Random Perturbation Of A Self-Adjoint Operator With A Multiple Eigenvalue

by

George Gaines, BS, MS

A Dissertation

In

Mathematics and Statistics

Submitted to the Graduate Faculty
of Texas Tech University in
Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy

Approved

Frits Ruymgaart, PhD
Chairman of the Committee

David Gilliam, PhD

Alex Trindade, PhD

Peggy Gordon Miller
Dean of the Graduate School

May, 2012

ACKNOWLEDGMENTS

I thank first Dr. Frits Ruymgaart, my adviser, for his generosity with his vast knowledge of mathematics and statistics. All of the main ideas in this dissertation came from Dr. Ruymgaart. I learned by figuring out the small steps connecting the main ideas. Previously, I had three courses with Dr. Ruymgaart and learned then many of the basic notions in this dissertation.

I also thank Dr. David Gilliam, who introduced me to functional analysis, served on the dissertation committee and, more importantly, showed interest in my efforts early on when I could easily have quit. I thank Dr. Alex Trindade for his excellently taught time series course and for serving on the committee. I thank the many professors in the Texas Tech Mathematics and Statistics Department for their teaching and patience, and the several staff members for their friendly help on numerous occasions. I thank Dr. James Valles, my classmate and officemate until he graduated two semesters before me, for countless conversations and suggestions.

And I thank most profoundly the memory of my parents, Matthew and Mollie Gaines, who never had an opportunity to go to college but who accumulated great wisdom from a lifetime of working class living in Brooklyn, New York. Would that I had learned to appreciate their wisdom years ago.

TABLE OF CONTENTS

ABSTRACT

In this dissertation, we first consider a bounded self-adjoint operator on a Hilbert space with a multiple eigenvalue as its largest eigenvalue. Here a multiple eigenvalue is defined as having an eigenspace with dimension two or more. We perturb the operator so that the perturbed operator has a cluster of eigenvalues in the neighborhood of the original multiple eigenvalue. First, nonrandom perturbation is studied and then random perturbation is introduced. We consider the convergence of the cluster of scattered eigenvalues to the original eigenvalue as the perturbation is made smaller. The results are applied to the special case of the covariance operator in statistics.

We then construct a concrete example, a Gaussian random vector with a multiple eigenvalue, and perform a computer simulation. Random samples are simulated and the resulting sample covariance matrices and their eigenvalues are computed. The sample covariance matrices are considered as a perturbation of the covariance matrix of the original random vector. A cluster of eigenvalues of the sample covariance matrix is observed to correspond to the original multiple eigenvalue. An average of the cluster of sample covariance matrix eigenvalues is computed according to results obtained in earlier chapters, and the convergence of the cluster average towards the original multiple eigenvalue is observed to be consistent with theory developed in earlier chapters.

Finally, we show a method of approximating Brownian motion with a step function chosen so that the resulting covariance matrix has eigenvalues and eigenprojections within predetermined tolerances of the true values.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

A simple eigenvalue of a linear operator has a one-dimensional eigenspace. A multiple eigenvalue is defined as having an eigenspace with dimension greater than one. In most practical cases, all the eigenvalues are simple, and that is often assumed in studies (see, for instance, [1], [2], [7], [12]).

In practice, the largest eigenvalue is often the most important eigenvalue. The approach here starts with a Hilbert space and a bounded self-adjoint operator whose largest eigenvalue is a multiple eigenvalue. Then the operator is perturbed so that, with a high probability, the perturbed operator has a cluster of eigenvalues of lower multiplicity (most likely simple eigenvalues) in a neighborhood of the original multiple eigenvalue. Related results for matrices are shown in [16]. The more general context of of Hilbert space is discussed in [3], [5] and [9].

Chapter 2 gives basic definitions and notation. Chapter 3 introduces nonrandom perturbation, gives some theorems to be used later, and discusses the notion of the scattering of eigenvalues to form the cluster mentioned above. Chapter 4 discusses an additional assumption introduced to make the perturbation random in the context of Hilbert space, and applies probabilities to the problem of the multiple eigenvalue. At this point, we have not had enough information about the original operator to develop an estimator. This is addressed in Chapter 5, where we consider an important special case in statistics, the covariance operator. There a random sample is assumed, and certain other assumptions are made, so that an approach for developing estimators can be discussed.

Since multiple eigenvalues are uncommon in practice, it is natural to consider how a multiple eigenvalue may occur, or how an operator with a multiple eigenvalue may be constructed. That is addressed in Chapter 6. An operator Σ with only simple eigenvalues is assumed to be known, and then an operator A is constructed so that the largest eigenvalue of the operator $A\Sigma A$ has a prescribed value with a prescribed multiplicity.

In Chapter 7, a concrete example is considered, starting with a Gaussian random vector in the space $L^2[0,1]$ such that its covariance matrix has prescribed eigenvalues, including a largest eigenvalue with a prescribed multiplicity. A computer simulation then generates a random sample of N such vectors, constructs

the sample covariance matrix and then computes the largest eigenvalues of the sample covariance matrix. This is done for several values of the sample size N, and for each N the simulation calculates an average of the sample covariance matrix eigenvalues following the method developed in earlier chapters. The difference between the original prescribed multiple eigenvalue and the average of the corresponding sample covariance matrix eigenvalues decreases as the sample size N increases, as expected from the results in earlier chapters. The computer code is reproduced in the Appendix.

Finally, Brownian motion is considered in Chapter 8 for the time interval $[0, 1]$. This is often approximated with the Wiener process. While the Wiener process has only simple eigenvalues, Chapter 6 has shown that there may be a direct relationship between a process that has only simple eigenvalues and a process that has a multiple eigenvalue. A step function approximation is considered with M subintervals of equal length. An analysis is done to show how large M should be chosen so that the approximation meets specified accuracy requirements.

CHAPTER 2
BASIC NOTATION AND TOOLS

2.1 Definitions and Notation

Throughout $\mathbb{H}$ will denote a separable Hilbert space over the real numbers with inner product $< \cdot, \cdot >$ and norm $|| \cdot ||$. The Banach space of all bounded operators on $\mathbb{H}$ will be denoted by $\mathcal{L}$, with norm $|| \cdot ||_{\mathcal{L}}$. A *self-adjoint* operator will be a bounded operator that is equal to its own adjoint. The subspace of $\mathcal{L}$ consisting of self-adjoint operators will be denoted by $\mathcal{L}_{SA}$. The subspace of $\mathcal{L}$ consisting of Hilbert-Schmidt operators will be denoted by $\mathcal{L}_{HS}$, where $U \in \mathcal{L}_{HS}$ if and only if $\sum_{k=1}^{\infty} ||Ue_k||^2 < \infty$, for an orthonormal basis $e_1, e_2, \ldots$ of $\mathbb{H}$. This sum is independent of the choice of orthonormal basis, for suppose $f_1, f_2, \ldots$ is another orthonormal basis. Then

$$\begin{aligned}
\sum_{k=1}^{\infty} ||Ue_k||^2 &= \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |\, \langle Ue_k, f_j \rangle \,|^2 \\
&= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |\, \langle Ue_k, f_j \rangle \,|^2 \\
&= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |\, \langle e_k, U^* f_j \rangle \,|^2 \\
&= \sum_{j=1}^{\infty} ||U^* f_j||^2.
\end{aligned}$$

The terms in the sums are all positive, so the change in order of summation is justified. Repeating this calculation, using only the $f_1, f_2, \ldots$ basis, gives

$$\sum_{k=1}^{\infty} ||Uf_k||^2 = \sum_{j=1}^{\infty} ||U^* f_j||^2.$$

Hence,

$$\sum_{k=1}^{\infty} ||Ue_k||^2 = \sum_{k=1}^{\infty} ||Uf_k||^2.$$

This is discussed in [14].

The space $\mathcal{L}_{HS}$ is itself a separable Hilbert space with inner product

$$\langle U, V \rangle_{HS} = \sum_{k=1}^{\infty} \langle Ue_k, Ve_k \rangle . \qquad U, V \in \mathcal{L}_{HS} \tag{2.1}$$

Also $< U, V >_{HS}$ as defined in (2.1) is independent of the choice of basis (see [11]). The corresponding norm will be denoted by $|| \cdot ||_{HS}$.

For any $a, b \in \mathbb{H}$ the operator $a \otimes b$ is defined by

$$(a \otimes b)x =< x, a > b. \qquad x \in \mathbb{H} \tag{2.2}$$

We have

$$\begin{aligned} ||a \otimes b||_{HS}^2 &= \langle a \otimes b, a \otimes b \rangle_{HS} \\ &= \sum_{k=1}^{\infty} \langle (a \otimes b)e_k, (a \otimes b)e_k \rangle \\ &= \sum_{k=1}^{\infty} \langle \langle e_k, a \rangle b, \langle e_k, a \rangle b \rangle \\ &= \sum_{k=1}^{\infty} | \langle e_k, a \rangle |^2 \langle b, b \rangle \\ &= \left(\sum_{k=1}^{\infty} | \langle e_k, a \rangle |^2 \right) ||b||^2 \\ &= ||a||^2 ||b||^2. \end{aligned}$$

Then $a \otimes b \in \mathcal{L}_{HS}$ and

$$||a \otimes b||_{HS} = ||a|| \cdot ||b||. \tag{2.3}$$

The set $\{e_j \otimes e_k, j, k \in \mathbb{N}\}$ is an orthonormal basis for $\mathcal{L}_{HS}$. We have

$||e_j \otimes e_k||_{HS} = 1$ from (2.3). For orthogonality,

$$\begin{aligned}
\langle e_j \otimes e_k, e_n \otimes e_m \rangle_{HS} &= \sum_{i=1}^{\infty} \langle (e_j \otimes e_k)e_i, (e_n \otimes e_m)e_i \rangle \\
&= \sum_{i=1}^{\infty} \langle \langle e_i, e_j \rangle e_k, \langle e_i, e_n \rangle e_m \rangle \\
&= \sum_{i=1}^{\infty} \langle e_i, e_j \rangle \langle e_i, e_n \rangle \langle e_k, e_m \rangle \\
&= \langle e_j, e_n \rangle \langle e_k, e_m \rangle \\
&= \begin{cases} 1 \text{ if } (j,k) = (n,m) \\ 0 \text{ otherwise} \end{cases}
\end{aligned}$$

For spanning, let $U \in \mathcal{L}_{HS}$ such that U is orthogonal to $e_j \otimes e_k$ for all $j, k \in \mathbb{N}$. Then

$$\begin{aligned}
0 &= \langle U, e_j \otimes e_k \rangle_{HS} \\
&= \sum_{n=1}^{\infty} \langle U e_n, (e_j \otimes e_k)e_n \rangle \\
&= \sum_{n=1}^{\infty} \langle U e_n, \langle e_n, e_k \rangle e_j \rangle \\
&= \sum_{n=1}^{\infty} \langle e_n, e_k \rangle \langle U e_n, e_j \rangle \\
&= \langle U e_k, e_j \rangle .
\end{aligned}$$

So Ue_k is orthogonal to all e_j; hence $Ue_k = 0$. This is true for all k, so $U = 0$. Thus we have that $\{e_j \otimes e_k, j, k \in \mathbb{N}\}$ spans and is an orthonormal basis for $\mathcal{L}_{HS}$.

Let P_k be the orthogonal projection onto the span of e_k. Then, in the notation of (2.2),

$$\begin{aligned}
P_k x &=< x, e_k > \, e_k = (e_k \otimes e_k)\, x, \qquad\qquad \text{all } x \in \mathbb{H}, \quad k = 1, 2, \ldots, \\
P_k &= e_k \otimes e_k. \qquad\qquad (2.4)
\end{aligned}$$

We let E_k be the orthogonal projection onto span$\{e_1, \ldots, e_k\}$. Then $E_k \in \mathcal{L}_{HS}$ and

$$E_k x = \sum_{j=1}^{k} \langle x, e_j \rangle \, e_j = \sum_{j=1}^{k} (e_j \otimes e_j) \, x = \sum_{j=1}^{k} P_j \, x,$$
$$E_k = \sum_{j=1}^{k} e_j \otimes e_j = \sum_{j=1}^{k} P_j. \tag{2.5}$$

We also have

$$||E_k||_{HS}^2 = k \tag{2.6}$$

from

$$\begin{aligned}
||E_k||_{HS}^2 &= \langle E_k, E_k \rangle_{HS} \\
&= \sum_{j=1}^{\infty} \langle E_k e_j, E_k e_j \rangle \\
&= \sum_{j=1}^{\infty} \left\langle \sum_{n=1}^{k} (e_n \otimes e_n) e_j, \sum_{m=1}^{k} (e_m \otimes e_m) e_j \right\rangle \\
&= \sum_{j=1}^{\infty} \sum_{n=1}^{k} \sum_{m=1}^{k} \langle (e_n \otimes e_n) e_j, (e_m \otimes e_m) e_j \rangle \\
&= \sum_{j=1}^{\infty} \sum_{n=1}^{k} \sum_{m=1}^{k} \langle \langle e_j, e_n \rangle \, e_n, \langle e_j, e_m \rangle \, e_m \rangle \\
&= \sum_{j=1}^{\infty} \sum_{n=1}^{k} \sum_{m=1}^{k} \langle e_j, e_n \rangle \, \langle e_j, e_m \rangle \, \langle e_n, e_m \rangle \\
&= \sum_{j=1}^{k} \sum_{n=1}^{k} \langle e_j, e_n \rangle^2 \\
&= k.
\end{aligned}$$

So if P is the orthogonal projection operator to a subspace of dimension k, then

$$||P||_{HS}^2 = k. \tag{2.7}$$

Let $T \in \mathcal{L}_{SA}$, not invertible, with isolated eigenvalue $\lambda_0 \neq 0$ of finite multiplicity

ν_0. Now let the projection operator to the eigenspace of λ_0 be E_0. Then

$$||E_0||_{HS}^2 = \nu_0. \tag{2.8}$$

Also define

$$T_0 = \lambda_0 E_0, \tag{2.9}$$
$$T_1 = T - T_0. \tag{2.10}$$

2.2 Resolvent Operators

T is not invertible, so $0 \in \sigma(T)$, the spectrum of T. Since T is self-adjoint, $\sigma(T) \subset \mathbb{R}$. Also

$$\sigma(T_0) = \{0, \lambda_0\}, \tag{2.11}$$
$$\sigma(T_1) = \sigma(T) \setminus \{\lambda_0\}. \tag{2.12}$$

For any operator S, let $\rho(S)$ be the resolvent set of S, i.e., $\rho(S) = \mathbb{C} \setminus \sigma(S)$. We define the resolvent operators

$$\begin{aligned} R(z) &= (zI - T)^{-1} && z \in \rho(T), \\ R_0(z) &= (zI - T_0)^{-1} && z \in \rho(T_0), \\ R_1(z) &= (zI - T_1)^{-1} && z \in \rho(T_1). \end{aligned} \tag{2.13}$$

From the spectral theorem, we have

$$\begin{aligned}
I &= \int_{\sigma(T)} dE(\lambda), \\
T &= \int_{\sigma(T)} \lambda \, dE(\lambda), \\
zI - T &= \int_{\sigma(T)} (z - \lambda) \, dE(\lambda), \\
R(z) &= \int_{\sigma(T)} \frac{1}{z - \lambda} \, dE(\lambda) \\
&= \frac{1}{z - \lambda_0} E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z - \lambda} \, dE(\lambda), \qquad (2.14) \\
R_0(z) &= \frac{1}{z - \lambda_0} E_0 + \frac{1}{z}(I - E_0), \qquad (2.15) \\
R_1(z) &= \frac{1}{z} E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z - \lambda} \, dE(\lambda). \qquad (2.16)
\end{aligned}$$

Because of orthogonality,

$$\begin{aligned}
&T_1 E_0 = 0, \qquad (2.17) \\
&T_0 E_k = 0. \qquad\qquad E_k \neq E_0
\end{aligned}$$

Using (2.15) and (2.16), this gives

$$\begin{aligned}
T_1 R_0(z) &= \frac{1}{z} T_1, \qquad (2.18) \\
T_0 R_1(z) &= \frac{1}{z} T_0. \qquad (2.19)
\end{aligned}$$

Using (2.14) - (2.16),

$$\begin{aligned}
R(z) &= \frac{1}{z - \lambda_0} E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z - \lambda} \, dE(\lambda) \\
&= \left[R_0(z) - \frac{1}{z}(I - E_0) \right] + \left[R_1(z) - \frac{1}{z} E_0 \right] \\
&= R_0(z) + R_1(z) - \frac{1}{z} I. \qquad (2.20)
\end{aligned}$$

We can verify (2.20) by showing that the identities $(zI - T)R(z) = I$ and $R(z)(zI - T) = I$ are true for $R(z)$ as given by (2.20). In that case,

$$\begin{aligned}
&(zI - T)\left[R_0(z) + R_1(z) - \frac{1}{z}I\right] \\
&\qquad = [(zI - T_0) - T_1][R_0(z) + R_1(z)] - (zI - T)\left(\frac{1}{z}\right)I \\
&\qquad = (zI - T_0)R_0(z) + (zI - T_0)R_1(z) - T_1R_0(z) - T_1R_1(z) - I + \frac{1}{z}T \\
&\qquad = I + zR_1(z) - T_0R_1(z) - T_1R_0(z) - T_1R_1(z) - I + \frac{1}{z}T \\
&\qquad = (zI - T_1)R_1(z) - T_0R_1(z) - T_1R_0(z) + \frac{1}{z}T \\
&\qquad = I - T_0R_1(z) - T_1R_0(z) + \frac{1}{z}T \\
&\qquad = I - \frac{1}{z}T_0 - \frac{1}{z}T_1 + \frac{1}{z}T \qquad \text{use (2.18) and (2.19)} \\
&\qquad = I - \frac{1}{z}[T_0 + T_1 - T] \\
&\qquad = I. \qquad \text{use (2.10)}
\end{aligned}$$

Similarly, $\left[R_0(z) + R_1(z) - \frac{1}{z}I\right](zI - T) = I$. This verifies (2.20) for $z \in \rho(T) \cap \rho(T_0) \cap \rho(T_1)$.

2.3 A Special Case for Large $|z|$

For large $|z|$, (2.18) and (2.19) can be derived by a different method. From (2.9) and (2.10),

$$\begin{aligned}
&||T_0||_{\mathcal{L}} = |\lambda_0|, \\
&||T_1||_{\mathcal{L}} = ||T - T_0||_{\mathcal{L}} \leq ||T||_{\mathcal{L}} + |\lambda_0|.
\end{aligned}$$

Suppose $|z| > ||T||_{\mathcal{L}} + |\lambda_0|$. In this special case, we can develop a geometric series approach.

$$\begin{aligned}
R_0(z) &= (zI - T_0)^{-1} \\
&= \left[z\left(I - \frac{1}{z} T_0 \right) \right]^{-1} \\
&= \frac{1}{z}\left(I - \frac{\lambda_0}{z} E_0 \right)^{-1} \\
&= \frac{1}{z}\left[I + \frac{\lambda_0}{z} E_0 + \left(\frac{\lambda_0}{z} E_0 \right)^2 + \ldots \right] \\
&= \frac{1}{z} I + \frac{1}{z}\left[\frac{\lambda_0}{z} + \left(\frac{\lambda_0}{z} \right)^2 + \ldots \right] E_0. \\
T_1 R_0(z) &= \frac{1}{z} T_1. \qquad \text{use (2.17)}
\end{aligned}$$

This is consistent with (2.18). Again using a geometric series approach in this special case,

$$\begin{aligned}
T_0 T_1 &= T_0(T - T_0) = \lambda_0^2 E_0 - \lambda_0^2 E_0 = 0. \\
R_1(z) &= (zI - T_1)^{-1} \\
&= \left[z\left(I - \frac{1}{z} T_1 \right) \right]^{-1} \\
&= \frac{1}{z}\left[I + \frac{1}{z} T_1 + \left(\frac{1}{z} T_1 \right)^2 + \left(\frac{1}{z} T_1 \right)^3 + \ldots \right]. \\
T_0 R_1(z) &= \frac{1}{z} T_0 + \left[\frac{1}{z^2} T_0 T_1 + \frac{1}{z^3} T_0 T_1^2 + \frac{1}{z^4} T_0 T_1^3 + \ldots \right] \\
&= \frac{1}{z} T_0.
\end{aligned}$$

This is consistent with (2.19).

2.4 Contour Integrals Using the Resolvent Operators

Let Γ_0 be a contour in the complex plane with λ_0 inside Γ_0, but no other point of $\sigma(T)$ inside Γ_0, and let Γ_1 be a contour with every point of $\sigma(T) \setminus \{\lambda_0\}$ inside Γ_1, such that Γ_0 and Γ_1 do not intersect and there is no point inside both of them. Note

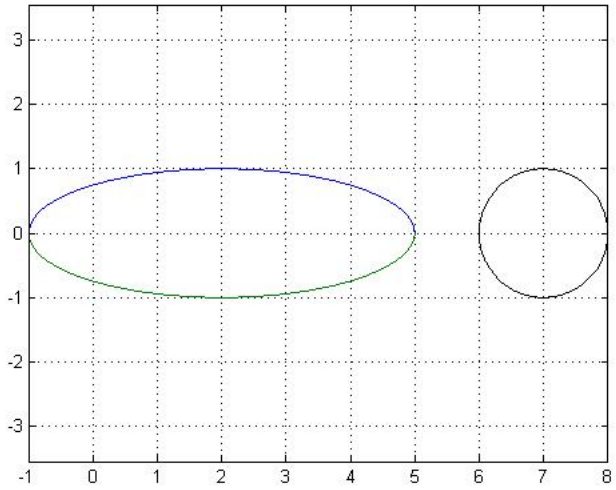

Figure 2.1. Contours for integrals.

$0 \in \sigma(T)$ and $\lambda_0 \neq 0$ (page 6), so 0 is inside Γ_1. In figure 2.1, Γ_1 is the ellipse on the left and Γ_0 is the circle on the right. Define

1. D_j is a simply connected region that contains $\Gamma_j,\ j = 1, 2$, such that D_0 and D_1 are disjoint.
2. ϕ is a complex-valued function analytic on both D_0 and D_1.
3. $\Gamma = \Gamma_0 \cup \Gamma_1$.
4. Ω_j is the simply connected open region with boundary $\Gamma_j,\ j = 1, 2$.

For the operator T, the function $\phi(T)$ is defined by

$$\phi(T) = \frac{1}{2\pi i} \oint_{\Gamma} \phi(z) R(z)\, dz, \tag{2.21}$$

where $R(z)$ is defined in (2.13).

Theorem 2.1.

$$\oint_{\Gamma_0} \phi(z) R(z)\, dz = \oint_{\Gamma_0} \phi(z) R_0(z)\, dz. \tag{2.22}$$

Proof. Using (2.20),

$$\begin{aligned}\oint_{\Gamma_0} \phi(z)R(z)\,dz &= \oint_{\Gamma_0} \phi(z)\left[R_0(z) + R_1(z) - \frac{1}{z}I\right]\,dz \\ &= \oint_{\Gamma_0} \phi(z)R_0(z)\,dz + \oint_{\Gamma_0} \phi(z)R_1(z)\,dz - \oint_{\Gamma_0} \frac{\phi(z)}{z}\,dz\,I.\end{aligned}$$

Since the second and third integrands are analytic on D_0, those integrals are both zero. □

Theorem 2.2.

$$\oint_{\Gamma_1} \phi(z)R(z)\,dz = \oint_{\Gamma_1} \phi(z)R_1(z)\,dz - 2\pi\,i\,\phi(0)\,E_0. \tag{2.23}$$

Proof. Using (2.20),

$$\begin{aligned}\oint_{\Gamma_1} \phi(z)R(z)\,dz &= \oint_{\Gamma_1} \phi(z)\left[R_0(z) + R_1(z) - \frac{1}{z}I\right]\,dz \\ &= \oint_{\Gamma_1} \phi(z)R_0(z)\,dz + \oint_{\Gamma_1} \phi(z)R_1(z)\,dz - \oint_{\Gamma_1} \frac{\phi(z)}{z}I\,dz. \qquad (2.24)\end{aligned}$$

For the first integral in (2.24), using (2.15),

$$\begin{aligned}\oint_{\Gamma_1} \phi(z)R_0(z)\,dz &= \oint_{\Gamma_1} \phi(z)\left[\frac{1}{z-\lambda_0}E_0 + \frac{1}{z}(I - E_0)\right]\,dz \\ &= \oint_{\Gamma_1} \frac{\phi(z)}{z-\lambda_0}\,dz\,E_0 + \oint_{\Gamma_1} \frac{\phi(z)}{z}\,dz\,(I - E_0) \qquad (2.25) \\ &= 0 + 2\pi i\,\phi(0)(I - E_0).\end{aligned}$$

The first integral in (2.25) is zero because z is on Γ_1 and λ_0 is outside Γ_1. For the second integral in (2.25), note that 0 is contained in Γ_1, so there is a simple pole at $z = 0$. For the third integral in (2.24),

$$\oint_{\Gamma_1} \frac{\phi(z)}{z}I\,dz = 2\pi i\,\phi(0)\,I.$$

Applying these values in (2.24) gives (2.23). □

2.5 Preliminary Result

We define

$$\chi_0(z) = \begin{cases} 1 & z \in D_0 \\ 0 & z \in D_1 \end{cases} \tag{2.26}$$

$$\psi_0(z) = z\,\chi_0(z) \qquad z \in D_0 \cup D_1. \tag{2.27}$$

We have

$$\chi_0(T) = E_0 \tag{2.28}$$

from

$$\begin{aligned}
\chi_0(T) &= \frac{1}{2\pi i}\oint_\Gamma \chi_0(z) R(z)\,dz && \text{use (2.21)} \\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} R(z)\,dz \\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} \left[\frac{1}{z-\lambda_0}E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\right] dz && \text{use (2.14)} \\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} \frac{dz}{z-\lambda_0}E_0 + \frac{1}{2\pi i}\int_{\sigma(T)\setminus\{\lambda_0\}} \left[\oint_{\Gamma_0} \frac{dz}{z-\lambda}\right] dE(\lambda) \\
&= E_0.
\end{aligned}$$

The last contour integral above is zero because z is on Γ_0 and $\lambda \in \sigma(T) \setminus \{\lambda_0\}$, which is entirely outside Γ_0. We also have

$$\psi_0(T) = T_0, \tag{2.29}$$

from

$$\begin{aligned}
\psi_0(T) &= \frac{1}{2\pi i}\oint_{\Gamma}\psi_0(z)R(z)\,dz && \text{use (2.21)}\\
&= \frac{1}{2\pi i}\oint_{\Gamma} z\chi_0(z)R(z)\,dz\\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} zR(z)\,dz\\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} z\left[\frac{1}{z-\lambda_0}E_0+\int_{\sigma(T)\setminus\{\lambda_0\}}\frac{1}{z-\lambda}\,dE(\lambda)\right]dz && \text{use (2.14)}\\
&= \frac{1}{2\pi i}\oint_{\Gamma_0}\frac{z\,dz}{z-\lambda_0}E_0+\frac{1}{2\pi i}\int_{\sigma(T)\setminus\{\lambda_0\}}\left[\oint_{\Gamma_0}\frac{z\,dz}{z-\lambda}\right]dE(\lambda)\\
&= \lambda_0 E_0\\
&= T_0.
\end{aligned}$$

The last contour integral above is zero because z is on Γ_0 and $\lambda \in \sigma(T)\setminus\{\lambda_0\}$, which is entirely outside Γ_0.

CHAPTER 3
NONRANDOM PERTURBATION

3.1 Theorem on Perturbation of Operator

We will use the following theorem.

Theorem 3.1. *There exists a number $0 < K < \infty$, independent of all the relevant parameters, such that if $\Pi \in \mathcal{L}_{SA}$ with $||\Pi||_{\mathcal{L}} < K$, then*

$$\sigma(T + \Pi) \subset \Omega_0 \cup \Omega_1, \tag{3.1}$$

$$\phi(T + \Pi) = \phi(T) + \dot{\phi}_T\Pi + \textit{remainder}, \tag{3.2}$$

where

$$\dot{\phi}_T\Pi = \frac{1}{2\pi i}\oint_\Gamma \phi(z)R(z)\Pi R(z)\,dz, \tag{3.3}$$

$$||\textit{remainder}||_{\mathcal{L}} = \mathcal{O}(||\Pi||_{\mathcal{L}}^2). \tag{3.4}$$

The regions Ω_0 and Ω_1 are defined on page 11. A proof of theorem 3.1 is given in [8]. Results related to this theorem can be found in [6] in the case where T and Π commute.

3.2 Background for the Theorems

Let

$$Q = \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{\lambda_0 - \lambda}\,dE(\lambda). \tag{3.5}$$

Then, because of orthogonality, range(Q) is orthogonal to range(E_0), so

$$QE_0 = E_0Q = 0. \tag{3.6}$$

For $\Pi \in \mathcal{L}_{SA}$ with $||\Pi||_{\mathcal{L}} < K$ as in Theorem 3.1, define

$$\tilde{T} = T + \Pi, \tag{3.7}$$

$$\tilde{E}_0 = \chi_0(\tilde{T}), \tag{3.8}$$

$$\tilde{T}_0 = \psi_0(\tilde{T}). \tag{3.9}$$

3.3 Approximation Theorem for a Projection Operator

Theorem 3.2. *For $\Pi \in \mathcal{L}_{SA}$ with $||\Pi||_{\mathcal{L}} < K$ as in Theorem 3.1, and Q given by* (3.5),

$$\tilde{E}_0 = E_0 + E_0\,\Pi\,Q + Q\,\Pi\,E_0 + \textit{remainder}, \tag{3.10}$$

where $||\textit{remainder}||_{\mathcal{L}} = \mathcal{O}(||\Pi||^2_{\mathcal{L}})$, and $\tilde{E}_0$ is a projection into a subspace of dimension $\tilde{\nu}_0 = \nu_0$ (page 7), which is invariant under $\tilde{T}$.

Proof. Using Theorem 3.1,

$$\begin{aligned}
\tilde{E}_0 &= \chi_0(\tilde{T}) \\
&= \chi_0(T + \Pi) \\
&= \chi_0(T) + \dot{\chi}_{0,T}\Pi + \text{ remainder} \\
&= E_0 + \dot{\chi}_{0,T}\Pi + \text{ remainder.} \qquad\qquad use(2.28)
\end{aligned}$$

The Fréchet derivative above is calculated as in Theorem 3.1:

$$\begin{aligned}
\dot{\chi}_{0,T}\Pi &= \frac{1}{2\pi i}\oint_{\Gamma} \chi_0(z)R(z)\Pi R(z)\,dz \\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} R(z)\Pi R(z)\,dz.
\end{aligned} \tag{3.11}$$

The last integral above is calculated using (2.14) (a similar calculation is done in [8]):

$$\begin{aligned}
&\oint_{\Gamma_0} R(z)\Pi R(z)\,dz = \\
&\oint_{\Gamma_0} \left[\frac{1}{z-\lambda_0}E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\right]\Pi\left[\frac{1}{z-\lambda_0}E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\right]\,dz.
\end{aligned} \tag{3.12}$$

Then calculate the four integrals (3.13) - (3.16). First,

$$\oint_{\Gamma_0} \frac{1}{z-\lambda_0}E_0\Pi\frac{1}{z-\lambda_0}E_0\,dz = \oint_{\Gamma_0} \frac{dz}{(z-\lambda_0)^2}E_0\Pi E_0 = 0. \tag{3.13}$$

The second integrand in (3.13) has a double pole at $z = \lambda_0$ with residue 0. Second,

$$\begin{aligned}
&\oint_{\Gamma_0} \frac{1}{z-\lambda_0} E_0 \Pi \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\, dE(\lambda)\, dz \qquad (3.14)\\
&\quad = E_0 \Pi \int_{\sigma(T)\setminus\{\lambda_0\}} \left[\oint_{\Gamma_0} \frac{dz}{(z-\lambda_0)(z-\lambda)} \right] dE(\lambda)\\
&\quad = E_0 \Pi \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{2\pi i}{\lambda_0 - \lambda}\, dE(\lambda)\\
&\quad = 2\pi i E_0 \Pi Q.
\end{aligned}$$

The last contour integral above has a simple pole at $z = \lambda_0$. Third,

$$\begin{aligned}
&\oint_{\Gamma_0} \left[\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\, dE(\lambda) \right] \Pi \frac{1}{z-\lambda_0} E_0\, dz \qquad (3.15)\\
&\quad = \int_{\sigma(T)\setminus\{\lambda_0\}} \left[\oint_{\Gamma_0} \frac{dz}{(z-\lambda)(z-\lambda_0)} \right] dE(\lambda) \Pi E_0\\
&\quad = \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{2\pi i}{\lambda_0 - \lambda}\, dE(\lambda) \Pi E_0\\
&\quad = 2\pi i Q \Pi E_0.
\end{aligned}$$

The last contour integral above has a simple pole at $z = \lambda_0$. Fourth and last,

$$\begin{aligned}
&\oint_{\Gamma_0} \left[\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\, dE(\lambda) \right] \Pi \left[\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\mu}\, dE(\mu) \right] dz \qquad (3.16)\\
&\quad = \int_{\sigma(T)\setminus\{\lambda_0\}} \int_{\sigma(T)\setminus\{\lambda_0\}} \left[\oint_{\Gamma_0} \frac{dz}{(z-\lambda)(z-\mu)} \right] dE(\lambda) \Pi\, dE(\mu)\\
&\quad = 0.
\end{aligned}$$

The last contour integral above is zero because z is on Γ_0 and
$\lambda, \mu \in \sigma(T) \setminus \{\lambda_0\}$, which is entirely outside Γ_0. Using (3.13) - (3.16) in (3.12),

$$\oint_{\Gamma_0} R(z) \Pi R(z)\, dz = 2\pi i (E_0 \Pi Q + Q \Pi E_0).$$

Then (3.11) becomes

$$\dot{\chi}_{0,T} \Pi = E_0 \Pi Q + Q \Pi E_0,$$

and (3.10) is proved.

We have, from (2.26), $\chi_0^2(z) = \chi_0(z)$ for $z \in D_0 \cup D_1$, so $\tilde{E}_0^2 = \chi_0^2(\tilde{T}) = \chi_0(\tilde{T}) = \tilde{E}_0$. As $\tilde{E}_0$ is also self-adjoint, it follows that $\tilde{E}_0$ is a projection operator. Further, we have from (3.10) that $||\tilde{E}_0 - E_0||_{\mathcal{L}} \to 0$ as $||\Pi||_{\mathcal{L}} \to 0$. Then the dimension of range$(\tilde{E}_0)$ equals the dimension of range(E_0), i.e., $\tilde{\nu}_0 = \nu_0$ for sufficiently small $||\Pi||_{\mathcal{L}}$ (see [13]).

Finally, suppose $x \in$ range$(\tilde{E}_0)$, so $\tilde{E}_0 x = x$. Then $\tilde{T}x = \tilde{T}\tilde{E}_0 x = \tilde{T}\chi_0(\tilde{T})x = \chi_0(\tilde{T})\tilde{T}x = \tilde{E}_0\tilde{T}x$. Then $\tilde{T}x \in$ range$(\tilde{E}_0)$. Hence, range$(\tilde{E}_0)$ is invariant under $\tilde{T}$.

This completes the proof of theorem 3.2. □

3.4 Approximation Theorem for a Scalar Multiple of a Projection Operator

We now consider $\tilde{T}_0$ (page 15).

Theorem 3.3. *For $\Pi \in \mathcal{L}_{SA}$ with $||\Pi||_{\mathcal{L}} < K$ as in Theorem 3.1, and Q given by (3.5),*

$$\tilde{T}_0 = T_0 + E_0\,\Pi\,E_0 + \lambda_0\,E_0\,\Pi\,Q + \lambda_0\,Q\,\Pi\,E_0 + \textit{remainder}, \tag{3.17}$$

where $||\textit{remainder}||_{\mathcal{L}} = \mathcal{O}(||\Pi||_{\mathcal{L}}^2)$.

Proof. Using Theorem 3.1 and equation (2.29):

$$\begin{aligned}
\tilde{T}_0 &= \psi_0(\tilde{T}) \\
&= \psi_0(T + \Pi) \\
&= \psi_0(T) + \dot{\psi}_{0,T}\Pi + \text{ remainder} \\
&= T_0 + \dot{\psi}_{0,T}\Pi + \text{ remainder}.
\end{aligned} \tag{3.18}$$

The Fréchet derivative above is calculated as in Theorem 3.1.

$$\begin{aligned}
\dot{\psi}_{0,T}\Pi &= \frac{1}{2\pi i}\oint_{\Gamma} \psi_0(z)R(z)\Pi R(z)\,dz \\
&= \frac{1}{2\pi i}\oint_{\Gamma} z\chi_0(z)R(z)\Pi R(z)\,dz \\
&= \frac{1}{2\pi i}\oint_{\Gamma_0} zR(z)\Pi R(z)\,dz.
\end{aligned} \tag{3.19}$$

The last integral above is calculated using (2.14) (a similar calculation is done

in [8]).

$$\oint_{\Gamma_0} zR(z)\Pi R(z)\,dz =$$
$$\oint_{\Gamma_0} z\left[\frac{1}{z-\lambda_0}E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\right]\Pi\left[\frac{1}{z-\lambda_0}E_0 + \int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\right]\,dz. \tag{3.20}$$

Then calculate the four integrals (3.21) - (3.24).

$$\oint_{\Gamma_0} \frac{z}{z-\lambda_0}E_0\Pi\frac{1}{z-\lambda_0}E_0\,dz = \oint_{\Gamma_0} \frac{z\,dz}{(z-\lambda_0)^2}E_0\Pi E_0 = 2\pi i E_0\Pi E_0. \tag{3.21}$$

The second integrand in (3.21) has a double pole at $z=\lambda_0$ with residue 1. Second,

$$\begin{aligned}
&\oint_{\Gamma_0} \frac{z}{z-\lambda_0}E_0\Pi\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\,dz \\
&\quad= E_0\Pi\int_{\sigma(T)\setminus\{\lambda_0\}}\left[\oint_{\Gamma_0}\frac{z\,dz}{(z-\lambda_0)(z-\lambda)}\right]dE(\lambda) \\
&\quad= E_0\Pi\int_{\sigma(T)\setminus\{\lambda_0\}}\frac{2\pi i\lambda_0}{\lambda_0-\lambda}\,dE(\lambda) \\
&\quad= 2\pi i\lambda_0 E_0\Pi Q.
\end{aligned} \tag{3.22}$$

The last contour integral above has a simple pole at $z=\lambda_0$. Third,

$$\begin{aligned}
&\oint_{\Gamma_0} z\left[\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\,dE(\lambda)\right]\Pi\frac{1}{z-\lambda_0}E_0\,dz \\
&\quad= \int_{\sigma(T)\setminus\{\lambda_0\}}\left[\oint_{\Gamma_0}\frac{z\,dz}{(z-\lambda)(z-\lambda_0)}\right]dE(\lambda)\Pi E_0 \\
&\quad= \int_{\sigma(T)\setminus\{\lambda_0\}}\frac{2\pi i\lambda_0}{\lambda_0-\lambda}\,dE(\lambda)\Pi E_0 \\
&\quad= 2\pi i\lambda_0 Q\Pi E_0.
\end{aligned} \tag{3.23}$$

The last contour integral above has a simple pole at $z = \lambda_0$. Fourth and last,

$$\begin{aligned}\oint_{\Gamma_0} z \left[\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\lambda}\, dE(\lambda)\right] \Pi \left[\int_{\sigma(T)\setminus\{\lambda_0\}} \frac{1}{z-\mu}\, dE(\mu)\right] dz \\ = \int_{\sigma(T)\setminus\{\lambda_0\}} \int_{\sigma(T)\setminus\{\lambda_0\}} \left[\oint_{\Gamma_0} \frac{z\, dz}{(z-\lambda)(z-\mu)}\right] dE(\lambda)\Pi\, dE(\mu) \\ = 0. \end{aligned} \tag{3.24}$$

The last contour integral above is zero because z is on Γ_0 and $\lambda, \mu \in \sigma(T) \setminus \{\lambda_0\}$, which is entirely outside Γ_0.

Using (3.21) - (3.24) in (3.20),

$$\oint_{\Gamma_0} zR(z)\Pi R(z)\, dz = 2\pi i(E_0\Pi E_0 + \lambda_0 E_0 \Pi Q + \lambda_0 Q \Pi E_0).$$

Then (3.19) becomes

$$\dot{\psi}_{0,T}\Pi = E_0\Pi E_0 + \lambda_0 E_0 \Pi Q + \lambda_0 Q \Pi E_0.$$

This completes the proof of theorem 3.3. □

3.5 Scattering of Eigenvalues Due to Perturbation

Let $||\Pi||_{\mathcal{L}}$ be sufficiently small so that, by Theorem 3.2, range$(\tilde{E}_0)$ is invariant under $\tilde{T}$, and let the domain of $\tilde{T}$ be restricted to range$(\tilde{E}_0)$. Now consider, using (3.9), $\tilde{T}_0 = \psi_0(\tilde{T})$ with $\tilde{T}$ so restricted. Then

$$\tilde{T}_0(\text{range}(\tilde{E}_0)) \subset \text{range}(\tilde{T}_0) = \text{range}(\psi_0(\tilde{T})) \subset \text{range}(\tilde{T}) \subset \text{range}(\tilde{E}_0).$$

The last set containment is valid because, as noted at the start of this section, under the current assumptions, range$(\tilde{E}_0)$ is invariant under $\tilde{T}$. We now have that range$(\tilde{E}_0)$ is invariant under $\tilde{T}_0$. By Theorem 3.2, range$(\tilde{E}_0)$ has dimension $\tilde{\nu}_0 = \nu_0$, the dimension of range(T_0) (page 7).

$\tilde{T}_0$ has $\tilde{k}_0 \leq \nu_0$ distinct eigenvalues $\tilde{\lambda}_0(1), \ldots, \tilde{\lambda}_0(\tilde{k}_0)$ with finite multiplicities

$\tilde{\nu}_0(1), \ldots, \tilde{\nu}_0(\tilde{k}_0)$, and eigenprojections $\tilde{E}_0(1), \ldots, \tilde{E}_0(\tilde{k}_0)$, such that

$$\tilde{\nu}_0 = \sum_{j=1}^{\tilde{k}_0} \tilde{\nu}_0(j), \tag{3.25}$$

$$\tilde{E}_0 = \sum_{j=1}^{\tilde{k}_0} \tilde{E}_0(j), \tag{3.26}$$

$$\tilde{T}_0 = \sum_{j=1}^{\tilde{k}_0} \tilde{\lambda}_0(j)\, \tilde{E}_0(j). \tag{3.27}$$

For $i \neq j$, range$(\tilde{E}_0(i))$ is orthogonal to range$(\tilde{E}_0(j))$, so

$$\left\langle \tilde{E}_0(i),\, \tilde{E}_0(j) \right\rangle_{HS} = \sum_{k=1}^{\infty} \left\langle \tilde{E}_0(i)\, e_k,\, \tilde{E}_0(j)\, e_k \right\rangle = 0.$$

Also, using (2.7),

$$\left\langle \tilde{E}_0(j),\, \tilde{E}_0(j) \right\rangle_{HS} = \left|\left| \tilde{E}_0(j) \right|\right|_{HS}^2 = \tilde{\nu}_0(j).$$

Then, using (3.27),

$$\begin{aligned}
\left|\left| \tilde{T}_0 \right|\right|_{HS}^2 &= \left\langle \tilde{T}_0,\, \tilde{T}_0 \right\rangle_{HS} \\
&= \left\langle \sum_{i=1}^{\tilde{k}_0} \tilde{\lambda}_0(i)\, \tilde{E}_0(i),\, \sum_{j=1}^{\tilde{k}_0} \tilde{\lambda}_0(j)\, \tilde{E}_0(j) \right\rangle_{HS} \\
&= \sum_{i=1}^{\tilde{k}_0} \sum_{j=1}^{\tilde{k}_0} \tilde{\lambda}_0(i) \tilde{\lambda}_0(j) \left\langle \tilde{E}_0(i),\, \tilde{E}_0(j) \right\rangle_{HS} \\
&= \sum_{j=1}^{\tilde{k}_0} \tilde{\lambda}_0^2(j) \tilde{\nu}_0(j).
\end{aligned} \tag{3.28}$$

CHAPTER 4
RANDOM PERTURBATION

4.1 Additional Assumption

We can define a probability measure on a Banach space by the following procedure. Here the Banach space is $\mathcal{L}_{SA}$ (page 3).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, i.e., Ω is a sample space, $\mathcal{F}$ is the σ-field consisting of measurable subsets of Ω, and $\mathbb{P}$ is a probability measure defined on Ω. The open subsets of $\mathcal{L}_{SA}$ are defined using the norm $||\cdot||_{\mathcal{L}}$. The σ-field $\mathcal{B}_{SA}$ of Borel sets is then generated from the open subsets of $\mathcal{L}_{SA}$. A random variable X is then defined by a measurable function $X : \Omega \longrightarrow \mathcal{L}_{SA}$. Finally, the probability of $B \in \mathcal{B}_{SA}$ is defined by $P(B) = \mathbb{P}(X^{-1}(B))$.

We will assume that there exist a sequence of random operators $\hat{T}_n \in \mathcal{L}_{SA}$ and a random operator $\mathbb{G} \in \mathcal{L}_{SA}$ such that

$$\sqrt{n}\,(\hat{T}_n - T) \xrightarrow{d} \mathbb{G} \qquad \text{as } n \to \infty, \tag{4.1}$$

where T is defined on page 6. Then

$$||\hat{T}_n - T||_{\mathcal{L}} \xrightarrow{d} \mathcal{O}p\left(\frac{1}{\sqrt{n}}\right). \tag{4.2}$$

4.2 Applying Probabilities

We have sequences of random operators $\hat{T}_n$ and $\hat{\Pi}_n = \hat{T}_n - T$ such that

$$||\hat{\Pi}_n||_{\mathcal{L}} = ||\hat{T}_n - T||_{\mathcal{L}} \xrightarrow{d} \mathcal{O}p\left(\frac{1}{\sqrt{n}}\right). \tag{4.3}$$

In chapter 3 we used $\tilde{}$ (tilde) to indicate a perturbation. We now use $\hat{}$ (hat) to indicate a random perturbation, and consider sequences of operators. From (3.8) - (3.9),

$$\hat{E}_{0,n} = \chi_0(\hat{T}_n), \tag{4.4}$$

$$\hat{T}_{0,n} = \psi_0(\hat{T}_n). \tag{4.5}$$

In the current notation, (3.25)-(3.27) become

$$\hat{\nu}_{0,n} = \sum_{j=1}^{\hat{k}_{0,n}} \hat{\nu}_{0,n}(j), \tag{4.6}$$

$$\hat{E}_{0,n} = \sum_{j=1}^{\hat{k}_{0,n}} \hat{E}_{0,n}(j), \tag{4.7}$$

$$\hat{T}_{0,n} = \sum_{j=1}^{\hat{k}_{0,n}} \hat{\lambda}_{0,n}(j)\, \hat{E}_{0,n}(j). \tag{4.8}$$

Theorem 4.1. *Under the current assumptions we have, for analytic function ϕ,*

$$\sqrt{n}\left[\phi(\hat{T}_n) - \phi(T)\right] \xrightarrow{d} \dot{\phi}_T \mathbb{G} \qquad \text{as } n \to \infty, \tag{4.9}$$

where $\dot{\phi}_T \mathbb{G}$ is calculated as in theorem 3.1. A proof of Theorem 4.1 is given in [4]. From Theorem 3.2 with $||\Pi_n||_{\mathcal{L}}$ sufficiently small and Q given by (3.5),

$$\hat{E}_{0,n} - E_0 = E_0\hat{\Pi}_n Q + Q\hat{\Pi}_n E_0 + \mathcal{O}p(||\hat{\Pi}_n||^2),$$

where Q is given by (3.5). Then, using (4.1),

$$\sqrt{n}\left(\hat{E}_{0,n} - E_0\right) \xrightarrow{d} E_0\mathbb{G}Q + Q\mathbb{G}E_0 \qquad \text{in } \mathcal{L}_{HS} \qquad \text{as } n \to \infty. \tag{4.10}$$

Also from theorem 3.2,

$$P(\hat{\nu}_{0,n} = \nu_0) \longrightarrow 1 \text{ as } n \to \infty. \tag{4.11}$$

From Theorem 3.3 with $||\Pi_n||_{\mathcal{L}}$ sufficiently small and Q given by (3.5),

$$\hat{T}_{0,n} - T_0 = E_0\hat{\Pi}_n E_0 + \lambda_0 E_0\hat{\Pi}_n Q + \lambda_0 Q\hat{\Pi}_n E_0 + \mathcal{O}p(||\hat{\Pi}_n||^2).$$

Then, using (4.1),

$$\sqrt{n}\left(\hat{T}_{0,n} - T_0\right) \xrightarrow{d} E_0\mathbb{G}E_0 + \lambda_0 E_0\mathbb{G}Q + \lambda_0 Q\mathbb{G}E_0 \qquad \text{as } n \to \infty. \tag{4.12}$$

From (2.8) and (2.9), we have

$$||T_0||^2_{HS} = ||\lambda_0 E_0||^2_{HS} = \lambda_0^2 \, \nu_0. \tag{4.13}$$

Also from (2.8) and (2.9), T_0 has finite dimension, so $T_0 \in \mathcal{L}_{HS}$. If $\phi(S) = \langle S, \, S \rangle_{HS} = ||S||^2_{HS}$ for $S \in \mathcal{L}_{HS}$ then, using (4.12)-(4.13) and a delta method described in [15],

$$\begin{aligned}
\sqrt{n} \, \left[||\hat{T}_{0,n}||^2_{HS} - \lambda_0^2 \nu_0 \right] &= \sqrt{n} \left[\phi \left(\hat{T}_{0,n} \right) - \phi \left(T_0 \right) \right] \\
&\xrightarrow{d} 2 \left\langle E_0 \mathbb{G} E_0 + \lambda_0 E_0 \mathbb{G} Q + \lambda_0 Q \mathbb{G} E_0, \, T_0 \right\rangle_{HS} \\
&= 2\lambda_0 \left\langle E_0 \mathbb{G} E_0, \, E_0 \right\rangle_{HS} + 2\lambda_0^2 \left\langle E_0 \mathbb{G} Q, E_0 \right\rangle_{HS} + 2\lambda_0^2 \left\langle Q \mathbb{G} E_0, E_0 \right\rangle_{HS} \\
&= 2\lambda_0 \left\langle \mathbb{G} E_0, \, E_0 \right\rangle_{HS} + 2\lambda_0^2 \left\langle \mathbb{G} Q, \, E_0 \right\rangle_{HS} + 2\lambda_0^2 \left\langle \mathbb{G} E_0, \, Q E_0 \right\rangle_{HS}.
\end{aligned} \tag{4.14}$$

Here we used (2.9), that E_0 and Q are self-adjoint and that $E_0^2 = E_0$. We also used that E_0 is finite rank, and therefore $E_0 \mathbb{G} E_0$, $E_0 \mathbb{G} Q$ and $Q \mathbb{G} E_0$ are all Hilbert-Schmidt. We can simplify (4.14) further:

$$\begin{aligned}
\left\langle \mathbb{G} Q, E_0 \right\rangle_{HS} &= \sum_{k=1}^{\infty} \left\langle \mathbb{G} Q \, e_k, \, E_0 \, e_k \right\rangle \\
&= \sum \{ \left\langle \mathbb{G} Q \, e_k, \, e_k \right\rangle : e_k \in \text{range}(E_0) \} \\
&= 0. \qquad \text{See (3.5) and the statement following it.}
\end{aligned}$$

Also, (3.6) gives that $\left\langle \mathbb{G} E_0, \, Q E_0 \right\rangle_{HS} = 0$. Therefore, (4.14) simplifies to

$$\sqrt{n} \, \left(||\hat{T}_{0,n}||^2_{HS} - \lambda_0^2 \nu_0 \right) \xrightarrow{d} 2\lambda_0 \left\langle \mathbb{G} E_0, \, E_0 \right\rangle_{HS} \qquad \text{as } n \to \infty. \tag{4.15}$$

From (3.28),

$$\left|\left| \hat{T}_{0,n} \right|\right|^2_{HS} = \sum_{j=1}^{\hat{k}_{0,n}} \hat{\lambda}^2_{0,n}(j) \, \hat{\nu}_{0,n}(j). \tag{4.16}$$

Using (4.15)-(4.16),

$$\sqrt{n} \left(\sum_{j=1}^{\hat{k}_{0,n}} \hat{\lambda}^2_{0,n}(j) \, \hat{\nu}_{0,n}(j) - \lambda_0^2 \, \nu_0 \right) \xrightarrow{d} 2\lambda_0 \left\langle \mathbb{G} E_0, \, E_0 \right\rangle_{HS} \qquad \text{as } n \to \infty.$$

Recalling (4.6) and (4.11),

$$\sqrt{n}\left(\sum_{j=1}^{\hat{k}_{0,n}} \hat{\lambda}_{0,n}^2(j)\,\frac{\hat{\nu}_{0,n}(j)}{\hat{\nu}_{0,n}} - \lambda_0^2\,\frac{\nu_0}{\hat{\nu}_{0,n}}\right) \xrightarrow{d} 2\,\frac{\lambda_0}{\nu_0}\,\langle \mathbb{G}E_0,\, E_0\rangle_{HS}\,. \qquad \text{as } n\to\infty.$$

$$\frac{1}{\hat{\nu}_{0,n}}\sum_{j=1}^{\hat{k}_{0,n}} \hat{\lambda}_{0,n}^2(j)\,\hat{\nu}_{0,n}(j) - \lambda_0^2 \xrightarrow{d} \mathcal{O}p\left(\frac{1}{\sqrt{n}}\right). \tag{4.17}$$

If $\phi(z)=\sqrt{z}$ for $z>0$, then $\phi'(z)=\frac{1}{2\sqrt{z}}$. Again using a delta method, along with (4.15) and (4.16),

$$\begin{aligned}\sqrt{n}\left(\left|\left|\hat{T}_{0,n}\right|\right|_{HS} - |\lambda_0|\,\sqrt{\nu_0}\right) &= \sqrt{n}\,\left[\phi\left(||\hat{T}_{0,n}||_{HS}^2\right) - \phi\left(\lambda_0^2\,\nu_0\right)\right] \\ &\xrightarrow{d} \frac{1}{2\,|\lambda_0|\,\sqrt{\nu_0}}\cdot 2\lambda_0\,\langle \mathbb{G}E_0,\, E_0\rangle_{HS} \qquad \text{as } n\to\infty \\ &= \frac{\lambda_0}{|\lambda_0|}\cdot\frac{1}{\sqrt{\nu_0}}\,\langle \mathbb{G}E_0,\, E_0\rangle_{HS}\,. \end{aligned} \tag{4.18}$$

Using (4.11) and (4.16) in (4.18),

$$\sqrt{n}\left(\sqrt{\frac{1}{\hat{\nu}_{0,n}}\sum_{j=1}^{\hat{k}_{0,n}} \hat{\lambda}_{0,n}^2(j)\,\hat{\nu}_{0,n}(j)} \;-\; |\lambda_0|\,\sqrt{\frac{\nu_0}{\hat{\nu}_{0,n}}}\right) \xrightarrow{d} \frac{\lambda_0}{|\lambda_0|}\cdot\frac{1}{\nu_0}\,\langle \mathbb{G}E_0,\, E_0\rangle_{HS}$$

$$\text{as } n\to\infty. \tag{4.19}$$

For calculation, we can make the inner product explicit. Let

$$\mathbb{K}_0 = \{k\in\mathbb{N} : e_k \in \text{range}(E_0)\}. \tag{4.20}$$

Then

$$\langle \mathbb{G}E_0,\, E_0\rangle_{HS} = \sum_{k=1}^{\infty}\langle \mathbb{G}E_0\,e_k,\, E_0\,e_k\rangle = \sum_{k\in\mathbb{K}_0}\langle \mathbb{G}\,e_k,\, e_k\rangle\,.$$

$$\sqrt{n}\left(\sqrt{\frac{1}{\hat{\nu}_{0,n}}\sum_{j=1}^{\hat{k}_{0,n}}\hat{\lambda}_{0,n}^2(j)\,\hat{\nu}_{0,n}(j)}\;-\;|\lambda_0|\sqrt{\frac{\nu_0}{\hat{\nu}_{0,n}}}\right)\xrightarrow{d}\frac{\lambda_0}{|\lambda_0|}\cdot\frac{1}{\nu_0}\sum_{k\in\mathbb{K}_0}\langle\mathbb{G}\,e_k,\,e_k\rangle$$

$$\text{as } n\to\infty. \tag{4.21}$$

CHAPTER 5
APPLICATION TO SAMPLE COVARIANCE OPERATOR

What we have done so far assumes that we have sufficient information about the operator T (page 6) and its eigenvalue λ_0, etc., and the contours Γ_0 and Γ_1 (page 10) to do the various calculations we have done. In practice, we do not have such information. Therefore, quantities such as those described in (4.4) - (4.5) are not proper estimators for the corresponding quantities of T. We address that problem for the special case of T being a covariance operator, which we define in what follows.

We use the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ given on page 22. Suppose $X \in \mathbb{H}$ is a random variable. $\mathbb{E}[X]$ is defined as the vector $\mu \in \mathbb{H}$ such that

$$\mathbb{E}\,\langle a,\, X\rangle = \langle a,\, \mu\rangle \qquad \text{for all } a \in \mathbb{H}.$$

That μ exists and is unique is proved with the Riesz representation theorem (see, e.g., [10]). We assume that $\mathbb{E}||X||^4 < \infty$.

The covariance operator of X is an operator Σ defined by

$$\mathbb{E}\,\langle a,\, X-\mu\rangle\,\langle X-\mu,\, b\rangle = \langle a,\, \Sigma\, b\rangle \qquad \text{for all } a, b \in \mathbb{H}.$$

That Σ exists and is unique is proved with the Riesz representation theorem (see, e.g., [10]), where $\Sigma \in \mathcal{L}_{HS}$, which is defined on page 3, and the inner product is given in (2.1). This entails

$$\Sigma = \mathbb{E}\,(X-\mu)\otimes(X-\mu), \tag{5.1}$$

where $\otimes$ is defined in (2.2). A covariance operator such as Σ is self-adjoint, finite trace and nonnegative. We assume, without loss of generality in practice, that Σ is positive, so that it is one-to-one and its eigenvalues are all positive. That Σ is finite trace means that

$$\sum_{k=1}^{\infty} \langle \Sigma\, e_k,\, e_k\rangle < \infty.$$

This, together with Σ being positive, assures that Σ is Hilbert-Schmidt and therefore compact (see [14]).

The eigenvalues of Σ are variances and so will be denoted by σ_k^2 in decreasing order

$$\sigma_0^2 > \sigma_1^2 > \ldots \downarrow 0.$$

Except as otherwise specified (e.g., $\Sigma = T$, $\sigma_k^2 = \lambda_k$), we continue with the notation and ideas previously established. Now (3.5) becomes

$$Q = \sum_{k=1}^{\infty} \frac{1}{\sigma_0^2 - \sigma_k^2} E(k),$$

where $E(k)$ is the orthogonal projection into the eigenspace of σ_k^2. Let $X_1, \ldots, X_n$ be independent random variables in $\mathbb{H}$ with the same distribution as X, and let the sample mean and sample covariance operator be

$$\bar{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i,$$
$$\hat{\Sigma}_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}) \otimes (X_i - \bar{X}).$$

It is shown in [5] that the central limit theorem for a separable Hilbert space gives

$$\sqrt{n}(\hat{\Sigma}_n - \Sigma) \xrightarrow{d} \mathbb{G} \qquad \text{in } \mathcal{L}_{HS} \quad \text{as } n \to \infty,$$

where $\mathbb{G}$ is Gaussian. Then (4.1) - (4.2) apply with $T = \Sigma$ and $\hat{T}_n = \hat{\Sigma}_n$. Now (4.4) - (4.5) become

$$\hat{E}_{0,n} = \chi_0(\hat{\Sigma}_n),$$
$$\hat{\Sigma}_{0,n} = \psi_0(\hat{\Sigma}_n).$$

Equations (4.6) - (4.7) remain unchanged but (4.8) becomes

$$\hat{\Sigma}_{0,n} = \sum_{j=1}^{\hat{k}_{0,n}} \hat{\sigma}_{0,n}^2(j)\, \hat{E}_{0,n}(j).$$

Now the distinct eigenvalues within the contour Γ_0 are $\hat{\sigma}_{0,n}^2(1) > \ldots > \hat{\sigma}_{0,n}^2(\hat{k}_{0,n})$ and their respective multiplicities are $\hat{\nu}_{0,n}(1) \ldots \hat{\nu}_{0,n}(\hat{k}_{0,n})$, which sum to $\hat{\nu}_{0,n}$ as in (4.6).

So (4.21) becomes

$$\sqrt{n}\left(\sqrt{\frac{1}{\hat{\nu}_{0,n}}\sum_{j=1}^{\hat{k}_{0,n}}\hat{\sigma}^4_{0,n}(j)\,\hat{\nu}_{0,n}(j)}\;-\;\sigma_0^2\sqrt{\frac{\nu_0}{\hat{\nu}_{0,n}}}\right)\xrightarrow{d}\frac{1}{\nu_0}\sum_{k\in\mathbb{K}_0}\langle\mathbb{G}\,e_k,\,e_k\rangle \qquad \text{as } n\to\infty, \tag{5.2}$$

where $\mathbb{K}_0$ is defined as in (4.20). Also (4.17) is now

$$\frac{1}{\hat{\nu}_{0,n}}\sum_{j=1}^{\hat{k}_{0,n}}\hat{\sigma}^4_{0,n}(j)\,\hat{\nu}_{0,n}(j)-\sigma_0^4\xrightarrow{d}\mathcal{O}p\left(\frac{1}{\sqrt{n}}\right).$$

We then have

$$\hat{\sigma}^4_{0,n}(j)-\sigma_0^4\xrightarrow{d}\mathcal{O}p\left(\frac{1}{\sqrt{n}}\right), \tag{5.3}$$

for $j=1,\ldots,\hat{k}_{0,n}$.

In practice, σ_0^2 and Γ_0 are not known. Then the various quantities with 0 subscript are also not known. So the square root of the average in (5.2) is not a proper estimator for σ_0^2; likewise $\hat{\nu}_{0,n}$ and $\hat{E}_{0,n}$ are not proper estimators for ν_0 and E_0, respectively. To develop a proper estimator, note that $\sigma(\hat{\Sigma})\subset\Omega_0\cup\Omega_1$ for sufficiently large n, where Ω_0 and Ω_1 are as defined on page 11. Now let

$$\begin{aligned}\hat{\sigma}^2(1)&=\text{ largest eigenvalue of }\hat{\Sigma}_n,\\ \hat{\Gamma}_n&=\text{ circle centered at }\hat{\sigma}^2(1)\text{ with radius }\frac{\rho}{n^\delta},\end{aligned} \tag{5.4}$$

for some $\rho\in(0,\infty)$ and $\delta\in(0,\frac{1}{2})$. We need to ensure that as n increases, the eigenvalues converge faster than the circle shrinks, so that the eigenvalues are always inside the circle; hence the requirement that $\delta\in(0,\frac{1}{2})$. Then for any $\epsilon>0$,

$$\begin{aligned}&P(|\hat{\sigma}^2(1)-\hat{\sigma}^2_{0,n}(j)|<\epsilon)\longrightarrow 1 \qquad && j=1,\ldots,\hat{k}_{0,n},\\ &P(|\hat{\sigma}^2(1)-\sigma_0^2|<\epsilon)\longrightarrow 1 && \text{as } n\to\infty.\end{aligned}$$

Hence, for sufficiently large n, the contour $\hat{\Gamma}_n$ is contained in the contour Γ_0 defined on page 10 (with λ_0 replaced by σ_0^2). Also for sufficiently large n, given the choice of δ in (5.4), the contour $\hat{\Gamma}_n$ contains $\hat{\sigma}^2_{0,n}(j)$, $j=1,\ldots,\hat{k}_{0,n}$. Recall Γ_0 contains σ_0^2 but

no other eigenvalue of Σ. Since $\hat{\Gamma}_n$ is contained in Γ_0, the former does not contain any eigenvalues of $\hat{\Sigma}_n$ other than $\hat{\sigma}^2_{0,n}(j)$, $j = 1, \ldots, \hat{k}_{0,n}$.

Now the distinct eigenvalues of $\hat{\Sigma}_n$ within the contour $\hat{\Gamma}_n$ are $\hat{\sigma}^2(1) > \ldots > \hat{\sigma}^2(\hat{k}_{0,n})$ and their respective multiplicities are $\hat{\nu}_{0,n}(1) \ldots \hat{\nu}_{0,n}(\hat{k}_{0,n})$, which sum to $\hat{\nu}_{0,n}$ as in (4.6). We let

$$\begin{aligned} \hat{k} &= \hat{k}_{0,n}, \\ \hat{\nu}(j) &= \hat{\nu}_{0,n}(j), \qquad\qquad j = 1, \ldots, \hat{k}, \\ \hat{\nu}_n &= \sum_{j=1}^{\hat{k}} \hat{\nu}(j). \end{aligned}$$

Then $\hat{\nu}_n$ is an estimator of ν_0, and

$$P(\hat{\nu}_n = \nu_0) \longrightarrow 1 \qquad\qquad \text{as } n \to \infty.$$

Then

$$\sqrt{\frac{1}{\hat{\nu}_n} \sum_{j=1}^{\hat{k}} \hat{\sigma}^4(j) \hat{\nu}(j)}$$

is an estimator for σ_0^2 and, using (5.2),

$$\sqrt{n} \left(\sqrt{\frac{1}{\hat{\nu}_n} \sum_{j=1}^{\hat{k}} \hat{\sigma}^4(j) \hat{\nu}(j)} \; - \; \sigma_0^2 \right) \xrightarrow{d} \frac{1}{\nu_0} \sum_{k \in \mathbb{K}_0} \langle \mathbb{G}\, e_k,\, e_k \rangle \,, \tag{5.5}$$

where $\mathbb{K}_0$ is defined as in (4.20).

We also need to estimate the eigenprojection E_0. Let $\hat{e}_1(j) \ldots \hat{e}_{\hat{\nu}(j)}(j)$ be an orthonormal basis of eigenvectors for the eigenspace of $\hat{\sigma}^2(j)$, where $j = 1, \ldots \hat{k}$. Now let

$$\hat{E}_n = \sum_{j=1}^{\hat{k}} \sum_{i=1}^{\hat{\nu}(j)} \hat{e}_i(j) \otimes \hat{e}_i(j).$$

We also have

$$\hat{E}_n \xrightarrow{d} \hat{E}_{0,n} \qquad \text{in } \mathcal{L}_{HS} \qquad \text{as } n \to \infty.$$

Then (4.10) gives

$$\sqrt{n}\left(\hat{E}_n - E_0\right) \xrightarrow{d} E_0\mathbb{G}Q + Q\mathbb{G}E_0 \qquad \text{in } \mathcal{L}_{HS} \qquad \text{as } n \to \infty.$$

CHAPTER 6
HOW AN OPERATOR WITH A MULTIPLE EIGENVALUE MAY OCCUR

A Hermitian matrix with random entries has zero probability of having a multiple eigenvalue, as we can see by considering the characteristic polynomial and the probability that two roots would be equal. Likewise, it is uncommon in practice for an infinite dimensional operator to have a multiple eigenvalue. In fact, it is a common assumption in the literature that all eigenvalues are simple, as noted in the Introduction. We consider in this chapter how a multiple eigenvalue may arise in practice as a result of a transformation applied to an operator that has only simple eigenvalues.

As before let $\mathbb{H}$ be a separable Hilbert space with orthonormal basis $e_1, e_2, \ldots$, and let $Y \in \mathbb{H}$ be a random vector with zero mean Gaussian distribution as given by

$$Y = \sum_{j=1}^{\infty} \sigma_j Z_j \, e_j \tag{6.1}$$

where

1. $\sigma_1 > \sigma_2 > \ldots \downarrow 0$
2. $\sum_{j=1}^{\infty} \sigma_j^2 < \infty$
3. $Z_1, Z_2, \ldots$ are iid $\mathrm{N}(0, 1)$.

Let Σ be the covariance operator (page 27) for Y. Using (5.1),

$$\begin{aligned}
\Sigma &= \mathbb{E}\,(Y \otimes Y) \\
&= \mathbb{E}\left(\sum_{j=1}^{\infty} \sigma_j Z_j\, e_j\right) \otimes \left(\sum_{k=1}^{\infty} \sigma_k Z_k\, e_k\right) \\
&= \mathbb{E} \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \sigma_j\sigma_k\, Z_j Z_k\, (e_j \otimes e_k) \\
&= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \sigma_j\sigma_k\, \mathbb{E}(Z_j Z_k)(e_j \otimes e_k) \\
&= \sum_{j=1}^{\infty} \sigma_j^2\, \mathbb{E}\left(Z_j^2\right)(e_j \otimes e_j) && Z_j, Z_k \text{ iid } \mathrm{N}(0,1) \\
&= \sum_{j=1}^{\infty} \sigma_j^2\,(e_j \otimes e_j) && Z_j^2 \overset{d}{=} \chi^2(1) \\
&= \sum_{j=1}^{\infty} \sigma_j^2\, P_j. && \text{use (2.4)} \qquad (6.2)
\end{aligned}$$

Then Σ has eigenvalues $\sigma_1^2 > \sigma_2^2 > \ldots \downarrow 0$ with corresponding eigenvectors $e_1, e_2 \ldots$, our orthonormal basis for $\mathbb{H}$.

Now let A be a positive compact operator on $\mathbb{H}$ with eigenvalues $\alpha_1, \alpha_2, \ldots$ and the same corresponding eigenvectors $e_1, e_2, \ldots$, so that

$$A = \sum_{j=1}^{\infty} \alpha_j P_j,$$

where

$$\sum_{j=1}^{\infty} \alpha_j^2 < \infty.$$

We now have

$$\begin{aligned} A\Sigma A &= \left(\sum_{j=1}^{\infty} \alpha_j P_j\right)\left(\sum_{k=1}^{\infty} \sigma_k^2 P_k\right)\sum_{\ell=1}^{\infty} \alpha_\ell P_\ell \\ &= \left(\sum_{j=1}^{\infty} \alpha_j P_j\right)\sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \sigma_k^2 \alpha_\ell P_k P_\ell \\ &= \left(\sum_{j=1}^{\infty} \alpha_j P_j\right)\sum_{k=1}^{\infty} \sigma_k^2 \alpha_k P_k \\ &= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \alpha_j \sigma_k^2 \alpha_k P_j P_k \\ &= \sum_{j=1}^{\infty} \alpha_j^2 \sigma_j^2 \, P_j. \end{aligned} \tag{6.3}$$

We also have

$$\begin{aligned} AY &= \left(\sum_{j=1}^{\infty} \alpha_j P_j\right)\left(\sum_{k=1}^{\infty} \sigma_k Z_k e_k\right) \\ &= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \alpha_j \sigma_k Z_k P_j e_k \\ &= \sum_{j=1}^{\infty} \alpha_j \sigma_j Z_j e_j. \end{aligned}$$

Now let S be the covariance operator for the random vector AY. Note

$$\begin{aligned} \mathbb{E}(AY) &= \mathbb{E}\left(\sum_{j=1}^{\infty} \alpha_j \sigma_j Z_j e_j\right) \\ &= \sum_{j=1}^{\infty} \alpha_j \sigma_j \mathbb{E}(Z_j) e_j \\ &= 0. \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad Z_j \overset{d}{=} N(0,1) \end{aligned}$$

Using (5.1),

$$\begin{aligned}
S &= \mathbb{E}\,(AY \otimes AY) \\
&= \mathbb{E}\left(\sum_{j=1}^{\infty} \alpha_j \sigma_j Z_j e_j\right) \otimes \left(\sum_{k=1}^{\infty} \alpha_k \sigma_k Z_k e_k\right) \\
&= \mathbb{E} \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \alpha_j \alpha_k \sigma_j \sigma_k Z_j Z_k (e_j \otimes e_k) \\
&= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \alpha_j \alpha_k \sigma_j \sigma_k\, \mathbb{E}(Z_j Z_k)\,(e_j \otimes e_k) \\
&= \sum_{j=1}^{\infty} \alpha_j^2 \sigma_j^2\, \mathbb{E}(Z_j^2)\,(e_j \otimes e_j) && Z_j, Z_k \text{ iid } \mathrm{N}(0,1) \\
&= \sum_{j=1}^{\infty} \alpha_j^2 \sigma_j^2\,(e_j \otimes e_j) && Z_j^2 \stackrel{d}{=} \chi^2(1) \\
&= \sum_{j=1}^{\infty} \alpha_j^2 \sigma_j^2\, P_j. && \text{use (2.4)}
\end{aligned} \tag{6.4}$$

Comparing (6.3) and (6.4),

$$S = A\Sigma A,$$

where Σ is the covariance operator for the random vector Y and S is the covariance operator for AY.

Now we can construct an operator with a multiple eigenvalue. Let $J > 1$ be an integer, let $\beta > 0$ and let

$$\alpha_j = \frac{\sqrt{\beta}}{\sigma_j}, \qquad\qquad j = 1, 2, \ldots, J.$$

So $\alpha_j^2 \sigma_j^2 = \beta, j = 1, 2, \ldots, J$. Then

$$S = \beta \sum_{j=1}^{J} P_j + \sum_{j=J+1}^{\infty} \alpha_j^2 \sigma_j^2 P_j.$$

Now β is an eigenvalue of S and this eigenvalue has multiplicity $J > 1$.

CHAPTER 7
COMPUTER SIMULATIONS

Here we show a computer simulation of random vectors, their covariance operator and its eigenvalues as described in (6.1) - (6.2). For computer simulation, we consider the Hilbert space $L^2[0,1]$ consisting of real valued square integrable functions on $[0,1]$ with the inner product $< x, y >= \int_0^1 x(t)y(t)\,dt$. We use the orthonormal basis functions

$$\check{\phi}_k(t) = \sqrt{2}\,\sin\left(\left(k - \frac{1}{2}\right)\pi t\right), \qquad k = 1, 2, \ldots$$

We partition $[0,1]$ into M intervals of equal length, where M is a positive integer parameter. We define the partition points

$$t_m = \frac{m}{M} \qquad m = 0, \ldots, M.$$

Define

$$1_m(t) = \begin{cases} 1 & t_{m-1} < t \le t_m \\ 0 & \text{otherwise} \end{cases} \qquad m = 1, \ldots, M, \quad 0 \le t \le 1$$

We approximate the functions $\check{\phi}_k$ with step functions:

$$\phi_k(t) = \sum_{m=1}^{M} 1_m(t)\,\check{\phi}_k(t_m) \qquad 0 \le t \le 1$$

We then represent ϕ_k in the computer as the vector

$$\left[\phi_k(t_1) \quad \ldots \quad \phi_k(t_M)\right].$$

This kind of representation of a function is done throughout the computer simulation.

Let ν be the desired multiplicity of the largest eigenvalue, let $\beta > 0$ be the desired

largest eigenvalue, let $p > 0$ and let

$$\sigma_k^2 = \begin{cases} \beta & 1 \le k \le \nu, \\ \frac{\beta}{(k-\nu+p)^2} & k > \nu. \end{cases} \tag{7.1}$$

If we could consider a function

$$Y = \sum_{k=1}^{\infty} \sigma_k \, Z_k \, \check{\sigma}_k$$

defined by an infinite series as in (6.1), the covariance operator $\Sigma = E\,(Y \otimes Y)$ would have eigenvalues σ_k^2 with corresponding eigenfunctions $\check{\phi}_k$, but we truncate the series for the computer. Let K be the number of terms in the series. We calculate sample covariance operators and their eigenvalues. Because we truncate the series and represent the functions as $\mathbb{R}^M$ vectors, the sample covariance operators are $\mathbb{R}^{M\times M}$ matrices and their eigenvalues are calculated using linear algebra routines.

We will simulate a random sample of N functions, i.e., the sample size will be N and the sample functions will be

$$X_n(t_m) = \sum_{k=1}^{K} \sigma_k \, Z_k \, \phi_k(t_m), \qquad\qquad m = 1, \ldots, M;\ n = 1, \ldots, N.$$

where $Z_1, \ldots, Z_K$ are iid $\mathrm{N}(0, 1)$, so each X_n is Gaussian. The sample mean is

$$\bar{X}(t_m) = \frac{1}{N} \sum_{n=1}^{N} X_n(t_m) \qquad\qquad m = 1, \ldots, M.$$

Define

$$Y_n(t_m) = X_n(t_m) - \bar{X}(t_m) \qquad\qquad m = 1, \ldots, M;\ n = 1, \ldots, N.$$

so that each Y_n is zero mean Gaussian. Define the sample covariance matrix

$$\widehat{\Sigma}_N = \frac{1}{N} \sum_{n=1}^{N} (Y_n \otimes Y_n).$$

Table 7.1. Values of parameters for simulation.

M	ν	β	p	K
1000	3	10	.05	800

Table 7.2. Predetermined eigenvalues σ_k^2.

k=1	k=2	k=3	k=4	k=5	k=6
10.0000	10.0000	10.0000	9.0703	2.3795	1.0750

In the computer, Y_n is the vector

$$\begin{bmatrix} Y_n(t_1) & \dots & Y_n(t_M) \end{bmatrix} \qquad\qquad n = 1, \dots, N.$$

In the linear algebra (and the computer),

$$\widehat{\Sigma}_N = \frac{1}{N}\frac{1}{M}\sum_{n=1}^{N} Y_n^T\, Y_n. \tag{7.2}$$

The factor $\dfrac{1}{M}$ is necessary because in $L^2[0,1]$, for a function f represented as step function with the same partition points t_m,

$$(Y_n \otimes Y_n)f(t) =< f, Y_n > Y_n(t) = \left(\int_0^1 f(t)Y_n(t)\,dt\right) Y_n(t).$$

and the factor $\frac{1}{M}$ is the Δt in the inner product for the $L^2[0,1]$ step function represented in the computer as an $\mathbb{R}^M$ vector.

We now use the linear algebra routine to calculate the eigenvalues of the sample covariance operator (7.2) and compare them with the predetermined eigenvalues given by (7.1). The simulation uses the values of the parameters shown in Table 7.1. The six largest predetermined eigenvalues σ_k^2 are shown in Table 7.2. The six largest eigenvalues $\hat{\sigma}_{k,N}^2$ of the sample covariance operator are then shown in Table 7.3 for several sample sizes N.

Table 7.3. Eigenvalues $\hat{\sigma}^2_{k,N}$ of $\widehat{\Sigma}_N$.

	k=1	k=2	k=3	k=4	k=5	k=6
N						
100	12.5177	12.4486	9.6112	6.6312	2.6081	.9134
200	11.4552	10.8425	10.1459	7.7930	3.0847	1.0169
400	10.8989	10.5813	8.7254	8.1262	2.0180	1.1835
800	11.0396	10.1195	9.5905	8.9821	2.4020	1.0325
1600	10.4867	10.3125	9.4335	9.0500	2.3316	1.0453
3200	10.5023	10.0906	9.7350	8.9725	2.4073	1.1123
6400	10.2726	10.0370	9.7703	9.1625	2.3099	1.1058
12800	10.0616	9.9644	9.7662	9.1069	2.3266	1.0685
25600	10.2714	10.0446	9.8430	9.2279	2.4304	1.0829
51200	10.0406	10.0037	9.8917	9.0756	2.4057	1.0694

Referring to (5.5) and simplifying notation, define A_N by

$$A_N = \sqrt{\frac{1}{\nu}\sum_{k=1}^{\nu} \hat{\sigma}^4_{k,N}}.$$

The values A_N and β, and the differences between them, are shown in Table 7.4. The computer code is show in the Appendix.

As for the eigenfunctions of the sample covariance operator $\hat{\Sigma}_N$, nothing we have done implies that they converge to the corresponding eigenfunctions ϕ_k of the original covariance operator. While $\{\check{\phi}_k\}_{k=1}^{\infty}$ is an orthonormal basis for $L^2[0,1]$, in the computer simulation we only use the step function approximations $\{\phi_k\}_{k=1}^{K}$. The eigenvectors ϕ_1, ϕ_2, ϕ_3 are shown as functions in Figure 7.1. The corresponding eigenvectors of the eigenvalues $\hat{\sigma}_{1,N}, \hat{\sigma}_{2,N}, \hat{\sigma}_{3,N}$ of $\hat{\Sigma}_N$ are shown in Figure 7.2 for $N = 51200$. While the eigenfunctions of the sample covariance operator do not converge to the corresponding eigenfunctions of the original covariance operator, they do preserve much of the original sinusoidal character.

Table 7.4. Average eigenvalue $\hat{\sigma}^2_{k,N}, k = 1, \ldots, \nu$ vs. original eigenvalue β.

N	A_N	β	$\lvert A_N - \beta \rvert$
100	11.6051	10.0000	1.6051
200	10.8242	10.0000	0.8242
400	10.1141	10.0000	0.1141
800	10.2673	10.0000	0.2673
1600	10.0881	10.0000	0.0881
3200	10.1142	10.0000	0.1142
6400	10.0287	10.0000	0.0287
12800	9.9315	10.0000	0.0685
25600	10.0545	10.0000	0.0545
51200	9.9789	10.0000	0.0211

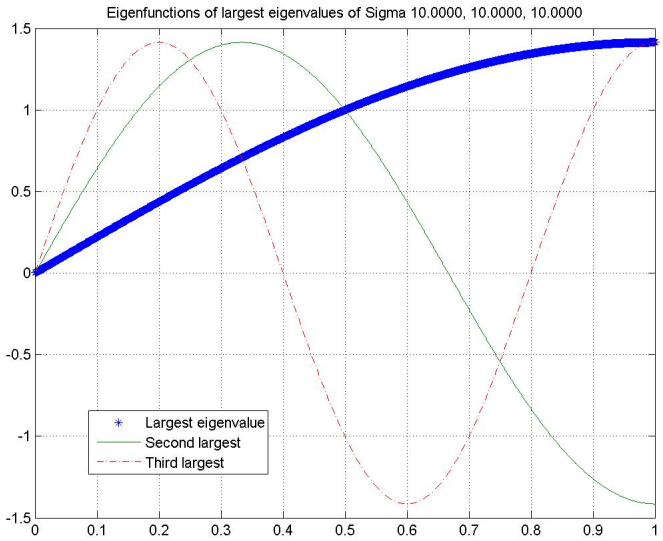

Figure 7.1. Eigenfunctions ϕ_1, ϕ_2, ϕ_3.

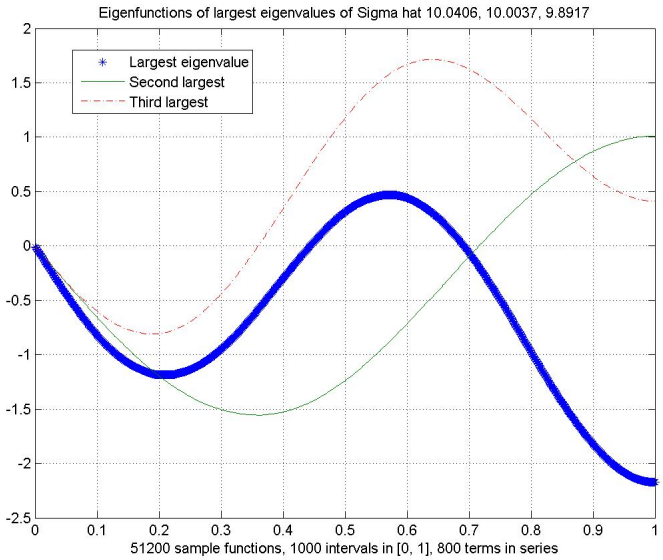

Figure 7.2. Eigenfunctions of $\hat{\sigma}^2_{1,N}, \hat{\sigma}^2_{2,N}, \hat{\sigma}^2_{3,N}, N = 51200$.

CHAPTER 8
APPROXIMATING BROWNIAN MOTION

8.1 Preliminary

The Wiener process is used to model Brownian motion. We consider each possible path for the motion as a function in the vector space $L^2[0,1]$, where $[0,1]$ is the time interval. A path W represents Brownian motion if

1. $W(0) = 0$,
2. $[W(t+\Delta t) - W(t)]$ and $[W(s+\Delta s) - W(s)]$ are stochastically independent for all $0 \le s < s+\Delta s < t < t+\Delta t \le 1$,
3. $W(t+\Delta t) - W(t)$ has distribution $N(0, \Delta t)$ for all $0 \le t < t+\Delta t \le 1$.

Two typical Wiener paths are shown in Figure 8.1.

We are interested in a step function approximation for Brownian motion. Perhaps the simplest version is obtained by partitioning $[0,1]$ into $[0,0]$ and M intervals I_m of equal length:

$$t_m = \frac{m}{M} \qquad m = 0, 1, \ldots, M, \tag{8.1}$$

$$I_m = (t_{m-1}, t_m] \qquad m = 1, 2, \ldots, M. \tag{8.2}$$

Let $W_M(0) = 0$ and

$$W_M(t_m) = W_M(t_{m-1}) + \frac{1}{\sqrt{M}} Z_m \qquad m = 1, 2, \ldots, M,$$

where the Z_m are iid $\mathrm{N}(0,1)$ random variables. Then

$$W_M(t_m) = \frac{1}{\sqrt{M}} \sum_{\ell=1}^{m} Z_\ell, \qquad m = 1, 2, \ldots, M,$$

$$\frac{1}{\sqrt{M}} Z_\ell \stackrel{d}{=} N\left(0, \frac{1}{M}\right),$$

$$W_M(t_m) \stackrel{d}{=} N\left(0, \frac{m}{M}\right).$$

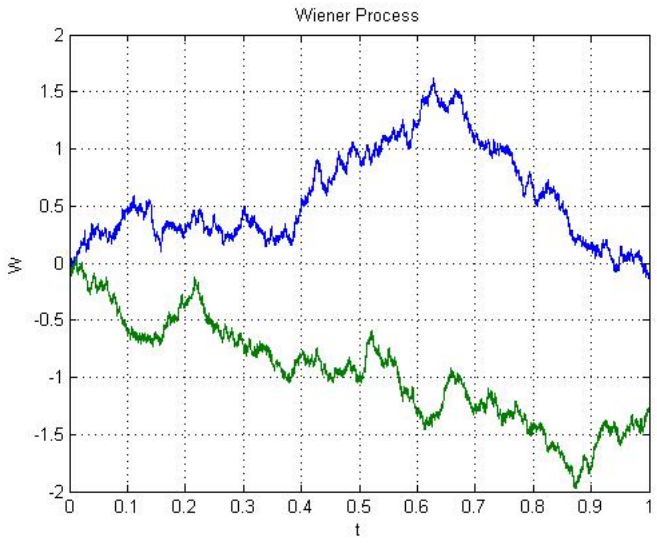

Figure 8.1. Two typical Wiener paths.

Define

$$1_m(t) = \begin{cases} 1 & t \in I_m, \\ 0 & \text{otherwise.} \end{cases} \qquad 0 \le t \le 1, \quad m = 1, \ldots, M.$$

Now we can define a step function:

$$W_M(t) = \sum_{m=1}^{M} 1_m(t)\, W_M(t_m) \qquad 0 \le t \le 1.$$

Let $W \in L^2[0,1]$ denote Brownian motion. The expected value of $W(t) = 0$ for all t. Let Σ be the covariance operator. Then

$$\begin{aligned} &\Sigma : L^2[0,1] \longrightarrow L^2[0,1], \\ &\Sigma = \mathbb{E}\,(W \otimes W). \qquad \text{use (5.1).} \end{aligned}$$

The eigenvalues σ_j^2 and corresponding eigenfunctions ϕ_j of Σ are (see [7])

$$\sigma_j^2 = \frac{1}{\left(j - \frac{1}{2}\right)^2 \pi^2}, \qquad j = 1, 2, \ldots \tag{8.3}$$

$$\phi_j(t) = \sqrt{2}\, \sin\left(\left(j - \frac{1}{2}\right) \pi\, t\right), \qquad 0 \le t \le 1. \tag{8.4}$$

The functions $\phi_j, j = 1, 2, \ldots$ form an orthonormal basis for $L^2[0,1]$. The question arises, can we find an M_0 such that $M \geq M_0$ ensures that the step function approximation is sufficiently accurate?

8.2 Conditions for Approximation

Let P_j be the orthogonal projection to $\phi_j, j = 1, 2, \ldots$. Then, for $x \in L^2[0,1]$,

$$(P_j x)(t) =< x, \phi_j > \phi_j(t) = \left(\int_0^1 x(s)\, \phi_j(s)\, ds \right) \phi_j(t). \qquad 0 \leq t \leq 1$$

For the step function approximation, we will have that

1. Σ_M is the covariance operator for the approximation to Brownian motion given by W_M,
2. J is a positive integer to be specified,
3. $\sigma^2_{M,1}, \ldots, \sigma^2_{M,J}$ are the J largest eigenvalues of Σ_M, in descending order,
4. $P_{M,j}$ is the eigenprojection for $\sigma^2_{M,j}, \quad j = 1, 2, \ldots, J$.

We have that both Σ and Σ_M are positive compact operators. Then the spectral theorem gives

$$\begin{aligned}
\Sigma_M &= \sum_{j=1}^{\infty} \sigma^2_{M,j} P_{M,j}, \\
\Sigma &= \sum_{j=1}^{\infty} \sigma^2_j P_j, \\
\Sigma_M - \Sigma &= \sum_{j=1}^{\infty} \left[\left(\sigma^2_{M,j} P_{M,j} - \sigma^2_j P_{M,j} \right) + \left(\sigma^2_j P_{M,j} - \sigma^2_j P_j \right) \right], \\
&= \sum_{j=1}^{\infty} \left[\left(\sigma^2_{M,j} - \sigma^2_j \right) P_{M,j} + \sigma^2_j \left(P_{M,j} - P_j \right) \right]. \qquad (8.5)
\end{aligned}$$

We seek a Σ_M that is a sufficiently accurate approximation for Σ. For this we specify a J and a_J. We will calculate b_J in (8.9) and require that

$$||P_j - P_{M,j}||_{\mathcal{L}} \leq a_J < 1 \qquad j = 1, 2, \ldots, J, \qquad (8.6)$$

$$|\sigma^2_j - \sigma^2_{M,j}| \leq b_J \qquad j = 1, 2, \ldots, J. \qquad (8.7)$$

Let $\rho_J > 0$ be a constant such that

$$\begin{aligned}\rho_J &< \frac{1}{2}\left(\sigma_J^2 - \sigma_{J+1}^2\right) \\ &= \frac{1}{2}\left[\frac{1}{\left(J-\frac{1}{2}\right)^2\pi^2} - \frac{1}{\left(J+\frac{1}{2}\right)^2\pi^2}\right] \\ &= \frac{J}{\left(J^2-\frac{1}{4}\right)^2\pi^2}. \end{aligned} \tag{8.8}$$

Let

$$b_J = \frac{J}{\left(J^2-\frac{1}{4}\right)^2\pi^2} \tag{8.9}$$

and

$$\Pi_M = \Sigma_M - \Sigma. \tag{8.10}$$

We will choose M so that

$$||\Pi_M||_{\mathcal{L}} = ||\Sigma_M - \Sigma||_{\mathcal{L}} \leq \frac{1}{4}\,\rho_J\,a_J. \tag{8.11}$$

We will show that M can be chosen this way.

8.3 Main Argument

Let $\Gamma_j,\ j = 1, \ldots, J$, be circles in the complex plane with centers σ_j^2 and equal radius ρ_J. Also, let Γ_{J+1} be the circle with center 0 and radius $\sigma_{J+1}^2 + \rho_j$. Let $\Omega_j, j = 1, \ldots, J+1$, be the simply connected open region whose boundary is Γ_j, i.e., all points inside but not on Γ_j. The $\Omega_j, j = 1, \ldots, J+1$ are pairwise disjoint. See Figure 8.2 (not to scale) for an example with $J = 3$, where Γ_1 is furthest to the right and the circles proceed until Γ_4, which is on the left and centered at 0. Note Ω_{J+1} contains all points of $\sigma(\Sigma)$ except $\sigma_j^2,\ j = 1, \ldots, J$. Now let

$$\begin{aligned}\Gamma &= \cup_{j=1}^{J+1}\Gamma_j, \\ \Omega &= \cup_{j=1}^{J+1}\Omega_j.\end{aligned}$$

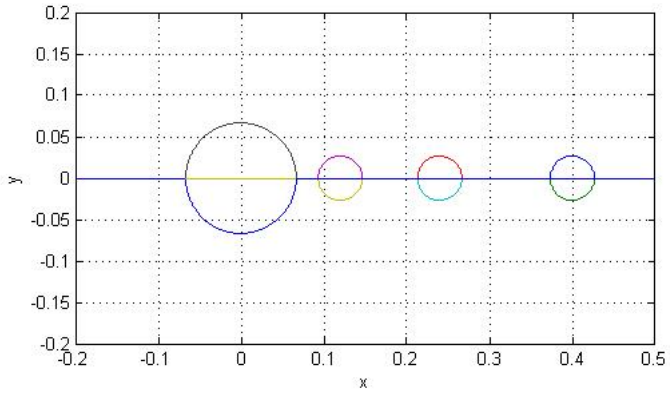

Figure 8.2. Example circles for approximating Brownian motion.

Then the boundary of Ω is Γ. For $z \in \rho(\Sigma)$, the resolvent set of Σ, recall as in (2.13), the resolvent operator is

$$R(z) = (zI - \Sigma)^{-1}. \tag{8.12}$$

Then, using the spectral theorem,

$$\begin{aligned} zI - \Sigma &= \sum_{k=1}^{\infty} (z - \sigma_k^2) P_k, \\ R(z) &= \sum_{k=1}^{\infty} \frac{1}{z - \sigma_k^2} P_k, \end{aligned} \tag{8.13}$$

and

$$\begin{aligned} ||R(z)||_{\mathcal{L}} &= \left|\left|\sum_{k=1}^{\infty} \frac{1}{z-\sigma_k^2}\, P_k\right|\right|_{\mathcal{L}} \\ &= \sup_k \left\{ \frac{1}{|z-\sigma_k^2|} \right\} \\ &= \frac{1}{\inf\limits_k \{|z-\sigma_k^2|\}} \\ &= \frac{1}{\text{dist}(z, \sigma(\Sigma))}. \end{aligned}$$

Note

$$\text{dist}(z, \sigma(\Sigma)) = \begin{cases} |z| & \text{if Re } z \leq 0 \\ \min\{|z-\sigma_k^2| : \ k = 1, 2, \dots\} & \text{if Re } z > 0 \end{cases}$$

Then

$$\begin{aligned} &\text{dist}(z, \sigma(\Sigma)) = \rho_J, && z \in \cup_{j=1}^{J} \Gamma_j \\ &\text{dist}(z, \sigma(\Sigma)) \geq \rho_J, && z \in \Gamma_{J+1} \text{ or outside } \Gamma \end{aligned}$$

so

$$\begin{aligned} &||R(z)||_{\mathcal{L}} = \frac{1}{\rho_J}, && z \in \cup_{j=1}^{J} \Gamma_j \\ &||R(z)||_{\mathcal{L}} \leq \frac{1}{\rho_J}. && z \in \Gamma_{J+1} \text{ or outside } \Gamma \end{aligned}$$

Then

$$||R(z)||_{\mathcal{L}} \leq \frac{1}{\rho_J}. \qquad z \in \mathbb{C} \setminus \Omega \tag{8.14}$$

This, combined with (8.11), gives

$$||\Pi_M R(z)||_{\mathcal{L}} \leq ||\Pi_M||_{\mathcal{L}}\, ||R(z)||_{\mathcal{L}} \leq \frac{1}{4}\, a_J \leq \frac{1}{4}. \qquad z \in \mathbb{C} \setminus \Omega \tag{8.15}$$

Therefore, $(I - \Pi_M R(z))^{-1}$ is well defined for all $z \in \mathbb{C} \setminus \Omega$, and

$$\begin{aligned} R(z)\,(I - \Pi_M\, R(z))^{-1} &= (zI - \Sigma)^{-1}(I - \Pi_M R(z))^{-1} && \text{Use (8.12)} \\ &= [(I - \Pi_M R(z))\,(zI - \Sigma)]^{-1} \\ &= [zI - \Sigma - \Pi_M R(z)(zI - \Sigma)]^{-1} \\ &= [zI - \Sigma - \Pi_M]^{-1} \\ &= [zI - \Sigma - (\Sigma_M - \Sigma)]^{-1} && \text{Use (8.10)} \\ &= (zI - \Sigma_M)^{-1}. \end{aligned}$$

Now we can define

$$R_M(z) = (zI - \Sigma_M)^{-1}, \qquad z \in \mathbb{C} \setminus \Omega \tag{8.16}$$

and we have that $\mathbb{C} \setminus \Omega \subset \rho(\Sigma_M)$, so

$$\sigma(\Sigma_M) = \mathbb{C} \setminus \rho(\Sigma_M) \subset \Omega.$$

Using (8.15) with a Neumann series as in [8], for $z \in \mathbb{C} \setminus \bar{\Omega}$,

$$\begin{aligned} R_M(z) &= R(z)[I - \Pi_m R(z)]^{-1} \\ &= R(z)[I + \Pi_M R(z) + (\Pi_m R(z))^2 + (\Pi_M R(z))^3 + \ldots] \\ &= R(z)[I + \Pi_M R(z)(I + \Pi_M R(z) + (\Pi_M R(z))^2 + \ldots)] \\ &= R(z)[I + \Pi_M R(z)(I - \Pi_M R(z))^{-1}]. \end{aligned}$$

Then

$$R_M(z) - R(z) = R(z)\Pi_M R(z)(I - \Pi_M R(z))^{-1}. \tag{8.17}$$

For any open $D \subset \mathbb{C} \setminus \bar{\Omega}$ and analytic function $f : D \longrightarrow \mathbb{C}$, define

$$f(\Sigma) = \frac{1}{2\pi i} \oint_\Gamma f(z) R(z)\, dz.$$

We also have, using (8.13),

$$\begin{aligned}
\frac{1}{2\pi i}\oint_{\Gamma_j} R(z)\,dz &= \frac{1}{2\pi i}\oint_{\Gamma_j}\left(\sum_{k=1}^{\infty}\frac{1}{z-\sigma_k^2}\,P_k\right)\,dz \\
&= \frac{1}{2\pi i}\sum_{k=1}^{\infty}\left(\oint_{\Gamma_j}\frac{1}{z-\sigma_k^2}\,dz\right)\,P_k \\
&= \frac{1}{2\pi i}\left(\oint_{\Gamma_j}\frac{1}{z-\sigma_j^2}\,dz\right)\,P_j \\
&= P_j. \qquad\qquad j=1,2,\ldots,J
\end{aligned}$$

So the orthogonal projection to the eigenspace of σ_j^2 is

$$P_j = \frac{1}{2\pi i}\oint_{\Gamma_j} R(z)\,dz. \qquad\qquad j=1,2,\ldots,J \tag{8.18}$$

Define (see item 4, page 44),

$$P_{M,j} = \frac{1}{2\pi i}\oint_{\Gamma_j} R_M(z)\,dz. \tag{8.19}$$

Using (8.15) and a standard inequality,

$$||(I-\Pi_M R(z))^{-1}||_{\mathcal{L}} \le \frac{1}{1-||\Pi_M R(z)||_{\mathcal{L}}} \le \frac{4}{3}. \tag{8.20}$$

We then have, for $j = 1, 2, \ldots, J$,

$$\begin{aligned}
||P_j - P_{M,j}||_{\mathcal{L}} &= \left|\left|\frac{1}{2\pi i}\oint_{\Gamma_j}(R(z) - R_M(z))\,dz\right|\right|_{\mathcal{L}} && \text{Use (8.18), (8.19)} \\
&= \frac{1}{2\pi}\left|\left|\oint_{\Gamma_j} R(z)\Pi_M R(z)(I - \Pi_M R(z))^{-1}\,dz\right|\right|_{\mathcal{L}} && \text{Use (8.17)} \\
&\leq \frac{1}{2\pi}\oint_{\Gamma_j} ||R(z)||^2_{\mathcal{L}}\,||\Pi_M||_{\mathcal{L}}\,||(I - \Pi_M R(z))^{-1}||_{\mathcal{L}}\,|dz| \\
&\leq \frac{1}{2\pi}\cdot\frac{1}{\rho_J^2}\cdot||\Pi_M||_{\mathcal{L}}\cdot\frac{4}{3}\oint_{\Gamma_j}|dz| && \text{Use (8.14), (8.20)} \\
&\leq \frac{1}{2\pi}\cdot\frac{1}{\rho_J^2}\cdot\frac{1}{4}\rho_J\,a_J\cdot\frac{4}{3}\cdot 2\pi\rho_J && \text{Use (8.11), see page 45} \\
&= \frac{1}{3}\cdot a_J < 1. && (8.21)
\end{aligned}$$

Then (8.6) is satisfied if (8.11) is satisfied.

We have that P_j is an orthogonal projection for the eigenvalue σ_j^2 to its one-dimensional eigenspace, for the operator Σ, and we have that $P_{M,j}$ is a projection (page 49). Then (8.21) and a result in [13] give that $P_{M,j}$ is also one-dimensional projection. It follows that Σ_M has only the one eigenvalue $\sigma^2_{M,j}$ inside Γ_j, and $P_{M,j}$ is the eigenprojection for $\sigma^2_{M,j}$ to the unique (up to a factor of -1) unit eigenfunction, for $j = 1, \ldots, J$. The eigenvalues $\sigma^2_{M,j}$ are those referred to in item 3 on page 44. We then have (see page 45)

$$|\sigma_j^2 - \sigma^2_{M,j}| \leq \rho_J. \qquad\qquad j = 1, 2, \ldots, J$$

Then (8.7) is satisfied if $\rho_J \leq b_J$, which we will specify using (8.8).

8.4 Choosing M

We now consider the requirement that we can satisfy (8.11) by choosing M sufficiently large. A general discussion of convergence such as $||\Sigma_m - \Sigma||_{\mathcal{L}} \longrightarrow 0$ can be found in [3](pages 247-249). We know from (8.3) - (8.4) that Σ is compact and $\Sigma = \mathbb{E}\,(W \otimes W)$ has only simple eigenvalues.

Here we will calculate only for our special case. Using the notation

$a \wedge b = \min(a, b)$, we have

$$\Sigma_M(s,t) = \sum_{j=1}^{M}\sum_{k=1}^{M} 1_j(s)\, 1_k(t) \left(\frac{j \wedge k}{M}\right),$$
$$\Sigma(s,t) = \sum_{j=1}^{M}\sum_{k=1}^{M} 1_j(s)\, 1_k(t)(s \wedge t).$$

Using the notation in (8.1) - (8.2),

$$\sup\{|\Sigma(s,t) - \Sigma_M(s,t)| : s \in I_j,\, t \in I_k\} = \frac{1}{M},$$

for any $j \in \{1, \ldots, M\}, \quad k \in \{1, \ldots, M\}$. Then

$$\sup\{|\Sigma(s,t) - \Sigma_M(s,t)| : s \in [0,1],\, t \in [0,1]\} = \frac{1}{M}.$$

We now have

$$\begin{aligned}
||\Sigma - \Sigma_M||_{\mathcal{L}}^2 &= \sup\, \{||\Sigma - \Sigma_M)f||_{L^2}^2 :\, ||f||_{L^2} = 1\} \\
&\leq \sup\, \left\{ \int_0^1 \int_0^1 \left[(\Sigma(s,t) - \Sigma_M(s,t))\, f(t)\right]^2 \, dt\, ds :\, ||f||_{L^2} = 1 \right\} \\
&\leq \sup\, \left\{ \frac{1}{M^2} \int_0^1 |f(t)|^2 \, dt :\, ||f||_{L^2} = 1 \right\} \\
&= \frac{1}{M^2}.
\end{aligned}$$

so that

$$||\Pi_M||_{\mathcal{L}} = ||\Sigma - \Sigma_M||_{\mathcal{L}} \leq \frac{1}{M}. \tag{8.22}$$

Then

$$M \geq \frac{4}{\rho_J a_J} \Longrightarrow \frac{1}{M} \leq \frac{1}{4}\rho_J a_J \Longrightarrow ||\Pi_M||_{\mathcal{L}} \leq \frac{1}{4}\rho_J a_J,$$

so, having chosen J and ρ_J as in (8.8), we can choose M to satisfy (8.11).

8.5 Summary

To make Σ_M a sufficiently accurate approximation for Σ, i.e., satisfying (8.6) and (8.7), first choose J as the desired number of eigenvalues (the largest J eigenvalues) to be considered. Then choose a_J for the desired accuracy in (8.6). Calculate b_J as

in (8.9). If the resulting b_J is not small enough for the desired accuracy in (8.7), choose a larger J and recalculate b_J as in (8.9). Then choose $\rho_J \in (0, b_J)$. Finally, choose integer M large enough so that

$$M \geq \frac{4}{\rho_J a_J}.$$

BIBLIOGRAPHY

[1] Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, Wiley, 1984.

[2] Bilodeau, M. and Brenner, D., *Theory of Multivariate Statistics*, Springer, 1999.

[3] Chatelin, F., *Spectral Approximation of Linear Operators*, Society for Industrial and Applied Mathematics, 2011 (first published by Academic Press, 1983).

[4] Cupidon, J., Gilliam, D., Eubank, R. and Ruymgaart, F., The Delta Method for Analytic Functions of Random Operators with Application to Functional Data, *Bernoulli* 13, 2007.

[5] Dauxois, J., Pousse, A. and Romain, Y., Asymptotic Theory for the Principal Component Analysis of a Vector Random Function: Some Applications to Statistical Inference, *Journal of of Multivariate Analysis*, 12, 1982.

[6] Dunford, N. and Schwartz, J.T., *Linear Operators, Part I: General Theory*, Wiley Classics Library, Wiley-Interscience, 1988.

[7] Gaines, G., Kaphle, K. and Ruymgaart, F., Application of a Delta-Method for Random Operators to Testing Equality of Two Covariance Operators, *Mathematical Methods of Statistics*, Vol. 20, No. 3, 2011

[8] Gilliam, D., Hohage, T., Ji, X. and Ruymgaart, F., The Fréchet Derivative of an Analytic Function of a Bounded Operator with some Applications, *International Journal of Mathematical Sciences*, 2009.

[9] Halmos, P.R., *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, American Mathematical Society, Second Edition, 1988 (first published by Chelsea Publishing Company, 1951 and 1957).

[10] Laha, R.G. and Rohatgi, V.K., *Probability Theory*, Wiley, 1979.

[11] Lax, P.D., *Functional Analysis*, Wiley, 2002.

[12] Muirhead, R.J., *Aspects of Multivariate Statistical Theory*, Wiley, 1982.

[13] Riesz, F. and Sz.-Nagy, B. *Functional Analysis*, Dover, 1990.

[14] Rynne, B. and Youngson, M., *Linear Functional Analysis*, Second Edition, Springer, 2008.

[15] van der Vaart, A.W., *Asymptotic Statistics*, Cambridge University Press, 1998.

[16] Watson, G.S., *Statistics on Spheres*, Wiley, 1983.

APPENDIX: COMPUTER CODE

```
function sim1

M = 1000;          % M intervals, t = 1/M, 2/M, ... 1, delta t = 1/M
nu = 3;            % The largest nu eigenvalues will be set equal. nu <= K
beta = 10;         % Largest eigenvalue, multiplicity nu
p = 0.05;          % Used in computer eigenvalue after th nu-th eigenvalue
K = 800;           % K eigen values/functions
N = 100;           % N sample functions
select_graph = 2; % 0 for no graph
                   % 1 for graphs of eigfuns of largest 3 true eigvals
                   % 2 for graphs of eigfuns of largest 3 approx eigvals
t = (0:M)./M;
phi = zeros(K, M+1);

for k = 1: K
  phi(k,:) = sqrt(2) * sin((k-.5)*pi.*t);   % eigenfunctions
end

lambda(1:nu) = beta;
for k = nu+1: K
  lambda(k) = beta/((k-nu+p)^2);      % eigenvalues of Sigma
end
X = zeros(N, M+1);           % sample functions
X_mean = zeros(1, M+1);
for n = 1: N
  for k = 1: K
    z = randn(1);
    X(n,:) = X(n,:) + sqrt(lambda(k))*z.*phi(k,:);
  end
X_mean = X_mean + X(n,:);    % This is sum. Correct to mean after loop.
end

X_mean = X_mean./N;          % Now X_mean really is mean.

for n = 1: N
  X(n,:) = X(n,:) - X_mean;
end
```

```
Sigma_hat = zeros(M,M);      % Sample covariance matrix
for n = 1: N
  Sigma_hat = Sigma_hat  +  X(n, 2:M+1)' * X(n, 2:M+1);
end
Sigma_hat = (1/N)*(1/M)*Sigma_hat;

% Use eigs to get 6 largest eigenvalues of Sigma_hat

[V D] = eigs (Sigma_hat);
lambda_hat = diag (D);

% Make eigenfunctions unit vectors (L2 norm = 1)

L2norm = zeros(1,3);
for k = 1: 3
  L2norm(k) = sqrt ((1/M) * V(:, k)' * V(:, k));
end
v1 = V(:, 1)' ./ L2norm(1);
v2 = V(:, 2)' ./ L2norm(2);
v3 = V(:, 3)' ./ L2norm(3);

% L2 inner products of eigenfunctions of 3 largest eigenvectors

in_prod = zeros(1, 3);
in_prod(1) = (1/M) * v1*(v2');
in_prod(2) = (1/M) * v2*(v3');
in_prod(3) = (1/M) * v3*(v1');

s = sprintf('Inner products of eigenfunctions (sample): 1*2, 2*3, 3*1 = ');
s1 = sprintf ('%s %f,%f,%f', s, in_prod(1), in_prod(2), in_prod(3))
s =sprintf('M = %d, nu = %d, beta = %f', M, nu, beta)
s1 =sprintf('p = %f, K = %d, N = %d', p, K, N)

evals = [lambda(1:6); lambda_hat'] % report original and approx. evals

if (select_graph == 1)
% Graph eigenfunctions of 3 largest eigenvalues of Sigma
  t1 = t(2:M+1);
  plot (t1, phi(1,2:M+1),'*', t1, phi(2,2:M+1),'-', t1, phi(3,2:M+1),'-.')
  grid on
  s1 = sprintf ('Eigenfunctions of largest eigenvalues of Sigma');
  s = sprintf ('%s %.4f, %.4f, %.4f', s1, beta, beta, beta);
```

```
  title (s)
  legend('Largest eigenvalue', 'Second largest', 'Third largest')

elseif (select_graph == 2)
% Graph eigenfunctions of 3 largest eigenvalues of Sigma_hat
  t1 = t(2:M+1);
  plot (t1, v1, '*', t1, v2, '-', t1, v3, '-.')
  grid on
  s1 = sprintf ('Eigenfunctions of largest eigenvalues of Sigma hat');
  s = sprintf ('%s %.4f, %.4f, %.4f', s1, lambda_hat(1:3));
  title (s)
  s1 = sprintf ('%d sample functions', N);
  s = sprintf ('%s, %d intervals in [0, 1], %d terms in series', s1, M, K);
  xlabel (s)
  legend('Largest eigenvalue', 'Second largest', 'Third largest')

end
```