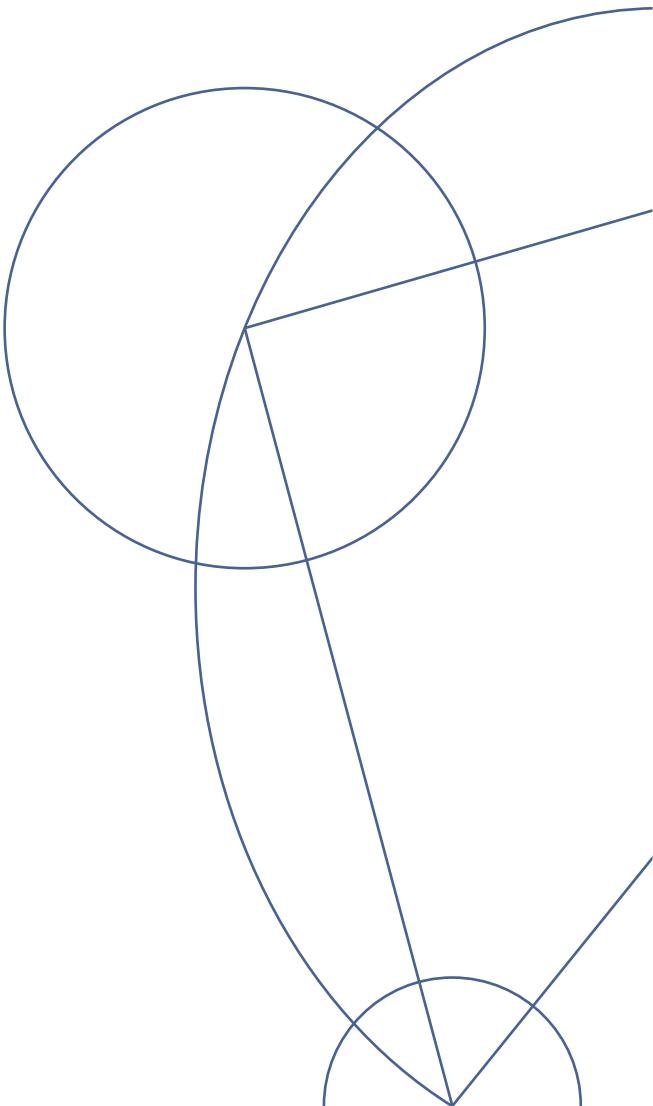
National Technical University of Athens
School of Electrical and Computer Engineering
Inter-Faculty Postgraduate Studies Programme
Data Science and Machine Learning

Stochastic Processes

The second assignment written in partial fulfillment of the requirements for the completion of the Data Driven Models in Engineering elective course of the Data Science & Machine Learning post-graduate studies programme.

Instructor
Prof. V. Papadopoulos

Written by
[redacted]



May 06, 2022

1 Stochastic Fields & the Karhunen-Loève Series Expansion

The present exercise deals with the stochastic field

$$E(x) = 10\,(1 + f(x))\,, \tag{1.1}$$

where $f(x)$ is a zero-mean, stationary Gaussian field with unit variance and domain $x \in [0, 5]$. Its autocorrelation function is given as

$$R_f(\tau) = \exp\left(-\frac{|\tau|}{2}\right) \tag{1.2}$$

and belongs to the broader family of exponential covariance kernels[1] of the form

$$C(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|}{b}\right), \tag{1.3}$$

for $b = 2$. The eigenvalue problem corresponding to such kernels is given by the integral equation

$$\int_{-a}^{a} C(x_i, x_j)\,\phi(x_j)\,dx_j = \lambda\phi(x_i)\,, \tag{1.4}$$

where $[-a, a]$ is the (symmetric) interval over which the process is defined and λ, $\phi(x)$ are the eigenvalues and eigenfunctions of the kernel, respectively. After introducing the variable

$$\omega^2 = \frac{2b - \lambda}{\lambda b^2}, \tag{1.5}$$

Eq. (1.4) leads to the transcendental equations

$$\frac{1}{b} - \omega_n \tan(\omega_n\, a) = 0 \quad \text{and} \quad \tilde{\omega}_n + \frac{1}{b}\tan(\tilde{\omega}_n\, a) = 0, \tag{1.6}$$

the solutions of which lie in the intervals $[(n-1)\,\pi/a, (n-0.5)\,\pi/a]$ and $[(n-0.5)\,\pi/a, n\pi/a]$, respectively [1], hence the $n \in \mathbb{N}$ index. The two families of solutions corresponding to Eqs. (1.6) lead to the two families of eigenvalues

$$\lambda_n = \frac{2b}{1 + \omega_n^2\, b^2} \quad \text{and} \quad \tilde{\lambda}_n = \frac{2b}{1 + \tilde{\omega}_n^2\, b^2}, \tag{1.7}$$

which are obtained by inverting Eq. (1.5), and two families of eigenfunctions given by [2]

$$\phi_n(x) = \frac{\cos(\omega_n\, x)}{\sqrt{a + \frac{\sin(2\omega_n\, a)}{2\omega_n}}} \quad \text{and} \quad \tilde{\phi}_n(x) = \frac{\sin(\tilde{\omega}_n\, x)}{\sqrt{a - \frac{\sin(2\tilde{\omega}_n\, a)}{2\tilde{\omega}_n}}}. \tag{1.8}$$

Note at this point that these results are obtained for symmetric domains of the form $[-a, a]$. In the case of non-symmetric domains such as the present one, the domain can first be transformed into a symmetric one via the change of variables $x \to x - a$ with $a = 2.5$. According to the Karhunen-Loève theorem, after the solutions of Eqs. (1.8), (1.9) are obtained, each realization of the stochastic field[2] can be expanded as

[1] The terms "autocorrelation" function and "covariance" function are used interchangeably herein, given the mean and variance of the Gaussian field $f(x)$.

[2] The dependency $\xi = \xi(\theta)$ is omitted for brevity. For the same reason, the convention $f(x, \theta) = f(x)$ is also followed.

$$f(x) = \sum_{n=1}^{\infty} \left[\sqrt{\lambda_n}\, \phi_n(x-a)\ \xi_n + \sqrt{\tilde{\lambda}_n}\, \tilde{\phi}_n(x-a)\ \tilde{\xi}_n \right], \tag{1.9}$$

where the sets $\{\xi_n\}$, $\left\{\tilde{\xi}_n\right\}$ correspond to random variables that follow the standardized Gaussian distribution. It is stressed here that the orthonormality relations

$$\int_{-a}^{a} \phi_n(x)\,\phi_m(x)\,dx = \delta_{nm}, \qquad \int_{-a}^{a} \phi_n(x)\,\tilde{\phi}_m(x)\,dx = 0, \qquad \int_{-a}^{a} \tilde{\phi}_n(x)\,\tilde{\phi}_m(x)\,dx = \delta_{nm} \tag{1.10}$$

hold for the eigenfunctions corresponding to Eq. (1.4).

For the purposes of simulations, only a finite number of terms can be retained in the sum of Eq. (1.9). Supposing that the autocovariance kernel's eigenvalues are ordered in descending order, keeping only the first p terms of the expansion induces an approximation error which can be quantified in terms of the explained variance. The explained variance, EV, is defined as

$$EV = \frac{\int_{\mathscr{D}} \mathrm{Var}\left[f^p(x)\right] dx}{\int_{\mathscr{D}} \mathrm{Var}\left[f(x)\right] dx}, \tag{1.11}$$

where $\mathscr{D}$ is the domain over which the process is defined and

$$f^p(x) = \sum_{n=1}^{p} \left[\sqrt{\lambda_n}\, \phi_n(x-a)\ \xi_n + \sqrt{\tilde{\lambda}_n}\, \tilde{\phi}_n(x-a)\ \tilde{\xi}_n \right]. \tag{1.12}$$

In the present case, the variance of $f(x)$ is equal to 1 and $\mathscr{D} = [0, 5]$, therefore the denominator of Eq. (1.11) is equal to 5. As far as the numerator is concerned, given the interchangeability of the sum and the integral which is due to Fubini's theorem, one may write

$$\begin{aligned} \int_{\mathscr{D}} \mathrm{Var}\left[f^p(x)\right] dx &= \sum_{n=1}^{p} \int_0^5 \mathrm{Var}\left[\sqrt{\lambda_n}\, \phi_n(x-a)\ \xi_n + \sqrt{\tilde{\lambda}_n}\, \tilde{\phi}_n(x-a)\ \tilde{\xi}_n \right] dx \\ &= \sum_{n=1}^{p} \sum_{m=1}^{p} \left\{ \sqrt{\lambda_n}\sqrt{\lambda_m}\, \delta_{nm}\, \mathrm{Var}\left[\xi_n\right] + \sqrt{\tilde{\lambda}_n}\sqrt{\tilde{\lambda}_m}\, \delta_{nm}\, \mathrm{Var}\left[\tilde{\xi}_n\right] \right\} \\ &= \sum_{n=1}^{p} \left(\lambda_n + \tilde{\lambda}_n\right), \end{aligned} \tag{1.13}$$

where Eqs. (1.10) as well as the fact that $\mathrm{Var}\left[\xi_n\right] = \mathrm{Var}\left[\tilde{\xi}_n\right] = 1$ were used in the second equality of Eq. (1.13). Combining these results, the explained variance can be expressed as

$$EV = \frac{1}{5} \sum_{n=1}^{p} \left(\lambda_n + \tilde{\lambda}_n\right). \tag{1.14}$$

The explained variance as a function of p for the present problem is shown in the graph of Fig. 1. The vertical dashed line corresponds to the value of p for which the explained variance becomes higher than $t = 0.99$ for the first time. This value is $p = 26$ and it corresponds to the number of terms that will be retained in the expansion of Eq. (1.12) for the following simulations.

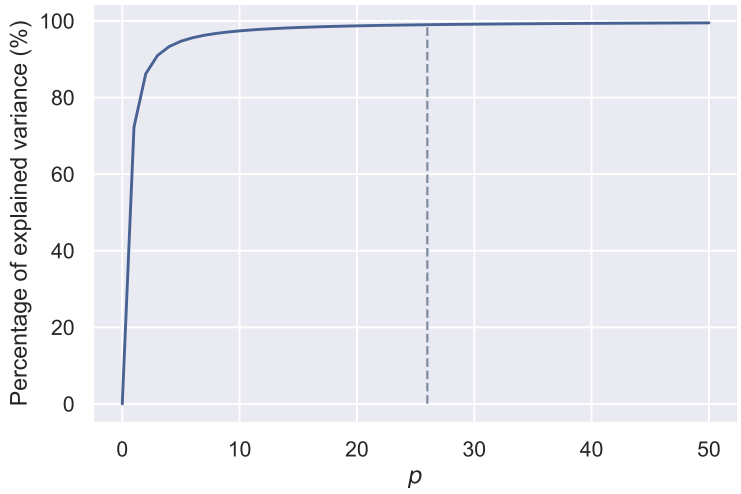


Figure 1: Explained variance as a function of p for the field $f(x)$.

The code developed in Python and shown in Snippet 1 calculates the first p solutions for ω and $\tilde{\omega}$ using SciPy's `fsolve` function and uses them to calculate the sum of Eq. (1.14) and thus determine the first value of $p \in \mathbb{N}$ for which $EV \geq t$.

```
import numpy as np
from scipy.optimize import fsolve

# Equations defined to use fsolve
def cos_eq(w,a=2.5,b=2.0):
    return (1.0/b) - (w*np.tan(w*a))

def sin_eq(w,a=2.5,b=2.0):
    return w+((1.0/b)*np.tan(w*a))

def omegas(wtype,a=2.5,b=2.0,p=100):
    ws = np.zeros(p)
    equation = cos_eq if wtype == 'cos' else sin_eq
    subt = 1.0 if wtype =='cos' else 0.5
    for idx in range(1,p+1):
        ws[idx-1] = fsolve(equation,(((idx-subt)*np.pi)/a)+1e-4)[0]
    return ws

def lambdas(ws,b=2.0):
    return (2.0*b)/(1 + (ws*b)**2)

# find p such that EV >= 99%
EV, thresh = 0.0, 0.01
p = 0
while EV < 1.0-thresh:
    p += 1
    w_cos = omegas(wtype='cos',p=p)
    w_sin = omegas(wtype='sin',p=p)
    EV = 0.2*(lambdas(w_cos)+lambdas(w_sin)).sum()
```

Python Code Snippet 1: Code for the solution of Eqs. (1.6) and calculation of p

Moving on to the simulations, it is stressed that the previous analysis corresponds to the field $f^p(x)$ of Eq. (1.12). In order to acquire results for $E^p(x)$, it suffices to simulate the realizations of $f^p(x)$ and substitute the results in Eq. (1.1). The code developed for this purpose is shown in Snippet 2. Fig. 2

depicts 5 such realizations, drawn from an ensemble of 5000 samples for $p = 26$.

```
# Eigenfunctions
def phi_cos(x,ws,a=2.5):
    coeff = 1.0/(np.sqrt(a+(np.sin(2.0*a*ws)/(2.0*ws))))
    return coeff*np.cos(ws*x)

def phi_sin(x,ws,a=2.5):
    coeff = 1.0/(np.sqrt(a-(np.sin(2.0*a*ws)/(2.0*ws))))
    return coeff*np.sin(ws*x)

# Simulation of field realizations
def realize(a=2.5,b=2.0,p=26,N=5000):
    dots = 500
    realizations = np.zeros((N,dots))
    Xs = np.linspace(-a,a,dots) # draw from [-a,a]
    # Precalculations to avoid calculating in each run
    w_cos = omegas(wtype='cos',a=a,b=b,p=p)
    w_sin = omegas(wtype='sin',a=a,b=b,p=p)
    l_cos, l_sin = np.sqrt(lambdas(w_cos,b=b)), np.sqrt(lambdas(w_sin,b=b))
    part_prod_cos, part_prod_sin = np.zeros((p,dots)), np.zeros((p,dots))
    for j in range(p):
        part_prod_cos[j,:] = l_cos[j]*phi_cos(Xs,w_cos[j],a=a)
        part_prod_sin[j,:] = l_sin[j]*phi_sin(Xs,w_sin[j],a=a)
    # Realizations
    for idx in range(N):
        xi_cos = np.random.standard_normal(p)
        xi_sin = np.random.standard_normal(p)
        costerms, sinterms = np.zeros((p,dots)), np.zeros((p,dots))
        for j in range(p):
            costerms[j,:] = part_prod_cos[j,:]*xi_cos[j]
            sinterms[j,:] = part_prod_sin[j,:]*xi_sin[j]
        realizations[idx,:] = 10.0*(1.0 + costerms.sum(0) + sinterms.sum(0))
    return realizations
```

Python Code Snippet 2: Code to simulate realizations of the stochastic process

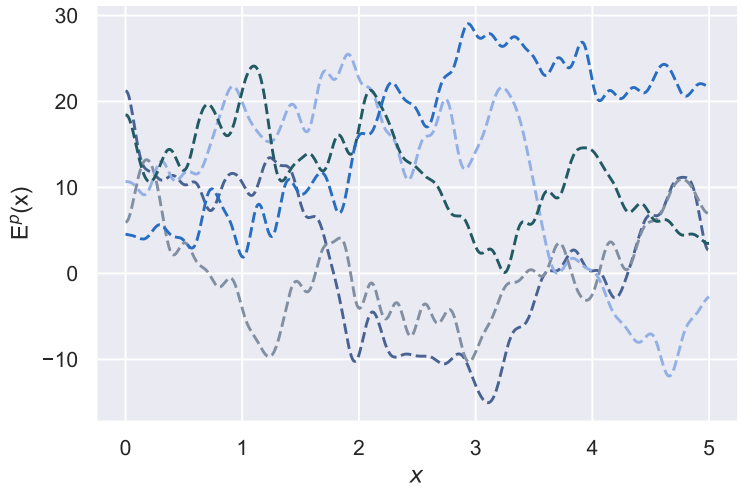


Figure 2: Random realizations of the field $E^p(x)$ of Eqs. (1.1) and (1.12) for $p = 26$.

As far as the ensemble mean and variance are concerned, the values to which they are expected to

converge can be analytically calculated. Using Eqs. (1.1) and (1.12) one obtains

$$\begin{aligned}\mathbb{E}\left[E^p\left(x\right)\right] &= 10\left\{1+\mathbb{E}\left[f^p\left(x\right)\right]\right\}\\ &= 10\left(1+\sum_{n=1}^{p}\left\{\sqrt{\lambda_n}\,\phi_n\left(x-a\right)\mathbb{E}\left[\xi_n\right]+\sqrt{\tilde{\lambda}_n}\,\tilde{\phi}_n\left(x-a\right)\mathbb{E}\left[\tilde{\xi}_n\right]\right\}\right)=10,\end{aligned} \tag{1.15}$$

since $\mathbb{E}\left[\xi_n\right]=\mathbb{E}\left[\tilde{\xi}_n\right]=0$ for all n. Similarly,

$$\begin{aligned}\mathrm{Var}\left[E^p\left(x\right)\right] &= 100\mathrm{Var}\left[f^p\left(x\right)\right]\\ &= 100\sum_{n=1}^{p}\left\{\lambda_n\,\phi_n^2\left(x-a\right)\mathrm{Var}\left[\xi_n\right]+\tilde{\lambda}_n\,\tilde{\phi}_n^2\left(x-a\right)\mathrm{Var}\left[\tilde{\xi}_n\right]\right\}+\\ &+100\sum_{n=1}^{p}\sum_{m=1}^{p}\sqrt{\lambda_n\tilde{\lambda}_m}\,\phi_n\left(x-a\right)\tilde{\phi}_m\left(x-a\right)\mathrm{Cov}\left[\xi_n,\tilde{\xi}_m\right]=\\ &= 100\sum_{n=1}^{p}\left[\lambda_n\,\phi_n^2\left(x-a\right)+\tilde{\lambda}_n\,\tilde{\phi}_n^2\left(x-a\right)\right],\end{aligned} \tag{1.16}$$

since $\mathrm{Var}\left[\xi_n\right]=\mathrm{Var}\left[\tilde{\xi}_n\right]=1$ and $\mathrm{Cov}\left[\xi_n,\tilde{\xi}_m\right]=0$ for all n,m. Note that the ensemble mean is expected to be equal to 10 no matter the choice of p, however the same does not hold for the ensemble variance, which tends to 100 in the limit $p\to\infty$, but is otherwise smaller than 100 (in fact, it is expected to be slightly higher than 99 in the present case, based on Fig. 1). In Fig. 3 the ensemble mean (dashed lines) is depicted for different ensemble volumes assuming $p=26$, with the solid line corresponding to the target mean (equal to 10).

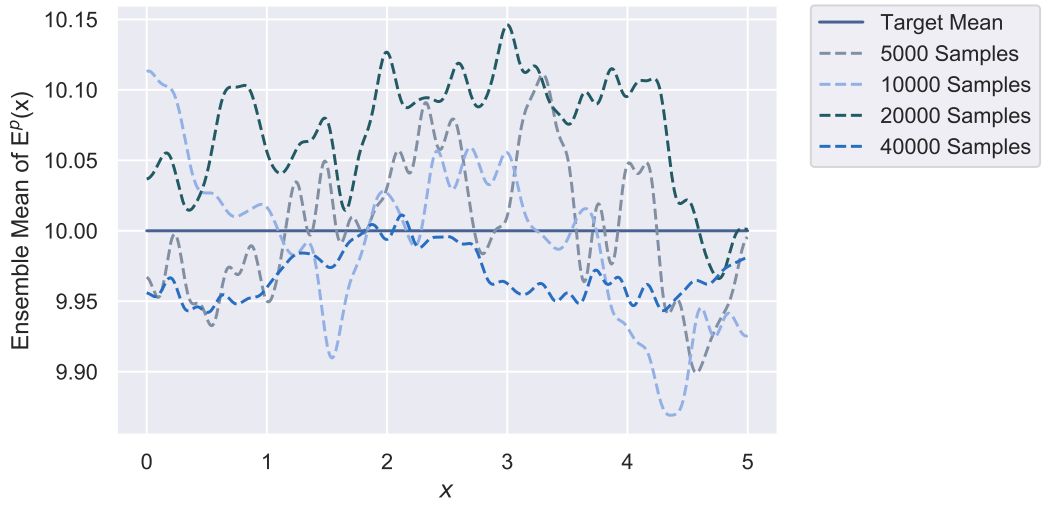


Figure 3: Ensemble mean for different ensemble volumes assuming $p=26$.

As expected, $\mathbb{E}\left[E^p\left(x\right)\right]\approx 10$ for all $x\in[0,5]$. In addition, as the number of samples increases, the graph's mean distance from the target mean decreases. In the case of the ensemble variance, the corresponding graphs can be seen in Fig. 4. The solid line corresponds to the target variance given by Eq. (1.16), which is calculated using the code shown in Snippet 3. Again, it appears that the higher the number of samples that constitute the ensemble, the lower the graph's mean distance from the target variance. As a closing remark, it is interesting to point out that the oscillatory behavior of the target variance is an example of the Gibbs phenomenon, which occurs due to the fact that p is finite, introducing

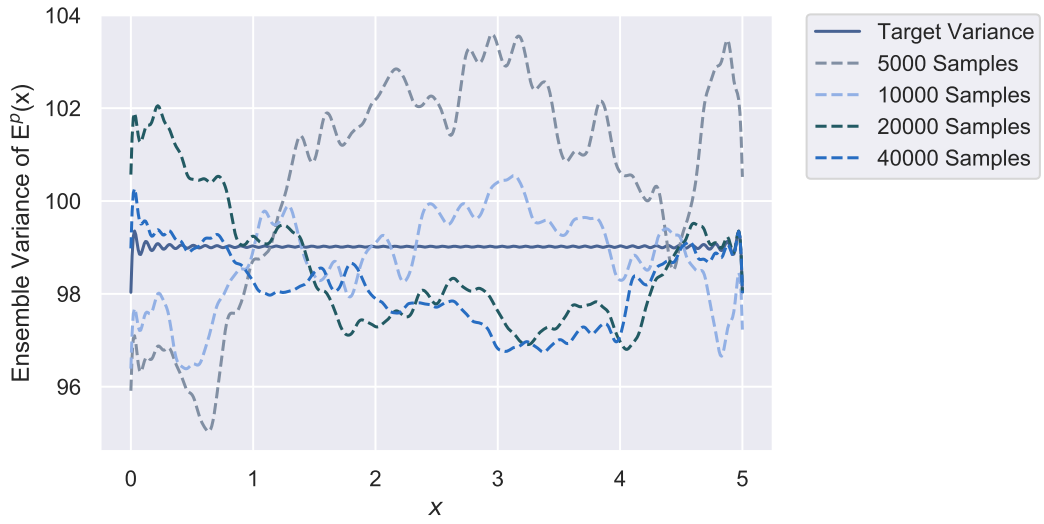


Figure 4: Ensemble variance for different ensemble volumes assuming $p = 26$.

a cutoff of sine and cosine terms in the expansion of Eq. (1.12). Theoretically, the Gibbs phenomenon is expected to disappear in the limit $p \to \infty$, however in practice it disappears for $p \gtrsim 100$.

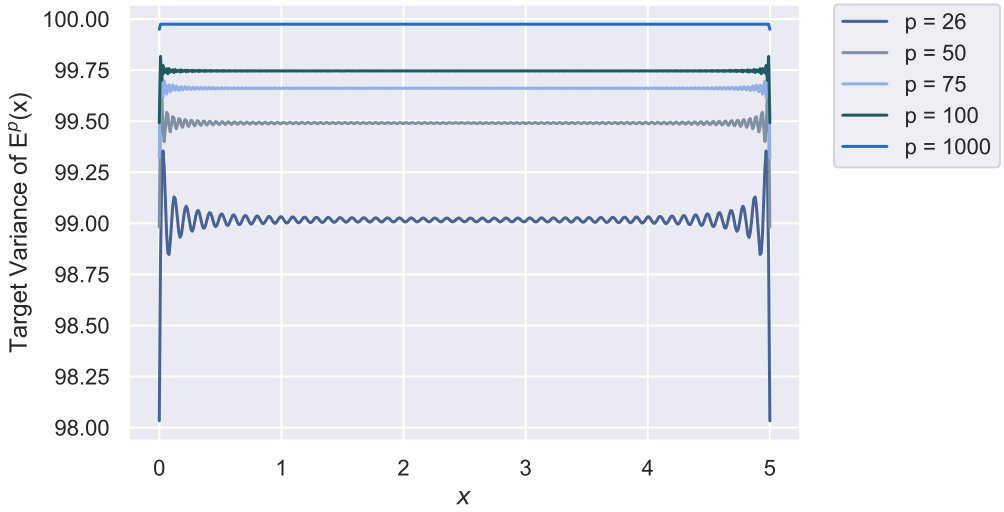


Figure 5: The gradual disappearance of the Gibbs phenomenon as p increases.

```
def conv_var(a=2.5,b=2.0,p=26):
    dots = 500
    Xs = np.linspace(-a,a,dots)
    w_cos = omegas(wtype='cos',a=a,b=b,p=p)
    w_sin = omegas(wtype='sin',a=a,b=b,p=p)
    l_cos, l_sin = lambdas(w_cos,b=b), lambdas(w_sin,b=b)
    costerms, sinterms = np.zeros((p,dots)), np.zeros((p,dots))
    for j in range(p):
        costerms[j,:] = l_cos[j]*(phi_cos(Xs,w_cos[j],a=a)**2)
        sinterms[j,:] = l_sin[j]*(phi_sin(Xs,w_sin[j],a=a)**2)
    return 100.0*(costerms.sum(0) + sinterms.sum(0))
```

Python Code Snippet 3: Code for the calculation of the target variance

2 Zero-mean Gaussian Process

The studied problem in this case involves a zero-mean Gaussian process, $X(t)$, with $t \in [0, 10]$, which has the one-sided power spectrum $G(\omega)$ given by

$$G(\omega) = \begin{cases} \omega - 1, & 1 \leq \omega \leq 2 \\ 3 - \omega, & 2 < \omega \leq 3 \\ 0, & \text{otherwise} \end{cases}. \tag{2.1}$$

According to the Wiener–Khinchin theorem, this power spectrum is related to the process' autocorrelation function, $R_X(\tau)$, via the (real) Fourier transform

$$\begin{aligned} R_X(\tau) &= \int_{-\infty}^{\infty} G(\omega)\cos(\omega\tau)\,d\omega = \int_1^2 (\omega-1)\cos(\omega\tau)\,d\omega + \int_2^3 (3-\omega)\cos(\omega\tau)\,d\omega \\ &= \frac{1}{\tau}\Big[(\omega-1)\sin(\omega\tau)\Big]_1^2 + \frac{1}{\tau^2}\Big[\cos(\omega\tau)\Big]_1^2 + \frac{1}{\tau}\Big[(3-\omega)\sin(\omega\tau)\Big]_2^3 - \frac{1}{\tau^2}\Big[\cos(\omega\tau)\Big]_2^3 \\ &= \frac{1}{\tau^2}\left[2\cos(2\tau) - \cos(3\tau) - \cos(\tau)\right] = \frac{4}{\tau^2}\sin^2\left(\frac{\tau}{2}\right)\cos(2\tau), \end{aligned} \tag{2.2}$$

where in the last equality trivial trigonometric properties were used. Given the autocorrelation function (which coincides with the covariance function, since the process is zero-mean), two interesting limits can be calculated. The first is the limit of R_X for $\tau \to 0$, which corresponds to the process' variance and is equal to

$$\mathrm{Var}[X(t)] = \lim_{\tau\to 0} R_X(\tau) = \lim_{\tau\to 0}\left\{\left[\frac{\sin(\tau/2)}{\tau/2}\right]^2 \cdot \cos(2\tau)\right\} = 1 \cdot 1 = 1. \tag{2.3}$$

The other interesting limit is that of R_X for $\tau \to \infty$, which is evaluated using

$$|R_X(\tau)| = \frac{4}{\tau^2}\left|\sin^2\left(\frac{\tau}{2}\right)\cos(2\tau)\right| \leq \frac{4}{\tau^2}\cdot 1 \xrightarrow[\tau\to\infty]{} 0 \tag{2.4}$$

and therefore, by the squeeze theorem,

$$\lim_{\tau\to\infty} R_X(\tau) = 0. \tag{2.5}$$

This result in conjunction with the fact that the process is stationary indicates that it is also ergodic in the autocorrelation.

Moving on to the simulation using the Spectral Representation method, since the process is stationary, it can be expanded as [1]

$$X(t) = \sqrt{2}\sum_{n=0}^{\infty} A_n \cos(\omega_n t + \Phi_n). \tag{2.6}$$

In the above expression[3],

$$A_n = \sqrt{G(\omega_n)\,\Delta\omega}, \quad n \in \mathbb{N}, \tag{2.7}$$

[3] Note that Eq. (2.55) of [1] includes a factor of 2 under the square root, which is inherited from Eq. (1.37), where the power spectrum was considered to be double-sided and even. The power spectrum studied here is one-sided, therefore this factor must not be included.

with $\Delta\omega$ being the step by which ω increases, i.e. $\omega_n = \omega_0 + n\Delta\omega$. Since $\omega \in [1, 3]$, ω_0 is set equal to 1. As far as the Φ_n terms are concerned, they correspond to random phase variables[4] following the uniform distribution $\mathcal{U}(0, 2\pi)$. Obviously, computing the infinite series of Eq. (2.6) is not feasible, therefore only a finite number of terms, p, is retained, inducing an approximation error as in Exercise 1. The approximated process is, therefore

$$X^p(t) = \sqrt{2} \sum_{n=0}^{p-1} A_n \cos(\omega_n t + \Phi_n), \tag{2.8}$$

where the upper limit of the sum is equal to $p - 1$ since the indexing starts from 0. As far as $\Delta\omega$ is concerned[5], it is set equal to $2/(p-1)$, since ω_{p-1} must be equal to 3. The code developed to perform the simulations is shown in Snippet 4 and Fig. 6 shows five realizations of the field of Eq. (2.8) for $p = 100$ as an example.

```
import numpy as np

def G(w):
    if 1.0 <= w <= 2.0: return w-1
    elif 2.0 < w <= 3.0: return 3-w
    else: return 0

# Simulation of field realizations
def realize(w_0=1,w_u=3,t_i=0.0,t_f=10.0,p=100,N=5000):
    dots = 500
    realizations = np.zeros((N,dots))

    Ws = np.linspace(w_0,w_u,p)
    Dw = Ws[1]-Ws[0] # (w_u-w_0)/(p-1)

    Ts = np.linspace(t_i,t_f,dots)
    Coeffs = np.zeros(p) # sqrt(2)*An
    for j in range(p):
        Coeffs[j] = np.sqrt(2.0*G(Ws[j])*Dw)

    for idx in range(N):
        Phi = np.random.uniform(0.0,2.0*np.pi,p)
        for j in range(p):
            realizations[idx,:] += Coeffs[j]*np.cos(Ws[j]*Ts + Phi[j])

    return realizations
```

Python Code Snippet 4: Code to simulate realizations of the field

As in Exercise 1, given N realizations of the approximated field, the basic goal is the investigation of its ensemble mean and variance. To this end, one may write

$$\mathbb{E}[X^p(t)] = \sqrt{2} \sum_{n=0}^{p-1} A_n \mathbb{E}[\cos(\omega_n t + \Phi_n)] = 0, \tag{2.9}$$

because

[4] As in Exercise 1, the dependence $\Phi_n = \Phi_n(\theta)$ is omitted for brevity.

[5] Note that in [1] $\Delta\omega$ is set equal to ω_u/p, where ω_u is the cutoff frequency, above which the power spectrum is considered to be practically equal to zero. This definition is reasonable in cases where $\omega \in [0, +\infty)$ and therefore a cutoff needs to be introduced in order to perform computations. However, the length of the ω-domain studied here is finite and equal to 2, since $\omega \in [1, 3]$.

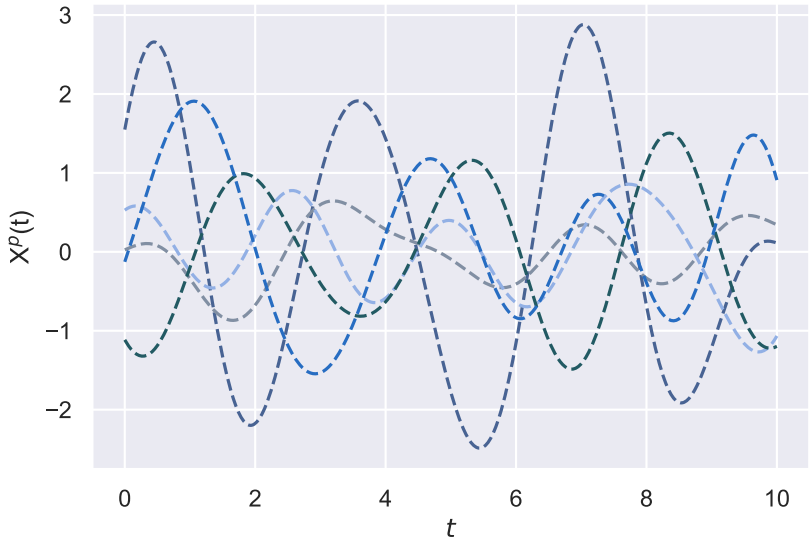


Figure 6: Random realizations of the field of Eq. (2.8) for $p = 100$.

$$\mathbb{E}\left[\cos\left(\omega_n t + \Phi_n\right)\right] = \frac{1}{2\pi}\int_0^{2\pi} \cos\left(\omega_n t + \Phi_n\right) d\Phi_n = \frac{1}{2\pi}\Big[\sin\left(\omega_n t + \Phi_n\right)\Big]_0^{2\pi} = 0. \tag{2.10}$$

Similarly,

$$\mathrm{Var}\left[X^p\left(t\right)\right] = 2\sum_{n=0}^{p-1} A_n^2 \mathrm{Var}\left[\cos\left(\omega_n t + \Phi_n\right)\right] = \sum_{n=0}^{p-1} A_n^2, \tag{2.11}$$

because of the independence between Φ_n and Φ_m for $n \neq m$ and the fact that

$$\begin{aligned}\mathrm{Var}\left[\cos\left(\omega_n t + \Phi_n\right)\right] &= \mathbb{E}\left[\cos^2\left(\omega_n t + \Phi_n\right)\right] - \mathbb{E}\left[\cos\left(\omega_n t + \Phi_n\right)\right]^2 = \frac{1}{2\pi}\int_0^{2\pi} \cos^2\left(\omega_n t + \Phi_n\right) d\Phi_n - 0 \\ &= \frac{1}{2\pi}\left[\frac{\Phi_n}{2} + \frac{\sin\left(2\omega_n t + 2\Phi_n\right)}{4}\right]_0^{2\pi} = \frac{1}{2}.\end{aligned} \tag{2.12}$$

Note at this point that if Eq. (2.7) is substituted into the sum of Eq. (2.11), the result is

$$\mathrm{Var}\left[X^p\left(t\right)\right] = \sum_{n=0}^{p-1} G\left(\omega_n\right)\Delta\omega \xrightarrow[p\to\infty]{} \int G\left(\omega\right) d\omega = 1, \tag{2.13}$$

which is consistent with the result obtained in Eq. (2.3), since $X^p\left(t\right) \to X\left(t\right)$ in the limit $p \to \infty$. These results indicate that the expected ensemble mean is equal to zero and the expected ensemble variance is practically equal to 1 for large values of p. In view of this fact, Figures 7 and 8 depict the ensemble mean and variance, respectively, for two studied cases: in the first case, corresponding to the left graphs, the number of ensemble samples is kept fixed at $N = 5000$ and p varies. In the second case, corresponding to the right graphs, the number of ensemble samples varies, while p is kept fixed at $p = 100$. As far as the ensemble mean is concerned, variations of p do not seem to significantly affect the results, which are close to 0 with double decimal accuracy for all $t \in [0, 10]$. However, increasing the sample size, N, leads to a decrease in the distance between the line $y = 0$ and the ensemble mean. In other words, larger numbers of samples lead to an ensemble mean that is closer to its expected value.

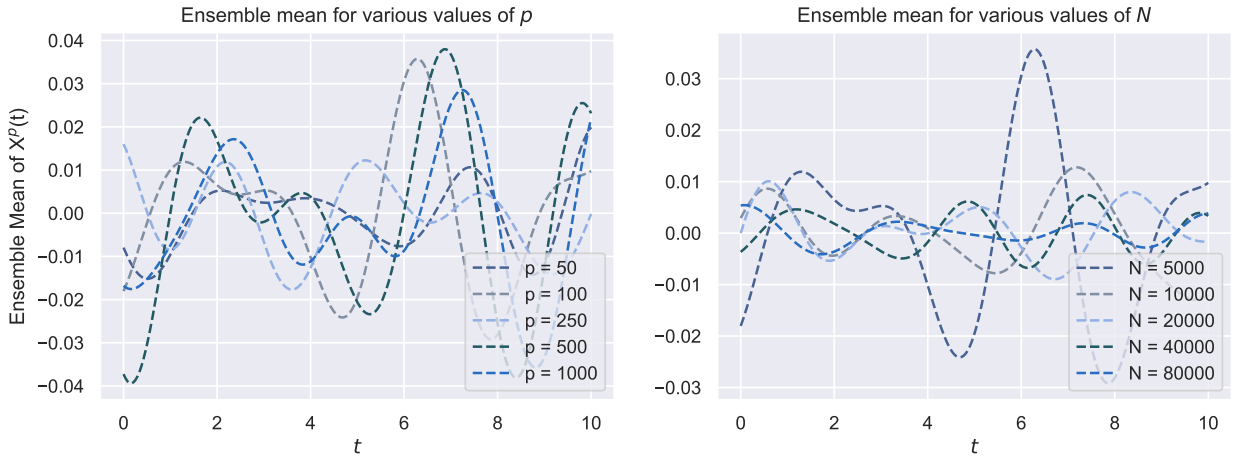


Figure 7: The ensemble mean for various values of p and $N = 5000$ (left) and various values of N and $p = 100$ (right).

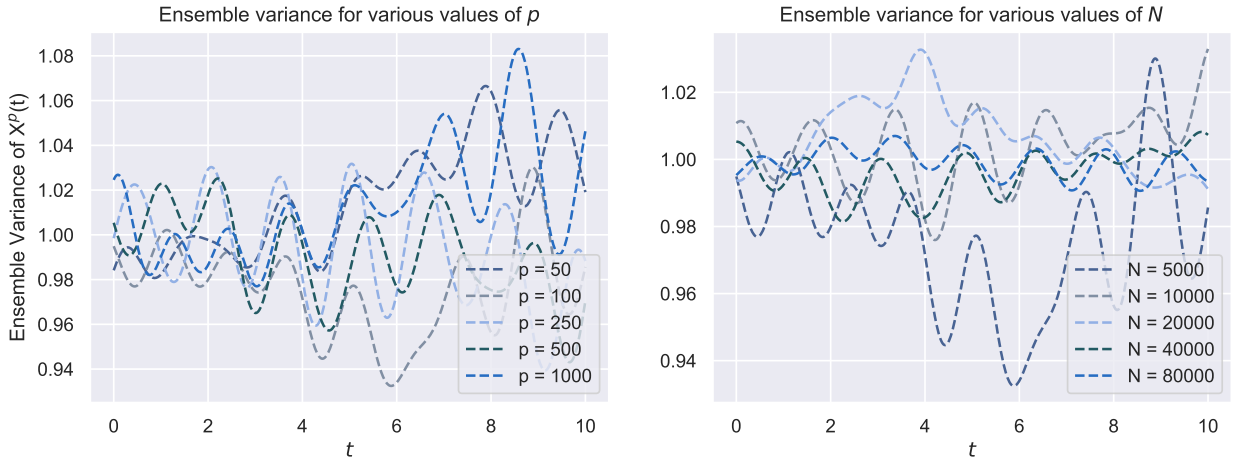


Figure 8: The ensemble variance for various values of p and $N = 5000$ (left) and various values of N and $p = 100$ (right).

The same holds in the case of the ensemble variance, which approaches the line $y = 1$ more as the number of ensemble realizations increases. In this case, not only the value of N, but also the value of p is expected to affect the results, as $\mathrm{Var}\left[X^p(t)\right] < 1$ for finite values of p. However, the left graph of Fig. 8 does not point to one such dependence. A possible reason for this is that even the lowest chosen value for p, i.e. $p = 50$, corresponds to a large number of retained terms in the expansion of Eq. (2.8), especially due to the fact that the ω-domain has a relatively small length. The fact that $G(\omega)$ is a linear function of ω might also be a reason why small values of p are sufficient to approximate $X(t)$ with a small error margin. It is possible that the ensemble variance would show significant divergences from its expected value for values of p lower than 50.

Closing the present analysis, it is worth discussing the temporal mean and variance, which are expected to converge to the corresponding ensemble statistics, due to the aforementioned ergodicity of the process. A single realization is not sufficient to provide information on that matter, especially as far as the variance is concerned. In the case $p = 100$ and $N = 5000$, randomly drawn realizations from the ensemble showed temporal means close to zero, with second decimal accuracy. However, when it came to the temporal variance, the results varied, reaching values as high as +3. Nonetheless, the mean

values of the temporal mean and average were indeed close to the expected ensemble metrics, with the mean temporal mean equal to 0.0001 and the mean temporal variance equal to 0.9987.

References

[1] V. Papadopoulos and D. G. Giovanis, *Stochastic Finite Element Methods*. Springer, 2018.

[2] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*. Springer, 1991.

[3] I. Kalogeris, *Lecture Notes on Stochastic Finite Element Methods & Data-driven models in Engineering Applications*. 8.04.2021.