Notes on Kernels for Normalized Functions

Carlo Graziani

May 4, 2023

1 Introduction

Suppose that we want a GP model for a function $n(k)$ that satisfies a momentum-space normalization of the type

$$\int_0^\infty \frac{4\pi dk}{(2\pi)^3}\, k^2 n(k) = N$$

or, in general

$$\int_0^\infty dk\, k^2 n(k) = A$$

where $A = 2\pi^2 N$. If we have some data on $n(k)$, we could fit a vanilla GP to it, but there would be no guarantee that any function that one sampled from it would satisfy the normalization condition.

Instead, suppose we start from a zero-mean GP with base covariance kernel $\langle n(k_1)n(k_2)\rangle = C(k_1, k_2)$. We can proceed by updating this kernel based on a zero-noise observation of the normalization integral

$$I[n(\cdot)] \equiv \int_0^\infty dk\, k^2 n(k). \tag{1.1}$$

The required covariance relations are

$$\begin{aligned}
\langle I \times n(k_1)\rangle &= \int_0^\infty dk_2\, k_2^2 \,\langle n(k_2)n(k_1)\rangle \\
&= \int_0^\infty dk_2\, k_2^2 \times C(k_1,k_2) \\
&\equiv C_1(k_1), \qquad (1.2)\\
\langle I^2\rangle &= \int_0^\infty dk_1\, k_1^2 \int_0^\infty dk_2\, k_2^2 \,\langle n(k_2)n(k_1)\rangle \\
&= \int_0^\infty dk_1 \int_0^\infty dk_2\, k_1^2 k_2^2 C(k_1,k_2) \\
&\equiv C_0. \qquad (1.3)
\end{aligned}$$

Given a choice of $C(\cdot,\cdot)$ that permits these integrals to be computed, the updated GP conditioned on the observation $I[n(\cdot)] = A$ has a non-zero mean function $\mu(k)$ given by the standard formula

$$\mu(k) = C_1(k) \times C_0^{-1} \times A, \tag{1.4}$$

and a covariance kernel $K(k_1, k_2)$ given by the equally-standard formula

$$K(k_1, k_2) = C(k_1, k_2) - C_1(k_1)C_0^{-1}C_1(k_2). \tag{1.5}$$

Note that $\int_0^\infty dk\, \mu(k) = C_0 C_0^{-1} A = A$, and functions $n(k)$ sampled from $GP\left[\mu(\cdot), K(\cdot,\cdot)\right]$ have normalizations $I[n(\cdot)]$ that satisfy $\left\langle I^2 \right\rangle = \int_0^\infty dk_1 \int_0^\infty dk_2\, k_1^2 k_2^2 K(k_1, k_2) = C_0 - C_0 = 0$. So I is fixed at its mean value A with zero uncertainty. In effect, Equation (1.5) shows that $K(\cdot,\cdot)$ is just $C(\cdot,\cdot)$ with the function $C_1(k)$ projected out, so that it has a zero eigenvalue corresponding to this eigenfunction.

In order for this to work, the integrals in Equations (1.2) and (1.3) must exist. This is awkward, because it precludes us from choosing a stationary kernel $C(k_1, k_2) = g(k_1 - k_2)$, for which the integral in Equation (1.3) diverges. This divergence is expected, because a stationary kernel choice in effect asserts that the statistical behavior of the function is the same throughout $\mathbb{R}$, whereas any function possessing a finite normalization integral in Equation (1.1) must necessarily go to zero faster than k^{-1} as $k \to \infty$.

Fortunately, we can conveniently taper covariance kernels such as $g(k_1 - k_2)$ using a rapidly-decaying function $f(k)$, so that $C(k_1, k_2) = f(k_1) g(k_1 - k_2) f(k_2)$—such a kernel is obviously non-negative definite if $g(k_1 - k_2)$ is non-negative definite. This gives hope that we can locate some kernel models for which the required integrals are doable.

2 A Proof-of-Concept Kernel

Let us begin with the popular squared-exponential kernel

$$g(k_1 - k_2) = \exp\left[-\frac{(k_1 - k_2)^2}{\sigma^2}\right], \tag{2.1}$$

and taper it using Gaussians, i.e.

$$f(k) = \exp\left[-\frac{k^2}{\gamma^2}\right], \tag{2.2}$$

so that our covariance model will be

$$\begin{aligned} C(k_1, k_2) &= f(k_1) g(k_1 - k_2) f(k_2) \\ &= \exp\left[-k^T P k\right], \end{aligned} \tag{2.3}$$

where we define $k^T \equiv \begin{bmatrix} k_1 & k_2 \end{bmatrix}$ and

$$P \equiv \begin{bmatrix} a & b \\ b & a \end{bmatrix}, \tag{2.4}$$

with

$$a \equiv \sigma^{-2} + \gamma^{-2} \tag{2.5}$$

$$b \equiv -\sigma^{-2}. \tag{2.6}$$

The required integral $C_1(k_1)$ is given by

$$\begin{aligned} C_1(k_1) &= \int_0^\infty dk_2\, k_2^2 \exp\left[-k^T P k\right] \\ &= \int_0^\infty dk_2\, k_2^2 \exp\left[-\left(a k_1^2 + 2 b k_1 k_2 + a k_2^2\right)\right] \\ &= \int_0^\infty dk_2\, k_2^2 \exp\left[-a\left(k_2 + b k_1/a\right)^2 - a\left(1 - b^2/a^2\right) k_1^2\right] \\ &= \exp\left[-a\left(1 - b^2/a^2\right) k_1^2\right] a^{-3/2} \int_{a^{-1/2} b k_1}^\infty ds\, \left(s - a^{-1/2} b k_1\right)^2 \exp\left[-s^2\right] \\ &= \exp\left[-a\left(1 - b^2/a^2\right) k_1^2\right] a^{-3/2} \int_{a^{-1/2} b k_1}^\infty ds\, \left(s^2 - 2a^{-1/2} b k_1 s + a^{-1} b^2 k_1^2\right) \exp\left[-s^2\right]. \end{aligned} \tag{2.7}$$

We may use

$$
\begin{aligned}
\int_l^\infty ds\, s^2 \exp\left[-s^2\right] &= \left\{\left(-\frac{d}{d\lambda}\right)\int_l^\infty ds\, \exp\left[-\lambda s^2\right]\right\}_{\lambda=1} \\
&= \left\{\left(-\frac{d}{d\lambda}\right)\lambda^{-1/2}\int_{\lambda^{1/2}l}^\infty du\, \exp\left[-u^2\right]\right\}_{\lambda=1} \\
&= \left\{\frac{1}{2}\lambda^{-3/2}\int_{\lambda^{1/2}l}^\infty du\, \exp\left[-u^2\right] + \frac{1}{2}\lambda^{-1} l \exp\left[-\lambda l^2\right]\right\}_{\lambda=1} \\
&= \frac{\pi^{1/2}}{4}\mathrm{erfc}\,[l] + \frac{l}{2}\exp\left[-l^2\right],
\end{aligned} \tag{2.8}
$$

$$
\begin{aligned}
\int_l^\infty ds\, s \exp\left[-s^2\right] &= \frac{1}{2}\int_{l^2}^\infty du\, \exp\left[-u\right] \\
&= \frac{1}{2}\exp\left[-l^2\right],
\end{aligned} \tag{2.9}
$$

and

$$
\int_l^\infty ds\, \exp\left[-s^2\right] = \frac{\pi^{1/2}}{2}\mathrm{erfc}\,[\mathrm{l}]\,. \tag{2.10}
$$

The result is

$$
\begin{aligned}
C_1(k_1) &= \exp\left[-a\left(1-b^2/a^2\right)k_1^2\right] a^{-3/2} \\
&\quad \times \Bigg\{\frac{\pi^{1/2}}{4}\mathrm{erfc}\left[a^{-1/2}bk_1\right] + \frac{a^{-1/2}bk_1}{2}\exp\left[-b^2k_1^2/a\right] \\
&\quad -2a^{-1/2}bk_1 \times \frac{1}{2}\exp\left[-b^2k_1^2/a\right] + a^{-1}b^2k_1^2\frac{\pi^{1/2}}{2}\mathrm{erfc}\left[a^{-1/2}bk_1\right]\Bigg\} \\
&= \frac{\pi^{1/2}}{4}a^{-3/2}\left(1+2\frac{b^2k_1^2}{a}\right)\exp\left[-a\left(1-b^2/a^2\right)k_1^2\right]\mathrm{erfc}\left[a^{-1/2}bk_1\right] \\
&\quad -\frac{bk_1}{2a^2}\times\exp\left[-ak_1^2\right].
\end{aligned} \tag{2.11}
$$

To get an expression for C_0, one could integrate this expression with respect to k_1, but the algebra would be very tedious, and fortunately there's a (slightly) better way. Define

$$
p(k_1, k_2) \equiv k^T P k, \tag{2.12}
$$

and the spheroidal coordinates

$$
R = p(k_1, k_2)^{1/2} \tag{2.13}
$$

$$
t = k_2/k_1. \tag{2.14}
$$

The inverse transformation is easily obtained by $R = p(k_1, tk_1)^{1/2} = k_1 p(1,t)$, whence

$$
k_1 = Rp(1,t)^{-1/2}. \tag{2.15}
$$

Similarly, $R = p(t^{-1}k_2, k_2)^{1/2} = k_2 p(t^{-1}, 1)^{1/2} = k_2 t^{-1} p(1,t)^{1/2}$, from which

$$
k_2 = Rtp(1,t)^{-1/2}. \tag{2.16}
$$

The determinant of the Jacobian of the transformation is

$$
\begin{aligned}
\det J &= \det \begin{bmatrix} p(1,t)^{-1/2} & tp(1,t)^{-1/2} \\ -Rp(1,t)^{-3/2}(at+b) & R\left(p(1,t)^{-1/2} - p(1,t)^{-3/2}t(at+b)\right) \end{bmatrix} \\
&= Rp(1,t)^{-2}\det\begin{bmatrix} 1 & t \\ -(at+b) & (bt+a) \end{bmatrix} \\
&= Rp(1,t)^{-1}.
\end{aligned} \tag{2.17}
$$

We therefore have

$$\begin{aligned} C_0 &= \int_0^\infty dk_1 \int_0^\infty dk_2\, k_1^2 k_2^2 \exp\left[-p(k_1,k_2)\right] \\ &= \int_0^\infty R^5 dR\, \exp\left[-R^2\right] \times \int_0^\infty dt\, \frac{t^2}{\left(at^2+2bt+a\right)^3}. \end{aligned} \tag{2.18}$$

The first integral is

$$\begin{aligned} \int_0^\infty R^5 dR\, \exp\left[-R^2\right] &= \frac{1}{2}\int_0^\infty ds\, s^2 \exp\left[-s\right] \\ &= 1. \end{aligned} \tag{2.19}$$

So we're left with

$$\begin{aligned} C_0 &= \int_0^\infty dt\, \frac{t^2}{\left(at^2+2bt+a\right)^3} \\ &= \int_0^\infty dt\, \frac{t^2}{\left((a^{1/2}t + a^{-1/2}b)^2 + a^{-1}(a^2-b^2)\right)^3}. \end{aligned} \tag{2.20}$$

Now substitute w for t, where

$$a^{-1/2}(a^2-b^2)^{1/2} w = a^{1/2}t + a^{-1/2}b$$

so that

$$\begin{aligned} w &= \left(a^2-b^2\right)^{-1/2}(at+b) \\ t &= a^{-1}\left(a^2-b^2\right)^{1/2}\left[w - \left(a^2-b^2\right)^{-1/2} b\right]. \end{aligned}$$

We then have

$$\begin{aligned} C_0 &= \left(a^2-b^2\right)^{-3/2} \int_{b\left(a^2-b^2\right)^{-1/2}}^\infty dw\, \frac{\left[w - \left(a^2-b^2\right)^{-1/2} b\right]^2}{\left(w^2+1\right)^3} \\ &= \left(a^2-b^2\right)^{-3/2} I_2 - 2\left(a^2-b^2\right)^{-2} bI_1 + \left(a^2-b^2\right)^{-5/2} b^2 I_0, \end{aligned} \tag{2.21}$$

where

$$I_n \equiv \int_{b\left(a^2-b^2\right)^{-1/2}}^\infty dw\, \frac{w^n}{\left(w^2+1\right)^3}. \tag{2.22}$$

With the substitution $w = \tan\psi$, and the fact that $\arctan\left[b\left(a^2-b^2\right)^{-1/2}\right] = \arcsin(b/a) = \arccos\left[\left(1-b^2/a^2\right)^{1/2}\right]$ we have

$$\begin{aligned} \frac{w^2 dw}{\left(w^2+1\right)^3} &= \sin^2\psi \cos^2\psi\, d\psi \\ &= \frac{1}{4}\sin^2 2\psi\, d\psi \\ &= \frac{1}{8}\left(1-\cos 4\psi\right) d\psi \\ &= \frac{1}{8} d\left[\psi - \frac{1}{4}\sin 4\psi\right] \\ &= \frac{1}{8} d\left[\psi - \frac{1}{2}\sin 2\psi \cos 2\psi\right] \\ &= \frac{1}{8} d\left[\psi - \sin\psi\cos\psi\left(1 - 2\sin^2\psi\right)\right] \end{aligned}$$

so that

$$\begin{aligned} I_2 &= \frac{1}{8}\int_{\psi=\arctan\left[b\left(a^2-b^2\right)^{-1/2}\right]}^{\psi=\pi/2} d\left[\psi - \sin\psi\cos\psi\left(1-2\sin^2\psi\right)\right] \\ &= \frac{1}{8}\left\{\pi/2 - \arcsin\left(b/a\right) + \left(b/a\right)\left(1-b^2/a^2\right)^{1/2}\left(1-2b^2/a^2\right)\right\} \\ &= \frac{1}{8}\left\{\pi/2 - \arcsin\left(b/a\right) + a^{-4}b\left(a^2-b^2\right)^{1/2}\left(a^2-2b^2\right)\right\}. \end{aligned} \tag{2.23}$$

Similarly,

$$\begin{aligned} \frac{w\,dw}{\left(w^2+1\right)^3} &= \sin\psi\cos^3\psi\,d\psi \\ &= \frac{1}{4}\sin 2\psi\left(1+\cos 2\psi\right)\,d\psi \\ &= \left(\frac{1}{4}\sin 2\psi + \frac{1}{8}\sin 4\psi\right)\,d\psi \\ &= d\left[-\frac{1}{8}\cos 2\psi - \frac{1}{32}\cos 4\psi\right] \\ &= -d\left[\frac{1}{8}\left(1-2\sin^2\psi\right) + \frac{1}{32}\left(1-2\sin^2 2\psi\right)\right] \\ &= -d\left[\frac{1}{8}\left(1-2\sin^2\psi\right) + \frac{1}{32}\left(1-8\sin^2\psi\cos^2\psi\right)\right], \end{aligned}$$

so that

$$\begin{aligned} I_1 &= \frac{1}{32}\int_{\psi=\arctan\left[b\left(a^2-b^2\right)^{-1/2}\right]}^{\psi=\pi/2} d\left[-4\left(1-2\sin^2\psi\right) - \left(1-8\sin^2\psi\cos^2\psi\right)\right] \\ &= \frac{1}{32}\left\{3+4\left(1-2b^2/a^2\right)+\left(1-8\left(b^2/a^2\right)\left(1-b^2/a^2\right)\right)\right\} \\ &= \frac{1}{32}\left\{8-16b^2/a^2+8b^4/a^4\right\} \\ &= \frac{1}{4a^4}\left\{a^4-2a^2b^2+b^4\right\} \\ &= \frac{\left(a^2-b^2\right)^2}{4a^4}. \end{aligned} \tag{2.24}$$

Finally,

$$
\begin{aligned}
\frac{dw}{\left(w^2+1\right)^3} &= \cos^4\psi\, d\psi \\
&= \frac{1}{4}\left(1+\cos 2\psi\right)^2 d\psi \\
&= \frac{1}{4}\left[1+2\cos 2\psi+\frac{1}{2}\left(1+\cos 4\psi\right)\right] d\psi \\
&= \frac{1}{4}d\left[\frac{3}{2}\psi+\sin 2\psi+\frac{1}{8}\sin 4\psi\right] \\
&= \frac{1}{4}d\left[\frac{3}{2}\psi+2\sin\psi\cos\psi+\frac{1}{4}\sin 2\psi\cos 2\psi\right] \\
&= \frac{1}{4}d\left[\frac{3}{2}\psi+2\sin\psi\cos\psi+\frac{1}{2}\sin\psi\cos\psi\left(1-2\sin^2\psi\right)\right] \\
&= \frac{1}{4}d\left[\frac{3}{2}\psi+\frac{5}{2}\sin\psi\cos\psi-\sin^3\psi\cos\psi\right],
\end{aligned}
$$

and we find

$$
\begin{aligned}
I_0 &= \frac{1}{4}\int_{\psi=\arctan\left[b\left(a^2-b^2\right)^{-1/2}\right]}^{\psi=\pi/2} d\left[\frac{3}{2}\psi+\frac{5}{2}\sin\psi\cos\psi-\sin^3\psi\cos\psi\right] \\
&= \frac{1}{4}\left\{\frac{3\pi}{4}-\frac{3}{2}\arcsin\left(b/a\right)-\frac{5}{2}\left(b/a\right)\left(1-b^2/a^2\right)^{1/2}+\left(b/a\right)^3\left(1-b^2/a^2\right)^{1/2}\right\} \\
&= \frac{1}{8}\left\{\frac{3\pi}{2}-3\arcsin\left(b/a\right)-\left(b/a\right)\left(1-b^2/a^2\right)^{1/2}\left(5-2b^2/a^2\right)\right\} \\
&= \frac{1}{8}\left\{\frac{3\pi}{2}-3\arcsin\left(b/a\right)-a^{-4}b\left(a^2-b^2\right)^{1/2}\left(5a^2-2b^2\right)\right\}.
\end{aligned}
\tag{2.25}
$$

Final assembly:

$$
\begin{aligned}
C_0 &= \left(a^2-b^2\right)^{-3/2}\times\frac{1}{8}\left\{\pi/2-\arcsin\left(b/a\right)+a^{-4}b\left(a^2-b^2\right)^{1/2}\left(a^2-2b^2\right)\right\} \\
&\quad -2\left(a^2-b^2\right)^{-2}b\times\frac{\left(a^2-b^2\right)^2}{4a^4} \\
&\quad +\left(a^2-b^2\right)^{-5/2}b^2\times\frac{1}{8}\left\{\frac{3\pi}{2}-3\arcsin\left(b/a\right)-a^{-4}b\left(a^2-b^2\right)^{1/2}\left(5a^2-2b^2\right)\right\} \\
&= \frac{1}{8}\left(a^2-b^2\right)^{-5/2}\left(a^2+2b^2\right)\left[\frac{\pi}{2}-\arcsin\left(b/a\right)\right]-\frac{b}{2a^4} \\
&\quad +\frac{1}{8}\left(a^2-b^2\right)^{-2}a^{-4}b\left[\left(a^2-b^2\right)\left(a^2-2b^2\right)-b^2\left(5a^2-2b^2\right)\right] \\
&= \frac{1}{8}\left(a^2-b^2\right)^{-5/2}\left(a^2+2b^2\right)\left[\frac{\pi}{2}-\arcsin\left(b/a\right)\right]-\frac{b}{2a^4} \\
&\quad +\frac{1}{8}\left(a^2-b^2\right)^{-2}a^{-4}b\left[a^4-8a^2b^2+4b^4\right] \\
&= \frac{1}{8}\left(a^2-b^2\right)^{-5/2}\left(a^2+2b^2\right)\left[\frac{\pi}{2}-\arcsin\left(b/a\right)\right] \\
&\quad +\frac{1}{8}\left(a^2-b^2\right)^{-2}a^{-4}b\left[a^4-8a^2b^2+4b^2-4\left(a^2-b^2\right)^2\right] \\
&= \frac{1}{8}\left(a^2-b^2\right)^{-5/2}\left(a^2+2b^2\right)\left[\frac{\pi}{2}-\arcsin\left(b/a\right)\right]+\frac{1}{8}\left(a^2-b^2\right)^{-2}a^{-4}b\left[-3a^4\right] \\
&= \frac{1}{8}\left(a^2-b^2\right)^{-5/2}\left(a^2+2b^2\right)\left[\frac{\pi}{2}-\arcsin\left(b/a\right)\right]-\frac{3}{8}\left(a^2-b^2\right)^{-2}b.
\end{aligned}
\tag{2.26}
$$