December 10, 2018

Spectral Mixture Kernel: for two points x, x' let $\tau = x - x'$. Then the SM kernel for a 1-d problem is defined by

$$k(x, x') = k(\tau) = \sum_{q=1}^{Q} k_q(\tau) = \sum_{q=1}^{Q} w_q \exp(-2\pi^2\tau^2\nu_q)\cos(2\pi\tau\mu_q)$$

where the spectral density of the kernel is a mixture of Q Gaussians, with means μ_q and variances ν_q. Note that a derivative of this kernel with respect to x is identical to its derivative with respect to τ, and that the same can be said of a derivative with respect to x' if multiplied by -1 to the power of the order of the derivative, i.e.

$$\frac{d^n}{dx^n}k(x, x') = \frac{d^n}{d\tau^n}k(\tau) \text{ and } \frac{d^n}{dx'^n}k(x, x') = (-1)^n \frac{d^n}{d\tau^n}k(\tau)$$

So, we express the general nth derivative of k with respect to τ. Here we consider each term separately, since differentiation is linear:

$$k_q^{(n)}(\tau) = e^{a\tau^2}\big(\cos(b\tau)P_c^n(\tau) + \sin(b\tau)P_s^n(\tau)\big)$$

where $P_c^n(\tau)$ and $P_s^n(\tau)$ are polynomial functions of τ. Note that here we have made the substitutions $a = -2\pi^2\nu_q$ and $b = 2\pi\mu_q$. From the above structure we can see by the product rule how the polynomial functions are related from derivative to derivative:

$$P_c^{n+1} = 2a\tau P_c^n + \frac{d}{d\tau}P_c^n + bP_s^n, \quad P_s^{n+1} = 2a\tau P_s^n + \frac{d}{d\tau}P_s^n - bP_c^n$$

So we compute the first eight such polynomials in order to analytically represent up to eight derivatives of k:

$$P_c^0(\tau) = 1$$
$$P_s^0(\tau) = 0$$

$$P_c^1(\tau) = 2a\tau$$
$$P_s^1(\tau) = -b$$

$$P_c^2(\tau) = 4a^2\tau^2 + (2a - b^2)$$
$$P_s^2(\tau) = -4ab\tau$$

$$P_c^3(\tau) = 8a^3\tau^3 + (12a^2 - 6ab^2)\tau$$
$$P_s^3(\tau) = -12a^2b\tau^2 + (-6ab + b^3)$$

$$P_c^4(\tau) = 16a^4\tau^4 + (48a^3 - 24a^2b^2)\tau^2 + (12a^2 - 12ab^2 + b^4)$$
$$P_s^4(\tau) = -32a^3b\tau^3 + (-48a^2b + 8ab^3)\tau$$

$$P_c^5(\tau) = 32a^5\tau^5 + (160a^4 - 80a^3b^2)\tau^3 + (120a^3 - 120a^2b^2 + 10ab^4)\tau$$
$$P_s^5(\tau) = -80a^4b\tau^4 + (-240a^3b + 40a^2b^3)\tau^2 + (-60a^2b + 20ab^3 - b^5)$$

$$\begin{aligned} P_c^6(\tau) = {} & 64a^6\tau^6 + (480a^5 - 240a^4b^2)\tau^4 + (720a^4 - 720a^3b^2 + 60a^2b^4)\tau^2 + \cdots \\ & + (120a^3 - 180a^2b^2 + 30ab^4 - b^6) \end{aligned}$$
$$P_s^6(\tau) = -192a^5b\tau^5 + (-960a^4b + 160a^3b^3)\tau^3 + (-960a^3b + 240a^2b^3 - 12ab^5)\tau$$

$$\begin{aligned} P_c^7(\tau) = {} & 128a^7\tau^7 + (1344a^6 - 672a^5b^2)\tau^5 + (3360a^5 - 3360a^4b^2 + 280a^3b^4)\tau^3 + \cdots \\ & + (1680a^4 - 2760a^3b^2 + 420a^2b^4 - 14ab^6)\tau \end{aligned}$$
$$\begin{aligned} P_s^7(\tau) = {} & -488a^6b\tau^6 + (-3360a^5b + 560a^4b^3)\tau^4 + (-5520a^4b + 1680a^3b^3 - 84a^2b^5)\tau^2 + \cdots \\ & + (-1080a^3b + 420a^2b^3 - 42ab^5 + b^7) \end{aligned}$$

$$\begin{aligned} P_c^8(\tau) = {} & 256a^8\tau^8 + (3584a^7 - 1832a^6b^2)\tau^6 + (13440a^6 - 13440a^5b^2 + 1120a^4b^4)\tau^4 + \cdots \\ & + (13440a^5 - 21120a^4b^2 + 3360a^3b^4 - 112a^2b^6)\tau^2 + \cdots \\ & + (1680a^4 - 3840a^3b^2 + 840a^2b^4 - 56ab^6 + b^8) \end{aligned}$$
$$\begin{aligned} P_s^8(\tau) = {} & -1104a^7b\tau^7 + (-10992a^6b + 1792a^5b^3)\tau^5 + (-27840a^5b + 8960a^4b^3 - 448a^3b^5)\tau^3 + \cdots \\ & + (-14880a^4b + 6960a^3b^3 - 672a^2b^5 + 16ab^7)\tau \end{aligned}$$

In addition, we compute the form of the PDE structured kernel for several example problems. First, to solve the Kuramoto-Sivashinsky equation, we use the following kernel structure: Note that the governing equation in this case is

$$u_t + \lambda_1 u u_x + \lambda_2 u_{xx} + \lambda_3 u_{xxxx} = 0$$

After applying the backwards Euler formula we obtain

$$u_{n-1} = (I - \Delta t \mathcal{N}_x) u_n, \quad \mathcal{N}_x u = -\lambda_1 u u_x - \lambda_2 u_{xx} - \lambda_3 u_{xxxx}$$

This operator $\mathcal{N}_x$ is linearized for our purposes as

$$\mathcal{L}_x u_n = -\lambda_1 u_{n-1} \frac{d}{dx} u_n - \lambda_2 \frac{d^2}{dx^2} u_n - \lambda_3 \frac{d^4}{dx^4} u_n$$

After placing a GP prior $u_n \sim \mathcal{GP}(0, k(x, x'))$ we obtain the joint GP prior

$$\begin{bmatrix} u_n \\ u_{n-1} \end{bmatrix} = \mathcal{GP}\left(0, \begin{bmatrix} k_{1,1} & k_{1,2} \\ k_{2,1} & k_{2,2} \end{bmatrix}\right)$$

where, considering $k^{(n)}$ to be the nth derivative of k with respect to $\tau = x - x'$,

$$k_{1,1} = k$$

$$\begin{aligned} k_{1,2} &= (I - \mathcal{L}_{x'})k \\ &= k - \Delta t \lambda_1 u'_{n-1} k^{(1)} + \Delta t \lambda_2 k^{(2)} + \Delta t \lambda_3 k^{(4)} \end{aligned}$$

$$\begin{aligned} k_{2,2} &= (I - \Delta t \mathcal{L}_x)(I - \Delta t \mathcal{L}_{x'})k \\ &= k + \Delta t \lambda_1 (u_{n-1} - u'_{n-1}) k^{(1)} + (2\Delta t \lambda_2 - (\Delta t \lambda_1)^2 u_{n-1} u'_{n-1}) k^{(2)} + (\Delta t^2 \lambda_1 \lambda_2 (u_{n-1} - u'_{n-1})) k^{(3)} \\ &\quad + (2\Delta t \lambda_3 + \Delta t^2 \lambda_2^2) k^{(4)} + (\Delta t^2 \lambda_1 \lambda_3 (u_{n-1} - u'_{n-1})) k^{(5)} + 2\Delta t^2 \lambda_2 \lambda_3 k^{(6)} + \Delta t^2 \lambda_3^2 k^{(8)} \end{aligned}$$

where $u_{n-1} = u(x, t_{n-1})$, and $u'_{n-1} = u(x', t_{n-1})$. In order to optimize the coefficients λ_1 and λ_2 as hyperparameters of the kernel, we differentiate:

$$
\begin{aligned}
\frac{\partial k_{1,1}}{\partial \lambda_1} &= \frac{\partial k_{1,1}}{\partial \lambda_2} = \frac{\partial k_{1,1}}{\partial \lambda_3} = 0 \\
\frac{\partial k_{1,2}}{\partial \lambda_1} &= -\Delta t u'_{n-1} k^{(1)} \\
\frac{\partial k_{1,2}}{\partial \lambda_2} &= \Delta t k^{(2)} \\
\frac{\partial k_{1,2}}{\partial \lambda_3} &= \Delta t k^{(4)} \\
\frac{\partial k_{2,2}}{\partial \lambda_1} &= \Delta t (u_{n-1} - u'_{n-1}) k^{(1)} - 2\Delta t^2 \lambda_1 u_{n-1} u'_{n-1} k^{(2)} + \Delta t^2 \lambda_2 (u_{n-1} - u'_{n-1}) k^{(3)} \cdots \\
&\quad + \Delta t^2 \lambda_3 (u_{n-1} - u'_{n-1}) k^{(5)} \\
\frac{\partial k_{2,2}}{\partial \lambda_2} &= 2\Delta t k^{(2)} + \Delta t^2 \lambda_1 (u_{n-1} - u'_{n-1}) k^{(3)} + 2\Delta t^2 \lambda_2 k^{(4)} + 2\Delta t^2 \lambda_3 k^{(6)} \\
\frac{\partial k_{2,2}}{\partial \lambda_3} &= 2\Delta t k^{(4)} + \Delta t^2 \lambda_1 (u_{n-1} - u'_{n-1}) k^{(5)} + 2\Delta t^2 \lambda_2 k^{(6)} + 2\Delta t^2 \lambda_3 k^{(8)}
\end{aligned}
$$

Second, for the Kortweg-de Vries equation,

$$u_t + \lambda_1 u u_x + \lambda_2 u_{xxx} = 0$$

via the same process as above we obtain the approximative linear operator

$$\mathcal{L}_x u_n = -\lambda_1 u_{n-1} \frac{d}{dx} u_n - \lambda_2 \frac{d^3}{dx^3} u_n$$

and so obtain the kernel structure

$$k_{1,1} = k$$

$$
\begin{aligned}
k_{1,2} &= (I - \mathcal{L}_{x'})k \\
&= k - \Delta t \lambda_1 u'_{n-1} k^{(1)} + \Delta t \lambda_2 k^{(3)}
\end{aligned}
$$

$$
\begin{aligned}
k_{2,2} &= (I - \Delta t \mathcal{L}_x)(I - \Delta t \mathcal{L}_{x'})k \\
&= k + \Delta t \lambda_1 (u_{n-1} - u'_{n-1}) k^{(1)} - \Delta t^2 \lambda_1^2 u'_{n-1} u_{n-1} k^{(2)} - \Delta t^2 \lambda_1 \lambda_2 (u_{n-1} + u'_{n-1}) k^{(4)} - \Delta t^2 \lambda_2^2 k^{(6)}
\end{aligned}
$$

where again $u_{n-1} = u(x, t_{n-1})$, and $u'_{n-1} = u(x', t_{n-1})$. Again, in order to optimize the coefficients λ_1 and λ_2 as hyperparameters of the kernel, we differentiate:

$$\frac{\partial k_{1,1}}{\partial \lambda_1} = \frac{\partial k_{1,1}}{\partial \lambda_2} = 0$$
$$\frac{\partial k_{1,2}}{\partial \lambda_1} = -\Delta t u'_{n-1} k^{(1)}$$
$$\frac{\partial k_{1,2}}{\partial \lambda_2} = \Delta t k^{(3)}$$
$$\frac{\partial k_{2,2}}{\partial \lambda_1} = \Delta t(u_{n-1} - u'_{n-1})k^{(1)} - 2\Delta t^2 \lambda_1 u'_{n-1} u_{n-1} k^{(2)} - \Delta t^2 \lambda_2 (u_{n-1} + u'_{n-1})k^{(4)}$$
$$\frac{\partial k_{2,2}}{\partial \lambda_2} = -\Delta t^2 \lambda_1 (u_{n-1} + u'_{n-1})k^{(4)} - 2\Delta t^2 \lambda_2 k^{(6)}$$