Chapter 2 Karhunen-Loeve expansion

2.1 Introduction

The mathematical series representation of a random process using the Karhunen-Loeve (K-L) expansion is first presented. Important considerations in using this expansion to generate stationary Gaussian, non-Gaussian and non-stationary processes are discussed in Sections 2.3 to 2.5. A key feature of K-L expansion in simulation is the spectral decomposition of the covariance function describing the process. This involves solving the Fredholm integral equation either analytically or numerically, as presented in Sections 2.6 and 2.7, and illustrated with an example based on the first-order Markov process. Section 2.8 gives a brief introduction of the most widely used spectral representation method for simulation. This chapter concludes with Section 2.9 showing that the K-L expansion converges to the well-known spectral representation method in the case of stationary process with a long record.

2.2 K-L Series Representation of a Random Process

Consider a random process $\varpi(x,\theta)$ defined on a probability space (Ω, **A**, P) and indexed on a bounded domain D. Assume that the process has a mean $\overline{\varpi}(x)$ and a finite variance σ^2 or $E[\varpi(x,\theta)-\overline{\varpi}(x)]^2$, that is bounded for all $x \in D$. The process can be expressed as (Van Trees, 1968):

$$\varpi(x,\theta) = \overline{\varpi}(x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i}\,\xi_i(\theta) f_i(x) \tag{2.1}$$

in which λ_i and $f_i(x)$ are the eigenvalues and eigenfunctions of the covariance function $C(x_1, x_2)$. By definition, $C(x_1, x_2)$ is bounded, symmetric and positive definite. Following Mercer's Theorem (Van Trees, 1968), it has the following spectral or eigen-decomposition:

$$C(x_1, x_2) = \sum_{i=1}^{\infty} \lambda_i f_i(x_1) f_i(x_2) \tag{2.2}$$

and its eigenvalues and eigenfunctions are the solution of the homogenous Fredholm integral equation of the second kind given by:

$$\int_D C(x_1, x_2) f_i(x_2) dx_2 = \lambda_i f_i(x_1) \tag{2.3}$$

Equation (2.3) arises from the fact that the eigenfunctions form a complete orthogonal set satisfying the equation:

$$\int_D f_i(x) f_j(x) dx = \delta_{ij} \tag{2.4}$$

where δ_{ij} is the Kronecker-delta function.

The parameter $\xi_i(\theta)$ in equation (2.1) is a set of uncorrelated random variables which can be expressed as:

$$\xi_i(\theta) = \frac{1}{\sqrt{\lambda_i}} \int_D [\varpi(x, \theta) - \overline{\varpi}(x)] f_i(x) dx \tag{2.5}$$

with mean and covariance function given by:

$$E[\xi_i(\theta)] = 0 \tag{2.6a}$$

$$E[\xi_i(\theta) \xi_j(\theta)] = \delta_{ij} \tag{2.6b}$$

The series expansion in equation (2.1), referred to as the K-L expansion, provides a second-moment characterization in terms of uncorrelated random variables and

deterministic orthogonal functions. It is known to converge in the mean square sense for any distribution of $\varpi(x,\theta)$ (Van Trees, 1968). For practical implementation, the series is approximated by a finite number of terms, say M, giving:

$$\varpi_M(x,\theta) = \overline{\varpi}(x) + \sum_{i=1}^{M} \sqrt{\lambda_i}\,\xi_i(\theta) f_i(x) \tag{2.7}$$

The corresponding covariance function is given by:

$$C_M(x_1,x_2) = \sum_{i=1}^{M} \lambda_i f_i(x_1) f_i(x_2) \tag{2.8}$$

Ghanem and Spanos (1991) have shown this truncated series to be optimal; that is, the mean square approximation error is minimised.

2.3 Gaussian Process

If $\varpi(x,\theta)$ is further restricted to a zero-mean Gaussian process, then the appropriate choice of $\{\xi_1(\theta),\xi_2(\theta)\Lambda\}$ is a vector of zero-mean uncorrelated Gaussian random variables. The variables can be generated by available established pseudo–random number generating subroutines and then multiplied by the eigenfunctions and eigenvalues derived from eigen-decomposition [equation (2.3)] of the target covariance model. The value of M is governed by the accuracy of the eigen-pairs in representing the covariance function. In view of the fact that equation (2.7) is a summation of Gaussian random variables, the process would be Gaussian distributed. Taking the expectation of equation (2.7), the mean of the simulated process is $\overline{\varpi}(x)$. The second-order statistics of the target process are also reproduced by the simulated process since the ensemble covariance function of the simulated random process $\hat{\varpi}(x,\theta)$ can be written as:

$$\hat{C}_M(x_1,x_2) = E[\{\hat{\varpi}_M(x_1,\theta)-\overline{\varpi}(x)\}\{\hat{\varpi}_M(x_2,\theta)-\overline{\varpi}(x)\}]$$

$$= E[\sum_{i=1}^{M}\sqrt{\lambda_i}\xi_i(\theta)f_i(x_1)\sum_{j=1}^{M}\sqrt{\lambda_j}\xi_j(\theta)f_j(x_2)]$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{M}\sqrt{\lambda_i}\sqrt{\lambda_j}f_i(x_1)f_j(x_2)E[\xi_i(\theta)\xi_j(\theta)] \tag{2.9}$$

By virtue of equation (2.6b),

$$\hat{C}_M(x_1,x_2) = \sum_{i=1}^{M}\lambda_i f_i(x_1)f_i(x_2) = C_M(x_1,x_2) \tag{2.10}$$

As $M \to \infty$, equation (2.10) converges to equation (2.2).

The key to K-L expansion for Gaussian processes is therefore to obtain the eigenvalues and eigenfunctions of the covariance function by solving the homogeneous Fredholm integral equation of the second kind in equation (2.3). The integral equation can be solved analytically only under special circumstances. In most cases, analytical solution of the integral equation is not tractable and numerical method is the only recourse. To apply K-L as a general simulation tool, it is therefore very important to study the practicality and accuracy of solving the Fredholm integral equation numerically. Detailed study for solving the Fredholm integral equation evolved in K-L expansion is presented in Sections 2.6 and 2.7 and Chapter 5.

2.4 Non-Gaussian Process

Simulation of non-Gaussian random processes using the K-L expansion basically requires: (a) solving the homogeneous Fredholm integral equation of the second kind in equation (2.3) to obtain the eigenvalues and eigenfunctions of the covariance function and

(b) selecting uncorrelated random variables $\xi_i(\theta)$ such that the finite sum [equation (2.7)] produces the desired non-Gaussian distribution. Note that for a random process $\varpi(x,\theta)$ with arbitrary distributions, the probability distribution of $\xi_i(\theta)$ may be determined using equation (2.5). The integral can be evaluated numerically using any of the quadrature schemes involving linear combination of the integrand values at discrete points within the integration intervals. However, the integrand is obviously an unknown. Hence, an iterative procedure is required which will be presented in Chapter 4.

2.5 Non-stationary Process

Note that in the above discussion, the covariance function need not be limited to stationary processes. Hence the discussions in Sections 2.3 and 2.4 are valid for second-order non-stationary processes. Details on simulating non-stationary Gaussian processes will be given in Chapter 3 while that for non-stationary non-Gaussian processes in Chapter 4.

2.6 Analytical Solution of Fredholm Integral Equation

For some classes of covariance function, equation (2.3) can be differentiated twice with respect to x_1 analytically. The resulting differential equation is then solved analytically. To satisfy the boundary conditions, the solution is substituted into equation (2.3) to yield the eigenvalues (Van Trees, 1968). The commonly used random processes that fall into this category are:

(1) Stationary process with rational spectra. An example in this class is the first-order Markov process that has the covariance function:

$$C(x_1, x_2) = \sigma^2 e^{-|x_1 - x_2|/b} \tag{2.11}$$

where σ^2 is the variance and b is the correlation parameter.

(2) Band-limited stationary process. The spectrum of such random process is not rational but the procedure described above can be used.

(3) Special class of non-stationary processes. The Wiener-Levy and the Brown-Bridge random processes are two known cases where equation (2.3) can be solved by converting to a boundary-value problem.

The procedure will be illustrated using two cases: (a) first-order Markov process and (b) non-stationary Wiener process.

Case 1:

As an illustration, the integral equation for the first-order Markov process given by equation (2.11) will be solved analytically. The process is assumed to be defined over the interval [-*a*, *a*]. The eigenfunctions and eigenvalues of equation (2.11) are the solutions to the following integral equation (Van Trees, 1968):

$$\int_{-a}^{a} \sigma^2 e^{-c|x_1 - x_2|} f(x_2) dx_2 = \lambda f(x_1) \tag{2.12}$$

where c=1/b. Equation (2.12) can be written as:

$$\int_{-a}^{x_1} \sigma^2 e^{-c(x_1 - x_2)} f(x_2) dx_2 + \int_{x_1}^{a} \sigma^2 e^{-c(x_2 - x_1)} f(x_2) dx_2 = \lambda f(x_1) \tag{2.13}$$

Differentiating equation (2.13) with respect to x_1 and re-arranging gives:

$$\lambda f'(x_1) = -\sigma^2 c e^{-cx_1} \int_{-a}^{x_1} e^{cx_2} f(x_2) dx_2 + \sigma^2 c e^{cx_1} \int_{x_1}^{a} e^{-cx_2} f(x_2) dx_2 \tag{2.14}$$

Differentiating once more with respect to x_1, the following equation is obtained:

$$\lambda f''(x) = (\lambda c^2 - 2\sigma^2 c) f(x) \tag{2.15}$$

Introducing the new variable:

$$\omega^2 = \frac{-\lambda c^2 + 2\sigma^2 c}{\lambda} \tag{2.16}$$

Equation (2.15) becomes:

$$f''(x) + \omega^2 f(x) = 0 \tag{2.17}$$

To find the boundary conditions associated with the differential equation (2.17), equations (2.13) and (2.14) are evaluated at x= -a and x = +a. After re-arrangement, the boundary conditions become:

$$cf(a) + f'(a) = 0 \tag{2.18}$$

$$cf(-a) - f'(-a) = 0 \tag{2.19}$$

Thus, the integral equation given by equation (2.12) is transformed into the ordinary differential equation and boundary conditions given by equations (2.17), (2.18) and (2.19).

Equation (2.17) is solved only for the range $\omega^2 > 0$, the solution being given by:

$$f(x) = a_1 \cos(\omega x) + a_2 \sin(\omega x) \tag{2.20}$$

Further, applying the boundary conditions specified by equations (2.18) and (2.19) gives

$$a_1(c - \omega \tan(\omega a)) + a_2(\omega + c \tan(\omega a)) = 0 \tag{2.21a}$$

$$a_1(c - \omega \tan(\omega a)) - a_2(\omega + c \tan(\omega a)) = 0 \tag{2.21b}$$

Non-trivial solutions exist only if the determinant of equation (2.21) is equal to zero. This results in the following transcendental equations:

$$c - \omega \tan(\omega a) = 0 \tag{2.22}$$

$$\omega^* + c \tan(\omega^* a) = 0 \tag{2.23}$$

The eigenvalues are

$$\lambda_i = \frac{2\sigma^2 c}{\omega_i^2 + c^2} \text{ , i = 1,3,5 } \ldots \tag{2.24}$$

and

$$\lambda_i = \frac{2\sigma^2 c}{\omega^{*2}_i + c^2} \text{ , i = 2,4,6...} \tag{2.25}$$

where ω and ω^* satisfy equations (2.22) and (2.23) respectively.

The resulting eigenfunctions are

$$f_i(x) = \frac{\cos(\omega_i x)}{\sqrt{a + \frac{\sin(2\omega_i a)}{2\omega_i}}} \text{ , i = 1,3,5 } \ldots \tag{2.26}$$

and

$$f_i(x) = \frac{\sin(\omega^*_i x)}{\sqrt{a - \frac{\sin(2\omega^*_i a)}{2\omega^*_i}}} \text{ , i = 2,4,6...} \tag{2.27}$$

for odd *i* and even *i* respectively.

Case 2:

For the non-stationary Wiener process, the covariance function is given by

$$C(x_1, x_2) = \sigma^2 \min(x_1, x_2) \tag{2.28}$$

In this case equation (2.3) becomes

$$\lambda f(x_1) = \int_0^a C(x_1, x_2) f(x_2) dx_2 \text{ , } 0 \le x \le a \tag{2.29}$$

Substituting equation (2.28) into equation (2.29)

$$\lambda f(x_1) = \sigma^2 \int_0^x x_2 f(x_2) dx_2 + \sigma^2 x_1 \int_x^a f(x_2) dx_2 \tag{2.30}$$

Proceeding as in case 1, differentiating equation (2.30) gives

$$\lambda f'(x_1) = \sigma^2 \int_x^a f(x_2) dx_2 \tag{2.31}$$

Differentiating equation (2.31) again yields

$$\lambda f''(x_1) = -\sigma^2 f(x_1) \tag{2.32}$$

The boundary conditions become:

$$f'(0) - f'(a) = 0 \tag{2.33}$$

$$f(a) - f(0) = af'(0) \tag{2.34}$$

The eigenvalues are

$$\lambda_i = \frac{\sigma^2 a^2}{(n-\frac{1}{2})^2 \pi^2}, \; i=1,2,\ldots \tag{2.35}$$

and the eigenfunctions are

$$f_i(x) = (\frac{2}{a})^{\frac{1}{2}} \sin[(i-\frac{1}{2})\frac{\pi}{a}x], \; 0 \le x \le a \tag{2.36}$$

2.7 Numerical Solution of Fredholm Integral Equation

For random processes where analytical solution of the integral equation for K-L expansion is not tractable, numerical solution is necessary. There are two commonly used classes of numerical methods:

(1) Integration formulae based methods. Examples of such simple and direct methods include the quadrature method and product-integration method (Gutierrez et al 1992).

(2) Expansion methods. Expansion methods are best suited for kernels that are analytically defined. Each eigenfunction is approximated by a linear combination of chosen basis functions. This will result in an error in equation (2.3), which can be

minimized using methods such as Galerkin, collocation or Rayleigh-Ritz. In this study, the Galerkin method is first used as the error is minimized over the entire domain rather than at pre-selected collocated points. For the case where the kernel $C(x_1, x_2)$ is symmetric, bounded and positive definite, Galerkin method and the Rayleigh-Ritz method are in fact identical. The numerical solution of Fredholm integral equation using the Galerkin method is briefly summarized below:

a. Select N basis functions, say ψ_i, $i = 0, \ldots, N\text{-}1$

For polynomial basis function, Legendre polynomials up to (N-1)th degree on the interval [-1, 1] are defined as:

$$\psi_0 = 1, \quad \psi_1 = x, \quad \psi_i = \frac{2i-1}{i} x\psi_{i-1} - \frac{i-1}{i} x\psi_{i-2}, \quad i = 2, \ldots P \tag{2.37}$$

which can be normalized and transformed to the interval [-a, a] by

$$\psi_i(x) = \sqrt{\frac{2i+1}{2a}} \psi_i\left(\frac{x}{a}\right) \tag{2.38}$$

Numerical solutions obtained using this set of basis functions will be referred to as the Pth polynomial solutions, where $P = N - 1$.

For trigonometric basis function, functions up to (N-1)/2 harmonics on the interval [-a, a] are defined as:

$$\psi_0 = 1, \quad \psi_1 = \cos\left(\frac{\pi x}{a}\right), \quad \psi_2 = \sin\left(\frac{\pi x}{a}\right), \ldots$$

$$\psi_{2i-1} = \cos\left(\frac{i\pi x}{a}\right), \quad \psi_{2i} = \sin\left(\frac{i\pi x}{a}\right), \quad i = 2, \ldots T \tag{2.39}$$

Numerical solutions obtained using this set of basis functions will be referred to as T harmonics solutions, where $T = (N - 1)/2$.

For wavelet basis function, the Haar wavelets are used and defined in Section 5.2.2.

b. Approximate each eigenfunction as a linear combination of these basis functions:

$$f_k(x) = \sum_{i=0}^{N-1} d_i^{(k)} \psi_i(x) = \Psi^{\mathrm{T}}(x) D^{(k)} \tag{2.40}$$

where $d_i^{(k)}$ are constant coefficients for the ith eigenfunction, $D^{(k)}$ is a Nx1 coefficient vector and $\Psi(x)$ is the vector of basis functions.

c. Substitute equation (2.40) into equation (2.3) and set the error in equation (2.3) to be orthogonal to the each basis function to obtain the following equation:

$$\sum_{i=0}^{N-1} d_i^{(k)} \left[\int_D \int_D C(x_1, x_2) \psi_i(x_1) \psi_j(x_2) \mathrm{d}x_1 \mathrm{d}x_2 \right] - \lambda_k \sum_{i=0}^{N-1} d_i^{(k)} \left[\int_D \psi_i(x_2) \psi_j(x_2) \mathrm{d}x_2 \right] = 0 \tag{2.41}$$

or $AD^{(k)} = \lambda_k B D^{(k)}$ (2.42)

in which the components in all the N x N matrices are:

$$A_{kj} = \int_D \int_D C(x_1, x_2) \psi_k(x_1) \psi_j(x_2) dx_1 dx_2 \tag{2.43}$$

$$B_{kj} = \int_D \psi_k(x_2) \psi_j(x_2) dx_2 \tag{2.44}$$

The integrals in equation (2.43) and equation (2.44) are evaluated numerically using the function D01FCF in MATLAB toolbox NAG (The Math Works Inc., 1995), which is based on multi-dimensional adaptive quadrature over a hyper-rectangle (i.e. with constant limits). The routine returns an estimate of a multi-dimensional integral, and also an estimate of the relative error. The user sets the relative accuracy required and supplies the integrand as a function subprogram. In this study, the relative accuracy is set to 10^{-4}. The

generalized algebraic eigenvalue problem of equation (2.42) may be solved to obtain the eigenvalues λ_i and coefficients $d_i^{(k)}$.

For most applications, use of classical bases (e.g. polynomial basis and trigonometric bases) produces a dense A matrix (most of its elements are nonzero). There may be substantial cost in computing all elements A_{ij}. Even after the elements are computed the solution of the eigenvalue problem [equation (2.42)] requires order $O(N^3)$, if a direct scheme such as QR or Gauss elimination is used. If equation (2.42) is solved by an iterative method which requires an order $O(N^2)$ on each iteration, the number of the iteration may be large, depending on the conditioning of the original integral equation. For large scale problems, these cost are often prohibitive.

Alternatively, the wavelet-Galerkin method is employed, which is based on the same general framework except that: (a) basis functions ψ_i are wavelets, and (b) integral in equation (2.43) is 2-D wavelet transform that does not require quadrature. Details for wavelet solution are described in Chapter 5.

In the following examples, 5^{th} to 9^{th} degree polynomials are chosen as basis functions, depending on the number of terms in K-L expansion [M in equation (2.8)]. The Galerkin scheme described above is known to produce lower bound solutions for the eigenvalues and is more accurate for eigenvalues than eigenfunctions (Ghanem and Spanos, 1991).

To illustrate, the first-order Markov process discussed in Section 2.6 is used with a=1. Legendre polynomials [equation (2.37)] on the interval [-1,1] are selected as a set of basis functions to approximate the eigenfunctions. Following the procedure described above, the numerical results for the eigenvalues and eigenfunctions are computed. Figure 2.1 shows the good agreement between the numerical and exact analytical eigenvalues. Figure

2.2 shows the first six eigenfunctions obtained using the algorithm described above. Note that the approximate results oscillate around exact results, which is typical of Galerkin approach. The Galerkin scheme described above can be shown to be equivalent to a variational treatment of the problem. This property ensures that the computed eigenvalues are lower bound, with monotonic convergence to the exact values.

2.8 Spectral Representation Method

The spectral representation method is a widely used simulation technique for simulation of stationary Gaussian processes (Grigoriu, 1993a; Shinozuka, 1970; Shinozuka, 1972; Shinozuka and Deodatis, 1988; Shinozuka and Deodatis, 1991), non-stationary Gaussian processes (Grigoriu, 1993c) and stationary non-Gaussian processes (; Deodatis et al., 1998;Yamazaki and Shinozuka, 1988). To compare the K-L expansion method with the spectral representation method, a brief introduction of the spectral method is described below.

2.8.1 Stationary Gaussian Processes

The spectral representation method requires that the power spectral density of the simulated random process be known. Consider a zero-mean, real valued, stationary random process with theoretical covariance function $C(\tau)$ and one-sided power spectral density $G(\omega)$. The parameter τ is the interval between two points in the process and ω is the circular frequency. Two simulation models are given as follows:

Sine-cosine series representation:

$$\varpi_M(x,\theta) = \sum_{i=1}^{M} \sigma_i [V_i(\theta)\cos\omega_i x + W_i(\theta)\sin\omega_i x] \tag{2.45}$$

where $V_i(\theta)$ and $W_i(\theta)$ are independent Gaussian random variables with zero means and unit variances.

Cosine series representation:

$$\varpi_M(x,\theta) = \sum_{i=1}^{M} \sqrt{2}\sigma_i \cos[\omega_i x + U_i(\theta)] \tag{2.46}$$

where $U_i(\theta)$ are random variables uniformly distributed over $(0, 2\pi)$.

The parameters common to both models are described in the following. The variance of the truncated process is the sum of the shaded strips associated with each discretized frequency ω_i

$$\sigma^2 = \sum_{i=1}^{M} \sigma_i^2 = \sum_{i=1}^{M} G(\omega_i)\Delta\omega \tag{2.47}$$

where

$$\omega_i = i\Delta\omega - \Delta\omega/2 \tag{2.48}$$

$$\Delta\omega = \omega_{\max}/M \tag{2.49}$$

and $\omega_{\max}$ is the maximum frequency included in the simulation function, M is the number of intervals between 0 and $\omega_{\max}$, and $\Delta\omega$ is the discretization width. Figure 2.3 shows the discretization of the one-sided power spectral density $G(\omega)$ and the derivation of various key parameters. The simulation procedure using equation (2.46) involves generating M random variables.

The ensemble variance will approach the target variance, which is given by the total area under $G(\omega)$, when both $\omega_{\max} \to \infty$ and $M \to \infty$. Depending on the tolerable error ε, $\omega_{\max}$ is selected based on the relationship

$$\int_0^{\omega_{\max}} G(\omega)d\omega = (1-\varepsilon)\int_0^{\infty} G(\omega)d\omega \tag{2.50}$$

In this study, cosine series representation is used because it has been used extensively in application and it depends on only M random variables rather than $2M$ variables as for sine-cosine model.

2.8.2 Non-stationary Gaussian Processes

The generation of realizations of the non-stationary Gaussian process is currently based on several probabilistic models: autoregressive series with time dependent coefficients, inhomogeneous filtered Poisson processes, oscillatory processes, and processes with modulated amplitude and frequency (Deodatis and Shinozuka, 1988; Grigoriu et al, 1988; Shinozuka and Sato, 1967; Shinozuka and Jan, 1972). The calibration of these models to a target non-stationary process can be complex because it requires finding one or more functions of time and may encounter theoretical difficulties because the models cannot represent arbitrary non-stationary processes. The alternative model based on spectral method is proposed by Grigoriu for generating general realizations of non-stationary Gaussian processes. The model consists of trigonometric polynomials with random Gaussian coefficients that are correlated. The random trigonometric polynomial is

$$\varpi_M(x,\theta) = \frac{J_0}{2} + \sum_{i=1}^{M}[J_i(\theta)\cos\omega_i x + K_i(\theta)\sin\omega_i x] \tag{2.51}$$

defined on [0,a] in which $\omega_1 = 2\pi / a, \omega_i = i\omega_i$; i=1, 2 , ...M, and

$$J_i(\theta) = \int_0^a \varpi_M(x,\theta)\cos\omega_i x dx \tag{2.52}$$

$$K_i(\theta) = \int_0^a \varpi_M(x,\theta)\sin\omega_i x dx \tag{2.53}$$

are random coefficients which are correlated. Details for the simulation algorithm are in (Grigoriu, 1993c). The results from this spectral representation will be compared with those from K-L expansion in Section 3.3.4.

2.8.3 Stationary Non-Gaussian Processes

A method for digitally generating sample functions of multidimensional non-Gaussian stationary processes has been developed by Yamazaki and Shinozuka. (1988). First, the two-sided target spectral density $S_{BB}^{(T)}(\omega)$ and one dimensional distribution function of the non-Gaussian process $F_B(B)$ are specified. $F_B(B)$ denotes a one dimensional distribution function of a non-Gaussian stochastic field, B , with zero mean and unit variance. The method is described as follows:

Step 1: The first step of the method is assume the spectral density $S_{GG}^{(1)}(\omega)$ of the Gaussian process to be

$$S_{GG}^{(1)}(\omega) = S_{BB}^{(T)}(\omega) \tag{2.54}$$

and generate sample function of the Gaussian process $G^{(1)}(x)$ by equation (2.46).

Step2: The simulated process is transformed into non-Gaussian process by translation

$$B^{(i)}(x_k) = F_B^{-1}\{F_G[G^{(i)}(x_k)]\} \tag{2.55}$$

If the number of data points is large enough, equation (2.55) generate sample functions that follow closely the target non-Gaussian distribution function. However, the spectral density of the translated process $S_{BB}^{(1)}(\omega)$ does not coincide with the target spectral density $S_{BB}^{(T)}(\omega)$ because the transformation represented by equation (2.55) is nonlinear.

Step 3: Therefore, an iterative algorithm is introduced in order to satisfy the target spectral density function. The spectral density for generating sample functions of the Gaussian process in the (i+1)th iteration is assumed to be

$$S_{GG}^{(i+1)}(\omega)=\frac{S_{GG}^{(i)}(\omega)}{S_{BB}^{(i)}(\omega)}S_{BB}^{(T)}(\omega) \tag{2.56}$$

Following the same procedure used in the first iteration, the Gaussian process of the ith iteration is generated using the updated power spectral density and the same random phase angles. The procedure is iterated until both the target spectral density $S_{BB}^{(T)}(\omega)$ and one dimensional distribution function of the non-Gaussian process $F_B(B)$ are achieved. However, the procedure cannot converge to the prescribed power spectral density for highly non-Gaussian processes (Deodatis, 2000).

2.9 Relation between K-L Expansion and Spectral Representation

Consider the case where the process to be represented is stationary and the observation interval is infinite. Equation (2.3) becomes:

$$\lambda f(x_2)=\int_{-\infty}^{\infty}C(x_2-x_1)f(x_1)dx_1 \qquad -\infty<x_2<\infty \tag{2.57}$$

Equation (2.57) is analogous to the convolution integral with input $f(x)$ and impulse response $C(x_2 - x_1)$ such that the output is $f(x_2)$ with a change in gain given by λ. It is well known that this requirement can be met by making $f(x) = e^{i\omega x}$. Substituting this into equation (2.57) yields:

$$\lambda = \int_{-\infty}^{\infty} C(x_2 - x_1) e^{-i\omega(x_2 - x_1)} dx_1 = S(\omega) \tag{2.58}$$

Thus the eigenvalue is given by the two-sided power spectral density $S(\omega)$, which is continuous. For the case of a finite length process defined in [-a,a] where a is large, it can be shown that (Van Trees, 1968):

$$\lambda_k \approx S(\omega_k) = S(\frac{\pi k}{a}) \tag{2.59}$$

$$f_k(x) \approx \frac{1}{\sqrt{2a}} e^{i(\pi k/a)x} \tag{2.60}$$

Substituting equations (2.59) and (2.60) into equations (2.7) and (2.8) gives:

$$\varpi(x,\theta) = \overline{\varpi}(x) + \sum_{k=1}^{M} \sqrt{\frac{1}{2a} S(\frac{\pi k}{a})}\, e^{i(\pi k/a)x}\, \xi_k(\theta) \tag{2.61}$$

$$C(x_1, x_2) = \sum_{k=1}^{M} \frac{1}{a} S(\frac{\pi k}{a}) \cos\left(\frac{\pi k(x_1 - x_2)}{a}\right) \tag{2.62}$$

which is the commonly used spectral representation used for the simulation of stationary Gaussian random process (Grigoriu, 1993; Shinozuka and Deodatis, 1991). Thus for large a, the K-L expansion reduces to the spectral representation method. This will be demonstrated numerically in Section 3.2.4.

2.10 Summary

This chapter presents the theoretical background for the thesis. K-L expansion as a representation and simulation method for Gaussian, non-Gaussian and non-stationary processes is presented. Simulation of random processes using the K-L expansion basically requires: (a) solving the Fredholm integral equation to obtain the eigenvalues and eigenfunctions of the covariance function and (b) selecting uncorrelated random variables such that the simulated process produces the desired distribution. A key feature of K-L expansion in simulation is the spectral decomposition of the covariance function describing the process. This involves solving homogeneous Fredholm integral equation of the second kind either analytically or numerically. The efficiency of K-L expansion for simulating random processes hinges on the availability of accurate solutions for the eigenvalues and eigenfunctions of the covariance function. For most covariance functions, analytical solution is not available so numerical methods are needed to solve the integral equation. For the convenience of comparison and reader's understanding, a brief introduction of the most widely used spectral representation method is given. The spectral method has been used for stationary Gaussian, non-stationary Gaussian and stationary non-Gaussian processes. Section 2.9 shows that the K-L expansion converges to the spectral representation method in the case of stationary process with a long observation record.

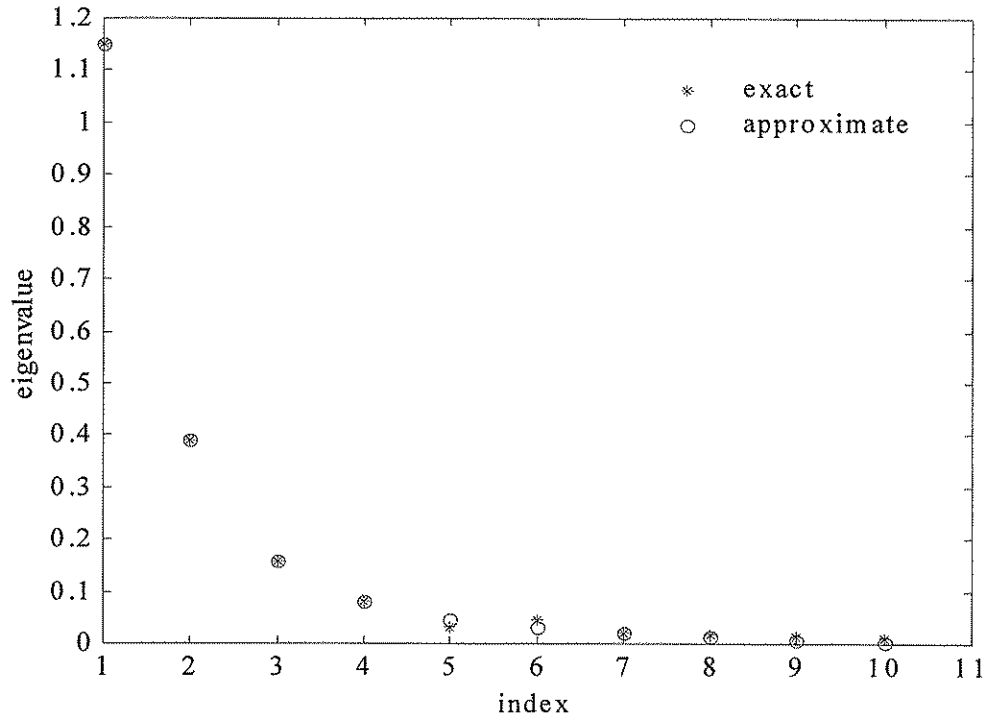

Figure 2.1 Exact (analytical) and approximate (numerical) eigenvalues for exponential covariance model (first-order Markov process)

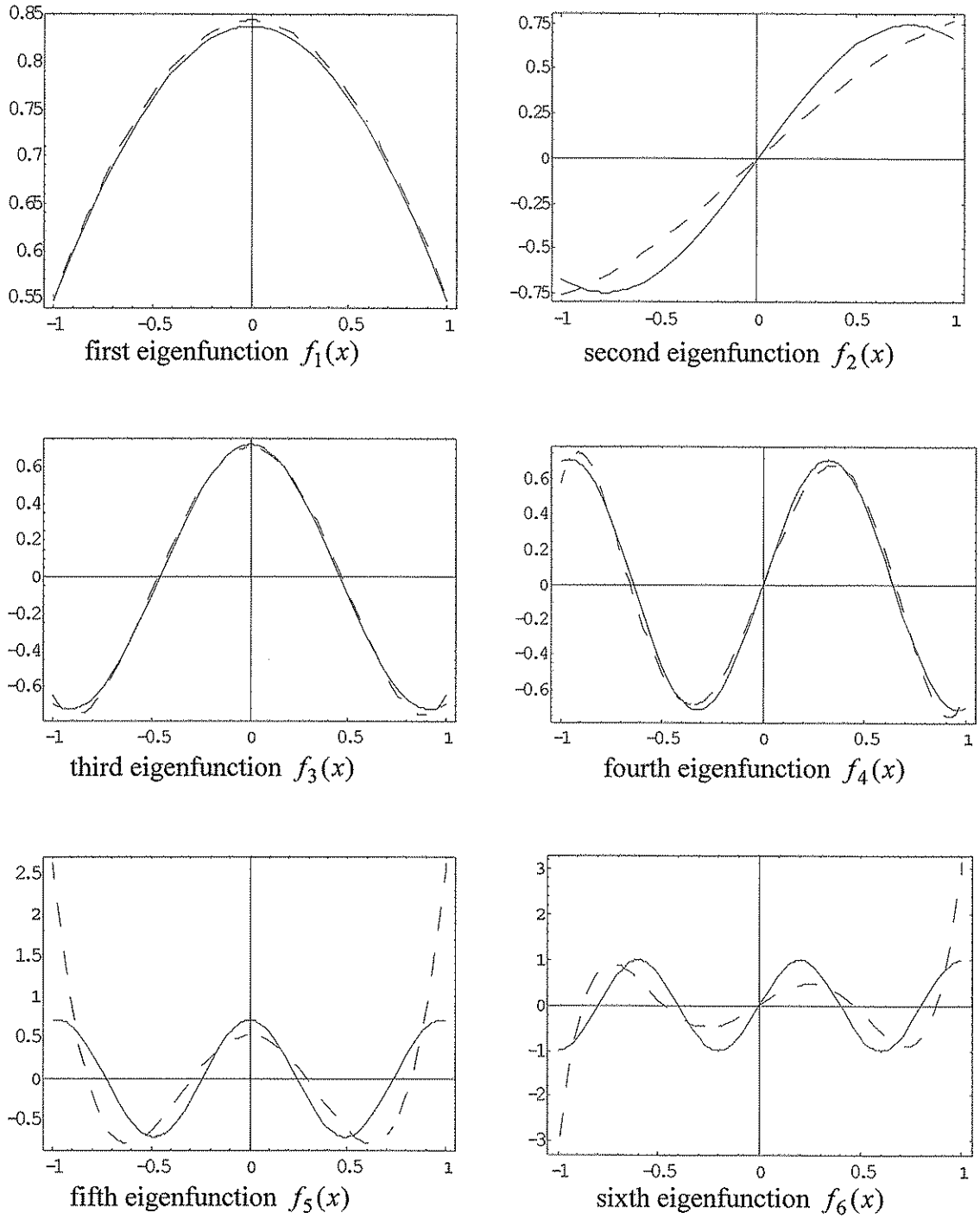

Figure 2.2 Exact (analytical; solid lines) and approximate (numerical; dash lines) eigenfunctions for the exponential model (first-order Markov process)

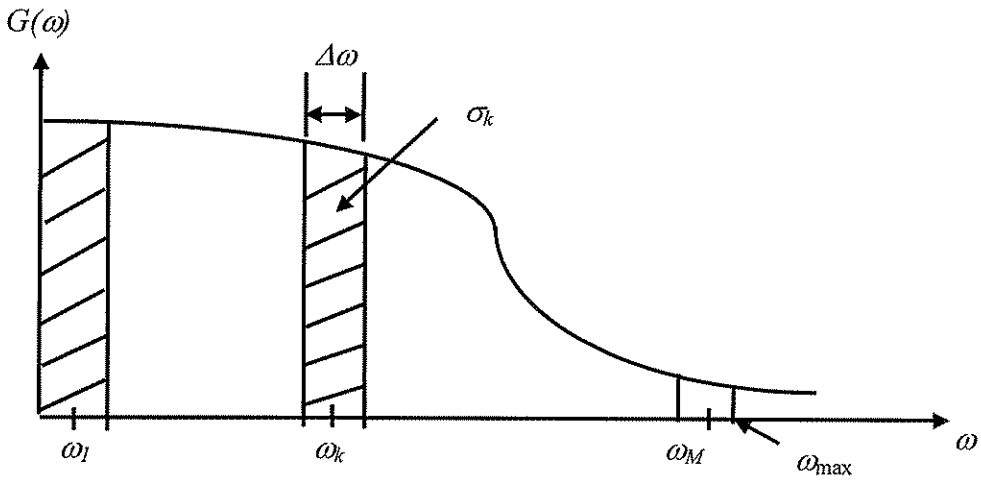

Figure 2.3 Discretization of power spectral density in spectral representation method