
Deterministic Fourier Features

Suppose we have time-series data, with a PSD stationary kernel k . Consider the following process:

- (i) Evaluate the kernel from lags $-T$ to T , and construct vector

$$\xi := (k(0), k(T), \dots, k(1), k(1), \dots, k(T))$$

For convenience, we define $\omega := \frac{2\pi}{N}$ with $N := 2T + 1$.

- (ii) Take the DFT of ξ :

$$S_j := \sum_{t=0}^{N-1} \xi_t \exp(-i\omega j t)$$

for $j = 0, \dots, N - 1$. Since the true spectral density of k is nonnegative (and the DFT is an approximation), take T large enough so that all $S_j \geq 0$. If not possible, clip small negative values to zero.

- (iii) Let \mathcal{I} be the set of all indices $j \in \{0, \dots, N - 1\}$ such that S_j is among the $2r$ largest values (excluding S_0). This forms the sequence

$$\tilde{S}_j = \begin{cases} S_0, & j = 0 \\ S_j, & j \in \mathcal{I} \\ 0, & \text{o.w.} \end{cases}$$

Note that by symmetry, $S_j = S_{N-j}$, for $1 \leq j \leq N - 1$, so that we always take pairs, ensuring the same symmetry in \tilde{S} .

- (iv) Take the inverse DFT of \tilde{S} :

$$\begin{aligned} \tilde{\xi}_t &:= \frac{1}{N} \sum_{j=0}^{N-1} \tilde{S}_j \exp(i\omega j t) \\ &= \frac{1}{N} \left[S_0 + 2 \sum_{j \in \mathcal{I}} S_j \cos(\omega j t) \right] \end{aligned}$$

for $t = 0, \dots, N - 1$. Some remarks:

- The sequence is of the form (Lemma 1)

$$\tilde{\xi} = (\tilde{k}(0), \tilde{k}(T), \dots, \tilde{k}(1), \tilde{k}(1), \dots, \tilde{k}(T))$$

- The equality using cosines follows from Lemma 2.

- (v) Define the low-rank kernel (as shown in $\tilde{\xi}$ above)

$$\tilde{k}(t) = \begin{cases} \tilde{\xi}_0, & t = 0 \\ \tilde{\xi}_{T+t}, & t \in \{1, \dots, T\} \\ \tilde{k}(-t), & t \in \{-T, \dots, -1\} \end{cases}$$

Note that this kernel satisfies (Lemma 3)

$$\tilde{k}(t - t') = \langle \phi(t), \phi(t') \rangle \quad (1)$$

for $t - t' \in \{-T, \dots, T\}$, where

$$\phi(t) := \frac{1}{\sqrt{N}} \begin{bmatrix} \sqrt{S_0} \\ \sqrt{2S_j} \cos(\omega j t) \\ \sqrt{2S_j} \sin(\omega j t) \end{bmatrix}_{j \in \mathcal{I}} \in \mathbb{R}^{2r+1}$$

Note that we need $S \geq 0$ to take the square roots.

We can then do GP inference on this approximated kernel, which reduces the computation to $O(r^3)$. Note that if we can compute the DTFT, we can directly pick the top r modes from it, instead of using the DFT.

1 Auxiliary results

Lemma 1

We have that $\tilde{\xi}_t = \tilde{\xi}_{N-t}$ for $t = 1, \dots, T-1$.

Proof. Ignoring the normalization,

$$\begin{aligned} \tilde{\xi}_{N-t} &\propto \sum_{j=0}^{N-1} \tilde{S}_j \exp(i\omega j(N-t)) \\ &= \sum_{j=0}^{N-1} \tilde{S}_j \exp(-i\omega j t) \\ &= \tilde{S}_0 + \sum_{j=1}^{N-1} \tilde{S}_{N-j} \exp(i\omega(N-j)t) \\ &= \sum_{j=0}^{N-1} \tilde{S}_j \exp(i\omega j t) \\ &\propto \tilde{\xi}_t \end{aligned}$$

■

Lemma 2

We have that

$$\tilde{\xi}_t = \frac{1}{N} \left[S_0 + 2 \sum_{j \in \mathcal{I}} S_j \cos(\omega j t) \right]$$

for $t = 0, \dots, N-1$.

Proof. Let us begin by noting that

$$\begin{aligned} \sum_{j=T+1}^{N-1} \tilde{S}_j \exp(i\omega j t) &= \sum_{j=T+1}^{N-1} \tilde{S}_{N-j} \exp(-i\omega(N-j)t) \\ &= \sum_{j=1}^T \tilde{S}_j \exp(-i\omega j t) \end{aligned}$$

Hence,

$$\begin{aligned} \tilde{\xi}_t &= \frac{1}{N} \sum_{j=0}^{N-1} \tilde{S}_j \exp(i\omega j t) \\ &= \frac{1}{N} \left[\tilde{S}_0 + 2 \sum_{j=1}^T \tilde{S}_j \cos(\omega j t) \right] \end{aligned}$$

and the claim follows from the definitions of \tilde{S} and \mathcal{I} . ■

Lemma 3: Proof of Eq. (1)

For $0 \leq t - t' \leq N - 1$, recall that

$$\begin{aligned} \tilde{\xi}_{t-t'} &= \frac{1}{N} \left[S_0 + 2 \sum_{j \in \mathcal{I}} S_j \cos(\omega j (t - t')) \right] \\ &= \frac{1}{N} \left[S_0 + 2 \sum_{j \in \mathcal{I}} S_j [\cos(\omega j t) \cos(\omega j t') + \sin(\omega j t) \sin(\omega j t')] \right] \\ &= \langle \phi(t), \phi(t') \rangle \end{aligned}$$