It turns out that the spatial fluctuations of $T(\vec{x})$ are very well described by Gaussian Random Fields.

GRF intro

Let R be some region in $\mathbb{R}^d$.

e.g. $R = [0,1]^2 \subset \mathbb{R}^2$

e.g. $R = S^2 := \{x \in \mathbb{R}^3 : \|x\| = 1\}$

e.g. $R = \mathbb{R}$.

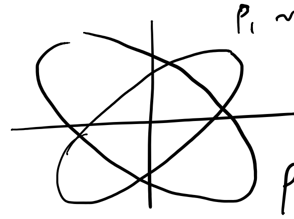
$f(x) : R \to \mathbb{R}$ denotes a function mapping R into $\mathbb{R}$.

Informally a GRF on R is a random function $f(x) : R \to \mathbb{R}$ with the property that all finite dimensional distributions (fdd) are multivariate Gaussian, i.e. that $\forall n \geq 1$, $\forall x_1, \dots, x_n \in R$ the random vector $(f(x_1), \dots, f(x_n))^T$ is jointly Gaussian.

Note: A random vector $(X_1, \dots, X_n)^T$ can have all marginals X_i Gaussian but not be jointly Gaussian.

Here is an example

$p_1 \sim N\left(0, \begin{pmatrix} 1 & \frac{1}{2} \\ & 1 \end{pmatrix}\right)$



$p_2 \sim N\left(0, \begin{pmatrix} 1 & -\frac{1}{2} \\ & 1 \end{pmatrix}\right)$

The density $\frac{1}{2}(p_1 + p_2)$ has $N(0,1)$ marginals but is not jointly Gaussian.

Thm: $X \in \mathbb{R}^n$ is jointly Gaussian iff $\alpha^T X$ is univariate Gaussian $\forall \alpha \in \mathbb{R}^n$.

Proof: use characteristic functions.

It is sometimes more natural to think of a GRF on R as a collection of univariate R.V.s indexed by R.

$$\{ f(x) : x \in R \}$$

For a fixed x, $f(x)$ is just a R.V.

Need that all these R.V.s are defined on the same probability space $(\Omega, \mathcal{F}, P)$.

i.e. the R.V.s $f(x)$ are actually a function of x & ω

$f(x) \equiv f(x, \omega)$.

2 ways to Model GRF's

First way: Direct map level construction.

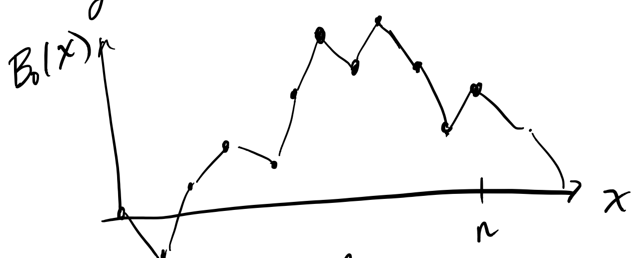
e.g. Let $\xi_1, \xi_2, \dots \xi_n$ be iid $N(0,1)$ & $\lambda_1, \dots, \lambda_n \in \mathbb{R}^d$ be non-random..

random coefs.

Then $f(x) := \sum_{i=1}^{n} \xi_i \underbrace{\cos(\langle \lambda_i, x_i \rangle)}_{\text{basis fns}}$

is a GRF on $\mathbb{R}^d$

e.g. Construct Brownian Motion by limiting scaled Random walks.



where $B_0(n) = \sum_{i=1}^{n} Z_i$ &

$Z_1, \dots$ iid $N(0,1)$.

$\therefore$ $B(x)$ is defined as the $L_2(\Omega, \mathcal{F}, P)$ limit $\lim_{k \to \infty} \frac{1}{\sqrt{k}} B_0(kx)$.

Second way: Implicitly by specifying the mean function $\mu(x)$, the covariance function $C(x,y)$ and postulating the existence of a GRF $\{f(x) : x \in R\}$ which satisfies

$$E(f(x)) = \mu(x)$$

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

Does such an $f(x)$ always exist?

Yes if $C(x,y) : R \times R \to R$ is positive definite.

Def: $C(x,y)$ is positive definite on $R \times R$ if $\forall n \geq 1$, $\forall x_1, \dots, x_n \in R$, $\forall b_1, \dots, b_n \in \mathbb{R}$ one has:

i) $\sum_{i,j=1}^{n} b_i b_j C(x_i, x_j) \geq 0$

ii) $C(x,y) = C(y,x)$, $\forall x, y \in R$.

Note that i) is required since

$$LHS = \mathrm{var}\left(\sum_{i=1}^{n} b_i f(x_i)\right) \geq 0.$$

e.g. $C(x,y) := \min(x,y)$ is p.d. on $[0,\infty)$

$\therefore$ $\exists$ a GRF $B(x)$ on $x \in [0,\infty)$ s.t.

$$E(B(x)) = 0$$

$$\mathrm{cov}(B(x), B(y)) = C(x,y)$$

This is the exact same Brownian motion we constructed before.

Under either construction a (separable) GRF is completely characterized by $\mu(x)$ & $C(x,y)$.

$\therefore$ the fdd satisfy

$$\begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(x_1) \\ \vdots \\ \mu(x_n) \end{pmatrix}, \begin{array}{cc} & j \\ & \downarrow \\ i \to & \begin{pmatrix} C(x_i, x_j) \end{pmatrix} \end{array} \right)$$

This gives you a way to simulate predict & do likelihood inference when working on a finite number of observation points $x_1, \dots, x_n \in R$ without worrying about the rest of R.

e.g. Simulating a GRF f with a given $\mu(x)$ & $C(x,y)$ at points $x_1, \dots, x_n \in R$.

Let $\Sigma = \left(C(x_i, x_j) \right)_{i,j=1}^n$

$\vec{\mu} = \left(\mu(x_i) \right)_{i=1}^n$

$\vec{f} = \left(f(x_i) \right)_{i=1}^n$

Cholesky $\Sigma = L L^T$, L is lower triangular

SVD $\Sigma = U \Lambda U^T$, U is orthonormal, Λ is diag with pos. entries

follows from p.d. of Σ

simulate $\vec{f}$ by

$$\vec{f} := L \underbrace{\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}}_{\text{iid } N(0,1)} + \vec{\mu}$$

or by $\vec{f} := U \Lambda^{1/2} U^T \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} + \vec{\mu}$

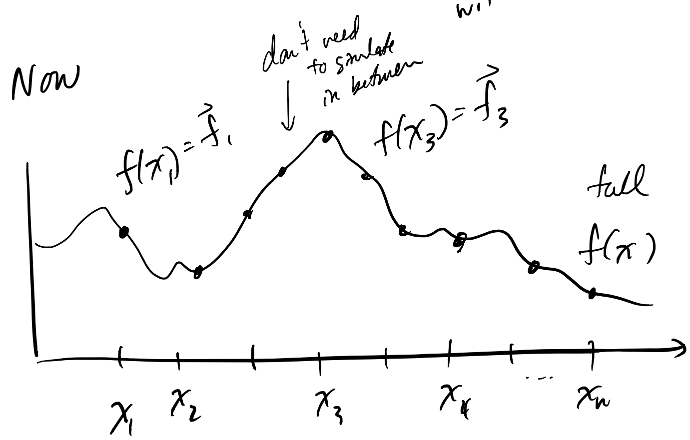
$$\overset{\mathcal{D}}{=} U \Lambda^{1/2} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} + \vec{\mu}$$

$$= \sum_{i=1}^n u_i \, \lambda_i^{1/2} z_i + \vec{\mu}$$

eigen basis of Σ

indep random coeffs with var $= \lambda_i$.

Now



Physics perspective:

$$\{ T(x) : x \in \mathbb{R}^3 \}$$

is predicted, under standard models, to be a GRF... note the Randomness & Gaussianity is predicted via Quantum mechanics

Stationary & Isotropic GRFs

Def: For two RFs f, g write

$$\{f(x) : x \in R\} \stackrel{\mathcal{D}}{=} \{g(x) : x \in R\}$$

if the fdds are the same, i.e.

$$(f(x_1), \ldots, f(x_n))^T \stackrel{\mathcal{D}}{=} (g(x_1), \ldots, g(x_n))^T$$

$\forall n \geq 1, \forall x_1, \ldots, x_n \in R$.

Def: A RF f on $\mathbb{R}^d$ is **stationary** if

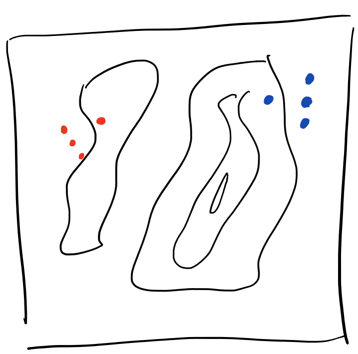
$$\{f(x) : x \in \mathbb{R}^d\} \stackrel{\mathcal{D}}{=} \{f(x+v) : x \in \mathbb{R}^d\}$$

$\forall v \in \mathbb{R}^d$ i.e. distributionally translation invariant.

Def: A RF f on $\mathbb{R}^d$ is **isotropic** if

$$\{f(x) : x \in \mathbb{R}^d\} \stackrel{\mathcal{D}}{=} \{f(\mathcal{U}x + v) : x \in \mathbb{R}^d\}$$

$\forall$ orthogal matrix $\mathcal{U}$ & $v \in \mathbb{R}^d$.



Contours of isotropic $f : \mathbb{R}^d \to \mathbb{R}$

- f evaluated at red dots has the same dist as f evaluated at blue dots.

Thm: Let $\mu(x)$ & $C(x,y)$ be the mean & cov fun for a GRF $f(x)$ on $\mathbb{R}^d$.

f is stationary $\Longrightarrow$ $C(x,y) = C(x-y)$, $\mu(x) = c \in \mathbb{R}$

f is isotropic $\Longrightarrow$ $C(x,y) = C(|x-y|)$, $\mu(x) = c \in \mathbb{R}$

$C(x-y)$ or $C(|x-y|)$ are called **auto covariance functions** in this case.

Spectral Representation for stationary GRF

Thm (Bochner): If f is a stationary GRF on $\mathbb{R}^d$ with $\mu(x) = 0$ iff $\exists\, C_k^{ff}$ mapping $k \in \mathbb{R}^d$ to $\mathbb{R}^+$ s.t.

$$\{f(x) : x \in \mathbb{R}^d\} \stackrel{\mathcal{D}}{=} \left\{ \int_{\mathbb{R}^d} e^{ix \cdot k} \sqrt{C_k^{ff}}\, W_k \frac{dk}{(2\pi)^{d/2}} \right\}$$

where W_k is complex white noise with unit variance.

Remark 1

usually we will just write

$$f(x) = \int e^{ix \cdot k} \underbrace{\sqrt{C_k^{ff}}\, W_k}_{} \frac{dk}{(2\pi)^{d/2}}$$

Think of this as the Fourier transform of $f(x)$.

Remark 2

For an ordinary (non random) function $g(x)$

let $$g_k := \int e^{-ix \cdot k} g(x) \frac{dx}{(2\pi)^{d/2}} = \text{Fourier transform}$$

$$g(x) = \int e^{ix \cdot k} g_k \frac{dk}{(2\pi)^{d/2}} = \text{inverse Fourier transform}$$

we'll Not worry about when these transforms exist & when $\mathcal{F}^{-1} = I\mathcal{F}$ (generally need $L_2(\mathbb{R}^d)$ or $L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$)

Remark 3.

Think of f_k as the random coeffs in a basis expansion of $f(x)$ where k indexes the basis elements:

$$f(x) \text{ "=" } \sum_k \underbrace{f_k}_{\text{random coeffs } \sqrt{C^{ff}_k}\, W_k} \underbrace{u_k(x)}_{\text{basis funs } e^{ix\cdot k}\frac{dk}{(2\pi)^{d/2}}}$$

Recall the SVD simulation of $\vec f = (f(x_1), \ldots, f(x_n))^T$

$$= \underbrace{U \Lambda U^T}_{(c(x_i - x_j))_{ij}} \vec Z$$

$\lambda_i \geq 0$ by pos. def cov mat.

$$= \sum_i \underbrace{\lambda_i^{1/2} z_i}_{\text{random coeffs } \overset{\text{indep}}{\sim} N(0, \lambda_i)} \vec u_i$$

basis eigenvectors of $(c(x_i - x_j))_{ij}$

$\therefore$ think of $e^{ix\cdot k} \frac{dk}{(2\pi)^{d/2}}$ as basis "eigenvectors" of $C(x-y)$ & $\sqrt{C^{ff}_k}\, W_k$ as indep coeffs $\sim N(0, C^{ff}_k)$.

$C^{ff}_k \geq 0$ by pos. def. cov fun (Bochners' Thm)

The Physics way of characterizing a mean zero stationary GRF is

$$E(f_k \bar f_{k'}) = \delta_{k-k'}\, C^{ff}_k$$

dirac delta

Sometimes written $\langle f_k f^*_{k'} \rangle = \delta_{k-k'} C^{ff}_k$

physics notation for expected value.

Heuristically:

$$E(f_k \bar f_{k'}) = E\left(W_k \sqrt{C^{ff}_k}\, \overline{W_{k'} \sqrt{C^{ff}_{k'}}} \right)$$

$$= \sqrt{C^{ff}_k} \sqrt{C^{ff}_{k'}} \underbrace{E(W_k \bar W_{k'})}_{\delta_{k-k'}}$$

only non-zero when $k = k'$

$$= C^{ff}_k\, \delta_{k-k'}$$

Also

$$E(f_k \bar f_{k'}) = E\left(\int e^{-ix\cdot k} f(x) \frac{dx}{(2\pi)^{d/2}} \overline{\int e^{-iy\cdot k'} f(y) \frac{dy}{(2\pi)^{d/2}}} \right)$$

$$= \iint e^{-ix\cdot k + iyk'} \underbrace{E(f(x) \overline{f(y)})}_{C(x-y)} \frac{dx}{(2\pi)^{d/2}} \frac{dy}{(2\pi)^{d/2}}$$

$\begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} x-y \\ y \end{pmatrix}$, $dz\,dy = dx\,dy$

$$= \iint e^{-i(z+y)k + iyk'} C(z) \frac{dz}{(2\pi)^{d/2}} \frac{dy}{(2\pi)^{d/2}}$$

$$= \underbrace{\int e^{-iz\cdot k} C(z) \frac{dz}{(2\pi)^{d/2}}}_{= C_k \text{ i.e. the fourier transform of the a.c.f } C(x)} \underbrace{\int e^{-iy(k-k')} \frac{dy}{(2\pi)^{d/2}}}_{= 2\pi^{d/2} \delta_{k-k'} \; (*)}$$

where (*) holds since

$$\int g_k \left[\int e^{iy(k-k')} \frac{dy}{(2\pi)^d} \right] dk = \int \underbrace{\left(\int e^{iyk} g_k \frac{dk}{(2\pi)^{d/2}} \right)}_{g(y)} e^{-iy\cdot k'} \frac{dy}{(2\pi)^{d/2}}$$

$$= \int g(y) e^{-iy\cdot k'} \frac{dy}{(2\pi)^{d/2}}$$

$$= g_{k'}$$

$$\therefore \quad E(f_k \overline{f}_{k'}) = C_k^{ff} \delta_{k-k'}$$

$$= (2\pi)^{d/2} C_k \delta_{k-k'}$$

Def: For a mean zero stationary GRF f, C_k^{ff} is called the spectral density & equals $(2\pi)^{d/2} C_k$ where

$$\mathrm{cov}(f(x), f(y)) = C(x-y).$$

Remark 4.

If $f(x)$ is also isotropic then $C_k^{ff} = C_{|k|}^{ff}$ i.e. C_k^{ff} is rotationally symmetric about the origin.

White Noise & the dirac delta

Think of two different dirac delta functions:

1) δ_k defined in Fourier space $k \in \mathbb{R}^d$.

$$\int_{\mathbb{R}^d} \delta_k g_k \, dk = g_0$$

2) $\delta(x)$ defined in "pixel space" $x \in \mathbb{R}^d$

$$\int_{\mathbb{R}^d} \delta(x) g(x) dx = g(0)$$

Heuristically:

$$\delta_k = I_{\{k=0\}} \frac{1}{dk} = \begin{cases} 0 & \text{if } k \neq 0 \\ \frac{1}{dk} & \text{if } k = 0 \end{cases}$$

$$\delta(x) = I_{\{x=0\}} \frac{1}{dx} = \begin{cases} 0 & \text{if } x \neq 0 \\ \frac{1}{dx} & \text{if } x = 0 \end{cases}$$

where dx & dk are infinitesimal area elements & $I_{\{\dots\}}$ is an indicator.

Real Gaussian white Noise in pixel space $W(x)$, $x \in \mathbb{R}^d$ with unit variance statisfies

$$E\big(W(x)W(y)\big) = \delta(x-y).$$

$$E\big(W_k \overline{W}_{k'}\big) = \delta_{k-k'}.$$

This is very "Non rigourous" but can be made rigourous.

Since $\delta(x-y) = I_{x=y} \dfrac{1}{dx}$

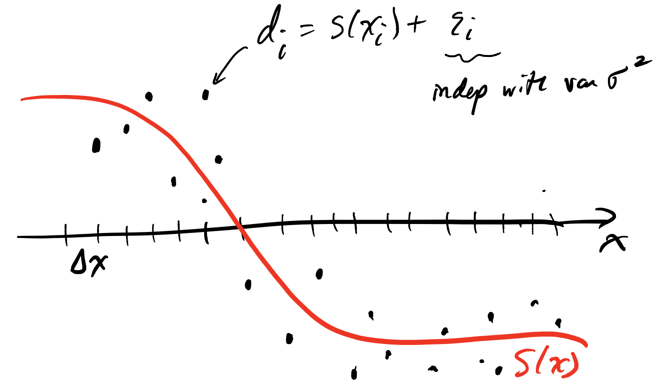
You can think of $W(x)$ as $W(x) = Z(x) \cdot \dfrac{1}{\sqrt{dx}}$

when at each x, $Z(x)$ is an independent $N(0,1)$ R.V.

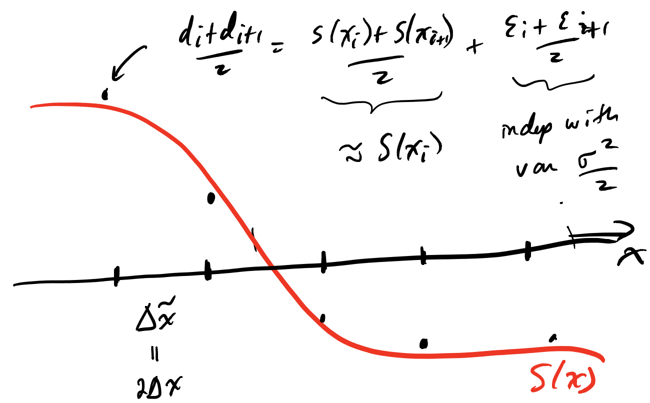
If you observe $W(x)$ on a grid then $\dfrac{1}{\sqrt{dx}} = \dfrac{1}{\sqrt{\text{grid pixel area}}}$.

The brilliant thing about $W(x)$ is that it gives "grid invariant" quantifation of information or signal-to-noise ratio.

e.g. suppose you observe a stationary GRF $S(x)$ (i.e. the signal) with iid $N(0,\sigma^2)$ Noise on a grid with area Δx



Now consider constructing a new data set $\tilde{d}$ obtained by averaging 2 pixels.



If $S(x)$ is sufficiently smooth Both of the above inferences (one based on d, the other on $\tilde{d}$) should be about the same.

i.e. the information content is the same (i.e. SNR) but the Noise level changes.

if we define $\varepsilon(x) = \sigma \underbrace{W(x)}_{\text{white Noise}}$

then $E(\varepsilon(x)\varepsilon(y)) = \sigma^2 \delta(x-y)$

$$= \sigma^2 \frac{I_{\{x=y\}}}{dx}$$

This suggest we can think of

$\varepsilon(x_1), \varepsilon(x_2), \ldots$

grid points with grid area Δx.

as iid $N\left(0, \frac{\sigma^2}{\Delta x}\right)$

This scales just like in the last exaple.

the variace gets smaller as $\Delta x \uparrow$.

Now I can write my observations

$$(\bullet) \quad d(x) = s(x) + \varepsilon(x)$$

which adapts the Noise level to whichever grid I decide to measure it on.

Another way to think about $(\bullet)$ is that it expresses the amont of information $d(x)$ carries for $s(x)$ in a observation grid independent way.