Gaussian vectors

Gaussian vector

• A random vector $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ is Gaussian iff any linear combination has normal distribution.

• A Gaussian random vector $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ (write $\boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$) has characteristic function

$$\mathbb{E}[e^{i\boldsymbol{\lambda}^T \boldsymbol{Z}}] = \int_{\mathbb{R}^d} e^{i\boldsymbol{\lambda}^T \boldsymbol{z}} p(\boldsymbol{z}) d\boldsymbol{z} = \exp\Big(i\boldsymbol{\lambda}^T \boldsymbol{\mu} - \frac{\boldsymbol{\lambda}^T \mathbf{C} \boldsymbol{\lambda}}{2}\Big), \qquad \boldsymbol{\lambda} \in \mathbb{R}^n$$

• A Gaussian random vector $\boldsymbol{Z} = (Z_1, \dots, Z_n)$ with mean $\boldsymbol{\mu}$ and *invertible* covariance matrix $\mathbf{C}$ has the pdf

$$p(\boldsymbol{z}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \mathbf{C}}} \exp\Big(-\frac{(\boldsymbol{z}-\boldsymbol{\mu})^T \mathbf{C}^{-1}(\boldsymbol{z}-\boldsymbol{\mu})}{2}\Big)$$

• The expectations of high-order moments of a zero-mean Gaussian vector can be expressed as a sum of second-order moments.
For instance, if $\boldsymbol{Z} = (Z_1, Z_2, Z_3, Z_4)$ is a zero-mean Gaussian vector, then

$$\mathbb{E}\Big[\prod_{j=1}^{4} Z_j\Big] = \mathbb{E}\big[Z_1 Z_2\big]\mathbb{E}\big[Z_3 Z_4\big] + \mathbb{E}\big[Z_1 Z_3\big]\mathbb{E}\big[Z_2 Z_4\big] + \mathbb{E}\big[Z_1 Z_4\big]\mathbb{E}\big[Z_2 Z_3\big]$$

Gaussian conditioning

Let $\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}$ be a Gaussian random vector (with $\boldsymbol{X}_1$ of size n_1 and $\boldsymbol{X}_2$ of size n_2):

$$\mathcal{L}\left(\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \right) = \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} \right)$$

with the means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ of sizes n_1 and n_2, the covariance matrices $\mathbf{R}_{11}$ of size $n_1 \times n_1$, $\mathbf{R}_{12}$ of size $n_1 \times n_2$, $\mathbf{R}_{21} = \mathbf{R}_{12}^T$ of size $n_2 \times n_1$, and $\mathbf{R}_{22}$ of size $n_2 \times n_2$. Then the distribution of $\boldsymbol{X}_1$ conditionally to $\boldsymbol{X}_2$ is Gaussian:

$$\mathcal{L}\big(\boldsymbol{X}_1 | \boldsymbol{X}_2 = \boldsymbol{x}_2\big) = \mathcal{N}\big(\boldsymbol{\mu}_1 + \mathbf{R}_{12}\mathbf{R}_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2), \mathbf{R}_{11} - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}\big)$$

Proof: The posterior pdf of $\boldsymbol{X}_1$ given that $\boldsymbol{X}_2 = \boldsymbol{x}_2$ is $p_{\rm post}(\boldsymbol{x}_1) = \dfrac{p(\boldsymbol{x}_1, \boldsymbol{x}_2)}{\int p(\boldsymbol{x}_1', \boldsymbol{x}_2) d\boldsymbol{x}_1'}$.

Use

$$\begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{Q}^{-1} & -\mathbf{Q}^{-1}\mathbf{R}_{12}\mathbf{R}_{22}^{-1} \\ -\mathbf{R}_{22}^{-1}\mathbf{R}_{21}\mathbf{Q}^{-1} & \mathbf{R}_{22}^{-1} + \mathbf{R}_{22}^{-1}\mathbf{R}_{21}\mathbf{Q}^{-1}\mathbf{R}_{12}\mathbf{R}_{22}^{-1} \end{pmatrix},$$

where $\mathbf{Q} = \mathbf{R}_{11} - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}$ is the Schur complement.

Example. Let us consider the Gaussian vector $(n_1 = n_2 = 1)$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

with $\rho \in [-1, 1]$ correlation coefficient of X_1 and X_2.

The distribution of X_1 is

$$\mathcal{L}\big(X_1\big) = \mathcal{N}\big(0, 1\big)$$

If ones observes $X_2 = x_2$, then:

$$\mathcal{L}\big(X_1 | X_2 = x_2\big) = \mathcal{N}\big(\rho x_2, 1 - \rho^2\big)$$

- the mean of X_1 is attracted (if $\rho > 0$) or repeled (if $\rho < 0$) by the observation.
- the variance of X_1 is reduced (if $\rho \neq 0$).

Extreme cases: $\rho = 0$ (one learns nothing from the observation of X_2) and $\rho = 1$ (ones knows everything once X_2 is observed).

Limit theorems

Limit theorems for sums (1/2)

Let $(X_n)_{n\geq 0}$ be a sequence of independent and identically distributed (i.i.d.) random variables (square integrable). The empirical mean is the random variable:

$$\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$$

Its mean is $\mathbb{E}[\bar{X}_n] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mathbb{E}[X_1]$, and its variance is

$$\mathrm{Var}(\bar{X}_n) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{\sigma^2}{n}$$

with $\sigma^2 = \mathrm{Var}(X_1) = \cdots = \mathrm{Var}(X_n)$. This quantity goes to 0 as $n \to \infty$, which means that $\bar{X}_n$ concentrates on the deterministic value $\mathbb{E}[X_1]$:

$$\mathbb{P}\left(|\bar{X}_n - \mathbb{E}[X_1]| \geq \varepsilon\right) \leq \frac{\mathbb{E}[|\bar{X}_n - \mathbb{E}[X_1]|^2]}{\varepsilon^2} = \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon}$$

- (Strong) law of large numbers.

Let $(X_n)_{n\geq 0}$ be a sequence of independent and identically distributed random variables (square integrable). The empirical mean $\bar{X}_n$ converges to $\mathbb{E}[X_1]$ with probability one:

$$\bar{X}_n \overset{n\to\infty}{\longrightarrow} \mathbb{E}[X_1] \quad a.s. \qquad \text{or } \mathbb{P}\left(\lim_{n\to\infty} \bar{X}_n = \mathbb{E}[X_1]\right) = 1$$

Limit theorems for sums (2/2)

- Central limit theorem

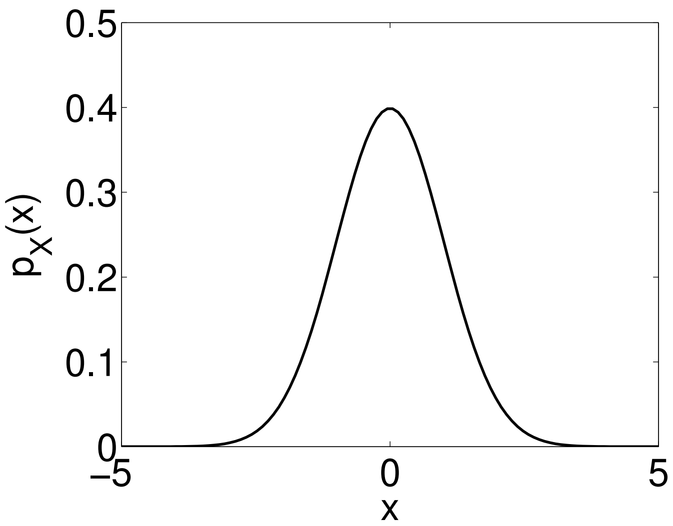
Let $(X_n)_{n\geq 0}$ be a sequence of independent and identically distributed random variables (square integrable), with the mean $\mu \in \mathbb{R}$ and the variance σ^2, $\sigma \in (0,\infty)$. Then

$$\sqrt{n}(\bar{X}_n - \mu) \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0,\sigma^2)$$

This means that, for any interval $I \subset \mathbb{R}$,

$$\mathbb{P}\left(\sqrt{n}(\bar{X}_n - \mu) \in I\right) \overset{n\to\infty}{\longrightarrow} \int_I p_{0,\sigma^2}(x)dx$$

where $p_{0,\sigma^2}(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/(2\sigma^2))$.



Limit theorems for maxima (1/2)

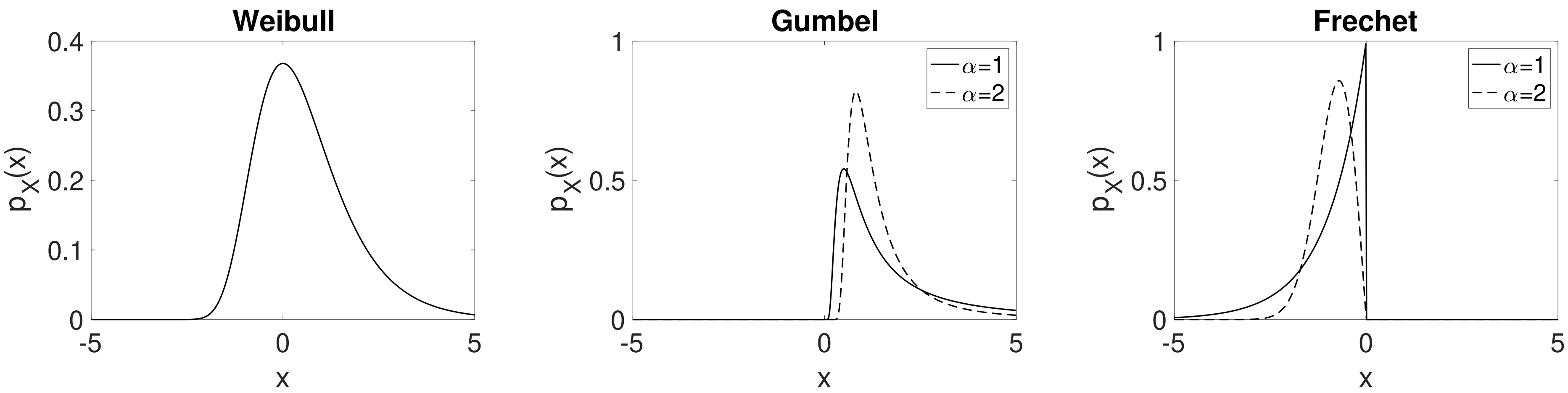
Let $M_n = \max(X_1, \ldots, X_n)$ where the X_i are i.i.d. with pdf $p(x)$.

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(X_1 \leq x, \ldots, X_n \leq x) = \mathbb{P}(X \leq x)^n = \left(1 - \int_x^{\infty} p(s)ds\right)^n$$

Theorem (Fisher-Tippett-Gnedenko): Assume that there exist a_n and b_n such that $\mathbb{P}(M_n \leq a_n + b_n x)$ converges to a cdf $F(x)$. Then $F(x)$ belongs to one of the three following types (up to affine scaling):

- $F_1(x) = e^{-e^{-x}}$ with support $\mathbb{R}$ (Gumbel).
- $F_{2,\alpha}(x) = e^{-1/x^{\alpha}}$ with support $\mathbb{R}^+$, $\alpha > 0$ (Weibull).
- $F_{3,\alpha}(x) = e^{-(-x)^{\alpha}}$ with support $\mathbb{R}^-$, $\alpha > 0$ (Fréchet).

Conclusion: As $n \to \infty$, $M_n = a_n + b_n Y$ where Y follows one of the three "extreme" distributions.



Limit theorems for maxima (2/2)

Main idea to determine the asymptotic distribution of $M_n = \max(X_1, \ldots, X_n)$:
1) Denote $x_{n,\alpha}$ such that $\int_{x_{n,\alpha}}^{\infty} p(s)ds = \frac{\alpha}{n}$. Then $\mathbb{P}(M_n \leq x_{n,\alpha}) = (1 - \frac{\alpha}{n}) \to e^{-\alpha}$.
2) Look for the expansion of $x_{n,\alpha}$ as $n \to \infty$.

Example 1: $X_i \sim \mathcal{U}(0,1)$, i.e. $p(x) = \mathbf{1}_{[0,1]}(x)$.
Then $a_n = 1$, $b_n = 1/n$, and $F = F_{3,1}$.

Example 2: $X_i \sim \mathcal{E}(1)$, i.e. $p(x) = e^{-x}\mathbf{1}_{[0,\infty)}(x)$.
Then $a_n = \ln(n)$, $b_n = 1$, and $F = F_1$.

Example 3: $X_i \sim \mathcal{N}(0,1)$, i.e. $p(x) = (2\pi)^{-1/2} e^{-x^2/2}$.
Then $a_n = \sqrt{2\ln(n)} - \ln\ln n/\sqrt{8\ln(n)} - \ln(4\pi)/\sqrt{8\ln(n)}$, $b_n = 1/\sqrt{2\ln(n)}$, and $F = F_1$.

Random processes

Random process

• Random variable Z = random number = application from a probability space to $\mathbb{R}$.
A realization of the random variable = a real number.
Distribution of a continuous random variable Z is characterized by its pdf $p_Z(z)$.
Example: $Z \sim \mathcal{N}(0,1) \mapsto p_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$.

• Random process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ = random function = application from a probability space to $\mathbb{R}^{\mathbb{R}^d}$.
A realization of the process = a function from $\mathbb{R}^d$ to $\mathbb{R}$.
Distribution of $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ characterized by the finite-dimensional distributions $(Z(\boldsymbol{x}_1),\ldots,Z(\boldsymbol{x}_n))$, for any n, $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \mathbb{R}^d$.

Gaussian process

- We say that a random process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is Gaussian if any linear combination $Z_\lambda = \sum_{i=1}^n \lambda_i Z(\boldsymbol{x}_i)$ has Gaussian distribution.

→ $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))$ is a Gaussian vector with

$$\mathbb{E}\big[e^{i\boldsymbol{\lambda}^T \boldsymbol{Z}}\big] = \exp\Big(i\boldsymbol{\lambda}^T\boldsymbol{\mu} - \frac{\boldsymbol{\lambda}^T\mathbf{C}\boldsymbol{\lambda}}{2}\Big)$$

with $\boldsymbol{\mu} = (\mathbb{E}[\mu(\boldsymbol{x}_j)])_{j=1}^n$ and $\mathbf{C} = (\mathbb{E}[Z(\boldsymbol{x}_i)Z(\boldsymbol{x}_j)] - \mathbb{E}[Z(\boldsymbol{x}_i)]\mathbb{E}[Z(\boldsymbol{x}_j)])_{i,j=1}^n$.

→ Z_λ has Gaussian distribution with pdf

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma_\lambda} \exp\Big(-\frac{(z-\mu_\lambda)^2}{2\sigma_\lambda^2}\Big)$$

where $\mu_\lambda = \sum_{i=1}^n \lambda_i \mathbb{E}[Z(\boldsymbol{x}_i)]$ and $\sigma_\lambda^2 = \sum_{i,j=1}^n \lambda_i\lambda_j \mathbb{E}[Z(\boldsymbol{x}_i)Z(\boldsymbol{x}_j)] - \mu_\lambda^2$, provided $\sigma_\lambda > 0$.

→ The first two moments $\mu(\boldsymbol{x}_1) = \mathbb{E}[Z(\boldsymbol{x}_1)]$ and $R(\boldsymbol{x}_1, \boldsymbol{x}_2) = \mathbb{E}[Z(\boldsymbol{x}_1)Z(\boldsymbol{x}_2)]$ characterize the finite-dimensional distribution of the process $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$.

→ The distribution of a Gaussian process is characterized by its first two moments $\mu(\boldsymbol{x}_1) = \mathbb{E}[Z(\boldsymbol{x}_1)]$ and $R(\boldsymbol{x}_1, \boldsymbol{x}_2) = \mathbb{E}[Z(\boldsymbol{x}_1)Z(\boldsymbol{x}_2)]$.

→ Any linear transform of a Gaussian process is a Gaussian process.

• Simulation: in order to simulate $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))$:

- compute the mean vector $M_i = \mathbb{E}[Z(\boldsymbol{x}_i)]$ and the covariance matrix $C_{ij} = \mathbb{E}[Z(\boldsymbol{x}_i)Z(\boldsymbol{x}_j)] - \mathbb{E}[Z(\boldsymbol{x}_i)]\mathbb{E}[Z(\boldsymbol{x}_j)]$.
- generate a random vector $\boldsymbol{G} = (G_1, \dots, G_n)$ of n independent Gaussian random variables with mean 0 and variance 1.
- compute $\boldsymbol{Z} = \boldsymbol{M} + \mathbf{C}^{1/2}\boldsymbol{G}$. The vector $\boldsymbol{Z}$ has the distribution of $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))$.

Note: the computation of the square root is expensive (use Cholesky method).

Proof. $\boldsymbol{Z}$ is a Gaussian vector.
The mean of $\boldsymbol{Z}$ is the mean of the vector $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))$. The covariance matrix of $\boldsymbol{Z}$ is the covariance matrix of the vector $(Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n))$.

Brownian motion

• Brownian motion $(W_t)_{t\geq 0}$= Gaussian process with mean 0 and covariance function

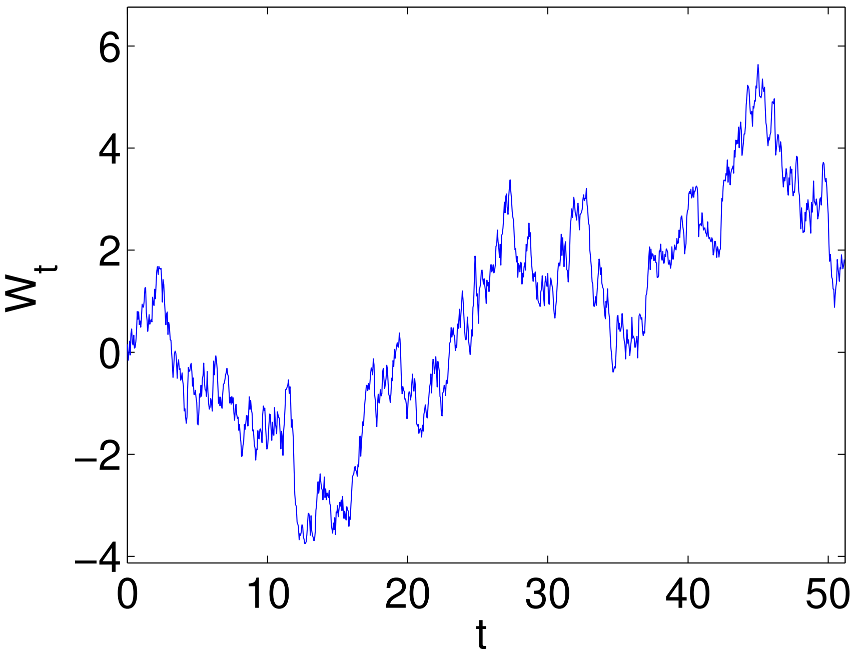
$$\mathbb{E}[W_t W_{t'}] = t \wedge t'$$

The increments of the Brownian motion are independent:
if $t_n \geq t_{n-1} \geq \cdots \geq t_1 \geq t_0 = 0$, then $(W_{t_n} - W_{t_{n-1}}, \ldots, W_{t_2} - W_{t_1}, W_{t_1})$ are independent Gaussian random variables with mean 0 and variance

$$\mathbb{E}\big[(W_{t_j} - W_{t_{j-1}})^2\big] = t_j - t_{j-1}$$

• Simulation: in order to simulate $(W_h, W_{2h}, ..., W_{nh})$, one can use the Cholesky method, but the following method is more rapid:
- generate a random vector $X = (X_1, ..., X_n)$ of n independent Gaussian random variables with mean 0 and variance 1.
- compute $Y_j = \sqrt{h}\sum_{i=1}^{j} X_i$. The vector $\boldsymbol{Y}$ has the distribution of $(W_h, W_{2h}, ..., W_{nh})$.

• Simulation: in order to simulate $(W_h, W_{2h}, ..., W_{nh})$:
- generate a random vector $\boldsymbol{G} = (G_1, ..., G_n)$ of n independent Gaussian random variables with mean 0 and variance 1.
- compute $Y_j = \sqrt{h} \sum_{i=1}^{j} G_i$. The vector $\boldsymbol{Y}$ has the distribution of $(W_h, W_{2h}, ..., W_{nh})$.



Stationary random process

• $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ is stationary if $(Z(\boldsymbol{x}+\boldsymbol{x}_0))_{\boldsymbol{x}\in\mathbb{R}^d}$ has the same distribution as $(Z(\boldsymbol{x}))_{\boldsymbol{x}\in\mathbb{R}^d}$ for any $\boldsymbol{x}_0 \in \mathbb{R}^d$.
Sufficient and necessary condition:

$$\mathbb{E}\big[\phi(Z(\boldsymbol{x}_1),\ldots,Z(\boldsymbol{x}_n))\big] = \mathbb{E}\big[\phi(Z(\boldsymbol{x}_0+\boldsymbol{x}_1),\ldots,Z(\boldsymbol{x}_0+\boldsymbol{x}_n))\big]$$

for any n, $\boldsymbol{x}_0,\ldots,\boldsymbol{x}_n \in \mathbb{R}^d$, $\phi \in \mathcal{C}_b(\mathbb{R}^n,\mathbb{R})$.

• Let $Z(\boldsymbol{x})$ be a stationary process (with finite second-order moments). Then its mean is constant and its covariance function $\mathrm{Cov}(Z(\boldsymbol{x}), Z(\boldsymbol{x}'))$ depends on $\boldsymbol{x}-\boldsymbol{x}'$ only.

• A Gaussian process $Z(\boldsymbol{x})$ is stationary iff its mean is constant and its covariance function $\mathrm{Cov}(Z(\boldsymbol{x}), Z(\boldsymbol{x}'))$ is of the form $c(\boldsymbol{x}-\boldsymbol{x}')$.
This function c is even, maximal at $\mathbf{0}$, and its Fourier transform is nonnegative.
• Bochner's theorem: a function $c(\boldsymbol{x})$ is a covariance function of a stationary process iff its Fourier transform is nonnegative.

Spectral representation of a real-valued stationary Gaussian process

$$Z(\boldsymbol{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \sqrt{\hat{c}(\boldsymbol{k})} \hat{n}_{\boldsymbol{k}} d\boldsymbol{k}$$

with $\hat{n}_{\boldsymbol{k}}$ complex white noise, i.e.:
$\hat{n}_{\boldsymbol{k}}$ complex-valued, Gaussian, $\hat{n}_{-\boldsymbol{k}} = \overline{\hat{n}_{\boldsymbol{k}}}$, $\mathbb{E}[\hat{n}_{\boldsymbol{k}}] = 0$, and $\mathbb{E}\left[\hat{n}_{\boldsymbol{k}} \overline{\hat{n}_{\boldsymbol{k}'}}\right] = (2\pi)^d \delta(\boldsymbol{k} - \boldsymbol{k}')$.
(the representation is formal, one should use stochastic integrals $d\hat{W}_{\boldsymbol{k}} = \hat{n}_{\boldsymbol{k}} d\boldsymbol{k}$).

We have $\hat{n}_{\boldsymbol{k}} = \int e^{i\boldsymbol{k}\cdot\boldsymbol{x}} n(\boldsymbol{x}) d\boldsymbol{x}$ where $n(\boldsymbol{x})$ is a real white noise, i.e.:
$n(\boldsymbol{x})$ real-valued, Gaussian, $\mathbb{E}[n(\boldsymbol{x})] = 0$, and $\mathbb{E}[n(\boldsymbol{x})n(\boldsymbol{x}')] = \delta(\boldsymbol{x} - \boldsymbol{x}')$.
(formal representation, one should use Brownian motions and fields;
in 1D, $n(x)dx = dW_x$, $\hat{n}(k)dk = dW_k^{(1)} + i dW_k^{(2)}$).

- Simulation ($d = 1$): in order to simulate $(Z(x_1), \ldots, Z(x_n))$, $x_j = (j-1)h$:

- compute the covariance vector $\boldsymbol{C} = (c(x_1), \ldots, c(x_n))$.
- generate a random vector $\boldsymbol{G} = (G_1, \ldots, G_n)$ of n independent Gaussian random variables with mean 0 and variance 1.
- filter with the square root of the Fourier transform of $\boldsymbol{C}$:

$$\boldsymbol{Z} = \mathrm{IFT}\big(\sqrt{\mathrm{DFT}(\boldsymbol{C})} \times \mathrm{DFT}(\boldsymbol{G})\big)$$

$\hookrightarrow$ $\boldsymbol{Z}$ is a realization of $(Z(x_1), \ldots, Z(x_n))$ (in practice, use FFT and IFFT).

Complement: Wiener integral

- Let W_t be a Brownian motion.

We want to define $I_f := \int_0^\infty f(t)dW_t$ for all $f \in L^2(0,+\infty)$.

- Let $\mathcal{S}$ be the set of simple functions. If $f \in \mathcal{S}$, i.e. $f(t) = \sum_{i=1}^n \alpha_i \mathbf{1}_{[a_i,a_{i+1})}(t)$, then

$$I_f := \sum_{i=1}^n \alpha_i \big(W_{a_{i+1}} - W_{a_i}\big)$$

- The application $f \in \mathcal{S} \mapsto I_f \in L^2(\Omega)$ is linear.
- $I_f \sim \mathcal{N}(0, \int_0^\infty f(t)^2 dt)$.
- Itô's isometry: $\mathbb{E}[I_f^2] = \displaystyle\int_0^\infty f(t)^2 dt$.
- Polar form: If $f, g \in L^2(0,\infty)$, then $\mathbb{E}[I_f I_g] = \int_0^\infty f(t)g(t)dt$.

- Extension to $L^2(0,\infty)$: If $f \in L^2(0,\infty)$, then $\exists f_n \in \mathcal{S}$ such that $f_n \to f$ in $L^2(0,\infty)$. We have

$$\mathbb{E}\big[(I_{f_n} - I_{f_p})^2\big] = \int_0^\infty (f_n - f_p)(t)^2 dt$$

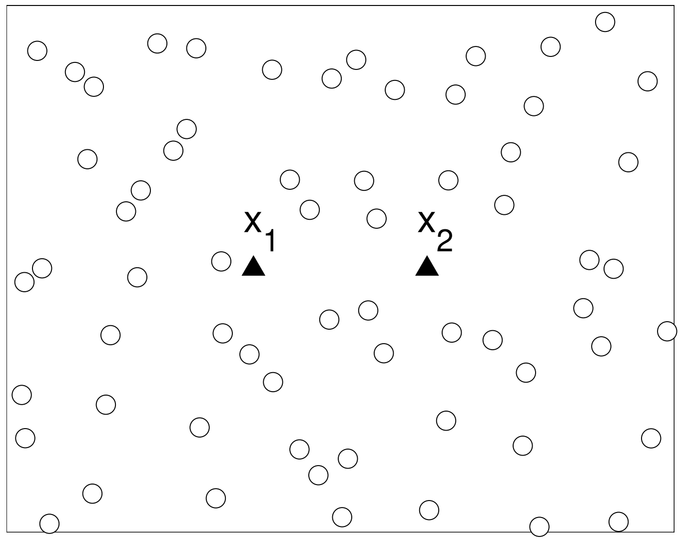
which shows that I_{f_n} is a Cauchy sequence in $L^2(\Omega)$, hence it converges to a limit denoted I_f. The limit still has the above properties.

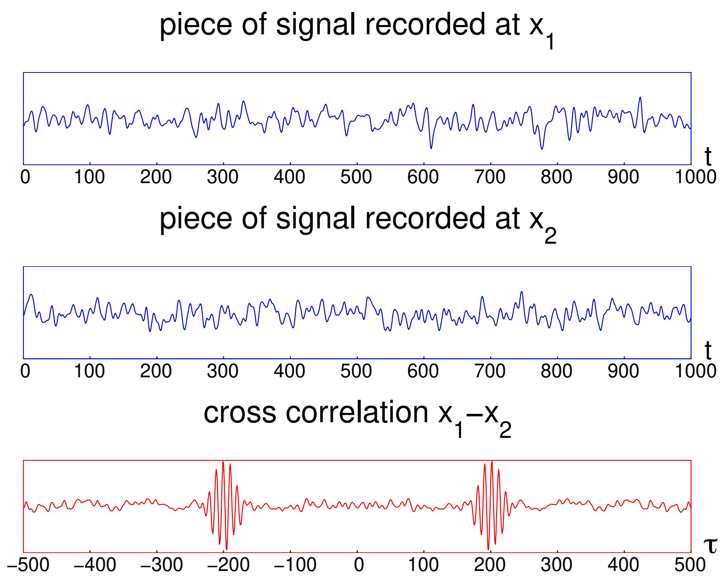
- If $f_i \in L^2(0,\infty)$, $i = 1,\ldots,n$, then $(I_{f_i})_{i=1}^n \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$, with $C_{jl} = \int_0^\infty f_j(t) f_l(t) dt$.

Passive imaging with noise sources

Green's function estimation by cross correlation of ambient noise signals

- Ambient noise sources (∘) emit stationary random signals.
- The waves propagate in the (inhomogeneous) medium.
- The signals $u(t, \boldsymbol{x}_1)$ and $u(t, \boldsymbol{x}_2)$ are recorded at two sensors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.





- Compute the empirical cross correlation:

$$C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{T}\int_0^T u(t, \boldsymbol{x}_1) u(t+\tau, \boldsymbol{x}_2) dt$$

- $C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)$ is related to the Green's function from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$!

Wave equation

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 u}{\partial t^2} - \Delta_{\boldsymbol{x}} u = n(t,\boldsymbol{x})$$

- Three-dimensional inhomogeneous medium (background velocity $c(\boldsymbol{x})$).

- Sources $n(t,\boldsymbol{x})$: Gaussian process, with mean zero, with covariance

$$\langle n(t_1,\boldsymbol{y}_1) n(t_2,\boldsymbol{y}_2)\rangle = F(t_2-t_1)\Gamma(\boldsymbol{y}_1,\boldsymbol{y}_2)$$

- Stationary in time: $\big(n(t,\boldsymbol{y})\big)_{t,\boldsymbol{y}}$ and $\big(n(t+h,\boldsymbol{y})\big)_{t,\boldsymbol{y}}$ have the same statistical distribution for any $h \Longrightarrow$ the time correlation function F depends only on t_2-t_1.

- The spatial distribution of the sources is characterized by $\Gamma(\boldsymbol{y}_1,\boldsymbol{y}_2)$. To simplify:

$$\Gamma(\boldsymbol{y}_1,\boldsymbol{y}_2) = K(\boldsymbol{y}_1)\delta(\boldsymbol{y}_1-\boldsymbol{y}_2)$$

The function K characterizes the spatial support of the sources.
[For more general Γ, use multiscale or microlocal analysis, see below.]

Remember: the Green's function

Solution u of the wave equation:

$$u(t,\boldsymbol{x}) = \int\int G(s,\boldsymbol{x},\boldsymbol{y}) n(t-s,\boldsymbol{y}) ds d\boldsymbol{y}$$

Green's function:

$$\frac{1}{c^2(\boldsymbol{x})}\frac{\partial^2 G}{\partial t^2} - \Delta_{\boldsymbol{x}} G = \delta(t)\delta(\boldsymbol{x}-\boldsymbol{y})$$

starting from $G(0,\boldsymbol{x},\boldsymbol{y}) = \partial_t G(0,\boldsymbol{x},\boldsymbol{y}) = 0 \quad (G(t,\boldsymbol{x},\boldsymbol{y}) = 0$ for $t<0)$.

Time-harmonic Green's function:

$$\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}) = \int G(t,\boldsymbol{x},\boldsymbol{y}) e^{i\omega t} dt$$

solution of

$$\frac{\omega^2}{c^2(\boldsymbol{x})}\hat{G} + \Delta_{\boldsymbol{x}}\hat{G} = -\delta(\boldsymbol{x}-\boldsymbol{y})$$

+ Sommerfeld radiation condition.

Example: If $c(\boldsymbol{x}) \equiv c_0$ (and in dimension 3)

$$G(t,\boldsymbol{x},\boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|}\delta\Big(\frac{|\boldsymbol{x}-\boldsymbol{y}|}{c_0} - t\Big)$$

$$\hat{G}(\omega,\boldsymbol{x},\boldsymbol{y}) = \frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{y}|} e^{i\omega\frac{|\boldsymbol{x}-\boldsymbol{y}|}{c_0}}$$

Empirical cross correlation and statistical cross correlation

Empirical cross correlation:

$$C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{T} \int_0^T u(t, \boldsymbol{x}_1) u(t+\tau, \boldsymbol{x}_2) dt$$

with $u(t, \boldsymbol{x}) = \iint G(s, \boldsymbol{x}, \boldsymbol{y}) n(t-s, \boldsymbol{y}) ds d\boldsymbol{y}$ and $G =$ causal time-dependent Green's function.

1. The expectation of C_T (with respect to the distribution of the sources) is independent of the integration time T:

$$\langle C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) \rangle = C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)$$

where the statistical cross correlation $C^{(1)}$ is given by

$$C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{2\pi} \int d\boldsymbol{y} \int d\omega \overline{\hat{G}}(\omega, \boldsymbol{x}_1, \boldsymbol{y}) \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}) K(\boldsymbol{y}) \hat{F}(\omega) e^{-i\omega\tau}$$

Proof. We have

$$u(t, \boldsymbol{x}) = \iint G(s, \boldsymbol{x}, \boldsymbol{y}) n(t - s, \boldsymbol{y}) ds d\boldsymbol{y}$$

By the stationarity of the process n, the product $u(t, \boldsymbol{x}_1) u(t + \tau, \boldsymbol{x}_2)$ is itself a stationary in t. Therefore the mean of C_T is independent of T:

$$\begin{aligned}
\langle C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)\rangle &= \langle u(0, \boldsymbol{x}_1) u(\tau, \boldsymbol{x}_2)\rangle \\
&= \iint d\boldsymbol{y}_1 d\boldsymbol{y}_2 \iint ds' ds G(s, \boldsymbol{x}_1, \boldsymbol{y}_1) G(s', \boldsymbol{x}_2, \boldsymbol{y}_2) \langle n(-s, \boldsymbol{y}_1) n(\tau - s', \boldsymbol{y}_2)\rangle \\
&= \iint d\boldsymbol{y}_1 d\boldsymbol{y}_2 \iint ds' ds G(s, \boldsymbol{x}_1, \boldsymbol{y}_1) G(s', \boldsymbol{x}_2, \boldsymbol{y}_2) F(\tau - s' + s) \Gamma(\boldsymbol{y}_1, \boldsymbol{y}_2) \,.
\end{aligned}$$

Using the spatial delta-correlation property:

$$\begin{aligned}
\langle C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)\rangle &= \int d\boldsymbol{y} \iint ds ds' G(-s, \boldsymbol{x}_1, \boldsymbol{y}) G(s', \boldsymbol{x}_2, \boldsymbol{y}) K(\boldsymbol{y}) F(\tau - s - s') \\
&= \frac{1}{2\pi} \int d\boldsymbol{y} \int d\omega \overline{\hat{G}}(\omega, \boldsymbol{x}_1, \boldsymbol{y}) \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}) K(\boldsymbol{y}) \hat{F}(\omega) e^{-i\omega\tau}
\end{aligned}$$

Empirical cross correlation and statistical cross correlation

Empirical cross correlation:

$$C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{T} \int_0^T u(t, \boldsymbol{x}_1) u(t+\tau, \boldsymbol{x}_2) dt$$

with $u(t, \boldsymbol{x}) = \iint G(s, \boldsymbol{x}, \boldsymbol{y}) n(t-s, \boldsymbol{y}) ds d\boldsymbol{y}$ and $G =$ causal time-dependent Green's function.

1. The expectation of C_T (with respect to the distribution of the sources) is independent of the integration time T:

$$\langle C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) \rangle = C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)$$

where the statistical cross correlation $C^{(1)}$ is given by

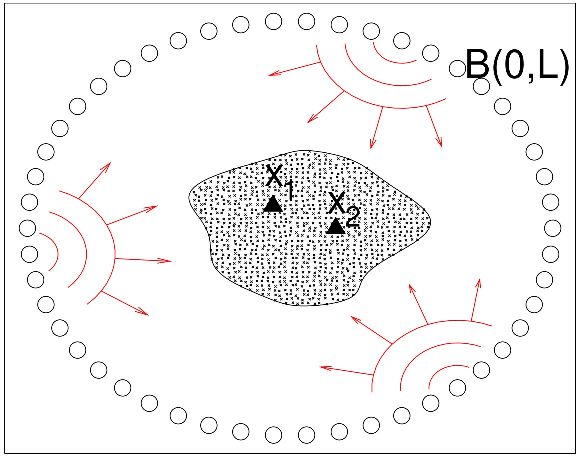
$$C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{2\pi} \int d\boldsymbol{y} \int d\omega \overline{\hat{G}}(\omega, \boldsymbol{x}_1, \boldsymbol{y}) \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}) K(\boldsymbol{y}) \hat{F}(\omega) e^{-i\omega\tau}$$

2. The empirical cross correlation is a self-averaging quantity:

$$C_T(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) \stackrel{T\to\infty}{\longrightarrow} C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2)$$

in probability.
More precisely, the fluctuations of C_T around its expectation $C^{(1)}$ are of order $T^{-1/2}$.

Emergence of the Green's function for an extended distribution of sources in an inhomogeneous open medium



Cross correlation with noise sources distributed on a closed surface $\partial B(\mathbf{0}, L)$:

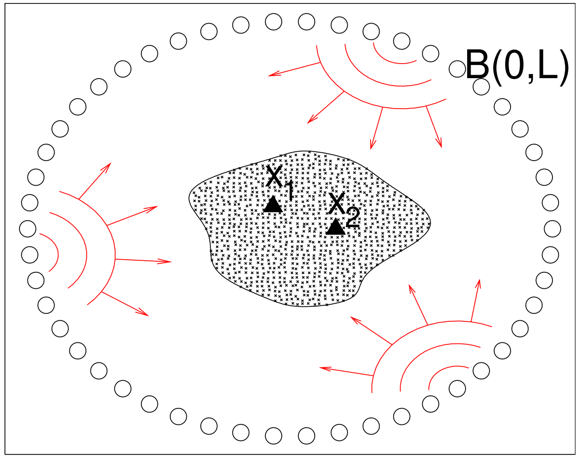
$$C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{2\pi} \int d\omega \int_{\partial B(\mathbf{0},L)} d\sigma(\boldsymbol{y}) \overline{\hat{G}}(\omega, \boldsymbol{x}_1, \boldsymbol{y}) \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y}) \hat{F}(\omega) e^{-i\omega\tau}$$

By Helmholtz-Kirchhoff identity,

$$\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2) - \overline{\hat{G}}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{2i\omega}{c_0} \int_{\partial B(\mathbf{0},L)} d\sigma(\boldsymbol{y}) \overline{\hat{G}}(\omega, \boldsymbol{x}_1, \boldsymbol{y}) \hat{G}(\omega, \boldsymbol{x}_2, \boldsymbol{y})$$

we have

$$C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{c_0}{2\pi} \int \frac{\hat{F}(\omega)}{\omega} \mathrm{Im}\big(\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2)\big) e^{-i\omega\tau} d\omega$$



$$\partial_\tau C^{(1)}(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) = -\frac{ic_0}{2\pi} \int \hat{F}(\omega) \mathrm{Im}\big(\hat{G}(\omega, \boldsymbol{x}_1, \boldsymbol{x}_2)\big) e^{-i\omega\tau} d\omega$$
$$= -\frac{c_0}{2} \Big(F *_\tau G(\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) - F *_\tau G(-\tau, \boldsymbol{x}_1, \boldsymbol{x}_2) \Big)$$

• The cross correlation of noise signals recorded by two passive sensors is related to the Green's function between the sensors.
↪ this is the classical proof that passive sensors can be transformed into virtual sources (known in seismology).

• This proof requires the noise sources to surround the region of interest.
The previous proofs also establish exact relations in ideal conditions.
Other proofs can justify approximate relations (enough for travel time estimation) in realistic conditions.