Proceedings of Machine Learning Research 71:Supplement, 2017 KDD 2017: Workshop on Anomaly Detection in Finance

Appendix for "Binned Kernels for Anomaly Detection in Multi-timescale Data using Gaussian Processes"

Matthew van Adelsberg MATTHEW.VANADELSBERG@CAPITALONE.COM
Christian Schwantes CHRISTIAN.SCHWANTES@CAPITALONE.COM
1680 Capital One Drive McLean, VA 22102

Appendix A. Binned Kernel Spectral Identity

We can define a sequence of kernels through the Fourier transform:

$$\hat{k}_n(f, f') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\pi^2(t+t')^2/4n} k(t, t') e^{-i2\pi ft} e^{-i2\pi f't'} dt dt', \tag{1}$$

where $k(t, t')$ is an arbitrary stationary kernel of interest. The reason for defining this sequence with the $e^{-\pi^2(t+t')^2/4n}$ term is in anticipation of a change of variables to $u \equiv t - t'$, $v \equiv t + t'$, which results in:

$$\hat{k}_n(f, f') = \frac{1}{2} \int_{-\infty}^{\infty} e^{-\pi^2 v^2/4n} e^{-i2\pi v(f+f')/2} dv \int_{-\infty}^{\infty} k(u) e^{-i2\pi u(f-f')/2} du. \tag{2}$$

Without the squared-exponential term, the Fourier transform in the first integral only exists as a distribution, rather than a function. We can explicitly evaluate the Fourier transform of the Gaussian function in the first integral:

$$\hat{k}_n(f, f') = \sqrt{\frac{n}{\pi}} e^{-n(f+f')^2} \int_{-\infty}^{\infty} k(u) e^{-i2\pi u(f-f')/2} du. \tag{3}$$

Now consider the binned kernel:

$$\begin{aligned} \bar{k}\left(t_\Delta, t'_{\Delta'}\right) = \frac{1}{\Delta\Delta'} \int_{t-\Delta/2}^{t+\Delta/2} \int_{t'-\Delta'/2}^{t'+\Delta'/2} k(\xi, \xi') d\xi d\xi' = \\ \frac{1}{\Delta\Delta'} \int_{-\Delta/2}^{\Delta/2} \int_{-\Delta'/2}^{\Delta'/2} k(\xi + t, \xi' + t') d\xi d\xi' = \\ \frac{1}{\Delta\Delta'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{rect}(\xi, \xi' | \Delta, \Delta') k(\xi + t, \xi' + t') d\xi d\xi', \end{aligned} \tag{4}$$

where 'rect' denotes the two-dimensional rectangle function. Thus, $\bar{k}(t_\Delta, t'_{\Delta'}) = (\mathrm{rect} \star k)(t, t' | \Delta, \Delta')$, where $\star$ is the cross-correlation operator. Using the cross-correlation theorem for Fourier transforms (e.g., Butz, 2015):

$$\hat{\bar{k}}(f, f') = \hat{\text{rect}}(f, f'|\Delta, \Delta')\hat{k}(f, f') = \text{sinc}(\Delta f)\text{sinc}(\Delta' f')\hat{k}(f, f'). \tag{5}$$

If we substitute Equation (3) into (5) and take the inverse Fourier transform, we obtain:

$$\bar{k}_n(t_\Delta, t'_{\Delta'}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sqrt{\frac{n}{\pi}} e^{-n(f+f')^2} \text{sinc}(\Delta f)\text{sinc}(\Delta' f') \times \left[\int_{-\infty}^{\infty} k(u) e^{-i2\pi u(f-f')/2} du\right] e^{i2\pi f t} e^{i2\pi f' t'} df' df. \tag{6}$$

We note that in Equation (1), taking the limit $n \to \infty$ recovers the expression for the Fourier transform of the original kernel of interest. We can therefore compute the inverse Fourier transform for the binned kernel as:

$$\bar{k}(t_\Delta, t'_{\Delta'}) \equiv \lim_{n\to\infty} \bar{k}_n(t_\Delta, t'_{\Delta'}) = \int_{-\infty}^{\infty} \lim_{n\to\infty} \int_{-\infty}^{\infty} \sqrt{\frac{n}{\pi}} e^{-n(f+f')^2} \text{sinc}(\Delta f)\text{sinc}(\Delta' f') \times \left[\int_{-\infty}^{\infty} k(u) e^{-i2\pi u(f-f')/2} du\right] e^{i2\pi f t} e^{i2\pi f' t'} df' df = \int_{-\infty}^{\infty} \text{sinc}(\Delta f)\text{sinc}(\Delta' f) \left[\int_{-\infty}^{\infty} k(u) e^{-i2\pi u f} du\right] e^{i2\pi f(t-t')} df, \tag{7}$$

where we have used the fact that the result of the inner integral over f' satisfies the Lebesgue dominated convergence theorem and the functions $\sqrt{n/\pi}\exp(-n[f+f']^2)$ form a Delta sequence (e.g., Arfken and Weber, 2005).

Thus, the Fourier transform of the binned kernel becomes:

$$\hat{\bar{k}}(f) = \text{sinc}(\Delta f)\text{sinc}(\Delta' f)\hat{k}(f). \tag{8}$$

Appendix B. Fourier Transform of the Locally Periodic Kernel

The locally periodic kernel is defined as:

$$k(u) = \eta^2 e^{-\sin^2(\pi\nu u)/w^2} e^{-u^2/2\ell^2}, \tag{9}$$

where $u \equiv t - t'$ and η, ν, w, and ℓ are hyperparameters. We begin by Taylor expanding the first exponential term:

$$k(u) = \eta^2 \left[\sum_{n=0}^{\infty} \frac{1}{n!} \left(-\sin^2(\pi\nu u)/w^2\right)^n\right] e^{-u^2/2\ell^2}. \tag{10}$$

Expanding the $\sin^{2n}(\pi\nu u)$ terms and doing some algebra results in:

$$k(u) = \sum_{n=0}^{\infty} \frac{1}{n!}(-1)^n \left(\frac{1}{w^2}\right)^n \frac{1}{2^{2n}} \left[\binom{2n}{n} + \sum_{k=0}^{n-1}(-1)^{n-k}\binom{2n}{k}\cos\left(2(n-k)\pi\nu u\right)\right] \times e^{-u^2/2\ell^2}. \quad (11)$$

If we reindex with $m = n - k$, some algebra yields:

$$k(u) = \left[\sum_{n=0}^{\infty} \frac{1}{n!2^{2n}}\binom{2n}{n}\left(-\frac{1}{w^2}\right)^n + \sum_{n=0}^{\infty} \frac{2}{n!2^{2n}}\left(-\frac{1}{w^2}\right)^n \sum_{m=1}^{n}(-1)^m\binom{2n}{n-m}\cos(2m\pi\nu u)\right] e^{-u^2/2\ell^2} \quad (12)$$

For a given m, we can evaluate the sum over n in terms of a modified Bessel function of the first kind (e.g., Abramowitz and Stegun, 1972):

$$\sum_{n=m}^{\infty} \frac{1}{n!2^{2n}}\binom{2n}{n-m}\left(-\frac{1}{w^2}\right)^n = e^{-1/2w^2} I_m(-1/2w^2). \quad (13)$$

Thus, the locally periodic kernel can be written:

$$k(u) = \left[a_0/2 + \sum_{n=1}^{\infty} a_n \cos(2\pi n\nu u)\right] e^{-u^2/2\ell^2}, \quad (14)$$

where $a_n \equiv 2e^{-1/2w^2} I_n(1/2w^2)$.

References

M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover Publications, Inc., 1972.

G. Arfken and H. Weber. *Mathematical Methods for Physicists*. Elsevier Academic Press Publications, 2005.

T. Butz. *Fourier Transformation for Pedestrians*. Springer, 2015.