1-dimension Discrete Fourier Transform (DFT) and circulent covariance matrices

Warm-up: Observing a cos/sin basis on a grid.

Start by considering a function $f:[0,P) \to \mathbb{R}$ with a basis expansion given as follows

$$f(x) = \sum_{k=1}^{\infty} \alpha_k \cos(x\tau_k) + \beta_k \sin(x\tau_k).$$

Lets examine how α_k, β_k and τ_k relate to properties of f and the vector

$$v := (v_1, v_2, \dots, v_N)^T$$

defined by

$$v_n := f\Big(\underbrace{(n-1)\frac{P}{N}}\Big)$$

for $n \in \{1, \dots, N\}$, this span through a uniform grid on $[0,P)$

τ_k characterizes the frequency of $\sin(x w_k)$ and $\cos(x\tau_k)$.

The corresponding wavelength, also called wavenumber, is given by $\frac{2\pi}{\tau_k}$

as x moves from 0 to $\frac{2\pi}{\tau_k}$ $\sin(x\tau_k)$ and $\cos(x\tau_k)$ traverse a single wave

Note if $\forall k\ \exists q \in \mathbb{Z}$ s.t. $P\tau_k = 2\pi q$,

i.e. $\tau_k \in \frac{2\pi}{P}\cdot\mathbb{Z} \quad \forall k$

Then f is periodic with period P.

For now lets assume

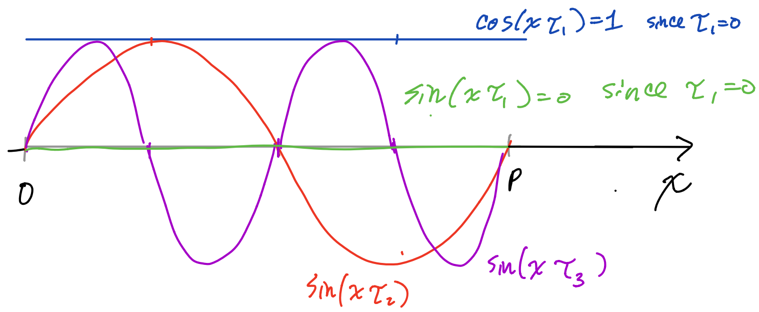
$$\tau_k \in \frac{2\pi}{P}\cdot\mathbb{Z}, \qquad \forall k$$

To make things even simpler lets assume

$$\tau_k = (k-1)\frac{2\pi}{P}$$

These k's span $\{1, 2, 3, \dots\}$

Lets plot some of the first few basis functions



α_1 : mean value of f

β_1 : has no impact on f

$\vdots$

α_k & β_k : quantifies the magnitude of f variation at length scale given by wavelength

$$\frac{2\pi}{\tau_k} = \frac{2\pi}{(k-1)\frac{2\pi}{P}} = \frac{P}{k-1}$$

as $k \to \infty$ the wavelengths get shorter, i.e. higher frequency

Now consider the case that we only observe f on a grid of size N.

Set

$$v = (v_1, \ldots, v_N)^T$$

$$v_n = f(x_n)$$

$$x_n = (n-1)\frac{P}{N}$$

$\{x_1, \ldots, x_N\}$ are N equally spaced grid points in $[0, P)$ with $x_1 = 0$ & spacing $\Delta x = \frac{P}{N}$

The basis expansion of f gives an "over parameterized" basis expansion of v

$$v = \alpha_1 \begin{bmatrix} | \\ \cos(x_n \tau_1) \\ | \end{bmatrix} + \beta_1 \begin{bmatrix} | \\ \sin(x_n \tau_1) \\ | \end{bmatrix}$$

At $k=1$ these two basis vectors are constant ... just like the continuous versions

$$+ \alpha_2 \begin{bmatrix} | \\ \cos(x_n \tau_2) \\ | \end{bmatrix} + \beta_2 \begin{bmatrix} | \\ \sin(x_n \tau_2) \\ | \end{bmatrix} +$$

$$\vdots$$

A $k=2$ these look similar to the continuous basis functions

At $k \approx \frac{N}{2}+1$ (i.e half the grid size) something interesting happens.

To make life easier suppose N is even.

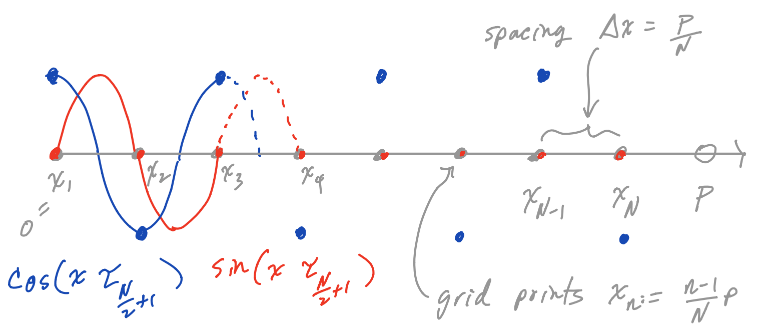
We can simplify the basis function corresponding to $k = \frac{N}{2}+1$ by noticing

$$\text{basis freq} = \tau_{\frac{N}{2}+1} = \left(\frac{N}{2}+1-1\right)\frac{2\pi}{P} = \frac{2\pi}{2\left(\frac{P}{N}\right)} = \frac{2\pi}{2\Delta x}$$

$$\text{basis wavelength} = \frac{2\pi}{\tau_{\frac{N}{2}+1}} = 2\Delta x$$

twice the grid spacing

Here is a picture:



Notice that the sin term is indistinguishable from zero and the cos term is at maximum ossillation.

Def: frequency $= \frac{2\pi}{2\Delta x}$ is called the Nyquist frequency

What happens for terms with $k > \frac{N}{2}+1$?

The magic of cos & sin result in an "aliasing" so that

$\cos(x\tau_k)$ & $\sin(x\tau_k)$

indistinguishable on $\{x_1, \dots, x_n\}$

$\cos(x\tau)$ & $\sin(x\tau)$

for some τ closer to zero.

Example:

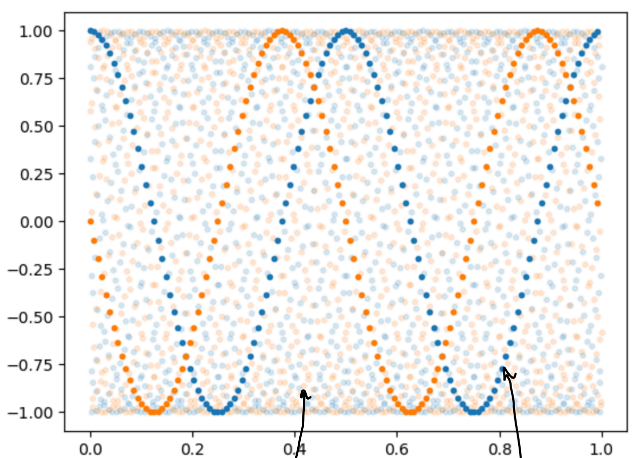
Plot of $\cos(x\tau_k)$ & $\sin(x\tau_k)$ on two grids.

Filled dots show grid

$$x_n = (n-1)\frac{1}{128}$$

Open & transparent dots show grid

$$y_m = (m-1)\frac{1}{1280}$$



$\cos(y_m\tau_k)$ $\cos(x_n\tau_k)$

So the end result is that the infinite expansion...

$$V_n = \sum_{k=1}^{\infty} \alpha_k \cos(x_n\tau_k) + \beta_k \sin(x_n\tau_k)$$

Collapses into a sum with $N/2$ terms

$$V_n = \sum_{k=1}^{N/2+1} \alpha'_k \cos(x_n\tau_k) + \beta'_k \sin(x_n\tau_k)$$

So the degrees of freedom match both LHS & RHS.

higher terms folding into lower freq terms... called aliasing

This is just a warm up to see how the continuous Fourier transform and discrete Fourier transform interact

Now lets move to the discrete Fourier Transform.

Discrete Fourier Transform. (1-d)

Sometimes written DFT or FFT

Definition:

For every positive integer $N \geq 2$ let

$$w_N := exp\left(-i \frac{2\pi}{N}\right)$$

where $i = \sqrt{-1}$

An important feature of w_N is that it is an "n^{th} root of unity"

i.e.
$$\begin{aligned} w_N^{\pm N} &= exp\left(-i \frac{2\pi}{N}(\pm N)\right) \\ &= exp(\pm i 2\pi) \\ &= cos(\pm 2\pi) + i \sin(\pm 2\pi) \\ &= 1 \end{aligned}$$

This will be important soon!

Definition:

The DFT is the $\mathbb{C}^{N\times N}$ matrix defined by

$$W_N[k,n] := w_N^{(k-1)(n-1)}$$

rows indexed by $k \in \{1,\dots,N\}$

columns indexed by $n \in \{1,\dots,N\}$

The DFT transform of $v \in \mathbb{C}^N$ or $v \in \mathbb{R}^N$ is simply

$$z = W_N v$$

Fundamental fact:

adjoint or "transpose & conj"

$$W_N^{-1} = \frac{1}{N} W_N^H$$

cols indexed by k

$$= \frac{1}{N} \begin{bmatrix} \ddots & & \\ & \bar{w}_N^{(n-1)(k-1)} & \\ \vdots & & \end{bmatrix}$$

rows indexed by n

Side Remark: Notice $W_N^T = W_N$ so

$$W_N^H = \bar{W}_N = N W^{-1}.$$

Written out explicitly:

$$W_N = \begin{bmatrix} \ddots & & \\ & exp\left(-i \frac{2\pi}{N}(n-1)(k-1)\right) & \\ \cdots & & \end{bmatrix}$$

$$W_N^{-1} = \frac{1}{N} \begin{bmatrix} \ddots & & \\ & exp\left(i \frac{2\pi}{N}(k-1)(n-1)\right) & \\ \cdots & & \end{bmatrix}$$

Notice that due to $W_N^{-1} = \frac{1}{N} W_N^H$ we have

$$\begin{aligned} I &= W_N W_N^{-1} \\ &= W_N\left(\frac{1}{N} W^H\right) \\ &= \left(\frac{1}{\sqrt{N}} W_N\right)\left(\frac{1}{\sqrt{N}} W\right)^H \end{aligned}$$

Therefore if we set $U_N = \frac{1}{\sqrt{N}} W_N$ then U_N is unitary, i.e.

$$U_N U_N^H = I$$

Written explicitly

$$\left.\begin{aligned} \mathcal{U}_N &= \frac{1}{\sqrt{N}}\begin{bmatrix} \ddots & & \\ & \exp\left(-i\frac{2\pi}{N}(n-1)(k-1)\right) & \\ & & \ddots \end{bmatrix} \\ \mathcal{U}_N^H &= \frac{1}{\sqrt{N}}\begin{bmatrix} \ddots & & \\ & \exp\left(i\frac{2\pi}{N}(n-1)(k-1)\right) & \\ & & \ddots \end{bmatrix}. \end{aligned}\right\} \text{Unitary}$$

Notice that we are free to choose other scalings ...

the following will become important for relating DCT and the continuous Fourier transform when the column indices of W_N & $\mathcal{U}_N$ (i.e. "n") are associated to the grid $x_n = (n-1)\frac{P}{N}$

$$\mathcal{F}_N = \frac{\Delta x}{\sqrt{2\pi}}\begin{bmatrix} \ddots & & \\ & \exp\left(-i\frac{2\pi}{N}(n-1)(k-1)\right) & \\ & & \ddots \end{bmatrix}$$

$$\mathcal{F}_N^{-1} = \frac{\sqrt{2\pi}}{\Delta x\, N}\begin{bmatrix} \ddots & & \\ & \exp\left(i\frac{2\pi}{N}(n-1)(k-1)\right) & \\ & & \ddots \end{bmatrix}$$

with $\Delta x = \frac{P}{N}$

Note: The notation $\mathcal{F}_N$ leaves the dependence on Δx implicit.

In Julia the FFTW package gives access to these operators.

$W_N \equiv$ plan_fft(...) $\leftarrow$ operator form of fft(·)

$\mathcal{U}_N \equiv \frac{1}{\sqrt{N}}$ * plan_fft(...)

$\mathcal{F}_N \equiv \frac{(P/N)}{\sqrt{2\pi}}$ * plan_fft(...)

$W_N^{-1} \equiv$ plan_ifft(...) $\leftarrow$ operator form of ifft(·)

$N\,W_N^{-1} \equiv W_N^H \equiv$ plan_bfft(...)

operator form of bfft(·)

So, now given $v \in \mathbb{R}^N$ or $\mathbb{C}^N$ we can easily ask Julia to compute

$$z = \text{fft}(v) = W_N v$$

where

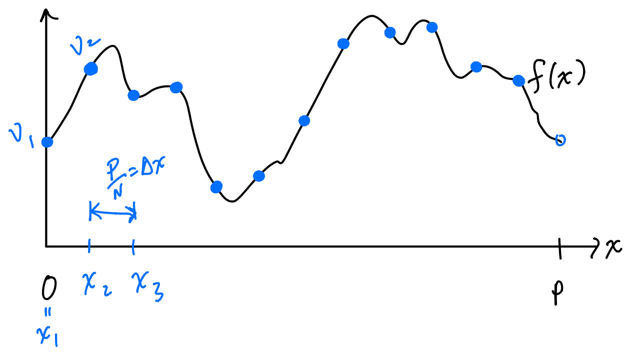
$$z_k = \sum_{n=1}^{N} w_N^{(n-1)(k-1)} v_n .$$

The goal in this part of the notes is to show how to interpret z_k in the case that

$$v = \begin{bmatrix} \vdots \\ v_n \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ f(x_n) \\ \vdots \end{bmatrix}$$

for $x_n = \frac{(n-1)}{N}P$, $n = 1, \dots, N$

for some smooth $f: [0,P) \to \mathbb{R}$ or $\mathbb{C}$



Notice there are two things the above picture gives us

(1) The indicies of $v_1, \dots, v_n$ map one-to-one to gridpoints $x_1, \dots, x_n \in \mathbb{R}$

(2) The pairs (x_n, v_n) give a discrete approximation of the continuous map $x \mapsto f(x)$

The implication of (1) is that we can interprete $z = W_N v$ as the coefficients in a sin/cos basis expansion of v

The question is: what are the corresponding sin/cos frequencies?

To derive them first notice that

$$v = W_N^{-1} z = \frac{1}{N} W_N^H z$$

Therefore

$$\begin{aligned} v_n &= \frac{1}{N} \sum_{k=1}^{N} \bar{w}_N^{(n-1)(k-1)} z_k \\ &= \frac{1}{N} \sum_{k=1}^{N} \exp\left(i \tfrac{2\pi}{N} (n-1)(k-1)\right) z_k \\ &= \frac{1}{N} \sum_{k=1}^{N} \exp\Big(i \underbrace{\left[(n-1)\tfrac{P}{N}\right]}_{=: x_n} \underbrace{\left[(k-1)\tfrac{2\pi}{P}\right]}_{=: \tau_k}\Big) z_k \\ &= \frac{1}{N} \sum_{k=1}^{N} \Big(\cos(x_n \tau_k) + i \sin(x_n \tau_k)\Big) z_k \end{aligned}$$

P is the period of the interval $x_1, \dots, x_N$ sampled

So for each $k = 1, \dots, N$ the DFT value z_k is the coefficient (with a $\frac{1}{N}$ factor) on the harmonic basis with "frequency"

$$\tau_k = (k-1) \frac{2\pi}{P}$$

However recall that when $\cos(x \tau_k)$ and $\sin(x \tau_k)$ are observed on a grid $x_1, \dots, x_n$ and $\tau_k > \tau_{N/2+1} = \text{Nyquist}$ the effective frequency on the grid is closer to zero.

To derive the effective frequency of

$$x \mapsto \exp(i x \tau_k)$$

when observed on grid $x_1, \dots, x_N$ we use the N^{th} root of unity property

$$w_N^{\pm N} = 1 \qquad \text{(derived above)}$$

$$\therefore \; 1 = \left(w_N^{\pm N}\right)^{n-1}$$

$$= \exp\left(i \frac{2\pi}{N}(n-1)(\pm N)\right)$$

$$= \exp\left(i \left(\frac{n-1}{N} P\right)\left(\pm N \frac{2\pi}{P}\right)\right)$$

$$= \exp\Big(i \underbrace{\left((n-1)\frac{P}{N}\right)}_{= x_n} \underbrace{\left(\pm \frac{2\pi}{P/N}\right)}_{= \pm \frac{2\pi}{\Delta x}}\Big)$$

$$= \exp\left(i x_n (\pm 2 \, nyq)\right)$$

$nyq = \frac{2\pi}{2\Delta x}$

Therefore

Nominal freq.

$q \in \mathbb{Z}$

$$\exp(i x_n \tau_k)$$

$$= \exp\left(i x_n (\tau_k + q \, 2 \, nyq)\right)$$

$$= \exp\left(i x_n \left(\tau_k \bmod \frac{2\pi}{\Delta x}\right)\right)$$

Definition:

The effective frequency τ^{nyq} associated with τ and a grid

$$x_n := (n-1)\frac{P}{N}, \quad n \in \{1, \dots, N\}$$

is defined to be the number which satisfies the following two conditions:

(i) $\tau^{nyq} \in \{\tau + q \, 2 \, nyq : q \in \mathbb{Z}\}$

(ii) $\tau^{nyq} \in (-nyq, nyq]$

Notice it is just my convention to choose left open/Right closed here. You can choose left closed/Right open if you want.

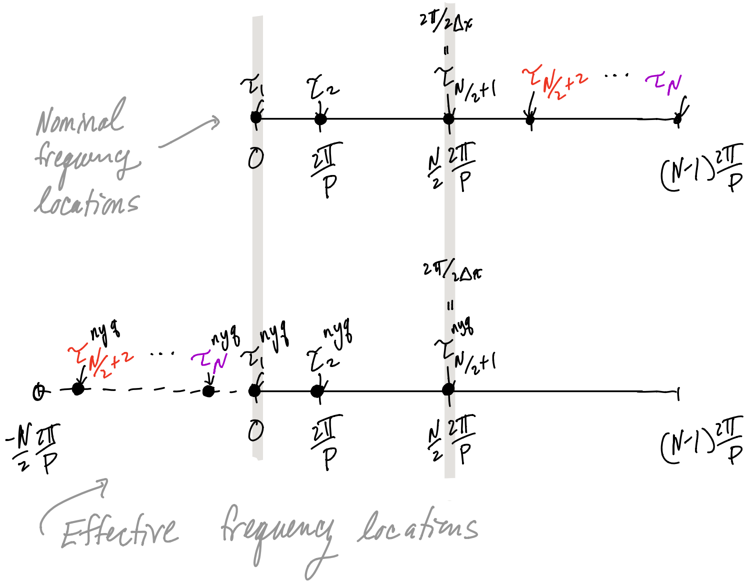
$nyq = \frac{\pi}{\Delta x} = \frac{N}{2}\frac{2\pi}{P} = \frac{N}{2}\Delta k$

Going back to our DFT decomposition of v we have

$$v_n = \frac{1}{N}\sum_{n=1}^{N} \exp(i x_n \tau_k) z_k$$

$$= \frac{1}{N}\sum_{n=1}^{N} \exp(i x_n \tau_k^{nyq}) z_k$$

Here is the picture when N is even



When N is odd we have

(letting $N \div 2$ denote $\lfloor N/2 \rfloor$)

$$\tau_1 = 0 \quad < nyq$$

$$\tau_2 = \frac{2\pi}{P} \quad < nyq$$

$$\vdots$$

$$\tau_{N\div 2+1} = \left\lfloor \frac{N}{2} \right\rfloor \frac{2\pi}{P} \quad < \quad \frac{N}{2}\frac{2\pi}{P} = nyq$$

$$\tau_{N\div 2+2} = \left(\left\lfloor \frac{N}{2} \right\rfloor + 1\right)\frac{2\pi}{P} \quad > nyq$$

$$\vdots$$

$$\tau_N = (N-1)\frac{2\pi}{P} \quad > nyq$$

So the picture of where the effective frequencies move to doesn't change, except that no frequency τ_k attains the exact nyqist limit in the case N is odd.

Now lets look at ...

(2) The pairs (x_n, v_n) give a discrete approximation of the continuous map $x \mapsto f(x)$

... which suggests that (τ_k^{nyq}, z_k) approximates the continuous Fourier transform (CFT) of f. Without worrying about rigour lets define the CFT of f as follows

$$\tilde{f}(\omega) = \int e^{-ix\cdot\omega} f(x)\frac{dx}{\sqrt{2\pi}} \qquad \text{(CFT)}$$

and the inverse transform

$$f(x) = \int e^{ix\cdot\omega}\tilde{f}(\omega)\frac{d\omega}{\sqrt{2\pi}} \qquad \text{(ICFT)}.$$

If one wanted to approximate $\tilde{f}(\omega)$ with a Riemann sum over $x_1, \dots, x_N \in [0, P)$ with $x_n = (n-1)\frac{P}{N}$

$$\tilde{f}(\omega) \approx \sum_{n=1}^{N} e^{-ix_n\cdot\omega} f(x_n)\frac{\Delta x}{\sqrt{2\pi}} \qquad (\ast)$$

If we restrict ω to effective frequencies, i.e.

$$\omega \in \{\tau_1^{nyq}, \dots, \tau_N^{nyq}\}$$

Then $(\ast)$ simplifies to

$$z = \frac{\Delta x}{\sqrt{2\pi}} W_N v = \mathcal{F}_N v$$

under the correspondance $z_k \approx \tilde{f}(\omega)$ when $\tau_k^{nyq} = \omega$

Likewise

$$V = \mathcal{F}_N^{-1} z = \frac{\sqrt{2\pi}}{N\Delta x} \overline{W}_N z$$

$$= \frac{\sqrt{2\pi}}{N\, P/N} \overline{W}_N z$$

$$= \frac{(2\pi)/P}{\sqrt{2\pi}} \overline{W}_N z$$

grid spacing of the effective fourier frequencies. Call it $\Delta\tau$

$\therefore$

$$v_n = \frac{(2\pi)/P}{\sqrt{2\pi}} \sum_{k=1}^{N} \exp(i x_n \tau_k^{nyq}) z_k$$

$$= \sum_{k=1}^{k} \exp(i x_n \tau_k^{nyq}) z_k \frac{\Delta\tau}{\sqrt{2\pi}}$$

which is clearly a Reimann Sum discretization of

$$f(x_n) = \int \exp(i x_n w) \tilde{f}(w) \frac{dw}{\sqrt{2\pi}}$$

over gridded w's in $\{\tau_1^{nyq}, \dots, \tau_n^{nyq}\}$.

To summarize for

$k, n \in \{1, \dots, N\}$

$w_k := \tau_k^{nyq} \in \left(\frac{2\pi}{P}\mathbb{Z}\right) \cap \left(-N\frac{\pi}{P}, N\frac{\pi}{P}\right]$

$\Delta w = \frac{2\pi}{P}$

$x_n \in \left(\frac{P}{N}\mathbb{Z}\right) \cap [0, P)$, $\Delta x = \frac{P}{N}$

Discrete approx of CFT	DFT
$\tilde{f}(w_k) = \sum_{n=1}^{N} e^{-i x_n w_k} f(x_n) \frac{\Delta x}{\sqrt{2\pi}}$	$z = \frac{\Delta x}{\sqrt{2\pi}} W_N v$
$f(x_n) = \sum_{k=1}^{N} e^{i x_n w_k} \tilde{f}(w_k) \frac{\Delta w}{\sqrt{2\pi}}$	$v = \frac{\sqrt{2\pi}}{\Delta x} W_N^{-1} z$

These are the same under re-labeling

$v_n = f(x_n)$, $z_k = \tilde{f}(w_k)$

Circulant Covariance Matrices

One of the reasons the Fourier transform is important for us: it diagonalizes stochastic processes which are stationary.

In this section we derive a discrete version of this fact.

Definition:

A N-by-N matrix Σ is circulant if it has the form

$$\Sigma = \begin{bmatrix} c_1 & c_N & \cdots & & c_2 \\ c_2 & c_1 & \ddots & & \\ \vdots & c_2 & \ddots & \ddots & c_N \\ & \vdots & \ddots & \ddots & \\ c_N & c_{N-1} & & c_2 & c_1 \end{bmatrix} \qquad (*)$$

Theorem:

If Σ has the form $(*)$ then

$$\Sigma = \mathcal{U}_N^H \Lambda \mathcal{U}_N$$

$$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$$

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{bmatrix} = W_N \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix}$$

If, in addition, Σ is real and symmetric, then $\lambda_1, \ldots, \lambda_N$ are real and are the eigenvalues of Σ. Also in this case we additionally have:

$$\Sigma = \mathcal{U}_N \Lambda \mathcal{U}_N^H.$$

Proof:

The proof is based on an analysis of the circular shift permutation matrix:

row 1 = [0, ..., 0 1]

$$P_N := \begin{bmatrix} 0 & & & 0 & 1 \\ 1 & \ddots & & & 0 \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{bmatrix}$$

row 2 = [1, 0, ..., 0]

Some facts about P_N:

(I) $P_N \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} x_N \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix}$

i.e. left multiply by P_N does downward circular shift.

(II) $P_N^T = P_N^{-1}$

(III)

$$P_N^m = \begin{bmatrix} & & 0 & 1 & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ 0 & & & & & 0 \\ 1 & \ddots & & & & \\ & \ddots & \ddots & & & \\ & & 1 & 0 & & \end{bmatrix}$$

row m = [0, ..., 0 1]

row m+1 = [1, 0, ..., 0]

for example $P_N^N = I$

(IV) $P_N = \mathcal{U}_N^H \, \mathrm{diag}(w_N^0, \ldots, w_N^{N-1}) \, \mathcal{U}_N$

unitary DFT

Recall $w_N = \exp(-i \frac{2\pi}{N})$

for a reference to (IV) see P.J. Davis, "Circulant Matrices", 1979

(V)

$$X \sim N(0,\Sigma) \text{ and } P_N X \overset{\mathcal{D}}{=} X$$

$$\Rightarrow \Sigma = P_N \Sigma P_N^T \underset{\text{by (II)}}{=} P_N \Sigma P_N^{-1}$$

$$\Rightarrow P_N \Sigma = \Sigma P_N$$

i.e. Σ and P_N commute

(VI)

$$\Sigma = c_1 I + c_2 P_N + c_3 P_N^2 + \cdots + c_N P_N^{N-1}$$

$$= \sum_{n=1}^{N} c_n P_N^{(n-1)}$$

This follows by direct calculation.

Note: We only need (IV) and (VI) to prove the theorem ... but we will use the other facts later so I've listed them here.

Now to prove the theorem notice that

$$\Sigma = \sum_{n=1}^{N} c_n P_N^{(n-1)} \quad \text{by (VI)}$$

$$= U_N^H \underbrace{\left[\sum_{n=1}^{N} c_n \left(\mathrm{diag}(w_N^0, \ldots, w_N^{N-1}) \right)^{(n-1)} \right]}_{=: \Lambda \;\leftarrow\; \text{diagonal}} U_N$$

with
$$\Lambda_{k,k} = \sum_{n=1}^{N} c_n w_N^{(k-1)(n-1)}$$

$$= k^{th} \text{ coordinate of } W_N \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix}$$

as was to be shown.

Now suppose in addition that Σ is real and symmetric. Then

$$\Sigma = \Sigma^T = U_N \Lambda U_N^H \leftarrow \text{as was to be shown}$$

and

$$\Sigma = \Sigma^H = U_N^H \bar{\Lambda} U_N$$

so that

$$\Sigma = \tfrac{1}{2}(\Sigma + \Sigma^H) = U_N^H \underbrace{Re(\Lambda)} U_N$$

i.e. $\Lambda = Re(\Lambda)$ as was to be shown.

Finally the fact that

$$\Sigma = U_N^H \Lambda U_N$$
$$\Sigma = U_N \Lambda U_N^H$$

and that $U_N^T = U_N$ means

$$\Sigma\, U_N[:,k] = \lambda_k U_N[:,k]$$
$$\Sigma\, \overline{U_N[:,k]} = \lambda_k \overline{U_N[:,k]}$$

which implies

$$\underbrace{\begin{aligned} \Sigma\, Re(U_N[:,k]) &= \lambda_k Re(U_N[:,k]) \\ \Sigma\, Im(U_N[:,k]) &= \lambda_k Im(U_N[:,k]) \end{aligned}}$$

i.e.

$$n \mapsto \cos(2\pi(n-1)(k-1)/N)/\sqrt{N}$$
$$n \mapsto \sin(-2\pi(n-1)(k-1)/N)/\sqrt{N}$$

are the eigenvectors of Σ with eigenvalue λ_k.

Now we immediately get a result for Gaussian vectors with a circular shift invariance.

Corollary:

If $(X_1, \dots, X_N)^T \sim \mathcal{N}(0, \Sigma)$ has the property that

cov matrix $\in \mathbb{R}^{N\times N}$

$$(X_1, \dots, X_N)^T \overset{\mathcal{D}}{=} P_N (X_1, \dots, X_N)^T$$

then

$$\Sigma = \mathcal{U}_N^H \Lambda \, \mathcal{U}_N$$

where $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_N)$

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{bmatrix} = W_N \begin{bmatrix} \operatorname{cov}(X_1, X_1) \\ \vdots \\ \operatorname{cov}(X_N, X_1) \end{bmatrix}$$

all real. These are the eigenvalues of Σ

first column of Σ

Note: This corollary can be written in different scalings...

$$\Sigma = \mathcal{U}_N^H \operatorname{diag}(W_N \Sigma[:,1]) \, \mathcal{U}_N \qquad (\Sigma\mathcal{U})$$

$$= \cancel{\sqrt{N}} W_N^{-1} \operatorname{diag}(W_N \Sigma[:,1]) \frac{1}{\cancel{\sqrt{N}}} W_N \qquad (\Sigma W)$$

$$= \mathcal{F}_N^{-1} \operatorname{diag}\left(\frac{\sqrt{2\pi}}{\Delta x} \mathcal{F}_N \Sigma[:,1]\right) \mathcal{F}_N \qquad (\Sigma F)$$

... and the form useful for cov modeling in the spectral domain:

$$\mathcal{F}_N \Sigma \mathcal{F}_N^H \overset{\text{by } (\Sigma F)}{=} \mathcal{F}_N \mathcal{F}_N^{-1} \operatorname{diag}\left(\frac{\sqrt{2\pi}}{\Delta x} \mathcal{F}_N \Sigma[:,1]\right) \mathcal{F}_N \mathcal{F}_N^H$$

where $\mathcal{F}_N = \frac{\Delta x}{\sqrt{2\pi}} W_N = \frac{\Delta x}{\sqrt{2\pi}} \sqrt{N}\, \mathcal{U}_N$ implies

$$\mathcal{F}_N \mathcal{F}_N^H = \frac{(\Delta x)^2}{2\pi} N I$$
$$= \frac{(P/N)^2 N}{2\pi} I = \frac{P/N}{(2\pi/P)} I$$

Therefore

$$\mathcal{F}_N \mathcal{F}_N^H = \frac{\Delta x}{\Delta \xi} I \qquad (FF^H)$$

$$\mathcal{F}_N \Sigma \mathcal{F}_N^H = \frac{1}{\Delta \xi} \operatorname{diag}\left(\sqrt{2\pi}\, \mathcal{F}_N \Sigma[:,1]\right) \qquad (F\Sigma)$$

Remark:

Notice that in the case Σ approximates a kernel integral operator $\mathcal{K}$:

$$\Sigma \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{bmatrix} \approx \begin{bmatrix} \int_0^P K(x_1 - x) f(x)\, dx \\ \\ \int_0^P K(x_n - x) f(x)\, dx \end{bmatrix}$$

i.e.

$$\Sigma = \begin{bmatrix} K(x_1 - x_1)\Delta x & \cdots & K(x_1 - x_N)\Delta x \\ \vdots & \ddots & \vdots \\ K(x_N - x_1)\Delta x & \cdots & K(x_N - x_N)\Delta x \end{bmatrix}$$

Then we have

$$\Sigma = \mathcal{F}_N^{-1} \operatorname{diag}\left(\sqrt{2\pi}\, \mathcal{F}_N v_K\right) \mathcal{F}_N \qquad (KF)$$

where $v_K := (K(x_1), \dots, K(x_N))^T$

log likelihood of DFT $\cdot X$

Again suppose

Random vector of length N — circ shift left by 1

$$X \sim N(0, \Sigma) \text{ with } P_N X \overset{\mathcal{D}}{=} X.$$

Lets see how to compute $\log P_r(X=x)$

Nominally we have

$$\log P_r(X=x) = -\frac{1}{2} x^T \Sigma^{-1} x - \frac{1}{2}\log|\Sigma|$$

both these are expensive to compute So lets see how to simplify with the DCT.

Let

$$\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_N) = \operatorname{diag}\left(W_N \begin{bmatrix} \Sigma_{1,1} \\ \vdots \\ \Sigma_{N,1} \end{bmatrix}\right)$$

so that

$$x^T \Sigma^{-1} x = x^T \mathcal{U}_N^H \Lambda^{-1} \mathcal{U}_N x$$

$\mathcal{U}_N^H \Lambda \mathcal{U}_N$

$$= (\mathcal{U}_N x)^H \Lambda (\mathcal{U}_N x)$$

This is z — since $x \in \mathbb{R}^N$

$$= \sum_{k=1}^{N} \frac{|z_k|^2}{\lambda_k} \quad \text{where } z = \mathcal{U}_N x$$

and

$$\log|\Sigma| = \cancel{\log|\mathcal{U}_N|}^{\,0} + \cancel{\log|\mathcal{U}_N^H|}^{\,0} + \sum_{k=1}^{N} \log\lambda_k$$

Therefore

$$\log P(X=x) = -\frac{1}{2}\sum_{k=1}^{N} \frac{|z_k|^2}{\lambda_k} - \frac{1}{2}\sum_{k=1}^{N}\log\lambda_k$$

Remark: Since X takes values in $\mathbb{R}^N$ there is extra structure in $z = \mathcal{U}_N x$ for $X=x$, that can be exploited to cut CPU time in half. Let $x \in \mathbb{R}^N$ and $z = W_N x$ (or $\mathcal{U}_N x$ or $\Theta_N x$).

For the first coordinate of z:

$$\bar{z}_1 = \sum_{n=1}^{N} \bar{x}_n = \sum_{n=1}^{N} x_n = z_1$$

since $x \in \mathbb{R}^n$

For z_k for $k \geq 2$ we have

$$\bar{z}_k = \sum_{n=1}^{N} \bar{w}_N^{(n-1)(k-1)} \bar{x}_n$$

$$= \sum_{n=1}^{N} (w_N)^{-(n-1)(k-1)} x_n$$

Since $x_n \in \mathbb{R}$

$$= \sum_{n=1}^{N} (w_N)^{-(n-1)[(k-1)-N]} x_n$$

Since w_N is an N^{th} root of 1

$$= \sum_{n=1}^{N} w_N^{(n-1)[N-k+2-1]} x_n$$

$$= z_{N-k+2}$$

In particular the "first half" of $(z_2, \dots, z_N)^T$ is a mirror (conj) reflection of the "second half" of $(z_2, \dots, z_N)^T$

Here is a diagram:

Case 1: N is odd

implies $z_1 \in \mathbb{R}$

even number of terms after z_1

halfway

$$\begin{bmatrix} \bar{z}_1 \\ \bar{z}_2 \\ \vdots \\ \bar{z}_{N \div 2+1} \\ \hline \bar{z}_{N \div 2+2} \\ \vdots \\ \bar{z}_N \end{bmatrix} = \begin{bmatrix} z_1 \\ z_N \\ \vdots \\ z_{N \div 2+2} \\ \hline z_{N \div 2+1} \\ \vdots \\ z_2 \end{bmatrix}$$

In particular

$$(z_2, \dots, z_{N \div 2+1})^T = (z_N, \dots, z_{N \div 2+2})^H$$

which means the terms $z_{N \div 2+2}, \dots, z_N$ are redundant.

In Julia one can save computation time using rfft:

takes about ½ the time as fft

$$(z_1, z_2, \dots, z_{N \div 2+1})^T = \mathrm{rfft}(\underset{\in \mathbb{R}^N}{x})$$

$$x = \mathrm{irfft}((z_1, z_2, \dots, z_{N \div 2+1})^T, \underset{N \text{ odd}}{N})$$

To use this rfft output for $\log P(X=x)$ notice

$$\log P(X=x)$$

$$= -\frac{1}{2} \sum_{k=1}^{N} \left(\frac{|z_k|^2}{\lambda_k} + \log \lambda_k \right)$$

$$= \underbrace{-\frac{1}{2} \left(\frac{|z_1|^2}{\lambda_1} + \log \lambda_1 \right)}_{z_1 \text{ is real \& gets a } \frac{1}{2} \text{ coeff}} - \underbrace{\sum_{k=2}^{N \div 2+1} \left(\frac{|z_k|^2}{\lambda_k} + \log \lambda_k \right)}_{\text{these are complex and get a 1 coeff}}$$

Case 2: N is even

implys $z_1 \in \mathbb{R}$

odd number of terms after z_1

halfway

$$\begin{bmatrix} \bar{z}_1 \\ \bar{z}_2 \\ \vdots \\ \bar{z}_{N \div 2} \\ \boxed{\bar{z}_{N \div 2+1}} \\ \bar{z}_{N \div 2+2} \\ \vdots \\ \bar{z}_N \end{bmatrix} = \begin{bmatrix} z_1 \\ z_N \\ \vdots \\ z_{N \div 2+2} \\ \boxed{z_{N \div 2+1}} \\ z_{N \div 2} \\ \vdots \\ z_2 \end{bmatrix}$$

implies $z_{N \div 2+1} \in \mathbb{R}$

Now terms $z_{N \div 2+2}, \dots, z_N$ are redundant and we have the additional restriction that $z_{N \div 2+1} \in \mathbb{R}$. Julia's rfft still works the same, i.e.

$$(z_1, z_2, \dots, z_{N \div 2+1})^T = \mathrm{rfft}(x)$$

$$x = \mathrm{irfft}((z_1, z_2, \dots, z_{N \div 2+1})^T, \underset{N \text{ even}}{N})$$

Also, when N is even we have

$$\log P(X=x)$$

$$= -\underbrace{\frac{1}{2}\left(\frac{|z_1|^2}{\lambda_1} + \log\lambda_1\right)}_{\text{z_1 is real \& gets a $\frac{1}{2}$ coeff}}$$

limit $N \div 2$ instead of $N \div 2+1$

$$-\sum_{k=2}^{N\div 2}\left(\frac{|z_k|^2}{\lambda_k} + \log\lambda_k\right)$$

$$-\frac{1}{2}\underbrace{\left(\frac{|z_{N\div 2+1}|^2}{\lambda_{N\div 2+1}} + \log\lambda_{N\div 2+1}\right)}_{\text{$z_{N\div 2+1}$ is real and gets a $\frac{1}{2}$ weight}}$$

complex terms get a 1 weighting

One advantage of examining $\log P(X=x)$ in both $N=$even and $N=$odd cases is that examining the likelihood as a function of real and imaginary parts of X one can derive the equivalent simulation model for the unique modes in $z = \mathcal{U}_N X$ as follows:

all independent

$$\begin{cases} z_1 \sim N(0,\lambda_1) \\ z_2 \sim N(0,\frac{\lambda_2}{2}) + iN(0,\frac{\lambda_2}{2}) \\ \vdots \\ z_{N\div 2} \sim N(0,\frac{\lambda_{N\div 2}}{2}) + iN(0,\frac{\lambda_{N\div 2}}{2}) \\ z_{N\div 2+1} \sim \begin{cases} N(0,\frac{\lambda_{N\div 2+1}}{2}) + iN(0,\frac{\lambda_{N\div 2+1}}{2}) & \text{if N is odd} \\ N(0,\lambda_{N\div 2+1}) & \text{if N is even} \end{cases} \end{cases}$$

Application to stationary processes on the unit circle

Let f denote a mean 0 GRF on S^1 where

$$S^1 := \text{unit circle} \equiv \{ x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1^2 \}.$$

The field f can be indexed by $x \in S^1$ or equiv by angles $\varphi \in (0, 2\pi]$:

$$\big(f(\varphi) : \varphi \in (0,2\pi]\big) \sim GRF_{S^1}(0, K)$$

$$\equiv$$

$$\big(f(x) : x = (x_1, x_2)^T,\ x_1^2 + x_2^2 = 1^2\big) \sim GRF_{S^1}(0, K)$$

with correspondence given by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix}$$

In addition suppose f is stationary in the sense that

$$\big(f(x) : x \in S^1\big) \overset{fdd\ \theta}{=} \big(f(R_\theta x) : x \in S^1\big)$$

For any rotation $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

This implies that $\mathrm{cov}(f(x), f(y))$ depends only on the angular distance $\theta_{x,y}$ between $x, y \in S^1$, i.e.

$$\mathrm{cov}(f(x), f(y)) = K(\theta_{x,y}).$$

where $\langle x, y\rangle = \cos(\theta_{x,y})$

In addition note that

$$\|x-y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 - 2\langle x, y\rangle$$

$$= 2 - 2\cos(\theta_{x,y})$$

$$= 2\big(1 - \cos(\theta_{x,y})\big)$$

$$\therefore \cos(\theta_{xy}) = 1 - \tfrac{1}{2}\|x-y\|^2$$

So that

$$\mathrm{cov}(f(x), f(y)) = K(\theta_{x,y}) \quad \leftarrow \text{Function of geodesic distance}$$

$$= K\Big(\cos^{-1}\big(1 - \tfrac{1}{2}\|x-y\|^2\big)\Big) \quad \leftarrow \text{Function of cordal distance}$$

$$= K\big(\cos^{-1}(\langle x, y\rangle)\big) \quad \leftarrow \text{Function of } \langle x,y\rangle$$

Now it is easy to see that when $f(\varphi)$ is observed on a grid

$$\varphi_1, \dots, \varphi_N \in S^1 \quad \text{s.t.}$$

$$\varphi_n := (n-1)\frac{2\pi}{N} \text{ for } n \in \{1, \dots, N\}$$

Then

$$v^f := \big(f(\varphi_1), \dots, f(\varphi_N)\big)^T \sim N(0, \Sigma)$$

where Σ is circulant with entries

$$\Sigma_{n,m} = K\big(\cos^{-1}(\langle x_n, x_m\rangle)\big)$$

$$x_n = \begin{bmatrix} \cos(\varphi_n) \\ \sin(\varphi_n) \end{bmatrix}$$

and setting

$$z^f := \mathcal{F}_N v^f$$

we have

$$E\big((z^f)(z^f)^H\big) = \underbrace{\frac{1}{\Delta\Sigma}\mathrm{diag}\big(\sqrt{2\pi}\,\mathcal{F}_N \Sigma[:,1]\big)}_{\mathcal{F}_N \Sigma \mathcal{F}_N^H \text{ by } (F\Sigma)}$$

Remark:

The continuum version of this (as $N \to \infty$ and replacing Reimann sums with integrals) suggests

$$E(f_\tau \bar{f}_{\tau'}) = \delta_{\tau-\tau'} C^{ff}_\tau$$

where

approximates $\mathcal{F}_N v^f$

$$f_\tau := \int_0^{2\pi} \exp(-i\tau\varphi) f(\varphi) \frac{d\varphi}{\sqrt{2\pi}}$$

approximates $\sqrt{2\pi}\, \mathcal{F}_N \Sigma[:,1]$

$$C^{ff}_\tau := \int_0^{2\pi} \exp(-i\tau\varphi) \operatorname{cov}(f(\varphi), f(0))\, d\varphi$$

can replace with $\exp(i\tau\varphi)$ or $\cos(\tau\varphi)$ by symmetry of $\operatorname{cov}(f(\varphi), f(0))$.

$\delta_\tau :=$ dirac delta function of frequency

$$\equiv \begin{cases} \frac{1}{d\tau} & \text{if } \tau = 0 \\ 0 & \text{o.w.} \end{cases}$$

so that $\int \delta_\tau g(\tau) d\tau = \frac{1}{d\tau} g(0)\, d\tau = g(0)$

In this case C^{ff}_τ is called the spectral density of f and yeilds a continous approximation to the discrete result: i.e.

$$E((z^f)(z^f)^T) = \frac{1}{\Delta\tau} \operatorname{diag}\left(\sqrt{2\pi}\, \mathcal{F}_N \Sigma[:,1]\right)$$

$$\Downarrow \text{ as } N \to \infty$$

$$E(f_\tau \bar{f}_{\tau'}) = \delta_{\tau-\tau'} C^{ff}_\tau$$

White Noise

Usually the process $(f(\varphi) : \varphi \in [0, 2\pi))$ as above, is observed with noise:

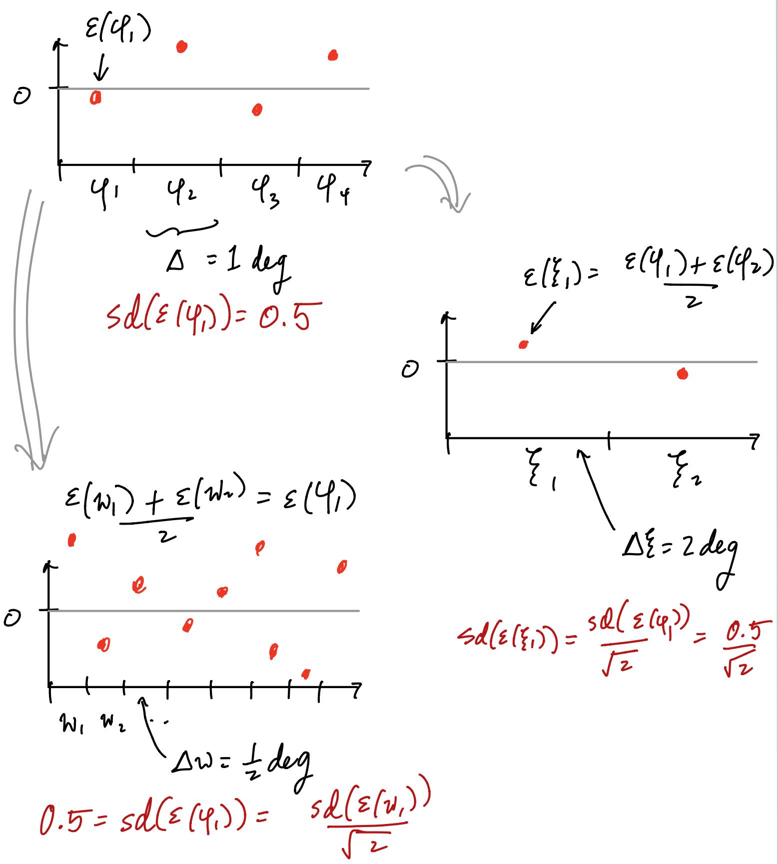
In cosmology often noise is quantified by standard deviation per unit pixel volumn (i.e length in 1-d, area in 2-d).

For example one might write

$$\underbrace{d(\varphi)}_{\text{data}} = \underbrace{f(\varphi)}_{\text{signal}} + \underbrace{\varepsilon(\varphi)}_{\text{noise}}$$

where $\varepsilon(\varphi)$ is white noise with $sd(\varepsilon(\varphi_n)) = 0.5$ when observing on a pixel grid with spacing $\Delta\varphi = 1$ deg.

The idea is that one can now compute the noise for other grids by up/down sampling and the formulas for variance of averages.



Lets compare this notion of white noise with the more traditional definition:

white noise $W(\varphi)$ satisfies

$$\int f(\varphi) g(\varphi) d\varphi = cov\left(\int f(\varphi) W(\varphi) d\varphi, \int g(\varphi) W(\varphi) d\varphi\right)$$

To relate $W(\varphi)$ to $\varepsilon(\varphi)$ replace the integral with a Reimann sum over a grid φ_n with spacing $\Delta\varphi$

$$\sum_n f(\varphi_n) g(\varphi_n) \Delta\varphi$$
$$= cov\left(\sum_n f(\varphi_n) W(\varphi_n) \Delta\varphi, \sum_n g(\varphi_n) W(\varphi_n) \Delta\varphi\right)$$
$$= \sum_{n,m} f(\varphi_n) g(\varphi_m) E(W(\varphi_n) W(\varphi_m)) (\Delta\varphi)^2$$

Match left and right sides above gives

$$E(W(\varphi_n) W(\varphi_m)) = \begin{cases} \frac{1}{\Delta\varphi} & \text{if } n = m \\ 0 & \text{o.w.} \end{cases}$$

which has the same scaling in $\Delta\varphi$ as defined for $\varepsilon(\varphi)$

Example:

Let $var(\varepsilon(\varphi_n)) = \sigma^2$ with $\Delta\varphi = 1$ deg.

$$\Delta\varphi = 1 \text{ deg} \Rightarrow var(\sigma W(\varphi_n)) = \frac{\sigma^2}{\Delta\varphi} = \sigma^2$$

$$\Delta\xi = 2 \text{ deg} \Rightarrow var(\sigma W(\xi_n)) = \frac{\sigma^2}{\Delta\xi} = \frac{\sigma^2}{2\Delta\varphi} = \frac{\sigma^2}{2}$$

$$\Delta w = \tfrac{1}{2} \text{ deg} \Rightarrow var(\sigma W(w_n)) = \frac{\sigma^2}{\Delta w} = \frac{\sigma^2}{(\Delta\varphi/2)} = 2\sigma^2$$

$$\text{i.e.} \quad \varepsilon(\varphi) \overset{\mathcal{D}}{\equiv} \sigma W(\varphi).$$

Now we know if we have noise $\varepsilon(\varphi)$ with $\mathrm{var}(\varepsilon(\tilde{\varphi}_n)) = \sigma^2$ for a reference grid spacing $\Delta\tilde{\varphi} = 1$ then

$$v^\varepsilon = \begin{bmatrix} \varepsilon(\varphi_1) \\ \vdots \\ \varepsilon(\varphi_N) \end{bmatrix} \Rightarrow E\left((v^\varepsilon)(v^\varepsilon)^T\right) = \frac{\sigma^2}{\Delta\varphi} I$$

in units of $\Delta\tilde{\varphi}$

Lets find $E\left((z^\varepsilon)(z^\varepsilon)^H\right)$ when $z_\varepsilon = \mathcal{F}_N v_\varepsilon$

$$E\left((z^\varepsilon)(z^\varepsilon)^H\right) = \mathcal{F}_N E\left((v^\varepsilon)(v^\varepsilon)^H\right) \mathcal{F}_N^H \qquad \frac{\sigma^2}{\Delta\varphi} I$$

$$= \left(\frac{\sigma^2}{\Delta\varphi} I\right) \underbrace{\mathcal{F}_N \mathcal{F}_N^H}_{\frac{\Delta\varphi}{\Delta\tau} I \text{ by } (FF^H)}$$

$$= \frac{\sigma^2}{\Delta\tau} I$$

so

$$E\left((v^\varepsilon)(v^\varepsilon)^H\right) = \frac{\sigma^2}{\Delta\varphi} I$$

pixel grid spacing

$$E\left((z^\varepsilon)(z^\varepsilon)^H\right) = \frac{\sigma^2}{\Delta\tau} I$$

Fourier grid spacing.

This explains why it is called white noise...

$$\mathrm{var}(z_1^\varepsilon) = \cdots = \mathrm{var}(z_N^\varepsilon) = \frac{\sigma^2}{\Delta\tau}$$

i.e. the variance of the fourier coefficients is constant (like pure white light)

Here is a <u>summary</u> for a generic noise R.V. X

$$\underbrace{X \sim N\left(0, \frac{\sigma^2}{\Delta x} I\right)}_{\text{white noise on a grid}} \Rightarrow \begin{cases} E(zz^H) = \dfrac{\sigma^2}{\Delta\tau} I \\ \text{where } \quad z = \mathcal{F}_N X \end{cases}$$

Which implies

$$X \sim N(0, \sigma^2 I) \Rightarrow \begin{cases} E(zz^H) = \sigma^2 \dfrac{\Delta x}{\Delta\tau} I \\ \text{where } \quad z = \mathcal{F}_N X \end{cases}$$

This is mostly used when $\mathcal{F}_N$ is meaningful for the application at hand (i.e. when the signal spectral density is known, so using $\mathcal{F}_N$, rather than $\mathcal{U}_N$, makes sense)

Otherwise using $\mathcal{U}_N$ is easier to remember

$$X \sim N(0, \sigma^2 I) \Rightarrow \begin{cases} E(ww^H) = \sigma^2 I \\ \text{where } \quad w = \mathcal{U}_N X \end{cases}$$

Beams

Another common feature in random field observations taken with an optical lense is a beam, which models the way light spreads over the pixel detectors as a convolution with a blurring kernel

The beamed process $(f(\varphi) : \varphi \in [0, 2\pi))$ is written:

$$\underset{\text{beam, i.e. convolution, operator}}{B}\; \underset{\text{signal}}{f(\varphi)}$$

where $\underset{\text{azmuth angle } \in [0,2\pi)}{B f(\varphi)} = \int_0^{2\pi} \underbrace{b(\xi - \varphi)}_{\text{centered at } \varphi} f(\xi)\, d\xi$

for some $b(\varphi) \geq 0$ s.t. $\int_0^{2\pi} b(\varphi)\, d\varphi = 1$

where $b(\xi)$ is 2π-periodic.

Notice that the discrete matrix version of this operator is a circulant matrix

$$\underset{\text{discrete version on a grid of size } N}{B_N^b} = \begin{bmatrix} b(\varphi_1 - \varphi_1)\Delta\varphi & \cdots & b(\varphi_1 - \varphi_N)\Delta\varphi \\ b(\varphi_2 - \varphi_1)\Delta\varphi & \ddots & \vdots \\ \vdots & & \\ b(\varphi_N - \varphi_1)\Delta\varphi & \cdots & b(\varphi_N - \varphi_N)\Delta\varphi \end{bmatrix}$$

Let $v^b :=$ first column of B_N^b

$$= (b(\varphi_1), \dots, b(\varphi_N))^T$$

$$v^f := (f(\varphi_1), \dots, f(\varphi_N))^T$$

Now our theory for circulant matrices implies

$$B_N^b v^f = U_N^H \operatorname{diag}(\underset{\text{to account for } d\varphi \text{ in the conv}}{\Delta\varphi}\, W_N v^b) U_N v^f$$

$$= \underbrace{F_N^{-1} \operatorname{diag}\left(\sqrt{2\pi}\, F_N v^b\right) F_N v^f}_{\text{by (KF)}}$$

Example: Gaussian beams.

Recall that the Gaussian density on $\mathbb{R}^d$

$$b(x) = \prod_{i=1}^{d} \frac{\exp\left(-\left(\frac{1}{\sigma_b^2}\right) x_i^2 / 2\right)}{\sqrt{2\pi}}$$

has characteristic function $\phi_b(k)$ given by

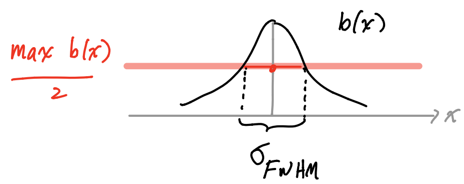
$$\phi^b(k) = \prod_{i=1}^{d} \exp(-\sigma_b^2 k_i^2 / 2)$$

which corresponds to the continuous Fourier transform of b.

Indeed one can use $\phi^b(k)$ to model $F_N v^b$ (i.e. the diagonal of B_N^b in the Fourier basis). In particular an approx Gaussian beam in 1-d is given by

$$\sqrt{2\pi}\, F_N v^b \approx \exp\left(-\sigma_b^2 (\underset{\text{vector of effective frequencies}}{\tau^{nyq}})^2 / 2\right)$$

Note: Often these beams are parametrized by σ_{FWHM} = "Full width at half max"



The conversion is: $\sigma_{FWHM}^2 = \sigma_b^2\, 8 \ln(2)$

Full model: Pixel to Spectral

High level continuum pixel model

$$\underbrace{d(\varphi)}_{\text{data}} = \underbrace{B}_{\substack{\text{beam,}\\ \text{i.e. convolution,}\\ \text{operator}}} \underbrace{f(\varphi)}_{\text{signal}} + \underbrace{\varepsilon(\varphi)}_{\text{white noise}}$$

$f \sim GRF_{S^1}(0, K)$ w/autocor $K(\varphi)$

$var(\varepsilon(\varphi_n)) = \sigma_\varepsilon^2$ when $\Delta\varphi = 1$

B is a beam operator with kernel $b(\varphi)$

Discrete pixel model specific to grid

For grid $\varphi_1, \ldots, \varphi_N \in [0, 2\pi)$ with $\varphi_n := \underbrace{(n-1)\frac{2\pi}{N}}$
the discretized fields are given by

grid spacing
$\Delta\varphi = \frac{\text{Period}}{N} = \frac{2\pi}{N}$

$$v^d = B_N^b v^f + v^\varepsilon$$

$$v^\varepsilon \sim N(0, \Sigma_N^{\varepsilon\varepsilon})$$

$$v^f \sim N(0, \Sigma_N^{ff})$$

where

$$B_N^b = \left[b(\varphi_i - \varphi_j)\Delta\varphi\right]_{i,j}$$

$$\Sigma_N^{\varepsilon\varepsilon} = \frac{\sigma_\varepsilon^2}{\Delta\varphi} I$$

$$\Sigma_N^{ff} = \left[cov(f(\varphi_i), f(\varphi_j))\right]_{ij}$$

$$v^f := (f(\varphi_1), \ldots, f(\varphi_N))^T$$

$$v^d := (d(\varphi_1), \ldots, d(\varphi_N))^T$$

$$v^\varepsilon := (\varepsilon(\varphi_1), \ldots, \varepsilon(\varphi_N))^T$$

$$v^b := (b(\varphi_1), \ldots, b(\varphi_N))^T$$

The corresponding discrete spectral model is given as follows

$$z^d = diag(\sqrt{2\pi}\, z^b) z^f + z^\varepsilon$$

where

$$z^d = \mathcal{F}_N v^d$$

$$z^f = \mathcal{F}_N v^f$$

$$z^\varepsilon = \mathcal{F}_N v^\varepsilon$$

$$z^b = \mathcal{F}_N v^b$$

$$E\left((z^\varepsilon)(z^\varepsilon)^H\right) = \frac{\sigma_\varepsilon^2}{\Delta\tau} I$$

$$E\left((z^f)(z^f)^H\right) = \frac{1}{\Delta\tau} diag\left(\sqrt{2\pi}\,\mathcal{F}_N \Sigma_N^{ff}[:,1]\right)$$

$$\Delta\tau := \frac{2\pi}{\text{period}} = 1$$

Multiple periodic Gaussian vectors

Consider the case of two random vectors

$$X = (X_1, \dots, X_N)^T \sim N(0, \Sigma^{XX})$$
$$Y = (Y_1, \dots, Y_N) \sim N(0, \Sigma^{YY})$$

jointly Gaussian s.t. the f.d.d of the random matrix $[X\ Y]$ is column circulant shift invariant, i.e.

$$\begin{bmatrix} X_1 & Y_1 \\ \vdots & \vdots \\ X_N & Y_N \end{bmatrix} \overset{\mathcal{D}}{=} \begin{bmatrix} X_N & Y_N \\ X_1 & Y_1 \\ & \\ X_{N-1} & Y_{N-1} \end{bmatrix} = P_N \begin{bmatrix} X_1 & Y_1 \\ \vdots & \vdots \\ X_N & Y_N \end{bmatrix}$$

Let $\Sigma = \mathrm{var}\left(\begin{bmatrix} X \\ Y \end{bmatrix}\right)$ be the full covariance matrix of the vector containing the X_i's & Y_i's.

$E(XY^T)$

$$\therefore \quad \Sigma = \begin{bmatrix} \Sigma^{XX} & \Sigma^{XY} \\ \Sigma^{YX} & \Sigma^{YY} \end{bmatrix}$$

Notice that the column shift invariance implies the 2x2 sub-blocks of Σ are circulant: i.e

$$\underbrace{\begin{bmatrix} P_N \Sigma^{XX} & P_N \Sigma^{XY} \\ P_N \Sigma^{YX} & P_N \Sigma^{YY} \end{bmatrix} = \begin{bmatrix} \Sigma^{XX} P_N & \Sigma^{XY} P_N \\ \Sigma^{YX} P_N & \Sigma^{YY} P_N \end{bmatrix}}$$

can be written compactly as

$$(I_2 \otimes P_N)\Sigma = \Sigma (I_2 \otimes P_N)$$

where

$$A \otimes B = \begin{bmatrix} A_{11} B & \cdots & A_{1m} B \\ \vdots & & \\ A_{m1} B & \cdots & A_{mm} B \end{bmatrix}$$

so that $I_2 \otimes P_N = \begin{bmatrix} P_N & 0 \\ 0 & P_N \end{bmatrix}$.

Therefore by our results on circulant matrices we have

$$\Sigma = \begin{bmatrix} U_N^H \Lambda^{XX} U_N & U_N^H \Lambda^{YX} U_N \\ U_N^H \Lambda^{YX} U_N & U^H \Lambda^{YY} U_N \end{bmatrix}$$

$$= (I_2 \otimes U_N^H) \begin{bmatrix} \Lambda^{XX} & \Lambda^{YX} \\ \Lambda^{YX} & \Lambda^{YY} \end{bmatrix} (I_2 \otimes U_N)$$

where

$AB = XX, XY, YX$, or YY

$$\Lambda^{AB} = \mathrm{diag}\left(W_N \Sigma^{AB}[:,1]\right)$$

The interpretation of the diagonal entries of $\Lambda^{XX}, \Lambda^{YY}, \Lambda^{XY}$ & Λ^{YX} is

$$E(z_n^A \bar{z}_m^B) = \begin{cases} \Lambda^{AB}_{n,n} & n = m \\ 0 & n \neq m \end{cases}$$

where $z^X = U_N X$, $z^Y = U_N Y$

i.e. the only dependence in

$$\begin{bmatrix} z_1^X & z_1^Y \\ \vdots & \vdots \\ z_N^X & z_N^Y \end{bmatrix}$$

is pairwise within each row.

If we want to fully diagonalize the dependence, we can go one step further and diagonalize each 2×2 matrix block.

Due to $\Sigma^H = \Sigma$ we have that Λ^{XX} & Λ^{YY} are real and $\Lambda^{YX} = \overline{\Lambda^{XY}}$.

Now we can use formulas for 2×2 Hermitian matrices:

$$\underbrace{\begin{bmatrix} 1 & re^{-i\theta} \\ re^{i\theta} & 1 \end{bmatrix}}_{\substack{\in \mathbb{C}^{2\times 2} \\ r \leq 1}} = \begin{bmatrix} e^{-i\theta} & -e^{-i\theta} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1-r & \\ & 1+r \end{bmatrix} \begin{bmatrix} e^{i\theta} & 1 \\ -e^{i\theta} & 1 \end{bmatrix}$$

Setting

$$r_n = \frac{|\Lambda^{YX}_{n,n}|}{\sqrt{\Lambda^{XX}_{n,n}}\sqrt{\Lambda^{YY}_{n,n}}}$$

θ_n = phase angle of $\Lambda^{YX}_{n,n}$

we have

$$\underbrace{\begin{bmatrix} e^{i\theta_n} & 1 \\ -e^{i\theta_n} & 1 \end{bmatrix} \begin{bmatrix} \Lambda^{XX}_{nn} & \\ & \Lambda^{YY}_{n,n} \end{bmatrix}^{-1/2} \begin{bmatrix} \Lambda^{XX}_{n,n} & \Lambda^{XY}_{n,n} \\ \Lambda^{YX}_{n,n} & \Lambda^{YY}_{n,n} \end{bmatrix} \begin{bmatrix} \Lambda^{XX}_{nn} & \\ & \Lambda^{YY}_{n,n} \end{bmatrix}^{-1/2} \begin{bmatrix} e^{-i\theta_n} & -e^{-i\theta_n} \\ 1 & 1 \end{bmatrix}}_{= \begin{bmatrix} 1-r_n & \\ & 1+r_n \end{bmatrix}}$$

In particular

$$\tilde{Z}^+_n := \frac{Z^Y_n}{\sqrt{\Lambda^{YY}_{n,n}}} + e^{i\theta_n} \frac{Z^X_n}{\sqrt{\Lambda^{XX}_{n,n}}}$$

$$\tilde{Z}^-_n := \frac{Z^Y_n}{\sqrt{\Lambda^{YY}_{n,n}}} - e^{i\theta_n} \frac{Z^X_n}{\sqrt{\Lambda^{XX}_{n,n}}}$$

are independent with

$$E|\tilde{Z}^+_n|^2 = 1 - r_n$$

$$E|\tilde{Z}^-_n|^2 = 1 + r_n$$

Side Comment

The above derivation is a particular case of a broader technique for constraining eigenspaces by finding commuting operators. The basic story is that the distributional symmetry

$$I_2 \otimes P_N \begin{bmatrix} X \\ Y \end{bmatrix} \overset{d}{=} \begin{bmatrix} X \\ Y \end{bmatrix}$$

implies

$$(I_2 \otimes P_N) \Sigma (I_2 \otimes P_N)^T = \Sigma$$

Also recall that $P_N^T = P_N^{-1}$ so

$$(I_2 \otimes P_N) \Sigma (I_2 \otimes P_N)^{-1} = \Sigma$$

i.e.

$$(I_2 \otimes P_N) \Sigma = \Sigma (I_2 \otimes P_N) \qquad (a)$$

Notice the following thm

<u>Theorem:</u> Commuting Hermitian matrices share a common eigenbasis.

This doesn't immediately apply since $I_2 \otimes P_N$ isn't Hermitian.

However notice that (a) also implies

$$(I_2 \otimes P_N)^H \Sigma = \Sigma (I_2 \otimes P_N)^H$$

which gives:

$$S_N \Sigma = \Sigma S_N$$

where

$$S_N = \frac{1}{2}\left[(I_2 \otimes P_N) + (I_2 \otimes P_N^H)\right]$$

Now we can apply the "Commuting Hermitian operator Theorem" to conclude that

Σ and S_N share a common eigenbasis

So we need to find the eigenbasis for S_N. Note that if

$$A = \mathcal{U} \Lambda \mathcal{U}^H \text{ and } B = V \Xi V^H$$

Then

$$A \otimes B = \underbrace{(\mathcal{U}^H \otimes V^H)}_{(\mathcal{U} \otimes V)^H} (\Lambda \otimes \Xi)(\mathcal{U} \otimes V)$$

Also note that $(A \otimes B)^T = (A^T \otimes B^T)$

Now recall

$$P_N = \mathcal{U}_N^H \operatorname{diag}(w_N^0, \ldots, w_N^{N-1}) \mathcal{U}_N$$

so that

$$I_2 \otimes P_N = (I_2 \otimes \mathcal{U}_N^H)(I_2 \otimes \operatorname{diag}(w_N^0, \ldots, w_N^{N-1}))(I_2 \otimes \mathcal{U}_N)$$

and

$$S_N = (I_2 \otimes \mathcal{U}_N^H) \underbrace{\left(I_2 \otimes \operatorname{diag}\left(\tfrac{w_N^0 + \bar{w}_N^0}{2}, \ldots, \tfrac{w_N^{N-1} + \bar{w}_N^{N-1}}{2}\right)\right)}_{\overset{\shortparallel}{I_2 \otimes \operatorname{diag}(Re(w_N^0), \ldots, Re(w_N^{N-1}))}} (I_2 \otimes \mathcal{U}_N)$$

Since the eigenvalues of S_N come in duplicate pairs

$$\underbrace{Re(w_N^0), Re(w_N^0)}, \ \ldots \ \underbrace{Re(w_N^{N-1}), Re(w_N^{N-1})}$$

the eigenvector basis for S_N is not unique and can be "rotated" on each 2-dim eigenspace to match the eigenvectors of Σ.

In a bit more detail, the "Commuting Hermitian operator Theorem" says that for each $n \in \{1, 2, \ldots, N\}$ we can find common eigenvectors

$$\varphi_1, \ldots, \varphi_N, \ \psi_1, \ldots, \psi_N \qquad (*)$$

such that

- $\varphi_n, \psi_n \in \mathbb{V}_n$
- $\mathbb{V}_n = \text{span}\left\{ \begin{bmatrix}1\\0\end{bmatrix} \otimes \bar{u}_{N,n}, \ \begin{bmatrix}0\\1\end{bmatrix} \otimes \bar{u}_{N,n} \right\}$

eigen vectors from I_2

eigen vector from S_N... i.e. the n^{th} column of U_N

- $\mathbb{V}_n$ is the eigenspace corresponding to eigenvalue $Re(w_N^{n-1})$

Remark: Since $w_N^{n-1} = e^{-i\frac{2\pi}{N}(n-1)}$,

$$Re(w_N^{n-1}) = \cos\left(\tfrac{2\pi}{N}(n-1)\right)$$

so that

$$\sum_{n=1}^{N} v_n \, \text{diag}\left(Re(w_N^0), \ldots, Re(w_N^{N-1})\right)^{n-1}$$

$$\shortparallel$$

the cosine transform of $(v_1, \ldots, v_n)^T$