to the parametric and nonparametric Fisher geometries, respectively. Finally, Lemma 3 guarantees that the finite-dimensional spherical geodesics (used in practice) pass in the limit to the geodesics of the sphere S^∞ and hence of $(\mathcal{P}, g_F(\cdot,\cdot))$. $\square$

3 Initializing the Markov chain: Newton's method on the sphere

Starting with a Riemannian manifold $\mathcal{Q}$ isometrically embedded in Euclidean space, we consider function $F : \mathcal{Q} \to \mathbb{R}$.

Definition 1. *Given point $q_0 \in \mathcal{Q}$ and initial velocity $\dot{q}_0 \in T_{q_0}\mathcal{Q}$, we follow Edelman, Arias, and Smith (1998) and define the Hessian of function F along $\dot{q}_0$ as the matrix satisfying*

$$Hess\, F(\dot{q}_0, \dot{q}_0) = \frac{d^2}{dt^2}\Big|_{t=0} F(q(t))\,. \tag{27}$$

Proposition 3. *Hess F on the sphere is given by*

$$Hess\, F = F_{qq} - F_q^T q_0\, \mathcal{I}\,, \tag{28}$$

where F_q and F_{qq} are the Jacobian and usual Hessian matrices.

Proof. We need the formula for the geodesic on the sphere given q_0 and $\dot{q}_0$. Letting α be the Euclidean norm of $\dot{q}_0$, the geodesic is given by:

$$q(t) = q_0 \cos(\alpha t) + \frac{\dot{q}_0}{\alpha} \sin(\alpha t)\,. \tag{29}$$

It is easy to verify that

$$\ddot{q}(t) = -\alpha^2\, q(t)\,. \tag{30}$$

Next the derivatives are given by:

$$\frac{d}{dt} F(q(t)) = \frac{\partial F}{\partial q}(y(t))\dot{q}(t)\,, \tag{31}$$

and

$$\frac{d^2}{dt^2} F(q(t)) = \dot{q}(t)^T \frac{\partial^2 F}{\partial q^2} \dot{q}(t) + \frac{\partial F}{\partial q}^T \ddot{q}(t)\,. \tag{32}$$

Combining (30) with (32) gives:

$$\begin{aligned} \frac{d^2}{dt^2} F(q(t)) &= \dot{q}(t)^T \frac{\partial^2 F}{\partial q^2} \dot{q}(t) - \alpha^2 \frac{\partial F}{\partial q}^T q(t) \\ &= \dot{q}(t)^T \Big(F_{qq} - F_q^T q(t)\, \mathcal{I} \Big) \dot{q}(t)\,. \end{aligned} \tag{33}$$

Evaluating at $t = 0$ gives the result. □

Hess F is the Hessian matrix at point q_0 in direction $\dot{q}_0$. Newton's method on the sphere is achieved by Algorithm 1.

Algorithm 1 A single iteration of Newton's method on the sphere

---

1: Given point q on sphere:

2: Calculate F_q

3: Calculate Hess $F = F_{qq} - F_q^T q_0 \mathcal{I}$

4: Calculate $W = (\mathcal{I} - qq^T)\, Hess^{-1} F\, (\mathcal{I} - qq^T)$

5: $V \leftarrow -W\, F_q$

6: Progress along geodesic (29) with initial velocity V for time 1.

7: $q \leftarrow q(1)$

---

4 Moments of data distribution

We now provide formulas for calculating probability of an interval along with the first and second moments for data distribution described by square-root density function

$$\sqrt{p(x)} = q(x) = \sum_{i=0}^{I} q_i \sqrt{2} \cos(\pi i x)\,.$$

Based on the formulas we provide, one can easily program all functionals and compute them in time $\mathcal{O}(I^2)$.

4.1 Measure of an interval

We wish to calculate the probability unobserved data will lie in an interval $[a, b]$ for $0 \leq a, b \leq 1$ or $\Pr(x \in [a, b])$:

$$\begin{aligned}\int_a^b p(x)\,\mathrm{d}x = \int_a^b q^2(x)\,\mathrm{d}x &= \int_a^b \left(\sum_{i=0}^{I} q_i \sqrt{2} \cos(\pi i x)\right)^2 \mathrm{d}x \\ &= 2 \sum_{i,j=0}^{I} q_i q_j \int_a^b \cos(\pi i x) \cos(\pi j x)\,\mathrm{d}x\end{aligned}$$

We must therefore be able to compute

$$\int_a^b \cos(\pi i x) \cos(\pi j x)\,\mathrm{d}x$$

for four scenarios:

a $i = j = 0$

b $i = j > 0$

c $i = 0$ and $j > i$

d $0 < i < j$.

Trivially, when $i = j = 0$:

$$\int_a^b \cos(\pi i x) \cos(\pi j x)\,\mathrm{d}x = b - a\,.$$

For the second scenario:

$$\int_a^b \cos^2(\pi i x)\,\mathrm{d}x = \frac{1}{4\pi i}\left(\sin(2\pi i b) - \sin(2\pi i a) + 2\pi i(b - a)\right)\,.$$

For the third scenario:

$$\int_a^b \cos(\pi j x)\,\mathrm{d}x = \left(\sin(\pi j b) - \sin(\pi j a)\right)/(\pi j)\,.$$

And for the fourth scenario:

$$\int_a^b \cos(\pi i x)\cos(\pi j x)\,\mathrm{d}x = \frac{1}{2\pi}\left(\frac{\sin(\pi x(i-j))}{i-j} + \frac{\sin(\pi x(i+j))}{i+j}\right)\Big|_a^b$$

4.2 First moment

We would like to compute

$$E_p(x) = \int_0^1 x\,p(x)\,\mathrm{d}x = \int_0^1 x\,q^2(x)\,\mathrm{d}x = 2\sum_{i,j=0}^{I} q_i q_j \int_0^1 x\,\cos(\pi i x)\cos(\pi j x)\,\mathrm{d}x$$

Again we consider the four scenarios. When $i = j = 0$

$$\int_0^1 x\,\cos(\pi i x)\cos(\pi j x)\,\mathrm{d}x = \frac{1}{2}\,.$$

For $i = j > 0$,

$$\int_0^1 x\,\cos^2(\pi i x)\,\mathrm{d}x = \frac{1}{4}\,.$$

For $i = 0$ and $j > i$

$$\int_0^1 x\,\cos(\pi j x)\,\mathrm{d}x = \begin{cases} 0 & j \text{ even} \\ -\frac{2}{\pi^2 j^2} & j \text{ odd} \end{cases}\,.$$

Finally, for $0 < i < j$:

$$\int_0^1 x\, \cos(\pi i x) \cos(\pi j x)\, \mathrm{d}x = \begin{cases} 0 & j, i \text{ both even or both odd} \\ -\frac{1}{\pi^2}\left(\frac{1}{j-i} + \frac{1}{j+i}\right) & j, i \text{ different parity} \end{cases} .$$

4.3 Second moment

We would like to compute

$$E_p(x^2) = \int_0^1 x^2\, p(x)\, \mathrm{d}x = \int_0^1 x^2\, q^2(x)\, \mathrm{d}x = 2 \sum_{i,j=0}^{I} q_i q_j \int_0^1 x^2\, \cos(\pi i x) \cos(\pi j x)\, \mathrm{d}x$$

Again we consider the four scenarios. For $i = j = 0$,

$$\int_0^1 x^2\, \cos(\pi i x) \cos(\pi j x)\, \mathrm{d}x = \frac{1}{3} .$$

For $i = j > 0$,

$$\int_0^1 x^2\, \cos^2(\pi i x)\, \mathrm{d}x = \frac{1}{6} + \frac{1}{4\pi^2 j^2} .$$

For $i = 0$ and $j > i$

$$\int_0^1 x^2\, \cos(\pi j x)\, \mathrm{d}x = \begin{cases} \frac{2}{\pi^2 j^2} & j \text{ even} \\ -\frac{2}{\pi^2 j^2} & j \text{ odd} \end{cases} .$$

Finally, for $0 < i < j$:

$$\int_0^1 x^2\, \cos(\pi i x) \cos(\pi j x)\, \mathrm{d}x = \begin{cases} \frac{1}{\pi^2 (j-i)^2} + \frac{1}{\pi^2 (j+i)^2} & j, i \text{ both even or both odd} \\ \frac{-1}{\pi^2 (j-i)^2} + \frac{-1}{\pi^2 (j+i)^2} & j, i \text{ different parity} \end{cases} .$$

References

Srivastava, Anuj, Ian Jermyn, and Shantanu Joshi (2007). "Riemannian analysis of probability density functions with applications in vision". In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, pp. 1–8.

Chen, Tian, Jeffrey Streets, and Babak Shahbaba (2015). "A Geometric View of Posterior Approximation". In: *arXiv preprint arXiv:1510.00861*.

Itoh, Mitsuhiro and Hiroyasu Satoh (2015). "Geometry of Fisher information metric and the barycenter map". In: *Entropy* 17.4, pp. 1814–1849.

Kurtek, Sebastian and Karthik Bharath (2015). "Bayesian sensitivity analysis with the Fisher–Rao metric". In: *Biometrika* 102.3, pp. 601–616.

Srivastava, Anuj and Eric P Klassen (2016). *Functional and shape data analysis*. Springer.

Peter, Adrian M, Anand Rangarajan, and Mark Moyou (2017). "The Geometry of Orthogonal-Series, Square-Root Density Estimators: Applications in Computer Vision and Model Selection". In: *Computational Information Geometry*. Springer, pp. 175–215.

Shahbaba, Babak, Shiwei Lan, Wesley O Johnson, and Radford M Neal (2014). "Split Hamiltonian Monte Carlo". In: *Statistics and Computing* 24.3, pp. 339–349.

Byrne, Simon and Mark Girolami (2013). "Geodesic Monte Carlo on embedded manifolds". In: *Scandinavian Journal of Statistics* 40.4, pp. 825–845.

Edelman, Alan, Tomás A Arias, and Steven T Smith (1998). "The geometry of algorithms with orthogonality constraints". In: *SIAM journal on Matrix Analysis and Applications* 20.2, pp. 303–353.