(Gaussian) random fields

statistics and data analysis (chapter 12)

Björn Malte Schäfer

Graduate School for Fundamental Physics
Fakultät für Physik und Astronomie, Universität Heidelberg

June 26, 2016

outline

random fields

- consider a randomly fluctuating smooth field where the amplitude φ at position y depends on that at position x
- description with a probability density

$$p(\varphi(y), \varphi(x)) = p(\varphi(y)|\varphi(x)) \tag{1}$$

- examples
  - noise in an electric circuit
  - waves on the ocean surface
  - matter distribution in the universe

Gaussian random fields

- Gaussian probability density for the field amplitudes

$$p(\varphi(x),\varphi(y)) = \frac{1}{\sqrt{(2\pi)^2 \det C}} \exp\left(-\frac{1}{2}\begin{pmatrix}\varphi(x)\\ \varphi(y)\end{pmatrix}^t C^{-1}\begin{pmatrix}\varphi(x)\\ \varphi(y)\end{pmatrix}\right) \tag{2}$$

- with covariance

$$C = \begin{pmatrix}\langle\varphi(x)^2\rangle & \langle\varphi(x)\varphi(y)\rangle\\ \langle\varphi(x)\varphi(y)\rangle & \langle\varphi(y)^y\rangle\end{pmatrix} \tag{3}$$

- where $\langle\ldots\rangle$ are ensemble averages with fixed positions x and y
- correlation function

$$\xi(x,y) = \langle\varphi(x)\varphi(y)\rangle \tag{4}$$

- obviously, $\langle\varphi(x)\varphi(y)\rangle \to \langle\varphi(x)^2\rangle$ as $x \to y$
- think of $\langle\varphi(x)\varphi(y)\rangle$ as the memory of the field on its value at x

homogeneity

- Cauchy-Schwarz bound:

$$|\xi(x,y)| \leq \sqrt{\langle\varphi(x)^2\rangle\,\langle\varphi(y)^2\rangle} \quad (5)$$

  making sure that $\det C \geq 0$

- separation of the Gaussian distribution occurs if $\langle\varphi(x)\varphi(y)\rangle = 0$
- **homogeneity of the random field**: statistical properties are identical everywhere, i.e. the correlation function does not depend on x, just on r:

$$\langle\varphi(x)\varphi(x+r)\rangle = \xi(r) \quad (6)$$

- for homogeneous fields, the Cauchy-Schwarz bound becomes:

$$|\xi(r)| \leq \langle\varphi(x)^2\rangle = \sigma^2 \quad (7)$$

  with a universal $\sigma^2 = \langle\varphi(x)^2\rangle = \langle\varphi(y)^2\rangle$

homogeneity

- homogeneous fields: $\sigma^2 = \langle\varphi(x)^2\rangle = \langle\varphi(y)^2\rangle$ and $\xi(r) = \langle\varphi(x)\varphi(x+r)\rangle$
- probability density

$$p(\varphi(x),\varphi(x+r)) = \frac{1}{\sqrt{(2\pi)^2\det C}}\exp\left(-\frac{1}{2}\left(\begin{array}{c}\varphi(x)\\ \varphi(x+r)\end{array}\right)^t C^{-1}\left(\begin{array}{c}\varphi(x)\\ \varphi(x+r)\end{array}\right)\right) \tag{8}$$

- with covariance

$$C = \left(\begin{array}{cc}\sigma^2 & \xi(r)\\ \xi(r) & \sigma^2\end{array}\right) \tag{9}$$

correlation functions and spectra

- spectrum of a random field: **only** for homogeneous fields
- Fourier-transform

$$\varphi(k) = \int \mathrm{d}x\, \varphi(x) \exp(-\mathrm{i}kx) \quad \leftrightarrow \quad \varphi(x) = \int \mathrm{d}k\, \varphi(k) \exp(+\mathrm{i}kx) \tag{10}$$

- is there a correlation between different Fourier-modes?

$$\langle \varphi(k)\varphi(k')^* \rangle = \int \mathrm{d}x \int \mathrm{d}x' \,\langle \varphi(x)\varphi(x') \rangle \exp(-\mathrm{i}kx + \mathrm{i}k'x') \tag{11}$$

- assume homogeneity: $x' = x + r$, $\mathrm{d}x' = \mathrm{d}r$ at fixed x:

$$= \int \mathrm{d}x \int \mathrm{d}r\, \langle \varphi(x)\varphi(x+r) \rangle \exp(-\mathrm{i}kx + \mathrm{i}k'x + \mathrm{i}k'r) \tag{12}$$

correlation functions and spectra

- correlation function $\xi(r) = \langle\varphi(x)\varphi(x+r)\rangle$:

$$= \int \mathrm{d}r\, \xi(r) \exp(\mathrm{i}k'r) \int \mathrm{d}x\ \exp(-\mathrm{i}(k-k')x) \tag{13}$$

  if the correlation function does not depend on position x

- $\mathrm{d}x$-integration yields the Dirac-δ function

$$= \int \mathrm{d}r\, \xi(r) \exp(\mathrm{i}kr) \times 2\pi\delta_D(k-k') \tag{14}$$

- define **spectrum**:

$$C(k) = \int \mathrm{d}r\, \xi(r) \exp(\mathrm{i}kr) \tag{15}$$

  as the Fourier-transform of the correlation function

- no correlation between Fourier-modes of a homogeneous random field:

$$\langle\varphi(k)\varphi(k')\rangle = 2\pi C(k)\, \delta_D(k-k') \tag{16}$$

random fields in more dimensions: isotropy

- derivation of the spectrum is identical in n dimensions:

$$C(k_i) = \int \mathrm{d}^n r \, \xi(r_i) \exp(\mathrm{i} k_j r_j) \tag{17}$$

- yields an identical orthogonality relation

$$\langle \varphi(k_i) \varphi(k_i') \rangle = (2\pi)^n C(k_i) \, \delta_D(k_i - k_i')) \tag{18}$$

- if the field is isotropic, the spectra for different k_i are identical, so C depends only on the modulus $\sqrt{k_i k_i}$

units of spectra and correlation functions

- assume that φ has the unit A, which is some combination of length, time, mass and temperature
- correlation function
$$\xi(r) = \langle\varphi(x)\varphi(x+r)\rangle \tag{19}$$
as units of A^2
- then, $\varphi(k)$ has units of AL^n in n dimensions
- the Dirac-δ function has units of L^n, because $\int \mathrm{d}^n x\, \delta_D(x) = 1$
- collect all results:
$$\langle \underbrace{\varphi(k)}_{AL^n} \underbrace{\varphi(k')}_{AL^n} \rangle \propto C(k) \underbrace{\delta_D(k-k')}_{L^n} \tag{20}$$
such that $C(k)$ needs to have units of A^2L^n

operations on random fields: smoothing

- smoothing: by convolution with a convolution kernel p:

$$\bar{\varphi}(x) = \int \mathrm{d}y\, \varphi(y)p(x-y) = \int \mathrm{d}y\, \varphi(x-y)p(y) = \varphi * p \tag{21}$$

- slide p over the field φ and average around x weighted by p
- $p(x)$ is normalised, $\int \mathrm{d}x\, p(x) = 1$
- the property

$$\int \mathrm{d}y\, \varphi(y)\delta_D(x-y) = \varphi(x) \tag{22}$$

  is defined through convolution, in particular with $\int \mathrm{d}y\, \delta_D(y) = 1$
- properties of convolutions
  - $\varphi * p = p * \varphi$
  - $\varphi * (p * q) = (\varphi * p) * q$
  - $\varphi * (\alpha p + \beta q) = \alpha\varphi * p + \beta\varphi * q$

convolutions in Fourier-space

- convolutions in real space are products in Fourier-space:

$$\varphi * p = \int \mathrm{d}y\, \varphi(x-y)p(y) \tag{23}$$

$$= \int \mathrm{d}y \int \mathrm{d}k\varphi(k)\exp(\mathrm{i}k(x-y)) \int \mathrm{d}k'\, p(k')\exp(\mathrm{i}k'y) \tag{24}$$

$$= \int \mathrm{d}y \int \mathrm{d}k\, \varphi(k)p(k') \int \mathrm{d}k' \exp(\mathrm{i}kx)\exp(\mathrm{i}(k-k')y) \tag{25}$$

$$= \int \mathrm{d}k \int \mathrm{d}k'\, \varphi(k)p(k')\exp(\mathrm{i}kx)\delta_D(k-k') \tag{26}$$

$$= \int \mathrm{d}k\, \varphi(k)p(k')\exp(\mathrm{i}kx) \tag{27}$$

- Fourier-transform is symmetric: convolutions in Fourier-space are products in real space

correlation

- correlationn with **lag** y:

$$\varphi \times \varphi = \int \mathrm{d}x\, \varphi(x)\varphi(x+y) \tag{28}$$

- estimate for an **ergodic** process:

$$\varphi \times \varphi \simeq \lim_{y\to\infty} \frac{1}{y} \int_0^y \mathrm{d}x\, \varphi(x)\varphi(x+y) \tag{29}$$

- correlations in Fourier-space

$$\varphi \times \varphi = \int \mathrm{d}x\, \varphi(x)\varphi(x+y) \tag{30}$$

$$= \int \mathrm{d}x \int \mathrm{d}k\, \varphi(k) \exp(\mathrm{i}kx) \int \mathrm{d}k'\, \varphi(k') \exp(\mathrm{i}k'(x+y)) \tag{31}$$

$$= \int \mathrm{d}k \int \mathrm{d}k'\, \varphi(k)\varphi(k')^* \exp(\mathrm{i}(k-k')x) \exp(-\mathrm{i}k'y) \tag{32}$$

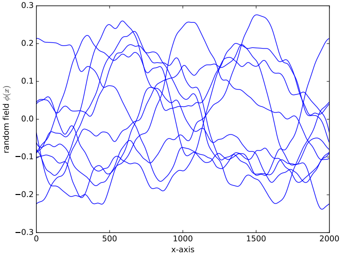
$$= \int \mathrm{d}k\, \varphi(k)\varphi(k)^* \exp(\mathrm{i}ky) \tag{33}$$

generation of a (homogeneous) random field

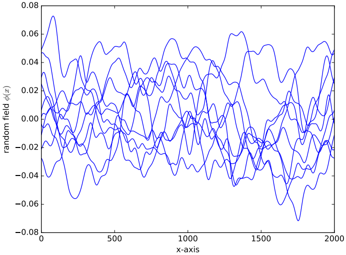
- use the fact that Fourier-modes of homogeneous random fields are independent
- and that the variance in Fourier-space is the spectrum $C(k)$
- algorithm:
  - draw Gaussian random numbers for $\varphi(k)$ on a grid in Fourier-space
  - from a distribution with $\langle\varphi(k)\rangle = 0$ and $\langle\varphi(k)^2\rangle = C(k)$
  - impose hermiticity: $\varphi(k)^* = \varphi(-k)$ for a real-valued field
  - Fourier-transform into real space

$$\varphi(x) = \int \frac{\mathrm{d}k}{2\pi}\,\varphi(k)\exp(+\mathrm{i}kx) \tag{34}$$

examples of Gaussian random fields



$C(k) \propto k^{-2}$

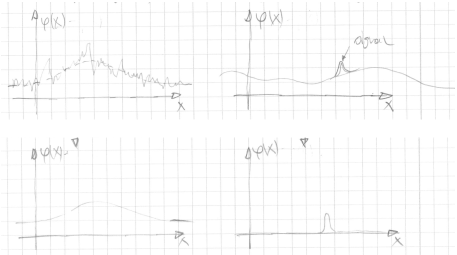


$C(k) \propto k^{-1}$

- 10 realisations of Gaussian random fields (different phases)

matched filtering: idea

- matched filtering: find weak signal **with known shape** in noisy data:
  - high-frequency noise on top of a slowly varying signal: low-pass
  - small-scale signal on top of a slowly varying background: high-pass



- should be possible if signal and noise look differently in Fourier-space, two competiting effects: lowering of the noise φ vs. dispersing the signal a

construction of a matched filter

- filter p should operate by convolution on a field $\varphi(x)$ where the signal is hidden

$$\varphi'(x) = \int \mathrm{d}y\, \varphi(y)p(x-y) \quad \leftrightarrow \quad \varphi'(k) = p(k)\varphi(k) \tag{35}$$

- require
  - the convolved field should on average conserve the amplitude of the signal:

$$\langle\varphi'\rangle = A\int \mathrm{d}k\, a(k)p(k) \tag{36}$$

    with the signal strength A and the signal profile a(x)
  - noise fluctuations σ^2 are minimised

$$\langle(\varphi' - \langle\varphi'\rangle)^2\rangle = \sigma^2 = \int \mathrm{d}k\, p^2(k)C(k) \tag{37}$$

derivation by variation

- $\sigma^2[p]$ is a functional which maps the function p onto a number
- it's possible to optimise $\sigma^2[p]$ for being as small as possible, because then, the signal stands out
- one needs an additional assumption: clearly, the smallest σ^2 is obtained by setting $p = 0$, but then the signal is lost as well
- require unbiasedness

$$\langle\varphi'\rangle = A\int \mathrm{d}k\, a(k)p(k) = \mu \tag{38}$$

  when performing the variation for minimising the noise
- introduce unbiasedness with a Lagrange-multiplier:

$$\sigma^2[p] \rightarrow \sigma^2[p] + \lambda G[p] \quad \text{with} \quad G[p] = A\int \mathrm{d}k\, a(k)p(k) - \mu \tag{39}$$

derivation by variation

- variation:

$$\frac{\mathrm{d}\sigma^2}{\mathrm{d}p} + \lambda\frac{\mathrm{d}G}{\mathrm{d}p} = 0 = \int \mathrm{d}k\ \left(2pC(k) + \lambda Aa\right) \quad \rightarrow \quad p = -\frac{\lambda}{2}\frac{a(k)}{C(k)} \quad (40)$$

  if one assumes a signal profile $a(k)$ and a noise spectrum $C(k)$
- λ is determined by the boundary conditions
- if the true profile or the noise spectrum are different, the signal is not enhanced

deconvolution

- convolution is an averaging process, so it is not possible to recover the orginial field φ from the convolved field $\varphi' = \varphi * p$
- it is **not possible** to divide out the filter in the Fourier-expression $\varphi' = \varphi p \rightarrow \varphi = \varphi'/p$ because the Fourier-transform of the filter is zero somewhere
- deconvolution can only be done approximatively

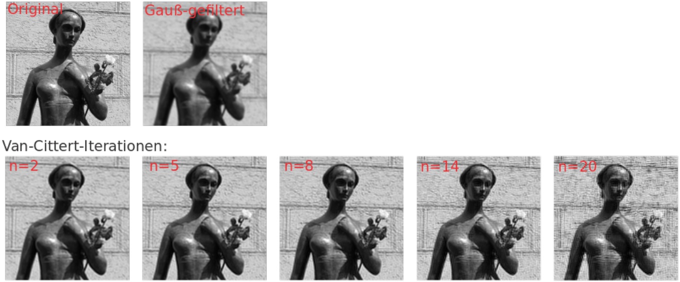
van Cittert deconvolution

- idea: control noise by iterative partial deconvolution, and stop at the right moment
- $\varphi = \varphi' p^{-1}$ inverts $\varphi' = \varphi p$
- introduce $\bar{p} = 1 - p$,
- use the geometric series:

$$\varphi = \frac{\varphi'}{p} = \frac{\varphi'}{1-\bar{p}} \simeq \varphi' \left(1 + \bar{p} + \bar{p}^2 + \ldots\right) = \varphi' \left(1 + \bar{p}(1 + \bar{p}(\ldots))\right) \quad (41)$$

  and rewrite it as a ”telescopic” sums
- then, define an iterative procedure:
  - $\varphi_0 = \varphi'$ starting point is the noisy signal
  - $\varphi_{n+1} = \varphi' + (1-p)\varphi_n$

van Cittert deconvolution



van Cittert deconvolution (source: wikipedia)

- picks up artefacts if n is too large

summary

- random fields are a generalisation to random variables, they're indexed by position
- correlation of field amplitudes given by correlation function
- homogeneous random fields have uncorrelated Fourier-modes
- convolutions and correlations can be carried out in Fourier-space
- filtering lets you search for a signal of known shape
- deconvolution is possible in an approximate way