Chapter 2

Stochastic Processes and Random Fields

A topic of intrinsic interest in this course is stochastic partial differential equations (SPDE). Although a rigorous treatment of this topic goes beyond our goals, in order to introduce the fundamental ideas of the basic theory of SPDE, it is necessary to discuss the basic concepts of stochastic processes and random fields. These two concepts are nothing more than extensions of the random variable concept, treated in the previous lecture, for cases in which these variables have temporal or spatial dependence, respectively. As a matter of fact there is much similarity between the two concepts at a fundamental level, with some particular nomenclature differences.

2.1 Definition and Probabilistic Concepts

In the previous lecture, the symbol $\boldsymbol{x}(\omega)$ referred to the value of a vector random variable $\mathbf{x}$, resulting from the realization of an experiment ω. Stochastic processes are those in which the random variable is also a function of time, that is, the random variables are defined on the product space $\Omega \times T$, where T represents the real time line. In this case, we denote by $\boldsymbol{x}(\omega, t)$ the result of a stochastic process $\mathbf{x}(t)$. In what follows, we utilize the most common abbreviation of a stochastic process, denoting the stochastic variables as $\boldsymbol{x}(t)$, where ω will be implicit in the notation. Stochastic processes are referred to as *discrete–time* or *continuous–time* depending whether the time domain is discrete or continuous, respectively.

In stochastic processes, an event B in the probability space is denoted by

$$B = \{\omega \in \Omega : \mathbf{x}(\omega, t) \leq \boldsymbol{x}\} . \tag{2.1}$$

The distribution function for a discrete–time process with N random n–vectors $\mathbf{x}(t_1), \cdots, \mathbf{x}(t_N)$, is defined as:

$$F_{\mathbf{x}(t_1)\cdots\mathbf{x}(t_N)}[\boldsymbol{x}(t_1), \cdots, \boldsymbol{x}(t_N)] \equiv P(\{\omega \in \Omega : \mathbf{x}(\omega, t_1) \leq \boldsymbol{x}_1, \cdots, \mathbf{x}(\omega, t_N) \leq \boldsymbol{x}_N\}) . \tag{2.2}$$

with probability density function (if it exits):

$$p_{\mathbf{x}(t_1),\cdots,\mathbf{x}(t_N)}[\boldsymbol{x}(t_1),\cdots,\boldsymbol{x}(t_N)] = \frac{\partial^{nN} F_{\mathbf{x}(t_1)\cdots\mathbf{x}(t_N)}[\boldsymbol{x}(t_1),\cdots,\boldsymbol{x}(t_N)]}{\partial\boldsymbol{x}(t_1)\cdots\partial\boldsymbol{x}(t_N)} \tag{2.3}$$

where we recall that $\partial^n/\partial\boldsymbol{x} = \partial^n/\partial x_1\cdots\partial x_n$. Consequently, we can write

$$F_{\mathbf{x}(t_1)\cdots\mathbf{x}(t_N)}[\boldsymbol{x}(t_1),\cdots,\boldsymbol{x}(t_N)] = \int_{-\infty}^{\boldsymbol{x}(t_1)}\cdots\int_{-\infty}^{\boldsymbol{x}(t_N)} p_{\mathbf{x}(t_1)\cdots\mathbf{x}(t_N)}(\boldsymbol{x}'_1,\cdots,\boldsymbol{x}'_N)\, d\boldsymbol{x}'_1\cdots d\boldsymbol{x}'_N\,. \tag{2.4}$$

In case of a continuous–time process, the probability distribution and probability density functions are defined for all times t, and can be symbolically written as

$$F_{\mathbf{x}}(\boldsymbol{x},t) \equiv P(\{\omega\in\Omega : \mathbf{x}(\omega,t)\le\boldsymbol{x}\}) \tag{2.5a}$$

$$p_{\mathbf{x}}(\boldsymbol{x},t) = \frac{\partial^n F_{\mathbf{x}}(\boldsymbol{x},t)}{\partial\boldsymbol{x}} \tag{2.5b}$$

respectively.

The concepts of mean, variance and correlation introduced in the previous lecture can be extended directly to the case of stochastic processes. Therefore, we define concisely these quantities for this case:

- *Mean vector*:

$$\mu_{\mathbf{x}}(t) \equiv \mathcal{E}\{\mathbf{x}(t)\} = \int_{-\infty}^{\infty} \boldsymbol{x}\, p_{\mathbf{x}(t)}(\boldsymbol{x})\, d\boldsymbol{x}\,, \tag{2.6}$$

- *Stationary mean value vector*: defined when the mean is independent of time, that is

$$\bar{\mathbf{x}} \equiv \lim_{t_f\to\infty}\frac{1}{2t_f}\int_{-t_f}^{t_f}\boldsymbol{x}(t)\,dt\,. \tag{2.7}$$

  For the case in which the stationary mean value coincides with the ensemble mean μ, the process is called ergodic in the mean.

- *Mean for discrete–time processes*:

$$\bar{\mathbf{x}} \equiv \lim_{K\to\infty}\frac{1}{2K+1}\sum_{k=-K}^{K}\boldsymbol{x}(kT) \tag{2.8}$$

  where T is the sampling period.

- *Quadratic mean value matrix*:

$$\boldsymbol{\Gamma}_{\mathbf{x}}(t) \equiv \mathcal{E}\{\mathbf{x}(t)\mathbf{x}(t)^T\} = \int_{-\infty}^{\infty}\boldsymbol{x}\,\boldsymbol{x}^T\, p_{\mathbf{x}(t)}(\boldsymbol{x})\; d\boldsymbol{x}\,. \tag{2.9}$$

  We can still define the stationary quadratic mean value based on the definition of stationary mean value, as we can define the stationary quadratic mean value for a discrete–time process utilizing the corresponding definition given above.

- *Auto–correlation matrix*:

$$\mathbf{\Gamma}_{\mathbf{x}}(t,\tau) \equiv \mathcal{E}\{\mathbf{x}(t)\mathbf{x}^T(\tau)\}\,, \tag{2.10}$$

or explicitly written,

$$\mathbf{\Gamma}_{\mathbf{x}}(t,\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{z}\,\boldsymbol{y}^T\, p_{\mathbf{x}(t)\mathbf{x}(\tau)}(\boldsymbol{z},\boldsymbol{y})\; d\boldsymbol{z}d\boldsymbol{y}\,. \tag{2.11}$$

where the word *auto* refers to the stochastic process $\mathbf{x}(t)$.

- *Cross–correlation matrix*:

$$\mathbf{\Gamma}_{\mathbf{xy}}(t,\tau) \equiv \mathcal{E}\{\mathbf{x}(t)\mathbf{y}^T(\tau)\}\,. \tag{2.12}$$

where the word *cross* refers to the two stochastic processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$.

- *Auto–covariance matrix of a stochastic process*:

$$\mathbf{C}_{\mathbf{x}}(t,\tau) = cov\{\mathbf{x}(t),\mathbf{x}(\tau)\} \equiv \mathcal{E}\{[\mathbf{x}(t)-\boldsymbol{\mu}_{\mathbf{x}}(t)][\mathbf{x}(\tau)-\boldsymbol{\mu}_{\mathbf{x}}(\tau)]^T\}\,, \tag{2.13}$$

where the designation *auto* refers to the stochastic process in question, in this case, only $\mathbf{x}(t)$. It is simple to show that

$$\mathbf{C}_{\mathbf{x}}(t,\tau) = \mathbf{\Gamma}_{\mathbf{x}}(t,\tau) - \mu_{\mathbf{x}}(t)\mu_{\mathbf{x}}^T(\tau) \tag{2.14}$$

When $t=\tau$, we define the covariance matrix as:

$$\mathbf{P}_{\mathbf{x}}(t) \equiv \mathbf{C}_{\mathbf{x}}(t,t)\,, \tag{2.15}$$

which is sometimes referred to as the variance matrix.

- *Cross–covariance matrix of a stochastic process*:

$$\mathbf{C}_{\mathbf{xy}}(t,\tau) \equiv \mathcal{E}\{[\mathbf{x}(t)-\boldsymbol{\mu}_{\mathbf{x}}(t)][\mathbf{y}(\tau)-\boldsymbol{\mu}_{\mathbf{y}}(\tau)]^T\}\,, \tag{2.16}$$

where the designation *cross* refers to the stochastic processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$. We can easily show that

$$\mathbf{C}_{\mathbf{xy}}(t,\tau) = \mathbf{\Gamma}_{\mathbf{xy}}(t,\tau) - \mu_{\mathbf{x}}(t)\mu_{\mathbf{y}}^T(\tau) \tag{2.17}$$

Correlation matrices can be defined analogously to the definitions given in the previous lecture.

2.2 Independent Process

We say that a stochastic process $\mathbf{x}(t)$ is independent when for all t and τ the probability density $p_{\mathbf{x}(t),\mathbf{x}(\tau)}(\boldsymbol{x}(t),\boldsymbol{x}(\tau)) = p_{\mathbf{x}(t)}p_{\mathbf{x}(\tau)}$. In this way, according to (2.11) it follows that

$$\begin{aligned}
\mathbf{\Gamma}_{\mathbf{x}}(t,\tau) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{z}\,\boldsymbol{y}^T\, p_{\mathbf{x}(t)}(\boldsymbol{z})p_{\mathbf{x}(\tau)}(\boldsymbol{y})\; d\boldsymbol{z}d\boldsymbol{y}\,. \\
&= \int_{-\infty}^{\infty} d\boldsymbol{z}\,\boldsymbol{z}\, p_{\mathbf{x}(t)}(\boldsymbol{z}) \int_{-\infty}^{\infty} d\boldsymbol{y}\,\boldsymbol{y}^T\, p_{\mathbf{x}(\tau)}(\boldsymbol{y}) \\
&= \mathcal{E}\{\mathbf{x}(t)\}\,\mathcal{E}\{\mathbf{x}^T(\tau)\}
\end{aligned} \tag{2.18}$$

Therefore, from the definition of auto–covariance matrix it follows that $\mathbf{C}_{\mathbf{x}}(t,\tau) = \mathbf{0}$, for $t \neq \tau$, that is, an independent stochastic process is uncorrelated in time. In an entirely analogous way, we can show that if two stochastic processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are independent, they are also uncorrelated, that is, $\mathbf{C}_{\mathbf{xy}}(t,\tau) = \mathbf{0}$, for any t and τ. As in the case of random variables, the contrary of this relation is not necessarily true, that is, two uncorrelated processes are not necessarily independent.

2.3 Markov Process

As in stochastic processes in general, a Markov process can be continuous or discrete depending on whether the time parameter is continuous or discrete, respectively. A discrete stochastic process (i.e., stochastic sequence) $\{\mathbf{x}(t_k)\}$, for $t_k > t_0$, or a continuous stochastic process $\mathbf{x}(t)$, for $t > t_0$, is said to be a Markov process if, for all $\tau \leq t$,

$$p_{\mathbf{x}(t)|\Xi(\tau)}[\boldsymbol{x}(t)|\xi(\tau)] = p_{\mathbf{x}(t)|\mathbf{x}(\tau)}[\boldsymbol{x}(t)|\boldsymbol{x}(\tau)] \tag{2.19}$$

where $\Xi(\tau) = \{\mathbf{x}(s), t_0 \leq s \leq \tau \leq t\}$, and analogously $\xi(\tau) = \{\boldsymbol{x}(s), t_0 \leq s \leq \tau \leq t\}$. More specifically for the discrete case, a first–order Markov process, also referred to as a Markov–1 process, is one for which

$$p_{\mathbf{x}_k|\mathbf{x}_{k-1}\cdots\mathbf{x}_1\mathbf{x}_0}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1},\cdots,\boldsymbol{x}_1,\boldsymbol{x}_0) = p_{\mathbf{x}_k|\mathbf{x}_{k-1}}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \tag{2.20}$$

That is to say, a Markov–1 process is one for which the probability density at time t, given all states up to t, in the interval $[t_0,\tau]$, depends only on the state at the final time, τ, of the interval. This is nothing more than a way of stating the causality principle: the state of a process at a particular moment in time is sufficient for us to determine the future states of the process, without us having to know its complete history.

In the discrete case, we can write for the joint probability density

$$\begin{aligned} p_{\Xi_k}(\xi_k) &= p_{\mathbf{x}_k\cdots\mathbf{x}_1\mathbf{x}_0}(\boldsymbol{x}_k,\cdots,\boldsymbol{x}_1,\boldsymbol{x}_0) \\ &= p_{\mathbf{x}_k|\mathbf{x}_{k-1}\cdots\mathbf{x}_1\mathbf{x}_0}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1},\cdots,\boldsymbol{x}_1,\boldsymbol{x}_0)\, p_{\mathbf{x}_{k-1}\cdots\mathbf{x}_1\mathbf{x}_0}(\boldsymbol{x}_{k-1},\cdots,\boldsymbol{x}_1,\boldsymbol{x}_0) \end{aligned} \tag{2.21}$$

where we utilize the property (1.77). Assuming that the stochastic process is first–order Markov, according to (2.20) we have that

$$p_{\Xi_k}(\xi_k) = p_{\mathbf{x}_k|\mathbf{x}_{k-1}}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\, p_{\mathbf{x}_{k-1}\cdots\mathbf{x}_1\mathbf{x}_0}(\boldsymbol{x}_{k-1},\cdots,\boldsymbol{x}_1,\boldsymbol{x}_0) \tag{2.22}$$

and utilizing repeatedly the definition (2.20) we obtain

$$p_{\Xi_k}(\xi_k) = p_{\mathbf{x}_k|\mathbf{x}_{k-1}}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\, p_{\mathbf{x}_{k-1}|\mathbf{x}_{k-2}}(\boldsymbol{x}_{k-1}|\boldsymbol{x}_{k-2})\, \cdots\, p_{\mathbf{x}_1|\mathbf{x}_0}(\boldsymbol{x}_1|\boldsymbol{x}_0)\, p_{\mathbf{x}_0}(\boldsymbol{x}_0) \tag{2.23}$$

Therefore, the joint probability density of a Markov–1 process can be determined from the initial marginal probability density $p_{\mathbf{x}(0)}[\boldsymbol{x}(0)]$, and from the probability density $p_{\mathbf{x}(t)|\mathbf{x}(s)}[\boldsymbol{x}(t)|\boldsymbol{x}(s)]$, for $t \geq s \in [t_0,\tau)$, and $t < \tau$.The quantity $p_{\mathbf{x}(t)|\mathbf{x}(s)}[\boldsymbol{x}(t)|\boldsymbol{x}(s)]$ is known as the *transition probability density* of a Markov process.

The concept of Markovian process can be extended to define Markov processes of different orders. For example, a discrete stochastic process for which the probability density at time t_k depends on the process at times t_{k-1} and t_{k-2} can be defined as those for which we have:

$$p_{\mathbf{x}_k|\mathbf{x}_{k-1}\cdots\mathbf{x}_1\mathbf{x}_0}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}\cdots,\boldsymbol{x}_0) = p_{\mathbf{x}_k|\mathbf{x}_{k-1}\mathbf{x}_{k-2}}(\boldsymbol{x}_k|\boldsymbol{x}_{k-1},\boldsymbol{x}_{k-2}) \tag{2.24}$$

which in this case is called a second–order Markov process, or Markov–2. Analogously, we can define k^{th}–order Markov processes. In those cases considered in this course, the definition given in (2.20) for first–order Markov processes is sufficient.

2.4 Gaussian Process

A stochastic process in n dimensions $\{\mathbf{x}(t), t \in T\}$, where T is an arbitrary interval of time, is said to be Gaussian if for any N instants of time $t_1, t_2, \cdots, t_N$ in T, its density function, distribution function, or characteristic function, is normal. In other words, the process is Gaussian if the vectors $\mathbf{x}(t_1), \mathbf{x}(t_2), \cdots, \mathbf{x}(t_N)$ are jointly Gaussian distributed. According to what was seen in the previous lecture we can write the density function of this process as:

$$p_{\mathbf{z}}(\boldsymbol{z}) = \frac{1}{(2\pi)^{Nn/2}|\mathbf{P}_{\mathbf{z}}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu}_{\mathbf{z}})^T\mathbf{P}_{\mathbf{z}}^{-1}(\boldsymbol{z}-\boldsymbol{\mu}_{\mathbf{z}})\right], \tag{2.25}$$

where the vector $\boldsymbol{z}$, of dimension $Nn = N \times n$, is defined as:

$$\boldsymbol{z} \equiv \begin{pmatrix} \boldsymbol{x}(t_1) \\ \boldsymbol{x}(t_2) \\ \vdots \\ \boldsymbol{x}(t_N) \end{pmatrix}, \tag{2.26}$$

the mean vectors $\boldsymbol{\mu}_{\mathbf{z}}(t_i)$, of dimension Nn are given by

$$\boldsymbol{\mu}_{\mathbf{z}}(t_i) \equiv \mathcal{E}\{\mathbf{z}(t_i)\}, \tag{2.27}$$

for $i = 1, 2, \cdots, N$, and the covariance $\mathbf{P}_{\mathbf{z}}$, of dimension $(N \times n)^2 = Nn \times Nn$, has elements which are the sub–matrices

$$\mathbf{P}_{\mathbf{z}} = [\mathbf{P}_{\mathbf{x}}]_{ij} \equiv \mathcal{E}\{[\mathbf{x}(t_i) - \boldsymbol{\mu}_{\mathbf{x}}(t_i)][\mathbf{x}(t_j) - \boldsymbol{\mu}_{\mathbf{x}}(t_j)]^T\} \tag{2.28}$$

for $i, j = 1, 2, \cdots, N$. In this way, a Gaussian process is completely determined by its mean and its autocovariance. A process which is simultaneously Gaussian and Markovian is said to be a Gauss–Markov process.

2.5 Stationary Process

A precise definition of the concept of stationary process can be given by returning to the concept of probability. However, for what interests us, it is sufficient to utilize wide–sense

stationary processes, which only requires that the first two moments be time–independent. In this sense, a stationary process is one for which the mean is independent of time:

$$\boldsymbol{\mu}_{\mathbf{x}}(t) \;=\; \boldsymbol{\mu}_{\mathbf{x}}\,, \tag{2.29}$$

and for which the correlation only depends on the time interval τ between events:

$$\boldsymbol{\Gamma}_{\mathbf{x}}(t,\tau) \;=\; \boldsymbol{\Gamma}_{\mathbf{x}}(t-\tau)\,, \tag{2.30}$$

which can be written as:

$$\boldsymbol{\Gamma}_{\mathbf{x}}(\tau) \;=\; \boldsymbol{\Gamma}_{\mathbf{x}}(t+\tau,t) \;=\mathcal{E}\{\mathbf{x}(t+\tau)\mathbf{x}^T(t)\}\,. \tag{2.31}$$

An even weaker concept of stationary process is defined when the covariance is stationary. In this case,

$$\mathbf{C}_{\mathbf{x}}(\tau) \;=\; \mathbf{C}_{\mathbf{x}}(t+\tau,t) \;=\; \boldsymbol{\Gamma}_{\mathbf{x}}(t+\tau,t) \;-\; \boldsymbol{\mu}_{\mathbf{x}}(t+\tau)\boldsymbol{\mu}_{\mathbf{x}}^T(t) \tag{2.32}$$

These concepts apply similarly to "cross–quantities", that is, the cross–correlation and cross–covariance

$$\boldsymbol{\Gamma}_{\mathbf{xy}}(\tau) \;=\; \boldsymbol{\Gamma}_{\mathbf{xy}}(t+\tau,t) \;=\mathcal{E}\{\mathbf{x}(t+\tau)\mathbf{y}^T(t)\}\,. \tag{2.33a}$$

$$\mathbf{C}_{\mathbf{xy}}(\tau) \;=\; \mathbf{C}_{\mathbf{xy}}(t+\tau,t) \;=\; \boldsymbol{\Gamma}_{\mathbf{xy}}(t+\tau,t) \;-\; \boldsymbol{\mu}_{\mathbf{x}}(t+\tau)\boldsymbol{\mu}_{\mathbf{y}}^T(t) \tag{2.33b}$$

respectively, are stationary. In this case, it is simple to show that

$$\boldsymbol{\Gamma}_{\mathbf{xy}}(\tau) \;=\; \boldsymbol{\Gamma}_{\mathbf{yx}}(-\tau) \tag{2.34a}$$

$$\mathbf{C}_{\mathbf{xy}}(\tau) \;=\; \mathbf{C}_{\mathbf{yx}}(-\tau) \tag{2.34b}$$

since stationary covariances and correlations are invariant under a time translation of $-\tau$.

2.6 Wiener–Khintchine Relation

A definition that follows from the concept of stationary process introduced above is given by the Wiener–Khintchine relation. This relation defines the spectral density of the stationary covariance as being the Fourier transform of the covariance. For a continuous stochastic process the power spectrum of the covariance can be written as:

$$\hat{\mathbf{C}}_{\mathbf{x}}(\omega) \equiv \int_{-\infty}^{\infty} \mathbf{C}_{\mathbf{x}}(\tau)e^{-i\omega\tau}\,d\tau\,, \tag{2.35}$$

and consequently, by the inverse Fourier transform we have that

$$\mathbf{C}_{\mathbf{x}}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\mathbf{C}}_{\mathbf{x}}(\omega)e^{i\omega\tau}\,d\omega\,. \tag{2.36}$$

For discrete stationary processes the discrete Fourier transform defines the corresponding power spectrum.

2.7 White Noise Process

The simplest power spectrum that we can think of is the one given by a constant, that is, one for which $\hat{\mathbf{C}}(\omega) \equiv \hat{\mathbf{Q}}_{\mathbf{w}}(\omega) = \mathbf{Q}_{\mathbf{w}}$, where the stochastic process $\mathbf{w}$ is called white noise. In this case, the covariance becomes a Dirac delta:

$$\begin{aligned} \mathbf{Q}_{\mathbf{w}}(\tau) &= \frac{1}{2\pi}\mathbf{Q}_{\mathbf{w}} \int_{-\infty}^{\infty} e^{i\omega\tau}\, d\omega \\ &= \mathbf{Q}_{\mathbf{w}}\delta(\tau)\,. \end{aligned} \tag{2.37}$$

Even if this noise is completely non–physical, because it is infinite at the origin, it is of great importance in the development of stochastic differential equations.

2.8 Wiener Process

A Wiener process, also called Brownian motion, denoted by $\mathbf{b}(t)$, is defined as the integral of a stationary, Gaussian white noise process $\mathbf{w}(t)$ with zero mean:

$$\mathbf{b}(t) = \int_0^t \mathbf{w}(t)\, dt\,, \tag{2.38}$$

where

$$cov\{\mathbf{w}(t), \mathbf{w}(\tau)\} = \mathbf{Q}_{\mathbf{w}}\delta(t-\tau)\,, \tag{2.39}$$

as we saw above. Some of the properties of this process are listed below:

1. $\mathbf{b}(t)$ is normally distributed.
2. $\mathcal{E}\{\mathbf{b}(t)\} = \mathbf{0}$, for all $t \leq 0$.
3. $P\{\mathbf{b}(0) = 0\} = 1$.
4. $\mathbf{b}(t)$ has independent and stationary increments, that is, independent of time. We refer to increments as being the differences $\mathbf{b}(t_1) - \mathbf{b}(t_2), \cdots, \mathbf{b}(t_{n-1}) - \mathbf{b}(t_n)$, where $t_{i+1} < t_i$, with $t_i \in T$.
5. $\mathbf{b}(t)$ is a Markov process.

Moreover the variance of a Wiener process increases linearly in time:

$$\begin{aligned} var\{\mathbf{b}(t)\} = \mathcal{E}\{\mathbf{b}(t)\mathbf{b}^T(t)\} &= \int_0^t \int_0^t \mathcal{E}\{\mathbf{w}(t_1)\mathbf{w}^T(t_2)\}\, dt_1\, dt_2 \\ &= \mathbf{Q}_{\mathbf{w}} \int_0^t \int_0^t \delta(t_1 - t_2)\, dt_1\, dt_2 \\ &= \mathbf{Q}_{\mathbf{w}} \int_0^t dt_1 \;=\; \mathbf{Q}_{\mathbf{w}}\, t\,, \end{aligned} \tag{2.40}$$

where we used the following definition of a delta function:

$$\int_0^t f(s)\delta(t-s)\,ds =, f(t)\,. \tag{2.41}$$

It is important to notice that the difficulty encountered in the description of a Gaussian white noise process, the problem of infinite variance, does not exist for the Wiener process. That means, the latter is a well–behaved process.

2.9 Spatial Random Fields

The literature on random stochastic fields is relatively smaller than that on stochastic processes. Still, there are several treatments, such as those of Vanmarcke [132] and Yaglom [140]. A more recent treatment, directed toward earth science applications is the one of Christakos [24]. In what follows, we will be as concise as possible, keeping in mind that the main purpose of this section is to introduce the concepts of homogeneity and isotropy for random fields.

The concept of random fields can be introduced similarly to the way we introduced stochastic processes. In this case, we associate with each random variable $\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_n$ the points $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n$ in the space R^n. A random spatial field can be considered a function of events $\omega \in \Omega$, where Ω is the sample space introduced in the previous lecture, and also a function of the spatial position $\mathbf{r} \in R^n$, that is, $\mathrm{x}(\mathbf{r}) = \mathrm{x}(\omega, \mathbf{r})$. When we write $\mathrm{x}(\mathbf{r})$ we are simplifying the notation in a manner entirely analogous to what we did in the previous section, when the variable was the time. This concept can be extended to several random variables depending on space in order to motivate the introduction to vector random spatial fields. We denote by $\mathbf{x}(\mathbf{r})$ the vector random field which represents the set of random spatial fields $\mathrm{x}_1(\mathbf{r}), \mathrm{x}_2(\mathbf{r}), \cdots, \mathrm{x}_m(\mathbf{r})$, that is,

$$\mathbf{x}(\mathbf{r}) \;=\; [\mathrm{x}_1(\mathbf{r}), \mathrm{x}_2(\mathbf{r}), \cdots, \mathrm{x}_m(\mathbf{r})]^T \tag{2.42}$$

The distribution function of a vector random spatial field is then defined as:

$$F_{\mathbf{x}}(\boldsymbol{x}, \mathbf{r}) = P(\{\omega : \mathbf{x}(\mathbf{r}) \leq \boldsymbol{x}; \mathbf{r} \in R^n\})\,. \tag{2.43}$$

We emphasize once more that the concept of random fields is an extension of the concept of stochastic process. A stochastic process is a random field for which the spatial argument $\mathbf{r} \in R^n$, is introduced for $n = 1$ and $\mathbf{r} \to r \to t$ so that the random variable becomes $\mathbf{x}(t)$, as before.

The distribution function is related to the probability density by means of the expression

$$p_{\mathbf{x}}(\boldsymbol{x}, \mathbf{r}) = \frac{\partial^n F_{\mathbf{x}}(\boldsymbol{x}, \mathbf{r})}{\partial \boldsymbol{x}} \tag{2.44}$$

and consequently

$$F_{\mathbf{x}}(\boldsymbol{x}, \mathbf{r}) = \int_{-\infty}^{\boldsymbol{x}} p_{\mathbf{x}}(\boldsymbol{x}', \mathbf{r})\, d\boldsymbol{x}'\,. \tag{2.45}$$

The concepts of mean, variance and correlation can be extended directly for the case of random spatial fields. Therefore we define concisely these quantities in this case:

- *Mean value of a random field*:

$$\mu_{\mathbf{x}}(\mathbf{r}) \equiv \mathcal{E}\{\mathbf{x}(\mathbf{r})\} = \int_{-\infty}^{\infty} \boldsymbol{x} p_{\mathbf{x}(\mathbf{r})}(\boldsymbol{x}, \mathbf{r})\, d\boldsymbol{x}\,, \tag{2.46}$$

- *Auto–covariance matrix of a random spatial field*:

$$\mathbf{C}_{\mathbf{x}}(\mathbf{r}_i, \mathbf{r}_j) = cov\{\mathbf{x}(\mathbf{r}_i), \mathbf{x}(\mathbf{r}_j)\} \equiv \mathcal{E}\{[\mathbf{x}(\mathbf{r}_i) - \boldsymbol{\mu}_{\mathbf{x}}(\mathbf{r}_i)][\mathbf{x}(\mathbf{r}_j) - \boldsymbol{\mu}_{\mathbf{x}}(\mathbf{r}_j)]^T\}\,, \tag{2.47}$$

  for two spatial points $\mathbf{r}_i$ and $\mathbf{r}_j$, where we made an analogy with what we saw in stochastic processes; *auto* refers to the random field in question, in this case $\mathbf{x}(\mathbf{r})$. Then, we have that

$$\mathbf{C}_{\mathbf{x}}(\mathbf{r}_i, \mathbf{r}_j) = \mathcal{E}\{\mathbf{x}(\mathbf{r}_i)\mathbf{x}^T(\mathbf{r}_j)\} - \mu_{\mathbf{x}}(\mathbf{r}_i)\mu_{\mathbf{x}}^T(\mathbf{r}_j) \tag{2.48}$$

  When $\mathbf{r}_i = \mathbf{r}_j$, we have the variance matrix which describes the local behavior of the random field.

In order to simplify and more easily demonstrate the notation, consider the case of a scalar random field $\mathrm{x}(\mathbf{r})$. The mean introduced above becomes a scalar quantity $\mu(\mathbf{r})$, that is, a function of "one" spatial variable $\mathbf{r} \in R^n$. The covariance becomes a function (no longer a matrix) of "two" spatial variables $\mathbf{r}_i, \mathbf{r}_j$. The variance is a function given by

$$\sigma_{\mathrm{X}}^2(\mathbf{r}) \;=\; C_{\mathrm{X}}(\mathbf{r}, \mathbf{r}_j = \mathbf{r}) \tag{2.49}$$

for $\mathbf{r} = \mathbf{r}_i$. We can still introduce the spatial correlation function $\rho_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j)$ between two points as:

$$\rho_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j) \;=\; \frac{C_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j)}{\sigma_{\mathrm{X}}(\mathbf{r}_i)\sigma_{\mathrm{X}}(\mathbf{r}_j)} \tag{2.50}$$

A scalar random spatial field is said to be uncorrelated when

$$C_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \sigma_{\mathrm{X}}^2(\mathbf{r})\,, & \text{for } \mathbf{r}_i = \mathbf{r}_j = \mathbf{r} \\ 0\,, & \text{otherwise} \end{cases} \tag{2.51}$$

and in fact, such a random field is said to be a white field (analogously to the white process seen previously).

Basically all the concepts defined for random processes can be generalized for spatial random fields:

- Markovian *process* $\rightarrow$ Markovian *field*
- Gaussian *process* $\rightarrow$ Gaussian *field*
- white *process* $\rightarrow$ white *field*

as for the concepts of characteristic function, conditional probability function, conditional mean, conditional covariance, etc.

A very important generalization is that of the concept of stationarity of a stochastic process, which for spatial random fields translates into the concept of *homogeneity*. In the wide sense, a spatial random field is said to be homogeneous when its mean value is independent of the spatial variable, and its covariance depends only on the distance between two points in space. For the scalar case, this can be written as:

$$\mu_{\mathrm{X}}(\mathbf{r}) = \mu \tag{2.52a}$$

$$C_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j) = C_{\mathrm{X}}(\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j) \tag{2.52b}$$

Another fundamental concept is that of the *isotropic* spatial random field, which is defined as a field for which

$$C_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j) = C_{\mathrm{X}}(r = |\mathbf{r}_i - \mathbf{r}_j|) \tag{2.53}$$

is satisfied. That is, a spatial random field is said to be isotropic when its covariance depends only on the magnitude of the distance between two points in space.

It is possible to show (e.g., Christakos [24]) that for a homogeneous random field, not necessarily isotropic, we can write the covariance function as

$$C_{\mathrm{X}}(\mathbf{r}) = \int_{R^n} \exp(i\mathbf{w}^T\mathbf{r})\, \hat{C}_{\mathrm{X}}(\mathbf{w})\, d\mathbf{w} \tag{2.54}$$

where $\hat{C}_{\mathrm{X}}(\mathbf{w})$ is the spectral density function that, by the inverse Fourier transform, can be written as:

$$\hat{C}_{\mathrm{X}}(\mathbf{w}) = \frac{1}{(2\pi)^n} \int_{R^n} \exp(-i\mathbf{w}^T\mathbf{r})\, C_{\mathrm{X}}(\mathbf{r})\, d\mathbf{r} \tag{2.55}$$

This result can be generalized for the case of vector random fields. Notice that for real random fields, the covariance and spectral density can in fact be expressed in terms of Fourier cosine integrals.

$$C_{\mathrm{X}}(\mathbf{r}) = \int_{R^n} \cos(\mathbf{w}^T\mathbf{r})\, \hat{C}_{\mathrm{X}}(\mathbf{w})\, d\mathbf{w} \tag{2.56a}$$

$$\hat{C}_{\mathrm{X}}(\mathbf{w}) = \frac{1}{(2\pi)^n} \int_{R^n} \cos(\mathbf{w}^T\mathbf{r})\, C_{\mathrm{X}}(\mathbf{r})\, d\mathbf{r} \tag{2.56b}$$

The importance of these results lies in the fact that they provide a relatively simple criterion to determine whether a continuous and symmetric function in R^n can be a covariance function. In fact, the necessary and sufficient condition for a continuous function $C\mathrm{x}(\mathbf{r}_i, \mathbf{r}_j)$ in R^n to be a covariance function is that it be a positive–semidefinite function, that is,

$$\int_{R^n} \int_{R^n} C_{\mathrm{X}}(\mathbf{r}_i, \mathbf{r}_j) f(\mathbf{r}_i) f(\mathbf{r}_j)\, d\mathbf{r}_i\, d\mathbf{r}_j \geq 0 \tag{2.57}$$

for any function $f(\mathbf{r})$. This criterion is generally very difficult to verify, even for homogeneous random fields. However, utilizing the spectral representation above, Bochner's [15]

theorem says that the criterion for a continuous and symmetric function in R^n to be a covariance function is that its spectral function be positive–semidefinite

$$\hat{C}_{\mathrm{X}}(\mathbf{w}) \geq 0 \tag{2.58}$$

for $\mathbf{w} \in R^n$.

A relevant result that appears in atmospheric data assimilation concerns the isotropic case with $n = 2$, that is, in R^2. In this case, $C_{\mathrm{X}}(\mathbf{r}) = C_{\mathrm{X}}(r)$, where $r = |\mathbf{r}|$. Introducing polar coordinates: $\mathbf{r} = (x, y) = (r\cos\theta, r\sin\theta)$ and $\mathbf{w} = (w\cos\varphi, w\sin\varphi)$; and recalling the change of variables in integrals means that we should calculate the determinant of the Jacobian matrix that corresponds to the transformation, that is

$$|\mathrm{Jac}(r,\theta)| \equiv \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r \tag{2.59}$$

where the notation $|.|$ is used for the determinant. In this way, using the fact that the integral over R^2 for any function $f(x, y)$ is transformed into an integral over the circle C as

$$\int_0^\infty \int_0^\infty f(x,y)\, dx\, dy = \int_0^\infty \int_0^{2\pi} f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta \tag{2.60}$$

(e.g., Apostol [4], pp. 479–485), the integral in (2.56b) becomes

$$\begin{aligned} \hat{C}_{\mathrm{X}}(\mathbf{w}) = \hat{C}_{\mathrm{X}}(w) &= \frac{1}{(2\pi)^2} \int_0^\infty C_{\mathrm{X}}(r)\, r \int_0^{2\pi} \cos(\mathbf{w}^T\mathbf{r})\, d\theta \\ &= \frac{1}{(2\pi)^2} \int_0^\infty C_{\mathrm{X}}(r)\, r \int_0^{2\pi} \cos[wr\cos(\theta-\varphi)]\, d\theta \end{aligned} \tag{2.61}$$

where the last equality is obtained by treating the inner product explicitly:

$$\begin{aligned} \mathbf{w}^T\mathbf{r} &= rw\cos\theta\cos\varphi + rw\sin\theta\sin\varphi \\ &= rw\cos(\theta-\varphi) \end{aligned} \tag{2.62}$$

where $w = |\mathbf{w}|$. Now performing the transformation, $\theta \to \theta + \varphi + \pi/2$, we have that $\cos(\theta - \varphi) \to -\sin\theta$, and therefore the integral of the expression above is independent of φ. This means that the result of the integral is also independent of φ, as should be the case for isotropic covariances. Introducing the Bessel function of order zero:

$$J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \cos(x\sin\theta)\, d\theta \tag{2.63}$$

(e.g., Arfken [5], pp. 579–580), we have that in two dimensions

$$\hat{C}_{\mathrm{X}}(w) = \frac{1}{2\pi} \int_0^\infty J_0(wr) C_{\mathrm{X}}(r) r\, dr \tag{2.64}$$

Utilizing the orthogonality of the Bessel function of order zero:

$$\int_0^\infty r J_0(wr) J_0(w'r)\, dr = \frac{1}{w}\delta(w - w') \tag{2.65}$$

(e.g., Arfken [5], p. 594), we obtain for the isotropic covariance function in two dimensions the formula:

$$C_{\mathrm{X}}(r) = 2\pi \int_0^\infty J_0(wr)\hat{C}_{\mathrm{X}}(w)w\,dw \tag{2.66}$$

It is interesting to mention that the concept of ergodicity can also be extended to spatial random fields. In an entirely analogous way to what can be done for stochastic processes, a spatial random field is said to be ergodic if its *spatial* mean and covariance coincide with its ensemble mean and covariance, respectively.

Exercises

1. (Problem 4.3, Meditch [103]) Assuming that three scalar stochastic process $\{\mathrm{x}(t), t \in T\}$, $\{\mathrm{y}(t), t \in T\}$ and $\{\mathrm{z}(t), t \in T\}$ are pairwise independent, show that they are not necessarily triplewise (simultaneously) independent.

2. Calculate the power spectrum for stationary processes having the following autocorrelation functions:

   (a) Gaussian pulse: $\Gamma(\tau) = \sigma^2 e^{-\tau^2/T^2}$

   (b) Damped cosine wave: $\Gamma(\tau) = \sigma^2 e^{-\beta|\tau|}\cos\omega_0\tau$

   (c) Triangular pulse:

   $$\Gamma(\tau) = \begin{cases} 1 - |\tau|\,, & \text{for } |\tau| \le 1 \\ 0\,, & \text{otherwise} \end{cases}$$

3. (Problem 2.17, Brown [19]) The stationary process $x(t)$ has mean $\mu = const.$ and an autocorrelation function of the form

   $$\Gamma(t, t+\tau) \;=\; \Gamma(\tau) \;=\; \sigma^2 e^{-\tau^2/T^2}$$

   Another process $y(t)$ is related to $x(t)$ by the deterministic equation

   $$y(t) \;=\; a\,x(t) + b$$

   where a and b are known constants.

   (a) What is the auto–correlation function for $y(t)$?

   (b) What is the cross–correlation function $\Gamma_{xy}(\tau)$?

4. (Problem 2.20, Brown [19]) Two random processes are defined by

   $$\begin{aligned} x(t) &= a\,\sin(\omega t + \theta) \\ y(t) &= b\,\sin(\omega t + \theta) \end{aligned}$$

   where θ is a random variable with uniform distribution between 0 and 2π, and ω is a known constant. The coefficients a and b are both normal random variables $\mathcal{N}\{0, \sigma^2\}$ and are correlated with a correlation coefficient ρ. What is the cross–correlation function $\Gamma_{xy}(\tau)$? (Assume a and b are independent of θ.)

5. Show that the following are admissible candidates for a covariance function:

   (a) $C_{\rm X}(\mathbf{r}) = a\delta(\mathbf{r})$, for $a > 0$ and $\mathbf{r} \in R^n$

   (b) $C_{\rm X}(x,y) = C_{\rm X}(r = |x-y|) = \pi \exp(-r^2)$, for $x, y \in R^1$. (Hint: In this case, the proof can be obtained by either showing that (2.57) is true, or showing that (2.58) is satisfied. Use (2.58) and expand $\exp(2xy)$ in Taylor series.)

6. Show that in R^3 the isotropic spectral density function can be expressed as

$$\hat{C}_{\rm X}(w) = \frac{1}{2\pi^2} \int_0^\infty \frac{\sin(wr)}{w} C_{\rm X}(r) r \, dr$$

   and that consequently the corresponding covariance function is given by

$$C_{\rm X}(r) = 4\pi \int_0^\infty \frac{\sin(wr)}{r} \hat{C}_{\rm X}(w) w \, dw$$

7. (Problem 7.14, Maybeck [101]) In Monte Carlo analyses and other type of system simulations (e.g., non–identical twin experiments), it is often desired to generate samples of a discrete–time white Gaussian noise vector process, described by mean zero and covariance

$$\mathcal{E}\{\mathbf{w}_k \mathbf{w}_k^T\} = \mathbf{Q}_k$$

   with $\mathbf{Q}_k$ nondiagonal. Independent scalar white Gaussian noises can be simulated readily through use of pseudorandom codes (as we have seen in our first computer assignment), but the question remains, how does one properly provide for cross–covariances of the scalar noises?

   (a) Let $\mathbf{v}_k$ be a vector process composed of independent scalar white Gaussian noises of zero mean and unit variance:

$$\mathcal{E}\{v_{j,k}\} = 0 \quad \mathcal{E}\{v_{j,k}^2\} = 1 \quad \text{for } k = 1, 2, \ldots$$

   where $v_{j,k}$ is the j–th component of $\mathbf{v}_k$, at time t_k. Show that

$$\mathbf{w}_k = \mathbf{L}_k \mathbf{v}_k \quad \text{for all } k$$

   properly models the desired characteristics. The matrix $\mathbf{L}_k$ above corresponds to the Cholesky lower triangular square root of $\mathbf{Q}_k$, that is, $\mathbf{Q}_k = \mathbf{L}_k \mathbf{L}_k^T$. Notice, that if the Cholesky upper triangular square root had been used instead, the corresponding expression for generating $\mathbf{w}_k$ would be

$$\mathbf{w}_k = \mathbf{U}_k \mathbf{v}_k \quad \text{for all } k$$

   where, in this case, $\mathbf{Q}_k = \mathbf{U}_k \mathbf{U}_k^T$.

   (b) If $\mathbf{U}_k$ and $\mathbf{D}_k$ are the U–D factors of $\mathbf{Q}_k$, that is, if $\mathbf{Q}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^T$, where $\mathbf{U}_k$ are upper triangular and unitary matrices and $\mathbf{D}_k$ are diagonal matrices, show that,

$$\mathbf{w}_k = \mathbf{U}_k \mathbf{u}_k \quad \text{for all } k$$

also provides the desired model if $\mathbf{u}_k$ is a vector process composed of independent scalar white Gaussian noises of mean zero and variance:

$$\mathcal{E}\{u_{j,k}^2\} = d_{j,j;k}\,,$$

where $u_{j,k}$ is the j–th component of $\mathbf{u}_k$, at time t_k, and $d_{j,j;k}$ is the (j,j) element of $\mathbf{D}_k$, at time t_k (i.e., the j–th element along its diagonal).

8. *Computer Assignment:* We want to use the results of the previous problem to perform Monte Carlo experiments for a given correlation and/or covariance structure. As a preparation for that, in this problem we are going to create a Matlab function that generates a homogeneous correlation on a grid defined over R^1 and examine some of its properties. Let us start by consider the interval $(-L_x, L_x]$, and let us divide it in a uniform grid of J points. Consider also the homogeneous and isotropic, Gaussian correlation function in $R^1 \times R^1$, that is,

$$Q(x,y) = Q(r = |x-y|) = \exp(-\frac{1}{2}(x-y)^2/L_d^2)$$

where $r = |x-y|$ is the distance between any two points in the domain, and L_d is the (de)correlation length. Therefore the points in the discrete domain can be defined as

$$x_j = j\Delta x$$

where $\Delta x = 2L_x/J$, for $j \in \{-J/2+1, J/2\}$, and the elements of the homogeneous, isotropic correlation matrix $\mathbf{Q}$ are given by

$$Q_{ij} = Q(x_i, y_j)$$

   (a) Construct a Matlab function that returns the covariance matrix $\mathbf{Q}$, given the half–length of the domain L_x, the number of grid points J, and the (de)correlation length L_d. For $(L_x, L_d, J) = (1, 0.2, 32)$, compute $\mathbf{Q}$ using this function. Make a contour plot of the correlation array. (Note: A real convenient way of generating this matrix in Matlab is using the intrinsic function meshgrid.)

   (b) For the parameters of the previous item, plot the correlation function at the following two specific locations: $x_j \in \{0, L_x\}$.

   (c) Is the $\mathbf{Q}$ obtained above an *acceptable* correlation matrix? Explain it. (Hint: Check its eigenvalues.)

   (d) From the figures constructed in the previous items, we see that the correlation decreases very quickly toward values that are nearly zero. (You can actually print out the values in $\mathbf{Q}$ to check it further). It could be computationally advantageous, particularly to reduce storage requirements, to approximate this correlation "function" (matrix) by one that neglects correlation values beyond a certain cut–off length L_c. In this way, only the elements of the matrix corresponding to $|r| \leq L_c$ would need to be stored. Without worrying about the storages savings, modify the function of item (a), to construct a new matrix $\mathbf{Q}_c$, by replacing the values of $\mathbf{Q}$ for which $|r| > L_c$ by zero. Using the same parameters as in item (a), and a cut–off value of $L_c = 3L_d$, make a contour plot of the resulting correlation structure. Also, repeat item (b).

(e) A visual comparison of the plots in items (a) and (b) with those of item (d) seem to indicate that our approximation is fairly reasonable. Is the correlation of the previous item, an acceptable correlation matrix?

9. *Computer Assignment*: The result obtained in the last item of the previous exercise makes us wonder what is the correct way of constructing a correlation field that has the structure that we want, but is zero beyond a certain correlation length. The procedure to generate what are called compact–support correlation functions is through the use of convolution of functions. (Note: see Gaspari and Cohn (1996) for the details on constructing these correlation functions in R^2 and R^3, that are of primary importance in modeling covariances for data assimilation). Another way of looking at convolution of functions is to think on the Hadamard product for the case of matrices. Without getting into the mathematical details, this problem has the intention to guide you though the steps of building an actual correlation matrix for the function of the previous problem. Consider then the compact–support triangular correlation function discussed earlier in the text:

$$T(x,y) == \begin{cases} 1 - |x-y|/L_c\,, & \text{for } |x-y| \leq L_c \\ 0 & \text{otherwise} \end{cases}$$

Then perform the following tasks:

(a) Repeat items (a)–(c) of the previous exercise, but now for the compact–support function $T(r) = T(|x-y|)$. (Note: The matrix $\mathbf{T}$ has elements $T_{ij} = T(x_i, y_j)$.)

(b) Construct a matrix $\bar{\mathbf{Q}}$ as the Hadamard product of the matrix $\mathbf{Q}$, of item (a) in the previous exercise, and $\mathbf{T}$ from the previous item, corresponding to the function $T(r)$. That is, let $\bar{\mathbf{Q}}$ be given by

$$\bar{\mathbf{Q}} \equiv \mathbf{Q} \circ \mathbf{T} = [Q_{ij} T_{ij}]$$

(Note: Matlab does the Hadamard product trivially). Make a contour plot $\bar{\mathbf{Q}}$, and repeat item (c) of the previous exercise.

(c) To get yet another visual representation of what the correlations from $\mathbf{Q}$, $\mathbf{T}$ and $\bar{\mathbf{Q}}$ are like, plot the correlation functions obtained from these three matrices at point $x = 0$. (Please, have all three curves on the same frame).

10. *Computer Assignment:* Using the Matlab function created in item (a) of Exercise 8 (i.e., without a cut–off length), let us apply the results of Exercise 7 to understand better what correlated noise actually is.

(a) Create a Matlab function that performs a Monte Carlo experiment given the number of samples. We want the output of this function to be the sampled correlation matrix, obtained from a weighted sum of outer products of the vectors $\mathbf{w}_k$ of Exercise 7. To obtain the Cholesky decomposition of the correlation matrix $\mathbf{Q}$ of Exercise 8, use the Matlab function chol. Make contour plots for the three sampled correlation matrices obtained by using 100, 1000, 10000 samples.

(b) Now using the identity matrix, in the 32–dimensional space of Exercise 8, perform a Monte Carlo run, with 1000 samples, assuming this identity matrix is the correlation matrix of interest. Make a contour plot of the resulting sampled correlation matrix. Compare this result with those obtained in the previous item. In particular, explain the meaning of using an identity correlation matrix.