Some Topics Concerning the Singular Value Decomposition and Generalized Singular Value Decomposition

by

Qing Huang

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2012 by the
Graduate Supervisory Committee:

Randy Eubank, Co-chair
Rosemary Renaut, Co-chair
Douglas Cochran
Anne Gelb
Dennis Young

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

This dissertation involves three problems that are all related by the use of the singular value decomposition (SVD) or generalized singular value decomposition (GSVD). The specific problems are (i) derivation of a generalized singular value expansion (GSVE), (ii) analysis of the properties of the χ^2 method for regularization parameter selection in the case of nonnormal data and (iii) formulation of a partial canonical correlation concept for continuous time stochastic processes.

The finite dimensional SVD has an infinite dimensional generalization to compact operators. However, the form of the finite dimensional GSVD developed in, e.g., Van Loan [40] does not extend directly to infinite dimensions as a result of a key step in the proof that is specific to the matrix case. Thus, the first problem of interest is to find an infinite dimensional version of the GSVD. One such GSVE for compact operators on separable Hilbert spaces is developed.

The second problem concerns regularization parameter estimation. The χ^2 method for nonnormal data is considered. A form of the optimized regularization criterion that pertains to measured data or signals with nonnormal noise is derived. Large sample theory for φ-mixing processes is used to derive a central limit theorem for the χ^2 criterion that holds under certain conditions. Departures from normality are seen to manifest in the need for a possibly different scale factor in normalization rather than what would be used under the assumption of normality. The consequences of our large sample work are illustrated by empirical experiments.

For the third problem, a new approach is examined for studying the relationships between a collection of functional random variables. The idea is based on the work of Sunder [36] that provides mappings to connect the ele-

ments of algebraic and orthogonal direct sums of subspaces in a Hilbert space. When combined with a key isometry associated with a particular Hilbert space indexed stochastic process, this leads to a useful formulation for situations that involve the study of several second order processes. In particular, using our approach with two processes provides an independent derivation of the functional canonical correlation analysis (CCA) results of Eubank and Hsing [13]. For more than two processes, a rigorous derivation of the functional partial canonical correlation analysis (PCCA) concept that applies to both finite and infinite dimensional settings is obtained.

To my Dad, Mom and Mengqi

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisors, Dr. Randy Eubank and Dr. Rosemary Renaut, for introducing me to areas of functional data analysis and discrete inverse problems and for their helpful advice, inspiration and support in my life of study. Without their continuous insight and guidance, this dissertation would not have been possible.

Many thanks to Dr. Douglas Cochran, Dr. Anne Gelb and Dr. Dennis Young for their willingness to be my committee members and for their time and efforts in reviewing my proposal and dissertation, in setting up the defense during the summer, and in giving helpful suggestions and discussions. I am also grateful to Dr. Ming-Hung Kao for his temporary service as my committee member for the oral exam.

Sincere thanks go to all the faculty who have taught me during my graduate study for their patience, inspiration and help and to my fellow graduate students for their friendship, help and support.

Last but not least, I wish to express my deepest gratitude to my family, my father Shisheng Huang, my mother Lin Zhou and Mengqi Hu for their endless love, understanding and support.

This research was partially supported by grants from National Science Foundation and block research funding from the School of Mathematical and Statistical Sciences, Arizona State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

Acronym		Page
SVD	singular value decomposition	1
GSVD	generalized singular value decomposition	1
CCA	canonical correlation analysis	1
SVE	singular value expansion	1
GSVE	generalized singular value expansion	1
UPRE	unbiased predictive risk estimation	21
GCV	generalized cross validation	21
FDA	Functional data analysis	31
RKHS	reproducing kernel Hilbert space	34
PCCA	partial canonical correlation analysis	34
CONS	complete orthonormal system	39
KS	Kolmogorov Smirnov	60
CDF	cumulative distribution function	60

LIST OF NOMENCLATURE

CHAPTER 1

OVERVIEW

1.1 Introduction

The singular value decomposition (SVD) of a matrix is one of the most important tools in mathematics. It has a long history dating back to the work of Sylvester [37], Autonne [2], Eckart and Young [10] and many others. There are now a number of extensions of the SVD involving factorization of more than one matrix, that are collectively termed generalized singular value decompositions: i.e., GSVDs. Of particular interest for the work in this dissertation is the one developed by Van Loan [40] and Paige and Saunders [29]; we will refer to it as the generalized singular value decomposition (GSVD). Other extensions include the cosine sine decomposition of a partitioned unitary matrix by Stewart [35], the product SVD proposed by Fernando and Hammarling [14] and the restricted SVD of three matrices introduced by Zha [42] and further developed by De Moor and Golub [9].

The SVD and GSVD have diverse applications involving areas such as signal processing, numerical computation, and statistics. In particular, canonical correlation analysis (CCA) in statistics is closely related to the SVD. The SVD and GSVD can also be applied to analyze and to solve least-squares problems in numerical analysis.

Strictly speaking, the SVD and GSVD refer to decompositions of finite dimensional matrices. Our interest is in similar decompositions for infinite dimensional compact operators. In that setting, we will refer to them as the singular value expansion (SVE) and generalized singular value expansion (GSVE), respectively. The SVE is an important theoretical tool with practical application in, e.g., the solution of integral equations.

This dissertation provides solutions for three distinct problems from mathematics and statistics. These problems are related through the principal tool that is used for their solution and their analysis: i.e., a SVD/ SVE or a GSVD.

1.2 The SVE, SVD and GSVD

Given the importance of the SVD/ SVE and GSVD for what follows, it will be worthwhile to first give explicit derivations of these decompositions. This is the subject of the current section. The derivation of a GSVE is the topic of the next chapter.

1.2.1 The SVE and SVD

Let us initially focus our attention on the SVD/ SVE. We will first derive the SVE of a compact linear operator in a Hilbert space. This will then allow us to derive the matrix version as a special case for a finite dimensional operator.

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be separable Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_{\mathcal{H}_i}$, and norms $|| \cdot ||_{\mathcal{H}_i}$, $i = 1, 2$. The set of bounded operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ will be denoted by $B(\mathcal{H}_1, \mathcal{H}_2)$.

Definition 1.2.1. *For an operator $A \in B(\mathcal{H}_1, \mathcal{H}_2)$, the adjoint A^* of A is the element of $B(\mathcal{H}_2, \mathcal{H}_1)$ that satisfies $\langle Ag, f \rangle_{\mathcal{H}_2} = \langle g, A^* f \rangle_{\mathcal{H}_1}$ for any $g \in \mathcal{H}_1$ and $f \in \mathcal{H}_2$.*

Definition 1.2.2. *An operator $A \in B(\mathcal{H}, \mathcal{H})$ is self-adjoint if $A = A^*$.*

Definition 1.2.3. *An operator $A \in B(\mathcal{H}, \mathcal{H})$ is unitary if $A^* A = A A^* = I$.*

Definition 1.2.4. *$A : \mathcal{H}_1 \to \mathcal{H}_2$ is compact if, for any bounded sequence $\{g_n\} \in \mathcal{H}_1$, the sequence $\{Ag_n\} \in \mathcal{H}_2$ contains a convergent subsequence.*

Our interest will be directed toward the case where A is compact. In that event, A^*A is compact, nonnegative definite and self-adjoint. As such, it has a pure point spectrum with nonzero eigenvalues λ_j^2 that provide eigenvalue-eigenfunction pairs (λ_j^2, g_j), $j = 1, 2, \ldots$, where $\lambda_1^2 \geq \lambda_2^2 \geq \cdots > 0$ and the g_j's are orthonormal [11].

Now, AA^* is also compact, nonnegative definite and selfadjoint with

$$A(A^*Ag_j) = (AA^*)Ag_j = \lambda_j^2 Ag_j, j = 1, 2, \ldots.$$

Thus, by letting $f_j = Ag_j/\lambda_j$, we obtain the pairs (λ_j^2, f_j), $j = 1, 2, \ldots$, that form the eigenvalue-eigenfunction system for AA^*. This follows from observing that

$$\begin{aligned} \langle f_i, f_j \rangle_{\mathcal{H}_2} &= \langle Ag_i/\lambda_i, Ag_j/\lambda_j \rangle_{\mathcal{H}_2} = (1/(\lambda_i\lambda_j))\langle g_i, A^*Ag_j \rangle_{\mathcal{H}_1} \\ &= (\lambda_j/\lambda_i)\langle g_i, g_j \rangle_{\mathcal{H}_1} = \delta_{ij}, \end{aligned}$$

where $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$.

Let $\mathrm{Im}(A) = \{Ag \in \mathcal{H}_2 : g \in \mathcal{H}_1\}$ and $\mathrm{Ker}(A) = \{g \in \mathcal{H}_1 : Ag = \mathbf{0}\}$ be the range and kernel of A, respectively. Since $\mathrm{Im}(AA^*) = \mathrm{Im}(A)$, for any $g \in \mathrm{Im}(A^*)$

$$\begin{aligned} Ag &= A\sum_{j=1}^{\infty} \langle g, g_j \rangle_{\mathcal{H}_1} g_j = \sum_{j=1}^{\infty} \langle g, g_j \rangle_{\mathcal{H}_1} Ag_j \\ &= \sum_{j=1}^{\infty} \lambda_j \langle g, g_j \rangle_{\mathcal{H}_1} f_j = \sum_{j=1}^{\infty} \lambda_j (f_j \otimes g_j) g, \end{aligned}$$

where the tensor product notation $(f \otimes g)h$ is defined by

$$(f \otimes g)h = \langle h, g \rangle_{\mathcal{H}} f. \tag{1.1}$$

We have just proved the following result.

Theorem 1.2.1. *[11] Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces and let A be a compact operator from $\mathcal{H}_1$ into $\mathcal{H}_2$. Then,*

$$A = \sum_{j=1}^{\infty} \lambda_j \left(f_j \otimes g_j \right) \tag{1.2}$$

with

(i) $\{\lambda_j^2\}$ *the non-zero eigenvalues of* A^*A *and* AA^*,

(ii) $\{g_j\}$ *orthonormal eigenfunctions of* A^*A *and*

(iii) $\{f_j\}$ *orthonormal eigenfunctions of* AA^* *satisfying* $Ag_j = \lambda_j f_j$.

The representation of A in (1.2) is the SVE of the operator. The triples (λ_j, f_j, g_j), $j = 1, 2, \ldots$ are sometimes called a singular system for A [11]. The $\{\lambda_j\}$ are termed singular values while the $\{f_j\}$ and $\{g_j\}$ are the left and right singular functions, respectively.

The SVE plays a fundamental role in the analysis and solution of least squares problems through its connection to an operator's generalized inverse. A Moore-Penrose generalized inverse A^- of a linear operator A satisfies

$$AA^-A = A, A^-AA^- = A^-, (AA^-)^* = AA^- \text{ and } (A^-A)^* = A^-A.$$

Definition 1.2.5. *[11] Let $\tilde{A} := A|_{\mathrm{Ker}(A)^\perp}$. The Moore-Penrose generalized inverse of a linear operator A is the unique linear extension of $\tilde{A}^{-1}$ to*

$$\mathrm{D}(A^-) := \mathrm{Im}(A) \oplus \mathrm{Im}(A)^\perp.$$

A characterization of A^- using the SVE of A is provided by the Theorem 1.2.2.

Theorem 1.2.2. *[11] Let (λ_j, f_j, g_j) be a singular system for the compact linear operator $A : \mathcal{H}_1 \to \mathcal{H}_2$. Then,*

(i) $f \in \mathrm{D}(A^-)$ if and only if the Picard condition $\sum_{j=1}^{\infty} \frac{|\langle f, f_j\rangle_{\mathcal{H}_2}|^2}{\lambda_j^2} < \infty$ is satisfied.

(ii) for $f \in \mathrm{D}(A^-)$,

$$A^- f = \sum_{j=1}^{\infty} \frac{\langle f, f_j\rangle_{\mathcal{H}_2}}{\lambda_j} g_j. \tag{1.3}$$

Proof. For (i) first suppose $f \in \mathrm{D}(A^-)$ and that P is the orthogonal projector onto $\overline{\mathrm{Im}(A)}$. Then, $Pf \in \mathrm{Im}(A)$ and there exists $g \in \mathcal{H}_1$ such that $Ag = Pf$. For $g \in \mathrm{Ker}(A)^{\perp}$,

$$Pf = \sum_{j=1}^{\infty} \langle f, f_j\rangle_{\mathcal{H}_2} f_j = Ag = \sum_{j=1}^{\infty} \lambda_j \langle g, g_j\rangle_{\mathcal{H}_1} f_j \tag{1.4}$$

so that $\langle f, f_j\rangle_{\mathcal{H}_2} = \lambda_j \langle g, g_j\rangle_{\mathcal{H}_1}$. Since $\langle g, g_j\rangle_{\mathcal{H}_1}$ are the generalized Fourier coefficients for g under the $\{g_j\}$ basis for $\overline{\mathrm{Im}(A)}$, $\{\langle f, f_j\rangle_{\mathcal{H}_2}/\lambda_j\} \in l^2$, the Hilbert space of square summable sequences. Consequently, the Picard condition holds.

To go in the other direction, write $g = \sum_{j=1}^{\infty} \frac{\langle f, f_j\rangle_{\mathcal{H}_2}}{\lambda_j} g_j \in \mathcal{H}_1$ to see that

$$Ag = \sum_{j=1}^{\infty} \langle f, f_j\rangle_{\mathcal{H}_2} f_j = Pf \in \mathrm{Im}(A)$$

and $f \in \mathrm{D}(A^-)$. Finally, for part (ii) use (1.4) to obtain

$$g = A^- Pf = A^- f = \sum_{j=1}^{\infty} \frac{\langle f, f_j\rangle_{\mathcal{H}_2}}{\lambda_j} g_j.$$

□

The matrix SVD is a special case of the operator SVE. We state this formally as the next corollary. We use $\mathbf{V}^H$ to denote the complex-conjugate transpose of any matrix $\mathbf{V}$.

Corollary 1.2.1. *If* $\mathbf{A} \in \mathbb{C}^{m\times n}$ *has rank* $k \leq \min(m,n) < \infty$*, there exist unitary matrices*

$$\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_m] \in \mathbb{C}^{m\times m}$$

and

$$\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_n] \in \mathbb{C}^{n\times n}$$

such that

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H = \sum_{j=1}^{k} \lambda_j \mathbf{u}_j \mathbf{v}_j^H,$$

where $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_k, 0, \ldots) \in \mathbb{R}^{m\times n}$ *is a diagonal matrix for which* $\lambda_1, \ldots, \lambda_k$ *satisfy* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$.

Proof. Since $\mathbf{A}$ is of finite rank, it is necessarily compact. Thus, by Theorem 1.2.1, $\mathbf{A} = \sum_{j=1}^{k} \lambda_j (\mathbf{u}_j \otimes \mathbf{v}_j)$ with $\mathbf{u}_j \in \mathbb{C}^m$ and $\mathbf{v}_j \in \mathbb{C}^n$. The tensor-product operator $\otimes$ in this case is just the vector outer product: i.e., $\mathbf{u}_j \otimes \mathbf{v}_j = \mathbf{u}_j \mathbf{v}_j^H$. □

Actually, an adjoint of an operator between finite-dimensional spaces is related with the complex conjugate transpose in the following way. Let an operator $A \in B(\mathbb{C}^n, \mathbb{C}^m)$ with standard basis for $\mathbb{C}^n$ and $\mathbb{C}^m$ and $M(A)$ be the matrix representation of A. Then, $M(A^*) = M(A)^H$.

1.2.2 An Illustration of the SVE

The SVE is an important tool that arises in, e.g., the solution of Fredholm integral equations of the first kind: i.e.,

$$\int_0^1 K(s,t)g(t)dt = f(s), 0 \leq s \leq 1, \tag{1.5}$$

where real valued $f, g \in L^2[0,1]$, the Hilbert space of square integrable functions on $[0,1]$. The $L^2[0,1]$ inner product is

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt$$

and the corresponding norms for $g \in L^2[0,1]$ and $K \in L^2([0,1] \times [0,1])$ are

$$\|g\|^2 = \langle g, g \rangle, \|K\|^2 = \int_0^1 \int_0^1 |K(s,t)|^2 dtds.$$

The SVE gives

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_j f_j(s) g_j(t).$$

Then, by Theorem 1.2.2, under the Picard condition, we have

$$g(t) = \sum_{i=1}^{\infty} \frac{\langle f_j, f \rangle}{\lambda_j} g_j(t). \tag{1.6}$$

For computational purposes, (1.5) is typically approximated using a finite discrete approximation in place of the integral equation. For example, if we use a quadrature approximation for (1.5), this produces the linear equations

$$\sum_{j=1}^{n} w_j K(s_i, t_j) g(t_j) = f(s_i), i = 1, \ldots, m,$$

with $\{w_j\}$ the quadrature weights. Equivalently, this can be written in the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ with elements $a_{ij} = w_j K(s_i, t_j)$, $\mathbf{x} \in \mathbb{R}^n$ with elements $x_j = g(t_j)$ and $\mathbf{b} \in \mathbb{R}^m$ with elements $b_i = f(s_i)$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Then, the SVD of $\mathbf{A}$ can be employed to solve or analyze the linear system.

Hansen [16] discusses the relationship between the SVD and SVE for (1.5) when K is square integrable. He shows how to use the SVD to compute an approximation to the SVE by a universal expansion method. Specifically, let

$\{f_1^0, \ldots, f_n^0\}$ be orthonormal functions in $L^2[0,1]$ and let the matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ have elements

$$a_{ij} = \langle f_i^0, Kf_j^0 \rangle, i, j = 1, \ldots, n. \tag{1.7}$$

If the SVD of $\mathbf{A}$ gives the singular system $(\tilde{\lambda}_j, \mathbf{u}_j, \mathbf{v}_j)$ for $j = 1, \ldots, n$, then the singular system (λ_j, f_j, g_j) for K is approximated by

$$(\tilde{\lambda}_j, \sum_{i=1}^{n} u_{ij} f_i^0(s), \sum_{i=1}^{n} v_{ij} f_i^0(t)), j = 1, \ldots, n.$$

As an example, consider the first kind Fredholm integral equation

$$\int_0^a \frac{1}{1+s^2t^2} g(t)dt = f(s), 0 \le s \le a. \tag{1.8}$$

Here, the kernel function is $K(s,t) = (1+s^2t^2)^{-1}$. If $a = \infty$, we have that

$$\int_0^\infty \int_0^\infty K(s,t)^2 dsdt = \infty, \tag{1.9}$$

in which case $K(s,t)$ is not square integrable and hence there exists no SVE. However, if we take $a = 5$ for instance, $K(s,t)$ is square integrable restricted to $[0,5] \times [0,5]$ and then we can use the SVD to compute the approximation to the SVE of the integral operator corresponding to $K(s,t)$.

Now, let the interval $[0,5]$ be divided into n subintervals with the same length $h = 5/n$. The orthonormal functions $\{f_i^0\}$ are chosen to be

$$f_i^0(x) = \begin{cases} h^{-1/2}, & x \in ((i-1)h, ih], i = 1, \ldots, n, \\ 0, & \text{otherwise.} \end{cases}$$

Thus, (1.7) can be written as

$$a_{ij} = h^{-1} \int_{(i-1)h}^{ih} \int_{(j-1)h}^{jh} K(s,t)dsdt, i, j = 1, \ldots, n, \tag{1.10}$$

and we can apply Simpson's rule to approximate the double integral in (1.10) to produce a matrix $\mathbf{A}$ for use by the SVD. We have listed some computed singular values $\tilde{\lambda}_j$ for different choices of n in Table 1.1. For comparison, we also list some computed singular values with increasing a at $n = 128$ in Table 1.2. The effect of (1.9) is seen here in the lack of convergence for the singular values. We do not expect convergence since the SVE does not exist as $a \to \infty$.

n	$\tilde{\lambda}_1$	$\tilde{\lambda}_2$	$\tilde{\lambda}_3$	$\tilde{\lambda}_4$
16	1.6352	0.7046	0.2323	0.0649
32	1.6408	0.7253	0.2591	0.0877
64	1.6421	0.7300	0.2643	0.0916
128	1.6424	0.7311	0.2656	0.0926

Table 1.1: Computed singular values for $a = 5$ with increasing n

a	$\tilde{\lambda}_1$	$\tilde{\lambda}_2$	$\tilde{\lambda}_3$	$\tilde{\lambda}_4$
5	1.6424	0.7311	0.2656	0.0926
10	1.7768	0.9827	0.4575	0.2026
20	1.8579	1.1408	0.5772	0.2598
40	1.8696	1.1134	0.4814	0.1567

Table 1.2: Computed singular values for $n = 128$ with increasing a for example (1.8)

If we consider another first kind Fredholm integral equation

$$\int_0^\infty e^{-s^2+0.5st-t^2} g(t)dt = f(s), 0 \le s \le a, \tag{1.11}$$

where the kernel function is $K(s,t) = e^{-s^2+0.5st-t^2}$. Since this $K(s,t)$ is square integrable even at $a = \infty$, then the SVE exists. Table 1.3 shows some computed singular values with increasing a at $n = 128$ and convergence is indicated.

a	$\tilde{\lambda}_1$	$\tilde{\lambda}_2$	$\tilde{\lambda}_3$	$\tilde{\lambda}_4$
5	0.6851	0.0364	0.0018	0.0001
10	0.6849	0.0363	0.0018	0.0001
20	0.6840	0.0359	0.0018	0.0001
40	0.6805	0.0342	0.0015	0.0001

Table 1.3: Computed singular values for $n = 128$ with increasing a for example (1.11)

1.2.3 The GSVD

The GSVD was first established by Van Loan [40] and is sometimes referred to as the B-singular value decomposition. It arises in the problem of finding $\zeta \geq 0$ for which $\det(\mathbf{A}^T\mathbf{A} - \zeta^2\mathbf{B}^T\mathbf{B}) = 0$, where $\mathbf{A} \in \mathbb{C}^{m_a \times n}$ and $\mathbf{B} \in \mathbb{C}^{m_b \times n}$. The result and proof of the GSVD for the finite dimensional case follows the developments in Paige and Saunders [30]. Here, and subsequently, we use I_n to denote an n-dimensional identity matrix.

Theorem 1.2.3. *[30]For matrices $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{B} \in \mathbb{C}^{p\times n}$, let $\mathbf{C}^H = (\mathbf{A}^H, \mathbf{B}^H)$ with $\mathbf{C}$ having rank k. Let $\mathbf{D_C} \in \mathbb{R}^{k\times k}$ be a diagonal matrix with the nonzero singular values of $\mathbf{C}$ as its diagonal elements. Then, there exist unitary matrices $\mathbf{U_A} \in \mathbb{C}^{m\times m}$, $\mathbf{U_B} \in \mathbb{C}^{p\times p}$, $\mathbf{V} \in \mathbb{C}^{n\times n}$ and $\mathbf{Q} \in \mathbb{C}^{k\times k}$ such that*

$$\mathbf{U}_\mathbf{A}^H\mathbf{A}\mathbf{V} = \mathbf{S_A}(\mathbf{Q}^H\mathbf{D_C}, 0), \qquad \mathbf{U}_\mathbf{B}^H\mathbf{B}\mathbf{V} = \mathbf{S_B}(\mathbf{Q}^H\mathbf{D_C}, 0),$$

where

$$\mathbf{S_A} = \begin{pmatrix} I_r & & \\ & \mathbf{D_A} & \\ & & O_\mathbf{A} \end{pmatrix}, \qquad \mathbf{S_B} = \begin{pmatrix} O_\mathbf{B} & & \\ & \mathbf{D_B} & \\ & & I_{k-r-s} \end{pmatrix},$$

for integers r *and* s *satisfying* $0 \leq r, s \leq k$, $O_{\mathbf{A}}$ *and* $O_{\mathbf{B}}$ *matrices of all zeros and* $\mathbf{D_A}$ *and* $\mathbf{D_B}$ *diagonal matrices with diagonal elements* $\lambda_{\mathbf{A},(r+1)}, \ldots, \lambda_{\mathbf{A},(r+s)}$ *and* $\lambda_{\mathbf{B},(r+1)}, \ldots, \lambda_{\mathbf{B},(r+s)}$, *respectively, satisfying*

$$1 > \lambda_{\mathbf{A},(r+1)} \geq \cdots \geq \lambda_{\mathbf{A},(r+s)} > 0, \quad 0 < \lambda_{\mathbf{B},(r+1)} \leq \cdots \leq \lambda_{\mathbf{B},(r+s)} < 1, \tag{1.12}$$

and

$$\lambda_{\mathbf{A},i}^2 + \lambda_{\mathbf{B},i}^2 = 1, i = r+1, \ldots, r+s. \tag{1.13}$$

Proof. The SVD of $\mathbf{C}$ allows us to write

$$\mathbf{C} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} = \mathbf{U} \begin{pmatrix} \mathbf{D_C} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{V}^H$$

with $\mathbf{U}$ and $\mathbf{V}$ unitary matrices and $\mathbf{D_C}$ a diagonal matrix with the nonzero singular values of $\mathbf{C}$ as its diagonal elements. Now, we can partition $\mathbf{U}$ as

$$\mathbf{U} = \begin{pmatrix} \mathbf{U_{A1}} & \mathbf{U_{A2}} \\ \mathbf{U_{B1}} & \mathbf{U_{B2}} \end{pmatrix}.$$

Let the SVD of $\mathbf{U_{A1}}$ be

$$\mathbf{U_{A1}} = \mathbf{U_A S_A Q}^H$$

and write

$$\mathbf{U_{B1}} = \mathbf{U_{B1} Q Q}^H = \mathbf{U_B L Q}^H,$$

where $\mathbf{U_{B1} Q} = \mathbf{U_B L}$ with $\mathbf{U_B}$ unitary and $\mathbf{L}$ lower triangular. Then,

$$\begin{pmatrix} \mathbf{U_{A1}} \\ \mathbf{U_{B1}} \end{pmatrix} = \begin{pmatrix} \mathbf{U_A} & 0 \\ 0 & \mathbf{U_B} \end{pmatrix} \begin{pmatrix} \mathbf{S_A} \\ \mathbf{L} \end{pmatrix} \mathbf{Q}^H, \tag{1.14}$$

which has orthonormal columns since $\mathbf{U}$ is unitary. Thus,

$$\begin{pmatrix} \mathbf{U}_{\mathbf{A}1}^H & \mathbf{U}_{\mathbf{B}1}^H \end{pmatrix} \begin{pmatrix} \mathbf{U}_{\mathbf{A}1} \\ \mathbf{U}_{\mathbf{B}1} \end{pmatrix} = I_k. \tag{1.15}$$

Combining (1.15) with (1.14) gives

$$\mathbf{Q} \begin{pmatrix} \mathbf{S}_{\mathbf{A}}^H & \mathbf{L}^H \end{pmatrix} \begin{pmatrix} \mathbf{U}_{\mathbf{A}}^H & 0 \\ 0 & \mathbf{U}_{\mathbf{B}}^H \end{pmatrix} \begin{pmatrix} \mathbf{U}_{\mathbf{A}} & 0 \\ 0 & \mathbf{U}_{\mathbf{B}} \end{pmatrix} \begin{pmatrix} \mathbf{S}_{\mathbf{A}} \\ \mathbf{L} \end{pmatrix} \mathbf{Q}^H = I_k.$$

From this we obtain $\mathbf{S}_{\mathbf{A}}^H\mathbf{S}_{\mathbf{A}} + \mathbf{L}^H\mathbf{L} = I_k$; i.e., $\mathbf{L}^H\mathbf{L} = I_k - \mathbf{S}_{\mathbf{A}}^H\mathbf{S}_{\mathbf{A}} = \mathbf{S}_{\mathbf{B}}^H\mathbf{S}_{\mathbf{B}}$ and, hence, $\mathbf{L} = \mathbf{S}_{\mathbf{B}}$. □

1.3 Regularization

In this section, we give a directed introduction to the use of regularization for the solution of ill-posed inverse problems. Such problems arise in many application areas such as medical imaging. For example, regularization is essential when attempting to reconstruct a sharp image from an observed blurred image.

In a general sense, inverse problems concern recovery of the interior information (input/source) from the observed information (output/data) through some connecting system. Taking (1.5) for instance, if $K(s,t)$ and $f(s)$ are given as the connecting system and output, finding $g(t)$ becomes an inverse problem. Most inverse problems are ill-posed which forces us to use techniques that return stable, approximate solutions. A problem is defined to be ill-posed if it is not well-posed. The latter concept is defined by Hadamard [11] as follows.

Definition 1.3.1. *A linear problem is well-posed if*

(i) for all admissible data, a solution exists,

(ii) for all admissible data, the solution is unique and

(iii) the solution depends continuously on the data.

If any one of the three properties is violated, we call the problem ill-posed.

In this dissertation, we consider the discrete linear inverse problem which can be expressed as

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{1.16}$$

for $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $m \geq n$. The known right hand side of (1.16) consists of unknown true data $\mathbf{b}_{\text{true}} \in \mathbb{R}^m$ and noise $\boldsymbol{\varepsilon} \in \mathbb{R}^m$; i.e.,

$$\mathbf{b} = \mathbf{b}_{\text{true}} + \boldsymbol{\varepsilon}. \tag{1.17}$$

The matrix $\mathbf{A}$ in (1.16) derives from the underlying connective system and is assumed to be known. The vector $\mathbf{x}$ is the solution we want to obtain given the noisy data and the system $\mathbf{A}$. Ideally, the solution should be close to the true solution $\mathbf{x}_{\text{true}}$ that satisfies

$$\mathbf{A}\mathbf{x}_{\text{true}} = \mathbf{b}_{\text{true}}. \tag{1.18}$$

In practice, all we are given is the noisy data instead of the true data. In that case, it is possible that (i) no solution exists for (1.16), (ii) the solution of (1.16) is not unique or (iii) small perturbations of the data lead to a large perturbations in the solution. Problem (iii) is often reflected in a large condition number for $\mathbf{A}$. In all three cases, it is necessary to use regularization to find a best (approximate) solution.

1.3.1 Regularization Methods

There is a large literature on regularization techniques, e.g., [11], [17] and [41]. Here we illustrate the impact of regularization through the truncated SVD and the Tikhonov approach.

1.3.1.1 Truncated SVD

Writing the SVD (from Corollary 1.2.1) of $\mathbf{A}$ in (1.16) as $\mathbf{A} = \sum_{j=1}^{k} \lambda_j \mathbf{u}_j \mathbf{v}_j^T$ with k being the numerical rank of $\mathbf{A}$, then,

$$\mathbf{x} = \sum_{j=1}^{k} \frac{\mathbf{u}_j^T \mathbf{b}}{\lambda_j} \mathbf{v}_j = \sum_{j=1}^{k} \frac{\mathbf{u}_j^T \mathbf{b}_{\text{true}}}{\lambda_j} \mathbf{v}_j + \sum_{j=1}^{k} \frac{\mathbf{u}_j^T \boldsymbol{\varepsilon}}{\lambda_j} \mathbf{v}_j. \tag{1.19}$$

From (1.19), it becomes obvious that the solution is contaminated by the second term of the last expression, especially when λ_j is much less than $\mathbf{u}_j^T \boldsymbol{\varepsilon}$.

Recall the Picard condition in Theorem 1.2.2 for the existence of the solution:

$$\sum_{j=1}^{\infty} \frac{|\mathbf{u}_j^T \mathbf{b}|^2}{\lambda_j^2} < \infty. \tag{1.20}$$

This leads to the introduction of the discrete Picard condition for discrete linear inverse problems, which is often given us.

Definition 1.3.2. *[17] Let τ denote the level at which the computed singular values λ_j level off due to rounding errors. Then the discrete Picard condition is satisfied if, for all singular values greater than τ, the corresponding coefficients $|\mathbf{u}_j^T \mathbf{b}|$, on average, decay faster than the λ_j.*

So the idea behind the truncated SVD [19] is to remove those components in (1.19) that are dominated by the noise, or that violate the discrete Picard

condition. Thus, the approximate solution is

$$\mathbf{x}_\kappa = \sum_{j=1}^{\kappa} \frac{\mathbf{u}_j^T \mathbf{b}}{\lambda_j} \mathbf{v}_j. \tag{1.21}$$

The integer parameter κ in (1.21), called a regularization parameter, needs to be chosen so that the noise-dominated terms in (1.19) are discarded to keep them from unduly perturbing the true solution.

To illustrate using the truncated SVD we work through a Shaw test problem. The Shaw matrix arises from discretization of the Fredholm integral equation of the first kind on $[-\pi/2, \pi/2]$ with

$$\begin{aligned} K(s,t) &= (\cos(s) + \cos(t))^2 \left(\frac{\sin(u)}{u} \right)^2, \\ u &= \pi(\sin(s) + \sin(t)) \end{aligned}$$

and

$$g(t) = 2e^{-6(t-0.8)^2} + e^{-2(t+0.5)^2}.$$

After producing $\mathbf{A}$ and $\mathbf{x}_{\text{true}}$, $\mathbf{b}_{\text{true}}$ is found by multiplying $\mathbf{A}$ and $\mathbf{x}_{\text{true}}$. $\mathbf{A}$, $\mathbf{x}_{\text{true}}$ and $\mathbf{b}_{true}$ are output using the regularization toolbox [18] from Matlab. Now we use the Shaw matrix for $\mathbf{A} \in \mathbb{R}^{40\times 40}$ with $\mathbf{x}_{\text{true}}$ and $\mathbf{b}_{\text{true}}$ shown in Figure 1.1.

The right-hand side of (1.16) is obtained by $\mathbf{b}_{\text{true}} + \boldsymbol{\varepsilon}$ where the noise vector $\boldsymbol{\varepsilon} \sim N_{40}(0, 10^{-6}I)$ as shown in Figure 1.2. Here, $\boldsymbol{\varepsilon} \sim N_{40}(0, 10^{-6}I)$ denotes that $\boldsymbol{\varepsilon}$ follows 40-variate normal distribution with mean 0 and covariance $10^{-6}I$. Even for such low noise levels, the ill-posedness is reflected in the solutions.

Figure 1.3 gives the approximate solutions $\mathbf{x}_\kappa$ produced by the truncated SVD for different values of κ along with the true solution. Figure 1.4 gives the

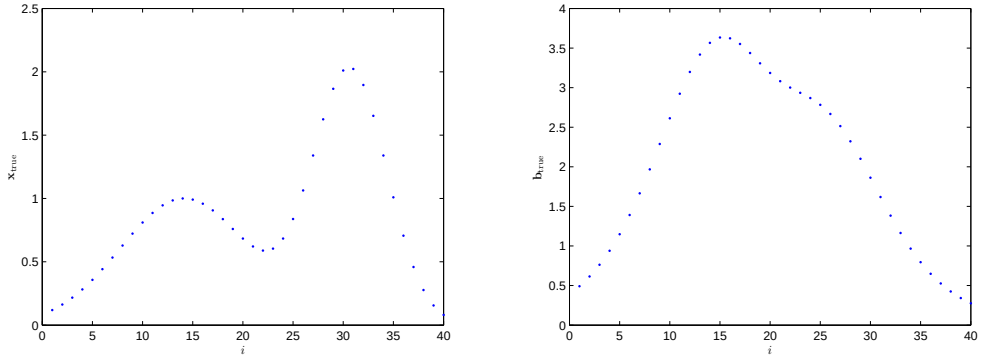


Figure 1.1: $\mathbf{x}_{\text{true}}$ and $\mathbf{b}_{\text{true}}$

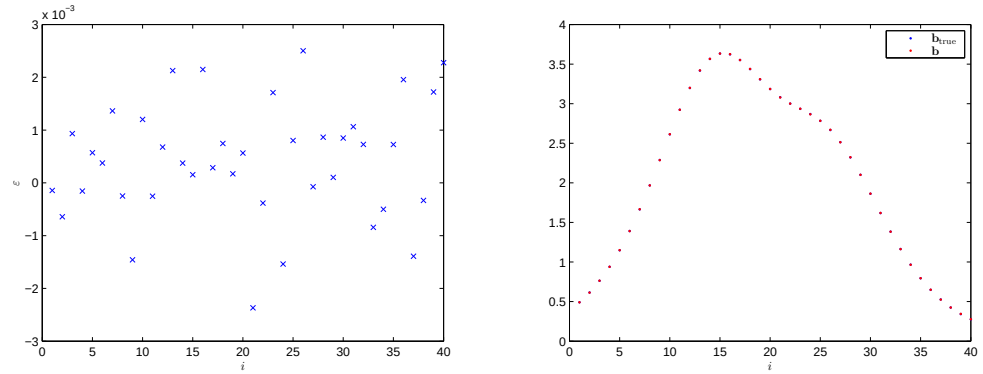
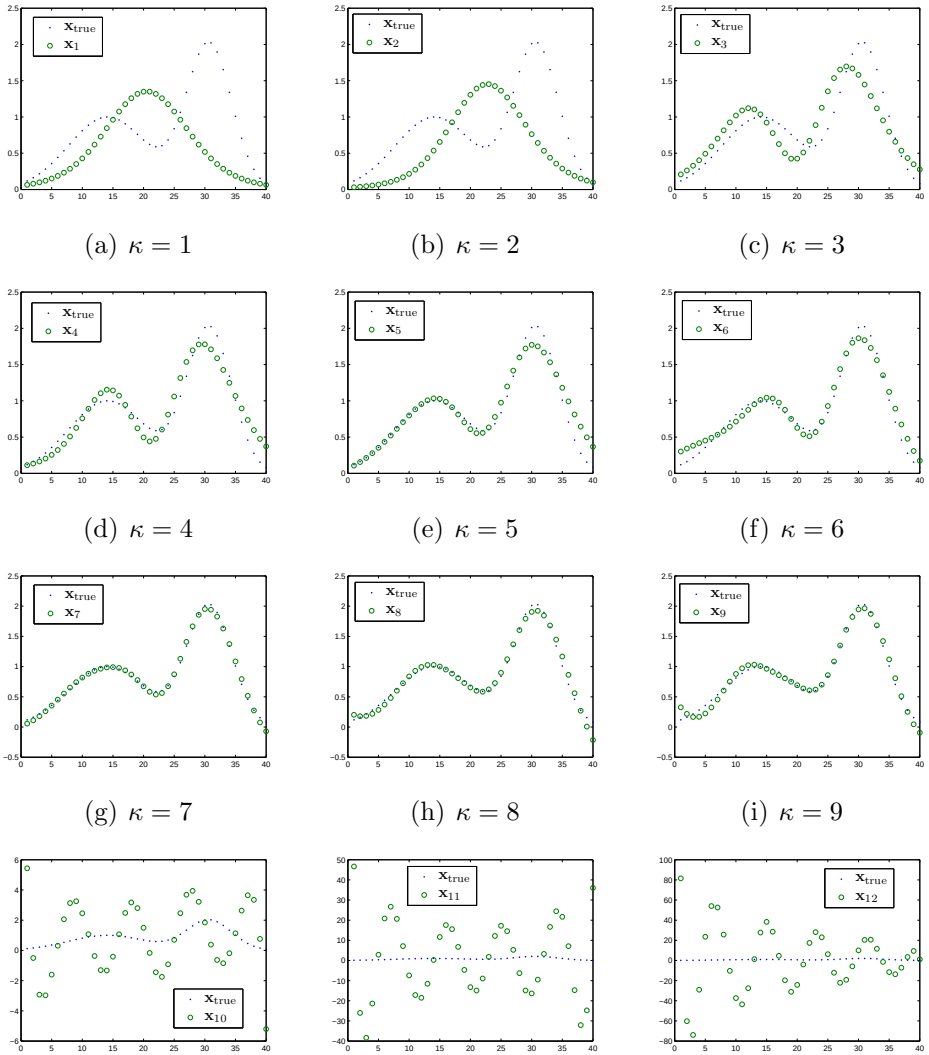


Figure 1.2: $\boldsymbol{\varepsilon}$ and $\mathbf{b}$

corresponding $\mathbf{b}_\kappa = \mathbf{A}\mathbf{x}_\kappa$ as compared with $\mathbf{b}_{\text{true}}$. From the two figures, we can see that $\mathbf{b}_\kappa$ is still a good approximation of $\mathbf{b}_{\text{true}}$ even when $\mathbf{x}_\kappa$ is a poor approximation to $\mathbf{x}_{\text{true}}$. This phenomenon is indicative of the ill-posedness of the problem and demonstrates that it is important to make a good choice for the regularization parameter κ. In Figure 1.3, a good approximate solution is achieved at $\kappa = 9$ and the impact of noise on the solution is clearly detectable for $\kappa = 10, 11, 12$ where these approximate solutions oscillate far from the true solution.



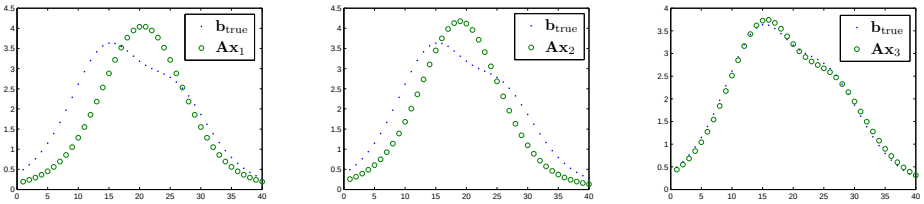
(a) $\kappa = 1$ (b) $\kappa = 2$ (c) $\kappa = 3$

(d) $\kappa = 4$ (e) $\kappa = 5$ (f) $\kappa = 6$

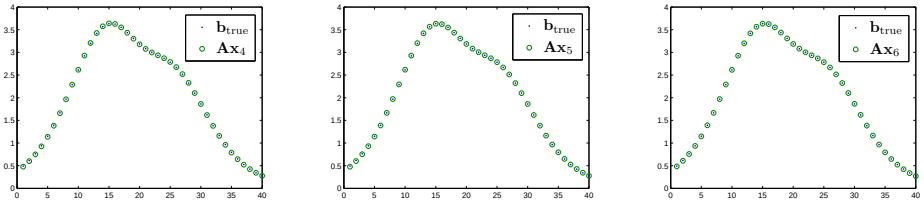
(g) $\kappa = 7$ (h) $\kappa = 8$ (i) $\kappa = 9$

(j) $\kappa = 10$ (k) $\kappa = 11$ (l) $\kappa = 12$

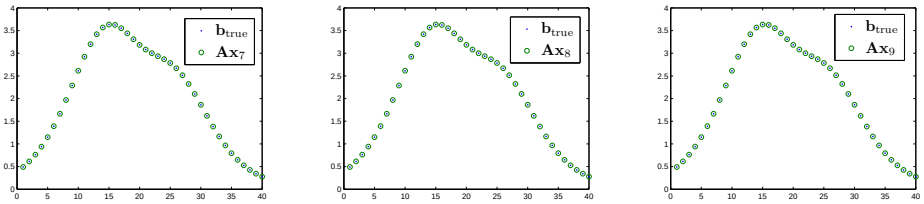
Figure 1.3: Truncated SVD solutions for different values of the parameter κ



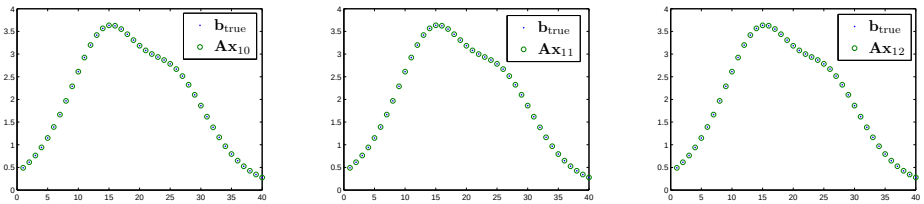
(a) $\kappa = 1$ (b) $\kappa = 2$ (c) $\kappa = 3$



(d) $\kappa = 4$ (e) $\kappa = 5$ (f) $\kappa = 6$



(g) $\kappa = 7$ (h) $\kappa = 8$ (i) $\kappa = 9$



(j) $\kappa = 10$ (k) $\kappa = 11$ (l) $\kappa = 12$

Figure 1.4: $\mathbf{Ax}_\kappa$ compared with $\mathbf{b}_{\text{true}}$

For further illustration of the discrete Picard condition, we use the discrete Picard plot shown in Figure 1.5 that displays the λ_j, $|\mathbf{u}_j^T\mathbf{b}|^2$ and $|\mathbf{u}_j^T\mathbf{b}|^2/\lambda_j$. It indicates that $|\mathbf{u}_j^T\mathbf{b}|$ decays faster than λ_j when $j \leq 9$. For $j \geq 10$, $|\mathbf{u}_j^T\mathbf{b}|$ decays much slower than λ_j, thereby producing large values of the $|\mathbf{u}_j^T\mathbf{b}|/\lambda_j$ that are dominated by the noise and violating the discrete Picard condition. Hence, choosing $\kappa = 9$ in (1.21) agrees with our visual perception from Figure 1.3.

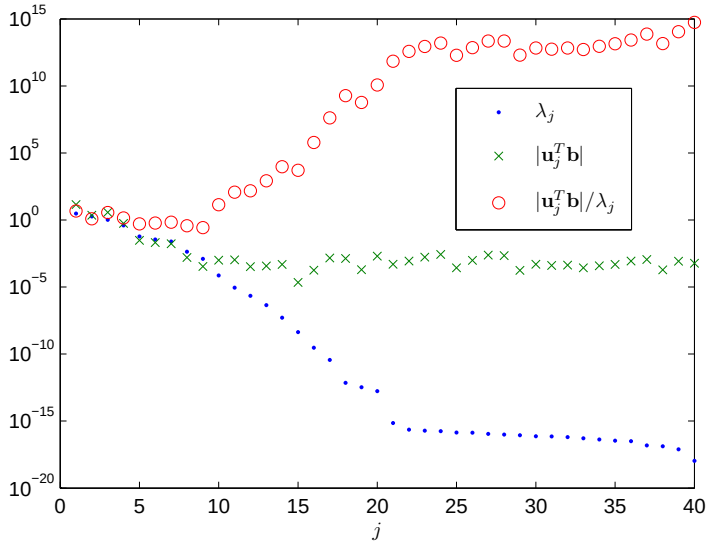


Figure 1.5: Discrete Picard plot of a test problem

1.3.1.2 Tikhonov Regularization

As we have seen, the truncated SVD method relies on computing the singular values and singular vectors of the matrix $\mathbf{A}$. The resulting computational task can be heavy or not feasible for large-scale problems. In contrast, the Tikhonov regularization method [39] does not require the calculation of the

SVD. Instead, we solve the problem

$$\min_{\mathbf{x}} \left\{ ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2^2 + \kappa^2 ||\mathbf{x}||_2^2 \right\} \tag{1.22}$$

for $\mathbf{x}$, where $\kappa > 0$ is a parameter that governs the weight of the regularization or penalty term in (1.22). That is,

$$\mathbf{x}_\kappa = \arg\min_{\mathbf{x}} \left\{ ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2^2 + \kappa^2 ||\mathbf{x}||_2^2 \right\}. \tag{1.23}$$

The first term of the right hand side of (1.22), i.e., the fidelity term, measures the fit of the solution to the noisy data and the second term controls the norm of the solution as a means of governing the noise distortion. There is a trade off between these two aspects of the criterion and we want to attain a suitable balance through adjusting the parameter κ.

To obtain a more explicit form for (1.23), we can write it as

$$\mathbf{x}_\kappa = \arg\min_{\mathbf{x}} \left\| \begin{pmatrix} \mathbf{A} \\ \kappa I \end{pmatrix} \mathbf{x} - \begin{pmatrix} \mathbf{b} \\ 0 \end{pmatrix} \right\|_2^2$$

which is now just an ordinary least-squares problem with the consequence that

$$\begin{aligned} \mathbf{x}_\kappa &= \left(\begin{pmatrix} \mathbf{A} \\ \kappa I \end{pmatrix}^T \begin{pmatrix} \mathbf{A} \\ \kappa I \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{A} \\ \kappa I \end{pmatrix}^T \begin{pmatrix} \mathbf{b} \\ 0 \end{pmatrix} \\ &= (\mathbf{A}^T\mathbf{A} + \kappa^2 I)^{-1}\mathbf{A}^T\mathbf{b}. \end{aligned} \tag{1.24}$$

Somewhat more generally, a Tikhonov criterion can be written as

$$||\mathbf{A}\mathbf{x} - \mathbf{b}||_{W_\mathbf{b}}^2 + ||\mathbf{L}(\mathbf{x} - \mathbf{x}_0)||_{W_L}^2. \tag{1.25}$$

Here the weighted norm is $||\mathbf{y}||_W^2 = \mathbf{y}^T W \mathbf{y}$, for a vector $\mathbf{y}$ and non-negative definite weighting matrix W, $\mathbf{L} \in \mathbb{R}^{p \times n}$ is, e.g., an approximate derivative

operator and $\mathbf{x}_0 \in \mathbb{R}^n$ represents a priori information about the unknown "signal" vector $\mathbf{x}$. Now the regularization parameter is the weighting matrix $W_{\mathbf{b}}$ (or $W_{\mathbf{L}}$) when $W_{\mathbf{L}}$ (or $W_{\mathbf{b}}$) is given. The least squares solution $\hat{\mathbf{x}}$ for (1.25) is seen to be

$$\hat{\mathbf{x}} = (\mathbf{A}^T W_{\mathbf{b}} \mathbf{A} + \mathbf{L}^T W_{\mathbf{L}} \mathbf{L})^{-1} \mathbf{A}^T W_{\mathbf{b}} (\mathbf{b} - \mathbf{A}\mathbf{x}_0) + \mathbf{x}_0 \tag{1.26}$$

under the invertibility condition for the matrix $(\mathbf{A}^T W_{\mathbf{b}} \mathbf{A} + \mathbf{L}^T W_{\mathbf{L}} \mathbf{L})$, i.e., $\mathrm{Ker}(\mathbf{A}) \cap \mathrm{Ker}(\mathbf{L}) = 0$.

1.3.2 Methods for Choosing the Regularization Parameter

A regularization method is not completely specified without a proper choice for its regularization parameter, e.g., the number of terms in the truncated SVD or the weight parameter for the Tikhonov criterion (1.22) as described in Section 1.3.1. There are various methods of accomplishing this; see [19], [17] and [41], for example. Here we emphasize three statistical methods: namely, unbiased predictive risk estimation (UPRE) ([41]), generalized cross validation (GCV) ([41] and [19]) and the χ^2 method ([28]). The χ^2 method is studied in more detail in Chapter 3.

1.3.2.1 UPRE

Substituting (1.18) into (1.17) produces the model

$$\mathbf{b} = \mathbf{A}\mathbf{x}_{\mathrm{true}} + \boldsymbol{\varepsilon}. \tag{1.27}$$

The predictive error is defined as

$$\mathrm{PE}_\kappa = \mathbf{A}\mathbf{x}_\kappa - \mathbf{A}\mathbf{x}_{\mathrm{true}}$$

and the predictive risk is the average of the mean squared norm of the predictive error; i.e.,

$$\frac{1}{m}\mathrm{E}\left(\|\mathrm{PE}_\kappa\|_2^2\right) = \frac{1}{m}\mathrm{E}\left(\|\mathbf{A}\mathbf{x}_\kappa - \mathbf{A}\mathbf{x}_{\mathrm{true}}\|_2^2\right).$$

Both the UPRE and GCV methods are based on estimators of the predictive risk.

The UPRE method was first developed for model selection in regression (e.g., [26]) and later used in regularization parameter estimation (see [41] and references therein). The idea is based on minimization of an unbiased estimator of the predictive risk criterion.

The case of interest is where $\mathbf{x}_\kappa$ is a linear estimator for a fixed choice of κ. In that event, we can write the regularized solution in the form $\mathbf{x}_\kappa = R_\kappa \mathbf{b}$ for some regularization matrix $R_\kappa \in \mathbb{R}^{n\times m}$. We then define the influence matrix as

$$\mathbf{A}_\kappa = \mathbf{A}R_\kappa.$$

Now we need to introduce the trace lemma wherein we will use tr to denote the matrix trace.

Lemma 1.3.1. *Let $\mathbf{h} \in \mathbb{R}^n$ be a deterministic vector, let $\boldsymbol{\varepsilon} \in \mathbb{R}^m$ be a random vector with mean 0 and variance-covariance $C_{\mathbf{b}}$ and let $B \in \mathbb{R}^{n\times m}$. Then,*

$$\begin{aligned}\mathrm{E}(\|\mathbf{h} + B\boldsymbol{\varepsilon}\|_2^2) &= \|\mathbf{h}\|_2^2 + \sum_{i=1}^{m}\sum_{j=1}^{m}(B^TB)_{ij}(C_{\mathbf{b}})_{ij} \\ &= \|\mathbf{h}\|_2^2 + \mathrm{tr}(BC_{\mathbf{b}}B^T).\end{aligned}$$

Proof. We have

$$\begin{aligned}
\mathrm{E}(\|\mathbf{h}+B\boldsymbol{\varepsilon}\|_2^2) &= \|\mathbf{h}\|_2^2 + 2\mathrm{E}(\mathbf{h}^T B\boldsymbol{\varepsilon}) + \mathrm{E}(\boldsymbol{\varepsilon}^T B^T B\boldsymbol{\varepsilon}) \\
&= \|\mathbf{h}\|_2^2 + 2\mathbf{h}^T B\mathrm{E}(\boldsymbol{\varepsilon}) + \sum_{i=1}^{m}\sum_{j=1}^{m}(B^T B)_{ij}\mathrm{E}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j) \\
&= \|\mathbf{h}\|_2^2 + \sum_{i=1}^{m}\sum_{j=1}^{m}(B^T B)_{ij}\mathrm{E}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j) \\
&= \|\mathbf{h}\|_2^2 + \mathrm{tr}(BC_{\mathbf{b}}B^T)
\end{aligned}$$

with $(B^T B)_{ij}$ the ijth element of $B^T B$. □

In model (1.27), we assume that $\mathrm{E}(\boldsymbol{\varepsilon}) = 0$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}) = C_{\mathbf{b}}$. Since PE_κ can be written as

$$\mathrm{PE}_\kappa = (\mathbf{A}_\kappa - I)\mathbf{A}\mathbf{x}_{\mathrm{true}} + \mathbf{A}_\kappa\boldsymbol{\varepsilon},$$

by Lemma 1.3.1, an explicit expression for the predictive risk is seen to be

$$\mathrm{E}\left(\frac{1}{m}\|\mathrm{PE}_\kappa\|_2^2\right) = \frac{1}{m}\|(\mathbf{A}_\kappa - I)\mathbf{A}\mathbf{x}_{\mathrm{true}}\|_2^2 + \frac{\mathrm{tr}(\mathbf{A}_\kappa C_{\mathbf{b}}\mathbf{A}_\kappa^T)}{m}. \tag{1.28}$$

Let $\mathbf{r}_\kappa = \mathbf{A}\mathbf{x}_\kappa - \mathbf{b}$ be the regularized residual vector

$$\mathbf{r}_\kappa = (\mathbf{A}_\kappa - I)\mathbf{A}\mathbf{x}_{\mathrm{true}} + (\mathbf{A}_\kappa - I)\boldsymbol{\varepsilon}.$$

Then, the average of the mean squared sum for the residuals is

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{m}\|\mathbf{r}_\kappa\|_2^2\right) &= \frac{1}{m}\|(\mathbf{A}_\kappa - I)\mathbf{A}\mathbf{x}_{\mathrm{true}}\|_2^2 \\
&+ \frac{\mathrm{tr}(\mathbf{A}_\kappa C_{\mathbf{b}}\mathbf{A}_\kappa^T)}{m} - \frac{2\mathrm{tr}(\mathbf{A}_\kappa C_{\mathbf{b}})}{m} + \frac{\mathrm{tr}(C_{\mathbf{b}})}{m}.
\end{aligned} \tag{1.29}$$

Combining (1.28) and (1.29) leads to

$$\mathrm{E}\left(\frac{1}{m}\|\mathrm{PE}_\kappa\|_2^2\right) = \mathrm{E}\left(\frac{1}{m}\|\mathbf{r}_\kappa\|_2^2\right) + \frac{2\mathrm{tr}(\mathbf{A}_\kappa C_{\mathbf{b}})}{m} - \frac{\mathrm{tr}(C_{\mathbf{b}})}{m}.$$

The UPRE is then defined as

$$\text{UPRE}(\kappa) = \frac{1}{m}\|\mathbf{r}_\kappa\|_2^2 + \frac{2\text{tr}(\mathbf{A}_\kappa C_\mathbf{b})}{m} - \frac{\text{tr}(C_\mathbf{b})}{m} \tag{1.30}$$

and satisfies $\text{E}\left(\text{UPRE}(\kappa)\right) = \text{E}\left(\frac{1}{m}\|\text{PE}_\kappa\|_2^2\right)$. The corresponding choice for κ is

$$\hat{\kappa}_{\text{UPRE}} = \arg\min_\kappa \text{UPRE}(\kappa).$$

In particular, if $C_\mathbf{b} = \sigma_\mathbf{b}^2 I$, we have the well-known trace lemma [41].

Lemma 1.3.2. *Let $\mathbf{h} \in \mathbb{R}^n$ be a deterministic vector, let $\boldsymbol{\varepsilon} \in \mathbb{R}^m$ be a random vector with mean 0 and variance-covariance matrix $\sigma^2 I$ and let $B \in \mathbb{R}^{n\times m}$. Then,*

$$\text{E}(\|\mathbf{h} + B\boldsymbol{\varepsilon}\|_2^2) = \|\mathbf{h}\|_2^2 + \sigma^2\text{tr}(B^T B).$$

Applying Lemma 1.3.2 leads to the result that

$$\text{E}\left(\frac{1}{m}\|\text{PE}_\kappa\|_2^2\right) = \text{E}\left(\frac{1}{m}\|\mathbf{r}_\kappa\|_2^2\right) + \frac{2\sigma_\mathbf{b}^2}{m}\text{tr}(\mathbf{A}_\kappa) - \sigma_\mathbf{b}^2.$$

Then, the UPRE becomes

$$\text{UPRE}(\kappa) = \frac{1}{m}\|\mathbf{r}_\kappa\|_2^2 + \frac{2\sigma_\mathbf{b}^2}{m}\text{tr}(\mathbf{A}_\kappa) - \sigma_\mathbf{b}^2 \tag{1.31}$$

and the corresponding choice for κ is

$$\hat{\kappa}_{\text{UPRE}} = \arg\min_\kappa \text{UPRE}(\kappa).$$

As a specific example, consider Tikhonov regularization. From (1.24), the regularization matrix is $R_\kappa = (\mathbf{A}^T\mathbf{A} + \kappa^2 I)^{-1}\mathbf{A}^T$ and the influence matrix is $\mathbf{A}_\kappa = \mathbf{A}(\mathbf{A}^T\mathbf{A} + \kappa^2 I)^{-1}\mathbf{A}^T$. An application of Corollary 1.2.1 to $\mathbf{A}$ produces

$$\begin{aligned}
\mathbf{A}_\kappa - I &= \mathbf{A}(\mathbf{A}^T\mathbf{A} + \kappa^2 I)^{-1}\mathbf{A}^T - I \\
&= \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T(\mathbf{V}\boldsymbol{\Lambda}^T\mathbf{U}^T\mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T + \kappa^2\mathbf{V}\mathbf{V}^T)^{-1}\mathbf{V}\boldsymbol{\Lambda}^T\mathbf{U} - \mathbf{U}\mathbf{U}^T \\
&= \mathbf{U}(\boldsymbol{\Lambda}(\boldsymbol{\Lambda}^T\boldsymbol{\Lambda} + \kappa^2 I_n)^{-1}\boldsymbol{\Lambda}^T - I_m)\mathbf{U}^T.
\end{aligned} \tag{1.32}$$

Then, we have

$$\mathrm{tr}(\mathbf{A}_\kappa) = \sum_{i=1}^{n} \frac{\lambda_i^2}{\lambda_i^2 + \kappa^2} \tag{1.33}$$

and

$$(\mathbf{A}_\kappa - I)\mathbf{b} = \sum_{i=1}^{m} (\mathbf{u}_i^T \mathbf{b}) \frac{-\kappa^2}{\lambda_i^2 + \kappa^2} \mathbf{u}_i. \tag{1.34}$$

Thus, by (1.31), (1.33) and (1.34), we obtain

$$\mathrm{UPRE}(\kappa) = \frac{1}{m} \sum_{i=1}^{m} |\mathbf{u}_i^T \mathbf{b}|^2 \left(\frac{\kappa^2}{\lambda_i^2 + \kappa^2} \right)^2 + \lambda_{\mathbf{b}}^2 \left(\frac{2}{m} \sum_{i=1}^{n} \frac{\lambda_i^2}{\lambda_i^2 + \kappa^2} - 1 \right).$$

From (1.30) and (1.31) we see that the UPRE method relies on the noise variance $C_{\mathbf{b}}$ or $\sigma_{\mathbf{b}}^2$ which will generally not be known exactly in practice. Thompson et al. [38] provide several ways of estimating $\sigma_{\mathbf{b}}^2$ that can be used in the UPRE function. The basic structure is to consider estimators of the form of $\mathbf{b}^T K \mathbf{b}/\mathrm{tr}(K)$, where K is chosen to be some symmetric, nonnegative-definite matrix that approximately annihilates $\mathbf{A}\mathbf{x}_{\mathrm{true}}$.

1.3.2.2 GCV

The GCV method furnishes an alternative to UPRE which does not need prior information on (or estimation of) $\sigma_{\mathbf{b}}^2$. The GCV functional is defined to be

$$\mathrm{GCV}(\kappa) = \frac{\frac{1}{m}\|\mathbf{r}_\kappa\|_2^2}{\left[\frac{1}{m}\mathrm{tr}(I - \mathbf{A}_\kappa)\right]^2}. \tag{1.35}$$

The corresponding choice for κ is

$$\hat{\kappa}_{\mathrm{GCV}} = \arg\min_{\kappa} \mathrm{GCV}(\kappa).$$

One justification for this estimator is the so-called GCV Theorem ([15], [12]) which can be stated as follows for the case of a symmetric $\mathbf{A}_\kappa$.

Theorem 1.3.1. *Let* $\tau_j(\kappa) = (1/m)\mathrm{tr}(\mathbf{A}_\kappa^j)$, $j = 1, 2$, *and assume that* $\tau_1(\kappa) < 1$. *Then,*

$$\frac{\left|\mathrm{E}(\mathrm{GCV}(\kappa)) - (\frac{1}{m}\mathrm{E}(\|\mathrm{PE}_\kappa\|_2^2) + \sigma_{\boldsymbol{b}}^2)\right|}{\frac{1}{m}\mathrm{E}(\|\mathrm{PE}_\kappa\|_2^2)} \leq g(\kappa),$$

where $g(\kappa) = \left[2\tau_1(\kappa) + \tau_1(\kappa)^2/\tau_2(\kappa)\right]/(1 - \tau_1(\kappa))^2$.

This theorem implies that when $g(\kappa)$ is small, the distance between $\mathrm{E}(\mathrm{GCV}(\kappa))$ and $(1/m)\mathrm{E}(\|\mathrm{PE}_\kappa\|_2^2) + \sigma_{\mathbf{b}}^2$ is also small relative to the predictive risk

$$(1/m)\mathrm{E}(\|\mathrm{PE}_\kappa\|_2^2).$$

In this case, $\mathrm{GCV}(\kappa)$ can be roughly regarded as an unbiased estimator of

$$(1/m)\mathrm{E}(\|\mathrm{PE}_\kappa\|_2^2) + \sigma_{\mathbf{b}}^2.$$

1.3.2.3 χ^2 Method

Mead and Renaut [28] have proposed a method for selection of regularization parameters that will be the focus of Chapter 3. For an ill-posed linear system (1.16), they consider the generalized Tikhonov regularization criterion (1.25) and assume that $m \geq n \geq p$, $\mathrm{Ker}(\mathbf{A}) \cap \mathrm{Ker}(\mathbf{L}) = 0$, $\mathbf{L}$ has rank p and the augmented matrix $(\mathbf{A}; \mathbf{L})$ has rank n.

The solution is obtained by solving the weighted least-squares problem

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} J(\mathbf{x}),$$

where $J(\mathbf{x})$ can be equivalently written as

$$J(\mathbf{x}) = (\mathbf{b} - \mathbf{A}\mathbf{x})^T W_{\mathbf{b}}(\mathbf{b} - A\mathbf{x}) + (\mathbf{x} - \mathbf{x}_0)^T W_{\mathbf{x}}(\mathbf{x} - \mathbf{x}_0), \tag{1.36}$$

with

$$W_{\mathbf{x}} = \mathbf{L}^T W_{\mathbf{L}} \mathbf{L}.$$

Mead [27] considered the case $\mathbf{L} = I$ and proposed the χ^2 method for estimating $W_{\mathbf{b}}$ given $W_{\mathbf{x}}$ or $W_{\mathbf{x}}$ given $W_{\mathbf{b}}$. This idea was extended by Mead and Renaut [28] for the general case of arbitrary $\mathbf{L}$, using the GSVD. The technique is based on the distribution of the minimum value of J. We use $\mathbf{x} \sim N_n(\mathbf{x_0}, C_{\mathbf{x}})$ to denote that $\mathbf{x}$ follows an n-variate normal distribution with mean $\mathbf{x}_0$ and variance-covariance matrix $C_{\mathbf{x}}$ and use $\mathbf{b}|\mathbf{x}$ to denote $\mathbf{b}$ conditional on $\mathbf{x}$.

Theorem 1.3.2. *[28] Let $J(\mathbf{x})$ be defined as in (1.36). Assume that $\mathbf{x} \sim N_n(\mathbf{x_0}, C_{\mathbf{x}})$ where $C_{\mathbf{x}} = W_{\mathbf{x}}^-$ with $W_{\mathbf{x}}$ a symmetric nonnegative definite weighting matrix and that $\mathbf{b}|\mathbf{x} \sim N_m(\mathbf{Ax}, W_{\mathbf{b}}^{-1})$ for a symmetric positive definite weighting matrix $W_{\mathbf{b}}$. Then, the minimum value of J is a random variable that has a χ^2 distribution with $m - n + p$ degrees of freedom.*

The theorem states that, for normally distributed data and known $W_{\mathbf{x}}$, $W_{\mathbf{b}}$, the minimized values of the criterion $J(\hat{\mathbf{x}})$ follows a χ^2 distribution. In instances where one of the two is unknown, $J(\hat{\mathbf{x}})$ can be used like a test statistic and inverted to construct a confidence region for the unknown matrix. As a byproduct one obtains a suitably regularized final choice for $\hat{\mathbf{x}}$ corresponding to a matrix that falls inside the confidence region.

If $W_{\mathbf{b}}$ and W_L are known, Theorem 1.3.2 has the consequence that for large $(m-n+p)$ and $z_{\alpha/2}$ the $100(1-\alpha/2)$ percentile of the standard normal distribution, the probability that

$$m - n + p - \sqrt{2(m-n+p)}z_{\alpha/2} < J(\hat{\mathbf{x}}) < m - n + p + \sqrt{2(m-n+p)}z_{\alpha/2} \quad (1.37)$$

will be approximately $1 - \alpha$. Mead [27] proposed choosing $W_{\mathbf{b}}$ (or W_L when $L = I$) so that these bounds are realized. For instance, suppose we want to

estimate W_L which is assumed to be of the form $W_L^{-1} = \sigma_{\mathbf{x}}^2 I$ given $W_{\mathbf{b}}$. Then, $\sigma_{\mathbf{x}}$ can be found by a single-variable root finding Newton method which solves

$$F(\sigma_{\mathbf{x}}) = J(\hat{\mathbf{x}}) - (m - n + p) = 0, \tag{1.38}$$

subject to the tolerance $|F(\sigma_{\mathbf{x}})| < tol$ where $tol = \sqrt{2(m-n+p)}z_{\alpha/2}$. Experimental results in [27] indicate that this approach can be more effective than GCV, for example.

1.3.3 An Example

Here, we apply the UPRE, GCV, and χ^2 methods for Tikhonov regularization to the same Shaw test problem that was described in Section 1.3.1.1.

Figures 1.6 and 1.7 show the UPRE(κ) and GCV(κ) functions, respectively. The Tikhonov parameter estimate minimizing the UPRE is obtained at $\hat{\kappa}_{\text{UPRE}} = 0.0033$ while the parameter estimate minimizing GCV is obtained at $\hat{\kappa}_{\text{GCV}} = 0.0036$. In terms of the χ^2 method, in this instance, $m = n = p = 40$ for $L = I$ and $W_{\mathbf{b}} = 10^6 I$. So, we are looking for $\sigma_{\mathbf{x}}$ such that (1.38) is satisfied with $m - n + p = 40$. We pick $\alpha = 0.95$ such that $tol = 0.5609$. Figure 1.8 shows $F(\sigma_{\mathbf{x}})$ and that the Newton method converges in 5 iterations at $\sigma_{\mathbf{x}} = 2.0387$ and $F(\sigma_{\mathbf{x}}) = 0.4900$. Since in this case, (1.36) can then be written as

$$J(\mathbf{x}) = 10^6 \left(||\mathbf{b} - \mathbf{A}\mathbf{x}||_2^2 + \frac{10^{-6}}{\sigma_{\mathbf{x}}^2} \|\mathbf{x} - \mathbf{x}_0\|_2^2 \right),$$

the corresponding Tikhonov parameter in (1.22) is estimated at

$$\hat{\kappa} = \sqrt{\frac{10^{-6}}{\hat{\sigma_{\mathbf{x}}}^2}} = \frac{10^{-3}}{2.0387} = 0.00049.$$

Figure 1.9 illustrates the approximate solutions using these three methods compared with the true solution.

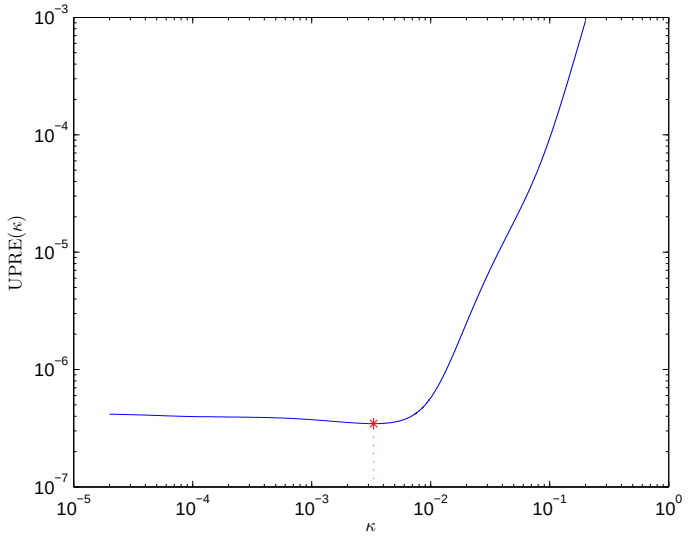


Figure 1.6: UPRE function: $\hat{\kappa}_{\mathrm{UPRE}} = 0.0033$. (Logarithmic scales are used for the x and y axes.)

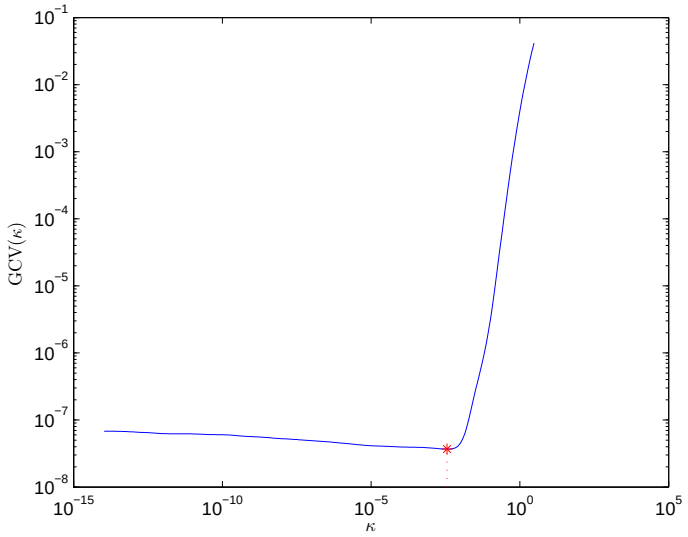


Figure 1.7: GCV function: $\hat{\kappa}_{\mathrm{GCV}} = 0.0036$. (Logarithmic scales are used for the x and y axes.)

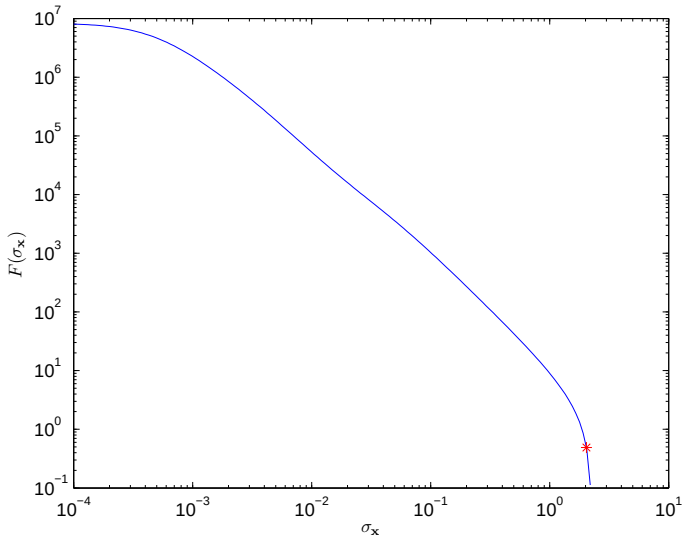


Figure 1.8: χ^2 method: $\hat{\sigma}_{\mathbf{x}} = 2.0387$. (Logarithmic scales are used for the x and y axes.)

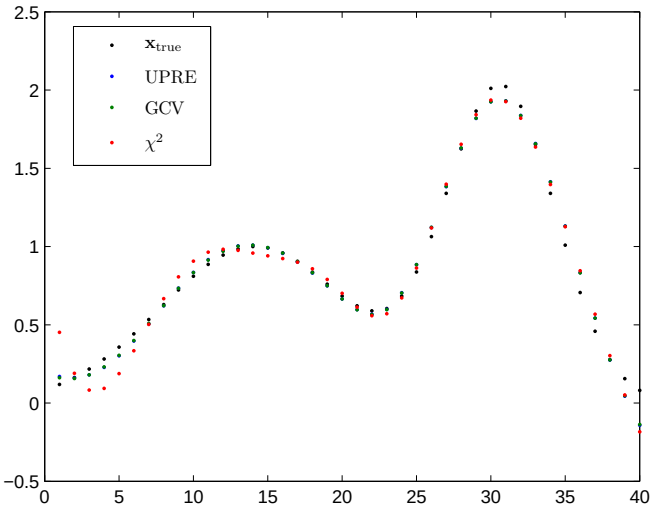


Figure 1.9: Solutions obtained with the UPRE, GCV and χ^2 methods.

1.4 Functional Data Analysis

Functional data can be viewed as collections of sample paths from a stochastic process that take values in a Hilbert function space. Unlike classic statistical data, each observation is a function rather than a scalar or vector value. Data of this type are now relatively common. Many physical processes evolve smoothly over time and modern computing technology allows for storage of an extensive digitized record of the outcomes.

Functional data analysis (FDA) is concerned with the development of extensions of classical multivariate data analysis methods such as principal components analysis and canonical correlation analysis (CCA) to the functional domain. Our emphasis is on functional CCA which concerns investigating the association between continuous time stochastic processes. A review of classic finite dimensional multivariate CCA is given first before we discuss the functional case.

1.4.1 Finite Dimensional CCA

The finite dimensional version of CCA in multivariate analysis stems from Hotelling [21]. His original CCA development was aimed toward summarizing the relationship between two sets of variables. The basic idea is to find linear combinations of each set of variables such that these new variables, called canonical variables, provide a simple representation of the multidimensional correlation structure. The first pair of canonical variables maximizes the squared correlation between a linear combination of the first set of variables and that of the second set of variables. The second pair of canonical variables maximizes the squared correlation between linear combinations of

each set but are chosen from those that are uncorrelated with the first pair of variables. This same process is then repeated to construct subsequent pairs of canonical variables.

To be more explicit, let $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$ be random vectors with zero means and variance-covariance matrices

$$\mathrm{Var}(X) = \Sigma_{XX}, \mathrm{Var}(Y) = \Sigma_{YY}, \mathrm{Cov}(X,Y) = \Sigma_{XY} = \Sigma_{YX}^T.$$

Then, the first canonical correlation ρ_1 and associated weight vectors a_1 and b_1 are defined as

$$\rho_1^2 = \sup_{a \in \mathbb{R}^p, b \in \mathbb{R}^q} \mathrm{Cov}^2(a^T X, b^T Y) = \mathrm{Cov}^2(a_1^T X, b_1^T Y),$$

where a and b are subject to

$$\mathrm{Var}(a^T X) = \mathrm{Var}(b^T Y) = 1. \tag{1.39}$$

For $i > 1$, the ith canonical correlation ρ_i and the associated weight vectors a_i and b_i are defined as

$$\rho_i^2 = \sup_{a \in \mathbb{R}^p, b \in \mathbb{R}^q} \mathrm{Cov}^2(a^T X, b^T Y) = \mathrm{Cov}^2(a_i^T X, b_i^T Y),$$

where a and b are subject to (1.39) and

$$\begin{aligned} & \mathrm{Cov}(a^T X, a_j^T X) = \mathrm{Cov}(b^T Y, b_j^T Y) \\ = \; & \mathrm{Cov}(a^T X, b_j^T Y) = \mathrm{Cov}(a_j^T X, b^T Y) = 0, j < i. \end{aligned}$$

Kshirsagar [22] connected CCA with the SVD. He showed that the canonical correlations ρ_i are the singular values of $\Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}$ and the canonical variables are obtained from its singular vectors.

1.4.2 Functional CCA

Several approaches have been developed for functional CCA. The work from [7] relies on the angles of two subspaces, but it only applies to finite dimensional covariance operators, and in that sense, is covered by Hotelling's original treatise.

He et al. [20] consider functional CCA for second order processes in the sense of Section 4.1. Specifically, let $\{X_i(t) : t \in [0,1]\}$, $i = 1, 2$, be two second order processes taking values in $L^2[0,1]$ with zero means, covariance kernels $K_{11}(t,s) = \mathrm{E}(X_1(t)X_1(s))$, $K_{22}(t,s) = \mathrm{E}(X_2(t)X_2(s))$ and cross covariance function $K_{12}(t,s) = \mathrm{E}(X_1(t)X_2(s))$ with $K_{ij} \in L^2([0,1] \times [0,1])$ for $i, j = 1, 2$. He et al. [20] define the integral operator R_{11} as

$$R_{11}u(t) = \int_0^1 K_{11}(t,s)u(s)ds$$

for $u \in L^2[0,1]$ and define analogous operators R_{22} and R_{12}. Their version of canonical correlation involves the singular values of $R_{11}^{-1/2}R_{12}R_{22}^{-1/2}$. Since R_{11} and R_{22} are compact, they impose range restrictions to ensure that the inverses are defined. In general, when these conditions do not hold, this may produce solutions that are suboptimal.

A typical functional CCA formulation focuses on the study of the induced random variables $\langle X, f\rangle_{\mathcal{H}}$, $f \in \mathcal{H}$ for some separable Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$. Roughly speaking, given two second order processes X_1, X_2 taking values in Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$, functional CCA looks for $f_i^* \in \mathcal{H}_i$ such that

$$\sup_{f_i \in \mathcal{H}_i} |\mathrm{Corr}(\langle X_1, f_1\rangle_{\mathcal{H}_1}, \langle X_2, f_2\rangle_{\mathcal{H}_2})| = |\mathrm{Corr}(\langle X_1, f_1^*\rangle_{\mathcal{H}_1}, \langle X_2, f_2^*\rangle_{\mathcal{H}_2})|.$$

Using an extended version of this framework, Kupresanin et al. [23] obtain canonical variables and correlations via spectral decomposition of an operator on the reproducing kernel Hilbert space (RKHS) generated by the process' covariance kernels. Their work will be discussed in Chapter 4.

1.4.3 PCCA

Roy [33] extended the canonical correlation idea from two sets of random variables to three sets in the following way. Let X, Y and Z be three random vectors. Roy defined the partial canonical correlation of Y and Z to be the ordinary canonical correlation of $\tilde{Y}$ and $\tilde{Z}$, where

$$\tilde{Y} = Y - P_X Y \text{ and } \tilde{Z} = Z - P_X Z$$

with P_X denoting the projection on the linear space spanned by X. In other words, for partial canonical correlation analysis (PCCA) we are interested in the relationship between Y and Z after removing X's influence on both of them.

Dauxois and Nkiet [7] have developed an abstract version of PCCA that generalizes Roy's approach. However, it is applicable to only finite dimensional covariance operators and, in that sense, fails to incorporate the instance of functional data.

1.5 Outline of the Thesis

The remainder of the dissertation is laid out as follows. In the next chapter, we develop a new version of a GSVE. Chapter 3 gives an analysis of the χ^2 method for choosing regularization parameters for nonnormal data. Both large sample and finite dimensional developments suggest the method may be sensitive to departures from the assumption of normally distributed data.

Chapter 4 gives a new methodology for analyzing the relationship between two or more stochastic processes in a functional data setting. The results of Eubank and Hsing [13] are re-derived using this framework. Then, a general notion of PCCA is derived that is valid in both the finite dimensional and functional case. Chapter 5 summarizes our findings and discusses plans for future research.

CHAPTER 2

GSVE FOR COMPACT OPERATORS

2.1 Introduction

One question of interest is whether the matrix GSVD described in Chapter 1 has an extension similar to Theorem 1.2.1 that holds for infinite dimensions. A direct extension of Theorem 1.2.1 does not appear to be feasible for reasons described subsequently. However, at least one type of generalized singular value expansion (GSVE) can be derived as demonstrated in the next section.

2.2 A GSVE

Here we present and prove a form of GSVE that can be applied to two compact operators. For this purpose we need the Hilbert space l^2 of square-summable sequences. If $l = (l_1, l_2, \ldots)$ and $l' = (l'_1, l'_2, \ldots)$ are in l^2, their inner product is $\langle l, l' \rangle_{l^2} = \sum_{j=1}^{\infty} l_j \overline{l'_j}$ with $\overline{l'_j}$ the complex conjugate of l'_j.

Theorem 2.2.1. *Suppose that there are two compact operators A and B on a separable Hilbert space $\mathcal{H}$, such that $A : \mathcal{H} \to \mathcal{H}_A$ and $B : \mathcal{H} \to \mathcal{H}_B$ for separable Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$.*

(i) Let $\{d_{jA}, f_{jA}, g_{jA}\}$ and $\{d_{jB}, f_{jB}, g_{jB}\}$ be the singular systems of A and B, respectively, and let $\{d_j, f_j, g_j\}$ be the singular system of

$$C = \begin{bmatrix} A \\ B \end{bmatrix}.$$

(ii) Define $D_A : l^2 \to l^2$ by

$$D_A l = (d_{1A} l_1, d_{2A} l_2, \ldots)$$

for $l = (l_1, l_2, \ldots)$, *and* $D_B : l^2 \to l^2$ *by*

$$D_B l = (d_{1B} l_1, d_{2B} l_2, \ldots).$$

Then, there exist unitary operators Q, Q_{AB} *in* $B(l^2)$ *and unitary operators* G, F_A *and* F_B *such that*

$$F_A^* A G = D_A Q^*,$$

$$F_B^* B G = D_B Q_{AB}^* Q^*,$$

where we have the following definitions:

(i) $G^* : \mathcal{H} \to l^2$ *is defined by*

$$G^* g = (\langle g, g_1 \rangle_{\mathcal{H}}, \langle g, g_2 \rangle_{\mathcal{H}}, \ldots) \tag{2.1}$$

for $g \in \mathcal{H}$,

(ii) $Q^* : l^2 \to l^2$ *is defined by*

$$Q^* l = \left(\sum_{i=1}^{\infty} \langle g_i, g_{1A} \rangle_{\mathcal{H}} l_i, \sum_{i=1}^{\infty} \langle g_i, g_{2A} \rangle_{\mathcal{H}} l_i, \ldots \right), \tag{2.2}$$

(iii) $F_A : l^2 \to \mathcal{H}_A$ *is defined by*

$$F_A l = \sum_j l_j f_{jA},$$

(iii) $F_B : l^2 \to \mathcal{H}_B$ *is defined by*

$$F_B l = \sum_{j=1}^{\infty} l_j f_{jB},$$

(iv) $Q^*_{AB} : l^2 \to l^2$ *is defined by*

$$Q^*_{AB} l = \left(\sum_{k=1}^{\infty} \langle g_{kA}, g_{1B} \rangle_{\mathcal{H}} l_k, \sum_{k=1}^{\infty} \langle g_{kA}, g_{2B} \rangle_{\mathcal{H}} l_k, \ldots \right). \tag{2.3}$$

Proof. Write the SVE of C as

$$C = FDG^*$$

with $G^* : \mathcal{H} \to l^2$, $D : l^2 \to l^2$ and $F : l^2 \to \mathcal{H}_A \otimes \mathcal{H}_B$ defined by (2.1), $Dl = (d_1 l_1, d_2 l_2, \ldots)$ and $Fl = \sum_{j=1}^{\infty} l_j f_j$, respectively. We begin by showing that G is unitary.

We have

$$\begin{aligned} \langle l, G^* g \rangle_{l^2} &= \sum_{j=1}^{\infty} l_j \overline{\langle g, g_j \rangle}_{\mathcal{H}} = \sum_{j=1}^{\infty} l_j \langle g_j, g \rangle_{\mathcal{H}} \\ &= \left\langle \sum_{j=1}^{\infty} l_j g_j, g \right\rangle_{\mathcal{H}}. \end{aligned}$$

So,

$$Gl = \sum_{j=1}^{\infty} l_j g_j,$$

$$G^* G l = G^* \left(\sum_{j=1}^{\infty} l_j g_j \right) = (l_1, l_2, \ldots) = l$$

and

$$GG^* g = G\left(\langle g, g_1 \rangle_{\mathcal{H}}, \ldots \right) = \sum_{j=1}^{\infty} \langle g, g_j \rangle_{\mathcal{H}} g_j = g$$

as required for a unitary operator.

Now, the SVE for A gives

$$A = F_A D_A Q^*_A.$$

Without loss of generality, we assume that $\{g_j\}$ provides a complete orthonormal system (CONS) $\{g_j\}$ for $\mathcal{H}$. Then,

$$g_{jA} = \sum_{i=1}^{\infty} \langle g_{jA}, g_i \rangle_{\mathcal{H}} g_i$$

and, hence,

$$A = \sum_{j=1}^{\infty} d_{jA} f_{jA} \otimes \left(\sum_{i=1}^{\infty} \langle g_{jA}, g_i \rangle_{\mathcal{H}} g_i \right) = F_A D_A Q^* G^*$$

with Q^* defined in (2.2).

We claim that Q is unitary. To see this, write

$$\begin{aligned} \langle l_2, Q^* l_1 \rangle_{l^2} &= \sum_{j=1}^{\infty} l_{2j} \left\{ \sum_{i=1}^{\infty} \overline{\langle g_i, g_{jA} \rangle_{\mathcal{H}} l_{1i}} \right\} \\ &= \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} l_{2j} \langle g_{jA}, g_i \rangle_{\mathcal{H}} \right) \overline{l_{1i}}, \end{aligned}$$

from which it follows that

$$Ql = \left(\sum_{j=1}^{\infty} l_j \langle g_{jA}, g_1 \rangle_{\mathcal{H}}, \sum_{j=1}^{\infty} l_j \langle g_{jA}, g_2 \rangle_{\mathcal{H}}, \dots \right)$$

and

$$\begin{aligned} Q^* Ql &= Q^* \left(\sum_{j=1}^{\infty} l_j \langle g_{jA}, g_1 \rangle_{\mathcal{H}}, \dots \right) \\ &= \left(\sum_{i=1}^{\infty} \langle g_i, g_{1A} \rangle_{\mathcal{H}} (\sum_{j=1}^{\infty} l_j \langle g_{jA}, g_i \rangle_{\mathcal{H}}), \dots \right). \end{aligned}$$

But, for each k,

$$\begin{aligned} \sum_{i=1}^{\infty} \langle g_i, g_{kA} \rangle_{\mathcal{H}} \sum_{j=1}^{\infty} l_j \langle g_{jA}, g_i \rangle_{\mathcal{H}} &= \sum_{j=1}^{\infty} l_j \sum_{i=1}^{\infty} \langle g_i, g_{kA} \rangle_{\mathcal{H}} \langle g_{jA}, g_i \rangle_{\mathcal{H}} \\ &= \sum_{j=1}^{\infty} l_j \sum_{i=1}^{\infty} \langle g_i, g_{kA} \rangle \overline{\langle g_i, g_{jA} \rangle}_{\mathcal{H}} \\ &= l_k, \end{aligned}$$

because

$$\begin{aligned}\sum_{i=1}^{\infty}\langle g_i, g_{kA}\rangle_{\mathcal{H}}\overline{\langle g_i, g_{jA}\rangle}_{\mathcal{H}} &= \sum_{i=1}^{\infty}\langle g_i, g_{jA}\rangle_{\mathcal{H}}\overline{\langle g_i, g_{kA}\rangle}_{\mathcal{H}} \\ &= \langle g_{kA}, g_{jA}\rangle_{\mathcal{H}} = \delta_{kj}\end{aligned}$$

by Parseval's relation. Similarly,

$$\begin{aligned}QQ^*l &= Q\left(\sum_{i=1}^{\infty}\langle g_i, g_{1A}\rangle_{\mathcal{H}} l_i, \dots\right) \\ &= \left(\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\langle g_i, g_{jA} >_{\mathcal{H}} l_i < g_{jA}, g_1\rangle_{\mathcal{H}}, \dots\right) \\ &= \left(\sum_{j=1}^{\infty}\langle g_{jA}, g_1\rangle_{\mathcal{H}}\sum_{i=1}^{\infty} l_i\langle g_i, g_{jA}\rangle_{\mathcal{H}}, \dots\right) \\ &= l.\end{aligned}$$

At this point we have

$$A = F_A D_A Q^* G^*$$

and need to show that F_A is unitary. But,

$$\begin{aligned}\langle F_A l, f\rangle_{\mathcal{H}_A} &= \sum_{j=1}^{\infty} l_j\langle f_{jA}, f\rangle_{\mathcal{H}_A} \\ &= \sum_{j=1}^{\infty} l_j\overline{\langle f, f_{jA}\rangle}_{\mathcal{H}_A} = \langle l, F_A^* f\rangle_{l^2}\end{aligned}$$

with

$$F_A^* f = \left(\langle f, f_{1A}\rangle_{\mathcal{H}_A}, \dots\right).$$

Thus,

$$F_A F_A^* f = \sum_{j=1}^{\infty}\langle f, f_{jA}\rangle_{\mathcal{H}_A} f_{jA} = f$$

and

$$\begin{aligned} F_A^* F_A l &= F_A^* \sum_{j=1}^{\infty} l_j f_{jA} = \left(\sum_{j=1}^{\infty} l_j \langle f_{jA}, f_{1A} \rangle_{\mathcal{H}_A}, \dots \right) \\ &= (l_1, l_2, \dots) = l. \end{aligned}$$

Finally, the SVE of B gives us

$$B = \sum_{j=1}^{\infty} d_{jB} f_{jB} \otimes \left(\sum_{i=1}^{\infty} \langle g_{jB}, g_i \rangle_{\mathcal{H}} g_i \right).$$

However,

$$\sum_{i=1}^{\infty} \langle g_{jB}, g_i \rangle_{\mathcal{H}} g_i = \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} \langle g_{jB}, g_{kA} \rangle_{\mathcal{H}} \langle g_{kA}, g_i \rangle_{\mathcal{H}} g_i.$$

Therefore, for any $g \in \mathcal{H}$,

$$\begin{aligned} Bg &= \sum_{j=1}^{\infty} d_{jB} \left\langle g, \sum_{i=1}^{\infty} \langle g_{jB}, g_i \rangle_{\mathcal{H}} g_i \right\rangle_{\mathcal{H}} f_{jB} \\ &= \sum_{j=1}^{\infty} d_{jB} f_{jB} \left\langle g, \sum_{k=1}^{\infty} \langle g_{jB}, g_{kA} \rangle_{\mathcal{H}} \sum_{i=1}^{\infty} \langle g_{kA}, g_i \rangle_{\mathcal{H}} g_i \right\rangle_{\mathcal{H}} \\ &= \sum_{j=1}^{\infty} d_{jB} f_{jB} \sum_{k=1}^{\infty} \langle g_{kA}, g_{jB} \rangle_{\mathcal{H}} \sum_{i=1}^{\infty} \langle g_i, g_{kA} \rangle_{\mathcal{H}} \langle g, g_i \rangle_{\mathcal{H}} \\ &= F_B D_B Q_{AB}^* Q^* G^*. \end{aligned}$$

And we can show that F_B and Q_{AB} are unitary in the same way as for F_A and Q. □

Theorem 2.2.1 provides a relationship between the singular values of the different operators that can be summarized as follows.

Corollary 2.2.1. *Under the assumptions of Theorem 2.2.1, the singular values of C, A and B satisfy*

$$D^2 = Q \left[D_A^2 + Q_{AB} D_B^2 Q_{AB}^* \right] Q^*.$$

Proof. Since

$$\begin{bmatrix} A \\ B \end{bmatrix} = FDG^* = \begin{bmatrix} F_A D_A Q^* G^* \\ F_B D_B Q_{AB}^* Q^* G^* \end{bmatrix},$$

it follows that

$$\begin{aligned} A^*A + B^*B &= GD^2G^* \\ &= GQD_A^2Q^*G^* + GQQ_{AB}D_B^2Q_{AB}^*Q^*G^*. \end{aligned}$$

□

Note that the form of our GSVE for compact operators differs from the usual matrix GSVD and this fact is reflected in Corollary 2.2.1 as compared to identities $(1.12) - (1.13)$. The problem arises when one attempts to directly extend the proof of Theorem 1.2.3. A key step in that argument was partitioning the matrix operator $\mathbf{U}$ that plays the role of F here and arose from the ordinary SVD of $\mathbf{C}$. Since the resulting sub-operators are finite dimensional, they will also admit SVDs whose components then appear in subsequent constructions.

While it is certainly possible to partition F in a similar manner to the Paige and Saunders [30] proof, the fact that F is a projection operator ensures that neither sub-operator can be compact. This forces the proof to go in a different direction.

2.3 Applications

The objective of this chapter was to develop a version of a GSVE. In that sense, Theorem 2.2.1 fulfills our stated goal. In the finite dimensional case, the GSVD has numerous applications. The utility of our GSVE is, on the other hand, much less clear. We touch on that topic briefly in this section.

A simple illustration of the information provided by our GSVE is the following well-known result.

Corollary 2.3.1. *Under the assumptions of Theorem 2.2.1,* $d_1^2 \leq d_{1A}^2 + d_{1B}^2$.

Proof. The inequality is a consequence of Corollary 2.2.1 since

$$||D^2||_{B(l^2)} \leq ||QD_A^2Q^*||_{B(l^2)} + ||QQ_{AB}D_B^2Q_{AB}^*Q^*||_{B(l^2)}.$$

□

In Van Loan's [40] derivation of the GSVD, he uses it to solve the generalized eigenvalue problem of finding ζ^2 such that $\det(\mathbf{A}^T\mathbf{A} - \zeta^2\mathbf{B}^T\mathbf{B}) = 0$ for real matrices $\mathbf{A}$ and $\mathbf{B}$. A similar development is possible with this GSVE. Suppose we wish to find ζ^2 and a $g \in \mathcal{H}$ such that

$$A^*Ag = \zeta^2 B^*Bg.$$

From Theorem 2.2.1, this becomes

$$GQD_A^2Q^*G^*g = \zeta^2 GQQ_{AB}D_B^2Q_{AB}^*Q^*G^*g.$$

Now let $\tilde{g} = Q^*G^*g$ to obtain

$$D_A^2\tilde{g} = \zeta^2 Q_{AB}D_B^2Q_{AB}^*\tilde{g}. \tag{2.4}$$

Since

$$\begin{aligned} \tilde{g} &= Q^*(\langle g, g_1\rangle_{\mathcal{H}}, \langle g, g_2\rangle_{\mathcal{H}}, \ldots) \\ &= (\langle g, g_{1A}\rangle_{\mathcal{H}}, \langle g, g_{2A}\rangle_{\mathcal{H}}, \ldots), \end{aligned}$$

take $g = g_{iB}$ and let

$$l_i = (\langle g_{iB}, g_{1A}\rangle_{\mathcal{H}}, \langle g_{iB}, g_{2A}\rangle_{\mathcal{H}}, \ldots). \tag{2.5}$$

Then, $l_i \in l^2$ satisfies $Q^*_{AB} l_i = e_i$ with $e_i \in l^2$ consisting of all zeros except for a one as its ith element. Thus, let $\tilde{g} = l_i$ and (2.4) becomes

$$Q^*_{AB} D^2_A l_i = \zeta_i^2 D^2_B e_i.$$

The last relation provides a characterization of the generalized eigenvalues that we formally state below.

Corollary 2.3.2. *For l_i defined as in (2.5), $\zeta_i^2 = \frac{\langle D^2_A l_i, l_i \rangle_{l^2}}{d^2_{iB}}, i = 1, 2, \ldots.$*

CHAPTER 3

THE GSVD AND THE χ^2 METHOD

3.1 Introduction

We begin this chapter by taking another look at regularization parameter estimation using the χ^2 method discussed in Section 1.3.2.3. By revisiting the method, we seek to analyze its relevance for data with nonnormal noise distributions. Recall from Section 1.3.2.3 that the setting is that $\mathbf{b}|\mathbf{x} \sim N_m(\mathbf{A}\mathbf{x}_0, W_{\mathbf{b}}^{-1})$ and $\mathbf{x} \sim N_n(\mathbf{x}_0, W_{\mathbf{x}}^{-})$. We observe $\mathbf{b}$ and wish to use that information to estimate $\mathbf{x}$.

It will be convenient to use the GSVD in the form given below.

Corollary 3.1.1. *Suppose that $m \geq n \geq p$, $\mathrm{Ker}(\mathbf{A}) \cap \mathrm{Ker}(\mathbf{L}) = 0$, $\mathbf{L}$ has rank p and the augmented matrix $(\mathbf{A}; \mathbf{L})$ has rank n. Then, for matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{L} \in \mathbb{R}^{p \times n}$, there exist orthogonal matrices $\mathbf{U} \in \mathbb{R}^{m \times m}$, $\mathbf{V} \in \mathbb{R}^{p \times p}$ and a nonsingular matrix $X \in \mathbb{R}^{n \times n}$ such that*

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} \mathbf{D_A} & O_{p\times(n-p)} \\ O_{(n-p)\times p} & I_{n-p} \\ O_{(m-n)\times p} & O_{n\times(n-p)} \end{bmatrix} X^{-1}, \mathbf{L} = \mathbf{V}[\ \mathbf{D_L}\ \ O_{p\times(n-p)}\ ]X^{-1},$$

$$\mathbf{D_A} = \mathrm{diag}(\lambda_{\mathbf{A}1}, \ldots, \lambda_{\mathbf{A}p}), \mathbf{D_L} = \mathrm{diag}(\lambda_{\mathbf{L}1}, \ldots, \lambda_{\mathbf{L}p}),$$

and

$$\lambda_{\mathbf{A}i}^2 + \lambda_{\mathbf{L}i}^2 = 1, i = 1, \ldots, p,$$

with

$$0 \leq \lambda_{\mathbf{A}1} \leq \cdots \leq \lambda_{\mathbf{A}p} \leq 1, 1 \geq \lambda_{\mathbf{L}1} \geq \cdots \geq \lambda_{\mathbf{L}p} > 0.$$

This form is a special case of that in Theorem 1.2.3 for $\mathbf{L} = \mathbf{B}$, $n = k$ and $X^{-1} = \mathbf{Q}^T \mathbf{D}_C \mathbf{V}^T$.

A proof of Theorem 1.3.2 can now be obtained in the following fashion. First, standard least-squares theory gives the minimizer of $J(\mathbf{x})$ as

$$\hat{\mathbf{x}} = (\mathbf{A}^T W_{\mathbf{b}} \mathbf{A} + \mathbf{L}^T W_{\mathbf{L}} \mathbf{L})^{-1} \mathbf{A}^T W_{\mathbf{b}} \mathbf{r} + \mathbf{x}_0$$

with

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}_0.$$

The optimized criterion function is therefore

$$\begin{aligned} J(\hat{\mathbf{x}}) &= \mathbf{r}^T [W_{\mathbf{b}} - W_{\mathbf{b}} \mathbf{A} (\mathbf{A}^T W_{\mathbf{b}} \mathbf{A} + \mathbf{L}^T W_{\mathbf{L}} \mathbf{L})^{-1} \mathbf{A}^T W_{\mathbf{b}}] \mathbf{r} \\ &= \tilde{\mathbf{r}}^T [I_m - \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^T \tilde{\mathbf{A}} + \tilde{\mathbf{L}}^T \tilde{\mathbf{L}})^{-1} \tilde{\mathbf{A}}^T] \tilde{\mathbf{r}} \end{aligned}$$

with

$$\tilde{\mathbf{r}} = W_{\mathbf{b}}^{1/2} \mathbf{r}, \tilde{\mathbf{A}} = W_{\mathbf{b}}^{1/2} \mathbf{A}, \text{and } \tilde{\mathbf{L}} = W_{\mathbf{L}}^{1/2} \mathbf{L}.$$

An application of Corollary 3.1.1 to $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{L}}$ produces

$$\tilde{\mathbf{A}} = \mathbf{U} \begin{bmatrix} \mathbf{D}_{\mathbf{A}} & O_{p \times (n-p)} \\ O_{(n-p) \times p} & I_{n-p} \\ O_{(m-n) \times p} & O_{(m-n) \times (n-p)} \end{bmatrix} X^{-1}$$

and

$$\tilde{\mathbf{L}} = \mathbf{V} \begin{bmatrix} \mathbf{D}_{\mathbf{L}(p \times p)} & O_{p \times (n-p)} \end{bmatrix} X^{-1}.$$

So

$$\begin{aligned}
\tilde{\mathbf{A}}^T\tilde{\mathbf{A}} &= (X^{-1})^T \begin{bmatrix} \mathbf{D}_{\mathbf{A}}^2 & O_{p\times(n-p)} \\ O_{(n-p)\times p} & I_{n-p} \end{bmatrix} X^{-1}, \\
\tilde{\mathbf{L}}^T\tilde{\mathbf{L}} &= (X^{-1})^T \begin{bmatrix} \mathbf{D}_{\mathbf{L}} \\ O_{(n-p)\times p} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{\mathbf{L}} & O_{p\times(n-p)} \end{bmatrix} X^{-1} \\
&= (X^{-1})^T \begin{bmatrix} \mathbf{D}_{\mathbf{L}}^2 & O_{p\times(n-p)} \\ O_{(n-p)\times p} & O_{n-p} \end{bmatrix} X^{-1}
\end{aligned}$$

and

$$\tilde{\mathbf{A}}^T\tilde{\mathbf{A}} + \tilde{\mathbf{L}}^T\tilde{\mathbf{L}} = (X^T)^{-1} \begin{bmatrix} I_p & O_{p\times(n-p)} \\ O_{(n-p)\times p} & I_{n-p} \end{bmatrix} X^{-1},$$

since $\mathbf{D}_{\mathbf{A}}^2 + \mathbf{D}_{\mathbf{L}}^2 = I_p$.

Combining our representations produces

$$\begin{aligned}
&\tilde{\mathbf{A}}(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}} + \tilde{\mathbf{L}}^T\tilde{\mathbf{L}})^{-1}\tilde{\mathbf{A}}^T \\
&= \mathbf{U} \begin{bmatrix} \mathbf{D}_{\mathbf{A}}^2 & O_{p\times(n-p)} & O_{p\times(m-n)} \\ O_{(n-p)\times p} & I_{n-p} & O_{(n-p)\times(m-n)} \\ O_{(m-n)\times p} & O_{(m-n)\times(n-p)} & O_{m-n} \end{bmatrix} \mathbf{U}^T.
\end{aligned}$$

Hence,

$$\begin{aligned}
&I_m - \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}} + \tilde{\mathbf{L}}^T\tilde{\mathbf{L}})^{-1}\tilde{\mathbf{A}}^T \\
&= \mathbf{U} \begin{bmatrix} \mathbf{D}_{\mathbf{L}}^2 & O_{p\times(n-p)} & O_{p\times(m-n)} \\ O_{(n-p)\times p} & O_{n-p} & O_{(n-p)\times(m-n)} \\ O_{(m-n)\times p} & O_{(m-n)\times(n-p)} & I_{m-n} \end{bmatrix} \mathbf{U}^T.
\end{aligned}$$

Now, since $\mathbf{b}|\mathbf{x} \sim N_m(\mathbf{A}\mathbf{x}, W_{\mathbf{b}}^{-1})$ and $\mathbf{x} \sim N_n(\mathbf{x}_0, C_{\mathbf{x}})$, we have

$$\mathbf{b} \sim N(\mathbf{A}\mathbf{x}_0, W_{\mathbf{b}}^{-1} + \mathbf{A}C_{\mathbf{x}}\mathbf{A}^T).$$

Consequently,

$$\tilde{\mathbf{r}} = W_{\mathbf{b}}^{1/2}[\mathbf{b} - \mathbf{A}\mathbf{x}_0] \sim N(0, I_m + \tilde{\mathbf{A}} C_{\mathbf{x}} \tilde{\mathbf{A}}^T). \tag{3.1}$$

But,

$$C_{\mathbf{x}} = (\mathbf{L}^T W_{\mathbf{L}} \mathbf{L})^{-1} = (\tilde{\mathbf{L}}^T \tilde{\mathbf{L}})^{-1}$$

for

$$\tilde{\mathbf{L}}^T \tilde{\mathbf{L}} = (X^{-1})^T \begin{bmatrix} \mathbf{D}_{\mathbf{L}}^2 & O_{p\times(n-p)} \\ O_{(n-p)\times p} & O_{n-p} \end{bmatrix} X^{-1}.$$

Thus,

$$C_{\mathbf{x}} = X \begin{bmatrix} \mathbf{D}_{\mathbf{L}}^{-2} & O_{p\times(n-p)} \\ O_{(n-p)\times p} & O_{n-p} \end{bmatrix} (X^T)^{-1}$$

and

$$I_m + \tilde{\mathbf{A}} C_{\mathbf{x}} \tilde{\mathbf{A}}^T = \mathbf{U} \begin{bmatrix} I_p + \mathbf{D}_{\mathbf{A}} \mathbf{D}_{\mathbf{L}}^{-2} \mathbf{D}_{\mathbf{A}} & O_{p\times(m-p)} \\ O_{(m-p)\times p} & I_{m-p} \end{bmatrix} \mathbf{U}^T.$$

From

$$\begin{aligned} I_p + \mathbf{D}_{\mathbf{A}} \mathbf{D}_{\mathbf{L}}^{-2} \mathbf{D}_{\mathbf{A}} &= \operatorname{diag}(1 + \lambda_{\mathbf{A}1}^2/\lambda_{\mathbf{L}1}^2, \ldots, 1 + \lambda_{\mathbf{A}p}^2/\lambda_{\mathbf{L}p}^2) \\ &= \operatorname{diag}((\lambda_{\mathbf{L}1}^2 + \lambda_{\mathbf{A}1}^2)/\lambda_{\mathbf{L}1}^2, \ldots, (\lambda_{\mathbf{L}p}^2 + \lambda_{\mathbf{A}p}^2)/\lambda_{\mathbf{L}p}^2) \\ &= \operatorname{diag}(1/\lambda_{\mathbf{L}1}^2, \ldots, 1/\lambda_{\mathbf{L}p}^2), \end{aligned}$$

we can conclude that

$$I_m + \tilde{\mathbf{A}} C_{\mathbf{x}} \tilde{\mathbf{A}}^T = \mathbf{U} \begin{bmatrix} \mathbf{D}_{\mathbf{L}}^{-2} & O_{p\times(m-p)} \\ O_{(m-p)\times p} & I_{m-p} \end{bmatrix} \mathbf{U}^T,$$

and

$$\mathbf{U}^T\tilde{\mathbf{r}} \sim N\left(0, \begin{bmatrix} \mathbf{D}_\mathbf{L}^{-2} & O_{p\times(m-p)} \\ O_{(m-p)\times p} & I_{m-p} \end{bmatrix}\right).$$

Therefore,

$$J(\hat{\mathbf{x}}) = (\mathbf{U}^T\tilde{\mathbf{r}})^T \begin{bmatrix} \mathbf{D}_\mathbf{L}^2 & O_{p\times(n-p)} & O_{p\times(m-n)} \\ O_{(n-p)\times p} & O_{n-p} & O_{(n-p)\times(m-n)} \\ O_{(m-n)\times p} & O_{(m-n)\times(n-p)} & I_{m-n} \end{bmatrix} \mathbf{U}^T\tilde{\mathbf{r}}.$$

Let

$$\begin{aligned} \mathbf{z} &= \begin{bmatrix} \mathbf{D}_\mathbf{L} & O_{p\times(m-p)} \\ O_{(m-p)\times p} & I_{m-p} \end{bmatrix} \mathbf{U}^T\tilde{\mathbf{r}} \\ &= (z_1, \ldots, z_m)^T. \end{aligned} \tag{3.2}$$

Then, by (3.1), $z_1, \ldots, z_m$ are independent and identically distributed (i.i.d.) standard normal random variables and

$$J(\hat{\mathbf{x}}) = \sum_{j=1}^{p} z_j^2 + \sum_{j=n+1}^{m} z_j^2. \tag{3.3}$$

Since $z_j \sim N(0,1)$, z_j^2 has a χ^2 distribution with one degree of freedom and Theorem 1.3.2 follows.

3.2 Nonnormal Data

The assumption of normality is quite restrictive. In this section, we relax the normality assumption and instead suppose only that

$$\mathrm{E}(\mathbf{b}|\mathbf{x}) = \mathbf{A}\mathbf{x}, \mathrm{Cov}(\mathbf{b}|\mathbf{x}) = W_\mathbf{b}^{-1}, \mathrm{E}(\mathbf{x}) = \mathbf{x}_0,$$

and

$$\mathrm{Cov}(\mathbf{x}) = C_\mathbf{x} = (\mathbf{L}^T W_\mathbf{L} \mathbf{L})^-,$$

or, equivalently, that

$$W_{\mathbf{L}} = (\mathbf{L}C_{\mathbf{x}}\mathbf{L}^T)^-.$$

Since the arguments in the previous section used only second moment properties, representation (3.3) continues to hold with $z_1, \ldots, z_m$ now being uncorrelated random variables that have zero means and unit variances.

To proceed further, we will assume that the sequence of random variables $\{z_m\}$ is ϕ-mixing as defined below.

Definition 3.2.1. *A σ-field on a set E is a family $\mathcal{M}$ of subsets of E such that:*

(i) $\emptyset \in \mathcal{M}$;

(ii) if $\mathcal{A} \in \mathcal{M}$, *then* $\mathcal{A}^c \in \mathcal{M}$;

(iii) if $\mathcal{A}_1, \mathcal{A}_2, \ldots \in \mathcal{M}$, *then* $\cup_{k=1}^{\infty} \mathcal{A}_k \in \mathcal{M}$.

Definition 3.2.2. *[5] Let $\{\xi_k\}_{k=-\infty}^{\infty}$ be a sequence of random variables on a probability space $(\Omega, \mathcal{A}, \mathrm{P})$, denote the σ-field generated by $\ldots, \xi_{k-1}, \xi_k$ as μ_k and take μ^k to be the σ-field generated by $\xi_k, \xi_{k+1}, \ldots$. Then, $\{\xi_k\}$ is ϕ-mixing if there exists a nonnegative function φ such that*

$$|\mathrm{P}(E_1 \cap E_2) - \mathrm{P}(E_1)\mathrm{P}(E_2)| \leq \varphi(n)\mathrm{P}(E_1)$$

for any $E_1 \in \mu_k$, $E_2 \in \mu^{k+n}$ and

$$\lim_{n\to\infty} \varphi(n) = 0.$$

Here, P denotes the probability measure for the probability space corresponding to the $\{\xi_k\}$.

Example 3.2.1. *The most obvious example of ϕ-mixing is independence in which case $\varphi(m) \equiv 0$.*

Example 3.2.2. *[5] Let $\{y_k\}$ be a stationary Markov process with finite state space. Denote the stationary probability for the Markov process as p_u, the 1-step transition probability as $p_{uv} = \mathrm{P}(y_{k+1} = v | y_k = u) > 0$ and the n-step transition probability as $p_{uv}^{(n)} = \mathrm{P}(y_{n+k} = v | y_k = u)$. For $i, j \geq 0$, $i \neq j$, let $\mathbb{H}_i$ be a set of $(i+1)$-tuples of states and $\mathbb{H}_j$ be a set of $(j+1)$-tuples of states. Then, for $E_1 = \{(y_{k-i}, \ldots, y_k) \in \mathbb{H}_i\}$ and $E_2 = \{(y_{k+n}, \ldots, y_{k+n+j}) \in \mathbb{H}_j\}$, we have*

$$
\begin{aligned}
& |\mathrm{P}(E_1 \cap E_2) - \mathrm{P}(E_1)\mathrm{P}(E_2)| \\
\leq & \sum_{u_0,\ldots,u_i,v_0,\ldots,v_j} p_{u_0} p_{u_0 u_1} \cdots p_{u_{i-1} u_i} |p_{u_i v_0}^{(n)} - p_{v_0}| p_{v_0 v_1} \cdots p_{v_{j-1} v_j} \\
= & \sum_{u_0,\ldots,u_i,v_0,\ldots,v_{j-1}} p_{u_0} p_{u_0 u_1} \cdots p_{u_{i-1} u_i} \left| \frac{p_{u_i v_0}^{(n)} - p_{v_0}}{p_{v_0}} \right| p_{v_0} p_{v_0 v_1} \cdots p_{v_{j-2} v_{j-1}} \sum_{v_j} p_{v_{j-1} v_j} \\
= & \ \ldots = \sum_{u_0,\ldots,u_i,v_0} p_{u_0} p_{u_0 u_1} \cdots p_{u_{i-1} u_i} \left| \frac{p_{u_i v_0}^{(n)} - p_{v_0}}{p_{v_0}} \right| p_{v_0},
\end{aligned}
$$

using the property for a Markov process that $\sum_v p_{uv} = 1$. Taking $\varphi(n) = \max_{u_i, v_0} \left| \frac{p_{u_i v_0}^{(n)} - p_{v_0}}{p_{v_0}} \right|$ gives

$$
|\mathrm{P}(E_1 \cap E_2) - \mathrm{P}(E_1)\mathrm{P}(E_2)| \leq \varphi(n)\mathrm{P}(E_1).
$$

If we assume the transition matrix is irreducible and aperiodic, we have $p_{u_i v_0}^{(n)} \to p_{v_0}$ when $n \to \infty$. Thus, $\varphi(n) \to 0$ and $\{y_k\}$ is ϕ-mixing.

If we can assume the z_j in (3.3) are ϕ-mixing, the large sample properties of $J(\hat{\mathbf{x}})$ for the general case can be established using a functional central limit theorem from Billingsley [5] (Theorem 21.1 on page 184) that we restate here for completeness of exposition.

Theorem 3.2.1. *Assume that $\{\xi_k\}$ is a sequence of mean zero random variables that is ϕ-mixing with*

$$\sum_{n=1}^{\infty} \varphi(n)^{1/2} < \infty.$$

Then,

$$\sigma^2 = \mathrm{Var}(\xi_1) + 2\sum_{k=2}^{\infty} \mathrm{Cov}(\xi_1, \xi_k) \tag{3.4}$$

converges absolutely. If $\sigma^2 > 0$, then $\left(\sum_{j=1}^{m} \xi_j\right)/(\sqrt{m}\sigma)$ has a standard normal limiting distribution as m goes to infinity.

Subsequently we will use $\xrightarrow{d}$ to indicate convergence in distribution. Now take $\xi_j = z_j^2 - 1$ and apply Theorem 3.2.1 to obtain

$$\frac{\sum_{j=1}^{m} z_j^2 - m}{\sqrt{m}\sigma} \xrightarrow{d} N(0,1),$$

as $m \to \infty$. In terms of $J(\hat{\mathbf{x}})$, we will employ Theorem 3.2.1 in conjunction with the following result that is sometimes referred to as Slutsky's Theorem.

Theorem 3.2.2. *[25] If a sequence of random variables Y_n converges in distribution to a random variable Y, and A_n and B_n tend in probability to constants a and b, respectively, then $A_n + B_n Y_n$ converges in distribution to $a + bY$.*

Our main result concerning the large sample properties of $J(\hat{\mathbf{x}})$ is given below.

Corollary 3.2.1. *Under the conditions of Theorem 3.2.1 for $\xi_j = z_j^2 - 1$ with z_j in (3.3), if $(n-p)/\sqrt{m} \to 0$ as $m \to \infty$, then,*

$$\frac{1}{\sigma\sqrt{m-n+p}}[J(\hat{\boldsymbol{x}}) - (m-n+p)] \xrightarrow{d} N(0,1).$$

Proof. By Markov's inequality, for any $\epsilon > 0$,

$$\mathrm{P}\left(\frac{1}{\sqrt{m}}\sum_{j=p+1}^{n} z_j^2 > \epsilon\right) \leq \frac{n-p}{\epsilon\sqrt{m}}.$$

By assumption, $(n-p)/(\epsilon\sqrt{m}) \to 0$ so that $(1/\sqrt{m})\sum_{j=p+1}^{n} z_j^2$ converges to zero in probability. Hence,

$$\begin{aligned}
&\frac{1}{\sigma\sqrt{m-n+p}}[J(\hat{\mathbf{x}}) - (m-n+p)] \\
&= \frac{\sqrt{m}}{\sqrt{m-n+p}}\frac{1}{\sqrt{m}\sigma}\left(\sum_{j=1}^{m} z_j^2 - m - \sum_{j=p+1}^{n} z_j^2 + (n-p)\right) \xrightarrow{d} N(0,1)
\end{aligned}$$

by Theorem 3.2.2. □

Another immediate result is

Corollary 3.2.2. *Assume the conditions of Theorem 3.2.1 hold for $\xi_j = z_j^2 - 1$ with z_j in (3.3), and that $(n-p)/\sqrt{m} \to 0$ as $m \to \infty$. If*

$$\mathrm{Cov}(z_i^2, z_j^2) = 2\delta_{ij}, \text{ for all } i, j,$$

$$T_m = \frac{J(\hat{x}) - (m-n+p)}{\sqrt{2(m-n+p)}}$$

has the same limiting distribution as

$$\tilde{T_m} = \frac{V_m - (m-n+p)}{\sqrt{2(m-n+p)}}$$

with V_m distributed as a χ^2 random variable with $m-n+p$ degrees of freedom; i.e., $\tilde{T_m}$, T_m both have limiting standard normal distributions.

Corollary 3.2.2 applies to the normal theory case, for example. More generally, it states that the normal theory formulation for the χ^2 method can be expected to work effectively in large samples if the squares of the z_j in (3.3)

are uncorrelated. Of course, there is no reason to expect this to be true in general. When the z_j^2 are correlated, the correct scaling factor in the central limit theorem is provided by (3.4) rather than the 2 that applies under independence. The prediction interval used by the χ^2 method then needs to have the form

$$m - n + p \pm \sigma\sqrt{m - n + p}z_{\alpha/2}. \tag{3.5}$$

We discuss the consequences of this fact in the next section.

3.3 Implications

In this section we present the results of both analytic and empirical work that have been used to explore the implications of our findings in the previous section. Here we focus on the case where $\mathbf{b}$ derives from a Poisson distribution which is often a relevant assumption in the areas of astronomy, microscopy and medical imaging. (See [41].)

3.3.1 A Simple Example

In this section we examine a toy model where it is possible to obtain an analytic expression for the variance of $J(\hat{\mathbf{x}})$. Specifically, assume that the response vector is

$$\mathbf{b} = I\mathbf{x} + \boldsymbol{\varepsilon}$$

with $\mathbf{x} \sim N(0, \sigma_{\mathbf{x}}^2 I)$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_m)^T$ is a vector of i.i.d. random variables from a mean centered Poisson distribution with parameter $\sigma_{\mathbf{b}}^2$ that are also independent of the components of $\mathbf{x}$.

We need the moments of ε_1 up to order 4 for what follows. These can be obtained from the moment generating function

$$\begin{aligned} M_{\varepsilon_1}(t) &= \mathrm{E}(e^{t\varepsilon_1}) \\ &= e^{-\sigma_{\mathbf{b}}^2 t} M_{\varepsilon_1+\sigma_{\mathbf{b}}^2}(t) \\ &= e^{-\sigma_{\mathbf{b}}^2 t} e^{\sigma_{\mathbf{b}}^2(e^t-1)} \\ &= e^{\sigma_{\mathbf{b}}^2(e^t-t)-\sigma_{\mathbf{b}}^2}. \end{aligned}$$

From this we obtain

$$\begin{aligned} M'_{\varepsilon_1}(t) &= \sigma_{\mathbf{b}}^2(e^t-1)M_{\varepsilon_1}(t), \\ M''_{\varepsilon_1}(t) &= \sigma_{\mathbf{b}}^2(e^t-1)M'_{\varepsilon_1}(t) + \sigma_{\mathbf{b}}^2 e^t M_{\varepsilon_1}(t), \\ M'''_{\varepsilon_1}(t) &= \sigma_{\mathbf{b}}^2(e^t-1)M''_{\varepsilon_1}(t) + 2\sigma_{\mathbf{b}}^2 e^t M'_{\varepsilon_1}(t) + \sigma_{\mathbf{b}}^2 e^t M_{\varepsilon_1}(t), \\ M^{(iv)}_{\varepsilon_1}(t) &= \sigma_{\mathbf{b}}^2(e^t-1)M'''_{\varepsilon_1}(t) + 3\sigma_{\mathbf{b}}^2 e^t M''_{\varepsilon_1}(t) + 3\sigma_{\mathbf{b}}^2 e^t M'_{\varepsilon_1}(t) + \sigma_{\mathbf{b}}^2 e^t M_{\varepsilon_1}(t). \end{aligned}$$

Thus,

$$\mathrm{E}(\varepsilon_1) = M'_{\varepsilon_1}(t)|_{t=0} = 0, \tag{3.6}$$

$$\mathrm{E}(\varepsilon_1^2) = M''_{\varepsilon_1}(t)|_{t=0} = \sigma_{\mathbf{b}}^2, \tag{3.7}$$

$$\mathrm{E}(\varepsilon_1^3) = M'''_{\varepsilon_1}(t)|_{t=0} = \sigma_{\mathbf{b}}^2, \tag{3.8}$$

$$\mathrm{E}(\varepsilon_1^4) = M^{(iv)}_{\varepsilon_1}(t)|_{t=0} = 3\sigma_{\mathbf{b}}^4 + \sigma_{\mathbf{b}}^2. \tag{3.9}$$

Under our simple model,

$$\begin{aligned} J(\hat{\mathbf{x}}) &= \frac{\mathbf{b}^T\mathbf{b}}{\sigma_{\mathbf{b}}^2+\sigma_{\mathbf{x}}^2} \\ &= \frac{\mathbf{x}^T\mathbf{x} + 2\mathbf{x}^T\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}{\sigma_{\mathbf{b}}^2+\sigma_{\mathbf{x}}^2}. \end{aligned} \tag{3.10}$$

Since $\mathbf{x} = (x_1, \ldots, x_m)^T$ has i.i.d. $N(0, \sigma_{\mathbf{x}}^2)$ components, we know that

$$\mathrm{E}(\mathbf{x}^T\mathbf{x}) = \sum_{i=1}^{m} \mathrm{E}(x_i^2) = m\sigma_{\mathbf{x}}^2 \tag{3.11}$$

and

$$\begin{aligned} \mathrm{E}(\mathbf{x}^T\mathbf{x})^2 &= \sum_{i,j=1}^{m} \mathrm{E}(x_i^2 x_j^2) \\ &= 3m\sigma_{\mathbf{x}}^4 + m(m-1)\sigma_{\mathbf{x}}^4 = m^2\sigma_{\mathbf{x}}^4 + 2m\sigma_{\mathbf{x}}^4 \end{aligned} \tag{3.12}$$

because the fourth moment of the standard normal distribution is 3. In particular, we can now verify directly that $\mathrm{E}(J(\hat{\mathbf{x}})) = m$ by (3.10), (3.11) and (3.7).

To obtain the variance of $J(\hat{\mathbf{x}})$, we first compute its second moment from (3.10) as

$$\begin{aligned} &\mathrm{E}(J(\hat{\mathbf{x}})^2) \\ =\ &\mathrm{E}\left(\frac{(\mathbf{x}^T\mathbf{x})^2 + 4(\mathbf{x}^T\boldsymbol{\varepsilon})^2 + (\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})^2 + 4\mathbf{x}^T\mathbf{x}\mathbf{x}^T\boldsymbol{\varepsilon} + 4\mathbf{x}^T\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} + 2\mathbf{x}^T\mathbf{x}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon}}{(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{b}}^2)^2}\right). \end{aligned}$$

This expression can be simplified by application of the relations

$$\mathrm{E}((\mathbf{x}^T\mathbf{x})(\mathbf{x}^T\boldsymbol{\varepsilon})) = \mathrm{E}((\mathbf{x}^T\boldsymbol{\varepsilon})(\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})) = 0, \tag{3.13}$$

$$\mathrm{E}((\mathbf{x}^T\mathbf{x})(\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})) = m^2\sigma_{\mathbf{x}}^2\sigma_{\mathbf{b}}^2, \tag{3.14}$$

$$\mathrm{E}((\mathbf{x}^T\boldsymbol{\varepsilon})^2) = m\sigma_{\mathbf{x}}^2\sigma_{\mathbf{b}}^2, \tag{3.15}$$

$$\begin{aligned} \mathrm{E}((\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon})^2) &= m\mathrm{E}(\varepsilon_1^4) + m(m-1)\mathrm{E}(\varepsilon_1\varepsilon_2) \\ &= m(3\sigma_{\mathbf{b}}^4 + \sigma_{\mathbf{b}}^2) + m(m-1)\sigma_{\mathbf{b}}^4 \\ &= (m^2 + 2m)\sigma_{\mathbf{b}}^4 + m\sigma_{\mathbf{b}}^2. \end{aligned} \tag{3.16}$$

This results in

$$\mathrm{E}(J(\hat{\mathbf{x}})^2) = m^2 + 2m + \frac{m\sigma_{\mathbf{b}}^2}{(\sigma_{\mathbf{b}}^2 + \sigma_{\mathbf{x}}^2)^2}.$$

Thus,

$$\begin{aligned} \mathrm{Var}(J(\hat{\mathbf{x}})) &= \mathrm{E}(J(\hat{\mathbf{x}})^2) - (\mathrm{E}(J(\hat{\mathbf{x}})))^2 \\ &= 2m + \frac{m\sigma_{\mathbf{b}}^2}{(\sigma_{\mathbf{b}}^2 + \sigma_{\mathbf{x}}^2)^2} > 2m. \end{aligned}$$

Note that $\text{Var}(J(\hat{\mathbf{x}}))$ does not behave like $2m$ even in an asymptotic sense as

$$\frac{\text{Var}(J(\hat{\mathbf{x}}))}{m} \to 2 + \frac{\sigma_{\mathbf{b}}^2}{(\sigma_{\mathbf{b}}^2 + \sigma_{\mathbf{x}}^2)^2} = 2 + \frac{1}{(\sigma_{\mathbf{b}} + \frac{\sigma_{\mathbf{x}}^2}{\sigma_{\mathbf{b}}})^2}$$

as $m \to \infty$. However, large values of $\sigma_{\mathbf{b}}^2$, small values of $\sigma_{\mathbf{b}}^2$ relative to $\sigma_{\mathbf{x}}^2$, or large values of $\sigma_{\mathbf{x}}^2$ relative to $\sigma_{\mathbf{b}}^2$ will provide cases where a choice of $2m$ for the variance of $J(\hat{\mathbf{x}})$ may be effective. These cases are consistent with what we would expect. When $\sigma_{\mathbf{b}}^2$ ($\sigma_{\mathbf{x}}^2$) is small (large) relative to $\sigma_{\mathbf{x}}^2$ ($\sigma_{\mathbf{b}}^2$), $\mathbf{b}$ is dominated by $\mathbf{x}$ and will be roughly $N(0, \sigma_{\mathbf{x}}^2 I)$. On the other hand, if $\sigma_{\mathbf{b}}^2$ is large, $\mathbf{b}^T\mathbf{b}$ is basically a sum of squares of independent, mean centered Poisson random variables and the central limit theorem provides the normal approximation for $J(\hat{\mathbf{x}})$.

With this as an introduction, we now turn to empirical examples with problems of a more complicated and realistic nature. That is the subject of the next section.

3.3.2 Empirical Study

We now describe the outcome of some Monte Carlo experiments. These were carried out using both normal and Poisson data for comparison purposes. For both instances, we took $m = n = p$. That is, we generate data from the model

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}$$

with $\mathbf{x}$ having a $N(0, \sigma_{\mathbf{x}}^2 I)$ distribution and $\boldsymbol{\varepsilon}$ either having $N(0, \sigma_{\mathbf{b}}^2 I)$ or a mean centered Poisson distribution with variance $\sigma_{\mathbf{b}}^2$. The matrix $\mathbf{A}$ is chosen from one of two matrix types that are known to be ill conditioned: namely,

(i) the Phillips [18] matrix. This matrix arises from discretization of the Fredholm integral equation of the first kind from Section 1.2.2 on $[-6, 6]$.

The Kernel K is given by

$$K(s,t) \;=\; \psi(s-t)$$

with

$$\psi(x) = \begin{cases} 1+\cos\left(\frac{\pi x}{3}\right), & |x|<3, \\ 0, & |x|\geq 3. \end{cases}$$

(ii) the Shaw [18] matrix. This also arises from discretization of the Fredholm integral equation of the first kind on $[-\pi/2, \pi/2]$ with

$$\begin{aligned} K(s,t) &= (\cos(s)+\cos(t))^2 \left(\frac{\sin(u)}{u}\right)^2, \\ u &= \pi(\sin(s)+\sin(t)). \end{aligned}$$

These matrices are evaluated using the regularization toolbox [18] from Matlab.

The error vector $\boldsymbol{\varepsilon}$ is generated from a given distribution (either normal or Poisson) and then added to $\mathbf{Ax}$ after generating $\mathbf{x}$ from $N(0, \sigma_{\mathbf{x}}^2 I)$. Then, the simulation scheme is set out in Algorithm 3.3.1.

ALGORITHM 3.3.1

1: For $i = 1, \ldots, i_0$
2: For $j = 1, \ldots, j_0$
3: Calculate $J(\hat{\mathbf{x}})_{ij}$
4: Calculate $sJ(\hat{\mathbf{x}})_{ij} = (J(\hat{\mathbf{x}})_{ij} - m)/\sqrt{m}$
5: End j
6: Use the j_0 $sJ(\hat{\mathbf{x}})_{ij}$ values to assess the goodness-of-fit for the $N(0,2)$ distribution with the Kolmogorov Smirnov statistic. Denote the p-value of the test by pv_i
7: End i
8: Test that the i_0 p-values pv_i come from a uniform distribution on $[0,1]$

The basic premise is that the $sJ(\hat{\mathbf{x}})_{ij}$ should behave like values from a normal distribution with variance 2 when both $\mathbf{x}$ and $\boldsymbol{\varepsilon}$ have normal distributions.

The objective is to see how this changes when $\boldsymbol{\varepsilon}$ is Poisson. For this purpose, we keep track of various statistics over the course of the simulation: namely, the sample mean of the $\{sJ(\hat{\mathbf{x}})_{ij}\}_{j=1}^{j_0}$,

$$\overline{sJ(\hat{\mathbf{x}})_{i\cdot}} = \frac{\sum_{j=1}^{j_0} sJ(\hat{\mathbf{x}})_{ij}}{j_0},$$

the sample mean of the $\{\overline{sJ(\hat{\mathbf{x}})_{i\cdot}}\}_{i=1}^{i_0}$,

$$\overline{\overline{sJ(\hat{\mathbf{x}})_{\cdot\cdot}}} = \frac{\sum_{i=1}^{i_0} \overline{sJ(\hat{\mathbf{x}})_{i\cdot}}}{i_0},$$

the simulation variance of the $\{sJ(\hat{\mathbf{x}})_{ij}\}_{j=1}^{j_0}$,

$$\widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})_{i\cdot}) = \sum_{j=1}^{j_0} (sJ(\hat{\mathbf{x}})_{ij} - \overline{sJ(\hat{\mathbf{x}})_{i\cdot}})^2,$$

and the sample mean of the $\{\widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})_{i\cdot})\}_{i=1}^{i_0}$

$$\overline{\widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})_{\cdot\cdot})} = \frac{\sum_{i=1}^{i_0} \widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})_{i\cdot})}{i_0}.$$

Similarly, the sample mean and variance of the p-values $\{pv_i\}_{i=1}^{i_0}$ of the Kolmogorov Smirnov test with $N(0,2)$ are $\overline{pv_\cdot}$ and $\widehat{\mathrm{Var}}(pv_\cdot)$ respectively.

Note that in step (iv), we do tests for the uniform distribution because under the null hypothesis H_0 that $J(\hat{\mathbf{x}})$ follows a $N(0,2)$ distribution, the p-values follow a uniform distribution on $[0,1]$. This comes from the following theorem.

Theorem 3.3.1. *Under the null hypothesis H_0 that a test statistic follows an assumed continuous distribution, the p-values of the test statistic follow a uniform distribution on* $[0,1]$.

Proof. If a random variable X has cumulative distribution function (CDF) $F(X)$, then $\mathrm{P}(F(X) \leq u) = \mathrm{P}(X \leq F^{-1}(u)) = F(F^{-1}(u)) = u$. In particular, if X is a test statistic, its p-value is $1 - F(x_0) = \mathrm{P}(X \geq x_0 | H_0 \text{ is true})$ with x_0 the observed value of X. The result now follows from the symmetry of the uniform distribution around 0.5. □

To evaluate the data produced by our Monte Carlo experiments we used the Kolmogorov Smirnov (KS) test [3]. This test is used to assess whether a sample of data is drawn from some specified distribution by measuring the distance between the empirical distribution and the one that was specified. That is, the null hypothesis H_0 is that the sample data comes from the given distribution while the alternative hypothesis H_a is that the sample data does not follow the given distribution.

Suppose $F(x)$ is the given distribution function, and $F_m(x)$ is its empirical counterpart from the sample data, which is defined as

$$F_m(x) = \frac{\text{number of elements in the sample} \leq x}{\text{sample size } m}.$$

Then, the KS test statistic is defined as $\sqrt{m} \sup_x |F_m(x) - F(x)|$, where m is the sample size [3]. The larger the test statistic, the more evidence there is to reject H_0. When $F(x)$ is continuous, under H_0, the test statistic converges to the Kolmogorov distribution. So, we reject H_0 when the test statistic is larger than critical values of the Kolmogorov distribution.

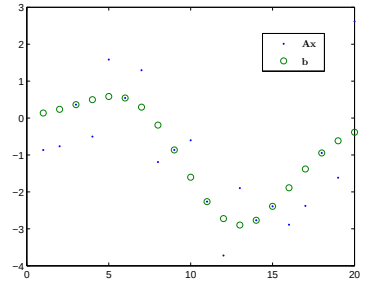
In our particular setting, the distribution will correspond to that of a $N(0, 2)$ distribution (i.e., for the $sJ(\hat{\mathbf{x}})$ data) or the uniform distribution on $[0, 1]$ (i.e., for the p-values). We chose $m \in \{20, 40, 80, 100, 200, 400, 800, 1600\}$, $j_0 = 50$ and $i_0 = 1000$. To simplify the plots, we indicate the values of m by $\{1, 2, \ldots, 8\}$ instead of $\{20, 40, 80, 100, 200, 400, 800, 1600\}$. For the normal

case, $\sigma_x^2 = 1.2^2$, $\sigma_b^2 = 1.5^2$ while for the Poisson case $\sigma_x^2 = 1.2^2$, $\sigma_b^2 = 1$. That is, we generate a random vector $\boldsymbol{\varepsilon}$ from the Poi(1)-1 distribution as the noise and form $\mathbf{b} = \mathbf{A}\mathbf{x} + \boldsymbol{\varepsilon}$ for each data set in the study.

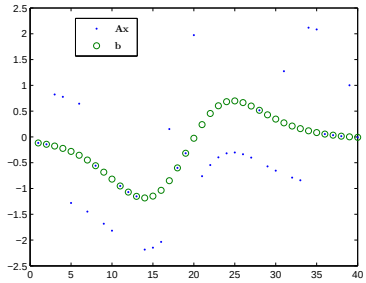
Figure 3.1 shows typical data sets for the Poisson case. Here a $\mathbf{b}$ and corresponding $\mathbf{A}\mathbf{x}$ for each size of Shaw matrix $\mathbf{A}$ illustrate how the Poisson noise is added. Now consider $J(\hat{\mathbf{x}})$ as a function of $\sigma_{\mathbf{x}}$ and denote it by $J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}})$. Then, Figure 3.2 plots the $J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}})$ functions for the data in Figure 3.1. The true value $\sigma_{\mathbf{x}} = 1.2$ is marked in the plots. The χ^2 method looks for the function's root to estimate $\sigma_{\mathbf{x}}$.

Figure 3.3 gives one example of relative histograms of $sJ(\hat{\mathbf{x}})$ for the normal and Poisson cases, respectively, that correspond to $i = 25$ and $m = 400$ in Algorithm 3.3.1. The p-values of the KS test with $N(0, 2)$ are 0.9709 for the normal case and 0.0261 for Poisson one, which indicates the normality of the distribution of $sJ(\hat{\mathbf{x}})$ in Figure 3.3 (a) and nonnormality of the distribution in (b).

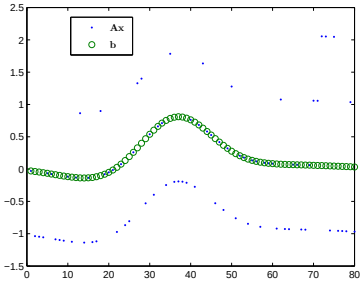
Summary plots of the statistics provided by our simulation are given in Figures 3.4-3.8. Figure 3.4 shows $\overline{\overline{sJ(\hat{\mathbf{x}})_{..}}}$ and $\overline{\widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})_{..})}$ as a function of m. We see that the mean is preserved but that the variance is increased as anticipated for the Poisson case. Figure 3.5 shows $\overline{pv.}$ and $\widehat{\mathrm{Var}}(pv.)$. Figure 3.6 provides the p-values of the KS test for the Unif$(0, 1)$ distribution. From these plots, we can see that $\overline{\widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})_{..})}$ is clearly larger than 2 and actually close to 3 in the Poisson case. Also, compared with the normal case, the KS test results show that $sJ(\hat{\mathbf{x}})$ does not fit the $N(0, 2)$ distribution for the Poisson case since the p-values are quite close to 0. For validation of the Poisson case, Figure 3.7 gives the results on p-values from the KS test with $N(0, \hat{\sigma}^2)$, where $\hat{\sigma}^2 = \overline{\widehat{\mathrm{Var}}(sJ(\hat{\mathbf{x}})..)}$ and Figure 3.8 shows the p-values of the KS test for the



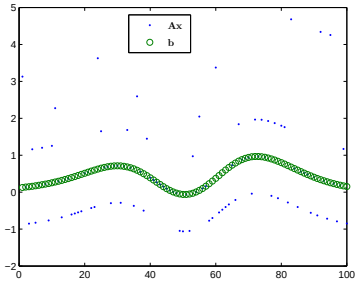
(a) m=20



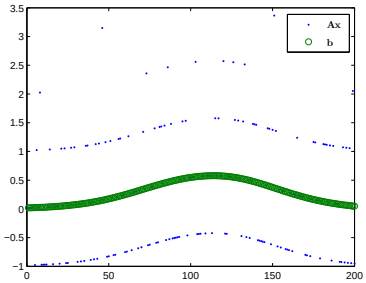
(b) m=40



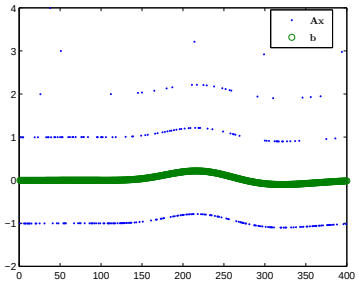
(c) m=80



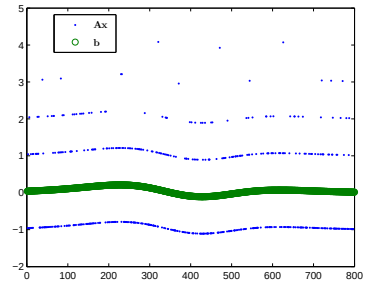
(d) m=100



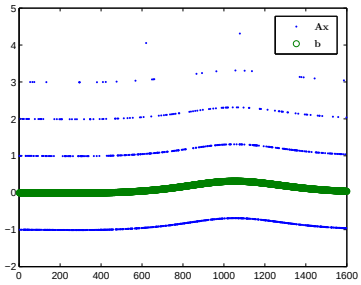
(e) m=200



(f) m=400

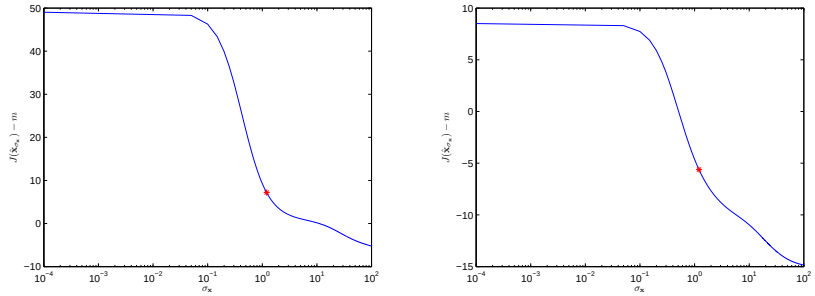


(g) m=800



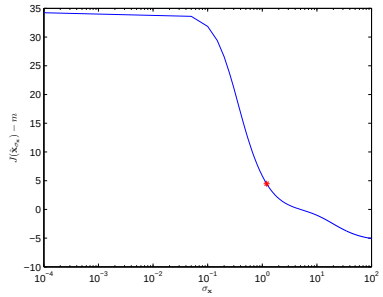
(h) m=1600

Figure 3.1: One example of generated **b** and **Ax** for different values of m

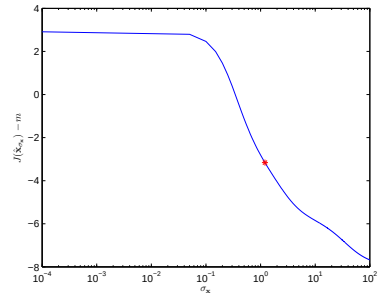


(a) m=20

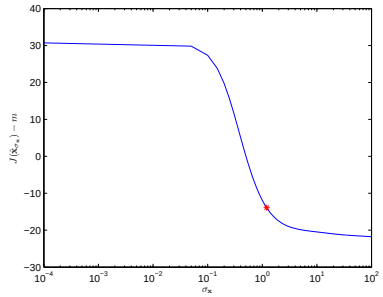
(b) m=40



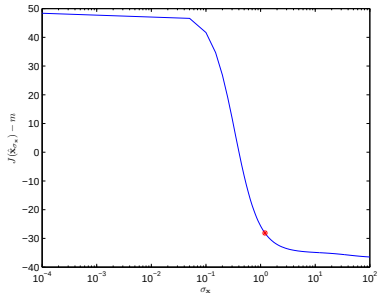
(c) m=80



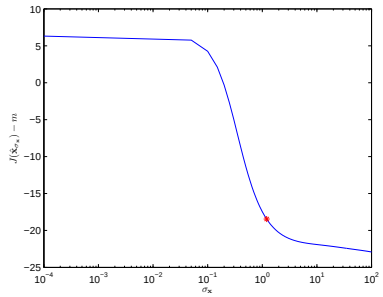
(d) m=100



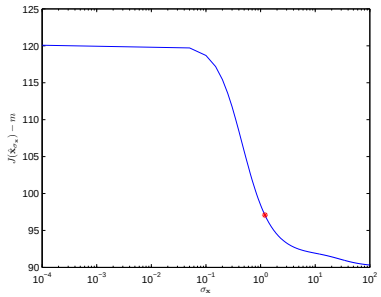
(e) m=200



(f) m=400

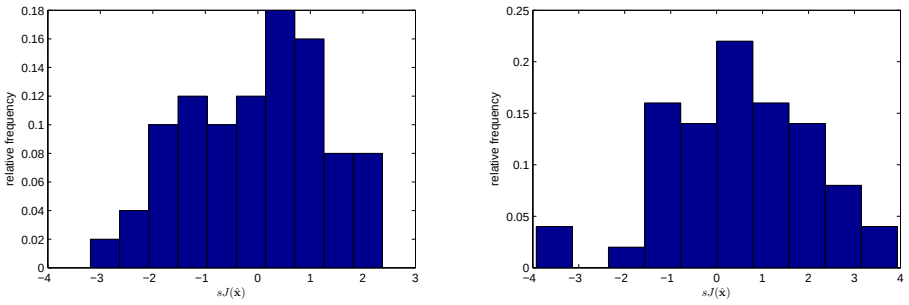


(g) m=800



(h) m=1600

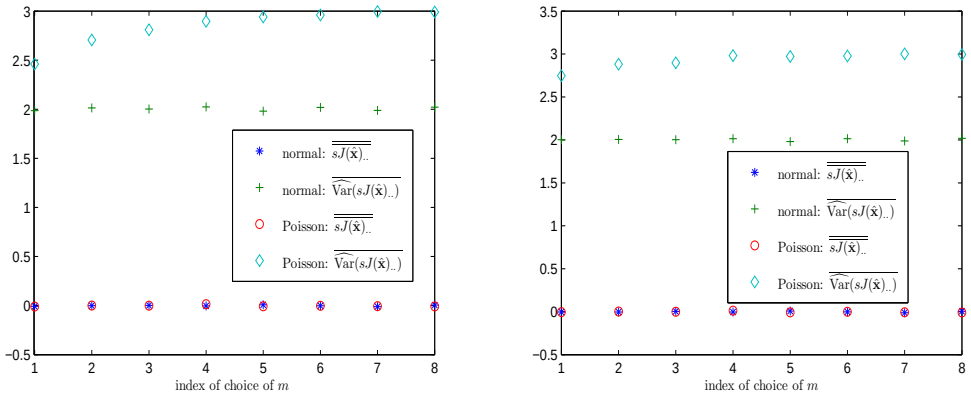
Figure 3.2: The $J(\hat{\mathbf{x}}_{\sigma_\mathbf{x}})$ function corresponding to data in Figure 3.1 (Logarithmic scale is used for x-axis.)



(a) Normal case (b) Poisson case

Figure 3.3: One example of the relative histograms of $sJ(\hat{\mathbf{x}})$

Unif$(0, 1)$ distribution. This confirms that at least in this particular instance, it is necessary to adjust the choice of σ^2 in (3.5) and use a value other than 2.

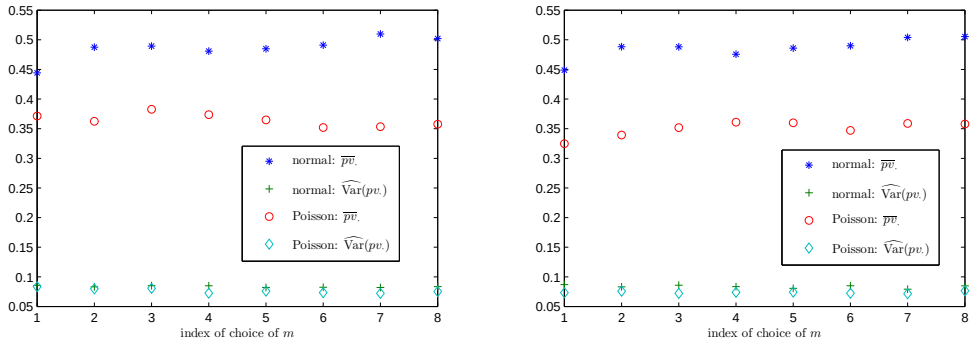


(a) Phillips matrix (b) Shaw matrix

Figure 3.4: Normal/Poisson case: mean of mean of $sJ(\hat{\mathbf{x}})$ and mean of variance of $sJ(\hat{\mathbf{x}})$

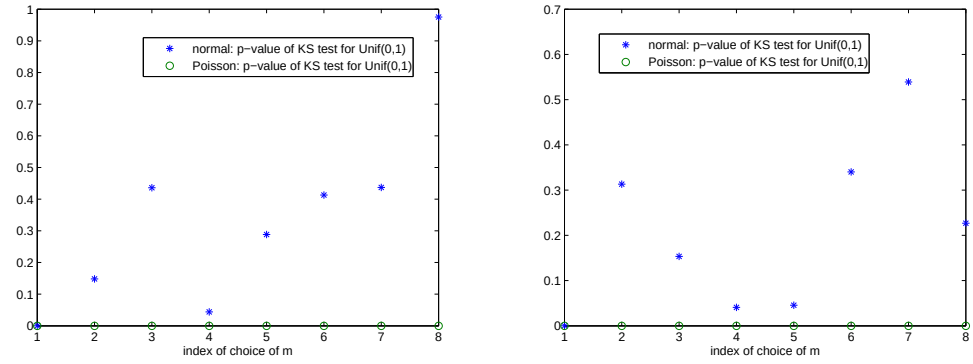
3.4 Parameter Estimation

From the work in the previous section we saw that the variance of $sJ(\hat{x})$ was better approximated by $\hat{\sigma}^2$ rather than 2 in the case of Poisson errors. The key question is whether or not this influences the χ^2 method for selection of the



(a) Phillips matrix (b) Shaw matrix

Figure 3.5: Normal/Poisson case: mean and variance of p-values from the KS test



(a) Phillips matrix (b) Shaw matrix

Figure 3.6: Normal/Poisson case: p-values of the KS test for Unif(0,1)

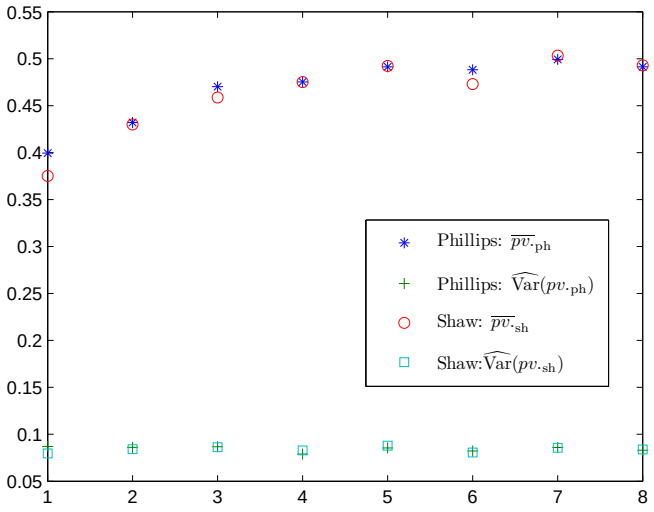


Figure 3.7: Poisson case with **A** Phillips / Shaw matrix: mean and variance of p-values from the KS test with normal distribution using adjusted variance $\hat{\sigma}^2$

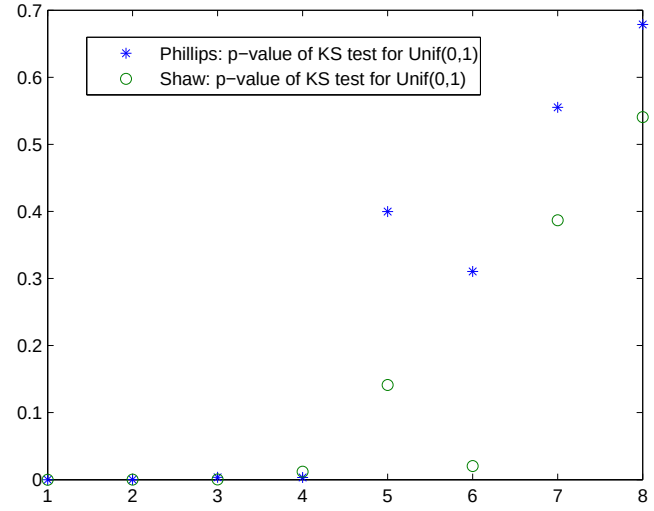


Figure 3.8: Poisson case with **A** Phillips / Shaw matrix: p-values of KS test for Unif$(0, 1)$

level of regularization. To investigate this issue, we conducted further experiments under the same basic design as in Section 3.3; i.e., we take $m = n = p$, $m \in \{20, 40, 80, 100, 200, 400, 800\}$, $\mathbf{x}_0 = 0$, $\mathbf{L} = I$, $W_{\mathbf{x}}^{-} = \sigma_{\mathbf{x}}^2 I = 1.2^2 I$ and $W_{\mathbf{b}}^{-1} = \sigma_{\mathbf{b}}^2 I = I$ in (1.36). We chose $\mathbf{A}$ to be the Phillips matrix and generated the same error vectors from the mean centered Poisson distribution as in Section 3.3. For the variance parameter we use the empirical approximation $\hat{\sigma}^2 \in \{2.4609, 2.7054, 2.8099, 2.8951, 2.9403, 2.9605, 2.9950\}$ with the order of these values corresponding to the different choices of m. We then estimate $\sigma_{\mathbf{x}}$ by the χ^2 method through finding the root of $J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m = 0$ with two different tolerances $\sqrt{2m}z_{\alpha/2}$ and $\sqrt{\hat{\sigma}^2 m}z_{\alpha/2}$. In order to have a more accurate estimation of $\sigma_{\mathbf{x}}$, we pick $\alpha = 0.95$ making a relatively small tolerance.

Algorithm 3.4.1 below describes the experimental procedure.

ALGORITHM 3.4.1

1: Generate $\mathbf{x}$ from $N(0, 1.2^2 I)$
2: For k=1,. . . ,1000
3: Generate $\mathbf{b}$ by adding centered noise $\boldsymbol{\varepsilon}$ to $\mathbf{Ax}$ such that $\mathbf{b} = \mathbf{Ax} + \boldsymbol{\varepsilon}$ where
$\boldsymbol{\varepsilon}$ is generated from $Poi(1) - 1$ distribution
4: For $\sigma^2 = 2, \hat{\sigma}^2$
5: Use the χ^2 method to estimate $\sigma_{\mathbf{x}}$
6: End σ^2
7: End k

The distributions of estimated $\sigma_{\mathbf{x}}$ for two cases $\sigma^2 = 2$ and $\sigma^2 = \hat{\sigma}^2$ are quite right-skewed. Thus, we show the median in Table 3.1 and find that there are not many differences between the two cases. In fact we are interested in the circumstance where the estimated $\sigma_{\mathbf{x}}$'s are different using $\sigma^2 = 2$ and $\sigma^2 = \hat{\sigma}^2$. r_D in Table 3.1 also indicates that around $5\% - 12\%$ out of 1000 estimated $\sigma_{\mathbf{x}}$ are different. Table 3.2 gives the five number summary for those cases having different estimated $\sigma_{\mathbf{x}}$ consisting of the minimum observation (min),

25th percentile (Q1), median, 75th percentile (Q3) and maximum observation (max). The range of the estimated $\sigma_{\mathbf{x}}$ is smaller for $\sigma^2 = \hat{\sigma}^2$. However, if we only look at the median, using $\sigma^2 = \hat{\sigma}^2$ only works better (closer to the true value $\sigma_{\mathbf{x}} = 1.2$) when $m = 40$ and $m = 80$.

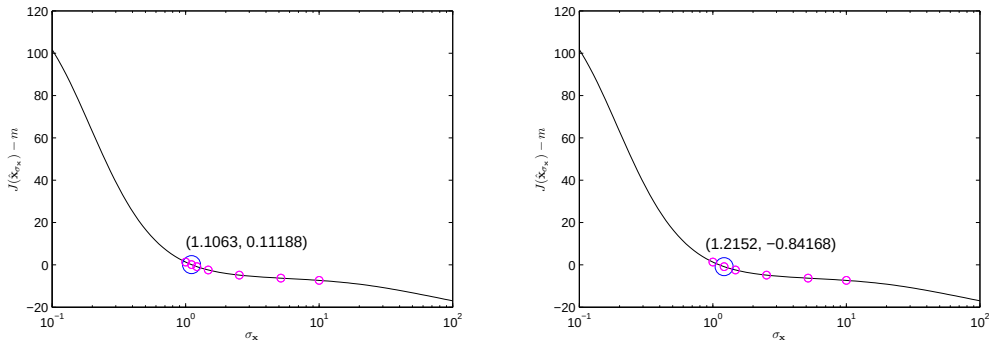
	$\sigma^2 = 2$	$\sigma^2 = \hat{\sigma}^2$	
m	median		r_D
20	0.8820	0.8793	4.7%
40	1.3966	1.3858	7.5%
80	1.0273	1.0354	7.5%
100	0.8022	0.8022	8.8%
200	0.6047	0.6047	9.1%
400	0.8407	0.7981	8.4%
800	1.4704	1.6330	11.6%

Table 3.1: The median of all the estimated $\sigma_{\mathbf{x}}$ for $\sigma^2 = 2$ and $\sigma^2 = \hat{\sigma}^2$ and percentage of cases where the estimates of $\sigma_{\mathbf{x}}$ differed

m	min	Q1	median	Q3	max
20($\sigma^2 = 2$)	0.473	0.64975	0.92854	2.048	39.522
20	0.46557	0.6314	0.89534	1.8975	35.454
40($\sigma^2 = 2$)	0.80496	0.91003	1.391	1.9101	54.875
40	0.79416	0.89134	1.3455	1.8181	51.165
80 ($\sigma^2 = 2$)	0.29563	0.59407	1.7439	3.9716	144.73
80	0.2893	0.57151	1.5645	4.7875	136.26
100 ($\sigma^2 = 2$)	0.28054	0.53812	1.4057	13.97	1712
100	0.27391	0.51398	1.4799	16.785	1639.1
200 ($\sigma^2 = 2$)	0.18826	0.37472	1.7039	80.524	2055.7
200	0.18096	0.35406	1.9049	100	1943.6
400 ($\sigma^2 = 2$)	0.20432	0.42302	1.9156	145.29	12728
400	0.19974	0.40399	3.5391	155.19	12096
800 ($\sigma^2 = 2$)	0.18952	0.71713	1.3066	146.39	11775
800	0.18434	0.66212	1.6922	181.97	11127

Table 3.2: Comparison of five number data summary of those different $\sigma_{\mathbf{x}}$'s using $\sigma^2 = 2$ and $\sigma^2 = \hat{\sigma}^2$ with sample size $1000r_D$

However, we find that the iteration times of the Newton method are less for cases where the estimators differ when we use $\sigma^2 = \hat{\sigma}^2$. This is due to the fact that the larger tolerance makes for faster convergence. We pick two examples for illustration. Figure 3.9 and Table 3.3 describe how the algorithm converges. In this case, $m = 80$ and $k = 35$. Figure 3.9 plots the function $J(\hat{\mathbf{x}}_{\sigma_\mathbf{x}}) - m$ as a function of $\sigma_\mathbf{x}$. Values at each iteration are shown by small circles with a large circle indicating the final estimator. Table 3.3 lists the values of $\sigma_\mathbf{x}$ and $(J(\hat{\mathbf{x}}_{\sigma_\mathbf{x}}) - m)_{it}$ at each iteration. With $\alpha = 0.95$, the tolerance is 0.7932 for $\sigma^2 = 2$ and 0.9402 for $\sigma^2 = \hat{\sigma}^2$. We can see that after iterating 6 steps, $J(\hat{\mathbf{x}}) - m$ reaches -0.8417 which satisfies the tolerance restriction for $\sigma^2 = \hat{\sigma}^2$ and concludes the iteration. Figure 3.10 and Table 3.4 show similar information for $m = 80$ and $k = 167$. Here, the number of iterations is reduced by two for $\sigma^2 = \hat{\sigma}^2$ as compared to the normal case with $\sigma^2 = 2$.

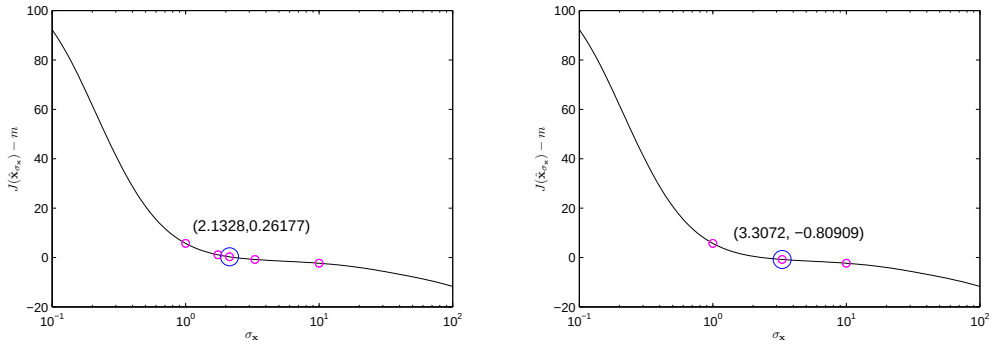


(a) $\sigma^2 = 2$ (b) $\sigma^2 = \hat{\sigma}^2$

Figure 3.9: An example for $k = 35$, $m = 80$

On the other hand, we are interested in whether the true $\sigma_\mathbf{x}$ falls in the confidence interval

$$\{\sigma_\mathbf{x} : -\sqrt{\sigma^2 m} z_{\alpha/2} < J(\hat{\mathbf{x}}_{\text{true } \sigma_\mathbf{x}}) - m < \sqrt{\sigma^2 m} z_{\alpha/2}\} \tag{3.17}$$



(a) $\sigma^2 = 2$ (b) $\sigma^2 = \hat{\sigma}^2$

Figure 3.10: Another example for $k = 167$, $m = 80$

	$\sigma^2 = 2$		$\sigma^2 = \hat{\sigma}^2$	
# of iterations	$(\sigma_{\mathbf{x}})_{it}$	$(J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m)_{it}$	$(\sigma_{\mathbf{x}})_{it}$	$(J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m)_{it}$
1	1	1.3152	1	1.3152
2	10	-7.342	10	-7.342
3	5.1695	-6.2719	5.1695	-6.2719
4	2.5220	-4.8339	2.5220	-4.8339
5	1.4771	-2.408	1.4771	-2.408
6	1.2152	-0.8417	1.2152	-0.8417
7	1.1063	0.1119		

Table 3.3: Values of $\sigma_{\mathbf{x}}$ and $(J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m)_{it}$ at each iteration for the example in Figure 3.9

	$\sigma^2 = 2$		$\sigma^2 = \hat{\sigma}^2$	
# of iterations	$(\sigma_{\mathbf{x}})_{it}$	$(J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m)_{it}$	$(\sigma_{\mathbf{x}})_{it}$	$(J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m)_{it}$
1	1	5.6967	1	5.6967
2	10	-2.3267	10	-2.3267
3	3.3072	-0.8091	3.3072	-0.8091
4	1.745	1.0904		
5	2.1328	0.2618		

Table 3.4: Values of $\sigma_{\mathbf{x}}$ and $(J(\hat{\mathbf{x}}_{\sigma_{\mathbf{x}}}) - m)_{it}$ at each iteration for the example in Figure 3.10

for $\sigma^2 = 2$ or $\sigma^2 = \hat{\sigma}^2$ and $\alpha = 0.95$. Let $tol = \sqrt{\sigma^2 m} z_{\alpha/2}$ then Table 3.5 shows the number of times out of 1000 that the true $\sigma_{\mathbf{x}}$ falls in the confidence interval (3.17). The use of $\sigma^2 = \hat{\sigma}^2$ as opposed to $\sigma^2 = 2$ results in a closer approximation of the expected 5%. For $m = 200$ and $m = 400$, the proportion of times that the true $\sigma_{\mathbf{x}}$ falls in the confidence interval for $\sigma^2 = \hat{\sigma}^2$ is significantly larger than that for $\sigma^2 = 2$ at the significance level 0.15. Table 3.5 lists the mean and standard deviation of the estimated $\sigma_{\mathbf{x}}$ values for those cases where the true $\sigma_{\mathbf{x}} = 1.2$ is inside the confidence interval. While there are no significant differences between the means for $\sigma^2 = 2$ and $\sigma^2 = \hat{\sigma}^2$, the standard deviations, when $\sigma^2 = \hat{\sigma}^2$, are in all but one instance larger than for $\sigma^2 = 2$. This suggests that the two choices for σ^2 produce similar $\sigma_{\mathbf{x}}$ values on the average but the estimated $\sigma_{\mathbf{x}}$ values exhibit more variation when σ^2 is chosen correctly.

If we consider the posterior uncertainty for $\mathbf{x}$, we have

$$\begin{aligned} \mathrm{Cov}(\mathbf{x}|\mathbf{b}) &= (\mathbf{A}^T W_{\mathbf{b}} \mathbf{A} + W_{\mathbf{x}})^{-1} \\ &= (\tilde{\mathbf{A}}^T \tilde{\mathbf{A}} + \sigma_{\mathbf{x}}^{-2} I)^{-1} \end{aligned}$$

for $\tilde{\mathbf{A}} = W_{\mathbf{b}}^{1/2} \mathbf{A}$. Applying SVD to $\tilde{\mathbf{A}}$ to have $\tilde{\mathbf{A}} = \mathbf{U}\mathrm{diag}(\sigma_{A1}, \ldots, \sigma_{Ak}, 0, \ldots, 0)\mathbf{V}^T$, where k is the rank of $\tilde{\mathbf{A}}$, we obtain

$$\mathrm{Cov}(\mathbf{x}|\mathbf{b}) = (\mathbf{V}^T)^{-1} \mathrm{diag}\left(\frac{1}{\sigma_{\mathbf{A}1}^2 + 1/\sigma_{\mathbf{x}}^2}, \ldots, \frac{1}{\sigma_{\mathbf{A}k}^2 + 1/\sigma_{\mathbf{x}}^2}, \sigma_{\mathbf{x}}^2, \ldots, \sigma_{\mathbf{x}}^2 \right) \mathbf{V}^{-1}.$$

It suggests that the posterior uncertainty for $\mathbf{x}$ is also more variable for $\sigma^2 = \hat{\sigma}^2$.

	$\sigma^2 = 2$	$\sigma^2 = \hat{\sigma}^2$	$\sigma^2 = 2$		$\sigma^2 = \hat{\sigma}^2$	
m	# of $\lvert J(\hat{\mathbf{x}}_{\text{true } \sigma_{\mathbf{x}}}) - m \rvert \leq tol$		mean	std	mean	std
20	47	51	1.1863	0.0381	1.1818	0.0447
40	48	53	1.1901	0.0259	1.1900	0.0305
80	30	36	1.1870	0.0609	1.1861	0.0597
100	32	40	1.2456	0.0869	1.2511	0.1855
200	42	53	1.4086	0.6295	1.5073	0.8112
400	45	56	1.2456	0.1372	1.2793	0.2583
800	47	54	1.2654	0.1586	1.3210	0.2534

Table 3.5: Experiment on whether the true $\sigma_{\mathbf{x}}$ is covered by interval (3.17)

CHAPTER 4

FUNCTIONAL CCA AND PCCA

4.1 Introduction

In this chapter we provide a new formulation of the functional CCA and PCCA concepts that were discussed in Chapter 1. We begin in this section with a few key definitions and some further literature review. In the next section, we describe the properties of the $\mathcal{H}$-valued random variables that arise in Section 4.2. Our main results for functional CCA are then laid out in Section 4.3. Section 4.4 extends the Section 4.3 work to include functional PCCA.

We begin with two important definitions.

Definition 4.1.1. *A random variable X on a probability space $\{\Omega, \mathcal{A}, \mathrm{P}\}$ is said to be second order if $\mathrm{E}|X|^2 = \int_\Omega |X|^2 d\mathrm{P} < \infty$.*

Definition 4.1.2. *A second-order stochastic process is a family of second-order random variables $\{X(t), t \in \mathbb{T}\}$ defined on a common probability space.*

Assume now that we have two second order stochastic processes

$$\{X_i(t) : t \in \mathbb{T}\}, i = 1, 2$$

for some index set $\mathbb{T}$. Then, provided all relevant variances are finite, we can define the covariance kernels

$$K_i(t, t') = \mathrm{Cov}(X_i(t), X_i(t'))$$

for $t, t' \in \mathbb{T}$, $i = 1, 2$. Inference is then based on the collection of random variables in $L^2_{X_i}$, $i = 1, 2$: the completion of the set of random variables of the

form of $\sum_{j=1}^{n} a_j X_i(t_j)$ for $a_j \in \mathbb{R}$, $t_j \in \mathbb{T}$ and $n \in \mathbb{Z}^+$, with the inner product

$$\langle \sum_{j=1}^{m} a_j X_i(t_j), \sum_{j'=1}^{n} b_{j'} X_i(t'_{j'}) \rangle_{L^2_{X_i}} = \sum_{j=1}^{m} \sum_{j'=1}^{n} a_i b_{j'} K_i(t_j, t'_{j'}).$$

Since covariance kernels are positive definite functions, they generate RKHSs in the sense defined below

Definition 4.1.3. *Let $\mathcal{H}$ be a Hilbert space of functions on some set $\mathbb{T}$ and denote by $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ the inner product in $\mathcal{H}$. A bivariate function K on $\mathbb{T} \times \mathbb{T}$ is said to be a reproducing kernel for $\mathcal{H}$ if*

(i) for every $t \in \mathbb{T}$, $K(\cdot, t) \in \mathcal{H}$ and

(ii) for every $t \in \mathbb{T}$ and $f \in \mathcal{H}$, $f(t) = \langle f, K(\cdot, t) \rangle_{\mathcal{H}}$.

When (i)–(ii) hold, $\mathcal{H}$ is said to be a RKHS with reproducing kernel K.

Aronszajn [1] shows that functions of the form $\sum_{i=1}^{m} a_i K(\cdot, t_i)$ for $a_i \in \mathbb{R}$, $t_i \in \mathbb{T}$ and $m = 1, 2, \ldots$, are dense in $\mathcal{H}$. So $\mathcal{H}$ is separable.

Now, each function K_1 and K_2 has a corresponding Hilbert function space $\mathcal{H}(K_1)$ and $\mathcal{H}(K_2)$ for which it is the reproducing kernel. The importance of this fact derives from the isometries between the $L^2_{X_i}$ and $\mathcal{H}(K_i)$ produced by the mappings $\Psi_i(K_i(\cdot, t)) = X_i(t)$ as demonstrated in numerous articles by Parzen (e.g., [32]).

Using these isometries, Eubank and Hsing [13] defined the first canonical correlation ρ and associated canonical variables $\Psi_1(f_1), \Psi_2(f_2)$ as

$$\begin{aligned} \rho^2 &= \sup_{a_1 \in L^2_{X_1}, a_2 \in L^2_{X_2}} \mathrm{Cov}^2(a_1, a_2) \\ &= \sup_{f_1 \in \mathcal{H}(K_1), f_2 \in \mathcal{H}(K_2)} \mathrm{Cov}^2(\Psi_1(f_1), \Psi_2(f_2)) \\ &= \mathrm{Cov}^2(\Psi_1(f_1), \Psi_2(f_2)), \end{aligned}$$

where f_1 and f_2 are subject to

$$\left\|f_1\right\|^2_{\mathcal{H}(K_1)} = \mathrm{Var}(\Psi_1(f_1)) = 1 = \mathrm{Var}(\Psi_2(f_2)) = \left\|f_2\right\|^2_{\mathcal{H}(K_2)}.$$

Additional canonical variables can then be obtained recursively by restricting attention to functions that are orthogonal to the previous ones in the sequence.

Observe that

$$\mathrm{Cov}(\Psi_1(f_1), \Psi_2(f_2)) = \langle f_1(\star), \langle K_{12}(\star, \cdot), f_2(\cdot)\rangle_{\mathcal{H}(K_2)}\rangle_{\mathcal{H}(K_1)}$$

with $K_{12}(t_1, t_2) = \mathrm{E}[X_1(t_1)X_2(t_2)]$ the X_1 and X_2 process cross-covariance kernel. Thus, the problem is equivalent to finding the singular system of the operator R_{12} defined by

$$(R_{12}f_2)(t_1) = \langle K_{12}(t_1, \cdot), f_2(\cdot)\rangle_{\mathcal{H}(K_2)}, f_2 \in \mathcal{H}(K_2). \tag{4.1}$$

Since R_{12} is compact, we may write

$$R_{12} = \sum_{j=1}^{\infty} \rho_j \varphi_{j2} \otimes \varphi_{j1}$$

with $\{\varphi_{ji}\}_{j=1}^{\infty}$ a CONS for $\mathcal{H}(K_i)$, and $1 \geq \rho_1 \geq \rho_2 \geq \ldots$. The Eubank and Hsing [13] solution returns the (first) canonical variable pair $(\Psi_1(\varphi_{11}), \Psi_2(\varphi_{12}))$ with canonical correlation coefficient $\rho = \rho_1$. We will give an independent derivation of this result in Section 4.3.

4.2 $\mathcal{H}$-valued Random Variables

Let $(\Omega, \mathcal{A}, \mathrm{P})$ be a probability space and let $\mathcal{H}$ represent a real, separable Hilbert space with norm and inner product $||\cdot||_{\mathcal{H}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{H}}$, respectively. The Borel σ-field generated by the class of all open subsets of $\mathcal{H}$ will be denoted by $\mathcal{B}$. A $\mathcal{B}$ measurable function on $(\mathcal{H}, \mathcal{B})$ is then defined as follows.

Definition 4.2.1. *A mapping $X : \Omega \to \mathcal{H}$ is called an $\mathcal{H}$-valued random variable if X is $\mathcal{B}$ measurable; that is, for every set $E \in \mathcal{B}$*

$$X^{-1}(E) = \{\omega : X(\omega) \in E\} \in \mathcal{A}.$$

Our attention will be restricted to second order random variables that satisfy $\mathrm{E}||X||^2_{\mathcal{H}} < \infty$ with expectation being relative to P. Associated with such a random variable, we can define the Hilbert space indexed process

$$\{\langle X, f\rangle_{\mathcal{H}} : f \in \mathcal{H}\}. \tag{4.2}$$

Then, from [24] and [4], there exists an element $h \in \mathcal{H}$ and a covariance operator S such that for all f, $f' \in \mathcal{H}$,

$$\mathrm{E}[\langle X, f\rangle_{\mathcal{H}}] = \langle h, f\rangle_{\mathcal{H}}$$

and

$$\mathrm{E}[\langle X - h, f\rangle_{\mathcal{H}}\langle X - h, f'\rangle_{\mathcal{H}}] = \langle f, Sf'\rangle_{\mathcal{H}}. \tag{4.3}$$

Here h is the mean of the process and for simplicity we assume that $||h||_{\mathcal{H}} = 0$. In that case, (4.3) simplifies to

$$\mathrm{E}[\langle X, f\rangle_{\mathcal{H}}\langle X, f'\rangle_{\mathcal{H}}] = \langle f, Sf'\rangle_{\mathcal{H}}.$$

It is known that S is a Hilbert-Schmidt operator: e.g., [24]. In particular, this means it is compact and admits the eigenvalue-eigenvector decomposition

$$S = \sum_{j=1}^{\infty} \lambda_j \varphi_j \otimes \varphi_j,$$

where $\lambda_1 \geq \lambda_2 \geq \ldots > 0$.

4.3 Functional CCA

In this and the subsequent section, our goals are twofold:

(i) to obtain a novel, rigorous derivation of CCA for Hilbert space valued processes that coincides with the developments in Section 1.4 and

(ii) to extend this CCA notion to include an infinite dimensional analog of PCCA as defined in Section 1.4.

The setting to be studied can be described as follows. There are two separable Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ with norms and inner products $||\cdot||_i$ and $\langle\cdot,\cdot\rangle_i$, $i = 1, 2$, and a probability space $(\Omega, \mathcal{A}, \mathrm{P})$. Then, X_1 and X_2 are, respectively, $\mathcal{H}_1$ and $\mathcal{H}_2$-valued random variables with associated covariance operators S_1 and S_2. From [4] it may be concluded that there are also cross-covariance operators

$$S_{12} : \mathcal{H}_2 \to \mathcal{H}_1 \text{ and } S_{21} : \mathcal{H}_1 \to \mathcal{H}_2$$

with, e.g.,

$$\mathrm{E}\langle X_1, f_1\rangle_1\langle X_2, f_2\rangle_2 = \langle f_1, S_{12}f_2\rangle_1$$

and

$$S_{12} = S_{21}^*.$$

The functional CCA problem addressed in the work of [8] and [20] is based on finding $f_1 \in \mathcal{H}_1$, $f_2 \in \mathcal{H}_2$ to maximize

$$|\mathrm{Corr}(\langle X_1, f_1\rangle_1, \langle X_2, f_2\rangle_2)|.$$

As stated, a solution does not exist since the maximizers need not to be in $\mathcal{H}_1$ or $\mathcal{H}_2$ ([13], [23] and [6]). This problem is ignored in [8] while He et al. [20] impose additional restrictions to insure that the maximizers f_1 and f_2 are attained within $\mathcal{H}_1$ and $\mathcal{H}_2$.

Let $\{\lambda_{ji}, \varphi_{ji}\}_{j=1}^{\infty}$ for $i = 1, 2$ be the eigenvalue-eigenvector pairs for the S_i. Then, rather than working with the random variables of the form (4.2), we deal with the more general formulation of the processes

$$Z_i(f_i) = \sum_{j=1}^{\infty} f_{ji} \langle X_i, \varphi_{ji} \rangle_i, \tag{4.4}$$

that are indexed by the Hilbert spaces

$$\mathcal{H}_i(S_i) = \left\{ f_i : f_i = \sum_{j=1}^{\infty} \lambda_{ji} f_{ji} \varphi_{ji}, ||f_i||^2_{\mathcal{H}_i(S_i)} = \sum_{j=1}^{\infty} \lambda_{ji} f_{ji}^2 = ||S_i^{-1/2} f_i||_i^2 \right\}. \tag{4.5}$$

In the case of a random variable $\langle X_i, f_i \rangle_{\mathcal{H}_i}$ with $f_i \in \mathcal{H}_i$, this reduces to (4.4) with $f_{ji} = \langle f_i, \varphi_{ji} \rangle_i$. The Hilbert space

$$\begin{aligned} L^2_{Z_i} \;=\; & \left\{ Z_i(f_i) : f_i \in \mathcal{H}_i(S_i), \right. \\ & \left. ||Z_i(f_i)||^2_{L^2_{Z_i}} = \mathrm{Var}(Z_i(f_i)) = \sum_{j=1}^{\infty} \lambda_{ji} f_{ji}^2 = ||f_i||^2_{\mathcal{H}_i(S_i)} \right\} \end{aligned} \tag{4.6}$$

is clearly congruent (isometric) to $\mathcal{H}_i(S_i)$.

To connect the development thus far with the Eubank and Hsing [13] approach in Section 1.4, observe that the Karhunen-Loeve theorem gives a more direct interpretation of the X_i processes when they have meaningful pointwise values: e.g., when the processes are continuous with probability one. In that case,

$$X_i(t) = \sum_{j=1}^{\infty} \langle X_i, \varphi_{ji} \rangle \varphi_{ji}(t)$$

for which the covariance kernel is

$$K_i(t,t') = \mathrm{Cov}(X_i(t), X_i(t')) = \sum_{j=1}^{\infty} \lambda_{ji}\varphi_{ji}(t)\varphi_{ji}(t')$$

and the RKHS corresponding to $X_i(\cdot)$ is $\mathcal{H}_i(S_i)$. For our particular formulation using Hilbert space indexed processes, Parzen [32] called $\mathcal{H}_i(S_i)$ a congruent RKHS with reproducing kernel K_i.

Now we construct a new Hilbert space

$$\mathcal{H}_0 = \left\{h = (f_1, f_2) : f_i \in \mathcal{H}_i(S_i), i = 1, 2, ||h||_0^2 = ||f_1||_{\mathcal{H}_1(S_1)}^2 + ||f_2||_{\mathcal{H}_2(S_2)}^2\right\}.$$

From this we obtain the $\mathcal{H}_0$ indexed process

$$Z(h) = Z_1(f_1) + Z_2(f_2)$$

with covariance function

$$\begin{aligned}
\mathrm{Cov}(Z(h), Z(h')) &= \mathrm{Cov}(Z_1(f_1), Z_1(f_1')) + \mathrm{Cov}(Z_2(f_2), Z_2(f_2')) \\
&\quad + \mathrm{Cov}(Z_1(f_1), Z_2(f_2')) + \mathrm{Cov}(Z_1(f_1'), Z_2(f_2)) \\
&= \langle f_1, f_1' \rangle_{\mathcal{H}_1(S_1)} + \langle f_2, f_2' \rangle_{\mathcal{H}_2(S_2)} \\
&\quad + \mathrm{Cov}(Z_1(f_1), Z_2(f_2')) + \mathrm{Cov}(Z_1(f_1'), Z_2(f_2)). \quad (4.7)
\end{aligned}$$

In order to avoid the degenerate setting where perfect prediction is possible, we make the assumption

Assumption A1 *There exist no* $(f_1, f_2) \in \mathcal{H}_0$ *such that*

$$|\mathrm{Corr}(Z_1(f_1), Z_2(f_2))| = 1.$$

The cross-covariance terms in (4.7) can be characterized as deriving from operators between $\mathcal{H}_1(S_1)$ and $\mathcal{H}_2(S_2)$. To see this, define the functional

$$l_{f_2}(f_1) = \mathrm{Cov}(Z_1(f_1), Z_2(f_2))$$

on $\mathcal{H}_1(S_1)$. Clearly, l_{f_2} is linear since covariance is bilinear and, from (4.4),

$$\begin{aligned} Z_1(\alpha f_1 + \alpha' f_1') &= \sum_{j=1}^{\infty} [\alpha f_{j1} + \alpha' f_{j1}'] \langle X_1, \varphi_{j1} \rangle_1 \\ &= \alpha Z_1(f_1) + \alpha' Z_1(f_1') \end{aligned}$$

for any scalars α, α' and any f_1, $f_1' \in \mathcal{H}_1(S_1)$. Also, by the Cauchy-Schwarz inequality,

$$\begin{aligned} |l_{f_2}(f_1)| &\leq \sqrt{\mathrm{Var} Z_1(f_1) \mathrm{Var} Z_2(f_2)} \\ &= ||f_1||_{\mathcal{H}_1(S_1)} ||f_2||_{\mathcal{H}_2(S_2)}. \end{aligned}$$

Thus, l_{f_2} is a bounded linear functional on $\mathcal{H}_1(S_1)$ and by the Riesz representation theorem, there is a bounded operator $C_{12} : \mathcal{H}_2(S_2) \to \mathcal{H}_1(S_1)$ satisfying

$$\mathrm{Cov}(Z_1(f_1), Z_2(f_2)) = \langle f_1, C_{12} f_2 \rangle_{\mathcal{H}_1(S_1)}.$$

Similarly, there is a bounded operator $C_{21} : \mathcal{H}_1(S_1) \to \mathcal{H}_2(S_2)$ with $C_{21} = C_{12}^*$, which satisfies

$$\mathrm{Cov}(Z_1(f_1), Z_2(f_2)) = \langle C_{21} f_1, f_2 \rangle_{\mathcal{H}_2(S_2)}. \tag{4.8}$$

Proposition 4.3.1. *Under Assumption A1,* $||C_{12}|| = ||C_{21}|| < 1$.

Proof. By definition, we have

$$||C_{12}||^2 = \sup_{f_2 \in \mathcal{H}_2(S_2), ||f_2||_{\mathcal{H}_2(S_2)} = 1} ||C_{12} f_2||^2_{\mathcal{H}_1(S_1)}. \tag{4.9}$$

An application of the Cauchy-Schwarz inequality produces

$$\begin{aligned} |\mathrm{Cov}(Z_1(f_1), Z_2(f_2))| &= |\langle f_1, C_{12} f_2 \rangle_{\mathcal{H}_1(S_1)}| \\ &< \sqrt{\mathrm{Var} Z_1(f_1) \mathrm{Var} Z_2(f_2)} \\ &= ||f_1||_{\mathcal{H}_1(S_1)} ||f_2||_{\mathcal{H}_2(S_2)} \end{aligned}$$

with the strict inequality coming from Assumption A1. Now take $f_1 = C_{12} f_2$ to obtain $||C_{12} f_2||^2_{\mathcal{H}_1(S_1)} < ||C_{12} f_2||_{\mathcal{H}_1(S_1)} ||f_2||_{\mathcal{H}_2(S_2)}$. □

The operators C_{12} and S_{12} are, of course, related as we now explain. For this purpose, define

$$\tilde{\mathcal{H}}(S_i) = \left\{ \tilde{f}_i : \tilde{f}_i = \sum_{j=1}^{\infty} \tilde{f}_{ji} \phi_{ij}, ||\tilde{f}_i||^2_{\tilde{\mathcal{H}}(S_i)} = \sum_{j=1}^{\infty} \lambda_{ij} \tilde{f}^2_{ij} = ||S_i^{1/2} \tilde{f}_i||^2 < \infty \right\}, i = 1, 2.$$

Then, S_i is an isometric mapping from $\tilde{\mathcal{H}}(S_i)$ onto $\mathcal{H}(S_i)$; i.e., $\tilde{\mathcal{H}}(S_i) = S_i^{-1} \mathcal{H}(S_i)$. This leads us to

Lemma 4.3.1. *S_{12} is an operator from $\tilde{\mathcal{H}}(S_2)$ into $\mathcal{H}(S_1)$ with $||S_{12}|| < 1$.*

Proof. For any $\tilde{f}_2 \in \tilde{\mathcal{H}}(S_2)$ and $f_1 \in \mathcal{H}(S_1)$

$$\begin{aligned} \mathrm{Cov}(Z_1(f_1), Z_2(S_2 \tilde{f}_2)) &= \sum_{i,j} f_{1i} \tilde{f}_{2j} \langle \phi_{1i}, S_{12} \phi_{2j} \rangle \\ &= \sum_{i,j} f_{1i} \tilde{f}_{2j} \langle S_1^{1/2} \phi_{1i}, S_1^{1/2} S_{12} \phi_{2j} \rangle_{\mathcal{H}(S_1)} \\ &= \sum_{i,j} \lambda_{1i} f_{1i} \tilde{f}_{2j} \langle \phi_{1i}, S_{12} \phi_{2j} \rangle_{\mathcal{H}(S_1)} \\ &= \langle f_1, S_{12} \tilde{f}_2 \rangle_{\mathcal{H}(S_1)}. \end{aligned}$$

Now use the Cauchy-Schwarz inequality and $||S_2 \tilde{f}_2||_{\mathcal{H}(S_2)} = ||\tilde{f}_2||_{\tilde{\mathcal{H}}(S_2)}$. □

Lemma 4.3.1 provides the means to characterize C_{12}. Specifically, observe that

$$\begin{aligned}\mathrm{Cov}(Z_1(f_1), Z_2(S_2\tilde{f}_2)) &= \langle f_1, S_{12}\tilde{f}_2\rangle_{\mathcal{H}(S_1)} \\ &= \langle f_1, S_{12}S_2^{-1}S_2\tilde{f}_2\rangle_{\mathcal{H}(S_1)} \\ &= \langle f_1, C_{12}S_2\tilde{f}_2\rangle_{\mathcal{H}(S_1)}.\end{aligned}$$

In addition, the fact that S_{12} is compact on $\mathcal{H}$ along with an argument similar to that of Lemma 4.3.1 reveals that C_{12} is the limit of a sequence of finite dimensional operators. We summarize these findings as follows.

Theorem 4.3.1. *$C_{12} = S_{12}S_2^{-1}$ is a compact operator from $\mathcal{H}(S_2)$ into $\mathcal{H}(S_1)$.*

Referring once again to the Eubank and Hsing approach of Section 1.4, we know in that instance that $K_{12}(\cdot, t_2) \in \mathcal{H}_1(S_1)$, and $K_{12}(t_1, \cdot) \in \mathcal{H}_2(S_2)$; so, R_{12} in (4.1) is a bounded operator from $\mathcal{H}_2(S_2)$ into $\mathcal{H}_1(S_1)$ with the property that

$$\mathrm{Cov}(Z_1(f_1), Z_2(f_2)) = \langle f_1, R_{12}f_2\rangle_{\mathcal{H}_1(S_1)}.$$

Therefore, $R_{12} = C_{12}$.

For $h \in \mathcal{H}_0$, define

$$Qh = (f_1 + C_{12}f_2, f_2 + C_{21}f_1) \tag{4.10}$$

or, equivalently, it will be convenient to use the matrix form

$$Qh = \begin{bmatrix} I & C_{12} \\ C_{21} & I \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \tag{4.11}$$

with the convention that we view the resulting vector as an element of $\mathcal{H}_0$. Observe that

$$\mathrm{Cov}(Z(h), Z(h')) = \langle h, Qh'\rangle_0.$$

This leads to the following proposition.

Proposition 4.3.2. *$Q : \mathcal{H}_0 \to \mathcal{H}_0$ is invertible with inverse defined by*

$$Q^{-1}(h) = (C_{11.2}^{-1} f_1 - C_{12} C_{22.1}^{-1} f_2, C_{22.1}^{-1} f_2 - C_{21} C_{11.2}^{-1} f_1), \tag{4.12}$$

where $h = (f_1, f_2) \in \mathcal{H}_0$ and

$$C_{ii.k} = I - C_{ik} C_{ki} = (I - C_{ik} C_{ki})^*$$

for $i, k = 1, 2$.

Analogous to (4.11), (4.12) will also be expressed as

$$Q^{-1} h = \begin{bmatrix} C_{11.2}^{-1} & -C_{12} C_{22.1}^{-1} \\ -C_{21} C_{11.2}^{-1} & C_{22.1}^{-1} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

Proof. The form of the inverse follows directly once we have shown all the relevant inverses exist. Thus, let us concentrate on the latter task.

We can write $Q = I - T$ with

$$Th = (-C_{12} f_2, -C_{21} f_1) = - \begin{bmatrix} 0 & C_{12} \\ C_{21} & 0 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}.$$

Then,

$$\begin{aligned} ||Th||_0^2 &= ||C_{12} f_2||_{\mathcal{H}_1(S_1)}^2 + ||C_{21} f_1||_{\mathcal{H}_2(S_2)}^2 \\ &\leq ||C_{12}||^2 |||f_2||_{\mathcal{H}_2(S_2)}^2 + ||C_{21}||^2 ||f_1||_{\mathcal{H}_1(S_1)}^2 \\ &= ||C_{12}||^2 [|||f_1||_{\mathcal{H}_1(S_1)}^2 + ||f_2||_{\mathcal{H}_2(S_2)}^2] \\ &= ||C_{12}||^2 ||h||_0^2 \\ &< ||h||_0^2 \end{aligned}$$

by Proposition 4.3.1. Theorem 4.40 of [34] now has the consequence that $I - T = Q$ is invertible.

To complete the proof, we need to show that $C_{11.2}$ and $C_{22.1}$ are invertible. This again follows from Theorem 4.40 of [34] because $C_{11.2} = I - C_{12}C_{21}$ with $||C_{21}|| = ||C_{12}|| < 1$ from Proposition 4.3.1. □

Now define

$$\mathcal{H}(Q) = \left\{ h : h = Q \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, f_i \in \mathcal{H}_i(S_i), i = 1, 2, ||h||^2_{\mathcal{H}(Q)} = ||Q^{-1/2}h||^2_0 \right\}.$$

We have the following proposition.

Proposition 4.3.3. *$\mathcal{H}(Q)$ is congruent to*

$$L^2_Z = \left\{ Z(h) : h \in \mathcal{H}_0, ||Z(h)||^2_{L^2_Z} = \mathrm{Var}(Z(h)) \right\}$$

under the mapping $\Psi(h) = Z(Q^{-1}h)$.

Proof. Clearly, for $h \in \mathcal{H}_0$, $||Z(h)||^2_{L^2_Z} = \langle h, Qh \rangle_0 = ||Qh||^2_{\mathcal{H}(Q)}$. □

With Proposition 4.3.3 in hand we can now give our formulation of CCA. Specially, we seek $f_i \in \mathcal{H}_i(S_i)$ to maximize $|\mathrm{Cov}(Z_1(f_1), Z_2(f_2))|$. But,

$$\begin{aligned} \mathrm{Cov}(Z_1(f_1), Z_2(f_2)) &= \mathrm{Cov}(Z(f_1, 0), Z(0, f_2)) \\ &= \left\langle Q \begin{bmatrix} f_1 \\ 0 \end{bmatrix}, Q \begin{bmatrix} 0 \\ f_2 \end{bmatrix} \right\rangle_{\mathcal{H}(Q)} \end{aligned}$$

which leads to the conclusion that it is equivalent to find $f_i \in \mathcal{H}_i(S_i)$ to maximize

$$\left| \left\langle Q \begin{bmatrix} f_1 \\ 0 \end{bmatrix}, Q \begin{bmatrix} 0 \\ f_2 \end{bmatrix} \right\rangle_{\mathcal{H}(Q)} \right|.$$

The corresponding canonical variables are then recovered via the congruence mapping that links the two spaces.

The analysis from this point is driven by the results of [36] as described in Section 4.5. For that purpose, we express $\mathcal{H}(Q)$ as

$$\mathcal{H}(Q) = M_1 + M_2$$

with

$$M_1 = \left\{ h \in \mathcal{H}(Q) : h = Q \begin{bmatrix} f_1 \\ 0 \end{bmatrix} := (f_1, C_{21} f_1), f_1 \in \mathcal{H}_1(S_1) \right\} \quad (4.13)$$

and

$$M_2 = \left\{ h \in \mathcal{H}(Q) : h = Q \begin{bmatrix} 0 \\ f_2 \end{bmatrix} := (C_{12} f_2, f_2), f_2 \in \mathcal{H}_2(S_2) \right\}. \quad (4.14)$$

Regarding M_1 and M_2, we have the following result.

Proposition 4.3.4. *$\mathcal{H}(Q) = M_1 + M_2$ with "+" indicating an algebraic direct sum.*

Proof. Clearly any element of $\mathcal{H}_0$ can be written as the sum of elements in M_1 and M_2. We therefore need only show that $M_1 \cap M_2 = \{0\}$. Thus, suppose there exist $f_1 \in \mathcal{H}_1(S_1)$ and $f_2 \in \mathcal{H}_2(S_2)$ such that

$$\langle f_1, C_{21} f_1 \rangle_{\mathcal{H}_1(S_1)} = \langle C_{12} f_2, f_2 \rangle_{\mathcal{H}_2(S_2)}.$$

Then, from (4.7)

$$\mathrm{Cov}(Z_1(f_1), Z_2(f_2)) = \langle f_1, C_{12} f_2 \rangle_{\mathcal{H}_1(S_1)},$$

$$\mathrm{Var}(Z_1(f_1)) = \langle f_1, f_1 \rangle_{\mathcal{H}_1(S_1)} = \langle f_1, C_{12} f_2 \rangle_{\mathcal{H}_1(S_1)},$$

and

$$\mathrm{Var}(Z_2(f_2)) = \langle f_2, f_2 \rangle_{\mathcal{H}_2(S_2)} = \langle f_2, C_{21} f_1 \rangle_{\mathcal{H}_2(S_2)} = \langle C_{12} f_2, f_1 \rangle_{\mathcal{H}_1(S_1)}.$$

But, these relations have the consequence that $|\text{Corr}(Z_1(f_1), Z_2(f_2))| = 1$ which contradicts Assumption A1. □

To relate this to Sunder's scheme in Section 4.5, let $L_1 = M_1$ and $L_2 = M_2 \cap M_1^{\perp}$. Then, for $h_1 = Q \begin{bmatrix} f_1 \\ 0 \end{bmatrix} \in M_1$ and $h_2 = Q \begin{bmatrix} 0 \\ f_2 \end{bmatrix} \in M_2$ with $||h_i||_{\mathcal{H}(Q)} = 1$, $i = 1, 2$, the first canonical correlation can be characterized as

$$\begin{aligned}
\rho &= \sup_{\substack{h_1 \in M_1, h_2 \in M_2 \\ ||h_i||_{\mathcal{H}(Q)}=1, i=1,2}} |\langle h_1, h_2 \rangle_{\mathcal{H}(Q)}| \\
&= \sup_{\substack{h_1 \in L_1, \tilde{h}_2 \in L_2 \\ ||h_1||_{\mathcal{H}(Q)}=1, ||\tilde{h}_2 + B\tilde{h}_2||_{\mathcal{H}(Q)}=1}} |\langle h_1, B\tilde{h}_2 + \tilde{h}_2 \rangle_{\mathcal{H}(Q)}| \\
&= \sup_{\substack{h_1 \in L_1, \tilde{h}_2 \in L_2 \\ ||h_1||_{\mathcal{H}(Q)}=1, ||\tilde{h}_2 + B\tilde{h}_2||_{\mathcal{H}(Q)}=1}} |\langle h_1, B\tilde{h}_2 \rangle_{\mathcal{H}(Q)}| \\
&\leq \sup_{\substack{h_1 \in L_1, \tilde{h}_2 \in L_2 \\ ||h_1||_{\mathcal{H}(Q)}=1, ||\tilde{h}_2 + B\tilde{h}_2||_{\mathcal{H}(Q)}=1}} ||h_1||_{\mathcal{H}(Q)} ||B\tilde{h}_2||_{\mathcal{H}(Q)} \\
&= \sup_{\substack{\tilde{h}_2 \in L_2 \\ ||\tilde{h}_2 + B\tilde{h}_2||_{\mathcal{H}(Q)}=1}} ||B\tilde{h}_2||_{\mathcal{H}(Q)}
\end{aligned}$$

for $B = P_{L_1|M_2}(P_{L_2|M_2})^{-1}$. But, by taking $h_1 = B\tilde{h}_2 / ||B\tilde{h}_2||_{\mathcal{H}(Q)}$, we obtain $\langle h_1, B\tilde{h}_2 \rangle_{\mathcal{H}(Q)} = ||B\tilde{h}_2||_{\mathcal{H}(Q)}$. So, the bound is attainable and holds with equality. Thus, we have shown that

$$\rho = \sup_{\tilde{h}_2 \in L_2} ||B\tilde{h}_2||_{\mathcal{H}(Q)}$$

subject to

$$\begin{aligned}
1 &= ||h_2||^2_{\mathcal{H}(Q)} = ||B\tilde{h}_2 + \tilde{h}_2||^2_{\mathcal{H}(Q)} \\
&= \langle B\tilde{h}_2 + \tilde{h}_2, B\tilde{h}_2 + \tilde{h}_2 \rangle_{\mathcal{H}(Q)} \\
&= \langle \tilde{h}_2, \tilde{h}_2 \rangle_{\mathcal{H}(Q)} + 2\langle B\tilde{h}_2, \tilde{h}_2 \rangle_{\mathcal{H}(Q)} + \langle B\tilde{h}_2, B\tilde{h}_2 \rangle_{\mathcal{H}(Q)} \\
&= \langle \tilde{h}_2, \tilde{h}_2 \rangle_{\mathcal{H}(Q)} + \langle B\tilde{h}_2, B\tilde{h}_2 \rangle_{\mathcal{H}(Q)} \\
&= \langle \tilde{h}_2, (I + B^*B)\tilde{h}_2 \rangle_{\mathcal{H}(Q)},
\end{aligned}$$

since $B\tilde{h}_2 \in L_1$ is orthogonal to $\tilde{h}_2 \in L_2$.

The operator $I + B^*B$ is self-adjoint and positive. Therefore, it has a self-adjoint square root $(I + B^*B)^{1/2}$. Moreover, $(I + B^*B)$ and $(I + B^*B)^{1/2}$ are invertible meaning that we can equivalently work with $\tilde{h}_2' = (I + B^*B)^{1/2}\tilde{h}_2$ and maximize

$$||B\tilde{h}_2||_{\mathcal{H}(Q)} = ||B(I + B^*B)^{-1/2}\tilde{h}_2'||_{\mathcal{H}(Q)}$$

subject to $\tilde{h}_2' \in L_2$ and $||\tilde{h}_2'||^2_{\mathcal{H}(Q)} = 1$. A consequence of Theorem 4.3.2 below is that B^*B is compact. Hence, the maximizer is the eigenvector for the largest eigenvalue of the operator

$$T = (I + B^*B)^{-1/2}B^*B(I + B^*B)^{-1/2}.$$

In general, the eigenvalues and eigenvectors for T are the values $\tilde{\alpha}^2 > 0$ and vectors $\tilde{h}_2' \in L_2$ of unit norm that satisfy

$$T\tilde{h}_2' = \tilde{\alpha}^2\tilde{h}_2'.$$

Some algebra reveals that this is equivalent to finding a vector $\tilde{h}_2 \in L_2$ that solves

$$B^*B\tilde{h}_2 = \alpha^2\tilde{h}_2, \tag{4.15}$$

where

$$\alpha^2 = \frac{\tilde{\alpha}^2}{1 - \tilde{\alpha}^2} \tag{4.16}$$

and

$$||\tilde{h}_2||^2_{\mathcal{H}(Q)} = 1.$$

From (4.16) it follows that

$$\rho = \tilde{\alpha} = \frac{\alpha}{\sqrt{1+\alpha^2}}. \tag{4.17}$$

Now suppose that $\tilde{h}_2 \in L_2$ is any vector that satisfies (4.17). Its M_1 component is $B\tilde{h}_2$ and its M_2 component is $B\tilde{h}_2 + \tilde{h}_2$. These correspond to the canonical variables $\Psi(B\tilde{h}_2)$, $\Psi(\tilde{h}_2 + B\tilde{h}_2)$ of the Z_1 and Z_2 spaces. As such, both need to have unit variance. At present, we have

$$||B\tilde{h}_2||^2_{\mathcal{H}(Q)} = \langle \tilde{h}_2, B^* B\tilde{h}_2 \rangle_{\mathcal{H}(Q)} = \alpha^2,$$

and

$$||\tilde{h}_2 + B\tilde{h}_2||^2_{\mathcal{H}(Q)} = 1 + \alpha^2.$$

Therefore, the canonical variable for the Z_1 space is

$$\Psi\left(\frac{1}{\alpha} B\tilde{h}_2\right), \tag{4.18}$$

and the canonical variable for the Z_2 space is

$$\Psi\left(\frac{1}{\sqrt{1+\alpha^2}}(\tilde{h}_2 + B\tilde{h}_2)\right). \tag{4.19}$$

The correlation between these two random variables is

$$\begin{aligned}
\rho &= \mathrm{Cov}\left(\Psi\left(\frac{1}{\alpha} B\tilde{h}_2\right), \Psi\left(\frac{1}{\sqrt{1+\alpha^2}}(\tilde{h}_2 + B\tilde{h}_2)\right)\right) \\
&= \langle \frac{1}{\alpha} B\tilde{h}_2, \frac{1}{\sqrt{1+\alpha^2}}(\tilde{h}_2 + B\tilde{h}_2) \rangle_{\mathcal{H}(Q)} \\
&= \frac{1}{\alpha\sqrt{1+\alpha^2}} \langle B\tilde{h}_2, B\tilde{h}_2 \rangle_{\mathcal{H}(Q)} \\
&= \frac{\alpha^2}{\alpha\sqrt{1+\alpha^2}} \\
&= \frac{\alpha}{\sqrt{1+\alpha^2}}
\end{aligned} \tag{4.20}$$

as expected from (4.17).

It remains to characterize B^*B. This is accomplished in the following propositions and corollaries.

Proposition 4.3.5. *If $h = (C_{12}f_2, f_2) \in M_2$, then $P_{L_1|M_2}h = (C_{12}f_2, C_{21}C_{12}f_2)$.*

Proof. Let $h_2 = (C_{12}f_2, f_2) \in M_2$ and $h_1 = (f_1, C_{21}f_1) \in M_1 = L_1$. Then,

$$\langle P_{L_1|M_2}h_2, h_1\rangle_{\mathcal{H}(Q)} = \langle h_2, h_1\rangle_{\mathcal{H}(Q)} \tag{4.21}$$

for every $h_1 \in M_1$. Writing $P_{L_1|M_2}h_2 = (f_1^\star, C_{21}f_1^\star)$ leads to

$$\begin{aligned}
\langle P_{L_1|M_2}h_2, h_1\rangle_{\mathcal{H}(Q)} &= \langle (f_1^\star, C_{21}f_1^\star), (f_1, 0)\rangle_0 \\
&= \langle f_1^\star, f_1\rangle_{\mathcal{H}_1(S_1)} \\
&= \langle (C_{12}f_2, f_2), h_1\rangle_{\mathcal{H}(Q)} \\
&= \langle (C_{12}f_2, f_2), (f_1, 0)\rangle_0 \\
&= \langle C_{12}f_2, f_1\rangle_{\mathcal{H}_1(S_1)}
\end{aligned}$$

for every $f_1 \in \mathcal{H}_1(S_1)$. So, $f_1^\star = C_{12}f_2$. □

Proposition 4.3.5 has the following immediate corollaries.

Corollary 4.3.1. *If $h = (C_{12}f_2, f_2) \in M_2$, then $P_{L_2|M_2}h = (I - P_{L_1|M_2})h = (0, C_{22.1}f_2)$.*

Corollary 4.3.2. *If $h = (0, \tilde{f}_2) \in L_2$, $(P_{L_2|M_2})^{-1}h = (C_{12}C_{22.1}^{-1}\tilde{f}_2, C_{22.1}^{-1}\tilde{f}_2)$.*

Corollary 4.3.3. *For $h = (0, \tilde{f}_2) \in L_2$, we have*

$$Bh := P_{L_1|M_2}(P_{L_2|M_2})^{-1}h = (C_{12}C_{22.1}^{-1}\tilde{f}_2, C_{21}C_{12}C_{22.1}^{-1}\tilde{f}_2).$$

Corollary 4.3.4. *Let $h, h' \in L_2$, then*

$$\langle h, h'\rangle_{\mathcal{H}(Q)} = \langle (0, \tilde{f}_2), Q^{-1}(0, \tilde{f}_2{}')\rangle_0 = \langle \tilde{f}_2, C_{22.1}^{-1}\tilde{f}_2{}'\rangle_{\mathcal{H}_2(S_2)}. \tag{4.22}$$

Corollary 4.3.5. $B^*(f_1, C_{21}f_1) = (0, C_{21}f_1)$.

Proof. For $h = (f_2, C_{21}f_1) \in M_1 = L_1$ and $\tilde{h} = (0, \tilde{f}_2) \in L_2$,

$$\begin{aligned}
\langle h, B\tilde{h}\rangle_{\mathcal{H}(Q)} &= \langle Q^{-1}h, B\tilde{h}\rangle_0 \\
&= \langle (f_1, 0), (C_{12}C_{22.1}^{-1}\tilde{f}_2, C_{21}C_{12}C_{22.1}^{-1}\tilde{f}_2)\rangle_0 \\
&= \langle f_1, C_{12}C_{22.1}^{-1}\tilde{f}_2\rangle_{\mathcal{H}_1(S_1)} \\
&= \langle C_{22.1}^{-1}C_{21}f_1, \tilde{f}_2\rangle_{\mathcal{H}_2(S_2)} \\
&= \langle B^*h, \tilde{h}\rangle_{\mathcal{H}(Q)}.
\end{aligned}$$

An application of Corollary 4.3.4 completes the proof. □

Corollary 4.3.3 and 4.3.5 in combination give us the desired characterization as follows.

Theorem 4.3.2. *For $h = (0, \tilde{f}_2) \in L_2$,*

$$\begin{aligned}
B^*B(0, \tilde{f}_2) &= B^*(C_{12}C_{22.1}^{-1}\tilde{f}_2, C_{21}C_{12}C_{22.1}^{-1}\tilde{f}_2) \\
&= (0, C_{21}C_{12}C_{22.1}^{-1}\tilde{f}_2).
\end{aligned}$$

The eigenvalue-eigenvector problem for B^*B can now be formulated in various ways; e.g., for a generic u that does not need to be the same on each line, we can write

(i) $C_{21}C_{12}C_{22.1}^{-1}u = \alpha^2 u$
$\Leftrightarrow C_{21}C_{12}u = \alpha^2 C_{22.1}u$
$\Leftrightarrow (1+\alpha^2)C_{21}C_{12}u = \alpha^2 u$
$\Leftrightarrow C_{21}C_{12}u = \rho^2 u$, or

(ii) $C_{21}C_{12}C_{22.1}^{-1}u = \alpha^2 u$
$\Leftrightarrow (C_{22.1} - I)C_{22.1}^{-1}u = \alpha^2 u$

$$\Leftrightarrow u = \alpha^2 u + C_{22.1}^{-1} u$$
$$\Leftrightarrow (1-\alpha^2)u = C_{22.1}^{-1} u$$
$$\Leftrightarrow C_{22.1} u = \tfrac{1}{1-\alpha^2} u.$$

We illustrate special cases through the examples below.

Example 4.3.1. *Suppose that S_1, S_2 and S_{12} are all full rank, finite-dimensional matrices. Then,*

$$C_{12} = S_{12} S_2^{-1}, C_{21} = S_{21} S_1^{-1}$$

and

$$C_{22.1} = I - S_{21} S_1^{-1} S_{12} S_2^{-1} = [S_2 - S_{21} S_1^{-1} S_{12}] S_2^{-1}.$$

Thus,

$$C_{21} C_{12} u = \rho^2 u$$
$$\Leftrightarrow [S_{21} S_1^{-1} S_{12} S_2^{-1}] u = \rho^2 u$$
$$\Leftrightarrow S_{21} S_1^{-1} S_{12} u = \rho^2 S_2 u$$
$$\Leftrightarrow S_2^{-1/2} S_{21} S_1^{-1/2} S_1^{-1/2} S_{12} S_2^{-1/2} S_2^{1/2} u = \rho^2 S^{1/2} u$$
$$\Leftrightarrow \tilde{S}^* \tilde{S} u = \rho^2 u.$$

Therefore, our formulation in this instance is equivalent to the SVD of $\tilde{S} = S_1^{-1/2} S_{12} S_2^{-1/2}$ which in turn, is equivalent to Hotelling's classic solution for the finite dimensional case as established in [22].

Example 4.3.2. *For the pointwise FDA setting, we have*

$$(C_{12} f_2)(t_1) = \langle K_{12}(t_1, \cdot), g(\cdot) \rangle_{\mathcal{H}_2(S_2)} \tag{4.23}$$

and the eigenvalue-eigenvector decomposition of $C_{21}C_{12} = C_{12}^*C_{12}$ *is equivalent to the SVE of* C_{12}*. Thus, our solution coincides with the Eubank/Hsing solution for this case.*

4.4 PCCA

A similar approach to that of the previous section can be used to address the PCCA setting. There are now three Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$ with norms and inner products $\|\cdot\|_i$ and $\langle\cdot,\cdot\rangle_i$, $i = 1, 2, 3$. Then, the X_i are $\mathcal{H}_i$-valued random variables with associated covariance operators S_i, $i = 1, 2, 3$. As in Section 4.3, we can also define the cross-covariance operators S_{12}, S_{13}, S_{23} and their adjoints.

For $i = 1, 2, 3$, the Hilbert spaces $L^2_{Z_i}$ spanned by the process $Z_i(f_i)$, $f_i \in \mathcal{H}_i$, and their congruent Hilbert space $\mathcal{H}_i(S_i)$ with indexed process $Z_i(f_i)$ are defined exactly the same as (4.6) and (4.5). Hence, by the Riesz representation theorem, there are bounded operators $C_{ij} : \mathcal{H}_j(S_j) \to \mathcal{H}_i(S_i)$ satisfying

$$\mathrm{Cov}(Z_i(f_i), Z_j(f_j)) = \langle f_i, C_{ij} f_j \rangle_{\mathcal{H}_i(S_i)}$$

for $i, j = 1, 2, 3$ and $i \neq j$. Also, we have that $C_{ij} = C_{ji}^*$.

We construct the new Hilbert space

$$\begin{aligned}\mathcal{H}_0 = \{ \quad & h = (f_1, f_2, f_3) : \\ & f_i \in \mathcal{H}_i, i = 1, 2, 3, \|h\|_0^2 = \|f_1\|^2_{\mathcal{H}_1(S_1)} + \|f_2\|^2_{\mathcal{H}_2(S_2)} + \|f_3\|^2_{\mathcal{H}_3(S_3)}\}.\end{aligned}$$

The corresponding $\mathcal{H}_0$ indexed process is obtained as

$$Z(h) = Z_1(f_1) + Z_2(f_2) + Z_3(f_3).$$

For $h \in \mathcal{H}_0$, define

$$Qh = (f_1 + C_{12}f_2 + C_{13}f_3, C_{21}f_1 + f_2 + C_{23}f_3, C_{31}f_1 + C_{32}f_2 + f_3)$$

which can be written equivalently in the matrix form

$$Qh = \begin{bmatrix} I & C_{12} & C_{13} \\ C_{21} & I & C_{23} \\ C_{31} & C_{32} & I \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$$

and

$$\mathrm{Cov}(Z(h), Z(h')) = \langle h, Qh' \rangle_0.$$

Now we want to show that Q is invertible. We write Q as

$$Q = \left[\begin{array}{c|cc} I & C_{12} & C_{13} \\ \hline C_{21} & I & C_{23} \\ C_{31} & C_{32} & I \end{array}\right] := \begin{bmatrix} I & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix},$$

where $\tilde{B} = \begin{bmatrix} C_{12} & C_{13} \end{bmatrix}$, $\tilde{C} = \begin{bmatrix} C_{21} \\ C_{31} \end{bmatrix}$ and $\tilde{D} = \begin{bmatrix} I & C_{23} \\ C_{32} & I \end{bmatrix}$. If the inverse of Q exists, standard results on the form of the inverse of a block matrix tell us it should be of the form

$$Q^{-1} = \begin{bmatrix} I + \tilde{B}\tilde{E}^{-1}\tilde{C} & -\tilde{B}\tilde{E}^{-1} \\ -\tilde{E}^{-1}\tilde{C} & \tilde{E}^{-1} \end{bmatrix}, \tag{4.24}$$

where

$$
\begin{aligned}
\tilde{E} &= \tilde{D} - \tilde{C}\tilde{B} \\
&= \begin{bmatrix} I & C_{23} \\ C_{32} & I \end{bmatrix} - \begin{bmatrix} C_{21} \\ C_{31} \end{bmatrix} \begin{bmatrix} C_{12} & C_{13} \end{bmatrix} \\
&= \begin{bmatrix} I - C_{21}C_{12} & C_{23} - C_{21}C_{13} \\ C_{32} - C_{31}C_{12} & I - C_{31}C_{13} \end{bmatrix} \\
&= \begin{bmatrix} C_{22.1}^{1/2} & 0 \\ 0 & C_{33.1}^{1/2} \end{bmatrix} \\
&\quad \cdot \left(I - \begin{bmatrix} 0 & -C_{22.1}^{-1/2}(C_{23} - C_{21}C_{13})C_{33.1}^{-1/2} \\ -C_{33.1}^{-1/2}(C_{32} - C_{31}C_{12})C_{22.1}^{-1/2} & 0 \end{bmatrix} \right) \\
&\quad \cdot \begin{bmatrix} C_{22.1}^{1/2} & 0 \\ 0 & C_{33.1}^{1/2} \end{bmatrix}.
\end{aligned}
$$

From Proposition 4.3.2, $C_{22.1}$ and $C_{33.1}$ are invertible. Then, in order to show the invertibility of Q, it suffices to show the norm of

$$
\begin{bmatrix} 0 & -C_{22.1}^{-1/2}(C_{23} - C_{21}C_{13})C_{33.1}^{-1/2} \\ -C_{33.1}^{-1/2}(C_{32} - C_{31}C_{12})C_{22.1}^{-1/2} & 0 \end{bmatrix} \tag{4.25}
$$

is less than 1 by Theorem 4.40 of [34]. To see that this is the case, we first establish

Lemma 4.4.1. *The projection of $Z_2(f_2)$ onto $L^2_{Z_1}$ is $Z_1(C_{12}f_2)$.*

Proof. If $P_{Z_1}Z_2(f_2)$ denotes the projection, it must satisfy

$$
\mathrm{Cov}(Z_1(f_1), P_{Z_1}Z_2(f_2)) = \mathrm{Cov}(Z_1(f_1), Z_2(f_2))
$$

for every $f_1 \in \mathcal{H}_1(S_1)$. Since there is some $f_1^* \in \mathcal{H}_1(S_1)$ such that $P_{Z_1} Z_2(f_2) = Z_1(f_1^*)$, then

$$\begin{aligned} \mathrm{Cov}(Z_1(f_1), Z_2(f_2)) &= \langle f_1, C_{12} f_2 \rangle_{\mathcal{H}_1(S_1)} \\ &= \mathrm{Cov}(Z_1(f_1), Z_1(f_1^*)) \\ &= \langle f_1, f_1^* \rangle_{\mathcal{H}_1(S_1)}. \end{aligned}$$

Therefore, $f_1^* = C_{12} f_2$. □

Similarly, we have

Lemma 4.4.2. *The projection of $Z_3(f_3)$ onto $L^2_{Z_1}$ is $Z_1(C_{13} f_3)$.*

Besides having Assumption A1 apply to both the pairs Z_1, Z_2 and Z_1, Z_3, we also need the following assumption for PCCA.

Assumption A2 *There exist no $f_2 \in \mathcal{H}_2(S_2)$ and $f_3 \in \mathcal{H}_3(S_3)$ such that*

$$|\mathrm{Corr}(Z_2(f_2) - P_{Z_1} Z_2(f_2), Z_3(f_3) - P_{Z_1} Z_3(f_3))| = 1.$$

Now we have the following lemma.

Lemma 4.4.3. $\|C_{22.1}^{-1/2}(C_{23} - C_{21}C_{13})C_{33.1}^{-1/2}\|_{\mathcal{H}_2(S_2)} < 1$.

Proof. Observe that

$$\begin{aligned} & |\mathrm{Cov}(Z_2(f_2) - Z_1(C_{12} f_2), Z_3(f_3) - Z_1(C_{13} f_3))| \\ = \; & |\langle f_2, C_{23} f_3 \rangle_{\mathcal{H}_2(S_2)} - \langle f_2, C_{21} C_{13} f_3 \rangle_{\mathcal{H}_2(S_2)}| \\ < \; & (\mathrm{Var}(Z_2(f_2) - Z_1(C_{12} f_2)))^{1/2} (\mathrm{Var}(Z_3(f_3) - Z_1(C_{13} f_3)))^{1/2} \\ = \; & \langle f_2, C_{22.1} f_2 \rangle_{\mathcal{H}_2(S_2)}^{1/2} \langle f_3, C_{33.1} f_3 \rangle_{\mathcal{H}_3(S_3)}^{1/2} \\ = \; & \|C_{22.1}^{1/2} f_2\|_{\mathcal{H}_2(S_2)} \|C_{33.1}^{1/2} f_3\|_{\mathcal{H}_3(S_3)}. \end{aligned}$$

Letting $\tilde{f}_2 = C_{22.1}^{1/2} f_2$ and $\tilde{f}_3 = C_{33.1}^{1/2} f_3$, we then obtain

$$\langle \tilde{f}_2, C_{22.1}^{-1/2}(C_{23} - C_{21}C13)C_{33.1}^{-1/2}\tilde{f}_3\rangle_{\mathcal{H}_2(S_2)} < \|\tilde{f}_2\|_{\mathcal{H}_2(S_2)}\|\tilde{f}_3\|_{\mathcal{H}_3(S_3)}.$$

Finally, take $\tilde{f}_2 = C_{22.1}^{-1/2}(C_{23} - C_{21}C_{13})C_{33.1}^{-1/2}\tilde{f}_3$ to see that

$$\|C_{22.1}^{-1/2}(C_{23} - C_{21}C_{13})C_{33.1}^{-1/2}\tilde{f}_3\|_{\mathcal{H}_2(S_2)} < \|\tilde{f}_3\|_{\mathcal{H}_3(S_3)}.$$

□

An identical argument shows that

Lemma 4.4.4. $\|C_{33.1}^{-1/2}(C_{32} - C_{31}C_{12})C_{22.1}^{-1/2}\|_{\mathcal{H}_3(S_3)} < 1.$

In combination, Lemmas 4.4.3 and 4.4.4 show that the norm of (4.25) is less than 1. Thus,

$$I - \begin{bmatrix} 0 & -C_{22.1}^{-1/2}(C_{23} - C_{21}C_{13})C_{33.1}^{-1/2} \\ -C_{33.1}^{-1/2}(C_{32} - C_{31}C_{12})C_{22.1}^{-1/2} & 0 \end{bmatrix}$$

is invertible and Q^{-1} exists with the form given in (4.24).

Now define

$$\mathcal{H}(Q) = \{h : h = Q \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}, f_i \in \mathcal{H}_i(S_i), i = 1, 2, 3, \|h\|^2_{\mathcal{H}(Q)} = \|Q^{-1/2}h\|^2_0\}.$$

The three process version of Proposition 4.3.3 is provided by

Proposition 4.4.1. $\mathcal{H}(Q)$ *is congruent to*

$$L^2_Z = \{Z(h) : h \in \mathcal{H}_0, \|Z(h)\|^2_{L^2_Z} = \mathrm{Var}(Z(h))\}$$

under the mapping $\Psi(h) = Z(Q^{-1}h)$.

For the PCCA formulation, we wish to find $f_2 \in \mathcal{H}_2(S_2)$ and $f_3 \in \mathcal{H}_3(S_3)$ to maximize

$$|\mathrm{Cov}(Z_2(f_2) - Z_1(C_{12}f_2), Z_3(f_3) - Z_1(C_{13}f_3))|.$$

Since

$$\begin{aligned} & \mathrm{Cov}(Z_2(f_2) - Z_1(C_{12}f_2), Z_3(f_3) - Z_1(C_{13}f_3)) \\ = \; & \mathrm{Cov}(Z(-C_{12}f_2, f_2, 0), Z(-C_{13}f_3, 0, f_3)) \\ = \; & \left\langle Q \begin{bmatrix} -C_{12}f_2 \\ f_2 \\ 0 \end{bmatrix}, Q \begin{bmatrix} -C_{13}f_3 \\ 0 \\ f_3 \end{bmatrix} \right\rangle_{\mathcal{H}(Q)}, \end{aligned}$$

it suffices to find $f_2 \in \mathcal{H}_2(S_2)$ and $f_3 \in \mathcal{H}_3(S_3)$ to maximize

$$\left| \left\langle Q \begin{bmatrix} -C_{12}f_2 \\ f_2 \\ 0 \end{bmatrix}, Q \begin{bmatrix} -C_{13}f_3 \\ 0 \\ f_3 \end{bmatrix} \right\rangle_{\mathcal{H}(Q)} \right|.$$

The corresponding canonical variables are gained via the congruence mapping Ψ as in the CCA case.

Again, we apply the results from Sunder [36] described in Section 4.5. We express $\mathcal{H}(Q)$ as

$$\mathcal{H}(Q) = M_1 + M_2 + M_3$$

with

$$M_1 = \left\{ h \in \mathcal{H}(Q) : h = Q \begin{bmatrix} f_1 \\ 0 \\ 0 \end{bmatrix} := (f_1, C_{21}f_1, C_{31}f_1) \right\}$$

$$M_2 = \left\{ h \in \mathcal{H}(Q) : h = Q \begin{bmatrix} 0 \\ f_2 \\ 0 \end{bmatrix} := (C_{12}f_2, f_2, C_{32}f_2) \right\}$$

and

$$M_3 = \left\{ h \in \mathcal{H}(Q) : h = Q \begin{bmatrix} 0 \\ 0 \\ f_3 \end{bmatrix} := (C_{13}f_3, C_{23}f_3, f_3) \right\}.$$

Regarding M_1, M_2 and M_3, we have the following proposition.

Proposition 4.4.2. *$\mathcal{H}(Q) = M_1 + M_2 + M_3$ with "+" indicating an algebraic direct sum.*

Proof. The proof is similar to that of Proposition 4.3.4. We need to show that $M_1 \cap M_2 \cap M_3 = \{0\}$. Suppose there exist $f_i \in \mathcal{H}(S_i)$, $i = 1, 2, 3$, such that $(f_1, C_{21}f_1, C_{31}f_1) = (C_{12}f_2, f_2, C_{32}f_2) = (C_{13}f_3, C_{23}f_3, f_3)$. Then, we can conclude that

$$\begin{aligned} & \mathrm{Cov}(Z_2(f_2) - Z_1(C_{12}f_2), Z_3(f_3) - Z_1(C_{13}f_3)) \\ = \; & \langle f_1, f_1 \rangle_{\mathcal{H}_1(S_1)} - \langle f_3, f_3 \rangle_{\mathcal{H}_3(S_3)} = \langle f_1, f_1 \rangle_{\mathcal{H}_1(S_1)} - \langle f_2, f_2 \rangle_{\mathcal{H}_2(S_2)}, \end{aligned}$$

$$\mathrm{Var}(Z_2(f_2) - Z_1(C_{12}f_2)) = \langle f_1, f_1 \rangle_{\mathcal{H}_1(S_1)} - \langle f_2, f_2 \rangle_{\mathcal{H}_2(S_2)}$$

and

$$\mathrm{Var}(Z_3(f_3) - Z_1(C_{13}f_3)) = \langle f_1, f_1 \rangle_{\mathcal{H}_1(S_1)} - \langle f_3, f_3 \rangle_{\mathcal{H}_3(S_3)}.$$

Hence,

$$\mathrm{Corr}(Z_2(f_2) - Z_1(C_{12}f_2), Z_3(f_3) - Z_1(C_{13}f_3)) = 1$$

which contradicts Assumption A2. □

If we let $L_1 = M_1$, $L_2 = M_2 \cap M_1^{\perp}$ and $L_3 = M_3 \cap M_2^{\perp} \cap M_1^{\perp}$, then for

$$\hat{h}_2 = Q \begin{bmatrix} -C_{12} f_2 \\ f_2 \\ 0 \end{bmatrix} \in M_2 - P_{L_1} M_2 \text{ and } \hat{h}_3 = Q \begin{bmatrix} -C_{13} f_3 \\ 0 \\ f_3 \end{bmatrix} \in M_3 - P_{L_1} M_3$$

with $\|\hat{h}_i\|^2 = 1$, $i = 2, 3$, the first partial canonical correlation can be characterized as

$$\begin{aligned}
\rho &= \sup_{\substack{\hat{h}_2 \in M_2 - P_{L_1} M_2, \hat{h}_3 \in M_3 - P_{L_1} M_3 \\ ||\hat{h}_i||_{\mathcal{H}(Q)} = 1, i=2,3}} \left| \langle \hat{h}_2, \hat{h}_3 \rangle_{\mathcal{H}(Q)} \right| \\
&= \sup_{\substack{\tilde{h}_2 \in L_2, \tilde{h}_3 \in L_3 \\ ||\tilde{h}_2||_{\mathcal{H}(Q)} = 1, ||B\tilde{h}_3 + \tilde{h}_3||_{\mathcal{H}(Q)} = 1}} \left| \langle \tilde{h}_2, B\tilde{h}_3 + \tilde{h}_3 \rangle_{\mathcal{H}(Q)} \right| \\
&= \sup_{\substack{\tilde{h}_2 \in L_2, \tilde{h}_3 \in L_3 \\ ||\tilde{h}_2||_{\mathcal{H}(Q)} = 1, ||B\tilde{h}_3 + \tilde{h}_3||_{\mathcal{H}(Q)} = 1}} \left| \langle \tilde{h}_2, B\tilde{h}_3 \rangle_{\mathcal{H}(Q)} \right| \\
&\leq \sup_{\substack{\tilde{h}_2, \tilde{h}_3 \\ ||\tilde{h}_2||_{\mathcal{H}(Q)} = 1, ||B\tilde{h}_3 + \tilde{h}_3||_{\mathcal{H}(Q)} = 1}} \|\tilde{h}_2\|_{\mathcal{H}(Q)} \|B\tilde{h}_3\|_{\mathcal{H}(Q)} \\
&= \sup_{\tilde{h}_3 \in L_3, ||B\tilde{h}_3 + \tilde{h}_3||_{\mathcal{H}(Q)} = 1} \|B\tilde{h}_3\|_{\mathcal{H}(Q)}
\end{aligned}$$

for $B = P_{L_2|M_3}(P_{L_3|M_3})^{-1}$. The bound is attainable and holds with equality by taking $\tilde{h}_2 = B\tilde{h}_3 / \|B\tilde{h}_3\|_{\mathcal{H}(Q)}$. Thus, we see that

$$\rho = \sup_{\tilde{h}_3 \in L_3} \|B\tilde{h}_3\|_{\mathcal{H}(Q)}$$

subject to

$$1 = \|\hat{h}_3\|^2_{\mathcal{H}(Q)} = \langle \tilde{h}_3, (I + B^* B)\tilde{h}_3 \rangle_{\mathcal{H}(Q)}.$$

Also, similarly as before, it is equivalently to maximize

$$\|B\tilde{h}_3\|_{\mathcal{H}(Q)} = \|B(I + B^* B)^{-1/2} \tilde{h}_3'\|_{\mathcal{H}(Q)}$$

subject to $\tilde{h}_3' \in L_3$ and $\|\tilde{h}_3'\|^2_{\mathcal{H}(Q)} = 1$. The maximizer is the eigenvector for the largest eigenvalue of the operator

$$T = (I + B^*B)^{-1/2} B^*B (I + B^*B)^{-1/2}.$$

Let $\tilde{\alpha}^2 > 0$ and $\tilde{h}_3' \in L_3$ with unit norm be eigenvalues and eigenvectors for T, and $\alpha^2 > 0$ and $\tilde{h}_3 \in L_3$ be eigenvalues and eigenvectors of B^*B. Then, it can be seen that

$$\rho = \tilde{\alpha} = \frac{\alpha}{\sqrt{1+\alpha^2}}.$$

If $\tilde{h}_3 \in L_3$ is any eigenvector of B^*B, then its $M_2 - P_{L_1} M_2$ component is $B\tilde{h}_3$ and its $M_3 - P_{L_1} M_3$ component is $B\tilde{h}_3 + \tilde{h}_3$. The corresponding canonical variables are $\Psi(B\tilde{h}_3)$ and $\Psi(\tilde{h}_3 + B\tilde{h}_3)$ of the $Z_2 - P_{Z_1} Z_2$ and $Z_3 - P_{Z_1} Z_3$ spaces, respectively. Due to the required unit variance, we have that the partial canonical variable for Z_2 space is

$$\Psi\left(\frac{1}{\alpha} B\tilde{h}_3\right),$$

and the partial canonical variable for the Z_3 space is

$$\Psi\left(\frac{1}{\sqrt{1+\alpha^2}} (\tilde{h}_3 + B\tilde{h}_3)\right).$$

Thus, the partial canonical correlation is

$$\rho = \frac{1}{\sqrt{1+\alpha^2}}.$$

Now we characterize B^*B through the following propositions and corollaries.

Proposition 4.4.3. *If $h = (C_{12} f_2, f_2, C_{32} f_2) \in M_2$, then*

$$P_{L_1 | M_2} h = (C_{12} f_2, C_{21} C_{12} f_2, C_{31} C_{12} f_2).$$

Proof. For $h_1 = (f_1, C_{21}f_1, C_{31}f_1) \in M_1 = L_1$, we have the relation

$$\langle P_{L_1|M_2}h, h_1\rangle_{\mathcal{H}(Q)} = \langle h, h_1\rangle_{\mathcal{H}(Q)}.$$

Writing $P_{L_1|M_2}h = (f_1^\star, C_{21}f_1^\star, C_{31}f_1^\star)$ leads to

$$\begin{aligned}
\langle P_{L_1|M_2}h, h_1\rangle_{\mathcal{H}(Q)} &= \langle (f_1^\star, C_{21}f_1^\star, C_{31}f_1^\star), (f_1, 0, 0)\rangle_0 \\
&= \langle f_1^\star, f_1\rangle_{\mathcal{H}_1(S_1)} \\
&= \langle h, h_1\rangle_{\mathcal{H}(Q)} \\
&= \langle (C_{12}f_2, f_2, C_{32}f_2), (f_1, 0, 0)\rangle_0 \\
&= \langle C_{12}f_2, f_1\rangle_{\mathcal{H}_1(S_1)}
\end{aligned}$$

for every $f_i \in \mathcal{H}_i(S_i)$ with $i = 1, 2$. So $f_1^\star = C_{12}f_2$. □

Proposition 4.4.3 has the following corollaries.

Corollary 4.4.1. *If $h = (C_{12}f_2, f_2, C_{32}f_2) \in M_2$, then*

$$P_{L_2|M_2}h = (I - P_{L_1|M_2})h = (0, C_{22.1}f_2, (C_{32} - C_{31}C_{12})f_2).$$

Corollary 4.4.2. *If $h = (C_{13}f_3, C_{23}f_3, f_3) \in M_3$, then*

$$P_{L_1|M_3}h = (C_{13}f_3, C_{21}C_{13}f_3, C_{31}C_{13}f_3).$$

Corollary 4.4.3. *If $h = (C_{13}f_3, C_{23}f_3, f_3) \in M_3$, then*

$$P_{L_2|M_3}h = (0, (C_{23} - C_{21}C_{13})f_3, (C_{32} - C_{31}C_{12})C_{22.1}^{-1}(C_{23} - C_{21}C_{13})f_3).$$

Proof. For $\tilde{h}_2 = (0, C_{22.1}f_2, (C_{32} - C_{31}C_{12})f_2) \in L_2$ and $h \in M_3$, we have the relation

$$\langle P_{L_2|M_3}h, \tilde{h}_2\rangle_{\mathcal{H}(Q)} = \langle h, \tilde{h}_2\rangle_{\mathcal{H}(Q)}.$$

If we write $P_{L_2|M_3}h = (0, C_{22.1}f_2^\star, (C_{32} - C_{31}C12)f_2^\star)$, then,

$$\begin{aligned}
& \langle P_{L_2|M_3}h, \tilde{h}_2\rangle_{\mathcal{H}(Q)} \\
= \; & \langle (0, C_{22.1}f_2^\star, (C_{32} - C_{31}C12)f_2^\star), (-C_{12}f_2, f_2, 0)\rangle_0 \\
= \; & \langle C_{22.1}f_2^\star, f_2\rangle_{\mathcal{H}_2(S_2)} \\
= \; & \langle h, \tilde{h}_2\rangle_{\mathcal{H}(Q)} \\
= \; & \langle (0, 0, f_3), (0, C_{22.1}f_2, (C_{32} - C_{31}C_{12})f_2)\rangle_0 \\
= \; & \langle f_3, (C_{32} - C_{31}C_{12})f_2\rangle_{\mathcal{H}_3(S_3)} \\
= \; & \langle (C_{23} - C_{21}C_{13})f_3, f_2\rangle_{\mathcal{H}_2(S_2)}.
\end{aligned}$$

So, $f_2^\star = C_{22.1}^{-1}(C_{23} - C_{21}C_{13})f_3$. □

Corollary 4.4.4. *If $h = (C_{13}f_3, C_{23}f_3, f_3) \in M_3$, then*

$$P_{L_3|M_3}h = (0, 0, [C_{33.1} - (C_{32} - C_{31}C_{12})C_{22.1}^{-1}(C_{23} - C_{21}C_{13})]f_3).$$

Proof. $P_{L_3|M_3}h = (I - P_{L_1|M_3} - P_{L_2|M_3})h$. □

For notational simplicity, let $C_0 = C_{33.1} - (C_{32} - C_{31}C_{12})C_{22.1}^{-1}(C_{23} - C_{21}C_{13})$.

Corollary 4.4.5. *For $h = (0, 0, \tilde{f}_3) \in L_3$,*

$$\begin{aligned}
Bh \; &:= \; P_{L_2|M_3}(P_{L_3|M_3})^{-1}h \\
&= \; (0, (C_{23} - C_{21}C_{13})C_0^{-1}\tilde{f}_3, (C_{32} - C_{31}C_{12})C_{22.1}^{-1}(C_{23} - C_{21}C_{13})C_0^{-1}\tilde{f}_3).
\end{aligned}$$

Corollary 4.4.6. *If $h = (0, C_{22.1}f_2, (C_{32} - C_{31}C_{12})f_2) \in L_2$, then*

$$B^*h = (0, 0, (C_{32} - C_{31}C_{12})f_2).$$

Proof. For $h = (0, C_{22.1}f_2, (C_{32} - C_{31}C_{12})f_2) \in L_2$ and $\tilde{h}_3 = (0, 0, \tilde{f}_3) \in L_3$,

$$\begin{aligned}
\langle B\tilde{h}_3, h\rangle_{\mathcal{H}(Q)} &= \langle B\tilde{h}_3, Q^{-1}h\rangle_0 \\
&= \langle B\tilde{h}_3, (-C_{12}f_2, f_2, 0)\rangle_0 \\
&= \langle (C_{23} - C_{21}C_{13})C_0^{-1}\tilde{f}_3, f_2\rangle_{\mathcal{H}_2(S_2)} \\
&= \langle C_0^{-1}\tilde{f}_3, (C_{32} - C_{31}C_{12})f_2\rangle_{\mathcal{H}_3(S_3)} \\
&= \langle \tilde{h}_3, B^*h\rangle_{\mathcal{H}(Q)} \\
&= \langle Q^{-1}\tilde{h}_3, B^*h\rangle_0 \\
&= \langle ([C_{21}C_{22.1}^{-1}(C_{23} - C_{21}C_{13}) - C_{13}]C_0^{-1}\tilde{f}_3, \\
&\quad -C_{22.1}^{-1}(C_{23} - C_{21}C_{13})C_0^{-1}\tilde{f}_3, C_0^{-1}\tilde{f}_3), B^*h\rangle_0.
\end{aligned}$$

Therefore,

$$B^*h = (0, 0, (C_{32} - C_{31}C_{12})f_2).$$

□

Now, through Corollaries 4.4.5 and 4.4.6, we finally obtain

Theorem 4.4.1. *For $h = (0, 0, \tilde{f}_3) \in L_3$,*

$$B^*Bh = (0, 0, (C_{32} - C_{31}C_{12})C_{22.1}^{-1}(C_{23} - C_{21}C_{13})C_0^{-1}\tilde{f}_3)).$$

4.5 Technical Appendix

The essential tool that is used for our work in Section 4.3 and Section 4.4 derives from the results in Sunder [36]. We summarize the key aspects that are needed for our purpose in this appendix.

Assume that a Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $|| \cdot ||$ can be written as the algebraic direct sum of n closed subspaces $M_1, \ldots, M_n$.

That is,

$$\mathcal{H} = \sum_{i=1}^{n} M_i,$$

where $M_1, \ldots, M_n$ are closed subspaces of $\mathcal{H}$ that satisfy

$$M_i \cap \sum_{j \neq i} M_j = \{0\}.$$

Now, for $1 \leq k \leq n$, define

$$R_k = \sum_{i=1}^{k} M_i, L_k = R_k \cap R_{k-1}^{\perp}.$$

Then, $L_k \perp M_i$, for $i = 1, \ldots, k-1$, and by construction $\mathcal{H} = \sum_{i=1}^{n} L_i$. Similarly,

$$\sum_{i=1}^{k} L_i = \sum_{i=1}^{k} M_i.$$

Let P_{M_k} and P_{L_k} be the orthogonal projection operators onto M_k and L_k, respectively. Then, for $1 \leq k \leq n$ and $1 \leq j \leq k \leq n$, we define the restriction of P_{L_j} to M_k by

$$P_{L_j|M_k} x = P_{L_j} x$$

for $x \in M_k$. Similarly,

$$P_{M_k|L_j} y = P_{M_k} y$$

for $y \in L_j$. Our first result states that $P_{L_k|M_k}$ is actually invertible.

Lemma 4.5.1. *$P_{L_k|M_k}$ is one-to-one and onto.*

Proof. By definition, $P_{L_k|M_k}$ is the projection operator P_{L_k} restricted to M_k. In general, P_{L_k} maps $\sum_{i=1}^{k} M_i$ onto L_k. But, $\oplus_{i=1}^{k-1} L_i = \sum_{i=1}^{k-1} M_i \perp L_k$. So, P_{L_k} maps M_k onto L_k. If $x \in M_k$ satisfies $P_{L_k} x = 0$, it must be that $x \in M_k \cap L_k^{\perp} = M_k \cap \sum_{i=1}^{k-1} M_i$. By assumption, the only element of this set is the 0 vector. $\square$

The adjoint for $P_{L_j|M_k}$ is determined in the next lemma.

Lemma 4.5.2. $(P_{L_j|M_k})^* = P_{M_k|L_j}$.

Proof. Let $x \in L_j$ and $y \in M_k$. Then,

$$\langle P_{M_k}x, y\rangle = \langle x, P_{M_k}y\rangle = \langle x, y\rangle = \langle x, P_{L_j}y\rangle = \langle x, P_{L_j|M_k}y\rangle.$$

□

For any $x \in \mathcal{H}$, we can write $x = \sum_{j=1}^{n} x_j$ for some unique set of vectors $x_1, x_2, \ldots, x_n$ with $x_j \in L_j$. In particular,

$$P_{M_k}x = \sum_{j=1}^{k} P_{M_k}x_j$$

since $P_{M_k}x \in M_k = \oplus_{i=1}^{k} L_i \perp L_j$ for $j > k$. But, if $P_{M_k}x \in M_k$, it must be that

$$P_{M_k}x_j = \sum_{i=1}^{k} P_{L_i|M_k}P_{M_k}x_j.$$

That leads to

$$P_{M_k}x = \sum_{j=1}^{k}\sum_{i=1}^{k} P_{L_i|M_k}P_{M_k}x_j = \sum_{j=1}^{k}\sum_{i=1}^{k} P_{L_i|M_k}P_{M_k|L_j}P_{L_j}x,$$

because $P_{L_j}^* = P_{L_j}$ and $P_{L_j}^2 = P_{L_j}$. We have therefore proved the following result.

Theorem 4.5.1. $P_{M_k} = \sum_{j=1}^{k}\sum_{i=1}^{k} P_{L_i|M_k}P_{M_k|L_j}P_{L_j}$.

Theorem 1 of Sunder (1988) is an immediate corollary of our expression for P_{M_k}.

Corollary 4.5.1. *For $x \in M_k$, we can write M_k as*

$$\begin{aligned} M_k &= \left\{(P_{L_1|M_k}x, \ldots, P_{L_{k-1}|M_k}x, P_{L_k|M_k}x, 0, \ldots, 0)\right\} \\ &= \left\{(P_{L_1|M_k}(P_{L_k|M_k})^{-1}P_{L_k|M_k}x, \ldots, P_{L_{k-1}|M_k}(P_{L_k|M_k})^{-1}P_{L_k|M_k}x, P_{L_k|M_k}x, 0, \ldots, 0)\right\} \\ &= \left\{(A_{L_1|L_k}z, \ldots, A_{L_{k-1}|L_k}z, z, 0, \ldots, 0)\right\}, \end{aligned}$$

where $z = P_{L_k|M_k}x \in L_k$ and $A_{L_j|L_k} = P_{L_j|M_k}(P_{L_k|M_k})^{-1}$ for $1 \leq j \leq k \leq n$.

Corollary 4.5.1 has the consequence that problems involving optimization over M_k can instead be formulated in terms of equivalent problems on L_k.

CHAPTER 5

SUMMARY AND CONCLUSIONS

The research presented in this dissertation addresses theory and application of the SVD and GSVD for both finite and infinite dimensional problems, the latter in the form of the SVE and the GSVE. The first problem we addressed was that of finding one possible extension of the GSVD as a GSVE for compact operators. We found that it is not possible to directly extend the work of Van Loan [40] on the finite dimensional GSVD because a key step is only valid for matrix operators. So we constructed our GSVE in a new way.

Second, we focused on the χ^2 method for regularization parameter estimation. Previous work has suggested that the normal theory version of the χ^2 method could also be used for nonnormal data. We proved that under the condition of ϕ-mixing, the variance factor used to normalize the penalized least-squares criterion is not 2 for nonnormal data. An analytic example is presented where the actual variance factor is larger than 2 for a simple Poisson case. Since the ϕ-mixing property is only an asymptotic condition, we presented an empirical study to estimate the new variance factor for selected finite sample experiments. The simulation results verified that the variance factor is larger than 2 in a more involved context that uses Poisson errors with normal signals that are blurred with Phillips and Shaw matrix transformations. We also explored how this affects parameter estimation for solving the regularization problems. Though the estimated values of the parameter do not differ much, we found that using an alternative variance factor that is tailored to the simulation produces confidence intervals with a coverage of the true parameter that is closer to the nominal level and larger than using the original variance factor. Also choosing the alternative variance factor pro-

duced different regularization parameter choices between 5% and 12% of the time and saved computation time when these choices differed.

Third, we provide a rigorous derivation for canonical correlation and partial canonical correlation for certain Hilbert space indexed stochastic processes. This work removes these restrictions that existed in previous work such as Dauxois et al. [7] [8] largely focus on the finite dimensional covariance operators whose range are closed and He et al. [20] impose restrictions on the cross-covariances of coefficients in the two process' Karhunen- Loeve expansions to insure that the canonical variables are elements of $\mathcal{H}$. For two processes, our derivation produces the same solution as in Eubank and Hsing [13]. While it appears to be difficult to use their approach to obtain results for the PCCA framework, the approach here extends readily to the PCCA and more general settings. It relies on a key congruence mapping between the space spanned by a second order, $\mathcal{H}$-valued process and a particular Hilbert function space derived from the process' covariance operator. It is an application of methodology for constructing orthogonal direct sums from algebraic direct sums of closed subspaces.

Topics to be addressed in future work include the following.

(i) It is important to explore more applications of the new GSVE derived in Chapter 2. One possible direction is the time series regression model of Parzen [31] which provides the infinite dimensional analog of the finite dimensional linear model. It will be interesting to see whether the new GSVE can be used to study time series regression similarly to the use of the GSVD for the study of linear models as presented by Van Loan [40]. Another possible direction is to find a new form of GSVD to obtain

estimates of GSVE as Hansen [16] used SVD to approximate SVE.

(ii) While our results in Chapter 3 show that we cannot assume the same variance factor for normal and nonnormal measurements, more work needs to be done to evaluate the impact of the variance factor in regularization parameter estimation. Our current results do not show that the results are significantly better when a more appropriate variance is used, although the use of the large variance factor makes the tolerance on the iteration less severe thereby allowing the algorithm to converge more quickly. It would be interesting to consider other distributions for error term such as the gamma distribution. It is also of interest to determine how to choose the variance factor for real data.

(iii) In Chapter 4, we have developed a framework that can be used to study the correlation properties of groups of Hilbert space indexed stochastic processes. Our applications have been restricted to groups of size two or three; however, it is clear that similar analyses are possible with any finite number of processes. For example, the partial canonical correlation work of Section 4.4 extends in principle to examination of pairs of residual processes after correcting for projections onto several other processes.

BIBLIOGRAPHY

[1] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[2] L. Autonne. *Sur les groupes lineaires, reels et orthogonaux*, volume 30. 1902.

[3] L. J. Bain and M. Engelhardt. *Introduction to Probability and Mathematical Statistics*. Duxbury, 1987.

[4] C. R. Baker. Mutual information for Gaussian processes. *SIAM J. Appl. Math.*, 19:451–458, 1970.

[5] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, Inc., New York, 1968.

[6] J. Cupidon, R. L. Eubank, D. Gilliam, and F. Ruymgaart. Some properties of canonical correlations and variates in infinite demensions. *J. Multivariate Anal.*, 99:1083–1104, 2008.

[7] J. Dauxois and G. M. Nkiet. Canonical analysis of two Euclidean subspaces and its applications. *Linear Algebra Appl.*, 264:355–388, 1997.

[8] J. Dauxois, G. M. Nkiet, and Y. Romain. Canonical analysis relative to a closed subspace. *Linear Algebra Appl.*, 388:119–145, 2004.

[9] B. De Moor and G. H. Golub. The restricted singular value decomposition: properties and applications. *SIAM. J. Matrix Anal. & Appl.*, 12:401–425, 1991.

[10] C. Eckart and G. Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1:211–218, 1936.

[11] W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Netherlands, 2000.

[12] R. L. Eubank. *Nonparametric Regression and Spline Smoothing*. Marcel Dekker, Inc., New York, 1999.

[13] R. L. Eubank and T. Hsing. Canonical correlation for stochastic processes. *Stochastic Process. Appl.*, 118:1634–1661, 2007.

[14] K. V. Fernando and S. Hammarling. A product induced singular value decomposition for two matrices and balanced realization. *NAG Technical Report*, TR8/87, 1987.

[15] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.

[16] P. C. Hansen. Computation of the singular value expansion. *Computing*, 40:185–199, 1988.

[17] P. C. Hansen. *Rank-deficient and discrete ill-posed problems.* SIAM, Philadelphia, 1998.

[18] P. C. Hansen. Regularization tools version 4.0 for Matlab 7.3. *Numer. Algo.*, 46:189–194, 2007.

[19] P. C. Hansen. *Discrete inverse problems: insight and algorithms.* SIAM, Philadelphia, 2010.

[20] G. He, H. G. Muller, and J. L. Wang. Functional canonical correlation analysis for square integrable stochastic processes. *J. Multivariate Anal.*, 85:54–77, 2003.

[21] H. Hotelling. Relations between two sets of variates. *Biometrika*, 28:321–377, 1936.

[22] A. Kshirsagar. *Multivariate Analysis.* Marcel Dekker, New York, 1972.

[23] A. Kupresanin, H. Shin, D. King, and R. L. Eubank. An RKHS framework for functional data analysis. *J. Statist. Plann. Inference*, 140:3627–3637, 2010.

[24] R. G. Laha and V. K. Rohatgi. *Probability Theory.* John Wiley & Sons, Inc., New York, 1979.

[25] E. L. Lehmann. *Elements of Large-Sample Theory*. Springer, New York, 1999.

[26] C. L. Mallows. Some comments on C_p. *Technometrics*, 15:661–675, 1973.

[27] J. L. Mead. Parameter estimation: A new approach to weighting a priori information. *J. Inv. Ill-Posed Problems*, 15:1–21, 2007.

[28] J. L. Mead and R. A. Renaut. A Newton root-finding algorithm for estimating the regularization parameter for solving ill-conditioned least squares problems. *Inverse Problems*, 25, 2009.

[29] C. C. Paige. The general linear model and the generalized singular value decomposition. *Linear Algebra Appl.*, 70:269–284, 1985.

[30] C. C. Paige and M. A. Saunders. Towards a generalized singular value decomposition. *SIAM J. Numer. Anal.*, 18:398–405, 1981.

[31] E. Parzen. Regression analysis of continuous parameter time series. In *Proceedings of the fourth Berkeley Symposium on Mathematical Statistics and Probability*, pages 469–489. University of California Press, Berkeley, 1961.

[32] E. Parzen. Statistical inference on time series by RKHS methods. Canadian Mathematical Congress, 1970.

[33] S. N. Roy. *Some Aspects of Multivariate Analysis*. Wiley, New York, 1958.

[34] B. P. Rynne and M. A. Youngson. *Linear Functional Analysis*. Springer, London, 2000.

[35] G. W. Stewart. Computing the CS-decomposition of a partitioned orthogonal matrix. *Numer. Math.*, 40:297–306, 1982.

[36] V. Sunder. N subspaces. *Canad. J. Math.*, XL:38–54, 1988.

[37] J. J. Sylvester. A new proof that a general quadric may be reduced to its canonical form by means of a real orthogonal substitution. *Messenger of Math.*, 19:1–5, 1889.

[38] A. M. Thompson, J. W. Kay, and D. M. Titterington. Noise estimation in signal restoration using regularization. *Biometrika*, 78:475–488, 1991.

[39] A. N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Mathematics-Doklady*, 4:1624–1627, 1963.

[40] C. F. Van Loan. Generalizing the singular value decomposition. *SIAM J. Numer. Anal.*, 13:76–83, 1976.

[41] C. R. Vogel. *Computational Methods for Inverse Problems.* SIAM, Philadelphia, 2002.

[42] H. Zha. The restricted singular value decomposition of matrix triplets. *SIAM J. Matrix Anal. Appl.*, 12:172–194, 1991.